Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”

Comparação de modelos com censura intervalar em análise de
sobrevivência

Elizabeth Strapasson

Tese apresentada para obtenção do título de Doutor
em Agronomia. Área de concentração: Estatística e
Experimentação Agronômica

Piracicaba

2007

Elizabeth Strapasson
Licenciada em Ciências - Habilitação Matemática

Comparação de modelos com censura intervalar em análise de sobrevivência

Orientadora:

Prof^ª Dr^a **CLARICE G. B. DEMÉTRIO**

Co-orientador:

Prof. Dr. **ENRICO A. COLOSIMO**

Tese apresentada para obtenção do título de Doutor
em Agronomia. Área de concentração: Estatística e
Experimentação Agronômica

Piracicaba

2007

Dados Internacionais de Catalogação na Publicação (CIP)
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP

Strapasson, Elizabeth

Comparação de modelos com censura intervalar em análise de sobrevivência / Elizabeth Strapasson. - - Piracicaba, 2007.
135 p.

Tese (Doutorado) - - Escola Superior de Agricultura Luiz de Queiroz, 2007.
Bibliografia.

1. Análise de sobrevivência 2. Método de Monte Carlo 3. Modelos matemáticos 4. Simulação (Estatística) I. Título

CDD 630.2195

DEDICATÓRIA

A Deus por me conceder a capacidade de converter sonhos em realidade.
Sem "Ele" não teria conseguido realizar esse sonho.

À minha família, em especial aos meus pais Maria e Pedro,
o amor e o incentivo constantes.

AGRADECIMENTOS

À Prof^a Dr^a Clarice Garcia Borges Demétrio, a orientação, a amizade e a confiança depositada.

Ao Prof. Dr. Enrico Antônio Colosimo, a orientação, a amizade e o incentivo.

Ao CNPq, o fundamental suporte financeiro concedido.

À Universidade Estadual de Londrina e aos professores do Departamento de Estatística, o afastamento concedido para a realização deste trabalho.

Aos professores e funcionários do Departamento de Ciências Exatas da ESALQ/USP que me propiciaram condições para a realização deste trabalho.

Aos colegas e amigos de doutorado, em especial a Ana Maria Souza de Araujo, Genevile Carife Bergamo, Pedro Ferreira Filho, a força, a amizade, a troca de conhecimentos e a atenção recebida em todos os momentos.

A todos que, de forma direta ou indireta, contribuíram para a realização deste trabalho.

SUMÁRIO

RESUMO	7
ABSTRACT	8
LISTA DE FIGURAS	9
LISTA DE TABELAS	11
1 INTRODUÇÃO	15
2 DESENVOLVIMENTO	17
2.1 Conceitos Básicos	17
2.1.1 Funções do tempo de Sobrevivência	17
2.1.2 Tipos de Censura	18
2.2 Modelo de Regressão de Cox	21
2.2.1 Estimação dos Parâmetros no Modelo de Regressão de Cox	22
2.3 Modelo de Weibull	23
2.3.1 Distribuição de Weibull	23
2.3.2 Estimação dos Parâmetros	25
2.3.3 Modelo de Regressão Weibull	27
2.4 Modelos de Regressão para Dados de Sobrevivência Agrupados	28
2.4.1 Modelos Discretos de Regressão	30
2.4.1.1 Modelo Discreto de Riscos Proporcionais de Cox	32
2.4.1.2 Modelo Discreto Weibull	34
2.5 Discriminação Entre os Modelos de Cox e de Weibull para Dados com Censura Intervalar	36
2.5.1 Teste da Razão de Máxima Verossimilhança	36
2.5.2 Teste Escore	37
2.6 Comparação de Métodos para tratar Dados com Censura Intervalar - Um Estudo de Simulação	40
2.7 Avaliação do Teste da Razão de Máxima Verossimilhança e do Escore por Simulação	43
2.8 Resultados e Discussão dos Estudos de Simulação	44
2.8.1 Comparação de Métodos para tratar Dados com Censura Intervalar	44
2.8.1.1 Caso sem covariável	44

2.8.1.2 Caso com covariável	53
2.8.2 Avaliação do Teste da Razão de Máxima Verossimilhança e do Escore por Simulação	60
2.8.2.1 Teste da Razão de Máxima Verossimilhança	60
2.8.2.2 Teste Escore	63
2.9 Aplicações	68
2.9.1 Modelo Discreto Weibull	68
2.9.2 Comparação de Métodos para tratar Dados com Censura Intervalar	71
3 CONCLUSÕES	73
REFERÊNCIAS	74
APÊNDICE	77
ANEXOS	80

RESUMO

Comparação de modelos com censura intervalar em análise de sobrevivência

Censura intervalar resulta quando os tempos de sobrevivência não são exatamente conhecidos, sabe-se apenas que eles ocorreram dentro de um intervalo. Dados de sobrevivência agrupados são casos particulares de censura intervalar quando os indivíduos são avaliados nos mesmos intervalos de tempo, ocasionando um grande número de empates.

Um procedimento comum para a análise desse tipo de dados é ignorar a natureza de censura intervalar dos dados, ou seja, tratar a variável aleatória tempo como contínua e assumir que o evento ocorreu no início, no ponto médio ou no final do intervalo e, então, usar um método padrão de análise de sobrevivência. Neste estudo, simulações de Monte Carlo, com o modelo de Weibull, foram realizadas para comparar esses três procedimentos e um método novo proposto que é uma combinação desses três métodos e é orientado pela observação do histograma do tempo versus a frequência de cada intervalo para a decisão de qual valor a ser usado. Considera-se também a análise dos dados como censura intervalar. Os resultados mostram que analisar os dados exatamente como censura intervalar é a forma correta. Entretanto, quando a taxa de falha aumenta o ponto médio poderia ser usado.

A natureza discreta dos tempos de falha deve ser reconhecida quando existe um grande número de empates. Métodos de regressão para tratar dados agrupados são apresentados por Lawless (2003) e Collett (2003), cuja estrutura é especificada em termos da probabilidade de um indivíduo falhar em um intervalo, condicionada à sua sobrevivência ao intervalo anterior. Os modelos considerados na literatura são o de riscos proporcionais de Cox ou o logístico.

O modelo de Weibull é proposto, neste trabalho, como uma alternativa ao modelo de Cox para ajustar dados de sobrevivência com censura intervalar no contexto de modelos discretos. Através de simulações foram construídas as estatísticas da razão de verossimilhança e do teste escore para a discriminação entre esses dois modelos.

Para ilustrar as simulações duas aplicações em dados agrônômicos foram utilizadas.

Palavras-chave: Censura intervalar; Simulação de Monte Carlo; Modelo Weibull

ABSTRACT

Comparison of interval-censored models in survival analysis

Interval-censored results when survival times are not exactly known, knowing only that they occur in an interval. Grouped survival data are particular cases of interval-censored when individuals are evaluated in the same time-intervals, causing a great number of ties.

A common procedure for the analysis this type of data is to ignore the nature of interval-censored data, or rather, treat the random variable time as continuous, and assume that the event occurred in the beginning, midpoint or interval end, and then use a standard method of survival analysis.

In this study, Monte Carlo simulations according to Weibull model, were performed in order to compare these three procedures and a new method proposed which is a combination of the three, and is directed by the observation of time histogram versus each interval frequency in order to decide which value be used. Interval-censored data is also considered. The results show that to analyse the data exactly as interval-censored is the correct form. However, when the failure rate increase the midpoint could be used.

The discrete nature of failure time must be recognized when there are a great number of ties. Regression methods to treat grouped data are presented by Lawless (2003) and Collett (2003), whose structure is specified in terms of the probability of an individual failing in an interval, conditioned to his survival to previous interval. The models considered in literature are either those of Cox proportional hazards or the logistic one.

Weibull model is proposed in this study as an alternative to Cox model in order to adjust survival data with interval-censored in the context of discrete models. Through simulations were built the statistics of ratio likelihood and score test to distinguish between these two models.

To illustrate the simulations two applications in agronomy data were used.

Keywords: Interval-censored; Monte Carlo simulation; Weibull model

LISTA DE FIGURAS

Figura 1 - Dados Censurados do Tipo I dos Tempos para o desenvolvimento de tumor em ratos	19
Figura 2 - Dados Censurados do Tipo II dos Tempos para o desenvolvimento de tumor em ratos	20
Figura 3 - Dados Censurados do Tipo Aleatório dos Tempos de remissão de pacientes com leucemia aguda	20
Figura 4 - Funções de densidade de probabilidade e respectivas funções de taxa de falha da distribuição de Weibull para $\lambda = 100, 200$ e 300	25
Figura 5 - Funções de densidade de probabilidade e respectivas funções de taxa de falha da distribuição de Weibull para $\lambda = 1$ and $\lambda = 2.72$, usadas no estudo de simulação	43
Figura 6 - Estimativas do parâmetro $\hat{\lambda}$, considerando os dados simulados a partir da distribuição de Weibull sem covariável e 0% de censura	47
Figura 7 - Erro Quadrático Médio (EQM) para $\hat{\lambda}$, considerando os dados simulados a partir da distribuição de Weibull sem covariável e 0% de censura	48
Figura 8 - Estimativas do parâmetro $\hat{\lambda}$, considerando os dados simulados a partir da distribuição de Weibull sem covariável e 20% de censura	49
Figura 9 - Erro Quadrático Médio (EQM) para $\hat{\lambda}$, considerando os dados simulados a partir da distribuição de Weibull sem covariável e 20% de censura	50
Figura 10 - Estimativas do parâmetro $\hat{\lambda}$, considerando os dados simulados a partir da distribuição de Weibull sem covariável e 40% de censura	51
Figura 11 - Erro Quadrático Médio (EQM) para $\hat{\lambda}$, considerando os dados simulados a partir da distribuição de Weibull sem covariável e 40% de censura	52
Figura 12 - Estimativas do parâmetro $\hat{\beta}_1$, considerando os dados simulados a partir da distribuição de Weibull com covariável e 0% de censura	54
Figura 13 - Erro Quadrático Médio (EQM) para $\hat{\beta}_1$, considerando os dados simulados a partir da distribuição de Weibull com covariável e 0% de censura	55
Figura 14 - Estimativas do parâmetro $\hat{\beta}_1$, considerando os dados simulados a partir da distribuição de Weibull com covariável e 20% de censura	56

Figura 15 - Erro Quadrático Médio (EQM) para $\hat{\beta}_1$, considerando os dados simulados a partir da distribuição de Weibull com covariável e 20% de censura	57
Figura 16 - Estimativas do parâmetro $\hat{\beta}_1$, considerando os dados simulados a partir da distribuição de Weibull com covariável e 40% de censura	58
Figura 17 - Erro Quadrático Médio (EQM) para $\hat{\beta}_1$, considerando os dados simulados a partir da distribuição de Weibull com covariável e 40% de censura	59
Figura 18 - Q-Qplot para a estatística do teste da razão de verossimilhança, $\gamma = 0,5$. . .	61
Figura 19 - Q-Qplot para a estatística do teste da razão de verossimilhança, $\gamma = 1$. . .	62
Figura 20 - Q-Qplot para a estatística do teste da razão de verossimilhança, $\gamma = 2$. . .	63
Figura 21 - Q-Qplot para a estatística do teste escore, $\gamma = 0,5$	65
Figura 22 - Q-Qplot para a estatística do teste escore, $\gamma = 1$	66
Figura 23 - Q-Qplot para a estatística do teste escore, $\gamma = 2$	67
Figura 24 - Curva de Sobrevivência estimada obtida pela Tabela de Vida para os dados da mangueira	68
Figura 25 - Curvas de sobrevivência estimadas através do modelo de Weibull para cada tipo de variedade de copa considerando a média dos blocos.	71

LISTA DE TABELAS

Tabela 1 - Valores dos tempos considerados para os cinco métodos	42
Tabela 2 - Tamanho empírico em porcentagem para o teste da razão de verossimilhança com nível nominal de 0,05	60
Tabela 3 - Tamanho empírico em porcentagem para o teste escore com nível nominal de 0,05	64
Tabela 4 - Função de sobrevivência estimada e erro padrão de $\hat{S}(t)$ obtidos por meio da Tabela de vida para os dados da mangueira	69
Tabela 5 - Valores estimados para o intercepto, parâmetro de forma, tipo de copa e blocos	70
Tabela 6 - Estimativas (E) dos parâmetros (P) e erros padrões (EP), considerando os métodos de estimação M1, M2, M3, M4 e M5	72
Tabela 7 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 0,5$ e $\lambda = 1, 4$ grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 0% de censura sem covariável	81
Tabela 8 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 1$ e $\lambda = 1, 4$ grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 0% de censura sem covariável	82
Tabela 9 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 2$ e $\lambda = 1, 4$ grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 0% de censura sem covariável	83

Tabela 10 -Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 0,5$ e $\lambda = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 20% de censura sem covariável	84
Tabela 11 -Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 1$ e $\lambda = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 20% de censura sem covariável	85
Tabela 12 -Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 2$ e $\lambda = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 20% de censura sem covariável	86
Tabela 13 -Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 0,5$ e $\lambda = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 40% de censura sem covariável	87
Tabela 14 -Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 1$ e $\lambda = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 40% de censura sem covariável	88
Tabela 15 -Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 2$ e $\lambda = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 40% de censura sem covariável	89

Tabela 16 -Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 0,5$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 0% de censura com covariável	90
Tabela 17 -Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 1$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 0% de censura com covariável	91
Tabela 18 -Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 2$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 0% de censura com covariável	92
Tabela 19 -Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 0,5$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 20% de censura com covariável	93
Tabela 20 -Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 1$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 20% de censura com covariável	94
Tabela 21 -Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 2$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 20% de censura com covariável	95

Tabela 22 -Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 0,5$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 40% de censura com covariável	96
Tabela 23 -Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 1$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 40% de censura com covariável	97
Tabela 24 -Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 2$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 40% de censura com covariável	98
Tabela 25 -Resumo das estimativas do parâmetro λ do modelo Weibull para dados com censura intervalar - sem covariável	99
Tabela 26 -Resumo das estimativas do parâmetro β_1 do modelo Weibull para dados com censura intervalar - com covariável	100
Tabela 27 -Resumo das estimativas do parâmetro γ do modelo Weibull para dados com censura intervalar - sem covariável	101
Tabela 28 -Resumo das estimativas do parâmetro γ do modelo Weibull para dados com censura intervalar - com covariável	102
Tabela 29 -Dados referentes ao tempo de vida das mangueiras	103
Tabela 30 -Dados referentes ao tempo de murcha do cultivar de linho	110

1 INTRODUÇÃO

A análise de sobrevivência é usada quando a variável resposta corresponde a um tempo até a ocorrência de algum evento de interesse. Os indivíduos que estão sob estudo podem ser plantas, animais, seres humanos, produtos alimentícios, equipamentos etc. Entende-se por tempo de falha o período de tempo decorrido para o evento ocorrer e a falha é a ocorrência de um evento pré-especificado.

Uma característica importante dos dados de sobrevivência é a presença de censuras causadas pela retirada dos indivíduos mais cedo do estudo ou pelo término do experimento sem que o evento tenha ocorrido, levando a observações incompletas. Assim, para os indivíduos cujo evento não ocorreu, observa-se o tempo censurado que corresponde ao período de tempo observado desde o início do experimento até, por exemplo, seu término. Nesse caso, sabe-se somente que o tempo de falha é maior do que o tempo de censura. A presença de censura nos conjuntos de dados amostrais, requer técnicas estatísticas especializadas para acomodar a informação contida nessas observações.

Muitos experimentos ou estudos observacionais são conduzidos de modo que não é possível observar o tempo exato de ocorrência do evento, e sim somente o intervalo em que o mesmo ocorreu obtendo-se nesse caso, respostas com censura intervalar. Um caso particular da censura intervalar aparece quando os indivíduos são avaliados nos mesmos tempos, ocasionando um grande número de empates. Dados desse tipo são conhecidos como agrupados. É muito comum a ocorrência de censura intervalar em estudos agrônômicos quando as visitas às unidades de campo são especificadas entre datas distantes (COLOSIMO; CHALITA; DEMÉTRIO, 2000), em avaliações de rebanho, como por exemplo, o tempo para um ganho de peso desejável (GIOLO; HENDERSON; DEMÉTRIO, 2003), em entomologia quando a morte dos insetos são observadas diariamente (PETKAU; SITTER, 1989), e em estudos clínicos longitudinais quando a ocorrência do evento de interesse é monitorada em visitas médicas de rotina (SUN, 1996, KIM; DE GRUTTOLA; LAGAKOS, 1993) .

A análise de dados agrupados tem sido feita, ajustando-se o Modelo de Cox (1972) e considerando-se uma aproximação para a Verossimilhança Parcial, quando o número de empates é pequeno. No caso de se ter um número grande de empates, o tempo é considerado como discreto e ajustam-se modelos à probabilidade de o indivíduo falhar em um certo

intervalo, dado que ele não falhou no intervalo anterior. Nessa situação, pode-se utilizar o modelo de riscos proporcionais (PRENTICE; GLOECKLER, 1978) ou o logístico.

Essencialmente, as aplicações envolvendo dados de sobrevivência intervalar utilizam o modelo de riscos proporcionais de Cox cuja flexibilidade e robustez justificam seu uso intensivo. Uma alternativa ao modelo de Cox, pouco explorada é a utilização do modelo paramétrico Weibull e é interessante verificar a relação entre ambos.

Especificamente, os principais objetivos deste trabalho são:

1. utilizar o modelo paramétrico Weibull para modelar dados de sobrevivência com censura intervalar, sendo que esse modelo terá $(p + 2)$ parâmetros, isto é, p covariáveis observadas nos indivíduos, o intercepto e o parâmetro de forma;
2. mostrar que o modelo Weibull é um caso particular do modelo de Cox para dados de sobrevivência intervalar;
3. comparar através de simulações, métodos para tratar dados com censura intervalar considerando a distribuição Weibull;
4. implementar os métodos em programas computacionais.

2 DESENVOLVIMENTO

2.1 Conceitos Básicos

2.1.1 Funções do tempo de Sobrevivência

Em estudos de análise de sobrevivência, para cada indivíduo i , $i = 1, \dots, n$, as observações são representadas pelo par (t_i, δ_i) , em que t_i é o tempo observado de falha ou de censura e δ_i é uma variável indicadora de censura, em que $\delta_i = 1$, se o tempo observado corresponde a uma falha ou $\delta_i = 0$, se corresponde a uma censura. Se para cada indivíduo forem observadas covariáveis, $\mathbf{x}_i$, como, por exemplo, sexo, idade, tratamento, de interesse na análise, tem-se, então, o vetor $(t_i, \delta_i, \mathbf{x}_i)$.

A variável aleatória não negativa T , que representa o tempo de falha, é, usualmente, especificada pela sua função de sobrevivência ou pela função de taxa de falha ou de risco. Assumindo-se que T tem distribuição de probabilidade contínua com função de densidade de probabilidade $f(t)$, a função de distribuição acumulada é dada por:

$$F(t) = P(T < t) = \int_0^t f(u)du$$

enquanto que a função de sobrevivência, denotada por $S(t)$, é definida como a probabilidade de um indivíduo sobreviver além de um certo tempo t , isto é,

$$S(t) = P(T \geq t) = 1 - F(t) = \int_t^\infty f(u)du,$$

sendo que $S(t) = 1$ quando $t = 0$ e $S(t) = 0$ quando $t \rightarrow \infty$.

A função de taxa de falha descreve a forma com que a taxa instantânea muda com o tempo, isto é, mostra o risco de um indivíduo falhar no tempo $t + \Delta t$, com $\Delta t \rightarrow 0$, dado que ele sobreviveu ao tempo t , e é definida por:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)} = -\frac{d}{dt} \log(S(t))$$

enquanto que a função de taxa de falha acumulada é definida por:

$$H(t) = \int_0^t h(u)du.$$

2.1.2 Tipos de Censura

É importante ressaltar que, mesmo censurados, todos os resultados provenientes de um estudo de sobrevivência devem ser analisados, pois mesmo incompletas, as observações censuradas fornecem informações sobre o tempo de vida do objeto sob estudo, e sua omissão no cálculo das estatísticas de interesse acarretará em conclusões viciadas.

Existem três tipos de censura:

1. **Censura do tipo I:** é aquela em que o estudo terminará após um período pré-estabelecido de tempo. Assim, por exemplo, suponha que seis ratos foram expostos a um agente causador de câncer através da injeção de células cancerígenas. Foram observados os tempos para o desenvolvimento do tumor, sendo que a pesquisa terminou após 30 semanas. A Figura 1 mostra que os tempos de desenvolvimento dos tumores nos ratos A, B, D e F ocorreram após 10, 15, 25 e 19 semanas, respectivamente, enquanto que os ratos C e E não desenvolveram nenhum tumor até o final dos estudos. Os dados de sobrevivência foram, então 10, 15, 30+, 25, 30+, 19 semanas, com o sinal “+” representando observação censurada.
 2. **Censura do tipo II:** é aquela em que o estudo terminará após ter ocorrido o evento de interesse em um número pré-estabelecido de indivíduos. Assim, por exemplo, considerando o mesmo experimento com os seis ratos, o pesquisador decide terminar o estudo somente quando quatro dos seis ratos tiverem desenvolvido os tumores. Observando a Figura 2, os tempos de sobrevivência são 10, 15, 35+, 25, 35, 35+ semanas.
- Para as censuras do tipo I e do tipo II todos os indivíduos entram no estudo ao mesmo tempo.
3. **Censura do tipo aleatório:** acontece quando um indivíduo é retirado no decorrer do estudo sem ter ocorrido a falha. Nesse caso, os indivíduos podem entrar no estudo em diferentes tempos e as censuras podem ocorrer pelas seguintes razões:

Figura 1 - Dados Censurados do Tipo I dos Tempos para o desenvolvimento de tumor em ratos

- o paciente muda de cidade, então seu tempo de sobrevivência é maior do que ou igual ao valor observado;
- o paciente recusa-se a continuar o tratamento;
- o paciente morre de causa diferente da estudada;
- o paciente ainda está vivo no final do estudo, como acontece com a censura dos tipos I e II.

Assim, por exemplo, suponha que seis pacientes com leucemia aguda entram em um estudo pelo período de um ano, e também que todos eles respondam ao tratamento e alcancem a remissão, ou seja, após o tratamento eles ficam livres da doença. Pela Figura 3, observa-se que os pacientes A, C e E atingem a remissão no início do segundo, quarto e nono mês, e têm recidiva (volta da doença) no sexto, décimo e décimo segundo mês, respectivamente. A remissão do paciente B ocorre no início do terceiro mês, mas ele sai do estudo após quatro meses, a duração da remissão é, portanto, pelo menos de quatro meses. Para os pacientes D e F, a remissão se dá no início do quinto e décimo mês, respectivamente, e permanecem em remissão até o final do estudo, assim, seus tempos de remissão são pelo menos oito e três meses. Os respectivos tempos de remissão são 4, 4+, 6, 8+, 3, 3+ meses.

Esses três mecanismos de censura apresentados são conhecidos por censura à direita, pois o tempo de ocorrência do evento de interesse está à direita do tempo registrado. É o tipo mais frequentemente encontrado, porém podem ocorrer a censura à esquerda e a censura intervalar.

Figura 2 - Dados Censurados do Tipo II dos Tempos para o desenvolvimento de tumor em ratos

Figura 3 - Dados Censurados do Tipo Aleatório dos Tempos de remissão de pacientes com leucemia aguda

A censura à esquerda ocorre se o evento de interesse já aconteceu quando o indivíduo foi observado. Colosimo (2001) apresenta o exemplo de um estudo que foi realizado para determinar a idade em que as crianças aprendem a ler em uma determinada comunidade. No início da pesquisa, algumas crianças já sabiam ler e não se lembravam com que idade isso tinha acontecido, caracterizando, assim, censuras à esquerda. O tempo registrado para essas crianças é, portanto, maior do que o tempo real de falha. Nesse mesmo estudo, pode ocorrer simultaneamente censura à direita para crianças que não sabiam ler quando os dados foram coletados. Os tempos de vida nesse caso são chamados de duplamente censurados.

De acordo com Giolo (2003), um tipo mais geral de censura acontece quando o tempo de sobrevivência de um indivíduo, T_i , ocorre entre dois valores, isto é, $T_i \in [L_i, U_i]$, sendo que $L_i \leq T_i \leq U_i$. Para indivíduos cujo evento de interesse ocorreu entre esses dois valores, sabe-se que, o respectivo tempo de sobrevivência do indivíduo i é, no mínimo, L_i e,

no máximo, U_i . Por outro lado, sabe-se que, para aqueles indivíduos cujo evento de interesse não ocorreu até o último contato feito com os mesmos, o respectivo tempo de sobrevivência do indivíduo i pertence ao intervalo $[L_i, \infty]$ em que L_i é o tempo decorrido desde sua entrada no estudo até o último contato mantido com o mesmo e $U_i \rightarrow \infty$. Note que tempos de falha exatos bem como tempos censurados à direita e à esquerda, são casos especiais de dados de sobrevivência intervalar.

Dessa forma, na censura intervalar não se sabe o tempo exato de ocorrência do evento de interesse, sabe-se apenas que ele ocorreu dentro de um intervalo. Por exemplo, considere um estudo cujo interesse é o tempo de recorrência de um particular câncer seguido de uma cirurgia para remoção do tumor. Suponha que três meses após a cirurgia um paciente é observado e é constatado que ele está livre da doença, mas quando ele é examinado seis meses depois da cirurgia verifica-se a recorrência da mesma. O tempo de recorrência verdadeiro desse paciente não é conhecido, sabe-se apenas que ele está entre três e seis meses, e, portanto, um tempo com censura intervalar.

2.2 Modelo de Regressão de Cox

O modelo de regressão de Cox (1972), o mais importante na literatura para a análise de dados de sobrevivência, permite que a análise dos tempos de vida até a ocorrência de um evento seja realizada considerando-se as covariáveis de interesse, através da função de taxa de falha.

Assume-se, nesse modelo, que os tempos t_i , $i = 1, \dots, n$, são independentes e que a função de taxa de falha do indivíduo i , dado o vetor $\mathbf{x} = (x_1, \dots, x_p)'$ de covariáveis, é dada por:

$$h(t | \mathbf{x}_i) = h_0(t) \exp\{\boldsymbol{\beta}'\mathbf{x}_i\}, \quad (1)$$

sendo que $h_0(t)$ é o componente não-paramétrico, conhecido como função de base, uma vez que $h(t) = h_0(t)$ quando $\mathbf{x} = \mathbf{0}$; $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ é o vetor de dimensão p de coeficientes de regressão desconhecidos e $\mathbf{x}_i$ é o vetor de dimensão p de covariáveis observadas para o indivíduo i .

O modelo definido em (1) é composto pelo produto de dois componentes, um não-paramétrico e outro paramétrico. O componente não-paramétrico, $h_0(t)$, é uma função

não-negativa do tempo e nenhuma forma paramétrica é assumida para ele. O componente paramétrico, $\exp(\beta' \mathbf{x}_i)$, é usado na forma multiplicativa. Observe que a constante β_0 não aparece no componente paramétrico, pois é absorvida pelo componente não-paramétrico.

2.2.1 Estimação dos Parâmetros no Modelo de Regressão de Cox

Os coeficientes β das covariáveis no modelo de Cox são estimados a partir das observações amostrais. Entretanto, o uso do método da máxima verossimilhança fica inviabilizado, devido à presença do componente não-paramétrico $h_0(t)$.

Cox (1975) propôs a função de verossimilhança parcial, $L(\beta)$, para a estimação do vetor de parâmetros desconhecidos β . Ela é construída para uma amostra de n indivíduos em que $k \leq n$ falhas distintas ocorrem nos tempos $t_1, \dots, t_k$ e que a probabilidade condicional da i -ésima observação vir a falhar no tempo t_i , conhecendo quais as observações que estão sob risco em t_i , é:

$$\begin{aligned} \frac{h_i(t_i | \mathbf{x}_i)}{\sum_{j \in R(t_i)} h_j(t_i | \mathbf{x}_j)} &= \frac{h_0(t_i) \exp\{\beta' \mathbf{x}_i\}}{\sum_{j \in R(t_i)} h_0(t_i) \exp\{\beta' \mathbf{x}_j\}} = \\ &= \frac{h_0(t_i) \exp\{\beta' \mathbf{x}_i\}}{h_0(t_i) \sum_{j \in R(t_i)} \exp\{\beta' \mathbf{x}_j\}} = \frac{\exp\{\beta' \mathbf{x}_i\}}{\sum_{j \in R(t_i)} \exp\{\beta' \mathbf{x}_j\}}, \end{aligned} \quad (2)$$

sendo que $R(t_i) = R_i$ é o conjunto dos índices dos indivíduos sob risco no tempo t_i .

Observar que, em (2), o componente não-paramétrico $h_0(t_i)$ desaparece.

Então, a função de verossimilhança parcial a ser utilizada é dada por:

$$L(\beta) = \prod_{i=1}^n \left(\frac{\exp\{\beta' \mathbf{x}_i\}}{\sum_{j \in R_i} \exp\{\beta' \mathbf{x}_j\}} \right)^{\delta_i} \quad (3)$$

que depende apenas de β .

Os indivíduos censurados entram na função de verossimilhança parcial $L(\beta)$ através do conjunto de risco $R(t_i)$ e contribuem, para essa função, somente enquanto permanecem sob risco. Os valores de β que maximizam a função de verossimilhança parcial, $L(\beta)$, são obtidos resolvendo o sistema de equações escores, isto é, $\frac{\partial l(\beta)}{\partial \beta_k} = 0$, em que

$$\frac{\partial l(\boldsymbol{\beta})}{\partial \beta_k} = \sum_{i=1}^n \delta_i \left[x_{ik} - \frac{\sum_{j \in R_i} x_{jk} \exp\{\boldsymbol{\beta}' \mathbf{x}_j\}}{\sum_{j \in R_i} \exp\{\boldsymbol{\beta}' \mathbf{x}_j\}} \right],$$

sendo $k = 1, \dots, p$ e $l(\boldsymbol{\beta}) = \log L(\boldsymbol{\beta})$.

As propriedades assintóticas dos estimadores de máxima verossimilhança parcial são necessárias para a construção de intervalos de confiança e testes de hipóteses. Andersen e Gill (1982) apresentaram provas mais gerais sobre a consistência dos estimadores e mostraram que, assintoticamente, têm distribuição normal sob certas condições de regularidade. Podem-se, dessa forma, usar as estatísticas de Wald e da razão de verossimilhança para fazer inferências no modelo de regressão de Cox.

2.3 Modelo de Weibull

2.3.1 Distribuição de Weibull

A análise estatística de dados de sobrevivência pode ser realizada, utilizando-se os modelos probabilísticos para o tempo de falha, sendo que os mais usados para descrever a variável tempo até a ocorrência de um evento são: exponencial, Weibull e log-normal.

Segundo Collet (2003), uma distribuição de probabilidade que tem um papel central em dados de análise de sobrevivência é a distribuição de Weibull, introduzida em 1951 por Weibull no contexto de teste de confiabilidade industrial. O autor ainda ressalta que essa distribuição é tão importante para a análise paramétrica de dados de sobrevivência quanto a distribuição normal o é em modelos lineares.

Uma variável aleatória T com distribuição de Weibull com parâmetro de escala $\lambda > 0$ e parâmetro de forma $\gamma > 0$, isto é, $T \sim \text{Weibull}(\gamma, \lambda)$ tem função de densidade de probabilidade, função de sobrevivência e função de taxa de falha dadas, respectivamente, por

$$f(t) = \frac{\gamma}{\lambda^\gamma} t^{\gamma-1} \exp \left\{ - \left(\frac{t}{\lambda} \right)^\gamma \right\}, \quad t \geq 0,$$

$$S(t) = \exp \left\{ - \left(\frac{t}{\lambda} \right)^\gamma \right\}, \quad t \geq 0,$$

e

$$h(t) = \frac{\gamma}{\lambda} t^{\gamma-1}, \quad t \geq 0.$$

Note que o parâmetro λ tem a mesma unidade de medida de t e γ não tem unidade.

Uma das características importantes dessa distribuição é que ela apresenta uma grande variedade de formas, por exemplo, quando $\gamma = 1$ a função de taxa de falha é constante e os tempos de sobrevivência têm uma distribuição exponencial. Para outros valores de γ , a função de taxa de falha cresce ou decresce monotonicamente. Algumas formas das funções de densidade de probabilidade e de taxa de falha de uma variável T com distribuição de Weibull são mostradas na Figura 4. Pode-se observar, que para $\gamma = 0,5$, tem-se função de taxa de falha decrescente, enquanto que para $\gamma > 1$ as funções de taxa de falha são crescentes.

Pode-se ainda mostrar que se $T \sim \text{Weibull}$, então, $Y = \log T \sim \text{Gumbel}$, ou seja, Y tem distribuição do valor extremo, com função de densidade de probabilidade, função de sobrevivência e função de taxa de falha dadas, respectivamente, por

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

$$S(y) = \exp \left[- \exp \left(\frac{y - \mu}{\sigma} \right) \right],$$

e

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

em que y e $\mu \in \Re$ e $\sigma > 0$. Se $\mu = 0$ e $\sigma = 1$ tem-se a distribuição do valor extremo padrão. Os parâmetros μ e σ são chamados de parâmetros de localização e escala, respectivamente e relacionam-se com os parâmetros da distribuição de Weibull do seguinte modo:

$$\lambda = \left(\exp \left\{ - \left(\frac{\mu}{\sigma} \right) \right\} \right)^{-\sigma} = \exp \{ \mu \} \quad e \quad \gamma = \left(\frac{1}{\sigma} \right).$$

Na análise de dados de sobrevivência, muitas vezes é conveniente trabalhar com o logaritmo dos tempos de vida, também os pacotes estatísticos produzem, em geral, estimativas para o modelo com distribuição do valor extremo, pois pode-se escrever $\log T = \mu + \sigma w$, sendo que $w \sim \text{Gumbel}(0,1)$

Figura 4 - Funções de densidade de probabilidade e respectivas funções de taxa de falha da distribuição de Weibull para $\lambda = 100, 200$ e 300

2.3.2 Estimação dos Parâmetros

Seja uma amostra aleatória $t_1, \dots, t_n$ de uma variável aleatória T com distribuição de Weibull, em que t_i , $i = 1, \dots, n$, indica o tempo de falha ou de censura. Seja δ_i , $i = 1, \dots, n$ a variável indicadora de falha ($\delta_i = 1$) ou de censura ($\delta_i = 0$). Então, a função de verossimilhança pode ser escrita como:

$$L(\boldsymbol{\theta}; \mathbf{t}) = L(\gamma, \lambda; t_1, \dots, t_n) = \prod_{i=1}^n (f(t_i; \boldsymbol{\theta}))^{\delta_i} (S(t_i; \boldsymbol{\theta}))^{1-\delta_i}$$

$$\begin{aligned}
&= \prod_{i=1}^n \left[\frac{\gamma}{\lambda^\gamma} t_i^{\gamma-1} \exp \left\{ - \left(\frac{t_i}{\lambda} \right)^\gamma \right\} \right]^{\delta_i} \left[\exp \left\{ - \left(\frac{t_i}{\lambda} \right)^\gamma \right\} \right]^{1-\delta_i} \\
&= \prod_{i=1}^n \left[\frac{\gamma}{\lambda^\gamma} t_i^{\gamma-1} \right]^{\delta_i} \exp \left\{ - \left(\frac{t_i}{\lambda} \right)^\gamma \right\}
\end{aligned}$$

cujo logaritmo é

$$\begin{aligned}
l(\gamma, \lambda) = \log L(\gamma, \lambda) &= \sum_{i=1}^n \delta_i \log(\gamma) - \sum_{i=1}^n \delta_i \gamma \log(\lambda) + (\gamma - 1) \sum_{i=1}^n \delta_i \log(t_i) - \lambda^{(-\gamma)} \sum_{i=1}^n t_i^\gamma \\
&= r \log(\gamma) - r \gamma \log(\lambda) + (\gamma - 1) \sum_{i=1}^n \delta_i \log(t_i) - \lambda^{(-\gamma)} \sum_{i=1}^n t_i^\gamma, \quad (4)
\end{aligned}$$

pois, $r = \sum_{i=1}^n \delta_i$ é o número de falhas.

Derivando-se a expressão (4) em relação aos parâmetros λ e γ tem-se:

$$\frac{\partial l(\gamma, \lambda)}{\partial \lambda} = -\frac{r\gamma}{\lambda} + \frac{\gamma \lambda^{(-\gamma)} \sum_{i=1}^n t_i^\gamma}{\lambda} = \frac{\gamma}{\lambda} \left[-r + \lambda^{(-\gamma)} \sum_{i=1}^n t_i^\gamma \right], \quad (5)$$

$$\frac{\partial l(\gamma, \lambda)}{\partial \gamma} = \frac{r}{\gamma} - r \log(\lambda) + \sum_{i=1}^n \delta_i \log(t_i) + \lambda^{(-\gamma)} \log(\lambda) \sum_{i=1}^n t_i^\gamma - \lambda^{(-\gamma)} \sum_{i=1}^n t_i^\gamma \log(t_i). \quad (6)$$

Fazendo-se $\frac{\partial l}{\partial \gamma} = 0$ e $\frac{\partial l}{\partial \lambda} = 0$, obtém-se:

$$\hat{\lambda} = \left(\frac{\sum_{i=1}^n t_i^{\hat{\gamma}}}{r} \right)^{\frac{1}{\hat{\gamma}}}$$

e

$$\frac{1}{\hat{\gamma}} - \frac{\sum_{i=1}^n t_i^{\hat{\gamma}} \log(t_i)}{\sum_{i=1}^n t_i^{\hat{\gamma}}} = \frac{\sum_{i=1}^n \delta_i \log(t_i)}{r}$$

cujas soluções são obtidas por um método iterativo do tipo Newton-Raphson.

2.3.3 Modelo de Regressão Weibull

Suponha, no caso anterior, que os valores de p covariáveis $x_1, x_2, \dots, x_p$ são registrados para cada um dos n indivíduos. O modelo de riscos proporcionais de Weibull é dado por:

$$h(t | \mathbf{x}_i) = h_0(t) \exp\{\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}\} = h_0(t) \exp\{\boldsymbol{\beta}' \mathbf{x}_i\}, \quad (7)$$

sendo que $h_0(t) = \frac{\gamma}{\lambda^\gamma} t^{\gamma-1}$, $t \geq 0$ e, portanto,

$$h(t | \mathbf{x}_i) = \frac{\gamma}{\lambda^\gamma} t^{\gamma-1} \exp\{\boldsymbol{\beta}' \mathbf{x}_i\}. \quad (8)$$

O modelo da equação (7) tem uma aparência similar ao modelo de regressão de Cox (1), porém existe uma diferença fundamental em relação à especificação da função de base $h_0(t)$. No modelo de regressão de Cox, $h_0(t)$ é arbitrária, e no modelo em consideração, assume-se que os tempos de sobrevivência têm uma distribuição de Weibull, e isso impõe uma forma paramétrica particular sobre $h_0(t)$.

A partir da forma da função de taxa de falha (8), pode-se observar que o tempo de sobrevivência do i -ésimo indivíduo tem uma distribuição de Weibull com parâmetro de escala $\frac{\exp(\boldsymbol{\beta}' \mathbf{x}_i)}{\lambda^\gamma}$ e parâmetro de forma γ . Esse resultado mostra que o efeito das covariáveis no modelo é para alterar o parâmetro de escala da distribuição, enquanto que o parâmetro de forma se mantém constante. Pela expressão (8) tem-se que $\frac{h(t | \mathbf{x}_i)}{h_0(t | \mathbf{x}_i)} = k$, então,

$$h(t | \mathbf{x}_i) = k h_0(t | \mathbf{x}_i) = k \frac{\gamma}{\lambda^\gamma} t^{\gamma-1}.$$

Mas, chamando $\frac{1}{\lambda} = \lambda^*$, vem que $h(t | \mathbf{x}_i) = k (\lambda^*)^\gamma \gamma t^{\gamma-1}$ e supondo que $\lambda^* = \exp(\boldsymbol{\beta}' \mathbf{x}_i)$, tem-se

$$\begin{aligned} h(t | \mathbf{x}_i) &= k [\exp(\boldsymbol{\beta}' \mathbf{x}_i)]^\gamma \gamma t^{\gamma-1} \\ &= k \exp(\gamma \boldsymbol{\beta}' \mathbf{x}_i) \gamma t^{\gamma-1} \\ &= \exp(\log k) \exp(\gamma \boldsymbol{\beta}' \mathbf{x}_i) \gamma t^{\gamma-1} \\ &= \exp[\log(k) + \gamma(\boldsymbol{\beta}' \mathbf{x}_i)] \gamma t^{\gamma-1} \\ &= \exp[\gamma \beta_0 + \gamma(\boldsymbol{\beta}' \mathbf{x}_i)] \gamma t^{\gamma-1} \\ &= \exp(\boldsymbol{\beta}^{*'} \mathbf{x}_i^*) \gamma t^{\gamma-1}. \end{aligned} \quad (9)$$

A função de sobrevivência correspondente é:

$$\begin{aligned} S(t \mid \mathbf{x}_i) &= \exp \left\{ - \int_0^t \exp(\boldsymbol{\beta}^{*'} \mathbf{x}_i^*) \gamma u^{\gamma-1} du \right\} \\ &= \exp \left\{ - \exp(\boldsymbol{\beta}^{*'} \mathbf{x}_i^*) t^\gamma \right\}, \end{aligned} \quad (10)$$

enquanto que a função densidade de probabilidade pode ser obtida, diferenciando-se (10), em relação a t , e multiplicando-se o resultado por (-1) , ou a partir do resultado $f(t) = h(t)S(t)$, isto é,

$$f(t \mid \mathbf{x}_i) = \exp(\boldsymbol{\beta}^{*'} \mathbf{x}_i^*) \gamma t^{\gamma-1} \exp \left\{ - \exp(\boldsymbol{\beta}^{*'} \mathbf{x}_i^*) t^\gamma \right\}.$$

O modelo de riscos proporcionais de Weibull é ajustado pela construção da função de verossimilhança das n observações e maximizando esta função em relação aos $(p+2)$ parâmetros desconhecidos, $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ e γ .

Freitas; Borges e Ho (2003) apresentam uma abordagem para modelar dados oriundos de avaliações sensoriais visando à determinação do tempo de vida do produto em prateleira. Os autores assumem que o tempo de falha T_{ij} da j -ésima unidade avaliada no tempo τ_i (fixado) tem distribuição de Weibull. Para cada tempo fixo τ_i tem-se uma amostra aleatória de tamanho n_i de variáveis Y_{ij} , em que Y_{ij} são variáveis aleatórias Bernoulli independentes com probabilidade de falha (sucesso) p_{ij} , modelada pela função de confiabilidade do modelo paramétrico de Weibull.

Gomes (2005) utiliza o modelo de riscos proporcionais para modelar dados oriundos de avaliações sensoriais, visando estimar o tempo de vida de produto em prateleira. A autora modela a probabilidade de falha p_{ij} pela função de taxa de falha. Verifica que o modelo proposto por Freitas; Borges e Ho (2003) que é baseado no modelo de Weibull é um caso particular da modelagem proposta por ela e mostra a equivalência de parâmetros entre os dois modelos.

2.4 Modelos de Regressão para Dados de Sobrevivência Agrupados

Dados de sobrevivência agrupados são casos particulares de censura intervalar quando os indivíduos são avaliados nos mesmos tempos. Tais estudos envolvem o acompa-

nhamento de grupos de populações grandes sobre certo período de tempo para estabelecer a causa e a taxa de mortes e ou para comparar taxas de morte entre diferentes grupos de populações. Dados agrupados a partir de amostras grandes são apresentados em tabelas de vida, uma vez que isso fornece um formato para apresentação e resumo das informações. O agrupamento também poderia ser intencional, como por exemplo, para armazenagem dos dados, para proteger a privacidade de registros individuais, ou para prestar contas das limitações de instrumentos de medidas.

Uma outra razão importante para o agrupamento de dados é a dificuldade ou mesmo a impossibilidade de se obter o tempo exato de falha, por questões físicas, éticas ou restrições econômicas no delineamento da pesquisa, permitindo que o acompanhamento dos indivíduos seja feito somente periodicamente.

Duas situações em que ocorreram dados agrupados foram analisados por Chalita (1997). A primeira envolve um ensaio em delineamento em blocos ao acaso, com cinco repetições cujo objetivo era verificar a resistência das mangueiras à doença seca da mangueira. O experimento foi visitado 12 vezes durante o período de 1972 a 1992 e foi registrado se a planta estava viva ou morta. Os dados obtidos são censurados por intervalo, uma vez que o evento de interesse ocorreu entre duas visitas consecutivas e o tempo exato da morte é desconhecido. A segunda situação é de um estudo experimental realizado com um cultivar de linho, susceptível ao patógeno *Fusarium oxysporum*, plantado sobre quatro substratos. A variável resposta era o tempo até a ocorrência da murcha. A visita era feita de forma alternada, a cada três e quatro dias, proporcionando um número grande de empates.

Um procedimento comum para a análise de dados desse tipo é ignorar a natureza de censura intervalar dos dados, ou seja, tratar a variável aleatória tempo como contínua e assumir que o evento ocorreu no final (DUCROCQ 1999 e LAWLESS 2003), início (LINDSEY; RYAN 1998) ou no ponto médio (ODELL; ANDERSON; D'AGOSTINO 1992) do intervalo e então usar um método padrão de análise de sobrevivência. Alguns autores como Rücker e Messerer (1988), Odell; Anderson e D'Agostino (1992), Dorey; Little e Schenker (1993) e Lindsey e Ryan (1998) ressaltam que assumir tempos de falha intervalares como tempos exatos de falha pode conduzir a inferências inválidas.

Lindsey e Ryan (1998) comparam alguns métodos, para análise de dados com

censura intervalar, utilizando dois conjuntos de dados, um de pacientes com câncer de mama e outro com AIDS. Para testar o efeito das covariáveis, consideram os modelos não-paramétricos de Finkelstein (1986) e o de regressão de Cox (1972), assumindo tempos de sobrevivência exatos (início, ponto médio e final do intervalo), os modelos paramétricos exponencial, Weibull e log-normal e o modelo exponencial “piecewise”. Para o exemplo do câncer de mama, as análises apresentaram resultados qualitativamente similares, isto é, tratar os dados como tempos exatos de falha versus censura intervalar faz pouca diferença. Quando os dados apresentam um número maior de censuras, como o exemplo da AIDS, usar métodos não-paramétricos, como a regressão de Cox, pode induzir a conclusões imprecisas. Pelos exemplos utilizados, as autoras observaram que os métodos paramétricos são satisfatórios, especialmente se uma das escolhas for a família Weibull ou log-normal, e para permitir uma modelagem mais flexível com suposições paramétricas fracas, sugeriram o uso do modelo “piecewise” com risco constante.

Segundo Odell; Anderson e D’Agostino (1992), os métodos usuais de regressão com censura à direita, para determinar a dependência do tempo de falha T sobre as covariáveis $(x_1, \dots, x_p)$, estão amplamente disponíveis para os modelos paramétricos e não-paramétricos, e há relativamente poucas publicações para o caso de censura intervalar, em particular, quando se assume uma distribuição paramétrica.

A forma alternativa correta para tratar dados de sobrevivência com censura intervalar é considerar modelos discretos (COLLETT 2003, LAWLESS 2003).

2.4.1 Modelos Discretos de Regressão

A natureza discreta dos tempos de falha deve ser explicitamente reconhecida quando existe um grande número de empates. Métodos de regressão para tratar dados agrupados são apresentados por Lawless (2003) e Collett (2003). A estrutura de regressão é especificada em termos da probabilidade de um indivíduo falhar em um certo intervalo, condicionada à sua sobrevivência ao intervalo anterior. Os dois modelos considerados na literatura são:

1. assumir que os tempos latentes de falha vêm de um modelo de riscos proporcionais contínuo (PRENTICE; GLOECKLER, 1978), ou
2. assumir que os tempos latentes de falha vêm de um modelo de chances proporcionais (modelo logístico).

Colosimo; Chalita e Demétrio (2000) abordam essa situação do modo que se segue. Suponha que as observações são tomadas sobre n indivíduos, com variável tempo de vida T e um vetor $\mathbf{x} = (x_1, \dots, x_p)'$ de covariáveis associadas a cada indivíduo. Os tempos de vida são agrupados em k intervalos iguais, $I_j = [a_{j-1}, a_j)$, $j = 1, 2, \dots, k$, sendo que $0 = a_0 < a_1 < \dots < a_k = \infty$, assumindo que todas as censuras ocorrem no final dos intervalos. Seja R_j o conjunto sob risco no tempo a_{j-1} e δ_{ij} , uma variável indicadora, em que $\delta_{ij} = 1$, se o tempo de vida do indivíduo i ocorre dentro de I_j e $\delta_{ij} = 0$, caso contrário, $p_j(\mathbf{x}_i)$ é a probabilidade do i -ésimo indivíduo morrer (falhar) até a_j , dado que ele sobreviveu até a_{j-1} e o vetor de covariáveis regressoras $\mathbf{x}_i$, ou seja,

$$\begin{aligned} p_j(\mathbf{x}_i) &= P[a_{j-1} \leq T_i < a_j \mid T_i \geq a_{j-1}, \mathbf{x}_i] = \frac{P[a_{j-1} \leq T_i < a_j \cap T_i \geq a_{j-1}]}{P[T_i \geq a_{j-1}]} \\ &= \frac{P[a_{j-1} \leq T_i < a_j]}{P[T_i \geq a_{j-1}]} = \frac{S(a_{j-1} \mid \mathbf{x}_i) - S(a_j \mid \mathbf{x}_i)}{S(a_{j-1} \mid \mathbf{x}_i)} \end{aligned}$$

e, portanto,

$$p_j(\mathbf{x}_i) = 1 - \frac{S(a_j \mid \mathbf{x}_i)}{S(a_{j-1} \mid \mathbf{x}_i)}. \quad (11)$$

Então, a contribuição da observação não-censurada em I_j para a função de verossimilhança é:

$$\begin{aligned} P[a_{j-1} \leq T_i < a_j \mid \mathbf{x}_i] &= S(a_{j-1} \mid \mathbf{x}_i) - S(a_j \mid \mathbf{x}_i) \\ &= [\{1 - p_1(\mathbf{x}_i)\} \dots \{1 - p_{j-1}(\mathbf{x}_i)\}] p_j(\mathbf{x}_i); \end{aligned} \quad (12)$$

enquanto que a contribuição de uma observação censurada em a_j , para a função de verossimilhança é:

$$\begin{aligned} P[T_i \geq a_j \mid \mathbf{x}_i] &= S(a_j \mid \mathbf{x}_i) \\ &= [\{1 - p_1(\mathbf{x}_i)\} \dots \{1 - p_j(\mathbf{x}_i)\}]. \end{aligned} \quad (13)$$

Logo, usando-se (12) e (13), a função de verossimilhança é expressa por:

$$\prod_{j=1}^k \prod_{i \in R_j} [S(a_j | \mathbf{x}_i)]^{1-\delta_{ij}} [S(a_{j-1} | \mathbf{x}_i) - S(a_j | \mathbf{x}_i)]^{\delta_{ij}} = \prod_{j=1}^k \prod_{i \in R_j} \{p_j(\mathbf{x}_i)\}^{\delta_{ij}} \{1 - p_j(\mathbf{x}_i)\}^{1-\delta_{ij}}. \quad (14)$$

em que, $[S(a_{j-1}, \mathbf{x}_i, \boldsymbol{\beta}) - S(a_j, \mathbf{x}_i, \boldsymbol{\beta})]$, para observações não-censuradas e $[S(a_{j-1}, \mathbf{x}_i, \boldsymbol{\beta})]$, para observações censuradas à direita.

A função de verossimilhança (14) corresponde a uma distribuição de Bernoulli, cuja variável resposta δ_{ij} é binária (o indivíduo morreu, ou não, no intervalo I_j) com probabilidade de sucesso $p_j(\mathbf{x}_i)$ que pode ser modelada considerando o modelo de riscos proporcionais de Cox ou o logístico, dentre outros.

2.4.1.1 Modelo Discreto de Riscos Proporcionais de Cox

Considerando o modelo de riscos proporcionais de Cox, para o tempo de vida T , a função de taxa de falha para t dado $\mathbf{x}_i$ é:

$$h(t | \mathbf{x}_i) = h_0(t) \exp\{\boldsymbol{\beta}' \mathbf{x}_i\}$$

em que $[\exp\{\boldsymbol{\beta}' \mathbf{x}_i\} = \exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip})]$, e a função de sobrevivência é dada por:

$$\begin{aligned} S(t | \mathbf{x}_i) &= \exp\left(-\int_0^t h(u | \mathbf{x}_i) du\right) = \exp\left[-\int_0^t \exp\{\boldsymbol{\beta}' \mathbf{x}_i\} h_0(u) du\right] \\ &= \exp\left[-\exp\{\boldsymbol{\beta}' \mathbf{x}_i\} \int_0^t h_0(u) du\right] = \left[\exp\left(-\int_0^t h_0(u) du\right)\right]^{\exp\{\boldsymbol{\beta}' \mathbf{x}_i\}} \\ &= [S_0(t)]^{\exp\{\boldsymbol{\beta}' \mathbf{x}_i\}}, \end{aligned} \quad (15)$$

sendo $S_0(t)$ a função de sobrevivência base.

Logo, usando-se (15) em (11) tem-se:

$$p_j(\mathbf{x}_i) = 1 - \left[\frac{S_0(a_j)}{S_0(a_{j-1})} \right]^{\exp\{\boldsymbol{\beta}'\mathbf{x}_i\}}. \quad (16)$$

Substituindo-se (16) em (14), encontra-se a função de verossimilhança:

$$\prod_{j=1}^k \prod_{i \in R_j} \left[1 - \left(\frac{S_0(a_j)}{S_0(a_{j-1})} \right)^{\exp\{\boldsymbol{\beta}'\mathbf{x}_i\}} \right]^{\delta_{ij}} \left[\left(\frac{S_0(a_j)}{S_0(a_{j-1})} \right)^{\exp\{\boldsymbol{\beta}'\mathbf{x}_i\}} \right]^{1-\delta_{ij}}, \quad (17)$$

cujo logaritmo é dado por:

$$\log L(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \sum_{j=1}^k \sum_{i \in R_j} (\delta_{ij} \log(1 - \alpha_j^{\exp\{\boldsymbol{\beta}'\mathbf{x}_i\}}) + (1 - \delta_{ij}) \log(\alpha_j^{\exp\{\boldsymbol{\beta}'\mathbf{x}_i\}})),$$

sendo $\alpha_j = S_0(a_j)/S_0(a_{j-1})$, que pode ser maximizado em relação aos parâmetros de interesse. Entretanto, Prentice e Gloeckler (1978) sugeriram o uso da reparametrização $\tau_j = \log(-\log(\alpha_j))$ e assim, a partir de (16) tem-se:

$$1 - p_j(\mathbf{x}_i) = \left[\frac{S_0(a_j)}{S_0(a_{j-1})} \right]^{\exp\{\boldsymbol{\beta}'\mathbf{x}_i\}},$$

e, portanto,

$$\log[1 - p_j(\mathbf{x}_i)] = \exp\{\boldsymbol{\beta}'\mathbf{x}_i\} \log \left[\frac{S_0(a_j)}{S_0(a_{j-1})} \right].$$

Então,

$$\log \{-\log[1 - p_j(\mathbf{x}_i)]\} = \boldsymbol{\beta}'\mathbf{x}_i + \log \left\{ -\log \left[\frac{S_0(a_j)}{S_0(a_{j-1})} \right] \right\} = \boldsymbol{\beta}'\mathbf{x}_i + \tau_j.$$

e

$$p_j(\mathbf{x}_i) = 1 - \exp[-\exp\{\boldsymbol{\beta}'\mathbf{x}_i + \tau_j\}], \quad (18)$$

sendo $\tau_j = \log\left\{-\log\left[\frac{S_0(a_j)}{S_0(a_{j-1})}\right]\right\} = \log(-\log(\alpha_j))$; $\boldsymbol{\beta}'\mathbf{x}_i = \beta x_{i1} + \beta x_{i2} + \dots + \beta x_{ip}$ é a combinação linear das p covariáveis em $\mathbf{x}_i$; $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ é o vetor de dimensão p de coeficientes de regressão; $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})'$ é o vetor de dimensão p de covariáveis observadas para o i -ésimo indivíduo; $\boldsymbol{\tau} = (\tau_1, \dots, \tau_k)'$ é o vetor de parâmetros associados aos k intervalos de tempo.

O logaritmo da função de verossimilhança reparametrizada fica:

$$l_C(\boldsymbol{\beta}, \boldsymbol{\tau}) = \sum_{j=1}^k \sum_{i \in R_j} [-(1 - \delta_{ij}) \exp(\tau_j + \boldsymbol{\beta}'\mathbf{x}_i) + \delta_{ij} \log(1 - \exp(-\exp(\tau_j + \boldsymbol{\beta}'\mathbf{x}_i)))]. \quad (19)$$

Isso mostra que o modelo de Cox pode ser ajustado a dados com censura intervalar supondo-se que δ_{ij} tem distribuição Bernoulli com função de ligação complemento log-log sendo que o preditor linear $\eta_{ij} = \boldsymbol{\beta}'\mathbf{x}_i + \tau_j$ tem $(p + k)$ parâmetros, $\tau_1, \tau_2, \dots, \tau_k, \beta_1, \dots, \beta_p$. É um modelo bastante flexível, mas pode ficar superparametrizado quando houver muitos intervalos.

2.4.1.2 Modelo Discreto Weibull

Usando-se a expressão (10) para $S(t_j \mid \mathbf{x}_i)$ tem-se:

$$\begin{aligned} p_j(x_i) &= 1 - \frac{S(a_j \mid x_i)}{S(a_{j-1} \mid x_i)} = 1 - \frac{\exp\left\{-\left[\exp(\boldsymbol{\beta}^*'\mathbf{x}_i^*)a_j^\gamma\right]\right\}}{\exp\left\{-\left[\exp(\boldsymbol{\beta}^*'\mathbf{x}_i^*)a_{j-1}^\gamma\right]\right\}} \\ &= 1 - \exp\left\{-\exp(\boldsymbol{\beta}^*'\mathbf{x}_i^*)(a_j^\gamma - a_{j-1}^\gamma)\right\} \end{aligned}$$

e, portanto,

$$p_j(x_i) = 1 - \exp\left[-\exp(\boldsymbol{\beta}^*'\mathbf{x}_i^* + \log(a_j^\gamma - a_{j-1}^\gamma))\right] \quad (20)$$

Substituindo-se (20) em (14), encontra-se a função de verossimilhança

$$\prod_{j=1}^k \prod_{i \in R_j} \{1 - \exp[-\exp(\boldsymbol{\beta}^{*'} \mathbf{x}_i^* + \log(a_j^\gamma - a_{j-1}^\gamma))]\}^{\delta_{ij}} \{ \exp[-\exp(\boldsymbol{\beta}^{*'} \mathbf{x}_i^* + \log(a_j^\gamma - a_{j-1}^\gamma))]\}^{1-\delta_{ij}} \quad (21)$$

cujo logaritmo é dado por:

$$l_W(\boldsymbol{\beta}^*, \gamma) = \sum_{j=1}^k \sum_{i \in R_j} \{ \delta_{ij} \log [1 - \exp[-\exp(\boldsymbol{\beta}^{*'} \mathbf{x}_i^* + \log(a_j^\gamma - a_{j-1}^\gamma))]] + (1 - \delta_{ij}) [-\exp(\boldsymbol{\beta}^{*'} \mathbf{x}_i^* + \log(a_j^\gamma - a_{j-1}^\gamma))] \} \quad (22)$$

que deve ser maximizado em relação aos $(p+2)$ parâmetros, $\gamma, \beta_0^*, \beta_1^*, \dots, \beta_p^*$.

Note que, a partir de (19), tem-se:

$$\begin{aligned} 1 - p_j(x_i) &= \exp[-\exp(\boldsymbol{\beta}^{*'} \mathbf{x}_i^* + \log(a_j^\gamma - a_{j-1}^\gamma))] \\ \log[1 - p_j(x_i)] &= [-\exp(\boldsymbol{\beta}^{*'} \mathbf{x}_i^* + \log(a_j^\gamma - a_{j-1}^\gamma))] \\ \log - \log[1 - p_j(x_i)] &= \boldsymbol{\beta}^{*'} \mathbf{x}_i^* + \log(a_j^\gamma - a_{j-1}^\gamma) \\ &= \gamma \boldsymbol{\beta}' \mathbf{x}_i + \gamma \beta_0 + \log(a_j^\gamma - a_{j-1}^\gamma). \end{aligned}$$

Isso mostra que o modelo de Weibull pode ser ajustado a dados com censura intervalar, supondo-se que δ_{ij} tem distribuição Bernoulli, com função de ligação complemento log-log, sendo que o preditor linear $\eta_{ij} = \log(a_j^\gamma - a_{j-1}^\gamma) + \gamma \beta_0 + \gamma \boldsymbol{\beta}' \mathbf{x}_i$ é uma função parcialmente não linear com $(p+2)$ parâmetros.

Fica assim, verificado que o modelo de Weibull é um caso particular do modelo de riscos proporcionais de Cox, isto é, ambos podem ser ajustados usando-se δ_{ij} como variável aleatória com distribuição Bernoulli, função de ligação complemento log-log, com preditor linear $\eta = \boldsymbol{\beta}' \mathbf{x}_i + \tau_j$ em que τ_j é olhado como fator, no caso do modelo de Cox, e $\tau_j = \gamma \beta_0 + \log(a_j^\gamma - a_{j-1}^\gamma)$ é olhado como um modelo de regressão parcialmente não linear no modelo de Weibull.

2.5 Discriminação Entre os Modelos de Cox e de Weibull para Dados com Censura Intervalar

Como demonstrado na Seção 2.4.1.2, o modelo de regressão de Weibull com $(p + 2)$ parâmetros é um caso particular do modelo de regressão de riscos proporcionais de Cox com $(p + k)$ parâmetros, que é um modelo mais amplo. Para discriminação entre esses dois modelos serão usados o teste da razão de verossimilhança e a estatística escore.

As estatísticas da razão de verossimilhança, RV e do teste escore, S , serão obtidas para as hipóteses de interesse dadas por:

H_0 : O modelo de Weibull é adequado aos dados

H_A : O modelo de Weibull não é adequado aos dados, ou seja,

$$\begin{aligned} H_0 : \tau_j &= \gamma\beta_0 + \log(a_j^\gamma - a_{j-1}^\gamma) \\ H_A : &\text{pelo menos um } \tau_j \neq \gamma\beta_0 + \log(a_j^\gamma - a_{j-1}^\gamma). \end{aligned} \quad (23)$$

2.5.1 Teste da Razão de Máxima Verossimilhança

Dado que o modelo de Cox, embora semi-paramétrico, pode ser representado por um modelo paramétrico como descrito em 2.4.1.1 e, como mostrado em (19), e tem como sub-modelo o modelo de Weibull, pode-se usar o teste da razão de verossimilhança para a comparação entre eles. Tem-se portanto,

$$\begin{aligned} RV &= -2 \log \left\{ \frac{L(Weibull)}{L(Cox)} \right\} \\ &= 2 \{ \log L(Cox) - \log L(Weibull) \} \\ &= 2 l_C(\beta, \tau) - 2 l_W(\beta^*, \gamma) \end{aligned} \quad (24)$$

em que $l_C(\beta, \tau)$ e $l_W(\beta^*, \gamma)$ são dadas, respectivamente em (19) e (22).

A estatística de teste RV tem distribuição assintótica Qui-Quadrado sob a hipótese nula, com $(k - 2)$ graus de liberdade.

A hipótese H_0 será rejeitada para valores grandes de RV comparado com o valor crítico obtido da distribuição $\chi_{(k-2)}^2$ para um nível de significância nominal α fixado.

No caso de não rejeição de H_0 , o modelo mais parcimonioso, ou seja, o modelo mais simples com um número menor de parâmetros deve ser preferido.

Usando o teste da razão de verossimilhança, Gomes (2005) faz uma discriminação entre o modelo proposto por Freitas; Borges e Ho (2003) que é baseado no modelo Weibull e o modelo de riscos proporcionais, proposto por ela. Com base nos resultados obtidos, a autora constata que esse teste estatístico pode ser utilizado para discriminar os dois modelos, indicando neste caso, que o modelo Weibull foi mais adequado.

2.5.2 Teste Escore

Seja $l_C(\boldsymbol{\beta}, \boldsymbol{\tau})$ o logaritmo da função de verossimilhança para o modelo de Cox dado em (19) com $\boldsymbol{\theta} = (\tau_1, \dots, \tau_k, \beta_1, \dots, \beta_p)'$. A estatística escore é obtida pela avaliação de $\frac{\partial l_C(\boldsymbol{\beta}, \boldsymbol{\tau})}{\partial \boldsymbol{\theta}}$ sob H_0 .

Dado que $\tau_j = \gamma\beta_0 + \log(a_j^\gamma - a_{j-1}^\gamma) = \gamma\beta_0 + d_j$ em que $d_j = \log(a_j^\gamma - a_{j-1}^\gamma)$, seguindo Cox e Hinkley (1990, p. 324, 325) considere a reparametrização,

$$\tau_j = (\gamma\beta_0 + \psi_j) + d_j, \quad j \leq k-2$$

$$\tau_{k-1} = \gamma\beta_0 + d_{k-1}$$

$$\tau_k = \gamma\beta_0 + d_k$$

em que $\boldsymbol{\psi} = (\psi_1, \psi_2, \dots, \psi_{k-2})'$ com $(k-2)$ parâmetros.

Portanto, sob essa reparametrização o vetor de parâmetros passa a ser $\boldsymbol{\theta} = (\psi_1, \psi_2, \dots, \psi_{k-2}, \gamma, \beta_0, \beta_1, \dots, \beta_p)'$ e as hipóteses definidas em (23) ficam,

$$H_0: \psi_1 = \psi_2 = \dots = \psi_{k-2} = 0$$

$$H_A: \text{pelo menos um } \psi_j \neq 0, \text{ para } j = 1, \dots, k-2.$$

Logo, o logaritmo da função de verossimilhança para o modelo de Cox fica,

$$\begin{aligned}
l(\boldsymbol{\theta}) = & \sum_{j=1}^{k-2} \sum_{i \in R_j} [-(1 - \delta_{ij}) \exp((\beta_0 \gamma + \psi_j) + d_j + \boldsymbol{\beta}' \mathbf{x}_i) + \\
& \delta_{ij} \log(1 - \exp(-\exp((\beta_0 \gamma + \psi_j) + d_j + \boldsymbol{\beta}' \mathbf{x}_i)))] + \\
& \sum_{i \in R_{(k-1)}} [-(1 - \delta_{i(k-1)}) \exp(\beta_0 \gamma + d_{k-1} + \boldsymbol{\beta}' \mathbf{x}_i) + \\
& \delta_{i(k-1)} \log(1 - \exp(-\exp(\beta_0 \gamma + d_{k-1} + \boldsymbol{\beta}' \mathbf{x}_i)))] + \\
& \sum_{l \in R_k} [-(1 - \delta_{ik}) \exp(\beta_0 \gamma + d_k + \boldsymbol{\beta}' \mathbf{x}_i) + \\
& \delta_{ik} \log(1 - \exp(-\exp(\beta_0 \gamma + d_k + \boldsymbol{\beta}' \mathbf{x}_i)))]. \tag{25}
\end{aligned}$$

Inicialmente, são obtidas as estatísticas escores, $U_\psi(\boldsymbol{\theta})$, $U_{\beta_0}(\boldsymbol{\theta})$, $U_\gamma(\boldsymbol{\theta})$ e $U_\beta(\boldsymbol{\theta})$, calculando as derivadas parciais de primeira ordem da equação (25), em relação aos parâmetros $\boldsymbol{\psi}$, γ , β_0 e $\boldsymbol{\beta}$, representadas, respectivamente, por:

$$U_\psi(\boldsymbol{\theta}) = \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \psi_1}, \dots, \frac{\partial l(\boldsymbol{\theta})}{\partial \psi_{k-2}} \right)', \quad U_{\beta_0}(\boldsymbol{\theta}) = \frac{\partial l(\boldsymbol{\theta})}{\partial \beta_0}$$

e

$$U_\gamma(\boldsymbol{\theta}) = \frac{\partial l(\boldsymbol{\theta})}{\partial \gamma} \quad e \quad U_\beta(\boldsymbol{\theta}) = \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \beta_1}, \dots, \frac{\partial l(\boldsymbol{\theta})}{\partial \beta_p} \right)'.$$

Seja $I(\boldsymbol{\theta})$ a matriz de informação observada dada por:

$$I(\boldsymbol{\theta})_{(p+k)x(p+k)} = \begin{bmatrix} I_{\psi\psi}(\boldsymbol{\theta}) & I_{\psi\beta_0}(\boldsymbol{\theta}) & I_{\psi\gamma}(\boldsymbol{\theta}) & I_{\psi\beta}(\boldsymbol{\theta}) \\ I_{\beta_0\psi}(\boldsymbol{\theta}) & I_{\beta_0\beta_0}(\boldsymbol{\theta}) & I_{\beta_0\gamma}(\boldsymbol{\theta}) & I_{\beta_0\beta}(\boldsymbol{\theta}) \\ I_{\gamma\psi}(\boldsymbol{\theta}) & I_{\gamma\beta_0}(\boldsymbol{\theta}) & I_{\gamma\gamma}(\boldsymbol{\theta}) & I_{\gamma\beta}(\boldsymbol{\theta}) \\ I_{\beta\psi}(\boldsymbol{\theta}) & I_{\beta\beta_0}(\boldsymbol{\theta}) & I_{\beta\gamma}(\boldsymbol{\theta}) & I_{\beta\beta}(\boldsymbol{\theta}) \end{bmatrix}.$$

A estatística do teste escore é obtida pela avaliação de $\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ sob H_0 , isto é, a partir de

$$S = U_\psi'(0, \hat{\gamma}, \hat{\beta}_0, \hat{\boldsymbol{\beta}}) I^{\psi\psi}(\hat{\boldsymbol{\theta}}_0) U_\psi(0, \hat{\gamma}, \hat{\beta}_0, \hat{\boldsymbol{\beta}}) \tag{26}$$

em que $\hat{\boldsymbol{\theta}}_0 = (\hat{\gamma}, \hat{\beta}_0, \hat{\beta})$ é a estimativa de máxima verossimilhança obtida para o modelo de Weibull; o elemento $I^{\psi\psi}(\hat{\boldsymbol{\theta}}_0)$ segundo Rao (1973) é dado por:

$$I^{\psi\psi}(\hat{\boldsymbol{\theta}}_0) = I_{\psi\psi}(\hat{\boldsymbol{\theta}}_0)_{(k-2) \times (k-2)} - I_{\psi\phi}(\hat{\boldsymbol{\theta}}_0)_{(k-2) \times (p+2)} I_{\phi\phi}^{-1}(\hat{\boldsymbol{\theta}}_0)_{(p+2) \times (p+2)} I_{\phi\psi}(\hat{\boldsymbol{\theta}}_0)_{(p+2) \times (k-2)}$$

sendo,

$$I_{\psi\phi}(\hat{\boldsymbol{\theta}}_0) = \begin{bmatrix} I_{\psi\beta_0}(\hat{\boldsymbol{\theta}}_0) & I_{\psi\gamma}(\hat{\boldsymbol{\theta}}_0) & I_{\psi\beta}(\hat{\boldsymbol{\theta}}_0) \end{bmatrix},$$

$$I_{\phi\phi}(\hat{\boldsymbol{\theta}}_0) = \begin{bmatrix} I_{\beta_0\beta_0}(\hat{\boldsymbol{\theta}}_0) & I_{\beta_0\gamma}(\hat{\boldsymbol{\theta}}_0) & I_{\beta_0\beta}(\hat{\boldsymbol{\theta}}_0) \\ I_{\gamma\beta_0}(\hat{\boldsymbol{\theta}}_0) & I_{\gamma\gamma}(\hat{\boldsymbol{\theta}}_0) & I_{\gamma\beta}(\hat{\boldsymbol{\theta}}_0) \\ I_{\beta\beta_0}(\hat{\boldsymbol{\theta}}_0) & I_{\beta\gamma}(\hat{\boldsymbol{\theta}}_0) & I_{\beta\beta}(\hat{\boldsymbol{\theta}}_0) \end{bmatrix}$$

e

$$I_{\phi\psi}(\hat{\boldsymbol{\theta}}_0) = \begin{bmatrix} I_{\beta_0\psi}(\hat{\boldsymbol{\theta}}_0) \\ I_{\gamma\psi}(\hat{\boldsymbol{\theta}}_0) \\ I_{\beta\psi}(\hat{\boldsymbol{\theta}}_0) \end{bmatrix}.$$

Mas $I_{\psi\psi}(\hat{\boldsymbol{\theta}}_0)$, $I_{\psi\beta_0}(\hat{\boldsymbol{\theta}}_0)$, $I_{\psi\gamma}(\hat{\boldsymbol{\theta}}_0)$, $I_{\psi\beta}(\hat{\boldsymbol{\theta}}_0)$, $I_{\beta_0\beta_0}(\hat{\boldsymbol{\theta}}_0)$, $I_{\beta_0\gamma}(\hat{\boldsymbol{\theta}}_0)$, $I_{\beta_0\beta}(\hat{\boldsymbol{\theta}}_0)$, $I_{\gamma\gamma}(\hat{\boldsymbol{\theta}}_0)$, $I_{\gamma\beta}(\hat{\boldsymbol{\theta}}_0)$ e $I_{\beta\beta}(\hat{\boldsymbol{\theta}}_0)$ são as derivadas parciais de segunda ordem da expressão (25), em relação aos parâmetros $\boldsymbol{\psi}$, β_0 , γ e β . Então, tem-se:

$$I_{\psi\psi}(\hat{\boldsymbol{\theta}}_0) = \begin{bmatrix} -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \psi_1 \partial \psi_1} & \dots & -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \psi_1 \partial \psi_{k-2}} \\ \dots & \dots & \dots \\ -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \psi_{k-2} \partial \psi_1} & \dots & -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \psi_{k-2} \partial \psi_{k-2}} \end{bmatrix},$$

$$I_{\psi\beta_0}(\hat{\boldsymbol{\theta}}_0) = \left(-\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \psi_1 \partial \beta_0}, \dots, -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \psi_{k-2} \partial \beta_0} \right)', \quad I_{\psi\gamma}(\hat{\boldsymbol{\theta}}_0) = \left(-\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \psi_1 \partial \gamma}, \dots, -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \psi_{k-2} \partial \gamma} \right)',$$

$$I_{\psi\beta}(\hat{\boldsymbol{\theta}}_0) = \begin{bmatrix} -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \psi_1 \partial \beta_1} & \dots & -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \psi_1 \partial \beta_p} \\ \dots & \dots & \dots \\ -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \psi_{k-2} \partial \beta_1} & \dots & -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \psi_{k-2} \partial \beta_p} \end{bmatrix},$$

$$\begin{aligned}
I_{\beta_0\beta_0}(\hat{\boldsymbol{\theta}}_0) &= -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \beta_0 \partial \beta_0}, & I_{\gamma\gamma}(\hat{\boldsymbol{\theta}}_0) &= -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \gamma \partial \gamma}, & I_{\beta_0\gamma}(\hat{\boldsymbol{\theta}}_0) &= -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \beta_0 \partial \gamma}, \\
I_{\beta_0\beta}(\hat{\boldsymbol{\theta}}_0) &= \left(-\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \beta_0 \partial \beta_1}, \dots, -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \beta_0 \partial \beta_P} \right), & I_{\gamma\beta}(\hat{\boldsymbol{\theta}}_0) &= \left(-\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \gamma \partial \beta_1}, \dots, -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \gamma \partial \beta_p} \right) e \\
I_{\beta\beta}(\hat{\boldsymbol{\theta}}_0) &= \begin{bmatrix} -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \beta_1 \partial \beta_1} & \dots & -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \beta_1 \partial \beta_p} \\ \dots & \dots & \dots \\ -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \beta_p \partial \beta_1} & \dots & -\frac{\partial^2 l(\hat{\boldsymbol{\theta}}_0)}{\partial \beta_p \partial \beta_p} \end{bmatrix}.
\end{aligned}$$

Para essa situação a estatística do teste escore, S , tem, segundo a hipótese nula H_0 , distribuição assintótica Qui-Quadrado com $(k - 2)$ graus de liberdade. A hipótese H_0 será rejeitada para valores grandes de S comparado com o valor crítico obtido da distribuição $\chi^2_{(k-2)}$ para um nível de significância nominal α fixado.

A estatística do teste escore requer a estimação de máxima verossimilhança dos parâmetros apenas sob a hipótese H_0 , enquanto que a estatística da razão de verossimilhança exige a estimação sob ambas as hipóteses.

Segundo Cordeiro (1992), as estatísticas da razão de verossimilhança (RV) e do teste escore (S) são assintoticamente equivalentes, sob a hipótese nula H_0 , à distribuição Qui-Quadrado. O problema de escolha entre elas surge quando a estimação segundo ambas as hipóteses apresenta o mesmo grau de dificuldade. As estatísticas (RV) e (S) são invariantes em relação à parametrização da distribuição dos dados.

Chalita (1997) construiu o teste escore partindo da família assimétrica Aranda-Ordaz (1981), para discriminar entre o modelo de riscos proporcionais e o modelo logístico para o caso de dados agrupados, concluindo que os dois modelos ajustam-se bem ao conjunto de dados utilizado.

2.6 Comparação de Métodos para tratar Dados com Censura Intervalar - Um Estudo de Simulação

Apesar de alguns pacotes estatísticos acomodarem dados com censura intervalar, normalmente na prática o usuário assume que o evento ocorreu no início, no ponto

médio ou no final do intervalo, e então, aplica métodos clássicos de análise de sobrevivência.

Cinco métodos serão considerados para verificar se realmente há diferenças na obtenção das estimativas dos parâmetros. São eles:

M1- considerar que o evento ocorreu no início do intervalo;

M2- considerar que o evento ocorreu no ponto médio do intervalo;

M3- considerar que o evento ocorreu no final do intervalo;

M4- tratar os dados exatamente como sendo censura intervalar;

M5- considerar que o evento ocorreu no início, no ponto médio ou no final do intervalo.

Os métodos M1 a M3 e M4 já foram apresentados na Seção 2.4. O método M5 é uma combinação dos três primeiros métodos, sendo orientado pela observação do histograma do tempo versus a frequência de cada intervalo para a decisão de qual valor a ser usado. Para cada intervalo de tempo a proposta é observar o valor da frequência do intervalo seguinte (f_s) e do anterior (f_a) em relação à frequência daquele intervalo, adotando o seguinte critério: se $f_s - f_a > 0$, usar o valor do limite superior do intervalo; se $f_s - f_a < 0$, usar o valor do limite inferior do intervalo; se $f_s - f_a = 0$, usar o valor do ponto médio do intervalo. Considere-se, ainda, a frequência anterior ao primeiro intervalo como sendo igual a sua própria, e a frequência do seguinte em relação ao último intervalo como sendo igual a sua própria frequência.

Na Tabela 1 é apresentado um exemplo de quais valores seriam utilizados para os cinco métodos.

Simulações de Monte Carlo foram realizadas para avaliar o comportamento dos cinco métodos propostos. Inicialmente, foi gerada uma amostra a partir de uma distribuição Weibull (γ, λ) , em que γ é o parâmetro de forma, assumindo os valores 0,5; 1 e 2, e λ é o parâmetro de escala com valor igual a 1 para o caso sem covariável e $\lambda(x) = \exp(\beta_0 + \beta_1 x)$ em que $x = 0$ ou 1, $\beta_0 = 0$, $\beta_1 = 1$ para o caso com covariável. Os valores para x foram gerados a partir da distribuição Bernoulli ($p = 0.5$). Considerando cada combinação do parâmetro de forma e os casos sem e com covariável foram simulados valores de T usando 0%, 20% e 40% de censura à direita para três tamanhos amostrais 50, 100 e 200. Os valores de γ foram escolhidos para taxa de falha decrescente, constante e crescente (ver Figura 5).

Tabela 1 - Valores dos tempos considerados para os cinco métodos

Intervalo-M4	Frequência	M1	M2	M3	M5
[0; 2)	2	0	1	2	2
[2; 4)	5	2	3	4	4
[4; 6)	6	4	5	6	4
[6; 8)	4	6	7	8	6
[8; 10)	4	8	9	10	8
[10; 12)	1	10	11	12	10
[12; 14)	1	12	13	14	13

Para cada situação foram construídas quatro grades, isto é, grades com 10, 6, 4 e 3 intervalos, sendo que para todas elas o tempo inicial considerado foi o valor zero. A seguir os valores de T foram classificados e foram obtidas as frequências para cada intervalo, observando-se a ocorrência de tempos empatados. Para cada combinação tamanho de amostra, γ , sem e com covariável e proporção de censura foram feitas 1.000 simulações de Monte Carlo. Usando a proporção de censura, foi determinado o valor máximo observado de T para cada caso.

Utilizando o modelo de regressão de Weibull para cada um dos cinco métodos considerados foram obtidas as estimativas dos parâmetros $(\hat{\gamma}, \hat{\lambda})$ e $(\hat{\gamma}, \hat{\beta}_0, e \hat{\beta}_1)$, o vício $(\hat{\theta} - \theta)$ e o Erro Quadrático Médio (EQM) $((\hat{\theta} - \theta)^2 + Var(\hat{\theta}))$ das 1.000 simulações, em que θ representa qualquer um dos parâmetros. Os valores do EQM foram usados para verificar a eficiência de cada estimador em relação ao valor verdadeiro.

Todas as simulações e o ajuste do modelo de regressão de Weibull, foram realizadas usando-se o pacote estatístico R.

Figura 5 - Funções de densidade de probabilidade e respectivas funções de taxa de falha da distribuição de Weibull para $\lambda = 1$ and $\lambda = 2.72$, usadas no estudo de simulação

2.7 Avaliação do Teste da Razão de Máxima Verossimilhança e do Escore por Simulação

Para verificar o comportamento das estatísticas RV e S foram simuladas 1.000 amostras de uma distribuição de Weibull com parâmetro de forma, γ , assumindo os valores 0,5; 1 e 2 e parâmetro de escala $\lambda(x) = \exp(\beta_0 + \beta_1 x)$ em que $\beta_0 = 0$, $\beta_1 = 1$ e a covariável binária x assumia os valores zero ou um, isto é, $x = 0$ ou $x = 1$. Consideraram-se três tamanhos amostrais 50, 100 e 200 e três situações de censura 0%, 20% e 40%. Os valores gerados foram classificados em uma grade de 10 intervalos. Para cada combinação tamanho

de amostra (n), proporção de censura (pc) e γ calcularam-se as estatísticas RV e S utilizando as equações 24 e 26, respectivamente.

A seguir calculou-se o tamanho empírico dos testes em porcentagem, ou seja, foi calculado o percentual de vezes que a hipótese nula (H_0) foi rejeitada. Também foram construídos os Q-Qplots para os dois testes para verificar a adequação de uma distribuição χ^2 .

Na obtenção dos logaritmos das funções de verossimilhanças de cada modelo para o cálculo da estatística RV e estimação dos parâmetros para obtenção de S usou-se uma função para otimização baseada no método quasi-Newton “BFGS” (NOCEDAL; WRIGHT 1999) que está disponível no pacote estatístico R.

2.8 Resultados e Discussão dos Estudos de Simulação

2.8.1 Comparação de Métodos para tratar Dados com Censura Intervalar

2.8.1.1 Caso sem covariável

Os resultados obtidos para o estudo de simulação proposto em 2.6, para o parâmetro λ , encontram-se resumidos nas Figuras 6 a 11 e nas Tabelas 7 a 15 do Anexo A e 25 do Anexo C. Na Figura 7 não foi colocado os $EQMs$ dos métodos $M2$ e $M3$ para $\gamma=0,5$ por apresentarem valores grandes, o que pode ser observado na Tabela 7 do Anexo A.

Considerações para 0% de censura

O método $M4$ apresentou o melhor resultado para o Erro Quadrático Médio (EQM), pode-se verificar ainda que o EQM aumenta quando o número de intervalos diminui, o que pode ser uma consequência do aumento do número de empates, e diminui quando o tamanho amostral aumenta. Isso também ocorre para os métodos $M2$ e $M5$, quando $\gamma = 2$. Para os métodos $M1$, $M2$, $M3$ e $M5$, quando $\gamma=0,5$ e 1, o EQM aumenta quando o número de intervalos diminui e com o aumento do tamanho amostral. Nesse caso, observa-se que as amostras geradas pela distribuição de Weibull têm uma amplitude grande, concentrando a maior parte dos empates no primeiro intervalo das grades, fazendo com que o vício e a variância sejam muito grandes. Por exemplo, para uma das amostras geradas verifica-se que o mínimo é 0,0000464 e o máximo 18,48, sendo que de 50 valores gerados, 37 estavam

concentrados no primeiro intervalo da primeira grade. Estudos mais detalhados desses casos necessitam ser realizados.

Os métodos $M1$ e $M5$ têm viés negativo enquanto que o método $M3$ tem viés positivo. Os métodos $M2$ e $M4$ têm viés pequeno quando $\gamma = 2$.

Considerações para 20% de censura

De forma geral, o EQM aumenta quando o número de intervalos diminui, e também diminui com o aumento do tamanho amostral. O método $M4$ apresentou o melhor resultado para o EQM.

Os métodos $M1$ e $M5$ têm viés negativo para $\gamma=0,5$ e 1 enquanto que o método $M3$ tem viés positivo. Os métodos $M2$ e $M4$ têm viés pequeno quando $\gamma = 2$.

Considerações para 40% de censura

Para os métodos $M1$ e $M5$, o EQM aumenta quando o número de intervalos diminui e também diminui com o aumento do tamanho amostral, quando $\gamma = 0,5$ e 2. O EQM do método $M3$ apresenta resultados menores do que o método $M4$, principalmente, para $\gamma=0,5$ e 1, sendo que a variância do $M3$ foi menor do que o $M4$. Observa-se que o método $M4$ sofre influência da constante adicionada ao valor inicial zero.

O método $M1$ tem viés negativo para $\gamma=0,5$ e positivo para $\gamma=1$ e 2 enquanto que o método $M5$ tem viés negativo $\gamma=0,5$ e 1. Os métodos $M2$ e $M4$ têm viés pequeno quando $\gamma = 2$.

Considerações para 0%, 20% e 40% de censura

Quando γ aumenta existem mais semelhanças entre os resultados para os métodos $M4$ e $M2$, o que está de acordo com Odell; Anderson e D'Agostino (1992). Quando γ decresce a semelhança maior ocorre entre os métodos $M1$ e $M5$.

De forma geral, o EQM diminui com o aumento da proporção de censura para todos os métodos. Isso pode ser explicado, provavelmente, porque o viés aumenta e a variância diminui à medida que o número de intervalos diminui, provavelmente, em função do aumento do número de empates. Por outro lado, com o aumento da proporção de censuras, o número de empates diminui e com isso o viés diminui. Essa situação também necessita de um estudo

mais detalhado. Observa-se, também, uma influência grande da constante que é adicionada ao valor inicial zero, principalmente para os métodos $M1$ e $M5$ e para $\gamma = 0,5$.

Para taxa de falha crescente o método $M3$ apresenta menor EQM do que o método $M1$, ocorrendo o inverso para taxa de falha decrescente para 0% e 20% de censura. Isto ocorre porque para $\gamma = 0,5$ as falhas ocorrem mais no início dos intervalos e para $\gamma = 2$ ocorrem mais no final, devido a distribuição dos tempos de falhas para cada valor de γ .

Figura 6- Estimativas do parâmetro $\hat{\lambda}$, considerando os dados simulados a partir da distribuição de Weibull sem covariável e 0% de censura

Figura 7 - Erro Quadrático Médio (EQM) para $\hat{\lambda}$, considerando os dados simulados a partir da distribuição de Weibull sem covariável e 0% de censura

Figura 8 - Estimativas do parâmetro $\hat{\lambda}$, considerando os dados simulados a partir da distribuição de Weibull sem covariável e 20% de censura

Figura 9 - Erro Quadrático Médio (EQM) para $\hat{\lambda}$, considerando os dados simulados a partir da distribuição de Weibull sem covariável e 20% de censura

Figura 10 - Estimativas do parâmetro $\hat{\lambda}$, considerando os dados simulados a partir da distribuição de Weibull sem covariável e 40% de censura

Figura 11 - Erro Quadrático Médio (EQM) para $\hat{\lambda}$, considerando os dados simulados a partir da distribuição de Weibull sem covariável e 40% de censura

2.8.1.2 Caso com covariável

Para esse caso, os resultados obtidos para o estudo de simulação proposto em 2.6, para o parâmetro β_1 , encontram-se resumidos nas Figuras 12 a 17 e nas Tabelas 16 a 24 do Anexo B e 26 do Anexo C.

Como esperado o método $M4$ que trata os dados como censura intervalar é o melhor. À medida que γ aumenta aumentam as semelhanças entre os resultados para os métodos $M4$ e $M2$, o que está de acordo com Odell; Anderson e D'Agostino (1992). Quando γ decresce a semelhança maior ocorre entre os métodos $M1$ e $M5$.

De forma geral, o EQM aumenta quando o número de intervalos diminui, o que pode ser uma consequência do aumento do número de empates. O EQM também diminui com o aumento do tamanho amostral. Para taxa de falha crescente o método $M3$ apresenta menor EQM do que o método $M1$. Isso porque a distribuição dos tempos de falhas é assimétrica e para $\gamma > 1$ as falhas, normalmente, concentram-se no final da cauda.

O EQM aumenta com o aumento da proporção de censura para os métodos $M2$, $M3$ e $M4$, enquanto que isso nem sempre é verdadeiro para os métodos $M1$ e $M5$. Isto pode ser explicado, provavelmente, porque o viés e a variância do estimador variam em sentidos opostos, provavelmente, em função do aumento do número de empates quando o número de intervalos diminui. Por outro lado, com o aumento da proporção de censura, o número de empates diminui e com isso o viés diminui. Esse resultado também foi obtido para o caso sem covariável, e da mesma forma que lá há necessidade de um estudo mais detalhado. Observa-se também uma influência grande da constante que é adicionada ao valor inicial zero, principalmente, para os métodos $M1$ e $M5$ e para $\gamma = 0, 5$. Verifica-se ainda que quando a variável x é igual a zero tem-se um número menor de valores censurados do que quando x é igual a 1.

Os métodos $M1$ e $M5$ têm viés negativo para 0% e viés positivo para 20% e 40% de censura enquanto que o método $M3$ tem viés negativo para 20% e 40% de censura.

Os métodos $M4$ e $M2$ têm viés pequeno para 0% de censura e também para 20% e 40% de censura quando $\gamma = 2$.

Figura 12 - Estimativas do parâmetro $\hat{\beta}_1$, considerando os dados simulados a partir da distribuição de Weibull com covariável e 0% de censura

Figura 13- Erro Quadrático Médio (EQM) para $\hat{\beta}_1$, considerando os dados simulados a partir da distribuição de Weibull com covariável e 0% de censura

Figura 14- Estimativas do parâmetro $\hat{\beta}_1$, considerando os dados simulados a partir da distribuição de Weibull com covariável e 20% de censura

Figura 15 - Erro Quadrático Médio (EQM) para $\hat{\beta}_1$, considerando os dados simulados a partir da distribuição de Weibull com covariável e 20% de censura

Figura 16 - Estimativas do parâmetro $\hat{\beta}_1$, considerando os dados simulados a partir da distribuição de Weibull com covariável e 40% de censura

Figura 17 - Erro Quadrático Médio (EQM) para $\hat{\beta}_1$, considerando os dados simulados a partir da distribuição de Weibull com covariável e 40% de censura

2.8.2 Avaliação do Teste da Razão de Máxima Verossimilhança e do Escore por Simulação

2.8.2.1 Teste da Razão de Máxima Verossimilhança

A Tabela 2 apresenta os resultados do estudo de simulação para o tamanho empírico do teste da razão de verossimilhança em porcentagem, para cada combinação do tamanho da amostra, proporção de censura e do parâmetro de forma γ . Pode-se observar que o tamanho empírico do teste decresce com o aumento do tamanho da amostra e cresce com o aumento da proporção de censura para os respectivos valores de γ , estes resultados estão de acordo com Chalita (1997). Constata-se também que com o aumento do parâmetro de forma, para um mesmo tamanho de amostra, o teste tende a rejeitar a hipótese H_0 um menor número de vezes. Provavelmente com o aumento do parâmetro de forma ocorre um melhor ajuste do modelo Weibull e conseqüentemente o desempenho do teste fica melhor, o que está de acordo com Gomes (2005).

Tabela 2 - Tamanho empírico em porcentagem para o teste da razão de verossimilhança com nível nominal de 0,05

Tamanho da amostra	Proporção de censura	Tamanho do teste (%)		
		$\gamma = 0,5$	$\gamma = 1$	$\gamma = 2$
50	0	7,0	6,9	6,7
	20	7,1	7,1	6,9
	40	7,7	7,5	7,2
100	0	6,5	6,4	6,3
	20	6,9	6,7	6,5
	40	7,0	6,9	6,7
200	0	5,8	5,7	5,7
	20	6,0	5,8	5,6
	40	6,1	6,0	5,9

Pelos Q-Qplots (Figuras 18 a 20), verifica-se que a estatística RV tende a ter uma distribuição aproximadamente qui-quadrado com $(k - 2)$ graus de liberdade. Nota-se que a qualidade do ajuste melhora com o tamanho amostral, o que está de acordo com a teoria assintótica.

Figura 18 - Q-Qplot para a estatística do teste da razão de verossimilhança, $\gamma = 0,5$

Figura 19 - Q-Qplot para a estatística do teste da razão de verossimilhança, $\gamma = 1$

Figura 20 - Q-Qplot para a estatística do teste da razão de verossimilhança, $\gamma = 2$

2.8.2.2 Teste Escore

A Tabela 3 apresenta os resultados do estudo de simulação para o tamanho empírico do teste escore em porcentagem, para cada combinação do tamanho da amostra, proporção de censura e do parâmetro de forma γ . Observa-se que o tamanho empírico do teste decresce com o aumento do tamanho da amostra e cresce com o aumento da proporção

de censura para os respectivos valores de γ . Pode-se observar que o tamanho empírico do teste decresce com o aumento do tamanho da amostra e cresce com o aumento da proporção de censura para os respectivos valores de γ , estes resultados estão de acordo com Chalita (1997). Constata-se também que com o aumento do parâmetro de forma, para um mesmo tamanho de amostra, o teste tende a rejeitar a hipótese H_0 um menor número de vezes. Provavelmente com o aumento do parâmetro de forma ocorre um melhor ajuste do modelo Weibull e consequentemente o desempenho do teste fica melhor, o que está de acordo com Gomes (2005).

Pelos Q-Qplots apresentados nas Figuras 21 a 23, verifica-se que a estatística S tende a ter uma distribuição aproximadamente qui-quadrado com $(k - 2)$ graus de liberdade. Nota-se que a qualidade do ajuste melhora com o tamanho amostral, o que está de acordo com a teoria assintótica.

Tabela 3 - Tamanho empírico em porcentagem para o teste escore com nível nominal de 0,05

Tamanho da amostra	Proporção de censura	Tamanho do teste (%)		
		$\gamma = 0,5$	$\gamma = 1$	$\gamma = 2$
50	0	6,8	6,7	6,5
	20	7,0	6,9	6,8
	40	7,5	7,2	7,0
100	0	6,4	6,1	6,1
	20	6,7	6,5	6,4
	40	7,2	7,0	6,8
200	0	5,9	5,8	5,6
	20	6,3	6,0	5,9
	40	6,4	6,2	6,1

Figura 21 - Q-Qplot para a estatística do teste escore, $\gamma = 0,5$

Figura 22 - Q-Qplot para a estatística do teste escore, $\gamma = 1$

Figura 23 - Q-Qplot para a estatística do teste escore, $\gamma = 2$

2.9 Aplicações

2.9.1 Modelo Discreto Weibull

Esse exemplo foi utilizado por Chalita (1997) e refere-se a um ensaio no delineamento em blocos ao acaso com cinco repetições, no esquema fatorial 6x7, cujo objetivo era verificar a resistência das mangueiras à doença seca da mangueira. As seis variedades de copas (Extrema, Oliveira, Pahiri, Imperial, Carlota e Bourbon) foram enxertadas sobre 7 porta-enxertos. A variável resposta, tempo até a morte da mangueira, foi medida em anos e apresentou muitos empates. Foram acompanhadas 210 mangueiras no período de 1972 a 1992 e aconteceram 154 mortes em 12 visitas. No final do estudo, 56 mangueiras ainda estavam vivas.

Uma análise exploratória inicial dos dados de resistência de mangueira, sem considerar as covariáveis, através de uma tabela de vida (Tabela 4) e a correspondente figura (Figura 24) mostra que o único intervalo que não apresenta tempos de vidas empatados é o terceiro, e que os demais apresentam vários empates. A proporção de empates foi de 67,6%, e de acordo com Chalita; Colosimo e Demétrio (2002) deve-se ajustar a esses dados um modelo discreto.

Figura 24 - Curva de Sobrevivência estimada obtida pela Tabela de Vida para os dados da mangueira

Tabela 4 - Função de sobrevivência estimada e erro padrão de $\hat{S}(t)$ obtidos por meio da Tabela de vida para os dados da mangueira

Intervalo [$a_{i-1}; a_i$)	n^o de Mortes	n^o sob Risco	$\hat{S}(t)$	Erro padrão de $\hat{S}(t)$
[72; 73)	12	198	1,000	0,0160
[73; 74)	8	190	0,943	0,0203
[74; 75)	1	189	0,905	0,0207
[75; 81)	8	181	0,900	0,0238
[81; 83)	2	179	0,862	0,0245
[83; 85)	23	156	0,852	0,0302
[85; 86)	13	143	0,743	0,0322
[86; 87)	16	127	0,681	0,0337
[87; 88)	28	99	0,605	0,0344
[88; 89)	10	89	0,471	0,0341
[89; 90)	27	62	0,424	0,0315
[90; 92)	6	56	0,295	0,0305

Chalita (1997), analisando esses dados usando o modelo de Cox verificou que não existe efeito da interação entre porta-enxerto e copa e nem do porta-enxerto.

O modelo adotado aqui tem efeito de blocos, intervalos de tempo e copas, isto é,

$$p_{jsr} = 1 - [\exp(-\exp(\gamma\beta_0 + \log(a_j^\gamma - a_{j-1}^\gamma)))]^{\exp(\gamma(\alpha_s + \omega_r))}$$

sendo, $a_j - a_{j-1}$ o comprimento do j -ésimo intervalo, $j = 1, \dots, 12$; α_s é o efeito da s -ésima copa, $s = 1, \dots, 6$; ω_r é o efeito do r -ésimo bloco, $r = 1, \dots, 5$; β_0 é o intercepto e γ é o parâmetro de forma da distribuição de Weibull.

As estimativas dos parâmetros para o modelo de regressão Weibull encontram-se na Tabela 5. Observa-se que o parâmetro de forma γ é igual a 0,35, indicando uma taxa de falha decrescente.

A estatística score é $S=4,12$, o quadrado desta estatística, 16,97 concorda

com a estatística da razão de verossimilhança. Neste caso, ambos os testes evidenciam que o modelo de Weibull é adequado aos dados. Situação semelhante também foi obtida por Ridout; Hinde e Demétrio (2001) para comparar o modelo de Poisson e binomial negativa inflacionada de zero, porém, o teste escore indicava que o modelo de Poisson era inconveniente para os dados utilizados.

Tabela 5 - Valores estimados para o intercepto, parâmetro de forma, tipo de copa e blocos

Termos		Modelo de Weibull									
parâmetros	β_0	γ	α_2	α_3	α_4	α_5	α_6	ω_2	ω_3	ω_4	ω_5
estimativas	1,85	0,35	-0,81	-0,33	-0,46	-0,50	0,79	0,41	0,25	0,53	0,53

A Figura 25 mostra o gráfico das probabilidade de sobrevivência estimadas $S(t | x_i)$ versus o tempo, para cada tipo de variedades de copa ao longo do ensaio, usando-se a média dos blocos. Pode-se observar que são formados três grupos:

1. a variedade de copa de maior resistência à seca é a Oliveira;
2. o grupo de copas de resistência intermediária é composto pelas variedades Carlota, Imperial, Extrema e Pahiri;
3. a variedade de copa de menor resistência à seca é a Bourbon;

Esses resultados estão de acordo com Chalita (1997), que ajustou o modelo discreto de riscos proporcionais a esse conjunto de dados, indicando a possibilidade de se utilizar o modelo de Weibull para dados com censura intervalar.

Figura 25 - Curvas de sobrevivência estimadas através do modelo de Weibull para cada tipo de variedade de copa considerando a média dos blocos.

2.9.2 Comparação de Métodos para tratar Dados com Censura Intervalar

Essa aplicação faz parte de um estudo experimental realizado com um cultivar de linho, susceptível ao patógeno *Fusarium oxysporum* (HÖPER; STEINBERG; ALABOU-VETTE, 1995) em que 286 cultivares foram plantados em 4 substratos: M (solo natural), MI (solo natural + ilita), MK (solo natural + caolinita) e MM (solo natural + montmorilonita). O evento de interesse foi o tempo até a ocorrência da murcha causada pelo patógeno. Cada cultivar foi avaliado duas vezes por semana durante o período de 52 dias, resultando em 10 intervalos de tempo.

A Tabela 6 apresenta os resultados obtidos para o ajuste do modelo de Weibull considerando substrato como covariável e os cinco métodos de estimação. Os valores estimados para o parâmetro de forma γ são 4,93; 5,35; 5,78; 5,35 e 5,95, respectivamente, para os métodos $M1$, $M2$, $M3$, $M4$ e $M5$, ou seja, a taxa de falha é crescente. Observa-se que o método exato ($M4$) e o ponto médio ($M2$) apresentam resultados semelhantes, o ponto final do intervalo ($M3$) e o método proposto ($M5$) apresentam erros padrões menores do que o início do intervalo ($M1$). Estes resultados estão de acordo com as discussões em 2.8.1.2., isto é, o $M2$ é o melhor método aproximado uma vez que seus resultados estão muito próximos

do método exato $M4$.

Tabela 6 - Estimativas (E) dos parâmetros (P) e erros padrões (EP), considerando os métodos de estimação M1, M2, M3, M4 e M5

	M1		M2		M3		M4		M5	
P	E	E.P	E	E.P	E	E.P	E	E.P	E	E.P
β_0	3,570	0,025	3,620	0,023	3,667	0,021	3,620	0,023	3,609	0,020
β_1	0,115	0,035	0,105	0,032	0,097	0,030	0,106	0,032	0,103	0,029
β_2	0,050	0,035	0,045	0,032	0,040	0,030	0,045	0,032	0,050	0,029
β_3	0,133	0,034	0,124	0,032	0,116	0,029	0,124	0,032	0,120	0,029
$\log(\sigma)$	-1,592	0,046	-1,675	0,047	-1,752	0,047	-1,679	0,047	-1,781	0,045

3 CONCLUSÕES

Os resultados obtidos neste trabalho permitem as seguintes conclusões:

- (a) o método $M4$ (censura intervalar) é a forma correta para analisar dados com censura intervalar, considerando os três tipos de censura com covariável e sem covariável;
- (b) quando a taxa de falha é crescente, o método $M2$ (ponto médio do intervalo) é o melhor método aproximado, considerando os três tipos de censura com e sem covariável;
- (c) os métodos aproximados $M1$ e $M3$ (o início e o final do intervalo) não são recomendáveis para analisar dados com censura intervalar, o que está de acordo com Lindsey e Ryan (1998), Dorey; Little e Schenker (1993) e Odell; Anderson e D'Agostino (1992);
- (d) o método proposto $M5$ também não apresentou uma boa performance;
- (e) maiores estudos devem ser feitos a respeito da influência da constante que é adicionada ao valor inicial zero e também a respeito do viés e variância, principalmente, para $\gamma=0,5$;
- (f) a partir das estatísticas da razão de verossimilhança e escore propostas, foi possível discriminar os modelos de risco proporcionais de Cox e o de Weibull.

REFERÊNCIAS

ANDERSEN, P.K.; GILL, R.D. Cox's regression model for counting processes: a large sample study. **Annals of Statistics**, Hayward, v. 10, p. 1100 – 1200, 1982.

ARANDA-ORDAZ, F.J. On two families of transformations to additivity for binary response data. **Biometrika**, Oxford, v. 68, p. 357 – 363, 1981.

CHALITA, L.V.A.S. **Modelos para dados agrupados e censurados**. 1997. 135 p. Tese (Doutorado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1997

CHALITA, L.V.A.S.; COLOSIMO, E.A.; DEMÉTRIO, C.G.B. Likelihood approximations and discrete models for tied survival data. **Communication in Statistics - Theory and Methods**, New York, v. 31, p. 1215 – 1229, 2002.

COLLETT, A. **Modelling survival data in medical research**. 2.ed. London: Chapman and Hall, 2003. 391 p.

COLOSIMO, E.A.; CHALITA, L.V.A.S.; DEMÉTRIO, C.G.B. Tests of proportional hazards and proportional odds models for grouped survival data. **Biometrics**, Washington, DC, v. 56, p. 1233 – 1240, 2000.

COLOSIMO, E.A. Análise de sobrevivência aplicada. In: REUNIÃO ANUAL DA RBRAS, 46.; SEAGRO, 9., 2001, Piracicaba. **Resumos** Piracicaba: ESALQ/USP, 2001. 145 p.

CORDEIRO, G. Introdução à teoria de verossimilhança. In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 10., 1992, Rio de Janeiro. **Resumos** Rio de Janeiro: UFRJ, 1992. 174 p.

COX, D.R. Regression models and life-tables (with discussion). **Journal of the Royal Statistical Society, Series B**, Oxford, v. 34, p. 187 – 220, 1972.

COX, D.R. Partial likelihood. **Biometrika**, Oxford, v. 62, p. 269 – 276, 1975.

COX, D.R.; HINKLEY, D.V. **Theoretical statistics**. London: Chapman and Hall, 1990. 511 p.

DOREY, F.J.; LITTLE, R.J.A.; SCHENKER, N. Multiple imputation for threshold-crossing data with interval censoring. **Statistics in Medicine**, Chichester, v. 12, p. 1589 – 1603, 1993.

DUCROCQ, V. Extension of survival analysis models to discrete measures of longevity. In: WOKSHOP INFORMATION, Jouy en Josas, p.41 – 47, 1999. Disponível em: <<http://www.boku.ac.at/nuwi/gift/Vincent1.pdf>>. Acesso em: 18 março 2004.

FINKELSTEIN, M.D. A proportional hazards model for interval-censored failure time data. **Biometrics**, Washington, DC, v. 42, p. 845 – 854, 1986.

FREITAS, M.A.; BORGES, W.; HO, L.L. A statistical model for shelf life estimation using sensory evaluations scores. **Communications in Statistics - Theory and Methods**, New York, v. 32, n. 8, p. 1559 – 1589, 2003.

GIOLO, S.R. **Introdução à análise de sobrevivência**. Curitiba, 2003. 101 p. UFP/PR.

GIOLO, S.R.; HENDERSON, R.; DEMÉTRIO, C.G.B. Um critério para a seleção de touros nelore usando modelos de sobrevivência. **Revista de Matemática e Estatística**, São Paulo, v. 21, n. 3, p. 115 – 123, 2003.

GOMES, R.C.D. **Estimando o tempo de vida de produto em prateleira utilizando modelo de riscos proporcionais em dados oriundos de avaliações sensoriais**. 2005. 178 p. Dissertação (Mestrado em Estatística) - Universidade Federal de Minas Gerais, Belo Horizonte, 2005.

HÖPER, H.; STEINBERG, C.; ALABOUVETTE, C. Importance of physical and chemical soil properties in the suppressiveness of soils to plant diseases. **European Journal of Soil Biology**, Gauthier Villars, v. 32, p. 41 – 58, 1995.

KIM, M. Y.; DE GRUTTOLA, V. G.; LAGAKOS, S. W. Analyzing doubly censored data with covariates, with application of AIDS. **Biometrics**, Washington, DC, v. 49, p. 13 – 22, 1993.

LAWLESS, J.F. **Statistical models and methods for lifetime data**. 2.ed. New Jersey: John Wiley & Sons, 2003. 630 p.

LINDSEY, J.C.; RYAN, L.M. Tutorial in biostatistics methods for interval-censored data. **Statistics in Medicine**, Chichester, v. 17, p. 219 – 238, 1998.

NOCEDAL, J.; WRIGHT, S.J. **Numerical optimization**. New York: Springer Series in Operations Research, 1999. 636 p.

ODELL, P.M.; ANDERSON, K.M.; D'AGOSTINHO, R.B. Maximum likelihood estimation for interval-censored data using a weibull-based failure time model. **Biometrics**, Washington, DC, v. 48, p. 951 – 959, 1992.

PETKAU, A. J.; SITTER, R. R. Models for quantal response experiments over time. **Biometrics**, Washington, DC, v. 45, p. 2299 – 1308, 1989.

PRENTICE, R.L.; GLOECKLER, L.A. Regression analysis of grouped survival data with application to breast cancer data. **Biometrics**, Washington, DC, v. 34, p. 57 – 67, 1978.

R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria, 2005. Disponível em: <<http://www.r-project.org>>. Acesso em: 05 fevereiro 2005.

RAO, C.R. **Linear statistical inference and its applications**. 2.ed. New York: John Wiley & Sons, 1973. 625 p.

RIDOUT, M.; HINDE, J.; DEMÉTRIO, C.G.B. A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. **Biometrics**, Washington, DC, v. 57, p. 219 – 223, 2001.

RÜCKER, G.; MESSERER, D. Remission duration: an example of interval-censored observations. **Statistics in Medicine**, Chichester, v. 7, p. 1139 – 1145, 1988.

SUN, J. A. Nonparametric test for interval-censored failure data with application to AIDS studies. **Statistics in Medicine**, Chichester, v. 15, p. 1387 – 1395, 1996.

APÊNDICE

APÊNDICE A - Método Iterativo de Newton-Raphson

Seja $U(\boldsymbol{\theta})$ o vetor de dimensões $p \times 1$ de derivadas de primeira ordem, do logaritmo da função de verossimilhança, em relação ao vetor de parâmetros $\boldsymbol{\theta}$. Essa quantidade é conhecida como função escore, cuja notação é:

$$U(\boldsymbol{\theta}) = l'(\boldsymbol{\theta}) = \frac{\partial \log L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}.$$

Seja também, $\mathfrak{S}(\boldsymbol{\theta})$ a matriz de dimensões $p \times p$ formada pelas derivadas de segunda ordem com sinal negativo, do logaritmo da função de verossimilhança, em relação ao vetor de parâmetros $\boldsymbol{\theta}$. A matriz $\mathfrak{S}(\boldsymbol{\theta})$ é conhecida como a matriz de informação observada, com notação dada por:

$$\mathfrak{S}(\boldsymbol{\theta}) = l''(\boldsymbol{\theta}) = -\frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}.$$

De acordo com o procedimento de Newton-Raphson, uma estimativa para o vetor de parâmetros $\boldsymbol{\theta}$, no passo $(j + 1)$ do processo iterativo, é:

$$\hat{\boldsymbol{\theta}}_{j+1} = \hat{\boldsymbol{\theta}}_j + \mathfrak{S}^{-1}(\hat{\boldsymbol{\theta}}_j)U(\hat{\boldsymbol{\theta}}_j),$$

para $j = 0, 1, 2, \dots$, em que $U(\hat{\boldsymbol{\theta}}_j)$ é a função escore e $\mathfrak{S}^{-1}(\hat{\boldsymbol{\theta}}_j)$ é o inverso da matriz de informação, sendo ambos avaliados em $\hat{\boldsymbol{\theta}}_j$.

O processo pode ser iniciado, tomando-se $\hat{\boldsymbol{\theta}}_0 = 0$ e repete-se até se obter a convergência.

Para o modelo Weibull, $U(\boldsymbol{\theta})$ é um vetor de dimensão 2×1 e $\mathfrak{S}(\boldsymbol{\theta})$ é uma matriz simétrica 2×2 , consistindo dos seguintes elementos:

$$\mathfrak{S}_{11}(\gamma, \lambda) = -\frac{\partial^2 l(\gamma, \lambda)}{\partial \lambda^2} = \frac{\gamma}{\lambda^2} \left(-r + \gamma \lambda^{(-\gamma)} \sum_{i=1}^n t_i^\gamma + \lambda^{(-\gamma)} \sum_{i=1}^n t_i^\gamma \right),$$

$$\mathfrak{S}_{22}(\gamma, \lambda) = -\frac{\partial^2 l(\gamma, \lambda)}{\partial \gamma^2} =$$

$$\frac{r}{\gamma^2} + \lambda^{(-\gamma)} (\log \lambda)^2 \sum_{i=1}^n t_i^\gamma - 2\lambda^{(-\gamma)} \log(\lambda) \sum_{i=1}^n t_i^\gamma \log(t_i) + \lambda^{(-\gamma)} \sum_{i=1}^n t_i^\gamma (\log t_i)^2,$$

$$\begin{aligned} \mathfrak{S}_{12}(\lambda, \gamma) &= \mathfrak{S}_{21}(\lambda, \gamma) = -\frac{\partial^2 l(\lambda, \gamma)}{\partial \lambda \partial \gamma} \\ &= \frac{r}{\lambda} + \frac{\lambda^{(-\gamma)} \gamma \log(\lambda) \sum_{i=1}^n t_i^\gamma}{\lambda} - \frac{\lambda^{(-\gamma)} \sum_{i=1}^n t_i^\gamma}{\lambda} - \frac{\lambda^{(-\gamma)} \gamma \sum_{i=1}^n t_i^\gamma \log(t_i)}{\lambda}. \end{aligned}$$

A estimação dos parâmetros λ e γ pode ser simplificada igualando a derivada de primeira ordem em relação ao parâmetro λ a zero, que resulta em

$$\hat{\lambda} = \left(\frac{\sum_{i=1}^n t_i^{\hat{\gamma}}}{r} \right)^{\frac{1}{\hat{\gamma}}}, \quad (27)$$

que pode ser substituído na derivada de primeira ordem em relação ao parâmetro γ resultando em,

$$U(\gamma) = \frac{r}{\gamma} + \sum_{i=1}^n \delta_i \log(t_i) - \frac{r}{\sum_{i=1}^n t_i^\gamma} \sum_{i=1}^n t_i^\gamma \log(t_i).$$

e

$$\mathfrak{S}(\gamma) = \frac{r}{\gamma^2} + \frac{r}{\sum_{i=1}^n t_i^\gamma} \sum_{i=1}^n t_i^\gamma (\log t_i - \log \gamma)^2.$$

Dessa forma, pode-se aplicar o método de Newton-Raphson apenas para o parâmetro γ e substituir o valor obtido na expressão (27) para obter $\hat{\lambda}$. O passo $(j+1)$ do processo iterativo, para obtenção de $\hat{\gamma}_{j+1}$ é:

$$\hat{\gamma}_{j+1} = \hat{\gamma}_j + \mathfrak{S}^{-1}(\hat{\gamma}_j) U(\hat{\gamma}_j).$$

ANEXOS

ANEXO A - As Tabelas 7 a 9, 10 a 12 e 13 a 15 apresentam os resultados obtidos nas simulações para 0%, 20% e 40% de censura, respectivamente, sem covariável.

Tabela 7 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 0,5$ e $\lambda = 1,4$ grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 0% de censura sem covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	λ	0,22	0,625	2,49	3,386	3,96	12,609	1,03	0,182	0,22	0,619
		γ	0,35	0,024	0,98	0,238	1,28	0,617	0,53	0,012	0,34	0,025
	2	λ	0,13	0,756	3,25	7,697	5,47	29,072	1,14	0,296	0,13	0,752
		γ	0,34	0,025	1,19	0,485	1,63	1,279	0,55	0,019	0,34	0,026
	3	λ	0,09	0,836	4,15	15,116	7,27	57,695	1,26	0,502	0,09	0,833
		γ	0,35	0,023	1,45	0,924	2,08	2,512	0,59	0,033	0,34	0,025
	4	λ	0,06	0,877	5,02	24,659	9,02	94,864	1,56	1,005	0,06	0,875
		γ	0,36	0,019	1,74	1,559	2,58	4,382	0,67	0,063	0,36	0,021
100	1	λ	0,16	0,714	2,84	4,431	4,68	17,307	1,02	0,117	0,16	0,712
		γ	0,34	0,026	1,06	0,322	1,40	0,815	0,52	0,007	0,34	0,027
	2	λ	0,09	0,824	3,82	10,508	6,63	41,234	1,09	0,198	0,09	0,822
		γ	0,34	0,024	1,31	0,661	1,79	1,695	0,53	0,012	0,34	0,025
	3	λ	0,06	0,883	4,99	21,383	8,99	84,174	1,23	0,378	0,06	0,882
		γ	0,36	0,020	1,61	1,256	2,31	3,320	0,57	0,021	0,36	0,021
	4	λ	0,05	0,912	6,15	35,745	11,33	141,450	1,52	0,891	0,04	0,911
		γ	0,38	0,015	1,94	2,105	2,88	5,767	0,63	0,043	0,37	0,017
200	1	λ	0,12	0,779	3,18	5,739	5,37	22,759	1,02	0,076	0,12	0,778
		γ	0,34	0,026	1,15	0,440	1,54	1,203	0,51	0,004	0,34	0,027
	2	λ	0,06	0,866	4,36	13,696	7,74	54,777	1,07	0,142	0,07	0,865
		γ	0,35	0,032	1,46	1,162	2,06	3,042	0,53	0,007	0,35	0,032
	3	λ	0,05	0,911	5,80	28,059	10,65	112,758	1,21	0,298	0,05	0,910
		γ	0,40	0,077	1,92	3,498	2,80	8,322	0,55	0,013	0,40	0,073
	4	λ	0,03	0,932	7,22	47,345	13,49	189,508	1,48	0,724	0,03	0,932
		γ	0,48	0,210	2,58	10,327	3,97	25,096	0,60	0,028	0,47	0,189

Tabela 8 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 1$ e $\lambda = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 0% de censura sem covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	λ	0,55	0,222	1,06	0,027	1,38	0,185	1,00	0,023	0,61	0,201
		γ	0,60	0,171	1,16	0,034	1,46	0,220	1,04	0,023	0,68	0,182
	2	λ	0,35	0,438	1,12	0,043	1,60	0,433	0,99	0,028	0,37	0,425
		γ	0,49	0,267	1,27	0,079	1,71	0,513	1,05	0,030	0,52	0,281
	3	λ	0,20	0,642	1,22	0,086	1,88	0,909	0,99	0,040	0,21	0,637
		γ	0,43	0,325	1,43	0,192	2,04	1,105	1,05	0,047	0,44	0,336
	4	λ	0,13	0,767	1,32	0,159	2,17	1,572	1,01	0,057	0,13	0,766
		γ	0,42	0,342	1,62	0,400	2,42	2,059	1,14	0,084	0,41	0,343
100	1	λ	0,49	0,272	1,08	0,018	1,43	0,214	1,00	0,012	0,51	0,264
		γ	0,55	0,206	1,16	0,031	1,49	0,245	1,02	0,011	0,57	0,209
	2	λ	0,29	0,518	1,16	0,040	1,69	0,533	0,99	0,016	0,29	0,516
		γ	0,45	0,298	1,29	0,089	1,77	0,598	1,03	0,016	0,46	0,301
	3	λ	0,16	0,716	1,27	0,099	2,03	1,160	1,00	0,024	0,16	0,716
		γ	0,41	0,343	1,48	0,239	2,14	1,324	1,05	0,027	0,41	0,344
	4	λ	0,09	0,824	1,40	0,203	2,37	2,052	1,02	0,038	0,09	0,824
		γ	0,41	0,348	1,70	0,509	2,56	2,477	1,09	0,049	0,41	0,348
200	1	λ	0,44	0,322	1,09	0,015	1,48	0,250	1,00	0,006	0,44	0,320
		γ	0,52	0,232	1,18	0,034	1,53	0,282	1,01	0,006	0,52	0,233
	2	λ	0,24	0,583	1,18	0,043	1,78	0,644	1,00	0,008	0,24	0,582
		γ	0,44	0,319	1,32	0,108	1,83	0,702	1,02	0,008	0,44	0,320
	3	λ	0,12	0,769	1,32	0,122	1,38	1,426	1,00	0,015	0,12	0,769
		γ	0,41	0,351	1,53	0,296	2,24	1,567	1,03	0,016	0,41	0,352
	4	λ	0,07	0,862	1,47	0,259	1,53	2,545	1,02	0,029	0,07	0,862
		γ	0,41	0,345	1,78	0,638	2,70	2,969	1,07	0,032	0,41	0,345

Tabela 9 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 2$ e $\lambda = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 0% de censura sem covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	λ	0,86	0,026	0,99	0,005	1,12	0,019	1,00	0,005	0,99	0,006
		γ	1,66	0,202	2,06	0,067	2,35	0,189	2,09	0,074	2,26	0,180
	2	λ	0,74	0,073	0,99	0,005	1,19	0,044	0,99	0,005	0,97	0,009
		γ	1,31	0,570	2,05	0,070	2,53	0,351	2,11	0,093	2,62	0,672
	3	λ	0,58	0,185	0,99	0,006	1,29	0,091	0,99	0,006	0,92	0,041
		γ	0,92	1,231	2,02	0,063	2,71	0,579	2,14	0,114	3,08	2,362
	4	λ	0,42	0,350	0,99	0,006	1,38	0,156	0,99	0,006	0,72	0,161
		γ	0,67	1,792	2,01	0,056	2,89	0,878	2,22	0,171	2,68	3,466
100	1	λ	0,84	0,027	1,00	0,003	1,13	0,019	1,00	0,003	0,99	0,004
		γ	1,57	0,223	2,01	0,027	2,32	0,130	2,04	0,029	2,33	0,172
	2	λ	0,72	0,082	1,00	0,003	1,21	0,048	1,00	0,003	0,97	0,004
		γ	1,19	0,692	1,99	0,028	2,49	0,267	2,06	0,037	2,69	0,568
	3	λ	0,55	0,214	1,00	0,003	1,31	0,104	0,99	0,003	0,91	0,035
		γ	0,81	1,428	1,96	0,026	2,66	0,471	2,08	0,046	2,89	1,589
	4	λ	0,37	0,409	1,00	0,003	1,42	0,182	0,99	0,003	0,64	0,224
		γ	0,60	1,973	1,94	0,025	2,82	0,710	2,12	0,073	2,13	1,918
200	1	λ	0,84	0,029	1,00	0,001	1,14	0,020	1,00	0,001	0,98	0,002
		γ	1,52	0,252	1,99	0,013	2,32	0,112	2,02	0,014	2,37	0,173
	2	λ	0,70	0,090	1,00	0,001	1,23	0,053	1,00	0,001	0,98	0,002
		γ	1,12	0,790	1,97	0,014	2,48	0,243	2,03	0,017	2,75	0,606
	3	λ	0,51	0,242	1,00	0,001	1,34	0,117	1,00	0,001	0,88	0,032
		γ	0,75	1,565	1,93	0,016	2,66	0,445	2,05	0,023	2,70	0,967
	4	λ	0,33	0,461	1,00	0,002	1,45	0,205	1,00	0,002	0,56	0,297
		γ	0,56	2,084	1,90	0,019	2,80	0,657	2,08	0,036	1,78	1,999

Tabela 10 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 0,5$ e $\lambda = 1, 4$ grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 20% de censura sem covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	λ	0,83	0,135	1,16	0,098	1,33	0,169	1,04	0,109	0,86	0,132
		γ	0,43	0,008	0,74	0,061	0,94	0,197	0,52	0,010	0,42	0,008
	2	λ	0,72	0,181	1,22	0,113	1,46	0,260	1,04	0,111	0,73	0,177
		γ	0,39	0,014	0,83	0,114	1,13	0,411	0,52	0,011	0,39	0,014
	3	λ	0,61	0,245	1,28	0,140	1,60	0,396	1,05	0,115	0,62	0,239
		γ	0,36	0,020	0,94	0,203	1,40	0,817	0,53	0,014	0,36	0,021
	4	λ	0,52	0,307	1,35	0,176	1,72	0,556	1,06	0,119	0,53	0,301
		γ	0,34	0,025	1,06	0,319	1,70	1,455	0,54	0,017	0,34	0,025
100	1	λ	0,81	0,086	1,15	0,057	1,32	0,134	1,03	0,052	0,83	0,082
		γ	0,42	0,007	0,73	0,055	0,93	0,187	0,51	0,005	0,42	0,007
	2	λ	0,69	0,140	1,21	0,076	1,45	0,230	1,03	0,054	0,70	0,135
		γ	0,38	0,014	0,82	0,107	1,13	0,396	0,51	0,006	0,38	0,014
	3	λ	0,58	0,216	1,28	0,107	1,59	0,374	1,03	0,057	0,59	0,211
		γ	0,36	0,021	0,94	0,193	1,39	0,794	0,52	0,007	0,36	0,021
	4	λ	0,49	0,289	1,34	0,144	1,72	0,537	1,04	0,058	0,50	0,284
		γ	0,34	0,025	1,05	0,306	1,69	1,419	0,53	0,008	0,34	0,026
200	1	λ	0,80	0,065	1,14	0,037	1,32	0,116	1,02	0,025	0,81	0,061
		γ	0,42	0,007	0,73	0,053	0,93	0,183	0,51	0,002	0,42	0,007
	2	λ	0,68	0,125	1,20	0,057	1,45	0,214	1,02	0,026	0,69	0,121
		γ	0,38	0,014	0,82	0,104	1,12	0,390	0,51	0,003	0,38	0,014
	3	λ	0,57	0,207	1,27	0,089	1,59	0,360	1,02	0,028	0,57	0,204
		γ	0,36	0,021	0,93	0,190	1,38	0,785	0,52	0,003	0,36	0,021
	4	λ	0,48	0,286	1,34	0,128	1,72	0,525	1,03	0,028	0,48	0,285
		γ	0,34	0,026	1,05	0,301	1,68	1,404	0,52	0,004	0,34	0,026

Tabela 11 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 1$ e $\lambda = 1, 4$ grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 20% de censura sem covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	λ	0,92	0,041	1,01	0,023	1,09	0,026	1,02	0,021	0,97	0,033
		γ	0,77	0,070	1,08	0,027	1,32	0,127	1,17	0,046	0,92	0,074
	2	λ	0,84	0,066	1,02	0,022	1,14	0,034	1,02	0,021	0,91	0,054
		γ	0,64	0,141	1,13	0,035	1,50	0,275	1,21	0,062	0,81	0,151
	3	λ	0,75	0,111	1,02	0,021	1,19	0,050	1,03	0,020	0,80	0,097
		γ	0,53	0,222	1,20	0,056	1,75	0,589	1,26	0,088	0,69	0,248
	4	λ	0,64	0,176	1,03	0,020	1,25	0,072	1,04	0,020	0,68	0,163
		γ	0,46	0,288	1,27	0,088	2,02	1,083	1,31	0,012	0,57	0,314
100	1	λ	0,91	0,025	1,01	0,012	1,09	0,017	1,02	0,011	0,96	0,018
		γ	0,75	0,068	1,07	0,013	1,30	0,102	1,15	0,030	0,88	0,060
	2	λ	0,84	0,046	1,01	0,011	1,14	0,026	1,02	0,011	0,88	0,038
		γ	0,63	0,144	1,11	0,021	1,48	0,244	1,19	0,046	0,73	0,138
	3	λ	0,74	0,093	1,02	0,011	1,19	0,044	1,03	0,011	0,76	0,084
		γ	0,52	0,228	1,18	0,041	1,73	0,544	1,25	0,070	0,60	0,232
	4	λ	0,63	0,160	1,03	0,010	1,25	0,068	1,04	0,011	0,64	0,155
		γ	0,46	0,293	1,26	0,073	2,01	1,027	1,30	0,098	0,48	0,297
200	1	λ	0,91	0,017	1,01	0,006	1,08	0,012	1,01	0,006	0,95	0,012
		γ	0,75	0,067	1,06	0,008	1,30	0,094	1,15	0,025	0,85	0,056
	2	λ	0,84	0,037	1,01	0,006	1,13	0,022	1,02	0,006	0,86	0,032
		γ	0,62	0,144	1,11	0,016	1,48	0,233	1,19	0,040	0,68	0,136
	3	λ	0,73	0,084	1,02	0,006	1,19	0,041	1,03	0,006	0,74	0,080
		γ	0,52	0,230	1,18	0,035	1,72	0,526	1,24	0,062	0,54	0,229
	4	λ	0,62	0,154	1,03	0,006	1,25	0,065	1,04	0,007	0,62	0,153
		γ	0,46	0,295	1,25	0,066	2,00	1,001	1,29	0,089	0,46	0,295

Tabela 12 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 2$ e $\lambda = 1, 4$ grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 20% de censura sem covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	λ	0,95	0,010	1,00	0,006	1,05	0,007	1,00	0,006	1,00	0,006
		γ	1,71	0,169	2,05	0,098	2,38	0,261	2,06	0,101	2,14	0,157
	2	λ	0,92	0,016	1,00	0,006	1,07	0,010	1,00	0,006	1,00	0,007
		γ	1,46	0,370	2,03	0,096	2,60	0,498	2,06	0,105	2,32	0,326
	3	λ	0,88	0,029	1,00	0,006	1,11	0,015	1,00	0,006	1,01	0,010
		γ	1,16	0,762	2,02	0,095	2,89	0,975	2,06	0,113	2,59	0,696
	4	λ	0,82	0,050	0,99	0,006	1,14	0,023	1,00	0,006	1,04	0,014
		γ	0,91	1,236	2,00	0,091	3,19	1,653	2,05	0,124	3,01	1,538
100	1	λ	0,95	0,006	1,00	0,003	1,05	0,005	1,00	0,003	1,00	0,003
		γ	1,68	0,143	2,01	0,041	2,34	0,164	2,02	0,042	2,15	0,088
	2	λ	0,92	0,011	1,00	0,003	1,08	0,008	1,00	0,003	1,00	0,004
		γ	1,42	0,366	2,00	0,041	2,55	0,365	2,02	0,044	2,34	0,227
	3	λ	0,88	0,022	1,00	0,003	1,11	0,014	1,00	0,003	1,00	0,007
		γ	1,13	0,782	1,98	0,041	2,84	0,788	2,02	0,049	2,57	0,489
	4	λ	0,82	0,041	0,99	0,003	1,14	0,022	1,00	0,003	1,04	0,009
		γ	0,88	1,275	1,96	0,040	3,14	1,391	2,02	0,053	3,03	1,200
200	1	λ	0,95	0,004	1,00	0,001	1,05	0,003	1,00	0,002	1,00	0,002
		γ	1,67	0,129	2,00	0,020	2,33	0,131	2,01	0,020	2,19	0,070
	2	λ	0,92	0,009	1,00	0,002	1,08	0,007	1,00	0,002	1,00	0,002
		γ	1,41	0,362	1,98	0,019	2,54	0,319	2,01	0,021	2,38	0,197
	3	λ	0,87	0,019	1,00	0,002	1,11	0,013	1,00	0,002	0,99	0,005
		γ	1,12	0,790	1,97	0,020	2,82	0,717	2,01	0,023	2,53	0,380
	4	λ	0,82	0,037	0,99	0,002	1,14	0,022	1,00	0,002	1,03	0,006
		γ	0,87	1,286	1,95	0,021	3,12	1,302	2,01	0,025	3,01	1,076

Tabela 13 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 0,5$ e $\lambda = 1, 4$ grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 40% de censura sem covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	λ	1,00	0,171	0,97	0,074	0,98	0,047	0,95	0,054	1,02	0,172
		γ	0,48	0,004	0,70	0,047	0,89	0,157	0,83	0,117	0,48	0,005
	2	λ	0,98	0,200	0,96	0,060	0,98	0,032	0,95	0,046	1,00	0,204
		γ	0,44	0,006	0,78	0,085	1,07	0,329	0,90	0,168	0,44	0,008
	3	λ	0,95	0,233	0,95	0,048	0,99	0,022	0,96	0,040	0,96	0,236
		γ	0,40	0,011	0,88	0,146	1,30	0,656	0,98	0,234	0,40	0,012
	4	λ	0,91	0,263	0,95	0,040	1,01	0,015	0,96	0,034	0,92	0,268
		γ	0,38	0,016	0,97	0,227	1,58	1,177	1,05	0,308	0,38	0,016
100	1	λ	0,97	0,077	0,95	0,037	0,97	0,024	0,94	0,028	0,99	0,078
		γ	0,47	0,002	0,69	0,040	0,88	0,146	0,82	0,108	0,47	0,002
	2	λ	0,94	0,092	0,95	0,031	0,98	0,017	0,94	0,025	0,96	0,092
		γ	0,43	0,006	0,77	0,077	1,05	0,312	0,89	0,158	0,43	0,006
	3	λ	0,90	0,109	0,94	0,026	0,99	0,011	0,95	0,021	0,91	0,110
		γ	0,40	0,011	0,87	0,138	1,29	0,633	0,97	0,223	0,40	0,011
	4	λ	0,86	0,126	0,94	0,022	1,01	0,008	0,96	0,018	0,87	0,127
		γ	0,37	0,016	0,96	0,216	1,57	1,143	1,04	0,295	0,37	0,016
200	1	λ	0,95	0,038	0,94	0,020	0,96	0,013	0,94	0,016	0,96	0,038
		γ	0,47	0,002	0,69	0,038	0,88	0,142	0,82	0,105	0,47	0,002
	2	λ	0,92	0,047	0,94	0,018	0,97	0,009	0,94	0,014	0,93	0,047
		γ	0,43	0,005	0,77	0,075	1,05	0,306	0,89	0,155	0,43	0,005
	3	λ	0,88	0,061	0,94	0,015	0,99	0,006	0,94	0,012	0,88	0,060
		γ	0,40	0,011	0,86	0,134	1,26	0,624	0,97	0,219	0,40	0,011
	4	λ	0,83	0,077	0,94	0,013	1,00	0,004	0,95	0,011	0,83	0,076
		γ	0,37	0,016	0,96	0,213	1,56	1,131	1,04	0,290	0,37	0,016

Tabela 14 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 1$ e $\lambda = 1, 4$ grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 40% de censura sem covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	λ	1,03	0,058	1,00	0,030	1,01	0,023	0,99	0,025	0,69	0,103
		γ	0,82	0,056	1,08	0,033	1,31	0,130	1,18	0,056	1,53	0,429
	2	λ	1,05	0,083	1,00	0,028	1,01	0,016	0,99	0,023	0,66	0,127
		γ	0,70	0,109	1,12	0,040	1,49	0,278	1,23	0,078	1,41	0,458
	3	λ	1,06	0,126	0,99	0,026	1,01	0,012	0,98	0,022	0,60	0,175
		γ	0,58	0,185	1,17	0,051	1,73	0,577	1,28	0,105	1,26	0,518
	4	λ	1,04	0,161	0,98	0,024	1,02	0,009	0,98	0,020	0,54	0,229
		γ	0,51	0,249	1,23	0,074	2,01	1,083	1,34	0,140	1,14	0,650
100	1	λ	1,02	0,027	1,00	0,015	1,00	0,010	0,98	0,013	0,68	0,103
		γ	0,80	0,050	1,06	0,015	1,29	0,097	1,16	0,036	1,47	0,314
	2	λ	1,04	0,038	0,99	0,014	1,01	0,008	0,98	0,012	0,64	0,132
		γ	0,68	0,110	1,10	0,020	1,46	0,231	1,21	0,055	1,30	0,271
	3	λ	1,04	0,054	0,99	0,013	1,01	0,006	0,98	0,011	0,58	0,184
		γ	0,57	0,189	1,15	0,032	1,70	0,511	1,27	0,081	1,11	0,258
	4	λ	1,01	0,069	0,98	0,012	1,02	0,005	0,98	0,010	0,52	0,243
		γ	0,50	0,255	1,21	0,053	1,98	0,988	1,32	0,113	0,96	0,292
200	1	λ	1,02	0,013	1,00	0,007	1,00	0,005	0,98	0,006	0,68	0,104
		γ	0,79	0,048	1,05	0,008	1,28	0,086	1,16	0,029	1,43	0,258
	2	λ	1,03	0,019	0,99	0,007	1,00	0,004	0,98	0,006	0,64	0,137
		γ	0,67	0,110	1,09	0,013	1,46	0,216	1,21	0,048	1,24	0,185
	3	λ	1,02	0,026	0,98	0,007	1,01	0,003	0,97	0,006	0,57	0,193
		γ	0,57	0,189	1,14	0,025	1,69	0,492	1,26	0,073	1,01	0,129
	4	λ	1,00	0,033	0,97	0,006	1,02	0,003	0,97	0,005	0,50	0,256
		γ	0,49	0,256	1,21	0,047	1,97	0,963	1,32	0,105	0,84	0,131

Tabela 15 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 2$ e $\lambda = 1, 4$ grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 40% de censura sem covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	λ	1,00	0,012	1,00	0,008	1,01	0,006	1,00	0,007	1,00	0,008
		γ	1,72	0,181	2,06	0,130	2,42	0,352	2,09	0,130	2,13	0,189
	2	λ	1,00	0,016	1,00	0,008	1,02	0,005	1,00	0,007	1,00	0,008
		γ	1,48	0,362	2,04	0,127	2,67	0,647	2,11	0,131	2,32	0,394
	3	λ	1,01	0,025	1,00	0,008	1,03	0,004	1,00	0,007	1,01	0,009
		γ	1,20	0,722	2,02	0,129	3,01	1,352	2,12	0,141	2,65	0,986
	4	λ	1,03	0,041	1,00	0,008	1,05	0,004	1,00	0,007	1,02	0,009
		γ	0,96	1,134	2,01	0,122	3,43	2,537	2,15	0,153	3,17	2,122
100	1	λ	1,00	0,006	1,00	0,004	1,01	0,003	1,00	0,004	1,00	0,004
		γ	1,68	0,150	2,01	0,054	2,36	0,205	2,05	0,053	2,14	0,096
	2	λ	1,00	0,008	1,00	0,004	1,02	0,003	1,00	0,004	1,00	0,004
		γ	1,44	0,356	1,99	0,054	2,61	0,463	2,06	0,054	2,35	0,270
	3	λ	1,01	0,013	1,00	0,004	1,04	0,003	1,00	0,004	1,00	0,005
		γ	1,16	0,732	1,98	0,053	2,94	1,018	2,08	0,058	2,67	0,713
	4	λ	1,03	0,021	1,00	0,004	1,05	0,003	1,00	0,004	1,01	0,005
		γ	0,94	1,171	1,96	0,053	3,34	1,965	2,11	0,066	3,12	1,552
200	1	λ	1,00	0,003	1,00	0,002	1,01	0,002	1,00	0,002	1,00	0,002
		γ	1,67	0,133	2,00	0,025	2,35	0,156	2,03	0,025	2,18	0,073
	2	λ	1,00	0,004	1,00	0,002	1,02	0,002	1,00	0,002	0,99	0,002
		γ	1,43	0,349	1,98	0,026	2,59	0,394	2,05	0,026	2,40	0,229
	3	λ	1,00	0,006	1,00	0,002	1,03	0,002	1,00	0,002	1,00	0,003
		γ	1,15	0,733	1,97	0,026	2,93	0,919	2,07	0,029	2,72	0,656
	4	λ	1,02	0,010	1,00	0,002	1,05	0,003	1,00	0,002	1,01	0,003
		γ	0,92	1,182	1,95	0,027	3,31	1,793	2,09	0,034	3,11	1,409

ANEXO B - As Tabelas 16 a 18, 19 a 21 e 22 a 24 mostram os resultados obtidos nas simulações para 0%, 20% e 40% de censura, respectivamente, com covariável.

Tabela 16 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 0,5$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 0% de censura com covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	β_1	0,72	0,788	1,02	0,447	1,03	0,522	1,00	0,349	0,72	0,806
		γ	0,33	0,029	0,93	0,183	1,19	0,482	0,55	0,011	0,33	0,030
	2	β_1	0,64	0,967	1,03	0,513	1,04	0,591	1,01	0,390	0,64	0,978
		γ	0,32	0,032	1,11	0,371	1,50	0,999	0,59	0,012	0,32	0,034
	3	β_1	0,57	1,061	1,04	0,579	1,04	0,651	1,02	0,453	0,57	1,052
		γ	0,32	0,032	1,34	0,703	1,90	1,957	0,65	0,013	0,32	0,034
	4	β_1	0,53	1,076	1,04	0,617	1,05	0,685	1,03	0,506	0,52	1,050
		γ	0,33	0,030	1,59	1,188	2,35	3,423	0,74	0,016	0,32	0,032
100	1	β_1	0,70	0,763	1,01	0,264	1,01	0,325	1,01	0,183	0,70	0,771
		γ	0,32	0,033	0,99	0,239	1,28	0,618	0,54	0,001	0,32	0,034
	2	β_1	0,61	0,915	1,01	0,321	1,01	0,382	1,00	0,208	0,61	0,914
		γ	0,32	0,034	1,19	0,486	1,63	1,280	0,56	0,011	0,31	0,035
	3	β_1	0,52	1,071	1,01	0,370	1,02	0,427	1,00	0,243	0,52	1,049
		γ	0,32	0,031	1,46	0,923	2,08	2,508	0,60	0,022	0,32	0,033
	4	β_1	0,46	1,043	1,01	0,409	1,02	0,458	1,00	0,289	0,46	1,011
		γ	0,34	0,027	1,74	1,559	2,59	4,381	0,68	0,051	0,33	0,029
200	1	β_1	0,66	0,727	1,00	0,169	1,00	0,226	1,00	0,085	0,66	0,732
		γ	0,31	0,035	1,06	0,315	1,39	0,798	0,53	0,004	0,31	0,035
	2	β_1	0,56	0,897	1,00	0,222	1,00	0,280	1,00	0,102	0,56	0,892
		γ	0,32	0,033	1,30	0,644	1,78	1,658	0,55	0,006	0,32	0,034
	3	β_1	0,49	0,930	1,00	0,271	1,00	0,323	1,00	0,130	0,48	0,911
		γ	0,33	0,028	1,60	1,220	2,29	3,237	0,59	0,014	0,33	0,029
	4	β_1	0,44	0,937	1,00	0,305	1,00	0,350	1,00	0,165	0,43	0,906
		γ	0,35	0,023	1,92	2,037	2,86	5,599	0,65	0,034	0,35	0,024

Tabela 17 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 1$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 0% de censura com covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	β_1	0,95	0,148	1,00	0,078	1,01	0,087	1,00	0,077	0,94	0,173
		γ	0,60	0,176	1,16	0,035	1,43	0,200	1,05	0,026	0,70	0,156
	2	β_1	0,89	0,316	1,01	0,080	1,01	0,096	1,00	0,078	0,88	0,371
		γ	0,48	0,277	1,26	0,075	1,67	0,459	1,08	0,035	0,53	0,260
	3	β_1	0,80	0,567	1,01	0,090	1,01	0,111	1,00	0,085	0,79	0,621
		γ	0,41	0,347	1,40	0,168	1,98	0,974	1,13	0,054	0,42	0,343
	4	β_1	0,70	0,836	1,01	0,099	1,01	0,122	1,00	0,095	0,69	0,775
		γ	0,39	0,375	1,58	0,337	2,33	1,792	1,22	0,102	0,39	0,377
100	1	β_1	0,95	0,117	1,00	0,038	1,00	0,044	1,00	0,038	0,96	0,137
		γ	0,54	0,214	1,16	0,028	1,45	0,212	1,03	0,012	0,59	0,192
	2	β_1	0,88	0,293	1,00	0,040	1,00	0,051	1,00	0,039	0,87	0,315
		γ	0,44	0,318	1,27	0,075	1,71	0,507	1,05	0,017	0,45	0,313
	3	β_1	0,77	0,533	1,00	0,047	1,00	0,063	1,00	0,044	0,77	0,541
		γ	0,39	0,374	1,43	0,189	2,05	1,102	1,08	0,027	0,39	0,375
	4	β_1	0,69	0,771	1,01	0,053	1,01	0,071	1,00	0,048	0,69	0,760
		γ	0,38	0,389	1,62	0,396	2,43	2,048	1,15	0,053	0,38	0,390
200	1	β_1	0,94	0,098	1,00	0,018	1,00	0,024	1,00	0,018	0,94	0,107
		γ	0,50	0,250	1,16	0,028	1,49	0,024	1,02	0,006	0,52	0,242
	2	β_1	0,86	0,280	1,00	0,020	1,00	0,030	1,00	0,018	0,86	0,283
		γ	0,41	0,346	1,29	0,086	1,76	0,059	1,03	0,008	0,41	0,347
	3	β_1	0,75	0,533	1,00	0,024	1,00	0,039	1,00	0,020	0,74	0,531
		γ	0,38	0,388	1,48	0,023	2,13	1,297	1,06	0,015	0,38	0,389
	4	β_1	0,64	0,717	1,00	0,030	1,00	0,048	1,00	0,023	0,64	0,716
		γ	0,37	0,392	1,69	0,049	2,55	2,415	1,11	0,031	0,37	0,392

Tabela 18 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 2$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 0% de censura com covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	β_1	1,00	0,020	1,00	0,019	1,00	0,019	1,00	0,019	1,00	0,020
		γ	1,69	0,189	2,10	0,080	2,38	0,219	2,13	0,089	2,15	0,141
	2	β_1	0,99	0,024	1,00	0,019	1,00	0,020	1,00	0,019	1,00	0,026
		γ	1,34	0,541	2,10	0,083	2,57	0,400	2,17	0,111	2,39	0,458
	3	β_1	0,98	0,053	1,00	0,020	1,01	0,021	1,00	0,020	0,98	0,081
		γ	0,92	1,218	2,08	0,074	2,76	0,663	2,22	0,143	2,67	1,776
	4	β_1	0,95	0,140	1,00	0,021	1,01	0,023	1,00	0,021	0,91	0,376
		γ	0,67	1,804	2,09	0,074	2,97	1,039	2,33	0,239	2,49	3,637
100	1	β_1	1,00	0,010	1,00	0,009	1,00	0,009	1,00	0,009	1,00	0,010
		γ	1,59	0,209	2,04	0,033	2,34	0,145	2,07	0,037	2,22	0,112
	2	β_1	1,00	0,015	1,00	0,009	1,00	0,010	1,00	0,010	1,00	0,013
		γ	1,21	0,672	2,03	0,032	2,51	0,293	2,09	0,046	2,54	0,454
	3	β_1	0,98	0,041	1,00	0,010	1,00	0,011	1,00	0,010	0,99	0,075
		γ	0,81	1,441	2,00	0,029	2,69	0,511	2,13	0,060	2,77	1,553
	4	β_1	0,95	0,134	1,00	0,011	1,00	0,012	1,00	0,011	0,97	0,792
		γ	0,58	2,010	1,99	0,026	2,87	0,801	2,21	0,102	2,03	1,814
200	1	β_1	1,00	0,005	1,00	0,004	1,00	0,004	1,00	0,004	1,00	0,005
		γ	1,52	0,252	2,01	0,013	2,31	0,112	2,03	0,015	2,31	0,126
	2	β_1	1,00	0,009	1,00	0,004	1,00	0,005	1,00	0,004	1,00	0,006
		γ	1,11	0,808	1,98	0,014	2,48	0,244	2,05	0,019	2,65	0,484
	3	β_1	0,98	0,032	1,00	0,005	1,00	0,006	1,00	0,005	1,00	0,049
		γ	0,73	1,617	1,95	0,014	2,66	0,447	2,07	0,025	2,68	1,061
	4	β_1	0,93	0,123	1,00	0,005	1,00	0,007	1,00	0,005	0,87	0,619
		γ	0,53	2,155	1,93	0,015	2,82	0,680	2,12	0,043	1,63	1,651

Tabela 19 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 0,5$ e $\beta_1 = 1, 4$ grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 20% de censura com covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	β_1	1,34	0,784	0,76	0,267	0,60	0,291	0,74	0,267	1,34	0,783
		γ	0,39	0,013	0,76	0,070	0,96	0,222	0,77	0,078	0,40	0,013
	2	β_1	1,50	1,094	0,68	0,272	0,50	0,344	0,70	0,270	1,50	1,100
		γ	0,36	0,021	0,86	0,133	1,17	0,462	0,82	0,110	0,36	0,021
	3	β_1	1,64	1,416	0,60	0,294	0,40	0,418	0,65	0,281	1,64	1,432
		γ	0,33	0,028	0,98	0,235	1,45	0,915	0,88	0,153	0,33	0,028
	4	β_1	1,74	1,692	0,54	0,323	0,33	0,492	0,61	0,294	1,74	1,710
		γ	0,32	0,032	1,10	0,371	1,76	1,620	0,94	0,202	0,32	0,032
100	1	β_1	1,33	0,463	0,75	0,172	0,59	0,233	0,74	0,175	1,33	0,462
		γ	0,39	0,013	0,74	0,062	0,95	0,205	0,76	0,069	0,39	0,013
	2	β_1	1,49	0,681	0,67	0,196	0,49	0,304	0,69	0,189	1,49	0,686
		γ	0,36	0,021	0,84	0,120	1,15	0,432	0,81	0,099	0,35	0,022
	3	β_1	1,63	0,930	0,60	0,232	0,40	0,391	0,65	0,208	1,64	0,937
		γ	0,33	0,029	0,96	0,216	1,43	0,863	0,87	0,138	0,33	0,029
	4	β_1	1,73	1,141	0,53	0,274	0,33	0,474	0,60	0,230	1,74	1,155
		γ	0,32	0,033	1,08	0,342	1,73	1,536	0,92	0,182	0,32	0,034
200	1	β_1	1,34	0,280	0,75	0,112	0,60	0,194	0,74	0,116	1,34	0,282
		γ	0,39	0,013	0,74	0,058	0,94	0,196	0,75	0,064	0,39	0,013
	2	β_1	1,49	0,451	0,67	0,147	0,49	0,277	0,69	0,138	1,50	0,455
		γ	0,35	0,022	0,84	0,114	1,14	0,417	0,80	0,093	0,35	0,022
	3	β_1	1,63	0,652	0,60	0,193	0,40	0,372	0,65	0,162	1,64	0,657
		γ	0,33	0,029	0,95	0,207	1,41	0,837	0,86	0,131	0,33	0,029
	4	β_1	1,73	0,826	0,53	0,243	0,33	0,463	0,61	0,190	1,74	0,833
		γ	0,32	0,034	1,07	0,329	1,72	1,495	0,91	0,173	0,32	0,034

Tabela 20 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 1$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 20% de censura com covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	β_1	1,39	0,376	0,97	0,091	0,80	0,100	0,94	0,088	1,26	0,297
		γ	0,70	0,103	1,10	0,029	1,36	0,152	1,15	0,042	0,85	0,085
	2	β_1	1,69	0,802	0,93	0,087	0,70	0,135	0,92	0,087	1,59	0,724
		γ	0,57	0,193	1,17	0,045	1,56	0,340	1,20	0,059	0,70	0,166
	3	β_1	2,07	1,631	0,88	0,087	0,60	0,194	0,89	0,088	2,04	1,625
		γ	0,47	0,283	1,25	0,076	1,82	0,700	1,24	0,080	0,54	0,254
	4	β_1	2,40	2,571	0,83	0,093	0,51	0,266	0,86	0,090	2,43	2,650
		γ	0,43	0,338	1,34	0,134	2,12	1,283	1,29	0,112	0,45	0,309
100	1	β_1	1,38	0,252	0,96	0,047	0,79	0,073	0,94	0,047	1,27	0,200
		γ	0,69	0,102	1,09	0,016	1,34	0,126	1,14	0,027	0,81	0,074
	2	β_1	1,67	0,616	0,93	0,047	0,70	0,115	0,91	0,048	1,62	0,579
		γ	0,56	0,196	1,15	0,030	1,54	0,299	1,18	0,040	0,65	0,160
	3	β_1	2,05	1,347	0,88	0,051	0,60	0,180	0,89	0,051	2,05	1,353
		γ	0,46	0,287	1,23	0,059	1,79	0,640	1,22	0,058	0,50	0,256
	4	β_1	2,38	2,224	0,83	0,062	0,51	0,256	0,86	0,056	2,41	2,285
		γ	0,42	0,341	1,32	0,112	2,09	1,202	1,27	0,084	0,44	0,317
200	1	β_1	1,38	0,196	0,96	0,023	0,80	0,056	0,94	0,024	1,31	0,163
		γ	0,68	0,105	1,08	0,010	1,33	0,111	1,13	0,019	0,77	0,075
	2	β_1	1,68	0,533	0,93	0,025	0,70	0,102	0,91	0,026	1,65	0,521
		γ	0,55	0,200	1,14	0,022	1,52	0,276	1,16	0,031	0,61	0,171
	3	β_1	2,06	1,226	0,88	0,031	0,60	0,171	0,89	0,030	2,06	1,235
		γ	0,46	0,290	1,22	0,050	1,77	0,608	1,21	0,047	0,48	0,269
	4	β_1	2,39	2,066	0,83	0,044	0,51	0,249	0,86	0,037	2,40	2,107
		γ	0,41	0,345	1,31	0,100	2,07	1,155	1,26	0,070	0,43	0,330

Tabela 21 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 2$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 20% de censura com covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	β_1	1,23	0,104	1,01	0,028	0,87	0,037	1,01	0,028	1,03	0,033
		γ	1,58	0,264	2,06	0,092	2,42	0,290	2,09	0,104	2,24	0,288
	2	β_1	1,47	0,319	1,01	0,029	0,79	0,059	1,01	0,028	1,06	0,064
		γ	1,20	0,714	2,03	0,084	2,62	0,517	2,10	0,116	2,55	0,865
	3	β_1	1,94	1,100	1,01	0,031	0,71	0,100	1,01	0,031	1,36	0,462
		γ	0,81	1,443	1,98	0,076	2,84	0,859	2,10	0,152	2,36	1,397
	4	β_1	2,59	2,920	1,01	0,033	0,63	0,148	1,02	0,034	2,27	2,544
		γ	0,59	1,989	1,94	0,050	3,04	1,221	2,14	0,205	1,79	4,645
100	1	β_1	1,22	0,076	1,01	0,015	0,87	0,028	1,00	0,014	1,02	0,016
		γ	1,54	0,251	2,01	0,040	2,37	0,186	2,04	0,046	2,26	0,175
	2	β_1	1,46	0,265	1,01	0,015	0,79	0,052	1,01	0,015	1,02	0,022
		γ	1,16	0,731	1,99	0,038	2,56	0,373	2,04	0,051	2,61	0,600
	3	β_1	1,91	0,945	1,00	0,016	0,70	0,095	1,00	0,015	1,26	0,234
		γ	0,80	1,463	1,95	0,035	2,78	0,675	2,05	0,061	2,41	0,731
	4	β_1	2,57	2,636	1,00	0,016	0,63	0,145	1,01	0,016	2,29	2,343
		γ	0,58	2,003	1,91	0,031	2,98	1,022	2,06	0,082	1,44	2,444
200	1	β_1	1,22	0,061	1,00	0,007	0,87	0,022	1,00	0,007	1,02	0,008
		γ	1,52	0,246	1,99	0,018	2,34	0,138	2,02	0,019	2,29	0,133
	2	β_1	1,46	0,236	1,01	0,007	0,79	0,048	1,01	0,007	0,99	0,008
		γ	1,14	0,748	1,96	0,018	2,53	0,302	2,02	0,021	2,67	0,530
	3	β_1	1,91	0,884	1,00	0,007	0,70	0,092	1,00	0,007	1,17	0,092
		γ	0,78	1,481	1,92	0,020	2,75	0,587	2,02	0,026	2,54	0,518
	4	β_1	2,56	2,529	1,00	0,008	0,63	0,142	1,01	0,008	2,35	2,304
		γ	0,58	2,013	1,89	0,022	2,94	0,917	2,03	0,034	1,21	1,880

Tabela 22 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 0,5$ e $\beta_1 = 1, 4$ grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 40% de censura com covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	β_1	1,12	0,692	0,77	0,364	0,62	0,344	0,67	0,341	1,10	0,667
		γ	0,47	0,005	0,70	0,047	0,88	0,155	0,82	0,109	0,49	0,007
	2	β_1	1,24	0,863	0,70	0,342	0,51	0,371	0,62	0,341	1,22	0,840
		γ	0,43	0,008	0,78	0,083	1,06	0,324	0,89	0,156	0,44	0,009
	3	β_1	1,36	1,095	0,63	0,341	0,42	0,426	0,57	0,350	1,35	1,082
		γ	0,39	0,013	0,87	0,142	1,29	0,644	0,96	0,215	0,40	0,014
	4	β_1	1,47	1,324	0,57	0,349	0,35	0,486	0,54	0,360	1,46	1,314
		γ	0,37	0,018	0,96	0,218	1,57	1,157	1,02	0,282	0,38	0,018
100	1	β_1	1,11	0,353	0,76	0,214	0,61	0,253	0,66	0,232	1,10	0,343
		γ	0,47	0,003	0,69	0,040	0,87	0,143	0,81	0,099	0,47	0,003
		β_1	1,22	0,460	0,69	0,223	0,51	0,310	0,61	0,249	1,22	0,454
		γ	0,42	0,007	0,77	0,075	1,05	0,305	0,88	0,145	0,43	0,007
	3	β_1	1,34	0,612	0,62	0,246	0,42	0,387	0,57	0,272	1,34	0,610
		γ	0,39	0,013	0,86	0,132	1,28	0,617	0,95	0,202	0,39	0,013
	4	β_1	1,45	0,776	0,57	0,273	0,34	0,462	0,53	0,294	1,45	0,775
		γ	0,37	0,018	0,95	0,206	1,55	1,115	1,01	0,266	0,37	0,018
200	1	β_1	1,11	0,185	0,76	0,135	0,61	0,202	0,66	0,173	1,10	0,181
		γ	0,46	0,002	0,68	0,036	0,86	0,135	0,80	0,093	0,46	0,002
	2	β_1	1,23	0,259	0,69	0,158	0,51	0,275	0,62	0,199	1,23	0,258
		γ	0,42	0,007	0,76	0,070	1,04	0,292	0,87	0,137	0,42	0,007
	3	β_1	1,35	0,368	0,62	0,193	0,42	0,363	0,57	0,228	1,35	0,368
		γ	0,39	0,013	0,85	0,125	1,27	0,596	0,94	0,193	0,39	0,013
	4	β_1	1,45	0,492	0,57	0,229	0,34	0,446	0,53	0,256	1,45	0,494
		γ	0,36	0,019	0,94	0,196	1,54	1,083	1,00	0,255	0,36	0,019

Tabela 23 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 1$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 40% de censura com covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	β_1	1,32	0,403	0,98	0,151	0,81	0,136	0,93	0,134	1,14	0,278
		γ	0,80	0,065	1,09	0,036	1,32	0,139	1,17	0,054	0,96	0,061
	2	β_1	1,57	0,756	0,96	0,142	0,71	0,158	0,90	0,131	1,34	0,539
		γ	0,67	0,127	1,12	0,042	1,50	0,288	1,21	0,069	0,86	0,112
	3	β_1	1,88	1,385	0,92	0,134	0,61	0,207	0,87	0,130	1,62	1,070
		γ	0,56	0,207	1,18	0,057	1,74	0,601	1,26	0,093	0,76	0,187
	4	β_1	2,18	2,149	0,87	0,128	0,52	0,273	0,83	0,130	1,94	1,812
		γ	0,48	0,274	1,24	0,081	2,03	1,115	1,30	0,121	0,66	0,258
100	1	β_1	1,31	0,244	0,97	0,076	0,80	0,089	0,92	0,071	1,14	0,168
		γ	0,78	0,060	1,07	0,019	1,30	0,108	1,15	0,035	0,93	0,045
	2	β_1	1,55	0,501	0,95	0,072	0,71	0,124	0,89	0,072	1,35	0,363
		γ	0,65	0,128	1,11	0,024	1,47	0,245	1,19	0,049	0,82	0,097
	3	β_1	1,85	1,018	0,91	0,071	0,60	0,185	0,86	0,075	1,64	0,802
		γ	0,55	0,209	1,16	0,038	1,72	0,543	1,24	0,071	0,71	0,166
	4	β_1	2,15	1,699	0,86	0,074	0,51	0,258	0,83	0,081	1,99	1,470
		γ	0,47	0,279	1,22	0,061	2,00	1,032	1,29	0,095	0,59	0,238
200	1	β_1	1,30	0,164	0,97	0,038	0,80	0,065	0,91	0,039	1,15	0,104
		γ	0,77	0,059	1,06	0,009	1,28	0,089	1,14	0,024	0,90	0,035
	2	β_1	1,54	0,397	0,94	0,038	0,70	0,107	0,88	0,043	1,38	0,291
		γ	0,64	0,130	1,09	0,014	1,46	0,220	1,18	0,037	0,77	0,090
	3	β_1	1,85	0,860	0,90	0,040	0,60	0,173	0,85	0,049	1,71	0,725
		γ	0,54	0,214	1,15	0,028	1,70	0,504	1,23	0,057	0,64	0,168
	4	β_1	2,14	1,484	0,86	0,046	0,51	0,249	0,82	0,056	2,06	1,350
		γ	0,47	0,283	1,21	0,050	1,98	0,979	1,27	0,080	0,53	0,246

Tabela 24 - Estimativas (E) dos parâmetros (P) e erro quadrático médio (EQM) considerando amostras simuladas de tamanho (n) 50, 100 e 200 de uma distribuição Weibull com $\gamma = 2$ e $\beta_1 = 1$, 4 grades (G, 10, 6, 4 e 3 intervalos) e os métodos de estimação M1, M2, M3, M4 e M5 para 40% de censura com covariável

n	G	P	M1		M2		M3		M4		M5	
			E	EQM	E	EQM	E	EQM	E	EQM	E	EQM
50	1	β_1	1,28	0,816	1,06	0,503	0,90	0,384	1,05	0,437	1,06	0,569
		γ	1,74	0,193	2,08	0,141	2,41	0,336	2,11	0,141	2,12	0,202
	2	β_1	1,51	1,161	1,07	0,461	0,82	0,325	1,04	0,400	1,04	0,438
		γ	1,47	0,395	2,06	0,138	2,63	0,596	2,12	0,145	2,23	0,379
	3	β_1	1,88	2,342	1,07	0,482	0,72	0,329	1,03	0,409	1,01	0,450
		γ	1,17	0,779	2,05	0,136	2,94	1,130	2,15	0,161	2,47	0,770
	4	β_1	2,40	4,604	1,09	0,468	0,63	0,330	1,03	0,383	1,02	0,644
		γ	0,90	1,265	2,01	0,122	3,23	1,811	2,15	0,165	2,78	1,954
100	1	β_1	1,22	0,122	1,02	0,043	0,86	0,050	1,01	0,044	0,99	0,044
		γ	1,70	0,158	2,04	0,071	2,37	0,223	2,05	0,074	2,13	0,120
		β_1	1,44	0,308	1,02	0,045	0,78	0,076	1,02	0,045	0,96	0,057
		γ	1,43	0,382	2,02	0,068	2,59	0,444	2,04	0,075	2,28	0,257
	3	β_1	1,80	0,820	1,03	0,045	0,68	0,122	1,01	0,045	0,91	0,080
		γ	1,13	0,780	2,00	0,065	2,88	0,907	2,05	0,081	2,55	0,563
	4	β_1	2,30	1,986	1,04	0,048	0,60	0,177	1,02	0,048	0,85	0,210
		γ	0,87	1,300	1,97	0,062	3,18	1,556	2,04	0,088	2,92	1,454
200	1	β_1	1,22	0,081	1,01	0,020	0,86	0,035	1,01	0,020	0,97	0,022
		γ	1,66	0,143	2,00	0,030	2,33	0,149	2,01	0,031	2,14	0,072
	2	β_1	1,43	0,237	1,02	0,020	0,77	0,064	1,01	0,020	0,93	0,028
		γ	1,40	0,382	1,99	0,029	2,55	0,347	2,01	0,032	2,32	0,196
	3	β_1	1,79	0,708	1,02	0,021	0,68	0,114	1,01	0,021	0,89	0,040
		γ	1,10	0,823	1,97	0,029	2,84	0,767	2,02	0,035	2,55	0,415
	4	β_1	2,28	1,778	1,04	0,022	0,60	0,171	1,01	0,022	0,77	0,105
		γ	0,85	1,332	1,94	0,029	3,14	1,368	2,01	0,037	3,04	1,333

ANEXO C - As Tabelas 25 e 26 apresentam um resumo das estimativas do parâmetro β_1 do modelo Weibull para os resultados obtidos nas simulações para 0%, 20% e 40% de censura, sem e com covariável, respectivamente.

Tabela 25 - Resumo das estimativas do parâmetro λ do modelo Weibull para dados com censura intervalar - sem covariável

γ	Métodos	Censura		
		0%	20%	40%
0,5	$M1$ e $M5$	viés negativo	viés negativo	viés negativo
	$M2$, $M3$ e $M4$	viés positivo	viés positivo	viés negativo
1,0	$M1$	viés negativo	viés negativo	viés positivo
	$M5$	viés negativo	viés negativo	viés negativo
	$M2$, $M3$	viés positivo	viés positivo	viés pequeno
	$M4$	viés pequeno	viés positivo	viés pequeno
2,0	$M1$	viés negativo	viés positivo	viés positivo
	$M5$	viés negativo	viés pequeno	viés pequeno
	$M2$ e $M4$	viés pequeno	viés pequeno	viés pequeno
	$M3$	viés positivo	viés positivo	viés positivo

Tabela 26 - Resumo das estimativas do parâmetro β_1 do modelo Weibull para dados com censura intervalar - com covariável

γ	Métodos	Censura		
		0%	20%	40%
0,5	$M1$ e $M5$	viés negativo	viés positivo	viés positivo
	$M2$, $M3$ e $M4$	viés pequeno	viés negativo	viés negativo
1,0	$M1$ e $M5$	viés negativo	viés positivo	viés positivo
	$M2$, $M3$ e $M4$	viés pequeno	viés negativo	viés negativo
2,0	$M1$, $M5$	viés negativo	viés positivo	viés positivo
	$M2$ e $M4$	viés pequeno	viés pequeno	viés pequeno
	$M3$	viés pequeno	viés negativo	viés negativo

ANEXO D - As Tabelas 27 e 28 apresentam um resumo das estimativas do parâmetro γ do modelo Weibull para os resultados obtidos nas simulações para 0%, 20% e 40% de censura, sem e com covariável, respectivamente.

Tabela 27 - Resumo das estimativas do parâmetro γ do modelo Weibull para dados com censura intervalar - sem covariável

γ	Métodos	Censura		
		0%	20%	40%
0,5	$M1$ e $M5$	viés negativo	viés negativo	viés negativo
	$M2$, $M3$ e $M4$	viés positivo	viés positivo	viés positivo
1,0	$M1$	viés negativo	viés negativo	viés negativo
	$M5$	viés negativo	viés negativo	viés positivo
	$M2$, $M3$ e $M4$	viés positivo	viés positivo	viés positivo
2,0	$M1$	viés negativo	viés negativo	viés negativo
	$M5$	viés positivo	viés positivo	viés positivo
	$M2$	viés positivo	viés pequeno	viés pequeno
	$M3$	viés positivo	viés positivo	viés positivo
	$M4$	viés positivo	viés pequeno	viés positivo

Tabela 28 - Resumo das estimativas do parâmetro γ do modelo Weibull para dados com censura intervalar - com covariável

γ	Métodos	Censura		
		0%	20%	40%
0,5	$M1$ e $M5$	viés negativo	viés negativo	viés negativo
	$M2$, $M3$ e $M4$	viés positivo	viés positivo	viés positivo
1,0	$M1$ e $M5$	viés negativo	viés negativo	viés negativo
	$M2$, $M3$ e $M4$	viés positivo	viés positivo	viés positivo
2,0	$M1$	viés negativo	viés negativo	viés negativo
	$M5$	viés positivo	viés positivo	viés positivo
	$M2$	viés positivo	viés negativo	viés positivo
	$M3$ e $M4$	viés positivo	viés positivo	viés positivo

ANEXO E- Conjunto dos Dados referentes ao tempo de vida das mangueiras

Tabela 29 - Dados referentes ao tempo de vida das mangueiras

(continua)

Ano	Censura	Copas	Cavalos	Blocos
85	1	1	1	3
85	1	1	1	4
88	1	1	1	5
90	1	1	1	1
92	0	1	1	2
85	1	1	2	3
85	1	1	2	4
85	1	1	2	5
88	1	1	2	1
88	1	1	2	2
88	1	1	3	5
89	1	1	3	1
90	1	1	3	4
92	0	1	3	3
81	1	1	4	4
88	1	1	4	1
88	1	1	4	5
92	0	1	4	2
92	0	1	4	3
73	1	1	5	1
85	1	1	5	3
88	1	1	5	5
90	1	1	5	4
92	0	1	5	2
81	1	1	6	4
87	1	1	6	5
89	1	1	6	1
90	1	1	6	2
90	1	1	6	3
73	1	1	7	5

Tabela 29 - Dados referentes ao tempo de vida das mangueiras

(continuação)				
Ano	Censura	Copas	Cavalos	Blocos
87	1	1	7	4
90	1	1	7	3
92	0	1	7	2
92	0	1	7	1
89	1	2	1	5
90	1	2	1	4
92	0	2	1	3
92	0	2	1	2
92	0	2	1	1
87	1	2	2	4
88	1	2	2	1
90	1	2	2	5
92	0	2	2	2
92	0	2	2	3
92	0	2	3	1
92	0	2	3	2
92	0	2	3	3
92	0	2	3	4
92	0	2	3	5
81	1	2	4	3
88	1	2	4	5
89	1	2	4	2
92	1	2	4	4
73	1	2	5	1
74	1	2	5	2
74	1	2	5	3
88	1	2	5	4
92	0	2	5	5
83	1	2	6	5
90	1	2	6	1
90	1	2	6	4

Tabela 29 - Dados referentes ao tempo de vida das mangueiras

(continuação)				
Ano	Censura	Copas	Cavalos	Blocos
92	0	2	6	2
92	0	2	6	3
88	1	2	7	2
88	1	2	7	5
90	1	2	7	3
90	1	2	7	4
92	0	2	7	1
73	1	3	1	5
89	1	3	1	4
92	0	3	1	3
92	0	3	1	2
92	0	3	1	1
74	1	3	2	4
74	1	3	2	5
88	1	3	2	2
92	0	3	2	1
92	0	3	2	3
73	1	3	3	3
73	1	3	3	5
88	1	3	3	4
92	0	3	3	2
92	0	3	3	1
74	1	3	4	5
75	1	3	4	3
87	1	3	4	4
90	1	3	4	1
92	0	3	4	2
74	1	3	5	2
74	1	3	5	4
87	1	3	5	5
90	1	3	5	3

Tabela 29 - Dados referentes ao tempo de vida das mangueiras

(continuação)				
Ano	Censura	Copas	Cavalos	Blocos
92	0	3	5	1
73	1	3	6	1
86	1	3	6	3
90	1	3	6	4
90	1	3	6	5
92	0	3	6	2
73	1	3	7	2
81	1	3	7	5
90	1	3	7	1
90	1	3	7	4
92	1	3	7	3
88	1	4	1	5
90	1	4	1	3
92	0	4	1	4
92	0	4	1	2
92	0	4	1	1
73	1	4	2	2
86	1	4	2	1
86	1	4	2	4
92	0	4	2	3
92	0	4	2	5
86	1	4	3	1
88	1	4	3	3
88	1	4	3	4
90	1	4	3	2
92	1	4	3	5
81	1	4	4	2
85	1	4	4	5
87	1	4	4	4
92	0	4	4	1
92	0	4	4	3

Tabela 29 - Dados referentes ao tempo de vida das mangueiras

(continuação)				
Ano	Censura	Copas	Cavalos	Blocos
73	1	4	5	2
81	1	4	5	4
85	1	4	5	5
92	0	4	5	1
92	0	4	5	3
87	1	4	6	2
90	1	4	6	4
90	1	4	6	5
92	1	4	6	3
92	1	4	6	1
87	1	4	7	4
87	1	4	7	5
89	1	4	7	3
92	0	4	7	1
92	0	4	7	2
73	1	5	1	5
85	1	5	1	1
89	1	5	1	2
90	1	5	1	3
92	0	5	1	4
85	1	5	2	2
85	1	5	2	4
89	1	5	2	3
92	0	5	2	1
92	0	5	2	5
86	1	5	3	1
86	1	5	3	2
88	1	5	3	4
92	0	5	3	3
92	0	5	3	5
81	1	5	4	5

Tabela 29 - Dados referentes ao tempo de vida das mangueiras

(continuação)				
Ano	Censura	Copas	Cavalos	Blocos
85	1	5	4	2
86	1	5	4	3
87	1	5	4	4
92	0	5	4	1
73	1	5	5	2
86	1	5	5	1
89	1	5	5	5
92	0	5	5	3
92	0	5	5	4
86	1	5	6	1
88	1	5	6	4
92	1	5	6	5
92	0	5	6	2
92	0	5	6	3
74	1	5	7	2
88	1	5	7	3
92	0	5	7	1
92	0	5	7	4
92	0	5	7	5
85	1	6	1	2
86	1	6	1	1
87	1	6	1	4
88	1	6	1	5
89	1	6	1	3
85	1	6	2	1
85	1	6	2	3
85	1	6	2	5
86	1	6	2	2
90	1	6	2	4
85	1	6	3	4
87	1	6	3	5

Tabela 29 - Dados referentes ao tempo de vida das mangueiras

(conclusão)				
Ano	Censura	Copas	Cavalos	Blocos
88	1	6	3	3
88	1	6	3	2
88	1	6	3	1
85	1	6	4	1
86	1	6	4	3
87	1	6	4	2
88	1	6	4	4
88	1	6	4	5
83	1	6	5	4
85	1	6	5	1
85	1	6	5	2
85	1	6	5	3
85	1	6	5	5
85	1	6	6	4
87	1	6	6	1
88	1	6	6	3
90	1	6	6	5
81	1	6	7	4
86	1	6	7	3
87	1	6	7	5
88	1	6	7	1
90	1	6	7	2

ANEXO F- Conjunto dos Dados referentes ao tempo de murcha do cultivar de linho

Tabela 30 - Dados referentes ao tempo de murcha do cultivar de linho

(continua)

Tempo	Substratos	Número de Plantas Murchas	Censura
21	M	5	66
24	M	8	58
28	M	13	45
31	M	10	35
35	M	11	24
38	M	16	8
42	M	4	4
45	M	2	2
49	M	1	1
52	M	0	1
21	MI	2	70
24	MI	6	64
28	MI	7	57
31	MI	12	45
35	MI	10	35
38	MI	12	23
42	MI	8	15
45	MI	8	7
49	MI	3	4
52	MI	0	4
21	MK	1	70
24	MK	14	56
28	MK	8	48
31	MK	11	37
35	MK	10	27
38	MK	14	13
42	MK	4	9
45	MK	6	3
52	MK	0	3

Tabela 30 - Dados referentes ao tempo de murcha do cultivar de linho

Tempo	Substratos	Número de Plantas Murchas	(conclusão)
			Censura
21	MM	1	71
24	MM	2	69
28	MM	3	66
31	MM	12	54
35	MM	13	41
38	MM	16	25
42	MM	8	17
45	MM	9	8
49	MM	5	3
52	MM	0	3

ANEXO G - Programa escrito em R para obter as estimativas dos parâmetros do modelo discreto Weibull considerando uma covariável

```

fn <- function(p){
d=p[3]*p[1] + p[3]*p[2]*X4
e=exp(d)
f=(X2**p[3]) - (X1**p[3])
g=-(e*f)
G=(1-X3)*g
h=1-exp(g)
H=X3*log(h)
-sum(H+G) }

dv<- deriv(~-(X3*log(1-exp(-(exp(p3*p1 + p3*p2*X4)))*(X2**p3 -
X1**p3)))+(1-X3)*(-(exp(p3*p1 + p3*p2*X4))*(X2**p3 -
X1**p3))),c("p1","p2","p3"), function(p1,p2,p3) NULL)

grd=function(p){ p1=p[1] p2=p[2] p3=p[3]
colSums(attr((dv(p1,p2,p3)),"gradient"))
}

out<-optim(c(beta0,beta1,gama),fn,grd,hessian=TRUE)

```

ANEXO H - Programa escrito em R para obtenção da estatística do teste escore

```
fescore <- function(p){
  (-(1-delta)*exp((p[10]*p[9]+p[1])+log(b^p[10]-g^p[10])+p[11]*co)+
    delta*log(1-exp(-exp((p[10]*p[9]+p[1])+log(b^p[10]-g^p[10])+p[11]*co))))*Int1+
  (-(1-delta)*exp((p[10]*p[9]+p[2])+log(b^p[10]-g^p[10])+p[11]*co)+
    delta*log(1-exp(-exp((p[10]*p[9]+p[2])+log(b^p[10]-g^p[10])+p[11]*co))))*Int2+
  (-(1-delta)*exp((p[10]*p[9]+p[3])+log(b^p[10]-g^p[10])+p[11]*co)+
    delta*log(1-exp(-exp((p[10]*p[9]+p[3])+log(b^p[10]-g^p[10])+p[11]*co))))*Int3+
  (-(1-delta)*exp((p[10]*p[9]+p[4])+log(b^p[10]-g^p[10])+p[11]*co)+
    delta*log(1-exp(-exp((p[10]*p[9]+p[4])+log(b^p[10]-g^p[10])+p[11]*co))))*Int4+
  (-(1-delta)*exp((p[10]*p[9]+p[5])+log(b^p[10]-g^p[10])+p[11]*co)+
    delta*log(1-exp(-exp((p[10]*p[9]+p[5])+log(b^p[10]-g^p[10])+p[11]*co))))*Int5+
  (-(1-delta)*exp((p[10]*p[9]+p[6])+log(b^p[10]-g^p[10])+p[11]*co)+
    delta*log(1-exp(-exp((p[10]*p[9]+p[6])+log(b^p[10]-g^p[10])+p[11]*co))))*Int6+
  (-(1-delta)*exp((p[10]*p[9]+p[7])+log(b^p[10]-g^p[10])+p[11]*co)+
    delta*log(1-exp(-exp((p[10]*p[9]+p[7])+log(b^p[10]-g^p[10])+p[11]*co))))*Int7+
  (-(1-delta)*exp((p[10]*p[9]+p[8])+log(b^p[10]-g^p[10])+p[11]*co)+
    delta*log(1-exp(-exp((p[10]*p[9]+p[8])+log(b^p[10]-g^p[10])+p[11]*co))))*Int8+
  (-(1-delta)*exp((p[10]*p[9])+log(b^p[10]-g^p[10])+p[11]*co)+
    delta*log(1-exp(-exp((p[10]*p[9])+log(b^p[10]-g^p[10])+p[11]*co))))*Int9+
  (-(1-delta)*exp((p[10]*p[9])+log(b^p[10]-g^p[10])+p[11]*co)+
    delta*log(1-exp(-exp((p[10]*p[9])+log(b^p[10]-g^p[10])+p[11]*co))))*Int10
}

descore<-deriv(~(-(1-delta)*exp((p10*p9+p1)+log(b^p10-g^p10)+p11*co)+
  delta*log(1-exp(-exp((p10*p9+p1)+log(b^p10-g^p10)+p11*co))))*Int1+
  (-(1-delta)*exp((p10*p9+p2)+log(b^p10-g^p10)+p11*co)+
    delta*log(1-exp(-exp((p10*p9+p2)+log(b^p10-g^p10)+p11*co))))*Int2+
  (-(1-delta)*exp((p10*p9+p3)+log(b^p10-g^p10)+p11*co)+
    delta*log(1-exp(-exp((p10*p9+p3)+log(b^p10-g^p10)+p11*co))))*Int3+
  (-(1-delta)*exp((p10*p9+p4)+log(b^p10-g^p10)+p11*co)+
    delta*log(1-exp(-exp((p10*p9+p4)+log(b^p10-g^p10)+p11*co))))*Int4+
  (-(1-delta)*exp((p10*p9+p5)+log(b^p10-g^p10)+p11*co)+
    delta*log(1-exp(-exp((p10*p9+p5)+log(b^p10-g^p10)+p11*co))))*Int5+
  (-(1-delta)*exp((p10*p9+p6)+log(b^p10-g^p10)+p11*co)+
    delta*log(1-exp(-exp((p10*p9+p6)+log(b^p10-g^p10)+p11*co))))*Int6+
  (-(1-delta)*exp((p10*p9+p7)+log(b^p10-g^p10)+p11*co)+
    delta*log(1-exp(-exp((p10*p9+p7)+log(b^p10-g^p10)+p11*co))))*Int7+
  (-(1-delta)*exp((p10*p9+p8)+log(b^p10-g^p10)+p11*co)+
    delta*log(1-exp(-exp((p10*p9+p8)+log(b^p10-g^p10)+p11*co))))*Int8+
  (-(1-delta)*exp((p10*p9)+log(b^p10-g^p10)+p11*co)+
    delta*log(1-exp(-exp((p10*p9)+log(b^p10-g^p10)+p11*co))))*Int9+
  (-(1-delta)*exp((p10*p9)+log(b^p10-g^p10)+p11*co)+
    delta*log(1-exp(-exp((p10*p9)+log(b^p10-g^p10)+p11*co))))*Int10),c("p1","p2","p3","p4","p5","p6",
  "p7","p8","p9","p10","p11"),function(p1,p2,p3,p4,p5,p6,p7,p8,p9,p10,p11)NULL,hessian=TRUE)
```

```

egrd1=function(p){
p1=p[1]
p2=p[2]
p3=p[3]
p4=p[4]
p5=p[5]
p6=p[6]
p7=p[7]
p8=p[8]
p9=p[9]
p10=p[10]
p11=p[11]
colSums(attr((descore(p1,p2,p3,p4,p5,p6,p7,p8,p9,p10,p11)),"gradient"))
}
attr((descore(psi1,psi2,psi3,psi4,psi5,psi6,psi7,psi8,betazero,gama2,betaum)),"gradient")

primderiv=egrd1(c(psi1,psi2,psi3,psi4,psi5,psi6,psi7,psi8,betazero,gama2,betaum))

egrd2=function(p){ p1=p[1] p2=p[2] p3=p[3] p4=p[4] p5=p[5] p6=p[6]
p7=p[7] p8=p[8] p9=p[9] p10=p[10] p11=p[11]
colSums(attr((descore(p1,p2,p3,p4,p5,p6,p7,p8,p9,p10,p11)),"hessian"))
}
segderiv=egrd2(c(psi1,psi2,psi3,psi4,psi5,psi6,psi7,psi8,betazero,gama2,betaum))

minform=-segderiv ipsipsi=minform[1:8,1:8]
ipsiphi=minform[1:8,9:11] iphipsi=t(ipsiphi)
iphiphi=minform[9:11,9:11] inviphiphi=solve(iphiphi)
Inv=ipsipsi-ipsiphi%*%inviphiphi%*%iphipsi
upsi=matrix(primderiv[1:8])
tupsi=t(upsi)
Escore=tupsi%*%Inv%*%upsi

```

ANEXO I - Programa escrito em R para obtenção da estatística do teste da razão da máxima verossimilhança

```

M<-1000
k<-10
RV<-2*(likcox-likweibull)

pvalor<-rep(0,M)
for (i in 1:M){
  k<-10

  pvalor[i]<-1-pchisq(RV[i],(k-2))

  if(pvalor[i]<0.05){ cat("\n", "Hipoteses Testadas:", "\n")
    cat("H0: O modelo Weibull é adequado aos dados", "\n") cat("HA: O
    modelo Weibull não é adequado aos dados", "\n") cat("\n",
    "Rejeitar H0, p-valor:", pvalor[i], "<0.05", "\n")} else{
    cat("\n", "Hipoteses Testadas:", "\n") cat("H0: O modelo Weibull é
    adequado aos dados", "\n") cat("HA: O modelo Weibull não é
    adequado aos dados", "\n") cat("\n", "Não Rejeitar H0, p-valor:",
    pvalor[i], ">=0.05", "\n")}
  }

  p<-ifelse(pvalor<0.05,1,0) percentual<-mean(p)*100

  cat("\n", "Teste da Razão de Verossimilhança", "\n", "Hipoteses
  Testadas:", "\n", "H0:O modelo Weibull é adequado aos dados",
  "\n", "HA:O modelo Weibull não é adequado aos dados", "\n", "\n",
  "Percentual de vezes que H0 foi rejeitada (tamanho do teste):",
  percentual, "%", "\n")

  RV<-RV[order(RV)] RV

  quantis<- qchisq(ppoints(length(RV)),df=(k-2))
  qqplot(quantis,RV)
  abline(0,1)

  legend(18,9, lty=c(1), col=c("white"), c("pc=20"), lwd=2, cex=0.9,
  bty="n")

```

Programas em R Utilizados nas Simulações de 2.6 para 20% de censura e com covariável

ANEXO J - Método M1: considerando que o evento ocorreu no início do intervalo para amostra de tamanho 50, $\gamma = 0.5$

```

n1<-26
n2<-24
n<-50
a<-0.5
b1<-1
b2<-2.72
d<-1000
f<-4000
M1<-matrix(0,n1,d)
for(k in 1:d){
w1<-rweibull(n1,a,b1)
w1<-w1[order(w1)]
for(i in 1:n1){
M1[i,k]=M1[i,k]+w1[i] } }
A1<-matrix(0,n1,(d*4))
for(colunaM in 1:d) {
for(i in 1:4) {
if(i== 1) { fracao<-(4.1/10)
} else
{
if(i == 2) { fracao<-(4.1/6)
} else
{
if(i == 3) { fracao<-(4.1/4)
}else
{
if(i == 4) { fracao<-(4.1/3) }
}
}
}
}
valor<-fracao linha=1 colunaA=i+(colunaM-1)*4
while(linha<=n1)
{
if(M1[linha,colunaM]>4.1)
{
A1[linha,colunaA]=4.1
} else
{
if(M1[linha,colunaM]<valor)

```



```

if(M2[linha,colunaM]<valor) { A2[linha,colunaA]=(valor-fracao) }
else
{
if ((M2[linha,colunaM]>valor) && (M2[linha,colunaM]<(valor+fracao)
)) { A2[linha,colunaA]=valor valor<-valor+fracao } else
{
if ( (M2[linha,colunaM]>(valor+fracao) ) &&
(M2[linha,colunaM]<4.1)) {
valor<-valor+fracao
linha<-linha-1
}
}
}

} linha<-linha+1
}
} }

A<-rbind(A1,A2)

B<-A[,]
B<-ifelse(A[,]==0,0.01,A[,])

require(survival)
ce<-B[,]
ce<-ifelse(B[,]<4.1,1,0)

trat<-c(rep(0,26),rep(1,24))
trat<-matrix(trat)

for (k in 1:f){
fit<-survreg(Surv(B[,k],ce[,k])~trat,dist='weibull')
}

```

ANEXO K - Método M2: considerando que o evento ocorreu no ponto médio do intervalo para amostra de tamanho 50, $\gamma = 0.5$

```

n1<-26 n2<-24
n<-50
<-0.5
b1<-1
b2<-2.72
d<-1000
f<-4000
M1<-matrix(0,n1,d) )

```

```

for(k in 1:d){
w1<-rweibull(n1,a,b1)
w1<-w1[order(w1)]
for(i in 1:n1){
M1[i,k]=M1[i,k]+w1[i] } }

A1<-matrix(0,n1,(d*4))
for(colunaM in 1:d) {
  for(i in 1:4)
  {
    if(i == 1)
    {
      fracao<-(4.1/10)
    } else
    {
      if(i == 2)
      {
        fracao<-(4.1/6)
      } else
      {
        if(i == 3)
        {
          fracao<-(4.1/4)
        }else
        {
          if(i == 4)
          {
            fracao<-(4.1/3)
          }
        }
      }
    }
    valorf<-fracao
    valori<-valorf-fracao
    linha=1
    colunaA=i+(colunaM-1)*4
    while(linha<=n1)
    {

      if(M1[linha,colunaM]>4.1)
      {
        A1[linha,colunaA]=4.1
      } else

    {

      if(M1[linha,colunaM]<valorf)

```



```

        {
            A1[linha,colunaA]=(valorf+valori)/2
        } else
            {
if ((M1[linha,colunaM]>=valorf) &&
(M1[linha,colunaM]<(valorf+fracao) )) {
                valorf<-valorf+fracao
                valori<-valorf-fracao
                A1[linha,colunaA]=(valorf+valori)/2
            } else
                {
if ( (M1[linha,colunaM]>(valorf+fracao) ) && (M1[linha,colunaM]<4.1)) {
                    valorf<-valorf+fracao
                    valori<-valorf-fracao
                    linha<-linha-1
                }
            }
        }

        linha<-linha+1
    }
} A1

M2<-matrix(0,n2,d)
for(k in 1:d){
w2<-rweibull(n2,a,b2)
w2<-w2[order(w2)]
for(i in 1:n2){
M2[i,k]=M2[i,k]+w2[i] } }

A2<-matrix(0,n2,(d*4)) for(colunaM in 1:d) {
    for(i in 1:4)
    {
        if(i == 1)
        {
            fracao<-(4.1/10)
        } else
        {
            if(i == 2)
            {
                fracao<-(4.1/6)
            } else
            {
                if(i == 3)

```

```

        {
            fracao<-(4.1/4)
        }else
        {
            if(i == 4)
            {
                fracao<-(4.1/3)
            }
        }
    }
}
valorf<-fracao
valori<-valorf-fracao
linha=1
colunaA=i+(colunaM-1)*4
while(linha<=n2)
{

    if(M2[linha,colunaM]>4.1)
    {
        A2[linha,colunaA]=4.1
    } else

    {

        if (M2[linha,colunaM]<valorf)
        {
            A2[linha,colunaA]=(valorf+valori)/2
        } else

        {
            if ((M2[linha,colunaM]>=valorf) &&
            (M2[linha,colunaM]<(valorf+fracao) )) {
                valorf<-valorf+fracao
                valori<-valorf-fracao
                A2[linha,colunaA]=(valorf+valori)/2
            } else

            {
                if ( (M2[linha,colunaM]>(valorf+fracao) ) && (M2[linha,colunaM]<4.1)) {
                    valorf<-valorf+fracao
                    valori<-valorf-fracao
                    linha<-linha-1
                }

            }

        }

    }

    linha<-linha+1
}

```

```

    }
  }
}

A<-rbind(A1,A2)

require(survival)
ce<-A[,]
ce<-ifelse(A[,]<4.1,1,0)

trat<-c(rep(0,26),rep(1,24))
trat<-matrix(trat)

for (k in 1:f){
  fit<-survreg(Surv(A[,k],ce[,k])~trat,dist='weibull')
}

```

ANEXO L - Método M3: considerando que o evento ocorreu no final do intervalo para amostra de tamanho 50, $\gamma = 0.5$

```

n1<-26
n2<-24
n<-50
a<-0.5
b1<-1
b2<-2.72
d<-1000
f<-4000
M1<-matrix(0,n1,d)
for(k in 1:d){
  w1<-rweibull(n1,a,b1)
  w1<-w1[order(w1)]
  for(i in 1:n1){
    M1[i,k]=M1[i,k]+w1[i] } }

M2<-matrix(0,n2,d)
for(k in 1:d){
  w2<-rweibull(n2,a,b2)
  w2<-w2[order(w2)]
  for(i in 1:n2){
    M2[i,k]=M2[i,k]+w2[i] } }

A1<-matrix(0,n1,(d*4)) for(colunaM in 1:d) {
  for(i in 1:4)
    {

```

```

if(i == 1)
{
    fracao<-(4.10/10)
} else
{
    if(i == 2)
    {
        fracao<-(4.10/6)
    } else
    {
        if(i == 3)
        {
            fracao<-(4.10/4)
        }else
        {
            if(i == 4)
            {
                fracao<-(4.10/3)
            }
        }
    }
}
valor<-fracao
linha=1
colunaA=i+(colunaM-1)*4
while(linha<=n1)
{
    if(M1[linha,colunaM]<valor)
    {
        A1[linha,colunaA]=valor
    } else
    {
        if ((M1[linha,colunaM]>valor) && (M1[linha,colunaM]<(valor+fracao)
        )) {
            A1[linha,colunaA]=valor+fracao
            valor<-valor+fracao
        } else
        {
            if ( (M1[linha,colunaM]>(valor+fracao) ) &&
            (M1[linha,colunaM]<4.10)) {
                valor<-valor+fracao
                linha<-linha-1
            }
        }
    }
    if(M1[linha,colunaM]>4.10)

```

```

        {
            A1[linha,colunaA]=4.11
        }
        linha<-linha+1
    }
}

ce1<-A1[,]
ce1<-ifelse(ce1[,]<4.11,1,0)
A1<-ifelse(A1[,]>4.10,4.10,A1[,])

A2<-matrix(0,n2,(d*4)) for(colunaM in 1:d) {
    for(i in 1:4)
    {
        if(i == 1)
        {
            fracao<-(4.10/10)
        } else
        {
            if(i == 2)
            {
                fracao<-(4.10/6)
            } else
            {
                if(i == 3)
                {
                    fracao<-(4.10/4)
                }else
                {
                    if(i == 4)
                    {
                        fracao<-(4.10/3)
                    }
                }
            }
        }
    }
    valor<-fracao
    linha=1
    colunaA=i+(colunaM-1)*4
    while(linha<=n2)
    {
        if(M2[linha,colunaM]<valor)
        {
            A2[linha,colunaA]=valor
        } else
    }
}

```

```

      {
if ((M2[linha,colunaM]>valor) && (M2[linha,colunaM]<(valor+fracao)
)) {

      A2[linha,colunaA]=valor+fracao
      valor<-valor+fracao
    } else
      {
if ( (M2[linha,colunaM]>(valor+fracao) ) &&
(M2[linha,colunaM]<4.10)) {
      valor<-valor+fracao
      linha<-linha-1
    }
      }
      }
      if(M2[linha,colunaM]>4.10)
      {
      A2[linha,colunaA]=4.11
      }
      linha<-linha+1
    }
  }
} A2

ce2<-A2[,]
ce2<-ifelse(ce2[,]<4.11,1,0)
A2<-ifelse(A2[,]>4.10,4.10,A2[,])

A<-rbind(A1,A2)
ce<-rbind(ce1,ce2)

trat<-c(rep(0,26),rep(1,24))
trat<-matrix(trat)

require(survival)

for (k in 1:f){
fit<-survreg(Surv(A[,k],ce[,k])~trat,dist='weibull')
}

```

ANEXO M - Método M4: considerando que o evento ocorreu no intervalo para amostra de tamanho 50, $\gamma = 0.5$

```

n1<-26
n2<-24
n<-50

```

```

a<-0.5
b1<-1
b2<-2.72
d<-1000
f<-4000
M1<-matrix(0,n1,d)
for(k in 1:d){
w1<-rweibull(n1,a,b1)
w1<-w1[order(w1)]
for(i in 1:n1){
M1[i,k]=M1[i,k]+w1[i] } }

M2<-matrix(0,n2,d)
for(k in 1:d){
w2<-rweibull(n2,a,b2)
w2<-w2[order(w2)]
for(i in 1:n2){
M2[i,k]=M2[i,k]+w2[i] } }

A1<-matrix(0,n1,(d*4)) for(colunaM in 1:d) {
  for(i in 1:4)
  {
    if(i == 1)
    {
      fracao<-(4.10/10)
    } else
    {
      if(i == 2)
      {
        fracao<-(4.10/6)
      } else
      {
        if(i == 3)
        {
          fracao<-(4.10/4)
        }else
        {
          if(i == 4)
          {
            fracao<-(4.10/3)
          }
        }
      }
    }
  }
  valor<-fracao
  linha=1

```

```

colunaA=i+(colunaM-1)*4
while(linha<=n1)
{
  if(M1[linha,colunaM]<valor)
  {
    A1[linha,colunaA]=(valor-fracao)
  } else
  {
    if ((M1[linha,colunaM]>valor) && (M1[linha,colunaM]<(valor+fracao)
)) {
      A1[linha,colunaA]=valor
      valor<-valor+fracao
    } else
    {
      if ( (M1[linha,colunaM]>(valor+fracao) ) && (M1[linha,colunaM]<4.10)) {
        valor<-valor+fracao
        linha<-linha-1
      }
    }
  }
  if(M1[linha,colunaM]>4.10)
  {
    A1[linha,colunaA]=4.10
  }
  linha<-linha+1
}
}

A2<-matrix(0,n2,(d*4)) for(colunaM in 1:d) {
  for(i in 1:4)
  {
    if(i == 1)
    {
      fracao<-(4.10/10)
    } else
    {
      if(i == 2)
      {
        fracao<-(4.10/6)
      } else
      {
        if(i == 3)
        {
          fracao<-(4.10/4)
        }else

```



```

        {
            if(i == 4)
            {
                fracao<-(4.10/3)
            }
        }
    }
}

valor<-fracao
linha=1
colunaA=i+(colunaM-1)*4
while(linha<=n2)
{
    if(M2[linha,colunaM]<valor)
    {
        A2[linha,colunaA]=(valor-fracao)
    } else
    {
        if ((M2[linha,colunaM]>valor) && (M2[linha,colunaM]<(valor+fracao)
        )) {
            A2[linha,colunaA]=valor
            valor<-valor+fracao
        } else
        {
            if ( (M2[linha,colunaM]>(valor+fracao) ) &&
            (M2[linha,colunaM]<4.10)) {
                valor<-valor+fracao
                linha<-linha-1
            }
        }
    }

    if(M2[linha,colunaM]>4.10)
    {
        A2[linha,colunaA]=4.10
    }
    linha<-linha+1
}
}

A3<-rbind(A1,A2)

B<-A3[,] B<-ifelse(A3[,]==0,0.01,A3[,])

A4<-matrix(0,n1,(d*4)) for(colunaM in 1:d) {
    for(i in 1:4)

```

```

{
  if(i == 1)
  {
    fracao<-(4.10/10)
  } else
  {
    if(i == 2)
    {
      fracao<-(4.10/6)
    } else
    {
      if(i == 3)
      {
        fracao<-(4.10/4)
      }else
      {
        if(i == 4)
        {
          fracao<-(4.10/3)
        }
      }
    }
  }
  valor<-fracao
  linha=1
  colunaA=i+(colunaM-1)*4
  while(linha<=n1)
  {
    if(M1[linha,colunaM]<valor)
    {
      A4[linha,colunaA]=valor
    } else
    {
      if ((M1[linha,colunaM]>valor) && (M1[linha,colunaM]<(valor+fracao)
    )) {
        A4[linha,colunaA]=valor+fracao
        valor<-valor+fracao
      } else
      {
        if ( (M1[linha,colunaM]>(valor+fracao) ) &&
        (M1[linha,colunaM]<4.10)) {
          valor<-valor+fracao
          linha<-linha-1
        }
      }
    }
  }
}

```

```

        if(M1[linha,colunaM]>4.10)
        {
            A4[linha,colunaA]=4.11
        }
        linha<-linha+1
    }
}

ce1<-A4[,]
ce1<-ifelse(ce1[,]<4.11,1,0)
A4<-ifelse(A4[,]<4.11,A4[,],NA)

A5<-matrix(0,n2,(d*4)) for(colunaM in 1:d) {
    for(i in 1:4)
    {
        if(i == 1)
        {
            fracao<-(4.10/10)
        } else
        {
            if(i == 2)
            {
                fracao<-(4.10/6)
            } else
            {
                if(i == 3)
                {
                    fracao<-(4.10/4)
                }else
                {
                    if(i == 4)
                    {
                        fracao<-(4.10/3)
                    }
                }
            }
        }
        valor<-fracao
        linha=1
        colunaA=i+(colunaM-1)*4
        while(linha<=n2)
        {
            if(M2[linha,colunaM]<valor)
            {
                A5[linha,colunaA]=valor
            }
        }
    }
}

```

```

    } else
    {
if ((M2[linha,colunaM]>valor) && (M2[linha,colunaM]<(valor+fracao)
)) {
        A5[linha,colunaA]=valor+fracao
        valor<-valor+fracao
    } else
    {
if ( (M2[linha,colunaM]>(valor+fracao) ) &&
(M2[linha,colunaM]<4.10)) {
        valor<-valor+fracao
        linha<-linha-1
    }
    }
    }
    if (M2[linha,colunaM]>4.10)
    {
        A5[linha,colunaA]=4.11
    }
    linha<-linha+1
}
}

ce2<-A5[,]
ce2<-ifelse(ce2[,]<4.11,1,0)
A5<-ifelse(A5[,]<4.11,A5[,],NA)

A6<-rbind(A4,A5)
ce<-rbind(ce1,ce2)
ce<-ifelse(ce==1,3,0)

require(survival)
trat<-c(rep(0,26),rep(1,24))
trat<-matrix(trat)

for (k in 1:f){
fit<-survreg(Surv(B[,k],A6[,k],type="interval2")~trat,dist='weibull')
}

```

ANEXO N - Método M5: considerando que o evento ocorreu no início, ponto médio ou final do intervalo para amostra de tamanho 50, $\gamma = 0.5$

```
n1<-26
```

```
n2<-24
```

```

n<-50
a<-0.5
b1<-1
b2<-2.72
d<-1000
f<-4000
M1<-matrix(0,n1,d)
for(k in 1:d){
  w1<-rweibull(n1,a,b1)
  w1<-w1[order(w1)]
  for(i in 1:n1){
    M1[i,k]=M1[i,k]+w1[i] } }

M2<-matrix(0,n2,d)
for(k in 1:d){
  w2<-rweibull(n2,a,b2)
  w2<-w2[order(w2)]
  for(i in 1:n2){
    M2[i,k]=M2[i,k]+w2[i] } }

A1<-matrix(0,n1+1,(d*4)) for(colunaM in 1:d) { for(i in 1:4) {
  if(i == 1) { fracao<-(4.10/10) } else
    {
  if(i == 2) { fracao<-(4.10/6) } else
    {
  if(i == 3) { fracao<-(4.10/4) }else
    {
  if(i == 4) { fracao<-(4.10/3) }
    }
    }
    }
  valor<-fracao linha=1 colunaA=i+(colunaM-1)*4
    while(linha<=n1)
  {
    if(M1[linha,colunaM]>4.10)
    {
      A1[linha,colunaA]=4.11
    }else
  {
    if(M1[linha,colunaM]<valor)
    {
      A1[linha,colunaA]=(valor-fracao)
    } else
    {
      if (M1[linha,colunaM]>valor)
      {

```

```

valor<-valor+fracao linha<-linha-1
}

}
} linha<-linha+1 }

A1[n1+1,colunaA]<-(A1[n1,colunaA]+fracao) } }

for(colunaM in 1:d) { for(i in 1:4) { if(i == 1) {
fracao<-(4.10/10) gride<-11 } else
{
if(i == 2) { fracao<-(4.10/6) gride<-7 } else
{
if(i == 3) { fracao<-(4.10/4) gride<-5 }else
{
if(i == 4) { fracao<-(4.10/3) gride<-4 }
}
}
}
}
s<-0 freq<-matrix(0,1,(gride+2)) freqval<-matrix(0,1,(gride+2))
for(fr in 1:(gride))
{
freqval[fr+1]<-s s<-s+fracao
}
valor<-(fracao-0.01) soma<-0 z<-0 linha<-1
colunaA=i+(colunaM-1)*4 while(linha<=(n1+1)) {
if(A1[linha,colunaA]<valor)
{
soma<-soma+1
linha<-linha+1
}else
{ z<-z+1 freq[z+1]<-soma valor<-valor+fracao soma<-0 } }
freq[1,1]<-freq[1,2] freq[1,gride+1]<-freq[1,gride] sf<-0
for(i in 1:(gride-1)) { if(freq[1,i+1]>0) {
if((freq[1,i+2]-freq[1,i])==0) { for(j in 1:(freq[1,i+1])) {
sf<-sf+1 A1[sf,colunaA]<-(freqval[i+1]+freqval[i+2])/2 } }

if((freq[1,i+2]-freq[1,i])>0) { for(j in 1:(freq[1,i+1])) {
sf<-sf+1 A1[sf,colunaA]<-freqval[i+2] } }

if((freq[1,i+2]-freq[1,i])<0) { for(j in 1:(freq[1,i+1])) {
sf<-sf+1 } } } } } }

D1<-matrix(0,n1,(d*4)) for(j in 1:(d*4))
for(i in 1:26)
D1[i,j]<-A1[i,j]

```

```

ce1<-D1[,] ce1<-ifelse(ce1[,]<4.11,1,0)
D1<-ifelse(D1[,]>4.10,4.10,D1[,])

A2<-matrix(0,n2+1,(d*4)) for(colunaM in 1:d) { for(i in 1:4) {
if(i == 1) { fracao<-(4.10/10) } else
{
if(i == 2) { fracao<-(4.10/6) } else
{
if(i == 3) { fracao<-(4.10/4) }else
{
if(i == 4) { fracao<-(4.10/3) }
}
}
}
valor<-fracao linha=1 colunaA=i+(colunaM-1)*4
while(linha<=n2)
{
if(M2[linha,colunaM]>4.10)
{
A2[linha,colunaA]=4.11
}else
{
if(M2[linha,colunaM]<valor)
{
A2[linha,colunaA]=(valor-fracao)
} else
{
if (M2[linha,colunaM]>valor)
{
valor<-valor+fracao linha<-linha-1
}

}
} linha<-linha+1 }

A2[n2+1,colunaA]<-(A2[n2,colunaA]+fracao) } }

for(colunaM in 1:d) { for(i in 1:4) { if(i == 1) {
fracao<-(4.10/10) gride<-11 } else
{
if(i == 2) { fracao<-(4.10/6) gride<-7 } else
{
if(i == 3) { fracao<-(4.10/4) gride<-5 }else
{
if(i == 4) { fracao<-(4.10/3) gride<-4 }

```

```

    }
  }
}
s<-0 freq<-matrix(0,1,(gride+2)) freqval<-matrix(0,1,(gride+2))
  for(fr in 1:(gride))
  {
freqval[fr+1]<-s s<-s+fracao
  }
valor<-(fracao-0.01) soma<-0 z<-0 linha<-1
colunaA=i+(colunaM-1)*4 while(linha<=(n2+1)) {
  if(A2[linha,colunaA]<valor)
  {
    soma<-soma+1
    linha<-linha+1
  }else
  { z<-z+1 freq[z+1]<-soma valor<-valor+fracao soma<-0 } }
freq[1,1]<-freq[1,2] freq[1,gride+1]<-freq[1,gride] sf<-0
for(i in 1:(gride-1)) { if(freq[1,i+1]>0) { if((freq[1,i+2]-freq[1,i])==0)
{ for(j in 1:(freq[1,i+1])) { sf<-sf+1
A2[sf,colunaA]<-(freqval[i+1]+freqval[i+2])/2 } }

if((freq[1,i+2]-freq[1,i])>0) { for(j in 1:(freq[1,i+1])) {
sf<-sf+1 A2[sf,colunaA]<-freqval[i+2] } }

if((freq[1,i+2]-freq[1,i])<0) { for(j in 1:(freq[1,i+1])) {
sf<-sf+1 } } } } } }

D2<-matrix(0,n2,(d*4)) for(j in 1:(d*4))
  for(i in 1:24)
    D2[i,j]<-A2[i,j]

ce2<-D2[,] ce2<-ifelse(ce2[,]<4.11,1,0)
D2<-ifelse(D2[,]>4.10,4.10,D2[,])

ce<-rbind(ce1,ce2) D<-rbind(D1,D2)
B<-D[,]
B<-ifelse(D[,]==0,0.01,D[,])

require(survival)

trat<-c(rep(0,26),rep(1,24))
trat<-matrix(trat)
for (k in 1:f){
fit<-survreg(Surv(B[,k],ce[,k])~trat,dist='weibull')
}

```