

UNIVERSIDADE ESTADUAL DE CAMPINAS
DEPARTAMENTO DE ESTATÍSTICA - IMECC

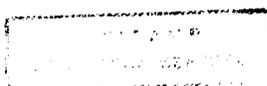
O Modelo de Regressão Gama Generalizada para
Discriminar entre Modelos Paramétricos de
Tempo de Vida

Dione Maria Valença *DMV*

Orientação
Prof. Dr. Jonathan Biele *J. Biele*

Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência de Computação da
Universidade Estadual de Campinas para obtenção de título de Mestre em Estatística
Campinas - S.P.

1994



Este exemplar corresponde à redação final da tese devidamente corrigida e defendida pela srta. **Dione Maria Valença**, e aprovada pela Comissão Julgadora.

Campinas, 25 de março de 1994.

Prof. Dr. 
Jonathan Biele

Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência da Computação, UNICAMP, como requisito parcial para obtenção do Título de Mestre em Estatística.

Dedico este trabalho aos meus pais
Jose Erisberto Valença e
Marinete Ribeiro Valença

Gostaria de expressar meus agradecimentos

- Ao CNPq, pelo apoio financeiro
- Ao professor Jonathan, pela amizade e pela enorme dedicação com que me orientou durante todas as etapas deste trabalho
- Aos professores da UFRN pela boa vontade que tiveram, em me conceder um tempo hábil para o término desta dissertação
- À professora Marcília, da UFPE, pelo estímulo
- Aos amigos que conquistei durante o mestrado, em especial a Suely e Silvia
- A Pledson, pelo companheirismo e incentivo em todas as horas.

Sumário

0.1	Introdução	7
1	Dados de Tempo de Vida - Conceitos Básicos	9
1.1	Introdução	9
1.1.1	Descrição dos Dados	9
1.1.2	Características da Distribuição dos Dados	10
1.1.3	Amostras Censuradas	12
1.1.4	Estimação Não Paramétrica da Função de Sobrevivência	13
1.2	Principais Modelos Paramétricos	14
1.2.1	Distribuição Exponencial	15
1.2.2	Distribuição Weibull	16
1.2.3	Distribuição Gama	17
1.2.4	Distribuição Log-Normal	19
1.2.5	Modelos de Regressão	20
1.3	Inferência com Modelos de Tempo de Vida	23
1.3.1	Função de Verossimilhança Incorporando Censura	23
1.3.2	Estimação pelo Método da Máxima Verossimilhança	25
1.3.3	Teoria Assintótica	26
1.4	Aplicação - Testes Acelerados	31
1.5	Discussão Sobre a Adequacidade de um Modelo	32
1.5.1	Análise de Resíduos	32
1.5.2	A Escolha de um Modelo	33
2	A Distribuição Gama Generalizada	35

2.1	Introdução	35
2.2	Propriedades da Distribuição Gama Generalizada	35
2.2.1	Função de Densidade	36
2.2.2	Análise Gráfica da Densidade	38
2.2.3	Função de distribuição	40
2.2.4	Função de Sobrevivência e Função Risco	41
2.2.5	Função Geratriz de Momentos	42
2.2.6	Momentos	42
2.2.7	Distribuição de Algumas Funções de Variáveis Gama Generalizada	44
2.3	O Modelo Log-Gama Generalizada Reparametrizado	53
2.3.1	Descrição do Modelo	53
2.3.2	Simplificação do Modelo Reparametrizado	56
2.4	O Modelo de Regressão Gama Generalizada	59
2.4.1	Descrição do Modelo	60
3	Procedimentos Inferenciais para o Modelo Gama Generalizada	62
3.1	Introdução	62
3.1.1	Breve Resumo Bibliográfico	62
3.2	Inferências com Amostras Completas	65
3.2.1	Estimação Pontual	65
3.2.2	Intervalos de Confiança e Testes de Hipóteses	70
3.3	Inferências com Dados Censurados e Variáveis Regressoras	74
3.3.1	Procedimentos Inferenciais com Amostras Censuradas Sem Variáveis Regressoras	74
3.3.2	Procedimentos Inferenciais para o Modelo de Regressão	80
4	Ajuste do Modelo Gama Generalizada a Dados de Vida	91
4.1	Análise de Dados Reais	91
4.1.1	Exemplo 1- Teste de Vida de Isolantes	92

4.1.2	Exemplo 2 - Teste de Vida de Capacitores	97
4.1.3	Exemplo 3 - Sobrevivência de Portadores de Mieloma	100
4.1.4	Exemplo 4 - Sobrevivência de Portadores da Doença de Hodgkin	104
4.1.5	Exemplo 5 - Corrosão de Metais	109
4.2	Comentários	113
5	Considerações Finais	115
A	Função Gama e Distribuição Beta	117
B	Propriedades da Distribuição Log-Gama	121
C	O Método de Newton Raphson	125
D	Métodos Computacionais	127

Lista de Figuras

1-1	Curva de Kaplan Meier para os tempos de vida de dois tipos de isolantes.	15
2-1	Algumas formas assumidas pela densidade gama generalizada $f(t; \alpha, \beta, k)$.	39
2-2	Algumas formas assumidas pela função de risco da gama generalizada, $h(t; \alpha, \beta, k)$.	43
2-3	Formas da densidade log gama generalizada (l.g.g) estendida, para $q = -2, 0$ e 2 .	57
4-1	À esquerda, o log dos tempos de vida de isolantes por nível de temperatura, e à direita, por nível transformado de temperatura (relacionamento Arrhenius).	94
4-2	Curva de sobrevivência estimada para a distribuição log-gama ($q = 0.17$) e gráfico de Kaplan-Meier para resíduos estimados. Exemplo 1 - Isolantes de motores.	96
4-3	Contorno da verossimilhança relativa maximizada $R_{\max}(q)$, para dados de teste de vida de isolantes de motores.	96
4-4	Curva de sobrevivência estimada para a distribuição log-gama ($q = 0.52$) e função de sobrevivência empírica para resíduos estimados. Exemplo 2 - Capacitores.	99
4-5	Contorno da verossimilhança relativa maximizada $R_{\max}(q)$, para dados de teste de vida de capacitores.	100
4-6	Contorno da verossimilhança relativa maximizada $R_{\max}(q)$, para dados de portadores de mieloma múltipla.	104

4-7	Curva de sobrevivência estimada para a distribuição log-gama ($q = -0.6$) e o gráfico de Kaplan-Meier para os resíduos estimados. Exemplo 3 - portadores de mieloma múltipla.	105
4-8	Curva de sobrevivência para a distribuição normal padrão e o gráfico de Kaplan - Meier para resíduos estimados. Exemplo 3 - Portadores de mieloma múltipla.	105
4-9	Curva de sobrevivência para a distribuição valor extremo padrão e o gráfico de Kaplan - Meier para resíduos estimados. Exemplo 3 - Portadores de mieloma múltipla.	106
4-10	Curva de sobrevivência estimada para a distribuição log-gama e função de Kaplan - Meier para resíduos estimados. Exemplo 4 - portadores da doença de Hodgkin	107
4-11	Contorno da verossimilhança relativa maximizada $R_{\max}(q)$, para dados de portadores da doença de Hodgkin.	108
4-12	À esquerda, mudança de peso de metais em atmosfera corrosiva, por níveis de umidade relativa, e à direita o mesmo gráfico com a transformação logarítima para as mudanças de peso.	111
4-13	Curva de sobrevivência estimada para a distribuição log-gama ($q = -0.09$) e função de sobrevivência empírica dos resíduos estimados. Exemplo 5 - corrosão de metais.	112
4-14	Contorno da verossimilhança relativa maximizada $R_{\max}(q)$, para dados de corrosão de metais.	112

Lista de Tabelas

2.1	Casos particulares da d.g.g., com densidade $f(t; \alpha, \beta, k)$ - Stacy e Miram (1965).	36
4.1	Tempo de vida de isolantes (horas) para cada temperatura, em graus centígrados, com + indicando censura (Nelson, 1990).	93
4.2	Etapas para a estimação dos parâmetros para dados do teste de vida de isolante de motores.	94
4.3	Resultados do teste de vida de capacitores de vidro (Zelen, 1959).	97
4.4	Etapas para estimação dos parâmetros referentes a dados de teste de vida de capacitores.	98
4.5	Tempos de sobrevivência (em meses) de pacientes portadores de mieloma múltipla (Lawless, 1982).	102
4.6	Etapas para estimação dos parâmetros para dados de sobrevivência de portadores de mieloma.	103
4.7	Tempos de sobrevivência de portadores da doença de Hodgkin tratados com terapia de gás nitrogênio. Grupo A sem terapia prévia e grupo B com terapia prévia (Bartolucci e Dickey, 1977).	107
4.8	Etapas para estimação dos parâmetros para dados de sobrevivência de portadores da doença de Hodgkin.	109
4.9	Perda de peso de metais em uma atmosfera corrosiva (Nelson, 1990).	110

0.1 Introdução

A análise de dados de tempo de vida representa o principal interesse de duas áreas de estudo, a teoria da Confiabilidade e a Análise de Sobrevivência. Estes dados são em geral estudados com base em modelos estatísticos paramétricos, que são supostamente apropriados. A escolha correta de um modelo que ajuste um específico conjunto de dados é de grande importância, tendo em vista que a precisão dos resultados da análise dependem desta decisão.

Esta dissertação trata basicamente do estudo de uma família de distribuições que pode ser usada para discriminar dentre alguns *modelos de tempo de vida*, o mais adequado para o ajuste de uma amostra. Especificamente, tratamos da família **gama generalizada com 3 parâmetros**, que possui como casos particulares os principais modelos paramétricos comumente utilizados no ajuste deste tipo de dados. A discriminação é feita através de testes sobre parâmetros desta distribuição.

Abordamos inicialmente os principais conceitos com respeito à teoria desenvolvida para a análise de dados de tempo de vida, incluindo características especiais dos dados (censuras), modelos mais usados, procedimentos inferenciais, e uma breve discussão sobre o processo de escolha de um modelo.

No capítulo 2 apresentamos definições e demonstramos diversas propriedades, da distribuição gama generalizada, e também, descrevemos o processo de reparametrização e extensão que leva sua densidade a uma forma operacionalmente mais simples.

No terceiro capítulo estudamos os procedimentos inferenciais de estimação e teste dos parâmetros, principalmente para o modelo reparametrizado. Tratamos também de populações heterogêneas através do modelo de regressão gama generalizada reparametrizado e de procedimentos inferenciais específicos para a análise de amostras censuradas.

O capítulo 4 mostra aplicações de parte da teoria estudada através da análise de dados reais encontrados na literatura. O objetivo principal é, com base no ajuste e teste do modelo que estudamos, julgar a adequação dos modelos previamente adotados. O quinto e último capítulo apresenta algumas considerações finais sobre esta dissertação.

Fornecemos também 4 apêndices que são necessários ao entendimento de alguns resultados apresentados nos capítulos anteriores. Nos apêndices A e B mostramos algumas definições e características importantes sobre as funções gama e beta e sobre a distribuição log-gama, respectivamente. No apêndice C descrevemos brevemente o método iterativo de Newton Raphson. No apêndice D relatamos os métodos computacionais adotados e fornecemos o programa elaborado em linguagem S-Plus, para estimar e testar os parâmetros do modelo de regressão gama generalizada reparametrizado em escala logaritmica.

Capítulo 1

Dados de Tempo de Vida - Conceitos Básicos

1.1 Introdução

Este trabalho trata da modelagem adequada de dados que em geral correspondem ao tempo de ocorrência de algum evento, ou mais genericamente, a valores de uma variável aleatória positiva. Dados deste tipo são chamados de *dados de tempo de vida*, mas são também referidos como dados de tempo de sobrevivência ou tempo até a falha, ou simplesmente como dados de vida. Neste capítulo apresentamos conceitos, abordando características específicas da distribuição dos dados, descrição de modelos probabilísticos e procedimentos inferenciais adequados. Na última seção discutimos sobre a importância da modelagem correta destes dados e introduzimos o procedimento de escolha de um modelo através da família de distribuição gama generalizada .

1.1.1 Descrição dos Dados

A teoria da Confiabilidade e a Análise de Sobrevivência têm como objetivo principal a análise de dados de tempo de vida. Na Confiabilidade, temos como exemplo os **testes de vida** onde os dados utilizados referem-se ao tempo até a falha de um item, enquanto que na Análise de Sobrevivência os dados representam em geral o tempo de vida ou tempo de remissão de uma doença, de pacientes ou de animais sob experimentação biológica. Neste

texto usamos com frequência uma linguagem direcionada a aplicações de Confiabilidade, sendo a unidade de estudo chamada de **ítem**. Contudo, as técnicas específicas e modelos estatísticos usados são desenvolvidos para qualquer área de aplicação que possua dados deste tipo.

Tomemos os exemplos abaixo como ilustração:

1. Frequentemente são realizados testes de vida em ítems de manufatura, como componentes eletrônicos ou mecânicos, com o objetivo de obter informações sobre a sua durabilidade. Para isto, os ítems são colocados em operação e são observados até o momento em que falhem.
2. Em estudos médicos que tratam de doenças fatais, o pesquisador tem como objetivo analisar o tempo de sobrevivência de indivíduos com esta doença, medindo o tempo de vida dos pacientes a partir da data do diagnóstico, ou de algum outro ponto de início. Um exemplo comum é a comparação de tratamentos de uma doença em termos do tempo de sobrevivência dos pacientes que receberam tratamentos diferentes.

Em geral examinamos estes dados com base em modelos estatísticos que são supostamente apropriados. A seguir, consideramos algumas definições importantes sobre distribuições de dados de tempo de vida para o caso de modelos contínuos univariados.

1.1.2 Características da Distribuição dos Dados

Seja T uma variável aleatória (v.a.) contínua e positiva, representando em geral o tempo até a falha de um ítem. Três importantes representações da distribuição de T são dadas abaixo.

- Função de Densidade

A função de densidade de probabilidade (f.d.p.) de T é definida como

$$f(t) = \lim_{\Delta t \rightarrow 0} p(t \leq T \leq t + \Delta t) = \frac{\partial F(t)}{\partial t},$$

onde $F(t) = P(T \leq t)$ é a função de distribuição de T, e temos que

$$f(t) \geq 0 \quad e \quad \int_0^{\infty} f(t)dt = 1.$$

- Função de Sobrevivência

Definimos a função de sobrevivência de T como sendo

$$S(t) = P(T \geq t) = \int_t^{\infty} f(x)dx.$$

Esta função representa a probabilidade de um ítem (ou indivíduo) sobreviver pelo menos até o tempo t . Algumas vezes $S(t)$ é referida como *função de confiabilidade*. Pode-se mostrar que $S(t) = 1 - F(t)$, sendo uma função monótona decrescente, contínua, com $S(0) = 1$ e $\lim_{t \rightarrow \infty} S(t) = 0$.

- Função Risco

A função risco, denotada por $h(t)$ é definida como segue,

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{p[(t \leq T \leq t + \Delta t) \setminus T \geq t]}{\Delta t} = \frac{f(t)}{S(t)}.$$

Pode-se mostrar que

$$h(t) \geq 0 \quad e \quad \int_0^{\infty} h(t)dt = \infty.$$

A função risco-especifica a taxa de falha instantânea a um tempo t , dado que o ítem sobreviveu até o tempo t . Uma distribuição de tempo de vida (ou *distribuição de vida*), pode ser melhor compreendida através da função risco, também chamada de taxa de falha, pois ela reflete o efeito do tempo sobre ítems sobreviventes. Em alguns casos existem informações sobre o comportamento da taxa de falha com relação ao tempo, e estas informações podem ser usadas na modelagem de $h(t)$. As implicações para $S(t)$ e

$f(t)$ podem ser facilmente obtidas através das relações abaixo, cujas demonstrações são diretas.

$$S(t) = \exp\left(-\int_0^t h(x)dx\right), \quad e$$

$$f(t) = h(t) \exp\left(-\int_0^t h(x)dx\right).$$

Um outro conceito útil a respeito da distribuição de T é o seu p -ésimo quantil, que corresponde ao valor t_p satisfazendo $P(T \leq t_p) = p$, isto é, $t_p = F^{-1}(p)$.

1.1.3 Amostras Censuradas

Geralmente experimentos de tempo de vida apresentam problemas especiais na análise de dados, devido a presença de observações censuradas, isto é, para alguns elementos em estudo não sabemos seu tempo exato de vida, mas apenas que este excede um certo valor. A necessidade de obter métodos de análises que acomodem censuras é, provavelmente a principal razão para o desenvolvimento de modelos e procedimentos especiais para a análise de dados de vida.

As censuras surgem de diversas maneiras. Consideremos três dos principais tipos:

- Censura Tipo II

Uma amostra censurada do tipo II é aquela onde são observados os r menores tempos de uma amostra aleatória de n ítems ($1 \leq r \leq n$). Consideremos como exemplo um teste de vida, onde um total de n ítems são colocados em teste e este terminará no momento em que o r -ésimo ítem falhar.

- Censura Tipo I

Os dados apresentam censura tipo I, se o experimento é realizado em um período de tempo fixo, de tal maneira que o tempo de vida do elemento só será conhecido se ele é

menor que um limite pré-determinado. Neste tipo de censura o número de falhas é uma variável aleatória. Especificamente temos que os elementos $1, 2, \dots, n$ de uma amostra são sujeitos a períodos de observações $T_1^C, T_2^C, \dots, T_n^C$ de forma que o tempo de vida T_i^F é observado apenas se $T_i^F \leq T_i^C$. Se $T_1^C = \dots = T_n^C$, temos a *censura tipo I simples*.

- Censura Aleatória

Neste tipo de censura, os tempos de censura T_i^C com $i = 1, \dots, n$, são considerados aleatórios. Um processo simples de censura aleatória supõe que cada elemento tem um tempo de vida T_i^F e um tempo de censura T_i^C , e que estas são v.a.'s independentes e contínuas.

Exemplo A. Os dados abaixo representam tempos até a falha de dois tipos de isolantes elétricos, obtidos de um teste acelerado de vida. Das 10 espécimes testadas de cada tipo, 9 falharam, e uma de cada tipo ainda funcionava quando o experimento foi encerrado. Estas duas observações (marcadas com *) representam censuras tipo I.

Isolante 1 : 5.1, 9.2, 9.3, 11.8, 17.7, 19.4, 22.1, 26.7, 37.3, 60*;

Isolante 2 : 11.0, 15.1, 18.3, 24.0, 29.1, 38.4, 44.2, 45.1, 50.9, 70*.

1.1.4 Estimação Não Paramétrica da Função de Sobrevivência

É comum sumarizar dados de vida em termos da função de sobrevivência empírica, que representa também uma estimativa não paramétrica da função de sobrevivência. Se não existem observações censuradas, a função de sobrevivência empírica é definida como

$$\hat{S}(t) = \frac{\text{n}^\circ \text{ de observações } \geq t}{\text{tamanho da amostra}} \quad t \geq 0.$$

Contudo, quando trabalhamos com dados censurados precisamos de uma metodologia conveniente para tratar disto. Entre alguns possíveis estimadores de $S(t)$ (ver Chiang, 1968), consideremos o mais utilizado, que é o estimador Produto-Limite, também conhecido como estimador de Kaplan-Meier, por estes autores terem sido os primeiros a discutirem suas propriedades (Kaplan e Meier, 1958).

Seja $t_1 < \dots < t_k$ os tempos de falha observados em uma amostra de tamanho n com função de sobrevivência $S(t)$. Suponha que d_j ítems falham no instante t_j ($j = 1, \dots, k$) e m_j ítems são censurados no intervalo $[t_j, t_{j+1})$ nos tempos $t_{j_1}, \dots, t_{j_{m_j}}$ com $j = 0, \dots, k$, onde $t_0 = 0$ e $t_{k+1} = \infty$. O estimador Produto-Limite de $S(t)$, conforme dado em Kalbfleisch e Prentice (1980), é definido como

$$\hat{S}(t) = \prod_{j: t_j \leq t} \left(\frac{n_j - d_j}{n_j} \right), \quad (1.1)$$

onde $n_j = (m_j + d_j) + \dots + (m_k + d_k)$, é o número de ítems em risco em t_j , ou seja, número de ítems que não falharam nem censuraram no instante exatamente anterior a t_j .

Um possível estimador da variância de $S(t)$ é

$$\hat{V}(\hat{S}(t)) = \hat{S}(t)^2 \sum_{j: t_j < t} \left(\frac{d_j}{n_j(n_j - d_j)} \right),$$

O estimador Produto-Limite possui várias propriedades importantes. A principal é que $\hat{S}(t)$ é um estimador consistente de $S(t)$ sob condições muito gerais. Algumas destas propriedades são examinados em Lawless(1982), que fornece referências de diversos trabalhos nesta área. Como ilustração temos na Figura 1-1 as curvas de Kaplan-Meier para os 2 tipos de isolantes dados no exemplo A.

1.2 Principais Modelos Paramétricos

Apesar de existirem técnicas não paramétricas que tratam de dados de vida, em diversas situações deseja-se ajustar modelos paramétricos aos dados. Na literatura existem alguns modelos que são usados com bastante frequência. Estes são os modelos exponencial, Weibull, gama e log-normal. Consideremos brevemente propriedades e características teóricas destes modelos, usados como distribuição da variável T , e alguma discussão sobre a distribuição do log de T (logaritmo natural), tendo em vista que na prática, é conveniente trabalhar com o log do tempo de vida. Informações teóricas mais detal-

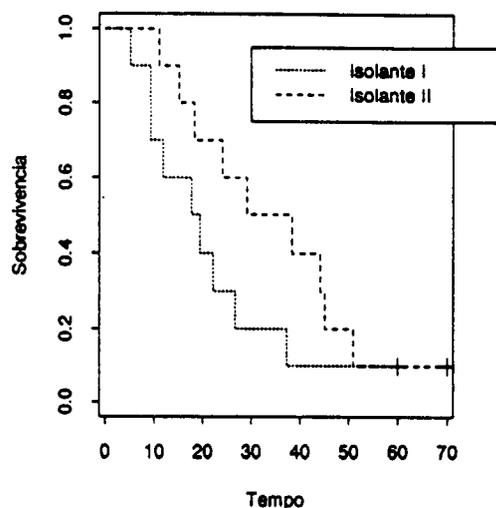


Figura 1-1: Curva de Kaplan Meier para os tempos de vida de dois tipos de isolantes.

hadas sobre estas distribuições são obtidas em Kotz e Johnson (1985). Aplicações destes modelos em Confiabilidade são encontradas em Nelson(1990).

1.2.1 Distribuição Exponencial

A distribuição exponencial de um parâmetro, por ter uma estrutura simples, é usada frequentemente tanto em áreas médicas como industriais. Se T tem distribuição exponencial com parâmetro $\alpha > 0$, a f.d.p. e a função de sobrevivência são dadas por

$$f(t) = \frac{1}{\alpha} \exp(-t/\alpha) \quad e \quad S(t) = e^{-t/\alpha} \quad t \geq 0.$$

Uma característica importante desta distribuição é o fato de possuir função risco constante,

$$h(t) = \frac{1}{\alpha} \quad t \geq 0.$$

Isto reflete a falta de memória da distribuição exponencial, e poderia ser visto da seguinte forma; para este modelo a chance condicional de falha em um intervalo de tempo de comprimento estipulado, é a mesma, independento do tempo em que o item

venha sendo testado. Desta forma, o envelhecimento do ítem não estaria influenciando no risco de falha instantânea.

A média e variância de T são respectivamente α e α^2 , e o p -ésimo quantil é $-\alpha \log(1 - p)$.

Tomando $Y = \log T$ temos que a f.d.p. de Y é dada por

$$f_Y(y) = \exp(y - \lambda - e^{y-\lambda}) \quad y \in \mathfrak{R}, \quad (1.2)$$

onde $\lambda = \log \alpha$. Podemos representar Y como $Y = \alpha + W$, sendo W uma variável aleatória com distribuição valor extremo padrão, com f.d.p e função de sobrevivência dadas abaixo,

$$f_W(w) = \exp(w - e^w) \quad e \quad S_W(w) = \exp(-e^w) \quad w \in \mathfrak{R}. \quad (1.3)$$

A média de W é $-\gamma$ ($\gamma = 0,5772\dots$ é a constante de Euler) e a variância é $\pi^2/6$. Consequentemente, a média e a variância de Y são $(\alpha - \gamma)$ e $\pi^2/6$.

1.2.2 Distribuição Weibull

A Distribuição Weibull representa uma generalização da distribuição exponencial e, de acordo com Lawless (1982), é possivelmente o modelo paramétrico mais usado no ajuste de dados de tempo de vida, tanto na área médica como na engenharia.

Se T tem distribuição Weibull com parâmetros $\alpha > 0$ e $\beta > 0$, a f.d.p. e a função de sobrevivência são dadas como segue,

$$f(t) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha}\right)^{\beta-1} \exp\left(-\left(t/\alpha\right)^\beta\right) \quad t > 0$$

$$S(t) = \exp\left(-\left(t/\alpha\right)^\beta\right) \quad t > 0.$$

A função risco de T é dada por

$$h(t) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha}\right)^{\beta-1}.$$

Esta função é monótona decrescente para $\beta < 1$, crescente para $\beta > 1$ e constante para $\beta = 1$. Percebe-se portanto que a distribuição exponencial é um caso particular da Weibull quando $\beta = 1$. A média e a variância para esta distribuição são, respectivamente

$$\alpha\Gamma(1 + 1/\beta) \quad e \quad \alpha^2[\Gamma(1 + 2/\beta) - \Gamma(1 + 1/\beta)^2].$$

Fazendo $Y = \log T$, temos que a v.a Y tem distribuição **valor extremo** com densidade e função de sobrevivência

$$f(y) = \frac{1}{\sigma} \exp \left[\frac{y - \delta}{\sigma} - \exp \left(\frac{y - \delta}{\sigma} \right) \right] \quad y \in \mathfrak{R} \quad (1.4)$$

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

onde $\sigma = 1/\beta$ e $\delta = \log \alpha$. Podemos também escrever $Y = \delta + \sigma W$ onde W tem distribuição valor extremo padrão dada em 1.3. Observamos que a distribuição dada em 1.3 é um caso particular desta quando $\delta = 0$ e $\sigma = 1$. A média e variância da distribuição representada por 1.4, são respectivamente $(\delta - \gamma\sigma)$ e $(\sigma^2\pi^2/6)$ onde $\gamma = 0.5772\dots$ (constante de Euler). O p -ésimo quantil de 1.4 é o valor $y_p = \delta + \sigma \log(-\log(1 - p))$. Mais informações sobre a distribuição Valor Extremo são dadas na página 606 do volume 2 Kotz e Johnson (1985).

1.2.3 Distribuição Gama

Em aplicações de Confiabilidade, a distribuição gama geralmente resulta da modelagem do tempo de falha de itens compostos de vários componentes, onde a falha do item ocorre quando todos os componentes falham, e a distribuição assumida para cada componente é exponencial (Tobias e Trindade, 1986). Gross e Clark (1975) aplicaram a distribuição

gama para dados em um estudo biomédico. Argumenta-se que a distribuição gama não é tão utilizada quanto a Weibull, apesar de se ajustar adequadamente a vários dados de tempo de vida, pela dificuldade operacional no seu uso.

Suponha T como distribuição gama com parâmetros α e k positivos. Suas funções de densidade e sobrevivência são dadas por

$$f(t) = \frac{1}{\alpha\Gamma(k)} \left(\frac{t}{\alpha}\right)^{k-1} e^{-t/\alpha} \quad t > 0 \quad (1.5)$$

$$S(t) = 1 - I(k, t/\alpha).$$

onde

$$I(k, x) = \frac{1}{\Gamma(k)} \int_0^x u^{k-1} e^{-u} du$$

é a função gama incompleta (ver apêndice A). Esta função relaciona-se com a função de distribuição de uma chi-quadrado, podendo ser calculada pelo uso de sua tabela (veja Lawless, 1982 - apêndice B). A média e a variância de T são respectivamente, αk e $\alpha^2 k$.

A função risco desta distribuição é dada por

$$h(t) = [1/\Gamma(k)](t^{k-1}/\alpha^k) \exp(-t/\alpha)/[1 - I(k, t/\alpha)].$$

Pode-se mostrar que $h(t)$ é monótona crescente para $k > 1$, com $h(0) = 0$ e $\lim_{t \rightarrow \infty} h(t) = 1/\alpha$ e para $k = 1$ é constante, isto é, $h(t) = 1/\alpha$. Para $0 < k < 1$, $h(t)$ é monótona decrescente, sendo $\lim_{t \rightarrow 0} h(t) = \infty$ e $\lim_{t \rightarrow \infty} h(t) = 1/\alpha$. Nelson (1982, pag. 46) apresenta um gráfico com algumas formas para esta função.

Se $\alpha = 1$, T tem distribuição gama de um parâmetro com densidade

$$f(t) = \frac{1}{\Gamma(k)} t^{k-1} e^{-t} \quad t > 0. \quad (1.6)$$

Se $k = 1$ temos a distribuição exponencial dada na seção (1.2.1). Uma importante relação existente entre a distribuição gama e a exponencial é o fato de a soma de expo-

nenciais independentes e identicamente distribuídas ter distribuição gama.

Se $Y = \log T$, podemos escrever $Y = \alpha + W$ onde W tem distribuição **log-gama**, com densidade dada por

$$f(w) = \exp(kw - e^w)/\Gamma(k) \quad w \in \mathfrak{R}. \quad (1.7)$$

Devido a importância desta distribuição neste trabalho, suas propriedades foram apresentadas com mais detalhes no Apêndice B.

1.2.4 Distribuição Log-Normal

A distribuição log-normal é muito usada para ajustar dados de vida. De acordo com Nelson (1990), existem diversas aplicações deste modelo em testes de vida de itens. Na área médica um exemplo de seu uso é o estudo do tempo de vida até o aparecimento de cancer de Pulmão em fumantes (Whitemore e Altschuler, 1976). Uma discussão detalhada sobre este modelo pode ser encontrada no livro de Crow e Shimizu (1988).

Se T tem distribuição log-normal, sua f.d.p. é dada por

$$f(t) = \frac{1}{(2\pi)^{1/2}\sigma t} \exp \left\{ -\frac{1}{2} \left(\frac{\log t - \mu}{\sigma} \right)^2 \right\} \quad t > 0$$

Uma importante característica de T é o fato de $Y = \log T$ possuir distribuição normal com média μ e variância σ^2 . Podemos então escrever $Y = \mu + \sigma Z$ onde Z tem distribuição normal padrão com função de distribuição dada por

$$\Phi(z) = \int_{-\infty}^z \frac{1}{(2\pi)^{1/2}} e^{-x^2/2} dx, \quad z \in \mathfrak{R}. \quad (1.8)$$

É fácil mostrar que a função de sobrevivência de uma log-normal é relacionada com $\Phi(z)$ da seguinte forma,

$$S(t) = 1 - \Phi \left(\frac{\log t - \mu}{\sigma} \right).$$

Como a função de Sobrevivência, a função risco também não apresenta uma forma fechada, mas pode-se mostrar que, $h(t)$ é zero para $t = 0$, cresce para um máximo e decresce aproximando-se de zero quando $t \rightarrow \infty$. Esta é uma característica pouco esperada para uma função risco (i.e., o risco de falha instantânea diminuir com o tempo). Apesar disto, esta distribuição é frequentemente usada quando os valores grandes de T não são de interesse.

A média e a variância de T são dadas por $\exp(\mu + \sigma^2/2)$ e $[\exp(\sigma^2) - 1] \exp(2\mu + \sigma^2)$, respectivamente. O p -ésimo quantil de T é $t_p = \exp[\mu + \sigma\Phi^{-1}(p)]$. Uma quantidade mais conhecida é o quantil de Y . $y_p = \mu + \sigma\phi^{-1}(p)$, que representa o quantil da distribuição normal. Poderíamos usar esta relação para estimar de maneira simples μ e σ , através do gráfico de probabilidade normal, como respectivamente, o intercepto e a inclinação de um linha reta ajustada aos pontos.

1.2.5 Modelos de Regressão

Discutimos até agora sobre modelagem para populações homogêneas. Contudo temos na prática muitas situações onde o tempo de vida é influenciado por uma ou mais *variáveis regressoras*. Por exemplo, o tempo de vida de um determinado isolante é naturalmente afetado pelo nível de voltagem a que ele é sujeito, e o tempo de sobrevivência de um paciente pode depender da idade, da quantidade de cigarros que ele fuma por dia, e de uma série de outros fatores.

Considere T uma v.a. representando tempo até a falha de um ítem e seja $\mathbf{x} = (x_0, x_1, \dots, x_{p-1})$ um vetor formado por observações de $p - 1$ variáveis regressoras (x_1, \dots, x_{p-1}) , também chamadas variáveis explanatórias ou covariáveis, que podem ser quantitativas ou qualitativas, e $x_0 = 1$. Por simplificação chamamos \mathbf{x} de vetor de variáveis regressoras. Uma maneira de determinar o relacionamento entre T e \mathbf{x} é através de um *Modelo de Regressão*. Duas classes importantes de modelos de regressão são:

- Modelos de riscos proporcionais para T , e

- Modelos de locação e escala para log de T .

Abordaremos aqui apenas a segunda classe. Uma descrição detalhada sobre modelos de Riscos proporcionais pode ser obtida em Kalbfleisch e Prentice (1980). Lee (1980), Cox e Oakes (1989), entre outros.

Modelos de Locação e Escala

A classe de modelos de locação e escala se caracteriza pelo fato de $Y = \log T$ ter uma distribuição com parâmetros de locação $\mu(\mathbf{x})$ dependendo das variáveis regressoras, e um parâmetro de escala σ constante. Podemos então escrever

$$Y = \mu(\mathbf{x}) + \sigma W, \quad (1.9)$$

onde $\sigma > 0$ e W tem uma distribuição que não depende de \mathbf{x} .

A função de sobrevivência de Y dado \mathbf{x} é da forma, $G\left(\frac{y-\mu(\mathbf{x})}{\sigma}\right)$ onde $G(\cdot)$ é a função de sobrevivência de W . Assim, é fácil ver que a função de sobrevivência de T dado \mathbf{x} é

$$S(t|\mathbf{x}) = G\left[\log(t/\alpha(\mathbf{x}))^\delta\right],$$

onde $\alpha(\mathbf{x}) = e^{\mu(\mathbf{x})}$ e $\delta = 1/\sigma$. Uma conhecida característica deste modelo é que as variáveis regressoras atuam multiplicativamente sobre T .

Podemos assumir várias formas para $\mu(\mathbf{x})$ dependendo do específico conjunto de dados que estudamos. Em muitas situações, a maneira mais natural é assumir $\mu(\mathbf{x}) = \mathbf{x}\beta$, onde $\beta = (\beta_0, \dots, \beta_{p-1})'$ é um vetor de parâmetros desconhecidos (tratamos aqui apenas desta forma para $\mu(\mathbf{x})$). Com esta suposição, 1.9 é um modelo log linear para T com resíduo W .

Consideremos brevemente os modelos de regressão log-linear mais importantes, que representam extensões das distribuições exponencial, weibull e log-normal, discutidas no início desta seção.

Modelo de Regressão Exponencial

Considere T dado \mathbf{x} com distribuição exponencial e seja $Y = \log T$. O modelo de regressão log-linear para T , representando uma generalização de 1.2, é dado por

$$Y = \mathbf{x}\beta + W,$$

onde W tem distribuição valor extremo padrão dada em 1.3. As funções de densidade e sobrevivência de Y dado \mathbf{x} são,

$$f_Y(y|\mathbf{x}) = \exp [y - \mathbf{x}\beta - \exp (y - \mathbf{x}\beta)] \quad y \in \mathfrak{R}, \quad (1.10)$$

e

$$S_Y(y|\mathbf{x}) = \exp [-\exp (y - \mathbf{x}\beta)].$$

Modelo de Regressão Weibull

Se T dado \mathbf{x} tem distribuição Weibull com parâmetros $\alpha(\mathbf{x}) = e^{\mathbf{x}\beta}$, e $b = \frac{1}{\sigma}$ o modelo de regressão log-linear, como generalização de 1.4 é da forma

$$Y = \mathbf{x}\beta + \sigma W,$$

onde W tem, como antes, distribuição valor extremo padrão dada por 1.3. A f.d.p. e a função de sobrevivência para Y dado \mathbf{x} , são

$$f_Y(y|\mathbf{x}) = \frac{1}{\sigma} \exp \left[\left(\frac{y - \mathbf{x}\beta}{\sigma} \right) - \exp \left(\frac{y - \mathbf{x}\beta}{\sigma} \right) \right] \quad y \in \mathfrak{R}, \quad (1.11)$$

$$S_Y(y|\mathbf{x}) = \exp - \exp \left(\frac{y - \mathbf{x}\beta}{\sigma} \right).$$

Modelo de Regressão Log-Normal

O modelo de regressão log-linear, cujo tempo de vida T dado \mathbf{x} tem distribuição log-normal, pode ser representado por

$$Y = \mathbf{x}\beta + \sigma Z,$$

onde Z tem distribuição normal padrão, com função de distribuição $\Phi(z)$ dada em 1.8. A densidade e a função de sobrevivência para este modelo são

$$f_Y(y|\mathbf{x}) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{1}{2} \left(\frac{y - \mathbf{x}\beta}{\sigma} \right)^2 \right] \quad y \in \mathfrak{R}, \quad (1.12)$$

$$S_Y(y|\mathbf{x}) = 1 - \Phi \left(\frac{y - \mathbf{x}\beta}{\sigma} \right).$$

Tratamos a seguir do problema da estimação dos parâmetros de um modelo de vida.

1.3 Inferência com Modelos de Tempo de Vida

Assumindo um modelo paramétrico como adequado para a análise dos dados, desejamos elaborar inferências com base neste modelo. Em geral esta análise torna-se mais complicada quando precisamos incorporar dados censurados, mesmo quando o mecanismo de censura é simples. Abordamos aqui métodos da verossimilhança para estimação e teste de modelos paramétricos de regressão.

1.3.1 Função de Verossimilhança Incorporando Censura

Consideramos que em um estudo de sobrevivência n ítems são colocados em teste, e que associado a cada ítem temos uma v.a. T_i que representa tempo de vida ou tempo até a censura, e um vetor de regressão $\mathbf{x}_i = (x_{i0}, x_{i1}, \dots, x_{ip-1})$, onde x_{i0} é identicamente 1.

Para tratarmos de populações homogêneas, consideramos $p = 1$, ou seja, um modelo sem variáveis regressoras.

Os dados observados podem ser representados da forma $(t_i, \delta_i, \mathbf{x}_i)$, com $i = 1, \dots, n$, onde

$$\delta_i = \begin{cases} 1 & \text{se o } i\text{-ésimo item falhou, ou seja, } i \in F \\ 0 & \text{se o } i\text{-ésimo item é censurado, ou } i \in C. \end{cases}$$

As notações $i \in F$ e $i \in C$ referem-se ao item para o qual t_i é tempo até a ocorrência de falha ou tempo até a ocorrência de censura, respectivamente. Seja $\theta = (\theta_1, \dots, \theta_r)$ o vetor de parâmetros desconhecidos que desejamos estimar, pertencendo a um espaço paramétrico Θ , e T_i^F uma v.a. representando **tempo de vida** do i -ésimo elemento que falhou, com funções de densidade e sobrevivência denotadas por

$$f(t_i; \theta, \mathbf{x}_i) \text{ e } S(t_i; \theta, \mathbf{x}_i), \quad \forall i \in F.$$

Para obter a função de verossilhança de θ , precisamos considerar o tipo de censura a que os dados estão sujeitos. Assumimos inicialmente censura aleatória, denotando por T_i^C o **tempo até a censura** do i -ésimo item, e considerando que esta é uma v.a. contínua com função de densidade e sobrevivência g_i e G_i , respectivamente. Assumimos também que $T_i^C, \forall i \in C$ são independentes entre si e dos tempos de falha $T_i^F, \forall i \in F$.

Temos então que,

$$\begin{aligned} P(T_i \in (t, t + dt), \delta = 1; \mathbf{x}_i, \theta) &= P(T_i^F \in (t, t + dt), T_i^C > t; \mathbf{x}_i, \theta) \\ &\doteq G_i(t) f(t; \theta, \mathbf{x}_i) dt, \end{aligned}$$

e

$$P(T_i \in (t, t + dt), \delta = 0; \mathbf{x}_i, \theta) \doteq g_i(t) S(t; \theta, \mathbf{x}_i) dt.$$

Assim, se a distribuição de T_i^C não envolve θ , a função de verossilhança sob os dados

(t_i, δ_i, x_i) , com $i = 1, \dots, n$ é

$$L(\theta) = \prod_{i=1}^n f(t_i; \theta, \mathbf{x}_i)^{\delta_i} S(t_i; \theta, \mathbf{x}_i)^{(1-\delta_i)}. \quad (1.13)$$

Embora este resultado tenha sido obtido com a suposição inicial de censura aleatória, no caso contínuo a função de verossimilhança dada em 1.13 é apropriada a uma ampla classe de tipos de censura, incluindo censura tipo I (que é um caso particular de censura aleatória) e censura tipo II, conforme foi mostrado em Kalbfleisch e Prentice (1980).

Uma outra representação bastante usada de 1.13 é dada por

$$L(\theta) = \prod_{i \in F} f(t_i; \theta, \mathbf{x}_i) \prod_{i \in C} S(t_i; \theta, \mathbf{x}_i). \quad (1.14)$$

1.3.2 Estimação pelo Método da Máxima Verossimilhança

Nas condições dadas acima, definimos o estimador de máxima verossimilhança (e.m.v.) para θ como o ponto $\hat{\theta}$, pertencente ao espaço paramétrico Θ , que maximiza $L(\theta)$ dado em 1.14. O processo de maximização é geralmente mais simples quando trabalhamos com o logaritmo natural de $L(\theta)$. Denotando a função de log-verossimilhança como $L^*(\theta) = \log L(\theta)$ e, considerando a representação 1.14, temos

$$L^*(\theta) = \sum_{i \in F} \log f(t_i; \theta, \mathbf{x}_i) + \sum_{i \in C} \log S(t_i; \theta, \mathbf{x}_i). \quad (1.15)$$

Quando $L^*(\theta)$ for contínua e diferenciável, o e.m.v. $\hat{\theta}$ pode ser obtido pela resolução do sistema de equações simultâneas

$$\frac{\partial L^*(\theta)}{\partial \theta_j} = 0, \quad j = 1, \dots, r. \quad (1.16)$$

Estimadores de máxima verossimilhança para os parâmetros dos modelos descritos na seção 1.2, assim como referência de trabalhos específicos desenvolvidos para cada modelo, são dados com detalhes em Lawless (1982).

Com muita frequência as equações 1.16 não podem ser resolvidas analiticamente, e a obtenção dos estimadores de máxima verossimilhança requer a utilização de métodos iterativos. Chambers (1977) descreve vários métodos que podem ser usados nesta situação. Um bastante aplicado é o *Método de Newton-Raphson*. No apêndice C apresentamos uma breve descrição deste método.

1.3.3 Teoria Assintótica

Consideramos aqui de forma resumida os principais resultados assintóticos da verossimilhança, que podem ser aplicados à verossimilhança gerada por modelos paramétricos de regressão, para dados com certos tipos de censura, como censura aleatória, tipo I e tipo II. A aplicabilidade da teoria assintótica a dados de sobrevivência é abordada principalmente em Kalbfleisch e Prentice (1980) e Lawless (1982) entre outros, e uma descrição detalhada da teoria clássica da verossimilhança é dada em Cox e Hinkley (1974).

Suponha t_1, \dots, t_n , observações de n variáveis aleatórias independentes (não necessariamente identicamente distribuídas), cuja distribuição envolve um vetor de parâmetros $\theta = (\theta_1, \dots, \theta_k)'$ tomando valores no espaço paramétrico Θ , e a função de verossimilhança é definida por

$$L(\theta) = \prod_{i=1}^n L_i(\theta),$$

onde $L_i(\theta)$ é a verossimilhança da observação i .

Vamos definir as principais estatísticas utilizadas para testar e criar intervalos de confiança aproximados para θ , quando possuímos grandes amostras.

Estatística de Escore de Rao

Como antes, considere $L^*(\theta) = \log L(\theta)$ e seja $U_i(\theta)$ definido como segue

$$U_i(\theta) = \frac{\partial L_i^*(\theta)}{\partial \theta} = \left[\frac{\partial L_i^*(\theta)}{\partial \theta_1}, \dots, \frac{\partial L_i^*(\theta)}{\partial \theta_k} \right]' \quad i = 1, \dots, n.$$

Se podemos intercambiar as operações de integração de t_i e derivação de θ , temos para $i = 1, \dots, n$ que a média e a variância de $\mathbf{U}_i(\theta)$ são dadas por

$$E(\mathbf{U}_i(\theta)) = 0 \quad e \quad V(\mathbf{U}_i(\theta)) = -E\left(\frac{\partial^2 L_i^*(\theta)}{\partial\theta\partial\theta'}\right) = \mathbf{I}_i(\theta).$$

Uma vez que $U_1(\theta), \dots, U_n(\theta)$, sendo funções da amostra, são independentes, temos que, sob certas condições de regularidade (Cox e Hinkley, 1974), o teorema central do limite de Lindeberg se aplica ao **vetor de escores**, dado por

$$\mathbf{U}(\theta) = \frac{\partial L^*(\theta)}{\partial\theta} = \sum_{i=1}^n \mathbf{U}_i(\theta)$$

e conseqüentemente, $\mathbf{U}(\theta)$ tem distribuição assintótica normal com média $\mathbf{0}$ e variância dada pela **matriz de informação de Fisher**,

$$\mathbf{I}(\theta) = \sum_{i=1}^n \mathbf{I}_i(\theta).$$

A descrição teórica sobre as condições para a validade da convergência normal para o caso geral, é dada, por exemplo, em Feller (1971).

Para testar a hipótese $H_0 : \theta = \theta_0$, temos a **estatística de escores de Rao**

$$\mathbf{U}(\theta)' \mathbf{I}^{-1}(\theta) \mathbf{U}(\theta), \tag{1.17}$$

que possui, sob H_0 , distribuição assintótica χ^2 com k graus de liberdade (a dimensão do vetor θ). Uma região de confiança aproximada para θ pode ser construída com o conjunto de valores θ_0 que produzem valores de 1.17 menores que um especificado quantil de uma χ_k^2 .

Estatística de Máxima Verossimilhança de Wald

Seja $\hat{\theta}$ o e.m.v. de θ . Se θ é um ponto interior de Θ , a função de verossimilhança $L(\theta)$ é três vezes diferenciável, e certas condições sobre a terceira derivada são satisfeitas (Cox

e Hinkley, 1974), pode-se mostrar que :

1. Assintoticamente $\hat{\theta}$ é a única solução de $\mathbf{U}(\theta) = \mathbf{0}$;
2. $\hat{\theta}$ é consistente para θ ;
3. $\hat{\theta}$ tem assintoticamente distribuição normal multivariada com média θ e variância $\mathbf{I}^{-1}(\theta)$.

Com base nestes resultados, podemos usar a **estatística de Wald**

$$(\hat{\theta} - \theta)' \mathbf{I}(\theta) (\hat{\theta} - \theta), \quad (1.18)$$

que tem distribuição assintótica χ^2 com k graus de liberdade para especificar regiões de confiança e testes sobre θ .

Uma observação importante é que podemos utilizar no lugar de $\mathbf{I}(\theta)$, seus estimadores consistentes $\mathbf{I}(\hat{\theta})$, $I(\theta)$ ou $I(\hat{\theta})$, sem afetar a distribuição assintótica das estatísticas 1.17 e 1.18. onde estes estimadores são definidos abaixo,

$$\mathbf{I}(\hat{\theta}) = -E \left(\frac{\partial^2 L^*(\theta)}{\partial \theta \partial \theta'} \Big|_{\theta=\hat{\theta}} \right) \quad (1.19)$$

$$I(\theta) = - \frac{\partial^2 L^*(\theta)}{\partial \theta \partial \theta'} \quad (1.20)$$

$$I(\hat{\theta}) = - \frac{\partial^2 L^*(\theta)}{\partial \theta \partial \theta'} \Big|_{\theta=\hat{\theta}}. \quad (1.21)$$

A matriz $I(\hat{\theta})$ 1.21 é chamada **matriz de informação observada** e é usada com muita frequência em aplicações.

A Estatística da Razão de Verossimilhança

A terceira classe de estatísticas de verossimilhança é baseada na **razão de verossimilhança** $R = L(\theta_0)/L(\hat{\theta})$. Pode-se mostrar que sob a hipótese $H_0 : \theta = \theta_0$, a estatística

$$A = -2 \log \left\{ \frac{L(\theta_0)}{L(\hat{\theta})} \right\}, \quad (1.22)$$

tem distribuição assintótica χ^2 com k graus de liberdade. Com base nesta estatística podemos então testar hipóteses e obter regiões de confiança de maneira similar ao que foi descrito anteriormente. Um método para computar intervalo de confiança com base no contorno da razão de verossimilhança é dado em Vezon e Moolgavkar (1988).

Podemos também usar uma modificação de 1.22 para fazer inferências sobre uma partição do vetor de parâmetros θ . Seja $\theta' = (\theta'_1, \theta'_2)$ e $\hat{\theta}_2(\theta'_1)$ o e.m.v. de θ_2 dado que $\theta_1 = \theta'_1$. Então,

$$-2 \log \left\{ \frac{L(\theta'_1, \hat{\theta}_2(\theta'_1))}{L(\hat{\theta})} \right\}, \quad (1.23)$$

tem distribuição χ^2 com graus de liberdade determinado pela dimensão de θ_1 , se a hipótese $H_0 : \theta_1 = \theta'_1$ é verdadeira. Esta estatística pode então ser utilizada para testar H_0 e para criar intervalos (regiões) de confiança aproximados para θ_1 .

Teoria Assintótica e Modelos de Tempo de Vida

Quando tratamos de modelos de tempo de vida com dados censurados, encontramos muitas dificuldades na obtenção de testes e intervalos de confiança exatos. Consequentemente os procedimentos baseados na teoria para grandes amostras são usados com muita frequência. Algumas restrições com respeito ao tempo de censura e às variáveis regressoras são necessárias para a validade dos resultados assintóticos. Por exemplo, os tempos potenciais de censura não devem convergir muito rápido para zero, e não devem existir valores extremos das variáveis regressoras que tenham efeito dominante na estimativa de β (Farewell e Prentice, 1977). Os métodos assintóticos mais utilizados são os baseados na normalidade assintótica de $\hat{\theta}$ (NA), por serem considerados mais simples.

Os testes e intervalos de confiança aproximados baseados nos métodos abordados aqui são considerados assintoticamente equivalentes. Contudo, intervalos baseados na NA não

possuem a propriedade da *invariância sob transformações de parâmetros* (Kalbfleisch e Prentice, 1980) que é válida para todos os procedimentos baseados na razão de verossimilhança (RV) e para os procedimentos dados pela estatística de escores. No entanto esta propriedade deixa de valer para a última estatística, se em 1.17 trocamos a matriz de informação de Fisher $I(\theta)$ por um de seus estimadores consistentes dados em 1.19, 1.20, 1.21. Isto representa uma dificuldade no uso da estatística de escores, tendo em vista que com amostras censuradas em geral é necessário a realização desta troca. Por outro lado, temos que o teste geral da razão de verossimilhança representa uma extensão do teste da razão de verossimilhança, que é ótimo para hipóteses simples, e que com o crescimento do tamanho da amostra, esse teste geral torna-se também ótimo.

A teoria assintótica para a obtenção de intervalos de confiança aproximados é baseada na forma quadrática e simétrica do log da função de verossimilhança. Os intervalos (regiões) de confiança sob a NA possuem comprimentos simétricos em torno de $\hat{\theta}$. Contudo a forma não quadrática e assimétrica da log verossimilhança em amostras pequenas e censuradas indica a imprecisão destes intervalos. Diversos autores, estudando modelos de tempo de vida, mostraram que intervalos baseados na NA dos estimadores de máxima verossimilhança não possuem boas propriedades (Thoman, Bain e Antle, 1969 ; Billmann e Antle, 1972, Ostrouchov e Meeker, 1988 ; Lawless, 1982 ; Doganaksoy e Schmee, 1993). De forma diferente, intervalos baseados na RV, usam a distribuição assintótica χ^2 para calcular os limites de uma função de log verossimilhança assimétrica **observada**. Doganaksoy e Schmee (1993) concluíram em um trabalho recente, através de extensivos resultados de Monte Carlo que intervalos baseados na RV (pura ou com correção) são muito melhores que os baseados na NA, para os parâmetros da Weibull (valor extremo) e log-normal (normal), em amostras censuradas, e recomendam que com dados de tempo de vida deveria-se utilizar intervalos baseados na NA, apenas quando o método da RV não poder ser obtido.

1.4 Aplicação - Testes Acelerados

Uma aplicação importante de regressão com dados de tempo de vida é a análise da durabilidade de itens de manufatura através de **testes acelerados de vida**. Um teste acelerado consiste basicamente de uma variedade de métodos para reduzir a vida de produtos ou provocar a degradação da performance destes mais rapidamente. O principal objetivo é obter rapidamente os dados a serem analisados e desta forma produzir informações sobre a vida do produto ou sua performance sob condições normais de uso. Estes testes economizam tempo e dinheiro.

A aceleração do término da vida dos produtos é obtida pela execução destes sob condições de **estresse**. Testar sob um super estresse consiste em fazer o produto trabalhar a níveis de estresses mais altos que o normal. Geralmente os estresses acelerados são temperatura, voltagem, carga mecânica, vibração, aumento da velocidade de execução, etc.. Um abrangente estudo deste ramo da confiabilidade, assim como numerosas referências são dadas em Nelson (1990).

Para realizar a análise estatística destes dados, assumimos um modelo que consiste na combinação da distribuição do tempo de vida e no relacionamento entre tempo e estresse. Existem estudos na literatura da engenharia sobre o relacionamento do tempo de vida com o estresse aplicado em determinados tipos de produtos. Considere t_1, \dots, t_n os tempos de vida observados de itens em estudo, sujeitos aos estresses v_1, \dots, v_s . O modelo de regressão log-linear comumente usado é da forma

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_s x_{is} + \sigma W_i \quad i = 1, \dots, n,$$

onde $y_i = \log t_i$, β_0, \dots, β_s e σ são parâmetros desconhecidos, W_i possui uma determinada distribuição e x_i é uma transformação do estresse v_i . Esta transformação é, em geral, determinada pela relação física entre tempo de vida do item estudado e os tipos de estresses aplicados. Duas importantes relações deste tipo são dadas abaixo.

1. Relação Arrhenius - Proveniente de uma lei sobre a dependência da reação química

- com a temperatura, sugere a transformação $x = 1/v$, onde v representa o estresse, geralmente sendo a temperatura absoluta.
2. Relação Potência inversa - Sugere uma transformação da forma $x = -\log v$, e aplica-se a vários tipos de estresse.

Exemplos numéricos de análise de dados resultantes de testes acelerados são fornecidos no capítulo 4.

1.5 Discussão Sobre a Adequacidade de um Modelo

Quando optamos por um procedimento paramétrico para a análise de dados de vida, a escolha de um modelo é sem dúvida uma etapa importante na análise. Existem muitos métodos informais para acessar a adequacidade do ajuste de um modelo, ou de selecionar um dentre vários modelos plausíveis. Consideramos a seguir uma breve discussão sobre a checagem de um modelo de regressão ajustado e sobre métodos para a escolha de um modelo.

1.5.1 Análise de Resíduos

O exame dos resíduos de um modelo de regressão ajustado é uma forma importante de checar as suposições deste modelo, e pode também revelar características especiais dos dados. Considerando um modelo de regressão log-linear $Y_i = \mathbf{x}_i\beta + \sigma W_i$, os resíduos estimados podem ser definidos como

$$w_i = (y_i - \mathbf{x}_i\hat{\beta})/\hat{\sigma} \quad i = 1, \dots, n. \quad (1.24)$$

Quando tratamos de dados completos, encontramos muitas referências que tratam da análise de resíduos para o modelo de regressão normal (veja por exemplo Draper e Smith, 1981). Para outros modelos de regressão, podemos construir gráficos de probabilidade dos w_i para checar a distribuição de W_i ou fazer um gráfico de resíduos versus alguma

variável regressora que pode revelar inadequidade do modelo. Com dados censurados, uma possível abordagem é tratar os resíduos observados w_i como uma amostra censurada da distribuição assumida para W_i (Lawless, 1982). Podemos então calcular as estimativas de $P(W > w_i)$ usando o estimador produto-limite (Kaplan-Meier), como dado em 1.1, para estimar a função de sobrevivência de W_i . Uma análise gráfica seria desenhar a curva de sobrevivência básica assumida para W_i , denotada S_1 , juntamente com o gráfico de Kaplan-Meier dos resíduos estimados (\hat{S}_{km}). Quando a distribuição de W_i depender de algum parâmetro desconhecido, podemos usar a estimativa de máxima verossimilhança deste parâmetro para obter uma estimativa paramétrica da função de sobrevivência (\hat{S}_1). Se a distribuição assumida é adequada para ajustar os dados (log dos tempos), esperamos que as curvas \hat{S}_1 e \hat{S}_{km} sejam semelhantes.

1.5.2 A Escolha de um Modelo

Existem diversos critérios formais e informais adotados para selecionar um modelo para ajustar dados de vida. Em algumas situações o uso de um modelo específico é sugerido por estudos anteriores com experimentos semelhantes, ou pelo conhecimento do mecanismo de falha. Para checar a adequação de um ajuste, é bastante comum a utilização de métodos gráficos. Por exemplo, para checar se o modelo Weibull é adequado ao ajuste de uma amostra de T, desenha-se o gráfico do $\log(-\log \hat{S}(t))$ versus $\log(t)$, onde $\hat{S}(t)$ é o estimador produto limite 1.1 da função de sobrevivência dos dados. Assim, consideramos o modelo Weibull adequado se o gráfico resultar uma linha reta. Isto ocorre porque implicitamente assumimos que $\hat{S}(t)$ é aproximadamente igual a $\exp(-(t_i/\alpha)^\beta)$, a função de sobrevivência de uma Weibull.

Em outras situações a utilização de um modelo é devida à sua simplicidade computacional. De acordo com Nelson (1990), com muita frequência o modelo exponencial é usado indevidamente em testes de vida, basicamente pelo fato de apresentar resultados inferenciais simples e bem conhecidos.

Existem vários testes formais de bondade de ajuste, que visam testar a adequidade

de um determinado modelo ajustado, através de hipóteses estatísticas. Os procedimentos mais bem conhecidos são os baseados na função de distribuição empírica .

Temos também alguns estudos que tratam da *escolha* de um modelo, como o artigo de Dyer (1973), que fornece procedimentos de discriminação entre 2 modelos de famílias separadas, e o trabalho de Dumonceaux e Antle (1973) que simula a distribuição da razão de verossimilhança, visando discriminar entre os modelos Weibull e log-normal. Contudo estes métodos não tratam de censuras, nem de variáveis regressoras.

Existem poucos trabalhos desenvolvidos sobre testes de bondade de ajuste para modelos de regressão, com exceção do caso da distribuição normal. Para tratar da escolha de um entre um pequeno número de modelos de regressão que melhor se adequem aos dados, podemos trabalhar com uma família de modelos que inclua como casos particulares as distribuições de interesse. Prentice(1974) utilizou a distribuição F-generalizada com este propósito. Sob este enfoque, abordamos neste trabalho a utilização do modelo **gama generalizada**, que possui como casos particulares os modelos apresentados na seção 1.2 (exponencial, Weibull, gama e log-normal), com a finalidade de discriminar o modelo mais indicado para o ajuste de um conjunto de dados. A utilização deste modelo na escolha de uma distribuição foi sugerida e estudada por vários autores. Entre eles temos, Hager e Bain (1970), Prentice (1974), Farewell e Prentice (1977), Lawless (1982), e Nelson (1990).

Capítulo 2

A Distribuição Gama Generalizada

2.1 Introdução

A distribuição gama generalizada (d.g.g.) foi introduzida por Stacy (1962) e despertou o interesse de diversos autores pelo fato de representar uma família paramétrica que possui como casos particulares outras distribuições amplamente utilizadas na análise de dados de tempo de vida. Parr e Webster (1965) estudaram o uso deste modelo no teste de adequação dos modelos Weibull e do modelo exponencial. Outros trabalhos importantes foram desenvolvidos com o objetivo de reduzir as dificuldades operacionais na análise estatística deste modelo. Prentice (1974), em um trabalho de grande relevância neste assunto, apresenta uma forma reparametrizada e estendida do log da d.g.g., que simplifica a solução de problemas inferenciais deste modelo.

Neste capítulo fazemos um estudo das propriedades da d.g.g., descrevemos o modelo reparametrizado de Prentice (1974) e o modelo de regressão baseado na distribuição log-gama.

2.2 Propriedades da Distribuição Gama Generalizada

Apresentamos nesta seção, com algum detalhamento, resultados e demonstrações sobre propriedades da d.g.g., incluindo formas de representação da sua distribuição e a dis-

Distribuição	Densidade
Exponencial	$f(t; \alpha, 1, 1)$
Gama	$f(t; \alpha, 1, k)$
Weibull	$f(t; \alpha, \beta, 1)$
Qui-quadrado (n g.l.)	$f(t; 2, n/2, 1)$
Qui (n g.l.)	$f(t; \sqrt{2}, n/2, 2)$
Half - Normal	$f(t; \sqrt{2}, 1/2, 2)$
Normal Circular	$f(t; \sqrt{2}, 1, 2)$
Normal Esférica	$f(t; \sqrt{2}, 3/2, 2)$
Rayleigh ($c > 0$)	$f(t; c\sqrt{2}, 1, 2)$

Tabela 2.1: Casos particulares da d.g.g., com densidade $f(t; \alpha, \beta, k)$ - Stacy e Miram (1965).

tribuição de funções de variáveis aleatórias independentes com d.g.g..

2.2.1 Função de Densidade

Seja T uma variável aleatória com distribuição gama generalizada. A sua função de densidade (f.d.p.), denotada por $f(t; \alpha, \beta, k)$, é dada por

$$f(t; \alpha, \beta, k) = \frac{\beta}{\alpha \Gamma(k)} (t/\alpha)^{\beta k - 1} \exp(-(t/\alpha)^\beta) \quad t > 0, \quad \alpha, \beta, k > 0. \quad (2.1)$$

Stacy e Mihram (1965) apresentaram uma reparametrização incluindo os casos onde o parâmetro β pode ser considerado negativo. Harter (1967) generalizou a densidade pela inclusão de um parâmetro de locação. A expressão 2.1, que é uma modificação natural da representação dos parâmetros da densidade dada originalmente por Stacy (1962), será utilizada no decorrer deste trabalho.

A densidade 2.1 pode ser usada para especificar uma variedade de distribuições de probabilidade. Alguns exemplos são apresentados na Tabela 2.1 (Stacy e Mirham, 1965).

Temos também que a distribuição log-normal surge como limite de 2.1 quando $k \rightarrow \infty$, conforme mostra o teorema a seguir.

Teorema 1 . *Seja T uma variável aleatória com distribuição gama generalizada, cuja densidade é dada em 2.1. Então, a distribuição de T aproxima-se de uma log-normal, quando $k \rightarrow \infty$.*

Demonstração :

Se T tem distribuição definida por 2.1, então $T = \alpha X^{\frac{1}{\beta}}$, onde X tem distribuição gama com um parâmetro k , dada em 1.6. Fazendo $Y = \log(T)$ temos,

$$Y = \log \alpha + \frac{1}{\beta} W,$$

onde $W = \log(X)$ tem distribuição log-gama com parâmetro k , cuja média e variância são dadas pelas funções digama ($\Psi(k)$) e trigama ($\Psi'(k)$) respectivamente (ver Apêndice B). Considerando as fórmulas assintóticas de $\Psi(k)$ e $\Psi'(k)$, dadas em A.2 e A.3 (Apêndice A), definimos a seguinte variável padronizada

$$Z = \sqrt{k}(W - \log(k)),$$

com densidade

$$f(z) = \frac{k^{k-1/2}}{\Gamma(k)} \exp\{\sqrt{k}z - ke^{z/\sqrt{k}}\} \quad z \in \mathfrak{R}.$$

Usando a expansão da função exponencial em série de Taylor,

$$\exp(z/\sqrt{k}) = \sum_{n=0}^{\infty} \frac{(z/\sqrt{k})^n}{n!},$$

temos como resultado que

$$\begin{aligned} \exp[\sqrt{k}z - ke^{z/\sqrt{k}}] &= \exp\left[zk^{1/2} - k \sum_{n=0}^{\infty} \frac{(z/\sqrt{k})^n}{n!}\right] \\ &= e^{-k} e^{-z^2/2} \exp\left[-\left(\frac{z^3}{3!k^{1/2}} + \frac{z^4}{4!k} + \frac{z^5}{5!k^{3/2}} + \dots\right)\right]. \end{aligned}$$

Pela fórmula de Stirling (Abramowitz e Stegun, 1965) temos,

$$\lim_{k \rightarrow \infty} \frac{k^{k-1/2} e^{-k}}{\Gamma(k)} = \frac{1}{\sqrt{2\pi}}.$$

Então,

$$\lim_{k \rightarrow \infty} f(z; k) = \lim_{k \rightarrow \infty} \left(\frac{k^{k-1/2} e^{-k}}{\Gamma(k)} \right) e^{-z^2/2} \exp \left[- \left(\frac{z^3}{3!k^{1/2}} + \dots \right) \right] = \frac{\exp(-z^2/2)}{\sqrt{2\pi}},$$

que representa a densidade de uma variável aleatória com distribuição normal padrão. isto é, pelo teorema de Sheffé (Barry James, 1981) Z converge em distribuição para uma $N(0, 1)$. Escrevendo Z em função de Y temos,

$$Z = \frac{Y - (\log \alpha + \beta^{-1} \log k)}{1/(\beta\sqrt{k})}.$$

Desta forma, com $k \rightarrow \infty$, $Y = \log T$ tem distribuição aproximadamente normal com média $(\log \alpha + \beta^{-1} \log k)$ e variância $1/(\beta^2 k)$ e, conseqüentemente, T tem distribuição aproximadamente log-normal. \square

2.2.2 Análise Gráfica da Densidade

A densidade 2.1 pode tomar uma variedade de formas, de acordo com o valor dos parâmetros. Na análise da forma da densidade, podemos, sem perda de generalidade, associar ao parâmetro α um valor constante, tendo em vista que este é um parâmetro de escala. A Figura 2-1 apresenta alguns exemplos da forma da densidade de T , para $\alpha = 2$. Percebe-se portanto, uma grande flexibilidade desta densidade e nota-se mudanças significativas da sua forma, acompanhando modificações do produto βk . Para valores de $\beta k < 1$ a função de densidade é monótona decrescente a partir do ponto $t = 0$, sendo $\lim_{t \rightarrow 0} f(t; \alpha, \beta, k) = \infty$. Quando $\beta k = 1$, $f(t; \alpha, \beta, k)$ é monótona decrescente e finita no ponto $t = 0$. Para valores de $\beta k > 1$, a função assume o valor zero quando $t = 0$ e

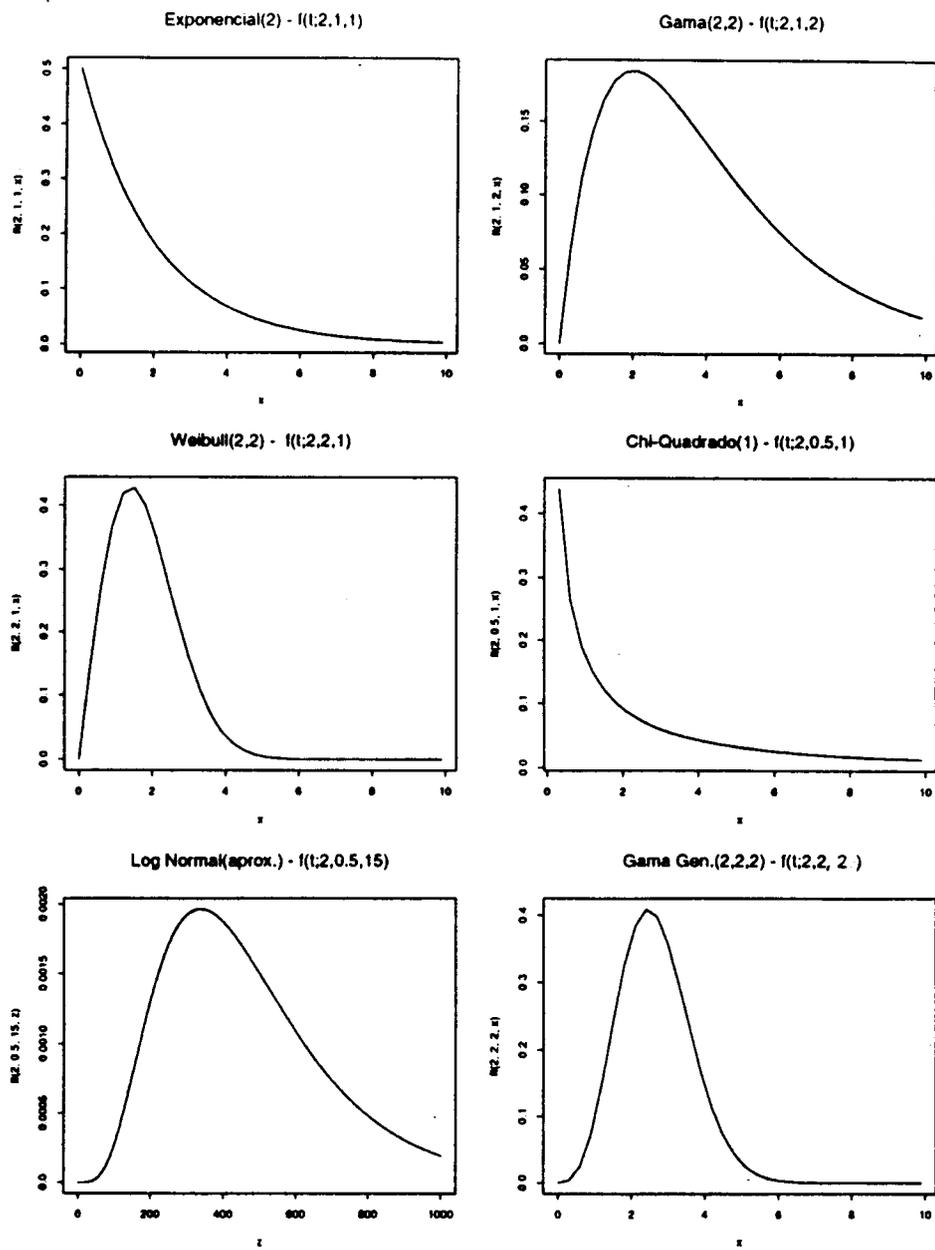


Figura 2-1: Algumas formas assumidas pela densidade gama generalizada $f(t; \alpha, \beta, k)$.

decrece para valores grandes de t .

Considere o logaritmo da função 2.1, dado por

$$\log f(t; \alpha, \beta, k) = \log c + (\beta k - 1) \log t - \left(\frac{t}{\alpha}\right)^\beta$$

onde $c = \frac{\beta}{\alpha^{\beta k} \Gamma(k)}$. Derivando esta expressão e igualando-a a zero, temos

$$\frac{\partial \log f(t; \alpha, \beta, k)}{\partial t} = \frac{\beta k - 1}{t} - \frac{\beta t^{\beta-1}}{\alpha^\beta} = 0$$

$$\Rightarrow t^* = \left(\frac{(\beta k - 1)\alpha^\beta}{\beta}\right)^{\frac{1}{\beta}} = \alpha(k - \beta^{-1})^{\frac{1}{\beta}}.$$

Como $f(t; \alpha, \beta, k)$ é definida apenas para $t > 0$,

$$t^* > 0 \Rightarrow (\beta k - 1) > 0 \Rightarrow \beta k > 1.$$

Desde que esta função é positiva e contínua em $t > 0$, os resultados acima implicam que ela tem um único ponto de máximo se $\beta k > 1$, e este ponto é dado por t^* .

2.2.3 Função de distribuição

Denotando por $F(t; \alpha, \beta, k)$ a função de distribuição acumulada (f.d.) de T temos, para $t > 0$

$$F(t; \alpha, \beta, k) = I(k, (t/\alpha)^\beta) \tag{2.2}$$

onde $I(k, x) = \frac{1}{\Gamma(k)} \int_0^x w^{k-1} e^{-w} dw$, é a **função gama incompleta** (ver Apêndice A. A-4).

Demonstração:

Por definição temos

$$F(t; \alpha, \beta, k) = \frac{1}{\Gamma(k)} \int_0^t \frac{\beta}{\alpha} (x/\alpha)^{\beta k - 1} \exp(-(x/\alpha)^\beta) dx. \quad (2.3)$$

Fazendo $w = (x/\alpha)^\beta$ a expressão dentro da integral se reduz da seguinte forma ;

$$\frac{\beta}{\alpha} (x/\alpha)^{\beta k - 1} \exp(-(x/\alpha)^\beta) dx = \frac{\beta x^{\beta k - 1} \alpha^\beta e^{-w}}{\alpha^{\beta k} \beta x^{\beta - 1}} dw = w^{k-1} e^{-w} dw.$$

Então

$$F(t; \alpha, \beta, k) = \frac{1}{\Gamma(k)} \int_0^{(t/\alpha)^\beta} w^{k-1} e^{-w} dw = I(k, (t/\alpha)^\beta). \quad \square$$

Podemos também escrever 2.3 como

$$\begin{aligned} F(t; \alpha, \beta, k) &= \frac{\beta}{\alpha^{\beta k} \Gamma(k)} \int_0^t x^{\beta k - 1} \sum_{j=0}^{\infty} \frac{(-1)^j (x/\alpha)^{\beta j}}{j!} dx \\ \Rightarrow F(t; \alpha, \beta, k) &= \frac{\beta}{\alpha^{\beta k} \Gamma(k)} \sum_{j=0}^{\infty} \left[\frac{(-1)^j (\alpha^{-\beta})^j}{j!} \int_0^t x^{\beta k + \beta j - 1} dx \right]. \end{aligned} \quad (2.4)$$

2.2.4 Função de Sobrevivência e Função Risco

Considerando a expressão 2.2 obtida no ítem anterior, a função de sobrevivência de T, denotada $S(t) = P(T \geq t)$ é dada por

$$S(t) = 1 - I(k, (t/\alpha)^\beta). \quad (2.5)$$

A função risco de T, definida por $h(t) = f(t)/S(t)$, é obtida pelas expressões 2.1 e 2.2 e dada por

$$h(t) = \frac{t^{\beta k - 1} e^{-(t/\alpha)^\beta}}{\int_t^\infty x^{\beta k - 1} e^{-(x/\alpha)^\beta} dx}. \quad (2.6)$$

Esta função, como a desidade da d.g.g. toma diversas formas. A Figura 2-2 apresenta

algumas das possíveis representações desta função, em correspondência com as densidades dadas na Figura 2-1.

2.2.5 Função Geratriz de Momentos

Denotando por $M(\theta)$ a função geratriz de momentos (f.g.m.) de T , verificamos que

$$M(\theta) = \sum_{r=0}^{\infty} \frac{(\theta\alpha)^r}{r!} \frac{\Gamma(k + r/\beta)}{\Gamma(k)}. \quad (2.7)$$

Demonstração :

Por definição temos

$$M(\theta) = E[e^{\theta T}] = \int_0^{\infty} \frac{e^{\theta t} \beta t^{\beta k - 1}}{\alpha^{\beta k} \Gamma(k)} \exp(-(\frac{t}{\alpha})^{\beta}) dt.$$

Usando a mesma transformação empregada para demonstrar 2.2, com $w = (t/\alpha)^{\beta}$, temos

$$\begin{aligned} M(\theta) &= \frac{1}{\Gamma(k)} \int_0^{\infty} w^{k-1} e^{-w} e^{\theta \alpha w^{1/\beta}} dw \\ &= \frac{1}{\Gamma(k)} \int_0^{\infty} w^{k-1} e^{-w} \sum_{r=0}^{\infty} \frac{(\theta \alpha w^{1/\beta})^r}{r!} dw \\ &= \frac{1}{\Gamma(k)} \sum_{r=0}^{\infty} \frac{(\theta \alpha)^r}{r!} \int_0^{\infty} (w^{1/\beta})^r w^{k-1} e^{-w} dw \\ &= \frac{1}{\Gamma(k)} \sum_{r=0}^{\infty} \frac{(\theta \alpha)^r}{r!} \Gamma(k + r/\beta). \quad \square \end{aligned}$$

2.2.6 Momentos

O momento de ordem n da d.g.g., para $n = 1, 2, \dots$ é dado por

$$E(T^n) = \frac{\alpha^n \Gamma(k + n/\beta)}{\Gamma(k)}. \quad (2.8)$$

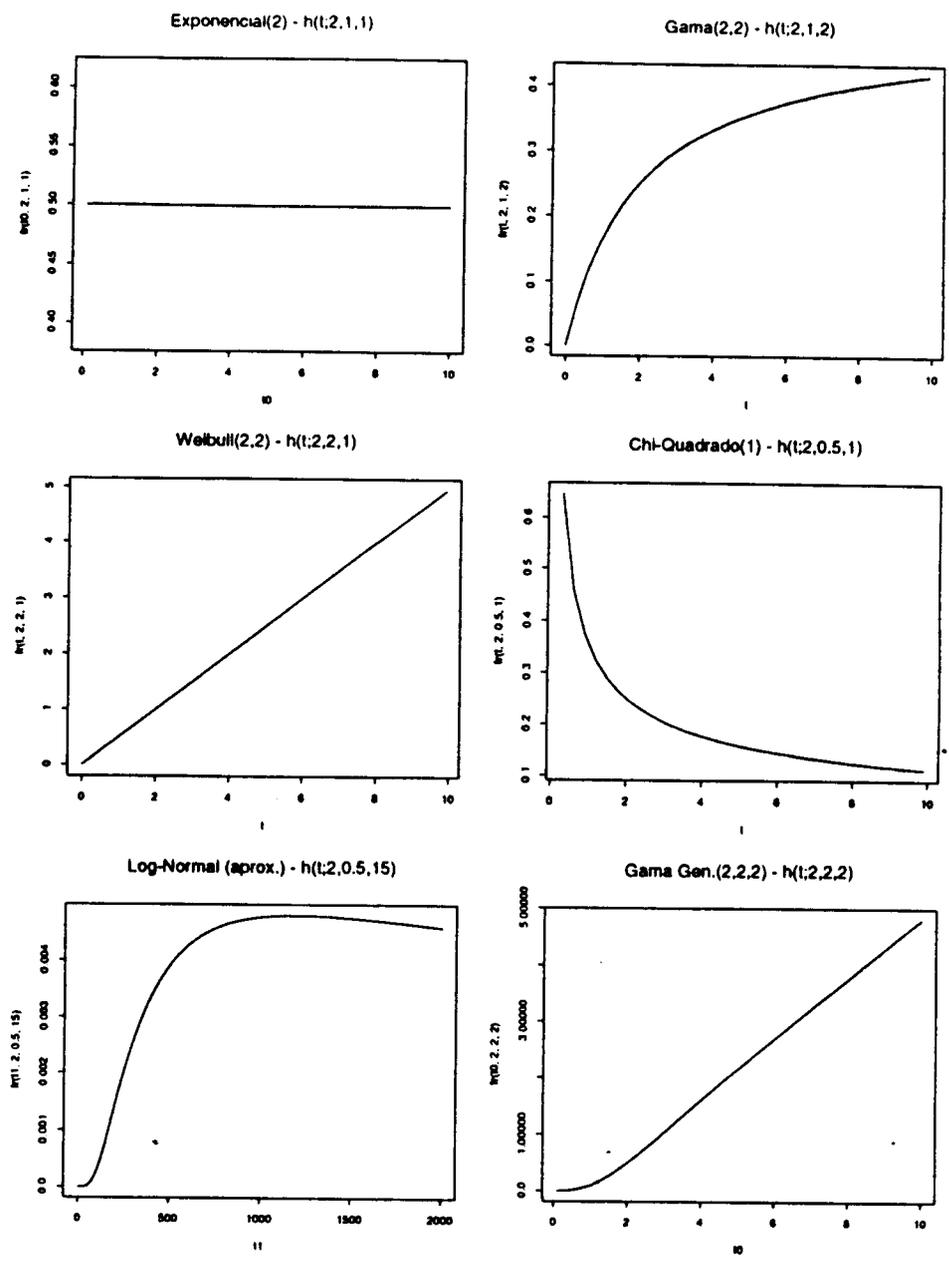


Figura 2-2: Algumas formas assumidas pela função de risco da gama generalizada, $h(t; \alpha, \beta, k)$.

Demonstração :

Obtemos este resultado pelo desenvolvimento da relação $E(T^n) = \frac{\partial^n M(\theta)}{\partial \theta^n} |_{\theta=0}$, e usando o resultado 2.7, conforme segue

$$\begin{aligned} \frac{\partial^n M(\theta)}{\partial \theta^n} &= \frac{\partial^n}{\partial \theta^n} \sum_{r=0}^{\infty} \frac{(\theta \alpha)^r}{r!} \frac{\Gamma(k+r/\beta)}{\Gamma(k)} \\ &= \sum_{r=0}^{\infty} \frac{\alpha^r}{r!} \frac{\Gamma(k+r/\beta)}{\Gamma(k)} \frac{\partial^n \theta^r}{\partial \theta^n} \\ &= \sum_{r=n}^{\infty} \frac{\alpha^r}{r!} \frac{\Gamma(k+r/\beta)}{\Gamma(k)} \frac{r! \theta^{r-n}}{(r-n)!}, \end{aligned}$$

então

$$E(T^n) = \frac{\partial^n M(\theta)}{\partial \theta^n} |_{\theta=0} = \frac{\alpha^n \Gamma(k+n/\beta)}{\Gamma(k)}. \square$$

Como consequência de 2.8, determinamos a esperança e a variância de T conforme segue.

$$E(T) = \frac{\alpha \Gamma((\beta k + 1)/\beta)}{\Gamma(k)}$$

$$V(T) = \frac{\alpha^2}{\Gamma(k)} \left\{ \Gamma((\beta k + 2)/\beta) - \frac{\Gamma^2((\beta k + 1)/\beta)}{\Gamma(k)} \right\}.$$

2.2.7 Distribuição de Algumas Funções de Variáveis Gama Generalizada

Consideraremos aqui dois resultados sobre funções de variáveis gama generalizada, apresentados em forma de teorema.

Teorema 2 . *Sejam X e Y variáveis aleatórias independentes com distribuição gama generalizada, com densidades dadas respectivamente por $f(x; \alpha_1, \beta, k_1)$ e $f(y; \alpha_2, \beta, k_2)$*

da forma 2.1. Considere a variável $V = X/Y$. Temos então que ¹

$$W = V^\beta / [V^\beta + (\alpha_1/\alpha_2)^\beta]$$

tem distribuição Beta padrão (ver Apêndice A), com parâmetros k_1 e k_2 .

Demonstração :

Para demonstrar este teorema, precisamos inicialmente de alguns resultados dados nos lemas abaixo.

Lema 1. Seja α e β constantes positivas. Se X_1 tem distribuição gama com um parâmetro k , então

$$X = \alpha X_1^{1/\beta}$$

tem distribuição gama generalizada com parâmetros α , β e k com densidade dada em 2.1.

Prova

Sejam F_X , f_X , F_{X_1} , f_{X_1} as respectivas funções de distribuição e de densidade de X e X_1 . Temos como resultado que

$$F_X(x) = P(\alpha X_1^{1/\beta} \leq x) = F_{X_1}((x/\alpha)^\beta).$$

Então, para $x > 0$, a densidade de X é dada por

$$\begin{aligned} f_X(x) &= f_{X_1}((x/\alpha)^\beta) \frac{\beta}{\alpha^\beta} x^{\beta-1} \\ &= \frac{\beta}{\alpha^\beta \Gamma(k)} \left(\frac{x}{\alpha}\right)^{\beta k-1} \exp\left(-\left(\frac{x}{\alpha}\right)^\beta\right), \end{aligned}$$

¹ Este resultado foi dado em Stacy (1962) sem demonstração.

que representa a densidade 2.1 .□

Lema 2. Se Z_1 e Z_2 são variáveis independentes, com distribuição gama com parâmetros (α, k_1) e (α, k_2) respectivamente, então Z_1/Z_2 tem distribuição Beta 1 com parâmetros (k_1, k_2) (ver Apêndice A), ou seja,

$$Z_1/Z_2 \sim B_1(k_1, k_2).$$

Prova

Seja $Z = Z_1/Z_2$ e f_z, f_{z_1}, f_{z_2} as respectivas densidades de Z, Z_1 e Z_2 . Então, conforme um resultado clássico da probabilidade, dado que Z_1 e Z_2 são positivas, temos para $z > 0$,

$$\begin{aligned} f_z(z) &= \int_0^\infty x f_{z_1}(zx) f_{z_2}(x) dx \\ &= \int_0^\infty \frac{x}{\alpha \Gamma(k_1)} \left(\frac{zx}{\alpha}\right)^{k_1-1} e^{-(zx/\alpha)} \frac{1}{\alpha \Gamma(k_2)} \left(\frac{x}{\alpha}\right)^{k_2-1} e^{-(x/\alpha)} dx \\ &= \frac{z^{k_1-1}}{\Gamma(k_1)\Gamma(k_2)\alpha^{k_1+k_2}} \int_0^\infty x^{k_1+k_2-1} e^{-x(1+z)/\alpha} dx . \end{aligned}$$

Fazendo $u = x(1+z)/\alpha$, temos

$$\begin{aligned} f_z(z) &= \frac{z^{k_1-1}}{\Gamma(k_1)\Gamma(k_2)\alpha^{k_1+k_2}} \int_0^\infty \left(\frac{\alpha}{z+1}\right)^{k_1+k_2-1} \left(\frac{\alpha}{z+1}\right) u^{k_1+k_2-1} e^{-u} du \\ &= \frac{z^{k_1-1}}{(z+1)^{k_1+k_2}} \frac{\Gamma(k_1+k_2)}{\Gamma(k_1)\Gamma(k_2)} \\ &= \frac{z^{k_1-1}}{(z+1)^{k_1+k_2}} \frac{1}{B(k_1, k_2)} , \end{aligned}$$

que representa a densidade de uma $B_1(k_1, k_2)$. \square

Lema 3. Se Z tem distribuição $B_1(k_1, k_2)$ então $Z/(1 + Z)$ tem distribuição Beta padrão com parâmetros k_1 e k_2 (ver Apêndice A), ou seja,

$$Z/(1 + Z) \sim B_p(k_1, k_2).$$

Prova

Seja $W = Z/(1 + Z)$ e F_w, f_w, F_z, f_z as respectivas f.d. e f.d.p de W e Z .

Temos

$$F_w(w) = F_z(w/(1 - w)).$$

Então

$$f_w(w) = f_z(w/(1 - w)) \frac{1}{(1 - w)^2}.$$

Para $0 < w/(1 - w) < \infty$, ou seja $0 < w < 1$, temos que

$$\begin{aligned} f_w(w) &= \frac{(w/(1 - w))^{k_1 - 1}}{(1 - w)^2 B(k_1, k_2) (1 + \frac{w}{1 - w})^{k_1 + k_2}} \\ &= \frac{w^{k_1 - 1} (1 - w)^{k_1 + k_2}}{B(k_1, k_2) (1 - w)^2 (1 - w)^{k_1 - 1}} \\ &= \frac{w^{k_1 - 1} (1 - w)^{k_2 - 1}}{B(k_1, k_2)}, \end{aligned}$$

que representa a densidade de uma $B_p(k_1, k_2)$. \square

Com os resultados dos lemas acima, podemos agora demonstrar facilmente o Teorema 2. Sejam X_1 e X_2 variáveis aleatórias independentes com distribuição gama de um parâmetro, onde X_1 tem parâmetro k_1 e X_2 tem parâmetro k_2 e sejam X e Y variáveis com d.g.g., cujas densidades estão apresentadas no teorema, isto é, $X = \alpha_1 X_1^{1/\beta}$ e $Y = \alpha_2 X_2^{1/\beta}$, para $\alpha_1, \alpha_2 > 0$. Seja $V = X/Y$. Então pelo Lema 1,

$$V = \frac{\alpha_1}{\alpha_2} \left(\frac{X_1}{X_2} \right)^{1/\beta} \quad \text{em distribuição.}$$

Pelo Lema 2 a razão X_1/X_2 tem distribuição $B_1(k_1, k_2)$. Fazendo

$$W = \frac{V^\beta}{V^\beta + \left(\frac{\alpha_1}{\alpha_2} \right)^\beta},$$

temos

$$W = \frac{\left(\frac{\alpha_1}{\alpha_2} \right)^\beta \frac{X_1}{X_2}}{\left(\frac{\alpha_1}{\alpha_2} \right)^\beta \frac{X_1}{X_2} + \left(\frac{\alpha_1}{\alpha_2} \right)^\beta} = \frac{(X_1/X_2)}{(X_1/X_2) + 1},$$

que pelo Lema 3 tem distribuição $B_p(k_1, k_2)$. \square

Teorema 3 . (Distribuição da soma de variáveis independentes). *Seja $Y = \sum_{i=1}^n T_i$ onde T_1, \dots, T_n são variáveis aleatórias independentes e cada T_i tem densidade $f(t; \alpha_i, \beta_i, k_i)$, conforme dado em 2.1. Denotemos a função de distribuição acumulada de Y por $G_n(y)$. Então,*

$$G_n(y) = \gamma_n y^{\beta k} \sum_{j=0}^{\infty} (-1)^j A_j \quad (2.9)$$

onde

$$\beta k = \sum_{i=1}^n \beta_i k_i, \quad \gamma_n = \prod_{i=1}^n \left[\frac{\beta_i}{\alpha_i^{\beta_i k_i} \Gamma(k_i)} \right]$$

$$\epsilon \quad A_j = \sum_{f_1 + \dots + f_n = j} \frac{y^{\sum_{i=1}^n \beta_i k_i}}{\Gamma(\beta k + \sum_{i=1}^n \beta_i f_i + 1)} \prod_{i=1}^n \frac{\Gamma(\beta_i k_i + \beta_i f_i)}{f_i! \alpha_i^{\beta_i f_i}}.$$

Este resultado foi apresentado, sem prova, por Stacy (1962), com uma parametrização equivalente. Para demonstrá-lo, precisamos utilizar o resultado da proposição abaixo, que foi desenvolvida e demonstrada neste trabalho, basicamente para possibilitar a demonstração desse teorema.

Proposição 1. *Sejam y e x_i , valores reais positivos, tais que $y > x_i$ para $i = 1, \dots, n$. Defina $\sum_{i=2}^m x_i = 0$ para o caso em que $m = 1$. Temos então que,*

$$\int_0^\infty \dots \int_0^\infty \int_0^{y - \sum_{i=2}^m x_i} \prod_{i=1}^m (x_i^{b_i - 1} dx_i) = \frac{y^{\sum_{i=1}^m b_i} \prod_{i=1}^m \Gamma(b_i)}{\Gamma(\sum_{i=1}^m b_i + 1)} \quad (2.10)$$

onde m é um inteiro positivo finito, $b_i > 0$ para $i = 1, \dots, m$ e $\Gamma(\cdot)$ é a função gama.

Prova :

Provamos esta proposição por indução matemática em m . Para $m = 1$, a relação dada em 2.10 torna-se

$$\int_0^y x_1^{b_1 - 1} dx_1 = \frac{y^{b_1}}{b_1} = \frac{y^{b_1} \Gamma(b_1)}{\Gamma(b_1 + 1)}.$$

Consideremos que 2.10 é válido para $m = k$, onde $k \geq 2$, isto é

$$\int_0^\infty \dots \int_0^\infty \int_0^{y - \sum_{i=2}^k x_i} \prod_{i=1}^k (x_i^{b_i - 1} dx_i) = \frac{y^{\sum_{i=1}^k b_i} \prod_{i=1}^k \Gamma(b_i)}{\Gamma(\sum_{i=1}^k b_i + 1)}. \quad (2.11)$$

Queremos mostrar que 2.10, vale para $m = k + 1$. Temos então que

$$\int_0^\infty \dots \int_0^\infty \int_0^{y - \sum_{i=2}^{k+1} x_i} \prod_{i=1}^{k+1} (x_i^{b_i-1} dx_i) =$$

$$\int_0^\infty x_{k+1}^{b_{k+1}-1} \left\{ \int_0^\infty \dots \int_0^\infty \int_0^{(y-x_{k+1}) - \sum_{i=2}^k x_i} x_1^{b_1-1} \dots x_k^{b_k-1} dx_1 \dots dx_k \right\} dx_{k+1} \quad (2.12)$$

Temos pela hipótese de indução (2.11), que a expressão 2.12 resulta,

$$\left\{ \int_0^\infty x_{k+1}^{b_{k+1}-1} (y - x_{k+1})^{\sum_{i=1}^k b_i} dx_{k+1} \right\} \frac{\prod_{i=1}^k \Gamma(b_i)}{\Gamma(\sum_{i=1}^k b_i + 1)} \quad (2.13)$$

Para resolver o primeiro termo de 2.13, consideremos uma notação simplificada;

$$a = b_{k+1}, \quad b = \sum_{i=1}^k b_i \quad x = x_{k+1} \quad e \quad dx = dx_{k+1}.$$

Portanto, temos que solucionar

$$\int_0^\infty x^{a-1} (y - x)^b dx = \int_0^\infty x^{a+b-1} (y/x - 1)^b dx \quad (2.14)$$

Fazendo $u = y/x - 1 \Rightarrow x = y/(u + 1)$. Quando $x \rightarrow \infty \Rightarrow u \rightarrow -1$, e para $x \rightarrow 0 \Rightarrow u \rightarrow \infty$. Desta forma, a integral 2.14 torna-se

$$- \int_\infty^{-1} \left(\frac{y}{u+1} \right)^{a+b-1} \frac{u^b}{(u+1)^2} y du = y^{a+b} \int_{-1}^\infty \frac{u^b}{(u+1)^{a+b+1}} du$$

$$= y^{a+b} \int_0^\infty \frac{u^b}{(u+1)^{a+b+1}} du$$

pois

$$\int_{-1}^0 \frac{u^b}{(u+1)^{a+b+1}} du = 0,$$

visto que $u > 0$.

Aplicando-se uma nova transformação da forma, $z = u/(1 + u)$, observamos que

$$y^{a+b} \int_0^\infty \frac{u^b}{(u+1)^{a+b+1}} du = y^{a+b} \int_0^1 z^b (1-z)^{a-1} dz = y^{a+b} B(b+1, a)$$

onde $B(b+1, a)$ é a função beta dada no apêndice A. Segue então que 2.14 resulta,

$$\frac{y^{a+b} \Gamma(b+1) \Gamma(a)}{\Gamma(a+b+1)}. \quad (2.15)$$

De 2.15 temos então que 2.13 fica da seguinte forma,

$$y^{\sum_{i=1}^{k+1} b_i} \frac{\Gamma\left(\sum_{i=1}^k b_i + 1\right) \Gamma(b_{k+1}) \prod_{i=1}^k \Gamma(b_i)}{\Gamma\left(\sum_{i=1}^{k+1} b_i + 1\right) \Gamma\left(\sum_{i=1}^k b_i + 1\right)} = y^{\sum_{i=1}^{k+1} b_i} \frac{\prod_{i=1}^{k+1} \Gamma(b_i)}{\Gamma\left(\sum_{i=1}^{k+1} b_i + 1\right)}.$$

Portanto a proposição 1 está provada. \square

Vamos agora à demonstração do Teorema 3. Seja $f_{T_i}(x_i)$ a densidade de T_i e $F_{Y_m}(y)$ a função de distribuição de $Y_m = \sum_{i=1}^m T_i$, para $m = 2, \dots, (n-1)$. Então temos

$$\begin{aligned} G_n(y) &= P((T_1 + \dots + T_{n-1}) + T_n \leq y) \\ &= \int_0^\infty F_{Y_{n-1}}(y - x_n) f_{T_n}(x_n) dx_n \\ &= \int_0^\infty \int_0^\infty F_{Y_{n-2}}(y - x_n - x_{n-1}) f_{T_{n-1}}(x_{n-1}) f_{T_n}(x_n) dx_{n-1} dx_n \\ &\quad \dots \\ &= \int_0^\infty \dots \int_0^\infty \int_0^{y - \sum_{i=2}^n x_i} f_{T_1}(x_1) f_{T_2}(x_2) \dots f_{T_n}(x_n) dx_1 \dots dx_n \end{aligned} \quad (2.16)$$

onde este resultado é obtido pela independência de T_1, \dots, T_n , pela propriedade hereditária de funções disjuntas de v.a.'s independentes (James, 1981), e pelo uso sucessivo da relação:

$$P((T_1 + T_2) \leq y) = \int_0^\infty F_{T_1}(y - t_2) f_{T_2}(t_2) dt_2 = \int_0^\infty \int_0^{y-t_2} f_{T_1}(t_1) dt_1 f_{T_2}(x) dx .$$

Usando o resultado de 2.16 e incluindo as representações das densidades de T_i conforme dado em 2.1. obtemos,

$$\begin{aligned} G_n(y) &= \int_0^\infty \dots \int_0^\infty \int_0^{y-\sum_{i=2}^n x_i} \prod_{i=1}^n \left\{ \frac{\beta_i}{\alpha_i^{\beta_i k_i} \Gamma(k_i)} x_i^{\beta_i k_i - 1} \exp(-(x_i/\alpha_i)^{\beta_i}) dx_i \right\} \\ &= \gamma_n \int_0^\infty \dots \int_0^\infty \int_0^{y-\sum_{i=2}^n x_i} \left(\prod_{i=1}^n x_i^{\beta_i k_i - 1} \right) \exp \left[- \sum_{i=1}^n (x_i/\alpha_i)^{\beta_i} \right] \prod_{i=1}^n dx_i \\ &= \gamma_n \int_0^\infty \dots \int_0^\infty \int_0^{y-\sum_{i=2}^n x_i} \prod_{i=1}^n x_i^{\beta_i k_i - 1} \sum_{j=0}^\infty \frac{(-1)^j \left(\sum_{i=1}^n (x_i/\alpha_i)^{\beta_i} \right)^j}{j!} \prod_{i=1}^n dx_i . \end{aligned}$$

Uma partição de um inteiro positivo r em d parcelas é um conjunto de d inteiros não negativos r_1, \dots, r_d , tal que $r_1 + \dots + r_d = r$. Considerando todas as possíveis partições de r temos, a generalização do teorema binomial (Mood, Graybill, Boes (1974), Apêndice A), dado por

$$\left(\sum_{i=1}^d a_i \right)^r = \sum_{r_1 + \dots + r_d = r} \left\{ \frac{r!}{\prod_{i=1}^d r_i!} \prod_{i=1}^d a_i^{r_i} \right\} . \quad (2.17)$$

Usando este resultado 2.17 , chegamos à seguinte expressão;

$$G_n(y) = \gamma_n \int_0^\infty \dots \int_0^\infty \int_0^{y-\sum_{i=2}^n x_i} \prod_{i=1}^n x_i^{\beta_i k_i - 1} \sum_{j=0}^\infty \left\{ \frac{(-1)^j}{j!} \sum_{f_1 + \dots + f_n = j} \left[\frac{j!}{\prod_{i=1}^n f_i!} \prod_{i=1}^n (x_i/\alpha_i)^{\beta_i f_i} \right] \right\} \prod_{i=1}^n dx_i$$

$$= \gamma_n \sum_{j=0}^{\infty} (-1)^j \sum_{f_1+\dots+f_n=j} \left(\frac{\prod_{i=1}^n \alpha_i^{-\beta_i f_i}}{\prod_{i=1}^n f_i!} \right) \int_0^{\infty} \dots \int_0^{\infty} \int_0^{y-\sum_{i=2}^n x_i} \prod_{i=1}^n x_i^{\beta_i k_i + \beta_i f_i - 1} dx_i.$$

Fazendo $b_i = \beta_i k_i + \beta_i f_i$, considerando 2.9, e usando o resultado da proposição 1, temos,

$$G_n(y) = \gamma_n \sum_{j=0}^{\infty} (-1)^j \sum_{f_1+\dots+f_n=j} \left(\frac{\prod_{i=1}^n \alpha_i^{-\beta_i f_i}}{\prod_{i=1}^n f_i!} \right) y^{\sum_{i=1}^n b_i} \frac{\prod_{i=1}^n \Gamma(b_i)}{\Gamma\left(\sum_{i=1}^n b_i + 1\right)}$$

$$\Rightarrow G_n(y) = \gamma_n y^{\beta k} \sum_{j=0}^{\infty} (-1)^j A_j,$$

conforme queríamos demonstrar. \square

2.3 O Modelo Log-Gama Generalizada Reparametrizada

Conforme será visto mais adiante, existe uma série de dificuldades na estimação dos parâmetros da d.g.g.. Por esta razão, Prentice (1974) estudou este modelo de uma forma reparametrizada, visando facilitar a sua aplicação. Descrevemos nesta seção o processo de reparametrização que leva à extensão do modelo log-gama generalizada, que denotamos aqui por l.g.g. estendido, assim como uma posterior simplificação apresentada em Farewell e Prentice (1977) e Lawless (1980).

2.3.1 Descrição do Modelo

Considere T uma variável aleatória com distribuição dada por 2.1. Observou-se inicialmente que o modelo dado em 2.1 é log-linear, uma vez que $Y = \log T$ pode ser representado como

$$Y = \log \alpha + \beta^{-1} Z, \tag{2.18}$$

onde a variável Z tem distribuição log-gama com densidade

$$f_Z(z) = \frac{1}{\Gamma(k)} \exp(zk - e^z) \quad z \in \mathfrak{R}. \quad (2.19)$$

Assim sendo, a variável Y tem uma distribuição log-gama generalizada, seguindo um modelo de locação e escala com resíduo Z (ver Apêndice B, expressão B-4).

Prentice (1974) considerou parâmetros de transformação do tipo $q = k^{-c}$, ($c > 0$), para possibilitar que o modelo normal, correspondente ao limite de 2.1 quando $k \rightarrow \infty$, (conforme demonstrado na seção 2.2.1) pudesse ser definido em um ponto finito ($q = 0$) no espaço paramétrico. Especificamente, o autor optou por $c = 1/2$ tendo em vista que a distribuição de Y aproxima-se da normalidade apenas para este valor.

A esperança e a variância de Z são respectivamente $\mu^* = \Psi(k)$ e $\sigma^{*2} = \Psi'(k)$ onde $\Psi(k)$ e $\Psi'(k)$ são as funções digama e trigama (ver Apêndice A), cujas fórmulas assintóticas, com $k \rightarrow \infty$ são,

$$\Psi(k) \simeq \log k - \frac{1}{2k} - \frac{1}{12k^2} + \frac{1}{120k^4} - \frac{1}{252k^6} + \dots \quad (2.20)$$

$$\Psi'(k) \simeq \frac{1}{k} + \frac{1}{2k^2} + \frac{1}{6k^3} - \frac{1}{30k^5} + \frac{1}{42k^7} - \frac{1}{30k^9} + \dots \quad (2.21)$$

Então, desde que $\sigma^{*2} = \Psi'(k) \simeq \frac{1}{k} = q^2$, uma padronização aproximada de Z é

$$\epsilon = (Z - \mu^*)q^{-1}.$$

Desta forma, ϵ tem média zero e variância aproximadamente 1 quando $q \rightarrow 0$ e o modelo dado por 2.18 e 2.19 pode ser reparametrizado com segue,

$$\begin{aligned} Y &= \log \alpha + \beta^{-1}Z \\ &= \log \alpha + \beta^{-1}(q\epsilon + \mu^*) \\ &= \log \alpha + \mu^*\beta^{-1} + q\beta^{-1}\epsilon. \end{aligned}$$

Portanto,

$$Y = \gamma + \sigma\epsilon, \quad (2.22)$$

onde $\gamma = \log \alpha + \mu^* \beta^{-1}$ e $\sigma = q\beta^{-1}$. A f.d.p. de Y pode ser facilmente obtida, como mostrado abaixo.

$$f_Y(y) = f_Z[z(y)]|J|$$

sendo

$$z(y) = \frac{y - \gamma}{\sigma}q + \mu^* \quad e \quad J = \frac{\partial z(y)}{\partial y} = \frac{q}{\sigma}$$

onde J é o jacobiano da transformação. Considerando então a densidade de Z dada em 2.19, a f.d.p. de Y resulta

$$f_Y(y; \gamma, \sigma, q) = \frac{q}{\sigma \Gamma(q^{-2})} \exp \left[\left(\frac{y - \gamma}{\sigma} q + \mu^* \right) q^{-2} - \exp \left(\frac{y - \gamma}{\sigma} q + \mu^* \right) \right] \quad \begin{array}{l} y \in \mathfrak{R} \\ \gamma, \sigma, q > 0 \end{array} \quad (2.23)$$

Prentice (1974) também estendeu a família 2.23 para incluir distribuições com $q \leq 0$. Este modelo estendido é representado por

$$f_Y(y; \gamma, \sigma, q) = \begin{cases} \frac{|q|}{\Gamma(q^{-2})\sigma} \exp \left\{ q^{-2} \left(\frac{y - \gamma}{\sigma} q + \mu^* \right) - \exp \left(\frac{y - \gamma}{\sigma} q + \mu^* \right) \right\} & \text{se } q \neq 0 \\ \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2} \left(\frac{y - \gamma}{\sigma} \right)^2 \right\} & \text{se } q = 0 \end{cases} \quad (2.24)$$

onde $\mu^* = \Psi(q^{-2})$. Esta é uma família de 3 parâmetros (γ, σ e q), centrada na distribuição normal (quando $q = 0$), que inclui entre outras, as distribuições valor extremo para $q = 1$ e $q = -1$ (Kotz e Johnson, 1985, Vol .2, pag 606).

2.3.2 Simplificação do Modelo Reparametrizado

Uma versão simplificada de 2.24 foi apresentada em Farewell e Prentice (1977), e mais tarde por Lawless (1980). Neste modelo simplificado, com base em 2.18, a padronização considerada para o resíduo Z foi

$$W = \frac{Z - \log k}{k^{-1/2}} = \frac{Z - \log q^{-2}}{q},$$

que corresponde a aproximar $\Psi(k) = \Psi(q^{-2})$ para o primeiro termo de sua forma assintótica 2.20, ou seja $\Psi(q^{-2}) \simeq \log(q^{-2})$. A distribuição de W é obtida como segue,

$$\begin{aligned} f_W(w) &= f_Z\left(\frac{w}{\sqrt{k}} + \log k\right) \frac{1}{\sqrt{k}} \\ &= \frac{1}{\Gamma(k)} \exp\left\{\left(\frac{w}{\sqrt{k}} + \log k\right)k - \exp\left(\frac{w}{\sqrt{k}} + \log k\right)\right\} \frac{1}{\sqrt{k}} \\ &= \frac{1}{\Gamma(k)} \exp\left\{k^{1/2}w + k \log k - \exp\left(\frac{w}{\sqrt{k}}\right)k\right\} \frac{1}{\sqrt{k}} \\ &= \frac{1}{\Gamma(k)} k^{k-1/2} \exp\left\{w\sqrt{k} - k e^{w/\sqrt{k}}\right\}, \end{aligned} \tag{2.25}$$

e, substituindo k por q^{-2} na última expressão de 2.25, obtemos

$$f_W(w; q) = \frac{1}{\Gamma(q^{-2})} (q^{-2})^{q^{-2}-\frac{1}{2}} \exp\left\{q^{-1}w - q^{-2}e^{qw}\right\}, \quad \begin{array}{l} w \in \mathfrak{R} \\ q > 0. \end{array} \tag{2.26}$$

De forma análoga à reparametrização anterior, o modelo 2.18 torna-se

$$\begin{aligned} Y &= \log \alpha + \beta^{-1}Z \\ &= \log \alpha + \beta^{-1}(qW + \log q^{-2}) \\ &= \log \alpha + \beta^{-1} \log q^{-2} + q\beta^{-1}W. \end{aligned}$$

Portanto,

$$Y = \mu + \sigma W, \tag{2.27}$$

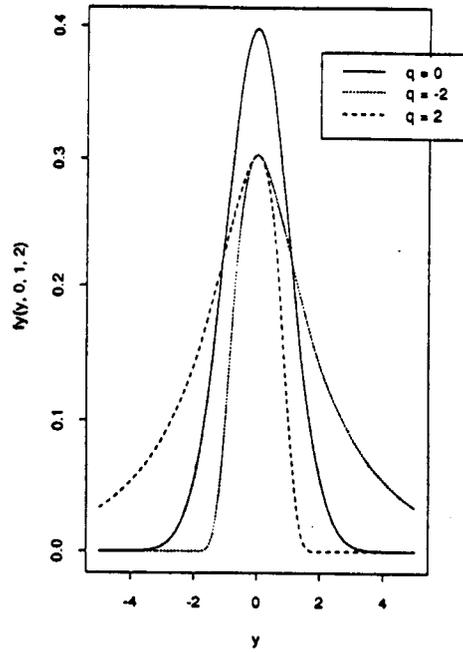


Figura 2-3: Formas da densidade log gama generalizada (l.g.g) estendida, para $q = -2, 0$ e 2 .

onde $\mu = \log \alpha + \beta^{-1} \log q^{-2}$, $\sigma = q\beta^{-1}$, e W representa o resíduo, com densidade dada em 2.26.

Uma extensão do modelo acima é obtida se consideramos $q \in \mathfrak{R}$. Temos então o modelo que chamamos **log-gama generalizada (l.g.g.) estendido**, dado abaixo

$$f_Y(y; \mu, \sigma, q) = \begin{cases} \frac{|q|}{\Gamma(\frac{|q|}{q-2})\sigma} (q^{-2})^{q-2} \exp \left\{ q^{-2} \left[\frac{y-\mu}{\sigma} q - \exp\left(\frac{y-\mu}{\sigma} q\right) \right] \right\} & \text{se } q \neq 0 \\ \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2} \left(\frac{y-\mu}{\sigma} \right)^2 \right\} & \text{se } q = 0. \end{cases} \quad (2.28)$$

Algumas Propriedades do Modelo Log-Gama Generalizada Estendido

A Figura 2-3 mostra algumas representações da densidade 2.28. A média e variância deste modelo, para o caso em que $q \neq 0$, são dadas respectivamente por,

$$\mu + \sigma \left(\frac{\Psi(q^{-2}) - \log(q^{-2})}{|q|} \right) e^{-\frac{\sigma^2 \Psi/(q^{-2})}{q^2}}.$$

Para $q = 0$, temos que $E(Y) = \mu$ e $V(Y) = \sigma^2$.

Teorema 4 .A função de sobrevivência de Y é dada por

$$S(y) = \begin{cases} Q(q^{-2}, q^{-2} \exp(wq)) & q > 0 \\ 1 - Q(q^{-2}, q^{-2} \exp(wq)) & q < 0 \\ 1 - \Phi(w) & q = 0 \end{cases}, \quad (2.29)$$

onde $w = (y - \mu)/\sigma$ e,

$$Q(k, a) = \int_a^\infty \frac{x^{k-1} e^{-x}}{\Gamma(k)} dx \quad (2.30)$$

é a integral gama incompleta.

Demonstração:

Para $q > 0$ a densidade de W é dada em 2.26. Temos então que

$$\begin{aligned} S(y) &= P(Y \geq y) = P(W \geq w) \\ &= \int_w^\infty \frac{1}{\Gamma(q^{-2})} (q^{-2})^{q^{-2}-1/2} \exp\{q^{-1}u - q^{-2}e^{qu}\} du \\ &= \int_w^\infty \frac{1}{\Gamma(q^{-2})} (q^{-2})^{q^{-2}-1} q^{-1} \exp(q^{-2}qu) \exp\{-q^{-2}e^{qu}\} du \\ &= \int_w^\infty \frac{(q^{-2})^{q^{-2}-1}}{\Gamma(q^{-2})} [\exp(qu)]^{q^{-2}-1} \exp\{-q^{-2}e^{qu}\} q^{-1} \exp(qu) du. \end{aligned}$$

Fazendo $x = q^{-2} \exp(qu)$ temos

$$S(y) = \frac{1}{\Gamma(q^{-2})} \int_{q^{-2} \exp(qw)}^{\infty} x^{q^{-2}-1} e^{-x} dx = Q(q^{-2}, q^{-2} e^{qw}) \quad q > 0.$$

Para $q < 0$, a densidade de W é da seguinte forma

$$f_W(w; q) = \frac{-q}{\Gamma(q^{-2})} (q^{-2})^{q^{-2}} \exp\{q^{-1}w - q^{-2}e^{qw}\} \quad \begin{array}{l} w \in \mathfrak{R} \\ q < 0. \end{array} \quad (2.31)$$

Usando a mesma transformação da demonstração anterior, chegamos a

$$\begin{aligned} S(y) &= -\frac{1}{\Gamma(q^{-2})} \int_{q^{-2} \exp(qw)}^0 x^{q^{-2}-1} e^{-x} dx \\ &= \frac{1}{\Gamma(q^{-2})} \int_0^{q^{-2} \exp(qw)} x^{q^{-2}-1} e^{-x} dx. \end{aligned} \quad (2.32)$$

Considerando que 2.32 representa a função de distribuição de uma gama de um parâmetro (1.6) aplicada no ponto $q^{-2} \exp(qw)$, concluímos que

$$S(y) = 1 - Q(q^{-2}, q^{-2} e^{qw}) \quad q < 0.$$

Para $q = 0$ o resultado é direto. \square

2.4 O Modelo de Regressão Gama Generalizada

Consideremos o modelo de regressão baseado na distribuição log-gama generalizada (l.g.g.) estendida, dada em 2.28, e nos conceitos dados no capítulo 1 sobre modelos de regressão do tipo locação e escala.

De forma similar ao que vem sendo atribuído ao modelo gama generalizada (ou l.g.g. estendido), o modelo de regressão baseado nesta distribuição pode ajustar adequadamente dados de vida, por ser um modelo bastante flexível, ou pode ser utilizado apenas para discriminar, dentre os modelos de regressão estudados no capítulo anterior, que são casos particulares deste, o mais apropriado para um determinado conjunto de dados.

2.4.1 Descrição do Modelo

Considere um modelo de regressão onde assumimos uma relação linear entre o log do tempo de vida, denotado por Y , e um vetor de covariáveis $\mathbf{x} = (x_0, x_1, \dots, x_{p-1})$. Y dado \mathbf{x} tem distribuição l.g.g. estendida se pode ser representado da forma.

$$Y = \mathbf{x}\beta + \sigma Z, \quad (2.33)$$

onde $\beta = (\beta_0, \beta_1, \dots, \beta_{p-1})$ e $\sigma > 0$, são parâmetros desconhecidos, e Z tem distribuição ($z \in \Re$), dada por

$$f_Z(z, q) = \begin{cases} \frac{|q|}{\Gamma(q^{-2})} (q^{-2})^{q-2} \exp\{q^{-1}z - q^{-2} \exp(qz)\} & q \neq 0 \\ \frac{1}{\sqrt{2\pi}} \exp\{-\frac{z^2}{2}\} & q = 0. \end{cases} \quad (2.34)$$

A função de sobrevivência de Y dado \mathbf{x} , analogamente a 2.29, é representada por

$$S(y|\mathbf{x}) = \begin{cases} Q(q^{-2}, q^{-2} \exp(q(y - \mathbf{x}\beta)/\sigma)) & q > 0 \\ 1 - Q(q^{-2}, q^{-2} \exp(q(y - \mathbf{x}\beta)/\sigma)) & q < 0 \\ 1 - \Phi[(y - \mathbf{x}\beta)/\sigma] & q = 0, \end{cases} \quad (2.35)$$

onde $Q(k, a)$ é a integral gama incompleta dada em 2.30.

Como casos particulares do modelo dado por 2.33 e 2.34, temos por exemplo os modelos de regressão exponencial dado em 1.10 para $q = \sigma = 1$, o Weibull dado em 1.11 para $q = 1$, (ou Weibull recíproco para $q = -1$), o log-normal para $q = 0$ da forma 1.12, e para $q > 0$ temos um modelo de regressão gama generalizada reparametrizado.

Para concluir esta seção, observamos que a família l.g.g. estendida 2.28 vem sendo utilizada para a modelagem de dados, no lugar da d.g.g. original, devido às simplificações adquiridas na estimação dos parâmetros. Como exemplo de uma aplicação recente, temos o artigo de Yamaguchi (1992), que utiliza o modelo log-gama generalizada estendido

(modificado pela introdução de *frações de sobrevivências*), para analisar dados referentes à mobilidade de emprego entre firmas no Japão.

Capítulo 3

Procedimentos Inferenciais para o Modelo Gama Generalizada

3.1 Introdução

As dificuldades básicas na análise estatística do modelo gama generalizada, ocorrem pela presença de três parâmetros desconhecidos, e pela complexidade computacional dos procedimentos inferenciais. Vários autores estudaram procedimentos de estimação e testes para os parâmetros usando o modelo original e posteriormente para o modelo reparametrizado log-gama generalizada (l.g.g.) estendido. No modelo original foram encontradas dificuldades, pela própria estrutura da sua função de verossimilhança. Devido às simplificações obtidas pelo uso do modelo l.g.g., este vem sendo utilizado no lugar do modelo gama generalizada original.

Considere inicialmente, um breve resumo bibliográfico que enfoca os principais trabalhos sobre os procedimentos inferenciais para esse modelo.

3.1.1 Breve Resumo Bibliográfico

Parr e Webster (1965) obtiveram as equações de máxima verossimilhança para os parâmetros, e sugeriram a solução destas através de métodos iterativos, tratando o parâmetro de locação como conhecido. Os autores utilizaram a normalidade assintótica dos estimadores

de máxima verossimilhança e apresentaram a densidade assintótica de $\hat{\beta}$ e \hat{d} , (onde $d = \beta k$, sob a parametrização dada em 2.1) utilizada para testar a adequação dos modelos Weibull ou exponencial.

Stacy e Mihran (1965) derivaram técnicas com base no método dos momentos, obtendo equações simultâneas sem fórmulas fechadas para sua solução. Apresentaram também três estimadores não viciados com respectivas variâncias para o caso onde apenas o parâmetro de escala era desconhecido.

Harter (1967) incluiu um quarto parâmetro de locação e formulou um procedimento iterativo para determinar os estimadores de máxima verossimilhança para amostras completas e censuradas.

Hager e Bain (1970) obtiveram as equações de máxima verossimilhança (sob a parametrização dada na expressão 2.1), e reduziram o sistema a uma única equação não linear. Notaram que o método de Newton Raphson não funcionava bem na solução desta equação, e que os e.m.v's de α, β, k deixam de convergir para amostras maiores que 200. Os autores apresentaram estatísticas com propriedades de funções pivotaes dos estimadores de máxima verossimilhança, e proporam estimadores máximo verossímeis para α e β , considerando k fixo. Como resultado importante notaram, através de simulação, que a aproximação normal para a distribuição do e.m.v. do parâmetro k pode ser inadequada mesmo para amostras de tamanho 400.

Em um trabalho importante Prentice (1974) desenvolveu a reparametrização e extensão do log da d.g.g., descrito no capítulo anterior, de forma a tornar as propriedades desta distribuição mais claras. Forneceu as equações de máxima verossimilhança e utilizou o método de Newton Raphson para obter os estimadores de máxima verossimilhança dos novos parâmetros de locação e escala (γ e σ) e $q = k^{-1/2}$. Para isso, usou um procedimento onde q é tratado como fixo em um primeiro estágio. O autor também estudou através de simulação a distribuição exata de \hat{q} e, com base em amostras de tamanho 25 e 50, observou que a aproximação normal para \hat{q} apresenta melhor concordância que a aproximação normal de \hat{k} para amostras de tamanho 400 no estudo de Hager e Bain

(1970).

Farewell e Prentice (1977) utilizaram o modelo reparametrizado e estendido de Prentice (1974) de forma simplificada, para discriminar entre os diversos modelos que representam casos particulares da d.g.g., na modelagem de dados de tempo de vida. Apresentaram várias aplicações em dados da indústria e da medicina que haviam sido ajustados por outros modelos em artigos publicados. Utilizaram um procedimento da razão de verossimilhança e argumentaram que possivelmente o tamanho amostral adequado para esta aplicação seria menor que o requerido para a aproximação normal da distribuição de \hat{q} dada em Prentice (1974).

Lawless (1980) considerou para amostras completas a utilização do modelo reparametrizado de Prentice (1974) também de forma simplificada, e formalizou em seu trabalho a utilização do método da razão de verossimilhança para a estimação dos parâmetros. Apresentou também procedimentos inferenciais exatos para estimação dos parâmetros, dos quantis e da função de sobrevivência, considerando k conhecido.

Em seu livro Lawless (1982) fornece as equações de máxima verossimilhança do modelo log-gama generalizada reparametrizado para $q > 0$, considerando também amostras censuradas. Descreveu as propriedades, testes e equações de máxima verossimilhança do modelo de regressão log-gama reparametrizado também apenas para o parâmetro q positivo.

Diciccio (1987) utilizou o modelo log-gama reparametrizado (Prentice, 1974 ; Farewell e Prentice, 1977 ; Lawless, 1980) para desenvolver um método inferencial que corresponde a um ajustamento do método da razão de verossimilhança que visa melhorar a previsão do método para pequenas amostras.

Rao, Katan e Narosimham (1991) consideraram a d.g.g. com 4 parâmetros da forma originalmente dada por Harter (1967), e forneceram coeficientes para o cálculo das estimativas BLUE (best linear unbiased estimators) dos parâmetros de locação e escala deste modelo.

Nesta seção abordaremos basicamente os resultados obtidos por Hager e Bain (1970),

Prentice (1974), Farewell e Prentice (1977) e Lawless (1980,1982).

3.2 Inferências com Amostras Completas

Esta seção trata apenas de procedimentos inferenciais para o caso de amostras completas, de populações homogêneas, e além de sua aplicação a dados de vida sem censura, serve como base para o desenvolvimento da próxima seção que considera amostras censuradas. Estudamos a estimação pontual sob o modelo original, e posteriormente com o modelo reparametrizado l.g.g., e apresentamos a construção de testes e intervalos baseados no modelo l.g.g. estendido.

3.2.1 Estimação Pontual

O Modelo Original Gama Generalizada

Considere que as observações t_1, \dots, t_n representam uma amostra aleatória de uma d.g.g. com densidade dada em 2.1, denotada por $f(t_i, \theta)$, sendo $\theta = (\alpha, \beta, k)$ o vetor de parâmetros desconhecidos. A função de log verossimilhança para θ , é dada por

$$L^*(\theta) = \log \left[\prod_{i=1}^n f(t_i, \theta) \right]$$

$$\Rightarrow L^*(\theta) = n \log \beta - n \log \alpha - n \log \Gamma(k) + (\beta k - 1) \sum_{i=1}^n \log(t_i/\alpha) - \sum_{i=1}^n (t_i/\alpha)^\beta. \quad (3.1)$$

As equações de log verossimilhança para α, β e k são :

$$\begin{cases} -n\hat{k} + \sum_{i=1}^n (t_i/\hat{\alpha})^{\hat{\beta}} & = 0 \\ n/\hat{\beta} + \hat{k} \sum_{i=1}^n \log(t_i/\hat{\alpha}) - \sum_{i=1}^n (t_i/\hat{\alpha})^{\hat{\beta}} \log(t_i/\hat{\alpha}) & = 0 \\ -n\Psi(\hat{k}) + \hat{\beta} \sum_{i=1}^n \log(t_i/\hat{\alpha}) & = 0 \end{cases} \quad (3.2)$$

onde $\Psi(k)$ é a função digama definida no apêndice A. O sistema de equações 3.2 pode ser reduzido a uma única equação não linear com uma única incógnita, da seguinte forma :

Da primeira equação temos

$$\hat{\alpha} = \left[\sum_{i=1}^n t_i^{\hat{\beta}} / (n\hat{k}) \right]^{1/\hat{\beta}}. \quad (3.3)$$

Substituindo este resultado na 2ª equação de 3.2, segue que,

$$\begin{aligned} \hat{k} \sum_{i=1}^n \log t_i - \frac{n\hat{k}}{\hat{\beta}} \log \left(\sum_{i=1}^n t_i^{\hat{\beta}} / (n\hat{k}) \right) - \frac{n\hat{k} \sum_{i=1}^n t_i^{\hat{\beta}} \log t_i}{\sum_{i=1}^n t_i^{\hat{\beta}}} + \frac{n\hat{k}}{\hat{\beta}} \log \left(\sum_{i=1}^n t_i^{\hat{\beta}} / n\hat{k} \right) &= \frac{-n}{\hat{\beta}} \\ \Rightarrow \hat{k} \left\{ \hat{\beta} \left[\frac{\sum_{i=1}^n \log t_i}{n} - \left(\sum_{i=1}^n t_i^{\hat{\beta}} \log t_i \right) / \sum_{i=1}^n t_i^{\hat{\beta}} \right] \right\} &= -1 \\ \Rightarrow \hat{k} &= - \left\{ \hat{\beta} \left[\frac{\sum_{i=1}^n \log t_i}{n} - \left(\sum_{i=1}^n t_i^{\hat{\beta}} \log t_i \right) / \sum_{i=1}^n t_i^{\hat{\beta}} \right] \right\}^{-1}. \end{aligned} \quad (3.4)$$

A terceira equação torna-se

$$-\Psi(\hat{k}) + \hat{\beta} \frac{\sum_{i=1}^n \log t_i}{n} - \log \left(\sum_{i=1}^n t_i^{\hat{\beta}} \right) + \log n\hat{k} = 0. \quad (3.5)$$

Substituindo 3.4 em 3.5 temos uma equação não linear de $\hat{\beta}$ (denotamos $g(\hat{\beta})$). A

determinação do ponto $\hat{\beta}$ que satisfaz a igualdade, $g(\hat{\beta}) = 0$, implica na obtenção de \hat{k} e $\hat{\alpha}$ respectivamente. Hager e Bain (1970) argumentaram que o método iterativo de Newton Raphson não obtém resultados satisfatórios na resolução de 3.5.

Estimação de Máxima Verossimilhança para o Modelo l.g.g. Estendido.

Estudando o modelo reparametrizado l.g.g. descrito no capítulo anterior, dado por 2.26 e 2.27, observou-se que a sua f.d.p. muda muito pouco quando $k = q^{-2}$ cresce de 1 para ∞ , ou equivalentemente, q decresce (cresce) de ± 1 para 0 (ver Lawless, 1982) e, a menos que a amostra seja muito grande, seria em geral difícil estimar de maneira precisa o parâmetro k . Contudo os problemas são bem maiores se a parametrização original α e β é utilizada ao invés de μ e σ do modelo 2.27. Se duas distribuições com valores similares para μ e σ , mas valores de k muito diferentes podem ser bem parecidas, em termos de α e β as duas distribuições poderiam ter também valores muito diferentes de α e β . Desta forma, uma justificativa para utilizarmos o modelo reparametrizado é que, com a parametrização (α, β, k) as dificuldades no exame da função de verossimilhança ou obtenção dos e.m.v.'s são maiores do que com a parametrização $(\mu, \sigma, k = q^{-2})$. Descrevemos a seguir o procedimento para estimação dos parâmetros μ, σ e q pelo método da máxima verossimilhança com amostras completas, utilizando o modelo reparametrizado l.g.g. estendido.

Seja y_1, \dots, y_n uma amostra de uma distribuição com densidade 2.28, e $\theta = (\mu, \sigma, q)$ o vetor de parâmetros desconhecidos. O log da função de verossimilhança, que chamaremos de *função de log-verossimilhança*, para $q \neq 0$ é dada por,

$$L^*(\theta) = \sum_{i=1}^n \log f(y_i; \mu, \sigma, q)$$

$$= n \left[\log |q| + q^{-2} \log(q^{-2}) - \log(\Gamma(q^{-2})\sigma) \right] + \sum_{i=1}^n \left\{ q^{-1} \left(\frac{y_i - \mu}{\sigma} \right) - q^{-2} \exp \left[q \left(\frac{y_i - \mu}{\sigma} \right) \right] \right\}. \quad (3.6)$$

Uma maneira de obter os estimadores de máxima verossimilhança é maximizar $L^*(\theta)$

pela resolução simultânea das equações $\frac{\partial L^*(\theta)}{\partial \mu} = 0$, $\frac{\partial L^*(\theta)}{\partial \sigma} = 0$ e $\frac{\partial L^*(\theta)}{\partial q} = 0$. Contudo, pelas dificuldades computacionais deste procedimento, optamos por um método menos trabalhoso que consiste na realização de iterações em 2 estágios, com q tratado como fixo no 1º estágio. Para um dado valor de q encontramos os valores $\tilde{\mu}(q)$ e $\tilde{\sigma}(q)$ que maximizam $L^*(\theta)$, pela resolução de $\frac{\partial L^*(\theta)}{\partial \mu} = 0$ e $\frac{\partial L^*(\theta)}{\partial \sigma} = 0$. Sendo feito isso para alguns valores de q , podemos obter a função de verossimilhança para q e localizar \hat{q} (ver Lawless, 1982). As derivadas de $L^*(\theta)$ com respeito a μ e a σ são dadas abaixo (para $q \neq 0$)

$$\frac{\partial L^*(\theta)}{\partial \mu} = -\frac{n}{q\sigma} + \frac{1}{q\sigma} \sum_{i=1}^n \exp \left[q \left(\frac{y_i - \mu}{\sigma} \right) \right]$$

$$\frac{\partial L^*(\theta)}{\partial \sigma} = -\frac{n}{\sigma} - \sum_{i=1}^n \left(\frac{y_i - \mu}{q\sigma^2} \right) + \sum_{i=1}^n \left(\frac{y_i - \mu}{q\sigma^2} \right) \exp \left[q \left(\frac{y_i - \mu}{\sigma} \right) \right].$$

Temos então as equações

$$\sum_{i=1}^n \exp \left[q \left(\frac{y_i - \mu}{\sigma} \right) \right] = n \tag{3.7}$$

$$\sum_{i=1}^n \left(\frac{y_i - \mu}{\sigma} \right) \exp \left[q \left(\frac{y_i - \mu}{\sigma} \right) \right] - \sum_{i=1}^n \left(\frac{y_i - \mu}{\sigma} \right) = nq.$$

Após manipulações e substituições adequadas em 3.7, necessitamos resolver uma única equação, ou seja, as equações de 3.7 são equivalentes a

$$e^{\tilde{\mu}} = \left\{ \frac{1}{n} \sum_{i=1}^n \exp \left(\frac{qy_i}{\sigma} \right) \right\}^{\sigma/q} \tag{3.8}$$

e

$$\frac{\sum_{i=1}^n y_i e^{y_i q/\sigma}}{\sum_{i=1}^n e^{y_i q/\sigma}} - \bar{y} - \tilde{\sigma} q = 0. \tag{3.9}$$

Desta forma, encontramos $\tilde{\sigma}(q)$ pela resolução de 3.9 iterativamente, e com este resultado, obtemos $\tilde{\mu}(q)$ de 3.8. Vejamos a demonstração destes resultados.

Demonstração:

De 3.7 obtemos

$$\begin{aligned} \sum_{i=1}^n \exp \left[q \left(\frac{y_i - \mu}{\sigma} \right) \right] &= n \Rightarrow \frac{e^{\mu q/\sigma}}{\sum_{i=1}^n e^{y_i q/\sigma}} = \frac{1}{n} \\ \Rightarrow e^{\tilde{\mu}} &= \left\{ \frac{\sum_{i=1}^n e^{y_i q/\sigma}}{n} \right\}^{\sigma/q}. \end{aligned} \quad (3.10)$$

Para substituir o resultado 3.10 na segunda equação de 3.7 temos o seguinte desenvolvimento,

$$\begin{aligned} \sum_{i=1}^n \left(\frac{y_i - \mu}{\sigma} \right) \exp \left[q \left(\frac{y_i - \mu}{\sigma} \right) \right] - \sum_{i=1}^n \left(\frac{y_i - \mu}{\sigma} \right) &= nq \\ \Rightarrow e^{-\mu q/\sigma} \sum_{i=1}^n \left(\frac{y_i - \mu}{\sigma} \right) e^{y_i q/\sigma} - \sum_{i=1}^n \left(\frac{y_i - \mu}{\sigma} \right) &= nq \\ \Rightarrow e^{-\mu q/\sigma} \left\{ \sum_{i=1}^n \frac{y_i}{\sigma} e^{y_i q/\sigma} - \frac{\mu}{\sigma} \sum_{i=1}^n e^{y_i q/\sigma} \right\} - \sum_{i=1}^n \left(\frac{y_i - \mu}{\sigma} \right) &= nq \\ \Rightarrow e^{-\mu q/\sigma} \left\{ \sum_{i=1}^n \frac{y_i}{\sigma} e^{y_i q/\sigma} - \frac{n\mu}{\sigma} \left[\left(\frac{\sum_{i=1}^n e^{y_i q/\sigma}}{n} \right)^{\sigma/q} \right]^{q/\sigma} \right\} - \sum_{i=1}^n \left(\frac{y_i - \mu}{\sigma} \right) &= nq. \end{aligned}$$

Por 3.10 temos

$$\begin{aligned} \Rightarrow e^{-\mu q/\sigma} \left\{ \sum_{i=1}^n \frac{y_i}{\sigma} e^{y_i q/\sigma} - \frac{n\mu}{\sigma} e^{\mu q/\sigma} \right\} - \sum_{i=1}^n \left(\frac{y_i - \mu}{\sigma} \right) &= nq \\ \Rightarrow \left(\frac{1}{e^{\mu}} \right)^{q/\sigma} \sum_{i=1}^n \frac{y_i}{\sigma} e^{y_i q/\sigma} - \frac{n\mu}{\sigma} - \sum_{i=1}^n \left(\frac{y_i - \mu}{\sigma} \right) &= nq \end{aligned}$$

$$\Rightarrow \frac{\frac{1}{\sigma} \sum_{i=1}^n y_i e^{y_i q / \sigma}}{\left\{ \left(\frac{\sum_{i=1}^n e^{y_i q / \sigma}}{n} \right)^{\sigma/q} \right\}^{q/\sigma}} - \frac{n\mu}{\sigma} - \frac{1}{\sigma} n\bar{y} + \frac{n\mu}{\sigma} = nq$$

$$\Rightarrow \frac{\sum_{i=1}^n y_i e^{y_i q / \tilde{\sigma}}}{\sum_{i=1}^n e^{y_i q / \tilde{\sigma}}} - \bar{y} - \tilde{\sigma}q = 0. \quad \square$$

Para o caso em que $q = 0$, os e.m.v.'s são os bem conhecidos estimadores para a distribuição normal, isto é,

$$\tilde{\mu}(0) = \bar{y} \quad e \quad \tilde{\sigma}(0) = \left\{ \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \right\}^{1/2}.$$

Com $\tilde{\mu}(q)$ e $\tilde{\sigma}(q)$ calculados para uma série de valores de q , a *função de máxima verossimilhança maximizada* dada por,

$$L_{\max}(q) = L(\tilde{\mu}(q), \tilde{\sigma}(q), q) \quad (3.11)$$

pode ser considerada suficientemente precisa para obter o e.m.v. \hat{q} , que é o valor que maximiza $L_{\max}(q)$. Como resultado temos os estimadores de máxima verossimilhança $\hat{\mu} = \tilde{\mu}(\hat{q})$, $\hat{\sigma} = \tilde{\sigma}(\hat{q})$ e \hat{q} .

3.2.2 Intervalos de Confiança e Testes de Hipóteses

Tendo em vista que desejamos generalizar os resultados obtidos aqui, para trabalharmos com amostras censuradas e variáveis regressoras, abordamos nesta seção os métodos da razão de verossimilhança, que permitem esta generalização.

Lawless (1980), estudou testes exatos para os parâmetros de locação e escala do modelo dado por 2.26 e 2.27, mas este procedimento não permite a generalização para amostras censuradas, não sendo portanto abordado aqui. A distribuição assintótica dos

e.m.v.'s pode também ser utilizada, embora existam algumas evidências de que para modelos de tempo de vida este método é em geral menos preciso que o método da razão de verossimilhança (Doganaksoy e Shemee, 1993).

Verossimilhança Relativa Maximizada

Considere a função de máxima verossimilhança maximizada 3.11, utilizada para obter os estimadores de máxima verossimilhança $\hat{\mu}$, $\hat{\sigma}$ e \hat{q} . Definimos a *função de verossimilhança relativa maximizada* como

$$R_{\max}(q) = L_{\max}(q)/L(\hat{\mu}, \hat{\sigma}, \hat{q}).$$

Note que $0 < R_{\max}(q) \leq 1$ e $R_{\max}(\hat{q}) = 1$. Esta função é utilizada para a construção de testes da razão de verossimilhança e seu gráfico contra os valores de q oferece informações importantes na obtenção de intervalos de confiança para q , e na escolha de um modelo mais simples para os dados (Weibull ou log-normal).

Teste de Hipóteses e Intervalo de Confiança para o Parâmetro de Forma (q)

Suponha que desejamos testar hipóteses do tipo $H_0 : q = q_0$ contra $H_1 : q \neq q_0$. Para isso, podemos usar a estatística da razão de verossimilhança equivalente a 1.23, que é dada por

$$\Lambda = -2 \log R_{\max}(q_0).$$

Se o valor q_0 é um ponto interior do espaço paramétrico, a distribuição de Λ sob a hipótese H_0 é assintoticamente χ^2 com 1 grau de liberdade (Lawless, 1980). Neste momento, pode-se perceber a importância da parametrização $q = k^{-1/2}$ associada à construção do modelo estendido 2.28 que admite $q \in \mathfrak{R}$, tendo em vista que para este modelo é possível testar o importante caso $q = 0$ (modelo log-normal para $T = \exp(Y)$). Isto não seria possível na parametrização original, pois neste caso precisaríamos testar um limitante do

espaço paramétrico ($k = \infty$), o que representaria a violação de uma condição necessária à aplicação de métodos assintóticos (Cordeiro, 1992).

Teste para o Modelo Weibull (Valor Extremo) Testar a adequação do modelo valor extremo para T , corresponde a testar as hipóteses $H_0 : q = 1$ contra $H_1 : q \neq 1$. Usamos então a estatística

$$\Lambda = -2 \log \frac{L_{\max}(\hat{\mu}(1), \hat{\sigma}(1), 1)}{L(\hat{\mu}, \hat{\sigma}, \hat{q})}.$$

Para n suficientemente grande, Λ tem distribuição aproximadamente $\chi^2(1)$. Considerando o χ^2_{ξ} o ponto da distribuição qui-quadrado com 1 grau de liberdade, tal que $P(\chi^2(1) \leq \chi^2_{\xi}) = \xi$, ou seja, o ξ -ésimo quantil de uma χ^2_1 (qui-quadrado com 1 grau de liberdade), e seja Λ_{obs} o valor observado da variável Λ . Admitiremos a rejeição da hipótese H_0 , correspondendo à rejeição do modelo valor extremo, se $\Lambda_{obs} \geq \chi^2_{\xi}$, a um nível $(1 - \xi)100\%$.

Teste para o Modelo Log-normal (Normal)

Para testar a adequacidade do modelo normal para $\log T$, formulamos as hipóteses $H_0 : q = 0$ contra $H_1 : q \neq 0$, e de forma similar, usamos a estatística

$$\Lambda = -2 \log \frac{L_{\max}(\hat{\mu}(0), \hat{\sigma}(0), 0)}{L(\hat{\mu}, \hat{\sigma}, \hat{q})},$$

e adotando a mesma notação dada acima, rejeitamos a um nível $(1 - \xi)100\%$ a hipótese H_0 , ou seja, o modelo normal para Y , se $\Lambda_{obs} \geq \chi^2_{\xi}$.

Intervalo de Confiança para q Um intervalo de confiança aproximado de nível $\xi\%$ para q é obtido pelo conjunto de valores de q_0 para o qual $\Lambda = \Lambda(q_0)$ tem valores menores que o ponto χ^2_{ξ} dado acima. Isto pode ser resolvido através do gráfico $R_{\max}(q)$ ou $-2 \log R_{\max}(q)$ como função de q , ou alternativamente, através de um método iterativo

de pesquisa.

Teste para Outros Parâmetros

Suponha agora que temos interesse em testar outras hipóteses paramétricas com o método da razão de verossimilhança, como por exemplo, hipóteses sobre parâmetros originais α e β de 2.1. Para o caso geral, considerando hipóteses bilaterais H_o contra H_1 , precisamos obter o máximo da função de verossimilhança $L(\mu, \sigma, q)$ sob a hipótese H_o , denotado por $L_{\max}(H_o)$. Usamos a estatística

$$R_{\max}(q) = L_{\max}(H_o) / L(\hat{\mu}, \hat{\sigma}, \hat{q}),$$

para estabelecer testes de interesse.

Teste para o Modelo Gama O teste da adequação do modelo gama é dado pelas hipóteses $H_o : \beta = 1$ contra $H_1 : \beta \neq 1$, onde β é o parâmetro do modelo original 2.1. Considerando $L_{\max}(H_o)$ a função de verossimilhança $L(\mu, \sigma, q)$ maximizada sob a hipótese H_o , temos então a estatística

$$\Lambda = -2 \log \frac{L_{\max}(H_o)}{L(\hat{\mu}, \hat{\sigma}, \hat{q})}, \quad (3.12)$$

que sob H_o , tem distribuição assintótica χ^2 com 1 grau de liberdade. Podemos testar as hipóteses da maneira usual.

Teste para o Modelo Exponencial O modelo exponencial ocorre quando $\beta = k = 1$ no modelo original 2.1, que corresponde a $\sigma = q = 1$ no modelo reparametrizado. Testamos então as hipóteses

$$H_o : \left\{ \begin{pmatrix} \sigma \\ q \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \quad \text{versus} \quad H_1 : \left\{ \begin{pmatrix} \sigma \\ q \end{pmatrix} \neq \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\},$$

através da estatística 3.12 que aqui tem distribuição assintótica χ^2 com 2 graus de liber-

dade, sob a hipótese H_0 .

3.3 Inferências com Dados Censurados e Variáveis Regressoras

Discutimos até aqui procedimentos inferenciais para amostras completas e sem a presença de covariáveis. Fornecemos a seguir o procedimento de estimação dos parâmetros pelo método da máxima verossimilhança, assim como métodos assintóticos para construção de testes e intervalos de confiança aproximados sob o modelo de regressão l.g.g. estendido para o caso de amostras censuradas, de forma similar ao que foi dado na seção anterior.

Em uma primeira etapa consideramos a estimação dos parâmetros sob o modelo l.g.g. estendido sem covariáveis, e depois tratamos do caso mais geral, ou seja, o modelo de regressão com amostras censuradas.

3.3.1 Procedimentos Inferenciais com Amostras Censuradas Sem Variáveis Regressoras

Estimação Pontual pelo Método da Máxima Verossimilhança

Sejam y_1, \dots, y_n observações de uma amostra onde y_i representa o logaritmo do tempo de vida ou do tempo de censura de cada um dos n itens em estudo. Suponha que esta amostra é proveniente de uma distribuição l.g.g. estendida, representada por 2.28. Denotemos ainda o número de falhas observadas por r , os conjuntos de itens que falharam e foram censurados respectivamente por, F e C , e seja $\theta = (\mu, \sigma, q)$.

A função de log-verossimilhança para este modelo, com base no resultado dado em 1.15 com $p = 1$, é dada por

$$L^*(\theta) = \begin{cases} \sum_{i \in F} \log f(y_i, \theta) + \sum_{i \in C} \log S(y_i) & q \neq 0 \\ -r \log \sigma - \frac{1}{2} \sum_{i \in F} w_i^2 + \sum_{i \in C} \log(1 - \Phi(w_i)) & q = 0, \end{cases} \quad (3.13)$$

onde $w_i = \frac{y_i - \mu}{\sigma}$, $f(y_i, \theta)$ é a densidade da log-gama para $q \neq 0$ dada em 2.28, $S(y_i)$ é a sua função de sobrevivência dada em 2.29 para $q \neq 0$, e $\Phi(x)$ é a função de distribuição da normal padrão.

Para obtermos os estimadores de máxima verossimilhança $\hat{\mu}$, $\hat{\sigma}$ e \hat{q} precisaríamos derivar a função de log-verossimilhança $L^*(\theta)$ dada em 3.13, com respeito a cada um dos parâmetros μ , σ e q (no caso em que $q \neq 0$). Contudo a dificuldade existente na derivação de $L^*(\theta)$ com respeito a q é ainda maior neste caso, onde admitimos censuras. Usamos então o método descrito na seção anterior, que sugere a estimação em 2 estágios, considerando inicialmente q fixo. Desta forma, precisamos apenas das derivadas com respeito a μ e a σ .

Derivadas de Primeira Ordem para Dados Censurados sem Variáveis Regressoras

(i) Para $q = 0$

Neste caso temos a distribuição normal gerando as observações y_1, \dots, y_n . As respectivas derivadas da função de log-verossimilhança são,

$$\frac{\partial L^*(\theta)}{\partial \mu} = \frac{1}{\sigma} \sum_{i \in F} w_i + \frac{1}{\sigma} \sum_{i \in C} \phi(w_i)/(1 - \Phi(w_i)). \quad (3.14)$$

$$\frac{\partial L^*(\theta)}{\partial \sigma} = \frac{-r}{\sigma} + \frac{1}{\sigma} \sum_{i \in F} w_i^2 + \frac{1}{\sigma} \sum_{i \in C} w_i \phi(w_i)/(1 - \Phi(w_i)). \quad (3.15)$$

(ii) Para $q \neq 0$

Consideremos $u_i = q^{-2} \exp [(y_i - \mu)q/\sigma]$ e $S(y_i)$ a função de sobrevivência para $q \neq 0$, dada como segue

$$S(y_i) = \begin{cases} Q(q^{-2}, u_i) & q > 0 \\ 1 - Q(q^{-2}, u_i) & q < 0, \end{cases} \quad (3.16)$$

onde $Q(k, a)$ é integral gama incompleta. As derivadas de $L^*(\theta)$ com relação a μ e σ são,

$$\frac{\partial L^*(\theta)}{\partial \mu} = \frac{-1}{q\sigma} \left(r - \sum_{i \in F} e^{qw_i} \right) + \sum_{i \in C} \frac{|q| u_i^{q-2} e^{-u_i}}{\sigma \Gamma(q^{-2}) S(y_i)} \quad (3.17)$$

$$\frac{\partial L^*(\theta)}{\partial \sigma} = \frac{-1}{q\sigma} \left(qr + \sum_{i \in F} w_i - \sum_{i \in F} w_i e^{qw_i} \right) + \sum_{i \in C} \frac{|q| w_i u_i^{q-2} e^{-u_i}}{\sigma \Gamma(q^{-2}) S(y_i)}. \quad (3.18)$$

Isto pode ser visto usando 3.6 e as relações $\frac{\delta u_i}{\delta \mu} = -u_i q/\sigma$ e $\frac{\delta u_i}{\delta \sigma} = -w_i u_i q/\sigma$, para a parcela de falha, e derivando $S(y_i)$ que representa a parcela de censura.

Os estimadores de máxima verossimilhança para μ e σ são obtidos pela resolução do sistema

$$\begin{cases} \frac{\partial L^*(\theta)}{\partial \mu} = 0 \\ \frac{\partial L^*(\theta)}{\partial \sigma} = 0, \end{cases} \quad (3.19)$$

com as expressões dadas em 3.14 e 3.15 ou 3.17 e 3.18 de acordo com o valor de q . Contudo, não conseguimos obter analiticamente os estimadores de máxima verossimilhança, e precisamos de um método iterativo para solucionar o sistema 3.19. O método de Newton Raphson, que é comumente empregado para este tipo de problema, utiliza as derivadas de segunda ordem da equação de log-verossimilhança. Com isso em vista, e pelo fato de precisarmos destas derivadas para construir a matriz de informação observada 1.21, frequentemente usada na aplicação de resultados assintóticos da verossimilhança, fornecemos a seguir, as expressões das segundas derivadas com respeito aos parâmetros μ e σ .

Derivadas de Segunda Ordem para Dados Censurados sem Variáveis Regressoras

(i) Para $q = 0$

Para o caso da distribuição normal ($q = 0$), usando 3.14 e 3.15, temos abaixo as expressões para as segundas derivadas com respeito aos parâmetros μ e σ ,

$$\begin{aligned}\frac{\partial^2 L^*(\theta)}{\partial \mu^2} &= -\frac{r}{\sigma^2} - \frac{1}{\sigma^2} \sum_{i \in C} \lambda(w_i) \\ \frac{\partial^2 L^*(\theta)}{\partial \sigma^2} &= \frac{r}{\sigma^2} - \frac{3}{\sigma^2} \sum_{i \in F} w_i^2 - \frac{1}{\sigma^2} \left(\sum_{i \in C} 2w_i V(w_i) + \sum_{i \in C} w_i^2 \lambda(w_i) \right) \\ \frac{\partial^2 L^*(\theta)}{\partial \mu \partial \sigma} &= -\frac{2}{\sigma^2} \sum_{i \in F} w_i - \frac{1}{\sigma^2} \left(\sum_{i \in C} V(w_i) + \sum_{i \in C} w_i \lambda(w_i) \right),\end{aligned}\tag{3.20}$$

onde definimos as funções

$$V(w_i) = \phi(w_i)/(1 - \Phi(w_i)) = -\frac{\delta \log(1 - \Phi(w_i))}{\delta w_i} \quad e\tag{3.21}$$

$$\lambda(w_i) = V(w_i)[V(w_i) - w_i] = \frac{\delta V(w_i)}{\delta w_i},$$

sendo que $\phi(\cdot)$ e $\Phi(\cdot)$ representam como antes, a densidade e a função de distribuição da normal padrão.

(ii) Para $q \neq 0$

Neste caso temos expressões um pouco mais complicadas. Por esta razão consideramos a seguinte notação,

$$L^*(\theta) = L_F^*(\theta) + L_C^*(\theta)\tag{3.22}$$

onde $L^*(\theta)$ é a log-verossimilhança 3.13 para $q \neq 0$ e $L_F^*(\theta)$ e $L_C^*(\theta)$ representam as parcelas da log-verossimilhança para dados de falha e para dados censurados, respectivamente.

Mostramos separadamente as segundas derivadas para cada uma destas parcelas.

Considere inicialmente as expressões abaixo, para possibilitar o desenvolvimento das derivadas parciais pela regra da cadeia, onde por simplificação usamos a notação adcional $Q = Q(q^{-2}, u_i)$ e como antes, $u_i = q^{-2} \exp[(y_i - \mu)q/\sigma] = q^{-2} \exp(w_i q)$.

$$\frac{\partial u_i}{\partial \mu} = \frac{-u_i q}{\sigma}$$

$$\frac{\partial u_i}{\partial \sigma} = \frac{-q w_i u_i}{\sigma}$$

$$\frac{\partial^2 u_i}{\partial \mu^2} = \left(\frac{q}{\sigma}\right)^2 u_i$$

$$\frac{\partial^2 u_i}{\partial \sigma^2} = \left(\frac{q}{\sigma}\right)^2 u_i w_i^2 + \frac{2q u_i w_i}{\sigma^2}$$

$$\frac{\partial^2 u_i}{\partial \mu \partial \sigma} = \frac{q^2 u_i w_i}{\sigma^2} + \frac{q u_i}{\sigma^2}$$

$$\frac{\partial Q}{\partial u_i} = -\frac{\partial(1-Q)}{\partial u_i} = -\frac{u_i^{q^{-2}-1} e^{-u_i}}{\Gamma(q^{-2})}$$

$$\frac{\partial^2 Q}{\partial u_i^2} = -\frac{\partial^2(1-Q)}{\partial u_i^2} = \frac{u_i^{q^{-2}-2} e^{-u_i} (u_i - q^{-2} + 1)}{\Gamma(q^{-2})}$$

$$\frac{\partial \log Q}{\partial u_i} = \frac{1}{Q} \frac{\partial Q}{\partial u_i}$$

$$\frac{\partial \log(1-Q)}{\partial u_i} = \frac{1}{(1-Q)} \frac{\partial(1-Q)}{\partial u_i}$$

$$\frac{\partial^2 \log Q}{\partial u_i^2} = \frac{1}{Q} \frac{\partial^2 Q}{\partial u_i^2} - \frac{1}{Q^2} \left(\frac{\partial Q}{\partial u_i}\right)^2$$

$$\frac{\partial^2 \log(1-Q)}{\partial u_i^2} = \frac{1}{(1-Q)} \frac{\partial^2(1-Q)}{\partial u_i^2} - \frac{1}{(1-Q)^2} \left(\frac{\partial(1-Q)}{\partial u_i} \right)^2$$

A seguir temos as segundas derivadas de $L_F^*(\theta)$ e $L_C^*(\theta)$ respectivamente.

Observações de Tempo de Falha

$$\frac{\partial^2 L_F^*(\theta)}{\partial \mu^2} = - \sum_{i \in F} \frac{e^{qw_i}}{\sigma^2}$$

$$\frac{\partial^2 L_F^*(\theta)}{\partial \mu \partial \sigma} = \frac{1}{q\sigma^2} \left(r - \sum_{i \in F} e^{qw_i} - q \sum_{i \in F} w_i e^{qw_i} \right)$$

$$\frac{\partial^2 L_F^*(\theta)}{\partial \sigma^2} = \frac{1}{q\sigma^2} \left(rq + 2 \sum_{i \in F} w_i - 2 \sum_{i \in F} w_i e^{qw_i} - q \sum_{i \in F} w_i^2 e^{qw_i} \right).$$

Observações Censuradas

Com base em 3.17 e 3.18, temos que as expressões resultantes para dados censurados são,

$$\frac{\partial^2 L_C^*(\theta)}{\partial \mu^2} = \sum_{i \in C} \frac{\partial^2 \log S(y_i)}{\partial u_i^2} \left(\frac{\partial u_i}{\partial \mu} \right)^2 + \sum_{i \in C} \frac{\partial \log S(y_i)}{\partial u_i} \frac{\partial^2 u_i}{\partial \mu^2}$$

$$\frac{\partial^2 L_C^*(\theta)}{\partial \mu \partial \sigma} = \sum_{i \in C} \frac{\partial^2 \log S(y_i)}{\partial u_i^2} \frac{\partial u_i}{\partial \mu} \frac{\partial u_i}{\partial \sigma} + \sum_{i \in C} \frac{\partial \log S(y_i)}{\partial u_i} \frac{\partial^2 u_i}{\partial \mu \partial \sigma} \quad (3.23)$$

$$\frac{\partial^2 L_C^*(\theta)}{\partial \sigma^2} = \sum_{i \in C} \frac{\partial^2 \log S(y_i)}{\partial u_i^2} \left(\frac{\partial u_i}{\partial \sigma} \right)^2 + \sum_{i \in C} \frac{\partial \log S(y_i)}{\partial u_i} \frac{\partial^2 u_i}{\partial \sigma^2},$$

onde $S(y_i)$ é a função de sobrevivência para $q \neq 0$ dada em 3.16.

Temos então que as segundas derivadas de $L^*(\theta)$ para $q \neq 0$ são obtidas por,

$$\frac{\partial^2 L^*(\theta)}{\partial \theta_i \partial \theta_j} = \frac{\partial^2 L_F^*(\theta)}{\partial \theta_i \partial \theta_j} + \frac{\partial^2 L_C^*(\theta)}{\partial \theta_i \partial \theta_j}, \quad (3.24)$$

fazendo $\theta_1 = \mu$ e $\theta_2 = \sigma$ ($i, j = 1, 2$).

Com base nos resultados acima, construímos a matriz de informação observada $I(\hat{\theta})$ para este modelo, conforme segue,

$$I(\hat{\theta}) = - \left[\begin{array}{cc} \frac{\partial^2 L^*(\theta)}{\partial \mu^2} & \frac{\partial^2 L^*(\theta)}{\partial \mu \partial \sigma} \\ \frac{\partial^2 L^*(\theta)}{\partial \sigma \partial \mu} & \frac{\partial^2 L^*(\theta)}{\partial \sigma^2} \end{array} \right]_{\theta = \hat{\theta}},$$

onde as expressões das derivadas são dadas por 3.20 e 3.24, respectivamente para $q = 0$ e $q \neq 0$.

3.3.2 Procedimentos Inferenciais para o Modelo de Regressão

Consideremos aqui métodos da máxima verossimilhança para estimar e testar parâmetros do modelo de regressão baseado na distribuição l.g.g estendida. A base da teoria inferencial desenvolvida para este modelo, foi estudada por Farewell e Prentice (1977), que fornecem as derivadas da log-verossimilhança do modelo l.g.g. estendido, e Lawless(1982) que apresenta diversos resultados sobre estimação e testes dos parâmetros para o modelo com $q > 0$.

Sejam $(y_1, \mathbf{x}_1), \dots, (y_n, \mathbf{x}_n)$ n observações independentes, provenientes de uma distribuição dada por 2.34, onde y_i representa logaritmo do tempo de vida ou tempo até a censura, e $\mathbf{x}_i = (x_{i0}, x_{i1}, \dots, x_{is})$ o vetor de covariáveis associadas. Como antes, F e C denotam conjuntos de itens que obtiveram falha e censura, respectivamente. A função de log-verossimilhança para este modelo é da forma,

$$L^*(\theta; \mathbf{x}) = \begin{cases} \sum_{i \in F} \log \frac{1}{\sigma} f(z_i, q) + \sum_{i \in C} \log S(y_i | \mathbf{x}_i) & \text{se } q \neq 0 \\ \sum_{i \in F} \log \frac{1}{\sigma} \phi(z_i) + \sum_{i \in C} \log(1 - \Phi(z_i)) & \text{se } q = 0, \end{cases} \quad (3.25)$$

onde $\theta = (\beta, \sigma, q)$, $z_i = (y_i - \mathbf{x}_i\beta)/\sigma$, $f(z_i, q)$ e $S(y_i|\mathbf{x}_i)$ são respectivamente a densidade dada em 2.34 e a função de sobrevivência dada em 2.35, para $q \neq 0$, e as funções $\phi(z_i)$ e $\Phi(z_i)$ são a densidade e função de distribuição da normal padrão.

Quando $q = 0$ em 3.25, temos a log-verossimilhança do modelo de regressão normal para Y (ou o log-normal para T). Existem vários estudos do modelo normal quando temos amostras completas. Como exemplo temos, Draper e Smith (1981), Montgomery e Peck (1992) e Seber (1977). Para obtermos os estimadores de máxima verossimilhança de θ no caso geral ($q \in \mathfrak{R}$), usamos aqui o mesmo procedimento dado antes, ou seja, estimamos β e σ para valores fixos de q e, fazendo isto para uma faixa de valores de q , determinamos a *função de log-verossimilhança maximizada* para este modelo,

$$\log L_{\max}(q) = \log L(\tilde{\beta}(q), \tilde{\sigma}(q), q).$$

Obtendo o valor \hat{q} que maximiza esta função, temos os e.m.v.'s $\hat{\beta}(\hat{q})$, $\hat{\sigma}(\hat{q})$ e \hat{q} . Precisamos portanto, das derivadas da log-verossimilhança com respeito ao vetor $\beta = (\beta_0, \dots, \beta_{p-1})$, e a σ , apresentadas a seguir.

Derivadas de Primeira Ordem para Amostras Censuradas em Modelos de Regressão

(i) Para $q = 0$

As derivadas de $L^*(\theta; \mathbf{x})$ são dadas a seguir,

$$\frac{\partial L^*(\theta; \mathbf{x})}{\partial \beta_j} = \frac{1}{\sigma} \left[\sum_{i \in F} x_{ij} z_i + \sum_{i \in C} x_{ij} \frac{\phi(z_i)}{1 - \Phi(z_i)} \right] \quad \text{para } j = 0, 1, \dots, p-1$$

$$\frac{\partial L^*(\theta; \mathbf{x})}{\partial \sigma} = -\frac{r}{\sigma} + \frac{1}{\sigma} \left[\sum_{i \in F} z_i^2 + \sum_{i \in C} z_i \frac{\phi(z_i)}{1 - \Phi(z_i)} \right].$$

(ii) Para $q \neq 0$

Como antes denotamos por r o número de itens que falharam na amostras e fazemos

$m_i = q^{-2} \exp(qz_i)$. A log-verossimilhança deste modelo pode ser escrita como

$$L^*(\theta; \mathbf{x}) = r \left\{ \log |q| + q^{-2} \log q^{-2} - \log [\sigma \Gamma(q^{-2})] \right\} + \frac{1}{q} \sum_{i \in F} z_i - q^{-2} \sum_{i \in F} e^{qz_i} + \sum_{i \in C} \log S(y_i | \mathbf{x}_i).$$

onde $S(y_i | \mathbf{x}_i)$ denota a função de sobrevivência para $q \neq 0$, dada por

$$S(y_i | \mathbf{x}_i) = \begin{cases} Q(q^{-2}, m_i) & q > 0 \\ 1 - Q(q^{-2}, m_i) & q < 0, \end{cases}$$

sendo $Q(q^{-2}, m_i) = \int_{m_i}^{\infty} \frac{u^{q^{-2}-1} e^{-u}}{\Gamma(q^{-2})} du$, a integral gama incompleta.

Temos então que as primeiras derivadas desta função são dadas abaixo,

$$\frac{\partial L^*(\theta; \mathbf{x}_i)}{\partial \beta_j} = -\frac{1}{\sigma q} \left(\sum_{i \in F} x_{ij} - \sum_{i \in F} x_{ij} e^{qz_i} \right) + \frac{|q|}{\Gamma(q^{-2})\sigma} \sum_{i \in C} \frac{m_i^{q^{-2}} x_{ij} e^{-m_i}}{S(y_i | \mathbf{x}_i)}, \quad j = 0, \dots, p-1$$

$$\frac{\partial L^*(\theta; \mathbf{x})}{\partial \sigma} = -\frac{r}{\sigma} - \frac{1}{\sigma q} \left(\sum_{i \in F} z_i - \sum_{i \in F} z_i e^{qz_i} \right) + \frac{|q|}{\Gamma(q^{-2})\sigma} \sum_{i \in C} \frac{m_i^{q^{-2}} z_i e^{-m_i}}{S(y_i | \mathbf{x}_i)},$$

onde, para obter estes resultados usamos a regra da cadeia, e as derivadas $\frac{\partial m_i}{\partial \beta_j} =$

$$-\frac{qx_{ij}m_i}{\sigma} e \frac{\partial m_i}{\partial \sigma} = -\frac{qz_i m_i}{\sigma}.$$

Os estimadores de máxima verossimilhança $\tilde{\beta}(q)$ e $\tilde{\sigma}(q)$, são os valores que solucionam o sistema,

$$\left\{ \begin{array}{l} \frac{\partial L^*(\theta; \mathbf{x})}{\partial \beta_0} = 0 \\ \vdots \\ \frac{\partial L^*(\theta; \mathbf{x})}{\partial \beta_{p-1}} = 0 \\ \frac{\partial L^*(\theta; \mathbf{x})}{\partial \sigma} = 0. \end{array} \right. \quad (3.26)$$

Como antes, fornecemos as segundas derivadas da log-verossimilhança, tendo em vista que estas são necessárias na utilização do método de Newton Raphson para resolver 3.26.

Derivadas de Segunda Ordem para Amostras Censuradas em Modelos de Regressão

(i) Para $q = 0$

Considerando o modelo de regressão normal com censuras, as segundas derivadas da verossimilhança são,

$$\frac{\partial^2 L^*(\theta; \mathbf{x})}{\partial \beta_j \partial \beta_l} = -\frac{1}{\sigma^2} \sum_{i \in F} x_{ij} x_{il} - \frac{1}{\sigma^2} \sum_{i \in C} x_j x_l \lambda(z_i) \quad j, l = 0, \dots, p-1$$

$$\frac{\partial^2 L^*(\theta; \mathbf{x})}{\partial \beta_j \partial \sigma} = -\frac{2}{\sigma^2} \sum_{i \in F} x_{ij} z_i - \frac{1}{\sigma^2} \left(\sum_{i \in C} x_{ij} V(z_i) + \sum_{i \in C} x_{ij} z_i \lambda(z_i) \right) \quad j = 0, \dots, p-1$$

$$\frac{\partial^2 L^*(\theta; \mathbf{x})}{\partial \sigma^2} = \frac{r}{\sigma^2} - \frac{3}{\sigma^2} \sum_{i \in F} z_i^2 - \frac{1}{\sigma^2} \left(\sum_{i \in C} 2z_i V(z_i) + \sum_{i \in C} z_i^2 \lambda(z_i) \right).$$

onde $V(\cdot)$ e $\lambda(\cdot)$, são as funções definidas em 3.21.

(ii) Para $q \neq 0$

Mostramos aqui as segundas derivadas da log-verossimilhança para $q \neq 0$, seguindo uma notação similar à usada na seção anterior, para o modelo sem variáveis regressoras. Consideramos então que

$$L^*(\theta; \mathbf{x}) = L_F^*(\theta; \mathbf{x}) + L_C^*(\theta; \mathbf{x}) \quad (3.27)$$

onde $L^*(\theta; \mathbf{x})$ é dada por 3.25 para $q \neq 0$ e $L_F^*(\theta; \mathbf{x})$ e $L_C^*(\theta; \mathbf{x})$ representam respectivamente as parcela da log-verossimilhança para dados de falha e para dados censurados. A seguir temos as expressões usadas para o cálculo das segundas derivadas, pela regra da cadeia, onde novamente por simplificação fazemos $Q = Q(q^{-2}, m_i)$ e $m_i = q^{-2} \exp(q(y_i - \mathbf{x}_i\beta)/\sigma)$.

$$\frac{\partial m_i}{\partial \beta_j} = -\frac{qx_{ij}m_i}{\sigma}$$

$$\frac{\partial m_i}{\partial \sigma} = -\frac{qz_i m_i}{\sigma}$$

$$\frac{\partial^2 m_i}{\partial \beta_j \partial \beta_l} = \frac{q^2 x_{ij} x_{il} m_i}{\sigma^2}$$

$$\frac{\partial^2 m_i}{\partial \sigma^2} = \frac{q^2 m_i z_i^2}{\sigma^2} + \frac{2qz_i m_i}{\sigma^2}$$

$$\frac{\partial^2 m_i}{\partial \beta_j \partial \sigma} = \frac{qx_{ij} m_i}{\sigma^2} + \frac{q^2 x_{ij} m_i z_i}{\sigma^2}$$

$$\frac{\partial Q}{\partial m_i} = -\frac{\partial(1-Q)}{\partial m_i} = -\frac{m_i^{q^{-2}-1} e^{-m_i}}{\Gamma(q^{-2})}$$

$$\frac{\partial^2 Q}{\partial m_i^2} = -\frac{\partial^2(1-Q)}{\partial m_i^2} = \frac{m_i^{q^{-2}-2} e^{-m_i} (m_i - q^{-2} + 1)}{\Gamma(q^{-2})}$$

$$\frac{\partial \log Q}{\partial m_i} = \frac{1}{Q} \frac{\partial Q}{\partial m_i} \quad ; \quad \frac{\partial \log(1-Q)}{\partial m_i} = \frac{1}{(1-Q)} \frac{\partial(1-Q)}{\partial m_i}$$

$$\frac{\partial^2 \log Q}{\partial m_i^2} = \frac{1}{Q} \frac{\partial^2 Q}{\partial m_i^2} - \frac{1}{Q^2} \left(\frac{\partial Q}{\partial m_i} \right)^2$$

$$\frac{\partial^2 \log(1-Q)}{\partial m_i^2} = \frac{1}{(1-Q)} \frac{\partial^2(1-Q)}{\partial m_i^2} - \frac{1}{(1-Q)^2} \left(\frac{\partial(1-Q)}{\partial m_i} \right)^2$$

As derivadas parciais para cada parcela de 3.27 são dadas a seguir.

Observações de Tempo de Falha

$$\frac{\partial^2 L_F^*(\theta; \mathbf{x})}{\partial \beta_j \beta_l} = -\frac{1}{\sigma^2} \sum_{i \in F} x_{ij} x_{il} e^{qz_i} \quad j, l = 0, \dots, p-1$$

$$\frac{\partial^2 L_F^*(\theta; \mathbf{x})}{\partial \beta_j \partial \sigma} = \frac{1}{q\sigma^2} \left(\sum_{i \in F} x_{ij} - \sum_{i \in F} x_{ij} e^{qz_i} - q \sum_{i \in F} x_{ij} z_i e^{qz_i} \right) \quad j = 0, \dots, p-1$$

$$\frac{\partial^2 L_F^*(\theta; \mathbf{x})}{\partial \sigma^2} = \frac{1}{q\sigma^2} \left(rq + 2 \sum_{i \in F} z_i - 2 \sum_{i \in F} z_i e^{qz_i} - q \sum_{i \in F} z_i^2 e^{qz_i} \right).$$

Observações Censuradas

Consideramos a mesma notação adotada para as primeiras derivadas, temos como

resultado as expressões abaixo.

$$\frac{\partial^2 L_C^*(\theta; \mathbf{x})}{\partial \beta_j \partial \beta_l} = \sum_{i \in C} \frac{\partial^2 \log S(y_i \setminus x_i)}{\partial m_i^2} \frac{\partial m_i}{\partial \beta_j} \frac{\partial m_i}{\partial \beta_l} + \sum_{i \in C} \frac{\partial \log S(y_i \setminus x_i)}{\partial m_i} \frac{\partial^2 m_i}{\partial \beta_j \partial \beta_l} \quad j, l = 0, \dots, p-1$$

$$\frac{\partial^2 L_C^*(\theta; \mathbf{x})}{\partial \beta_j \partial \sigma} = \sum_{i \in C} \frac{\partial^2 \log S(y_i \setminus x_i)}{\partial m_i^2} \frac{\partial m_i}{\partial \beta_j} \frac{\partial m_i}{\partial \sigma} + \sum_{i \in C} \frac{\partial \log S(y_i \setminus x_i)}{\partial m_i} \frac{\partial^2 m_i}{\partial \beta_j \partial \sigma} \quad j = 0, \dots, p-1$$

$$\frac{\partial^2 L_C^*(\theta; \mathbf{x})}{\partial \sigma^2} = \sum_{i \in C} \frac{\partial^2 \log S(y_i \setminus x_i)}{\partial m_i^2} \left(\frac{\partial m_i}{\partial \sigma} \right)^2 + \sum_{i \in C} \frac{\partial \log S(y_i \setminus x_i)}{\partial m_i} \frac{\partial^2 m_i}{\partial \sigma^2},$$

Temos portanto que as derivadas para a função $L^*(\theta; \mathbf{x})$ são dadas por

$$\frac{\partial^2 L^*(\theta; \mathbf{x})}{\partial \theta_i \partial \theta_j} = \frac{\partial^2 L_F^*(\theta; \mathbf{x})}{\partial \theta_i \partial \theta_j} + \frac{\partial^2 L_C^*(\theta; \mathbf{x})}{\partial \theta_i \partial \theta_j},$$

onde usamos a notação $\theta_i = \beta_{i-1}$, para $i = 1, \dots, p$ e $\theta_{p+1} = \sigma$, ($i, j = 1, \dots, p+1$).

Podemos agora construir a matriz de informação observada 1.21 para este modelo de regressão, conforme segue,

$$I(\hat{\theta}) = - \begin{bmatrix} \frac{\partial^2 L^*(\theta; \mathbf{x})}{\partial \beta_0^2} & \frac{\partial^2 L^*(\theta; \mathbf{x})}{\partial \beta_0 \partial \beta_1} & \dots & \frac{\partial^2 L^*(\theta; \mathbf{x})}{\partial \beta_0 \partial \sigma} \\ & \frac{\partial^2 L^*(\theta; \mathbf{x})}{\partial \beta_1^2} & \dots & \frac{\partial^2 L^*(\theta; \mathbf{x})}{\partial \beta_1 \partial \sigma} \\ & & \ddots & \vdots \\ \text{simétrica} & & & \frac{\partial^2 L^*(\theta; \mathbf{x})}{\partial \sigma^2} \end{bmatrix}_{\theta = \hat{\theta}}$$

Intervalos de Confiança e Testes

Tratamos agora de alguns procedimentos inferenciais sob o modelo de regressão l.g.g. estendido em amostras censuradas que representam uma generalização dos resultados apresentados na seção 3.2.2, que trata apenas de amostras completas.

Descrevemos métodos baseados na teoria assintótica da verossimilhança para elaboração de testes e intervalos de confiança para os parâmetros do modelo, que implicam na indicação ou não de um determinado modelo paramétrico dentre os principais modelos que compõem a família l.g.g. estendida. Consideramos o método assintótico da razão de verossimilhança para testar o parâmetro de forma (q) do modelo de regressão, com base no que foi dado em Farewell e Prentice (1977), e Lawless (1982) que usou o modelo l.g.g. apenas para $q \geq 0$. Similarmente ao que foi dado antes, definimos a verossimilhança relativa maximizada conforme segue

$$R_{\max}(q) = \frac{L_{\max}(q)}{L(\hat{\beta}, \hat{\sigma}, \hat{q})}, \quad (3.28)$$

onde $L_{\max}(q) = L(\tilde{\beta}(q), \tilde{\sigma}(q), q)$ é o máximo da verossimilhança para o dado valor de q , sendo $\tilde{\beta}(q) = (\tilde{\beta}_0(q), \tilde{\beta}_1(q), \dots, \tilde{\beta}_{p-1}(q))$. Utilizamos esta estatística para discriminar entre alguns modelos, conforme dado a seguir.

Teste para os Modelos Weibull e Log- Normal Sob o modelo l.g.g. estendido, testar a validade dos modelos Weibull e log-normal, é equivalente a testar o parâmetro de forma q . Com hipóteses do tipo $H_0 : q = q_0$ versus $H_1 : q \neq q_0$, podemos usar a estatística da razão de verossimilhança, que é dada por

$$\Lambda = -2 \log R_{\max}(q_0),$$

e para grandes amostras, tem distribuição aproximadamente χ^2 com 1 grau de liberdade, sob a hipótese H_0 .

- **Teste para o Modelo de Regressão Weibull (Valor Extremo)**

Se consideramos $q_0 = 1$ no teste dado acima, estaremos testando a adequação do modelo de regressão Weibull para os tempos de vida. Usamos especificamente a estatística

$$\Lambda = -2 \log \left\{ \frac{L(\tilde{\beta}(1), \tilde{\sigma}(1), 1)}{L(\hat{\beta}, \hat{\sigma}, \hat{q})} \right\},$$

onde $L(\tilde{\beta}(1), \tilde{\sigma}(1), 1)$ corresponde ao máximo de $L(\beta(1), \sigma(1), 1)$, dado por

$$L(\beta(1), \sigma(1), 1) = \prod_{i \in F} \frac{1}{\sigma} \exp[z_i - e^{z_i}] \prod_{i \in C} e^{-e^{z_i}}$$

sendo $z_i = (y_i - \mathbf{x}_i \beta) / \sigma$.

• **Teste para o Modelo de Regressão Log-normal (Normal)**

Considerar as hipóteses $H_0 : q = 0$ versus $H_1 = q \neq 0$, equivale a testar se os dados admitem o modelo de regressão log-normal para T, e a estatística usada para o teste é

$$\Lambda = -2 \log \left\{ \frac{L(\tilde{\beta}(0), \tilde{\sigma}(0), 0)}{L(\hat{\beta}, \hat{\sigma}, \hat{q})} \right\},$$

onde $\tilde{\beta}(0)$ e $\tilde{\sigma}(0)$ são os valores que maximizam a verossimilhança condicionada ao valor de $q = 0$, e

$$L(\beta(0), \sigma(0), 0) = \prod_{i \in F} \frac{1}{\sigma} \exp(z_i^2/2) \prod_{i \in C} (1 - \Phi(z_i)),$$

sendo $z_i = (y_i - \mathbf{x}_i \beta) / \sigma$ e $\Phi(\cdot)$ a função de distribuição da normal padrão.

Da mesma forma dada antes, usamos a estatística Λ , que tem assintoticamente distribuição χ^2 com um grau de liberdade, para testar a adequação do modelo.

• **Intervalo de Confiança para q**

Uma maneira de checar simultaneamente a adequação dos modelo Weibull e log-normal é construir um intervalo de confiança para q, e verificar se os valores $q = 1$ e

$q = 0$ pertencem a este intervalo. Podemos construir um intervalo de confiança aproximado para q através da estatística $\Lambda = -2 \log R_{\max}(q_0)$. Para isso obtemos os valores de q_0 para o qual Λ seja menor que o ξ -ésimo quantil de um v.a. χ^2 com 1 grau de liberdade. O conjunto de valores satisfazendo esta condição fornece um intervalo de confiança aproximado de nível $(1 - \xi)100\%$ para o parâmetro q .

Teste para o Modelo de Regressão Exponencial Testar a adequacidade do modelo de regressão exponencial aos dados corresponde a testar se, sob a parametrização original da d.g.g. 2.1, $b = k = 1$ (trocando β por b para evitar problemas de notação). Equivalentemente, sob o modelo reparametrizado dado em 2.34 para $q > 0$, testamos as hipóteses

$$H_0 : \left\{ \begin{pmatrix} \sigma \\ q \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \text{ versus } H_1 : \left\{ \begin{pmatrix} \sigma \\ q \end{pmatrix} \neq \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}.$$

A estatística usada para testar H_0 é

$$\Lambda = -2 \log \left\{ \frac{L(\tilde{\beta}(1), 1, 1)}{L(\hat{\beta}, \hat{\sigma}, \hat{q})} \right\}, \quad (3.29)$$

onde $L(\tilde{\beta}(1), 1, 1)$ é o máximo de $L(\beta, \sigma, q)$ sob a hipótese H_0 . A função de verossimilhança $L(\beta(1), 1, 1)$ é dada por

$$L(\beta(1), 1, 1) = \prod_{i \in F} \exp[(y_i - \mathbf{x}_i \beta) - \exp(y_i - \mathbf{x}_i \beta)] \prod_{i \in C} \exp(-\exp(y_i - \mathbf{x}_i \beta)).$$

Sob H_0 a estatística Λ tem assintoticamente distribuição χ^2 com 2 graus de liberdade.

Testes e Intervalos para os Coeficientes de Regressão Em geral as hipóteses de interesse sobre $\beta' = (\beta_0, \beta_1, \dots, \beta_{p-1})$ são da forma $H_0 : \beta_1 = \beta_1^0$, onde β_1 é um vetor $s \times 1$

($s < p$), representando uma partição de $\beta' = (\beta'_1, \beta'_2)$ e β_1^0 é um vetor dado. Podemos usar a estatística da razão de verossimilhança para testar H_0 contra $H_1 : \beta_1 \neq \beta_1^0$. dada por

$$\Lambda = -2 \log \left\{ \frac{L(\beta_1^0, \tilde{\beta}_2(\tilde{q}), \tilde{\sigma}(\tilde{q}), \tilde{q})}{L(\hat{\beta}, \hat{\sigma}, \hat{q})} \right\}, \quad (3.30)$$

onde $\tilde{\beta}_2(\tilde{q}), \tilde{\sigma}(\tilde{q})$ e \tilde{q} são os estimadores de máxima verossimilhança de β_2, σ e q , sob a hipótese H_0 , e $\hat{\sigma}, \hat{\beta} = (\hat{\beta}'_1, \hat{\beta}'_2)'$ e \hat{q} são os e.m.v.'s de σ, β e q para o modelo completo. Como antes usamos a distribuição assintótica de Λ , que aqui é χ^2 com s graus de liberdade (a dimensão de β_1), para calcular níveis de significância aproximados.

Um outro procedimento que pode ser adotado para testar as hipóteses dadas acima é baseada na estatística de Wald dada em 1.18. Consideramos para isto a seguinte estatística

$$\Lambda_1 = (\hat{\beta}_1 - \beta_1^0)' C_{11}^{-1} (\hat{\beta}_1 - \beta_1^0), \quad (3.31)$$

sendo que $C = I^{-1}(\hat{\beta}, \hat{\sigma})$ (o inverso da matriz de informação observada com respeito aos parâmetros β e σ), é particionada como

$$\begin{pmatrix} C_{11} & C_{12} \\ C'_{12} & C_{22} \end{pmatrix},$$

onde C_{11} é a matriz de covariâncias assintótica para β_1 . Sob a hipótese H_0 , Λ_1 tem distribuição χ^2 com s graus de liberdade.

Conforme discutido no capítulo 1, embora as estatísticas 3.30 e 3.31 sejam assintoticamente equivalentes, quando trabalhamos com modelos de vida a estatística da razão de verossimilhança em geral é mais indicada. Contudo o seu uso requer a maximização da função de verossimilhança sob a hipótese H_0 , o que representa maior dificuldade computacional. Por esta razão nas aplicações, a estatística 3.31 é usada com mais frequência para inferir sobre os coeficientes de regressão.

Capítulo 4

Ajuste do Modelo Gama Generalizada a Dados de Vida

Neste capítulo temos como objetivo a aplicação de alguns dos procedimentos estudados nos capítulos anteriores, através da análise de 5 conjuntos de dados. As análises seguiram principalmente o enfoque da seleção de um modelo paramétrico apropriado aos dados, dentre os principais modelos que compõem a família l.g.g. estendida (2.33, 2.34). Os métodos computacionais utilizados para realizar tais procedimentos são descritos em detalhes no apêndice D. Apresentamos no final, alguns comentários sobre os resultados obtidos.

4.1 Análise de Dados Reais

Apresentamos aqui aplicações do modelo l.g.g. estendido a alguns conjuntos de dados que já haviam sido estudados, sob um dos modelos que representam casos particulares desta família. A finalidade é avaliar graficamente e pelo teste da razão de verossimilhança a adequadade do modelo adotado.

Analisamos inicialmente se a suposição de uma distribuição log-gama generalizada parece ser adequada para o log dos tempos, através de um gráfico que compara a curva da distribuição suposta para o resíduo, com o gráfico de Kaplan-Meier dos resíduos estimados (seção 1.5.1), conforme sugerido por Lawless (1982).

Estudamos 5 conjuntos de dados, sendo os dois primeiros dados da engenharia, que tratam de testes acelerados de vida (seção 1.4). Os exemplos 3 e 4 são estudos médicos sobre tempo de vida de pacientes portadores de determinada doença. Estes exemplos mostram que o modelo l.g.g. estendido pode ser uma alternativa paramétrica para a análise de dados da medicina, que em geral é feita com base em modelos de riscos proporcionais. O último conjunto de dados é considerado basicamente para ilustrar a aplicabilidade dos métodos a uma variável positiva que não representa especificamente "tempo até a falha". Na maioria dos casos não fornecemos intervalos de confiança para o parâmetro de forma (q), conforme sugerido anteriormente, devido a problemas de convergência do método iterativo para valores extremos deste parâmetro. Contudo isto não dificultou a análise, tendo em vista que foi possível realizar os testes de hipóteses de interesse sobre este parâmetro.

4.1.1 Exemplo 1- Teste de Vida de Isolantes

Os dados da tabela 4.1 são resultados de um teste acelerado de vida para isolantes de sistemas de uma nova classe II e referem-se ao tempo em horas que os isolantes levaram para tornar-se defeituosos. Os testes foram realizados em pequenos motores, a temperaturas elevadas. Dez motores foram colocados para trabalhar a temperaturas de 100, 220, 240 e 260 °C e inspecionados periodicamente para detectar a ocorrência de falha. Os tempos da Tabela 4.1 referem-se à média entre o tempo em que a falha foi observada e o tempo da inspeção anterior. Os tempos entre checagem foram 7, 4, 2, e 2 dias para as respectivas temperaturas. Para estes dados, temos 7 censuras, correspondendo ao tempo de vida dos isolantes que ainda estavam em boas condições no final do experimento.

Estes dados foram analisados por Nelson (1990), que considerou o modelo de regressão log-normal para o ajuste dos tempos de vida. Vamos assumir o modelo de regressão log-gama estendido para o log dos tempos, para checar a adequação deste ajuste.

Transformando a temperatura (variável independente) para assumir a relação Arrhenius, conforme sugerido na literatura de engenharia, usamos o modelo

190	220	240	260
7228	1764	1175	1128
7228	2436	1521	1464
7228	2436	1569	1512
8448	2436 +	1617	1608
9167	2436	1665	1632 +
9167	2436	1665	1632 +
9167	3108	1713	1632 +
9167	3108	1761	1632 +
10511	3108	1831 +	1632 +
10511	3108	1953	1896

Tabela 4.1: Tempo de vida de isolantes (horas) para cada temperatura, em graus centígrados, com + indicando censura (Nelson, 1990).

$$Y_i = \beta_0 + \beta_1 x_i + \sigma Z_i \quad i = 1, 2, \dots, 40, \quad (4.1)$$

onde $Y_i = \log(\text{tempo de vida do isolante } i)$, $x_i = 1000/(\text{temperatura } ^\circ\text{C do } i\text{-ésimo motor} + 273.16)$,¹ β_0 , β_1 , e σ são parâmetros desconhecidos e Z_i tem distribuição dada em 2.34, com parâmetro q .

A Figura 4-1 mostra respectivamente a relação entre o log dos tempos e a temperatura, e a relação entre o log dos tempos e x_i . Apesar de, aparentemente, a transformação na temperatura não favorecer ao relacionamento linear desta variável transformada com o log dos tempos, optamos pelo seu uso.

Considerando a log-verossimilhança dada em 3.25 temos que o estimador de máxima verossimilhança do parâmetro $\theta = (\beta, \sigma, q)$ é obtido pela resolução do sistema de equações $\frac{\partial L^*(\theta)}{\partial \beta} = 0$ e $\frac{\partial L^*(\theta)}{\partial \sigma} = 0$ para alguns valores fixos de q . Isto é feito através do programa desenvolvido no software S-Plus para $q \neq 0$, e do SAS, versão 6.08, módulo Lifereg para $q = 0$ (ver apêndice D). A Tabela 4.2 mostra algumas das etapas realizadas para obter os estimadores de máxima verossimilhança de β e σ para cada valor de q dado, onde cada linha representa o resultado de um processo iterativo de maximização.

¹A temperatura absoluta Kelvin é igual a temperatura em graus centígrados mais 273.16

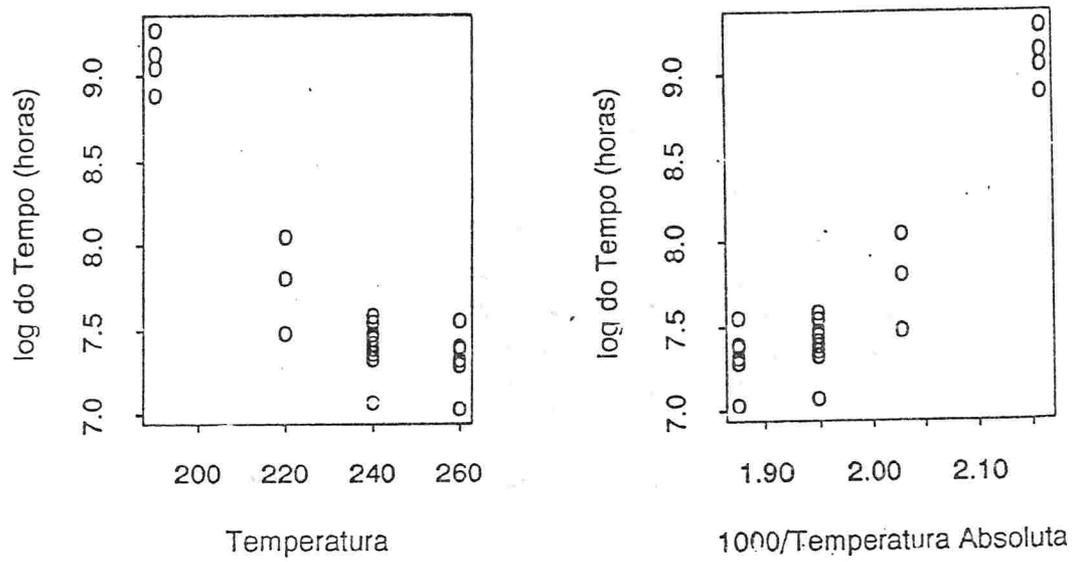


Figura 4-1: À esquerda, o log dos tempos de vida de isolantes por nível de temperatura, e à direita, por nível transformado de temperatura (relacionamento Arrhenius).

q	$\hat{\beta}_0(q)$	$\hat{\beta}_1(q)$	$\hat{\sigma}(q)$	$\log L_{\max}(q)$	$R_{\max}(q)$
-1.00	-4.8511	6.3305	0.2835	-12.0352	0.1568
-0.76	-4.6734	6.2577	0.2848	-11.3581	0.3086
-0.52	-4.4988	6.1864	0.2847	-10.8421	0.5168
-0.40	-4.4177	6.1538	0.2839	-10.6481	0.6277
-0.16	-4.2714	6.0963	0.2806	-10.3491	0.8464
-0.13	-4.2546	6.0899	0.2801	-10.3431	0.8515
0.00	-4.1860	6.0640	0.2772	-10.2697	0.9164
0.17	-4.1059	6.0348	0.2723	-10.1823	1.0000
0.20	-4.0926	6.0300	0.2714	-10.2327	0.9508
0.32	-4.0431	6.0129	0.2670	-10.2655	0.9202
0.44	-3.9977	5.9977	0.2621	-10.3255	0.8666
0.68	-3.9202	5.9737	0.2505	-10.5438	0.6967
0.80	-3.8873	5.9645	0.2438	-10.6926	0.6003
1.00	-3.8415	5.9535	0.2317	-10.2317	0.4512

Tabela 4.2: Etapas para a estimação dos parâmetros para dados do teste de vida de isolante de motores.

Assim, as estimativas globais são obtidas para $q = 0.17$, e os resultados são:

$$\hat{\beta}_0 = -4.1059 (0.8386) \quad \hat{\beta}_1 = 6.0348 (0.4158) \quad \hat{\sigma} = 0.2723 (0.0373)$$
$$\log L_{\max}(0.17) = -10.1824.$$

Cada valor entre parênteses ao lado das estimativas representa a raiz quadrada do correspondente elemento de $I^{-1}(\hat{\beta}, \hat{\sigma})$, onde $I(\hat{\beta}, \hat{\sigma})$ é a matriz de informação observada 1.21, condicionada em $q = 0.17$. Esta é o negativo da matriz de 2ª's derivadas parciais com respeito a β e σ , somada para todos os pontos da amostra e avaliada nos máximos estimados ($\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)$ e $\hat{\sigma}$). Se admitimos a normalidade assintótica dos estimadores para a amostra estudada, $I^{-1}(\hat{\beta}, \hat{\sigma})$ corresponde à estimativa da matriz de covariâncias assintóticas, e pode ser utilizada para as inferências sobre estes parâmetros.

Na Figura 4-2 temos o gráfico Kaplan-Meier dos resíduos estimado sob o modelo 4.1, juntamente com a curva de sobrevivência da log-gama assumindo $q = 0.17$, mostrando que o modelo l.g.g. parece ajustar bem os dados. Para estudar a possibilidade da utilização do modelo log-normal ($q = 0$), temos na Figura 4-3 o contorno da função $R_{\max}(q)$, que mostra a existência de fortes evidências favoráveis ao ajuste deste modelo. A estatística da razão de verossimilhança para testar $H_0 : q = 0$ é $-2 \log R_{\max}(0) = -2 \log(0.9164) = 0.1747$, e indica a aceitação do modelo log-normal a um nível de significância 0.05 ($\chi^2_{(1,0.05)} = 3.8441$). Testando o modelo Weibull ($H_0 : q = 1$) temos $-2 \log R_{\max}(1) = -2 \log(0.4512) = 1.5917$, e este resultado implica que o modelo Weibull também é estatisticamente favorável para ajustar os dados. Contudo a Figura 4-3_x mostra que o ajuste log-normal é bem mais adequado. Acreditamos então que o modelo usado por Nelson (1990) é apropriado para o ajuste.

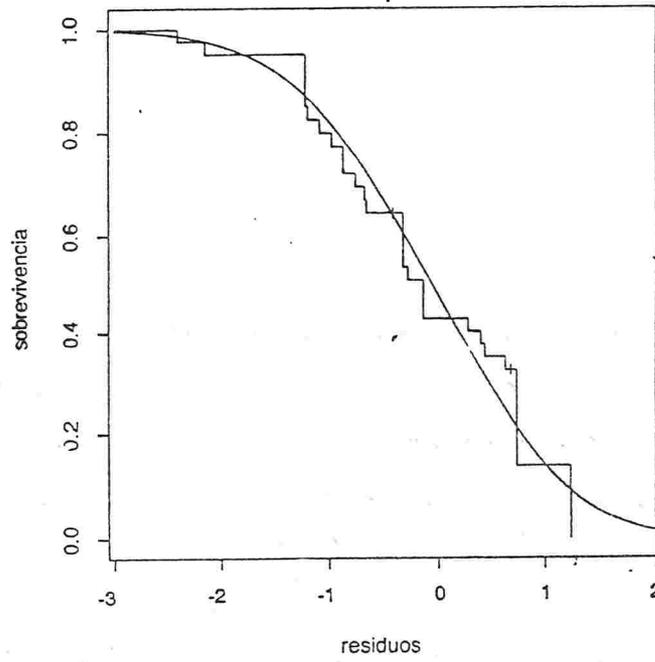


Figura 4-2: Curva de sobrevivência estimada para a distribuição log-gama ($q = 0.17$) e gráfico de Kaplan-Meier para resíduos estimados. Exemplo 1 - Isolantes de motores.

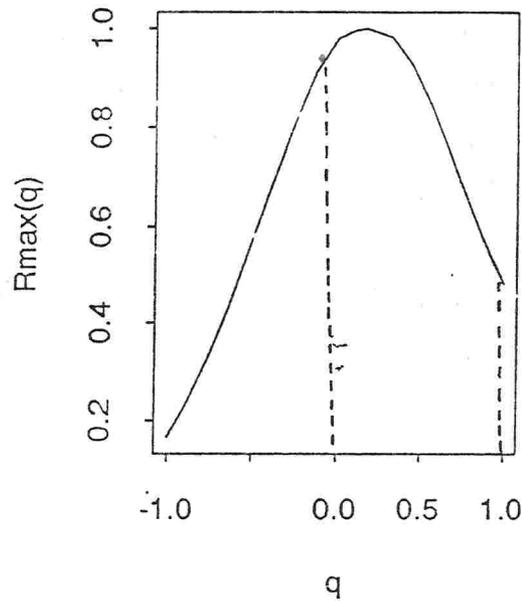


Figura 4-3: Contorno da verossimilhança relativa maximizada $R_{\max}(q)$, para dados de teste de vida de isolantes de motores.

Temperatura (C)	Volt. Aplicada (KW)			
	200	250	300	350
170	439	572	315	258
	904	690	315	258
	1092	904	439	347
	1105	1090	628	588
180	959	216	241	241
	1065	315	315	241
	1065	455	332	435
	1087	473	380	455

Tabela 4.3: Resultados do teste de vida de capacitores de vidro (Zelen, 1959).

4.1.2 Exemplo 2 - Teste de Vida de Capacitores

Em Zelen (1959) temos o estudo dos resultados de um teste de vida de capacitores² sujeitos a variações de temperatura e voltagem, onde a suposição básica é que o calor e a alta voltagem atuam como estresses, reduzindo a vida dos capacitores. Os testes foram aplicados a oito combinações de temperatura e voltagem, envolvendo as voltagens de 200, 250, 300 e 350 kilowatts, e temperaturas de 170 e 180° C. Para cada combinação aplicou-se um teste de vida com censura tipo II, ou seja, oito capacitores foram simultaneamente colocados em teste e avaliados até a ocorrência de quatro falhas. Os dados referem-se apenas aos tempos de falha, e estão sumarizados na tabela 4.3.

Zelen (1959) analisou estes dados conforme um experimento fatorial e assumiu a distribuição exponencial para os tempos de vida. Lawless e Singhal (1980) estudaram também estes dados sob a suposição de um modelo exponencial. Tratamos aqui o log dos tempos sob o modelo de regressão l.g.g. estendido, para avaliar se a modelagem exponencial é adequada. Assumimos então que o log dos tempos de vida tem distribuição dada por 2.33 e 2.34 com parâmetro de forma q , de escala σ e de locação $\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$ com $i = 1, 2, \dots, 32$, onde as variáveis regressoras são definidas como

$$x_1 = 1000 / (\text{temperatura} + 273,16)$$

$$x_2 = -\log(\text{voltagem}).$$

²Capacitores são componentes eletrônicos de um sistema que tem como função reter energia potencial.

q	$\hat{\beta}_0(q)$	$\hat{\beta}_1(q)$	$\hat{\beta}_2(q)$	$\hat{\sigma}(q)$	$\log L_{\max}(q)$	$R_{\max}(q)$
-0.60	5.0989	4.8341	1.7510	0.3672	-14.0107	0.0789
-0.48	5.5758	4.7378	1.7939	0.3659	-13.5273	0.1280
-0.24	6.3045	4.5943	1.8591	0.3604	-12.6805	0.2986
0.00	6.7418	4.5082	1.8954	0.3525	-12.0367	0.5684
0.17	6.8903	4.4767	1.9040	0.3460	-11.7248	0.7764
0.24	6.9171	4.4694	1.9037	0.3432	-11.6503	0.8365
0.45	6.8938	4.4655	1.8915	0.3342	-11.5223	0.9507
0.49	6.8738	4.4674	1.8875	0.3324	-11.4754	0.9963
0.52	6.8559	4.4695	1.8842	0.3311	-11.4718	1.0000
0.57	6.8210	4.4738	1.8782	0.3287	-11.4770	0.9947
1.00	6.3166	4.5548	1.8076	0.3044	-11.8060	0.7159
1.56	5.3502	4.7785	1.7084	0.26150	-12.4887	0.3617
1.92	4.6463	5.0059	1.6648	0.23157	-12.7556	0.2770

Tabela 4.4: Etapas para estimação dos parâmetros referentes a dados de teste de vida de capacitores.

As transformações nos estresses implicam na suposição da relação Arrhenius para a temperatura e poder inverso para a voltagem, conforme sugerido em literatura de engenharia (ver seção 1.2.6).

A estimação dos parâmetros é feita pelo mesmo procedimento dado antes, onde escolhemos uma faixa de valores convenientes para q e realizamos um processo iterativo de maximização para cada um destes valores. Algumas etapas deste procedimento são dadas na Tabela 4.4.

Encontramos que $\log L_{\max}(q)$ é maximizado aproximadamente para $q = 0.52$ e os resultados são,

$$\begin{aligned} \hat{\beta}_0 &= 6.8559(5.4090) & \hat{\beta}_1 &= 4.4695(2.3521) & \hat{\beta}_2 &= 1.8842(0.2665) \\ \hat{\sigma} &= 0.3311(0.0429) & \log L_{\max}(0.52) &= -11.4718. \end{aligned}$$

A Figura 4-4, que apresenta a função de sobrevivência empírica para os resíduos estimados ($z_i = (y_i - \mathbf{x}_i\hat{\beta})/\hat{\sigma}$), juntamente com a curva de sobrevivência da distribuição log-gama, com base no e.m.v. $\hat{q} = 0.52$, mostra que o modelo parece apresentar um bom ajuste para os dados.

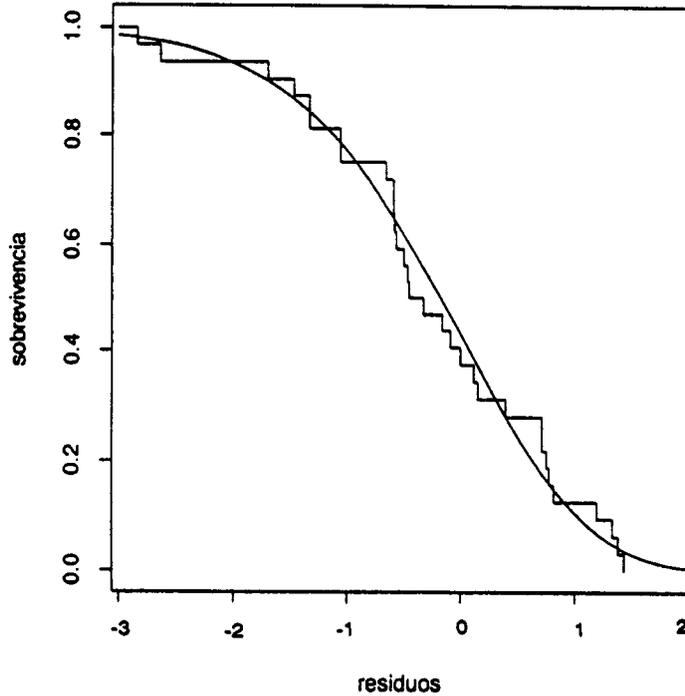


Figura 4-4: Curva de sobrevivência estimada para a distribuição log-gama ($q = 0.52$) e função de sobrevivência empírica para resíduos estimados. Exemplo 2 - Capacitores.

A Figura 4-5 mostra o contorno da verossimilhança relativa maximizada $R_{\max}(q) = L_{\max}(q)/L_{\max}(0.52)$. Esta curva que indica os valores mais prováveis de q , mostra evidências favoráveis tanto para o modelo log-normal ($q = 0$), como para o modelo Weibull ($q = 1$). As estatísticas da razão de verossimilhança para testar estes modelos tem distribuição $\chi^2_{(1)}$ e resultam $-2 \log R_{\max}(0) = 0.669$ e $-2 \log R_{\max}(1) = 1.130$, indicando que a um nível de significância de 5%, estes modelos são adequados para ajustar os dados. Para testar o ajuste exponencial, a estatística da razão de verossimilhança, conforme dada em 3.29, tem distribuição $\chi^2_{(2)}$ e como resultado encontramos $-2 \log R_{\max}(H_0) = 4.830$. Isto mostra que a um nível de significância 5%, não existem evidências contra este modelo ($\chi^2_{(2,0.05)} = 5.99$). Temos então, de acordo com o critério aqui adotado, que o modelo de regressão exponencial usado por Zelen (1959) parece ser apropriado para os dados em estudo.

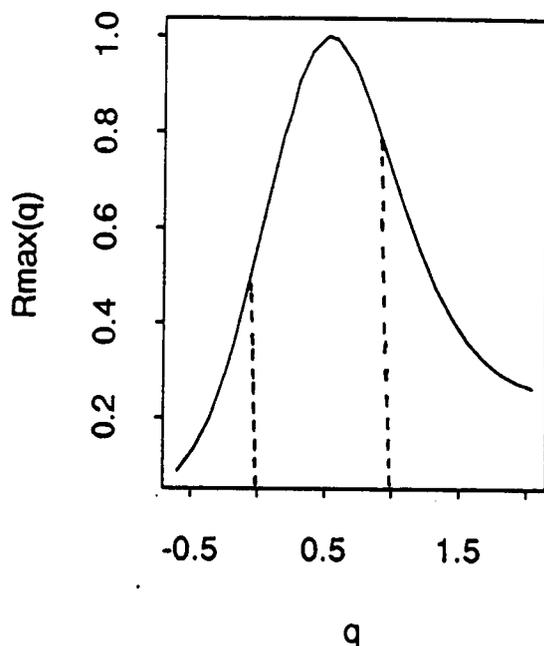


Figura 4-5: Contorno da verossimilhança relativa maximizada $R_{\max}(q)$, para dados de teste de vida de capacitores.

4.1.3 Exemplo 3 - Sobrevivência de Portadores de Mieloma

Estudando um procedimento de seleção de variáveis regressoras em dados de sobrevivência, Krall, et al. (1975) analisaram os tempos de sobrevivência de 65 pacientes com um determinado tipo de cancer na medula óssea chamado **mieloma múltipla**. Foram também consideradas diversas variáveis associadas a cada indivíduo. A Tabela 4.5 mostra os dados conforme apresentados em Lawless (1982), que considera apenas 5 destas variáveis. Os tempos de vida são dados em meses, contados a partir do dia do diagnóstico, e as ocorrências de censura são seguidas de um asterisco (*). As variáveis regressoras consideradas são dadas abaixo e referem-se a informações obtidas na época do diagnóstico.

Notação	Variáveis (medidas no dia do diagnóstico)
x_1	Logaritimo natural de uma química do sangue
x_2	Medida da hemoglobina
x_3	Idade
x_4	Sexo
x_5	Soro cálcio

Krall et al. (1975) assumiram o modelo de regressão exponencial para realizar a análise destes dados. Utilizamos aqui o modelo de regressão l.g.g. estendido para o log dos tempos, para checar a adequação deste ajuste. Assumimos que o log dos tempos de vida Y_i para $i = 1, 2, \dots, 65$ tem distribuição l.g.g. estendida, podendo ser representada por

$$Y_i = \mathbf{x}_i\beta + \sigma Z_i,$$

onde $\mathbf{x}_i = (x_{i0}, x_{i1}, \dots, x_{i5})$, é o vetor de variáveis regressoras do indivíduo i , sendo $x_{i0} \equiv 1$, $\beta = (\beta_0, \beta_1, \dots, \beta_5)$ e σ são parâmetros do modelo e Z_i tem distribuição log-gama dada em 2.34, com parâmetro de forma q . Para estimar os parâmetros, consideramos o mesmo procedimento dado antes e temos na Tabela 4.6 alguns resultados dos estimadores de β e σ para valores fixos de q .

Como resultado temos os estimadores de máxima verossimilhança e a função de máxima verossimilhança, observados para $q = -0.6$, que são mostrados abaixo,

$$\begin{aligned} \hat{\beta}_0 &= 4.6908 (1.30434) & \hat{\beta}_3 &= 0.0027 (0.01256) & \hat{\sigma} &= 0.9706 (0.1031) \\ \hat{\beta}_1 &= -1.5692 (0.38519) & \hat{\beta}_4 &= 0.2550 (0.27707) \\ \hat{\beta}_2 &= 0.15410 (0.04980) & \hat{\beta}_5 &= -0.1664 (0.0667) \\ \log L_{\max}(-0.6) &= -76.7443. \end{aligned}$$

A Figura 4-6 apresenta o gráfico da verossimilhança relativa maximizada em função de q e este mostra evidências favoráveis ao modelo log-normal para os tempos de vida, e

tempo	x_1	x_2	x_3	x_4	x_5	tempo	x_1	x_2	x_3	x_4	x_5
1	2.218	9.4	67	0	10	35	1.114	7.0	48	0	10
1	1.940	12.0	38	0	18	37	1.602	11.0	63	0	9
2	1.519	9.8	81	0	15	41	1.000	10.2	69	0	10
2	1.748	11.3	75	0	12	42	1.146	5.0	70	1	9
2	1.301	5.1	57	0	9	51	1.568	7.7	74	0	13
3	1.544	6.7	46	1	10	52	1.000	10.1	60	1	10
5	2.236	10.1	50	1	9	54	1.255	9.0	49	0	10
5	1.681	6.5	74	0	9	58	1.204	12.1	42	1	10
6	1.362	9.0	77	0	8	66	1.447	6.6	59	0	9
6	2.114	10.2	70	1	8	67	1.322	12.8	52	0	10
6	1.114	9.7	60	0	10	88	1.176	10.6	47	1	9
6	1.415	10.4	67	1	8	89	1.322	14.0	63	0	9
7	1.978	9.5	48	0	10	92	1.431	11.0	58	1	11
7	1.041	5.1	61	1	10	4*	1.945	10.2	59	0	10
7	1.176	11.4	53	1	13	4*	1.924	10.0	49	1	13
9	1.724	8.2	55	0	12	7*	1.114	12.4	48	1	10
11	1.114	14.0	61	0	10	7*	1.532	10.2	81	0	11
11	1.230	12.0	43	0	9	8*	1.079	9.9	57	1	8
11	1.301	13.2	65	0	10	12*	1.146	11.6	46	1	7
11	1.508	7.5	70	0	12	11*	1.613	14.0	60	0	9
11	1.079	9.6	51	1	9	12*	1.398	8.8	66	1	9
13	0.778	5.5	60	1	10	13*	1.663	4.9	71	1	9
14	1.398	14.6	66	0	10	16*	1.146	13.0	55	0	9
15	1.602	10.6	70	0	11	19*	1.322	13.0	59	1	10
16	1.342	9.0	48	0	10	19*	1.322	10.8	69	1	10
16	1.322	8.8	62	1	10	28*	1.230	7.3	82	1	9
17	1.230	10.0	53	0	9	41*	1.756	12.8	72	0	9
17	1.591	11.2	68	0	10	53*	1.114	12.0	66	0	11
18	1.447	7.5	65	1	8	57*	1.255	12.5	66	0	11
19	1.079	14.4	51	0	15	77*	1.079	14.0	60	0	12
19	1.255	7.5	60	1	9						
24	1.301	14.6	56	1	9						
25	1.000	12.4	67	0	10						
26	1.230	11.2	49	1	11						
32	1.322	10.6	46	0	9						

Tabela 4.5: Tempos de sobrevivência (em meses) de pacientes portadores de mieloma múltipla (Lawless, 1982).

q	$\hat{\beta}_0(q)$	$\hat{\beta}_1(q)$	$\hat{\beta}_2(q)$	$\hat{\beta}_3(q)$	$\hat{\beta}_4(q)$	$\hat{\beta}_5(q)$	$\hat{\sigma}(q)$	$\log L_{\max}(q)$	$R_{\max}(q)$
-2.00	3.9446	-1.5525	0.1703	-0.0024	0.3577	-0.1437	0.7168	-77.6608	0.3999
-1.80	4.0375	-1.5480	0.1690	-0.0020	0.3566	-0.1477	0.7596	-77.5616	0.4416
-1.40	4.2541	-1.5448	0.1653	-0.0009	0.3416	-0.1555	0.8445	-77.2458	0.6056
-1.00	4.4865	-1.5525	0.1603	0.0006	0.3068	-0.1623	0.9185	-76.9022	0.8539
-0.60	4.6908	-1.5692	0.1541	0.0027	0.2550	-0.1664	0.9706	-76.7443	1.0000
-0.40	4.7693	-1.5793	0.1505	0.0042	0.2243	-0.1669	0.9859	-76.7922	0.9532
-0.25	4.8136	-1.5870	0.1475	0.0054	0.1999	-0.1665	0.9924	-76.8954	0.8598
0.0	4.8528	-1.5994	0.1420	0.0077	0.1578	-0.1636	0.9934	-77.2119	0.6265
0.20	4.8477	-1.6082	0.13726	0.0098	0.1245	-0.1591	0.9854	-77.6306	0.4122
0.35	4.8216	-1.6140	0.13362	0.0115	0.1009	-0.1543	0.9743	-77.9703	0.2935
0.50	4.7780	-1.6192	0.12998	0.0133	0.0796	-0.1482	0.9590	-78.4091	0.1892
0.65	4.7207	-1.6237	0.1264	0.01503	0.0615	-0.1411	0.9398	-78.9130	0.1143
1.00	4.5674	-1.6323	0.1183	0.0187	0.0356	-0.1212	0.8816	-80.2782	0.0292

Tabela 4.6: Etapas para estimação dos parâmetros para dados de sobrevivência de portadores de mieloma.

poucas evidências para o modelo Weibull. De fato, temos $R_{\max}(0) = 0.6265$ e $R_{\max}(1) = 0.02919$, e com base na estatística de razão de verossimilhança (RV), rejeitamos $H_0 : q = q_0$ a um nível 0.05, para valores de R_{\max} menores que 0.1466 ($-2 \log R_{\max}(q) > 3.8441 \Rightarrow R_{\max}(q) < 0.1466$). Isto indica que o modelo log-normal é um modelo adequado, enquanto que o modelo Weibull, e o modelo exponencial não são admitidos por este critério para ajustar os dados (pelo teste da RV para o modelo de regressão exponencial, encontramos o valor $p < 0.025$).

A seguir temos gráficos com respeito aos resíduos $z_i = \frac{y_i - x_i \hat{\beta}}{\hat{\sigma}}$. Para a Figura 4-7 foi assumido o modelo l.g.g. estendido, ou seja, usamos as estimativas globais $\hat{\beta}$ e $\hat{\sigma}$ dadas acima para o cálculo dos z_i 's. A proximidade entre a curva de Kaplan-Meier (K-M) destes resíduos e a função de sobrevivência da log-gama para $q = -0.6$, mostra que o ajuste aparenta ser adequado. Na Figura 4-8 foi assumido o modelo normal para o log dos tempos, e as estimativas $\tilde{\beta}(0) = (\tilde{\beta}_0(0), \tilde{\beta}_1(0), \dots, \tilde{\beta}_5(0))$ e $\hat{\sigma}(0)$, dadas na Tabela 4.6 para $q = 0$, foram usadas para calcular os resíduos. A curva de K-M destes resíduos é comparada com a função de sobrevivência da normal padrão. De maneira similar, na figura 4-9 os resíduos são calculados com base na suposição do modelo valor extremo

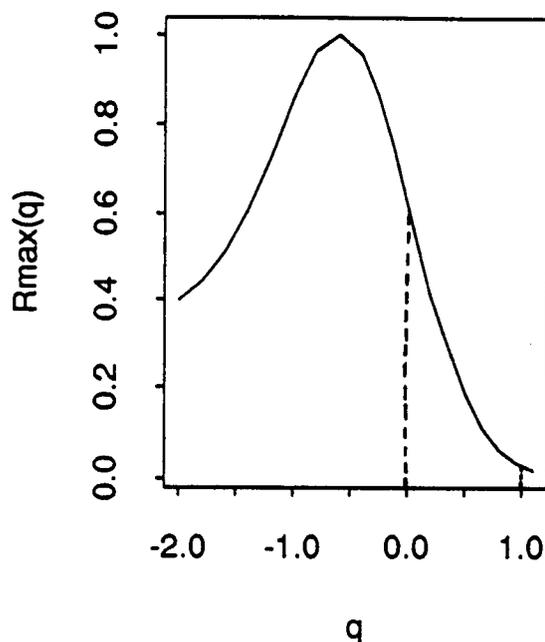


Figura 4-6: Contorno da verossimilhança relativa maximizada $R_{\max}(q)$, para dados de portadores de mieloma múltipla.

para o log dos tempos (exponencial para os tempos), e a curva de K-M destes resíduos é mostrada junto com a curva de sobrevivência da distribuição valor extremo padrão. Comparando estas Figuras 4-8 e 4-9, constatamos uma melhor concordância no ajuste log-normal que no ajuste exponencial.

4.1.4 Exemplo 4 - Sobrevivência de Portadores da Doença de Hodgkin

Os dados da Tabela 4.7 são resultados de um ensaio clínico e representam o tempo de sobrevivência de 35 pacientes (em meses) com a **doença de Hodgkin** em estado avançado, tratados com terapia de gás nitrogênio. Um grupo de pacientes (grupo A) recebeu pouca ou nenhuma terapia prévia, enquanto que um outro grupo (grupo B) havia recebido forte terapia antes de começar esse tratamento. As censuras, denotadas com um asterisco (*), correspondem aos tempos decorridos até o último contato com

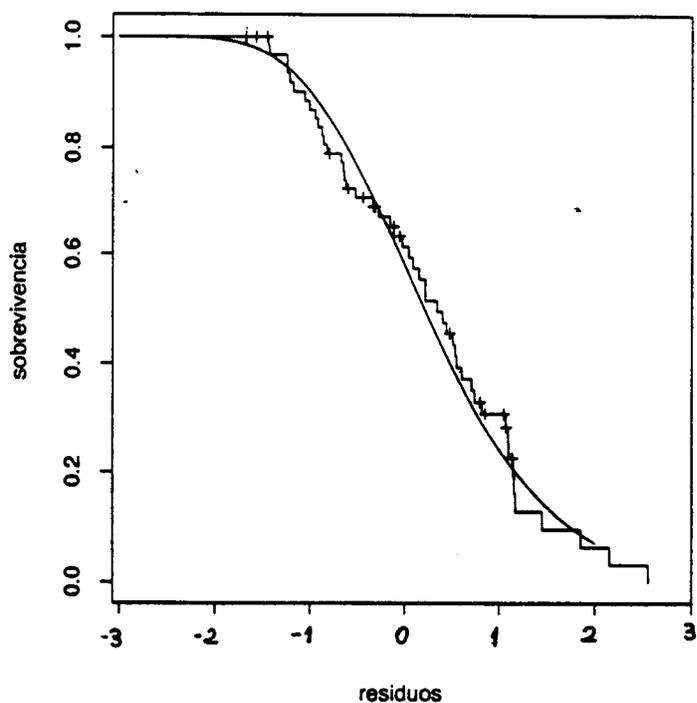


Figura 4-7: Curva de sobrevivência estimada para a distribuição **log-gama** ($q = -0.6$) e o gráfico de Kaplan-Meier para os resíduos estimados. Exemplo 3 - portadores de mieloma múltipla.

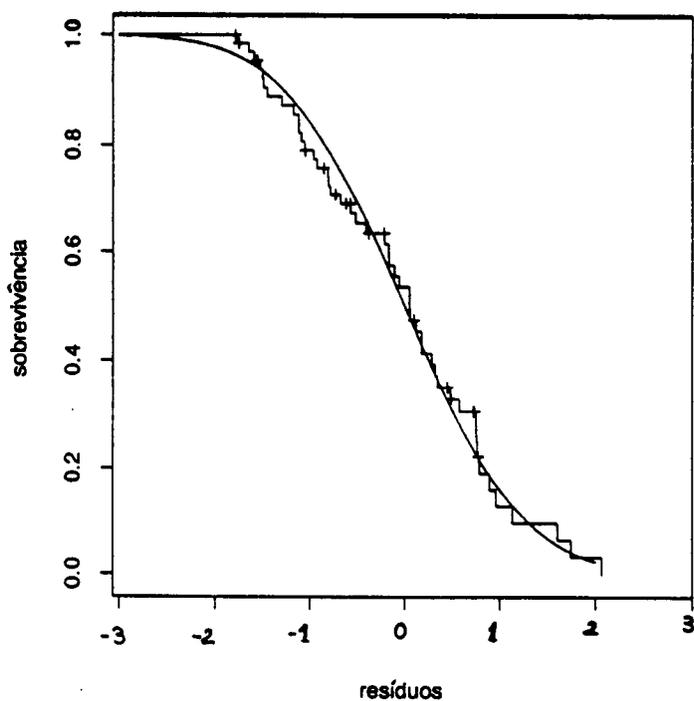


Figura 4-8: Curva de sobrevivência para a distribuição **normal padrão** e o gráfico de Kaplan - Meier para resíduos estimados. Exemplo 3 - Portadores de mieloma múltipla.

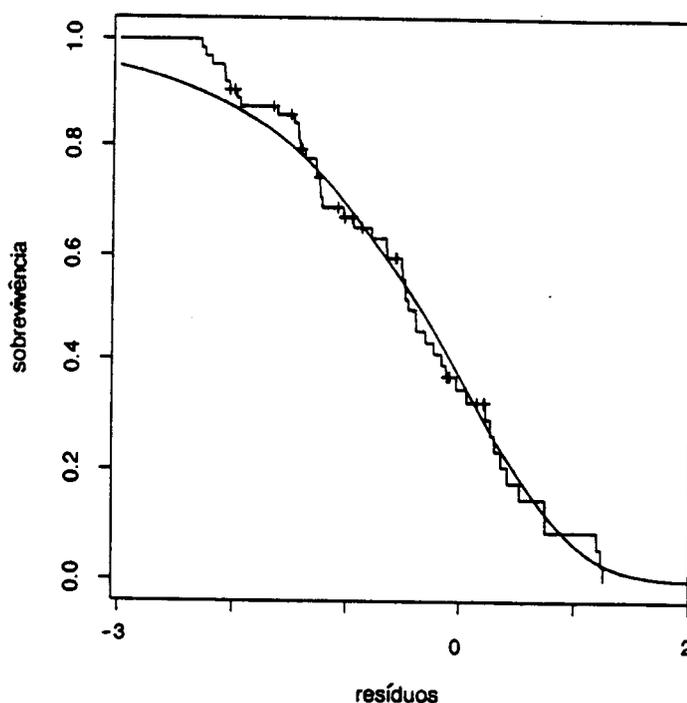


Figura 4-9: Curva de sobrevivência para a distribuição **valor extremo padrão** e o gráfico de Kaplan - Meier para resíduos estimados. Exemplo 3 - Portadores de mieloma múltipla.

pacientes que saíram do tratamento, ou daqueles ainda vivos até o final do ensaio. O principal objetivo do estudo foi avaliar a diferença da sobrevivência nos grupos.

Bartolucci e Dickey (1977) utilizaram estes dados para comparar métodos inferenciais bayesianos com os tradicionais e assumiram para os dados completos (incluindo falhas e censuras), a distribuição exponencial para cada grupo. Consideramos aqui um modelo de regressão l.g.g. estendido para o log dos tempos, $Y_i = \beta_0 + \beta_1 x_i + \sigma Z_i$, com uma covariável indicadora x representando o grupo a que o indivíduo pertence.

Realizamos o procedimento iterativo para obtenção dos estimadores de máxima verossimilhança, e encontramos $q = -0.4$, $\hat{\beta}_0 = 2.5948(0.7229)$, $\hat{\beta}_1 = -0.2026(0.4888)$ e $\hat{\sigma} = 1.2634(0.1698)$. A verossimilhança para estes valores é $\log L_{\max}(-0.4) = -50.5484$. Checando graficamente o ajuste, através da comparação da sobrevivência da log-gama ($q = -0.4$) com o K-M dos resíduos estimados (Figura 4-10), vemos que o ajuste do modelo log-gama parece adequado.

	Tempos de Vida (meses)					
Grupo A	1.25	1.41	4.98	5.25	5.38	6.92
	8.89	10.98	11.18	13.11	13.21	16.33
	19.77	21.08	21.84*	22.07	31.38*	32.62*
	37.18*	42.92				
Grupo B	1.05	2.92	3.61	4.20	4.49	6.72
	7.31	9.08	9.11	14.49*	16.85	18.82*
	26.59*	30.26*	41.34*			

Tabela 4.7: Tempos de sobrevivência de portadores da doença de Hodgkin tratados com terapia de gás nitrogênio. Grupo A sem terapia prévia e grupo B com terapia prévia (Bartolucci e Dickey, 1977).

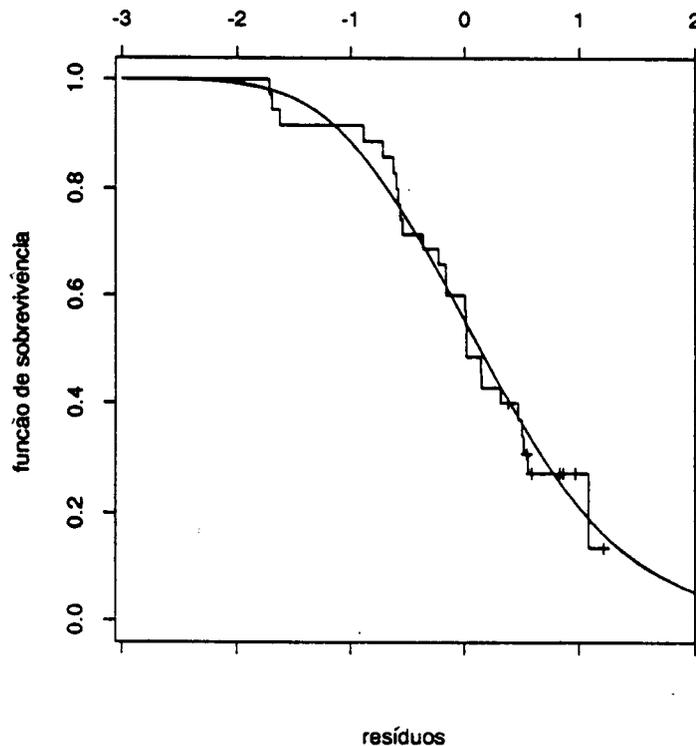


Figura 4-10: Curva de sobrevivência estimada para a distribuição log-gama e função de Kaplan - Meier para resíduos estimados. Exemplo 4 - portadores da doença de Hodgkin

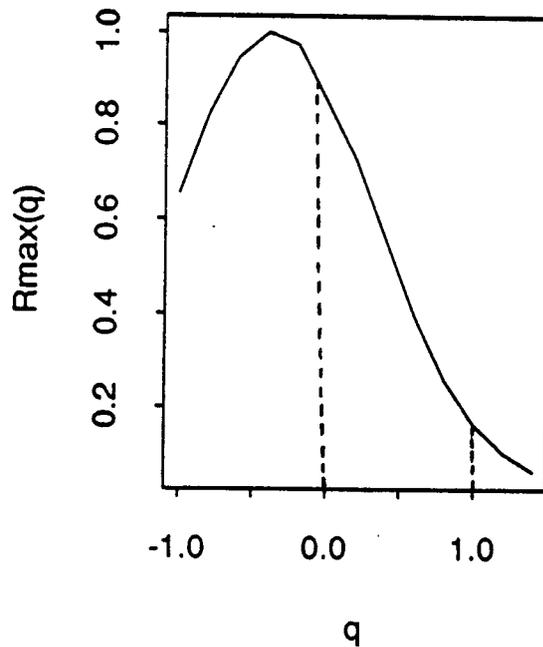


Figura 4-11: Contorno da verossimilhança relativa maximizada $R_{\max}(q)$, para dados de portadores da doença de Hodgkin.

A Tabela 4.8 mostra resultados obtidos para alguns valores fixo de q , e fornece os correspondentes valores de $R_{\max}(q)$, cujo gráfico em função de q é dado na Figura 4-11. O teste da razão de verossimilhança admite, para um nível significância 0.05, os modelos Weibull ($-2 \log R_{\max}(1) = 3.6247$), exponencial ($-2 \log R_{\max}(H_0 : \sigma = q = 1) = 3.6279$) e o log-normal ($-2 \log R_{\max}(0) = 0.2637$). Contudo o exame da figura 4-11 nos mostra que o modelo log-normal é o mais forte candidato para ajustar os dados (o valor p do teste de ajuste exponencial é $p_1 > 0.15$, e para o ajuste log-normal é $p_2 > 0.50$). Acreditamos então que embora seja admitido o modelo exponencial, este não é o mais apropriado para a amostra estudada.

Com respeito à diferença entre os grupos, consideremos o teste da hipótese $H_0 : \beta_1 = 0$. Supondo a normalidade assintótica de $\hat{\beta}_1$, usamos a estatística $(\hat{\beta}_1 / dp(\hat{\beta}_1))^2$ que sob H_0 tem distribuição $\chi^2_{(1)}$, onde $dp(\hat{\beta}_1)$ é a raiz quadrada da variância assintótica estimada de $\hat{\beta}_1$. Sob o modelo log-gama estendido, isto resulta em $(-0.2026/0.4888)^2 = 0.1718$, indicando que a um nível 0.05 não existem evidências para rejeitar H_0 , ou seja, não

q	$\beta_0(q)$	$\beta_1(q)$	$\hat{\sigma}(q)$	$\log L_{\max}(q)$	$R_{\max}(q)$
-1.0	2.2619	-0.2021	1.2686	-50.9698	0.6561
-0.8	2.3820	-0.2079	1.2737	-50.7433	0.8229
-0.6	2.4930	-0.2078	1.2721	-50.6046	0.9453
-0.4	2.5948	-0.2026	1.2634	-50.5484	1.0000
-0.2	2.6868	-0.1922	1.2473	-50.5757	0.9731
0.0	2.7691	-0.1768	1.2238	-50.6794	0.8764
0.2	2.8424	-0.15686	1.1932	-50.8713	0.7240
0.4	2.9069	-0.13244	1.1560	-51.1356	0.5559
0.6	2.9636	-0.10408	1.1125	-51.4835	0.3925
0.8	3.0135	-0.07247	1.0634	-51.8992	0.2590
1.0	3.0582	-0.03890	1.0093	-52.3640	0.1627
1.2	3.09940	-0.0049	0.9513	-52.8306	0.1021
1.4	3.1386	0.0275	0.8910	-53.2874	0.0646

Tabela 4.8: Etapas para estimação dos parâmetros para dados de sobrevivência de portadores da doença de Hodgkin.

existe diferença estatisticamente significativa entre os grupos de pacientes com respeito ao tempo de sobrevivência. Usando estatística da razão de verossimilhança dada em 3.30 sob o mesmo modelo, encontramos $\Lambda = 0.1033$, que também indica a rejeição da hipótese a um nível de significância 0.05.

4.1.5 Exemplo 5 - Corrosão de Metais

Consideremos como ilustração um conjunto de dados utilizado em Nelson (1990), onde a variável resposta não representa "tempo de vida", mas a diferença de peso de um tipo de metal. Especificamente temos uma amostra de trinta espécimes de metal ("permalloy") que foram testados em uma atmosfera corrosiva a seis umidades relativas (% U.R.). A perda de peso de cada espécime sobre um tempo estabelecido foi registrada. Os resultados do experimento estão dados na Tabela 4.9.

O principal objetivo deste estudo seria acessar se a mudança de peso a 10% e 20% de umidade relativa está abaixo do valor especificado 0.0050.

Na Figura 4-12 temos representações gráficas dos níveis de UR versus a mudança de peso, e versus o logaritmo natural da mudança de peso das espécimes. Através destes

U.R. (%)	Corrosão (Mudança de Peso)							
30	0.0144	0.0153	0.0092	0.0120	0.0111	0.0163	0.0193	0.024
40	0.0221	0.0280	0.0287	0.0301	0.0330			
50	0.0744	0.0684						
60	0.1050	0.1110	0.1160	0.1185				
70	0.1665	0.2065	0.2220	0.2540	0.2930	0.3008	0.3408	
78	0.3807	0.6549	0.6660					

Tabela 4.9: Perda de peso de metais em uma atmosfera corrosiva (Nelson, 1990).

gráficos percebemos que a transformação logarítmica, sempre sugerida na literatura, favorece a aceitação de uma relação linear entre estas variáveis. Tratamos então estes dados sob um modelo de regressão log-linear da forma

$$Y_i = x_i\beta + \sigma Z_i \quad i = 1, \dots, 30,$$

onde a variável resposta Y_i é o log da mudança de peso da espécime i , a covariável x_i é o nível de umidade relativa, β e σ são parâmetros do modelo, e assumimos como nos exemplos anteriores, que o resíduo Z_i tem distribuição log-gama com parâmetro de forma q .

Obtemos como antes, os estimadores de máxima verossimilhança, em um processo de maximização em várias etapas, e os resultados são dados abaixo,

$$\hat{q} \doteq -0.09$$

$$\hat{\beta}_0 = -6.497(0.1346) \quad \hat{\beta}_1 = 0.074(0.0025) \quad \hat{\sigma} = 0.234(0.0302)$$

$$\log L_{\max}(-0.09) = 0.9927.$$

Vemos na Figura 4-13 que o ajuste parece bastante adequado, ou seja, a suposição de que os resíduos possuem distribuição log-gama de um parâmetro, parece razoável quando comparamos esta curva ($q = -0.09$) com a função de sobrevivência empírica dos resíduos estimados.

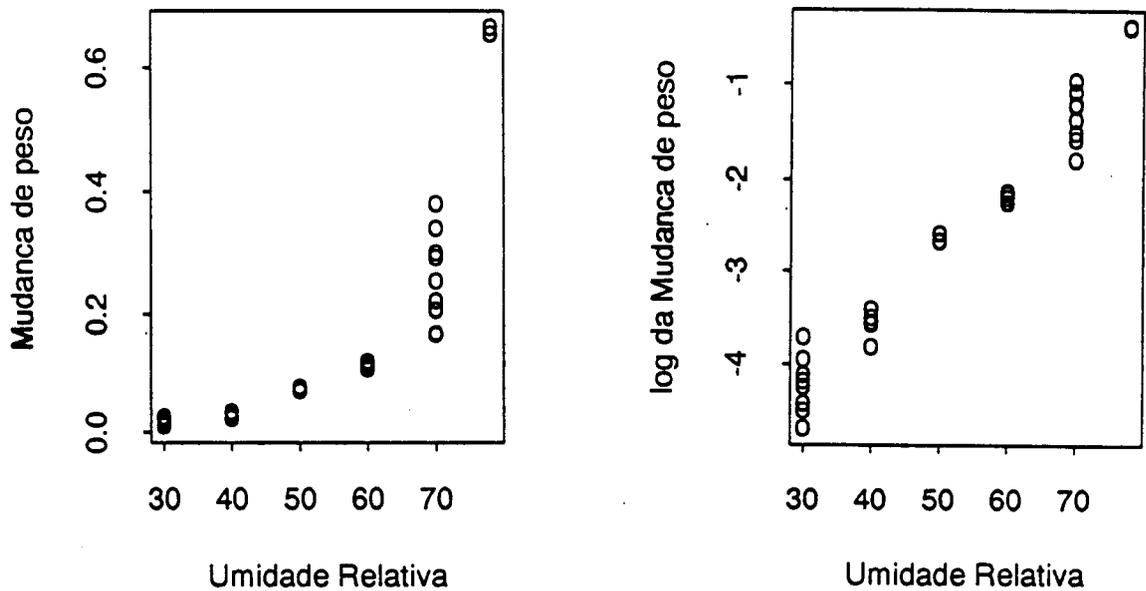


Figura 4-12: À esquerda, mudança de peso de metais em atmosfera corrosiva, por níveis de umidade relativa, e à direita o mesmo gráfico com a transformação logarítima para as mudanças de peso.

A Figura 4-14 apresenta o contorno da verossimilhança relativa maximizada, e este mostra que o modelo log-normal ($q = 0$) é fortemente indicado para o ajuste dos dados. Um intervalo de confiança aproximado de nível 0.95 para o parâmetro q dado por este gráfico é $[-1.02, 0.91]$, e confirma a adequação do modelo log-normal, além de mostrar que, a este nível de confiança o modelo Weibull ($q = 1$) não pertence ao conjunto de modelos favoráveis para ajustar estes dados.

Admitindo então o modelo normal para Y , consideramos os estimadores de máxima verossimilhança obtidos para $q = 0$, ou seja,

$$\hat{\beta}_0 = -6.4870(0.1339), \quad \hat{\beta}_1 = 0.0741(0.0024) \quad e \quad \hat{\sigma} = 0.2341(0.0302).$$

Desta forma, os valores correspondentes às médias do log das mudanças de peso estimadas, para as umidades relativas de 10% e 20%, são calculados a seguir,

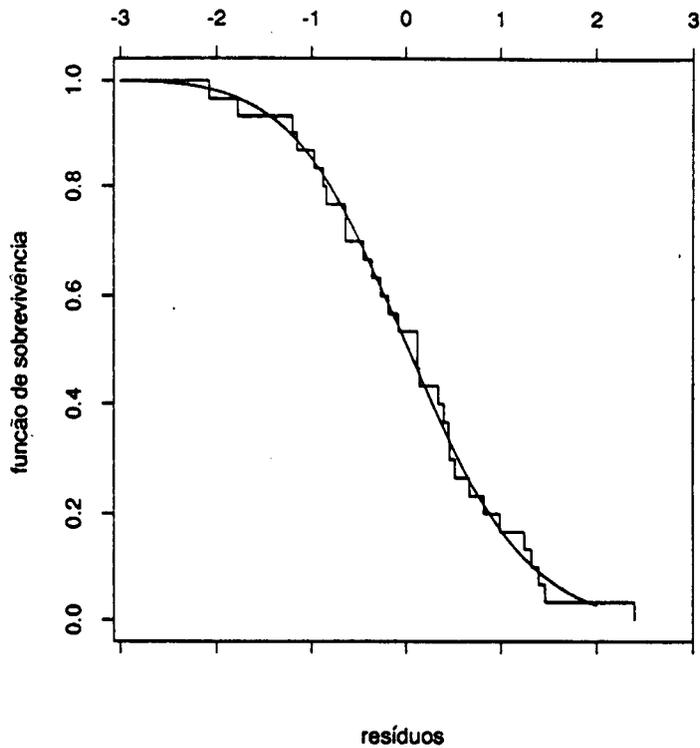


Figura 4-13: Curva de sobrevivência estimada para a distribuição log-gama ($q = -0.09$) e função de sobrevivência empírica dos resíduos estimados. Exemplo 5 - corrosão de metais.

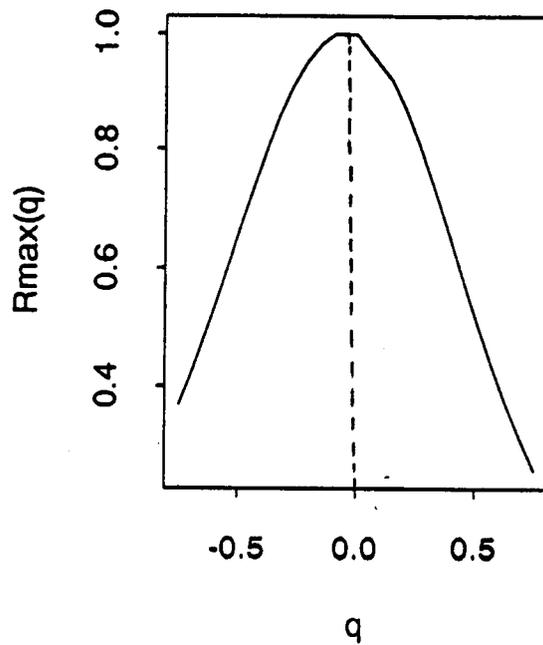


Figura 4-14: Contorno da verossimilhança relativa maximizada $R_{\max}(q)$, para dados de corrosão de metais.

$$\hat{y}_{(10\%)} = -6.4870 + 0.0741(10) = -5.746$$

$$\hat{y}_{(20\%)} = -6.4870 + 0.0741(20) = -5.005.$$

Passando estes resultados para a escala original, temos $\exp(\hat{y}_{(10\%)}) = 0.003$ e $\exp(\hat{y}_{(20\%)}) = 0.007$. Podemos também construir intervalos de confiança para estas estimativas, com base na normalidade assintótica de $\hat{y}_{(x_i)}$, isto é, $\hat{y}_{(x_i)} \sim N(y_{(x_i)}, \mathbf{x}_i I^{-1}(\hat{\beta}) \mathbf{x}_i')$ (veja Lawless(1982), seção 6.5). Desta forma, intervalos de 95% de confiança para mudanças médias de peso estimadas, a umidades relativas de 10% e 20% são respectivamente $[0.0024, 0.0042]$ e $[0.0032, 0.0138]$. Observamos então que apenas para 20% de umidade relativa a perda de peso pode ser considerada acima do valor estabelecido 0.005.

4.2 Comentários

Nos exemplos estudados encontramos diferentes formas para a função $R_{\max}(q)$, como mostram as figuras 4-3, 4-5, 4-6, 4-11, 4-14. Notamos uma assimetria com cauda longa à esquerda nas curvas obtidas para os dados de tempo de vida de portadores de Mieloma, e de pacientes com doença de Hodgkin, assimetria à direita para os exemplos 1 e 2, que tratam de testes acelerados de vida, e uma quase simetria para o exemplo 5, sobre perda de peso de metais. Isto mostra que, conforme citado por Farewell e Prentice (1977), não parece haver uma valor fixo provável para q , ou seja, este parâmetro assim como sua função de verossimilhança, variam sensivelmente com os dados. Observamos também que valores negativos de q ocorreram tão naturalmente quanto os positivos, indicando que a extensão do modelo l.g.g. para valores negativos de q fornece um modelo mais flexível e útil ao ajuste de dados de tempo de vida.

Dentre os quatro exemplos onde foi checada a adequacidade do modelo que havia sido adotado anteriormente, notamos que para os exemplos da engenharia (exemplos 1 e 2), os modelos foram considerados apropriados, e para os 2 exemplos da área médica, o modelo adotado originalmente no exemplo 3 foi rejeitado, e o adotado no exemplo 4, embora

tenha sido admitido, não foi considerado o mais indicado, conforme nosso critério. Estes resultados servem como uma indicação de que é necessária a adoção de algum critério consistente para a escolha de uma modelo na análise de dados de vida, e que esta escolha não deveria se basear apenas no que sugere a literatura para uma determinado tipo de experimento.

Capítulo 5

Considerações Finais

Concluimos esta dissertação acreditando na aplicabilidade do modelo gama generalizada, ou mais especificamente, do modelo l.g.g. estendido, no tratamento de dados de vida.

A teoria desenvolvida nos capítulos 2 e 3, juntamente com as referências citadas, possibilitam o uso deste modelo não apenas para discriminar entre outras distribuições, mais como um modelo flexível de vida para o ajuste de dados.

Os problemas computacionais encontrados não foram insuperáveis, além de já existirem devidamente implantados em softwares conhecidos, procedimentos para estimação e teste desse modelo.

Do ponto de vista prático, um procedimento simples que deveria ser adotado quando se desejasse escolher entre os modelos Weibull, exponencial e log-normal para ajustar dados de vida, seria obter a máxima verossimilhança de cada um destes modelos, assim como a do modelo l.g.g. estendido (por exemplo no Lifereg do SAS a partir da versão 6.04), e aplicar o teste da razão de verossimilhança para decidir sobre o modelo mais adequado.

Sabemos contudo, que apesar do l.g.g. estendido ser um modelo bastante flexível, ele próprio pode não ser um bom modelo para um determinado conjunto de dados. Desta forma seria necessário fazer o diagnóstico do seu ajuste antes de prosseguir na discriminação de outros modelos.

Finalmente, como sugestão para trabalhos posteriores, acreditamos que seria de grande interesse a realização de um estudo para esse modelo, semelhante ao que foi feito para os

modelos log-normal e Weibull no artigo de Doganaksoy e Schemee (1993), para avaliar o grau de imprecisão de procedimentos inferenciais baseados na normalidade assintótica dos estimadores de máxima verossimilhança, quando comparados com os procedimentos utilizando a distribuição assintótica da razão de verossimilhança. Sob o aspecto computacional, seria importante a implantação de procedimentos para a obtenção de intervalos de confiança com base na razão de verossimilhança, tendo em vista que estes são considerados por vários autores como mais precisos na análise de sobrevivência.

Apêndice A

Função Gama e Distribuição Beta

Apresentamos sumariamente algumas definições e resultados sobre a função gama e funções relacionadas, e descrevemos alguns membros da família de distribuições beta. Estudos mais detalhados sobre estas funções são obtidos em Jonhson e Kotz (1970) e Abramowitz e Stegun (1965) entre outros. Referências adicionais são dadas em Lawless (1982).

• Função Gama

A função gama é definida como

$$\Gamma(k) = \int_0^{\infty} u^{k-1} e^{-u} du \quad k > 0. \quad (\text{A.1})$$

Alguns resultados importantes relacionados com esta função são dados abaixo;

$$\Gamma(k+1) = k\Gamma(k) \quad k > 0$$

$$\Gamma(1/2) = \sqrt{\pi}$$

$$\log \Gamma(k) = (k-1/2) \log k - k + (1/2) \log(2\pi) + \frac{1}{12k} - \frac{1}{360k^3} + \frac{1}{1260k^5} - \dots$$

$$\Gamma(k) = (k-1)! \quad \text{se } k \text{ é um inteiro positivo.}$$

Função Digama e Poligama

A função digama é definida como

$$\Psi(k) = \frac{\partial \log \Gamma(k)}{\partial k} = \frac{\Gamma'(k)}{\Gamma(k)} \quad k > 0.$$

Definimos também as funções poligama como,

$$\Psi^{(n)}(k) = \frac{\partial^n \Psi(k)}{\partial k^n} \quad n = 1, 2, \dots$$

Temos abaixo alguns resultados de interesse para estas funções,

$$\Psi(k+1) = \Psi(k) + \frac{1}{k} \quad k > 0$$

$$\Psi(1) = -\gamma = -0.577215$$

$$\Psi'(k+1) = \Psi'(k) - 1/k^2$$

$$\Psi'(1) = \pi^2/6,$$

onde γ é a constante de Euler.

As fórmulas assintóticas para Ψ e Ψ' , com $k \rightarrow \infty$, são

$$\Psi(k) = \log k - \frac{1}{2k} - \frac{1}{12k^2} + \frac{1}{120k^4} - \frac{1}{252k^6} + \dots \quad (\text{A.2})$$

$$\Psi'(k) = \frac{1}{k} + \frac{1}{2k^2} + \frac{1}{6k^3} - \frac{1}{30k^5} + \frac{1}{42k^7} + \dots \quad (\text{A.3})$$

Função Gama Incompleta

A gama incompleta é uma função definida por

$$I(k, x) = (1/\Gamma(k)) \int_0^x u^{k-1} e^{-u} du \quad k > 0 \quad x > 0. \quad (\text{A.4})$$

Definimos também a *integral gama incompleta* como

$$Q(k, x) = 1 - I(k, x). \quad (\text{A.5})$$

Função Beta

A função beta é definida como

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = B(b, a),$$

• A Família Beta

A família de distribuições beta é formada por todas as distribuições com f.d.p. da forma

$$f(y) = \frac{1}{B(a, b)} \frac{(y-u)^{a-1}(v-y)^{b-1}}{(v-u)^{a+b-1}} \quad \begin{array}{l} u \leq y \leq v \\ a, b > 0. \end{array}$$

Distribuição Beta Padrão

Fazendo a transformação $X = (Y - u)/(v - u)$, a densidade de X é dada por

$$f(x) = \frac{1}{B(a, b)} x^{a-1}(1-x)^{b-1} \quad \begin{array}{l} 0 \leq x \leq 1 \\ a, b > 0. \end{array} \quad (\text{A.6})$$

Esta representa a distribuição beta padrão, com parâmetros a e b (notação : $B_p(a, b)$).

Distribuição Beta 1

Se X tem distribuição dada por A.6, então fazendo $T = X/(1 + X)$, obtemos a distribuição abaixo,

$$f(t) = \frac{1}{B(a, b)} \frac{t^{a-1}}{(1-t)^{a+b}} \quad t > 0.$$

Esta é chamada distribuição beta 1 com parâmetros a e b (notação : $B_1(a, b)$).

Apêndice B

Propriedades da Distribuição

Log-Gama

Tratamos aqui da descrição de algumas propriedades da distribuição log-gama (Bartlett e Kendall, 1946), que representa a distribuição do log de uma distribuição gama. Referências e maior detalhamento teórico podem ser obtidos em Kotz e Johnson (1985).

Seja X uma variável com distribuição gama com um parâmetro cuja densidade é da forma $(1/\Gamma(k))x^{k-1} \exp(-x)$, $x > 0$. A variável $Y = \log X$ tem distribuição log-gama com densidade

$$f_Y(y) = \frac{1}{\Gamma(k)} \exp[yk - e^y] \quad y \in \mathfrak{R}, \quad k > 0. \quad (\text{B.1})$$

Função de Distribuição

A função de distribuição de uma log-gama é dada por

$$F_Y(y) = I(k, e^y),$$

onde $I(k, a)$ é a função gama incompleta definida em A.4,

Demonstração:

$$\begin{aligned}F_Y(y) &= \int_{-\infty}^y \frac{1}{\Gamma(k)} \exp [tk - e^t] dt \\&= \frac{1}{\Gamma(k)} \int_{-\infty}^y (e^t)^k e^{-e^t} dt \\&= \frac{1}{\Gamma(k)} \int_0^{e^y} u^{k-1} e^{-u} du \\&= I(k, e^y),\end{aligned}$$

sendo $u = e^t$. □

Função Geratriz de Momentos

A função geratriz de momentos de uma log-gama é dada por

$$M(t) = \Gamma(t + k)/\Gamma(k). \quad (\text{B.2})$$

Demonstração:

$$\begin{aligned}M(t) &= E(e^{tY}) = \int_{-\infty}^{\infty} e^{ty} \frac{1}{\Gamma(k)} \exp [yk - e^y] dy \\&= \frac{1}{\Gamma(k)} \int_{-\infty}^{\infty} \exp [y(t + k) - e^y] dy \\&= \frac{1}{\Gamma(k)} \int_{-\infty}^{\infty} (e^y)^{(t+k)} e^{-e^y} dy \\&= \frac{1}{\Gamma(k)} \int_0^{\infty} u^{(t+k-1)} e^{-u} du = \Gamma(t + k)/\Gamma(k),\end{aligned}$$

onde consideramos que $u = e^y$. □

Momentos

O momento de ordem r da distribuição log-gama é dado por

$$E(Y^r) = \frac{\Gamma^{(r)}(k)}{\Gamma(k)}. \quad (\text{B.3})$$

Demonstração:

Considere inicialmente o lema abaixo

Lema 4 - Sejam $k, t > 0$ e $\Gamma(\cdot)$ a função gama, definida no apêndice A. Então

$$\frac{\partial^r \Gamma(k+t)}{\partial k^r} = \frac{\partial^r \Gamma(k+t)}{\partial t^r}.$$

Prova.

Observamos que a prova é trivial, pois existe uma dualidade entre os símbolos k e t . Trocando a posição destes símbolos obtemos o mesmo resultado. \square

Podemos demonstrar a relação B.3 conforme segue. Por definição sabemos que

$$E(Y^r) = \frac{\partial^r M(t)}{\partial t^r} \Big|_{t=0}.$$

Usando a relação B.2 temos,

$$E(Y^r) = \frac{1}{\Gamma(k)} \left(\frac{\partial^r \Gamma(k+t)}{\partial t^r} \Big|_{t=0} \right),$$

e pelo lema 4, segue que

$$E(Y^r) = \frac{1}{\Gamma(k)} \left(\frac{\partial^r \Gamma(k+t)}{\partial k^r} \Big|_{t=0} \right) = \frac{\Gamma^{(r)}(k)}{\Gamma(k)}. \square$$

Com base neste resultado, temos que a média e a variância da log-gama são dadas por,

$$E(Y) = \Psi(k) \quad e \quad Var(Y) = \Psi'(k),$$

onde Ψ e Ψ' são respectivamente as funções digama e trigama, definidas no apêndice A.

Distribuição log-gama generalizada

Se Y tem distribuição log-gama, a variável aleatória $Z = (1/\beta)Y + \log \alpha$ tem dis-

tribuição log-gama generalizada, com densidade

$$f_Z(z) = \frac{\beta}{\alpha^{\beta k} \Gamma(k)} \exp \left\{ \beta k z - \frac{e^{z\beta}}{\alpha^\beta} \right\} \quad \begin{array}{l} z \in \mathfrak{R} \\ \alpha, \beta, k > 0. \end{array} \quad (\text{B.4})$$

Esta representa a distribuição do log de uma gama generalizada.

Apêndice C

O Método de Newton Raphson

Em muitas situações, quando desejamos encontrar o estimador de máxima verossimilhança de um parâmetro $\theta = (\theta_1, \dots, \theta_k)$, em um espaço paramétrico Θ , através da maximização da função de log-verossimilhança $L^*(\theta)$, não encontramos soluções algébricas. Nestes casos, precisamos de um método numérico iterativo para obter o estimador de máxima verossimilhança $\hat{\theta}$. Descrevemos a seguir, de forma resumida, o método de Newton Raphson, que é bastante usado na prática. Informações mais detalhadas sobre este e outros métodos iterativos são obtidas por exemplo em Chambers (1977) e Kennedy e Gentle (1980).

Seja o **vetor de escores**, denotado por $\mathbf{U}(\theta)$, definido como

$$\mathbf{U}(\theta) = \frac{\partial L^*(\theta)}{\partial \theta},$$

e a **matriz de informação observada**, $\mathbf{I}(\hat{\theta})$ dada por

$$\mathbf{I}(\hat{\theta}) = -\frac{\partial^2 L^*(\theta)}{\partial \theta \partial \theta'} \Big|_{\theta = \hat{\theta}} = -\frac{\partial \mathbf{U}(\theta)}{\partial \theta} \Big|_{\theta = \hat{\theta}}.$$

Temos então que as equações de máxima verossimilhança são expressas por

$$\mathbf{U}(\hat{\theta}) = \mathbf{0}. \tag{C.1}$$

Expandindo-se $\mathbf{U}(\hat{\theta})$ em série multivariada de Taylor até a 1ª ordem, ao redor de um ponto θ_0 pertencente a uma vizinhança de $\hat{\theta}$, obtemos a aproximação

$$\begin{aligned}
\mathbf{U}(\hat{\theta}) &\doteq \mathbf{U}(\theta_0) + \left. \frac{\partial \mathbf{U}(\theta)}{\partial \theta} \right|_{\theta = \theta_0} (\hat{\theta} - \theta_0) \\
&= \mathbf{U}(\theta_0) - \mathbf{I}(\theta_0) (\hat{\theta} - \theta_0).
\end{aligned}$$

Considerando C.1 temos,

$$\hat{\theta} \doteq \theta_0 + \mathbf{I}^{-1}(\theta_0) \mathbf{U}(\theta_0). \quad (\text{C.2})$$

O método de Newton-Raphson consiste em usar a equação C.2 iterativamente para calcular o e.m.v. $\hat{\theta}$. Para isso, na primeira iteração, dado um valor inicial $\theta_0^{(1)}$, calculamos $\hat{\theta}^{(1)}$, e repetimos o processo, usando como novo valor inicial, o resultado da iteração anterior, isto é, fazemos $\theta_0^{(j+1)}$ receber o valor de $\hat{\theta}^{(j)}$, para $j = 1, 2, \dots$, até que a distância entre $\theta_0^{(j)}$ e $\hat{\theta}^{(j)}$ seja desprezível (menor que um valor pequeno, previamente estabelecido). Se $L^*(\theta)$ é uma função bem comportada em $\hat{\theta}$, e o valor inicial θ_0 não estiver muito longe de $\hat{\theta}$, a sequência de aproximações geradas $\{\hat{\theta}^{(m)}, m \geq 1\}$ irá convergir para o e.m.v. $\hat{\theta}$.

Apêndice D

Métodos Computacionais

Existem diversos pacotes estatísticos com técnicas para tratar dados de tempo de vida. Nelson (1990) fornece referências e um resumo dos principais métodos desenvolvidos nos pacotes CENSOR, GRAFSTAT, LINDEP, SAS, STAR, STATPAC, SURVICALC, SURVIVAL e SURVIREG. De acordo com este autor, dentre estes pacotes apenas o SAS, o STAR, e o SURVICALC utilizam procedimentos sob o modelo gama generalizada. De fato o SAS, módulo LIFEREG (a partir da versão 5.0)¹ trabalha com o modelo l.g.g. estendido, como dado em 2.33 e 2.34, e estima simultaneamente os parâmetros $\beta = (\beta_0, \beta_1, \dots, \beta_s), \sigma$ e q .

Neste trabalho optamos por fazer um programa na linguagem S-Plus versão 2.0 (Becker e Chambers, 1988) que, fornece além das estimativas de máxima verossimilhança dos parâmetros e suas respectivas variâncias assintóticas, os valores da função $R_{\max}(q)$ (3.28), como dada em Lawless (1982).

O S-Plus é um software estatístico com diversos métodos de análise previamente desenvolvidos, que também pode ser utilizado como uma linguagem de programação. O módulo que trata de análise de sobrevivência possui apenas procedimentos não paramétricos. Contudo as poderosas ferramentas de programação da linguagem S-Plus, que suporta a *programação orientada para objetos*, possibilitaram o desenvolvimento de programas que trabalham com o modelo paramétrico que estudamos. As rotinas elaboradas nesta lin-

¹Foram encontrados resultados inconsistentes no ajuste de dados sob este modelo, na versão 6.03 do SAS, divergindo dos resultados obtidos pelo SAS versão 6.04 e 6.08

guagem utilizam manipulações de matrizes e aritmética vetorial, ao invés de laços (loops) comentados usados em linguagens estruturadas, visando otimizar o tempo de execução dos procedimentos e a memória utilizada, além de reduzir sensivelmente o número de operações criadas. Todas os cálculos foram realizadas em dupla precisão e foram consideradas 14 casas decimais para os valores calculados.

O programa principal calcula especificamente os estimadores de máxima verossimilhança do modelo de regressão l.g.g. estendido para $q \neq 0$ (para $q = 0$ usamos os resultados do SAS - LIFEREG, versão 6.08 sob o modelo log-normal). Isto é feito através do método iterativo de Newton Raphson (ver apêndice C), que usa as derivadas de 1ª e 2ª ordem da log-verossimilhança, dadas na seção 3.3.2. Para cada valor fixo de q o programa realiza um processo iterativo para obter as estimativas $\tilde{\beta}(q)$ e $\tilde{\sigma}(q)$. Os valores adotados para inicializar o primeiro processo iterativo são as estimativas de mínimos quadrados ordinários do modelo $\log t_i = \mathbf{x}_i\beta + \sigma z_i$, sob toda a amostra ($i = 1, \dots, n$), tratando igualmente os tempos de falha e de censura. Para os demais processos, os valores iniciais usados são os resultados obtidos na maximização anterior. Dado o nome do arquivo contendo os dados, o número de covariáveis, e uma faixa de valores adequadamente escolhida para q , os principais resultados fornecidos por este programa são, uma tabela com os estimadores de β e σ e as respectivas verossimilhanças obtidas para a sequência de valores de q , os estimadores de máxima verossimilhança globais $\hat{\beta}$ e $\hat{\sigma}$, e a matriz de variância assintótica (o inverso da matriz de informação observada para $\hat{\beta}$ e $\hat{\sigma}$, condicionada ao valor de $q = \hat{q}$). Na tabela fornecida, cada linha representa o resultado de um processo de maximização da log-verossimilhança para um dado q . Após encontrados os e.m.v.'s globais, são calculados os valores de $R_{\max}(q)$, e incluídos na tabela. Com estes resultados podemos construir o gráfico de $R_{\max}(q)$ versus q que ilustra principalmente a possibilidade de escolha entre o modelo Weibull e o log-normal, e com o seu contorno podemos obter um intervalo de confiança aproximado para q .

Para os dados utilizados no capítulo 4 o método iterativo de Newton Raphson geralmente apresentou convergência em no máximo 5 iterações, e as dificuldades encontradas

na convergência foram devido aos valores iniciais atribuídos ao parâmetro q , principalmente quando estes encontravam-se distantes do ponto maximizando a verossimilhança. Estas ocorrências estão de acordo com a teoria deste método iterativo.

Fôram elaborados também na linguagem S-Plus, outros programas que fornecem os gráficos apresentados no corpo da tese. Mostramos a seguir, o programa principal com comentários (procedidos do símbolo #), para facilitar o seu entendimento.

```

# Nome do programa : lgama.s
#.....
#   O objetivo principal deste programa e calcular os estimadores
# de maxima verossimilhanca dos parametros do modelo Log-gama
# generalizada estendido, conforme parametrizacao dada em Farewell
# e Prentice(1977).
#.....
if(search()[[5]]!="d:/jon/survival/_Data")attach("d:/jon/survival/_Data",pos=5)
if (dev.cur()==1 && dev.cur() !=3) vga()
#.....
# ENTRADA DE DADOS
# Para cada conjunto de dados preciso modificar :
# - O nome do arquivo com os dados : 'dados.dta'
# - O numero de covariaveis : s
#.....
s <- 5
p<-s+1
coluna <- s+2
dados<-matrix(scan('dados.dta'),byrow=T,ncol=coluna)
tempo <- dados[,1] ; n<- length(tempo)
falha <-dados[,2]
Covariaveis
X<-NULL
if (s>0)
{ X <- dados[,3:coluna]
  X <- as.matrix(X) }
xo <- rep(1,n) ; X<-cbind(xo,X)

```

```

#Discriminando dados de falha e de censura.
#-----
tf<-tempo[falha==1]
tc<-tempo[falha==0]
nf<- length(tf)
if (s>0)
{Xf<-X[falha==1]; Xf<-matrix(Xf,ncol=p)
  Xc<-X[falha==0]
}
nc<- length(tc)
if (nc>0)
{ if (s>0)
{ Xc<-matrix(Xc,ncol=p)
  }
  yc <- log(tc)
}
if(nc==0)
  { Xc <- 0 ; yc <- 0 }
yf <- log(tf)
y <- log(tempo)

# Funcao w = (y-xb)/s
#-----
gw <- function(py,pbeta,ps,pX)
  {if ( length(py) <= 1) 0
    else
{ Xbeta <- pX%*%pbeta

```

```

        (py - Xbeta)/ps
    }
}

# Funcao Q (gama integral incompleta)
#-----
Q <- function(pk,px){ 1-pgamma(px,pk) }

# Funcao de densidade da Log-gama generalizada
#-----
fdp <- function(pq,ps,pw)
  { if (pq == 0) dnorm(pw,0,1)/ps
    else
      { k <- pq ^(-2)
        r <- pq*pw
        const <- (abs(pq)*(k^k))/(ps*gamma(k))
        const*exp( k * (r - exp(r)))
      }
  }
}

# Funcao de sobrevivencia
#-----
fs <- function(pq,pw)
  { if (pq == 0) sobre <- (1-pnorm(pw,0,1))
    else
      { k <- pq ^(-2)
        Qv <- Q(k,k*exp(pq*pw))
        if (pq > 0) {sobre <- Qv }
        if (pq < 0) {sobre <- (1-Qv) }
      }
  }

```

```

}
sobre
}
# Funcao de log-verossimilhanca
#-----
logL <- function(pq,ps,pwfalha,pwcensura)
{ logLf <- sum(log(fdp(pq,ps,pwfalha)))
  logLc <- 0
  if (nc > 0) { logLc <- sum(log(fs(pq,pwcensura))) }
  logLf + logLc
}
#.....
# Derivadas da funcao de log-verossimilhanca com respeito
# aos parametros Beta=(b0,...,bp), e sigma.
#.....

# 1. Derivadas para tempos de FALHA:
# -----
derivaf<- function(pq,ps,pwf,pXf)
{ eqwf <- as.vector(exp(pq*pwf))
  pwf <- as.vector(pwf)
  s2 <- ps^2
  invsq <- 1/(ps*pq)
  # 1.1. Primeiras derivadas
# -----
  # Beta
  dBf <- (-invsq) * t(pXf[,1]) %*% (pwf * (1-eqwf))
  # Sigma

```

```

dsf <- (-invsq)* sum(pq+pwf-(pwf*eqwf))
# 1.2. Segundas derivadas
# -----
# Beta2
d2Bf <- (-1/s2)*( t(pXf) %*% (pXf*eqwf) )
# Sigma2
d2sf <- sum( ( 1/ (pq*s2) ) *( pq+2*pwf-2*pwf*eqwf-(pwf^2)*eqwf*pq ))
# Beta e Sigma
d2Bsf <- (1/(pq*s2))* ( t(pXf[,1]) %*% ( pXf-(pXf*eqwf)-(pq*pXf*pwf*eqwf) ) )
# Vetor de primeiras derivadas para tempo de falha
#-----
Utetaf <<- rbind(t(dBf),dsf)
# Menos a matriz de 2a's derivadas para tempos de falha
#-----
Iobsf <<- - cbind(rbind(d2Bf,d2Bsf),rbind(t(d2Bsf),d2sf) )
}

# 2. Derivadas para tempos de CENSURA :
# -----
derivac <- function(pq,ps,pwc,pXc)
{ pwc <- as.vector(pwc)
k <- pq^(-2)
s2 <- ps^2
m <- as.vector(k*exp(pwc*pq))
Qinv <- Q(k,m)^(-1)
lQinv <- (1-Q(k,m))^(-1)
dQm <- (-1/gamma(k))*m^(k-1)*exp(-m)
d2Qm <- exp(-m)*(m^(k-2))*(m-k+1)/gamma(k)
if ( pq > 0)

```

```

{ dlogSm <- Qinvs * dQm
  d2logSm <- Qinvs*d2Qm - (Qinvs^2)*(dQm^2)
}
if (pq < 0)
{ dlogSm <- lQinvs *(-dQm)
  d2logSm <- lQinvs*(-d2Qm) - (lQinvs^2)*(dQm^2)
}
dMB <- -(pq/ps)*m*pXc
d2MBs <- (pq/s2)*( m*pXc + pq*m*psc*pXc)
dms <- (-pq/ps)*m*psc
d2ms <- (psc*m*pq)*(2+psc*pq)
# 2.1. Primeiras derivadas
#-----
# Beta
dBc <- dlogSm %*% dMB
# Sigma
dsc <- sum(dlogSm*dms)
# 2.2. Segundas derivadas
#-----
# Beta2
  d2Bc <- matrix(nrow=p,ncol=p)
  for (j in (1:p))
    {for ( l in (1:p) )
      {if (j <= l)
        {d2mb <- (1/k*s2)*(pXc[,j]*pXc[,l]*m)
          vetor2 <- d2logSm*dMB[,j]*dMB[,l] +(dlogSm*d2mb)
            d2Bc[l,j] <- sum(vetor2)
              d2Bc[j,l] <- d2Bc[l,j]
        }
      }
    }

```

```

}
}
}
# Beta.Sigma
d2Bsc <- ( (d2logSm*dms)**%dMB ) + (dlogSm**%d2MBs)
# Sigma2
d2sc <- (d2logSm **% t(dms^2)) + (dlogSm**%t(d2ms))
# Vetor de primeiras derivadas para tempos de censura
#-----
Utetac <<- rbind(t(dBc),dsc)
# Menos a matriz de 2a's derivadas para tempos de censura
#-----
Iobsc <<- -cbind(rbind(d2Bc,d2Bsc),rbind(t(d2Bsc),d2sc))
}

# 0 Metodo Iterativo de Newton Raphson
# -----
nrapsion <- function(q0,beta0,sigma0)
{ i<- 0 ; q<- q0
  beta <- beta0
  sigma <- sigma0
  teta1 <- c(beta0,sigma0)
  teta <- 1.01*teta1
  while (max(abs(teta1-teta)) > 10e-4)
{ i <- i+1
  wf <<- gw(yf,beta,sigma,invXf) #( invXf =1000/Xf)
  derivaf(q,sigma,wf,invXf)
  if (nc>0)

```

```

      {wc <- gw(yc,beta,sigma,invXc) #(invXc =1000/Xc)
        derivac(q,sigma,wc,invXc)
        Uteta <- Utetaf + Utetac
        Iobs <- Iobsf + Iobsc
      }
    if (nc == 0)
{wc <- 0
  Uteta <- Utetaf
  Iobs <- Iobsf
}

  Iobs
  Iobsinv <- solve(Iobs)
  teta <- teta1
  if(length(teta)!=length(teta1)) stop("numero 1")
    if(ncol(Iobsinv)!=length(Uteta)) stop("numero 2")
    # *****>> Principal procedimento do Newton Raphson
#-----
  teta1 <- teta + Iobsinv %*% Uteta
  beta <- teta1[1:p]
  sigma <- teta1[p+1]
    print(cat("i= ",i,"q : ",q,"beta : ",beta,"sigma : ",sigma,"\n",
      "Vt. Escores : ", Uteta,"\n"))
    if (sigma < -1) stop("sigma<-1")
}

var.ass <- Iobsinv
  logLmaxq <- logL(q,sigma,wf,wc)
list(q=q,betaq=teta1[1:p], sigmaq=teta1[p+1],logLmaxq=logLmaxq)
}

```

```

# Funcao para calcular o maximo global de teta
#-----

maxteta <- function(qseq,beta0,sigma0)
{ lista <- NULL
  for (q in qseq)
  { nr <- nrphson(q,beta0,sigma0)
li <- c(q,nr$betaq,nr$sigmaq,nr$logLmaxq)
      lista <- rbind(lista,li)
beta0 <- nr$betaq ; sigma0 <-nr$sigmaq
  }

  logL <- lista[,s+4]
  maxlogL <- max(logL)
  maxq <-lista[logL==maxlogL,1]
  maxbeta <-lista[logL==maxlogL,2:(s+2)]
  maxsigma <-lista[logL==maxlogL,(s+3)]
  Rmax <- exp(logL-maxlogL)
  Lamb <- -2*log(Rmax)
  lista <<- cbind(lista,Rmax,Lamb)

  nrphson(maxq,maxbeta,maxsigma) # Para obter a var.ass do tetamax.
  list(maxq=maxq,maxbeta=maxbeta,maxsigma=maxsigma,maxlogL=maxlogL)
}

#.....
# Valores iniciais para o metodo iterativo: beta0 e sigma0 .
# "qs" e uma sequencia de valores dados.
#.....

qs <- seq(q1,q2,by=passo)

```

```
reg <- lm(y ~ X[,2:p])
beta0=coef(reg)
sigma0=sqrt(anova(reg)$"Mean Sq"[2])
maxteta(qs,beta0,sigma0)
```

Referências Bibliográficas

- [1] Abramowitz, M., e I. A. Stegun, Eds. (1965). *Handbook of Mathematical Functions*. New York: Dover.
- [2] Bartlett, M. S., e D. G. Kendall (1946). The statistical analysis of variance - heterogeneity and logarithmic transformation. *J. R. Stat. Soc. Suppl.*, **8**, 128-138.
- [3] Bartolucci, A. A., e J. M. Dickey (1977). Comparative Bayesian and traditional inference for gamma-modelled survival data. *Biometrics*, **33**, 343-354.
- [4] Becker, R. A., J. M. Chambers, e A.R. Wilks (1988). *The New S Language - A programming Environment for Data Analysis and Graphics*. California: Wadsworth & Brooks/Cole.
- [5] Billmann, B., e C. E. Antle (1972). Statistical inference from censored Weibull samples. *Technometrics*, **14**, 831-840.
- [6] Chambers, J. M. (1977). *Computational Methods for Data Analysis*. New York: Wiley.
- [7] Chiang, C. L. (1968). *Introduction to Stochastic Processes in Biostatistics*. New York: Wiley.
- [8] Cordeiro, G. M. (1992). *Introdução à Teoria da Verossimilhança*. Rio de Janeiro: Livro texto do 10º Simpósio Nacional de Probabilidade e Estatística.
- [9] Cox, D. R., e D. V. Hinkley (1974). *Theoretical Statistics*. London: Chapman & Hall.

- [10] Cox, D. R. e D. Oakes (1984). *Analysis of Survival Data*. Methuen. New York: Chapman & Hall.
- [11] Crow, E. L., e K. Shimizu. Eds. (1988). *Lognormal Distribution*. New York : Marcel Dekker.
- [12] Doganaksoy, N. e J.Schemee (1993). Comparisons of approximate confidence intervals for distributions used in life-data analysis. *Technometrics*, **35**, 175-184.
- [13] Diccio, T. J. (1987). Approximate inference for the generalized gamma distributions. *Technometrics*, **29**, 33-40.
- [14] Draper, N. R. e H. Smith, Jr. (1981). *Applied Regression Analysis 2. ed.* New York : Wiley.
- [15] Dumonceaux, R. e C.E. Antle (1973). Discrimination between the log-normal and the Weibull distributions. *Technometrics*, **15**, 923-926.
- [16] Dumonceaux, R., C.E. Antle e G. Hass (1973). Likelihood ratio test for discrimination between two models with unknown location and scale parameters. *Technometrics*, **15**, 19-27.
- [17] Dyer, A. R. (1973). Discrimination procedures for separate families of hypotheses. *J. Am. Stat. Assoc.*, **68**, 970-974.
- [18] Farewell, V. T. e R. L. Prentice (1977). A study of distributional shape in life testing. *Technometrics*, **19**, 69-75.
- [19] Feller, W. (1971). *An Introduction to Probability Theory. Vol II, 2nd ed.*. New York: Wiley.
- [20] Gross, A. J., e V. A. Clark (1975). *Survival Distributions - Reliability Applications in the Biomedical Sciences*. New York: Wiley.

- [21] Hager, H. W., e L.J. Bain (1970). Inferential procedures for the generalized gamma distributions. *J. Am. Stat. Assc.*, **65**, 1601-1609.
- [22] Harter, H. L. (1967). Maximum-likelihood estimation of the parameters of a four-parameter generalized gamma population for complete and censored samples. *Technometrics*, **9**, 159-165.
- [23] James, B. R. (1981). *Probabilidade : Um Curso em Nível Intermediário*. Rio de Janeiro: IMPA- Projeto Euclides.
- [24] Johnson, N. L. e S. Kotz (1970a). *Distributions in Statistics - Continuous Univariate Distributions, Vol I*. Boston: Houghton Mifflin.
- [25] Kalbfleisch, J. D., e R. I. Prentice (1980). *The Statistical Analysis of Failure Time Data*. New York: Wiley.
- [26] Kaplan, E. L. e P. Meier (1958). Nonparametric estimation from incomplete observations. *J. Am. Stat. Assoc.*, **53**, 457-481.
- [27] Kennedy, W. J. Jr., J.E. Gentle (1980). *Statistical Computing*. New York : Marcel Dekker.
- [28] Kotz, S. e N. L. Johnson (1985). *Encyclopedia of Statistical Sciences*. New York: Wiley.
- [29] Lawless, J. F. (1980). Inference in the generalized gamma and log-gama distributions. *Technometrics*, **22**, 409-419.
- [30] Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. New York: Wiley.
- [31] Lawless, J. F., e K. Singhal (1980). Analysis of data from life-test experiments under an exponential model. *Nav. Res. Log. Q.*, **27**, 323-334.

- [32] Lee, E. T. (1980). *Statistical Methods for Survival Data Analysis*. Lifetime Learning (Wadsworth). CA.: Belmont.
- [33] Montgomery, D. C. e Peck, E. A. (1992). *Introduction to Linear Regression Analysis*, 2. ed. New York: Wiley.
- [34] Mood, A. M., F. A. Graybill, D. C. Boes (1974). *Introduction to the Theory of Statistics*. New York: McGraw-Hill.
- [35] Nelson, W. (1982). *Applied Life Data Analysis*. New York: Wiley.
- [36] Nelson, W. (1990). *Accelerated Testing - Statistical Models Test Plans, and Data Analysis*. New York: Wiley.
- [37] Ostrouchov, G. e W. Q. Meeker, Jr. (1988). Accuracy of approximate confidence bounds computed from interval censored Weibull and lognormal data. *J. of Statist. Computation and Simulation*, **29**, 43-76.
- [38] Parr, V. B. e J. T. Webster (1965). A method for discriminating between failure density functions used in reliability predictions. *Technometrics*, **7** , 1-10.
- [39] Prentice, R. L. (1974). A log-gama model and its maximum likelihood estimation. *Biometrika*, **61** , 539-544.
- [40] Rao, A. V., R. R. Katam e V. L. Narasimham (1991). Linear estimation of location and scale parameters in the generalized gamma distribution. *Commun. Stat. - Theory Meth.*, **20(12)**, 3823-3848.
- [41] SAS Institute (1985). *SAS User's Guide: Statistics*, version 5 edition. Cary, NC: SAS Institute Inc.
- [42] Seber, G. A. F. (1977). *Linear Regression Analysis*. New York: Wiley.
- [43] Stacy, E. W. (1962). A generalization of the gamma distribution. *Ann. Math. Stat.*, **33**, 1187-1192.

- [44] Stacy, E. W. e G. A. Mihram (1965). Parameter estimation for a generalized gama distribution. *Technometrics*, **7**, 349-358.
- [45] Thoman, D. R., L. J. Bain, e C. E. Antle (1969). Inferences on the parameters of the Weibull distribution. *Technometrics*, **11**, 445-460.
- [46] Tobias, P. A. e D. Trindade (1986). *Applied Reliability*. New York: Van Nostrand Reinhold Co.
- [47] Whittemore, A. e B. Altschuler (1976). Lung cancer incidence in cigarette smokers: further analysis of Doll and Hill's data for British physicians. *Biometrics*, **32**, 805-816.
- [48] Vezon, D. J. e S. H. Moolgavkar (1988). A method for computing profile-likelihood-based confidence intervals. *Appl. Statist.*, **37**, 87-94.
- [49] Zelen, M. (1959). Factorial experiments in life testing. *Technometrics*, **1**, 269-288.