



UNIVERSIDADE FEDERAL DE LAVRAS

**ANÁLISE DO AJUSTE DE MODELOS  
UTILIZANDO OS RESÍDUOS  
GENERALIZADOS DE COX-SNELL**

**ANA LÚCIA SOUZA DA SILVA**

**2003**



55574

MF004426

**ANA LÚCIA SOUZA DA SILVA**

**ANÁLISE DO AJUSTE DE MODELOS UTILIZANDO OS RESÍDUOS  
GENERALIZADOS DE COX-SNELL**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-graduação em Agronomia, área de concentração em Estatística e Experimentação Agropecuária, para obtenção do título de "Mestre".

Orientador

Prof. Dr. Mário Javier Ferrua Vivanco

LAVRAS

MINAS GERAIS - BRASIL

2003

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da  
Biblioteca Central da UFLA**

**Silva, Ana Lúcia Souza**

**Análise do ajuste de modelos utilizando resíduos generalizados de Cox-Snell /  
Ana Lúcia Souza da Silva. -- Lavras : UFLA, 2003.**

**68 p. : il.**

**Orientador: Mário Javier Ferrua Vivanco.**

**Dissertação (Mestrado) – UFLA.**

**Bibliografia.**

**1. Estatística. 2. Resíduo generalizado de Cox-Snell. 3. Adequacidade de  
modelos. 4. Análise de resíduos. 5. Teste de hipótese. I. Universidade Federal de  
Lavras. II. Título.**

**CDD-519.56**

**ANA LÚCIA SOUZA DA SILVA**

**ANÁLISE DO AJUSTE DE MODELOS UTILIZANDO OS RESÍDUOS  
GENERALIZADOS DE COX-SNELL**

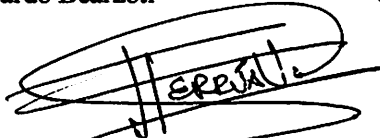
Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-graduação em Agronomia, área de concentração em Estatística e Experimentação Agropecuária.

**APROVADA em sexta-feira, 28 de fevereiro de 2003.**

Prof. Dr. Fortunato Silva de Menezes                      UFLA

Prof. Dr. Ruben Delly Veiga                                      UFLA

Prof. Dr. Eduardo Bearzoti                                      UFLA



Prof. Dr. Mário Javier Ferrua Vivanco  
UFLA  
(Orientador)

LAVRAS  
MINAS GERAIS - BRASIL

*A Deus,*

*por ter me oferecido esta oportunidade e me dado as forças necessárias para que este sonho se realizasse.*

### **OFEREÇO**

À minha inesquecível avó, **Maria José de Souza** ("*in memoriam*"), pelo amor, dedicação e carinho;

À minha maravilhosa mãe, **Vilma de Souza**, pelos ensinamentos de vida, amor e confiança;

Ao meu pai, **Ismel Silva**, pelo apoio e incentivo.

### **DEDICO**

## AGRADECIMENTOS

A Deus pela saúde, tranquilidade e força de vontade a mim concedidas para completar mais uma etapa de estudo;

À Universidade Federal de Lavras, em especial ao Departamento de Ciências Exatas, pela oportunidade de realizar este curso;

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES pela concessão de bolsa de pesquisa;

Ao coordenador do curso de mestrado em Agronomia – Estatística e Experimentação Agropecuária, Prof. Joel Augusto Muniz, pelo pleno critério e esforço frente ao programa e pela amizade e motivação durante todo o curso;

Ao Prof. Mário Javier Ferrua Vivanco pela orientação, amizade, compreensão e por sempre confiar na minha capacidade para a realização deste trabalho;

Ao Prof. Fortunato Silva de Menezes pela ajuda e colaboração essenciais para o andamento e conclusão deste trabalho;

Aos professores do Departamento de Ciências Exatas, Augusto, Daniel Furtado, Delly Veiga, Eduardo Bearzoti, Fortunato, Joel, Júlio, Lucas, Luis Henrique, Marcelo, Mário Vivanco, Paulo César e Telma, pelos ensinamentos fundamentais para minha formação profissional e humana;

Aos funcionários do Departamento de Exatas, Andréia, Edila, Ester, Rogéria e Maristela, em especial a Maria, pela simpatia, eficiência e atenção dispensada;

Aos colegas de curso, Alessandra, Carla, Cirilo, Élcio, Lúcia, Luis, João Marcos, José Ermelino, José Otaviano, Waldemar e Washington, em especial ao José Airton, pela colaboração no desenvolvimento deste trabalho, pela amizade, pelo companheirismo e incentivo.

Às grandes amigas, Polyanna e Deise, pela amizade e incentivo;

Ao meu namorado Wilson pela força, pelo carinho, pelo apoio, pelo amor e cumplicidade;

Ao meu querido pai pelo esforço, pela coragem e vontade de vencer que depositou em mim;

À minha querida mãe (“*in memoriam*”) pelo apoio, pelo amor eterno, pelo sacrifício e confiança;

À minha maravilhosa tia pela força, pela coragem e incentivo;

Aos meus queridos irmãos, Shaine e Luis Cláudio, pela união, pelo incentivo e carinho;

As minhas queridas sobrinhas, Gleice e Franciene, pela simpatia e pela alegria que nos proporciona;

Enfim, aos familiares e demais que contribuíram para a realização deste trabalho.



# SUMÁRIO

RESUMO.....	i
ABSTRACT .....	ii
1 INTRODUÇÃO .....	01
2 REFERENCIAL TEÓRICO .....	04
2.1 Análise dos Resíduos.....	04
2.2 Teste de Hipótese com Resíduos.....	05
2.2.1 Ausência de Normalidade dos Erros.....	06
2.2.2 Não Independência dos Erros.....	09
2.2.3 Diagnósticos de Observações Individuais.....	12
2.2.4 Heterogeneidade de Variâncias ...	16
2.3 Resíduos na Análise de Sobrevida.....	18
2.3.1 Resíduos de Cox-Snell.....	18
2.3.2 Resíduos Modificados de Cox-Snell.....	19
2.3.3 Resíduos de Martingale.....	20
2.3.4 Resíduos Score .....	21
2.3.5 Resíduos Deviance.....	22
3 RESÍDUOS GENERALIZADOS DE COX-SNELL.....	23
3.1 Definição Geral .....	23
3.2 Vício para o Estimador de Máxima Verossimilhança .....	25
3.2.1 Caso de um único parâmetro.....	25
3.2.2 Caso multiparamétrico .....	33
3.3 Expressão Geral para os Resíduos Generalizados de Cox-Snell .....	40
3.4 Resíduos Modificados.....	45
3.5 Resíduos de Cox-Snell para modelos de Análise de Sobrevida.....	48

4 APLICAÇÕES DOS RESÍDUOS GENERALIZADOS PARA TESTAR A ADEQUACIDADE DE AJUSTE DE MODELOS .....	51
4.1 Aplicação 01.. .....	51
4.2 Aplicação 02 .....	56
4.3 Aplicação 03.....	60
5 CONCLUSÃO .....	65
REFERÊNCIAS BIBLIOGRÁFICAS .....	66

## RESUMO

SILVA, Ana Lúcia Souza da. *Análise do ajuste de modelos utilizando os resíduos generalizados de Cox-Snell*. 2003. 68 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras, MG.\*

A partir da adoção de um modelo para explicar um conjunto de dados, o pesquisador está aceitando as hipóteses de que os dados em questão satisfazem às pressuposições inerentes ao modelo proposto. Desse modo, torna-se importante a análise da adequação do modelo aos dados, para que os resultados, em termos de inferências, sejam confiáveis. Uma das técnicas usadas para essa verificação é a análise de resíduos. Vários tipos diferentes de resíduos têm sido propostos para modelos de sobrevivência; dentre eles, os mais adequados são resultado dos resíduos generalizados de Cox-Snell (1968). O objetivo deste trabalho é avaliar a adequacidade de modelos por meio de gráficos de diagnósticos gerados a partir dos resíduos generalizados de Cox-Snell e mostrar detalhadamente todo o desenvolvimento para o entendimento e aplicação desta teoria. Para ilustrar a teoria foram feitas três aplicações. A primeira aplicação visou ilustrar a lógica existente entre a plotagem dos resíduos ordenados de três distribuições, normal (0,1), logística (0,1) e valor extremo (0,1) versus as estatísticas de ordem esperadas destes resíduos de acordo com as distribuições assumidas. Para a segunda aplicação, foram utilizados dados de tempo de vida de isolantes, obtidos em Nelson (1990). A partir da verificação por meio dos gráficos de diagnósticos utilizando os resíduos generalizados de Cox-Snell, encontrou-se que o modelo apropriado para o tempo de vida dos isolantes era o log-normal. Para a terceira aplicação, foram analisados dados censurados referentes ao tempo de vida de pacientes, obtidos em Collett (1994). Avaliou-se a adequacidade de vários modelos por meio dos resíduos de Cox-Snell adaptados para dados de sobrevivência. Os resultados mostraram que o modelo Weibull foi o mais adequado.

---

\* Comitê Orientador: Prof. Dr. Mário Javier Ferrua Vivanco - UFLA (Orientador), Prof. Dr. Fortunato Silva de Menezes.

## ABSTRACT

SILVA, Ana Lúcia Souza da. **Fitting analysis of models using Cox-Snell generalized residuals.** 2003. 68 p. Dissertation (Master in Statistics and Agricultural Experimentation) – Universidade Federal de Lavras, Lavras, MG.\*

Once a model is chosen to explain a data set, the researcher is accepting the hypotheses that such data satisfy inherent assumptions of the proposed model. Then it, becomes important the evaluation of the fitted model to the data, so that the results, be reliable. One of the techniques used for such evaluation is residual analysis. Several kinds of residuals have been proposed for survival models; among those, the most suitable for this purpose are Cox-Snell (1968) generalized residuals. The objective of this work was to evaluate the adequacy of models by graphical diagnostics using Cox-Snell generalized residuals and to show the whole development in full detail for understanding and application of this theory. To illustrate the theory three applications were considered. The first application sought to illustrate the existent logic by plotting of ordered residuals of three distributions: normal (0,1), logistics (0,1) and extreme value (0,1), versus the expected order statistics of these residuals in consonance with the assumed distributions. The second application consisted of lifetime data of electric insulating, obtained by Nelson (1990). Starting from the verification by graphical diagnostics using Cox-Snell generalized residuals, it was found that the model appropriate for lifetime of electric insulating was log-normal. The third application referred to censored data of lifetime of patients, obtained by Collett (1994). The adequacy of several models was evaluated by Cox-Snell generalized residuals adapted for survival data. The results show that Weibull model was the most appropriate.

---

\* Guidance Committee: Mário Javier Ferrua Vivanco - UFLA (Major Professor), Fortunato Silva de Menezes - UFLA.

# 1 INTRODUÇÃO

De maneira geral, a avaliação da adequacidade de ajuste de modelos a um conjunto de observações tem grande importância para validar a escolha do modelo utilizado. Muitos procedimentos para esta checagem são baseados em quantidades conhecidas como resíduos.

Na análise de regressão, por exemplo, a partir da análise dos resíduos pode-se detectar várias falhas no modelo tais como pressuposições distribucionais admitidas, a omissão incorreta do intercepto, ou a necessidade de termos de ordem superiores para a variável preditora.

Especificamente em Análise de Sobrevida, existem duas classes de modelos propostas na literatura: os modelos paramétricos e os modelos semi-paramétricos, que também podem ser chamados de modelos de regressão de Cox. Dentro da classe dos modelos paramétricos, encontra-se a classe dos modelos de Tempo de Falha Acelerados (AFT), que são dados pela seguinte expressão:

$$\text{Log}T_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \sigma \varepsilon_i \quad (1)$$

para  $i = 1, 2, \dots, n$ , em que  $T_i$  é o  $i$ -ésimo tempo até a ocorrência do evento correspondente à variável dependente,  $\beta_0, \beta_1, \dots, \beta_k$  e  $\sigma$  são os parâmetros desconhecidos,  $X_{ik}$  é a  $i$ -ésima observação correspondente à  $k$ -ésima variável independente e  $\varepsilon_i$  é o erro aleatório.

Existem diversas propostas para avaliar a adequacidade destes modelos na literatura. Allison (1995) apresenta duas propostas: o estudo da adequacidade

a partir do modelo Gama Generalizado e o estudo da adequacidade a partir dos gráficos de diagnósticos, isto é, por meio da análise de resíduos.

Várias espécies diferentes de resíduos têm sido propostos para modelos de sobrevivência, como os resíduos generalizados de Cox-Snell, os resíduos Deviance, os resíduos Martingale e os resíduos Score. Collett (1994) sugere como uma alternativa mais apropriada o uso dos Resíduos Generalizados de Cox-Snell.

O objetivo deste trabalho é avaliar a adequacidade de modelos por meio dos gráficos de diagnósticos gerados a partir dos resíduos generalizados de Cox-Snell e mostrar detalhadamente toda a abordagem matemática e estatística matemática que é envolvida para o entendimento e aplicação desta teoria.

Para alcançar os objetivos acima, no capítulo 2 será apresentado uma revisão de como se testa a adequacidade de modelos com os resíduos clássicos. Nesta seção serão apresentados vários testes de hipóteses utilizados para detectar problemas em relação ao modelo proposto e definições de alguns dos principais resíduos usados em Análise de Sobrevivência. No capítulo 3 será abordado todo o desenvolvimento da teoria dos resíduos generalizados, assim como o procedimento para a obtenção de sua expressão geral, a interpretação da análise gráfica destes resíduos e suas principais propriedades e aplicações. Também será determinada a forma particular destes resíduos, sugerida por Allison (1995), quando se trabalha com dados censurados.

No capítulo 4 serão apresentados os resultados de três aplicações por meio de gráficos; a primeira aplicação objetivou ilustrar a lógica da plotagem dos resíduos ordenados originados de três distribuições versus as estatísticas de ordem esperadas de acordo com as distribuições assumidas; a segunda aplicação envolve dados de tempo de vida de isolantes, obtidos por Nelson (1990), cujo modelo proposto será avaliado por meio dos resíduos generalizados de Cox-Snell. Para a terceira aplicação serão analisados dados censurados referentes ao

tempo de vida de pacientes, conforme Collett (1994), cujo modelo também será avaliado por meio dos resíduos de Cox-Snell adaptados para dados de sobrevivência.

No capítulo 5, serão apresentadas as conclusões mais importantes deste trabalho.

## 2 REFERENCIAL TEÓRICO

Embora este trabalho tenha como objetivo avaliar a adequacidade de modelos através dos Resíduos Generalizados de Cox-Snell, neste capítulo será apresentado uma revisão geral de como são utilizados os resíduos na análise de regressão.

### 2.1 Análise dos Resíduos

Considere-se um modelo de regressão linear dado por:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i \quad (2)$$

para  $i = 1, 2, \dots, n$ , em que  $Y_i$  é a  $i$ -ésima observação correspondente à variável aleatória dependente,  $\beta_0, \beta_1, \dots, \beta_k$  são os parâmetros desconhecidos,  $X_{ik}$  é a  $i$ -ésima observação correspondente à  $k$ -ésima variável aleatória independente e  $\varepsilon_i$  é o erro experimental. Sendo assim, os erros podem ser dados pela seguinte equação:

$$\varepsilon_i = Y_i - \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}$$

Quando se estima a equação de regressão dada por (2), o  $i$ -ésimo valor estimado, denotado por  $\hat{Y}_i$ , é dado por:  $\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_k X_{ik}$  para  $i = 1, 2, \dots, n$ , em que  $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$  são os respectivos estimadores de  $\beta_0, \beta_1, \dots, \beta_k$ . Dessa forma, os resíduos estimados, denotados por  $e_i$ , poderão



ser expressos da seguinte forma:  $Y_i - \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_k X_{ik}$ .

Conforme Anscombe & Tukey (1963),  $e_i = Y_i - \hat{Y}_i$ .

Os resíduos,  $e_i$ , contêm informações que permitem saber se o modelo proposto ajusta-se bem ou não aos dados. Então, deve-se checar os resíduos e observar qualquer peculiaridade que possa ter ocorrido no ajuste do modelo de regressão.

Segundo Parente (1984), de um modo geral a análise de resíduos tem a finalidade de detectar os seguintes problemas em relação ao modelo proposto:

- i. Se o modelo matemático proposto não é apropriado;
- ii. Se os erros não têm distribuição normal;
- iii. Se os erros não são independentes;
- iv. A presença de dados discrepantes (outliers);
- v. Se os erros não têm variância constante.

Vários métodos para examinar os resíduos depois de uma análise de variância convencional ou ajustando quadrados mínimos têm sido desenvolvidos. Uma variedade das técnicas mais utilizadas são apresentadas em Anscombe & Tukey (1963).

O exame de resíduos pode ser feito por gráficos básicos (análise gráfica dos resíduos) ou por métodos analíticos. Para a confirmação ou não da análise gráfica dos resíduos podem ser aplicados testes estatísticos.

Conforme Parente (1984), a análise gráfica de resíduos apresenta, sobre a inferência com resíduos, a vantagem de não depender da estrutura dos dados.

## 2.2 Testes de Hipóteses com Resíduos

Para checar os possíveis problemas em relação ao modelo proposto, foram desenvolvidos uma diversidade de testes baseados nos resíduos.

Segundo Seber (1977), Box et al. (1978) e Draper & Smith (1998), existem várias maneiras de plotar os resíduos. Porém, os autores aconselham

fazer a análise gráfica com os resíduos padronizados definidos por  $d_i = \frac{e_i}{S}$ ,

em que  $e_i$  é o  $i$ -ésimo resíduo e  $S$  é uma estimativa de sua variância. Assim, a dispersão dos  $e_i$  varia em sua quase totalidade de  $-3$  a  $3$ ; quando existem condições ideais para a análise dos dados, ou seja, quando os erros são independentes, apresentam variância constante, sobretudo quando é correta a suposição de normalidade dos erros.

Os testes de Hipóteses baseados nos resíduos são apresentados a seguir.

### 2.2.1 Ausência de Normalidade dos Erros

Shapiro & Wilk (1965) propuseram um teste de normalidade em que a estatística do teste, denotada por  $W$ , é obtida dividindo o quadrado de uma combinação linear apropriada dos valores esperados das estatísticas de ordem amostrais pelo quadrado dos desvios em relação à média. Inicialmente, associa-se um vetor de variáveis aleatórias  $\underline{y}$ , de modo que cada elemento será correspondente às amostras ordenadas do menor ao maior valor, ou seja,  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ , em que  $X_{(i)}$  denota a  $i$ -ésima estatística de ordem. A estatística do teste  $W$  é definida por:

$$W = \frac{b^2}{S^2},$$

em que:

$$S^2 = \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (X_{(i)} - \bar{X})^2 \quad e \quad b^2 = \left[ \sum_{i=1}^k a_{n-i+1} (y_{n-i+1} - y_i) \right]^2.$$

Os coeficientes  $(a_{n-i+1})$  são tabelados e o valor de  $k$  será dado da seguinte forma: se  $n$  for par,  $k = n/2$ , e se for ímpar,  $k = (n-1)/2$ . Definido o valor de  $k$ ,  $b^2$  pode ser dado por  $[a_n(y_n - y_1) + \dots + a_{k+2}(y_{k+2} - y_k)]^2$ , ou simplesmente  $[\sum_{i=1}^n a_i y_i]^2$ , pois por simetria a soma dos coeficientes é igual a 0.

Se o resultado da estatística  $W$  estiver na região crítica, estabelecida por valores tabelados ( $W_{\alpha,n}$ ), então pode-se dizer que os dados não se distribuem normalmente,  $\alpha$  é o nível de significância e  $n$  é o tamanho da amostra.

Por outro lado, Seber (1977) determinou que se  $d_{(1)} < d_{(2)} < \dots < d_{(n)}$  representam  $n$  resíduos padronizados ordenados definidos na seção 2, e para  $n$  moderadamente grande, um gráfico  $d_{(i)}$  versus  $(i-1/2)/n$ , no papel de probabilidade normal, detectará ausência ou não de normalidade. Quando os dados são normais, esses pontos devem estar dispersos em torno de uma linha reta.

Campos (1983) apresenta o teste de Kolmogorov-Smirnov (1933), através do qual adapta-se uma distribuição completamente conhecida  $F(x)$  a dados provenientes de uma distribuição desconhecida  $F_0(x)$ . Sendo:

$F(x)$ : proporção de valores esperados menores ou iguais a  $x$ ;

$S(x)$ : proporção de valores observados menores ou iguais a  $x$ ;

$F_0(x)$ : verdadeira função de distribuição de  $X$  e que é conhecida.

Observe-se que  $S(x)$  é uma função de distribuição empírica baseada na amostra aleatória  $X_1, X_2, \dots, X_n$ . O teste utiliza as seguintes estatísticas:

$$D = \sup_x |F(x) - S(x)|$$

$$D^+ = \sup_x [F(x) - S(x)]$$

$$D^- = \sup_x [S(x) - F(x)]$$

para testar, ao nível  $\alpha$  de significância:

$$H_0 : F(x) = F_0(x) \text{ versus}$$

$H_1 : F(x) \neq F_0(x), F(x) > F_0(x) \text{ ou } F(x) < F_0(x)$  , sendo que para pelo menos um valor de  $x$ , rejeita-se  $H_0$  se  $D \geq d$ ,  $D^+ \geq d_1$  ou  $D^- \geq d_1$  , respectivamente. Os valores de  $d$  e  $d_1$  são tabelados.

Lilliefors (1967) sugeriu uma modificação no teste de Kolmogorov-Smirnov, em que se ampliou o seu uso aos casos em que a média e a variância não são previamente especificadas, mas sim estimadas através dos dados da amostra, ou seja, a partir de

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad e \quad s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} .$$

Dessa forma, obtém-se a variável reduzida:

$$z_i = \frac{x_i - \bar{x}}{s}, \quad i = 1, 2, \dots, n .$$

A estrutura do teste é análoga ao teste de Kolmogorov-Smirnov, considerando os  $z_i$ , ao invés da variável original.

D'Agostino (1972) propôs um teste de normalidade baseado na estatística  $D$ , que envolve as estatísticas de ordem  $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ , obtidos a partir de uma amostra aleatória  $X_1, X_2, \dots, X_n$ . A estatística  $D$  é definida como:

$$D = T / (n^2 S),$$

em que  $T = \sum_{i=1}^n \{i - 1/2(n+1)\} X_{(i)}$ , e  $S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$ , sendo  $\bar{X}$  a média amostral. A estatística do teste para amostras maiores que 50 é denotado por  $Y$  e dada por:

$$Y = \frac{n^{1/2} (D - 0,28209479)}{0,02998598}.$$

A hipótese de nulidade, portanto de normalidade, deve ser rejeitada quando o valor calculado  $Y$  for maior que  $Y_{(n,\alpha)}$ , os quais são tabelados.

Draper & Smith (1998) consideram que uma alternativa para checar a normalidade dos erros é construir um gráfico de probabilidade normal ou construir um histograma com os resíduos.

### 2.2.2 Não Independência dos Erros

Campos (1983) apresenta dois testes de independência: o teste de Kendall e o teste de Spearman. O teste de Kendall serve para verificar o grau de dependência entre duas variáveis aleatórias  $X$  e  $Y$ . Considerando os  $n$  pares de observações  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ , provenientes de uma população contínua bivariada, pode-se formar um número  $C_n^2 = \frac{n(n-1)}{2}$  de combinações de produtos do tipo:

$$M_{ij} = (X_i - X_j)(Y_i - Y_j) \quad \begin{array}{l} i = 1, 2, \dots, n-1 \\ j = i+1, \dots, n \end{array}$$

O coeficiente de Kendall é dado por:

$$\tau = 2P(M_{ij} > 0) - 1$$

E as hipóteses a serem testadas são as seguintes:

$$H_0: \tau = 0 \quad \text{ou} \quad P(M_{ij} > 0) = 1/2 \quad \text{versus}$$

$$H_a: \tau > 0, \quad H_a: \tau < 0 \quad \text{ou} \quad H_a: \tau \neq 0$$

Rejeita-se  $H_0$  se  $K \geq \frac{n(n-1)}{2} - k(\alpha, n)$ ,  $K \leq k(\alpha, n)$  e se

$K \leq k(\alpha_1, n)$  ou  $K \geq \frac{n(n-1)}{2} - k(\alpha_2, n)$ , respectivamente.

Levando em conta os  $C_n^2$  produtos de  $M_{ij}$ , define-se a estatística  $K$  como o número de pares  $M_{ij} > 0$  (pares concordantes). Os valores de  $k(\alpha, n)$  são tabelados e  $\alpha_1 + \alpha_2 = \alpha$ ; geralmente, toma-se  $\alpha_1 = \alpha_2 = \alpha/2$ .

Finalmente, o teste de Spearman, que equivale ao teste de Pearson do campo paramétrico, é feito tomando-se  $n$  pares de observações  $(X_i, Y_i)$  provenientes de uma população contínua bivariada e, sendo  $R_i$  e  $S_i$  as ordens das observações  $X_i$  e  $Y_i$ , respectivamente, tem-se que o coeficiente de correlação do teste de Spearman é dado por:

$$r = \frac{\sum_{i=1}^n (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\left[ \sum_{i=1}^n (R_i - \bar{R})^2 \right] \left[ \sum_{i=1}^n (S_i - \bar{S})^2 \right]}}$$

em que  $\bar{R} = \frac{\sum_{i=1}^n R_i}{n}$  e  $\bar{S} = \frac{\sum_{i=1}^n S_i}{n}$ ,

que pode ser utilizado para testar as hipóteses:

$H_0$ : X e Y são independentes;

$H_1$ : X e Y são positivamente correlacionados, ou

$H_1$ : X e Y são negativamente correlacionados, ou

$H_1$ : X e Y são dependentes.

Assim podem-se realizar os seguintes testes:

a)  $H_0: \rho = 0$  vs  $H_a: \rho > 0$ , rejeitamos  $H_0$  se  $r \geq r(\alpha, n)$ ;

b)  $H_0: \rho = 0$  vs  $H_a: \rho < 0$ , rejeitamos  $H_0$  se  $r \leq -r(\alpha, n)$ ;

c)  $H_0: \rho = 0$  vs  $H_a: \rho \neq 0$ , rejeitamos  $H_0$  se  $r \leq -r(\alpha_2, n)$  ou  $r \geq r(\alpha_1, n)$ .

em que  $\alpha_1 + \alpha_2 = \alpha$  ou simplesmente  $\alpha_1 = \alpha_2 = \alpha/2$ . Os valores de  $r(\alpha, n)$  são tabelados. Para verificar a independência ou não dos erros, basta substituir X e Y pelos erros testados.

Anscombe & Tukey (1963), Box et al. (1978) e Draper & Smith (1998) aconselham que se existe alguma correlação dos erros no tempo, esta pode ser verificada sempre quando o tempo poder ser levado em conta (como variável regressora ou não). Esta verificação da correlação serial é examinada mediante plotagem dos resíduos versus o tempo. Quando os valores positivos dos resíduos são seguidos de um negativo ou vice-versa, evidencia-se a presença de correlação serial negativa, enquanto, na correlação serial positiva, sucessivos resíduos numa seqüência no tempo tendem a ser mais parecidos. Gráficos que detalham ambos os tipos de correlação (positiva e negativa) encontram-se em Draper & Smith (1998).

Um dos métodos utilizados para checar a dependência seqüencial dos erros é o teste de Durbin-Watson, apresentado por Draper & Smith (1998). O teste é baseado na seguinte estatística:

$$d = \frac{\sum_{u=2}^n (e_u - e_{u-1})^2}{\sum_{u=1}^n e_u^2} .$$

Compara-se a estatística  $d$  (ou  $4-d$ , que quase sempre é próximo de 0) com os valores tabelados de  $d_L$  e  $d_U$ . Se  $d < d_L$ , conclui-se que há possibilidade de uma correlação serial positiva; se  $d > d_U$ , conclui-se que existe um indicativo de não haver correlação. Se  $4-d < d_L$ , conclui-se que há possibilidade de uma correlação serial negativa; se  $4-d > d_U$ , conclui-se que existe um indicativo de não haver correlação. Caso os valores de  $d$  ou  $4-d$  estejam entre  $d_L$  e  $d_U$ , o teste é inconclusivo.

Por outro lado, Draper & Smith (1998) apresentam um teste de seqüências não paramétrico baseado no posto dos sinais dos resíduos.

O SAS (1996) realiza o teste de Durbin-Watson através do PROC REG com o comando Model /  $y=x$  / DW; já os outros testes, como o de Kendall e o de Spearman, são realizados por meio do PROC CORR.

### 2.2.3 Diagnósticos de Observações Individuais

Algumas técnicas mais sofisticadas são chamadas de diagnósticos que têm por objetivo analisar influências ou discrepâncias mediante exame dos resíduos estimados e das observações, os quais estão fora dos conjuntos padrões de outras observações ou daqueles que influenciam fortemente os resultados de regressão. Existem várias destas técnicas na literatura, dentre elas são citadas algumas neste trabalho.

Weisberg (1985) e Draper & Smith (1998) apresentam o projetor ortogonal, ou seja, a matriz  $H$  dada por  $X(X'X)^{-1} X'$ , que projeta o vetor  $Y$  no espaço gerado pelas colunas da matriz  $X$ , obtendo o vetor  $Y$  estimado:  $\hat{Y} = HY$ . O vetor de resíduos estimados, por sua vez, nada mais é do que a projeção no espaço ortogonal a este primeiro, ou seja,  $e = (I-H)Y$ . Esta grandeza indica a influência da observação.



O traço de H é igual a p, o qual corresponde a  $\sum_{i=1}^n h_{ii}$  ( $h_{ii}$  é o i-ésimo elemento da diagonal de H. São consideradas observações que influenciam aquelas cujo  $h_{ii}$  supera 2 ou 3 vezes a sua média, ou seja,  $\left(\frac{2p}{n}\right)$  ou  $\left(\frac{3p}{n}\right)$ , sendo que n é o número de observações e p é o número de parâmetros.

Cook & Weisberg (1986) e Draper & Smith (1998) apresentam os Resíduos Estudentizados Internamente dados por:

$$s_i = \frac{e_i}{\sqrt{(1-h_{ii})s^2}},$$

em que  $s^2$  é o quadrado médio do resíduo que pode ser estimado por:

$$\sigma^2 = \sum_{i=1}^n e_i^2 / (n-p).$$

São chamados assim pelo fato de  $e_i$  ter entrado no cálculo de  $s^2$ , estando presente “internamente” em  $s^2$  e indicam a discrepância da observação quando se obtêm valores de  $s_i$  muito altos.

Também Draper & Smith (1998) apresentam os Resíduos Estudentizados Externamente como:

$$t_i = \frac{e_i}{\sqrt{(1-h_{ii})s_{(t)}^2}},$$

em que  $S_{(i)}^2$  é dado por: 
$$S_{(i)}^2 = \frac{1}{n-p-1} \left[ (n-p)s^2 - \frac{e_i^2}{1-h_{ii}} \right].$$

É possível demonstrar que  $t_i$  tem distribuição  $t$ , pois está sendo dividido por uma estimativa da variância totalmente independente de  $e_i$  e provoca uma redução na soma de quadrado de resíduos. Tais resíduos ( $t_i$ ) indicam a discrepância da observação quando a estatística for maior que  $t_{(n-p-1, \alpha)}$ .

Belsley et al. (1980) definem dois critérios importantes para avaliar a influência ou discrepância de uma observação, o DFFITS e o DFBETAS. Considere a estatística DFFITS dada por:

$$DFFITS = \sqrt{\frac{h_{ii}}{1-h_{ii}}} t_i,$$

em que  $t_i$  é o resíduo studentizado externamente.

Alguns autores sugerem a regra para considerar as observações influentes ou discrepantes, para as quais  $|DFFITS| > 2\sqrt{\frac{p}{n}}$ , quando  $n$  é bem maior que  $p$ .

Cook (1977) sugere uma medida similar ao DFFITS para examinar a influência de uma observação.

Por outro lado, a estatística DFBETAS é definida como:

$$DFBETA_j = \frac{c_{ij}}{1 - h_{ii}} \frac{1}{\sqrt{w_j s_{(i)}^2}},$$

em que  $c_{ij}$  são os elementos da matriz  $X^+ = (X'X)^{-1} X'$ . Assim,  $(X'X)^{-1} x_j$  é a  $i$ -ésima coluna de  $X^+$  e  $w_j$  é o  $j$ -ésimo elemento da diagonal de  $(X'X)^{-1}$ . Deve-se

considerar como observações influentes, aquelas cujos  $|DFBETAS| > \frac{2}{\sqrt{n}}$ .

Conforme Cook (1977), a influência da  $i$ -ésima observação pode ser medida por informações combinadas a partir dos resíduos estudentizados, variâncias dos resíduos e valores preditos. A estatística é denotada por  $D_i$  e é definida como:

$$D_i = \frac{(\hat{Y}_i - \hat{Y}_{(i)})'(\hat{Y}_i - \hat{Y}_{(i)})}{ps^2},$$

em que  $(\hat{Y}_i - \hat{Y}_{(i)})$  é o comprimento normalizado do vetor  $Y$ .

Se a observação afeta pouco o ajustamento, esperam-se pequenos  $D_i' s$ . Como se trata de uma grandeza muito relacionada com o DFFITS, o usuário poderá escolher entre uma ou outra (Vide Draper & Smith (1998)).

Uma outra técnica muito usada é a de matriz de covariância dos coeficientes estimados, cuja a estatística definida por Belsley et al. (1980) é dada por:

$$COVRATIO = \frac{1}{\left[ \frac{n-p-1}{n-p} + \frac{e_i^2}{n-p} \right]^p (1-h_{ii})}$$

Esta estatística avalia as observações influentes mediante a seguinte relação: se  $|COVRATIO - 1| \geq 3p/n$ , a observação é considerada influente. Os  $e_i$  são os resíduos estudentizados.

O SAS (1996) realiza todos os diagnósticos acima por meio do PROC REG, com o comando INFLUENCE. Com o PROC GLM, são realizados todos os testes com exceção do DFBETAS.

#### 2.2.4 Heterogeneidade de Variâncias

Tendo-se um conjunto de resíduos referente a cada  $Y_i$ , Draper & Smith (1998) sugerem verificar o efeito do tempo, da heterogeneidade das variâncias, ou até mesmo verificar a necessidade de uma alguma transformação em Y (variável dependente) ou a curvatura do modelo proposto, com a plotagem dos resíduos verticalmente versus, respectivamente,

- i. A ordem do tempo dos dados, se conhecida;
- ii. Os correspondentes valores ajustados  $\hat{Y}_i$ , usando o modelo ajustado;
- iii. Os correspondentes valores de  $X_i$  se existir somente uma variável preditora; ou, em geral, cada conjunto de  $X_{ji}$ ,  $j = 1, 2, \dots, k$ , que representam as covariáveis.

Milliken & Johnson (1984) apresentam o teste F-max de Hartley, o qual requer que todas as amostras, a partir dos quais são estimadas as variâncias,

sejam do mesmo tamanho, isto é  $m_1 = m_2 = \dots = m_t = m$ . O teste é baseado na estatística:

$$F_{\max} = \frac{\max_i \{\hat{\sigma}_i^2\}}{\min_i \{\hat{\sigma}_i^2\}}.$$

Com esta estatística, rejeita-se  $H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_t^2$  se  $F_{\max} > F_{\max, \alpha, v}$ , em que  $v = m-1$  graus de liberdade e  $F_{\max, \alpha, v}$  é um valor tabelado. Vale a pena salientar que este teste é muito sensível à falta de normalidade.

Snedecor & Cochran (1980) apresentam o teste de Bartlett da seguinte forma: se existem  $k$  estimativas  $s_i^2$ , cada uma com o mesmo número de graus de liberdade  $n$ , a estatística do teste é dada por:

$$B = \frac{M}{C},$$

em que  $M = n(k \ln \bar{s}^2 - \sum_{i=1}^k \ln s_i^2)$ ,  $\bar{s}^2 = \sum_{i=1}^k s_i^2 / k$  e  $C = 1 + \frac{k+1}{3kn}$ .

A hipótese nula é a de que cada  $s_i^2$  é uma estimativa do mesmo  $\sigma^2$ . Logo, o quociente  $M/C$  é distribuído aproximadamente como uma distribuição de  $\chi^2$  com  $(k-1)$  graus de liberdade se esta hipótese for verdadeira.

Se  $\chi^2$  calculado ( $M/C$ ) for maior que  $\chi^2_{(k-1)}$  com  $\alpha$  de probabilidade ou significância, rejeita-se  $H_0 = \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$ .

A razão  $M/C$  segue uma distribuição aproximada de  $\chi^2$  somente quando os dados têm distribuição normal, sendo que este teste é muito sensível à falta de normalidade.

$L(\beta) = \prod_{j=1}^n p_j(y_j; \beta)$ ; com isto, tem-se que a log-verossimilhança é dada

por  $l(\beta) = \sum_{j=1}^n \log p_j(Y_j, \beta)$ .

Em seguida, expande-se a derivada primeira da função de verossimilhança  $L'(\hat{\beta})$ , em uma série de Taylor de primeira ordem na vizinhança de  $\beta$ . A série de Taylor de primeira ordem para uma  $f(x)$  na vizinhança de  $a$  será dada por:  $f(x) = f(a) + f'(a)(x - a)$ , em que

$f'(a) = \left. \frac{\partial f(x)}{\partial x} \right|_{x=a}$ . Logo, a expansão em série Taylor da derivada primeira

da função de verossimilhança (função escore)  $L'(\hat{\beta}) = 0$  é:

$$L'(\beta) + (\hat{\beta} - \beta)L''(\beta) = 0 \quad (12).$$

Fazendo  $U^{(j)} = \frac{d \log p_j(Y_j, \beta)}{d\beta}$  e  $V^{(j)} = \frac{d^2 \log p_j(Y_j, \beta)}{d\beta^2}$ , pode-se

expressar  $L'(\beta)$  e  $L''(\beta)$  como:

$$\begin{cases} L'(\beta) = \sum_{j=1}^n U^{(j)} = U^{(.)} \\ L''(\beta) = \sum_{j=1}^n V^{(j)} = V^{(.)} \end{cases}$$

Por definição, a Informação de Fisher,  $I$ , é dada por:

aos resíduos de Cox-Snell, apresentados por Allison (1995). Trata-se de uma forma particular da definição geral dos resíduos sugeridos por Cox-Snell (1968). O  $i$ -ésimo resíduo é definido por:

$$e_i = -\log \hat{S}(t_i | \underline{x}_i) \quad (3),$$

em que  $t_i$  é o tempo de evento observado ou tempo censurado para o indivíduo  $i$ ,  $\underline{x}_i$  é o vetor de covariáveis para o indivíduo  $i$ , e  $\hat{S}(t_i | \underline{x}_i)$  é a probabilidade estimada de sobreviver ao tempo  $t_i$  dado o vetor de covariáveis  $\underline{x}_i$ , baseado no modelo ajustado. Se  $t_i$  é um tempo censurado, então  $e_i$  é também tratado com uma observação censurada. É possível demonstrar que, se o modelo ajustado é correto, os  $e_i$ s têm aproximadamente uma distribuição exponencial com parâmetro  $\lambda = 1$ .

Collett (1994) apresenta outros resíduos utilizados em Análise de Sobrevivência, descritos a seguir:

### 2.3.2 Resíduos Modificados de Cox-Snell

Os resíduos modificados de Cox-Snell são obtidos a partir da equação (3), adicionando-se uma constante positiva ( $\Delta$ ), que pode ser chamada de resíduo excesso, o novo resíduo tomará a seguinte forma:

$$e_i = \begin{cases} e_i & \text{para observações não censuradas,} \\ e_i + \Delta & \text{para observações censuradas.} \end{cases}$$

O valor esperado de  $\Delta$  é 1 devido ao fato de que os resíduos em (3) têm uma distribuição exponencial unitária. Desta forma, tem-se que:

$$e_i' = \begin{cases} e_i & \text{para observações não censuradas,} \\ e_i + 1 & \text{para observações censuradas.} \end{cases}$$

O  $i$ -ésimo resíduo modificado de Cox-Snell pode ser expresso em uma forma alternativa pela introdução de um indicador de censura,  $\delta_i$ , que pode tomar o valor de 0 se o tempo de sobrevivência do  $i$ -ésimo indivíduo é censurado e 1 se ele não é censurado. Desta forma, os resíduos modificados de Cox-Snell serão dados por:

$$e_i' = 1 - \delta_i + e_i \quad (4).$$

Crowley & Hu (1977) encontraram, por meio de evidências empíricas, que a adição de 1 para os resíduos de Cox-Snell para observações censuradas inflaciona os resíduos e sugere que o valor mediano do resíduo excesso é preferível em relação ao valor médio. Então, a segunda versão dos resíduos modificados de Cox-Snell será:

$$e_i' = \begin{cases} e_i & \text{para observações não censuradas,} \\ e_i + 0,693 & \text{para observações censuradas.} \end{cases}$$

Entretanto, se a proporção de observações censuradas não é tão grande, o conjunto de resíduos obtidos a partir destas duas formas de modificações não parecerão tão diferentes.

### 2.3.3 Resíduos de Martingale

Para a obtenção destes resíduos, serão necessárias algumas modificações em (4); primeiramente, considera-se que os resíduos em (4) apresentam uma média zero quando uma observação não é censurada, em seguida, multiplica-se



este resultado por  $-1$ . Dessa forma, os resíduos de Martingale ficam definidos por:

$$e_{Mi} = \delta_i - e_i \quad (5).$$

Estes resíduos têm valores compreendidos entre  $-\infty$  e  $1$ , sendo negativos quando  $\delta_i = 0$ , com resíduos para observações censuradas. Eles apresentam propriedades similares às dos resíduos obtidos em Análise de Regressão Linear, porém não são simetricamente distribuídos sobre  $0$ . Em grandes amostras, os resíduos de Martingale são não correlacionados e têm um valor esperado de zero.

#### 2.3.4 Resíduos Score

Os resíduos Score não tratam de um simples valor do resíduo para cada indivíduo, mas de um conjunto de valores, um para cada covariável no modelo ajustado. O  $i$ -ésimo resíduo score para  $X_j$ , a  $j$ -ésima covariável no modelo, é dado por:

$$e_{Sji} = \delta_i \{x_{ji} - a_{ji}\},$$

em que  $x_{ji}$  é o valor da  $j$ -ésima covariável,  $j = 1, 2, \dots, p$ , para o  $i$ -ésimo indivíduo

em estudo, e  $a_{ji} = \frac{\sum_{l \in R(t_i)} x_{jl} \exp(\hat{\beta}' x_l)}{\sum_{l \in R(t_i)} \exp(\hat{\beta}' x_l)}$ , sendo que  $R(t_i)$  é o conjunto de

todos os indivíduos em risco para um tempo  $t_i$ ,  $x_{jl}$  é o valor da  $j$ -ésima

covariável,  $j = 1, 2, \dots, p$ , para o  $i$ -ésimo indivíduo e  $\hat{\beta}'$  é o vetor de estimativas dos coeficientes das covariáveis  $x_1, x_2, \dots, x_p$ .

### 2.3.5 Resíduos Deviance

Os resíduos Deviance foram introduzidos por Therneau et al. (1990). A motivação original desses resíduos é de que eles são componentes de uma deviance; esta estatística é utilizada como medida de ajustamento de modelos e é dada por:  $D = -2\{\log \hat{L}_c - \hat{L}_s\}$ , em que  $\hat{L}_c$  é a verossimilhança parcial do modelo corrente e  $\hat{L}_s$  é a verossimilhança parcial do modelo saturado ou completo. O resíduo Deviance para o  $i$ -ésimo indivíduo será dado por:

$$e_{Di} = \text{sgn}(e_{Mi}) \{-2[e_{Mi} + \delta_i \log(\delta_i - e_{Mi})]\}^{1/2},$$

em que  $e_{Mi}$  é o resíduo dado em (5) e  $\text{sgn}$  indica o sinal da função. O sinal de  $e_{Mi}$  indica que o resíduo Deviance terá o mesmo sinal do resíduo de Martingale, ou seja, +1 se for positivo e -1 se for negativo.

Estes resíduos podem ser simetricamente distribuídos sobre zero quando um modelo apropriado tiver sido ajustado e sua soma não necessariamente resultará em zero.

Na próxima seção será abordado, passo a passo, o desenvolvimento da teoria dos resíduos generalizados de Cox-Snell (1968), incluindo sua definição, a expressão geral, algumas de suas propriedades, aplicações destas e a demonstração de sua forma particular adaptada para dados de sobrevivência.

### 3 RESÍDUOS GENERALIZADOS DE COX-SNELL

Esta seção tem como objetivo encontrar uma expressão para os resíduos generalizados de Cox-Snell. Após a sua obtenção, será apresentada uma avaliação da adequacidade de modelos por meio de uma análise gráfica de resíduos.

A partir dos resultados de Cox-Snell foram desenvolvidas as demonstrações resultados.

#### 3.1 Definição Geral

Em geral, no contexto de modelos lineares da teoria normal, o vetor  $n \times 1$  de variáveis aleatórias  $\underline{Y}$  tem a seguinte forma:

$$\underline{Y} = X\underline{\beta} + \underline{\xi} \quad (6),$$

em que  $X$  é a matriz de modelo,  $\underline{\beta}$  é um vetor de parâmetros desconhecidos e  $\underline{\xi}$  é um vetor  $n \times 1$  de variáveis aleatórias não observáveis distribuídas normal e independentemente com média zero e variância constante. Se  $\tilde{\underline{\beta}}$  é o vetor de quadrados mínimos, estimador de  $\underline{\beta}$ , os resíduos  $\underline{R}^*$  são definidos por:

$$\underline{R}^* = X\tilde{\underline{\beta}} + \underline{Y} \quad (7).$$

Sob a condição de que o número de parâmetros é pequeno comparado com  $n$ , muitas das propriedades de  $\underline{R}^*$  poderão ser consideradas como as mesmas de  $\underline{\varepsilon}$ , isto é,  $\underline{R}^*$  deveria ter aproximadamente as propriedades de um vetor de variáveis aleatórias distribuídas normalmente com média zero e variância constante.

Assim como com o modelo dado em (6), pode-se considerar que em qualquer outro modelo, cada observação  $Y_i$  depende exclusivamente de  $\varepsilon_i$ , então, pode-se escrever:

$$Y_i = g_i(\underline{\beta}, \varepsilon_i) \quad (i=1, \dots, n) \quad (8).$$

Considerando que  $\hat{\underline{\beta}}$  é o estimador de máxima verossimilhança de  $\underline{\beta}$ , então a expressão (8) pode ser escrita como:

$$Y_i = g_i(\hat{\underline{\beta}}, R_i) \quad (9),$$

sendo  $R_i$  ( $i=1, 2, \dots, n$ ) o  $i$ -ésimo residual.

Se a função  $g_i$  for inversível, pode-se escrever os residuais como:

$$R_i = h_i(Y_i, \hat{\underline{\beta}}) \quad (10).$$

Note-se que os erros podem ser escritos por:

$$\varepsilon_i = h_i(Y_i, \beta) \quad (11);$$

além disso, não se está assumindo nenhuma distribuição para eles.

**Os residuais dados em (10) são chamados de Resíduos Generalizados de Cox-Snell.**

Freqüentemente o número de parâmetros é pequeno comparado com o número de observações, e todas as combinações relevantes de parâmetros são estimadas com pequenos erros padrões da ordem de  $n^{-1/2}$ . Sendo assim,  $R_i$  diferenciará do  $\varepsilon_i$  em uma quantidade  $n^{-1/2}$  em probabilidade, isto é,  $P(|R_i - \varepsilon_i| < n^{-1/2}) \rightarrow 1$ . Logo, as propriedades estatísticas dos  $R$ 's serão aproximadamente as mesmas propriedades dos  $\varepsilon$ 's.

Na próxima seção, será abordado todo o procedimento para determinar o vício para o estimador de máxima verossimilhança, com a finalidade de utilizá-lo para encontrar a expressão geral dos resíduos generalizados e suas propriedades.

### 3.2. Vício para o Estimador de Máxima Verossimilhança

#### 3.2.1. Caso de um único parâmetro

Seja  $Y_1, \dots, Y_n$  um conjunto de observações independentes, mas não necessariamente identicamente distribuídas com função de densidade  $p_j(y_j; \beta)$ ; então a função de verossimilhança é dada por:

$L(\beta) = \prod_{j=1}^n p_j(y_j; \beta)$ ; com isto, tem-se que a log-verossimilhança é dada

por  $l(\beta) = \sum_{j=1}^n \log p_j(Y_j, \beta)$ .

Em seguida, expande-se a derivada primeira da função de verossimilhança  $L'(\hat{\beta})$ , em uma série de Taylor de primeira ordem na vizinhança de  $\beta$ . A série de Taylor de primeira ordem para uma  $f(x)$  na vizinhança de  $a$  será dada por:  $f(x) = f(a) + f'(a)(x - a)$ , em que

$$f'(a) = \left. \frac{\partial f(x)}{\partial x} \right|_{x=a}. \text{ Logo, a expansão em série Taylor da derivada primeira}$$

da função de verossimilhança (função score)  $L'(\hat{\beta}) = 0$  é:

$$L'(\beta) + (\hat{\beta} - \beta)L''(\beta) = 0 \quad (12).$$

Fazendo  $U^{(j)} = \frac{d \log p_j(Y_j, \beta)}{d\beta}$  e  $V^{(j)} = \frac{d^2 \log p_j(Y_j, \beta)}{d\beta^2}$ , pode-se

expressar  $L'(\beta)$  e  $L''(\beta)$  como:

$$\begin{cases} L'(\beta) = \sum_{j=1}^n U^{(j)} = U^{(.)} \\ L''(\beta) = \sum_{j=1}^n V^{(j)} = V^{(.)} \end{cases}$$

Por definição, a Informação de Fisher,  $I$ , é dada por:

$$I = -E \left[ \frac{d^2 \ln L(Y; \beta)}{d\beta^2} \right] = -E[L''(\beta)] \quad (13).$$

Isolando-se o termo  $(\hat{\beta} - \beta)$  da equação (12) e substituindo  $-L''(\beta)$  por sua esperança, obtém-se:

$$\left\{ (\hat{\beta} - \beta) = \frac{L'(\beta)}{-L''(\beta)} = \frac{U^{(1)}}{-E[L''(\beta)]} = \frac{U^{(1)}}{I} \right. \quad (14).$$

A partir da Informação de Fisher, sabe-se que  $V(\hat{\beta}) = \frac{1}{I}$ .

O termo  $-L''(\beta)$  que corresponde à matriz de informação observada pode ser substituído por  $-E[L''(\beta)]$ , dita como a matriz de informação esperada ou de Fisher, pelo fato de que  $-L''(\beta)$  é o estimador de  $-E[L''(\beta)]$  e a variância é dada pela inversa da matriz de informação de Fisher.

As expressões obtidas em (14) podem ser determinadas de uma forma mais exata quando se expande  $L'(\hat{\beta})$  em uma série de Taylor de segunda ordem,

como a seguir:  $f(x) = f(a) + f'(a)(x - a) + f''(a) \frac{(x - a)^2}{2!}$ , em que

$$f''(a) = \left. \frac{\partial^2 f(x)}{\partial x^2} \right|_{x=a}, \text{ e obtém-se:}$$

$$L'(\beta) + (\hat{\beta} - \beta)L''(\beta) + \frac{1}{2}(\hat{\beta} - \beta)^2 L'''(\beta) = 0 \quad (15).$$

Sabendo-se que  $E(XY) = E(X)E(Y) + \text{cov}(X, Y)$  e calculando as esperanças em (15), tem-se:

$$E(\hat{\beta} - \beta)E[L''(\beta)] + \text{cov}\{(\hat{\beta} - \beta), L''(\beta)\} + \frac{1}{2}E[(\hat{\beta} - \beta)^2]E[L'''(\beta)] + \text{cov}\left\{\frac{1}{2}(\hat{\beta} - \beta)^2, L'''(\beta)\right\} = 0 \quad (16).$$

A  $E[L'(\beta)]$  é igual a 0, pois a esperança de uma função escore, denotada por  $E(U)$ , é igual a 0, segundo demonstrado em Cordeiro (1992), portanto,  $E(U^{(j)}) = 0 \quad \forall j$ .

O segundo termo da equação (16) é dado por:

$$\begin{aligned} \text{cov}\{(\hat{\beta} - \beta), L''(\beta)\} &= \text{cov}\left\{\frac{U^{(2)}}{I}, V^{(2)}\right\} = \\ &= \text{cov}\left\{\frac{\sum_{j=1}^n U^{(j)}}{I}, \sum_{j=1}^n V^{(j)}\right\} = \frac{1}{I} \text{cov}\left\{\sum_{j=1}^n U^{(j)}, \sum_{j=1}^n V^{(j)}\right\} = \\ &= \frac{1}{I} \{E[(\sum_{j=1}^n U^{(j)})(\sum_{j=1}^n V^{(j)})] - E(\sum_{j=1}^n U^{(j)})E(\sum_{j=1}^n V^{(j)})\} = \end{aligned}$$



$$\sum_n \sum_{j=1}^{j \neq i} \text{cov}(U^{(i)}, A^{(j)}) = 0$$

A explicação para que a expressão acima seja igual a 0 é que, de acordo com a definição da função de verossimilhança dada na seção 3.2.1, pode-se dizer que se  $Y_i$  e  $Y_j$  são observações independentes, logo  $P_i(V_i; \beta)$  é independente de  $P_j(V_j; \beta)$ . Dessa forma tem-se que  $U^{(i)}$  também é independente de  $A^{(j)}$ ; e considerando-se a suposição de independência, tem-se

$$\begin{aligned} & + \text{cov}(U^{(2)}, A^{(2)}) + \text{cov}(U^{(3)}, A^{(3)}) + \dots + \text{cov}(U^{(n)}, A^{(n-1)}) = 0 \\ & + \sum_n \sum_{j=1}^{j \neq i} \text{cov}(U^{(i)}, A^{(j)}) = \text{cov}(U^{(1)}, A^{(1)}) + \text{cov}(U^{(2)}, A^{(2)}) + \dots + \end{aligned}$$

em que:

$$- \sum_n \sum_{j=1}^{j \neq i} \text{cov}(U^{(i)}, A^{(j)}),$$

Pode-se observar que:

$$\text{cov}(U^{(1)}, A^{(1)}) + \text{cov}(U^{(2)}, A^{(2)}) + \dots + \text{cov}(U^{(n)}, A^{(n)}) -$$

$$\begin{aligned} & + E(U^{(n)})E(A^{(n)}) \\ & + E(U^{(2)})E(A^{(2)}) + \dots + E(U^{(1)})E(A^{(1)}) + E(U^{(n)})E(A^{(2)}) + \dots + \\ & + E(U^{(1)})E(A^{(n)}) + E(U^{(2)})E(A^{(2)}) + \dots + \\ & + E(U^{(n)})E(A^{(1)}) - [E(U^{(1)})E(A^{(1)}) + E(U^{(2)})E(A^{(2)}) + \dots + \\ & + E(U^{(2)})E(A^{(2)}) + \dots + E(U^{(n)})E(A^{(1)}) + E(U^{(n)})E(A^{(2)}) + \dots + \\ & + E(U^{(1)})E(A^{(2)}) + \dots + E(U^{(1)})E(A^{(n)}) + E(U^{(2)})E(A^{(1)})] \end{aligned}$$

Logo,

$$\text{cov}(U^{(i)}, V^{(i)}) = \frac{1}{I} \left\{ \sum_{j=1}^n E(U^{(j)} V^{(j)}) - \sum_{j=1}^n [E(U^{(j)}) E(V^{(j)})] \right\}$$

$$\text{cov}(U^{(i)}, V^{(i)}) = \frac{J}{I} \quad (17),$$

em que  $\sum_{j=1}^n E(U^{(j)} V^{(j)}) = J$  e  $\sum_{j=1}^n [E(U^{(j)}) E(V^{(j)})] = 0$ , pois

$$E(U^{(j)}) = 0 \quad \forall j.$$

Já o terceiro termo em (16) pode ser expresso de uma forma mais simplificada, como será explicado a seguir. Para isto, faz-se

$$W^{(i)} = \frac{d^3 \log p_j(Y_j, \beta)}{d\beta^3}. \text{ Denotando-se } L''(\beta) \text{ por:}$$

$$L''(\beta) = \frac{\sum_{j=1}^n d^3 \log p_j(Y_j, \beta)}{d\beta^3} = W^{(i)} \text{ e calculando a esperança, pode-se}$$

escrever que:

$$E(W^{(i)}) = E[L''(\beta)] = K \quad (18).$$

Por definição, tem-se que:

$$E[(\hat{\beta} - \beta)^2] = \text{Var}(\hat{\beta}) = \frac{1}{I} \quad (19).$$

Substituindo (18) e (19) no terceiro termo dado em (16), pode-se escrever que:

$$\frac{1}{2} E[(\hat{\beta} - \beta)^2] E[L'''(\beta)] = \frac{K}{2I} \quad (20).$$

O último termo em (16) é desprezado, pois ele é no máximo da ordem  $n^{-1}$ , ou simplesmente ordem de magnitude  $n^{-1}$ , denotado por  $O(n^{-1})$ . Por definição, se “f” e “g” são duas funções reais de variável real, que assume  $g(n) > 0$  para um n suficientemente grande, diz-se que f(n) é no máximo da

ordem g(n) se  $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} < c$ ; sendo assim, f(n) é no máximo da  $O(n^{-1})$ , se

$\lim_{n \rightarrow \infty} \frac{f(n)}{(1/n)} < c$ , em que c é uma constante maior que 0. À medida que n

crece f(n) tende rapidamente a zero e f(n) é da  $O(n)$  se  $\lim_{n \rightarrow \infty} \frac{f(n)}{n} < c$ .

Dessa forma, aproximadamente pela expressão (14), tem-se que:

$$\begin{aligned} \text{cov} \left\{ \frac{1}{2} (\hat{\beta} - \beta)^2, L'''(\beta) \right\} &= \frac{1}{2} \text{cov} \{ (\hat{\beta} - \beta)^2, L'''(\beta) \} = \\ &= \frac{1}{2} \text{cov} \left\{ \left[ \frac{U^{(i)}}{I} \right]^2, W^{(i)} \right\} = \frac{1}{2} \frac{1}{I^2} \text{cov} \{ [U^{(i)}]^2, W^{(i)} \} = \\ &= \frac{1}{2I^2} \left\{ E \left[ \left( \sum_{j=1}^n U^{(j)} \right)^2 \left( \sum_{j=1}^n W^{(j)} \right) \right] - \left[ E \left( \sum_{j=1}^n U^{(j)} \right)^2 \right] E \left( \sum_{j=1}^n W^{(j)} \right) \right\} = \\ &= \frac{1}{2I^2} \left\{ E \left[ (U^{(1)^2} + U^{(2)^2} + \dots + U^{(n)^2} + 2 \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n U^{(i)} U^{(j)}) \right] \right\} \end{aligned}$$

$$(W^{(1)}, W^{(2)}, \dots, W^{(n)})] - [(EU^{(1)^2} + EU^{(2)^2} + \dots + EU^{(n)^2} + \\ + 2 \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n EU^{(i)} EU^{(j)}) (EW^{(1)}, EW^{(2)}, \dots, EW^{(n)})] \left. \right\}$$

Como  $E(U^{(j)}) = 0 \forall j$ , logo  $2 \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n E(U^{(i)}) E(U^{(j)}) = 0$  e

$$E \left[ 2 \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n U^{(i)} U^{(j)} W^{(i)} \right] = 2 \left[ E \sum_{i=1}^n (U^{(i)} W^{(i)}) E \sum_{j=1}^n (U^{(j)}) \right] = 0.$$

Dessa forma,

$$\Rightarrow \text{cov} \left\{ \frac{1}{2} (\hat{\beta} - \beta)^2, L^m(\beta) \right\} = \frac{1}{2} \frac{1}{I^2} \left\{ \text{cov}(U^{(1)^2}, W^{(1)}) + \dots + \right. \\ \left. \text{cov}(U^{(n)^2}, W^{(n)}) \right\} \quad (21).$$

Como  $I = \sum_{i=1}^n E(-V^{(j)})$ , de acordo com (13), e  $\lim_{n \rightarrow \infty} \frac{V^{(1)}}{1} = V^{(1)}$ ,

pode-se dizer que  $V^{(1)}$  é da ordem 1, ou simplesmente,  $O(1)$ . Pela primeira propriedade de ordem de magnitude apresentada por Rohatgi (1976), a soma de duas funções de ordem de magnitude é igual à soma de suas ordens de magnitude, logo  $I$  é de  $O(n)$ . Pela segunda propriedade, o produto de duas funções de ordem de magnitude é igual ao produto de suas ordens de magnitude; com isto,  $I^2$  é da  $O(n^2)$ .

Seguindo a mesma idéia, como  $\text{cov}(U^{(1)^2}, W^{(1)}) = E(U^{(1)^2} W^{(1)})$ ,

pois  $(EU^{(1)2}EW^{(1)})=0$ ,  $U^{(1)2}$  é de  $O(1)$  e  $W^{(1)}$  também; sendo assim, toda a covariância é de  $O(1)$  e o somatório de todas as covariâncias é de  $O(n)$ .

Portanto, se o numerador de (21) é de  $O(n)$  e o denominador é de  $O(n^2)$

a expressão  $\text{COV}\left\{\frac{1}{2}(\hat{\beta} - \beta)^2, L''(\beta)\right\}$  é de  $O(n^{-1})$ .

Substituindo (13), (17) e (20) em (16) tem-se que:

$$E(\hat{\beta} - \beta) \cdot (-I) + \frac{J}{I} + \frac{K}{2I} = 0 \Rightarrow E(\hat{\beta} - \beta) = \frac{J}{I^2} + \frac{K}{2I^2}.$$

Dessa forma, o vício do estimador de máxima verossimilhança de  $(\beta)$  será dado por:

$$b = E(\hat{\beta} - \beta) = \frac{1}{2I^2}(K + 2J) \quad (22).$$

### 3.2.2. Caso multiparamétrico

Para encontrar um vício para o estimador de máxima verossimilhança multiparamétrico, será apresentado a mesma metodologia utilizada na seção 3.2.1, envolvendo agora mais parâmetros  $\beta_1, \dots, \beta_p$ .

Para isto, expande-se a equação  $\left[\frac{\partial L}{\partial \beta_r}\right]_{\beta=\hat{\beta}} = 0$  em uma série de

Taylor de primeira ordem para o caso multivariado, dada por:

$$f(x_1, \dots, x_n) = f(a_1, \dots, a_n) + f'_{x_1}(a_1, \dots, a_n)(x_1 - a_1) + \dots + f'_{x_n}(a_1, \dots, a_n)(x_n - a_n) \quad , \text{ em que}$$

$$f'_{x_i}(a_1, \dots, a_n) = \frac{\partial f(x_1, \dots, x_n)}{\partial x_i} \bigg|_{\substack{x_1 = a_1 \\ \vdots \\ x_n = a_n}} \quad , \quad i = 1, 2, \dots, n.$$

A expansão de Taylor de primeira ordem da equação  $\left[ \frac{\partial L}{\partial \beta_r} \right]_{\beta = \hat{\beta}} = 0$

para o caso multivariado é dada por:

$$\frac{\partial L}{\partial \beta_r} + (\hat{\beta}_s - \beta_s) \sum_{r=1}^p \frac{\partial^2 L}{\partial \beta_r \partial \beta_s} = 0 \quad (23).$$

Denominando

$$U_r^{(j)} = \frac{\partial \log p_j(Y_j, \beta)}{\partial \beta_r} \quad \text{e} \quad V_{rs}^{(j)} = \frac{\partial^2 \log p_j(Y_j, \beta)}{\partial \beta_r \partial \beta_s} \quad ,$$

pode-se escrever  $\frac{\partial L}{\partial \beta_r}$  e  $\frac{\partial^2 L}{\partial \beta_r \partial \beta_s}$  por:

$$\frac{\partial L}{\partial \beta_r} = \sum_{j=1}^n U_r^{(j)} = U_r^{(\cdot)} \quad \text{e} \quad \frac{\partial^2 L}{\partial \beta_r \partial \beta_s} = \sum_{j=1}^n V_{rs}^{(j)} = V_{rs}^{(\cdot)}$$

Seguindo a definição da matriz de informação de Fisher dada em (13), tem-se que:

$$I_{rs} = -E \left[ \frac{\partial^2 L}{\partial \beta_r \partial \beta_s} \right] \quad (24).$$

Dessa forma, isolando-se o termo  $(\hat{\beta}_s - \beta_s)$  em (23), e substituindo

$\frac{\partial^2 L}{\partial \beta_r \partial \beta_s}$  por sua esperança, obtém-se:

$$(\hat{\beta}_s - \beta_s) = \sum_{r=1}^p \left( \frac{U_r^{(i)}}{-\frac{\partial^2 L}{\partial \beta_r \partial \beta_s}} \right) = \sum_{r=1}^p \left( \frac{U_r^{(i)}}{-E \left( \frac{\partial^2 L}{\partial \beta_r \partial \beta_s} \right)} \right).$$

Logo,

$$(\hat{\beta}_s - \beta_s) = \sum_{r=1}^p I^{rs} U_r^{(i)} \quad (25).$$

Da mesma forma realizada na seção anterior,  $\frac{\partial^2 L}{\partial \beta_r \partial \beta_s}$  será substituído

por  $-E \left[ \frac{\partial^2 L}{\partial \beta_r \partial \beta_s} \right]$ . E por definição, a matriz de covariâncias é obtida pela

inversa da matriz de informação ( $I^{rs}$ ), denotada pelo subscrito

$$\text{cov}(\hat{\beta}_r - \beta_r) = I^{rs}.$$

Para obter uma expressão mais correta, pode-se, a partir de (25),

expandir  $\left[ \frac{\partial L}{\partial \beta_r} \right]_{\beta=\hat{\beta}} = 0$  em uma série de Taylor de segunda ordem para o

caso multiparamétrico, dada por:

$$\begin{aligned}
 f(x_1, \dots, x_n) = & f(a_1, \dots, a_n) + f'_{x_1}(a_1, \dots, a_n)(x_1 - a_1) + \dots + \\
 & + f'_{x_n}(a_1, \dots, a_n)(x_n - a_n) + \frac{1}{2!} \left[ f'_{x_1 x_1}(a_1, \dots, a_n)(x_1 - a_1)^2 + \dots + \right. \\
 & + f'_{x_n x_n}(a_1, \dots, a_n)(x_n - a_n)^2 + 2f'_{x_1 x_2}(a_1, \dots, a_n)(x_1 - a_1)(x_2 - a_2) + \\
 & + \dots + 2f'_{x_1 x_n}(a_1, \dots, a_n)(x_1 - a_1)(x_n - a_n) + 2f'_{x_2 x_3}(a_1, \dots, a_n)(x_2 - a_2) \\
 & (x_3 - a_3) + \dots + 2f'_{x_2 x_n}(a_1, \dots, a_n)(x_2 - a_2)(x_n - a_n) + \dots + \\
 & \left. + 2f'_{x_{n-1} x_n}(a_1, \dots, a_n)(x_{n-1} - a_{n-1})(x_n - a_n) \right],
 \end{aligned}$$

em que  $f'_{x_i x_j}(a_1, \dots, a_n) = \frac{\partial^2 f(x_1, \dots, x_n)}{\partial x_i \partial x_j} \Big|_{\substack{x_1 = a_1 \\ \vdots \\ x_n = a_n}} \quad i=1,2,\dots,n \quad e$

$j=1,2,\dots,n$ .

A expansão de Taylor de segunda ordem da equação  $\left[ \frac{\partial L}{\partial \beta_r} \right]_{\beta=\hat{\beta}} = 0$

para o caso multivariado é dada por:



$$\begin{aligned} & \frac{\partial L}{\partial \beta_r} + (\hat{\beta}_s - \beta_s) \sum_{r=1}^p \frac{\partial^2 L}{\partial \beta_r \partial \beta_s} + \\ & + \frac{1}{2} (\hat{\beta}_t - \beta_t) (\hat{\beta}_u - \beta_u) \sum_{r=1}^p \frac{\partial^3 L}{\partial \beta_r \partial \beta_t \partial \beta_u} \end{aligned} \quad (26).$$

Sabendo-se que  $E(\underline{XY}') = E(\underline{X})E(\underline{Y}') + \text{cov}(\underline{X}, \underline{Y}')$  e aplicando-se esperança em (26), tem-se:

$$\begin{aligned} & E(\hat{\beta}_s - \beta_s) E \sum_{r=1}^p \frac{\partial^2 L}{\partial \beta_r \partial \beta_s} + \text{cov} \left[ (\hat{\beta}_s - \beta_s), \sum_{r=1}^p \frac{\partial^2 L}{\partial \beta_r \partial \beta_s} \right] + \\ & + \frac{1}{2} \left\{ E(\hat{\beta}_t - \beta_t) (\hat{\beta}_u - \beta_u) E \sum_{r=1}^p \frac{\partial^3 L}{\partial \beta_r \partial \beta_t \partial \beta_u} + \right. \\ & \left. + \text{cov} \left[ (\hat{\beta}_t - \beta_t) (\hat{\beta}_u - \beta_u), \sum_{r=1}^p \frac{\partial^3 L}{\partial \beta_r \partial \beta_t \partial \beta_u} \right] \right\} \end{aligned} \quad (27).$$

Como a esperança de uma função escore é igual a 0, logo  $E\left(\frac{\partial L}{\partial \beta_r}\right) = 0$ .

Da mesma forma, em relação a (17), o segundo termo da equação (27) é dado por:

$$\begin{aligned} & \text{cov} \left[ (\hat{\beta}_s - \beta_s), \sum_{r=1}^p \frac{\partial^2 L}{\partial \beta_r \partial \beta_s} \right] = \sum_{r=1}^p \text{cov} \left( \sum_{j=1}^n U_r^{(j)} I^{rs}, \sum_{j=1}^n V_{rs}^{(j)} \right) = \\ & = \sum_{r=1}^p I^{rs} \text{cov} \left( \sum_{j=1}^n U_r^{(j)}, \sum_{j=1}^n V_{rs}^{(j)} \right) = \\ & = \sum_{r=1}^p I^{rs} \left\{ E \left[ \left( \sum_{j=1}^n U_r^{(j)} \right) \left( \sum_{j=1}^n V_{rs}^{(j)} \right) \right] - E \left( \sum_{j=1}^n U_r^{(j)} \right) E \left( \sum_{j=1}^n V_{rs}^{(j)} \right) \right\} = \end{aligned}$$

$$\begin{aligned}
&= I^{rs} \{ [E(U_r^{(1)}V_{rs}^{(1)}) + E(U_r^{(1)}V_{rs}^{(2)}) + E(U_r^{(1)}V_{rs}^{(n)}) + \\
&+ E(U_r^{(2)}V_{rs}^{(1)}) + E(U_r^{(2)}V_{rs}^{(2)}) + \dots + E(U_r^{(2)}V_{rs}^{(n)}) + \\
&+ E(U_r^{(n)}V_{rs}^{(1)}) + E(U_r^{(n)}V_{rs}^{(2)}) + \dots + E(U_r^{(n)}V_{rs}^{(n)})] - \\
&- [E(U_r^{(1)})E(V_{rs}^{(1)}) + E(U_r^{(1)})E(V_{rs}^{(2)}) + \dots + E(U_r^{(1)})E(V_{rs}^{(n)}) + \\
&+ E(U_r^{(2)})E(V_{rs}^{(1)}) + E(U_r^{(2)})E(V_{rs}^{(2)}) + \dots + E(U_r^{(2)})E(V_{rs}^{(n)}) + \\
&+ \dots + E(U_r^{(n)})E(V_{rs}^{(1)}) + E(U_r^{(n)})E(V_{rs}^{(2)}) + \dots + \\
&+ E(U_r^{(n)})E(V_{rs}^{(n)})] \}.
\end{aligned}$$

Isto é,

$$\begin{aligned}
\text{cov}(U_r^{(i)}, V_{rs}^{(i)}) &= \text{cov}(U_r^{(1)}, V_{rs}^{(1)}) + \text{cov}(U_r^{(2)}, V_{rs}^{(2)}) + \dots + \\
&+ \text{cov}(U_r^{(n)}, V_{rs}^{(n)}) - \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n \text{cov}(U_r^{(i)}, V_{rs}^{(j)}).
\end{aligned}$$

Considerando a mesma idéia de suposição de independência abordada

na seção anterior, 
$$\sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n \text{cov}(U_r^{(i)}, V_{rs}^{(j)}) = 0.$$

Portanto,

$$\text{cov}(U_r^{(i)}, V_{rs}^{(i)}) = \sum_{s=1}^p I^{rs} \{ E[(\sum_{j=1}^n U_r^{(j)})(\sum_{j=1}^n V_{rs}^{(j)})] - E(\sum_{j=1}^n U_r^{(j)})E(\sum_{j=1}^n V_{rs}^{(j)}) \}$$

$$\text{cov}(U_r^{(i)}, V_{rs}^{(i)}) = \sum_{s=1}^p I^{rs} J_{r,rs},$$

em que 
$$\sum_{j=1}^n E(U_r^j, V_{rs}^j) = J_{r,rs} \quad \text{e} \quad E(\sum_{j=1}^n U_r^{(j)})E(\sum_{j=1}^n V_{rs}^{(j)}) = 0, \text{ pois}$$

$$E\left(\sum_{j=1}^n U_r^{(j)}\right) = 0.$$

Fazendo  $W_{ru}^{(j)} = \frac{\partial^3 \log p_j(Y_j, \beta)}{\partial \beta_r \partial \beta_t \partial \beta_u}$ , pode-se denotar  $\frac{\partial^3 L}{\partial \beta_r \partial \beta_t \partial \beta_u}$  da

seguinte forma:

$$\frac{\partial^3 L}{\partial \beta_r \partial \beta_t \partial \beta_u} = \sum_{j=1}^n W_{ru}^{(j)} = W^{(*)}; \text{ logo, } E(W^{(*)}) = K_{ru}$$

O terceiro termo em (27), é dado de acordo com (19) por:

$$\frac{1}{2} \left\{ E(\hat{\beta}_t - \beta_t)(\hat{\beta}_u - \beta_u) E \sum_{r=1}^p \frac{\partial^3 L}{\partial \beta_r \partial \beta_t \partial \beta_u} \right\} = \frac{1}{2} \sum_{r=1}^p K_{ru} I^{tu} \quad (29),$$

em que  $E[(\hat{\beta}_t - \beta_t)(\hat{\beta}_u - \beta_u)] = \text{cov}(\hat{\beta}_t, \hat{\beta}_u) = I^{tu}$  e  $E \frac{\partial^3 L}{\partial \beta_r \partial \beta_t \partial \beta_u} = K_{ru}$ .

O último termo da equação (27) é desprezado, sendo de  $O(n^{-1})$ .

Fazendo-se as devidas substituições em (27) e generalizado obteve-se a

seguinte equação:  $E(\hat{\beta}_s - \beta_s) I_{rs} = \frac{1}{2} \sum_{r=1}^p \sum_{t=1}^p \sum_{u=1}^p I^{tu} (K_{ru} + 2J_{t,ru})$ .

Finalmente, o vício para o estimador de máxima verossimilhança multiparamétrico, ou seja, para o parâmetro  $\beta_s$ , será:

$$b_s = E(\hat{\beta}_s - \beta_s) = \sum_{r,t,u} \left[ \frac{1}{2} I^{rs} I^{tu} (K_{ru} + 2J_{t,ru}) \right] \quad (30).$$

### 3.3 Expressão Geral para os Resíduos Generalizados de Cox-Snell

Para a obtenção da expressão geral dos resíduos, será feita uma expansão de  $(R_i - \varepsilon_i)$ , definidos em (10) e (11), em uma série de Taylor de segunda ordem para o caso bivariado em termos de  $(\hat{\beta}_s - \beta_s)$ , que é dada por:

$$f(x, y) = f(a, b) + f'_x(a, b)(x - a) + f'_y(a, b)(y - b) + \\ + \frac{1}{2!} [f''_{xx}(a, b)(x - a)^2 + f''_{yy}(a, b)(y - b)^2 + \\ + f''_{xy}(a, b)(x - a)(y - b)],$$

em que  $f'_x(a, b) = \left. \frac{\partial f(x, y)}{\partial x} \right|_{\substack{x=a \\ y=b}}$  e  $f''_{xx}(a, b) = \left. \frac{\partial^2 f(x, y)}{\partial x^2} \right|_{\substack{x=a \\ y=b}}$ , semelhantemente para os

outros. O resultado é dado por:

$$R_i = \varepsilon_i + \frac{\partial h_i(Y_i, \beta)}{\partial \beta_1} (\hat{\beta}_1 - \beta_1) + \dots + \frac{\partial h_i(Y_i, \beta)}{\partial \beta_p} (\hat{\beta}_p - \beta_p) + \\ + \frac{1}{2!} \left[ \frac{\partial^2 h_i(Y_i, \beta)}{\partial \beta_1^2} (\hat{\beta}_1 - \beta_1)^2 + \dots + \frac{\partial^2 h_i(Y_i, \beta)}{\partial \beta_p^2} (\hat{\beta}_p - \beta_p)^2 + \right. \\ + \frac{\partial^2 h_i(Y_i, \beta)}{\partial \beta_1 \partial \beta_2} (\hat{\beta}_1 - \beta_1)(\hat{\beta}_2 - \beta_2) + \dots + \frac{\partial^2 h_i(Y_i, \beta)}{\partial \beta_1 \partial \beta_p} (\hat{\beta}_1 - \beta_1) \\ (\hat{\beta}_p - \beta_p) + \frac{\partial^2 h_i(Y_i, \beta)}{\partial \beta_3 \partial \beta_2} (\hat{\beta}_3 - \beta_3)(\hat{\beta}_2 - \beta_2) + \dots + \frac{\partial^2 h_i(Y_i, \beta)}{\partial \beta_2 \partial \beta_p} \\ \left. (\hat{\beta}_2 - \beta_2)(\hat{\beta}_p - \beta_p) + \dots + \frac{\partial^2 h_i(Y_i, \beta)}{\partial \beta_p \partial \beta_{p-1}} (\hat{\beta}_p - \beta_p)(\hat{\beta}_{p-1} - \beta_{p-1}) \right].$$

Portanto,

$$R_i = \varepsilon_i + \sum_{r=1}^p (\hat{\beta}_r - \beta_r) H_r^{(i)} + \sum_{r=1}^p \sum_{s=1}^p \frac{1}{2} (\hat{\beta}_s - \beta_s) (\hat{\beta}_r - \beta_r) H_{rs}^{(i)} \quad (31),$$

em que  $H_r^{(i)} = \frac{\partial h_i(Y_i, \beta)}{\partial \beta_r}$  e  $H_{rs}^{(i)} = \frac{\partial^2 h_i(Y_i, \beta)}{\partial \beta_r \partial \beta_s}$ .

A expressão (31) é definida como a expressão geral para os resíduos generalizados de Cox-Snell, e a partir dela pode-se determinar algumas propriedades destes resíduos, como esperança, a variância e a covariância existente entre eles.

Para determinar as esperanças dos  $R_i$ 's, basta aplicar a esperança em (31), obtendo-se a seguinte expressão:

$$E(R_i) = E(\varepsilon_i) + \sum_{r=1}^p \{E(\hat{\beta}_r - \beta_r) E H_r^{(i)} + \text{cov}[(\hat{\beta}_r - \beta_r), H_r^{(i)}]\} + \sum_{s=1}^p \sum_{r=1}^p \left\{ \frac{1}{2} E(\hat{\beta}_r - \beta_r) (\hat{\beta}_s - \beta_s) E(H_{rs}^{(i)}) + \text{cov} \left[ \frac{1}{2} (\hat{\beta}_r - \beta_r) (\hat{\beta}_s - \beta_s), H_{rs}^{(i)} \right] \right\} \quad (32).$$

Em (32), o segundo termo pode ser dado, de acordo com (30), por:

$$\sum_{r=1}^p E(\hat{\beta}_r - \beta_r) E(H_r^{(i)}) = \sum_{r=1}^p b_r E(H_r^{(i)}) \quad (33),$$

em que  $E(\hat{\beta}_r - \beta_r) = b_r$ , isto é, o vício para o parâmetro  $\beta_r$ .

Para calcular o terceiro termo, considera-se de acordo com (25), que

$$(\hat{\beta}_r - \beta_r) = \sum_{s=1}^p I^{rs} U_s^{(i)}; \text{ portanto:}$$

$$\begin{aligned} \sum_{r=1}^p \text{cov}(\hat{\beta}_r - \beta_r, H_r^{(i)}) &= \sum_{r=1}^p \left( \sum_{s=1}^p I^{rs} U_s^{(i)}, H_r^{(i)} \right) = \\ &= \sum_{r=1}^p \sum_{s=1}^p \text{cov} \left( \sum_{i=1}^n I^{rs} U_s^{(i)}, \sum_{i=1}^n H_r^{(i)} \right) = \sum_{r=1}^p \sum_{s=1}^p I^{rs} \text{cov} \left( \sum_{i=1}^n U_s^{(i)}, \sum_{i=1}^n H_r^{(i)} \right) \\ &= \sum_{r=1}^p \sum_{s=1}^p I^{rs} \{ E[(\sum_{i=1}^n U_s^{(i)})(\sum_{i=1}^n H_r^{(i)})] - E(\sum_{i=1}^n U_s^{(i)}) E(\sum_{i=1}^n H_r^{(i)}) \} = \\ &= \sum_{r=1}^p \sum_{s=1}^p I^{rs} \{ [E(U_s^{(1)} H_r^{(1)}) + E(U_s^{(1)} H_r^{(2)}) + \dots + E(U_s^{(1)} H_r^{(n)}) + \\ &+ E(U_s^{(2)} H_r^{(1)}) + E(U_s^{(2)} H_r^{(2)}) + \dots + E(U_s^{(2)} H_r^{(n)}) + \\ &+ E(U_s^{(n)} H_r^{(1)}) + E(U_s^{(n)} H_r^{(2)}) + \dots + E(U_s^{(n)} H_r^{(n)})] - \\ &- [E(U_s^{(1)}) E(H_r^{(1)}) + E(U_s^{(1)}) E(H_r^{(2)}) + \dots + \\ &+ E(U_s^{(1)}) E(H_r^{(n)}) + E(U_s^{(2)}) E(H_r^{(1)}) + E(U_s^{(2)}) E(H_r^{(2)}) + \dots + \\ &+ E(U_s^{(2)}) E(H_r^{(n)}) + \dots + E(U_s^{(n)}) E(H_r^{(1)}) + E(U_s^{(n)}) E(H_r^{(2)}) + \\ &+ \dots + E(U_s^{(n)}) E(H_r^{(n)})] \}. \end{aligned}$$

Dessa forma,

$$\sum_{r=1}^p \sum_{s=1}^p \text{cov}(U_s^{(.)}, H_r^{(.)}) = \text{cov}(U_s^{(1)}, H_r^{(1)}) + \text{cov}(U_s^{(2)}, H_r^{(2)}) +$$

$$+ \dots + \text{cov}(U_s^{(n)}, H_r^{(n)}) - \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n \text{cov}(U_s^{(i)}, H_r^{(j)}),$$

em que:

$$\sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n \text{cov}(U_s^{(i)}, H_r^{(j)}) = \text{cov}(U_s^{(1)}, H_r^{(2)}) + \text{cov}(U_s^{(1)}, H_r^{(3)}) + \dots +$$

$$+ \text{cov}(U_s^{(2)}, H_r^{(1)}) + \text{cov}(U_s^{(2)}, H_r^{(3)}) + \dots + \text{cov}(U_s^{(n)}, H_r^{(n-1)}) = 0.$$

O  $U_s^{(i)}$  é uma função de observações independentes denotada pela primeira derivada da log-verossimilhança e os resíduos também constituem uma função destas observações e das estimativas de verossimilhança dos parâmetros.

Além disso, havendo propriedades aproximadas dos  $\varepsilon_i$ , tem-se que  $H_r^{(j)}$  e

$$U_s^{(i)} \text{ são independentes e } \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n \text{cov}(U_s^{(i)}, H_r^{(j)}) = 0.$$

Portanto,

$$\sum_{r=1}^p \sum_{s=1}^p \text{cov}(U_s^{(.)}, H_r^{(.)}) = \sum_{s=1}^p I^{sr} \{ E[(\sum_{j=1}^n U_s^{(j)})(\sum_{j=1}^n H_r^{(j)})] -$$

$$- E(\sum_{j=1}^n U_s^{(j)}) E(\sum_{j=1}^n H_r^{(j)}) \},$$

$$\text{e como } E(\sum_{j=1}^n U_s^{(j)}) E(\sum_{j=1}^n H_r^{(j)}) = 0.$$

$$\sum_{r=1}^p \sum_{s=1}^p \text{cov}(U_s^{(i)}, H_r^{(i)}) = \sum_{r=1}^p \sum_{s=1}^p I^{rs} E(U_s^{(i)} H_r^{(i)}) \quad (34),$$

O quarto termo é dado, de acordo com (29), por:

$$\sum_{r=1}^p \sum_{s=1}^p \frac{1}{2} \{E[(\hat{\beta}_r - \beta_r)(\hat{\beta}_s - \beta_s)]E(H_{rs}^{(i)})\} = \sum_{r=1}^p \sum_{s=1}^p \frac{1}{2} I^{rs} E(H_{rs}^{(i)}) \quad (35),$$

em que  $E[(\hat{\beta}_r - \beta_r)(\hat{\beta}_s - \beta_s)] = \text{cov}(\hat{\beta}_r, \hat{\beta}_s) = I^{rs}$  é o último termo de (32) será desprezado, sendo de  $O n^{-1}$ .

Substituindo (33), (34) e (35) em (32), obtém-se:

$$E(R_i) = E(\varepsilon_i) + \sum_{r=1}^p b_r E(H_r^{(i)}) + \sum_{s=1}^p \sum_{r=1}^p I^{sr} E(H_r^{(i)} U_s^{(i)}) + \frac{1}{2} H_{rs}^{(i)}$$

Denotando  $\sum_{r=1}^p b_r E(H_r^{(i)}) + \sum_{s=1}^p \sum_{r=1}^p I^{sr} E(H_r^{(i)} U_s^{(i)}) + \frac{1}{2} H_{rs}^{(i)}$  por  $a_i$ ,

finalmente tem-se que:

$$E(R_i) = E(\varepsilon_i) + a_i \quad (36).$$

Seguindo o mesmo caminho, elevando ao quadrado (31) e calculando suas esperanças, tem-se que:

$$\begin{aligned} E(R_i^2) &= E(\varepsilon_i^2) + \sum_{r=1}^p 2b_r E(\varepsilon_i H_r^{(i)}) + \\ &+ 2 \sum_{s=1}^p \sum_{r=1}^p I^{sr} E(\varepsilon_i H_r^{(i)} U_s^{(i)}) + \frac{1}{2} H_r^{(i)} H_s^{(i)} + \frac{1}{2} \varepsilon_i H_{rs}^{(i)} \end{aligned} \quad (37);$$



que, para  $i \neq j$ ,

$$E(R_i R_j) = \{E(\varepsilon_i)\}^2 + (a_i + a_j)E(\varepsilon_i) + \sum_{s=1}^p \sum_{r=1}^p I^{rs} E\{\varepsilon_i H_r^{(i)} U_s^{(i)} + \varepsilon_j H_r^{(j)} U_s^{(j)} + H_r^{(i)} H_s^{(j)}\} \quad (38),$$

e a  $\text{Var}(R_i)$  é dada por:

$$\text{Var}(R_i) = \text{Var}(\varepsilon_i) + c_{ii} \quad (39).$$

Dessa forma, aproximadamente tem-se que:

$$\begin{cases} E(R_i) = E(\varepsilon_i) + a_i \\ \text{Var}(R_i) = \text{Var}(\varepsilon_i) + c_{ii} \\ \text{cov}(R_i, R_j) = c_{ij} \end{cases} \quad (40).$$

A partir de (40) pode-se afirmar que os  $R_i$  têm as mesmas propriedades dos  $\varepsilon_i$ , pois  $a_i$ ,  $c_{ii}$ ,  $c_{ij}$  são constantes da ordem  $n^{-1}$ , sendo praticamente desprezíveis.

### 3.4 Resíduos Modificados

O objetivo desta seção é determinar, por meio das propriedades dos resíduos generalizados apresentados na seção 3.3, um resíduo modificado

$(R'_i)$  que possa ter as mesmas propriedades dos  $\varepsilon_i$ 's, isto é,  $E(R'_i) = E(\varepsilon_i)$  e  $\text{Var}(R'_i) = \text{Var}(\varepsilon_i)$ .

De um modo geral, existem três caminhos para os quais os resultados anteriores, obtidos em (40), podem ser usados:

1. Se todos os termos corretivos em (40) são numericamente pequenos, isto dá alguma segurança de que os  $R_i'$ s têm as mesmas propriedades estatísticas dos  $\varepsilon_i'$ s.
2. Para qualquer estatística T, as expressões dadas em (40) podem ser usadas para determinar sua esperança e variância. Considere-se que

$$T = \sum_{i=1}^n R_i z_i,$$

em que os  $z_i$ s são constantes, e sabendo que  $E[\sum_{i=1}^n X_i] = \sum_{i=1}^n E(X_i)$  e que

$$\begin{aligned} \text{Var}[\sum_{i=1}^n a_i X_i] &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j \text{cov}[X_i X_j] \\ &= \sum_{i=1}^n a_i^2 \text{Var}(X_i) + \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n a_i a_j \text{cov}[X_i X_j], \end{aligned}$$

então:

$$E(T) = E\left[\sum_{i=1}^n R_i z_i\right] = \sum_{i=1}^n E(R_i z_i) = \sum_{i=1}^n z_i E(R_i) = \sum_{i=1}^n z_i [E(\varepsilon_i) + a_i]$$

$$E(T) = E(\varepsilon_i) \sum_{i=1}^n z_i + \sum_{i=1}^n a_i z_i$$

$$\text{Var}(T) = \text{Var}\left(\sum_{i=1}^n R_i z_i\right) = \sum_{i=1}^n \text{Var}(R_i z_i) + \sum_{i=1}^n \sum_{j=1}^n z_i z_j \text{cov}(R_i, R_j)$$

$$\begin{aligned}
\text{Var}(T) &= \sum_{i=1}^n z_i^2 \text{Var}(R_i) + \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n z_i z_j c_{ij} = \\
&= \sum_{i=1}^n z_i^2 [\text{Var}(\varepsilon_i) + c_{ii}] + \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n z_i z_j c_{ij} = \\
\text{Var}(T) &= \sum_{i=1}^n z_i^2 \text{Var}(\varepsilon_i) + \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n z_i z_j c_{ij} .
\end{aligned}$$

Uma estatística usada para testar possíveis dependências entre  $z_i$  de  $\text{Var}$

$(\varepsilon_i)$  é dada por:  $T' = \sum_{i=1}^n R_i^2 (z_i - \bar{z})$ , em que  $\bar{z} = \sum_{i=1}^n z_i / n$ .

3. A partir de (40), pode-se definir um resíduo modificado, denotado por  $R'_i$ , com as mesmas propriedades dos erros aleatórios  $(\varepsilon_i)$ . Existem várias formas de escrever uma expressão para estes resíduos modificados; uma delas pode ser escrita por:

$$R'_i = (1 + K_i)R_i + l_i ,$$

em que  $K_i$ ,  $l_i$  são constantes pequenas da  $O(n^{-1})$  e  $R_i$  é o  $i$ -ésimo resíduo generalizado. Finalmente, tem-se que  $E(R'_i) = E(\varepsilon_i)$  e  $\text{Var}(R'_i) = \text{Var}(\varepsilon_i)$ .

A partir das propriedades enunciadas para os resíduos modificados de Cox-Snell, pode-se aplicá-los indiretamente para o exame da forma distribucional dos erros aleatórios, por meio da plotagem dos resíduos

modificados ordenados versus as estatísticas de ordem esperadas para a forma distribucional proposta para os  $\varepsilon_i'$ .

Dessa forma, os  $R_i'$  podem ter a mesma média e variância e distribuição bem próxima dos  $\varepsilon_i$ .

A partir da análise gráfica, pode-se concluir que se o modelo proposto pelos  $\varepsilon_i'$  estiver correto, o resultado será uma linha reta. Caso isso não aconteça, será necessário assumir uma outra forma distribucional para os  $\varepsilon_i'$ .

Na próxima seção será apresentada uma demonstração para a obtenção da forma particular dos resíduos generalizados definidos para Análise de Sobrevivência.

### 3.5 Resíduos de Cox-Snell para modelos de Análise de Sobrevivência

Após determinar a expressão geral dos resíduos de Cox-Snell, esta seção tem como finalidade apresentar a demonstração da forma particular destes resíduos segundo Allison (1995), dada por  $e_i = -\log \hat{S}(t_i | x_i)$ , em que  $e_i$  tem aproximadamente uma distribuição exponencial com  $\lambda=1$ .

A idéia para este desenvolvimento vem do fato de que a variável tempo de falha  $T$ , sendo contínua e não censurada, implica que a função de sobrevivência de  $T$ , denotada por  $S(t)$ , que é a probabilidade de um indivíduo não falhar até o tempo  $t$ ,  $P(T \geq t)$ , tem uma distribuição uniforme em  $(0,1)$ .

Para chegar a esta demonstração, as probabilidades logicamente devem ser determinadas no início do experimento. Dessa forma, como no início do

experimento não se sabe se a observação será censurada ou não, então não é apropriado fazer diferenças entre dados censurados e dados não censurados.

Para esta demonstração será utilizado o teorema da transformação integral da probabilidade, que é enunciado e demonstrado por Mood et al. (1974). Tal teorema diz o seguinte:

1ª) Se a variável aleatória contínua  $X \sim F_X(x)$  e  $U = F_X(X)$ , então  $U \sim U(0,1)$ ;

2ª) Se  $U \sim U(0,1)$  e  $X = F_X^{-1}(U)$ , então  $X \sim F_X(\cdot)$ .

Mediante o teorema integral da probabilidade pode-se chegar a seguinte conclusão:

Se  $t_i | \underline{x}_i \sim F_{t_i|\underline{x}_i}(\cdot)$ , então  $F_{t_i|\underline{x}_i}(t_i | \underline{x}_i) \sim U(0,1)$

Com esse argumento, pode-se encontrar a distribuição de  $-\log S(t_i | \underline{x}_i)$ , da seguinte forma:

$$\begin{aligned} P(-\log S(t_i | \underline{x}_i) < t) &= P(\log S(t_i | \underline{x}_i) > -t) = P(S(t_i | \underline{x}_i) > e^{-t}) = P(1 - F_{t_i|\underline{x}_i}(t_i | \underline{x}_i) > e^{-t}) = \\ &= P(-F_{t_i|\underline{x}_i}(t_i | \underline{x}_i) > -1 + e^{-t}) = P(F_{t_i|\underline{x}_i}(t_i | \underline{x}_i) < 1 - e^{-t}) = 1 - e^{-t} \end{aligned}$$

Portanto,  $-\log S(t_i | \underline{x}_i)$  tem uma distribuição exponencial ( $\lambda=1$ ).

Segundo Allison (1995), se  $t_i$  é um tempo censurado, então o  $i$ -ésimo resíduo também é tratado como uma grandeza censurada .

A explicação para isto vem de que, sendo a função de sobrevivência estimada pelo estimador de Kaplan-Meier (Kaplan & Meier, 1958), a observação não censurada terá a mesma função de sobrevivência de uma

observação censurada no mesmo intervalo de tempo. Este método é baseado em um número de intervalos igual ao número de tempos de falhas distintos, no qual  $n$  indivíduos estão em estudo, para os quais são registradas  $k$  falhas distintas, em que  $k \leq n$ , nos tempos  $t_1 < t_2 < \dots < t_k$  (não censurados). A expressão geral do estimador é dada por:

$$\hat{S}(t) = \prod_{v_{t_j} \leq t} \left( \frac{n_j - d_j}{n_j} \right) = \prod_{v_{t_j} \leq t} \left( 1 - \frac{d_j}{n_j} \right),$$

em que  $d_j$  é o número de indivíduos que falharam no tempo  $t_j$  e  $n_j$  é o número de indivíduos sob risco (não falhou e não foi censurado) até o tempo  $t_j$ .

No próximo capítulo serão apresentadas algumas aplicações dos Resíduos Generalizados de Cox-Snell, considerando também o caso particular apresentado na seção 3.5 para modelos de análise de sobrevivência.

## **4 APLICAÇÕES DOS RESÍDUOS GENERALIZADOS PARA TESTAR A ADEQUACIDADE DE AJUSTE DE MODELOS**

Este capítulo tem como finalidade ilustrar a teoria dos resíduos generalizados de Cox-Snell, por meio de quatro aplicações.

### **4.1 Aplicação 01**

A primeira aplicação tem como objetivo mostrar a lógica da plotagem dos resíduos ordenados originados de uma certa distribuição versus as estatísticas de ordem esperadas desta distribuição.

Para esta aplicação foram simuladas 4 amostras de tamanho 50; estas amostras foram consideradas como resíduos pertencentes a 3 distribuições, sendo uma amostra da distribuição normal (0,1), uma amostra da distribuição logística (0,1) e 2 amostras da distribuição valor extremo (0,1) (para valores mínimos).

De acordo com o modelo dado em (1), tem-se que:

- i. Se os erros têm uma distribuição normal, isto implica que  $T$  tem uma distribuição log-normal;
- ii. Se os erros têm uma distribuição logística, isto implica que  $T$  tem uma distribuição log-logística;
- iii. Se os erros têm uma distribuição valor extremo com 2 parâmetros, isto implica que  $T$  tem uma distribuição Weibull;
- iv. Se os erros têm uma distribuição valor extremo com 1 parâmetro (sendo  $\sigma$  fixado em 1), isto implica que  $T$  tem distribuição exponencial.

Para gerar estas amostras foi utilizado o método de Monte Carlo, que se aplica a processos de simulação que envolvem componentes aleatórios,

associado ao teorema da transformação integral da probabilidade, enunciado na seção 3.5. Na Tabela 1 estão os resíduos referentes a cada distribuição.

TABELA 1 Resíduos simulados para quatro distribuições propostas

Normal	Logística	Valor Extremo	Valor Extremo
1.3118	-0.5325	0.0056	0.5294
-0.6641	2.7547	2.7857	2.3643
0.5410	1.3841	1.4980	-0.4775
2.0420	0.0752	0.4212	1.3832
1.4392	1.6823	1.7688	1.3163
0.9047	2.7262	2.7581	2.3362
0.5534	-2.3640	-0.8977	0.7685
-0.3516	-2.6931	-1.0147	0.6557
-0.8022	-2.2018	-0.8358	-0.6425
-0.2952	-1.8735	-0.7013	-0.4768
0.6387	0.5695	0.8020	0.2472
-1.8486	-2.9931	-1.1125	0.5811
-0.1160	0.8798	1.0584	0.5456
0.6832	-1.8629	-0.6967	-0.4711
0.0480	2.3641	2.4094	1.9803
1.2476	0.3821	0.6535	2.1649
0.3421	0.1814	0.5000	1.6040
-1.5370	-0.4902	0.0326	-0.9792
0.6595	3.0214	3.0452	2.6276
-0.7377	-0.1885	0.2334	0.9510
-0.0726	-0.5046	0.0234	-1.0072
0.6913	1.5543	1.6517	-0.3607
0.4325	-2.0201	-0.7630	0.9363
-0.2853	0.3000	0.5900	1.9047
1.8385	2.9782	3.0031	0.1624
-0.6965	0.8000	0.9913	0.4691
0.0037	-0.9995	-0.2722	2.2887
-0.3561	-1.5114	-0.5369	1.3473
-1.9286	-1.6203	-0.5882	1.2355
0.4121	1.1202	1.2647	-0.7091
-0.9650	-0.1735	0.2438	-0.5229
-2.2211	-4.3137	-1.4649	0.4242

“...continua...”



“Tabela 1, cont.”

Normal	Logística	Valor Extremo	Valor Extremo
0.4038	1.8356	1.9105	1.4648
-0.4231	1.8008	1.8782	1.4310
-1.0471	-1.5210	-0.5415	1.3368
0.1416	0.3824	0.6537	2.1658
0.3646	1.2779	1.4033	-0.5617
-0.7218	2.8411	2.8696	0.1347
-0.3961	-0.5795	-0.0240	-1.1802
-1.9796	-0.6804	-0.0863	0.3788
0.5986	-0.6944	-0.0948	0.3653
1.6027	1.0575	1.2103	0.7156
0.5694	0.3954	0.6638	2.2127
0.4597	1.8946	1.9655	1.5221
-1.3802	1.3922	1.5052	1.0366
-0.7263	-1.8357	-0.6849	1.0566
-1.0828	-2.2074	-0.8380	-0.6451
-0.3846	-0.9887	-0.2662	0.1059
0.5196	-3.7653	-1.3319	0.4661
0.6813	1.8696	1.9421	-0.1867

Após a obtenção dos resíduos, o próximo passo foi ordená-los e, em seguida, calcular as estatísticas de ordem esperadas das distribuições propostas. As estatísticas de ordem esperadas das distribuições logística e o valor extremo foram obtidas por meio do programa matemático MAPLE<sup>®</sup>V (1997), devido à dificuldade de execução das integrais manualmente. Já as estatísticas de ordem esperadas da distribuição normal foram obtidas via simulação devido à facilidade comparada com o programa MAPLE<sup>®</sup>V (1997). Para isto, geraram-se 5000 amostras de uma distribuição normal padrão de tamanho 50, sendo que cada amostra foi gerada por meio da função rannor, do sistema SAS (1996), cuja sintaxe é dada por RANNOR(x), sendo x utilizado como “semente”, isto é, o valor referente à geração dos números aleatórios, e em seguida estas amostras foram ordenadas ascendentemente. Finalmente, calcularam as médias, obtendo-se, assim, as estimativas das estatísticas de ordem esperadas. Feita a plotagem

dos resíduos ordenados versus as estatísticas de ordem esperadas, obtiveram os seguintes resultados:

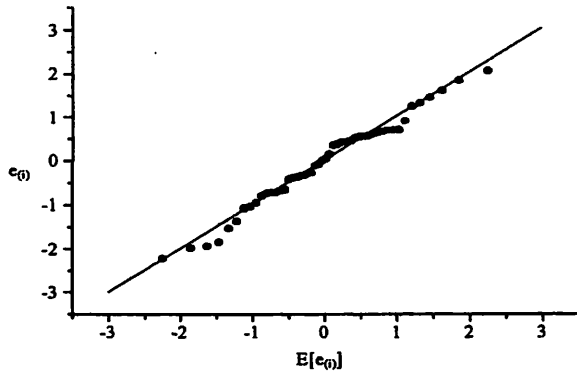


FIGURA 1 Resíduos ordenados versus os valores esperados das estatísticas de ordem, caso da distribuição normal (0,1). T tem distribuição Log-normal.

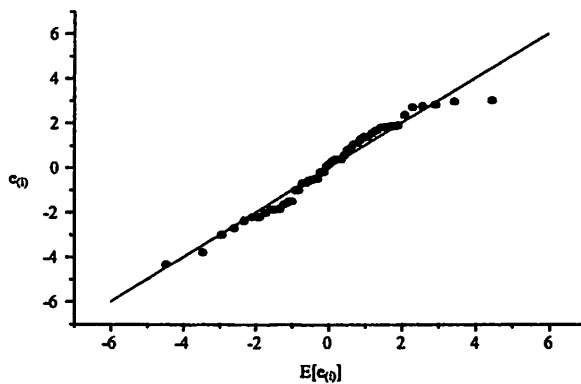


FIGURA 2 Resíduos ordenados versus os valores esperados das estatísticas de ordem, caso da distribuição logística (0,1). T tem distribuição Log-logística.

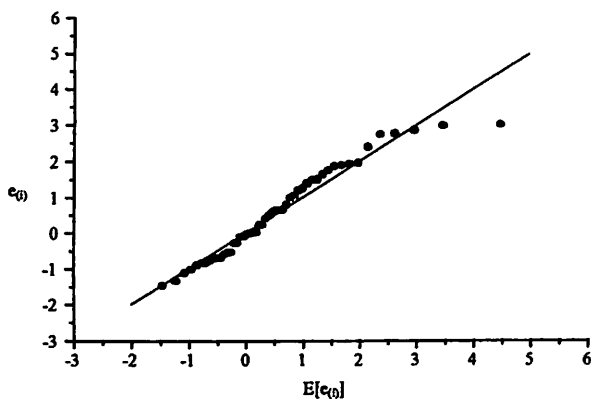


FIGURA 3 Resíduos ordenados versus os valores esperados das estatísticas de ordem, caso da distribuição valor extremo (0,1). T tem distribuição Weibull

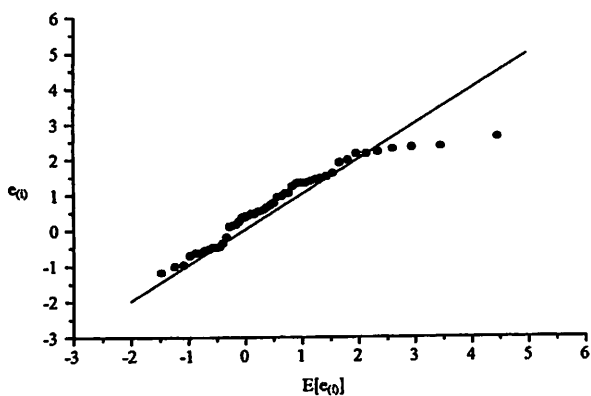


FIGURA 4 Resíduos ordenados versus os valores esperados das estatísticas de ordem, caso da distribuição valor extremo (0,1) de um parâmetro, T tem distribuição Exponencial

Pode-se observar, nas Figuras 1,2,3 e 4, que na plotagem dos conjuntos dos resíduos, para as distribuições sugeridas, versus seus respectivos valores esperados das estatísticas de ordem, o resultado obtido foi próximo a uma linha reta.

Na teoria apresentada na seção 3, sobre resíduos generalizados de Cox-Snell, determinou-se uma expressão para os resíduos de tal maneira que eles tenham uma distribuição muito próxima dos erros aleatórios, que são quantidades desconhecidas. Sendo assim, ao assumir uma distribuição para os erros, se a plotagem dos resíduos generalizados modificados de Cox-Snell versus as estatísticas de ordem esperadas da distribuição assumida fornece uma linha reta, isto quer dizer que a suposição distribucional para os erros aleatórios é correta. Caso contrário, deve-se testar uma outra distribuição (assumida) para os ditos erros. Isto será visto na próxima seção.

## **4.2 Aplicação 02**

A segunda aplicação tem como finalidade ilustrar com dados reais a lógica apresentada na seção 4.1 e aplicar toda a teoria apresentada no capítulo 3, a partir de dados de tempo de vida de isolantes de sistemas de uma nova classe H, analisados por Nelson (1990), em que se aplicou um teste acelerado de vida nestes isolantes. Tais dados referem-se ao tempo em horas que os isolantes levaram para tornarem-se defeituosos. Os testes foram realizados em pequenos motores, a temperaturas elevadas. Dez motores foram colocados para trabalhar a temperaturas de 190, 220, 240 e 260<sup>o</sup>C e inspecionados periodicamente para detectar a ocorrência de falha. Segundo Nelson (1990), para o ajuste dos tempo de vida, foi considerado o modelo de regressão log-normal, como é dado a seguir:

$$\text{Log}T_{ij} = \beta_0 + \beta_1 x_i + \varepsilon_{ij} ,$$

para  $i = 1,2,\dots,4$  e  $j = 1,2,\dots,10$ , em que  $\log T_{ij}$  é o log do tempo de vida dos isolantes,  $x_i$  é a temperatura em graus centígrados,  $\beta_0$  e  $\beta_1$  são os parâmetros do modelo e  $\varepsilon_{ij}$  tem distribuição normal.

Na Tabela 2, são apresentados os tempos que se referem 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.

TABELA 2 Tempo de vida de isolantes (em horas) para cada temperatura, em graus centígrados

190°C	220°C	240°C	260°C
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	1881	1632
10511	3108	1953	1896

Aplicando-se a teoria dos resíduos generalizados, apresentado na seção 3, foram obtidos os seguintes resultados.

TABELA 3 Resíduos modificados obtidos para o ensaio do tempo de vida dos isolantes

Obs.	$c_{ii}^+$	$k_i$	$R_i$	$R'_i$	Obs.	$c_{ii}^+$	$k_i$	$R_i$	$R'_i$
1	0.0041	-0.0267	-0.0647	-0.0630	21	0.0049	-0.0319	-0.3599	-0.3484
2	0.0041	-0.0267	-0.0647	-0.0630	22	0.0049	-0.0319	-0.3599	-0.3484
3	0.0041	-0.0267	-0.0647	-0.0630	23	0.0049	-0.0319	-0.1018	-0.0985
4	0.0041	-0.0267	0.0913	0.0888	24	0.0049	-0.0319	-0.0707	-0.0684
5	0.0041	-0.0267	0.1729	0.1683	25	0.0049	-0.0319	-0.0405	-0.0393
6	0.0041	-0.0267	0.1729	0.1683	26	0.0049	-0.0319	-0.0113	-0.0109
7	0.0041	-0.0267	0.1729	0.1683	27	0.0049	-0.0319	-0.0113	-0.0109
8	0.0041	-0.0267	0.1729	0.1683	28	0.0049	-0.0319	0.0171	0.0166
9	0.0041	-0.0267	0.3098	0.3015	29	0.0049	-0.0319	0.0448	0.0433
10	0.0041	-0.0267	0.3098	0.3015	30	0.0049	-0.0319	0.1482	0.1435
11	0.0045	-0.0295	-0.5622	-0.5456	31	0.0052	-0.0338	-0.4233	-0.4090
12	0.0045	-0.0295	-0.2394	-0.2323	32	0.0052	-0.0338	-0.2082	-0.2012
13	0.0045	-0.0295	-0.2394	-0.2323	33	0.0052	-0.0338	-0.2082	-0.2012
14	0.0045	-0.0295	-0.2394	-0.2323	34	0.0052	-0.0338	-0.2082	-0.2012
15	0.0045	-0.0295	-0.2394	-0.2323	35	0.0052	-0.0338	-0.0046	-0.0045
16	0.0045	-0.0295	-0.2394	-0.2323	36	0.0052	-0.0338	0.2079	0.2009
17	0.0045	-0.0295	0.0042	0.0041	37	0.0052	-0.0338	0.3651	0.3528
18	0.0045	-0.0295	0.0042	0.0041	38	0.0052	-0.0338	0.4687	0.4528
19	0.0045	-0.0295	0.0042	0.0041	39	0.0052	-0.0338	0.5625	0.5435
20	0.0045	-0.0295	0.0042	0.0041	40	0.0052	-0.0338	0.7272	0.7027

Em seguida foram, obtidas as estatísticas de ordem esperadas de uma normal via simulação, como foi relatado na seção 4.1, e de uma exponencial dada pela expressão  $E(e_{(i)}) = \sum_{l=1}^n (n - l + 1)^{-1}$ , segundo Lawless (1982).

Feita a plotagem dos resíduos modificados versus os valores esperados das estatísticas de ordem das distribuições normal e exponencial, obtiveram-se os resultados apresentados nas Figuras 5 e 6.

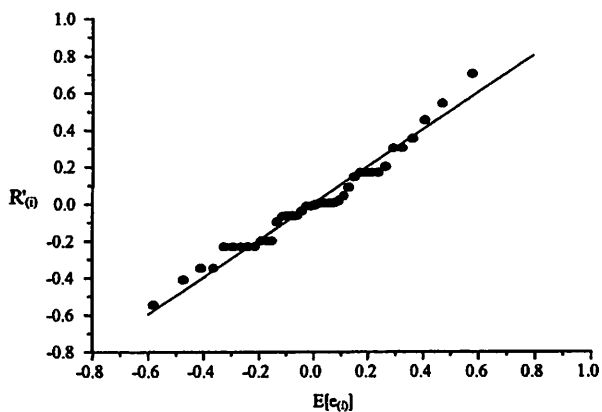


FIGURA 5 Resíduos modificados versus os valores esperados das estatísticas de ordem de uma distribuição normal

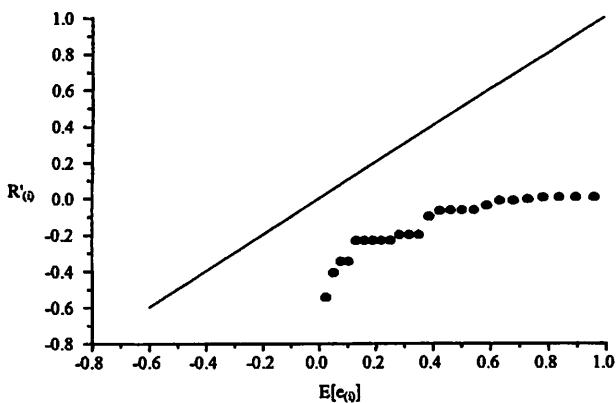


FIGURA 6 Resíduos modificados versus os valores esperados das estatísticas de ordem de uma distribuição exponencial

Pela Figura 6, nota-se claramente que os resíduos modificados

apresentam a mesma distribuição dos erros aleatórios, pois estes tendem a afastar-se da reta. Sendo assim, pode-se afirmar que os erros aleatórios não seguem uma distribuição exponencial.

É possível observar, pela Figura 5, que os resíduos modificados têm distribuição equivalente à distribuição dos erros do modelo proposto para os tempo de vida dos isolantes. Dessa forma, pode-se confirmar que o modelo proposto por Nelson (1990) está correto.

### 4.3 Aplicação 03

Para a terceira aplicação foram utilizados dados com presença de censura, em que, para a obtenção dos resíduos de Cox-Snell, foi utilizada uma sub-rotina apresentada em SAS® por Allison (1994).

Os dados tratam do estudo de tempo de vida de pacientes com mieloma múltipla, doença caracterizada pelo acúmulo de células plasmática anormais, na medula, realizado pelo Centro Médico da Universidade da Virgínia do Oeste, USA, cuja finalidade era examinar a associação entre os valores de certas covariáveis e o tempo de sobrevivência dos pacientes. No estudo, a variável resposta era o tempo, em meses, a partir do diagnóstico antes da morte do paciente devido à doença. Os dados foram obtidos por Krall et al.(1975), citados por Collett (1994), em que foram registrados 48 pacientes com idade entre 50 e 80 anos. Durante o tempo de diagnóstico, os valores de um número de covariáveis foram registrados para cada paciente, incluindo a idade dos pacientes em anos, o sexo, sendo 1 para homem e 2 para mulher, o nível de uréia nitrogenada no sangue, cálcio e hemoglobina, a porcentagem de células plasmática na medula e um indicador da variável BC que denota a presença ou não de proteína Bence-Jones, dada por 0 se estiver ausente na urina e por 1, caso contrário. Os pacientes que não morreram até o tempo em que o estudo foi



completado tiveram o tempo de sobrevivência censurado (censura tipo I à direita), indicado por 0, e os pacientes que morreram a partir da mieloma múltipla receberam 1.

O tempo de sobrevivência dos pacientes e as covariáveis estão apresentados na Tabela 4. O modelo ajustado é um AFT (tempo de falha acelerado) dado por:

$$\text{Log}T_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + \beta_6 X_{i6} + \beta_7 X_{i7} + \sigma \varepsilon_i \quad (41),$$

em que  $T_i$  é o  $i$ -ésimo tempo de sobrevivência para o  $i$ -ésimo paciente,  $\beta_0, \beta_1, \dots, \beta_7$  e  $\sigma$  são os parâmetros do modelo,  $X_{i1}, X_{i2}, \dots, X_{i7}$  correspondem às covariáveis do modelo e  $\varepsilon_i$  tem distribuição valor extremo.

TABELA 4 Tempo de sobrevivência de pacientes com mieloma múltipla e suas respectivas covariáveis

Tempo	Censura	Idade	Sexo	Nível de Uréia	Cálcio	Hemoglobina	% Células Plasmáticas	Proteína BC
13	1	66	1	25	10	14.6	18	1
52	0	66	1	13	11	12	100	0
6	1	53	2	15	13	11.4	33	1
40	1	69	1	10	10	10.2	30	1
10	1	65	1	20	10	13.2	66	0
7	0	57	2	12	8	9.9	45	0
66	1	52	1	21	10	12.8	11	1
10	0	60	1	41	9	14	70	1
10	1	70	1	37	12	7.5	47	0
14	1	70	1	40	11	10.6	27	0
16	1	68	1	39	10	11.2	41	0
4	1	50	2	172	9	10.1	46	1

“...continua...”

“Tabela 4, cont.”

Tempo	Censura	Idade	Sexo	Nível de Uréia	Cálcio	Hemoglobina	% Células Plasmática	Proteína BC
65	1	59	1	28	9	6.6	66	0
5	1	60	1	13	10	9.7	25	0
11	0	66	2	25	9	8.8	23	0
10	1	51	2	12	9	9.6	80	0
15	0	55	1	14	9	13	8	0
5	1	67	2	26	8	10.4	49	0
76	0	60	1	12	12	14	9	0
56	0	66	1	18	11	12.5	90	0
88	1	63	1	21	9	14	42	1
24	1	67	1	10	10	12.4	44	0
51	1	60	2	10	10	10.1	45	1
4	1	74	1	48	9	6.5	54	0
40	0	72	1	57	9	12.8	28	1
8	1	55	1	53	12	8.2	55	0
18	1	51	1	12	15	14.4	100	0
5	1	70	2	130	8	10.2	23	0
16	1	53	1	17	9	10	28	0
50	1	74	1	37	13	7.7	11	1
40	1	70	2	14	9	5	22	0
1	1	67	1	165	10	9.4	90	0
36	1	63	1	40	9	11	16	1
5	1	77	1	23	8	9	29	0
10	1	61	1	13	10	14	19	0
91	1	58	2	27	11	11	26	1
18	0	69	2	21	10	10.8	33	0
1	1	57	1	20	9	5.1	100	1
18	0	59	2	21	10	13	100	0
6	1	61	2	11	10	5.1	100	0
1	1	75	1	56	12	11.3	18	0
23	1	56	2	20	9	14.6	3	0
15	1	62	2	21	10	8.8	5	0
18	1	60	2	18	9	7.5	85	1
12	0	71	2	46	9	4.9	62	0
12	1	60	2	6	10	5.5	25	0
17	1	65	2	28	8	7.5	8	0
3	0	59	1	90	10	10.2	6	1

Como se trata de dados censurados, a adequacidade do ajuste do modelo (41) será testada a partir dos resíduos apresentados na seção 3.5. Os resultados são apresentados nas Figuras 7, 8 e 9. É possível perceber que o tempo de vida dos pacientes com mieloma múltipla demonstra ser melhor ajustados pelo modelo prababilístico Weibull (Figura 8).

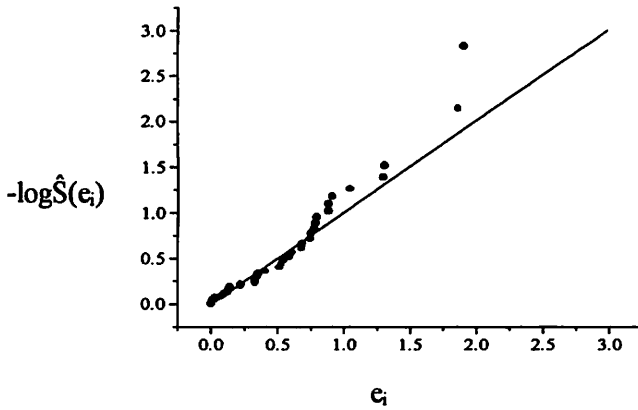


FIGURA 7 Ajuste log-normal para pacientes com mieloma múltipla

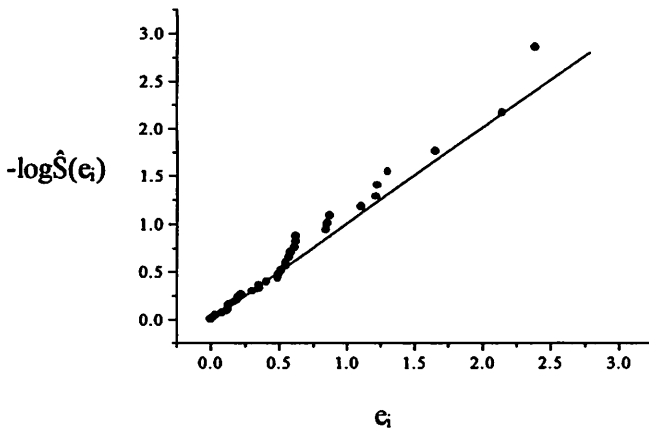


FIGURA 8 Ajuste Weibull para pacientes com mieloma múltipla

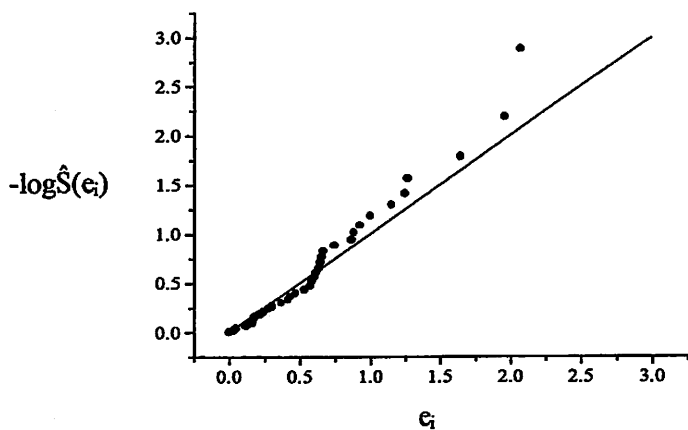


FIGURA 9 Ajuste exponencial para pacientes com mieloma múltipla

## 5 CONCLUSÃO

Para avaliar a adequacidade de ajuste de modelos de acordo com a teoria dos resíduos generalizados de Cox-Snell apresentada neste trabalho, foi determinada uma expressão para os resíduos de forma que eles tivessem uma distribuição muito próxima da distribuição dos erros aleatórios. Dessa forma, a adequacidade a partir dos gráficos de diagnósticos é feita com a plotagem dos resíduos generalizados modificados de Cox-Snell ordenados versus as estatísticas de ordem esperadas da distribuição assumida pelos erros no início da avaliação. Sendo assim, se o resultado da plotagem for uma linha reta, isto significa que a distribuição proposta para os erros aleatórios está correta. Caso o resultado não seja uma linha reta, deve-se assumir uma outra distribuição para estes erros aleatórios.

Modificando os resíduos generalizados de Cox-Snell pode-se obter os resíduos de Cox-Snell, conforme Allison (1995), para modelos de sobrevivência, em que qualquer que seja a distribuição dos erros, se ela for a distribuição correta, então tem distribuição exponencial com  $\lambda=1$ .

### **Considerações Finais**

Para um estudo futuro, este trabalho apresenta sugestões alternativas para a obtenção e avaliação dos resíduos generalizados de Cox-Snell. Entre estas alternativas, propõe-se a elaboração de um rotina que facilite o processo de avaliação de adequacidade de modelos por meio dos resíduos generalizados, bem como a confecção de tabelas referentes às estatísticas de ordem das distribuições mais conhecidas.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALLISON, P. D. **Survival analysis using the SAS system: a practical guide.** Cary, NC: SAS Institute, 1995. 292 p.
- ANSCOMBE, F. J.; TUKEY, W. J. The examination and analysis of residuals. **Technometrics**, Washington, v. 5, n. 2, p. 141-174, May 1963.
- BESLEY, D. A.; KUH, E.; WELSCH, R. E. **Regression diagnostics: identifying influential data and sources of collinearity.** New York: J. Wiley, 1980. 284 p.
- BOX, G. E. P.; HUNTER Jr. W. G.; HUNTER, J. S. **Statistics for experimenters: an introduction to design, data analysis and model building.** New York: J. Wiley, 1978. 653 p.
- CAMPOS, H. **Estatística experimental não-paramétrica.** 4. ed. Piracicaba: ESALQ, 1983. 349 p.
- COLLETT, D. **Modelling survival data medical research.** London: Chapman and Hall, 1994. 347 p.
- COOK, R. D. Detection of influential observations in linear regression. **Technometrics**, Washington, v. 19, n. 1, p. 15-18, Feb. 1977.
- COOK, R. D.; WEISBERG, S. Diagnostics for heteroscedasticity in regression. **Biometrika**, London, v. 70, n. 1, p. 1-10, Apr. 1983.
- COOK, R. D.; WEISBERG, S. **Residuals and influence in regression.** London: Chapman and Hall, 1986. 230 p.
- CORDEIRO, G. M. Introdução à teoria de verossimilhança. In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 10., 1992, Rio de Janeiro. **Anais...** Rio de Janeiro, 1992. p. 1-174.
- COX, D. R.; SNELL, E. J. A general definition of residuals. **Journal of the Royal Statistical Society B**, London, v. 30, n. 2, p. 248-254, Mar. 1968.

CROWLEY, J.; HU, M. Covariance analysis of heart transplant survival data. **Journal of the American Statistical Association**, Washington, v. 72, n. 357, p. 27-36, Mar. 1977.

D'AGOSTINO, R. B. Small sample probability points for the D test for normality. **Biometrika**, London, v. 59, n. 1, p. 219-221, Apr. 1972.

DRAPER, N. R.; SMITH, H. **Applied regression analysis**. 3. ed. New York: J. Wiley, 1998. 706 p.

KAPLAN, E. L. E.; MEIER, P. Nonparametric estimation from incomplete observations. **Journal of the American Statistical Association**, Washington, v. 53, n. 282, p. 457-481, June. 1958.

KRALL, J. M. ; UTHOFF, V. A.; HARLEY, J.B. A step-up procedure for selecting variables associated with survival. **Biometrics**, Washington, v. 31, n. 282, p. 49-57, June. 1975.

LAWLESS, J. F. **Statistical models and methods for lifetime data**. New York: J. Wiley, 1982. 580 p.

LILLIEFORS, H. W. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. **Journal of the American Statistical Association**, Washington, v. 62, n. 1, p. 399-402, June 1967.

MAPLE V. Versão 5. Waterloo Maple, 1981-1997.

MILLIKEN, G. A.; JOHNSON, D. E. **Analysis of messy data**. New York: Chapman & Hall, 1984. v. 1, 473 p.

MOOD, A. M.; GRAYBILL, F. A; BOES, D. C. **Introduction to the theory of statistics**. 3. ed. New York: John Wiley, 1974. 564 p.

NELSON, W. **Accelerated testing, statistical models, test plans and data analyses**. New York: J. Wiley, 1982. 621 p.

PARENTE, R. C. P. **Aspectos da análise de resíduos**. 1984. 139 p. Dissertação (Mestrado em Estatística e Experimentação Agronômica) – Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, SP.

ROHATGI, K.V. **An introduction to probability theory and mathematical statistics**. New York: J. Wiley, 1976. 684 p.

SAS INSTITUTE/ STAT. Versão 6.12. Copyright©. 1989-1996.

SEBER, G. A. F. **Linear regression analysis**. New York: J. Wiley, 1977. 465 p.

SHAPIRO, S. S.; WILK, M. B. An analysis of variance teste for normality (complete sample). **Biometrika**, London, v. 52, n. 3, p. 519-611, Dec. 1965.

SNEDECOR, G. W.; COCHRAN, W. G. **Statistical methods**. 7. ed. Iowa: The Iowa State University Press, 1980. 507 p.

THERNEAU, T. M.; GRAMBSCH, P. M.; FLEMING, T. R. Martingale-based residuals for survival models. **Biometrika**, London, v. 77, n. 1, p.147-160, Apr. 1990.

WEISBERG, S. **Applied linear regression**. 2. ed. New York: Wiley, 1985. 324 p.