



UNIVERSIDADE FEDERAL DE SANTA CATARINA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO  
CURSO DE DOUTORADO EM ENGENHARIA DE PRODUÇÃO

**Isolde T. S. Previdelli**

**ESTIMADORES CORRIGIDOS PARA MODELOS  
NÃO-LINEARES GENERALIZADOS SUPERDISPERSADOS**

**Tese de Doutorado**

Florianópolis - SC  
Outubro de 2005

**Isolde T. S. Previdelli**

**ESTIMADORES CORRIGIDOS PARA MODELOS  
NÃO-LINEARES GENERALIZADOS SUPERDISPERSADOS**

Tese apresentada ao Programa de Pós – Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do grau de Doutor em Engenharia de Produção.

Área de Concentração: **Gestão da Qualidade e Produtividade**  
Orientador: **Prof. Robert Wayne Samohyl, Ph.D.**  
Co-orientador: **Prof. Gauss Moutinho Cordeiro, Ph.D.**

Florianópolis - SC  
Outubro/2005

Isolde T. S. Previdelli

# ESTIMADORES CORRIGIDOS PARA MODELOS NÃO-LINEARES GENERALIZADOS SUPERDISPERSADOS

Esta tese foi julgada e aprovada para a obtenção do grau de **Doutor em Engenharia de Produção** no Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina.

Florianópolis - SC, 27 de outubro de 2005.

---

Prof. Edson Pacheco Paladini, Dr.  
Coordenador do Programa de Pós-Graduação em Engenharia de Produção

## Banca Examinadora

---

Prof. Robert Wayne Samohyl, *Ph.D.*  
Universidade Federal de Santa Catarina  
**Orientador**

---

Prof. Gauss Moutinho Cordeiro, *Ph.D.*  
Universidade Federal Rural de Pernambuco  
**Co-orientador**

---

Prof<sup>a</sup>. Dora Maria Orth, Dr<sup>a</sup>.  
Universidade Federal de Santa Catarina  
**Moderadora/Examinadora**

---

Prof<sup>a</sup>. Linda Lee Ho, Dr<sup>a</sup>.  
Universidade de São Paulo  
**Examinadora Externa**

---

Prof<sup>a</sup>. Lúcia Pereira Barroso, Dr<sup>a</sup>.  
Universidade de São Paulo  
**Examinadora Externa**

---

Prof. Dalton Francisco de Andrade, *Ph.D.*  
Universidade Federal de Santa Catarina  
**Examinador**

---

Prof. Pedro Alberto Barbeta, Dr.  
Universidade Federal de Santa Catarina  
**Examinador**

Aos meus pequenos grandes homens  
Lorenzo e Enrico,  
meus filhos.

## AGRADECIMENTOS

Uma tese de doutorado, pelas suas características e por representar uma soma de conhecimentos, sempre será um trabalho coletivo. Portanto, identificar as colaborações, críticas e sugestões recebidas, antes e durante sua realização, pode conduzir a omissão de nomes. Independentemente desta ressalva, ousarei nominar estes benfeitores tendo, entretanto, consciência de possíveis esquecimentos.

Ao Prof. *Ph.D.* Robert W. Samohyl, pelo apoio, capacidade e pela flexibilidade de pensamento proporcionado ao longo das orientações deste estudo;

Ao Prof. *Ph.D.* Gauss M. Cordeiro, pois foi uma honra tê-lo como professor e amigo. Sua orientação, críticas e sugestões, desde a fase embrionária desta pesquisa até o final da mesma, tiveram e terão um valor inestimável;

À Prof. Dr<sup>a</sup>. Lúcia P. Barroso pela gentileza da disponibilidade e competência das sugestões durante todo o estudo;

Muito obrigada às minhas queridas amigas e parceiras Rosângela e Margareth, para vocês um brinde, "valeu a pena!";

Ao Previdelli, meu marido "Mo", pelo companheirismo e incentivo constante. Deu certo o desafio!;

Silvia, Karen, Kelen, Cláudia e Cláudia, desejo tê-las como amigas para sempre;

Aos meus pais e irmãos, meu amor de sempre;

Ao Vanderly, obrigada pela amizade sincera e ao Josmar, pelas dicas nos programas;

Édis, obrigada pela oportunidade e por acreditar em mim. Serei sempre grata;

Aos professores Doherty Andrade, Humberto Milani e Ivo Neitzel, sempre dispostos a discutir os dados em estudo;

Ao Programa de Pós-Graduação em Engenharia da Produção da Universidade Federal de Santa Catarina por me proporcionar a oportunidade de realizar o doutorado;

Ao Departamento de Estatística e à Universidade Estadual de Maringá que viabilizaram, através de uma importante política interna de qualificação do corpo docente, as condições e os meios para a conclusão deste curso.

PREVIDELLI, Isolde T. S. **Estimadores corrigidos para modelos não-lineares generalizados superdispersados.** Tese (Doutorado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2005.

## RESUMO

A teoria dos modelos lineares e não-lineares da família exponencial vem encontrando espaço cada vez maior entre pesquisadores que querem explorá-la tanto na aplicação quanto na melhoria dos métodos usuais e alternativos. Uma classe mais ampla é a dos modelos lineares e não-lineares generalizados superdispersados, nos quais se modelam os parâmetros da média e dispersão e que, além disso, incorpora a dispersão na função de variância. Essa classe tem sido utilizada de forma expressiva principalmente para dados onde haja superdispersão, isto é, onde a variância real seja maior que a predita pelo modelo. Os estimadores dos parâmetros desses modelos têm vieses de  $O(n^{-1})$  e costumam ser ignorados. Entretanto, para amostras de tamanho moderado a pequeno, esses vieses podem ser significativos, podendo atingir o mesmo valor do respectivo erro-padrão. Dentro desse contexto, é plausível fazer melhorias nos estimadores em áreas de atuação onde nem sempre é possível obter grandes amostras, como, por exemplo, na produção industrial, no controle de qualidade, em segmentos de produção de animais, nas engenharias, na farmacologia, na saúde, entre outras. Neste estudo foram obtidas expressões para o viés de  $O(n^{-1})$  para corrigir os estimadores de máxima verossimilhança dos parâmetros dos modelos não-lineares generalizados superdispersados. Para validar essa correção, foram executadas simulações de Monte Carlo e aplicações de dados advindos da área de engenharia da produção. Os resultados mostraram que estimativas de  $O(n^{-2})$  devem ser utilizadas nos modelos, principalmente em amostras de tamanho pequeno a moderado, podendo-se evidenciar que, quanto menor o tamanho da amostra, maior a necessidade de se fazerem correções. Em termos práticos, isto é, do ponto de vista econômico e operacional, é altamente positivo, pois o fato de se trabalhar com modelos de maior precisão traz como resultado produtos mais uniformes e, conseqüentemente, redução significativa de custos.

**Palavras-chave:** modelos não-lineares generalizados superdispersados; vieses de  $O(n^{-1})$ ; estimadores corrigidos.

PREVIDELLI, Isolde T. S. **Corrected Estimators for Overdispersed Generalized Nonlinear Models.** Thesis (Doctorate in Production Engineering) – Post-graduation Program in Production Engineering, Universidade Federal de Santa Catarina, Florianópolis, 2005.

### ABSTRACT

The theory of linear and nonlinear models of the exponential class has been growing among researchers who wish to explore it, either as to the application or as to the improvement of ordinary and alternative methods. One broader class is the overdispersed generalized linear and nonlinear models, in which the mean and dispersion parameters are molded, and the dispersion in function of the variance is also incorporated. This class has been significantly used, mainly for dispersion data, that is, where the real variance is higher than the one previewed by the model. The parameters estimators of these models have  $O(n^{-1})$  biases and are commonly overlooked. However, for samples varying from moderate to small size, these biases may be significant, reaching the same value as the respective standard errors. In that context, it is plausible to improve the estimators, in the areas where large samples are not possible to be obtained, such as, in the industrial production, quality control, animal production sector, engineering, pharmacology, health, among others. In this study, expressions for the  $O(n^{-1})$  bias were obtained to correct the estimators maximum likelihood of the parameters of generalized overdispersed nonlinear models. In order to validate such correction, Monte Carlo simulations and data applications from the production engineering area were carried out. Results showed that the estimates of  $O(n^{-2})$  must be used in the models, mainly in the samples of small and moderate size, making clear that the smaller the size of the sample, the bigger the necessity of correction. In a practical view, that is, from the economical and operational point of view, working with more precise models is very positive, because it results in more uniform products and, consequently, significant reduction of costs.

**Key words:** overdispersed generalized nonlinear models;  $O(n^{-1})$  biases; corrected estimators.

# Lista de Figuras

7.1	Procedimento da simulação para obtenção das estimativas pela distribuição passeio aleatório e seus respectivos vieses de $O(n^{-1})$ . . . . .	59
7.2	Estimativas de $O(n^{-1})$ e $O(n^{-2})$ para $\beta_1$ . . . . .	61
7.3	Estimativas de $O(n^{-1})$ e $O(n^{-2})$ para $\beta_2$ . . . . .	61
7.4	Estimativas de $O(n^{-1})$ e $O(n^{-2})$ para $\beta_3$ . . . . .	61
7.5	Estimativas de $O(n^{-1})$ e $O(n^{-2})$ para $\gamma_1$ . . . . .	62
7.6	Estimativas de $O(n^{-1})$ e $O(n^{-2})$ para $\gamma_2$ . . . . .	62
7.7	Representação esquemática do labirinto radial de oito braços . . . . .	64
7.8	Erros cometidos pelos ratos no tempo. . . . .	66
7.9	Erros cometidos pelos ratos não-lesionados no tempo. . . . .	67
7.10	Erros cometidos pelos ratos lesionados no tempo. . . . .	67
7.11	Modelo Poisson duplo ajustado . . . . .	74



# Lista de Tabelas

2.1	Condições para que as distribuições pertençam à família Exponencial . . . . .	16
4.1	Estimativas de $O(n^{-1})$ e $O(n^{-2})$ para o modelo inversa gaussiana. . . . .	30
4.2	Estimativas de $O(n^{-1})$ e $O(n^{-2})$ para a distribuição gama com função ligação recíproca. . . . .	33
5.1	Estimativas do submodelo Média da Poisson Dupla . . . . .	40
5.2	Estimativas do submodelo dispersão da Poisson dupla . . . . .	40
7.1	Estimativas de $O(n^{-1})$ e $O(n^{-2})$ da distribuição passeio aleatório . . . . .	60
7.2	Estimativas da média para o modelo Poisson . . . . .	72
7.3	Estimativas do submodelo média Poisson dupla . . . . .	72
7.4	Estimativas do submodelo dispersão Poisson dupla . . . . .	73
7.5	Vieses e estimativas corrigidas do modelo Poisson dupla . . . . .	74
7.6	Submodelo média da normal, estimativas, erros-padrão e vieses. . . . .	82
7.7	Submodelo dispersão da normal, estimativas, erros-padrão e vieses. . . . .	82
7.8	Submodelo média do passeio aleatório com réplica e sem réplica, estimativas, erros-padrão e vieses. . . . .	83
7.9	Submodelo dispersão do passeio aleatório com réplica e sem réplica, estimativas, erros-padrão e vieses. . . . .	83
F.1	Média dos erros cometidos pelos ratos isquêmicos(lesionados) . . . . .	118
F.2	Média dos erros cometidos pelos ratos não-lesionados . . . . .	118
G.1	Resistência à compressão (MPa) de corpos-de-prova de concreto. . . . .	124

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>1</b>
1.1	Definição do problema . . . . .	4
1.2	Objetivos . . . . .	6
1.3	Metodologia . . . . .	7
1.4	Estrutura . . . . .	7
<b>2</b>	<b>ASPECTOS TEÓRICOS</b>	<b>9</b>
2.1	Teoria assintótica . . . . .	9
2.1.1	Identities de Bartlett . . . . .	10
2.1.2	Expressão do viés dos EMV devida a Cox e Snell . . . . .	11
2.2	Família exponencial . . . . .	12
2.2.1	Família exponencial uniparamétrica . . . . .	12
2.2.2	Família exponencial de dispersão . . . . .	14
2.3	Estimação . . . . .	15
<b>3</b>	<b>FAMÍLIA EXPONENCIAL SUPERDISPERSADA</b>	<b>18</b>
3.1	Função densidade de probabilidade . . . . .	18
3.2	Parametrização da família exponencial superdispersada . . . . .	19
3.3	Distribuições pertencentes à família superdispersada . . . . .	19
<b>4</b>	<b>CÁLCULO DO VIÉS DE <math>O(n^{-1})</math> PARA OS MLG E OS MODELOS NÃO LINEARES</b>	<b>25</b>
4.1	Modelos lineares generalizados . . . . .	25
4.1.1	Estimadores de $O(n^{-1})$ . . . . .	26
4.1.2	Avaliação do MLG . . . . .	27

4.1.3	Cálculo do viés de $O(n^{-1})$ para os MLG . . . . .	28
4.1.4	Significado do viés . . . . .	28
4.2	Aplicação 1: vendas de telhas . . . . .	29
4.3	Modelos não-lineares da família exponencial (MNLFE) . . . . .	30
4.3.1	Estimadores de $O(n^{-1})$ . . . . .	30
4.3.2	Cálculo do viés do $O(n^{-1})$ para os MNLFE . . . . .	31
4.4	Aplicação 2: demanda de energia elétrica . . . . .	32
<b>5</b>	<b>MODELOS LINEARES GENERALIZADOS SUPERDISPERSADOS</b>	<b>34</b>
5.1	Modelo . . . . .	34
5.2	Estimadores de $O(n^{-1})$ de $\beta$ e $\gamma$ . . . . .	35
5.3	Cálculo do viés de $O(n^{-1})$ para os MLGS . . . . .	37
5.4	Aplicação 3: endoparasitos em ovinos . . . . .	38
5.4.1	Delineamento do experimento . . . . .	38
5.4.2	Modelo adotado . . . . .	39
5.4.3	Resultados . . . . .	39
<b>6</b>	<b>VIESES DE <math>O(n^{-1})</math> DOS MODELOS NÃO-LINEARES GENERALIZADOS SUPERDISPERSADOS</b>	<b>41</b>
6.1	Modelo . . . . .	41
6.1.1	Estimadores de MV para $\beta$ e $\gamma$ . . . . .	42
6.2	Vieses de $O(n^{-1})$ para $\hat{\beta}$ e $\hat{\gamma}$ . . . . .	43
6.2.1	Viés de $\hat{\beta}, \hat{\eta}$ e $\hat{\mu}$ . . . . .	44
6.2.2	Viés de $\hat{\gamma}, \hat{\tau}$ e $\hat{\phi}$ . . . . .	47
6.3	Caso especial . . . . .	52
<b>7</b>	<b>SIMULAÇÃO E APLICAÇÃO</b>	<b>54</b>
7.1	Simulação do modelo passeio aleatório . . . . .	54
7.1.1	Modelo não-linear generalizado superdispersado . . . . .	54
7.1.2	Procedimento da geração do modelo . . . . .	55
7.1.3	Obtenção dos vieses . . . . .	56
7.1.4	Resultados . . . . .	60

7.2	Aplicação dos MNLGS corrigidos a dados reais . . . . .	62
7.2.1	Aplicação 1: Modelando os erros cometidos num treinamento de aprendizagem e avaliando as estimativas de $O(n^{-2})$ . . . . .	62
7.2.2	Aplicação 2: Modelando a resistência do concreto e avaliando as estimativas de $O(n^{-2})$ . . . . .	75
<b>8</b>	<b>CONCLUSÕES</b>	<b>85</b>
8.1	Considerações finais . . . . .	85
8.2	Recomendações . . . . .	86
	<b>REFERÊNCIAS</b>	<b>86</b>
	<b>APÊNDICES</b>	<b>90</b>
<b>A</b>	<b>CÁLCULO DOS CUMULANTES PARA OS MNLGS</b>	<b>91</b>
A.1	Obtenção dos cumulantes para o cálculo do $B(\hat{\beta})$ . . . . .	91
A.2	Obtenção dos cumulantes para o cálculo do $B(\hat{\gamma})$ . . . . .	93
<b>B</b>	<b>PROGRAMA DO CÁLCULO DO VIÉS DE <math>O(n^{-1})</math> PARA OS MLG.</b>	<b>96</b>
<b>C</b>	<b>PROGRAMA DO CÁLCULO DO VIÉS <math>O(n^{-1})</math> PARA OS MODELOS NÃO LINEARES</b>	<b>98</b>
<b>D</b>	<b>PROGRAMA DO CÁLCULO DO VIÉS <math>O(n^{-1})</math> PARA OS MLGS</b>	<b>101</b>
<b>E</b>	<b>PROGRAMA DE SIMULAÇÃO</b>	<b>105</b>
<b>F</b>	<b>PROGRAMA PARA O CÁLCULO DO VIÉS DOS MNLGS</b>	<b>114</b>
<b>G</b>	<b>PROGRAMA PARA O CÁLCULO DO VIÉS DOS MNLGS</b>	<b>119</b>
<b>H</b>	<b>PROGRAMA PARA O CÁLCULO DO VIÉS DOS MNLGS</b>	<b>125</b>

# Lista de siglas

MLG	Modelo Linear Generalizado
MNLFE	Modelo Não Linear da Família Exponencial
MLGS	Modelo Linear Generalizado Superdispersado
MNLGS	Modelo Não Linear Generalizado Superdispersadado
EMV	Estimador de Máxima Verossimilhança
Viés $O(n^{-1})$	Viés de Primeira Ordem
Viés $O(n^{-2})$	Viés de Segunda Ordem

# Lista de símbolos

$\beta$	Parâmetro beta
$\hat{\beta}$	Estimador beta
$\tilde{\beta}$	Estimador beta corrigido
$\gamma$	Parâmetro gama
$\hat{\gamma}$	Estimador gama
$\tilde{\gamma}$	Estimador gama corrigido
$\mu$	Parâmetro média
$\phi$	Parâmetro dispersão phi
$\eta$	eta
$\tau$	tau
$\xi$	qsi
$\theta$	teta
$\delta$	delta
$\psi$	psi
$(s, t, u, \dots)$	Denotam as derivadas da log-verossimilhança das componentes do vetor $\beta$ do submodelo da média.
$(R, S, T, U, \dots)$	Denotam as derivadas da log-verossimilhança das componentes do vetor $\gamma$ do sub modelo de dispersão.
$k_s$	São todos os cumulantes de $O(n^{-1})$ .
$B(\hat{\beta}_r)$	Viés de primeira ordem de $\hat{\beta}_r$ .
$B(\hat{\gamma}_R)$	Viés de primeira ordem de $\hat{\gamma}_R$ .
$K^{-1}$	Inversa da matriz informação de Fisher
$K_{\beta,\beta}$ e $K_{\gamma,\gamma}$	Matriz informação de $\beta$ e $\gamma$

# CAPÍTULO 1

## INTRODUÇÃO

A teoria dos modelos lineares generalizados e não-lineares vem encontrando espaço cada vez maior entre pesquisadores, estudiosos e profissionais de diversas áreas do conhecimento que querem aplicá-la tanto em situações empíricas quanto na melhoria dos modelos teóricos existentes. Como consequência desse esforço intelectual, é possível identificar uma quantidade expressiva de publicações que se encontram atualmente disponíveis em livros, revistas científicas, periódicos de estatística e áreas correlatas, bem como na Internet.

Os modelos não-lineares têm contribuído de forma expressiva para patrocinar avanços do conhecimento da área de engenharia de produção. Como ilustração da importância e da frequência do uso desses conhecimentos, podem ser citadas as seguintes áreas, nas quais o emprego das técnicas e modelos tem sido comum: na indústria química, particularmente nos setores de refino de petróleo e petroquímica, para otimizar a extração dos produtos, subprodutos e derivados; nos segmentos da indústria de papel e celulose, servindo para subsidiar o controle e a melhoria dos processos produtivos; na economia, onde o uso das ferramentas e dos conhecimentos da engenharia de produção contribuem significativamente para a formulação de políticas nacionais, construção de modelos econométricos utilizados no mercado financeiro e mercado de capitais, controle de banco de dados sobre a qualidade do crédito utilizado por organizações como a SERASA, entre outras; na agroindústria, onde o uso das informações sobre curvas de crescimento otimizam as técnicas de manejo de rebanhos especializados, especificamente na produção de carne, leite e ovos, aumentando a competitividade desse segmento econômico; na engenharia civil, onde várias pesquisas têm contribuído para desenvolver materiais alternativos, reduzindo custos, aumentando a qualidade e proporcionando maior segurança para o usuário; na indústria farmacêutica, onde os modelos não-lineares são utilizados constantemente para determinar a quantidade certa das doses, para minimizar os efeitos colaterais, controlar a variabilidade, além de reduzir os custos de produção da droga, já que esta utiliza processo produtivo de larga escala. Os exemplos são quase infindáveis, e mostrar todas as possibilidades seria inviável para um estudo desta natureza.

Assim, a questão relevante que se materializa como escopo principal desta pesquisa pode ser

formulada da seguinte maneira: Por que melhorar os estimadores de máxima verossimilhança para os modelos não-lineares generalizados superdispersados (*MNLGS*)?

Uma resposta adequada para essa questão pode ser estruturada da seguinte forma: ao melhorar os estimadores de máxima verossimilhança, será possível trabalhar com modelos mais precisos, de maneira que, quanto menor o erro nas estimativas, maior será a chance de essas mesmas estimativas estarem mais próximas dos verdadeiros valores dos parâmetros.

Em termos práticos, essa melhoria nos estimadores resulta em economia de recursos, redução de riscos, aumento de eficiência, de produtividade e lucratividade nas atividades produtivas, podendo ser traduzida economicamente em milhares de dólares, beneficiando a economia, a sociedade e, por extensão, a qualidade de vida do cidadão.

A correção das estimativas nasceu com Bartlett (1937), que melhorou a aproximação da distribuição da estatística da razão de verossimilhanças mediante a utilização da distribuição qui-quadrado. A partir daí, todas as demais pesquisas nessa área são conhecidas como “correção de Bartlett”, sendo o pesquisador brasileiro Gauss M. Cordeiro, seu precursor no Brasil.

Os avanços e desenvolvimentos ocorridos nessa área podem ser encontrados em revistas e periódicos como *Biometrika*, *Annals of Statistics*, *Journal of the Royal Statistical Society*, *Econometric Reviews*, entre outros.

Uma classe-padrão de modelos muito utilizada para análise de dados tem sido a dos modelos lineares generalizados desenvolvidos por Nelder e Wedderburn (1972). Esses autores modelam os dados não-normais, ou seja, desenvolvem analogias com os modelos de regressão normal, tornando-os lineares por meio de um preditor linear numa função de ligação adequada para os valores esperados das observações, e além disso, permitindo a heterocedasticidade dos dados por meio da relação entre a média e a variância. Wei (1998), por sua vez, assegura que, se a suposição de linearidade na componente sistemática for relegada e se for permitido que qualquer valor esperado seja função dos parâmetros de regressão, têm-se assim, os modelos não-lineares generalizados, ou seja, modelos não-lineares da família exponencial.

Portanto, os modelos não-lineares da família exponencial são extensões dos modelos lineares generalizados e dos modelos de regressão normal não-lineares.

Muitos pesquisadores têm trabalhado na inferência desses modelos. Entre eles podem ser destacados Jørgensen (1983), McCullagh (1983), Cordeiro e Paula (1989), Cook e Tsai (1990), Pazman (1991), Wei e Shi (1994). É comum encontrar nos modelos de regressão a ocorrência da superdispersão, visto que em muitas aplicações há uma heterogeneidade muito grande na amostra, ou seja, verifica-se que a variabilidade dos dados é maior do que o valor predito pelo modelo, não podendo, portanto, tal heterogeneidade ser explicada pela função de variância. Isso ocorre, principalmente em dados na forma de proporções e contagens, em alguns campos de aplicação, tais como estudos epidemiológicos, econômicos, industriais, agrônômicos, análise de processos, entre outros. Para Jørgensen (1997), a principal característica está na distribuição dos erros, em particular no estudo das funções de variância para a família exponencial de



dispersão.

Considerando que há superdispersão nos dados, haverá perda de eficiência significativa na estimação da média quando se usam modelos de dispersão constante. Logo, é necessário modelar a dispersão por meio de metodologias adequadas para refletir essa característica, obtendo-se, dessa forma, modelos bem mais ajustados, com erros-padrão e intervalos de confiança mais estreitos.

As causas da presença da superdispersão, segundo Cox (1983) e Hinde e Demétrio (1998), podem ser resumidas da seguinte forma:

- i) as probabilidades de ocorrência não são constantes entre os indivíduos;
- ii) as observações de ocorrência não são independentes entre os indivíduos;
- iii) o processo adotado na coleta de dados não é adequado;
- iv) variáveis são omitidas na investigação do estudo;
- v) a variabilidade do material utilizado no experimento é uma característica da ocorrência do fenômeno.

As principais conseqüências de não se considerar a superdispersão no modelo levam a:

- i) obter erros-padrão incorretos, tendo como conseqüência direta a existência de erros na significância do modelo;
- ii) os critérios de seleção do modelo serão afetados, pois a *deviance* será grande e isso conduzirá a modelos complexos e, conseqüentemente, a interpretação e a precisão também estarão incorretas.

A investigação da superdispersão foi primeiramente vista por Fisher (1950), seguida nos últimos vinte anos por pesquisadores como Cox (1983), Pregibon (1984), Efron (1986, 1991, 1992), McCullagh e Nelder (1989), Dean e Lawless (1989), Zelterman e Chen (1989), Breslow (1990), Gelfand e Dalal (1990), Smyth e Gelfand (1993), Wei *et al.* (1997), Hinde e Demétrio (1998) e Smyth e Verbyla (1999), além de outros que também evidenciaram a importância dessa linha de pesquisa, tanto no diagnóstico quanto na apresentação de metodologias para acomodar a superdispersão nos modelos lineares, mistos e não-lineares, da família exponencial.

As principais propostas para acomodar a superdispersão foram apresentadas pelos autores abaixo e seguem a seguinte cronologia: Efron (1986) considera uma família exponencial dupla, enquanto que Lindsay (1986) propõe utilizar os modelos lineares generalizados mistos, ou seja, adicionar estruturas de efeitos aleatórios no preditor linear. Já Smyth (1989) trata os modelos lineares generalizados como covariáveis de dispersão. Mais tarde, Hinde e Demétrio (1998) afirmam que a superdispersão pode ser acomodada incluindo mais um parâmetro na função de variância, e Dey *et al.* (1997) consideram um modelo de regressão adicional para o parâmetro de dispersão e incorporam o parâmetro de dispersão na função de variância, enquanto que Smyth e Verbyla (1999) consideram um modelo de regressão adicional para o parâmetro de dispersão e

demonstram que o submodelo para a dispersão é um modelo linear generalizado gama e utiliza a aproximação de ponto de sela na log-verossimilhança.

Há várias propostas para ajustar os métodos de verossimilhança em modelos de regressão, entre as quais se destacam as apresentadas por Cox e Reid (1987) e McCullagh e Tibshirani (1990). Para os modelos de regressão normal linear, Smyth e Verbyla (1996) fornecem estimadores de máxima verossimilhança restrita ou residual (EMVR), que ajustam as estimativas dos parâmetros para serem aproximadamente não-viesadas e que podem ser consideradas mais confiáveis para pequenas amostras. Novamente, Smyth e Verbyla, em 1999, propõem uma generalização do EMVR para os modelos generalizados duplos e para os modelos não-lineares em 2001.

Uma abordagem bayesiana é utilizada por Gelman *et al.* (1995) e Dey *et al.* (1997) entre outros, afirmando que a principal vantagem é a facilidade computacional e, também, uma redução no viés das estimativas.

Outra proposta de se obter os estimadores de máxima verossimilhança aproximadamente não-viesados é feita por Cordeiro e McCullagh (1991), Paula (1992), Cordeiro (1993), Botter e Cordeiro (1998), Cordeiro e Botter (2001) e Cordeiro e Barroso (2003), seguindo a fórmula de Cox e Snell (1968).

Pode-se, ainda, encontrar uma expressão geral para o viés de  $O(n^{-1})$  em modelos de regressão não-lineares multivariados em Box (1971) e para a regressão logística em Pike *et al.* (1980). Dentro dessa ótica, Young e Bakir (1987) contribuem argumentando que a correção do viés melhora significativamente a estimação de regressão log-gama em MLG.

## 1.1 Definição do problema

Dey *et al.* (1997) definiram a classe dos modelos lineares generalizados superdispersados (*MLGS*), que tem sido utilizada, principalmente, para dados em que há superdispersão, isto é, onde a variância da variável resposta é maior que a especificada ou predita pelo modelo adotado. Essa classe adiciona um modelo de regressão para o parâmetro de dispersão, além de introduzir o parâmetro de dispersão na função de variância. Desse modo, a superdispersão é modelada através da componente sistemática.

Os *MLGS* seguem a mesma estrutura dos modelos lineares generalizados, que, nesse caso, são apresentados com duas componentes sistemáticas, uma para o submodelo da média e outro para o submodelo de dispersão. É comum, na prática, os parâmetros da média e dispersão assumirem modelos de regressão diferentes; além disso, a metodologia permite que as covariáveis sejam ou não excludentes.

É necessário observar, no entanto, que, se for permitido que essas componentes sistemáticas sejam funções não-lineares, ou seja, que as restrições de linearidade dos parâmetros do submodelo da média e dispersão, forem flexibilizados, têm-se então modelos não-lineares generalizados

superdispersados.

Os estimadores de máxima verossimilhança para os parâmetros desses modelos têm vieses de  $O(n^{-1})$ , que costumam ser ignorados na prática para grandes amostras; contudo, para amostras de tamanho moderado a pequeno, os valores destes vieses podem ser significativos, podendo chegar a ter os mesmos valores que os seus respectivos erros-padrão.

Para os MLGS, Cordeiro e Botter (2001) apresentam as fórmulas para os cálculos dos vieses de  $O(n^{-1})$  para os estimadores de máxima verossimilhança dos parâmetros. Os resultados das simulações demonstram que as correções dos estimadores têm menor impacto para amostras grandes, e as estimativas corrigidas têm erros-padrão menores para as amostras de tamanho reduzido a moderado. Concluem que, os estimadores de segunda ordem deveriam ser usados para conseguir estimativas de máxima verossimilhança melhores e mais confiáveis para amostras pequenas.

Dentro desse contexto, é muito comum trabalhar com amostras pequenas, principalmente em áreas de atuação onde nem sempre é possível obter grandes amostras, quer seja pelo custo, quer seja pela dificuldade inerente ao experimento. Situações reais com essas características podem ser encontradas nas áreas de engenharia de produção, principalmente no setor de melhoramento de qualidade, onde um dos objetivos é minimizar a variação do produto enquanto se controla a média, bem como nos segmentos de produção industrial e animal, na ciência florestal, farmacologia, nas áreas da saúde, nas engenharias, entre outras. Em consequência, a utilização desses estimadores corrigidos é mais adequada; além disso, quanto menores os erros nas estimativas, maior será a chance destas estarem próximas dos verdadeiros valores dos parâmetros, permitindo assim trabalhar com modelos mais precisos.

Segundo Ratkowsky (1983) e Paula (1991), as correções conduzem a bons melhoramentos em termos do viés e do erro quadrático médio; além do que, o cálculo do viés também serve como critério de eleição para melhor modelo ajustado.

Entretanto, é necessário que esses avanços teóricos tenham evidências empíricas e possam contribuir de forma inequívoca para melhoramentos de correções de erros em situações reais.

Dessa forma, esta pesquisa apresenta uma expressão matemática para o viés de  $O(n^{-1})$  dos estimadores de máxima verossimilhança para os parâmetros dos *MNLGS* e um estudo de simulação para diferentes tamanhos amostrais para avaliar os estimadores corrigidos.

Mostra também a correção do viés de primeira ordem dos estimadores em duas situações reais. Primeiramente, foram utilizadas os dados experimentais obtidos durante um teste de aprendizagem e memória espacial aplicado em ratos portadores de lesão cerebral isquêmica, ou seja, lesão causada pela falta de sangue no cérebro.

Esse experimento foi realizado no Laboratório de Farmacologia da Universidade Estadual de Maringá e serviu para consolidar a base teórica do estudo que procura verificar os efeitos colaterais provocados pela isquemia. É importante salientar que esta pesquisa tem como escopo nortear novas propostas de tratamento, contribuindo para ensaios clínicos úteis na cura de

pacientes vítimas de injúria isquêmica do cérebro.

A segunda aplicação, por sua vez, utilizou os dados oriundos de experimentos fatoriais com poucas replicações na área de engenharia civil, mais especificamente na avaliação da resistência do concreto aos efeitos de alguns parâmetros do processo produtivo com adição de cinza de casca de arroz. Esse estudo vem sendo desenvolvido por pesquisadores, desde 1996, no Laboratório de Resistência de Materiais da Universidade Católica de Pelotas.

Para a produção de um concreto de alta qualidade em que a resistência é a variável de maior importância, é necessário conhecer a dosagem de seus componentes. Essa qualidade leva em consideração diversos parâmetros, entre os quais a heterogeneidade da mistura dos componentes utilizados na produção do concreto. Uma grande variabilidade dessa mistura tem como consequência direta um aumento do consumo de cimento, elevando, portanto, os custos de produção. Além desses aspectos, quando da elaboração de cálculos estruturais para a construção de obras civis, leva-se em consideração o coeficiente mínimo de resistência que o concreto deve apresentar.

Portanto, esse experimento pode contribuir nas pesquisas que buscam desenvolver a utilização de materiais alternativos na produção de novos conceitos de concreto. Além de diminuir os problemas ambientais, visto ser uma prática comum a queima da casca de arroz, também pode reduzir os custos de produção.

## 1.2 Objetivos

A principal contribuição deste estudo foi obter uma fórmula geral analítica para os vieses de primeira ordem  $O(n^{-1})$  dos estimadores de máxima verossimilhança para os parâmetros dos modelos não-lineares generalizados superdispersados, validando por meio da simulação de Monte Carlo e ainda demonstrando a utilidade dessas correções em problemas típicos de Engenharia de Produção.

Em paralelo, e no sentido de legitimar o objetivo principal desta pesquisa, algumas ações foram desenvolvidas, caracterizando-se como objetivos específicos, podendo ser assim enumerados:

- i) Estender as propostas contidas nos estudos de Cordeiro e Botter (2001);
- ii) Implementar as fórmulas para o cálculo dos vieses de  $O(n^{-1})$  dos estimadores de máxima verossimilhança dos parâmetros dos modelos lineares generalizados, modelos não-lineares da família exponencial e modelos lineares generalizados superdispersados em situações reais no aplicativo SAS;
- iii) Disponibilizar os programas do item I com o intuito de popularizar estimativas de  $O(n^{-2})$ , por meio de uma home-page hospedada no site do Departamento de Estatística (DES) da Universidade Estadual de Maringá (UEM);

iv) Fazer estudos, via simulação de Monte Carlo, para os vieses de  $O(n^{-1})$  dos estimadores de máxima verossimilhança dos modelos não-lineares generalizados superdispersados com diferentes tamanhos amostrais;

v) Aplicar as correções dos vieses  $O(n^{-1})$  dos *MNLGS* na área de engenharia de produção, consolidando, portanto a teoria apresentada.

### 1.3 Metodologia

Para Becker (1998), Booth *et al.* (1995) e mais especificamente Köche (1997), este estudo insere-se na tipologia de pesquisa bibliográfica, a qual deve ser desenvolvida através do conhecimento disponível em teorias publicadas em livros, artigos e outras mídias disponíveis. Logo, identificam-se as teorias produzidas fazendo análises e avaliando-se as contribuições teóricas, no sentido de fornecer suporte para desenvolver, no caso, a expressão do viés de  $O(n^{-1})$  para os modelos superdispersados não-lineares.

No decorrer desse estudo são evidenciados os avanços das pesquisas na melhoria dos estimadores de máxima verossimilhança, tanto para os modelos lineares generalizados e generalizados superdispersados, bem como para os não-lineares. Esses desenvolvimentos podem ser vistos em Cordeiro e McCullagh, (1991), Paula (1992), Botter e Cordeiro (1998) e Cordeiro e Botter (2001) que apresentam os vieses de segunda ordem.

Os cálculos dos vieses para os modelos citados acima são todos obtidos pela fórmula de Cox e Snell (1968), por meio de expansão em série de Taylor e alguns cumulantes da função densidade de probabilidade do modelo em estudo.

### 1.4 Estrutura

Este estudo está organizado obedecendo à seguinte estrutura: o Capítulo 2 apresenta um resumo dos principais aspectos teóricos que serão utilizados no decorrer do estudo: teoria assintótica e família exponencial. O Capítulo 3 apresenta algumas distribuições pertencentes à família superdispersada. Por sua vez, o Capítulo 4 tem como objetivo apresentar os estimadores de  $O(n^{-2})$ , para alguns modelos de regressão, subdividindo-se em duas seções: a primeira seção descreve os vieses dos MLG dados por Cordeiro e McCullagh (1990); a segunda mostra os modelos não-lineares da família exponencial, apresentados por Paula (1991). Como ilustração, duas aplicações na área de engenharia de produção são discutidas, evidenciando os vieses. O Capítulo 5 enfoca a classe dos modelos lineares generalizados superdispersados (MLGS), permeando a metodologia de Dey *et al.* (1997), e também uma aplicação de dados superdispersados foi ajustada enfocando as suas respectivas correções. No Capítulo 6 é demonstrada passo a passo a fórmula geral para os vieses de  $O(n^{-1})$  para os estimadores de máxima verossimilhança dos parâmetros para os modelos não-lineares generalizados superdispersados. O penúltimo, enfoca um estudo de simulação pelo método de Monte Carlo, bem como a aplicação da fórmula dos

vieses de  $O(n^{-1})$  a dados reais nas áreas de análises clínicas/farmacologia e engenharia civil. Finalmente, o último capítulo apresenta as conclusões e as considerações finais dos vieses de primeira ordem para MNLGS.

Todas as estimativas dos parâmetros dos modelos e os seus respectivos vieses, bem como as simulações, foram obtidos por meio de programas implementados no aplicativo SAS, versão 8.5, disponível no Departamento de Estatística da Universidade Estadual de Maringá, no endereço [www.des.uem.br/~isolde](http://www.des.uem.br/~isolde).

# CAPÍTULO 2

## ASPECTOS TEÓRICOS

Este capítulo apresenta um resumo da teoria assintótica dos estimadores de máxima verossimilhança e um grande número de distribuições pertencentes à família exponencial, que tem importância fundamental em probabilidade e estatística. Muitas vezes, uma só distribuição é considerada, pois pode ser que esta ocorra como limite de muitas outras distribuições. Em outros casos, a família exponencial é usada em uma ampla variedade de fenômenos aleatórios.

### 2.1 Teoria assintótica

A teoria assintótica da verossimilhança é de grande valia, pois fornece simplificações significativas para problemas que envolvem grandes amostras, o que já não ocorre para amostras de tamanho pequeno, uma vez que, para obter soluções exatas, existe um alto grau de complexidade. Desse modo, é inevitável a obtenção de soluções aproximadas.

Há várias técnicas de aproximação; no entanto, para essa teoria a mais utilizada é feita por meio das expansões em série de Taylor, que consistem em linearizar um estimador  $\hat{\theta}$ . Com isso, apresentam-se uma variedade de métodos e testes estatísticos equivalentes.

A teoria assintótica produz simplificações importantes nos cálculos, conseqüentemente produzindo resultados simples e fáceis de serem interpretados. No entanto, em pequenas amostras, isso não necessariamente ocorre, podendo tornar-se complexa a obtenção de alguns resultados.

Segundo Cordeiro (1999), as soluções aproximadas mostram-se razoáveis para problemas de grandes amostras diante das seguintes situações:

i) quando, inicialmente, não se tem uma solução exata para o problema ou não é viável devido à complicação matemática;

ii) quando “realmente” não se tem solução exata, ou seja, não há uma região de confiança exata para o parâmetro de interesse.

Desse modo, é razoável obter soluções aproximadas.

As propriedades assintóticas dos estimadores de máxima verossimilhança de maior interesse são: não-viesada, consistência, unicidade, normalidade, eficiência e suficiência, sendo que estas são válidas somente quando o número de observações tende a infinito e essa é a base da teoria assintótica de primeira ordem –  $O(n^{-1})$ .

A teoria assintótica de segunda ordem -  $O(n^{-2})$ - é um melhoramento nos vieses associados as estatísticas de  $O(n^{-1})$ . Essa teoria tem crescido rapidamente nas correções dos vieses das estimativas de máxima verossimilhança e extensões para as correções tipo-Bartlett de outras estatísticas de testes.

### 2.1.1 Identidades de Bartlett

Considere  $\ell = \ell(\theta)$  a log-verossimilhança total de um problema regular e  $\theta$  um vetor de  $p$  parâmetros.

Definem-se os momentos conjuntos de derivadas de  $\ell(\theta)$  por:

$$\begin{aligned}\mu_r &= E(U_r) = E\left(\frac{\partial \ell}{\partial \theta_r}\right), & \mu_{rs} &= E(U_{rs}) = E\left(\frac{\partial^2 \ell}{\partial \theta_r \partial \theta_s}\right), \\ \mu_{r,s} &= E(U_r U_s) = E\left(\frac{\partial \ell}{\partial \theta_r} \frac{\partial \ell}{\partial \theta_s}\right), \\ \mu_{r,st} &= E(U_r U_{st}) = E\left(\frac{\partial \ell}{\partial \theta_r} \frac{\partial^2 \ell}{\partial \theta_s \partial \theta_t}\right), \text{ e assim por diante.}\end{aligned}$$

Os cumulantes conjuntos são expressos em termos dos momentos e assumem a notação  $\kappa$ . O cumulante  $\kappa_{r,s}$  é obtido como se segue :

$$\begin{aligned}\kappa_{r,s} &= E(U_r U_s) - E(U_r) \cdot E(U_s) \text{ e como} \\ &\quad \text{a esperança da função escore é igual a zero, isto é,} \\ E(U_r) &= 0 \text{ e portanto } \mu_r = 0, \text{ e assim,} \\ \kappa_{r,s} &= \mu_{r,s} .\end{aligned}$$

Os demais cumulantes são obtidos de maneira análoga com a seguinte notação:

$$\begin{aligned}\kappa_{rs} &= \mu_{rs}, & \kappa_{rs,t} &= \mu_{rs,t}, \\ \kappa_{rs,tu} &= \mu_{rs,tu} - \mu_{rs} \mu_{tu}, & \kappa_{r,s,t} &= \mu_{r,s,t}, \\ \kappa_{r,s,t,u} &= \mu_{r,s,t,u} - \sum_{(k)} \mu_{rs} \mu_{tu},\end{aligned}$$

onde  $\sum_{(k)}$  é o somatório que cobre todas as  $k$  combinações de índices.

Não há independência entre os cumulantes e os momentos, no entanto satisfazem algumas



equações que auxiliam muito nos cálculos, são válidas sob condições de regularidade e são denominadas “identidades de Bartlett”

O ponto principal na dedução das identidades de Bartlett é que só se pode inverter as operações de diferenciação com respeito a  $\theta$  e a integração com respeito a  $y$ , isto é,

$$\frac{\partial}{\partial \theta} E \{t(y)\} = \int t(y) \frac{\partial f(y; \theta)}{\partial \theta} dy,$$

que é válido para qualquer estatística  $t(Y)$ .

As principais identidades de Bartlett podem ser vistas em detalhes em Cordeiro (1999) e são utilizadas para facilitar a obtenção dos cumulantes.

McCullagh e Nelder (1989) afirmam que, para os cumulantes nos modelos lineares generalizados, há uma propriedade importante, que é a validade da permutação dos índices, ou seja, a invariância dos cumulantes no sentido de que  $\kappa_{rs,t} = \kappa_{rt,s} = \kappa_{st,r}$ . Essa propriedade também é válida para os modelos lineares generalizados com covariáveis de dispersão e os MLGS, conforme comentam Cordeiro *et al* (2003).

### 2.1.2 Expressão do viés dos EMV devida a Cox e Snell

As correções foram originalmente propostas por Bartlett (1937), que fez uma transformação escalar aplicada à estatística da razão de verossimilhanças para inferência em modelos paramétricos. Mas foi no ano de 1953 que Bartlett deu uma contribuição considerável para a teoria estatística, ao deduzir uma fórmula para o viés de  $O(n^{-1})$  para os estimadores de máxima verossimilhança no caso iid (observações independentes e identicamente distribuídas) para os modelos uniparamétricos. Esse resultado foi estendido posteriormente por Cox e Snell (1968) para os modelos multiparamétricos, cujos dados não são necessariamente identicamente distribuídos.

A expressão do viés é dada por

$$B(\hat{\theta}_r) = \sum \kappa^{ar} \kappa^{tu} \{\kappa_{rtu}/2 + \kappa_{rt,u}\}, \quad (2.1)$$

onde  $B(\hat{\theta}_r)$  é o viés de primeira ordem de  $\hat{\theta}_r$  e  $\sum$  representa o somatório para todos os  $p$  parâmetros, isto é,  $\theta_1, \dots, \theta_r, \dots, \theta_p$ .

Todos os vieses apresentados neste estudo são originados a partir da fórmula (2.1).

O estimador de  $O(n^{-2})$  é dado pela diferença entre o EMV,  $\hat{\theta}_r$ , e o viés da estimativa, denotado por  $B(\hat{\theta}_r)$ , e o resultado será denominado por  $\tilde{\theta}_r$ , ou seja,

$$\tilde{\theta}_r = \hat{\theta}_r - B(\hat{\theta}_r),$$

sendo portanto, o estimador de segunda ordem. Os demais EMV corrigidos seguem a mesma notação acima.

## 2.2 Família exponencial

A família exponencial reúne uma rica coleção de distribuições, com um pequeno número de parâmetros (usualmente 1 ou 2) que descrevem adequadamente um fenômeno dado.

### 2.2.1 Família exponencial uniparamétrica

Essa família é importante e bastante utilizada devido às inúmeras aplicações e às facilidades de suas propriedades matemáticas na teoria estatística.

Considere a função densidade de probabilidade

$$\pi(\mathbf{y}; \boldsymbol{\theta}) = \frac{1}{\zeta(\boldsymbol{\theta})} \exp\{-\alpha(\boldsymbol{\theta}) d(\mathbf{y}) + v(\mathbf{y})\}, \quad (2.2)$$

onde  $\theta$  é um parâmetro escalar e  $\zeta(\cdot)$ ,  $\alpha(\cdot)$ ,  $d(\cdot)$  e  $v(\cdot)$  são funções conhecidas e consideradas contínuas nas quatro primeiras derivadas com relação a  $\theta$ ,  $\zeta(\cdot)$  é um valor positivo, e  $\frac{d\alpha(\boldsymbol{\theta})}{d\boldsymbol{\theta}}$  e  $\frac{d\beta(\boldsymbol{\theta})}{d\boldsymbol{\theta}}$  são diferentes de zero para todo no espaço paramétrico, onde  $\beta(\boldsymbol{\theta})$  é definida como

$$\beta(\boldsymbol{\theta}) = (d\zeta(\boldsymbol{\theta})/d\boldsymbol{\theta}) \boldsymbol{\theta} (\zeta(\boldsymbol{\theta}) d\alpha(\boldsymbol{\theta})/d\boldsymbol{\theta})^{-1}.$$

A família de distribuições em (2.2) tem muitas distribuições, como casos especiais, que são apresentadas abaixo, conforme Cordeiro et al. (1995).

**i)** Binomial ( $0 < \theta < 1$ ,  $m \in N$ ,  $m$  conhecidos,  $y = 0, 1, 2, \dots, m$ )

$$\alpha(\theta) = -\log\{\theta/(1-\theta)\}, \zeta(\theta) = (1-\theta)^{-m}, d(y) = y, v(y) = \log\binom{m}{y}.$$

**ii)** Binomial negativa ( $\theta > 0$ ,  $\gamma > 0$ ,  $\gamma$  conhecidos,  $y = 0, 1, 2, \dots$ ) :

$$\alpha(\theta) = -\log\theta, \zeta(\theta) = (1-\theta)^{-\gamma}, d(y) = y, v(y) = \log\binom{\gamma+y-1}{y}.$$

**iii)** Poisson ( $\theta > 0$ ,  $y = 0, 1, 2, \dots$ ) :  $\alpha(\theta) = -\log(\theta)$ ,  $\zeta(\theta) = \exp\{\theta\}$ ,

$$d(y) = y, v(y) = -\log(y!).$$

**iv)** Poisson truncada ( $\theta > 0$ ,  $y = 0, 1, 2, \dots$ ) :  $\alpha(\theta) = -\log(\theta)$ ,

$$\zeta(\theta) = e^\theta(1-e^{-\theta}), d(y) = y, v(y) = -\log y!.$$

**v)** Série logarítmica ( $0 < \theta < 1$ ,  $y = 0, 1, 2, \dots, m$ ) :  $\alpha(\theta) = -\log\theta$ ,

$$\zeta(\theta) = -\log(1-\theta), d(y) = y, v(y) = -\log(y).$$

**vi)** Série de potência ( $\theta > 0$ ,  $\alpha_y > 0$ ,  $y = 0, 1, 2, \dots$ ) :  $\alpha(\theta) = -\log\theta$ ,

$$\zeta(\theta) = \sum_{y=0}^{\infty} \alpha_y \theta^y, \quad d(y) = y, \quad v(y) = \log(\alpha_y).$$

**vii)** Zeta ( $\theta > 0, y = 0, 1, 2, \dots$ ) :  $\alpha(\theta) = \theta + 1, \zeta(\theta) = Zeta(\theta + 1),$

$$d(y) = \log(y), \quad v(y) = 0.$$

**viii)** Hipergeométrica não-centrada ( $\theta > 0, m_1, m_2, r$  inteiros positivos e conhecidos,  $\alpha = \max\{0, r - m_2\} \leq y \leq \min\{m_1, r\} = b$ ) :

$$\alpha(\theta) = \theta, \zeta(\theta) = D_o(\theta), \quad d(y) = -y, \quad v(y) = \log \left\{ \binom{m_1}{y} \binom{m_2}{r-y} \right\}.$$

**ix)** Maxwell ( $\theta > 0, y > 0$ ) :  $\alpha(\theta) = (2\theta^2)^{-1}, \zeta(\theta) = \theta^3, \quad d(y) = y^2,$

$$v(y) = \log \left( y^2 \sqrt{2/\pi} \right).$$

**x)** Gama ( $\kappa > 0, \theta > 0, y > 0$ ) :

a)  $\kappa$  conhecido:  $\alpha(\theta) = \theta, \zeta(\theta) = \theta^{-\kappa}, \quad d(y) = y,$

$$v(y) = (\kappa - 1) \log y - \log \{ \Gamma(\kappa) \};$$

b)  $\theta$  conhecido:  $\alpha(\kappa) = -(\kappa - 1), \zeta(\kappa) = \theta^{-\kappa} \Gamma(\kappa), \quad d(y) = \log y,$

$$v(y) = -\theta y, \quad \text{onde } \Gamma \text{ é função gama.}$$

**xi)** Distribuições do sistema de Burr ( $\theta > 0, b > 0, b$  conhecido,  $y > 0$ ) :

$$\zeta(\theta) = c(\theta) / \theta, \quad d(y) = -\log G(y), \quad v(y) = \log \{ |d \log G(y) / dy| \},$$

onde as funções  $c(\cdot)$  e  $G(\cdot)$  têm valores reais positivos.

Diferentes escolhas para  $c(\theta)$  e  $G(y)$  implicam diferentes distribuições.

**xii)** Rayleigh ( $\theta > 0, \theta > 0, y > 0$ ) :  $\alpha(\theta) = \theta^{-2}, \zeta(\theta) = \theta^2, \quad d(y) = y^2,$

$$v(y) = \log(2y).$$

**xiii)** Pareto ( $\theta > 0, \kappa > 0, y > \kappa$ , com  $\kappa$  conhecido) :  $\alpha(\theta) = \theta + 1,$

$$\zeta(\theta) = (\theta \kappa^\theta)^{-1}, \quad d(y) = \log(y), \quad v(y) = 0.$$

**xiv)** Weibull ( $\theta > 0, \phi > 0$ , sendo  $\phi$  conhecido,  $y > 0$ ) :  $\alpha(\theta) = \theta^{-\phi},$

$$\zeta(\theta) = \theta^\phi, \quad d(y) = y^\phi, \quad v(y) = \log(\phi) + (\phi - 1) \log y.$$

**xv)** Potência ( $\theta > 0, \phi > 0$ , sendo  $\phi$  conhecido,  $0 < y < \phi$ ) :  $\alpha(\theta) = 1 - \theta^\phi,$

$$\zeta(\theta) = \theta^{-1} \theta^\phi, \quad d(y) = \log y, \quad v(y) = 0.$$

**xvi)** Laplace ( $\theta > 0, -\infty < \kappa < \infty$ , sendo  $\kappa$  conhecido,  $y > 0$ ) :  $\alpha(\theta) = \theta^{-1},$

$$\zeta(\theta) = 2\theta, \quad d(y) = |y - \kappa|, \quad v(y) = 0.$$

**xvii)** Valor extremo ( $-\infty < \theta < \infty, \phi > 0$ , sendo  $\phi$  conhecido,  $-\infty < y < \infty$ ) :

$$\alpha(\theta) = \exp(\theta/\phi), \quad \zeta(\theta) = \phi \exp(-\theta/\phi), \quad d(y) = \exp(-y/\phi), \quad v(y) = -y/\phi.$$

**xviii)** Valor extremo truncado ( $\theta > 0, y > 0$ ) :  $\alpha(\theta) = \theta^{-1}, \zeta(\theta) = \theta,$

$d(y) = \exp(y) - 1, v(y) = y.$

**xix)** Lognormal( $\theta > 0, \mu > 0$ , sendo  $\mu$  conhecido,  $y > 0$ ) :  $\alpha(\theta) = \theta^{-2},$

$\zeta(\theta) = \theta, d(y) = (\log y - \mu^2/2), v(y) = -\log y + \{\log(2\pi)\}/2.$

**xx)** Normal ( $\theta > 0, -\infty < \mu < \infty, -\infty < y < \infty,$ ) :

a)  $\mu$  conhecido:  $\alpha(\theta) = (2\theta)^{-1}, \zeta(\theta) = \theta^{\frac{1}{2}}, d(y) = (y - \mu)^2,$

$v(y) = -\{\log(2\pi)\}/2.$

b)  $\theta$  conhecido:  $\alpha(\mu) = -\mu/\theta, \zeta(\mu) = \exp\{\mu^2/(2\theta)\}, d(y) = y,$

$v(y) = -\{y^2 + \log(2\pi\theta)\}/2.$

**xxi)** Inversa gaussiana ( $\theta > 0, \mu > 0, y > 0$ ) :

a)  $\mu$  conhecido:  $\alpha(\theta) = \theta, \zeta(\theta) = \theta^{-\frac{1}{2}}, d(y) = (y - \mu)^2 / (2\mu^2 y),$

$v(y) = -\{\log(2\pi y^3)\}/2.$

b)  $\theta$  conhecido:  $\alpha(\mu) = (2\mu^2)^{-1}, \zeta(\mu) = \exp\{-\mu/\theta\}, d(y) = y,$

$v(y) = -\theta(2y)^{-1} + [\log\{\theta/(2\pi y^3)\}]/2.$

**xxii)** McCullagh ( $\theta > -1/2, -1 \leq \mu \leq 1$ , sendo  $\mu$  conhecido,  $0 < y < 1$ ) :

$\alpha(\theta) = -\theta, \zeta(\theta) = 4\theta B(\theta + 1/2, 1/1), d(y) = \log[y(1-y) / \{(1+\mu)^2 - 4\mu y\}],$

$v(y) = -[\log\{y(1-y)\}]/2.$

**xxiii)** Von Mises ( $\theta > 0, \theta < \mu < 2\pi$ , sendo  $\mu$  conhecido,  $0 < y < 2\pi$ ) :

$\alpha(\theta) = -\theta, \zeta(\theta) = 2\pi I_0(\theta), d(y) = \cos(y - \mu), v(y) = 0.$

**xxiv)** Secante hiperbólica generalizada ( $-\pi/2 \leq \theta \leq \pi/2, 0 < y < 1, r > 0$ ,

sendo  $r$  conhecido):  $\alpha(\theta) = \theta, \zeta(\theta) = \pi(\sec \theta)^r, d(y) = -\pi^{-1} -$

$\log\{y/(1-y)\}, v(y) = -(1/2)\log\{y/(1-y)\}.$

Existem muitas outras que poderiam ser citadas.

## 2.2.2 Família exponencial de dispersão

Segundo Wei (1998), uma família exponencial de dispersão  $P_{\theta, \phi}$  de um vetor de variáveis aleatórias  $Y = (y_1, \dots, y_n)^T$ , tem função de densidade ou função de probabilidade igual a

$$\pi(\mathbf{y}; \boldsymbol{\theta}, \phi) = \exp\{\phi[\mathbf{y}^T \boldsymbol{\theta} - b(\boldsymbol{\theta}) - c(\mathbf{y})] - \frac{1}{2}s(\mathbf{y}, \phi)\}, \quad (2.3)$$

onde  $b(\cdot)$ ,  $c(\cdot)$  e  $s(\cdot, \cdot)$  são funções específicas,  $\theta$  é o parâmetro natural e  $\phi$  é o parâmetro de dispersão, sendo que a média e a variância são dadas por

$$\boldsymbol{\mu} = E(\mathbf{Y}) = \frac{\partial b(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \text{ e } Var(\mathbf{Y}) = \phi \frac{\partial^2 b(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}.$$

Uma distribuição não é membro da família exponencial quando:

- i) a distribuição não pode ser escrita na forma definida em (2.3); ou
- ii) o domínio depende de algum subconjunto de parâmetros.

A família exponencial de dispersão reúne muitas das distribuições comumente usadas na estatística e tem importância cada vez maior na teoria estatística, sendo que, caso  $\phi$  seja conhecido em (2.3), representa a família exponencial de um parâmetro. No entanto, se  $\phi$  for desconhecido, poderá ou não pertencer à família exponencial de dispersão.

Algumas distribuições são apresentadas por Casella e Berger (2001), conforme Tabela 2.1 a seguir.

Há muitas outras distribuições que poderiam ser listadas.

## 2.3 Estimação

Dentre os métodos de estimação, o de máxima verossimilhança é um dos mais utilizados. Algumas de suas propriedades podem ser vistas em Nelder e Wedderburn (1972), que apresentam discussões sobre a unicidade das estimativas de máxima verossimilhança, e Cordeiro (1986), que declara terem esses estimadores propriedades ótimas, consistência e eficiência assintótica.

Alguns autores, como Morris (1982, 1983), Jørgensen (1987), McCullagh e Nelder (1989) e Wei (1998), têm utilizado o logaritmo da verossimilhança para os parâmetros  $\theta$  e  $\phi$ , denotado por  $\ell(y; \theta, \phi)$ , para um conjunto de observações no qual, dependendo da distribuição o parâmetro  $\theta$  pode ser a média  $\mu$  e  $\phi$  o parâmetro de perturbação ou dispersão.

Os EMV para os parâmetros  $\theta$  e  $\phi$ , que serão denotados por  $\hat{\theta}$  e  $\hat{\phi}$  respectivamente, são obtidos solucionando-se o sistema de equações a seguir:

$$\begin{cases} \frac{\partial \ell(y; \theta, \phi)}{\partial \theta} = U(\theta) = 0 \\ \frac{\partial \ell(y; \theta, \phi)}{\partial \phi} = U(\phi) = 0, \end{cases}$$

ou seja, é preciso encontrar as derivadas parciais de primeira ordem de  $\ell(\mathbf{y}; \theta, \phi)$  com relação a  $\theta$  e  $\phi$  e em seguida solucionar o sistema de equações.

Cordeiro (1986) e Paula (2000) demonstram em detalhes que os parâmetros  $\theta$  e  $\phi$  são ortogonais, ou seja, que  $E\left(\frac{\partial^2 \ell}{\partial \theta \partial \phi}\right) = 0$ , logo podem ser estimados separadamente.

Para resolver esse sistema de equações é necessário recorrer a métodos numéricos, visto que

Tabela 2.1: Condições para que as distribuições pertençam à família Exponencial

Distribuições comuns	Pertence a família exponencial	Comentários
Bernoulli	sim	Se $p \in (0,1)$
Binomial	sim	Se $p \in (0,1)$ , parâmetro único (isto é, $n$ fixo)
Uniforme Discreta	não	O domínio depende dos parâmetros $N$
Geométrica	sim	Se $p \in (0,1)$
Hipergeométrica	não	Distribuição não pode ser escrita da forma requerida
Binomial Negativa	sim	Se $p \in (0,1)$ , parâmetro único (isto é, $r$ fixo)
Poisson	sim	Se $\lambda > 0$
Beta	sim	Pode ser da família com um ou dois parâmetros
Cauchy	não	A distribuição não pode ser escrita da forma requerida
$\chi^2$	sim	$\omega(p) = p$ , deve ser considerado como contínua
Exponencial Dupla	sim	Se $\sigma > 0$ parâmetro único (isto é, $\mu$ fixo)
Exponencial	sim	Se $\beta > 0$
F	não	A distribuição não pode ser escrita da forma
Gama	sim	Pode ser da família com 1 ou 2 parâmetros
Logística	não	A distribuição não pode ser escrita da forma requerida
Log-normal	sim	Caso especial, desde que FGM não exista
Normal	sim	Pode ser da família com 1 ou 2 parâmetros
Pareto	sim	Se $\beta > 0$ , parâmetro único, isto é, $\alpha$ fixo
t	não	A distribuição não pode ser escrita da forma requerida
Uniforme	não	O domínio depende dos parâmetros $N$
Weibull	sim	Se $\beta > 0$ , parâmetro único, isto é, $\gamma$ fixo
Hipergeométrica não central	sim	Se $\theta > 0$ , desde que $N$ , $M$ e $K$ seja considerados como fixos

geralmente as equações são não-lineares. Há vários métodos matemáticos, sendo que nesse estudo foi utilizado o de Newton-Raphson, uma vez que é disponibilizado nas procedures NLP e NLMIXED do aplicativo SAS. Esse método consiste em expandir em série de Taylor até primeira ordem a equação de verossimilhança  $U(\boldsymbol{\theta}) = 0$ , na vizinhança de uma tentativa inicial  $\boldsymbol{\theta}_0$ . A convergência é rápida, no entanto depende fortemente dos valores iniciais. Formalizando tem-se:

$$\frac{\partial \ell(y; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \approx \frac{\partial \ell(y; \boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} + \frac{\partial^2 \ell(y; \boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}^T \partial \boldsymbol{\theta}} (\boldsymbol{\theta} - \boldsymbol{\theta}_0) = 0,$$

cuja solução na  $i$ -ésima iteração é :

$$\boldsymbol{\theta}^i = \boldsymbol{\theta}^{i-1} + \left( -\frac{\partial^2 \ell(y; \boldsymbol{\theta}^{i-1})}{\partial \boldsymbol{\theta}^T \partial \boldsymbol{\theta}} \right)^{-1} \frac{\partial \ell(y; \boldsymbol{\theta}^{i-1})}{\partial \boldsymbol{\theta}}. \quad (2.4)$$

Cabe ressaltar que, caso não se tenha um aplicativo com esse método, usualmente utiliza-se o escore de Fisher, que substitui  $-\frac{\partial^2 \ell(y; \boldsymbol{\theta}^{i-1})}{\partial \boldsymbol{\theta}^T \partial \boldsymbol{\theta}}$  na equação (2.4) por  $E\left(-\frac{\partial^2 \ell(y; \boldsymbol{\theta}^{i-1})}{\partial \boldsymbol{\theta}^T \partial \boldsymbol{\theta}}\right)$ . De modo análogo é obtido o processo iterativo para  $\phi$ .

Para os modelos da família exponencial, são de interesse as funções escore e as matrizes de informação de Fisher que são definidas e apresentadas aqui.

### Função escore de $\theta$ e $\phi$

Denotada por  $U(\boldsymbol{\theta})$  e  $U(\phi)$ , que são definidas como sendo as derivadas de ordem um da log-verossimilhança. Logo,

$$U(\boldsymbol{\theta}) = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \text{ e } U(\phi) = \frac{\partial \ell(\phi)}{\partial \phi}$$

Note que se  $\boldsymbol{\mu}$  é função de  $\theta$ , então  $U(\boldsymbol{\mu}) = \frac{\partial \ell}{\partial \boldsymbol{\theta}} \frac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{\mu}}$ .

### Matriz de informação de Fisher de $\theta$ e $\phi$ .

Serão denotadas por  $K(\boldsymbol{\theta})$  e  $K(\phi)$ , definidas como a esperança negativa para a segunda derivada da log-verossimilhança, ou seja,

$$K(\boldsymbol{\theta}) = E\left(-\frac{\partial^2 \ell}{\partial \boldsymbol{\theta}^T \partial \boldsymbol{\theta}}\right) \text{ e } K(\phi) = E\left(-\frac{\partial^2 \ell}{\partial \phi^2}\right).$$

# CAPÍTULO 3

## FAMÍLIA EXPONENCIAL SUPERDISPERSADA

A família exponencial de dispersão incorpora somente o parâmetro média na função de variância e, em alguns problemas práticos, pode não acomodar a superdispersão. Gelfand e Dalal (1990) estendem essa família propondo incorporar o parâmetro de dispersão na função de variância e com isso capturar a superdispersão.

O objetivo deste capítulo é exibir algumas distribuições pertencentes à família exponencial superdispersada.

### 3.1 Função densidade de probabilidade

Considere as variáveis aleatórias independentes  $Y_1, \dots, Y_n$  com uma função de probabilidade para cada  $Y_i$  na família exponencial de dois parâmetros apresentada por Gelfand e Dalal (1990) da seguinte forma:

$$\pi(y_i; \theta_i, \phi_i) = A(y_i) \exp\{\theta_i y_i + \phi_i T(y_i) - \rho(\theta_i, \phi_i)\}. \quad (3.1)$$

Se  $T(Y)$  for convexa, então, para uma média comum, a variância de  $Y$  será aumentada em  $\phi$ . Em outras palavras, se  $\phi$  cresce, logo a  $Var(Y)$  também cresce. Por esse motivo, o parâmetro  $\phi$  é, comumente, denominado parâmetro de dispersão, pois ele captura a noção de superdispersão.

Gelfand e Dalal (1990) demonstram que, sob condições de regularidade usuais, têm-se as seguintes propriedades:

$$\text{seja } \rho^{(r,s)}(\boldsymbol{\theta}, \boldsymbol{\phi}) = \frac{\partial \rho^{r+s}(\boldsymbol{\theta}, \boldsymbol{\phi})}{\partial \boldsymbol{\theta}^r \partial \boldsymbol{\phi}^s}, \text{ então } \boldsymbol{\mu} = \rho^{(1,0)}(\boldsymbol{\theta}, \boldsymbol{\phi}) \text{ e } Var(\mathbf{y}) = \rho^{(2,0)}(\boldsymbol{\theta}, \boldsymbol{\phi}).$$



### 3.2 Parametrização da família exponencial superdispersada

Por conveniência, Dey *et al.* (1997) parametrizaram a expressão (3.1) por meio da média, obtendo

$$\pi(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\phi}) = A(\mathbf{y}) \exp \left\{ (\mathbf{y} - \boldsymbol{\mu}) \Psi^{(1,0)}(\boldsymbol{\mu}, \boldsymbol{\phi}) + \boldsymbol{\phi} T(\mathbf{y}) + \Psi(\boldsymbol{\mu}, \boldsymbol{\phi}) \right\}, \quad (3.2)$$

com :

- as funções  $A(\cdot)$ ,  $T(\cdot)$ ,  $\Psi(\cdot, \cdot)$  conhecidas;
- média  $E(Y) = \mu$  e variância  $Var(Y) = (\Psi^{(2,0)})^{-1}$ ,
- $E\{T(Y)\} = -\Psi^{(0,1)}$ ,  $Var\{T(Y)\} = -\Psi^{(0,2)}$  e  $Cov\{Y, T(Y)\} = 0$ .

A notação  $\Psi^{(r,s)}(\boldsymbol{\mu}, \boldsymbol{\phi}) = \frac{\partial^{(r,s)} \Psi(\boldsymbol{\mu}, \boldsymbol{\phi})}{\partial \mu^r \partial \phi^s}$ , na qual, para  $r = 1$  e  $s = 0$ , tem-se  $\Psi^{(1,0)}$  que é a primeira derivada da função  $\Psi(\boldsymbol{\mu}, \boldsymbol{\phi})$  com a relação a  $\mu$ ; de modo análogo, para  $r = 0$  e  $s = 1$ , tem-se que  $\Psi^{(0,1)}$  é a primeira derivada da função  $\Psi(\boldsymbol{\mu}, \boldsymbol{\phi})$  com relação a  $\phi$ .

Importante salientar que a função  $\Psi(\boldsymbol{\mu}, \boldsymbol{\phi})$  da expressão (5.1) é equivalente a  $\rho(\boldsymbol{\mu}, \boldsymbol{\phi})$  da expressão (3.1). E ainda, por comparação, tem-se que

$$\boldsymbol{\theta} = \Psi^{(1,0)}(\boldsymbol{\mu}, \boldsymbol{\phi}) \quad \text{e} \quad \rho(\boldsymbol{\theta}, \boldsymbol{\phi}) = -\Psi(\boldsymbol{\mu}, \boldsymbol{\phi}) + \boldsymbol{\mu} \Psi^{(1,0)}(\boldsymbol{\mu}, \boldsymbol{\phi})$$

e pode-se demonstrar que  $E\left(\frac{\partial^2 \log \pi}{\partial \boldsymbol{\theta} \partial \boldsymbol{\mu}}\right) = 0$ , isto é, que  $\mu$  e  $\phi$  são parâmetros ortogonais conforme Cox e Reid (1987).

### 3.3 Distribuições pertencentes à família superdispersada

Na seqüência, algumas distribuições de probabilidades foram verificadas como pertencentes aos MLGS.

**Distribuição inversa Gaussiana generalizada** Introduzida por Good (1953) e mais tarde estudada com interesse por Wise (1971, 1975) e Jørgensen (1983), a distribuição IGG abaixo é dada por Johnson, Kotz e Balakrishnam (1994, vol 1), que utilizam três parâmetros para definir a função densidade como

$$\pi_y = (y; \alpha, \delta, r) = \frac{\left(\frac{\alpha}{\delta}\right)^{\frac{r}{2}}}{2K_r\left(\sqrt{\alpha\delta}\right)} y^{r-1} \exp\left\{-\frac{1}{2}(\delta y^{-1} + \alpha y)\right\}, \quad y, \alpha, \delta > 0, \quad (3.3)$$

onde  $K_r(\cdot)$  é uma função de Bessel modificada de terceira ordem.

É utilizada quando a variável resposta tem assimetria considerável e intervalo de variação positivo. Ela é uma concorrente das distribuições Gama e passeio aleatório.

A expressão acima pode ser escrita como

$$\pi_y = (y; \alpha, \delta, r) = y^{r-1} \exp \left\{ -\frac{1}{2} \delta y^{-1} - \frac{1}{2} \alpha y + \log \left[ \frac{\left(\frac{\alpha}{\delta}\right)^{\frac{r}{2}}}{2K_r \left(\sqrt{\alpha\delta}\right)} \right] \right\}.$$

Fazendo  $r$  conhecido e considerando que

$$\begin{aligned} \theta &= -\frac{\alpha}{2}, & T(y) &= y^{-1} \\ \tau &= -\frac{\delta}{2} & \text{e } A(y) &= y^{r-1}, \end{aligned}$$

lembrando que os MLGS tem origem na forma da família de Gelfand e Dalal (1990) dada por

$$f(y; \mu, \tau) = A(y) \exp \{ \theta y + \tau T(y) - \rho(\theta, \tau) \},$$

e que a partir de  $\rho(\theta, \tau)$  pode-se determinar a  $\Psi(\mu, \phi, )$ , então

$$\begin{aligned} \rho(\theta, \tau) &= -\log \left(\frac{\alpha}{\delta}\right)^{\frac{r}{2}} + \log 2 + \log K_r \left(\sqrt{\alpha\delta}\right) \text{ e lembrando que } \alpha\delta = 4\theta\tau \\ \rho(\theta, \tau) &= -\frac{r}{2} \log \left(\frac{\theta}{\tau}\right) + \log 2 + \log K_r \left(2\sqrt{\theta\tau}\right), \end{aligned}$$

deriva-se  $\rho(\theta, \tau)$  com relação a  $\theta$  e tem-se  $\Psi$

$$\begin{aligned} \frac{\partial \rho(\theta, \tau)}{\partial \theta} &= -\frac{r}{2\theta} + \frac{K'_r \left(2\sqrt{\theta\tau}\right)}{K_r \left(2\sqrt{\theta\tau}\right)} \cdot \frac{1}{\sqrt{\theta\tau}} \\ &= -\frac{r}{2\theta} + M(\theta, \tau) \\ &= \mu. \end{aligned}$$

Gelfand e Dalal (1990) apresentam que  $\theta = \Psi^{(1,0)}(\mu, r)$ , ou seja,

$$\begin{aligned} \frac{\partial \Psi(\mu, \tau)}{\partial \mu} &= \theta \Rightarrow \Psi(\mu, \tau) = \int \theta du \\ &= \Psi(\mu, \tau) = \int \Psi^{(1,0)}(\mu, r) du, \end{aligned}$$

e que só pode ser calculada numericamente, ou seja, não tem solução fechada.

Desse modo, fica demonstrado que a distribuição inversa gaussiana generalizada pode ser escrita sob a forma dos modelos lineares generalizados superdispersados. E ainda, quando os parâmetros  $\alpha, \delta$  e  $r$  assumem alguns valores na IGG, essa distribuição torna-se caso especial,

como:

- i) Distribuição inversa gaussiana , se  $r = -\frac{1}{2}$ .
- ii) Distribuição recíproca da inversa gaussiana, se  $r = \frac{1}{2}$ .
- iii) Distribuição gama, se  $\delta = 0$  e  $r > 0$ .
- iv) Distribuição recíproca gama, se  $\alpha = 0$  e  $r < 0$ .
- v) Distribuição hiperbólica, se  $r = 0$ .

Desse modo, pode-se fazer a extensão de que a distribuição inversa gaussiana generalizada é uma subfamília dos modelos lineares generalizados superdispersados.

### Passeio aleatório

O embrião dessa distribuição começou em 1827, com o botânico inglês Robert Brown, ao observar o movimento errático de pequenas partículas no interior de grãos de pólen.

O modelo do passeio aleatório tem sido associado ao desenvolvimento de inúmeros modelos financeiros, como por exemplo os preços no mercado futuro brasileiro, e é muito comum também adotá-lo para valores imobiliários.

A função densidade de probabilidade do modelo passeio aleatório é da seguinte forma:

$$\pi(\mathbf{y}; \boldsymbol{\theta}, \boldsymbol{\delta}) = \left( \frac{\delta}{2\pi y} \right)^{\frac{1}{2}} \exp \left\{ -\frac{\delta y}{2} + \frac{\delta}{\theta} - \frac{\delta}{2\theta^2 y} \right\}, \quad (3.4)$$

com  $y, \theta, \delta > 0$ , com a média e a variância

$$E(y) = \mu = \frac{1}{\theta} + \frac{1}{\delta} \text{ e } Var(y) = \frac{1}{(\theta\delta)} + \frac{2}{\delta^2},$$

respectivamente.

Essa distribuição é considerada especial, pois pode ser escrita sob a forma dos MLGS, no entanto não pertence aos MLG e modelos lineares duplos.

### Família exponencial dupla

Para facilitar o entendimento, a família exponencial uniparamétrica de Efron (1986) pode ser escrita como

$$g(y; \theta, n) = \exp \{ n(y\theta - \chi^2(\theta))h(y; n) \}, \quad (3.5)$$

considere que  $\theta = E(Y) = \mu$ ,  $n$  é parâmetro de dispersão e  $\Psi(\mu) = \chi^2(\theta)$  e ainda  $h(y; n)$  e  $n$  conhecidos. A partir disso, definiu-se a família exponencial dupla pela função densidade

$$\pi(y; \mu, \theta, n) = c(\mu, \theta, n) \theta^{\frac{1}{2}} \{g(y; \mu, n)\}^{\theta} \{g(y; n)\}^{1-\theta} [dh(y; n)], \quad (3.6)$$

com parâmetros  $\mu, \theta, n$  e a constante  $c(\mu, \theta, n)$  aproximadamente igual a 1 e média e variância aproximadamente  $\mu$  e  $V(\mu)/n\theta$ .

Essa família de distribuição desfruta das propriedades da família exponencial tanto para o parâmetro da média como para o parâmetro de dispersão.

Seguindo a notação usada no artigo Gelfand e Dalal (1990), denota-se por  $\theta = \rho$  e substituindo na expressão (3.6), tem-se

$$\pi(y; \theta, \rho, n) = c(\theta, \rho, n) \rho^{\frac{1}{2}} \{g(y; \mu, n)\}^{\rho} \{g(y; n)\}^{1-\rho} [dh(y; n)]$$

e agora, substituindo a função  $g(y; \mu, n)$  dada por (3.5) na expressão acima, passa a ser escrita como

$$\pi(y; \theta, \rho, n) = c(\theta, \rho, n) \rho^{\frac{1}{2}} \{\exp\{n(y\theta - \chi^2(\theta))\} h(y; n)\}^{\rho} \quad (3.7)$$

$$\begin{aligned} & \{\exp\{n(y\theta - \chi^2(\theta))\} h(y; n)\}^{1-\rho} [dh(y; n)] \\ &= c(\theta, \rho, n) \rho^{\frac{1}{2}} \exp\{n\rho(y\theta - \chi^2(\theta))\} \\ & \quad + n(1 - \rho)(y\theta - \chi^2(\theta)) h(y; n). \end{aligned} \quad (3.8)$$

Fazendo

$$T(y) = y\theta - \chi^2(\theta),$$

$$\begin{aligned} n(1 - \rho) &= \phi \\ \rho &= 1 - \frac{\phi}{n}, \end{aligned}$$

temos

$$\begin{aligned} \alpha &= n\rho\theta \\ \alpha &= n\left(1 - \frac{\phi}{n}\right)\theta \\ \theta &= \frac{\alpha}{n - \phi}. \end{aligned}$$

Agora basta substituir essas quantidades na expressão (3.7), que passa a ser

$$\begin{aligned} \pi(y; \theta, \rho, n) &= c(\theta, \rho, n) \rho^{\frac{1}{2}} \exp\{n\rho y\theta - n\rho\chi^2(\theta) \\ & \quad + n(1 - \rho)T(y) h(y; n)\} \\ &= c(\theta, \rho, n) \rho^{\frac{1}{2}} \exp\{\alpha y + \phi T(y) - \rho(\alpha, \phi) h(y; n)\}, \end{aligned}$$

e assim fica demonstrado que a família exponencial dupla apresentada por Efron (1986) pode ser escrita sob a forma da família exponencial de dois parâmetros apresentada por Gelfand e Dalal

(1990) na expressão (3.1), que derivou os modelos lineares generalizados superdispersados.

Desse modo, conclui-se que os MLGS englobam a família exponencial dupla apresentada por Efron (1986) e que, portanto, todas as distribuições denominadas família binomial, família gama e Poisson dupla podem ser escritas sob a forma dos MLGS.

Na seqüência, a família Poisson dupla foi escrita, com certo rigor nos detalhes, sob a forma dos MLGS, uma vez que, foi utilizada para ajustar dados de aplicação nos Capítulos 5 e 7.

### Família Poisson dupla

Essa família é baseada na distribuição de Poisson, sendo que a variável resposta é inteira não negativa. Pode ser utilizada em controle de qualidade, visto que é comum haver interesse em determinar o número de falhas de equipamentos em função do tempo.

Segundo Efron (1986), assume-se que a variável resposta  $Y$  foi reescalada, com função densidade de probabilidade dada por

$$\pi(\mathbf{y}; \boldsymbol{\mu}, \phi) = \left( \phi^{\frac{1}{2}} e^{-\phi\mu} \right) \left( \frac{e^{-y} y^y}{y!} \right) \left( \frac{e\mu}{y} \right)^{\phi y}, \quad (3.9)$$

com  $y = 0, 1, 2, 3, \dots$  e média  $\mu$  e  $Var(y) = \frac{\mu}{\phi}$ .

Considerando que  $A(y) = \left( \frac{e^{-y} y^y}{y!} \right)$  e aplicando  $\exp$  da log na função densidade de probabilidade tem-se que:

$$\begin{aligned} \pi(y; \mu, \phi) &= A(y) \exp \left\{ \log \left[ \phi^{\frac{1}{2}} e^{-\phi\mu} \left( \frac{e\mu}{y} \right)^{\phi y} \right] \right\} \\ &= A(y) \exp \left\{ -\phi\mu + \frac{1}{2} \log \phi + \phi y \log \mu - \phi y \log y \right\} \\ &= A(y) \exp \left\{ (\phi + \phi \log \mu) y - \phi y \log y - \phi\mu + \frac{1}{2} \log \phi \right\}. \end{aligned} \quad (3.10)$$

Denotando por

$$\begin{aligned} \delta &= \phi + \phi \log \mu, \text{ temos que} \\ \mu &= \exp \left( \frac{\delta}{\phi} - 1 \right) \end{aligned} \quad (3.11)$$

Substituindo essas igualdades na expressão (3.9), tem-se

$$\pi(y; \mu, \phi) = A(y) \exp \left\{ y(\delta) - \phi y \log y + \frac{1}{2} \log(\phi) - \phi \exp \left( \frac{\delta}{\phi} - 1 \right) \right\},$$

assim, a equação acima está na forma da expressão (3.1), onde

$$\theta = \delta \quad \text{e} \quad T(y) = -y \log y,$$

$$\rho(\theta, \tau) = \rho(\delta, \phi) = - \left\{ \frac{1}{2} \log(\phi) - \phi \exp\left(\frac{\delta}{\phi} - 1\right) \right\}.$$

Considerando que

$$\theta = \delta = \Psi^{(1,0)}(\mu, \phi) = \phi + \phi \log \mu, \text{ e substituindo em}$$

$$\begin{aligned} \rho(\delta, \phi) &= -\Psi(\mu, \phi) + \mu\psi^{(1,0)}(\mu, \phi), \text{ tem-se que} \\ \Psi(\mu, \phi) &= +\mu\psi^{(1,0)}(\mu, \phi) - \left\{ -\frac{1}{2} \log(\phi) + \phi \exp\left(\frac{\delta}{\phi} - 1\right) \right\} \\ \Psi(\mu, \phi) &= \mu(\phi + \phi \log \mu) + \frac{1}{2} \log(\phi) - \phi \exp\left(\frac{\delta}{\phi} - 1\right) \\ \Psi(\mu, \phi) &= \mu\phi \log \mu + \frac{1}{2} \log(\phi). \end{aligned}$$

Dessa forma, determinaram-se as componentes da família Poisson dupla sob a forma MLGS.

No contexto dos modelos lineares generalizados, cabe ressaltar que, a distribuição binomial negativa é bastante utilizada para dados que apresentam superdispersão, no entanto, não pertence a classe apresentada por Dey et al. (1997), em outras palavras, esta distribuição não pode ser escrita sob a forma dos MLGS.

# CAPÍTULO 4

## CÁLCULO DO VIÉS DE $O(n^{-1})$ PARA OS MLG E OS MODELOS NÃO LINEARES

Nos modelos de regressão, houve um avanço com a teoria dos modelos lineares generalizados, visto que a variável resposta pode assumir qualquer distribuição de probabilidade pertencente à família exponencial e, além disso, pode apresentar heterocedasticidade, sendo que esta passa a ser uma função conhecida da média. Portanto, no contexto dos MLG não são mais necessárias as suposições de normalidade e variância constante.

A popularidade dos MLG é devida à sua utilização em todas as áreas da ciência, como no caso da engenharia de produção, visto ser comum encontrar dados dessa natureza, e também pela disponibilidade nos aplicativos estatísticos para estimação dos parâmetros.

Este capítulo contém duas seções que apresentam as expressões para os cálculos dos vieses de  $O(n^{-1})$ . A primeira para os modelos lineares generalizados e a segunda para os modelos não-lineares da família exponencial. O principal objetivo desse capítulo é apresentar uma aplicação para cada modelo na área de engenharia de produção. Dessa forma, pretende-se disseminar e consolidar o uso desses modelos visando obter estimativas corrigidas e, assim, evidenciar a melhoria principalmente quando estas se referem a custos, produção, etc.

### 4.1 Modelos lineares generalizados

Nelder e Wedderburn (1972) contribuíram com a teoria dos MLG desenvolvendo uma classe de modelos baseados na família exponencial.

Os modelos lineares generalizados ficam caracterizados pelo trinômio: componente aleatória da variável resposta, componente sistemática dos parâmetros e função de ligação, que vincula a média ao preditor linear, isto é,

i) componente aleatória,  $Y$ , pertencente à família exponencial;

ii) componente sistemática  $\eta = g(\boldsymbol{\mu}) = X\boldsymbol{\beta}$ ,

sendo que

- $\eta$  é o preditor linear,
- $g(\cdot)$  a função de ligação conhecida, monótona e diferenciável, geralmente sendo não-linear,
- $X$  a matriz modelo e
- $\boldsymbol{\beta}$  o vetor de parâmetros.

Admite-se que o parâmetro de dispersão  $\phi$  seja o mesmo para todas as observações, ou seja, constante. Caso não seja conhecido, é necessário estimá-lo.

Outros detalhes desses modelos, bem como identificadores da família exponencial para as distribuições mais utilizadas (normal, Poisson, binomial, gama e inversa gaussiana), com suas respectivas funções de ligação canônica, podem ser vistas em Nelder e Wedderburn (1972), Cordeiro (1986), Dobson (1990), Demétrio (1993), McCulloch (2000) e Paula (2001), entre outros.

Há alguns modelos considerados como casos especiais de modelos lineares generalizados, que são apresentados por Cordeiro (1986).

#### 4.1.1 Estimadores de $O(n^{-1})$

O método de estimação para os parâmetros  $\boldsymbol{\beta}$  e  $\phi$  é o escore de Fisher, descrito no Capítulo 2. Existem vários aplicativos disponíveis para estimação dos parâmetros dos MLG, tais como GLIM, SAS, Statistica e S-Plus, baseados no método de máxima verossimilhança, sendo que o GLIM apresentado por Baker e Nelder (1978) foi o primeiro a ser desenvolvido.

Os resultados assintóticos podem ser encontrados em McCullagh e Nelder (1983), Cordeiro (1983, 1987) e Fahrmeir e Kaufmann (1985). Demonstração dada por Cordeiro (1986, *pág.39*), que  $\hat{\boldsymbol{\beta}}$  é consistente para  $\boldsymbol{\beta}$  e expandindo  $U(\hat{\boldsymbol{\beta}})$  em série de Taylor ao redor do parâmetro verdadeiro  $\boldsymbol{\beta}$ , isto é,

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} + O(n^{-1}) \text{ e } Cov(\hat{\boldsymbol{\beta}}) = E\left[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T\right] = O(n^{-1}),$$

estas expressões refletem que para pequenas amostras,  $\hat{\boldsymbol{\beta}}$  é tendencioso e têm poucas propriedades que são satisfeitas para todos os tamanhos de amostras, como suficiência e invariância.

Argumentos baseados no teorema do limite central adaptados a função escore  $U(\boldsymbol{\beta})$ , e a lei fraca dos grandes números, aplicados à matriz de informação de Fisher  $K(\boldsymbol{\beta})$  e sob as



condições de regularidades apresentadas em detalhes em Wei (1998), demonstra-se que  $\hat{\beta}$  é assintoticamente normal, com média  $\beta$  e matriz de covariância  $Cov(\hat{\beta})$ .

Intervalos aproximados de confiança e diversos testes são construídos através de  $\hat{\beta}$  ou de estatísticas de  $\hat{\beta}$ , tais como a estatística da razão de máxima verossimilhanças, a estatística escore ou a estatística de Wald. A estatística da razão de máxima verossimilhanças é mais utilizada, uma vez que, a mesma é invariante com reparametrizações, o que não ocorre com as outras duas. Estas estatísticas são apresentadas por Demétrio (1993) entre outros.

### 4.1.2 Avaliação do MLG

Uma estatística bastante utilizada é a deviance que serve para avaliar a qualidade do ajuste de um MLG. Para obtê-la basta fazer o logaritmo da verossimilhança. Segue a seguinte notação para a log-verossimilhança dos MLG:

$$\ell(\mu; y) = \sum_{i=1}^n \ell(\mu_i; y_i), \text{ onde } \mu_i = g^{-1}(\eta_i) \text{ e } \eta_i = X\beta.$$

Considere como modelo saturado, aquele em que o número de parâmetros  $p$  é igual ao número de observações  $n$ , isto é,  $p = n$ . Portanto, para este modelo a  $\ell(\mu; y)$  é estimada por:

$$\ell(y; y) = \sum_{i=1}^n \ell(y_i; y_i),$$

ou seja, a estimativa de máxima verossimilhança de  $\mu_i$  fica no modelo saturado dada por  $\hat{\mu}_i = y_i$ . Quando  $p < n$ , a estimativa de  $\ell(\mu; y)$ , será denotada por  $\ell(\hat{\mu}; y)$ . Neste caso, as estimativas de máxima verossimilhança de  $\hat{\mu}_i = g^{-1}(\hat{\eta}_i)$ , sendo  $\hat{\eta} = X\hat{\beta}$ .

A deviance é considerada como uma distância entre a log-verossimilhança do modelo saturado e o modelo sob investigação avaliado na estimativa de máxima verossimilhança de  $\beta$ , é dada por:

$$\begin{aligned} D^*(y; \hat{\mu}) &= \phi D(y; \hat{\mu}) = 2 \{ \ell(y; y) - \ell(\hat{\mu}; y) \} \\ &= 2\phi \sum_{i=1}^n \{ v(y_i) - v(\hat{\mu}_i) + (\hat{\mu}_i - y_i) q(\hat{\mu}_i) \}, \end{aligned} \quad (4.1)$$

sendo  $v(\mu) = \mu q(\mu) - b[q(\mu)]$  e  $q(\mu) = \theta = \int V^{-1} d\mu$ , definidos em Wei (1998).

Para interpretar o valor obtido da deviance, se o valor for pequeno, isto indica que, para um número menor de parâmetros, obtém-se um ajuste tão bom quanto o ajuste para o modelo saturado.

Quando o parâmetro  $\phi$  for estimado, utiliza-se para avaliar MLG, a deviance scaled, que é dado pelo quociente entre a deviance e o parâmetro de dispersão. Uma forma de avaliar o

modelo é usar a deviance scaled dividida pelos graus de liberdade e caso este estiver próximo de um, diz-se que, o modelo está bem ajustado, conforme afirma Demétrio (1993).

### 4.1.3 Cálculo do viés de $O(n^{-1})$ para os MLG

Cordeiro e McCullagh(1991) apresentam uma fórmula para o viés de primeira ordem para os modelos lineares generalizados por meio de uma regressão linear ponderada. Usando a fórmula devida a Cox e Snell (1968) dada pela expressão (2.1) e após algumas manipulações algébricas, foi obtida a seguinte expressão na forma matricial para o viés do estimador do parâmetro  $\beta_r$ ,

$$B(\hat{\beta}) = (X^T W X)^{-1} X^T W \xi, \quad (4.2)$$

onde

$$\xi = -\frac{1}{2\phi} W^{-1} Z_d F \mathbf{1} \quad (4.3)$$

sendo que  $r = 1, \dots, p$  parâmetros, cujas componentes são obtidas como os coeficientes de regressão linear, onde  $\xi$  é o vetor resposta e  $\mathbf{1}$  é um vetor  $n \times 1$  de uns e  $W$ ,  $Z_d$  e  $F$  são matrizes diagonais definidas da seguinte forma:

$$W = \left( \frac{\partial \mu}{\partial \eta} \right)^2, \quad F = V^{-1} \frac{d\mu}{d\eta} \frac{d^2 \mu}{d\eta^2} \text{ e } V = \frac{\partial^2 b(\theta)}{\partial \theta^2},$$

onde  $\theta$  é o parâmetro canônico da família exponencial.

$$\text{E } Z_d = \text{diag} \{z_{11}, \dots, z_{nn}\} = \{z_{ij}\} = X (X^T W X)^{-1} X^T, \quad .$$

A matriz  $Z_d$  é a matriz de covariância assintótica dos estimadores dos preditores lineares do modelo.

### 4.1.4 Significado do viés

O viés nada mais é do que a variação na resposta média, ou seja, mede o impacto da variação da média da resposta. Escrevendo isso matematicamente, tem-se

$$\begin{aligned} \nabla \hat{\mu} &= \hat{\beta} \\ \nabla \tilde{\mu}_{\text{corrigido}} &= \tilde{\beta}_{\text{corrigido}} \\ \nabla \tilde{\mu}_{\text{corrigido}} - \nabla \hat{\mu} &= \tilde{\beta}_{\text{corrigido}} - \hat{\beta} \\ \frac{\nabla \tilde{\mu}_{\text{corrigido}} - \nabla \hat{\mu}}{\nabla \hat{\mu}} &= \frac{\text{viés}}{\hat{\beta}}. \end{aligned}$$

O termo antes da igualdade significa a elasticidade marginal da média da resposta e o termo depois da igualdade significa a variação em percentual. Cabe ressaltar que essa interpretação

só é válida para os casos dos modelos lineares.

O ajuste do modelo e as análises contida neste capítulo foi realizada no aplicativo SAS, mais especificamente a procedure GENMOD.

## 4.2 Aplicação 1: vendas de telhas

O objetivo desta seção é somente ilustrar a aplicação dos modelos lineares generalizados corrigidos na engenharia de produção e, em particular, implementar as expressões dos vieses, evidenciando que o cálculo dos vieses nada mais é do que uma regressão ponderada, portanto de fácil implementação.

Para determinar as estimativas de máxima verossimilhança dos parâmetros do modelo adotado, utilizou-se o aplicativo SAS, mais especificamente a Proc GENMOD (modelos lineares generalizados) e a Proc IML (operações com matrizes).

O banco de dados com 26 observações foi obtido na home page e autorizado pelo Prof. *Ph.D.* Julio M. Singer, no endereço [www.ime.usp.br/~julio](http://www.ime.usp.br/~julio).

### Modelo adotado

Pela natureza dos dados, ou seja, por meio de uma análise preliminar, observou-se que a variável resposta (total de vendas de telhas), apresenta assimetria e intervalo de variação positivo e que, portanto, adequa-se o uso da distribuição de probabilidade inversa gaussiana, da seguinte forma:

A variável resposta como :

$\mathbf{Y}$  → total = total de vendas de telhas.

A matriz  $X$  é composta das seguintes covariáveis:

$\mathbf{X}_1$  → gasto = gasto com propaganda,

$\mathbf{X}_2$  → clientes = número de clientes,

$\mathbf{X}_3$  → concor = número de concorrentes e

$\mathbf{X}_4$  → potencial = clientes em potencial.

Componente sistemática,

$$\boldsymbol{\eta} = \log(\boldsymbol{\mu}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$$

Como o objetivo é simplesmente ilustrar o cálculo do viés, não foram efetuadas análises quanto ao ajuste do modelo adotado com outras funções de ligações.

Na Tabela 4.1 estão apresentadas as estimativas do modelo com suas respectivas correções. Cabe ressaltar que as estatísticas de Wald são para as estimativas de  $O(n^{-1})$  e que todos os níveis descritivos para as covariáveis foram significativos, exceto para a covariável potencial.

Tabela 4.1: Estimativas de  $O(n^{-1})$  e  $O(n^{-2})$  para o modelo inversa gaussiana.

Estimativas	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$
Não corrigidas- $O(n^{-1})$	5.3873	-0.0874	0.0328	-0.1578	-0.0126
Erros-padrão	0.2487	0.0235	0.0032	0.0177	0.0098
Vieses	-0.1362	-0.0009	0.0040	0.0278	0.0012
Corrigidas- $O(n^{-2})$	5.5236	-0.0883	0.0368	-0.1848	-0.0113
IC- 95% - Wald	(4.89; 5.87)	(-0.13; -0.04)	(0.02; 0.03)	(-0.19; -0.12)	(-0.03; 0.01)

Os valores dos vieses dos estimadores  $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$  quando comparadas ao seus erros-padrão representam 11%, 13,95% e 11,39% respectivamente. Portanto, houve uma correção considerável nas estimativas do modelo inversa gaussiana, pois representam em torno de 11% da variação média dos totais de vendas de telhas.

O valor obtido da deviance é de 0.0065 e a deviance scaled é de 26, que, dividido pelos graus de liberdade, é próximo de 1, o que, portanto, indica um bom ajuste.

No Apêndice B está disponível o programa implementado no SAS para obtenção das estimativas, bem como os cálculos dos vieses.

### 4.3 Modelos não-lineares da família exponencial (MNLFE)

Jørgensen (1983) e McCullagh (1983) discutiram alguns modelos de regressão não-lineares, tomando-os como caso particular dos modelos não-lineares da família exponencial. Nos últimos anos, muitos pesquisadores têm direcionado estudos para o melhoramento da inferência nesses modelos. Entre eles, Cordeiro e Paula (1989), Pazman (1991), Wei e Shi (1994) e outros.

O modelo não-linear da família exponencial segue a mesma estrutura dos MLG e só difere no que diz respeito ao preditor  $\eta$ , que é não-linear.

As condições de regularidade para esses modelos podem ser vistas em Wei (1998).

#### 4.3.1 Estimadores de $O(n^{-1})$

Para calcular o estimador  $\hat{\beta}$ , pode-se utilizar os métodos iterativos de Gauss-Newton, Gauss-Newton modificado ou Newton Raphson, que são apresentados por Wei (1998), inclusive com alguns exemplos. Estes iniciam pela expansão em série de Taylor de primeira ordem para a

equação de verossimilhança  $U(\boldsymbol{\beta}) = 0$ . A convergência é rápida, no entanto, depende fortemente dos valores iniciais.

O processo iterativo para obter as estimativas desse modelo é equivalente aos dos MLG, com as seguintes adaptações:

- A matriz modelo considerada passa a ser  $\tilde{X} = \tilde{X}(\boldsymbol{\beta}) = \frac{\partial \eta}{\partial \boldsymbol{\beta}}$ , avaliada em cada iteração.

O vetor resposta ajustado como

$$y^* = \eta - \tau + \frac{\partial \eta}{\partial \mu} (y - \mu),$$

observa-se que é o mesmo que os MLG, exceto por  $\tau$ . Cordeiro e Paula (1989), com o objetivo de usar o mesmo algoritmo dos MLG, consideram que  $\tau = f(X; \boldsymbol{\beta}) - \tilde{X}\boldsymbol{\beta}$ , é uma medida local de não-linearidade da componente sistemática, denominada offset.

Com essas adaptações, implementou-se uma macro no SAS, em que em cada iteração são obtidas a matriz  $\tilde{X}$  e o vetor  $\tau$ . Para calcular a estimativa de  $\boldsymbol{\beta}$  utilizou-se a proc GENMOD, usando o recurso da offset.

### 4.3.2 Cálculo do viés do $O(n^{-1})$ para os MNLFE

Paula (1991) apresenta uma fórmula para o cálculo do viés de  $O(n^{-1})$ , em que, após determinar alguns cumulantes e usando a fórmula de Cox e Snell (1968) com algumas operações algébricas, obteve a seguinte expressão para o viés do estimador do parâmetro  $\boldsymbol{\beta}_r$ :

$$B(\hat{\boldsymbol{\beta}}_r) = (\tilde{X}W\tilde{X})^{-1} \tilde{X}^T W (\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2), \quad (4.4)$$

onde

$$\boldsymbol{\xi}_1 = -(2\phi)^{-1} Z_d W^{-1} F \mathbf{1} \quad \text{e} \quad \boldsymbol{\xi}_2 = -(2\phi)^{-1} D \mathbf{1},$$

tendo  $\mathbf{1}$  como um vetor  $n \times 1$  de uns,  $Z_d = \text{diag}(Z)$  de dimensão  $n \times n$ , onde  $Z = \tilde{X}^T (\tilde{X}^T W \tilde{X})^{-1} \tilde{X}$  e, ainda,  $D = \text{diag}\{d_1, \dots, d_n\}$ , sendo que  $d_i = \text{tr} \left( \tilde{X}_i (\tilde{X}^T W \tilde{X})^{-1} \right)$  e  $\tilde{X}_i^T$  é uma matriz  $p \times p$  das segundas derivadas com elementos  $\partial^2 \eta_i / \partial \beta_r \partial \beta_s = (rs)_i$  e  $F = \text{diag}\{f_1, \dots, f_n\}$  sendo  $f_i = V^{-1} \frac{d\mu_i}{d\eta_i} \frac{d^2 \mu_i}{d\eta_i^2}$ .

Desse modo, o viés de  $O(n^{-1})$  para os modelos não-lineares da família exponencial é a regressão de mínimos quadrados ponderados de  $\boldsymbol{\xi} = \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2$  sobre as colunas de  $\tilde{X}$  com a matriz peso  $W$ .

## 4.4 Aplicação 2: demanda de energia elétrica

Muitos trabalhos têm sido desenvolvidos na área de energia elétrica, principalmente pela preocupação atual de expandir e modernizar todos os segmentos ligados à energia elétrica, diante de uma situação iminente de racionamento, segundo Schimidt *et al.* (2002); além disso, vinte milhões de brasileiros ainda não têm acesso à energia elétrica.

A demanda residencial por energia elétrica pode ser descrita por um conjunto de variáveis; no entanto, algumas parecem ter um peso maior, como, por exemplo, a restrição orçamentária e o preço.

A modelagem da demanda de energia elétrica tem por finalidade evidenciar o uso da correção na área de engenharia de produção, visto que é um problema de maximizar a utilidade, sujeita a restrições e que, portanto, uma correção nas estimativas é sempre bem-vinda, visto que envolve consumo, o que equivale dizer produção de energia e conseqüentemente custos.

O banco de dados de demanda de energia elétrica foi disponibilizado por Cordeiro (1988) e coletado no primeiro trimestre de 1961 até o quarto trimestre de 1983, com uma amostra de 92 observações.

### Modelo adotado

Como a variável resposta é sempre positiva, o modelo adotado foi a distribuição gama com função de ligação recíproca.

A variável resposta :

$\mathbf{Y}$  → Elar = demanda de eletricidade agregada per capita para o setor residencial.

A matriz  $X$  é composta das seguintes covariáveis:

$\mathbf{X}_1$  → Per = preço médio da eletricidade para o setor residencial,

$\mathbf{X}_2$  → Pge = preço do gás natural para o setor residencial,

$\mathbf{X}_3$  → Rea = renda per capita,

Componente sistemática,

$$\boldsymbol{\eta} = \boldsymbol{\mu}^{-1} = X_1^{\beta_1} * X_2^{\beta_2} * X_3^{\beta_3}.$$

As estimativas foram determinadas usando-se o método de Newton Raphson na procedure GENMOD, no aplicativo SAS.

Observa-se pela Tabela 4.2, que as estimativas de  $O(n^{-1})$  foram todas significativas e a deviance scaled igual a 92,3429, evidenciando um bom ajuste do modelo.

Tabela 4.2: Estimativas de  $O(n^{-1})$  e  $O(n^{-2})$  para a distribuição gama com função ligação recíproca.

Estimativas	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\phi}$
Não corrigidas- $O(n^{-1})$	2.3466	-1.2349	0.4703	44.7079
Erros-padrão	0.1521	0.0683	0.0541	6.5674
Vieses	0.036	0.023	0.010	
Corrigidas - $O(n^{-2})$	2.3106	1.2579	0.4603	
IC - 95% - Wald	(2.05; 2.64)	(-1.37; -1.11)	(0.36; 0.57)	

Os vieses dos estimadores  $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$  correspondem a 23% , 33% e 18%, respectivamente, dos erros-padrão, evidenciando valores importantes para vieses e, assim, indicando a necessidade de obter estimativas de  $O(n^{-2})$ .

No Apêndice C está disponível a macro no SAS dos cálculos das estimativas e os seus respectivos vieses para esse modelo.

# CAPÍTULO 5

## MODELOS LINEARES GENERALIZADOS SUPERDISPERSADOS

Encontra-se, rotineiramente, nos modelos de regressão a presença de superdispersão, ou seja, a ocorrência do fenômeno em que os dados apresentam uma variabilidade maior do que a prevista pelo modelo adotado. Isso é muito comum na área de engenharia de produção, visto que uma das preocupações em programas de melhoria da qualidade é encontrar respostas às causas/fatores que influenciam a variabilidade no processo de produção. Visto ser comum também, nos ensaios de controle biológico, agrônômicos, etc. Para Jørgensen (1997), a principal característica está na distribuição dos erros.

Há várias propostas para acomodar a superdispersão, como podem ser vistas em Efron (1986) e Smyth (1989), que consideram uma família exponencial dupla, e Lindsay (1986), que propõe utilizar os modelos lineares generalizados mistos, ou seja, adicionar estruturas de efeitos aleatórios no preditor linear. O mecanismo para Hinde e Demétrio (1998) é incluir mais um parâmetro na função de variância, principalmente para os dados de contagens e binários. Já Smyth e Verbyla (1999) consideram um modelo de regressão adicional para o parâmetro de dispersão no contexto de modelos lineares generalizados, e Dey *et al.* (1997) definem a classe dos MLGS, visto que a partir dela obtiveram-se os modelos não-lineares generalizados superdispersados e os seus respectivos vieses de  $O(n^{-1})$ .

### 5.1 Modelo

Pela dificuldade de modelar a dispersão, Dey *et al.* (1997) estenderam a metodologia dos MLG numa classe maior, ou seja, considerando um modelo de regressão adicional para o parâmetro de dispersão e incorporando-o na função de variância.

Seguindo a mesma estrutura MLG, esses modelos ficam caracterizados por uma componente



aleatória, no entanto, com duas componentes sistemáticas e duas funções de ligação.

Formalizando, têm-se:

- Componente aleatória é a que descreve o vetor resposta  $Y$ , cuja função densidade de probabilidade é apresentada por Dey et al. (1997), que teve origem em Gelfand e Dalal (1990), e esse resultado foi mostrado detalhadamente no Capítulo 3.

Portanto, a família de distribuições que é utilizada nos MLGS é a que segue:

$$\pi(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\phi}) = A(\mathbf{y}) \exp \{ (\mathbf{y} - \boldsymbol{\mu}) \Psi^{(1,0)}(\boldsymbol{\mu}, \boldsymbol{\phi}) + \boldsymbol{\phi} T(\mathbf{y}) + \Psi(\boldsymbol{\mu}, \boldsymbol{\phi}) \}, \quad (5.1)$$

sendo as funções  $A(\cdot)$ ,  $T(\cdot)$ ,  $\Psi(\cdot, \cdot)$  conhecidas, com média e variância,  $E(Y) = \boldsymbol{\mu}$  e  $Var(Y) = \Psi^{(2,0)^{-1}}$ , respectivamente. E, ainda,  $E\{T(Y)\} = -\Psi^{(0,1)}$ ,  $Var\{T(Y)\} = -\Psi^{(0,2)}$  e  $Cov\{Y, T(Y)\} = 0$ . A notação  $\Psi^{(1,0)}$  é a primeira derivada da função  $\Psi(\boldsymbol{\mu}, \boldsymbol{\phi})$  com relação a  $\boldsymbol{\mu}$  e  $\Psi^{(0,1)}$  é a primeira derivada da função  $\Psi(\boldsymbol{\mu}, \boldsymbol{\phi})$  com relação a  $\boldsymbol{\phi}$  e assim por diante em relação a outras ordens.

- Componentes sistemáticas

para o submodelo média

$$g(\boldsymbol{\mu}) = \boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta},$$

para o submodelo dispersão

$$h(\boldsymbol{\phi}) = \boldsymbol{\tau} = \mathbf{S}\boldsymbol{\gamma},$$

sendo que  $X$  e  $S$  são matrizes  $n \times p$  e  $n \times q$ , respectivamente, e não necessitam ser exclusivas;  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_n)^T$  e  $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_n)^T$  são os vetores dos parâmetros e as funções  $g(\cdot)$  e  $h(\cdot)$  são chamadas de função de ligação da média e dispersão, respectivamente.

Os MLG são um caso particular dos MLGS, tomando  $\boldsymbol{\phi}$  constante.

## 5.2 Estimadores de $O(n^{-1})$ de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$

Como nos capítulos anteriores, a teoria da verossimilhança é adotada. Para obtenção dos estimadores dos MLGS, para utilizar o processo iterativo de Newton Raphson, é necessária a log verossimilhança e, no caso do método score de Fisher, são necessárias a função score e a matriz de informação de Fisher.

## Log-verossimilhança dos MLGS

A função verossimilhança é a densidade conjunta das  $n$  observações dada pela expressão (5.1). Aplicando-se o  $\log$ , toma a seguinte forma:

$$\ell(\beta, \gamma) = \sum_{i=1}^n \left\{ (y_i - \mu_i) \Psi^{(1,0)}(\mu_i, \phi_i) + \phi_i T(y_i) + \Psi(\mu_i, \phi_i) \right\} + \sum_{i=1}^n \log A(y_i) \quad (5.2)$$

suposta regular e derivável até a terceira ordem.

Denota-se por

$$\begin{aligned} m_{li} &= \frac{\partial^l \mu_i}{\partial \eta_i^l} \quad \text{e} \quad \phi_{li} = \frac{d^l \phi_i}{d\tau_i^l}, \\ M_l &= \text{diag} \{m_{l1}, \dots, m_{ln}\} \\ \Phi_l &= \text{diag} \{\phi_{l1}, \dots, \phi_{ln}\} \\ \Psi^{(2,0)} &= \text{diag} \left\{ \psi_1^{(2,0)}, \dots, \psi_n^{(2,0)} \right\}, \\ \Psi^{(0,2)} &= \text{diag} \left\{ \psi_1^{(0,2)}, \dots, \psi_n^{(0,2)} \right\} \end{aligned}$$

sendo que  $l = 1, 2$  e  $i = 1, \dots, n$

## Função escore de $\beta$ e $\gamma$

A função escore pode ser particionada em duas submatrizes, uma devida a  $\beta$  e outra a  $\gamma$ , isto é,

$$U = U(\beta, \gamma) = \begin{pmatrix} \frac{\partial \ell(\beta, \gamma)}{\partial \beta} \\ \frac{\partial \ell(\beta, \gamma)}{\partial \gamma} \end{pmatrix} = \begin{pmatrix} X^T \Psi^{(0,2)} M_1 (\mathbf{y} - \boldsymbol{\mu}) \\ -S^T \Phi_1 \boldsymbol{\nu} \end{pmatrix}, \quad (5.3)$$

sendo  $\boldsymbol{\nu} = (v_1, \dots, v_n)^T$  com

$$v_i = \psi_i^{(1,1)} (y_i - \mu_i) + T(y_i) + \psi_i^{(0,1)}.$$

## Matriz de informação de $\beta$ e $\gamma$

A matriz informação total é bloco-diagonal dada por

$$K(\beta, \gamma) = \begin{pmatrix} K_{\beta, \beta} & 0 \\ 0 & K_{\gamma, \gamma} \end{pmatrix}, \quad (5.4)$$

onde as matrizes informações de  $\beta$  e  $\gamma$  são

$$K_{\beta, \beta} = X^T \Psi^{(2,0)} M_1^2 X \quad \text{e} \quad K_{\gamma, \gamma} = -S^T \Psi^{(0,2)} \Phi_1^2 S,$$

respectivamente. Desse modo os vetores  $\beta$  e  $\gamma$  são globalmente ortogonais e assim os estimadores são assintoticamente independentes.

### Estimadores de $\hat{\beta}$ e $\hat{\gamma}$ pelo método escore de Fisher

Como se sabe,  $\hat{\beta}$  e  $\hat{\gamma}$  satisfazem as equações não-lineares  $U(\hat{\beta}, \hat{\gamma}) = 0$ . Expandindo em série de Taylor, em um ponto inicial  $\beta_0$ , o procedimento iterativo no  $m$ -ésimo +1 passo torna-se

$$\begin{pmatrix} \hat{\beta}^{(m+1)} \\ \hat{\gamma}^{(m+1)} \end{pmatrix} = \begin{pmatrix} \hat{\beta}^{(m)} \\ \hat{\gamma}^{(m)} \end{pmatrix} + (K(\beta, \gamma)^{-1})^{(m)} (U(\beta, \gamma))^m$$
, e através de cálculos algébricos obtém-se o sistema de equações para estimar  $\beta$  e  $\gamma$  iterativamente

$$\begin{aligned} X^T \Psi^{(0,2)^m} M_1^{(m)^2} X \beta^{(m+1)} &= X^T \Psi^{(2,0)^m} M_1^{(m)^2} \xi_1^{(m)}, \\ -S^T \Psi^{(0,2)^m} \Phi_1^{(m)^2} S \gamma^{(m+1)} &= -S^T \Psi^{(0,2)^m} \Phi_1^{(m)^2} \xi_2^{(m)} \end{aligned} \quad (5.5)$$

sendo  $\xi_1 = \eta + M_1^{-1} (\mathbf{y} - \boldsymbol{\mu})$  e o  $\xi_2 = \tau - \Psi^{(0,2)^{-1}} \phi_1^{-1} \nu 1$  são vetores  $n \times 1$ , que aqui funcionam como variável dependente ajustada, fazendo uma analogia aos MLG e MNLFE.

### 5.3 Cálculo do viés de $O(n^{-1})$ para os MLGS

O artigo de Cordeiro e Botter (2001) apresenta as fórmulas gerais para os cálculos dos vieses de primeira ordem dos estimadores de máxima verossimilhança de  $\hat{\beta}$  e  $\hat{\gamma}$  para os MLGS. De maneira análoga aos demais já apresentados, estes partem da fórmula de Cox e Snell (1968), e após cálculos dos cumulantes e algumas manipulações algébricas obtém-se o viés de  $\hat{\beta}$  do submodelo média dado pela expressão

$$\mathbf{B}(\hat{\beta}) = (X^T \Psi^{(2,0)} M_1^2 X)^{-1} X^T \Psi^{(2,0)} M_1^2 \xi, \quad (5.6)$$

sendo  $\xi = -\frac{1}{2} Z_{\beta d} M_1^{-1} M_2 1$  e a matriz  $Z_{\beta d}$  é a diagonal de  $Z_\beta = X K_{\beta\beta}^{-1} X^T$ .

A principal diferença desse viés para com os modelos lineares generalizados é a função  $\Psi^{(r,s)}(\boldsymbol{\mu}, \boldsymbol{\phi})$ , que, por sua vez, depende da função densidade de probabilidade utilizada, a qual, ao fazer o produto com a matriz  $M_1^2$ , passa a ser considerada como a matriz de peso, igual a  $W$  nos modelos lineares generalizados.

E o viés de  $\hat{\gamma}$  do submodelo de dispersão com notação matricial é dado por:

$$B(\hat{\gamma}) = (S^T \Psi^{(0,2)} \Phi_1^2 S)^{-1} S^T \Psi^{(0,2)} \Phi_1^2 \boldsymbol{\delta}, \quad (5.7)$$

sendo  $\boldsymbol{\delta}$  um vetor  $n \times 1$ , cuja expressão é

$$\boldsymbol{\delta} = \frac{1}{2} (\Psi^{(0,2)} \Phi_1^2)^{-1} \{ Z_{\gamma d} (\Psi^{(0,3)} \Phi_1^3 + \Psi^{(0,2)} \Phi_1 \Phi_2) + Z_{\beta d} \Psi^{(2,1)} M_1^2 \Phi_1 \} 1,$$

e a matriz  $Z_{\gamma d}$  é a diagonal de  $Z_{\gamma} = SK_{\gamma\gamma}^{-1}S^T$ . Dessa forma o  $B(\hat{\gamma})$  é a regressão de mínimos quadrados ponderados de  $\delta$  em (5.7) sobre as colunas de  $S$  com a matriz peso  $\Psi^{(0,2)}\Phi_1^2$ .

## 5.4 Aplicação 3: endoparasitos em ovinos

Com o objetivo de disponibilizar os cálculos dos vieses de  $O(n^{-1})$ , descritos nas seções anteriores para os MLGS, foram implementadas no aplicativo SAS as expressões nessa aplicação.

A produção de carne ovina tem sido limitada principalmente devido à contaminação das pastagens por endoparasitas. Essa população de ovinos localiza-se principalmente nos estados do Rio Grande do Sul e Bahia e no estado do Paraná após o ano de 1992, quando o Governo do Estado criou um programa de incentivo à ovinocultura.

A criação de ovinos, numa região onde o clima sofre grandes variações e em áreas com dimensões restritas em face do tamanho do rebanho, tem sido a causa do elevado índice de disseminação de larvas infectantes, tanto na massa fecal quanto no capim, levando as pastagens a se tornarem um meio propício para o aumento de parasitoses gastrointestinais nos hospedeiros.

O sistema de criação em pequenas áreas usando altas taxas de lotação determinou um extraordinário nível de contaminação das pastagens, levando os proprietários a aumentarem a frequência de aplicação de anti-helmínticos, sendo em algumas situações a cada 15 dias, o que leva à resistência a diversos anti-helmínticos.

Entre as possíveis ações para dar solução ao problema, estão o estabelecimento de esquemas de controles profiláticos rígidos para retardar o agravamento da ocorrência de resistência, sistemas de manejo que reduzam a reinfestação dos animais e programa de seleção visando uma maior qualidade na produção, visto que há uma demanda do consumo por essa carne.

Este estudo é resultado de uma pesquisa de iniciação científica que resultou numa apresentação no 16° *SINAPE* - Simpósio Nacional de Probabilidade e Estatística, 2004.

### 5.4.1 Delineamento do experimento

A coleta do material seguiu um traçado, previamente definido na pastagem, em forma de “W”, no sentido da maior dimensão do piquete, simulando o comportamento de pastejo dos animais. Manualmente, uma amostra a cada quatro passos foi retirada das partes aéreas das forrageiras das espécies Tanzânia, Coast cross e Pensacola.

No laboratório, as amostras foram processadas de acordo com Amarante e Barbosa (1995), para a determinação da contagem de larvas de endoparasitas.

### Variáveis

A variável resposta

$Y \rightarrow op$  = quantidade de endoparasitas;

A matriz  $X$  e  $S$ , composta das seguintes covariáveis:

$X_1 \rightarrow gg$  = variável espécie do capim com três categorias, que, para facilitar a interpretação, foi transformada em duas variáveis indicadoras pertencentes à matriz modelo  $X$ ;

$S_1 \rightarrow mes2$  = variável mês com duas categorias, ou seja, meses chuvosos e secos, pertencente à matriz modelo  $S$ .

### 5.4.2 Modelo adotado

**Componente aleatória** A variável resposta é resultado de contagem e como apresenta superdispersão adotou-se como função densidade de probabilidade a Poisson dupla, isto é,

$$Y \sim PD(\boldsymbol{\mu}, \boldsymbol{\phi}).$$

#### Componente sistemática

- Para o submodelo média,

$$\eta_i = \log(\mu_i) = \beta_0 + \beta_1 gg1_i + \beta_2 gg2_i,$$

- Para o submodelo dispersão,

$$\tau_i = \log(\phi_i) = \gamma mes2_i.$$

#### Observações

- Não houve análises sobre as interações entre as espécies de capim e os meses.
- Não foram avaliadas outras funções de ligação para o modelo adotado.
- Para efeito ilustrativo de como a correção nas estimativas, se torna necessária em amostras pequenas, ajustou-se o modelo em dois tamanhos diferentes, isto é, a amostra obtida pelo experimento contendo 541 observações e uma subamostra aleatória de tamanho 35.

### 5.4.3 Resultados

Os resultados da magnitude dos vieses podem ser comparados nas Tabelas 5.1 e 5.2 para o submodelo da média e dispersão.

Tabela 5.1: Estimativas do submodelo Média da Poisson Dupla

$\hat{\beta}_r$	$n = 541$	Vieses $O(n^{-1})$	$n = 35$	Vieses $O(n^{-1})$
$\hat{\beta}_0$	6.3159 (0.00009)	$-4.55E - 6$	6.6939 (0.0003)	-0.00004
$\hat{\beta}_1$	0.3687 (0.00023)	-0.00002	0.1077 (0.0015)	-0.00034
$\hat{\beta}_2$	-0.3088 (0.00026)	$3.028E - 6$	-0.1545 (0.0012)	0.0004

Para ambos os tamanhos amostrais, os valores dos níveis descritivos para a estatística  $t$  foram menores que 0,0001, em todas estimativas do submodelo média, evidenciando que estas são significativas.

Os vieses usando o tamanho amostral de  $n = 541$ , quando comparados com os erros-padrão (entre parenteses), apresentam os seguintes percentuais: 0, 11% , 0, 44% e 8,69% , enquanto que para a amostra de tamanho 35 ficaram com a grandeza de: 13%, 22%, 33% , respectivamente, evidenciando que a correção de  $O(n^{-1})$  é necessária para pequenas amostras.

Na Tabela 5.2 observam-se os resultados do submodelo dispersão com seus vieses para os dois diferentes tamanhos de amostra.

Tabela 5.2: Estimativas do submodelo dispersão da Poisson dupla

$\hat{\gamma}_R$	$n = 541$	Vieses $O(n^{-1})$	$n = 35$	Vieses $O(n^{-1})$
$\hat{\gamma}_1$	-6.3308 (0.09)	0.0046	-6.15 (0.3701)	-0.075

Nesse submodelo dispersão as estimativas também mostraram-se significativas, com valores descritivos menores que 0,0001, para os dois tamanhos diferentes de amostra. Quanto ao viés para a amostra igual a 541, este representa 5,1% do erro-padrão, enquanto usando amostra de tamanho  $n = 35$  ficou com magnitude de 20%.

Portanto, por meio de uma aplicação os resultados vieram corroborar os obtidos na simulação por Cordeiro e Botter (2001), indicando que as estimativas de  $O(n^{-2})$  deveriam ser usadas para amostras pequenas.

O modelo indica que há evidência de que a quantidade média de infestação de endoparasitas nas pastagens se deve às espécies do capim. A grande variabilidade da quantidade de endoparasitas é possivelmente em decorrência de condições climáticas. Isso está de acordo com Amarante e Barbosa (1995), que afirmam que nos tempos chuvosos há facilitação na condução dessas larvas, contaminando a pastagem.

No Apêndice D encontra-se disponível o programa desse estudo.

# CAPÍTULO 6

## VIESES DE $O(n^{-1})$ DOS MODELOS NÃO-LINEARES GENERALIZADOS SUPERDISPERSADOS

No decorrer deste estudo, ficou evidenciado que em diversas aplicações na área de engenharia da produção a redução do viés nos modelos proporciona ganhos econômicos, pois os produtos podem apresentar maior uniformidade.

Nesse contexto, desenvolveu-se uma fórmula geral para o viés de  $O(n^{-1})$  para os MNLGS, caracterizando-se como a principal contribuição desta pesquisa.

A obtenção das expressões dos cálculos dos vieses de primeira ordem foi demonstrada em detalhes, vindo em seguida um caso especial dos modelos não lineares generalizados superdispersados.

### 6.1 Modelo

Seguindo a metodologia dos MLGS e permitindo que as componentes sistemáticas sejam funções não-lineares, ou seja, se as restrições de linearidade dos parâmetros média ( $\mu$ ) e dispersão ( $\phi$ ), forem flexibilizadas têm-se então os modelos não-lineares generalizados superdispersados.

Podem ser caracterizados da seguinte forma:

- A componente aleatória é pertencente à família superdispersada apresentada no Capítulo 3 com função densidade de probabilidade dada por Dey et al. (1997)

$$\pi(y; \mu, \phi) = A(y) \exp \left\{ (y - \mu) \Psi^{(1,0)}(\mu, \phi) + \phi T(y) + \Psi(\mu, \phi) \right\}, \quad (6.1)$$

- Duas componentes sistemáticas, porém com funções não-lineares, as quais são como segue.

Uma para a média

$$g(\boldsymbol{\mu}) = \boldsymbol{\eta} = f_1(\mathbf{X}; \boldsymbol{\beta}), \quad (6.2)$$

onde  $g(\cdot)$  é denominada função ligação da média, conhecida e biunívoca, sendo  $X$  a matriz conhecida de covariáveis e  $\beta$  o vetor de  $p$  parâmetros desconhecidos a ser estimado, com  $p < n$  observações e  $f_1$  uma função contínua e diferenciável, de modo que  $X$  e  $\tilde{X} = \frac{\partial f_1(X; \beta)}{\partial \beta}$  tenham posto completo.

E um outro modelo para a dispersão

$$h(\boldsymbol{\phi}) = \boldsymbol{\tau} = f_2(\mathbf{S}; \boldsymbol{\gamma}) \quad (6.3)$$

onde  $h(\cdot)$  é denominada função ligação da dispersão, conhecida e biunívoca, sendo  $S$  a matriz conhecida de covariáveis e  $\gamma$  o vetor de  $q$  parâmetros desconhecidos a ser estimado, com  $q < n$  observações e  $f_2$  uma função contínua e diferenciável com  $S$  e  $\tilde{S} = \frac{\partial f_2(S; \gamma)}{\partial \gamma}$  de posto completo.

### 6.1.1 Estimadores de MV para $\beta$ e $\gamma$

O procedimento para obter os EMV de  $\beta$  e  $\gamma$  é análogo ao apresentado no Capítulo 5, mudando somente as matrizes modelos  $X$  e  $S$  por  $\tilde{X}$  e  $\tilde{S}$  respectivamente.

Denotam-se os valores amostrais  $y_1, y_2, \dots, y_n$  e a função de log-verossimilhança por

$$\ell(\beta, \gamma) = \sum_i^n \left\{ (y_i - \mu_i) \Psi^{(1,0)}(\mu_i, \phi_i) + \phi_i T(y_i) + \Psi(\mu_i, \phi_i) \right\} + \sum_i^n \log A(y_i),$$

com  $T(y)$ ,  $\Psi(\mu, \phi)$  e  $A(y)$  são funções conhecidas.

Vetores escores para  $\beta$  e  $\gamma$

Os vetores escores para  $\beta$  e  $\gamma$  são:

$$U = U(\beta, \gamma) = \begin{pmatrix} \frac{\partial \ell(\beta, \gamma)}{\partial \beta} \\ \frac{\partial \ell(\beta, \gamma)}{\partial \gamma} \end{pmatrix} = \begin{pmatrix} \tilde{X}^T \Psi^{(0,2)} M_1 (\mathbf{y} - \boldsymbol{\mu}) \\ -\tilde{S}^T \boldsymbol{\phi}_1 \mathbf{v} \end{pmatrix},$$

$$v_i = \Psi_i^{(1,1)}(y_i - \mu_i) + T(y_i) + \Psi_i^{(0,1)}.$$

Sendo que  $M_1 = \frac{\partial \mu}{\partial \eta}$ ,  $\phi_1 = \frac{\partial \phi}{\partial \tau}$  e  $\Psi^{(0,2)}$  é a segunda derivada da função  $\Psi(\boldsymbol{\mu}, \boldsymbol{\phi})$  com relação a  $\phi$  e as demais ordens das derivadas conforme sobreescritos da função como já apresentado no Capítulo 3.

Matrizes de informação de  $\beta$  e  $\gamma$



A matriz de informação de Fisher em relação a  $\mu$ , por definição, é a derivada segunda do logaritmo da verossimilhança em relação a  $\beta$  e  $\gamma$ , ou seja,

$$K(\beta) = \frac{\partial^2 \ell}{\partial \beta \partial \beta^T} \quad \text{e} \quad K(\gamma) = \frac{\partial^2 \ell}{\partial \gamma \partial \gamma^T},$$

e a matriz de informação total é dada pela matriz bloco-diagonal

$$K(\beta, \gamma) = \begin{pmatrix} K_{\beta, \beta} & 0 \\ 0 & K_{\gamma, \gamma} \end{pmatrix}$$

onde as matrizes informações de  $\beta$  e  $\gamma$  são

$$K_{\beta, \beta} = \tilde{X}^T \Psi^{(2,0)} M_1^2 \tilde{X} \quad \text{e} \quad K_{\gamma, \gamma} = -\tilde{S}^T \Psi^{(0,2)} \Phi_1^2 \tilde{S},$$

respectivamente. Como os vetores  $\beta$  e  $\gamma$  são globalmente ortogonais, os estimadores são assintoticamente independentes e podem ser estimados separadamente.

Para determinar as estimativas desse modelo pode-se utilizar a procedure NLMIXED do aplicativo SAS, que depende fortemente dos valores iniciais.

Um fluxograma está apresentado no próximo capítulo, para melhor entendimento do processo do uso do aplicativo SAS para determinar as estimativas e o cálculos dos vieses para o modelo em estudo.

## 6.2 Vieses de $O(n^{-1})$ para $\hat{\beta}$ e $\hat{\gamma}$

Aqui foram demonstrados com certo rigor os detalhes sobre como determinar as expressões dos cálculos dos vieses para os estimadores dos submodelos média e dispersão, visto que estas são as principais contribuições desta pesquisa.

Assume-se que  $\hat{\beta}$  e  $\hat{\gamma}$  existem, são finitos e únicos dados pela solução de  $U(\hat{\beta}, \hat{\gamma}) = 0$ , que convergem para os verdadeiros valores dos parâmetros quando  $n \rightarrow \infty$  e que suas distribuições assintóticas são:

i)  $\hat{\beta} \sim N_p(\beta, k_{\beta, \beta}^{-1});$

ii)  $\hat{\gamma} \sim N_q(\gamma, k_{\gamma, \gamma}^{-1}).$

Denotam-se: as letras minúsculas ( $r, s, t, u, \dots$ ) para as derivadas da log-verossimilhança das componentes do vetor  $\beta$  do submodelo média e as letras maiúsculas ( $R, S, T, U, \dots$ ) as das componentes do vetor  $\gamma$  do submodelo dispersão, por  $B(\hat{\beta}_r)$  e  $B(\hat{\gamma}_R)$  os vieses de  $O(n^{-1})$  dos estimadores de máxima verossimilhança  $\hat{\beta}_r$  e  $\hat{\gamma}_R$ , para  $r = 1, 2, \dots, p$  e  $R = 1, 2, \dots, q$ .

### 6.2.1 Viés de $\hat{\beta}$ , $\hat{\eta}$ e $\hat{\mu}$

A fórmula geral dada por Cox e Snell (1968) para os vieses é a que segue

$$B(\hat{\beta}_r) = \sum_s^p \sum_t^p \sum_u^p \kappa^{rs} \kappa^{tu} \left[ \frac{\kappa_{rst}}{2} + \kappa_{rs,t} \right],$$

que usando a relação  $\kappa_{rs,t} = \kappa_{rst} - \kappa_{rs}^{(t)}$ , obtida pela identidade de Bartlett nos MNLGS torna-se:

$$B(\hat{\beta}_r) = \sum_s^p \sum_t^p \sum_u^p \kappa^{rs} \kappa^{tu} \left\{ \kappa_{st}^{(u)} - \frac{1}{2} \kappa_{stu} \right\} \quad (6.4)$$

Observa-se que são necessários os cumulantes  $\kappa_{st}^{(u)}$  e  $\kappa_{stu}$  para a obtenção do viés, os quais estão demonstrados no Apêndice A cujas expressões são:

Cumulante  $\kappa_{rst}$

$$\kappa_{rst} = \sum \left\{ \begin{array}{l} -2\Psi^{(3,0)} M_1^3(s, t, u) - 3\Psi^{(2,0)} M_1 M_2(s, t, u) - \\ \Psi^{(2,0)} M_1^2[(su, t) + (tu, s) + (st, u)] \end{array} \right\}.$$

Cumulante  $\kappa_{st}^{(u)}$

$$\kappa_{st}^{(u)} = - \sum \Psi^{(3,0)} M_1^3(s, t, u) - 2\Psi^{(2,0)} M_1 M_2(s, t, u) - \Psi^{(2,0)} M_1^2[(su, t) + (tu, s)].$$

Substituindo esses cumulantes na expressão (6.4) obtém-se:

$$B(\hat{\beta}_r) = \sum_{\beta} \kappa^{rs} \kappa^{tu} \left\{ \begin{array}{l} \sum (-\Psi^{(3,0)} M_1^3(s, t, u) - 2\Psi^{(2,0)} M_1 M_2(s, t, u) \\ - \Psi^{(2,0)} M_1^2[(su, t) + (tu, s)]) \\ - \frac{1}{2} \sum (-2\Psi^{(3,0)} M_1^3(s, t, u) - 3\Psi^{(2,0)} M_1 M_2(s, t, u) \\ - \Psi^{(2,0)} M_1^2[(su, t) + (tu, s) + (st, u)]) \end{array} \right\}$$

$$B(\hat{\beta}_r) = \sum_{\beta} \kappa^{rs} \kappa^{tu} \left\{ \begin{array}{l} \sum (-\frac{1}{2} \Psi^{(2,0)} M_1 M_2(s, t, u) - \frac{1}{2} \Psi^{(2,0)} M_1^2(su, t)) \\ - \frac{1}{2} \Psi^{(2,0)} M_1^2(tu, s) + \frac{1}{2} \Psi^{(2,0)} M_1^2(st, u) \end{array} \right\}.$$

Fazendo o somatório, permutação de índices e alguns cálculos, tem-se que

$$\begin{aligned}
B(\hat{\beta}_r) &= \sum_{stu} \kappa^{rs} \kappa^{tu}(s, t, u) \sum_i -\frac{1}{2} \Psi^{(2,0)} M_1 M_2 \\
&\quad + \sum_{stu} \kappa^{rs} \kappa^{tu}(tu, s) \sum_i -\frac{1}{2} \Psi^{(2,0)} M_1^2. \\
B(\hat{\beta}_r) &= \sum_{stu} \kappa^{rs}(s) \kappa^{tu}(t, u) \sum_i -\frac{1}{2} \Psi^{(2,0)} M_1 M_2 \\
&\quad + \sum_{stu} \kappa^{rs} \kappa^{tu}(tu, s) \sum_i -\frac{1}{2} \Psi^{(2,0)} M_1^2.
\end{aligned} \tag{6.5}$$

Para o cálculo do viés será denotado por  $D_\beta = \text{diag}\{d_1, \dots, d_n\}$  com cada

$$\begin{aligned}
d_i &= \text{tr} \left( \tilde{X}_i \left( \tilde{X}^T \Psi^{(2,0)} M_1^2 \tilde{X} \right)^{-1} \right), \text{ e ainda, por} \\
Z_\beta &= \tilde{X} K_{\beta, \beta}^{-1} \tilde{X}^T, \quad Z_{\beta d} = \text{diag}\{z_{\beta 11}, \dots, z_{\beta nn}\} \text{ e } \tilde{X} = \frac{\partial^2 \eta}{\partial \beta \partial \beta^T}.
\end{aligned}$$

Substituindo na expressão (6.5), rearranjando os termos e colocando na forma matricial, tem-se

$$\begin{aligned}
B(\hat{\beta}) &= -\frac{1}{2} K_{\beta, \beta}^{-1} \tilde{X}^T \tilde{X} K_{\beta, \beta}^{-1} \tilde{X}^T \Psi^{(2,0)} M_1 M_2 - \frac{1}{2} \Psi^{(2,0)} M_1^2 K_{\beta, \beta}^{-1} \tilde{X}^T K_{\beta, \beta}^{-1} \tilde{X}^T \\
&= -\frac{1}{2} K_{\beta, \beta}^{-1} \tilde{X}^T \Psi^{(2,0)} \left( \tilde{X} K_{\beta, \beta}^{-1} \tilde{X}^T M_1 M_2 + K_{\beta, \beta}^{-1} \tilde{X}^T M_1^2 \right) \\
&= -\frac{1}{2} K_{\beta, \beta}^{-1} \tilde{X}^T \Psi^{(2,0)} (Z_{\beta d} M_1 M_2 + D_\beta M_1^2) \\
&= -\frac{1}{2} K_{\beta, \beta}^{-1} \tilde{X}^T \Psi^{(2,0)} (Z_{\beta d} M_1 M_2 + D_\beta M_1^2).
\end{aligned}$$

Cordeiro e Botter (2001) propõem que o cálculo do viés pode ser considerado como uma regressão linear ponderada. Pode-se, então reescrever a expressão acima da seguinte maneira

$$\begin{aligned}
B(\hat{\beta}) &= -\frac{1}{2} \left[ \tilde{X}^T \Psi^{(2,0)} M_1^2 \tilde{X} \right]^{-1} \tilde{X}^T \Psi^{(2,0)} M_1 M_1 M_1^{-1} Z_{\beta d} M_2 \mathbf{1} \\
&\quad - \frac{1}{2} \left[ \tilde{X}^T \Psi^{(2,0)} M_1^2 \tilde{X} \right]^{-1} \tilde{X}^T \Psi^{(2,0)} D_\beta M_1^2 \\
B(\hat{\beta}) &= \left[ \tilde{X}^T \Psi^{(2,0)} M_1^2 \tilde{X} \right]^{-1} \tilde{X}^T \Psi^{(2,0)} M_1^2 \left\{ -\frac{1}{2} Z_{\beta d} M_1^{-1} M_2 - \frac{1}{2} D_\beta \right\} \mathbf{1}.
\end{aligned}$$

Fazendo  $\xi_1 = -\frac{1}{2} Z_{\beta d} M_1^{-1} M_2$  e  $\xi_2 = -\frac{1}{2} D_\beta$ . A expressão do viés de  $\hat{\beta}$  é:

$$B(\hat{\beta}) = \left[ \tilde{X}^T \Psi^{(2,0)} M_1^2 \tilde{X} \right]^{-1} \tilde{X}^T \Psi^{(2,0)} M_1^2 \{\xi_1 + \xi_2\} \mathbf{1}, \tag{6.6}$$

onde  $\mathbf{1}$  é um vetor de uns e  $\xi_2$  é a contribuição devida à parte não-linear do modelo.

Portanto, o viés é obtido simplesmente como um vetor de coeficientes na regressão linear e não-linear clássica de  $\xi_1$  e  $\xi_2$  respectivamente sobre colunas da matriz  $\tilde{X}$  e com a matriz pesos  $\Psi^{(2,0)} M_1^2$ .

### Viés de $\hat{\eta}$

Para obter esse viés, expande-se o preditor não-linear  $\hat{\eta}$  em série de Taylor, até segunda ordem, na vizinhança de  $\beta$ , isto é,

$$\hat{\eta} = f(\mathbf{X}, \hat{\beta}) = f(X, \beta) + (\beta - \hat{\beta}) \frac{\partial f(X, \beta)}{\partial \beta} + \frac{(\beta - \hat{\beta})(\beta - \hat{\beta})^T}{2!} \frac{\partial^2 f(X, \beta)}{\partial \beta \partial \beta^T},$$

que rearranjando os termos tem-se

$$f(X, \hat{\beta}) - f(X, \beta) = (\beta - \hat{\beta}) \frac{\partial f(X, \beta)}{\partial \beta} + \frac{(\beta - \hat{\beta})(\beta - \hat{\beta})^T}{2!} \frac{\partial^2 f(X, \beta)}{\partial \beta \partial \beta^T}.$$

Calculando a esperança, tem-se

$$\begin{aligned} E\left(f(X, \hat{\beta}) - f(X, \beta)\right) &= E\left(\beta - \hat{\beta}\right) \frac{\partial f(X, \beta)}{\partial \beta} + E\left[\frac{(\beta - \hat{\beta})(\beta - \hat{\beta})^T}{2!}\right] \frac{\partial^2 f(X, \beta)}{\partial \beta \partial \beta^T} \\ B(\hat{\eta}) &= B(\hat{\beta}) \tilde{X} + \frac{1}{2} V(\hat{\beta}) \tilde{\tilde{X}} \\ B(\hat{\eta}) &= B(\hat{\beta}) \tilde{X} + \frac{1}{2} \left(\tilde{X}^T \Psi^{(2,0)} M_1^2 \tilde{X}\right) \tilde{\tilde{X}} \\ &= B(\hat{\beta}) \tilde{X} + \frac{1}{2} D_\beta. \end{aligned}$$

Substituindo o viés de  $\hat{\beta}$  na expressão acima

$$\begin{aligned} B(\hat{\eta}) &= \tilde{X} \left[ \tilde{X}^T \Psi^{(2,0)} M_1^2 \tilde{X} \right]^{-1} \tilde{X}^T \Psi^{(2,0)} M_1^2 \{\xi_1 + \xi_2\} + \frac{1}{2} D_\beta \\ &= Z_\beta \Psi^{(2,0)} M_1^2 \{\xi_1 + \xi_2\} + \frac{1}{2} D_\beta \\ &= Z_\beta \Psi^{(2,0)} M_1^2 \left\{ \xi_1 - \frac{1}{2} D_\beta \right\} + \frac{1}{2} D_\beta \\ B(\hat{\eta}) &= Z_\beta \Psi^{(2,0)} M_1^2 \xi_1 - \frac{1}{2} Z_\beta \Psi^{(2,0)} M_1^2 D_\beta + \frac{1}{2} D_\beta \end{aligned} \tag{6.7}$$

Finalizando, o viés de  $\hat{\eta}$  depende do viés  $B(\hat{\beta})$  e da matriz  $D_\beta$ .

Portanto, a expressão de  $B(\hat{\boldsymbol{\eta}})$  é

$$B(\hat{\boldsymbol{\eta}}) = Z_{\beta} \Psi^{(2,0)} M_1^2 \boldsymbol{\xi}_1 - \frac{1}{2} \{ Z_{\beta} \Psi^{(2,0)} M_1^2 - I \} D_{\beta}.$$

### Viés de $\hat{\mu}$

Nesse caso expande-se em série de Taylor  $f_1^{-1}(\hat{\boldsymbol{\eta}}) = \hat{\mu}$ , isto é,

$$f_1^{-1}(\hat{\boldsymbol{\eta}}) = f_1^{-1}(\boldsymbol{\eta}) + \frac{df_1^{-1}(\boldsymbol{\eta})}{d\eta_r} (\hat{\eta}_r - \eta_r) + \frac{1}{2} \frac{d^2 f_1^{-1}(\boldsymbol{\eta})}{d\eta_r d\eta_s} (\hat{\eta}_r - \eta_r) (\hat{\eta}_s - \eta_s),$$

e fazendo  $f_1^{-1}(\hat{\boldsymbol{\eta}}) - f_1^{-1}(\boldsymbol{\eta})$  e em seguida calculando a esperança, tem-se

$$\begin{aligned} B(\hat{\boldsymbol{\mu}}) &= B(\hat{\boldsymbol{\eta}}) \frac{\partial \mu}{\partial \boldsymbol{\eta}} + \frac{1}{2} \text{var}(\hat{\boldsymbol{\eta}}) \frac{\partial^2 \mu}{\partial \boldsymbol{\eta}^2} \\ &= B(\hat{\boldsymbol{\eta}}) M_1 + \frac{1}{2} \text{var}(\hat{\boldsymbol{\eta}}) M_2. \text{ Como } \text{Var}(\hat{\boldsymbol{\eta}}) = Z_{\beta d} \text{ tem-se que} \\ &= M_1 \left[ Z_{\beta} \Psi^{(2,0)} M_1^2 \boldsymbol{\xi}_1 - \frac{1}{2} (Z_{\beta} \Psi^{(2,0)} M_1^2 - I) D_{\beta} \right] + \frac{1}{2} M_2 Z_{\beta d} \mathbf{1} \\ &= M_1 \left[ Z_{\beta} \Psi^{(2,0)} M_1^2 \boldsymbol{\xi}_1 - \frac{1}{2} Z_{\beta} \Psi^{(2,0)} M_1^2 D_{\beta} + \frac{1}{2} D_{\beta} \right] + \frac{1}{2} Z_{\beta d} M_2 \mathbf{1} \\ &= M_1 Z_{\beta} \Psi^{(2,0)} M_1^2 \boldsymbol{\xi}_1 - \frac{1}{2} M_1 Z_{\beta} \Psi^{(2,0)} M_1^2 D_{\beta} + \frac{1}{2} M_1 D_{\beta} + \frac{1}{2} Z_{\beta d} M_2. \end{aligned}$$

O viés de  $\hat{\boldsymbol{\mu}}$  depende do viés  $B(\hat{\boldsymbol{\eta}})$  e da variância de  $\hat{\boldsymbol{\eta}}$ . Portanto, a expressão é a que segue:

$$B(\hat{\boldsymbol{\mu}}) = \frac{1}{2} (M_2 Z_{\beta d} \mathbf{1} + 2 M_1 Z_{\beta} \Psi^{(2,0)} M_1^2 \boldsymbol{\xi}_1) - \frac{1}{2} M_1 (Z_{\beta} \Psi^{(2,0)} M_1^2 - I) D_{\beta}. \quad (6.8)$$

### 6.2.2 Viés de $\hat{\gamma}$ , $\hat{\tau}$ e $\hat{\phi}$

Para obtenção do viés de  $\hat{\gamma}$  a fórmula de Cox e Snell (1968) nos MNLGS é dada por:

$$B(\hat{\boldsymbol{\gamma}}_R) = \sum_{\gamma} \kappa^{RS} \kappa^{TU} \left( \kappa_{ST}^{(U)} - \frac{1}{2} \kappa_{STU} \right) - \frac{1}{2} \sum_{\beta, \gamma} \kappa^{RS} \kappa^{tu} \kappa_{Stu}. \quad (6.9)$$

Os cumulantes necessários para calcular esse viés são:  $\kappa_{ST}$ ,  $\kappa_{ST}^{(U)}$ ,  $\kappa_{STU}$  e  $\kappa_{Stu}$ , que são obtidos no Apêndice A cujas expressões são apresentadas a seguir:

- i)  $\kappa_{ST} = \sum \Psi^{(0,2)} \phi_1^2(S, T)$
- ii)  $\kappa_{ST}^{(U)} = \sum \{ \Psi^{(0,3)} \phi_1^3(S, T, U) + 2 \Psi^{(0,2)} \phi_1 \phi_2(S, T, U) + \Psi^{(0,2)} \phi_1^2(ST, U + UT, S) \}$

$$\text{iii) } \kappa_{STU} = \sum \{ \Psi^{(0,3)} \phi_1^3 + 3\Psi^{(0,2)} \phi_1 \phi_2 (S, T, U) \\ + \Psi^{(0,2)} \phi_1^2 [(ST, U) + (SU, T) + (UT, S)] \}$$

$$\text{iv) } \kappa_{stu} = \sum -\Psi^{(2,1)} \phi_1 M_1^2 (S, t, u).$$

Substituindo os cumulantes na expressão (6.9), tem-se

$$B(\hat{\gamma}_R) = \sum_{\gamma} \kappa^{RS} \kappa^{TU} \{ \sum \Psi^{(0,3)} \phi_1^3 (S, T, U) + \sum 2\Psi^{(0,2)} \phi_1 \phi_2 (S, T, U) \\ + \sum \Psi^{(0,2)} \phi_1^2 (ST, U + UT, S) - \frac{1}{2} \sum \Psi^{(0,3)} \phi_1^3 \\ + 3\Psi^{(0,2)} \phi_1 \phi_2 (S, T, U) + \Psi^{(0,2)} \phi_1^2 [(ST, U) \\ + (SU, T) + (UT, S)] - \frac{1}{2} \sum_{\beta, \gamma} \kappa^{RS} \kappa^{tu} [-\Psi^{(2,1)} \phi_1 M_1^2 (S, t, u)] \},$$

fazendo algumas operações algébricas, obtém-se

$$B(\hat{\gamma}_R) = \sum_{\gamma} \kappa^{RS} \kappa^{TU} \sum_i \left\{ \frac{1}{2} \Psi^{(0,3)} \phi_1^3 (S, T, U) \right. \quad (6.10) \\ \left. + \frac{1}{2} \Psi^{(0,2)} \phi_1 \phi_2 (S, T, U) \right. \\ \left. + \frac{1}{2} \Psi^{(0,2)} \phi_1^2 (ST, U + UT, S) \right. \\ \left. - \frac{1}{2} \Psi^{(0,2)} \phi_1^2 (SU, T) \right\} \\ + \sum_{\beta, \gamma} \kappa^{RS} \kappa^{tu} \frac{1}{2} \Psi^{(2,1)} \phi_1 M_1^2 (S, t, u).$$

E agora, fazendo as permutações dos somatórios e índices e passando para a notação matricial,

$$B(\hat{\gamma}) = \sum_i \sum_{RSTu} \kappa^{RS} \kappa^{TU} (S, T, U) \left( \frac{1}{2} \Psi^{(0,3)} \phi_1^3 + \frac{1}{2} \Psi^{(0,2)} \phi_1 \phi_2 \right) + \\ \sum_i \sum_{RSTU} \kappa^{RS} \kappa^{TU} (ST, u) \frac{1}{2} \Psi^{(0,2)} \phi_1^2 + \\ \sum_i \sum_{RSTU} \kappa^{RS} \kappa^{TU} (uT, S) \frac{1}{2} \Psi^{(0,2)} \phi_1^2 - \\ \sum_i \sum_{RSTU} \kappa^{RS} \kappa^{TU} (SU, T) \frac{1}{2} \Psi^{(0,2)} \phi_1^2 + \\ \sum_i \sum_{RStu} \kappa^{RS} \kappa^{tu} (S, t, u) \frac{1}{2} \Psi^{(2,1)} \phi_1 M_1^2 (S, t, u).$$

Após algumas operações

$$\begin{aligned}
B(\hat{\gamma}) &= \sum_i \sum_{RSTU} \kappa^{RS} \kappa^{TU}(S, T, U) \left( \frac{1}{2} \Psi^{(0,3)} \phi_1^3 + \frac{1}{2} \Psi^{(0,2)} \phi_1 \phi_2 \right) + \quad (6.11) \\
&\quad \sum_i \sum_{RStu} \kappa^{RS} \kappa^{tu}(S, t, u) \frac{1}{2} \Psi^{(2,1)} \phi_1 M_1^2 + \\
&\quad \sum_i \sum_{RSTU} \kappa^{RS} \kappa^{TU}(UT, S) \frac{1}{2} \Psi^{(0,2)} \phi_1^2.
\end{aligned}$$

Lembrando que

$$\begin{aligned}
(S) &= \frac{\partial \tau}{\partial \gamma_S}, \quad (T, U) = \frac{\partial \tau}{\partial \gamma_T} \frac{\partial \tau}{\partial \gamma_U}, \\
(TU) &= \frac{\partial^2 \tau}{\partial \gamma_T \partial \gamma_U} \text{ e } (t, u) = \frac{\partial \eta}{\partial \beta_t} \frac{\partial \eta}{\partial \beta_u} \\
\tilde{S} &= \frac{\partial \tau}{\partial \gamma} \text{ e } \tilde{\tilde{S}} = \frac{\partial^2 \tau}{\partial \gamma^T \partial \gamma}
\end{aligned}$$

e ainda que

$$\sum_{RSTU} \kappa^{RS} \kappa^{TU}(S, T, U) = \sum_{RSTU} \kappa^{RS}(S) \kappa^{TU}(T, U),$$

e

$$\sum_{RSTU} \kappa^{RS}(S) = -K_{\gamma, \gamma}^{-1} \tilde{S}^T \text{ e } \sum_{RSTU} \kappa^{TU}(T, U) = -\tilde{S} K_{\gamma, \gamma}^{-1} \tilde{S}^T,$$

e no segundo somatório da expressão (6.10)

$$\sum_{RStu} \kappa^{RS} \kappa^{tu}(S, t, u) = \sum_{RStu} \kappa^{RS}(S) \kappa^{tu}(t, u),$$

tem-se  $\sum_{RStu} \kappa^{RS}(S) = -K_{\gamma, \gamma}^{-1} S^T$  e  $\sum_{RStu} \kappa^{tu}(t, u) = -\tilde{X} K_{\beta, \beta}^{-1} \tilde{X}^T$ .

E por último tem-se

$$\sum_{RSTU} \kappa^{RS} \kappa^{TU}(UT, S) = \sum_{RSTU} \kappa^{RS}(S) \kappa^{TU}(TU),$$

sendo que  $\sum_{RSTU} \kappa^{RS}(S) = -K_{\gamma, \gamma}^{-1} S^T$  e  $\sum_{RSTU} \kappa^{TU}(TU) = -K_{\gamma, \gamma}^{-1} \tilde{S}$ .

Substituindo os cumulantes na expressão (6.11) e como a inversa da matriz de informação é  $K_{\gamma, \gamma}^{-1} = \left( -\tilde{S}^T \Psi^{(0,2)} \Phi_1^2 \tilde{S} \right)^{-1}$ , a matriz peso é igual a  $\Psi^{(0,2)} \Phi_1^2$  e ainda considerando que

$$\begin{aligned}
\tilde{S}K_{\gamma,\gamma}^{-1}\tilde{S}^T &= Z_{\gamma d}, & \tilde{X}K_{\beta,\beta}^{-1}\tilde{X}^T &= Z_{\beta d} \\
K_{\gamma,\gamma}^{-1}\tilde{\tilde{S}} &= D_\gamma, \text{ sendo que } D_\gamma = \text{diag}(d_i, \dots, d_n), \\
\text{com } d_i &= \text{tr} \left( \tilde{\tilde{S}}_i \left( \tilde{S}^T \Psi^{(0,2)} \Phi_1^2 \tilde{S} \right)^{-1} \right)
\end{aligned}$$

desse modo, o viés de  $\hat{\gamma}$  assume a forma matricial abaixo

$$\begin{aligned}
B(\hat{\gamma}) &= -\frac{1}{2}(\tilde{S}^T \Psi^{(0,2)} \Phi_1^2 \tilde{S})^{-1} \tilde{S}^T (\Psi^{(0,2)} \Phi_1^2) (\Psi^{(0,2)} \Phi_1^2)^{-1} \\
&\quad Z_{\gamma d} (\Psi^{(0,3)} \Phi_1^3 + \Psi^{(0,2)} \Phi_1 \Phi_2) \\
&\quad -\frac{1}{2}(\tilde{S}^T \Psi^{(0,2)} \Phi_1^2 \tilde{S})^{-1} \tilde{S}^T D_\gamma (\Psi^{(0,2)} \Phi_1^2) \\
&\quad -\frac{1}{2}(\tilde{S}^T \Psi^{(0,2)} \Phi_1^2 \tilde{S})^{-1} \tilde{S}^T (\Psi^{(0,2)} \Phi_1^2) (\Psi^{(0,2)} \Phi_1^2)^{-1} Z_{\beta d} (\Psi^{(2,1)} \Phi_1 M_1^2).
\end{aligned}$$

Simplificando

$$B(\hat{\gamma}) = (\tilde{S}^T \Psi^{(0,2)} \Phi_1^2 \tilde{S})^{-1} \tilde{S}^T (\Psi^{(0,2)} \Phi_1^2) \delta \quad (6.12)$$

sendo  $\delta = \delta_1 + \delta_2$ ,

$$\begin{aligned}
\delta_1 &= \left\{ -\frac{1}{2} (\Psi^{(0,2)} \Phi_1^2)^{-1} [Z_{\gamma d} (\Psi^{(0,3)} \Phi_1^3 + \Psi^{(0,2)} \Phi_1 \Phi_2) + Z_{\beta d} (\Psi^{(2,1)} \Phi_1 M_1^2)] \right\} \\
\delta_2 &= -\frac{1}{2} D_\gamma, \text{ sendo a contribuição devida a não-linearidade do modelo.}
\end{aligned}$$

O viés do estimador  $\hat{\gamma}$  é um conjunto de coeficientes de regressão de mínimos quadrados reponderados de  $\delta$  sobre as colunas de  $\tilde{S}$  com matriz de peso  $\Psi^{(0,2)} \Phi_1^2$ .

## Viés de $\hat{\tau}$

Expandindo  $\hat{\tau}$  em série de Taylor em torno de  $\hat{\gamma}$ , obtém-se que

$$B(\hat{\tau}) = \tilde{S}B(\hat{\gamma}) + \frac{1}{2}D_\gamma$$

substituindo o viés de  $\hat{\gamma}$



$$\begin{aligned}
B(\hat{\boldsymbol{\tau}}) &= \tilde{S}(\tilde{S}^T \Psi^{(0,2)} \Phi_1^2 \tilde{S})^{-1} \tilde{S}^T (\Psi^{(0,2)} \Phi_1^2) \boldsymbol{\delta} + \frac{1}{2} D_\gamma \\
&= Z_\gamma \Psi^{(0,2)} \Phi_1^2 (\boldsymbol{\delta}_1 + \boldsymbol{\delta}_2) + \frac{1}{2} D_\gamma \\
&= Z_\gamma \Psi^{(0,2)} \Phi_1^2 \boldsymbol{\delta}_1 + Z_\gamma \Psi^{(0,2)} \Phi_1^2 \boldsymbol{\delta}_2 + \frac{1}{2} D_\gamma \\
&= Z_\gamma \Psi^{(0,2)} \Phi_1^2 \boldsymbol{\delta}_1 + Z_\gamma \Psi^{(0,2)} \Phi_1^2 \left( -\frac{1}{2} D_\gamma \right) + \frac{1}{2} D_\gamma \\
B(\hat{\boldsymbol{\tau}}) &= Z_\gamma \Psi^{(0,2)} \Phi_1^2 \boldsymbol{\delta}_1 - \frac{1}{2} (Z_\gamma \Psi^{(0,2)} \Phi_1^2 - I) D_\gamma, \tag{6.13}
\end{aligned}$$

sendo que o vetor  $\boldsymbol{\delta}_1$  é o mesmo da expressão do viés de  $\hat{\boldsymbol{\beta}}$ .

### Viés de $\hat{\boldsymbol{\phi}}$

Como o viés de  $\hat{\boldsymbol{\mu}}$ , esse também é obtido expandindo em série de Taylor  $f_2^{-1}(\hat{\boldsymbol{\tau}}) = \boldsymbol{\phi}$ , isto é,

$$\begin{aligned}
B(\hat{\boldsymbol{\phi}}) &= B(\hat{\boldsymbol{\tau}}) \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{\tau}} + \frac{1}{2} \text{var}(\hat{\boldsymbol{\tau}}) \frac{\partial^2 \boldsymbol{\phi}}{\partial \boldsymbol{\tau}^2} \\
\text{sendo que } \text{Var}(\hat{\boldsymbol{\tau}}) &= \tilde{S} V(\hat{\boldsymbol{\gamma}}) \tilde{S}^T = Z_{\gamma d} \\
B(\hat{\boldsymbol{\phi}}) &= B(\hat{\boldsymbol{\tau}}) \Phi_1 + \frac{1}{2} \Phi_2 Z_{\gamma d} \mathbf{1} \\
&= \Phi_1 Z_\gamma \Psi^{(0,2)} \Phi_1^2 \boldsymbol{\delta}_1 - \frac{1}{2} (Z_\gamma \Psi^{(0,2)} \Phi_1^2 - I) D_\gamma + \frac{1}{2} \Phi_2 Z_{\gamma d} \\
B(\hat{\boldsymbol{\phi}}) &= \frac{1}{2} (Z_{\gamma d} \Phi_2 \mathbf{1} + 2 \Phi_1 Z_\gamma \Psi^{(0,2)} \Phi_1^2 \boldsymbol{\delta}_1) - \frac{1}{2} (Z_\gamma \Psi^{(0,2)} \Phi_1^2 - I) D_\gamma. \tag{6.14}
\end{aligned}$$

As expressões obtidas neste capítulo para o cálculo dos vieses dos estimadores do submodelo da média e dispersão alcançam o principal objetivo do estudo. E com isso, pode-se determinar as estimativas de  $O(n^{-2})$  para a classe dos MNLGS.

Portanto, pode-se fazer algumas afirmações como:

i) Os vieses são funções das matrizes modelos  $\tilde{X}$  e  $\tilde{S}$ .

ii) Para avaliar esses modelos é necessário determinar as covariâncias de cada submodelo, ou seja,  $Z_\beta = \tilde{X} K_{\beta\beta}^{-1} \tilde{X}^T$  e  $Z_\gamma = \tilde{S} K_{\gamma\gamma}^{-1} \tilde{S}^T$  dos estimadores  $\hat{\boldsymbol{\eta}}$  e  $\hat{\boldsymbol{\tau}}$  e as matrizes diagonais  $Z_{\beta d}$ ,  $Z_{\gamma d}$ ,  $\Psi^{(2,0)}$ ,  $\Psi^{(0,2)}$ ,  $\Psi^{(2,1)}$ ,  $\Psi^{(0,3)}$ ,  $M_1$ ,  $M_2$ ,  $\Phi_1$  e  $\Phi_2$ .

iii) É fácil obter os vieses  $B(\hat{\boldsymbol{\beta}})$  e  $B(\hat{\boldsymbol{\gamma}})$  e implementá-los com auxílio de um aplicativo.

### 6.3 Caso especial

Nesta seção foi demonstrado que os MLGS são um caso particular dos MNLGS, ou seja, que os estimadores de segunda ordem de Botter e Cordeiro (2001) passam a ser um caso especial dos aqui apresentados.

Assume-se que nos MLGS as matrizes modelos  $X = S = 1$ , para os parâmetros do submodelo média e dispersão, ou seja, uma coluna de uns.

Considerando que a função de ligação para a componente sistemática da média é a identidade,  $\eta = g(\boldsymbol{\mu}) = X\beta = \beta\mathbf{1}$  e lembrando que

$$\begin{aligned} M_1 &= \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\eta}}, & M_2 &= \frac{\partial^2 \boldsymbol{\mu}}{\partial \boldsymbol{\eta}^2} \\ \tilde{X} &= \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\beta}} = X = 1, & \tilde{\tilde{X}} &= \frac{\partial^2 \boldsymbol{\eta}}{\partial \boldsymbol{\beta}^2} = 0 \end{aligned}$$

e substituindo na expressão do viés para  $\hat{\boldsymbol{\beta}}$  dos modelos não-lineares generalizados superdispersados dada por

$$B(\hat{\boldsymbol{\beta}}) = \left[ \tilde{X}^T \Psi^{(2,0)} M_1^2 \tilde{X} \right]^{-1} \tilde{X}^T \Psi^{(2,0)} M_1^2 \{\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2\} \mathbf{1},$$

$\boldsymbol{\xi}_1 = -\frac{1}{2} Z_{\beta d} M_1^{-1} M_2$ ,  $\boldsymbol{\xi}_2 = -\frac{1}{2} D_\beta$ , é a parte não-linear do modelo, onde  $\mathbf{1}$  é um vetor de uns e  $D_\beta = \text{diag}\{d_1, \dots, d_n\}$  com  $d_i = \text{tr} \left( \tilde{\tilde{X}}_i \left( \tilde{X}^T \Psi^{(2,0)} M_1^2 \tilde{X} \right)^{-1} \right)$  e ainda  $Z_\beta = \tilde{X} K_{\beta, \beta}^{-1} \tilde{X}^T$  com  $Z_{\beta d} = \text{diag}\{z_{\beta 11}, \dots, z_{\beta nn}\}$ . Portanto, a expressão do viés reduz-se somente à parte linear do modelo que coincide com a expressão dada por Cordeiro e Botter (2001), ou seja,

$$B(\hat{\boldsymbol{\beta}}) = (X^T \Psi^{(2,0)} M_1^2 X)^{-1} X^T \Psi^{(2,0)} M_1^2 \boldsymbol{\xi}, \quad (6.15)$$

onde  $\boldsymbol{\xi} = -\frac{1}{2} Z_{\beta d} M_1^{-1} M_2 \mathbf{1}$ .

O mesmo ocorre para o  $B(\hat{\boldsymbol{\gamma}})$  com função de ligação identidade para a componente sistemática da dispersão. Assim,  $\tau = h(\boldsymbol{\phi}) = S\boldsymbol{\gamma} = \boldsymbol{\gamma}\mathbf{1}$  e a segunda derivada da matriz  $\tilde{\tilde{S}} = \frac{\partial^2 \boldsymbol{\phi}}{\partial \boldsymbol{\gamma}^2} = 0$ .

Novamente, após substituir na expressão do viés  $B(\hat{\boldsymbol{\gamma}})$  dada por (6.12) e como a parte não-linear do modelo é igual a zero, resta do modelo somente a parte linear, ou seja, o viés dos estimadores dos parâmetros do submodelo da dispersão

$$B(\hat{\boldsymbol{\gamma}}) = (S^T \Psi^{(0,2)} \Phi_1^2 S)^{-1} S^T \Psi^{(0,2)} \Phi_1^2 \boldsymbol{\delta}, \quad (6.16)$$

onde denota-se por  $\boldsymbol{\delta}$  o vetor  $n \times 1$  igual a

$$\boldsymbol{\delta} = \frac{1}{2} (\Psi^{(0,2)} \Phi_1^2)^{-1} \{Z_{\gamma d} (\Psi^{(0,3)} \Phi_1^3 + \Psi^{(0,2)} \Phi_1 \Phi_2) + Z_{\beta d} \Psi^{(2,1)} M_1^2 \Phi_1\} \mathbf{1},$$

que está de acordo com Cordeiro e Botter (2001) para os MLGS.

# CAPÍTULO 7

## SIMULAÇÃO E APLICAÇÃO

Este capítulo apresenta um processo de simulação, cujo interesse foi observar o comportamento da magnitude dos vieses para diferentes tamanhos amostrais. Além disso, foram obtidos os vieses em dois experimentos e, portanto, colocando em prática a teoria desenvolvida.

### 7.1 Simulação do modelo passeio aleatório

O modelo simulado pertencente à classe dos MNLGS foi o da distribuição passeio aleatório, que é muito utilizado em economia, em finanças corporativas, na área industrial, entre outras.

Os resultados para comparar os estimadores de máxima verossimilhança  $\hat{\beta}$  e  $\hat{\gamma}$ , seus correspondentes vieses e seus respectivos estimadores corrigidos  $\tilde{\beta}$  e  $\tilde{\gamma}$  foram efetuados via simulação de Monte Carlo.

Para evidenciar que a correção do viés  $O(n^{-1})$  se torna tanto mais apreciável quanto menor for o tamanho da amostra, geraram-se diferentes tamanhos amostrais, que foram de  $n = 15$ ,  $n = 30$  e  $n = 80$  com 10.000 repetições para cada caso.

Para cada amostra de tamanho  $n$  foram estimados os parâmetros e computadas as médias e os desvios-padrão dessas estimativas nas 10.000 repetições com os seus respectivos vieses  $B(\hat{\beta})$  e  $B(\hat{\gamma})$  dadas pelas expressões (6.6) e (6.12).

#### 7.1.1 Modelo não-linear generalizado superdispersado

##### Componente aleatória

Considerou-se a variável resposta aleatória  $Y$ , com distribuição passeio aleatório cuja função densidade é dada pela expressão (3.4).

Para que se possa escrever sob a forma dos MNLGS, é necessário definir o parâmetro de dispersão  $\phi$ . Utilizou-se a parametrização dada por Cordeiro e Botter (2001) pela expressão

$$\phi = -\frac{\delta}{2\theta^2}.$$

Com essa parametrização a log-verossimilhança dos MNLGS toma a seguinte forma

$$p(\mathbf{y}; \boldsymbol{\mu}, \phi) = \exp(\mathbf{y} - \boldsymbol{\mu})\Psi^{(1,0)} + \phi T(\mathbf{y}) + \Psi(\boldsymbol{\mu}, \phi), \quad (7.1)$$

onde as funções são:

$$\begin{aligned} \Psi(\boldsymbol{\mu}, \phi) &= 2(-\phi)^{\frac{1}{2}} \left\{ (2\boldsymbol{\mu} - \phi) - (-\phi)^{\frac{1}{2}} \right\}^{-1} + \frac{1}{2} \log 2 \\ &\quad - \log \left\{ (2\boldsymbol{\mu} - \phi)^{\frac{1}{2}} - (-\phi)^{\frac{1}{2}} \right\} - \boldsymbol{\mu} \left\{ (2\boldsymbol{\mu} - \phi)^{\frac{1}{2}} - (-\phi)^{\frac{1}{2}} \right\}^{-1}, \end{aligned}$$

com

$$T(y) = \frac{1}{y} \quad \text{e} \quad \Psi^{(1,0)}(\boldsymbol{\mu}, \phi) = -\frac{1}{(-2\boldsymbol{\mu} - \phi)^{\frac{1}{2}} + (-\phi)^{\frac{1}{2}}}$$

### 7.1.2 Procedimento da geração do modelo

A variável resposta segue uma distribuição passeio aleatório e as covariáveis independentes geradas de uma distribuição uniforme. As duas componentes sistemáticas têm matrizes excluídas, ou seja, para o submodelo média a matriz  $X$  é composta de três covariáveis e para a componente sistemática do submodelo dispersão a matriz  $S$  com duas covariáveis.

Os verdadeiros valores dos parâmetros são:

- Submodelo média

$$\beta_1 = 1, \quad \beta_2 = 2, \quad \text{e} \quad \beta_3 = 3,$$

- Submodelo dispersão

$$\gamma_1 = 0.1 \quad \text{e} \quad \gamma_2 = 1.2.$$

Para a obtenção da variável resposta com distribuição passeio aleatório, foi necessário primeiro gerar uma distribuição inversa gaussiana, conforme sugerem Cordeiro e Botter (2001), ou seja,  $X \sim IG(\theta, \delta)$ , sendo que os parâmetros  $\theta$  e  $\delta$  são considerados como média e dispersão.

- Se  $X \sim IG(\theta, \delta)$  e fazendo  $Y = \frac{1}{X}$ , então a variável aleatória  $Y$  tem distribuição passeio aleatório, isto é,  $Y \sim PA(\boldsymbol{\mu}, \phi)$  com parâmetros média e dispersão dados por

$$\boldsymbol{\mu} = \frac{1}{\theta} + \frac{1}{\delta} \quad \text{e} \quad \phi = -\frac{\delta}{(2\theta^2)},$$

isto é, estes são função da média e dispersão da distribuição inversa gaussiana, que são substituídos na expressão (7.1)

## Componente sistemática

Para melhor entendimento, usou-se a notação  $\mu_{IG}$  e  $\phi_{IG}$  para média e dispersão para a distribuição inversa gaussiana no lugar de  $\theta$  e  $\delta$ , respectivamente.

i) Para os parâmetros do submodelo média

$$\eta = g(\boldsymbol{\mu}_{IG}) = \exp(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3),$$

ii) Para os parâmetros do submodelo dispersão

$$\tau = \log(\phi_{IG}) = \exp(\gamma_1 s_1 + \gamma_2 s_2).$$

Substituindo os parâmetros da distribuição inversa gaussiana no passeio aleatório, tem-se que

$$\mu = \frac{1}{\exp(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3)} + \frac{1}{\exp(\exp(\gamma_1 s_1 + \gamma_2 s_2))}$$

e

$$\phi = -\frac{\exp(\exp(\gamma_1 s_1 + \gamma_2 s_2))}{(2 \exp(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3))^2}.$$

## Estimativas de máxima verossimilhança

As estimativas foram obtidas no aplicativo SAS, que dispõe de um procedimento no qual a verossimilhança é maximizada usando-se o método de Newton Raphson. Para isto basta fornecer uma tentativa inicial e a equação da log-verossimilhança do modelo.

Para todos os diferentes tamanhos de amostras foram criados arquivos contendo as médias das estimativas de cada parâmetro das componentes sistemáticas da média e dispersão com seus respectivos desvios-padrão.

### 7.1.3 Obtenção dos vieses

Para a obtenção dos valores dos vieses das estimativas do modelo passeio aleatório foi necessário determinar as primeiras e segundas derivadas das matrizes modelos  $X$  e  $S$  e as derivadas até a terceira ordem da função  $\Psi(\mu, \phi)$ .

As matrizes do submodelo da média com as primeiras e segundas derivadas tomam os seguintes valores :

$$\tilde{X} = \frac{\partial f_1(X; \beta)}{\partial \beta_r} = [x_1 \mu_{IG}, x_2 \mu_{IG}, x_3 \mu_{IG}]_{3 \times n}$$

e

$$\tilde{X}_i = \frac{\partial^2 f_1(X; \beta)_i}{\partial \beta_r^2} = \begin{bmatrix} x_1^2 \mu_{IG} & x_1 x_2 \mu_{IG} & x_1 x_3 \mu_{IG} \\ x_2 x_1 \mu_{IG} & x_2^2 \mu_{IG} & x_2 x_3 \mu_{IG} \\ x_3 x_1 \mu_{IG} & x_3 x_2 \mu_{IG} & x_3^2 \mu_{IG} \end{bmatrix}_{n \text{ matrizes } 3 \times 3},$$

onde  $\mu_{IG} = \exp(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3)$ ,  $D_\beta = \text{diag}(d_1, \dots, d_n)$  e  $d_i = \text{tr}((\tilde{X}_i (\tilde{X}^T \Psi^{(2,0)} M_1^2 \tilde{X}))^{-1})$ .

Para obter as matrizes  $M_1$  e  $M_2$  foi necessário utilizar a regra da cadeia e, como se adotou a função de ligação identidade para o submodelo média, foram obtidas as expressões :

$$M_1 = \frac{\partial \mu_{PA}}{\partial \eta} = \frac{\partial \mu_{PA}}{\partial \mu_{IG}} \cdot \frac{\partial \mu_{IG}}{\partial \eta} = -\frac{1}{\mu_{IG}^2}$$

$$M_2 = \frac{\partial^2 \mu_{PA}}{\partial \eta^2} = \frac{\partial^2 \mu_{PA}}{\partial \eta \partial \mu_{IG}} \cdot \frac{\partial \mu_{IG}}{\partial \eta} + \frac{\partial^2 \mu_{IG}}{\partial \eta^2} \cdot \frac{\partial \mu_{PA}}{\partial \mu_{IG}} = \frac{2}{\mu_{IG}^3}$$

As matrizes do submodelo de dispersão com as primeiras e segundas derivadas tomam os seguintes valores:

$$\tilde{S} = \frac{\partial f_2(S; \gamma)}{\partial \gamma_R} = [s_1 \exp(\exp(\gamma_1 s_1 + \gamma_2 s_2)), s_2 \exp(\exp(\gamma_1 s_1 + \gamma_2 s_2))]_{2 \times n}$$

e

$$\tilde{S}_i = \frac{\partial^2 f_2(S; \gamma)_i}{\partial \gamma_R^2} = \begin{bmatrix} s_1^2 \phi_{IG} & s_1 s_2 \phi_{IG} \\ s_1 s_2 \phi_{IG} & s_2^2 \phi_{IG} \end{bmatrix}_{n \text{ matrizes } 2 \times 2}$$

sendo que  $\Phi_{IG} = \exp(\exp(\gamma_1 s_1 + \gamma_2 s_2))$  e  $D_\gamma = \text{diag}(d_1, \dots, d_n)$ , com  $d_i = \text{tr}((\tilde{S}_i (\tilde{S}^T \Psi^{(0,2)} \Phi_1^2 \tilde{S}))^{-1})$ .

Para este submodelo de dispersão a função de ligação foi a *log* tendo-se então:

$$\Phi_1 = \frac{\partial \phi_{PA}}{\partial \tau} = \frac{\partial \phi_{PA}}{\partial \phi_{IG}} \frac{\partial \phi_{IG}}{\partial \tau} = -\frac{\phi_{IG}}{2\mu_{IG}^2}$$

$$\Phi_2 = \frac{\partial^2 \phi}{\partial \tau^2} = -\frac{\phi_{IG}}{2\mu_{IG}^2}$$

Para o cálculo das derivadas das matrizes  $\tilde{X}$ ,  $\tilde{X}$ ,  $\tilde{S}$ ,  $\tilde{S}$  e da função  $\Psi(\mu, \phi)$  foi utilizado o aplicativo matemático *MAPLE*.

As derivadas da função  $\Psi(\mu, \phi)$  estão apresentadas a seguir:

$$\begin{aligned}
\Psi^{(2,0)}(\mu, \phi) &= \frac{2}{(\sqrt{2\mu - \phi} - \sqrt{-\phi})^3 \sqrt{2\mu - \phi}} \\
\Psi^{(2,1)}(\mu, \phi) &= \frac{2(-4\phi\sqrt{2\mu - \phi} + 4\phi\sqrt{-\phi} + 3\mu\sqrt{2\mu - \phi} - 7\mu\sqrt{-\phi})}{\sqrt{2\mu - \phi}(-2\mu + \phi)\sqrt{-\phi}(-\sqrt{2\mu - \phi} + \sqrt{-\phi})^5} \\
\Psi^{(0,2)}(\mu, \phi) &= \frac{-2\phi\sqrt{2\mu - \phi} + 2\phi\sqrt{-\phi} + \mu\sqrt{2\mu - \phi} - 3\sqrt{-\phi}\mu}{\phi(-\sqrt{2\mu - \phi} + \sqrt{-\phi})^3 \sqrt{-\phi}\sqrt{2\mu - \phi}} \\
\Psi^{(0,3)}(\mu, \phi) &= \frac{-2(4\phi^4 - 18\phi^3\mu + 4\phi^3\sqrt{2\mu - \phi}\sqrt{-\phi} - 14\sqrt{-\phi}\mu\phi^2\sqrt{2\mu - \phi} \\
&\quad + 26\phi^2\mu^2 - 12\mu^3\phi + 14\sqrt{-\phi}\mu^2\phi\sqrt{2\mu - \phi} - 3\mu^3\sqrt{2\mu - \phi}\sqrt{-\phi})/}{\left(\phi^3(-2\mu + \phi)^2(-\sqrt{2\mu - \phi} + \sqrt{-\phi})^4\right)}
\end{aligned}$$

Portanto, com as funções e as derivadas calculadas, basta substituí-las nas expressões dos cálculos dos vieses dos estimadores de  $O(n^{-1})$ , isto é,  $B(\hat{\beta})$  e  $B(\hat{\gamma})$  dos MNLGS.

A seguir é apresentado o fluxograma da simulação, que permite uma visualização do procedimento adotado.



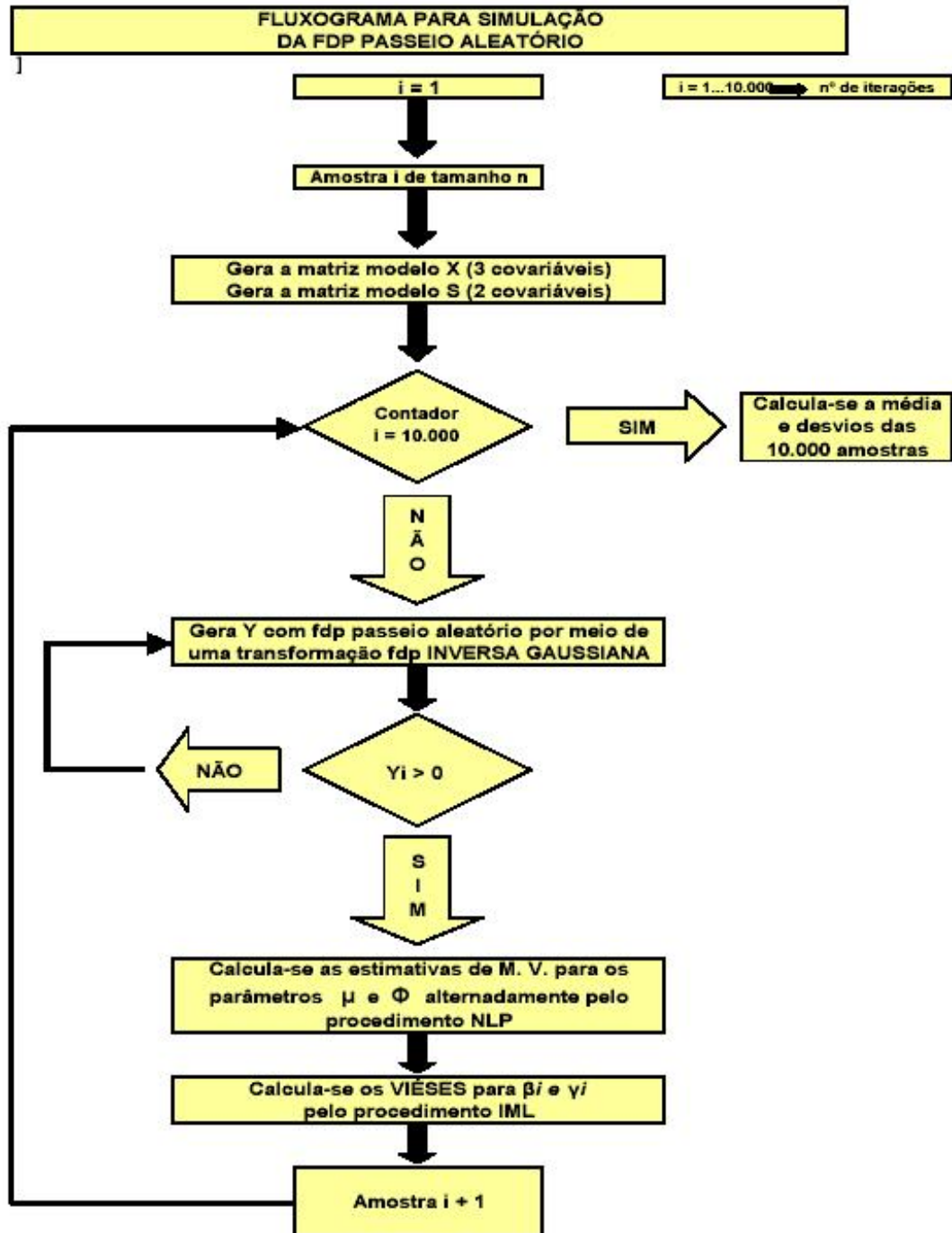


Figura 7.1: Procedimento da simulação para obtenção das estimativas pela distribuição passeio aleatório e seus respectivos vieses de  $O(n^{-1})$

Desse modo, obtêm-se as estimativas de  $\hat{\beta}$  e  $\hat{\gamma}$  iterativamente. Os valores iniciais de  $\beta_0$  e  $\gamma_0$  são dados na primeira iteração. Esse ciclo é repetido até a convergência. O processo converge rapidamente para os verdadeiros valores dos parâmetros do modelo adotado. Com as estimativas obtidas, calcula-se os vieses no procedure IML.

O processo de simulação para o modelo passeio aleatório e o cálculo dos vieses foi implementado no SAS, e está disponível no Apêndice E.

### 7.1.4 Resultados

Os resultados das simulações da distribuição passeio aleatório para as amostras de tamanho 15, 30 e 80 estão apresentados na Tabela 7.1 com as estimativas de  $O(n^{-1})$  e  $O(n^{-2})$  dos parâmetros do submodelo da média e da dispersão com seus respectivos desvios-padrão entre parênteses.

Tabela 7.1: Estimativas de  $O(n^{-1})$  e  $O(n^{-2})$  da distribuição passeio aleatório

$n$	estimadores	$\beta_1 = 1$	$\beta_2 = 2$	$\beta_3 = 3$	$\gamma_1 = 0.1$	$\gamma_2 = 1.2$
15	$O(n^{-1})$	0.751003 (1.554807)	2.143848 (1.201387)	2.612902 (1.990701)	0.079744 (0.690793)	1.247107 (0.531046)
	$O(n^{-2})$	0.980931 (1.300309)	1.977587 (1.060599)	3.00564 (1.967098)	0.088137 (0.66819)	1.247424 (0.517579)
30	$O(n^{-1})$	0.885744 (0.689121)	2.058922 (0.82072)	2.824105 (0.787896)	0.091284 (0.463308)	1.228039 (0.328041)
	$O(n^{-2})$	0.953472 (0.539257)	2.014259 (0.701625)	2.946344 (0.580152)	0.092265 (0.454755)	1.227805 (0.316279)
80	$O(n^{-1})$	0.989479 (0.377063)	1.985778 (0.421254)	2.953042 (0.474062)	0.100426 (0.219949)	1.209943 (0.173201)
	$O(n^{-2})$	0.991624 (0.367242)	1.991859 (0.405744)	2.965836 (0.367248)	0.100125 (0.21844)	1.209746 (0.164866)

Pode-se observar que as estimativas de  $O(n^{-2})$  dos parâmetros do submodelo da média e dispersão ficaram mais próximos dos verdadeiros parâmetros e com desvios-padrão menores que os das estimativas de  $O(n^{-1})$ .

A seguir alguns gráficos são apresentados para melhor visualização do comportamento das estimativas de  $O(n^{-1})$  e  $O(n^{-2})$  com relação ao tamanho das amostras nos submodelos média e dispersão.

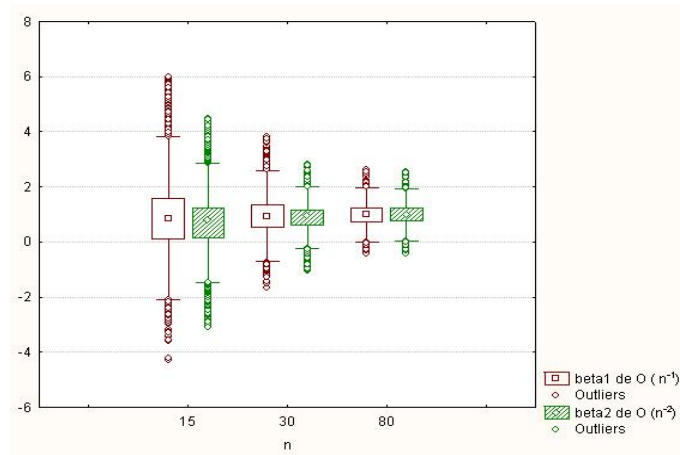


Figura 7.2: Estimativas de  $O(n^{-1})$  e  $O(n^{-2})$  para  $\beta_1$ .

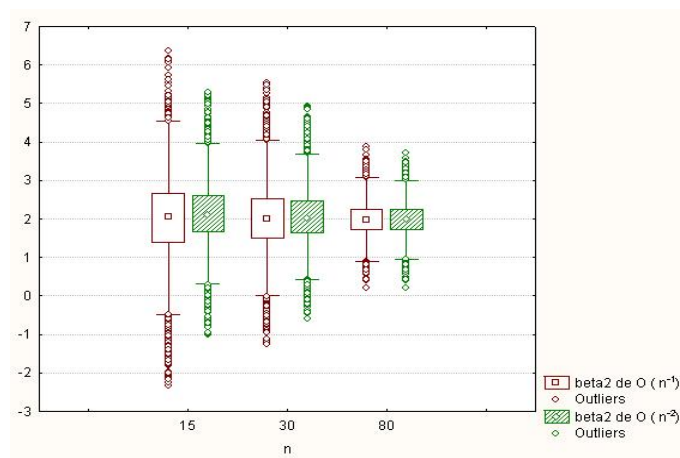


Figura 7.3: Estimativas de  $O(n^{-1})$  e  $O(n^{-2})$  para  $\beta_2$ .

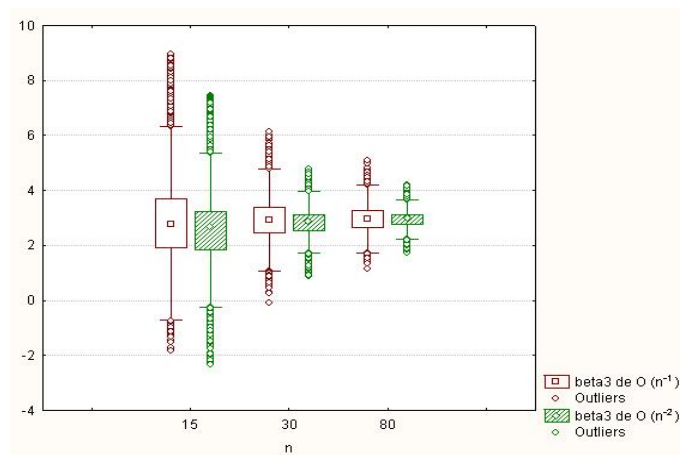


Figura 7.4: Estimativas de  $O(n^{-1})$  e  $O(n^{-2})$  para  $\beta_3$ .

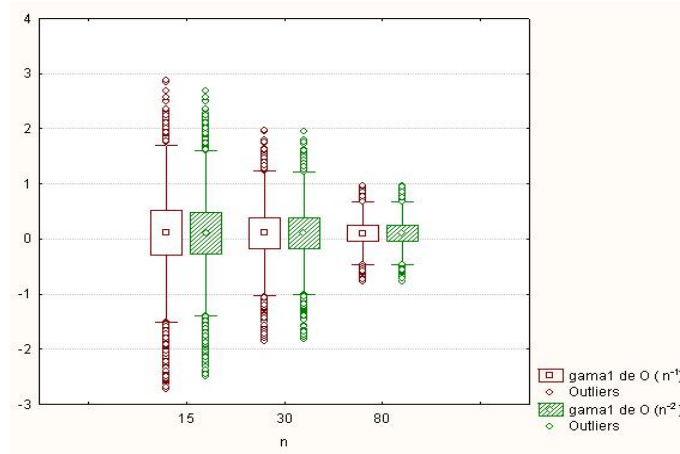


Figura 7.5: Estimativas de  $O(n^{-1})$  e  $O(n^{-2})$  para  $\gamma_1$ .

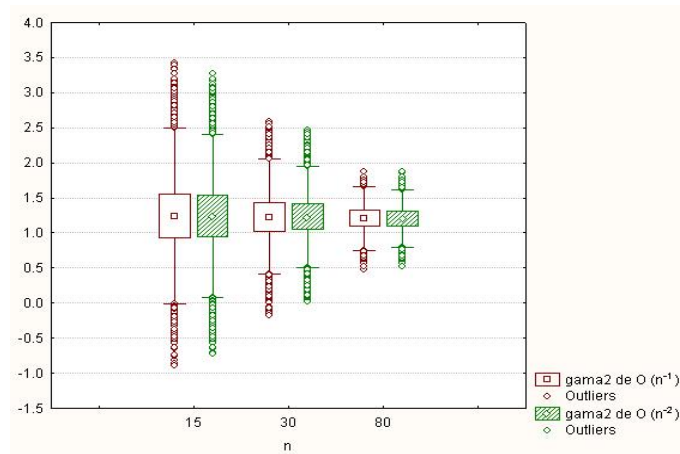


Figura 7.6: Estimativas de  $O(n^{-1})$  e  $O(n^{-2})$  para  $\gamma_2$ .

As figuras evidenciam que as estimativas dos parâmetros do submodelo da média e dispersão de  $O(n^{-1})$ , comparadas com as de  $O(n^{-2})$ , têm diferenças consideráveis para as amostras de tamanho 15 e, quando as amostras aumentam, as diferenças entre as estimativas não corrigidas e corrigidas vão ficando cada vez mais próximas, ou seja, corroborando a teoria de que a correção das estimativas é indicada para amostras pequenas.

A magnitude dos vieses para os MNLGS estão de acordo com Box (1971) para os modelos não lineares.

## 7.2 Aplicação dos MNLGS corrigidos a dados reais

### 7.2.1 Aplicação 1: Modelando os erros cometidos num treinamento de aprendizagem e avaliando as estimativas de $O(n^{-2})$

Para testar a aplicação das estimativas de  $O(n^{-2})$  dos MNLGS, empregaram-se os dados experimentais obtidos durante um teste de aprendizagem e memória espaciais aplicado em

ratos portadores de lesão cerebral isquêmica, ou seja, pela falta de sangue no cérebro. Esse teste utiliza o modelo do labirinto radial de 8-braços conforme Figura 7.7, usado pelo Laboratório da Isquemia Cerebral e Neuroproteção, do Departamento de Farmácia e Farmacologia da Universidade Estadual de Maringá.

## **Introdução**

A necessidade de desenvolvimento de estratégias farmacológicas para o tratamento e/ou prevenção de doenças neurodegenerativas agudas (lesão cerebral de origem isquêmica, infecciosa, traumática, etc) ou crônicas (doença de Parkinson, doença de Alzheimer, entre outras) é urgente. A população idosa do globo cresce em escala geométrica e essas doenças afetam fundamentalmente pessoas da terceira idade, ou seja, acima de 65 anos. Modelos animais que reproduzem tanto as doenças quanto as seqüelas das mesmas nos seres humanos representam o ponto de partida para a investigação dessas questões.

Na experimentação pré-clínica, o rato é o animal mais freqüentemente utilizado, inclusive como modelo para as doenças neurodegenerativas e suas sintomatologias. Disfunções de aprendizagem e memória podem ser medidas nesses modelos animais mediante o uso de testes e/ou modelos comportamentais, tais como os chamados labirintos. Estes medem, de modo geral, a capacidade do animal para executar uma determinada tarefa que demanda “atenção”, “raciocínio”, “motivação”, ou seja, funções cognitivas (aprendizagem e memória). Além disso, o teste de labirinto é suficientemente sensível para detectar os efeitos das drogas no rato, isto é, na eficácia terapêutica em facilitar ou dificultar a capacidade de aprendizagem ou dos efeitos causados por lesões cerebrais, além dos efeitos do envelhecimento, entre outros.

O labirinto radial de oito braços (Figura 7.7) apetitivo tem sido um dos modelos de aprendizagem e memória espacial usados mais freqüentemente para se estudar os mecanismos neurobiológicos das funções de aprendizagem e memória. Caracteristicamente, o teste do labirinto radial requer o treinamento repetitivo do animal (usualmente em dias consecutivos), fornecendo assim uma curva de crescimento do desempenho cognitivo.

## **O labirinto radial aversivo**

No experimento em questão foi utilizado um labirinto radial de oito braços aversivo conforme a Figura 7.7, e com sistema de confinamento.

Os braços (55 cm x 15 cm) projetam-se a partir de uma plataforma poligonal de 16 lados (arena central com 71 cm na transversal), estando os braços dispostos alternadamente em relação às faces do polígono. A comunicação dos braços com a arena central tem trânsito livre entre essas partes do labirinto. Foram feitas portas tipo guilhotina (em acrílico transparente), controladas pelo experimentador. Ao final de cada braço, uma abertura (9 cm de diâmetro) permite o acesso do animal a uma caixa escura (23 x 11 x 9,5 cm) localizada logo abaixo de

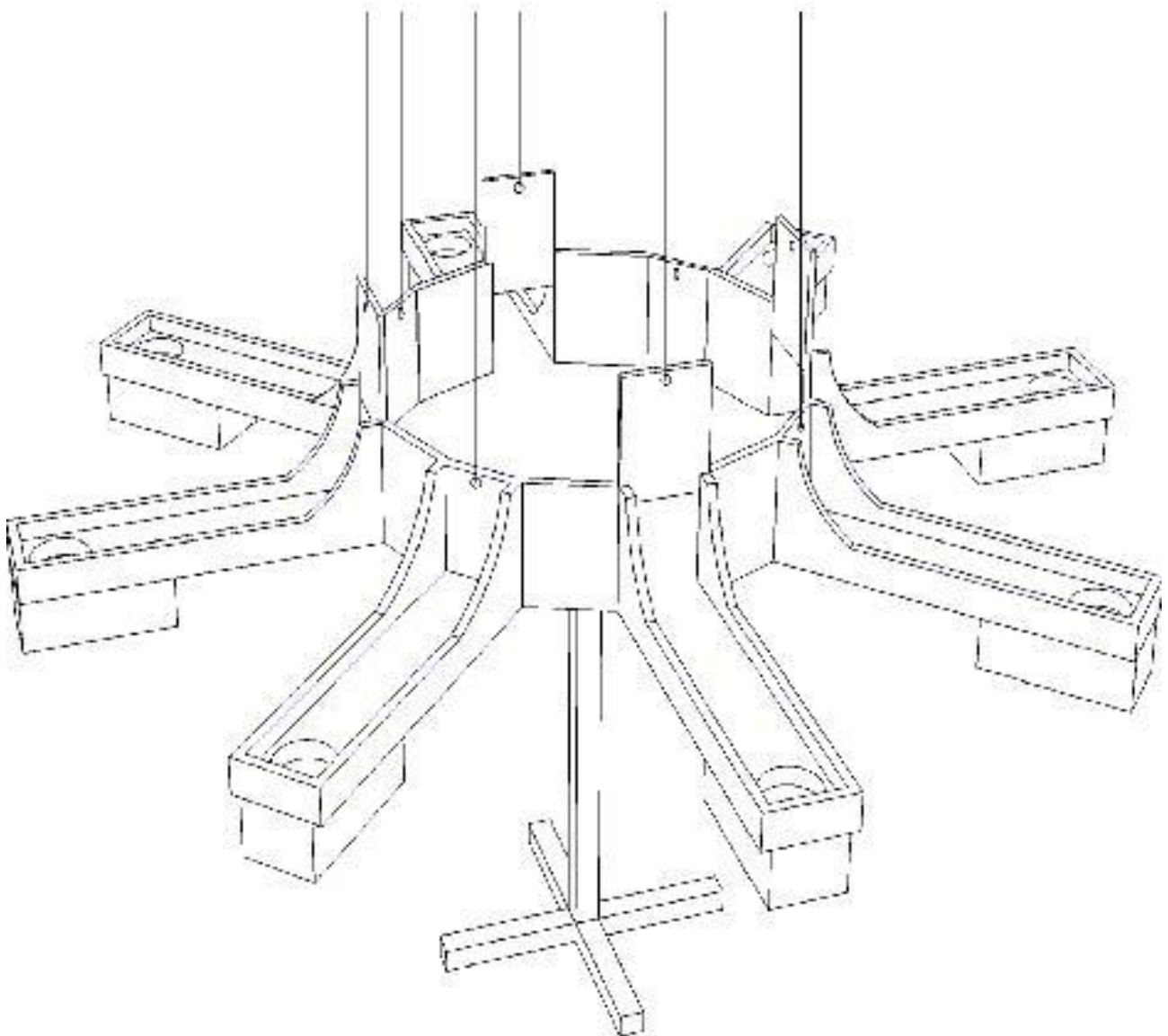


Figura 7.7: Representação esquemática do labirinto radial de oito braços

cada orifício, a qual pode ser inserida e removida como uma gaveta abaixo, servindo como refúgio para o rato em relação às áreas iluminadas do labirinto. Dentre os oito braços, somente um contém um refúgio verdadeiro, sendo que nos demais braços os esconderijos são de fundo falso. Nas bordas de cada braço existe uma parede também de acrílico (2,5 cm), que serve para prevenir uma eventual queda do animal. O labirinto encontra-se elevado a 90 cm do solo, podendo ser livremente girado em qualquer sentido. Várias pistas tridimensionais extra-labirinto estão dispostas na sala (porta fechada, janela e alguns objetos tridimensionais). Um pequeno circulador de ar, localizado no chão, gera um ruído de fundo constante na sala de experimentação. Duas lâmpadas incandescentes (200 W cada), mais um par de lâmpadas

fluorescentes (40 W cada), estão fixas no teto, a uma altura de 1,80 cm da plataforma do labirinto. Uma câmera de vídeo posicionada na porta de entrada da sala é usada para que o experimentador tenha controle de acesso do animal aos braços e para gravar o comportamento, possibilitando posteriormente uma análise e quantificação da performance.

## **Delineamento do experimento**

### **i) Animais**

Foram utilizados 51 ratos Wistar machos, adultos, pesando entre 280 e 300 gramas (aproximadamente 3 meses de idade). Eles permaneceram em condições-padrão de alojamento, com ciclo de luz claro/escuro, temperatura controlada ( $22 \pm 1^\circ\text{C}$ ) e receberam água e ração à vontade.

### **ii) Isquemia**

Vinte e cinco animais foram submetidos à isquemia cerebral global e transitória conforme o modelo de oclusão dos quatro vasos (Pulsinelli & Brierley, 1979). Os outros 26 animais, designados como grupo falso isquêmico (não-lesionado), também foram submetidos ao mesmo procedimento cirúrgico, porém sem a oclusão das artérias vertebrais e carótidas. Esse procedimento resulta numa redução de 95% do fluxo sanguíneo cerebral. Os procedimentos operacionais para indução da isquemia seguiram a rotina do Laboratório de Isquemia Cerebral e Neuroproteção (Giordani *et al.*, 1999; Paganelli *et al.*, 2004; Benetolli *et al.*, 2004) e também seguiram os “Princípios básicos para a utilização animal”, conforme aprovado pelo Comitê de Ética em Experimentação Animal, da Universidade Estadual de Maringá.

**Procedimento do teste** No teste do labirinto radial avaliou-se a capacidade de aprendizagem e memória espacial no rato. O desempenho comportamental do animal foi expresso da seguinte forma:

- i)** latência para encontrar a recompensa (esconderijo);
- ii)** número de erros de memória de referência (memória a longo prazo), e
- iii)** número de erros de memória operacional (memória a curto prazo).

A cada dia de teste foram dadas três tentativas ao animal para encontrar o esconderijo. Dentro de cada tentativa, um erro de referência é contado toda vez que o animal entrar num braço falso do esconderijo, pela primeira vez. Um erro operacional é registrado toda vez que o rato torna a entrar num braço previamente visitado. Por essa definição, o animal utiliza-se da memória de referência para se lembrar de que apenas um dos oito braços contém a recompensa (esconderijo), e que a sua relativa localização espacial se manteve inalterada ao longo das várias sessões (dias) de teste, indicando, portanto, uma memória de longo-prazo. Diferentemente, a memória operacional implica que o animal deve se recordar de qualquer dos outros sete braços (sem recompensa), os quais foram previamente visitados durante uma mesma tentativa, definida

por memória de curto-prazo.

Cerca de vinte dias após a indução da isquemia cerebral, os animais são testados no labirinto radial aversivo. Eles recebem treinamento durante 15 dias consecutivos, com três tentativas por dia. Os resultados são expressos como a média entre a latência e o número de erros em função do tempo de treinamento. Os 15 dias de treino são representados em cinco blocos de três dias cada, conforme Tabela F1 e F2 no Apêndice F.

### Modelo não-linear

O delineamento descrito é do tipo medidas repetidas, ou seja, a contagem do número de erros é observada no mesmo rato em tempos diferentes. Esses dados já foram analisados pela metodologia de modelos lineares mistos, utilizando uma matriz de covariância não estruturada. No entanto, segundo Dobson (2001), uma análise alternativa é considerar os modelos da família exponencial, como, por exemplo, um modelo de regressão Poisson.

Para visualizar o comportamento da variável resposta (número de erros cometidos pelos ratos), alguns gráficos foram executados com os ratos lesionados e não-lesionados ao longo do tempo, como, por exemplo, os diagramas de dispersão com o número total de erros por tipo de tratamento (lesionado e não-lesionado) e o comportamento de cada animal ao longo do tempo.

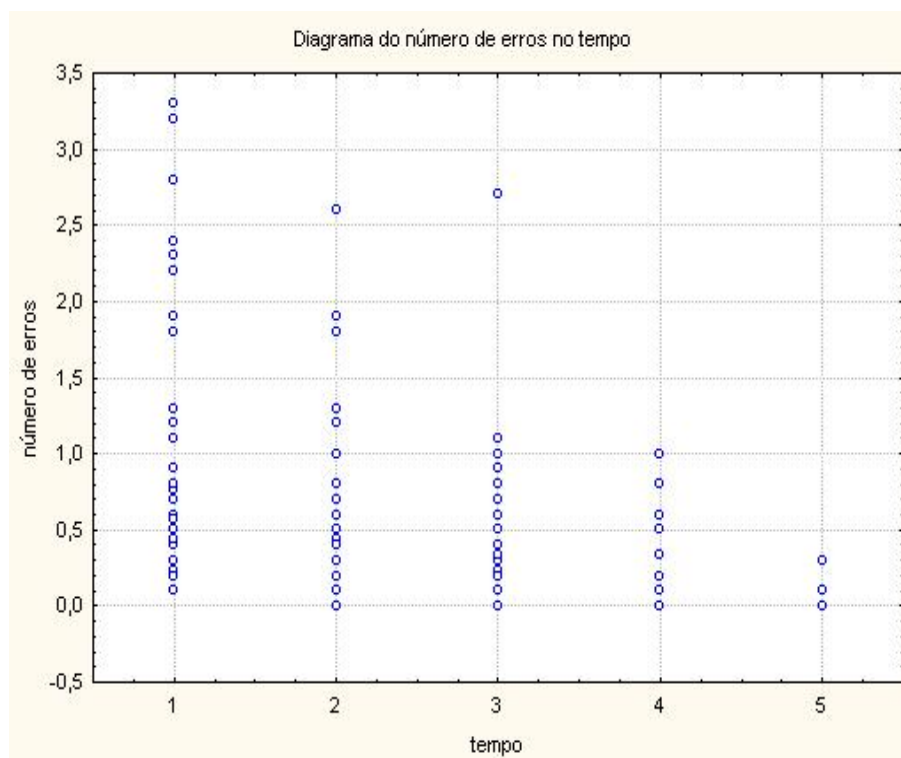


Figura 7.8: Erros cometidos pelos ratos no tempo.



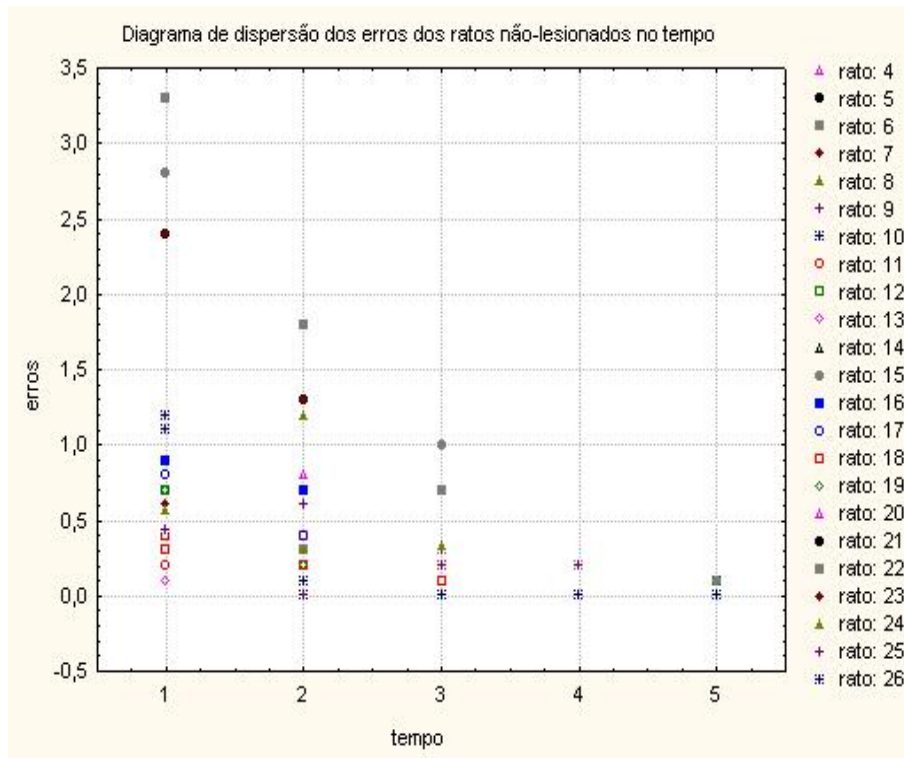


Figura 7.9: Erros cometidos pelos ratos não-lesionados no tempo.

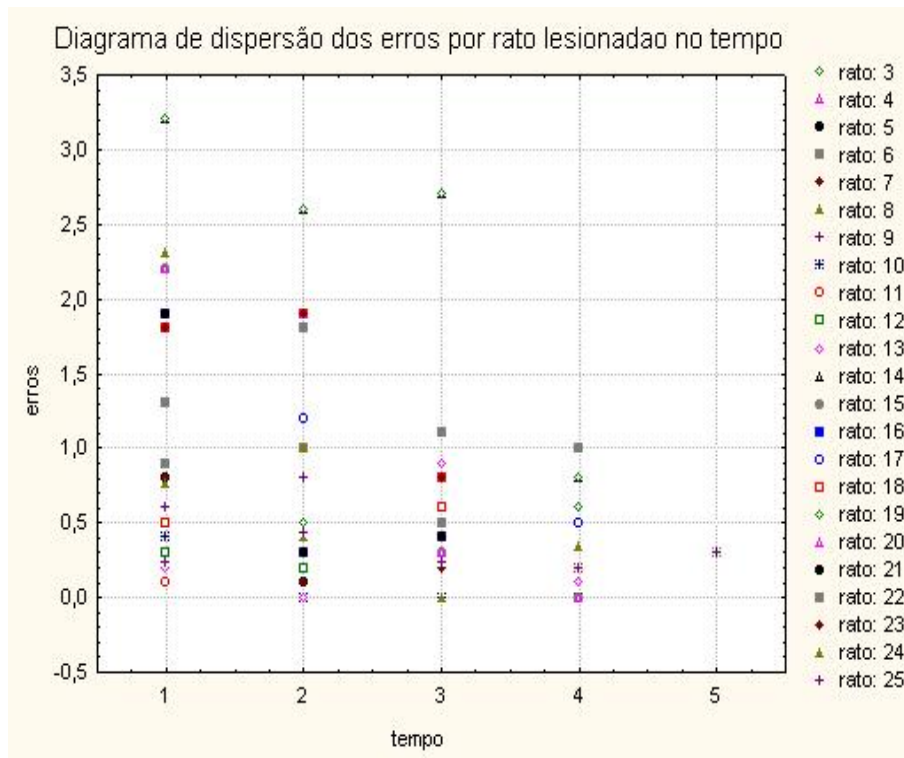


Figura 7.10: Erros cometidos pelos ratos lesionados no tempo.

Observa-se pelos gráficos que a variável-resposta “número médio de erros cometidos pelos ratos” tem um comportamento não-linear, indicando que esses dados podem ser ajustados por meio de um modelo de regressão Poisson.

Na seqüência, dois modelos de regressão foram ajustados. O primeiro ajuste por um modelo de regressão Poisson considerando o parâmetro de dispersão constante e o outro modelo pela regressão Poisson dupla seguida do cálculo dos vieses de primeira ordem e as estimativas de  $O(n^{-2})$  de máxima verossimilhança dos parâmetros dos submodelos da média e dispersão são apresentados.

### Modelo de regressão Poisson

Para esse modelo, a variável resposta é o número médio de erros cometidos pelos ratos e as covariáveis: número de ratos lesionados e número de ratos não lesionados e o tempo (dias). Para essa regressão, modelam-se somente os parâmetros do submodelo média e parâmetro de dispersão  $\phi$  fixo igual a 1.

$$\begin{aligned}\hat{\mu}_i &= f_1(X, \hat{\beta}) \\ &= \hat{\beta}_0 * lesionado_i + não-lesionado_i * (\exp(\hat{\beta}_1 * tempo_i)) + \\ &\quad lesionado_i * (\exp(\hat{\beta}_2 * tempo_i)),\end{aligned}\tag{7.2}$$

Após algumas discussões com o pesquisador da área, que afirma, que a contribuição na média do efeito devido a lesão tem comportamento linear, enquanto que, o efeito devido ao tempo tem comportamento exponencial (não-linear), os quais são confirmados pelos gráficos 7.8 , 7.9 e 7.10, e que, portanto, justifica-se adoção do seguinte modelo: com  $i = 1, \dots, 51$  .

A função log-verossimilhança dos modelos não-lineares generalizados superdispersados depende das funções  $\Psi(\mu, \phi)$  e  $T(y)$  , que nesse caso tem a forma abaixo

$$\Psi(\mu, 1) = \mu \log \mu \quad \text{e} \quad T(y) = y \log y.$$

$$\ell(\beta) = (erros_i - \hat{\mu}_i) * (1 + \log \hat{\mu}_i) - erros_i * \log(erros_i) + \hat{\mu}_i * \log(\hat{\mu}_i) + \frac{1}{2}.$$

Para calcular os valores das estimativas dos parâmetros do submodelo média utilizou-se o procedimento *NLMIXED* do *SAS*.

### Modelo de regressão Poisson dupla

A componente sistemática da média e a função  $T(y)$  continuam sendo as mesmas usadas na regressão Poisson; no entanto, agora o parâmetro de dispersão é estimado pela componente sistemática dada por

$$\hat{\phi}_i = f_2(S, \hat{\gamma}) = \exp(\hat{\gamma}_0 + \hat{\gamma}_1 * tempo_i),\tag{7.3}$$

a função  $\Psi(\mu, \phi)$  passa a ser

$$\Psi(\mu, \phi) = \mu\phi \log \mu + \frac{1}{2} \log \phi$$

e a log-verossimilhança

$$\begin{aligned} \ell(\beta, \gamma) = & (\text{erros}_i - \hat{\mu}_i) * \left( \hat{\phi}_i + \hat{\phi}_i \log \hat{\mu}_i \right) - \\ & \hat{\phi}_i * \text{erros}_i * \log(\text{erros}_i) + \hat{\mu}_i * \hat{\phi}_i * \log(\hat{\mu}_i) + \frac{1}{2} * \log \hat{\phi}_i. \end{aligned}$$

Como os parâmetros  $\beta$  e  $\gamma$  são ortogonais, as estimativas foram calculadas separadamente, com as tentativas iniciais encontradas pelo modelo não-linear normal no procedimento *NLMIXED* do aplicativo SAS.

### Cálculo dos vieses $O(n^{-1})$ para o modelo Poisson dupla

São necessárias para determinar os vieses de  $O(n^{-1})$  dos estimadores de máxima verossimilhança para o modelo Poisson dupla, as primeiras e segundas derivadas das matrizes  $X$  e  $S$  dos submodelos da média e dispersão, bem como as derivadas da função  $\Psi(\mu, \phi)$  e das funções de ligação das componentes sistemáticas.

**Vieses para os estimadores do submodelo média** A expressão para o cálculo do viés é dada por

$$\begin{aligned} B(\hat{\beta}) &= (\tilde{X}^T \Psi^{(2,0)} M_1^2 \tilde{X})^{-1} \tilde{X}^T \Psi^{(2,0)} M_1^2 \xi, \\ \xi_1 &= -\frac{1}{2} Z_{\beta d} M_1^{-1} M_2 \\ \xi_2 &= -\frac{1}{2} D_{\beta}. \end{aligned}$$

Utilizou-se a função de ligação identidade para  $\mu$  logo, as matrizes

$$M_1 = \frac{\partial \mu}{\partial \eta} = 1, \quad M_1^2 = \left( \frac{\partial \mu}{\partial \eta} \right)^2 = 1 \quad \text{e} \quad M_2 = \frac{\partial^2 \mu}{\partial \eta^2} = 0$$

e as primeiras e segundas derivadas da  $f_1$  são:

$$\tilde{X}_i = \frac{\partial f_1(X, \beta)_i}{\partial \beta_r} = \begin{bmatrix} \text{lesionado}_i & \text{não-lesionado}_i * \text{tempo}_i * \exp(\beta_1 * \text{tempo}_i) \\ \text{lesionado}_i * \text{tempo}_i * \exp(\beta_2 * \text{tempo}_i) \end{bmatrix}_{1 \times 3},$$

$$\begin{aligned} \tilde{X}_i &= \frac{\partial^2 f_2(X, \beta)_i}{\partial \beta^T \partial \beta} \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & \text{n\~{a}o lesionado}_i * \text{tempo}_i^2 * \exp(\beta_1 * \text{tempo}_i) & 0 \\ 0 & 0 & \text{lesionado}_i * \text{tempo}_i^2 * \exp(\beta_2 * \text{tempo}_i) \end{bmatrix}_{n \text{ matrizes } 3 \times 3}, \end{aligned}$$

e as matrizes

$$D_\beta = \text{diag}(d_i, \dots, d_n) \text{ com cada } d_i = \text{tr}((\tilde{X}_i (\tilde{X}^T \Psi^{(2,0)} M_1^2 \tilde{X}))^{-1}), \text{ ou seja,}$$

$$\begin{aligned} d_i &= \text{tr}(\text{n\~{a}o-lesionado} * \text{tempo}_i^2 (\exp(\beta_1 * \text{tempo}_i) * K_{\beta\beta}^{-1} (2, 2) \\ &\quad + \text{lesionado} * \text{tempo}_i^2 * (\exp(\beta_2 * \text{tempo}_i) * K_{\beta\beta}^{-1} (3, 3))), \end{aligned}$$

lembrando que  $K_{\beta\beta}^{-1}(i, i)$  s\~{a}o os respectivos elementos da inversa da matriz de informa\~{c}o.

As derivadas da fun\~{c}o  $\Psi(\mu, \phi)$  com rela\~{c}o a m\~{e}dia s\~{a}o:

$$\begin{aligned} \Psi^{(1,0)}(\mu, \phi) &= \phi \log \mu + \phi \\ \Psi^{(2,0)}(\mu, \phi) &= \phi \frac{1}{\mu}. \end{aligned}$$

**Vieses para os estimadores do submodelo dispers\~{a}o** A express\~{a}o para o c\~{a}lculo \~{e} dada por

$$B(\hat{\gamma}) = (\tilde{S}^T \Psi^{(0,2)} \Phi_1^2 \tilde{S})^{-1} \tilde{S}^T (\Psi^{(0,2)} \phi_1^2) \delta$$

sendo  $\delta = \delta_1 + \delta_2$

$$\begin{aligned} \delta_1 &= \left\{ -\frac{1}{2} (\Psi^{(0,2)} \Phi_1^2)^{-1} [Z_{\gamma d} (\Psi^{(0,3)} \Phi_1^3 + \Psi^{(0,2)} \Phi_1 \Phi_2) + Z_{\beta d} (\Psi^{(2,1)} \Phi_1 M_1^2)] \right\} \\ \delta_2 &= -\frac{1}{2} D_\gamma, \end{aligned}$$

onde  $D_\gamma = \text{diag}(d_i, \dots, d_n)$  com  $d_i = \text{tr}((\tilde{S}_i (\tilde{S}^T \Psi^{(0,2)} \Phi_1^2 \tilde{S}))^{-1})$ .

Utilizou-se tamb\~{e}m a fun\~{c}o de liga\~{c}o identidade para a componente sistem\~{a}tica da dispers\~{a}o, ou seja,

$$\tau = \phi = \exp(\mathbf{S}\gamma) = \exp(\gamma_0 + \gamma_1 * \text{tempo}_i).$$

Então as derivadas são:

$$\Phi_1 = \frac{\partial \phi}{\partial \tau} = 1, \quad \Phi_1^2 = \left(\frac{\partial \phi}{\partial \tau}\right)^2 = 1 \quad e \quad \Phi_2 = \frac{\partial^2 \phi}{\partial \tau^2} = 0, \quad (7.4)$$

sendo que as primeiras e segundas derivadas da matriz  $S$  são:

$$\tilde{S} = \frac{\partial f_2(S; \gamma)}{\partial \gamma_R} = [\exp(\gamma_0 * \gamma_1 * tempo_i), \quad tempo_i * \exp(\gamma_0 * \gamma_1 * tempo_i)]_{2 \times 1}$$

e

$$\tilde{S}_i = \frac{\partial^2 f_2(S; \gamma)_i}{\partial \gamma_R^2} = \begin{bmatrix} \phi & tempo_i * \phi \\ tempo_i * \phi & tempo_i^2 * \phi \end{bmatrix}_{n \text{ matrizes } 2 \times 2}$$

$$d_i = tr(\exp(\gamma_0 * \gamma_1 * tempo_i) * K_{\gamma\gamma}^{-1}(1, 1) + 2 * (tempo_i * \exp(\gamma_0 * \gamma_1 * tempo_i)) * K_{\gamma\gamma}^{-1}(2, 1) + tempo_i^2 * (\exp(\gamma_0 * \gamma_1 * tempo_i)) K_{\gamma\gamma}^{-1}(3, 3))$$

As derivadas da função  $\Psi(\mu, \phi)$  com relação à dispersão são:

$$\begin{aligned} \Psi^{(0,1)}(\mu, \phi) &= \phi \log \mu + \frac{1}{2} \frac{1}{\phi} \\ \Psi^{(0,2)}(\mu, \phi) &= -\frac{1}{2\phi^2} \\ \Psi^{(2,1)}(\mu, \phi) &= \frac{1}{\mu} \\ \Psi^{(0,3)}(\mu, \phi) &= \frac{1}{\phi^3}. \end{aligned}$$

Com as derivadas calculadas das matrizes modelos  $X$  e  $S$  e da função  $\Psi(\mu, \phi)$ , basta substituir nas expressões dos vieses  $B(\hat{\beta}_r)$  e  $B(\hat{\gamma})$  e executar no procedimento *IML* do SAS.

O modelo não foi avaliado com outras funções de ligação para os submodelos da média e dispersão.

## Resultados

Estão apresentadas por meio de tabelas as estimativas dos parâmetros dos modelos não-lineares, bem como os vieses de  $O(n^{-1})$ .

### Ajuste do modelo de regressão Poisson

A Tabela 7.2 apresenta as estimativas de máxima verossimilhança para o modelo de regressão Poisson com parâmetro de dispersão igual a 1 e seus respectivos erros-padrão.

Tabela 7.2: Estimativas da média para o modelo Poisson

estimadores	estimativas	erros-padrão	Wald	<i>p</i> -valor
$\hat{\beta}_0$	0.2392	0.1526	1.57	0.1182
$\hat{\beta}_1$	-0.5164	0.0700	-7.38	< .0001
$\hat{\beta}_2$	-0.3762	0.1586	-2.37	0.0185

Pode-se observar que o estimador  $\hat{\beta}_0$  foi não significativo, ou seja, há evidências de que a média dos erros dos ratos lesionados e a dos não-lesionados inicialmente não diferem, no entanto, isso vai mudando ao longo do tempo.

Substituindo as estimativas no submodelo média dos erros, tem-se se que

$$\hat{\mu}_i = 0.2392 * lesionado_i + não-lesionado_i * (\exp(-0.5164 * tempo_i)) + lesionado_i * (\exp(-0.3762 * tempo_i))$$

Portanto, a estimativa  $\hat{\beta}_1 = -0.5164$  corresponde ao déficit de aprendizagem para os ratos não-lesionados e  $\hat{\beta}_2 = -0.3762$  para os ratos lesionados. Apesar que neste estudo não houve interesse em testar a significância de  $(\beta_1 - \beta_2)$ , a Figura 7.11 evidencia que o decréscimo dos erros para os ratos lesionados é menor.

O modelo foi de fácil convergência quando executado pelo procedimento *NLMIXED* do SAS.

### Ajuste do modelo de regressão Poisson dupla

Nesse caso, para ajustar a média dos erros, a dispersão não é constante. A Tabela 7.3 contém as estimativas dos parâmetros do submodelo da média com seus respectivos erros-padrão.

Tabela 7.3: Estimativas do submodelo média Poisson dupla

estimadores	estimativas	erros-padrão	Wald	<i>p</i> -valor
$\hat{\beta}_0$	0.1344	0.1129	1.19	0.2351
$\hat{\beta}_1$	-0.7333	0.0522	-14.05	< .0001
$\hat{\beta}_2$	-0.3650	0.1282	-2.85	0.0048

e substituindo na equação (7.2) da média tem-se :

$$\hat{\mu} = 0.1344 * lesionado_i + não lesionado_i * (\exp(-0.7333 * tempo_i)) + lesionado_i * (\exp(-0.3650 * tempo_i))$$

Observa-se que os valores das estimativas não foram muito diferentes do modelo Poisson, inclusive concordando no  $p$  - *valor* maior que 0.05 para  $\hat{\beta}_0$ ; no entanto, para os estimadores que foram significativos, seus erros-padrão foram bem menores.

Podem-se observar na Tabela 7.4 as estimativas para o submodelo dispersão com seus respectivos erros-padrão.

Tabela 7.4: Estimativas do submodelo dispersão Poisson dupla

estimadores	estimativas	erros-padrão	Wald	$p$ -valor
$\hat{\gamma}_0$	-0.8387	0.2368	-3.54	0.0005
$\hat{\gamma}_1$	0.5676	0.0732	7.75	< .0001

Assim, substituindo-se as estimativas pela expressão (7.3), obtém-se a equação

$$\hat{\phi}_i = \exp(-0.8387 + 0.5676 * tempo_i).$$

Também para esse modelo foi fácil a convergência, quando executado no procedimento *NLMIXED* do SAS. Para as estimativas do submodelo da média na 17<sup>a</sup> iteração e para as estimativas do submodelo dispersão na 10<sup>a</sup>. O programa executado para essas estimativas podem ser vistas no Apêndice F .

Para comparar dois ou mais modelos para o mesmo conjunto de dados, utilizou-se o critério de *AIC* ( critério de informação de Akaike), que penaliza o número de parâmetros na função da *log* -verossimilhança dada pela expressão

$$AIC = -2 \log \ell (\beta, \gamma) + 2n_{par} ,$$

onde  $n_{par}$  é o número de parâmetros do modelo ajustado. Sob essa definição, prefere-se o modelo com menor *AIC*.

Para o ajuste da Poisson obteve-se  $AIC = 141.323$  e para a Poisson dupla a estatística  $AIC = 34.57$ . Pode-se observar pelos valores das estatísticas *AIC* que o modelo Poisson dupla se ajusta melhor aos dados, pois quanto menor o valor do *AIC* melhor o modelo conforme Pinheiro e Bates (2000, p. 84)

A Figura 7.11 abaixo apresenta o submodelo da média ajustado pelo modelo Poisson dupla.

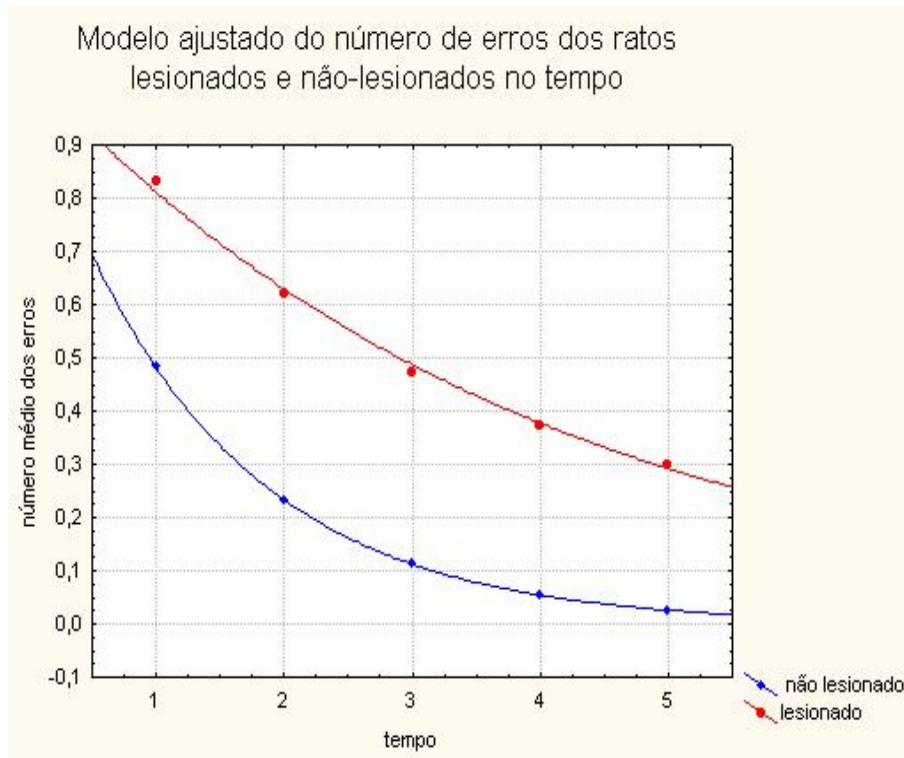


Figura 7.11: Modelo Poisson duplo ajustado

Pela representação gráfica dos modelos ajustados, tanto para os ratos lesionados como para os não-lesionados, pode-se afirmar que estes aprendem ao executar uma tarefa. No entanto, há uma declividade maior no número médio dos erros para os ratos não lesionados, isto é, estes aprendem numa velocidade maior ao longo do tempo.

Os valores dos vieses e as estimativas corrigidas podem ser vistas na Tabela 7.5 abaixo para cada estimativa dos parâmetros que compõem os submodelos para a média e dispersão.

Tabela 7.5: Vieses e estimativas corrigidas do modelo Poisson dupla

Valores dos vieses	estimativas corrigidas
$B(\hat{\beta}_0) = -0.022542$	$\tilde{\beta}_0 = 0.156955$
$B(\hat{\beta}_1) = -0.004613$	$\tilde{\beta}_1 = -0.728681$
$B(\hat{\beta}_2) = -0.002850$	$\tilde{\beta}_2 = -0.362185$
$B(\hat{\gamma}_0) = 0.007078$	$\tilde{\gamma}_0 = -0.362185$
$B(\hat{\gamma}_1) = -0.007346$	$\tilde{\gamma}_1 = 0.5749664$

## Conclusão

Os valores do vieses das estimativas quando, comparados com os seus respectivos erros-padrão, correspondem a um percentual de 20%, 9% e 2%, respectivamente, para as estimativas



do submodelo média, enquanto que para as estimativas do submodelo dispersão esses valores são de 3% e 10%.

Usualmente, nesse tipo de experimento, as amostras são pequenas e, além disso, é muito comum ocorrer, durante o experimento, o descarte de animais por morte ou por algum comportamento atípico. Para efeito ilustrativo, ajustou-se novamente o mesmo modelo com um sorteio dos elementos da amostra (10 ratos isquêmicos lesionados e 11 ratos não-lesionados com cinco repetições) e obtiveram-se os valores quando comparados com seus erros-padrão para o submodelo média de 62%, 25% e 7% e para a dispersão 76% e 119%. Logo, pode-se afirmar que, quanto menor o tamanho da amostra, maior a necessidade de se fazer correções.

### 7.2.2 Aplicação 2: Modelando a resistência do concreto e avaliando as estimativas de $O(n^{-2})$

Esse experimento foi estudado por Mattos (2004) para avaliar os efeitos de alguns parâmetros do processo produtivo do concreto de alto desempenho, moldado com a adição de cinza de casca de arroz, sobre sua resistência à compressão.

Segundo Mattos (2004), embora a resistência à compressão seja uma característica do tipo maior-é-melhor, na produção do concreto é realizado um estudo de dosagem para o produto atingir determinada resistência. O resultado considera, entre outros fatores, a heterogeneidade da mistura. Uma grande variabilidade resulta em um maior consumo de cimento e, portanto, elevação do custo do concreto. Além disso, no cálculo estrutural, considera-se que o concreto deve apresentar uma resistência mínima. Se existir grande variabilidade, é mais difícil garantir a ocorrência desse mínimo, podendo dar-se o colapso da estrutura. Para evitar tal possível colapso, os calculistas adotam um coeficiente de segurança (40%), aumentando consideravelmente os custos. Todas as estruturas ficam superdimensionadas. Por isso, reduzir variabilidade é fundamental para a redução de custos.

Adicionou-se cinza da casca de arroz, queimada a 850°C, proveniente de uma indústria beneficiadora de arroz de Pelotas/RS. Teoricamente, espera-se que a utilização da cinza de casca de arroz, por ser rica em sílica (entre 90 e 95%), aumente a resistência do concreto.

A metodologia utilizada por Mattos (2004) foi a de desenvolver uma estratégia para identificar efeitos de dispersão em experimentos com poucas replicações. Esta inicia com a identificação de efeitos de locação pelos métodos clássicos, ou seja, calcula o teste F para identificar os efeitos de locação e o método de mínimos quadrados ordinários (MQO) para construir o modelo. Após a avaliação da qualidade do modelo encontrado, feita por meio de técnicas formais e técnicas gráficas, é avaliada a necessidade de mudança da métrica da resposta, que no caso, foi a transformação logarítmica. Após a transformação dos dados, a existência de efeitos de dispersão é avaliada por meio de métodos que se utiliza dos resíduos quadráticos médio. Em função de alguns efeitos falsos poderem ser detectados, realiza-se o refinamento do modelo por

meio do método dos mínimos quadrados generalizados (MQG).

## Delimitação amostral

O delineamento do estudo foi planejado na forma de experimento fatorial completo com 5 fatores (A, B, C, D e E) ensaiados em dois níveis (-1 e +1) com três replicações. Portanto, nessas condições, a amostra foi composta por 96 ensaios em corpos-de-prova cilíndricos, de dimensões 10 x 20 cm, conforme Tabela G1 do Apêndice G.

## Variáveis de interesse

**Resposta**  $\rightarrow y =$  Resistência à compressão do concreto.

**Fator A**  $\rightarrow A =$  Relação água/cimento indica a quantidade de água a ser adicionada à mistura. As relações utilizadas foram: 0,35, menor quantidade de água (-1), e 0,60, maior quantidade de água (1).

**Fator B**  $\rightarrow B =$  Adição da cinza de casca de arroz que, em caso positivo, foi utilizada com o teor de 10% representada por (1) ou não utilizada por (-1).

**Fator C**  $\rightarrow C =$  Tempo de mistura, medido em segundos. Os tempos considerados foram de 100 segundos (-1) e 300 segundos (1).

**Fator D**  $\rightarrow D =$  Tipo de adensamento, realizado manualmente (-1) ou mecanicamente, com a utilização de um vibrador de imersão (1).

**Fator E**  $\rightarrow E =$  Idade dos corpos-de-prova, medida em dias, rompidos aos 7 dias (-1) e aos 28 dias (1).

## Análise de dados

O objetivo foi apresentar uma alternativa de modelar o experimento delineado por Mattos (2004), por meio dos MNLGS.

Essa aplicação segue a seguinte estrutura de modelagem:

i) Modelo linear generalizado com distribuição normal e ligação *log* com interações até quinta ordem do plano fatorial e parâmetro de dispersão fixo, para identificar efeitos significativos, norteados pela análise feita por Mattos (2004);

ii) Modelo linear generalizado superdispersado normal com função de ligação identidade;

iii) Modelo com distribuição passeio aleatório com o mesmo planejamento fatorial e função de ligação identidade;

iv) Cálculo dos vieses para as estimativas dos parâmetros do sub modelo da média e dispersão, com distribuição Normal e passeio aleatório;

v) Avaliação dos modelos ajustados;

vi) Avaliação dos estimadores de  $O(n^{-2})$  no experimento com uma e três replicações.

### Algoritmo para obtenção das estimativas de $O(n^{-1})$ e $O(n^{-2})$ do modelo Normal Superdispersado

A componente aleatória resistência à compressão do concreto, denotada por  $Y$ , com função densidade de probabilidade normal com parâmetros  $(\mu_i, \phi_i)$ , sendo  $\mu_i$  a média de  $y_i$  e  $\phi_i$  o parâmetro de dispersão para  $i = 1, 2, \dots, n$ .

Submodelo da média:

$$\eta_i = \mu_i = \exp(\beta_0 + \beta_A A_i + \beta_D D_i + \beta_E E_i + \beta_{AC} (AC)_i + \beta_{AE} (AE)_i + \beta_{BC} (BC)_i).$$

Foram considerados neste modelo somente os fatores e as interações que foram significativos ao nível de 5% no MLG normal.

Submodelo de dispersão:

$$\tau_i = \phi_i = \exp(\gamma_0 + \gamma_E E_i).$$

A adoção deste fator deve-se a Mattos (2004).

Os passos a seguir descrevem como foi elaborado o programa do SAS para modelar a resistência do concreto, com função densidade de probabilidade Normal.

#### 1º passo: obter as estimativas de $O(n^{-1})$

A Proc NLMIXED fornece essas estimativas maximizando a log-verossimilhança diretamente, isto é, obtendo as raízes do sistema de equações da função escore pelo método de Newton Raphson. Para isso, é necessário fornecer algumas equações que são enumeradas a seguir:

- → a função log-verossimilhança do modelo normal sob a forma dos MNLGS, que é dado por:

$$l(\boldsymbol{\mu}, \boldsymbol{\phi}; \mathbf{y}) = -\frac{1}{2} \log\left(\frac{1}{\boldsymbol{\phi}}\right) - \frac{\boldsymbol{\phi}}{2} (\mathbf{y} - \boldsymbol{\mu});$$

- → o submodelo média dada pela componente sistemática  $\eta$ .

- → o submodelo dispersão dada pela componente sistemática  $\tau$ .

## 2º passo: tentativa inicial

Foi utilizada a procedure GENMOD, para obter a tentativa inicial, pois a convergência do método de Newton Raphson, depende dos valores iniciais.

## 3º passo: resultados da procedure NLMIXED

As estimativas de ordem  $O(n^{-1})$  dos parâmetros dos submodelos, são disponibilizados em arquivos, que serão utilizados na Proc IML, para obter as estimativas de  $O(n^{-2})$ .

## 4º passo: obter as estimativas de $O(n^{-2})$

A Proc IML, permite ao usuário operar com matrizes; neste passo obtêm-se as matrizes que estão na expressão dos vieses, para os submodelos média e dispersão:

### Estimativas de $O(n^{-2})$ do submodelo da média

$\tilde{X}$  → cada elemento dessa matriz é um vetor de  $n$  elementos, isto é:

$$\tilde{X} = \left[ \frac{\partial \boldsymbol{\eta}}{\partial \beta_0}, \frac{\partial \boldsymbol{\eta}}{\partial \beta_A}, \frac{\partial \boldsymbol{\eta}}{\partial \beta_D}, \frac{\partial \boldsymbol{\eta}}{\partial \beta_E}, \frac{\partial \boldsymbol{\eta}}{\partial \beta_{AC}}, \frac{\partial \boldsymbol{\eta}}{\partial \beta_{AE}}, \frac{\partial \boldsymbol{\eta}}{\partial \beta_{BC}} \right], \text{ sendo } \left( \frac{\partial \boldsymbol{\eta}}{\partial \beta_0} \right)^T = \left[ \frac{\partial \eta_1}{\partial \beta_0} \quad \frac{\partial \eta_2}{\partial \beta_0} \quad \dots \quad \frac{\partial \eta_n}{\partial \beta_0} \right].$$

Neste problema  $\tilde{X} = [\boldsymbol{\eta}, \mathbf{A}\boldsymbol{\eta}, \mathbf{D}\boldsymbol{\eta}, \mathbf{E}\boldsymbol{\eta}, \mathbf{AC}\boldsymbol{\eta}, \mathbf{AE}\boldsymbol{\eta}, \mathbf{BC}\boldsymbol{\eta}]$ , com dimensão de  $(n \times 7)$ .

$\tilde{\tilde{X}}$  → são calculadas  $n$  matrizes  $\tilde{\tilde{X}}$ , cuja dimensão é  $(7 \times 7)$ ,

$$\tilde{\tilde{X}}_i = \begin{bmatrix} \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_0^2} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_0 \partial \beta_A} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_0 \partial \beta_D} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_0 \partial \beta_E} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_0 \partial \beta_{AC}} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_0 \partial \beta_{AE}} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_0 \partial \beta_{BC}} \\ & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_A^2} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_A \partial \beta_D} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_A \partial \beta_E} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_A \partial \beta_{AC}} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_A \partial \beta_{AE}} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_A \partial \beta_{BC}} \\ & & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_D^2} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_D \partial \beta_E} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_D \partial \beta_{AC}} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_D \partial \beta_{AE}} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_D \partial \beta_{BC}} \\ & & & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_E^2} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_E \partial \beta_{AC}} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_E \partial \beta_{AE}} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_E \partial \beta_{BC}} \\ & & & & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_{AC}^2} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_{AC} \partial \beta_{AE}} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_{AC} \partial \beta_{BC}} \\ & & & & & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_{AE}^2} & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_{AE} \partial \beta_{BC}} \\ & & & & & & \frac{\partial^2 \boldsymbol{\eta}_i}{\partial \beta_{BC}^2} \end{bmatrix},$$

a matriz de derivada de segunda ordem, para este problema é dada por:

$$\tilde{X}_i = \begin{bmatrix} \eta_i & \eta_i A_i & \eta_i D_i & \eta_i E_i & \eta_i (AC)_i & \eta_i (AE)_i & \eta_i (BC)_i \\ & \eta_i A_i^2 & \eta_i A_i D_i & \eta_i A_i E_i & \eta_i A_i (AC)_i & \eta_i A_i (AE)_i & \eta_i A_i (BC)_i \\ & & \eta_i D_i^2 & \eta_i D_i E_i & \eta_i D_i (AC)_i & \eta_i D_i (AE)_i & \eta_i D_i (BC)_i \\ & & & \eta_i E_i^2 & \eta_i E_i (AC)_i & \eta_i E_i (AE)_i & \eta_i E_i (BC)_i \\ & & & & \eta_i (AC)_i^2 & \eta_i (AC)_i (AE)_i & \eta_i (AC)_i (BC)_i \\ & & & & & \eta_i (AE)_i^2 & \eta_i (AE)_i (BC)_i \\ & & & & & & \eta_i (BC)_i^2 \end{bmatrix},$$

para  $i = 1, 2, \dots, n$ .

**Matriz**  $\Psi^{(2,0)}$ , cujos elementos da diagonal são:

$$diag(\Psi^{(2,0)}) = \phi_i.$$

**Matriz**  $M_1^2$ , cujos elementos da diagonal são 1:

$$M_1^2 = \left( \frac{\partial \mu}{\partial \eta} \right)^2, \text{ com função de ligação identidade.}$$

$$M_1^2 = I, \text{ matriz identidade.}$$

**Matriz**  $(K_{\beta\beta})$ , matriz de informação de Fisher essa matriz é dada por:

$$\tilde{X}^T \psi^{(2,0)} M_1^2 \tilde{X}.$$

**Traço da matriz de informação de Fisher pela**  $\tilde{X}$ , que é dada por :

$$tr(K_{\beta\beta}^{-1} * \tilde{X}_i).$$

**Vetor**  $\xi_1$ , os elementos desse vetor são:

$$\xi_1 = -\frac{1}{2} Z_{\beta d} M_1^{-1} M_2, \text{ onde}$$

$$Z_{\beta d} = \tilde{X} K_{\beta,\beta}^{-1} \tilde{X}^T \text{ e } M_2 = 0.$$

**Vetor**  $\xi_2$ , os elementos desse vetor são:

$$\xi_2 = -\frac{1}{2} D_\beta, \text{ onde } D_\beta \text{ é a matriz cuja diagonal}$$

$$\text{têm como elementos dado por } tr(K_{\beta\beta}^{-1} * \tilde{X}_i).$$

## Estimativas de $O(n^{-2})$ do submodelo da dispersão

$\tilde{S}$ , cada elemento desta matriz é um vetor de  $n$  elementos, sua expressão é dada por:

$\tilde{S} = \left[ \frac{\partial \boldsymbol{\tau}}{\partial \gamma_0}, \frac{\partial \boldsymbol{\tau}}{\partial \gamma_E} \right]$ , sendo  $\left( \frac{\partial \boldsymbol{\tau}}{\partial \gamma_0} \right)^T = \left[ \frac{\partial \tau_1}{\partial \gamma_0} \quad \frac{\partial \tau_2}{\partial \gamma_0} \quad \dots \quad \frac{\partial \tau_n}{\partial \gamma_0} \right]$ . Neste problema  $\tilde{S} = [\boldsymbol{\tau}, E\boldsymbol{\tau}]$ , com dimensão de  $(n \times 2)$

$\tilde{S}$  Matrizes das derivadas segundas, que foram calculadas  $n$  matrizes  $\tilde{S}_i$ , cuja dimensão é  $(2 \times 2)$ ,

$$\tilde{S}_i = \begin{bmatrix} \frac{\partial^2 \boldsymbol{\tau}}{\partial \gamma_{0i}^2} & \frac{\partial^2 \boldsymbol{\tau}_i}{\partial \gamma_0 \partial \gamma_E} \\ & \frac{\partial^2 \boldsymbol{\tau}_i}{\partial \gamma_E^2} \end{bmatrix} = \begin{bmatrix} \tau_i & \tau_i E_i \\ & \tau_i E_i^2 \end{bmatrix}$$

**Matriz  $\Psi^{(0,2)}$** , cujos os elementos da diagonal são:

$$diag(\Psi^{(0,2)}) = -0.5\phi_i^{-2}$$

**Matriz  $\Psi^{(0,3)}$** , cujos os elementos da diagonal são:

$$diag(\Psi^{(0,3)}) = \phi_i^{-3}$$

**Matriz  $\Psi^{(2,1)}$** , cujos os elementos da diagonal são:

$$diag(\Psi^{(2,1)}) = 1$$

**Matriz  $\phi_1^2$** , cujos os elementos da diagonal são 1:

$$\begin{aligned} \Phi_1^2 &= \left( \frac{\partial \phi}{\partial \boldsymbol{\tau}} \right)^2, \text{ como a função ligação é a identidade} \\ \Phi_1^2 &= I, \text{ matriz identidade.} \end{aligned}$$

**Matriz  $\phi_2$** , cujos os elementos da diagonal são *zero*:

$$\begin{aligned} \Phi_2 &= \left( \frac{\partial^2 \phi}{\partial \boldsymbol{\tau}^2} \right), \text{ como a função ligação é a identidade} \\ \Phi_2 &= 0, \text{ matriz nula.} \end{aligned}$$

**Matriz de informação de Fisher** ( $K_{\gamma\gamma}$ ) essa matriz é dada por:

$$\tilde{S}^T \Psi^{(0,2)} \Phi_1^2 \tilde{S}.$$

**Matriz**  $tr(K_{\gamma\gamma}^{-1} * \tilde{S}_i)$ , é a matriz traço do produto da matriz inversa de informação de Fisher pela  $\tilde{S}_i$ .

**Vetor**  $\delta_1$  os elementos desse vetor são:

$$\delta_1 = \left\{ -\frac{1}{2} (\Psi^{(0,2)})^{-1} [Z_{\gamma d} (\Psi^{(0,3)}) + Z_{\beta d}] \right\} \text{ onde}$$

$$Z_{\beta d} = \tilde{X} K_{\beta,\beta}^{-1} \tilde{X}^T \text{ e } Z_{\gamma d} = \tilde{S} K_{\gamma,\gamma}^{-1} \tilde{S}^T$$

**Vetor**  $\delta_2$  os elementos desse vetor são:

$$\delta_2 = -\frac{1}{2} D_\gamma, \text{ onde } D_\gamma$$

é a matriz cuja diagonal têm como elemento  $tr(K_{\gamma\gamma}^{-1} *) \tilde{S}_i$

### 5º passo: Proc Model

$B(\hat{\beta})$  - O viés de  $\hat{\beta}$  é a solução do sistema de equações normais de mínimos quadrados ponderados, isto é, que nada mais é do que as equações normais com matriz modelo  $\tilde{X}$ , peso  $\Psi^{(2,0)} M_1^2$  e o vetor resposta  $\xi_1 + \xi_2$ .

$B(\hat{\gamma})$  - O viés de  $\hat{\gamma}$  é a solução do sistema de equações normais de mínimos quadrados ponderados, isto é, que nada mais é do que as equações normais com matriz modelo  $\tilde{S}$ , peso  $\Psi^{(0,2)} \Phi_1^2$  e vetor resposta  $\delta_1 + \delta_2$ .

No Apêndice G se encontra o programa do procedimento descrito acima para o modelo normal com as estimativas de  $O(n^{-1})$  e  $O(n^{-2})$ .

### Algoritmo para obtenção das estimativas de $O(n^{-1})$ e $O(n^{-2})$ do modelo passeio aleatório

Levando em consideração que a característica da variável resposta resistência à compressão do concreto são sempre positivos, adotou-se a distribuição passeio aleatório.

Para a distribuição passeio aleatório a função verossimilhança dos modelos não lineares generalizados superdispersados segue a expressão (7.1), com as duas componentes sistemáticas representadas iguais ao do modelo normal.

Os passos para a obtenção das estimativas seguem os mesmos apresentados para o modelo normal, porém com as matrizes derivadas do modelo passeio aleatório que estão disponíveis na primeira seção deste capítulo.

No Apêndice H pode ser visto o programa executado para as estimativas de  $O(n^{-1})$  e  $O(n^{-2})$  do passeio aleatório

## Resultados

Na seqüência estão apresentadas tabelas com os resultados dos modelos normal e passeio aleatório.

Para a distribuição Normal a Tabela 7.6 apresenta as estimativas dos parâmetros do sub-modelo média com seus erros-padrão e seus respectivos vieses de  $O(n^{-1})$ .

Tabela 7.6: Submodelo média da normal, estimativas, erros-padrão e vieses.

Estimadores	Estimativas (erros)	Vieses de $O(n^{-1})$	$p$ -valor
$\hat{\beta}_0$	3.7116 (0.0118)	-0.0004	< .0001
$\hat{\beta}_A$	-0.2552 (0.0113)	-0.0001	< .0001
$\hat{\beta}_D$	-0.0539 (0.0087)	$-8.1E - 6$	< .0001
$\hat{\beta}_E$	0.0550 (0.0118)	-0.0002	< .0001
$\hat{\beta}_{AC}$	0.0210 (0.0087)	$3.3E - 6$	0.0181
$\hat{\beta}_{AE}$	0.0481 (0.0118)	-0.0003	< .0001
$\hat{\beta}_{BC}$	0.0247 (0.0087)	$4.1E - 6$	0.0055

Para o submodelo média, os fatores principais A, D e E foram significativos com as interações AC, AE e BC, evidenciando que influenciam na resistência média à compressão do concreto. Quanto aos vieses, estes representam 1%, 1,7% e 2,5% do erros-padrão para os fatores A, E e AE e 3,4 % para o erro-padrão de  $\hat{\beta}_0$ . Nos demais fatores e interações os vieses tomaram valores não importantes.

As estimativas do submodelo dispersão encontram-se na Tabela 7.7 a seguir.

Tabela 7.7: Submodelo dispersão da normal, estimativas, erros-padrão e vieses.

Estimadores	Estimativas (erros)	Vieses de $O(n^{-1})$	$p$ -valor
$\hat{\gamma}_0$	-2.8370 (0.1443)	-0.0209	< .0001
$\hat{\gamma}_E$	-0.7282 (0.1455)	-0.0018	< .0001

O valor do viés representa 13,8% do erro-padrão para o estimador  $\hat{\gamma}_0 = -2.8370$  e para  $\hat{\gamma}_1$  o valor do viés representa 1,2% de seu erro-padrão. O valor do  $AIC$  para o modelo normal foi de 386.

Usando a log-verossimilhança da distribuição Normal no proc NLMIXED do SAS, este convergiu na 95ª iteração.



No entanto, usando a metodologia dos MNLGS a estimativa foi encontrada na 21<sup>a</sup>. Portanto, houve uma melhoria da convergência.

Os resultados do modelo estimado passeio aleatório para a média da resistência do concreto à compressão estão apresentados nas tabelas que se seguem.

Tabela 7.8: Submodelo média do passeio aleatório com réplica e sem réplica, estimativas, erros-padrão e vieses.

	<i>Estimadores</i>	<i>Estimativas</i>	<i>Erros – padrão</i>	<i>Vieses de <math>O(n^{-1})</math></i>	<i>p – valor</i>
$n = 96$	$\hat{\beta}_0$	17.6242	0.0097	0.00047	< .0001
	$\hat{\beta}_A$	-3.4004	0.0096	0.00004	< .0001
	$\hat{\beta}_D$	0.0592	0.0191	0.00002	0.0024
	$\hat{\beta}_E$	-21.3757	0.0096	0.00047	< .0001
	$\hat{\beta}_{AC}$	-0.0154	0.0189	-5.88E - 6	0.4177
	$\hat{\beta}_{AE}$	3.5995	0.0096	0.00004	< .0001
	$\hat{\beta}_{BC}$	-0.019	0.0189	-7.25E - 6	0.3190
$n = 32$	$\hat{\beta}_0$	17.1093	0.0153	0.00117	< .0001
	$\hat{\beta}_A$	-3.2296	0.0154	0.00012	< .0001
	$\hat{\beta}_D$	0.0678	0.0303	0.00006	0.032
	$\hat{\beta}_E$	-20.8907	0.0154	0.00117	< .0001
	$\hat{\beta}_{AC}$	-0.0014	0.0301	-1.31E - 6	0.9631
	$\hat{\beta}_{AE}$	3.4704	0.0154	0.00011	< .0001
	$\hat{\beta}_{BC}$	-0.0008	0.0301	-8.25E - 7	0.9768

Para o submodelo média com distribuição passeio aleatório, os fatores principais A, D e E também mostraram-se significativos, no entanto somente a interação AE influencia na resistência média à compressão do concreto.

Os resultados das estimativas do submodelo dispersão encontram-se na tabela abaixo.

Tabela 7.9: Submodelo dispersão do passeio aleatório com réplica e sem réplica, estimativas, erros-padrão e vieses.

	<i>Estimadores</i>	<i>Estimativas</i>	<i>Erros-padrão</i>	<i>Vieses de <math>O(n^{-1})</math></i>	<i>p-valor</i>
$n = 96$	$\hat{\gamma}_0$	-1.7052	0.1443	0.0042	< .0001
	$\hat{\gamma}_E$	1.9985	0.1444	0.0043	< .0001
$n = 32$	$\hat{\gamma}_0$	-1.6282	0.2500	0.0196	< .0001
	$\hat{\gamma}_E$	2.0712	0.2501	0.0198	< .0001

No submodelo dispersão, este manteve-se significativo com os mesmos fatores encontrados na distribuição normal.

Para o tamanho de amostra igual a 96, os vieses das estimativas do submodelo média representam de 0,5% a 5% dos erros-padrão e para o submodelo dispersão igual a 3%, enquanto que para amostra com apenas uma repetição (32 observações) os vieses ficaram num intervalo

de 0,2% a 12% quando comparados com seus respectivos erros-padrão no submodelo média e 8% no submodelo dispersão.

Mais uma vez, observa-se pela magnitude dos vieses que em amostras pequenas, estes tomam valores consideráveis quando comparados aos erros-padrão.

Para o ajuste do modelo passeio aleatório obteve-se um valor para o  $AIC$  igual a 277.

Diante dos valores obtidos do  $AIC$  nos modelos adotados, o melhor foi com distribuição passeio aleatório, pois apresenta menor valor do  $AIC$ . Observou-se, também, que os estimadores obtiveram magnitudes superiores quando comparado pela distribuição normal, indicando a intensidade maior nos efeitos na resposta média e dispersão da resistência à compressão do concreto. O sinal da estimativa indica a direção do efeito, ou seja, se a resposta média da resistência do concreto aumenta ou diminui com o nível de (+1) para (-1).

Pode-se ainda afirmar que a metodologia dos modelos não-lineares generalizados superdispersados apresenta algumas vantagens consideráveis para modelar a resistência à compressão do concreto :

- i) A variável resposta pode ter qualquer distribuição de probabilidade pertencente aos MNLGS;
- ii) Modelam-se os dados originais, não tendo necessidade de transformações ;
- iii) Considera-se o parâmetro de dispersão na própria função densidade e na função de variância do modelo;
- iv) É fácil implementar em qualquer aplicativo e, além disso, a convergência é rápida

Na metodologia dos modelos não-lineares generalizados superdispersados detectou-se que os principais fatores que influenciaram na resistência à compressão do concreto foram água/cimento (Fator A), com a forma de adensamento (Fator D) e idade (Fator E), concordando com os encontrados por Mattos (2004). Portanto, há evidências de que maiores valores para a resistência média para corpos-de-prova moldados são para as rupturas aos 28 dias, com menor quantidade de água e adensados manualmente. E a variabilidade (dispersão) da resistência à compressão do concreto se deve a idade dos corpos-de-prova rompidos aos 7 e 28 dias, concordando com os resultados obtidos por Mattos (2004) e também quanto a hipótese de que a adição de cinza de casca de arroz poderia aumentar a resistência à compressão do concreto é rejeitada, exceto quando interage com o tempo de mistura (fator B).

Quanto aos vieses de  $O(n^{-1})$  para o modelo passeio aleatório, estes apresentaram valores significativos para amostra com apenas uma repetição.

# CAPÍTULO 8

## CONCLUSÕES

O modelo-não linear ajustado para o comportamento do número médio de erros de aprendizagem foi significativo, evidenciando o déficit de aprendizagem entre os ratos lesionados e os não-lesionados. Portanto, ele descreve de modo satisfatório os resultados biológicos, quantificando de modo fidedigno as diferenças observadas entre ratos lesionados e não-lesionados. E quanto as estimativas de segunda ordem, estas devem ser utilizadas, uma vez que apresentam magnitude aproximadamente na casa de 10% dos seus respectivos erros-padrão, principalmente neste tipo de experimento, em que as amostras (grupo de animais) são geralmente pequenas, o que resulta em estimativas viesadas.

Para o experimento da resistência à compressão do concreto, o MNLGS mostrou -se eficiente para identificar os fatores que influenciam na resistência à compressão do concreto, e quanto aos vieses de  $O(n^{-1})$  para os estimadores do submodelo da média e dispersão, evidenciaram mais uma vez que, quando comparados com os erros-padrão respectivos, tomam valores importantes em amostras de tamanho pequeno.

Desse modo, esses resultados vieram corroborar com as simulações de que o tamanho amostral interfere na magnitude dos vieses.

Quanto aos estimadores corrigidos  $\tilde{\beta}$  e  $\tilde{\gamma}$ , estes demonstraram que ficaram em média mais próximos dos verdadeiros valores dos parâmetros do que os não-corrigidos  $\hat{\beta}$  e  $\hat{\gamma}$ .

### 8.1 Considerações finais

Evidenciou-se por meio da simulação de Monte Carlo e em problemas com dados reais que os estimadores de segunda ordem são mais precisos, isto é, apresentam erros-padrão menores. Além disso, a magnitude dos vieses dos estimadores dos MNLGS, têm maior impacto para as amostras pequenas a moderadas e que as expressões dos vieses de  $O(n^{-1})$  são suficientemente sensíveis para detectar diferenças entre os estimadores e os parâmetros de máxima verossimilhança em amostras pequenas e isso traz conseqüências práticas positivas, tanto do ponto de vista econômico quanto operacional, na engenharia de produção, visto ser comum na prática

obter amostras em tempo real e, que portanto, o tamanho desta tende a ser pequeno, o que justifica utilizar a correção do viés nos modelos de produção, proporcionando ganhos econômicos significativos, pois os produtos finais podem apresentar maior uniformidade, reduzindo, conseqüentemente seus custos de produção. Assim, sugere-se usar a expressão do cálculo do viés de  $O(n^{-1})$  obtida nesta pesquisa para os modelos não lineares generalizados superdispersados.

## 8.2 Recomendações

Recomenda-se para o desenvolvimento de pesquisas futuras no contexto dos modelos não-lineares generalizados superdispersados, obter uma expressão fechada para a distribuição densidade de probabilidade da inversa gaussiana generalizada, a qual pertence a esta classe de modelos, conforme foi demonstrada nesta pesquisa.

Recomenda-se também que nas aplicações do Capítulo 7, seja feita uma análise criteriosa quanto aos testes de diagnósticos, uma vez que, nesta tese preocupou-se somente com os ajustes de primeira e segunda ordem.

# REFERÊNCIAS

- [1] BAUER, C. (2000), Value at Risk Using Hyperbolic Distributions - Journal of Economics and Business, 52,455-467.
- [2] BARTLETT, M.S. (1953). Approximate Confidence Intervals II. Biometrika, 40, 306-317.
- [3] BOTTER, D. A. (1993). Correções de Bartlett e Poderes de Testes em Algumas Classes de Modelos de Regressão. Tese (Doutorado)- Instituto de Matemática e Estatística - Universidade de São Paulo.
- [4] BOTTER, D. A. e CORDEIRO, G. M. (1998). Improved Estimators for Generalized Linear Models with Dispersion Covariates. Journal of Statistical Computation and Simulation, 62, 91-104.
- [5] BOX, M.J. Bias in Nonlinear Estimation. (1971). Journal of the Royal Statistical Society B, 33, 171-201.
- [6] COOK, R.D., TSAI, C.L. e WEI, B.C. (1986). Bias in Nonlinear Regression. Biometrika, 73, 615-623.
- [7] COOK, R.D. e TSAI, C.L. (1983). Diagnostic for Assessing the Accuracy of Normal Approximations in Exponential Family Nonlinear Models. Biometrika, 70, 1-10.
- [8] CORDEIRO, G.M. (1983). Improved Likelihood Ratio Statistics for Generalized Linear Models. Journal of the Royal Statistical Society B, 45, 404-413.
- [9] CORDEIRO, G.M. (1986). Modelos Lineares Generalizados. VII Simpósio Nacional de Probabilidade e Estatística, UNICAMP, Campinas, São Paulo.
- [10] CORDEIRO, G.M. (1987). On the Correction to the likelihood ratio statistics. Biometrika, 74, 265-274.
- [11] CORDEIRO, G.M. (1999). Introdução à Teoria Assintótica. 22<sup>o</sup> Colóquio Brasileiro de Matemática, IMPA, Rio de Janeiro.
- [12] CORDEIRO, G.M. e BOTTER, D.A. (2001). Second-Order Biases of Maximum Likelihood Estimates in Overdispersed Generalized Linear Models. Statistics and Probability Letters, 55, 269-280.

- [13] CORDEIRO, G.M. e MCCULLAGH, P. (1991). Bias Correction in Generalized Linear Models. *Journal of the Royal Statistical Society B*, 53, 629-643.
- [14] CORDEIRO, G.M. e BARROSO, L.P. (2003). A Third-Order Bias Corrected Estimate in Generalized Linear Models. Relatório Técnico.RT-MAE.
- [15] CORDEIRO, G.M. e VASCONCELLOS, K.L.P. (1997). Bias Correction for a Class of Multivariate Nonlinear Regression Models. *Statistics and Probability Letters*, 35, 155-164.
- [16] COX, D.R. e SNELL, E.J. (1968). A General Definition of Residuals (with discussion). *Journal of the Royal Statistical Society B*, 30, 248-275.
- [17] DAVIDIAN, M. e GILTINAN, D.M. (1995). *Nonlinear Models for Repetead Measurement Data*. 1ª Edição. Chapman & Hall. London.
- [18] DOBSON, A.J. (2001). *An Introduction to Generalized Linear Models*. 2ª Edição. Chapman & Hall/CRC. Boca Raton, Florida.
- [19] EFRON, B. (1986). Double Exponential Families and Their Use in Generalized Linear Regression. *Journal of the American Statistical Association*, 81, 709-721.
- [20] FAHMEIR, L. e KAUFMANN, H. (1985). Consistency and Asymptotic Normality of the Maximum Likelihood Estimator in Generalized Linear Models. *Ann. Statist.*, 13, 342-368.
- [21] FISHER, R. A. (1912). On an Absolute Criterion for Fitting Frequency Curves. *Messenger of Mathematics* 41, 155 - 160.
- [22] JORGENSEN, B. (1983). Maximum Likelihood Estimation and Large-Sample Inference for Generalized Linear and Nonlinear Regression Models. *Biometrika*, 70, 19-28.
- [23] JORGENSEN, B. (1987). Small Dispersion Asymptotics. *Revista Brasileira de probabilidade e Estatística*, 1, 59-90.
- [24] JORGENSEN, B. (1997). *The Theory of Dispersion Models*. 1ª Edição. Chapman & Hall. Great Britain.
- [25] KÖCHE, J.C. (1997). *Fundamentos de Metodologia Científica*. 14ª Edição. Editora Vozes. Petrópolis, Rio de Janeiro.
- [26] LEITE, J. G. e SINGER J. da M. (1990). Métodos Assintóticos em Estatística - Fundamentos e Aplicações. 9º Simpósio Nacional de Probabilidade e Estatística, IME - USP, São Paulo.
- [27] LINDSEY, J. K. et al. (2000). Generalized Nonlinear Models for Pharmacokinetics Data. *Biometrics* 56, 81 - 88.

- [28] Manual do SASa. (13/07/05) [www-fp.mcs.anl.gov/otc/Guide/SoftwareGuide/Blurbs/procnlp.html](http://www-fp.mcs.anl.gov/otc/Guide/SoftwareGuide/Blurbs/procnlp.html)
- [29] Manual do SASb (13/07/05) [www.id.unizh.ch/software/unix/statmath/sas/sasdoc/stat/chap46/sect4.htm](http://www.id.unizh.ch/software/unix/statmath/sas/sasdoc/stat/chap46/sect4.htm)
- [30] McCULLAGH, P. e NELDER, J. A. (1983). *Generalized Linear Models*. 1ª Edição. Chapman & Hall. London.
- [31] McCULLAGH, P. e NELDER, J. A. (1989). *Generalized Linear Models*. 2ª Edição. Chapman & Hall. London.
- [32] NELDER, J.A. e WEDDERBURN, R.W.M. (1972). *Generalized Linear Models*. *Journal of the Royal Statistical Society A*, 135, 370-384.
- [33] OGLIARI, P. J. (1998). *Modelos Não-Lineares para Dados Longitudinais Provenientes de Experimentos em Blocos Casualizados*. Tese (Doutorado)- Escola Superior de Agricultura "Luiz de Queiroz" - Universidade de São Paulo.
- [34] PAZMAN, A. (1991). *Curvatures and the Distribution of the Maximum Likelihood Estimator in Nonlinear Exponential Models*. *Brazilian Journal of Probability and Statistics*, 5, 43-63.
- [35] PAULA, G.A. (1992). *Bias Correction for Exponential Family Nonlinear Models*. *Journal of Statistical Computation and Simulation*, 40, 43-54.
- [36] PREGIBON, D. (1984). *Review of Generalized Linear Models*. *Annals of Statistics*, 12, 1589-1596.
- [37] SMITH, G.K. (1989). *Generalized Linear Models With Varying Dispersion*. *Journal of the Royal Statistical Society B*, 51, 47-60.
- [38] SMITH, G.K. e VERBYLA, A.P. (1999). *Adjusted Likelihood Methods for Modelling Dispersion in Generalized Linear Models*. *Environmetrics*, 10, 696-709.
- [39] SOWDEN, R.R. (1971). *Bias and Accuracy of Parameter Estimates in a Quantal Response Model*. *Biometrika*, 58, 595-603.
- [40] Wei, B.C. (1998). *Exponential Family Nonlinear Models*. Springer-Verlag Singapore Pte. Ltd. China. 1998.
- [41] Weisberg, S. (1985). *Applied Linear Regression*. 1ª Edição. Wiley. New York.
- [42] Whitmore, G.A. (1986). *Inverse Gaussian Ratio Estimation*. *Applied Statistics*, 35, 8-15.
- [43] Wolfinger, R. D. (2004) *Fitting Nonlinear Mixed Models with the New NLMIXED Procedure*. SAS Institute Inc., Cary, NC.

[44] [www-fp.mcs.anl.gov/otc/Guide/SoftwareGuide/Blurbs/procnlp.html](http://www-fp.mcs.anl.gov/otc/Guide/SoftwareGuide/Blurbs/procnlp.html)



# APÊNDICE A

## CÁLCULO DOS CUMULANTES PARA OS MNLGS

É importante lembrar algumas notações para o cálculo dos cumulantes, tais como:

$$\begin{aligned} \kappa_{rtu} &= E \left[ \frac{\partial^3 \ell(\beta, \phi)}{\partial \beta_r \partial \beta_t \partial \beta_u} \right] \quad \text{e} \quad \kappa_{rt,u} = E \left[ \left\{ \frac{\partial^2 \ell(\beta, \phi)}{\partial \beta_r \partial \beta_t} \right\} \frac{\partial \ell(\beta, \phi)}{\partial \beta_\mu} \right], \\ (r) &= \frac{\partial \eta}{\partial \beta_r}, \quad (r, t) = \frac{\partial \eta}{\partial \beta_r} \frac{\partial \eta}{\partial \beta_t}, \quad (rt) = \frac{\partial^2 \eta}{\partial \beta_r \partial \beta_t}, \quad \tilde{X} = \frac{\partial \eta}{\partial \beta} \\ \kappa_{rt,u} &= \kappa_{rtu} - \kappa_{rs}^{(t)}, \quad \tilde{\tilde{X}} = \frac{\partial^2 \eta}{\partial \beta \partial \beta^T} \end{aligned}$$

Os  $\kappa$ 's são todos de  $O(n^{-1})$ .

A inversa da matriz informação de Fisher para  $\beta$  é dada por  $K^{-1} = \{-\kappa^{rs}\}$ .

E as derivadas dos cumulantes são por

$$\kappa_{rs}^{(t)} = \frac{\partial \kappa_{rs}}{\partial \beta_t}, \quad \kappa_{rs}^{(T)} = \frac{\partial \kappa_{rs}}{\partial \gamma^T}, \quad \text{etc.}$$

Observe as relações  $\kappa_{rs,t} = \kappa_{rst} - \kappa_{rs}^{(t)}$ , que dá a covariância entre as primeiras derivadas de  $\ell(\beta, \gamma)$ . Note ainda que  $\kappa_{r,s} = -\kappa_{rs}$  e  $\kappa_{R,S} = -\kappa_{RS}$  são os elementos das matrizes de informação de Fisher  $K_{\beta,\beta}$  e  $K_{\gamma,\gamma}$  para os parâmetros  $\beta$  e  $\gamma$ , respectivamente. Tem-se também que  $\kappa^{r,s} = -\kappa^{rs}$  e  $\kappa^{R,S} = -\kappa^{RS}$  são os elementos de suas inversas  $K_{\beta,\beta}^{-1}$  e  $K_{\gamma,\gamma}^{-1}$  os quais são da ordem  $O(n^{-1})$  e, ainda, que nem todos os  $k_s$  são independentes.

Alguns cumulantes já estão disponíveis em Cordeiro e Barroso (2004), Cordeiro et al (2003).

### A.1 Obtenção dos cumulantes para o cálculo do $B(\hat{\beta})$

Os cumulantes necessários para calcular o viés do estimador  $\hat{\beta}_r$  são:  $\kappa_s$ ,  $\kappa_{st}$ ,  $\kappa_{st}^{(u)}$ ,  $\kappa_{rst}$  e  $\kappa_{rs,t}$ .

Esses foram demonstrados como se seguem:

$$1. \text{ i) } \kappa_{st} \implies E[U_{st}] = E\left[\frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta_s \partial \beta_t}\right].$$

Considerando que  $\ell(\beta, \gamma)$  é dada pela equação

$$\begin{aligned} \ell &= \sum \left[ (y - \mu) \psi^{(1,0)}(\mu, \phi) + \phi T(y) + \psi(\mu, \phi) \right], \text{ então a} \\ \frac{\partial \ell}{\partial \beta_s} &= \sum \left[ (y - \mu) \frac{\partial \mu}{\partial \eta} \frac{\partial \eta}{\partial \beta_s} \psi^{(1,0)} + \psi^{(2,0)} \frac{\partial \mu}{\partial \eta} \frac{\partial \eta}{\partial \beta_s} (y - \mu) \right. \\ &\quad \left. + \psi^{(1,0)} \frac{\partial \mu}{\partial \eta} \frac{\partial \eta}{\partial \beta_s} \right], \\ &= \sum \psi^{(2,0)} \frac{\partial \mu}{\partial \eta} \frac{\partial \eta}{\partial \beta_s} (y - \mu). \end{aligned}$$

Calculando  $\frac{\partial^2 \ell}{\partial \beta_s \partial \beta_t}$

$$\frac{\partial^2 \ell}{\partial \beta_s \partial \beta_t} = \sum \frac{\partial}{\partial \beta_t} \left( \psi^{(2,0)} (y - \mu) \frac{\partial \mu}{\partial \eta} \frac{\partial \eta}{\partial \beta_s} \right)$$

Para simplificar denotam-se as seguintes parcelas:

$P1 = \psi^{(2,0)} (y - \mu)$ ,  $P2 = \frac{\partial \mu}{\partial \eta}$  e  $P3 = \frac{\partial \eta}{\partial \beta_s}$ , logo tem-se que

$$\begin{aligned} \frac{\partial^2 \ell}{\partial \beta_s \partial \beta_t} &= \sum \frac{\partial}{\partial \beta_t} [P1.P2.P3] \\ &= \sum \left( \frac{\partial P1}{\partial \beta_t} . P2 . P3 + \frac{\partial P2}{\partial \beta_t} . P3 . P1 + \frac{\partial P3}{\partial \beta_t} . P2 . P1 \right). \end{aligned}$$

Calculando  $\frac{\partial P1}{\partial \beta_t} . P2 . P3$

$$\begin{aligned} \frac{\partial P1}{\partial \beta_t} . P2 . P3 &= \sum (y - \mu) \psi^{(3,0)} \left( \frac{\partial \mu}{\partial \eta} \right)^2 \frac{\partial \eta}{\partial \beta_s} \frac{\partial \eta}{\partial \beta_t} \\ &\quad - \psi^{(2,0)} \left( \frac{\partial \mu}{\partial \eta} \right)^2 \frac{\partial \eta}{\partial \beta_s} \frac{\partial \eta}{\partial \beta_t}. \end{aligned}$$

Calculando  $\frac{\partial P2}{\partial \beta_t} . P1 . P3$

$$\frac{\partial P2}{\partial \beta_t} . P1 . P3 = \sum (y - \mu) \psi^{(2,0)} \frac{\partial^2 \mu}{\partial \eta^2} \frac{\partial \eta}{\partial \beta_s} \frac{\partial \eta}{\partial \beta_t}$$

Calculando  $\frac{\partial P3}{\partial \beta_t} . P1 . P2$

$$\frac{\partial P3}{\partial \beta_t} . P1 . P2 = \sum (y - \mu) \psi^{(2,0)} \frac{\partial \mu}{\partial \eta} \frac{\partial^2 \eta}{\partial \beta_s \partial \beta_t}.$$

Substituindo estas derivadas em  $\frac{\partial^2 \ell}{\partial \beta_s \partial \beta_t}$  e calculando a esperança, tem-se o valor do cumulante

$$\begin{aligned}\kappa_{st} &= \left\{ -\psi^{(2,0)} \left( \frac{\partial \mu}{\partial \eta} \right)^2 \frac{\partial \eta}{\partial \beta_s} \frac{\partial \eta}{\partial \beta_t} \right\}, \text{ ou seja,} \\ \kappa_{st} &= -\sum \psi^{(2,0)} m_1^2(s, t).\end{aligned}$$

1. ii) Obtendo a derivada do cumulante de 2ª ordem  $\kappa_{st}^{(u)}$

$$\begin{aligned}\kappa_{st}^{(u)} &= \frac{\partial \kappa_{st}}{\partial \beta_u} \\ &= -\sum \psi^{(3,0)} \left( \frac{\partial \mu}{\partial \eta} \right)^3 \frac{\partial \eta}{\partial \beta_s} \frac{\partial \eta}{\partial \beta_t} \frac{\partial \eta}{\partial \beta_u} \\ &\quad - 2\psi^{(2,0)} \frac{\partial \mu}{\partial \eta} \frac{\partial^2 \mu}{\partial \eta^2} \frac{\partial \eta}{\partial \beta_s} \frac{\partial \eta}{\partial \beta_t} \frac{\partial \eta}{\partial \beta_u} \\ &\quad - \psi^{(2,0)} \left( \frac{\partial \mu}{\partial \eta} \right)^2 \left[ \frac{\partial^2 \eta}{\partial \beta_s \partial \beta_u} \frac{\partial \eta}{\partial \beta_t} + \frac{\partial^2 \eta}{\partial \beta_t \partial \beta_u} \frac{\partial \eta}{\partial \beta_s} \right],\end{aligned}$$

ou seja,

$$\begin{aligned}\kappa_{st}^{(u)} &= -\sum \psi^{(3,0)} m_1^3(s, t, u) - 2\psi^{(2,0)} m_1 m_2(s, t, u) \\ &\quad - \psi^{(2,0)} m_1^2[(su, t) + (tu, s)].\end{aligned}$$

iii) Obtendo o cumulante  $\kappa_{rst}$

$$\begin{aligned}\kappa_{rst} &= E(U_{stu}) = E\left(\frac{\partial^3 \ell}{\partial \beta_s \partial \beta_t \partial \beta_u}\right) = E\left(\frac{\partial}{\partial \beta_u} \left(\frac{\partial^2 \ell}{\partial \beta_s \partial \beta_t}\right)\right) \\ \kappa_{rst} &= E\left\{ \sum \frac{\partial}{\partial \beta_u} \left( (y - \mu) \psi^{(3,0)} m_1^2(s, t) - \psi^{(2,0)} m_1^2(s, t) + \right. \right. \\ &\quad \left. \left. (y - \mu) \psi^{(2,0)} m_2(s, t) + (y - \mu) \psi^{(2,0)} m_1(st) \right) \right\} \\ \kappa_{rst} &= \sum \left\{ \begin{aligned} &-2\psi^{(3,0)} m_1^3(s, t, u) - 3\psi^{(2,0)} m_1 m_2(s, t, u) - \\ &\psi^{(2,0)} m_1^2[(su, t) + (tu, s) + (st, u)] \end{aligned} \right\}.\end{aligned}$$

Alguns cumulantes, como o

$$\kappa_{ST}^{(U)} = 0 \quad \text{e} \quad \kappa_{sTU} = 0$$

foram apresentados em Cordeiro et al (2003)

## A.2 Obtenção dos cumulantes para o cálculo do $B(\hat{\gamma})$

Os cumulantes necessários para calcular esse viés são:  $\kappa_{ST}$ ,  $\kappa_{ST}^{(u)}$ ,  $\kappa_{STU}$  e  $\kappa_{Stu}$

## Obtenção dos cumulantes

$$i) \kappa_{ST} = E \left( \frac{\partial \ell}{\partial \gamma_s} \left( \frac{\partial \ell}{\partial \gamma_T} \right) \right)$$

$$\begin{aligned} \ell &= \sum \left[ (y - \mu) \psi^{(1,0)}(\mu, \phi) + \phi T(y) + \psi(\mu, \phi) \right] \\ \frac{\partial \ell}{\partial \gamma_T} &= \sum \left[ \psi^{(1,1)}(y - \mu) \frac{\partial \phi}{\partial \tau} \frac{\partial \tau}{\partial \gamma_s} \right. \\ &\quad \left. + T(y) \frac{\partial \phi}{\partial \tau} \frac{\partial \tau}{\partial \gamma_s} + \psi^{(0,1)} \frac{\partial \phi}{\partial \tau} \frac{\partial \tau}{\partial \gamma_s} \right], \end{aligned}$$

e agora derivando com relação  $\gamma_s$ , tem-se

$$\begin{aligned} \frac{\partial \ell}{\partial \gamma_s} &= \left\{ \sum \left[ \psi^{(1,1)}(y - \mu) \frac{\partial \phi}{\partial \tau} \frac{\partial \tau}{\partial \gamma_s} \right. \right. \\ &\quad \left. \left. + T(y) \frac{\partial \phi}{\partial \tau} \frac{\partial \tau}{\partial \gamma_s} + \psi^{(0,1)} \frac{\partial \phi}{\partial \tau} \frac{\partial \tau}{\partial \gamma_s} \right] \right\}, \end{aligned}$$

e após obter as derivadas de cada termo, a Esperança e, sabendo por Dey et al (1997) que a  $E(T(y)) = -\psi^{(0,1)}(\mu, \phi)$ , o cumulante é dada por:

$$\kappa_{ST} = \sum \psi^{(0,2)} \phi_1^2(S, T).$$

1. ii)  $\kappa_{ST}^{(U)} = \frac{\partial \kappa_{ST}}{\partial \gamma_U} = \sum \psi^{(0,2)} \left( \frac{\partial \phi}{\partial \tau} \right)^2 \frac{\partial \tau}{\partial \gamma_R} \frac{\partial \tau}{\partial \gamma_S}$  e após o cálculo das derivadas obtém-se

$$\kappa_{ST}^{(U)} = \sum \left\{ \begin{aligned} &\psi^{(0,3)} \phi_1^3(S, T, U) + 2\psi^{(0,2)} \phi_1 \phi_2(S, T, U) \\ &+ \psi^{(0,2)} \phi_1^2(ST, U + UT, R) \end{aligned} \right\}$$

$$iii) \kappa_{STU} = E \left[ \frac{\partial \ell^3(\beta, \phi)}{\partial \gamma_s \partial \gamma_T \partial \gamma_U} \right] = E \left[ \frac{\partial \ell}{\partial \gamma_s} \left( \frac{\partial^2 \ell}{\partial \gamma_T \partial \gamma_U} \right) \right] = E \left[ \frac{\partial \ell}{\partial \gamma_s} (\mathbf{U}_{TU}) \right].$$

Portanto faz-se necessário calcular a  $U_{TU}$ , a qual está apresentada a seguir

$$\begin{aligned} \frac{\partial \ell}{\partial \gamma_s} &= \frac{\partial}{\partial \gamma_s} (\mathbf{U}_{TU}) \\ &= \frac{\partial}{\partial \gamma_s} \left\{ \sum T(y) \left[ \frac{\partial^2 \phi}{\partial \tau^2} \frac{\partial \tau}{\partial \gamma_T} \frac{\partial \tau}{\partial \gamma_U} + \frac{\partial \phi}{\partial \tau} \frac{\partial^2 \tau}{\partial \gamma_T \partial \gamma_U} \right] \right. \\ &\quad \left. + \sum \psi^{(0,2)} \left( \frac{\partial \phi}{\partial \tau} \right)^2 \frac{\partial \tau}{\partial \gamma_T} \frac{\partial \tau}{\partial \gamma_U} \right. \\ &\quad \left. + \sum \psi^{(0,1)} \left[ \frac{\partial^2 \phi}{\partial \tau^2} \frac{\partial \tau}{\partial \gamma_T} \frac{\partial \tau}{\partial \gamma_U} + \frac{\partial \phi}{\partial \tau} \frac{\partial^2 \tau}{\partial \gamma_T \partial \gamma_U} \right] \right\}, \end{aligned}$$

fazendo as exaustivas derivadas com relação a  $\gamma_s$  de cada termo e, em seguida obter as suas respectivas esperanças, tem-se o cumulante

$$\begin{aligned} \kappa_{STU} = & \sum \{ \psi^{(0,3)} \phi_1^3 + 3\psi^{(0,2)} \phi_1 \phi_2 (S, T, U) \\ & + \psi^{(0,2)} \phi_1^2 [(ST, U) + (SU, T) + (UT, R)] \}. \end{aligned}$$

1. iv)  $\kappa_{Stu} = E \left[ \frac{\partial \ell^3(\beta, \phi)}{\partial \gamma_s \partial \beta_t \partial \beta_u} \right] = E \left[ \frac{\partial \ell}{\partial \gamma_s} \left( \frac{\partial^2 \ell}{\partial \beta_t \partial \beta_u} \right) \right]$ , logo é necessário calcular as derivadas com relação aos  $\beta^i$ s, as quais são descritas como se segue:

$$\begin{aligned} \frac{\partial \ell}{\partial \gamma_s} \left( \frac{\partial^2 \ell}{\partial \beta_t \partial \beta_u} \right) = & \frac{\partial \ell}{\partial \gamma_s} \left\{ \sum (\psi^{(3,0)} (y - \mu) \left( \frac{\partial \mu}{\partial \eta} \right)^2 \frac{\partial \eta}{\partial \beta_t} \frac{\partial \eta}{\partial \beta_u} \right. \\ & - \psi^{(2,0)} \left( \frac{\partial \mu}{\partial \eta} \right)^2 \frac{\partial \eta}{\partial \beta_s} \frac{\partial \eta}{\partial \beta_t} \\ & + \psi^{(2,0)} (y - \mu) \frac{\partial^2 \mu}{\partial \eta^2} \frac{\partial \eta}{\partial \beta_s} \frac{\partial \eta}{\partial \beta_t} \\ & \left. + \psi^{(2,0)} (y - \mu) \frac{\partial \mu}{\partial \eta} \frac{\partial^2 \eta}{\partial \beta_s \partial \beta_t} \right\}, \end{aligned}$$

e calculando as derivadas com relação  $\gamma_s$  para cada termo do somatório e em seguida obter suas respectivas esperanças resulta no cumulante

$$\kappa_{Stu} = \sum -\psi^{(2,1)} \phi_1 m_1^2 (S, t, u).$$

# APÊNDICE B

## PROGRAMA DO CÁLCULO DO VIÉS DE $O(n^{-1})$ PARA OS MLG.

```
/* Programa Executado para Ilustrar o Uso da Expressão do
Cálculo do Viés de Primeira Ordem dos Estimadores dos
Modelos Lineares Generalizados
Aplicação: Vendas de Telhas
Modelo: Inversa Gaussiana
Função de ligação: Log */

data vendas;
  input total gasto clientes concor potencial ;

cards;

DADOS DA APLICAÇÃO ;

proc genmod data = vendas; MAKE
'PARAMETERESTIMATES' OUT = EST(KEEP=ESTIMATE);
  model total = gasto clientes concor potencial / dist = ig
          link = log ;
output out = peso hesswgt = matw xbeta = neta;
quit;
run;

/* Implementação das Matrizes Necessárias para o
Cálculo do Viés */
```

```

proc iml;
use est;
read all;
estimativas = estimate;
close est;
print estimativas;
use peso; read all; print matw;
close peso;
um = j(26,1,1);
use vendas;
read all;
x = um ||gasto||clientes|| concor|| potencial;
close vendas;
inffisher = t(X)*diag(matw)*X;
invfisher = inv(inffisher);
ZD = X*invfisher*t(X);
mu = exp(neta);
parte1 = -(1/2)#invfisher;
parte2 = t(x)*diag(zd)*diag(mu);
biasbeta = parte1*parte2*um;
print biasbeta;

```

# APÊNDICE C

## PROGRAMA DO CÁLCULO DO VIÉS $O(n^{-1})$ PARA OS MODELOS NÃO LINEARES

```
/*Programa Executado para Ilustrar o Uso da Expressão do  
Cálculo do Viés de Primeira Ordem dos Estimadores dos  
Modelos Não Lineares da Família Exponencial  
Aplicação: Energia Elétrica  
Modelo Adotado: Gama  
Função de Ligação: Recíproca */
```

```
data eletrica;
```

```
input Elar Per Pgr Reca ;  
cards; DADOS DA APLICAÇÃO ;
```

```
data betas;
```

```
A0 = 0;  
beta0 = 2.3465951;  
beta1 = -1.234926;  
beta2 = 0.4703;
```

```
RUN;
```

```
%macro execucao(dados1,dados2);
```

```
%do i = 1 %to 25;
```

```
/* Obtenção das estimativas utilizando o procedimento OFFSET apresentado  
por Cordeiro e Paula (1989)*/
```

```
proc iml; use betas; read all var{A0 beta0 beta1 beta2}; close
```

```
betas; use eletrica; read all var{Elar Per Pgr Reca }; close
```



```

eletrica;

ETA = PER##BETA0#PGR##BETA1#RECA##BETA2;
XTIL2 = PER##BETA0#LOG(PER)#PGR##BETA1#RECA##BETA2;
XTIL3 = PER##BETA0#LOG(PGR)#PGR##BETA1#RECA##BETA2;
XTIL4 = PER##BETA0#LOG(RECA)#PGR##BETA1#RECA##BETA2;
MU = 1/ETA;
JBETA = BETA0#XTIL2+BETA1#XTIL3+beta2#xtil4;
TAL = ETA-JBETA;
nomes = {"ELAR" "xtil2" "xtil3" "xtil4" "tal"};
vetor = ELAR || XTIL2 || XTIL3 ||xtil4|| tal;
create dados from vetor[colname = nomes];
append from vetor;
quit;

proc genmod data = dados;

make 'Parameterestimates' out = EST(KEEP = Estimate); model ELAR =
XTIL2 XTIL3 xtil4 / dist = GAMMA noint
link = POW(-1)
offset = TAL;

run;

proc iml;
use est;
read all;
name = {"A0" "beta0" " beta1" " beta2" "FI"};
estimativas = t(estimate);
print (&i) estimativas;
create betas from estimativas[colname=name];
append from estimativas;
%end;
%mend execucao;
%execucao(eletrica, betas);
proc genmod data = dados;
make 'Parameterestimates' out = EST(KEEP = Estimate);
model ELAR = XTIL2 XTIL3 xtil4/ dist = GAMMA noint
link = POW(-1)
offset = TAL;
output out = peso hesswgt = matw xbeta = neta;

```

```

quit; run;
/* Implementação das Matrizes Necessárias para o
    Cálculo do Viés */

proc iml; use ELETRICA; read all var{Elar Per Pgr Reca }; close
ELETRICA; n = 92; um = j(n,1,1); traco = j(n,1,0); use est; read
all; name = {"beta0" " beta1" "beta2" " beta3" "fi"}; estimativas =
t(estimate); betaest =
estimativas[2]||estimativas[3]||estimativas[4]; fi =
estimativas[5]; close est; use peso; read all; close peso; matxtil
= XTIL2|| XTIL3||xtil4; fisher = t(matxtil)*diag(matw)*matxtil;
invfisher = inv(fisher); zd = matxtil*invfisher*t(matxtil); F =
-1/neta##3; dernetab11 =
per##estimativas[2]#log(per)##2#pgr#estimativas[3]#reca#estimativas[4];
dernetab12 =
per##estimativas[2]#log(per)#pgr##estimativas[3]#log(pgr)#reca##estimativas[4];
dernetab13 =
per##estimativas[2]#log(per)#pgr##estimativas[3]#log(reca)#reca##estimativas[4];
dernetab22 =
per##estimativas[2]#pgr##estimativas[3]#log(pgr)##2#reca##estimativas[4];
dernetab23 =
per##estimativas[2]#log(reca)#pgr##estimativas[3]#log(pgr)#reca##estimativas[4];
dernetab33 =
per##estimativas[2]#pgr##estimativas[3]#log(reca)##2#reca##estimativas[4];

DO I=1 TO N;
    traco[I] = DERNETAB11[I]*INVFISHER[1,1]+DERNETAB12[I]*INVFISHER[2,1]
        +DERNETAB13[I]*INVFISHER[3,1]+DERNETAB12[I]*INVFISHER[1,2]
        +DERNETAB22[I]*INVFISHER[2,2]+DERNETAB13[I]*INVFISHER[3,1]
        +DERNETAB13[I]*INVFISHER[1,3]+DERNETAB33[I]*INVFISHER[3,3]
        +DERNETAB23[I]*INVFISHER[3,2];
end;
d = diag(traco); dw = diag(matw); invdw = 1/vecdiag(dw); zdinvdw =
vecdiag(zd)#invdw#F; quisi1 = -0.5/fi#zdinvdw; quisi2 =
-0.5/fi#d*um; quisi = quisi1 + quisi2; bias =
invfisher*t(matxtil)*diag(matw)*quisi;

corrigido = t(betaest)-bias; beta=t(betaest); print beta bias
corrigido; quit

```

# APÊNDICE D

## PROGRAMA DO CÁLCULO DO VIÉS $O(n^{-1})$ PARA OS MLGS

```
/* Programa Executado para Ilustrar o Uso da Expressão do
Cálculo do Viés de Primeira Ordem dos Estimadores dos
Modelos Lineares Generalizados Superdispersados
Aplicação: Contagem do Número de Endoparasitas em Ovinos
Modelo Adotado: Passeio Aleatório
Função de Ligação: Logarítmica para ambos Submodelos */
```

```
data ovinos(drop = beta0 beta1 beta2 gama1);
input gg gg1 gg2 mes mes2 op ;
```

```
cards; DADOS DA APLICAÇÃO ;
run;
```

```
/* Determinar as estimativas do
sub modelo média na Poisson Dupla*/
```

```
proc nlmixed data = ovinos cov ; parms beta0 = 6.3, beta1 = .16,
beta2 = -.52; gama1 = 7;
mi = exp(beta0 + beta1*gg1 + beta2*gg2)+.0001 ;
fi = exp( gama1*mes2)+.00001;
L = (op-mi)*(fi+fi*log(mi))-fi*op*log(op)
+mi*fi*log(mi)+.5*log(fi);
model op~general(1);
```

```
ods output ParameterEstimates = emvsbeta; ods output
```

```
CovMatParmEst= covarbeta; run;
```

```
/* Determinar as estimativas do submodelo dispersão na Poisson  
Dupla */
```

```
proc nlmixed data=ovinos cov ; parms  
  gama1 = 7; beta0 = 6.3; beta1 = .36; beta2 = -.30;  
    mi = exp(beta0 + beta1*gg1 + beta2*gg2)+.0001 ;  
    fi = exp( gama1*mes2)+.0001;  
    L = (op-mi)*(fi+fi*log(mi))-fi*op*log(op)  
        +mi*fi*log(mi)+.5*log(fi);  
  model op~general(1);
```

```
ods output ParameterEstimates = emvsgama; ods output CovMatParmEst  
= covargama; run;
```

```
* Matrizes necessárias para o cálculo dos vieses das estimativas dos  
  betas e gama *;
```

```
proc iml; use ovinos; read all ; close ovinos;  
  use emvsbeta; read all into betas;  
  use emvsgama; read all into gamas;  
  use covarbeta; read all into invdebeta;  
  use covargama; read all into invdegama;  
  n=541;
```

```
    um      =    j(n,1,1);  
    s2      =    mes2;  
  
    beta    =    betas[ ,1];  
    beta0   =    beta[1];  
    beta1   =    beta[2];  
    beta2   =    beta[3];  
    gama    =    gamas[ ,1];  
    gama1   =    gama[1];  
    tbeta   =    t(beta);  
    tgama   =    t(gama);  
    matx    =    um||gg1||gg2;  
    mats    =    s2;  
    beta    =    beta0//beta1//beta2;  
    tbeta   =    t(beta);
```

```

gama      = gama0//gama1;
tgama     = t(gama);
sgama     = mats*gama;
xbeta     = matx*beta;
fi        = exp(sgama);
mi        = exp(xbeta);

L = (op-mi)#(fi+fi#log(mi))-fi#op#log(op)
    +mi#fi#log(mi)+.5#log(fi);
soma=t(um)*1;
aicsoma = -2*soma+10; bic = aicsoma+3*log(38); print aicsoma bic;

* Obter os vieses dos betas *;
pisi20    = diag(fi/mi);
miqua     = diag(mi##2);
invkdebeta = inv(t(matx)*pisi20*miqua*matx);
zdebeta   = diag(matx*invkdebeta*t(matx));
quisivies = -0.5#(zdebeta)*um;

quisivies = -0.5#(zdebeta)*um;
xpisi20miqua = t(matx)*pisi20*diag(mi##2);
viesdebeta = invkdebeta*xpisi20miqua*quisivies;

betacorrigido=beta-viesdebeta;
print beta viesdebeta betacorrigido;
* Obter o vies do gama *;
fiqua     = fi##2;
pisi02    = -.5/fiqua;
pisi02fiqua = diag(fiqua#pisi02);
invpisi02fiqua = inv(pisi02fiqua);
invkdegama = inv(t(mats)*pisi02fiqua*mats);
zdegama   = diag(mats*invkdegama*t(mats));
filcubo   = diag(fi##3);
psi03     = inv(filcubo);
pisi21    = diag(1/mi);
mifi      = diag(mi#fi);
quartzgama = zdegama#(.25);

zbetapisi21mifi = .5*zdebeta*mifi;

```

```
soma          = quartozgama + zbetapisi21mifi;
delta         = invpisi02fiqua*soma*um;
viesdegama    = invkdegama*t(mats)*pisi02fiqua*delta;
gamacorrigido=gama-viesdegama;
print gama viesdegama gamacorrigido;

quit;
```

# APÊNDICE E

## PROGRAMA DE SIMULAÇÃO

```
/* Programa de Simulação para a Geração e Cálculo
   dos Vieses dos Modelos Não Lineares Generalizados Superdispersados
   Modelo Adotado: Passeio aleatório
   Função de Ligação: Identidade para submodelo média e
                       log para o submodelo dispersão.  */

options nonotes nodetails; libname libref "c:\saidas"; libname out
"c:\saidas"; data out.estimativas; beta1 = 0; beta2 = 0; beta3 =
0; g1 = 0; g2 = 0; vbeta1 = 0; vbeta2 = 0; vbeta3 = 0; vg1 = 0;
vg2 = 0; cbeta1 = 0; cbeta2 = 0; cbeta3 = 0; cg1 = 0; cg2 = 0;
run;
* Geração das Matrizes Modelo X e S*;
* quantidade: Número de Amostra *;
* m: Tamanho da Amostra*;
%macro geracao(quantidade =, m =);
data out.dados(drop=i);
  do i=1 to &m;
    x1      = uniform(i+12);
    x2      = uniform(i+34);
    x3      = uniform(i+56);
    s1      = uniform(i+78);
    s2      = uniform(i+91);
    mu      = exp(1*x1+2*x2+3*x3);
    phi     = exp(exp(.1*s1+1.2*s2));
    output;
  end;
run;
%do i = 1 %to &quantidade;
```

```
/* A geração da distribuição Passeio Aleatório Y segue uma Inversa
Gaussiana com a transformação dada por  $y = 1/ig$  */
```

```
proc iml;
  use out.dados;
  read all;
  close out.dados;
  y = j(&m,1,0);
  do j = 1 to &m;
    y[j] = normal(j+&m+&i+123838318)**2;
  end;
  p1 = mu+mu##2#y/(2#phi);
  p2 = -mu/(2#phi)#sqrt(4*mu#phi#y+mu##2#y##2);
  x = p1+p2;
  ig = j(&m,1,0); u = ig;
  do j = 1 to &m;
    u[j] = uniform(j+&m+&i);
  end;
  do j = 1 to &m;
    if u[j] <= mu[j]/(mu[j]+x[j]) then ig [j] = x[j];
    else ig[j] = mu[j]##2/x[j];
    y[j]=1/ig[j];
  end;
  saida = ig||y||x1||x2||x3||s1||s2;
  create out.saida from saida;
  append from saida;
  quit;
```

```
/* Obtenção das estimativas do submodelo média */
```

```
proc nlp
  data = out.saida tech=tr cov=h vardef=n pcov noprint
  outest = out.emvsbeta(keep =beta1 beta2 beta3
    _type_ where =(_type_ in("PARMS","COV2: H")));
  max l;
  parms
  beta1 = 1,
  beta2 = 2,
  beta3 = 3;
  g1 = .1;
```



```

g2 = 1.2;
y = col1;
yrw = col2;
x1 = col3;
x2 = col4;
x3 = col5;
s1 = col6;
s2 = col7;
murw = 1/exp((beta1*x1+beta2*x2+beta3*x3))
      +1/exp(exp(g1*s1+g2*s2));
phirw = - exp(exp(g1*s1+g2*s2))/(2*exp(beta1*x1+beta2*x2+beta3*x3)**2);
pisimufi = 2*(-phirw)**.5*((2*murw-phirw)**.5-(-phirw)**.5)**(-1)
          +0.5*log(2)-log((2*murw-phirw)**.5-(-phirw)**.5)
          - murw*((2*murw-phirw)**.5-(-phirw)**.5)**(-2);
pisi10 = -1/((-2*murw-phirw)**.5+(-phirw)**.5)**2);
L = (yrw-murw)*pisi10+phirw*(1/yrw)+pisimufi;
run;

```

/\* Obtenção das estimativas do submodelo dispersão\*/

```

proc nlp data = out.saida tech = tr cov = h vardef = n pcov
noprnt outest = out.emvsgama(keep = g1 g2 _type_ where=( _type_
in("PARMS", "COV2: H")));
max 1;
parms
g1 = .1,
g2 = 1.2;
beta1 = 1;
beta2 = 2;
beta3 = 3;
y = col1;
yrw = col2;
x1 = col3;
x2 = col4;
x3 = col5;
s1 = col6;
s2 = col7;
murw = 1/exp((beta1*x1+beta2*x2+beta3*x3))
      +1/exp(exp(g1*s1+g2*s2));
phirw = - exp(exp(g1*s1+g2*s2))/(2*exp(beta1*x1

```

```

        +beta2*x2+beta3*x3)**2);
pisimufi = 2*(-phirw)**.5*((2*murw-phirw)**.5
        -(-phirw)**.5)**(-1)+0.5*log(2)
        -log((2*murw-phirw)**.5-(-phirw)**.5)
        -murw*((2*murw-phirw)**.5-(-phirw)**.5)**(-2);
pisi10 = -1/((-2*murw-phirw)**.5+(-phirw)**.5)**2);
        L = (yrw-murw)*pisi10+phirw*(1/yrw)+pisimufi;
run;
data out.estimatbeta(drop =_type_);
set out.emvsbeta; if _type_ ne "PARMS" then delete;
run;
data out.estimatgama(drop =_type_);
set out.emvsgama; if _type_ ne "PARMS" then delete;
run;
data out.hessian(drop =_type_);
set out.emvs; if _type_ = "PARMS" then delete;
run;

/*cálculo dos vieses dos estimativas betas e gamas*/
proc iml;
    use out.estimatbeta;
    read all into betas; /* vetor das estimativas dos betas*/

    use out.estimatgama;
    read all into gamas; /* vetor das estimativas dos gamas*/

    use out.hessian;
    read all into H; /* inversa da matriz hessiana*/

    use out.saida;
    read all into dados; /* dados gerados*/
    y      = dados[ ,1];
    yrw   = dados[ ,2];
    x1    = dados[ ,3];
    x2    = dados[ ,4];
    x3    = dados[ ,5];
    s1    = dados[ ,6];
    s2    = dados[ ,7];
    matx  = x1||x2||x3;
    mats  = s1||s2;

```

```

beta1 = betas[1];
beta2 = betas[2];
beta3 = betas[3];
g1 = gamas[1];
g2 = gamas[2];
beta = beta1//beta2//beta3;
tbeta = t(beta);
gama = g1//g2;
tgama = t(gama);
um = j(&m,1,1);
p=3;          q=2;
dimensao = p**2;
dimensaoq = q**2;
x2til = j(&m,dimensao,0);
vecfisher = j(dimensao,1,0);
traco = j(&m,1,0);

/* Cálculo dos vieses das estimativas dos betas*/

eta = matx*beta;
tal = mats*gama;
miig = exp(eta);
fiig = exp(exp(tal));
phirw = -fiig/(2#miig##2);
murw = (1/miig)+(1/fiig);
pisi20 = 2/(((2#murw-phirw)##.5
            -(-phirw)##.5)##3#(2#murw-phirw)##.5);
m1 = -1/miig##2;
m2 = 2/miig##3;
m1qua = m1##2;
xtil1 = x1#miig;
xtil2 = x2#miig;
xtil3 = x3#miig;
xtil = xtil1||xtil2||xtil3;
xtiltpisi20 = t(xtil)*diag(pisi20#m1qua);
kdebeta = xtiltpisi20*xtil;
detfisherbeta = abs(det(kdebeta));

```

```

if detfisherbeta <= 0.1 then
do;

```

```

    invkdebeta = ginv(kdebeta);
end;
else do;
    invkdebeta = inv(kdebeta);
end;
    zdebeta = diag(xtil*invkdebeta*t(xtil));
    K=0;
do i = 1 to p;
    do j = 1 to p;
        k = k+1;
        x2til[ ,k] = (matx[ ,i]#matx[ ,j])#exp(eta);
        vecfisher[k] = invkdebeta[i,j];
    end;
end;
do i = 1 to &m;
    traco[i] = x2til[i, ]*vecfisher;
end;
    invm1 = 1/m1;
    quisi1 = -.5*zdebeta*diag(invm1#m2)*um;
    quisi2 = -.5#traco;
    quisi = quisi1+quisi2;
    biasdebeta = invkdebeta*xtiltpisi20*quisi;
betacorrigido= beta- biasdebeta;
tbiasdebeta = t(biasdebeta);
tbetacorrigido = t(betacorrigido);

/* Cálculo do vieses das estimativas dos gamas */

fiqua = phirw##2;
ficubo = phirw##3;
fiquatro = phirw##4;
miqua = murw##2;
micubo = murw##3;
sqrtdoismuphi = (2#murw-phirw)##.5;
sqrtphi = (-phirw)##.5;
miigqua = miig##2;
fi1 = -.5*fiig/miigqua;
fi2 = fi1;
filcubo = fi1##3;
filqua = fi1##2;

```

```

pisi02          =  -((-2#sqrtdoismuphi#phirw+2#sqrtphi#phirw
                    +murw#sqrtdoismuphi-3#sqrtphi#murw)/(phirw#
                    (-sqrtdoismuphi+sqrtphi))##3#sqrtphi#sqrtdoismuphi));
pisi02filqua   =  pisi02#filqua;
diagpisi02filqua=  diag(pisi02filqua);
pisi03          =  -2#(4#fiquatro-18#ficubo#murw+4#ficubo#
                    sqrtdoismuphi#sqrtphi-14#sqrtphi#murw#
                    fiqua#sqrtdoismuphi+26#fiqua#miqua-12#
                    micubo#phirw+14#sqrtphi#miqua#phirw#
                    sqrtdoismuphi-3#micubo#sqrtdoismuphi#
                    sqrtphi)/(ficubo#(-2#murw+phirw))##2#
                    (-sqrtdoismuphi+sqrtphi))##4);
pisi03filcubo  =  pisi03#filcubo;
diagpisi03filcubo=  diag(pisi03filcubo);
pisi21          =  -(2#(-4#phirw#sqrtdoismuphi+4#phirw#
                    sqrtphi+3#murw#sqrtdoismuphi-7#murw#
                    sqrtphi))/(sqrtdoismuphi#(-2#murw+phirw)
                    #sqrtphi#(-sqrtdoismuphi+sqrtphi))##5);
pisi21fi       =  pisi21#phirw;
stil1          =  fiig#s1;
stil2          =  fiig#s2;
stil           =  stil1||stil2;
prodpis02filqua =  pisi02#filqua;
invprodpis02filqua = 1/prodpis02filqua;
stiltpisi02    =  t(stil)*diag(prodpis02filqua);
kdegama        =  -stiltpisi02*stil;
detfishergama  =  abs(det(kdegama));
    if          detfishergama <=0.1
        then do;
            invkdegama = ginv(kdegama);
            end;
        else do;
            invkdegama =  inv(kdegama);
            end;
zdegama        =  diag(stil*invkdegama*t(stil));
film1qua      =  diag(fil#m1qua);
delta1         =  -.5*diag(invprodpis02filqua)*(zdegama*
                    (diagpisi03filcubo+diagpisi02filqua)
                    +zdebeta*diag(pisi21)*film1qua);
s2til         =  j(&m,dimensaoq,0);

```

```

vecfishergama = j(dimensaoq,1,0);
tracogama     = j(&m,1,0);
um            = j(&m,1,1);
K=0;
do i = 1 to q;
  do j = 1 to q;
    k = k+1;
    s2til[ ,k] = (mats[ ,i]#mats[ ,j])#fiig;
    vecfishergama[k]= invkdegama[i,j];
  end;
end;
do i = 1 to &m;
  tracogama[i] = s2til[i, ]*vecfishergama;
end;
delta2 = -.5*diag(tracogama);
delta  = (delta1+delta2)*um;
biasdegama = invkdegama*stiltpsi02*delta;
gamacorrigido = gama- biasdegama;
tbiasdegama = t(biasdegama);
tgamacorrigido = t(gamacorrigido);

nomes = {"beta1" "beta2" "beta3" "g1" "g2" "vbeta1"
         "vbeta2" "vbeta3" "vg1" "vg2" "cbeta1"
         "cbeta2" "cbeta3" "cg1" "cg2"};
dados1 = tbeta||tgama||tbiasdebeta||tbiasdegama||
         tbetacorrigido||tgamacorrigido;

create out.saidapar from dados1[colname=nomes]; append from
dados1; quit; proc append base = out.estimativas data =
out.saidapar;
run;
data out.est;
set out.estimativas;
if beta1 = 0 then delete;
run;
%put &i;
%end;

/* Geração das 10.000 simulações para os tamanhos de amostra 15 */
%mend geracao;

```

```
%geracao(quantidade = 10000, m = 15);  
proc means data = out.est; run;
```

# APÊNDICE F

## PROGRAMA PARA O CÁLCULO DO VIÉS DOS MNLGS

Aplicação 1 - Contagem de erros

```
/* Programa Executado para Cálculo dos Vieses dos Modelos
   Não Lineares Generalizados Superdispersados
   Aplicação 1: Número de erros cometidos por
               ratos lesionados e não lesionados
   Modelo adotado: Poisson dupla
   Função de Ligação: Identidade para o submodelo média e
                     log para o submodelo dispersão*/
```

```
data ratos;
```

```
input tempo erros doente sadio;
```

```
cards; DADOS DA APLICACAO ; proc nlmixed data=ratos cov; parms
beta0=5,beta2 = 1.278, beta1=0.52;
```

```
mi = beta0*doente+((sadio*(exp(beta1*tempo)))+(doente*(exp(beta2*tempo))))
```

```
pisi10=log(mi)+1;
```

```
pisimifi= mi*log(mi);
```

```
erro1=erros+.0000001;
```

```
L = (erro1-mi)*pisi10+pisimifi;
```

```
model erro1~general(1);
```

```
run;
```

```
proc nlmixed data = ratos cov;
```

```
parms beta0 = 5, beta2 = 1.278, beta1 = 0.52;
```



```

gama0 = 1.847; gama1 = -.00661;
mi      = beta0*doente+((sadio*(exp(beta1*tempo)))
      +(doente*(exp(beta2*tempo)))) ;
fi      = exp(gama0+gama1*tempo);
erro1   = erros+.0000001;
L       = (erro1-mi)*(fi+fi*log(mi))+fi*erro1*log(erro1)
      +mi*fi*log(mi)+.5*log(fi);
      model erro1~general(1);
ods output ParameterEstimates = emvsbeta; ods output CovMatParmEst
= invbeta; run; proc nlmixed data = ratos cov; parms
gama0 = 1.847, gama1 = -.00661;
beta0 = 5; beta2 = 1.278; beta1 = 0.52;
      mi = beta0*doente+((sadio*(exp(beta1*tempo)))
      +(doente*(exp(beta2*tempo)))) ;
      fi = exp(gama0+gama1*tempo);
      erro1 = erros+.0000001;
      L   = (erro1-mi)*(fi+fi*log(mi))+fi*erro1*log(erro1)
      +mi*fi*log(mi)+.5*log(fi);
      model erro1~general(1);
ods output ParameterEstimates = emvsgama; ods output CovMatParmEst
= invgama; run;
* C\U{e1}lculo dos vieses para o modelo Poisson Dupla*;
proc iml;
  use ratos; read all ; close ratos;
  use emvsbeta ; read all into betas;
  use emvsgama;read all into gamas;
  use invbeta; read all into invdebeta;
  use invgama; read all into invdegama;

n=105;

invkdegama = invdegama[,2]||invdegama[,3];
  s1      = j(n,1,1);
  s2      = tempo;
invkdebeta = invdebeta[ ,2]||invdebeta[ ,3]||invdebeta[ ,4];
invkdegama = invdegama[ ,2]||invdegama[ ,3];
  beta    = betas[ ,1];
  beta0   = beta[1];
  beta1   = beta[2];
  beta2   = beta[3];

```

```

gama      =  gamas[ ,1];
gama0     =  gama[1];
gama1     =  gama[2];
tbeta     =  t(beta);
tgama     =  t(gama);
matx      =  doente||sadio||tempo;
mats      =  s1||s2;
beta      =  beta0//beta1//beta2;
tbeta     =  t(beta);
gama      =  gama0//gama1;
tgama     =  t(gama);
sgama     =  mats*gama;
fi        =  exp(sgama);
mi        =  beta0#doente+((sadio#(exp(beta1#tempo)))
                    +(doente#(exp(beta2#tempo)))) ;
xtil1     =  doente;
xtil2     =  sadio#tempo#exp(beta1#tempo);
xtil3     =  doente#tempo#exp(beta2#tempo);
xtil      =  xtil1||xtil2||xtil3;
pisi20    =  fi/mi      ;
xtiltpisi20  =  t(xtil)*diag(pisi20);
kdebeta   =  xtiltpisi20*xtil;
zdebeta   =  diag(xtil*invkdebeta*t(xtil));
x2til22   =  sadio#(tempo##2)#exp(beta1#tempo);
x2til33   =  doente#(tempo##2)#exp(beta2#tempo);
traco     =  x2til22*invkdebeta[2,2]+x2til33*invkdebeta[3,3];
quisi2    =  -.5#traco;
biasdebeta =  invkdebeta*xtiltpisi20*quisi2;
betacorrigido =  beta- biasdebeta;
tbiasdebeta =  t(biasdebeta);
tbetacorrigido =  t(betacorrigido);
fiqua     =  fi##2;
pisi02    =  -1/(2#fiqua);
invpisi02 =  1/pisi02;
pisi02fiqua =  pisi02#fiqua;
pisi03    =  1/(fiqua#fi);
pisi21    =  1/mi;
stil1     =  fi;
stil2     =  fi#s2;
stil      =  stil1||stil2;

```

```

stiltpisi02    =  t(stil)*diag(pisi02);
kdegama       =  stiltpisi02*stil;
zdegama       =  diag(stil*invkdegama*t(stil));
delta1        =  -.5*inv(diag(pisi02))*(zdegama*(diag(pisi03)
+zdebeta*diag(pisi21)));

s2til11       =  fi;
s2til12       =  s2#fi;
s2til22       =  s2##2#fi;
tracogama     =  s2til11#invkdegama[1,1]
                +2*s2til12*invkdegama[1,2]
                +s2til22*invkdegama[2,2];

delta2        =  -.5*diag(tracogama);
delta         =  (delta1+delta2)*s1;
biasdegama    =  invkdegama*stiltpisi02*delta;
gamacorrigido =  gama- biasdegama;
tbiasdegama   =  t(biasdegama);
tgamacorrigido =  t(gamacorrigido);
print beta biasdebeta betacorrigido;
print gama biasdegama gamacorrigido;

quit;

```

Tabela F.1: Média dos erros cometidos pelos ratos isquêmicos(lesionados)

Rato	Bl <sub>1</sub>	Bl <sub>2</sub>	Bl <sub>3</sub>	Bl <sub>4</sub>	Bl <sub>5</sub>	Rato	Bl <sub>1</sub>	Bl <sub>2</sub>	Bl <sub>3</sub>	Bl <sub>4</sub>	Bl <sub>5</sub>
1	0,5	1,2	0,3	0,5	0,3	14	3,2	2,6	2,7	0,8	0,3
2	0,5	1	0,6	0	0	15	2,2	1,9	0,3	0	0
3	0,3	0,5	0,2	0,6	0,1	16	1,9	0,3	0,4	0,2	0,2
4	0,3	0,4	0	0	0	17	0,9	1,8	1,1	1	0,6
5	0,8	0,1	0	0	0,1	18	1,8	1,9	0,8	0,2	0,6
6	1,3	1	0,5	0	0,6	19	3,2	2,6	2,7	0,8	0,3
7	0,8	0,1	0,2	0	0	20	2,2	1,9	0,3	0	0
8	2,3	0,4	0	0	0	21	1,9	0,3	0,4	0,2	0,2
9	0,6	0,8	0	0	0	22	0,9	1,8	1,1	1	0,6
10	0,4	0	0	0	0	23	1,8	1,9	0,8	0,2	0,6
11	0,1	0,2	0,6	0	0	24	0,8	1	0	0,3	0
12	0,3	0,2	0,8	0	0	25	0,2	0,5	0,2	0,2	0,1
13	0,2	0	0,9	0,1	0						

Tabela F.2: Média dos erros cometidos pelos ratos não-lesionados

Rato	Bl <sub>1</sub>	Bl <sub>2</sub>	Bl <sub>3</sub>	Bl <sub>4</sub>	Bl <sub>5</sub>	Rato	Bl <sub>1</sub>	Bl <sub>2</sub>	Bl <sub>3</sub>	Bl <sub>4</sub>	Bl <sub>5</sub>
1	0,8	0	0	0	0	14	0,9	0	0,1	0	0
2	0,3	0,4	0,2	0	0	15	2,8	1,8	1	0	0,1
3	0,7	0	0	0	0	16	0,9	0,7	0	0	0
4	0,7	0,8	0	0	0	17	0,8	0,4	0,2	0	0
5	0,8	0,4	0,1	0	0	18	0,4	0,2	0,1	0	0
6	0,9	0,3	0,1	0	0	19	0,7	0,2	0	0	0
7	0,6	0,1	0	0	0	20	3,3	1,8	0,7	0,2	0
8	1,1	1,2	0,3	0,2	0	21	2,4	1,3	0	0	0
9	0,2	0,6	0,3	0	0	22	3,3	1,8	0,7	0,2	0
10	1,1	0	0	0	0	23	2,4	1,3	0	0	0
11	0,2	0,1	0	0	0	24	0,6	0,3	0,3	0	0
12	0,7	0,1	0	0	0,1	25	0,4	0	0,2	0,2	0
13	0,1	0	0,1	0	0	26	1,2	0,1	0	0	0

# APÊNDICE G

## PROGRAMA PARA O CÁLCULO DO VIÉS DOS MNLGS

Aplicação 2 - Concreto (normal)

```
/* Programa Executado para Cálculo dos Vieses dos Modelos
   Não Lineares Generalizados Superdispersados
   Aplicação 2: Resistência à Compressão do Concreto
   Modelo adotado: Normal
   Função de Ligação: Identidade e Log para ambos submodelos*/
```

```
data dados;
```

```
  input A B C D E Y;
```

```
  cards;
```

```
dados com uma repetição
```

```
; /* Estimando parâmetros usando como verosimilhança a normal direto*/
```

```
proc nlmixed data=dados maxiter=7000;
```

```
  parms
```

```
    betainterc=3.71,
```

```
    betaA    =-0.253,
```

```
    betaD    =-.055,
```

```
    betaE    =  .053,
```

```
    betaAC   =  .023,
```

```
    betaAE   = -0.051,
```

```
    betaBC   =  .022,
```

```
    gamainterc = -5.15,
```

```

gamaE      =   .39,
gamaAE     =   -.53;
AC         =   A*C;
AE         =   A*E;
BC         =   B*C;
mi         =   exp(betainterc+betaA*A+betaD*D+betaE*E+betaAC*AC
                  +betaAE*AE+betaBC*BC);
fi         =   exp(gamainterc+gamaE*E+gamaAE*AE);

L1=-.5*log(fi)-(.5/fi)* (y-mi)**2;
model y ~ general(L1);
run;
/* Estimando parâmetros usando como vero a normal com enfoque dos
MNLGS*/

proc nlmixed data=dados cov maxiter=7000;
  parms
    betainterc=3.71,
    betaA    =-0.253,
    betaD    =-.055,
    betaE    =  .053,
    betaAC   =  .023,
    betaAE   = -0.051,
    betaBC   =  .022,
    gamainterc = -5.15,
    gamaE    =  .39;
    AC      =   A*C;
    AE      =   A*E;
    BC      =   B*C;
    mi      =   exp(betainterc+betaA*A+betaD*D+betaE*E+betaAC*AC
                  +betaAE*AE+betaBC*BC);
    fi      =   exp(gamainterc+gamaE*E);
    pisimufi=.5*fi*mi**2+.5*log(fi);
    pisi10=mi*fi;
    L      = (y-mi)*pisi10-.5*fi*y**2+pisimufi;
  model y ~ general(L);
  ods output ParameterEstimates=emvsbeta;
  ods output CovMatParmEst=covarbeta;
run;
/* Calculando o vi\U{ea}s*/ proc iml; use dados; read all ; close

```

```

dados;
use emvsbeta ; read all into betas;
use covarbeta; read all into covbetagama;
n=32;p=7;q=2;invkdebetasas=j(p,p,0);invkdegamasas=j(q,q,0);j=0;
beta=j(p,1,0);gama=j(q,1,0);
do i = 1 to p;
    beta[i] = betas[i,1];
    do j = i to p;
        invkdebetasas[i,j] = covbetagama[i,j+1];
        invkdebetasas[j,i] = invkdebetasas[i,j];
    end;end;
do i = 1 to q;
    gama[i] = betas[i+p,1];
    do j = i to q;
        invkdegamasas[i,j] = covbetagama[i+p,j+p+1];
        invkdegamasas[j,i] = invkdegamasas[i,j];
    end;end;
tbeta = t(beta);
tgama = t(gama);
um = J(N,1,1);
ac = a#c;
ae = a#e;
bc = b#c;
matx = um||a||d||e||ac||ae||bc;
mats = um||e;
sgama = mats*gama;
xbeta = matx*beta;
fi = exp(sgama);
mi = exp(xbeta) ;
xtil = j(n,p,0);stil = j(n,q,0);
/* Calculando matriz xtil*/
do i=1 to p;
    xtil[,i] = matx[,i]#mi;
end;
/* Calculando matriz stil*/
do i=1 to q;
    stil[,i] = mats[,i]#fi;
end;

/* Calculando matriz inversa de Fisher para beta*/

```

```

pisi20 = fi      ;
xtiltpisi20 = t(xtil)*diag(pisi20);
kdebeta = xtiltpisi20*xtil;
invkdebeta = inv(kdebeta);
zdebeta = diag(xtil*invkdebeta*t(xtil));
zdebetasas = diag(xtil*invkdebetasas*t(xtil));
/* Calculando matriz inversa de fisher para gama*/
pisi02 = -.5/fi##2;
stiltpisi02 = t(stil)*diag(pisi02);
kdegama = -stiltpisi02*stil;
invkdegama = inv(kdegama);

/* Calculando matriz x2til*/
pqua = p**2;
qqua = q**2;
x2til = j(n,pqua,0);s2til = j(n,qqua,0);
vecfisherbeta = j(pqua,1,0); vecfishergama = j(qqua,1,0);
tracobeta = j(n,1,0); tracogama = j(n,1,0);
vecfisherbetasas = j(pqua,1,0); vecfishergamasas = j(qqua,1,0);
tracobetasas = j(n,1,0); tracogamasas = j(n,1,0);
K=0;
do i = 1 to p;
  do j= 1 to p;
    k = k+1;
    x2til[ ,k] = (matx[ ,i]#matx[ ,j])#mi;
    vecfisherbeta[k]= invkdebeta[i,j];
    vecfisherbetasas[k]= invkdebetasas[i,j];

  end;
end;
do i = 1 to n;
  tracobeta[i] = x2til[i, ]*vecfisherbeta;
  tracobetasas[i] = x2til[i, ]*vecfisherbetasas;
end;
/* Calculando matriz s2til*/
K=0;
do i = 1 to q;
  do j = 1 to q;
    k = k+1;
    s2til[ ,k] = (mats[ ,i]#mats[ ,j])#fi;
  end;
end;

```



```

        vecfishergama[k]=   invkdegama[i,j];
        vecfishergamasas[k]=   invkdegamasas[i,j];
    end;
end;
do   i= 1   to  n;
        tracogama[i] =   s2til[i,]*vecfishergama;
        tracogamasas[i] =   s2til[i,]*vecfishergamasas;
end;
/* Calculando vies de beta*/
quisi2           =   -.5#tracobeta;
biasdebeta       =   invkdebeta*xtiltpisi20*quisi2;
betacorrigido    =   beta- biasdebeta;
quisi2sas        =   -.5#tracobetasas;
biasdebetasas   =   invkdebetasas*xtiltpisi20*quisi2sas;
betacorrigidosas =   beta- biasdebetasas;
tbiasdebeta     =   t(biasdebeta);
tbetacorrigido  =   t(betacorrigido);
invpisi02       =   1/pisi02;
pisi02fiqua     =   pisi02#fiqua;
pisi03          =   1/(fi##3);
zdegama         =   diag(stil*invkdegama*t(stil));
zdegamasas      =   diag(stil*invkdegamasas*t(stil));
delta1          =   -.5*inv(diag(pisi02))*(zdegama*(diag(pisi03)));
delta2          =   -.5*diag(tracogamasas);
delta           =   (delta1+delta2)*um;
biasdegama      =   invkdegama*stiltipisi02*delta;
gamacorrigido   =   gama- biasdegama;

delta1sas       =   -.5*inv(diag(pisi02))*(zdegamasas*(diag(pisi03)));
delta2sas       =   -.5*diag(tracogamasas);
deltasas       =   (delta1sas+delta2sas)*um;
biasdegamasas  =   invkdegamasas*stiltipisi02*deltasas;
gamacorrigidosas =   gama- biasdegamasas;

tbiasdegama     =   t(biasdegama);
tgamacorrigido  =   t(gamacorrigido);

print  beta biasdebeta betacorrigido;
print  gama biasdegama gamacorrigido;

```

```

print beta biasdebetasas betacorrigidosas;
print gama biasdegamasas gamacorrigidosas;
quit;

```

Tabela G.1: Resistência à compressão (MPa) de corpos-de-prova de concreto.

CE	Fatores <sup>(1)</sup>					Y <sub>1</sub>	Y <sub>2</sub>	Y <sub>3</sub>
	A	B	C	D	E			
1	-1	-1	-1	-1	-1	57,30	55,90	59,84
2	-1	-1	-1	-1	1	65,57	60,10	54,11
3	-1	-1	-1	1	-1	48,38	50,93	54,75
4	-1	-1	-1	1	1	52,33	54,37	40,74
5	-1	-1	1	-1	-1	53,35	49,85	45,45
6	-1	-1	1	-1	1	62,71	38,71	59,21
7	-1	-1	1	1	-1	45,20	45,58	56,34
8	-1	-1	1	1	1	60,73	48,06	54,75
9	-1	1	-1	-1	-1	56,79	56,53	53,35
10	-1	1	-1	-1	1	49,53	55,39	46,47
11	-1	1	-1	1	-1	52,71	48,38	51,95
12	-1	1	-1	1	1	57,61	40,87	60,35
13	-1	1	1	-1	-1	60,03	57,04	56,66
14	-1	1	1	-1	1	66,84	60,81	62,39
15	-1	1	1	1	-1	46,60	44,56	53,99
16	-1	1	1	1	1	40,74	43,42	43,29
17	1	-1	-1	-1	-1	27,50	32,47	34,82
18	1	-1	-1	-1	1	38,20	38,20	37,56
19	1	-1	-1	1	-1	28,65	26,36	26,10
20	1	-1	-1	1	1	33,23	35,01	35,65
21	1	-1	1	-1	-1	32,78	27,88	29,92
22	1	-1	1	-1	1	35,90	41,51	36,16
23	1	-1	1	1	-1	26,10	29,28	26,35
24	1	-1	1	1	1	35,01	37,30	37,69
25	1	1	-1	-1	-1	28,52	31,07	26,10
26	1	1	-1	-1	1	37,56	32,47	35,20
27	1	1	-1	1	-1	23,94	22,92	26,74
28	1	1	-1	1	1	27,69	31,83	25,21
29	1	1	1	-1	-1	27,76	32,72	32,59
30	1	1	1	-1	1	34,12	34,06	39,98
31	1	1	1	1	-1	31,19	28,65	26,86
32	1	1	1	1	1	36,61	35,65	33,74

# APÊNDICE H

## PROGRAMA PARA O CÁLCULO DO VIÉS DOS MNLGS

Aplicação 2 - Concreto (passeio aleatório)

```
/* Programa executado para os Modelos
   Não Lineares Generalizados Superdispersados
   Aplicação 2: resistência à compressão do concreto
   Modelo adotado: Passeio aleatório
   Função de ligação: identidade para ambos submodelos
   Cálculo do vieses: Primeira ordem para os estimadores */
```

```
data dados;
  input A B C D E Y @@;
  cards;
    DADOS DO CONCRETO;
proc nlmixed data=dados cov maxiter=7000;
  parms
    betainterc=18,
    betaA   =-3,
    betaD   =.03,
    betaE   = -21,
    betaAC  = -.09,
    betaAE  = 4,
    betaBC  = -.11,
    gamainterc = -203,
    gamaE   = 206;
  AC      = A*C;
  AE      = A*E;
```

```

BC      = B*C;
teta    = exp(betainter+c+betaA*A+betaD*D+betaE*E+betaAC*AC
              +betaAE*AE+betaBC*BC);
delta   = exp(gamainter+c+gamaE*E);
mi      = 1/teta+1/delta;
fi      = -delta/(2*teta**2);
pisimufi = 2*(-fi)**.5*((2*mi-fi)**.5-(-fi)**.5)**(-1)
          +0.5*log(2)-log((2*mi-fi)**.5-(-fi)**.5)
          -mi*((2*mi-fi)**.5-(-fi)**.5)**(-2);
pisi10  = -1/((-2*mi-fi)**.5+(-fi)**.5)**2);
L1      = (y-mi)*pisi10+fi*(1/y)+pisimufi;
model y ~ general(L1);
ods output ParameterEstimates = emvs;

run;

```

\* Cálculo dos vieses para o modelo passeio aleatório\*;

```

proc iml;
  use dados;
  read all ; close dados; /* dados*/
  use emvs;
  read all into betas; /* vetor das estimativas*/
  n      = 96;
  p      = 7;
  q      = 2;
  beta   = j(p,1,0);
  gama   = j(q,1,0);
  AC     = A#C;
  AE     = A#E;
  BC     = B#C;
  um     = j(n,1,1);
  matx   = um||A||D||E||AC||AE||BC;
  mats   = UM||E;
  do i   = 1 to p;
    beta[i] = betas[i,1];
  end;
  do i   = 1 to q;
    gama[i] = betas[i+p,1];
  end;
  print beta;

```

```

print gama;
dimensao      = p**2;
dimensaoq     = q**2;
xtil          = j(n,p,0);
sttil         = j(n,q,0);
x2til        = j(n,dimensao,0);
vecfisher    = j(dimensao,1,0);
traco        = j(n,1,0);
eta           = matx*beta;
tal           = mats*gama;
teta         = exp(eta);
delta        = exp(tal);
phirw        = -delta/(2#teta##2);
murw         = (1/teta)+(1/delta);
doismu       = (2#murw-phirw)##.5;
sqrtfi       = (-phirw)##.5;
dif          = (doismu-sqrtfi)##3#doismu;
m1           = diag(-1/teta##2);
m1qua       = m1*m1;
m2           = diag(2/teta##3);
pisi20       = 2/dif;
xtil1        = um#teta;
xtil2        = a#teta;
xtil3        = d#teta;
xtil4        = e#teta;
xtil5        = ac#teta;
xtil6        = ae#teta;
xtil7        = bc#teta;
xtil         = xtil1||xtil2||xtil3||xtil4||xtil5||xtil6||xtil7;
xtiltpisi20  = t(xtil)*diag(pisi20)*m1qua;
kdebeta     = xtiltpisi20*xtil;
invkdebeta   = ginv(kdebeta);
zdebeta      = diag(xtil*invkdebeta*t(xtil));
K=0;
do i = 1 to p;
  do j = 1 to p;
    k = k+1;
    x2til[ ,k] = (matx[ ,i]#matx[ ,j])#exp(eta);
    vecfisher[k] = invkdebeta[i,j];
  end;
end;

```

```

end;
do i      = 1 to n;
  traco[i] = x2til[i, ]*vecfisher;
end;
prodm1m2 = diag(-2/teta);
quisi1    = -.5#zdebeta*prodm1m2*um;
quisi2    = -.5#traco;
quisi     = quisi1+quisi2;
biasdebeta = invkdebeta*xtiltpisi20*quisi;
betacorrigido = beta- biasdebeta;
tbiasdebeta = t(biasdebeta);
tbetacorrigido = t(betacorrigido);
fiqua      = phirw##2;
ficubo     = phirw##3;
fiquatro   = phirw##4;
miqua      = murw##2;
micubo     = murw##3;
sqrtdoismuphi = (2#murw-phirw)##.5;
sqrtphi    = (-phirw)##.5;
pisi02     = -((-2#sqrtdoismuphi#phirw+2#sqrtphi#phirw
+murw#sqrtdoismuphi-3#sqrtphi#murw)/(phirw#(
-sqrtdoismuphi+sqrtphi)##3#sqrtphi#sqrtdoismuphi));
fi1        = (-.5/teta##2);
filqua     = fi1#fi1;
prodpisi02 = (pisi02#filqua);
invpisi02  = prodpisi02;
diaginvpisi02 = diag( invpisi02);
pisi03     = -2*(4#fiquatro-18#ficubo#murw+4#ficubo#
sqrtdoismuphi#sqrtphi-14#sqrtphi#murw#
fiqua#sqrtdoismuphi+26#fiqua#miqua-12#
micubo#phirw+14#sqrtphi#miqua#phirw#
sqrtdoismuphi-3#micubo#sqrtdoismuphi#
sqrtphi)/(ficubo#(-2#murw+phirw)##2#
(-sqrtdoismuphi+sqrtphi)##4);

diagpisi03 = diag(pisi03);
pisi21     = -(2#(-4#phirw#sqrtdoismuphi+4#phirw#
sqrtphi+3#murw#sqrtdoismuphi-7#murw#
sqrtphi))/(sqrtdoismuphi#(-2#murw+phirw)
#sqrtphi#(-sqrtdoismuphi+sqrtphi)##5);

```

```

stil1          = delta#um;
stil2          = delta#e;
stil           = stil1||stil2;
stiltpisi02    = t(stil)*diag(pisi02)*diag(filqua);
kdegama        = -stiltpisi02*stil;
invkdegama     = ginv(kdegama);
filcubo        = diag(fil#fiqua);
prodfilm1qua   = -.5/teta##6;
zdegama        = diag(stil*invkdegama*t(stil));
delta1         = -.5*diaginvpisi02*(zdegama*(diagpisi03)*filcubo
+zdebeta*diag(pisi21#prodfilm1qua));

s2til          = j(n,dimensaoq,0);
vecfishergama  = j(dimensaoq,1,0);
tracogama      = j(n,1,0);
K=0;
do i = 1 to q;
  do j = 1 to q;
    k = k+1;
    s2til[ ,k] = (mats[ ,i]#mats[ ,j])#delta;
    vecfishergama[k] = invkdegama[i,j];
  end;
end;

do i = 1 to n;
  tracogama[i] = s2til[i, ]*vecfishergama;
end;

delta2         = -.5*diag(tracogama);
deltatot       = (delta1+delta2)*um;
biasdegama     = invkdegama*stiltpisi02*deltatot;
gamacorrigido  = gama- biasdegama;
varibeta       = vecdiag(invkddebeta);
stddevbeta     = varibeta##.5;
varigama       = vecdiag(invkdegama);
stddevgama     = varigama##.5;
tbeta          = beta/stddevbeta;
tgama          = gama/stddevgama;
print beta biasdebeta betacorrigido stddevbeta tbeta;
print gama biasdegama gamacorrigido stddevgama tgama;

```

```
quit;
```

```
{\Large  
}
```