

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

**MELHORIA DOS ESTIMADORES DE MÁXIMA
VEROSSIMILHANÇA DOS SUBMODELOS NÃO-LINEARES DA
FAMÍLIA EXPONENCIAL DE DISPERSÃO**

TESE DE DOUTORADO

MARGARETH C. TOYAMA UDO

FLORIANÓPOLIS
SANTA CATARINA
2005

Margareth C. Toyama Udo

**MELHORIA DOS ESTIMADORES DE MÁXIMA
VEROSSIMILHANÇA DOS SUBMODELOS NÃO-
LINEARES DA FAMÍLIA EXPONENCIAL DE DISPERSÃO**

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 de qualidade e produtividade
Orientador: Prof. Paulo J. Ogliari, Dr.
Co-orientador: Prof. Gauss M. Cordeiro, PhD.

Florianópolis
Santa Catarina – Brasil
Dezembro/2005

Margareth C. Toyama Udo

MELHORIA DOS ESTIMADORES DE MÁXIMA VEROSSIMILHANÇA DOS SUBMODELOS NÃO- LINEARES DA FAMÍLIA EXPONENCIAL DE DISPERSÃO

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, 16 de dezembro de 2005.

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

Banca Examinadora

Prof. Paulo J. Ogliari, Dr.
Universidade Federal de Santa Catarina
Orientador

Prof. Gauss M. Cordeiro, PhD.
Universidade Federal Rural de Pernambuco
Co-orientador

Prof. Pedro A. Barbeta, Dr.
Universidade Federal de Santa Catarina
Examinador/Moderador

Profa. Lúcia P. Barroso, Dra.
Universidade de São Paulo
Examinadora Externa

Profa. Denise A. Botter, Dra.
Universidade de São Paulo
Examinadora Externa

Prof. Robert W. Samohyl, PhD.
Universidade Federal de Santa Catarina
Examinador

Prof. Adriano F. Borgatto, Dr.
Universidade Federal de Santa Catarina
Examinador

**Ao meu esposo Paulo e aos meus filhos
João Paulo, Pedro Luís e Cláudia Marina
que os amo igualmente.**

Agradecimentos

Ao Prof. Dr. Paulo José Ogliari, pelo tema da tese, apoio, compreensão e amizade.

Ao Prof. Dr. Gauss M. Cordeiro, exemplo de profissionalismo, competência e dedicação, pelo problema da tese e ter se deslocado de Recife a Maringá, tantas vezes, para sanar as nossas dúvidas.

À minha querida amiga Professora Rosangela G. Santana, por seus conselhos, incentivo e ajuda constantes durante estes anos de convívio. E ainda, na fase final da tese, me fez acreditar que era possível mesmo quando tudo se apresentava improvável.

À minha querida amiga Professora Isolde Previdelli que somou ao nosso grupo tornando-o melhor e mais alegre.

À Profa. Dra. Lúcia P. Barroso, pelas sugestões apresentadas e por nos ajudar tantas vezes via telefone e e-mails.

Ao Prof. Dr. Josmar Mazucheli, pelo auxílio na elaboração dos programas do SAS.

Ao Prof. Dr. Doherty Andrade, do Departamento de Matemática, pela iniciação e auxílio no uso do aplicativo *Maple*.

Ao Prof. Dr. Dalton Andrade e Prof. Dr. Pedro Barbetta, pelas sugestões apresentadas.

Ao Prof. PhD. Robert W. Samohyl, pelo apoio e amizade.

Ao Prof. PhD. James Lindsey, por ter cedido os dados, da droga *flosequinan*, da Knoll Pharmaceuticals.

À Profa. Dra. Denise A. Botter, pelo envio de material bibliográfico mesmo sem conhecêr-me.

À amiga Cláudia Barbiero, pela ajuda, incentivo e carinho.

Aos amigos do Departamento de Estatística, da UEM, pelo incentivo e apoio.

RESUMO

Em diversas áreas, particularmente em controle de qualidade, é de interesse modelar não só o parâmetro média, mas também o de dispersão. Um modelo desse tipo auxilia na melhoria contínua dos produtos e processos, sem aumentar os custos de produção. Um aprimoramento de qualidade de um processo industrial é verificado quando a dispersão é minimizada e o valor médio permanece sob controle. Nesta tese, são apresentados os modelos não-lineares da família exponencial de dispersão linear, que ficam caracterizados por duas componentes sistemáticas não-lineares, uma para a média e outra para a dispersão, além de uma componente aleatória, cuja função densidade de probabilidade pertence à família exponencial de dispersão linear. O método de estimação adotado foi o de máxima verossimilhança, que tem propriedades assintóticas de ordem n^{-1} , onde n é o tamanho da amostra; portanto, esses estimadores são não-viesados com erro aproximado de n^{-1} . Na prática, geralmente um viés dessa ordem é ignorado, contudo, para amostras pequenas, o viés pode ser apreciável e ter o mesmo valor que o seu correspondente erro padrão; neste caso, torna-se útil se ter uma estimativa de sua magnitude. Em geral, a redução do viés não necessariamente apresenta um melhor estimador, ela será benéfica somente quando a variância é reduzida. O objetivo desta tese é apresentar uma expressão para os vieses e, com isto, melhorar as inferências. Como aplicação, modelou-se e obteve-se o estimador corrigido da concentração de *flosequinan* no plasma, com submodelos *one-compartment* para a média e dispersão e componentes aleatórias: normal, gama e inversa gaussiana. É importante salientar que as expressões dos vieses são de fácil implementação. Pelo estudo de simulação, os estimadores corrigidos foram avaliados e comparados aos estimadores de máxima verossimilhança restrita e também os de máxima verossimilhança. Os resultados evidenciaram que os estimadores corrigidos são mais precisos do que os obtidos pela máxima verossimilhança restrita e, também, do que máxima verossimilhança, recomendando, assim, o uso de estimativas corrigidas em amostras de tamanho pequeno ou moderado.

Palavras-chave: método de máxima verossimilhança, família exponencial de dispersão, viés.

ABSTRACT

In several areas, mainly in quality control, it is from model interest not only the mean, but also the dispersion parameter, which allows the achievement of product and process continuous improvement without increasing the production costs, that is, the quality of an industrial process that occurs when the dispersion is minimized and the mean value remains under control. In this work, the non-linear models of the dispersion exponential family are presented, which are characterized by two systematic components, one for the mean and another for the dispersion, in addition to a random component whose probability density function belongs to the dispersion exponential family. The estimate method was the maximum likelihood that has asymptotic properties of n^{-1} order, where n is the sample size, thus, these estimators are not biased with approximate error of n^{-1} . In practice, the bias is usually ignored; nevertheless, for small samples it can be appreciable and have the same value as its correspondent standard error, in this case, it becomes useful to have an estimate of its magnitude. As a general rule, the reduction in the bias does not necessarily present a better estimator; it will be beneficial only when the variance is reduced. The objective of this work is to present an expression for the bias and, in this way, improve the inferences. As application, it was modeled and obtained the corrected estimator of *flosequinan* concentration in plasma, with *one-compartment* sub-models for the mean and dispersion and random components: normal, gamma and Gaussian inverse. It is important to highlight that the bias expressions are easily implemented. Through the simulation study, the corrected estimators were evaluated and compared to the restricted maximum likelihood estimators. The results show that the corrected estimators are more accurate than the ones obtained through the restricted maximum likelihood, besides the necessity of obtaining corrected estimates in samples from small to moderate sizes.

Key-words: maximum likelihood, dispersion exponential family, bias.

Sumário

1	Introdução	8
1.1	Importância	9
1.2	Objetivos	10
1.2.1	Geral	10
1.2.2	Específicos	10
1.3	Estrutura	11
2	Modelos lineares generalizados duplos	12
2.1	Modelo	12
2.2	EMV e estimadores de $O(n^{-2})$	14
2.2.1	Função log-verossimilhança, função escore e matriz de informação de Fisher de β e ϕ	14
2.2.2	EMV de β e γ	16
2.2.3	Estimador de $O(n^{-2})$	18
2.2.4	Aplicação: modelando a quantidade de gases hidrocarbonos liberados	20
2.3	Procedimento de estimação dos MLGD por Smyth e Verbyla (1999)	22
2.3.1	Estudo de simulação: comparando EMVR e estimadores de $\mathbf{O}(n^{-2})$	24
2.3.2	Resultados	26
3	Modelos não-lineares da família exponencial	28
3.1	O modelo	28
3.2	Função log-verossimilhança, função escore e matriz de informação de Fisher de β e ϕ	29
3.3	EMV de β	30
3.4	Estimadores de $O(n^{-2})$	31

3.4.1	Vieses do estimador de β, η e ϕ	31
3.5	Aplicação: modelando a quantidade de vendas de produtos ao consumidor . . .	32
4	Estimadores dos modelos não-lineares da família exponencial de dispersão linear	36
4.1	Modelo	36
4.2	Funções escore e informação de Fisher de β e ϕ	37
4.3	EMV de β e γ	39
4.4	Vieses de $\hat{\beta}$ e $\hat{\gamma}$	40
4.5	Caso especial	44
4.6	Viés de $\hat{\eta}, \hat{\tau}, \hat{\mu}, \hat{\phi}$ e $\hat{\phi}^{-1}$	44
4.7	Aplicação: modelando a quantidade de vendas de produtos ao consumidor . . .	46
4.7.1	Modelo	47
4.7.2	Resultados	48
5	Avaliando estimadores de $O(n^{-2})$ através de um estudo de simulação em MNLFED	49
5.1	Simulação da variável resposta com distribuição inversa gaussiana	49
5.1.1	Submodelos média e dispersão	50
5.1.2	Resultados	52
5.2	Simulação da variável resposta com distribuição gama	55
5.2.1	Submodelos média e dispersão	55
5.2.2	Resultados	56
6	Aplicação em modelos <i>one-compartment</i>	58
6.1	Modelo	59
6.1.1	Variáveis de interesse	59
6.1.2	Delineamento experimental	59
6.1.3	Componente aleatória	59
6.1.4	Componentes sistemáticas:	60
6.2	Algoritmo do SAS	61
6.3	Resultados	62

7	Considerações finais	65
	Referências bibliográficas	65
A	Identities de Bartlett	70
B	Propriedade dos estimadores de máxima verossimilhança	73
C	Viés dos EMV-Cox e Snell (1968)	75
D	Cumulantes dos MLGD	78
D.1	Cálculo de U_r e κ_r	78
D.2	Cálculo de U_{rs} e κ_{rs}	79
D.3	Cálculo de U_{rst} e κ_{rst}	80
D.4	Cálculo do $\kappa_{rs}^{(T)}$	81
D.5	Cálculo de U_{stR} e κ_{stR}	81
D.6	Cálculo do $\kappa_{rs}^{(t)}$	81
D.7	Cálculo da U_R	82
D.8	Cálculo de U_{RS} e κ_{RS}	82
D.8.1	Cálculo do Valor Esperado da $\mathbf{c}(y)$	82
D.9	Cálculo do $\kappa_{RS}^{(t)}$	83
D.10	Cálculo do $\kappa_{RS}^{(T)}$	83
D.11	Cálculo de U_{rS} e κ_{rS}	84
D.12	Cálculo do $\kappa_{rS}^{(T)}$	84
D.13	Cálculo do κ_{rST}	84
D.14	Cálculo do κ_{RST}	85
E	Cumulantes dos MNLFED	86
E.1	Cálculo de U_r e κ_r	86
E.2	Cálculo de U_{rs} e κ_{rs}	87
E.3	Cálculo do $\kappa_{rs}^{(t)}$	87
E.4	Cálculo de U_{rst} e κ_{rst}	88
E.5	Cálculo do $\kappa_{rs,t}$	89

E.6	Cálculo de U_R	90
E.7	Cálculo de U_{RS} e κ_{RS}	90
E.8	Cálculo de $U_{RS}^{(T)}$ e do cumulante $\kappa_{RS}^{(T)}$	91
E.9	Cálculo de U_{RST} e do cumulante κ_{RST}	91
E.10	Cálculo de U_{Rst} e do cumulante κ_{Rst}	92
F	Programa de simulação para inversa gaussiana	94

Lista de Figuras

2.1	Fluxograma do MLGD com distribuição inversa gaussiana.	25
2.2	Diagrama de caixas da distribuição das estimativas de γ_1	27
2.3	Diagrama de caixas da distribuição das estimativas de γ_2	27
5.1	Diagrama de caixas da estimativa do parâmetro β_1	54
5.2	(a): Diagrama de caixas da estimativa do parâmetro γ_1 ; (b): Diagrama de caixas da estimativa do parâmetro γ_2	54
6.1	Modelos <i>one-compartment</i>	58
6.2	(a): Comportamento da log-verossimilhança quando k_{am} é fixado para dose 50; (b): Comportamento da log-verossimilhança quando k_{am} é fixado para dose 100; (c): Comportamento da log-verossimilhança quando k_{am} é fixado para dose 150.	61
6.3	Diagrama de dispersão da concentração versus tempo.	63

Lista de Tabelas

2.1	Estimativas de MV, de $O(n^{-2})$ e desvios padrão do submodelo média.	22
2.2	Estimativas de MV, de $O(n^{-2})$ e desvios padrão do submodelo dispersão.	22
2.3	Componentes do submodelo dispersão para estimar γ	23
2.4	Tabela das médias e desvios padrão.	26
3.1	Dados das quantidades de vendas reais e estimadas.	33
3.2	Estimativa de MV, de $O(n^{-2})$, desvios padrão, quociente e AIC.	35
4.1	Estimativas de MV, de $O(n^{-2})$ e erros padrão.	48
5.1	Estimativas de MV, de $O(n^{-2})$ e desvios padrão.	53
5.2	Quociente do viés do parâmetro pelo seu respectivo desvio padrão.	53
5.3	Estimativas de MV, de $O(n^{-2})$ e desvios padrão.	56
5.4	Quociente do viés do parâmetro pelo seu respectivo desvio padrão.	57
6.1	Estimativa de MV, de $O(n^{-2})$, erros padrão para modelo gama.	63
6.2	Estimativas de MV, de $O(n^{-2})$, erros padrão para modelo inversa gaussiana. . .	63
6.3	Estimativas de MV, de $O(n^{-2})$, erros padrão para modelo normal.	64
6.4	Porcentagem do quociente entre o viés e seu respectivo erro padrão para resposta inversa gaussiana.	64

Lista de siglas

MLG	Modelo Linear Generalizado
MNLFE	Modelo Não-Linear da Família Exponencial
MLGD	Modelo Linear Generalizado Duplo
MNLFED	Modelo Não-Linear da Família Exponencial de Dispersão Linear
EMV	Estimador de Máxima Verossimilhança
MMV	Método de Máxima Verossimilhança
MV	Máxima Verossimilhança
EMVR	Estimador de Máxima Verossimilhança Restrita
$O(n^{-1})$	Primeira Ordem
$O(n^{-2})$	Segunda Ordem

Lista de símbolos

β	Parâmetro beta
$\hat{\beta}$	Estimador de beta
$\tilde{\beta}$	Estimador de beta de $O(n^{-2})$
γ	Parâmetro gama
$\hat{\gamma}$	Estimador de gama
$\tilde{\gamma}$	Estimador de gama de $O(n^{-2})$
$\check{\gamma}^*$	Estimador de máxima verossimilhança restrita
μ	Parâmetro média
ϕ	Parâmetro dispersão
η	Preditor eta
τ	Preditor tau
k_s	Cumulantes de $O(n^{-1})$.
$B(\hat{\beta})$	Viés de $\hat{\beta}$.
$B(\hat{\gamma})$	Viés de $\hat{\gamma}$.
U_β e U_γ	Função escore de β e γ .
$K_{\beta\beta}$ e $K_{\gamma\gamma}$	Matriz informação de Fisher de β e γ
K^{-1}	Inversa da matriz informação de Fisher

Capítulo 1

Introdução

A modelagem estatística tem um importante papel em aplicações na Engenharia de Produção, como em uma indústria, quando se deseja identificar os fatores que afetam o nível do processo de produção ou manufatura, ou ainda quando se deseja otimizar a produção diária. Assim, a modelagem fornece um conjunto de informações por meio das quais é possível prever futuros resultados. Além disso, servem de base para o entendimento das próprias observações. A modelagem iniciou-se com os modelos clássicos de regressão, das quais uma das restrições é a suposição de que a variável resposta apresenta uma distribuição normal. Posteriormente, com o desenvolvimento teórico nesta área e os avanços computacionais, foi possível a construção de modelos cada vez mais sofisticados. Uma das contribuições se deve a Nelder e Wedderburn (1972) que introduziram os Modelos Lineares Generalizados (MLG), nos quais supõe-se que a variável resposta apresenta uma distribuição de probabilidade membro da família exponencial, um conjunto de covariáveis que descrevem uma estrutura linear do modelo e uma função de ligação que relaciona a média da variável resposta à estrutura linear.

Embora os MLG sejam flexíveis quanto à distribuição de probabilidade da variável resposta, estes não permitem que a componente sistemática seja não-linear; contudo, se esta suposição é válida, têm-se os Modelos Não-Lineares da Família Exponencial (MNLFE). A diferença entre eles é que nos MLG o preditor linear é usado para funções adequadas dos valores esperados das observações e nos MNLFE os valores esperados são funções gerais dos parâmetros de regressão. Os MNLFE têm sido estudados, por exemplo, por Cordeiro e Paula (1989), os quais apresentam a melhoria das estatísticas da razão de verossimilhança, Cook e Tsai (1990), que discutem a aproximação cúbica das regiões de confiança, Pazman (1991), que apresenta a aproximação do ponto de sela da distribuição do estimador de máxima verossimilhança do ponto de vista geométrico e Wei (1998), que introduz uma base para o estudo do comportamento estatístico desta classe de modelos.

Tradicionalmente, tanto nos MLG como nos MNLFE modelam-se apenas a média, enquanto a dispersão é supostamente conhecida ou desconhecida mas, fixada. Pregibon (1984), em sua resenha sobre a primeira edição do livro de McCullagh e Nelder (1983), foi o primeiro a

apresentar uma proposta para modelar o efeito da dispersão. A partir desta, vários trabalhos têm surgido para se modelar a dispersão. Entre eles, os MLG com covariáveis de dispersão ou MLG duplos (MLGD), que modelam os submodelos média e dispersão pela função densidade de probabilidade, além de uma outra técnica baseada nos conceitos de Taguchi, que utiliza a metodologia de superfície de resposta para modelar a média e variância de características de qualidade para se obter a otimização experimental de produtos e processos, como apresentada por Barbetta (1998) e Barbetta *et al.* (1999). Em ambos, a meta é modelar a variabilidade na busca da qualidade de um processo industrial, e, se a partir destes modelos a dispersão for minimizada, enquanto o valor médio ou esperado permanece sob controle, é alcançada a melhoria contínua dos produtos e processos sem aumentar os custos de produção.

Uma aplicação dos Modelos Não-Lineares da Família Exponencial de Dispersão Linear (MNLFED) é utilizada em indústrias farmacêuticas, em particular em farmacocinética, que trata do estudo do desenvolvimento de uma droga (medicamento), no processo da absorção, da distribuição do metabolismo e eliminação de algumas substâncias em um organismo. Nesta Tese, são utilizados voluntários sadios, o que significa um número reduzido dos mesmos, pois quanto menor a amostra, menor é o custo para as indústrias farmacêuticas. Uma aplicação deste estudo é apresentado no Capítulo 6.

1.1 Importância

Os estimadores dos parâmetros de todos os modelos citados geralmente são obtidos através do Método de Máxima Verossimilhança (MMV), pois, segundo Cordeiro (1986, p. 25), este apresenta propriedades de consistência e eficiência assintótica, mas são geralmente viesados para pequenas amostras. Portanto, é importante reduzir este viés, para melhorar os estimadores, e, para isso, geralmente, é utilizada a correção proposta por Cox e Snell (1968) originada de Bartlett (1953).

O estudo do viés de Estimadores de Máxima Verossimilhança (EMV) em modelos não-lineares tem apresentado grandes avanços e vários autores têm estudado sobre o assunto. Destacando-se: Box (1971) apresentou uma expressão geral do viés de primeira ordem em modelos não-lineares multivariados com matrizes de covariâncias conhecidas; Sowden (1971) estudou o viés em modelos com resposta quantal; Pike *et al.* (1980) investigaram o viés do modelo linear logístico; Cook, Tsai e Wei (1986) apresentaram para os modelos não-lineares o viés para a posição da covariável no espaço das amostras; Young e Bakir (1987) mostraram que a correção do viés pode melhorar a estimação em modelos de regressão log-gama generalizado; Cordeiro e McCullagh (1991) demonstraram a fórmula geral do viés em MLG; Paula (1991) deduziu o viés de primeira ordem para os modelos não-lineares da família exponencial; Cordeiro e Klein (1994) apresentaram a fórmula geral do viés em modelos ARMA; Cordeiro e Vasconcellos (1997) obtiveram a fórmula matricial geral para a correção do viés em modelos de regressão não-linear multivariado com erros normais; Botter e Cordeiro (1998) deduziram fórmulas gerais

para o viés de primeira ordem dos estimadores de máxima verossimilhança para os parâmetros dos MLGD; Cordeiro e Vasconcellos (1999) apresentaram o viés de segunda ordem para a estimativa de máxima verossimilhança nos modelos de regressão von-Mises; Cordeiro e Botter (2001) demonstraram a utilidade do viés com o objetivo de melhorar as propriedades estatísticas de suas estimativas e apresentam as fórmulas do viés de segunda ordem para estimativas de máxima verossimilhança em MLG superdispersados; finalmente, Cordeiro e Barroso (2003) apresentaram a correção de terceira ordem das estimativas dos parâmetros, isto é, o viés de segunda ordem para os estimadores dos parâmetros dos MLG.

Segundo Botter e Cordeiro (1998), o uso do viés de primeira ordem é ignorado na prática, pois, conforme alguns pesquisadores, este valor é muito pequeno quando comparado ao erro padrão da média. Esta afirmativa pode ser considerada verdadeira para grandes amostras, mas, se a amostra é pequena, o viés pode apresentar o mesmo tamanho do seu correspondente erro padrão, sendo, neste caso, importante se ter uma idéia de sua magnitude. Portanto, cabe salientar que o uso do viés é útil em estudos farmacocinéticos, nos quais a amostra é constituída de voluntários sadios, e, também, em controle de qualidade, como pode ser visto em Montgomery (2000, p. 146) e também Mattos (2004), que apresenta um estudo de experimentos com poucas repetições. Assim, nestes casos, o objetivo é ter amostras de tamanho reduzido, visto que isto implica em menor custo e, conseqüentemente, maior lucro para as empresas. Quando o número de elementos na amostra é reduzido, Botter e Cordeiro (1998) enfatizam que o cálculo do viés de primeira ordem, $O(n^{-1})$, é a mais importante de todas as aproximações na teoria de estimação pelo método de máxima verossimilhança em modelos de regressão em geral.

Portanto, como todo produto apresenta uma variabilidade, uma vez que não é possível obter dois produtos exatamente iguais, a idéia é controlar a mesma, tentando manter a distância próxima de zero com relação ao valor-alvo. Então, quanto mais precisas as estimativas dos parâmetros dos submodelos média e dispersão, melhor é o resultado obtido, contribuindo para a melhoria dos estimadores de máxima verossimilhança para os MNLFED.

1.2 Objetivos

1.2.1 Geral

Obter o viés dos estimadores dos parâmetros de máxima verossimilhança em MNLFED.

1.2.2 Específicos

- 1) Divulgar a facilidade da obtenção dos estimadores de segunda ordem, $O(n^{-2})$, em alguns modelos de regressão.
- 2) Estender, para os MNLFED, as propostas contidas no estudo de Botter e Cordeiro (1998).

- 3) Analisar e avaliar, via simulação, submodelos lineares na média e na dispersão os estimadores de $O(n^{-2})$, e a metodologia proposta por Smyth e Verbylla (1999).
- 4) Obter expressões analíticas dos vieses de $O(n^{-1})$ dos estimadores de máxima verossimilhança (EMV) dos submodelos não-lineares: média e dispersão na classe de modelos da família exponencial de dispersão.
- 5) Analisar e avaliar o viés obtido através de simulação em submodelos não-lineares na média e dispersão para os EMV e estimadores de $O(n^{-2})$.

1.3 Estrutura

Este trabalho apresenta a seguinte estrutura:

Capítulo 1: Introdução, que apresenta a importância, o problema a ser resolvido, os objetivos geral e específicos e uma estruturação geral.

Capítulo 2: Uma revisão da literatura sobre os MLGD é apresentada com uma aplicação, e ainda um estudo de simulação que compara estimadores de $O(n^{-2})$ e a metodologia proposta por Smyth e Verbylla (1999), para contemplar os objetivos específicos 1 e 3. Não é apresentado o MLG, pois esse é análogo ao submodelo média do MLGD.

Capítulo 3: Continuando a revisão, são apresentados os MNLFE para contemplar o objetivo específico 1.

Capítulo 4: A contribuição teórica na melhoria dos EMV para os MNLFE, a qual foi desenvolvida por meio do conhecimento disponível em teorias publicadas em livros e artigos, identificando as teorias produzidas, a fim de contribuir, apresentando expressões para o viés de 1ª ordem, satisfazendo os objetivos específicos 2 e 4.

Capítulo 5: Análise e avaliação da simulação, considerando variáveis respostas com distribuição: inversa gaussiana e gama, realizada baseada nos resultados do Capítulo 4, satisfazendo o objetivo específico 5.

Capítulo 6. Aplicação dos resultados obtidos em um experimento real, na qual alguma característica da droga *flosequinan* é modelada através de modelos *one-compartment*, muito utilizadas em farmacocinética. Estes dados foram cedidos por Lindsey, os quais foram usados no artigo Lindsey *et al.* (2000).

Capítulo 2

Modelos lineares generalizados duplos

Os Modelos Lineares Generalizados Duplos (MLGD) são comumente utilizados quando a variância da variável resposta excede a variância explicada pelo modelo ajustado, que considera o parâmetro de precisão ϕ constante. A suposição de que esse parâmetro seja conhecido ou desconhecido, mas igual para todas as observações, será eliminada, permitindo que a média e a dispersão sejam modeladas.

Esses modelos podem ser aplicados em controle de qualidade de uma indústria, onde é de interesse minimizar a dispersão, enquanto o valor médio ou esperado permanece sob controle.

Para estimar os parâmetros dos MLGD, dois procedimentos são apresentados:

- A) pelo MMV para os dois submodelos. Como os EMV são geralmente viesados, é proposto um método no qual, tanto para o submodelo média como para o submodelo dispersão, estes sejam corrigidos por meio do cálculo da expressão do viés, apresentado no Apêndice C, devido a Cox e Snell (1968);
- B) pelo método creditado a Smyth e Verbyla (1999), que estimam o submodelo média no contexto dos MLG e, para o submodelo dispersão, a função densidade de probabilidade é aproximada pelo ponto de sela.

2.1 Modelo

Os MLGD ficam caracterizados por uma componente aleatória, pelas componentes sistemáticas para média e dispersão e pelas funções de ligação, ou seja:

- a) componente aleatória: $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)^T$, supostas variáveis aleatórias independentemente distribuídas, das quais os Y_l 's são provenientes da família exponencial de dispersão, cuja função densidade de probabilidade é dada por:

$$p(y; \theta_l, \phi_l) = \exp [\phi_l \{y\theta_l - b(\theta_l) + c(y)\} + d_1(y) + d_2(\phi_l)], \quad (2.1)$$

sendo $b(\cdot)$, $c(\cdot)$, $d_1(\cdot)$ e $d_2(\cdot)$ funções conhecidas. Têm-se ainda que θ é o parâmetro canônico e ϕ o parâmetro de precisão, em que $E(Y_l) = \frac{\partial b(\theta_l)}{\partial \theta_l} = \mu_l$, $V(Y_l) = \phi_l^{-1}V_l$ com $V_l = V(\mu_l) = \frac{d\mu_l}{d\theta_l}$ uma função de variância.

b1) componente sistemática do submodelo média: as covariáveis supostamente fixadas $x_l = (x_{l1}, \dots, x_{lp})^T$ formam uma estrutura linear, isto é,

$$\eta_l = g(\mu_l) = \sum_{j=1}^p x_{lj}\beta_j, \quad l = 1, \dots, n \quad \text{ou} \quad \boldsymbol{\eta} = X\boldsymbol{\beta}, \quad (2.2)$$

no qual

- $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)^T$ é o preditor linear $n \times 1$;
- $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ é o vetor de parâmetros desconhecidos $\boldsymbol{\beta} \subset R^p, p < n$ e
- $X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{12} & x_{22} & \cdots & x_{2p} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}$ é a matriz modelo das covariáveis $n \times p$.

b2) componente sistemática do submodelo dispersão: as covariáveis supostamente fixadas $s_l = (s_{l1}, \dots, s_{lq})^T$ formam uma estrutura linear, isto é,

$$\tau_l = h(\phi_l) = \sum_{k=1}^q s_{lk}\gamma_k, \quad l = 1, \dots, n \quad \text{ou} \quad \boldsymbol{\tau} = S\boldsymbol{\gamma}, \quad (2.3)$$

no qual

- $\boldsymbol{\tau} = (\tau_1, \dots, \tau_n)^T$ é o preditor linear de dispersão $n \times 1$;
- $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_q)^T$ é o vetor de parâmetros desconhecidos $\boldsymbol{\gamma} \subset R^q, q < n$ e
- $S = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1q} \\ s_{12} & s_{22} & \cdots & s_{2q} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ s_{n1} & s_{n2} & \cdots & s_{nq} \end{pmatrix}$ é a matriz modelo das covariáveis $n \times q$.

c) funções de ligação da média (g) e dispersão (h) relacionam a componente aleatória às componentes sistemáticas dos submodelos média e dispersão, respectivamente, ou seja, associam os submodelos média e dispersão aos seus respectivos preditores lineares, isto é:

$$\eta_l = g(\mu_l) \quad \text{e} \quad \tau_l = h(\phi_l),$$

sendo $g(\cdot)$ e $h(\cdot)$ funções monótonas e diferenciáveis.

Desse modo, tem-se que o modelo apresenta $p + q$ parâmetros a serem estimados, uma vez que a matriz X é de posto p e a matriz S de posto q .

A função de ligação é dita canônica se $\boldsymbol{\theta} = \boldsymbol{\eta}$ e não há restrição quanto ao uso. No entanto, é conveniente a sua utilização, pois, segundo Cordeiro (1986, p. 20) e Paula (2001, p. 10), existe a vantagem de que a mesma garante a concavidade da função log-verossimilhança e, conseqüentemente, muitos resultados assintóticos são obtidos mais facilmente.

Uma observação importante é que a matriz modelo S de covariáveis não é necessariamente um subconjunto das covariáveis da matriz X .

É essencial salientar que as duas componentes sistemáticas são parametrizadas como $\mu = \mu(\beta)$ e $\phi = \phi(\gamma)$, onde a média e a dispersão podem apresentar estruturas lineares diferentes.

2.2 EMV e estimadores de $O(n^{-2})$

Nessa seção é apresentada a metodologia proposta por Botter e Cordeiro (1998) que subsidiará a obtenção dos estimadores de $O(n^{-2})$ para os MNLFED.

2.2.1 Função log-verossimilhança, função escore e matriz de informação de Fisher de β e ϕ

O logaritmo da função de verossimilhança de (2.1) para os parâmetros μ ou θ e ϕ , dado y_l é:

$$\begin{aligned} \ell(\mu, \phi) &= \sum_{l=1}^n [\phi_l \{y_l \theta_l - b(\theta_l) + c(y_l)\} + d_1(y_l) + d_2(\phi_l)] \quad \text{ou} \\ \ell(\boldsymbol{\mu}, \boldsymbol{\phi}) &= \boldsymbol{\phi} \{\mathbf{y}^T \boldsymbol{\theta} - b(\boldsymbol{\theta}) + c(\mathbf{y})\} + d_1(\mathbf{y}) + d_2(\boldsymbol{\phi}). \end{aligned}$$

A função escore e a matriz de informação de Fisher de β e γ são apresentadas, pois são necessárias para a obtenção dos seus estimadores.

Função escore de β

A função escore de β , que será denotada por $U_{\beta} = \frac{\partial \ell}{\partial \beta}$, é obtida a partir da equação apresentada a seguir, a qual está demonstrada no Apêndice D:

$$U_{\beta_r} = \sum_{l=1}^n \phi_l (y_l - \mu_l) (V(\mu_l))^{-1} \frac{\partial \mu_l}{\partial \eta_l} x_{lr}, \quad r = 1, \dots, p$$

que, após realizar os somatórios, a sua forma matricial resulta em:

$$U_{\beta} = \left(V^{-1} \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\eta}} X^T \right)^T \Phi (\mathbf{y} - \boldsymbol{\mu}) \quad (2.4)$$

sendo $V = \text{diag} \{V_1, V_2, \dots, V_n\}$ e $\Phi = \text{diag} \{\phi_1, \phi_2, \dots, \phi_n\}$.

Substituindo a matriz de pesos $W = \text{diag} \{w_1, w_2, \dots, w_n\}$, $w_l = \frac{1}{V_l} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2$ e manipulando algebricamente, a expressão (2.4) torna-se:

$$U_{\beta} = X^T \Phi W^{1/2} V^{-1/2} (\mathbf{y} - \boldsymbol{\mu}). \quad (2.5)$$

Função escore de γ

A função escore de γ , que será denotada por U_{γ} , é obtida de modo análogo a U_{β} e é dada por:

$$U_{\gamma_R} = \sum_{l=1}^n \left\{ y_l \theta_l - b(\theta_l) + c(y_l) + \frac{\partial d_2(\phi_l)}{\partial \phi_l} \right\} \frac{\partial \phi_l}{\partial \tau_l} s_{lR}, \quad R = 1, \dots, q,$$

conforme apresentada no Apêndice D. Considerando $\nu_l = y_l \theta_l - b(\theta_l) + c(y_l) + \frac{dd_2(\phi_l)}{d\phi_l}$, U_{γ_R} resulta em:

$$U_{\gamma_R} = \sum_{l=1}^n \nu_l \frac{\partial \phi_l}{\partial \tau_l} s_{lR}. \quad (2.6)$$

A expressão (2.6), após aplicar a soma sob todos os elementos, resulta na seguinte forma matricial:

$$U_{\gamma} = S^T \Phi_1 \boldsymbol{\nu}, \quad (2.7)$$

sendo $\Phi_1 = \text{diag} \{\phi_{11}, \phi_{12}, \dots, \phi_{1n}\}$, com $\phi_{1l} = \frac{\partial \phi_l}{\partial \tau_l}$.

Funções escore de $(\boldsymbol{\beta}, \boldsymbol{\gamma})$

A partir das funções escore de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ obtém-se este vetor que é dado por:

$$U(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \begin{pmatrix} \frac{\partial \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\beta}} \\ \frac{\partial \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\phi}} \end{pmatrix} = \begin{pmatrix} X^T \Phi W^{1/2} V^{-1/2} (\mathbf{y} - \boldsymbol{\mu}) \\ S^T \Phi_1 \boldsymbol{\nu} \end{pmatrix}.$$

Matriz de informação de Fisher para $\boldsymbol{\beta}$, $K_{\beta\beta}$

O elemento (r, s) dessa matriz é definido como sendo

$$K_{\beta_r \beta_s} = E \left(-\frac{\partial^2 \ell}{\partial \beta_r \partial \beta_s} \right) = \sum_{l=1}^n \phi_l \frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 x_{lr} x_{ls}, \quad r = 1, \dots, p,$$

(mais detalhes da obtenção de K_{β_r, β_s} é apresentado no Apêndice D) que, após cálculos algébricos, resulta na seguinte forma matricial:

$$K_{\beta\beta} = X^T \Phi W X. \quad (2.8)$$

Matriz de informação de Fisher para γ , $K_{\gamma\gamma}$

O elemento (R, S) dessa matriz é dado por:

$$K_{\gamma_R \gamma_S} = E \left(-\frac{\partial^2 \ell}{\partial \gamma_R \partial \gamma_S} \right) = - \sum_{l=1}^n \frac{\partial^2 d_2(\phi_l)}{\partial \phi_l^2} \left(\frac{\partial \phi_l}{\partial \tau_l} \right)^2 s_{lR} s_{lS}, \quad R = 1, \dots, q,$$

(veja detalhes da demonstração no Apêndice D). Matricialmente, torna-se:

$$K_{\gamma\gamma} = S^T (-D_2 \Phi_1^2) S,$$

em que $D_2 = \text{diag} \{d_{21}, d_{22}, \dots, d_{2n}\}$ sendo $d_{2l} = \frac{\partial^2 d_2(\phi)}{\partial \phi_l^2}$.

As matrizes $K_{\beta\gamma} = K_{\gamma\beta} = 0$ (como demonstrada no Apêndice D). Portanto, a matriz de informação de Fisher para os parâmetros β e γ é dada por:

$$K(\beta, \gamma) = \begin{pmatrix} K_{\beta\beta} & K_{\beta\gamma} \\ K_{\gamma\beta} & K_{\gamma\gamma} \end{pmatrix} = \begin{pmatrix} X^T \Phi W X & \mathbf{0} \\ \mathbf{0} & S^T (-D_2 \Phi_1^2) S \end{pmatrix}. \quad (2.9)$$

Logo, os parâmetros β e γ são globalmente ortogonais implicando que os EMV $\hat{\beta}$ e $\hat{\gamma}$ são assintoticamente independentes.

2.2.2 EMV de β e γ

Justifica-se estimar separadamente os parâmetros β e γ , uma vez que os estimadores são assintoticamente independentes.

Para obtenção dos EMV, conforme Apêndice B, deve-se resolver o sistema de equações $U(\hat{\beta}, \hat{\gamma}) = 0$ que é, geralmente, não-linear, sendo necessário um processo iterativo para sua solução.

EMV de β

Para obter as estimativas de β através do MMV podem ser utilizados os algoritmos:

- de Newton-Raphson que está disponível no aplicativo SAS nas *procedures* NLP e NLMIXED. Para utilizar este método, é necessário expandir $\frac{\partial \ell(\beta)}{\partial \beta}$ em série de Taylor até primeira

ordem para uma dada vizinhança de β^i . Após alguns cálculos, obtém-se que para uma dada iteração $i + 1$ o valor de $\hat{\beta}^{i+1}$ é dado por:

$$\hat{\beta}^{i+1} = \hat{\beta}^i + \left\{ -\frac{\partial^2 \ell(\beta^i)}{\partial \beta \partial \beta^T} \right\}^{-1} \frac{\partial \ell(\beta^i)}{\partial \beta}. \quad (2.10)$$

- escore de Fisher que está disponibilizado na *procedure* GENMOD, do aplicativo SAS. Este método consiste em substituir $-\frac{\partial^2 \ell(\beta)}{\partial \beta^2}$ por $E\left(-\frac{\partial^2 \ell(\beta)}{\partial \beta^2}\right)$ na expressão (2.10). Após alguns cálculos, resulta no seguinte processo iterativo:

$$\hat{\beta}^{i+1} = \left[(X^T \Phi W X)^{-1} X^T \Phi W \left\{ \eta + \frac{\partial \eta}{\partial \mu} (\mathbf{y} - \mu) \right\} \right]^i,$$

que pode ser reescrito como:

$$\hat{\beta}^{i+1} = \left\{ (X^T \Phi W X)^{-1} X^T \Phi W z_\beta \right\}^i, \quad (2.11)$$

sendo o vetor de respostas ajustadas z_β , dado por:

$$z_\beta = \left(\eta + \frac{\partial \eta}{\partial \mu} (\mathbf{y} - \mu) \right)_{n \times 1}$$

no qual $\frac{\partial \eta}{\partial \mu} = \text{diag} \left\{ \frac{\partial \eta_1}{\partial \mu_1}, \frac{\partial \eta_2}{\partial \mu_2}, \dots, \frac{\partial \eta_n}{\partial \mu_n} \right\}$.

O sistema de equações (2.11) pode ser visto como uma regressão de mínimos quadrados ponderados, com variável resposta ajustada z_β e matriz peso ΦW .

EMV de γ

O estimador de γ é obtido de forma análoga a β , resultando no processo iterativo

$$\hat{\gamma}^{i+1} = \left\{ (S^T (-D_2 \Phi_1^2) S)^{-1} S^T (-D_2 \Phi_1^2) z_\gamma \right\}^i, \quad (2.12)$$

sendo o vetor de respostas ajustadas z_γ , dado por:

$$z_\gamma = \left(\tau - D_2^{-1} \Phi_1^{-1} \nu \right)_{n \times 1},$$

em que $\nu = \mathbf{y}^T \theta + b(\theta) + c(\mathbf{y}) + \frac{\partial d_2(\phi)}{\partial \phi}$.

O sistema de equações (2.12) pode ser visto como uma regressão de mínimos quadrados ponderados, com variável resposta ajustada z_γ e matriz peso $-D_2 \Phi_1^2$.

2.2.3 Estimador de $O(n^{-2})$

Para a obtenção dos vieses dos EMV do MLGD, são necessários alguns cumulantes que estão demonstrados no Apêndice D.

Os estimadores de $O(n^{-2})$, conforme Apêndice C, nada mais são do que os estimadores obtidos, na subseção anterior, subtraídos dos respectivos vieses.

Viés de $\hat{\beta}$

O viés de $\hat{\beta}_a$, denotado por $B(\hat{\beta}_a)$, para $a = 1, 2, \dots, p$, é obtido a partir da fórmula de Cox e Snell (1968), que para MLGD é a soma de todas as combinações dos $p + q$ parâmetros β_1, \dots, β_p e $\gamma_1, \dots, \gamma_q$ que resulta em:

$$B(\hat{\beta}_a) = \sum_{r,s,t} \kappa^{ar} \kappa^{st} \left(\kappa_{rs}^{(t)} - \frac{1}{2} \kappa_{rst} \right) + \sum_r \sum_{S,T} \kappa^{ar} \kappa^{ST} \left(\kappa_{rS}^{(T)} - \frac{1}{2} \kappa_{rST} \right). \quad (2.13)$$

Os cumulantes são obtidos a partir das derivadas da log-verossimilhança os subscritos minúsculos r, s, t, u, \dots denotarão as componentes do vetor β e subscritos maiúsculos R, S, T, U, \dots para o vetor γ .

No Apêndice D, demonstram-se que:

- $k_{rS} = 0; \kappa_{rS}^{(T)} = 0, \kappa_{rST} = 0,$
- $\kappa_{rst} = - \sum_{l=1}^n \phi_l (f + 2g) x_{lr} x_{ls} x_{lt}$ e
- $\kappa_{rs}^{(t)} = - \sum_{l=1}^n \phi_l (f + g)_l x_{lr} x_{ls} x_{lt}.$

Substituindo esses resultados na expressão (2.13), tem-se:

$$B(\hat{\beta}_a) = \sum_{r,s,t} \kappa^{ar} \kappa^{st} \left[- \sum_{l=1}^n \phi_l (f_l + g_l) x_{lr} x_{ls} x_{lt} + \frac{1}{2} \sum_{l=1}^n \phi_l (f_l + 2g_l) x_{lr} x_{ls} x_{lt} \right]. \quad (2.14)$$

Para obter $B(\hat{\beta})$ na forma matricial, considere:

- $F = \text{diag} \{f_1, \dots, f_n\}, f_l = \frac{1}{V_l} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial^2 \mu_l}{\partial \eta_l^2},$
- $Z_\beta = X K_{\beta\beta}^{-1} X^T = \{z_{\beta in}\}; Z_{\beta d} = \text{diag} \{z_{\beta 11}, \dots, z_{\beta nn}\}$ e
- $\mathbf{1} = (1, \dots, 1)^T.$

Rearranjando os termos da expressão (2.14), tem-se:

$$B(\hat{\beta}) = (X^T \Phi W X)^{-1} X^T \Phi W \xi, \quad (2.15)$$

sendo

$$\xi = -\frac{1}{2} W^{-1} Z_{\beta d} F \mathbf{1}.$$

Observê-se que a expressão (2.15) é uma regressão de mínimos quadrados ponderados, cuja variável resposta é ξ e a matriz peso é ΦW .

Viés de $\hat{\gamma}$

Analogamente ao caso do $B(\hat{\beta})$, é obtido o viés de $\hat{\gamma}$, denotado por $B(\hat{\gamma}_A)$, para $A = 1, 2, \dots, q$. A expressão de Cox e Snell, neste caso, considerando o somatório da combinação dos $p + q$ parâmetros, é dada por:

$$B(\hat{\gamma}_A) = \sum_{R,S,T} \kappa^{AR} \kappa^{ST} \left(\kappa_{RS}^{(T)} - \frac{1}{2} \kappa_{RST} \right) - \frac{1}{2} \sum_R \sum_{s,t} \kappa^{AR} \kappa^{st} \kappa_{Rst}. \quad (2.16)$$

As expressões a seguir são demonstradas no Apêndice D:

- $\kappa_{rS} = \kappa^{sR} = 0$,
- $\kappa_{RS}^{(T)} = \sum_{l=1}^n \left\{ \frac{d^3 d_2(\phi_l)}{d\phi_l^3} \left(\frac{\partial \phi_l}{\partial \tau_l} \right)^3 + 2 \frac{\partial^2 d_2(\phi_l)}{\partial \phi_l^2} \frac{\partial^2 \phi_l}{\partial \tau_l^2} \frac{\partial \phi_l}{\partial \tau_l} \right\} s_{lR} s_{lS} s_{lT}$,
- $\kappa_{RST} = \sum_{l=1}^n \left[\frac{d^3 d_2(\phi_l)}{d\phi_l^3} \left(\frac{\partial \phi_l}{\partial \tau_l} \right)^3 + 3 \frac{d^2 d_2(\phi_l)}{d\phi_l^2} \frac{\partial^2 \phi_l}{\partial \tau_l^2} \frac{\partial \phi_l}{\partial \tau_l} \right] s_{lR} s_{lS} s_{lT}$ e
- $\kappa_{Rst} = - \sum_{l=1}^n w_l \frac{\partial \phi_l}{\partial \tau_l} x_{ls} x_{st} s_{lR}$.

Substituindo esses resultados na expressão (2.16), e fazendo algumas manipulações algébricas, tem-se:

$$B(\hat{\gamma}_A) = \frac{1}{2} \sum_{l=1}^n \left\{ (d_{3l} \phi_{1l}^3 + d_{2l} \phi_{1l} \phi_{2l}) z_{\gamma l} + (\phi_{1l} w_l) z_{\beta l} \right\} \sum_{R=1}^q \kappa^{AR} s_{lR}, \quad (2.17)$$

em que:

- $d_{2l} = \frac{\partial^2 d_2(\phi_l)}{\partial \phi_l^2}$; $d_{3l} = \frac{\partial^3 d_2(\phi_l)}{\partial \phi_l^3}$,
- $\phi_{2l} = \frac{\partial^2 \phi_l}{\partial \tau_l^2}$,

- $z_{\gamma ll}$ é o elemento da matriz $Z_\gamma = SK_{\gamma\gamma}^{-1}S^T$ e
- $z_{\beta ll}$ é o elemento da matriz $Z_\beta = XK_{\beta\beta}^{-1}X^T$.

Então, a expressão (2.17) em notação matricial é dada por

$$B(\hat{\gamma}) = \{S^T (-D_2\Phi_1^2) S\}^{-1} S^T (-D_2\Phi_1^2) \Psi, \quad (2.18)$$

sendo:

- $D_2 = \text{diag} \{d_{2ll}\}; D_3 = \text{diag} \{d_{3ll}\},$
- $\Phi_2 = \text{diag} \{\phi_{2ll}\}$ e
- $\Psi = -\frac{1}{2} (D_2\Phi_1^2)^{-1} \{Z_{\gamma d} (D_3\Phi_1^3 + D_2\Phi_1\Phi_2) + Z_{\beta d}W\Phi_1\} \mathbf{1}.$

A expressão (2.18) é obtida como os coeficientes da regressão de mínimos quadrados ponderados com variável resposta Ψ e $-D_2\Phi_1^2$ sendo a matriz peso.

Viés de $\hat{\eta}$ e $\hat{\tau}$

Os vieses dos preditores lineares η e τ são obtidos a partir das expressões (2.15) e (2.18), respectivamente, que, após algumas álgebras, são dadas por:

$$B(\hat{\eta}) = Z_\beta\Phi W\xi \quad \text{e} \quad B(\hat{\tau}) = -Z_\gamma D_2\Phi_l^2\Psi. \quad (2.19)$$

Estimadores de $O(n^{-2})$ de β, γ, η e τ

Os estimadores corrigidos são denotados por um til sobre o parâmetro ($\tilde{}$) e dados por:

$$\tilde{\beta} = \hat{\beta} - B(\hat{\beta}), \quad \tilde{\gamma} = \hat{\gamma} - B(\hat{\gamma}), \quad \tilde{\eta} = \hat{\eta} - B(\hat{\eta}) \quad \text{e} \quad \tilde{\tau} = \hat{\tau} - B(\hat{\tau}).$$

2.2.4 Aplicação: modelando a quantidade de gases hidrocarbonos liberados

Como ilustração, os estimadores de $O(n^{-2})$ de β e γ são obtidos utilizando o aplicativo SAS, para o problema apresentado em Weisberg (1985). É importante salientar que o MLGD adotado é o mesmo que o apresentado em Smyth e Verbyla (1989) apenas para ilustrar as fórmulas obtidas.

Descrição do problema

Sabe-se que quando a gasolina é bombeada em um tanque, os vapores de hidrocarbono são forçados para o tanque e liberados para a atmosfera. Com o objetivo de reduzir esta fonte de poluição do ar, são instalados aparelhos para capturar esse vapor. Para avaliar sua eficiência, realizou-se um experimento em um laboratório, no qual a quantidade de vapor liberado foi medido sob condições controladas.

Modelo

Variável Resposta

\mathbf{Y} → **hc**: quantidade de hidrocarbonos liberados, em gramas, segue a distribuição inversa gaussiana, ou seja,

$$\mathbf{Y}_l \sim IG(\mu_l, \phi_l), \quad l = 1, 2, \dots, 125.$$

Covariáveis

\mathbf{X}_1 → **ttanque**: temperatura inicial no tanque, em $^{\circ}F$,

\mathbf{X}_2 → **tgasolina**: temperatura distribuída na gasolina, em $^{\circ}F$,

\mathbf{X}_3 → **ptanque**: pressão inicial no tanque, em *psi*,

\mathbf{X}_4 → **pgasolina**: pressão distribuída na gasolina, em *psi*.

Componentes sistemáticas do submodelo média e do submodelo dispersão

Os preditores lineares $\boldsymbol{\eta}$ e $\boldsymbol{\tau}$ são:

$$\boldsymbol{\eta} = \log(\boldsymbol{\mu}) = \beta_0 + \beta_1 \mathbf{ttanque} + \beta_2 \mathbf{tgasolina} + \beta_3 \mathbf{ptanque} + \beta_4 \mathbf{pgasolina},$$

$$\boldsymbol{\tau} = \log(\boldsymbol{\phi}) = \gamma_0 + \gamma_1 \mathbf{tgasolina} + \gamma_2 \mathbf{pgasolina}.$$

Algoritmo do SAS

I) *procedure nlmixed* : aplicado para obter as estimativas de MV dos parâmetros dos submodelos média e dispersão.

II) *procedure iml* : aplicado para obter as estimativas de $O(n^{-2})$ dos parâmetros dos submodelos média e dispersão.

Resultados

A convergência do método ocorreu na 27^a iteração. Pela Tabela 2.1, observa-se que apenas a temperatura do tanque não foi significativa no submodelo média. Quanto aos vieses, quando comparados com o seu respectivo erro padrão, observam-se que eles apresentam valores pequenos. Logo, pelos resultados do quociente pode-se dizer que os estimadores de $O(n^{-2})$ não contribuíram para a melhoria dos EMV.

Tabela 2.1: Estimativas de MV, de $O(n^{-2})$ e desvios padrão do submodelo média.

Parâmetro	Estimativa $\hat{\beta}$	Erro padrão	Viés $B(\hat{\beta})$	Quociente, % $\frac{ Viés }{Erro\ padrão}$	Estimativa $\tilde{\beta}$	Valor p Normal
Intercepto	2,47786	0,04149	-0,00011	0,26	2,47798	<0,0001
Temp. Tanque	0,00014	0,00163	5,43E-6	0,33	0,00013	0,9327
Temp. Gasolina	0,00607	0,00119	6,13E-6	0,51	0,00606	<0,0001
Pres. Tanque	-0,17460	0,04523	-0,00013	0,29	-0,17447	0,0002
Pres. Gasolina	0,30891	0,04431	-0,00003	0,07	0,30894	<0,0001

Pela Tabela 2.2, quando se comparam os vieses com seus respectivos erros padrão, observa-se que a variação foi grande. E, nesse caso, pode-se dizer que os estimadores de $O(n^{-2})$ contribuíram para a melhoria dos EMV.

Ao contrário do submodelo média, observa-se que a temperatura da gasolina não foi significativa no submodelo dispersão.

Tabela 2.2: Estimativas de MV, de $O(n^{-2})$ e desvios padrão do submodelo dispersão.

Parâmetro	Estimativa $\hat{\gamma}$	Erro padrão	Viés $B(\hat{\gamma})$	Quociente, (%) $\frac{ Viés }{Erro\ padrão}$	Estimativa $\tilde{\gamma}$	Valor p Normal
Intercepto	6,77624	0,75420	-0,05304	7,03	6,82927	<0,0001
Temp. Gasolina	-0,01929	0,01949	-0,00347	17,80	-0,01582	0,3243
Pres. Gasolina	0,59456	0,21480	0,07420	34,54	0,52035	0,0065

Pelos resultados desta aplicação, observa-se que o uso dos estimadores de $O(n^{-2})$ não deve ser ignorado quando os dois submodelos são considerados.

2.3 Procedimento de estimação dos MLGD por Smyth e Verbyla (1999)

Este método foi apresentado por Smyth e Verbyla (1999) para a família exponencial de dispersão linear, dos quais os parâmetros do submodelo média são estimados pelo MMV e do submodelo dispersão pelo Método de Máxima Verossimilhança Restrita (MMVR).

O MMVR, segundo os autores, produz estimadores para os parâmetros do submodelo dispersão com viés menor do que a estimativa de MV. Para comparar os estimadores de $O(n^{-2})$ com os EMVR, é apresentado um estudo de simulação no final dessa seção.

O procedimento apresentado, por Smyth e Verbyla (1999), consiste em reescrever a função densidade de probabilidade da variável resposta em função de sua *deviance*, isto é,

$$p(y_l; \theta_l, \phi_l) = c(y_l, \phi_l) \exp \left\{ -\frac{1}{2} d(y_l, \mu_l) \right\}, \quad (2.20)$$

no qual d é uma medida de distância entre y e μ , e para a maioria das distribuições d é obtido como $d(y_l, \mu_l) = 2 \{t(y_l, y_l) - t(y_l, \mu_l)\}$ e $t(y_l, \mu_l) = y_l \theta_l - b(\theta_l)$.

Jørgensen (1997, p. 103) estudou a família de distribuições dada em (2.20) para diferentes valores de d , demonstrando que, para a maioria das distribuições de \mathbf{Y} , μ pode ser interpretado como parâmetro de locação e ϕ como parâmetro de dispersão. Os autores utilizam a aproximação ponto de sela, em que a variável aleatória d_l é distribuída como uma gama, isto é,

$$d_l \sim G(\phi_l, 2)$$

sendo $\phi_l = E(d_l)$ e 2 o parâmetro de dispersão. Portanto, a componente aleatória do submodelo de dispersão tem as características apresentadas na Tabela 2.3.

Tabela 2.3: Componentes do submodelo dispersão para estimar γ .

Componente	Valor assumido
Variável Resposta	d_l
Média	ϕ_l
Função de Variância	$V_d(\phi) = \phi^2$
Função de Ligação	$h = \log(\phi)$
Dispersão	2

EMV de γ usando aproximação ponto de sela

Assim, o EMV dos parâmetros do submodelo dispersão, denotados por $\tilde{\gamma}$, é obtido solucionando o seguinte sistema de equações, iterativamente:

$$S^T W_\gamma S \tilde{\gamma} = S^T W_\gamma \mathbf{z}_\gamma, \quad (2.21)$$

sendo a matriz de pesos W_γ dada por:

$$W_\gamma = \text{diag} \left\{ \frac{1}{\left(\frac{\partial h(\phi_l)}{\partial \phi_l} \right)^2 V_\gamma(\phi_l)} \right\}$$

e o vetor de respostas ajustadas

$$\mathbf{z}_\gamma = \frac{1}{\phi_l} (d_l - \phi_l) + \log(\phi_l).$$

EMVR

O MMVR consiste em maximizar a variância não das observações originais, mas de um conjunto de contrastes de média zero, com o objetivo de ajustar os graus de liberdade disponíveis e produzir estimadores que são aproximadamente não-viesados. Segundo Smyth e Verbyla (1999), a generalização do método de máxima verossimilhança restrita ou residual para modelos não normais não é óbvia, uma vez que não existem contrastes com média zero. Nesse caso, a matriz chapéu é dada por:

$$H = W_m^{1/2} X (X^T W_m X)^{-1} X^T W_m^{1/2}.$$

A partir dessa matriz, os autores obtêm a matriz peso e o vetor de respostas ajustado como sendo:

- $W_\gamma^* = W_\gamma - 2diag \left\{ \frac{h_l}{\phi_l^2 \left(\frac{dh(\phi_l)}{d\phi_l} \right)^2} \right\} + H^2$, h_l são os elementos da matriz diagonal de H e $h(\phi_l)$ representa a função de ligação da dispersão e
- $z_{dl}^* = \frac{\partial h(\phi_l)}{\partial \phi_l} \{d_l - (1 - h_l) \phi_l\} + h(\phi_l)$.

Assim sendo, o EMVR, denotado por $\check{\gamma}^*$, é a solução do seguinte sistema de equações:

$$S^T W_\gamma^* S \check{\gamma}^* = S^T W_\gamma^* z_\gamma^*.$$

2.3.1 Estudo de simulação: comparando EMVR e estimadores de $O(n^{-2})$

Nesta seção, o interesse é comparar os EMVR com os estimadores de $O(n^{-2})$ e, além disso, ilustrar o procedimento proposto por Smyth e Verbyla (1999).

Simulação da inversa gaussiana

Para amostras de tamanho 20, 30, 40 e 45 não foi possível obter as estimativas de $\check{\gamma}$, pela *procedure* NLMIXED, do SAS, devido a não convergência do processo, sendo que isto ocorreu para amostra de tamanho 50. Portanto, foram simuladas 10.000 amostras, de tamanho 50, considerando a variável resposta com distribuição inversa gaussiana que, como a gama, é muito usada:

- no mercado de finanças, quando o objetivo é descrever a volatilidade financeira diária, como em Forsberg (2002);

- em cartas de controle, pois no processo de produção de uma indústria as distribuições são positivas e assimétricas, como em Hawkins e Olwell (1997), entre outros.

Considerou-se a componente aleatória $Y \sim IG(\mu, \phi)$, sendo μ a média e ϕ a dispersão.

O submodelo da média toma a seguinte forma:

$$\eta = \log(\mu) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3,$$

e, para o submodelo da dispersão,

$$\tau = \log(\phi) = \gamma_1 s_1 + \gamma_2 s_2.$$

As matrizes modelo X e S foram simuladas de uma distribuição $U(0, 1)$, sendo fixadas para todas as 10.000 amostras.

Foram utilizados os seguintes valores para os parâmetros $\beta = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$ e $\gamma = \begin{pmatrix} 2 & 3 \end{pmatrix}$. A Figura 2.1 ilustra o processo de simulação.

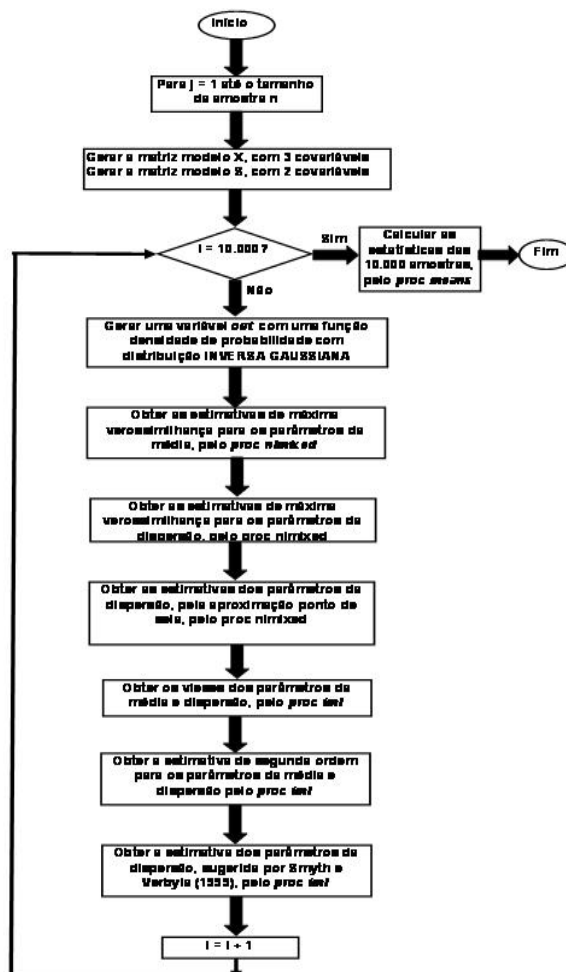


Figura 2.1: Fluxograma do MLGD com distribuição inversa gaussiana.

Portanto, serão obtidos:

- EMV de $O(n^{-1})$ $\hat{\beta}$, $\hat{\gamma}$,
- viés de $\hat{\beta}$, $\hat{\gamma}$,
- estimador de $O(n^{-2})$ de β , γ e
- EMVR, $\tilde{\gamma}^*$.

2.3.2 Resultados

Para amostras de tamanho pequeno, não foi possível a obtenção das estimativas de $\tilde{\gamma}$, pela *procedure* NLMIXED, devido a não convergência do processo. E isto ocorreu para amostras a partir do tamanho 50. Assim, realizou-se a simulação apenas para este tamanho, uma vez que este pode ser considerado o tamanho de amostra moderado.

A Tabela 2.4 apresenta a média e o desvio padrão das estimativas para as 10.000 amostras. Observa-se que todas as estimativas convergem para o verdadeiro parâmetro e que, em média, os desvios padrão das estimativas de $O(n^{-2})$ são menores que as demais.

Tabela 2.4: Tabela das médias e desvios padrão.

Estimativa	Média	Desvio padrão
$\hat{\beta}_1$	0,950667	0,745817
$\hat{\beta}_2$	1,996136	0,059512
$\hat{\beta}_3$	3,028843	0,478166
$\tilde{\beta}_1$	0,952256	0,739449
$\tilde{\beta}_2$	1,996146	0,059500
$\tilde{\beta}_3$	3,028419	0,477618
$\hat{\gamma}_1$	1,973722	0,516332
$\hat{\gamma}_2$	2,970330	0,594779
$\tilde{\gamma}_1$	1,983251	0,470063
$\tilde{\gamma}_2$	2,982076	0,468009
$\tilde{\gamma}_1^*$	1,932108	0,560401
$\tilde{\gamma}_2^*$	3,098265	0,661342

Os diagramas de caixa, apresentados nas Figuras 2.2 e 2.3, mostram que, para os parâmetros do submodelo dispersão, as estimativas de $O(n^{-2})$ são melhores, isto é, os desvios padrão destas são menores quando comparadas às outras.

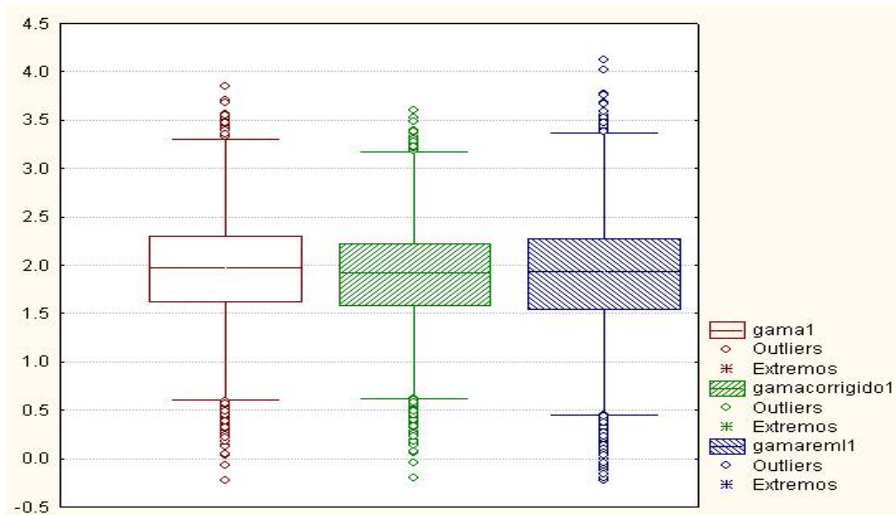


Figura 2.2: Diagrama de caixas da distribuição das estimativas de γ_1 .

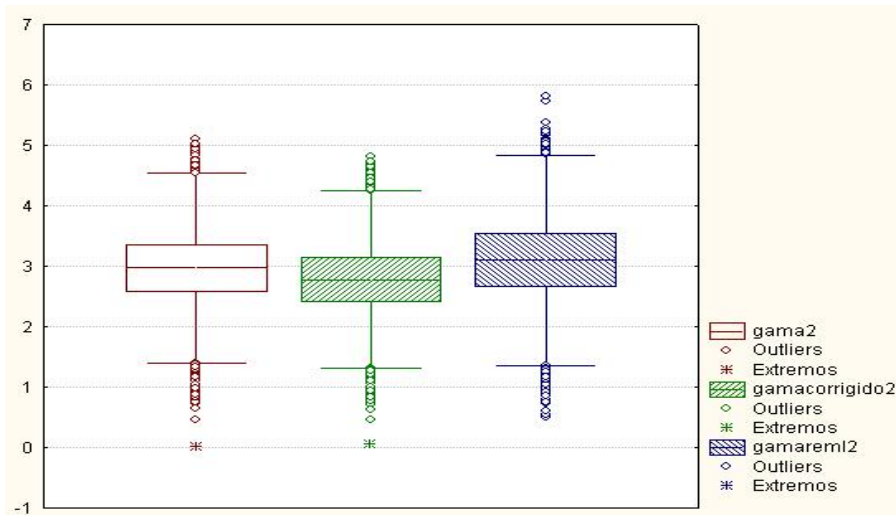


Figura 2.3: Diagrama de caixas da distribuição das estimativas de γ_2 .

Considerando este estudo de simulação, verificou-se que a proposta de Smyth e Verbyla (1999) é de difícil convergência para amostras de tamanho pequeno. E, ainda, mesmo com uma amostra de tamanho 50, considerada moderado, as estimativas se equiparam às de MV, concluindo, portanto, que as estimativas de $O(n^{-2})$ são melhores, isto é, são mais homogêneas.

Capítulo 3

Modelos não-lineares da família exponencial

Neste capítulo, é apresentada uma continuação da revisão de literatura, não só para facilitar o desenvolvimento teórico desta Tese, mas também para mostrar o uso prático dos estimadores de $O(n^{-2})$.

Os MNLFE são extensões dos MLG, ou seja, o preditor $\boldsymbol{\eta}$, que nos MLG é linear, neste modelo é permitido que seja não-linear.

Estes modelos são aplicados na farmacocinética, ecologia e produção industrial, entre outros.

3.1 O modelo

Os MNLFE ficam caracterizados por uma componente aleatória, por uma componente sistemática da média e por uma função de ligação, isto é,

a) componente aleatória: $\mathbf{Y} = (Y_1, \dots, Y_n)^T$, no qual Y_l tem uma função densidade de probabilidade pertencente à família exponencial:

$$p(y_l; \theta_l, \phi) = \exp [\{\phi \{y_l \theta_l - b(\theta_l) + c(y_l)\} + d_1(y_l) + d_2(\phi)\}], \quad (3.1)$$

com parâmetros canônicos θ_l e de precisão ϕ , sendo $E(Y_l) = \mu_l = \frac{\partial b(\theta_l)}{\partial \theta_l}$, $V(Y_l) = \phi^{-1} V_l$ e $V_l = V(\mu_l) = \frac{\partial \mu_l}{\partial \theta_l}$ que é chamada função de variância.

b) componente sistemática:

$$\eta_l = g(\mu_l) = f(x_{jl}; \boldsymbol{\beta}), \quad l = 1, \dots, n \quad \text{e} \quad j = 1, \dots, p, \quad (3.2)$$

sendo

- $f(\cdot; \cdot)$ uma função não-linear conhecida com relação a um vetor $\boldsymbol{\beta}$,

- $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ um vetor de parâmetros desconhecidos, $\boldsymbol{\beta} \in R^p$, $p < n$,
- X uma matriz modelo de posto p .

Logo, esse modelo fica caracterizado por sua aplicabilidade para qualquer membro da família exponencial linear, e a presença da função de ligação que conecta o preditor $\boldsymbol{\eta}$ ao valor esperado de \mathbf{Y} .

As condições de regularidade requeridas por (3.2) são apresentadas em Wei (1998, p. 15).

3.2 Função log-verossimilhança, função escore e matriz de informação de Fisher de $\boldsymbol{\beta}$ e ϕ

Os estimadores dos parâmetros serão obtidos pelo MMV e os de interesse são $\boldsymbol{\beta}$, ϕ , θ e μ .

A função log-verossimilhança para os parâmetros desse modelo é dada por:

$$\ell(\boldsymbol{\beta}) = \phi \{ \mathbf{y}^T \boldsymbol{\theta} - b(\boldsymbol{\theta}) + c(\mathbf{y}) \} + d_1(\mathbf{y}) + d_2(\phi) \quad (3.3)$$

ou equivalentemente

$$\ell(\boldsymbol{\beta}) = \phi \sum_{l=1}^n [\{ y_l \theta_l - b(\theta_l) + c(y_l) \} + d_1(y_l) + d_2(\phi)].$$

A função escore de $\boldsymbol{\beta}$, denotada por $U_{\boldsymbol{\beta}}$, está deduzida no Apêndice E, considerando ϕ constante, e apresenta o seguinte resultado:

$$U_{\beta_r} = \sum_{l=1}^n \phi (y_l - \mu_l) \frac{1}{V_l} \frac{\partial \mu_l}{\partial \eta_l}(r),$$

sendo $(r) = \frac{\partial \eta_l}{\partial \beta_r}$. E, fazendo as somas, tem-se a seguinte forma matricial:

$$U_{\boldsymbol{\beta}} = \phi \tilde{X}^T W^{1/2} V (\mathbf{y} - \boldsymbol{\mu}), \quad (3.4)$$

sendo

- $\tilde{X} = \frac{\partial f(X; \boldsymbol{\beta})}{\partial \boldsymbol{\beta}}$ e
- $W = \text{diag}\{w_1, w_2, \dots, w_n\}$, $w_l = \frac{1}{V_l} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2$.

A matriz de informação de Fisher de $\boldsymbol{\beta}$, denotada por $K_{\boldsymbol{\beta}\boldsymbol{\beta}}$, deduzida no Apêndice E, considerando ϕ constante é dada por:

$$K_{\beta_r \beta_s} = - \sum_{l=1}^n \phi w_l (r, s),$$

sendo $(r, s) = \frac{\partial \eta_r}{\partial \beta_r} \frac{\partial \eta_s}{\partial \beta_s}$. Portanto, a matriz de informação de Fisher para $\boldsymbol{\beta}$ é dada por:

$$K_{\boldsymbol{\beta}\boldsymbol{\beta}} = \phi \tilde{X}^T W \tilde{X}. \quad (3.5)$$

A função escore e a matriz de informação de Fisher de ϕ , denotadas por U_ϕ e $K_{\phi\phi}$, respectivamente, são dadas por:

- $U_\phi = \frac{\partial \ell}{\partial \phi} = \mathbf{y}^T \boldsymbol{\theta} - b(\boldsymbol{\theta}) + c(\mathbf{y}) + \frac{\partial d_2(\phi)}{\partial \phi}$ e
- $K_{\phi\phi} = -\frac{\partial^2 d_2(\phi)}{\partial \phi \partial \phi^T}$.

A matriz de informação de Fisher de $\boldsymbol{\beta}$ e ϕ é:

$$K(\boldsymbol{\beta}, \phi) = \begin{pmatrix} K_{\boldsymbol{\beta}\boldsymbol{\beta}} & K_{\boldsymbol{\beta}\phi} \\ K_{\phi\boldsymbol{\beta}} & K_{\phi\phi} \end{pmatrix} = \begin{pmatrix} \phi \tilde{X}^T W \tilde{X} & 0 \\ 0 & -\frac{\partial^2 d(\mathbf{y}; \phi)}{\partial \phi \partial \phi^T} \end{pmatrix}.$$

Portanto, $\boldsymbol{\beta}$ e ϕ são ortogonais, logo, podem ser estimados separadamente.

3.3 EMV de $\boldsymbol{\beta}$

O EMV de $\boldsymbol{\beta}$ é obtido igualando a expressão (3.4) a zero, isto é,

$$\tilde{X}^T W^{1/2} V^{-1/2} (\mathbf{y} - \boldsymbol{\mu}) = 0, \quad (3.6)$$

pois ϕ é uma constante. Todavia, este sistema de equações não é linear, logo, só poderá ser resolvido numericamente por um processo iterativo, que pode ser:

- A)** método de Newton-Raphson, que está disponibilizado na *procedure* NLP e NLMIXED, do aplicativo SAS. Este método consiste em expandir a função log-verossimilhança em série de Taylor, em uma vizinhança de $\boldsymbol{\beta}_0$, como apresentado no Apêndice B, que resulta em:

$$\boldsymbol{\beta}^{i+1} = \boldsymbol{\beta}^i + \left(-\frac{\partial^2 \ell(\boldsymbol{\beta}^i)}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} \right)^{-1} \frac{\partial \ell(\boldsymbol{\beta}^i)}{\partial \boldsymbol{\beta}^T}, \quad i = 1, 2, 3, \dots$$

onde $\boldsymbol{\beta}^i$ é a i -ésima iteração.

- B)** método escore de Fisher, que está disponibilizado na *procedure* GENMOD do aplicativo SAS. Esse método é uma aproximação do Newton-Raphson, no qual o valor $-\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T}$ é substituído por $E\left(-\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T}\right)$. O EMV para este método na iteração, $i + 1$, consiste em resolver um sistema de equações reponderado, isto é,

$$\hat{\boldsymbol{\beta}}^{i+1} = \left\{ \left(\tilde{X}^T W \tilde{X} \right)^{-1} \tilde{X}^T W \mathbf{z} \right\}^i,$$

sendo

$$\mathbf{z} = \boldsymbol{\eta} + \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\mu}} (\mathbf{y} - \boldsymbol{\mu}),$$

$$\text{em que } \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\mu}} = \text{diag} \left\{ \frac{\partial \eta_1}{\partial \mu_1}, \frac{\partial \eta_2}{\partial \mu_2}, \dots, \frac{\partial \eta_n}{\partial \mu_n} \right\}.$$

A convergência do processo iterativo é rápida, mas depende fortemente da escolha do valor inicial $\boldsymbol{\beta}_0$. Quanto mais próximo o valor inicial do verdadeiro parâmetro, mais fácil a sua convergência.

Cordeiro e Paula (1989b) apresentam um procedimento que permite que a estimativa dos parâmetros dos MNLFE seja feita através da *procedure* GENMOD do SAS, utilizando o comando *offset*.

3.4 Estimadores de $O(n^{-2})$

Paula (1991) apresenta os vieses do EMV e os estimadores de $O(n^{-2})$ dos parâmetros $\boldsymbol{\beta}$, $\boldsymbol{\eta}$, $\boldsymbol{\mu}$ e ϕ , que serão abordados nessa seção, uma vez que essa Tese trata de modelos cujo preditor $\boldsymbol{\eta}$ é não-linear.

3.4.1 Vieses do estimador de $\boldsymbol{\beta}$, $\boldsymbol{\eta}$ e ϕ

Viés de $\hat{\boldsymbol{\beta}}$

A expressão de Cox e Snell (1968) é utilizada para a obtenção do viés de $\hat{\beta}_a$, denotado por $B(\hat{\beta}_a)$, para $a = 1, \dots, p$, como no Capítulo 2. Nesta expressão, são necessários os cumulantes dados a seguir, que são calculados no Apêndice E, considerando ϕ constante:

$$\kappa_{rs,t} = \phi \sum_l [g_l(r, s, t)_l + w_l(t, rs)_l];$$

$$\kappa_{rst} = - \sum_{l=1}^n \phi (f_l + 2g_l)(r, s, t)_l + w_l \{ (r, st)_l + (s, tr)_l + (t, rs)_l \}.$$

Substituindo esses cumulantes na expressão de Cox e Snell, e fazendo algumas manipulações algébricas, tem se:

$$B(\hat{\boldsymbol{\beta}}) = \left(\tilde{X}^T W \tilde{X} \right)^{-1} \tilde{X}^T W \boldsymbol{\xi}, \quad (3.7)$$

no qual:

- $\boldsymbol{\xi} = \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2;$
- $\boldsymbol{\xi}_1 = -(2\phi)^{-1} Z_d W^{-1} F \mathbf{1}$ e $\boldsymbol{\xi}_2 = -(2\phi)^{-1} D_{\boldsymbol{\beta}} \mathbf{1};$

- $f_{ll} = V^{-1} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial^2 \mu_l}{\partial \eta_l^2}$ é o elemento diagonal da matriz F ;
- Z_d é a diagonal da matriz $Z = \tilde{X} \left(\tilde{X}^T W \tilde{X} \right)^{-1} \tilde{X}^T$;
- $D_\beta = \text{diag} \{d_1, d_2, \dots, d_n\}$ e $d_l = \text{tr} \left\{ \tilde{X}_l \left(\tilde{X}_l^T W \tilde{X}_l \right)^{-1} \right\}$, \tilde{X}_l é uma matriz $p \times p$ cujos elementos são $\frac{\partial^2 \eta}{\partial \beta_r \partial \beta_s}$.

Observa-se que a expressão (3.7) nada mais é que uma regressão de mínimos quadrados ponderada, cuja variável resposta é $\boldsymbol{\xi}$ e a matriz peso é W .

Viés de $\hat{\boldsymbol{\eta}}$

Para o cálculo do viés de $\hat{\boldsymbol{\eta}} = f \left(X, \hat{\boldsymbol{\beta}} \right)$, aplica-se a expansão em série de Taylor; em seguida, a esperança, e obtém-se,

$$B(\hat{\boldsymbol{\eta}}) = - (2\phi)^{-1} \{ Z Z_d F + (Z W - I) D_\beta \} \mathbf{1}. \quad (3.8)$$

Estimadores de $O(n^{-2})$

Substituindo um estimador consistente de ϕ nas expressões (3.7) e (3.8), obtém-se os estimadores corrigidos de $\boldsymbol{\beta}$ e $\boldsymbol{\eta}$, denotados $\tilde{\boldsymbol{\beta}}$ e $\tilde{\boldsymbol{\eta}}$, dados, respectivamente, por

$$\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}} - B(\hat{\boldsymbol{\beta}}) \quad \text{e} \quad \tilde{\boldsymbol{\eta}} = \hat{\boldsymbol{\eta}} - B(\hat{\boldsymbol{\eta}}).$$

3.5 Aplicação: modelando a quantidade de vendas de produtos ao consumidor

Como ilustração, serão obtidos os estimadores corrigidos de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$, utilizando o aplicativo SAS, para o problema apresentado em Whitmore (1986), onde o interesse é modelar as quantidades de vendas reais e as projetadas para 20 produtos de uma companhia.

Descrição do problema

Uma organização produz um inventário de mercado e vende um relatório, o qual apresenta projeções, isto é, estimativas da venda anual, em dólares, para todos os produtos de todas as companhias, em particular para a indústria de produtos ao consumidor. As projeções são feitas monitorando as quantidades de vendas em um painel da saída de vendas a varejo. Qualquer companhia em uma atividade que compra o relatório do inventário é capaz de comparar a quantidade real de vendas para esses n produtos ($Y_l, l = 1, 2, \dots, n$) com as quantidades projetadas no relatório ($x_l, l = 1, 2, \dots, n$).

Na Tabela 3.1, estão apresentados os dados fornecidos por Whitmore (1986) sobre as quantidades de vendas de produtos de uma certa companhia.

Tabela 3.1: Dados das quantidades de vendas reais e estimadas.

Produto (i)	Estimado (x_i)	Real (y_i)	Produto (i)	Estimado (x_i)	Real (y_i)
1	5.950	5.673	11	527	487
2	3.534	3.659	12	353	463
3	2.641	2.565	13	331	225
4	1.965	2.182	14	290	257
5	1.738	1.839	15	253	311
6	1.182	1.236	16	193	212
7	667	918	17	156	166
8	613	902	18	133	123
9	610	756	19	122	198
10	549	500	20	114	99

Variáveis

\mathbf{Y}_l → resposta: quantidade real vendida, em dólares;

\mathbf{X}_l → covariável: quantidade projetada para venda, em dólares.

Modelo Componente aleatório

Whitmore (1986) sugere a distribuição inversa gaussiana para descrever a variável resposta com:

$$\mathbf{Y}_l \sim IG\left(\beta_1 x_l^{\beta_2}, \phi\right), \quad l = 1, 2, \dots, 20.$$

Neste caso, ϕ foi considerado constante, mas desconhecido.

Componente Sistemática

$$\eta_l = \frac{1}{\mu_l} = \beta_1 x_l^{\beta_2}$$

Algoritmo do SAS

Passo 1) *procedure nlin*: aplicado para obter os valores iniciais.

Passo 2) *macro*: obtenção das estimativas

a) *procedure iml* utilizada para estimar β , e em cada iteração i ,

a1) dar as primeiras derivadas da matriz modelo, $\tilde{X} = \frac{\partial \eta}{\partial \beta}$;

a2) considerar $\boldsymbol{\tau} = \boldsymbol{\eta} - \tilde{X}\boldsymbol{\beta}$, para o uso do comando *offset*, de Cordeiro e Paula (1989), permitindo o uso do *procedure genmod* com vetor de respostas ajustadas dado por $z^* = \boldsymbol{\eta} + \text{diag} \left\{ \frac{\partial \eta_l}{\partial \mu_l} \right\} [\mathbf{y} - \boldsymbol{\mu}] - \boldsymbol{\tau}$; $l = 1, 2, \dots, 20$, onde \tilde{X} é dada por:

$$\tilde{X} = \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\beta}} = \begin{pmatrix} \frac{\partial \eta_l}{\partial \beta_1} & \frac{\partial \eta_l}{\partial \beta_2} \end{pmatrix} = \begin{pmatrix} x_l^{\beta_2} & \beta_1 x_l^{\beta_2} \log(x_l) \end{pmatrix}_{n \times 2};$$

b) em cada iteração i , atualizar os valores de \tilde{X} , z^* , $\boldsymbol{\eta}$ para estimar $\boldsymbol{\beta}^i$.

Passo 3) *procedure iml*: calcula-se o viés de $\hat{\boldsymbol{\beta}}$, utilizando um arquivo saída, do *procedure genmod*, onde os resultados da matriz peso W e do vetor de estimativas do preditor linear $\boldsymbol{\eta}$ são retirados. Agora são necessárias:

- matriz de peso $W = \text{diag} \{ \mu_l \}$;
- matriz $F = \text{diag} \left\{ -\frac{2}{\eta_l^2} \right\}$;
- cada uma das 20 matrizes das derivadas segundas, que é dada por:

$$\tilde{X}_l = \begin{pmatrix} \frac{\partial^2 \eta_l}{\partial \beta_1^2} & \frac{\partial^2 \eta_l}{\partial \beta_1 \partial \beta_2} \\ \frac{\partial^2 \eta_l}{\partial \beta_2 \partial \beta_1} & \frac{\partial^2 \eta_l}{\partial \beta_2^2} \end{pmatrix} = \begin{pmatrix} 0 & x_l^{\beta_2} \log(x_l) \\ x_l^{\beta_2} \log(x_l) & \beta_1 x_l^{\beta_2} [\log(x_l)]^2 \end{pmatrix}.$$

Nota: Uma medida para verificar a qualidade do ajuste é o AIC, critério de informação de Akaike, quando se deseja comparar modelos ajustados que pertençam a distribuições diferentes e é penalizada pelo número de parâmetros. É calculado por:

$$AIC = -2\log\text{-verossimilhança} + 2p,$$

sendo p número de parâmetros. Segundo Pinheiro e Bates (2000, p. 84), Ogliari (1998, p. 93) e Davidian e Giltman (1995, p. 156), o melhor modelo é aquele que apresenta menor valor de AIC.

Resultados

O processo convergiu na 3ª iteração e os resultados estão na Tabela 3.2, em que pode-se observar que os parâmetros β_1 e β_2 apresentam erros padrão pequenos e, para avaliar a magnitude do viés, calculou-se o quociente deste com o seu respectivo erro padrão.

Tabela 3.2: Estimativa de MV, de $O(n^{-2})$, desvios padrão, quociente e AIC.

Parâmetro	Estimativa $\hat{\beta}$	Erro padrão	Viés	Quociente $\frac{ Viés }{Erro\ padrão} \times 100$	Estimativa $\tilde{\beta}$
β_1	0,9242	0,1095	0,0056	5,14	0,9186
β_2	-0,9868	0,0713	-0,0047	6,60	-0,9822

$AIC = 88,7867$

Pelo resultado, as estimativas de $O(n^{-2})$ parecem influenciar pouco nas conclusões.

Capítulo 4

Estimadores dos modelos não-lineares da família exponencial de dispersão linear

Neste capítulo, são obtidos os EMV e os estimadores de $O(n^{-2})$ dos MNLFED. Os estimadores de $O(n^{-2})$ são a contribuição teórica desta Tese.

Como referenciado no capítulo da Introdução, esses modelos são importantes, pois, tanto na área tecnológica como na biológica, observam-se fenômenos que apresentam estruturas não-lineares, distribuições assimétricas, e o parâmetro de dispersão pode depender de covariáveis. A resolução desse problema é complexa, porém, pode ser superada por meio do avanço dos computadores e do uso de aplicativos não só estatísticos, mas também matemáticos.

Para enfatizar os resultados obtidos nesse capítulo, são apresentadas, a seguir, as expressões dos vieses para o submodelo média e dispersão:

$$B(\hat{\beta}) = (\tilde{X}\Phi W\tilde{X})^{-1} \tilde{X}\Phi W\xi$$

e

$$B(\hat{\gamma}) = (\tilde{S}(-D_2\Phi_1^2)\tilde{S})^{-1} \tilde{S}(-D_2\Phi_1^2)\psi.$$

4.1 Modelo

Os MNLFED ficam caracterizados por uma componente aleatória, pelas componentes sistemáticas para média e dispersão e pelas funções de ligação, isto é,

a) componente aleatória: $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)^T$, variáveis aleatórias independentemente distribuídas, nos quais os Y_i 's são provenientes da família exponencial de dispersão, cuja

função densidade de probabilidade é dada por:

$$p(y; \theta_l, \phi_l) = \exp [\phi_l \{y\theta_l - b(\theta_l) + c(y)\} + d_1(y) + d_2(\phi_l)],$$

sendo $b(\cdot)$, $c(\cdot)$, $d_1(\cdot)$ e $d_2(\cdot)$ funções conhecidas. A média e a variância de Y_l são $E(Y_l) = \frac{\partial b(\theta_l)}{\partial \theta_l} = \mu_l$, $V(Y_l) = \phi^{-1}V_l$ com $V_l = V(\mu_l) = \frac{d\mu_l}{d\theta_l}$, como uma função de variância. Os parâmetros θ_l e $\phi_l > 0$ são chamados canônico e de precisão, respectivamente.

b) componentes sistemáticas do:

submodelo média: as covariáveis supostamente fixadas x_j , $j = 1, \dots, p$ apresentam uma estrutura não-linear, isto é,

$$\eta_l = g(\mu_l) = f_1(x_{jl}, \beta_l), \quad l = 1, \dots, n;$$

submodelo dispersão: as covariáveis supostamente fixadas s_k , $k = 1, \dots, q$, formam uma estrutura não-linear, isto é,

$$\tau_l = h(\phi_l) = f_2(s_{kl}, \gamma_l), \quad l = 1, \dots, n.$$

4.2 Funções score e informação de Fisher de β e ϕ

Como as componentes sistemáticas dos submodelos da média e dispersão são não-lineares, isso torna o cálculo dos cumulantes das derivadas da função log-verossimilhança mais complexo, como observa-se no Apêndice E, em que são calculadas as derivadas da log-verossimilhança e os cumulantes dos MNLFED.

Função score de β

O procedimento para obter a função score de β , denotada por U_β , a partir do Apêndice E, é análogo ao apresentado no Capítulo 3 e é dado por:

$$U_\beta = \tilde{X}^T \Phi W^{1/2} V^{-1/2} (\mathbf{y} - \boldsymbol{\mu}), \quad (4.1)$$

sendo $\Phi = \text{diag}\{\phi_1, \dots, \phi_n\}$.

Função score de γ

A função score de γ é obtida com mais detalhes por ser a primeira vez que é referida nesta Tese. Seu cálculo parte da derivada primeira da função log-verossimilhança com relação a γ , que é demonstrado no Apêndice E, cuja expressão é dada por:

$$\begin{aligned}
U_{\gamma_R} &= \frac{\partial \ell(\beta_l, \gamma_l)}{\partial \phi_l} \frac{\partial \phi_l}{\partial \tau_l} \frac{\partial \tau_l}{\partial \gamma_R} \\
&= \sum_{l=1}^n \left[\{y_l \theta_l - b(\theta_l) + c(y_l)\} + \frac{dd_2(\phi_l)}{d\phi_l} \right] \frac{\partial \phi_l}{\partial \tau_l} \frac{\partial \tau_l}{\partial \gamma_R} \\
&= \sum_{l=1}^n [\{y_l \theta_l - b(\theta_l) + c(y_l)\} \phi_{1l}(R)_l + d_{1l} \phi_{1l}(R)_l],
\end{aligned}$$

sendo:

- $\frac{\partial \tau_l}{\partial \gamma_R} = (R)_l$; $\frac{\partial^2 \tau_l}{\partial \gamma_R \partial \gamma_S} = (RS)_l$ e $\frac{\partial \tau_l}{\partial \gamma_R} \frac{\partial \tau_l}{\partial \gamma_S} = (R, S)_l$;
- $\frac{\partial^i \phi_l}{\partial \tau_l^i} = \phi_{il}$ e $\Phi_i = \text{diag} \{\phi_{i1}, \dots, \phi_{in}\}$, para $i = 1, 2$;
- $\frac{\partial d_2(\phi_l)}{\partial \phi_l} = d_{1l}$ e $\mathbf{d}_1 = \{d_{11}, \dots, d_{1n}\}^T$, somando sob todos os l tem-se, em notação matricial, que:

$$U_{\gamma} = \tilde{S}^T \Phi_1 \boldsymbol{\nu},$$

sendo

- $\boldsymbol{\nu} = \mathbf{y}^T \boldsymbol{\theta} + b(\boldsymbol{\theta}) + c(\mathbf{y}) + \mathbf{d}_1$;
- $\tilde{S} = \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\gamma}} = \frac{\partial f(S; \boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}}$.

Função escore $(\boldsymbol{\beta}, \boldsymbol{\gamma})$

A função escore $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ é representada por um vetor particionado como segue:

$$U(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \begin{pmatrix} \frac{\partial \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\beta}} \\ \frac{\partial \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}} \end{pmatrix} = \begin{pmatrix} \tilde{X}^T \Phi W^{1/2} V^{-1/2} (\mathbf{y} - \boldsymbol{\mu}) \\ \tilde{S}^T \Phi_1 \boldsymbol{\nu} \end{pmatrix}.$$

Matriz de informação de Fisher de $\boldsymbol{\beta}$

A matriz de informação de Fisher de $\boldsymbol{\beta}$ para os MNFED, a partir do resultado do Apêndice E, é obtida de modo análogo aos MNLFE, sendo dada por:

$$K_{\beta\beta} = \tilde{X}^T \Phi W \tilde{X}.$$

Matriz de informação de Fisher de $\boldsymbol{\gamma}$

No Apêndice E, demonstra-se que:

$$\begin{aligned}
-U_{\gamma_R \gamma_S} &= -\frac{\partial^2 \ell}{\partial \gamma_R \partial \gamma_S} = -\sum_l \{(y_l \theta_l - b(\theta_l) + c(y_l)) (\phi_{2l}(R, S)_l + \phi_{1l}(RS)_l)\} \\
&\quad - \sum_l \{d_{2l} (\phi_{1l})^2 (R, S)_l\} - \sum_l \{\phi_{1l} \phi_{2l} (R, S)_l\} \\
&\quad - \sum_l \{d_{1l} \phi_{1l} (RS)_l\},
\end{aligned}$$

sendo

- $\frac{\partial^i d_2(\phi_l)}{\partial \phi_l^i} = d_{il}$ e $D_i = \text{diag} \{d_{i1}, \dots, d_{in}\}$, para $i = 2, 3$;
- $\phi_{2l} = \frac{\partial^2 \phi_l}{\partial \gamma_l^2}$ e $\Phi_2 = \text{diag} \{\phi_{21}, \dots, \phi_{2n}\}$.

Para obter a matriz $K_{\gamma\gamma} = E\left(-\frac{\partial^2 \ell}{\partial \gamma_R \partial \gamma_S}\right)$, aplica-se a esperança nessa equação e após algumas manipulações algébricas obtém-se:

$$K_{\gamma\gamma} = \tilde{S}^T (-D_2 \Phi_1^2) \tilde{S}.$$

Para esses modelos, demonstra-se que $K_{\beta\gamma} = K_{\gamma\beta} = \mathbf{0}$, e para mais detalhes ver Apêndice E. Portanto, a matriz de informação de Fisher particionada de β e γ forma a seguinte matriz bloco diagonal:

$$K(\beta, \gamma) = \begin{pmatrix} K_{\beta\beta} & K_{\beta\gamma} \\ K_{\beta\gamma} & K_{\gamma\gamma} \end{pmatrix} = \begin{pmatrix} \tilde{X}^T \Phi W \tilde{X} & \mathbf{0} \\ \mathbf{0} & \tilde{S}^T (-D_2 \Phi_1^2) \tilde{S} \end{pmatrix}. \quad (4.2)$$

Logo, os parâmetros β e γ são globalmente ortogonais.

4.3 EMV de β e γ

Para estimar os parâmetros β e γ , basta solucionar o sistema de equações $U(\hat{\beta}, \hat{\gamma}) = 0$. Esse sistema de equações não é linear, logo, só poderá ser resolvido numericamente por um processo iterativo.

Submodelo média

O método de Newton-Raphson está disponibilizado nas *procedures* NLP e NLMIXED do aplicativo SAS. Este método consiste em expandir a função log-verossimilhança em série de Taylor, em uma vizinhança de β_0 , como pode ser visto no Apêndice B, para uma dada iteração $i + 1$, tem-se:

$$\beta^{i+1} = \beta^i + \left(-\frac{\partial^2 \ell(\beta^i)}{\partial \beta \partial \beta^T}\right)^{-1} \frac{\partial \ell(\beta^i)}{\partial \beta^T}.$$

O método escore de Fisher está disponibilizado na *procedure* GENMOD do aplicativo SAS. Esse método é uma aproximação do Newton-Raphson no qual o valor $-\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T}$ é substituído por $E\left(-\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T}\right)$. Assim, tem-se:

$$\boldsymbol{\beta}^{i+1} = \left(\hat{\boldsymbol{\beta}} - K_{\boldsymbol{\beta}\boldsymbol{\beta}}^{-1} U_{\boldsymbol{\beta}}\right)^i,$$

substituindo $K_{\boldsymbol{\beta}\boldsymbol{\beta}}^{-1}$ e $U_{\boldsymbol{\beta}}$ e fazendo algumas manipulações algébricas obtém-se o EMV, para a iteração $i + 1$, que é dado por:

$$\hat{\boldsymbol{\beta}}^{i+1} = \left\{ \left(\tilde{X}^T \Phi W \tilde{X} \right)^{-1} \tilde{X}^T \Phi W z_{\boldsymbol{\beta}} \right\}^i,$$

sendo

$$z_{\boldsymbol{\beta}} = \tilde{X} \boldsymbol{\beta} + \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\mu}} (\mathbf{y} - \boldsymbol{\mu}),$$

em que $\frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\mu}} = \text{diag} \left(\frac{\partial \eta_l}{\partial \mu_l} \right)$.

Submodelo dispersão

O método de Newton-Raphson é a solução do sistema

$$\boldsymbol{\gamma}^{i+1} = \boldsymbol{\gamma}^i + \left(-\frac{\partial^2 \ell(\boldsymbol{\gamma}^i)}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}^T} \right)^{-1} \frac{\partial \ell(\boldsymbol{\gamma}^i)}{\partial \boldsymbol{\gamma}^T}.$$

Para o método escore de Fisher, a quantil $-\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T}$ é substituída por $E\left(-\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T}\right)$. Logo,

$$\boldsymbol{\gamma}^{i+1} = \left(\boldsymbol{\gamma} - K_{\boldsymbol{\gamma}\boldsymbol{\gamma}}^{-1} U_{\boldsymbol{\gamma}}\right)^i,$$

em seguida, substituindo os valores de $K_{\boldsymbol{\gamma}\boldsymbol{\gamma}}^{-1}$ e $U_{\boldsymbol{\gamma}}$, e após algumas manipulações algébricas tem-se:

$$\boldsymbol{\gamma}^{i+1} = \left\{ \left(\tilde{S}^T (-D_2 \Phi_1^2) \tilde{S} \right)^{-1} \tilde{S}^T (-D_2 \Phi_1^2) z_{\boldsymbol{\gamma}} \right\}^i$$

sendo

$$z_{\boldsymbol{\gamma}} = \tilde{S} \boldsymbol{\gamma} - D_2^{-1} \Phi_1^{-1} \boldsymbol{\nu}, \quad (4.3)$$

em que $\boldsymbol{\nu} = \mathbf{y}^T \boldsymbol{\theta} + b(\boldsymbol{\theta}) + c(\mathbf{y}) + \mathbf{d}_1$.

4.4 Vieses de $\hat{\boldsymbol{\beta}}$ e $\hat{\boldsymbol{\gamma}}$

Esta seção é a contribuição teórica desta Tese; logo, os vieses de $\hat{\boldsymbol{\beta}}$ e $\hat{\boldsymbol{\gamma}}$ são demonstrados com certo rigor.

Viés de $\hat{\beta}$

O viés de $\hat{\beta}_a$ é obtido por meio da expressão (2.13), do Capítulo 2, para $a = 1, \dots, p$. Assim, é dado por:

$$B(\hat{\beta}_a) = \sum_{r,s,t=1}^p \kappa^{ar} \kappa^{st} \left(\kappa_{rs}^{(t)} - \frac{1}{2} \kappa_{rst} \right) + \sum_{r=1}^p \sum_{S,T=1}^q \kappa^{ar} \kappa^{ST} \left(\kappa_{rS}^{(T)} - \frac{1}{2} \kappa_{rST} \right). \quad (4.4)$$

Os cumulantes do MNLFED são demonstrados no Apêndice E e apresentam os seguintes resultados:

- $\kappa_{rs}^{(t)} = - \sum_{l=1}^n \phi_l [(f_l + g_l)(r, s, t)_l + w_l \{(r, st) + (s, rt)\}]$;
- $\kappa_{rst} = - \sum_{l=1}^n \phi_l (f_l + 2g_l)(r, s, t)_l + w_l \{(r, st)_l + (s, tr)_l + (t, rs)_l\}$;
- $\kappa_{rS}^{(T)} = 0$ e $\kappa_{rST} = 0$.

Portanto, substituindo esses resultados na expressão (4.4) obtém-se

$$B(\hat{\beta}_a) = -\frac{1}{2} \sum_l \phi_l \left\{ \left(\sum_r \kappa^{ar} (r)_l \right) \left(\sum_{t,u} \kappa^{st} (s, t)_l f_l \right) + \sum_r \sum_{s,t} w_l \kappa^{ar} \kappa^{st} \{(r, st)_l + (s, rt)_l - (t, rs)_l\} \right\}.$$

Após alguns cálculos e rearranjando convenientemente os somatórios, o viés de $\hat{\beta}$, na notação matricial, é dado por:

$$\begin{aligned} B(\hat{\beta}) &= -\frac{1}{2} \left(\tilde{X}^T \Phi W \tilde{X} \right)^{-1} \tilde{X}^T Z_{\beta d} F \Phi \mathbf{1} - \frac{1}{2} \left(\tilde{X}^T \Phi W \tilde{X} \right)^{-1} \tilde{X}^T \Phi W D_{\beta} \\ &= \left(\tilde{X}^T \Phi W \tilde{X} \right)^{-1} \tilde{X}^T \Phi W \left\{ -\frac{1}{2} W^{-1} Z_{\beta d} F \mathbf{1} - \frac{1}{2} D_{\beta} \mathbf{1} \right\}, \end{aligned} \quad (4.5)$$

sendo:

- $D_{\beta} = \text{diag} \{d_l\}$, $d_l = \text{traço} \left\{ \tilde{X}_l \left(\tilde{X}^T \Phi W \tilde{X} \right)^{-1} \right\}$ e \tilde{X}_l uma matriz $p \times p$ cujos elementos $(st)_l = \frac{\partial^2 \eta_l}{\partial \beta_s \partial \beta_t}$,
- $Z_{\beta} = \tilde{X} \left(\tilde{X}^T \Phi W \tilde{X} \right)^{-1} \tilde{X}^T$ e $Z_{\beta d} = \text{diag} (Z_{\beta})$.

Na expressão (4.5) substitui-se $\left\{ -\frac{1}{2} W^{-1} Z_{\beta d} F \mathbf{1} - \frac{1}{2} D_{\beta} \mathbf{1} \right\}$ por $\boldsymbol{\xi}$ resultando em:

$$B(\hat{\beta}) = \left(\tilde{X}^T \Phi W \tilde{X} \right)^{-1} \tilde{X}^T \Phi W \boldsymbol{\xi}. \quad (4.6)$$

É importante salientar que $\boldsymbol{\xi}$ pode ser decomposta em duas componentes, isto é, $\boldsymbol{\xi} = \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2$, sendo

$$\boldsymbol{\xi}_1 = -\frac{1}{2}W^{-1}Z_{\beta d}F\mathbf{1} \quad \text{e} \quad \boldsymbol{\xi}_2 = -\frac{1}{2}D_{\beta}\mathbf{1}. \quad (4.7)$$

Portanto, o viés é um conjunto de coeficientes de uma regressão de mínimos quadrados ponderados, cuja variável resposta é $\boldsymbol{\xi} = \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2$, com a matriz \tilde{X} sendo a matriz de covariável e ΦW a matriz peso. Logo, para sua obtenção, basta um aplicativo com módulo de regressão linear ponderado.

Viés de $\hat{\gamma}$

O viés de $\hat{\gamma}_A$, denotado por $B(\hat{\gamma}_A)$, para $A = 1, \dots, q$, será obtido a partir de Cox e Snell (1968), que, nesse caso, como nos MLGD, resulta em:

$$B(\hat{\gamma}_A) = \sum_{R,S,T} \kappa^{AR} \kappa^{ST} \left(\kappa_{RS}^{(T)} - \frac{1}{2} \kappa_{RST} \right) - \frac{1}{2} \sum_R \sum_{s,t} \kappa^{AR} \kappa^{st} \kappa_{Rst}, \quad (4.8)$$

cujos cumulantes desta expressão são obtidos no Apêndice E e são dados por:

- $\kappa_{RS}^{(T)} = \sum_l \{d_{3l}(\phi_{1l})^3 + 2d_{2l}\phi_{1l}\phi_{2l}\} (R, S, T)_l + \sum_l d_{2l}(\phi_{1l})^2 \{(R, ST)_l + (S, RT)_l\}$;
- $\kappa_{RST} = \sum_l \{d_{3l}(\phi_{1l})^3 + 3d_{2l}\phi_{1l}\phi_{2l}\} (R, S, T)_l + \sum_l \{d_{2l}(\phi_{1l})^2\} \{(R, ST)_l + (S, RT)_l + (T, RS)_l\}$;
- $\kappa_{Rst} = -\sum_{l=1}^n w_l \phi_{1l} (R, s, t)_l$.

Para obter $B(\hat{\gamma}_A)$, é necessário substituir estes cumulantes na expressão (4.8). Para facilitar, foi decomposta a equação em duas parcelas, sendo a primeira:

$$\begin{aligned} & \sum_{R,S,T} \kappa^{AR} \kappa^{ST} \left(\kappa_{RS}^{(T)} - \frac{1}{2} \kappa_{RST} \right) = \\ & = \sum_{R,S,T} \left[\kappa^{AR} \kappa^{ST} \sum_l (d_{3l}(\phi_{1l})^3 + d_{2l}\phi_{1l}\phi_{2l}) (R, S, T)_l + d_{2l}(\phi_{1l})^2 \{(R, ST)_l + (S, RT)_l - (T, RS)_l\} \right], \end{aligned}$$

e a segunda:

$$\frac{1}{2} \sum_{\beta, \gamma} \kappa^{AR} \kappa^{st} \kappa_{Rst} = \frac{1}{2} \sum_{s,t} \sum_R \kappa^{AR} \kappa^{st} \sum_l \phi_{1l} w_l (R, s, t)_l.$$

Em seguida, subtraindo a primeira da segunda parcela e rearranjando os somatórios obtém-se:

$$\begin{aligned}
B(\hat{\gamma}_A) &= \frac{1}{2} \sum_{l=1}^n \left\{ (d_{3l}(\phi_{1l})^3 + d_{2l}\phi_{1l}\phi_{2l}) \sum_{R,S,T} \kappa^{AR}(R)_l \kappa^{ST}(S,T)_l \right\} \\
&\quad + \frac{1}{2} \sum_{l=1}^n d_{2l}(\phi_{1l})^2 \sum_{R,S,T} \kappa^{AR}(R)_l \kappa^{ST}(ST)_l + \frac{1}{2} \sum_{l=1}^n \phi_{1l} w_l \sum_{s,t} \sum_R \kappa^{AR}(R)_l \kappa^{st}(s,t)_l.
\end{aligned} \tag{4.9}$$

Fazendo os somatórios e rearranjando obtém-se a notação matricial,

$$B(\hat{\gamma}) = \frac{1}{2} \left[\left(K_{\gamma}^{-1} \tilde{S}^T \right) Z_{\gamma d} (D_3 \Phi_1^3 + D_2 \Phi_1 \Phi_2) + \left(K_{\gamma\gamma}^{-1} \tilde{S}^T \right) Z_{\beta d} W \Phi_1 + \left(K_{\gamma\gamma}^{-1} \tilde{S}^T \right) D_2 \Phi_1^2 D_{\gamma} \right],$$

sendo:

- $Z_{\gamma} = \tilde{S} K_{\gamma\gamma}^{-1} \tilde{S}^T$ e $Z_{\gamma d} = \text{diag} \{z_{\gamma 11}, \dots, z_{\gamma mn}\}$;
- $D_{\gamma} = \text{diag} \{d_{ll}\}$ em que $d_{ll} = \text{traço} \left\{ \tilde{S}_l \left(\tilde{S}^T (-D_2 \Phi_1^2) \tilde{S} \right)^{-1} \right\}$ e \tilde{S}_l são os elementos da matriz $\frac{\partial^2 \tau}{\partial \gamma_s \partial \gamma_T} = (ST)_l$, $l = 1, 2, \dots, n$.

Para simplificar, considere

$$\Psi = \Psi_1 + \Psi_2, \tag{4.10}$$

sendo:

- $\Psi_1 = -\frac{1}{2} (D_2 \Phi_1^2)^{-1} [Z_{\gamma d} (D_3 \Phi_1^3 + D_2 \Phi_1 \Phi_2) + Z_{\beta d} W \Phi_1] \mathbf{1}$;
- $\Psi_2 = -\frac{1}{2} D_{\gamma} \mathbf{1}$.

Portanto, a expressão simplificada para o viés de $\hat{\gamma}$ resulta em:

$$B(\hat{\gamma}) = \left(\tilde{S}^T (-D_2 \Phi_1^2) \tilde{S} \right)^{-1} \tilde{S}^T (-D_2 \Phi_1^2) \Psi. \tag{4.11}$$

Assim, o viés de $\hat{\gamma}$ é um conjunto de coeficientes de uma regressão de mínimos quadrados ponderados, cuja variável resposta é $\Psi = \Psi_1 + \Psi_2$, sendo \tilde{S} a matriz modelo e $(-D_2 \Phi_1^2)$ a matriz peso. Logo, para o seu cálculo, basta um aplicativo com módulo de regressão linear ponderada.

Estimador de $O(n^{-2})$ de β e γ

Como nos capítulos anteriores, estes estimadores nada mais são do que a diferença entre o EMV e o viés, pelo Apêndice C, e serão denotados por um til sobre o parâmetro, isto é,

$$\tilde{\beta} = \hat{\beta} - B(\hat{\beta}) \quad \text{e} \quad \tilde{\gamma} = \hat{\gamma} - B(\hat{\gamma}).$$

4.5 Caso especial

Nesta seção, é demonstrado que as expressões obtidas na seção 4 estendem os resultados de Cordeiro e Botter (1998), que é um dos objetivos específicos dessa Tese.

Para os MLGD, as componentes sistemáticas são lineares, portanto,

$$\eta_l = g(\mu_l) = f_1(x_{jl}, \beta_l) = X\boldsymbol{\beta} \quad e \quad \tau_l = h(\phi_l) = f_2(s_{jl}, \gamma_l) = S\boldsymbol{\gamma}.$$

Observa-se que para estes preditores:

- $\tilde{X} = X \Rightarrow Z_{\beta d} = X (X^T \Phi W X)^{-1} X^T$;
- $\tilde{X}_l = 0 \Rightarrow D_\beta = 0 \Rightarrow \boldsymbol{\xi}_2 = \mathbf{0}$, portanto a componente $\boldsymbol{\xi} = \boldsymbol{\xi}_1 = -\frac{1}{2} W^{-1} Z_{\beta d} F \mathbf{1}$;
- $\tilde{S} = S \Rightarrow Z_{\gamma d} = S (S^T (-D_2 \Phi_1^2) S)^{-1} S^T$;
- $\tilde{S}_l = 0 \Rightarrow D_\gamma = 0 \Rightarrow \Psi_2 = 0$;
- $\Psi = \Psi_1 = -\frac{1}{2} (D_2 \Phi_1^2)^{-1} [Z_{\gamma d} (D_3 \Phi_1^3 + D_2 \Phi_1 \Phi_2) + Z_{\beta d} W \Phi_1] \mathbf{1}$.

Substituindo nas expressões dos vieses de $\hat{\boldsymbol{\beta}}$ e $\hat{\boldsymbol{\gamma}}$, isto é, nas expressões (4.6) e (4.11), obtêm-se:

$$B(\hat{\boldsymbol{\beta}}) = (X^T \Phi W X)^{-1} X^T \Phi W \boldsymbol{\xi}_1 \quad e \quad B(\hat{\boldsymbol{\gamma}}) = (S^T (-D_2 \Phi_1^2) S)^{-1} S^T (-D_2 \Phi_1^2) \Psi_1.$$

Estas expressões são as mesmas apresentadas por Botter e Cordeiro (1998) para os vieses de $\hat{\boldsymbol{\beta}}$ e $\hat{\boldsymbol{\gamma}}$, respectivamente.

4.6 Viés de $\hat{\boldsymbol{\eta}}$, $\hat{\boldsymbol{\tau}}$, $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\phi}}$ e $\hat{\boldsymbol{\phi}}^{-1}$

Viés de $\hat{\boldsymbol{\eta}}$

Para se obter o viés de $\hat{\boldsymbol{\eta}}$, denotado por $B(\hat{\boldsymbol{\eta}})$, expande-se $\hat{\boldsymbol{\eta}} = f(X, \hat{\boldsymbol{\beta}})$ em série de Taylor, na vizinhança de $\hat{\boldsymbol{\beta}}$, isto é,

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

Rearranjando os termos

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

Aplicando a esperança, tem-se:

$$B(\hat{\boldsymbol{\eta}}) = \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\beta}} B(\hat{\boldsymbol{\beta}}) + \frac{1}{2} \frac{\partial^2 \boldsymbol{\eta}}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} V(\hat{\boldsymbol{\beta}}).$$

Substituindo $\mathbf{V}(\hat{\boldsymbol{\beta}})$ por $\text{diag}(K_{\beta\beta}^{-1})$ e $B(\hat{\boldsymbol{\beta}})$ pela expressão (4.6), resulta em:

$$B(\hat{\boldsymbol{\eta}}) = -\frac{1}{2} \{Z_{\beta} Z_{\beta d} F \mathbf{1} + (Z_{\beta} W \Phi - I) D_{\beta}\} \mathbf{1}. \quad (4.12)$$

Viés de $\hat{\boldsymbol{\tau}}$

O viés de $\hat{\boldsymbol{\tau}}$ é obtido analogamente ao anterior, isto é, expande-se $f(S, \hat{\boldsymbol{\gamma}})$ em série de Taylor, na vizinhança de $\boldsymbol{\gamma}$, aplica-se a esperança e considera-se $V(\hat{\boldsymbol{\gamma}}) = \text{diag}(K_{\gamma\gamma}^{-1})$. Após alguns cálculos algébricos, tem-se:

$$\mathbf{B}(\hat{\boldsymbol{\tau}}) = Z_{\gamma} (-D_2 \Phi_1^2) \Psi_1 + \{Z_{\gamma} (-D_2 \Phi_1^2) - I\} \Psi_2. \quad (4.13)$$

Viés de $\hat{\boldsymbol{\mu}}$

Obtém-se, agora, o viés de $\hat{\boldsymbol{\mu}}$ expandindo $g^{-1}(\hat{\boldsymbol{\eta}}) = \hat{\boldsymbol{\mu}}$ em série de Taylor, na vizinhança de $\boldsymbol{\eta}$, isto é,

$$g^{-1}(\hat{\boldsymbol{\eta}}) - g^{-1}(\boldsymbol{\eta}) = \frac{\partial g^{-1}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} (\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}) + \frac{1}{2} \frac{\partial^2 g^{-1}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}^T} (\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}) (\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})^T.$$

Após aplicar a esperança, considera-se somente os termos da diagonal da matriz de covariância $E[(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})^T]$, isto é, $V(\hat{\boldsymbol{\eta}}) = \text{diag}\{E[(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})^T]\} = Z_{\beta d}$, uma vez que é o viés de cada componente. Agora, considerando $G_1 = \text{diag}\left\{\frac{\partial \mu_l}{\partial \eta_l}\right\}$, $G_2 = \text{diag}\left\{\frac{\partial^2 \mu}{\partial \eta^2}\right\}$ e substituindo em (4.12), o viés de $\hat{\boldsymbol{\mu}}$ resulta em:

$$B(\hat{\boldsymbol{\mu}}) = \frac{1}{2} (G_2 Z_{\beta d} \mathbf{1} + 2G_1 Z_{\beta} W \Phi \xi_1) - \frac{1}{2} G_1 (Z_{\beta} W \Phi - I) D_{\beta} \mathbf{1}. \quad (4.14)$$

Viés de $\hat{\boldsymbol{\phi}}$

Para o cálculo do viés do parâmetro de precisão $\boldsymbol{\phi}$, $h^{-1}(\hat{\boldsymbol{\tau}}) = \hat{\boldsymbol{\phi}}$, é necessário expandir $h^{-1}(\hat{\boldsymbol{\tau}})$ em série de Taylor, na vizinhança de $\boldsymbol{\tau}$, isto é,

$$h^{-1}(\hat{\boldsymbol{\tau}}) - h^{-1}(\boldsymbol{\tau}) = \frac{\partial h^{-1}(\boldsymbol{\tau})}{\partial \boldsymbol{\tau}} (\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}) + \frac{1}{2} \frac{\partial^2 h^{-1}(\boldsymbol{\tau})}{\partial \boldsymbol{\tau} \partial \boldsymbol{\tau}^T} (\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}) (\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})^T.$$

Em seguida, aplicando esperança, considera-se somente os termos da diagonal da matriz de covariância $E[(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})^T]$, isto é, $V(\hat{\boldsymbol{\tau}}) = \text{diag}\{E[(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})^T]\} = Z_{\gamma d}$, uma

vez que é o viés de cada componente. Ao substituir (4.13) na expressão anterior, e realizar algumas manipulações algébricas, obtém-se:

$$B(\hat{\phi}) = \frac{1}{2} (\Phi_2 Z_{\gamma d} \mathbf{1} - 2\Phi_1 Z_{\gamma} \Phi_l^2 D_2 \Psi_1) - (Z_{\gamma} D_2 \phi_1^2 + I) \Psi_2. \quad (4.15)$$

Viés de $\hat{\phi}^{-1}$

Finalmente, o viés do parâmetro de dispersão ϕ^{-1} é obtido também expandindo $(g^{-1}(\hat{\tau}))^{-1} = \hat{\phi}^{-1}$, em série de Taylor, na vizinhança de τ . Aplicando a esperança, resulta em:

$$\begin{aligned} B(\hat{\phi}^{-1}) &= \frac{\partial (h(\tau)^{-1})^{-1}}{\partial \tau} B(\hat{\tau}) + \frac{1}{2} \frac{\partial^2 (h(\tau)^{-1})^{-1}}{\partial \tau \partial \tau^T} V(\hat{\tau}) \\ &= -\frac{\partial \phi}{\partial \tau} (h(\tau)^{-1})^{-2} B(\hat{\tau}) + \frac{1}{2} \left\{ \frac{2 \left(\frac{d\phi}{d\tau} \right)^2 - \frac{d^2 \phi}{d\tau d\tau^T} h(\tau)^{-1}}{(h(\tau)^{-1})^3} \right\} V(\hat{\tau}). \end{aligned}$$

Em seguida, substituindo (4.13) e considerando $V(\hat{\tau}_l) = Z_{\gamma d}$, na expressão anterior, obtém-se a seguinte forma matricial:

$$\mathbf{B}(\hat{\phi}^{-1}) = \frac{1}{2} \{ (2\Phi_1^2 - \Phi_2 \Phi) \Phi^{-3} Z_{\gamma d} \mathbf{1} + 2\Phi_1 \Phi^{-2} Z_{\gamma} D_2 \Phi_1^2 \Psi_1 \} + \frac{1}{2} (Z_{\gamma} D_2 \Phi_1^2 + I) D_{\gamma} \mathbf{1},$$

que resulta em:

$$\mathbf{B}(\hat{\phi}^{-1}) = \frac{1}{2} \{ (2\Phi_1^2 - \Phi_2 \Phi) \Phi^{-3} Z_{\gamma d} \mathbf{1} + 2\Phi_1 \Phi^{-2} Z_{\gamma} D_2 \Phi_1^2 \Psi_1 \} - (Z_{\gamma} D_2 \Phi_1^2 + I) \Psi_2 - (Z_{\gamma} D_2 \Phi_1^2 + I) \Psi_2. \quad (4.16)$$

Estimadores de $O(n^{-2})$

Concluindo, os estimadores de $O(n^{-2})$ para os parâmetros η , τ , $\tilde{\mu}$, $\tilde{\phi}$ e $\tilde{\phi}^{-1}$ são obtidos, respectivamente, por:

$$\tilde{\eta} = \hat{\eta} - B(\hat{\eta}), \quad \tilde{\tau} = \hat{\tau} - B(\hat{\tau}), \quad \tilde{\mu} = \hat{\mu} - \mathbf{B}(\hat{\mu}), \quad \tilde{\phi} = \hat{\phi} - B(\hat{\phi}) \quad \text{e} \quad \tilde{\phi}^{-1} = \hat{\phi}^{-1} - B(\hat{\phi}^{-1}).$$

4.7 Aplicação: modelando a quantidade de vendas de produtos ao consumidor

Apresenta-se aqui a mesma aplicação do Capítulo 3, considerando, agora, os submodelos da média e dispersão.

4.7.1 Modelo

Componente aleatório

Neste Capítulo, seguindo a sugestão de Whitmore (1986), é adotada como variável resposta a distribuição inversa gaussiana:

$$Y_l \sim IG\left(\beta_1 x_l^{\beta_2}, \gamma_1^{-1} x_l^{\gamma_2}\right), \quad l = 1, 2, \dots, 20.$$

Componentes Sistemáticas:

a) Submodelo média:

$$\eta_l = \frac{1}{\mu_l} = \beta_1 x_l^{\beta_2}.$$

b) Submodelo dispersão:

$$\tau_l = \phi_l = \frac{x_l^{\gamma_2}}{\gamma_1}.$$

Estimativas de $O(n^{-1})$:

- O estimador $\tilde{\beta}$ foi obtido de forma análoga ao do Capítulo 3.
- Calcula-se, inicialmente, o viés de $\hat{\gamma}$ e, em seguida, $\tilde{\gamma}$. Para isso, é preciso:

matriz dos parâmetros

$$\gamma^T = \begin{pmatrix} \gamma_1 & \gamma_2 \end{pmatrix}_{20 \times 2},$$

matriz da primeira derivada

$$\tilde{S}^T = \begin{pmatrix} -\frac{s^{\gamma_2}}{\gamma_1^2} & \frac{s^{\gamma_2} \log(s)}{\gamma_1^2} \end{pmatrix}_{20 \times 2}, \text{ sendo } S = X,$$

matriz peso • $-D_2 \phi_1^2 = \text{diag} \left\{ -\left(-\frac{1}{2\phi^2} \right) \right\}_{20 \times 20}$,

matriz das segundas derivadas, $\tilde{\tilde{S}}_l$

$$\tilde{\tilde{S}}_l = \begin{pmatrix} \frac{2s^{\gamma_2}}{\gamma_1^3} & -\frac{s^{\gamma_2}}{\gamma_1^2} \\ -\frac{s^{\gamma_2}}{\gamma_1^2} & \frac{s^{\gamma_2} \log(s)}{\gamma_1} \end{pmatrix} \text{ e } l = 1, 2, \dots, 20.$$

4.7.2 Resultados

Os resultados das estimativas encontram-se na Tabela 4.1, juntamente com os quocientes dos seus respectivos erros padrão, para avaliar sua magnitude. Pode-se observar que o submodelo média apresenta para o quociente entre o viés e o erro padrão valores próximos de zero, enquanto que no submodelo de dispersão o mesmo não ocorre. Logo, neste último caso, as estimativas de $O(n^{-2})$ parecem necessárias.

Tabela 4.1: Estimativas de MV, de $O(n^{-2})$ e erros padrão.

Parâmetro	Estimativa de MV	Erro padrão	Quociente	p valor	Estimativa de $O(n^{-2})$
β_1	1,09701	0,03808	0,3	<0,0001	1,09689
β_2	0,92716	0,02616	0,2	<0,0001	0,92710
γ_1	0,00984	0,00403	23,1	0,0241	0,01212
γ_2	2,48626	0,44290	36,1	<0,0001	2,32586
$AIC = 27,65135$					

E, ainda, comparando o AIC deste modelo com o do Capítulo 3, considerando ϕ constante, observa-se que este está melhor ajustado, pois aqui o valor do AIC foi 27,65 e, no outro caso, pela Tabela 3.2, foi de 88,78. Concluindo, a necessidade de modelar a dispersão.

Capítulo 5

Avaliando estimadores de $O(n^{-2})$ através de um estudo de simulação em MNLFED

Neste capítulo, é realizado um estudo de simulação para mostrar a importância dos estimadores de $O(n^{-2})$ em experimentos onde o tamanho da amostra é pequeno. Isto, ocorre quando se deseja controlar a qualidade de um produto, como em processos industriais que necessitam ser caracterizados em amostra de tamanho pequeno, para que defeitos nos produtos sejam detectados logo, permitindo, assim, a eliminação de perdas e, conseqüentemente, aumento do lucro.

Foram realizadas simulações para variáveis respostas com distribuição inversa gaussiana e gama que podem ser utilizadas para descrever a volatilidade financeira diária, segundo Forsberg (2002).

Para se obter as estimativas de MV, foram utilizadas as *procedures* NLP e NLMIXED do SAS, que usam o método de Newton-Raphson. Segundo Wolfinger (2004), apesar da *procedure* NLMIXED do SAS ser utilizada para ajustar modelos com efeito aleatório, pode também ser utilizada quando isto não ocorre.

5.1 Simulação da variável resposta com distribuição inversa gaussiana

O primeiro modelo simulado, pertencente à classe dos MNFED, foi a distribuição inversa gaussiana, que é muito utilizada para fazer a previsão do risco de vida em atuária ou em cartas de controle, pois o processo de produção de uma indústria apreta distribuição positiva e assimétrica, segundo Hawkins e Olwell (1997).

Os resultados para comparar os EMV $\hat{\beta}$ e $\hat{\gamma}$, seus correspondentes vieses e os estimadores

corrigidos $\tilde{\beta}$ e $\tilde{\gamma}$, foram efetuados via simulação de Monte Carlo.

5.1.1 Submodelos média e dispersão

Adotou-se para a componente sistemática da média o modelo Gompertz, que, como as demais curvas de crescimento, apresenta dois parâmetros biologicamente relevantes. Um deles, β_1 , a assíntota da função, interpretada como o tamanho médio de maturidade, e o outro, β_3 , é a taxa de crescimento até o tamanho adulto, sendo chamado índice de maturidade. O terceiro parâmetro, β_2 , não tem significado especial. Este modelo foi utilizado por Ogliari (1998) para descrever a média da variável resposta: volume sólido com casca de árvores de eucalipto em diversos anos.

As componentes sistemáticas adotadas são dadas como se segue.

Submodelo média

O modelo Gompertz é dado por:

$$\eta = \mu = \beta_1 \exp(-\exp(\beta_2 - \beta_3 x)). \quad (5.1)$$

Submodelo dispersão

E, considerou-se uma variabilidade como sendo:

$$\tau = \phi = \exp(\gamma_1 s_1 + \gamma_2 s_2).$$

Algoritmo do SAS

Passo 1) *Macro*: para gerar 10.000 amostras de tamanho n , no qual os valores de n foram 20, 30, 40 e 60.

Passo 2) Gerar matrizes modelo X e S com distribuição uniforme, $U(0, 1)$, sendo fixadas para todas as simulações, com igual tamanho de amostra.

Passo 3) Os parâmetros do modelo foram fixados como sendo

$$\beta = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \text{ e } \gamma = \begin{pmatrix} 1,5 & 2,5 \end{pmatrix}.$$

Passo 4) Gerar amostras de tamanho n , cuja variável resposta, $Y \sim IG(\mu, \phi)$, isto é, a sua função densidade de probabilidade, é dada por:

$$p(y; \mu, \phi) = \left(\frac{\phi}{2\pi y^3}\right)^{1/2} \exp\left(-\phi \frac{(y - \mu)^2}{2\mu^2 y}\right).$$

Para isto, utilizou-se a transformação apresentada por Chhikara e Folks (1989, p. 53).

Passo 5) *Procedure NLP*: para estimar os parâmetros β e γ .

Passo 6) *Procedure IML*: para obter os vieses.

Após a simulação de cada amostra, são obtidos os vieses e as estimativas de $O(n^{-2})$. Para obter os vieses, são necessárias algumas matrizes, apresentadas a seguir:

- Matriz $\tilde{X} = \begin{pmatrix} \frac{\partial \eta_l}{\partial \beta_1} & \frac{\partial \eta_l}{\partial \beta_2} & \frac{\partial \eta_l}{\partial \beta_3} \end{pmatrix}_{n \times 3}$, onde cada $\begin{pmatrix} \frac{\partial \eta_l}{\partial \beta_j} \end{pmatrix}$ é um vetor de n elementos e $j = 1, 2$ e 3 , que é dado por:

$$\Rightarrow \frac{\partial \eta_l}{\partial \beta_1} = \exp(-\exp(\beta_2 - \beta_3 x)),$$

$$\Rightarrow \frac{\partial \eta_l}{\partial \beta_2} = -\beta_1 \exp(\beta_2 - \beta_3 x) \exp(-\exp(\beta_2 - \beta_3 x)) \text{ e}$$

$$\Rightarrow \frac{\partial \eta_l}{\partial \beta_3} = \beta_1 x \exp(\beta_2 - \beta_3 x) \exp(-\exp(\beta_2 - \beta_3 x)).$$

- Matriz peso

$$\Phi W = \text{diag} \left(\phi_l \frac{1}{\mu_l^3} \right).$$

- Matriz F , cujo elemento diagonal é

$$f_l = \frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial^2 \mu_l}{\partial \eta_l^2} = 0,$$

uma vez que a ligação é identidade, isto é, $\frac{\partial \mu_l}{\partial \eta_l} = 1$ e, conseqüentemente, $\frac{\partial^2 \mu_l}{\partial \eta_l^2} = 0$.

- Matriz \tilde{X}_l é dada por:

$$\tilde{X}_l = \begin{pmatrix} \frac{\partial^2 \eta_l}{\partial \beta_1^2} & \frac{\partial^2 \eta_l}{\partial \beta_1 \partial \beta_2} & \frac{\partial^2 \eta_l}{\partial \beta_1 \partial \beta_3} \\ \frac{\partial^2 \eta_l}{\partial \beta_1 \partial \beta_2} & \frac{\partial^2 \eta_l}{\partial \beta_2^2} & \frac{\partial^2 \eta_l}{\partial \beta_2 \partial \beta_3} \\ \frac{\partial^2 \eta_l}{\partial \beta_1 \partial \beta_3} & \frac{\partial^2 \eta_l}{\partial \beta_2 \partial \beta_3} & \frac{\partial^2 \eta_l}{\partial \beta_3^2} \end{pmatrix}$$

no qual cada elemento da matriz é dado por:

$$\Rightarrow \frac{\partial^2 \eta_l}{\partial \beta_1^2} = 0;$$

$$\Rightarrow \frac{\partial^2 \eta_l}{\partial \beta_1 \partial \beta_2} = -\exp(\beta_2 - \beta_3 x) \exp(-\exp(\beta_2 - \beta_3 x));$$

$$\Rightarrow \frac{\partial^2 \eta_l}{\partial \beta_1 \partial \beta_3} = x \exp(\beta_2 - \beta_3 x) \exp(-\exp(\beta_2 - \beta_3 x));$$

$$\begin{aligned} \Rightarrow \frac{\partial^2 \eta_l}{\partial \beta_2^2} &= -\beta_1 \exp(\beta_2 - \beta_3 x) \exp(-\exp(\beta_2 - \beta_3 x)) + \beta_1 \{\exp(\beta_2 - \beta_3 x)\}^2 \exp(-\exp(\beta_2 - \beta_3 x)); \\ \Rightarrow \frac{\partial^2 \eta_l}{\partial \beta_2 \partial \beta_3} &= \beta_1 x \exp(\beta_2 - \beta_3 x) \exp(-\exp(\beta_2 - \beta_3 x)) - \beta_1 x \{\exp(\beta_2 - \beta_3 x)\}^2 \exp(-\exp(\beta_2 - \beta_3 x)); \\ \Rightarrow \frac{\partial^2 \eta_l}{\partial \beta_3^2} &= -\beta_1 x^2 \exp(\beta_2 - \beta_3 x) \exp(-\exp(\beta_2 - \beta_3 x)) + \beta_1 x^2 \{\exp(\beta_2 - \beta_3 x)\}^2 \exp(-\exp(\beta_2 - \beta_3 x)). \end{aligned}$$

- $\tilde{S} = \left(\frac{\partial \tau_l}{\partial \gamma_1}, \frac{\partial \tau_l}{\partial \gamma_2} \right)$, no qual cada $\left(\frac{\partial \eta_l}{\partial \gamma_k} \right)$ é um vetor de n elementos e $k = 1$ e 2 , logo, para o submodelo dispersão, a matriz da primeira derivada é dada por:

$$\tilde{S} = \begin{pmatrix} s1 \exp(\gamma_1 s1 + \gamma_2 s2) & s2 \exp(\gamma_1 s1 + \gamma_2 s2) \end{pmatrix}_{n \times 2}.$$

- Matrizes:

$$\Rightarrow \text{peso: } -D_2 \phi_1^2 = \text{diag} \left(\frac{1}{2\phi_1^2} \right);$$

$$\Rightarrow D_3 = \text{diag} \left(\frac{1}{\phi_l^3} \right);$$

$$\Rightarrow \Phi_1 = I_{n \times n}, \text{ sendo } I \text{ a matriz identidade e}$$

$$\Rightarrow \Phi_2 = 0_{n \times n}, \text{ pois a ligação é identidade.}$$

- Em seguida, para o cálculo de Ψ_2 , são necessários os cálculos das n matrizes 2×2 das derivadas segundas:

$$\tilde{\tilde{S}}_l = \begin{pmatrix} s1^2 \exp(\gamma_1 s1 + \gamma_2 s2) & s1 s2 \exp(\gamma_1 s1 + \gamma_2 s2) \\ s1 s2 \exp(\gamma_1 s1 + \gamma_2 s2) & s2^2 \exp(\gamma_1 s1 + \gamma_2 s2) \end{pmatrix}_{2 \times 2}.$$

5.1.2 Resultados

As estimativas de MV e de $O(n^{-2})$, bem como os desvios padrão para as 10.000 amostras, com tamanhos 20, 30, 40 e 60, estão apresentados na Tabela 5.1. Observa-se que os desvios padrão das estimativas de $O(n^{-2})$ são menores quando comparados com os respectivos desvios padrão obtidos pela MV, e que $\hat{\beta}_3$ é praticamente o valor do parâmetro para todos os tamanhos da amostra. Tem-se, ainda, que o submodelo média está melhor estimado do que o de dispersão e, além disso, quanto maior o tamanho da amostra, as estimativas de MV e as de $O(n^{-2})$ tendem ao parâmetro.

Tabela 5.1: Estimativas de MV, de $O(n^{-2})$ e desvios padrão.

n	estimativas	parâmetros				
		$\beta_1 = 1$	$\beta_2 = 2$	$\beta_3 = 3$	$\gamma_1 = 1,5$	$\gamma_2 = 2,5$
20	de MV	1,006403	2,000003	3,001460	1,580471	2,584113
	desvio padrão	(0,110406)	(0,012272)	(0,097340)	(0,945028)	(0,945961)
	de $O(n^{-2})$	1,005261	2,000002	3,001463	1,492191	2,510794
	desvio padrão	(0,105140)	(0,012233)	(0,095937)	(0,752975)	(0,820877)
30	de MV	1,004672	1,999965	3,001460	1,550399	2,563342
	desvio padrão	(0,097658)	(0,011048)	(0,079667)	(0,795413)	(0,767448)
	de $O(n^{-2})$	1,004009	1,999961	3,001468	1,492046	2,507641
	desvio padrão	(0,093978)	(0,011029)	(0,078330)	(0,686448)	(0,649970)
40	de MV	1,003828	1,999996	3,000856	1,555018	2,531408
	desvio padrão	(0,085336)	(0,009716)	(0,068270)	(0,648638)	(0,629063)
	de $O(n^{-2})$	1,003361	1,999995	3,000866	1,518822	2,499339
	desvio padrão	(0,082407)	(0,009709)	(0,067325)	(0,558200)	(0,549016)
60	de MV	1,002441	1,999970	3,000795	1,536999	2,523925
	desvio padrão	(0,071417)	(0,008260)	(0,057046)	(0,469749)	(0,477901)
	de $O(n^{-2})$	1,002162	1,999969	3,000795	1,521429	2,507760
	desvio padrão	(0,069417)	(0,008248)	(0,056396)	(0,424031)	(0,400079)

Na Tabela 5.2, observa-se que o quociente do viés pelo seu respectivo desvio padrão diminui conforme cresce o tamanho da amostra. Assim, por esta Tabela, tem-se que a magnitude dos vieses é considerável para tamanho da amostra pequeno.

Tabela 5.2: Quociente do viés do parâmetro pelo seu respectivo desvio padrão.

tamanho da amostra	quociente				
	$\frac{ \text{viés } \beta_1 }{s\beta_1}$	$\frac{ \text{viés } \beta_2 }{s\beta_2}$	$\frac{ \text{viés } \beta_3 }{s\beta_3}$	$\frac{ \text{viés } \gamma_1 }{s\gamma_1}$	$\frac{ \text{viés } \gamma_2 }{s\gamma_2}$
20	1,10%	0,01%	0,02%	10,31%	8,63%
30	0,71%	0,06%	0,01%	7,85%	7,84%
40	0,56%	0,02%	0,02%	5,90%	5,41%
60	0,40%	0,01%	0,00%	3,47%	3,56%

Pela Figura 5.1, que representa o diagrama de caixas da estimativa do parâmetro β_1 , pelos dois métodos, observa-se que os resultados estão praticamente empatados.

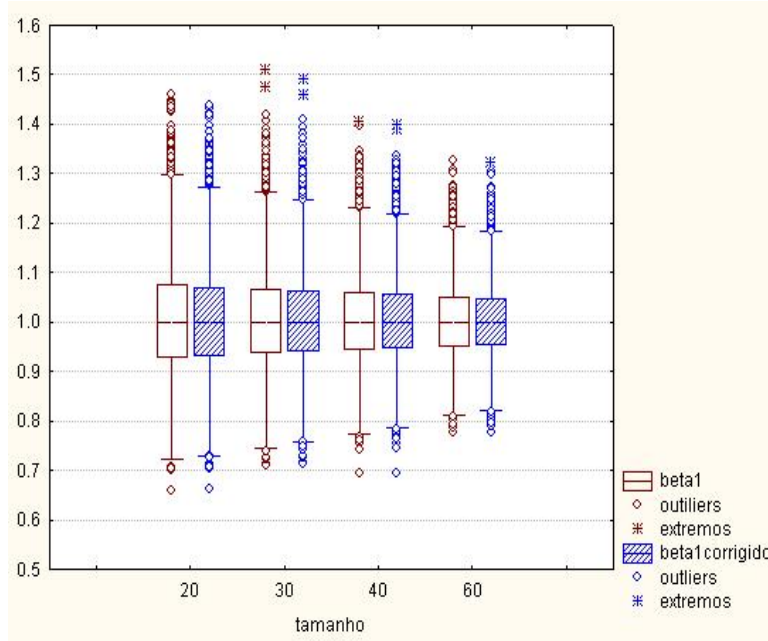
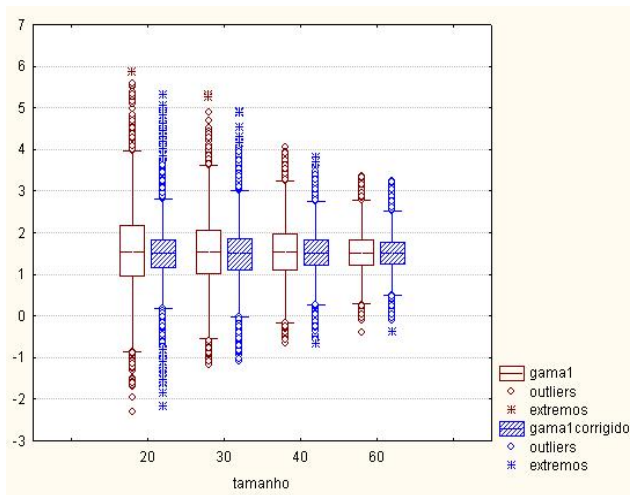
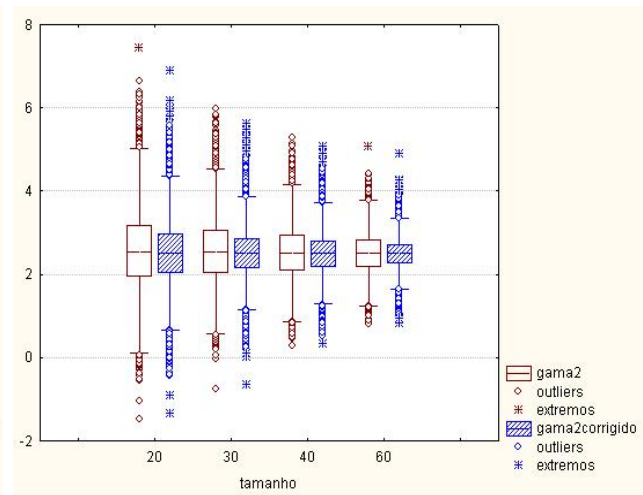


Figura 5.1: Diagrama de caixas da estimativa do parâmetro β_1 .

Considerando, agora, as estimativas dos parâmetros γ_1 e γ_2 , construído os diagramas de caixas para as suas estimativas, observa-se que as estimativas de $O(n^{-2})$ são mais precisas.



(a)



(b)

Figura 5.2: (a): Diagrama de caixas da estimativa do parâmetro γ_1 ; (b): Diagrama de caixas da estimativa do parâmetro γ_2 .

Os resultados foram os esperados, uma vez que, para amostras de tamanho pequeno e moderado, o viés é considerável, enquanto que, para amostras grandes, o viés é praticamente nulo.

O programa deste estudo de simulação encontra-se no Apêndice F.

5.2 Simulação da variável resposta com distribuição gama

A segunda simulação realizada foi a distribuição gama, que, assim como a inversa gaussiana, é muito utilizada na prática. Também, para esta, o uso do viés foi estudado através da simulação das amostras via Monte Carlo.

Como no aplicativo SAS a simulação da distribuição gama é feita considerando apenas um parâmetro, dado por:

$$p(x; \phi) = \frac{1}{\Gamma(\phi)} x^{\phi-1} \exp(-x), \quad (5.2)$$

em que ϕ é o parâmetro *shape* e $E(X) = V(X) = \phi$, porém, nesta Tese é necessária a distribuição gama de dois parâmetros, que é dado por:

$$p(y; \mu, \phi) = \frac{1}{y\Gamma(\phi)} \left(\frac{\phi y}{\mu}\right)^{\phi} \exp\left(-\frac{\phi y}{\mu}\right), \quad (5.3)$$

sendo $E(X) = \mu$, $V(X) = \frac{\mu^2}{\phi}$ em que μ é a média. Para passar da distribuição (5.2) para (5.3) aplica-se o teorema da transformação da variável aleatória, considerando os parâmetros de escala igual a $\frac{\mu}{\phi}$, *shape* igual a ϕ e a parametrização:

$$y = \frac{\mu}{\phi} x.$$

As matrizes X e S foram geradas a partir da distribuição uniforme $U(0, 1)$, sendo fixadas para as 10.000 simulações para amostras de mesmo tamanho. Também, utilizou-se a restrição de que os valores de ϕ deveriam ser todas maiores que 1, pois, caso isso não ocorra, o processo de estimação apresenta instabilidade.

5.2.1 Submodelos média e dispersão

As componentes sistemáticas, tanto para a média como para a dispersão, foram as mesmas da distribuição inversa gaussiana.

Os valores dos parâmetros dados foram para: $\beta_1 = 1$, $\beta_2 = 2$, $\beta_3 = 3$, $\gamma_1 = 1,5$ e $\gamma_2 = 2,5$.

Inicialmente, foram geradas as matrizes X e S, consideradas fixadas e, em seguida, a variável resposta Y. Após, a geração da variável resposta Y, para cada amostra, pelo *procedure NLP*, foram obtidas as estimativas dos parâmetros pelo método de máxima verossimilhança: β_1 , β_2 e β_3 e, alternadamente, γ_1 e γ_2 , uma vez que β e γ são globalmente ortogonais, por (4.2). E, posteriormente, para se obter as estimativas de $O(n^{-2})$, é necessário calcular os vieses de β , por (4.6) e γ , por (4.11), pelo *procedure IML*.

Para o cálculo dos vieses, são necessárias as matrizes das primeira e segunda derivadas, obtidas pelo aplicativo matemático *Maple*.

a) Para o submodelo média, a matriz \tilde{X} e \tilde{X}_l são iguais ao caso da distribuição inversa gaussiana, no qual a única diferente é a matriz peso, que é dada por:

$$\phi W = \text{diag} \left(\phi \frac{1}{\mu^2} \right).$$

Em seguida, todos os passos são iguais aos descritos para a distribuição inversa gaussiana.

b) Para o submodelo de dispersão, as matrizes \tilde{S} e \tilde{S}_l são iguais ao caso da inversa gaussiana, no qual a diferença é a matriz peso dada por:

$$-D_2 \phi_1^2 = \text{diag} \left(\frac{1}{\phi} - \text{digamma}(\phi) \right).$$

Todos os passos são os mesmos daquelas descritos para a distribuição inversa gaussiana.

5.2.2 Resultados

Para cada tamanho de amostra 20, 30, 40 e 60, realizaram-se 10.000 simulações, cujos resultados estão apresentados na Tabela 5.3.

Tabela 5.3: Estimativas de MV, de $O(n^{-2})$ e desvios padrão.

n	<i>estimativas</i>	<i>parâmetros</i>				
		$\beta_1 = 1$	$\beta_2 = 2$	$\beta_3 = 3$	$\gamma_1 = 1,5$	$\gamma_2 = 2,5$
20	<i>de MV</i>	1,051734	2,005798	3,006257	1,495976	2,680128
	desvio padrão	(0, 255802)	(0, 031210)	(0, 356724)	(0, 900227)	(0, 853378)
	<i>de $O(n^{-2})$</i>	1,027184	2,005114	3,005271	1,495976	2,679933
	desvio padrão	(0, 197087)	(0, 028896)	(0, 357752)	(0, 900162)	(0, 849476)
30	<i>de MV</i>	1,026738	2,002912	3,007338	1,532217	2,577719
	desvio padrão	(0, 189736)	(0, 025147)	(0, 258882)	(0, 675256)	(0, 649081)
	<i>de $O(n^{-2})$</i>	1,017699	2,002728	3,006740	1,532230	2,577823
	desvio padrão	(0, 165064)	(0, 022938)	(0, 259201)	(0, 675229)	(0, 647034)
40	<i>de MV</i>	1,017300	2,001662	3,004823	1,522720	2,555748
	desvio padrão	(0, 151543)	(0, 021089)	(0, 210565)	(0, 596970)	(0, 574742)
	<i>de $O(n^{-2})$</i>	1,012602	2,001580	3,004629	1,522732	2,555781
	desvio padrão	(0, 136374)	(0, 019603)	(0, 211415)	(0, 596938)	(0, 573347)
60	<i>de MV</i>	1,010848	2,001105	3,006253	1,526725	2,525774
	desvio padrão	(0, 130747)	(0, 017461)	(0, 172457)	(0, 478318)	(0, 446653)
	<i>de $O(n^{-2})$</i>	1,008302	2,001060	3,006111	1,526729	2,525826
	desvio padrão	(0, 120840)	(0, 016553)	(0, 173152)	(0, 478309)	(0, 445773)

Na Tabela 5.4, é apresentado o quociente do viés pelo seu respectivo desvio padrão, cujos resultados confirmam a teoria, isto é, quanto maior o tamanho da amostra menor ela é.

Tabela 5.4: Quociente do viés do parâmetro pelo seu respectivo desvio padrão.

tamanho da amostra	quociente				
	$\frac{ \text{viés } \beta_1 }{s\beta_1}$	$\frac{ \text{viés } \beta_2 }{s\beta_2}$	$\frac{ \text{viés } \beta_3 }{s\beta_3}$	$\frac{ \text{viés } \gamma_1 }{s\gamma_1}$	$\frac{ \text{viés } \gamma_2 }{s\gamma_2}$
20	12,46%	2,39%	0,28%	0,01%	0,16%
30	5,48%	0,78%	0,23%	0,00%	0,02%
40	3,45%	0,41%	0,09%	0,00%	0,01%
60	2,11%	0,01%	0,08%	0,00%	0,01%

Portanto, por este estudo de simulação, tanto para a distribuição inversa gaussiana como a gama, conclui-se que os vieses são necessários para amostras de tamanho pequeno e que, para tamanhos de amostras grandes, as estimativas de segunda ordem são praticamente iguais às de MV, concordando com a teoria, isto é, conforme o tamanho da amostra vai aumentando, as estimativas tendem ao verdadeiro parâmetro e uma observação importante é que as estimativas de $O(n^{-2})$ apresentam menor desvio padrão, ou seja, são mais precisas.

Capítulo 6

Aplicação em modelos *one-compartment*

Estes modelos são usualmente adotados pela indústria farmacêutica na Fase I de experimentos em farmacocinética, cujo interesse é conhecer o comportamento geral de uma nova droga.

Os modelos desse tipo consideram que, após a administração da droga, a mesma se distribui em todo o organismo instantaneamente. Supondo um bloco homogêneo para a droga, é como se ela estivesse sendo dissolvida em um béquer contendo um único solvente. A Figura 6.1 esquematiza a relação entre a concentração da droga no plasma e o tempo após a aplicação, que pode ser descrito como a soma de parcelas exponenciais.



Figura 6.1: Modelos *one-compartment*.

Lindsey *et al.* (2000) analisam a concentração do *flosequinan* no plasma, que é uma droga para o tratamento de insuficiência cardíaca, ressaltando que:

- na Fase I, a amostra é constituída de voluntários sadios, e isto implica para as indústrias farmacêuticas um custo muito alto quando o tamanho da amostra é grande;
- geralmente a transformação logarítmica é aplicada na variável resposta, induzindo uma implícita relação entre a média e a variância;

- ajustaram três modelos: com variância constante, para a média e para dispersão e por último, os mistos considerando os voluntários sadios como medidas repetidas, e isto, pode provocar superdispersão, conforme Dobson (2002, p. 167), uma vez que, as observações dentre os indivíduos estão correlacionadas. Os resultados evidenciaram que os modelos mistos, modelos para média e para dispersão quando comparados aos modelos com variância constante, são melhores.

Diante do exposto, justifica-se o uso deste experimento, uma vez que, o intuito é somente ilustrar a viabilidade prática da metodologia proposta.

6.1 Modelo

Os dados analisados estão disponibilizados no *site* <http://www.popgen.unimaas.nl/~jlindsey>, com autorização deste, e referem-se à Fase I do estudo da droga *flosequinan*.

6.1.1 Variáveis de interesse

Y_l → resposta: concentração do *flosequinan*, em microg/ml;

t_l → covariável: tempo, em h e

d_l → covariável: dose, em mg.

6.1.2 Delineamento experimental

O delineamento experimental é do tipo medidas repetidas, isto é, as amostras de sangue foram coletadas a 0; 0,5; 1; 1,5; 2; 3; 4; 6; 8; 10; 12; 24 horas, em cada voluntário sadio.

As doses da droga usadas foram 50 mg, 100 mg e 150 mg.

A amostra consistiu de 18 voluntários sadios, com idade média de 34,3 anos, com desvio padrão de 10,3 anos, peso médio de 63,6 kg e desvio padrão de 10,2 kg.

6.1.3 Componente aleatória

Utilizou-se para variável resposta as distribuições gama, inversa gaussiana e normal, que serão denotadas com subscrito G , IG e N , respectivamente, isto é,

$$Y_{Gl} \sim G(\mu_l, \phi_l), \quad Y_{IGl} \sim IG(\mu_l, \phi_l) \quad \text{e} \quad Y_{Nl} \sim N(\mu_l, \phi_l) \quad \text{sendo } l = 1, \dots, 648.$$

6.1.4 Componentes sistemáticas:

Submodelo média

A função que descreve a média dos modelos *one-compartment* apresentam três parâmetros k_a , k_e e v , no entanto, nesta análise, será considerado que $k_a = k_e$, pois, caso contrário, a função log-verossimilhança da concentração será bimodal, dificultando a obtenção das estimativas; portanto, o preditor não-linear η_l é dado por:

$$\eta_l = \mu_l = \frac{k_{am}d_l t_l}{v_m} (\exp(-k_{am}t_l)), \quad (6.1)$$

sendo:

- k_a interpretado como razão de absorção;
- k_e interpretado como razão de eliminação da droga;
- v interpretado como volume que converte a dose total do compartimento em concentração;
- índice m colocados como subscrito, denota parâmetros da média.

Submodelo dispersão

$$\tau_l = \phi_l = \frac{\exp(-k_{ad}t_l)}{\nu_d}, \quad (6.2)$$

sendo:

- índice d colocado como subscrito denota o parâmetro dispersão.

Aqui também considerou-se $k_a = k_e$.

É importante salientar que neste modelo as estimativas dos parâmetros obtidos devem ser:

1. positivas, uma vez que a razão de absorção, razão de eliminação e volume são valores positivos
2. $\hat{v} > \hat{k}_{am}$ e $\hat{v} > \hat{k}_{em}$.

Observações:

- As variáveis aleatórias com distribuições gama ou inversa gaussiana assumem valores positivos, então, para adotá-las, adicionou-se uma constante à concentração do *flosequinan*

no plasma, uma vez que neste banco de dados há várias observações iguais a zero, isto é,

$$Y = \text{concentração} + 10; \quad (6.3)$$

e o submodelo de dispersão para distribuições:

- gama: $\phi_{Gl} = \phi_l \left(\frac{1}{\mu_l^2} \right) (\mu_l + 10)^2 = \frac{\exp(-k_{ad}t_l)}{\nu_d} \left(\frac{1}{\mu_l^2} \right) (\mu_l + 10)^2$;
- inversa gaussiana: $\phi_{IGl} = \phi_l \left(\frac{1}{\mu_l^3} \right) (\mu_l + 10)^3 = \frac{\exp(-k_{ad}t_l)}{\nu_d} \left(\frac{1}{\mu_l^3} \right) (\mu_l + 10)^3$.

6.2 Algoritmo do SAS

Passo 1) As estimativas de MV foram obtidas pela *procedure* NLMIXED. Nesta *procedure*, deve-se fornecer a função log-verossimilhança, os submodelos média e dispersão e a tentativa inicial.

Para obter a tentativa inicial, construiu-se gráficos com vários valores de k_{am} ou v_m ou k_{ad} ou v_d no eixo X, e a função log-verossimilhança no eixo Y. Um desses gráficos é apresentado a seguir:

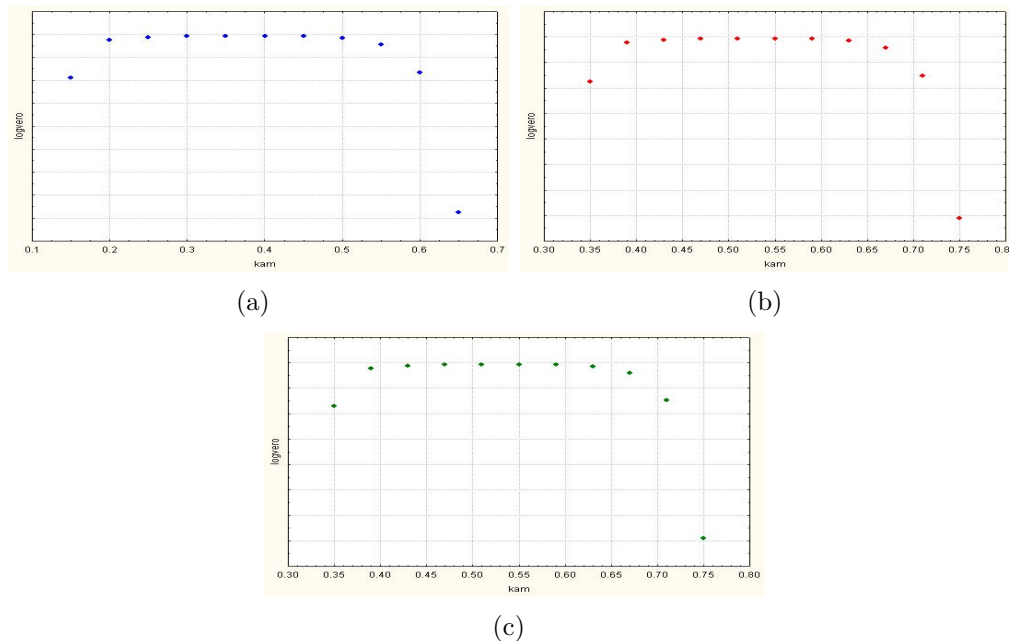


Figura 6.2: (a): Comportamento da log-verossimilhança quando k_{am} é fixado para dose 50; (b): Comportamento da log-verossimilhança quando k_{am} é fixado para dose 100; (c): Comportamento da log-verossimilhança quando k_{am} é fixado para dose 150.

Passo 2) As estimativas de $O(n^{-2})$ foram obtidas pela *procedure* IML e, para isto, é necessária a obtenção das seguintes matrizes:

matriz das primeiras derivadas

$$\tilde{X}^T = \begin{bmatrix} \frac{dt \exp(-k_{am}t) - k_{am} dt^2 \exp(-k_{am}t)}{v_m} & \frac{k_{am} dt \exp(-k_{am}t)}{v_m^2} \end{bmatrix}_{n \times 2}^T.$$

matrizes peso do submodelo média

- $\Phi W_G = \text{diag} \left(\phi_l \frac{1}{\mu_l^2} \right)_{n \times n}$,
- $\Phi W_{IG} = \text{diag} \left(\phi \frac{1}{\mu^3} \right)_{n \times n}$ e
- $\Phi W_N = \text{diag} (\phi \mathbf{1})_{n \times n}$.

matriz das segundas derivadas, \tilde{X}_l para $l = 1, 2, \dots, 648$,

$$\tilde{X}_l = \begin{bmatrix} -\frac{2dt^2 \exp(-k_{am}t)}{v_m} + \frac{k_{am} dt^3 \exp(-k_{am}t)}{v_m} & -\frac{dt \exp(-k_{am}t)}{v_m^2} + \frac{k_{am} dt^2 \exp(-k_{am}t)}{v_m^2} \\ -\frac{dt \exp(-k_{am}t)}{v_m^2} + \frac{k_{am} dt^2 \exp(-k_{am}t)}{v_m^2} & \frac{2k_{am} dt \exp(-k_{am}t)}{v_m^3} \end{bmatrix}_{2 \times 2}.$$

matriz F é nula.

matriz das primeiras derivadas do submodelo dispersão

$$\tilde{S}^T = \begin{bmatrix} -\frac{t \exp(-k_{ad}t)}{v_d} & -\frac{\exp(-k_{ad}t)}{v_d^2} \end{bmatrix}_{n \times 2}^T.$$

matrizes peso do submodelo dispersão

- $(-D_2 \phi_1^2)_G = \text{diag} \left\{ -\left(\frac{1}{\phi} - \text{digamma}(\phi) \right) \right\}_{n \times n}$,
- $(-D_2 \phi_1^2)_{IG} = \text{diag} \left\{ -\left(-\frac{1}{2\phi^2} \right) \right\}_{n \times n}$ e
- $(-D_2 \phi_1^2)_N = \text{diag} \left\{ -\left(-\frac{1}{2\phi^2} \right) \right\}_{n \times n}$.

matriz das segundas derivadas, \tilde{S}_l para $l = 1, 2, \dots, 648$,

$$\tilde{S}_l = \begin{bmatrix} \frac{t^2 \exp(-k_{ad}t)}{v_d} & \frac{t \exp(-k_{ad}t)}{v_d^2} \\ \frac{t \exp(-k_{ad}t)}{v_d^2} & \frac{2 \exp(-k_{ad}t)}{v_d^2} \end{bmatrix}.$$

matriz F é uma matriz nula.

6.3 Resultados

O diagrama de dispersão, apresentado na Figura 6.3, mostra que o comportamento da média decresce exponencialmente. Portanto, a função que descreve a média é razoável.

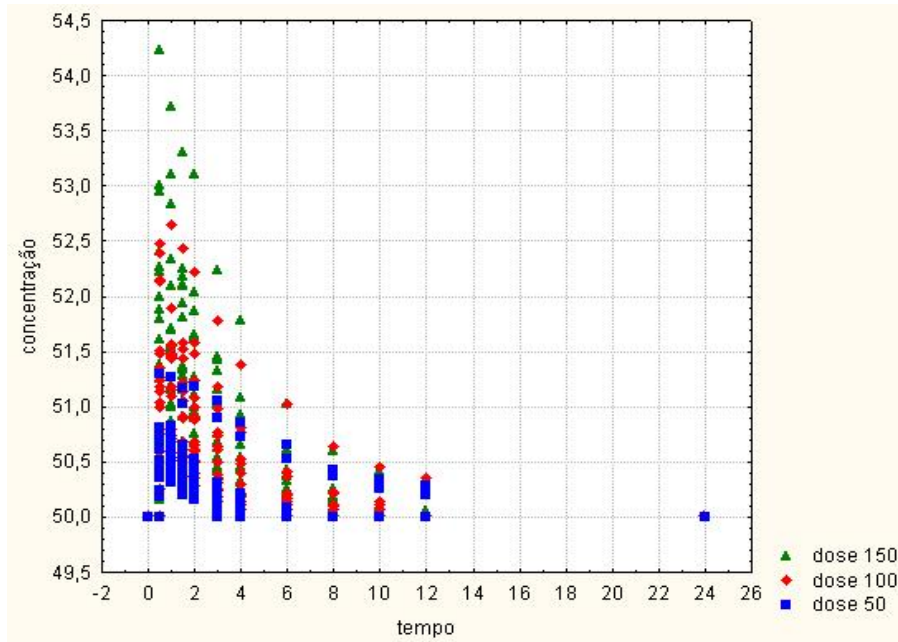


Figura 6.3: Diagrama de dispersão da concentração versus tempo.

As Tabelas 6.1, 6.2 e 6.3 apresentam as estimativas obtidas para as componentes aleatórias gama, inversa gaussiana e normal.

Pode-se observar que o modelo normal é o que tem menor AIC, porém, neste, \hat{k}_{ad} é negativo. Logo, este modelo é descartado. Portanto, utilizando o critério AIC, o modelo a ser adotado deve ser inversa gaussiana.

Tabela 6.1: Estimativa de MV, de $O(n^{-2})$, erros padrão para modelo gama.

Parâmetro	Estimativa de MV	Erro Padrão	p-nível	Estimativa de $O(n^{-2})$
k_{am}	0,4684	0,0078	<0,0001	0,4066
v_m	6,4839	0,2462	<0,0001	6,0756
k_{ad}	0,8305	0,0191	<0,0001	0,7997
v_d	5,6587	0,6804	<0,0001	5,9077
AIC	3621,3105			

Tabela 6.2: Estimativas de MV, de $O(n^{-2})$, erros padrão para modelo inversa gaussiana.

Parâmetro	Estimativa de MV	Erro Padrão	p-nível	Estimativa de $O(n^{-2})$
k_{am}	0,5528	0,0035	<0,0001	0.5484
v_m	84,4798	2,0047	<0,0001	83,5095
k_{ad}	0,4782	0,0101	<0,0001	0,4788
v_d	11,1950	0,9534	<0,0001	11,1771
AIC	3369,8256			

Tabela 6.3: Estimativas de MV, de $O(n^{-2})$, erros padrão para modelo normal.

Parâmetro	Estimativa de MV	Erro Padrão	p-nível	Estimativa de $O(n^{-2})$
k_{am}	0,9075	0,4349	<0,0001	0,9068
v_m	30,9131	8,7391	<0,0001	30,8909
k_{ad}	-0,4601	0,0134	<0,0001	-0,4599
v_d	0,8316	0,0850	<0,0001	0,8336
AIC	1011,2265			

Para avaliar os vieses, o quociente entre estes e seus respectivos erros padrão, para o modelo adotado, estão apresentados na Tabela (6.4). Observa-se que para o submodelo média, a variação é grande; o mesmo não ocorre para o submodelo de dispersão. Portanto, os vieses devem ser considerados neste caso.

Tabela 6.4: Porcentagem do quociente entre o viés e seu respectivo erro padrão para resposta inversa gaussiana.

Quociente	Modelo Inversa Gaussiana
$\frac{visk_{am}}{sk_{am}}$	128,2
$\frac{visv_m}{sv_m}$	48,3
$\frac{visk_{ad}}{sk_{ad}}$	6,4
$\frac{visv_d}{sv_d}$	1,8

Conclui-se que, para estes dados, adotando a componente aleatória inversa gaussiana, os estimadores de $O(n^{-2})$ parecem relevantes.

Capítulo 7

Considerações finais

Como a área de atuação de um Engenheiro de Produção é a gerência dos sistemas de produção, nas indústrias manufatureiras, nas instituições financeiras, nas instituições de saúde, entre outras e, como compete a ele prever e avaliar os resultados destes sistemas, é importante, portanto, a tomada de decisão em tempo real, o que induz amostras de tamanho pequeno. E, nesses casos, o estudo de simulação das distribuições inversa gaussiana e gama evidenciou que o uso dos estimadores de $O(n^{-2})$ é relevante. A obtenção dos vieses, aqui, foi calculada na *procedure* IML, com o objetivo de mostrar que pode ser facilmente implementada, pois é solução de um sistema de equações de uma regressão de mínimos quadrados ponderados. A maior dificuldade para isto é resolver as derivadas de primeira e segunda ordem dos submodelos média e dispersão, mas isto pode ser sanado por qualquer aplicativo matemático, como o *Maple*.

Ainda, enfatizou-se o uso não só do submodelo média, mas também do submodelo dispersão para uma função densidade de probabilidade pertencente à família exponencial de dispersão linear, já que a meta na Engenharia de Qualidade é desvio nulo com relação à média. Contudo, isto não ocorre na prática, e é necessário modelar a dispersão para controlar a variabilidade de um dado produto, de tal forma que ela seja a mais reduzida possível.

Desse modo, recomendam-se:

- o uso dos estimadores de $O(n^{-2})$ sempre que o tamanho da amostra for pequeno, pois é de fácil implementação e sua magnitude pode ser considerável nestes casos;
- o estudo da distribuição amostral dos estimadores de $O(n^{-2})$, uma vez que, nesta Tese, foi obtida a expressão e avaliada através de um estudo de simulação.

Todos os programas utilizados nesta Tese serão disponibilizados no site

www.uem.br/des~mctudo.

Referências Bibliográficas

- [1] BARBETTA, P.A. (1998). **Construção de Modelos para Médias e Variâncias na Otimização Experimental de Produtos e Processos**. Tese (Doutorado)-Programa de Pós-Graduação em Engenharia de Produção - UFSC, Florianópolis.
- [2] BARBETTA, P.A., RIBEIRO, J.L.D. e BORNIA, A.C. (1999). **Construção de Modelo para Variância em Estudos Experimentais**. Produto & Produção, **2**, 56-65.
- [3] BARTLETT, M.S. (1953). **Approximate Confidence Intervals II**. Biometrika, **40**, 306-317.
- [4] 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 - USP, São Paulo.
- [5] 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.
- [6] BOX, M.J. (1971). **Bias in Nonlinear Estimation**. Journal of the Royal Statistical Society B, **33**, 171-201.
- [7] CHHIKARA, R.S. e FOLKS, J.L. (1989). **The Inverse Gaussian Distribution**. Marcell Dekker. New York.
- [8] COOK, R.D., TSAI, C.L. e WEI, B.C. (1986). **Bias in Nonlinear Regression**. Biometrika, **73**, 615-623.
- [9] COOK, R.D. e TSAI, C.L. (1990). **Diagnostic for Assessing the Accuracy of Normal Approximations in Exponential Family Nonlinear Models**. Journal of the American Statistical Association, **85**, 770-777.
- [10] CORDEIRO, G.M. (1983). **Improved Likelihood Ratio Statistics for Generalized Linear Models**. Journal of the Royal Statistical Society B, **45**, 404-413.
- [11] CORDEIRO, G.M. (1986). **Modelos Lineares Generalizados**. VII Simpósio Nacional de Probabilidade e Estatística, UNICAMP, Campinas, São Paulo.

- [12] CORDEIRO, G.M. (1987). **On the Correction to the Likelihood Ratio Statistics.** *Biometrika*, **74**, 265-274.
- [13] CORDEIRO, G.M. (1999). **Introdução à Teoria Assintótica.** 22^o Colóquio Brasileiro de Matemática, IMPA, Rio de Janeiro.
- [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 BOTTER, D.A. (2001). **Second-Order Biases of Maximum Likelihood Estimates in Overdispersed Generalized Linear Models.** *Statistics and Probability Letters*, **55**, 269-280.
- [16] CORDEIRO, G.M. e KLEIN, R. (1994). **Bias Correction in ARMA Models.** *Statistics and Probability Letters*, **19**, 169-176.
- [17] CORDEIRO, G.M. e McCULLAGH, P. (1991). **Bias Correction in Generalized Linear Models.** *Journal of the Royal Statistical Society B*, **53**, 629-643.
- [18] CORDEIRO, G.M. e PAULA, G.A. (1989). **Improved Likelihood Ratio Statistics for Exponential Family Nonlinear Models.** *Biometrika*, **76**, 93-100.
- [19] CORDEIRO, G.M. e PAULA, G.A. (1989 *b*). **Fitting Non-Exponential Family Nonlinear Models in GLIM by Using the Offset Facility.** *Lecture Notes in Statistics*, **57**, 105-114.
- [20] 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.
- [21] CORDEIRO, G.M. e VASCONCELLOS, K.L.P. (1999). **Second-Order Biases of the Maximum Likelihood Estimates in von Mises Regression Models.** *Australian and New Zealand Journal of Statistics*, **41**, 189-198.
- [22] 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.
- [23] DAVIDIAN, M. e GILTINAN, D.M. (1995). **Nonlinear Models for Repetead Measurement Data.** 1^a Edição. Chapman & Hall. London.
- [24] DOBSON, A.J. (2002). **An Introduction to Generalized Linear Models.** 2^a Edição. Chapman & Hall. United States of America.
- [25] FISHER, R. A. (1912). **On an Absolute Criterion for Fitting Frequency Curves.** *Messenger of Mathematics* **41**, 155 - 160.

- [26] FORSBERG, L. (2002). **On the Normal Inverse Gaussian Distribution in Modeling Volatility in the Financial Markets**. Uppsala University Press.
- [27] HAWKINS, D. e OLWELL, D. (1997). **Inverse Gaussian Cumulative Sum Control Charts for Location and Shape**. Journal of the Royal Statistical Society D, **46**, 3 - 323.
- [28] JØRGENSEN, B. (1983). **Maximum Likelihood Estimation and Large-Simple Inference for Generalized Linear and Nonlinear Regression Models**. Biometrika, **70**, 19-28.
- [29] JØRGENSEN, B. (1987). **Small Dispersion Asymptotics**. Revista Brasileira de Probabilidade e Estatística, **1**, 59-90.
- [30] JØRGENSEN, B. (1997). **The Theory of Dispersion Models**. 1ª Edição. Chapman & Hall. Great Britain.
- [31] LINDSEY, J. K., BYROM, W. D., WANG, J., JARVIS, P. e JONES, B. (2000). **Generalized Nonlinear Models for Pharmacokinetics Data**. Biometrics, **56**, 81 - 88.
- [32] MATTOS, V.L.D. (2004). **Identificação de Efeitos de Dispersão em Experimentos com Poucas Replicações**. Tese (Doutorado) - Programa de Pós-Graduação em Engenharia de Produção - UFSC, Florianópolis.
- [33] McCULLAGH, P. e NELDER, J. A. (1983). **Generalized Linear Models**. 1ª Edição. Chapman & Hall. London.
- [34] McCULLAGH, P. e NELDER, J. A. (1989). **Generalized Linear Models**. 2ª Edição. Chapman & Hall. London.
- [35] MONTGOMERY, D.C. (2000). **Introduction to Statistical Quality**. John Wiley. New York.
- [36] NELDER, J.A. e WEDDERBURN, R.W.M. (1972). **Generalized Linear Models**. Journal of the Royal Statistical Society A, **135**, 370-384.
- [37] 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” - USP, São Paulo.
- [38] 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.
- [39] PAULA, G.A. (1991). **Bias Correction for Exponential Family Nonlinear Models**. Journal of Statistical Computation and Simulation, **40**, 43-54.

- [40] PAULA, G.A. (2001). **Modelos de Regressão com Apoio Computacional**. Apostila. IME/USP. www.ime.usp/~giapaula.
- [41] PIKE, M.C., HILL A.P. e SMITH P.G. (1980). **Bias and Efficiency in Logistic Analyses of Stratified Case-Control Studies**. *International Journal of Epidemiologie*, **9**, 89-95.
- [42] PINHEIRO, J.C. e BATES, D.M. (2000). **Mixed-Effects Models in S and S-PLUS**. Springer-Verlag. New York.
- [43] PREGIBON, D. (1984). **Review of Generalized Linear Models**. *Annals of Statistics*, **12**, 1589-1596.
- [44] RATKOWSKY, D. A. (1983). **Nonlinear Regression Analisis**. Marcel Decker Inc. New York.
- [45] SMYTH, G.K. (1989). **Generalized Linear Models With Varying Dispersion**. *Journal of the Royal Statistical Society B*, **51**, 47-60.
- [46] SMYTH, G.K. e VERBYLA, A.P. (1999). **Adjusted Likelihood Methods for Modelling Dispersion in Generalized Linear Models**. *Evironmetrics*, **10**, 696-709.
- [47] SOWDEN, R.R. (1971). **Bias and Accuracy of Parameter Estimates in a Quantal Response Model**. *Biometrika*, **58**, 595-603.
- [48] WEI, B.C. (1998). **Exponential Family Nonlinear Models**. Springer-Verlag Singapore Pte. Ltd. Singapore.
- [49] WEISBERG, S. (1985). **Applied Linear Regression**. 1^a Edição. John Wiley. New York.
- [50] WHITEMORE, G.A. (1986). **Inverse Gaussian Ratio Estimation**. *Applied Statistics*, **35**, 8-15.
- [51] WOLFINGER, R.D. (2004). **Fitting Nonlinear Mixed Models With the New NLMIXED Procedure**. SAS Institute Inc., Cary, NC.
- [52] YOUNG, D.H. e BAKIR, S.T. (1987). **Bias Correction for a Generalized Log-gama Regression Models**. *Tecnometrics*. **29**, 183-191.

Apêndice A

Identidades de Bartlett

Para a obtenção dessas identidades são necessárias as condições de regularidade definidas a seguir.

Condições de regularidade

Considere Y o valor observado de uma variável aleatória $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)^T$ caracterizada por uma função de densidade dada por $f(y; \theta)$ conhecida, porém dependente de um vetor de parâmetros conhecidos $\theta = (\theta_1, \theta_2, \dots, \theta_p)^T$. Seja $\Theta \subset R^p$ o espaço paramétrico representando o conjunto de valores possíveis para o vetor θ . Assim, para se ter regularidade devem ser satisfeitas as seguintes condições:

- a) Θ é fechado, compacto e tem dimensão finita onde o parâmetro verdadeiro θ_0 é um ponto interior de Θ ;
- b) $f(y; \theta)$, a função densidade, é uma função biunívoca de θ_i ;
- c) $\ell(\theta)$, a função log-verossimilhança do parâmetro θ , apresenta derivadas de no mínimo até a terceira ordem na vizinhança de θ_0 ;
- d) K_θ , a matriz de informação de Fisher, é finita e positiva definida na vizinhança de θ_0 ;
- e) para modelos contínuos a igualdade $\frac{\partial}{\partial \theta} E \{t(Y)\} = \int t(y) \frac{\partial p(y; \theta)}{\partial \theta} dy$ é válida para qualquer estatística $t(y)$. Para modelos discretos basta substituir a integral pelo somatório.

Identidades de Bartlett

Seja $L = L(\theta)$ a verossimilhança total de um problema regular e adotado e a seguinte notação para as derivadas da função log-verossimilhança $\ell = \ell(\theta) = \log L(\theta)$,

$$U_r = \frac{\partial \ell}{\partial \theta_r}, \quad U_{rs} = \frac{\partial^2 \ell}{\partial \theta_r \partial \theta_s}, \quad U_{r,s} = \frac{\partial \ell}{\partial \theta_r} \frac{\partial \ell}{\partial \theta_s} \text{ etc.}$$

onde todos índices variam de 1 a p , sendo p o número de parâmetros.

Os momentos conjuntos de derivadas de $\ell(\theta)$ são

$$\mu_r = E(U_r) \ , \ \mu_{rs} = E(U_{rs}) \ , \ \mu_{r,s} = E(U_r U_s) \ , \ \mu_{r,st} = E(U_r U_{st})$$

e assim por diante.

Como $\mu_r = 0$ os correspondentes cumulantes conjuntos (κ' s) expresso em termos dos momentos são:

$$\begin{aligned} \kappa_r &= \mu_r, & \kappa_{r,s} &= \mu_{r,s}, & \kappa_{rs} &= \mu_{rs}, \\ \kappa_{rs,t} &= \mu_{rs,t}, & \kappa_{r,s,t} &= \mu_{r,s,t}, \\ \kappa_{rs,tu} &= \mu_{rs,tu} - \mu_{rs}\mu_{tu} \quad \text{e} \quad \kappa_{r,s,t,u} = \mu_{r,s,t,u} - \sum_{(3)} \mu_{r,s}\mu_{t,u}, \end{aligned}$$

onde $\sum_{(k)}$ representa o somatório sobre todas as k combinações de índices.

Os momentos e cumulantes definidos não são independentes, mas satisfazem a certas equações que facilitam seus cálculos. Estas equações são denominadas de identidades de Bartlett e para a sua dedução utiliza-se a propriedade:

$$\frac{\partial}{\partial \theta} E \{t(Y)\} = \int t(y) \frac{\partial p(y; \theta)}{\partial \theta} dy \quad (\text{A.1})$$

que é válida em problemas regulares.

Se

$$\kappa_r = 0,$$

para sua dedução tem-se que $U_r = \frac{L_r}{L}$ onde $L_r = \frac{\partial L}{\partial \theta_r}$. Por outro lado, $\int L dy = 1$. Diferenciando com relação a θ_r esta igualdade tem-se $\int \frac{\partial L}{\partial \theta_r} dy = 0$. Logo, $\int L_r dy = 0$, mas como $L_r = U_r L$, tem-se que $\int U_r L dy = 0$. Conclui-se, então, que $E(U_r) = 0$, logo $\kappa_r = 0$.

Ou ainda, se

$$\kappa_{rs} + \kappa_{r,s} = 0,$$

para sua dedução basta diferenciar $\int L_r dy = 0$ com relação a θ_s obtendo-se $\int \frac{\partial}{\partial \theta_s} L_r dy = 0$, substituindo L_r por $U_r L$ tem-se que $\int \frac{\partial}{\partial \theta_s} (U_r L) dy = \int \left\{ \left(\frac{\partial U_r}{\partial \theta_s} \right) L + \left(\frac{\partial L}{\partial \theta_s} \right) U_r \right\} dy = 0$, assim $\int (U_{rs} L + U_s U_r) dy = 0$, ou seja, $\kappa_{rs} + \kappa_{r,s} = 0$ ou ainda, $\kappa_{r,s} = -\kappa_{rs}$.

A seguir, mais identidades de Bartlett, denotando-se as derivadas dos cumulantes com sobrescritos, isto é,

$$\kappa_{rs}^{(t)} = \frac{\partial \kappa_{rs}}{\partial \theta_t}, \quad \kappa_{rs}^{(tu)} = \frac{\partial^2 \kappa_{rs}}{\partial \theta_t \partial \theta_u}, \quad \kappa_{rst}^{(u)} = \frac{\partial \kappa_{rst}}{\partial \theta_u}, \quad \text{etc...}$$

e estas podem ser deduzidas de forma análoga aos anteriores: $\kappa_{r,s,t} + \kappa_{rst} + \sum_{(3)} \kappa_{r,st} = 0$, $\kappa_{r,st} + \kappa_{rst} - \kappa_{st}^{(r)} = 0$, $\kappa_{r,s,t} - 2\kappa_{rst} + \sum_{(3)} \kappa_{rs}^{(t)} = 0$, $\kappa_{rst}^{(u)} = \kappa_{rst,u}$, $\kappa_{r,stu} + \kappa_{rstu} - \kappa_{stu}^{(r)} = 0$,

$$\kappa_{rstu} + \sum_{(4)} \kappa_{r,stu} + \sum_{(3)} \kappa_{rs,tu} + \sum_{(6)} \kappa_{r,s,tu} + \kappa_{r,s,t,u} = 0, \quad \kappa_{r,s,t,u} = -3\kappa_{rstu} + 2 \sum_{(4)} \kappa_{rst}^{(u)} - \sum_{(6)} \kappa_{rs}^{(tu)} + \sum_{(3)} \kappa_{rs,tu}, \quad \kappa_{r,s,tu} = \kappa_{rstu} - \kappa_{rtu}^{(s)} - \kappa_{stu}^{(r)} + \kappa_{tu}^{(rs)} - \kappa_{rs,tu}.$$

Além das identidades de Bartlett a invariância entre os cumulantes é muito utilizado, para o cálculo de cumulantes, como apresentado em Cordeiro e Barroso (2003) e McCullagh e Nelder (1989, p. 462). Essa propriedade permite a troca da ordem da diferenciação e da integração, isto é, $\kappa_{rs,t} = \kappa_{st,r} = \kappa_{rt,s}$, cujo resultado é válido para os MLGs, mas não é uma regra geral.

Apêndice B

Propriedade dos estimadores de máxima verossimilhança

Teoria Assintótica de Primeira Ordem

A importância da teoria assintótica da verossimilhança é que produz simplificações consideráveis para problemas onde o tamanho da amostra pode ser considerado como grandes, ao contrário da teoria de amostras de tamanho pequeno a moderado onde as soluções exatas apresentam alto grau de complexidade.

As soluções aproximadas são necessárias em dois casos distintos. O primeiro quando não se tem a solução exata para o problema estatístico ou ela é complicada. Então, poderá ser de interesse obter uma aproximação simples em grandes amostras. O segundo caso, o mais freqüente, revela o papel central da teoria assintótica da verossimilhança na inferência estatística e ocorre quando realmente não se tem uma solução exata, sendo inevitável a obtenção da solução aproximada. Os resultados da teoria assintótica de primeira ordem são válidos quando n tende a infinito, onde n é o tamanho da amostra, e são decorrentes de técnicas de linearização local baseadas nas expansões da série de Taylor e nos teoremas centrais do limite.

Método de máxima verossimilhança

Em 1922, Fisher introduziu o método de máxima verossimilhança que constitui uma das técnicas mais utilizadas em estimação paramétrica, pois suas equações são práticas para serem usadas computacionalmente, e além disso, por Cordeiro (1999, p.94) os estimadores apresentam as propriedades tais como: consistência, unicidade, normalidade e eficiência, quando n tende a infinito. Porém, o problema é que elas são usualmente viesadas de $O(n^{-1})$.

Em geral, os estimadores de máxima verossimilhança não podem ser expressos analiticamente, e neste caso, são necessários métodos iterativos para a obtenção das raízes da equação de verossimilhança. Quando as duas primeiras derivadas da log-verossimilhança existem, aplica-

se a expansão de Taylor, em torno de alguma estimativa inicial θ_{n-1} , para obter as estimativas de máxima verossimilhança, isto é,

$$\frac{\partial \ell(\theta)}{\partial \theta} = \frac{\partial \ell(\theta_{n-1})}{\partial \theta} + (\hat{\theta}_n - \theta_{n-1}) \frac{\partial^2 \ell(\theta_{n-1})}{\partial \theta^2} = 0.$$

Então,

$$\hat{\theta}_n = \theta_{n-1} - \frac{\partial \ell(\theta_{n-1})}{\partial \theta} \left\{ \frac{\partial^2 \ell(\theta_{n-1})}{\partial \theta^2} \right\}^{-1}. \quad (\text{B.1})$$

O método de estimação de Newton-Raphson para a obtenção dos estimadores de máxima verossimilhança consiste numa dada iteração i , (B.1), substituir θ_{i-1} por θ_i , obtido no passo anterior. Mas, quando o valor de $\frac{\partial^2 \ell(\theta_{i-1})}{\partial \theta^2}$ é substituído por $E\left(-\frac{\partial^2 \ell(\theta_{i-1})}{\partial \theta^2}\right)$ em cada iteração tem-se o método escore de Fisher.

Segundo Cordeiro (1992, p.16) em muitas situações é mais fácil obter a inversa em forma analítica de $E\left(-\frac{\partial^2 \ell(\theta_{i-1})}{\partial \theta^2}\right)$ do que $\frac{\partial^2 \ell(\theta_{i-1})}{\partial \theta^2}$.

Teoria Assintótica de Segunda Ordem

Com objetivo de melhorar os resultados da teoria assintótica de primeira ordem alguns resultados referentes à teoria assintótica de segunda ordem são apresentados, e neste caso os erros associados as propriedades estatísticas são de ordem $O(n^{-2})$.

Para o cálculo do viés de primeira ordem, $O(n^{-1})$, são necessários cumulantes e as identidades de Bartlett, pois estas facilitam a obtenção dos cumulantes, uma vez que para determinada parametrização pode-se conduzir a um cálculo direto e simples de alguns cumulantes, sendo os demais calculados indiretamente através destas identidades. Portanto, a seguir será apresentado as identidades de Bartlett e o viés o estimador de máxima verossimilhança apresentados por Cox e Snell (1968).

Apêndice C

Viés dos EMV-Cox e Snell (1968)

Como foi visto no comportamento assintótico dos estimadores de máxima verossimilhança, em geral, estes estimadores são viesadas, porém, segundo Cordeiro (1986) e Botter e Cordeiro (1998) o viés é algumas vezes ignorado na prática, uma vez que comparando a ordem de seu viés, que é n^{-1} , com a do erro padrão da estimativa de máxima verossimilhança, que é $n^{-1/2}$, observa-se que o viés é desprezível se n for grande, mas em pequenas amostras este viés pode apresentar uma magnitude comparável ao do erro padrão de sua estimativa, assim, é interessante estimá-lo nestes casos.

A fórmula do viés de ordem n^{-1} , do estimador do parâmetro pelo método de máxima verossimilhança considerando observações identicamente distribuídas para modelos uniparamétricos foi deduzida por Bartlett (1953), e, para modelos multiparamétricos com observações não necessariamente identicamente distribuídos a dedução é apresentada por Cox e Snell (1968).

Para determinar o viés de ordem $O(n^{-1})$ do estimador de máxima verossimilhança, Cox e Snell (1968) expandem o sistema de p equações $r = 0$ até o termo de segunda ordem da série de Taylor mas, para isso, é necessário estimar inicialmente $(\hat{\theta} - \theta)$ através da expansão até primeira ordem de $r = 0$, isto é,

$$r = \frac{\partial \ell(\theta_r)}{\partial \theta} + \frac{\partial^2 \ell(\theta_r)}{\partial \theta^2} (\hat{\theta}_r - \theta_r) = U_r + \sum_s U_{rs} (\hat{\theta}_s - \theta_s) + O_p(1),$$

que na forma matricial fica,

$$U = K (\hat{\theta}_s - \theta_s) + O_p(1).$$

Resultando,

$$(\hat{\theta}_s - \theta_s) = K^{-1}U + O_p(1). \quad (\text{C.1})$$

Agora, expandindo r até segunda ordem, tem-se,

$$r = U_r + \sum_s U_{rs} (\hat{\theta}_s - \theta_s) + \frac{1}{2} \sum_{s,t} U_{rst} (\hat{\theta}_s - \theta_s) (\hat{\theta}_t - \theta_t) + o_p(1). \quad (\text{C.2})$$

Aplicando a esperança em (C.2) e como $r = 0$, tem-se

$$E(U_r) + \sum_s E\{U_{rs}(\hat{\theta}_s - \theta_s)\} + \frac{1}{2} \sum_s E\{U_{rst}(\hat{\theta}_s - \theta_s)(\hat{\theta}_t - \theta_t)\} + o_p(1) = 0. \quad (C.3)$$

Por outro lado a segunda expressão de (C.3) fica,

$$\sum_s E\{U_{rs}(\hat{\theta}_s - \theta_s)\} = \sum_s E(U_{rs}) \sum_s E(\hat{\theta}_s - \theta_s) + Cov\{U_{rs}, (\hat{\theta}_s - \theta_s)\}, \quad (C.4)$$

e por (C.1) tem-se que $\hat{\theta}_s - \theta_s = -\sum_t \kappa^{st} U_t$, então (C.4) resulta,

$$\begin{aligned} Cov\{U_{rs}, (\hat{\theta}_s - \theta_s)\} &= \sum_s E\left\{U_{rs} \left(-\sum_t \kappa^{st} U_t\right)\right\} \\ &= -\sum_t (\kappa^{st}) E(U_{rs} U_t) \\ &= -\sum_t \kappa^{st} \kappa_{rs,t}. \end{aligned} \quad (C.5)$$

E o terceiro termo da expressão (C.2) é calculada analogamente, resultando

$$\frac{1}{2} \sum_s E\{U_{rst}(\hat{\theta}_s - \theta_s)(\hat{\theta}_t - \theta_t)\} = \frac{1}{2} \sum_{s,t} \kappa_{rst} (-\kappa^{st}). \quad (C.6)$$

Assim, a equação (C.3), lembrando que $E(U_r) = 0$, é a soma das parciais (C.4), (C.5) e (C.6) resultando

$$\sum_s \kappa_{rs} \sum_s E(\hat{\theta}_s - \theta_s) - \sum_t \kappa^{st} \kappa_{rs,t} + \frac{1}{2} \sum_{s,t} \kappa_{rst} (-\kappa^{st}) + o_p(1) = 0,$$

e, como $\sum_s E(\hat{\theta}_s - \theta_s) = B(\hat{\theta})$ e como $o_p(1)$ é um valor desprezível, obtém-se

$$\sum_s \kappa_{rs} B(\hat{\theta}) = \left(\sum_{s,t} \kappa^{st} \kappa_{rs,t} + \frac{1}{2} \sum_{s,t} \kappa_{rst} \right), \quad (C.7)$$

cuja inversão resulta em

$$B(\hat{\theta}) = \sum_{s,t,u} \kappa^{ru} \kappa^{st} \left(\kappa_{rs,t} + \frac{1}{2} \kappa_{rst} \right), \quad (C.8)$$

para $r = 1, 2, \dots, p$ até ordem n^{-1} . Esta fórmula (C.8) é devida a Cox e Snell (1968) e através desta obtém-se o viés de ordem n^{-1} , $O(n^{-1})$, da estimativa de máxima verossimilhança em modelos multiparamétricos. Para a utilização da fórmula (C.8) basta conhecer a matriz inversa da informação de Fisher, os cumulantes com relação a todos os parâmetros e as identidades de Bartlett.

Estimativa corrigida de segunda ordem

A grande utilidade da equação (C.8) é definir um estimador de máxima verossimilhança corrigida até ordem dois, $O(n^{-2})$, que é dada por

$$\tilde{\theta}_r = \hat{\theta} - \hat{B}(\hat{\theta}_r), \quad (\text{C.9})$$

onde $\hat{B}(\cdot)$ é o viés $B(\cdot)$ avaliado em $\hat{\theta}$. O estimador de máxima verossimilhança corrigida $\tilde{\theta}_r$ é de ordem dois, $O(n^{-2})$.

Apêndice D

Cumulantes dos MLGD

Função densidade de probabilidade

$$p(y_l; \theta, \phi) = \exp[\phi_l \{y_l \theta_l - b(\theta_l) + c(y_l)\} + d_1(y_l) + d_2(\phi_l)]$$

Log-verossimilhança

$$\ell(\beta, \gamma) = \sum_{l=1}^n \phi_l \{y_l \theta_l - b(\theta_l) + c(y_l)\} + d_1(y_l) + d_2(\phi_l)$$

D.1 Cálculo de U_r e κ_r

Para obter o cumulante κ_r calcula-se o valor esperado da derivada da log-verossimilhança com relação a β_r , isto é,

$$\kappa_r = E\left(\frac{\partial \ell(\beta_l, \gamma_l)}{\partial \beta_r}\right) = E(U_r).$$

Então, pela regra da cadeia em,

$$U_r = \frac{\partial \ell(\beta_l, \gamma_l)}{\partial \theta_l} \frac{\partial \theta_l}{\partial \mu_l} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial \eta_l}{\partial \beta_r}.$$

Em ordem, a primeira parcela,

$$\frac{\partial \ell(\beta_l, \gamma_l)}{\partial \theta_l} = \sum_{l=1}^n \phi_l \left\{ y_l - \frac{\partial b(\theta_l)}{\partial \theta_l} \right\},$$

mas como $\frac{\partial b(\theta_l)}{\partial \theta_l} = \mu$

$$\frac{\partial \ell(\beta_l, \gamma_l)}{\partial \theta_l} = \sum_{l=1}^n \phi_l \{y_l - \mu_l\};$$

a segunda parcela,

$$\frac{\partial \theta_l}{\partial \mu_l} = V(\mu_l)^{-1};$$

a quarta parcela,

$$\frac{\partial \eta_l}{\partial \beta_r} = x_{lr}, \quad (\text{D.1})$$

pois,

$$\eta_l = \beta_0 + \beta_1 x_{l1} + \dots + \beta_r x_{lr} + \dots + \beta_p x_{lp}.$$

Portanto,

$$U_r = \sum_{l=1}^n \phi_l \{y_l - \mu_l\} V(\mu_l)^{-1} \frac{\partial \mu_l}{\partial \eta_l} x_{lr}. \quad (\text{D.2})$$

Portanto, o cumulante κ_r será,

$$\kappa_r = 0, \text{ pois } E(y_l - \mu_l) = 0.$$

D.2 Cálculo de U_{rs} e κ_{rs}

$$\kappa_{rs} = E\left(\frac{\partial^2 \ell(\beta_l, \gamma_l)}{\partial \beta_r \partial \beta_s}\right) = E\left(\frac{\partial}{\partial \beta_s} \left(\frac{\partial \ell(\beta_l, \gamma_l)}{\partial \beta_r}\right)\right) = E(U_{rs}).$$

Aplicando a regra da cadeia em

$$\begin{aligned} U_{rs} &= \frac{\partial}{\partial \beta_s} \left(\sum_{l=1}^n \phi_l \{y_l - \mu_l\} V(\mu_l)^{-1} \frac{\partial \mu_l}{\partial \eta_l} x_{lr} \right) \\ &= \sum_{l=1}^n \phi_l x_{lr} \left\{ -\frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} - (y_l - \mu_l) \frac{V^{(1)}}{\{V(\mu_l)\}^2} \frac{\partial \mu_l}{\partial \eta_l} \right. \\ &\quad \left. + \frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} \right\} x_{ls} \\ &= \sum_{l=1}^n \phi_l x_{lr} x_{ls} \left\{ -\frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l}\right)^2 - (y_l - \mu_l) \frac{V^{(1)}}{\{V(\mu_l)\}^2} \left(\frac{\partial \mu_l}{\partial \eta_l}\right)^2 \right. \\ &\quad \left. + (y_l - \mu_l) \frac{1}{V(\mu_l)} \frac{\partial^2 \mu_l}{\partial \eta_l^2} \right\} \\ &= \sum_{l=1}^n \phi_l \left\{ -\frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l}\right)^2 x_{lr} x_{ls} - (y_l - \mu_l) \frac{V^{(1)}}{\{V(\mu_l)\}^2} \left(\frac{\partial \mu_l}{\partial \eta_l}\right)^2 x_{lr} x_{ls} + \right. \\ &\quad \left. (y_l - \mu_l) \frac{1}{V(\mu_l)} \frac{\partial^2 \mu_l}{\partial \eta_l^2} x_{lr} x_{ls} \right\} \end{aligned} \quad (\text{D.3})$$

onde $V^{(1)} = \frac{\partial V(\mu_l)}{\partial \mu_l}$.

Aplicando a esperança em (D.3) tem-se

$$\kappa_{rs} = \sum_{l=1}^n \phi_l \left\{ -\frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l}\right)^2 x_{lr} x_{ls} \right\},$$

denotando

$$\frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 = w_l, \quad (\text{D.4})$$

conclui-se que

$$\kappa_{rst} = - \sum_{l=1}^n \phi_l w_l x_{lr} x_{ls}. \quad (\text{D.5})$$

D.3 Cálculo de U_{rst} e κ_{rst}

$$\kappa_{rst} = E \left(\frac{\partial^3 \ell(\beta_l, \gamma_l)}{\partial \beta_r \partial \beta_s \partial \beta_t} \right) = E \left(\frac{\partial}{\partial \beta_t} \left(\frac{\partial^2 \ell(\beta_l, \gamma_l)}{\partial \beta_r \partial \beta_s} \right) \right) = E(U_{rst}).$$

Derivando (D.3) tem-se

$$\begin{aligned} U_{rst} &= \sum_{l=1}^n \phi_l x_{lr} x_{ls} \left\{ \frac{-V^{(1)}}{(V(\mu_l))^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 + 2 \frac{1}{V(\mu_l)} \frac{\partial^2 \mu_l}{\partial \eta_l^2} \right\} \frac{\partial \mu}{\partial \eta} x_{lt} \\ &\quad - \sum_{l=1}^n \phi_l x_{lr} x_{ls} \left[\left\{ \begin{aligned} &\frac{-V^{(1)}}{(V(\mu_l))^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 \\ &+ (y_l - \mu_l) \frac{V^{(2)} V(\mu_l) - 2V^{(1)}}{(V(\mu_l))^3} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 \\ &+ 2 \frac{\partial^2 \mu_l}{\partial \eta_l^2} (y_l - \mu_l) \frac{V^{(1)}}{(V(\mu_l))^2} \end{aligned} \right\} \right] \frac{\partial \mu}{\partial \eta} x_{lt} \\ &\quad + \sum_{l=1}^n \phi_l x_{lr} x_{ls} \left[\begin{aligned} &-\frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial^2 \mu_l}{\partial \eta_l^2} + (y_l - \mu_l) \frac{V^{(1)}}{(V(\mu_l))^2} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial^2 \mu_l}{\partial \eta_l^2} \\ &+ (y_l - \mu_l) \frac{1}{V(\mu_l)} \frac{\partial^3 \mu_l}{\partial \eta_l^3} \end{aligned} \right] x_{lr} x_{ls} x_{lt} \end{aligned}$$

A seguir, calculando o valor esperado de U_{rst} obtém-se

$$\kappa_{rst} = - \sum_{l=1}^n \phi_l \left\{ \begin{aligned} &\left\{ -\frac{V^{(1)}}{(V(\mu_l))^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^3 + 2 \frac{1}{V(\mu_l)} \left(\frac{\partial^2 \mu_l}{\partial \eta_l^2} \right) \frac{\partial \mu_l}{\partial \eta_l} \right\} \\ &-\left\{ -\frac{V^{(1)}}{(V(\mu_l))^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^3 \right\} + \left(\frac{1}{V(\mu_l)} \frac{\partial^2 \mu_l}{\partial \eta_l^2} \frac{\partial \mu_l}{\partial \eta_l} \right) \end{aligned} \right\} x_{lr} x_{ls} x_{lt},$$

e denotando

$$f = \frac{1}{V(\mu)} \frac{\partial \mu}{\partial \eta} \frac{\partial^2 \mu}{\partial \eta^2} \quad \text{e} \quad (\text{D.6})$$

$$g = f - \frac{V^{(1)}}{(V(\mu_l))^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^3, \quad (\text{D.7})$$

resulta em

$$\kappa_{rst} = - \sum_{l=1}^n \phi_l (f + 2g) x_{lr} x_{ls} x_{lt}. \quad (\text{D.8})$$

D.4 Cálculo do $\kappa_{rs}^{(T)}$

$$\kappa_{rs}^{(T)} = \frac{\partial \kappa_{rs}}{\partial \gamma_T} = \frac{\partial \kappa_{rs}}{\partial \phi_l} \frac{\partial \phi_l}{\partial \tau_l} \frac{\partial \tau_l}{\partial \gamma_T}$$

Aplicando a regra da cadeia em (D.5) tem-se que

$$\begin{aligned} \kappa_{rs}^{(T)} &= - \sum_{l=1}^n w_l x_{lr} x_{ls} \frac{\partial \phi_l}{\partial \tau_l} s_{lT} \\ &= - \sum_{l=1}^n w_l \frac{\partial \phi_l}{\partial \tau_l} x_{lr} x_{ls} s_{lT} \end{aligned} \quad (D.9)$$

D.5 Cálculo de U_{stR} e κ_{stR}

$$\kappa_{stR} = E \left(\frac{\partial^3 \ell(\beta_l, \gamma_l)}{\partial \beta_r \partial \beta_s \partial \gamma_R} \right) = E(U_{stR})$$

Assim, derivando (D.3)

$$U_{stR} = - \sum_{l=1}^n \left[\begin{array}{c} \left\{ \frac{1}{V(\mu_l)} + (y_l - \mu_l) \frac{V^{(1)}}{(V(\mu_l))^2} \right\} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 \\ - (y_l - \mu_l) \frac{1}{V(\mu_l)} \frac{\partial^2 \mu_l}{\partial \eta_l^2} \frac{\partial \mu_l}{\partial \eta_l} \end{array} \right] \frac{\partial \phi_l}{\partial \tau_l} x_{ls} x_{st} s_{lR},$$

agora, aplicando a esperança na equação anterior obtém-se,

$$\kappa_{stR} = - \sum_{l=1}^n \frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 \frac{\partial \phi_l}{\partial \tau_l} x_{ls} x_{st} s_{lR}$$

e de (D.4) resulta que

$$\kappa_{stR} = - \sum_{l=1}^n w_l \frac{\partial \phi_l}{\partial \tau_l} x_{ls} x_{st} s_{lR}. \quad (D.10)$$

D.6 Cálculo do $\kappa_{rs}^{(t)}$

$$\kappa_{rs}^{(t)} = \frac{\partial \kappa_{rs}}{\partial \beta_t} = \frac{\partial \kappa_{rs}}{\partial \mu_l} \frac{\partial \mu_l}{\partial \eta_l} x_{lt}$$

Utilizando (D.5),

$$\begin{aligned}\kappa_{rs}^{(t)} &= \frac{\partial}{\partial \mu_l} \left[- \sum_{l=1}^n \phi_l \left\{ - \frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 x_{lr} x_{ls} \right\} \right] \frac{\partial \mu_l}{\partial \eta_l} x_{lt} \\ &= - \sum_{l=1}^n \phi_l \left\{ \frac{V^{(1)}}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^3 + 2 \frac{1}{V(\mu_l)} \frac{\partial^2 \mu_l}{\partial \eta_l^2} \frac{\partial \mu_l}{\partial \eta_l} \right\} x_{lr} x_{ls} x_{lt}\end{aligned}$$

e utilizando (D.6) e (D.7) na equação anterior tem-se

$$\kappa_{rs}^{(t)} = - \sum_{l=1}^n \phi_l (f + g)_l x_{lr} x_{ls} x_{lt}. \quad (D.11)$$

D.7 Cálculo da U_R

$$U_R = \frac{\partial \ell(\beta_l, \gamma_l)}{\partial \gamma_R}$$

Derivando $\ell(\beta_l, \gamma_l)$ com relação a γ_R obtém-se

$$\begin{aligned}U_R &= \frac{\partial}{\partial \phi_l} \left(\sum_{l=1}^n \phi_l \{y_l \theta_l - b(\theta_l) + c(y_l)\} + d_1(y_l) + d_2(\phi_l) \right) \frac{\partial \phi_l}{\partial \tau_l} \frac{\partial \tau_l}{\partial \gamma_R} \\ &= \sum_{l=1}^n \left[\{y_l \theta_l - b(\theta_l) + c(y_l)\} + \frac{dd_2(\phi_l)}{d\phi_l} \right] \frac{\partial \phi_l}{\partial \tau_l} s_{lR}.\end{aligned} \quad (D.12)$$

D.8 Cálculo de U_{RS} e κ_{RS}

$$k_{RS} = E \left(\frac{\partial}{\partial \gamma_S} \left(\frac{\partial \ell(\beta_l, \gamma_l)}{\partial \gamma_R} \right) \right) = E(U_{RS})$$

Para este cálculo é necessário obter o valor esperado de $\mathbf{c}(y)$, obtido a seguir.

D.8.1 Cálculo do Valor Esperado da $\mathbf{c}(y)$

De (D.12) e pelas condições de regularidade $\mu_R = E(U_R) = 0$ e como $E(y_l) = \mu_l$ tem-se que

$$\begin{aligned}E \left[\{y_l \theta_l - b(\theta_l) + c(y_l)\} \frac{\partial \phi_l}{\partial \tau_l} s_{lR} + \frac{dd_2(\phi_l)}{d\phi_l} \frac{\partial \phi_l}{\partial \tau_l} s_{lR} \right] &= 0 \\ \left[E(y_l \theta_l) - b(\theta) + E(c(y_l)) \right] + \frac{dd_2(\phi_l)}{d\phi_l} \frac{\partial \phi_l}{\partial \tau_l} s_{lR} &= 0 \\ \mu_l \theta_l - b(\theta_l) + E(c(y_l)) + \frac{dd_2(\phi_l)}{d\phi_l} &= 0,\end{aligned}$$

e portanto,

$$E(c(y_l)) = - \left\{ \mu_l \theta_l - b(\theta_l) + \frac{dd_2(\phi_l)}{d\phi_l} \right\}. \quad (D.13)$$

Para obter o cumulante κ_{RS} será calculado a seguir U_{RS} .

$$\begin{aligned} U_{RS} &= \sum_{l=1}^n \{y_l \theta_l - b(\theta_l) + c(y_l)\} \frac{\partial^2 \phi_l}{\partial \tau_l^2} s_{lR} s_{lS} \\ &\quad + \frac{d^2 d_2(\phi_l)}{d\phi_l^2} \left(\frac{\partial \phi_l}{\partial \tau_l} \right)^2 s_{lR} s_{lS} + \frac{dd_2(\phi_l)}{d\phi_l} \frac{\partial^2 \phi_l}{\partial \tau_l^2} s_{lR} s_{lS} \end{aligned} \quad (D.14)$$

Aplicando esperança em (D.14) e utilizando (D.13) o cumulante κ_{RS} resulta em

$$\kappa_{RS} = \sum_{l=1}^n \frac{d^2 d_2(\phi_l)}{d\phi_l^2} \left(\frac{\partial \phi_l}{\partial \tau_l} \right)^2 s_{lR} s_{lS}. \quad (D.15)$$

D.9 Cálculo do $\kappa_{RS}^{(t)}$

$$\frac{\partial \kappa_{RS}}{\partial \beta_t} = \frac{\partial}{\beta_t} \left(\sum_{l=1}^n \frac{d^2 d_2(\phi_l)}{d\phi_l^2} \left(\frac{\partial \phi_l}{\partial \tau_l} \right)^2 s_{lR} s_{lS} \right),$$

conclui-se que

$$\kappa_{RS}^{(t)} = 0.$$

D.10 Cálculo do $\kappa_{RS}^{(T)}$

$$\kappa_{RS}^{(T)} = \frac{\partial \kappa_{RS}}{\partial \gamma_T}$$

Aplicando a regra da cadeia

$$\begin{aligned} \kappa_{RS}^{(T)} &= \frac{\partial}{\partial \gamma_T} \left(\frac{d^2 d_2(\phi_l)}{d\phi_l^2} \left(\frac{\partial \phi_l}{\partial \tau_l} \right)^2 s_{lR} s_{lS} \right) \\ &= \sum_{l=1}^n \left\{ \frac{d^3 d_2(\phi_l)}{d\phi_l^3} \left(\frac{\partial \phi_l}{\partial \tau_l} \right)^2 + 2 \frac{d^2 d_2(\phi_l)}{d\phi_l^2} \frac{\partial^2 \phi_l}{\partial \tau_l^2} \right\} \frac{\partial \phi_l}{\partial \tau_l} s_{lR} s_{lS} s_T \\ &= \sum_{l=1}^n \left\{ \frac{d^3 d_2(\phi_l)}{d\phi_l^3} \left(\frac{\partial \phi_l}{\partial \tau_l} \right)^3 + 2 \frac{\partial^2 d_2(\phi_l)}{\partial \phi_l^2} \frac{\partial^2 \phi_l}{\partial \tau_l^2} \frac{\partial \phi_l}{\partial \tau_l} \right\} s_{lR} s_{lS} s_T \end{aligned} \quad (D.16)$$

D.11 Cálculo de U_{rS} e κ_{rS}

$$\kappa_{rS} = E \left(\frac{\partial^2 \ell(\beta_l, \gamma_l)}{\partial \beta_r \partial \gamma_S} \right) = E \left(\frac{\partial}{\partial \gamma_S} \left(\frac{\partial \ell(\beta_l, \gamma_l)}{\partial \beta_r} \right) \right) = E(U_{rS}).$$

derivando pela regra da cadeia a (D.2),

$$\begin{aligned} U_{rS} &= \frac{\partial}{\partial \phi_l} \left(\sum_{l=1}^n \phi_l \{y_l - \mu_l\} V(\mu_l)^{-1} \frac{\partial \mu_l}{\partial \eta_l} x_{lr} \right) \frac{\partial \phi_l}{\partial \tau_l} s_{lS} \\ &= \sum_{l=1}^n (y_l - \mu_l) V(\mu_l)^{-1} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial \phi_l}{\partial \tau_l} x_{lr} s_{lS}. \end{aligned} \quad (D.17)$$

Portanto, aplicando esperança em (D.17) obtém-se

$$\kappa_{rS} = 0. \quad (D.18)$$

D.12 Cálculo do $\kappa_{rS}^{(T)}$

$$\kappa_{rS}^{(T)} = \frac{\partial \kappa_{rS}}{\partial \gamma_T}$$

Portanto, pela equação (D.18) observa-se que

$$\kappa_{rS}^{(T)} = 0. \quad (D.19)$$

D.13 Cálculo do κ_{rST}

$$\kappa_{rST} = E \left(\frac{\partial^3 \ell(\beta_l, \gamma_l)}{\partial \beta_r \partial \gamma_S \partial \gamma_T} \right) = E \left(\frac{\partial}{\partial \gamma_T} \left(\frac{\partial^2 \ell(\beta_l, \gamma_l)}{\partial \beta_r \partial \gamma_S} \right) \right) = E(U_{rST})$$

Logo, derivando pela regra da cadeia (D.17) com relação a γ_T ,

$$\begin{aligned} U_{rST} &= \frac{\partial}{\partial \tau_l} \left(\sum_{l=1}^n (y_l - \mu_l) \frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial \phi_l}{\partial \tau_l} x_{lr} s_{lS} \right) \frac{\partial \tau_l}{\partial \gamma_T} \\ &= \sum_{l=1}^n (y_l - \mu_l) \frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial^2 \phi_l}{\partial \tau_l^2} x_{lr} s_{lS} s_{lT}. \end{aligned}$$

Portanto, aplicando esperança tem-se que

$$\kappa_{rST} = 0. \quad (D.20)$$

D.14 Cálculo do κ_{RST}

$$\kappa_{RST} = E \left(\frac{\partial^3 \ell(\beta_l, \gamma_l)}{\partial \gamma_R \partial \gamma_S \partial \gamma_T} \right) = E(U_{RST})$$

Assim, calculando

$$\begin{aligned} U_{RST} &= \frac{\partial}{\partial \gamma_T} \left[\sum_{l=1}^n \left[\begin{aligned} &\{y_l \theta_l - b(\theta_l) + c(y_l)\} \frac{\partial^2 \phi_l}{\partial \tau_l^2} s_{lR} s_{lR} \\ &+ \frac{d^2 d_2(\phi_l)}{d\phi_l^2} \left(\frac{\partial \phi_l}{\partial \tau_l} \right)^2 s_{lR} s_{lS} + \frac{d d_2(\phi_l)}{d\phi_l} \frac{\partial^2 \phi_l}{\partial \tau_l^2} s_{lR} s_{lS} \end{aligned} \right] \right] \\ &= \sum_{l=1}^n \left[\begin{aligned} &\{y_l \theta_l - b(\theta_l) + c(y_l)\} \frac{\partial^3 \phi_l}{\partial \tau_l^3} + \frac{d^3 d_2(\phi_l)}{d\phi_l^3} \left(\frac{\partial \phi_l}{\partial \tau_l} \right)^3 \\ &+ 3 \frac{d^2 d_2(\phi_l)}{d\phi_l^2} \frac{\partial \phi_l}{\partial \tau_l} \frac{\partial^2 \phi_l}{\partial \tau_l^2} + \frac{d d_2(\phi_l)}{d\phi_l} \frac{\partial^3 \phi_l}{\partial \tau_l^3} \end{aligned} \right] s_{lR} s_{lS} s_{lT}, \end{aligned}$$

e aplicando a esperança e utilizando (D.13) resulta em

$$\kappa_{RST} = \sum_{l=1}^n \left[\frac{d^3 d_2(\phi_l)}{d\phi_l^3} \left(\frac{\partial \phi_l}{\partial \tau_l} \right)^3 + 3 \frac{d^2 d_2(\phi_l)}{d\phi_l^2} \frac{\partial^2 \phi_l}{\partial \tau_l^2} \frac{\partial \phi_l}{\partial \tau_l} \right] s_{lR} s_{lS} s_{lT}. \quad (D.21)$$

Apêndice E

Cumulantes dos MNLFED

Função densidade de probabilidade

$$p(y_l; \theta, \phi) = \exp[\phi_l \{y_l \theta_l - b(\theta_l) + c(y_l)\} + d_1(y_l) + d_2(\phi_l)]$$

Log-verossimilhança

$$\ell(\beta, \gamma) = \sum_{l=1}^n [\phi_l \{y_l \theta_l - b(\theta_l) + c(y_l)\} + d_1(y_l) + d_2(\phi_l)]$$

E.1 Cálculo de U_r e κ_r

Para obter o cumulante κ_r calcula-se o valor esperado da derivada da log-verossimilhança com relação a β_r , isto é, $\kappa_r = E\left(\frac{\partial \ell(\beta_l, \gamma_l)}{\partial \beta_r}\right) = E(U_r)$. Assim, pela regra da cadeia obtém-se

$$U_r = \frac{\partial \ell(\beta_l, \gamma_l)}{\partial \theta_l} \frac{\partial \theta_l}{\partial \mu_l} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial \eta_l}{\partial \beta_r},$$

logo,

$$\frac{\partial \ell(\beta_l, \gamma_l)}{\partial \beta_r} = \sum_{l=1}^n \phi_l \left\{ y_l - \frac{\partial b(\theta_l)}{\partial \theta_l} \right\} \frac{\partial \theta_l}{\partial \mu_l} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial \eta_l}{\partial \beta_r},$$

como $\frac{\partial b(\theta_l)}{\partial \theta_l} = \mu$, $\frac{\partial \theta_l}{\partial \mu_l} = \frac{1}{V(\mu)}$ e denotando

$$\frac{\partial \eta_l}{\partial \beta_r} = (r)$$

tem-se que

$$U_r = \sum_{l=1}^n \phi_l (y_l - \mu_l) \frac{1}{V(\mu)} \frac{\partial \mu_l}{\partial \eta_l} (r). \quad (\text{E.1})$$

Portanto, aplicando a esperança o cumulante

$$\kappa_r = 0.$$

E.2 Cálculo de U_{rs} e κ_{rs}

$$\kappa_{rs} = E \left(\frac{\partial^2 \ell(\beta_l, \gamma_l)}{\partial \beta_r \partial \beta_s} \right) = E \left(\frac{\partial}{\partial \beta_s} (U_r) \right).$$

Por (E.1) tem-se que

$$\begin{aligned} U_{rs} &= \sum_{l=1}^n \frac{\partial}{\partial \beta_s} \left(\phi_l (y_l - \mu_l) \frac{1}{V(\mu)} \frac{\partial \mu_l}{\partial \eta_l} (r) \right) \\ &= \sum_{l=1}^n \phi_l \left\{ \begin{aligned} &-\frac{1}{V(\mu_l)} - (y_l - \mu_l) \frac{V^{(1)}}{\{V(\mu_l)\}^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 (r, s)_l \\ &+ (y_l - \mu_l) \frac{1}{V(\mu_l)} \frac{\partial^2 \mu_l}{\partial \eta_l^2} (r, s)_l \\ &+ (y_l - \mu_l) \frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} (rs)_l \end{aligned} \right\}, \end{aligned} \quad (E.2)$$

em que

$$\frac{\partial \eta_l}{\partial \beta_s} = (s), \quad \frac{\partial \eta_l}{\partial \beta_r} \frac{\partial \eta_l}{\partial \beta_s} = (r, s), \quad \frac{\partial^2 \eta_l}{\partial \beta_r \partial \beta_s} = (rs), \quad \text{e} \quad V^{(1)} = \frac{\partial V(\mu_l)}{\partial \mu_l}.$$

Portanto,

$$\kappa_{rs} = E \left(\sum_{l=1}^n -\phi_l \left\{ -\frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 (r, s) \right\} \right),$$

logo, como (D.4) conclui-se que

$$\kappa_{rs} = - \sum_{l=1}^n -\phi_l w_l (r, s). \quad (E.3)$$

E.3 Cálculo do $\kappa_{rs}^{(t)}$

$$\kappa_{rs}^{(t)} = \frac{\partial \kappa_{rs}}{\partial \beta_t}$$

Portanto, derivando (E.2) com relação a β_t tem-se

$$\begin{aligned}
\kappa_{rs}^{(t)} &= - \sum_{l=1}^n \frac{\partial}{\partial \beta_t} \left[\phi_l \left\{ -\frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 (r, s) \right\} \right] \\
&= - \sum_{l=1}^n \phi_l \left[\left\{ -\frac{V^{(1)}}{[V(\mu_l)]^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^3 + \frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial^2 \mu_l}{\partial \eta_l^2} + \frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial^2 \mu_l}{\partial \eta_l^2} \right\} (r, s, t) \right. \\
&\quad \left. + \frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 \{(r, st) + (s, rt)\} \right].
\end{aligned}$$

e por (D.6) e (D.7) tem-se que

$$\kappa_{rs}^{(t)} = - \sum_{l=1}^n \phi_l [(f_l + g_l)(r, s, t)_l + w_l \{(r, st) + (s, rt)\}] \quad (\text{E.4})$$

E.4 Cálculo de U_{rst} e κ_{rst}

$$\kappa_{rst} = E \left(\frac{\partial^3 \ell(\beta_l, \gamma_l)}{\partial \beta_r \partial \beta_s \partial \beta_t} \right) = E \left\{ \frac{\partial}{\partial \beta_t} \left(\frac{\partial^2 \ell(\beta_l, \gamma_l)}{\partial \beta_r \partial \beta_s} \right) \right\} = E(U_{rst})$$

Derivando (E.2) e denotando $\frac{\partial^2 V(\mu_l)}{\partial \mu_l^2} = V^{(2)}$ tem-se

$$\begin{aligned}
U_{rst} &= \sum_{l=1}^n \phi_l \left\{ \begin{aligned} &\frac{V^{(1)}}{(V(\mu_l))^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^3 (r, s, t)_l - 2 \frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial^2 \mu_l}{\partial \eta_l^2} (r, s, t)_l \\ &- \frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 \{(rt, s) + (st, r)\} \end{aligned} \right\} \\
&+ \sum_{l=1}^n \phi_l \left\{ \begin{aligned} &-(y_l - \mu_l) \frac{V^{(2)}V(\mu_l) - 2V^{(1)}}{(V(\mu_l))^3} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^3 (r, s, t)_l \\ &+ \frac{-V^{(1)}}{(V(\mu_l))^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^3 (r, s, t)_l \\ &- 2(y_l - \mu_l) \frac{V^{(1)}}{(V(\mu_l))^2} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial^2 \mu_l}{\partial \eta_l^2} (r, s, t)_l \\ &-(y_l - \mu_l) \frac{V^{(1)}}{(V(\mu_l))^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 \{(rt, s)_l + (st, r)_l\} \end{aligned} \right\} \\
&+ \sum_{l=1}^n \phi_l \left\{ \begin{aligned} &\left\{ -\frac{1}{V(\mu_l)} - (y_l - \mu_l) \frac{V^{(1)}}{(V(\mu_l))^2} \right\} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial^2 \mu_l}{\partial \eta_l^2} (r, s, t)_l \\ &+ (y_l - \mu_l) \frac{1}{V(\mu_l)} \frac{\partial^3 \mu_l}{\partial \eta_l^2 \partial \mu_l} (r, s, t)_l \\ &-(y_l - \mu_l) \frac{1}{V(\mu_l)} \frac{\partial^2 \mu_l}{\partial \eta_l^2} \{(rt, s)_l + (st, r)_l\} \end{aligned} \right\} \\
&+ \sum_{l=1}^n \phi_l \left\{ \begin{aligned} &-\frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 (rs, t) - (y_l - \mu_l) \frac{V^{(1)}}{(V(\mu_l))^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 (rs, t)_l \\ &+ (y_l - \mu_l) \frac{1}{V(\mu_l)} \frac{\partial^2 \mu_l}{\partial \eta_l^2} (rs, t)_l + (y_l - \mu_l) \frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} (rst)_l \end{aligned} \right\}.
\end{aligned}$$

Portanto,

$$\begin{aligned}
\kappa_{rst} &= \sum_{l=1}^n \phi_l \left\{ \begin{aligned} &\frac{V^{(1)}}{(V(\mu_l))^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^3 (r, s, t)_l - 2 \frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial^2 \mu_l}{\partial \eta_l^2} (r, s, t)_l \\ & - \frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 [(s, rt) + (r, st)] + \frac{-V^{(1)}}{(V(\mu_l))^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^3 (r, s, t)_l \\ & - \frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial^2 \mu_l}{\partial \eta_l^2} (r, s, t)_l - \frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 (t, rs) \end{aligned} \right\} \\
&= \sum_{l=1}^n \phi_l \left\{ \begin{aligned} &\left[2 \frac{V^{(1)}}{(V(\mu_l))^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^3 - 2 \frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial^2 \mu_l}{\partial \eta_l^2} - \frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial^2 \mu_l}{\partial \eta_l^2} \right] (r, s, t)_l \\ & - \frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 [(s, rt) + (r, st) + (t, rs)] \end{aligned} \right\}.
\end{aligned}$$

De (D.6) e (D.7) κ_{rst} resulta em

$$\kappa_{rst} = - \sum_{l=1}^n \phi_l (f_l + 2g_l) (r, s, t)_l + w_l \{ (r, st)_l + (s, tr)_l + (t, rs)_l \} \quad (E.5)$$

E.5 Cálculo do $\kappa_{rs,t}$

$$\kappa_{rs,t} = E(U_{rs}U_t)$$

De (E.1) e (E.2) tem-se que

$$U_{rs}U_t = \sum_l \sum_j \phi_l \phi_j \left\{ \begin{aligned} & - \frac{1}{V(\mu_l)} \frac{1}{V(\mu_j)} (y_j - \mu_j) \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 \frac{\partial \mu_j}{\partial \eta_j} (r, s)_l (t)_j \\ & + (y_l - \mu_l) (y_j - \mu_j) \frac{V^{(1)}}{(V(\mu_l))^2} \frac{1}{V(\mu_j)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 \frac{\partial \mu_j}{\partial \eta_j} (r, s)_l (t)_j \\ & + (y_l - \mu_l) (y_j - \mu_j) \frac{1}{V(\mu_l)} \frac{1}{V(\mu_j)} \frac{\partial^2 \mu_l}{\partial \eta_l^2} \frac{\partial \mu_l}{\partial \eta_l} (rs)_l (t)_j \\ & + (y_l - \mu_l) (y_j - \mu_j) \frac{1}{V(\mu_l)} \frac{1}{V(\mu_j)} \frac{\partial \mu_l}{\partial \eta_l} \frac{\partial \mu_j}{\partial \eta_j} (rs)_l (t)_j \end{aligned} \right\}$$

e portanto, o seu cumulante resulta em

$$\kappa_{rs,t} = \sum_l \phi_l^2 \left\{ \begin{aligned} & -V(y_l) \frac{V^{(1)}}{(V(\mu_l))^3} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^3 (r, s, t)_l + V(y_l) \frac{1}{(V(\mu_l))^2} \frac{\partial^2 \mu_l}{\partial \eta_l^2} \frac{\partial \mu_l}{\partial \eta_l} (r, s, t)_l \\ & + V(y_l) \frac{1}{(V(\mu_l))^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 (rs, t)_l \end{aligned} \right\}$$

mas, como $V(y_l) = \phi^{-1}V(\mu_l)$,

$$\kappa_{rs,t} = \sum_l \phi_l \left[\begin{aligned} & \left\{ - \frac{V^{(1)}}{(V(\mu_l))^2} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^3 + \frac{1}{V(\mu_l)} \frac{\partial^2 \mu_l}{\partial \eta_l^2} \frac{\partial \mu_l}{\partial \eta_l} \right\} (r, s, t)_l \\ & + \frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 (t, rs)_l \end{aligned} \right],$$

que de (D.6), (D.7) e (D.4) resulta em

$$\kappa_{rs,t} = \sum_l \phi_l [g_l(r, s, t)_l + w_l(t, rs)_l] \quad (\text{E.6})$$

E.6 Cálculo de U_R

$$U_R = \frac{\partial \ell(\beta_l, \gamma_l)}{\partial \gamma_R}$$

Pela regra da cadeia,

$$\begin{aligned} U_R &= \frac{\partial \ell(\beta_l, \gamma_l)}{\partial \phi_l} \frac{\partial \phi_l}{\partial \tau_l} \frac{\partial \tau_l}{\partial \gamma_R} \\ &= \sum_{l=1}^n \left[\{y_l \theta_l - b(\theta_l) + c(y_l)\} + \frac{dd_2(\phi_l)}{d\phi_l} \right] \frac{\partial \phi_l}{\partial \tau_l} \frac{\partial \tau_l}{\partial \gamma_R} \end{aligned}$$

denotando

$$\frac{\partial \tau_l}{\partial \gamma_R} = (R)_l, \quad \frac{\partial \phi_l}{\partial \tau_l} = \phi_{1l} \quad \text{e} \quad \frac{dd_2(\phi_l)}{d\phi_l} = d_{1l} \quad (\text{E.7})$$

obtém-se

$$U_R = \sum_{l=1}^n [\{y_l \theta_l - b(\theta_l) + c(y_l)\} \phi_{1l} (R)_l + d_{1l} \phi_{1l} (R)_l]. \quad (\text{E.8})$$

E.7 Cálculo de U_{RS} e κ_{RS}

$$\kappa_{RS} = E \left(\frac{\partial^2 \ell(\beta_l, \gamma_l)}{\partial \gamma_R \partial \gamma_S} \right) = E \left\{ \frac{\partial}{\partial \gamma_S} \left(\frac{\partial \ell(\beta_l, \gamma_l)}{\partial \gamma_R} \right) \right\} = E(U_{RS})$$

Derivando (E.8),

$$U_{RS} = \frac{\partial}{\partial \gamma_S} \sum_l \left\{ (y_l \theta_l - b(\theta_l) + c(y_l)) \frac{\partial \phi_l}{\partial \tau_l} \frac{\partial \tau_l}{\partial \gamma_R} + \frac{dd_2(\phi_l)}{d\phi_l} \frac{\partial \phi_l}{\partial \tau_l} \frac{\partial \tau_l}{\partial \gamma_R} \right\}.$$

Denotando

$$\frac{\partial \tau_l}{\partial \gamma_R} \frac{\partial \tau_l}{\partial \gamma_S} = (R, S)_l, \quad \frac{\partial^2 \tau_l}{\partial \gamma_R \partial \gamma_S} = (RS)_l, \quad \frac{\partial^2 \phi_l}{\partial \tau_l^2} = \phi_{2l} \quad \text{e} \quad \frac{d^2 d_2(\phi_l)}{d\phi_l^2} = d_{2l} \quad (\text{E.9})$$

e por (E.7) obtém-se

$$\begin{aligned}
U_{RS} &= \sum_l \{(y_l \theta_l - b(\theta_l) + c(y_l)) (\phi_{2l}(R, S)_l + \phi_{1l}(RS)_l)\} \\
&+ \sum_l \{d_{2l}(\phi_{1l})^2(R, S)_l\} + \sum_l \{\phi_{1l} \phi_{2l}(R, S)_l\} \\
&+ \sum_l \{d_{1l} \phi_{1l}(RS)_l\}.
\end{aligned}$$

E aplicando a esperança e de (D.13) resulta em

$$\kappa_{RS} = \sum_l d_{2l}(\phi_{1l})^2(R, S)_l. \quad (\text{E.10})$$

E.8 Cálculo de $U_{RS}^{(T)}$ e do cumulante $\kappa_{RS}^{(T)}$

$$\kappa_{RS}^{(T)} = \frac{\partial \kappa_{RS}}{\partial \gamma_T}$$

Derivando (E.10),

$$\begin{aligned}
\kappa_{RS}^{(T)} &= \sum_l \frac{\partial}{\partial \gamma_T} [d_{2l}(\phi_{1l})^2(R, S)_l] \\
&= \sum_l \{d_{3l}(\phi_{1l})^3 + 2d_{2l}\phi_{1l}\phi_{2l}\}(R, S, T)_l \\
&+ \sum_l d_{2l}(\phi_{1l})^2 \{(R, ST)_l + (S, RT)_l\}
\end{aligned} \quad (\text{E.11})$$

E.9 Cálculo de U_{RST} e do cumulante κ_{RST}

$$\kappa_{RST} = E \left(\frac{\partial^3 \ell(\beta_l, \gamma_l)}{\partial \gamma_R \partial \gamma_S \partial \gamma_T} \right) = E \left(\frac{\partial}{\partial \gamma_T} \left(\frac{\partial^2 \ell(\beta_l, \gamma_l)}{\partial \gamma_R \partial \gamma_S} \right) \right) = E(U_{RST})$$

Denotando

$$\frac{\partial^3 \phi_l}{\partial \tau_l^3} = \phi_{3l} \quad \text{e} \quad \frac{d^3 d_2(\phi_l)}{d\phi_l^3} = d_{3l},$$

obtém-se

$$\begin{aligned}
U_{RST} &= \sum_l \left\{ (y_l \theta_l - b(\theta_l) + c(y_l)) \begin{pmatrix} \phi_{3l}(R, S, T)_l + \phi_{2l}(RT, S)_l \\ + \phi_{2l}(ST, R)_l + \phi_{2l}(RS, T)_l \\ + \phi_{1l}(RST)_l \end{pmatrix} \right\} \\
&+ \sum_l \left\{ \begin{matrix} d_{3l}(\phi_{1l})^3(R, S, T)_l + 2d_{2l}\phi_{1l}\phi_{2l}(R, S, T)_l \\ + d_{2l}(\phi_{1l})^2 \{(R, ST) + (S, RT)\} \end{matrix} \right\} \\
&+ \sum_l \left\{ \begin{matrix} d_{2l}\phi_{1l}\phi_{2l}(R, S, T) + d_{1l}\phi_{3l}(R, S, T) \\ + d_{1l}\phi_{2l} \{(R, ST)_l + (S, RT)_l\} \end{matrix} \right\} \\
&+ \sum_l \left\{ \begin{matrix} d_{1l}\phi_{2l}(S, RT)_l + d_{1l}\phi_{2l}(R, ST)_l \\ + d_{2l}\phi_{1l}\phi_{2l}(R, S, T)_l + d_{1l}\phi_{1l}(RST)_l \end{matrix} \right\} \\
&= \sum_l \left\{ \begin{matrix} [y_l \theta_l - b(\theta_l) + c(y_l)](\phi_{3l}) + d_{3l}(\phi_{1l})^3 \\ + 2d_{2l}\phi_{1l}\phi_{2l} + d_{2l}\phi_{1l}\phi_{2l} + d_{1l}\phi_{3l} \end{matrix} \right\} (R, S, T)_l \\
&+ \sum_l \left\{ \begin{matrix} (y_l \theta_l - b(\theta_l) + c(y_l))\phi_{2l} + d_{1l}\phi_{2l} \\ + d_{2l}(\phi_{1l})^2 \end{matrix} \right\} \begin{pmatrix} (R, ST)_l \\ + (S, RT)_l \\ + (T, RS)_l \end{pmatrix} \\
&+ \sum_l \{(y_l \theta_l - b(\theta_l) + c(y_l))\phi_{1l} + d_{1l}\phi_{1l}\} (RST)_l
\end{aligned}$$

Agora, aplicando a esperança obtém-se

$$\begin{aligned}
\kappa_{RST} &= \sum_l \{d_{3l}(\phi_{1l})^3 + 3d_{2l}\phi_{1l}\phi_{2l}\} (R, S, T)_l \\
&+ \sum_l \{d_{2l}(\phi_{1l})^2\} \{(R, ST)_l + (S, RT)_l + (T, RS)_l\} \quad (E.12)
\end{aligned}$$

E.10 Cálculo de U_{Rst} e do cumulante κ_{Rst}

$$\kappa_{Rst} = E \left(\frac{\partial^3 \ell(\beta_l, \gamma_l)}{\partial \gamma_R \partial \beta_s \partial \beta_t} \right) = E \left(\frac{\partial}{\partial \beta_t} \left(\frac{\partial^2 \ell(\beta_l, \gamma_l)}{\partial \gamma_R \partial \beta_s} \right) \right) = E(U_{Rst})$$

Derivando (E.8) tem-se que

$$\begin{aligned}
U_{Rs} &= \sum_{l=1}^n \frac{\partial}{\partial \beta_s} [\{y_l \theta_l - b(\theta_l) + c(y_l)\} \phi_{1l}(R)_l + d_{1l}\phi_{1l}(R)_l] \\
&= \sum_{l=1}^n (y_l - \mu_l) \frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l}(s)_l \phi_{1l}(R)_l.
\end{aligned}$$

E aplicando a esperança tem-se que

$$\kappa_{Rs} = 0. \quad (E.13)$$

Em seguida derivando com relação a β_t ,

$$\begin{aligned}
U_{Rst} &= \sum_{l=1}^n \frac{\partial}{\partial \beta_t} \left\{ (y_l - \mu_l) \frac{1}{V(\mu_l)} \frac{\partial \mu_l}{\partial \eta_l} (s)_l \phi_{1l}(R)_l \right\} \\
&= \sum_{l=1}^n \phi_{1l}(R)_l \left\{ \begin{aligned} &(y_l - \mu_l) \frac{1}{V(\mu_l)} \left[\frac{\partial^2 \mu_l}{\partial \eta_l^2} (s, t)_l + \frac{\partial \mu_l}{\partial \eta_l} (st)_l \right] \\ &- \left[\frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 + (y_l - \mu_l) \frac{V^{(1)}}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 \right] (s, t)_l. \end{aligned} \right\}
\end{aligned}$$

Portanto, o cumulante será

$$\kappa_{Rst} = - \sum_{l=1}^n \frac{1}{V(\mu_l)} \left(\frac{\partial \mu_l}{\partial \eta_l} \right)^2 \phi_{1l}(R, s, t)_l$$

e de (D.4)

$$\kappa_{Rst} = - \sum_{l=1}^n w_l \phi_{1l}(R, s, t)_l. \tag{E.14}$$

Apêndice F

Programa de simulação para inversa gaussiana

Listagem F.1: Programa de simulacao para variável resposta inversa gaussiana

```
1
2  /*****
3  /*          SIMULACAO: MODELO INVERSA GAUSSIANA
4  /*          MNLFED
5  /*  VRESPOSTA DE Y : Inversa Gaussiana
6  /*          SUBMODELO DA MEDIA: MODELO DE GOMPERTZ
7  /*          SUBMODELO DA DISPERSAO: MODELO EXPONENCIAL
8  /*          COVARIAVEL DO SUBMODELO MEDIA: X : U(0,1)
9  /*          COVARIAVEL DO SUBMODELODISPERSAO: S1 : U(0,1)
10 /*          S2 : U(0,1)
11 /*****/
12 options nonotes nodetails; libname libref "f:\marga";
13 libname out f:\marga";
14 %macro geracao(quantidade=m);
15     /* Inicializacao do Vetor das Estimativas */
16 data out.estimativas; beta1 = 0; beta2 = 0; beta3 = 0; gama1 = 0;
17 gama2 = 0; vbeta1=0; vbeta2=0; vbeta3=0;vg1=0;vg2=0;
18 cbeta1=0;cbeta2=0;cbeta3=0;cg1=0;cg2=0;stddevb1=0;
19 stddevb2=0;stddevb3=0; stddevg1=0;stddevg2=0;
20 run;
21 /* Simulacao das Matrizes dos Modelo */
22     data out.dados(drop=i);
23     do i=1 to &m;
24         x1 = uniform(i+12);
25         s1 = uniform(i+78);
```

```

26             s2                =      uniform(i+91);
27 /*          Submodelo da Media mu dispersao phi          */
28             mu                =      1*exp(-exp(2-3*x1));
29             phi                =      exp(1.5*s1+2.5*s2);
30             output;
31         end;
32 run;
33 %do i = 1 %to &quantidade;
34 /*          Simulacao da Variavel Resposta Dado X e S Com
35             Distribuicao de Probabilidade inversa gaussiana */
36 proc iml;
37     use out.dados;
38     read all;
39     close out.dados;
40     y = j(&m,1,0);
41     do j = 1 to &m;
42         y[j] = normal(j+&m+&i+62579)**2;
43     end;
44     p1 = mu+mu##2#y/(2#phi);
45     p2 = -mu/(2#phi)#sqrt(4*mu#phi#y+mu##2#y##2);
46     x = p1+p2;
47     out = j(&m,1,0); u = out;
48     do j = 1 to &m;
49         u[j] = uniform(j+&m+&i);
50     end;
51     do j = 1 to &m;
52         if u[j] <= mu[j]/(mu[j]+x[j]) then out[j] = x[j];
53         else out[j] = mu[j]##2/x[j];
54     end;
55     saida                =      out || x1 || s1 || s2;
56     create out.saida from saida;
57     append from saida;
58     quit;
59 /*          Estimativa de Maxima Verossimilhanca para Beta
60             Independentemente de Gama          */
61     proc nlp data = out.saida tech=tr noprint
62         outest=out.emvsbeta(keep=beta1 beta2 beta3 _type_
63                             where=( _type_ in("PARMS"))));
64     max l;
65     parms

```

```

66     beta1      = 1,
67     beta2      = 2,
68     beta3      = 3;
69     gama1      = 1.5;
70     gama2      = 2.5;
71
72     y          = col1;
73     x1         = col2;
74     s1         = col3;
75     s2         = col4;
76     mu         = beta1*exp(-exp(beta2-beta3*x1));
77     phi        = exp(gama1*s1+gama2*s2);
78
79     teta       = -1/(2*mu**2);
80     bedeteta  = -(-2*teta)**.5;
81     cdey      = -1/(2*y);
82     L         = phi*(y*teta-bedeteta+cdey)+.5*log(phi);
83
84 /*      Estimativa de maxima Verossimilhanca para Gama
85           Independientemente de Beta          */
86 proc nlp data = out.saida tech=tr cov=2 vardef=n pcov noprint
87 outest=out.emvsgama(keep=gama1 gama2 _type_ where=
88                       (_type_ in("PARMS","COV2:H")));
89
90 max l;
91
92     parms
93     gama1 = 1.5,
94     gama2 = 2.5;
95     beta1 = 1;
96     beta2 = 2;
97     beta3 = 3;
98
99     y          =      col1;
100    x1         =      col2;
101    s1         =      col3;
102    s2         =      col4;
103    mu         =      beta1*exp(-exp(beta2-beta3*x1));
104    phi        =      exp(gama1*s1+gama2*s2);
105    teta       =      -1/(2*mu**2);
106    bedeteta  =      -(-2*teta)**.5;
107    cdey      =      -1/(2*y);
108    L         =      phi*(y*teta-bedeteta+cdey)+.5*log(phi);
109
110 run;

```

```

106     data out.estimatbeta(drop=_type_);
107     set out.emvsbeta; if _type_ ne "PARMS" then delete;
108     run;
109     data out.estimatgama(drop=_type_);
110     set out.emvsgama; if _type_ ne "PARMS" then delete;
111     run;
112     proc iml;
113 /* Chamando os Arquivos das Matrizes que Serao Usadas */
114     use out.estimatbeta;
115     read all into betas;
116     use out.estimatgama;
117     read all into gamas;
118     use out.saida;
119     read all into dados;
120 /* Nomeando as Variaveis e Matrizes */
121     y = dados[,1];
122     x1 = dados[,2];
123     s1 = dados[,3];
124     s2 = dados[,4];
125     matx = x1;
126     mats = s1 || s2;
127     beta1 = betas[1];
128     beta2 = betas[2];
129     beta3 = betas[3];
130     g1 = gamas[1];
131     g2 = gamas[2];
132     beta = beta1//beta2//beta3;
133     gama = g1//g2;
134     tbeta = t(beta);
135     tgama = t(gama);
136     p=3; q=2;
137     dimensao = p**2;
138     dimensaoq = q**2;
139     sgama = mats*gama;
140     expsgama = exp(sgama);
141 /* Nomeando c1 e c2 para Facilitar o Produto de Matrizes */
142     c1 = exp(beta2-beta3#x1);
143     c2 = exp(-exp(beta2-beta3#x1));
144     phi = expsgama;
145     mu = beta1#c2;

```



```

146         wdebeta           =      1/mu###3;
147         wfi                =      diag(phi#wdebeta);
148         pesobeta          =      phi#wdebeta;
149     /*      Obtencao da Matriz das Primeiras Derivadas de Eta      */
150         xtil1              =      c2;
151         xtil2              =      -beta1#c1#c2;
152         xtil3              =      x1#beta1#c1#c2;
153         xtil               =      xtil1 || xtil2 || xtil3;
154         xtiltwfi           =      t(xtil)*wfi;
155         kdebeta            =      xtiltwfi*xtil;
156         detkdebeta         =      det(kdebeta);
157         invkdebeta         =      inv(kdebeta);
158     /*      Termino do Calculo da Matriz Inversa de Fisher para Beta      */
159         zdebeta            =      (xtil*invkdebeta*t(xtil));
160         diagzdebeta        =      diag(zdebeta);
161     /*      Nomeando produtoc1c2 e produtoc1c2qua para Facilitar o Calculo
162           da Matriz das Segundas Derivadas de Eta      */
163         produtoc1c2        =      c1#c2;
164         produtoc1c2qua     =      produtoc1c2#c1;
165     /*      Obtendo a Matriz das Segundas Derivadas de Eta      */
166         x2til1             =      j(&m,1,0);
167         x2til2             =      (-1)#produtoc1c2;
168         x2til3             =      x1#produtoc1c2;
169         x2til4             =      beta1#(-produtoc1c2+produtoc1c2qua);
170         x2til5             =      x1#beta1#(produtoc1c2-produtoc1c2qua);
171         x2til6             =      x1##2#beta1#(-produtoc1c2+produtoc1c2qua);
172     /*      Obtendo o çTrao da Matriz Xtiltil      */
173         traco              =      x2til1#invkdebeta[1,1]+2# x2til2#
174                               invkdebeta[1,2]+2# x2til3#invkdebeta[1,3]+
175                               x2til4#invkdebeta[2,2]+2# x2til5#
176                               invkdebeta[2,3]+x2til6#invkdebeta[3,3];
177         quisi2             =      -.5#traco;
178     /*      Obtendo o Vies de Beta      */
179         biasdebeta         =      invkdebeta*xtiltwfi*quisi2;
180     /*      Obtendo Beta Corrigido      */
181         betacorrigido      =      beta - biasdebeta;
182         tbiasdebeta        =      t(biasdebeta);
183         tbetacorrigido     =      t(betacorrigido);
184     /*      Inicio do Calculo do Vies de Gama      */
185         fica               =      phi##2;

```

```

186         ficubo           =         phi###3;
187         d3               =         1/ficubo;
188         d2               =         -.5/fiqua;
189 /*  Obtendo a Matriz das Primeiras Derivadas de Tau      */
190         stil1            =         expsgama#s1;
191         stil2            =         expsgama#s2;
192         stil             =         stil1 || stil2;
193 /*  Inicio do Calculo da Matriz Inversa de Fisher para Gama*/
194         wdagama          =         -d2;
195         stiltwdagama     =         t(stil)*diag(wdagama);
196         kdegama          =         stiltwdagama*stil;
197         invkdegama       =         inv(kdegama);
198 /*  Termino do Calculo da Matriz Inversa de Fisher para Gama */
199         zdegama          =         diag(stil*invkdegama*t(stil));
200         pisi1            =         .5*inv(diag(wdagama))*(zdegama*
201                                     (diag(d3))+diagzdebeta*diag(wdebeta));
202 /*  Obtendo a Matriz das Segundas Derivadas de Tau      */
203         s2til1           =         s1##2#expsgama;
204         s2til2           =         s1#s2#expsgama;
205         s2til3           =         s2##2#expsgama;
206 /*  Obtendo o çTrao da Matriz Stiltil                    */
207         tracogama        =         s2til1#invkdegama[1,1]+2#s2til2#
208                                     invkdegama[1,2]+s2til3#invkdegama[2,2];
209         um                =         j(&m,1,1);
210         pisi2            =         -.5*diag(tracogama);
211         pisi             =         (pisi1+pisi2)*um;
212 /*  Obtendo o Vies de Gama                               */
213         biasdegama        =         invkdegama*stiltwdagama*pisi;
214         gamacorrigido     =         gama- biasdegama;
215         tbiasdegama       =         t(biasdegama);
216         tgamacorrigido    =         t(gamacorrigido);
217         varibeta          =         vecdiag(invkddebeta);
218         erropadbeta       =         t(varibeta##.5);
219         varigama          =         vecdiag(invkdegama);
220         erropadgama=t(varigama##.5);
221 nomes                    =         {"beta1" "beta2" "beta3" "gama1" "gama2" "vbeta1"
222                                     "vbeta2" "vbeta3" "vg1" "vg2" "cbeta1" "cbeta2"
223                                     "cbeta3" "cg1" "cg2" "stddevb1" "stddevb2"
224                                     "stddevb3" "stddevg1" "stddevg2"};
225 dados1                   =         tbeta || tgama || tbiasdebeta || tbiasdegama ||

```

```
226             tbeta corrigido || tgamma corrigido || erro pad beta ||
227             erro pad gama ;
228     create out.saidapar from dados1[colname=nomes];
229     append from dados1;
230     quit;
231 proc append base=out.estimativas data=out.saidapar;
232     run;
233         data out.est;
234         set out.estimativas ;
235         if beta1 = 0 then delete;
236 run;
237 %put &i;
238 %end;
239 %mend geracao;
240 %geracao(quantidade=10000,m=60);
241 proc means data = out.est; run;
```
