UNIVERSIDADE ESTADUAL DE CAMPINAS - UNICAMP
Instituto de Matemática, Estatística e Ciência da Computação - IMECC

MODELOS DE DOSE-RESPOSTA NA AVALIAÇÃO DE RISCOS ASSOCIADOS A AGENTES TERATOGÊNICOS

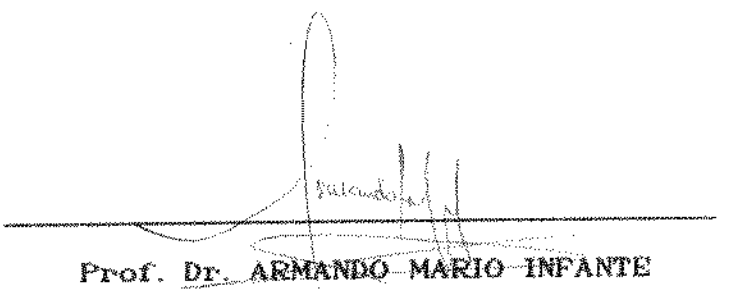
NÁDIA REGINA MASCARENHAS ROCHA

PROF. DR. ARMANDO MARIO INFANTE
Orientador

CAMPINAS - SÃO PAULO
1992

Este exemplar corresponde à redação final da tese devidamente corrigida e defendida por NADIA REGINA MASCARENHAS ROCHA e aprovada pela Comissão Julgadora.

Campinas, 28 de setembro de 1992.



Prof. Dr. ARMANDO MARIO INFANTE

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

Aos meus pais.

A G R A D E C I M E N T O S

À minha família, pelo apoio incondicional às minhas decisões profissionais, mesmo quando tais decisões significaram um recomeço ou a convivência com a saudade.

Ao meu orientador Armando Mario Infante, pelos grandes ensinamentos e pelo incentivo nas horas difíceis.

Ao professor Lúcio Tunes dos Santos, pela boa vontade e, principalmente, pela ajuda na parte computacional deste trabalho, essencial para a conclusão do mesmo. À Oscar Humberto Bustos, pelas valiosas sugestões apresentadas.

Ao Alvaro Vigo, pela força e carinho constantes.

Aos velhos amigos, pela torcida. Aos amigos do mestrado, pelo companheirismo que ajudou a superar as inúmeras dificuldades, e pela alegria de se fazer novas amizades.

Aos professores da UFMG e da UNICAMP que contribuíram para a minha formação estatística.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de mestrado por 30 meses. Ao Fundo de Apoio ao Ensino e à Pesquisa (FAEP/UNICAMP), pela contribuição para a finalização desta tese, através do auxílio-ponte nº 0210/92.

All models are wrong, but some are useful

G. E. P. BOX

R E S U M O

Neste trabalho são discutidos e comparados modelos de dose-resposta para a análise de dados binários gerados em experimentos teratológicos. Primeiramente são considerados modelos que generalizam o estudado por RAI & VAN RYZIN (1985), onde é incorporado o tamanho das ninhadas geradas no experimento, numa tentativa de considerar a presença do efeito de ninhada. Em particular, são considerados os modelos que supõem as distribuições de Poisson e binomial negativa para o tamanho da ninhada. Em segundo lugar, são considerados modelos logísticos lineares como os sugeridos por WILLIAMS (1987), onde o tamanho da ninhada e a dose podem atuar como covariáveis. O método de máxima verossimilhança é aplicado para estimar os parâmetros de todos os modelos. No caso do modelo geral de Rai & Van Ryzin, são demonstradas propriedades assintóticas dos estimadores, completando a prova sugerida por RAI & VAN RYZIN (1981). Dados provenientes de dois experimentos teratológicos apresentados na literatura são utilizados no ajuste dos modelos: os contidos em LÜNING et al (1966) e os de TYL et al (1983). Os métodos numéricos iterativos de Fletcher & Reeves, quase-Newton e de Newton são aplicados no ajustamento do modelo geral de Rai & Van Ryzin. Para os conjuntos de dados utilizados, os métodos produzem resultados concordantes, exceto para o modelo binomial negativo. Os modelos logísticos são ajustados através do pacote estatístico GLIM e, com os dois conjuntos de dados, os ajustes obtidos para os modelos com variação binomial pura não são bons. Porém, a suposição de variação extra-binomial produz uma melhora na qualidade do ajustamento para os dados de Tyl et al (1983)

A B S T R A C T

Dose-response models for the analysis of binary data from teratologic experiments are discussed and compared. Firstly, are considered models generalizing that studied by RAI & VAN RYZIN (1985), which incorporate the litter sizes as an attempt to account for litter effects. The models considered assume a Poisson or a negative binomial distribution for the litter sizes. Secondly, logistic linear models like those suggested by WILLIAMS (1987) are considered, which contain litter sizes and doses as covariables. The maximum likelihood method is applied to estimate model's parameters. For the Rai & Van Ryzin model, assymptotic properties of the estimators are demonstrated, completing thus the proof suggested by RAI & VAN RYZIN (1981). The models are fitted to data from two teratologic experiments: the one discussed by LÜNING et al (1966) and the experiment of TYL et al (1983). The iteractive Fletcher & Reeves, quasi-Newton and Newton methods are applied to the fitting of Rai & Van Ryzin's model. The results obtained are in general agreement for the models, excepting the negative binomial part. The logistic models are fitted using the software GLIM. Pure binomial variation models do not display a good fitting, but models incorporating extra binomial variation show a better fit for Tyl et al data.

S U M Á R I O

CAPÍTULO 1

INTRODUÇÃO

1.1 AGENTES TERATOGÊNICOS

No ambiente onde o homem moderno vive, estão presentes diversos agentes que produzem efeitos adversos à saúde. Em particular, eles podem afetar a capacidade reprodutiva de indivíduos expostos e gerar malformações congênitas em seus filhos.

A indução de defeitos na descendência de um indivíduo submetido a um agente exógeno é estudada pela Teratologia, o processo que leva à estes defeitos é denominado teratogênese, enquanto que o agente capaz de causá-los é dito teratogênico.

No feto, vários resultados adversos podem ser causados por agentes teratogênicos, incluindo malformações, retardamento do crescimento, desordens funcionais e eventualmente sua morte. Estes resultados dependem principalmente da susceptibilidade do feto no período da exposição, governada por fatores tais como seu genótipo,

estágio de desenvolvimento durante a exposição, natureza do agente, duração, intensidade e via de exposição. Crianças nascidas fisicamente perfeitas também podem apresentar alguns resultados adversos mais tarde, como por exemplo, distúrbios neurológicos.

Para um agente potencialmente teratogênico, pode-se avaliar o risco de que pessoas submetidas à diferentes níveis de exposição ao agente tenham suas capacidades reprodutivas reduzidas ou gerem filhos disformes. Nesta tese, discutimos uma forma de avaliar este risco, baseada no ajuste de curvas de dose-resposta. Porém, antes de apresentar com mais detalhe o processo de avaliação, é conveniente exibir alguns exemplos de agentes teratogênicos conhecidos.

A listagem de possíveis agentes teratogênicos inclui os do tipo físico, como as radiações ionizantes, os químicos e os biológicos. Exemplos de agentes com efeitos conhecidos são os seguintes.

i) A talidomida, um sedativo usado na Europa e nos Estados Unidos no final dos anos 50 e início dos 60. A administração de talidomida a mulheres grávidas produziu deformidades nos músculos e esqueleto dos filhos, afetando principalmente membros e face dos mesmos.

ii) O mercúrio, responsável pelo desastre ocorrido na Baía de Minamata, no Japão, onde no período de 1954 a 1960, a população consumiu peixe contaminado com metilmercúrio proveniente de uma usina que utilizava o mercúrio, jogando os resíduos do processo na baía. Crianças filhas de mães expostas tiveram vários níveis de sintomas neurológicos semelhantes à paralisia cerebral.

iii) O vírus da rubéola, cuja apresentação durante a gravidez materna está associada ao nascimento de crianças anormais. Alguns defeitos de nascimento apresentados são catarata, microcefalia, retardamento mental, cegueira, surdez e lesões cardíacas.

iv) As bifenilas policloradas ("polychlorinated biphenyls", ou

PCB em inglês). No sul do Japão, em 1968, um grande número de pessoas foi intoxicada pela ingestão de óleo de arroz contaminado por PCB. Algumas crianças filhas de mães intoxicadas nasceram mortas e outras apresentaram retardamento de crescimento, pigmentação anômala e prejuízos neurológicos.

Veja WILSON & FRASER (1977); LONGO (1980); KURZEL & CETRULO (1981) e SHANE (1989) para maiores informações sobre a relação entre agentes teratogênicos e resultados adversos no feto e também para detalhes suplementares sobre os exemplos apresentados.

1.2 AVALIAÇÃO DE RISCOS TERATOGÊNICOS

A presença de agentes possivelmente teratogênicos no ambiente faz com que seja necessário determinar cuidadosamente seus efeitos, bem como avaliá-los para diferentes níveis de exposição. Um procedimento consiste na avaliação quantitativa do risco associado aos diferentes níveis de exposição ao agente em estudo. Seu objetivo é estimar as probabilidades de que diferentes doses do agente resultem em efeitos adversos. VAN RYZIN (1980) define a avaliação quantitativa do risco como a estimação de níveis de exposição a uma substância tóxica que levam a aumentos especificados nas taxas de incidência ou na ocorrência de uma consequência adversa predeterminada. Esta avaliação permite eventualmente determinar "níveis aceitáveis" de exposição.

Segundo VAN RYZIN (1980), os dados utilizados para estimar estes riscos, geralmente são provenientes de:

i) Testes mutagênicos , que são testes rápidos, simples e não muito caros.

ii) Estudos epidemiológicos feitos em populações humanas, que ajudam a determinar associações entre possíveis agentes teratogênicos e efeitos reprodutivos adversos.

iii) Experimentos com animais que envolvem a avaliação controlada da teratogenicidade dos agentes em animais de laboratório.

A seguir, faremos uma breve apresentação dos estudos epidemiológicos e dos experimentos com animais. Não nos deteremos nos testes mutagênicos, remetendo o leitor a SHANE (1989) para uma breve discussão dos mesmos.

1.2.1 ESTUDOS EPIDEMIOLÓGICOS

A Epidemiologia pode ser definida como o estudo da distribuição e dos determinantes relacionados aos estados de saúde e eventos em populações, e a aplicação deste estudo no controle de problemas de saúde. Veja LAST (1988).

Assim, a Epidemiologia pode ser utilizada como uma metodologia para exploração de informações sobre os efeitos de agentes teratogênicos quando estudamos, por exemplo, a ocorrência de nascimentos disformes em populações humanas em um período determinado e em uma área geográfica fixa.

Nos estudos epidemiológicos são geralmente utilizadas duas técnicas, que denominaremos descritiva e analítica. Nos estudos descritivos são obtidas informações a respeito da distribuição e frequência de um resultado reprodutivo adverso. Já o estudo analítico, é planejado para testar ou gerar hipóteses a respeito da associação entre exposição e resultados reprodutivos adversos.

Em particular, duas estratégias utilizadas na pesquisa epidemiológica são os estudos de casos e controles e os estudos de coortes. No estudo de casos e controles, é feita uma avaliação retrospectiva dos fatores de exposição nos casos (que são indivíduos com a característica adversa de interesse) e controles (indivíduos sem essa característica). No estudo de coorte, os indivíduos expostos e os não expostos a um determinado fator são acompanhados por um período de tempo, usando técnicas retrospectivas ou prospectivas para observar a presença do resultado adverso.

A principal vantagem dos estudos epidemiológicos é que

seus resultados são diretamente aplicáveis para populações humanas. Lamentavelmente, eles possuem muitas desvantagens: dificuldades logísticas e éticas; ausência de documentação confiável sobre níveis de exposição individual; presença de fatores de confundimento (que são aqueles fatores que variam juntamente com as variáveis explicitamente consideradas e que por isso geram dificuldades na interpretação dos resultados); custo e tempo necessários para sua realização e, finalmente, a inexistência de dados referentes a novos agentes.

Uma discussão global e detalhada das diversas estratégias da pesquisa epidemiológica pode ser encontrada em FRIEDMAN (1987) e em KLEINBAUM et al (1982). Para uma breve apresentação dos estudos epidemiológicos na área de reprodução, veja SHANE (1989). Em BROWN (1983); BROWN & KOZIOL (1983) e VAN RYZIN (1980) são discutidas as vantagens e desvantagens dos estudos epidemiológicos na avaliação de riscos.

Numa tentativa de contornar as limitações dos estudos epidemiológicos são feitos experimentos em condições controladas. Porém, quando existem dados epidemiológicos confiáveis, eles devem ser usados para completar, comparar e/ou modificar os resultados dos estudos em animais.

1.2.2 EXPERIMENTOS COM ANIMAIS

O estudo dos efeitos de um agente teratogênico também pode ser realizado mediante experimentos com animais. Neles, o investigador intencionalmente altera um ou mais fatores sob condições controladas e estuda o efeito desta alteração na resposta dos animais. Este tipo de experimento é realizado com animais homogêneos, registrando em detalhe fatores tais como espécie, raça, idade e sexo, diferenças de alojamento, alimentação, via de administração do agente, dosagem, frequência da exposição e duração dos tratamentos.

Algumas vantagens destes experimentos são o controle que se tem sobre as condições experimentais, sua relativa estabilidade, a melhor precisão na medição dos níveis de exposição e das respostas

obtidas dos animais, seu menor custo e sua maior rapidez em comparação aos estudos epidemiológicos.

O principal ganho que se tem ao planejar um experimento é a possibilidade de eliminar fontes potenciais de erro sistemático, avaliando e reduzindo a variabilidade experimental. Para isto, utilizam-se os princípios do Planejamento de Experimentos propostos por R. A. Fisher: aleatorização, blocagem e replicação. Veja por exemplo, GART et al (1986) para considerações a respeito destes princípios.

Em um típico experimento teratológico com animais, um determinado número de fêmeas grávidas é aleatoriamente dividido em diferentes grupos. As fêmeas de cada grupo recebem uma dose do agente em estudo, de maneira que os grupos são tratados com doses crescentes, mantendo-se um grupo controle que nada recebe. É também possível que os machos sejam tratados e depois acasalados com as fêmeas e, neste caso, o estudo recebe o nome de estudo letal dominante. Em ambos os casos, as fêmeas grávidas sofrem a exposição (direta ou indireta) ao agente, registrando-se as respostas na sua descendência. Assim, a informação a respeito do efeito produzido por uma dada dose é extraída da viabilidade dos ovos fertilizados, da morte ou sobrevivência do feto, ou da completude de suas características. Exemplos de experimentos teratológicos podem ser encontrados em WEIL (1970); PAUL (1982); no Capítulo 2 desta tese e no Capítulo 4, onde serão descritos os dois experimentos que geraram os dados que serão utilizados.

Além das vantagens comuns aos experimentos com animais, os experimentos teratológicos apresentam outros benefícios, pois nesta situação muitos fetos podem ser examinados em pouco tempo a um custo relativamente baixo, existindo também a possibilidade de planejar o experimento para fazer coincidir as exposições com períodos de alta sensibilidade dos animais expostos.

Os experimentos com animais têm também dificuldades conceituais. As maiores são: (a) a necessidade de estender os resultados da espécie animal ao homem e (b) a extrapolação dos resultados obtidos nos níveis de exposição experimentais aos níveis

esperados em uma exposição humana. Faremos, a seguir, uma breve apresentação destes dois problemas.

1.2.2.1 EXTRAPOLAÇÃO DO ANIMAL PARA O HOMEM

A questão da conversão biológica dos resultados de um experimento com animais recebe, às vezes, o nome de extrapolação do animal para o homem.

Esta extrapolação tem como apelo intuitivo o fato de que se uma substância causa efeitos adversos nos animais do experimento, é provável que ela cause estes mesmos efeitos nos humanos. Na verdade, ela deve se basear em considerações de similaridade entre espécies no que diz respeito, por exemplo, à absorção, distribuição, metabolismo e excreção do agente em estudo.

Nesta discussão não nos deteremos neste problema, salientando que existe controvérsia em torno deste assunto, como pode ser visto em FREEDMAN & ZEISEL (1988).

1.2.2.2 CURVAS DE DOSE-RESPOSTA

A relação de dose-resposta é definida por LAST (1988), como aquela na qual uma mudança na quantidade, intensidade ou duração da exposição a um agente está associada com uma alteração no risco de um resultado especificado. Este resultado é medido através do registro de uma resposta.

Nos experimentos teratológicos a resposta de interesse geralmente é um efeito adverso bem definido, cuja presença é determinada em cada um dos filhotes gerados no experimento.

A resposta analisada pode ser quantitativa, ou quantal. No primeiro caso, ela é expressa mediante medidas contínuas ou contagens, como por exemplo o peso fetal ou o número de implantações observadas. A resposta é quantal quando ela é expressa em categorias. Em especial, ela pode ser dicotômica (ou binária), quando assume apenas

dois resultados possíveis, tais como morte ou sobrevivência do feto, ou presença ou ausência de uma determinada malformação. Nesta tese estudaremos o caso particular de experimentos com respostas quantais dicotômicas.

Para precisar a estrutura do experimento, considere n animais aleatoriamente divididos em m+1 grupos, sendo um deles o grupo controle formado por animais que não recebem nenhuma dose do agente. O i-ésimo grupo é composto por n_i animais que recebem a dose d_i do agente em estudo, onde $0 = d_0 < d_1 < \ldots < d_m$. Considere ainda que para cada nível d_i é registrada a resposta quantal de cada animal do i-ésimo grupo, $i = 0, 1, \ldots, m$.

Se P(d) é a probabilidade de que um animal submetido à dose d do agente apresente a resposta adversa de interesse, ela é uma função geralmente crescente de d. Nosso interesse é descrever P(d) matematicamente. Para tanto, se $f(d, \underset{\sim}{\theta})$ é uma função suposta conhecida, com valores entre 0 e 1 (geralmente não decrescente em d e dependente de um vetor de parâmetros $\underset{\sim}{\theta}$ desconhecido), podemos escrever

$$P(d) = f(d, \underset{\sim}{\theta}).$$

A curva de dose-resposta é obtida representando a probabilidade de resposta contra os correspondentes níveis de dose, para $\underset{\sim}{\theta}$ fixo. Na Figura 1.1 apresentamos uma típica curva de dose-resposta. Entretanto, as curvas de dose-resposta não são necessariamente sigmoidais como nesta figura, sendo possíveis outros formatos. Veja, por exemplo, GART et al (1986) p. 110.

O vetor de parâmetros $\underset{\sim}{\theta}$ da função matemática que descreve a curva de dose-resposta, deve ser estimado através dos dados observados no experimento, que levam assim, a uma estimativa $\hat{\underset{\sim}{\theta}}$. Com a curva estimada $\hat{P}(d) = f(d, \hat{\underset{\sim}{\theta}})$, podemos também estimar a probabilidade de resposta para uma dada dose. Depois da estimação, também é possível resolver o problema inverso, que consiste em determinar a dose correspondente a uma dada probabilidade de resposta.

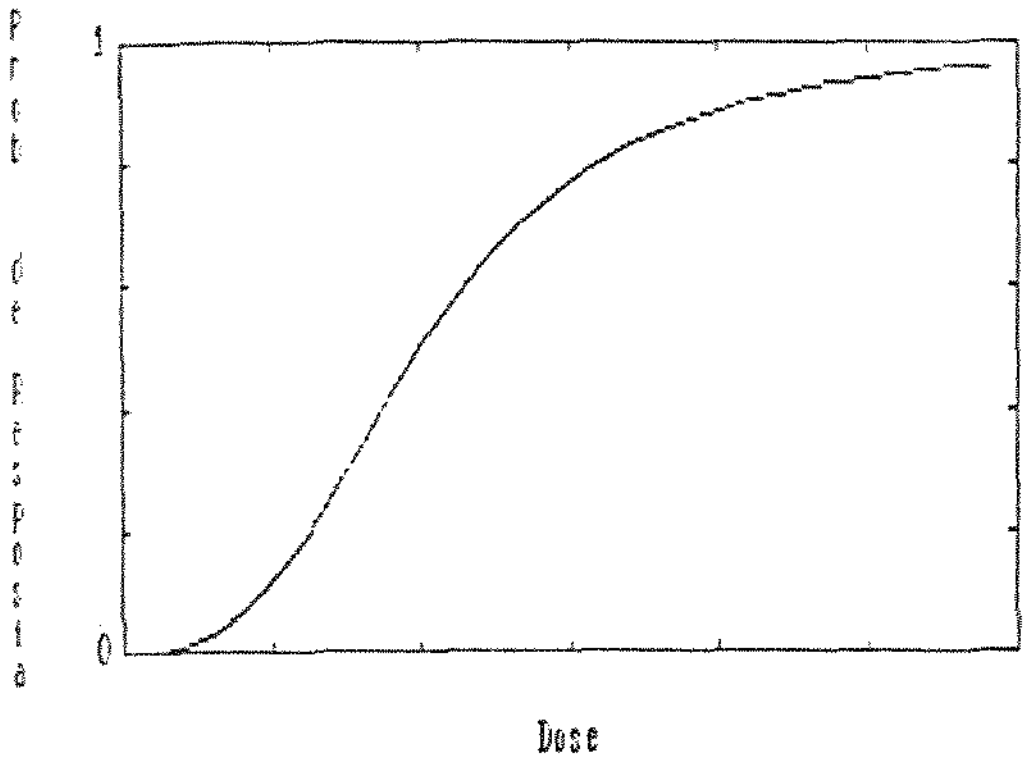


Figura 1.1: Curva de dose-resposta $P(d) = f(d, \underset{\sim}{\theta})$ obtida para um valor fixo do parâmetro $\underset{\sim}{\theta}$.

1.2.2.3 EXTRAPOLAÇÃO PARA DOSES BAIXAS

Numa situação comum de exposição humana, o nível da dose envolvida é muito baixo. Se o experimento com animais fosse conduzido nestes mesmos níveis, a proporção de respostas adversas observadas no experimento seria muito pequena. Assim, para que efeitos reduzidos fossem adequadamente detectados, seria necessário um número muito grande de animais, o que tornaria difícil e cara a execução do experimento. Estas afirmações serão exemplificadas a seguir com um raciocínio semelhante ao seguido em WARE & LOUIS (1983).

Considere que uma população homogênea de animais será utilizada em um experimento planejado para determinar se uma dada substância é teratogênica. Para simplificar, considere que a resposta é a presença ou ausência de alguma malformação em cada animal e que a proporção de animais com a resposta é 0 e 50/10000, nos grupos controle

e de tratados, respectivamente. Nosso interesse é determinar qual o número necessário de animais no grupo de tratados, para que com alta probabilidade, a diferença entre as proporções acima possa ser detectada. Uma situação hipotética como esta dificilmente ocorre na prática; uma situação mais realista, na qual a proporção de animais com a resposta adversa no grupo controle é não nula, requer a formação dos dois grupos, aumentando assim o número de animais no experimento.

Seja, em geral, π a proporção populacional de animais que apresentam a resposta adversa e p uma estimativa de π, dada pela proporção de animais com a resposta adversa no grupo de tratados. Pelo Teorema Central do Limite,a variável aleatória padronizada $Z = \dfrac{p-\pi}{\sqrt{\dfrac{\pi(1-\pi)}{n}}}$ é aproximadamente distribuída segundo uma distribuição normal com média 0 e variância 1. Considere o teste da hipótese H_0: $\pi = 0$ que rejeita a hipótese nula se $p > 0$. Se sua potência (isto é, a probabilidade de rejeição sob uma hipótese alternativa H_1) é superior a $1-\beta$, temos

$$P_{H_1}(p > 0) = P_{H_1}\left(Z > \frac{-\pi}{\sqrt{\dfrac{\pi(1-\pi)}{n}}} \right) \geq 1-\beta.$$

Portanto, $Z_\beta \geq \dfrac{-\pi}{\sqrt{\dfrac{\pi(1-\pi)}{n}}}$ e o tamanho do grupo deve satisfazer a relação $n \geq (Z_\beta)^2 \dfrac{(1-\pi)}{\pi}$, onde Z_β é o percentil de ordem $100\beta\%$ da distribuição normal padrão.

Note que neste caso particular, o tamanho do grupo de tratados independe do nível de significância do teste.

Na Tabela 1.1 apresentamos o tamanho do grupo de tratados para diferentes valores da potência do teste, no caso particular $\pi = 0.005$.

Tabela 1.1: Tamanho do grupo de tratados em função da probabilidade de detecção P para controle de resposta espontânea nula e tratados com probabilidade de resposta 0.005.

	P (%)			
	99	95	90	80
Tamanho (n)	1081	539	327	141

Numa tentativa de contornar o problema do grande número de animais necessários em um experimento, este é conduzido em níveis de exposição altos o bastante para mostrar resultados positivos. Assim, a curva de dose-resposta e, consequentemente, o risco, são estimados através de dados provenientes de níveis de exposição consideráveis. Entretanto, como o objetivo do experimento é fornecer uma estimativa do risco associado a baixos níveis de exposição, comuns à população humana, é necessário realizar uma extrapolação para doses baixas, utilizando a curva de dose-resposta estimada com as doses maiores.

A principal dificuldade deste procedimento é que o risco estimado nas doses baixas é fortemente dependente da função matemática escolhida para descrever a curva de dose-resposta.

Os diversos modelos de dose-resposta utilizados não diferem muito quanto a qualidade com que ajustam os dados, sendo muitas vezes difícil escolher o mais adequado dos modelos. No entanto, quando a extrapolação é feita, obtém-se riscos estimados bastante diferentes. Para discussões e exemplos de ajustes de curvas de dose-resposta e comparações dos riscos estimados nas doses baixas associados aos diferentes modelos, veja VAN RYZIN (1980); KREWSKI & VAN RYZIN (1981); BROWN (1983) e FREEDMAN & ZEISEL (1988).

Até agora, foram apresentadas questões gerais relacionadas com a avaliação quantitativa de riscos teratogênicos. Nos capítulos seguintes, serão discutidas algumas características dos dados

provenientes de experimentos teratológicos, estudando depois dois tipos de modelos de dose-resposta. No que segue, será detalhada a estrutura deste trabalho.

1.3 ESTRUTURA DA TESE

Nos experimentos teratológicos, as respostas de interesse são observadas na descendência dos animais expostos ao agente em estudo. Uma característica especial destas respostas é que elas são geradas por animais relativamente homogêneos, que constituem ninhadas. Os fenômenos provocados pela existência de ninhadas serão analisados no Capítulo 2. Primeiramente, discutiremos a presença dos efeitos de ninhadas e algumas de suas consequências para a análise estatística. Nos modelos estudados nesta tese, estes efeitos são quase exclusivamente considerados mediante a introdução de informação sobre o tamanho da ninhada produzida por cada fêmea que participa do experimento. Assim, neste capítulo também é discutida a modelagem estatística do tamanho da ninhada, considerando que o mesmo pode ser representado mediante uma variável aleatória discreta que segue a distribuição de Poisson ou, alternativamente, a distribuição binomial negativa. Estes dois modelos serão depois incorporados à discussão de um modelo geral, e testados em conjuntos de dados da literatura.

No Capítulo 3 será estudado o modelo proposto por RAI & VAN RYZIN (1985). Numa tentativa de considerar o efeito de ninhada, este modelo incorpora o tamanho da ninhada, considerando-o representado por uma variável aleatória. Assim, a quantidade de respostas adversas é suposta binomialmente distribuída, condicional à observação do tamanho da ninhada produzida por cada fêmea. No caso especial considerado por RAI & VAN RYZIN (1985), esta variável tem uma distribuição de Poisson cuja intensidade depende exponencialmente da dose aplicada. Neste Capítulo, é proposta ainda uma modificação no modelo original de Rai & Van Ryzin, considerando que o tamanho da ninhada segue uma distribuição binomial negativa. Serão também calculadas as probabilidades incondicionais de resposta adversa e por último, serão demonstradas algumas propriedades assintóticas dos estimadores de máxima verossimilhança do modelo condicional de Rai & Van Ryzin, completando

assim uma demonstração apresentada em RAI & VAN RYZIN (1981).

No Capítulo 4, são apresentados os procedimentos numéricos de Fletcher & Reeves, de Newton e quase-Newton, utilizados para estimar os parâmetros dos modelos mediante máxima verossimilhança, descrevendo-se também as dificuldades numéricas encontradas. Em particular, serão apresentadas as estimativas dos parâmetros dos modelos e as probabilidades condicional e incondicional de resposta adversa para dois conjuntos de dados provenientes da literatura: o exibido por LÜNING et al (1966) e o exposto no estudo de TYL et al (1983). Estes conjuntos de dados têm características diferentes. O primeiro, provém de um estudo letal dominante, onde o total de fêmeas é extremamente elevado e apenas as respostas dos filhotes de fêmeas que tiveram de 5 a 10 implantes foram registradas. O segundo conjunto de dados foi gerado em um experimento teratológico onde o total de fêmeas é mais reduzido e as observações não foram selecionadas, registrando-se as respostas para os filhotes de fêmeas que produziram quaisquer tamanhos de ninhada.

No Capítulo 5, é explorado o argumento apresentado por WILLIAMS (1987). Diversos modelos logísticos lineares que utilizam o tamanho da ninhada e a dose como covariáveis são ajustados e comparados. Como eles fazem parte da classe dos modelos lineares generalizados, podem ser ajustados utilizando o pacote estatístico GLIM, desenvolvido especialmente para estes modelos. É também avaliada a qualidade do ajustamento destes modelos utilizando instrumentos gráficos de diagnóstico. A variação extra-binomial, frequente em dados de ninhada é também incorporada aos modelos utilizando o procedimento de estimação de WILLIAMS (1982) e a qualidade dos ajustamentos reavaliada. Por último, são comparados os modelos com variação binomial pura e extra-binomial.

No Capítulo 6, são feitas as considerações finais sobre o trabalho.

No Apêndice A estão contidos os lemas e definições utilizados principalmente nas demonstrações do Capítulo 3. No Apêndice

B são apresentados os dois conjuntos originais de dados já citados, e no Apêndice C é feita uma descrição dos procedimentos computacionais para a estimação dos parâmetros (a) do modelo condicional de Rai & Van Ryzin e (b) dos modelos de Poisson e binomial negativo para o tamanho da ninhada. No Apêndice D encontram-se tabelas com o desempenho dos três procedimentos numéricos aplicados na estimação, além das estimativas dos parâmetros dos modelos logísticos, obtidas pelo pacote estatístico GLIM. Finalmente, no Apêndice E são apresentados os programas computacionais utilizados nesta tese.

CAPÍTULO 2

EFEITO DE NINHADA

Nos experimentos teratológicos, as respostas adversas são observadas em todos os filhotes de fêmeas expostas. Portanto, os dados gerados neste tipo de experimento são provenientes de ninhadas. Como as ninhadas são formadas por animais geneticamente homogêneos e na mesma fase de desenvolvimento, pode-se esperar que os animais de uma mesma ninhada apresentem características comuns. Elas serão discutidas a seguir.

2.1 FENÔMENOS GERADOS PELA PRESENÇA DE NINHADAS

Em geral, os animais de uma ninhada respondem de maneira mais similar do que animais de diferentes ninhadas. Na literatura, este fenômeno de homogeneidade é denominado efeito de ninhada, veja HASEMAN & KUPPER (1979). Desta maneira, podemos também dizer que o efeito de ninhada é a ocorrência de dependência entre as respostas produzidas por animais de uma mesma ninhada. Esta dependência deve ser incorporada à modelagem estatística das respostas. Por este motivo, a seguir faremos

uma breve apresentação de duas formas de modelagem.

Considere um experimento teratológico onde um total de n fêmeas é aleatoriamente dividido em m+1 grupos experimentais, sendo m grupos de tratamento e um grupo controle. O i-ésimo grupo é composto por n_i fêmeas que recebem a dose d_i do agente em estudo com $0 = d_0 < d_1 < \ldots < d_m$, onde $d_0 = 0$ é a dose correspondente ao grupo controle. É registrada a resposta quantal de interesse de todos os filhotes das fêmeas que compõem cada grupo e esta resposta é modelada utilizando as variáveis aleatórias X_{ijk}, onde

$$X_{ijk} = \begin{cases} 1 \text{ se o k-ésimo filhote da j-ésima fêmea do i-ésimo grupo} \\ \quad \text{experimental apresenta o resultado adverso de interesse} \\ 0 \text{ em caso contrário} \end{cases}$$

para $k = 1, 2, \ldots, s_{ij}$; $j = 1, 2, \ldots, n_i$ e $i = 0, 1, \ldots, m$, onde s_{ij} é o tamanho observado da j-ésima ninhada do i-ésimo grupo.

Seja $y_{ij} = \sum_{k=1}^{s_{ij}} x_{ijk}$ uma realização da variável aleatória Y_{ij}, que representa o número de filhotes da j-ésima fêmea do i-ésimo grupo, apresentando o resultado adverso. Seja P_{ij} a correspondente probabilidade de resposta. Dois modelos muito usados para a distribuição de Y_{ij} são o binomial e o de Poisson.

Sob o primeiro modelo, condicional ao tamanho da ninhada, Y_{ij} é suposta binomialmente distribuída com parâmetros s_{ij} e P_{ij}, onde P_{ij} é constante e igual a P_i para todo $j = 1, 2, \ldots, n_i$. Isto é, para um gupo fixado, a probabilidade de que filhotes dentro deste grupo apresentem a resposta é a mesma.

No segundo modelo, Y_{ij} segue uma distribuição de Poisson com parâmetro λ_{ij} constante e igual a λ_i para todo $j = 1, 2, \ldots, n_i$, assumindo-se, portanto, que a intensidade da distribuição é a mesma

para todos os animais dentro de um particular grupo de tratamento. Assim, o número esperado $E(Y_{i\,j})$ de animais com a resposta adversa é o mesmo para todas as ninhadas dentro de um grupo. Note que o modelo de Poisson, ao contrário do binomial, não leva o tamanho da ninhada em consideração.

Os modelos binomial e de Poisson são geralmente inadequados para modelar a morte fetal em experimentos teratológicos, veja por exemplo, HASEMAN & SOARES (1976). Esta inadequação é devida a sobre-dispersão, frequente em organismos agregados e que é a presença de maior variabilidade do que seria prevista com base no modelo escolhido. É também possível, apesar de menos comum, a ocorrência da sub-dispersão, que é a existência de menor variabilidade do que a esperada pelo modelo considerado. A sub-dispersão pode surgir, por exemplo, como resultado da competição entre os organismos.

Nos experimentos teratológicos, as respostas são obtidas a partir de cada filhote, cuja variabilidade é afetada pela variabilidade das ninhadas que eles integram. Por este motivo, geralmente a variabiliadade observada entre os filhotes é maior do que aquela esperada com base nos modelos binomial ou de Poisson. Quando estamos lidando com estes dois modelos, esta sobre-dispersão é dita variação extra-Binomial ou variação extra-Poisson, respectivamente.

Vários modelos têm sido propostos para a análise estatística de dados binários observados em ninhadas. Uma proposta para utilizar a distribuição binomial negativa na modelagem da morte embriônica foi feita por McCAUGHRAN & ARNOLD (1976). Nesta tese, vamos utilizar esta distribuição para modelar o tamanho da ninhada. Outra proposta é o modelo beta-binomial, veja WILLIAMS (1975). Neste modelo, assume-se que dentro de cada ninhada as respostas são variáveis aleatórias de Bernoulli com parâmetro que varia entre as ninhadas do mesmo grupo segundo a distribuição beta. O modelo logístico linear proposto por WILLIAMS (1982) é semelhante ao anterior, mas não postula nenhuma distribuição para o parâmetro da Bernoulli. No Capítulo 5, vamos ajustar este último modelo a dois conjuntos de dados. Outra proposta é o modelo binomial correlacionado, veja KUPPER & HASEMAN

(1978), onde é incorporada uma forma de correlação entre as respostas de uma mesma ninhada.

Estudos do comportamento de modelos que consideram o efeito de ninhada foram realizados por vários autores. Podemos mencionar HASEMAN & KUPPER (1979) para uma revisão de modelos, procedimentos de estimação e métodos de análise e PAUL (1982) para uma comparação entre modelos utilizando dados reais e simulados. Estudos de de simulação podem ser vistos também em KUPPER et al (1986) e MARUBINI et al (1988). Para trabalhos relativos à vícios de estimação, veja WILLIAMS (1988) e YAMAMOTO & YANAGIMOTO (1988).

Será visto nos Capítulos 3 e 5, que uma forma de modelagem que considera os efeitos de ninhada, consiste em incorporar a modelos simples a informação sobre os tamanhos observados das ninhadas. A seguir, consideraremos esta importante variável que é frequentemente avaliada em experimentos teratológicos.

2.2 TAMANHO DA NINHADA

O agente que será estudado através de um experimento teratológico pode ter o efeito de diminuir o número de fetos implantados no útero ou o número de filhotes nascidos vivos, aumentando possivelmente o número de filhotes nascidos com alguma malformação. Assim, uma importante variável a ser considerada neste tipo de estudo é o tamanho da ninhada. Para cada resposta adversa de interesse, deve ser feita uma medida apropriada para este tamanho. Se a resposta é o número de implantes mortos, o tamanho da ninhada é o total de implantes. Se a resposta é malformação, o tamanho da ninhada é o total de fetos nascidos vivos.

Na Tabela 2.1, apresentamos um conjunto de dados de um experimento teratológico realizado pela Profa. Nancy Airoldi Teixeira do Departamento de Farmacologia da Faculdade de Ciências Médicas da UNICAMP, para que o leitor tenha uma idéia dos tamanhos de ninhada típicos.

No experimento foram utilizadas ratas Wistar de três meses de idade, mantidas durante três dias para acasalamento com ratos Wistar de cinco meses. No dia inicial da gravidez, isolou-se as fêmeas, dividindo-as em dois grupos experimentais. No grupo denominado Lítio, as ratas foram tratadas com cloreto de lítio (LiCl) numa concentração de 10 mN diluídos em água de torneira como única fonte de bebida. No grupo de controle, as ratas receberam água de torneira como bebida, na mesma quantidade média de líquido consumido pelo grupo tratado.

Tabela 2.1: Tabela de frequência do tamanho das ninhadas produzidas por ratas Wistar em um experimento teratológico para avaliação do lítio.

Tamanho da Ninhada	Grupo	
	Controle	Lítio
2	0	1
3	2	0
4	1	0
5	0	3
6	0	2
7	4	2
8	7	1
9	7	10
10	10	11
11	5	4
12	8	7
13	0	1
14	0	1
15	1	1
16	1	0

O tamanho da j-ésima ninhada ($j = 1, \ldots, n_i$) do i-ésimo grupo ($i = 0, 1, \ldots, m$) de um experimento teratológico será denotado por s_{ij} e pode ser modelado através de uma variável aleatória S_{ij}. Como S_{ij} corresponde a uma contagem, a primeira tentativa de ajuste será a distribuição de Poisson.

Neste caso, se S_{ij} segue uma distribuição de Poisson com parâmetro λ_{ij} constante e igual a λ_i para todas as ninhadas do i-ésimo

grupo, e se este parâmetro depende da dose, ele pode ser modelado na forma

$$\lambda_{ij} = \lambda_i = \phi_1 \exp(-\phi_2 d_i)$$

onde ϕ_1 e ϕ_2 são parâmetros com $\phi_1 > 0$ e $-\infty < \phi_2 < \infty$. Assim, a média e a variância de S_{ij} coincidem entre si e com o valor de λ_i, de maneira que podemos interpretar o parâmetro da distribuição de Poisson como o tamanho esperado das ninhadas do i-ésimo grupo (i = 0, 1, ..., m). Note que ϕ_1 é o tamanho esperado da ninhada no grupo controle e ϕ_2 representa o efeito da dose no tamanho da ninhada. Fixados ϕ_1 e ϕ_2, se $\phi_2 > 0$, este tamanho esperado diminui à medida que a dose aumenta. Se $\phi_2 < 0$, o tamanho esperado cresce com o aumento da dose, enquanto que para $\phi_2 = 0$, a distribuição do tamanho da ninhada é a mesma para qualquer valor da dose.

Como indicam ROSS & PREECE (1985) e JACKSON (1972), um modelo um pouco mais complexo que o de Poisson, e que muitas vezes resulta mais adequado para o ajuste de dados de contagem, é baseado na distribuição binomial negativa. Uma grande utilidade desta distribuição reside no fato dela estar relacionada com um número de outras distribuições discretas, servindo como boa aproximação em diversas situações. Faremos, a seguir, uma breve apresentação desta distribuição.

Como é sabido, a distribuição binomial negativa pode ser obtida utilizando o número X de ensaios de Bernoulli necessários para a obtenção de exatamente r sucessos. Se p é a probabilidade de sucesso dos ensaios, a probabilidade de que r + s ensaios sejam necessários pode ser escrita como

$$P(X = r + s) = \binom{r + s - 1}{r - 1} p^r (1 - p)^s \qquad s = 0, 1, 2, \ldots$$

ROSS & PREECE (1985) apresentam várias parametrizações utilizadas para a distribuição binomial negativa. Nesta tese consideraremos aquela devida a ANSCOMBE (1949): a variável aleatória R

tem distribuição binomial negativa com parâmetros m e k positivos mas não necessariamente inteiros, se

$$P(R = r) = \binom{r+k-1}{k-1}\left(\frac{m}{m+k}\right)^r \left(1 + \frac{m}{k}\right)^{-k} \qquad r = 0, 1, 2, \ldots$$

onde o número combinatório $\binom{r+k-1}{k-1}$ é definido por $\dfrac{\Gamma(r+k)}{r!\,\Gamma(k)}$ para k arbitrário, e $\Gamma(x)$ denota a função gama, $\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}dt$. Utilizamos a notação $R \sim BN(m, k)$ para denotar que R segue a distribuição binomial negativa com parâmetros m e k.

A média e a variância da distribuição binomial negativa são dadas por

$$E(R) = m \quad e \quad Var(R) = m + m^2/k \ .$$

Assim, o parâmetro k pode ser interpretado como uma medida do afastamento da distribuição binomial negativa em relação a distribuição de Poisson, obtida como caso limite da binomial negativa para m fixo e $k \to \infty$, veja o Lema A9 no Apêndice A.

Vários são os mecanismos estocásticos que geram a distribuição binomial negativa. Entretanto, nosso interesse se encontra em três que julgamos razoáveis no caso do tamanho da ninhada.

i) Mistura de Poisson. Se uma variável aleatória $X|\lambda$ segue uma distribuição de Poisson com parâmetro λ e se este parâmetro está distribuído segundo uma distribuição gama com parâmetros (k, m), então a distribuição de probabilidade de X é uma binomial negativa (m, k).

ii) Considere colônias ou grupos de indivíduos distribuídos aleatoriamente em uma área ou no tempo, de maneira que o número de colônias observadas em amostras de área ou duração fixas tenha a distribuição de Poisson. Então, obtemos a distribuição binomial negativa para a contagem total de indivíduos, se os números de indivíduos nas colônias são distribuídos independentemente segundo uma distribuição logarítmica. Veja, por exemplo, ANSCOMBE (1950).

iii) Um simples modelo de crescimento de população, no qual há taxas constantes de nascimento e morte por indivíduo e uma taxa de imigração constante, leva à distribuição binomial negativa para o tamanho da população. Este modelo foi aplicado ao crescimento de algumas populações vivas, por exemplo populações de bactérias e expansão de uma doença infecciosa na população, ANSCOMBE (1950).

Para maiores informações sobre os vários aspectos da distribuição binomial negativa e/ou outros modelos que induzem a ela, veja JOHNSON & KOTZ (1969) e TRIPATHI (1985).

A distribuição binomial negativa é aplicada como alternativa à distribuição de Poisson, quando se tem dúvidas sobre as suposições necessárias para sua aplicação, principalmente a independência. Como os dados de experimentos teratológicos frequentemente apresentam o efeito de ninhada e, consequentemente, dependência entre as respostas, a distribuição binomial negativa será escolhida para modelar o tamanho da ninhada.

Considere que S_{ij}, o tamanho aleatório da j-ésima ninhada do i-ésimo grupo de dose, siga uma distribuição binomial negativa com parâmetros

$m_i = \delta_1 \exp(-\delta_2 d_i)$ e k, com $\delta_1 > 0$; $-\infty < \delta_2 < \infty$; $k > 0$; $i = 0, \ldots, m$.

A justificativa para que a média de S_{ij} seja a mesma função matemática dependente do nível da dose para todas as ninhadas dentro do mesmo grupo da dose é a mesma apresentada para o caso da distribuição de Poisson. A mesma forma funcional da média da distribuição de Poisson é escolhida para efeito de comparação. O parâmetro k será considerado constante para todos os grupos, também para simplificar a modelagem.

Os modelos de Poisson e binomial negativo para o tamanho da ninhada serão detalhados na discussão do Capítulo 3. Seus parâmetros serão estimados para dois conjuntos de dados no Capítulo 4.

C A P Í T U L O 3

O MODELO GERAL DE RAI & VAN RYZIN

Neste capítulo, apresentaremos uma forma geral do modelo de dose-resposta proposto em RAI & VAN RYZIN (1985), para avaliação de risco de efeitos teratogênicos. Em toda esta tese ele será denominado modelo geral RVR.

Visto que este modelo foi construído em analogia aos de dose-resposta para avaliação de risco de câncer, na primeira seção deste capítulo discutiremos brevemente estes últimos modelos, dando ênfase ao de um impacto.

Em seguida, serão considerados alguns aspectos do modelo geral condicional RVR, onde a probabilidade de resposta adversa depende dos valores observados dos tamanhos de ninhada. Originalmente, os autores supõem que a variável aleatória que representa os tamanhos das ninhadas segue uma distribuição de Poisson cuja média é função da dose. Numa tentativa de tratar mais adequadamente a variabilidade destes tamanhos, sugerimos uma modificação no modelo original, que consiste em

supor uma distribuição binomial negativa para o tamanho de ninhada. Na terceira seção, será calculada a probabilidade incondicional de resposta adversa, sob ambas as distribuições para os tamanhos das ninhadas.

Na última seção, serão estudados os estimadores de máxima verossimilhança dos parâmetros dos modelos. Para o modelo de dose-resposta, vamos demonstrar que algumas propriedades assintóticas de seus estimadores, tais como existência, unicidade e normalidade são satisfeitas.

3.1 MODELOS PARA AVALIAÇÃO DO RISCO DE CÂNCER

Na modelagem estatística da origem do câncer, a relação entre o número de animais que desenvolvem um determinado tumor durante um experimento e o nível de exposição pode ser descrita por um modelo probabilístico. O mesmo relaciona a probabilidade de indução do tumor e a dose d mediante uma função P(d).

Alguns destes modelos postulam a existência de um limiar individual de intensidade, denominado tolerância do indivíduo, tal que a resposta só se apresenta se o estímulo for superior a este limiar. Parece então razoável descrever a distribuição das tolerâncias T mediante a probabilidade de que um animal selecionado ao acaso responda à dose d, obtendo-se então o modelo $P(d) = P(T \leq d)$.

Dois modelos clássicos de distribuição de tolerâncias são o probito e logístico, baseados respectivamente em uma distribuição normal ou logística para as tolerâncias. Suas equações se encontram na Tabela 3.1. Mais adiante, no Capítulo 5, serão discutidos e ajustados alguns modelos logísticos para dados de experimentos teratológicos.

A hipótese da existência de tolerâncias é bastante discutida, pois acredita-se que o processo de indução do câncer tenha um caráter estocástico mais complexo do que o proposto pela hipótese de existência de um limiar individual. Desta maneira, também têm sido propostos modelos de dose-resposta que consideram diversos mecanismos

estocásticos de indução. Eles são baseados na suposição de que, para cada animal, a resposta adversa é o resultado da ocorrência aleatória de um ou mais eventos biológicos.

Alguns destes modelos são os de um impacto ("one hit models"), de múltiplos impactos ("multihit models"), de Weibull, de múltiplos estágios ("multistage models"). Estes modelos são baseados na teoria de impactos, que coloca a iniciação do processo cancerígeno no nível celular. Segundo esta teoria, o processo de indução do câncer é desencadeado por um impacto inicial devido à dose d, continuando depois independentemente da dose aplicada. Na Tabela 3.1 encontram-se as equações dos modelos acima mencionados e os parâmetros contidos em cada um deles.

Tabela 3.1: Alguns modelos de dose-resposta

Modelo	Função P(d)	Parâmetros
probito	$\Phi\left(\frac{\log(d)-\mu}{\sigma}\right)$ (*)	$-\infty < \mu < \infty,\ \sigma > 0$
logístico	$(1+\exp[-(\log(\theta_0)+\theta_1 \log(d))])^{-1}$	$\theta_0,\ \theta_1 > 0$
de um impacto	$1 - \exp(-\lambda d)$	$\lambda > 0$
de múltiplos impactos	$\int_0^{\lambda d} \frac{x^{k-1}\exp(-x)}{(k-1)!}\,dx$	$\lambda,\ k > 0$
de Weibull	$1 - \exp(-\lambda d^k)$	$\lambda,\ k > 0$
de múltiplos estágios	$1 - \exp\{-\sum_{i=0}^{k} b_i d^i\}$	$b_0, \ldots, b_k \geq 0$

(*) $\Phi(x) = (2\Pi)^{-1/2} \int_{-\infty}^{x} \exp(-u^2/2)\,du$

Apesar de importantes, estes modelos não serão discutidos neste texto, com exceção do modelo de um impacto, que é a base do modelo geral RVR. Portanto, para informações adicionais sobre os modelos citados anteriormente e comparações entre estes modelos quanto a ajustes e estimativas de risco, enviamos o leitor a VAN RYZIN (1980); RAI & VAN RYZIN (1981); KREWSKI & VAN RYZIN (1981); BROWN (1983) e BROWN & KOZIOL (1983). Para detalhes sobre modelos para avaliacão de risco de câncer, veja GART et al (1986). Veja também em FREEDMAN & ZEISEL (1988) várias críticas sobre o procedimento de avaliação de risco de câncer via ajuste de modelos de dose-resposta.

No que segue, faremos uma breve discussão a respeito do modelo de um impacto, que como as equações da Tabela 3.1 mostram, é um caso particular dos modelos de múltiplos impactos, do modelo de Weibull e do modelo de múltiplos estágios.

No modelo de um impacto, assume-se que a resposta é induzida uma vez que o tecido alvo tenha sido atingido por uma única unidade de dose biologicamente efetiva, dentro de um intervalo de tempo especificado. Assumindo que o número X de impactos durante este período segue a distribuição de Poisson, a probabilidade de que um animal responda na dose d é dada por:

$$P(d) = P(X \geq 1) = 1 - P(X = 0) = 1 - \exp(-\lambda d),$$

onde a intensidade λ é positiva e λd é o número esperado de impactos durante este intervalo de tempo.

O modelo de um impacto tem diversos inconvenientes: a ausência de uma definição biológica precisa de impacto, a escassa evidência de que um número específico de impactos cause câncer e a suposição de que os impactos são descritos por um processo de Poisson.

Em relação ao parâmetro λ, temos que para d fixo, quando λ aumenta, também cresce a probabilidade de indução do câncer. Fixando λ, esta probabilidade também é crescente como função de d. Quando a dose aumenta, a probabilidade de iniciação do câncer se aproxima de 1 e

se λd é pequeno, a equação do modelo pode ser aproximada mediante um desenvolvimento em série de Taylor, produzindo

$$P(d) \cong \lambda d.$$

Por este motivo, o modelo de um impacto é às vezes chamado de linear.

A linearidade do modelo de um impacto nos níveis baixos de dose é uma característica importante dos modelos conservadores, que produzem grandes riscos estimados. Entretanto, esta linearidade não se apresenta com frequência no ajuste de dados típicos de experimentos em animais.

A partir de agora nos deteremos no estudo do modelo de dose-resposta e nos modelos de Poisson e binomial negativo para o tamanho da ninhada. Depois de apresentarmos o modelo geral RVR discutiremos a estimação dos parâmetros e estudaremos suas propriedades assintóticas. No capítulo 4 vamos estimar os parâmetros dos modelos para dois conjuntos de dados da literatura com a ajuda de diferentes procedimentos numéricos.

3.2 MODELO CONDICIONAL

Em um experimento teratológico, as fêmeas sofrem a exposição direta ou indireta ao agente em estudo, observando-se a resposta de interesse nos filhotes gerados após a exposição. O efeito da dose, no entanto, pode estar presente tanto nas fêmeas quanto em seus filhotes. Mais ainda, a probabilidade de que um filhote apresente a resposta a uma dada dose pode depender do efeito daquela dose no organismo materno.

Considere então um experimento teratológico para avaliação do risco associado a um determinado agente. Supondo o mecanismo de um impacto para o efeito da dose na fêmea, a probabilidade de ocorrência de uma resposta tóxica em uma fêmea que sofreu uma exposição ao nível d da dose pode ser modelada por:

$$\lambda_1(d) = 1 - \exp[-(\alpha + \beta d)] \qquad \alpha > 0,\ \beta > 0,\ d \geq 0 \qquad (3.1)$$

Note que a probabilidade "basal" de resposta tóxica na fêmea é incorporada ao modelo pela introdução do intercepto α na expressão linear do expoente do modelo (3.1).

Queremos avaliar a probabilidade de que um filhote apresente o resultado adverso de interesse, considerando que ela varia para filhotes de diferentes fêmeas dentro de um mesmo grupo de dose. Podemos lidar com estas diferentes probabilidades de resposta considerando algumas variáveis que podem estar afetando tal probabilidade. Neste caso, para efeito de simplificação, vamos considerar apenas o tamanho da ninhada.

Em RAI & VAN RYZIN (1985) é proposto que $P(d|s)$, a probabilidade condicional de resposta adversa para um filhote de uma fêmea que produziu uma ninhada de tamanho s quando exposta à dose d, seja dada por

$$P(d|s) = (1 - \exp[-(\alpha + \beta d)]) \exp[-s\theta(d)] = \lambda_1(d) \exp[-s\theta(d)], \qquad (3.2)$$

com $\theta(d) > 0$ e $s \geq 0$.

Certas características deste modelo o tornam bastante razoável. Destacaremos algumas delas.

A primeira característica importante é que o segundo fator em (3.2) pode ser interpretado como a probabilidade condicional de resposta adversa para um filhote, dado que a mãe recebeu uma dose d e produziu uma ninhada de tamanho s. No caso de $s = 0$, este fator vale 1, significando que se não houve a produção de filhotes, é certa a existência da resposta adversa. Assim, $\lambda_1(d)$ representa uma probabilidade "basal" de resposta adversa para o filhote.

A segunda característica a ser destacada é que o risco relativo (RR) de uma resposta adversa para filhotes de fêmeas que produziram tamanhos de ninhadas s_1 e s_2, para uma mesma dose d, é dado por

$$RR(s_1,s_2) = \frac{P(d|s_1)}{P(d|s_2)} = \exp(-\theta(d)[s_1 - s_2]). \qquad (3.3)$$

Assim, para $s_1 = s$ e $s_2 = 0$, temos que $RR(s,0) = \exp[-s\theta(d)]$, e o segundo fator de (3.2) pode também ser interpretado como o risco relativo de resposta adversa para os filhotes de uma fêmea que produziu uma ninhada de tamanho s em relação a uma fêmea que não produziu filhotes, dentro de um mesmo grupo experimental.

Outra característica aparentemente razoável de se ter em modelos para experimentos teratológicos, presente no modelo geral RVR, é que fixando a dose, quanto maior for o tamanho da ninhada produzida por uma fêmea, menor será o risco de que seus filhotes exibam a resposta adversa, em relação a uma fêmea que não produziu filhotes. Desta maneira, o modelo sugere que uma fêmea que produz mais filhotes é menos afetada pela dose e, consequentemente, seus filhotes terão menor chance de exibir a resposta adversa.

Voltando ao risco relativo, para $s_1 = s+1$ e $s_2 = s$, temos que $RR(s+1,s) = \exp[-\theta(d)]$ e, neste caso, o risco de resposta adversa para um filhote de uma fêmea com tamanho de ninhada s+1 em relação a uma fêmea com tamanho de ninhada s, é uma constante dependente da dose.

Assumindo que a função positiva $\theta(d)$ é linear para qualquer dose, obtemos

$$\theta(d) = \theta_1 + \theta_2 d \quad \text{com} \quad \theta_1 > 0 \quad e \quad \theta_1 + \theta_2 d > 0. \qquad (3.4)$$

Com esta suposição, resulta que

$$\log[RR(s_1,s_2)] = -(\theta_1 + \theta_2 d)(s_1 - s_2) \qquad (3.5)$$

e, portanto, o logaritmo do risco relativo para filhotes de fêmeas que produziram ninhadas de tamanhos s_1 e s_2 é linear na dose e não nulo no grupo controle. Como casos particulares, temos que

$$\log[RR(s+1,s)] = -(\theta_1 + \theta_2 d) \qquad e \qquad \log[RR(s,0)] = -s(\theta_1 + \theta_2 d).$$

Para muitas substâncias teratogênicas de origem química, é de se esperar que $\log[RR(s+1,s)]$ seja uma função não decrescente da dose e, portanto, $\theta_2 \leq 0$.

Incorporando a suposição (3.4), o modelo de dose-resposta (3.2) pode ser reescrito como

$$P(d|s) = \lambda_1(d) \exp[-s(\theta_1 + \theta_2 d)] \qquad (3.6)$$

com $\lambda_1(d)$ dado em (3.1).

Note que o modelo geral RVR depende do valor s observado para o tamanho da ninhada, que pode ser representado por uma variável aleatória não negativa e discreta S, com função de probabilidade denotada por $f(s,\underset{\sim}{\gamma},d)$, onde $\underset{\sim}{\gamma}$ é um vetor de parâmetros. Para generalizar o modelo RVR, deve ser postulada a função de probabilidade para S. Neste trabalho são consideradas duas distribuições para o tamanho da ninhada: a de Poisson, originalmente proposta em RAI & VAN RYZIN (1985) e a binomial negativa. No modelo de Poisson, o parâmetro de intensidade é dado por $E(S) = \phi_1 \exp(-\phi_2 d)$, resultando em $\underset{\sim}{\gamma} = (\gamma_1, \gamma_2)' = (\phi_1, \phi_2)'$. No caso do modelo binomial negativo, supomos que seus parâmetros são $E(S) = \delta_1 \exp(-\delta_2 d)$ e k e, portanto, $\underset{\sim}{\gamma} = (\gamma_1, \gamma_2, \gamma_3)' = (\delta_1, \delta_2, k)'$. Lembre-se que na seção 2.2 foi apresentada uma justificativa para a escolha destes modelos. No próximo capítulo, os parâmetros destes modelos serão estimados com o auxílio de métodos numéricos.

Já foi dito que o modelo geral RVR é um modelo condicional ao tamanho da ninhada. Entretanto, é possível obter o modelo incondicional, supondo conhecida a distribuição de probabilidade para o tamanho da ninhada. Isto é feito a seguir, para os dois casos estudados.

3.3 MODELO INCONDICIONAL

Se o tamanho da ninhada é desconhecido e $f(s,\underset{\sim}{\gamma},d)$, sua função de probabilidade, é conhecida exceto pelo parâmetro $\underset{\sim}{\gamma}$, a probabilidade marginal de resposta para qualquer filhote de uma fêmea exposta à dose d pode ser denotada por P(d).

Pelo teorema das probabilidades totais, da relação (3.6) segue que

$$P(d) = E_S[P(d|S)] = \lambda_1(d)\, E_S(\exp[-S(\theta_1 + \theta_2 d)]).$$

O valor esperado $E_S(\exp[-S(\theta_1+\theta_2 d)])$ para o caso em que o tamanho da ninhada S é modelada segundo as distribuições de Poisson e binomial negativa, resulta

$$E_S(\exp[-S(\theta_1 + \theta_2 d)]) = \sum_{s=0}^{\infty} \exp[-s(\theta_1 + \theta_2 d)]\, P(S = s).$$

No modelo de Poisson para o tamanho da ninhada S, temos

$$E_S(\exp[-S(\theta_1+\theta_2 d)]) = \exp[-\phi_1 \exp(-\phi_2 d)] \sum_{s=0}^{\infty} \frac{(\phi_1 \exp(-\phi_2 d)\exp[-(\theta_1+\theta_2 d)])^s}{s!}.$$

Lembrando que $e^a = \sum_{x=0}^{\infty} \frac{a^x}{x!}$ para $-\infty < a < \infty$ e fazendo $x = s$ e $a = \phi_1 \exp(-\phi_2 d)\exp[-(\theta_1 + \theta_2 d)]$ com $a > 0$, temos que

$$E_S(\exp[-S(\theta_1 + \theta_2 d)]) = \exp(-\phi_1 \exp(-\phi_2 d)\,[1 - \exp[-(\theta_1+\theta_2 d)]]).$$

Portanto,

$$P(d) = (1 - \exp[-(\alpha + \beta d)])\exp(-\phi_1 \exp(-\phi_2 d)\,[1 - \exp[-(\theta_1 + \theta_2 d)]]).$$

No modelo binomial negativo, temos que

$E_S(\exp[-S(\theta_1 + \theta_2 d)]) =$

$$= \left[1 + \frac{\delta_1 \exp(-\delta_2 d)}{k}\right]^{-k} \sum_{s=0}^{\infty} \binom{s+k-1}{s} \left[\frac{\exp[-(\theta_1+\theta_2 d)]\ \delta_1 \exp(-\delta_2 d)}{\delta_1 \exp(-\delta_2 d) + k}\right]^{s}.$$

Lembrando que $\frac{1}{(1-x)^{n+1}} = \sum_{j=0}^{\infty} \binom{n+j}{j} x^j$ para $-1 < x < 1$ e fazendo

$n = k - 1$ e $x = \frac{\exp[-(\theta_1+\theta_2 d)]\ \delta_1 \exp(-\delta_2 d)}{\delta_1 \exp(-\delta_2 d) + k}$ com $0 < x < 1$, temos que

$$E_S(\exp[-S(\theta_1 + \theta_2 d)]) = \left[1 + \frac{\delta_1 \exp(-\delta_2 d)\ (1 - \exp[-(\theta_1 + \theta_2 d)])}{k}\right]^{-k}.$$

E, assim,

$$P(d) = (1 - \exp[-(\alpha+\beta d)]) \left[1 + \frac{\delta_1 \exp(-\delta_2 d)\ (1 - \exp[-(\theta_1 + \theta_2 d)])}{k}\right]^{-k}.$$

No capítulo 4, após a estimação dos parâmetros dos modelos para dois conjuntos de dados, também serão estimados os valores de P(d).

Nosso interesse é estimar os parâmetros α, β, θ_1, θ_2 e $\underset{\sim}{\gamma}$, e com eles o risco de que um filhote da fêmea exposta a uma dose pré-fixada apresente o resultado adverso de interesse. Tal estimação será feita através dos dados observados no experimento utilizando o método de estimação por máxima verossimilhança, baseado na maximização da função de verossimilhança.

3.4 ESTUDO DA VEROSSIMILHANÇA

Considere um experimento teratológico como definido na seção 2.1 do Capítulo 2. Neste caso, a suposição de independência entre as variáveis aleatórias X_{ijk} correspondentes a diferentes grupos e a diferentes fêmeas parece razoável sempre que o experimento for

conduzido de maneira não viciada; isto é, sempre que for mantida a comparabilidade dos grupos experimentais e das fêmeas incluídas no experimento.

Para cada fêmea, vamos assumir que as respostas de seus filhotes podem ser modeladas mediante variáveis aleatórias independentes. Isto é, assumiremos que o fato de um filhote de uma determinada fêmea apresentar a resposta adversa não influencia a resposta dos outros filhotes da mesma fêmea.

Esta última suposição é bastante forte, pois uma correlação positiva entre respostas pode em geral ser aguardada neste caso, devido ao fenômeno do efeito de ninhada discutido no Capítulo 2. Entretanto, vale notar que o modelo geral RVR incorpora de certa forma aquela dependência existente - mas desconsiderada - entre as respostas dos animais de uma mesma ninhada, ao propor que a probabilidade de resposta seja função do valor observado de uma variável aleatória própria de cada fêmea: o tamanho da ninhada.

Com estas suposições, temos então que as respostas de todos os filhotes que participam do experimento são independentes. Isto é, as variáveis aleatórias X_{ijk} são independentes para todo $k = 1, 2, \ldots, s_{ij}$, para todo $j = 1, 2, \ldots, n_i$ e para todo $i = 0, 1, \ldots, m$.

Sob o modelo geral RVR, a probabilidade condicional com respeito ao tamanho da ninhada, de que o k-ésimo filhote da j-ésima fêmea do i-ésimo grupo apresente a resposta de interesse, dada pela equação (3.6), pode ser reescrita como

$$P(X_{ijk}=1|s_{ij}) = P_{\underset{\sim}{\psi}}(d_i|s_{ij})$$

$$= \{1 - \exp[-(\alpha + \beta d_i)]\} \exp[-s_{ij}(\theta_1 + \theta_2 d_i)] \qquad (3.7)$$

onde $\underset{\sim}{\psi} = (\alpha, \beta, \theta_1, \theta_2)'$ é um vetor de parâmetros com $\alpha > 0$, $\beta > 0$, $\theta_1 > 0$ e $\theta_1 + \theta_2 d_i > 0$ para todo $i = 0, \ldots, m$.

Desta maneira, a distribuição de X_{ijk} condicional sob $S_{ij} = s_{ij}$ é uma

distribuição de Bernoulli com parâmetro $P_{\underset{\sim}{\psi}}(d_i|s_{ij})$. Este parâmetro será denotado por $P_{ij,\underset{\sim}{\psi}}$ ou ainda por P_{ij}, se não houver possibilidade de confusão. Portanto, condicional ao tamanho da ninhada, a probabilidade de obter uma resposta adversa é a mesma para todos os filhotes de uma dada fêmea dentro de um determinado grupo, pois esta probabilidade depende apenas do nível da dose e do tamanho da ninhada produzida pela fêmea.

Seja $Y_{ij} = \sum_{k=1}^{s_{ij}} X_{ijk}$ a variável aleatória que representa o número de filhotes da j-ésima fêmea do i-ésimo grupo, que apresentam o resultado adverso de interesse. Pelas mesmas considerações feitas à respeito de X_{ijk}, temos que os Y_{ij} são independentes para $j = 1, 2, \ldots, n_i$ e $i = 0, 1, \ldots, m$.

Condicional ao tamanho s_{ij} da ninhada da j-ésima fêmea do i-ésimo grupo, a função de probabilidade de Y_{ij} é dada por

$$f(Y_{ij} = y_{ij}, \underset{\sim}{\psi} | S_{ij} = s_{ij}) = \begin{pmatrix} s_{ij} \\ y_{ij} \end{pmatrix} (P_{ij})^{y_{ij}} (Q_{ij})^{s_{ij} - y_{ij}} \qquad (3.8)$$

para $j = 1, 2, \ldots, n_i$; $i = 0, 1, \ldots, m$ onde $y_{ij} \in \{0, \ldots, s_{ij}\}$, $Q_{ij} = 1 - P_{ij}$ e $\begin{pmatrix} s_{ij} \\ y_{ij} \end{pmatrix} = \dfrac{s_{ij}!}{y_{ij}!\,(y_{ij} - s_{ij})!}$.

Ou seja, se S_{ij} é a variável aleatória que representa o tamanho da ninhada, a distribuição condicional (sob $S_{ij} = s_{ij}$) da variável aleatória Y_{ij} é binomial com parâmetros s_{ij} e P_{ij}. A notação $Y_{ij}|S_{ij} = s_{ij} \sim bin(s_{ij}, P_{ij})$ denota este fato.

Como as S_{ij} são variáveis aleatórias discretas com função

de probabilidade $f(s_{ij},\underset{\sim}{\gamma},d_i)$, a função de verossimilhança é dada por:

$$L = \prod_{i=0}^{m} \prod_{j=1}^{n_i} P(Y_{ij} = y_{ij},\ S_{ij} = s_{ij})$$

$$= \prod_{i=0}^{m} \prod_{j=1}^{n_i} P(Y_{ij} = y_{ij} | S_{ij} = s_{ij})\ P(S_{ij} = s_{ij})$$

$$= \prod_{i=0}^{m} \prod_{j=1}^{n_i} \binom{s_{ij}}{y_{ij}} P_{ij}^{y_{ij}}\ Q_{ij}^{s_{ij} - y_{ij}}\ f(s_{ij},\underset{\sim}{\gamma},d_i);$$

seu logaritmo pode ser escrito na forma $\log(L) = c + l(\underset{\sim}{\psi}) + l(\underset{\sim}{\gamma})$, onde

$c = \sum_{i=0}^{m} \sum_{j=1}^{n_i} \log\left[\binom{s_{ij}}{y_{ij}} \right]$ é uma constante,

$$l(\underset{\sim}{\psi}) = l(\alpha,\beta,\theta_1,\theta_2) = \sum_{i=0}^{m} \sum_{j=1}^{n_i} y_{ij} \log(P_{ij}) + (s_{ij} - y_{ij}) \log(Q_{ij}) \qquad (3.9)$$

e $\quad l(\underset{\sim}{\gamma}) = \sum_{i=0}^{m} \sum_{j=1}^{n_i} \log[f(s_{ij},\underset{\sim}{\gamma},d_i)].$

Se S_{ij}, o tamanho da j-ésima ninhada do i-ésimo grupo de dose, segue uma distribuição de Poisson com parâmetro $\phi_1 \exp(-\phi_2 d_i)$, então

$$f(s_{ij},\ \underset{\sim}{\gamma},\ d_i) = \frac{\exp[-\phi_1 \exp(-\phi_2 d_i)]\ [\phi_1 \exp(-\phi_2 d_i)]^{s_{ij}}}{s_{ij}\ !}$$

onde $\underset{\sim}{\gamma} = (\gamma_1,\ \gamma_2) = (\phi_1,\ \phi_2)$, $\phi_1 > 0$, $-\infty < \phi_2 < \infty$ e $s_{ij} = 0,\ 1,\ldots,\ \infty$ e

$$l(\underset{\sim}{\gamma}) = \sum_{i=0}^{m} \sum_{j=1}^{n_i} -\phi_1 \exp(-\phi_2 d_i) + s_{ij} \log(\phi_1) - \phi_2 s_{ij} d_i - \log(s_{ij}!). \qquad (3.10)$$

Mas se S_{ij} segue a distribuição binomial negativa com parâmetros $m_i = \delta_1 \exp(-\delta_2 d_i)$ e k, então,

$$f(s_{ij},\underset{\sim}{\gamma},d_i) =$$

$$= \binom{s_{ij}+k-1}{k-1}\left(\frac{\delta_1\exp(-\delta_2 d_i)}{\delta_1\exp(-\delta_2 d_i)+k}\right)^{s_{ij}}\left(1+\frac{\delta_1\exp(-\delta_2 d_i)}{k}\right)^{-k}$$

com $\underset{\sim}{\gamma} = (\gamma_1, \gamma_2, \gamma_3)' = (\delta_1, \delta_2, k)'$, $\delta_1 > 0$, $-\infty < \delta_2 < \infty$ e $k > 0$ e,

portanto, $$l(\underset{\sim}{\gamma}) = \sum_{i=0}^{m}\sum_{j=1}^{n_i}\log\binom{s_{ij}+k-1}{k-1} + s_{ij}\log[\delta_1\exp(-\delta_2 d_i)] -$$

$$- (s_{ij}+k)\log[\delta_1\exp(-\delta_2 d_i)+k] + k\log(k). \qquad (3.11)$$

O núcleo do logaritmo de L, denotado por $l(\underset{\sim}{\theta})$ é dado por

$$l(\underset{\sim}{\theta}) = l(\underset{\sim}{\psi}) + l(\underset{\sim}{\gamma})$$

onde a função $l(\underset{\sim}{\psi})$ depende dos parâmetros do modelo de dose-resposta e a função $l(\underset{\sim}{\gamma})$ depende dos parâmetros da distribuição de probabilidade do tamanho da ninhada e $\underset{\sim}{\theta} = (\underset{\sim}{\psi}, \underset{\sim}{\gamma})$.

Como $l(\underset{\sim}{\theta})$ é a soma de duas funções, a inferência estatística relativa ao parâmetro $\underset{\sim}{\psi}$ pode ser realizada separadamente da inferência relativa ao parâmetro $\underset{\sim}{\gamma}$. As estimativas de máxima verossimilhança de $\underset{\sim}{\psi}$ e $\underset{\sim}{\gamma}$ serão em geral obtidas maximizando-se $l(\underset{\sim}{\psi})$ e $l(\underset{\sim}{\gamma})$, respectivamente, por métodos numéricos, como será visto adiante no Capítulo 4.

A seguir, vamos obter as matrizes de covariâncias assintóticas dos estimadores de máxima verossimilhança de $\underset{\sim}{\psi}$ e $\underset{\sim}{\gamma}$, quando n e n_i tendem ao infinito de uma maneira que será precisada posteriormente.

3.4.1 MATRIZ DE COVARIÂNCIAS

Sob as condições de regularidade que serão apresentadas na próxima sub-seção, os estimadores de máxima verossimilhança de α, β, θ_1 e θ_2, denotados respectivamente por $\hat{\alpha}$, $\hat{\beta}$, $\hat{\theta}_1$, e $\hat{\theta}_2$, possuem uma matriz de covariância Σ com elementos σ_{kl} (k, l = 1, ...,4), que é obtida invertendo-se a matriz de informação de Fisher $I(\alpha,\beta,\theta_1,\theta_2)$ condicional aos tamanhos observados das ninhadas s_{ij} para j = 1, ..., n_i e i = 0, ..., m.

Por definição, o elemento i_{kl} da k-ésima linha e l-ésima coluna da matriz $I(\alpha,\beta,\theta_1,\theta_2)$, é dado por

$$i_{kl} = E\left[\frac{\partial l(\underset{\sim}{\psi})}{\partial \psi_k}\frac{\partial l(\underset{\sim}{\psi})}{\partial \psi_l}\right]\Bigg|_{\underset{\sim}{\psi}_0} = -E\left[\frac{\partial^2 l(\underset{\sim}{\psi})}{\partial \psi_k\,\partial \psi_l}\right]\Bigg|_{\underset{\sim}{\psi}_0}$$

para k, l = 1, ..., 4 e onde $\underset{\sim}{\psi} = (\psi_1, \psi_2, \psi_3, \psi_4) = (\alpha, \beta, \theta_1, \theta_2)$ e $\underset{\sim}{\psi}_0$ é o verdadeiro valor do vetor paramétrico.

Utilizando o fato de que Y_{ij} é $bin(s_{ij}, P_{ij})$ e, portanto, $E(Y_{ij}|S_{ij}) = s_{ij}P_{ij}$ para j = 1, ..., n_i e i = 0, ..., m, podemos calcular explicitamente a matriz de informação de Fisher. Derivando parcialmente a função de log-verossimilhança resulta

$$\frac{\partial l(\underset{\sim}{\psi})}{\partial \psi_l} = \sum_{i=0}^{m}\sum_{j=1}^{n_i}\left[y_{ij}\frac{1}{P_{ij}}\frac{\partial P_{ij}}{\partial \psi_l} + (s_{ij}-y_{ij})\frac{1}{Q_{ij}}\frac{\partial Q_{ij}}{\partial \psi_l}\right],$$

$$\frac{\partial^2 l(\underset{\sim}{\psi})}{\partial \psi_k\,\partial \psi_l} = \sum_{i=0}^{m}\sum_{j=1}^{n_i} - \left(\frac{y_{ij}}{P_{ij}^2} + \frac{s_{ij}-y_{ij}}{Q_{ij}^2}\right)\frac{\partial P_{ij}}{\partial \psi_k}\frac{\partial P_{ij}}{\partial \psi_l} +$$

$$+ \sum_{i=0}^{m} \sum_{j=1}^{n_i} \left(\frac{y_{ij}}{P_{ij}} - \frac{s_{ij} - y_{ij}}{Q_{ij}} \right) \frac{\partial^2 P_{ij}}{\partial \psi_k \, \partial \psi_l},$$

onde $Q_{ij} = 1 - P_{ij}$. Como a esperança do segundo termo é nula, temos

$$i_{kl} = -E\left(\frac{\partial^2 l(\psi)}{\partial \psi_k \, \partial \psi_l} \right) = \sum_{i=0}^{m} \sum_{j=1}^{n_i} \frac{s_{ij}}{P_{ij} Q_{ij}} \frac{\partial P_{ij}}{\partial \psi_k} \frac{\partial P_{ij}}{\partial \psi_l} \qquad (3.12)$$

para k, l = 1, ..., 4. Derivando P_{ij} com respeito aos parâmetros, obtemos as relações

$$\varphi_1(d_i) = \frac{\partial P_{ij}}{\partial \psi_1} = \frac{\partial P_{ij}}{\partial \alpha} = \exp[-(\alpha+\beta d_i)] \exp[-s_{ij}(\theta_1+\theta_2 d_i)] = \exp[-(\rho+\eta d_i)]$$

$$\varphi_2(d_i) = \frac{\partial P_{ij}}{\partial \psi_2} = \frac{\partial P_{ij}}{\partial \beta} = d_i \, \varphi_1(d_i) = d_i \exp[-(\rho+\eta d_i)]$$

$$\varphi_3(d_i) = \frac{\partial P_{ij}}{\partial \psi_3} = \frac{\partial P_{ij}}{\partial \theta_1} = -s_{ij}(\exp[-s_{ij}(\theta_1+\theta_2 d_i)] - \exp[-(\rho+\eta d_i)]) = -s_{ij} P_{ij}$$

$$\varphi_4(d_i) = \frac{\partial P_{ij}}{\partial \psi_4} = \frac{\partial P_{ij}}{\partial \theta_2} = d_i \, \varphi_3(d_i) = -s_{ij} d_i P_{ij}$$

com $\rho = \alpha + s\theta_1$ e $\eta = \beta + s\theta_2$; $\alpha > 0$, $\beta > 0$, $\theta_1 > 0$,

$\theta_1 + \theta_2 d_i > 0$ para todo i = 0, ..., m. (3.13)

Calcularemos a seguir a matriz de informação para os parâmetros da distribuição de Poisson suposta válida para o tamanho da ninhada, cuja função log-verossimilhança é exibida em (3.10). Note que como temos apenas dois parâmetros, é mais fácil calcular as derivadas parciais e, consequentemente, a matriz de informação diretamente. Entretanto, vamos escrever estas derivadas de uma maneira um pouco mais

geral.

De agora em diante, será usada a notação

$$D = \left(\frac{\partial}{\partial\theta_1}, \frac{\partial}{\partial\theta_2}, \ldots, \frac{\partial}{\partial\theta_t} \right)' = (D_1, D_2, \ldots, D_t)'$$

e

$$H = \left(\frac{\partial^2}{\partial\theta_k \partial\theta_l} \right) = (D_{kl}) \qquad k, l = 1, \ldots, t$$

onde D e H denotam respectivamente o vetor e a matriz cujos são os operadores gradiente e Hessiano, e t é a dimensão do vetor geral de parâmetros $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_t)'$.

A matriz de informação de Fisher $I(\underset{\sim}{\gamma}) = (i_{kl}) = \Sigma^{-1}$ k, l = 1, 2 coincide com $E(D[l(\underset{\sim}{\gamma})]\ D'[l(\underset{\sim}{\gamma})]) = -E(H[l(\underset{\sim}{\gamma})])$.

Mas
$$D_k[l(\underset{\sim}{\gamma})] = \sum_{i=0}^{m} n_i \frac{D_k[\phi_1 \exp(-\phi_2 d_i)]\ [s_{ij} - \phi_1 \exp(-\phi_2 d_i)]}{\phi_1 \exp(-\phi_2 d_i)}$$

e

$$D_{kl}[l(\underset{\sim}{\gamma})] = \sum_{i=0}^{m} n_i \left[\frac{D_{kl}[\phi_1 \exp(-\phi_2 d_i)]\ [s_{ij} - \phi_1 \exp(-\phi_2 d_i)]}{\phi_1 \exp(-\phi_2 d_i)} - \right.$$

$$\left. - \frac{s_{ij}\ D_k[\phi_1 \exp(-\phi_2 d_i)]\ D_l[\phi_1 \exp(-\phi_2 d_i)]}{[\phi_1 \exp(-\phi_2 d_i)]^2} \right].$$

Logo, como $E(S_{ij}) = \phi_1 \exp(-\phi_2 d_i)$, temos que para k, l = 1, 2

$$i_{kl} = -E(D_{kl}[l(\underset{\sim}{\gamma})]) = \sum_{i=0}^{m} n_i \frac{D_k[\phi_1 \exp(-\phi_2 d_i)]\ D_l[\phi_1 \exp(-\phi_2 d_i)]}{\phi_1 \exp(-\phi_2 d_i)} \qquad (3.14)$$

No caso da distribuição binomial negativa para o tamanho

da ninhada, a log-verossimilhança correspondente é exibida em (3.11). De maneira análoga, a matriz de informação de Fisher $I(\underset{\sim}{\gamma}) = \Sigma^{-1}$ é, neste caso, de dimensão 3x3 e composta pelos elementos

$$i_{kl} = \sum_{i=0}^{m}\sum_{j=1}^{n_i}\left\{ D_k(\gamma_3)D_l(\gamma_3)\left[\sum_{p=0}^{w_{ij}-1}(\gamma_3+p)^{-2} - \frac{\gamma_1\exp(-\gamma_2 d_i)}{[\gamma_1\exp(-\gamma_2 d_i)+\gamma_3]\gamma_3} \right] + \right.$$

$$\left. + D_k[\gamma_1\exp(-\gamma_2 d_i)]\, D_l[\gamma_1\exp(-\gamma_2 d_i)]\, \frac{\gamma_3}{\gamma_1\exp(-\gamma_2 d_i)[\gamma_1\exp(-\gamma_2 d_i)+\gamma_3]} \right\}. \qquad (3.15)$$

Na seção 3.4.3, demonstraremos três teoremas relacionados a propriedades assintóticas dos estimadores de máxima verossimilhança do modelo geral RVR. Na demonstração destes teoremas, certas condições de regularidade precisam ser satisfeitas.

A primeira delas diz respeito aos tamanhos n_i dos grupos experimentais e ao total $n = \sum_{i=0}^{m} n_i$ de fêmeas no experimento. Sob esta condição, a medida que o número n de animais presentes no experimento aumenta, o número n_i presente em cada grupo aumenta proporcionalmente. Outra condição a ser satisfeita é a identificabilidade do modelo de interesse. A terceira condição está relacionada com a matriz de informação e assegura que ela é positiva definida. A quarta e quinta condições são relativas às derivadas parciais do modelo em relação a seus parâmetros. Estas cinco condições de regularidade serão apresentadas a seguir.

3.4.2 CONDIÇÕES DE REGULARIDADE

As condições de regulariade mencionadas são

Condição 1: $\lim_{\substack{n\to\infty \\ n_i\to\infty}} \left(\frac{n_i}{n} \right) = c_i$ onde $0 < c_i < 1$, $i = 0, \ldots, m$. (C1)

Condição 2: O modelo $P(d,\underset{\sim}{\theta})$ de interesse é identificável. **(C2)**

Condição 3: A matriz de informação de Fisher do modelo de interesse é positiva definida. **(C3)**

Condição 4: Para todo $d \geq 0$, as derivadas parciais de segunda ordem $\dfrac{\partial^2 P(d,\underset{\sim}{\theta})}{\partial\theta_k \partial\theta_l}$ são contínuas para todo $\underset{\sim}{\theta} \in \Omega$, onde $k, l = 1, \ldots, t$ e t é o número de parâmetros. **(C4)**

Condição 5: $P(d,\underset{\sim}{\theta})$ é continuamente diferenciável em d e $\underset{\sim}{\theta}$ para todo $d > 0$ e $\underset{\sim}{\theta} \in \Omega$. **(C5)**

No Teorema 1 adiante será provado que no caso do modelo condicional geral RVR, as condições de regularidade C1 a C5 são satisfeitas. Antes porém, vamos apresentar a definição de identificabilidade de um modelo e um conceito necessário para seu estudo, além de enunciar e provar um lema. Estes resultados serão utilizados na demonstração do Teorema 1.

Nesta tese, o termo <u>modelo estatístico</u> P_θ designa uma função de dose-resposta ou uma medida de probabilidade que descreve a distribuição da variável aleatória que representa o tamanho da ninhada. Em ambos os casos, o modelo P_θ é dito identificável se valores diferentes do parâmetro θ produzem diferentes elementos da família $\{P_\theta : \theta \in \Omega\}$ de interesse.

<u>Definição 1</u>: O modelo $P(d,\underset{\sim}{\theta})$ será dito <u>identificável</u> se a relação $P(d,\underset{\sim}{\theta}) = P(d,\underset{\sim}{\theta}^*)$ para todo $d > 0$, implica em $\underset{\sim}{\theta} = \underset{\sim}{\theta}^*$.

<u>Definição 2</u>: Um conjunto de funções $\{\varphi_r(d), r = 1, \ldots, t\}$ forma um conjunto de Tchebycheff em $[0, D]$ se $\varphi(d, a) = \sum_{r=1}^{t} a_r \varphi_r(d)$ tem no máximo $t-1$ zeros em $[0, D]$ para todo $a = (a_1, \ldots, a_t)' \neq 0$.

Lema 1: O conjunto de funções $\{\varphi_r(d),\ r = 1, \ldots, t\}$ forma um conjunto de Tchebycheff em $[0, D]$ se e somente se a matriz $(\varphi_r(d_i))$, $r, i = 1, \ldots, t$ *é de posto completo para todo conjunto* $\{d_1, \ldots d_t\}$ *de* t *pontos distintos em* $[0, D]$.

Prova:

Suponha que $\{\varphi_r(d),\ r = 1, \ldots, t\}$ forma um conjunto de Tchebycheff em $[0, D]$, considere a matriz

$$A = \begin{bmatrix} \varphi_1(d_1) & \varphi_2(d_1) & \ldots & \varphi_t(d_1) \\ \vdots & \vdots & & \vdots \\ \varphi_1(d_{t-1}) & \varphi_2(d_{t-1}) & \ldots & \varphi_t(d_{t-1}) \\ \varphi_1(d_t) & \varphi_2(d_t) & \ldots & \varphi_t(d_t) \end{bmatrix}$$

e a função $\tau(d) = \det(B)$, onde B é a matriz obtida substituindo d_t por d na matriz A. Assim, $\tau(d) = \sum_{r=1}^{t} \varphi_r(d)B_{tr} = \varphi_1(d)B_{t1} + \ldots + \varphi_t(d)B_{tt}$, onde B_{tr} é o cofator do elemento da t-ésima linha e r-ésima coluna de B.

Desta maneira, $\tau(d_1) = \tau(d_2) = \ldots = \tau(d_{t-1}) = 0$, pois B tem duas linhas iguais. Por outro lado, suponha, por absurdo, que A é singular. Assim, $\det(A) = 0$ e $\tau(d_t) = 0$. Então, $\tau(d) = \sum_{r=1}^{t} \varphi_r(d)B_{tr} = \varphi_1(d)B_{t1} + \ldots + \varphi_t(d)B_{tt}$ tem t raízes e, portanto, $\{\varphi_r(d),\ r = 1, \ldots, t\}$ não forma um conjunto de Tchebycheff.

Reciprocamente, suponha que A é regular. Então, o sistema linear

$$\sum_{r=1}^{t} a_r\varphi_r(d_i) = y_i \text{ para } i = 1, \ldots, t,$$

que pode ser também escrito como $Aa = y$, tem solução única para todo vetor $y = (y_1, \ldots, y_t)'$. Se existissem t zeros de $\{\varphi_r(d),\ r = 1, \ldots,$

t) em [0, D] então Aa = 0 e, portanto, a = 0. Por definição, o conjunto $(\varphi_r(d), r = 1, \ldots, t)$ não seria de Tchebycheff.■

A definição de conjunto de Tchebycheff e o Lema 1 apresentado anteriormente serão úteis na demonstração de que as matrizes de informação são positivas definidas. Antes porém, vamos mostrar que o modelo condicional RVR cumpre as condições de regularidade C1 a C5.

TEOREMA 1. As condições de regularidade C1 a C5 são satisfeitas para o modelo condicional RVR dado por

$$P(d,\underset{\sim}{\psi}|s) = (1 - \exp[-(\alpha + \beta d)])\exp[-s(\theta_1 + \theta_2 d)]$$

onde $\alpha > 0$, $\beta > 0$, $\theta_1 > 0$, $\theta_1 + \theta_2 d > 0$, $\underset{\sim}{\psi} = (\alpha, \beta, \theta_1, \theta_2)'$ e $s \geq 0$ é o tamanho observado da ninhada.

Prova:

Em relacão a condição C1, não há o que ser provado. Para provar a identificabilidade do modelo, considere que se a relação $P(d,\underset{\sim}{\psi}|s) = P(d,\underset{\sim}{\psi}^*|s)$ é válida para todo $d \geq 0$, então a função

$$f(d) + g(d)[1 - h(d)] \qquad (3.16)$$

deve ser identicamente igual a 1, onde $f(d) = \exp[-(\alpha + \beta d)]$, $g(d) = \exp[-s(\nu_1 + \nu_2)]$, $h(d) = \exp[-(\alpha^* + \beta^* d)]$, $\nu_1 = \theta_1^* - \theta_1$, $\nu_2 = \theta_2^* - \theta_2$. Derivando (3.16) duas vezes em relação a d e utilizando as relações $f'(d) = -\beta f(d)$, $g'(d) = -s\nu_2 g(d)$ e $h'(d) = -\beta^* h(d)$, obtemos a identidade $Af(d) + Bg(d) = 0$, onde $A = \beta(-\beta + s\nu_2 + \beta^*)$ e $B = \beta^* s\nu_2$. Mas como f(d) e g(d) são funções linearmente independentes, temos que $A = B = 0$ e assim, $\underset{\sim}{\psi} = \underset{\sim}{\psi}^*$.■

A demonstração de que a matriz de informação de Fisher é positiva definida será feita em duas etapas, com o auxílio dos Lemas 2 e 3.

Lema 2: O conjunto de funções $T = (\varphi_1(d), \varphi_2(d), \varphi_3(d), \varphi_4(d))$ dado em

(3.13) forma um conjunto de Tchebycheff em [0, D] para todo s > 0.

Prova:

De acordo com o Lema 1, mostrar que o conjunto T é de Tchebycheff é equivalente a mostrar que a matriz $A = (\varphi_r(d_i))$ r, i = 1, ..., 4 é *de posto completo para todo conjunto de quatro pontos* distintos (d_1, d_2, d_3, d_4) em [0, D]. Ou seja, basta mostrar que para s > 0, $\det(A) = \det(A') = \frac{-\exp(\rho)}{s} \det(B) \neq 0$ e, portanto, que $\det(B) \neq 0$ onde

$$B = \begin{bmatrix} \exp(-\eta d_1) & \exp(-\eta d_2) & \exp(-\eta d_3) & \exp(-\eta d_4) \\ d_1\exp(-\eta d_1) & d_2\exp(-\eta d_2) & d_3\exp(-\eta d_3) & d_4\exp(-\eta d_4) \\ \exp(-s\theta_2 d_1) & \exp(-s\theta_2 d_2) & \exp(-s\theta_2 d_3) & \exp(-s\theta_2 d_4) \\ d_1\exp(-s\theta_2 d_1) & d_2\exp(-s\theta_2 d_2) & d_3\exp(-s\theta_2 d_3) & d_4\exp(-s\theta_2 d_4) \end{bmatrix}.$$

Ainda pelo Lema 1, a matriz B é regular se e somente se o conjunto

$$U = (\exp(-\eta d),\ d\exp(-\eta d),\ \exp(-s\theta_2 d),\ d\exp(-s\theta_2 d))$$

é de Tchebycheff. Mas, desde que as funções de um conjunto de Tchebycheff multiplicadas por uma função positiva formam ainda um conjunto de Tchebycheff, se mostrarmos que o conjunto $(1,\ d,\ \exp(-\beta d),\ d\exp(-\beta d))$ é de Tchebycheff, teremos mostrado que $\det(B) \neq 0$, que A é de posto completo e, portanto, que T é um conjunto de Tchebycheff.

O conjunto $(1,\ d,\ \exp(-\beta d),\ d\exp(-\beta d))$ forma um conjunto de Tchebycheff se a função

$$\nu(d) = a_1 + a_2 d + a_3\exp(-\beta d) + a_4 d\exp(-\beta d), \qquad \beta > 0$$

tem no máximo três raízes em [0, D] para todo $a = (a_1, a_2, a_3, a_4)' \neq 0$.

Se $a_4 = 0$, $\nu(d) = a_1 + a_2 d + a_3\exp(-\beta d)$ e o Lema 4 de KREWSKI & VAN RYZIN (1981) demonstra que $\nu(d)$ tem no máximo duas raízes. Se $a_4 \neq 0$, suponhamos, por absurdo, que $\nu(d)$ tem no mínimo quatro raízes em [0, D] para todo $a = (a_1, a_2, a_3, a_4)' \neq 0$. Assim, $\nu''(d)$

Prova:

Se $m \geq 3$, segue do Lema 1 que o posto da matriz A de dimensão 4xn, e elementos $a_{kij} = \frac{\partial P_{ij}}{\partial \psi_k}$ $k = 1, \ldots, 4$; $(i,j) \rightarrow r = 1, \ldots, n$ é quatro.

Seja B de dimensão nxn a matriz diagonal com (i,j) ésimo elemento diagonal $\left(\frac{s_{ij}}{P_{ij} Q_{ij}} \right)^{1/2}$ e $C = AB$. Como B é não singular, temos que o posto de C é quatro, veja RAO (1973) p. 30. Como Σ^{-1} é simétrica, C é não singular e $\Sigma^{-1} = CC'$, segue que Σ^{-1} é positiva definida, veja SEARLE (1971) p. 37.■

A quarta condição de regularidade é verificada, visto que as derivadas parciais de segunda ordem de $P(d,\underset{\sim}{\psi}|s)$ com respeito aos parâmetros são continuas. Para simplificar a notação, na apresentação destas derivadas faremos $P(d,\underset{\sim}{\psi}|s) = P$.

$$\frac{\partial^2 P}{\partial \alpha^2} = - \exp[-(\alpha+\beta d)] \exp[-s(\theta_1+\theta_2 d)],$$

$$\frac{\partial^2 P}{\partial \alpha \partial \beta} = d \frac{\partial^2 P}{\partial \alpha^2},$$

$$\frac{\partial^2 P}{\partial \alpha \partial \theta_1} = s \frac{\partial^2 P}{\partial \alpha^2},$$

$$\frac{\partial^2 P}{\partial \alpha \partial \theta_2} = sd \frac{\partial^2 P}{\partial \alpha^2},$$

$$\frac{\partial^2 P}{\partial \beta^2} = d^2 \frac{\partial^2 P}{\partial \alpha^2},$$

$$\frac{\partial^2 P}{\partial \beta \partial \theta_1} = sd \frac{\partial^2 P}{\partial \alpha^2},$$

$$\frac{\partial^2 P}{\partial \beta \partial \theta_2} = sd^2 \frac{\partial^2 P}{\partial \alpha^2},$$

$$\frac{\partial^2 P}{\partial \theta_1^2} = s^2 P,$$

$$\frac{\partial^2 P}{\partial \theta_1 \partial \theta_2} = s^2 dP$$

$$\frac{\partial^2 P}{\partial \theta_2^2} = s^2 d^2 P.$$

A quinta condição de regularidade também é satisfeita, pois $P(d,\underset{\sim}{\psi}|s)$ é continuamente diferenciável para todo $d \geq 0$ e para todo $\underset{\sim}{\psi} = (\alpha, \beta, \theta_1, \theta_2)$.

$= \exp(-\beta d)[a_4\beta^2 d + (a_3\beta^2 - 2a_4\beta)]$ tem - pelo Teorema de Rolle - no mínimo duas raízes em [0, D], o que é absurdo.■

Antes de prosseguir, é útil que se faça uma observação sobre y_{ij}, s_{ij} e P_{ij}. Em algumas ocasiões, será necessário representar as quantidades y_{ij}, s_{ij} e P_{ij} na forma matricial. Portanto, considere como exemplo o vetor "empilhado" ("stack vector", em inglês) y de dimensão nx1 onde $n = \sum_{i=0}^{m} n_i$, com componentes y_{ij}, ou seja

$$y = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{bmatrix} \quad \text{com} \quad y_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{in_i} \end{bmatrix} \quad \text{para } i = 0, 1, \ldots, m.$$

Adotaremos a seguinte notação para descrever a relação entre o par (grupo, fêmea) = (i, j) e o índice r:

$$(i, j) \longrightarrow f(i, j) = r,$$

onde $j = 1, \ldots, n_i$; $i = 0, \ldots, m$; $r = j$ para $i = 0$ e $r = \sum_{l=0}^{i-1} n_l + j$ para $i \neq 0$. Logo, $r = 1, \ldots, n$.

Assim, nas demonstrações a seguir, a referência a alguma componente (i, j) dos vetores y, s e P é equivalente à referência à componente r = f(i, j) dos mesmos vetores.

Lema 3: Se o conjunto $T = \{\varphi_1(d), \varphi_2(d), \varphi_3(d), \varphi_4(d)\}$ do Lema 2 é um conjunto de Tchebycheff em [0, D], se $s_{ij} > 0$ para todo $i = 0, \ldots, m$ e para todo $j = 1, \ldots, n_i$, se $m \geq 3$, e portanto, $n = \sum_{i=0}^{m} n_i \geq 4$, então a matriz de informação de Fisher sobre o parâmetro $\underset{\sim}{\psi}$ contida no vetor aleatório Y, $I(\underset{\sim}{\psi}) = \Sigma^{-1} = (i_{kl})$ dada em (3.12) é positiva definida.

Para o modelo incondicional, onde o tamanho da ninhada segue uma distribuição de Poisson ou uma distribuição binomial negativa, é também possível demonstrar que as condições de regularidade são satisfeitas. Para o leitor interessado, os Teoremas 2 e 3 a seguir auxiliam nesta demonstração, que não será feita nesta tese.

TEOREMA 2: Os modelos de Poisson e binomial negativo para o tamanho da ninhada e os modelos incondicionais correspondentes são identificáveis.

Prova:

Na demonstração de que o modelo de Poisson para o tamanho da ninhada $P(s,\underset{\sim}{\gamma}) = \dfrac{\exp[-\phi_1 \exp(-\phi_2 d)][\phi_1 \exp(-\phi_2 d)]^s}{s!}$ com $s = 0, 1, \ldots$ $\phi_1 > 0$, $-\infty < \phi_2 < \infty$ e $\underset{\sim}{\gamma} = (\phi_1,\phi_2)$ é identificável, considere válida a relação

$$P(s,\underset{\sim}{\gamma}) = P(s,\underset{\sim}{\gamma}^*) \qquad \forall\, d \geq 0. \qquad (3.17)$$

Podemos derivar o logaritmo da relação (3.17) duas vezes em relação a d, chegando à expressão $\phi_1\phi_2^2\exp(-\phi_2 d) = \phi_1^*\phi_2^{*2}\exp(-\phi_2^* d)$, cujo logaritmo será ainda derivado em relação a d, levando à $\underset{\sim}{\gamma} = \underset{\sim}{\gamma}^*$. Desta maneira, o modelo acima é identificável.■

No caso do modelo binomial negativo para o tamanho da ninhada dado por

$$P(s,\underset{\sim}{\gamma}) = \frac{\Gamma(s+k)}{s!\,\Gamma(k)} \left[\frac{\delta_1 \exp(-\delta_2 d)}{\delta_1 \exp(-\delta_2 d) + k}\right]^s \left[1 + \frac{\delta_1 \exp(-\delta_2 d_i)}{k}\right]^{-k},$$

com $s = 0, 1, \ldots$ $\delta_1 > 0$, $-\infty < \delta_2 < \infty$, $k > 0$ e $\underset{\sim}{\gamma} = (\delta_1, \delta_2, k)$, também provamos sua identificabilidade. Para tanto, considere a relação

$$P(s,\underset{\sim}{\gamma}) = P(s,\underset{\sim}{\gamma}^*) \qquad \forall\, d \geq 0. \qquad (3.18)$$

Derivando o logaritmo de (3.18) em relação a d, chegamos à expressão

$$A + B\exp(\delta_2^* d) + C\exp(\delta_2 d) + D\exp[(\delta_2 + \delta_2^*)d] = 0 \qquad \forall\, d \geq 0 \qquad (3.19)$$

onde $A = \delta_1\delta_1^*(k\delta_2 + k^*\delta_2^*)$ $\qquad B = k^*\delta_1(k\delta_2 - s\delta_2^*)$

$C = -k\delta_1^*(k^*\delta_2^* + s\delta_2)$ $\qquad D = -sk^*k(\delta_2 - \delta_2^*)$

Como a expressão (3.19) só é verificada se A = B = C = D = 0, chegamos a $\underset{\sim}{\gamma} = \underset{\sim}{\gamma}^{*}$.■

Podemos demonstrar agora a identificabilidade dos modelos incondicionais correspondentes

$$P(d,\underset{\sim}{\theta}) = \sum_{s=0}^{\infty} P(d,\underset{\sim}{\psi}|s)\, P(s,\underset{\sim}{\gamma}).$$

Neles, $\underset{\sim}{\psi} = (\alpha, \beta, \theta_1, \theta_2)'$, $P(d,\underset{\sim}{\psi}|s)$ é o modelo condicional RVR e o parâmetro $\underset{\sim}{\gamma} = (\phi_1, \phi_2)'$ ou $\underset{\sim}{\gamma} = (\delta_1, \delta_2, k)'$ se os modelos $P(s,\underset{\sim}{\gamma})$ para o tamanho da ninhada são, respectivamente, Poisson ou binomial negativo e $\underset{\sim}{\theta} = (\underset{\sim}{\psi}, \underset{\sim}{\gamma})$. Utilizando a identificabilidade do modelo condicional RVR e dos modelos de Poisson e binomial negativo para o tamanho da ninhada, pode então ser imediatamente demonstrada a identificabilidade dos modelos incondicionais.

TEOREMA 3: A matriz de informação de Fisher para os parâmetros do modelo de Poisson para o tamanho da ninhada é positiva definida.

A prova deste teorema será feita com o auxílio dos Lemas 4 e 5 a seguir.

Lema 4: O conjunto de funções $T = \{D_k[\phi_1 \exp(-\phi_2 d)],\ k = 1, 2\}$ forma um conjunto de Tchebycheff em $[0, D]$, onde $\phi_1 > 0$ e $-\infty < \phi_2 < \infty$.

Prova:

De acordo com o Lema 1, basta mostrar que a matriz A é de posto completo para todo conjunto de dois pontos distintos d_1 e d_2. Ou seja, basta mostrar que $\det(A) \neq 0$, onde

$$A = \begin{bmatrix} \exp(-\phi_2 d_1) & \exp(-\phi_2 d_2) \\ \phi_1 d_1 \exp(-\phi_2 d_1) & \phi_1 d_2 \exp(-\phi_2 d_2) \end{bmatrix}.$$

Mas $\det(A) = (d_2 - d_1)\, \phi_1 \exp[-\phi_2 (d_1 + d_2)] \neq 0$, pois $\phi_1 > 0$ e $d_1 \neq d_2$. Portanto, o conjunto T é de Tchebycheff.■

Lema 5: Se $T = \{D_k[\phi_1 \exp(-\phi_2 d)],\ k = 1, 2\}$ forma um conjunto de Tchebycheff em [0, D] e se $m \geq 1$, então a matriz de informação $I(\underset{\sim}{\gamma})$ dada em (3.14) é positiva definida.

Prova:

Para $m \geq 1$, prova-se que Σ^{-1} é positiva definida. Basta tomar a matriz A de dimensão 2xm+1 com elementos $a_{ki} = D_k[\phi_1 \exp(-\phi_2 d_i)]$ para $k = 1, 2$ e $i = 1, \ldots, m+1$ e fazer B a matriz diagonal de dimensão m+1xm+1 cujo elemento diagonal é $\left[\frac{n_i}{\phi_1 \exp(-\phi_2 d_i)}\right]^{1/2}$. O restante da demonstração é feita de maneira análoga à demonstração do Lema 3.■

Uma vez demonstrado que são satisfeitas as condições de regularidade para o modelo condicional geral RVR, podemos derivar o comportamento assintótico dos estimadores de máxima verossimilhança de seus parâmetros. Vamos portanto, demonstrar três teoremas relacionados com a existência, unicidade e normalidade assintótica destes estimadores. As provas destes teoremas completam as demonstrações citadas em RAI & VAN RYZIN (1985) e sugeridas no Apêndice de RAI & VAN RYZIN (1981).

3.4.3 PROPRIEDADES ASSINTÓTICAS

TEOREMA 4: Sob a condição C1, as equações de máxima verossimilhança têm assintoticamente uma raiz $\hat{\underset{\sim}{\psi}}$ com probabilidade 1. Esta raiz $\hat{\underset{\sim}{\psi}}$ é um estimador fortemente consistente do parâmetro verdadeiro $\underset{\sim}{\psi}_0$ e $\hat{\underset{\sim}{\psi}}$ é assintoticamente único com probabilidade 1.

Prova:

Seja $\delta > 0$ dado. Considere uma bola fechada de raio δ centrada em $\underset{\sim}{\psi}_0$, denotada por $B = B(\underset{\sim}{\psi}_0, \delta)$. Seja $\underset{\sim}{\psi}$ um ponto qualquer da bola e $\underset{\sim}{\psi}^*$ um ponto qualquer na sua fronteira, denotada por ∂B.

A demonstração deste teorema será dividida em seis partes, de modo a torná-la mais clara.

1ª Parte. Utilizaremos a notação $P_{ij,\underset{\sim}{\psi}_o} = P^o_{ij}$ e $P_{ij,\underset{\sim}{\psi}^*} = P^*_{ij}$ com $P_{ij,\underset{\sim}{\psi}_o}$ e $P_{ij,\underset{\sim}{\psi}^*}$ dados em (3.7), substituindo $\underset{\sim}{\psi}$ por $\underset{\sim}{\psi}_o$ e $\underset{\sim}{\psi}^*$, respectivamente. Seja a variável aleatória $U_{ij} = \dfrac{f(Y_{ij}, \underset{\sim}{\psi}^* \mid S_{ij}=s_{ij})}{f(Y_{ij}, \underset{\sim}{\psi}_o \mid S_{ij}=s_{ij})}$

onde $f(.)$ é a função de probabilidade de Y_{ij} condicional a $S_{ij} = s_{ij}$. Calculando a esperança de U_{ij} quando a distribuição condicional de $Y_{ij}|S_{ij}=s_{ij}$ é uma binomial $bin(s_{ij}, P^o_{ij})$, para todo $j = 1, 2, \ldots, n_i$ e para todo $i = 0, 1, \ldots, m$ e denotando por $E_o(U_{ij})$ tal esperança, resulta

$$E_o(U_{ij}) = E\left[\left(\frac{P^*_{ij}}{P^o_{ij}}\right)^{Y_{ij}}\left(\frac{Q^*_{ij}}{Q^o_{ij}}\right)^{s_{ij}-Y_{ij}}\right]$$

$$= \sum_{y=0}^{s_{ij}} \left(\frac{P^*_{ij}}{P^o_{ij}}\right)^{y}\left(\frac{Q^*_{ij}}{Q^o_{ij}}\right)^{s_{ij}-y}\binom{s_{ij}}{y}(P^c_{ij})^{y}(Q^c_{ij})^{s_{ij}-y}$$

$$= \sum_{y=0}^{s_{ij}} f(Y_{ij} = y, \underset{\sim}{\psi}^* \mid S_{ij} = s_{ij}) = 1.$$

2ª Parte. Mostraremos que se $V_{ij} = \log(U_{ij})$, $\dfrac{1}{n_i}\sum_{j=1}^{n_i} V_{ij} \xrightarrow[n_i\to\infty]{qc}$ $E_o(V_{ij}) < 0$, para $i = 0, \ldots, m$.

Considere a desigualdade de Jensen na versão exposta em RAO (1973) p. 149. Como $S = (0,\infty)$ é um conjunto convexo, U_{ij} é uma

variável aleatória com esperança finita e com distribuição não degenerada, e $f(U_{ij}) = -\log(U_{ij})$ é uma função estritamente convexa, podemos aplicar a desigualdade, o que leva a

$$E_o(-\log(U_{ij})) > -\log(E_o(U_{ij})), \quad \text{ou seja,} \quad E_o(V_{ij}) < 0$$

para $j = 1, 2, \ldots, n_i$ e $i = 0, 1, \ldots, m$.

Desde que para i fixo e s_{ij} conhecido, as variáveis aleatórias Y_{ij} são independentemente distribuídas segundo as leis binomiais $bin(s_{ij}, P_{ij})$, onde P_{ij} só depende da dose d_i, temos que as variáveis aleatórias V_{ij} são também independentemente distribuídas, pois dependem apenas das variáveis Y_{ij}. Além disso, a sequência $(V_{i1}, V_{i2}, \ldots)$ contém variáveis aleatórias com esperança finita e negativa. Então, utilizando uma versão adequada da Lei Forte dos Grandes Números, (veja, por exemplo em JAMES (1981) p. 205-208), temos que a sequência das médias converge com probabilidade 1 para $E_o(V_{ij})$, ou seja,

$$\frac{1}{n_i} \sum_{j=1}^{n_i} V_{ij} \xrightarrow[n_i \to \infty]{qc} E_o(V_{ij}) < 0.$$

3ª Parte. Aqui mostraremos que a sequência $\frac{1}{n}[l(\underset{\sim}{\psi}^*) - l(\underset{\sim}{\psi}_o)] =$

$$\frac{1}{n} \sum_{i=0}^{m} \sum_{j=1}^{n_i} \log[f(Y_{ij}, \underset{\sim}{\psi}^* | S_{ij} = s_{ij})] - \log[f(Y_{ij}, \underset{\sim}{\psi}_o | S_{ij} = s_{ij})] =$$

$$\frac{1}{n} \sum_{i=0}^{m} \sum_{j=1}^{n_i} V_{ij} = \sum_{i=0}^{m} \frac{n_i}{n} \frac{1}{n_i} \sum_{j=1}^{n_i} V_{ij}, \text{ converge quase certamente para um}$$

número negativo.

Utilizando o Lema A4 do Apêndice A e a condição C1, temos

que

$$\frac{1}{n}\,[l(\underset{\sim}{\psi}^*) - l(\underset{\sim}{\psi}_o)] = \sum_{i=o}^{m} \frac{n_i}{n}\,\frac{1}{n_i}\sum_{j=1}^{n_i} V_{ij} \xrightarrow[n_i\to\infty]{qc} \sum_{i=o}^{m} c_i\, E_o(V_{ij}) = \alpha < 0.$$

Incidentalmente, note-se que a quantidade $E_o(V_{ij})$ independe de j.

4ª Parte. Mostraremos a seguir que a convergência quase certa acima pode ser feita uniforme em um conjunto compacto na bola.

A variável aleatória $X_n(\underset{\sim}{\psi}^*) = \frac{1}{n}\,[l(\underset{\sim}{\psi}^*) - l(\underset{\sim}{\psi}_o)]$ depende apenas do ponto $\underset{\sim}{\psi}^*$ de ∂B, pois o centro $\underset{\sim}{\psi}_o$ da bola B é fixo. Assim, da convergência quase certa de $X_n(\underset{\sim}{\psi}^*)$ a $\alpha < 0$, podemos dizer que dado $\varepsilon > 0$, para todo $\lambda > 0$ existe $N = N(\underset{\sim}{\psi}^*, \varepsilon)$ tal que o conjunto $A(m) = \left\{\omega \in \Omega: \sup_{n\geq N} |X_n(\underset{\sim}{\psi}^*,\omega) - \alpha| > \varepsilon \right\}$ é um evento com probabilidade inferior a λ, para todo $m \geq N$, para todo ε e para cada $\underset{\sim}{\psi}^* \in \partial B$. Para eliminar a dependência do índice N com respeito ao parâmetro $\underset{\sim}{\psi}^*$, observemos que a fronteira ∂B da bola B é um conjunto compacto do espaço paramétrico incluído em $\mathbb{R}^4$. Sob a condição C1, as sequências numéricas $X_n(\underset{\sim}{\psi}^*,\omega)$ convergem para α quando ω percorre um evento $M = M(\underset{\sim}{\psi}^*)$ de probabilidade 1.

Vamos mostrar que $X_n(\underset{\sim}{\psi}, \omega) \xrightarrow[n\to\infty]{} \alpha$ uniformemente para todo $\underset{\sim}{\psi} \in S(\underset{\sim}{\psi}^*)$, onde $S(\underset{\sim}{\psi}^*)$ é uma superfície esférica fechada de área positiva $A = A(\delta)$ e centrada em $\underset{\sim}{\psi}^*$. Para cada $\omega \in \Omega$ e para $\underset{\sim}{\psi}$ em $S(\underset{\sim}{\psi}^*)$, $X_n(\underset{\sim}{\psi}, \omega)$ é uma sequência de funções contínuas de $\underset{\sim}{\psi}$. Como $S(\underset{\sim}{\psi}^*)$ é compacto e as sequências $X_n(\underset{\sim}{\psi}, \omega)$ verificam

$$X_1(\underset{\sim}{\psi}, \omega) \geq X_2(\underset{\sim}{\psi}, \omega) \geq X_3(\underset{\sim}{\psi}, \omega) \geq \ldots,$$

o Lema A5 do Apêndice A prova que $X_n(\underset{\sim}{\psi}, \omega) \xrightarrow[n\to\infty]{} \alpha$ uniformemente em $S(\underset{\sim}{\psi}^*)$. Assim, $X_n(\underset{\sim}{\psi}, \omega) \xrightarrow[n\to\infty]{} \alpha$ uniformemente para $\underset{\sim}{\psi} \in S(\underset{\sim}{\psi}_l^*)$ e para $\omega \in M_l$, onde $\{S(\underset{\sim}{\psi}_l^*): l = 1, 2, \ldots, r\}$ é um sub-recobrimento finito do compacto ∂B e M_l é um "conjunto de convergência" com $P(M_l) = 1$, $l = 1, 2, \ldots, r$. Desta forma, a convergência quase certa de $X_n(\underset{\sim}{\psi})$ para α é uniforme para $\underset{\sim}{\psi}^* \in \partial B$.

5ª Parte. Vejamos agora que a sequência $\sup_{\underset{\sim}{\psi} \in B} \frac{1}{n} [l(\underset{\sim}{\psi}) - l(\underset{\sim}{\psi}_o)]$ converge quase certamente a uma quantidade negativa.

Pelo Lema A6 do Apêndice A, $\sup_{\underset{\sim}{\psi} \in S(\underset{\sim}{\psi}_l^*)} X_n(\underset{\sim}{\psi}, \omega) \xrightarrow{n\to\infty} \alpha$ para todo $l = 1, 2, \ldots, r$ e para todo $\omega \in M_l$. Como $\{S(\underset{\sim}{\psi}_l^*): l = 1, 2, \ldots, r\}$ *define um sub-recobrimento finito de* ∂B, *temos que*

$$\sup_{\underset{\sim}{\psi} \in \partial B} X_n(\underset{\sim}{\psi}, \omega) = \max_{1 \le l \le r} \left\{ \sup_{\underset{\sim}{\psi} \in S(\underset{\sim}{\psi}_l^*)} X_n(\underset{\sim}{\psi}, \omega) \right\}$$

e ainda pelo Lema A6, temos que $\sup_{\underset{\sim}{\psi} \in \partial B} X_n(\underset{\sim}{\psi}, \omega) \xrightarrow{n\to\infty} \alpha$, para todo $\omega \in M$. Mas como $P(M) = 1$, resulta que

$$\sup_{\underset{\sim}{\psi} \in \partial B} X_n(\underset{\sim}{\psi}) = \sup_{\underset{\sim}{\psi} \in \partial B} \frac{1}{n} [l(\underset{\sim}{\psi}) - l(\underset{\sim}{\psi}_o)] \xrightarrow[n\to\infty]{qc} \alpha < 0.$$

Temos então que $l(\underset{\sim}{\psi}_o)$ é quase certamente maior que $l(\underset{\sim}{\psi})$ para todo $\underset{\sim}{\psi} \in \partial B$.

Desde que a bola B é compacta e $l(\underset{\sim}{\psi})$ contínua, temos que $l(\underset{\sim}{\psi})$ assume um valor máximo em algum ponto de B, veja BUCK (1965) p. 74. Considere que o máximo de $l(\underset{\sim}{\psi})$ é assumido no ponto denotado por $\hat{\underset{\sim}{\psi}}$ e, portanto, sobre algum raio de B. Mostramos anteriormente que $l(\underset{\sim}{\psi}_o)$

é quase certamente maior que $l(\underset{\sim}{\psi}^*)$ para todo $\underset{\sim}{\psi}^* \in \partial B$. Desta maneira, $\hat{\underset{\sim}{\psi}}$ não se encontra na fronteira da bola, mas sim entre $\underset{\sim}{\psi}_o$ e $\underset{\sim}{\psi}^*$, sobre algum raio. Como δ foi escolhido arbitrariamente, pode-se fazer $\delta \to 0$, obtendo

$$\hat{\underset{\sim}{\psi}} \xrightarrow[n\to\infty]{qc} \underset{\sim}{\psi}_o.$$

Ou seja, o estimador de máxima verossimilhança $\hat{\underset{\sim}{\psi}}$ é um estimador fortemente consistente do verdadeiro valor do parâmetro $\underset{\sim}{\psi}_o$.

6ª Parte. Finalmente, a unicidade do estimador $\hat{\underset{\sim}{\psi}}$ será provada.

Se $\hat{\hat{\underset{\sim}{\psi}}}$ é outra raiz consistente das equações de verossimilhança, temos que $0 = Dl(\underset{\sim}{\psi})\Big|_{\hat{\underset{\sim}{\psi}}} = Dl(\underset{\sim}{\psi})\Big|_{\hat{\hat{\underset{\sim}{\psi}}}}$. Pela extensão multivariada do teorema de Rolle exibida no Lema A7 do Apêndice A, existe um ponto $\underset{\sim}{\psi}^*$ no segmento de reta que une os pontos $\hat{\underset{\sim}{\psi}}$ e $\hat{\hat{\underset{\sim}{\psi}}}$, tal que o determinante de $Hl(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}^*} = 0$. Por outro lado, como $\underset{\sim}{\psi}^*$ é também consistente, $Hl(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}^*}$ converge a uma matriz negativa definida com probabilidade 1. Chegamos portanto a um absurdo, que provém da suposição de existência de outra raiz consistente das equações de verossimilhança. Desta maneira, fica provado que $\hat{\underset{\sim}{\psi}}$ é assintoticamente único com probabilidade 1.■

TEOREMA 5: Sob a condição C1, o vetor aleatório $n^{1/2}(\hat{\underset{\sim}{\psi}} - \underset{\sim}{\psi}_o)$ tem distribuição normal de dimensão 4 com vetor de médias 0 e matriz de covariância regular Σ, onde os elementos de Σ^{-1} são dados por i_{kl} com

$$i_{kl} = \sum_{i=0}^{m} \sum_{j=1}^{n_i} \frac{s_{ij}}{P_{ij} Q_{ij}} (D_k P_{ij}) (D_l P_{ij}) \qquad k, l = 1, \ldots, 4.$$

Prova:

A expansão em série de Taylor de $D l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\hat{\psi}}}$ em torno de $\underset{\sim}{\psi}_0$ é dada por

$$0 = D l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\hat{\psi}}} = D l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0} + H l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}^*} (\underset{\sim}{\hat{\psi}} - \underset{\sim}{\psi}_0) \qquad (3.20)$$

onde $\underset{\sim}{\psi}^*$ é um ponto qualquer no segmento de reta que une $\underset{\sim}{\psi}_0$ e $\underset{\sim}{\hat{\psi}}$.

Da convergência quase certa de $\underset{\sim}{\hat{\psi}}$ a $\underset{\sim}{\psi}_0$ e, do fato de $\underset{\sim}{\psi}^*$ ser um ponto no segmento de reta que une $\underset{\sim}{\psi}_0$ e $\underset{\sim}{\hat{\psi}}$, temos que $\underset{\sim}{\psi}^* \xrightarrow{qc} \underset{\sim}{\psi}_0$.

Pela continuidade de H, temos que $H l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}^*} \xrightarrow{qc} H l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0}$. (3.21)

Mas como a matriz de informação de Fisher é positiva definida, a matriz $H l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0}$ é negativa definida e, portanto, $\det H l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0} \neq 0$. E desde que $\det H l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}^*} \xrightarrow{qc} \det H l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0} \neq 0$, existe a inversa da matriz $H l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}^*}$.

Assim, da equação (3.20) temos

$$n^{1/2}(\underset{\sim}{\hat{\psi}} - \underset{\sim}{\psi}_0) = \left[-n^{-1} H l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}^*}\right]^{-1} n^{-1/2} D l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0}. \qquad (3.22)$$

Para qualquer $\mathbf{b} = (b_1, b_2, b_3, b_4)'$ arbitrário, o uso da versão de Liapunov para o Teorema Central do Limite, exibida em RAO (1973) p. 127, implica que sob a condição C1,

$$n^{-1/2} \mathbf{b}' D l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0} \xrightarrow{D} N(0,\ \mathbf{b}'\Sigma^{-1}\mathbf{b}).$$

Justificaremos a seguir, esta afirmação. Para isto, faça

$$X_n = n^{-1/2} \mathbf{b}' D l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0} = n^{-1/2} \sum_{t=1}^{4} b_t \frac{\partial l(\underset{\sim}{\psi})}{\partial \psi_t}\Big|_{\underset{\sim}{\psi}_0} = n^{-1/2} \sum_{t=1}^{4} b_t D_t\, l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0}$$

e $Z_t = Z_t(Y) = D_t l(\psi)\Big|_{\underset{\sim}{\psi}_0}$, onde para o vetor aleatório Y de dimensão nx1 com componentes y_{ij}, $D_t l(\psi)\Big|_{\underset{\sim}{\psi}_0} = \sum_{i=0}^{m}\sum_{j=1}^{n_i} D_t P_{ij}\,\frac{Y_{ij} - s_{ij}P_{ij}}{P_{ij}Q_{ij}}$ é a derivada da função log-verossimilhança com respeito à componente ψ_t de $\underset{\sim}{\psi}$, para t = 1, ..., 4. Desta forma, $X_n = n^{-1/2}\sum_{t=1}^{4} b_t Z_t$.

À medida que o número $n = \sum_{i=0}^{m} n_i$ de fêmeas presentes no experimento aumenta, X_n define uma sequência de variáveis aleatórias independentes, pois cada X_n é função de variáveis aleatórias independentes Y_{i1}, ..., Y_{in_i}.

No Lema A8 do Apêndice A, é exibida a demonstração de que as condições para a utilização do Teorema Central do Limite na versão de Liapunov são satisfeitas. Assim, aplicando este resultado, temos que

$$Y_n = \frac{\sum_{i=1}^{n} X_i}{\left(\mathrm{Var}\sum_{i=1}^{n} X_i\right)^{1/2}} \xrightarrow{D} N(0, 1) \text{ e assim,}$$

$$n^{-1/2} b' Dl(\psi)\Big|_{\underset{\sim}{\psi}_0} \xrightarrow{D} N(0,\ b'\Sigma^{-1}b).$$

Continuando com a demonstração do Teorema 2, vamos usar o Teorema Central do Limite multivariado de RAO (1973) p. 128, para concluir que

$$n^{-1/2} Dl(\psi)\Big|_{\underset{\sim}{\psi}_0} \xrightarrow{D} N(0,\ \Sigma^{-1}). \qquad (3.23)$$

Reescrevendo a matriz $-n^{-1}Hl(\psi)\Big|_{\underset{\sim}{\psi}^*}$ como

$$\left(n^{-1}Hl(\psi)\Big|_{\underset{\sim}{\psi}_0} - n^{-1}Hl(\psi)\Big|_{\underset{\sim}{\psi}^*}\right) - n^{-1}Hl(\psi)\Big|_{\underset{\sim}{\psi}_0},$$ e lembrando que se

$$W_{ij} = -\left(\frac{y_{ij}}{P_{ij}^2} + \frac{s_{ij} - y_{ij}}{Q_{ij}^2} \right) \frac{\partial P_{ij}}{\partial \psi_k} \frac{\partial P_{ij}}{\partial \psi_l} + \left(\frac{y_{ij}}{P_{ij}} - \frac{s_{ij} - y_{ij}}{Q_{ij}} \right) \frac{\partial^2 P_{ij}}{\partial \psi_k \partial \psi_l},$$

resulta que $-n^{-1}Hl(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0} = -n^{-1}\sum_{i=0}^{m}\sum_{j=1}^{n_i} W_{ij}$ é a média da sequência de variáveis aleatórias W_{ij}, e podemos aplicar novamente uma Lei Forte dos Grandes Números para concluir que $-n^{-1}Hl(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0} \xrightarrow{qc} E(W_{ij}) = \Sigma^{-1}$. Além disso, da convergência quase certa de $n^{-1}Hl(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}^*}$ para $n^{-1}Hl(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0}$, temos que

$$-n^{-1}Hl(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}^*} \xrightarrow{qc} \Sigma^{-1}. \qquad (3.24)$$

Do Teorema de Slutsky, resulta que se

$Z_n = n^{-1/2}Dl(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0} \xrightarrow{D} Z$, onde $Z \sim N(0, \Sigma^{-1})$ e se

$U_n = \left[-n^{-1}Hl(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}^*}\right]^{-1} \xrightarrow{qc} \Sigma$, onde Σ é uma constante, então

$U_n Z_n = \left[-n^{-1}Hl(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}^*}\right]^{-1} n^{-1/2}Dl(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0} \xrightarrow{D} \Sigma Z \sim N(0, \Sigma)$. Assim, finalmente, resulta de (3.22) que $n^{1/2}(\hat{\underset{\sim}{\psi}} - \underset{\sim}{\psi}_0) \xrightarrow{D} \Sigma Z \sim N(0, \Sigma)$.■

Neste capítulo, discutimos algumas propriedades assintóticas dos estimadores de máxima verossimilhança dos parâmetros do modelo geral RVR. No capítulo seguinte, estes resultados serão complementados pelo cômputo destes estimadores em algumas situações concretas.

CAPÍTULO 4

ESTIMAÇÃO DE MÁXIMA VEROSSIMILHANÇA

Para estimar os parâmetros dos modelos estatísticos considerados nesta tese, utilizaremos o método de máxima verossimilhança, que seleciona como estimativa de um parâmetro o valor paramétrico que maximiza a função de verossimilhança. Desta maneira, o procedimento estatístico de estimação por máxima verossimilhança apóia-se em recursos numéricos que permitem maximizar uma função de interesse: a função de verossimilhança.

Neste capítulo, serão brevemente apresentados três procedimentos numéricos iterativos para maximização, ou equivalentemente, minimização de funções. São eles: o método dos gradientes conjugados de Fletcher & Reeves, o método de Newton e o método quase-Newton de Broyden-Fletcher-Goldfarb-Shanno (BFGS). Nas seções seguintes, descreveremos algumas dificuldades numéricas encontradas quando o processo de estimação foi aplicado a dois conjuntos de dados extraídos da literatura. Por fim, serão apresentadas as estimativas dos parâmetros, de seus erros padrão e das

probabilidades de resposta adversa condicional e incondicional. Os Apêndice C, D e E estão diretamente relacionados com este capítulo, como será visto adiante.

4.1 ALGORITMOS

As funções de log-verossimilhança $l(\underset{\sim}{\psi})$ e $l(\underset{\sim}{\gamma})$ introduzidas na seção 3.4 deste trabalho, contêm informações necessárias para a estimação dos parâmetros $\underset{\sim}{\psi}$ do modelo condicional geral RVR e $\underset{\sim}{\gamma}$ do modelo para os tamanhos da ninhada, respectivamente.

Neste capítulo, estas funções serão consideradas casos especiais de uma função de log-verossimilhança geral, que denominaremos $l(\underset{\sim}{\theta})$. O vetor $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_n)'$ coincidirá assim com $\underset{\sim}{\psi} = (\alpha, \beta, \theta_1, \theta_2)$ para o modelo geral RVR, com $\underset{\sim}{\gamma} = (\gamma_1, \gamma_2)' = (\phi_1, \phi_2)'$ para o modelo de Poisson e com $\underset{\sim}{\gamma} = (\gamma_1, \gamma_2, \gamma_3)' = (\delta_1, \delta_2, k)'$ para o modelo binomial negativo.

A estimação do parâmetro $\underset{\sim}{\theta}$ pelo método de máxima verossimilhança equivale a encontrar o maximizador da função log-verossimilhança $l(.)$, satisfeitas as restrições que definem o espaço paramétrico Ω para $\underset{\sim}{\theta}$. No nosso caso, buscaremos solucionar problemas de maximização com canalizações, que consistem em buscar o vetor $\underset{\sim}{\theta}$ que maximiza a função real $l(\underset{\sim}{\theta})$, sujeito a restrições do tipo $\underset{\sim}{\theta} \in \Omega = \{\underset{\sim}{\theta} \in \mathbb{R}^n \mid l_i \leq \theta_i \leq u_i,\ i = 1, \ldots, n\}$. Para todas as funções desta tese, as componentes θ_i do vetor $\underset{\sim}{\theta}$ são irrestritas (ou livres) ou positivas. Na prática, a restrição de positividade de uma componente θ_i será substituída pelo intervalo $[l_i, u_i]$, com limite inferior $l_i = 1.0\text{x}10^{-6}$ e limite superior $u_i = 1.0\text{x}10^{6}$. Quando a componente é livre, seus limites inferiores e superiores serão respectivamente, $-1.0\text{x}10^{6}$ e $1.0\text{x}10^{6}$.

Lembrando que a maximização de $l(\underset{\sim}{\theta})$ é equivalente à minimização de $-l(\underset{\sim}{\theta})$, a partir de agora podemos, sem perda de generalidade, tratar apenas do problema de minimização de funções. Isto será feito porque os programas computacionais e as referências bibliográficas utilizadas adotam esta convenção.

Se $l(\underset{\sim}{\theta}) \geq l(\underset{\sim}{\theta}^*)$ para todo $\underset{\sim}{\theta} \in \Omega$, $\underset{\sim}{\theta}^*$ é dito minimizador global. Se $l(\underset{\sim}{\theta}) \geq l(\underset{\sim}{\theta}^*)$ para todo $\underset{\sim}{\theta}$ numa vizinhança de $\underset{\sim}{\theta}^*$, então $\underset{\sim}{\theta}^*$ é minimizador local. Em termos práticos, qualquer mínimo local será aceito como solução do procedimento iterativo de minimização de uma função.

Os algoritmos de minimização que serão considerados neste trabalho têm uma estrutura comum. A partir de um ponto inicial dado, todos eles determinam uma direção na qual a função objetivo $l(.)$ decresce, movimentando-se nesta direção até um ponto onde a função tenha um decréscimo suficiente (este processo é denominado busca linear). No novo ponto, outra direção é determinada e o processo é repetido, até que um ponto suficientemente perto da solução seja obtido.

De uma maneira mais formal, um algoritmo para minimização de uma função com canalizações pode ser formulado como segue, em termos do vetor gradiente $\nabla l(\underset{\sim}{\theta})$ da função.

Algoritmo básico: Se $\underset{\sim}{\theta}^* \in \mathbb{R}^n$ é uma solução e $\underset{\sim}{\theta}^k$ é uma estimativa de $\underset{\sim}{\theta}^*$ no k-ésimo passo, tal que $\nabla l(\underset{\sim}{\theta}^k) \neq 0$, uma nova estimativa $\underset{\sim}{\theta}^{k+1}$ pode ser assim obtida:

Passo 1: Escolha $\mathbf{d}^k \in \mathbb{R}^n$ tal que $\nabla l(\underset{\sim}{\theta}^k)\mathbf{d}^k < 0$ e exista $\bar{\lambda} > 0$ com $\underset{\sim}{\theta}^k + \lambda \mathbf{d}^k \in \Omega$, $\lambda \in [0, \bar{\lambda})$. Isto é, uma direção factível de descida é selecionada.

Passo 2: Determine o tamanho do passo. Isto é feito calculando um

número λ_k com $0 < \lambda_k < \bar{\lambda}$ tal que $f(\underset{\sim}{\theta}^k + \lambda_k \mathbf{d}^k) < f(\underset{\sim}{\theta}^k)$.

Passo 3: Calcule a nova estimativa, fazendo $\underset{\sim}{\theta}^{k+1} = \underset{\sim}{\theta}^k + \lambda_k \mathbf{d}^k$.

Os algoritmos para minimização de funções com canalizações são baseados em algoritmos sem restrições. Por este motivo, daremos uma noção destes últimos, introduzindo assim os métodos utilizados.

O algoritmo dos gradientes conjugados proposto por FLETCHER & REEVES (1964) toma, em cada passo, uma direção de movimento conjugada com a anterior e que é gerada por uma combinação linear do vetor gradiente no passo atual com a direção no passo anterior. Lembre-se que dada uma matriz simétrica Q, um conjunto finito de vetores $(\mathbf{d}_0, \ldots, \mathbf{d}_k)$ é dito Q-conjugado se $\mathbf{d}_i' Q \mathbf{d}_j = 0$ para todo $i \neq j$.

Apesar de simples, este algoritmo não é muito usado na estimação por máxima verossimilhança. Este fato pode ser devido ao método não fornecer o erro padrão das estimativas dos parâmetros, visto que a matriz Hessiana de derivadas parciais de segunda ordem, denotada por $H[l(\underset{\sim}{\theta})] = \nabla^2 l(\underset{\sim}{\theta})$ não é calculada. Em RAI & VAN RYZIN (1985), porém, este foi o procedimento selecionado para a estimação dos parâmetros do modelo condicional geral RVR e por isto, vamos utilizá-lo.

No caso do método de Newton, a idéia básica é, em cada passo, minimizar a função que é a aproximação quadrática da função objetivo $l(\underset{\sim}{\theta})$, obtida por expansão em série de Taylor ao redor do ponto corrente $\underset{\sim}{\theta}^k$. No k-ésimo passo, a direção de movimento $\mathbf{d}^k$ é obtida pela solução do sistema linear $\nabla^2 l(\underset{\sim}{\theta}^k)\mathbf{d}^k = -\nabla l(\underset{\sim}{\theta}^k)$. Se a matriz hessiana $\nabla^2 l(\underset{\sim}{\theta}^k)$ for positiva definida neste ponto, a direção é de descida.

Uma vantagem do método de Newton é a garantia de sua convergência para pontos iniciais próximos da solução. Uma outra vantagem é sua ordem de convergência ser quadrática; isto é, o erro no passo k+1 é proporcional ao quadrado do erro no passo k. Mais

formalmente, existe $c > 0$ com $\|\underset{\sim}{\theta}^{k+1} - \underset{\sim}{\theta}^{*}\| \leq c \|\underset{\sim}{\theta}^{k} - \underset{\sim}{\theta}^{*}\|^2$. Além disso, o cálculo da derivada segunda pode ser usado na obtenção do erro padrão das estimativas dos parâmetros. Por outro lado, uma desvantagem do método de Newton é seu alto custo computacional, envolvendo a solução de um sistema linear a cada passo iterativo. Além disso, a matriz hessiana pode não ser positiva definida e neste caso, alguma aproximação para a sua inversa deve ser usada na definição da direção do passo. Isto é precisamente o que é realizado pelos métodos quase-Newton.

Os procedimentos denominados quase-Newton são assim uma alternativa econômica ao método de Newton. Existem várias variantes deste método, sendo a forma de aproximação da inversa da hessiana o que varia entre elas. Nesta tese utilizamos o método quase-Newton proposto por Broyden-Fletcher-Goldfarb-Shanno (BFGS).

Uma desvantagem destes métodos para aplicações estatísticas é que a aproximação da matriz hessiana não pode ser usada como estimativa da matriz assintótica de covariâncias dos estimadores.

O leitor que desejar detalhes teóricos sobre estes algoritmos pode consultar, por exemplo, LUENBERGER (1984); GILL et al (1981) e FLETCHER (1987). Para um estudo das relações entre os algoritmos e o procedimento de estimação por máxima verossimilhança, veja THISTED (1988).

No que segue, serão brevemente fornecidas informações sobre o trabalho computacional necessário para obtenção das estimativas de máxima verossimilhança nos modelos estatísticos considerados nesta tese.

4.2 TRABALHO COMPUTACIONAL

Os programas que serão utilizados foram escritos na linguagem FORTRAN 77 e testados no VAX11/785-sistema operacional VMS. Para os métodos quase-Newton e de Newton, estes programas utilizam as subrotinas E04JBF, E04KBF e E04LBF da biblioteca NAG (1982). A

subrotina OPTIM, que é parte do programa MULTI80 desenvolvido por RAI & VAN RYZIN (1980a), foi utilizada para os cálculos com o método de Fletcher & Reeves. Na Seção C1 do Apéndice C, é feita uma breve descrição das subrotinas utilizadas no procedimento de estimação. Os programas correspondentes podem ser consultados no Apéndice E.

Na seção C2 do Apéndice C apresentamos a estrutura da matriz de entrada dos dados, incluindo uma mudança de notação que agiliza o cálculo das funções e a obtenção das estimativas de máxima verossimilhança. Na Seção C3 do Apéndice C encontram-se detalhadas as funções -log-verossimilhança e suas derivadas, na nova notação, para o modelo geral condicional RVR e para os modelos de Poisson e binomial negativo do tamanho da ninhada. Para o modelo de dose-resposta, além da mudança de notação, também foi feita uma transformação de variáveis com o objetivo de simplificar o tipo de restrição a ser satisfeita no procedimento de minimização. Esta transformação de variáveis encontra-se detalhada na seção C3 do Apéndice C.

No que segue, descreveremos resumidamente os experimentos que geraram os dois conjuntos de dados que serão utilizados nesta tese.

4.3 DOIS EXPERIMENTOS

O primeiro conjunto de dados foi apresentado por LÜNING et al (1966). Ele foi produzido em um experimento onde camundongos machos da raça CBA foram tratados com raios X nas doses de 300 e 600 rad (R), tendo sido mantido um grupo controle de animais que não receberam irradiação. Os camundongos foram acasalados com fêmeas da mesma raça nos sete primeiros dias após a irradiação. Note que este é um estudo letal dominante, visto que os machos é que foram diretamente expostos. A resposta adversa observada foi morte fetal.

Os dados são exibidos na Tabela B1 do Apêndice B, e foram também analisados por RAI & VAN RYZIN (1985) e WILLIAMS (1987).

O total de fêmeas no experimento, que foi de 1773 (sendo 683 no grupo controle, 604 no grupo com machos irradiados com 300R e

486 no grupo restante) é muito grande se comparado com a maioria dos experimentos teratológicos. Além disso, as fêmeas com menos de quatro ou mais de onze implantes não foram registradas por Lüning et al (1966), de maneira que o número de implantes é o mesmo (de 5 a 10) em todos os grupos experimentais. Por estes motivos, estes dados não podem ser considerados representativos do que ocorre em um experimento teratológico usual.

Na análise feita nesta tese, assim como na realizada por RAI & VAN RYZIN (1985), foi desconsiderada a fêmea que recebeu irradiação de 600 R e produziu dez implantes, sete dos quais com morte fetal. Isto foi feito porque as observações desta fêmea modificam bastante a proporção de fetos mortos, como pode ser visto na Figura 4.1, obtida com pacote gráfico Microsoft-CHART versão 2.0. Nesta figura, os grupos 0, 1 e 2 representam respectivamente, os grupos de controle e os irradiados com as doses de 300 R e 600 R.

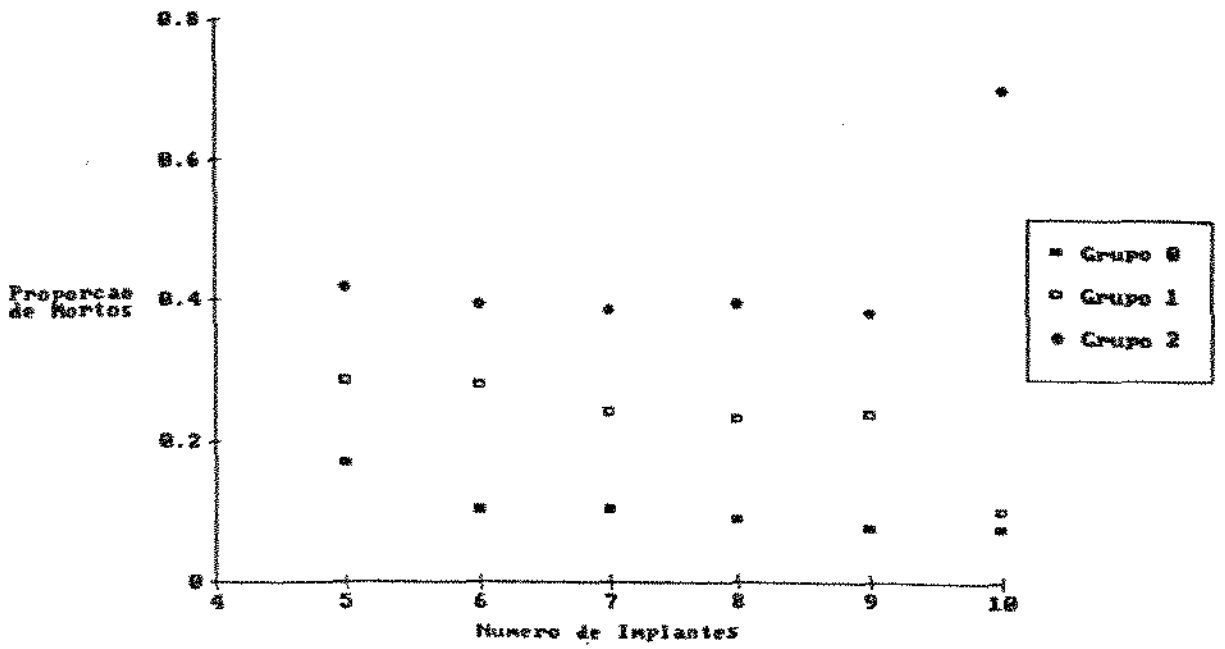


Figura 4.1: Proporção de filhotes mortos de acordo com o número de implantes e grupo experimental. Dados de Lüning et al (1966).

Em cada grupo experimental e para cada número de implantes, as proporções de mortos foram calculadas pelo quociente entre o total de implantes mortos e o total de implantes efetuados. Como exemplo, considere a Tabela B1 do Apêndice B, onde no grupo controle e com 5 implantes existem 71 fêmeas, o que totaliza 355 implantes efetuados. Destes, 27x1 + 9x2 + 5x3 = 60 são mortos. Assim, a proporção de implantes mortos provenientes de fêmeas com 5 implantes do grupo de controle é 60/355 = 0.1690. Procedendo desta maneira, obtemos as proporções restantes e com elas a Figura 4.1.

O segundo conjunto de dados aqui utilizado foi analisado sucessivamente por CHEN & KODELL (1989), FAUSTMAN et al (1989) e RYAN (1992) e gerado em um experimento de TYL et al (1983). Este experimento também pode ser encontrado em outra referência mais acessível, veja TYL et al (1988).

Os dados são provenientes de um experimento para avaliar o efeito da exposição à substância química ftalato de dietilexila (diethylhexyl phthalate ou DEHP, em inglês) em cinco grupos de camundongos fêmeas. O número de fêmeas em cada grupo e as respectivas doses em gramas por kilograma de peso corporal foram: 30 fêmeas no controle, 26 fêmeas em cada um dos grupos expostos às doses de 0.044 e 0.091, 24 fêmeas expostas à dose 0.191 e 25 fêmeas no grupo de dose 0.292. Note que neste caso, as fêmeas são diretamente expostas ao DEHP. O número de implantes foi registrado, assim como outras respostas de interesse: número de fetos ou filhotes afetados (incluindo morte, malformação e absorção, que é uma morte fetal prematura caracterizada por uma pequena marca na parede uterina). Os dados estão na Tabela B2 do Apêndice B.

Este segundo conjunto de dados pode ser considerado mais próximo da realidade do que o anterior: todas as ninhadas geradas foram registradas, e o total de fêmeas do experimento (131) é mais frequente de se ter em experimentos teratológicos.

Na Figura 4.2 pode ser visto o gráfico da proporção de filhotes afetados em função do número de implantes para os diferentes

grupos experimentais do experimento de Tyl et al (1983). Esta figura também foi obtida com o pacote gráfico CHART e as proporções foram calculadas da mesma maneira descrita para o primeiro conjunto de dados. A legenda grupo 0, 1, ..., 4 representa, na ordem, os grupos controle e os grupos expostos à doses crescentes do DEHP.

Examinado a Figura 4.2, não parecem existir pontos que possam estar modificando muito a proporção de afetados nos diferentes grupos. Entretanto, para o grupo exposto à maior dose, praticamente todos os filhotes são afetados, sugerindo que talvez todo este grupo pudesse ser eliminado da análise. Preferimos no entanto, manter todas as observações.

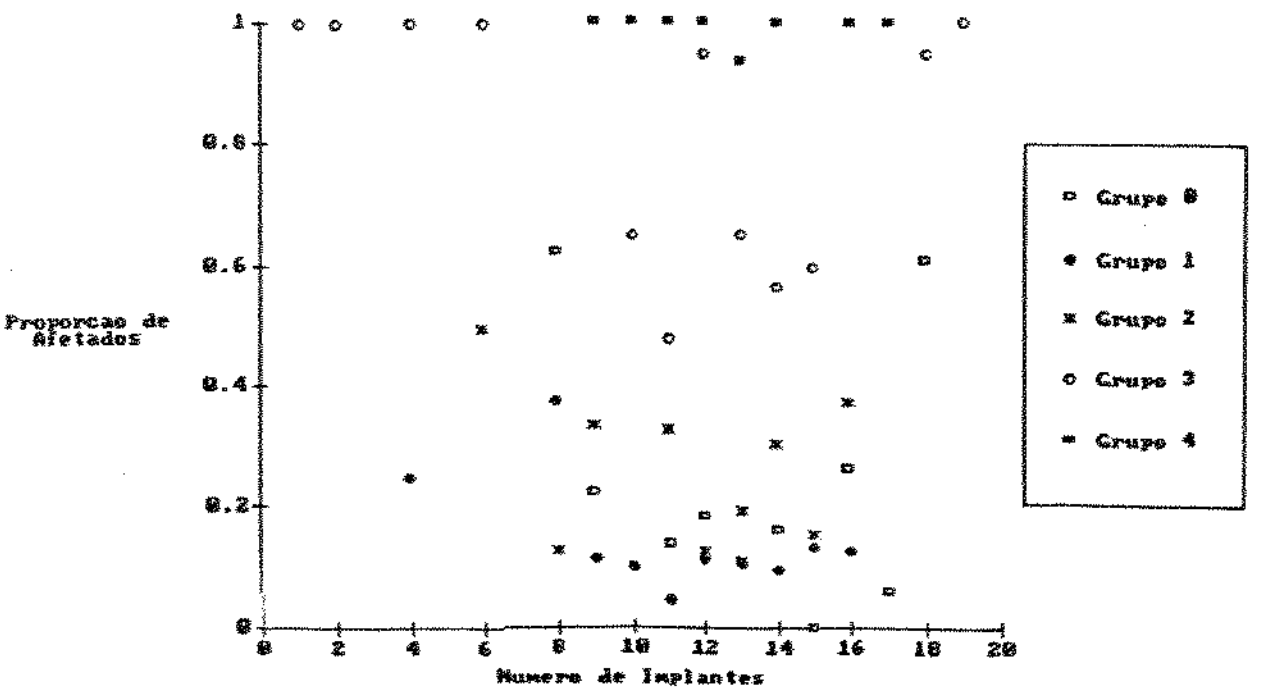


Figura 4.2: Proporção de filhotes afetados de acordo com o número de implantes e grupo experimental. Dados de Tyl et al (1983).

O Apêndice D contém tabelas que visam dar ao leitor uma visão geral dos resultados obtidos na estimação dos parâmetros dos modelos para os dois conjuntos de dados utilizados e os três algoritmos numéricos.

Em todos os casos, não adotamos critérios para a seleção da estimativa do ponto inicial e, na grande maioria das vezes, pontos iniciais bastante diferentes levaram às mesmas soluções. Além dos pontos iniciais considerados nas tabelas do Apêndice D, seção D1 outros foram utilizados mas não são mostrados, visto que os resultados obtidos são semelhantes.

4.4 RESULTADOS

Começaremos esta seção comentando os resultados obtidos aplicando os três procedimentos numéricos aos dados de Lüning et al (1966) para o modelo geral condicional **RVR** e para os modelos de Poisson e binomial negativo para o tamanho da ninhada. A seção D1 do Apêndice D contém as tabelas que resumem estes resultados.

No caso do modelo condicional de dose-resposta, as estimativas obtidas pelos métodos de Fletcher & Reeves, de Newton e quase-Newton, com diferentes pontos iniciais, são bastante próximas, pouco se alterando o mínimo da função -log-verossimilhança. O método de Newton converge mais rapidamente do que os outros, e pontos iniciais mais próximos da solução levam a uma convergência mais rápida em todos os métodos, como exibido na Tabela D1.1 do Apêndice D1.

No caso do modelo de Poisson para o tamanho da ninhada os três métodos levam a estimativas muito próximas, para os diferentes pontos iniciais, além de poucas iterações até a convergência. Mais uma vez o método de Newton é o mais rápido. Entretanto, para alguns pontos iniciais, este método não converge, gerando valores extremamente altos para a função, levando assim à "overflow". Veja a Tabela D1.2 do Apêndice D1.

Nas Figuras 4.3 e 4.4, temos respectivamente as superfícies de verossimilhança $z = -l(.)$ da distribuição de Poisson para o tamanho da ninhada e as correspondentes curvas de nível, obtidas com os dados de Lüning et al (1966). Note a existência de um mínimo desta superfície, e que é encontrado pelos diferentes algoritmos.

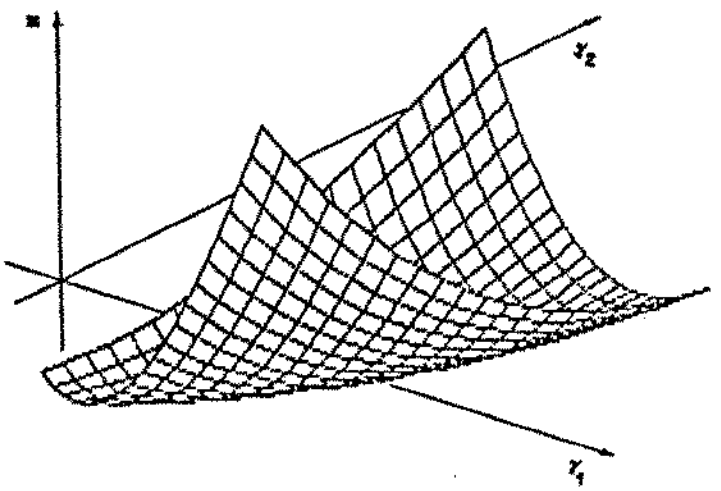


Figura 4.3: Superfície de verossimilhança

$$z = -l(\gamma_1,\gamma_2) = \sum_{k=1}^{n} \left\{ n_k \gamma_1 \exp(-\gamma_2 d_k) - \log(\gamma_1 \exp(-\gamma_2 d_k)) \sum_{i=1}^{i_k} x_{ki} m_{ki} \right\}$$

correspondente à distribuição de Poisson para o tamanho da ninhada. Dados de Lüning et al (1966).

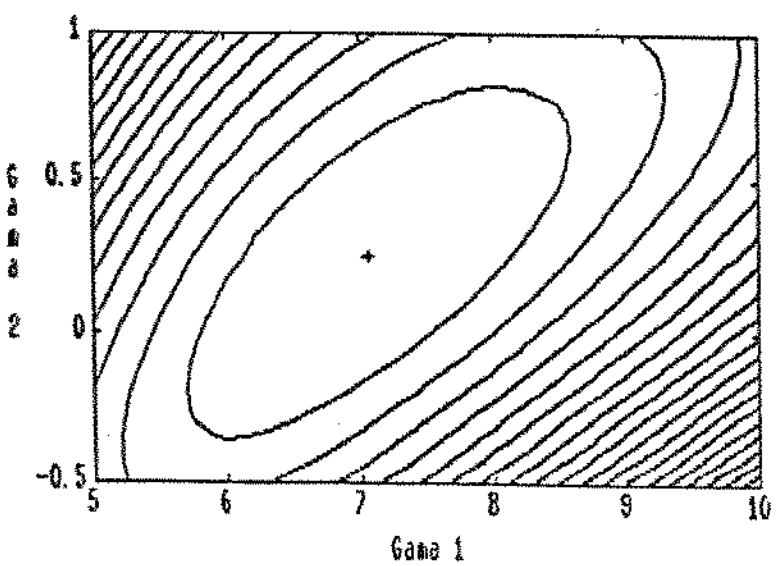


Figura 4.4: Curvas de nível da superfície de verossimilhança da Figura 4.3.

Nesta tese, todas as superfícies de verossimilhança e suas correspondentes curvas de nível foram obtidas através do software matemático PC-MATLAB versão 3.05.

Em relação ao modelo binomial negativo para o tamanho da ninhada, um comportamento especial foi verificado. Com o método de Fletcher & Reeves, os diferentes pontos iniciais praticamente levam às mesmas estimativas para a primeira e a segunda componente do vetor de parâmetros, correspondentes à média da distribuição. No entanto, a terceira componente, correspondente ao parâmetro k, é grande em relação às outras e varia bastante, sem grande variação no valor da função. O número máximo de avaliações da função (32500) é atingido e por este motivo o procedimento é interrompido em todos os casos.

Na tentativa de explicar o comportamento da função com este algoritmo, foram utilizados pontos iniciais com valores altos (de 10^3 a 10^6) para a última componente. Nestes casos, a convergência é rápida, as estimativas das duas primeiras componentes permanecem próximas das anteriores e o valor da última aumenta também, sem alteração notável no valor da função. Isto sugere que o ponto minimizador tem suas duas primeiras componentes próximas dos valores obtidos, com a terceira tendendo ao seu limite superior.

Com o método quase-Newton, as estimativas das duas primeiras componentes do vetor de parâmetros são relativamente próximas, diferindo pouco daquelas obtidas pelo método anterior. A terceira componente, apesar de ser ainda grande em relação às duas primeiras, também difere das obtidas pelo método de Fletcher & Reeves. Os valores da função nos pontos tidos como soluções são maiores do que aqueles obtidos pelo outro método. Isto nos faz crer que as estimativas obtidas não são realmente as melhores. O número de iterações e o número de avaliações da função são bem menores do que no caso anterior e alguns pontos iniciais levam à "overflow".

Em relação ao método de Newton, as estimativas obtidas para as duas primeiras componentes continuam próximas entre si, apesar da terceira variar bastante em função do ponto inicial. Mesmo sendo

pequeno o número de iterações e de avaliações da função, as estimativas obtidas não são aceitáveis, porque os minimizadores geram valores da função maiores do que os obtidos pelos dois métodos anteriores. Os resultados correspondentes ao modelo binomial negativo com os três algoritmos encontram-se na Tabela D1.3 do Apêndice D1.

Para este conjunto de dados, as estimativas obtidas pelos três procedimentos numéricos para o modelo binomial negativo não diferem muito quanto às duas primeiras componentes, apesar da terceira variar bastante. Além disso, as estimativas das duas primeiras componentes são próximas daquelas obtidas para o modelo de Poisson.

Seções da superfície de verossimilhança e as curvas de nível correspondentes à distribuição binomial negativa para o tamanho da ninhada podem ser vistas nas Figuras 4.5 a 4.8. Pelas Figuras 4.6 e 4.8 podemos perceber que inexiste um mínimo da superfície. Possivelmente, a inexistência do mínimo é a causa dos resultados diferentes para os três algoritmos utilizados.

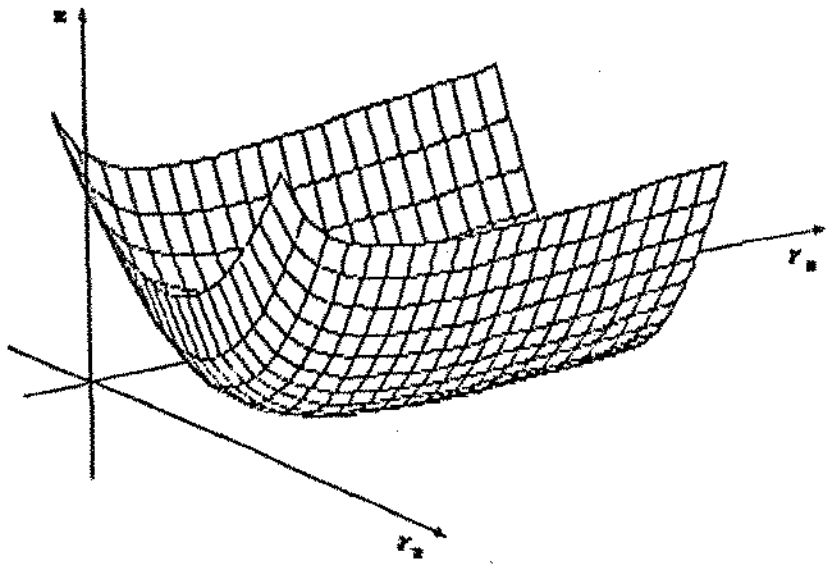


Figura 4.5: Seção $z = -l(7.04,\gamma_2,\gamma_1) = \sum_{k=1}^{n} \left\{ -\sum_{i=1}^{j_k} m_{ki} \log\left[\prod_{r=0}^{a_{ki}-1}(\gamma_1+r)\right] - \right.$

$\left. - \log\left(\frac{7.04\exp(-\gamma_2 d_k)}{7.04\exp(-\gamma_2 d_k)+\gamma_1}\right)\sum_{i=1}^{j_k} a_{ki} m_{ki} + \gamma_1 n_k \log\left(\frac{7.04\exp(-\gamma_2 d_k)+\gamma_1}{\gamma_1}\right)\right\}$

da superfície de verossimilhança correspondente à distribuição binomial negativa para o tamanho da ninhada. Dados de Lüning et al (1966).

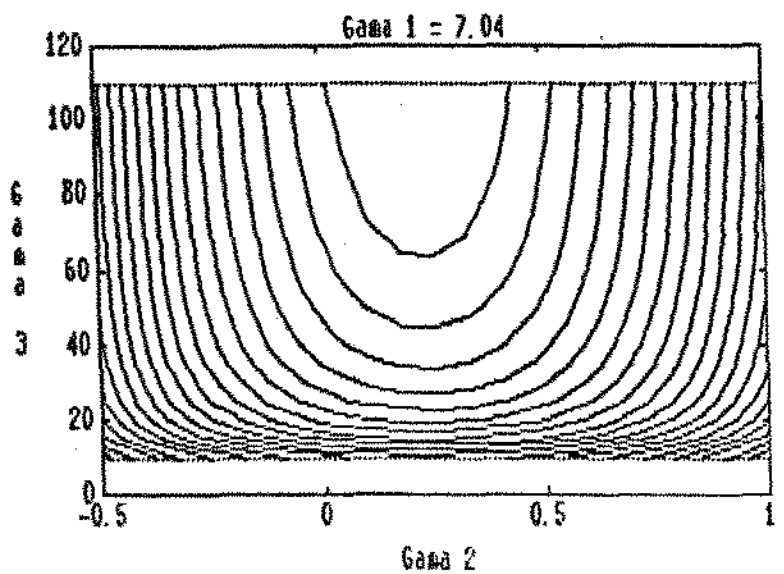


Figura 4.6: Curvas de nível da seção $z = -l(7.04,\gamma_2,\gamma_1)$ da superfície de verossimilhança da Figura 4.5.

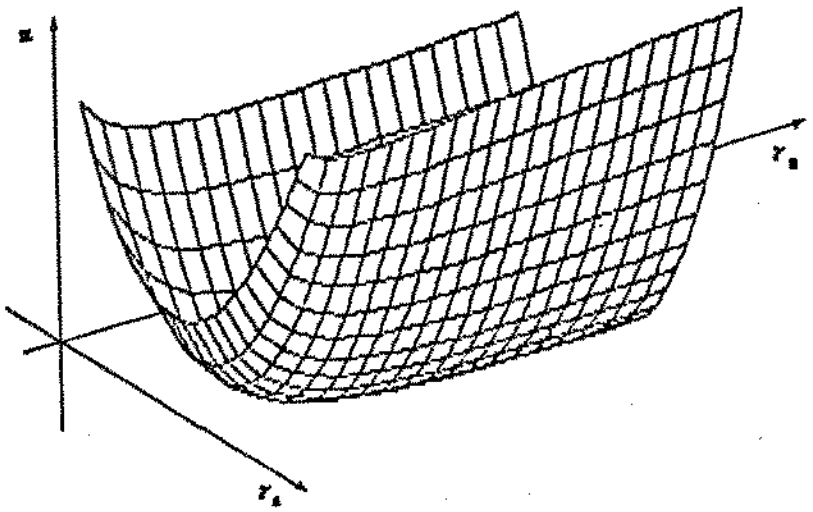


Figura 4.7: Secção $z = -l(\gamma_1, 0.23, \gamma_2) = \sum_{k=1}^{n} \left\{ -\sum_{i=1}^{l_k} m_{ki} \log\left[\prod_{r=0}^{a_{ki}-1}(\gamma_2+r)\right] - \right.$

$\left. - \log\left(\frac{\gamma_1 \exp(-0.23 d_k)}{\gamma_1 \exp(-0.23 d_k) + \gamma_2}\right) \sum_{i=1}^{l_k} a_{ki} m_{ki} + \gamma_2 n_k \log\left(\frac{\gamma_1 \exp(-0.23 d_k) + \gamma_2}{\gamma_2}\right)\right\}$

da superfície de verossimilhança correspondente à distribuição binomial negativa para o tamanho da ninhada. Dados de Lüning et al (1966).

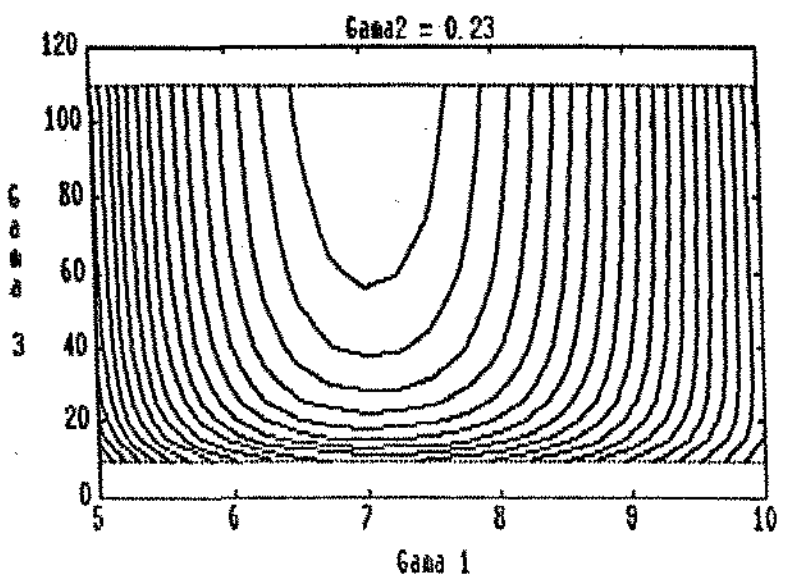


Figura 4.8: Curvas de nível da secção $z = -l(\gamma_1, 0.23, \gamma_2)$ da superfície de verossimilhança da Figura 4.7.

Para os dados de Tyl et al (1983), os resultados que serão condensados a seguir podem ser consultados nas tabelas da seção D1 do Apêndice D.

No caso do modelo de dose-resposta e o método de Fletcher & Reeves, os vários pontos iniciais levam a estimativas diferentes, sendo a última componente da estimativa obtida igual ao limite inferior considerado. Apesar dos diferentes valores da função nos pontos obtidos como soluções, o menor deles é próximo de 726.5, sugerindo que são melhores as estimativas correspondentes a estes valores. Com os métodos de Newton e quase-Newton, todas as estimativas são próximas entre si e daquelas obtidas com o método Fletcher & Reeves que leva ao menor valor da função. Neste caso, o método de Fletcher & Reeves não teve um bom desempenho e, em relação à convergência, o método de Newton é mais rápido. Veja a Tabela D1.4 do Apêndice D1.

Com os três métodos, diferentes pontos iniciais produzem as mesmas estimativas dos parâmetros do modelo de Poisson para o tamanho da ninhada. Pontos iniciais mais próximos da solução convergem mais rapidamente, sendo o método de Newton o mais rápido. Alguns pontos iniciais produzem "overflow" com o último método. Veja a Tabela D1.5 do Apêndice D1.

Nas figuras 4.9 e 4.10 podem ser vistas as superfícies de verossimilhança e as curvas de nível correspondentes à distribuição de Poisson para o tamanho da ninhada para os dados de Tyl et al (1983). O mínimo desta superfície existe e é obtido pelos três algoritmos.

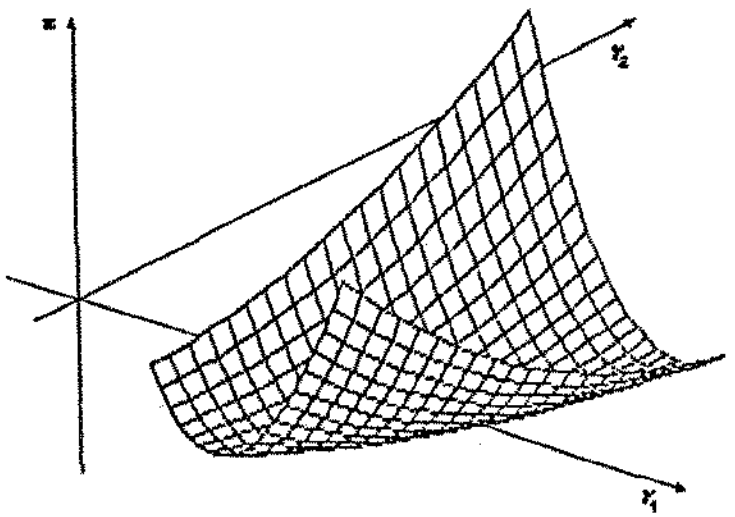


Figura 4.9: Superfície de verossimilhança

$$z = -l(\gamma_1,\gamma_2) = \sum_{k=1}^{n_d} \left\{ n_k \gamma_1 \exp(-\gamma_2 d_k) - \log(\gamma_1 \exp(-\gamma_2 d_k)) \sum_{i=1}^{i_k} m_{ki} \right\}$$

correspondente à distribuição de Poisson para o tamanho da ninhada. Dados de Tyl et al (1983).

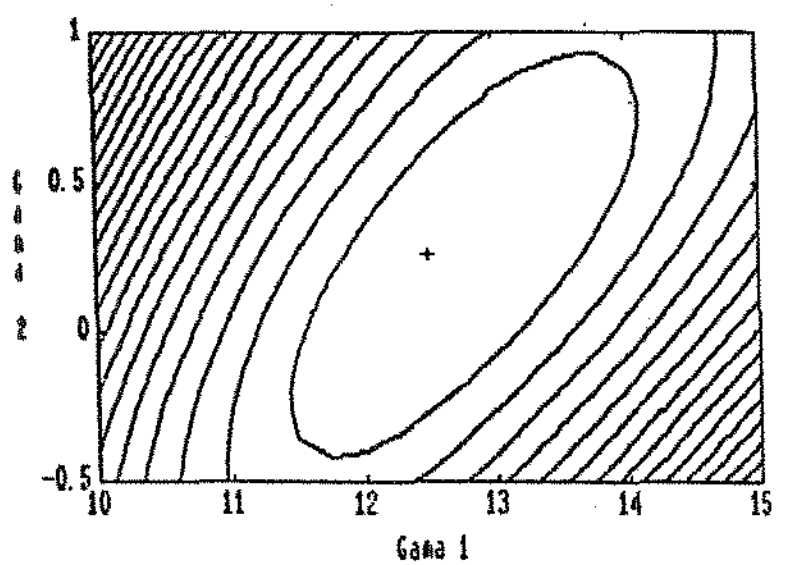


Figura 4.10: Curvas de nível da superfície de verossimilhança da Figura 4.9

Os resultados referentes ao modelo binomial negativo e que serão comentados a seguir, encontram-se na Tabela D1.6 do Apêndice D1.

No caso do método de Fletcher & Reeves, mais uma vez apresentam-se problemas numéricos. A subrotina OPTIM apresentada no Apêndice E e executada no VAX11/785 produz "overflow" para pontos iniciais quaisquer. Em um microcomputador XT com o compilador LAHEY PERSONAL FORTRAN 77 (LP77) versão 2.0, esta mesma subrotina foi executada sem problemas.

Os diferentes pontos iniciais produzem as mesmas estimativas para as duas primeiras componentes, mas a última varia bastante. O valor da função no entanto, praticamente não é alterado. Na maioria das vezes, o número máximo de avaliações da função (32500) é atingido. Pontos iniciais com a terceira componente variando entre 10^3 e 10^6 foram também utilizados: nestes casos a convergência é rápida, as duas primeiras componentes permanecem as mesmas de antes e a estimativa da terceira é o limite superior considerado. Assim, repetiu-se o mesmo comportamento verificado para o primeiro conjunto de dados.

Com o método quase-Newton, foi necessário utilizar a subrotina E04JBF da biblioteca NAG (1982), que calcula o vetor gradiente por diferenças finitas, porque a subrotina E04KBF, que o calcula diretamente e foi usada para todas as outras funções, produz "overflow" para quaisquer pontos iniciais. Os resultados obtidos com E04JBF mostram estimativas próximas daquelas obtidas pelo método de Fletcher & Reeves para as duas primeiras componentes do vetor de parâmetros; a terceira componente permanece igual à mesma componente do ponto inicial dado. O valor da função na suposta solução é maior do que o obtido com o método anterior.

Para o método de Newton, as estimativas das duas primeiras componentes são próximas entre si e daquelas obtidas com os dois métodos anteriores. O valor da terceira componente difere bastante, apesar de pouco mudar em relação à esta mesma componente do ponto inicial. Para todos os pontos iniciais o valor da função no ponto

tido como solução também é maior do que aqueles obtidos com o método de Fletcher & Reeves.

Mais uma vez, com os três algoritmos e para a distribuição binomial negativa, as estimativas obtidas para as duas primeiras componentes do vetor de parâmetros são próximas às correspondentes ao modelo de Poisson. No entanto, a estimativa do parâmetro k do modelo binomial negativo varia bastante. As maiores estimativas de k levam aos menores valores da função e como estamos em busca do seu minimizador, estas, que são obtidas pelo método de Fletcher & Reeves, podem ser consideradas melhores.

Muitos pontos iniciais produzem "overflow" para os métodos de Fletcher & Reeves e quase-Newton, indicando maior sensibilidade do modelo binomial negativo para este conjunto de dados.

Nas Figuras 4.11 a 4.14 são exibidas as seções da superfície de verossimilhança z = -l(.) e as curvas de nível correspondentes à distribuição binomial negativa para o tamanho da ninhada. Pelas Figuras 4.12 e 4.14 percebemos que não existe um mínimo desta superfície. Aqui se repete o comportamento verificado com os dados de Lüning et al (1966): a obtenção de diferentes estimativas para o vetor de parâmetros deve-se possivelmente à inexistência de mínimo.

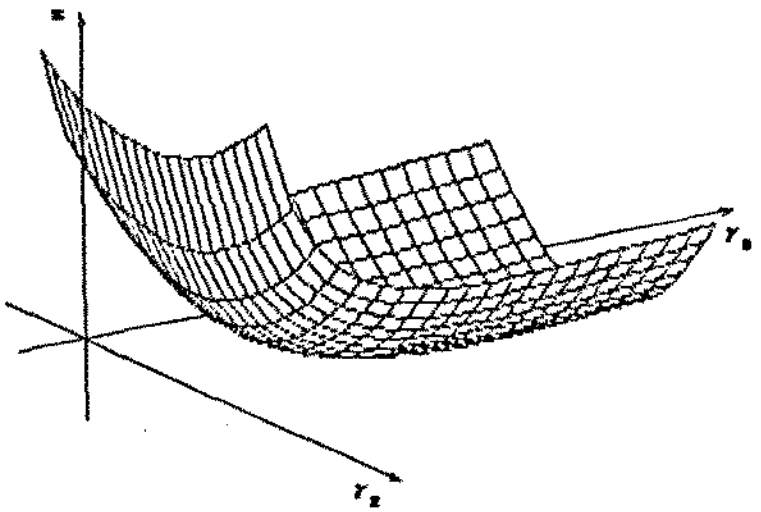


Figura 4.11: Seção $z = -l(12.72,\gamma_2,\gamma_3) = \sum_{k=1}^{n_d} \left\{ -\sum_{i=1}^{i_k} m_{ki} \log\left[\prod_{r=0}^{a_{ki}-1}(\gamma_3+r)\right] - \right.$

$\left. - \log\left(\frac{12.72\exp(-\gamma_2 d_k)}{12.72\exp(-\gamma_2 d_k)+\gamma_3}\right)\sum_{i=1}^{i_k} a_{ki} m_{ki} + \gamma_3 n_k \log\left(\frac{12.72\exp(-\gamma_2 d_k)+\gamma_3}{\gamma_3}\right)\right)$

da superfície de verossimilhança correspondente à distribuição binomial negativa para o tamanho da ninhada. Dados de Tyl et al (1983).

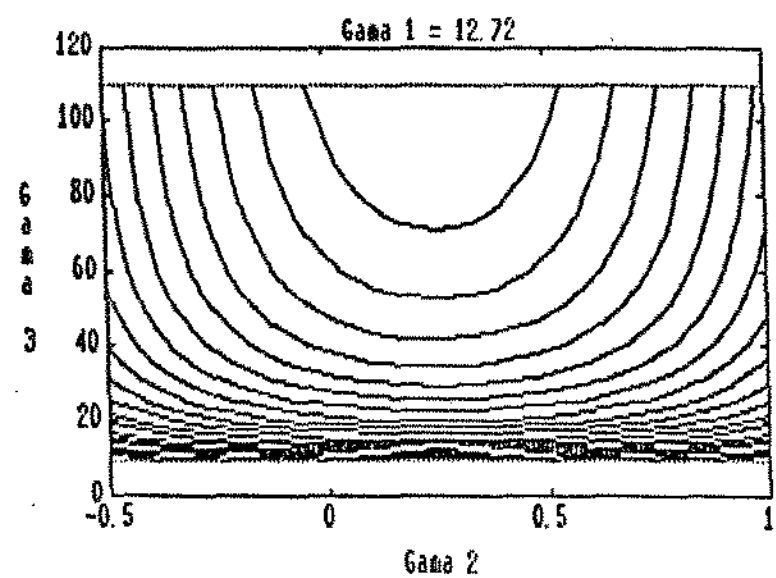


Figura 4.12: Curvas de nível da seção $z = -l(12.72,\gamma_2,\gamma_3)$ da superfície de verossimilhança da Figura 4.11.

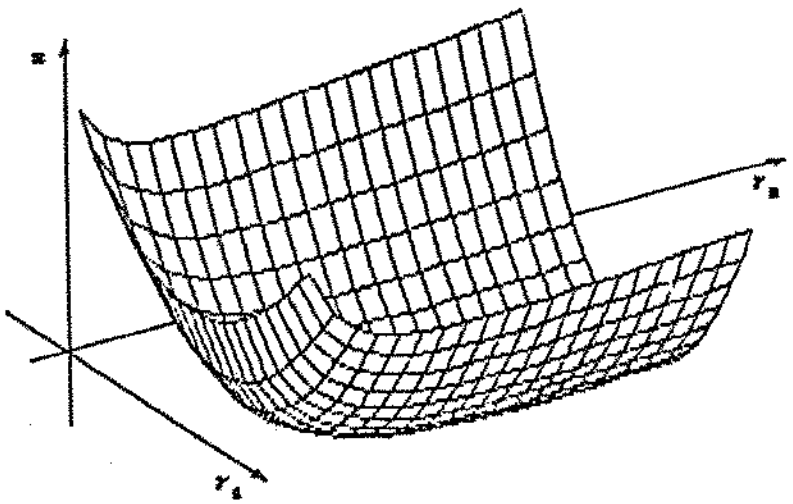


Figura 4.13: Seção $z = -l(\gamma_1, 0.25, \gamma_2) = \sum_{k=1}^{n_d} \left\{ -\sum_{i=1}^{l_k} m_{ki} \log \left[\prod_{r=0}^{a_{ki}-1} (\gamma_2 + r) \right] - \right.$

$\left. - \log \left(\frac{\gamma_1 \exp(-0.25 d_k)}{\gamma_2 \exp(-0.25 d_k) + \gamma_2} \right) \sum_{i=1}^{l_k} a_{ki} m_{ki} + \gamma_2 n_k \log \left(\frac{\gamma_1 \exp(-0.25 d_k) + \gamma_2}{\gamma_2} \right) \right)$

da superfície de verossimilhança correspondente à distribuição binomial negativa para o tamanho da ninhada. Dados de Tyl et al (1983).

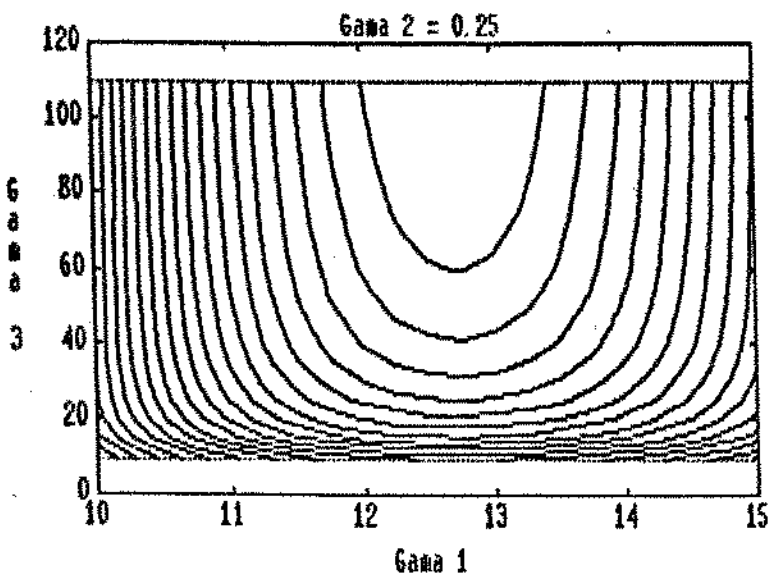


Figura 4.14: Curvas de nível da seção $z = -l(\gamma_1, 0.25, \gamma_2)$ da superfície de verossimilhança da Figura 4.13.

4.5 MODELO DE POISSON VERSUS BINOMIAL NEGATIVO

Na seção 2.2 sugerimos a distribuição binomial negativa para o tamanho da ninhada utilizando a mesma média da distribuição de Poisson proposta por RAI & VAN RYZIN (1985). Nossa sugestão foi feita com o intuito de explicar melhor a variabilidade presente nos dados de ninhada. Entretanto, esta proposta não manifestou-se boa, como pode ser visto a seguir.

Para os dois conjuntos de dados usados nesta tese, a distribuição binomial negativa para o tamanho da ninhada gerou um padrão comum de comportamento. Pelos três algoritmos, o valor estimado da última componente (γ_3 = k) variou bastante, com as duas primeiras (γ_1 e γ_2) bem próximas daquelas estimadas para a distribuição de Poisson, sendo que valores mais altos de γ_3 geraram os menores valores da função. Os algoritmos deram resultados divergentes, pois não detectaram da mesma maneira o ponto minimizador da função -log-verossimilhança, dado por $(\hat{\gamma}_1, \hat{\gamma}_2, \infty)$. Finalmente, seções da superfície de verossimilhança (que não constam neste texto) obtidas para valores altos de γ_3 produziram uma superfície semelhante à de Poisson mostrada nas Figuras 4.3 e 4.9. Lembre-se ainda que um resultado teórico conhecido confirma esta observação numérica: a distribuição de Poisson é um caso limite da distribuição binomial negativa quando o parâmetro k tende ao infinito; veja a demonstração deste fato no Lema A9 do Apêndice A.

Com base no exposto, concluimos que para estes dois conjuntos de dados, não faz sentido o ajuste da distribuição binomial negativa para o tamanho da ninhada, pois na verdade estamos ajustando uma distribuição de Poisson. Desta maneira, a sugestão da distribuição binomial negativa para o tamanho da ninhada não supera a proposta original de Rai & Van Ryzin. Portanto, de agora em diante, esta distribuição não mais será considerada.

Finalmente, vamos apresentar para os dois conjuntos de

dados, as estimativas e seus erros padrão estimados para os parâmetros dos modelos geral condicional **RVR** e de Poisson para o tamanho da ninhada, além de estimativas das probabilidades condicional e incondicional de resposta adversa.

4.6 ESTIMATIVAS

Considere que foram obtidos os estimadores de máxima verossimilhança $\hat{\alpha}$, $\hat{\beta}$, $\hat{\theta}_1$, $\hat{\theta}_2$, $\hat{\phi}_1$, e $\hat{\phi}_2$ dos parâmetros dos modelos de dose-resposta e de Poisson para o tamanho da ninhada. Uma estimativa da probabilidade de um filhote de uma fêmea que produziu uma ninhada de tamanho s_{ij} e foi exposta a uma dose d_i apresentar a resposta adversa é

$$\hat{P}(d_i|s_{ij}) = \{1 - \exp[-(\hat{\alpha} + \hat{\beta}d_i)]\} \exp[-s_{ij}(\hat{\theta}_1 + \hat{\theta}_2 d_i)].$$

Se nosso interesse está na predição do risco de um filhote de uma fêmea exposta a uma pequena dose apresentar a resposta, ou se ele está na estimação da dose correspondente a um risco pré-fixado, pode-se usar a equação acima, caso o tamanho da ninhada seja conhecido. Se o tamanho da ninhada é desconhecido, mas sua distribuição é a de Poisson já estudada, pode-se estimar a probabilidade incondicional de resposta a uma determinada dose d_i pela expressão abaixo, obtida na seção 3.3 deste texto:

$$\hat{P}(d_i) = \{1-\exp[-(\hat{\alpha}+\hat{\beta}d_i)]\} \exp(-\hat{\phi}_1 \exp(-\hat{\phi}_2 d_i)\{1-\exp[-(\hat{\theta}_1+\hat{\theta}_2 d_i)]\}).$$

Utilizando uma versão multivariada da Lei de Propagação de Erros (veja, por exemplo SRIVASTAVA & KHATRI (1979) p.45), podem ser estimadas as variâncias $\mathrm{Var}(\hat{P}(d_i|s_{ij}))$ e $\mathrm{Var}(\hat{P}(d_i))$. Também podem ser construídos intervalos de confiança individuais de coeficiente de $100(1-\alpha)\%$ para as probabilidades condicionais e incondicionais de resposta. Como estes intervalos não são simultâneos, eles não serão incorporados às Figuras 4.15-4.18 adiante. Na verdade, neste trabalho vamos estimar apenas os erros padrão das estimativas dos parâmetros dos modelos **RVR** e de dose-resposta.

A matriz de covariâncias dos estimadores de máxima verossimilhança dos parâmetros é obtida invertendo-se a matriz de informação de Fisher no valor alvo do vetor paramétrico. Assim, para o modelo condicional RVR, basta inverter $I(\underset{\sim}{\psi})$ dada pela expressão (3.12), enquanto que para o modelo de Poisson a matriz a ser invertida é $I(\underset{\sim}{\gamma})$, dada por (3.14).

Tendo sido obtidos $\hat{\underset{\sim}{\psi}}$ e $\hat{\underset{\sim}{\gamma}}$, as estimativas de máxima verossimilhança dos parâmetros, uma estimativa para a matriz de covariância será $I^{-1}(\hat{\underset{\sim}{\psi}})$ para o modelo RVR e $I^{-1}(\hat{\underset{\sim}{\gamma}})$ para a distribuição de Poisson para o tamanho da ninhada. Desta maneira, os erros padrão estimados das componentes de $\hat{\underset{\sim}{\psi}}$ e $\hat{\underset{\sim}{\gamma}}$ são as raízes quadradas dos elementos diagonais de $I^{-1}(\hat{\underset{\sim}{\psi}})$ e $I^{-1}(\hat{\underset{\sim}{\gamma}})$.

Veja na Tabela 4.1, as estimativas dos parâmetros dos modelos RVR e de Poisson, com seus erros padrão estimados, para os dois conjuntos de dados considerados nesta tese.

Tabela 4.1: Estimativas dos parâmetros do modelo condicional RVR e do modelo de Poisson para os dados de Lüning et al (1966) (Conjunto A) e para os dados de Tyl et al (1983) (Conjunto B). Os erros padrão correspondentes figuram entre parênteses.

		Conjunto de Dados	
Modelo	Parâmetro	A	B
RVR	α	0.2238 (0.0365)	0.7918 (0.2859)
	β	1.0377 (0.0167)	9.8035 (4.7277)
	θ	0.0963 (0.0190)	0.1060 (0.1537)
	θ_2^1	-0.0668 (0.0251)	-0.3630 (0.0444)
Poisson	ϕ	7.0412 (0.0945)	12.7223 (0.4685)
	ϕ_2^1	0.2262 (0.0385)	0.2499 (0.2363)

Uma vez obtidas as estimativas dos parâmetros do modelo condicional **RVR**, podemos avaliar a qualidade do ajustamento deste modelo aos conjuntos de dados utilizados, através da estatística de qui-quadrado $T = \sum_{i=0}^{m} \sum_{j=1}^{n_i} \frac{(Y_{ij} - s_{ij}\hat{P}_{ij})^2}{s_{ij}\hat{P}_{ij}\hat{Q}_{ij}}$, apresentada em RAI & VAN RYZIN (1985). Ao nível de significância α, o teste será feito comparando-se T com o correspondente valor crítico $\chi^2_{n-4}(\alpha)$ da distribuição de qui-quadrado com n-4 graus de liberdade. Para os dados de Lüning et al (1966) e Tyl et al (1983), as respectivas estatísticas T são 1983.6891 e 505.6624, e o modelo não se ajusta bem a nenhum dos dois conjuntos de dados, pois as estatíticas T excedem os correspondentes pontos críticos da distribuição de qui-quadrado. Os níveis de significância atingidos são menores que 0.1%.

Na Tabela 4.2 podem ser vistas as estimativas das probabilidades condicionais de resposta adversa, para os dados de Lüning et al (1966). Na Figura 4.15 é apresentado o gráfico desta probabilidade estimada, em função do tamanho de ninhada observado para os diferentes níveis da dose. A Tabela 4.3 e a Figura 4.16 apresentam as mesmas probabilidades, estimadas para os dados de Tyl et al (1983).

Tabela 4.2: Estimativa da probabilidade condicional P(d|s) de resposta adversa em função do nível da dose. Dados de Lüning et al (1966).

	Dose d		
Tamanho s	0.0	0.3	0.6
5	0.1239	0.2830	0.4311
6	0.1125	0.2622	0.4075
7	0.1022	0.2430	0.3853
8	0.0928	0.2251	0.3642
9	0.0843	0.2086	0.3443
10	0.0765	0.1933	0.3255

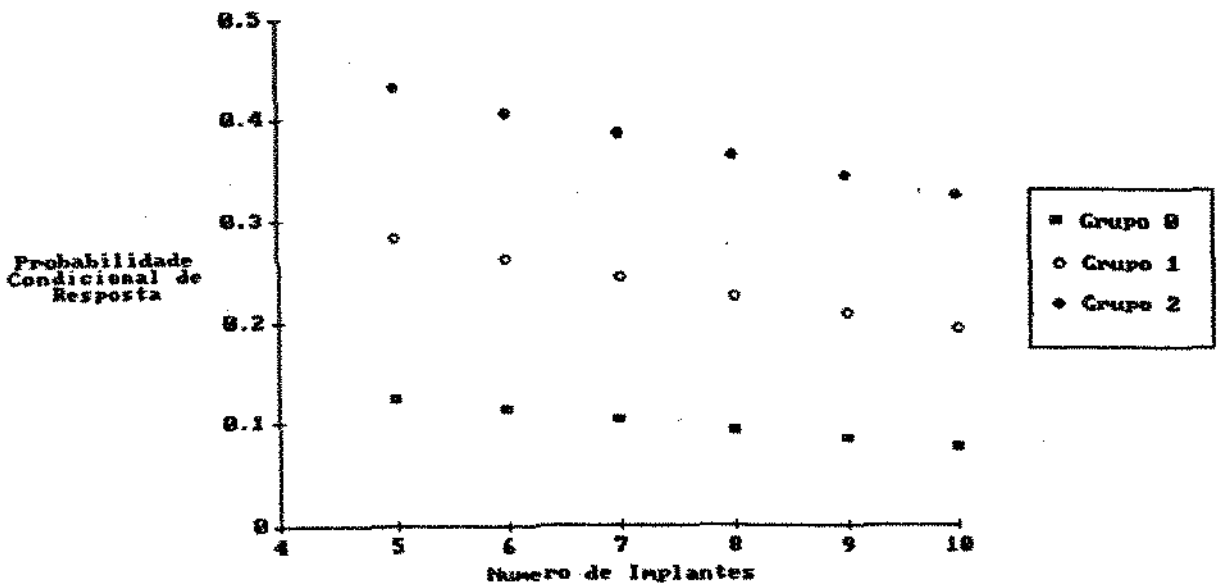


Figura 4.12: Probabilidade de resposta adversa, condicional ao número de implantes efetuados, estimada para os diferentes grupos experimentais. Dados de Lüning et al (1966).

Tabela 4.3: Estimativa da probabilidade condicional P(d|s) de resposta adversa em função do nível da dose. Dados de Tyl et al (1983).

	Dose d				
Tamanho s	0.0	0.044	0.091	0.191	0.292
1				0.8968	
2				0.8646	
4		0.4923		0.8034	
6			0.5256	0.7466	
8	0.2342	0.3434	0.4542		
9	0.2107	0.3138	0.4223		0.97409
10	0.1895	0.2868		0.6448	0.97408
11	0.1704	0.2621	0.3649	0.6215	0.97408
12	0.1533	0.2396	0.3393	0.5992	0.97408
13	0.1379	0.2189	0.3154	0.5776	0.97407
14	0.1240	0.2001	0.2932	0.5568	0.97407
15	0.1115	0.1829	0.2726	0.5368	
16	0.1003	0.1671	0.2534		0.97406
17	0.0902				0.97406
18	0.0811			0.4808	
19				0.4635	

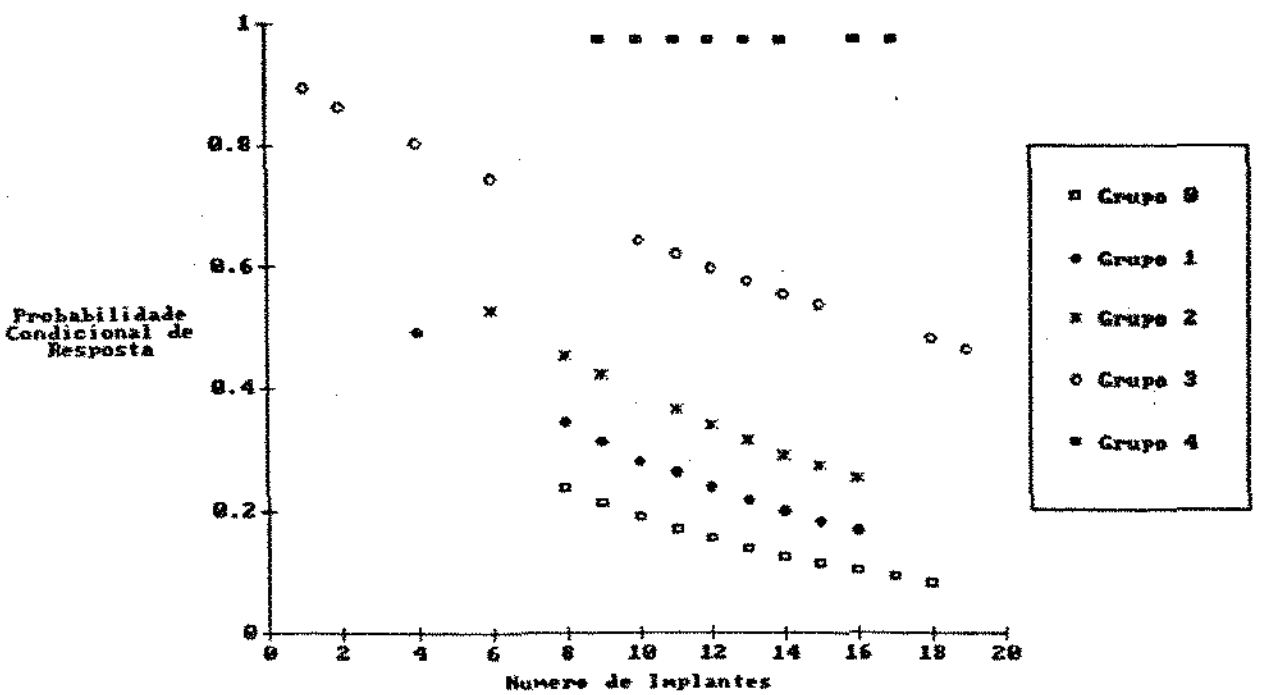


Figura 4.16: Probabilidade de resposta adversa, condicional ao número de implantes efetuados, estimada para os diferentes grupos experimentais. Dados de Tyl et al (1983).

Finalmente, na Tabela 4.4 estão as estimativas das probabilidades de resposta adversa em função do nível da dose, para os dois conjuntos de dados. As Figuras 4.17 e 4.18 apresentam o gráfico destas probabilidades estimadas, em função do nível da dose.

Tabela 4.4: Estimativa da probabilidade incondicional P(d) de resposta adversa para os diferentes níveis de dose d. Dados de Lüning et al (1966) (conjunto A) e Tyl et al (1983), (conjunto B).

Conjunto de Dados	d	$\hat{P}(d)$
A	0.0	0.1050
	0.3	0.2556
	0.6	0.4080
B	0.0	0.1521
	0.044	0.2388
	0.091	0.3394
	0.191	0.6012
	0.292	0.9741

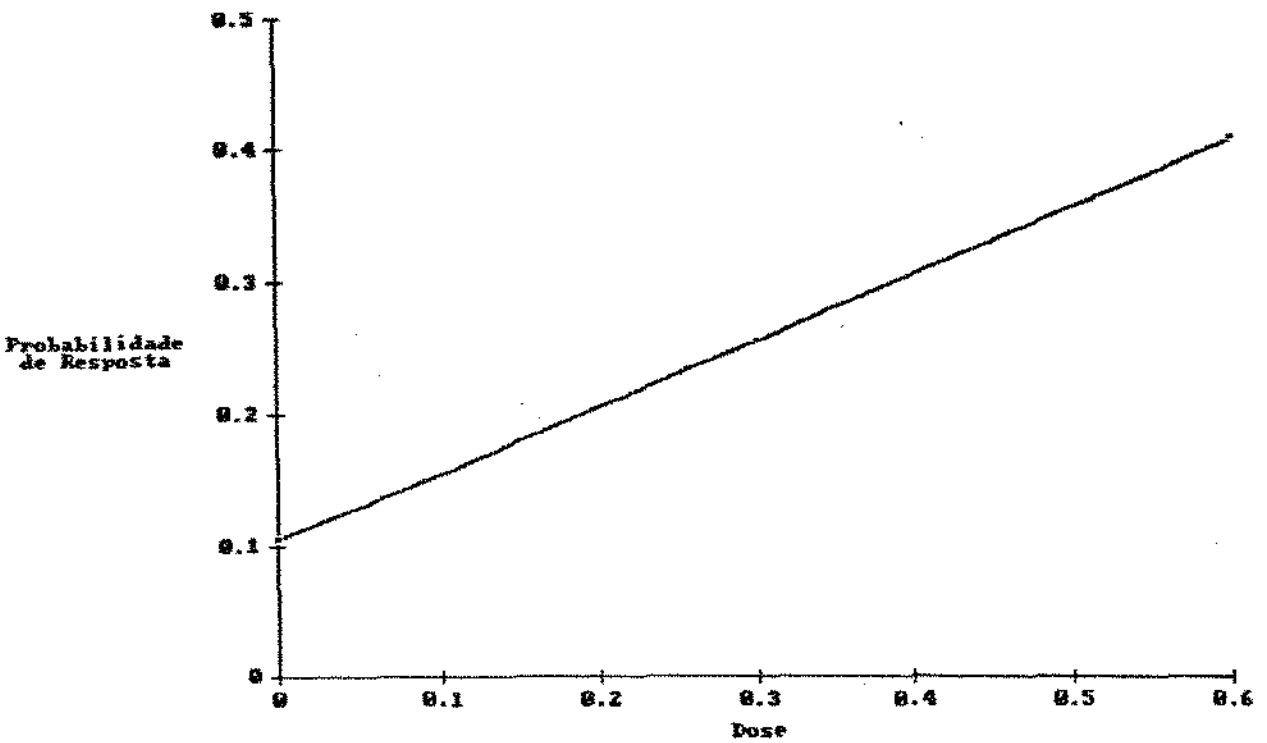


Figura 4.17: Curva estimada da probabilidade incondicional de resposta adversa em função do nível da dose, expressa em rads. Dados de Lüning et al (1966).

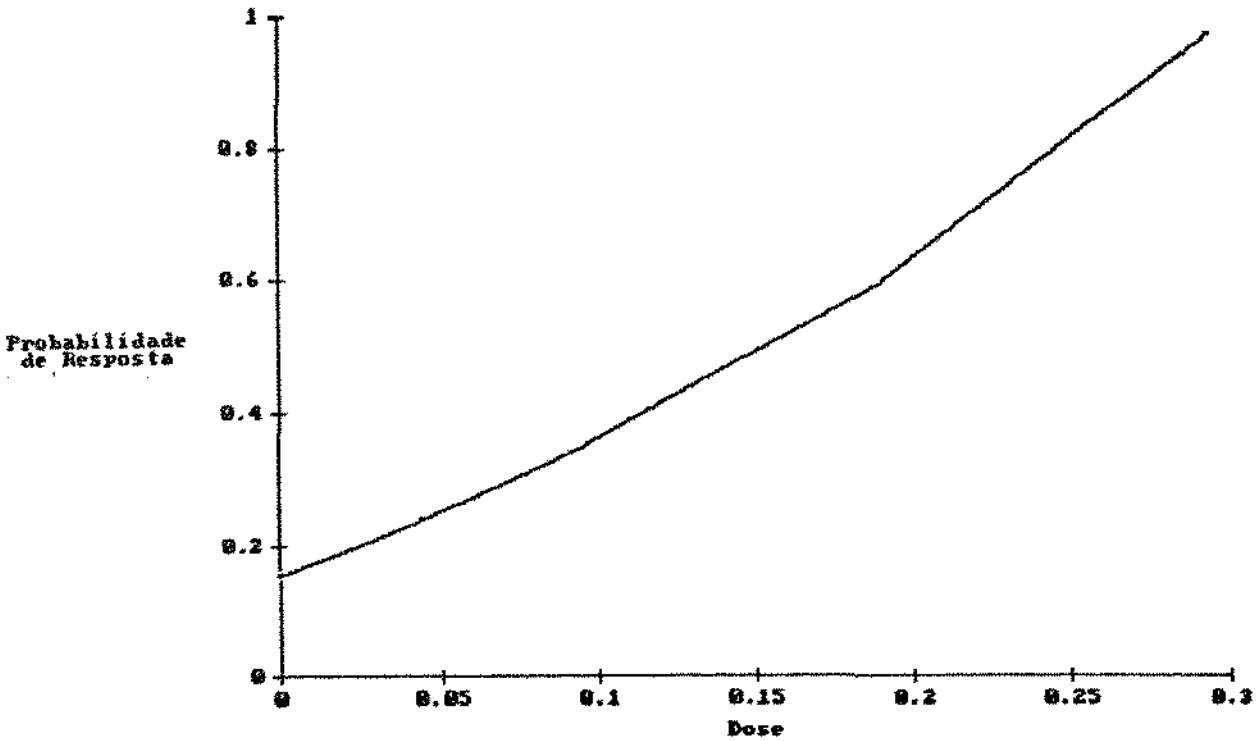


Figura 4.10: Curva estimada da probabilidade incondicional de resposta adversa em função do nível da dose, expressa em g/kg. Dados de Tyl et al (1983).

No próximo capítulo, serão ajustados modelos logísticos que utilizam o tamanho da ninhada como covariável. Estes modelos fazem parte da classe dos modelos lineares generalizados. A existência de um pacote estatístico específico para ajustar esta classe de modelos, facilitará o trabalho computacional correspondente.

CAPÍTULO 5

MODELAGEM LOGÍSTICA

Nos capítulos anteriores, foi discutido o modelo condicional geral RVR proposto por RAI & VAN RYZIN (1985) para análise de dados binários gerados em experimentos teratológicos. Na verdade, ajustar o modelo geral RVR equivale a estimar a probabilidade de que um filhote de mãe exposta a uma dada dose do agente em estudo apresente a resposta adversa, levando em consideração o efeito de ninhada mediante a introdução do correspondente tamanho como covariável.

Neste capítulo, nosso interesse é estimar esta probabilidade de resposta adversa mediante modelos logísticos lineares. No ajuste de tais modelos, será utilizada a teoria de modelos lineares generalizados desenvolvida inicialmente por NELDER & WEDDERBURN (1972) e o pacote estatístico GLIM. Esta teoria tem a vantagem de permitir o tratamento unificado de vários modelos estatísticos, antes estudados separadamente.

Na primeira parte do capítulo, será apresentado um

esquema da teoria dos modelos lineares generalizados, discutindo-se brevemente o correspondente procedimento de estimação, testes de qualidade do ajustamento dos modelos e também alguns resultados assintóticos.

Os modelos logísticos ajustados incluem os sugeridos por WILLIAMS (1987). O refinamento progressivo dos modelos induz uma ordem hierárquica, que será discutida na segunda seção deste capítulo, onde os modelos serão ainda ajustados aos dois conjuntos de dados já utilizados. Além disto, será avaliada a qualidade do ajustamento dos modelos utilizando não apenas o procedimento de Williams, mas também instrumentos gráficos.

O procedimento de estimação apresentado por WILLIAMS (1982) permite ajustar uma variante dos modelos logísticos comuns, que consiste em incorporar a variação extra-binomial. Modelos com esta característica são ajustados na terceira seção do capítulo, onde é também reavaliada a qualidade do ajustamento. Vale notar que estimamos o parâmetro de variação extra-binomial individualmente para cada modelo, enquanto que Williams realiza a estimação apenas para o modelo com maior número de parâmetros, reutilizando a estimativa nos modelos restantes. Também será feita uma comparação entre os modelos logísticos com variação binomial pura e os que possuem variação extra-binomial.

No que segue, começaremos por definir a classe dos modelos lineares generalizados, que inclui os modelos logísticos lineares.

5.1 MODELOS LINEARES GENERALIZADOS

Modelos lineares generalizados formam uma classe de modelos estatísticos especificados por três componentes. São elas: uma componente aleatória, caracterizada pelo fato da distribuição de probabilidade da variável resposta pertencer à família exponencial, uma componente sistemática, que especifica uma estrutura linear com respeito a um conjunto de variáveis explicativas, e uma função de ligação, que relaciona a média da variável resposta com a componente

sistemática.

Na componente aleatória do modelo linear generalizado, supõe-se que o vetor de resposta $Y = (Y_1, \ldots, Y_n)'$ é composto por n respostas Y_i independentes, cada uma das quais possui uma densidade de probabilidade f pertencente à família exponencial, expressa na forma

$$f(y_i; \theta_i, \phi_i) = \exp(\phi_i[y_i\theta_i - b(\theta_i)] + c(y_i, \phi_i)) \qquad (5.1)$$

onde b(.) e c(.) são funções conhecidas, $\phi_i > 0$ é um parâmetro de escala suposto conhecido para cada observação, θ_i é o parâmetro natural (ou canônico) da distribuição e $i = 1, \ldots, n$. A média e a variância de cada resposta Y_i são dadas respectivamente por

$$E(Y_i) = \mu_i = \frac{db(\theta_i)}{d\theta_i} \quad \text{e} \quad Var(Y_i) = \frac{1}{\phi_i}\frac{d^2b(\theta_i)}{d\theta_i^2} \qquad i = 1, \ldots, n.$$

A variância pode também ser descrita mediante $Var(Y_i) = \frac{V_i}{\phi_i}$, onde $V = \frac{d\mu}{d\theta}$ é chamada função de variância.

A componente sistemática relaciona o vetor $\underset{\sim}{\eta} = (\eta_1, \ldots, \eta_n)'$ a um conjunto x_{ij}, $i = 1, \ldots, n$ e $j = 1, \ldots, p$ de variáveis explicativas através da estrutura linear $\underset{\sim}{\eta} = X\underset{\sim}{\beta}$, onde X é uma matriz de elementos x_{ij}, denominada matriz do modelo, de dimensão nxp e suposta de posto completo, e $\underset{\sim}{\beta} = (\beta_1, \ldots, \beta_p)'$ é o vetor de parâmetros de dimensão $p < n$. O vetor $\underset{\sim}{\eta}$ é chamado de preditor linear.

A terceira componente do modelo linear generalizado é uma função de ligação g(.), monótona e diferenciável, que conecta as componentes aleatória e sistemática na forma

$$\eta_i = g(\mu_i) = \sum_{j=1}^{p} x_{ij}\beta_j \qquad i = 1, \ldots, n.$$

Como exemplo, considere a variável resposta $Y = (Y_1, \ldots, Y_n)'$ onde Y_i corresponde ao número de sucessos em m_i ensaios independentes, com $Y_i \sim bin(m_i, \pi_i)$ para todo $i = 1, \ldots, n$. Neste caso, a função de probabilidade da variável aleatória Y_i para $i = 1, \ldots, n$ é dada por

$$f(y_i, \pi_i) = \exp\left[\log\binom{m_i}{y_i} + y_i \log\left(\frac{\pi_i}{1-\pi_i}\right) + m_i \log(1-\pi_i)\right], \qquad (5.2)$$

que pertence à família exponencial com $\phi_i = 1$, parâmetro natural $\theta_i = \text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right)$, $b(\theta_i) = -m_i \log(1-\pi_i)$ e $c(y_i, \phi_i) = \log\binom{m_i}{y_i}$. Tomando o parâmetro natural como função de ligação $g(\mu_i)$, temos o <u>modelo logístico linear</u>

$$\text{logit}(\pi_i) = \sum_{j=1}^{p} x_{ij}\beta_j, \qquad i = 1, \ldots, n. \qquad (5.3)$$

No que segue, vamos descrever resumidamente algumas características da inferência em modelos lineares generalizados.

A estimação dos parâmetros de um modelo deste tipo é usualmente feita pelo método de máxima verossimilhança. As soluções das equações de verossimilhança podem ser obtidas numericamente mediante a iteração de um procedimento de mínimos quadrados ponderados, como demonstrado em NELDER & WEDDERBURN (1972). Felizmente, não é necessário programar este procedimento, pois o pacote estatístico GLIM ("Generalized Linear Interactive Modelling", em inglês) desenvolve esta tarefa. Ele foi desenvolvido pelo Grupo de Computação da Royal Statistical Society, com o objetivo de facilitar o ajuste e a investigação de modelos lineares generalizados. Veja em HEALY (1988), AITKIN et al (1989) e no manual GLIM (1985) esclarecimentos e exemplos sobre a utilização do GLIM.

Quando utilizamos a função de ligação $g(\mu_i) = \text{logit}(\pi_i)$ para a variável resposta binomial, o estimador de máxima

verossimilhança do vetor de parâmetros $\underset{\sim}{\beta} = (\beta_1, \ldots, \beta_p)'$ sempre existe, é único, finito e restrito ao interior do espaço paramétrico $\mathbb{R}$. Tais resultados são provados em WEDDERBURN (1976).

A inferência em um modelo linear generalizado é baseada em resultados assintóticos pois, em geral, não é possível obter as distribuições exatas de estimadores e estatísticas de teste. As condições de regularidade necessárias para garantir os resultados assintóticos são satisfeitas para os modelos lineares generalizados. No caso do modelo binomial, em CORDEIRO (1986) é demonstrado que assintoticamente o estimador $\hat{\underset{\sim}{\beta}}$ tem uma distribuição normal com média $\underset{\sim}{\beta}$ e matriz de covariância $\mathbf{I}^{-1}(\underset{\sim}{\beta})$, onde $\underset{\sim}{\beta}$ é o verdadeiro valor do parâmetro estimado, e $\mathbf{I}(\underset{\sim}{\beta}_0) = - \mathrm{E}\left(\frac{\partial^2 l(\underset{\sim}{\beta})}{\partial \underset{\sim}{\beta}^2} \right) \Big|_{\underset{\sim}{\beta}=\underset{\sim}{\beta}_0}$ é a matriz de informação de Fisher avaliada em $\underset{\sim}{\beta}_0$. A matriz $\mathbf{I}(\underset{\sim}{\beta})$ é consistentemente estimada pela matriz $\mathbf{I}(\hat{\underset{\sim}{\beta}})$, que é não singular quando a matriz **X** é de posto completo. Nestas condições, as estatísticas de Wald

$$(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta})'\mathbf{I}^{-1}(\underset{\sim}{\beta})(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) \quad \text{e} \quad (\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta})'\mathbf{I}^{-1}(\hat{\underset{\sim}{\beta}})(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta})$$

são assintoticamente equivalentes e convergem em distribuição para uma variável aleatória qui-quadrado com p graus de liberdade.

Um modelo linear generalizado é denominado <u>saturado</u> se possui um vetor paramétrico $\underset{\sim}{\beta}_{max}$ de dimensão n. Este modelo atribui toda a variação dos dados à componente sistemática e assim, os reproduz por completo. Os modelos investigados têm vetores paramétricos de dimensão p inferior a n, e são contrastados com o modelo saturado. A qualidade do ajustamento de um modelo linear generalizado é testada pela estatística denominada <u>desvio</u> (<u>deviance</u>, em inglês), que depende da log-verossimilhança $l(\hat{\underset{\sim}{\beta}}_{max};\mathbf{y})$ do modelo saturado e da log-verossimilhança do modelo em investigação $l(\hat{\underset{\sim}{\beta}};\mathbf{y})$. O desvio é definido como

$$D = 2[l(\hat{\beta}_{max};y) - l(\hat{\beta};y)] \qquad (5.4)$$

e sua distribuição pode ser aproximada assintoticamente por uma distribuição de qui-quadrado com n-p graus de liberdade.

Para o modelo binomial (5.2) com índices m_i conhecidos, o desvio é dado por

$$D = 2[l(\hat{\pi}_{max},y) - l(\hat{\pi},y)] = 2\sum_{i=1}^{n} y_i \log\left(\frac{y_i}{m_i\hat{\pi}_i}\right) + (m_i-y_i)\log\left(\frac{m_i-y_i}{m_i(1-\hat{\pi}_i)}\right)$$

onde $\hat{\pi}_i$ é o estimador de máxima verossimilhança de π_i obtido para o modelo em investigação.

Assim, na prática, para testar a qualidade do ajustamento de um modelo com nível de significância α, pode-se comparar D com o correspondente valor crítico $\chi^2_{n-p}(\alpha)$ da distribuição de qui-quadrado com n-p graus de liberdade. Entretanto, esta comparação é somente aproximada.

A análise gráfica é outro procedimento útil na verificação da qualidade do ajustamento de um modelo. Para o modelo binomial, serão utilizados os gráficos (a) do número predito contra o número observado de "sucessos", (b) do número predito de "sucessos" contra os resíduos de Pearson, $rp_i = (y_i - m_i\hat{\pi}_i)/(m_i\hat{\pi}_i(1-\hat{\pi}_i))^{1/2}$, $i = 1, \ldots, n$. Outros gráficos são ainda possíveis, mas não serão utilizados nesta tese.

Se o primeiro gráfico possui um padrão que se aproxima da reta bissetriz do primeiro quadrante, há o indício de bom ajustamento. Se o segundo gráfico não apresentar nenhuma tendência e os resíduos forem pequenos, teremos mais um indício de que o modelo binomial ajustado é adequado. Veja no Capítulo 12 de McCULLAGH & NELDER (1989), detalhes suplementares sobre diagnósticos em modelos lineares generalizados.

Podemos ainda testar hipóteses à respeito do vetor de parâmetros para modelos com a mesma distribuição e função de ligação. Assim, suponha dois modelos do mesmo tipo: o modelo M_0 com parâmetro $\underset{\sim}{\beta}_0 = (\beta_1, \ldots, \beta_q)'$ e desvio D_0, e o modelo M_1 com parâmetros $\underset{\sim}{\beta}_1 = (\beta_1, \ldots, \beta_q, \beta_{q+1}, \ldots, \beta_p)'$ e desvio D_1. O modelo M_0 é dito encaixado no modelo M_1. Para testar a validade do modelo M_0 no contexto de M_1 ou, equivalentemente, a hipótese

$$H_0: \beta_{q+1} = \ldots = \beta_p = 0,$$

pode-se utilizar a diferença entre seus desvios, que coincide com a estatística do teste da razão de verossimilhança. Aqui, se H_0 é verdadeira, a distribuição assintótica de D é uma χ^2_{p-q}. Assim, calculamos a diferença

$$D = D_1 - D_0 = 2[l(\underset{\sim}{\hat{\beta}}_1; y) - l(\underset{\sim}{\hat{\beta}}_0; y)] \qquad (5.5)$$

e se $D > \chi^2_{p-q}(\alpha)$, rejeitamos a hipótese nula ao nível α de significância.

Não é objetivo deste capítulo aprofundar a teoria de modelos lineares generalizados, mas sim apresentá-la resumidamente e utilizá-la. Por este motivo, não nos deteremos mais neste assunto. O leitor que deseja uma introdução à teoria, pode consultar DOBSON (1983). Maiores detalhes podem ser obtidos em CORDEIRO (1986) e McCULLAGH & NELDER (1989).

5.2 HIERARQUIA DE MODELOS LOGÍSTICOS

Nesta seção, será apresentada uma hierarquia de modelos logísticos que inclui os propostos por WILLIAMS (1987). Estes modelos serão ajustados aos dois conjuntos de dados já utilizados nesta tese, sendo também avaliada a qualidade do ajustamento de cada modelo.

Na seção 2.1 está descrita a estrutura do experimento teratológico que gera os dados aos quais os modelos logísticos serão

ajustados. Lembre-se que $Y_{ij}|S_{ij}=s_{ij} \sim bin(s_{ij}, P_{ij})$, onde Y_{ij} é a variável aleatória que representa o número de respostas adversas entre os s_{ij} filhotes produzidos pela j-ésima fêmea do i-ésimo grupo, exposta à dose d_i, para $j = 1, \ldots, n_i$; $i = 0, \ldots, m$ e $0 = d_0 < d_1 < \ldots < d_m$.

Antes de apresentar os modelos logísticos que serão ajustados, é bom lembrar que, da expressão (5.4), podemos calcular o desvio para o modelo condicional RVR estudado no Capítulo 3, onde $P_{ij} = (1 - \exp[-(\alpha+\beta d_i)])\exp[-s_{ij}(\theta_1+\theta_2 d_i)]$ e $Y_{ij}|S_{ij}=s_{ij} \sim bin(s_{ij}, P_{ij})$. Desta maneira,

$$D = 2\sum_{i=0}^{m}\sum_{j=1}^{n_i} y_{ij}\log\left(y_{ij}/s_{ij}\hat{P}_{ij}\right) + (s_{ij}-y_{ij})\log\left((s_{ij}-y_{ij})/s_{ij}\hat{Q}_{ij}\right) \qquad (5.6)$$

e $\hat{P}_{ij}$ indica o valor do parâmetro P_{ij} avaliado nos estimadores de máxima verossimilhança de α, β, θ_1 e θ_2 e $\hat{Q}_{ij} = 1 - \hat{P}_{ij}$.

Incorporando o efeito da dose e do grupo experimental, e utilizando o tamanho da ninhada como covariável, podemos ajustar vários modelos na tentativa de obter bons estimadores da probabilidade de resposta adversa. Nesta tese, vamos ajustar os modelos que serão agora apresentados. Em todos eles, $j = 1, \ldots, n_i$ é o índice correspondente à fêmea do i-ésimo grupo e $i = 0, \ldots, m$ é o índice que identifica o grupo experimental.

O <u>modelo I</u> ajusta um parâmetro para cada combinação grupo e tamanho da ninhada. Ele pode ser representado como

$$\text{logit}(P_{ij}) = \tau_{it} \qquad [I]$$

onde t é o total de diferentes combinações de dose e tamanho de ninhada. Note que este modelo é diferente do saturado, que especifica um parâmetro para cada combinação de grupo e fêmea, ou seja, para cada observação.

Supondo que no modelo anterior, o efeito conjunto da

combinação grupo e tamanho da ninhada pode ser caracterizado por uma regressão linear na covariável tamanho da ninhada, onde intercepto e inclinação variam com o grupo, temos o modelo II

$$\text{logit}(P_{ij}) = \tau_i + \gamma_i s_{ij}. \qquad \text{[II]}$$

Se supomos, por sua vez, que o parâmetro γ_i depende linearmente do nível da dose ($\gamma_i = \gamma + \delta d_i$ para todo i), temos o modelo III dado por

$$\text{logit}(P_{ij}) = \tau_i + \gamma s_{ij} + \delta d_i s_{ij}. \qquad \text{[III]}$$

Também podemos ajustar o modelo IV dado por

$$\text{logit}(P_{ij}) = \tau_i + \gamma s_{ij}, \qquad \text{[IV]}$$

quando as regressões de [III] são paralelas, caracterizando o efeito de cada grupo mediante os interceptos τ_i.

Se supomos que os parâmetros do modelo II dependem linearmente da dose, podemos ajustar o modelo V

$$\text{logit}(P_{ij}) = \alpha + \beta d_i + \gamma s_{ij} + \delta d_i s_{ij}. \qquad \text{[V]}$$

Supondo que a dependência linear da dose está presente no parâmetro τ_i do modelo IV, chegamos ao modelo VI

$$\text{logit}(P_{ij}) = \alpha + \beta d_i + \gamma s_{ij}. \qquad \text{[VI]}$$

Podemos ainda ajustar o modelo VII, com apenas o efeito do grupo

$$\text{logit}(P_{ij}) = \tau_i. \qquad \text{[VII]}$$

Por fim, se este efeito depende linearmente da dose, obtemos o modelo VIII

$$\text{logit}(P_{ij}) = \alpha + \beta d_i. \qquad \text{[VIII]}$$

Os modelos dados pelas expressões [I], [III], [IV] e [V]

foram considerados por WILLIAMS (1987).

Nos modelos acima, observe que o grupo é considerado como fator qualitativo enquanto que a dose é um fator quantitativo. Além disso, a representação do efeito do grupo mediante funções lineares da dose permite reduzir o número de parâmetros do modelo. Poderíamos ainda supor que os parâmetros possuíssem algum tipo de dependência não linear com a dose e desta maneira, outros modelos poderiam ser ajustados. No entanto, nesta tese optamos por considerar apenas a estrutura de dependência linear entre parâmetros e dose.

A Figura 5.1 a seguir apresenta a hierarquia correspondente aos modelos acima definidos, exibindo o número de parâmetros de cada um. Lembre-se que o experimento teratológico que gera os dados aos quais estes modelos serão ajustados tem m+1 grupos experimentais.

Modelo	Número de Parâmetros	Preditor
I	total de diferentes combinações de grupo e tamanho da ninhada	τ_{ij}
II	$2(m+1)$	$\tau_i + \gamma_i s_{ij}$
III	$(m+1)+2$	$\tau_i + \gamma s_{ij} + \delta d_i s_{ij}$
IV	$(m+1)+1$	$\tau_i + \gamma s_{ij}$
V	4	$\alpha + \beta d_i + \gamma s_{ij} + \delta d_i s_{ij}$
VI	$m+1$	τ_i
VII	3	$\alpha + \beta d_i + \gamma s_{ij}$
VIII	2	$\alpha + \beta d_i$

Figura 5.1: Hierarquia de modelos logísticos para o ajustamento de dados de experimentos teratológicos.

Agora, procederemos ao ajuste dos modelos anteriores aos dados de (a) Lüning et al (1966) e (b) Tyl et al (1983) que se encontram no Apêndice B desta tese. Nosso objetivo principal, porém, não é encontrar o melhor ajuste de cada um destes dois conjuntos de dados, senão buscar modelos logísticos parcimoniosos, que com apenas a covariável tamanho da ninhada e o valor da dose estimem razoavelmemte bem a probabilidade de resposta adversa para um filhote de mãe exposta àquela dose.

Observe que o desvio para o modelo condicional RVR pode ser calculado pela expressão (5.6), usando as estimativas de máxima verossimilhança $\hat{\alpha} = 0.2238$, $\hat{\beta} = 1.0377$, $\hat{\theta}_1 = 0.0963$ e $\hat{\theta}_2 = -0.0668$ encontradas no Capítulo 4. Temos portanto, o desvio D = 2244.2 com 1768 graus de liberdade. Mais adiante, este valor calculado para o desvio será usado para testar a qualidade do ajustamento do modelo RVR, além de compará-lo com os modelos logísticos.

5.2.1 AJUSTAMENTO AOS DADOS DE LÜNING et al

As estimativas dos parâmetros dos modelos I a VIII e os desvios padrão das estimativas foram obtidas aplicando o pacote estatístico GLIM versão 3.77. Elas encontram-se na Tabela D.2.1 do Apêndice D.

A Tabela 5.1 a seguir mostra desvios e graus de liberdade obtidos no ajustamento dos modelos logísticos aos dados de Lüning et al (1966).

Tabela 5.1: Modelos logísticos, seus desvios e correspondentes graus de liberdade obtidos para os dados de Lüning et al (1966). Os níveis de significância atingidos são em todos os casos inferiores a 0.1%.

Modelo	Desvio	Graus de Liberdade
I	2227.5	1755
II	2239.0	1766
III	2239.0	1767
IV	2244.8	1768
V	2255.6	1768
VI	2270.1	1769
VII	2266.9	1769
VII	2292.2	1770

Se calculamos (veja PEARSON & HARTLEY (1970)) os valores críticos da distribuição de qui-quadrado com mais de 100 graus de liberdade mediante a expressão

$$\chi^2_\nu(\alpha) = \nu\left[1 - \frac{2}{9\nu} + X_\alpha\sqrt{\frac{2}{9\nu}}\right]^3, \qquad \nu > 100 \qquad (5.7)$$

onde X_α é o ponto da distribuição normal padrão correspondente ao nível de significância α, concluímos que nenhum dos modelos logísticos propostos, nem o modelo condicional RVR, se ajustam bem aos dados de Lüning et al (1966), pois todos os desvios excedem os correspondentes pontos críticos da qui-quadrado.

Examinando os gráficos do número predito $\hat{y}_{ij}$ contra o observado y_{ij} de respostas adversas, percebemos que os modelos não parecem adequados, pois em todos os casos, há um afastamento considerável da reta $\hat{y}_{ij} = y_{ij}$. Analogamente, os gráficos de $\hat{y}_{ij}$ contra os resíduos de Pearson mostram também a baixa qualidade do ajustamento dos diversos modelos: os resíduos são grandes e os pontos $(\hat{y}_{ij}, rp_{ij})$ mostram três aglomerações diferenciadas.

Nas Figuras 5.2 e 5.3 são exibidos estes dois gráficos para o modelo VII. Desde que eles podem ser considerados representativos do que acontece com os outros modelos, serão os únicos apresentados.

Todos os gráficos deste capítulo foram obtidos pela aplicação do pacote estatístico GLIM, onde os pontos únicos são representados por um asterisco, dois ou mais pontos coincidentes são representados pela frequência de pontos coincidentes e o número 9 indica nove ou mais pontos coincidentes.

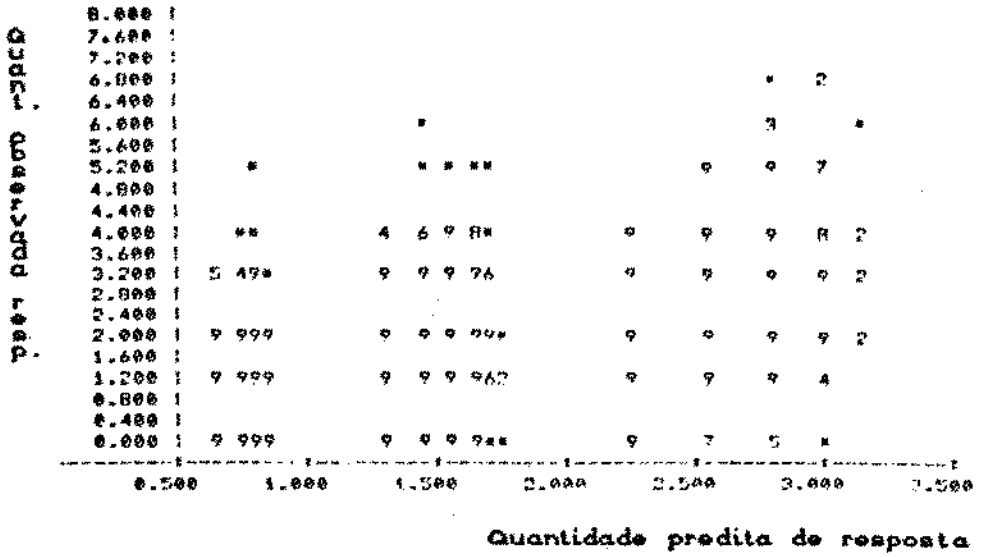


Figura 5.2: Gráfico da quantidade predita contra a quantidade observada de respostas adversas para o modelo $\text{logit}(P_{ij}) = \alpha + \beta d_i + \gamma z_{(j)}$ ajustado aos dados de Lüning et al (1966).

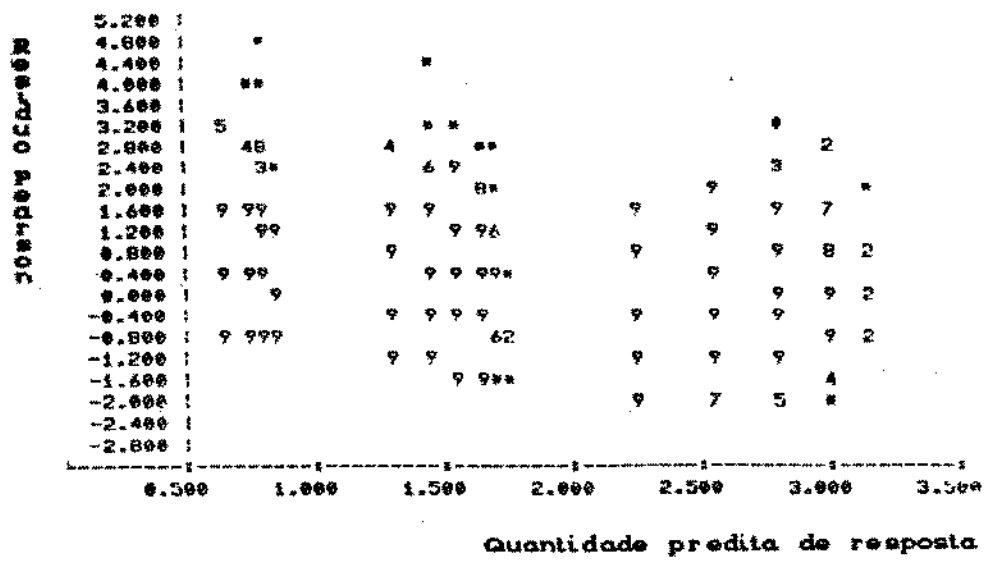


Figura 5.3: Gráfico da quantidade predita de respostas adversas contra o resíduo de Pearson para o modelo $\text{logit}(P_{ij}) = \alpha + \beta d_i + \gamma z_{(j)}$ ajustado aos dados de Lüning et al (1966).

É bom ressaltar que a rigor, só poderíamos fazer testes de hipóteses à respeito do vetor de parâmetros de um modelo encaixado em outro se o ajuste de ambos os modelos fosse considerado adequado. Entretanto, apesar dos modelos não se ajustarem bem aos dados, testaremos alguns modelos contra outros (encaixados ou não), na tentativa de reproduzir a análise feita em WILLIAMS (1987).

Na Tabela 5.2 encontram-se os modelos correspondentes às hipóteses testadas, o valor da diferença entre os desvios, os graus de liberdade e os correspondentes níveis de significância atingidos.

Tabela 5.2: Resultados dos testes dos modelos ajustados aos dados de Lüning et al (1966).

Modelo M_0	Modelo M_1	Diferença entre Desvios	Graus de Liberdade	Nível de Significância
RVR	I	16.7	13	5% < P < 50%
IV	III	5.8	1	1% < P < 2.5%
V	III	16.6	1	P < 0.5%
VI	IV	25.3	1	P < 0.5%
VII	V	11.3	1	P < 0.5%
VIII	VII	25.3	1	P < 0.5%

Testando o modelo RVR contra o modelo de 17 parâmetros (um parâmetro para cada combinação de dose e tamanho de ninhada), verificamos informalmente a existência de evidência a favor do modelo RVR. O teste do modelo IV contra o modelo III, parece favorecer a presença do termo com a covariável $d_i s_{ij}$. Consequentemente, as regressões no tamanho da ninhada não podem ser consideradas paralelas. Comparando os modelos V e III, as evidências são contra o primeiro modelo, indicando que o efeito do grupo do modelo III parece não depender linearmente da dose. Da comparação entre o modelo VI e o modelo IV concluímos que apenas o efeito de grupo não descreve bem os dados, sendo necessária a presença do tamanho da ninhada como regressor. Testando o modelo VII contra o modelo V, vimos que é importante a contribuição de $d_i s_{ij}$. Do teste do modelo VIII contra o modelo VII concluímos que a covariável s_{ij} também parece necessária,

havendo portanto, evidência favorável ao modelo VII.

Seria conveniente analisar o comportamento destes modelos logísticos para outros conjuntos de dados. Por isto, os ajustes anteriores serão aplicados aos dados de Tyl et al (1983).

Antes porém, indiquemos o desvio para o modelo condicional RVR: com as estimativas de máxima verossimilhança $\hat{\alpha}$ = 0.7918, $\hat{\beta}$ = 9.8035, $\hat{\theta}_1$ = 0.1060 e $\hat{\theta}_2$ = -0.3630, obtem-se um desvio de 462.56 com 127 graus de liberdade. Utilizando este critério, concluimos portanto, que o modelo RVR não se ajusta bem a estes dados.

5.2.2 AJUSTAMENTO AOS DADOS DE TYL et al

Para os dados de TYL et al (1983), as correspondentes estimativas dos parâmetros e seus desvios padrão para os modelos I a VIII figuram na tabela D2.2 do Apêndice D.

Na Tabela 5.3 a seguir encontram-se os valores dos desvios e dos graus de liberdade correspondentes para os modelos logísticos ajustados aos dados de Tyl et al (1983).

Tabela 5.3: Modelos logísticos, seus desvios e correspondentes graus de liberdade obtidos para os dados de Tyl et al (1983). Os níveis de significância atingidos são em todos os casos inferiores a 0.1%.

Modelo	Desvio	Graus de Liberdade
I	244.34	81
II	376.49	121
III	379.76	124
IV	379.93	125
V	441.34	127
VI	380.86	126
VII	445.28	128
VIII	447.94	129

Utilizando o desvio para testar se os modelos são adequados, temos que também para este conjunto de dados, nenhum dos modelos logísticos propostos parece fornecer um bom ajuste, pois os desvios superam os correspondentes pontos críticos da distribuição de qui-quadrado.

Em todos os gráficos do número predito $\hat{y}_{ij}$ versus o observado y_{ij} de respostas adversas, os pontos apresentam praticamente a mesma tendência. Podemos considerar que os pontos estão distribuídos ao redor da reta $\hat{y}_{ij} = y_{ij}$, mas não muito próximos dela, indicando um ajuste sofrível. Por outro lado, os gráficos dos resíduos de Pearson contra o número predito de respostas também apresentam pontos com praticamente o mesmo padrão. Há muitos resíduos grandes e podemos notar uma tendência entre os pontos $(\hat{y}_{ij}, rp_{ij})$ de formarem duas aglomerações. Em particular, para os modelos II, III, IV e VI (que envolvem o efeito do i-ésimo grupo) a observação número 121 (i = 4, d_i = 0.292, s_{ij} = 13 e y_{ij} = 8) produz valores muito elevados do resíduo. Quando esta observação é retirada e o ajustamento refeito, não é obtida uma melhora perceptível.

Estes dois gráficos correspondentes aos modelos I, IV e V encontram-se nas Figuras 5.4 a 5.9 a seguir.

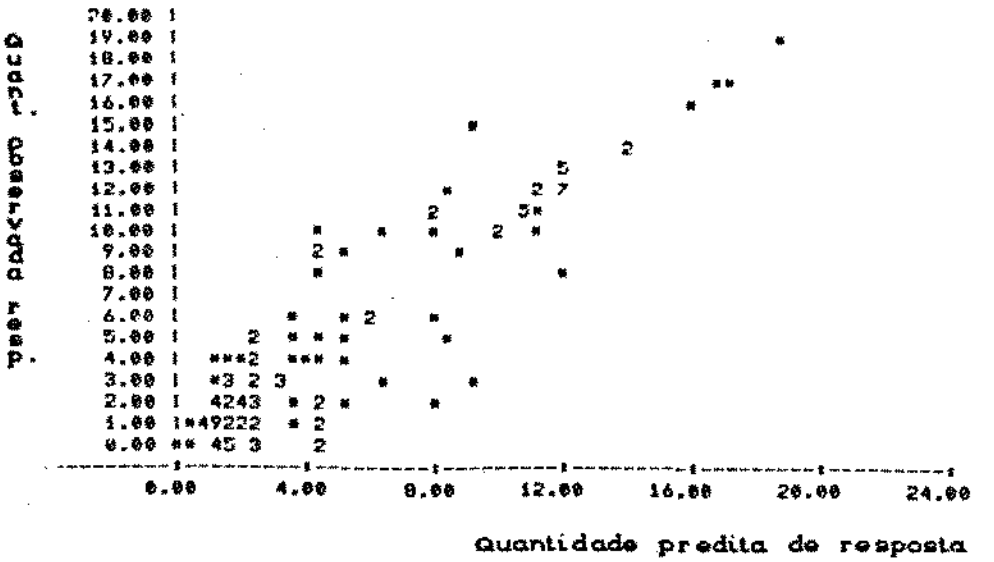


Figura 5.4: Gráfico da quantidade predita de respostas adversas contra a quantidade observada para o modelo $\text{logit}(P_{ij}) = \tau_{ij}$ ajustado aos dados de Tyl et al (1983).

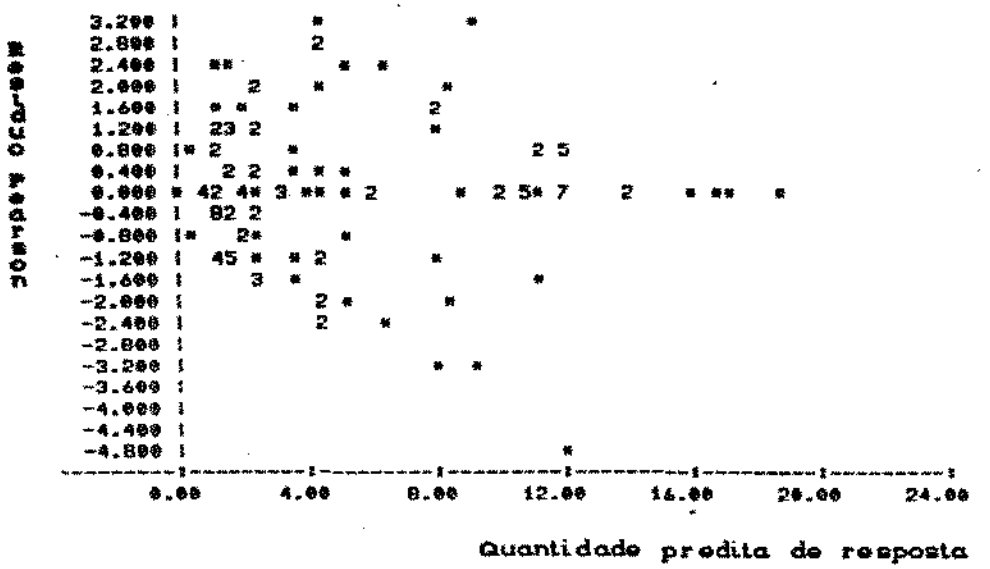


Figura 5.5: Gráfico da quantidade predita de respostas adversas contra o resíduo de Pearson para o modelo $\text{logit}(P_{ij}) = \tau_{ij}$ ajustado aos dados de Tyl et al (1983).

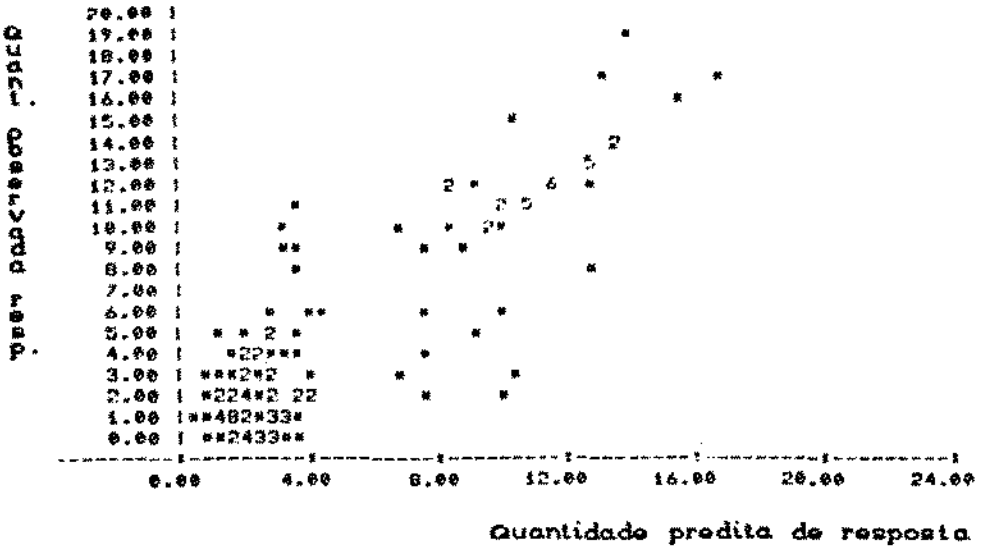


Figura 5.6: Gráfico da quantidade predita de respostas adversas contra a quantidade observada para o modelo $\text{logit}(P_{ij}) = \tau_i + \gamma x_{ij}$ ajustado aos dados de Tyl et al (1983).

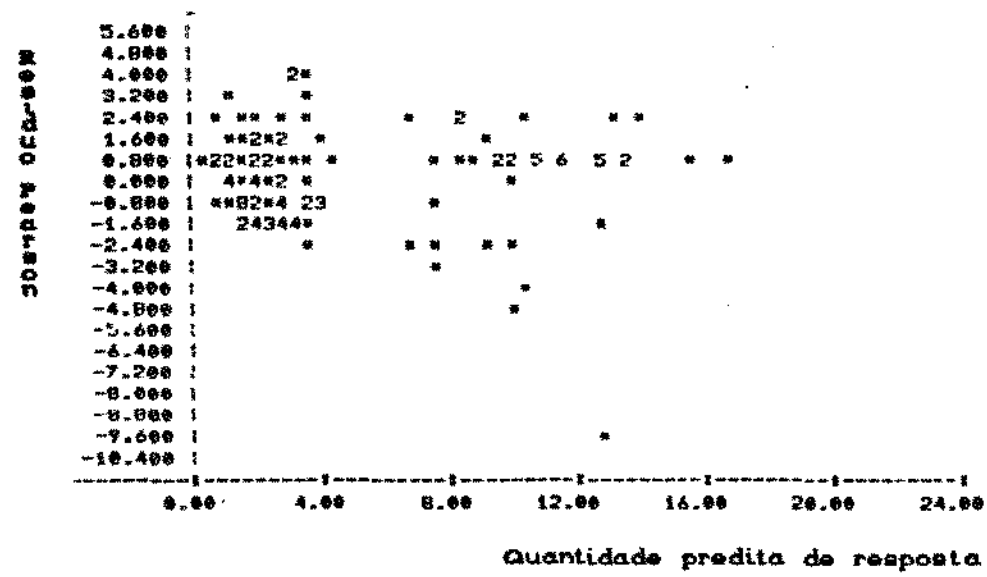


Figura 5.7: Gráfico da quantidade predita de respostas adversas contra o resíduo de Pearson para o modelo $\text{logit}(P_{ij}) = \tau_i + \gamma x_{ij}$ ajustado aos dados de Tyl et al (1983)

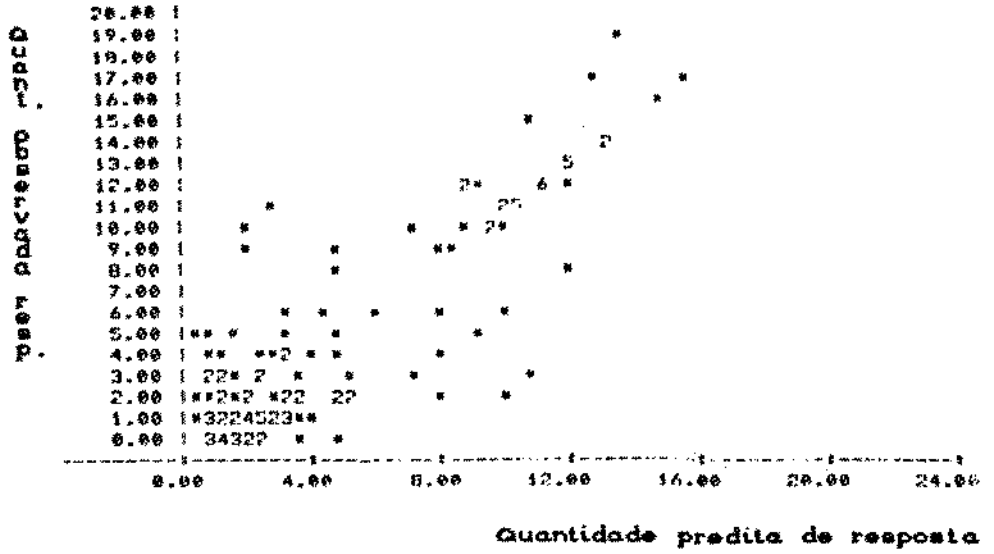


Figura 5.8: Gráfico da quantidade predita de respostas adversas contra a quantidade observada para o modelo $\text{logit}(P_{ij}) = \alpha + \beta d_i + \gamma m_{ij} + \delta d_i m_{ij}$ ajustado aos dados de Tyl et al (1983).

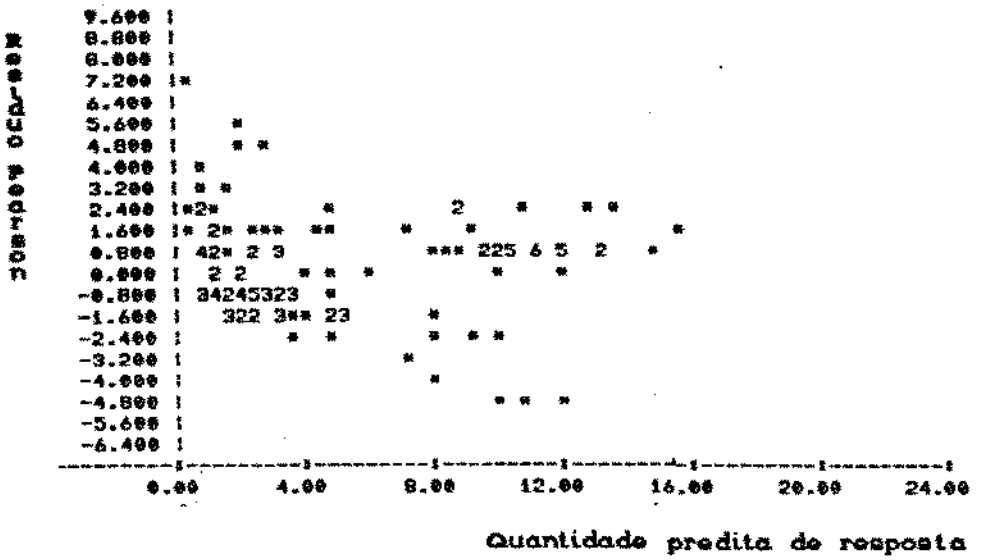


Figura 5.9: Gráfico da quantidade predita de respostas adversas contra o resíduo de Pearson para o modelo $\text{logit}(P_{ij}) = \alpha + \beta d_i + \gamma m_{ij} + \delta d_i m_{ij}$ ajustado aos dados de Tyl et al (1983).

Na aplicação dos testes de hipóteses relativos aos modelos, os valores das diferenças entre os desvios, os graus de liberdade e os níveis de significância atingidos P podem ser vistos na Tabela 5.4.

Tabela 5.4: Resultados dos testes dos modelos ajustados aos dados de Tyl et al (1983).

Modelo M_0	Modelo M_1	Diferença entre Desvios	Graus de Liberdade	Nível de Significância
RVR	I	218.22	46	P < 0.5%
IV	III	0.17	1	P > 50%
V	III	61.58	3	P < 0.5%
VI	IV	67.08	1	P < 0.5%
VII	V	3.94	1	2.5% < P < 5%
VIII	VII	2.66	1	5% < P < 50%
VIII	V	6.60	2	2.5% < P < 5%

Para os dados de Tyl et al (1983), o modelo RVR é rejeitado em favor do modelo que ajusta um total de 50 parâmetros, um para cada combinação de dose e tamanho de ninhada. Entretanto, este modelo é pouco parcimonioso e de difícil interpretação. No teste do modelo IV contra o modelo III, existem evidências de que regressões paralelas com respeito ao tamanho da ninhada são satisfatórias, não sendo necessária a contribuição do termo $\delta d_i s_{ij}$. Por outro lado, não há evidências da dependência linear de τ_i com a dose e o modelo $\alpha + \beta d_i + \gamma s_{ij} + \delta d_i s_{ij}$ não pode ser considerado adequado. Há evidências de que o modelo que ajusta apenas o efeito do grupo é inadequado, sendo necessária a covariável s_{ij}. Mais uma vez, temos evidências de regressões paralelas com respeito ao tamanho da ninhada. Da comparação entre os modelos VII e V, as evidências estão a favor do último. E, por fim, o teste do modelo VIII contra o VII não leva a abandonar o primeiro destes modelos. Como os modelos VIII e V são ainda encaixados, podemos testar um contra o outro, favorecendo assim o modelo V. Desta maneira, dentre os modelos considerados para este conjunto de dados, o modelo $\text{logit}(P_{ij}) = \tau_i + \gamma s_{ij}$ $i = 0, \ldots, m$ e $j = 1, \ldots, n_i$ parece ser

o mais razoável.

Para os dois conjuntos de dados, a estimação de todos os modelos foi refeita, retirando aqueles parâmetros cujas estimativas poderiam ser consideradas nulas. Entretanto, este procedimento não melhorou os resultados do teste de qualidade do ajustamento, nem os gráficos de diagnóstico. Por estes motivos, não apresentaremos os resultados.

A estimação dos parâmetros dos modelos logísticos e o cálculo dos desvios foram feitos supondo a presença de uma variação binomial pura nos dados. Entretanto, proporções de ninhadas observadas em experimentos teratológicos frequentemente apresentam variação extra-binomial. Por este motivo, na próxima seção serão reajustados os modelos I a VIII utilizando o procedimento apresentado por WILLIAMS (1982), onde o autor modifica o procedimento usual de estimação em modelos logísticos de modo a incorporar uma componente de variação extra-binomial.

5.3 VARIAÇÃO EXTRA-BINOMIAL

Um modelo logístico linear comum como o apresentado na seção 5.1, pode acomodar a variação extra-binomial ao introduzir variáveis aleatórias P_i absolutamente contínuas, independentes e não observáveis, distribuídas no intervalo (0, 1) com

$$E(P_i) = \pi_i \qquad e \qquad Var(P_i) = \phi\pi_i(1 - \pi_i),$$

onde ϕ é o parâmetro de variação extra-binomial. Assumindo ainda que a distribuição condicional $Y_i|P_i = p_i$ é uma binomial $bin(m_i, p_i)$, resultam os momentos incondicionais

$$E(Y_i) = E\,[E(Y_i|P_i = p_i)] = m_i\pi_i$$

$$Var(Y_i) = Var[E(Y_i|P_i = p_i)] + E[Var(Y_i|P_i = p_i)] = v_i w_i^{-1}$$

onde $v_i = m_i\pi_i(1 - \pi_i)$, $w_i^{-1} = [1 + \phi(m_i - 1)]$ e $i = 1, \ldots, n$. Observe que a variação extra-binomial não afeta a média da distribuição

de Y_i, mas apenas sua variância. Além disto, o procedimento de estimação que será descrito depende apenas da média e da variância de Y_i, sem uma especificação da distribuição correspondente.

A estimação dos parâmetros do modelo logístico com variação extra-binomial é feita de maneira diferente, dependendo do conhecimento existente sobre o parâmetro ϕ.

Se ϕ é conhecido, e desde que a distribuição de Y_i é incompletamente especificada, pelo conhecimento de apenas seus dois primeiros momentos, não é possível utilizar o procedimento de estimação por máxima verossimilhança. Por este motivo, a estimação será feita recorrendo à função de quase-verossimilhança discutida em WEDDERBURN (1974) e McCULLAGH & NELDER (1989). Desta maneira, a estimação é feita pelo uso iterativo das equações de mínimos quadrados ponderados, com pesos dados por $w_i^{-1} = [1 + \phi(m_i - 1)]$. No pacote estatístico GLIM, basta calcular o vetor W com componentes w_i como vetor de pesos para a diretiva $WEIGHT W, ajustando o modelo logístico da maneira usual.

Se ϕ é desconhecido, será necessário estimar não apenas o vetor de parâmetros do modelo, mas também o próprio parâmetro ϕ, o que será feito iterativamente pelo método dos momentos. Como o valor esperado da estatística χ^2 de Pearson envolve ϕ, igualando esta esperança aos graus de liberade correspondentes, deriva-se uma equação de estimação para ϕ.

Na prática, o procedimento apresentado por WILLIAMS (1982) para estimar o parâmetro ϕ e o modelo logístico com variação extra-binomial pode ser descrito pelo seguinte algoritmo.

Passo 1: Assuma que $\phi = 0$, estime o modelo logístico comum por máxima verossimilhança e calcule a estatística χ^2 de Pearson.

Passo 2: Compare χ^2 com a distribuição de referência χ^2_{n-p}. Se χ^2 é grande, conclua que $\phi > 0$ e calcule a estimativa

$\hat{\phi} = \dfrac{\chi^2 - (n-p)}{\sum_{i=1}^{n}(m_i - 1)(1 - v_i^* q_i)}$, onde q_i é o i-ésimo elemento diagonal de $X(X'WV^*X)^{-1}X'$, que é a matriz de variância do preditor η_i $i = 1, \ldots, n$.

Passo 3: Com a nova ponderação $w_i = [1 + \hat{\phi}(m_i - 1)]^{-1}$, estime $\hat{\beta}$ iterativamente usando $X'W\,V^*X\,\hat{\beta} = X'W\,V^*Y^*$ e recalcule χ^2.

Passo 4: Se χ^2 é próximo de n-p a estimativa de ϕ é satisfatória. Caso contrário, reestime ϕ como $\hat{\phi} = \dfrac{\chi^2 - \sum_{i=1}^{n} w_i(1 - w_i v_i^* q_i)}{\sum_{i=1}^{n} w_i(m_i - 1)(1 - w_i v_i^* q_i)}$ e volte ao passo 3.

Os passos necessários à este procedimento de estimação são facilmente programados no pacote estatístico GLIM. A correspondente sequência de comandos é apresentada no Apêndice E5.

O procedimento anterior foi aplicado para os dois conjuntos de dados já analisados. Na Tabela 5.5 encontram-se, para os modelos I a VIII e para os dados de Lüning et al (1966), as estimativas de ϕ, os valores dos novos desvios e a porcentagem de redução estes desvios em relação aos desvios dos modelos sem variação extra-binomial. A Tabela 5.6 mostra as mesmas quantidades, para os dados de Tyl et al (1983).

Em WILLIAMS (1982), o parâmetro ϕ é estimado para o modelo com maior número de parâmetros e esta estimativa é utilizada nos modelos restantes. Desta maneira, a estimativa de ϕ é dada pela primeira linha das Tabelas 5.5 e 5.6. Entretanto, neste trabalho optamos por estimar ϕ para cada modelo, visto que o procedimento de

estimação depende precisamente do modelo ajustado, através das covariáveis consideradas nele.

Tabela 5.5: Estimativas do parâmetro ϕ de variação extra-binomial, novos desvios e porcentagem de redução dos novos desvios em relação aos desvios do modelo logístico usual obtidos para os dados de Lüning et al (1966).

Modelo	$\hat{\phi}$	Novo Desvio	Redução(%)
I	0.0209	1995.01	10.4
II	0.0205	2009.96	12.2
III	0.0204	2011.10	10.2
IV	0.0214	2005.46	10.7
V	0.0199	2030.37	10.0
VI	0.0268	1975.66	13.0
VII	0.0214	2026.03	10.6
VIII	0.0266	1996.95	12.9

Tabela 5.6: Estimativas do parâmetro ϕ de variação extra-binomial, novos desvios e porcentagem de redução dos novos desvios em relação aos desvios do modelo logístico usual obtidos para os dados de Tyl et al (1983).

Modelo	$\hat{\phi}$	Novo Desvio	Redução(%)
I	0.1270	96.24	60.6
II	0.1995	113.58	69.8
III	0.1966	115.35	69.6
IV	0.1972	115.36	69.6
V	0.2331	118.98	73.0
VI	0.1958	116.31	69.5
VII	0.2392	117.92	73.5
VIII	0.2349	119.55	73.9

Em relação aos dados de Lüning et al (1966), em todos os modelos, as novas estimativas e os novos erros padrão tiveram um mudança muito pequena em relação aos modelos com variação binomial

pura. Além disso, as conclusões à respeito da qualidade do ajustamento dos modelos permanecem inalteradas.

Considerando a Tabela 5.6 e o modelo I, após várias iterações no procedimento de estimação de ϕ, o valor obtido mais próximo dos graus de liberdade (81) que obtivemos para a estatística χ^2 foi de 86.45. Entretanto, como não foi possível reduzi-lo, aceitamos 0.1270 como uma estimativa apenas razoável de ϕ. Numa tentativa de obter uma estimativa melhor, pode-se retirar do modelo I todos aqueles parâmetros cujas estimativas possam ser consideradas nulas, reestimando ϕ. Procedendo desta maneira, é obtido um valor χ^2 = 107.1558 com 107 graus de liberdade, uma estimativa para ϕ igual a 0.3224 e um devio D = 136.5688.

Para os dados de Tyl et al (1983), houve mudanças não apenas nas estimativas dos parâmetros e dos desvios-padrão para todos os modelos, mas também nas conclusões sobre a qualidade do ajustamento. Com estes novos desvios, todos os modelos têm um bom ajuste quando comparamos seus desvios com os respectivos valores críticos da distribuição qui-quadrado com nível de significância 5%. Entretanto, os gráficos diagnósticos continuam apresentando o padrão anteriormente descrito.

Vamos então refazer os testes à respeito dos vetores dos parâmetros destes modelos, exibindo os resultados na Tabela 5.7.

Tabela 5.7: Testes de hipóteses dos modelos com variação extra-binomial para os dados de Tyl et al (1983).

Modelo M_0	Modelo M_1	Diferença entre Desvios	Graus de Liberdade	Nível de Significância
IV	III	0.01	1	P > 50%
V	III	3.63	3	5% < P < 50%
VI	IV	0.96	1	5% < P < 50%
VII	V	1.06	1	5% < P < 50%
VIII	VII	1.63	1	5% < P < 50%

No teste do modelo IV contra o modelo III, existem evidências a favor de regressões paralelas no tamanho da ninhada. Por outro lado, há também evidências da dependência linear de τ_i com a dose d_i e, portanto, favoráveis ao modelo V. Também parece que o modelo que ajusta apenas o efeito do grupo é melhor do que as regressões paralelas no tamanho da ninhada. Comparando os modelos VII e V, as evidências estão a favor do modelo VII, sendo portanto desnecessário o termo $\delta d_i s_{ij}$. Do teste do modelo VIII contra o VII não podemos desprezar o modelo VIII.

As conclusões gerais deste exercício de comparação de modelos serão brevemene apresentadas no Capítulo seguinte, bem como as considerações finais relativas à este trabalho.

CAPÍTULO 6

CONSIDERAÇÕES FINAIS

Neste trabalho, procuramos discutir alguns aspectos inerentes ao problema da avaliação do risco de que um indivíduo exposto a um agente teratogénico tenha sua capacidade reprodutiva afetada. Desde que estudos epidemiológicos para avaliação destes riscos apresentam diversas dificuldades, experimentos teratológicos em animais de laboratório podem ser utilizados para gerar evidências sobre os agentes teratogénicos.

Com base nestes experimentos, pode-se avaliar o risco de que filhotes de fêmeas direta ou indiretamente expostas a diferentes doses de um agente apresentem resultados adversos tais como morte e malformações. No nosso caso, isto foi feito através da introdução de dois tipos de modelos estatísticos para dados de respostas binárias observadas em experimentos teratológicos: o modelo de dose-resposta proposto por RAI & VAN RYZIN (1985) e modelos logísticos lineares semelhnates aos de WILLIAMS (1987).

Dados gerados em experimentos teratológicos têm como principal característica o fato de serem coletados nos filhotes que formam as ninhadas de cada fêmea. Por este motivo, é de se esperar uma tendência de similaridade entre as respostas dos animais de uma mesma ninhada, denominada efeito de ninhada. Este efeito é responsável pela sobre-dispersão neste tipo de dados, que por afetar o ajuste dos modelos e, consequentemente, a estimativa do risco, deve ser considerado. Neste trabalho, tal efeito foi incorporado aos modelos pela introdução da informação sobre o tamanho da ninhada produzida por cada fêmea.

O modelo de dose-resposta proposto por RAI & VAN RYZIN (1985) foi construído sob a suposição básica de que o risco de um filhote apresentar a resposta adversa depende da toxicidade da exposição materna, por sua vez modelada segundo a teoria de um impacto. Este modelo é condicional ao tamanho da ninhada gerada por cada fêmea. Este tamanho foi também modelado, supondo as distribuições de Poisson e binomial negativa com mesma média, que depende exponencialmente da dose aplicada. Sugerimos a última distribuição com o intuito de tratar mais adequadamente a variabilidade destes tamanhos. Entretanto, para os dados aqui utilizados, as duas distribuições se mostraram equivalentes, devendo ser escolhida a de Poisson, por sua simplicidade. A expressão para o modelo de dose-resposta incondicional também foi obtida sob as duas distribuições.

No estudo da verossimilhança dos modelos, calculamos a matriz de informação de Fisher dos parâmetros de cada modelo; provamos que são satisfeitas as condições de regularidade para o modelo condicional, além de resultados necessários à demonstração de que tais condições são satisfeitas também para o modelo incondicional. Provamos ainda em detalhe a unicidade, existência, consistência e normalidade assintótica dos estimadores de máxima verossimilhança dos parâmetros do modelo condicional, completando assim a demonstração esquematizada em RAI & VAN RYZIN (1981).

Na estimação dos parâmetros foram utilizados três métodos numéricos de minimização de funções que produziram resultados

concordantes, exceto para o modelo binomial negativo, onde a discordância mais acentuada foi verificada na estimativa do parâmetro k. Os resultados diferentes podem ser atribuídos à equivalência entre as distribuições de Poisson e binomial negativa, para os conjuntos de dados utilizados. Em relação aos algoritmos, optamos pela utilização do método de Newton com canalizações, que encontra-se programado na subrotina E04LBF da NAG (1982). Desde que é necessário calcular as derivadas parciais de segunda ordem da função log-verossimilhança para a obtenção da matriz de covariâncias, é aconselhável utilizá-las na minimização da função. Dos métodos utilizados, o de Newton é aquele que, além de mais rápido, utiliza esta informação. O leitor que não tiver acesso à biblioteca NAG poderá usar a subrotina OPTIM (e, portanto, o método dos gradientes conjugados), que apesar de mais lento é também adequado.

Em relação ao modelo RVR, algumas questões devem ser ainda estudadas, tais como a construção de estimadores para extrapolar o risco para doses baixas, a estimação de doses virtualmente seguras e testes de hipóteses relativos aos parâmetros. Possibilidades de melhora deste modelo também devem ser avaliadas, incluindo a consideração de outras covariáveis, além do tamanho da ninhada para cada fêmea.

Também avaliamos o risco de resposta adversa para filhotes gerados em experimentos teratológicos através do ajuste de modelos logísticos lineares. Os modelos estudados incluiram aqueles propostos por WILLIAMS (1987) e por fazerem parte da classe dos modelos lineares generalizados, foram ajustados através do pacote estatístico GLIM. Nestes modelos, o tamanho da ninhada atuou como covariável e foram incorporados os efeitos da dose e do grupo. Entretanto, para os dois conjuntos de dados utilizados, os ajustes obtidos não parecem bons. O ajuste de todos os modelos foi ainda refeito, eliminando aqueles parâmetros cujas estimativas foram consideradas nulas, sem que houvesse melhora notável na qualidade do ajustamento.

Supondo a existência de variação extra-binomial nos dados considerados, utilizamos finalmente o procedimento de estimação apresentado por WILLIAMS (1982), de modo a incorporar uma componente de

variação extra-binomial nos modelos logísticos. Estes modelos foram reajustados e reavaliada a qualidade do ajustamento de cada um. Para os dados de Lüning et al (1966) as conclusões relativas à qualidade do ajustamento permaneceram inalteradas. Para os dados de Tyl et al (1983), os modelos com variação extra-binomial produziram um ajuste melhor.

Em relação aos modelos logísticos, uma análise mais aprofundada, baseada em instrumentos gráficos de diagnóstico pode permitir uma melhor avaliação da qualidade do ajustamento, exibindo modelos parcimoniosos (em princípio, diversos) com ajuste razoável. Além disto, talvez a formulação de modelos um pouco mais complexos possa melhorar a qualidade do ajustamento, sendo razoável a incorporação de regressores suplementares biologicamente aceitáveis e a suposição de que o efeito do grupo dependa da dose de forma não linear.

APÊNDICE A

LEMAS

Neste Apêndice são apresentados alguns resultados necessários para demonstrar o Teorema 4 na seção 3.4.3, referente a existência, unicidade e consistência dos estimadores de máxima verossimilhança dos parâmetros do modelo condicional de dose-resposta proposto em RAI & VAN RYZIN (1985). Também é demonstrado que são satisfeitas as condições para a utilização do Teorema Central do Limite na versão de Liapunov, necessário na demonstração do Teorema 5, da seção 3.4.3, relativo à normalidade assintótica dos mesmos estimadores. Por fim, é demonstrado que a distribuição de Poisson é obtida como caso limite da distribuição binomial negativa.

Definição A1: A sequência de variáveis aleatórias reais (X_n) $n = 1, 2, \ldots$ definida sobre um espaço de probabilidade (Ω, A, P) converge quase certamente à constante $c \in \mathbb{R}$ se $P\left(\left\{ \lim_{n\to\infty} X_n = c \right\} \right) = 1$ ou equivalentemente, se $\lim_{N\to\infty} \left(P\left\{ \sup_{n\geq N} |X_n - c| > \varepsilon \right\} \right) = 0$ para todo $\varepsilon > 0$. Uma sequência de variáveis aleatórias (X_n) converge à variável aleatória X no sentido acima se a sequência $(X_n - X)$ $n = 1, 2, \ldots$ converge quase certamente a zero. Utilizaremos a notação $X_n \xrightarrow{qc} X$ para denotar este fato.

As propriedades básicas de convergência quase certa podem ser encontradas por exemplo, no Capítulo 2 de RAO (1973). No que segue, demonstraremos alguns resultados que não podem ser encontrados ali.

Lema A2 (Lema de Fréchet-Slutsky): Se $X_n^i \xrightarrow{qc} X^i$ para $i = 1, 2, \ldots, k$ e se $g: \mathbb{R}^k \to \mathbb{R}$ é uma função contínua, então

$$g(X_n^1, \ldots, X_n^k) \xrightarrow{qc} g(X^1, \ldots, X^k).$$

Prova:

Para cada $i = 1, 2, \ldots, k$ existe um conjunto $N_i \subseteq \Omega$ com $P(N_i) = 0$ onde $N_i = \left\{ \omega: \lim_{n\to\infty} X_n^i(\omega) \neq X^i(\omega) \right\}$. Então, se $M = \bigcap_{i=1}^{k} N_i^C$ temos que $P(M^C) = P\left(\bigcup_{i=1}^{k} N_i\right) \leq \sum_{i=1}^{k} P(N_i) = 0$, $P(M) = 1$ e se $\omega \in M$, $X_n^1(\omega), \ldots, X_n^k(\omega) \xrightarrow[n\to\infty]{} X^1(\omega), \ldots, X^k(\omega)$. Como g é uma função contínua em $\mathbb{R}^k$, de acordo com o Teorema 1, p. 57 em BUCK (1965), temos que

$$\lim_{n\to\infty} g(X_n^1(\omega), \ldots, X_n^k(\omega)) = g(X^1(\omega), \ldots, X^k(\omega)),$$

mas como $M \subseteq \left\{\omega: \lim_{n\to\infty} g(X_n^1(\omega), \ldots, X_n^k(\omega)) = g(X^1(\omega), \ldots, X^k(\omega))\right\}$ resulta que

$$P(M) \leq P\left[\left\{\omega: \lim_{n\to\infty} g(X_n^1(\omega), \ldots, X_n^k(\omega)) = g(X^1(\omega), \ldots, X^k(\omega))\right\}\right] \leq 1.$$

E assim, $g(X_n^1(\omega), \ldots, X_n^k(\omega)) \xrightarrow{qc} g(X^1(\omega), \ldots, X^k(\omega))$.■

Lema A3: Seja (α_n) $n = 1, 2, \ldots$ uma sequência numérica convergente com $\lim_{n\to\infty} \alpha_n = \alpha$. Então, $\alpha_n \xrightarrow{qc} \alpha$.

Prova:

Como α_n é uma variável aleatória certa, o conjunto Ω e os conjuntos $\left\{\omega: \lim_{n\to\infty} \alpha_n(\omega) = \alpha(\omega)\right\}$ coincidem. Assim, $\alpha_n \xrightarrow{qc} \alpha$.■

Lema A4: Uma combinação linear finita de sequências quase certamente convergentes é quase certamente convergente.

Prova:

Seja $X_n^i \xrightarrow{qc} X^i$ e $\alpha_n^i \xrightarrow[n\to\infty]{} \alpha^i$. Pelo Lema A3, temos que $\alpha_n^i \xrightarrow{qc} \alpha^i$, pelo Lema A2, $\alpha_n^i + X_n^i \xrightarrow{qc} \alpha^i + X^i$ e $\alpha_n^i X_n^i \xrightarrow{qc} \alpha^i X^i$ para todo $i = 1, \ldots, k$. Portanto, a combinação linear finita $\sum_{i=1}^{k} \alpha_n^i X_n^i$ converge quase certamente para $\sum_{i=1}^{k} \alpha^i X^i$.■

Lema A5: Se $(f_n(x))$ é uma sequência de funções contínuas sobre o compacto $E \subseteq \mathbb{R}^k$ com $f_1(x) \geq f_2(x) \geq \ldots$, se $f_n(x) \xrightarrow[n\to\infty]{} f(x)$ pontualmente e se $f(x)$ é contínua, então $(f_n(x))$ converge uniformemente a $f(x)$.

Prova:

A sequência $g_n(x) = f_n(x) - f(x)$ é uma sequência de funções contínuas com $g_n(x) \geq 0$; $g_1(x) \geq g_2(x) \geq \ldots$ e $\lim_{n\to\infty} g_n(x) = 0$.

Mostraremos que $g_n(x) \xrightarrow[n\to\infty]{} 0$ uniformemente.

Seja $x \in E$. De $\lim_{n\to\infty} g_n(x) = 0$, tem-se que para todo $\varepsilon > 0$, existe um $N(x,\varepsilon)$ tal que $g_n(x) < \varepsilon/2$ para todo $n \geq N(x,\varepsilon)$. Em particular, $g_{N(x,\varepsilon)}(x) < \varepsilon/2$. Então, $g_{N(x,\varepsilon)}(y) < \varepsilon$ para todo $y \in C_x$, onde C_x é uma bola aberta de centro x contida em E. Se isto não fosse verdadeiro, existiria $y_0 \in C_x$ tal que $g_{N(x,\varepsilon)}(y_0) > \varepsilon$. Então, $g_{N(x,\varepsilon)}(y_0) - g_{N(x,\varepsilon)}(x) > \varepsilon - \varepsilon/2 = \varepsilon/2 > 0$. Assim, $g_{N(x,\varepsilon)}(y_0) - g_{N(x,\varepsilon)}(x) = |g_{N(x,\varepsilon)}(y_0) - g_{N(x,\varepsilon)}(x)|$. Portanto, dado $\varepsilon > 0$, dado $x \in E$ e dada uma bola aberta arbitrária C_x, existiria $y_0 \in C_x$ tal que $|g_{N(x,\varepsilon)}(y_0) - g_{N(x,\varepsilon)}(x)| > \varepsilon/2$, de maneira que $g_{N(x,\varepsilon)}(.)$ não seria uma função contínua.

Pelo Lema de Heine-Borel (veja BUCK (1965), p. 39) o compacto E pode ser recoberto por uma coleção infinita de bolas abertas C_x com centros $x \in E$, existindo ainda um sub-recobrimento finito de E formado por bolas abertas centradas em $x_1, \ldots, x_r$ e denotado por $M = \{C_{x_1}, \ldots, C_{x_r}\}$. Para este subrecobrimento, existe um índice $N(x_l,\varepsilon)$ com $g_{N(x_l,\varepsilon)}(x_l) < \varepsilon/2$ para todo $l = 1, \ldots, r$. Seja $N(\varepsilon) = \sup\{N(x_1,\varepsilon), \ldots N(x_r,\varepsilon)\}$. Então, para todo $y \in E$ existe um índice l com $y \in C_{x_l}$ e $g_{N(\varepsilon)}(y) \leq g_{N(x_l,\varepsilon)}(y)$, pois $g_1(y) \geq g_2(y) \geq \ldots$ e $N(\varepsilon) \geq N(x_l,\varepsilon)$ para todo $l = 1, \ldots, r$. E como $0 \leq g_{N(\varepsilon)}(y) < \varepsilon$, isto implica que $0 \leq g_n(y) < \varepsilon$ para todo $n \geq N(\varepsilon)$ e para todo $y \in E$. Logo, $(g_n(x))$ converge a zero uniformemente para todo $x \in E$.■

Lema A6: Se $f_n(x)$: $E \to \mathbb{R}$ é uma sequência uniformemente convergente de funções contínuas sobre o compacto $E \subseteq \mathbb{R}^k$ e se $f(x) = \lim_{n\to\infty} f_n(x)$, então

$$\sup_{x\in E} f_n(x) \xrightarrow[n\to\infty]{} \sup_{x\in E} f(x).$$

Prova:

Como f(x) é limite uniforme de funções contínuas, temos que f(x) é contínua. Além disso, de acordo com o Lema de BUCK (1965) p. 74, $\sup_{x\in E} f_n(x)$ e $\sup_{x\in E} f(x)$ existem, pois $f_n(x)$ e f(x) são contínuas e E é compacto. Sejam $a_n = f_n(x_n) = \sup_{x\in E} f_n(x)$ e $a_o = f(x_o) = \sup_{x\in E} f(x)$, onde x_n e x_o são os pontos em E onde $f_n(x)$ e f(x) respectivamente atingem seus valores máximos. Como $f_n(x) \xrightarrow[n\to\infty]{} f(x)$ uniformemente, temos que para todo $\varepsilon > 0$ existe $N(\varepsilon) \in \mathbb{N}$ independente de $x \in E$ tal que

$$|f_n(x) - f(x)| \leq \varepsilon/3 \text{ para todo } n \geq N(\varepsilon) \text{ e para todo } x \in E. \qquad \text{(A6.1)}$$

Em particular, $$f_n(x_n) - f(x_n) \leq |f_n(x_n) - f(x_n)| \leq \varepsilon/3 \qquad \text{(A6.2)}$$

Além disso, $f_n(x) \leq f_n(x_n)$ para todo $x \in E$. Em particular, se $x = x_o$,

$$f_n(x_o) \leq f_n(x_n). \qquad \text{(A6.3)}$$

De A6.2 temos que $$f_n(x_n) \leq \varepsilon/3 + f(x_n). \qquad \text{(A6.4)}$$

Em A6.1, se $x = x_o$, temos que $|f_n(x_o) - f(x_o)| \leq \varepsilon/3$ e, portanto, $f_n(x_o) - f(x_o) \leq |f_n(x_o) - f(x_o)| \leq \varepsilon/3$, o que leva a

$$f_n(x_o) \leq \varepsilon/3 + f(x_o). \qquad \text{(A6.5)}$$

De A6.3, A6.4 e A6.5, temos que

$$f_n(x_o) \leq f_n(x_n) \leq \varepsilon/3 + f(x_n) \leq \varepsilon/3 + f(x_o). \qquad \text{(A6.6)}$$

Assim, $|a_n - a_o| = |f_n(x_n) - f(x_o)| \leq f_n(x_n) - f_n(x_o) + |f_n(x_o) - f(x_o)|$.

De A6.6, $f_n(x_n) - f_n(x_o) \leq \varepsilon/3 + f(x_o) - f_n(x_o)$ e $f_n(x_o) \leq \varepsilon/3 + f(x_o)$.

Mas com $\varepsilon > 0$, temos que $$f_n(x_o) - f(x_o) \leq 0. \qquad \text{(A6.7)}$$

Desta maneira, $|a_n - a_o| \leq \varepsilon/3 + 2\,[f(x_o) - f_n(x_o)]$. De A6.1 e A6.7 temos que $|f_n(x) - f(x)| \leq \varepsilon/3$ e $f_n(x) - f(x) \leq 0$ e, portanto, $f(x_o) - f_n(x_o) \leq \varepsilon/3$. Assim, $|a_n - a_o| = \left| \sup_{x\in E} f_n(x) - \sup_{x\in E} f(x) \right| \leq \varepsilon/3 + 2\varepsilon/3 =$

ε e, portanto, $\sup_{x\in E} f_n(x) \xrightarrow[n\to\infty]{} \sup_{x\in E} f(x)$.■

O lema a seguir é uma extensão multivariada do Teorema de Rolle. Ele foi provado e discutido em RAI & VAN RYZIN (1980b) e será incluído aqui, pois foi publicado em um relatório técnico de difícil acesso.

Lema A7: Seja $g: S \to \mathbb{R}$ uma função contínua de um subconjunto convexo $S \subseteq \mathbb{R}^k$, com derivadas parciais de primeira e segunda ordens contínuas em S. Seja D(.) o vetor coluna das derivadas parciais de primeira ordem de g(x) avaliadas em x = . e H(.) a matriz hessiana de ordem k de g(x), avaliada em x = .. Se há dois pontos x_1 e x_2 em S tais que $D(x_1) = D(x_2)$ e se H(.) é positiva ou negativa semi definida em S, então existe um ponto x_3 no segmento de linha que une os pontos x_1 e x_2 tal que o determinante de $H(x_3)$ é nulo.

Prova:

Como x_1 e x_2 são dois pontos distintos, podemos escrever o segmento de linha em S que une estes dois pontos como

$$x(\lambda) = \lambda x_1 + (1 - \lambda)x_2 \quad \text{com } 0 \le \lambda \le 1.$$

Seja $F(\lambda)$ o vetor coluna das derivadas parciais de primeira ordem de $x(\lambda)$ com respeito a λ. Então, $F(\lambda) = x_1 - x_2 \neq 0$. Faça $\phi(\lambda) = g(x(\lambda))$.

Então, $$\frac{d\phi(\lambda)}{d\lambda} = [D(x(\lambda))]'(x_1 - x_2). \qquad \text{(A7.1)}$$

Mas como $D(x_1) = D(x_2)$, de A7.1 temos que $\left.\frac{d\phi(\lambda)}{d\lambda}\right|_{\lambda=0} = \left.\frac{d\phi(\lambda)}{d\lambda}\right|_{\lambda=1}$.

Do Teorema de Rolle univariado, infere-se a existência de pelo menos um ponto λ_0, com $0 < \lambda_0 < 1$ tal que

$$0 = \left.\frac{d^2\phi(\lambda)}{d\lambda^2}\right|_{\lambda=\lambda_0} = (x_1 - x_2)'H(x(\lambda_0))(x_1 - x_2).$$

Como $x_1 - x_2 \neq 0$ e a matriz H(.) é positiva ou negativa semi definida em S, isto implica que o determinante de $H(x_3)$ é zero, para $x_3 = x(\lambda_0)$.■

Lema A8: As condições para a aplicação do Teorema Central do Limite na versão de Liapunov exibida em RAO (1973) p. 127, são satisfeitas para a sequência de variáveis aleatórias independentes X_n, onde

$$X_n = n^{-1/2}\mathbf{b}'\mathbf{D}l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0} = n^{-1/2}\sum_{t=1}^{4} b_t \frac{\partial l(\underset{\sim}{\psi})}{\partial \psi_t}\Big|_{\underset{\sim}{\psi}_0} = n^{-1/2}\sum_{t=1}^{4} b_t D_t l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0},$$

$\mathbf{b} = (b_1, b_2, b_3, b_4)'$ é um vetor arbitrário e $Z_t = Z_t(Y) = D_t l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0}$, onde para o vetor aleatório Y de dimensão nx1 com componentes y_{ij}, $D_t l(\underset{\sim}{\psi})\Big|_{\underset{\sim}{\psi}_0}$ $= \sum_{i=0}^{m}\sum_{j=1}^{n_i} D_t P_{ij} \dfrac{Y_{ij} - s_{ij}P_{ij}}{P_{ij}Q_{ij}}$ é a derivada da função log-verossimilhança com respeito à componente ψ_t de $\underset{\sim}{\psi}$, para t = 1, ..., 4. Desta forma, $X_n = n^{-1/2}\sum_{t=1}^{4} b_t Z_t$.

Prova:

Precisamos calcular $\mu_n = E(X_n)$, $\sigma_n^2 = Var(X_n)$, $\beta_n = E|X_n - \mu_n|^3$, $B_n = \left(\sum_{i=1}^{n}\beta_i\right)^{1/3}$ e $C_n = \left(\sum_{i=1}^{n}\sigma_i^2\right)^{1/2}$. Assim,

i) $E(X_n) = n^{-1/2}\sum_{t=1}^{4} b_t[E(Z_t)] = 0$, pois $E(Z_t) = 0$ para t = 1, ..., 4, e

ii) $Var(X_n) = Var\left[n^{-1/2}\mathbf{b}'\mathbf{D}l(\underset{\sim}{\psi})\Big|_{\psi_0}\right] = n^{-1}\mathbf{b}'Var[Z(Y)]\mathbf{b}$. Mas $Var[Z(Y)] = Var\left[\mathbf{D}l(\underset{\sim}{\psi})\Big|_{\psi_0}\right] = E\left[\left(\mathbf{D}l(\underset{\sim}{\psi})\right)\left(\mathbf{D}l(\underset{\sim}{\psi})\right)'\Big|_{\psi_0}\right] = E(\mathbf{W})$, onde $\mathbf{W}$ é a matriz de dimensão 4x4 com elementos $w_{rs} = \dfrac{\partial l(\underset{\sim}{\psi})}{\partial \psi_r}\dfrac{\partial l(\underset{\sim}{\psi})}{\partial \psi_s}$ r, s = 1, ..., 4. Logo, $Var[Z(Y)] = I(\alpha,\beta,\theta_1,\theta_2) = \Sigma^{-1}$, com Σ^{-1} positiva definida. Desta maneira, $\sigma_n^2 = Var(X_n) = n^{-1}\mathbf{b}'\Sigma^{-1}\mathbf{b}$ e $0 < Var(X_n) < \infty$.

Assim, o valor da expressão $C_n = \left(\sum_{i=1}^{n} \sigma_i^2 \right)^{1/2}$ neste caso é dado por $\left(b'\Sigma^{-1}b \sum_{i=1}^{n} i^{-1} \right)^{1/2}$.

Como o cômputo de $E|X_n|^3$ é complicado, calcularemos $E(X_n^4)$ e utilizaremos a relação $E|X_n|^3 \leq [E(X_n^4)]^{3/4}$, que pode ser derivada da desigualdade $(E|X|^t)^{1/t} \leq (E|X|^r)^{1/r}$ $0 < t \leq r$. Veja RAO (1973) p.149. Assim,

iii) $E(X_n^4) = E\left(n^{-1/2} \sum_{i=1}^{4} b_i Z_i\right)^4 = n^{-2} \sum_r \sum_s \sum_t \sum_u b_r b_s b_t b_u E(Z_r Z_s Z_t Z_u)$.

Para simplificar a notação, escreveremos $s_{ij} = s$, $P_{ij} = P$, $Q_{ij} = Q$, $Y_{ij} = Y$, $D_r P_{ij} = D_r P = D_r$ e $\sum_{i=0}^{m} \sum_{j=1}^{n_i} = \sum$ e assim sucessivamente até $s_{i'j'} = s'''$, $P_{i'j'} = P'''$, $Q_{i'j'} = Q'''$, $Y_{i'j'} = Y'''$, $D_r P_{i'j'} = D_r P''' = D_r'''$ e $\sum_{i'''=0}^{m} \sum_{j'''=1}^{n_{i'''}} = \sum'''$, denotando também a somatória $\sum_r \sum_s \sum_t \sum_u$ mediante $\sum_{r,s,t,u}$. Desta forma, temos

$$E(X_n^4) = n^{-2} \sum_{r,s,t,u} b_r b_s b_t b_u \sum \sum' \sum'' \sum''' D_r D_s' D_t'' D_u''' \times$$

$$\times E\left(\frac{(Y-sP)(Y'-s'P')(Y''-s''P'')(Y'''-s'''P''')}{PQP'Q'P''Q''P'''Q'''} \right).$$

Mas a esperança acima é sempre nula, exceto no caso em que todos os índices são iguais ou iguais dois a dois. No primeiro caso, $E\left(\frac{Y-sP}{PQ}\right)^4 = \frac{s}{(PQ)^3} + \frac{3s(s-2)}{(PQ)^2}$ e no segundo caso, $E\left(\frac{(Y-sP)^2(Y'-s'P')^2}{(PQ)^2(P'Q')^2} \right) = \frac{ss'}{PQP'Q'}$.

Temos então que

$$E(X_n^4) = n^{-2}\left\{\sum\left(\frac{s}{(PQ)^3}+\frac{3s(s-2)}{(PQ)^2}\right)\sum_{r,s,t,u} b_r b_s b_t b_u D_r D_s D_t D_u + \right.$$

$$+ \sum\sum{}' \frac{ss'}{PQP'Q'}\left(\sum_{r,s,t,u} b_r b_s b_t b_u D_r D_s D'_t D'_u + \sum_{r,s,t,u} b_r b_s b_t b_u D_r D'_s D_t D'_u + \right.$$

$$\left.\left. + \sum_{r,s,t,u} b_r b_s b_t b_u D_r D'_s D'_t D_u \right)\right\}.$$

Isto é,

$$E(X_n^4) = n^{-2}\left\{\sum\left(\frac{s}{(PQ)^3}+\frac{3s(s-2)}{(PQ)^2}\right)\left(\sum_{t=1}^{4} b_t D_t\right)^4 + \right.$$

$$+ \sum\sum{}' \frac{ss'}{PQP'Q'}\left[(b_1D_1+b_2D_2+b_3D'_3+b_4D'_4)^4 + (b_1D'_1+b_2D'_2+b_3D'_3+b_4D_4)^4 + \right.$$

$$\left.\left. (b_1D_1+b_2D'_2+b_3D'_3+b_4D_4)^4\right]\right\}$$

Nosso interesse é encontrar um limite superior para $E(X_n^4)$. Sendo assim, vale lembrar que para todo $j = 1, \ldots, n_i$ e para todo $i = 0, \ldots, m$, temos:

a) $D_1 = \exp[-(\alpha+\beta d_i)]\exp[-s_{ij}(\theta_1+\theta_2 d_i)] < 1, \quad D'_1 < 1,$

b) $D_2 = d_i D_1 < d_i \leq d_m \quad D'_2 < d_m$ onde d_m é a dose máxima

c) $D_3 = -s_{ij}P_{ij}, \quad |D_3| < s_{ij} \quad e \quad |D'_3| < s_{ij}.$

d) $D_4 = -d_i s_{ij} P_{ij}, \quad |D_4| < d_m s_{ij} \quad e \quad |D'_4| < d_m s_{ij}.$

Também são válidas as relações

e) $\left(\sum_{t=1}^{4} b_t D_t\right)^4 = \left|\sum_{t=1}^{4} b_t D_t\right|^4 \leq t_{ij}^4 \quad$ com $t_{ij}^4 = (|b_1| + |b_2|d_m + |b_3|s_{ij} + |b_4|d_m s_{ij})^4 = As_{ij}^4 + Bs_{ij}^3 + Cs_{ij}^2 + Ds_{ij} + E$, com os coeficientes A, B, C, D e E dependendo apenas da dose máxima d_m e de $b = (b_1, b_2, b_3, b_4)'$.

f) $(b_1D_1+b_2D_2+b_3D'_3+b_4D'_4)^4 \leq u_{ij}^4 \quad$ com $u_{ij}^4 = (|b_1| + |b_2|d_m +$

$+ |b_3|s_{i'j'} + |b_4|d_m s_{i'j'})^4 = As^4_{i'j'} + Bs^3_{i'j'} + Cs^2_{i'j'} + Ds_{i'j'} + E,$

com os mesmos coeficientes A, B, C, D e E de (e).

g) $(b_1D'_1+b_2D_2+b_3D'_3+b_4D_4)^4 = (b_1D_1+b_2D'_2+b_3D'_3+b_4D_4)^4 \leq v^4_{iji'j'}$ onde

$v^4_{iji'j'} = (|b_1| + |b_2|d_m + |b_3|s_{i'j'} + |b_4|d_m s_{ij})^4 = Fs^4_{ij} + Gs^3_{ij} + Hs^2_{ij} + Is_{ij} + Js^4_{i'j'} + Ks^3_{i'j'} + Ls^2_{i'j'} + Ms_{i'j'} + Ns_{i'j'}s^3_{ij} + Os^3_{i'j'}s_{ij} + Rs^2_{i'j'}s^2_{ij} + Ts_{i'j'}s_{ij} + Us^2_{i'j'}s_{ij} + Vs_{i'j'}s^2_{ij} + W.$ Os coeficientes F, G, ..., W também só dependem de d_m e do vetor b.

h) $0 < P_{ij}(d_m) < P_{ij}(d_i) < P_{ij}(0) < 1$ e

$0 < Q_{ij}(0) < Q_{ij}(d_i) < Q_{ij}(d_m) < 1.$

Portanto, $P_{ij}(d_i)Q_{ij}(d_i) > P_{ij}(d_m)Q_{ij}(0) = a(\underset{\sim}{\psi})$ onde $a(\underset{\sim}{\psi}) = a$ é uma constante. Mas se $P_{ij}(d_i)Q_{ij}(d_i) > a$, então $[P_{ij}(d_i)Q_{ij}(d_i)]^{-k} < a^{-k}$ onde k é um inteiro positivo.

Usando as relações (e) a (g), podemos escrever

$$E(X_n^4) < n^{-2}\left\{\sum\left(\frac{s}{a^3} + \frac{3s|s-2|}{a^2}\right)t^4_{ij} + \sum\sum{}' \frac{ss'}{aa'}(u^4_{i'j'} + 2v^4_{iji'j'})\right\}$$

ou seja, $E(X_n^4) < n^{-2}c$ e $[E(X_n^4)]^{3/4} < n^{-3/2}c^{3/4}$, onde c é uma constante positiva que depende apenas do vetor de parâmetros $\underset{\sim}{\psi}$, pois d_m, s_{ij}, $s_{i'j'}$ e b são conhecidos.

Seja $B_n = \left(\sum_{i=1}^n \beta_i\right)^{1/3} = \left(\sum_{i=1}^n E|X_i|^3\right)^{1/3}$. Mas desde que $E|X_i|^3 \leq [E(X_i^4)]^{3/4}$, temos que $B_n \leq \left\{\sum_{i=1}^n [E(X_i^4)]^{3/4}\right\}^{1/3}$ e, portanto,

$$B_n \leq c^{1/4}\left(\sum_{i=1}^n i^{-3/2}\right)^{1/3}.$$

Como as sequências B_n e C_n são não negativas, $\lim_{n\to\infty} \frac{B_n}{C_n} \geq 0$

se o limite existe. Por outro lado, temos que $\lim_{n\to\infty} \frac{B_n}{C_n} \leq 0$, visto que

$$\lim_{n\to\infty} \frac{B_n}{C_n} \leq \lim_{n\to\infty} \frac{c^{1/4}\left(\sum_{i=1}^{n} i^{-3/2}\right)^{1/3}}{(b'\Sigma^{-1}b)^{1/2}\left(\sum_{i=1}^{n} i^{-1}\right)^{1/2}} = 0,$$ pela divergência da série

harmónica. Portanto, $\lim_{n\to\infty} \frac{B_n}{C_n} = 0$.

Assim, as condições para a utilização do Teorema Central do Limite na versão de Liapunov são satifeitas.■

Lema A9: Se a variável aleatória X segue a distribuição binomial negativa com parâmetros m e k, e $k \to \infty$, então $\lim_{k\to\infty} P(X = x) = \frac{m^x e^{-m}}{x!}$.

Prova:

Se $X \sim BN(m, k)$, então

$$P(X = x) = \binom{x+k-1}{k-1}\left(\frac{m}{m+k}\right)^x \left(1 + \frac{m}{k}\right)^{-k}$$

$$= \frac{(x+k-1)(x+k-2)\ldots(k-2)\,k^x}{k^x} \frac{m^x}{x!}\left(\frac{1}{m+k}\right)^x \left(1 + \frac{m}{k}\right)^{-k}$$

$$= \left(1 + \frac{x-1}{k}\right)\left(1 + \frac{x-2}{k}\right)\ldots\left(1 - \frac{2}{k}\right)\left(\frac{1}{1+(m/k)}\right)^x \frac{m^x}{x!}\left(1 + \frac{m}{k}\right)^{-k}.$$

Desta maneira,

$$\lim_{k\to\infty} P(X = x) = \frac{m^x}{x!} \lim_{k\to\infty} \left(1 + \frac{x-1}{k}\right)\left(1 + \frac{x-2}{k}\right)\ldots\left(1 - \frac{2}{k}\right) \times$$

$$\times \left(\frac{1}{1+(m/k)}\right)^x \left(1 + \frac{m}{k}\right)^{-k}$$

$$= \frac{m^x}{x!} \lim_{k\to\infty} \left(1 + \frac{m}{k}\right)^{-k} = \frac{m^x e^{-m}}{x!}.■$$

Observação: Seja uma sequência de variáveis aleatórias (X_n), $n \geq 1$ com $X_n \sim BN(m, k_n)$ onde (k_n), $n \geq 1$ é uma sequência tal que $(k_n) \xrightarrow[n \to \infty]{} \infty$. Seja X uma variável aleatória distribuída segundo uma distribuição de Poisson com parâmetro m. Então, o Teorema 1 p. 242 em ROHATGI (1976) mostra que $X_n \xrightarrow{D} X$.

APÊNDICE B

DADOS ORIGINAIS

Neste Apéndice encontram-se os conjuntos de dados gerados nos dois experimentos teratológicos analisados nesta tese. O primeiro deles, exibido na Tabela B1, é proveniente do experimento de LÜNING et al (1966) e foi analisado por RAI & VAN RYZIN (1985) e WILLIAMS (1987). O segundo conjunto de dados é exibido na Tabela B2, foi gerado em um experimento de TYL et al (1983) e estudado por CHEN & KODELL (1989), FAUSTMAN et al (1989) e RYAN (1992).

Como foi exposto na seção 4.3, os conjuntos de dados são bastante diferentes. No primeiro conjunto, apenas as respostas dos filhotes de fêmeas que geraram ninhadas de tamanhos 5 a 10 foram registradas e o total de fêmeas no experimento é muito grande. O segundo conjunto de dados apresenta tamanhos variados de ninhadas e o total de fêmeas no experimento é mais próximo ao usual em experimentos teratológicos.

Tabela B1: Quantidade de fêmeas expostas à irradiação por raios X e classificadas de acordo com o número de implantes realizados e número de filhotes mortos. Dados de Lüning et al (1966).

Dose ($Rx10^{-3}$)	Número de implantes	Nº de filhotes mortos								Total de fêmeas
		0	1	2	3	4	5	6	7	
0.0	5	30	27	9	5					71
	6	86	51	14	4	1				156
	7	111	73	31	8	1				224
	8	79	44	23	3	0	1			150
	9	32	29	8	1					70
	10	5	5	2						12
0.3	5	27	41	32	17	4				121
	6	28	47	59	28	6	1	1		170
	7	31	61	54	20	19	1			186
	8	12	32	24	22	8	1			99
	9	1	6	9	6	1	1			24
	10	1	2	1						4
0.6	5	16	32	48	49	15				160
	6	7	35	45	37	20	9			153
	7	5	22	27	36	17	9	3	1	120
	8	1	4	12	11	8	7	0	2	45
	9	0	0	2	2	2	0	1		7
	10	0	0	0	0	0	0	0	1	1

Tabela B2: Quantidade de fêmeas expostas ao ftalato de dietilexila (DEHP), classificadas de acordo com o número de implantes realizados e número de fetos ou filhotes afetados (mortos, absorvidos e malformados. Dados de Tyl et al (1983).

Dose (g/kg)	Nº de implantes	Nº de filhotes afetados																				Nº de fêmeas
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
0.0	8	0	0	0	0	0	1															1
	9	0	0	1																		1
	10	0	2																			2
	11	1	0	0	1																	2
	12	2	0	1	0	1	1															5
	13	3	1	1	1	1																7
	14	1	1	0	1	0	1															4
	15	1																				1
	16	1	2	0	0	0	0	0	0	0	1	1										5
	17	0	1																			1
	18	0	0	0	0	0	0	0	0	0	0	0	1									1
0.044	4	0	1																			1
	8	0	0	0	1																	1
	9	1	0	1																		2
	10	0	1																			1
	11	1	1																			2
	12	0	2	1																		3
	13	2	4	0	1	1																8
	14	0	2	1																		3
	15	0	2	1	0	1																4
	16	0	0	1																		1
0.091	6	0	0	0	1																	1
	8	1	0	1																		2
	9	0	0	0	1																	1
	11	0	1	1	0	1	1	1														5
	12	1	1	1	1																	4
	13	0	1	0	0	1																2
	14	1	0	2	0	1	1	0	0	1	1											7
	15	0	0	2	1																	3
	16	0	0	0	0	0	0	1														1
0.191	1	0	1																			1
	2	0	0	1																		1
	4	0	0	0	0	1																1
	6	0	0	0	0	0	0	1														1
	10	0	0	0	1	0	0	0	0	0	0	1										2
	11	0	0	1	0	1	0	1	0	0	1											4
	12	0	0	0	0	0	0	0	0	0	0	1	0	2								3
	13	0	0	0	0	0	1	0	0	0	0	0	0	1								2
	14	0	0	1	0	0	0	1	0	0	0	1	2									5
	15	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1					2
	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1			1
	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
0.292	9	0	0	0	0	0	0	0	0	0	1											1
	10	0	0	0	0	0	0	0	0	0	0	2										2
	11	0	0	0	0	0	0	0	0	0	0	0	5									5
	12	0	0	0	0	0	0	0	0	0	0	0	0	6								6
	13	0	0	0	0	0	0	0	0	1	0	0	0	1	5							7
	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2						2
	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1				1
	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1			1

A P Ê N D I C E C

ESTIMAÇÃO: ASPECTOS COMPUTACIONAIS

Neste Apêndice foram coletadas informações sobre os aspectos computacionais correspondentes ao método de estimação por máxima verossimilhança dos parâmetros dos modelos relacionados no Capítulo 3.

Os procedimentos numéricos discutidos no Capítulo 4 encontram-se programados nas subrotinas E04JBF, E04KBF e E04LBF da biblioteca NAG (1982) e na subrotina OPTIM, que é parte do programa MULTI80 desenvolvido por RAI & VAN RYZIN (1980a). Estas subrotinas são descritas na seção C1 e apresentadas no Apêndice E. Na seção C2, é apresentada a estrutura necessária à matriz de entrada dos dados. Finalmente, na seção C3, as funções de verossimilhança e suas derivadas são escritas com uma nova notação, objetivando maior rapidez no procedimento de estimação.

C1. APRESENTAÇÃO DAS SUBROTINAS

Na subrotina OPTIM, veja RAI & VAN RYZIN (1980a), encontra-se programada uma modificação do método de Fletcher & Reeves, onde as derivadas são estimadas por diferenças finitas e a busca linear é feita dentro da região factível (no caso, a de positividade). O processo pára quando a mudança relativa no valor da função log-verossimilhança é menor que 10^{-11} ou quando as diferenças relativas nos valores dos parâmetros são menores que 10^{-10}. Para utilizar esta subrotina, o usuário fornece apenas a função de interesse.

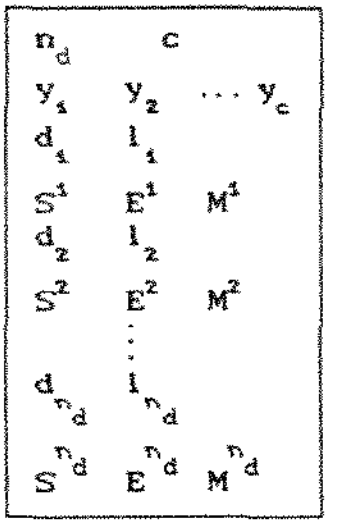
As subrotinas E04JBF e E04KBF pertencentes à biblioteca NAG (1982) resolvem problemas de minimização de funções de várias variáveis com canalizações, pelo método quase-Newton. A primeira subrotina aproxima, por diferenças finitas, as derivadas da função. Assim, o usuário fornece apenas a função de interesse. Na segunda subrotina, além da função de interesse, o usuário fornece também o vetor gradiente. Como as derivadas da função são calculadas de maneira exata, esta subrotina foi a preferida em relação a E04JBF. Na verdade, E04JBF foi usada apenas na estimação dos parâmetros da distribuição binomial negativa para o tamanho da ninhada com os dados de Tyl et al (1983), pois neste caso, a utilização de E04KBF produziu "overflow".

Outra subrotina da NAG (1982), denominada E04LBF, minimiza funções de várias variáveis sujeitas a canalizações, pelo método de Newton. Para utilizá-la, o usuário fornece a função, o vetor gradiente e a matriz hessiana.

Em seguida, vamos apresentar a estrutura da matriz de entrada dos dados.

C2. ENTRADA DOS DADOS

Com o objetivo de otimizar a leitura dos dados, à matriz de entrada dos dados será atribuída a seguinte estrutura:



A descrição é feita com o auxílio de três índices i, j e k, dois vetores S^k e M^k e uma matriz E^k.

O índice $i = 1, \ldots, l_k$ identifica o número de linhas de cada grupo em uma tabela análoga à Tabela B1 do Apêndice B, e l_k é a quantidade de diferentes tamanhos de ninhadas no k-ésimo grupo (ou o número de linhas da tabela para cada grupo) . O índice $j = 1, \ldots, c$ descreve o número máximo de colunas correspondentes à quantidade de filhotes com a resposta adversa, e c é o total de diferentes valores assumidos pela variável y_j, que é o número de respostas adversas diferentes observadas entre os filhotes da i-ésima ninhada no k-ésimo grupo experimental. O índice $k = 1, \ldots, n_d$ identifica grupo experimental, sendo n_d a quantidade de grupos experimentais. Observe que o índice j é hierarquicamente subordinado aos índices i e k.

A dose aplicada às fêmeas do k-ésimo grupo é d_k. O vetor S^k de dimensão l_k, tem como componentes as quantidades s_{ki}, que indicam os tamanhos das ninhadas produzidas pela i-ésima fêmea do k-ésimo grupo. A matriz E^k tem dimensão $l_k$xc, e elementos e_{kij} que denotam o número de fêmeas do k-ésimo grupo com tamanho de ninhada s_{ki} e com y_j

respostas adversas. O vetor M^k possui dimensão l_k, e componentes m_{ki}, que indicam o total de fêmeas do k-ésimo grupo que geraram uma ninhada de tamanho s_{ki}.

Para exemplificar a estrutura da matriz de entrada de dados, considere os dados de Lüning et al (1966), exibidos na Tabela B1 do Apêndice B. Neste caso, resultam

$n_d = 3$; $c = 8$; $y_1 = 0$, $y_2 = 1$, ..., $y_8 = 7$

e para o grupo controle,

$d_1 = 0.0$, $l_1 = 6$;

$s_{11} = 5$, $s_{12} = 6$, ..., $s_{16} = 10$;

$e_{111} = 30$, $e_{112} = 27$, ..., $e_{163} = 2$;

$m_{11} = 71$, $m_{12} = 156$, ..., $m_{16} = 12$.

Para agilizar o trabalho computacional, a função de verossimilhança e suas derivadas serão reescritas com a notação acima.

C3. MUDANÇA DE NOTAÇÃO

No processo de estimação, as funções log-verossimilhança $l(\underset{\sim}{\psi})$ e $l(\underset{\sim}{\gamma})$ dadas nas seções 3.2.1 serão reescritas utilizando os índices da leitura dos dados ($i = 1, ..., l_k$; $j = 1, ..., c$; $k = 1, ..., n_d$). Além disso, para o modelo geral RVR será feita uma transformação na variável θ_2, substituindo-a por $(\theta_2^* - \theta_1)/d_m$ onde d_m é a dose máxima aplicada, como explicado abaixo.

Para o modelo geral RVR, nosso objetivo é minimizar $-l(\underset{\sim}{\psi})$, o núcleo da função oposta à log-verossimilhança, sujeita às restrições $\alpha > 0$, $\beta > 0$, $\theta_1 > 0$ e $\theta_1 + \theta_2 d_k > 0$ para todo $k = 1, ..., n_d$. Com a nova notação, temos que esta função é

$$-l(\underset{\sim}{\psi}) = -\sum_{k=1}^{n_d}\sum_{i=1}^{l_k}\log(P_{ki})\sum_{j=1}^{c}e_{kij}y_j + \log(Q_{ki})\sum_{j=1}^{c}e_{kij}(s_{ki} - y_j),$$

onde $P_{ki} = P_{ki}(\underset{\sim}{\psi}) = [1 - \exp[-(\alpha + \beta d_k)]] \exp[-s_{ki}(\theta_1 + \theta_2 d_k)]$ e $Q_{ki} = Q_{ki}(\underset{\sim}{\psi}) = 1 - P_{ki}$.

A restrição $\theta_1 + \theta_2 d_k > 0$, para todo $k = 1, \ldots, n_d$ equivale à restrição $\theta_1 + \theta_2 d_m > 0$. Assim, nosso problema agora é a minimização de $-l(\underset{\sim}{\psi})$ sujeito às restrições $\alpha > 0$, $\beta > 0$, $\theta_1 > 0$ e $\theta_1 + \theta_2 d_m > 0$.

Fazendo $\theta_2^* = \theta_1 + \theta_2 d_m$, temos que $\theta_2 = (\theta_2^* - \theta_1)/d_m > 0$ e podemos reescrever $-l(\underset{\sim}{\psi})$ como

$$-l(\underset{\sim}{\psi}^*) = -\sum_{k=1}^{n_d}\sum_{i=1}^{l_k} \log(P_{ki}) \sum_{j=1}^{c} e_{kij} y_j + \log(Q_{ki}) \sum_{j=1}^{c} e_{kij}(s_{ki} - y_j),$$

onde

$P_{ki} = P_{ki}(\underset{\sim}{\psi}^*) = [1 - \exp[-(\alpha + \beta d_k)]] \exp[-s_{ki}[\theta_1 + d_k(\theta_2^* - \theta_1)/d_m]]$,

$Q_{ki} = Q_{ki}(\underset{\sim}{\psi}^*) = 1 - P_{ki}$, $\underset{\sim}{\psi}^* = (\alpha, \beta, \theta_1, \theta_2^*)$, $\theta_2^* = \theta_1 + \theta_2 d_m$.

Desta maneira, nosso problema agora é encontrar os pontos $\hat{\alpha}$, $\hat{\beta}$, $\hat{\theta}_1$ e $\hat{\theta}_2^*$ que minimizam $-l(\underset{\sim}{\psi}^*)$, sujeita às restrições $\alpha > 0$, $\beta > 0$, $\theta_1 > 0$ e $\hat{\theta}_2 > 0$. Tendo obtido estes pontos, basta fazer $\hat{\theta}_2 = (\hat{\theta}_2^* - \hat{\theta}_1)/d_m$ e teremos encontrado os valores que maximizam a função log-verossimilhança original.

Vamos agora apresentar, na nova notação, as funções de verossimilhança dos modelos considerados no Capítulo 3 e suas respectivas derivadas. Isto será feito porque estas expressões são necessárias nos procedimentos numéricos de estimação.

Antes porém de apresentar as derivadas da função $-l(\underset{\sim}{\psi}^*)$, será conveniente introduzir as quantidades T_{ki}, U_{ki}, V_{ki} e W_{ki}, onde

$$T_{ki} = \frac{1}{P_{ki}} \sum_{j=1}^{c} e_{kij} y_j - \frac{1}{Q_{ki}} \sum_{j=1}^{c} e_{kij} (s_{ki} - y_j),$$

$$U_{ki} = \frac{1}{P_{ki}^2} \sum_{j=1}^{c} e_{kij} y_j - \frac{1}{Q_{ki}^2} \sum_{j=1}^{c} e_{kij} (s_{ki} - y_j),$$

$$V_{ki} = T_{ki} - P_{ki} U_{ki} \qquad e$$

$$W_{ki} = T_{ki} + U_{ki} \exp[-(\alpha + \beta d_k)] \exp(-s_{ki}[\theta_1 + d_k(\theta_2^* - \theta_1)/d_m]).$$

As derivadas são da função $-l(\underset{\sim}{\psi}^*)$ são então:

$$\frac{-\partial l(\underset{\sim}{\psi})}{\partial \alpha} = - \sum_{k=1}^{n_d} \sum_{i=1}^{l_k} T_{ki} \exp[-(\alpha + \beta d_k)] \exp(-s_{ki}[\theta_1 + d_k(\theta_2^* - \theta_1)/d_m])$$

$$\frac{-\partial l(\underset{\sim}{\psi})}{\partial \beta} = - \sum_{k=1}^{n_d} \sum_{i=1}^{l_k} d_k T_{ki} \exp[-(\alpha + \beta d_k)] \exp(-s_{ki}[\theta_1 + d_k(\theta_2^* - \theta_1)/d_m])$$

$$\frac{-\partial l(\underset{\sim}{\psi})}{\partial \theta_1} = \sum_{k=1}^{n_d} \sum_{i=1}^{l_k} (1 - d_k/d_m) s_{ki} P_{ki} T_{ki}$$

$$\frac{-\partial l(\underset{\sim}{\psi})}{\partial \theta_2} = \sum_{k=1}^{n_d} \sum_{i=1}^{l_k} (d_k/d_m) s_{ki} P_{ki} T_{ki}$$

$$\frac{-\partial^2 l(\underset{\sim}{\psi})}{\partial \alpha^2} = \sum_{k=1}^{n_d} \sum_{i=1}^{l_k} W_{ki} \exp[-(\alpha + \beta d_k)] \exp(-s_{ki}[\theta_1 + d_k(\theta_2^* - \theta_1)/d_m])$$

$$\frac{-\partial^2 l(\underset{\sim}{\psi})}{\partial \alpha \partial \beta} = \sum_{k=1}^{n_d} \sum_{i=1}^{l_k} d_k W_{ki} \exp[-(\alpha + \beta d_k)] \exp(-s_{ki}[\theta_1 + d_k(\theta_2^* - \theta_1)/d_m])$$

$$\frac{-\partial^2 l(\underset{\sim}{\psi})}{\partial \alpha \partial \theta_1} = \sum_{k=1}^{n_d} \sum_{i=1}^{l_k} (1-d_k/d_m) V_{ki} s_{ki} \exp[-(\alpha+\beta d_k)] \exp(-s_{ki}[\theta_1+d_k(\theta_2^*-\theta_1)/d_m])$$

$$\frac{-\partial^2 l(\underset{\sim}{\psi})}{\partial \alpha \partial \theta_2} = \sum_{k=1}^{n_d} \sum_{i=1}^{l_k} (d_k/d_m) V_{ki} s_{ki} \exp[-(\alpha+\beta d_k)] \exp(-s_{ki}[\theta_1+d_k(\theta_2^*-\theta_1)/d_m])$$

$$\frac{-\partial^2 l(\psi)}{\partial \beta^2} = \sum_{k=1}^{n_d} \sum_{i=1}^{l_k} d_k^2 \, W_{ki} \exp[-(\alpha + \beta d_k)] \exp(-s_{ki}[\theta_1 + d_k(\theta_2^* - \theta_1)/d_m])$$

$$\frac{-\partial^2 l(\psi)}{\partial \beta \partial \theta_1} = \sum_{k=1}^{n_d} \sum_{i=1}^{l_k} d_k (1 - d_k/d_m) V_{ki} s_{ki} \exp[-(\alpha+\beta d_k)] \exp(-s_{ki}[\theta_1 + d_k(\theta_2^* - \theta_1)/d_m])$$

$$\frac{-\partial^2 l(\psi)}{\partial \beta \partial \theta_2} = \sum_{k=1}^{n_d} \sum_{i=1}^{l_k} d_k (d_k/d_m) V_{ki} s_{ki} \exp[-(\alpha+\beta d_k)] \exp(-s_{ki}[\theta_1 + d_k(\theta_2^* - \theta_1)/d_m])$$

$$\frac{-\partial^2 l(\psi)}{\partial \theta_1^2} = - \sum_{k=1}^{n_d} \sum_{i=1}^{l_k} [(1 - d_k/d_m) \, s_{ki}]^2 \, P_{ki} \, V_{ki}$$

$$\frac{-\partial^2 l(\psi)}{\partial \theta_1 \partial \theta_2} = - \sum_{k=1}^{n_d} \sum_{i=1}^{l_k} (1 - d_k/d_m) \, d_k/d_m \, s_{ki}^2 \, P_{ki} \, V_{ki}$$

$$\frac{-\partial^2 l(\psi)}{\partial \theta_2^2} = - \sum_{k=1}^{n_d} \sum_{i=1}^{l_k} (d_k/d_m \, s_{ki})^2 \, P_{ki} \, V_{ki}$$

No modelo de Poisson, o núcleo da função -log-verossimilhança e suas derivadas estão dadas pelas expressões:

$$-l(\gamma) = \sum_{k=1}^{n_d} n_k \phi_1 \exp(-\phi_2 d_k) - \log[\phi_1 \exp(-\phi_2 d_k)] \sum_{i=1}^{l_k} s_{ki} m_{ki}$$

$$\frac{-\partial l(\gamma)}{\partial \phi_1} = \sum_{k=1}^{n_d} n_k \exp(-\phi_2 d_k) - \frac{1}{\phi_1} \sum_{i=1}^{l_k} s_{ki} m_{ki}$$

$$\frac{-\partial l(\gamma)}{\partial \phi_2} = \sum_{k=1}^{n_d} -n_k d_k \phi_1 \exp(-\phi_2 d_k) + d_k \sum_{i=1}^{l_k} s_{ki} m_{ki}$$

$$\frac{-\partial^2 l(\gamma)}{\partial \phi_1^2} = \sum_{k=1}^{n_d} \frac{1}{\phi_1^2} \sum_{i=1}^{l_k} s_{ki} m_{ki}$$

$$\frac{-\partial^2 l(\gamma)}{\partial \phi_1 \partial \phi_2} = \sum_{k=1}^{n_d} -n_k d_k \exp(-\phi_2 d_k)$$

$$\frac{-\partial^2 l(\gamma)}{\partial \phi_2^2} = \sum_{k=1}^{n_d} n_k d_k^2 \phi_1 \exp(-\phi_2 d_k)$$

Finalmente, as correspondentes expressões para o modelo binomial negativo, são as seguintes:

$$-l(\gamma) = -l(\delta_1, \delta_2, k) = \sum_{k=1}^{n_d} \left\{ -\sum_{i=1}^{l_k} m_{ki} \log\left[\prod_{r=0}^{s_{ki}-1}(k+r)\right] - \right.$$

$$\left. - \log\left(\frac{\delta_1 \exp(-\delta_2 d_k)}{\delta_1 \exp(-\delta_2 d_k) + k}\right) \sum_{i=1}^{l_k} s_{ki} m_{ki} + k n_k \log\left(\frac{\delta_1 \exp(-\delta_2 d_k) + k}{k}\right) \right\}$$

$$\frac{-\partial l(\gamma)}{\partial \delta_1} = \sum_{k=1}^{n_d} \frac{k}{\delta_1 \exp(-\delta_2 d_k) + k} \left[\frac{-1}{\delta_1} \sum_{i=1}^{l_k} s_{ki} m_{ki} + n_k \exp(-\delta_2 d_k) \right]$$

$$\frac{-\partial l(\gamma)}{\partial \delta_2} = \sum_{k=1}^{n_d} \frac{k d_k}{\delta_1 \exp(-\delta_2 d_k) + k} \left[\sum_{i=1}^{l_k} s_{ki} m_{ki} - n_k \delta_1 \exp(-\delta_2 d_k) \right]$$

$$\frac{-\partial l(\gamma)}{\partial k} = \sum_{k=1}^{n_d} \left\{ -\sum_{i=1}^{l_k} m_{ki} \sum_{r=0}^{s_{ki}-1} \frac{1}{k+r} + \frac{1}{\delta_1 \exp(-\delta_2 d_k) + k} \sum_{i=1}^{l_k} s_{ki} m_{ki} + \right.$$

$$\left. + n_k \log\left(\frac{\delta_1 \exp(-\delta_2 d_k) + k}{k}\right) - \frac{n_k \delta_1 \exp(-\delta_2 d_k)}{\delta_1 \exp(-\delta_2 d_k) + k} \right\}$$

$$\frac{-\partial^2 l(\gamma)}{\partial \delta_1^2} = \sum_{k=1}^{n_d} \left\{ \frac{k}{[\delta_1 \exp(-\delta_2 d_k) + k]\delta_1)^2} \left[[2\delta_1 \exp(-\delta_2 d_k) + k] \sum_{i=1}^{l_k} s_{ki} m_{ki} - \right. \right.$$

$$\left. \left. - n_k \exp(-\delta_2 d_k)^2 \right] \right\}$$

$$\frac{-\partial^2 l(\gamma)}{\partial \delta_1 \partial \delta_2} = \sum_{k=1}^{n_d} \frac{-k d_k \exp(-\delta_2 d_k)}{[\delta_1 \exp(-\delta_2 d_k) + k]^2} \left[\sum_{i=1}^{l_k} s_{ki} m_{ki} + n_k k \right]$$

$$\frac{-\partial^2 l(\gamma)}{\partial \delta_1 \partial k} = \sum_{k=1}^{n_d} \frac{\delta_1 \exp(-\delta_2 d_k)}{[\delta_1 \exp(-\delta_2 d_k) + k]^2} \left[\frac{-1}{\delta_1} \sum_{i=1}^{l_k} s_{ki} m_{ki} + n_k \exp(-\delta_2 d_k) \right]$$

$$\frac{-\partial^2 l(\gamma)}{\partial \delta_2^2} = \sum_{k=1}^{n_d} \frac{k d_k^2 \delta_1 \exp(-\delta_2 d_k)}{[\delta_1 \exp(-\delta_2 d_k) + k]^2} \left[\sum_{i=1}^{l_k} s_{ki} m_{ki} + n_k k) \right]$$

$$\frac{-\partial^2 l(\gamma)}{\partial \delta_2 \partial k} = \sum_{k=1}^{n_d} \frac{\delta_1 \exp(-\delta_2 d_k) d_k}{[\delta_1 \exp(-\delta_2 d_k) + k]^2} \left[\sum_{i=1}^{l_k} s_{ki} m_{ki} - n_k \delta_1 \exp(-\delta_2 d_k) \right]$$

$$\frac{-\partial^2 l(\gamma)}{\partial k^2} = \sum_{k=1}^{n_d} \left\{ \sum_{i=1}^{l_k} m_{ki} \left(\sum_{r=0}^{a_{ki}-1} \frac{1}{(k+r)^2} \right) - \right.$$

$$\left. - \frac{1}{[\delta_1 \exp(-\delta_2 d_k) + k]^2} \left[\sum_{i=1}^{l_k} s_{ki} m_{ki} + \frac{n_k [\delta_1 \exp(-\delta_2 d_k)]^2}{k} \right] \right\}.$$

APÊNDICE D

RESULTADOS NUMÉRICOS

Neste Apêndice encontram-se tabelas correspondentes aos resultados numéricos obtidos com o auxílio

(a) dos métodos numéricos de Fletcher & Reeves, quase-Newton e de Newton, descritos no Capítulo 4 e utilizados na estimação dos parâmetros do modelo condicional geral RVR, e dos modelos de Poisson e binomial negativo para o tamanho da ninhada.

(b) do pacote estatístico GLIM versão 3.77, utilizado na estimação dos parâmetros dos modelos logísticos lineares com variação binomial pura discutidos no Capítulo 5.

Nas seções D1 e D2 encontram-se, respectivamente, os resultados de (a) e (b).

D1. ESTIMATIVAS DE MÁXIMA VEROSSIMILHANÇA: MODELOS DO CAPÍTULO 4

Tabela D1.1: Estimativas $\hat{\underset{\sim}{\psi}}$ de máxima verossimilhança para o parâmetro $\underset{\sim}{\psi} = (\alpha, \beta, \theta_1, \theta_2)'$ do modelo RVR ajustado aos dados de Lüning et al (1966), obtidas pelos métodos de Fletcher & Reeves, quase-Newton e de Newton, com ponto inicial $\underset{\sim}{\psi}_o$, número de iterações NI e número NF de avaliações da função. O máximo da função log-verossimilhança é denotado por $l(\hat{\underset{\sim}{\psi}})$.

Método	$\underset{\sim}{\psi}_o$	$\hat{\underset{\sim}{\psi}}$	$l(\hat{\underset{\sim}{\psi}})$	NI	NF
Fletcher & Reeves	(0.1 2.0 0.6 0.1)	(0.2238 1.0383 0.0963 0.0562)	5795.470	322	2794
	(0.6 0.5 0.4 0.1)	(0.2238 1.0359 0.0962 0.0560)	5795.470	145	1499
	(5.2 1.4 1.7 0.3)	(0.2235 1.0350 0.0962 0.0559)	5795.470	190	2011
	(1.5 3.0 0.6 0.4)	(0.2236 1.0352 0.0961 0.0559)	5795.470	191	1992
	(0.23 1.04 0.1 0.06)	(0.2241 1.0408 0.0965 0.0564)	5795.470	18	189
Quase Newton	(0.1 2.0 0.6 0.1)	(0.2231 1.0362 0.0958 0.0560)	5795.471	14	65
	(0.6 0.5 0.4 0.1)	(0.2234 1.0267 0.0959 0.0553)	5795.473	13	58
	(5.2 1.4 1.7 0.3)	(0.2239 1.0375 0.0963 0.0562)	5795.470	29	134
	(1.5 3.0 0.6 0.4)	(0.2240 1.0384 0.0964 0.0562)	5795.470	20	97
	(0.23 1.04 0.1 0.06)	(0.2238 1.0377 0.0963 0.0562)	5795.470	6	46
Newton	(0.1 2.0 0.6 0.1)	(0.2233 1.0366 0.0961 0.0560)	5795.469	8	13
	(0.6 0.5 0.4 0.1)	(0.2238 1.0377 0.0963 0.0562)	5795.470	7	13
	(5.2 1.4 1.7 0.3)	(0.2239 1.0383 0.0963 0.0562)	5795.471	17	39
	(1.5 3.0 0.6 0.4)	(0.2238 1.0380 0.0963 0.0562)	5795.470	11	28
	(0.23 1.04 0.1 0.06)	(0.2238 1.0378 0.0963 0.0562)	5795.470	3	5

Observação: As restrições aplicadas em todos os casos são da forma $1.0\text{x}10^{-6} \leq \psi_i \leq 1.0\text{x}10^{6}$ para $i = 1, 2, 3, 4$.

Tabela D1.2: Estimativas $\hat{\underset{\sim}{\gamma}}$ de máxima verossimilhança para o parâmetro $\underset{\sim}{\gamma} = (\gamma_1, \gamma_2)'$ da distribuição de Poisson para o tamanho da ninhada ajustada aos dados de Lüning et al (1966), obtidas pelos métodos de Fletcher & Reeves, quase-Newton e de Newton, com ponto inicial $\underset{\sim}{\gamma}_0$, número de iterações NI e número NF de avaliações da função. O máximo da função log-verossimilhança é $l(\hat{\underset{\sim}{\gamma}}) = -10523.78$.

Método	$\underset{\sim}{\gamma}_0$	$\hat{\underset{\sim}{\gamma}}$	NI	NF
Fletcher & Reeves	(2.0 0.9)	(7.0417 0.2264)	19	139
	(4.6 1.2)	(7.0417 0.2264)	13	99
	(3.4 0.8)	(7.0417 0.2264)	17	115
	(1.5 0.2)	(7.0417 0.2264)	20	127
	(7.0 0.2)	(7.0417 0.2264)	14	100
Quase Newton	(2.0 0.9)	(7.0422 0.2262)	7	42
	(4.6 1.2)	(7.0413 0.2267)	4	34
	(3.4 0.8)	(7.0427 0.2266)	6	43
	(1.5 0.2)	(7.0419 0.2264)	6	43
	(7.0 0.2)	(7.0416 0.2263)	3	30
Newton	(2.0 0.9)	(7.0411 0.2262)	6	12
	(4.6 1.2)	(7.0413 0.2263)	5	8
	(3.4 0.8)	(7.0417 0.2264)	5	7
	(1.5 0.2)	(7.0411 0.2262)	7	15
	(7.0 0.2)	(7.0418 0.2264)	2	5

Observações: i) As restrições aplicadas neste caso são da forma $1.0\text{x}10^{-6} \leq \gamma_1 \leq 1.0\text{x}10^{6}$ e $-1.0\text{x}10^{6} \leq \gamma_2 \leq 1.0\text{x}10^{6}$.

ii) Alguns pontos, tais como (0.1 0.2), (0.2 0.1) e (0.8 3.4) levaram a "overflow" para o método de Newton.

Tabela D1.3: Estimativa $\hat{\underset{\sim}{\gamma}}$ de máxima verossimilhança para o parâmetro $\underset{\sim}{\gamma} = (\gamma_1, \gamma_2, \gamma_3)'$ da distribuição binomial negativa para o tamanho da ninhada ajustada aos dados de Lüning et al (1966), obtidas pelos métodos de Fletcher & Reeves, quase-Newton e de Newton com ponto inicial $\underset{\sim}{\gamma}_o$, número de iterações NI e número NF de avaliações da função. O máximo da função log-verossimilhança é denotado por $l(\hat{\underset{\sim}{\gamma}})$.

Método	$\underset{\sim}{\gamma}_o$	$\hat{\underset{\sim}{\gamma}}$	$l(\hat{\underset{\sim}{\gamma}})$	NI	NF
Fletcher & Reeves	(10.0 4.0 9.0)	(7.0417 0.2264 2369.036)	-10521.77	3867	32500
	(1.0 2.0 3.0)	(7.0421 0.2265 1780.417)	-10521.00	3922	32494
	(1.9 0.7 2.4)	(7.0417 0.2264 2353.434)	-10521.76	3936	32496
	(6.0 0.4 100.0)	(7.0417 0.2264 1663.629)	-10520.93	3814	32493
	(7.0 0.2 1000.0)	(7.0418 0.2264 1809.712)	-10521.16	3796	32493
Quase Newton	(10.0 4.0 9.0)	(7.0396 0.2239 514.7524)	-10514.65	38	175
	(1.0 2.0 3.0)	(7.0705 0.2329 584.0160)	-10515.72	48	218
	(1.9 0.7 2.4)	(7.0484 0.2238 910.4848)	-10518.65	44	192
	(6.0 0.4 100.0)	(7.0488 0.2281 527.5726)	-10514.86	25	123
	(7.0 0.2 1000.0)	(7.0232 0.2211 1000.0000)	-10519.09	3	52
Newton	(10.0 4.0 9.0)	(7.0469 0.2274 10.4939)	-10151.88	14	28
	(1.0 2.0 3.0)	(7.0418 0.2264 8.1146)	-10065.10	29	64
	(1.9 0.7 2.4)	(7.0447 0.2270 8.0535)	-10062.33	47	118
	(6.0 0.4 100.0)	(7.0417 0.2264 100.0663)	-10447.44	3	4
	(7.0 0.2 1000.0)	(7.0417 0.2263 1000.0358)	-10519.08	2	8

Observações: i) As restrições aplicadas neste caso são da forma $1.0\text{x}10^{-6} \leq \gamma_1 \leq 1.0\text{x}10^{6}$, $-1.0\text{x}10^{6} \leq \gamma_2 \leq 1.0\text{x}10^{6}$ e $1.0\text{x}10^{-6} \leq \gamma_3 \leq 1.0\text{x}10^{6}$.

ii) O número máximo de avaliações da função permitido foi 32500 para o método de Fletcher & Reeves.

iii) O ponto (7.0 0.2 $1.0\text{x}10^{6}$) levou à "overflow" para o método Quase-Newton.

Tabela D1.4: Estimativas $\hat{\psi}$ de máxima verossimilhança para o parâmetro $\psi = (\alpha, \beta, \theta_1, \theta_2)'$ do modelo RVR ajustado aos dados de Tyl et al (1983), obtidas pelos métodos de Fletcher & Reeves, quase-Newton e de Newton, com ponto inicial ψ_o, número de iterações NI e número NF de avaliações da função. O máximo da função log-verossimilhança é denotado por $l(\hat{\psi})$.

Método	ψ_o				$\hat{\psi}$				$l(\hat{\psi})$	NI	NF
	(0.9	5.7	0.2	0.1)	(1.6576	6.7525	0.1237	1.0e-6)	730.0157	179	1962
Fletcher	(0.9	6.4	0.3	0.1)	(0.9090	6.4010	0.1038	1.0e-6)	732.7340	6	55
&	(0.4	3.9	1.5	0.01)	(0.7939	9.7314	0.1059	1.0e-6)	726.4706	1142	12794
Reeves	(0.4	0.1	0.2	0.1)	(3.6462	1.1420	0.1357	1.0e-6)	733.1726	98	1015
	(0.79	10.0	0.11	1.0e-6)	(0.7877	9.9943	0.1062	1.0e-6)	726.4831	28	315
	(0.9	5.7	0.2	0.1)	(0.7930	9.8505	0.1062	1.0e-6)	726.4685	10	43
	(0.9	6.4	0.3	0.1)	(0.7942	9.8021	0.1061	1.0e-6)	726.4684	8	43
Quase Newton	(0.4	3.9	1.5	0.01)	(0.7903	9.8010	0.1059	1.0e-6)	726.4684	16	73
	(0.4	0.1	0.2	0.1)	(0.7927	9.8047	0.1060	1.0e-6)	726.4685	15	58
	(0.79	10.0	0.11	1.0e-6)	(0.7919	9.8036	0.1060	1.0e-6)	726.4684	5	34
	(0.9	5.7	0.2	0.1)	(0.7896	9.8065	0.1059	1.0e-6)	726.4684	7	15
	(0.9	6.4	0.3	0.1)	(0.7922	9.8031	0.1060	1.0e-6)	726.4684	8	15
Newton	(0.4	3.9	1.5	0.01)	(0.7918	9.8036	0.1060	1.0e-6)	726.4684	8	11
	(0.4	0.1	0.2	0.1)	(0.7919	9.8034	0.1060	1.0e-6)	726.4684	13	16
	(0.79	10.0	0.11	1.0e-6)	(0.7919	9.8034	0.1060	1.0e-6)	726.4686	2	3

Observação: As restrições aplicadas em todos os casos são da forma $1.0\text{x}10^{-6} \leq \psi_i \leq 1.0\text{x}10^{6}$ para $i = 1, 2, 3, 4$.

Tabela D1.5: Estimativas $\hat{\underset{\sim}{\gamma}}$ de máxima verossimilhança para o parâmetro $\underset{\sim}{\gamma} = (\gamma_1, \gamma_2)'$ da distribuição de Poisson para o tamanho da ninhada ajustada aos dados de Tyl et al (1983), obtidas pelos métodos de Fletcher & Reeves, quase-Newton e de Newton, com ponto inicial $\underset{\sim}{\gamma}_0$, número de iterações NI e número NF de avaliações da função. O máximo da função log-verossimilhança é $l(\hat{\underset{\sim}{\gamma}}) = -2452.32$.

Método	$\underset{\sim}{\gamma}_0$	$\hat{\underset{\sim}{\gamma}}$	NI	NF
Fletcher & Reeves	(2.0 0.5)	(12.7228 0.2499)	20	128
	(4.2 0.7)	(12.7228 0.2499)	37	199
	(3.7 0.4)	(12.7228 0.2499)	19	114
	(10.5 3.8)	(12.7228 0.2499)	10	76
	(13.0 0.2)	(12.7228 0.2499)	13	93
Quase Newton	(2.0 0.5)	(12.7224 0.2495)	8	46
	(4.2 0.7)	(12.7223 0.2499)	6	38
	(3.7 0.4)	(12.7247 0.2507)	8	43
	(10.5 3.8)	(12.7218 0.2498)	6	42
	(13.0 0.2)	(12.7224 0.2500)	4	31
Newton	(2.0 0.5)	(12.7228 0.2500)	8	11
	(4.2 0.7)	(12.7223 0.2500)	6	8
	(3.7 0.4)	(12.7226 0.2499)	7	10
	(10.5 3.8)	(12.7121 0.2471)	4	13
	(13.0 0.2)	(12.7228 0.2500)	2	3

Observações: i) As restrições aplicadas neste caso são da forma $1.0\text{x}10^{-6} \leq \gamma_1 \leq 1.0\text{x}10^{6}$ e $-1.0\text{x}10^{6} \leq \gamma_2 \leq 1.0\text{x}10^{6}$.

ii) Alguns pontos, tais como (0.2 1.5), levaram a "overflow" para o método de Newton.

Tabela D1.6: Estimativa $\hat{\underset{\sim}{\gamma}}$ de máxima verossimilhança para o parâmetro $\underset{\sim}{\gamma} = (\gamma_1, \gamma_2, \gamma_3)'$ da distribuição binomial negativa para o tamanho da ninhada ajustada aos dados de Tyl et al (1983), obtidas pelos métodos de Fletcher & Reeves, quase-Newton e de Newton com ponto inicial $\underset{\sim}{\gamma}_0$, número de iterações NI e número NF de avaliações da função. O máximo da função log-verossimilhança é denotado por $l(\hat{\underset{\sim}{\gamma}})$.

Método	$\underset{\sim}{\gamma}_0$	$\hat{\underset{\sim}{\gamma}}$	$l(\hat{\underset{\sim}{\gamma}})$	NI	NF
Fletcher & Reeves	(6.7 1.3 4.3)	(12.7228 0.2500 2629.928)	-2452.218	3896	32494
	(9.3 0.1 4.9)	(12.7228 0.2500 2849.893)	-2452.218	3944	32493
	(15.0 3.0 8.0)	(12.7228 0.2500 2485.199)	-2452.204	3956	32494
	(6.9 1.2 9.0)	(12.7228 0.2500 3972.492)	-2452.247	2325	19417
	(12.0 0.2 10.0)	(12.7228 0.2500 3571.401)	-2452.239	3052	25374
Quase Newton	(6.7 1.3 4.3)	(12.7171 0.2460 4.3)	-2397.748	8	79
	(9.3 0.1 4.9)	(12.6714 0.2287 4.9)	-2402.768	5	64
	(15.0 3.0 8.0)	(12.7424 0.2500 8.0)	-2418.774	5	71
	(6.9 1.2 9.0)	(12.6812 0.2325 9.0)	-2421.965	8	89
	(12.0 0.2 10.0)	(12.7034 0.2431 10.0)	-2424.616	4	61
Newton	(6.7 1.3 4.3)	(12.7392 0.2517 4.8942)	-2402.722	16	35
	(9.3 0.1 4.9)	(12.7164 0.2457 5.4053)	-2406.338	10	20
	(15.0 3.0 8.0)	(12.7174 0.2463 8.4474)	-2420.288	7	13
	(6.9 1.2 9.0)	(12.7175 0.2464 9.0286)	-2422.049	5	6
	(12.0 0.2 10.0)	(12.7179 0.2467 10.7725)	-2426.374	5	11

Observações: i) As restrições aplicadas neste caso são da forma $1.0\text{x}10^{-6} \leq \gamma_1 \leq 1.0\text{x}10^{6}$, $-1.0\text{x}10^{6} \leq \gamma_2 \leq 1.0\text{x}10^{6}$ e $1.0\text{x}10^{-6} \leq \gamma_3 \leq 1.0\text{x}10^{6}$.

ii) Para o método de Fletcher & Reeves, o ponto (12.0 0.2 $1.0\text{x}10^{6}$) levou em 16 iterações e 128 avaliações da função, ao valor -2452.319 da função no ponto (12.7228, 0.25, $1.0\text{x}10^{6}$).

D2. ESTIMATIVAS DE MÁXIMA VEROSSIMILHANÇA: MODELOS DO CAPITULO 5

Tabela D2.1: Estimativas dos parâmetros dos modelos logísticos produzidas pelo pacote estatístico GLIM, para os dados de Lüning et al (1966).

Modelo	Estimativa	Erro-padrão	Parâmetro
I	-1.593	0.1416	τ_{01}
	-2.181	0.1080	τ_{02}
	-2.154	0.08253	τ_{03}
	-2.355	0.1021	τ_{04}
	-2.495	0.1495	τ_{05}
	-2.512	0.3451	τ_{06}
	-0.9232	0.09004	τ_{17}
	-0.9523	0.06981	τ_{18}
	-1.163	0.06506	τ_{19}
	-1.202	0.08430	τ_{11}
	-1.174	0.1602	τ_{11}
	-2.197	0.5266	τ_{11}
	-0.3279	0.07162	τ_{21}
	-0.4337	0.06751	τ_{21}
	-0.4754	0.07092	τ_{21}
	-0.4287	0.1078	τ_{21}
	-0.4855	0.2593	τ_{21}
II	-0.9706	0.2913	τ_0
	-0.3499	0.2181	τ_1
	-0.1492	0.2264	τ_2
	-0.1723	0.04073	γ_0
	-0.1090	0.03207	γ_1
	-0.04245	0.03540	γ_2
III	-0.9650	0.2556	τ_0
	-0.3565	0.1492	τ_1
	-0.1454	0.2058	τ_2
	-0.1731	0.03560	γ
	0.2167	0.08958	δ
IV	-1.463	0.1539	τ_0
	-0.3913	0.1424	τ_1
	0.2314	0.1345	τ_2
	-0.1029	0.02052	γ
V	-0.7156	0.2427	α
	0.8287	0.5836	β
	-0.1952	0.03450	γ
	0.2882	0.08618	δ
VI	-2.202	0.04802	τ_0
	-1.085	0.03649	τ_1
	-0.4172	0.03741	τ_2
VII	-1.352	0.1518	α
	2.755	0.1020	β
	-0.1031	0.02053	γ
VIII	-2.092	0.04042	α
	2.906	0.09783	β

Tabela D2.2: Estimativas dos parâmetros dos modelos logísticos produzidas pelo pacote estatístico GLIM, para os dados de Tyl et al (1983).

Modelo	Estimativa	Erro-padrão	Parâmetro
I	0.5108	0.7303	T01
	-1.253	0.8018	T02
	-2.197	0.7454	T03
	-1.846	0.6213	T04
	-1.494	0.3336	T05
	-2.092	0.3352	T06
	-1.653	0.3639	T07
	-13.48	132.7	T08
	-1.033	0.2541	T09
	-2.773	1.031	T010
	0.4520	0.4835	T011
	-1.099	1.155	T112
	-0.5108	0.7303	T113
	-2.079	0.7500	T114
	-2.197	1.054	T115
	-3.045	1.024	T116
	-2.079	0.5303	T117
	-2.135	0.3188	T118
	-2.251	0.5257	T119
	-1.872	0.3798	T120
	-1.946	0.7559	T121
	-8.167e-1	0.8165	T222
	-1.946	0.7559	T223
	-0.6931	0.7071	T224
	-0.7205	0.2874	T225
	-1.946	0.4364	T226
	-1.435	0.4976	T227
	-0.8183	0.2192	T228
	-1.692	0.4113	T229
	-0.5108	0.5164	T230
	10.23	100.8	T331
	10.26	72.59	T332
	10.56	59.52	T333
	10.82	55.26	T334
	0.6190	0.4688	T335
	-0.09097	0.3018	T336
	2.833	0.7276	T337
	0.6360	0.4122	T338
	0.2877	0.2415	T339
	0.4055	0.3727	T340
	2.833	1.029	T341
	11.75	49.49	T342
	11.12	52.44	T443
	11.20	36.68	T444
	11.28	23.00	T445
	11.35	20.84	T446
	2.651	0.4224	T447
	11.48	35.66	T448
	11.60	49.98	T449
	11.65	49.80	T450

Tabela D2.2: Estimativas dos parâmetros dos modelos logísticos produzidas pelo pacote estatístico GLIM, para os dados de Tyl et al (1983). Continuação.

Modelo	Estimativa	Erro-padrão	Parâmetro
II	-2.714	0.8034	τ_0
	-1.268	0.9948	τ_1
	-0.9631	0.7632	τ_2
	0.3405	0.5547	τ_3
	5.535	2.788	τ_4
	0.09154	0.05701	γ_1
	-0.06014	0.07751	γ_2
	-0.01031	0.05911	γ_3
	0.03721	0.04146	γ_4
	-0.1266	0.2123	γ_5
III	-2.024	0.6481	τ_0
	-2.494	0.4921	τ_1
	-1.467	0.3848	τ_2
	0.6244	0.5179	τ_3
	3.899	0.9302	τ_4
	0.04162	0.04620	γ
	-0.1371	0.3343	δ
IV	-1.815	0.3966	τ_0
	-2.375	0.3952	τ_1
	-1.432	0.3739	τ_2
	0.4811	0.3803	τ_3
	3.585	0.5238	τ_4
	0.02641	0.02739	γ
V	-3.899	0.6711	α
	25.41	4.489	β
	0.1251	0.04922	γ
	-0.6602	0.3328	δ
VI	-1.454	0.1283	τ_0
	-2.035	0.1748	τ_1
	-1.094	0.1292	τ_2
	0.8267	0.1308	τ_3
	3.919	0.4117	τ_4
VII	-2.826	0.3813	α
	16.73	0.7792	β
	0.04451	0.02745	γ
VIII	-2.237	0.1075	α
	16.63	0.7757	β

A P Ê N D I C E E

PROGRAMAS

Neste Apêndice encontram-se os programas computacionais escritos na linguagem FORTRAN 77 e utilizados na estimação dos parâmetros dos modelos discutidos no Capítulo 3, além da sequência de comandos do pacote estatístico GLIM que foi utilizada no ajuste dos modelos do Capítulo 5.

Na seção E1 são exibidos os programas referentes ao método dos gradientes conjugados proposto por Fletcher & Reeves. Primeiramente, é apresentada a subrotina OPTIM desenvolvida por RAI & VAN RYZIN (1980a), para minimização de funções com restrição de positividade nas variáveis. Em seguida é apresentado o programa mestre, responsável pela leitura de dados, execução de cálculos auxiliares e chamada da subrotina de minimização. Em seguida, são listadas as três subrotinas que permitem calcular o oposto da função log-verossimilhança para os modelos condicional de dose-resposta, de Poisson e binomial Negativo, nesta ordem.

Os programas necessários à utilização das subrotinas E04JBF e E04KBF da biblioteca NAG (1982), que minimizam funções com canalizações (no nosso caso, restrições de positividade) pelo método quase-Newton de Broyden-Fletcher-Goldfarb-Shanno são apresentados na seção E2. A diferença entre estas duas subrotinas é que a segunda calcula diretamente o vetor gradiente da função a ser minimizada, enquanto que a primeira o aproxima por diferenças finitas.

Na seção E3 são listados os programas que serão utilizados com a subrotina E04LBF da NAG (1982). Esta subrotina permite a minimização de funções com canalizações pelo método de Newton.

Nas seções E2 e E3 os programas são listados na mesma ordem. Primeiro, temos os o programa mestre, com a mesma finalidade do programa mestre para o método de Fletcher & Reeves. Em seguida, são apresentadas as subrotinas que calculam as funções -log-verossimilhança e suas derivadas. Para as subrotinas referentes aos métodos quase-Newton e de Newton, imediatamente após as subrotinas que calculam as funções a serem minimizadas, deve ser inserida a subrotina MONIT, responsável pelo monitoramento do processo de minimização. Ela é exibida ao final da seção E3.

Na seção E4 é apresentado o programa que calcula as estimativas das matrizes de informação de Fisher para os modelos de dose-resposta e de Poisson calculadas na seção 3.4.1. Este programa permite ainda calcular as probabilidades de resposta adversa condicional e incondicional, dadas respectivamente nas seções 3.2 e 3.3.

À medida que os programas forem sendo apresentados, faremos breves comentários sobre suas variáveis. Estes comentários não serão repetidos em todos os programas, acompanhando as variáveis apenas nas suas primeiras aparições. Para tornar os programas mais claros, o usuário deverá conhecer a estrutura da matriz de entrada de dados descrita no Apêndice C deste trabalho.

A sequência de comandos do pacote estatístico GLIM,

utilizada no ajuste dos modelos logísticos do Capítulo 5 pode ser vista na seção E5. Estão separados os programas para os dados de Lüning **et al** (1966) e Tyl **et al** (1983). Observe ainda que a última série de comandos é utilizada na estimação do parâmetro de variação extra-binomial da maneira como foi apresentado em WILLIAMS (1982).

E1. METODO DOS GRADIENTES CONJUGADOS DE FLETCHER & REEVES

```
C ==========================================================
C Subrotina OPTIM, de RAI & VAN RYZIN (1980). Minimiza uma
C funcao de varias variaveis pelo metodo de FLETCHER & REEVES,
C sob a restricao de positividade das variaveis.
C ==========================================================

      SUBROUTINE OPTIM(N,X,F,G,S,XMIN,XMAX,EPS,DELX,FACC,MAXFN,KF,
     * ITER,IFN)

C ----------------------------------------------------------
C Esta subrotina busca o ponto que minimiza uma funcao de
C interesse f(x), onde x = (x1, ..., xn) sujeita as restricoes
C do tipo li (= xi (= ui, i = 1, ... n. E parte de MULTI80, um
C programa maior desenvolvido por RAI & VAN RYZIN (1980a). Usa
C um metodo modificado de Fletcher & Reeves, onde a derivada
C primeira e estimada por diferencas finitas.
C ----------------------------------------------------------

      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION X(*),G(*),S(*),XMIN(*),XMAX(*)

C ----------------------------------------------------------
C O usuario devera fornecer:
C
C N: dimensao do vetor x; especifica o numero de variaveis
C    independentes da funcao objetivo. Exemplo: modelo RVR n =
C    4, Poisson n = 2 e Bin. Neg. n = 3;
C
C X = (x1, ..., xn): ponto a partir do qual sera iniciado o
C     procedimento de minimizacao da funcao. Na saida contem a
C     a estimativa de maxima verossimilhanca;
C
C XMIN: vetor de dimensao n cujas componentes especificam o
C       limite inferior da correspondente componente de x;
C
C XMAX: vetor de dimensao n cujas componentes especificam o
C       limite superior da correspondente componente de x;
C
C EPS: precisao requerida no procedimento de maximizacao;
C
C DELX: incremento para o calculo da derivada;
C
C FACC: precisao da maquina;
C
C MAXFN: numero maximo de avaliacoes da funcao;
C
C KF: periodo para a impressao de resultados parciais.
C
C Parametros de saida:
C
C F: funcao objetivo f(x), avaliada em cada ponto. Na saida,
C    contem o valor minimo da funcao objetivo;
C
C G: vetor gradiente de dimensao n, que contem as derivadas
C    primeiras de f(x), avaliada em qualquer ponto;
C
C S: vetor direcao de dimensao n;
C
C ITER: numero de iteracoes efetuadas;
C
C IFN: numero efetuado de avaliacoes da funcao.
C ----------------------------------------------------------

      IP=KF
      KF=0
      ITER=0
      ICENT=0
      NP1=N+1
      IFN=1
      CALL FUNCT(N,X,F)

C ----------------------------------------------------------
C FUNCT(N,X,F): subrotina fornecida pelo usuario para
C calcular a funcao objetivo em cada ponto.
C ----------------------------------------------------------
```

```
      DF=9.0D-1*DABS(F)
      IF (IP .NE. 0) PRINT 9
      DO 9 I=1,N
9     G(I)=1.0D3
10    DO 11 I=1,N
11    S(I)=0.0D0
      GGP=1.0D0
      DO 60 ICYC=1,NP1
      IF (IP .EQ. 0) GO TO 12
      IF (MOD(ITER,IP) .NE. 0) GO TO 12
      PRINT 500,ITER,IFN
      PRINT 501,F
      PRINT 502,(X(I),I=1,N)
      PRINT 503,(G(I),I=1,N)
12    ITER=ITER+1
      NF=IFN+N+2
      IF (NF .GE. MAXFN) GO TO 80
      IF (ICENT .EQ. 1 .AND. NF+N .GE. MAXFN) GO TO 80
      DO 14 I=1,N
      SX=X(I)
      X(I)=SX+DELX
      IFN=IFN+1
      CALL FUNCT(N,X,FY)
      DELF=FY-F
      IF (DABS(DELF/F) .GE. FACC) GO TO 13
      G(I)=0.0D0
      GO TO 14
13    G(I)=DELF/DELX
14    X(I)=SX
      IF (ICENT .EQ. 0) GO TO 19
16    DO 18 I=1,N
      SX=X(I)
      X(I)=SX-DELX
      IFN=IFN+1
      CALL FUNCT(N,X,FY)
      DELF=F-FY
      IF (DABS(DELF/F) .LT. FACC) GO TO 18
      G(I)=5.0D-1*(G(I)+DELF/DELX)
18    X(I)=SX
19    CALL VPROD(G,G,N,GG)
      IF (GG .GT. 0.0D0) GO TO 20
      IF (ICENT .EQ. 1) GO TO 82
      ICENT=1
      GO TO 16
20    Z=GG/GGP
      ALMAX=1.0D+10
      DO 228 I=1,N
      SI=Z*S(I)-G(I)
      S(I)=SI
      IF (SI) 223,228,224
223   SI=-SI
      XLIM=X(I)-XMIN(I)
      GO TO 225
224   XLIM=XMAX(I)-X(I)
225   IF (XLIM .LE. EPS) GO TO 227
      ALMAX=DMIN1(ALMAX,XLIM/SI)
      GO TO 228
227   S(I)=0.0D0
228   CONTINUE
      CALL VPROD(S,G,N,GS)
      IF (GS .LT. 0.0D0) GO TO 21
      IF (GGP .NE. 1.0D0) GO TO 10
      IF (ICENT .EQ. 1) GO TO 82
      ICENT=1
      GO TO 16
21    ALB=DMIN1(ALMAX,-2.0D0*DF/GS)
      STM=0.0D0
      DO 23 I=1,N
      STM=DMAX1(STM,DABS(S(I)))
23    X(I)=X(I)+ALB*S(I)
      IFN=IFN+1
      CALL FUNCT(N,X,FB)
24    TEMP=FB-F-GS*ALB
      IF (TEMP .LE. 0.0D0) GO TO 30
      ALC=-GS*ALB**2/(2.0D0*TEMP)
      IF (DABS(ALC-ALB)/ALB .LE. 5.0D-2) GO TO 50
      ALC=DMIN1(ALC,4.0D0*ALB)
      IF (ALC*STM .GT. EPS) GO TO 34
```

```
25      DO 26 I=1,N
26      X(I)=X(I)-ALB*S(I)
        IF (ICENT .EQ. 1) GO TO 82
        ICENT=1
        GO TO 16
30      ALC=DMIN1(2.0D0*ALB,ALMAX)
34      IF (ALC .LT. ALMAX) GO TO 36
        IF (ALB .EQ. ALMAX) GO TO 50
        ALC=5.0D-1*(ALB+ALMAX)
36      ADF=ALC-ALB
        DO 38 I=1,N
38      X(I)=X(I)+ADF*S(I)
        IFN=IFN+1
        CALL FUNCT(N,X,FC)
        IF (IFN .EQ. MAXFN) GO TO 80
        IF (FC .GE. FB) GO TO 46
        IF (ALB .LT. ALC) GO TO 40
        ALMAX=ALB
40      FB=FC
        ALB=ALC
        GO TO 24
46      ADF=ALB-ALC
        DO 48 I=1,N
48      X(I)=X(I)+ADF*S(I)
        IF (FB .LT. F) GO TO 50
        GO TO 25
50      GGP=GG
        DF=F-FB
60      F=FB
        GO TO 10
80      KF=1
82      IF (IP .EQ. 0) GO TO 84
        PRINT 83
83      FORMAT(1H0,24X,'RESULTADO FINAL DE OPTIM')
        PRINT 500,ITER,IFN
        PRINT 501,F
        PRINT 502,(X(I),I=1,N)
        PRINT 503,(G(I),I=1,N)
84      RETURN
C
8       FORMAT (19X, 'RESULTADOS PARCIAIS')
500     FORMAT (8X,'No ITER.',I5,18X,'No AVAL. DA FUNCAO',I5)
501     FORMAT (15X,'VALOR DA FUNCAO',E15.7)
502     FORMAT (8X,'NO PONTO',3X,7E15.7/(14X,7E15.7))
503     FORMAT (9X,'COM GRAD.',3X,7E15.7/(14X,7E15.7))
C
        END
C
        SUBROUTINE VPROD(G,S,N,GS)
        IMPLICIT REAL*8(A-H,O-Z)
        DIMENSION G(*),S(*)
        GS=0.0D0
        DO 10 I=1,N
10      GS=GS+G(I)*S(I)
        RETURN
        END
```

```
################################################################################
```

```
C       ================================================================
C       Programa MESTRE para a estimacao  por  maxima  verossimilhanca
C       dos parametros do modelo de DOSE RESPOSTA e das  distribuicoes
C       de POISSON e BINOMIAL NEGATIVA para o tamanho das ninhadas.  A
C       estimacao sera feita atraves da subrotina OPTIM que utiliza  o
C       metodo de FLETCHER & REEVES.
C       ================================================================
C
C       ----------------------------------------------------------------
C       Este programa le a matriz de dados,  faz calculos auxiliares e
C       chama a subrotina de minimizacao da funcao. Deve ser usado com
C       a subrotina OPTIM para qualquer funcao deste trabalho.
C       ----------------------------------------------------------------
C
C       Declaracao das variaveis
C
        INTEGER N,ND,MAXCAL,I,J,K,C,L(5),OP
        REAL*8 DD,X(4),F,G(4),FACC,DELX,DIR(4),EPS,M(5,20),Y(20),D(5)
```

```
      REAL*8 S(5,20),NK(5),XMIN(4),XMAX(4),E(5,20,20),S1(5)
      REAL*8 S2(5,20),S3(5,20),X4
      CHARACTER NOME*20
C
C-----------------------------------------------------------------
C     As variaveis utilizadas neste programa e que nao foram  ainda
C     descritas sao:
C
C     ND: numero  de  grupos  experimentais  ou  numero  de  doses
C         aplicadas;
C
C     MAXCAL  = MAXFN: numero maximo permitido de  avaliacoes  da
C                      funcao. Veja OPTIM;
C
C     I,  J, K: indices que especificam o numero de linhas, colunas
C               e grupos da tabela de entrada de dados.  O  numero
C               maximo aqui utilizado foi: I, J = 20, K = 5;
C
C     C: total de diferentes valores assumidos por Y(J);
C
C     L(K): no k-esimo grupo, indica o total de linhas da tabela de
C           entrada de dados;
C
C     OP: variavel que especifica a funcao a ser minimizada;
C
C     DD: dose maxima aplicada;
C
C     DIR(N) = S(N): vetor direcao de dimensao n. Veja OPTIM;
C
C     M(K,I): total de femeas do k-esimo grupo que geraram ninhada
C             de tamanho S(K,I);
C
C     Y(J): numero de diferentes respostas adversas observadas entre
C           os filhotes da i-esima femea do k-esimo grupo;
C
C     D(K): dose aplicada ao k-esimo grupo;
C
C     S(K,I): tamanho da ninhada produzida pela i-esima femea do k-
C             esimo grupo;
C
C     NK(K): numero de femeas no k-esimo grupo;
C
C     E(K,I,J): numero de femeas do k-esimo grupo  que  produziram
C               ninhada  de tamanho S(K,I),  sendo que Y(J)  destes
C               filhotes apresentam a resposta adversa;
C
C     S1(K), S2(K,I), S3(K,I): somatorios auxiliares;
C
C     X4 = (X(4) - X(3))/DD
C-----------------------------------------------------------------
C
      COMMON/DADOS/ND,D,L,S,M,NK,DD
      COMMON/SOMA/S1,S2,S3
C
C     Escolha da matriz de dados a ser lida
C
      WRITE(*,*)'ARQUIVO DE DADOS A SER LIDO:'
      READ(*,'(A20)')NOME
      OPEN(1,STATUS='UNKNOWN',FILE=NOME)
C
C     Leitura dos dados
C
C     K = Indice da dose (grupo)
C     I = Indice da linha do k-esimo grupo
C     J = Indice da coluna
C
      READ(1,*)ND,C
      READ(1,*) (Y(J),J=1,C)
      DO 1 K=1,ND
        NK(K)=0.D0
        READ(1,*) D(K),L(K)
        DO 2 I=1,L(K)
          READ(1,*) S(K,I),(E(K,I,J),J=1,C),M(K,I)
          NK(K)=NK(K)+M(K,I)
    2   CONTINUE
    1 CONTINUE
      CLOSE(1)
C
      WRITE(*,*)'ARQUIVO DE SAIDA DOS RESULTADOS:'
```

```
      READ(*,'(A20)')NOME
      OPEN(1,STATUS='UNKNOWN',FILE=NOME)
C
10    WRITE(*,*)'DIGITE A OPCAO DO MODELO A SER AJUSTADO'
      WRITE(*,*)'1=DOSE RESP. 2=POISSON  3=BIN. NEG.:'
      READ(*,*) OP
C
      IF((OP.GT.3).OR.(OP.LT.1)) GO TO 10
      IF(OP.EQ.1) THEN
         N=4
         XMIN(1)=1.0D-6
         XMIN(2)=1.0D-6
         XMIN(3)=1.0D-6
         XMIN(4)=1.0D-6
         ELSE IF(OP.EQ.2) THEN
                 N=2
                 XMIN(1)=1.0D-6
                 XMIN(2)=-1.0D+6
                 ELSE
                      N=3
                      XMIN(1)=1.0D-6
                      XMIN(2)=-1.0D+6
                      XMIN(3)=1.0D-6
      ENDIF
C
CC    Calculo de somatorios auxiliares
C
      DO 3 K=1,ND
        S1(K)=0.D0
        DO 4 I=1,L(K)
          S1(K)=S1(K)+M(K,I)*S(K,I)
          S2(K,I)=0.D0
          S3(K,I)=0.D0
          DO 5 J=1,C
            S2(K,I)=S2(K,I)+E(K,I,J)*Y(J)
            DIF=S(K,I)-Y(J)
            S3(K,I)=S3(K,I)+E(K,I,J)*DIF
5         CONTINUE
4       CONTINUE
3     CONTINUE
CC
CC    Determinacao da dose maxima
C
      DD=D(1)
      DO 8 I=2,ND
        IF(D(I).GT.DD) DD=D(I)
8     CONTINUE
CCC
CC    Inicializacao das variaveis
C
      WRITE(*,*) 'PONTO INICIAL :'
      READ(*,*) (X(I),I=1,N)
C
      WRITE(*,*)'MOSTRAR AS ITERACOES MULTIPLAS DE :'
      READ(*,*)K
C
      DO 11 I=1,N
        XMAX(I)=1.0D+6
11    CONTINUE
C
      MAXCAL=32500
      EPS=1.D-10
      DELX=1.0D-3
      FACC=1.D-11
C
CCC   Chamada da subrotina que calcula o minimo da funcao
C
      CALL OPTIM(N,X,F,G,DIR,XMIN,XMAX,EPS,DELX,FACC,MAXCAL,K,
     *  ITER,IFN)
C
      WRITE(1,*)'RESULTADO FINAL PRODUZIDO POR OPTIM'
      WRITE(1,*)
      WRITE(1,*)'No DE ITERACOES:',ITER
      WRITE(1,*)
      WRITE(1,*)'No DE AVAL. DA FUNCAO:',IFN
      WRITE(1,*)
      WRITE(1,*)'VALOR DA FUNCAO:',F
      WRITE(1,*)
      WRITE(1,*)'NO PONTO:',(X(I),I=1,N)
```

```
      WRITE(1,*)
      WRITE(1,*)'COM GRADIENTE:',(G(I),I=1,N)
      WRITE(1,*)
      IF(N.EQ.4) THEN
         X4=(X(4)-X(3))/DD
         WRITE(1,*)'X(4);',X4
      ENDIF
      CLOSE(1)
      STOP
      END
```

##

```
C     ==========================================================
C     Subrotina para calcular a funcao -log-verossimilhanca do
C     modelo de  DOSE RESPOSTA. Esta  funcao  sera  minimizada
C     pelo metodo de FLETCHER & REEVES, utilizando a subrotina
C     OPTIM.
C     ==========================================================
C
      SUBROUTINE FUNCT(N,X,F)
C
C     ----------------------------------------------------------------
C     Esta subrotina calcula o oposto da funcao  log-verossimilhanca
C     para  o  modelo  de Rai & Van Ryzin.  Deve ser  usada  com  a
C     subrotinas OPTIM e MESTRE.
C     ----------------------------------------------------------------
C
C     Declaracao das variaveis
C
      INTEGER N,ND,L(5),I,K
      REAL*8 EXAB,EXT,A,B,P,Q,R,X(4),M(5,20),NK(5),DD,
     *  S(5,20),D(5),X34,F,S1(5),S2(5,20),S3(5,20)
C
C     ----------------------------------------------------------------
C     As variaveis utilizadas neste programa e que nao foram  ainda
C     descritas sao:
C
C     EXAB, EXT, A, B, R, X34: variaveis auxiliares;
C
C     P: probabilidade condicional de resposta adversa;
C
C     Q = 1 - P.
C     ----------------------------------------------------------------
C
      COMMON/DADOS/ND,D,L,S,M,NK,DD
      COMMON/SOMA/S1,S2,S3
C
      F=0.D0
C
C     I = indice da linha do k-esimo grupo
C     J = indice da coluna
C     K = indice do grupo (dose)
C
      DO 1 K=1,ND
        EXAB=-X(1)-X(2)*D(K)
        EXAB=DEXP(EXAB)
        X34=X(3)+D(K)*(X(4)-X(3))/DD
        R=1.D0-EXAB
C
        DO 2 I=1,L(K)
          EXT=-S(K,I)*X34
          A=DLOG(R)+EXT
          P=R*DEXP(EXT)
          Q=1.D0-P
          B=DLOG(Q)
C
C         Calculo da funcao
C
          F=F-(S2(K,I)*A)-(S3(K,I)*B)
C
2       CONTINUE
1     CONTINUE
      RETURN
      END
```

##

```
C     ================================================================
C     Subrotina para calcular a funcao -log-verossimilhanca  para  a
C     distribuicao de POISSON para o tamanho da ninhada. Esta funcao
C     sera  minimizada  pelo  metodo  de   FLETCHER & REEVES, com  a
C     subrotina OPTIM.
C     ================================================================

      SUBROUTINE FUNCT(N,X,F)

C     ----------------------------------------------------------------
C     Esta subrotina calcula o oposto da funcao  log-verossimilhanca
C     para o modelo Poisson.  Deve ser usada com as subrotinas OPTIM
C     e MESTRE.
C     ----------------------------------------------------------------

C     Declaracao de variaveis

      INTEGER K,L(5),ND,N
      REAL*8 F,A,B,X(2),S(5,20),D(5),NK(5),DD,
     *  M(5,20),X2D,S1(5),S2(5,20),S3(5,20)

C     ----------------------------------------------------------------
C     X2D, A, B: variaveis auxiliares.
C     ----------------------------------------------------------------

      COMMON/DADOS/ND,D,L,S,M,NK,DD
      COMMON/SOMA/S1,S2,S3

      F=0.D0

C     I = indice da linha do k-esimo grupo
C     K = indice do grupo (dose)

      DO 1 K=1,ND
        X2D=-X(2)*D(K)
        A=NK(K)*X(1)*DEXP(X2D)
        B=DLOG(X(1))+X2D

C         Calculo da funcao

          F=F+A-(B*S1(K))

    1   CONTINUE
      RETURN
      END

################################################################################

C     ================================================================
C     Subrotina para calcular a funcao -log-verossimilhanca  corres-
C     pondente a distribuicao BINOMIAL NEGATIVA para  o  tamanho  da
C     ninhada.  Esta  funcao  sera  minimizada  pelo  metodo  de
C     FLETCHER & REEVES, com a subrotina OPTIM.
C     ================================================================

      SUBROUTINE FUNCT(N,X,F)

C     ----------------------------------------------------------------
C     Esta subrotina calcula o oposto da funcao  log-verossimilhanca
C     para  o  modelo  binomial  negativo.  Deve  ser  usada  com  as
C     subrotinas OPTIM e MESTRE.
C     ----------------------------------------------------------------

C     Declaracao de variaveis

      INTEGER K,L(5),N,ND,R,IS
      DOUBLE PRECISION F,A,X(3),S(5,20),D(5),NK(5),DD,
     *  M(5,20),X2D,X1E,X13,L1,L2,S1(5),S2(5,20),S3(5,20),S4,PR

C     ----------------------------------------------------------------
C     A, X2D, X1E, X13, L1, L2, PR: variaveis auxiliares;

C     S4: somatorio auxiliar.
C     ----------------------------------------------------------------
```

```
C
      COMMON/DADOS/ND,D,L,S,M,NK,DD
      COMMON/SOMA/S1,S2,S3
C
      F=0.D0
C
CC    I = indice da linha do k-esimo grupo
CC    K = indice do grupo (dose)
C
      DO 1 K=1,ND
C
        X2D=-X(2)*D(K)
        X1E=X(1)*DEXP(X2D)
        X13=X1E+X(3)
        L1=DLOG(X1E/X13)
        L2=DLOG(X13/X(3))
        A=X(3)*NK(K)*L2
        S4=0.D0
C
        DO 2 I=1,L(K)
          PR=1.D0
C         IS=IDINT(S(K,I))
          DO 3 R=1,IS
            PR=PR*(X(3)+DBLE(R)-1.D0)
3         CONTINUE
CC
          S4=S4+M(K,I)*DLOG(PR)
2       CONTINUE
CC
        F=F-S4-(S1(K)*L1)+A
1     CONTINUE
      RETURN
      END
```

##

E2. METODO QUASE-NEWTON

```
C     ================================================================
CC    Programa MESTRE para estimacao por maxima verossimilhanca dos
CC    parametros do modelo de DOSE RESPOSTA e das distribuicoes de
CC    POISSON e BINOMIAL NEGATIVA para o tamanho da ninhada. A
CC    estimacao sera feita atraves da subrotina E04KBF da NAG, que
CC    utiliza o metodo QUASE-NEWTON com canalizacao. O vetor
CC    gradiente e calculado diretamente.
C     ================================================================

C     ----------------------------------------------------------------
      Este programa le a matriz de dados, faz calculos auxiliares e
      chama a subrotina de minimizacao da funcao. Deve ser usado com
      a subrotina E04KBF para qualquer funcao deste trabalho.
C     ----------------------------------------------------------------

      Declaracao de escalares e matrizes

      INTEGER IBOUND,IFAIL,INTYPE,IPRINT,J,LH,LIW,LW,MAXCAL,N
      INTEGER ISTATE(4),IW(2)
      REAL ETA,F,FEST,STEPMX,XTOL,BL(4),BU(4)
      REAL G(4),HESD(4),HESL(6),W(36),X(4),X4
      LOGICAL LOCSCH
      EXTERNAL E04JBQ,FUNCT,MONIT
C     ----------------------------------------------------------------
      O usuario interessado deve consultar a documentacao da NAG
      para a subrotina E04KBF, para esclarecer o significado das
      variaveis acima.
C     ----------------------------------------------------------------

      INTEGER I,K,C,L(5),OP
      REAL DD,M(5,20),Y(20),S(5,20),D(5),NK(5),E(5,20,20),S1(5)
      REAL S2(5,20),S3(5,20)
      CHARACTER NOME*20
C
      COMMON/DADOS/ND,D,L,S,M,NK,DD
      COMMON/SOMA/S1,S2,S3
C
```

```
      WRITE(*,*)'ARQUIVO DE DADOS A SER LIDO:'
      READ(*,'(A20)')NOME
      OPEN(1,STATUS='UNKNOWN',FILE=NOME)
C
CCC   Leitura dos dados
CCCCC
CCCCC K = Indice da dose (grupo)
CCCCC I = Indice da linha do k-esimo grupo
CCCCC J = Indice da coluna
C
      READ(1,*)ND,C
      READ(1,*) (Y(J),J=1,C)
      DO 1 K=1,ND
        NK(K)=0.0
        READ(1,*) D(K),L(K)
        DO 2 I=1,L(K)
          READ(1,*) S(K,I),(E(K,I,J),J=1,C),M(K,I)
          NK(K)=NK(K)+M(K,I)
2       CONTINUE
1     CONTINUE
      CLOSE(1)
C
      WRITE(*,*)'ARQUIVO DE SAIDA DOS RESULTADOS:'
      READ(*,'(A20)')NOME
      OPEN(1,STATUS='UNKNOWN',FILE=NOME)
C
10    WRITE(*,*)'DIGITE A OPCAO DO MODELO A SER AJUSTADO:'
      WRITE(*,*)'1=DOSE-RESP.  2=POISSON 3=BIN. NEG.:'
      READ(*,*) OP
C
      IF((OP.GT.3).OR.(OP.LT.1)) GO TO 10
      IF(OP.EQ.1) THEN
         N=4
         LH=6
         LW=36
         BL(1)=1.0E-6
         BL(2)=1.0E-6
         BL(3)=1.0E-6
         BL(4)=1.0E-6
         ELSE IF(OP.EQ.2) THEN
                  N=2
                  LH=1
                  LW=18
                  BL(1)=1.0E-6
                  BL(2)=-1.0E+6
                  ELSE
                      N=3
                      LH=3
                      LW=27
                      BL(1)=1.0E-6
                      BL(2)=-1.0E+6
                      BL(3)=1.0E-6
      ENDIF
CCCC
CCCC  Calculo de somatorios auxiliares
C
      DO 3 K=1,ND
        S1(K)=0.0
        DO 4 I=1,L(K)
          S1(K)=S1(K)+M(K,I)*S(K,I)
          S2(K,I)=0.0
          S3(K,I)=0.0
          DO 5 J=1,C
            S2(K,I)=S2(K,I)+E(K,I,J)*Y(J)
            DIF=S(K,I)-Y(J)
            S3(K,I)=S3(K,I)+E(K,I,J)*DIF
5         CONTINUE
4       CONTINUE
3     CONTINUE
CCCC
CCCC  Determinacao da dose maxima
C
      DD=D(1)
      DO 6 I=2,ND
        IF(D(I).GT.DD) DD=D(I)
6     CONTINUE
CCCC
CCCC  Inicializacao das variaveis
CCC
```

```
      WRITE(*,*)'PONTO INICIAL:'
      READ(*,*) (X(I),I=1,N)
C
      LIW=2
      IPRINT=1
      LOCSCH=.TRUE.
      INTYPE=0
      MAXCAL=200*N
      ETA=0.5
      XTOL=0.0
      STEPMX=100000.0
      FEST=0.0
      IFAIL=1
      IBOUND=0
C
      DO 11 I=1,N
        BU(I)=1.0E+6
11    CONTINUE
CC
CCC   Chamada de E04KBF para calcular o minimo da funcao
CC
      CALL E04KBF(N,FUNCT,MONIT,IPRINT,LOCSCH,INTYPE,E04JBQ,
     *  MAXCAL,ETA,XTOL,STEPMX,FEST,IBOUND,BL,BU,X,HESL,
     *  LH,HESD,ISTATE,F,G,IW,LIW,W,LW,IFAIL)
C
CCC   Teste se IFAIL e diferente de zero na saida
CC
      IF (IFAIL.NE.0) WRITE(*,*) 'ERRO TIPO:',IFAIL
      IF (IFAIL.EQ.1) GO TO 60
C
      WRITE(1,*) 'VALOR DA FUNCAO NA SAIDA E:',F
      WRITE(1,*) 'NO PONTO:',(X(J),J=1,N)
      WRITE(1,*) 'O GRADIENTE CORRESPONDENTE E:',(G(J),J=1,N)
      IF(N.EQ.4) THEN
        X4=(X(4)-X(3))/DD
        WRITE(1,*)'X(4):',X4
      ENDIF
C
      CLOSE(1)
60    STOP
      END


######################################################################################


CC    ============================================================
CCC   Subrotina para calcular  a  funcao  -log-verossimilhanca  e  o
CCC   vetor gradiente para o modelo de  DOSE RESPOSTA.  Esta  funcao
CCC   sera minimizada pela subrotina E04KBF, que  utiliza  o  metodo
CCC   QUASE-NEWTON.
CCC   ============================================================
CC
      SUBROUTINE FUNCT(IFLAG,N,X,F,G,IW,LIW,W,LW)

CC
CCC   Declaracao das variaveis
CC
      INTEGER N,ND,L(5),I,K,IFLAG,LIW,LW,IW(LIW)
      REAL EXAB,EXT,A,B,P,Q,R,X(4),M(5,20),NK(5),DD,W(LW),D1,D2,T,
     *  S(5,20),D(5),X34,F,G(4),S1(5),S2(5,20),S3(5,20),A1,A2,A1E
C
      COMMON/DADOS/ND,D,L,S,M,NK,DD
      COMMON/SOMA/S1,S2,S3
C
      F=0.0
      DO 1 J=1,N
        G(J)=0.0
1     CONTINUE
CCC
CCCC  I = indice da linha do k-esimo grupo
CCCC  J = indice da coluna
CCCC  K = indice do grupo (dose)
CC
      DO 2 K=1,ND
        EXAB=-X(1)-X(2)*D(K)
        EXAB=EXP(EXAB)
        D1=D(K)/DD
        D2=1-D1
        X34=X(3)+D1*(X(4)-X(3))
```

```
            R=1.0-EXAB
            A1=0.0
            A2=0.0
C
            DO 3 I=1,L(K)
              EXT=-S(K,I)*X34
              A=ALOG(R)+EXT
              EXT=EXP(EXT)
              P=R*EXT
              Q=1.0-P
              B=ALOG(Q)
              T=(S2(K,I)/P)-(S3(K,I)/Q)
              A1=A1+T*EXT
              A2=A2+P*S(K,I)*T
C
C             Calculo da funcao
C
              F=F-(S2(K,I)*A)-(S3(K,I)*B)
C
3             CONTINUE
C
C           Calculo do vetor gradiente
C
            A1E=A1*EXAB
            G(1)=G(1)-A1E
            G(2)=G(2)-A1E*D(K)
            G(3)=G(3)+D2*A2
            G(4)=G(4)+D1*A2
C
2         CONTINUE
          RETURN
          END

C         Aqui entra a subrotina MONIT de monitoramento  do  procedimento
C         de minimizacao

##########################################################################################


C         ================================================================
C         Subrotina para o calculo da funcao -log-verossimilhanca  e  do
C         vetor gradiente do modelo de POISSON para o tamanho da ninhada.
C         Esta funcao sera minimizada pela subrotina E04KBF, que utiliza
C         o metodo QUASE-NEWTON.
C         ================================================================

          SUBROUTINE FUNCT(IFLAG,N,X,F,G,IW,LIW,W,LW)

C         Declaracao das variaveis

          INTEGER N,ND,L(5),I,K,IFLAG,LIW,LW,IW(LIW)
          REAL A,AS,B,X(2),M(5,20),NK(5),DD,W(LW),X2D,
         *S(5,20),D(5),F,G(2),S1(5),S2(5,20),S3(5,20)
C
          COMMON/DADOS/ND,D,L,S,M,NK,DD
          COMMON/SOMA/S1,S2,S3
C
          F=0.0
          DO 1 J=1,N
            G(J)=0.0
1         CONTINUE
C
C         I = indice da linha do k-esimo grupo
C         K = indice do grupo (dose)
C
          DO 2 K=1,ND
            X2D=-X(2)*D(K)
            A=NK(K)*X(1)*EXP(X2D)
            AS=A-S1(K)
            B=ALOG(X(1))+X2D
C
C             Calculo da funcao
C
              F=F+A-(B*S1(K))
C
C             Calculo do vetor gradiente
C
              G(1)=G(1)+AS/X(1)
```

```
            G(2)=G(2)-D(K)*AS

2         CONTINUE
        RETURN
        END
C
CCC     Aqui entra a subrotina MONIT de monitoramento  do  procedimento
C       de minimizacao

##########################################################################


C       ================================================================
CCC     Subrotina para calculo da  funcao  -log-verossimilhanca  e  do
CCCC    vetor gradiente para a distribuicao BINOMIAL NEGATIVA  para  o
CCCCC   tamanho da ninhada. Esta funcao sera minimizada pela subrotina
CCCCCC  E04KBF, que utiliza o metodo QUASE-NEWTON.
        ================================================================
C
        SUBROUTINE FUNCT(IFLAG,N,X,F,G,IW,LIW,W,LW)
C
CCC     Declaracao das variaveis
CC
        INTEGER N,ND,L(5),I,K,IFLAG,LIW,LW,IW(LIW),R,IS
        REAL X(3),M(5,20),NK(5),DD,W(LW),S(5,20),D(5),F,G(3),S1(5),
     *  S2(5,20),S3(5,20),A,AX,X1E,X13,L1,L2,PR,E,XR,S4,S5,S6
C
        COMMON/DADOS/ND,D,L,S,M,NK,DD
        COMMON/SOMA/S1,S2,S3
C
        F=0.0
        DO 1 J=1,N
          G(J)=0.0
1       CONTINUE
C
CCC     I = indice da linha do k-esimo grupo
CCC     K = indice do grupo (dose)
CC
        DO 2 K=1,ND
          E=EXP(-X(2)*D(K))
          X1E=X(1)*E
          X13=X1E+X(3)
          L1=ALOG(X1E/X13)
          L2=ALOG(X13/X(3))
          A=X(3)*NK(K)*L2
          AX=(-S1(K)+NK(K)*X1E)*X(3)/X13
          S5=0.0
          S6=0.0
C
          DO 3 I=1,L(K)
            PR=1.0
            S4=0.0
            IS=IFIX(S(K,I))
            DO 4 R=1,IS
              XR=X(3)+FLOAT(R)-1.0
              PR=PR*XR
              S4=S4+1/XR
4           CONTINUE
            S5=S5+M(K,I)*ALOG(PR)
            S6=S6+M(K,I)*S4
3         CONTINUE
C
CCC       Calculo da funcao
C
          F=F-S5-(S1(K)*L1)+A
C
CCC       Calculo do vetor gradiente
CC
          G(1)=G(1)+AX/X(1)
          G(2)=G(2)-D(K)*AX
          G(3)=G(3)-S6+S1(K)/X13+NK(K)*(L2-X1E/X13)
2         CONTINUE
        RETURN
        END
C
CCC     Aqui entra a subrotina MONIT de monitoramento  do  procedimento
C       de minimizacao
```

```
##########################################################################

C     ================================================================
C     Programa MESTRE para minimizar a  funcao  -log-verossimilhanca
C     da distribuicao BINOMIAL NEGATIVA para o tamanho  da  ninhada,
C     atraves da subrotina E04JBF  da  NAG,  que  utiliza  o  metodo
C     QUASE-NEWTON com canalizacao. O vetor  gradiente e  aproximado
C     por diferencas finitas.
C     ================================================================

C     ----------------------------------------------------------------
C     Este programa le a matriz de dados,  faz calculos auxiliares e
C     chama a subrotina de minimizacao da funcao. Deve ser usado com
C     a subrotina E04JBF para qualquer funcao deste trabalho.
C     ----------------------------------------------------------------

C     Declaracao de escalares e matrizes

      INTEGER IBOUND,IFAIL,INTYPE,IPRINT,J,LH,LIW,LW,MAXCAL,N
      INTEGER ISTATE(2),IW(2)
      REAL ETA,F,FEST,STEPMX,XTOL,BL(3),BU(3),DELTA(3)
      REAL G(3),HESD(3),HESL(3),W(27),X(3),ts
      LOGICAL LOCSCH
      EXTERNAL E04JBQ,FUNCT,MONIT

C     ----------------------------------------------------------------
C     O  usuario interessado deve consultar a documentacao  da  NAG
C     para  a subrotina E04JBF,  para esclarecer o significado  das
C     variaveis acima.
C     ----------------------------------------------------------------

      INTEGER I,K,C,L(5)
      REAL DD,M(5,20),Y(20),S(5,20),D(5),NK(5),E(5,20,20),S1(5)
      REAL S2(5,20),S3(5,20)
      CHARACTER NOME*20

C
      COMMON/DADOS/ND,D,L,S,M,NK,DD
      COMMON/SOMA/S1,S2,S3

C     Escolha da matriz de dados a ser lida

      WRITE(*,*)'ARQUIVO:'
      READ(*,'(A20)')NOME
      OPEN(1,STATUS='UNKNOWN',FILE=NOME)

C     Leitura dos dados

C     K = Indice da dose (grupo)
C     I = Indice da linha do k-esimo grupo
C     J = Indice da coluna

      READ(1,*)ND,C
      READ(1,*) (Y(J),J=1,C)
      DO 1 K=1,ND
        NK(K)=0.0
        READ(1,*) D(K),L(K)
        DO 2 I=1,L(K)
          READ(1,*) S(K,I),(E(K,I,J),J=1,C),M(K,I)
          NK(K)=NK(K)+M(K,I)
2       CONTINUE
1     CONTINUE
      CLOSE(1)

C     Calculo de somatorios auxiliares

      DO 3 K=1,ND
        S1(K)=0.0
        DO 4 I=1,L(K)
          S1(K)=S1(K)+M(K,I)*S(K,I)
          S2(K,I)=0.0
          S3(K,I)=0.0
          DO 5 J=1,C
            S2(K,I)=S2(K,I)+E(K,I,J)*Y(J)
            DIF=S(K,I)-Y(J)
            S3(K,I)=S3(K,I)+E(K,I,J)*DIF
5         CONTINUE
4       CONTINUE
```

```
3       CONTINUE
C
C       Determinacao da dose maxima
C
        DD=D(1)
        DO 8 I=2,ND
          IF(D(I).GT.DD) DD=D(I)
8       CONTINUE
C
C       Inicializacao das variaveis
C
        N=3
        LH=3
        LIW=2
        LW=27
        IPRINT=1
        LOCSCH=.TRUE.
        INTYPE=0
        MAXCAL=40*N*(N+5)
        ETA=0.5
        XTOL=0.0
        STEPMX=100000.0
        FEST=0.0
        IBOUND=0
        BL(1)=1.0E-6
        BU(1)=1.0E+6
        BL(2)=-1.0E+6
        BU(2)=BU(1)
        BL(3)=BL(1)
        BU(3)=BU(1)
        IFAIL=1
C
C       Leitura do ponto inicial e de delta
C
        WRITE(*,*)'PONTO INICIAL:'
        READ(*,*) (X(I),I=1,N)
C
        WRITE(*,*)'DELTA:'
        READ(*,*) DELTAX
        DO 7 I=1,N
          DELTA(I)=DELTAX
7       CONTINUE
C
C       Chamada de E04JBF para calcular o minimo da funcao
C
        CALL E04JBF(N,FUNCT,MONIT,IPRINT,LOCSCH,INTYPE,E04JBQ,
     *  MAXCAL,ETA,XTOL,STEPMX,FEST,DELTA,IBOUND,BL,BU,X,HESL,
     *  LH,HESD,ISTATE,F,G,IW,LIW,W,LW,IFAIL)
C
C       Teste se IFAIL e diferente de zero na saida
C
        IF (IFAIL.NE.0) WRITE(*,*) 'ERRO TIPO:',IFAIL
        IF (IFAIL.EQ.1) GO TO 60
C
        WRITE(*,*) 'VALOR DA FUNCAO NA SAIDA E:',F
        WRITE(*,*) 'NO PONTO:',(X(J),J=1,N)
C
60      STOP
        END

################################################################################


C       ================================================================
C       Subrotina para calculo da funcao -log-verossimilhanca  para  a
C       distribuicao BINOMIAL NEGATIVA para o tamanho da  ninhada  que
C       sera minimizada com a subrotina E04JBF, que utiliza  o  metodo
C       QUASE-NEWTON.
C       ================================================================
C
        SUBROUTINE FUNCT(IFLAG,N,X,F,G,IW,LIW,W,LW)
C
C       Declaracao das variaveis
C
        INTEGER N,ND,L(5),I,K,R,IS,IFLAG,LIW,LW,IW(LIW)
        REAL F,A,X(3),S(5,20),D(5),NK(5),DD,M(5,20),X2D,X1E,X13,L1
        REAL L2,S1(5),S2(5,20),S3(5,20),S4,PR
C
```

```
      COMMON/DADOS/ND,D,L,S,M,NK,DD
      COMMON/SOMA/S1,S2,S3
C
      F=0.0
C
C     I = indice da linha do k-esimo grupo
C     K = indice do grupo (dose)
C
      DO 1 K=1,ND
C
        X2D=-X(2)*D(K)
        X1E=X(1)*EXP(X2D)
        X13=X1E+X(3)
        L1=ALOG(X1E/X13)
        L2=ALOG(X13/X(3))
        A=X(3)*NK(K)*L2
        S4=0.0
C
        DO 2 I=1,L(K)
          PR=1.0
          IS=IFIX(S(K,I))
          DO 3 R=1,IS
            PR=PR*(X(3)+FLOAT(R)-1.0)
3         CONTINUE
C
          S4=S4+M(K,I)*ALOG(PR)
2         CONTINUE
C
          F=F-S4-(S1(K)*L1)+A
1       CONTINUE
      RETURN
      END
C
C     Aqui entra a subrotina MONIT de monitoramento  do  procedimento
C     de minimizacao
C
```

##

E3. METODO DE NEWTON

```
C     ================================================================
C     Programa MESTRE para a estimacao por maxima verossilhanca dos
C     parametros do modelo de DOSE RESPOSTA  e das distribuicoes de
C     POISSON  e  BINOMIAL NEGATIVA para o tamanho das ninhadas.  A
C     estimacao sera feita atraves da subrotina E04LBF  da NAG, que
C     utiliza o metodo de NEWTON com canalizacao.
C     ================================================================
C
C     ----------------------------------------------------------------
C     Este programa le a matriz de dados,  faz calculos auxiliares e
C     chama a subrotina de minimizacao da funcao. Deve ser usado com
C     a subrotina E04LBF para qualquer funcao deste trabalho.
C     ----------------------------------------------------------------
C
      Declaracao de escalares e matrizes

      INTEGER IBOUND,IFAIL,IPRINT,J,LH,LIW,LW,MAXCAL,N
      INTEGER ISTATE(4),IW(2)
      REAL ETA,F,STEPMX,XTOL,BL(4),BU(4),G(4),HESD(4),HESL(6),W(34)
      REAL X(4),X4
      EXTERNAL FUNCT,HESS,MONIT
C
C     ----------------------------------------------------------------
C     O  usuario interessado deve consultar a documentacao  da  NAG
C     para  a subrotina E04LBF,  para esclarecer o significado  das
C     variaveis acima.
C     ----------------------------------------------------------------
C
      INTEGER I,K,C,L(5),OP
      REAL DD,M(5,20),Y(20),S(5,20),D(5),NK(5),E(5,20,20),S1(5)
      REAL S2(5,20),S3(5,20)
      CHARACTER NOME*20
C
      COMMON/DADOS/ND,D,L,S,M,NK,DD
```

```
      COMMON/SOMA/S1,S2,S3
C
CC    Escolha da matriz de dados a ser lida
C
      WRITE(*,*)'ARQUIVO DE DADOS A SER LIDO:'
      READ(*,'(A20)')NOME
      OPEN(1,STATUS='UNKNOWN',FILE=NOME)
C
CC    Leitura dos dados
CC
CC    K = Indice da dose (grupo)
CC    I = Indice da linha do k-esimo grupo
CC    J = Indice da coluna
C
      READ(1,*)ND,C
      READ(1,*) (Y(J),J=1,C)
      DO 1 K=1,ND
        NK(K)=0.0
        READ(1,*) D(K),L(K)
        DO 2 I=1,L(K)
          READ(1,*) S(K,I),(E(K,I,J),J=1,C),M(K,I)
          NK(K)=NK(K)+M(K,I)
2       CONTINUE
1     CONTINUE
C
      CLOSE(1)
C
      WRITE(*,*)'ARQUIVO DE SAIDA DOS RESULTADOS:'
      READ(*,'(A20)')NOME
      OPEN(1,STATUS='UNKNOWN',FILE=NOME)
C
10    WRITE(*,*)'DIGITE A OPCAO DO MODELO A SER AJUSTADO:'
      WRITE(*,*)'1=DOSE RESP.  2=POISSON  3=BIN. NEG.:'
      READ(*,*) OP
C
      IF((OP.GT.3).OR.(OP.LT.1)) GO TO 10
      IF(OP.EQ.1) THEN
         N=4
         LH=6
         LW=34
         BL(1)=1.0E-6
         BL(2)=1.0E-6
         BL(3)=1.0E-6
         BL(4)=1.0E-6
         ELSE IF(OP.EQ.2) THEN
                 N=2
                 LH=1
                 LW=15
                 BL(1)=1.0E-6
                 BL(2)=-1.0E+6
                 ELSE
                      N=3
                      LH=3
                      LW=24
                      BL(1)=1.0E-6
                      BL(2)=-1.0E+6
                      BL(3)=1.0E-6
      ENDIF
C
CC    Calculo de somatorios auxiliares
C
      DO 3 K=1,ND
        S1(K)=0.0
        DO 4 I=1,L(K)
          S1(K)=S1(K)+M(K,I)*S(K,I)
          S2(K,I)=0.0
          S3(K,I)=0.0
          DO 5 J=1,C
            S2(K,I)=S2(K,I)+E(K,I,J)*Y(J)
            DIF=S(K,I)-Y(J)
            S3(K,I)=S3(K,I)+E(K,I,J)*DIF
5         CONTINUE
4       CONTINUE
3     CONTINUE
CC
CC    Determinacao da dose maxima
C
      DD=D(1)
      DO 6 I=2,ND
```

```
          IF(D(I).GT.DD) DD=D(I)
6       CONTINUE
C
C       Inicializacao das variaveis
C
        IFAIL=1
        LIW=2
        IPRINT=1
        MAXCAL=150*N
        ETA=0.9
        XTOL=0.0
        STEPMX=100000.0
        IBOUND=0
C
        WRITE(*,*) 'PONTO INICIAL :'
        READ(*,*) (X(I),I=1,N)
C
        DO 11 I=1,N
          BU(I)=1.0E+6
11      CONTINUE
C
C       Chamada da subrotina E04LBF
C
        CALL E04LBF(N,FUNCT,HESS,MONIT,IPRINT,MAXCAL,ETA,XTOL,STEPMX,
     *  IBOUND,BL,BU,X,HESL,LH,HESD,ISTATE,F,G,IW,LIW,W,LW,IFAIL)
C
C       Teste se IFAIL e diferente de zero na saida de E04LBF
C
        IF (IFAIL.NE.0) WRITE(*,*) 'ERRO TIPO:',IFAIL
        IF (IFAIL.EQ.1) GO TO 20
C
        WRITE(1,*) 'VALOR DA FUNCAO NA SAIDA E:',F
        WRITE(1,*) 'NO PONTO:',(X(J),J=1,N)
        WRITE(1,*) 'O GRADIENTE CORRESPONDENTE E:',(G(J),J=1,N)
        WRITE(1,*) 'ISTATE:',(ISTATE(J),J=1,N)
        IF(N.EQ.4) THEN
          X4=(X(4)-X(3))/DD
          WRITE(1,*)'X(4):',X4
        ENDIF
        CLOSE(1)
20      STOP
        END

################################################################################


C       ================================================================
C       Subrotina para calcular a funcao -log-verossimilhanca, o vetor
C       gradiente e a matriz hessiana para o modelo de DOSE  RESPOSTA.
C       Esta funcao sera  minimizada  com  a  subrotina  E04LBF,  pelo
C       metodo de NEWTON.
C       ================================================================
C
        SUBROUTINE FUNCT(IFLAG,N,X,F,G,IW,LIW,W,LW)
C
C       Declaracao das variaveis
C
        INTEGER N,ND,L(5),I,K,IFLAG,LIW,LW,IW(LIW)
        REAL EXAB,EXT,A,B,P,Q,R,X(4),M(5,20),NK(5),DD,W(LW),D1,D2,T,
     *  S(5,20),D(5),X34,F,G(4),S1(5),S2(5,20),S3(5,20),A1,A2,A1E
C
        COMMON/DADOS/ND,D,L,S,M,NK,DD
        COMMON/SOMA/S1,S2,S3
C
        F=0.0
        DO 1 J=1,N
          G(J)=0.0
1       CONTINUE
C
C       I = indice da linha do k-esimo grupo
C       J = indice da coluna
C       K = indice do grupo (dose)
C
        DO 2 K=1,ND
          EXAB=-X(1)-X(2)*D(K)
          EXAB=EXP(EXAB)
          D1=D(K)/DD
          D2=1.0-D1
```

```
      X34=X(3)+D1*(X(4)-X(3))
      R=1.0-EXAB
      A1=0.0
      A2=0.0
C
      DO 3 I=1,L(K)
        EXT=-S(K,I)*X34
        A=ALOG(R)+EXT
        EXT=EXP(EXT)
        P=R*EXT
        Q=1.0-P
        B=ALOG(Q)
        T=(S2(K,I)/P)-(S3(K,I)/Q)
        A1=A1+T*EXT
        A2=A2+P*S(K,I)*T
C
CC      Calculo da funcao
C
        F=F-(S2(K,I)*A)-(S3(K,I)*B)
C
3     CONTINUE
CCC
CC    Calculo do vetor gradiente
C
      A1E=A1*EXAB
      G(1)=G(1)-A1E
      G(2)=G(2)-A1E*D(K)
      G(3)=G(3)+D2*A2
      G(4)=G(4)+D1*A2
C
2     CONTINUE
      RETURN
      END
C
CC    Subrotina que calcula a matriz hessiana
C
      SUBROUTINE HESS(IFLAG,N,X,FHESL,LH,FHESD,IW,LIW,W,LW)
C
CC    Declaracao das variaveis
C
      INTEGER N,ND,L(5),I,K,IFLAG,LIW,LW,IW(LIW)
      REAL EXAB,EXT,P,P2,Q,Q2,R,X(4),M(5,20),NK(5),DD,W(LW),D1,D2,
     *  T,S(5,20),D(5),X34,S1(5),S2(5,20),S3(5,20),A1,A2,A3,E1,E2,
     *  FHESD(4),FHESL(6),U,TPU,TEU
C
      COMMON/DADOS/ND,D,L,S,M,NK,DD
      COMMON/SOMA/S1,S2,S3
C
      DO 1 J=1,LH
        FHESL(J)=0.0
1     CONTINUE
C
      DO 2 J=1,N
        FHESD(J)=0.0
2     CONTINUE
CCC
CC    I = indice da linha do k-esimo grupo
CC    K = indice do grupo (dose)
C
      DO 3 K=1,ND
        EXAB=-X(1)-X(2)*D(K)
        EXAB=EXP(EXAB)
        D1=D(K)/DD
        D2=1.0-D1
        X34=X(3)+D1*(X(4)-X(3))
        R=1.0-EXAB
        A1=0.0
        A2=0.0
        A3=0.0
C
        DO 4 I=1,L(K)
          EXT=-S(K,I)*X34
          EXT=EXP(EXT)
          P=R*EXT
          P2=P**2
          Q=1.0-P
          Q2=Q**2
          T=S2(K,I)/P-S3(K,I)/Q
          U=S2(K,I)/P2+S3(K,I)/Q2
```

```
            TPU=T-P*U
            TEU=T+EXAB*EXT*U
            A1=A1+TEU*EXT
            A2=A2+EXT*S(K,I)*T/U
            A3=A3+P*(S(K,I)**2)*TPU
4         CONTINUE
C
          E1=EXAB*A1
          E2=EXAB*A2

C         Calculo da matriz Hessiana
CC
          FHESL(1)=FHESL(1)+D(K)*E1
          FHESL(2)=FHESL(2)+E2*D2
          FHESL(3)=FHESL(3)+D(K)*E2*D2
          FHESL(4)=FHESL(4)+E2*D1
          FHESL(5)=FHESL(5)+D(K)*E2*D1
          FHESL(6)=FHESL(6)-D1*D2*A3
          FHESD(1)=FHESD(1)+E1
          FHESD(2)=FHESD(2)+(D(K)**2)*E1
          FHESD(3)=FHESD(3)-(D2**2)*A3
          FHESD(4)=FHESD(4)-(D1**2)*A3
3         CONTINUE
        RETURN
        END
C
CC      Aqui entra a subrotina MONIT de monitoramento  do  procedimento
C       de minimizacao

##########################################################################################

C       ============================================================
CC      Subrotina para calcular a funcao -log-verossimilhanca, o vetor
CC      gradiente e a matriz hessiana para a  distribuicao de  POISSON
CCC     para o tamanho da ninhada.  Esta funcao sera  minimizada  pelo
CCC     metodo de NEWTON, atraves da subrotina E04LBF.
CCCC    ============================================================
CC
        SUBROUTINE FUNCT(IFLAG,N,X,F,G,IW,LIW,W,LW)
C
CC      Declaracao das variaveis

        INTEGER N,ND,L(5),I,K,IFLAG,LIW,LW,IW(LIW)
        REAL A,AS,B,X(2),M(5,20),NK(5),DD,W(LW),X2D,
     *  S(5,20),D(5),F,G(2),S1(5),S2(5,20),S3(5,20)
C
        COMMON/DADOS/ND,D,L,S,M,NK,DD
        COMMON/SOMA/S1,S2,S3
C
        F=0.0
        DO 1 J=1,N
          G(J)=0.0
1       CONTINUE
CC
CCC     I = indice da linha do k-esimo grupo
CCC     K = indice do grupo (dose)
CC
        DO 2 K=1,ND
          X2D=-X(2)*D(K)
          A=NK(K)*X(1)*EXP(X2D)
          AS=A-S1(K)
          B=ALOG(X(1))+X2D
C
CCC         Calculo da funcao
CC
            F=F+A-(B*S1(K))
C
CCC         Calculo vetor gradiente
CC
            G(1)=G(1)+AS/X(1)
            G(2)=G(2)-D(K)*AS
C
2         CONTINUE
        RETURN
        END
C
CC      Subrotina que calcula a matriz hessiana
```

```
C
      SUBROUTINE HESS(IFLAG,N,X,FHESL,LH,FHESD,IW,LIW,W,LW)
C
C     Declaracao das variaveis
C
      INTEGER N,ND,L(5),I,K,IFLAG,LH,LIW,LW,IW(LIW)
      REAL X(2),M(5,20),NK(5),DD,W(LW),S(5,20),D(5),S1(5),
     *  S2(5,20),S3(5,20),E,NDE,FHESD(2),FHESL(1)
C
      COMMON/DADOS/ND,D,L,S,M,NK,DD
      COMMON/SOMA/S1,S2,S3
C
      FHESL(1)=0.0
      FHESD(1)=0.0
      FHESD(2)=0.0
C
C     I = indice da linha do k-esimo grupo
C     K = indice do grupo (dose)
C
      DO 1 K=1,ND
        E=EXP(-X(2)*D(K))
        NDE=NK(K)*D(K)*E
C
C       Calculo da matriz Hessiana
C
        FHESL(1)=FHESL(1)-NDE
        FHESD(1)=FHESD(1)+S1(K)/(X(1)**2)
        FHESD(2)=FHESD(2)+NDE*D(K)*X(1)
1       CONTINUE
      RETURN
      END
C
C     Aqui entra a subrotina MONIT de monitoramento  do  procedimento
C     de minimizacao

##########################################################################################

C     ================================================================
C     Subrotina para calcular a funcao -log-verossimilhanca, o vetor
C     gradiente e a matriz  hessiana  para  a distribuicao  BINOMIAL
C     NEGATIVA para o tamanho da ninhada.Esta funcao sera minimizada
C     pelo metodo de NEWTON, atraves da subrotina E04LBF.
C     ================================================================
C
      SUBROUTINE FUNCT(IFLAG,N,X,F,G,IW,LIW,W,LW)
C
C     Declaracao das variaveis
C
      INTEGER N,ND,L(5),I,K,IFLAG,LIW,LW,IW(LIW),R
      REAL X(3),M(5,20),NK(5),DD,W(LW),S(5,20),D(5),S1(5),F,G(3),
     *  S2(5,20),S3(5,20),A,AX,X1E,X13,L1,L2,PR,E,XR,S4,S5,S6,S7,S8
C
      COMMON/DADOS/ND,D,L,S,M,NK,DD
      COMMON/SOMA/S1,S2,S3
      COMMON/S8/S8
C
      F=0.0
      DO 1 J=1,N
        G(J)=0.0
1     CONTINUE
C
C     I = indice da linha do k-esimo grupo
C     K = indice do grupo (dose)
C
      DO 2 K=1,ND
        E=EXP(-X(2)*D(K))
        X1E=X(1)*E
        X13=X1E+X(3)
        L1=ALOG(X1E/X13)
        L2=ALOG(X13/X(3))
        A=X(3)*NK(K)*L2
        AX=(-S1(K)+NK(K)*X1E)*X(3)/X13
        S5=0.0
        S6=0.0
        S8=0.0
C
        DO 3 I=1,L(K)
```

```
            PR=1.0
            S4=0.0
            IS=IFIX(S(K,I))
            DO 4 R=1,IS
              XR=X(3)+FLOAT(R)-1.0
              PR=PR*XR
              S4=S4+1/XR
              S7=S7+(1/XR)**2
4           CONTINUE
          S5=S5+M(K,I)*ALOG(PR)
          S6=S6+M(K,I)*S4
          S8=S8+M(K,I)*S7
3         CONTINUE
C
C         Calculo da funcao
C
          F=F-S5-(S1(K)*L1)+A
C
C         Calculo do vetor gradiente
C
          G(1)=G(1)+AX/X(1)
          G(2)=G(2)-D(K)*AX
          G(3)=G(3)-S6+S1(K)/X13+NK(K)*(L2-X1E/X13)
2         CONTINUE
        RETURN
        END
C
C       Subrotina que calcula a matriz hessiana
C
        SUBROUTINE HESS(IFLAG,N,X,FHESL,LH,FHESD,IW,LIW,W,LW)
C
C       Declaracao das variaveis
C
        INTEGER N,ND,L(5),I,K,IFLAG,LH,LIW,LW,IW(LIW)
        REAL X(3),M(5,20),NK(5),DD,W(LW),S(5,20),D(5),S1(5),S8,
     *  S2(5,20),S3(5,20),A1,A2,X1E,X132,XX3,X1E2,E,FHESD(3),FHESL(3)
C
        COMMON/DADOS/ND,D,L,S,M,NK,DD
        COMMON/SOMA/S1,S2,S3
        COMMON/S8/S8
C
        DO 1 J=1,LH
          FHESL(J)=0.0
1       CONTINUE
C
        DO 2 J=1,N
          FHESD(J)=0.0
2       CONTINUE
C
C       I = indice da linha do k-esimo grupo
C       K = indice do grupo (dose)
C
        DO 3 K=1,ND
          E=EXP(-X(2)*D(K))
          X1E=X(1)*E
          X132=(X1E+X(3))**2
          A1=(-S1(K)+NK(K)*X1E)*X1E/X132
          A2=S1(K)+NK(K)*X(3)
          XX3=X(3)/X132
          X1E2=(X1E**2)*NK(K)
C
C         Calculo da matriz Hessiana
C
          FHESL(1)=FHESL(1)-X(3)*D(K)*E*A2/X132
          FHESL(2)=FHESL(2)+A1/X(1)
          FHESL(3)=FHESL(3)-A1*D(K)
          FHESD(1)=FHESD(1)+(XX3/(X(1)**2))*((2.0*X1E+X(3))*
     *            S1(K)-X1E2)
          FHESD(2)=FHESD(2)+XX3*(D(K)**2)*X1E*A2
          FHESD(3)=FHESD(3)+S8-(S1(K)+X1E2/X(3))/X132
3         CONTINUE
        RETURN
        END

C       Aqui entra a subrotina MONIT de monitoramento  do  procedimento
C       de minimizacao
```

##

```
C     ================================================
CCCCCC Subrotina MONIT de monitoramento do processo de
CCCCCC minimizacao. Esta subrotina estara localizada
CCCCCC imediatamente apos o final das subrotinas que calculam
CCCCCC as funcoes a serem minimizadas pelos programas da NAG.
CCCCC ================================================

      SUBROUTINE MONIT(N,X,F,G,ISTATE,GPJNRM,COND,POSDEF,
     *  NITER,NF,IW,LIW,W,LW)
C
      INTEGER LIW,LW,N,NF,NITER,ISTATE(N),IW(LIW),ISJ,J
      REAL COND,F,GPJNRM,G(N),W(LW),X(N)
      LOGICAL POSDEF

CCCCCC ------------------------------------------------------------
CCCCCC O usuario interessado deve consultar a documentacao da NAG
CCCCCC para a subrotina E04JBF, E04KBF ou E04LBF, para esclarecer o
CCCCCC significado das variaveis acima.
CCCCC ------------------------------------------------------------

      WRITE(*,99) NITER,NF,F,GPJNRM
      WRITE(*,98)
      DO 100 J=1,N
         ISJ=ISTATE(J)
         IF (ISJ.GT.0) GO TO 20
         ISJ=-ISJ
         GO TO (40,60,80),ISJ
20       WRITE(*,97) J,X(J),G(J)
         GO TO 100
40       WRITE(*,96) J,X(J),G(J)
         GO TO 100
60       WRITE(*,95) J,X(J),G(J)
         GO TO 100
80       WRITE(*,94) J,X(J),G(J)
100   CONTINUE
      IF (COND.EQ.0.0) RETURN
      IF (COND.LE.1.0E+6) GO TO 120
      WRITE(*,*) 'No DE CONDIC ESTIM DA HESS PROJ ) 1.0E+6'
      GO TO 140
120   WRITE(*,92) COND
C
140   IF (.NOT.POSDEF) WRITE(*,*) 'HESS PROJ NAO E POS DEF'
      RETURN
C
99    FORMAT(/5H0ITER, 5X, 8HAVALS FN, 11X, 8HVALOR FN, 11X,
     *  18HNORMA DO GRAD PROJ/I4, 6X, I5, 2(6X,1PE20.4))
98    FORMAT(3H0 J, 11X, 4HX(J), 16X, 4HG(J), 13X, 6HSTATUS)
97    FORMAT(1H , I2, 1X, 1P2E20.4, 5X, 5HLIVRE)
96    FORMAT(1H , I2, 1X, 1P2E20.4, 5X, 7HLIM SUP)
95    FORMAT(1H , I2, 1X, 1P2E20.4, 5X, 7HLIM INF)
94    FORMAT(1H , I2, 1X, 1P2E20.4, 5X, 5HCONST)
92    FORMAT(34H0No DE CONDIC ESTIM DA HESS PROJ =, 1H, 1PE10.2)
      END
```

##

E4. CALCULO DA MATRIZ DE INFORMACAO E PROBABILIDADES DE RESPOSTA CONDICIONAL E INCONDICIONAL

```
C     ================================================================
CCCCCC Programa para calcular a matriz de informacao estimada para
CCCCCC os modelos de dose-resposta e de Poisson. Calcula tambem as
CCCCCC estimativas das probabilidades de resposta condicional e
CCCCCC incondicional.
CCCCCC ================================================================

CCCCC Declaracao de escalares e matrizes
C
      INTEGER ND,I,K,C,L(5)
      REAL*8 M(5,20),Y(20),S(5,20),D(5),NK(5),E(5,20,20),S1(5),
     *  S2(5,20),S3(5,20),AUX1,AUX2,PDK
      CHARACTER NOME*20
      REAL*8 EXAB,EXT,P,Q,R,XD(4),X34,ID(4,4),EU,EA1,EA2,EA3,EE1,
```

```
*  EE2,ETPU,ETEU
   REAL*8 XP(2),EX,NDE,IP(2,2)

   ------------------------------------------------------------
   AUX1,  AUX2, EU, EA1, EA2, EA3, EE1, EE2, ETPU, ETEU, EX, NDE:
   variaveis auxiliares;

   PDK: probabilidade incondicional de resposta adversa;

   XD(4): vetor  de  dimensao 4 com as  estimativas  de  maxima
          verossimilhanca do modelo de dose-resposta;

   ID(4,4): estimativa  da matriz de informacao de  Fisher  dos
            parametros do modelo de dose-resposta;

   XP(2): estimativas de maxima verossimilhanca dos parametros do
          modelo de Poisson;

   IP(2,2): estimativa  da matriz de informacao de  Fisher  dos
            parametros do modelo de Poisson.
   ------------------------------------------------------------

   Escolha da matriz de dados a ser lida

   WRITE(*,*)'ARQUIVO DE DADOS A SER LIDO:'
   READ(*,'(A20)')NOME
   OPEN(1,STATUS='UNKNOWN',FILE=NOME)

   Leitura dos dados

   K = Indice da dose (grupo)
   I = Indice da linha do k-esimo grupo
   J = Indice da coluna

   READ(1,*)ND,C
   READ(1,*) (Y(J),J=1,C)
   DO 1 K=1,ND
     NK(K)=0.D0
     READ(1,*) D(K),L(K)
     DO 2 I=1,L(K)
       READ(1,*) S(K,I),(E(K,I,J),J=1,C),M(K,I)
       NK(K)=NK(K)+M(K,I)
     CONTINUE
   CONTINUE

   Calculo de somatorios auxiliares

   DO 3 K=1,ND
     S1(K)=0.D0
     DO 4 I=1,L(K)
       S1(K)=S1(K)+M(K,I)*S(K,I)
       S2(K,I)=0.D0
       S3(K,I)=0.D0
       DO 5 J=1,C
         S2(K,I)=S2(K,I)+E(K,I,J)*Y(J)
         DIF=S(K,I)-Y(J)
         S3(K,I)=S3(K,I)+E(K,I,J)*DIF
       CONTINUE
     CONTINUE
   CONTINUE

   Inicializacao das variaveis

   WRITE(*,*) 'ALFA^ BETA^ TETA1^ TETA2^:'
   READ(*,*) (XD(I),I=1,4)

C
   WRITE(*,*) 'GAMA1^ GAMA2^:'
   READ(*,*) XP(1),XP(2)

C
   DO 11 I=1,4
     DO 12 J=1,4
       ID(I,J)=0.D0
12   CONTINUE
11 CONTINUE
C
   DO 13 I=1,2
     DO 14 J=1,2
       IP(I,J)=0.D0
14   CONTINUE
```

```
13      CONTINUE
C
        CLOSE(1)
C
        WRITE(*,*)'ARQUIVO DE SAIDA DOS RESULTADOS:'
        READ(*,'(A20)')NOME
        OPEN(1,STATUS='UNKNOWN',FILE=NOME)
C
C       Inicio dos calculos das matrizes de covariancia
C
        DO 6 K=1,ND
C
C       Modelo de dose-resposta
C
          EXAB=-XD(1)-XD(2)*D(K)
          EXAB=DEXP(EXAB)
          X34=XD(3)+XD(4)*D(K)
          R=1.D0-EXAB
          EA1=0.D0
          EA2=0.D0
          EA3=0.D0
C
C       Modelo de Poisson
C
          EX=DEXP(-XP(2)*D(K))
          NDE=NK(K)*D(K)*EX
C
          AUX1=1.D0-DEXP(-X34)
          AUX2=-XP(1)*EX*AUX1
          PDK=R*DEXP(AUX2)
          WRITE(1,*)'D, P(D)',D(K), PDK
          WRITE(1,*)
C
          IP(1,2)=IP(1,2)-NDE
          IP(1,1)=IP(1,1)+EX*NK(K)/XP(1)
          IP(2,2)=IP(2,2)+NDE*D(K)*XP(1)
C
C       Modelo de dose-resposta
C
          DO 7 I=1,L(K)
            EXT=-S(K,I)*X34
            EXT=DEXP(EXT)
            P=R*EXT
            WRITE(1,*)'S, P',S(K,I),P
            WRITE(1,*)
            Q=1.D0-P
            EU=S(K,I)*M(K,I)/(P*Q)
            ETPU=-P*EU
            ETEU=EXAB*EXT*EU
            EA1=EA1+ETEU*EXT
            EA2=EA2+EXT*S(K,I)*ETPU
            EA3=EA3+P*(S(K,I)**2)*ETPU
7         CONTINUE
C
          EE1=EXAB*EA1
          EE2=EXAB*EA2
C
C         Matriz de informacao - modelo de dose-resposta
C
          ID(1,2)=ID(1,2)+D(K)*EE1
          ID(1,3)=ID(1,3)+EE2
          ID(2,3)=ID(2,3)+D(K)*EE2
          ID(1,4)=ID(2,3)
          ID(2,4)=ID(2,4)+(D(K)**2)*EE2
          ID(3,4)=ID(3,4)-D(K)*EA3
          ID(1,1)=ID(1,1)+EE1
          ID(2,2)=ID(2,2)+(D(K)**2)*EE1
          ID(3,3)=ID(3,3)-EA3
          ID(4,4)=ID(4,4)-(D(K)**2)*EA3
C
6         CONTINUE
C
        WRITE(1,*)'MATRIZ DE INFORMACAO - DOSE-RESP.'
        WRITE(1,*)
        WRITE(1,*)'I11=',ID(1,1), 'I12=',ID(1,2)
        WRITE(1,*)'I13=',ID(1,3), 'I14=',ID(2,3)
        WRITE(1,*)'I22=',ID(2,2), 'I23=',ID(2,3)
        WRITE(1,*)'I24=',ID(2,4), 'I33=',ID(3,3)
        WRITE(1,*)'I34=',ID(3,4), 'I44=',ID(4,4)
```

```
        WRITE(1,*)
C
        WRITE(1,*)'MATRIZ DE INFORMACAO - POISSON'
        WRITE(1,*)
        WRITE(1,*)'I11=',IP(1,1), 'I12=',IP(1,2)
        WRITE(1,*)'I22=',IP(2,2)

        CLOSE(1)
        STOP
        END
```

##

E5. COMANDOS GLIM

```
!============================================================
!
!Ajuste dos modelos logisticos propostos por WILLIAMS (1987)
!Dados de LUNING et al (1966)
!
!============================================================
!
$SUBFILE DADLUN
$UNITS 1772$
$DATA I DI YIJ SIJ$
$READ
  0  0.0  0   5
  0  0.0  0   5
  .   .   .   .
  .   .   .   .
  .   .   .   .

  2  0.6  4   9
  2  0.6  6   9 $
!
!Calculo do desvio do modelo de RAI & VAN RYZIN (1985)
!
$CALC %A=0.2238 : %B=1.0377 : %T=0.0963 : %R=-0.0668:
   E=%EXP(-(%A+%B*DI)): F=%EXP(-SIJ*(%T+%R*DI)): P=(1-E)*F: Q=1-P:
   AX=YIJ/(SIJ*P): AZ=(SIJ-YIJ)/(SIJ*Q): L=%LOG(AX): LL=%LOG(AZ):
   DD=YIJ*L+(SIJ-YIJ)*LL: %D=%CU(DD)$
$PRI %D$
$DEL %A: %B: %T: %R: E: F: P: Q: AX: AZ: L: LL: DD: %D$
!
!Ajuste dos modelos logisticos
!
$YVAR YIJ$ $ERROR B SIJ$ $LINK G$
$CALC  S05=%E0(I,0)*%E0(SIJ,5):  S06=%E0(I,0)*%E0(SIJ,6):
       S07=%E0(I,0)*%E0(SIJ,7):  S08=%E0(I,0)*%E0(SIJ,8):
       S09=%E0(I,0)*%E0(SIJ,9):  S010=%E0(I,0)*%E0(SIJ,10):
       S15=%E0(I,1)*%E0(SIJ,5):  S16=%E0(I,1)*%E0(SIJ,6):
       S17=%E0(I,1)*%E0(SIJ,7):  S18=%E0(I,1)*%E0(SIJ,8):
       S19=%E0(I,1)*%E0(SIJ,9):  S110=%E0(I,1)*%E0(SIJ,10):
       S25=%E0(I,2)*%E0(SIJ,5):  S26=%E0(I,2)*%E0(SIJ,6):
       S27=%E0(I,2)*%E0(SIJ,7):  S28=%E0(I,2)*%E0(SIJ,8):
       S29=%E0(I,2)*%E0(SIJ,9):
 F=S05+2*S06+3*S07+4*S08+5*S09+6*S010+7*S15+8*S16+9*S17+10*S18+11*S19+12*S110:
 F=F+13*S25+14*S26+15*S27+16*S28+17*S29$
$DEL S05:S06:S07:S08:S09:S010:S15:S16:S17:S18:S19:S110:S25:S26:S27:S28:S29$
!
!Ajuste do modelo de 17 parametros (um para cada combinacao DI e SIJ)
!
$FACTOR F 17$ $FIT F-1$
$DIS E $  $PLOT YIJ %FV '*'$
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)$
$PLOT RP %FV '*'$
$CALC I0=%E0(I,0): I1=%E0(I,1): I2=%E0(I,2): DS=DI*SIJ$
!
!Ajuste do modelo TALi
!
$FIT I0+I1+I2-1$ $DIS E$ $PLOT YIJ %FV '*'$
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)$
$PLOT RP %FV '*'$
!
!Ajuste do modelo TALi + GAMA SIJ
```

```
$FIT I0+I1+I2+SIJ-1$ $DIS E$ $PLOT YIJ %FV '*'$
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)$
$PLOT RP %FV '*'$

!Ajuste do modelo TALi + GAMA SIJ + DELTA DI*SIJ

$FIT I0+I1+I2+SIJ+DS-1$ $DIS E$ $PLOT YIJ %FV '*'$
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)$
$PLOT RP %FV '*'$

!Ajuste do modelo ALFA + BETA DI

$FIT DI$ $DIS E$ $PLOT YIJ %FV '*'$
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)$
$PLOT RP %FV '*'$

!Ajuste do modelo ALFA + BETA DI + GAMA SIJ

$FIT DI+SIJ$ $DIS E$ $PLOT YIJ %FV '*'$
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)$
$PLOT RP %FV '*'$

!Ajuste do modelo ALFA + BETA DI + GAMA SIJ + DELTA DI*SIJ

$FIT DI+SIJ+DS$ $DISP E $ $PLOT YIJ %FV '*'$
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)$
$PLOT RP %FV '*'$

!Ajuste do modelo TALi + GAMAi SIJ

$CALC S0=I0*SIJ: S1=I1*SIJ: S2=I2*SIJ$
$FIT I0+I1+I2+S0+S1+S2-1$  $DIS E$ $PLOT YIJ %FV '*'$
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)$
$PLOT RP %FV '*'$

$RETURN


!##############################################################################


!============================================================

!Ajuste dos modelos logisticos propostos por WILLIAMS (1987)
!Dados de TYL et al (1983)

!============================================================

$SUBFILE DADLUN
$UNITS 131$
$DATA I DI YIJ SIJ$
$READ
  0  0.0     5   8
  0  0.0     2   9
  .   .      .   .
  .   .      .   .
  .   .      .   .

  4  0.292  16  16
  4  0.292  17  17 $
!
!I = grupo
!DI = nivel da dose
!YIJ = numero de respostas na j-esima ninhada do i-esimo grupo
!SIJ = tamanho da j-esima ninhada do i-esimo grupo
!
!Calculo do desvio para o modelo de RAI & VAN RYZIN (1985)
!
$CALC %A=0.7918 : %B=9.8035 : %T=0.1060 : %R=-0.3630:
   E=%EXP(-(%A+%B*DI)): F=%EXP(-SIJ*(%T+%R*DI)): P=(1-E)*F: Q=1-P:
   AX=YIJ/(SIJ*P): AZ=(SIJ-YIJ)/(SIJ*Q): L=%LOG(AX): LL=%LOG(AZ):
   DD=YIJ*L+(SIJ-YIJ)*LL: %D=%CU(DD)$
$PRI %D$
$DEL %A: %B: %T: %R: E: F: P: Q: AX: AZ: L: LL: DD: %D$
!
!Ajuste dos modelos logisticos
!
$YVAR YIJ$ $ERROR B SIJ$ $LINK G$
```

```
!
!Ajuste do modelo de 50 parametros (um para cada combinacao DI e SIJ)
!
$calc Z8=%EQ(I,0)*%EQ(SIJ,8): Z9=%EQ(I,0)*%EQ(SIJ,9): Z10=%EQ(I,0)*%EQ(SIJ,10):
  Z11=%EQ(I,0)*%EQ(SIJ,11): Z12=%EQ(I,0)*%EQ(SIJ,12): Z13=%EQ(I,0)*%EQ(SIJ,13):
  Z14=%EQ(I,0)*%EQ(SIJ,14): Z15=%EQ(I,0)*%EQ(SIJ,15): Z16=%EQ(I,0)*%EQ(SIJ,16):
  Z17=%EQ(I,0)*%EQ(SIJ,17): Z18=%EQ(I,0)*%EQ(SIJ,18): U4 =%EQ(I,1)*%EQ(SIJ, 4):
  U8 =%EQ(I,1)*%EQ(SIJ, 8): U9 =%EQ(I,1)*%EQ(SIJ, 9): U10=%EQ(I,1)*%EQ(SIJ,10):
  U11=%EQ(I,1)*%EQ(SIJ,11): U12=%EQ(I,1)*%EQ(SIJ,12): U13=%EQ(I,1)*%EQ(SIJ,13):
  U14=%EQ(I,1)*%EQ(SIJ,14): U15=%EQ(I,1)*%EQ(SIJ,15): U16=%EQ(I,1)*%EQ(SIJ,16):
  D6 =%EQ(I,2)*%EQ(SIJ, 6): D8 =%EQ(I,2)*%EQ(SIJ, 8): D9 =%EQ(I,2)*%EQ(SIJ, 9):
  D11=%EQ(I,2)*%EQ(SIJ,11): D12=%EQ(I,2)*%EQ(SIJ,12): D13=%EQ(I,2)*%EQ(SIJ,13):
  D14=%EQ(I,2)*%EQ(SIJ,14): D15=%EQ(I,2)*%EQ(SIJ,15): D16=%EQ(I,2)*%EQ(SIJ,16):
  T1 =%EQ(I,3)*%EQ(SIJ, 1): T2 =%EQ(I,3)*%EQ(SIJ, 2): T4 =%EQ(I,3)*%EQ(SIJ, 4):
  T6 =%EQ(I,3)*%EQ(SIJ, 6): T10=%EQ(I,3)*%EQ(SIJ,10): T11=%EQ(I,3)*%EQ(SIJ,11):
  T12=%EQ(I,3)*%EQ(SIJ,12): T13=%EQ(I,3)*%EQ(SIJ,13): T14=%EQ(I,3)*%EQ(SIJ,14):
  T15=%EQ(I,3)*%EQ(SIJ,15): T18=%EQ(I,3)*%EQ(SIJ,18): T19=%EQ(I,3)*%EQ(SIJ,19):
  Q9 =%EQ(I,4)*%EQ(SIJ, 9): Q10=%EQ(I,4)*%EQ(SIJ,10): Q11=%EQ(I,4)*%EQ(SIJ,11):
  Q12=%EQ(I,4)*%EQ(SIJ,12): Q13=%EQ(I,4)*%EQ(SIJ,13): Q14=%EQ(I,4)*%EQ(SIJ,14):
  Q16=%EQ(I,4)*%EQ(SIJ,16): Q17=%EQ(I,4)*%EQ(SIJ,17)%
!
$CALC F=Z8+2*Z9+3*Z10+4*Z11+5*Z12+6*Z13+7*Z14+8*Z15+9*Z16+10*Z17+11*Z18:
   f=f+12*U4+13*U8+14*U9+15*U10+16*U11+17*U12+18*U13+19*U14+20*U15+21*U16:
   f=f+22*D6+23*D8+24*D9+25*D11+26*D12+27*D13+28*D14+29*D15+30*D16:
   f=f+31*T1+32*T2+33*T4+34*T6+35*T10+36*T11+37*T12+38*T13+39*T14+40*T15:
   f=f+41*T18+42*T19+43*Q9+44*Q10+45*Q11+46*Q12+47*Q13+48*Q14+49*Q16+50*Q17%
!
$DEL Z8:Z9:Z10:Z11:Z12:Z13:Z14:Z15:Z16:Z17:Z18:U4:U8:U9:U10:U11:U12:U13:U14:
   U15:U16:D6:D8:D9:D11:D12:D13:D14:D15:D16:T1:T2:T4:T6:T10:T11:T12:T13:T14:
   T15:T18:T19:Q9:Q10:Q11:Q12:Q13:Q14:Q16:Q17%
!
$FACTOR F 50% $FIT F-1%
$DIS E % $PLOT YIJ %FV '*'%
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)%
$PLOT RP %FV '*'%
!
$CALC I0=%EQ(I,0): I1=%EQ(I,1): I2=%EQ(I,2): I3=%EQ(I,3): I4=%EQ(I,4):
      DS=DI*SIJ%
!
!Ajuste do modelo TALi
!
$FIT I0+I1+I2+I3+I4-1% $DIS E% $PLOT YIJ %FV '*'%
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)%
$PLOT RP %FV '*'%
!
!Ajuste do modelo TALi + GAMA SIJ
!
$FIT I0+I1+I2+I3+I4+SIJ-1% $DIS E% $PLOT YIJ %FV '*'%
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)%
$PLOT RP %FV '*'%
!
!Ajuste do modelo TALi + GAMA SIJ + DELTA DI*SIJ
!
$FIT I0+I1+I2+I3+I4+SIJ+DS-1% $DIS E% $PLOT YIJ %FV '*'%
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)%
$PLOT RP %FV '*'%
!
!Ajuste do modelo TALi + GAMAi SIJ
!
$CALC S0=I0*SIJ: S1=I1*SIJ: S2=I2*SIJ: S3=I3*SIJ: S4=I4*SIJ%
$FIT I0+I1+I2+I3+I4+S0+S1+S2+S3+S4-1%  $DIS E% $PLOT YIJ %FV '*'%
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)%
$PLOT RP %FV '*'%
$DEL I0: I1: I2: I3: I4: S0: S1: S2: S3: S4%
!
!Ajuste do modelo ALFA + BETA DI
!
!$FIT DI% $DIS E% $PLOT YIJ %FV '*'%
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)%
$PLOT RP %FV '*'%
!
!Ajuste do modelo ALFA + BETA DI + GAMA SIJ
!
$FIT DI+SIJ% $DIS E% $PLOT YIJ %FV '*'%
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)%
$PLOT RP %FV '*'%
!
!Ajuste do modelo ALFA + BETA DI + GAMA SIJ
!
```

```
$FIT DI+SIJ+OS$ $DISP E $ $PLOT YIJ %FV '*'$
$CALC VA=%FV*(%BD-%FV)/%BD: RP=(YIJ-%FV)/%SQRT(VA)$
$PLOT RP %FV '*'$
!
$RETURN


##################################################################################


!================================================================================
!
!Sequencia de comandos apresentada por  WILLIAMS (1982)  e  utilizada  no
!ajuste dos modelos logisticos com variacao extra-binomial.
!
!================================================================================
!
!Apos a declaracao das variaveis e leitura dos dados como de costume,
!declare a variavel resposta, o erro e a ligacao
$YVAR YIJ$ $ERROR B SIJ$ $LINK G$
!
!Declare uma matriz de pesos
$WEIG W$
!
!Faca a matriz de pesos igual a identidade
$ACC 8$ $CALC W=1$
!
!Ajuste o modelo desejado (aqui e ajustado o modelo comum W = 1)
$FIT o modelo desejado $  $DIS E$
!
!Imprima a estatistica qui-quadrado de Pearson
$PRI %X2$
!
!Estime e imprima fi e calcule os novos pesos
$EXTR %VL$
$CAL V=%PW*%WT*%VL: %P=(%X2-%CU(%PW*(1-V)))/(%CU((%BD-1)*%PW*(1-V)))$
$PRI %P$ $CAL W=1/(1+%P*(%BD-1))$
!
!Ajuste novamente o modelo desejado (aqui e ajustado o modelo modificado)
$FIT o modelo desejado$
!
!Repita os comandos abaixo ate que a estatistica qui-quadrado seja
!suficientemente proxima dos graus de liberdade do desvio
$PRI %X2$
$EXTR %VL$
$CAL V=%PW*%WT*%VL: %P=(%X2-%CU(%PW*(1-V)))/(%CU((%BD-1)*%PW*(1-V)))$
$PRI %P$ $CAL W=1/(1+%P*(%BD-1))$
!
$RETURN
```

REFERÊNCIAS BIBLIOGRÁFICAS

AITKIN, M., ANDERSON, D., FRANCIS, B., HINDE, J. (1989). **Statistical Modelling in GLIM.** New York: Oxford University Press.

ANCOMBE, F. J. (1949). The statistical analysis of insect counts based on the negative binomial distribution. **Biometrics 5,** 165-173.

ANSCOMBE, F. J. (1950). Sampling theory of the negative binomial and logarithmic series distributions. **Biometrika 37,** 358-382.

BROWN, C. C. (1983). Learning about toxicity in humans from studies on animals. **Chemtech 13,** 350-358.

BROWN, C. C., KOZIOL, J. A. (1983). **Statistical aspects of the estimation of human risk from suspected environmental carcinogens. SIAM Review 25,** 151-181.

BUCK, R. C. (1965). **Advanced Calculus,** Second Edition. New York: Mc Graw-Hill.

CHEN, J. J., KODELL, R. L. (1989). **Quantitative risk assessment for teratological effects. Journal of the American Statistical Association 84,** 966-971.

CORDEIRO, G. M. (1986). **Modelos Lineares Generalizados:** VII Simpósio Nacional de Probabilidade e Estatística. Campinas: UNICAMP-ABE.

DOBSON, A. J. (1983). **Introduction to Statistical Modelling.** London: Chapman and Hall.

FAUSTMAN, E. M., WELLINGTON, D. G., SMITH, W. P., KIMMEL, C. A. (1989).

Characterization of a developmental toxicity dose-response model. **Environmental Health Persectives 79**, 229-241.

FLETCHER, R. (1987). **Practical Methods of Optimization**, Second Edition. Chichester: John Wiley & Sons.

FLETCHER, R., REEVES, C. M. (1964). Function minimization by conjugate gradients. **The Computer Journal 7**, 149-154.

FREEDMAN, D. A., ZEISEL, H. (1988). Cancer risk assessment: from mouse to man. **Statistical Science 3**, 3-56.

FRIEDMAN, G. D. (1987). **Primer of Epidemiology**, Third Edition. New York: Mc Graw-Hill.

GART, J. J., KREWSKI, D., LEE, P. N., TARONE, R. E., WAHRENDORF, J., (1986). **Statistical Methods in Cancer Research. Vol. 3 - The Design and Analysis of Long Term Animal Experiments.** Lyon: International Agency for Research on Cancer.

GILL, P. E., MURRAY, W., WRIGHT, M. E. (1981). **Practical Optimization.** London: Academic Press.

GLIM (1985). **The GLIM System, Release 3.77 Manual.** Oxford: Numerical Algorithms Group.

HASEMAN, J. K., SOARES, E. R. (1976). The distribution of fetal death in control mice and its implications on statistical tests for dominant lethal effects. **Mutation Research 41**, 277-288.

HASEMAN, J. K., KUPPER, L. L. (1979). Analysis of dichotomous response data from certain toxicological experiments. **Biometrics 35**, 281-293.

HEALY, M. J. R. (1988). **GLIM: An Introduction.** New York: Oxford University Press.

HEHL, M. E. (1986). **Linguagem de Programação Estruturada FORTRAN 77.**

São Paulo: Mc Graw-Hill.

JACKSON, J. E. (1972). All count distributions are not alike. **Journal of Quality Technology 4**, 86-92.

JAMES, B. R. (1981). **Probabilidade: Um Curso em Nível Intermediário.** Rio de Janeiro: Instituto de Matemática Pura e Aplicada/CNPQ, Projeto Euclides

JOHNSON, N. L., KOTZ, S. (1969). **Distributions in Statistics - Discrete Distributions.** Boston: Houghton Mifflin Company.

KLEINBAUM D. G., KUPPER, L. L., MORGENSTERN H., (1982). **Epidemiologic Research - Principles and Quantitative Methods.** New York: Van Nostrand Reinhold.

KREWSKI, D., VAN RYZIN, J. (1981). Dose-response models for quantal response toxicity data. Em: CSÖRGÖ, M., DAWSON, D. A., RAO, J. N. K., SALEH, A. K. Md. E. (Eds). **Statistics and Related Topics**, p. 201-231. New York: North Holland Publishing Company.

KUPPER, L. L., HASEMAN, J. K. (1978). The use of a correlated binomial model for the analysis of certain toxicological experiments. **Biometrics 34**, 69-76.

KUPPER, L. L., PORTIER, C., HOGAN, M. D., YAMAMOTO, E. (1986). The impact of litter effects on dose-response modeling in teratology. **Biometrics 42**, 85-98.

KURZEL, R. B., CETRULO, C. L. (1981). The effect of environmental pollutants on human reproduction, including birth defects. **Environmental Science & Technology 15**, 626-640.

LAST, J. M. (Ed), (1988). **A dictionary of Epidemiology,** Second Edition. New York: Oxford University Press.

LONGO, L. D. (1980). Environmental pollution and pregnancy: Risks and

uncertainties for the fetus and infant. **American Journal of Obstetrics and Gynecology 137**, 162-173.

LUENBERGER, D. G. (1984). **Introduction to Linear and Nonlinear Programming.** Menlo Park: Addison-Wesley Publishing Company.

LÜNING, K. G., SHERIDAN, W., YTTERBORN, K. H., GULLBERG, U. (1966). The relation between the number of implantations and the rate of intra-uterine death in mice. **Mutation Research 3**, 444-451.

MARUBINI, E., LEITE, M. L. C., MILANI, S. (1988). Analysis of dichotomous response variables in teratology. **Biometrical Journal 30**, 965-974.

McCAUGHRAN, D. A., ARNOLD, D. W. (1976). Statistical models for number of implantation sites and embryonic deaths in mice. **Toxicology and Applied Pharmacology 38**, 325-333.

McCULLAGH, P., NELDER, J. A. (1989). **Generalized Linear Models.** London: Chapman and Hall.

NAG (1982). FORTRAN Library Manual, Mark 9, Vol. 3. Minimising and Maximising a Function. Illinois: Numerical Algorithms Group.

NELDER, J. A., WEDDERBURN, R. W. M. (1972). Generalized linear models. **Journal of the Royal Statistical Society, Series A 135**, 370-384.

PAUL, S. R. (1982). Analysis of proportions of affected foetuses in teratological experiments. **Biometrics 38**, 361-370.

PEARSON, E. S., HARTLEY, H.O. (Eds) (1970). **Biometrika Tables for Statisticians**. Volume 1, Third Edition. Cambridge University Press.

RAI, K., VAN RYZIN, J. (1980a). **MULTI80: A Computer Program for Risk Assessment of Toxic Substances**. Technical Report N-1512-NIEHS, Rand Corporation: Santa Monica, California.

RAI, K., VAN RYZIN, J. (1980b). **A Note on a Multivariate version of Rolle's Theorem and Uniqueness of Maximum Likelihood Roots.** Technical Report N-1513-NIEHS, Rand Corporation: Santa Monica, California.

RAI, K., VAN RYZIN, J. (1981). A generalized multihit dose-response model for low dose extrapolation. **Biometrics 37,** 341-352.

RAI, K., VAN RYZIN, J. (1985). A dose-response model for teratological experiments involving quantal responses. **Biometrics 41,** 1-9.

RAO, C. R. (1973). **Linear Statistical Inference and its Applications.** New York: John Wiley & Sons.

ROHATGI, V. K. (1976). **An Introduction to Probability Theory and Mathematical Statistics.** New York: John Wiley & Sons.

ROSS, G. J. S., PREECE, D. A. (1985). The negative binomial distribution. **The Statistician 34,** 323-336.

RYAN, L. (1982). Quantitative risk assessment for developmental toxicity. **Biometrics 48,** 163-174.

SEARLE, S. R. (1971). **Linear Models.** New York: John Wiley & Sons.

SHANE, B. S. (1989). Human reproductive hazards. **Environmental Science & Technology 23,** 1187-1195.

SRIVASTAVA, M. S., KHATRI, C. G. (1979). **An Introduction to Multivariate Statistics.** New York: North - Holland.

THISTED, R. A. (1988). **Elements of Statistical Computing: Numerical Computation.** New York: Chapman and Hall.

TRIPATHI, R. C. (1985). Negative binomial distribution. Em: KOTZ, S. JOHNSON, N. E. (Eds). **Encyclopedia of Statistical Sciences,** vol. 6. p. 169-177. New York: John Wiley & Sons.

TYL, R. W., JONES-PRICE, C., MARR, M. C., KIMMEL, C. A. (1983). Teratologic evaluation of diethylhexyl phthalate (CAS N° 111-81-7) in CD-1 mice. Technical Report (NCTR/NTP Contract 222-80-2031 (C), NTIS PB85105674), Springfield, VA: National Technical Information Service.

TYL, R. W., JONES-PRICE, C., MARR, M. C., KIMMEL, C. A. (1988). Developmental toxicity evaluation of dietary di(2-ethylhexyl) phthalate in Fisher 344 rats and CD-1 mice. **Fundamental and Applied Toxicology 10,** 395-412.

VAN RYZIN, J. (1980). Quantitative risk assessment. **Journal of Occupational Medicine 22,** 321-326.

WARE, J. H., LOUIS, T. A. (1983). Statistical problems in environmental research. **The Canadian Journal of Statistics 11,** 51-70.

WEDDERBURN, R. W. M. (1974). Quasi-likelihood functions, generalized linear models and the Gauss Newton method. **Biometrika 61,** 439-447.

WEDDERBURN, R. W. M. (1976). On the existence and uniqueness of maximum likelihood estimates for certain generalized linear models. **Biometrika 63,** 27-32.

WEIL, C. S. (1970). Selection of the valid number of sampling units and a consideration of their combination in toxicological studies involving reproduction, teratogenesis or carcinogenesis. **Food and Cosmetics Toxicology 8,** 177-182.

WILLIAMS, D. A. (1975). The analysis of binary responses from toxicological experiments involving reproduction and teratogenicity. **Biometrics 31,** 949-952.

WILLIAMS, D. A. (1982). Extra-binomial variation in logistic linear models. **Applied Statistics 31,** 144-148.

WILLIAMS, D. A. (1987). Dose-response models for teratological experiments. **Biometrics 43,** 1013-1016.

WILLIAMS, D. A. (1988). Estimation bias using the beta-binomial distribution in teratology. **Biometrics 44,** 305-309.

WILSON, J. G., FRASER, F. C. (Eds) (1977). **Handbook of Teratology, Volume 1: General Principles and Etiology.** New York: Plenum Press.

YAMAMOTO, E. YANAGIMOTO, T. (1988). Litter effects to dose-response curve estimation. **Journal of the Japan Statistical Society 18,** 97-106.