

**Universidade de São Paulo**  
**Escola Superior de Agricultura “Luiz de Queiroz”**

**Modelagem para dados de parasitismo**

**João Maurício Araújo Mota**

Tese apresentada para obtenção do título de Doutor  
em Agronomia. Área de concentração: Estatística e  
Experimentação Agronômica

**Piracicaba**  
**Julho - 2005**

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

João Maurício Araújo Mota

Estatístico

## **Modelagem para dados de parasitismo**

Orientadora:

Prof<sup>ª</sup> Dr<sup>ª</sup> **CLARICE GARCIA BORGES DEMÉTRIO**

Tese apresentada para obtenção do título de Doutor  
em Agronomia. Área de concentração: Estatística e  
Experimentação Agronômica

**Piracicaba**

**Julho - 2005**

**Dados Internacionais de Catalogação na Publicação (CIP)  
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP**

Mota, João Maurício Araújo

Modelagem para dados de parasitismo / João Maurício Araújo. - - Piracicaba,  
2005.

153 p.

Tese (Doutorado) - - Escola Superior de Agricultura Luiz de Queiroz, 2005.  
Bibliografia.

1. Modelagem de dados 2. Parasitismo 3. Processos de Poisson I. Título

CDD 595.700212

**“Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor”**

## DEDICATÓRIA

À minha mãe

**Inah Araújo Mota**

Mulher virtuosa, criou 12 filhos sozinha e cuja crença em Deus é tão forte que se irradia por todos os seus descendentes. Meu exemplo de vida.

À minha esposa,

**Rosa Maria,**

o imenso amor, a suprema dedicação, o apoio incondicional na batalha do doutorado. Se mil vidas, eu tivesse, mil vezes me casaria com você. Claro que só se você ainda me quisesse!!!!

Aos meus três desdobramentos celulares **Bruna, Sula** e **André**, o orgulho da paternidade, a certeza de um mundo melhor na evolução e no proceder cotidiano de cada um. Ah, se no universo existissem mais pessoas como vocês. Que maravilha viver!!!. Quanta paixão!!!!!!!

## Agradecimentos

À Prof<sup>ª</sup> Dr<sup>ª</sup> Clarice Garcia Borges Demétrio, a orientação, o constante estímulo, a amizade, a infinita paciência, não apenas as suas palavras certas mais também às do Dema nas minhas inúmeras tentativas de desistir e, finalmente, a infinita confiança depositada na realização desse nosso trabalho.

Ao Prof. Roberto Cláudio Frota Bezerra, meu exemplo de mestre, por sempre acreditar em mim desde a primeira aula da graduação.

Ao Prof. Doutor Dalton Francisco de Andrade, a orientação, a amizade compartilhada desde o IME-USP.

Ao Prof. Dr. Décio Barbin, o exemplo de vida, a amizade, sua disponibilidade para ouvir e discutir.

Ao Prof. Dr. Silvio Sandoval Zocchi e à Prof<sup>ª</sup> Dr<sup>ª</sup> Roseli Aparecida Leandro, a amizade, o carinho e a convivência.

Ao Prof. Dr. Carlos Tadeu dos Santos Dias e à Prof<sup>ª</sup> Dr<sup>ª</sup> Sonia Maria de Stefano Piedade , a amizade e os ensinamentos recebidos.

Ao Prof. Dr. Edwin Moisés Marcos Ortega, as valiosas sugestões durante o meu exame de qualificação e a oportunidade de conviver com sua bela família.

À Prof<sup>ª</sup> Dr<sup>ª</sup> Silvia Maria de Freitas, da UFC, a leitura e a discussão do material dessa tese.

À Prof<sup>ª</sup> Ana Maria Sousa de Araújo, da UFC, a amizade, a moradia, o eterno incentivo, a ajuda no Látex e as inúmeras discussões das angústias teóricas de minha tese.

Ao Prof. Dr. Renato Ribeiro de Lima, da UFAL, a orientação e a discussão de vários programas computacionais que muito me ajudaram a compreender a parte teórica de minha tese além de me permitir a convivência com sua família o que originou uma bela amizade.

Ao Prof. Dr. Silvano Cesar da Costa, da Universidade Estadual de Londrina, a imensa paciência em me ensinar o abc do Látex, a amizade familiar e as inúmeras discussões científicas durante a minha passagem por Piracicaba.

Ao Prof. Dr. Adriano Ferreti Borgatto, da Universidade Federal de Santa Catarina, a amizade, o companheirismo, o incentivo e a ajuda computacional.

À Universidade Federal do Ceará, meu grande orgulho, o afastamento concedido para a realização deste trabalho e aos professores do meu Departamento de Estatística e Matemática Aplicada (DEMA), Rosa, Sílvia, Ana, Welliandre, Robson, Ronald, Nelson, André, Vicente, Manoel, Lassance e Rafael o incentivo e as ameaças para que eu terminasse o doutorado. As secretárias do Dema, Márgeri, Luisa e Mariluse, a enorme compreensão e a amizade que sempre tiveram com este vestusto lente.

Aos funcionários do Departamento de Ciências Exatas da ESALQ/USP, Luciane, Solange, Expedita e Jorge que sempre estavam prontos para um socorro, uma palavra amiga e dessa maneira propiciaram condições para a realização deste trabalho.

Aos colegas e amigos de doutorado da turma 2001: Renato, Adriano, Osmir, Antonio Carlos, Telde, a força, a amizade, a troca de conhecimentos e a atenção recebida em todos os momentos. Formamos um grande time!!!

Aos colegas e amigos do mestrado e doutorado: Moita, Geneville, Beth, Betti, Juliana, Osmar, Melissa, Lúcio, Hélio, Denise, a força, a amizade, a troca de conhecimentos e a atenção recebida em todos os momentos.

A todos que, de forma direta ou indireta, contribuíram para a realização deste trabalho.

## SUMÁRIO

	Página
RESUMO . . . . .	7
ABSTRACT . . . . .	8
1 INTRODUÇÃO . . . . .	9
2 DESENVOLVIMENTO . . . . .	13
2.1 Modelos . . . . .	13
2.1.1 Modelos de Bakker - M1 a M4 . . . . .	14
2.1.2 Modelo de Rogers - M5 . . . . .	21
2.1.3 Modelo de Griffiths - M6 . . . . .	22
2.1.4 Modelos de Hemerik - M7 a M13 . . . . .	26
2.1.5 Modelo de Janardan - M14 . . . . .	29
2.1.6 Modelos de Daley e Maindonald - M15 a M18 . . . . .	32
2.1.7 Modelo generalizado de Poisson - M19 . . . . .	36
2.2 Estimação . . . . .	38
2.3 Equivalência e relação entre modelos . . . . .	42
2.4 Modelo de Faddy - M20 . . . . .	46
2.4.1 Aproximação para o modelo de Faddy . . . . .	51
2.4.2 Modelo de Faddy como generalização . . . . .	52
2.4.3 Modelagem das taxas de transição . . . . .	59
2.4.4 Diagrama de Podlich . . . . .	65
2.4.5 Aplicações . . . . .	67
3 CONSIDERAÇÕES FINAIS . . . . .	73
REFERÊNCIAS . . . . .	75
APÊNDICES . . . . .	78
ANEXOS . . . . .	119



## RESUMO

### Modelagem para dados de parasitismo

Experimentos com diferentes objetivos têm sido conduzidos a fim de se estudar o mecanismo do parasitismo, sendo muito comuns os bioensaios para encontrar condições ótimas para a produção de parasitas e para definir estratégias para liberações inundativas no campo. Assim, por exemplo, o número de ovos parasitados depende de fatores como: espécie, tipo e densidade do hospedeiro, longevidade do adulto e densidade do parasita, tipo de alimentação, temperatura, umidade etc. Logo, o objetivo de um determinado ensaio pode ser, então, estudar o comportamento da variável resposta como função do número de parasitoides ou do número de hospedeiros ou ainda do tipo de alimentação. Pode-se verificar que, em geral, as variáveis observáveis são contagens ou somas aleatórias de variáveis aleatórias ou proporções com denominadores fixos ou aleatórios. A distribuição padrão para modelar contagens é a Poisson enquanto que para proporções é a binomial. Em geral, elas não se ajustam dados oriundos do processo de parasitismo, pois suas pressuposições não são satisfeitas, e surgiram modelos alternativos e que levam em consideração o mecanismo de evitar o superparasitismo. Alguns deles supõem que a probabilidade de fuga (evitar o superparasitismo) é função do número de ovos presentes no hospedeiro (Bakker, 1967,1972; Rogers, 1975; Griffiths, 1977), outros não consideram tal processo (Daley; Maindonald, 1989; Griffiths, 1977). Outros, ainda, incluem o comportamento seletivo do parasita na escolha do hospedeiro e a habilidade do hospedeiro em atrair o parasita (Hemerik et al, 2002). Alguns deles surgiram independentemente, outros como generalizações, sendo, portanto, de interesse um estudo adicional para ressaltar pontos comuns entre eles. No presente trabalho, são estudados 19 modelos probabilísticos para explicar a distribuição do número de ovos postos por um parasita em um determinado hospedeiro. Foi mostrada a equivalência entre alguns deles e, além disso, foi provado que um modelo usado por Faddy (1997) na estimação do tamanho de população animal generaliza-os. As propriedades desse modelo são apresentadas e discutidas. O uso do modelo de Faddy para a distribuição do número de ovos no sistema parasita-hospedeiro é o principal resultado teórico dessa tese.

Palavras-chave: Parasitismo; Evitar superparasitismo; Processo de Poisson Estendido

## ABSTRACT

### Modelling parasitism data

Experiments with distinct aims have been conducted in order to study the parasitism mechanism. Bioassays to find optimum conditions for parasite production and to define strategies for application in the field are very common. In this way, for example, the number of parasitized eggs depends on factors such as: species, type and host density, adult longevity and parasite density, type of food, temperature, humidity etc. Hence, the objective of a specific assay can be to study the response variable behavior as a function of the number of parasitoids, number of hosts or type of food. In general, the observable variables are counts or random sums of random variables or proportions with fixed or random denominators. Poisson is the standard distribution used to model counts, while the distribution used for proportions is the binomial. In general, they do not fit to standard distributions don't fit to data generated by the parasitism process, because their assumptions are not satisfied and alternative models that consider the mechanism of avoidance of superparasitism have appeared in the literature. Some of them consider that the refuse probability (avoidance of superparasitism) is a function of the number of eggs in the host (Bakker, 1967, 1972; Rogers, 1975; Griffiths, 1977) others don't consider such process (Daley and Maindonald 1989; Griffiths, 1977). Some include the selective behavior of the parasite in the choice of the host and the host ability in attract the parasite (Hemerick et al., 2002). Some of the models have appeared independently, others as generalizations. There is therefore some interest in making a study of the common points of the models. In this work 19 probability models found in the literature are presented to explain the distribution of the number of eggs laid by a parasite on a specific host. As an initial result, the equivalence between some of these models is shown. Also it is shown that a model developed by Faddy (1997) can be considered as a generalization of 18 of the models. The properties of this model are presented and discussed. The equivalence between the Janardan and Faddy models is an original and interesting result. The use of Faddy's model as a general probability model for the distribution of the number of eggs in a parasite-host system is the main theoretical result of this thesis.

Keywords: Parasitism; Avoidance of superparasitism; Extended Poisson Processes

## 1 INTRODUÇÃO

Um importante ramo da Entomologia que teve um grande avanço após a década de 60 com o desenvolvimento de técnicas de criação de insetos de forma intensiva, segundo Parra; Zucchi (1997), foi o controle biológico aplicado de pragas (CBA). Muitas pesquisas foram feitas nessa área, utilizando inimigos naturais de pragas (parasitóides e/ou predadores) e tendo, também, como objetivo a preservação do meio ambiente e do ecossistema. No Brasil, por exemplo, a broca-da-cana (*Diatraea saccharalis*), a principal praga da cana-de-açúcar, e a lagarta do cartucho (*Spodoptera frugiperda*), uma das pragas mais importantes da cultura do milho, vêm sendo combatidas pelo método CBA.

A Embrapa Agropecuária Oeste está desenvolvendo tecnologia para controlar biologicamente os percevejos da soja através do inimigo natural *Telenomus podisi*, mais conhecido como vespinha. Os percevejos diminuem a quantidade de óleo nos grãos, provocando o desenvolvimento de fungos e causando retenção foliar, o que dificulta a colheita. A tecnologia ajudará a reduzir o uso de inseticidas nas lavouras de pequenos e médios produtores, agricultores familiares e produtores orgânicos. A empresa cria anualmente percevejos em gaiolas para copularem e produzirem ovos. Estes são coletados, armazenados em “freezer” durante 6 meses e depois colados em cartelas para serem oferecidos ao parasitóide por 24 horas. Depois de 10 dias, o parasitóide completa o ciclo, sendo as cartelas colocadas no campo no dia anterior. Após o experimento, a equipe analisa o índice de parasitismo natural nos ovos e da população de percevejos.

Há muitas classes de métodos de controle, embora uma classificação geral possa incorporar os sistemas clássicos predador-presa a parasita-hospedeiro (Payne, 1988; Faddy e Fenlon, 1999). Dentro do último grupo estão os macroparasitas (por exemplo, insetos himenópteros que parasitam ovos, larvas ou pupas e os nematóides entomológicos) e os microparasitas (por exemplo, baculovirus e *Bacillus thuringensis*). Os insetos parasitóides põem seus ovos interna ou externamente nos hospedeiros, os responsáveis por fornecer alimentação para os descendentes dos parasitóides. Quando a ovoposição no hospedeiro é feita internamente, devem ser consideradas três fases:

- i) encontro e exame externo do hospedeiro pelo parasitóide, durante um determinado tempo (verificação de tamanho do hospedeiro, viabilidade etc) - um encontro significa uma oportunidade de penetração do ovopositor que pode, ou não, acontecer;
- ii) penetração do ovopositor e reconhecimento interno - significa uma oportunidade de colocação de ovos que pode, ou não, acontecer;
- iii) oviposição (colocação de ovos propriamente dita) caracterizada por movimentos verticais do abdome.

Quando a oviposição no hospedeiro é feita externamente, devem ser consideradas duas fases:

- i) encontro e exame externo do hospedeiro pelo parasitóide, durante um determinado tempo (verificação de tamanho do hospedeiro, viabilidade etc) - um encontro significa uma oportunidade de colocação de ovos que pode, ou não, acontecer;
- ii) oviposição (colocação de ovos propriamente dita).

Os parasitóides podem colocar um ou mais ovos por hospedeiro. A colocação de mais de um ovo por hospedeiro é conhecida como superparasitismo (Smith, 1916) e é, em geral, evitada, para não haver competição por alimento.

No caso em que o parasitóide é um nematóide, devem ser consideradas duas fases:

- i) encontro e exame externo do hospedeiro pelo nematóide, durante um determinado tempo (verificação de tamanho do hospedeiro, viabilidade etc) - um encontro significa uma oportunidade de invasão que pode, ou não, acontecer;
- ii) invasão do hospedeiro pelo nematóide.

Experimentos com diferentes objetivos têm sido conduzidos a fim de se estudar o mecanismo do parasitismo, sendo muito comuns os bioensaios para encontrar condições ótimas para a produção de parasitas e para definir estratégias para liberações inundativas no campo. Assim, por exemplo, o número de ovos parasitados depende de fatores como:

espécie, tipo e densidade do hospedeiro, longevidade do adulto e densidade do parasita, tipo de alimentação, temperatura, umidade etc. Logo, o objetivo de um determinado ensaio pode ser, então, estudar o comportamento da variável resposta como função do número de parasitóides ou como função do número de hospedeiros ou ainda como função do tipo de alimentação. Em geral,  $K$  parasitas (usualmente,  $K = 1$ ) são confinados, por um período de tempo  $T$ , com  $H$  hospedeiros (na maioria das vezes  $H \geq 5$ ). Dependendo do objetivo do experimento, diferentes variáveis podem ser observadas:

- números de encontros  $R_j$  vivenciados pelo  $j$ -ésimo, ( $j = 1, 2, \dots, H$ ), hospedeiro, sendo  $R = \sum_{j=1}^H R_j$  o número total de encontros;
- números de ovoposições  $V_j$  vivenciadas pelo  $j$ -ésimo, ( $j = 1, 2, \dots, H$ ), hospedeiro, o que pode ser registrado através de observação de filmagem dos movimentos do abdômen do inseto, sendo  $N = \sum_{j=1}^H V_j$  o número total de ovoposições;
- números de ovos  $S_j$ ,  $j = 1, 2, \dots, H$  do parasita no  $j$ -ésimo,  $j = 1, 2, \dots, H$ , hospedeiro, que podem ser observados após o tempo de confinamento parasita-hospedeiro (processo destrutivo);
- os números de hospedeiros  $X_i$  com  $i$  ovos, sendo o número de hospedeiros  $H = \sum_{i=0}^{\infty} X_i$  e o número total de ovos  $S = \sum_{j=1}^H V_j = \sum_{i=0}^{\infty} iX_i$ ; na maioria dos casos  $X_i = 0$ , para  $i \geq 3$ , de acordo com Daley e Maindonald (1989).
- proporção de ovos parasitados ( $\frac{S}{H}$ ) e, em determinados casos, é facilmente observada, pois os ovos mudam de cor após serem parasitados;
- proporção de ovos com sinal de emergência ( $\frac{Y}{S}$ );
- número de descendentes ( $Z$ );
- número de fêmeas ( $W$ );
- proporção de fêmeas entre os descendentes ( $\frac{W}{Z}$ ), chamada razão sexual.

Borgatto (2004), por exemplo, discute um conjunto de dados de um experimento completamente casualizado com 10 repetições com duas espécies de um parasita (*Trichogramma galloi*). Os tratamentos utilizados foram números de fêmeas (2, 4, 8, 16, 32, 64, 128) que foram colocadas durante 24 horas para parasitar 128 ovos do *Anagasta kuehniella*, um hospedeiro alternativo economicamente mais viável do que o hospedeiro natural *Diatraea saccharalis*, sendo a variável resposta o número de ovos parasitados. O interesse do pesquisador era comparar as espécies e determinar o número ótimo de fêmeas para maximizar a proporção de ovos parasitados.

Pode-se verificar que, em geral, as variáveis observáveis são contagens ou somas aleatórias de variáveis aleatórias ou proporções com denominadores fixos ou aleatórios. A distribuição padrão utilizada para modelar contagens é a Poisson enquanto que para proporções é uma função da binomial. Em geral, essas distribuições padrões não se ajustam a esses tipos de dados, pois as pressuposições delas não são satisfeitas.

Surgiram na literatura vários modelos alternativos para a análise de dados oriundos do processo de parasitismo e que levam em consideração o mecanismo de evitar o superparasitismo. Alguns deles consideram que a probabilidade de fuga (evitar o superparasitismo) é função do número de ovos presentes no hospedeiro (Bakker, 1967,1972; Rogers, 1975; Griffiths, 1977), outros como o modelo nulo de Daley e Maindonald (1989) e um particular modelo de Griffiths (1977) não consideram tal processo. Outros, ainda, incluem o comportamento seletivo do parasita na escolha do hospedeiro e o poder do hospedeiro em atrair o parasita (Hemerik et al, 2002).

Alguns desses modelos surgiram independentemente, outros como generalizações, sendo, portanto, de interesse um estudo adicional para ressaltar pontos comuns entre eles e, além disso, obter testes para o uso de modelos mais simples. Especificamente, os principais objetivos deste trabalho foram:

- (i) fazer uma revisão de literatura sobre esses modelos com ênfase naqueles em que existe o mecanismo para evitar o superparasitismo, uniformizando a notação;
- (ii) comparar os modelos existentes, e verificar se algum deles generaliza os demais;
- (iii) implementar os métodos em programas computacionais.

## 2 DESENVOLVIMENTO

### 2.1 Modelos

A distribuição do número de ovos de parasitas em hospedeiros é resultado de um processo que envolve duas etapas: a procura pelo hospedeiro e a ovoposição. Um comportamento observado é que os parasitas tendem a evitar o superparasitismo e isso depende das condições em que ocorre: laboratório ou campo. Estudos quantitativos descrevendo superparasitismo em insetos, segundo Rogers (1975), começaram com Fiske (1910)<sup>1</sup> que concluiu que havia uma relação envolvendo a proporção de hospedeiros parasitados ( $\pi$ ) e o número médio de ovos postos ( $\lambda$ ) em cada hospedeiro supondo que cada ovoposição é unitária, pois ele acreditava que superparasitismo é um fenômeno do tipo determinístico (tudo ou nada). Essa relação, de acordo com Rogers (1975), foi expressa por Thompson (1924)<sup>2</sup> através da fórmula de Poisson

$$\pi = 1 - e^{-\lambda}.$$

Rogers (1975) afirmou que, posteriormente, a distribuição de Poisson foi usada para analisar o número de ovos postos em cada hospedeiro e que estudos da qualidade do ajuste da distribuição mostraram que, embora o superparasitismo possa ocorrer entre os parasitas solitários, é raramente encontrado. Isso sugere que os parasitas (fêmeas) tendem a evitar o superparasitismo mas suas habilidades de como detectar hospedeiros já parasitados ou de refrear ovoposição neles é incompleta. Ainda, de acordo com Rogers (1975) estudos que envolvem tal comportamento para evitar o superparasitismo foram feitos por Salt (1932) e Walker (1937), porém, o mecanismo sugerido não é realístico. O modelo permitia uma defasagem entre a ocorrência de hospedeiro e parasita dentro de uma geração por assumir que a quantidade diária de ovos postos era dividida entre hospedeiros parasitados, ou não, de acordo com a proporção de não parasitados e parasitados.

A modelagem de dados provenientes de estudos de parasitismo de ovos, porém, só teve avanços significativos a partir do final da década de 60. A seguir, serão apresentados

---

<sup>1</sup>**Fiske, W.F. (1910).** Superparasitism: an important factor in the natural control of insects. J. econ. Ent. 3,88-97.

<sup>2</sup>**Thompson, W.R. (1924).** La theorie mathématique de l'action des parasites entophages et le facteur du hasard. Annals Fac. Sci. Marseille, Ser.11,2,1-89.

os modelos que surgiram na literatura nas últimas quatro décadas.

### 2.1.1 Modelos de Bakker - M1 a M4

Modelos para descrever a distribuição do número de ovos do parasitóide tipo vespinha *Pseudeucoila bochei* Weld. em *Drosophila larvae* foram apresentados por Bakker et al. (1967, 1972), pois a distribuição de Poisson não se ajusta a esse tipo de dados (Tabela 1).

Tabela 1 - Distribuição do número de ovos de *Pseudeucoils bochei* (vespa) sobre larva de *Drosophila melanogaster*

Ensaio		Número de ovos por hospedeiro						Total
		0	1	2	3	4	> 4	
1	Observado	24	30	5	1	0	0	60
	Poisson(0,72)	29,30	21,00	7,52	1,80	0,32	0,06	60
2	Observado	8	30	9	2	0	0	49
	Poisson(1,10)	16,28	17,94	9,88	3,63	1,00	0,77	49
3	Observado	3	36	41	6	0	0	86
	Poisson(1,58)	17,69	27,97	22,11	11,66	4,61	1,96	86
4	Observado	1	23	40	25	2	0	91
	Poisson(2,04)	11,78	24,09	24,62	16,77	8,57	5,17	91
5	Observado	1	12	42	28	5	0	88
	Poisson(2,27)	9,07	20,61	23,42	17,74	10,08	7,08	88
6	Observado	1	10	51	40	6	0	109
	Poisson(2,37)	10,09	23,92	28,35	22,40	13,28	10,96	109

As observações laboratoriais mostraram que as parasitas fêmeas reconheciam os hospedeiros já parasitados e evitavam a ovoposição. Assim, quatro modelos foram propostos, assumindo-se que o mecanismo sob o processo de superparasitismo está sujeito às condições:



- i. os dois estágios de parasitismo, a procura pelo hospedeiro e a ovoposição, são independentes;
- ii. o processo de procura, conduz cada hospedeiro a um número aleatório de encontros com os parasitóides, segundo a distribuição de Poisson com média  $\lambda$ , isto é, a probabilidade de o número de encontros  $R_j$  de um dado hospedeiro ser igual a  $r$  é dada por

$$q_r = P(R_j = r) = \frac{e^{-\lambda} \lambda^r}{r!} I_{\{0,1,2,\dots\}}(r); \quad (1)$$

- iii. no primeiro encontro entre hospedeiro e os parasitóides, é sempre posto um único ovo com probabilidade  $\delta_0 = 1$ ;
- iv. a cada encontro, subsequente ao primeiro, com um hospedeiro que tenha  $i$  ovos, o parasitóide depositará um único ovo com probabilidade  $\delta_i < 1$  ou evitará o superparasitismo com probabilidade  $1 - \delta_i$ ;
- v. o comportamento de evitar o superparasitismo ocorre a uma proporção fixada de encontros múltiplos e tem a probabilidade assumida independente do número e da idade de parasitóides imaturos já presentes no hospedeiro e das condições fisiológicas do parasita.

Bakker (1972) propõe a distribuição geral de probabilidade de  $S_j$ , o número de ovos postos no  $j$ -ésimo hospedeiro, que para  $s = 0$  é dada por:

$$P(S_j = 0) = P(R_j = 0) = e^{-\lambda}, \quad (2)$$

e para  $s \geq 1$ , por

$$\begin{aligned} P(S_j = s) &= \sum_{i=s}^{\infty} P(R_j = i, S_j = s) \\ &= \sum_{i=s}^{\infty} P(R_j = i) P(S_j = s | R_j = i) \\ &= \sum_{i=s}^{\infty} \frac{e^{-\lambda} \lambda^i}{i!} P(S_j = s | R_j = i). \end{aligned} \quad (3)$$

A probabilidade de que ocorra exatamente um ovo, dado que haja  $i$  encontros, é  $(1 - \delta_1)^{i-1}$ , pois o parasita põe um ovo no primeiro encontro e evita por ovos nos demais

$(i - 1)$  encontros. Portanto, a probabilidade que um hospedeiro contenha exatamente um ovo é dada por:

$$\begin{aligned}
P(S_j = 1) &= \sum_{i=1}^{\infty} \frac{e^{-\lambda} \lambda^i}{i!} P(S_j = 1 | R_j = i) \\
&= \sum_{i=1}^{\infty} \frac{e^{-\lambda} \lambda^i}{i!} (1 - \delta_1)^{i-1} \\
&= \frac{e^{-\lambda}}{(1 - \delta_1)} \sum_{i=1}^{\infty} \frac{[\lambda(1 - \delta_1)]^i}{i!} \\
&= \frac{e^{-\lambda}}{(1 - \delta_1)} \left[ \sum_{i=0}^{\infty} \frac{[\lambda(1 - \delta_1)]^i}{i!} - 1 \right] \\
&= \frac{e^{-\lambda}}{(1 - \delta_1)} [e^{\lambda(1 - \delta_1)} - 1] \\
&= \frac{1}{(1 - \delta_1)} [e^{-\lambda \delta_1} - e^{-\lambda}],
\end{aligned}$$

e assim,

$$P(S_j = 1) = \frac{e^{-\lambda \delta_1} - e^{-\lambda}}{1 - \delta_1}. \quad (4)$$

Fazendo  $s = 2$  em (3) e lembrando que a probabilidade condicional de o parasita por 2 ovos, dado que aconteceram  $i$  encontros, é:

$$P(S_j = 2 | R_j = i) = \sum_{k=2}^i \delta_1 (1 - \delta_1)^{k-2} (1 - \delta_2)^{i-k}, \quad (5)$$

pois o primeiro ovo é posto com probabilidade 1 no primeiro encontro e o segundo ovo é posto no encontro  $k = 2, 3, \dots, i$  com probabilidade  $\delta_1$ . Assim, do segundo até o  $(k - 1)$ -ésimo encontro, o parasita evitou o superparasitismo em  $(k - 2)$  encontros com probabilidade  $(1 - \delta_1)^{k-2}$ . Daí em diante, evitou o superparasitismo do  $(k + 1)$  ao  $i$ -ésimo encontro com probabilidade  $(1 - \delta_2)^{i-k}$ , isto é, durante  $(i - k)$  encontros com um hospedeiro com 2 ovos. Assim,

$$\begin{aligned}
P(S_j = 2) &= \sum_{i=2}^{\infty} \frac{e^{-\lambda} \lambda^i}{i!} P(S_j = 2 | R_j = i) \\
&= \sum_{i=2}^{\infty} \frac{e^{-\lambda} \lambda^i}{i!} \sum_{k=2}^i \delta_1 (1 - \delta_1)^{k-2} (1 - \delta_2)^{i-k} \\
&= \frac{\delta_1 e^{-\lambda}}{(1 - \delta_1)^2} \sum_{i=2}^{\infty} \frac{[\lambda(1 - \delta_2)]^i}{i!} \sum_{k=2}^i \left[ \frac{(1 - \delta_1)}{(1 - \delta_2)} \right]^k \\
&= \frac{\delta_1 e^{-\lambda}}{(1 - \delta_1)^2} \sum_{i=2}^{\infty} \frac{[\lambda(1 - \delta_2)]^i}{i!} \sum_{k=2}^i a^k \\
&= \frac{\delta_1 e^{-\lambda}}{(1 - \delta_1)^2} \sum_{i=2}^{\infty} \frac{[\lambda(1 - \delta_2)]^i}{i!} \frac{a(a - a^i)}{1 - a} \\
&= \frac{\delta_1 a e^{-\lambda}}{(1 - \delta_1)^2 (1 - a)} \left[ a \sum_{i=2}^{\infty} \frac{[\lambda(1 - \delta_2)]^i}{i!} - \sum_{i=2}^{\infty} \frac{[\lambda a (1 - \delta_2)]^i}{i!} \right] \\
&= \frac{a}{1 - a} \frac{\delta_1 e^{-\lambda}}{(1 - \delta_1)^2} [a e^{\lambda(1 - \delta_2)} - a - \lambda(1 - \delta_2)a - e^{\lambda(1 - \delta_1)} + 1 + \lambda(1 - \delta_1)] \\
&= \frac{a}{1 - a} \frac{\delta_1 e^{-\lambda}}{(1 - \delta_1)^2} [a e^{\lambda(1 - \delta_2)} + (1 - a) - e^{\lambda(1 - \delta_1)}] \\
&= \frac{1 - \delta_1}{(\delta_1 - \delta_2)} \frac{\delta_1 e^{-\lambda}}{(1 - \delta_1)^2} \left[ \frac{1 - \delta_1}{1 - \delta_2} e^{\lambda(1 - \delta_2)} + \frac{\delta_1 - \delta_2}{1 - \delta_2} - e^{\lambda(1 - \delta_1)} \right]
\end{aligned} \tag{6}$$

em que  $a = \left[ \frac{(1 - \delta_1)}{(1 - \delta_2)} \right]$ ,  $1 - a = \left[ \frac{(\delta_1 - \delta_2)}{(1 - \delta_2)} \right]$  e  $\frac{a}{1 - a} = \left[ \frac{(1 - \delta_1)}{(\delta_1 - \delta_2)} \right]$ , para  $\delta_1 \neq \delta_2$ .

Assim,

$$P(S_j = 2) = \delta_1 \left[ \frac{e^{-\lambda \delta_2}}{(1 - \delta_2)(\delta_1 - \delta_2)} + \frac{e^{-\lambda \delta_1}}{(1 - \delta_1)(\delta_2 - \delta_1)} + \frac{e^{-\lambda}}{(\delta_1 - 1)(\delta_2 - 1)} \right]. \tag{7}$$

De uma maneira geral,  $P(S_j = s)$ ,  $s = 1, 2, \dots$ , pode ser escrita como:

$$P(S_j = s) = (-1)^s \delta_0 \delta_1 \dots \delta_{s-1} \sum_{i=0}^s \frac{e^{-\lambda \delta_i}}{\prod_{\substack{j=0 \\ j \neq i}}^s (\delta_i - \delta_j)}, \tag{8}$$

com  $\delta_0 = 1$ . A média e a variância de  $S_j$  não serão apresentadas para o caso geral embora sejam facilmente estimadas pois, na prática, é muito raro aparecerem dados com número de ovos superior a quatro por hospedeiro. Casos particulares são obtidos, através de suposições feitas sobre os parâmetros  $\delta_j$ .

**M1. Caso 1 -  $\delta_i = \delta, i \geq 1$**

No primeiro modelo, Bakker (1972) supõe que a fêmea só é capaz de distinguir se o hospedeiro  $j$  foi parasitado, ou não, mas não distingue quantos ovos estão presentes. Dessa forma, tem-se que  $\delta_i = \delta, i \geq 1$ , isto é, as probabilidades de ovoposição, após o primeiro encontro, são constantes. Portanto, há dois parâmetros a serem estimados,  $\lambda$  e  $\delta$ . Assim, a distribuição de  $S_j$ , o número total de ovos postos em um hospedeiro  $j$ , é dada por  $p_0 = e^{-\lambda}$ , para  $s = 0$ , e para  $s \geq 1$  por

$$P(S_j = s) = \delta^{s-1} \left[ \frac{(-1)^s e^{-\lambda}}{(1-\delta)^s} + \sum_{i=0}^s (-1)^{s-i} \frac{e^{-\lambda\delta} \lambda^{s-i}}{(1-\delta)^i (s-i)!} \right], \quad (9)$$

com uma expressão alternativa dada por:

$$\begin{aligned} p_s &= \delta^{s-1} (\delta-1)^{-s} e^{-\lambda\delta} \sum_{i=s}^{\infty} \frac{\lambda^i (\delta-1)^i}{i!} \\ &= \delta^{s-1} (\delta-1)^{-s} \left[ e^{-\lambda} - e^{-\lambda\delta} \sum_{i=0}^{s-1} \frac{\lambda^i (\delta-1)^i}{i!} \right] \\ &= \sum_{k=0}^{\infty} q_{s+k} \binom{s+k-1}{s} \delta^{s-1} (1-\delta)^k, \end{aligned} \quad (10)$$

em que  $q_r$  é definida em (1).

Usando-se (6), a  $P(S_j = 2)$ , para o caso em que  $\delta_1 = \delta_2 = \delta$  e, portanto,  $a = \left[ \frac{(1-\delta_1)}{(1-\delta_2)} \right] = 1$  e  $\sum_{k=2}^i a^k = (i-1)$ , é dada por:

$$\begin{aligned} P(S_j = 2) &= \frac{\delta e^{-\lambda}}{(1-\delta)^2} \sum_{i=2}^{\infty} \frac{[\lambda(1-\delta)]^i}{i!} \sum_{k=2}^i a^k \\ &= \frac{\delta e^{-\lambda}}{(1-\delta)^2} \left[ \sum_{i=2}^{\infty} \frac{[\lambda(1-\delta)]^i}{(i-1)!} - \sum_{i=2}^{\infty} \frac{[\lambda(1-\delta)]^i}{i!} \right] \\ &= \frac{\delta e^{-\lambda}}{(1-\delta_1)^2} \left[ \lambda(1-\delta) \sum_{i=2}^{\infty} \frac{[\lambda(1-\delta)]^{i-1}}{(i-1)!} - \sum_{i=2}^{\infty} \frac{[\lambda(1-\delta)]^i}{i!} \right] \end{aligned}$$

$$\begin{aligned}
&= \frac{\delta e^{-\lambda}}{(1-\delta)^2} \left[ \lambda(1-\delta) \sum_{i=1}^{\infty} \frac{[\lambda(1-\delta)]^i}{i!} - \sum_{i=2}^{\infty} \frac{[\lambda(1-\delta)]^i}{i!} \right] \\
&= \frac{\delta e^{-\lambda}}{(1-\delta)^2} [\lambda(1-\delta)e^{\lambda(1-\delta)} - e^{\lambda(1-\delta)} + 1] \\
&= \delta \left[ \frac{\lambda e^{-\lambda\delta}}{1-\delta} - \frac{e^{-\lambda\delta}}{(1-\delta)^2} + \frac{e^{-\lambda}}{(1-\delta)^2} \right], \tag{11}
\end{aligned}$$

que é a mesma expressão que se obtém ao fazer  $s = 2$  em (9).

A média e a variância de  $S_j$ , o número de ovos por hospedeiro, são dadas por:

$$\mu = E(S_j) = \lambda\delta + (1 - e^{-\lambda})(1 - \delta) = \frac{E(R_j)}{H} \tag{12}$$

e

$$\begin{aligned}
Var(S_j) &= \lambda^2\delta^2 - \mu^2 + (1 - e^{-\lambda})(1 - \delta)(1 - 2\delta) + \lambda\delta(3 - 2\delta) \\
&= \lambda\delta(\lambda\delta + 2) + E(S_j)(1 - 2\delta - E(S_j)) \\
&= \frac{var(R_j)}{H}, \tag{13}
\end{aligned}$$

em que  $R_j$  é o número de encontros experimentados pelo hospedeiro  $j$  e  $H$  é o número de hospedeiros.

Na prática, segundo os autores, quando a média do número de ovos for superior a 1,1 ovos por hospedeiro, esse modelo não descreve satisfatoriamente a distribuição do número de ovos.

## M2. Caso 2 - $\delta_i = \delta_2$ , $i \geq 2$

No segundo modelo, Bakker (1972) supõe que a fêmea é capaz de distinguir se o hospedeiro foi parasitado, ou não, e, além disso, se um ovo está presente ou estão dois ou mais. Dessa forma, tem-se que a probabilidade  $\delta_1$  de um parasita colocar um ovo em um hospedeiro que já tenha um ovo é arbitrária enquanto que porá um ovo em um hospedeiro com pelo menos dois ovos com probabilidade  $\delta_i = \delta_2$ ,  $i \geq 2$ , com  $\delta_2 < \delta_1$ . Portanto, há três parâmetros a serem estimados:  $\lambda$ ,  $\delta_1$  e  $\delta_2$ . As probabilidades de que haja 0, 1 ou 2 ovos são dadas pelas fórmulas (2), (4) e (7), respectivamente. A probabilidade de que haja 3 ovos é dada por:

$$\begin{aligned}
P(S_j = 3) &= \delta_1 \delta_2 \left\{ e^{-\lambda \delta_2} \left[ \frac{\lambda}{(1 - \delta_2)(\delta_1 - \delta_2)} + \frac{2\delta_2 - \delta_1 - 1}{(1 - \delta_2)^2(\delta_1 - \delta_2)^2} \right] + \right. \\
&\quad \left. + \frac{e^{-\lambda \delta_1}}{(1 - \delta_1)(\delta_2 - 1)^2} + \frac{e^{-\lambda}}{(\delta_1 - 1)(\delta_2 - 1)^2} \right\},
\end{aligned} \tag{14}$$

enquanto que para 4 ovos a probabilidade é dada por:

$$\begin{aligned}
P(S_j = 4) &= \delta_1 \delta_2^2 \left\{ e^{-\lambda \delta_2} \left[ \frac{\lambda^2/2}{(1 - \delta_2)(\delta_1 - \delta_2)} + \frac{\lambda(2\delta_2 - \delta_1 - 1)}{(1 - \delta_2)^2(\delta_1 - \delta_2)^2} + \right. \right. \\
&\quad \left. \left. + \frac{\delta_1^2 + 3\delta_2^2 - 3\delta_1 \delta_2 + \delta_1 - 3\delta_2 + 1}{(1 - \delta_2)^3(\delta_1 - \delta_2)^3} \right] + \right. \\
&\quad \left. + \frac{e^{-\lambda \delta_1}}{(1 - \delta_1)(\delta_2 - 1)^3} + \frac{e^{-\lambda}}{(\delta_1 - 1)(\delta_2 - 1)^3} \right\}.
\end{aligned} \tag{15}$$

Expressões explícitas para 5 ou mais ovos não serão apresentadas, pois dificilmente são usadas na prática.

### M3. Caso 3 - $\delta_i = \frac{\delta}{i}$

No terceiro modelo, Bakker (1972) supõe que a fêmea é capaz de distinguir se o hospedeiro foi parasitado, ou não, e, além disso, a probabilidade de ovoposição é inversamente proporcional ao número de ovos  $i$  presentes no hospedeiro. Dessa forma, tem-se que  $\delta_i = \frac{\delta}{i}$ ,  $i = 1, 2, \dots$ . Portanto, há dois parâmetros a serem estimados,  $\lambda$  e  $\delta$ . A distribuição de probabilidade de  $S_j$  é obtida fazendo a substituição de  $\delta_i$  por  $\frac{\delta}{i}$ , na expressão (8), isto é,

$$P(S_j = s) = \frac{(-1)^s}{(s-1)!} \sum_{i=0}^s \frac{e^{-\lambda \frac{\delta}{i}}}{\prod_{\substack{j=1 \\ j \neq i}}^s \left( \frac{1}{i} - \frac{1}{j} \right)}. \tag{16}$$

#### M4. Caso 4 - $\delta_i = \delta^i$

No quarto modelo, Bakker (1972) supõe que a fêmea é capaz de distinguir se o hospedeiro foi parasitado, ou não, e, além disso, a probabilidade de ovoposição decresce geometricamente com o número de ovos. Dessa forma, tem-se que a probabilidade  $\delta_i = \delta^i$ ,  $i = 0, 1, \dots$ . A distribuição de probabilidade de  $S_j$  é obtida pela substituição de  $\delta_i$  por  $\delta^i$ ,  $i = 1, 2, \dots$  na expressão (8), isto é,

$$P(S_j = s) = (-1)^s \delta^{\frac{(s-1)s}{2}} \sum_{i=0}^s \frac{e^{-\lambda \delta^i}}{\prod_{\substack{j=0 \\ j \neq i}}^s (\delta^i - \delta^j)}. \quad (17)$$

#### 2.1.2 Modelo de Rogers - M5

Uma variação do modelo M1 de Bakker, motivada pelo fato que a distribuição de Poisson não é apropriada para pequenas amostras, foi proposta por Rogers (1975). Ele supõe que o número  $R_j$  de encontros experimentados por um particular hospedeiro, dado que houve  $R = r_0$  encontros entre os parasitóides e os  $H$  hospedeiros, segue uma distribuição binomial “condicional” e, portanto,

$$P(R_j = r) = q_r = \binom{r_0}{r} p^r (1-p)^{r_0-r} I_{\{0,1,\dots,r_0\}}(r)$$

em que  $p = \frac{1}{H}$  é a probabilidade de que um parasitóide escolha esse particular hospedeiro. Diferentemente do usual,  $p$  é conhecido e  $r_0$  é um parâmetro desconhecido.

Assim, a função geradora de probabilidade (f.g.p)(ver apêndice A) do número de encontros é dada por

$$G_{R_j}(s) = \left[ \frac{H-1}{H} + \frac{s}{H} \right]^{r_0}$$

com média  $\nu = E(R_j) = \frac{r_0}{H}$ . Usando as probabilidades de ovoposição  $\delta_0 = 1$  e  $\delta_i = \delta$ ,  $i = 1, 2, \dots$ , como no modelo M1 de Bakker, tem-se que a f.g.p. de  $S_j$ , o número total de ovos postos em um hospedeiro, é dada por:

$$G_{S_j}(s) = \frac{s(1 - \delta p + s\delta p)^{r_0} + (1 - \delta)(1 - s)(1 - p)^{r_0}}{1 - \delta + \delta s}, \quad (18)$$

e a probabilidade,  $p_s = P(S_j = s)$ , de que um dado hospedeiro  $j$  contenha  $s$  ovos é  $p_0 = (1 - \frac{1}{H})^{r_0} = q_0$ , para  $s = 0$ , e para  $s \geq 1$  é

$$p_s = \sum_{k=0}^{\infty} q_{s+k} \binom{s+k-1}{s} \delta^{s-1} (1-\delta)^k, \quad (19)$$

em que  $q_{s+k} = P(R_j = s+k)$ .

A média e a variância de  $S_j$ , o número de ovos por hospedeiro, são:

$$\mu = E(S_j) = \delta \frac{r_0}{H} + (1-\delta) [1 - (1 - \frac{1}{H})^{r_0}] \quad (20)$$

e

$$Var(S_j) = \frac{r_0(r_0-1)\delta^2}{H^2} + \frac{r_0\delta}{H}(3-2\delta) + (1-\delta)(1-2\delta)(1 - (1 - \frac{1}{H})^{r_0}) - \mu^2. \quad (21)$$

### 2.1.3 Modelo de Griffiths - M6

Estudos sobre superparasitismo de *Spalangia drosophilae* em pupas de *Drosophila melanogaster* foram feitos por Simmonds (1956). Ele observou que somente um indivíduo chegava à idade adulta desenvolvendo-se e que somente um ovo é posto em cada ovoposição. Os dados da Tabela 2 referem-se a um ensaio em que 5 fêmeas de *Spalangia* foram deixadas com 25 hospedeiros durante 24 horas e foi contado o número de pupas contendo  $y$  ovos. Verifica-se que a distribuição de Poisson não se ajusta a esse conjunto de dados.

Tabela 2 - Distribuição do número de ovos ( $Y$ ) de *Spalangia drosophilae* em pupas de *Drosophila melanogaster*

Frequência	Número de ovos por pupa							Total
	0	1	2	3	4	5	> 5	
Observada	19	169	50	6	1	1	0	246
Poisson(1,20)	73,85	88,86	53,46	21,44	6,45	1,55	0,39	246

Griffiths (1977a, b) ajustou a esses dados os modelos M1 de Bakker e M5 de Rogers e mostrou que eles são casos particulares de um modelo mais geral em que o número de



ovos postos no primeiro encontro,  $W_1$ , tem uma distribuição com f.g.p.  $G_1(s)$  e a distribuição do número de ovos postos nos encontros subsequentes ( $W_i, i > 1$ ) tem uma distribuição com f.g.p.  $G_2(s)$ . Além disso, o número total de encontros de um dado hospedeiro,  $R_j$ , é uma variável aleatória com uma distribuição com f.g.p.  $G_{R_j}(s)$  e função de probabilidade (f.p.)  $q_r = P(R_j = r), r = 0, 1, 2, \dots$ . Logo, o número total de ovos postos em um hospedeiro  $j$  é  $S_j = \sum_{i=1}^{R_j} W_i$  com f.g.p.  $G_{S_j}(s)$  dada por:

1.  $G_{S_j}(s) = G_{R_j}(G_1(s))$ , se não houver o comportamento de evitar o superparasitismo e
2. em caso contrário, por

$$G_{S_j}(s) = G_1(s) \frac{G_{R_j}[G_2(s)]}{G_2(s)} + q_0 \left[ 1 - \frac{G_1(s)}{G_2(s)} \right]. \quad (22)$$

Para verificar isso, suponha que cada hospedeiro tenha um número aleatório,  $R_j$ , de encontros. Em cada encontro o parasitóide deposita  $W_i$  ovos,  $i = 1, 2, \dots$ . A quantidade total de ovos postos,  $S_j$ , é definida por:

$$S_j = \begin{cases} 0, & \text{se } R_j = 0; \\ W_1 + W_2 + \dots + W_{R_j}, & \text{se } R_j = 1, 2, \dots \end{cases}$$

Assim

$$\begin{aligned} G_{S_j}(s) &= E(S_j^s) = E[E(S_j^s | R_j)] \\ &= \sum_{r=0}^{\infty} E(S_j^s | R_j = r) P(R_j = r) = \sum_{r=0}^{\infty} G_{S_j | R_j=r}(s) P(R_j = r). \end{aligned} \quad (23)$$

Mas,  $G_{S_j | R_j=0}(s) = 1$  e, para  $r \geq 1$ , tem-se

$$\begin{aligned} G_{S_j | R_j=r}(s) &= E(s^{W_1+W_2+\dots+W_r}) = E(s^{W_1}) E(s^{W_2+\dots+W_r}) \\ &= G_{W_1}(s) \prod_{i=2}^r G_{W_i}(s) = G_{W_1}(s) \prod_{i=2}^r G_{W_2}(s) = G_1(s) [G_2(s)]^{r-1} \end{aligned}$$

e, portanto,

$$\begin{aligned}
G_{S_j}(s) &= q_0 + \sum_{r=1}^{\infty} G_1(s)[G_2(s)]^{r-1}q_r \\
&= q_0 + \frac{G_1(s)}{G_2(s)} \sum_{r=1}^{\infty} [G_2(s)]^r q_r \\
&= q_0 + \frac{G_1(s)}{G_2(s)} \left[ \sum_{r=0}^{\infty} [G_2(s)]^r q_r - q_0 \right] \\
&= q_0 + \frac{G_1(s)}{G_2(s)} [G_{R_j}(G_2(s)) - q_0] \\
&= G_1(s) \frac{G_{R_j}(G_2(s))}{G_2(s)} + q_0 \left[ 1 - \frac{G_1(s)}{G_2(s)} \right].
\end{aligned} \tag{24}$$

O modelo, agora, é específico para o caso em que no máximo um ovo é colocado em cada encontro, como nos modelos de Bakker (1967) e Rogers (1975). Considere um caso particular do modelo de Griffiths em que  $W_1 \sim \text{Bernoulli}(1)$ , isto é, sempre no primeiro encontro um ovo é posto. Tem-se que  $W_i \sim \text{Bernoulli}(\delta)$ , significando que o parasitóide põe um ovo com probabilidade  $\delta$  ou evitará o superparasitismo com probabilidade  $(1 - \delta)$ . Assim, tem-se:

$$G_1(s) = E(s^{W_1}) = s^0 P(W_1 = 0) + s^1 P(W_1 = 1) = s$$

e

$$G_2(s) = E(s^{W_i}) = s^0 P(W_i = 0) + s^1 P(W_i = 1) = 1 - \delta + \delta s$$

e portanto, a f.g.p. para o número total de ovos colocados  $S_j$ , dada por (24) ficará

$$G_{S_j}(s) = \frac{sG_{R_j}(1 - \delta + \delta s)}{1 - \delta + \delta s} + q_0 \left[ 1 - \frac{s}{1 - \delta + \delta s} \right]. \tag{25}$$

Para se encontrar a probabilidade de que nenhum ovo tenha sido posto, substitui-se  $s = 0$  em (25), obtendo-se

$$G_{S_j}(0) = q_0 = p_0.$$

Para se obterem as demais probabilidades, basta derivar  $G_{S_j}(s)$  em relação a  $s$  e fazer  $s = 0$ , isto é,

$$G'_{S_j}(0) = \frac{G(1 - \delta) - q_0}{1 - \delta} = p_1,$$

$$G''_{S_j}(0) = -2 \frac{\delta[G'(1-\delta) + G'(1-\delta)\delta + G(1-\delta) - q_0]}{(\delta-1)^2} = p_2$$

e assim por diante.

Para se obterem os momentos fatoriais, deriva-se  $G_{S_j}(s)$  em relação a  $s$  e faz-se  $s = 1$ . Logo,

$$G'_{S_j}(1) = G(1) + \delta G'(1) - \delta G(1) - q_0 - q_0\delta = 1 + \delta E(N) - \delta - q_0(1-\delta)$$

em que  $\nu = E(R_j)$  é o número médio de encontros e, portanto,

$$\mu_{[1]} = \mu = E(S_{R_j}) = \nu\delta + (1-\delta)(1-q_0).$$

Os demais momentos fatoriais são obtidos por derivadas de ordens superiores e fazendo-se alguns cálculos algébricos tem-se,

$$G''_{S_j}(1) = \nu_{[2]}\delta^2 + 2\delta(1-\delta) \left[ -(1-q_0) + \frac{\nu_{[2]}}{2} \right] = \mu_{[2]},$$

$$G'''_{S_j}(1) = \nu_{[3]}\delta^3 + 6\delta^2(1-\delta) \left[ (1-q_0) - \frac{\nu_{[2]}}{2} + \frac{\nu_{[3]}}{6} \right] = \mu_{[3]}$$

e assim por diante.

Sejam  $\mu_{[r]}$  e  $\nu_{[r]}$  os momentos fatoriais das distribuições de  $S_{R_j}$  e  $R_j$ , respectivamente. Então, vê-se que

$$\mu_{[r]} = \nu_{[r]}\delta^r + \delta^{r-1}(1-\delta)r! \left[ (-1)^{r+1}(1-q_0) + \sum_{i=1}^{r-1} (-1)^{r+1-i} \frac{\nu_{[i]}}{i!} \right].$$

Considere agora o Modelo de Bakker em que a distribuição do número de encontros é Poisson com média  $\lambda$ . Substituindo  $q_0 = e^{-\lambda}$  e  $G_{R_j} = e^{\lambda(s-1)}$  na equação (25) tem-se:

$$G_{S_{R_j}}(s) = \frac{se^{-\lambda\delta(1-s)} + (1-\delta)(1-s)e^{-\lambda}}{1-\delta+\delta s}, \quad (26)$$

que é a f.g.p. da variável aleatória do modelo M1 de Bakker.

### 2.1.4 Modelos de Hemerik - M7 a M13

Sete modelos, combinando (i.) modos diferentes de busca do parasita pelo hospedeiro e (ii.) habilidade do parasitóide em reconhecer a qualidade do hospedeiro, foram propostos por Hemerik et al. (2002). Os autores supuseram que os parasitóides procuram os hospedeiros aleatoriamente ou reagindo a vibrações causadas pelo hospedeiro (conceito de vibrotaxis), levando, respectivamente, às distribuições Poisson e binomial negativa para o número  $R_j$  de encontros. Supuseram, ainda, que os parasitóides: (i.) tratam todos os hospedeiros igualmente, (ii.) são capazes de distinguir entre hospedeiros parasitados, ou não, (iii.) estão aptos a perceber o número exato de ovos de cada hospedeiro e (iv.) reconhecem seus próprios ovos.

Os quatro primeiros modelos de Hemerik et al. (2002) foram construídos sob a suposição que o número  $R_j$  de encontros tem distribuição Poisson e os  $K$  parasitóides escolhem os hospedeiros aleatoriamente (não se evidenciando qualquer preferência ou rejeição) com a mesma eficiência de procura  $\gamma$ .

#### M7. Caso 1 - Os parasitóides tratam todos os hospedeiros igualmente.

Os parasitóides não distinguem se os hospedeiros estão, ou não, parasitados. Portanto, a probabilidade de parasitar um hospedeiro contendo  $s$  ovos,  $\delta_s$ , no encontro é independente de  $s$  para  $s \geq 0$ . Nesse caso, supõe-se que  $\delta_s = 1$ . Logo, o número  $S_{R_j}$  de ovos postos, após um tempo  $T$ , tem distribuição Poisson com parâmetro  $\lambda = \gamma KT$ , isto é,

$$P(S_{R_j} = s) = \frac{(\gamma KT)^s e^{-\gamma KT}}{s!} I_{\{0,1,\dots\}}(s). \quad (27)$$

#### M8. Caso 2 - Os parasitóides são capazes de distinguir entre hospedeiros parasitados, ou não.

Hospedeiros não parasitados são parasitados com certeza, isto é, com probabilidade  $\delta_0 = 1$ , enquanto que para um encontro com um hospedeiro contendo  $s \geq 1$  ovos, existe uma probabilidade  $0 < \delta_s = \delta < 1$  de haver a ovoposição, depois do encontro. Logo, a distribuição do número  $S_j$  de ovos postos nos hospedeiros tem dois parâmetros,  $\delta$  e  $\lambda = \gamma KT$ , e é dada por:

$$P(S_j = s) = e^{-\gamma KT} I_{\{0\}}(s) + \delta^{s-1} \left[ \frac{(-1)^s e^{-\gamma KT}}{(1-\delta)^s} + e^{-\delta \gamma KT} \sum_{i=1}^s \frac{(-1)^{i-1} (\gamma KT)^{s-i}}{(1-\delta)^i (s-i)!} \right] I_{\{1,2,\dots\}}(s) \quad (28)$$

A distribuição limite (convergência em Distribuição) quando  $\delta \rightarrow 1$  é Poisson  $\sim (\lambda = \gamma KT)$ .

**M9. Caso 3 - Os parasitóides estão aptos a perceber o número exato de ovos de cada hospedeiro.**

Nesse caso, supõe-se que a probabilidade de o parasita ovopositar, em um dado encontro, em um hospedeiro depende do número  $s$  de ovos já postos nesse hospedeiro. Hospedeiros não parasitados são parasitados com certeza, isto é, com probabilidade  $\delta_0 = 1$ , enquanto que para um encontro com um hospedeiro contendo  $s \geq 1$  ovos, existe uma probabilidade  $0 < \delta_s < 1$  de haver a ovoposição, depois do encontro. Além disso, supõe-se que  $\delta_s$  é uma função decrescente de  $s$ . Se um hospedeiro contém poucos ovos, a probabilidade de cada ovo produzir um inseto adulto é maior do que em um hospedeiro com muitos ovos. Dessa maneira,  $\delta_s < \delta_{s-1}$  para todo  $s \geq 1$ . Considere  $\delta_s = \delta^s$ , ( $0 < \delta < 1$ ). Logo, a distribuição do número  $S_j$  de ovos postos nos hospedeiros tem dois parâmetros,  $\delta$  e  $\lambda = \gamma KT$ , e é dada por:

$$P(S_j = s) = \delta^{s(s-1)/2} \sum_{j=0}^s \frac{e^{-\gamma KT \delta^j}}{s \prod_{\substack{i=0 \\ i \neq j}}^s (\delta^i - \delta^j)} I_{\{0,1,\dots\}}(s). \quad (29)$$

A distribuição limite (convergência em Distribuição) quando  $\delta \rightarrow 1$  é Poisson  $\sim (\lambda = \gamma KT)$ .

**M10. Caso 4 - Os parasitóides reconhecem seus próprios ovos.**

Nesse caso, o parasitóide está apto a reconhecer hospedeiros saudáveis entre os parasitados e, também, perceber se algum dos ovos do hospedeiro é seu, ou não. Sejam  $\delta_s^0$  e  $\delta_s^1$  as probabilidades de ovoposição de um hospedeiro que tenha  $s$  ovos se, respectivamente, nenhum ou um ovo do parasitóide está presente. Supõe-se, ainda, que  $\delta_0^0 = 1$ ,  $\delta_s^0 = \beta$  e  $\delta_s^1 = 0$ , ( $s \geq 1, 0 < \beta < 1$ ), isto é, os parasitóides sempre refugam o hospedeiro quando reconhecem seus próprios ovos. Logo, a distribuição do número  $s$  de ovos postos nos hospedeiros tem dois parâmetros,  $\beta$  e  $\lambda = \gamma KT$ , e é dada por:

$$P(S_j = s) = \begin{cases} e^{-\gamma KT}, & \text{se } s = 0; \\ \binom{K-1}{s-1} K e^{-\gamma KT} \sum_{j=0}^{s-1} \binom{s-1}{j} (-1)^j \frac{e^{-\gamma T[\beta(s-j)+K(1-\beta)]} - 1}{\beta(s-j) + K(1-\beta)}, & \text{se } s = 1, 2, \dots, K. \end{cases}$$

Os três últimos modelos de Hemerik et al. (2002) foram construídos sob a suposição de que o número  $R$  de encontros tem distribuição binomial negativa e os  $K$  parasitóides escolhem os hospedeiros aleatoriamente e reagem a alguma característica deles, por exemplo, ao tamanho ou a movimentos mandibulares (conceito de vibrotaxis), com eficiência de procura  $m\gamma$ , sendo esse tempo de escolha descrito por uma variável aleatória que tem distribuição gama com parâmetro de agrupamento  $m$  e valor esperado  $\gamma KT$ .

**M11. Caso 5 - Os parasitóides tratam todos os hospedeiros igualmente.**

Os parasitóides não distinguem se os hospedeiros estão, ou não, parasitados. Portanto, a probabilidade de parasitar um hospedeiro contendo  $s$  ovos,  $\delta_s$ , depois do encontro é independente de  $s$  para  $s \geq 0$ . Nesse caso, supõe-se que  $\delta_s = 1$ . Logo, a distribuição do número  $S_j$  de ovos postos nos hospedeiros, durante o período  $T$ , é Binomial Negativa com parâmetros  $m$  e  $\lambda = \gamma KT$ , e é dada por:

$$P(S_j = s) = \frac{\Gamma(s + m)}{s! \Gamma(m)} \left( \frac{m}{m + m\gamma KT} \right)^m \left( \frac{m\gamma KT}{m + m\gamma KT} \right)^s I_{\{0,1,\dots\}}(s). \quad (30)$$

Esse modelo reduz-se ao modelo M7 se  $\gamma KT$  é fixado e  $m \rightarrow \infty$ .

**M12. Caso 6 - Os parasitóides são capazes de distinguir entre hospedeiros parasitados, ou não.**

Hospedeiros não parasitados são parasitados com certeza, isto é, com probabilidade  $\delta_0 = 1$ , enquanto que para um encontro com um hospedeiro contendo  $s \geq 1$  ovos, existe uma probabilidade  $0 < \delta_s = \delta < 1$  de haver a ovoposição, depois do encontro. Logo, a distribuição do número  $S_j$  de ovos postos nos hospedeiros, durante o período  $T$ , tem três parâmetros  $\delta$ ,  $m$  e  $\lambda = \gamma KT$ , e sua função de probabilidade (f.p.) é dada por:

$$P(S_j = s) = \left( \frac{m}{m + m\gamma KT} \right)^m I_{\{0\}}(s) + \left[ \sum_{i=1}^s \frac{\Gamma(m + s - i)(-\delta)^{i-1}}{(s - i)! \Gamma(m)(1 - \delta)^i} \left( \frac{m}{m + m\delta\gamma KT} \right)^m \right. \\ \left. \left( \frac{m\delta\gamma KT}{m + m\delta\gamma KT} \right)^{s-i} + (-1)^s \frac{\delta^{s-1}}{(1 - \delta)^s} \left( \frac{m}{m + m\gamma KT} \right)^m \right] I_{\{1,2,\dots\}}(s). \quad (31)$$

Quando  $\delta \rightarrow 1$  a distribuição tende (convergência em Distribuição) para uma binomial negativa com parâmetros  $m$  e  $\gamma KT$ .

**M13. Caso 7 - Os parasitóides estão aptos a perceber o número exato de ovos de cada hospedeiro.**

Nesse caso, supõe-se que a probabilidade de o parasita ovopositar em um dado encontro com um hospedeiro depende do número  $s$  de ovos já postos nesse hospedeiro. Hospedeiros não parasitados são parasitados com certeza, isto é, com probabilidade  $\delta_0 = 1$ , enquanto que para um encontro com um hospedeiro contendo  $s \geq 1$  ovos, existe uma probabilidade  $0 < \delta_s < 1$  de haver a ovoposição, depois do encontro. Além disso, supõe-se que  $\delta_s$  é uma função geometricamente decrescente de  $s$ , isto é,  $\delta_s = \delta^s$ , ( $0 < \delta < 1$ ). Logo, a distribuição do número  $S_j$  de ovos postos nos hospedeiros tem três parâmetros,  $\delta$ ,  $m$  e  $\lambda = \gamma KT$ , e é dada por:

$$P(S_j = s) = \left( \frac{m}{m + m\gamma KT} \right)^m I_{\{0\}}(s) + \delta^{s(s-1)/2} \sum_{j=0}^s \frac{\left( \frac{m}{m + m\delta^j \gamma KT} \right)^m}{\prod_{\substack{i=0 \\ j \neq i}}^s (\delta^i - \delta^j)} I_{\{1,2,\dots\}}(s). \quad (32)$$

Quando  $\delta \rightarrow 1$  a distribuição tende (convergência em Distribuição) para uma binomial negativa com parâmetros  $m$  e  $\gamma KT$ .

Os autores não discutiram o modelo em que o número de encontros  $R_j$  tem distribuição binomial negativa e os parasitas reconhecem seus próprios ovos. Para os demais, há estudos comparativos através de simulação. Comentaram, ainda, que o modelo de Bakker et al. (1972) não é totalmente confiável.

Através de simulação e usando o critério de Akaike, Hemerik e van der Hoven (2003) mostraram que os modelos  $M1$ ,  $M4$  e  $M7$  não são identificáveis quando aplicados a um conjunto de dados.

### 2.1.5 Modelo de Janardan - M14

Estudos de comportamento das técnicas de ovoposição do caruncho em feijão (tipo para produção de brotos) foram feitos em laboratório por Mitchel (1975). Observou que: (i.) as fêmeas dos gorgulhos põem seus ovos sobre os feijões e que após 8 a 10 dias as larvas saem e entram no feijão; (ii.) os carunchos são seletivos para o tamanho e a qualidade do feijão sobre o qual porá seus ovos, pois a larva não pode se mover de um feijão para outro,

alimentando-se, crescendo e maturando no feijão escolhido por sua mãe; (iii.) os gorgulhos preferem feijões sem ovos, pois a chance de sobrevivência da larva decresce com o aumento do número de ovos sobre o feijão e, portanto, encontrou poucos feijões com dois ou mais ovos. A Tabela 3 refere-se ao resultado de 3 ensaios para estudo da distribuição de *Callosobraccus maculatus* sobre feijões do tipo “mung” (Consul, 1989, pág.123). Concluiu, então, que uma distribuição de Poisson não se ajustava ao número de ovos por feijão.

Tabela 3 - Distribuição do número de ovos do gorgulho em feijão “mung”

		Número de ovos por feijão					
Ensaio		0	1	2	3	> 3	Total
1	Observado	138	46	0	0	0	184
	Poisson(0,25)	143,30	35,82	4,48	0,37	0,03	184
2	Observado	26	117	35	1	0	179
	Poisson(1,06)	61,92	65,73	34,88	12,34	4,13	179
3	Observado	5	68	88	32	0	193
	Poisson(1,76)	33,15	58,40	51,44	30,20	19,81	193

Motivados pelo fato de que o processo biológico do experimento de Mitchell não era um processo de Poisson homogêneo, Janardan et al. (1979) (ver Apêndice B) usaram a distribuição generalizada de Poisson (Consul, 1989), que é uma distribuição discreta associada a um processo estocástico de nascimento e morte, para descrever o número de ovos postos em cada feijão. Entretanto, Janardan (1980) frisou que essa variável representa apenas um processo de nascimento e desenvolveu um modelo assumindo uma taxa  $\delta = \lambda$  para a ovoposição em um hospedeiro não parasitado e uma taxa  $\delta = \mu$  para o hospedeiro já parasitado. Esse modelo foi usado, independentemente, por Morgan (1982) que discutiu uma classe de distribuições modificadas para um estudo da polispermia em ovos de ouriços-do-mar que incluía o modelo de Bakker (M1) como caso particular.

Um modelo estocástico alternativo foi proposto por Janardan et al. (1997) (ver Apêndice B) para descrever o número  $X(t)$  de ovos postos no tempo  $0 \leq t \leq T$ ,



incorporando o fato que a mudança da taxa descrita pode ocorrer depois de um ovo ser posto (probabilidade  $p$ ) ou depois de dois ovos serem postos (probabilidade  $q = 1 - p$ ), através de uma variável aleatória binária  $Y$  que assume, respectivamente, os valores  $m = 1$  ou  $m = 2$ . Além disso, supõe-se que haja independência entre  $X(t)$  e  $Y$  e que os seguintes postulados sejam verdadeiros:

- a.  $P(X(t+h) = k+1 | X(t)=k, Y = m) = \lambda h + o(h), 0 \leq k \leq m-1$ ;
- b.  $P(X(t+h) = k | X(t)=k, Y = m) = 1 - \lambda h + o(h), 0 \leq k \leq m-1$ ;
- c.  $P(X(t+h) = k+1 | X(t)=k, Y = m) = \mu h + o(h), k \geq m$ ;
- d.  $P(X(t+h) = k | X(t)=k, Y = m) = 1 - \mu h + o(h), k \geq m$ ;
- e.  $P(X(t+h) = k+2 | X(t)=k, Y = m) = o(h), k \geq 0, m = 1, 2, \dots$

Seja  $P_n(t) = P(X(t) = n)$ ,  $n \geq 0$ . A f.g.p. de  $X(t)$ ,  $G(z, t)$  de  $X(t)$  é dada por:

$$G(z, t) = \sum_{n=0}^{\infty} P_n(t) z^n = \frac{(z-1)(\beta p + z\lambda q)}{p(\mu z - \beta)} [e^{-\lambda t} - e^{-(1-z)\mu t}] - \frac{z(z-1)\lambda q}{p(\mu z - \beta q)} [e^{-\alpha t} - e^{-(1-z)\mu t}] + e^{-(1-z)\mu t}, \quad (33)$$

em que  $\beta = \mu - \lambda$ .

Sua função de probabilidade é dada por:

$$P_n(t) = e^{-\lambda t} I_{\{0\}}(n) + \frac{\lambda}{p\beta} [e^{-\lambda t} - e^{-\alpha t}] I_{\{1\}}(n) + \frac{\alpha \lambda \mu^{n-2} e^{-\mu t}}{p q^{n-1} \beta^n} [q^{n-1} (e^{\beta t} - u(t)) - (e^{q\beta t} - v(t))] I_{\{2,3,\dots,\infty\}}(n), \quad (34)$$

em que  $\alpha = \lambda q + \mu p$ ,  $u(t) = \sum_{j=0}^{n-2} \frac{(\beta t)^j}{j!}$  e  $v(t) = \sum_{j=0}^{n-2} \frac{(q\beta t)^j}{j!}$ .

Além disso,

$$m(t) = E[X(t)] = \mu t - \beta \left\{ \frac{1 - e^{-\lambda t}}{\lambda} - \frac{\lambda q}{\beta p} \left[ \frac{1 - e^{-\alpha t}}{\alpha} - \frac{1 - e^{-\lambda t}}{\lambda} \right] \right\}.$$

Para calcular a  $V[X(t)] = E[(X(t))^2] - [m(t)]^2 = m_2(t) - [m(t)]^2$ , o segundo momento de  $X(t)$  em relação à origem é dado por:

$$m_2(t) = \frac{[\alpha\mu p(\lambda - 2\beta) + 2\mu\lambda^2 q(\alpha - 1)] t}{\alpha\lambda p} + \mu^2 t^2 + \frac{\beta e^{-\lambda t}}{\lambda} + \quad (35)$$

$$+ \frac{1}{\lambda p} [2\mu(\beta p - \lambda^2 q) + 2\mu\lambda^2(\alpha - 1)] \left[ \frac{1 - e^{-\lambda t}}{\lambda} \right] + \quad (36)$$

$$+ \frac{\lambda q(3 - 2\mu)}{p} \left[ \frac{1 - e^{-\alpha t}}{\alpha} \right] \quad (37)$$

Quando  $\mu \rightarrow \lambda$  o modelo de Janardan reduz-se ao modelo de Poisson homogêneo de taxa  $\lambda$  e quando  $\mu < \lambda$ , a razão  $\mu/\lambda$  é interpretada como a probabilidade de repulsa.

### 2.1.6 Modelos de Daley e Maindonald - M15 a M18

Modelos estocásticos foram propostos por Daley e Maindonald (1989) para explicar a repulsa ao superparasitismo, numa primeira tentativa de unificação dos modelos existentes na literatura. Os autores mostraram que os modelos podem ser formulados através do comportamento temporal de um processo markoviano (discreto ou contínuo) e introduziram uma variante do modelo seqüencial de ocupação que origina uma nova versão dos modelos de Bakker, conforme mostrado a seguir.

#### M15. Modelo Nulo - Os parasitóides não evitam o superparasitismo.

Para esse modelo é suficiente supor que a ovoposição ocorre a cada encontro, e portanto, existem  $N = R$  ovoposições ao todo. A suposição de que a ovoposição ocorre aleatória e independentemente sobre os  $H$  hospedeiros implica que, condicional ao número total  $N$  de ovoposições, existe uma distribuição multinomial para o número de ovoposições  $V_1, V_2, \dots, V_H$  nos  $H$  hospedeiros. A distribuição marginal do número de ovos colocados em um dado hospedeiro é Binomial  $\left(N, p = \frac{1}{H}\right)$ . Assim, para um particular hospedeiro, tem-se

$$P(V_j = i) = \pi_i = \binom{N}{i} p^i (1-p)^{N-i} \simeq \frac{e^{-N/H} (N/H)^i}{i!}, \quad (38)$$

pela aproximação usual da Poisson com média  $\frac{N}{H}$ , e a média e a variância de  $X_0$ , número de hospedeiros não parasitados ou com nenhum ovo, são dadas por:

$$E(X_0) = H\pi_0 = H(1-p)^N \simeq H e^{-N/H}$$

e

$$\begin{aligned} \text{Var}(X_0) &= H(1-p)^N + H(H-1)(1-2p)^N - [H(1-p)^N]^2 \\ &\simeq He^{-N/H} [1 - (1 + N/H)e^{-N/H}] \end{aligned}$$

quando  $N$  e  $H$  são grandes e  $N/H$  é limitado.

Mesmo para os modelos mais simples, é bem complicado achar uma forma fechada para a função de probabilidade de  $X_i$  embora calcular sua esperança seja elementar. Uma aproximação pode ser obtida, considerando as ovoposições seqüencialmente, levando à seguinte distribuição de recorrência para  $X_0$ :

$$p_k^{n+1} = P(X_0 = k \mid N = n + 1) = (1 - kp)p_k^n + (k + 1)p p_{k+1}^n,$$

sendo  $p_k^1 = I_{\{H-1\}}(k)$ ,  $n = 1, 2, \dots$  e  $k = 0, 1, \dots, H - 1$

Logo, a função de probabilidade de  $X_i$  condicionada a  $N = n$ , em que  $\min(H, n/i) = H$ , quando  $i = 0$ , é dada por:

$$P(X_i = k \mid N = n) = (-1)^k \sum_{r=k}^{\min(H, n/i)} \frac{(-1)^r H!}{k!(r-k)!(H-r)!} \frac{n! p^{ir} (1-rp)^{n-ir}}{(n-ir)!(i!)^r},$$

que pode ser aproximada pela distribuição de Poisson com parâmetro  $\mu_i = He^{-n/H}(n/H)^i/i!$ , que é totalmente distinta da aproximação dada em (38) quando  $N = n$ .

## M16. Modelos de Bakker - Versão Ocupação Seqüencial

Daley e Maindonald (1989) consideram o processo de ovoposições sucessivas e  $X_i^{(n)}$  o número de hospedeiros com  $i$  ovoposições após  $n$  ovoposições, isto é, o vetor  $X^{(n)} = (X_0^{(n)}, X_1^{(n)}, \dots)$ . Supõem, ainda, que um parasita sempre deposita ovos em um encontro com um hospedeiro não parasitado e que independente de cada encontro, uma ovoposição ocorre em um hospedeiro com  $i$  ovos com probabilidade  $\delta$ ,  $i = 1, 2, \dots$ . Logo, a probabilidade,  $\rho_k$ , de a próxima ovoposição ocorrer em um hospedeiro não parasitado, dado que  $k$  hospedeiros são não parasitados, é obtida por:

$$\rho_k = \frac{k}{H} + (1 - \frac{k}{H})(1 - \delta)\rho_k = \frac{kp}{kp + \delta(1 - kp)}. \quad (39)$$

Então, para  $k = 0, 1, \dots, H - 1$  e  $n = 1, 2, \dots$ , tem-se

$$p_k^{n+1} = P(X_0^{n+1} = k) = (1 - \rho k)pk^n + \rho_{k+1}p_{k+1}^n \quad (40)$$

com  $p_{H-1}^1 = 1$  e  $p_k^1$  para  $k = 0, 1, 2, \dots, H - 2, H$ . Na realidade,  $\{X_0^n\}$  é Markoviano e a equação (40) é a equação prospectiva (*forward*) de Chapman-Kolmogorov.

Supondo  $\delta_0 = 1 > \delta_1 \neq \delta_2 = \delta_3 = \dots = \delta$ , o processo Markoviano  $\{(X_0^{(n)}, X_1^{(n)})\}$  foi apresentado pelos autores com o intuito de melhorar a modelagem das probabilidades de transição em um único passo. Na suposição de independência de comportamento dos parasitas em cada encontro, tem-se:

$$\begin{aligned} p_{kl}^n &= P[(X_0^{(n)}, X_1^{(n)}) = (k, l)] \\ \rho_{kl}^0 &= P[(X_0^{(n+1)}, X_1^{(n+1)}) = (k - 1, l + 1) \mid (X_0^n, X_1^n) = (k, l)] \\ &= kp + [lp(1 - \delta_1) + (1 - kp - lp)(1 - \delta)]\rho_{kl}^0 \\ &= kp/[kp + (1 - kp - lp)\delta + lp\delta_1] \\ \rho_{kl}^1 &= P[X_1^{n+1} = l - 1 \mid (X_0^n, X_1^n) = (k, l)] \\ &= lp\delta_1 + [lp(1 - \delta_1)(1 - kp - lp)(1 - \delta)]\rho_{kl}^1 \\ &= lp\delta_1/[kp + (1 - kp - lp)\delta + lp\delta_1] \\ p_{kl}^{n+1} &= (1 - \rho_{kl}^0 - \rho_{kl}^1)p_{kl}^n + \rho_{k+1, l-1}^0 p_{k+1, l-1}^n + \rho_{k, l+1}^1 p_{k, l+1}^n. \end{aligned} \quad (41)$$

Um processo ainda mais geral surge quando  $\delta_i = \delta$  ( $i \geq m$ ) e, então,  $\{(X_0^{(n)}, X_1^{(n)}, \dots, X_{m-1}^{(n)})\}$  é Markoviano.

### M17. Modelo para o número sucessivo de encontros sem ovoposição

Uma família mais geral de diferentes modelos de ocupação seqüencial do que M16 proposta por Daley e Maindonald (1989), possibilitou a modelagem da forma da repulsa ao superparasitismo dependendo de duas características:

1. se o hospedeiro encontrado é não parasitado ou parasitado;
2. do número de encontros sem ovoposição efetuado pelo parasita desde sua última ovoposição.

A segunda característica faz com que não haja conexão entre esse modelo e os modelos de Bakker (1972), Rogers (1975), Griffiths (1977b) e de Morgan (1982).

Os autores supuseram que se há encontro com um hospedeiro não parasitado haverá ovoposição. Um contador,  $m$ , foi introduzido pelos mesmos, assumindo o valor 0 imediatamente após qualquer ovoposição, sendo  $\alpha_m$  a probabilidade de o parasita depositar ovos no  $m$ -ésimo encontro sucessivo com um hospedeiro já parasitado dado, que não houve ovoposição nos  $(m - 1)$  encontros anteriores com o mesmo hospedeiro.

Sob estas condições  $\{X_0^{(n)}\}$  é Markoviano, com

$$1 - \rho_k = P\left(X_0^{(n+1)} = k \mid X_0^{(n)} = k\right) = \sum_{m=1}^{\infty} (1 - kp)^m \alpha_m \prod_{k=1}^{m-1} (1 - \alpha_k). \quad (42)$$

O caso  $\alpha_m = \delta$  ( $m = 1, 2, \dots$ ) coincide com o modelo *M16*. Outro conjunto simples de parâmetros  $\{\alpha_m\}$  aparece quando  $\alpha_m = 0$  ( $m = 1, 2, \dots, \tau - 1$ ) e  $\alpha_\tau = 1$  para algum inteiro positivo  $\tau$ , de forma que

$$1 - \rho_k = (1 - kp)^\tau. \quad (43)$$

Isso formaliza a abordagem usada por Simmonds (1956). Análises mais acuradas mostraram que o parâmetro  $\tau$  é um real positivo desde que  $\rho_k$  varie no intervalo  $[0, 1]$ , embora a obtenção do resultado seja possível somente através de (42), para valores inteiros de  $\tau$ , e sabe-se que  $\rho_k$  é maior ou menor do que esse valor sob ovoposições aleatórias (Modelo Nulo), quando  $\tau$  é maior ou menor do que 1. No modelo *M16*, o inverso  $\delta^{-1}$  pode ser interpretado como o número médio de hospedeiros já parasitados em encontros sucessivos até ocorrer ovoposição. Assim,  $\tau^{-1}$  em (43) tem uma interpretação análoga à de  $\delta$ .

## **M18. Formulação do modelo como uma Cadeia de Markov a Parâmetro**

### **Contínuo**

Um modelo, mais geral ainda, foi proposto por Daley e Maindonald (1989) com base em uma observação de Bakker et al. (1972) sobre a existência de um processo estocástico de nascimento ao descrever  $X(t) = (X_0(t), X_1(t), \dots)$  para os números  $X_i(t)$  de hospedeiros no tempo  $t$  ( $0 < t < T$ ) após o início do experimento com  $i = 0, 1, 2, \dots$  ovoposições. Então, observando que, desde que ovoposições em hospedeiros distintos ocorrem independentemente,

$X(t)$  é uma cadeia de Markov de tempo contínuo com taxas de transição não nulas para  $(X_i(t), X_{i+1}(t)) \rightarrow (X_i(t) - 1, X_{i+1}(t) + 1)$ , permanecendo todos os demais componentes de  $X(t)$  inalterados, iguais a  $\lambda\delta_i X_i(t)$ .

Se nos tempos  $(0 < t_1 < t_2 < \dots < t_N < t_{N+1} < \dots)$  ocorrem ovoposições, então o processo de salto  $\{X(t_n+0)\}$ , embutido em  $X(t)$ , é precisamente uma cadeia de Markov de tempo discreto como comentado anteriormente. Suas probabilidades de transição, semelhantes a  $\rho_k$ , são obtidas com razão de taxas de transição como  $\lambda\delta_i X_i(t)$ . Desde que  $X(\cdot)$  é a soma de  $H$  vetores de funções indicadoras  $I_0$  correspondendo a  $H$  processos de nascimento independentes, a distribuição de  $(X_0, N)$  é a soma de  $H$  variáveis aleatórias bivariadas independentes com f.g.p. comum  $G(w, s) = E(w^{I_0} s^{V_j})$ . A distribuição de probabilidade de  $(X_0, N)$ ,  $\{q_{kn}^H\}$ , pode ser encontrada recursivamente. Note que no modelo de Daley e Maindonald (1989) a variável estudada é o número de ovoposições e não o número de ovos. Segundo os autores, usar a variável número de ovos (em lugar do número de ovoposições) violaria seriamente a suposição de independência exigida pelo modelo, o que faz que não seja comparável com os demais.

### 2.1.7 Modelo generalizado de Poisson - M19

Um modelo alternativo para análise de dados de parasitismo é a distribuição generalizada de Poisson (G.P.D) de parâmetros  $\theta > 0$  e  $\lambda \in (-1, 1)$  que tem função de probabilidade definida matematicamente, para  $\lambda > 0$ , por

$$f_X(x | \theta, \lambda) = \frac{\theta(\theta + x\lambda)^{x-1} e^{-(\theta+x\lambda)}}{x!} I_{\{0,1,2,\dots\}}(x). \quad (44)$$

Essa distribuição permite que o valor de  $\lambda$  seja negativo e nesse caso há um truncamento nos valores de  $x$ . Quando  $\lambda$  for negativo precisa-se ter  $(\theta + x\lambda) > 0$  de modo que  $\lambda > \frac{-x}{\theta}$ . Seja  $m$  o maior inteiro positivo para o qual  $(\theta + m\lambda) > 0$ . Assim,  $\lambda > \frac{-m}{\theta}$  e  $\lambda > -1$  que implica uma nova restrição para esse parâmetro que é  $\lambda > \max(\frac{-m}{\theta}, -1)$ . O limite inferior para  $m$  é imposto para que haja pelo menos cinco classes com probabilidades não nulas quando  $\lambda$  é negativo. Nessa situação, a f.p. de  $X$  é dada por:

$$f_X(x | \theta, \lambda) \approx \frac{\theta(\theta + x\lambda)^{x-1} e^{-(\theta+x\lambda)}}{x!} I_{\{0,1,2,\dots,m\}}(x). \quad (45)$$

Quando  $\lambda$  é negativo o modelo *GPD* definido por (44) inclui um truncamento devido ao fato que  $P(X = x) = 0$  para  $x > m$ . Assim,

$$F_m(\theta, \lambda) = \sum_{x=0}^m f_X(x|\theta, \lambda)$$

poderá assumir um valor maior ou menor do que 1, dependendo dos valores dos parâmetros. Na prática, segundo Consul (1989), os valores de  $m$ ,  $\lambda$  e  $\theta$  que surgem nas aplicações tornam desnecessário o uso de qualquer fator de normalização. Os símbolos  $\theta$  e  $\lambda$  são chamados de primeiro e segundo parâmetros do modelo, respectivamente. O segundo parâmetro  $\lambda$  é um indicador de superdispersão, de subdispersão, ou não, em relação à distribuição de Poisson. Esse parâmetro pode ser pensado como uma medida de afastamento dos dados em relação à distribuição de Poisson. Ele, ainda, reflete uma taxa média de restituição do processo, um valor positivo indica que os indivíduos estão fazendo um esforço para acelerar o processo enquanto um valor negativo significa um retardo. Quando  $\lambda = 0$ , tem-se a distribuição de Poisson. O primeiro parâmetro  $\theta$  é a taxa média natural de trocas do processo para a ocorrência dos eventos analisados e é um indicador da intensidade do processo de Poisson. Em muitas situações práticas é sabido que o segundo parâmetro,  $\lambda$ , é linearmente proporcional ao primeiro  $\theta$ .

#### Momentos da distribuição de Poisson Generalizada:

- i.  $E(X) = \frac{\theta}{1 - \lambda}$ ;
- ii.  $\mu_2 = V(X) = \frac{\theta}{(1 - \lambda)^3}$ ;
- iii.  $\mu_3 = \frac{\theta(1 + 2\lambda)}{(1 - \lambda)^5}$ ;
- iv.  $\mu_4 = \frac{3\theta^2}{(1 - \lambda)^6} + \frac{\theta(1 + 8\lambda + 6\lambda^2)}{(1 - \lambda)^7}$ .

O coeficiente de assimetria e o coeficiente de curtose são dados, respectivamente, por:

- i.  $\gamma_1 = \frac{(1 + 2\lambda)}{\sqrt{(1 - \lambda)\theta}}$ ;

$$\text{ii. } \beta_2 = 3 + \frac{(1 + 8\lambda + 6\lambda^2)}{(1 - \lambda)\theta}$$

Fato: A f.g.p. de  $X \sim PoiGen(\theta, \lambda)$ ,  $G_X(u)$ , é dada por:

$$G_X(u) = e^{\theta(t-1)}, \text{ em que } t = ue^{\lambda(t-1)}. \quad (46)$$

Segundo Consul (1989), a distribuição de Poisson Generalizada é a distribuição limite da Binomial Negativa Generalizada, da distribuição Quase-binomial e da distribuição Generalizada de Markov-Polya.

Algumas das propriedades importantes da distribuição Generalizada de Poisson são demonstradas no Apêndice E.

## 2.2 Estimação

Os dois primeiros momentos amostrais da distribuição do total do número de ovos foram usados por Bakker et al. (1967, 1972) para estimar os parâmetros,  $\lambda$  e  $\delta$  de seus modelos. Esse processo requer o uso de métodos iterativos em sua determinação. Para evitá-lo, Griffiths (1977a) apresentou o método da média e frequência zero, que tem formas fechadas para os estimadores e é mais eficiente no sentido em que fornece menores variâncias. Sejam a média do número de ovos por hospedeiro,

$$\bar{Y} = \frac{\sum_{i=1}^H Y_i}{H}$$

e a frequência amostral de hospedeiros não parasitados,

$$f_0 = \frac{X_0}{H}.$$

Os estimadores pelo método da média usando (12) e da frequência zero são dados por:

$$e^{-\tilde{\lambda}} = f_0$$

e portanto,

$$\tilde{\lambda} = -\ln(f_0).$$



$$\tilde{\lambda}\tilde{\delta} + (1 - e^{-\tilde{\lambda}})(1 - \tilde{\delta}) = \bar{Y}.$$

e portanto,

$$\tilde{\delta} = \frac{\bar{Y} - 1 + f_0}{\tilde{\lambda} - 1 + f_0}.$$

Griffiths (1977a) utilizou um outro estimador proposto por Rogers (1975), baseado nas proporções observadas  $f_0$  e  $f_1$  de 0 e 1 ovo, respectivamente, para estimar os parâmetros do modelo M1 de Bakker.

Usando o método das frequências de zeros e uns, tem-se:

$$\begin{cases} e^{-\hat{\lambda}} = f_0 \\ \frac{e^{-\hat{\lambda}\hat{\delta}} - e^{-\hat{\lambda}}}{\hat{\delta} - 1} = f_1. \end{cases}$$

No método dos momentos, o sistema de equações é determinado pelas expressões:

$$\begin{cases} \lambda\delta + (1 - e^{-\lambda})(1 - \delta) = \bar{y} \\ \lambda^2\delta^2 + (1 - e^{-\lambda})(1 - \delta)(1 - 2\delta) + \lambda\delta(3 - 2\delta) = s^2 \end{cases}$$

e, a partir desse sistema, obtêm-se as equações recursivas. Além desses métodos, o autor usou o método da máxima verossimilhança.

Os estimadores para o modelo de Rogers, apresentado por Griffiths (1977b), pelo método da média e da frequência zero são dados por:

$$\tilde{y} = \frac{\ln(f_0)}{\ln(1 - 1/N)},$$

e

$$\tilde{\delta} = \frac{\bar{X} - 1 + f_0}{\tilde{y}/N - 1 + f_0}.$$

Para fazer a estimação dos parâmetros de seus modelos, Janardan (1997) considerou, sem perda de generalidade,  $t = 1$  e utilizou o método da média e da frequência de

zeros e uns e a reparametrização  $\theta = \frac{\alpha}{\lambda}$ . A solução do sistema:

$$\begin{aligned} f_0 &= e^{-\lambda}, \\ f_1 &= \frac{\lambda}{p\beta}[e^{-\lambda} - e^{-\alpha}], \\ \bar{X} &= m \end{aligned}$$

forneceu o estimador de  $\lambda$ :

$$\hat{\lambda} = -\ln(f_0).$$

O estimador de  $\theta$ ,  $\hat{\theta}$ , é a solução da equação

$$\frac{e^{(1-\hat{\theta})\hat{\lambda}} - 1}{1 - \hat{\theta}} = \frac{f_1}{f_0},$$

que tem solução única pois a função  $\frac{e^x - 1}{x}$  é uma função estritamente crescente para  $x \neq 0$ .

Logo, o estimador de  $p$  é dado por:

$$\hat{p} = \frac{\hat{b}_0}{\bar{X} - \hat{b}_1},$$

em que

$$\hat{b}_0 = \left[ \hat{\lambda}(\hat{\theta} - 1) - \hat{\theta} + \frac{1}{\hat{\theta}} + \hat{\theta}f_0 - \frac{f_0^{\hat{\theta}}}{\hat{\theta}} \right]$$

e

$$\hat{b}_1 = \left[ 1 + \hat{\lambda} - \frac{1}{\hat{\theta}} - f_0 + \frac{f_0^{\hat{\theta}}}{\hat{\theta}} \right]$$

e o de  $\mu$ , por:

$$\hat{\mu} = \hat{\lambda} \left[ \frac{\hat{\theta} - 1}{\hat{p}} + 1 \right].$$

Nos demais modelos, o processo de estimação utilizado foi o da máxima verossimilhança. Seja  $\boldsymbol{\theta}$  o vetor de parâmetros de um determinado modelo e seja uma amostra aleatória de tamanho  $n$  da variável aleatória desse modelo com os valores  $j = 0, 1, 2, \dots, y_{max}$  com freqüências absolutas  $f_0, f_1, \dots, f_{y_{max}}$ , respectivamente. A função de verossimilhança é dada por:

$$L(\boldsymbol{\theta}) = \prod_{j=0}^{y_{max}} [p_j(t)]^{f_j},$$

em que  $p_j(t) = P[Y(t) = j]$ .

Como uma aplicação, foi feita a estimação dos parâmetros do modelo  $M1$  de Bakker, considerando-se os diversos métodos de estimação, para os dados da Tabela 2, utilizando-se o procedimento IML do SAS (Anexo A). Ainda, para o método da máxima verossimilhança, foi utilizado também o procedimento NLMIXED do SAS que resolve diretamente a estimação dos parâmetros dos modelos bem como calcula sua matriz estimada de variâncias-covariâncias (Anexo A).

Tabela 4 - Estimativas do parâmetros do Modelo  $M1$  de Bakker para a distribuição do número de ovos da *Spalangia drosophilae* em pupas de *Drosophila melanogaster* ( $Y$ ), obtidas através de 5 métodos de estimação

Método	$\lambda$		$\delta$		Correl.
	Estimat.	E.P.	Estimat.	E.P.	
Met. 1 - Média e Freq. zero	2,5609	0,2204	0,1712	0,0277	-0,6526
Met. 2 - Freq. de zeros e uns	2,5609	0,2204	0,1696	0,0292	-0,6239
Met. 3 - Momentos	2,4479	0,2357	0,1887	0,0357	-0,7391
Met. 4 - Máx. Verossim.(IML)	2,5575	0,2199	0,1716	0,0278	-0,6523
Met. 5 - Máx. Verossim.(NLMixed)	2,5575	0,2199	0,1716	0,0278	-0,6523

Da Tabela 2 tem-se que a proporção observada de hospedeiros sem ovos é  $f_0=19/246=0,0772$  e que a proporção observada de hospedeiros com 1 ovo é  $f_1=169/246=0,6870$ . O número médio de ovos por hospedeiro é  $\bar{y}=1,203$ . Isto originou as estimativas dos parâmetros pelos métodos 1 e 2. As estimativas dos parâmetros, seus erros-padrões e a correlação são dadas na Tabela 4. Como se nota com exceção do método dos momentos, os outros métodos fornecem resultados bastante próximos.

A Tabela 5 traz, além do número esperado de hospedeiros usando o modelo de Poisson, as frequências esperadas pelos 5 métodos para o modelo  $M1$  de Bakker. Assim, a hipótese de que o parasita não evita o superparasitismo ( $\delta=1$ ) não pode confirmada através desses dados.

Tabela 5 - Número esperado de ovos ( $Y$ ) de *Spalangia drosophilae* em pupas de *Drosophila melanogaster*, obtidas através de cinco métodos de estimação

Número de ovos	Observada	Poisson(1, 20)	Frequências Estimadas - Bakker(M1)				
			Mét. 1	Mét. 2	Mét.3	Mét. 4	Mét. 5
0	19	73,85	19,00	19,00	21,27	19,06	19,06
1	169	88,86	168,53	169,00	164,84	168,46	168,46
2	50	53,46	49,13	48,82	49,90	49,13	49,13
3	6	21,44	8,25	8,13	8,77	8,26	8,26
4	1	6,45	0,98	0,96	1,10	0,99	0,99
5	1	1,55	0,09	0,09	0,11	0,09	0,09
>5	0	0,39	1,00	0,00	0,01	0,01	0,01
Total	246	246	246	246	246	246	246

### 2.3 Equivalência e relação entre modelos

Como visto no item 2.1, muitos modelos (M1 a M18) foram propostos para o estudo de parasitismo. Aqui, será provada a equivalência entre alguns dos modelos.

#### (i) Modelos M1 - Bakker e M8 - Hemerik

Os modelos M1-Bakker e M8-Hemerik têm funções de probabilidade são dadas por (9) e (28), respectivamente. A f.p. para o número total de ovos,  $S_j$ , é dada por:

$$P(S_j=s) = e^{-\gamma NT} I_{\{0\}}(s) + \delta^{s-1} \left[ \frac{(-1)^s e^{-\gamma NT}}{(1-\delta)^s} + e^{-\delta \gamma NT} \sum_{i=1}^s \frac{(-1)^{i-1} (\gamma NT)^{s-i}}{(1-\delta)^i (s-i)!} \right] I_{\{1,2,\dots\}}(s)$$

Tomando-se  $\lambda = \gamma NT$ , fixo, e  $\alpha = \delta$ , tem-se

$$P(S_j = s) = e^{-\lambda} I_{\{0\}}(s) + \delta^{s-1} \left[ \frac{(-1)^s e^{-\lambda}}{(1-\delta)^s} + e^{-\delta \lambda} \sum_{i=1}^s \frac{(-1)^{i-1} (\lambda)^{s-i}}{(1-\delta)^i (s-i)!} \right] I_{\{1,2,\dots\}}(s).$$

Mas  $\frac{(-1)^s e^{-\lambda}}{(1-\delta)^s} = \frac{(-1)^s (-1)^s e^{-\lambda}}{(-1)^s (1-\delta)^s} = \frac{e^{-\lambda}}{(\delta-1)^s}$  e fazendo  $j = s - i$ , tem-se

$$\begin{aligned}
\sum_{i=1}^s \frac{(-1)^{i-1} \lambda^{s-i}}{(1-\delta)^i (s-i)!} &= \sum_{j=0}^{s-1} \frac{(-1)^{s-j-1} \lambda^j}{(1-\delta)^{s-j} j!} \\
&= -\frac{(-1)^s}{(1-\delta)^s} \sum_{j=0}^{s-1} \frac{(-1)^{-j} \lambda^j}{(1-\delta)^{-j} j!} \\
&= -\frac{(-1)^s}{(1-\delta)^s} \sum_{j=0}^{s-1} \frac{\lambda^j}{(\delta-1)^{-j} j!} \\
&= -\frac{(-1)^s}{(1-\delta)^s} \sum_{j=0}^{s-1} \frac{\lambda^j (\delta-1)^j}{j!} \\
&= -\frac{1}{(\delta-1)^s} \sum_{j=0}^{s-1} \frac{\lambda^j (\delta-1)^j}{j!}.
\end{aligned}$$

Então,

$$\begin{aligned}
P(S_j = s) &= e^{-\lambda} I_{\{0\}}(s) + \delta^{s-1} \left[ \frac{e^{-\lambda}}{(\delta-1)^s} - \frac{e^{-\delta\lambda}}{(\delta-1)^s} \sum_{j=0}^{s-1} \frac{\lambda^j (\delta-1)^j}{j!} \right] I_{\{1,2,\dots\}}(s) \\
&= e^{-\lambda} I_{\{0\}}(s) + \delta^{s-1} (\delta-1)^{-s} \left[ e^{-\lambda} - e^{-\delta\lambda} \sum_{j=0}^{s-1} \frac{\lambda^j (\delta-1)^j}{j!} \right] I_{\{1,2,\dots\}}(s) \\
&= p_s,
\end{aligned}$$

que é a equação (12) de Griffiths (1977, página 106) e o primeiro modelo de Bakker (1967, 1972). Esse modelo com  $N = K$  é o modelo *M8* de Hemerik.

## (ii) Modelo M12 de Hemerik como uma modificação do modelo M1 de Bakker

Bakker considerou o número de encontros,  $R_j$ , entre um parasitóide e um hospedeiro com uma variável aleatória tendo distribuição de Poisson com parâmetro  $\lambda$ . Quando se supõe  $\lambda$  aleatório com distribuição gama, isso equivale a supor que o número de encontros  $R_j$  tem distribuição binomial negativa com média  $\mu$  e variância  $\mu + \sigma^2$ . Esse modelo foi utilizado por Hemerik et al. (2002) com uma parametrização diferente, conforme pode ser visto a seguir.

Fazendo  $\lambda = \gamma NT$  aleatório, isto é,  $\Lambda \sim \text{Gama}(m, \beta)$  e  $\alpha = \delta$ , tem-se uma modificação do modelo *M1* de Bakker (1967, 1972) com função de probabilidade

$$P(S_j = s | \Lambda = \lambda) = e^{-\lambda} I_{\{0\}}(s) + \delta^{s-1} \left[ \frac{(-1)^s e^{-\lambda}}{(1-\delta)^s} + e^{-\delta\lambda} \sum_{i=1}^s \frac{(-1)^{i-1} (\lambda)^{s-i}}{(1-\delta)^i (s-i)!} \right] I_{\{1,2,\dots\}}(s).$$

Portanto, incondicionalmente,

$$P(S_j = 0) = P(S_j = 0 | \Lambda = \lambda) = \left(\frac{1}{1 + \beta}\right)^\alpha = \left(\frac{1}{1 + \gamma NT}\right)^m = \left(\frac{m}{m + m\gamma NT}\right)^m$$

que é a equação (6a) of Hemerik et al. (2002, página 174). Para  $s > 0$ , tem-se:

$$\begin{aligned} P(S_j = s) &= \delta^{s-1} \left[ \int_0^\infty \left[ \frac{(-1)^s e^{-\lambda}}{(1-\delta)^s} + e^{-\delta\lambda} \sum_{i=1}^s \frac{(-1)^{i-1} (\lambda)^{s-i}}{(1-\delta)^i (s-i)!} \right] \frac{\lambda^{m-1} e^{-\lambda/\beta}}{\beta^m \Gamma(m)} d\lambda \right] I_{\{1,2,\dots,\infty\}}(s) \\ &= \delta^{s-1} \int_0^\infty \frac{(-1)^s e^{-\lambda}}{(1-\delta)^s} \frac{\lambda^{m-1} e^{-\lambda/\beta}}{\beta^m \Gamma(m)} d\lambda + \\ &\quad \delta^{s-1} \int_0^\infty e^{-\delta\lambda} \sum_{i=1}^s \frac{(-1)^{i-1} (\lambda)^{s-i}}{(1-\delta)^i (s-i)!} \frac{\lambda^{m-1} e^{-\lambda/\beta}}{\beta^m \Gamma(m)} d\lambda \\ &= \delta^{s-1} \frac{(-1)^s}{(1-\delta)^s \beta^m} \int_0^\infty \frac{\lambda^{m-1}}{\Gamma(m)} e^{-(1+\frac{1}{\beta})\lambda} d\lambda + \\ &\quad \delta^{s-1} \sum_{i=1}^s \frac{(-1)^{i-1}}{(1-\delta)^i (s-i)! \beta^m} \int_0^\infty \frac{\lambda^{m+s-i-1}}{\Gamma(m)} e^{-(\delta+\frac{1}{\beta})\lambda} d\lambda \\ &= \delta^{s-1} \frac{(-1)^s}{(1-\delta)^s \beta^m} \frac{1}{(1+\frac{1}{\beta})^m} + \\ &\quad \delta^{s-1} \sum_{i=1}^s \frac{(-1)^{i-1}}{(1-\delta)^i (s-i)! \Gamma(m) \beta^m} \frac{\Gamma(m+s-i)}{(\delta+\frac{1}{\beta})^{m+s-i}} \end{aligned}$$

$$\begin{aligned}
&= \delta^{s-1} \frac{(-1)^s}{(1-\delta)^s} \frac{1}{(1+\beta)^m} + \\
&\quad \sum_{i=1}^s \frac{(-1)^{i-1} \Gamma(m+s-i)}{\Gamma(m)} (1-\delta)^i (s-i)! \frac{1}{(1+\delta\beta)^m} \left[ \frac{\beta}{1+\delta\beta} \right]^{s-i} \delta^{s-i+i-1} \\
&= \delta^{s-1} \frac{(-1)^s}{(1-\delta)^s} (1+\beta)^{-m} + \\
&\quad \sum_{i=1}^s \frac{\Gamma(m+s-i)(-\delta)^{i-1}}{\Gamma(m)(1-\delta)^i (s-i)!} (1+\delta\beta)^{-m} \left[ \frac{\delta\beta}{1+\delta\beta} \right]^{s-i} \\
&= \left[ \delta^{s-1} \frac{(-1)^s}{(1-\delta)^s} (1+\beta)^{-m} + \sum_{i=1}^s \frac{\Gamma(m+s-i)(-\delta)^{i-1}}{\Gamma(m)(1-\delta)^i (s-i)!} \left[ \frac{1}{1+\delta\beta} \right]^m \left[ \frac{\delta\beta}{1+\delta\beta} \right]^{s-i} \right]
\end{aligned}$$

que é a equação (6b) de Hemerik et al (2002, página 174, modelo *M2b*) e também o modelo *M12* apresentado no item 2.1 com  $N = K$ . Há um erro de impressão nesse artigo, página 174 - equação (6b) ao invés de  $(-\alpha^{i-1})$  é  $(-\alpha)^{i-1}$ .

### (iii) Modelos *M4* - Bakker e *M9* - Hemerik

A distribuição do número,  $S_j$ , de ovos postos nos hospedeiros do modelo *M9* de Hemerik tem dois parâmetros,  $\delta$  e  $\lambda = \gamma KT$ , e é dada por (29), enquanto que no modelo *M4* de Bakker tem dois parâmetros,  $\delta$  e  $\lambda$ , e é dada por (17).

Considere a expressão (29) de Hemerik e faça  $\lambda = \gamma KT$  e a transformação  $i = j$ . Então,

$$\begin{aligned}
P(S_j = s) &= \delta^{\frac{(s-1)s}{2}} \sum_{j=0}^s \frac{e^{-\lambda\delta^j}}{s \prod_{\substack{i=0 \\ i \neq j}} (\delta^j - \delta^i)} I_{\{0,1,\dots\}}(s) \\
&= \delta^{\frac{(s-1)s}{2}} \sum_{j=0}^s \frac{e^{-\lambda\delta^j}}{s \prod_{\substack{i=0 \\ i \neq j}} -(\delta^i - \delta^j)} I_{\{0,1,\dots\}}(s)
\end{aligned}$$

$$\begin{aligned}
&= \frac{\delta^{\frac{(s-1)s}{2}}}{(-1)^s} \sum_{j=0}^s \frac{e^{-\lambda\delta^j}}{\prod_{\substack{i=0 \\ i \neq j}}^s (\delta^i - \delta^j)} I_{\{0,1,\dots\}}(s) \\
&= (-1)^s \delta^{\frac{(s-1)s}{2}} \sum_{j=0}^s \frac{e^{-\lambda\delta^j}}{\prod_{\substack{i=0 \\ i \neq j}}^s (\delta^i - \delta^j)} I_{\{0,1,\dots\}}(s),
\end{aligned}$$

que é a expressão (17) de Bakker.

#### (iv) Modelos M12 - Hemerik e M6 - Griffiths

Considere agora o Modelo de Hemerik em que a distribuição do número de encontros é Binomial negativa com parâmetros  $m$  e  $p$ , e substituindo-se  $q_0 = (1-p)^{-m}$  e  $G_{R_j} = [\frac{p}{1-(1-p)s}]^m$  na equação (25), tem-se

$$G_{S_j}(s) = \frac{sp^m}{(1-\delta+\delta s)[p+\delta(1-\delta)(1-s)]^m} + \frac{(1-p)^{-m}(1-\delta)(1-s)}{1-\delta+\delta s}$$

que é a f.g.p. da variável aleatória do modelo M12 de Hemerik.

Esse mesmo modelo foi usado por Morgan (1975) para descrever o processo de fertilização dos ovos do ouriço-do-mar em que a variável resposta estudada era o número de espermatozoides entrando em um único ovo.

## 2.4 Modelo de Faddy - M20

Modelos estendidos de processos de Poisson fornecem uma estrutura bastante ampla para a análise de dados discretos (Faddy, 1997a, 1997b, 1998). Esses processos representam uma distribuição discreta para o número de eventos ocorridos em um intervalo de tempo finito como um processo de nascimento puro de Markov. Nesse processo, as taxas de ocorrências dos eventos, conhecidas como taxas de transição, são modeladas como função de  $y$ , o número de eventos acumulados. Isso é chamado de dependência de estado ou forma  $y$ -dependente das taxas as quais determinam as propriedades de dispersão da distribuição discreta resultante. A generalidade dessa abordagem reside no fato de que qualquer distribuição discreta com suporte não negativo tem uma representação do tipo Faddy (1997a)



e, portanto, outros modelos para dados discretos podem ser vistos como casos especiais. Entretanto, os modelos de Poisson estendidos envolvem um esforço computacional bem maior e mais difícil do que os outros modelos tradicionais para dados discretos.

Podlich (1999), em sua tese de doutorado, estendeu uma classe geral de modelos proposta por Faddy (1997a) que permitia o afastamento do modelo de Poisson com particular interesse em dados de abundância de espécies. Podlich, Faddy e Smith (2002) propuseram um método geral para análise de dados ecológicos, usando uma representação em termos de um processo nascimento puro de Markov para as distribuições de contagem envolvidas. Podlich & Smyth (2002) propuseram uma aproximação em “saddlepoint” para a distribuição binomial negativa usando um processo de nascimento geral. Essa classe de modelos, segundo Podlich et al (2004), pode ser aplicada a qualquer problema de regressão com resposta do tipo contagens mas é especialmente apropriada quando as contagens acumulam-se através do tempo. Nesses casos, a forma da dependência de estados, taxas de transição podem ser diretamente interpretadas e fornecem uma explicação intuitiva para qualquer superdispersão ou subdispersão que tenha ocorrido.

A forma com que se escolhem as taxas de transição da dependência em  $y$  determina a maneira como a distribuição resultante se afasta da distribuição de Poisson. Uma das maneiras é a alta frequência de um particular valor (geralmente zero ou um) em relação ao modelo Poisson. Outra maneira de afastamento é em relação à razão média-variância que na Poisson é igual a 1. Dados para os quais a variância observada é significativamente maior do que a média é indicativo de uma variabilidade maior do que a distribuição de Poisson e representa uma superdispersão. Se a variância é significativamente menor do que a média, tem-se subdispersão. Modelar o efeito de super ou subdispersão em relação ao modelo Poisson é muito importante para tornar confiável a significância de covariáveis. Muitos modelos de mistura associados ao modelo de Poisson têm sido amplamente empregados na literatura mas a estrutura geral dos modelos de Poisson estendidos não necessita de especificar nenhuma delas. Essa classe de modelos será denominada de agora em diante de modelos de Faddy (ver Apêndice C) e será contextualizada para dados oriundos de estudos de parasitismo.

Supondo que o número de ovos postos até um instante  $0 \leq t \leq T$ ,  $Y(t)$  com

$Y(0) = 0$ , segue um processo de nascimento puro, tem-se

$$P(Y(t + \Delta t) = y + 1 \mid Y(t) = y) = \delta_y \Delta t + o(\Delta t).$$

Se  $\delta_y = \delta$  for independente de  $y$  a distribuição do número de ovos postos é Poisson  $\delta t$ .

Se  $\delta_y$  cresce linearmente com  $y$ , então a distribuição é binomial negativa. De uma forma geral, qualquer distribuição de contagem pode ser obtida a partir de uma escolha adequada da sequência  $\{\delta_y\}$ . Em particular, a sequência

$$\delta_y = \begin{cases} \delta_0 & \text{se } y = 0; \\ \delta_1 & \text{se } y = 1, 2, \dots, \end{cases}$$

em que  $\delta_1 < \delta_0$  pode ser usada para descrever o fato de que a taxa com que novos ovos são postos em hospedeiros já parasitados decresce. Uma formulação ainda mais geral modela essa forma de crescimento como função do número de ovos presentes no hospedeiro antes da próxima ovoposição através de

$$\delta_y = \begin{cases} \delta_0 & \text{se } y = 0; \\ \delta_1 y^c & \text{se } y = 1, 2, \dots, \end{cases}$$

sendo  $c$  o parâmetro que controla a variabilidade. Se  $c = 0$ , tem-se a distribuição de Poisson modificada no ponto zero, enquanto que  $c > 0$  e  $c < 0$  correspondem, respectivamente, a maior variabilidade (superdispersão) ou menor (subdispersão). Os outros parâmetros,  $\delta_0$  e  $\delta_1$ , podem ser modelados como funções log-lineares das covariáveis,

$$\delta_0 = \exp[\mathbf{z}'_0 \boldsymbol{\beta}_0]$$

$$\delta_1 = \exp[\mathbf{z}'_1 \boldsymbol{\beta}_1].$$

em que  $\mathbf{z}$  é o vetor das covariáveis e  $\boldsymbol{\beta}_0$  e  $\boldsymbol{\beta}_1$  são vetores de parâmetros ligados às covariáveis. O uso de funções log-lineares garante que as taxas de transição permanecem positivas.

Para construir a distribuição de probabilidade,  $p_y(t)$ , através de um processo de nascimento Markoviano com taxas  $\delta_y$ , é necessário resolver um sistema de equações diferenciais utilizando as equações prospectivas (*forward*) de Chapman-Kolmogorov (Cox; Miller, 1965),

$$\dot{p}'_y(t) = \begin{cases} -\delta_0 p_0(t), & \text{com } p_0(0) = 1; \\ \delta_{y-1} p_{y-1}(t) - \delta_y p_y(t), & \text{com } p_y(0) = 0 \text{ se } y = 1, 2, \dots \end{cases} \quad (47)$$

A solução do sistema é dada por:

$$p_y(t) = \begin{cases} e^{-\delta_0 t}, & \text{se } y = 0; \\ \delta_{y-1} e^{-\delta_y t} \int_0^t p_{y-1}(u) e^{\delta_y u} du, & \text{se } y = 1, 2, \dots \end{cases} \quad (48)$$

Sem qualquer perda de generalidade na solução desse sistema, pode-se tomar  $t = 1$ .

Para a solução do sistema de equações diferenciais usado no modelo de Faddy (47), considere inicialmente a equação para  $y=0$ :

$$p_0'(t) = \frac{dp_0(t)}{dt} = -\delta_0 p_0(t),$$

logo,

$$\frac{dp_0(t)}{p_0(t)} = -\delta_0 dt,$$

de onde se tem:

$$\int_0^t \frac{dp_0(u)}{p_0(u)} du = -\delta_0 \int_0^t du.$$

Logo

$$\ln(p_0(t)) - \ln(p_0(0)) = -\delta_0 t,$$

e assim

$$p_0(t) = e^{-\delta_0 t}.$$

Para  $y = 1, 2, \dots$ , com  $p_y(0) = 0$ , tem-se

$$p_y'(t) = \delta_{y-1} p_{y-1}(t) - \delta_y p_y(t),$$

e, assim,

$$\frac{dp_y(t)}{dt} + \delta_y p_y(t) = \delta_{y-1} p_{y-1}(t)$$

que multiplicada por  $e^{\delta_y t}$  fica

$$\frac{dp_y(t)}{dt} e^{\delta_y t} + p_y(t) \delta_y e^{\delta_y t} = \delta_{y-1} p_{y-1}(t) e^{\delta_y t}.$$

Assim, a derivada de  $p_y(t) e^{\delta_y t}$  é o lado esquerdo da equação anterior, isto é,

$$\frac{d}{dt} [p_y(t) e^{\delta_y t}] = \delta_{y-1} p_{y-1}(t) e^{\delta_y t}.$$

Logo,

$$\int_0^t \frac{d}{du} [p_y(u)e^{\delta_y u}] du = \delta_{y-1} \int_0^t p_{y-1}(u)e^{\delta_y u} du$$

e, portanto,

$$p_y(t)e^{\delta_y t} - p_y(0) = \delta_{y-1} \int_0^t p_{y-1}(u)e^{\delta_y u} du$$

ou ainda,

$$p_y(t) = \delta_{y-1} e^{-\delta_y t} \int_0^t p_{y-1}(u)e^{\delta_y u} du,$$

cuja solução, para  $\delta_i \neq \delta_j$  e  $y \geq 1$ , é dada por Biswas (1995), página 31,

$$p_y(t) = (-1)^y \delta_0 \delta_1 \dots \delta_{y-1} \sum_{i=0}^y \frac{e^{-\delta_i t}}{\prod_{\substack{j=0 \\ j \neq i}}^y (\delta_i - \delta_j)}. \quad (49)$$

Essa mesma expressão é apresentada de modo diferente por Podlich (1999)

$$p_y(t) = \frac{1}{\delta_y} \sum_{i=0}^y \delta_i e^{-\delta_i t} \prod_{\substack{j=0 \\ j \neq i}}^y \frac{\delta_j}{\delta_j - \delta_i}. \quad (50)$$

Uma outra maneira de resolver esse sistema é reescrevendo as equações de Chapman-Kolmogorov (Cox-Miller, 1965):

$$\frac{dp(t)}{dt} = p(t)Q,$$

em que  $p(t) = (p_0(t), p_1(t), \dots, p_y(t))$  e  $Q$  é a matriz das taxas de transição,

$$Q = \begin{bmatrix} -\delta_0 & \delta_0 & 0 & \dots & 0 & 0 \\ 0 & -\delta_1 & \delta_1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\delta_{y-1} & \delta_{y-1} \\ 0 & 0 & 0 & \dots & 0 & -\delta_y \end{bmatrix}.$$

A solução pode ser escrita em termos da função matricial exponencial (ver Apêndice D),

$$\begin{aligned}\mathbf{p}(t) &= \mathbf{p}(0) \sum_{i=0}^{\infty} \frac{t^i}{i!} \mathbf{Q}^i \\ &= \mathbf{p}(0) \exp(t\mathbf{Q}) \\ &= (1, 0, 0, \dots, 0) \exp(\mathbf{Q}t),\end{aligned}$$

o vetor  $\mathbf{p}(0) = (1, 0, 0, \dots, 0)$  é o vetor de probabilidade inicial pois que  $Y(0) = 0$  ou fazendo  $t = 1$

$$(p_0, p_1, \dots, p_y) = (1, 0, 0, \dots, 0, 0) \exp(\mathbf{Q}).$$

Assim, a probabilidade de se obter uma contagem de  $y$  ovos,  $p_y$  é calculada tomando-se o último elemento do vetor.

Modelos desse tipo foram utilizados por Faddy (1996) para analisar a abundância de espécies animais. Faddy (1998) mostrou que qualquer distribuição discreta com suporte finito tem uma representação em termos de um processo de morte de Markov com taxas de transição  $\delta_y$ ,  $y \geq 0$ , sendo que formas lineares para  $\delta_y$ ,  $y \geq 0$ , originam a distribuição binomial. Fenlon e Faddy (1999) modelaram o processo de invasão de  $N$  nematóides em larvas de mosca através de um processo estocástico de nascimento puro de taxa  $\lambda_y = (N - y)a_y$ , sendo que  $a_y$  pode assumir quatro formas distintas.

#### 2.4.1 Aproximação para o modelo de Faddy

Embora as contagens resultantes de ensaios em que há o comportamento de evitar o superparasitismo não sejam grandes, uma aproximação em “saddlepoint”, bastante acurada quando  $y$  é grande, segundo Daniels (1982) é dada por:

$$P(Y = y) \approx \frac{\prod_{j=0}^{y-1} \delta_j e^{-\tilde{s}}}{\prod_{j=0}^y (\delta_j - \tilde{s}) \left[ 2\pi \sum_{j=0}^y \frac{1}{(\delta_j - \tilde{s})^2} \right]^{1/2}} \left[ 1 + \frac{1}{8} \rho_4 - \frac{5}{24} \rho_3^2 \right] \quad (51)$$

com  $\tilde{s}$  satisfazendo

$$\sum_{j=0}^y \frac{1}{(\delta_j - \tilde{s})} = 1$$

e

$$\rho_r = \left[ \sum_{j=0}^y \frac{1}{(\delta_j - \tilde{s})^2} \right]^{-r/2} (r-1)! \sum_{j=0}^y \frac{1}{(\delta_j - \tilde{s})^r}.$$

Segundo Podlich (2004), esta aproximação mantém um erro relativo pequeno para todos valores de  $y$  o que é importante no cálculo do logaritmo da função de log-verossimilhança. Uma aproximação mais precisa foi desenvolvida por Smith & Podlich (2002) e tem uma vantagem adicional que os casos particulares Poisson, binomial e binomial negativo tornam-se exatos.

### 2.4.2 Modelo de Faddy como generalização

Será demonstrado, a seguir, que o modelo de Faddy (M20), usualmente empregado no estudo de tamanho de população animal, é uma generalização dos vários modelos usados na literatura em análise de dados de contagem, exceto o de Poisson Generalizado (M19), especificando a particular seqüência  $\delta_y$  de coeficientes que os geram. Assim, têm-se os casos:

- i.  $\delta_y = \delta \Rightarrow$  Poisson;
- ii.  $\delta_y = \lambda(H - y) \Rightarrow$  Binomial( $H, p = 1 - e^{-\lambda t}$ );
- iii.  $\delta_y = \lambda(r + y) \Rightarrow$  Binomial Negativa( $r, p = e^{-\lambda t}$ ) com  $\lambda > 0$ ;
- iv.  $\delta_0 = \lambda$  e  $\delta_1 = \lambda \delta \Rightarrow$  Modelo M1 de Bakker;
- v.  $\delta_0 = \lambda$  e  $\delta_1 = \alpha$  e  $\delta_y = \mu$  se  $y=2,3,\dots \Rightarrow$  Modelo de Janardan;
- vi. se a seqüência  $\delta_y$  for crescente gera uma distribuição com dispersão maior do que a associada à Poisson;
- vii. se a seqüência  $\delta_y$  for decrescente gera uma distribuição com dispersão menor do que a associada à Poisson;
- viii. se a seqüência  $\delta_y$  for crescente convexa gera uma distribuição com dispersão maior do que a da binomial negativa (Poisson-Normal);

ix. se a seqüência  $\delta_y$  for crescente côncava gera uma distribuição com dispersão entre a Poisson e a binomial negativa (Poisson-Lognormal, Poisson-Normal-Inversa).

(i) **Distribuição de Poisson**

A distribuição de Poisson  $p_y(t) = \frac{e^{-\delta t}(\delta t)^y}{y!}$  é um caso particular do modelo de Faddy, pois é solução do sistema (48), usando-se a seqüência  $\delta_y = \delta_0 = \delta$ .

Assim,

$$p_y(t) = \begin{cases} e^{-\delta t}, & \text{se } y = 0; \\ \delta e^{-\delta t} \int_0^t p_{y-1}(u) e^{\delta u} du, & \text{se } y = 1, 2, \dots \end{cases}$$

mas

$$\begin{aligned} \delta e^{-\delta t} \int_0^t p_{y-1}(u) e^{\delta u} du &= \delta e^{-\delta t} \int_0^t \frac{e^{-\delta u}(\delta u)^{y-1}}{(y-1)!} e^{\delta u} du \\ &= \frac{\delta^y e^{-\delta t}}{(y-1)!} \int_0^t u^{y-1} du \\ &= \frac{\delta^y e^{-\delta t}}{(y-1)!} \frac{t^y}{y} \\ &= \frac{(\delta t)^y e^{-\delta t}}{y!}, \end{aligned}$$

o que prova a afirmação.

(ii) **Distribuição Binomial**

A seqüência  $\delta_y = \lambda(H - y)$  gera a distribuição Binomial  $(H, p = 1 - e^{-\lambda t})$  cuja f.p. é dada por:

$$p_y(t) = \binom{H}{y} (1 - e^{-\lambda t})^y e^{-\lambda(H-y)t} I_{\{0,1,\dots,H\}}(y).$$

Assim

$$p_0(t) = e^{-\lambda H t} = e^{-\delta_0 t}.$$

Logo  $\delta_0 = \lambda H = \lambda(H - 0)$ . Para  $y = 1, 2, \dots, H$ , tem-se:

$$\begin{aligned}
p'_y(t) &= \binom{H}{y} [y(1 - e^{-\lambda t})^{y-1} \lambda e^{\lambda t} e^{-\lambda(H-y)t} - \lambda(H-y)(1 - e^{-\lambda t})^y e^{-\lambda(H-y)t}] \\
&= -\lambda(H-y) \binom{H}{y} (1 - e^{-\lambda t})^y e^{-\lambda(H-y)t} + \lambda y \binom{H}{y} (1 - e^{-\lambda t})^{y-1} \lambda e^{-\lambda(H-y+1)t} \\
&= -\lambda(H-y)p_y(t) + \lambda y \frac{H!}{y!(H-y)!} (1 - e^{-\lambda t})^{y-1} \lambda e^{-\lambda(H-y+1)t} \\
&= -\lambda(H-y)p_y(t) + \lambda \frac{H!}{(y-1)!(H-y)!} (1 - e^{-\lambda t})^{y-1} \lambda e^{-\lambda(H-y+1)t} \\
&= -\lambda(H-y)p_y(t) + \lambda \frac{H!(H-y+1)}{(y-1)!(H-y+1)!} (1 - e^{-\lambda t})^{y-1} \lambda e^{-\lambda(H-y+1)t} \\
&= -\lambda(H-y)p_y(t) + \lambda(H-y+1) \frac{H!}{(y-1)!(H-y+1)!} (1 - e^{-\lambda t})^{y-1} \lambda e^{-\lambda(H-y+1)t} \\
&= -\lambda(H-y)p_y(t) + \lambda(H-y+1) \binom{H}{y-1} (1 - e^{-\lambda t})^{y-1} \lambda e^{-\lambda(H-y+1)t} \\
&= -\lambda(H-y)p_y(t) + \lambda(H-y+1)p_{y-1}(t) \\
&= -\delta_y p_y(t) + \delta_{y-1} p_{y-1}(t).
\end{aligned}$$

### (iii) Distribuição Binomial Negativa

A seqüência  $\delta_y = \lambda(r+y)$  gera a distribuição binomial negativa de parâmetros  $r$  e  $p = e^{-\lambda t}$ , com  $\lambda > 0$ . A f. p. de  $Y(t)$  é dada por:

$$p_y(t) = \frac{\Gamma(r+y)}{\Gamma(r)y!} e^{-\lambda r t} (1 - e^{-\lambda t})^y I_{\{0,1,\dots,\infty\}}(y).$$

Assim,

$$p_0(t) = e^{-\lambda r t} = e^{-\delta_0 t}.$$

Logo,  $\delta_0 = \lambda r = \lambda(r+0)$ . Para  $y = 1, 2, \dots, \infty$ , vai-se usar a relação  $\delta_y - \lambda r = \lambda y$ . Assim,

$$\begin{aligned}
p'_y(t) &= \frac{\Gamma(r+y)}{\Gamma(r)y!} [-\lambda r e^{-\lambda r t} (1 - e^{-\lambda t})^y + \lambda y e^{-\lambda t} e^{-\lambda r t} (1 - e^{-\lambda t})^{y-1}] \\
&= \frac{\Gamma(r+y)}{\Gamma(r)y!} [-(\delta_y - \lambda y) e^{-\lambda r t} (1 - e^{-\lambda t})^y + \lambda y e^{-\lambda t} e^{-\lambda r t} (1 - e^{-\lambda t})^{y-1}] \\
&= -\delta_y \frac{\Gamma(r+y)}{\Gamma(r)y!} e^{-\lambda r t} (1 - e^{-\lambda t})^y + \frac{\Gamma(r+y)}{\Gamma(r)y!} \lambda y [e^{-\lambda r t} (1 - e^{-\lambda t})^y + e^{-\lambda t} e^{-\lambda r t} (1 - e^{-\lambda t})^{y-1}]
\end{aligned}$$



$$\begin{aligned}
&= -\delta_y p_y(t) + \frac{\Gamma(r+y)}{\Gamma(r)y!} \lambda y e^{-\lambda r t} (1 - e^{-\lambda t})^{y-1} [1 - e^{-\lambda t} + e^{-\lambda t}] \\
&= -\delta_y p_y(t) + \lambda \frac{(r+y-1)\Gamma(r+y-1)}{\Gamma(r)(y-1)!} e^{-\lambda r t} (1 - e^{-\lambda t})^{y-1} \\
&= -\delta_y p_y(t) + \delta_{y-1} p_{y-1}(t).
\end{aligned}$$

(iv) **Modelo M1-Bakker**

Agora considere a seqüência

$$\delta_y = \begin{cases} \delta_0, & \text{se } y = 0; \\ \delta_1, & \text{se } y = 1, 2, \dots, \end{cases}$$

em que  $\delta_1 < \delta_0$ , e vai-se provar que se obtém a distribuição de Poisson modificada no ponto zero que é a distribuição do modelo M1 de Baker. Introduzindo o tempo  $t$  na distribuição de Baker, dada por (2), (4) e (9), tem-se:

$$p_y(t) = \begin{cases} e^{-\lambda t}, & \text{se } y = 0; \\ \frac{e^{-\lambda t} - e^{-\lambda \delta t}}{1 - \delta}, & \text{se } y = 1; \\ \delta^{y-1} (\delta - 1)^{-y} e^{-\lambda \delta t} \sum_{i=y}^{\infty} \frac{(\lambda t)^i (\delta - 1)^i}{i!}, & \text{se } y = 2, 3, \dots \end{cases}$$

Assim, a solução do sistema é dada por:

$$p_y(t) = \begin{cases} e^{-\delta_0 t}, & \text{se } y = 0; \\ \delta_0 e^{-\delta_1 t} \int_0^t p_0(u) e^{\delta_1 u} du, & \text{se } y = 1; \\ \delta_1 e^{-\delta_1 t} \int_0^t p_{y-1}(u) e^{\delta_1 u} du, & \text{se } y = 2, 3, \dots \end{cases}$$

e para  $y=1$ , tem-se

$$\begin{aligned}
\int_0^t p_0(u) e^{\delta_1 u} du &= \int_0^t e^{-\delta_0 u} e^{\delta_1 u} du \\
&= \int_0^t e^{(\delta_1 - \delta_0)u} du \\
&= \frac{e^{(\delta_1 - \delta_0)t} - 1}{(\delta_1 - \delta_0)}.
\end{aligned}$$

Dessa maneira,

$$\begin{aligned}
p_1(t) &= \delta_1 e^{-\delta_1 t} \int_0^t p_0(u) e^{\delta_1 u} du \\
&= \delta_1 e^{-\delta_1 t} \frac{e^{(\delta_1 - \delta_0)t} - 1}{(\delta_1 - \delta_0)}
\end{aligned}$$

$$\begin{aligned}
&= \frac{\delta_1(e^{-\delta_0 t} - e^{-\delta_1 t})}{(\delta_1 - \delta_0)} \\
&= \frac{e^{-\delta_0 t} - e^{-\delta_1 t}}{1 - \frac{\delta_0}{\delta_1}} \\
&= \frac{e^{-\lambda t} - e^{-\lambda \delta t}}{1 - \delta},
\end{aligned}$$

usando-se a reparametrização  $\lambda = \delta_0$  e  $\delta = \frac{\delta_1}{\delta_0}$ .

Para o caso  $y = 2, 3, \dots$ , deve-se verificar que

$$p_y(t) = \delta^{y-1}(\delta - 1)^{-y} e^{-\lambda \delta t} \sum_{i=y}^{\infty} \frac{(\lambda t)^i (\delta - 1)^i}{i!}$$

é solução de:

$$\begin{aligned}
p'_y(t) &= \delta_{y-1} p_{y-1}(t) - \delta_y p_y(t) \\
&= \delta_1 p_{y-1}(t) - \delta_1 p_y(t) \\
&= \delta_1 [p_{y-1}(t) - p_y(t)].
\end{aligned}$$

Assim,

$$\begin{aligned}
p'_y(t) &= \delta^{y-1}(\delta - 1)^{-y} \left[ -\lambda \delta e^{-\lambda \delta t} \sum_{i=y}^{\infty} \frac{(\lambda t)^i (\delta - 1)^i}{i!} + e^{-\lambda \delta t} \sum_{i=y}^{\infty} \frac{\lambda i (\lambda t)^{i-1} (\delta - 1)^i}{i!} \right] \\
&= \delta^{y-1}(\delta - 1)^{-y} \left[ -\lambda \delta e^{-\lambda \delta t} \sum_{i=y}^{\infty} \frac{(\lambda t)^i (\delta - 1)^i}{i!} + e^{-\lambda \delta t} \sum_{i=y}^{\infty} \frac{\lambda (\lambda t)^{i-1} (\delta - 1)^i}{(i-1)!} \right] \\
&= -\lambda \delta p_y(t) + \lambda \delta \delta^{y-2} (\delta - 1)^{-(y-1)} \sum_{i=y-1}^{\infty} \frac{\lambda i (\lambda t)^{i-1} (\delta - 1)^i}{i!} \\
&= -\lambda \delta p_y(t) + \lambda \delta p_{y-1}(t) \\
&= \lambda \delta [p_{y-1}(t) - p_y(t)] \\
&= \delta_1 [p_{y-1}(t) - p_y(t)],
\end{aligned}$$

o que demonstra completamente a afirmação. Logo, também o modelo  $M8$  de Hemerik é um caso particular do modelo de Faddy com  $\delta_0 = \lambda = \gamma K T$  e  $\delta_1 = \lambda \delta = \gamma K T \delta$ , pois ele é equivalente ao modelo  $M1$  de Bakker conforme já verificado.

(v) **Modelos M3 - M4 de Bakker**

A distribuição de probabilidade de  $S_j$  nos modelos de Bakker é dada por  $P(S_j = 0) = e^{-\lambda}$  e, de uma maneira geral,  $P(S_j = s)$ ,  $s = 1, 2, \dots$ , pode ser escrita como:

$$P(S_j = s) = (-1)^s \delta_1 \dots \delta_{s-1} \sum_{i=0}^s \frac{e^{-\lambda \delta_i}}{\prod_{\substack{j=0 \\ j \neq i}}^s (\delta_i - \delta_j)},$$

já que  $\delta_0 = 1$ . Usando-se a reparametrização  $\delta$  por  $\phi$ , tem-se:

$$P(S_j = s) = (-1)^s \phi_1 \dots \phi_{s-1} \sum_{i=0}^s \frac{e^{-\lambda \phi_i}}{\prod_{\substack{j=0 \\ j \neq i}}^s (\phi_i - \phi_j)},$$

enquanto a forma geral da distribuição de Faddy é dada por:

$$p_y(t) = (-1)^y \delta_0 \delta_1 \dots \delta_{y-1} \sum_{i=0}^y \frac{e^{-\delta_i t}}{\prod_{\substack{j=0 \\ j \neq i}}^y (\delta_i - \delta_j)}.$$

Assim, fazendo no modelo de Faddy  $\phi_i = \frac{\delta_i}{\delta_0}$ ,  $i = 1, 2, \dots$ , têm-se os modelos de Bakker com probabilidade de evitar o superparasitismo  $\phi_i$ .

(vi) **Modelo M14 - Janardan**

Vai-se mostrar que o modelo M14 de Janardan é um caso particular do modelo de Faddy. Como  $P_0(t) = e^{-\lambda t}$ , então  $\delta_0 = \lambda$  e

$$P_1(t) = \frac{\lambda}{p\beta} [e^{-\lambda t} - e^{-\alpha t}] = \frac{\delta_0 [e^{-\delta_0 t} - e^{-\delta_1 t}]}{(\delta_1 - \delta_0)},$$

e, portanto,  $\delta_1 - \delta_0 = p\beta$ . Tem-se que provar que  $\delta_1 = \alpha$ .

Mas  $\alpha = \lambda q + \mu p$  e  $\beta = \mu - \lambda$ . Assim,

$$\begin{aligned}\delta_1 &= \delta_0 + p\beta = \lambda + p(\mu - \lambda) \\ &= \delta_0 + p\mu - p\lambda \\ &= \lambda + \alpha - \lambda q - \lambda p \\ &= \lambda + \alpha - \lambda(p + q) \\ &= \alpha,\end{aligned}$$

já que  $p + q = 1$  e  $\alpha - \lambda q = \mu p$ .

Para  $y = 2$ , tem-se  $u_2(t) = v_2(t) = 1$  e, assim:

$$p_2(t) = \frac{\alpha\lambda e^{-\mu t}}{pq\beta^2} [q(e^{\beta t} - u_2(t)) - (e^{q\beta t} - v_2(t))] = \frac{\alpha\lambda e^{-\mu t}}{pq\beta^2} [q(e^{\beta t} - 1) - (e^{q\beta t} - 1)].$$

Derivando  $p_2(t)$  em relação a  $t$ , tem-se

$$\begin{aligned}p_2'(t) &= -\mu \frac{\alpha\lambda}{pq\beta^2} e^{-\mu t} [q(e^{\beta t} - 1) - (e^{q\beta t} - 1)] + \frac{\alpha\lambda}{pq\beta^2} e^{-\mu t} [q\beta e^{\beta t} - q\beta e^{q\beta t}] \\ &= -\mu p_2(t) + \alpha \frac{\lambda}{pq\beta^2} q\beta e^{-\mu t} [e^{\beta t} - e^{q\beta t}] \\ &= -\mu p_2(t) + \alpha \frac{\lambda}{p\beta} e^{-\mu t} [e^{\beta t} - e^{q\beta t}] \\ &= -\mu p_2(t) + \alpha p_1(t).\end{aligned}$$

Logo,  $\delta_2 = \mu$ . Mas para  $y = 3, 4, \dots$

$$p_y(t) = \frac{\alpha\lambda\mu^{y-2}e^{-\mu t}}{pq^{y-1}\beta^y} [q^{y-1}(e^{\beta t} - u_y(t)) - (e^{q\beta t} - v_y(t))],$$

em que,  $u_y(t) = \sum_{j=0}^{y-2} \frac{\beta^j t^j}{j!}$  e  $v_y(t) = \sum_{j=0}^{y-2} \frac{(q\beta)^j t^j}{j!}$ .

Inicialmente, serão calculadas as derivadas de  $u_y(t)$  e  $v_y(t)$  em relação a  $t$ .

Assim,

$$\begin{aligned}u_y'(t) &= \sum_{j=1}^{y-2} \frac{j\beta^j t^{j-1}}{j!} \\ &= \beta \sum_{j=1}^{y-2} \frac{\beta^{j-1} t^{j-1}}{(j-1)!} \\ &= \beta \sum_{j=0}^{y-3} \frac{\beta^j t^j}{j!} = \beta u_{y-1}(t),\end{aligned}$$

e de maneira similar  $v'_y(t) = q\beta v_{y-1}(t)$ .

Além disso, a  $p_{y-1}(t)$  é dada por:

$$p_{y-1}(t) = \frac{\alpha\lambda\mu^{y-3}e^{-\mu t}}{pq^{y-2}\beta^{y-1}} [q^{y-2}(e^{\beta t} - u_{y-1}(t)) - (e^{q\beta t} - v_{y-1}(t))],$$

em que ,  $u_{y-1}(t) = \sum_{j=0}^{y-3} \frac{\beta^j t^j}{j!}$  e  $v_{y-1}(t) = \sum_{j=0}^{y-3} \frac{(q\beta)^j t^j}{j!}$ .

Finalmente, será mostrado que  $p_y(t)$  satisfaz ao sistema de equações diferenciais do modelo de Faddy. Logo,

$$\begin{aligned} p'_y(t) &= -\mu \frac{\alpha\lambda\mu^{y-2}}{pq^{y-1}\beta^y} e^{-\mu t} [q^{y-1}(e^{\beta t} - u_y(t)) - (e^{q\beta t} - v_y(t))] + \\ &\quad + \frac{\alpha\lambda\mu^{y-2}}{pq^{y-1}\beta^y} e^{-\mu t} [q^{y-1}(\beta e^{\beta t} - u'_y(t)) - (q\beta e^{q\beta t} - v'_y(t))] \\ &= -\mu p_y(t) + \frac{\alpha\lambda\mu^{y-2}}{pq^{y-1}\beta^y} e^{-\mu t} [q^{y-1}(\beta e^{\beta t} - \beta u_{y-1}(t)) - (q\beta e^{q\beta t} - q\beta v_{y-1}(t))] \\ &= -\mu p_y(t) + \mu \frac{\alpha\lambda\mu^{y-2}}{pq^{y-1}\beta^y} q\beta e^{-\mu t} [q^{y-2}(\beta e^{\beta t} - u_{y-1}(t)) - (e^{q\beta t} - v_{y-1}(t))] \\ &= -\mu p_y(t) + \mu \frac{\alpha\lambda\mu^{y-3}}{pq^{y-2}\beta^{y-1}} e^{-\mu t} [q^{y-2}(e^{\beta t} - u_{y-1}(t)) - (e^{q\beta t} - v_{y-1}(t))] \\ &= -\mu p_y(t) + \mu p_{y-1}(t). \end{aligned}$$

Assim, para  $y = 3, 4, \dots, \delta_y = \mu$ .

Portanto, a seqüência dos coeficientes  $\delta$ 's que torna o modelo de Janardan um caso particular do modelo de Faddy é:

$$\delta_y = \begin{cases} \lambda & \text{se } y = 0; \\ \alpha & \text{se } y = 1; \\ \mu & \text{se } y = 2, 3, \dots \end{cases}$$

O modelo de Janardan generaliza o modelo  $M1$  de Bakker, quando  $\mu = \alpha$ .

### 2.4.3 Modelagem das taxas de transição

Para construir um modelo mais geral para as taxas de transição é preciso incorporar duas formas de dependências:

- i. a dependência em  $y$  para controlar as propriedades de dispersão do modelo;

ii. o efeito de covariáveis.

Sejam  $\mathbf{z}'_i$ , o vetor de covariáveis da  $i$ -ésima observação e  $\boldsymbol{\beta}$  o vetor de coeficientes associados às covariáveis. Considere  $a(\cdot)$ , uma função dos efeitos lineares das covariáveis e  $h(y)$  uma função para introduzir no modelo a forma da dependência em  $y$  das taxas de transição. Podlich (1999) propôs a seguinte estrutura multiplicativa.

$$\delta_y = a(\mathbf{z}'_i\boldsymbol{\beta})h(y). \quad (52)$$

Além disso, para evitar superparametrização, Podlich (1999) sugere que se omita o termo constante da dependência das covariáveis ou se imponha que  $h(0) = 1$ . Para a forma paramétrica, ela também sugere que uma função exponencial possa ser usada levando a uma forma log-linear para as covariáveis. Assim, a expressão (52) toma a forma

$$\delta_y = \exp(\mathbf{z}'_i\boldsymbol{\beta})h(y). \quad (53)$$

Se  $h(y) = 1$ , tem-se o modelo log-linear de Poisson para a média. Uma formulação totalmente paramétrica poderia também parametrizar a função  $h(y)$  em (53) com a introdução de pelo menos um parâmetro para controlar a dispersão do modelo. No modelo com taxas de transição em que  $\delta_y = \lambda(r + y)^c$  pode-se fazer  $\lambda = \exp(\mathbf{z}'_i\boldsymbol{\beta})$  e  $h(y) = (r + y)^c$ . Casos especiais desse modelo serão especificados.

- i. Se  $c = 0$  tem-se o modelo de Poisson;
- ii. Se  $r > 0$ ,  $c = 1$  tem-se o modelo binomial negativo;
- iii. Se a seqüência for crescente côncava com  $0 < c < 1$ , a distribuição resultante tem dispersão entre a Poisson e a Binomial negativa;
- iv. Se a seqüência for decrescente côncava com  $c < 0$ , a distribuição resultante tem dispersão menor do que a Poisson.

Uma situação muito comum em análise de dados biológicos é a existência de excessos de contagens nulas (zeros). Para incorporar esse fato é preciso apenas modelar a

taxa de transição quando  $y = 0$  separadamente das taxas de transição para  $y \geq 1$ . Para um modelo com dependência log-linear nas covariáveis, pode-se ter:

$$\begin{aligned}\delta_0 &= \exp[\mathbf{z}'_0 \boldsymbol{\beta}_0] \\ \delta_y &= \exp[\mathbf{z}'_1 \boldsymbol{\beta}_1] h(y), y = 1, 2, 3, \dots,\end{aligned}$$

em que  $\boldsymbol{\beta}_0$  e  $\boldsymbol{\beta}_1$  são vetores dos coeficientes de covariáveis permitindo uma estrutura diferente para  $y = 0$  e  $y \geq 1$ .

Vai-se agora analisar a estrutura :

$$\delta_y = \begin{cases} \delta_0, & \text{se } y = 0; \\ \delta_1, & \text{se } y = 1, 2, \dots, \end{cases} \quad (54)$$

Seja  $G(s, t) = \sum_{y=0}^{\infty} p_y(t) s^y$ , a f.g.p do processo estocástico de nascimento puro cuja seqüência da taxas de transição é dada por (54). Assim,

$$\frac{\partial G}{\partial t} = \sum_{y=0}^{\infty} p'_y(t) s^y. \quad (55)$$

Usando-se as equações de Chapman-Kolmogorov, tem-se

$$\begin{aligned}\frac{\partial G}{\partial t} &= p'_0(t) + p'_1(t)s + \sum_{y=2}^{\infty} p'_y(t)s^y \\ &= -\delta_0 p_0(t) + [\delta_0 p_0(t) - \delta_1 p_1(t)]s + \sum_{y=2}^{\infty} [\delta_1 p_{y-1}(t) - \delta_1 p_y(t)]s^y \\ &= -\delta_0 p_0(t) + [\delta_0 p_0(t) - \delta_1 p_1(t)]s + \delta_1 \left[ \sum_{y=2}^{\infty} p_{y-1}(t) s^y \right] - \left[ \sum_{y=2}^{\infty} p_y(t) s^y \right] \\ &= \delta_0 p_0(t) + [\delta_0 p_0(t) - \delta_1 p_1(t)]s + \delta_1 \left[ s \sum_{y=2}^{\infty} p_{y-1}(t) s^{y-1} \right] - \delta_1 \left[ \sum_{y=2}^{\infty} p_y(t) s^y \right] \\ &= \delta_0 p_0(t) + [\delta_0 p_0(t) - \delta_1 p_1(t)]s + \delta_1 \left[ s \sum_{y=1}^{\infty} p_y(t) s^y \right] - \delta_1 \left[ \sum_{y=2}^{\infty} p_y(t) s^y \right] \\ &= \delta_0 p_0(t) + [\delta_0 p_0(t) - \delta_1 p_1(t)]s + \delta_1 [sG(s, t) - s p_0(t)] - \delta_1 [G(s, t) - p_0(t) - p_1(t)s] \\ &= \delta_0 p_0(t) + [\delta_0 p_0(t) - \delta_1 p_1(t)]s + \delta_1 [(s-1)G(s, t) - (s-1)p_0(t) + p_1(t)s] \\ &= (s-1)[\delta_1 G(s, t) + (\delta_0 - \delta_1)p_0(t)] \\ &= (s-1)[\delta_1 G(s, t) + (\delta_0 - \delta_1)e^{-\delta_0 t}]. \quad (56)\end{aligned}$$

Derivando-se (56) em relação a  $s$ , tem-se:

$$\frac{\partial^2 G}{\partial s \partial t} = \delta_1 G(s, t) + (\delta_0 - \delta_1)e^{-\delta_0 t} + (s - 1)\delta_1 G'(s, t). \quad (57)$$

Fazendo  $s = 1$  em (57), tem-se a equação diferencial para a média  $m_1(t)$

$$\frac{dm_1(t)}{dt} = (\delta_0 - \delta_1)e^{-\delta_0 t} + \delta_1, \quad (58)$$

com  $m_1(0) = 0$ . Diferenciando-se novamente (57) com relação a  $s$ , obtém-se:

$$\frac{\partial^3 G}{\partial^2 s \partial t} = 2\delta_1 G'(s, t) + (s - 1)\delta_1 G''(s, t). \quad (59)$$

Fazendo  $s = 1$  em (59), tem-se a equação diferencial para o segundo momento fatorial  $m_2(t)$ ,

$$\frac{dm_2(t)}{dt} = 2\delta_1 m_1(t), \quad (60)$$

com  $m_2(0) = 0$ .

Resolvendo-se a equação diferencial (58), tem-se

$$\begin{aligned} \int_0^t \frac{dm_1(u)}{du} du &= \int_0^t [(\delta_0 - \delta_1)e^{-\delta_0 u} + \delta_1] du \\ m_1(t) - m_1(0) &= \delta_1 t + \left(1 - \frac{\delta_1}{\delta_0}\right)(1 - e^{-\delta_0 t}) \\ m_1(t) &= \delta_1 t + \left(1 - \frac{\delta_1}{\delta_0}\right)(1 - e^{-\delta_0 t}). \end{aligned} \quad (61)$$

Resolvendo-se a equação diferencial (60), tem-se

$$\begin{aligned} \int_0^t \frac{dm_2(u)}{du} du &= \int_0^t 2\delta_1 m_1(u) du \\ m_2(t) - m_2(0) &= \int_0^t 2\delta_1 \left[ \delta_1 u + \left(1 - \frac{\delta_1}{\delta_0}\right)(1 - e^{-\delta_0 u}) \right] du \\ m_2(t) &= \delta_1^2 t^2 + 2\delta_1 \left(1 - \frac{\delta_1}{\delta_0}\right) \left[ t + \frac{1}{\delta_0} [e^{-\delta_0 t} - 1] \right] \\ &= \delta_1^2 t^2 + 2\frac{\delta_1}{\delta_0} \left(1 - \frac{\delta_1}{\delta_0}\right) [\delta_0 t + e^{-\delta_0 t} - 1] \end{aligned} \quad (62)$$

Usando-se (61) e (62), tem-se que a  $V[Y(t)]$  é dada por



$$V[Y(t)] = \sigma^2 = (1 - e^{-\delta_0 t}) \left(1 - \frac{\delta_0}{\delta_1}\right) \left[1 - 2\frac{\delta_1}{\delta_0} - \left(1 - \frac{\delta_0}{\delta_1}\right) (1 - e^{-\delta_0 t}) - 2\delta_1 t\right] + \delta_1 t \left(3 - 2\frac{\delta_1}{\delta_0}\right). \quad (63)$$

Considere agora  $D(t) = Var[Y(t)] - E[Y(t)]$

$$D(t) = (1 - e^{-\delta_0 t}) \left\{ \left(1 - \frac{\delta_0}{\delta_1}\right) \left[1 - 2\frac{\delta_1}{\delta_0} - \left(1 - \frac{\delta_0}{\delta_1}\right) (1 - e^{-\delta_0 t}) - 2\delta_1 t\right] + \frac{\delta_1}{\delta_0} - 1 \right\} + 2\delta_1 t \left(1 - \frac{\delta_1}{\delta_0}\right). \quad (64)$$

### Relação média-variância

De (61), tem-se  $m_1(t) - \delta_1 t = \mu - \delta_1 t = (1 - \frac{\delta_1}{\delta_0})(1 - e^{-\delta_0 t})$ , que substituído em (63) resulta em:

$$\begin{aligned} V[Y(t)] = \sigma^2 &= (\mu - \delta_1 t) \left[1 - 2\frac{\delta_1}{\delta_0} - \mu + \delta_1 t - 2\delta_1 t\right] + \delta_1 t \left(3 - 2\frac{\delta_1}{\delta_0}\right) \\ &= -\mu^2 + \left(1 - 2\frac{\delta_1}{\delta_0}\right)\mu + 2\delta_1 t + \delta_1^2 t^2. \end{aligned}$$

Assim, a variância é uma função quadrática da média. Pode-se mostrar diretamente que para  $\delta_0 < \delta_1$ , a variância é maior do que a média (superdispersão em relação a Poisson) enquanto que para  $\delta_0 > \delta_1$  a variância é menor do que a média (subdispersão em relação a Poisson). Vai-se agora analisar a estrutura :

$$\delta_y = \begin{cases} \delta_0 & \text{se } y = 0; \\ \delta_1 y & \text{se } y = 1, 2, \dots \end{cases}, \quad (65)$$

que é um caso particular com  $c = 1$  da estrutura

$$\delta_y = \begin{cases} \delta_0 & \text{se } y = 0; \\ \delta_1 y^c & \text{se } y = 1, 2, \dots \end{cases}$$

Seja  $G(s, t) = \sum_{y=0}^{\infty} p_y(t) s^y$  a f.g.p do processo estocástico de nascimento puro cuja seqüência da taxas de transição é dada por (65). Então,

$$\begin{aligned}
\frac{\partial G}{\partial t} &= \sum_{y=0}^{\infty} p'_y(t) s^y \\
&= p'_0(t) + p'_1(t) + \sum_{y=2}^{\infty} [\delta_{y-1} p_{y-1}(t) - \delta_y p_y(t)] s^y \\
&= -\delta_0 p_0(t) + [\delta_0 p_0(t) - \delta_1 p_1(t)] s + \sum_{y=2}^{\infty} [\delta_1 (y-1) p_{y-1}(t) - \delta_1 y p_y(t)] s^y \\
&= (s-1) \delta_0 p_0(t) - \delta_1 p_1(t) s + \sum_{y=2}^{\infty} \delta_1 (y-1) p_{y-1}(t) s^y - \sum_{y=2}^{\infty} \delta_1 y p_y(t) s^y \\
&= (s-1) \delta_0 p_0(t) - \delta_1 p_1(t) s + \delta_1 \left[ s \sum_{y=1}^{\infty} y p_y(t) s^y - \sum_{y=1}^{\infty} y p_y(t) s^y + p_1(t) s \right] \\
&= (s-1) \delta_0 p_0(t) - \delta_1 p_1(t) s + \delta_1 [(s-1) \sum_{y=1}^{\infty} y p_y(t) s^y + p_1(t) s] \\
&= (s-1) \delta_0 p_0(t) + \delta_1 (s-1) \sum_{y=1}^{\infty} y p_y(t) s^y \\
&= (s-1) \delta_0 p_0(t) + \delta_1 (s-1) s \sum_{y=1}^{\infty} y p_y(t) s^{y-1} \\
&= (s-1) \delta_0 p_0(t) + \delta_1 (s-1) s \sum_{y=1}^{\infty} [p_y(t) s^y]' \\
&= (s-1) \delta_0 p_0(t) + \delta_1 (s-1) s \frac{\partial G}{\partial t} \\
&= (s-1) \left[ \delta_0 e^{-\delta_0 t} + \delta_1 s \frac{\partial G}{\partial t} \right], \tag{66}
\end{aligned}$$

pois

$$\frac{\partial G}{\partial t} = \sum_{y=0}^{\infty} p'_y(t) s^y.$$

Pode ser mostrado ainda que os dois primeiros momentos fatoriais satisfazem as seguintes equações diferenciais:

$$\frac{dm_1(t)}{dt} = \delta_0 e^{-\delta_0 t} + \delta_1 m_1(t) \tag{67}$$

$$\frac{dm_2(t)}{dt} = 2\delta_1 (m_1(t) + m_2(t)), \tag{68}$$

e portanto,

$$m_1(t) = \frac{\delta_0}{\delta_0 + \delta_1} (e^{-\delta_1 t} - e^{-\delta_0 t})$$

$$m_2(t) = \frac{2\delta_0}{\delta_0 + 2\delta_1} e^{2\delta_1 t} + \frac{2\delta_0}{\delta_0 + \delta_1} \left[ \frac{\delta_1}{\delta_0 + 2\delta_1} e^{-\delta_0 t} - e^{\delta_1 t} \right],$$

sendo a média e a variância dadas por:

$$m_1(t) = \frac{\delta_0}{\delta_0 + \delta_1} (e^{-\delta_1 t} - e^{-\delta_0 t})$$

$$\sigma^2(t) = \frac{2\delta_0}{\delta_0 + 2\delta_1} e^{2\delta_1 t} - \frac{\delta_0}{\delta_0 + \delta_1} \left( \frac{\delta_0}{\delta_0 + 2\delta_1} e^{-\delta_0 t} + e^{\delta_1 t} \right) - \frac{\delta_0^2}{(\delta_0 + \delta_1)^2} (e^{\delta_1 t} - e^{-\delta_0 t})^2.$$

Agora, a seqüência,

$$\delta_y = \begin{cases} \delta_0 & \text{se } y = 0; \\ \delta_1 y & \text{se } y = 1, 2, \dots, \end{cases}$$

é convexa crescente para  $\delta_0 < \delta_1$ . Assim, pela conjectura de Faddy tem-se:

$$\sigma^2 > \mu \frac{\delta(\mu)}{\delta(0)} = \mu \frac{\delta_1(\mu)}{\delta_0} = \frac{\delta_1}{\delta_0} \mu^2.$$

#### 2.4.4 Diagrama de Podlich

Usando essa teoria, é proposto um método gráfico para a seleção de modelos. Como já demonstrado, se a seqüência de taxas de transição  $\delta_y, y = 0, 1, \dots$ , for constante, origina a distribuição de Poisson, enquanto que se a seqüência de taxas de transição for linear crescente aparece a distribuição binomial negativa e se for linear crescente a binomial. Parece razoável, então, segundo Podlich (1999), sugerir que uma seqüência de taxas de transição  $\delta_y, y = 0, 1, \dots$ , que cresce a uma taxa entre uma função constante e uma função afim possa corresponder a uma distribuição com dispersão que de alguma maneira esteja “entre” a Poisson e a binomial negativa, enquanto que se a seqüência de taxas de transição  $\delta_y, y = 0, 1, \dots$  cresce a uma taxa menor do que a linear corresponde a uma distribuição com

uma dispersão maior do que a associada à binomial negativa. Faddy (1997b) conjecturou que qualquer seqüência crescente côncava  $\delta_y, y = 0, 1, \dots$  (uma seqüência que tem diferenças de segunda ordem negativas) corresponde a uma dispersão entre a Poisson e a binomial negativa, enquanto que, qualquer seqüência crescente convexa  $\delta_y, y = 0, 1, \dots$ , (uma seqüência que tem diferenças de segunda ordem positivas) corresponde a uma dispersão maior do que aquela associada à binomial negativa. Esta conjectura ainda está em aberto no mundo científico esperando uma prova formal, mas na prática em análise de dados de campo ela sempre tem sido verificada de acordo com Faddy (1997b) e Podlich(1999).

Uma distribuição binomial negativa com uma seqüência de taxas de transição linear crescente  $\delta_y = \lambda(r + y), y = 0, 1, \dots$ , com  $p = e^{-\lambda t}$ , tem a variância relacionada com a média  $\mu$ , isto é,

$$V(Y) = \frac{rq}{p^2} = \mu \frac{1}{p}.$$

Como  $\mu = r \frac{1-p}{p}$  tem-se que  $\frac{1}{p} = 1 + \frac{\mu}{r}$ . Assim:

$$V(Y) = \frac{rq}{p^2} = \mu \frac{1}{p} = \mu \left(1 + \frac{\mu}{r}\right) = \mu \frac{\lambda(r + \mu)}{\lambda r} = \mu \frac{\delta_\mu}{\delta_0}.$$

A conjectura de Faddy diz, então que, para a seqüência de taxas de transição crescente convexa  $\delta_y, y = 0, 1, 2, \dots$ , a variância da distribuição resultante é maior do que a função da média  $\mu \frac{\delta_\mu}{\delta_0}$ , e para seqüência de taxas de transição crescente côncava  $\delta_y, y = 0, 1, 2, \dots$ , com  $\mu < V(Y) < \mu \frac{\delta_\mu}{\delta_0}$ .

Com dados discretos de campo, oriundos de contagem, há uma única seqüência de taxas de transição que corresponde às freqüências relativas das contagens observadas,  $f_0, f_1, \dots, f_{y_{max}}$ , em que  $y_{max}$  é o maior valor observado da variável na amostra. Essa seqüência é chamada de taxas de transição empíricas,  $\hat{\delta}_y$ , que são obtidas usando a exponencialização da matriz  $Q$ . Inicia-se o processo com  $\hat{\delta}_0 = -\ln(f_0)$  e, então, seqüencialmente, resolve-se para  $\hat{\delta}_y, y = 1, 2, \dots, y_{max}$ , com  $\hat{\delta}_{y_{max}} = 0$ . Por outro lado, pode-se pensar cada  $\delta_y$  como um parâmetro e maximizar o logaritmo da função de máxima verossimilhança sobre o espaço paramétrico e as estimativas obtidas correspondem à seqüência empírica (Podlich 1999). Essas estimativas podem ser obtidas, utilizando-se uma subrotina em S-Plus escrita pela autora. Estimadas as taxas de transição pode-se usar a expressão (50) para calcular as probabilidades

esperadas sob o modelo de Faddy.

A representação gráfica, chamada “lambda-grams” por Podlich (1999) e aqui deltagrama, das taxas empíricas de transição,  $\hat{\delta}_y$ , como função dos  $y$ 's,  $y=0,1,2,\dots$ , dá uma idéia aproximada sobre a possível dependência entre eles e do tipo de modelo a ser usado.

### 2.4.5 Aplicações

Como exemplos de aplicação, foram estimadas as taxas de transição, as frequências esperadas para o modelo de Faddy e feitos os diagramas de Poodlich para os dados das Tabelas 1, 2 e 3 e os programas utilizados estão no anexo B. De uma forma geral, verifica-se que o diagrama de Poodlich tem sua interpretação comprometida em função do número de pontos para esse tipo de dados.

As taxas de transição estimadas para os dados de cada ensaio de Bakker (Tabela 1) estão na Tabela 6, enquanto que as frequências esperadas por diferentes modelos de Bakker e pelo modelo de Faddy estão nas Tabelas 7 a 12. Verifica-se que as frequências esperadas pelo modelo de Faddy, praticamente, reproduzem as frequências observadas. Os diagramas de Podlich são apresentados nas Figuras 1 a 6.

Tabela 6 - Estimativas das taxas de transição para os dados da Tabela 1

Ensaio	Estimativas das taxas de nascimento					
	$\hat{\delta}_0$	$\hat{\delta}_1$	$\hat{\delta}_2$	$\hat{\delta}_3$	$\hat{\delta}_4$	$\hat{\delta}_5$
1	0,9162	0,3243	0,5165	0,0000	—	—
2	1,8124	0,4997	0,5271	0,0000	—	—
3	3,3557	1,1854	0,3035	0,0000	—	—
4	4,5108	1,8225	1,0828	0,2260	0,0000	—
5	4,4773	2,7486	1,1259	0,4683	0,0010	0,0000
6	4,6913	3,3298	1,2091	0,4425	0,6064	0,0000

Tabela 7 - Freqüências esperadas para os dados da Tabela 1 - Ensaio 1

ovos	observada	Poisson	M1	Faddy
0	24	29,2	23,9	24,0022
1	30	21,0	29,6	29,9978
2	5	7,7	5,7	4,9978
$\geq 3$	1	2,1	1,0	1,0022

Tabela 8 - Freqüências esperadas para os dados da Tabela 1 - Ensaio 2

ovos	observada	Poisson	M1	Faddy
0	8	15,9	8	7,9998
1	30	18,0	29,9	30,0006
2	9	9,9	9,2	8,9992
$\geq 3$	2	5,2	1,9	2,0004

Tabela 9 - Freqüências esperadas para os dados da Tabela 1 - Ensaio 3

ovos	observada	Poisson	M1	M2	Faddy
0	3	17,8	2,5	3,0	3,0001
1	36	28,0	42,3	35,9	36,0008
2	41	22,2	27,7	41,2	40,9984
$\geq 3$	6	18,0	13,5	5,9	6,0007

Estimando-se as taxas de transição, usando-se os dados da Tabela 2, obtém-se  $\hat{\delta}_0 = 2,5609$ ,  $\hat{\delta}_1 = 0,4343$ ,  $\hat{\delta}_2 = 0,3722$ ,  $\hat{\delta}_3 = 1,0769$ ,  $\hat{\delta}_4 = 3,8533$ , enquanto que as freqüências esperadas ajustando uma distribuição de Poisson e o modelo de Faddy estão nas Tabelas 13 e percebe-se claramente a adequação do modelo. O diagrama de Podlich é apresentado na Figura 7.

As taxas de transição estimadas para os dados da (Tabela 3) estão na Tabela 14,

Tabela 10 - Freqüências esperadas para os dados da Tabela 1 - Ensaio 4

ovos	observada	Poisson	$M1$	$M2$	$M3$	$M4$	Faddy
0	1	11,7	0,7	0,9	1,0	1,3	1,0006
1	23	24,0	29,6	22,1	21,9	21,0	23,0003
2	40	24,6	31,7	43,1	43,1	43,9	32,5950
3	25	16,8	18,5	18,4	20,3	21,5	24,9997
$\geq 4$	2	3,9	10,5	6,5	4,7	3,3	9,4044

Tabela 11 - Freqüências esperadas para os dados da Tabela 1 - Ensaio 5

ovos	observada	Poisson	$M1$	$M2$	$M3$	$M4$	Faddy
0	1	8,8	0,5	0,9	0,8	0,2	1,0000
1	12	20,3	24,1	11,5	14,4	11,5	12,0007
2	42	23,3	29,6	45,2	39,5	42,4	42,0010
3	28	17,9	19,9	21,9	25,3	28,0	27,9997
$\geq 4$	5	17,7	13,9	8,5	6,0	5,9	4,9986

Tabela 12 - Freqüências esperadas para os dados da Tabela 1 - Ensaio 6

ovos	observada	Poisson	$M1$	$M2$	$M3$	$M4$	Faddy
0	1	9,5	0,5	0,9	0,7	0,1	1,0000
1	10	23,2	27,2	9,3	14,4	10,3	10,0000
2	51	28,1	35,9	56,4	47,0	49,7	51,0007
3	40	22,7	25,7	29,8	34,4	39,8	39,9989
$\geq 4$	7	25,5	19,7	12,6	12,5	9,3	7,0004

enquanto que as freqüências esperadas ajustando uma distribuição de Poisson e os modelos de Janardan e de Faddy estão nas Tabelas 15 e 16 e os deltagramas nas Figuras 8 e 9. Mais uma vez fica evidenciada a qualidade do ajuste do modelo de Faddy.

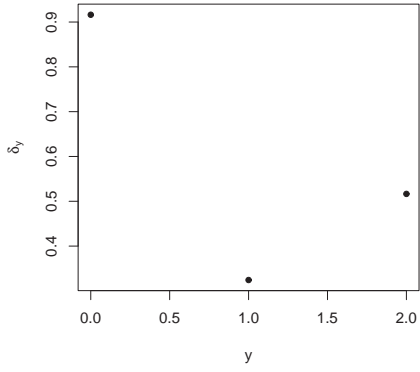


Figura 1 - Deltagrama, Tabela 1 - E1

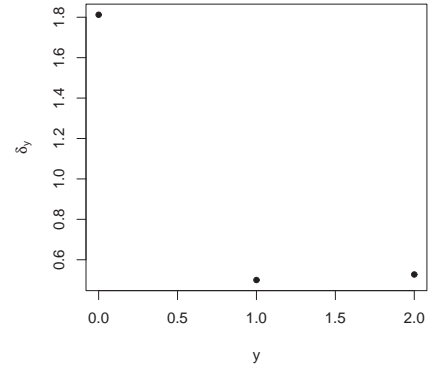


Figura 2 - Deltagrama, Tabela 1 - E2

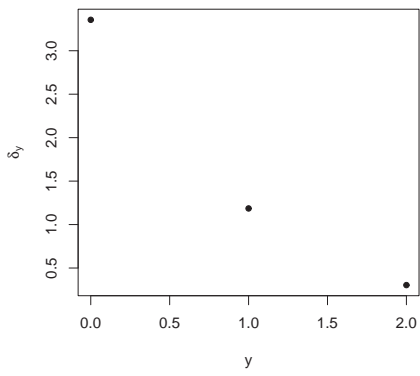


Figura 3 - Deltagrama, Tabela 1 - E3

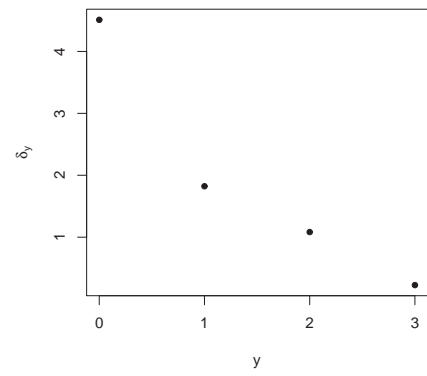


Figura 4 - Deltagrama, Tabela 1 - E4

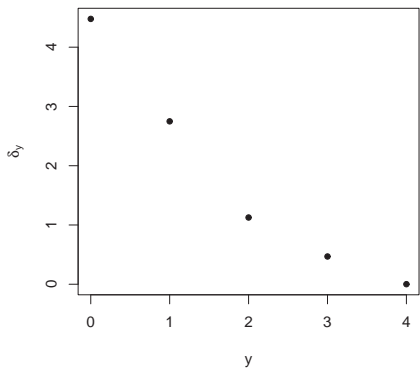


Figura 5 - Deltagrama, Tabela 1 - E5

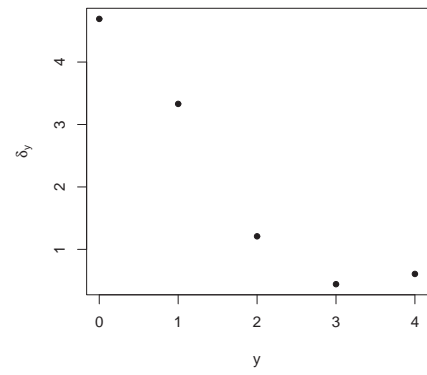


Figura 6 - Deltagrama, Tabela 1 - E6



Tabela 13 - Frequências esperadas para os dados da Tabela 2

ovos	Observados	Poisson	Faddy
0	19	73,85	19
1	169	88,86	169
2	50	53,46	50
3	6	21,44	6
4	1	6,45	1
$\geq 5$	1	1,94	1
Total	246	246	246

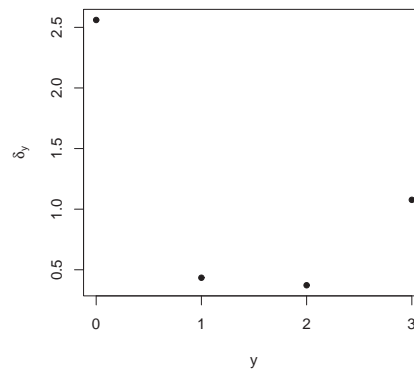


Figura 7 - Deltagrama, Tabela 2

Tabela 14 - Estimativas das taxas de transição para os dados da Tabela 3

Ensaio	Estimativas das taxas de nascimento				$\hat{p}$
	$\hat{\delta}_0 = \hat{\lambda}$	$\hat{\delta}_1$	$\hat{\delta}_2 = \hat{\mu}$	$\hat{\delta}_3$	
2	1,99293	0,4214	0,0715	0,0000	0,8116
3	3,6533	1,4229	0,6825	0,0000	0,7506

Tabela 15 - Frequências esperadas para os dados da Tabela 3 - Ensaio 2

ovos	Observados	Poisson	Modelo Janardan	Modelo de Faddy
0	26	62,00	26,00	26,0000
1	117	65,70	117,00	116,9999
2	35	34,90	35,98	34,9999
$\geq 3$	1	26,40	0,92	1,0012
Total	179	179	179	179

Tabela 16 - Frequências esperadas para os dados da Tabela 3 - Ensaio 3

ovos	Observados	Poisson	Modelo Janardan	Modelo de Faddy
0	5	33,10	5,00	5,0000
1	68	58,40	68,00	67,9998
2	88	51,40	92,87	88,0000
$\geq 3$	32	50,10	27,13	32,0002
Total	193	193	193	193

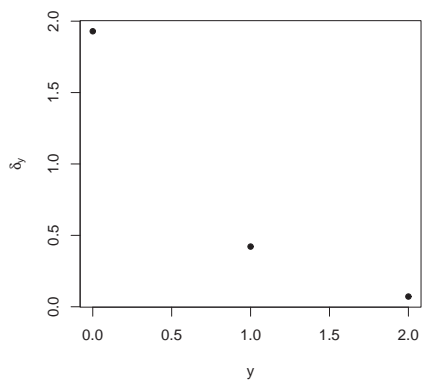


Figura 8 - Deltagrama, Tabela 3 - E2

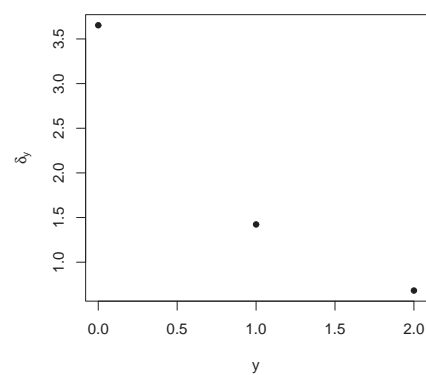


Figura 9 - Deltagrama, Tabela 3 - E3

### 3 Considerações Finais

Nesse trabalho, foram descritos 19 modelos para analisar dados oriundos do estudo de parasitismo em que a variável resposta era o número de ovos postos por um parasita em um determinado hospedeiro. Esses modelos diferem quanto à distribuição do número de encontros,  $(R_J)$ , que pode ser Poisson (Bakker, 1972; Griffiths, 1977a,b; Hemerick, Dayle; Maindonald, 1989), Binomial Negativa (Hemerick, ) e Binomial (Rogers, 1975; Griffiths, 1977) e do modo como o parasita seleciona o hospedeiro.

Foi mostrado que o modelo de Hemerick corresponde ao modelo de Bakker quando se supõe que o parâmetro  $\lambda$  da distribuição de Poisson referente ao número de encontros é uma variável aleatória com distribuição gama. Assim, tem-se uma explicação probabilística para tais modelos, além da explicação biológica apresentada pelos autores.

Ainda foi encontrado na literatura o modelo de Faddy (1997), geralmente, empregado no estudo do tamanho de população animal, que foi adaptado para generalizar esses modelos. Seu uso tem uma vantagem adicional pois permite a incorporação de covariáveis. Na realidade, ao se generalizar o processo de Poisson tem-se uma maneira diferente de se analisar dados discretos que exibam comportamento distante da Poisson. A natureza desse afastamento é explicada através da forma da dependência de  $y$  em relação às taxas de transição do processo de nascimento envolvido. Qualquer distribuição discreta tem representação única em termos de um processo de Poisson estendido. Isso gera uma grande flexibilidade de aplicação desse modelo. Foi ainda feita a obtenção dos coeficientes do processo de Poisson estendido para os diversos modelos.

Extensões deste trabalho devem levar em consideração a dependência do tempo nos coeficientes  $\delta'_s$  do modelo de Faddy, isto é, supor que a probabilidade de um nascimento no intervalo  $(t, t + \Delta t)$  é  $\lambda_y \Delta t + o(\Delta t)$  em que  $\lambda_y$  depende de  $y$  e  $t$ . Isso equivale a supor um processo de Polya no lugar do processo de nascimento puro. Outra abordagem interessante, seria provar mais uma conjectura de Faddy (1997): “Se a seqüência  $\delta_y$  for crescente convexa gera uma distribuição com dispersão maior do que a da binomial negativa enquanto se a seqüência  $\delta_y$  for crescente côncava gera uma distribuição com dispersão entre a Poisson e a binomial negativa”.

Ainda, um desafio teórico será obter a função de probabilidade do modelo de

Hemerik para o caso em que o número de encontros tenha uma distribuição binomial negativa e os parasitas reconheçam seus próprios ovos.

Finalmente, seria de grande interesse incorporar esses modelos na modelagem de outras variáveis respostas envolvidas no estudo de parasitismo, como, por exemplo, número de insetos e proporção de fêmeas.

## REFERÊNCIAS

- APOSTOL, T.M. **Calculus**. 2.ed. Singapore: John Wiley, 1969. v. 2, 673 p.
- BAKKER, K.; BAGCHEE, S.N.; VAN ZWET, W.R.; MEELIS, E. Host discrimination in *Pseudocoila Bochei* (Hymenoptera: Cynipidae). **Entomologia Experimentalis et Applicata**, Amsterdam, v. 10, p. 295 – 311, 1967.
- BAKKER, K.; EIJSACKERS, H.J.P.; VAN LENTEREN, J.C.; MEELIS, E. Some models describing the distribution of eggs of the parasite *Pseudocoila Bocheii* (Hymenoptera: Cynipidae) over its host, larvae of *Drosophila melanogaster*. **Oecologia**, Berlin, v.10, p. 29 – 57, 1972.
- BORGATTO, A.F. **Modelos para proporções com superdispersão e excessos de zeros** : um procedimento Bayesiano. 2004. 90 p. Tese (Doutorado em Estatística e Experimentação Agrônômica) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2004.
- COX, D.R.; MILLER, H.D. **The theory of stochastic processes**. London: Methuen, 1965. 500 p.
- DALEY, D.J.; MAINDONALD, J.H. A unified view of models describing the avoidance of superparasitism. **IMA Journal of Mathematics in Medicine & Biology**, London, v. 6, p. 161 – 178, 1989.
- DANIELS, H.E. The saddlepoint approximation for general birth process. **Journal of Applied Probability**, Oxford, v. 19, p. 20 – 28, 1982.
- DEMÉTRIO, C.G.B. **Modelos lineares generalizados na experimentação agrônômica**. Piracicaba: ESALQ, Departamento de Ciências Exatas, 2001. 113 p.
- FADDY, M.J. On variation in Poisson processes. **Mathematical Scientist**, Melbourne, v. 19, p. 47 – 51, 1994.
- FADDY, M.J. Extended Poisson process modelling and analysis of count data. **Biometrical Journal**, Berlin, v. 39, p. 431 – 440, 1997a.
- FADDY, M.J. On extending the negative binomial distribution and the number of weekly winners of the UK national lottery. **Mathematical Scientist**, Melbourne, v. 22, p. 77 – 82, 1997b.
- FADDY, M.J. Markov process modelling and analysis of discrete data. **Applied Stochastic Models and Data Analysis**, Chichester, v. 13, p. 217 – 223, 1998a.
- FADDY, M.J. Stochastic models for analysis of species abundance data. In: FLETCHER, D.J.; KAVALIERIS, L.; MANLY, B.F.J. (Ed.). **Statistics in ecology and environmental monitoring 2**. Dunedin: University of Otago Press, 1998b. p. 33 – 40.
- FADDY, M.J. Markov death process modelling and analysis of binary data. **Statistics & Probability Letters**, Amsterdam, v. 40, p. 9 – 13, 1998c.
- FADDY, M.J.; FENLON, J.S. Stochastic modeling of the invasion process of nematodes in fly larvae. **Applied Statistics**, London, v. 48, pt. 1, p. 31 – 37, 1999.

- FADDY, M.J.; BOSCH, R.J. Likelihood-based modeling and analysis of data underdispersed relative to the Poisson distribution. **Biometrics**, Washington, v. 57, p. 620 – 624, 2001.
- FADDY, M.J.; BROWN, J.; COMMINS, P. Likelihood-based modeling and analysis of possum trapping data. **The Journal of Agricultural, Biological, and Environmental Statistics**, Alexandria, v. 6, n. 2, p. 235 – 242, 2001.
- GRIFFITHS, D. Avoidance-modified generalised distributions and their application to studies of superparasitism. **Biometrics**, Washington, v. 33, p. 103 – 112, 1977a.
- GRIFFITHS, D. Models for avoidance of superparasitism. **Journal of Animal Ecology**, London, v. 46, p. 59 – 62, 1977b.
- HEMERIK, L.; VAN DER HOEVEN, N.; VAN ALPHEN, J.J.M. Egg distributions and the information a solitary parasitoid has and uses for its oviposition decisions. **Acta Biotheoretica**, Leiden, v. 50, p. 167 – 188, 2002.
- HEMERIK, L.; VAN DER HOEVEN, N. Egg distributions of solitary parasitoids revisited. **The Netherlands Entomological Society Entomologia Experimentalis et Applicata**, Amsterdam, v. 107, p. 81 – 86, 2003.
- JANARDAN, K.G.; MOHAN, N.R.; VASUDEVA, R.; HEBBAR, H.V. A stochastic model with random change point for the study of oviposition tactics of weevils on beans. **Journal of Statistical Planning and Inference**, Amsterdam, v. 61, p. 193 – 201, 1997.
- JENSEN, J.L.W.V. Sur une identité et sur d'autres formules analogues. **Acta Mathematica**, Stockholm, v. 26, p. 307 – 318, 1902.
- MORGAN, B.J.T. **Applied stochastic modelling**. New York: Oxford University Press, 2000. 297 p.
- MORGAN, R.W. Some stochastic models to describe the fertilization of an egg. **Applied Statistics**, London, v. 24, p. 137 – 138, 1975.
- PARRA, J.R.P. Técnicas de criação de *Anagasta kueiella* hospedeiro alternativo para a produção de *Trichogramma*. In: PARRA, J.R.P.; ZUCCHI, R.A. (Ed.). **Trichogramma e o controle biológico aplicado**. Piracicaba: FEALQ, 1997. cap. 4, p. 121-150.
- PAYNE, C.C. Pathogens for the controls of insects: where next? **Philosophical Transactions of the Royal Society London**. Series B, London, v. 318, p. 225-248, 1988.
- PODLICH, H.M. **Extended Poisson models for count data**. 1999. 204p. Thesis (Ph.D) - Department of Mathematics, The University of Queensland, Brisbane, 1999.
- PODLICH, H.M.; FADDY, M.J.; SMYTH, G.K. Likelihood computations for extended Poisson process models. **InterStat**, n. 1, p. 15, 1999. Disponível em: <<http://InterStat.StalJournals.net/>>. Acesso em: 13 out. 2004.
- PODLICH, H.M.; FADDY, M.J.; SMYTH, G.K. A general approach to modeling and analysis of species abundance data with extra zeros. **The Journal of Agricultural, Biological, and Environmental Statistics**, Alexandria, v. 7, n. 3, p. 324 – 334, 2002.

- PODLICH, H.M.; FADDY, M.J.; SMYTH, G.K. Semi-parametric extended Poisson process models for count data. **Statistics and Computing**, London, v. 14, p. 311 – 321, 2004.
- ROGERS, D. A model for avoidance of superparasitism by solitary insect parasitoids. **Journal of Animal Ecology**, London, v. 44, p. 623 – 628, 1975.
- ROHATGI, V.K. **Statistical inference**. New York: John Wiley, 1984. 939 p.
- SAS INSTITUTE. **SAS/STAT 9.1: user's guide**. Cary, 2004. 5136 p
- SIMMONDS, F.J. Superparasitism by *Spalangia drosophilae* ashm. **Bulletin of Entomological Research**, London, v. 47, p. 361 – 376, 1956.
- SMITH, H.S. An attempt to redefine the host relationships exhibited by entomophagous insects. **Journal of Economic Entomology**, College Park, v. 9, p. 447 – 486, 1916.
- SMYTH, G.H.; PODLICH, H.M. An improved saddlepoint approximation based on the negative binomial for the general birth process. **Computational Statistics**, Heidelberg, v. 17, p. 17 – 28, 2002.
- TOSCAS, P.J.; FADDY, M.J. A likelihood-based analysis of longitudinal count data using a generalised Poisson model. **Statistical Modelling**, London, v. 3, p. 99 – 108, 2003.

## APÊNDICES



## APÊNDICE A - Função Geradora de Probabilidades.

### Definição

De acordo com Rohatgi (1984), seja uma sequência de números reais  $\{q_x\}$  ( $x = 0, 1, 2, \dots$ ). Usando esses coeficientes, pode-se formar uma série de potências para alguma variável  $s$ .

$$R(s) = \sum_{x=0}^{\infty} q_x s^x.$$

Se  $R(s)$  converge em algum intervalo aberto contendo  $s = 0$ , então  $R(s)$  é chamada função geradora da sequência  $\{q_x\}$ .

Em particular, uma variável aleatória  $X$ , com valores inteiros não negativos, em que

$$q_x = P(X = x), \quad x = 0, 1, 2, \dots$$

tem sua função geradora de probabilidades  $G(s)$  definida por:

$$G(s) = q_0 + q_1 s + q_2 s^2 + \dots = \sum_{x=0}^{\infty} q_x s^x = E[s^X]$$

com  $G(s)$  definida para  $|s| \leq 1$ .

### Propriedades

As principais propriedades das funções geradoras de probabilidades são:

1.  $G(1) = \sum_{x=0}^{\infty} q_x = 1$ ;
2.  $G_Y(s) = s^b G(s^a)$  se  $Y = aX + b$ ;
3.  $G(s) = M_X[\ln(s)]$  em que  $M_X[\ln(s)]$  é a função geradora de momentos de  $X$ ;

Prova:

$$M_X[\ln(s)] = E(e^{\ln(s)X}) = E(e^{\ln(s^X)}) = E(s^X) = G(s)$$

4. Se  $S_n = \sum_{i=1}^n X_i$ , então,  $M_{S_n}(s) = \prod_{i=1}^n G_{X_i}(s)$  se  $X_i$ ,  $i = 1, 2, \dots, n$ , são variáveis aleatórias independentes com f.g.p.  $G_{X_i}(s)$ . Se também forem identicamente distribuídas, isto é, elas constituem uma amostra aleatória de  $X$ , então  $M_{S_n}(s) = [G(s)]^n$ ;

5. A função geradora de probabilidade é infinitamente diferenciável para  $|s| < 1$  e a  $r$ -ésima derivada de  $G$  é dada por

$$G^{(r)}(s) = \sum_{x=r}^{\infty} x(x-1)\cdots(x-r+1)P(X=x)s^{x-r}.$$

Em particular,

$$G^{(r)}(0) = r(r-1)\cdots(r-r+1)P(X=r) = r!P(X=r)$$

e, portanto, pode-se determinar  $P(X=r)$  através da  $r$ -ésima derivada de  $G$  calculada no ponto zero e assim:

$$P(X=r) = (r!)^{-1}G^{(r)}(0);$$

6. Se  $E(X) < \infty$  e  $E(X^2) < \infty$ , então,

$$E(X) = G'(1) \quad e \quad E(X^2) = G'(1) + G''(1).$$

Logo,

$$Var(X) = E(X^2) - [E(X)]^2 = G'(1) + G''(1) - [G'(1)]^2.$$

Em geral, se  $\mu_r = E[X(X-1)\cdots(X-r+1)] < \infty$  é o  $r$ -ésimo momento fatorial de  $X$ , então

$$E[X(X-1)\cdots(X-r+1)] = P^{(r)}(1).$$

7. Sejam  $X$  e  $Y$  variáveis aleatórias discretas com funções geradoras de probabilidades  $G(s)$  e  $H(s)$ , respectivamente. Se  $G(s) = H(s)$ , para todo  $|s| < h$ , então as variáveis  $X$  e  $Y$  possuem a mesma distribuição, isto é,  $P(X=x) = P(Y=x)$ .

Prova. Sejam  $G(s) = \sum_{x=0}^{\infty} q_x s^x$  e  $H(s) = \sum_{x=0}^{\infty} q_x s^x$ . Então pela propriedade (5):

$P(X=r) = (r!)^{-1}G^{(r)}(0) = (r!)^{-1}H^{(r)}(0) = P(Y=r), r = 0, 1, \dots$ . Portanto,  $X$  e  $Y$  têm a mesma distribuição de probabilidade.

8. A transformada de Laplace,  $L(s)$ , é dada por  $L(s) = E(e^{-sX}) = G(e^{-s})$ .

Nos exemplos, a seguir, serão calculadas a f.g.p das principais distribuições discretas bem como suas médias e variâncias.

*Exemplo 1. Poisson*

Seja  $X \sim \text{Poisson}(\lambda)$  com f.p.

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!} I_{\{0,1,2,\dots\}}(x).$$

Assim, dado que

$$\sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} = 1 \Leftrightarrow e^{-\lambda} = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!},$$

então, pode-se obter

$$e^{\lambda s} = \sum_{x=0}^{\infty} \frac{(\lambda s)^x}{x!}$$

e, portanto,

$$G(s) = \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x s^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda s)^x}{x!} = e^{-\lambda} e^{\lambda s} = e^{\lambda(s-1)}$$

para  $s \in \mathfrak{R}$ . Como  $G^{(r)}(s) = \lambda^r \exp[\lambda(s-1)]$ , o  $r$ -ésimo momento fatorial de  $X$  é dado por  $\mu_{[r]} = \lambda^r$ .

Assim  $G'(1) = \lambda$  e  $G''(1) = \lambda^2$  e portanto,  $E(X) = \text{Var}(X) = \lambda$ .

*Exemplo 2. Binomial*

Seja  $X \sim \text{Binomial}(m, p)$  com f.p.

$$q_x = P(X = x) = \binom{m}{x} p^x (1-p)^{m-x} I_{\{0,1,\dots,m\}}(x).$$

Assim, dado que

$$\sum_{x=0}^m \binom{m}{x} p^x (1-p)^{m-x} = 1,$$

então, pode-se obter

$$G(s) = \sum_{x=0}^m s^x \binom{m}{x} p^x (1-p)^{m-x} = \sum_{x=0}^m \binom{m}{x} (sp)^x (1-p)^{m-x} = [1-p+ps]^m$$

para  $s \in \mathfrak{R}$ . Como  $G^{(r)}(s) = m(m-1)\cdots(m-r+1)p^r [1-p+ps]^{m-r}$ , o  $r$ -ésimo momento fatorial de  $X$  é dado por  $\mu_{[r]} = m(m-1)\cdots(m-r+1)p^r$ .

Assim  $G'(1) = mp$  e  $G''(1) = m(m-1)p^2$  e, portanto,  $E(X) = mp$  e  $\text{Var}(X) = mp(1-p)$ .

*Exemplo 3.* Binomial Negativa

Seja  $X \sim \text{BinomialNegativa}(m, p)$

$$q_x = P(X = x) = \binom{x+m-1}{m-1} p^m (1-p)^x I_{\{0,1,\dots,\infty\}}(x).$$

Assim, dado que

$$\sum_{x=0}^{\infty} \binom{x+m-1}{m-1} p^m (1-p)^x = 1,$$

e, portanto,

$$\sum_{x=0}^{\infty} \binom{x+m-1}{m-1} (1-p)^x = p^{-m},$$

então, pode-se obter

$$\begin{aligned} G(s) &= \sum_{x=0}^{\infty} s^x \binom{x+m-1}{m-1} p^m (1-p)^x = \sum_{x=0}^{\infty} \binom{x+m-1}{x} p^m [(1-p)s]^x \\ &= p^m \sum_{x=0}^{\infty} \binom{x+m-1}{x} [(1-p)s]^x = p^m [1 - (1-p)s]^{-m} \\ &= \left[ \frac{p}{1 - (1-p)s} \right]^m, \end{aligned}$$

com  $s < (1-p)^{-1}$ .

Como

$$G^{(r)}(s) = m(m+1) \cdots (m+r-1) p^m [1 - p + ps]^{-(m+r)},$$

o  $r$ -ésimo momento fatorial de  $X$  é dado por

$$\begin{aligned} \mu_{[r]} &= m(m+1) \cdots (m+r-1) p^m (1-p)^r p^{-(m+r)} \\ &= m(m+1) \cdots (m+r-1) p^{-r} (1-p)^{-r}. \end{aligned}$$

Assim  $G'(1) = \frac{mq}{p}$  e  $G''(1) = m(m-1)(1-p)^2 p^{-2}$  e, portanto,

$$E(X) = \frac{m(1-p)}{p} \text{ e } Var(X) = \frac{m(1-p)}{p^2}.$$

*Exemplo 4.* Poisson truncada no ponto zero

Seja  $X \sim \text{PoissonTruncada}(\lambda)$

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{(1 - e^{-\lambda}) x!} I_{\{1,2,\dots\}}(x).$$

Assim, dado que

$$\sum_{x=1}^{\infty} \frac{e^{-\lambda} \lambda^x}{(1 - e^{-\lambda}) x!} = 1 \Leftrightarrow e^{\lambda} (1 - e^{-\lambda}) = \sum_{x=1}^{\infty} \frac{\lambda^x}{x!},$$

então, pode-se obter

$$e^{\lambda s} (1 - e^{-\lambda}) = \sum_{x=1}^{\infty} \frac{(\lambda s)^x}{x!}$$

e, portanto,

$$G(s) = \sum_{x=1}^{\infty} \frac{e^{-\lambda} \lambda^x s^x}{(1 - e^{-\lambda}) x!} = \frac{e^{-\lambda}}{(1 - e^{-\lambda})} \sum_{x=1}^{\infty} \frac{(\lambda s)^x}{x!} = \frac{e^{\lambda(s-1)}}{(1 - e^{-\lambda})},$$

para  $s \in \mathfrak{R}$ . Como

$$G^{(r)}(s) = \frac{\lambda^r e^{\lambda s}}{(1 - e^{-\lambda})},$$

o  $r$ -ésimo momento fatorial de  $X$  é dado por

$$\mu_{[r]} = \frac{\lambda^r e^{\lambda s}}{1 - e^{-\lambda}}.$$

$$\text{Assim } G'(1) = \frac{\lambda}{1 - e^{-\lambda}} \quad \text{e} \quad G''(1) = \frac{\lambda^2}{1 - e^{-\lambda}} \quad \text{e, portanto,}$$

$$E(X) = \frac{\lambda}{(1 - e^{-\lambda})} = \mu \quad \text{e} \quad V(X) = \mu(1 + \lambda - \mu).$$

*Exemplo 5.* Logarítmica

Seja  $X \sim \text{logarítmica}(p)$  com f.p.

$$q_x = P(X = x) = \frac{(1-p)^x}{-x \ln p} I_{\{1,2,3,\dots,\infty\}}(x).$$

Assim, dado que

$$\sum_{x=1}^{\infty} \frac{(1-p)^x}{-x \ln p} = 1$$

tem-se

$$\sum_{x=1}^{\infty} \frac{[(1-p)s]^x}{x} = -\ln p$$

então:

$$G(s) = \sum_{x=1}^{\infty} s^x \frac{(1-p)^x}{-x \ln p} = \frac{1}{-\ln p} \sum_{x=1}^{\infty} \frac{[(1-p)s]^x}{x} = \frac{\ln[1 - s(1-p)]}{\ln p}.$$

## Distribuição da soma de um número aleatório de variáveis aleatórias

É cada vez mais freqüente em Entomologia o aparecimento de situações práticas em que se precisa calcular a distribuição de probabilidade de uma variável aleatória discreta que representa a soma aleatória de um número aleatório de variáveis aleatórias. Particularmente, suponha que cada fêmea de um número aleatório,  $N$ , de insetos deposita  $X_i$  ovos,  $i = 1, 2, \dots$ . A quantidade total de ovos postos  $S_N$ , é definida por :

$$S_N = \begin{cases} 0 & \text{se } N = 0; \\ X_1 + X_2 + \dots + X_N & \text{se } N = 1, 2, \dots \end{cases}$$

Como se nota o número de parcelas em  $S_N$  depende de um particular valor assumido pela variável de contagem  $N$ . Seja inicialmente, o caso em que as variáveis aleatórias  $X_i$   $i = 1, 2, \dots$  são inteiras não negativas, independente e identicamente distribuídas com a mesma f.g.p  $G_1(s)$  e independe da variável aleatória inteira não negativa  $N$  com f.g.p.  $G_N(s)$ .

Fato 1. A função geradora de probabilidade de  $S_N$ ,  $G_{S_N}(s)$ , é dada por  $G_{S_N}(s) = G_N[G_1(s)]$ .

Prova:

$$G_{S_N}(s) = E[s^{S_N}] = E[E(s^{S_N}|N)]$$

tem-se pela independência entre  $X_1, X_2, \dots$  e delas com  $N$  que:

$$\begin{aligned} E(s^{S_N}|N = n) &= E[s^{X_1+X_2+\dots+X_n}] = E[s^{X_1} s^{X_2} \dots s^{X_n}] = \prod_{i=1}^n [E(s^{X_i})] \\ &= \prod_{i=1}^n G_1(s) = [G_1(s)]^n \end{aligned}$$

e portanto,

$$G_{S_N}(s) = E[[G_1(s)]^N] = \sum_{n=0}^{\infty} q_n [G_1(s)]^n = G_N[G_1(s)].$$

Será mostrada agora a relação que há entre as médias de  $S_N$ ,  $N$  e  $X_1$  conhecida como equação de Wald.

Fato 2.  $E(S_N) = E(N).E(X_1)$ .

Prova: Diferenciando  $G_{S_N}(s)$ , em relação a  $s$ , tem-se;

$$G'_{S_N}(s) = G'_N[G_1(s)]G'_1(s).$$

Fazendo  $s = 1$  e usando as propriedades 1 e 6,

$$G'_{S_N}(1) = G'_N[G_1(1)]G'_1(1)$$

logo,

$$G_{S_N}^{(1)}(1) = G_N^{(1)}(1)G_1^{(1)}(1)$$

e assim  $E(S_N) = E(N) \cdot E(X_1)$ .

O próximo passo é calcular a variância de  $S_N$ .

Fato 3.  $Var(S_N) = [E(X_1)]^2 Var(N) + E(N) \cdot E(X_1)$ .

Prova: Diferenciando  $G'_{S_N}(s)$ , em relação a  $s$ , tem-se;

$$G''_{S_N}(s) = G''_N[G_1(s)][G'_1(s)]^2 + G'_1(s)[G''_N(G_1(s))].$$

Fazendo  $s = 1$  e usando as propriedades 1 e 6,

$$\begin{aligned} G''_{S_N}(1) &= G''_N[G_1(1)][G'_1(1)]^2 + G'_1(1)G''_N[G_1(1)] \\ &= G''_N(1)E^2(X_1) + G'_1(1)E(N) \\ &= E[N(N-1)]E^2(X_1) + E[X_1(X_1-1)]E(N) \\ &= [E(N^2) - E(N)]E^2(X_1) + E[X_1^2 - X_1]E(N) \\ &= [Var(N) + E^2(N) - E(N)]E^2(X_1) + [E(X_1^2) - E(X_1)]E(N) \\ &= Var(N)E^2(X_1) + E^2(N)E^2(X_1) - E(N)Var(X_1). \end{aligned}$$

Então,

$$\begin{aligned} Var(S_N) &= G''_{S_N}(1) + G'_{S_N}(1) - [G'_{S_N}(1)]^2 \\ &= Var(N)E^2(X_1) + E(N)Var(X_1). \end{aligned}$$

### Exemplo 6. Aplicação

Suponha que o número de insetos  $N \sim P(\lambda)$  e que cada ovo gere um novo inseto com probabilidade  $p$ , isto é,  $X_i \sim B(p)$ . Logo  $S_N = \sum_{i=1}^N X_i$ , o número de insetos gerados, tem f.g.p. dada por:

$$G_{S_N}(s) = G_N[G_1(s)] = G_N[1 - p + ps] = e^{\lambda(1-p+ps-1)} = e^{\lambda p(s-1)},$$

que é a f.g.p. de uma Poisson, e portanto  $S_N \sim P(\lambda p)$ .

*Exemplo 7. Aplicação*

Suponha que o número de insetos  $N \sim P(\lambda)$  e que cada inseto gere pelo menos um descendente segundo uma distribuição logarítmica de parâmetro  $p$ , isto é,  $X_i \sim \text{LogB}(p)$ .

Logo  $S_N = \sum_{i=1}^N X_i$ , o número de insetos gerados, tem f.g.p. dada por:

$$G_{S_N}(s) = G_N[G_1(s)] = G_N \left[ \frac{\ln[1 - s(1 - p)]}{\ln p} \right] = e^{\lambda \left( \frac{\ln[1 - s(1 - p)]}{\ln p} - 1 \right)} = \left[ \frac{p}{1 - s(1 - p)} \right]^{-\frac{\lambda}{\ln p}},$$

que é a f.g.p. de uma Binomial negativa, e portanto  $S_N \sim \text{BN} \left( r = -\frac{\lambda}{\ln p}, p \right)$ .



## APÊNDICE B - Modelo de Janardan

Janardan et al. (1997), para descrever o número  $X(t)$  de ovos postos no tempo  $0 \leq t \leq T$ , incorporam o fato de que a mudança da taxa descrita pode ocorrer depois de um ovo ser posto (probabilidade  $p$ ) ou depois de dois ovos serem postos (probabilidade  $q = 1 - p$ ), através de uma variável aleatória binária  $Y$  que assume, respectivamente, os valores  $m = 1$  ou  $m = 2$ . Supõem, ainda, independência entre  $X(t)$  e  $Y$  e que os seguintes postulados sejam verdadeiros:

- a.  $P(X(t+h) = k+1 | X(t)=k, Y = m) = \lambda h + o(h), 0 \leq k \leq m-1$ .
- b.  $P(X(t+h) = k | X(t)=k, Y = m) = 1 - \lambda h + o(h), 0 \leq k \leq m-1$ .
- c.  $P(X(t+h) = k+1 | X(t)=k, Y = m) = \mu h + o(h), k \geq m$ .
- d.  $P(X(t+h) = k | X(t)=k, Y = m) = 1 - \mu h + o(h), k \geq m$ .
- e.  $P(X(t+h) = k+2 | X(t)=k, Y = m) = o(h), k \geq 0, m = 1, 2, \dots$

Seja  $P_n(t) = P(X(t) = n)$ ,  $n \geq 0$ . Para a obtenção da f.p. de  $X(t)$  há necessidade de se resolver um sistema de equações diferenciais, cujos principais passos serão discutidos a seguir.

### Fato 1

$$\begin{aligned} P_0(t+h) &= \sum_{m=1}^2 P(X(t+h) = 0 | X(t) = 0, Y = m) P_0(t) P(Y = m) \\ &= (1 - \lambda h + o(h)) P_0(t), \end{aligned} \quad (69)$$

### Prova

$$\begin{aligned} P_0(t+h) &= P(X(t+h) = 0 | X(t) = 0, Y = 1) P_0(t) P(Y = 1) + \\ &\quad + P(X(t+h) = 0 | X(t) = 0, Y = 2) P_0(t) P(Y = 2) \\ &= (1 - \lambda h + o(h)) P_0(t) p + (1 - \lambda h + o(h)) P_0(t) q \\ &= (1 - \lambda h + o(h)) P_0(t), \end{aligned}$$

pois pela propriedade b tem-se  $P(X(t+h) = k | X(t) = k, Y = m) = 1 - \lambda h + o(h)$ ,  $0 \leq k \leq m-1$ . Assim, para  $m = 1$  acarreta  $k = 0$  e a propriedade b fica:

$$P(X(t+h) = 0 | X(t) = 0, Y = 1) = 1 - \lambda h + o(h).$$

Já para  $m = 2$ , os valores de  $k = 0, 1$  menores do que  $m - 1 = 1$  e, portanto,  $P(X(t + h) = 0 | X(t) = 0, Y = 2) = 1 - \lambda h + o(h)$  e usando  $P(Y = 1) = p$  e  $P(Y = 2) = q$ .

**Fato 2**

$$\begin{aligned}
 P_1(t + h) &= \sum_{m=1}^2 P(X(t + h) = 1, Y = m) \\
 &= \sum_{m=1}^2 \sum_{k=0}^1 P(X(t + h) = 1 | X(t) = k, Y = m) P_0(t) P(Y = m) \\
 &= \lambda h P_0(t) + (1 - \alpha h) P_1(t) + o(h),
 \end{aligned} \tag{70}$$

em que  $\alpha = \lambda q + \mu p$ .

**Prova**

$$\begin{aligned}
 P_1(t + h) &= \sum_{m=1}^2 P((X(t + h) = 1, Y = m) \\
 &= \sum_{m=1}^2 \sum_{k=0}^1 P(X(t + h) = 1 | X(t) = k, Y = m) P_k(t) P(Y = m) \\
 &= \sum_{k=0}^1 P(X(t + h) = 1 | X(t) = k, Y = 1) P_k(t) P(Y = 1) \\
 &\quad + \sum_{k=0}^1 P(X(t + h) = 1 | X(t) = k, Y = 2) P_k(t) P(Y = 2)
 \end{aligned}$$

mas

$$\sum_{k=0}^1 P(X(t + h) = 1 | X(t) = k, Y = 1) P_k(t) P(Y = 1) = p \sum_{k=0}^1 P(X(t + h) = 1 | X(t) = k, Y = 1) P_k(t)$$

$$\begin{aligned}
 P_2(t + h) &= \sum_{m=1}^2 P(X(t + h) = 2, Y = m) \\
 &= \sum_{m=1}^2 \sum_{k=0}^1 P(X(t + h) = 2 | X(t) = k, Y = m) P_0(t) P(Y = m) \\
 &= \alpha h P_1(t) + (1 - \mu h) P_2(t) + o(h),
 \end{aligned} \tag{71}$$

$$\begin{aligned}
P_n(t+h) &= \sum_{m=1}^2 P(X(t+h) = n, Y = m) \\
&= \sum_{m=1}^2 \sum_{k=0}^1 P(X(t+h) = n | X(t) = k, Y = m) P_0(t) P(Y = m) \\
&= \mu h P_{n-1}(t) + (1 - \mu h) P_n(t) + o(h), n \geq 3.
\end{aligned} \tag{72}$$

Através das equações (69) - (72) obtém-se o sistema de equações diferenciais:

$$P'_0(t) = -\lambda P_0(t), \tag{73}$$

$$P'_1(t) = \lambda P_0(t) - \alpha P_1(t), \tag{74}$$

$$P'_2(t) = \alpha P_1(t) - \mu P_2(t), \tag{75}$$

$$P'_n(t) = \mu P_{n-1}(t) - \mu P_n(t), n \geq 3. \tag{76}$$

Resolvendo a equação (73) tem-se  $P_0(t) = e^{-\lambda t}$ , e substituindo-a na equação (74) obtém-se:

$$P'_1(t) = \lambda e^{-\lambda t} - \alpha P_1(t).$$

Logo

$$P'_1(t) + \alpha P_1(t) = \lambda e^{-\lambda t},$$

que multiplicada por  $e^{\alpha t}$  resulta em:

$$[P'_1(t) + \alpha P_1(t)] e^{\alpha t} = \lambda e^{(\alpha-\lambda)t}.$$

Então,

$$\begin{aligned}
P'_1(t) e^{\alpha t} + \alpha P_1(t) e^{\alpha t} &= \lambda e^{(\alpha-\lambda)t} \\
[P_1(t) e^{\alpha t}]' &= \lambda e^{(\alpha-\lambda)t},
\end{aligned}$$

que integrada leva a:

$$\begin{aligned} \int_0^t [P_1(u)e^{\alpha u}]' du &= \int_0^t \lambda e^{(\alpha-\lambda)u} du \\ [P_1(u)e^{\alpha u}] \Big|_0^t &= \frac{\lambda}{\alpha-\lambda} [e^{(\alpha-\lambda)t} - 1] \\ P_1(t) - P_1(0) &= \frac{\lambda}{\alpha-\lambda} [e^{(\alpha-\lambda)t} - 1] \end{aligned}$$

Como  $P_1(0) = 0$  e multiplicando-se por  $e^{-\alpha t}$  tem-se:

$$P_1(t) = \frac{\lambda}{\alpha-\lambda} [e^{-\lambda t} - e^{-\alpha t}].$$

Mas

$$\begin{aligned} \frac{\lambda}{\alpha-\lambda} &= \frac{\lambda}{\lambda q + \mu p - \lambda} \\ &= \frac{\lambda}{\lambda(1-p) + \mu p - \lambda} \\ &= \frac{\lambda}{p(\mu - \lambda)} \\ &= \frac{\lambda}{p\beta}, \end{aligned}$$

pois  $q = 1 - p$  e  $\beta = \mu - \lambda$ .

Seja a f.g.p. de  $X(t)$ ,  $G(z, t) = \sum_{n=0}^{\infty} P_n(t) z^n$ . Derivando-a em relação a  $t$  tem-se:

$$\begin{aligned} G'(z, t) &= \sum_{n=0}^{\infty} P_n'(t) z^n \\ &= -\lambda P_0(t) + [\lambda P_0(t) - \alpha P_1(t)]z + [\alpha P_1(t) - \mu P_2(t)]z^2 \\ &\quad + \sum_{n=3}^{\infty} \mu [P_{n-1}(t) - P_n(t)](t) z^n \\ &= (z-1)\lambda P_0(t) + (z^2 - z)\alpha P_1(t) - z^2 \mu P_2(t) \\ &\quad + \mu \sum_{n=3}^{\infty} [P_{n-1}(t) - P_n(t)](t) z^n \end{aligned}$$

mas,

$$\begin{aligned}
\sum_{n=3}^{\infty} [P_{n-1}(t) - P_n(t)](t)z^n &= \sum_{n=3}^{\infty} [P_{n-1}(t)z^n - \sum_{n=3}^{\infty} [P_{n-1}(t)z^n \\
&= z \sum_{n=3}^{\infty} [P_{n-1}(t)z^{n-1} - \sum_{n=3}^{\infty} [P_{n-1}(t)z^n \\
&= z \sum_{n=2}^{\infty} [P_n(t)z^n - \sum_{n=3}^{\infty} [P_{n-1}(t)z^n \\
&= z[G(z, t) - P_0(t) - zP_1(t)] - [G(z, t) - P_0(t) - zP_1(t) - z^2P_2(t)] \\
&= (z - 1)G(z, t) - (z - 1)P_0(t) - (z^2 - z)P_1(t) + z^2P_2(t),
\end{aligned}$$

Substituindo na expressão anterior e lembrando que  $\alpha = \lambda + p\beta$  e  $\mu = \beta + \lambda$ , tem-se:

$$G'(z, t) = (z - 1)G(z, t) - (z - 1)\beta[P_0(t) + zqP_1(t)]. \quad (77)$$

Resolvendo-se a equação(77), obtém-se

$$\begin{aligned}
G(z, t) = \sum_{n=0}^{\infty} P_n(t)z^n &= \frac{(z - 1)(\beta p + z\lambda q)}{p(\mu z - \beta)} [e^{-\lambda t} - e^{-(1-z)\mu t}] \\
&\quad - \frac{z(z - 1)\lambda q}{p(\mu z - \beta q)} [e^{-\alpha t} - e^{-(1-z)\mu t}] + e^{-(1-z)\mu t},
\end{aligned}$$

em que  $\beta = \mu - \lambda$ .

Sua função de probabilidade é dada por:

$$\begin{aligned}
P_n(t) &= e^{-\lambda t} I_{\{0\}}(n) + \frac{\lambda}{p\beta} [e^{-\lambda t} - e^{-\alpha t}] I_{\{1\}}(n) \\
&\quad + \frac{\alpha \lambda \mu^{n-2} e^{-\mu t}}{pq^{n-1}\beta^n} [q^{n-1}(e^{\beta t} - u(t)) - (e^{q\beta t} - v(t))] I_{\{2,3,\dots,\infty\}}(n),
\end{aligned}$$

em que  $\alpha = \lambda q + \mu p$ ,  $u(t) = \sum_{j=0}^{n-2} \frac{(\beta t)^j}{j!}$  e  $v(t) = \sum_{j=0}^{n-2} \frac{(q\beta t)^j}{j!}$ .

Além disso,

$$E[X(t)] = \mu t - \beta \left\{ \frac{1 - e^{-\lambda t}}{\lambda} - \frac{\lambda q}{\beta p} \left[ \frac{1 - e^{-\alpha t}}{\alpha} - \frac{1 - e^{-\lambda t}}{\lambda} \right] \right\}.$$

**Prova**

Derivando (77) em relação a  $z$ , tem-se

$$\frac{\partial G}{\partial z} \frac{\partial G}{\partial t} = \mu [G(z, t) + (z - 1)G'(z, t)] - \beta [P_0(t) + zqP_1(t) - (z - 1)\beta qP_1(t)]. \quad (78)$$

Invertendo as ordens de integração em (78) e fazendo  $z = 1$ , tem-se a derivada de  $m(t) = E[X(t)]$  e, portanto,

$$\begin{aligned}
m'(t) &= \mu G(1, t) - \beta [P_0(t) + qP_1(t)] \\
&= \mu - \beta [P_0(t) + qP_1(t)] \\
&= \mu - \beta \left[ e^{-\lambda t} + q \frac{\lambda}{\alpha - \lambda} [e^{-\lambda t} - e^{-\alpha t}] \right] \\
&= \mu - \beta \left[ e^{-\lambda t} + \frac{\lambda q}{\beta p} [e^{-\lambda t} - e^{-\alpha t}] \right].
\end{aligned} \tag{79}$$

Integrando a equação (79), tem-se

$$\int_0^t m'(a) da = \int_0^t \mu da - \beta \left[ \int_0^t e^{-\lambda a} da + \frac{\lambda q}{\beta p} \left[ \int_0^t e^{-\lambda a} da - \int_0^t e^{-\alpha a} da \right] \right].$$

Assim,

$$m(t) - m(0) = \mu t + \beta \left\{ \frac{1 - e^{-\lambda t}}{\lambda} - \frac{\lambda q}{\beta p} \left[ \frac{1 - e^{-\alpha t}}{\alpha} - \frac{1 - e^{-\lambda t}}{\lambda} \right] \right\},$$

pois  $m(0) = 0$ .

Para calcular a  $V[X(t)] = E[(X(t))^2] - [m(t)]^2 = m_2(t) - [m(t)]^2$ , em que  $m_2(t) = E[X(t)[X(t) - 1]] + m(t)$ , precisa-se calcular o segundo momento fatorial de  $X(t)$ . Isto pode ser feito facilmente, mas o esforço algébrico é grande. Vai-se agora calcular a derivada parcial de (77) em relação a  $z$ , em seguida invertem-se as ordens de integração e faz-se  $z = 1$  obtendo-se a equação:

$$m'_2(t) = m'(t) + \mu(1 + m(t)) - 2\beta q P_1 t$$

que resolvida tem a seguinte expressão,

$$\begin{aligned}
m_2(t) &= \frac{[\alpha \mu p (\lambda - 2\beta) + 2\mu \lambda^2 q (\alpha - 1)] t}{\alpha \lambda p} + \mu^2 t^2 + \frac{\beta e^{-\lambda t}}{\lambda} + \\
&+ \frac{1}{\lambda p} [2\mu(\beta p - \lambda^2 q) + 2\mu \lambda^2 (\alpha - 1)] \left[ \frac{1 - e^{-\lambda t}}{\lambda} \right] + \\
&+ \frac{\lambda q (3 - 2\mu)}{p} \left[ \frac{1 - e^{-\alpha t}}{\alpha} \right].
\end{aligned} \tag{80}$$

## APÊNDICE C - Modelo de Faddy

### 1. Processo de Nascimento

Um processo estocástico a parâmetro contínuo  $\{Y(t); t \in T\}$  é uma coleção de variáveis indexadas por um conjunto contínuo  $T$ . Interpreta-se, geralmente,  $t$  como tempo e  $Y(t)$  como o estado do processo no tempo  $t$ . Usualmente, observa-se o número de ovos postos em um hospedeiro durante um período de 24 horas. Assim,  $Y(2)$  é o número de ovos postos no hospedeiro após 2 horas do início do experimento e o conjunto  $T = [0, 24]$ . Tal processo evolui através do tempo e usa-se o termo evento ou transição para denotar a troca do estado do processo. A palavra transição não significa necessariamente uma mudança física de estado pelo processo e sim uma mudança no tempo de observação do processo. Podlich (1999) considera que em um processo de nascimento a parâmetro contínuo um evento corresponda a um aumento de uma unidade no estado do processo. Assim,  $Y(t)$  é o número de nascimentos acumulados até o tempo  $t$ .

Seja  $\{Y(t); t \geq 0\}$  com  $Y(0) = 0$  um processo geral de nascimento sobre  $S = \{0, 1, 2, \dots, \infty\}$  conhecido como espaço de estados. As seguintes propriedades são válidas:

i. se  $s < t$ , então,  $Y(s) \leq Y(t)$ ;

ii.

$$P(Y(t + \Delta t) = (y + m) | Y(t) = y) = \begin{cases} \delta_y \Delta t + o(\Delta t) & \text{se } m = 1; \\ 1 - \delta_y \Delta t + o(\Delta t) & \text{se } m = 0; \\ o(\Delta t) & \text{se } m > 1 \end{cases}$$

iii. se  $s < t$ , então,  $Y(s) - Y(t)$  é independente de  $Y(s)$ .

A propriedade (i) diz que somente nascimentos podem ocorrer, isto é, o número atual de nascimentos é maior do que a de qualquer tempo anterior.

A propriedade (ii) diz que em um curto intervalo de tempo de comprimento  $\Delta t$  ou um nascimento ocorre com probabilidade aproximadamente proporcional a  $\Delta t$ , isto é,  $\delta_y \Delta t + o(\Delta t)$ , ou não ocorrem nascimentos. A quantidade  $\delta_y$  é a taxa de ocorrência de um próximo nascimento quando o processo está no estado  $y$ . Se  $\delta_y = 0$ , o estado  $y$  é dito absorvente pois se o processo entrar nele, nunca poderá deixá-lo, enquanto se  $\delta_y = \infty$  o estado

$y$  é chamado de instantâneo, desde que se ele for alcançado imediatamente, o processo deverá deixá-lo. As taxas de ocorrências dos eventos em Processos Estocásticos são conhecidas como taxas de transição do processo.

A propriedade (iii) diz que o número de nascimentos em qualquer intervalo de tempo é independente do número de eventos que já tenham ocorrido.

Uma outra importante propriedade diz que o tempo de permanência do sistema no estado  $y$ ,  $S_y$ , tem distribuição exponencial com parâmetro  $\delta_y$ . Assim,

$$P(S_y > t) = e^{-\delta_y t}$$

e

$$E(S_y) = \frac{1}{\delta_y}.$$

## 2. Equações de Chapman-Kolmogorov

Para se obter a distribuição de probabilidade das variáveis aleatórias  $\{Y(t); t \geq 0\}$  precisa-se resolver o sistema de equações diferenciais que governa o processo estocástico, conhecido como equações prospectivas de Chapman-Kolmogorov. Seja  $p_y(t) = P(Y(t) = y)$ , a probabilidade que  $y$  nascimentos tenham ocorrido até o tempo  $t$ . Pelo condicionamento de  $Y(t + \Delta t)$  sobre  $Y(t)$  e fazendo  $\Delta t \rightarrow 0$ , obtém-se o sistema de equações prospectivas (Cox ; Miller(1965))

$$p'_y(t) = \begin{cases} -\delta_0 p_0(t) & \text{com } p_0(0) = 1; \\ \delta_{y-1} p_{y-1}(t) - \delta_y p_y(t) & \text{com } p_y(0) = 0 \text{ se } y = 1, 2, \dots \end{cases}$$



**Prova**

Seja

$$\begin{aligned}
 p_0(t + \Delta) &= P(Y(t + \Delta) = 0) \\
 &= P(Y(t) = 0, Y(t + \Delta) - Y(t) = 0) \\
 &= P(Y(t) = 0)P(Y(t + \Delta) - Y(t) = 0) = p_0(t)[1 - \delta_0\Delta t + o(\Delta t)].
 \end{aligned}$$

Assim

$$\begin{aligned}
 \frac{p_0(t + \Delta) - p_0(t)}{\Delta t} &= p_0(t) \left[ -\delta_0 + \frac{o(\Delta t)}{\Delta t} \right] \\
 &= -\delta_0 p_0(t) + o(1),
 \end{aligned}$$

e fazendo-se  $\Delta t \rightarrow 0$  tem-se :

$$p'_0(t) = -\delta_0 p_0(t).$$

Por outro lado, para  $y = 1, 2, \dots$  tem-se que a probabilidade que no tempo  $(t + \Delta t)$  tenham ocorrido  $y$  nascimentos é a probabilidade que no tempo  $t$  tenham ocorrido  $(y - 1)$  nascimentos e que no intervalo  $(t, t + \Delta t)$  tenha ocorrido mais um nascimento ou que no tempo  $t$  tenham ocorrido  $y$  nascimentos e que no intervalo  $(t, t + \Delta t)$  não tenham ocorrido nascimentos.

$$\begin{aligned}
 p_y(t + \Delta) &= P(Y(t + \Delta) = y) \\
 &= P(Y(t) = y - 1, Y(t + \Delta) - Y(t) = 1) + P(Y(t) = y, Y(t + \Delta) - Y(t) = 1) \\
 &= P(Y(t) = y - 1)P(Y(t + \Delta) - Y(t) = 1) + P(Y(t) = y)P(Y(t + \Delta) - Y(t) = 0) \\
 &= p_{y-1}(t)[\delta_{y-1}\Delta t + o(\Delta t)] + p_y(t)[1 - \delta_y\Delta t + o(\Delta t)].
 \end{aligned}$$

Assim,

$$\begin{aligned}
 \frac{p_y(t + \Delta) - p_y(t)}{\Delta t} &= \delta_{y-1}p_{y-1}(t) - \delta_y p_y(t) + \frac{o(\Delta t)}{\Delta t} \\
 &= \delta_{y-1}p_{y-1}(t) - \delta_y p_y(t) + o(1).
 \end{aligned}$$

e fazendo-se  $\Delta t \rightarrow 0$ , tem-se :

$$p'_y(t) = \delta_{y-1}p_{y-1}(t) - \delta_y p_y(t).$$

### 3. Processo de Poisson

Um caso especial do processo de nascimento é o de Poisson. Seja  $T_i$ ,  $i = 1, 2, \dots$ , o tempo do  $i$ -ésimo nascimento e  $S_i = T_i - T_{i-1}$  o tempo de espera entre nascimentos sucessivos. O processo somente cresce, isto é, apenas nascimentos podem ocorrer e esse acréscimo é somente unitário e, portanto, a probabilidade de mais de um nascimento ocorrendo é zero. As taxas de nascimento são constantes, isto é,

$$P(Y(t + \Delta t) = y + m | Y(t) = y) = \begin{cases} \delta \Delta t + o(\Delta t) & \text{se } m = 1, \\ 1 - \delta \Delta t + o(\Delta t) & \text{se } m = 0 \\ o(\Delta t) & \text{se } m > 1 \end{cases}$$

em que  $\delta$  é a taxa de nascimento constante. Sob essa suposição de taxas de nascimentos constante,  $Y(t)$  tem uma distribuição de Poisson com média  $\delta t$ .

Diz-se que a distribuição de Poisson aparece de um processo de Poisson e para explicar a subdispersão e a superdispersão relativa à distribuição de Poisson vai-se generalizar a distribuição de Poisson, generalizando o processo de Poisson subjacente.

### 4. Generalização do Processo de Poisson

A generalização do processo de Poisson proposta por Faddy (1997a) para explicar a subdispersão ou a superdispersão já é bem conhecida (Feller 1971) e envolve a modelagem da seqüência das taxas de transição  $\delta_y$  como uma função de  $y$ . Assim, a taxa pela qual um nascimento ocorre é proporcional ao número acumulado de nascimentos já ocorridos.

Segundo Podlich (1999), a distribuição de probabilidade que surge de um processo de nascimento cujas taxas de nascimento não são constantes mas dependem do número de eventos já acumulados pode não se afastar muito da Poisson. As relações entre seqüências crescentes e decrescentes e o grau de afastamento foi conjecturado por Faddy (1994).

#### Conjectura de Faddy

Qualquer seqüência crescente  $\delta_y$ ,  $y = 0, 1, \dots$ , em  $y$  corresponde à distribuição  $p_y(t)$ ;  $y = 0, 1, \dots$ , que é superdispersa em relação à distribuição de Poisson e qualquer seqüência decrescente  $\delta_y$ ,  $y = 0, 1, \dots$ , em  $y$  corresponde à distribuição  $p_y(t)$ ,  $y = 0, 1, \dots$ , que é subdispersa em relação à distribuição de Poisson.

**Prova (Ball, 1995; Podlich, 1999)**

Seja  $D(t) = Var[Y(t)] - E[Y(t)]$  em que  $Y(0) = 0$  e  $Var[Y(0)] = E[Y(0)] = 0$  e, portanto,  $D[Y(0)] = 0$ . Mas,  $E[Y(t)] = \mu(t)$  é dada por:

$$\mu(t) = \sum_{y=0}^{\infty} yp_y(t) = \sum_{y=1}^{\infty} yp_y(t). \quad (81)$$

Derivando-se (81), em relação a  $t$ , e usando-se as equações de Chapman-Kolmogorov, tem-se:

$$\begin{aligned} \mu'(t) &= \sum_{y=0}^{\infty} yp'_y(t) \\ &= \sum_{y=1}^{\infty} y[\delta_{y-1}p_{y-1}(t) - \delta_y p_y(t)] \\ &= \sum_{y=1}^{\infty} y\delta_y p_y(t) - \sum_{y=1}^{\infty} y\delta_y p_y(t) \\ &= \sum_{y=0}^{\infty} (y+1)\delta_y p_y(t) - \sum_{y=0}^{\infty} y\delta_y p_y(t) \\ &= \sum_{y=0}^{\infty} \delta_y p_y(t). \end{aligned} \quad (82)$$

Considere, também,

$$\mu_2(t) = E[(Y(t))(Y(t-1))] = \sum_{y=2}^{\infty} y(y-1)\delta_y p_y(t). \quad (83)$$

Derivando-se (83), em relação a  $t$ , e usando-se as equações de Chapman-Kolmogorov, tem-se:

$$\begin{aligned}
\mu'_2(t) &= \sum_{y=2}^{\infty} y(y-1)p'_y(t) \\
&= \sum_{y=2}^{\infty} y(y-1)[\delta_{y-1}p_{y-1}(t) - \delta_y p_y(t)] \\
&= \sum_{y=2}^{\infty} y(y-1)\delta_{y-1}p_{y-1}(t) - \sum_{y=2}^{\infty} y(y-1)\delta_y p_y(t) \\
&= \sum_{y=1}^{\infty} (y+1)y\delta_y p_y(t) - \sum_{y=1}^{\infty} y\delta_y p_y(t) \\
&= \sum_{y=0}^{\infty} (y+1)y\delta_y p_y(t) - \sum_{y=0}^{\infty} y(y-1)\delta_y p_y(t) \\
&= 2 \sum_{y=0}^{\infty} y\delta_y p_y(t). \tag{84}
\end{aligned}$$

Como

$$\begin{aligned}
D(t) &= \text{Var}[Y(t)] - E[Y(t)] \\
&= E[Y^2(t)] - E^2[Y(t)] - E[Y(t)] \\
&= \mu_2(t) - [\mu(t)]^2, \tag{85}
\end{aligned}$$

tem-se, derivando-se (85), em relação a  $t$ ,

$$\begin{aligned}
D'(t) &= \mu'_2(t) - 2[\mu(t)][\mu'(t)] \\
&= 2 \left\{ \sum_{y=0}^{\infty} y\delta_y p_y(t) - \mu(t) \left[ \sum_{y=0}^{\infty} \delta_y p_y(t) \right] \right\}. \tag{86}
\end{aligned}$$

Assim  $D'(t) > 0$  ou  $D'(t) < 0$  para  $t \geq 0$  se:

$$\sum_{y=0}^{\infty} y\delta_y p_y(t) \succ \prec \mu(t) \left[ \sum_{y=0}^{\infty} \delta_y p_y(t) \right] \tag{87}$$

de acordo se a seqüência  $\delta_y$  é crescente ou decrescente, respectivamente. A conjectura de Faddy foi provada por Ball (1995) que utilizou uma variável aleatória  $Z$  cuja f.p. é dada por  $P(Z = z) = \mu^{-1}y p_y(t)$  em que  $\mu = \mu(t) = E[Y(t)]$  e usou comparações estocásticas para provar (87). Uma variável aleatória  $Z$  é dita estocasticamente maior do que  $Y$ ,  $Z \stackrel{\text{st}}{\geq} Y$ , se  $P(Z \geq n) \geq P(Y \geq n)$ . A variável  $Z$  definida anteriormente é estocasticamente maior do

que a variável  $Y$  definida por  $P(Y = n) = p_n$ . Agora, se  $Z \stackrel{\text{st}}{\geq} Y$ , então  $E[f(Z)] \geq E[f(Y)]$  se  $f$  é não decrescente enquanto  $E[f(Z)] \leq E[f(Y)]$  se  $f$  é não crescente. Definindo  $f$  como a sequência das taxas de transição e usando o resultado sobre comparações estocásticas tem-se:

$$\sum_{y=0}^{\infty} y \delta_y p_y(t) \geq \mu_t \sum_{y=0}^{\infty} \delta_y p_y(t)$$

e, então,  $V[X(t)] \geq E[X(t)]$  para seqüências crescentes e ocorrendo o contrário para seqüências decrescentes.

### Nota

1. Se  $X$  é uma variável aleatória inteira não negativa com f.p.  $p_x = P(X = x)$ ,  $x = 0, 1, 2, \dots$  e esperança finita, então  $E(X) = \sum_{x=1}^{\infty} P(X \geq x)$ .

### Prova.

$$\begin{aligned} E(X) &= \sum_{x=0}^{\infty} x P(X = x) \\ &= \sum_{x=1}^{\infty} x P(X = x) \\ &= \sum_{x=1}^{\infty} \sum_{y=1}^x P(X = x) \\ &= \sum_{y=1}^{\infty} \sum_{x=y}^{\infty} P(X = x) \\ &= \sum_{y=1}^{\infty} P(X \geq y). \end{aligned}$$

2. Se uma variável aleatória  $Z$  é estocasticamente maior do que  $Y$ ,  $Z \geq^{\text{st}} Y$ , então  $E(Z) \geq E(Y)$ .

**Prova.**

Como  $Z \geq^{\text{st}} Y$ , então  $P(Z \geq n) \geq P(Y \geq n)$ . Somando-se essa última expressão para  $n = 1, 2, \dots$ , tem-se:

$$\sum_{n=1}^{\infty} P(Z \geq n) \geq \sum_{n=1}^{\infty} P(Y \geq n)$$

$$E(Z) \geq E(Y),$$

usando o item 1.

3. Se  $X$  e  $Y$  são variáveis aleatórias independentes e identicamente distribuídas, então, para quaisquer funções crescentes (decrecentes)  $f$  e  $g$ , tem-se  $E[f(X)g(Y)] \geq E[f(X)]E[g(Y)]$ .

**Prova.**

Sejam  $x$  e  $y$  reais quaisquer, como  $f$  e  $g$  são funções crescentes tem-se:

$$[f(x) - f(y)][g(x) - g(y)] \geq 0,$$

pois, se  $x \geq y$  ( $x \leq y$ ), ambos os fatores têm o mesmo sinal e, portanto, o produto é positivo. Assim, para quaisquer variáveis aleatórias  $X$  e  $Y$ ,

$$[f(X) - f(Y)][g(X) - g(Y)] \geq 0$$

o que acarreta

$$E\{[f(X) - f(Y)][g(X) - g(Y)]\} \geq 0$$

ou equivalentemente,

$$E[f(X)g(X)] + E[f(Y)g(Y)] \geq E[f(X)g(Y)] + E[f(Y)g(X)].$$

Como  $X$  e  $Y$  são variáveis aleatórias independentes e identicamente distribuídas, então

$$E[f(X)g(X)] = E[f(Y)g(Y)],$$

e

$$E[f(X)g(Y)] = E[f(Y)g(X)] = E[f(X)]E[g(Y)]$$

e, portanto,

$$E[f(X)g(Y)] \geq E[f(X)]E[g(Y)].$$

Se  $f$  e  $g$  são funções decrescentes então,  $(-f)$  e  $(-g)$  são funções crescentes. Assim, aplicando o resultado anterior, tem-se:

$$E[-f(X)(-g(Y))] \geq E[-f(X)]E[-g(Y)] = E[f(X)]E[g(Y)]$$

e assim,

$$E[f(X)g(Y)] = E[-f(X)(-g(Y))] \geq E[f(X)]E[g(Y)].$$

O caso em que  $f$  é crescente e  $g$  decrescente, então  $f$  e  $(-g)$  são funções crescentes. Assim, aplicando o resultado anterior, tem-se:

$$E[f(X)(-g(Y))] \geq E[f(X)]E[-g(Y)] = -E[f(X)]E[g(Y)].$$

e assim,

$$-E[f(X)g(Y)] \geq -E[f(X)]E[g(Y)],$$

e multiplicando-se por  $(-1)$ , tem-se:

$$E[f(X)g(Y)] \leq E[f(X)]E[g(Y)].$$

Vai-se agora provar o resultado: Seja  $X$  uma variável aleatória discreta sobre os inteiros não negativos com média  $\mu$  e, seja  $Y$  uma variável aleatória sobre os inteiros positivos e cuja f.p. é dada por  $P(Y = y) = \frac{1}{\mu}yP(X = y)$ . Nessas condições,  $Y \stackrel{\text{st}}{\geq} X$ .

**Prova.**

Seja  $y = 1, 2, \dots$ . Vai-se provar que  $P(Y \geq y) \geq P(X \geq y)$  e para isso dois casos serão considerados.

Caso 1:  $0 < \mu < 1$ . Assim,  $\frac{1}{\mu} > 1$ . Logo

$$P(Y = i) = \frac{1}{\mu} i P(X = i) \geq i P(X = i) \geq P(X = i),$$

pois  $i \geq 1$ . Portanto,

$$\begin{aligned} \sum_{i=y}^{\infty} P(Y = i) &\geq \sum_{i=y}^{\infty} P(X = i) \\ P(Y \geq y) &\geq P(X \geq y). \end{aligned}$$

Caso 2.  $\mu \geq 1$

$$\begin{aligned} P(Y \geq y) &= \sum_{i=y}^{\infty} P(Y = i) \\ &= \sum_{i=y}^{\infty} \sum_{j=1}^i \frac{1}{\mu} i P(X = i) \\ \mu P(Y \geq y) &= \sum_{i=y}^{\infty} \sum_{j=1}^i P(X = i) \\ &= \left[ \sum_{j=1}^y \sum_{i=y}^{\infty} P(X = i) + \sum_{j=y+1}^{\infty} \sum_{i=j}^{\infty} P(X = i) \right] \\ &= \left[ \sum_{j=1}^y P(X \geq y) + \sum_{j=y+1}^{\infty} P(X \geq j) \right] \\ &= \left[ \sum_{j=1}^y P(X \geq y) + \mu - \sum_{j=1}^y P(X \geq j) \right] \end{aligned}$$



$$\begin{aligned}
&= \left[ \mu - \sum_{j=1}^{y-1} [P(X \geq y) - P(X \geq j)] \right] \\
&= \mu - \sum_{j=1}^y \sum_{i=j}^{y-1} P(X = i) \\
&= \mu - \sum_{i=1}^{y-1} \sum_{j=i}^{y-1} P(X = i) \\
&= \mu - \sum_{i=1}^{y-1} (y - i)P(X = i) \\
&\geq \mu - \sum_{i=1}^{y-1} P(X = i), \tag{88}
\end{aligned}$$

a última desigualdade se origina do fato que  $(y - i)P(X = i) \geq P(X = i)$ . Assim, tem-se de (88) que

$$P(Y \geq y) \geq \left[ 1 - \frac{\sum_{i=1}^{y-1} P(X = i)}{\mu} \right] = g(\mu).$$

Como  $g(\mu)$  é uma função estritamente crescente, então, como  $\mu \geq 1$ , por hipótese, tem-se:

$$P(Y \geq y) = g(\mu) \geq g(1) = 1 - \sum_{i=1}^{y-1} P(X = i) = \sum_{i=y}^{\infty} P(X = i) = P(X \geq y).$$

Logo,

$$P(Y \geq y) \geq P(X \geq y).$$

e, assim,  $Y \stackrel{\text{st}}{\geq} X$ .

Vai-se mostrar agora que para uma dada seqüência de taxas de transição  $\delta_y$  existe uma única distribuição discreta satisfazendo as equações prospectivas de Chapman-Kolmogorov. A recíproca também é verdadeira, isto é, uma dada distribuição de probabilidade tem um única representação em termos da seqüência das taxas de transição do processo estocástico de nascimento subjacente. Esse resultado será chamado de Teorema de Faddy.

### Teorema de Faddy

Para uma dada distribuição de probabilidade discreta  $\{\pi_y, y = 0, 1, 2, \dots, N\}$  existe uma única seqüência das taxas de transição  $\{\pi_y, y = 0, 1, 2, \dots, N\}$  tal que  $p_y(1) = \pi_y, y = 0, 1, 2, \dots, N$ , com  $\delta_0 = -\ln(p_0)$  e  $\delta_N = 0$ .

#### Prova.

(Faddy 1997a)

Sabe-se que  $p'_0(t) = -\delta_0 p_0(t)$  cuja solução é dada por  $p_0(t) = e^{-\delta_0(t)}$  e assim  $\delta_0 = -\frac{\ln(p_0(t))}{t}$ . Fazendo  $t = 1$  tem-se  $\delta_0 = -\ln(p_0(1)) = -\ln(p_0)$ . Vai-se agora provar por indução. Suponha que para algum  $m$  exista uma seqüência  $\delta_0, \delta_1, \dots, \delta_m$  tal que  $p_y(1) = \pi_y, y = 0, 1, 2, \dots, m$ . Será mostrado que isso se verifica para  $(m + 1)$  e, portanto, o resultado segue para todo  $m$ . Agora se  $\delta_{m+1} \rightarrow \infty$ , o estado  $(m + 1)$  é instantâneo e, portanto,  $p_{m+1} \rightarrow 0$ . Se  $\delta_{m+1} \rightarrow 0$ , o estado  $(m + 1)$  é absorvente e, portanto,  $p_{m+1}(1) = 1 - p_0(1) - p_1(1) - \dots - p_m(1)$ . Agora para  $0 < \pi_{m+1}(1) < 1 - p_0(1) - p_1(1) - \dots - p_m(1)$ , existe, então,  $\delta_{m+1} > 0$  com  $p_{m+1}(1) = \pi_{m+1}$ . Tal  $\pi_{m+1}$  é único, e satisfaz

$$p_{m+1}(1) = \int_0^1 \delta_{m+1} e^{-\delta_{m+1}(1-x)} p_m(x) dx.$$

Finalmente, se o suporte da distribuição discreta é  $0, 1, 2, \dots, N$ , então,  $N$  corresponde a um estado absorvente e, portanto,  $\delta_N = 0$ .

## APÊNDICE D - Exponencialização de uma Matriz

### 1. Introdução

Uma das mais usuais funções matriciais empregada na solução de sistemas lineares de equações diferenciais é a exponencialização de uma matriz que é definida, para qualquer matriz quadrada  $Q$  de ordem  $k$ , por:

$$e^Q = \exp(Q) = \sum_{j=0}^{\infty} \frac{Q^j}{j!}, \quad (89)$$

em que  $Q^0 = I_k$ .

A exponencialização de uma matriz aparece na equação diferencial

$$\frac{dx(t)}{dt} = Qx(t) \quad (90)$$

em que  $t \geq 0$  e  $x(t) = (x_1(t), x_2(t), \dots, x_k(t))'$  é uma função vetorial real. Se  $x_0 = a$ , então, a solução de (90) é:

$$x(t) = e^{Qt}a.$$

A definição da exponencialização de uma matriz é baseada nos polinômios de Taylor

$$T_n(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!}$$

que converge pontualmente e uniformemente para  $e^x$  quando  $x$  é limitado e  $n$  tende para o infinito. Assim, para qualquer matriz quadrada  $Q$ , seja a seqüência de matrizes

$$Q_n = I + \frac{Q}{1!} + \frac{Q^2}{2!} + \frac{Q^3}{3!} + \dots + \frac{Q^n}{n!}.$$

E assim, quando  $n$  cresce,  $\lim_{n \rightarrow \infty} Q_n = e^Q$ .

## 2. Exponencialização de uma matriz $Q$ diagonalizável

Se  $Q$  é diagonalizável existe uma matriz invertível  $B$  e uma matriz diagonal  $D = \text{diag}(d_1, d_2, \dots, d_k)$  tal que  $Q = B^{-1}DB$  então  $e^Q = Be^DB^{-1} = B\text{diag}(e^{d_1}, e^{d_2}, \dots, e^{d_k})B^{-1}$ , pois

$$Q^2 = B^{-1}DB \cdot B^{-1}DB = B^{-1}DB = Q,$$

e, mais geralmente

$$Q^j = B^{-1}D^jB^{-1}.$$

Logo,

$$e^{tQ} = \exp(tQ) = \sum_{j=0}^{\infty} \frac{t^j Q^j}{j!} = \sum_{j=0}^{\infty} \frac{t^j}{j!} B^{-1}D^jB^{-1} = B \left( \sum_{j=0}^{\infty} \frac{t^j D^j}{j!} \right) B^{-1} = Be^{tQ}B^{-1}.$$

## 3. Propriedades

1. Se  $AB = BA$ , então

$$e^{A+B} = e^A e^B = e^B e^A.$$

(Se  $AB \neq BA$ , a equação pode não se verificar.)

2. Para qualquer matriz quadrada  $Q$ ,  $e^Q$  é não singular e  $(e^Q)^{-1} = e^{-Q}$ .
3. Para quaisquer matrizes quadradas  $B$  e  $Q$ , não singulares de ordem  $k$ ,

$$e^{B^{-1}QB} = B^{-1}e^QB.$$

(Assim se  $A$  é semelhante a  $B$ , então  $e^A$  é semelhante a  $e^B$ ).

4. Se  $Q$  tem raízes características  $\lambda_1, \lambda_2, \dots, \lambda_k$  (não necessariamente distintos), então  $e^Q$  tem raízes características  $e^{\lambda_1}, e^{\lambda_2}, \dots, e^{\lambda_k}$ .
5.  $\det(e^Q) = e^{\text{tr}(Q)}$ , em que  $\text{tr}(Q)$  = traço da matriz  $Q$ .
6.  $Qe^Q = e^QQ$

**Exemplos:**

1. Calcular  $e^{tQ}$  para a matriz  $Q = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$ .

*Solução:* A matriz  $Q$  tem duas raízes características distintas  $\lambda_1 = 6$  e  $\lambda_2 = 1$ , assim existe uma matriz não singular  $B$  tal que  $B^{-1}QB = D$ , em que  $D = \text{diag}(\lambda_1, \lambda_2) = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}$ . Para determinar a matriz  $C$  pode-se usar que  $QB = BD$ , ou

$$\begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}$$

Portanto

$$\begin{bmatrix} 5a + 4c & 5b + 4d \\ a + 2c & b + 2d \end{bmatrix} = \begin{bmatrix} 6a & b \\ 6c & d \end{bmatrix}.$$

Da igualdade de matrizes tem-se que  $a = 4c$  e  $b = -d$ . Tomando-se  $c = d = 1$ , obtém-se:

$$B = \begin{bmatrix} 4 & -1 \\ 1 & 1 \end{bmatrix}, B^{-1} = \frac{1}{5} \begin{bmatrix} 4 & -1 \\ 1 & 1 \end{bmatrix}.$$

Assim,

$$e^{tQ} = Be^{tD}B^{-1} = \frac{1}{5} \begin{bmatrix} 4 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{6t} & 0 \\ 0 & e^t \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 4 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 4e^{6t} + e^t & 4e^{6t} - 4e^t \\ e^{6t} - e^t & e^{6t} + 4e^t \end{bmatrix}$$

2. Ache a solução do sistema linear

$$\begin{bmatrix} x_1' = 5x_1 + 4x_2 \\ x_2' = x_1 + 2x_2 \end{bmatrix},$$

sujeito às condições iniciais  $x_1(0) = 2$  e  $x_2(0) = 3$ .

*Solução:* Em linguagem matricial tem-se

$$X'(t) = QX(t), \quad X(0) = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \text{em que } Q = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}.$$

Assim a solução é dada por  $X(t) = e^{tQ} X(0)$ . Usando a matriz  $e^{tQ}$  calculada no exemplo 1 acha-se:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 4e^{6t} + e^t & 4e^{6t} - 4e^t \\ e^{6t} - e^t & e^{6t} + 4e^t \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix},$$

e finalmente,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4e^{6t} - 2e^t \\ e^{6t} + 2e^t \end{bmatrix}$$

Assim, até o momento sabe-se como calcular  $e^{tQ}$  para o caso em que  $Q$  é diagonalizável, vai-se então enunciar o teorema de Cayley-Hamilton que será usado para o cálculo de  $e^{tQ}$ ,  $Q$  geral.

### Teorema de Cayley-Hamilton

Seja  $Q$  uma matriz quadrada de ordem  $k$  e seja

$$f(\lambda) = \det(\lambda I - Q) = \lambda^k + c_{k-1}\lambda^{k-1} + \dots + c_1\lambda + c_0$$

seu polinômio característico. Então  $F(Q) = 0$ . Assim,  $Q$  satisfaz à equação matricial

$$Q^k + c_{k-1}Q^{k-1} + \dots + c_1Q + c_0I_k = 0.$$

O teorema de Cayley-Hamilton mostra como a  $k$ -ésima potência de qualquer matriz quadrada de ordem  $k$  pode ser expressa como combinação linear das potências de ordens menores do que  $k$ ,  $I, Q, Q^2, \dots, Q^{k-1}$ . Segue que as potências maiores  $Q^{k+1}, Q^{k+2}, \dots$ , também podem ser expressas como uma combinação linear de  $I, Q, Q^2, \dots, Q^{k-1}$ . Logo, na série infinita definidora de  $e^{tQ}$ , cada termo  $\frac{t^j Q^j}{j!}$ . Assim,

$$e^{tQ} = \sum_{j=0}^{k-1} q_j(t)Q^j,$$

em que os coeficientes escalares  $q_j(t)$  dependem de  $t$ .

**Teorema.** Sejam  $\lambda_1, \lambda_2, \dots, \lambda_k$  os auto-valores de  $Q$  e a seguinte sequência de polinômios em  $Q$ :

$$P_0(Q) = I, P_k(Q) = \prod_{l=1}^k (Q - \lambda_l I) \text{ para } j = 1, 2, \dots, k.$$

Logo, tem-se:

$$e^{tA} = \sum_{j=0}^{k-1} r_{j+1}(t) P_j(Q),$$

em que  $r_1(t), r_2(t), \dots, r_k(t)$  são recursivamente determinados a partir do sistema de equações diferenciais

$$r'_1(t) = \lambda_1 r_1(t), r_1(0) = 1, r'_{j+1}(t) = \lambda_{j+1} r_{j+1}(t) + r_j(t), r_{j+1}(0) = 0, j = 1, 2, \dots, k - 1.$$

**Exemplo.**

Expresse  $e^{tQ}$  como uma combinação linear de  $I$  e  $Q$ , se  $Q$  é uma matriz quadrada de ordem 2, com autovalores iguais.

*Solução:* Quando  $\lambda_1 = \lambda_2 = \lambda$  tem-se:

$$\begin{cases} r'_1(t) = \lambda r_1(t), & r_1(0) = 1 \\ r'_2(t) = \lambda r_2(t) + r_1(t), & r_2(0) = 0 \end{cases},$$

Resolvendo o sistema tem-se

$$r_1(t) = e^{\lambda t} \text{ e } r_2(t) = te^{\lambda t}.$$

Desde que  $P_0(Q) = I$  e  $P_1(Q) = Q - \lambda I$ , a fórmula para  $e^{tQ}$  é

$$e^{tQ} = e^{\lambda t} I + te^{\lambda t} (Q - \lambda I) = e^{\lambda t} (1 - \lambda t) I + te^{\lambda t} Q.$$

**Exemplo.**

Resolva o exemplo anterior se os auto-valores de  $Q$  são  $\lambda$  e  $\mu$ , em que  $\lambda \neq \mu$ .

$$r'_1(t) = \lambda r_1(t), r_1(0) = 1, r'_2(t) = \mu r_2(t) + r_1(t), r_2(0) = 0.$$

Resolvendo o sistema tem-se

$$r_1(t) = e^{\lambda t}, \quad r_2(t) = \frac{e^{\lambda t} - e^{\mu t}}{\lambda - \mu}.$$

Desde que  $P_0(Q) = I$  e  $P_1(Q) = Q - \lambda I$ , a fórmula para  $e^{tQ}$  é

$$e^{tQ} = e^{\lambda t}I + \frac{\lambda e^{\lambda t} - \mu e^{\mu t}}{\lambda - \mu}(Q - \lambda I) = \frac{\lambda e^{\lambda t} - \mu e^{\mu t}}{\lambda - \mu}I + \frac{e^{\lambda t} - e^{\mu t}}{\lambda - \mu}Q.$$

Se os auto-valores  $\lambda$  e  $\mu$  são complexos, as exponenciais  $e^{\lambda t}$  e  $e^{\mu t}$  também podem ser números complexos. Mas se  $\lambda$  e  $\mu$  são complexos conjugados, os coeficientes de  $I$  e  $Q$  serão reais. Se

$$\lambda = \alpha + i\beta, \quad \mu = \alpha - i\beta, \quad \beta \neq 0.$$

Logo,  $\lambda - \mu = 2i\beta$ , e, portanto

$$\begin{aligned} e^{tQ} &= e^{(\alpha+i\beta)t}I + \frac{e^{(\alpha+i\beta)t} - e^{(\alpha-i\beta)t}}{2i\beta}(Q - (\alpha + i\beta)I) \\ &= e^{\alpha t} \left[ e^{\beta i t}I + \frac{e^{\beta i t} - e^{-\beta i t}}{2i\beta}(Q - \alpha I - i\beta I) \right] \\ &= e^{\alpha t} \left[ (\cos \beta t + i \operatorname{sen} \beta t) + \frac{\operatorname{sen} \beta t}{\beta}(Q - \alpha I - i\beta I) \right] \\ &= \frac{e^{\alpha t}}{\beta} [(\beta \cos \beta t - \alpha \operatorname{sen} \beta t)I + \operatorname{sen} \beta t Q]. \end{aligned}$$

Fato 1. Se  $Q$  é uma matriz quadrada de ordem  $k$  com todos os auto-valores iguais a  $\lambda$ , então tem-se :

$$e^{tQ} = e^{\lambda t} \sum_{j=0}^{k-1} \frac{t^j}{j!} (Q - \lambda I)^j.$$

Fato 2. Se  $Q$  é uma matriz quadrada de ordem  $k$  com todos os auto-valores distintos  $\lambda_1, \lambda_2, \dots, \lambda_k$ , então, tem-se :

$$e^{tQ} = \sum_{j=0}^{k-1} e^{\lambda_j t} L_j(Q),$$

em que  $L_j(Q)$  é um polinômio de grau  $(k-1)$  dado por :

$$L_j(Q) = \prod_{l=1}^k \frac{Q - \lambda_l I}{\lambda_k - \lambda_l},$$



para  $j = 1, 2, \dots, k$ .

Observação: Os polinômios  $L_j(Q)$  são chamados de coeficientes de interpolação de Lagrange.

Fato 3. Se  $Q$  é uma matriz quadrada de ordem  $k \geq 3$  com dois auto-valores distintos  $\lambda$  com multiplicidade  $(k - 1)$  e  $\mu$  com multiplicidade 1. Assim, a exponencialização da matriz  $Q$  é dada por:

$$e^{tQ} = e^{\lambda t} \sum_{j=0}^{k-2} \frac{t^j}{j!} (Q - \lambda I)^j + \left[ \frac{e^{\mu t}}{(\mu - \lambda)^{k-1}} - \frac{e^{\lambda t}}{(\mu - \lambda)^{k-1}} \sum_{j=0}^{k-2} \frac{t^j}{j!} (Q - \lambda I)^j \right] (Q - \lambda I)^{k-1}.$$

A prova desse resultado é encontrada no livro do Apostol (1969), bem como outros resultados sobre a exponencialização de uma matriz quadrada.

A exponencialização de uma matriz  $Q$  pode ser feita no MAPPLE através dos seguintes comandos:

```
> with(linalg,exponential):
```

```
> Q:=array([[5,4],[1,2]]);
```

$$Q := \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$$

```
> e^Q=exponential(Q);
```

$$e = \begin{bmatrix} \frac{1}{5}e + \frac{4}{5}e^6 & \frac{4}{5}e^6 - \frac{4}{5}e \\ \frac{1}{5}e^6 - \frac{1}{5}e & \frac{4}{5}e + \frac{1}{5}e^6 \end{bmatrix}$$

```
> e^Qt=exponential(Q,t);
```

$$e^{Qt} = \begin{bmatrix} \frac{1}{5}e^t + \frac{4}{5}e^{(6t)} & \frac{4}{5}e^{(6t)} - \frac{4}{5}e^t \\ \frac{1}{5}e^{(6t)} - \frac{1}{5}e^t & \frac{4}{5}e^t + \frac{1}{5}e^{(6t)} \end{bmatrix}$$

### APÊNDICE E - Modelo de Poisson Generalizado

A distribuição generalizada de Poisson (G.P.D) de parâmetros  $\theta > 0$  e  $\lambda \in (-1, 1)$  tem função de probabilidade definida matematicamente, para  $\lambda > 0$ , por

$$f_X(x | \theta, \lambda) = \frac{\theta(\theta + x\lambda)^{x-1} e^{-(\theta+x\lambda)}}{x!} I_{\{0,1,2,\dots\}}(x). \quad (91)$$

Para mostrar que (91) é realmente uma função de probabilidade deve-se provar que:

$$e^\theta = \sum_{x=0}^{\infty} \frac{\theta(\theta + x\lambda)^{x-1} e^{-x\lambda}}{x!}. \quad (92)$$

Esse somatório é uma aplicação da fórmula de Lagrange apresentada na página 308 em Jensen (1902). Dadas duas funções holomorfas  $\phi(z)$  e  $g(z)$  então:

$$\phi(z) = \phi(0) + \sum_{x=1}^{\infty} \frac{1}{x!} \left[ \frac{d^{z-1}}{dx^{z-1}} [(g(z))^x \phi'(z)] \right]_{z=0} \left( \frac{z}{g(z)} \right)^x. \quad (93)$$

Considere as funções  $\phi(z) = e^{\theta z}$  e  $g(z) = e^{\lambda z}$ . Assim:

- i.  $\phi(0) = 1$  ;
- ii.  $\phi'(z) = \theta e^{\theta z}$  e  $(g(z))^x = e^{\lambda z x}$  ;
- iii.  $(g(z))^x \phi'(z) = \theta e^{(\lambda x + \theta)z}$  ;
- iv.  $\left[ \frac{d^{z-1}}{dx^{z-1}} [(g(z))^x \phi'(z)] \right] = \theta(\lambda x + \theta)^{x-1} e^{(\lambda x + \theta)z}$  ;
- v.  $\left[ \frac{d^{z-1}}{dx^{z-1}} [(g(z))^x \phi'(z)] \right]_{z=0} = \theta(\lambda x + \theta)^{x-1}$  ;
- vi.  $\left( \frac{z}{g(z)} \right)^x = z^x e^{-\lambda z x}$ .

Substituindo os resultados (i) a (vi) em (93), tem-se:

$$e^{\theta z} = 1 + \sum_{x=1}^{\infty} \frac{1}{x!} (\lambda x + \theta)^{x-1} z^x e^{-\lambda z x} \quad (94)$$

Fazendo  $z = 1$  em (94), tem-se :

$$e^\theta = 1 + \sum_{x=1}^{\infty} \frac{1}{x!} (\lambda x + \theta)^{x-1} e^{-\lambda x}, \quad (95)$$

e que ainda pode ser posto na forma

$$e^\theta = \sum_{x=0}^{\infty} \frac{1}{x!} (\lambda x + \theta)^{x-1} e^{-\lambda x}, \quad (96)$$

o que origina

$$\sum_{x=0}^{\infty} \frac{1}{x!} (\lambda x + \theta)^{x-1} e^{-\lambda x - \theta} = 1. \quad (97)$$

Essa distribuição permite que o valor de  $\lambda$  seja negativo e nesse caso há um truncamento nos valores de  $x$ . Quando  $\lambda$  for negativo, precisa-se ter  $(\theta + x\lambda) > 0$  de modo que  $\lambda > \frac{-x}{\theta}$ . Seja  $m$  o maior inteiro positivo para o qual  $(\theta + m\lambda) > 0$ . Assim,  $\lambda > \frac{-m}{\theta}$  e  $\lambda > -1$  implicam em uma nova restrição para esse parâmetro que é  $\lambda > \max(\frac{-m}{\theta}, -1)$ . O limite inferior para  $m$  é imposto para que haja pelo menos cinco classes com probabilidades não nulas quando  $\lambda$  é negativo. Nessa situação, a f.p. de  $X$  é dada por:

$$f_X(x | \theta, \lambda) \approx \frac{\theta(\theta + x\lambda)^{x-1} e^{-(\theta + x\lambda)}}{x!} I_{\{0,1,2,\dots,m\}}(x). \quad (98)$$

Quando  $\lambda$  é negativo, o modelo *GPD* definido por (91) inclui um truncamento devido ao fato que  $P(X = x) = 0$  para  $x > m$ . Assim,

$$F_m(\theta, \lambda) = \sum_{x=0}^m f_X(x | \theta, \lambda)$$

poderá assumir um valor maior ou menor do que 1 dependendo dos valores dos parâmetros. Na prática, segundo Consul (1989), os valores de  $m$ ,  $\lambda$  e  $\theta$  que surgem nas aplicações tornam desnecessário o uso de qualquer fator de normalização. Os símbolos  $\theta$  e  $\lambda$  são chamados de primeiro e segundo parâmetros do modelo, respectivamente. O segundo parâmetro  $\lambda$  é um indicador de superdispersão, de subdispersão, ou não, em relação à distribuição de Poisson. Esse parâmetro pode ser pensado como uma medida de afastamento dos dados em análise em relação à distribuição de Poisson. Ele, ainda, reflete uma taxa média de restituição do processo. Um valor positivo indica que os indivíduos estão fazendo um esforço para acelerar o processo enquanto um valor negativo significa um retardo. Quando  $\lambda = 0$ , tem-se a distribuição de Poisson. O primeiro parâmetro  $\theta$  é a taxa média natural de trocas

do processo para a ocorrência dos eventos analisados e é um indicador da intensidade do processo de Poisson.

Daqui por diante será usada a notação

$$(\theta, \lambda)_x = \frac{\theta(\theta + x\lambda)^{x-1}}{x!}, \quad (99)$$

e a f.p. de  $X \sim \text{PoissonGEN}(\theta, \lambda)$  é dada por:

$$f(x; \theta, \lambda) = (\theta, \lambda)_x e^{-\theta - \lambda x} I_{\{0,1,2,\dots\}}(x), \quad (100)$$

que é mais conveniente para provar vários resultados.

Em muitas situações práticas, é sabido que o segundo parâmetro,  $\lambda$ , é linearmente proporcional ao primeiro  $\theta$ . Fazendo  $\lambda = \alpha\theta$  no modelo GPD tem-se:

$$\begin{aligned} f_X(x|\theta, \alpha\theta) &= \frac{(1 + x\alpha)^{x-1} \theta^x e^{-(\theta + x\alpha\theta)}}{x!} I_{\{0,1,2,\dots\}}(x) \\ &= (1, \alpha)_x (\theta e^{-\alpha\theta})^x e^{-\theta} I_{\{0,1,2,\dots\}}(x), \end{aligned} \quad (101)$$

que é conhecido como modelo restrito.

O domínio de  $\alpha$  é  $\max(-\theta^{-1}, -m^{-1}) < \alpha < \theta^{-1}$ .

Vai-se utilizar o resultado (92), com  $\theta > 0$  e  $\lambda > 0$  para especificar as probabilidades para valores particulares da Poisson Generalizada bem como seus quatro primeiros momentos.

- i.  $P(X = 0) = e^{-\theta}$ ;
- ii.  $P(X = 1) = \theta e^{-\theta - \lambda} = \theta e^{-\theta} e^{-\lambda}$ ;
- iii.  $P(X = 2) = \frac{\theta(\theta + \lambda)e^{-\theta - 2\lambda}}{2}$ , e
- iv.  $P(X = x) = \frac{(\theta + x\lambda)^{x-1} e^{-\lambda}}{(\theta + x\lambda)^{x-1} x} P(X = x - 1), x = 1, 2, \dots$

#### Momentos da distribuição de Poisson Generalizada:

- i.  $E(X) = \frac{\theta}{1 - \lambda}$ ;
- ii.  $\mu_2 = V(X) = \frac{\theta}{(1 - \lambda)^3}$ ;

$$\text{iii. } \mu_3 = \frac{\theta(1+2\lambda)}{(1-\lambda)^5};$$

$$\text{iv. } \mu_4 = \frac{3\theta^2}{(1-\lambda)^6} + \frac{\theta(1+8\lambda+6\lambda^2)}{(1-\lambda)^7}.$$

O coeficiente de assimetria e o coeficiente de curtose são dados por:

$$\text{i. } \gamma_1 = \frac{(1+2\lambda)}{\sqrt{(1-\lambda)\theta}};$$

$$\text{ii. } \beta_2 = 3 + \frac{(1+8\lambda+6\lambda^2)}{(1-\lambda)\theta}.$$

### Valor Esperado no Modelo Restrito:

Desde que:

$$e^\theta = \sum_{x=0}^{\infty} (1, \alpha)_x \theta^x e^{-x\alpha\theta} \quad (102)$$

Derivando em relação a  $\theta$  ambos os membros de (102), tem-se:

$$\begin{aligned} e^\theta &= \sum_{x=0}^{\infty} [x\theta^{-1} - x\alpha](1, \alpha)_x \theta^x e^{-x\alpha\theta} \\ &= \sum_{x=0}^{\infty} [\theta^{-1} - \alpha]x(1, \alpha)_x \theta^x e^{-x\alpha\theta} \\ &= [\theta^{-1} - \alpha] \sum_{x=0}^{\infty} x(1, \alpha)_x \theta^x e^{-x\alpha\theta} \\ &= \frac{1-\alpha}{\theta} \sum_{x=0}^{\infty} x(1, \alpha)_x \theta^x e^{-x\alpha\theta}, \end{aligned} \quad (103)$$

e multiplicando-se (103) por  $e^{-\theta} \frac{\theta}{1-\alpha\theta}$ , tem-se:

$$\sum_{x=0}^{\infty} x(1, \alpha)_x \theta^x e^{\theta-x\alpha\theta} = \frac{\theta}{1-\alpha\theta}.$$

Assim,

$$E(X) = \sum_{x=0}^{\infty} x(1, \alpha)_x \theta^x e^{\theta-x\alpha\theta} = \frac{\theta}{1-\alpha\theta} = \frac{\theta}{1-\lambda}. \quad (104)$$

Para determinar  $E(X^2)$ , deriva-se (110) em relação a  $\theta$ ,

$$\begin{aligned}
\frac{1}{(1-\alpha\theta)^2} &= \sum_{x=0}^{\infty} x(1,\alpha)_x \theta^x e^{-\theta-x\alpha\theta} \left[ \frac{x}{\theta} - 1 - x\alpha \right] \\
&= \sum_{x=0}^{\infty} x(1,\alpha)_x \theta^x e^{-\theta-x\alpha\theta} \left[ -1 + x\left(\frac{1}{\theta} - \alpha\right) \right] \\
&= \sum_{x=0}^{\infty} x(1,\alpha)_x \theta^x e^{-\theta-x\alpha\theta} \left[ -1 + x\left(\frac{1-\theta\alpha}{\theta}\right) \right] \\
&= -\sum_{x=0}^{\infty} x(1,\alpha)_x \theta^x e^{-\theta-x\alpha\theta} + \frac{1-\theta\alpha}{\theta} \sum_{x=0}^{\infty} x^2(1,\alpha)_x \theta^x e^{-\theta-x\alpha\theta} \\
&= -E(X) + \frac{1-\theta\alpha}{\theta} E(X^2).
\end{aligned} \tag{105}$$

Assim,

$$E(X^2) = \frac{1}{(1-\alpha\theta)^3} + \frac{\theta}{(1-\alpha\theta)^2}$$

e dessa maneira,

$$V(X) = \frac{1}{(1-\alpha\theta)^3} + \frac{\theta}{(1-\alpha\theta)^2} - \frac{\theta^2}{(1-\alpha\theta)^2} = \frac{\theta}{(1-\lambda)^3} \tag{106}$$

### Esperança da Poisson Generalizada:

Seja  $E(X) = \mu(\theta, \lambda)$  o valor esperado da Poisson Generalizada com parâmetros  $\theta, \lambda$ . Então,

$$\begin{aligned}
 \mu(\theta, \lambda) &= \sum_{x=0}^{\infty} x \frac{\theta(\theta + x\lambda)^{x-1} e^{-(\theta+x\lambda)}}{x!} \\
 &= \sum_{x=1}^{\infty} x \frac{\theta(\theta + x\lambda)^{x-1} e^{-(\theta+x\lambda)}}{x!} \\
 &= \sum_{x=1}^{\infty} \frac{\theta(\theta + \lambda + (x-1)\lambda)^{x-1} e^{-(\theta+\lambda+(x-1)\lambda)}}{(x-1)!} \\
 &= \sum_{y=0}^{\infty} \frac{\theta(\theta + \lambda + y\lambda)^y e^{-(\theta+\lambda+y\lambda)}}{y!}, \tag{107}
 \end{aligned}$$

em que  $y=x-1$ . Seja a relação:

$$(\theta + \lambda + y\lambda)^y = (\theta + \lambda + y\lambda)(\theta + \lambda + y\lambda)^{y-1}.$$

que, substituída em (107), resulta em:

$$\begin{aligned}
 \mu(\theta, \lambda) &= \sum_{y=0}^{\infty} \frac{\theta(\theta + \lambda + y\lambda)(\theta + \lambda + y\lambda)^{y-1} e^{-(\theta+\lambda+y\lambda)}}{y!} \\
 &= \theta \left[ \sum_{y=0}^{\infty} \frac{(\theta+\lambda)(\theta+\lambda+y\lambda)^{y-1} e^{-(\theta+\lambda+y\lambda)}}{y!} + \lambda \sum_{y=0}^{\infty} y \frac{(\theta+\lambda+y\lambda)^{y-1} e^{-(\theta+\lambda+y\lambda)}}{y!} \right] \\
 &= \theta \left[ \sum_{y=0}^{\infty} \frac{(\theta+\lambda)(\theta+\lambda+y\lambda)^{y-1} e^{-(\theta+\lambda+y\lambda)}}{y!} + \frac{\lambda}{\theta+\lambda} \sum_{y=0}^{\infty} y \frac{(\theta+\lambda)(\theta+\lambda+y\lambda)^{y-1} e^{-(\theta+\lambda+y\lambda)}}{y!} \right] \\
 &= \theta \left[ \sum_{y=0}^{\infty} f(y; \theta + \lambda, \lambda) + \frac{\lambda}{\theta + \lambda} \sum_{y=0}^{\infty} y f(y; \theta + \lambda, \lambda) \right] \tag{108}
 \end{aligned}$$

Tem-se, então, a equação de recorrência

$$\mu(\theta, \lambda) = \theta + \frac{\theta\lambda}{\theta + \lambda} \mu(\theta + \lambda, \lambda), \tag{109}$$

em que  $\mu(\theta + \lambda, \lambda)$  é a média de uma Poisson Generalizada de parâmetros  $(\theta + \lambda, \lambda)$ .

Aplicando-se, sucessivamente, a equação de recorrência, tem-se:

$$\begin{aligned}
\mu(\theta, \lambda) &= \theta + \frac{\theta\lambda}{\theta + \lambda} \mu(\theta + \lambda, \lambda) \\
&= \theta + \frac{\theta\lambda}{(\theta + \lambda)} \left[ (\theta + \lambda) + \frac{(\theta + \lambda)\lambda}{\theta + 2\lambda} \mu(\theta + 2\lambda, \lambda) \right] \\
&= \theta + \theta\lambda + \frac{\theta\lambda^2}{\theta + 2\lambda} [\mu(\theta + 2\lambda, \lambda)] \\
&= \theta + \theta\lambda + \frac{\theta\lambda^2}{\theta + 2\lambda} \left[ (\theta + 2\lambda) + \frac{(\theta + 2\lambda)\lambda}{\theta + 3\lambda} \mu(\theta + 3\lambda, \lambda) \right] \\
&= \theta + \theta\lambda + \theta\lambda^2 + \frac{\theta\lambda^3}{\theta + 3\lambda} [\mu(\theta + 3\lambda, \lambda)] \\
&= \theta + \theta\lambda + \theta\lambda^2 + \theta\lambda^3 + \dots \\
&= \frac{\theta}{1 - \lambda} \text{ se } \lambda < 1.
\end{aligned}$$

Propriedade: A f.g.p. de  $X \sim PoiGen(\theta, \lambda)$ , no ponto  $u$ ,  $G_X(u)$ , é dada por:

$$G_X(u) = e^{\theta(t-1)}, \text{ em que } t = ue^{\lambda(t-1)}. \quad (110)$$

**Propriedade da convolução:** Se  $X \sim PoiGen(\theta_1, \lambda)$  e  $Y \sim PoiGen(\theta_2, \lambda)$ ,  $X$  e  $Y$  independentes, então  $S = X + Y \sim PoiGen(\theta_1 + \theta_2, \lambda)$ .

**Prova**

Seja  $G_S(u)$  a f.g.p. de  $S$ . Assim,  $G_S(u) = G_X(u)G_Y(u) = e^{\theta_1(t-1)}e^{\theta_2(t-1)} = e^{(\theta_1+\theta_2)(t-1)}$ , em que  $t = ue^{\lambda(t-1)}$ , e, portanto,

$$S = X + Y \sim PoiGen(\theta_1 + \theta_2, \lambda).$$



## ANEXOS

## ANEXO A

**Programa para obter as estimativas  
do Modelo M1 de Bakker- método 1: “média e frequência zero”.**

```

proc iml;
reset noname;

/*ESTIMATIVAS*/ ;
n=246;
f0=19/n;
f1=169/n;
f2=50/n;
f3=6/n;
f4=1/n;
f5=1/n;
y=0*f0+1*f1+2*f2+3*f3+4*f4+5*f5;
lambda_hat=-log(f0);
delta=(y-1+f0)/(lambda_hat-1+f0);
print 'ESTIMATIVAS PELO MÉTODO 1 DA MÉDIA E FREQUÊNCIA ZERO';
print 'lambda=' lambda ' ' 'delta=' delta;

/*FREQUÊNCIA ESPERADA*/
f0=0.07723576902; f1=0.6850773966; f2=0.3994505285/2;
f3=0.2013387220/6; f4=0.09607391862/24; f5=0.04461268004/120;
zero=f0*n; um=f1*n; dois=f2*n; tres=f3*n; quatro=f4*n; cinco=f5*n;
print ' ';
print 'NÚMERO DE OVOS ESPERADOS NOS HOSPEDEIROS';
print '0= 'zero; print '1= 'um; print '2= ' dois;
print '3= ' tres; print '4= ' quatro; print '5= ' cinco;

/*ESTIMATIVAS DAS VARIÂNCIAS E COVARIÂNCIA*/
freq0=exp(-lambda);
u=lambda*delta+(1-exp(-lambda))*(1-delta);

```

```

varf0=f0*(1-f0)/n;
varx=((lambda**2*delta**2)-u**2+((1-exp(-lambda))*(1-delta)*(1-2*delta))+
(lambda*delta*(3-2*delta)))/n;
cov=-f0*u/n;
dFlambda=-exp(-lambda);
dFdelta=0;
dUlambda=delta+exp(-lambda)*(1-delta);
dUdelta=lambda-1+exp(-lambda);
dFlambd2=dFlambda**2;
dFdelta2=dFdelta**2;
dUlambda2=dUlambda**2;
dUdelta2=dUdelta**2;
A=(dFlambd2||dFdelta2||2*dFlambda*dFdelta)//
(dUlambda2||dUdelta2||2*dUlambda*dUdelta)//
(dFlambda*dUlambda||dFdelta*dUdelta||((dFlambda*dUdelta)+(dFdelta*dUlambda)));
varobs=varf0//varx//cov;
VarEsp=inv(A)*varobs;
print 'ERROS-PADRÃO E CORRELAÇÃO DOS ESTIMADORES DE LAMBDA E DELTA';
SEL=sqrt(VarEsp[1]);
SED=sqrt(VarEsp[2]);
COR=VarEsp[3]/(SEL*SED);
print ' ';
print 'DESVIOS-PADRÃO E CORRELAÇÃO';
print 'Erro-Padrão ( $\hat{\lambda}$ )= ' SEL;
print 'Erro-Padrão ( $\hat{\delta}$ )= ' SED;
print 'Correlação Estimada= ' COR;
quit;

```

## ESTIMATIVAS MÉTODO DE MÉDIA E FREQUÊNCIA ZERO

lambda= 2.5608926

delta= 0.1712246

## NUMERO DE OVOS ESPERADOS NOS HOSPEDEIROS

0= 18.999999

1= 168.52904

2= 49.132415

3= 8.2548876

4= 0.9847577

5= 0.091456

## ERROS-PADRÃO E CORRELAÇÃO DE LAMBDA E DELTA

Erro-Padrão (Lambda)= 0.2203782

Erro-Padrão (Delta)= 0.0277311

Correlação= -0.652603

**Programa para obter as estimativas  
do Modelo M1 de Bakker- método 2: "frequência zero e um".**

```

proc iml;
reset noname;

/*chute inicial*/ ;
lambda=2.5;
delta=0.17;
p0=0.077236;
p1=0.686992;
f=exp(-lambda)-p0;
g=exp(-lambda*delta)-exp(-lambda)-p1+delta*p1;
dfl=-exp(-lambda);
dfd=0;
dgl=-delta*exp(-lambda*delta)+exp(-lambda);
dgd=-lambda*exp(-lambda*delta)+p1;
jacob=(dfl|dfd)/(dgl|dgd);
teta0=lambda//delta;
funcao0=f//g;
tetan=teta0-inv(jacob)*funcao0;
lambda=tetan[1]; delta=tetan[2];

/*inicio do processo iterativo*/
i=0;
do while (max(abs(tetan-teta0)) > 0.000001);
f=exp(-lambda)-p0;
g=exp(-lambda*delta)-exp(-lambda)-p1+delta*p1;
dfl=-exp(-lambda);
dfd=0;
dgl=-delta*exp(-lambda*delta)+exp(-lambda);
dgd=-lambda*exp(-lambda*delta)+p1;

```

```

jacob=(df1|dfd)/(dgl|dgd);
teta0=lambda//delta;
funcao0=f//g;
teta0=tetan;
tetan=teta0-inv(jacob)*funcao0;
lambda=tetan[1]; delta=tetan[2];
i=i+1;
end;
tetaf=tetan';
print 'ESTIMATIVAS MÉTODO DAS FREQUENCIAS ZERO E UM';
print 'lambda=' lambda ' ' 'delta=' delta;
print 'iterações' i ' ' 'convergencia' (max(abs(tetan-teta0))) ;

/*FREQUENCIA ESPERADA*/
f0=0.07723600073; f1=0.6869920567; f2=0.3969015934/2;
f3=0.1981805179/6; f4=0.09367392413/24; f5=0.04308576809/120;
zero=f0*246; um=f1*246; dois=f2*246; tres=f3*246;
quatro=f4*246; cinco=f5*246;
print ' ';
print 'NUMERO DE OVOS ESPERADOS NOS HOSPEDEIROS';
print '0= 'zero; print '1= 'um; print '2= ' dois;
print '3= ' tres; print '4= ' quatro; print '5= ' cinco;

/*ESTIMATIVAS DAS VARIÂNCIAS E COVARIÂNCIA*/
n=246;
freq0=exp(-lambda);
freq1=(exp(-lambda*delta)-exp(-lambda))/(1-delta);
varf0=f0*(1-f0)/n;
varf1=f1*(1-f1)/n;
cov=-f0*f1/n;
d0lambda=-exp(-lambda);
d0delta=0;
d1lambda=(-delta*exp(-lambda*delta)+exp(-lambda))/(1-delta);

```

```

d1delta=(exp(-lambda*delta)-exp(-lambda))/((1-delta)**2)-
lambda*exp(-lambda*delta)/(1-delta);
A=(d0lambda**2||d0delta**2||2*d0lambda*d0delta)//
(d1lambda**2||d1delta**2||2*d1lambda*d1delta)//
(d0lambda*d1lambda||d0delta*d1delta||((d0lambda*d1delta)+(d0delta*d1lambda)));
varobs=varf0//varf1//cov;
Matriz=inv(A)*varobs;
SEL=sqrt(Matriz[1]);
SED=sqrt(Matriz[2]);
COR=Matriz[3]/(SEL*SED);
print ' ';
print ' ERROS-PADRÃO E CORRELAÇÃO DE LAMBDA E DELTA ';
print 'Erro-Padrão (Lambda)= ' SEL;
print ' Erro-Padrão (Delta)= ' SED;
print ' Correlação= ' COR;
quit;

```

## ESTIMATIVAS MÉTODO DAS FREQUENCIAS ZERO E UM

lambda= 2.5608896

delta= 0.1695859

iterações 4

convergencia 5.872E-10

## NUMERO DE OVOS ESPERADOS NOS HOSPEDEIROS

0= 19.000056

1= 169.00005

2= 48.818896

3= 8.1254012

4= 0.9601577

5= 0.0883258

## ERROS-PADRÃO E CORRELAÇÃO DE LAMBDA E DELTA

Erro-Padrão (Lambda)= 0.2203778

Erro-Padrão (Delta)= 0.0292093

Correlação= -0.623914



**Programa para obter as estimativas  
do Modelo M1 de Bakker- método 3: método dos "momentos".**

```

proc iml;
reset noname; /*chute inicial*/ ;

n=246;
lambda=2.5;
delta=0.17;
m=1.203;
v=1.886;
f=lambda*delta+(1-exp(-lambda))*(1-delta)-m;
g=lambda**2*delta**2+(1-exp(-lambda))*(1-delta)*(1-2*delta)+
lambda*delta*(3-2*delta)-v;
dff=delta+(exp(-lambda)*(1-delta));
dfd=lambda-1+exp(-lambda);
dgl=2*lambda*delta**2+exp(-lambda)*(1-delta)*(1-2*delta)+delta*(3-2*delta);
dgd=2*delta*lambda**2-(1-exp(-lambda))*(1-2*delta)-2*(1-exp(-lambda))*
(1-delta)+lambda*(3-2*delta)-2*lambda*delta;
jacob=(dff||dfd)/(dgl||dgd);
teta0=lambda//delta;
funcao0=f//g;
tetan=teta0-inv(jacob)*funcao0;
lambda=tetan[1]; delta=tetan[2];

/*inicio do processo iterativo*/
i=0;
do while (max(abs(tetan-teta0)) >0.000001);
f=lambda*delta+(1-exp(-lambda))*(1-delta)-m;
g=lambda**2*delta**2+(1-exp(-lambda))*(1-delta)*(1-2*delta)+
lambda*delta*(3-2*delta)-v;
dff=delta+(exp(-lambda)*(1-delta));
dfd=lambda-1+exp(-lambda);

```

```

dgl=2*lambda*delta**2+exp(-lambda)*(1-delta)*(1-2*delta)+delta*(3-2*delta);
dgd=2*delta*lambda**2-(1-exp(-lambda))*(1-2*delta)-2*(1-exp(-lambda))*
(1-delta)+lambda*(3-2*delta)-2*lambda*delta;
jacob=(df1|dfd)/(dgl|dgd);
teta0=lambda//delta;
funcao0=f//g;
teta0=tetan;
tetan=teta0-inv(jacob)*funcao0;
lambda=tetan[1]; delta=tetan[2];
i=i+1;
end;
tetaf=tetan';
lambda=tetaf[1];
delta=tetaf[2];
print 'ESTIMATIVAS MÉTODO DOS MOMENTOS';
print 'lambda=' lambda ' ' 'delta=' delta;
print 'iterações' i ' ' 'convergencia' (max(abs(tetan-teta0))) ;

/*FREQUENCIA ESPERADA*/
f0=0.08647376551; f1=0.6700790898; f2=0.4057299040/2;
f3=0.2139000989/6; f4=0.1070376201/24; f5=0.05219250515/120;
zero=f0*246; um=f1*246; dois=f2*246; tres=f3*246; quatro=f4*246; cinco=f5*246;
print ' ';
print 'NUMERO DE OVOS ESPERADOS NOS HOSPEDEIROS';
print '0= 'zero; print '1= 'um; print '2= ' dois;
print '3= ' tres; print '4= ' quatro; print '5= ' cinco;

/* VARIÂNCIAS E COVARIÂNCIA DE LAMBDA E DELTA */
u1=lambda*delta+(1-exp(-lambda))*(1-delta);
u2=lambda**2*delta**2+2*delta*(1-delta)*(-1+exp(-lambda)+lambda);
u3=lambda**3*delta**3+6*delta**2*(1-delta)*(1-exp(-lambda)-lambda+
(lambda**2/2));
u4=lambda**4*delta**4+24*delta**3*(1-delta)*(-1+exp(-lambda)+lambda

```

```

-(lambda**2/2)+(lambda**3/6));
k2=u2+u1-u1**2;
k3=u3+3*u2+u1-3*u1*u2-3*u1**2+2*u1**3;
k4=u4+6*u3+7*u2+u1-4*u1*u3-12*u1*u2-4*u1**2+6*u1**3+6*u1**2*u2-3*u1**4;
varx=k2/n;
vars=(k4-((n-3)/(n-1))*k2**2)/n;
cov=k3/n;
dul=delta+exp(-lambda)*(1-delta);
dud=-1+exp(-lambda)+lambda;
du2l=2*lambda*delta**2+2*delta*(1-delta)*(1-exp(-lambda))+delta+
exp(-lambda)*(1-delta)-2*(lambda*delta+(1-exp(-lambda))*
(1-delta))*(delta+exp(-lambda)*(1-delta));
du2d=2*lambda**2*delta+2*(1-delta)*(-1+exp(-lambda)+lambda)-2*delta*
(-1+exp(-lambda)+lambda)+lambda-1+exp(-lambda)-2*(lambda*delta+
(1-exp(-lambda))*(1-delta))*(-1+exp(-lambda)+lambda);
A=(dul**2||dud**2||(2*dul*dud))/(du2l**2||du2d**2||(2*du2l*du2d))/
((dul*du2l)||((dud*du2d)||((dul*du2d)+(dud*du2l)));
varobs=varx//vars//cov;
PAR=inv(A)*varobs;
EPL=sqrt(PAR[1]);
EPD=sqrt(PAR[2]);
COR=PAR[3]/(EPL*EPD);  print ' ';
print ' ERROS-PADRÃO E CORRELAÇÃO DE LAMBDA E DELTA ';
print 'Desvio-Padrão(Lambda)= ' EPL;
print ' Desvio-Padrão(Delta)= ' EPD;
print ' Correlação= ' COR;
quit;

```

## ESTIMATIVAS MÉTODO DOS MOMENTOS

lambda= 2.4479142                      delta= 0.1886575

iterações    4                      convergencia    8.256E-11

## NUMERO DE OVOS ESPERADOS NOS HOSPEDEIROS

0= 21.272546

1= 164.83946

2= 49.904778

3= 8.7699041

4= 1.0971356

5= 0.1069946

## ERROS-PADRÃO E CORRELAÇÃO DE LAMBDA E DELTA

Desvio-Padrão(Lambda)= 0.2357379

Desvio-Padrão(Delta)= 0.0357267

Correlação= -0.739086

**Programa para obter as estimativas  
do Modelo M1 de Bakker- método 4:  
“máxima verossimilhança” usando o IML.**

```

proc iml;
reset noname;

/*chute inicial*/ ;
lambda=2.5;
delta=0.17;
f=-19+169*(-delta*exp(-lambda*delta)+exp(-lambda))/(exp(-lambda*delta)-
.....
lambda**2*(delta-1)**2+1/6*lambda**3*(delta-1)**3+1/24*lambda**4*
(delta-1)**4));
g=169*(-lambda*exp(-lambda*delta)/(1-delta)+(exp(-lambda*delta)-
.....
1/6*lambda**3*(delta-1)**3+1/24*lambda**4*(delta-1)**4));
dff=169*(delta**2*exp(-lambda*delta)-exp(-lambda))/(exp(-lambda*delta)-
.....
(delta-1)**3+1/24*lambda**4*(delta-1)**4)**2);
dfd=169*(-exp(-lambda*delta)+delta*lambda*exp(-lambda*delta))/
.....
+1/6*lambda**3*(delta-1)**3+1/24*lambda**4*(delta-1)**4)**2);
dgl=169*(-exp(-lambda*delta)/(1-delta)+lambda*delta*exp(-lambda*delta)/
.....
(1+lambda*(delta-1)+1/2*lambda**2*(delta-1)**2+1/6*lambda**3*
(delta-1)**3+1/24*lambda**4*(delta-1)**4)**2);
dgd=169*(lambda**2*exp(-lambda*delta)/(1-delta)-2*lambda*exp(-lambda*delta)/
.....
(1+lambda*(delta-1)+1/2*lambda**2*(delta-1)**2+1/6*lambda**3*(delta-1)
**3+1/24*lambda**4*(delta-1)**4)**2);
jacob=(dff||dfd)/(dgl||dgd);
teta0=lambda//delta;

```

```

funcao0=f//g;
tetan=teta0-inv(jacob)*funcao0;
lambda=tetan[1]; delta=tetan[2];

/*inicio do processo iterativo*/
i=0;
do while (max(abs(tetan-teta0)) >0.000001);
f=-19+169*(-delta*exp(-lambda*delta)+exp(-lambda))/(exp(-lambda*delta)-
.....
lambda**2*(delta-1)**2+1/6*lambda**3*(delta-1)**3+1/24*lambda**4*
(delta-1)**4);
g=169*(-lambda*exp(-lambda*delta)/(1-delta)+(exp(-lambda*delta)-
.....
1/6*lambda**3*(delta-1)**3+1/24*lambda**4*(delta-1)**4));
dff=169*(delta**2*exp(-lambda*delta)-exp(-lambda))/(exp(-lambda*delta)-
.....
(delta-1)**3+1/24*lambda**4*(delta-1)**4)**2);
dfd=169*(-exp(-lambda*delta)+delta*lambda*exp(-lambda*delta))/
.....
+1/6*lambda**3*(delta-1)**3+1/24*lambda**4*(delta-1)**4)**2);
dgl=169*(-exp(-lambda*delta)/(1-delta)+lambda*delta*exp(-lambda*delta)/
.....
(1+lambda*(delta-1)+1/2*lambda**2*(delta-1)**2+1/6*lambda**3*
(delta-1)**3+1/24*lambda**4*(delta-1)**4)**2);
dgd=169*(lambda**2*exp(-lambda*delta)/(1-delta)-2*lambda*exp(-lambda*delta)/
.....
(1+lambda*(delta-1)+1/2*lambda**2*(delta-1)**2+1/6*lambda**3*(delta-1)
**3+1/24*lambda**4*(delta-1)**4)**2);
jacob=(dff||dfd)/(dgl||dgd);
teta0=lambda//delta;
funcao0=f//g;
teta0=tetan;

```

```

tetan=teta0-inv(jacob)*funcao0;
lambda=tetan[1]; delta=tetan[2];
i=i+1;
end;
tetaf=tetan';
print 'ESTIMATIVAS MÉTODO DA MAXIMA VEROSSIMILHANÇA';
print 'lambda=' lambda ' ' 'delta=' delta;
print 'iterações'i ' ' 'convergencia' (max(abs(tetan-teta0))) ;

/*FREQUENCIA ESPERADA*/
f0=exp(-lambda);
f1=(exp(-lambda*delta)-exp(-lambda))/(1-delta);
f2=delta*(exp(-lambda)-exp(-lambda*delta)*(1+lambda*(delta-1)))/((delta-1)**2);
f3=delta**2*(exp(-lambda)-exp(-lambda*delta)*(1+lambda*(delta-1)+1/2*lambda**2*
(delta-1)**2))/((delta-1)**3);
f4=delta**3*(exp(-lambda)-exp(-lambda*delta)*(1+lambda*(delta-1)+1/2*lambda**2*
(delta-1)**2+1/6*lambda**3*(delta-1)**3))/((delta-1)**4);
f5=delta**4*(exp(-lambda)-exp(-lambda*delta)*(1+lambda*(delta-1)+1/2*lambda**2*
(delta-1)**2+1/6*lambda**3*(delta-1)**3+1/24*lambda**4*(delta-1)**4))/
((delta-1)**5);
zero=f0*246; um=f1*246; dois=f2*246; tres=f3*246; quatro=f4*246; cinco=f5*246;
print ' ';
print 'NUMERO DE OVOS ESPERADOS NOS HOSPEDEIROS';
print '0= 'zero; print '1= 'um; print '2= ' dois;
print '3= ' tres; print '4= ' quatro; print '5= ' cinco;

/*VARIÂNCIAS E COVARIÂNCIA DE LAMBDA E DELTA*/
q=(1||0)/(0||1);
inf=(-dff||-dfd)/(-dgl||-dgd);
var=inv(q*inf*q);
i11=sqrt(var[1,1]);
i22=sqrt(var[2,2]);
cor=var[1,2]/(i11*i22);

```

```
print ' ';
print ' ERROS-PADRÃO E CORRELAÇÃO DE LAMBDA E DELTA ';
print 'Desvio-Padrão(Lambda)= ' i11;
print ' Desvio-Padrão(Delta)= ' i22;
print ' Correlação=' cor;

/*INTERVALOS DE CONFIANÇA PARA DELTA E LAMBDA*/
lil=lambda-2*i11; lsl=lambda+2*i11;
lid=delta-2*i22; lsd=delta+2*i22;
print ' ';
print 'INTERVALOS DE CONFIANÇA';
print 'IC(lambda)= ['lil';'lsl']';
print ' IC(delta)= ['lid';'lsd']';
quit;
```



## ESTIMATIVAS MÉTODO DA MÁXIMA VEROSSIMILHANÇA

lambda= 2.5575296                      delta= 0.1715768  
 iterações    3                              convergencia    1.124E-11

## NUMERO DE OVOS ESPERADOS NOS HOSPEDEIROS

0= 19.064003  
 1= 168.4612  
 2= 49.130634  
 3= 8.259186  
 4= 0.9858877  
 5= 0.0916215

## ERROS-PADRÃO E CORRELAÇÃO DE LAMBDA E DELTA

Desvio-Padrão(Lambda)= 0.2199433  
 Desvio-Padrão(Delta)= 0.0277813  
 Correlação= -0.652348

## INTERVALOS DE CONFIANÇA

IC(lambda)= [ 2.117643 ; 2.9974162 ]  
 IC(delta)= [ 0.1160143 ; 0.2271393 ]

**Programa para obter as estimativas  
 do Modelo M1 de Bakker- método 4:  
 “máxima verossimilhança” usando o NLMIXED.**

```
data bakker; input x freq; N=246; datalines; 0 19 1 169 2 50 3 6 4 1
5 1 ;
```

```
/* Impressao dos dados observados */ proc print data=bakker;
```

```
title 'DADOS OBSERVADOS'; run;
```

```
/* Processo iterativo para a estimacao - Modelo Poisson*/
```

```
Proc nlmixed data=bakker maxiter=1000 corr;
```

```
title 'PROCESSO ITERATIVO - MODELO POISSON';
```

```
parms lambda=5; /* Valores iniciais dos parametros */
```

```
bounds lambda>0; /* Espaco parametrico */
```

```
model x ~ poisson(lambda); /* Modelo proposto */
```

```
p=pdf('poisson',x,lambda); /* Estimacao das frequências */
```

```
ep=N*p;
```

```
predict ep out=pred_p; /* Arquivo externo com frequencias
estimadas e intervalos de confianca */
```

```
run;
```

```
/*Impressao das frequencias estimadas e intervalos de confianca*/
```

```
proc print data=pred_p; title 'FREQUENCIAS ESTIMADAS - MODELO
POISSON';
```

```
run;
```

```
/* Processo iterativo para a estimacao - Modelo de Bakker*/
```

```
Proc nlmixed data=bakker maxiter=1000 corr;
```

```
title 'PROCESSO ITERATIVO - MODELO DE BAKKER';
```

```
parms lambda=5 delta=0.1; /* Valores iniciais dos parametros*/
```

```
bounds lambda>0,0<=delta<=1 /* Espaco parametrico */
```

```
/* Calculo do somatorio */ /* */
```

```
a1=1; /* */
```

```

a2=a1+(lambda**(2-1)*(delta-1)**(2-1)/fact(2-1)); /* */
a3=a2+(lambda**(3-1)*(delta-1)**(3-1)/fact(3-1)); /* */
a4=a3+(lambda**(4-1)*(delta-1)**(4-1)/fact(4-1)); /* */
a5=a4+(lambda**(5-1)*(delta-1)**(5-1)/fact(5-1)); /* */
/*Atribuicao do valor correto do somatorio para cada valor de x*/
/*MODELO PROBABILISTICO */
if x=1 then a=a1; /* */
if x=2 then a=a2; /* */
if x=3 then a=a3; /* */
if x=4 then a=a4; /* */
if x=5 then a=a5; /* */
/* Calculo das probabilidades */ /* */
if x=0 then p=exp(-lambda); /* */
else p=(delta**(x-1)*(delta-1)**(-x)*(exp(-lambda)
-exp(-lambda*delta)*a));
/* */
logver=freq*log(p); /* Calculo do log(verossimilhanca) */
model x ~ general(logver); /* Modelo proposto */
ep=N*p; /* Estimacao das frequencias */
predict ep out=pred; /* Arquivo externo com frequencias
estimadas e intervalos de confianca */
run;

/*Impressao das frequencias estimadas e intervalos de confianca*/
proc print data=pred;

title 'FREQUENCIAS ESTIMADAS - MODELO DE BAKKER'; run;

```

DADOS OBSERVADOS

Obs	x	freq	N
1	0	19	246
2	1	169	246
3	2	50	246
4	3	6	246
5	4	1	246
6	5	1	246

PROCESSO ITERATIVO - MODELO POISSON

The NLMIXED Procedure

Specifications

Data Set	WORK.BAKKER
Dependent Variable	x
Distribution for Dependent Variable	Poisson
Optimization Technique	Dual Quasi-Newton
Integration Method	None

Dimensions

Observations Used	6
Observations Not Used	0
Total Observations	6
Parameters	1

## Parameters

lambda    NegLogLike

5    16.3088835

## Iteration History

Iter	Calls	NegLogLike	Diff	MaxGrad	Slope
1	5	11.7521259	4.556758	0.446524	-15
2	7	11.7061299	0.045996	0.013597	-0.17951
3	8	11.7060915	0.000038	0.001096	-0.00008
4	9	11.7060912	2.502E-7	2.489E-6	-5.01E-7

NOTE: ABSGCONV convergence criterion satisfied.

## Fit Statistics

-2 Log Likelihood	23.4
AIC (smaller is better)	25.4
AICC (smaller is better)	26.4
BIC (smaller is better)	25.2

PROCESSO ITERATIVO - MODELO POISSON

## The NLMIXED Procedure

## Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value
lambda	2.5000	0.6455	6	3.87
Pr >  t	Alpha	Lower	Upper	Gradient
0.0082	0.05	0.9205	4.0795	2.489E-6

Correlation Matrix of  
Parameter Estimates

Row	Parameter	lambda
1	lambda	1.0000

## FRECUENCIAS ESTIMADAS - MODELO POISSON

Obs	x	freq	N	Pred	StdErr Pred	DF
-----	---	------	---	------	----------------	----

1	0	19	246	20.1929	13.0345	6
2	1	169	246	50.4822	19.5517	6
3	2	50	246	63.1028	8.1466	6
4	3	6	246	52.5857	6.7888	6
5	4	1	246	32.8661	12.7290	6
6	5	1	246	16.4330	10.6075	6

tValue	Probt	Alpha	Lower	Upper
1.54919	0.17231	0.05	-11.7013	52.0871
2.58199	0.04166	0.05	2.6409	98.3235
7.74595	0.00024	0.05	43.1689	83.0367
7.74598	0.00024	0.05	35.9742	69.1972
2.58199	0.04166	0.05	1.7194	64.0128
1.54919	0.17231	0.05	-9.5225	42.3886

PROCESSO ITERATIVO - MODELO DE BAKKER

The NLMIXED Procedure

Specifications

Data Set WORK.BAKKER Dependent  
Variable x Distribution for Dependent

Variable            General Optimization Technique Dual Quasi-Newton  
 Integration Method                            None

#### Dimensions

Observations Used                            6 Observations Not Used 0  
 Total Observations                         6 Parameters 2

#### Parameters

lambda	delta	NegLogLike
5	0.1	259.657751

#### Iteration History

Iter	Calls	NegLogLike	Diff	MaxGrad	Slope
1	4	256.698329	2.959422	42.85622	-1115.99
2	9	227.267496	29.43083	28.53644	-7.35075
3	11	227.15061	0.116886	34.23464	-0.58147
4	13	226.904505	0.246105	3.126502	-0.15475
5	14	226.902348	0.002157	0.188713	-0.00423
6	16	226.902304	0.000044	0.000109	-0.00008
7	18	226.902304	2.96E-10	4.189E-7	-591E-12



NOTE: GCONV convergence criterion satisfied.

### Fit Statistics

-2 Log Likelihood	453.8	AIC (smaller is better)
457.8 AICC (smaller is better)	461.8	BIC (smaller is better)
457.4		

### PROCESSO ITERATIVO - MODELO DE BAKKER

#### The NLMIXED Procedure

#### Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value
lambda	2.5575	0.2199	6	11.63
delta	0.1716	0.02778	6	6.18

Pr >  t	Alpha	Lower	Upper	Gradient
<.0001	0.05	2.0193	3.0957	6.21E-8
0.0008	0.05	0.1036	0.2396	4.189E-7

#### Correlation Matrix of Parameter Estimates

Row	Parameter	lambda	delta
-----	-----------	--------	-------

1	lambda	1.0000	-0.6523
2	delta	-0.6523	1.0000

## FRECUENCIAS ESTIMADAS - MODELO DE BAKKER

Obs	x	freq	N	Pred
1	0	19	246	19.064
2	1	169	246	168.461
3	2	50	246	49.131
4	3	6	246	8.259
5	4	1	246	0.986
6	5	1	246	0.092

## StdErr

Pred	DF	tValue	Probt	Alpha	Lower	Upper
4.19301	6	4.5466	0.003906	0.05	8.804	29.324
6.74575	6	24.9729	0.000000	0.05	151.955	184.967
4.13980	6	11.8679	0.000022	0.05	39.001	59.260
1.66781	6	4.9521	0.002572	0.05	4.178	12.340
0.31799	6	3.1004	0.021107	0.05	0.208	1.764
0.04071	6	2.2506	0.065384	0.05	-0.008	0.191

## ANEXO B

A estimação da seqüência das taxas de transição é um passo inicial muito importante de uma análise preliminar de um processo de Poisson estendido. Podlich (1999) desenvolveu a rotina “empir.lamb” para tal fim usando o SPlus. O comando `empir.lamb(f)`, em que  $f$  é o vetor das freqüências relativas observadas, retorna um vetor das taxas de nascimentos cujo comprimento é dado pelo número de elementos não nulos de  $f$ . Ele fornece ainda o delta-grama que é um valioso instrumento gráfico para a escolha do modelo apropriado.

### Modelo de Bakker - Experimento 1

```
> #####\par
> #####  Ovos por hospedeiro - Bakker (1972), Exp. 1\par
> #####\par
> ## Teste 1\par
> x<-0:3\par
> fx<-c(24,30,5,1)\par
> fx<-fx/sum(fx)\par
> ovos1<-list(x=x,freq.x=fx)\par
> out.bak1<-empir.lamb(ovos1$freq.x)\par
\par
  Solving for the birth rates..... \par
> empir.lamb(ovos1$freq.x)\par
\par
  Solving for the birth rates..... \par
$n:\par [1] 0 1 2 3\par
\par
$lambda.n:\par [1] 0.9162907 0.3242803 0.5164798 0.0000000\par
\par
> write(out.bak1$n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\\')
```

```

y_bak1.txt')\par
> write(out.bak1$lambda.n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\\
l_bak1.txt')\par
> #####\par

```

## Modelo de Bakker - Experimento 2

```

> #####\par
> #####  Ovos por hospedeiro - Bakker (1972), Exp. 2\par
> #####\par
> ## Teste 1\par
> x<-0:3\par
> fx<-c(8,30,9,2)\par
> fx<-fx/sum(fx)\par
> ovos2<-list(x=x,freq.x=fx)\par
> out.bak2<-empir.lamb(ovos2$freq.x)\par
\par
Solving for the birth rates..... \par
> empir.lamb(ovos2$freq.x)\par
\par
Solving for the birth rates..... \par
$n:\par [1] 0 1 2 3\par
\par
$lambda.n:\par [1] 1.8123788 0.4997261 0.5271197 0.0000000\par
\par
> write(out.bak2$n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\\
y_bak2.txt')\par
> write(out.bak2$lambda.n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\\
l_bak2.txt')\par
> #####\par

```

### Modelo de Bakker - Experimento 3

```

> #####\par
> #####  Ovos por hospedeiro - Bakker (1972), Exp. 3\par
> #####\par
> ## Teste 1\par
> x<-0:3\par
> fx<-c(3,36,41,6)\par
> fx<-fx/sum(fx)\par
> ovos3<-list(x=x,freq.x=fx)\par
> out.bak3<-empir.lamb(ovos3$freq.x)\par
\par
  Solving for the birth rates..... \par
> empir.lamb(ovos3$freq.x)\par
\par
  Solving for the birth rates..... \par
$n:\par [1] 0 1 2 3\par
\par
$lambda.n:\par [1] 3.3557350 1.1854412 0.3034648 0.0000000\par
\par
> write(out.bak3$n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
  y_bak3.txt')\par
> write(out.bak3$lambda.n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
  l_bak3.txt')\par
> #####\par

```

### Modelo de Bakker - Experimento 4

```

#####\par
> #####  Ovos por hospedeiro - Bakker (1972), Exp. 4\par
> #####\par
> ## Teste 1\par

```

```

> x<-0:4\par
> fx<-c(1,23,40,25,2)\par
> fx<-fx/sum(fx)\par
> ovos4<-list(x=x,freq.x=fx)\par
> out.bak4<-empir.lamb(ovos4$freq.x)\par
\par
Solving for the birth rates..... \par
> empir.lamb(ovos4$freq.x)\par
\par
Solving for the birth rates..... \par
$n:\par [1] 0 1 2 3 4\par
\par
$lambda.n:\par [1] 4.5108595 1.8225257 1.0827848 0.2259753
0.0000000\par
\par
> write(out.bak4$n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
y_bak4.txt')\par
> write(out.bak4$lambda.n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
l_bak4.txt')\par
> #####\par

```

## Modelo de Bakker - Experimento 5

```

#####\par
> ##### Ovos por hospedeiro - Bakker (1972), Exp. 5\par
> #####\par
> ## Teste 1\par
> x<-0:4\par
> fx<-c(1,12,42,28,5)\par
> fx<-fx/sum(fx)\par
> ovos5<-list(x=x,freq.x=fx)\par

```

```

> out.bak5<-empir.lamb(ovos5$freq.x)\par
\par
Warning: probability vector 'p' does not sum to 1 \par
Appending 1-sum(p) to make 'p' a proper probability distribution
\par
\par
Solving for the birth rates..... \par
> empir.lamb(ovos5$freq.x)\par
\par
Warning: probability vector 'p' does not sum to 1 \par
Appending 1-sum(p) to make 'p' a proper probability distribution
\par
\par
Solving for the birth rates..... \par
$n:\par [1] 0 1 2 3 4 5\par
\par
$lambda.n:\par [1] 4.4773368 2.7486626 1.1259367 0.4682627 0.0010000
0.0000000\par
\par
> write(out.bak5$n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
y_bak5.txt')\par
> write(out.bak5$lambda.n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
l_bak5.txt')\par
> #####\par

```

## Modelo de Bakker - Experimento 6

```

#####\par
> ##### Ovos por hospedeiro - Bakker (1972), Exp. 6\par
> #####\par
> ## Teste 1\par

```

```

> x<-0:5\par
> fx<-c(1,10,51,40,6,1)\par
> fx<-fx/sum(fx)\par
> ovos6<-list(x=x,freq.x=fx)\par
> out.bak6<-empir.lamb(ovos6$freq.x)\par
\par
Solving for the birth rates..... \par
> empir.lamb(ovos6$freq.x)\par
\par
Solving for the birth rates..... \par
$n:\par [1] 0 1 2 3 4 5\par
\par
$lambda.n:\par [1] 4.6913479 3.3298096 1.2091291 0.4424755 0.6064630
0.0000000\par
\par
> write(out.bak6$n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
y_bak6.txt')\par
> write(out.bak6$lambda.n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
l_bak6.txt')\par
> #####\par

```

### Modelo de Bakker - Datos de Spalangia

```

#####\par
> ### 7. Exemplo 1: Dist. Bakker\par
> ### Spalangia\par
> #####\par
> ## Teste 1\par
> x1<-0:5\par
> fx1<-c(19,169,50,6,1,1)\par
> fx1<-fx1/sum(fx1)\par

```



```

> spal<-list(x=x1,freq.x=fx1)\par
> out.bak7<-empir.lamb(spal$freq.x)\par
\par
Solving for the birth rates..... \par
> empir.lamb(spal$freq.x)\par
\par
Solving for the birth rates..... \par
$n:\par [1] 0 1 2 3 4 5\par
\par
$lambda.n:\par [1] 2.5608926 0.4342912 0.3722282 1.0769210 3.8532547
0.0000000\par
\par
> write(out.bak7$n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
y_spal.txt')\par
> write(out.bak7$lambda.n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
l_spal.txt')\par

```

### Modelo de Janardan - Feijão - Ensaio 1

```

> #####\par
> ### Exemplo 3: Dist. Janardan\par
> ### Feijao - Ensaio 1 (138,46)\par
> #####\par
> ## Teste 1\par
> x1<-0:1\par
> fx1<-c(138,46)\par
> fx1<-fx1/sum(fx1)\par
> jana1<-list(x=x1,freq.x=fx1)\par
> out.jan1<-empir.lamb(jana1$freq.x)\par
\par
Solving for the birth rates..... \par

```

```

> empir.lamb(jana1$freq.x)\par
\par
Solving for the birth rates..... \par
$n:\par [1] 0 1 0 1\par
\par
$lambda.n:\par [1] 0.2876821 0.0010000 0.0010000 0.0000000\par
\par
> write(out.jan1$n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
y_jan1.txt')\par
> write(out.jan1$lambda.n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
l_jan1.txt')\par
>\par

```

## Modelo de Janardan - Feijão - Ensaio 2

```

> #####\par
> ### 8. Exemplo 1: Dist. Janardan\par
> ### Feijao - Ensaio 2\par
> #####\par
> ## Teste 1\par
> x1<-0:3\par
> fx1<-c(26,117,35,1)\par
> fx1<-fx1/sum(fx1)\par
> jana2<-list(x=x1,freq.x=fx1)\par
> out.jan2<-empir.lamb(jana2$freq.x)\par
\par
Solving for the birth rates..... \par
> empir.lamb(jana2$freq.x)\par
\par
Solving for the birth rates..... \par
$n:\par [1] 0 1 2 3\par
\par

```

```

$lambda.n:\par [1] 1.92928927 0.42144127 0.07145438 0.00000000\par
\par
> write(out.jan2$n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
  y_jan2.txt')\par
> write(out.jan2$lambda.n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
  l_jan2.txt')\par

```

### Modelo de Janardan - Feijão - Ensaio 3

```

> #####\par
> ### 9. Exemplo 2: Dist. Janardan\par
> ###   Feijao - Ensaio 3 (5,68,88,32)\par
> #####\par
> ## Teste 1\par
> x1<-0:3\par
> fx1<-c(5,68,88,32)\par
> fx1<-fx1/sum(fx1)\par
> jana3<-list(x=x1,freq.x=fx1)\par
> out.jan3<-empir.lamb(jana3$freq.x)\par
\par
  Solving for the birth rates..... \par
> empir.lamb(jana3$freq.x)\par
\par
  Solving for the birth rates..... \par
$n:\par [1] 0 1 2 3\par
\par
$lambda.n:\par [1] 3.6532523 1.4229378 0.6825043 0.0000000\par
\par
> write(out.jan3$n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
  y_jan3.txt')\par
> write(out.jan3$lambda.n,file='c:\\\\mauricio\\\\tese\\\\graficos\\\\
  l_jan3.txt')\par

```

# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)