

**EDLEIDE DE BRITO**

**DISTRIBUIÇÕES EM SÉRIE DE POTÊNCIAS  
MODIFICADAS: ABORDAGEM CLÁSSICA E BAYESIANA**

RECIFE-PE - DEZ/2008



**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO**  
**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA**

## **DISTRIBUIÇÕES EM SÉRIE DE POTÊNCIAS MODIFICADAS: ABORDAGEM CLÁSSICA E BAYESIANA**

Dissertação apresentada ao Programa de Pós-Graduação em Biometria e Estatística Aplicada como exigência parcial à obtenção do título de Mestre.

**Área de Concentração: Modelagem Estatística e Computacional**

Orientador: Prof. Dr. Eufrázio de Souza Santos  
Co-orientador: Prof. Dr. Gauss Moutinho Cordeiro

RECIFE-PE - DEZ/2008

## FICHA CATALOGRÁFICA

B862d Brito, Edleide de  
Distribuições em série de potências modificadas : abordagem clássica e bayesiana / Edleide de Brito. -- 2008.  
70f. : il.

Orientador : Eufrázio de Souza Santos  
Dissertação (Mestrado em Biometria e Estatística Aplicada) -- Universidade Federal Rural de Pernambuco. Departamento de Estatística e Informática.  
Inclui apêndice e bibliografia.

CDD 574. 018 2

1. Distribuição discreta
  2. Função de probabilidade
  3. Função geradora de momentos
  4. Inferência bayesiana
  5. Função de verossimilhança
  6. Momentos
- I. Santos, Eufrázio de Souza
  - II. Título


**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO**  
**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA**

**DISTRIBUIÇÕES EM SÉRIE DE POTÊNCIAS MODIFICADAS: ABORDAGENS**  
**CLÁSSICA E BAYESIANA**


EDLEIDE DE BRITO

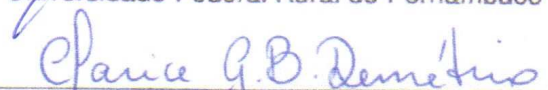
Dissertação julgada adequada para obtenção do título de mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 03/12/2008 pela Comissão Examinadora.

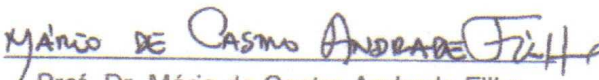
Orientador:

  
Prof. Dr. Eufrazio de Souza Santos  
Universidade Federal Rural de Pernambuco

Banca Examinadora:

  
Prof. Dr. Gauss Moutinho Cordeiro  
Universidade Federal Rural de Pernambuco

  
Prof. Dra. Clarice Garcia Borges Demétrio  
Universidade de São Paulo - ESALQ

  
Prof. Dr. Mário de Castro Andrade Filho  
Universidade de São Paulo - ICMC

## Dedicatória

*À minha mamãe Arcanja, aos meus irmãos  
Elisângela e Jorge, ao meu sobrinho Luis  
Philippe e ao meu amado Angelo.*

## Agradecimentos

*Agradeço primeiramente a Deus, por me fazer crer no Seu amor e na Sua justiça e por ter me guiado e encorajado na busca de mais um sonho.*

*A minha mamãe Arcanja que sempre nos dispensou carinho e atenção mesmo com todas as dificuldades que passou para manter nossa família. Por ter me apoiado, incondicionalmente, desde o princípio desta jornada e pelas orações e incentivos constantes. Obrigada Mamy.*

*Aos meus irmãos Elisângela e Jorge pelo carinho, apoio e incentivo principalmente nos momentos mais difíceis. Obrigada Manos.*

*Ao meu querido Angelo, companheiro de vida, pelo amor, carinho, paciência e incentivo, por entender minhas prioridades, me apoiar e ter ficado ao meu lado em todos os momentos, felizes ou não. Obrigada meu Amor.*

*À Maria Pereira de Jesus, minha tia-avó, lembrança sempre presente de alegrias e saudades eternas.(In memoriam)*

*Ao Professor Gauss Cordeiro pela sabedoria, pelo exemplo, pela dedicação, pelo apoio fundamental para a elaboração deste trabalho e pelo incentivo constante. Meu eterno carinho e gratidão.*

*Ao Professor Eufrázio de Souza Santos pelo apoio, contribuição e por ter confiado em meu trabalho.*

*A Professora Clarice Demétrio pelas valiosas contribuições para a elaboração e finalização deste trabalho e pelos ensinamentos transmitidos durante minha temporada na ESALQ/USP.*

*Ao Professor Mário de Castro pelas valiosas contribuições para a finalização deste trabalho, bem como, as dúvidas esclarecidas e pelo apoio, principalmente, na parte computacional.*

*A Professora Roseli Aparecida pelos ensinamentos transmitidos durante minha temporada na ESALQ/USP.*

*Ao pessoal do Núcleo de Nutrição e Epidemiologia da UFBA pelo apoio e incentivo,*

*desde o primeiro dia em que cheguei a esta família. Principalmente a Professora Ana Marlúcia Assis, a minha amiga Sandra Pinheiro e as nutricionistas Lucivalda, Maria da Conceição, Mônica, Nedja e Valterlinda.*

*A amiga Andrea Prudente pelo apoio em um momento tenso desta jornada, pelo carinho, compreensão e pelas horas de risos. Uma amiga para sempre ser lembrada. Obrigada por tudo Dea.*

*A amiga Silvia Lemos pelo apoio na primeira parte dessa jornada e pelas horas de descontração.*

*A Rejane Brito pelo apoio na segunda parte dessa jornada e pelas horas de alegrias que tivemos.*

*A amiga Alexandra Esteban pelo carinho, compreensão e paciência. Nossa convivência foi curta, mas suficiente para me fazer querer ser sua amiga para sempre.*

*A amiga Amanda Lira que conquistou minha amizade pela sua maneira carinhosa e alegre de tratar a todos.*

*A todos os colegas do mestrado, nomeadamente a Adilton, Dênis, Felipe, Juliana, Lenaldo, Luciano, Luiz, Vanessa, Katia e Vinicius.*

*A Zuleide pelas palavras de carinho e compreensão nos momentos difíceis dessa jornada.*

*A todos os meus professores que contribuíram com a minha formação.*

*Ao CNPq pela bolsa de estudos.*

*A todos que, de forma direta ou indireta, contribuíram para que este trabalho fosse possível.*

*“O que nós fazemos nunca é compreendido,  
apenas louvado ou condenado.”*

*Friedrich Nietzsche*



## Resumo

Neste trabalho, alguns resultados na classe de distribuições em série de potências modificadas (representada pela sigla MPSD) proposta por Gupta (1974) são apresentados. Importantes distribuições tais como a binomial, binomial negativa, Poisson e séries logarítmicas pertencem a esta classe de distribuições. Mais recentemente essa classe de distribuições foi estudada no contexto de modelos não-lineares generalizados por Cordeiro et al. (2008). Dentre estas distribuições, devido sua aplicabilidade em problemas práticos propusemos uma abordagem bayesiana para o modelo binomial negativo generalizado com parâmetros  $\mu$ ,  $\phi$  e  $m$ , em que através do método de Monte Carlo com Cadeias de Markov (MCMC) encontramos as estimativas bayesianas, considerando  $m$  conhecido, para os parâmetros  $\mu$  e  $\phi$  utilizando um conjunto de dados reais.

**Palavras-chave:** Distribuição discreta, Função de probabilidade, Função geradora de momentos, Inferência bayesiana, Função de verossimilhança, Momentos.

## Abstract

We present some mathematical properties of the class of modified power series distributions (MPSD) introduced by Gupta (1974). This class of distribution includes some important distributions such as the binomial, negative binomial, Poisson, logarithmic series, among others. More recently, this class of models was studied in the context of regression models by Cordeiro et al. (2008). It presents some methods for determination of *priors* and using real data sets some was accomplished analysis bayesian.

# Lista de Figuras

3.1	Gráficos da distribuição binomial . . . . .	25
3.2	Gráficos da distribuição binomial negativa . . . . .	27
3.3	Gráficos da distribuição binomial negativa generalizada . . . . .	28
3.4	Gráficos da distribuição de Borel . . . . .	30
3.5	Gráficos da distribuição de Borel-Tanner . . . . .	31
3.6	Gráficos da distribuição de Consul . . . . .	33
3.7	Gráficos da distribuição delta-binomial . . . . .	34
3.8	Gráficos da distribuição de Geeta . . . . .	35
3.9	Gráficos da distribuição de Haight . . . . .	36
3.10	Gráficos da distribuição “lost game” . . . . .	38
3.11	Gráficos da distribuição de Poisson . . . . .	39
3.12	Gráficos da distribuição de Poisson generalizada . . . . .	41
3.13	Função de probabilidade da distribuição em série logarítmica . . . . .	42
3.14	Função de probabilidade da distribuição em série logarítmica generalizada . . . . .	43
7.1	Gráficos dos parâmetros da binomial negativa generalizada . . . . .	62

# Lista de Tabelas

4.1	Casos especiais de distribuições em série de potências modificadas inflacionadas em zero . . . . .	47
4.2	Casos especiais de distribuições em série de potências modificadas truncadas em zero . . . . .	48
6.1	Desvios unitários para alguns modelos em série de potências modificadas .	59
7.1	Mortalidade de besouros segundo as doses de três inseticidas . . . . .	60
7.2	Estatísticas para verificar a qualidade do ajuste dos dados dos insetos . . .	61
7.3	Estimativas dos parâmetros do modelo binomial negativo generalizado utilizando os dados dos insetos . . . . .	62
7.4	Distribuição de duas espécies de lagartos segundo quatro fatores . . . . .	63
7.5	Estatísticas para verificar a qualidade do ajuste na abordagem clássica dos dados dos lagartos . . . . .	63
7.6	Estatísticas para verificar a qualidade do ajuste na abordagem bayesiana dos dados dos lagartos . . . . .	64
A.1	Assimetria e curtose de algumas distribuições na classe MPSD . . . . .	70

# Sumário

<b>1</b>	<b>Introdução</b>	<b>14</b>
<b>2</b>	<b>Propriedades da Classe de Distribuições em Série de Potências Modificadas</b>	<b>15</b>
2.1	Função de Probabilidade . . . . .	15
2.2	Momentos e Cumulantes . . . . .	16
2.2.1	Função Geradora de Momentos . . . . .	16
2.2.2	Média . . . . .	17
2.2.3	Fórmulas de Recorrência para os Momentos . . . . .	17
2.2.4	Relação de Recorrência entre Cumulantes e Momentos . . . . .	19
2.2.5	Momentos Negativos e Momentos Incompletos . . . . .	20
2.3	Estimação . . . . .	21
2.3.1	Estimador de Máxima Verossimilhança . . . . .	21
2.3.2	Estimação Intervalar . . . . .	22
<b>3</b>	<b>Alguns Casos Especiais da Classe MPSD</b>	<b>24</b>
3.1	Distribuições Parametrizadas na Média $\mu$ . . . . .	24
3.1.1	Distribuição Binomial . . . . .	24
3.1.2	Distribuição Binomial Negativa . . . . .	26
3.1.3	Distribuição Binomial Negativa Generalizada . . . . .	26
3.1.4	Distribuição de Borel . . . . .	29
3.1.5	Distribuição de Borel-Tanner . . . . .	29
3.1.6	Distribuição de Consul . . . . .	31
3.1.7	Distribuição Delta-binomial . . . . .	32

<i>Sumário</i>	xii
3.1.8 Distribuição de Geeta . . . . .	33
3.1.9 Distribuição de Haight . . . . .	35
3.1.10 Distribuição “Lost Game” . . . . .	37
3.1.11 Distribuição de Poisson . . . . .	37
3.1.12 Distribuição de Poisson Generalizada . . . . .	40
3.2 Distribuições em que $\mu = \tau(\theta)$ . . . . .	40
3.2.1 Distribuição em Série Logarítmica . . . . .	40
3.2.2 Distribuição Série Logarítmica Generalizada . . . . .	42
<b>4 Distribuições em Série de Potências Modificadas Inflacionadas e Truncadas</b>	<b>44</b>
4.1 Distribuições em Série de Potências Modificadas Inflacionadas . . . . .	44
4.1.1 Função de Probabilidade . . . . .	44
4.1.2 Momentos . . . . .	45
4.1.3 Estimação de Máxima Verossimilhança . . . . .	46
4.2 Distribuições em Série de Potências Modificadas Truncadas . . . . .	47
4.2.1 Função de Probabilidade . . . . .	47
4.2.2 Momentos . . . . .	47
<b>5 Inferência Bayesiana</b>	<b>49</b>
5.1 Teorema de Bayes . . . . .	49
5.2 Distribuições a Priori . . . . .	51
5.2.1 Distribuições a Priori Conjugadas . . . . .	51
5.2.2 Distribuições a Priori não Informativas . . . . .	51
5.3 Distribuição a Posteriori para o Modelo Binomial Negativo Generalizado . . . . .	54
5.3.1 Distribuição a Priori para os Parâmetros . . . . .	54
<b>6 Modelos Não-Lineares Generalizados em Série de Potências Modificadas</b>	<b>56</b>
6.1 Modelos Lineares Generalizados . . . . .	56

<i>Sumário</i>	xiii
6.2 Ajuste do Modelo . . . . .	57
6.3 Inferência . . . . .	58
<b>7 Aplicações</b>	<b>60</b>
7.1 Exemplo 1 . . . . .	60
7.1.1 Abordagem Clássica . . . . .	60
7.1.2 Abordagem Bayesiana . . . . .	61
7.2 Exemplo 2 . . . . .	62
7.2.1 Abordagem Clássica . . . . .	62
7.2.2 Abordagem Bayesiana . . . . .	63
<b>8 Conclusão</b>	<b>65</b>
<b>Referências Bibliográficas</b>	<b>66</b>
<b>Apêndice A – Assimetria e Curtose</b>	<b>69</b>

# 1 Introdução

A classe de distribuições em série de potências modificadas (representada pela sigla MPSD) proposta por Gupta (1974) e algumas de suas propriedades serão revisadas neste trabalho.

A classe de distribuições MPSD inclui várias distribuições importantes, tais como, binomial, binomial negativa, Poisson, série logarítmica, binomial negativa generalizada estudada por Jain e Consul (1971), Poisson generalizada de Consul e Jain (1973a), série logarítmica generalizada de Jain e Gupta (1973), Poisson Lagrangiana de Consul e Jain (1973b), binomial Lagrangiana de Jain e Consul (1971), Borel-Tanner e “lost game” de Kemp e Kemp (1968) e suas formas truncadas.

No Capítulo 2, apresentam-se propriedades gerais da classe MPSD, tais como, momentos, fórmulas recursivas para os momentos, momentos centrais e momentos fatoriais, estimador de máxima verossimilhança (EMV), entre outras. No Capítulo 3, apresentam-se algumas das principais distribuições que pertencem à classe de distribuição em série de potências modificada, bem como algumas das suas principais propriedades e aplicações. No Capítulo 4, algumas formas truncadas e inflacionadas das distribuições ilustradas no Capítulo 3 são apresentadas. Em seguida, no Capítulo 5, apresenta-se uma breve introdução à inferência bayesiana e usando-se o método de distribuição *a priori* conjugada, determinou-se a distribuição *a posteriori* para o modelo binomial negativo generalizado. As distribuições em série de potências modificadas foram estudadas, recentemente, no contexto de modelos não-lineares generalizados por Cordeiro et al. (2008) e que será apresentado no Capítulo 6. Por fim, no Capítulo 7, apresentam-se algumas aplicações da classe de distribuições em série de potências modificadas no contexto de modelos não-lineares generalizados e realiza-se uma comparação com a inferência bayesiana.



## 2 Propriedades da Classe de Distribuições em Série de Potências Modificadas

Neste capítulo algumas propriedades da classe de distribuições em série de potências modificadas são apresentadas.

### 2.1 Função de Probabilidade

Uma variável aleatória discreta  $Y$  é dita ter uma distribuição em série de potências modificada se a sua função de probabilidade for

$$P(Y = y) = \frac{a(y)g(\theta)^y}{f(\theta)}, \quad y \in T, \quad (2.1)$$

em que  $T$  é um subconjunto do conjunto dos inteiros não-negativos,  $a(y) > 0$ ,  $g(\theta)$  e

$f(\theta) = \sum_y a(y)g(\theta)^y$  são funções positivas, finitas e diferenciáveis. Essa classe de distribuições é uma extensão da da classe de distribuições em série de potências generalizada (GPSD), que nesse caso, tem  $g(\theta) = \theta$  em (2.1).

## 2.2 Momentos e Cumulantes

### 2.2.1 Função Geradora de Momentos

A função geradora de momentos (fgm) de uma variável aleatória  $Y$  pertencente à classe MPSD, se existe, é expressa como

$$M_Y(t) = E(e^{tY}) = \sum_y e^{ty} P(Y = y) = \sum_y e^{ty} \frac{a(y)g(\theta)^y}{f(\theta)}.$$

Logo,

$$\begin{aligned} M_Y(t) &= \sum_y (e^t - 1 + 1)^y \frac{a(y)g(\theta)^y}{f(\theta)} = \sum_y \sum_{i=0}^y \binom{y}{i} \frac{a(y)g(\theta)^y (e^t - 1)^{y-i}}{f(\theta)} \\ &= \sum_{i=0}^{\infty} \sum_y \binom{y+i}{i} y! \frac{a(y+i)g(\theta)^{y+i} (e^t - 1)^y}{f(\theta)y!}. \end{aligned}$$

Por expansão de  $(e^t - 1)^y$  como série de potências em  $t$ , Gupta e Singh (1982) obtiveram a fgm para a classe MPSD como

$$M_Y(t) = \sum_{s=0}^{\infty} \frac{t^s}{s!} \sum_y \sum_{i=0}^{\infty} \binom{y+i}{i} y! \frac{a(y+i)g(\theta)^{y+i} S(s, y)}{f(\theta)}, \quad y \in T,$$

em que  $T$  é um subconjunto do conjunto dos inteiros não-negativos e  $S(s, x)$  denota o número de Stirling de segunda ordem, definido por

$$S(s, x) = \frac{1}{x!} \sum_{r=0}^x (-1)^{x-r} \binom{x}{r} r^s,$$

sendo  $x = 0, 1, \dots, s$ .

Quando  $g(\theta)$  é inversível e  $\theta$  pode ser expresso como  $\theta = \psi[g(\theta)]$ , a função geradora de momentos (fgm)(JAIN; KHAN, 1979) pode ser expressa como

$$M_Y(t) = \frac{f\{\psi[g(\theta)e^t]\}}{f\{\psi[g(\theta)]\}}.$$

### 2.2.2 Média

Tem-se que

$$\sum_y a(y)g(\theta)^y = f(\theta).$$

Diferenciando em relação a  $\theta$ , vem

$$\frac{g'(\theta)}{g(\theta)} \sum_y y a(y)g(\theta)^y = f'(\theta)$$

e

$$\frac{g'(\theta)}{g(\theta)} f(\theta) \sum_y y \frac{a(y)g(\theta)^y}{f(\theta)} = f'(\theta).$$

Então, a média da classe MPSD é expressa por

$$E(Y) = \mu(\theta) = \mu'_1 = \frac{g(\theta) f'(\theta)}{g'(\theta) f(\theta)}.$$

### 2.2.3 Fórmulas de Recorrência para os Momentos

#### Momentos Ordinários

Tem-se que

$$E(Y^r) = \mu'_r = \sum_y y^r \frac{a(y)g(\theta)^y}{f(\theta)}.$$

Derivando em relação a  $\theta$  vem

$$\frac{d\mu'_r}{d\theta} = \sum_y y^r a(y) y g(\theta)^{y-1} \frac{g'(\theta)}{f(\theta)} - \sum_y y^r a(y) g(\theta)^y \frac{f'(\theta)}{f(\theta)^2}.$$

Multiplicando essa expressão por  $g(\theta)$  obtemos

$$g(\theta) \frac{d\mu'_r}{d\theta} = g'(\theta) \mu'_{r+1} - \frac{g(\theta) f'(\theta)}{f(\theta)} \mu'_r.$$

Logo, os momentos ordinários satisfazem a seguinte relação de recorrência:

$$\mu'_{r+1} = \frac{g(\theta)}{g'(\theta)} \frac{d\mu'_r}{d\theta} + \mu'_1 \mu'_r, \quad r = 0, 1, 2, \dots,$$

sendo  $\mu'_0 = 1$ .

A variância da classe MPSD é expressa como

$$\text{Var}(Y) = E(Y^2) - [E(Y)]^2 = \mu'_2 - \mu_1'^2 = \frac{g(\theta)}{g'(\theta)} \frac{d\mu'_1}{d\theta}. \quad (2.2)$$

### Momentos Centrais

Tem-se que

$$E[(Y - \mu'_1)^r] = \mu_r = \sum_y (y - \mu'_1)^r \frac{a(y)g(\theta)^y}{f(\theta)}.$$

Derivando em relação a  $\theta$  vem

$$\begin{aligned} \frac{d\mu_r}{d\theta} &= -\frac{1}{f(\theta)} \sum_y r(y - \mu'_1)^{r-1} \frac{d\mu'_1}{d\theta} a(y)g(\theta)^y \\ &\quad + \frac{1}{f(\theta)} \sum_y (y - \mu'_1)^r a(y)yg(\theta)^{y-1} g'(\theta) \\ &\quad - \frac{f'(\theta)}{f(\theta)^2} \sum_y (y - \mu'_1)^r a(y)g(\theta)^y. \end{aligned}$$

Logo,

$$g(\theta) \frac{d\mu_r}{d\theta} = \frac{g'(\theta)}{f(\theta)} \sum_y \left\{ (y - \mu'_1)^r a(y)g(\theta)^y \left[ y - \frac{f'(\theta)}{f(\theta)} \frac{g(\theta)}{g'(\theta)} \right] \right\} - rg(\theta) \frac{d\mu'_1}{d\theta} \mu_{r-1}$$

e

$$\mu_{r+1} = \frac{g(\theta)}{g'(\theta)} \frac{d\mu_r}{d\theta} + r \frac{g(\theta)}{g'(\theta)} \frac{d\mu'_1}{d\theta} \mu_{r-1}.$$

Mas,

$$\text{Var}(Y) = \mu_2 = E[(Y - \mu'_1)^2] = \frac{g(\theta)}{g'(\theta)} \frac{d\mu'_1}{d\theta},$$

portanto, os momentos centrais satisfazem

$$\mu_{r+1} = \frac{g(\theta)}{g'(\theta)} \frac{d\mu_r}{d\theta} + r \mu_2 \mu_{r-1}, \quad r = 1, 2, 3, \dots,$$

sendo  $\mu_0 = 1$  e  $\mu_1 = 0$ .

### Momentos Fatoriais

O  $r$ -ésimo momento fatorial de uma variável aleatória discreta  $Y$  é expresso como

$$E(Y^{[r]}) = E[Y(Y-1)\dots(Y-r+1)] = \sum_y y(y-1)\dots(y-r+1)P(Y=y).$$

Para  $Y$  na classe MPSD, tem-se que

$$E(Y^{[r]}) = \mu'_{[r]} = \sum_y y^{[r]} \frac{a(y)g(\theta)^y}{f(\theta)}.$$

Derivando em relação a  $\theta$ , tem-se

$$\frac{d\mu'_{[r]}}{d\theta} = \frac{1}{f(\theta)} \sum_y y^{[r]} a(y) y g(\theta)^{y-1} g'(\theta) - \frac{f'(\theta)}{f(\theta)^2} \sum_y y^{[r]} a(y) g(\theta)^y$$

ou

$$= \frac{g'(\theta)}{g(\theta)} \sum_y y^{[r]} y \frac{a(y)g(\theta)^y}{f(\theta)} - \frac{f'(\theta)}{f(\theta)} \sum_y \frac{y^{[r]} a(y) g(\theta)^y}{f(\theta)}$$

ou

$$= \frac{g'(\theta)}{g(\theta)} \sum_y y^{[r+1]} \frac{a(y)g(\theta)^y}{f(\theta)} + r \frac{g'(\theta)}{g(\theta)} \sum_y y^{[r]} \frac{a(y)g(\theta)^y}{f(\theta)} - \frac{f'(\theta)}{f(\theta)} \mu'_{[r]}.$$

Logo,

$$\frac{d\mu'_{[r]}}{d\theta} = \frac{g'(\theta)}{g(\theta)} \mu'_{[r+1]} + r \frac{g'(\theta)}{g(\theta)} \mu'_{[r]} - \frac{f'(\theta)}{f(\theta)} \mu'_{[r]}.$$

Finalmente,

$$\mu'_{[r+1]} = \frac{g(\theta)}{g'(\theta)} \frac{d\mu'_{[r]}}{d\theta} - r \mu'_{[r]} + \frac{f'(\theta)g(\theta)}{f(\theta)g'(\theta)} \mu'_{[r]}.$$

Portanto, os momentos fatoriais têm a seguinte fórmula de recorrência

$$\mu'_{[r+1]} = \frac{g(\theta)}{g'(\theta)} \frac{d\mu'_{[r]}}{d\theta} + \mu'_{[r]}(\mu'_{[1]} - r), \quad r = 1, 2, 3, \dots,$$

sendo  $\mu'_{[1]} = \mu'_1 = \mu$ .

## 2.2.4 Relação de Recorrência entre Cumulantes e Momentos

A relação entre os momentos ordinários e os cumulantes  $\kappa_j$ ,  $j = 1, 2, 3, \dots$ , para qualquer distribuição de probabilidade discreta foi obtida por Noack (1950) como

$$\mu'_r = \sum_{j=1}^r \binom{r-1}{j-1} \mu'_{r-j} \kappa_j, \quad r = 1, 2, 3, \dots \quad (2.3)$$

Gupta (1974), baseado em (2.3), obteve a relação de recorrência entre os cumulantes e os momentos ordinários da classe MPSD

$$\kappa_{r+1} = \frac{g(\theta)}{g'(\theta)} \sum_{j=1}^r \binom{r-1}{j-1} \mu'_{r-j} \frac{d\kappa_j}{d\theta} - \sum_{j=2}^r \binom{r-1}{j-2} \mu'_{r+1-j} \kappa_j, \quad r = 1, 2, 3, \dots$$

Jani (1978) mostrou que uma distribuição discreta é uma MPSD, se e somente se, os cumulantes satisfazem a relação de recorrência (CONSUL; FAMOYE, 2006)

$$\kappa_{r+1} = \frac{g(\theta)}{g'(\theta)} \frac{d\kappa_r}{d\theta}, r = 1, 2, 3, \dots$$

### 2.2.5 Momentos Negativos e Momentos Incompletos

Seja  $Y$  uma variável aleatória com distribuição na classe MPSD,  $g(0) = 0$  e  $k$  um número não-negativo, tal que,  $k + y \neq 0$  para  $y \in T$ . Define-se o  $r$ -ésimo momento negativo de  $Y$  por

$$M(r, k) = E(Y + k)^{-r} = \sum_y \frac{a(y)g(\theta)^y}{(y + k)^r f(\theta)}.$$

A relação de recorrência entre momentos negativos sucessivos da classe MPSD foi obtida por Kumar e Consul (1979) como

$$M(r, k) = \frac{1}{f(\theta)g(\theta)^k} \int_0^\theta M(r-1, k)g'(\theta)f(\theta)g(\theta)^{k-1} d\theta, \quad (2.4)$$

para  $r = 1, 2, 3, \dots$  e sendo  $M(0, k) = E(Y + k)^0 = 1$ .

De (2.4) vem

$$M(1, k) = E(Y + k)^{-1} = \frac{\int_0^\theta g'(\theta)f(\theta)g(\theta)^{k-1} d\theta}{f(\theta)g(\theta)^k}.$$

Seja  $Y$  uma variável aleatória definida em um sub-conjunto do conjunto dos números reais com função distribuição  $F(\cdot)$ . Então, o  $r$ -ésimo momento incompleto de  $Y$  é

$$\mu'_r(t) = \int_{-\infty}^t y^r dF(y).$$

O  $r$ -ésimo momento incompleto da classe MPSD em (2.1) é expresso por

$$\mu'_r(t) = \sum_{y=1}^t y^r \frac{a(y)g(\theta)^y}{f(\theta)}. \quad (2.5)$$

Diferenciando ambos os lados da equação (2.5) com relação a  $\theta$ , Tripathi et al. (1986) obtiveram a relação de recorrência entre os momentos incompletos da classe MPSD como

$$\mu'_{r+1}(t) = \frac{g(\theta)}{g'(\theta)} \frac{d\mu'_r(t)}{d\theta} + \mu'_1 \mu'_r(t).$$

O  $r$ -ésimo momento fatorial incompleto da classe MPSD é expresso por

$$\mu'_{[r]}(t) = \sum_{y=1}^t y^{[r]} \frac{a(y)g(\theta)^y}{f(\theta)}, \quad t \geq r \quad \text{e} \quad r = 1, 2, 3, \dots \quad (2.6)$$

Diferenciando ambos os lados de (2.6) em relação a  $\theta$ , a equação de recorrência entre os momentos fatoriais incompletos para a classe MPSD (TRIPATHI et al., 1986) é

$$\mu'_{[r+1]}(t) = \frac{g(\theta)}{g'(\theta)} \frac{d\mu'_{[r]}(t)}{d\theta} + \mu'_{[r]}(t)(\mu - r), \quad t \geq r \quad \text{e} \quad r = 1, 2, 3, \dots$$

## 2.3 Estimação

A classe MPSD como definida em (2.1) pode ter muitos parâmetros desconhecidos, além do parâmetro  $\theta$ , mas, neste trabalho, apenas o parâmetro  $\theta$  é considerado desconhecido. Portanto, deve-se encontrar estimadores para  $\theta$ .

### 2.3.1 Estimador de Máxima Verossimilhança

Seja  $Y_1, \dots, Y_n$  uma amostra aleatória de tamanho  $n$  da classe MPSD (2.1) e seja  $\bar{Y}$  a média amostral. A função de verossimilhança  $L = L(\theta)$  é expressa como

$$L = \prod_{i=1}^n \frac{a(y_i)g(\theta)^{y_i}}{f(\theta)}.$$

Ao diferenciar o logaritmo da função de verossimilhança,  $\log L$ , em relação a  $\theta$ , obtém-se a equação de verossimilhança para  $\theta$  como

$$\frac{1}{L} \frac{dL}{d\theta} = n \frac{g'(\theta)}{g(\theta)} [\bar{y} - \mu(\theta)] = 0. \quad (2.7)$$

A solução de (2.7) é a equação não-linear

$$\bar{y} = \mu(\theta). \quad (2.8)$$

Se  $\mu(\theta)$  for inversível, o EMV de  $\theta$ , obtido pela inversão de (2.8), é igual a

$$\hat{\theta} = \mu^{-1}(\bar{y}).$$

Se  $\mu(\theta)$  não for inversível, pode-se resolver (2.8) iterativamente utilizando o método de Newton-Raphson para obter  $\hat{\theta}$ . O viés,  $b(\hat{\theta})$ , do EMV do parâmetro  $\theta$  (CONSUL; FAMOYE, 2006) é expresso como

$$b(\hat{\theta}) = -\frac{B_1}{2n A_1^2}, \quad (2.9)$$

sendo

$$A_1 = \sum_y \frac{1}{P(Y=y)} \left[ \frac{d P(Y=y)}{d\theta} \right]^2$$

e

$$B_1 = \sum_y \frac{1}{P(Y=y)} \left[ \frac{d P(Y=y)}{d\theta} \right] \left[ \frac{d^2 P(Y=y)}{d\theta^2} \right].$$

Para a classe MPSD tem-se

$$A_1 = \left[ \frac{g'(\theta)}{g(\theta)} \right]^2 \mu_2 \quad (2.10)$$

e

$$B_1 = \left[ \frac{g'(\theta)}{g(\theta)} \right]^3 \left\{ \mu_3 + \frac{g(\theta)g''(\theta) - [g'(\theta)]^2}{[g'(\theta)]^2} \mu_2 \right\}. \quad (2.11)$$

Substituindo (2.10) e (2.11) em (2.9), o viés de  $\hat{\theta}$  é

$$b(\hat{\theta}) = -\frac{1}{2n\mu_2^2} \frac{g(\theta)}{g'(\theta)} \left\{ \mu_3 + \frac{g(\theta)g''(\theta) - [g'(\theta)]^2}{[g'(\theta)]^2} \mu_2 \right\}.$$

O EMV  $\hat{\theta}$  é não-viesado quando  $b(\hat{\theta}) = 0$ , ou seja,  $B_1 = 0$  e

$$\mu_3 = -\frac{g(\theta)g''(\theta) - [g'(\theta)]^2}{[g'(\theta)]^2} \mu_2.$$

A variância assintótica de  $\hat{\theta}$  é expressa como

$$Var(\hat{\theta}) = \frac{g(\theta)}{g'(\theta)} \bigg/ \left[ n \frac{d\mu}{d\theta} \right].$$

### 2.3.2 Estimação Intervalar

A estimação intervalar de  $\theta$  na classe MPSD foi desenvolvida por Famoye e Consul (1989a). Intervalos de confiança  $(\theta_l, \theta_u)$  para  $\theta$  em pequenas amostras são calculados pela resolução das equações não-lineares

$$\sum_{y=t}^{\infty} a(y) \frac{g(\theta_l)^y}{[f(\theta_l)]^n} = \frac{1}{2} \alpha \quad \text{e} \quad \sum_{y=0}^t a(y) \frac{g(\theta_u)^y}{[f(\theta_u)]^n} = \frac{1}{2} \alpha,$$



em que  $t = \sum_{i=1}^n y_i$ ,  $0 < \alpha < 1$  e coeficiente de confiança igual a  $1 - \alpha$ .

Para grandes amostras, o intervalo de confiança para  $\theta$  pode ser baseado na estatística

$$W = (\bar{Y} - \mu)\sqrt{n}/\sigma,$$

que converge para a distribuição normal com média zero e variância unitária. O limite superior  $\theta_u$  e o limite inferior  $\theta_l$  são as soluções das equações

$$\bar{y} - \frac{g(\theta)f'(\theta)}{g'(\theta)f(\theta)} \pm z_{\alpha/2} \left[ \frac{g(\theta)}{ng'(\theta)} \frac{d\mu}{d\theta} \right]^{1/2} = 0, \quad (2.12)$$

sendo  $z_{\alpha/2}$  o valor crítico tabelado da distribuição normal padrão.

O resultado em (2.12) é baseado na estatística  $\bar{Y}$ . Famoye e Consul (1989a), também, encontram um intervalo de confiança para  $\theta$  baseado em duas estatísticas, a média amostral  $\bar{Y}$  e a variância amostral  $S^2$ . Um intervalo de confiança  $100(1 - \alpha)\%$  bilateral para  $\theta$  em grandes amostras é obtido por

$$1 - \alpha = \Pr \left[ \bar{Y} - z_{\alpha/2} \frac{S}{\sqrt{n}} < \frac{g(\theta)f'(\theta)}{g'(\theta)f(\theta)} < \bar{Y} + z_{\alpha/2} \frac{S}{\sqrt{n}} \right]. \quad (2.13)$$

A vantagem de (2.13) sobre (2.12) é que a desigualdade (2.13) pode ser resolvida algebricamente para  $\theta$  e o resultado é expresso na forma

$$1 - \alpha = \Pr[\theta_l < \theta < \theta_u],$$

em que  $(\theta_l, \theta_u)$  é um intervalo de confiança de  $100(1 - \alpha)\%$  bilateral para  $\theta$ .

## 3 Alguns Casos Especiais da Classe MPSD

Neste capítulo algumas distribuições que pertencem à classe de distribuições em série de potências modificadas são apresentadas, bem como, algumas de suas propriedades, baseadas nas fórmulas encontradas no Capítulo 2. As distribuições foram subdivididas em dois grupos. Um primeiro grupo em que as distribuições foram parametrizadas em função da média, ou seja,  $\mu = \theta$  e outro em que a média da distribuição é uma função do parâmetro  $\theta$ , ou seja,  $\mu = \tau(\theta)$ , sendo que  $\tau(\cdot)$  não é a função identidade.

### 3.1 Distribuições Parametrizadas na Média $\mu$

Existem distribuições na classe MPSD que podem ser parametrizadas de tal forma que a média delas seja igual ao parâmetro  $\mu$ , ou seja,  $E(Y) = \mu(\theta) = \mu'_1 = \mu$ , essa parametrização é importante para estudar essas distribuições no contexto de modelos de regressão. Algumas distribuições com essa característica são apresentadas aqui.

#### 3.1.1 Distribuição Binomial

A distribuição binomial é a distribuição de probabilidade discreta que representa o número de sucessos (ou insucessos) em uma sequência de  $m$  experimentos independentes, em que, cada sucesso (ou insucesso), tem probabilidade  $p$  de ocorrer. Cada repetição do experimento, também, é chamada de ensaio Bernoulli. Na verdade, quando  $m = 1$ , a distribuição binomial é uma distribuição de Bernoulli.

A função de probabilidade da distribuição binomial é expressa como

$$P(Y = y) = \binom{m}{y} \left( \frac{\mu}{m - \mu} \right)^y \left( 1 + \frac{\mu}{m - \mu} \right)^{-m},$$

sendo  $y = 0, 1, \dots, m$ ,  $0 < \mu < m$  e  $m > 0$  conhecido. Tem-se ainda que  $f(\mu) = \left(1 + \frac{\mu}{m-\mu}\right)^m$  e  $g(\mu) = \frac{\mu}{m-\mu}$ . A variância da distribuição binomial é dada por  $\mu \left(1 - \frac{\mu}{m}\right)$ .

(a) Função de probabilidade

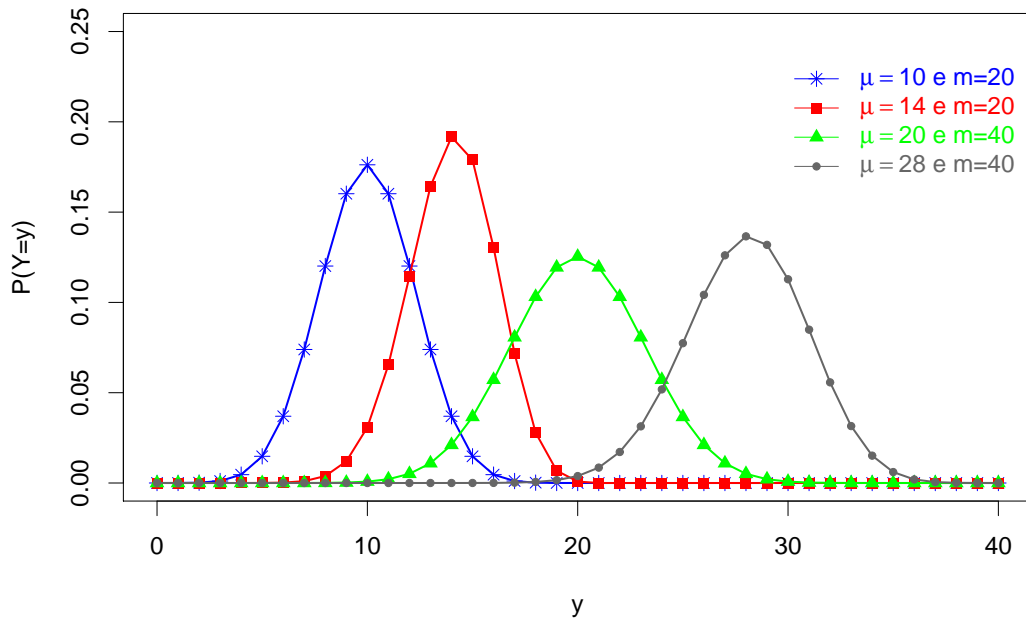
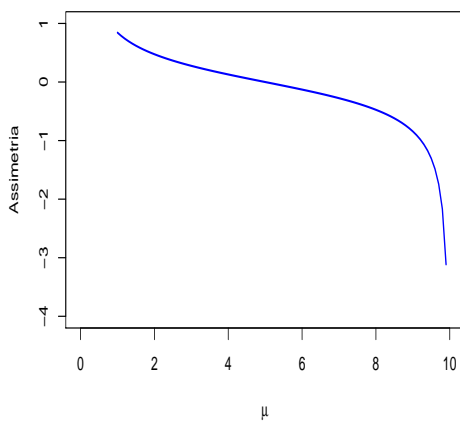
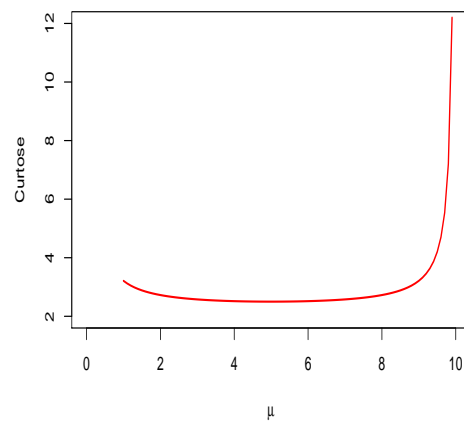
(b) Assimetria ( $m = 10$ )(c) Curtose ( $m = 10$ )

Figura 3.1: Gráficos da distribuição binomial

Observando-se a figura 3.1 nota-se pelo gráfico (a) que a distribuição binomial é simétrica em torno da média  $\mu$  e pelo gráfico (b) tem-se que a assimetria tem valores mais altos para valores pequenos da média  $\mu$  e vice-versa. A curtose, como pode ser visto no gráfico (c), tem um leve decréscimo para valores pequenos de  $\mu$ , mas para valores de  $\mu$  próximos a  $m$  tem crescimento acentuado.

### 3.1.2 Distribuição Binomial Negativa

A distribuição binomial negativa é, também, conhecida como distribuição de Pascal ou distribuição de Polya. A distribuição geométrica é um caso especial da distribuição binomial negativa. A distribuição binomial negativa modela o número de tentativas necessárias para se obter  $m$  resultados desejados (ou sucessos). A binomial negativa é uma variante da distribuição binomial. Os eventos continuam a ocorrer até o  $m$ -ésimo sucesso.

Esta distribuição foi a primeira distribuição, considerada em estatística, que tem variância maior do que a média.

A função de probabilidade da distribuição binomial negativa é expressa como

$$P(Y = y) = \frac{\Gamma(m+y)}{y!\Gamma(m)} \left(\frac{\mu}{m+\mu}\right)^y \left(1 - \frac{\mu}{m+\mu}\right)^m,$$

sendo  $y = 0, 1, 2, \dots$ ,  $\mu > 0$  e  $m > 0$  conhecido. Tem-se ainda que  $f(\mu) = \left(1 - \frac{\mu}{\mu+m}\right)^{-m}$  e  $g(\mu) = \frac{\mu}{\mu+m}$ . A variância da distribuição binomial negativa iguala  $\mu \left(1 + \frac{\mu}{m}\right)$ .

Observando-se a figura 3.2 nota-se pelo gráfico (a) que a distribuição binomial negativa é simétrica em torno da média  $\mu$ , para  $m \geq 2\mu$  e nos demais é assimétrica à esquerda. Pelo gráfico (b) tem-se que a assimetria tem valores mais altos para valores de  $\mu$  menores que  $m$ . A curtose, como pode ser visto no gráfico (c), tem crescimento acentuado para valores de  $\mu$  próximos a  $m$ .

### 3.1.3 Distribuição Binomial Negativa Generalizada

A distribuição binomial negativa generalizada tem muitas aplicações importantes em vários campos de estudo como em teoria das filas, em processos de reação de polimerização em química, para modelar o número de acidentes e para dados de contagens em que a variância é muito maior do que a média dos dados (JAIN; CONSUL, 1971). A distribuição binomial negativa generalizada recebeu a atenção de muitos pesquisadores, tais como, Consul e Shenton (1973c), Kumar e Consul (1979, 1980), Consul (1989), Consul e Famoye (1986, 1995), Famoye e Consul (1989b) e Famoye (1997), entre outros (CONSUL; FAMOYE, 2006).

A função de probabilidade da distribuição binomial negativa generalizada é expressa

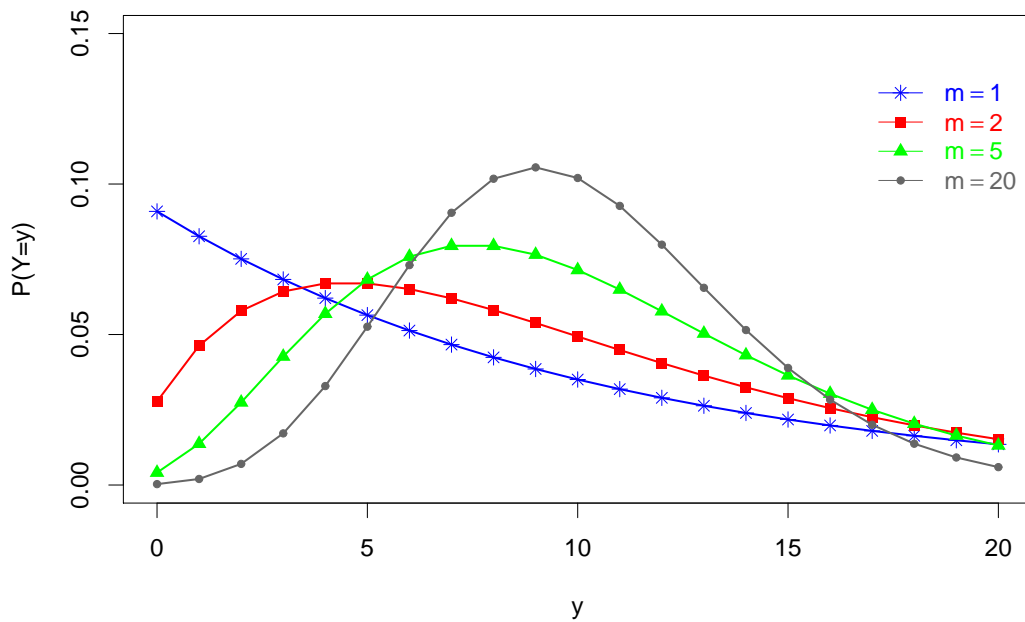
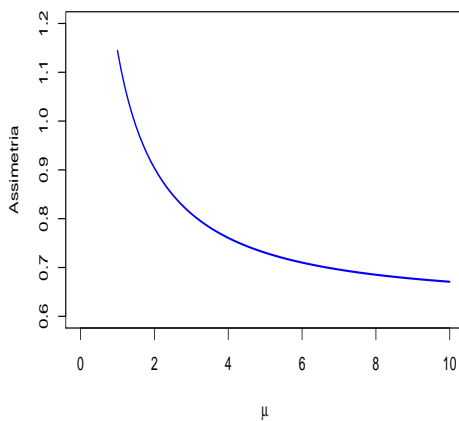
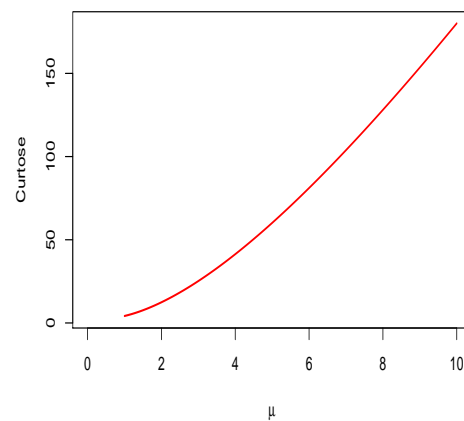
(a) Função de probabilidade ( $\mu = 10$ )(b) Assimetria ( $m = 10$ )(c) Curtose ( $m = 10$ )

Figura 3.2: Gráficos da distribuição binomial negativa

como

$$P(Y = y) = \frac{m\Gamma(m + \beta y)}{y!\Gamma(m + \beta y - y + 1)} \left\{ \left( \frac{\mu}{m + \beta\mu} \right) \left[ \frac{m + (\beta - 1)\mu}{m + \beta\mu} \right]^{\beta - 1} \right\}^y \left[ \frac{m + (\beta - 1)\mu}{m + \beta\mu} \right]^m,$$

sendo  $y = 0, 1, \dots$ ,  $\mu > 0$ ,  $\beta > 0$  e  $m > 0$  conhecido, mas não, necessariamente, inteiro.

Tem-se ainda que  $f(\mu) = \left[ \frac{m + (\beta - 1)\mu}{m + \beta\mu} \right]^{-m}$  e  $g(\mu) = \left( \frac{\mu}{m + \beta\mu} \right) \left[ \frac{m + (\beta - 1)\mu}{m + \beta\mu} \right]^{\beta - 1}$ . A variância da distribuição binomial negativa generalizada é expressa como  $\mu[1 + \beta\mu m^{-1}][1 + (\beta - 1)\mu m^{-1}]$ .

Observações: Para  $\beta = 0$  tem-se a distribuição binomial com  $f(\mu) = \left(1 + \frac{\mu}{m-\mu}\right)^m$  e  $g(\mu) = \frac{\mu}{m-\mu}$ . Para  $\beta = 1$  tem-se a distribuição binomial negativa com  $f(\mu) = \left(1 - \frac{\mu}{\mu+m}\right)^{-m}$  e  $g(\mu) = \frac{\mu}{\mu+m}$ .

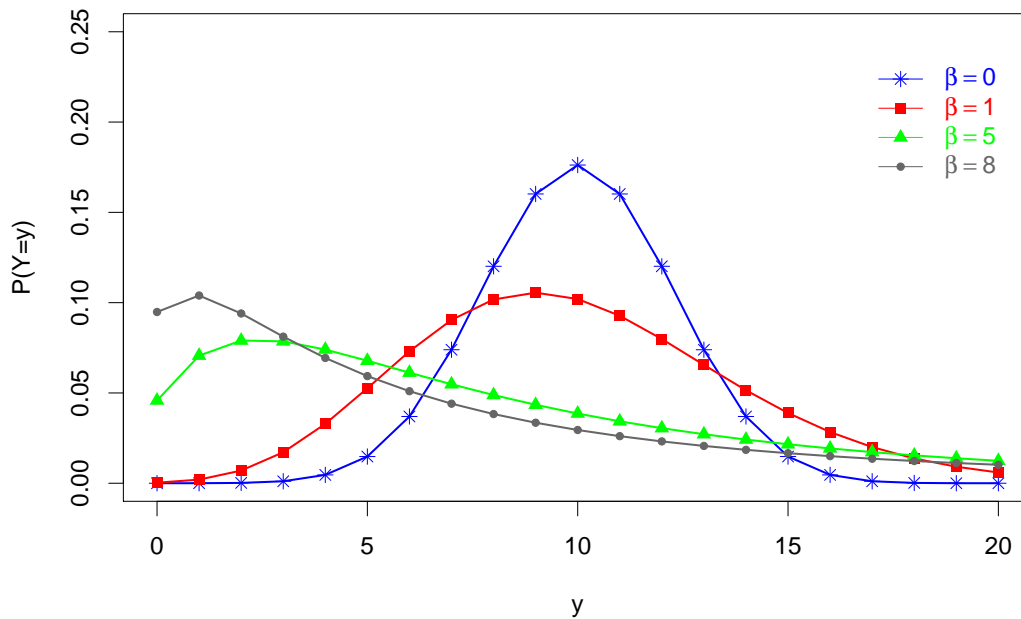
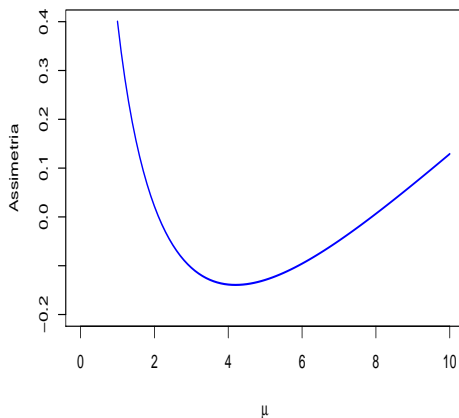
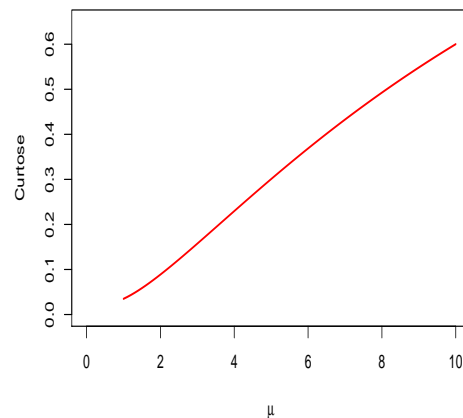
(a) Função de probabilidade ( $\mu = 10$  e  $m = 20$ )(b) Assimetria ( $\beta = 2$  e  $m = 10$ )(c) Curtose ( $\beta = 2$  e  $m = 10$ )

Figura 3.3: Gráficos da distribuição binomial negativa generalizada

Observando-se a figura 3.3 nota-se pelo gráfico (a) que a distribuição binomial negativa generalizada é simétrica em torno da média  $\mu$ , para  $\beta = 0$ , tem uma leve assimetria a esquerda para  $\beta = 1$  e nos demais casos é assimétrica à esquerda. Pelo gráfico (b) tem-se que a assimetria decresce para valores menores que  $\beta$  e cresce para valores maiores que  $\beta$ . A curtose, como pode ser visto no gráfico (c), é crescente para valores, também,

crecentes de  $\mu$ .

### 3.1.4 Distribuição de Borel

A distribuição de Borel foi, inicialmente, utilizada para representar o problema de fluxos de veículos em sentidos opostos em uma estrada, quando esta só comporta, em determinado trecho, um veículo de cada vez. Mas pode ser aplicada a outras situações envolvendo fluxo de objetos, como por exemplo, artigos produzidos em uma fábrica, o fluxo de pedestres em uma via com as mesmas características da estrada, entre outros (BOREL, 1942).

Uma variável aleatória  $Y$  tem distribuição de Borel se

$$P(Y = y) = \frac{y^{y-1} \left[ \left(1 - \frac{1}{\mu}\right) e^{-1 + \frac{1}{\mu}} \right]^y}{y! \left(1 - \frac{1}{\mu}\right)},$$

sendo  $y = 1, 2, \dots$  e  $\mu > 1$ . Tem-se ainda que  $f(\mu) = 1 - \frac{1}{\mu}$  e  $g(\mu) = \left(1 - \frac{1}{\mu}\right) e^{-1 + \frac{1}{\mu}}$ . A variância da distribuição de Borel é dada por  $\mu^2(\mu - 1)$ .

Observando-se a figura 3.4 nota-se pelo gráfico (a) que a distribuição de Borel é assimétrica a esquerda e decrescente. Pelo gráfico (b) tem-se que a assimetria tem um leve decrescimento para valores pequenos de  $\mu$  e, em seguida, tem comportamento crescente para valores crescentes de  $\mu$ . A curtose, como pode ser visto no gráfico (c), tem comportamento similar ao da assimetria.

### 3.1.5 Distribuição de Borel-Tanner

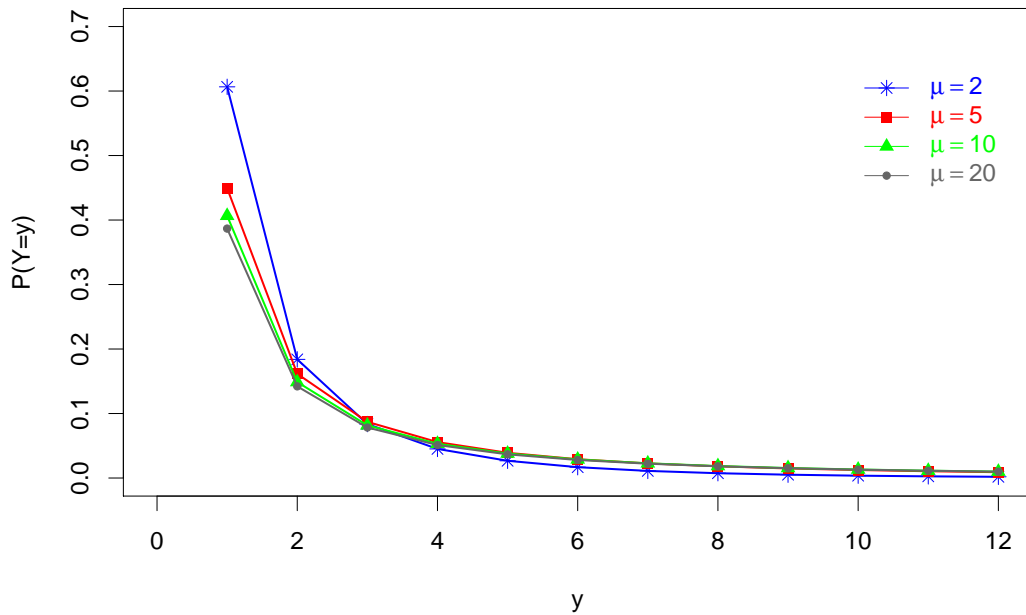
A distribuição de Borel-Tanner descreve o número total de clientes atendidos antes que a fila seja esvaziada, supondo que é fila única com chegadas de clientes com distribuição de Poisson (com taxa constante) e o tempo de atendimento de cada cliente é constante (TANNER, 1953).

A variável aleatória  $Y$  tem distribuição de Borel-Tanner se

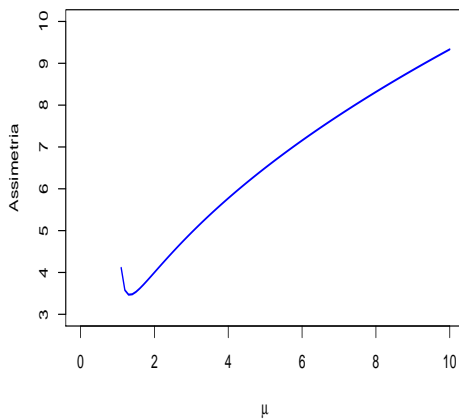
$$P(Y = y) = \frac{m y^{y-m-1}}{(y-m)!} \left[ \left(1 - \frac{m}{\mu}\right) e^{-1 + \frac{m}{\mu}} \right]^y \left(1 - \frac{m}{\mu}\right)^{-m},$$

sendo  $y = m, m + 1, \dots$ ,  $\mu > 1$  e  $m \geq 1$  um inteiro conhecido. Tem-se ainda que  $f(\mu) =$

(a) Função de probabilidade



(b) Assimetria



(c) Curtose

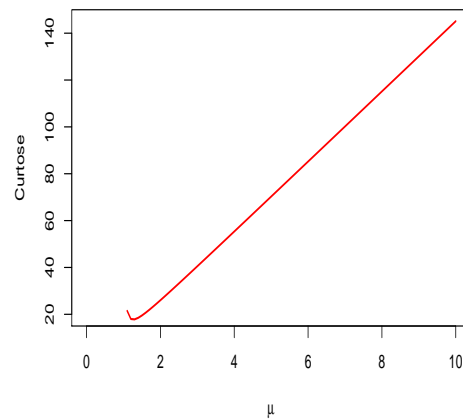


Figura 3.4: Gráficos da distribuição de Borel

$\left(1 - \frac{m}{\mu}\right)^m$  e  $g(\mu) = \left(1 - \frac{m}{\mu}\right)e^{-1+\frac{m}{\mu}}$ . A variância da distribuição de Borel-Tanner é igual a  $\frac{(\mu - m)\mu^2}{m^2}$ . Para  $m = 1$  tem-se a distribuição de Borel.

Observando-se a figura 3.5 nota-se pelo gráfico (a) que a distribuição de Borel-Tanner é assimétrica a esquerda. Pelo gráfico (b) tem-se que a assimetria tem um decréscimo para valores pequenos de  $\mu$  e, em seguida, tem comportamento crescente para valores crescentes de  $\mu$ . A curtose, como pode ser visto no gráfico (c), tem comportamento similar ao da assimetria.



(a) Função de probabilidade

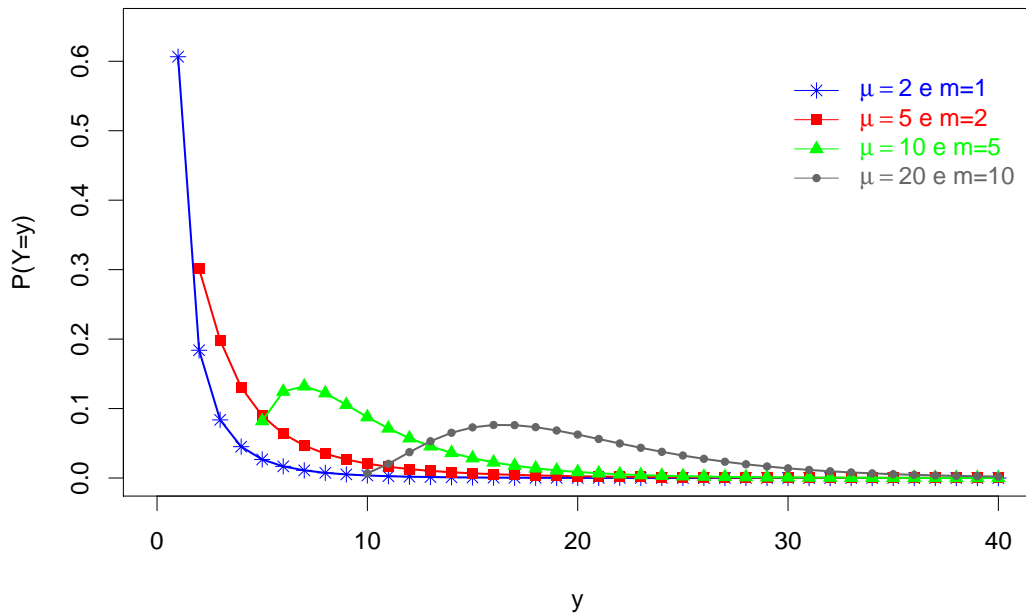
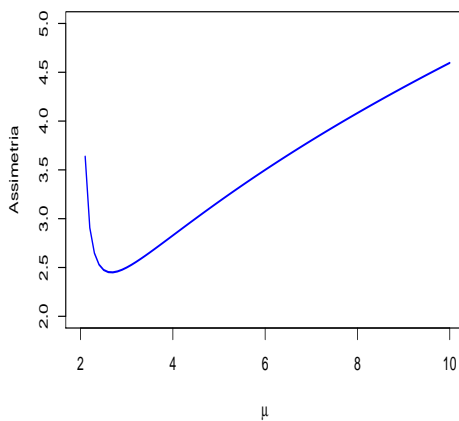
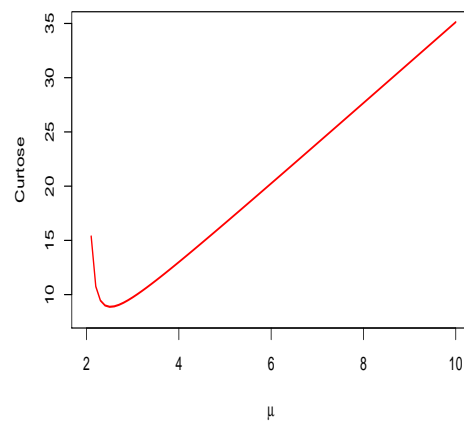
(b) Assimetria ( $m = 2$ )(c) Curtose ( $m = 2$ )

Figura 3.5: Gráficos da distribuição de Borel-Tanner

### 3.1.6 Distribuição de Consul

Suponha que uma fila é iniciada com um membro e tem número de chegadas com distribuição binomial e tempo de atendimento constante. A distribuição de Consul representa a probabilidade que exatamente  $y$  membros sejam atendidos antes que a fila se encerre.

Um importante problema em estudos de epidemias é encontrar a distribuição de probabilidade do número total de infectados,  $N(u)$ , em qualquer parte do *habitat*, sendo as pessoas infectadas a partir de uma única pessoa infectada em cada local,  $u$ , até o mo-

mento da extinção da epidemia. Nesse caso, supondo  $X_0 = 1$  com probabilidade um e o número de novas pessoas infectadas, entre as suscetíveis, tendo distribuição binomial, o número de pessoas suscetíveis é finito,  $m$ , como utilizado por Kumar (1981). Se o processo de infecção continua assim, indefinidamente, o número de pessoas infectadas em dado momento segue o modelo de Consul.

A função de probabilidade da distribuição de Consul (CONSUL, 1990b) é expressa como

$$P(Y = y) = \frac{\Gamma(\beta y + 1)}{y! \Gamma(\beta y - y)} \left[ \frac{(1 - \mu^{-1})(\beta - 1 + \mu^{-1})^{\beta - 1}}{\beta^\beta} \right]^y \frac{1}{\left[ \frac{\mu - 1}{\mu(\beta - 1) + 1} \right]},$$

sendo  $y = 1, 2, \dots$ ,  $\mu > 1$  e  $\beta \geq 1$ . Tem-se ainda que  $f(\mu) = \frac{\mu - 1}{\mu(\beta - 1) + 1}$  e  $g(\mu) = \beta^{-\beta} (1 - \mu^{-1})(\beta - 1 + \mu^{-1})^{\beta - 1}$ . A variância da distribuição de Consul é expressa como  $\frac{\mu(\mu - 1)[\mu(\beta - 1) + 1]}{\beta}$ .

Observando-se a figura 3.6 nota-se pelo gráfico (a) que a distribuição de Consul é assimétrica a esquerda. Pelo gráfico (b) tem-se que a assimetria é decrescente para valores crescentes de  $\mu$ . A curtose, como pode ser visto no gráfico (c), é crescente para valores, também, crescentes de  $\mu$ .

### 3.1.7 Distribuição Delta-binomial

Se  $X_1, \dots, X_m$  são variáveis aleatórias independentemente distribuídas com distribuição Consul, então a soma amostral,  $Y = \sum_{i=1}^m X_i$ , tem distribuição delta-binomial expressa como

$$P(Y = y) = \frac{m \Gamma(\beta y + 1)}{y(y - m)! \Gamma(\beta y - y + m + 1)} \frac{[\beta^{-\beta} (1 - \frac{m}{\mu})(\beta - 1 + \frac{m}{\mu})^{\beta - 1}]^y}{\left[ \frac{\mu - m}{\mu(\beta - 1) + m} \right]^m},$$

sendo  $y = m, m + 1, \dots$ ,  $\mu > m$  um inteiro conhecido,  $m \geq 1$  e  $\beta \geq 1$ . Tem-se ainda que  $f(\mu) = \left[ \frac{\mu - m}{\mu(\beta - 1) + m} \right]^m$  e  $g(\mu) = \beta^{-\beta} \left( 1 - \frac{m}{\mu} \right) \left( \beta - 1 + \frac{m}{\mu} \right)^{\beta - 1}$ . A variância da distribuição delta-binomial é igual a  $\frac{\mu(\mu - m)[\mu(\beta - 1) + m]}{\beta m^2}$ .

Observando-se a figura 3.7 nota-se pelo gráfico (a) que a distribuição delta-binomial é assimétrica a esquerda. Pelo gráfico (b) tem-se que a assimetria tem um decréscimo para valores pequenos de  $\mu$  e, em seguida, tem comportamento crescente para valores crescentes de  $\mu$ . A curtose, como pode ser visto no gráfico (c), tem comportamento similar ao da assimetria.

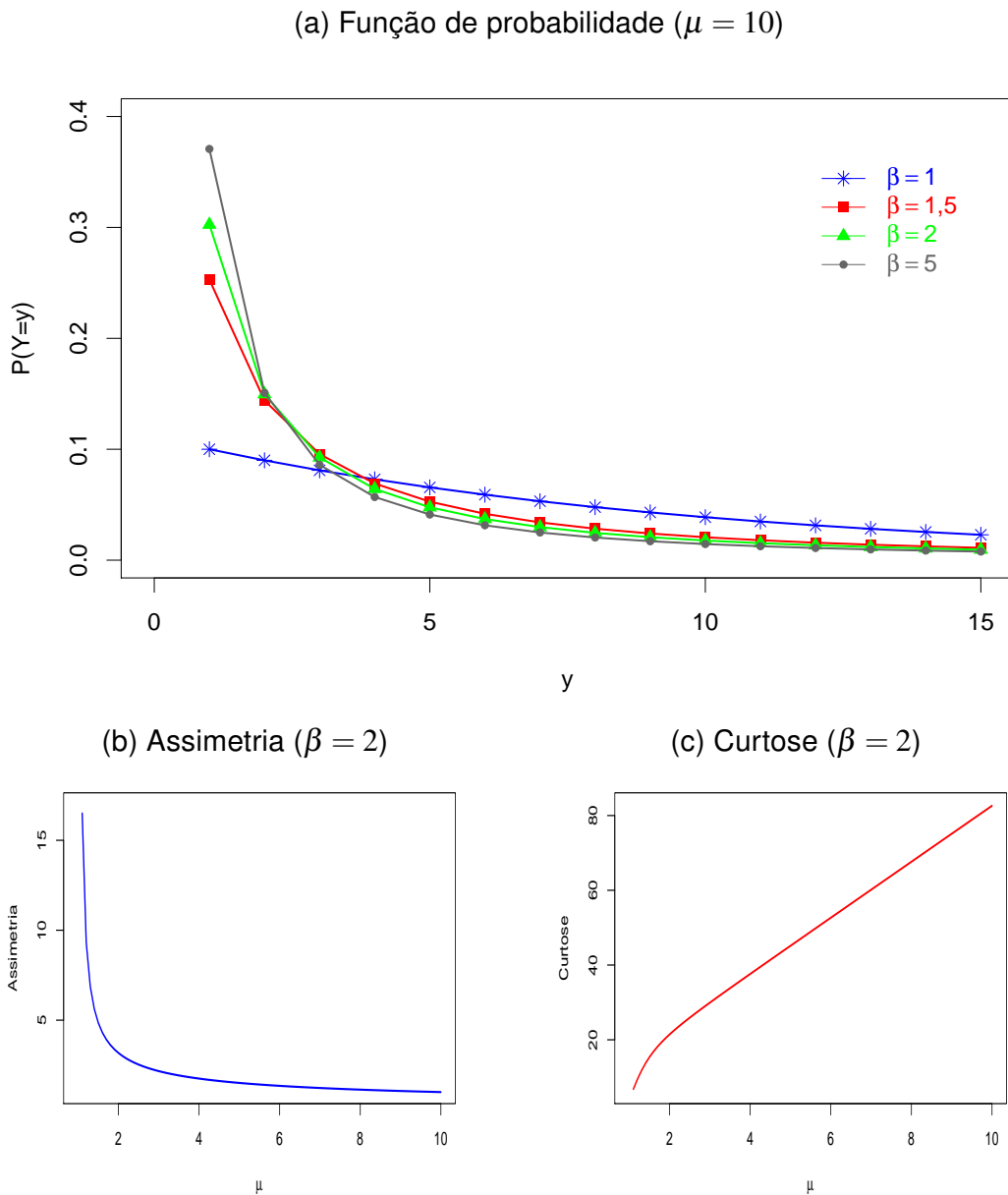


Figura 3.6: Gráficos da distribuição de Consul

### 3.1.8 Distribuição de Geeta

A distribuição de Geeta (CONSUL, 1990a) tem dois parâmetros de forma “L-shaped” e é mais versátil do que as distribuições em série logarítmica, de Pareto discreta e de Yale.

A distribuição de Geeta representa a probabilidade de que exatamente  $y$  membros sejam atendidos antes que a fila encerre, se a fila é iniciada com um membro e tem número de chegadas com distribuição binomial negativa e tempo constante.

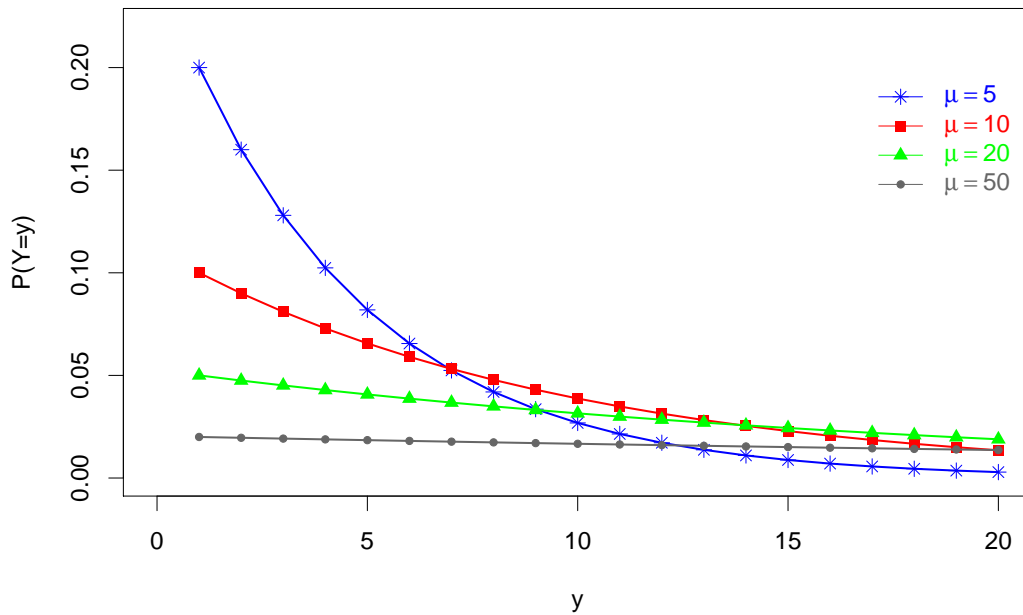
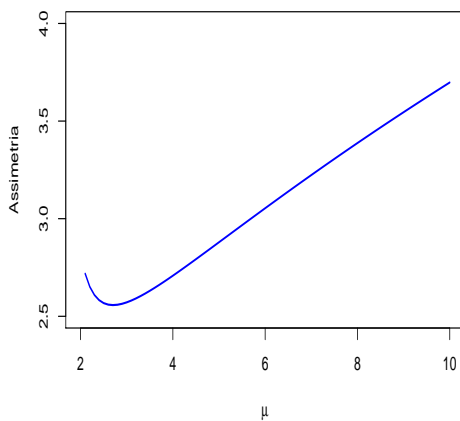
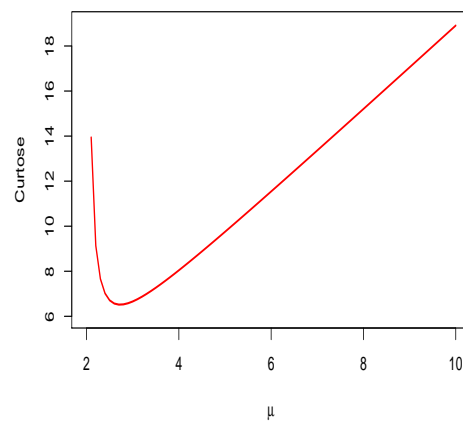
(a) Função de probabilidade ( $\beta = 1$  e  $m = 1$ )(b) Assimetria ( $\beta = 2$  e  $m = 2$ )(c) Curtose ( $\beta = 2$  e  $m = 2$ )

Figura 3.7: Gráficos da distribuição delta-binomial

A função de probabilidade da distribuição de Geeta é expressa como

$$P(Y = y) = \frac{\Gamma(\beta y - 1)}{y! \Gamma(\beta y - y)} \left\{ \left( \frac{\mu - 1}{\beta \mu - 1} \right) \left[ \frac{(\beta - 1)\mu}{\beta \mu - 1} \right]^{\beta - 1} \right\}^y \left( \frac{\mu - 1}{\beta \mu - 1} \right)^{-1},$$

sendo  $y = 1, 2, \dots$ ,  $\mu \geq 1$  e  $\beta > 1$ . Tem-se ainda que  $f(\mu) = \left( \frac{\mu - 1}{\beta \mu - 1} \right)$  e  $g(\mu) = \left( \frac{\mu - 1}{\beta \mu - 1} \right) \left[ \frac{(\beta - 1)\mu}{\beta \mu - 1} \right]^{\beta - 1}$ . A variância da distribuição de Geeta é igual a  $\mu(\mu - 1)(\beta \mu - 1)(\beta - 1)^{-1}$ .

Observando-se a figura 3.8 nota-se pelo gráfico (a) que a distribuição de Geeta é assimétrica a esquerda e em formato de "L". Pelo gráfico (b) tem-se que a assimetria é crescente para valores crescentes de  $\mu$ . A curtose, como pode ser visto no gráfico (c), tem

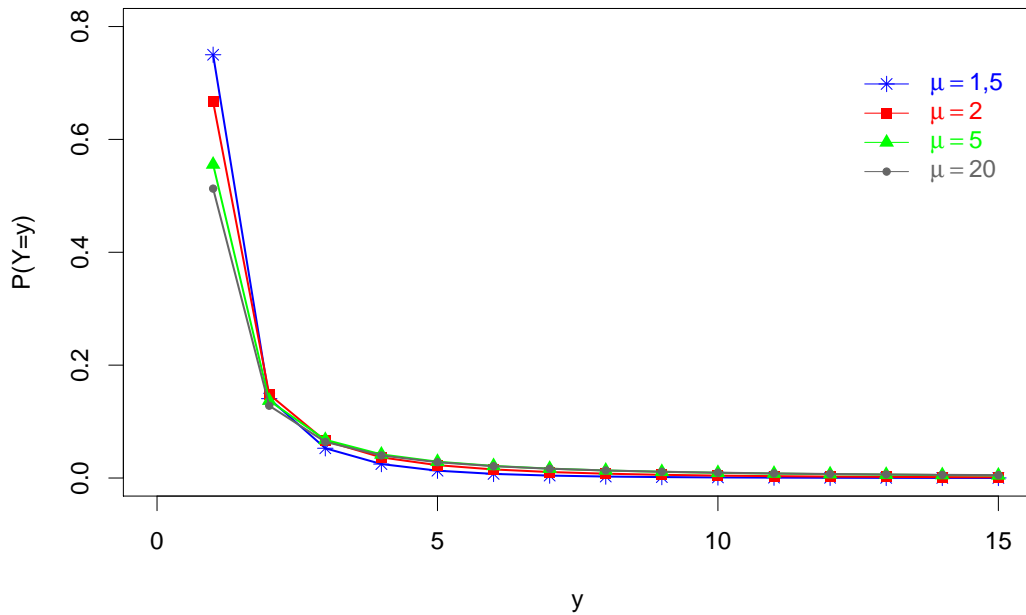
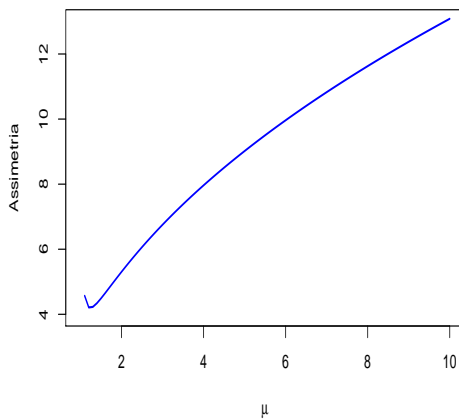
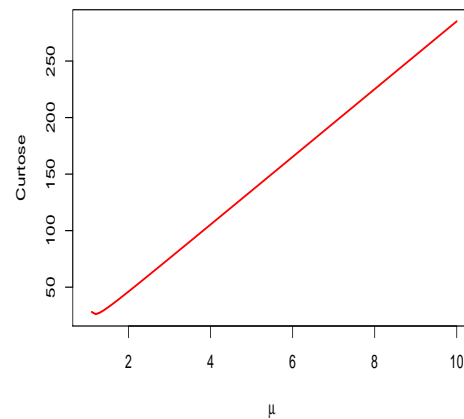
(a) Função de probabilidade ( $\beta = 2$ )(b) Assimetria ( $\beta = 2$ )(c) Curtose ( $\beta = 2$ )

Figura 3.8: Gráficos da distribuição de Geeta

comportamento similar a assimetria.

### 3.1.9 Distribuição de Haight

A distribuição de Haight é um caso especial da distribuição de Geeta quando  $\beta = 2$ . Em teoria das filas a distribuição de Haight representa a probabilidade de que exatamente  $y$  membros da fila sejam atendidos antes que a primeira fila encerre, começando com um membro, sendo  $1 - \mu^{-1}$  a intensidade do tráfego, com chegadas segundo a distribuição

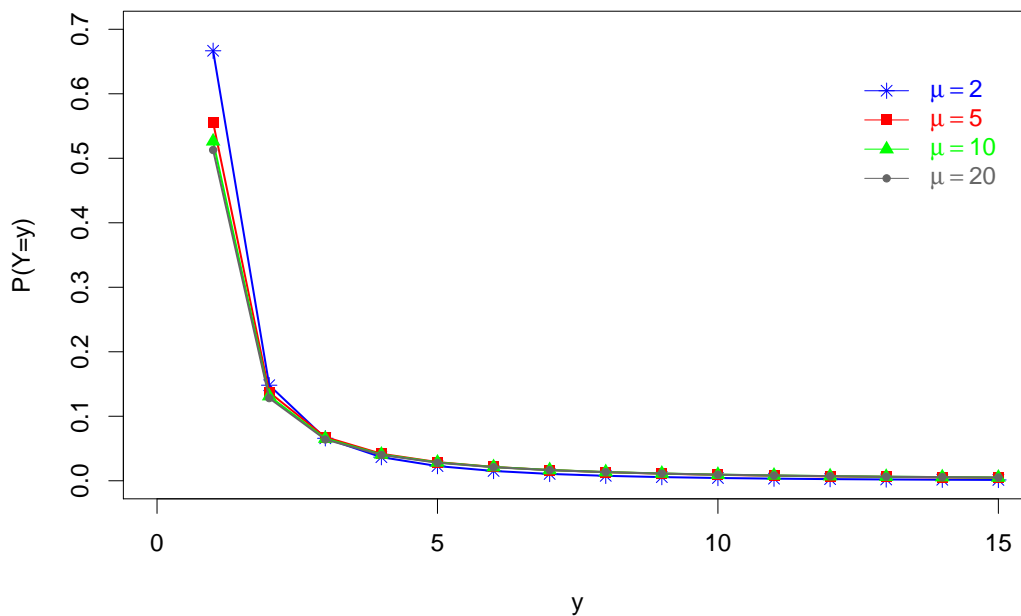
de Poisson e tempo de atendimento exponencial negativo. Se o tempo de atendimento é constante, a distribuição será Borel-Tanner.

A função de probabilidade da distribuição de Haight é expressa por

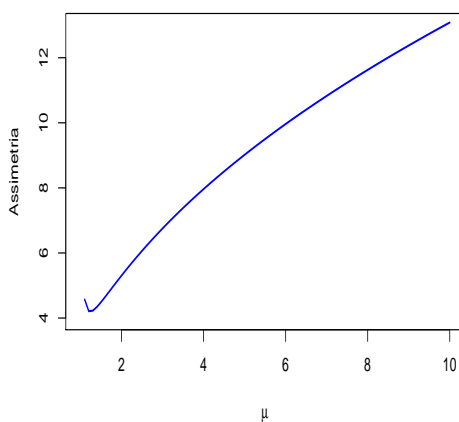
$$P(Y = y) = \frac{(2y - 2)!}{y!(y - 1)!} \left[ \frac{\mu(\mu - 1)}{(2\mu - 1)^2} \right]^y \left( \frac{\mu - 1}{2\mu - 1} \right)^{-1}$$

sendo  $y = 1, 2, \dots$  e  $\mu > 1$ . Tem-se ainda que  $f(\mu) = \frac{\mu - 1}{2\mu - 1}$  e  $g(\mu) = \frac{\mu(\mu - 1)}{(2\mu - 1)^2}$ . A variância da distribuição de Haight é igual a  $\mu(\mu - 1)(2\mu - 1)$ .

(a) Função de probabilidade



(b) Assimetria



(c) Curtose

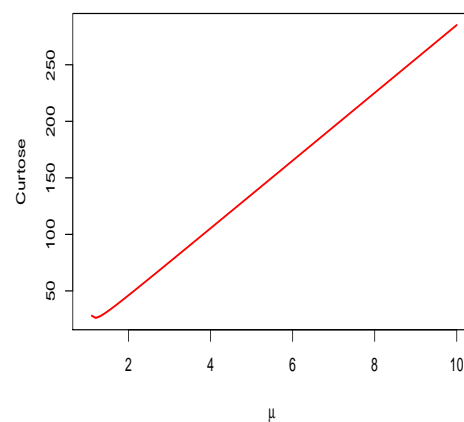


Figura 3.9: Gráficos da distribuição de Haight

Observando-se a figura 3.9 nota-se pelo gráfico (a) que a distribuição de Haight é assimétrica a esquerda. Pelo gráfico (b) tem-se que a assimetria é crescente para valores crescentes de  $\mu$ . A curtose, como pode ser visto no gráfico (c), tem comportamento similar a assimetria.

### 3.1.10 Distribuição “Lost Game”

A variável aleatória  $Y$  tem distribuição “lost game” se

$$P(Y = y) = \frac{m}{2y - m} \binom{2y - m}{y} \left[ \frac{\mu(\mu - m)}{(2\mu - m)^2} \right]^y \left( \frac{\mu - m}{2\mu - m} \right)^{-m},$$

sendo  $y = m, m + 1, \dots$ ,  $m \geq 1$  inteiro e conhecido e  $\mu > m$ . Tem-se ainda que  $f(\mu) = \left( \frac{\mu - m}{2\mu - m} \right)^m$  e  $g(\mu) = \frac{\mu(\mu - m)}{(2\mu - m)^2}$ . A variância da distribuição “lost game” é  $\frac{\mu^2(\mu - m)(2\mu - m)}{m^2}$ .

Observando-se a figura 3.10 nota-se pelo gráfico (a) que a distribuição “lost game” é assimétrica a esquerda. Pelo gráfico (b) tem-se que a assimetria é decrescente para valores pequenos de  $\mu$  e, em seguida, tem comportamento crescente para valores crescentes de  $\mu$ . A curtose, como pode ser visto no gráfico (c), tem comportamento similar ao da assimetria.

### 3.1.11 Distribuição de Poisson

A distribuição de Poisson foi descoberta por Siméon-Denis Poisson e publicada, conjuntamente com a sua teoria da probabilidade, em 1838 em “*Recherches sur la probabilité des jugements en matières criminelles et matière civile*” (Inquérito sobre a probabilidade em julgamentos sobre matérias criminais e civis). O trabalho concentrava-se em certas variáveis aleatórias  $N$  que representavam, entre outras coisas, o número de ocorrências discretas (por vezes chamadas de “chegadas”) que ocorriam durante um intervalo de tempo especificado.

A distribuição de Poisson surge em conexão com os processos de Poisson. Ela aplica-se aos vários fenômenos de natureza discreta (isto é, aqueles que podem assumir os valores  $0, 1, 2, 3, \dots$ , durante um determinado período de tempo ou numa dada área), sempre que a probabilidade do fenômeno acontecer seja constante no tempo ou no espaço. Exemplos de eventos que podem ser modelados com a distribuição de Poisson incluem:

(a) Função de probabilidade

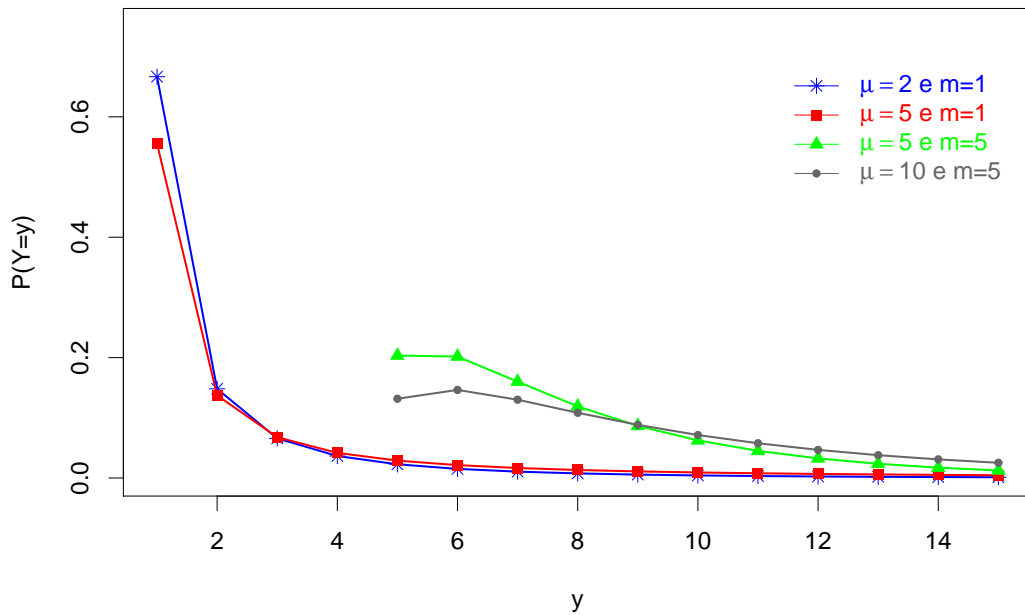
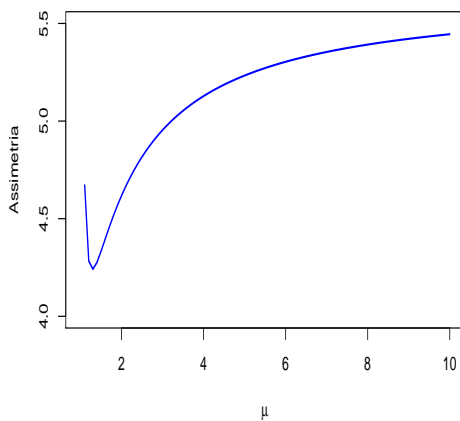
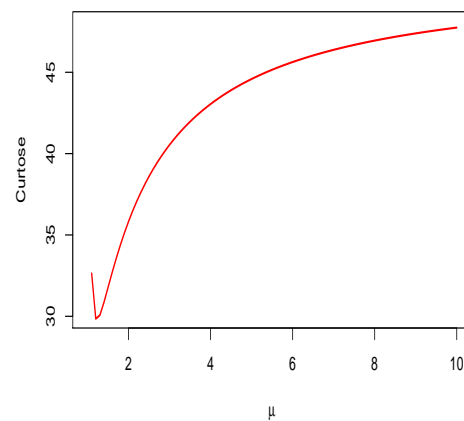
(b) Assimetria ( $m = 1$ )(c) Curtose ( $m = 1$ )

Figura 3.10: Gráficos da distribuição “lost game”

- O número de veículos que passam por um determinado ponto de uma estrada (suficientemente distante dos semáforos), durante um determinado período de tempo.
- O número de chamadas telefônicas que chegam a uma central por minuto.
- O número de vezes que um servidor web é acessado por minuto.
- O número de mutações em uma determinada sequência de DNA após uma determinada quantidade de radiação.
- O número de pinheiros por unidade de área de uma floresta.



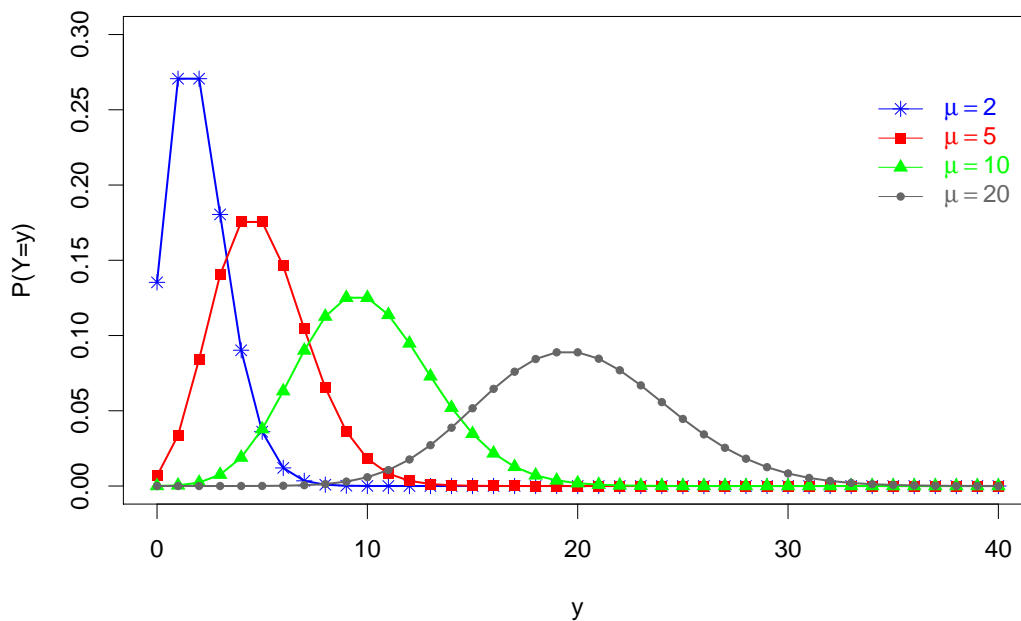
- O número de lâmpadas que queimam em um determinado período de tempo.

A função de probabilidade da distribuição de Poisson é expressa por

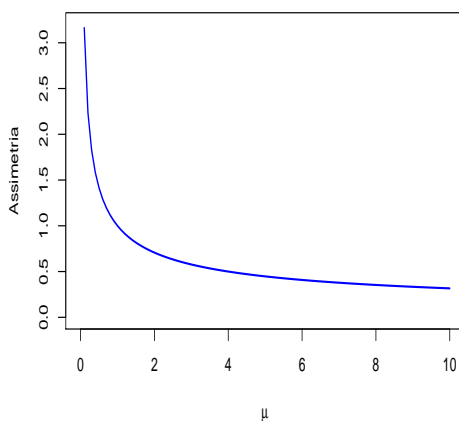
$$P(Y = y) = \frac{e^{-\mu} \mu^y}{y!},$$

sendo  $y = 0, 1, \dots$  e  $\mu > 0$ . Tem-se ainda que  $f(\mu) = e^\mu$  e  $g(\mu) = \mu$ . A variância da distribuição de Poisson é igual a  $\mu$ .

(a) Função de probabilidade



(b) Assimetria



(c) Curtose

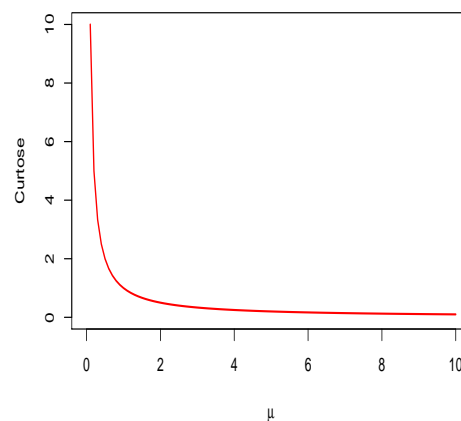


Figura 3.11: Gráficos da distribuição de Poisson

Observando-se a figura 3.11 nota-se pelo gráfico (a) que a distribuição de Poisson é simétrica para valores grandes de  $\mu$  e é levemente assimétrica a esquerda para valores

pequenos de  $\mu$ . Pelo gráfico (b) tem-se que a assimetria é decrescente e a curtose, como pode ser visto no gráfico (c), tem comportamento similar.

### 3.1.12 Distribuição de Poisson Generalizada

A distribuição de Poisson generalizada (CONSUL; JAIN, 1973a) é eficiente para modelar conjuntos de dados com diferentes situações de variância em relação à média (equidispersão, subdispersão e sobredispersão).

A função de probabilidade da distribuição de Poisson generalizada é expressa por

$$P(Y = y) = \frac{(1 - \beta y)^{y-1}}{y!} \left[ \frac{\mu e^{-\beta \mu (1 + \beta \mu)^{-1}}}{1 + \beta \mu} \right]^y e^{-\mu (1 + \beta \mu)^{-1}},$$

sendo  $y = 0, 1, \dots$ ,  $\mu > 0$  e  $\beta \geq 0$ . Tem-se ainda que  $f(\mu) = e^{\mu(1+\beta\mu)^{-1}}$  e  $g(\mu) = \frac{\mu e^{-\beta\mu(1+\beta\mu)^{-1}}}{1 + \beta\mu}$ . A variância da distribuição de Poisson generalizada é igual a  $\mu(1 + \beta\mu)^2$ .

Para  $\beta > 0$ , a variância de  $Y$  excede a média  $\mu$  (sobredispersão); para  $-\frac{2}{\mu} < \beta < 0$  a média  $\mu$  excede a variância de  $Y$  (subdispersão) e para  $\beta = 0$  a média  $\mu$  é igual a variância de  $Y$ ,  $\mu = Var(Y)$  (equidispersão) e, neste caso, tem-se a distribuição de Poisson.

Observando-se a figura 3.12 nota-se pelo gráfico (a) que a distribuição de Poisson generalizada é simétrica para  $\beta = 0$  e é assimétrica a esquerda para os demais valores de  $\beta$ . Pelo gráfico (b) tem-se que a assimetria é crescente e a curtose, como pode ser visto no gráfico (c), tem comportamento similar.

## 3.2 Distribuições em que $\mu = \tau(\theta)$

### 3.2.1 Distribuição em Série Logarítmica

A distribuição em série logarítmica é muito usada para representar o número de artigos adquiridos por um comprador em um determinado intervalo de tempo, em conjuntos de dados ecológicos para determinar o número de espécies e em aplicações para problemas de controle de inventário na indústria de aço.

A distribuição em série logarítmica é uma boa aproximação para a distribuição binomial negativa com o parâmetro  $m$  assumindo pequenos valores (por exemplo, valores menores

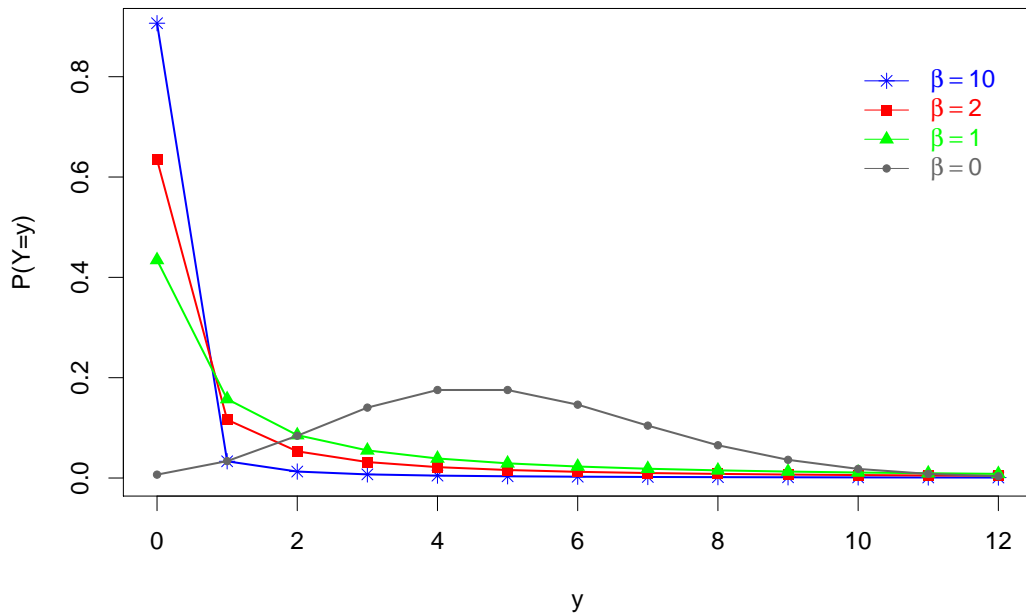
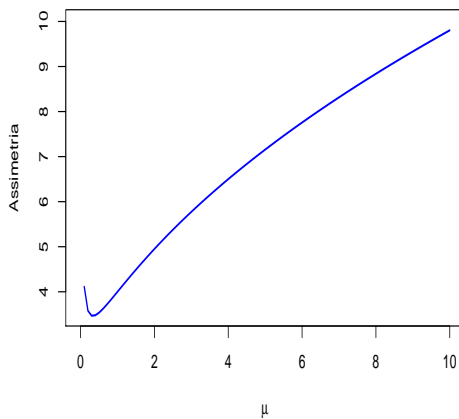
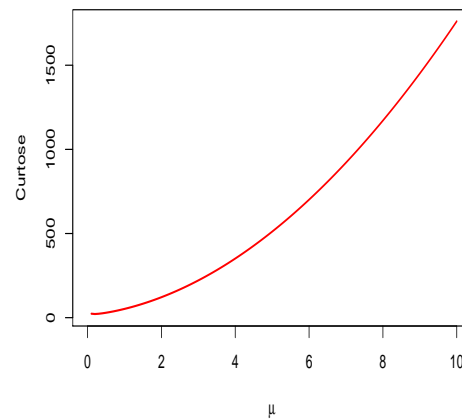
(a) Função de probabilidade ( $\mu = 5$ )(b) Assimetria ( $\beta = 1$ )(c) Curtose ( $\beta = 1$ )

Figura 3.12: Gráficos da distribuição de Poisson generalizada

do que 0, 1). A distribuição logarítmica tem a vantagem de depender de um só parâmetro,  $\theta$ , em vez de dois,  $m$  e  $\theta$ , na distribuição binomial negativa.

A variável aleatória  $Y$  tem distribuição em série logarítmica se

$$P(Y = y) = -\frac{\theta^y}{y \log(1 - \theta)},$$

sendo  $y = 1, 2, \dots$ ,  $0 < \theta < 1$ . Tem-se ainda que  $f(\theta) = -\log(1 - \theta)$  e  $g(\theta) = \theta$ . A média da distribuição em série logarítmica é  $\frac{\alpha\theta}{1 - \theta}$  e sua variância é  $Var(Y) = \frac{\alpha\theta(1 - \alpha\theta)}{(1 - \theta)^2}$ , sendo  $\alpha = [-\log(1 - \theta)]^{-1}$ .

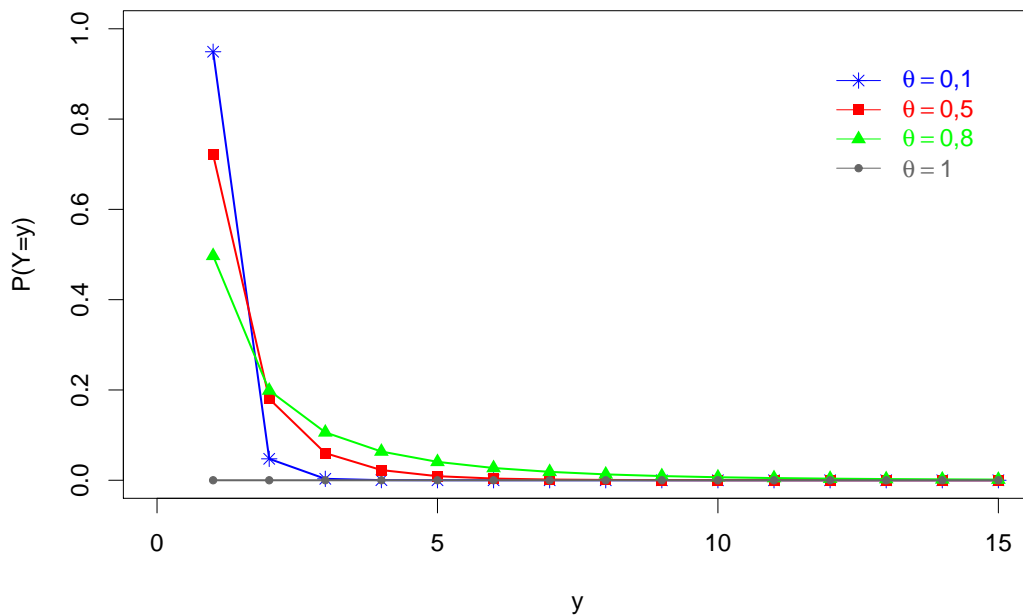


Figura 3.13: Função de probabilidade da distribuição em série logarítmica

Observando-se a figura 3.13 nota-se que a distribuição em série logarítmica é assimétrica a esquerda.

### 3.2.2 Distribuição Série Logarítmica Generalizada

A função de probabilidade da distribuição em série logarítmica generalizada (JAIN; GUPTA, 1973) é expressa por

$$P(Y = y) = \frac{\Gamma(\beta y)}{y\Gamma(y)\Gamma(\beta y - y + 1)} \frac{[\theta(1 - \theta)^{\beta-1}]^y}{-\log(1 - \theta)},$$

sendo  $y = 1, 2, \dots$ ,  $0 < \theta < 1$ ,  $\beta \geq 1$  e  $0 < \beta\theta < 1$ . Tem-se ainda que  $f(\theta) = -\log(1 - \theta)$  e  $g(\theta) = \theta(1 - \theta)^{\beta-1}$ . A média da distribuição em série logarítmica generalizada é  $\frac{\alpha\theta}{1 - \beta\theta}$

e sua variância é  $Var(Y) = -\frac{\alpha\theta(1 - \theta)^2}{(1 - \beta\theta)^3}$ , sendo  $\alpha = [-\log(1 - \theta)]^{-1}$ .

A distribuição em série logarítmica generalizada pode ser usada para modelar a distribuição de espécies animais e para representar crescimento populacional, por exemplo. Jain e Gupta (1973) usaram a distribuição em série logarítmica generalizada para modelar o número de publicações escritas por biólogos.

Observações: Para  $\beta = 1$  tem-se a distribuição em série logarítmica com  $f(\theta) = -\log(1 - \theta)$ ,  $g(\theta) = \theta$ ,  $E(Y) = \frac{\alpha\theta}{1-\theta}$  e  $Var(Y) = \frac{\alpha\theta(1-\theta)^2}{(1-\theta)^3}$ .

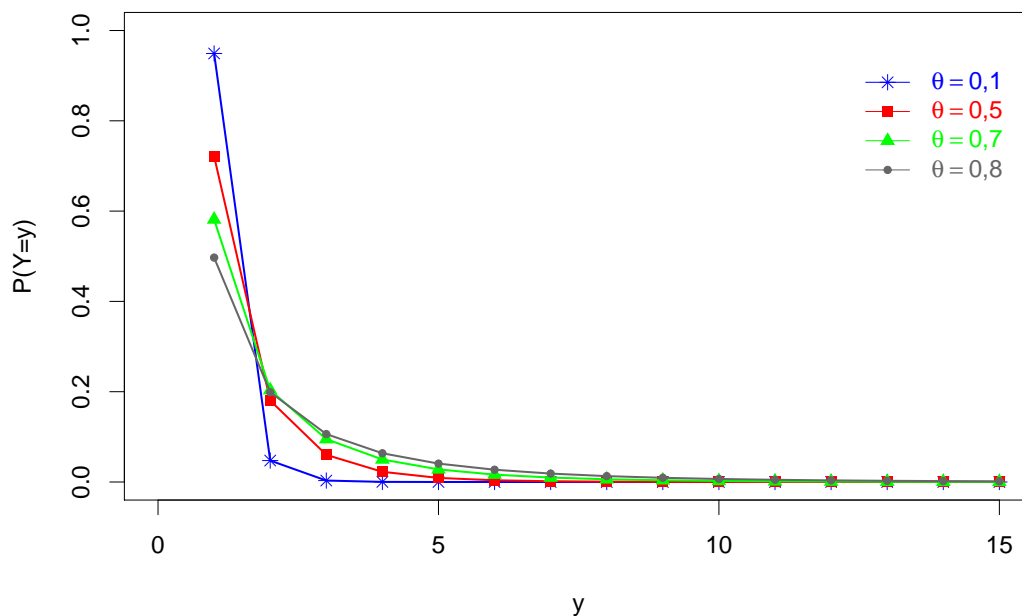


Figura 3.14: Função de probabilidade da distribuição em série logarítmica generalizada com  $\beta = 1$

Observando-se a figura 3.14 nota-se que a distribuição em série logarítmica generalizada é assimétrica a esquerda.

## 4 Distribuições em Série de Potências Modificadas Inflacionadas e Truncadas

Neste capítulo, descrevem-se algumas propriedades da classe de distribuições em série de potências modificadas inflacionadas e relacionam-se alguns casos especiais da classe MPSD inflacionada. Algumas propriedades da classe de distribuições em série de potências modificadas truncada e algumas distribuições na classe MPSD truncada também são apresentadas.

### 4.1 Distribuições em Série de Potências Modificadas Inflacionadas

#### 4.1.1 Função de Probabilidade

Uma variável aleatória discreta  $Y$  tem distribuição em série de potências modificada inflacionada no ponto  $s$  se

$$P(Y = y) = \begin{cases} \phi + (1 - \phi)a(y)g(\theta)^y/f(\theta), & y = s, \\ (1 - \phi)a(y)g(\theta)^y/f(\theta), & y \neq s, \end{cases} \quad (4.1)$$

sendo  $0 < \phi \leq 1$  e  $y \in T$ , em que  $T$  é um subconjunto dos inteiros positivos. Além disso,  $f(\theta) = \sum_y a(y)g(\theta)^y$  e  $g(\theta)$  são funções positivas, finitas e diferenciáveis e os coeficientes  $a(y)$  são não-negativos e não dependem de  $\theta$ .

Gupta et al. (1995) mostraram resultados para casos especiais em que o ponto de inflação é  $s = 0$ , enquanto Murat e Szynal (1998) apresentaram resultados quando o ponto de inflação,  $s$ , é qualquer ponto em  $T$ .

### 4.1.2 Momentos

Sejam  $\mu'_r, \mu'_{[r]}$  e  $\mu_r$ , respectivamente, os momentos ordinário, fatorial e central da classe de distribuições em série de potências modificadas, conforme o Capítulo 2, e  $m'_r, m'_{[r]}$  e  $m_r$  os momentos ordinário, fatorial e central, respectivamente, da classe MPSD inflacionada. Murat e Szynal (1998) provaram que

$$m'_r = \phi s^r + (1 - \phi) \mu'_r, \quad m'_{[r]} = \phi s^{[r]} + \phi \mu'_{[r]}$$

e

$$m_r = \phi(1 - \phi)(s - \mu'_1)^r [(1 - \phi)^{r-1} - (-\phi)^{r-1}] + (1 - \phi) \sum_{j=2}^r \binom{r}{j} [\phi(\mu'_1 - s)]^{r-j} \mu_j,$$

em que  $r \geq 1$ .

Sendo  $M_Y(t)$  a função geradora de momentos da variável aleatória  $Y$  com distribuição em série de potências modificada inflacionada no ponto  $s$  e  $M_X(t)$  a função geradora de momentos da variável aleatória  $X$  com distribuição em série de potências modificada, tem-se que

$$M_Y(t) = \phi e^{ts} + (1 - \phi)M_X(t).$$

Utilizando a fórmula de recorrência para os momentos ordinários da classe de distribuições em série de potências modificadas, tem-se que a fórmula de recorrência dos momentos ordinários para a classe MPSD inflacionada é

$$m'_{r+1} = \frac{g(\theta)}{g'(\theta)} \frac{dm'_r}{d\theta} + \frac{m'_1 m'_r}{1 - \phi} - \frac{\phi}{1 - \phi} s[m'_r + s^{r-1}(m'_1 - s)].$$

Para os momentos fatoriais tem-se

$$m'_{[r+1]} = \frac{g(\theta)}{g'(\theta)} \frac{dm'_{[r]}}{d\theta} - \left[ r - \frac{m'_1}{1 - \phi} \right] m'_{[r]} + \frac{\phi}{1 - \phi} [s^{[r]}(s - m'_1) + s m'_{[r]}].$$

A relação de recorrência dos momentos centrais é

$$m_{r+1} = \frac{g(\theta)}{g'(\theta)} \left[ \frac{dm_r}{d\theta} + r \frac{dm'_1}{d\theta} m_{r-1} \right] - \frac{\phi(s - m'_1)}{1 - \phi} m_r + \frac{\phi}{1 - \phi} (s - m'_1)^{r-1}.$$

### 4.1.3 Estimação de Máxima Verossimilhança

Fazendo  $\lambda = (1 - \phi)(1 - a(s)g(\theta)^s/f(\theta))$  em (4.1) pode-se escrever

$$P(Y = y) = \begin{cases} 1 - \lambda, & y = s, \\ \frac{\lambda a(y)[g(\theta)]^y}{f(\theta) - a(s)g(\theta)^s}, & y \neq s, \end{cases} \quad (4.2)$$

sendo  $0 < \phi \leq 1$  e  $y \in T$ , em que  $T$  é um subconjunto dos inteiros positivos.

Seja  $Y_1, \dots, Y_n$  uma amostra aleatória de (4.2) e seja  $n_i$  o número de observações iguais a  $i$ , tal que  $n = \sum_{i=0}^{\infty} n_i$ . A função de verossimilhança pode ser escrita como

$$L = \prod_{i=0}^{\infty} P(Y = i)^{n_i}.$$

O logaritmo da função de verossimilhança é

$$\log L = n_s \log(1 - \lambda) + \sum_{i=0, i \neq s}^{\infty} n_i \log \left[ \frac{\lambda a(i)g(\theta)^i}{f(\theta) - a(s)g(\theta)^s} \right].$$

Os estimadores de  $\lambda$  e  $\theta$  são calculados por

$$\hat{\lambda} = \frac{n - n_s}{n}$$

e

$$\frac{\mu(\hat{\theta})f(\hat{\theta}) - sa(s)g(\hat{\theta})^s}{f(\hat{\theta}) - a(s)g(\hat{\theta})^s} = \frac{1}{n - n_s} \sum_{i=1, i \neq s}^{\infty} in_i, \quad (4.3)$$

em que  $\mu(\theta) = \mu'_1$  é a média da classe MPSD.

O estimador  $\hat{\theta}$  pode ser obtido de (4.3) por técnicas numéricas iterativas.

Alguns casos especiais da classe MPSD inflacionada em zero,  $s = 0$ , são ilustrados na Tabela 4.1.



Tabela 4.1: Casos especiais de distribuições em série de potências modificadas inflacionadas em zero

Distribuição	$\mu(\theta)$	$Var(Y)$
Binomial Negativa Generalizada Inflacionada em Zero	$\theta(1-\phi)$	$(1-\phi)\{\theta[1+\beta\theta m^{-1}][1+(\beta-1)\theta m^{-1}]+\phi\theta^2\}$
Poisson Generalizada Inflacionada em Zero	$\theta(1-\phi)$	$(1-\phi)[\theta(1-\beta\theta)^2+\phi\theta^2]$
Série Logarítmica Generalizada Inflacionada em Zero*	$\frac{\alpha\theta(1-\phi)}{1-\beta\theta}$	$\alpha\theta(1-\phi)\left[\frac{\phi\alpha\theta(1-\beta\theta)-(1-\theta)^2}{(1-\beta\theta)^3}\right]$

\* $\alpha = [-\log(1-\theta)]^{-1}$

## 4.2 Distribuições em Série de Potências Modificadas Truncadas

### 4.2.1 Função de Probabilidade

Uma variável aleatória discreta  $Y$  tem distribuição em série de potências modificada truncada no ponto  $r$  se sua função de probabilidade é expressa como

$$P(Y = y) = \frac{P(Z = y)}{P(Z > r)},$$

sendo  $y = r, r+1, \dots$  e  $Z$  uma variável aleatória com distribuição em série de potências modificada.

### 4.2.2 Momentos

As fórmulas de recorrência para os momentos ordinários, fatoriais e centrais da classe MPSD truncada em  $r$  são idênticas às fórmulas de recorrência dadas no Capítulo 2.

Alguns casos especiais da classe MPSD truncada em zero são ilustrados na Tabela 4.2.

Tabela 4.2: Casos especiais de distribuições em série de potências modificadas truncadas em zero

Distribuição	$f(\theta)$	$g(\theta)$	$\mu(\theta)$	$Var(Y)$
Poisson Truncada em Zero	$\frac{1-e^{-\theta}}{e^{-\theta}}$	$\theta$	$\frac{\theta}{1-e^{-\theta}}$	$\left[\frac{\theta}{1-e^{-\theta}}\right] \left[\frac{1-\theta e^{-\theta}-e^{-\theta}}{1-e^{-\theta}}\right]$
Binomial Negativa Truncada em Zero	$\frac{1-\theta^m}{\theta^m}$	$1-\theta$	$\frac{m(1-\theta)}{\theta(1-\theta^m)}$	$\frac{m(1-\theta)(1-\theta^{m+1})}{[\theta(1-\theta^m)]^2}$

## 5 Inferência Bayesiana

A semente para a abordagem bayesiana a problemas de inferência foi lançada por Richard Price quando em 1763 publicou a obra póstuma do Rev. Thomas Bayes intitulada “*An Essay Towards Solving a Problem in the Doctrine of Chances*” (PAULINO et al., 2003).

A inferência estatística trata de problemas para os quais se têm quantidades de interesse, chamadas de parâmetros, associadas a algumas características populacionais ou situações práticas a respeito das quais se deseja fazer afirmações sobre elas com maior precisão possível. A abordagem bayesiana utiliza as informações fornecidas por uma amostra da população e, também, incorpora informações obtidas de experimentos passados ou experiências de especialistas para quantificar em termos de probabilidade o que é desconhecido, ou seja, os parâmetros.

A aproximação bayesiana (POLLARD, 1986) é um método que põe num fundo comum distribuição *a priori* e informação amostral especificando como essa informação *a priori* será atualizada pelos novos dados. A análise bayesiana com distribuições *a priori* não-informativas, isto é, distribuições *a priori* construídas por alguma regra formal, é muito comum quando pouca ou nenhuma informação está disponível. Uma das distribuições *a priori* não-informativas mais amplamente usada, formulada por Laplace, é a distribuição uniforme sobre o espaço paramétrico. Outra distribuição *a priori* muito citada é a proposta por Jeffreys que é dada pela raiz quadrada do determinante da matriz de informação de Fisher e que possui a propriedade da invariância.

### 5.1 Teorema de Bayes

Suponha que  $y^T = (y_1, \dots, y_n)$  é um vetor de  $n$  observações cuja distribuição de probabilidade  $\pi(y|\theta)$  depende dos  $k$  valores do parâmetro  $\theta^T = (\theta_1, \dots, \theta_k)$ . Suponha também que a distribuição de probabilidade de  $\theta$  é  $\pi(\theta)$ . Então,  $\pi(y, \theta) = \pi(y|\theta)\pi(\theta) = \pi(\theta|y)\pi(y)$ .

Dado  $y$  observado, a distribuição condicional de  $\theta$  é

$$\pi(\theta|y) = \frac{\pi(y|\theta)\pi(\theta)}{\pi(y)}. \quad (5.1)$$

Tem-se ainda que

$$\pi(y) = E\pi(y|\theta) = c^{-1} = \begin{cases} \int \pi(y|\theta)\pi(\theta) d\theta, & \theta \text{ contínuo,} \\ \sum \pi(y|\theta)\pi(\theta), & \theta \text{ discreto,} \end{cases}$$

sendo a soma ou a integral tomada sobre o domínio do parâmetro  $\theta$  e  $E[f(\theta)]$  é a esperança matemática de  $f(\theta)$  com relação à distribuição  $\pi(\theta)$ . Assim, pode-se escrever (5.1) como

$$\pi(\theta|y) = c \pi(y|\theta)\pi(\theta). \quad (5.2)$$

A expressão em (5.1), ou em (5.2), é referida como Teorema de Bayes. Nessas expressões,  $\pi(\theta)$ , a distribuição *a priori* de  $\theta$ , representa o que é conhecido a respeito do parâmetro  $\theta$  sem o conhecimento dos dados. Correspondentemente,  $\pi(\theta|y)$ , a distribuição *a posteriori* de  $\theta$  dado  $y$ , sintetiza o que é conhecido a respeito do parâmetro  $\theta$  dado o conhecimento dos dados. A quantidade  $c$  é simplesmente uma constante de “normalização” necessária para assegurar que a distribuição *a posteriori*  $\pi(\theta|y)$  quando integrada ou somada (em caso discreto) será unitária.

O Teorema de Bayes, também, pode ser representado não por uma função de  $y$ , mas de  $\theta$ , através do logaritmo da função de verossimilhança de  $\theta$  dado  $y$ , ou seja,  $\ell(\theta|y)$  e, portanto, (5.2) pode ser escrita como

$$\pi(\theta|y) = c \ell(\theta|y)\pi(\theta). \quad (5.3)$$

Portanto, o Teorema de Bayes diz que a distribuição de probabilidade de  $\theta$  *a posteriori* dado  $y$  é proporcional ao produto da distribuição *priori* de  $\theta$  e a verossimilhança de  $\theta$  dado  $y$ , isto é,

$$\text{Distribuição a posteriori} \propto \text{Verossimilhança} \times \text{Distribuição a priori}$$

Segundo Box e Tiao (1992), a função de verossimilhança  $\ell(\theta|y)$  desempenha um papel muito importante na fórmula de Bayes, pois ela é a função pela qual  $y$  modifica o conhecimento *a priori* sobre  $\theta$ . Essa função pode ser considerada como a representação da informação sobre  $\theta$  que vem dos dados.

## 5.2 Distribuições a Priori

A utilização de informação *a priori* em inferência bayesiana requer a especificação de uma distribuição *a priori* para a quantidade de interesse  $\theta$ . Essa distribuição deve representar (probabilisticamente) o conhecimento que se tem sobre  $\theta$  antes da realização do experimento. A seguir, serão discutidas diferentes formas de especificação da distribuição *a priori*.

### 5.2.1 Distribuições a Priori Conjugadas

A partir do conhecimento que se tem sobre  $\theta$ , pode-se definir uma família paramétrica de densidades. Nesse caso, a distribuição *a priori* é representada por uma forma funcional, cujos parâmetros devem ser especificados de acordo com esse conhecimento. Esses parâmetros indexadores da família de distribuições *a priori* são chamados de hiperparâmetros para distingui-los dos parâmetros de interesse  $\theta$ .

Essa abordagem em geral facilita a análise e o caso mais importante é o de distribuição *a priori* conjugada. A idéia é que as distribuições *a priori* e *a posteriori* pertençam à mesma classe de distribuições e assim a atualização do conhecimento que se tem de  $\theta$  envolve apenas uma mudança nos hiperparâmetros. Nesse caso, o aspecto sequencial do método bayesiano pode ser explorado definindo-se apenas a regra de atualização dos hiperparâmetros já que as distribuições permanecem as mesmas.

### 5.2.2 Distribuições a Priori não Informativas

O subjetivismo se tornou a fundação filosófica dominante para a inferência bayesiana. A maioria das análises bayesianas são executadas com as denominadas distribuições *a priori* não-informativas, ou seja, construídas por alguma regra. O aumento do poder computacional e o surgimento de métodos computacionais, tal como, o método de Monte Carlo com Cadeias de Markov (CHRISTIAN; CASELLA, 2004), ampliaram o interesse em técnicas bayesianas, mas novas aplicações continuam a levantar a questão da escolha da distribuição *a priori* (KASS; WASSERMAN, 1996).

Alguns métodos utilizados para identificar qual distribuição *a priori* deve ser aplicada aos parâmetros de uma distribuição, sem que haja um conhecimento prévio sobre a mesma, ou que a informação *a priori* seja vaga, e que gere os resultados esperados quando da realização de um experimento, serão apresentados a seguir.

### Método de Jeffreys

O conceito de selecionar uma distribuição *a priori* através de convenção, como um “padrão de referência” análogo a escolher um padrão em outras questões científicas é devido a Jeffreys. As regras que surgiram depois para seleção de distribuições *a priori* são modificações da regra de Jeffreys (KASS; WASSERMAN, 1996).

Jeffreys subscreveu o “princípio da razão insuficiente”. Esse princípio exige que a distribuição sobre o espaço paramétrico finito seja uniforme, a menos que haja alguma razão para considerar um evento mais provável do que o outro, ou seja, atribua probabilidades iguais a cada um dos possíveis valores.

A classe de distribuições *a priori* não informativas proposta por Jeffreys é invariante a transformações um a um, embora, em geral, seja imprópria e será definida a seguir.

A distribuição *a priori* de um parâmetro  $\theta$  é, aproximadamente, não-informativa se for tomada como proporcional à raiz quadrada da informação de Fisher (BOX; TIAO, 1992). Portanto, a priori não informativa de Jeffreys, ou a “regra geral” de Jeffreys, tem função de densidade dada por

$$\pi_{\theta}(\theta) \propto \det[\mathbf{I}(\theta)]^{1/2}, \quad (5.4)$$

em que  $\mathbf{I}(\theta) = E \left[ \frac{\partial^2 \ell(\theta|y)}{\partial \theta \partial \theta^T} \right]$  é a matriz de informação esperada de Fisher.

### Verossimilhança de Dados Transladados

Box e Tiao (1992) propuseram a noção de “verossimilhança de dados transladados” para motivar o uso de distribuição *a priori* uniforme. Se  $\phi(\theta)$  é uma transformação um a um de  $\theta$ , pode-se dizer que uma distribuição a priori para  $\theta$  é não-informativa se, em termos de  $\phi$ , a curva da verossimilhança apresenta os dados transladados, ou seja, os dados servem apenas para mudar a locação da função de verossimilhança  $L_y(\phi)$ .

E ainda, de acordo com eles, a função de verossimilhança de dados transladados é

$$L_y(\phi) \approx \tau[\phi - t(y)], \quad (5.5)$$

para alguns valores reais das funções  $\tau(\cdot)$  e  $t(\cdot)$ , com  $\tau(\cdot)$  não dependendo de  $y$ , para  $y$  um vetor de observações e  $L_y(\cdot)$  a função de verossimilhança no espaço paramétrico real unidimensional  $\Phi$ . O uso de uma distribuição *a priori* uniforme para duas amostras diferentes  $y$  e  $y^*$  produzirão distribuições *a posteriori* que diferem apenas no que diz respeito à locação (KASS; WASSERMAN, 1996).

Uma transformação para que a função de verossimilhança seja expressa, exatamente, como em (5.5), geralmente, não está disponível. No entanto, para amostras de tamanho moderado, devido à insensibilidade da distribuição *a posteriori* a pequenas alterações na distribuição *a priori*, é necessária uma transformação em  $\phi(\theta)$ , a fim de que a função de verossimilhança seja aproximada aos dados transladados. Ou seja, a função de verossimilhança para  $\phi$  é quase independente de  $y$ , exceto por sua locação.

### Método de Berger-Bernardo

Bernardo (1979) sugeriu um método para a construção de distribuição *a priori* que envolviam duas inovações. A primeira foi definir o conceito de informação faltante (*missing*) e a segunda foi desenvolver um processo gradual (*stepwise*) para a manipulação de parâmetros de incômodo. Existe uma série de trabalhos, principalmente por Berger e Bernardo, refinando o método e aplicando-o a vários problemas e por esse motivo é conhecido como o método de Berger-Bernardo.

Quando não existem parâmetros de incômodo e certas condições de regularidade são satisfeitas, a distribuição *a priori* de Bernardo será igual a (5.4). Quando existe uma partição no espaço paramétrico em parâmetros de interesse e parâmetros de incômodo, esse método muitas vezes produz distribuições *a priori* distintas de (5.4) (KASS; WASSERMAN, 1996).

#### Informação Faltante (*Missing*)

Seja  $Y_1^n = (Y_1, \dots, Y_n)$  formado por  $n$  variáveis aleatórias *iid* e seja  $K_n[\pi(\theta|y_1^n), \pi(\theta)]$  a distância de Kullback-Leibler entre as funções densidade *a posteriori* e *a priori*,

$$K_n[\pi(\theta|y_1^n), \pi(\theta)] = \int \pi(\theta|y_1^n) \log \left( \frac{\pi(\theta|y_1^n)}{\pi(\theta)} \right) d\theta. \quad (5.6)$$

Seja  $K_n^\pi = E[K_n(\pi(\theta|y_1^n), \pi(\theta))]$  o ganho esperado de informação, em que a esperança é com relação à densidade marginal  $m(y_1^n) = \int p(y_1^n|\theta)\pi(\theta)d\theta$ . A idéia de Bernardo foi pensar  $K_n^\pi$  para  $n$  grande como uma medida da informação faltante no experimento, uma idéia que tem as suas raízes nos trabalhos de Good (1960, 1966) e Lindley (1956).

Bernardo sugeriu encontrar a priori que maximize  $K_\infty^\pi = \lim_{n \rightarrow \infty} K_n^\pi$  e chamou o resultado de priori de referência. Mas este termo já tinha sido utilizado por Box e Tiao para a regra de Jeffreys.

#### Parâmetros de Incômodo (*Nuisance*)

Suponha  $\theta = (\omega, \nu)$  em que  $\omega$  é o parâmetro de interesse e  $\nu$  é o parâmetro de incômodo. Bernardo sugeriu uma metodologia diferente nesse caso e ignorando alguns problemas técnicos, primeiramente define-se  $\pi(\nu|\omega)$  pela distribuição *a priori* de Berger-Bernardo considerando  $\omega$  fixo. Depois encontra-se o modelo  $\pi(y|\omega) = \int \pi(x|\omega, \nu)\pi(\nu|\omega)d\nu$  e só, então, encontra-se  $\pi(\omega)$ , como a distribuição *a priori* de Berger-Bernardo, baseado nesse modelo. A distribuição *a priori* recomendada é então  $\pi(\omega)\pi(\nu|\omega)$ .

O método de Berger-Bernardo já foi aplicado a muitos exemplos, incluindo regressão exponencial (YE; BERGER, 1991), modelos multinomiais (BERGER; BERNARDO, 1992), modelos AR(1) (BERGER; YANG, 1994) e problemas que envolvem produto de médias de normais (BERGER; BERNARDO, 1989; SUN; YE, 1995).

### 5.3 Distribuição a Posteriori para o Modelo Binomial Negativo Generalizado

Seja uma amostra em que  $Y$  tem distribuição binomial negativa generalizada com parâmetros  $\mu$ ,  $\phi$  e  $m$  (ver Seção 3.1.3). A função de verossimilhança  $L = L(\theta)$  é expressa como

$$L = \left\{ m \left[ \frac{m + (\phi - 1)\mu}{m + \phi\mu} \right]^m \right\}^n \left\{ \left( \frac{\mu}{m + \phi\mu} \right) \left[ \frac{m + (\phi - 1)\mu}{m + \phi\mu} \right]^{\phi - 1} \right\}^{n\bar{y}} \prod_{i=1}^n \frac{\Gamma(m + \phi y_i)}{y_i! \Gamma(m + \phi y_i - y_i + 1)}.$$

em que  $\theta = (\mu, \phi)^T$ . Para utilizar a abordagem bayesiana precisa-se definir as distribuições *a priori* para os parâmetros  $\mu$  e  $\phi$ , admitindo  $m$  conhecido, e, em seguida, iremos propor uma distribuição *a posteriori* para os parâmetros utilizando o método de distribuição *a priori* conjugada.

#### 5.3.1 Distribuição a Priori para os Parâmetros

Considerando a independência entre os parâmetros, a distribuição *a priori* possui a seguinte estrutura

$$\pi(\theta) = \pi(\mu)\pi(\phi).$$

em que  $\theta = (\mu, \phi)^T$ .

Nesse trabalho, consideram-se as seguintes distribuições *a priori* para os parâmetros.



**Distribuição a priori para  $\mu$** 

Utilizou-se a distribuição gama para  $\mu$ , ou seja,  $\mu \sim \text{Gama}(\alpha, \delta)$

$$\pi(\mu) = \frac{\delta^\alpha}{\Gamma(\alpha)} \mu^{\alpha-1} e^{-\delta\mu},$$

em que  $\mu, \alpha, \delta > 0$ , pois  $\mu$  deve ser um número real positivo.

**Distribuição a priori para  $\phi$** 

Utilizou-se a distribuição exponencial para  $\phi$ , ou seja,  $\phi \sim \text{Exponencial}(\iota)$

$$\pi(\phi) = \iota e^{-\iota\phi}.$$

em que  $\phi, \iota > 0$ , pois  $\phi$  deve ser um número real positivo.

A distribuição *a priori* conjunta será

$$\pi(\theta) = \frac{\iota \delta^\alpha}{\Gamma(\alpha)} \mu^{\alpha-1} e^{-(\delta\mu + \iota\phi)}.$$

A distribuição *a posteriori*, admitindo independência de  $\mu$  e  $\phi$ , será

$$\begin{aligned} \pi(\theta|y, m) \propto & \mu^{\alpha-1} e^{-(\delta\mu + \iota\phi)} \left\{ n\bar{y} \left[ \log\left(\frac{\mu}{m + \phi\mu}\right) + (\phi - 1) \log\left(1 - \frac{\mu}{m + \phi\mu}\right) \right] \right. \\ & \left. + nm \log\left(1 - \frac{\mu}{m + \phi\mu}\right) \right\}. \end{aligned} \quad (5.7)$$

A distribuição *a posteriori* (5.7) não possui uma expressão com forma fechada. Assim, as estimativas dos parâmetros devem ser obtidas através de métodos aproximados e para isso utiliza-se o método de Monte Carlo com Cadeias de Markov (MCMC).

## 6 Modelos Não-Lineares Generalizados em Série de Potências Modificadas

Neste capítulo, uma breve revisão dos modelos não-lineares generalizados é apresentada e, em seguida, descreve-se a classe de distribuições em série de potências modificadas neste contexto.

### 6.1 Modelos Lineares Generalizados

Nelder e Wedderburn (1972) mostraram que uma série de técnicas estatísticas, comumente estudadas separadamente, podem ser formuladas, de uma maneira unificada, como uma classe de modelos de regressão. A essa teoria unificadora de modelagem estatística, uma extensão dos modelos clássicos de regressão, deram o nome de modelos lineares generalizados (MLG).

Os MLG permitem ampliar as hipóteses básicas dos modelos normais lineares, tais como, a variável resposta, componente aleatório do modelo, pode ter qualquer distribuição da família exponencial, deixando de ser obrigatoriamente uma distribuição normal. Além disso, a relação entre o valor médio da variável resposta e a combinação linear das variáveis explicativas, componente sistemático, pode ser qualquer função monótona e diferenciável, ao contrário dos modelos de regressão clássicos que só comportam a função identidade.

Objetiva-se utilizar a classe de distribuições em série de potências modificadas no contexto dos MLG. Então, seja  $y = (y_1, \dots, y_n)^T$  um conjunto de  $n$  observações independentes que satisfazem

$$P(Y = y) = \frac{a(y, \phi)g(\mu, \phi)^y}{f(\mu, \phi)}, \quad y \in T, \quad (6.1)$$

em que  $T$  é um subconjunto do conjunto dos inteiros não-negativos,  $a(y, \phi) > 0$ ,  $g(\mu, \phi)$  e  $f(\mu, \phi)$  são funções analíticas positivas, finitas e diferenciáveis e  $\phi > 0$ . E observa-se um conjunto de variáveis explanatórias em  $X$ . O componente sistemático para o vetor de médias  $\mu = E(Y)$  é expresso como

$$h(\mu) = \eta = \eta(X, \beta), \quad (6.2)$$

em que  $h(\cdot)$  é uma função monótona e diferenciável,  $\eta$  é um preditor linear se  $\eta(X, \beta) = X\beta$ , caso contrário diz-se que  $\eta$  é um preditor não-linear,  $X$  é a matriz modelo de dimensão  $n \times p$  de posto completo  $p < n$  e  $\beta = (\beta_1, \dots, \beta_p)^T$  é o conjunto de parâmetros desconhecidos a serem estimados. Os modelos não-lineares generalizados em série de potências modificadas são definidos por (6.1) e (6.2).

Os MLG aplicados a dados de contagem são restritos às distribuições de Poisson, binomial e binomial negativa. Cordeiro et al. (2008) formularam uma nova classe de modelos, a qual chamaram modelos não-lineares generalizados em série de potências (representada pela sigla PSGNLMS), que incluem como casos especiais outros modelos discretos, tais como, as distribuições em série de potências modificadas, e podem ser utilizados para analisar dados de contagens.

## 6.2 Ajuste do Modelo

Observa-se a matriz modelo  $X$  e assume-se que as observações  $y = (y_1, \dots, y_n)^T$  satisfazem (6.1) e (6.2). O logaritmo da função de verossimilhança  $\ell = \ell(\beta, \phi)$  para os parâmetros do modelo é (CORDEIRO et al., 2008)

$$\ell(\beta) = \sum_{i=1}^n \log\{a(y_i, \phi)\} + \sum_{i=1}^n [y_i \log\{g(\mu_i, \phi)\} - \log\{f(\mu_i, \phi)\}]. \quad (6.3)$$

Pode-se maximizar (6.3) incondicionalmente, por exemplo, usando algum programa estatístico padrão ou o pacote GAMLSS (STASINOPOULOS; RIGBY, 2007) no R (R Development Core Team, 2008). Alternativamente, pode-se assumir  $\phi$  fixo e obter estimativas para  $\beta$  através da função de log-verossimilhança condicional. Sejam  $\hat{\beta}^{(\phi)}$ ,  $\hat{\eta}^{(\phi)} = \phi \eta(X, \hat{\beta}^{(\phi)})$  e  $\hat{\mu}^{(\phi)} = h^{-1}(\hat{\eta}^{(\phi)})$  os estimadores de máxima verossimilhança condicional de  $\beta$ ,  $\eta$  e  $\mu$ , respectivamente, para  $\phi$  dado. O estimador condicional  $\hat{\beta}^{(\phi)}$  pode ser obtido iterativamente

através de mínimos quadrados ponderados (CORDEIRO et al., 2008).

Condicionalmente em  $\phi$ , a função escore para o vetor de parâmetros  $\beta$  é expressa como

$$U_{\beta_r} = \frac{\partial \ell}{\partial \beta_r} = \sum_{i=1}^n \frac{d\ell_i}{d\mu_i} \frac{d\mu_i}{d\eta_i} \frac{\partial \eta_i}{\partial \beta_r} = \sum_{i=1}^n \left[ y_i \frac{g'_i}{g_i} - \frac{f'_i}{f_i} \right] \frac{1}{h'_i} \tilde{x}_{ir},$$

sendo  $f_i = f(\mu_i, \phi)$ ,  $f'_i = \frac{df_i}{d\mu_i}$ , usa-se notação similar para  $g$ ,  $\frac{1}{h'_i} = \frac{d\mu}{d\eta}$  e  $\tilde{x}_{ir} = \frac{\partial \eta_i}{\partial \beta_r}$ .

Matricialmente tem-se

$$U_{\beta} = \tilde{X}^T (Ty - Q1),$$

em que  $T = \text{diag}\{t_i\}$  e  $Q = \text{diag}\{q_i\}$  são matrizes diagonais  $n \times n$  com elementos definidos por  $t_i = \frac{g'_i}{g_i h'_i}$  e  $q_i = \frac{f'_i}{f_i h'_i}$  e  $1$  é um vetor  $n \times 1$  de uns. A matriz de informação para  $\beta$  é  $K_{\beta} = \tilde{X}^T W \tilde{X}$ , em que  $W = \text{diag}\{w_i\}$  é uma matriz diagonal  $n \times n$  de pesos definidos por

$$w_i = \left( q'_i - \frac{f'_i g'_i}{f_i g_i} t'_i \right) \frac{1}{h'_i}.$$

## 6.3 Inferência

Após uma função de regressão ter sido especificada, diferentes modelos de regressão, tais como Poisson generalizada ou binomial negativa generalizada podem ser aplicados a conjuntos de dados de contagens. A distribuição em (6.1) pertence à família exponencial uniparamétrica quando  $\phi$  é conhecido e os parâmetros não-lineares seguem a abordagem do MLG. O estimador de máxima verossimilhança (EMV)  $\hat{\beta}$  de  $\beta$  é assintoticamente normal com média  $\beta$  e matriz de covariâncias consistentemente estimada por  $(\tilde{X}^T W \tilde{X})^{-1}$ .

O desvio do modelo sob investigação tem a mesma forma do desvio no contexto de MLG tradicional e é função apenas dos dados e dos EMV  $\hat{\beta}$  obtidos com os dados. Portanto, o desvio  $D$  do modelo não-linear modificado em série de potências é expresso por

$$D = 2 \sum_{i=1}^n D_i(y_i, \hat{\mu}_i),$$

em que o desvio unitário para a  $i$ -ésima observação é

$$D_i(y_i, \mu_i) = y_i \log \left\{ \frac{g(y_i)}{g(\mu_i)} \right\} + \log \left\{ \frac{f(\mu_i)}{f(y_i)} \right\}.$$

Na Tabela 6.1 estão relacionados os desvios unitários para alguns modelos discretos para dados de contagem (CORDEIRO et al., 2008).

Tabela 6.1: Desvios unitários para alguns modelos em série de potências modificadas

Distribuição	Desvio
Binomial	$2 \left[ y \log \left( \frac{y}{\mu} \right) + (m - y) \log \left\{ \frac{m-y}{m-\mu} \right\} \right]$
Binomial Negativa	$2 \left[ y \log \left( \frac{y}{\mu} \right) + (y + m) \log \left\{ \frac{\mu+m}{y+m} \right\} \right]$
Binomial Negativa Generalizada	$2 \left[ y \log \left\{ \frac{y}{\mu} \left( \frac{m+\beta\mu}{m+\beta y} \right)^\beta \left( \frac{m+\beta y-y}{m+\beta\mu-\mu} \right)^{\beta-1} \right\} + m \log \left\{ \frac{(m+\beta\mu)(m+\beta y-y)}{(m+\beta y)(m+\beta\mu-\mu)} \right\} \right]$
Borel	$2 \left[ \frac{\mu-y}{\mu} + (y-1) \log \left\{ \frac{\mu(y-1)}{y(\mu-1)} \right\} \right]$
Borel-Tanner	$2 \left[ m \left( 1 - \frac{y}{\mu} \right) + (y-m) \log \left\{ \frac{\mu(y-m)}{y(\mu-m)} \right\} \right]$
Consul	$2 \left[ y \log \left\{ \left( \frac{y-1}{\mu-1} \right) \left( \frac{\mu}{y} \right)^\beta \left( \frac{\beta y-y+1}{\beta\mu-\mu+1} \right)^{\beta-1} \right\} + \log \left\{ \left( \frac{\mu-1}{y-1} \right) \left( \frac{\beta y-y+1}{\beta\mu-\mu+1} \right) \right\} \right]$
Delta-Binomial	$2 \left[ y \log \left\{ \left( \frac{y-m}{\mu-m} \right) \left( \frac{\mu}{y} \right)^\beta \left( \frac{\beta y-y+m}{\beta\mu-\mu+m} \right)^{\beta-1} \right\} + m \log \left\{ \left( \frac{\mu-m}{y-m} \right) \left( \frac{\beta y-y+m}{\beta\mu-\mu+m} \right) \right\} \right]$
Geeta	$2 \left[ y \log \left\{ \left( \frac{y-1}{\mu-1} \right) \left( \frac{y}{\mu} \right)^{\beta-1} \left( \frac{\beta\mu-1}{\beta y-1} \right)^\beta \right\} + \log \left\{ \frac{(\mu-1)(\beta y-1)}{(y-1)(\beta\mu-1)} \right\} \right]$
Haight	$2y \left[ \log \left\{ \frac{y}{\mu} \left( \frac{y-1}{\mu-1} \right) \left( \frac{2\mu-1}{2y-1} \right)^2 \right\} + \log \left\{ \left( \frac{2y-1}{2\mu-1} \right) \left( \frac{\mu-1}{y-1} \right) \right\} \right]$
“Lost Game”	$2 \left[ y \log \left\{ \frac{y(m-y)(m-2\mu)^2}{\mu(m-\mu)(m-2y)^2} \right\} + m \log \left\{ \frac{(m-\mu)(m-2y)}{(m-y)(m-2\mu)} \right\} \right]$
Poisson	$2 \left[ y \log \left( \frac{y}{\mu} \right) - (y - \mu) \right]$
Poisson Generalizada	$2 \left[ y \log \left\{ \frac{y(1+\beta\mu)}{\mu(1+\beta y)} \right\} - \frac{(y-\mu)}{(1+\beta\mu)} \right]$

## 7 Aplicações

As distribuições em série de potências modificadas no contexto de modelos não-lineares generalizados, foram estudadas, recentemente, por Cordeiro et al. (2008) e neste capítulo serão apresentadas algumas aplicações utilizando esses modelos. Foram utilizados os programas SAS 9.1, R 2.8.0 e WinBugs 1.4 para efetuar as análises.

### 7.1 Exemplo 1

#### 7.1.1 Abordagem Clássica

Os dados utilizados nesta aplicação referem-se a um experimento em que três inseticidas (DDT,  $\gamma$ -BHC, DDT+ $\gamma$ -BHC) são aplicados a grupos de cerca de cinquenta besouros de farinha, *Tribolium Castaneum*, em seis dosagens distintas ( $mg/10cm^2$ ) e verifica-se o número de besouros mortos para cada dose aplicada (COLLETT, 1991; PAULA, 2004).

Tabela 7.1: *Mortalidade de besouros segundo as doses de três inseticidas.*

Inseticida	Dose					
	2,00	2,64	3,48	4,59	6,06	8,00
DDT	3/50	5/49	19/47	19/50	24/49	35/50
$\gamma$ -BHC	2/50	14/49	20/50	27/50	41/50	40/50
DDT+ $\gamma$ -BHC	28/50	37/50	46/50	48/50	48/50	50/50

Na Tabela 7.1 o numerador da taxa de mortalidade representa o número de besouros mortos e o denominador é o total de besouros expostos.

Os seguintes modelos foram utilizados:

Modelo Linear:  $\eta_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i}$ ;

Modelo Não-linear:  $\eta_i = x_{1i}^{\beta_1} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i}$ .

Em que  $x_1$  é a dose do inseticida e  $x_2, x_3$  e  $x_4$  identificam os inseticidas DDT,  $\gamma$ -BHC e

DDT+ $\gamma$ -BHC, respectivamente.

A análise dos dados deste experimento foi realizada através da PROC NLP do SAS 9.1. Utilizou-se os modelos de regressão de Poisson, binomial negativo, binomial negativo generalizado ( $m = 50$ ), Consul e delta-binomial ( $m = 0,5$ ) com função de ligação identidade. E, como pode ser visto na Tabela 7.2, o modelo binomial negativo generalizado não-linear seria o mais adequado para ajustar esses dados, pois apresentou o menor desvio e, também, o menor *AIC*.

Tabela 7.2: *Estatísticas para verificar a qualidade do ajuste dos dados dos insetos.*

Modelo	<i>AIC</i>		Desvio	
	Linear	Não linear	Linear	Não linear
Bin. Neg. Generalizada	13,53	17,77	3,94	3,59
Poisson	150,52	144,50	54,39	48,38
Consul	156,70	156,25	4,88	4,43
Binomial Negativa	158,72	158,35	3,71	3,34

### 7.1.2 Abordagem Bayesiana

Para estimar os parâmetros da posteriori dada em (5.7), que não possui forma fechada, utilizou-se o programa WinBugs 1.4, que utiliza o algoritmo Metropolis-Hastings para obtenção de uma amostra da respectiva distribuição e, através dela, as estimativas dadas pela média da amostra. Foi gerada uma cadeia com 10.000 valores para cada parâmetro, desprezando os primeiros 5.000 valores e, então, formando uma amostra final de 5.000 valores.

A convergência das estimativas dos parâmetros foi monitorada através da análise gráfica dos valores gerados (ver figura 7.1).

Na Tabela 7.3 tem-se a estimativa de cada parâmetro (média), o desvio padrão da estimativa, a mediana da amostra gerada e os percentis 2,5% e 97,5%.

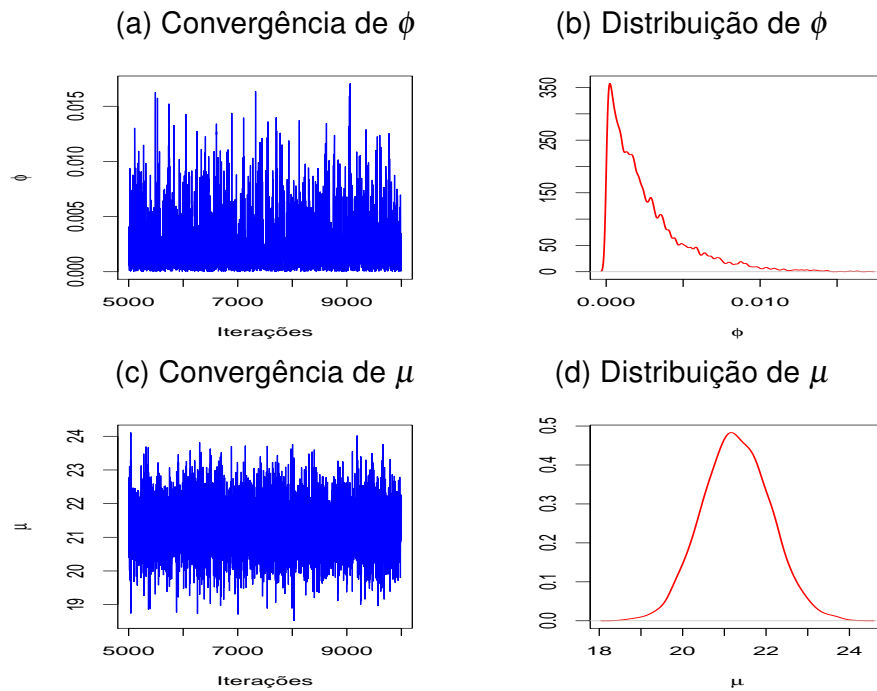


Figura 7.1: Gráficos dos parâmetros da binomial negativa generalizada

Tabela 7.3: Estimativas dos parâmetros do modelo binomial negativo generalizado utilizando os dados dos insetos.

Parâmetro	Média	Desvio padrão	2,5%	Mediana	97,5%
$\phi$	0,0025	0,0025	0,0661	0,0018	0,0092
$\mu$	21,2900	0,8095	19,7500	21,2700	22,9100

## 7.2 Exemplo 2

### 7.2.1 Abordagem Clássica

Os dados da Tabela 7.4 são referentes à distribuição de duas espécies de lagarto (grahani - G e opalinus - O) segundo quatro fatores: (i) período do dia (manhã, meio-dia, tarde), (ii) comprimento da madeira (curta, comprida), (iii) largura da madeira (estreita, larga) e (iv) local de ocupação (claro, escuro) (NELDER; WEDDERBURN, 1972; PAULA, 2004).

Para o conjunto de dados dos lagartos as análises foram feitas no programa R 2.8.0 e o modelo utilizado foi

$$\eta_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i},$$



Tabela 7.4: Distribuição de duas espécies de lagartos segundo quatro fatores.

Ocupação	Madeira		Período								
			Manhã			Meio-dia			Tarde		
	Comp.	Larg.	G	O	Total	G	O	Total	G	O	Total
Claro	$\leq 2$	$< 5$	20	2	22	8	1	9	4	4	8
		$\geq 5$	13	0	13	8	0	8	12	0	12
	$> 2$	$< 5$	8	3	11	4	1	5	5	3	8
		$\geq 5$	6	0	6	0	0	0	1	1	2
Escuro	$\leq 2$	$< 5$	34	11	45	69	20	89	18	10	28
		$\geq 5$	31	5	36	55	4	59	13	3	16
	$> 2$	$< 5$	17	15	32	60	32	92	8	8	16
		$\geq 5$	12	1	13	21	5	26	4	4	8

em que  $x_1$  representa o período do dia,  $x_2$  o comprimento da madeira,  $x_3$  a largura da madeira e  $x_4$  o local de ocupação.

Utilizou-se os modelos de regressão de Poisson, binomial negativo, binomial negativo generalizado ( $m = 2$ ), Consul e Poisson generalizado. E, como pode ser visto na Tabela 7.5, o modelo binomial negativo seria o mais adequado para ajustar esses dados, pois apresentou o menor *AIC*.

Tabela 7.5: Estatísticas para verificar a qualidade do ajuste na abordagem clássica dos dados dos lagartos.

Modelo	-2 log(L)	<i>AIC</i>	<i>BIC</i>
Bin. Negativo	155,0	167,0	173,8
Bin. Neg. Gen.	159,2	171,2	178,0
Consul	164,7	176,7	183,5
Poisson Gen.	185,5	197,5	204,4
Poisson	236,9	246,9	252,6

## 7.2.2 Abordagem Bayesiana

Para este conjunto de dados, também, foi realizada a análise bayesiana através do programa WinBugs 1.4. O modelo binomial negativo foi utilizado por ter sido o modelo selecionado pela análise clássica e também utilizou-se o modelo de Poisson por ser o mais utilizado para dados de contagens.

Como pode ser visto na Tabela 7.6, assim como na análise clássica, o modelo binomial negativo teve menor *AIC* que o modelo de Poisson. E, também, o *AIC* na abordagem

Tabela 7.6: *Estatísticas para verificar a qualidade do ajuste na abordagem bayesiana dos dados dos lagartos.*

Modelo	-2 log(L)	AIC	BIC
Binomial Negativo	130,1	138,1	142,6
Poisson	245,5	253,5	258,1

bayesiana foi menor que na abordagem clássica para o modelo binomial negativo, portanto a abordagem bayesiana mostrou-se melhor para este conjunto de dados.

## 8 Conclusão

A classe de distribuições em série de potências modificadas (MPSD) abrange inúmeras distribuições discretas, tais como, binomial, binomial negativa, Poisson, série logarítmica, entre outras. Existem artigos e livros que descrevem as propriedades desta classe de distribuições, bem como inúmeras aplicações. Porém, apenas recentemente Cordeiro et al. (2008) introduziram esta classe no contexto de modelos de regressão não-linear e com esta nova classe de modelos, a qual chamaram modelos não-lineares generalizados em série de potências (PSGNLMs), a análise de dados de contagem não fica restrita às distribuições de Poisson, binomial e binomial negativa, pode-se utilizar outros modelos discretos, tais como, binomial negativo generalizado, Borel, Borel-Tanner, Consul, delta-binomial, Haight ou Poisson generalizado.

Neste trabalho, fez-se uma revisão da classe MPSD, relacionando as principais propriedades. Aplicações no contexto de regressão foram realizadas a conjunto de dados reais. Também, fez-se uma abordagem bayesiana para o modelo binomial negativo generalizado, em que uma posteriori foi proposta, porém, como esta não obteve forma fechada, as estimativas dos parâmetros foram encontradas utilizando um dos bancos de dados analisados.

Para o conjunto de dados dos lagartos foi feita as duas abordagens, clássica e bayesiana, para os modelos de Poisson e binomial negativo. O modelo binomial negativo mostrou-se melhor nas duas abordagens, sendo que a abordagem bayesiana mostrou-se melhor para este conjunto de dados.

Uma sequência natural do trabalho seria aprimorar a abordagem bayesiana para os modelos não-lineares generalizados em série de potências.

## Referências Bibliográficas

- BERGER, J.; BERNARDO, J. Estimating a product of means: bayesian analysis with reference priors. **Journat of the American Statistical Association**, v. 84, p. 200 – 207, 1989.
- BERGER, J.; BERNARDO, J. Ordered group reference priors with application to the multinomial problem. **Biometrika**, v. 25, p. 25 – 37, 1992.
- BERGER, J. O.; YANG, R. Noninformative priors and bayesian testing for the ar(1) model. **Econometric Theory**, v. 10, p. 461 – 482, 1994.
- BERNARDO, J. M. Reference posterior distributions for bayesian inference. **Journal of the Royal Statistjcai Society, Serie B**, v. 41, p. 113 –147, 1979.
- BOREL, E. Sur l'emploi du théorème de bernoulli pour faciliter le calcul d'un infinité de coefficients. **Comptes Rendus, Académie des Sciences, Paris, Series A**, v. 214, p. 452–456, 1942.
- BOX, G. E. P.; TIAO, G. C. **Bayesian Inference in Statistical Analysis**. New York: Wiley Classics Library, John Wiley & Sons, 1992.
- CHRISTIAN, P. R.; CASELLA, G. **Monte Carlo statistical methods**. Second. New York: Springer, 2004.
- COLLETT, D. **Modelling Binary Data**. London: Chapman and Hall, 1991.
- CONSUL, P. C. On the differences of two generalized negative binomial variates. **Communications Statistics, Theory and Methods**, v. 18, p. 673–690, 1989.
- CONSUL, P. C. Geeta distribution and its properties. **Communications Statistics, Theory and Methods**, v. 19 (8), p. 3051–3068, 1990a.
- CONSUL, P. C. New class of location-parameter discrete probability distributions and their characterizations. **Communications Statistics, Theory and Methods**, v. 19 (12), p. 4653–4666, 1990b.
- CONSUL, P. C.; FAMOYE, F. On the unimodality of the generalized negative binomial distribution. **Statistica Neerlandica**, v. 40, p. 141–144, 1986.
- CONSUL, P. C.; FAMOYE, F. On the generalized negative binomial distribution. **Communications Statistics, Theory and Methods**, v. 24, p. 459–472, 1995.
- CONSUL, P. C.; FAMOYE, F. **Lagrangian Probability Distributions**. Boston: Birkhäuser, 2006.
- CONSUL, P. C.; JAIN, G. C. A generalization of the poisson distribution. **Technometrics**, v. 15, n. 4, p. 791–799, 1973a.

- CONSUL, P. C.; JAIN, G. C. On some interesting properties of the generalized poisson distribution. **Biometrische Zeitschrift**, v. 15, p. 495–500, 1973b.
- CONSUL, P. C.; SHENTON, L. R. Some interesting properties of lagrangian distributions. **Communications Statistics, Theory and Methods**, v. 2 (3), p. 263–272, 1973c.
- CORDEIRO, G. M.; ANDRADE, M. G. de; CASTRO, M. de. Power series generalized nonlinear models. **CSDA - Computational Statistics and Data Analysis**, Na impressão, p. –, 2008. Disponível em: <<http://www.sciencedirect.com/science/article/B6V8V-4TS6SFD-1/2/d90076c4c800a5f681ed9f121ca8211a>>.
- FAMOYE, F. Parameter estimation of generalized negative binomial distribution. **Communications Statistics, Theory and Methods**, v. 26, p. 269–279, 1997.
- FAMOYE, F.; CONSUL, P. C. Confidence interval estimation in the class of modified power series distributions. **Statistics**, v. 20, p. 141–148, 1989a.
- FAMOYE, F.; CONSUL, P. C. A stochastic urn model for the generalized negative binomial distribution. **Statistics**, v. 20, p. 607–613, 1989b.
- GOOD, I. J. Weight of evidence, corroboration, explanatory power, information and the utility of experiments. **Journal of the Royal Statistical Society, Serie B**, v. 22, p. 319 – 331, 1960.
- GOOD, I. J. A derivation of the probabilistic explication of information. **Journal of the Royal Statistical Society, Serie B**, v. 28, p. 578 –581, 1966.
- GUPTA, P. L.; GUPTA, R. C.; TRIPATHI, R. C. Inflated modified power series distributions with applications. **Communications Statistics, Theory and Methods**, v. 24 (9), p. 2355–2374, 1995.
- GUPTA, R. C. Modified power series distribution and some of its applications. **Sankhya Série B**, v. 36, p. 288–298, 1974.
- GUPTA, R. C.; SINGH, J. Estimation of probabilities in the class of modified power series distributions. **Statistics**, v. 13, p. 71–77, 1982.
- JAIN, G. C.; CONSUL, P. C. A generalized negative binomial distribution. **SIAM Journal on Applied Mathematics**, v. 21, n. 4, p. 501–513, 1971.
- JAIN, G. C.; GUPTA, R. C. A logarithmic type distribution. **Trab. Estadist.**, v. 24, p. 99–105, 1973.
- JAIN, G. C.; KHAN, S. H. On an exponential family. **Statistics**, v. 10, p. 153–168, 1979.
- JANI, P. On modified power series distributions. **Metron**, v. 36, p. 173–186, 1978.
- KASS, R. E.; WASSERMAN, L. The selection of prior distributions by formal rules. **Journal of the American Statistical Association**, v. 91, p. 1343–1370, 1996.
- KEMP, A. W.; KEMP, C. D. On a distribution associated with certain stochastic processes. **Journal of the Royal Statistical Society. Series B**, v. 30, n. 1, p. 160–163, 1968.

- KUMAR, A. Some applications of lagrangian distributions in queueing theory and epidemiology. **Communications Statistics, Theory and Methods**, v. 10, p. 1429–1436, 1981.
- KUMAR, A.; CONSUL, P. C. Negative moments of a modified power series distribution and bias of the maximum likelihood estimator. **Communications Statistics, Theory and Methods**, v. 8, p. 151–166, 1979.
- KUMAR, A.; CONSUL, P. C. Minimum variance unbiased estimation for modified power series distribution. **Communications Statistics, Theory and Methods**, v. 9 (12), p. 1261–1275, 1980.
- LINDLEY, D. V. On a measure of the information provided by an experiment. **Annals of Mathematical Statistics**, v. 27, p. 986 – 1005, 1956.
- MURAT, M.; SZYNAL, D. Non-zero inflated modified power series distributions. **Communications Statistics, Theory and Methods**, v. 27 (12), p. 3047–3064, 1998.
- NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear models. **Journal of the Royal Statistical Society Serie A**, v. 135, p. 370–384, 1972.
- NOACK, A. A class of random variables with discrete distributions. **The Annals of Mathematical Statistics**, v. 21, p. 127–132, 1950.
- PAULA, G. A. **Modelos de Regressão com Apoio Computacional**. São Paulo: Instituto de Matemática e Estatística - Universidade de São Paulo, 2004.
- PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística Bayesiana**. Lisboa: Fundação Calouste Gulbenkian, 2003.
- POLLARD, W. E. **Bayesian Statistics for Evaluation Research an Introduction**. New Delhi: Sage Publications, 1986.
- R Development Core Team. **R: A language and environment for statistical computing**. Vienna, Austria: R Foundation for Statistical Computing, 2008.
- STASINOPOULOS, D. M.; RIGBY, R. A. Generalized additive models for location, scale and shape (gamlss) in r. **Journal of Statistical Software**, v. 23, p. 1–46, 2007.
- SUN, D.; YE, K. Reference prior bayesian analysis for normal mean products. **Journal of the American Statistical Association**, v. 90, p. 589 – 597, 1995.
- TANNER, J. A problem of interference between two queues. **Biometrika**, v. 40, p. 58–69, 1953.
- TRIPATHI, R. C.; GUPTA, P. L.; GUPTA, R. C. Incomplete moments of modified power series distributions with applications. **Communications Statistics, Theory and Methods**, v. 15 (3), p. 999–1015, 1986.
- YE, K.; BERGER, J. Noninformative priors for inferences in exponential regression models. **Biometrika**, v. 78, p. 645 – 656, 1991.

## APÊNDICE A – Assimetria e Curtose

O presente apêndice tem por objetivo apresentar as expressões de assimetria e curtose das distribuições na classe MPSD parametrizadas na média apresentadas no Capítulo 3.

A assimetria e a curtose de uma variável aleatória  $Y$  na classe MPSD são, respectivamente, dadas por

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}}, \quad \gamma_2 = \frac{\mu_4}{\mu_2^2} - 3,$$

sendo  $\mu_4 = \frac{g(\theta)}{g'(\theta)} \frac{d\mu_3}{d\theta} + 3\mu_2^2$ ,  $\mu_3 = \frac{g(\theta)}{g'(\theta)} \frac{d\mu_2}{d\theta}$  e  $\mu_2$  a variância da classe MPSD (2.2).

Na Tabela A.1 estão relacionadas as expressões para assimetria e curtose para as distribuições apresentadas na Seção 3.1.

Tabela A.1: Assimetria e curtose de algumas distribuições na classe MPSD

Distribuição	Assimetria	Curtose
Binomial	$\frac{m-2\mu}{[m\mu(m-\mu)]^{1/2}}$	$\frac{m^2+3\mu(m-\mu)(m-3)}{m\mu(m-\mu)}$
Binomial Negativa	$\frac{m+2\mu}{[m\mu(m+\mu)]^{1/2}}$	$\frac{m^2+36m\mu^3}{m\mu(m+\mu)}$
Binomial Negativa Generalizada	$\frac{m^2-2m\beta\mu-2m\mu+3\beta^2\mu^2-3\beta\mu^2}{m\{\mu(m+\beta\mu)[m+(\beta-1)\mu]\}^{1/2}}$	$\frac{m^2+36m\mu^3}{15\mu^2-20\mu+6}$
Borel	$\frac{3\mu-2}{(\mu-1)^{1/2}}$	$\frac{t-1}{15\mu^2-20m\mu+6m^2}$
Borel-Tanner	$\frac{m(\mu-m)^{1/2}}{3\mu-2m}$	$\frac{m^2(\mu-m)}{15\mu^2-20m\mu+6m^2}$
Consul	$\frac{3\beta\mu^2-3\mu^2+4\mu-2\beta\mu-1}{\mu(\mu-1)[(\beta-1)\mu+1]^{1/2}}$	$\frac{15\beta^2\mu^4-30\beta\mu^4+60\beta\mu^3-20\beta^2\mu^3-36\beta\mu^2+15\mu^4-40\mu^3+36\mu^2-12\mu+6\beta^2\mu^2+6\beta\mu+1}{\beta\mu(\mu-1)[(\beta-1)\mu]+1}$
Delta-binomial	$\frac{3\beta\mu^2-3\mu^2+4m\mu-2m\beta\mu-m^2}{\{m^2\beta\mu(\mu-m)[(\beta-1)\mu+m]\}^{1/2}}$	$\frac{15\beta^2\mu^4-30\beta\mu^4+60m\beta\mu^3-20m\beta^2\mu^3-36m^2\beta\mu^2+15\mu^4-40m\mu^3+36m^2\mu^2-12m^2\mu+6m^2\beta^2\mu^2+6m^3\beta\mu+m^4}{m^2\beta\mu(\mu-m)[(\beta-1)\mu+m]}$
Geeta	$\frac{3\beta\mu^2-2\beta\mu-2\mu+1}{[(b-1)\mu(\mu-1)(\beta\mu-1)]^{1/2}}$	$\frac{15\beta^2\mu^4-20\beta^2\mu^3-20\beta\mu^3+24\beta\mu^2+6\beta^2\mu^2-6\beta\mu+6\mu^2-6\mu+1}{60\mu^4-120\mu^3+78\mu^2-18\mu+1}$
Haight	$\frac{[\mu(\mu-1)(2\mu-1)](0.5)}{2m^2-9m\mu+8\mu^2}$	$\frac{\mu(\mu-1)(2\mu-1)}{60\mu^4-120\mu^3+78\mu^2-18\mu+1}$
“Lost game”	$\frac{m\mu[(m-\mu)(m-2\mu)]^{1/2}}{\mu^{-1/2}}$	$\frac{6m^2\mu^4-9m^3\mu^3+3m^4\mu^2+4m^4+96\mu^4-210m\mu^3+156m^2\mu^2-45m^3\mu}{m^2\mu^2(m-\mu)(m-2\mu)}$
Poisson	$\mu^{-1/2}$	$\mu^{-1}$
Poisson generalizada	$\frac{3\beta^2\mu^2+4\beta\mu+1}{\mu^{1/2}(\beta\mu+1)}$	$\frac{15\beta^4\mu^4+40\beta^3\mu^3+36\beta^2\mu^2+12\beta\mu+1}{\mu(\beta\mu+1)}$