

**A DISTRIBUIÇÃO BETA BINOMIAL
NEGATIVA**

ANA PAULA COELHO MADEIRA

2009

ANA PAULA COELHO MADEIRA

A DISTRIBUIÇÃO BETA BINOMIAL NEGATIVA

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de "Mestre".

Orientador
Prof.Dr. Lucas Monteiro Chaves

**LAVRAS
MINAS GERAIS - BRASIL
2009**

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Madeira, Ana Paula Coelho.

A Distribuição Beta Binomial Negativa / Ana Paula Coelho
Madeira. – Lavras : UFLA, 2009.

81 p. : il.

Dissertação (Mestrado) – Universidade Federal de Lavras, 2009.
Orientador: Lucas Monteiro Chaves.

Bibliografia.

1. Mistura. 2. Distribuições generalizadas. 3. Distribuição
Binomial Negativa. 4. Distribuição Beta 5. Momentos fatoriais I.
Universidade Federal de Lavras. II. Título.

CDD – 519.53

ANA PAULA COELHO MADEIRA

A DISTRIBUIÇÃO BETA BINOMIAL NEGATIVA

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de "Mestre".

APROVADA em 26 de fevereiro de 2009

Prof. Dr. Daniel Furtado Ferreira

UFLA

Prof. Dr. Denismar Alves Nogueira

UNIFAL

Prof. Dr. João Domingos Scalon

UFLA

Prof. Dr. Lucas Monteiro Chaves
UFLA
(Orientador)

LAVRAS
MINAS GERAIS – BRASIL

Ao mestre dos mestres, Jesus Cristo,

Aos meus saudosos pais Sebastião (em memória) e Marlene,

Aos meus queridos irmãos Luiz Gustavo, Luiza e Antônio César,

Ao meu fiel companheiro Valdevino Júnior e minha amada filha Ana Luiza,

Aos meus dedicados mestres,

Dedico.

AGRADECIMENTOS

A Deus que se faz presente em minha vida, guiando meus passos e concedendo-me graças a cada dia.

À Universidade Federal de Lavras (UFLA), em especial ao Departamento de Ciências Exatas (DEX), pela realização de um sonho.

À FAPEMIG pela bolsa de estudo concedida.

Aos professores do DEX pelos valiosos ensinamentos e a professora Maria do Carmo pela amizade e carinho.

Ao meu orientador, Lucas Monteiro Chaves, pela arte de ensinar.

A todos os funcionários do DEX pelos serviços prestados.

Ao meu professor de Graduação Alex Erickson que muito contribuiu para minha vida acadêmica, incentivou-me e despertou em mim o hábito de estudar.

Aos meus colegas de curso Altemir, Ana Patrícia, Augusto, Denise, Edcarlos, Iron, Isabel, Paulo, Ricardo, Richardson, Stephânia e Tania pelo convívio e amizade.

Ao colega Devanil que muito acrescentou em minha formação.

Às minhas amigas Ana Patrícia e Tania pelas longas horas de estudo, brincadeiras, confidências, afeto,...

À minha amiga Lilyane pelo carinho e amizade de muitos anos.

Aos funcionários da Escola Estadual Firmino Costa, em especial à minha amiga Heloíza Ribeiro, pelo apoio, amizade e consideração.

Ao meu querido marido Valdevino Júnior pela compreensão, incentivo, zelo, amizade e amor.

À minha amada filha Ana Luiza por ser a razão de minhas conquistas.

Ao meu amado pai que, mesmo “entre nuvens”, olhou por mim e a quem sempre pude me espelhar. À minha prezada mãe pelo exemplo de vida, força, incentivo e dedicação.

Aos meus queridos irmãos Luiz Gustavo, Luiza e Antônio César que vibram com as minhas conquistas.

Enfim, a todos que contribuíram, direta ou indiretamente, meus eternos agradecimentos.

SUMÁRIO

	Página
LISTA DE TABELAS	i
LISTA DE FIGURAS	ii
RESUMO	iv
ABSTRACT	v
1 INTRODUÇÃO	1
2 REFERENCIAL TEÓRICO	3
2.1 Mistura de distribuições	3
2.1.1 Simulando misturas	12
2.2 Função geradora de probabilidade	14
2.2.1 Função geradora de probabilidade da mistura de distribuições	20
2.2.2 Função geradora de probabilidade da soma de um número aleatório de variáveis aleatórias independentes.....	21
2.3 Distribuições generalizadas.....	23
2.4 A mistura das distribuições binomial negativa e beta.....	32
2.4.1 Função geradora de momentos fatoriais da BBN.....	37
2.4.2 Média e variância.....	38
2.5 Unimodalidade.....	40
2.5.1 Unimodalidade da distribuição beta.....	40
2.5.2 Unimodalidade da distribuição binomial negativa.....	42
2.5.3 Unimodalidade de misturas.....	46
3 MATERIAL E MÉTODOS	50
3.1 Dados.....	50

3.2	Simulação de distribuições obtidas por mistura.....	50
3.3	Estimação e ajuste.....	50
4	RESULTADOS E DISCUSSÃO	52
4.1	Simulando mistura.....	52
4.2	Estimadores de máxima verossimilhança.....	55
4.2.1	Sistema de equações de aproximadas para o cálculo dos estimadores de máxima verossimilhança.....	56
4.3	Ajustes.....	62
5	CONCLUSÕES	69
	REFERÊNCIAS BIBLIOGRÁFICAS	70
	ANEXOS.....	72
	ANEXO A: Função gama.....	73
	ANEXO B: Rotinas.....	76

LISTA DE TABELAS

TABELA 1	Valores observados e esperados pelas distribuições binomial negativa e beta binomial negativa para dados do número de cáries em crianças de 12 anos.....	62
TABELA 2	Valores observados e esperados pelas distribuições binomial negativa e beta binomial negativa para dados do número de acidentes por maquinista num período de três meses.....	64
TABELA 3	Valores observados e esperados pelas distribuições binomial negativa e beta binomial negativa para dados do número de plantas, <i>Salicornia stricta</i> , por quadrado.....	65
TABELA 4	Valores observados e esperados pelas distribuições binomial negativa e beta binomial negativa para dados de contagem do número de ácaros em folhas de maçã.....	67

LISTA DE FIGURAS

FIGURA 1	Representação de um padrão espacial aleatório.....	7
FIGURA 2	Representação da mistura de normais.....	12
FIGURA 3	Representação gráfica da mistura das distribuições Poisson e gama.....	13
FIGURA 4	Representação de um padrão espacial agrupado.....	24
FIGURA 5	Distribuição beta binomial negativa $3, 5, 7$	35
FIGURA 6	Gráficos da distribuição beta binomial negativa para diferentes valores dos parâmetros a , b e r	36
FIGURA 7	Gráficos da distribuição beta.....	41
FIGURA 8	Distribuição binomial negativa $r = 5, p = 0,3$	43
FIGURA 9	Distribuição binomial negativa $r = 5, p = 0,08$	44
FIGURA 10	Distribuição binomial negativa $r = 5, p = 0,9$	45
FIGURA 11	Discretização da variável x e do parâmetro θ	52
FIGURA 12	Distribuição das frequências do número de cáries por criança. Ajuste pelas distribuições binomial negativa e beta binomial negativa	63
FIGURA 13	Distribuição das frequências de acidentes por maquinista, período de três meses. Ajuste pelas distribuições binomial negativa e beta binomial negativa	64
FIGURA 14	Distribuição do número de plantas <i>Salicornia stricta</i> por quadrado. Ajuste pelas distribuições binomial negativa e beta binomial negativa.....	66

FIGURA 15	Distribuição do número ácaros em folhas de maçã. Ajuste pelas distribuições binomial negativa e beta binomial negativa.....	67
-----------	---	----

RESUMO

MADEIRA, Ana Paula Coelho. **A Distribuição beta binomial negativa**. 2009. 81 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, MG.¹

Os conceitos de mistura e de distribuições generalizadas são trabalhados. A mistura da distribuição binomial negativa com a distribuição beta é obtida e denominada distribuição beta binomial negativa. A distribuição é unimodal, apresenta super dispersão e não possui todos os momentos. Uma aproximação das equações de log-verossimilhança é encontrada, obtendo-se estimativas dos parâmetros do modelo. Uma validação matemática para o algoritmo de simulação de misturas é apresentado. Como aplicações, são realizados os ajustes da distribuição beta binomial negativa a dados da literatura, analisados anteriormente pela distribuição binomial negativa.

¹ **Orientador:** Lucas Monteiro Chaves – UFLA

ABSTRACT

MADEIRA, Ana Paula Coelho. **The beta negative binomial distribution** 2009. 81p. Dissertation (Master of Statistics and Agricultural Experimentation) Federal University of Lavras, Lavras, MG.¹

The mixture concepts and of generalized distributions are worked. The mixture of the negative binomial distribution with the beta distribution is obtained and denominated negative beta binomial distribution. The distribution is unimode, it presents over dispersion and it doesn't present every moments. An approach of the log-likelihood equations is found, being obtained the estimator of the parameters of the model. A mathematical validation for the algorithm of simulation of mixtures is presented. As application, fittings are accomplished from the negative beta binomial distribution to data of the literature, previously analyzed by the negative binomial distribution.

¹ **Adviser:** Lucas Monteiro Chaves – UFLA

1 INTRODUÇÃO

Novas distribuições de probabilidade são obtidas quando é admitido que o parâmetro da distribuição também varie segundo uma distribuição de probabilidade, isto é, o parâmetro de interesse passa a ser considerado também uma variável aleatória com sua própria distribuição. Esse mecanismo de obtenção de distribuições de probabilidade é definido como mistura. É bem conhecido que a distribuição beta binomial surge quando o parâmetro p da binomial é permitido variar de grupo a grupo segundo uma distribuição beta. Outra distribuição obtida pelo processo de mistura é a binomial negativa, muito usada no ajuste de dados biológicos. Essa distribuição é resultado da mistura das distribuições Poisson e gama. O conceito de mistura é suficientemente flexível para se permitir obter novas distribuições, a partir de distribuições conhecidas, com propriedades desejáveis.

Outro processo similar para se obter novas distribuições é dado pelo conceito de distribuições generalizadas, resultado da soma de um número aleatório de variáveis aleatórias independentes e identicamente distribuídas. Novamente, obtém-se a distribuição binomial negativa como distribuição generalizada de uma distribuição de Poisson com uma distribuição logarítmica. Deste modo, tem-se que, a distribuição binomial negativa pode ser obtida por dois processos distintos, com pressupostos diferentes, mistura e distribuição generalizada. Algumas relações entre mistura de distribuições e distribuições generalizadas foram obtidas por Gurland (1957).

A distribuição binomial negativa tem sido muito empregada para descrever a distribuição de pragas em lavouras. Considerando uma plantação, define-se uma unidade espacial na qual é contado o número de insetos. A presença de insetos em uma unidade espacial é afetada pela maior ou menor

densidade nas unidades vizinhas, portanto, é razoável supor que o parâmetro p , probabilidade de ocorrência de um inseto, varia de unidade a unidade. Supondo uma distribuição beta para o parâmetro p e fazendo a mistura entre essas distribuições, obtém-se a distribuição discreta beta binomial negativa.

Alguns autores, entre eles, Jonhson & Kotz (1969), Gerstenkorn (2004) e Hassan & Bilal (2008) trabalham com a mistura das distribuições binomial negativa e beta.

O presente trabalho teve por objetivo estudar a mistura das distribuições binomial negativa e beta, a distribuição beta binomial negativa, estimar seus parâmetros segundo o método da máxima verossimilhança. Como aplicação, alguns ajustes a dados anteriormente analisados na literatura são realizados, no sentido de mostrar a aplicabilidade desta distribuição. Uma validação matemática do algoritmo para simular misturas é apresentada.

2 REFERENCIAL TEÓRICO

2.1 Mistura de distribuições

Seja $f_0(\cdot), f_1(\cdot), \dots, f_n(\cdot), \dots$ uma sequência de funções densidades de probabilidades e $p_0, p_1, \dots, p_n, \dots$ uma sequência dos números satisfazendo $p_i \geq 0$ e $\sum_{i=0}^{\infty} p_i = 1$, então $h(x) = \sum_{i=0}^{\infty} p_i f_i(x)$ é também uma densidade de probabilidade, pois

$$\begin{aligned} \sum_x h(x) &= \sum_x \sum_i p_i f_i(x) = \sum_i \sum_x p_i f_i(x) \\ &= \sum_i p_i \left(\sum_x f_i(x) \right) = \sum_i p_i = 1, \end{aligned}$$

e é denominada mistura. Tal procedimento é útil para se obter novas distribuições de probabilidade com propriedades adequadas. Um exemplo de interesse é:

Exemplo 1 Mistura de normais.

$$h(x) = p f_1(x; \mu_1) + (1-p) f_2(x; \mu_2)$$

em que

$$f_j(x; \mu_j, 1) = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} (x - \mu_j)^2 \right], \quad j = 1, 2$$

$h(x)$ é uma mistura de normais envolvendo dois componentes com médias diferentes μ_1 e μ_2 e variâncias iguais a um, onde p é chamado peso ou proporção da mistura.

Para $p = \frac{1}{3}$, tem-se

$$h(x) = \frac{1}{3} \left\{ \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2}(x - \mu_1)^2 \right] \right\} + \frac{2}{3} \left\{ \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2}(x - \mu_2)^2 \right] \right\}.$$

É interessante notar que, os estimadores de máxima verossimilhança (considerando uma amostra de tamanho $n=1$) dos parâmetros μ_1 e μ_2 , dados pela solução do sistema,

$$\begin{cases} \frac{\partial h(x)}{\partial \mu_1} = \frac{(x - \hat{\mu}_1)}{3\sqrt{2\pi}} \exp \left[-\frac{1}{2}(x - \hat{\mu}_1)^2 \right] = 0 \\ \frac{\partial h(x)}{\partial \mu_2} = \frac{2(x - \hat{\mu}_2)}{3\sqrt{2\pi}} \exp \left[-\frac{1}{2}(x - \hat{\mu}_2)^2 \right] = 0 \end{cases}$$

$$\frac{(x - \hat{\mu}_1)}{2(x - \hat{\mu}_2)} = \exp \left\{ -\frac{1}{2} \left[(x - \hat{\mu}_2)^2 - (x - \hat{\mu}_1)^2 \right] \right\}$$

não possuem expressões explícitas pelo fato das equações serem transcendentais. Assim, os valores estimados são obtidos por métodos numéricos iterativos, tais como: algoritmo EM, métodos de Newton-Raphson, entre outros.

Na mistura de normais, pode-se ainda considerar os casos em que as médias são iguais e variâncias diferentes, as médias e as variâncias são diferentes e, por fim, misturas envolvendo mais que duas normais.

Exemplo 2 Seja X uma variável aleatória com distribuição $N(\mu, 1)$.

Suponha que μ é também uma variável aleatória com distribuição $N(0, 1)$.

$$X \sim N(\mu, 1) \rightarrow f(x; \mu) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x - \mu)^2\right\}$$

$$\mu \sim N(0, 1) \rightarrow g(\mu) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\mu^2\right\}$$

A distribuição $h(x)$, é dada pela mistura $f(x; \mu) \wedge_{\mu} g(\mu)$

$$\begin{aligned} h(x) &= \int_{-\infty}^{\infty} f(x; \mu) g(\mu) d\mu \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x - \mu)^2\right\} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\mu^2\right\} d\mu \\ &= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(2\mu^2 - 2\mu x)\right\} d\mu \\ &= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(x^2 - \frac{x^2}{2}\right)\right\} \\ &\quad \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left((\sqrt{2}\mu)^2 - 2(\sqrt{2}\mu)\left(\frac{x}{\sqrt{2}}\right) + \left(\frac{x}{\sqrt{2}}\right)^2\right)\right\} d\mu \end{aligned}$$

Fazendo $y = \sqrt{2} \mu \rightarrow d\mu = \frac{1}{\sqrt{2}} dy$,

$$\begin{aligned}
 &= \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(x^2 - \frac{x^2}{2} \right) \right\} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(y - \frac{x}{\sqrt{2}} \right)^2 \right\} dy \\
 &= \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{x}{\sqrt{2}} \right)^2 \right\} \rightarrow N(0, 2)
 \end{aligned}$$

Silva (2003) destaca o uso da mistura de normais na Genética onde esses modelos são usados para estimar os parâmetros relativos a um gene de efeito maior e no mapeamento de QTLs.

O conceito de mistura pode ser estendido.

Seja $\{f(x; \theta); \theta \in \bar{\Theta}\}$ uma família de funções densidade de probabilidade parametrizadas por θ em um espaço paramétrico $\bar{\Theta}$. Se $g(\theta)$ é uma função densidade de probabilidade definida em $\bar{\Theta}$, então

$$h(x) = \int_{\bar{\Theta}} f(x; \theta) g(\theta) d\theta$$

é uma nova densidade de probabilidade, pois

$$\begin{aligned}
 \int h(x) dx &= \int \int_{\bar{\Theta}} f(x; \theta) g(\theta) d\theta dx = \int \int_{\bar{\Theta} \times x} f(x; \theta) g(\theta) dx d\theta \\
 &= \int_{\bar{\Theta}} g(\theta) \left(\int_x f(x; \theta) dx \right) d\theta \\
 &= \int_{\bar{\Theta}} g(\theta) d\theta = 1
 \end{aligned}$$

também é denominada mistura das distribuições $f(x; \theta)$ e $g(\theta)$. Considere a notação $f_{X \wedge Y}(x) = f_X(x; \theta) \wedge_{\theta} g_Y(\theta)$, apresentada por Gurland (1957), para representar a densidade de probabilidade da mistura das distribuições $X \sim f(x; \theta)$ com $Y \sim g(\theta)$, em que $f(x; \theta)$ representa a distribuição original dependendo do parâmetro θ e $g(\theta)$ a distribuição para o parâmetro.

Exemplo 3 Suponha que fêmeas de insetos depositem ovos em determinados locais que denominaremos de unidades, folhas, por exemplo. Suponha que se tenha uma uniformidade na postura dos ovos, isto é, as fêmeas depositam exatamente n ovos nas unidades (folhas). Essa situação encontra-se representada na Figura 1.



FIGURA 1 Padrão espacial aleatório.

Considere, agora, o seguinte modelo: com probabilidade p , cada ovo eclode dando origem a uma larva. Então, para cada unidade a quantidade de larvas será o resultado de n ensaios de Bernoulli. O número de larvas por folha (unidade) segue um modelo binomial, $X \sim \text{Bin}(n, p)$. A probabilidade de exatamente k ovos eclodirem é dada por

$$P[X = k] = \binom{n}{k} p^k (1-p)^{n-k} I_{\{0,1,2,\dots,n\}}(x).$$

Porém, sob certas condições, a pressuposição de que p é constante em todas as unidades não é válida, ou seja, a probabilidade de um ovo eclodir varia de unidade a unidade. Tal pressuposição faz sentido, pois uma folha mais exposta ao sol, por exemplo, pode ser menos propícia ao desenvolvimento dos ovos. Uma forma de modelar a variação de p é supor que p varia segundo uma distribuição beta

$$g(p) = \frac{1}{B(a,b)} p^{a-1} (1-p)^{b-1} I_{(0,1)}(p).$$

A distribuição resultante da mistura das distribuições binomial e beta, $\text{Bin}(n, p) \underset{p}{\wedge} \text{Beta}(a, b)$, é denominada beta binomial cuja distribuição de probabilidade é dada por

$$\begin{aligned} h(x) &= \int_0^1 f(x; p) g(p) dp \\ &= \int_0^1 \binom{n}{x} p^x (1-p)^{n-x} \frac{1}{B(a,b)} p^{a-1} (1-p)^{b-1} dp \\ &= \binom{n}{x} \frac{1}{B(a,b)} \int_0^1 p^{a+x-1} (1-p)^{n-x+b-1} dp \end{aligned}$$

$$= \binom{n}{x} \frac{B(a+x, b+n-x)}{B(a, b)} I_{\{0,1,2,\dots,n\}}(x)$$

Outra aplicação de mistura de distribuições de probabilidade é discutida em Pielou (1977). Na distribuição espacial de certos organismos, a probabilidade de uma unidade conter x indivíduos é modelada por uma

distribuição de Poisson(λ), $f(x; \lambda) = \frac{e^{-\lambda} \lambda^x}{x!} I_{\{0,1,2,\dots\}}(x)$. Considerando

que as unidades são heterogêneas, algumas provêm de ambientes mais favoráveis que outras, a média λ de indivíduos na unidade varia de unidade a unidade, isto é, λ é também uma variável aleatória. Considerando a distribuição

de λ uma gama, $g(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} I_{(0,\infty)}(\lambda)$, tem-se

$$\begin{aligned} h(x) &= \int_0^\infty f(x; \lambda) g(\lambda) d\lambda = \int_0^\infty \frac{e^{-\lambda} \lambda^x}{x!} \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} d\lambda \\ &= \frac{\beta^\alpha}{x! \Gamma(\alpha)} \int_0^\infty \lambda^{\alpha+x-1} e^{-(\beta+1)\lambda} d\lambda \\ &= \frac{\beta^\alpha}{x! \Gamma(\alpha) (\beta+1)^{\alpha+x}} \int_0^\infty \frac{(\beta+1)^{\alpha+x}}{\Gamma(\alpha+x)} \lambda^{\alpha+x-1} e^{-(\beta+1)\lambda} d\lambda \\ h(x) &= \binom{\alpha+x-1}{x} \left(\frac{\beta}{\beta+1} \right)^\alpha \left(\frac{1}{\beta+1} \right)^x I_{\{0,1,2,\dots\}}(x) . \end{aligned} \quad \#$$

A distribuição resultante para o número de indivíduos por unidade é, portanto, uma distribuição binomial negativa com parâmetros α e $p = \frac{\beta}{\beta+1}$. A média e a variância são, respectivamente, $E[X] = \frac{\alpha q}{p}$ e $Var[X] = \frac{\alpha q}{p^2}$.

Exemplo 4 Seja X_n uma variável aleatória com distribuição binomial com parâmetros n e p , $X_n = 0, 1, \dots, n$ e $b(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$. Defina

$\binom{n}{k} = 0$ para $k > n$ e estenda a definição de X_n para uma variável aleatória

assumindo valores em $\mathbb{N} = \{0, 1, 2, \dots, k, \dots, n, n+1, \dots\}$. Neste caso,

$$\sum_{k=0}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} = 1.$$

Considere n uma variável aleatória com distribuição Poisson com

$$\text{parâmetro } \lambda, \quad g(n) = \frac{e^{-\lambda} \lambda^n}{n!} I_{\{0, 1, 2, \dots\}}(n).$$

A mistura $b(k; n, p) \wedge_n g(n)$ é dada por

$$\begin{aligned} h(k) &= \sum_{n=0}^{\infty} b(k; n, p) g(n) = \sum_{n=k}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} \frac{e^{-\lambda} \lambda^n}{n!} \\ &= e^{-\lambda} \sum_{n=k}^{\infty} \frac{n!}{(n-k)! k!} \left(\frac{p}{q}\right)^k \frac{(q\lambda)^n}{n!} = e^{-\lambda} \left(\frac{p}{q}\right)^k \frac{1}{k!} \sum_{n=k}^{\infty} (q\lambda)^n \frac{1}{(n-k)!} \end{aligned}$$

$$\begin{aligned}
&= e^{-\lambda} \left(\frac{p}{q} \right)^k \frac{1}{k!} \sum_{n=k}^{\infty} (q\lambda)^{n-k} (q\lambda)^k \frac{1}{(n-k)!} \\
&= e^{-\lambda} \left(\frac{p}{q} \right)^k \frac{1}{k!} (q\lambda)^k \sum_{n=k}^{\infty} (q\lambda)^{n-k} \frac{1}{(n-k)!}
\end{aligned}$$

Fazendo a substituição $s = n - k$ tem-se que, quando n varia de k até ∞ ($n = k \rightarrow \infty$), s varia de 0 a ∞ ($s = 0 \rightarrow \infty$). Portanto,

$$\begin{aligned}
h(k) &= e^{-\lambda} \left(\frac{p}{q} \right)^k \frac{1}{k!} (q\lambda)^k \sum_{s=0}^{\infty} (q\lambda)^s \frac{1}{s!} = e^{-\lambda} \left(\frac{p}{q} \right)^k \frac{1}{k!} (q\lambda)^k e^{q\lambda} \\
&= (p\lambda)^k e^{\lambda(q-1)} \frac{1}{k!}
\end{aligned}$$

Portanto temos a distribuição de Poisson, com parâmetro λp

$$h(k) = \frac{e^{-\lambda p} (\lambda p)^k}{k!} I_{\{0,1,2,\dots\}}(k) . \quad \#$$

Os exemplos acima mostram, claramente, que o conceito de mistura é muito útil na modelagem de fenômenos aleatórios. Permite que se faça uma hierarquia entre os parâmetros e depois toma-se o efeito médio de todos eles. Como referência para o estudo das distribuições obtidas pelo processo de mistura, tem-se o livro clássico de Jonhson & Kotz (1969).

2.1.1 Simulando misturas

A seguir são apresentados três exemplos de simulação de misturas, obtidos via algoritmo mistura, Devroye (1986). Uma validação matemática para o algoritmo é apresentada na seção resultados e discussão.

Exemplo 5 Mistura de normais

Seja $x_1, x_2, \dots, x_n$ uma amostra aleatória de tamanho $n = 100000$, obtida por simulação (ver rotina no apêndice) da distribuição $h(x)$.

A distribuição $h(x) = \int_{-\infty}^{\infty} f(x; \mu)g(\mu)d\mu$ é resultado da mistura de normais, $X \sim N(\mu; 1)$ e $\mu \sim N(0; 1)$

Na Figura 2 apresenta-se o histograma de freqüências de x .

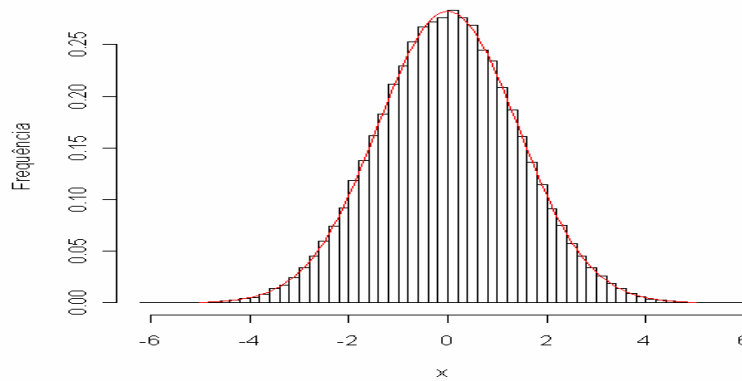


FIGURA 2 Função densidade de probabilidade que aproxima a distribuição de freqüências observadas.

Percebe-se que a distribuição $h(x)$ corresponde a uma distribuição Normal com média $\mu = 0$. O resultado obtido pode ser confirmado pelo exemplo 2, onde se obtém uma distribuição $N(0, 2)$ a partir da mistura de duas normais.

Exemplo 6 Simulando a distribuição binomial negativa.

Seja $x_1, x_2, \dots, x_n$ uma amostra aleatória de tamanho $n = 100000$, obtida por simulação, via algoritmo mistura, da distribuição $h(x)$. A distribuição $h(x)$ é resultado da mistura das distribuições Poisson(λ) e gama(10,1).

Um histograma de frequências de x é apresentado na Figura 3.

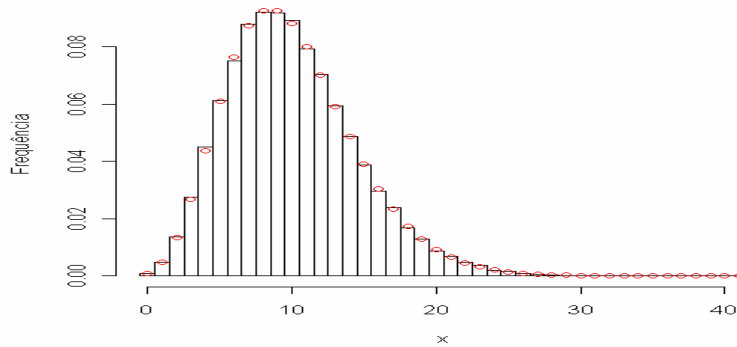


FIGURA 3 Densidade da mistura das distribuições Poisson e gama.

Constatou-se que a distribuição $h(x)$ apresenta comportamento semelhante ao de uma distribuição binomial negativa. De fato, a mistura das distribuições Poisson e gama resulta em uma distribuição binomial negativa (ver página 9), neste caso com parâmetros $(10, 0.5)$.

2.2 Função geradora de probabilidade

Seja X uma variável aleatória discreta assumindo valores inteiros não negativos, $X \in \{0, 1, 2, \dots\}$, $P\{X = x\} = p_x$ e $\sum_{x=0}^{\infty} p_x = 1$. A função geradora de probabilidade (f.g.p.) de X , denotada por $G_X(z)$, ou apenas $G(z)$ quando a variável aleatória a qual se refere a f.g.p estiver explicitada, é definida por

$$G_X(z) = \sum_{x=0}^{\infty} p_x z^x = E[z^X].$$

Sendo X uma variável aleatória contínua, a função geradora de probabilidade é dada por $G_X(z) = E[z^X] = \int_{-\infty}^{\infty} z^x f_X(x) dx$.

Exemplo 7 Seja X uma variável aleatória com distribuição binomial negativa com parâmetros r e p

$$f_X(x; r, p) = \binom{x+r-1}{x} p^r q^x I_{\{0,1,2,\dots\}}(x).$$

Reescrevendo o coeficiente binomial de acordo com Feller (1968) página 61,

$$\begin{aligned} \binom{-r}{x} &= \frac{(-r)(-r-1)\cdots(-r-(x-1))}{x!} = (-1)^x \frac{r(r+1)\cdots(r+x-1)}{x!} \\ &= (-1)^x \binom{r+x-1}{x} \text{ para } r > 0 \end{aligned}$$

obtém-se uma expressão alternativa para a densidade

$$f_X(x; r, p) = \binom{-r}{x} p^r (-q)^x I_{\{0,1,2,\dots\}}(x).$$

$G_X(z) = E[z^X]$ é dada por

$$\begin{aligned} G_X(z) &= \binom{-r}{0} p^r + \binom{-r}{1} p^r (-q) z + \binom{-r}{2} p^r (-q)^2 z^2 + \dots \\ &= p^r \left[\binom{-r}{0} + \binom{-r}{1} (-qz) + \binom{-r}{2} (-qz)^2 + \dots \right]. \end{aligned}$$

Como $\sum_{k=0}^{\infty} \binom{-r}{k} (-q)^k = (1-q)^{-r}$ (Mood et al., 1974), segue que

$$G_X(z) = p^r (1-qz)^{-r} = \left(\frac{p}{1-qz} \right)^r. \quad (2.1)$$

Definição 1 Se X é uma variável aleatória, o l -ésimo momento fatorial de X é definido como (l é um inteiro positivo):

$$E[X(X-1)\dots(X-(l-1))] = E[X^{(l)}]$$

em que $x^{(l)} = x(x-1)(x-2)\dots(x-(l-1))$ representa um fatorial descendente.

De uso mais corrente que a função geradora de probabilidade tem-se a função geradora de momentos definida por $m_X(t) = E[e^{tX}]$ (Mood et al., 1974).

Uma observação interessante é que, a função geradora de probabilidade e a função geradora de momentos estão relacionadas. Tal fato, não se encontra explicitado na maioria dos textos usuais de estatística matemática. Esta relação é:

$$G_X(z) = E[z^X] = E[e^{X \ln z}] = m_X(\ln z).$$

A função geradora de probabilidade é também chamada de função geradora de momentos fatoriais, pois é usada para gerar momentos fatoriais do mesmo modo como os momentos usuais são obtidos a partir da função geradora de momentos, $m_X(t) = E[e^{tX}]$, exceto que z se aproxima de 1 em vez de 0.

$$G(z) = E[z^X]$$

$$\begin{aligned} G'(z) &= \frac{d}{dz} E[z^X] = \frac{d}{dz} \int z^x f_X(x) dx = \int \frac{d}{dz} z^x f_X(x) dx \\ &= \int x z^{x-1} f_X(x) dx = E[X z^{X-1}] \end{aligned}$$

$$\lim_{z \rightarrow 1} G'(z) = E[X]$$

$$\begin{aligned} G''(z) &= \frac{d}{dz} \int x z^{x-1} f_X(x) dx = \int \frac{d}{dz} x z^{x-1} f_X(x) dx \\ &= \int x(x-1) z^{x-2} f_X(x) dx = E[X(X-1) z^{(X-1)(X-2)}] \end{aligned}$$

$$\lim_{z \rightarrow 1} G''(z) = E[X(X-1)]$$

$$G^3(z) = \int \frac{d}{dz} x(x-1)z^{x-2} f_X(x) dx$$

$$\begin{array}{ccc} \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \end{array}$$

$$G^l(z) = E \left[X(X-1) \dots (X-(l-1)) z^{(X-1)(X-2) \dots (X-l)} \right]$$

$$G^l(1) = \lim_{z \rightarrow 1} G^l(z) = E \left[X(X-1) \dots (X-(l-1)) \right]$$

$$G^l(1) = m_{(l)}$$

Observe que, a média e a variância de X são:

$$E[X] = G'(1) \quad (2.2)$$

$$Var[X] = E(X^2) - [E(X)]^2 \quad (2.3)$$

Sendo $G''(1) = E[X(X-1)] = E(X^2) - E(X)$, segue que

$$E(X^2) = G''(1) + E(X) = G''(1) + G'(1) \quad (2.4)$$

Substituindo (2.4) em (2.3)

$$\begin{aligned} Var[X] &= G''(1) + G'(1) - [G'(1)]^2 \\ &= G''(1) + G'(1)[1 - G'(1)] \end{aligned} \quad (2.5)$$

A importância da função geradora de momentos fatoriais fica evidente no seguinte teorema.

Teorema 1 Se X e Y são variáveis aleatórias independentes com função geradora de momentos fatoriais iguais, então X e Y tem a mesma distribuição de probabilidade (Mood et al., 1974).

Para algumas variáveis aleatórias (usualmente as discretas), os momentos fatoriais podem ser facilmente obtidos a partir dos momentos usuais e vice-versa.

Exemplo 8 Seja X uma variável aleatória com distribuição de Poisson com parâmetro λ . Sua função geradora de probabilidade é dada por

$$\begin{aligned} G(z) &= \sum_{x=0}^{\infty} p_x z^x = p_0 + p_1 z + p_2 z^2 + \dots \\ &= e^{-\lambda} + e^{-\lambda} (\lambda z) + \frac{e^{-\lambda} (\lambda z)^2}{2!} + \frac{e^{-\lambda} (\lambda z)^3}{3!} + \dots \\ &= e^{-\lambda} \left(1 + \lambda z + \frac{(\lambda z)^2}{2!} + \frac{(\lambda z)^3}{3!} + \dots \right) = e^{-\lambda} \cdot e^{\lambda z} = e^{\lambda(z-1)} \end{aligned} \quad (2.6)$$

Assim,

$$E[X] = G'(1) = \lambda e^{\lambda(z-1)} \Big|_{z=1} = \lambda.$$

$$\text{Var}[X] = G''(1) + G'(1)[1 - G'(1)].$$

Como $G''(1) = \lambda^2 e^{\lambda(z-1)} \Big|_{z=1} = \lambda^2$, segue que $\text{Var}[X] = \lambda^2 + \lambda[1 - \lambda] = \lambda$.

Exemplo 9 Seja X uma variável aleatória com distribuição logarítmica,

$f(x; p) = \frac{q^x}{-x \ln p} I_{\{1,2,3,\dots\}}(x)$. A função geradora de probabilidade da distribuição logarítmica é dada por

$$\begin{aligned} G_X(z) &= p_1 z + p_2 z^2 + p_3 z^3 + \dots \\ &= -\frac{qz}{\ln p} - \frac{(qz)^2}{2 \ln p} - \frac{(qz)^3}{3 \ln p} - \dots = -\frac{1}{\ln p} \left[qz + \frac{(qz)^2}{2} + \frac{(qz)^3}{3} + \dots \right]. \end{aligned}$$

Sendo $\ln\left(\frac{1}{1-qz}\right) = qz + \frac{(qz)^2}{2} + \frac{(qz)^3}{3} + \dots$ Anton (2000), obtém-se,

$$\begin{aligned} G_X(z) &= -\frac{1}{\ln p} \left[\ln\left(\frac{1}{1-qz}\right) \right] = -\frac{1}{\ln p} \left[\ln(1-qz)^{-1} \right] \\ &= -\frac{1}{\ln p} \left[-\ln(1-qz) \right] = \frac{\ln(1-qz)}{\ln p}. \end{aligned} \quad (2.7) \quad \#$$

Exemplo 10 Seja X uma variável aleatória contínua com distribuição gama(α, β). A função geradora de probabilidade da variável X é dada por

$$\begin{aligned} X \sim \text{Gama}(\alpha, \beta) &\rightarrow f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} I_{(0,\infty)}(x) \\ G_X(z) &= E[z^X] = \int_{-\infty}^{\infty} z^x f_X(x) dx = \int_0^{\infty} z^x \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx \end{aligned}$$

$$\begin{aligned}
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty z^x x^{\alpha-1} e^{-\beta x} dx = \beta^\alpha \int_0^\infty \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-(\beta - \ln z)x} dx \\
&= \frac{\beta^\alpha}{(\beta - \ln z)^\alpha} \int_0^\infty \frac{(\beta - \ln z)^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-(\beta - \ln z)x} dx \\
&= \left(\frac{\beta}{\beta - \ln z} \right)^\alpha. \tag{2.8}\#
\end{aligned}$$

2.2.1 Função geradora de probabilidade da mistura de distribuições

Seja X_θ uma variável aleatória com função de probabilidade $f(x; \theta)$ e função geradora de probabilidade $G_{X_\theta}(z)$. Considere a variável aleatória X com distribuição $h(x) = \int_{\bar{\Theta}} f(x; \theta) g(\theta) d\theta$ em que $g(\theta)$ é uma distribuição para o parâmetro θ .

A função geradora de probabilidade de X , $G_X(z)$ é dada por

$$G_X(z) = \int_{-\infty}^{\infty} G_{X_\theta}(z) g(\theta) d\theta.$$

$$\begin{aligned}
\textbf{Prova} \quad G_X(z) &= E[z^X] = \int_{-\infty}^{\infty} z^x h(x) dx = \int_{-\infty}^{\infty} z^x \left[\int_{\bar{\Theta}} f(x; \theta) g(\theta) d\theta \right] dx \\
&= \int_{\bar{\Theta}} \left[\int_{-\infty}^{\infty} z^x f(x; \theta) dx \right] g(\theta) d\theta = \int_{\bar{\Theta}} E[z^{X_\theta}] g(\theta) d\theta \\
&= \int_{\bar{\Theta}} G_{X_\theta}(z) g(\theta) d\theta \tag{\#}
\end{aligned}$$

Os momentos fatoriais das distribuições resultantes de mistura de distribuições são obtidos pela derivação de $G_X(z)$.

Outra forma de obter os momentos é dada por Gerstenkorn (2004).

$$m_{(l)} = E\left[X_{\theta}^{(l)}\right] = \int_{-\infty}^{\infty} E\left[X_{\theta}^{(l)}\right] g(\theta) d\theta \quad (2.9)$$

sendo representado por $E\left[X_{\theta}^{(l)}\right] \hat{=} g(\theta)$, em que $E\left[X_{\theta}^{(l)}\right]$ é o momento fatorial da distribuição original.

$$\begin{aligned} \textbf{Prova} \quad E\left[X_{\theta}^{(l)}\right] &= \int_{-\infty}^{\infty} X_{\theta}^{(l)} h(x) dx = \int_{-\infty}^{\infty} X_{\theta}^{(l)} \left[\int_{\Theta} f(x; \theta) g(\theta) d\theta \right] dx \\ &= \int_{\Theta} g(\theta) \left[\int_{-\infty}^{\infty} X_{\theta}^{(l)} f(x; \theta) dx \right] d\theta \\ &= \int_{\Theta} E\left[X_{\theta}^{(l)}\right] g(\theta) d\theta \quad \# \end{aligned}$$

2.2.2 Função geradora de probabilidade da soma de um número aleatório de variáveis aleatórias independentes

Sejam X e Y variáveis aleatórias independentes e $S = X + Y$.

$$G_S(z) = E\left[z^{X+Y}\right] = E\left[z^X z^Y\right] = E\left[z^X\right] E\left[z^Y\right] = G_X(z) G_Y(z)$$

No caso em que X e Y são variáveis aleatórias discretas não negativas, $a_j = P\{X = j\}$ e $b_k = P\{Y = k\}$, o evento $\{X = j, Y = k\}$ tem probabilidade $a_j b_k$. O evento $\{S = r\}$ é a união dos eventos mutuamente excludentes $\{S = r\} = \{X = 0, Y = r\} \cup \{X = 1, Y = r-1\} \cup \dots \cup \{X = r, Y = 0\}$.

Portanto,

$$c_r = P\{S = r\} = a_0 b_r + a_1 b_{r-1} + a_2 b_{r-2} + \dots + a_r b_0 \quad (2.10)$$

A sequência $\{c_r\}$ definida por (2.10) é dita convolução das sequências $\{a_k\}$ e $\{b_k\}$ e é usualmente denotada por $\{c_k\} = \{a_k\} * \{b_k\}$. Essa fórmula segue diretamente de $G_X(z)G_Y(z)$, pois

$$\begin{aligned} G_X(z)G_Y(z) &= (a_0 + a_1 z + a_2 z^2 + a_3 z^3 + \dots)(b_0 + b_1 z + b_2 z^2 + b_3 z^3 + \dots) \\ &= a_0 b_0 + (a_0 b_1 + a_1 b_0)z + (a_0 b_2 + a_1 b_1 + a_2 b_0)z^2 + \dots \end{aligned}$$

Uma situação importante em aplicações é a seguinte: seja N uma variável aleatória, assumindo valores inteiros, $N \geq 1$. Considere $S_N = X_1 + X_2 + \dots + X_N$ a soma de variáveis aleatórias X_i , independentes e identicamente distribuídas (*i.i.d.*) em que N , o número de termos da soma é também uma variável aleatória independente dos X_i . Seja $p_n = P\{N = n\}$ a distribuição de N e $G_N(z)$ sua função geradora de probabilidade. A função geradora de probabilidade de S_N é dada por

$$\begin{aligned}
G_{S_N}(z) &= E[z^{S_N}] = E[E[z^{S_N} | N]] \quad (\text{Mood et al., 1974}). \\
&= E[E[z^{X_1+X_2+\dots+X_N} | N]] = E[E[z^{X_1} \dots z^{X_N} | N]] \\
&= E[(G_X(z))^N] = G_N(G_X(z)) \tag{2.11}
\end{aligned}$$

Portanto, tem-se uma propriedade simples para o cálculo da função geradora de probabilidade de S_N , dada pela composição de funções,

$$G_{S_N}(z) = G_N(G_X(z)).$$

Exemplo 11 Seja X_i variáveis aleatórias com distribuição Bernoulli(p) e N com distribuição de Poisson(λ). Defina $S_N = X_1 + X_2 + \dots + X_N$.

$$\begin{aligned}
X_i \sim \text{Bernoulli}(p) &\rightarrow G_X(z) = q + pz \\
N \sim \text{Poisson}(\lambda) &\rightarrow G_N(z) = e^{\lambda(z-1)} \\
G_{S_N}(z) &= G_N(G_X(z)) \\
&= e^{\lambda((q+pz)-1)} = e^{\lambda p(z-1)} \\
G_{S_N}(z) &= e^{\lambda p(z-1)} \rightarrow S_N \sim \text{Poisson}(\lambda p)
\end{aligned}$$

2.3 Distribuições generalizadas

O importante conceito de soma de um número aleatório de variáveis aleatórias *i.i.d* nos leva ao seguinte conceito.

Considere X e Y variáveis aleatórias com funções geradoras de probabilidades $G_X(z)$ e $G_Y(z)$. Define-se a distribuição generalizada de X com Y como a variável aleatória $X \vee Y$ cuja função geradora de probabilidade é dada por $G_X(G_Y(z))$. Esse procedimento é também um mecanismo adequado para se gerar novas distribuições.

Uma aplicação deste conceito é descrito por Pielou (1977) e enunciada no exemplo abaixo. Suponha que fêmeas de insetos depositem grupos de ovos em unidades (folhas), representação na Figura 4. Considere que os grupos de indivíduos (larvas) constituem entidades tendo um padrão específico e que o número de indivíduos por grupo seja também uma variável aleatória com sua própria distribuição.

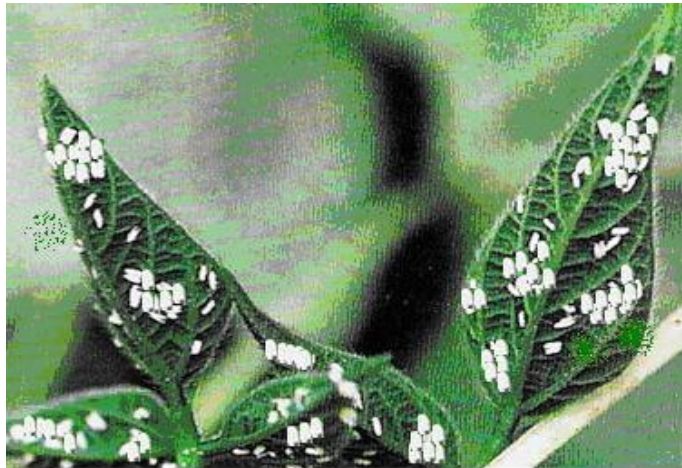


FIGURA 4 Padrão espacial agrupado.

Considerando $G_X(z)$ a função geradora de probabilidade relativa à variável X : número de grupos por unidade (folha), em que X é uma Poisson, e

$G_Y(z)$ a função geradora de probabilidade da variável Y : número de indivíduos por grupo, Y tendo distribuição logarítmica, a distribuição do número de indivíduos por unidade é dada por

$$X \sim \text{Poisson}(\lambda) \rightarrow G_X(z) = e^{\lambda(z-1)} \quad (\text{equação (2.6)})$$

$$Y \sim \text{Log}(p) \rightarrow G_Y(z) = \frac{\ln(1-qz)}{\ln(1-q)} \quad (\text{equação (2.7)})$$

$$G_X(G_Y(z)) = \exp \left\{ \lambda \left[\frac{\ln(1-qz)}{\ln(1-q)} - 1 \right] \right\}$$

Fazendo a reparametrização,

$$\lambda = -r \ln p \quad \text{em que } p = 1 - q$$

tem-se,

$$\begin{aligned} G_X(G_Y(z)) &= \exp \left\{ -r \ln p \left[\frac{\ln(1-qz)}{\ln(1-q)} - 1 \right] \right\} \\ &= \exp \left\{ -r \ln p \left[\frac{\ln(1-qz)}{\ln p} \right] \right\} \exp(r \ln p) \\ &= \exp[-r \ln(1-qz)] \exp(r \ln p) \\ &= (1-qz)^{-r} p^r = \left(\frac{p}{1-qz} \right)^r \end{aligned} \quad \#$$

que, como visto em (2.1), representa a função geradora de probabilidade da distribuição binomial negativa. Em termos de soma de um número aleatório de variáveis aleatórias, pode-se considerar que a distribuição binomial negativa resulta de uma soma de N variáveis aleatórias *i.i.d* com distribuição logarítmica(p) em que N tem distribuição Poisson(λ).

É interessante notar que, a distribuição binomial negativa pode ser obtida por dois processos diferentes, pela mistura das distribuições Poisson e gama, $\text{Poisson}(\lambda) \underset{\lambda}{\wedge} \text{Gama}(\alpha, \beta)$, como também pela generalizada das distribuições Poisson e logarítmica, $\text{Poisson}(\lambda) \vee \text{Log}(p)$. Pielou (1977) discute as consequências deste fato que serão analisadas na seção Resultados e Discussão.

Outra aplicação da distribuição generalizada surge se for considerado que, na distribuição espacial de larvas por unidade de área, a variação no número de grupos de ovos por unidade pode ser representada por uma distribuição de Poisson com parâmetro λ e o número de larvas que se desenvolve dentro de cada grupo como uma variável Poisson com parâmetro β .

$$X \sim \text{Poisson}(\lambda) \rightarrow G_X(z) = e^{\lambda(z-1)}$$

$$Y \sim \text{Poisson}(\beta) \rightarrow G_Y(z) = e^{\beta(z-1)}$$

A distribuição do número de larvas por unidade de área será dada então por $\text{Poisson}(\lambda) \vee \text{Poisson}(\beta)$, que é uma Neyman Tipo A ou Poisson-Poisson (Pielou, 1977).

A função geradora de probabilidade, definida por $H(z) = G_X(G_Y(z))$, é

$$H(z) = \exp\left\{\lambda\left[\exp(\beta(z-1))-1\right]\right\}.$$

Fazendo a expansão em série da função geradora de probabilidade, obtém-se a função densidade de probabilidade da distribuição Neyman tipo A, que é dada por

$$h(x) = \frac{e^{-\lambda}}{x!} \sum_{\beta=0}^{\infty} \frac{e^{-\beta} \beta^x \lambda^\beta}{\beta!} \quad x = 0, 1, 2, \dots$$

Os dois primeiros momentos da distribuição generalizada são,

$$\begin{aligned} E[X] &= G'(z)\Big|_{z=1} \\ &= \exp\left\{\lambda\left[\exp(\beta(z-1))-1\right]\right\} \lambda \left[\exp(\beta(z-1))\right] \beta \Big|_{z=1} = \lambda\beta \end{aligned}$$

Sendo $Var[X] = G''(1) + G'(1)[1 - G'(1)]$ e $G''(1) = (\lambda\beta)^2 + \lambda\beta^2$, obtém-se

$$Var[X] = (\lambda\beta)^2 + \lambda\beta^2 + \lambda\beta[1 - \lambda\beta] = \lambda\beta(1 + \beta). \quad \#$$

Definição 2 Considere X_1 e X_2 variáveis aleatórias com função de distribuição $F_1(x; \alpha)$, $F_2(x; \beta)$ respectivamente. Se para cada α existe um β tal que $F_1(x; \alpha) = F_2(x; \beta)$ qualquer que seja x , então as variáveis aleatórias X_1 e X_2 são ditas equivalentes, $X_1 \sim X_2$.

Para certas classes de distribuições, o resultado obtido pela mistura de $f(x; \theta) \underset{\theta}{\wedge} g(\theta)$ é essencialmente igual ao obtido pela generalizada de $g(\theta) \vee f(x; \theta)$. Tal fato é obtido pelo teorema.

Teorema 2 (Gurland, 1957) Seja X uma variável aleatória com função geradora de probabilidade $G_X(z) = [h(z)]^\theta$, onde θ é um parâmetro dado. Suponha que θ seja uma variável aleatória cY , tal que $\theta = cY$ onde c é uma constante e Y possui função densidade de probabilidade $f_Y(y)$ e função geradora de probabilidade $G_Y(z)$. Então, qualquer que seja a variável Y , a variável $X \wedge Y$ é equivalente à variável $Y \vee X$, $(X \wedge Y \sim Y \vee X)$.

Prova Tem-se que, $X \sim f(x; \theta)$, $G_X(z) = [h(z)]^\theta$. Considerando $\theta = cY$, $Y \sim g(y)$ e $G_Y(z)$ sua função geradora de probabilidade, a mistura das distribuições é dada por $h(x) = f(x) \underset{\theta/c}{\wedge} g(y) = \int_{-\infty}^{\infty} f(x; \theta) g(y) dy$.

Sendo $G_{X \wedge Y}(z)$ a função geradora de probabilidade da mistura, segue que

$$\begin{aligned} G_{X \wedge Y}(z) &= E[z^X] = \int_{-\infty}^{\infty} z^x h(x) dx = \int_{-\infty}^{\infty} z^x \left[\int_{-\infty}^{\infty} f(x; \theta) g(y) dy \right] dx \\ &= \int_{-\infty}^{\infty} g(y) \left[\int_{-\infty}^{\infty} z^x f(x; \theta) dx \right] dy = \int_{-\infty}^{\infty} g(y) [h(z)]^\theta dy \\ &= \int_{-\infty}^{\infty} [h(z)]^{cy} g(y) dy \end{aligned}$$

A função geradora de probabilidade de $Y \vee X$ é

$$\begin{aligned} G_Y(G_X(z)) &= G_Y\left([h(z)]^\theta\right) = E\left[\left([h(z)]^\theta\right)^y\right] \\ &= \int_{-\infty}^{\infty} \left([h(z)]^\theta\right)^y g(y) dy = \int_{-\infty}^{\infty} [h(z)]^{\theta y} g(y) dy \end{aligned}$$

Assim, quando $c = \theta$ temos que $X \wedge Y \sim Y \vee X$. #

Exemplo 12 Seja X uma variável aleatória com distribuição binomial negativa, $BN(r, p)$. Considere r uma variável aleatória com distribuição gama(α, β). Temos,

$$BN(r, p) \underset{r}{\wedge} \text{Gama}(\alpha, \beta)$$

$$X \sim BN(r, p) \rightarrow G_X(z) = \left(\frac{p}{1-qz}\right)^r \quad (\text{equação (2.1)})$$

Considere $r = cy$ em que $Y \sim \text{Gama}(\alpha, \beta)$. A função geradora de probabilidade da mistura é dada por

$$\begin{aligned} G(z) &= \int_0^{\infty} \left(\frac{p}{1-qz}\right)^{cy} \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y} dy \\ &= \int_0^{\infty} e^{cy[\ln p - \ln(1-qz)]} \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y} dy \end{aligned}$$

$$\begin{aligned}
&= \beta^\alpha \int_0^\infty \frac{1}{\Gamma(\alpha)} y^{\alpha-1} e^{-[\beta+c\ln(1-qz)-\ln p]y} dy \\
&= \left[\frac{\beta}{\beta+c\ln(1-qz)-\ln p} \right]^\alpha = \left[\frac{\beta}{\beta-c\ln\left(\frac{p}{1-qz}\right)} \right]^\alpha.
\end{aligned}$$

$$\text{Gama}(\alpha, \beta) \vee \text{BN}(r, p)$$

$$Y \sim \text{Gama}(\alpha, \beta) \rightarrow G_Y(z) = \left(\frac{\beta}{\beta - \ln z} \right)^\alpha \quad (\text{equação (2.8)})$$

$$X \sim \text{BN}(r, p) \rightarrow G_X(z) = \left(\frac{p}{1-qz} \right)^r$$

Então,

$$G_Y(G_X(z)) = G_Y\left(\left(\frac{p}{1-qz}\right)^r\right) = \left[\frac{\beta}{\beta - \ln\left(\left(\frac{p}{1-qz}\right)^r\right)} \right]^\alpha = \left[\frac{\beta}{\beta - r \ln\left(\frac{p}{1-qz}\right)} \right]^\alpha$$

Portanto, quando $r = c$ tem-se que

$$\text{BN}(r, p) \underset{r}{\wedge} \text{Gama}(\alpha, \beta) \sim \text{Gama}(\alpha, \beta) \vee \text{BN}(r, p). \quad \#$$

O uso das distribuições generalizadas torna-se muito útil como um recurso auxiliar na obtenção de determinadas misturas. Como para certas

distribuições, fazer a mistura é equivalente a fazer a generalizada, usa-se este artifício para evitar cálculos muito complexos.

Exemplo 13 Considere novamente a situação descrita no exemplo 3, porém, admita que o número de ovos observados em cada grupo i é também uma variável aleatória. Adote a extensão da variável binomial feita no exemplo 4.

Em cada grupo i , a quantidade de larvas será o resultado de n ensaios de Bernoulli, assim, o total de ovos Y_i que eclodem em cada grupo será dado por uma binomial, $Y_i \sim \text{Bin}(n, p)$. Então, em cada unidade tem-se uma soma de Binomiais e o total de larvas é dado por $Y_1 + Y_2 + \dots + Y_N$. Como cada Y_i representa uma soma de variáveis Bernoulli, pode-se representar o total de larvas por $X_1 + X_2 + \dots + X_{nN}$, que é uma soma aleatória de variáveis Bernoulli.

$$X_i \sim \text{Bernoulli}(p) \rightarrow G_x(z) = q + pz$$

Considere que o número de termos desta soma segue uma distribuição Poisson (λ) . Portanto,

$$\begin{aligned} G_{nN}(z) &= \sum_{k=0}^{\infty} P[nN = k] z^k = \sum_{k=0}^{\infty} P[N = k/n] z^k \\ &= \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^{k/n}}{(k/n)!} z^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k/n}}{(k/n)!} z^k \end{aligned}$$

Fazendo $s = k/n$ tem-se que, quando $s = k/n = 0 \rightarrow \infty$ ($s = 0, 1, 2, 3, \dots$), $k = 0 \rightarrow \infty$ ($k = 0, n, 2n, 3n, \dots$). Então,

$$G_{nN}(z) = e^{-\lambda} \sum_{s=0}^{\infty} \frac{\lambda^s}{s!} z^{sn} = e^{-\lambda} \sum_{s=0}^{\infty} \frac{(\lambda z^n)^s}{s!}$$

$$= e^{-\lambda} e^{\lambda z^n} = e^{\lambda(z^n - 1)}$$

$$G_{nN}(G_X(z)) = G_{nN}(q + pz) = e^{\lambda[(q-pz)^n - 1]} \quad \#$$

$G_{nN}(G_X(z))$ representa a função geradora de probabilidade da distribuição do número de larvas por unidade quando é assumido que N é uma variável Poisson.

2.4 A mistura das distribuições binomial negativa e beta

Considere uma sequência de ensaios de Bernoulli realizada de forma independente com probabilidade p constante e defina X como o número de fracassos anteriores ao r -ésimo sucesso. A variável aleatória X segue o modelo binomial negativo com parâmetros r e p , em que $0 < p < 1$ e $r > 0$, e tem função densidade de probabilidade dada por:

$$f(x; r, p) = \binom{r+x-1}{x} p^r (1-p)^x I_{\{0,1,2,\dots\}}(x) \quad (2.12)$$

Quando r é um número inteiro positivo, (2.12) pode ser interpretada como a distribuição do tempo de espera para a ocorrência do r -ésimo sucesso, sendo também conhecida por distribuição Pascal (Feller, 1968). Para $r = 1$ ela se reduz a distribuição geométrica. A distribuição binomial negativa (BN) apresenta uma propriedade importante, a variância excede a média. Essa característica é, às vezes, referida como superdispersão.

A distribuição binomial negativa tem exercido um importante papel para descrever a distribuição de pragas (insetos) em lavouras de determinadas culturas (Bearzoti, 1998). Dado uma plantação, define-se uma unidade espacial na qual conta-se o número de insetos presentes na unidade. Sabe-se que a presença de insetos em uma determinada unidade é afetada pela maior ou menor densidade nas unidades vizinhas (isso ocorre em geral com pulgões e colchonilhas). Se a presença de indivíduos em uma unidade depende das outras unidades, a pressuposição de que p é constante é violada, tornando a aplicação da binomial negativa inadequada. Uma forma de contornar o problema da variabilidade de p é supor que p varia segundo uma distribuição beta

$$g(p) = \frac{1}{B(a,b)} p^{a-1} (1-p)^{b-1} I_{(0,1)}(p).$$

Neste caso, obtém-se como mistura a distribuição discreta beta binomial

negativa (BBN) dada por $h(x; a, b, r) = \int_0^1 f(x; p) g(p) dp$.

$$\begin{aligned} h(x) &= \int_0^1 \binom{r+x-1}{x} p^r q^x \frac{1}{B(a,b)} p^{a-1} (1-p)^{b-1} dp \\ &= \binom{r+x-1}{x} \frac{1}{B(a,b)} \int_0^1 p^{r+a-1} (1-p)^{x+b-1} dp \\ &= \binom{r+x-1}{x} \frac{B(r+a, x+b)}{B(a,b)} \int_0^1 \frac{1}{B(r+a, x+b)} p^{r+a-1} (1-p)^{x+b-1} dp \\ &= \binom{r+x-1}{x} \frac{B(r+a, x+b)}{B(a,b)} I_{\{0,1,2,\dots\}}(x) \end{aligned} \quad (2.13) \#$$

Utilizando-se a função beta em termos da função gama $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$,
 $a > 0$ e $b > 0$ e as relações $\Gamma(t+1) = t\Gamma(t)$ e $\Gamma(t+1) = t!$, para t inteiro,
 equação (2.13) pode ser expressa em termos da função gama,

$$h(x; a, b, r) = \binom{r+x-1}{x} \frac{\Gamma(r+a)\Gamma(x+b)\Gamma(a+b)}{\Gamma(a+b+r+x)\Gamma(a)\Gamma(b)} \quad (2.14)$$

Considere $x^{[n]} = x(x+1)(x+2)\dots(x+n-1) = \frac{\Gamma(x+n)}{\Gamma(x)}$ em que $x^{[n]}$
 representa um fatorial ascendente. A equação (2.14) pode ser reescrita como,

$$\begin{aligned} h(x; a, b, r) &= \binom{r+x-1}{x} \frac{a(a+1)\dots(a+(r-1))b(b+1)\dots(b+(x-1))}{(a+b)(a+b+1)\dots(a+b+(r+x-1))} \\ &= \binom{r+x-1}{x} \frac{\prod_{i=0}^{r-1} (a+i) \prod_{i=0}^{x-1} (b+i)}{\prod_{i=0}^{r+x-1} (a+b)} \end{aligned}$$

que, também, pode ser representada em termos de fatoriais ascendentes

$$h(x; a, b, r) = \frac{r^{[x]} a^{[r]} b^{[x]}}{x! (a+b)^{[r+x]}} \quad (2.15)$$

Exemplo 14 Seja X uma variável aleatória com distribuição beta binomial negativa ($a = 3, b = 5, r = 7$).

$$h(x; 3, 5, 7) = \frac{\Gamma(10)\Gamma(x+5)\Gamma(8)}{\Gamma(x+15)\Gamma(3)\Gamma(5)} = \frac{(9!).105}{(x+14)(x+13)\dots(x+5)}$$

O gráfico de $h(x; 3, 5, 7)$ encontra-se na Figura 5.

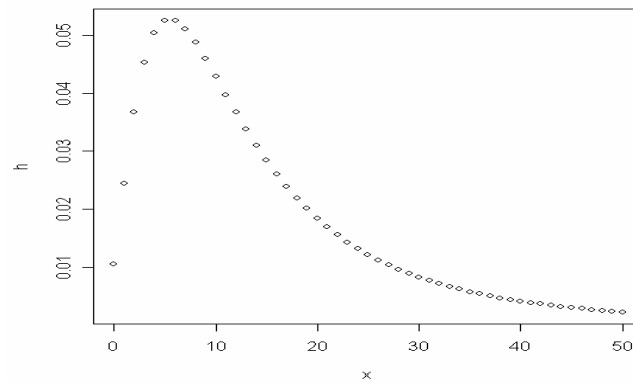
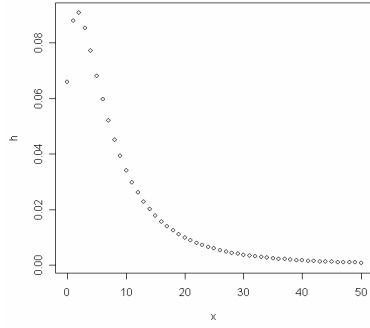
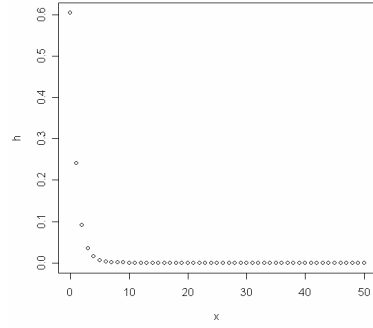


FIGURA 5 Beta binomial negativa ($3, 5, 7$).

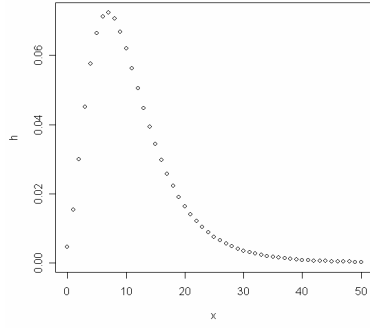
Outras representações gráficas da distribuição beta binomial negativa (BBN) são dadas na Figura 6.



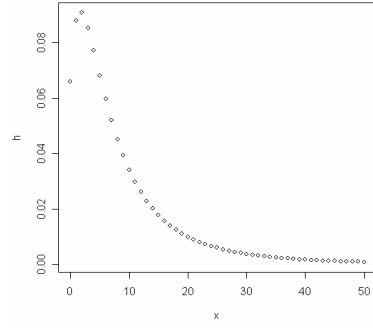
a) BBN ($a = 3, b = 10, r = 2$)



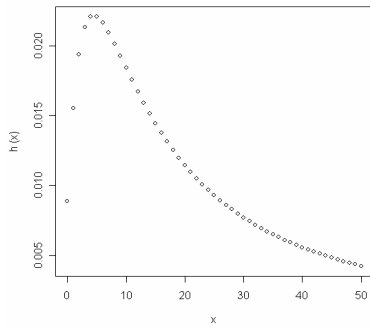
b) BBN ($a = 10, b = 3, r = 2$)



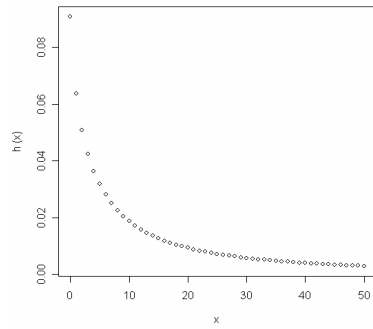
c) BBN ($a = 10, b = 10, r = 10$)



d) BBN ($a = 3, b = 2, r = 10$)



e) BBN ($a = 0,6, b = 5, r = 3$)



f) BBN ($a = 0,6, b = 0,8, r = 10$)

FIGURA 6 Gráficos da distribuição beta binomial negativa para diferentes valores dos parâmetros a , b e r .

2.4.1 Função geradora de momentos fatoriais da BBN

Tem-se que $m_{(l)} = E \left[X^{(l)} \right] = \int_0^1 E \left[X_p^{(l)} \right] g(p) dp$ (equação (2.9)).

Os momentos fatoriais da mistura são obtidos como uma média dos momentos fatoriais da distribuição original $m_{(l)} = E \left[X_p^{(l)} \right]$ é o l -ésimo momento fatorial da distribuição binomial negativa. Tem-se que:

$$G_X(z) = \left(\frac{p}{1-qz} \right)^r \quad (\text{equação (2.1)})$$

Diferenciando $G(z)$ e tomando $\lim_{z \rightarrow 1} G^{(l)}(z)$ obtém-se

$$\frac{d}{dz} G(z) = rqp^r (1-qz)^{-(r+1)}$$

$$\frac{d^2}{dz^2} G(z) = r(r+1)q^2 p^r (1-qz)^{-(r+2)}$$

$$\begin{array}{cc} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{array}$$

$$\frac{d^l}{dz^l} G(1) = r(r+1)(r+2)\dots(r+(l-1))q^l p^r (1-qz)^{-(r+l)} \Big|_{z=1}$$

$$\frac{d^l}{dz^l} G(1) = r(r+1)\dots(r+(l-1)) \left(\frac{q}{p} \right)^l = r^{[l]} \left(\frac{q}{p} \right)^l = m_{(l)}^{BN}(r, p).$$

Sendo, $m_{(l)}^{BBN}(a, b, r) = \int_0^1 E \left[X_p^{(l)} \right] g(p) dp$, segue que

$$\begin{aligned}
&= \int_0^1 r^{[l]} \left(\frac{q}{p} \right)^l \frac{1}{B(a, b)} p^{a-1} (1-p)^{b-1} dp \\
&= r^{[l]} \frac{1}{B(a, b)} \int_0^1 p^{(a-l)-1} (1-p)^{(b+l)-1} dp \\
&= r^{[l]} \frac{1}{B(a, b)} B(a-l, b+l) \int_0^1 \frac{1}{B(a-l, b+l)} p^{(a-l)-1} (1-p)^{(b+l)-1} dp
\end{aligned}$$

e, portanto, o l -ésimo momento fatorial da distribuição beta binomial negativa é dado por

$$m_{(l)}^{BBN}(a, b, r) = r^{[l]} \frac{B(a-l, b+l)}{B(a, b)} \quad (2.16)$$

2.4.2 Média e variância

Como $m_{(l)}^{BBN}(a, b, r) = r^{[l]} \frac{B(a-l, b+l)}{B(a, b)}$.

$$m_{(1)}^{BBN} = G'(1) = E(X)$$

$$\begin{aligned}
&= r \frac{B(a-1, b+1)}{B(a, b)} = r \left[\frac{\Gamma(a-1)\Gamma(b+1)}{\Gamma(a+b)} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \right] \\
&= \frac{rb}{(a-1)} \quad \text{para } a > 1.
\end{aligned}$$

É dado que, $Var[X] = G''(1) + G'(1)[1 - G'(1)]$. Se

$$\begin{aligned}
 G''(1) &= r(r+1) \frac{B(a-2, b+2)}{B(a, b)} \\
 &= r(r+1) \left[\frac{\Gamma(a-2)\Gamma(b+2)}{\Gamma(a+b)} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \right] \\
 &= r(r+1) \left[\frac{\Gamma(a-2)(b+1)b\Gamma(b)}{\Gamma(a+b)} \frac{\Gamma(a+b)}{(a-1)(a-2)\Gamma(a-2)\Gamma(b)} \right] \\
 &= \frac{rb(r+1)(b+1)}{(a-1)(a-2)}
 \end{aligned}$$

obtém-se

$$\begin{aligned}
 Var[X] &= \frac{rb(r+1)(b+1)}{(a-1)(a-2)} + \frac{rb}{(a-1)} - \left(\frac{rb}{a-1} \right)^2 \\
 &= \frac{rb[(r+a-1)(a+b-1)]}{(a-1)^2(a-2)} \quad \text{para } a > 2.
 \end{aligned}$$

Observa-se que, assim como a distribuição binomial negativa, a beta binomial negativa apresenta variância maior que a média.

Diferente do exposto em Hassan & Bilal (2008), a distribuição beta binomial negativa, referida pelos autores como “Negative Polya-Eggenberger Distribution”, não apresenta todos os momentos. Os momentos existem somente até a ordem $l < a$, onde l representa o momento a ser calculado e a , um dos parâmetros da distribuição. A distribuição beta binomial negativa $(2, 5, 3)$, por exemplo, apresenta apenas o primeiro momento.

2.5 Unimodalidade

Definição 3 Uma variável aleatória X é dita unimodal se existe um valor x_0 tal que $f_X(x)$ é estritamente crescente à esquerda de x_0 e estritamente decrescente à direita de x_0 . Caso $f_X(x)$ seja diferenciável, $f'_X(x_0) = 0$. O valor x_0 é chamado de moda.

2.5.1 Unimodalidade da distribuição beta

Seja P uma variável aleatória com distribuição beta

$$f(p) = f_P(p; a, b) = \frac{1}{B(a, b)} p^{a-1} (1-p)^{b-1} I_{(0,1)}(p).$$

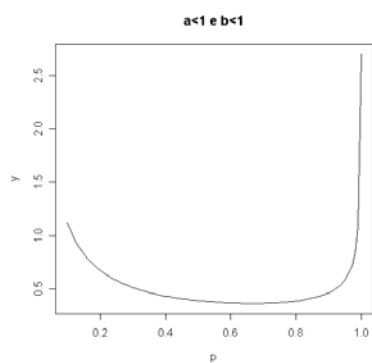
Como $f(p)$ é diferenciável, tem-se

$$\ln(f(p)) = \ln\left[\frac{1}{B(a, b)}\right] + (a-1)\ln p + (b-1)\ln(1-p)$$

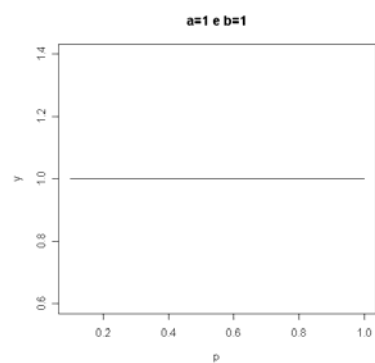
$$\frac{d \ln(f(p))}{dp} = \frac{a-1}{p} - \frac{b-1}{1-p} = 0$$

$$(a-1)(1-p) - (b-1)p = 0 \rightarrow p = \frac{a-1}{a+b-2} \quad a, b > 1.$$

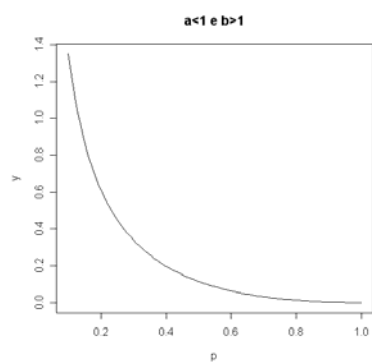
Um estudo desses pontos críticos nos dá que a distribuição é unimodal para $a > 1$ e $b > 1$. Na Figura 7 ilustram-se alguns possíveis casos da distribuição beta para diferentes valores dos parâmetros a e b .



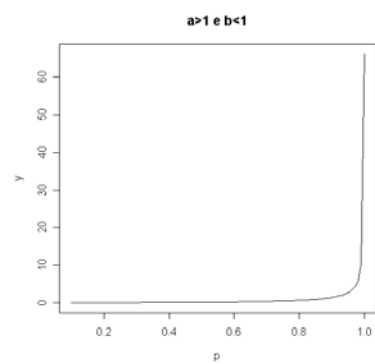
a) Beta ($a = 0,2, b = 0,6$)



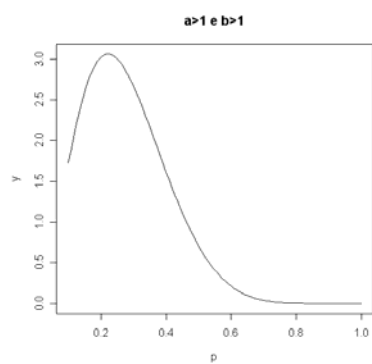
b) Beta ($a = 1, b = 1$)



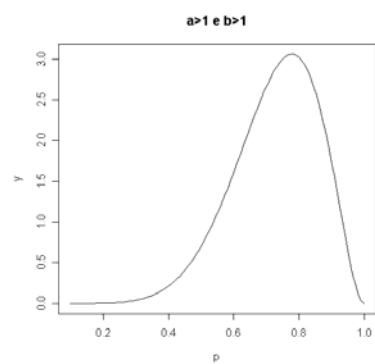
c) Beta ($a = 0,2, b = 3$)



d) Beta ($a = 3, b = 0,2$)



e) Beta ($a = 3, b = 8$)



f) Beta ($a = 8, b = 3$)

FIGURA 7 Gráficos da distribuição beta.

Na Figura 7, itens de **a** até **d**, a distribuição beta não apresenta moda. Porém, nos itens **e** e **f**, a distribuição é unimodal e a moda ocorre em $p = \frac{a-1}{a+b-2}$. Na beta (3,8) a moda é $p = 0.22$ e na beta (8,3) em $p = 0.78$.

2.5.2 Unimodalidade da distribuição binomial negativa

Considere X uma variável aleatória com distribuição binomial negativa,

$$f_X(x; r, p) = \binom{r+x-1}{x} p^r (1-p)^x I_{\{0,1,2,\dots\}}(x).$$

Se x_0 é uma moda da distribuição, tem-se:

$$\begin{cases} \frac{P(x_0)}{P(x_0-1)} \geq 1 \\ \frac{P(x_0+1)}{P(x_0)} < 1 \end{cases}$$

Resolvendo as desigualdades,

$$\frac{P(x_0)}{P(x_0-1)} \geq 1 \rightarrow \frac{\binom{r+x_0-1}{x_0} p^r (1-p)^{x_0}}{\binom{r+x_0-2}{x_0-1} p^r (1-p)^{x_0-1}} = \frac{(r+x_0-1)q}{x_0} \geq 1$$

e

$$\frac{P(x_0+1)}{P(x_0)} < 1 \rightarrow \frac{\binom{r+x_0}{x_0+1} p^r (1-p)^{x_0+1}}{\binom{r+x_0-1}{x_0} p^r (1-p)^{x_0}} = \frac{(r+x_0)(1-p)}{(x_0+1)} < 1$$

obtém-se

$$\begin{cases} \frac{P(x_0+1)}{P(x_0)} < 1 & \text{se } x_0 > \frac{rq-1}{p} \\ \frac{P(x_0)}{P(x_0-1)} \geq 1 & \text{se } x_0 \leq \frac{(r-1)q}{p} \end{cases}$$

Portanto a moda da distribuição binomial negativa ocorre no valor x_0 tal que

$$\frac{rq-1}{p} < x_0 \leq \frac{(r-1)q}{p}.$$

Uma análise desses pontos críticos realizada por Johnson et al., (1992), mostra que:

- Quando $\frac{(r-1)q}{p}$ não é um número inteiro, exemplo na Figura 8, a distribuição é unimodal;

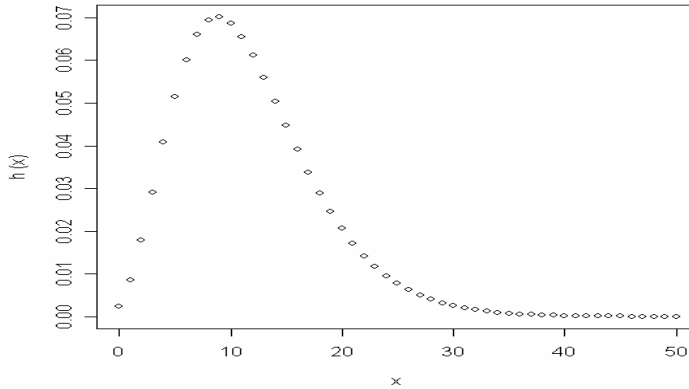


FIGURA 8 Binomial negativa $(r = 5, p = 0,3)$.

- Quando $\frac{(r-1)q}{p}$ é inteiro, a distribuição apresenta duas modas nos pontos $x_0 = \frac{(r-1)q}{p}$ e $x_0 = \frac{rq-1}{p}$, Figura 9.

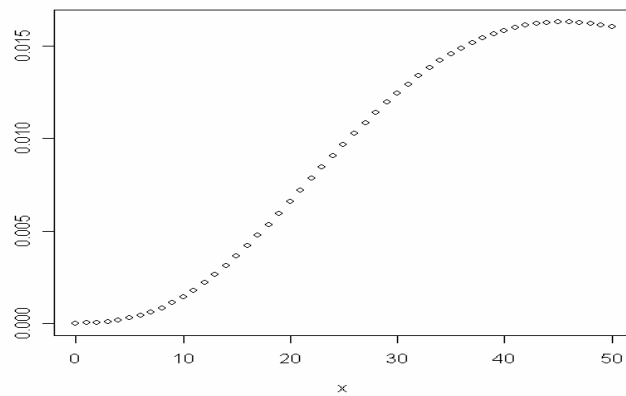


FIGURA 9 Binomial negativa ($r = 5, p = 0,08$).

Na distribuição binomial negativa (5, 0,08) os pontos $x_0 = 45$ e $x_0 = 46$ são os valores da moda da distribuição.

- Quando $rq < 1$, o ponto $x_0 = 0$ é a moda, Figura 10.

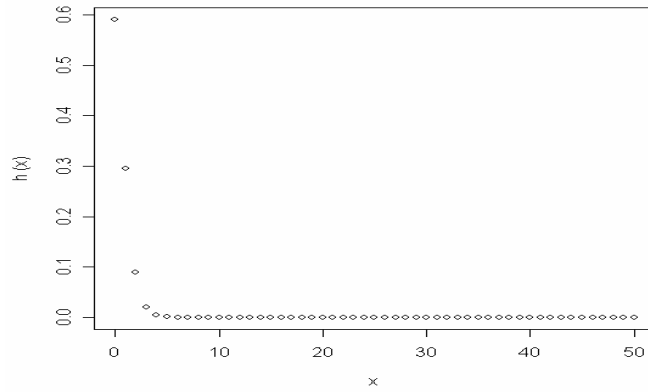


FIGURA 10 Binomial negativa $(r = 5, p = 0,9)$.

Outra maneira interessante de se obter a moda da binomial negativa é utilizar o fato de que a função gama generaliza o fatorial permitindo assim uma extensão contínua. Desta forma, é possível derivar obtendo-se uma aproximação do valor da moda.

$$f_X(x; r, p) = \binom{r+x-1}{x} p^r (1-p)^x I_{\{0,1,2,\dots\}}(x)$$

$$= \frac{\Gamma(r+x)}{\Gamma(r)\Gamma(x+1)} p^r (1-p)^x I_{\{0,1,2,\dots\}}(x)$$

$$\ln(f(x)) = \ln(\Gamma(r+x)) - \ln(\Gamma(r)) - \ln(\Gamma(x+1)) + r \ln p + x \ln(1-p)$$

Sendo

$$\frac{\partial \ln \Gamma(x+r)}{\partial x} = \psi(x+r) \quad (\text{ver apêndice}),$$

tem-se

$$\begin{aligned}
\frac{\partial \ln(f(x))}{\partial x} &= \psi(r+x) - \psi(x+1) + \ln(1-p) = 0 \\
&= \psi(x) + \sum_{k=1}^r (x+k-1)^{-1} - \psi(x) - \frac{1}{x} + \ln(1-p) = 0 \\
&= \sum_{k=1}^r (x+k-1)^{-1} - \frac{1}{x} + \ln(1-p) \\
&= \sum_{k=2}^r (x+k-1)^{-1} + \ln(1-p)
\end{aligned}$$

Conhecendo-se os valores dos parâmetros r e p é possível, com o auxílio de um software, obter uma aproximação para o valor da moda.

2.5.3 Unimodalidade de misturas

Holgate (1970) discute a unimodalidade de algumas misturas envolvendo a distribuição de Poisson.

Teorema 3 Seja $g(\lambda)$ a função densidade de probabilidade de uma variável aleatória positiva, contínua e unimodal.

A variável aleatória com função densidade de probabilidade dada por

$$h(x) = \frac{1}{x!} \int_0^{\infty} e^{-\lambda} \lambda^x g(\lambda) d\lambda$$

é uma variável unimodal.

Baseado no **Teorema 3** confirma-se novamente que a distribuição binomial negativa é unimodal pois

Binomial Negativa $\sim$ Poisson $\wedge$ Gama

e a distribuição gama assume valores positivos, é continua e unimodal.

A unimodalidade da distribuição beta binomial negativa é estudada por Hassan & Bilal (2008). Os autores justificam a unimodalidade da distribuição com base nos resultados obtidos por Holgate (1970). No entanto, para tal afirmação ser um corolário do **Teorema 3**, faz-se necessário provar a unimodalidade da distribuição resultante da mistura das distribuições gama e beta, pois

Beta Binomial Negativa $\sim$ Binomial Negativa $\wedge$ Beta

porém,

Binomial Negativa $\sim$ Poisson $\wedge$ Gama

então,

Beta Binomial Negativa $\sim$ Poisson $\wedge$ Gama $\wedge$ Beta.

Provar a unimodalidade da mistura das distribuições gama e beta parece ser um problema mais difícil que a própria prova da unimodalidade da beta binomial negativa. Os autores não apresentam uma discussão sobre essa questão. Um intervalo contendo a moda da beta binomial negativa é obtido no teorema abaixo.

Teorema 4 (Hassan & Bilal, 2008) A distribuição beta binomial negativa é unimodal para todos os valores de (a, b, r) e a moda ocorre em $x = 0$ se $rb < 1$ e para $rb > 1$ a moda é algum outro ponto $x = M$ tal que

$$\frac{r(b-1)-(a+b)}{(a+1)} < M < \frac{(r-1)(b+1)}{(a+1)}.$$

Prova Seja $h(x; a, b, r) = \binom{r+x-1}{x} \frac{B(r+a, x+b)}{B(a, b)} I_{\{0,1,2,\dots\}}(x)$.

A razão

$$\frac{P(x+1)}{P(x)} = \frac{(r+x)(b+x)}{(x+1)(a+b+r+x)} \quad (2.17)$$

é menor que um, isto é

$$\frac{P(x+1)}{P(x)} < 1 \text{ se } rb < 1 \quad \forall (a, b, r) > 0.$$

Então, para $rb < 1$, a razão $\frac{P(x+1)}{P(x)}$ é uma função monótona decrescente e a

moda existe e ocorre no ponto $x = 0$.

Suponha que para a moda seja igual a $x = M$. Então, (2.17) fornece duas desigualdades

$$\left\{ \begin{array}{l} \frac{P(M+1)}{P(M)} < 1 \\ \frac{P(M)}{P(M-1)} > 1 \end{array} \right\} \rightarrow \left\{ \begin{array}{l} \frac{(M+r)}{(M+1)} \frac{(M+b)}{(a+b+r+M)} < 1 \\ \frac{(M+r-1)}{M} \frac{(M+b-1)}{(a+b+r+M-1)} > 1 \end{array} \right.$$

Resolvendo o sistema em M , obtêm-se as seguintes desigualdades

$$\frac{r(b-1)-(a+b)}{(a+1)} < M \text{ e } \frac{(r-1)(b+1)}{(a+1)} > M .$$

Portanto,

$$\frac{r(b-1)-(a+b)}{(a+1)} < M < \frac{(r-1)(b+1)}{(a+1)} . \quad \#$$

3 MATERIAL E MÉTODOS

3.1 Dados

Os dados utilizados no trabalho foram analisados anteriormente pela distribuição binomial negativa. Um estudo semelhante é realizado utilizando-se a beta binomial negativa. Esses dados são relativos à contagem do número de número de Cáries por criança, Tabela 1, (Gurland, 1959); número de acidentes por maquinista, Tabela 2, (Bliss & Fisher, 1953); número de plantas *Salicornia stricta* por quadrado, Tabela 3, (Bliss & Fisher 1953) e número de ácaros por folha, Tabela 4, (Bliss & Fisher 1953).

3.2 Simulação de distribuições obtidas por misturas

Foram desenvolvidas rotinas em R, versão 2.8.0, com o objetivo de implementar o algoritmo mistura, Devroye (1986). Uma validação matemática para o algoritmo é apresentada posteriormente.

Algoritmo mistura:

- 1) Um valor do parâmetro θ é simulado a partir de $g(\theta)$;
- 2) Com o valor de θ obtido no passo anterior, um valor x é simulado a partir de $f(x; \theta)$;
- 3) Os passos (1) e (2) são repetidos n vezes resultando na amostra aleatória de tamanho n .

3.3 Estimação e ajuste

Um sistema de equações de máxima verossimilhança aproximado é obtido e os valores estimados para cada conjunto de dados são determinados

com o auxílio do software Maple versão 11.

A estatística utilizada para testar a qualidade do ajuste da distribuição beta binomial negativa aos dados é dada por

$$\chi_c^2 = \sum_{i=1}^k \frac{(Fo_i - Fe_i)^2}{Fe_i}$$

em que

$Fe_i \rightarrow$ Frequência esperada de elementos na categoria i ;

$Fo_i \rightarrow$ Frequência observada de elementos na categoria i ;

$k \rightarrow$ Número de categorias.

A distribuição da estatística χ_c^2 é aproximada por uma distribuição qui-quadrado com $(k-1)$ graus de liberdade. Quando o modelo a ser testado envolve a estimação de parâmetros desconhecidos χ_c^2 tem distribuição aproximada de qui-quadrado com $(k-t-1)$ graus de liberdade, em que t é o número de parâmetros estimados no modelo pelo método da máxima verossimilhança.

Um teste de ajuste é realizado considerando-se um nível de significância $\alpha = 0,05$.

Alguns autores, dentre eles Bliss & Fisher (1953), recomendam que as frequências esperadas nas categorias não sejam inferiores a 5. Caso isso ocorra, as categorias devem ser agrupadas em uma única categoria antes do cálculo de χ_c^2 . Desse modo, o número de graus de liberdade, para os dados das Tabelas 2, 3 e 4, é calculado segundo essa metodologia.

4 RESULTADOS E DISCUSSÃO

4.1 Simulando misturas

Uma validação matemática do algoritmo mistura é apresentada no teorema.

Teorema 5 O algoritmo mistura simula uma amostra da distribuição

$$h(x) = \int_{\Theta} f(x; \theta) g(\theta) d\theta .$$

Prova Suponha Θ e o domínio de $f(x; \theta)$ intervalos reais finitos, representados na Figura 11.

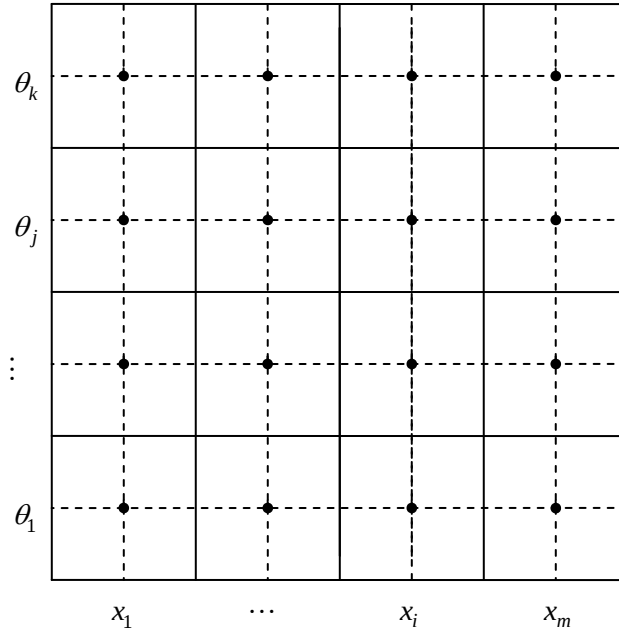


FIGURA 11 Discretização do parâmetro θ e da variável x .

Considere em Θ uma partição π em k subintervalos de tamanho $\Delta\theta$, nos quais toma-se o ponto central de cada subintervalos que serão indexados por $\theta_1, \dots, \theta_j, \dots, \theta_k$. Da mesma forma, uma partição η para o domínio de $f(x; \theta)$ é realizada, obtendo-se os pontos $x_1, \dots, x_i, \dots, x_m$. Assim, podem-se obter as distribuições discretas

$$P[\Theta = \theta_j] = \left\{ \frac{g(\theta_j) \Delta\theta}{S_\pi}, j = 1, \dots, k \right\}, \text{ em que } S_\pi = \sum_{j=1}^k g(\theta_j) \Delta\theta,$$

De forma similar, obtém-se

$$P[X = x_i] = \left\{ \frac{f(x_i; \theta_j) \Delta x}{S_\eta}, i = 1, \dots, m \right\}, \text{ em que } S_\eta = \sum_{i=1}^m f(x_i; \theta_j) \Delta x.$$

Essas distribuições aproximam as densidades $g(\theta)$ e $f(x; \theta_j)$. Portanto a distribuição $h(x)$ é aproximada por

$$\frac{1}{S_\pi S_\eta} \sum_j \left\{ f(x_i; \theta_j) g(\theta_j) \right\} \Delta x \Delta \theta, \quad i = 1, \dots, m$$

Utilizando a distribuição $\left\{ \frac{g(\theta_j) \Delta\theta}{S_\pi}, j = 1, \dots, k \right\}$, uma amostra de

tamanho n , com $n > k$ e $n > m$, é gerada. Ordenando-se essa amostra e observando que valores repetidos de θ_j são obtidos, tem-se, pela lei dos grandes números (Feller, 1968), que o número de vezes que um determinado θ_j aparece

na amostra é aproximadamente igual a $\left(\frac{g(\theta_j)\Delta\theta}{S_\pi} \right) n$. Tem-se então que

$$\underbrace{\overbrace{\theta_1 \quad \theta_1 \quad \cdots \quad \theta_1}^{S_\pi} \quad \overbrace{\theta_2 \quad \theta_2 \quad \cdots \quad \theta_2}^{S_\pi} \quad \cdots \quad \overbrace{\theta_k \quad \theta_k \quad \cdots \quad \theta_k}^{S_\pi}}_n$$

Para cada valor de θ_j , um valor x_i é simulado utilizando-se a distribuição $f(x_i; \theta_j)\Delta x$. Valores de x_i ocorrem repetidas vezes. Novamente, pela Lei dos grandes números que o número de vezes que um valor x_i aparece

na amostra de tamanho n é aproximadamente $\frac{\left(f(x_i; \theta_j)\Delta x \right) \left(g(\theta_j)\Delta\theta \right) n}{S_\pi S_\eta}$.

$$\begin{array}{ccccccc} & & \overbrace{S_\pi}^{(g(\theta_j)\Delta\theta)n} & & & & \\ & & \underbrace{\theta_j \quad \cdots \quad \theta_j} & \underbrace{\theta_j \quad \cdots \quad \theta_j} & \cdots & \underbrace{\theta_j \quad \cdots \quad \theta_j} & \\ & \downarrow & & \downarrow & & \downarrow & \downarrow \\ & \underbrace{x_1 \quad \cdots \quad x_1} & & \underbrace{x_2 \quad \cdots \quad x_2} & \cdots & \underbrace{x_m \quad \cdots \quad x_m} & \\ & \underbrace{(f(x_1; \theta_j)\Delta x)(g(\theta_j)\Delta\theta)n}_{S_\pi S_\eta} & & \underbrace{(f(x_2; \theta_j)\Delta x)(g(\theta_j)\Delta\theta)n}_{S_\pi S_\eta} & & \underbrace{(f(x_m; \theta_j)\Delta x)(g(\theta_j)\Delta\theta)n}_{S_\pi S_\eta} & \end{array}$$

Tal somatório é exatamente a soma de Riemann da função de duas variáveis $f(x|\theta)g(\theta)$ em relação à partição da região no plano definida por π e η . Fazendo $\Delta\theta \rightarrow 0$ e $\Delta x \rightarrow 0$, obtém-se

$$\lim_{j \rightarrow \infty} \left(\frac{1}{S_\pi S_\eta} \sum_j \{ f(x_i; \theta_j) g(\theta_j) \} \Delta \theta \right) \Delta x \approx \left(\int_{\Theta} f(x_i; \theta) g(\theta) d\theta \right) \Delta x = h(x_i) \Delta x$$

isto é, o número de vezes que x_i aparece na amostra é aproximadamente igual a $h(x_i) \Delta x$, de onde o resultado segue.

No caso em que θ e o domínio de $f(x; \theta)$ não são limitadas, (por exemplo, a média de uma normal), basta truncar as distribuições $g(\theta)$ e $f(x; \theta)$. isto é, o número de vezes que x_i aparece na amostra é aproximadamente igual a $h(x_i) \Delta x$. Fazendo $\Delta x \rightarrow 0$, o resultado segue.

4.2 Estimadores de máxima verossimilhança

Hassan & Bilal (2008) propõem um sistema de equações de verossimilhança aproximadas. No entanto, foram detectados alguns erros na dedução das fórmulas. A saber, foi utilizada a aproximação

$$\psi(x) \approx \ln(\Gamma(x)) - \frac{1}{2x}.$$

As aproximações adequadas são $\psi(x) \approx \ln x - \frac{1}{2x}$ ou $\psi(x) \approx \ln(x - 0,5)$.

E também no uso da relação de recorrência

$$\psi(x+n) \approx \psi(x) + \sum_{k=1}^n (x+k-1)^{-1}.$$

Neste trabalho, foram deduzidas novas aproximações das equações de máxima verossimilhança.

4.2.1 Sistema de equações aproximadas para o cálculo dos estimadores de máxima verossimilhança

Seja X uma variável aleatória $BBN(a, b, r)$ com função densidade de probabilidade dada por

$$h(x) = \binom{r+x-1}{x} \frac{B(r+a, x+b)}{B(a, b)} = \frac{r^{[x]} a^{[r]} b^{[x]}}{x! (a+b)^{[r+x]}} \mathbf{I}_{\{0,1,2,\dots\}}(x) \quad (4.1)$$

Porém,

$$(a+b)^{[r+x]} = \underbrace{(a+b)(a+b+1)(a+b+2)\dots(a+b+r-1)}_{(a+b)^{[r]}} \times \underbrace{(a+b+r)(a+b+r+1)\dots(a+b+r+x-1)}_{(a+b+r)^{[x]}} = (a+b)^{[r]} (a+b+r)^{[x]}$$

Assim, (4.1) pode ser escrita como

$$h(x; a, b, r) = \frac{r^{[x]} a^{[r]} b^{[x]}}{x! (a+b)^{[r]} (a+b+r)^{[x]}} \quad (4.2)$$

Considere $X_1 = x_1, \dots, X_n = x_n$ uma amostra aleatória desta população. Os estimadores de máxima verossimilhança dos parâmetros a, b e r são aqueles que maximizam a função de verossimilhança

$$L = L(a, b, r; x_1, \dots, x_n) = \prod_{i=1}^n h(x_i; a, b, r)$$

$$L = \prod_{i=1}^n \frac{r^{[x_i]} a^{[r]} b^{[x_i]}}{x_i! (a+b)^{[r]} (a+b+r)^{[x_i]}} = \left[\frac{a^{[r]}}{(a+b)^{[r]}} \right]^n \prod_{i=1}^n \frac{r^{[x_i]} b^{[x_i]}}{(x_i!)(a+b+r)^{[x_i]}}$$

A amostra $x_1, \dots, x_n$ é uma coleção de números inteiros com possíveis repetições. Coletam-se os valores iguais desta sequência e denomina-se de f_x o número de vezes que o inteiro x aparece na amostra. $f_x = 0$ indica que o inteiro x não apareceu na sequência. Desta forma, pode-se fazer o somatório em f_x para x variando de zero até infinito. Note que neste caso, $\sum_{x=0}^{\infty} f_x = n$.

Para obter os máximos da função de verossimilhança é comum o uso da função $\ln(L(a, b, r; x_1, \dots, x_n))$, pois facilita o processo algébrico nos procedimentos de maximização.

$$\begin{aligned} \ln(L) = n \left[\ln a^{[r]} - \ln(a+b)^{[r]} \right] + \sum f_x \ln r^{[x]} + \sum f_x \ln b^{[x]} \\ - \sum f_x \ln(a+b+r)^{[x]} - \sum f_x \ln(x!) \end{aligned} \quad (4.3)$$

Desenvolvendo-se os fatoriais em (4.3),

$$\begin{aligned} \ln(L) = n \left[\ln(a(a+1)\dots(a+r-1)) - \ln(a(a+b)(a+b+1)(a+b+2)\dots \right. \\ \left. \dots(a+b+r-1)) \right] + \sum f_x \ln(r(r+1)\dots(r+x-1)) + \sum f_x \ln(b\dots \\ \dots(b+x-1)) - \sum f_x \ln(x!) - \sum f_x \ln((a+b+r)(a+b+r+1)\dots \\ \dots(a+b+r+x-1)) \end{aligned}$$

$$\ln(L) = n \left[\sum_{k=0}^{r-1} \ln(a+k) - \sum_{k=0}^{r-1} \ln(a+b+k) \right] + \sum f_x \left(\sum_{k=0}^{x-1} \ln(r+k) \right) + \\ \sum f_x \left(\sum_{k=0}^{x-1} \ln(b+k) \right) - \sum f_x \ln(x!) - \sum f_x \left(\sum_{k=0}^{x-1} \ln(a+b+r+k) \right)$$

Para encontrar os estimadores de máxima verossimilhança basta tomar as derivadas parciais da função $\ln(L(a,b,r;x_1,...,x_N))$ em relação aos parâmetros a, b e r , igualar a zero e resolver as equações. Portanto,

$$1) \quad \frac{\partial \ln(L)}{\partial a} = n \sum_{k=0}^{r-1} \frac{1}{a+k} - n \sum_{k=0}^{r-1} \frac{1}{a+b+k} - \sum f_x \left(\sum_{k=0}^{x-1} \frac{1}{a+b+r+k} \right)$$

$$2) \quad \frac{\partial \ln(L)}{\partial b} = \sum f_x \left(\sum_{k=0}^{x-1} \frac{1}{b+k} \right) - n \sum_{k=0}^{r-1} \frac{1}{a+b+k} - \sum f_x \sum_{k=0}^{x-1} \frac{1}{a+b+r+k}$$

$$3) \quad \frac{\partial \ln(L)}{\partial r} = n \frac{\partial \ln a^{[r]}}{\partial r} - n \frac{\partial \ln(a+b)^{[r]}}{\partial r} + \sum f_x \frac{\partial \ln r^{[x]}}{\partial r} - \\ - \sum f_x \frac{\partial \ln(a+b+r)^{[x]}}{\partial r}$$

Convertendo a 3ª equação em termos da função gama tem-se:

$$\begin{aligned}
&= n \frac{\partial \ln}{\partial r} \left[\frac{\Gamma(a+r)}{\Gamma(a)} \right] - n \frac{\partial \ln}{\partial r} \left[\frac{\Gamma(a+b+r)}{\Gamma(a+b)} \right] + \sum f_x \frac{\partial \ln}{\partial r} \left[\frac{\Gamma(r+x)}{\Gamma(r)} \right] - \\
&\quad - \sum f_x \frac{\partial \ln}{\partial r} \left[\frac{\Gamma(a+b+r+x)}{\Gamma(a+b+r)} \right]
\end{aligned}$$

segue que,

$$\begin{aligned}
\frac{\partial \ln(L)}{\partial r} &= n \frac{\partial \ln \Gamma(a+r)}{\partial r} - n \frac{\partial \ln \Gamma(a)}{\partial r} - n \frac{\partial \ln \Gamma(a+b+r)}{\partial r} + n \frac{\partial \ln \Gamma(a+b)}{\partial r} + \\
&\quad + \sum f_x \frac{\partial \ln \Gamma(r+x)}{\partial r} - \sum f_x \frac{\partial \ln \Gamma(r)}{\partial r} - \sum f_x \frac{\partial \ln \Gamma(a+b+r+x)}{\partial r} + \\
&\quad + \sum f_x \frac{\partial \ln \Gamma(a+b+r)}{\partial r} = 0
\end{aligned}$$

portanto,

$$\begin{aligned}
\frac{\partial \ln(L)}{\partial r} &= n \frac{\partial \ln \Gamma(a+r)}{\partial r} - n \frac{\partial \ln \Gamma(a+b+r)}{\partial r} - \sum f_x \frac{\partial \ln \Gamma(a+b+r+x)}{\partial r} - \\
&\quad - \sum f_x \frac{\partial \ln \Gamma(r)}{\partial r} + \sum f_x \frac{\partial \ln \Gamma(r+x)}{\partial r} + \sum f_x \frac{\partial \ln \Gamma(a+b+r)}{\partial r} \quad (4.4)
\end{aligned}$$

A diferenciação de (4.4) não é direta e pode ser resolvida com o auxílio da relação de recorrência apresentada por Johnson et al., (1992), (ver apêndice),

$$\psi(x+n) = \psi(x) + \sum_{j=1}^n (x+j-1)^{-1}, \quad n=1,2,\dots \quad (4.5)$$

em que $\psi(x) = \frac{d}{dx} \{\ln \Gamma(x)\} = \frac{\Gamma'(x)}{\Gamma(x)}$ é chamada função digama ou função

psi. Uma boa aproximação para $\psi(x)$ é

$$\psi(x) \approx \ln x - \frac{1}{2x}. \quad (4.6)$$

Substituindo (4.6) em (4.5)

$$\psi(x+n) \approx \ln x - \frac{1}{2x} + \sum_{j=1}^n (x+j-1)^{-1}, \quad n=1, 2, \dots \quad (4.7)$$

Aplicando (4.7) em (4.4)

$$\begin{aligned} \frac{\partial \ln(L)}{\partial r} &= n \left[\psi(a+r) - \psi(a+b+r) \right] - \sum f_x \left[\psi(r) - \psi(a+b+r) - \right. \\ &\quad \left. - \psi(r+x) + \psi(a+b+r+x) \right] \\ &= n \left[\left(\ln r - \frac{1}{2r} + \sum_{k=1}^a \frac{1}{r+k-1} \right) - \left(\ln r - \frac{1}{2r} + \sum_{k=1}^{a+b} \frac{1}{r+k-1} \right) \right] - \sum f_x \left[\left(\ln r - \frac{1}{2r} \right) - \right. \\ &\quad \left. - \left(\ln r - \frac{1}{2r} + \sum_{k=1}^{a+b} \frac{1}{r+k-1} \right) - \left(\ln r - \frac{1}{2r} + \sum_{k=1}^x \frac{1}{r+k-1} \right) + \right. \\ &\quad \left. + \left(\ln r - \frac{1}{2r} + \sum_{k=1}^{a+b+x} \frac{1}{r+k-1} \right) \right] \end{aligned}$$

obtem-se a terceira equação da verossimilhança,

$$\begin{aligned} \frac{\partial \ln(L)}{\partial r} = n \left[\sum_{k=1}^a \frac{1}{r+k-1} - \sum_{k=1}^{a+b} \frac{1}{r+k-1} \right] + \sum f_x \left[\left(\sum_{k=1}^{a+b} \frac{1}{r+k-1} \right) + \right. \\ \left. + \left(\sum_{k=1}^x \frac{1}{r+k-1} \right) - \left(\sum_{k=1}^{a+b+x} \frac{1}{r+k-1} \right) \right]. \end{aligned}$$

O sistema obtido

$$\left\{ \begin{aligned} \frac{\partial \ln L}{\partial a} &= n \sum_{k=0}^{\hat{r}-1} \frac{1}{\hat{a}+k} - n \sum_{k=0}^{\hat{r}-1} \frac{1}{\hat{a}+\hat{b}+k} - \sum f_x \left(\sum_{k=0}^{x-1} \frac{1}{\hat{a}+\hat{b}+\hat{r}+k} \right) = 0 \\ \frac{\partial \ln L}{\partial b} &= \sum f_x \left(\sum_{k=0}^{x-1} \frac{1}{\hat{b}+k} \right) - n \sum_{k=0}^{\hat{r}-1} \frac{1}{\hat{a}+\hat{b}+k} - \sum f_x \left(\sum_{k=0}^{x-1} \frac{1}{\hat{a}+\hat{b}+\hat{r}+k} \right) = 0 \\ \frac{\partial \ln L}{\partial r} &= n \left[\sum_{k=1}^{\hat{a}} \frac{1}{\hat{r}+k-1} - \sum_{k=1}^{\hat{a}+\hat{b}} \frac{1}{\hat{r}+k-1} \right] + \sum f_x \left[\left(\sum_{k=1}^{\hat{a}+\hat{b}} \frac{1}{\hat{r}+k-1} \right) + \right. \\ &\quad \left. + \left(\sum_{k=1}^x \frac{1}{\hat{r}+k-1} \right) - \left(\sum_{k=1}^{\hat{a}+\hat{b}+x} \frac{1}{\hat{r}+k-1} \right) \right] = 0 \end{aligned} \right.$$

pode ser resolvido utilizando-se um software de computação algébrica. Neste trabalho utilizou-se o software Maple, versão 5.

4.3 Ajustes

Considere os dados, tabela 1, do número de cáries que se desenvolvem em crianças de 12 anos, analisados por Gurland (1959) pela binomial negativa. Sendo $\hat{a} = 2,5$, $\hat{b} = 2$ e $\hat{r} = 2$, estimativas de máxima verossimilhança, um ajuste via beta binomial negativa é também realizado, Figura 12.

TABELA 1 Distribuição do número de cáries em crianças de 12 anos.

Contagem por criança x	Frequência observada F_o	Frequência esperada BBN	Frequência esperada BN
0	63	56,2	61,2
1	29	35,1	29,2
2	12	21,3	18,8
3	15	13,5	13,1
4	8	9,0	9,6
5	9	6,2	7,1
6	5	4,4	5,4
7	4	3,3	4,1
8	6	2,5	3,2
9	2	1,9	2,4
10	3	1,5	1,9
11	3	1,2	1,6
12	2	1,0	1,2
13 ⁺	2	0,8	4,2
χ_c^2	---	20,3	10,0
$P(\chi^2 > \chi_c^2)$	---	0,03	0,53
$g.l.$	---	10	11

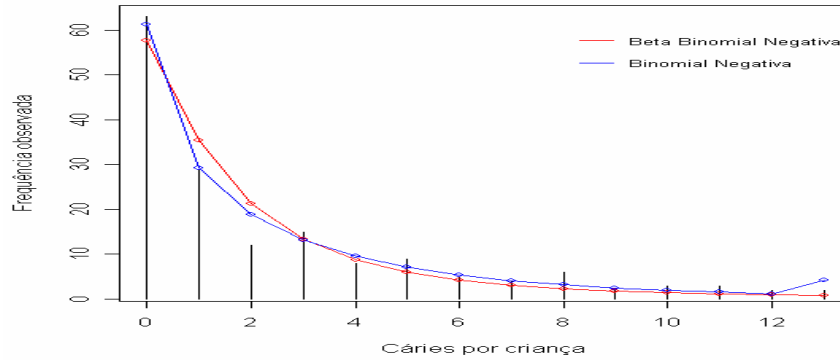


FIGURA 12 Distribuição das freqüências do número de cáries por criança. Ajuste pela distribuição binomial negativa e beta binomial negativa.

Como $P(\chi_c^2 > \chi^2) = 0,03$ pode-se concluir que a distribuição do número de cáries em crianças não pode ser descrita pela distribuição beta binomial negativa, indicando que o parâmetro p , probabilidade de um dente desenvolver cárie, não varia de dente para dente.

Os dados apresentados a seguir, Tabela 2, são referentes ao número de acidentes por maquinista num período de três meses, analisados por Bliss & Fisher (1953). Os valores observados e esperados pelas distribuições binomial negativa e beta binomial negativa ($\hat{a} = 5,4$, $\hat{b} = 1,5$ e $\hat{r} = 1,5$) são obtidos e um ajuste por estas distribuições é apresentado na Figura 13.

TABELA 2 Distribuição do número de acidentes por maquinista.

Acidentes por maquinistas x	Frequência observada Fo	Frequência esperada BBN	Frequência esperada BN
0	296	290,5	296,7
1	74	77,8	71,0
2	26	25,9	26,4
3	8	10,2	11,0
4	4	4,5	4,8
5	4	2,2	2,2
6	1	1,2	1,0
7	0	0,6	0,5
8	1	0,4	0,2
9 ⁺	0	0,2	0,2
χ^2_c	---	0,8	1,1
$P(\chi^2 > \chi^2_c)$	---	0,37	0,57
$g.l.$	---	1	2

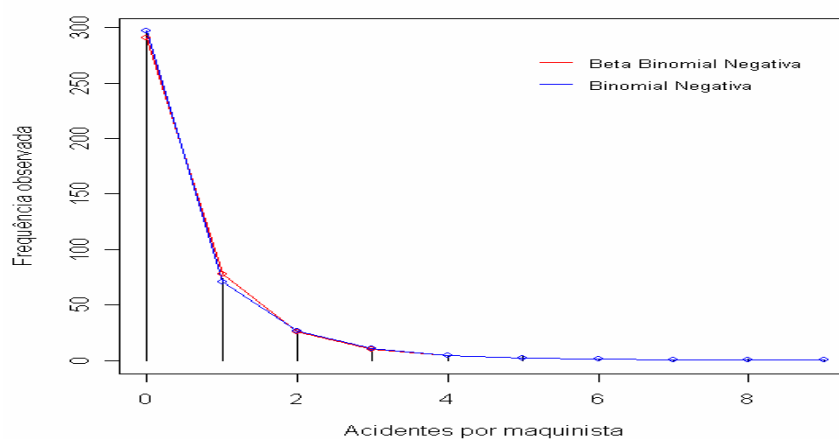


FIGURA 13 Ajuste via binomial negativa e beta binomial negativa.

Os dados da Tabela 3 referem-se à distribuição de plantas *Salicornia Stricta*, por quadrado, que se desenvolvem num pântano.

TABELA 3: Contagem do número de plantas por quadrante.

Plantas por quadrante x	Frequência observada Fo	Frequência esperada BBN	Frequência esperada BN
0	4	3,2	3,3
1	3	5,5	6,4
2	8	8,5	8,4
3	13	10,2	9,4
4	11	10,7	9,6
5	9	10,2	9,2
6	8	9,2	8,4
7	10	8,0	7,5
8	3	6,7	6,5
9	3	5,5	5,6
10	8	4,4	4,7
11	3	3,6	3,1
12	4	2,8	2,5
13	4	2,2	2,1
14	0	1,8	1,7
15	3	1,4	1,3
16	0	1,1	1
17	0	0,9	0,8
18	1	0,7	0,6
19	0	0,5	0
20 ⁺	3	0,4	5,9
χ_c^2	---	8,1	7,5
$P(\chi^2 > \chi_c^2)$	---	0,32	0,48
<i>g.l.</i>	---	7	8

O ajuste via distribuição binomial negativa e beta binomial negativa ($\hat{a} = 12,3$, $\hat{b} = 8,7$ e $\hat{r} = 8,7$) encontra-se na Figura 14.

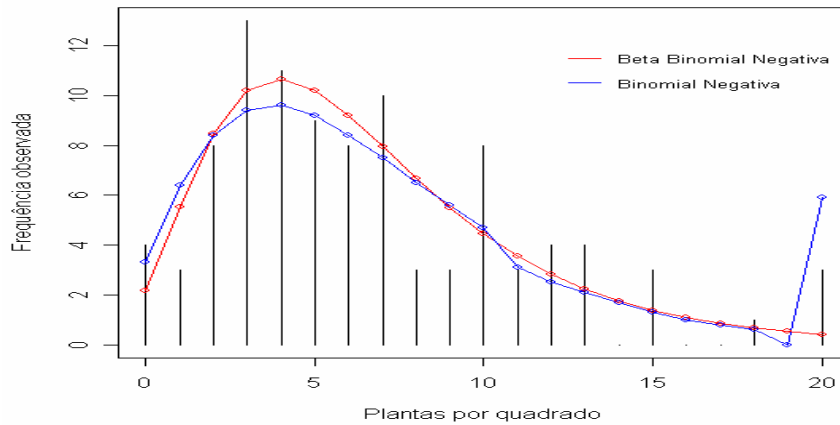


FIGURA 14 Distribuição do número de plantas *Salicornia stricta* por quadrado. Ajuste pela distribuição binomial negativa e beta binomial negativa.

Os dados da Tabela 4 referem-se à contagem do número de ácaros em folhas de maçã. As frequências esperadas pelas distribuições binomial negativa e beta binomial negativa ($\hat{a} = 8,1$, $\hat{b} = 2,9$ e $\hat{r} = 2,9$) são obtidas e o ajuste é apresentado na Figura 15.

TABELA 4 Distribuição de ácaros em folhas de maçã.

Ácaros por folha x	Frequência observada F_o	Frequência esperada BBN	Frequência esperada BN
0	70	66,67	69,49
1	38	40,34	37,6
2	17	20,59	20,1
3	10	10,36	10,7
4	9	5,34	5,69
5	3	2,84	3,02
6	2	1,56	1,6
7	1	0,89	0,85
8	0	0,52	0,95
χ^2	$N = 150$	3,46	2,48
$P(\chi^2)$	---	0,18	0,48
$g.l.$	---	2	3

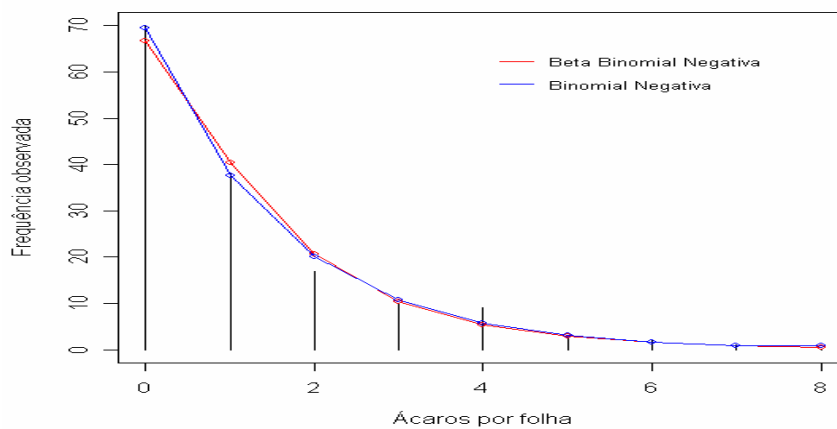


FIGURA 15 Distribuição do número de ácaros por folhas de maçã. Ajuste pela distribuição binomial negativa e beta binomial negativa.

Pelo exposto, tem-se que a distribuição beta binomial negativa se ajustou aos dados das Tabelas 2, 3 e 4, portanto foi adequado considerar a variabilidade do parâmetro p nesses casos.

Pielou (1977) desenvolve uma discussão sobre o ajuste de distribuições teóricas a dados observados. Os dados analisados são adequadamente ajustados pela binomial negativa. A binomial negativa pode ser obtida como distribuição generalizada de uma Poisson e uma logarítmica como também por mistura de uma Poisson com uma gama. Portanto, a distribuição binomial negativa surge de modelos matemáticos diferentes, com pressupostos biológicos diferentes. Com isto Pielou coloca o problema:

“,,, o ajuste de distribuições de frequência teóricas a dados observados não é suficiente para explicar o padrão natural de uma população”

O Ajuste obtido pela beta binomial negativa acrescenta mais um aspecto a esta discussão.

5 CONCLUSÕES

- 1) A distribuição beta binomial negativa é uma distribuição trabalhável e com boas propriedades;
- 2) As equações aproximadas para o cálculo dos estimadores de máxima verossimilhança foram obtidas;
- 3) Uma validação matemática do algoritmo mistura é apresentada;
- 4) A distribuição beta binomial negativa apresentou um bom ajuste aos dados analisados sendo, portanto um modelo alternativo para explicar os dados considerando-se a variabilidade do parâmetro p .

REFERÊNCIAS BIBLIOGRÁFICAS

- ANTON, H. **Cálculo**: um novo horizonte. 6.ed. Porto Alegre: Bookman, 2000.
- BEARZOTI, E. **Introdução à teoria de probabilidade e à inferência estatística**. Lavras: UFLA, 1998.
- BLISS, C.I.; FISHER, R.A. Fitting the negative binomial distribution to biological data. **Biometrics**, Washington, v.9, n.2, p.176-200, May 1953.
- DEVROYE, L. **Non-uniform random variate generation**. New York: Springer-Verlag, 1986.
- FELLER, W. **An introduction to probability theory and its applications**. New York: J. Wiley, 1968.
- GERSTENKORN, T. A compound of the generalized negative binomial distribution with the generalized beta distribution. **Central European Journal of Mathematics**, London, v.2, n.4, p.527-237, Aug. 2004.
- GURLAND, J. Some interrelations among compound and generalized distributions. **Biometrika**, London, v.44, n.1-2, p.265-268, June 1957.
- GURLAND, J. Some applications of the negative binomial distribution and other contagious distributios. **Biometrika**, London, v.49, n.10, p.1388-1399, Dec. 1959.
- HASSAN, A.; BILAL, S. On estimation of negative Polya-Eggenberger distribution and applications. **Journal of the Korean Society for Industrial and Applied Mathematics**, Seul, v.12, n.2, p.81-95, Aug. 2008,
- HOLGATE, P.B. The modality of compound Poisson distribution. **Biometrika**, London, v.57, n.3, p.665-667, Dec. 1970,
- JOHNSON, N.L.; KOTZ, S. **Discrete distributions**. New York: J. Wiley, 1969.
- JOHNSON, N.L.; KOTZ, S.; KEMP, A.W. **Univariate discrete distributions**. New York: J. Wiley, 1992.

MOOD, A.M.; GRAYBILL, F.A.; BOES, D.C. **Introduction to the theory of statistics**. New York: McGraw-Hill, 1974. (Probability and statistics).

PIELOU, E.C. **Mathematical ecology**. New York: J. Wiley, 1977.

R DEVELOPMENT CORE TEAM. **R**: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing 2007

SILVA, W.P. **Estimadores de máxima verossimilhança em misturas de densidades normais**: uma aplicação em genética. 2003. 60p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária)–Universidade Federal de Lavras, Lavras, MG.

ANEXOS

ANEXO A: Função gama.....	73
ANEXO B: Rotinas.....	76

ANEXO A

Função gama

A função gama, denotada por $\Gamma(\cdot)$, é definida por

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt \quad \text{para } x > 0, \quad (1)$$

Usando integração por partes, (1) fornece uma relação de recorrência para $\Gamma(x)$:

$$\Gamma(x+1) = x\Gamma(x) \quad (2)$$

e, se $x = n$ é um inteiro, $\Gamma(n+1) = n!$.

Utilizando a recorrência dada em (2), obtém-se uma relação entre a função gama e o fatorial ascendente,

$$a^{[x]} = a(a+1)\dots(a+x-1) = \frac{\Gamma(a+x)}{\Gamma(a)}.$$

A relação entre as funções beta e gama é

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \quad a, b > 0,$$

A derivada do logaritmo da função gama, denominada *função digama* ou *função psi*, é dada por

$$\psi(x) = \frac{d}{dx} [\ln \Gamma(x)] = \frac{\Gamma'(x)}{\Gamma(x)}.$$

A expressão (2) para a função gama produz a seguinte fórmula de recorrência para a função psi, (Johnson *et al.*, 1992):

$$\psi(x+1) = \psi(x) + \frac{1}{x}$$

Prova $\Gamma(x+1) = x\Gamma(x)$

$$\Gamma'(x+1) = \Gamma(x) + x\Gamma'(x)$$

$$\frac{\Gamma'(x+1)}{\Gamma(x)} = 1 + x \frac{\Gamma'(x)}{\Gamma(x)}$$

Mas, $\frac{\Gamma'(x)}{\Gamma(x)} = \psi(x)$, então

$$\frac{\Gamma'(x+1)}{\Gamma(x)} = 1 + x\psi(x)$$

$$\frac{x\Gamma'(x+1)}{\Gamma(x+1)} = 1 + x\psi(x)$$

$$x\psi(x+1) = 1 + x\psi(x)$$

$$\psi(x+1) = \psi(x) + \frac{1}{x} \quad \#$$

Deste fato, decorre que

$$\psi(x+1) = \psi(x) + \frac{1}{x}$$

$$\psi(x+2) = \psi((x+1)+1) = \psi(x+1) + \frac{1}{x+1} = \psi(x) + \frac{1}{x} + \frac{1}{x+1}$$

$$\psi(x+3) = \psi((x+2)+1) = \psi(x+2) + \frac{1}{x+2} = \psi(x) + \frac{1}{x} + \frac{1}{x+1} + \frac{1}{x+2}$$

$$\begin{array}{ccc} \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \end{array}$$

$$\psi(x+n) = \psi((x+n-1)+1) = \psi(x) + \frac{1}{x} + \frac{1}{x+1} + \dots + \frac{1}{x+n-1}$$

$$\psi(x+n) = \psi(x) + \sum_{j=1}^n \frac{1}{x+j-1} \quad .$$

A função $\psi(x+n)$ representa a derivada do logaritmo da função $\Gamma(x+n)$.

$$\begin{aligned} \textbf{Prova} \quad \psi(x+n) &= \psi(x) + \sum_{j=1}^n (x+j-1)^{-1} = \psi(x) + \frac{\partial \ln x^{[n]}}{\partial x} \\ &= \psi(x) + \frac{\partial}{\partial x} \ln \left[\frac{\Gamma(x+n)}{\Gamma(x)} \right] \end{aligned}$$

como $\psi(x) = \frac{d}{dx} [\ln \Gamma(x)]$, segue que

$$\psi(x+n) = \frac{\partial \ln \Gamma(x)}{\partial x} + \frac{\partial \ln \Gamma(x+n)}{\partial x} - \frac{\partial \ln \Gamma(x)}{\partial x} = \frac{\partial \ln \Gamma(x+n)}{\partial x}.$$

ANEXO B

Rotinas

As rotinas utilizadas na simulação das misturas, realizadas pelo software R versão 2.8.0, são apresentadas abaixo.

- **Simulando mistura de normais**

```
#      X~Normal(mu,1)  --  mu~Normal(1,0)
n = 10000
X = rnorm(n,rnorm(n),1)
hist(X, freq = F, breaks = 50, main="",
      xlab="x", ylab= "h (x)")
```

```
x = (-50:50)/10
Y = dnorm(x,0,sqrt(2))
lines(x,Y,col="red")
```

- **Simulando binomial negativa**

```
n = 1000000      # Tamanho da amostra
a = 10           # Parâmetro da distribuição gama
b = 1           # Parâmetro da distribuição gama
```

```
X = sort(rpois(n,rgamma(n,a,b)))    # Vetor de amostras da mistura Poisson
                                     gama
levX = as.numeric(levels(factor(X)))
nX = length(levX)
fX = matrix(0,nX,2)                 # Matriz de frequências
fX[1:nX,1] = levX[1:nX]
```

```

kk = 0
for (i in 1:nX){
  ref = fX[i,1]
  k = 0
  while ((kk+k) < n & X[kk+k+1] == ref){
    k = k+1
  }
  fX[i,2] = k
  kk = kk + k
}
ma = max(X)
brea = (0:(ma+1))-0,5
x = 0:ma
Y = dnbinom(x,a,(b/(b+1)))
hist(X,breaks = brea, freq = F, main="",
      xlab = "Eixo dos X", ylab = "Eixo dos Y" )
points(x,Y,col="red")

```

- **Simulando beta binomial negativa**

```

n = 10000          # Tamanho da amostra
X = 1:n            # Vetor de amostras
a = 4              # Parâmetro da beta
b = 4              # Parâmetro da beta
r = 3              # Parâmetro r da binomial negativa

X = sort(rnbinom(n,r,rbeta(n,a,b)))          # Gera o vetor de dados
flag = 1

```

```

k = 2
while (flag < 3 & k < n){
    k = k+1
    flag = X[k] - X[k-1]
}
max = X[k]-1
for (i in 1:n){
    if (X[i]>max) X[i] = max
}
levX = 0:max
nX = length(levX)
fX = matrix(0,2,nX)          # Matriz de frequências
fX[1,1:nX] = levX[1:nX]
for (i in 1:nX){
    ref = fX[1,i]
    k = 0
    for (j in 1:n) if (X[j] == ref) k = k+1
    fX[2,i] = k
}
for (i in 1:nX){
    ref = fX[1,i]
    k = 0
    for (j in 1:n) if (X[j] == ref) k = k+1
    fX[2,i] = k
}

```

- **Ajustando uma beta binomial negativa**

```
xx = 0:20                                     # Valores observados
yy = c(4,3,8,13,11,9,8,10,3,3,8,3,4,0,3,0,0,1,0,3) # Freqüência
observada
n = sum(yy)
plot(xx,yy, type = "h", xlab=" ", ylab=" ")
a = , b = , r =                               #Parâmetros do modelo
bbn <- function(x,a,b,r){
  num = gamma(r+x)*gamma(r+a)*gamma(x+b)*gamma(a+b)
  den = gamma(r)*gamma(r+a+b+x)*gamma(a)*gamma(b)*factorial(x)
  h = num/den
  return(h)
}
Ybbn = bbn(xx,a, b, r)*n
lines(xx,Ybbn,col="red")
points(xx,Ybbn,col="red")
legend(12,12,lty=c(1,1), c("beta binomial negativa"),
      lwd=1, bty="n",cex=0,8, col="red")
```

Na estimação dos parâmetros da distribuição beta binomial negativa pelo método da Máxima verossimilhança utilizou-se o software Maple versão 5. A rotina é dada abaixo,

```
> restart;
```

```
with(LinearAlgebra):
```

A matriz das freqüências observadas F_x

```

> fx :=
convert([7,16,23,20,23,24,12,13,9,9,8,10,8,7,2,12,3,5,4,2,2,5,5,2,1,16],Vector):
tam := Dimension(fx):
Fx := Matrix(tam,2):
N := 0:
for i from 1 to tam do
    Fx[i,1] := i-1:
    Fx[i,2] := fx[i]:
    N := N + fx[i]:
end do:

```

As equações de verossimilhança aproximadas

```

> f11 := (a,b,r)-> N*sum('1/(a+k)',k=0,,(r-1)):
f12 := (a,b,r)-> N*sum('1/(a+b+k)',k=0,,(r-1)):
f13 := (a,b,r)-> sum('Fx[(x+1),2]*sum(1/(a+b+r+k),
    k=0,,(x-1))',x=1,,(tam-1)):
f1 := (a,b,r) -> f11 - f12 - f13:

> f21 := (a,b,r)-> sum('Fx[(x+1),2]*sum(1/(b+k),
    k=0,,(x-1)))',x=1,,(tam-1)):
f22 := (a,b,r)-> N*sum('1/(a+b+k)',k=0,,(r-1)):
f23 := (a,b,r)-> sum('Fx[(x+1),2]*sum(1/(a+b+r+k), k=0,,(x-1))',x=1,,(tam-1)):
f2 := f21 - f22 - f23:

> f31 := (a,b,r)-> N*sum('1/(r+k-1)',k=1,,a):
f32 := (a,b,r)-> N*sum('1/(r+k-1)',k=1,,(a+b)):

```

```

f33 := (a,b,r)-> sum('Fx[(x+1),2]*(sum(1/(r+k-1),
                                k=1,,x))',x=1,,(tam-1)):
f34 := (a,b,r)-> sum('Fx[(x+1),2]',x=1,,(tam-1))
                                *sum('1/(r+k-1)',k=1,,(a+b)):
f35 := (a,b,r)-> sum('Fx[(x+1),2]*(sum(1/(r+k-1),
                                k=1,,(a+b+x))))',x=1,,(tam-1)):
f3 := f31 - f32 + f33 + f34 - f35:

> fsolve({f1(a,b,r)=0, f2(a,b,r)=0, f3(a,b,r)=0}, {a,b,r});
      {a= , b= ,r= }

```