ISSN: 2317-0840

Comportamento Unimodal da Distribuição Beta Binomial Negativa

Cícero Carlos F. Oliveira[1†], Cláudio T. Cristino[2], Pedro F. Lima[3]

[1] *Doutorando pelo Programa de Pós-Graduação em Biometria e Estatística Aplicada - UFRPE, professor de estatística do IFCE - Campus Crato.*
[2] *Professor do Programa de Pós-Graduação em Biometria e Estatística Aplicada na UFRPE - e-mail: ctcristino@gmail.com.*
[3] *Doutorando pelo Programa de Pós-Graduação em Biometria e Estatística Aplicada - UFRPE, professor de estatística da URCA - CE, e e-mail: limapf@yahoo.com.br.*

Resumo: *Neste artigo apresenta-se um estudo sobre o comportamento da distribuição Beta Binomial Negativa em relação a sua moda. Aplica-se este estudo quando essa distribuição é dada através de uma distribuição hipergeométrica equivalente multiplicada pela verossimilhança da distribuição Binomial Negativa. Essa representação, que não aparece costumeiramente na literatura, é uma interessante visão da distribuição Beta Binomial Negativa, tanto em termos didáticos, quanto computacionais. Este distribuição é obtida a partir de considerações segundo uma visão bayesiana. Sendo assim, duas análises são feitas de forma comparativa: a primeira usando a idéia do Holgate (1970) que também é observado no artigo de Hassan and Bilal (2008); e a segunda usando a idéia de que a moda é um ponto de máximo em relação a distribuição, este cálculo é feito usando o software livre* **R-software** *e ao mesmo tempo é observada graficamente.*

Palavras-chave: Distribuição Beta Binomial Negativa, Distribuição Binomial Negativa, Moda.

Abstract: *This article presents a study on the behavior of the Negative Binomial Beta distribution regarding their mode. Applies this study that when the distribution is given by a hypergeometric distribution equivalent multiplied by the likelihood of the Negative Binomial distribution. This representation, which does not ordinarily appear in the literature, is an interesting view of the Negative Binomial Beta distribution, both in didactic terms, as computational. This distribution is obtained from a Bayesian considerations. Thus, two analyzes are done on a comparative basis: the first using the idea of Holgate (1970) which is also observed in Hassan and Bilal (2008) article, and the second using the idea that mode is a maximum point regarding the distribution, this calculation is done using the software free R-software and at the same time is observed graphically.*

Keywords: Negative binomial beta distribution, Negative binomial distribution, Mode.

† Autor correspondente: cicerofelixoliveira@bol.com.br.

Introdução

A distribuição Polya-Eggenberger e sua inversa, a distribuição Polya-Eggenberger Negativo (ou distribuição Beta Binomial Negativa), foram introduzidas por Polya e Eggenberger (1923) através de um modelo de urna.

Pode-se verificar facilmente que a distribuição Beta Binomial Negativa é unimodal, ou seja, apresenta uma única moda. Serão mostradas duas maneiras de se obter a moda desse distribuição e depois ambas serão comparadas. Uma das maneiras de se obter a moda desta distribuição é utilizar o fato de que a função gama generaliza a função fatorial para os números inteiros positivos, permitido uma extensão contínua para os números reais positivos.

Sendo assim, sabemos por definição que a moda é calculada pela primeira derivada do logaritmo da distribuição e depois igualando a zero; outra maneira seria utilizando os resultados obtidos por Holgate (1970).

Material e Métodos

Distribuição Beta Binomial Negativa

Considere uma sequência de ensaios de Bernoulli realizada de forma independente com probabilidade θ constante e defina X como o número de fracassos anteriores ao r-ésimo sucesso. A variável aleatória X segue uma distribuição Binomial Negativa com parâmetros r e θ, em que $0 < \theta < 1$ e $r > 0$, e tem função densidade de probabilidade dada para $x = 0, 1, 2, \ldots$ por:

$$p(x \mid r, \theta) = \binom{x+r-1}{x} \theta^r (1-\theta)^x.$$

Do ponto de vista bayesino, pode-se introduzir a variabilidade de θ, no exposto acima, supondo que θ é uma variável aleatória seguindo uma distribuição Beta, isto é,

$$p(\theta) = \frac{1}{\beta(a,b)} \theta^{a-1} (1-\theta)^{b-1} \qquad 0 < \theta < 1,$$

em que $\beta(a, b)$ é a função beta da por

$$\beta(a,b) = \frac{\Gamma(a) \cdot \Gamma(b)}{\Gamma(a+b)} \qquad\qquad a > 0 \quad e \quad b > 0$$

e

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt.$$

Neste caso, obtém-se como mistura a distribuição discreta Beta Binomial Negativa (BBN) dada por

$$\begin{aligned} p(x \mid a, b, r, \theta) &= \int_0^1 p(x|r,\theta) p(\theta) d\theta \\ &= \binom{x+r-1}{x} \frac{\beta(r+a, x+b)}{\beta(a,b)} \\ &= \binom{x+r-1}{x} \frac{\Gamma(r+a)\Gamma(x+b)\Gamma(a+b)}{\Gamma(x+r+a+b)\Gamma(a)\Gamma(b)} \qquad x = 0, 1, 2, \ldots \end{aligned}$$

Deste modo, pode-se definir esta distribuição como segue.

A distribuição Beta Binomial Negativa é uma distribuição Binomial Negativa, cuja probabilidade de sucesso do parâmetro θ segue uma distribuição Beta com parâmetros a e b. Em outras palavras, se a probabilidade de sucesso do parâmetro θ de uma distribuição Binomial Negativa (com parâmetros r e θ) tem uma distribuição Beta com parâmetros a e b, então a distribuição resultante é referida como a distribuição Beta Binomial Negativa com parâmetros a, b, r e θ e denotada por $BBN(a, b, r, \theta)$. Para uma distribuição Binomial Negativa Padrão, θ é geralmente considerado como fixo para os ensaios sucessivos, mas o valor de θ muda para cada ensaio, para a distribuição Beta Binomial Negativa. A distribuição Beta Binomial Negativa é muitas vezes referida como uma distribuição de Markov Pólya Inversa.

Considere uma variável aleatória X tendo uma distribuição Beta Binomial Negativo. Então, para $x = 0, 1, \ldots$, esta distribuição de probabilidade,(JOHNSON et al. 2005; TEERAPABOLARN, 2008) é definida por

$$p(x \mid a, b, r, \theta) = \binom{x + r - 1}{x} \frac{\Gamma(r + a)\Gamma(x + b)\Gamma(a + b)}{\Gamma(x + r + a + b)\Gamma(a)\Gamma(b)}. \tag{1}$$

A expressão (1) pode ser representada também pela seguinte forma:

$$p(x \mid a, b, r, \theta) = \frac{r}{x + r} \frac{\binom{x + b - 1}{x}\binom{r + a - 1}{r}}{\binom{x + r + a + b - 1}{x + r}}. \tag{2}$$

Essa representação não aparece na literatura e é uma visão interessante da Beta Binomial Negativa. Implica também que, para valores reais de $a > 0$ e $b > 0$, pode ser avaliada facilmente em qualquer pacote de software que fornecem comandos para combinação.

A unimodalidade da distribuição Beta Binomial Negativa

A unimodalidade da distribuição Beta Binomial Negativa foi estudada por Hassan e Bilal (2008) e logo depois estudada por Madeira (2009). Os autores justificam a unimodalidade da distribuição com base nos resultados obtidos por Holgate (1970). Em Madeira (2009), tem o seguinte Teorema (HASSAN e 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* $x = x_M$ *tal que:*

$$\frac{r(b - 1) - (a + b)}{a + 1} < x_M < \frac{(r - 1)(b + 1)}{a + 1}. \tag{3}$$

Agora, obtém-se uma melhor aproximação para a moda da distribuição Beta Binomial Negativa utilizando o fato de que a função gama generaliza o fatorial (definido para os números naturais) permitindo assim uma extensão contínua. Assim, pode-se calcular a moda fazendo primeira derivada do logaritmo da expressão (1) e depois igualando a zero, isto é:

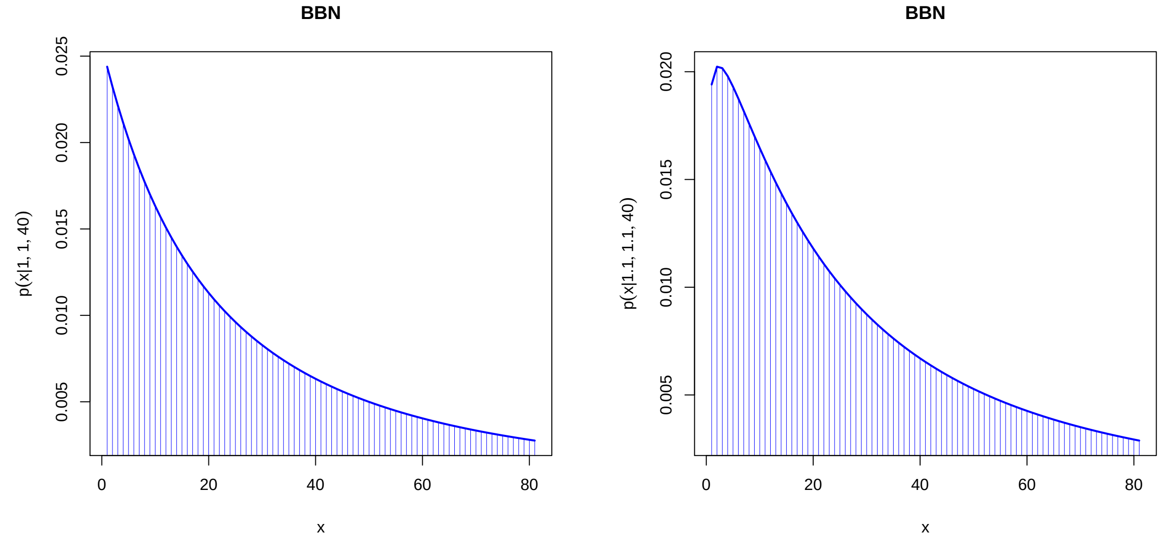
$$\Psi(x) = \frac{\partial}{\partial x} \log p(x|a, b, r, \theta) = \frac{\partial}{\partial x} \log \left(\frac{\Gamma(x + r)\Gamma(r + a)\Gamma(x + b)\Gamma(a + b)}{\Gamma(x + 1)\Gamma(r)\Gamma(x + r + a + b)\Gamma(a)\Gamma(b)} \right).$$

Pode-se verificar que:

$$\Psi(x + r) + \Psi(x + b) - \Psi(x + 1) - \Psi(x + r + a + b) = 0 \tag{4}$$

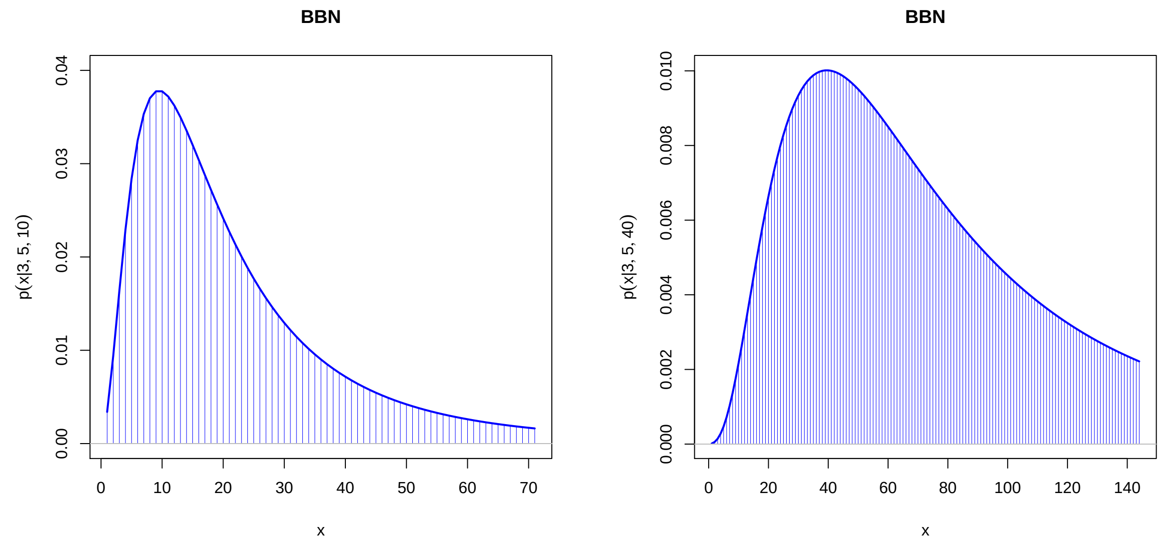
A construção dos gráficos abaixo foram baseados na distribuição Beta Binomial Negativa que corresponde a equação (2).

Para a Figura 1a, a moda calculada usando a expressão (4) é aproximadamente -40.50; para expressão (3), a moda está no intervalo aberto $(-1,\ 39)$ e observando na Figura 1a, a moda se aproxima a zero. Daí, pode-se concluir que a expressão (3) tem uma aproximação melhor, mas com um intervalo muito amplo.



(a) $a = 1,\ b = 1$ e $r = 40$ (b) $a = 1.1,\ b = 1.1$ e $r = 40$

Figura 1: Densidade da distribuição Beta Binomial Negativa.



(a) $a = 3,\ b = 5$ e $r = 10$. (b) $a = 3,\ b = 5$ e $r = 40$.

Figura 2: Densidade da distribuição Beta Binomial Negativa.

Para Figura 1b, a moda calculada usando a expressão (4) é aproximadamente 1.31, para expressão (3), a moda está no intervalo aberto $(0.86,\ 39)$ e observado na Figura 2, a moda se aproxima do valor encontrado na expressão (4). Daí, conclui-se que a moda encontrada na expressão (4) é melhor do que a encontrada na expressão (3), apesar da moda esta num intervalo muito amplo.

Para a Figura 2a, a moda calculada usando a expressão (4) é aproximadamente 8.49, para expressão (3), a moda está no intervalo aberto $(8, 13.5)$. Neste caso, a expressão (3) tem uma boa previsão da moda, mas a expressão (4) também tem uma boa aproximação.

Para a Figura 2b, a moda calculada usando a expressão (4) é aproximadamente 38.50, para expressão (3), a moda está no intervalo aberto $(38, 58.5)$. Neste caso, a expressão (4) tem uma aproximação da moda melhor do que na Figura 2a, mas a expressão (3) também tem uma boa aproximação, mesmo tendo um intervalo amplo.

Conclusões

Desta análise, pode-se concluir dois fatos importantes:

1. de acordo com as Figuras 1a e 1b, pode-se observar que a distribuição Beta Binomial Negativa é unimodal somente para $a > 1$ e $b > 1$. Quando a e b se aproxima de 1, o valor de r tem que crescer para garantir a afirmação anterior.

2. de acordo com as Figuras 2a e 2b, observa-se que o resultado da moda na expressão (4) se aproxima muito do resultado real da moda, quando o valor de r cresce.

Referências

EGGENBERGER, F; POLYA,G.; ANGEW,Z. Math. Mech, *Uber die Statistik verketteter vorgange*, 1(1923), 279-289.

HASSAN, A.; BILAL, S. *On Estimation of Negative Polya-Eggenberger Distribution and Its Applications.* J. Ksiam, vol. 12, N^o 2, 81 - 95, 2008.

JOHNSON, N. L.; KOTZ, S.; ADRIENNE, W. K. *Univariate Discrete Distributions.* Third edition. New York: Wiley, 2005.

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

HOLGATE,P. B. *The modality of compound Poisson distribution*,Biometrika, 57 (1970), 665-667.

R CORE TEAM. R: *A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. 2013. ISBN 3-900051-07-0, URL http://www.R-project.org/

TEERAPABOLARN, K. *Poisson approximation to the beta-negative binomial distribution.* International Journal of Contemporary Mathematical Sciences, vol.3, n.3,(2008), 457-461.