

DISTRIBUIÇÕES EM SÉRIE DE POTÊNCIAS MODIFICADAS

Edleide de BRITO¹
Gauss Moutinho CORDEIRO²
Clarice Garcia Borges DEMÉTRIO³

- RESUMO: Neste trabalho, apresentam-se alguns resultados na classe de distribuições em série de potências modificadas, introduzida por Gupta (1974, Sankhya B), tais como, momentos, fórmulas recursivas para os momentos, momentos centrais e momentos fatoriais e estimador de máxima verossimilhança. Essa classe de distribuições inclui algumas distribuições importantes, tais como binomial, binomial negativa, Poisson, série logarítmica, dentre outras. Mais recentemente, essa classe de distribuições foi estudada no contexto de modelos não-lineares generalizados por Cordeiro et al. (2009, CSDA).
- PALAVRAS-CHAVE: Distribuição discreta; função de probabilidade; função geradora de momentos; função de verossimilhança; momentos.

1 Introdução

A classe de distribuições em série de potências modificadas (MPSD), proposta por Gupta (1974), inclui várias distribuições importantes, tais como, binomial, binomial negativa, Poisson, série logarítmica, binomial negativa generalizada estudada por Jain e Consul (1971), Poisson generalizada de Consul e Jain (1973a), série logarítmica generalizada de Jain e Gupta (1973), Poisson Lagrangiana de Consul e Jain (1973b), binomial Lagrangiana de Jain e Consul (1971), Borel-Tanner e “lost game” de Kemp e Kemp (1968) e suas formas truncadas.

¹Universidade Federal da Bahia – UFBA, Departamento de Estatística, CEP: 41490-175, Salvador, Bahia, Brasil. E-mail: *edbnet@gmail.com*

²Universidade Federal Rural de Pernambuco – UFRPE, Departamento de Informática, CEP: 52171-900, Recife, Pernambuco, Brasil. E-mail: *gauss@deinfo.ufrpe.br*

³Universidade de São Paulo – USP, Escola Superior de Agricultura Luiz de Queiroz – ESALQ, Departamento de Ciências Exatas, CEP: 13418-900, Piracicaba, São Paulo, Brasil. E-mail: *clarice@esalq.usp.br*

Neste trabalho, apresentam-se as propriedades gerais da classe MPSD, tais como, momentos, fórmulas recursivas para os momentos, momentos centrais e momentos fatoriais, estimador de máxima verossimilhança (EMV), dentre outras. Algumas das principais distribuições que pertencem à essa classe de distribuições, bem como algumas das suas principais propriedades e aplicações, também, serão apresentadas. Essa classe de distribuições foi estudada, recentemente, no contexto de modelos não-lineares generalizados por Cordeiro et al. (2009) o que, também, será apresentado.

2 Propriedades da distribuição em série de potências modificada

Nesta seção, algumas propriedades da classe de distribuições em série de potências modificadas são apresentadas.

2.1 Função de probabilidade

Uma variável aleatória discreta Y tem distribuição em série de potências modificada se a sua função de probabilidade for expressa como

$$P(Y = y) = \frac{a(y, \phi)g(\mu, \phi)^y}{f(\mu, \phi)}, \quad y \in T, \quad (1)$$

em que T , o suporte de Y , não pode depender dos parâmetros μ ou ϕ e é um subconjunto do conjunto dos inteiros não-negativos, $a(y, \phi) > 0$, $\phi > 0$ e $g(\mu, \phi)$ e $f(\mu, \phi) = \sum_y a(y, \phi)g(\mu, \phi)^y$ são funções positivas, finitas e diferenciáveis. Essa classe de distribuições é uma extensão da classe de distribuições em série de potências generalizada (GPSD), que tem $g(\theta) = \theta$ em (1) e se, além disso, T é um subconjunto dos inteiros não-negativos essa classe será a de distribuições em série de potências.

2.2 Função geradora de momentos

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

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

Logo,

$$M_Y(t) = \sum_{i=0}^{\infty} \sum_{y=0}^{\infty} \binom{y+i}{i} y! \frac{a(y+i)g(\mu, \phi)^{y+i}(e^t - 1)^y}{f(\mu, \phi)y!}.$$

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

$$M_Y(t) = \sum_{s=0}^{\infty} \frac{t^s}{s!} \sum_{y=0}^{\infty} \sum_{i=0}^{\infty} \binom{y+i}{i} y! \frac{a(y+i)g(\mu, \phi)^{y+i} S(s, y)}{f(\mu, \phi)},$$

em que $S(s, x)$ representa o número de Stirling de segunda ordem, definido por

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

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

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

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

2.3 Fórmulas de recorrência para os momentos

Na Tabela 1, apresentam-se as fórmulas de recorrência para os momentos ordinários, centrais e fatoriais da classe de distribuições MPSD, sendo, $\mu'_1 = \frac{g(\mu, \phi)f'(\mu, \phi)}{g'(\mu, \phi)f(\mu, \phi)}$, $\mu_0 = 1$, $\mu_1 = 0$ e $\mu'_{[1]} = \mu'_1 = \mu$. A variância da classe de distribuições MPSD é expressa como $\mu'_2 - \mu_1^2 = \frac{g(\mu, \phi)}{g'(\mu, \phi)} \frac{d\mu'_1}{d\mu}$.

Tabela 1 - Fórmulas de recorrência para os momentos da classe MPSD

Momentos	Fórmula
Ordinários	$\mu'_{r+1} = \frac{g(\mu, \phi)}{g'(\mu, \phi)} \frac{d\mu'_r}{d\mu} + \mu'_1 \mu'_r$
Centrais	$\mu_{r+1} = \frac{g(\mu, \phi)}{g'(\mu, \phi)} \frac{d\mu_r}{d\mu} + r \mu_2 \mu_{r-1}$
Fatoriais	$\mu'_{[r+1]} = \frac{g(\mu, \phi)}{g'(\mu, \phi)} \frac{d\mu'_{[r]}}{d\mu} + \mu'_{[r]}(\mu'_{[1]} - r)$

3 Estimação

Considerando-se, em (1), ϕ conhecido, será usado o método de máxima verossimilhança para estimar o parâmetro μ . Seja Y_1, \dots, Y_n uma amostra aleatória de tamanho n de uma variável aleatória com distribuição na classe MPSD (1) e seja \bar{Y} a média amostral. A função de verossimilhança $L = L(\mu)$ é expressa como

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

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

$$\frac{1}{L} \frac{dL}{d\mu} = n \frac{g'(\mu, \phi)}{g(\mu, \phi)} (\bar{y} - \mu) = 0. \quad (2)$$

A solução de (2) é a equação

$$\bar{y} = \mu. \quad (3)$$

Como $\mu = \mu(\theta)$ e se o interesse estiver na estimação do parâmetro θ , então, se $\mu(\theta)$ for inversível, a estimativa de máxima verossimilhança (EMV) de θ , obtida pela inversão de (3), é igual a

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

Se μ não for inversível, pode-se resolver (3), iterativamente, utilizando-se o método de Newton-Raphson.

Consul e Famoye (2006) expressam o viés, $b(\hat{\theta})$, do EMV do parâmetro θ , como

$$b(\hat{\theta}) = -\frac{B_1}{2n A_1^2}, \quad (4)$$

sendo

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

e

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

Para a classe MPSD, tem-se

$$A_1 = \left[\frac{g'(\mu, \phi)}{g(\mu, \phi)} \right]^2 \mu_2 \quad (5)$$

e

$$B_1 = \left[\frac{g'(\mu, \phi)}{g(\mu, \phi)} \right]^3 \left\{ \mu_3 + \frac{g(\mu, \phi)g''(\mu, \phi) - [g'(\mu, \phi)]^2}{[g'(\mu, \phi)]^2} \mu_2 \right\}, \quad (6)$$

em que μ_2 e μ_3 são, respectivamente, o segundo e terceiro momentos centrais.

Substituindo (5) e (6) em (4), o viés de $\hat{\theta}$ é

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

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

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

4 Alguns casos especiais da classe MPSD

Existem distribuições na classe MPSD que podem ser parametrizadas de tal forma que a média delas seja igual ao parâmetro μ , ou seja, $E(Y) = \mu'_1 = \mu$ e que estão apresentadas na Tabela 2. Essa parametrização é importante para estudar essas distribuições no contexto de modelos de regressão. Algumas distribuições que pertencem à classe MPSD serão descritas a seguir, com algumas de suas propriedades. As expressões para variância dessas distribuições estão apresentadas na Tabela A1 (do Apêndice) e dos coeficientes de assimetria e curtose na Tabela A2 (do Apêndice).

Na Tabela 3, estão relacionados os desvios unitários para alguns modelos discretos para dados na forma de contagem formulados por Cordeiro et al. (2009).

4.1 Distribuição binomial negativa generalizada

A distribuição binomial negativa generalizada tem aplicações importantes em vários campos de estudo como em teoria das filas, em processos de reação de polimerização em química, para analisar o número de acidentes e para dados de contagens em que a variância observada é maior do que a média dos dados (JAIN; CONSUL, 1971).

A função de probabilidade da distribuição binomial negativa generalizada é expressa como

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

sendo $y = 0, 1, \dots$, $\mu > 0$, $\beta > 0$ e $m > 0$ conhecido, mas não, necessariamente, inteiro. Tem-se, ainda, que $f(\mu) = \left[\frac{m + (\beta - 1)\mu}{m + \beta\mu} \right]^{-m}$ e $g(\mu) =$

$\left(\frac{\mu}{m + \beta\mu} \right) \left[\frac{m + (\beta - 1)\mu}{m + \beta\mu} \right]^{\beta - 1}$. A variância da distribuição binomial negativa generalizada é expressa como $\mu[1 + \beta\mu m^{-1}] [1 + (\beta - 1)\mu m^{-1}]$.

Tabela 2 - Casos especiais da classe MPSD

Distribuição	$f(\mu, \phi)$	$g(\mu, \phi)$	$a(y, \phi)$
Binomial	$\left(1 + \frac{\mu}{m - \mu}\right)^m$	$\frac{\mu}{m - \mu}$	$\frac{\Gamma(m + 1)}{y! \Gamma(m - y + 1)}$
Binomial Negativa	$\left(1 - \frac{\mu}{\mu + m}\right)^{-m}$	$\frac{\mu}{\mu + m}$	$\frac{\Gamma(m + y)}{y! \Gamma(m)}$
Binomial Negativa Generalizada	$\left[\frac{m + (\phi - 1)\mu}{m + \phi\mu}\right]^{-m}$	$\left(\frac{\mu}{m + \phi\mu}\right) \left[\frac{m + (\phi - 1)\mu}{m + \phi\mu}\right]^{\phi - 1}$	$\frac{m\Gamma(m + \phi y + 1)}{y! \Gamma(m + \phi y - y + 1)}$
Borel	$1 - \frac{1}{\mu}$	$g(\mu) = \left(1 - \frac{1}{\mu}\right) e^{-1 + \frac{1}{\mu}}$	$\frac{y^{y-2}}{(y-1)!}$
Borel-Tanner	$\left(1 - \frac{m}{\mu}\right)^m$	$\left(1 - \frac{m}{\mu}\right) e^{-1 + \frac{m}{\mu}}$	$\frac{m y^{y-m-1}}{(y-m)!}$
Consul	$\frac{\mu - 1}{\mu(\phi - 1) + 1}$	$\phi^{-\phi} (1 - \mu^{-1})(\phi - 1 + \mu^{-1})^{\phi - 1}$	$\frac{\Gamma(\phi y + 1)}{y! \Gamma(\phi y - y)}$
Delta-binomial	$\left[\frac{\mu - m}{\mu(\phi - 1) + m}\right]^m$	$\phi^{-\phi} \left(1 - \frac{m}{\mu}\right) \left(\phi - 1 + \frac{m}{\mu}\right)^{\phi - 1}$	$\frac{m\Gamma(\phi y + 1)}{y(y-m)! \Gamma(\phi y - y + m + 1)}$
Geeta	$\left(\frac{\mu - 1}{\phi\mu - 1}\right)$	$\left(\frac{\mu - 1}{\phi\mu - 1}\right) \left[\frac{(\phi - 1)\mu}{\phi\mu - 1}\right]^{\phi - 1}$	$\frac{\Gamma(\phi y - 1)}{y! \Gamma(\phi y - y)}$
Haight	$\frac{\mu - 1}{2\mu - 1}$	$\frac{\mu(\mu - 1)}{(2\mu - 1)^2}$	$\frac{(2y - 2)!}{y!(y - 1)!}$
Poisson	e^μ	μ	$\frac{1}{y!}$
Poisson Generalizada	$e^{\mu(1 + \phi\mu)^{-1}}$	$\frac{\mu e^{-\phi\mu(1 + \phi\mu)^{-1}}}{1 + \phi\mu}$	$\frac{(1 - \phi y)^{y-1}}{y!}$

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

Observando-se a Figura 4.1, nota-se pelo gráfico (a) que a distribuição binomial negativa generalizada é simétrica em torno da média μ , para $\beta = 0$; tem uma leve assimetria à esquerda para $\beta = 1$ e, nos demais casos, é assimétrica à esquerda. Pelo gráfico (b), tem-se que a assimetria é decrescente para $\beta = 0$ e $\beta = 1$ e, para β com valores maiores a assimetria é crescente. A curtose, como pode ser visto no gráfico (c), é crescente para $\beta = 0$ e $\beta = 1$ e tende a zero para $\beta = 5$ e $\beta = 7$.

Tabela 3 - Desvios unitários para alguns modelos em série de potências modificadas

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

4.2 Distribuição de Borel-Tanner

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

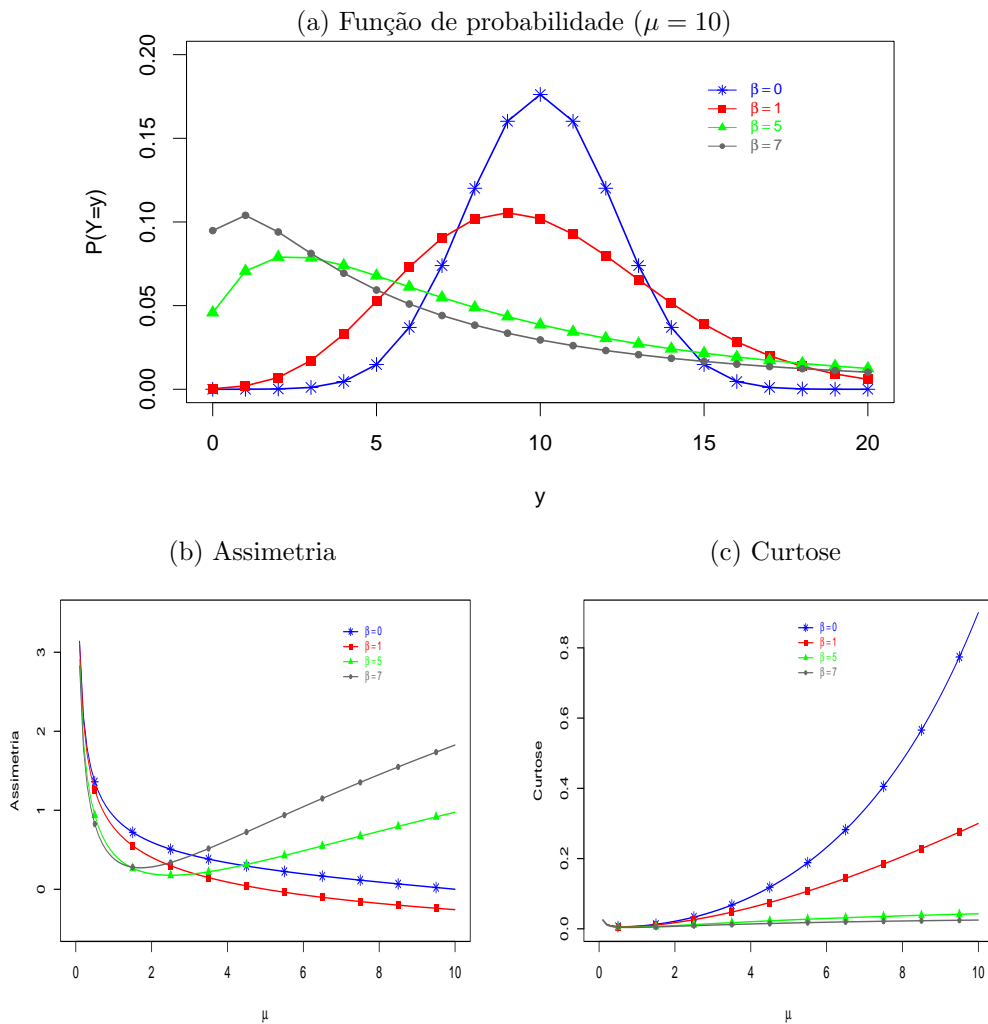


Figura 4.1 - Gráficos da distribuição binomial negativa generalizada para $m = 20$.

A variável aleatória Y tem distribuição de Borel-Tanner se

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

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

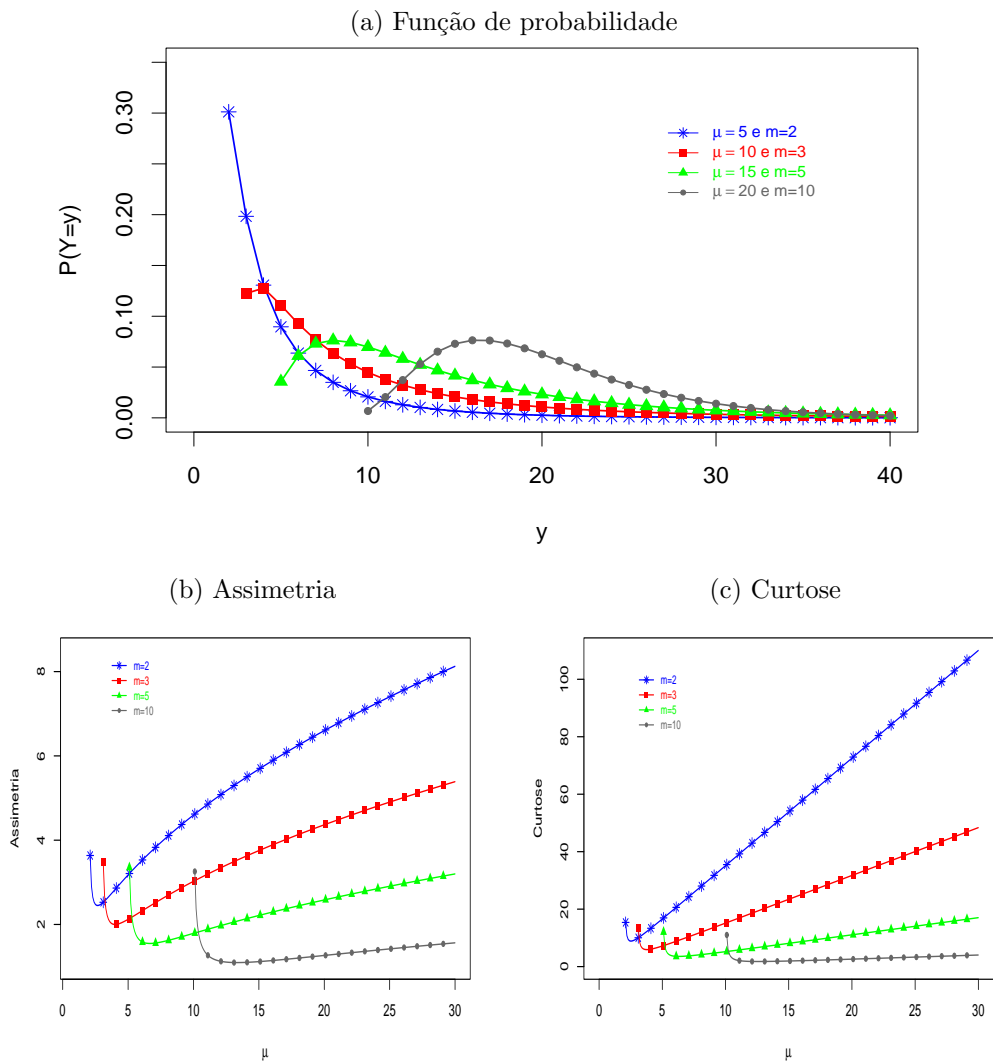


Figura 4.2 - Gráficos da distribuição de Borel-Tanner.

Observando-se a Figura 4.2, nota-se pelo gráfico (a) que a distribuição de Borel-Tanner é assimétrica à esquerda. Pelo gráfico (b), tem-se que a assimetria tem um decréscimo para valores pequenos de μ e, em seguida, tem comportamento crescente para valores crescentes de μ , tendo crescimento mais acentuado para $m = 2$. A curtose, como pode ser visto no gráfico (c), tem comportamento similar ao da assimetria.

4.3 Distribuição de Consul

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

Um importante problema em estudos de epidemias é encontrar a distribuição de probabilidade do número total de infectados, $N(u)$, em qualquer parte do *habitat*, sendo as pessoas infectadas a partir de uma única pessoa infectada em cada local, u , até o momento da extinção da epidemia. Nesse caso, supondo $X_0 = 1$ com probabilidade um e o número de novas pessoas infectadas, entre as suscetíveis, tendo distribuição binomial, o número de pessoas suscetíveis é finito, m , como utilizado por Kumar (1981). Se o processo de infecção continua assim, indefinidamente, o número de pessoas infectadas em dado momento pode ser analisado, usando-se o modelo de Consul.

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

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

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

Observando-se a Figura 4.3, nota-se pelo gráfico (a) que a distribuição de Consul é assimétrica à direita. Pelo gráfico (b), tem-se que a assimetria é decrescente para valores crescentes de μ . A curtose, como pode ser visto no gráfico (c), é crescente para valores crescentes de μ .

4.4 Distribuição delta-binomial

Se X_1, \dots, X_m são variáveis aleatórias independentemente distribuídas com distribuição Consul, então a soma amostral, $Y = \sum_{i=1}^m X_i$, tem distribuição delta-binomial, isto é, possui a seguinte função de probabilidade:

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

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

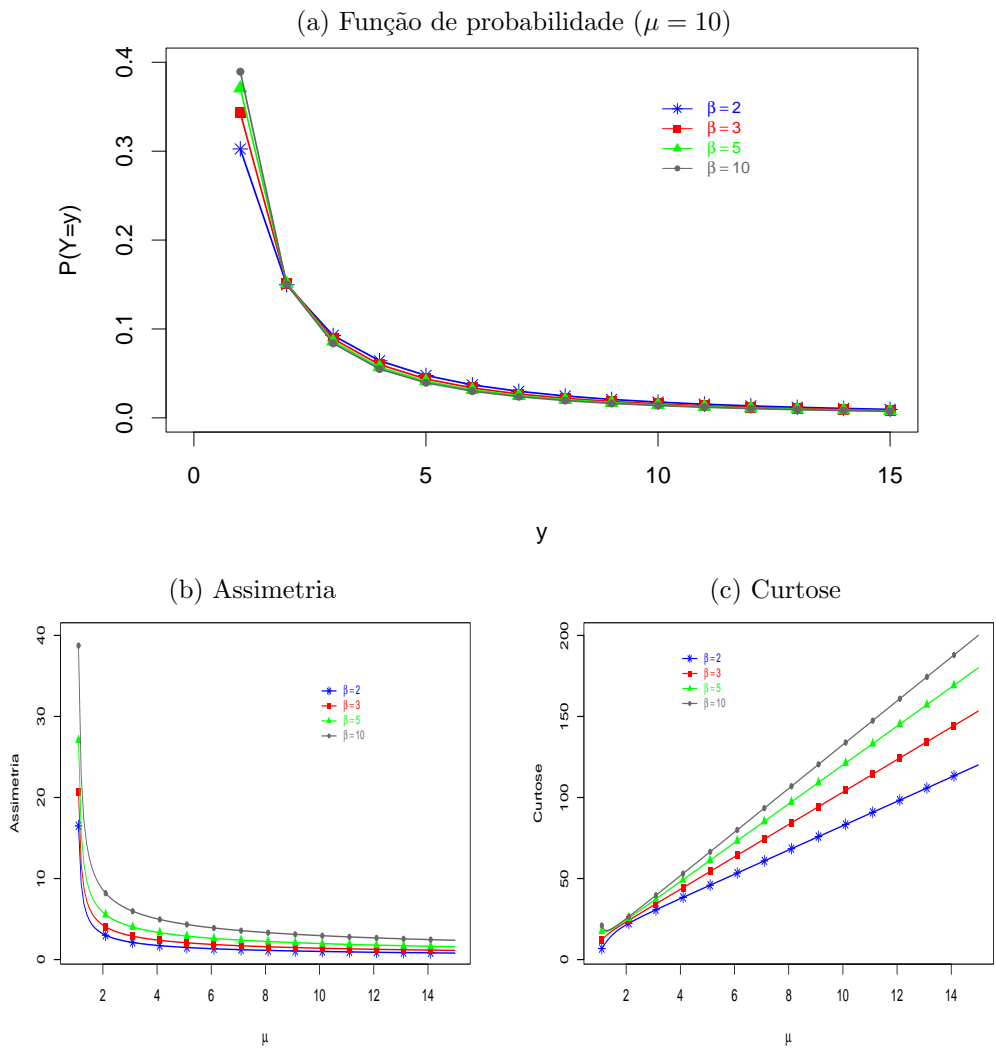


Figura 4.3 - Gráficos da distribuição de Consul.

Observando-se a Figura 4.4, nota-se pelo gráfico (a) que a distribuição delta-binomial para $m = 2$ é assimétrica à direita. Pelo gráfico (b), tem-se que a assimetria decresce para valores pequenos de μ e, em seguida, tem comportamento crescente para valores crescentes de μ . A curtose, como pode ser visto no gráfico (c), tem comportamento similar ao da assimetria.

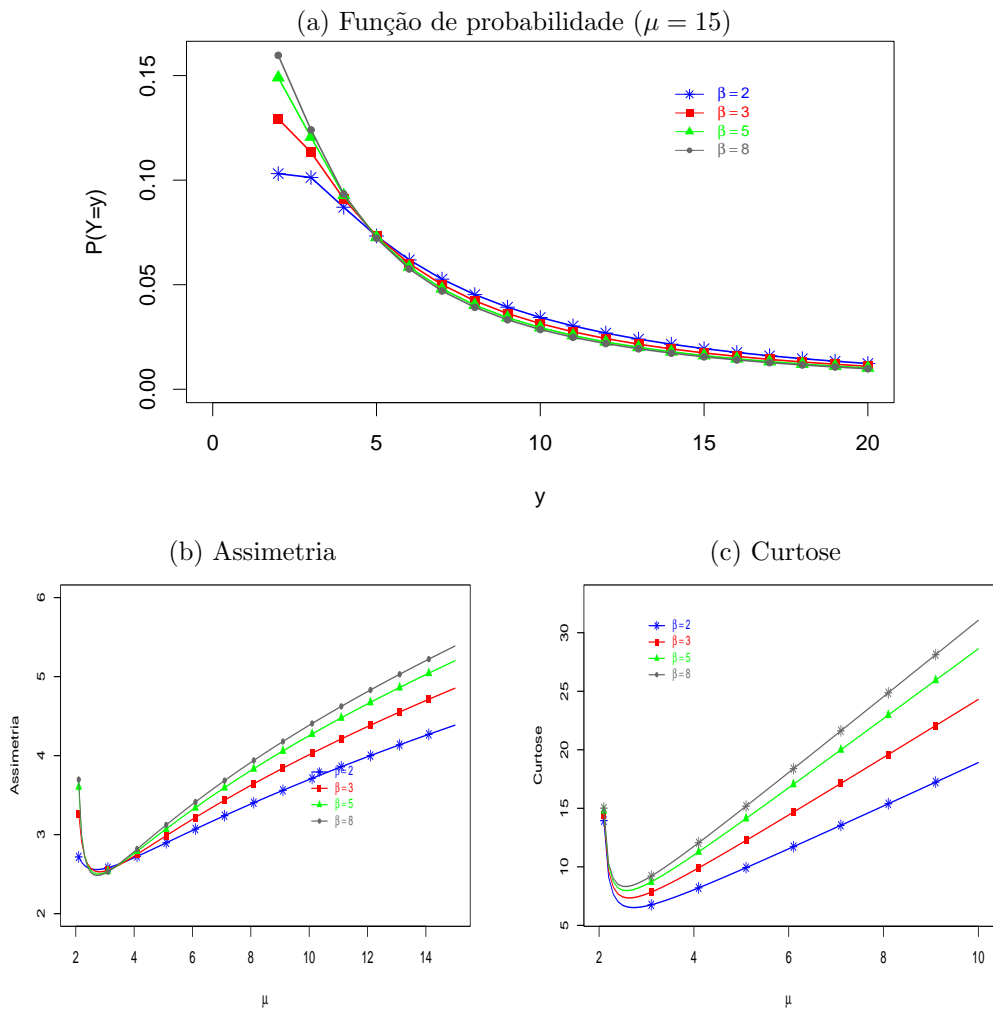


Figura 4.4 - Gráficos da distribuição delta-binomial para $m = 2$.

4.5 Distribuição de Geeta

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

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

A função de probabilidade da distribuição de Geeta é expressa como

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

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

Observando-se a Figura 4.5, nota-se pelo gráfico (a) que a distribuição de Geeta para $\beta = 2$ é assimétrica à esquerda e em formato de "L". Pelo gráfico (b), tem-se que a assimetria decresce para valores pequenos de μ e, em seguida, tem comportamento crescente para valores crescentes de μ . A curtose, como pode ser visto no gráfico (c), tem comportamento similar à assimetria.

4.6 Distribuição de Poisson generalizada

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

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

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

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

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

Observando-se a Figura 4.6, nota-se pelo gráfico (a) que a distribuição de Poisson generalizada é simétrica para $\beta = 0$ e é assimétrica à esquerda para os demais valores de β . Pelo gráfico (b), tem-se que a assimetria tende a zero para $\beta = 0$ e, para os demais valores de β , decresce inicialmente, mas em seguida tem comportamento crescente. A curtose, como pode ser visto no gráfico (c), tende a zero para $\beta = 0$ e, nos demais casos, é crescente para valores crescentes de μ .

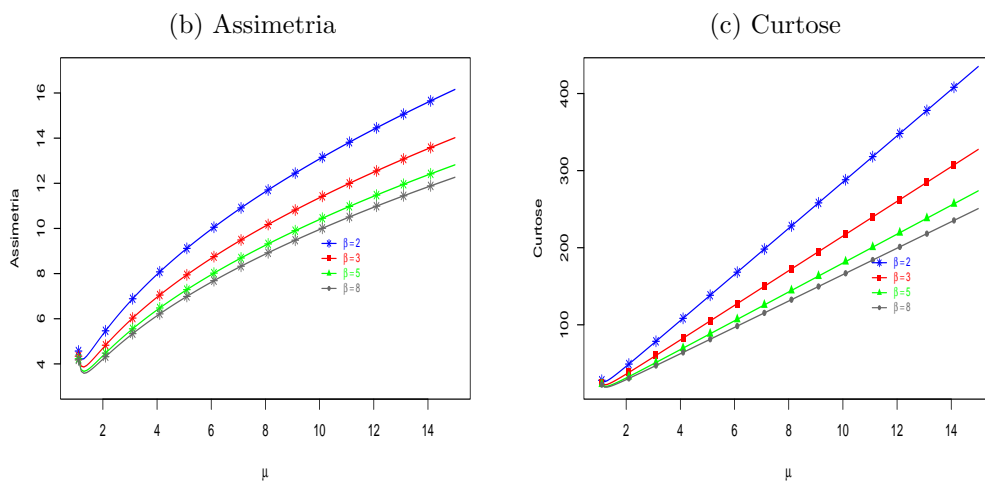
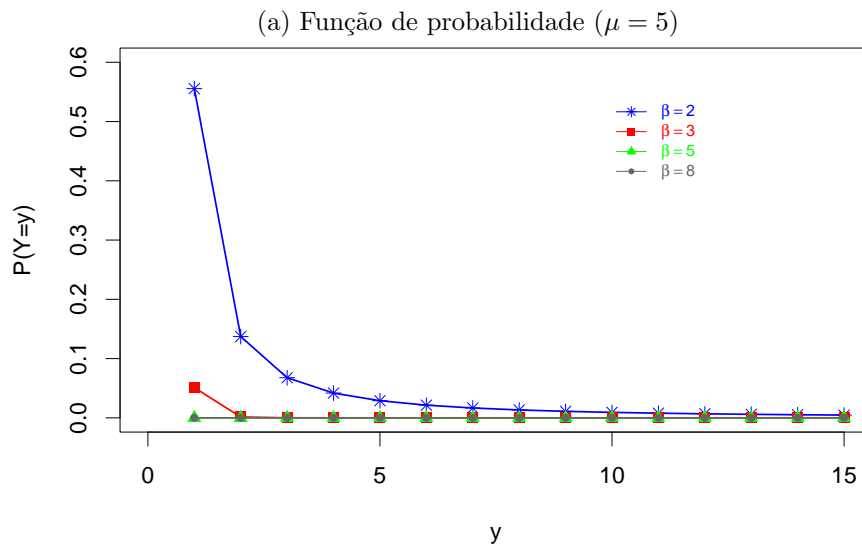


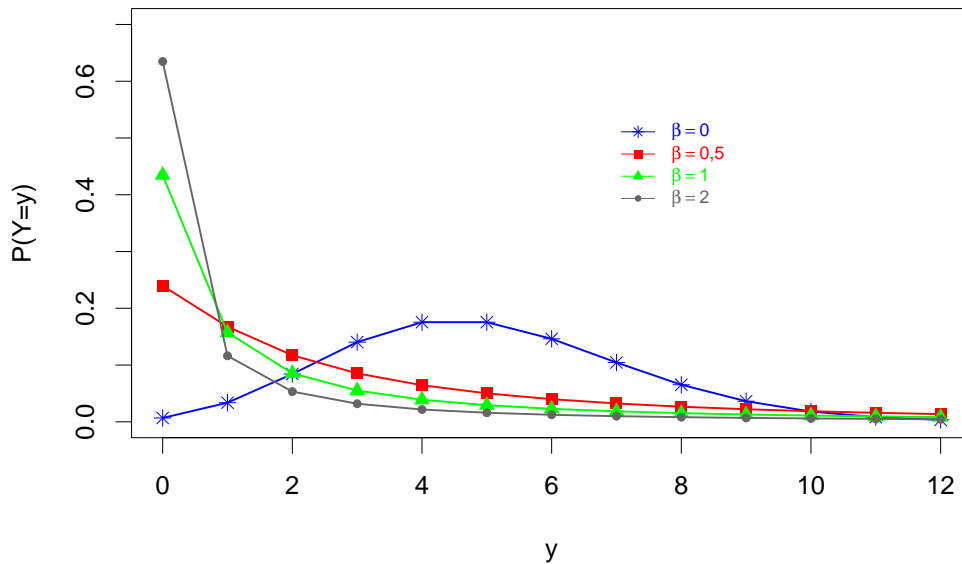
Figura 4.5 - Gráficos da distribuição de Geeta.

5 Modelos não-lineares generalizados em série de potências modificadas

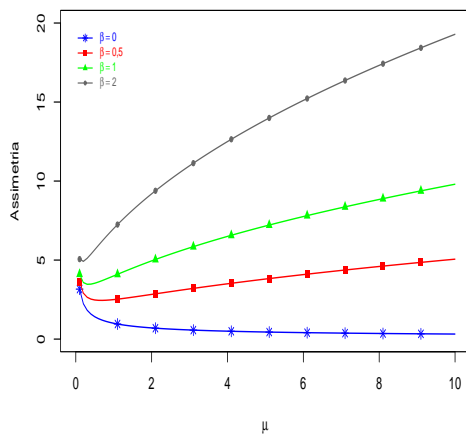
Seja $y = (y_1, \dots, y_n)^T$ um conjunto de n observações independentes de uma variável aleatória Y que tem distribuição com função de probabilidade

$$P(Y = y) = \frac{a(y, \phi)g(\mu, \phi)^y}{f(\mu, \phi)}, \quad y \in T, \quad (7)$$

(a) Função de probabilidade ($\mu = 5$)



(b) Assimetria



(c) Curtose

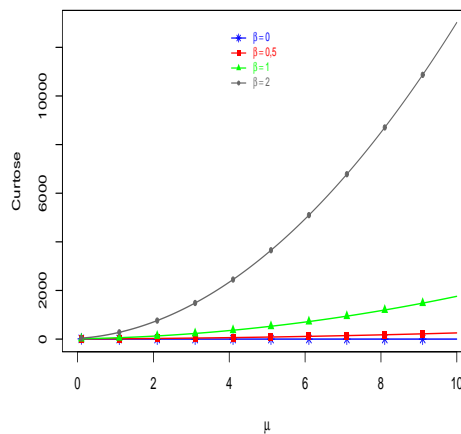


Figura 4.6 - Gráficos da distribuição de Poisson generalizada.

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

$$h(\mu) = \eta = \eta(X, \beta), \quad (8)$$

em que $h(\cdot)$ é uma função monótona e diferenciável, η é um preditor linear se $\eta(X, \beta) = X\beta$, caso contrário, diz-se que η é um preditor não-linear, X é a matriz modelo de dimensão $n \times p$ de posto completo $p < n$ e $\beta = (\beta_1, \dots, \beta_p)^T$ é o conjunto de parâmetros desconhecidos a serem estimados. Os modelos não-lineares generalizados em série de potências modificadas são definidos por (7) e (8) (CORDEIRO et al., 2009).

Os MLG usados para análise de dados na forma de contagens são restritos às distribuições de Poisson, binomial e binomial negativa. Cordeiro et al. (2009) introduziram uma nova classe de modelos, a que chamaram modelos não-lineares generalizados em série de potências (PSGNLMs), que incluem como casos especiais outros modelos discretos, tais como, as distribuições em série de potências modificadas, e podem ser utilizados para analisar dados na forma de contagens.

Após uma função de regressão ter sido especificada, diferentes modelos de regressão, tais como aqueles com distribuição de Poisson generalizada ou binomial negativa generalizada podem ser usados para analisar conjuntos de dados na forma de contagens. A distribuição em (7) pertence à família exponencial uniparamétrica quando ϕ é conhecido e os parâmetros não-lineares seguem a abordagem do MLG. O estimador de máxima verossimilhança (EMV) $\hat{\beta}$ é, assintoticamente, normal com média β e matriz de covariâncias, consistentemente, estimada por $(\hat{X}^T W \hat{X})^{-1}$, em que $\hat{X} = \partial\eta/\partial\beta$.

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

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

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

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

Considerações finais

A classe de distribuições em série de potências modificadas (MPSD) abrange inúmeras distribuições discretas, tais como, binomial, binomial negativa, Poisson, série logarítmica, dentre outras. Existem artigos e livros que descrevem as propriedades dessa classe de distribuições, bem como inúmeras aplicações. Porém, apenas recentemente, Cordeiro et al. (2009) introduziram essa classe no contexto de modelos de regressão não-linear e com essa nova classe de modelos, a qual chamaram modelos não-lineares generalizados em série de potências (PSGNLMs), a análise de dados de contagem não fica restrita às distribuições de Poisson, binomial e binomial negativa, podendo-se utilizar outros modelos discretos, tais

como, binomial negativo generalizado, Borel, Borel-Tanner, Consul, delta-binomial, Haight ou Poisson generalizado.

Neste trabalho, fez-se uma revisão da classe MPSD, relacionando as principais propriedades, e descrevendo alguns casos especiais.

BRITO, E.; CORDEIRO, G. M.; DEMÉTRIO, C. G. B. Modified power series distributions. *Rev. Bras. Biom.*, São Paulo, v.24, n.4, p.1-20, 2010.

- **ABSTRACT:** *We present some mathematical properties of the class of modified power series distributions (MPSD) introduced by Gupta (1974, Sankhya B). This class of distribution includes some important distributions such as the binomial, negative binomial, Poisson, logarithmic series, among others. More recently, this class of models was studied in the context of regression models by Cordeiro et al. (2009, CSDA).*
- **KEYWORDS:** *Discrete distribution; probability function; moments generating function; likelihood function; moments.*

Referências

CONSUL, P. C. *Geeta distribution and its properties. Commun. Stat., Theory Methods*, New York, v.19, n.8, p.3051-3068, 1990a.

CONSUL, P. C. New class of location-parameter discrete probability distributions and their characterizations. *Commun. Stat., Theory Methods*, New York, v.19, n.12, p.4653-4666, 1990b.

CONSUL, P. C.; FAMOYE, F. *Lagrangian probability distributions*. Boston: Birkhäuser, 2006. 352p.

CONSUL, P. C.; JAIN, G. C. A generalization of the poisson distribution. *Technometrics*, Washington, v.15, n.4, p.791-799, 1973a.

CONSUL, P. C.; JAIN, G. C. On some interesting properties of the generalized poisson distribution. *Biometr. Z.*, Berlin, v.15, p.495-500, 1973b.

CORDEIRO, G. M.; ANDRADE, M. G. de; CASTRO, M. de. Power Series Generalized Nonlinear Models. *CSDA - Comput. Stat. Data Anal.*, Amsterdam, v.53, n.4, p.1155-1166, 2009.

GUPTA, R. C. (1974). Modified power series distribution and some of its applications. *Sankhya Sér. B*, Calcutá, v.36, p.288-298, 1974.

GUPTA, R. C.; SINGH, J. Estimation of probabilities in the class of modified power series distributions. *Statistics*, Berlin, v.13, p.71-77, 1982.

JAIN, G. C.; CONSUL, P. C. A generalized negative binomial distribution. *SIAM J. Appl. Math.*, Philadelphia, v.21, n.4, p.501-513, 1971.

JAIN, G. C.; GUPTA, R. C. *A logarithmic type distribution. Trab. Estad. Invest. Oper.*, Madrid, v.24, n.1-2, p.99-105, 1973.

JAIN, G. C.; KHAN, S. H. On an exponential family. *Statistics*, Berlin, v.10, p.153-168, 1979.

KEMP, A. W.; KEMP, C. D. On a distribution associated with certain stochastic processes. *J. R. Stat. Soc., Ser. B, Methodol.*, London, v.30, n.1, p.160-163, 1968.

KUMAR, A. Some applications of lagrangian distributions in queueing theory and epidemiology. *Commun. Stat., Theory Methods*, New York, v.10, p.1429-1436, 1981.

TANNER, J. A problem of interference between two queues. *Biometrika*, London, v.40, p.58-69, 1953.

Recebido em 11.02.2009.

Aprovado após revisão em 11.11.2010.

APÊNDICE - Variância, Assimetria e Curtose

O presente apêndice tem por objetivo apresentar as expressões de variância, assimetria e curtose das distribuições na classe MPSD parametrizadas na média apresentadas na Seção 4.

A variância de uma variável aleatória Y na classe MPSD é dada por

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

Na Tabela A1 estão relacionadas as expressões para a variância das distribuições apresentadas na Seção 4.

Tabela A1 - Variância de algumas distribuições na classe MPSD	
Distribuição	Variância
Binomial	$\mu \left(1 - \frac{\mu}{m}\right)$
Binomial Negativa	$\mu \left(1 + \frac{m}{\mu}\right)$
Binomial Negativa Generalizada	$\mu[1 + \beta\mu m^{-1}] [1 + (\beta - 1)\mu m^{-1}]$
Borel	$\mu^2(\mu - 1)$
Borel-Tanner	$\frac{(\mu - m)\mu^2}{m^2}$
Consul	$\frac{\mu(\mu - 1)[\mu(\beta - 1) + 1]}{\beta}$
Delta-binomial	$\frac{\mu(\mu - m)[\mu(\beta - 1) + m]}{\beta m^2}$
Geeta	$\mu(\mu - 1)(\beta\mu - 1)(\beta - 1)^{-1}$
Haight	$\mu(\mu - 1)(2\mu - 1)$
Poisson	μ
Poisson generalizada	$\mu(1 + \beta\mu)^2$

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

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

sendo $\mu_4 = \frac{g(\theta)}{g'(\theta)} \frac{d\mu_3}{d\theta} + 3\mu_2^2$, $\mu_3 = \frac{g(\theta)}{g'(\theta)} \frac{d\mu_2}{d\theta}$ e μ_2 a variância da classe MPSD (1).

Na Tabela A2 estão relacionadas as expressões para assimetria e curtose das distribuições apresentadas na Seção 4.

Tabela A2 - Assimetria e curtose de algumas distribuições na classe MPSPD

Distribuição	Assimetria	Curtose
Binomial	$\frac{m-2\mu}{[m\mu(m-\mu)]^{1/2}}$	$\frac{m^2+3\mu(m-\mu)(m-3)}{m\mu(m-\mu)}$
Binomial Negativa	$\frac{m+2\mu}{[m\mu(m+\mu)]^{1/2}}$	$\frac{m^2+36m\mu^3}{m\mu(m+\mu)}$
Binomial Negativa Generalizada	$\frac{m^2-2m\beta\mu-2m\mu+3\beta^2\mu^2-3\beta\mu^2}{m\{\mu(m+\beta\mu)[m+(\beta-1)\mu]\}^{1/2}}$	$\frac{m^2+36m\mu^3}{m^2\mu(m+\beta\mu)[m+(\beta-1)\mu]}$
Borel	$\frac{3\mu-2}{(\mu-1)^{1/2}}$	$\frac{15\mu^2-20m\mu+6m^2}{t-1}$
Borel-Tanner	$\frac{m(\mu-m)^{1/2}}{3\mu-2m}$	$\frac{15\mu^2-20m\mu+6m^2}{m^2(\mu-m)}$
Consul	$\frac{3\beta\mu^2-3\mu^2+4\mu-2\beta\mu-1}{\mu(\mu-1)[(\beta-1)\mu+1]^{1/2}}$	$\frac{15\beta^2\mu^4-30\beta\mu^4+60\beta\mu^3-20\beta^2\mu^3-36\beta\mu^2+15\mu^4-40\mu^3+36\mu^2-12\mu+6\beta^2\mu^2+6\beta\mu+1}{15\beta^2\mu^4-30\beta\mu^4+60m\beta\mu^3-20m\beta^2\mu^3-36m^2\mu^2+15\mu^4-40m\mu^3+36m^2\mu^2-12m^3\mu+6m^3\beta\mu+m^4}$
Delta-binomial	$\frac{3\beta\mu^2-3\mu^2+4m\mu-2m\beta\mu-m^2}{[m^2\beta\mu(\mu-m)[(\beta-1)\mu+m]]^{1/2}}$	$\frac{m^2\beta\mu(\mu-m)[(\beta-1)\mu+m]}{(\beta-1)\mu(\mu-1)(\beta\mu-1)}$
Geeta	$\frac{3\beta\mu^2-2\beta\mu-2\mu+1}{[(\beta-1)\mu(\mu-1)(\beta\mu-1)]^{1/2}}$	$\frac{60\mu^4-120\mu^3+78\mu^2-18\mu+1}{\mu(\mu-1)(2\mu-1)}$
Haight	$\frac{6\mu^2-6\mu+1}{[\mu(\mu-1)(2\mu-1)]^{(0.5)}}$	μ^{-1}
Poisson	$\frac{\mu}{\mu^{-1/2}}$	μ^{-1}
Poisson generalizada	$\frac{3\beta^2\mu^2+4\beta\mu+1}{\mu^{1/2}(\beta\mu+1)}$	$\frac{15\beta^4t^4+40\beta^3\mu^3+36\beta^2\mu^2+12\beta\mu+1}{\mu(\beta\mu+1)}$