

Distribuição bivariada gama beta II: soma, produto e proporção das variáveis componentes

A.P.M. SILVA¹, UFSJ - Universidade Federal de São João del-Rei, Campus de Sete Lagoas, CSL, Cx. P. 56, 35701-970 Sete Lagoas, MG, Brasil

J.A. RODRIGUES², DEPEN - Departamento de Ensino, IFBA - Instituto Federal de Educação, Ciência e Tecnologia da Bahia, 45823-431 Eunápolis, BA, Brasil

L.M. CHAVES³, DEX - Departamento de Ciências Exatas, UFLA - Universidade Federal de Lavras, Cx.P. 3037, 37200-000 Lavras, MG, Brasil

D.J. SOUZA⁴, DEX - Departamento de Ciências Exatas, UFLA - Universidade Federal de Lavras, Cx.P. 3037, 37200-000 Lavras, MG, Brasil.

Resumo. Os modelos bivariados tem sido utilizados com sucesso na análise de processos hidrológicos. Neste trabalho, são deduzidas as distribuições exatas das variáveis $U = X + Y$, $P = XY$ e $Q = X/(X + Y)$ juntamente com seus respectivos momentos quando X e Y seguem o modelo bivariado gama beta II. Essas funções descrevem importantes variáveis hidrológicas. Os resultados obtidos são aplicados em dados de precipitações pluviométricas ocorridas na cidade de Passo Fundo - RS.

Palavras-chave. Distribuição gama beta II, combinação de variáveis aleatórias, precipitação pluviométrica.

1. Introdução

Motivados pelo crescente uso principalmente na análise de dados não normais, vários modelos bivariados podem ser encontrados na literatura [2]. Os trabalhos desenvolvidos por [1] impulsionaram o uso da distribuição gama bivariada na análise de processos hidrológicos, sendo atualmente um dos mais utilizados. Por exemplo, [15] estuda a aplicabilidade da distribuição gama bivariada de Smith na análise de frequência das variáveis hidrológicas duração e volume; [16] apresentam uma revisão de vários modelos gama, com diferentes parâmetros de escala e forma, apontando as vantagens e desvantagens de cada modelo no estudo de precipitação; [9] estudam o comportamento de dados de seca do Estado de Nebraska considerando a

¹anapaula@ufsj.edu.br

²jailsondearaujo@yahoo.com.br

³lucas@dex.ufla.br

⁴devaniljaques@dex.ufla.br

distribuição gama bivariada de Cherian e [10] realizam o mesmo estudo com a distribuição exponencial bivariada de Friday e Patil; [7] aplica um modelo bivariado gama exponencial na modelagem de dados de seca.

Considerando que X e Y se distribuem segundo um modelo bivariado, funções dessas variáveis aleatórias, expressas por $U = X + Y$, $P = XY$ e $Q = X/(X + Y)$, têm um significado físico importante de modo que diferentes autores têm trabalhado no sentido de caracterizar essas distribuições e aplicá-las em diferentes áreas do conhecimento, em particular em hidrologia [6, 11, 4, 8, 14]. Por exemplo, se X representa o período de chuva e Y o período contíguo sem ocorrência de chuva, $U = X + Y$ denota o período climático e $Q = X/(X + Y)$ a proporção de chuva. Neste contexto, o trabalho tem por objetivo apresentar a distribuição bivariada Gama Beta tipo II e deduzir a distribuição exata das variáveis $U = X + Y$, $P = XY$ e $Q = X/(X + Y)$ sob a pressuposição de que X e Y seguem esse modelo. Como aplicação é realizado o ajuste dessas distribuições a dados de precipitação pluviométrica do município de Passo Fundo, RS.

1.1. Distribuições beta

Distribuição beta tipo I: Uma variável aleatória X tem distribuição beta tipo I com parâmetros $\alpha > 0$ e $\beta > 0$ quando sua função densidade de probabilidade (fdp) é da forma:

$$f(x) = \frac{x^{\alpha-1} (1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad (1.1)$$

em que $0 < x < 1$ e $B(\alpha, \beta)$ representa a função beta,

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt. \quad (1.2)$$

Simbolicamente, quando X possui função densidade de probabilidade dada por (1.1), denotamos por $X \sim BI(\alpha, \beta)$, ou apenas, $X \sim B(\alpha, \beta)$.

Distribuição beta tipo II: Uma variável aleatória X tem distribuição beta tipo II com parâmetros $\alpha > 0$, $\beta > 0$ e $\lambda > 0$ quando sua função densidade de probabilidade (fdp) é da forma:

$$f(x) = \frac{\lambda^\beta}{B(\alpha, \beta)} x^{\alpha-1} (\lambda + x)^{-(\alpha+\beta)}, \quad (1.3)$$

em que $x > 0$. Quando X possui função densidade de probabilidade dada por (1.3), denotamos por $X \sim BII(\alpha, \beta)$.

As duas distribuições beta estão relacionadas intrinsecamente por meio de uma transformação. Várias características referentes aos dois modelos podem ser encontradas em [3, 5].

Além da função beta, os cálculos envolvidos no trabalho incluem o uso de outras funções especiais como a função gama,

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} \exp(-t) dt, \quad (1.4)$$

a função $\psi(\cdot)$,

$$\psi(\alpha) = \frac{d}{d\alpha} [\ln \Gamma(\alpha)], \quad (1.5)$$

a função hipergeométrica confluyente,

$${}_1F_1(a; b; x) = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)} \int_0^1 t^{a-1} (1-t)^{b-a-1} \exp(-xt) dt, \quad (1.6)$$

com $0 < a < b$ e $b \neq 0, -1, -2, \dots$ e a função cilíndrica parabólica,

$$D_\nu(x) = \frac{2^{\nu/2} \exp(-x^2/4)}{\Gamma(-\nu/2)} \int_0^\infty t^{-(1+\nu/2)} (1+t)^{\nu/2-1/2} \exp\left(\frac{-x^2 t}{2}\right) dt. \quad (1.7)$$

As propriedades dessas funções especiais podem ser vistas em [3, 12]. Serão ainda utilizados os importantes Lemas 1.1 e 1.2.

Lema 1.1. (Equação 2.3.6.1, [13]). Se $a > 0$,

$$\int_0^a x^{\alpha-1} (a-x)^{\beta-1} \exp(-px) dx = B(\alpha, \beta) a^{\alpha+\beta-1} {}_1F_1(\alpha; \alpha+\beta; -ap).$$

Lema 1.2. (Equação 2.3.15.1, [13]). Se $\alpha > 0$ e $\beta > 0$,

$$\int_0^\infty x^{\alpha-1} \exp(-rx^2 - qx) dx = \Gamma(\alpha)(2r)^{-\alpha/2} \exp\left(\frac{q^2}{8r}\right) D_{-\alpha}\left(\frac{q}{\sqrt{2r}}\right).$$

2. O modelo

A distribuição bivariada gama beta tipo II tem função densidade de probabilidade (fdp) conjunta dada por:

$$f_{X,Y}(x, y) = K x^{\alpha-1} y^{\beta-1} \exp[-(ax + cxy)] \quad (2.1)$$

com $x > 0$, $y > 0$, $a > 0$, $c > 0$, $0 < \beta < \alpha$ e K é a constante de normalização definida por

$$K = \frac{c^\beta a^{\alpha-\beta}}{\Gamma(\beta)\Gamma(\alpha-\beta)}.$$

Esse modelo pertence à família gama bivariada de Arnould (ver [2]) e com essa parametrização ainda não foi abordado na literatura.

2.1. Funções Densidade de Probabilidade

Considerando que X e Y têm distribuição conjunta dada pelo modelo (2.1), as fdps marginais de X e Y são respectivamente:

$$\begin{aligned} f_X(x) &= \int_0^{\infty} f(x, y) dy = \int_0^{\infty} Kx^{\alpha-1}y^{\beta-1} \exp[-(ax + cxy)] dy \\ &= \frac{a^{\alpha-\beta}}{\Gamma(\alpha-\beta)} x^{\alpha-\beta-1} \exp(-ax), \end{aligned} \quad (2.2)$$

$$\begin{aligned} f_Y(y) &= \int_0^{\infty} f(x, y) dx = \int_0^{\infty} Kx^{\alpha-1}y^{\beta-1} \exp[-(ax + cxy)] dx \\ &= \frac{c^{\beta} a^{\alpha-\beta} \Gamma(\alpha)}{\Gamma(\beta) \Gamma(\alpha-\beta)} y^{\beta-1} \left[c \left(\frac{a}{c} + y \right) \right]^{-\alpha} \\ &= \frac{(a/c)^{\alpha-\beta}}{B(\beta, \alpha-\beta)} y^{\beta-1} \left(\frac{a}{c} + y \right)^{-\alpha}. \end{aligned} \quad (2.3)$$

Portanto, $X \sim G(\alpha - \beta, a)$ e $Y \sim \text{BII}(\beta, \alpha - \beta, a/c)$.

Nos Teoremas 2.1, 2.2 e 2.3 são deduzidas as fdps das variáveis $U = X + Y$, $P = XY$ e $Q = X/(X + Y)$ quando X e Y seguem o modelo (2.1). Inicialmente apresentamos a fdp conjunta das variáveis U e Q (ver equação 2.5) e a fdp conjunta das variáveis X e P (ver equação 2.7), que são úteis nas demonstrações desses teoremas.

Considere $(U, Q) = (X + Y, X/(X + Y))$, onde X e Y são variáveis aleatórias com função densidade de probabilidade conjunta dada pela equação (2.1). Então, o Jacobiano da transformação é dado por:

$$J = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial q}{\partial x} & \frac{\partial q}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ \frac{y}{(x+y)^2} & -\frac{x}{(x+y)^2} \end{vmatrix} = -\frac{1}{x+y}. \quad (2.4)$$

Assim, a fdp conjunta de $(U, Q) = (X + Y, X/(X + Y))$ é dada por

$$\begin{aligned} f_{U,Q}(u, q) &= f_{X,Y}(uq, u(1-q)) \times |J|^{-1} \\ &= Ku^{\alpha+\beta-1} q^{\alpha-1} (1-q)^{\beta-1} \exp\{-auq - cu^2q(1-q)\} \\ &= Ku^{\alpha+\beta-1} q^{\alpha-1} (1-q)^{\beta-1} \exp\{-(a+cu)uq + cu^2q^2\} \\ &= Ku^{\alpha+\beta-1} q^{\alpha-1} (1-q)^{\beta-1} \exp\left\{cu^2(q-\theta)^2 - \frac{(a+cu)^2}{4c}\right\} \end{aligned} \quad (2.5)$$

Considere agora $(X, P) = (X, XY)$, onde X e Y são variáveis aleatórias com função densidade de probabilidade conjunta dada pela equação (2.1). Então, o Jacobiano da transformação é dado por:

$$J = \begin{vmatrix} \frac{\partial x}{\partial x} & \frac{\partial x}{\partial y} \\ \frac{\partial p}{\partial x} & \frac{\partial p}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ y & x \end{vmatrix} = x. \quad (2.6)$$

Assim, a fdp conjunta de (X, P) é dada por

$$\begin{aligned} f_{X,P}(x,p) &= f_{X,Y}\left(x, \frac{p}{x}\right) \times |J|^{-1} \\ &= Kx^{\alpha-\beta-1}p^{\beta-1} \exp[-(ax+cp)]. \end{aligned} \quad (2.7)$$

Teorema 2.1. *Sejam X e Y variáveis aleatórias com função densidade de probabilidade conjunta dada pela equação (2.1) e U variável aleatória dada por $U = X+Y$. Se $\alpha \geq 2$ e $\beta \geq 2$ são inteiros, então:*

$$\begin{aligned} f_U(u) &= Ku^{\alpha+\beta-1} \exp\left\{-\frac{(a+cu)^2}{4c}\right\} \sum_{m=0}^{\alpha-1} \sum_{n=0}^{\beta-1} \sum_{j=0}^n \binom{n}{j} \binom{\beta-1}{n} \binom{\alpha-1}{m} \times \\ &\quad \frac{\theta^{\alpha+n-(m+j+1)}}{m+j+1} \left[\theta^{m+j-1} {}_1F_1\left(\frac{m+j+1}{2}; \frac{m+j+3}{2}; cu^2\theta^2\right) + \right. \\ &\quad \left. (1-\theta)^{m+j-1} {}_1F_1\left(\frac{m+j+1}{2}; \frac{m+j+3}{2}; cu^2(1-\theta)^2\right) \right], \end{aligned}$$

para $0 \leq \theta \leq 1$ e

$$\begin{aligned} f_U(u) &= Ku^{\alpha+\beta-1} \exp\left\{-\frac{(a+cu)^2}{4c}\right\} \sum_{m=0}^{\alpha-1} \sum_{n=0}^{\beta-1} \sum_{j=0}^n \binom{n}{j} \binom{\beta-1}{n} \binom{\alpha-1}{m} \times \\ &\quad \frac{\theta^{\alpha+n-(m+j+1)}}{m+j+1} \left[(1-\theta)^{m+j-1} {}_1F_1\left(\frac{m+j+1}{2}; \frac{m+j+3}{2}; cu^2(1-\theta)^2\right) - \right. \\ &\quad \left. \theta^{m+j-1} {}_1F_1\left(\frac{m+j+1}{2}; \frac{m+j+3}{2}; cu^2\theta^2\right) \right], \end{aligned}$$

para $\theta > 1$, em que $u > 0$ e $\theta = (a-cu)/(2cu)$.

Demonstração. Considere U e P variáveis aleatórias com função densidade de probabilidade conjunta dada pela equação (2.5). Para $0 \leq \theta \leq 1$ então a fdp de U é dada por:

$$\begin{aligned} f_U(u) &= Ku^{\alpha+\beta-1} \int_0^1 q^{\alpha-1} (1-q)^{\beta-1} \exp\left\{cu^2(q-\theta)^2 - \frac{(a+cu)^2}{4c}\right\} dq \\ &= Ku^{\alpha+\beta-1} \exp\left\{-\frac{(a+cu)^2}{4c}\right\} \int_0^1 q^{\alpha-1} (1-q)^{\beta-1} \exp\left\{cu^2(q-\theta)^2\right\} dq. \end{aligned}$$

Fazendo $y = (w - \theta)^2$ temos:

$$\begin{aligned} f_U(u) &= 2^{-1} K u^{\alpha+\beta-1} \exp \left\{ -\frac{(a+cu)^2}{4c} \right\} \left(\int_0^{\theta^2} + \int_0^{(1-\theta)^2} \right) (\sqrt{y} + \theta)^{\alpha-1} \times \\ &\quad (1 - \sqrt{y} - \theta)^{\beta-1} \exp \{ cu^2 y \} y^{-1/2} dy \\ &= 2^{-1} K u^{\alpha+\beta-1} \exp \left\{ -\frac{(a+cu)^2}{4c} \right\} \left(\int_0^{\theta^2} + \int_0^{(1-\theta)^2} \right) \sum_{m=0}^{\alpha-1} \binom{\alpha-1}{m} (\sqrt{y})^m \theta^{\alpha-m-1} \times \\ &\quad \sum_{n=0}^{\beta-1} \binom{\beta-1}{n} (\sqrt{y} + \theta)^n \exp \{ cu^2 y \} y^{-1/2} dy \\ &= 2^{-1} K u^{\alpha+\beta-1} \exp \left\{ -\frac{(a+cu)^2}{4c} \right\} \sum_{m=0}^{\alpha-1} \sum_{n=0}^{\beta-1} \sum_{j=0}^n \binom{n}{j} \binom{\beta-1}{n} \binom{\alpha-1}{m} \times \\ &\quad \frac{\theta^{\alpha+n-(m+j+1)}}{m+j+1} \left(\int_0^{\theta^2} + \int_0^{(1-\theta)^2} \right) y^{\frac{m+j-1}{2}} \exp \{ cu^2 y \} dy. \end{aligned}$$

Aplicando o Lema (1.1) à integral acima, obtém-se:

$$\begin{aligned} f_U(u) &= K u^{\alpha+\beta-1} \exp \left\{ -\frac{(a+cu)^2}{4c} \right\} \sum_{m=0}^{\alpha-1} \sum_{n=0}^{\beta-1} \sum_{j=0}^n \binom{n}{j} \binom{\beta-1}{n} \binom{\alpha-1}{m} \times \\ &\quad \frac{\theta^{\alpha+n-(m+j+1)}}{m+j+1} \left[\theta^{m+j-1} {}_1F_1 \left(\frac{m+j+1}{2}; \frac{m+j+3}{2}; cu^2 \theta^{m+j-1} \right) + \right. \\ &\quad \left. (1-\theta)^{m+j-1} {}_1F_1 \left(\frac{m+j+1}{2}; \frac{m+j+3}{2}; cu^2 (1-\theta)^{m+j-1} \right) \right]. \end{aligned}$$

O resultado para $\theta > 1$ é obtido similarmente. \square

Teorema 2.2. *Sejam X e Y variáveis aleatórias com função densidade de probabilidade conjunta dada pela equação (2.1) e Q a variável aleatória dada por $Q = X/(X+Y)$. Se $\alpha \geq 2$ e $\beta \geq 2$ são inteiros, então:*

$$\begin{aligned} f_Q(q) &= K q^{\alpha-1} (1-q)^{\beta-1} \Gamma(\alpha+\beta) [2cq(1-q)]^{-\frac{\alpha+\beta}{2}} \exp \left[\frac{(aq)^2}{8cq(1-q)} \right] \times \\ &\quad D_{-(\alpha+\beta)} \left(\frac{aq}{\sqrt{2cq(1-q)}} \right), \end{aligned}$$

em que $0 < q < 1$.

Demonstração. De (2.5) tem-se que a distribuição de Q é dada por:

$$f_Q(q) = \int_0^{\infty} f(u, q) du$$

$$= Kq^{\alpha-1}(1-q)^{\beta-1} \int_0^{\infty} u^{\alpha+\beta-1} \exp\{-[auq + cq(1-q)u^2]\} du.$$

O resultado segue da aplicação do Lema (1.2) à integral. \square

Teorema 2.3. *Sejam X e Y variáveis aleatórias com função densidade de probabilidade conjunta dada pela equação (2.1) e P a variável aleatória dada por $P = XY$. Se $\alpha \geq 2$ e $\beta \geq 2$ são inteiros, então:*

$$f_P(p) = \frac{c^\beta}{\Gamma(\beta)} p^{\beta-1} \exp(-cp),$$

em que $p > 0$.

Demonstração. Considere X e P variáveis aleatórias com função densidade de probabilidade conjunta dada pela equação (2.7). A fdp de P é dada por

$$\begin{aligned} f_P(p) &= \int_0^{\infty} f(x, p) dx \\ &= \int_0^{\infty} Kx^{\alpha-\beta-1} p^{\beta-1} \exp[-(ax + cp)] dx \\ &= Kp^{\beta-1} \exp(-cp) \int_0^{\infty} x^{\alpha-\beta-1} \exp(-ax) dx \\ &= \frac{c^\beta}{\Gamma(\beta)} p^{\beta-1} \exp(-cp). \end{aligned}$$

\square

2.2. Momentos

Nesta seção são deduzidos os momentos das variáveis $U = X + Y$ e $P = XY$ quando X e Y seguem o modelo gama beta II (ver equação 2.1). Inicialmente apresentamos o Lema 2.1 o qual é muito útil nas demonstrações dos Teoremas 2.4 e 2.5.

Lema 2.1. *Sejam X e Y variáveis aleatórias com função densidade de probabilidade conjunta dada pela equação (2.1). Então:*

$$E[X^n Y^m] = \frac{a^{m-n} \Gamma(\beta + m) \Gamma(\alpha + n - \beta - m)}{c^m \Gamma(\beta) \Gamma(\alpha - \beta)},$$

para os inteiros $n \geq 1$ e $1 \leq m \leq n$.

Demonstração. Sob o modelo (2.1), tem-se que:

$$\begin{aligned}
 E[X^n Y^m] &= \int_0^\infty \int_0^\infty x^n y^m f(x, y) dy dx \\
 &= \int_0^\infty \int_0^\infty K x^{\alpha+n-1} y^{\beta+m-1} \exp[-(ax + cxy)] dy dx \\
 &= K \int_0^\infty x^{\alpha+n-1} \exp(-ax) \left[\int_0^\infty y^{\beta+m-1} \exp(-cxy) dy \right] dx \\
 &= K \int_0^\infty x^{\alpha+n-1} \exp(-ax) \frac{\Gamma(\beta+m)}{(cx)^{\beta+m}} dx \\
 &= \frac{a^{m-n} \Gamma(\beta+m) \Gamma(\alpha+n-\beta-m)}{c^n \Gamma(\beta) \Gamma(\alpha-\beta)}, \quad \text{com } m \leq n.
 \end{aligned}$$

□

Em particular, para $n = m = 1$ obtém-se:

$$E[XY] = \frac{\Gamma(\beta+1) \Gamma(\alpha-\beta)}{c \Gamma(\beta) \Gamma(\alpha-\beta)} = \frac{\beta}{c}.$$

Teorema 2.4. *Sejam X e Y variáveis aleatórias com função densidade de probabilidade conjunta dada pela equação (2.1). Então:*

$$E[U^n] = \sum_{j=0}^n \binom{n}{j} \frac{a^n \Gamma(\beta+n-j) \Gamma(\alpha+2j-\beta)}{c^{n-j} \Gamma(\beta) \Gamma(\alpha-\beta)},$$

para todo inteiro $n \geq 1$.

Demonstração. Temos que:

$$E[U^n] = E[(X+Y)^n] = \sum_{j=0}^n \binom{n}{j} E[X^j Y^{n-j}]. \quad (2.8)$$

O resultado segue da aplicação do Lema (2.1). □

Teorema 2.5. *Sejam X e Y variáveis aleatórias com função densidade de probabilidade conjunta dada pela equação (2.1). Então:*

$$E[P^n] = \frac{\Gamma(\beta+n)}{c^n \Gamma(\beta)},$$

para todo inteiro $n \geq 1$.

Demonstração. Temos que: $E[P^n] = E[(XY)^n] = E[X^n Y^n]$. O resultado segue da aplicação do Lema (2.1) para $m = n$. □

2.3. Estimação

Seja $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ uma amostra aleatória da variável aleatória (X, Y) com distribuição conjunta expressa por (2.1), de parâmetros a, c, α e β . A função de verossimilhança é:

$$\begin{aligned} L &= \prod_{i=1}^n f(x_i, y_i) = \prod_{i=1}^n K x_i^{\alpha-1} y_i^{\beta-1} \exp[-(ax_i + cx_i y_i)] \\ &= K^n \prod_{i=1}^n (x_i^{\alpha-1} y_i^{\beta-1}) \exp \left[- \left(a \sum_{i=1}^n x_i + c \sum_{i=1}^n x_i y_i \right) \right]. \end{aligned} \quad (2.9)$$

Aplicando a função logarítmica a ambos os lados da equação (2.9) obtém-se:

$$\ln(L) = n \ln(K) + (\alpha - 1) \sum_{i=1}^n \ln(x_i) + (\beta - 1) \sum_{i=1}^n \ln(y_i) - a \sum_{i=1}^n x_i - c \sum_{i=1}^n x_i y_i. \quad (2.10)$$

Assim,

$$\begin{aligned} \frac{\partial \ln(L)}{\partial a} &= \frac{n(\alpha - \beta)}{a} - \sum_{i=1}^n x_i, \\ \frac{\partial \ln(L)}{\partial c} &= \frac{n\beta}{c} - \sum_{i=1}^n x_i y_i, \\ \frac{\partial \ln(L)}{\partial \alpha} &= n \ln(a) - n\psi(\alpha - \beta) + \sum_{i=1}^n \ln(x_i), \\ \frac{\partial \ln(L)}{\partial \beta} &= n \ln(c) - n \ln(a) - n\psi(\beta) + n\psi(\alpha - \beta) + \sum_{i=1}^n \ln(y_i), \end{aligned}$$

em que $\psi(\cdot)$ denota a função psi dada pela equação (1.5).

Os estimadores de máxima verossimilhança dos parâmetros a, c, α e β , são os valores na qual a função de verossimilhança é máxima, ou seja, é a solução do sistema de equações:

$$\begin{cases} \frac{n(\hat{\alpha} - \hat{\beta})}{\hat{a}} - \sum_{i=1}^n x_i = 0 \\ \frac{n\hat{\beta}}{\hat{c}} - \sum_{i=1}^n x_i y_i = 0 \\ n \ln(\hat{a}) - n\psi(\hat{\alpha} - \hat{\beta}) + \sum_{i=1}^n \ln(x_i) = 0 \\ n \ln(\hat{c}) - n \ln(\hat{a}) - n\psi(\hat{\beta}) + n\psi(\hat{\alpha} - \hat{\beta}) + \sum_{i=1}^n \ln(y_i) = 0. \end{cases}$$

A solução do sistema de equações de máxima verossimilhança pode ser obtida numericamente através do software de computação algébrica MAPLE (versão 16.02).

Como não é possível obter as propriedades exatas do estimador de máxima verossimilhança, um estudo computacional é apresentado. Para os dados utilizados na seção 3 seguinte, as estimativas de máxima verossimilhança para os parâmetros do modelo são:

$$\left(\hat{a}, \hat{c}, \hat{\alpha}, \hat{\beta}\right) = (0.8654, 0.1899, 3.3920, 1.2842)$$

Fixando-se três dessas estimativas, é possível simular o gráfico do erro quadrático médio (EQM) do estimador de máxima verossimilhança da seguinte forma: Suponha fixados os valores de $\hat{c} = 0.1899$, $\hat{\alpha} = 3.3920$ e $\hat{\beta} = 1.2842$, e permita que a varie no intervalo $[0, 64905, 1.752436]$. Tomam-se valores a_i ($i = 1, 2, \dots, I$) nesse intervalo e, para cada um desses valores, considera-se que o vetor dos parâmetros reais é $\left(a_i, \hat{c}, \hat{\alpha}, \hat{\beta}\right) = (a_i, 0.1899, 3.3920, 1.2842)$. Substituindo na densidade (2.1), uma amostra de tamanho 101 é obtida. A partir desta, a solução do sistema das equações de máxima verossimilhança é obtido e o erro quadrático determinado. Esse procedimento, repetido 1000 vezes, permite a obtenção de uma estimativa para o erro quadrático médio. Com isto, é possível traçar o gráfico do erro quadrático médio como função do parâmetro a . Repetindo-se o processo para os outros parâmetros, obtêm-se gráficos representados pelas Figuras 1, 2, 3 e 4.

Supondo os parâmetros reais dados por $(a, c, \alpha, \beta) = (0.8654, 0.1899, 3.3920, 1.2842)$ foram simuladas 100.000 amostras de tamanho 101 e, para cada amostra, calculadas as estimativas de máxima verossimilhança dos parâmetros, $\left(\hat{a}, \hat{c}, \hat{\alpha}, \hat{\beta}\right)$. Com isso foi possível obter:

1. As esperanças de cada componente dos estimadores:

$$\left(E[\hat{a}], E[\hat{c}], E[\hat{\alpha}], E[\hat{\beta}]\right) = (1.52940, 0.45604, 4.23305, 0.74029) \quad (2.11)$$

2. O erro quadrático médio:

$$\begin{aligned} EQM &= \frac{1}{100000} \sum \left[(a - \hat{a})^2 + (b - \hat{b})^2 + (\alpha - \hat{\alpha})^2 + (\beta - \hat{\beta})^2 \right] \\ &= 0.49260 + 0.07933 + 1.02017 + 0.30527 = 1.89736 \end{aligned} \quad (2.12)$$

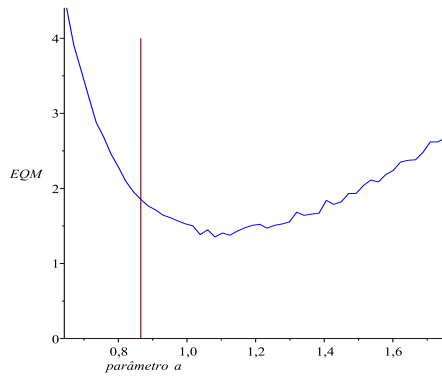
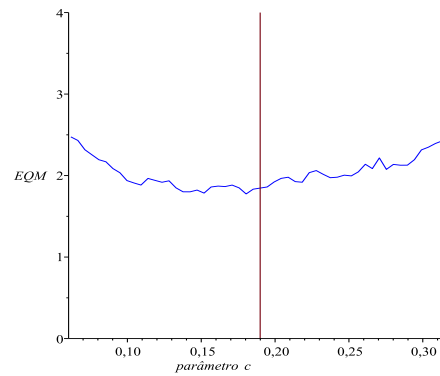
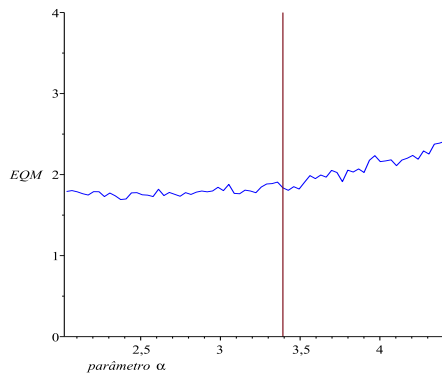
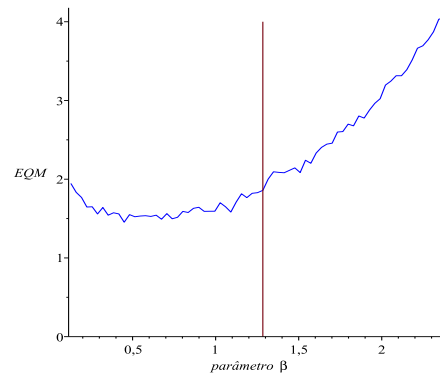
3. A Variância das estimativas:

$$\begin{aligned} Var &= \frac{1}{100000} \sum \left[(\hat{a} - E[\hat{a}])^2 + (\hat{b} - E[\hat{b}])^2 + (\hat{\alpha} - E[\hat{\alpha}])^2 + (\hat{\beta} - E[\hat{\beta}])^2 \right] \\ &= 0.05170 + 0.07933 + 1.02017 + 0.03053 = 0.38242 \end{aligned} \quad (2.13)$$

4. O Viés:

$$\begin{aligned} Vies &= (a - E[\hat{a}])^2 + (b - E[\hat{b}])^2 + (\alpha - E[\hat{\alpha}])^2 + (\beta - E[\hat{\beta}])^2 \\ &= 0.44089 + 0.07083 + 0.70741 + 0.29581 = 1.51494 \end{aligned} \quad (2.14)$$

Observe que, como tem que ser, $EQM = Var + Vies$.

Figura 1: $(a, 0.1899, 3.3920, 1.2842)$ Figura 2: $(0.8654, c, 3.3920, 1.2842)$ Figura 3: $(0.8654, 0.1899, \alpha, 1.2842)$ Figura 4: $(0.8654, 0.1899, 3.3920, \beta)$

3. Aplicação

Nesta seção é feita uma aplicação do modelo na análise de dados de precipitações pluviométricas ocorridas na cidade de Passo Fundo, Estado do Rio Grande do Sul.

Os dados explorados correspondem a medições diárias de precipitação pluviométrica (mm) no período de Julho de 2009 a Julho de 2011, totalizando 730 observações. Os dados foram coletados pelo Laboratório de Meteorologia Aplicada à Agricultura da Embrapa Trigo, Passo Fundo - RS, (latitude: $28^{\circ}15'46''$ S; longitude: $52^{\circ}24'24''$ W; altitude: 684m) e encontram-se disponíveis livremente para download no endereço eletrônico <http://www.cnpt.embrapa.br/pesquisa/agromet>.

Utilizando as medições do índice pluviométrico, obtém-se os dados sobre período de dias com ocorrência de precipitações (X) e período contíguo de dias sem ocorrência de precipitação (Y). O objetivo é modelar as variáveis X , Y e o período climático U . O período climático indica um ciclo climático, formado pela soma dos dias sem chuva com os dias contíguos com chuva, essa quantidade também indica o retorno do período de chuva.

O ajuste das distribuições foi feito via Método da Máxima Verossimilhança. Se $(x_1, y_1), \dots, (x_n, y_n)$ é uma amostra aleatória de (2.1), as estimativas dos parâmetros do modelo são: $\hat{\alpha} = 0.8654$, $\hat{c} = 0.1899$, $\hat{\alpha} = 3.3920$ e $\hat{\beta} = 1.2842$.

As fdps ajustadas de X , Y e U e os respectivos gráficos de probabilidades observadas versus probabilidades esperadas, são apresentadas nas Figuras de 5 a 10.

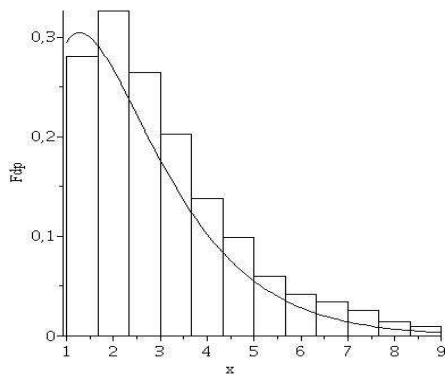


Figura 5: fdp ajustada de X

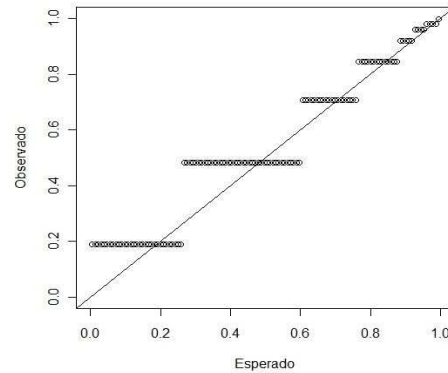


Figura 6: p-pplot para a variável X

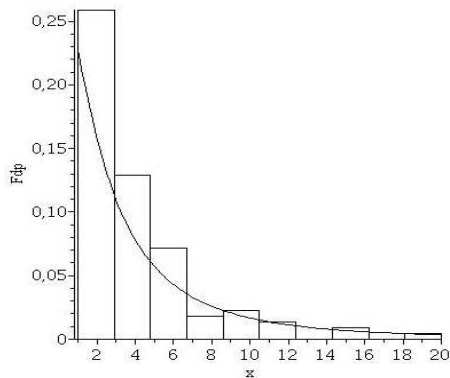


Figura 7: fdp ajustada de Y

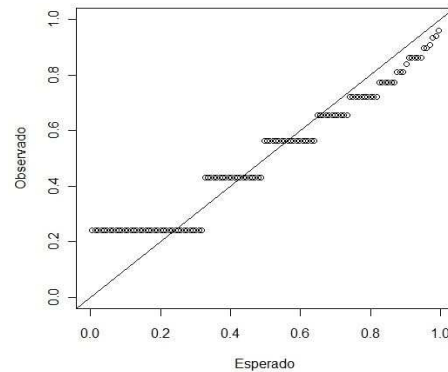
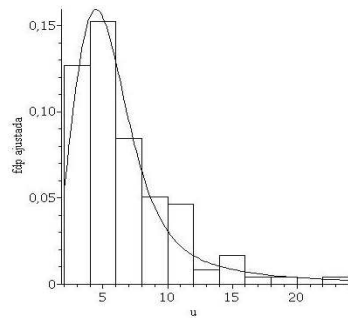
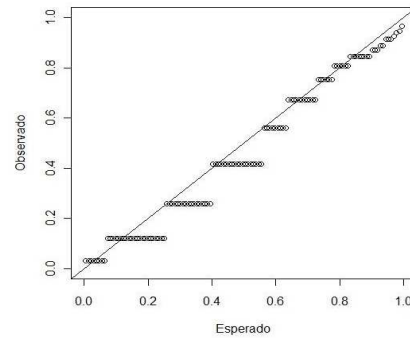


Figura 8: p-pplot para a variável Y

Os gráficos das distribuições ajustadas e de probabilidade, Figuras 5 a 10, sugerem um bom ajuste para as variáveis período de chuva (X), período sem ocorrência de chuva (Y) e período climático (U), pelo gama beta tipo II. Deste modo temos que o modelo bivariado gama beta II apresenta-se como uma alternativa na análise de dados de precipitação pluviométrica.

Figura 9: fdp ajustada de Y Figura 10: p-pplot para a variável Y

4. Conclusões

Considerando que X e Y seguem o modelo bivariado gama beta II, foi possível deduzir as distribuições exatas e os momentos das variáveis aleatórias $U = X + Y$, $P = XY$ e $Q = X/(X + Y)$ utilizando funções especiais. A estimação dos parâmetros do modelo pelo Método da Máxima Verossimilhança e suas propriedades, podem computadas utilizando métodos numéricos.

A aplicação do modelo na análise de dados de precipitações ocorridas na cidade de Passo Fundo do Estado do Rio Grande do Sul apresentou resultados satisfatórios, levando-se em consideração que os critérios gráficos de qualidade de ajuste indicaram uma boa adequação do modelo aos dados observados.

Abstract. Bivariate gamma distributions have been used successfully on modeling hydrological processes. In this work, supposing that X and Y follow the gamma beta II bivariate gamma model, we deduce the exact distributions of the functions $U = X + Y$, $P = XY$ and $Q = X/(X + Y)$, as well as their respective moments. Those functions describe important hidrological variables. An application of the results is provided to rainfall data from Passo Fundo - RS.

Keywords. Gamma beta II distribution, combination of random variables, rainfall.

Referências

- [1] T. Izawa, Two or multi-dimensional gamma-type distribution and its application to rainfall data, *Meteorological Research Institute*, **15** (1965), 167–200.
- [2] N. Balakrishnan, C.D. Lai, "Continuous bivariate distributions", Springer Verlag, 2009.
- [3] R. Beals, R. Wong, "Special functions", Cambridge University, New York, 2010.

-
- [4] A.K. Gupta, S. Nadarajah, Sums, products and ratios for Mckay's bivariate gamma distribution, *Mathematical and Computer Modelling*, **43** (2006), 185–193.
- [5] M.V. Jambunathan, Some properties of beta and gamma distributions, *The Annals of Math. Stat.*, **25** (1954), 401–405.
- [6] S. Nadarajah, Products and ratios for a bivariate gamma distribution, *Journal of the American Statistical Association Applied Mathematics and Computing*, **24**, No. 1 (2005), 581–595.
- [7] S. Nadarajah, Bivariate gamma exponential distribution with application to drought data, *Journal Applied Mathematical and Computer Modelling*, **24**, No. 1 (2007), 221–130.
- [8] S. Nadarajah, Bivariate pareto model for drought, *Stochastic Environmental Research and Risk Assessment*, **23**, No. 6 (2008), 811–822.
- [9] S. Nadarajah, A.K. Gupta, Cherian's bivariate gamma distributioins as a model for drought data, *Agrociencia*, **40**, No. 4 (2006), 483–490.
- [10] S. Nadarajah, A.K. Gupta, Friday and Patil's bivariate exponential distributiion with application to drought data, *Water Resource Management*, **20**, No. 5 (2006), 749–759.
- [11] S. Nadarajah, S. Kotz, Sums, products and ratios for downton's bivariate exponential distribution, *Stochastic Environmental Research and Risk Assessment*, **20**, No. 3 (2006), 164–170.
- [12] K. Oldham, J. Myland, J. Spanier, "An atlas of functions: with equator, the atlas function calculator", Springer Vergland, Canadá, 2009.
- [13] A.P. Prudnikov, O.I. Marichev, "Integrals and series", Gordon and Breach, Amsterdam, 1998.
- [14] A.P.C.M. Silva, J.A. Rodrigues, L.M. Chaves, D.J. Souza, Sums, Products and Ratios for Crovelli's Bivariate Gamma Distribution, *Water Resources Management*, **27**, No. 5 (2013), 1363–1376.
- [15] S. Yue, A bivariate gamma distribution for use in multivariate flood frequency analysis, *Journal of Hydrology*, **15**, No. 6 (2001), 1033–1045.
- [16] S. Yue, T.B.M.J. Ouarda, B. Bobee, A review of bivariate gamma distributions for hydrological application, *Journal of Hydrology*, **246**, No.1 (2001), 1–18.