

DISCUSSÃO E APLICAÇÃO DO MODELO DE REGRESSÃO BETA EM DADOS SOBRE A ANSIEDADE

Juliana Faria de Carvalho¹, Isolde Previdelli², Rosangela Santana²

Resumo: *Em estatística é de suma importância conhecer uma distribuição de probabilidade para descrever a chance de ocorrência da variável resposta. O interesse na distribuição beta deve-se à sua versatilidade para modelar experimentos aleatórios que produzem resultados no intervalo (0,1) devido à grande flexibilidade de ajuste dos seus parâmetros. Dando continuidade a um trabalho que começou como um projeto de iniciação científica: Momentos e Cumulantes da Distribuição Beta da Família Biexponencial, é apresentada aqui uma discussão sobre a função densidade, a distribuição acumulada, a função geratriz de momentos (f.g.m), a função escore, os cumulantes até ordem quatro, a Matriz de Informação de Fisher, a relação dessa distribuição com algumas outras e ainda, como a distribuição beta foi reparametrizada por Ferrari e Cribari Neto (2004), com o objetivo de desenvolver um modelo de regressão beta. Este é bastante útil para modelar taxas e proporções, restritos ao intervalo (0,1). Uma aplicação em dados sobre ansiedade é apresentada. O aplicativo Maple foi utilizado para a realização dos cálculos matemáticos e o software estatístico R para análises e gráficos.*

Palavras-chave: *Momentos, Regressão beta, Aplicação.*

Abstract: *The beta distribution has great flexibility to adjust its parameters to results in the interval (0,1). We discuss this distribution and its reparametrization by Ferrari and Cribari Neto (2004), to develop the beta regression model. An application for data on anxiety is presented.*

Keywords: *Beta Distribution, Beta Regression Model, Application.*

1 Introdução

Em estatística é de suma importância conhecer uma distribuição de probabilidade para descrever a chance de ocorrência da variável de interesse. Este trabalho é resultado do projeto de iniciação científica intitulado *Momentos e Cumulantes da Distribuição Beta da Família Biexponencial* e consiste num estudo da Distribuição Beta, bastante utilizada para modelar experimentos aleatórios que produzem resultados no intervalo (0,1), devido à grande flexibilidade de ajuste dos seus parâmetros. Ela está entre as mais frequentemente empregadas para modelar distribuições teóricas, pois muitas das distribuições finitas encontradas na prática podem ser facilmente transformadas na distribuição beta padronizada, *apud* (Brito, 2009).

A busca pela origem da distribuição beta remete a 1676, a uma carta de Isaac Newton para Henry Oldenberg, na qual foi encontrada a fórmula que mais tarde foi usada para calcular a

¹DES - UEM. Agradecimento ao CNPq pelo apoio financeiro.

²DES-UEM. e-mail: *isoldeprevidelli@gmail.com, rgsantana@uem.br.*

distribuição acumulada da beta. Um relato detalhado sobre a distribuição acumulada é apresentado por Dutka(1981), *apud* (Johnson, Kotz, Balakrishnan, 1995). Como esta é uma distribuição que tem relação com várias outras distribuições, a sua história acaba sendo construída em partes, ao longo dos anos na medida em que os pesquisadores foram variando os valores dos seus parâmetros e relacionando-a com várias outras distribuições, dentre as quais destacam-se: a distribuição arco-seno e arco-seno generalizada, distribuição uniforme (quando $p = q = 1$), Pareto e distribuição gama.

A distribuição beta foi reparametrizada por Ferrari e Cribari Neto (2004), com o objetivo de criar um novo modelo de Regressão para situações nas quais as variáveis de interesse são contínuas e restritas ao intervalo (0,1). Eles fornecem expressões de forma fechada para a Função Escore, para a Matriz de Informação de Fisher e sua inversa, e ainda realizam Testes de Hipóteses para os estimadores de máxima verossimilhança usando aproximações obtidas a partir da normalidade assintótica. A partir de então, a regressão beta passa a ser bastante utilizada para modelar taxas e proporções.

Na 13^ª Escola de Modelos de Regressão, Ferrari (2013) apresenta muitas aplicações recentes, dentre as quais, na área da saúde, destacam-se: a proporção de incidência de raios UVB em pessoas com e sem óculos (2006); a proporção de área de necrose miocárdica em pacientes com infarto agudo do miocárdio (2011); a qualidade de vida em pacientes com AVC medidos pelo *Stroke Impact Scale* SIS (2012); o percentual de densidade mamográfica em que densidade alta é um marcador de câncer de mama (2012); a pontuação do comprometimento cognitivo em pacientes com Alzheimer (2012); a diferença genética entre duas estirpes do vírus febre aftosa medido como a porcentagem de nucleótidos que diferem entre si por uma definida porção do genoma (2007); a fração de matéria orgânica em do total de matéria em suspensão em uma zona de amostragem de um rio (2009); dentre várias outras.

Uma aplicação foi utilizada nos dados de um estudo clínico sobre transtorno de ansiedade generalizada, conduzido pelo professor Dipankar Bandyopadhyay, do departamento de saúde pública, divisão de bioestatística, da Universidade de Minnesota. O estudo foi aleatorizado e controlado por placebo.

2 Material e métodos

Por se tratar de um trabalho teórico, foi feito um levantamento de referências sobre a distribuição beta e sua relação com outras distribuições, bem como suas aplicações em diversas áreas do conhecimento. Há várias referências relevantes, porém, as que mais se destacam são os livros *Continuous Univariate Distributions*, Johnson, Kotz e Balakrishnan (1995), e *Probability and Statistics*, De Groot (1989), e o artigo *Beta Regression for modelling rates and proportions*, Ferrari e Cribari Neto (2004).

Para os cálculos matemáticos foi utilizado o aplicativo *Maple* e para as análises estatísticas o aplicativo *R*, que é um *software* estatístico livre.

3 Resultados e discussões

A família de distribuições beta, representada aqui pela variável aleatória x e pelos parâmetros de forma e escala p e q , segundo Degroot(1989), é composta de todas as distribuições com função densidade de probabilidade da forma

$$f(x; p, q) = \frac{1}{B(p, q)} x^{p-1} (1-x)^{q-1}, \quad 0 \leq x \leq 1, \quad p, q > 0. \quad (1)$$

Para achar o valor de B parte-se da definição da função gama

$$\Gamma(z) = \int_0^{\infty} u^{z-1} e^{-u} du. \quad (2)$$

Assim, tem-se

$$\Gamma(p)\Gamma(q) = \int_0^\infty u^{p-1}e^{-u}du \int_0^\infty v^{q-1}e^{-v}dv = \int_0^\infty \int_0^\infty u^{p-1}v^{q-1}e^{-(u+v)}dudv.$$

Fazendo $x = \frac{u}{u+v}$ e $y = u + v$ então $u = xy$ e $v = y(1 - x)$, sendo que $u, v, y < 0$ e $0 < x < 1$. Calculando o Jacobiano obtêm-se $|J| = y$.

Substituindo na equação (2)

$$\Gamma(p)\Gamma(q) = \int_0^1 \int_0^\infty (xy)^{p-1}[y(1-x)]^{q-1}e^{-(xy+y-xy)}ydydx$$

$$\Gamma(p)\Gamma(q) = \int_0^1 \int_0^\infty x^{p-1}(1-x)^{q-1}y^{p+q-1}e^{-y}dydx$$

$$\Gamma(p)\Gamma(q) = \int_0^1 x^{p-1}(1-x)^{q-1}dx \underbrace{\int_0^\infty y^{p+q-1}e^{-y}dy}_{\text{Gama}(1,p+q)}$$

$$\Gamma(p)\Gamma(q) = \int_0^1 x^{p-1}(1-x)^{q-1}dx\Gamma(p+q)$$

$$\int_0^1 x^{p-1}(1-x)^{q-1}dx = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

Portanto

$$B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}. \quad (3)$$

Então, a função densidade dada pela equação (1) pode ser reescrita da seguinte forma

$$f(x; p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}x^{p-1}(1-x)^{q-1}, \quad 0 \leq x \leq 1 \quad p, q > 0. \quad (4)$$

De acordo com Degroot (1989), a densidade beta pode exibir formas diferentes dependendo dos valores assumidos pelos parâmetros. Como exemplo tem-se as distribuições uniforme e arcosseno que são simétricas em relação as suas respectivas médias, o que ilustra a propriedade de quase-simetria da distribuição beta. Nas figuras abaixo observa-se variando os parâmetros p e q obtêm-se diferentes formas:

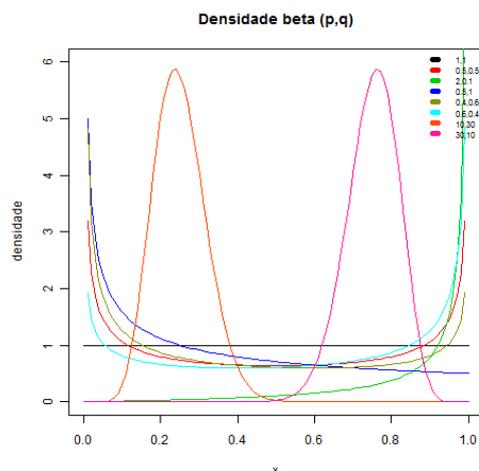


Figura 1: Densidade beta para diferentes valores para os parâmetros p e q

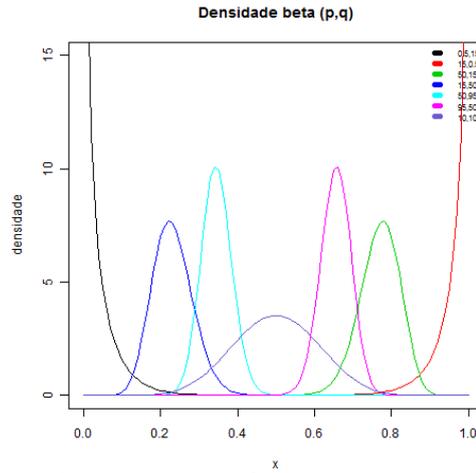


Figura 2: Densidade beta para diferentes valores para os parâmetros p e q

Uma família de densidades dependendo de um vetor de parâmetros $\theta_1, \theta_2, \dots, \theta_k$ é chamada família exponencial, \mathfrak{S} , se a função de densidade $f(x; \theta_1, \theta_2, \dots, \theta_k)$ puder ser expressa da forma

$$f(x; \theta_1, \theta_2, \dots, \theta_k) = \exp \left\{ \sum_{j=1}^k B_j(\theta_1, \theta_2, \dots, \theta_k) S_j(x) + C(\theta_1, \theta_2, \dots, \theta_k) + D(x) \right\}$$

para todo x real e $D(x), C(\theta_1, \theta_2, \dots, \theta_k), B_j(\theta_1, \theta_2, \dots, \theta_k)$ e $S_j(x)$ funções conhecidas.

Como a densidade da beta definida em (1) pode ser escrita por

$$f(x; p, q) = \exp(p-1) \ln x + (q-1) \ln(1-x) - \ln B(p, q),$$

observa-se então, que de acordo com a definição, a distribuição beta, de parâmetros p e q pertence à \mathfrak{S} .

A função distribuição acumulada da beta é definida por

$$F_X(x) = \frac{B_X(p, q)}{B(p, q)},$$

em que $B_X(p, q) = \int_0^x z^{p-1} (1-z)^{q-1} dz$ é a função beta incompleta, Johnson, Kotz e Balakrishnan (1995).

A função geratriz de momentos (f.g.m) permite gerar todos os momentos de uma função de distribuição como média, variância, assimetria e curtose. Considerando uma variável aleatória: x e um t (real) a f.g.m. é dada por: $M_x(t) = E[e^{tx}]$. Na distribuição beta para facilitar os cálculos considera-se $(-\log x)$, onde x é uma beta padrão, (Johnson, Kotz, Balakrishnan, 1995). Portanto

$$M_x(t) = E(e^{-t \log x}) = \frac{B(p-t, q)}{B(p, q)}.$$

A função geradora de cumulantes é dada por

$$K(t) = \log \left[\frac{\Gamma(p+q)}{\Gamma(p)} \right] - \log \left[\frac{\Gamma(p+q-t)}{\Gamma(p-t)} \right].$$

Muitos pesquisadores discutem sobre a estimação dos parâmetros pelo método dos momentos e pelo método de máxima verossimilhança, as vantagens e desvantagens de cada um, maiores detalhes ver Johnson, Kotz e Balakrishnan (1995). Como este último requer o uso de métodos iterativos de estimação, apresenta-se aqui, então, o cálculo dos quatro primeiros cumulantes pelo método dos momentos.

Quando a função densidade é dada pela equação (4), os momentos de ordem, (m_k) , podem ser calculados para $k=1,2,3,4\dots$

$$E(x^k) = \int_0^1 x^k f(x) dx.$$

Portanto

$$E(x^k) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \int_0^1 \underbrace{x^{p+k-1}(1-x)^{q-1}}_{\text{beta}(p+k,q)} dx.$$

Tem-se então

$$E(x^k) = \frac{\Gamma(p+q)\Gamma(p+k)\Gamma(q)}{\Gamma(p)\Gamma(q)\Gamma(p+k+q)}.$$

Simplificando

$$E(x^k) = \frac{p(p+1)\dots(p+k-1)}{(p+q)(p+q+1)\dots(p+q+k-1)}. \quad (5)$$

Os dois primeiros cumulantes são respectivamente a média e a variância

$$E(x) = \frac{p}{p+q} \quad (6)$$

$$Var(x) = E(x^2) - [E(x)]^2 = \frac{pq}{(p+q)^2(p+q+1)}. \quad (7)$$

O terceiro cumulante é a Assimetria, que é uma medida utilizada para avaliar a normalidade de uma distribuição. Quando a média coincide com a moda e a mediana, a distribuição é simétrica. A assimetria é equivalente ao momento central de ordem 3, e os momentos centrais são representado por

$$\mu_k = E(x - \bar{x})^k = \int_0^1 (x - \bar{x})^k f(x) dx.$$

Portanto

$$E(x - \bar{x})^3 = \int_0^1 (x - \bar{x})^3 x^{(p-1)}(1-x)^{(q-1)} dx$$

expandindo

$$E(x - \bar{x})^3 = \int_0^1 (x^3 - 3x^2\bar{x} + 3x\bar{x}^2 - \bar{x}^3)x^{(p-1)}(1-x)^{(q-1)} dx$$

distribuindo

$$\begin{aligned} E(x - \bar{x})^3 &= \underbrace{\int_0^1 x^3 x^{(p-1)}(1-x)^{(q-1)} dx}_{E(x^3)} - 3\bar{x} \underbrace{\int_0^1 x^2 x^{(p-1)}(1-x)^{(q-1)} dx}_{E(x^2)} \\ &\quad + 3\bar{x}^2 \underbrace{\int_0^1 x x^{(p-1)}(1-x)^{(q-1)} dx}_{E(x^1)} - \bar{x}^3 \underbrace{\int_0^1 x^0 x^{(p-1)}(1-x)^{(q-1)} dx}_{E(x^0)=1} \end{aligned}$$

simplificando

$$E(x - \bar{x})^3 = E(x^3) - 3\bar{x}E(x^2) + 3\bar{x}^2E(x) - \bar{x}^3.$$

Sabendo que $\bar{x} = E(x) = \frac{p}{p+q}$ então,

$$\begin{aligned} E(x - \bar{x})^3 &= \frac{p(p+1)(p+2)}{(p+q)(p+q+1)(p+q+2)} - 3 \left(\frac{p}{p+q} \right) \left(\frac{p(p+1)}{(p+q)(p+q+1)} \right) \\ &\quad + 3 \left(\frac{p}{p+q} \right)^2 \left(\frac{p}{p+q} \right) - \left(\frac{p}{p+q} \right)^3. \end{aligned}$$

Portanto

$$E(x - \bar{x})^3 = \frac{p^3 + 3p^2 + 2p}{(p+q)(p+q+1)(p+q+2)} - \frac{3(p^3 + p^2)}{(p+q)^2(p+q+1)} + \frac{2p^3}{(p+q)^3}. \quad (8)$$

A Curtose é uma medida do achatamento da distribuição de probabilidade e é calculada por meio do quarto cumulante. Essa medida indica se a curva da distribuição é mais aguda ou mais achatada do que uma curva normal padrão. Esta medida tem três classificações:

i)- quando a curva apresenta comportamento de uma normal padrão é denominada Mesocúrtica e tem o coeficiente de curtose $k = 0.263$;

ii)- quando ela tem o topo mais alto é denominada Leptocúrtica, isso acontece quando os valores observados se encontram agrupados em torno da moda e o coeficiente de curtose é $k < 0.263$;

iii)- e quando ela é mais achatada é denominada Platicúrtica, e isso significa que várias classes ou observações têm valores iguais, nesse caso o coeficiente de curtose é $k > 0.263$.

A representação da curtose pelo método dos momentos é dada por

$$k_4 = \mu_4 - 3\mu_2^2 = m_4 - 3 \underbrace{m_2^2}_{(E(x^2))^2} + 4\bar{x} \underbrace{\mu_3}_{Assimetria} + 12(\bar{x}^2) \underbrace{\mu_2}_{Var} - 6(\bar{x}^4), \quad (9)$$

onde μ_k representa os momentos centrados e m_k os momentos de ordem. Para obter m_4 , utiliza-se a equação (5) e obtém-se: $m_4 = \frac{p(p+1)(p+2)(p+3)}{(p+q)(p+q+1)(p+q+2)(p+q+3)}$. Os demais momentos já foram calculados anteriormente, então, fazendo as devidas substituições tem-se

$$k_4 = \left[\frac{p(p+1)(p+2)(p+3)}{(p+q)(p+q+1)(p+q+2)(p+q+3)} \right] - 3 \left[\frac{p(p+1)}{(p+q)(p+q+1)} \right]^2 + 4 \left(\frac{p}{p+q} \right) \times \\ \times \left[\frac{p^3 + 3p^2 + 2p}{(p+q)(p+q+1)(p+q+2)} - \frac{3(p^3 + p^2)}{(p+q)^2(p+q+1)} + \frac{2p^3}{(p+q)^3} \right] + \\ + 12 \left(\frac{p}{p+q} \right)^2 \left[\frac{pq}{(p+q)^2(p+q+1)} \right] - 6 \left[\left(\frac{p}{p+q} \right)^4 \right]. \quad (10)$$

A Função de Verossimilhança é a função densidade conjunta de uma amostra aleatória (x_1, x_2, \dots, x_n) sendo n o tamanho da amostra e cada x_i é proveniente de uma distribuição beta. Ela geralmente é utilizada para obter estimativas dos parâmetros, utilizando o princípio da máxima verossimilhança, e é representada por

$$L(p, q; x) = \prod_{i=1}^n (x_i; p, q).$$

Esse método consiste em maximizar $L(p, q; x)$, que no caso da distribuição beta tem-se

$$L(p, q; x) = \prod_{i=1}^n \frac{1}{B(p, q)} x_i^{p-1} (1-x_i)^{q-1},$$

expandindo o produtório obtém-se

$$L(p, q; x) = \frac{1^n}{[B(p, q)]^n} \prod_{i=1}^n x_i^{p-1} \prod_{i=1}^n (1-x_i)^{q-1}.$$

Geralmente adota-se o logaritmo dessa função para facilitar os cálculos e por isso é denominada log-verossimilhança, denotada por $l(p, q; x)$

$$l(p, q; x) = n(\log(1)) - n(\log(B(p, q))) + \log\left(\prod_{i=1}^n x_i^{p-1}\right) + \log\left(\prod_{i=1}^n (1-x_i)^{q-1}\right).$$

Desenvolvendo as propriedades logarítmicas

$$l(p, q; x) = 0 - n(\log(B(p, q))) + \sum_{i=1}^n (p-1)\log(x_i) + \sum_{i=1}^n (q-1)\log(1-x_i),$$

expandindo B

$$l(p, q; x) = -n\left(\log\left[\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}\right]\right) + (p-1)\sum_{i=1}^n \log(x_i) + (q-1)\sum_{i=1}^n \log(1-x_i),$$

obtém-se então a log-verossimilhança

$$l(p, q; x) = -n[\log(\Gamma(p)) + \log(\Gamma(q)) - \log(\Gamma(p+q))] + (p-1)\sum_{i=1}^n \log(x_i) + (q-1)\sum_{i=1}^n \log(1-x_i). \quad (11)$$

Para obter os Estimadores de Máxima Verossimilhança (EMV) deriva-se a log-verossimilhança com relação a cada um dos parâmetros (p, q) . Essas derivadas de primeira ordem são denominadas Função Escore e são representadas por: $U_p(p, q)$ e $U_q(p, q)$.

$$U_p(p, q) = \frac{\partial l(p, q; x)}{\partial p} = -n \left[\frac{\Gamma'(p)}{\Gamma(p)} \right] + n \left[\frac{\Gamma'(p+q)}{\Gamma(p+q)} \right] + \sum_{i=1}^n \log(x_i) = 0. \quad (12)$$

$$U_q(p, q) = \frac{\partial l(p, q; x)}{\partial q} = -n \left[\frac{\Gamma'(q)}{\Gamma(q)} \right] + n \left[\frac{\Gamma'(p+q)}{\Gamma(p+q)} \right] + \sum_{i=1}^n \log(1-x_i) = 0. \quad (13)$$

Não há uma forma fechada para a obtenção desses estimadores, logo é preciso obtê-los por algoritmos de aproximação numérica, ou de otimização não linear, como o de Newton Raphson. Os algoritmos de otimização requerem a especificação de valores iniciais para serem usados no processo iterativo (os chutes iniciais).

A derivada da função gama, $\Gamma'(\cdot)$, é usualmente representada por $\psi(\cdot)$ e conhecida como função digama, e a derivada de segunda ordem, $\Gamma''(\cdot)$, representada por $\psi'(\cdot)$, conhecida como função trigama.

De acordo com Fisher, qualquer função densidade da família $F((x/\theta), \forall \theta \in \Theta)$ tem toda sua **informação** presumível desde que satisfaça às seguintes condições de regularidade:

- i)- Θ é intervalo aberto da reta real, podendo mesmo coincidir com toda a reta real;
- ii)- Os conjuntos $\{x : F((x/\theta) > 0)\}$ são independentes de θ ;
- iii)- $\frac{\partial f(x/\theta)}{\partial \theta}$ existe e é finita, $\forall x$ e $\forall \theta \in \Theta$;
- iv)- $\forall \theta \in \Theta : 0 < E_{\theta} \left\{ \left(\frac{\partial \ln f(x/\theta)}{\partial \theta} \right)^2 \right\} < \infty$;

Uma vez que se verifiquem as quatro condições acima, pode-se obter a Matriz de Informação de Fisher, que é denotada por k e consiste basicamente na esperança matemática negativa das derivadas parciais, de segunda ordem, da função log-verossimilhança

$$k = \begin{bmatrix} -E \left(\frac{\partial^2 l(p, q)}{\partial p^2} \right) & -E \left(\frac{\partial^2 l(p, q)}{\partial p \partial q} \right) \\ -E \left(\frac{\partial^2 l(p, q)}{\partial q \partial p} \right) & -E \left(\frac{\partial^2 l(p, q)}{\partial q^2} \right) \end{bmatrix} = \begin{bmatrix} k_{pp} & k_{pq} \\ k_{qp} & k_{qq} \end{bmatrix}.$$

As derivadas de segunda ordem são dadas por

$$\begin{aligned} \frac{\partial^2 l}{\partial p^2} &= -n \left[-\frac{\Gamma'(p)}{\Gamma^2(p)} + \frac{\Gamma''(p)}{\Gamma(p)} + \frac{\Gamma'(p+q)}{\Gamma^2(p+q)} - \frac{\Gamma''(p+q)}{\Gamma(p+q)} \right] \\ \frac{\partial^2 l}{\partial q^2} &= -n \left[-\frac{\Gamma'(q)}{\Gamma^2(q)} + \frac{\Gamma''(q)}{\Gamma(q)} + \frac{\Gamma'(p+q)}{\Gamma^2(p+q)} - \frac{\Gamma''(p+q)}{\Gamma(p+q)} \right] \\ \frac{\partial^2 l}{\partial p \partial q} &= \frac{\partial^2 l}{\partial q \partial p} = -n \left[\frac{\Gamma'(p+q)}{\Gamma^2(p+q)} - \frac{\Gamma''(p+q)}{\Gamma(p+q)} \right]. \end{aligned}$$

O cálculo da Esperança foi realizado no aplicativo *Maple* e consiste na integral em todo o espaço em que a função densidade está definida, ou seja, $(0, 1)$ de x vezes a função: as derivadas de segunda ordem. Portanto tem-se

$$k_{pp} = \frac{1}{2}n \left[-\frac{\psi(p)}{\Gamma(p)} + \frac{\psi(1,p)\Gamma(p) + \psi(p)^2\Gamma(p)}{\Gamma(p)} + \frac{\psi(p+q)}{\Gamma(p+q)} - \frac{\psi(1,p+q)\Gamma(p+q) + \psi(p+q)^2\Gamma(p+q)}{\Gamma(p+q)} \right]$$

$$k_{qq} = \frac{1}{2}n \left[-\frac{\psi(q)}{\Gamma(q)} + \frac{\psi(1,q)\Gamma(q) + \psi(q)^2\Gamma(q)}{\Gamma(q)} + \frac{\psi(p+q)}{\Gamma(p+q)} - \frac{\psi(1,p+q)\Gamma(p+q) + \psi(p+q)^2\Gamma(p+q)}{\Gamma(p+q)} \right]$$

$$k_{pq} = k_{qp} = \frac{1}{2}n \left[\frac{\psi(p+q)}{\Gamma(p+q)} - \frac{\psi(1,p+q)\Gamma(p+q) + \psi(p+q)^2\Gamma(p+q)}{\Gamma(p+q)} \right].$$

E por fim, a Inversa da Matriz de Informação de Fisher

$$k^{-1} = k^{-1}(p, q) = \begin{pmatrix} k^{pp} & k^{pq} \\ k^{qp} & k^{qq} \end{pmatrix},$$

também foi obtida com o aplicativo *Maple*, que se encontra em anexo.

3.1 Relação da beta com outras distribuições

A distribuição beta pode ser transformada ou originada de outras distribuições, como por exemplo, fazendo $p=q=1$, obtém-se a distribuição Uniforme no intervalo $(0,1)$.

Uma possibilidade também, é que a beta surge naturalmente da "Teoria Normal". Fazendo: $V^2 = \frac{X_1^2}{X_1^2 + X_2^2}$, onde X_1^2 e X_2^2 são variáveis aleatórias independentes e X_j^2 é uma distribuição χ^2 com v_j graus de liberdade. A distribuição de V^2 é então uma distribuição beta padrão com parâmetros $p = \frac{1}{2}V_1$ $q = \frac{1}{2}V_2$.

Como V^2 e $(X_1^2 + X_2^2)$ são mutuamente independentes, esse resultado pode ser estendido para $X_1^2, X_2^2, \dots, X_k^2$, com X_j^2 sendo uma χ^2 com v_j graus de liberdade ($j= 1,2,\dots,k$). Então $V_1^2 = \frac{X_1^2}{X_1^2 + X_2^2}$, $V_2^2 = \frac{X_1^2 + X_2^2}{X_1^2 + X_2^2 + X_3^2}$ \dots $V_{k-1}^2 = \frac{X_1^2 + \dots + X_{k-1}^2}{X_1^2 + X_k^2}$ são variáveis aleatórias mutualmente independentes, cada uma com distribuição beta, os valores de p, q para V_j^2 sendo $\sum_{i=1}^j v_i$, $\frac{1}{2}v_j + 1$, respectivamente.

O pesquisador Kotlarski (1962) investigou as condições gerais do produto de variáveis aleatórias independentes com distribuição beta, *apud* Johnson, Kotz e Balakrishnan (1995). Ele afirmou que sob essas condições descritas acima, o produto de qualquer conjunto consecutivo de V_j^2 's também tem distribuição beta. Essa propriedade vale para quando v 's é qualquer número positivo (não necessariamente inteiro).

Uma relação importante é com a distribuição de Pareto. A variável aleatória X , com distribuição de Pareto de quatro parâmetros, tem função de densidade de probabilidade dada por

$$f(x, \alpha, \beta, \lambda, C) = \frac{C\alpha}{\beta} \left(\frac{x-\lambda}{\beta} \right)^{C-1} \left[1 + \left(\frac{x-\lambda}{\beta} \right)^C \right]^{-(\alpha+1)}, \quad (14)$$

onde $\lambda < x < \infty$, $\beta > 0$, $\alpha > 0$, $C > 0$; e onde λ é parâmetro de localização, β é parâmetro de escala, (α, C) são parâmetros de forma.

Os pesquisadores Abd Elfattah et al. (2007) mostram como a distribuição Pareto Generalizada de quatro parâmetros pode ser transformada na distribuição beta tipo II. Considerando os parâmetros $C = \beta = 1$ e $\lambda = 0$, a equação (14) reduz-se a uma distribuição beta tipo dois com fdp $f(x, \alpha) = \alpha(1+x)^{-(\alpha+1)}$, com $x > 0$, $\alpha > 0$.

Uma outra possibilidade é seguinte a transformação: $y = \left(\frac{x-\lambda}{\beta} \right)^C$, a equação (14) é transformada numa distribuição beta tipo dois $\beta(1, \alpha)$ com fdp $f(y, \alpha) = \alpha[1+y]^{-(\alpha+1)}$, com $y > 0$ $\alpha > 0$.

Uma outra distribuição que surge a partir dessa relação da Pareto Generalizada com a beta é a Distribuição Beta-Pareto (DBP). De acordo com Akinsete et al.(2008), muitos autores já haviam observado a versatilidade da Distribuição de Pareto Generalizada (DPG) na modelagem de diversos tipos de dados com cauda longa e isso o motivou a propor essa generalização da distribuição Pareto.

Seja $F(x)$ uma função de distribuição acumulada (fda) de uma variável aleatória x . A fda para uma classe generalizada de distribuição, para a variável aleatória x é gerada mediante a aplicação da função de distribuição acumulada inversa, para uma variável aleatória beta, para obter

$$G(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^{F(x)} t^{\alpha-1}(1-t)^{\beta-1} dt \quad 0 \leq \alpha \quad \beta \leq \infty. \quad (15)$$

Resolvendo essa integral e derivando em seguida, obtém-se a função densidade de probabilidade para $G(x)$, em função de $F(x)$

$$g(x) = \frac{1}{B(\alpha\beta)} [F(x)]^{\alpha-1} [1 - F(x)]^{\beta-1} F'(x), \quad k, \theta > 0, \quad y \geq \theta \quad (16)$$

onde

$$B(\alpha\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}.$$

Feito isso, substitui-se esse $F(x)$ que representa uma função de distribuição acumulada qualquer, pela distribuição acumulada da pareto de dois parâmetros

$$F(x) = 1 - \left(\frac{x}{\theta}\right)^{-k}.$$

O resultado dessa operação é a função densidade de probabilidade para a variável aleatória beta-Pareto, representada por

$$g(x) = \frac{k}{\theta B(\alpha\beta)} \left\{ 1 - \left(\frac{x}{\theta}\right)^{-k} \right\} \left(\frac{x}{\theta}\right)^{-k\beta-1}, \quad x \geq \theta; \quad \alpha, \beta, \theta, k > 0. \quad (17)$$

Definindo $y = \left(\frac{x}{\theta}\right)^{-k}$, na equação acima, não é difícil mostrar que $\int_{\theta}^{\infty} F(x)dx = 1$.

Se a variável aleatória x tem função densidade definida na equação(17) pode ser escrita como

$$X \sim BP(\alpha, \beta, \theta, k).$$

3.2 Reparametrização

Na estatística tem-se grande interesse em estudar a relação entre as variáveis predictoras e uma variável de interesse (resposta). Esse problema é abordado por meio de modelos de regressão, que ao serem ajustados, podem ser usados para realizar predições, calcular probabilidade e explicar o relacionamento entre as variáveis.

Para modelar a média da resposta é necessário um parâmetro de precisão. Esse parâmetro foi obtido por Ferrari e Cribari Neto (2004), com uma reparametrização da distribuição. Com isso, desenvolveu-se o modelo de Regressão Beta que desde então vem sendo bastante utilizado por ser muito eficaz para modelar taxas, proporções ou qualquer variável resposta que esteja no intervalo (0,1). E ainda, numa análise de regressão, é comum a variável resposta ser representada pela variável aleatória y , então a densidade reparametrizada é dada por

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma(1-\mu)\phi} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1}, \quad 0 < y < 1, \quad (18)$$

com $E(y) = \mu$ e fazendo $\phi = p + q$, obtém-se $Var(y) = \frac{\mu(1-\mu)}{\phi+1}$, conforme observa-se abaixo

$$E(y) = \frac{\left[\frac{\Gamma(\phi)}{\Gamma(\phi\mu)\Gamma(-(-1+\mu)\phi)} \right] \Gamma(-(-1+\mu)\phi)\Gamma(\phi\mu+1)}{\Gamma(\phi+1)}$$

simplificando e expandindo

$$E(y) = \frac{\left[\frac{\Gamma(\phi)}{\Gamma(\phi\mu)} \right] \Gamma(\phi\mu+1)}{\Gamma(\phi+1)} = \frac{\left[\frac{(\phi-1)!}{(\phi\mu-1)!} \right] (\phi\mu)(\phi\mu-1)!}{(\phi)!}$$

simplificando

$$E(y) = \frac{(\phi-1)!(\phi\mu)}{(\phi)!} = \frac{(\phi-1)!(\phi\mu)}{(\phi)(\phi-1)!} = \frac{\phi\mu}{\phi} = \mu$$

e portanto,

$$E(y) = \mu. \quad (19)$$

Para obter a variância, primeiro desenvolve-se o segundo momento ordinário dado por

$$E(y^2) = \frac{\left[\frac{\Gamma(\phi)}{\Gamma(\phi\mu)\Gamma(-(-1+\mu)\phi)} \right] \Gamma(-(-1+\mu)\phi)\Gamma(\phi\mu+2)}{\Gamma(\phi+2)},$$

que simplificando

$$E(y^2) = \frac{(\phi\mu+1)(\phi\mu)}{(\phi+1)(\phi)}.$$

Sabendo que a variância é dada por $E(y^2) - [E(y)]^2$, substitui-se então pelos valores obtidos

$$Var(y) = \frac{(\phi\mu+1)(\phi\mu)}{(\phi+1)(\phi)} - \mu^2,$$

desenvolvendo

$$Var(y) = \frac{(\phi\mu)^2 + (\phi\mu) - (\mu\phi)^2 - \mu^2\phi}{(\phi+1)(\phi)}$$

simplificando

$$Var(y) = \frac{(\phi\mu)(1-\mu)}{(\phi+1)(\phi)}$$

portanto

$$Var(y) = \frac{\mu(1-\mu)}{(\phi+1)}.$$

Para obter do terceiro cumulante basta desenvolver $E(y-\bar{y})^3 = E(y^3) - 3\bar{y}E(y^2) + 3\bar{y}^2E(y) - \bar{y}^3$. Desenvolvendo o terceiro momento ordinário obtém-se

$$E(y^3) = \frac{(\phi\mu+2)(\phi\mu+1)(\phi\mu)}{(\phi+2)(\phi+1)(\phi)}.$$

Substituindo pelos valores obtidos anteriormente

$$E(y-\bar{y})^3 = \left(\frac{(\phi\mu+2)(\phi\mu+1)(\phi\mu)}{(\phi+2)(\phi+1)(\phi)} \right) - 3\mu \left[\frac{(\phi\mu+1)(\phi\mu)}{(\phi+1)(\phi)} \right] + 3\mu^2\mu - \mu^3,$$

portanto a assimetria é dada por

$$E(y-\bar{y})^3 = \left(\frac{(\phi\mu+2)(\phi\mu+1)(\phi\mu)}{(\phi+2)(\phi+1)(\phi)} \right) - \left[\frac{3\mu(\phi\mu+1)(\phi\mu)}{(\phi+1)(\phi)} \right] + 2\mu.$$

Para obter o quarto cumulante substitui-se na equação (9) os momentos apropriados. Primeiro é preciso obter o momento ordinário de ordem quatro

$$m_4 = \frac{(\phi\mu + 3)(\phi\mu + 2)(\phi\mu + 1)(\phi\mu)}{(\phi + 3)(\phi + 2)(\phi + 1)(\phi)}.$$

Fazendo as devidas substituições tem-se a Curtose

$$k_4 = \left[\frac{(\phi\mu + 3)(\phi\mu + 2)(\phi\mu + 1)(\phi\mu)}{(\phi + 3)(\phi + 2)(\phi + 1)(\phi)} \right] - 3 \left[\frac{(\phi\mu + 1)(\phi\mu)}{(\phi + 1)(\phi)} \right]^2 + 4\mu \left[\left(\frac{(\phi\mu + 2)(\phi\mu + 1)(\phi\mu)}{(\phi + 2)(\phi + 1)(\phi)} \right) - \left[\frac{3\mu(\phi\mu + 1)(\phi\mu)}{(\phi + 1)(\phi)} \right] + 2\mu \right] + 12\mu^2 \left[\frac{\mu(1 - \mu)}{\phi + 1} \right] - 6\mu^4.$$

Para obter uma estrutura de regressão para a média da resposta, é necessário um parâmetro de precisão, representado aqui por $\frac{1}{\phi}$ que tem o seguinte significado: para μ fixado, o maior valor de ϕ implica na menor variância de y , ou seja, para μ fixado, a dispersão da distribuição diminui quando ϕ aumenta. Isso é bem representado pela Figura 3, onde para um valor de μ fixo e variando ϕ , pode observar com clareza que, aumentando o valor de ϕ , a dispersão da função densidade diminui.

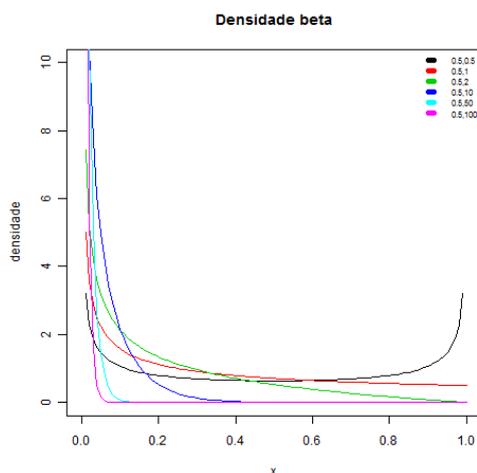


Figura 3: Fixado um valor para μ , aumentando ϕ , a dispersão da função densidade diminui.

O modelo de regressão beta é obtido assumindo que a média de y_t pode ser escrita como

$$g(\mu_t) = \sum_{i=1}^k x_{ti}\beta_i = \eta_t, \quad (20)$$

onde $\beta = (\beta_1, \beta_2, \dots, \beta_k)$ é um vetor de parâmetros de regressão desconhecidos ($\beta \in \mathbb{R}^n$) e (x_{t1}, \dots, x_{tk}) são as observações das k covariáveis fixas e conhecidas e $g(\cdot)$ é a função de ligação que mapeia $(0, 1)$, estritamente monótona e duas vezes diferenciável em \mathbb{R} . A variância de y_t é uma função de μ_t e como uma consequência dos valores das covariáveis, assim as variâncias da resposta, não constantes, são naturalmente acomodadas dentro do modelo.

Baseado em apenas uma única observação a log-verossimilhança é dada por

$$L(\mu, \phi) = \log P(y_i, \mu, \phi) = \log \Gamma(\phi) - \log \Gamma(\mu\phi) - \log \Gamma((1 - \mu)\phi) + (\mu\phi - 1) \log y + [(1 - \mu)\phi - 1] \log(1 - y).$$

A Função Escore

$$U_\mu(\mu, \phi) = \frac{\partial l(\mu, \phi)}{\partial \mu} = -\log \Psi(\phi\mu)\Gamma(\phi\mu)\phi + \Psi((1 - \mu)\phi)\phi + \phi \ln(y) - \ln(1 - y)$$

$$U_{\phi}(\mu, \phi) = \frac{\partial l_{\mu, \phi}}{\partial \phi} = \log \Psi(\phi) \Gamma(\phi) - \log \Psi(\phi \mu) \Gamma(\phi \mu) \mu - \Psi(1 - \mu) \phi (1 - \mu) + \mu \ln(y) + (1 - \mu) \ln(1 - y).$$

Assumindo que

$$\begin{aligned} y^* &= \log\left(\frac{y}{1-y}\right), \\ \mu^* &= \Psi(\mu \phi) - \Psi((1 - \mu) \phi), \\ \psi(\lambda) &= \frac{\partial \log \Gamma(\lambda)}{\partial \lambda} \text{ e} \\ \psi(\cdot) &\rightarrow \text{função digama.} \end{aligned}$$

As Funções Escores de μ e ϕ podem ser reescritas como

$$U_{\mu}(\mu, \phi) = \frac{\partial l(\mu, \phi)}{\partial \mu} = \phi(y^* - \mu^*) \quad (21)$$

$$U_{\phi}(\mu, \phi) = \frac{\partial l_{\mu, \phi}}{\partial \phi} = \mu(y^* - \mu^*) + \log(1 - y) - \psi(1 - \mu) \phi + \psi(\phi). \quad (22)$$

Ferrari e Cribari Neto (2004), apresentam esse resultado de forma matricial e para isso é necessário entender a seguinte definição

$$\frac{\partial l(\beta, \phi)}{\partial \beta_i} = \sum_{t=1}^n \frac{\partial l_t(\mu_t, \phi)}{\partial \mu_t} \frac{d\mu_t}{d\eta_t} \frac{\partial \eta_t}{\partial \beta_i}.$$

Observe que

$$\frac{d\mu_t}{d\eta_t} = \frac{1}{g'(\mu_t)}.$$

Desta forma a Função Escore pode ser reescrita matricialmente como segue

$$U_{\beta}(\beta, \phi) = \phi X^{\top} T (y^* - \mu^*) \quad (23)$$

com X sendo uma matriz cuja t -ésima linha é x_t , $T = \text{diag}\left\{\frac{1}{g'(\mu_1)}, \dots, \frac{1}{g'(\mu_n)}\right\}$, $y^* = (y_1^*, \dots, y_n^*)$, $\mu^* = (\mu_1^*, \dots, \mu_n^*)$ e

$$U_{\phi}(\beta, \phi) = \sum_{t=1}^n \{\mu_t(y^* - \mu^*) + \log(1 - y_t) - \psi((1 - \mu_t) \phi) + \psi(\phi)\}. \quad (24)$$

Os estimadores de máxima verossimilhança de β e ϕ são obtidos das equações $U_{\beta}(\beta, \phi) = 0$ e $U_{\phi}(\beta, \phi) = 0$ e não têm forma analítica, portanto é preciso obtê-los numericamente usando um algoritmo de otimização, geralmente não linear, tal como o algoritmo de Newton Raphson, ou um algoritmo de *quasi-Newton*. Ferrari e Cribari Neto (2004) sugerem usar como ponto de partida para estimar β os estimadores de mínimos quadrados ordinários do vetor de parâmetros obtido a partir de uma regressão linear das respostas transformadas $g(y_1), \dots, g(y_n)$ em X , isto é: $(X^{\top} X)^{-1} X_z^{\top}$, onde $z = (g(y_1), \dots, g(y_n))^{\top}$. E como palpite inicial para ϕ sugerem

$$\frac{1}{n} \sum_{t=1}^n \frac{\check{\mu}_t(1 - \check{\mu}_t)}{\check{\sigma}_t^2} - 1 \quad (25)$$

onde, $\check{\mu}_t$ é obtido aplicando $g^{-1}(\cdot)$ ao t -ésimo valor ajustado de regressão linear de $g(y_1), \dots, g(y_n)$ em X , isto é, $\check{\mu}_t = g^{-1}(X^{\top} (X^{\top} X)^{-1} X_z^{\top})$ e $\check{\sigma}_t^2 = \frac{\check{e}^{\top} \check{e}}{(n-k)\{g'(\check{\mu}_t)\}}$, aqui $\check{e} = z - X(X^{\top} X)^{-1} X_z^{\top}$ é o vetor de resíduos de mínimos quadrados de regressão linear da resposta transformada.

O próximo passo é a obtenção da Matriz de Informação de Fisher.

$$k = \begin{bmatrix} -E\left(\frac{\partial^2 l(\mu, \phi)}{\partial \mu^2}\right) & -E\left(\frac{\partial^2 l(\mu, \phi)}{\partial \mu \partial \phi}\right) \\ -E\left(\frac{\partial^2 l(\mu, \phi)}{\partial \phi \partial \mu}\right) & -E\left(\frac{\partial^2 l(\mu, \phi)}{\partial \phi^2}\right) \end{bmatrix} = \begin{bmatrix} k_{\mu\mu} & k_{\mu\phi} \\ k_{\phi\mu} & k_{\phi\phi} \end{bmatrix}.$$

As derivadas de segunda ordem foram obtidas no aplicativo *Maple* e o cálculo da esperança negativa das segundas derivadas com as respectivas simplificações fornece

$$k_{\mu\mu} = \phi^2[\psi'(\mu\phi) + \psi'((1-\mu)\phi)]$$

$$k_{\phi\phi} = \mu^2\psi'(\mu\phi) + (1-\mu)^2\psi'((1-\mu)\phi) - \psi'(\phi)$$

$$k_{\mu\phi} = k_{\phi\mu} = \phi[\mu\psi'(\mu\phi) - (1-\mu)\psi'((1-\mu)\phi)], \text{ sendo } \psi'(\cdot) \text{ a função trigama.}$$

Seja $w = \text{diag}\{w_1, w_2, \dots, w_n\}$ com $w_t = \phi\{\psi'(\mu_t\phi) + \psi'((1-\mu_t)\phi)\} \frac{1}{\{g'(\mu_t)\}^2}$, $c = (c_1, \dots, c_n)^\top$, com $c_t = \phi\{\psi'(\mu_t\phi)\mu_t - \psi'((1-\mu_t)\phi)(1-\mu_t)\}$, onde $\psi'(\cdot)$ a função trigama. Seja também $D = \text{diag}\{d_1, \dots, d_n\}$, com $d_t = \psi'(\mu_t\phi)\mu_t^2 + \psi'((1-\mu_t)\phi)(1-\mu_t)^2 - \psi'(\phi)$.

A Matriz de Informação é dada por

$$k = k(\beta, \phi) = \begin{bmatrix} k_{\beta\beta} & k_{\beta\phi} \\ k_{\phi\beta} & k_{\phi\phi} \end{bmatrix}$$

onde:

$$k_{\beta\beta} = \phi X^\top W X$$

$$k_{\beta\phi} = k_{\phi\beta}^\top = X^\top T c$$

$$k_{\phi\phi} = \text{tr}(D)$$

Os parâmetros β e ϕ não são ortogonais. E sob condições normais de regularidade para a estimação de máxima verossimilhança quando a amostra é grande,

$$\begin{pmatrix} \hat{\beta} \\ \hat{\phi} \end{pmatrix} \sim N_{k+1} \left(\begin{pmatrix} \beta \\ \phi \end{pmatrix}; k^{-1} \right),$$

onde $\hat{\beta}$ e $\hat{\phi}$ são os estimadores de máxima verossimilhança de β e ϕ respectivamente.

Portanto é útil obter uma expressão para a Inversa da Matriz de Informação de Fisher, representada por

$$k^{-1} = k^{-1}(\beta, \phi) = \begin{pmatrix} k^{\beta\beta} & k^{\beta\phi} \\ k^{\phi\beta} & k^{\phi\phi} \end{pmatrix},$$

onde

$$k^{\beta\beta} = \frac{1}{\phi} (X^\top W X)^{-1} \left\{ I_k + \frac{X^\top T c c^\top X (X^\top W X)^{-1}}{\gamma\phi} \right\}, \text{ com } \gamma = \text{tr}(D) - \phi^{-1} c^\top T^\top X (X^\top W X)^{-1} X^\top T c$$

$$k^{\beta\phi} = (k^{\phi\beta})^\top = \frac{-1}{\gamma\phi} (X^\top W X)^{-1} X^\top T c,$$

$$k^{\phi\phi} = \gamma^{-1}$$

I_k é uma matriz identidade $k \times k$.

3.3 Aplicação

Foi utilizado um estudo clínico sobre transtorno de ansiedade generalizada, conduzido pelo professor Dipankar Bandyopadhyay, do departamento de saúde pública, divisão de bioestatística, da Universidade de Minnesota. O estudo foi aleatorizado e controlado por placebo, envolveu 326 participantes em 5 centros clínicos. Estes participantes são acompanhados por até 8 semanas. O resultado primário é se os participantes responderam ao tratamento ou não. A Variável Resposta é definida como a porcentagem da redução da pontuação final do participante, medida pela linha de base na Escala de Ansiedade de Hamilton (HAM-A). O conjunto de dados está localizado na *webpage* <http://www.biostat.umn.edu/~dipankar/bmtry711.11/exam2prob1.txt>.

Foi aplicada a regressão beta nesses dados, utilizando o pacote *betareg*, do *software* estatístico *R*. A variável resposta é a razão entre o escore final e o inicial dos indivíduos participantes, e as covariáveis são $x1$: droga (ativo = 1 ou placebo = 0) e $x2$: número de semanas que participou do estudo. Os parâmetros estimados e seus respectivos erros padrão, acompanhados do valor da estatística z e o valor p são apresentados na Tabela 1 abaixo:

Como pode-se observar pela estimativa do parâmetro *Droga*, que o erro padrão é alto quando comparado com o valor da estimativa e ainda, há evidência de que esta não ajuda explicar a variabilidade da variável resposta, o que não ocorre com os demais, como apresentado na Tabela 1.

Tabela 1: Parâmetros estimados.

Parâmetro	Estimado	Erro Padrão	Estatística z	valor z
β_0	1,7574	0,3211	-5,64	¡0,001
<i>Droga</i>	-0,1294	0,1019	-1,27	0,2052
<i>Semanas</i>	-0,2329	0,04131	-5,64	¡0,001

A estimativa de ϕ é de 4,0492, com erro padrão 0,2982, um valor z 13,58 e a probabilidade $< 0,001$ e as estatísticas de ajuste do modelo foram: $-2 \text{ Log Likelihood} = -90,9$, $AIC = -82,9$, $AICC = -82,8$ e $BIC = -68,1$. Com base nestes critérios, recomenda-se o uso do modelo ajustado aos dados. Parece que o paciente que é acompanhado por mais tempo (semanas) apresenta melhores resultados, independente do uso da droga.

Abaixo seguem os gráficos do ajuste da regressão beta aos dados de ansiedade:

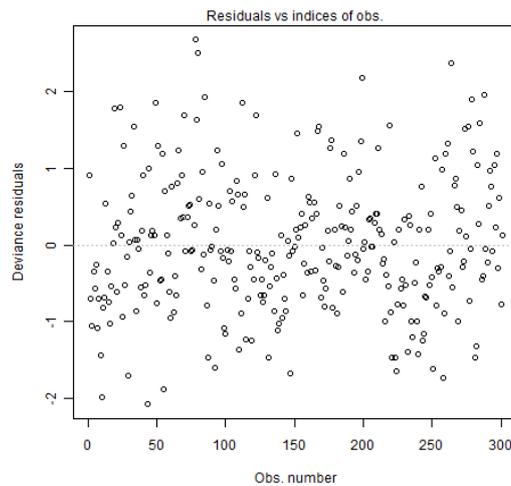


Figura 4: Gráficos para análise de influência da Regressão da resposta transformada.

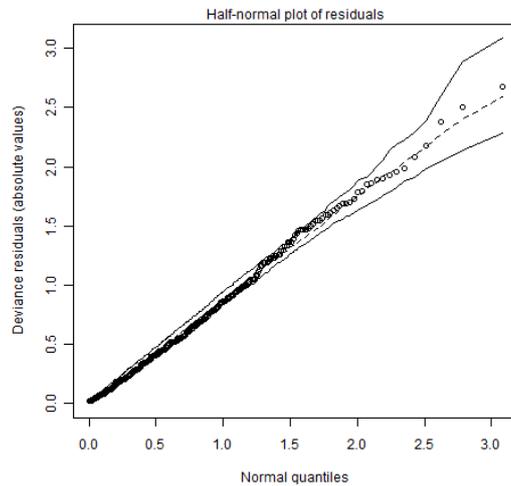


Figura 5: Gráficos para análise de influência da Regressão da resposta transformada.

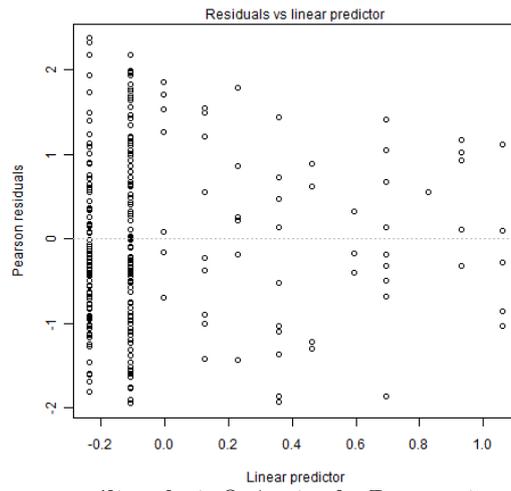


Figura 6: Gráficos para análise de influência da Regressão da resposta transformada.

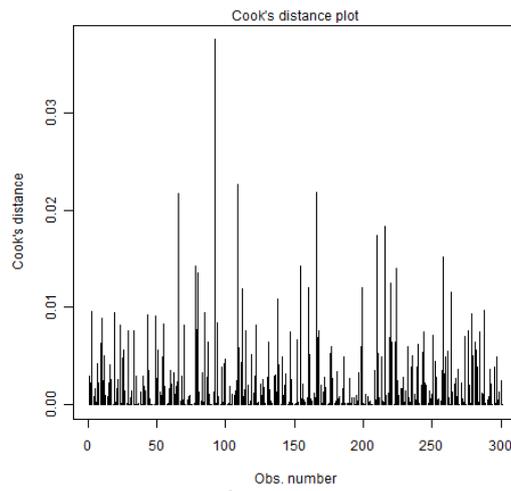


Figura 7: Gráficos para análise de influência da Regressão da resposta transformada.

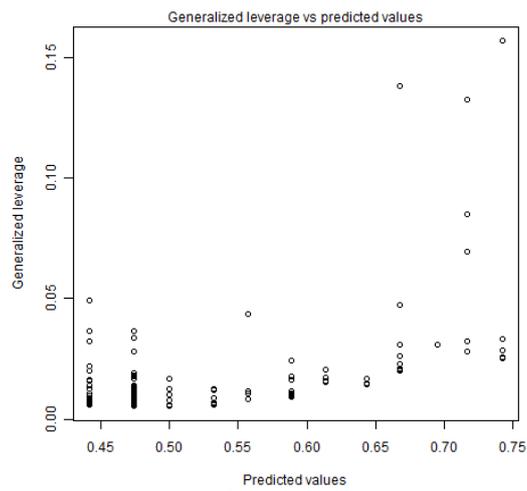


Figura 8: Gráficos para análise de influência da Regressão da resposta transformada.

4 Conclusões

Esta pesquisa de iniciação científica abordou a distribuição beta e seus principais cumulantes, bem como função escore e matriz de informação de Fisher, e obteve os mesmos resultados teóricos apresentados por Johnson, Kotz e Balakrishnan (1995) e adicionalmente apresentou o modelo de regressão beta que também obteve os mesmos resultados apresentados por Ferrari e Cribari Neto (2004). Com isso esta pesquisa contribuiu para a divulgação do modelo beta para adotá-lo em situações onde a variável resposta $y \in (0, 1)$ e, ainda, seus parâmetros podem ser interpretados em termos de *odds ratio* quando a função de ligação utilizada for a logit, e isto, tem um grande apelo nas áreas de saúde e biológicas.

Referências

- [1] AKINSETE, A; FAMOYEB, F.; LEEB C.; The beta-Pareto distribution. **Statistics**. v. 42, n.6 , 2008.
- [2] BRITO, R.S. **Estudo de expansões assintóticas, avaliação numérica de momentos das distribuições beta generalizadas, aplicações em modelos de regressão e análise discriminante**. 2009. 124 p. Dissertação (Mestrado em Biometria e Estatística Aplicada), Universidade Federal Rural de Pernambuco, Pernambuco, 2009.
- [3] DEGROOT, M. H. **Probability and Statistics**. Addison Wesley Publishing Company. 1989.
- [4] FERRARI, S.L.P. and CRIBARI-NETO, F. Beta regression for modelling rates and proportions. **Journal of Applied Statistics**. Elsevier. v. 7, n. 2, p.799-815, 2004.
- [5] FERRARI, S.L.P. **Beta Regression Modeling: Recent Advances in Theory and Applications**. 13th School of Regression Models, February 2013. Disponível em: <http://http://www.ime.usp.br/sferrari/>. Acesso em: 28 de fevereiro de 2014.
- [6] JOHNSON, N.L. and KOTZ, S. and BALAKRISHNAN, L. **Continuous Univariate Distributions**. Chapman Hall. 1995.
- [7] OLIVEIRA, M.S.. **Um modelo de regressão Beta: teoria e aplicações..** 2004. Dissertação de mestrado, IME-USP, São Paulo, 2004.
- [8] VENEZUELA, M.K. **Equação de Estimação generalizada e influência local para modelos de regressão beta com medidas repetidas..** 2008. Tese de Doutorado, IME-USP, São Paulo, 2008.
- [9] University of Minesota, School of Public Health, Biostatistics. **Dipankar Bandyopadhyay**. Disponível em: <http://www.biostat.umn.edu/dipankar/bmtry711.11/exam2prob1.txt>. Acesso em: 28 de fevereiro de 2014.