Distribuição Kumaraswamy
E suas Aplicações

Alisson Rosa, Digite seu nome aqui!

Resumo

Muitas vezes estamos interessados em modelar variáveis que estão definidas entre zero e um, como sabemos aonde nossa variável esta definida mas não sabemos qual dos valores será observado, temos portanto uma incerteza probabilística, que pode e deve ser modelada por medidas de probabilidade. Aqui portanto, introduziremos a distribuição Kumaraswamy, que é uma das muitas possibilidades para modelagem desse tipo de variável, encontraremos estimativas para os parâmetros da distribuição usando estimadores de máxima verossimilhança blah blah.

1 Introdução

Atualmente muitos fenômenos podem ser descritos como variáveis aleatórias (va) definidas no intervalo unitário (0,1) [1], assim é natural que pesquisadores desenvolvam distribuições de probabilidade que abarcam esse tipo de va. Uma dessas distribuições é a Kumaraswamy, que foi introduzida em Kumaraswamy (1980) como uma alternativa ao modelo beta para aplicações na´area de hidrologia. Em virtude deste fato, grande parte dos trabalhos empíricos desta distribuição concentra-se nessa área Nadarajah (2008). O presente trabalho visa contribuir na expansão e utilização da Kumaraswamy, empregando modelos incondicionais para a taxa de desfloresmento em diversos munícipios da Amazônia legal, disponilizados (http://www.dpi.inpe.br/prodesdigital/prodesmunicipal.php)[aqui] pelo projeto PRODES do INPE. Assim é possível mensurar a qualidade da Kumaraswamy para modelagem dos dados propostos, para isso estamos utilizando 6 métricas frequentistas estabelecidas: AIC, BIC, CAIC, Kolmogorov-Smirnov, Cramer-Von Mises e Anderson-Darling, em contraste com a Distribuição Normal e a Distribuição Beta.

2 A distribuição Kumaraswamy

Vamos nessa seção introduzir quantidades básicas da distribuição Kumaraswamy, sendo elas sua função densidade de probabilidade (pdf), função de distribuição acumulada (cdf), função quantilica (qf), função de verossimilhança (ll) e esperança (**E**)

2.1 Quantidade Básicas

Seja X uma variável aleatória que segue uma distribuição Kumaraswamy, então sua cdf é dada por:

$$F(x;\alpha,\beta) = 1 - (1 - x^{\alpha})^{\beta}, \quad 0 < x < 1 \tag{1}$$

Onde $\alpha, \beta > 0$. Sua pdf então fica definida como:

$$f(x;\alpha,\beta) = \frac{dF}{dx} = \alpha\beta x^{\alpha-1}(1 - x^{\alpha})^{\beta-1}, \quad 0 < x < 1 \tag{2}$$

Sua qf, que é a função inversa da cdf, fica definida como:

[1] Onde parenteses indica limites do intervalo abertos.

$$Q(u;\alpha,\beta) = \left(1-(1-u)^{1/\beta}\right)^{\dfrac{1}{\alpha}}, \quad 0 < u < 1 \tag{3}$$

É FÁCIL ver que que a esperança da distribuição Kumaraswamy é dada por

$$\mathrm{E}(X) = \frac{\beta\Gamma\left(1+\dfrac{1}{\alpha}\right)\Gamma(\beta)}{\Gamma\left(1+\dfrac{1}{\alpha}+\beta\right)} \tag{4}$$

A função de verossimilhança é dada por:

$$L(\alpha,\beta;x) = \prod_{i=1}^{n} f(x;\alpha,\beta) = \alpha^n\beta^n \prod_{i=1}^{n} x_i^{\alpha-1} \prod_{i=1}^{n} (1-x_i^{\alpha})^{\beta-1} \tag{5}$$

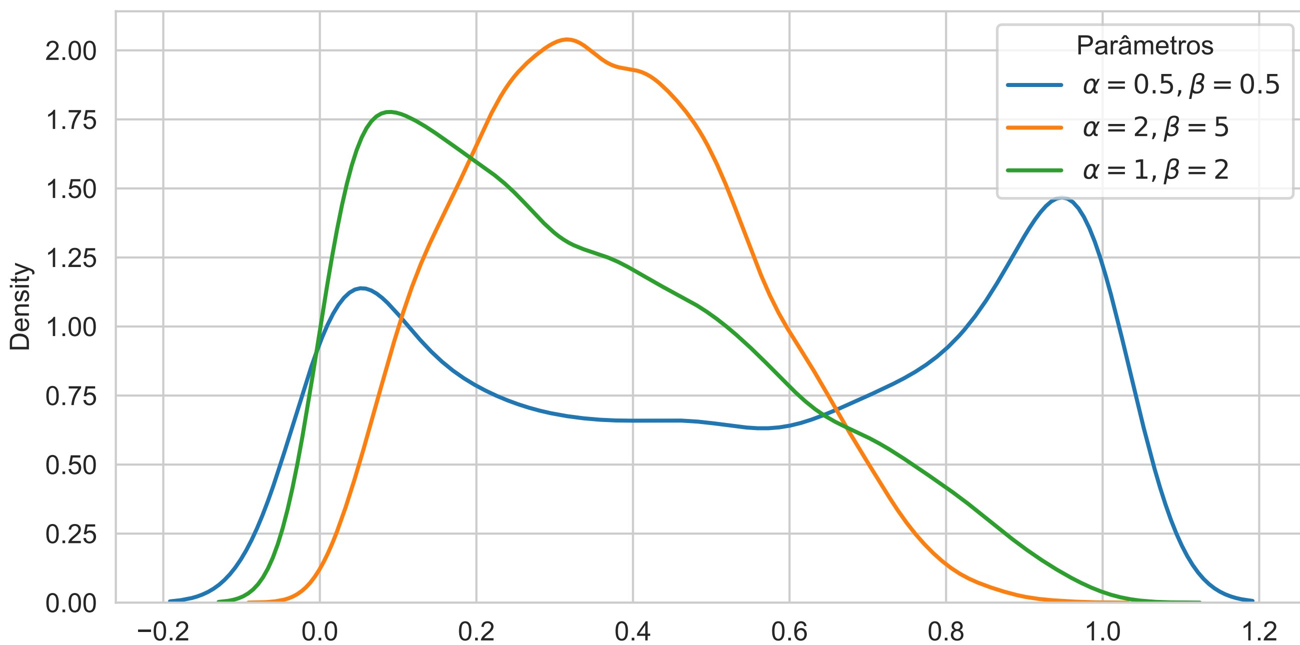


Figura 1: Função densidade da Kumaraswamy para alguns valores de parâmetros

2.2 Justificativa

O artigo de Kumaraswamy (1980) propõe e demonstra aplicações da distribuição Kumaraswamy para variáveis aleatórias e processos aleatórios derivados de processos hidrológicos. O artigo foi publicado na *Journal of Hydrology*, assim é perceptível que a distribuição foi concebida para se adequar a dados hidrológicos. Temos como casos de suas utilizações em precipitação diária, fluxo diário, reservatórios de água e análise das ondas do oceano, entre outras.

Para Nadarajah (2008) a utilização da Kumaraswamy para o campo da hidrologia, é consolidada. Sendo perceptível pelas inúmeras aplicações em diversos artigos como: Sundar e Subbiah (1989), Fletcher e Ponnambalam (2008) e Koutsoyiannis e Xanthopoulos (1989), além de se sobressair em relação a distribuição beta, a distribuição padrão para dados no (0,1), de acordo com Koutsoyiannis e Xanthopoulos (1989). Em Dey, Mazucheli, e Nadarajah (2018) a Kumaraswamy é utilizada para a quantidade de deslocamento de líquido metálico em duas máquinas diferentes, sendo superior as Distribuições Gumbel e Fréchet.

É notório também a disseminação recente do estudo da Kumaraswamy, em expansão para distribuições mais complexas, como em Lemonte, Barreto-Souza, e Cordeiro (2013), Mazucheli et al. (2020), Cribari-Neto e Santos (2019) e Sagrillo, Guerra, e Bayer (2021) ou em suas aplicações em regressão e séries temporais, visto em Pumi, Rauber, e Bayer (2020), Mitnik e Baek (2013), Fábio M. Bayer, Cribari-Neto, e Santos (2021) e Fábio Mariano Bayer, Bayer, e Pumi (2017).

Assim é factível a falta de ajustes da Distribuição Kumaraswamy em outras áreas de estudo, visto que a Kumaraswamy tem disseminação acadêmica exponencial recentemente, com admiráveis contribuições da UFSM, em artigos supracitados. Decidimos contribuir para um maior estudo da distribuição em outra área ambiental, o desmatamento, pensando especificamente na Amazônia brasileira.

O estudo e modelos com bons ajustes de desmatamento e desflorestamento, impactam a vida no mundo todo, tanto humana quanto não humana. Bons modelos conseguem prever e informar quais variáveis mais impactam nos desmatamento, possibilitando verificar sua evolução conforme o tempo. Sendo útil para construção de políticas públicas, visando tomadas de decisão mais eficientes.

O desflorestamento é questão com importância ambiental, social, econômica e até política, pois a floresta amazônica tem seu papel no armazenamento de carbono, evitando o aquecimento global, na reciclagem de água e na manutenção da biodiversidade. Além de fornecerem uma grande variedade de produtos materiais e de sustento para as populações locais. Mesmo com áreas com grandes partes preservadas, há impacto na biodiversidade, pois a distribuição das espécies não é uniforme. Muitas espécies têm áreas de ocorrência restritas a partes que já foram reduzida a pequenos fragmentos.

3 Análise Inicial

blah blah

3.1 Apresentação dos dados

3.2 Medidas Básicas

4 Ajuste do Modelo

Para a verificação da adequação da distribuição Kumaraswamy para modelagem de dados de desmatamento iremos realizar a comparação com duas distribuições amplamente utilizadas: Distribuição Normal e a Distribuição beta, a distribuição mais utilizada para variáveis aleatórias com suporte no (0,1).

A comparação foi construída utilizando 6 métricas: AIC, BIC, CAIC, Kolmogorov-Smirnov, Cramer-Von Mises e Anderson-Darling, que se baseiam principalmente na verossimilhança. A verossimilhança a considera os parâmetros variáveis, assim a função de verossimilhança indica os parâmetros mais plausíveis de terem gerado a amostra. Logo, podemos verificar entre todas as distribuições quais possuem as maiores verossimilhanças, tendo assim os parâmetros mais plausíveis para a geração da amostra. No entanto, nota-se que todos os critérios contam outros elementos para a sua construção.

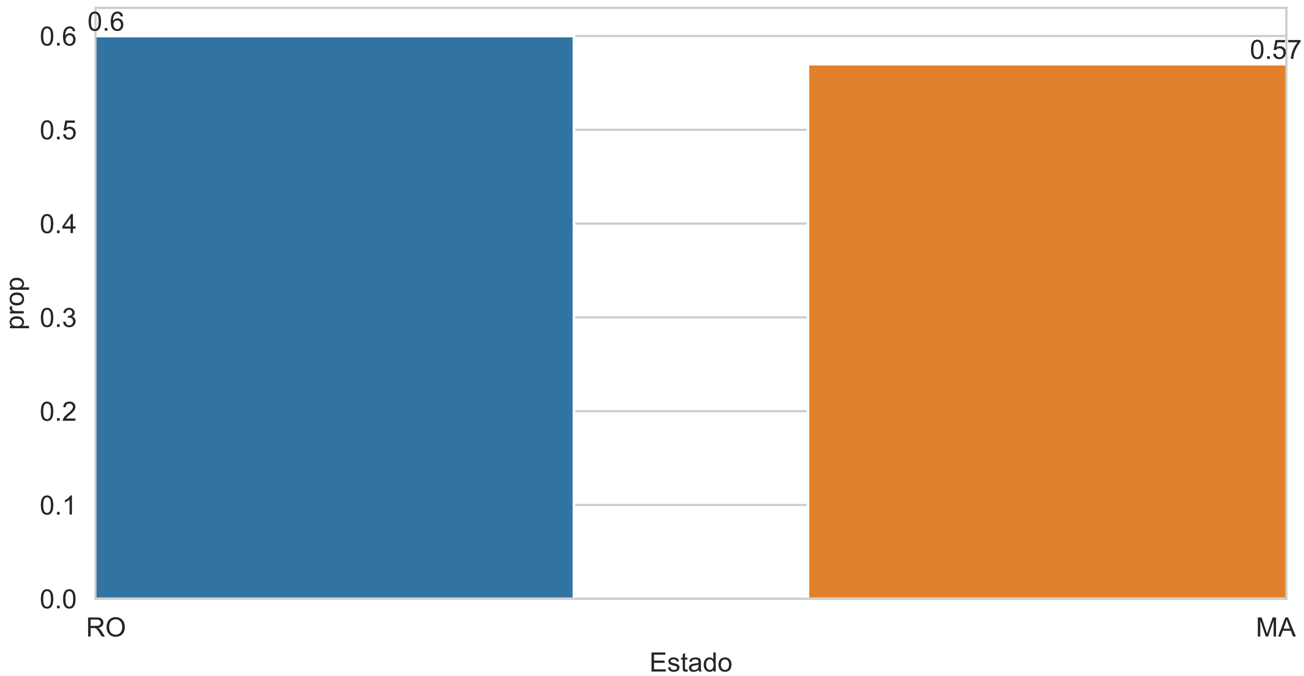
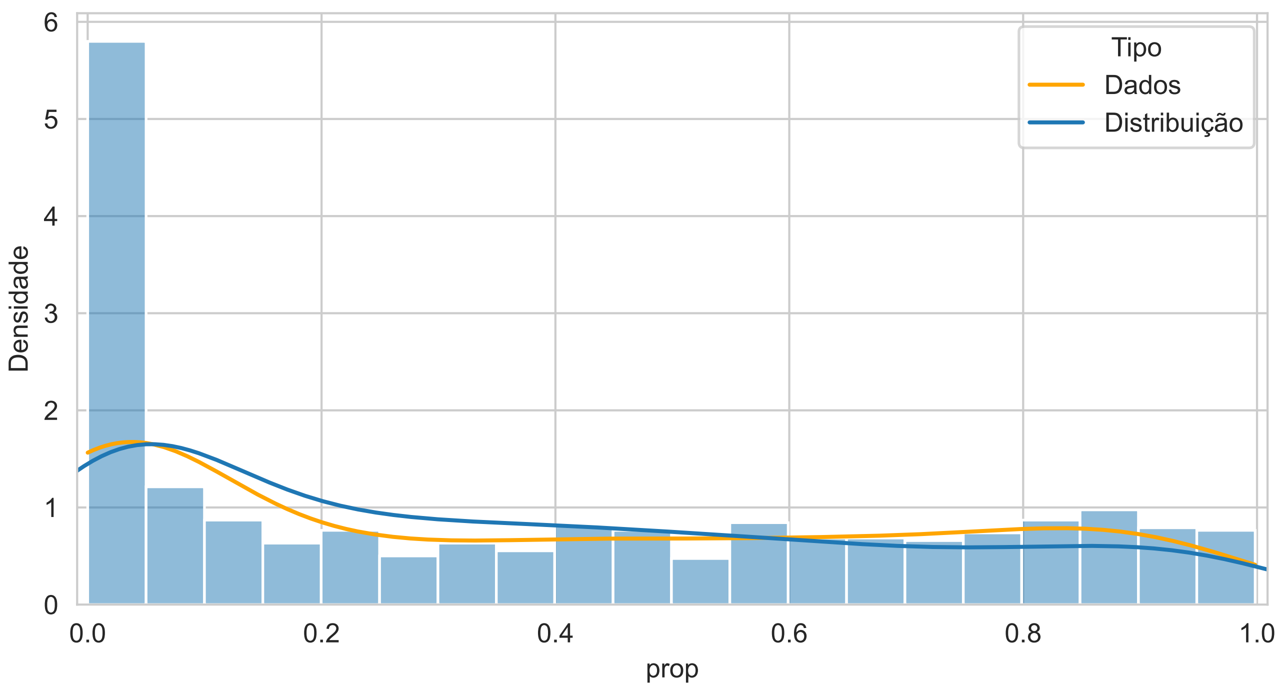


Figura 2: Proporção média dos Estados



	Kumaraswamy	Beta	Normal	Vencedor
0	-440.933936	-443.086140	-154637.091792	Normal
1	-440.918084	-433.819504	-154627.825155	Normal
2	-431.667299	-443.070288	-154637.075940	Normal
3	18.106283	9.736648	767.892814	Normal
4	9.691977	1.251566	64.378061	Normal
5	0.088746	0.089166	0.537076	Normal

5 Conclusão

6 Referências

Bayer, Fábio Mariano, Débora Missio Bayer, e Guilherme Pumi. 2017. "Kumaraswamy autoregressive moving average models for double bounded environmental data". *Journal of Hydrology* 555: 385–96.

Bayer, Fábio M, Francisco Cribari-Neto, e Jéssica Santos. 2021. "Inflated Kumaraswamy regressions with application to water supply and sanitation in Brazil". *Statistica Neerlandica* 75 (4): 453–81.

Cribari-Neto, Francisco, e Jessica Santos. 2019. "Inflated Kumaraswamy distributions". *Anais da Academia Brasileira de Ciências* 91.

Dey, Sanku, Josmar Mazucheli, e Saralees Nadarajah. 2018. "Kumaraswamy distribution: different methods of estimation". *Computational and Applied Mathematics* 37 (2): 2094–2111.

Fletcher, SG, e K Ponnambalam. 2008. "Stochastic control of reservoir systems using indicator functions: new enhancements". *Water Resources Research* 44 (12).

Ganji, A, K Ponnambalam, D Khalili, e M Karamouz. 2006. "Grain yield reliability analysis with crop water demand uncertainty". *Stochastic Environmental Research and Risk Assessment* 20 (4): 259–77.

Koutsoyiannis, Demetris, e Themistocle Xanthopoulos. 1989. "On the parametric approach to unit hydrograph identification". *Water resources management* 3 (2): 107–28.

Kumaraswamy, Ponnambalam. 1980. "A generalized probability density function for double-bounded random processes". *Journal of hydrology* 46 (1-2): 79–88.

Lemonte, Artur J, Wagner Barreto-Souza, e Gauss M Cordeiro. 2013. "The exponentiated Kumaraswamy distribution and its log-transform". *Brazilian Journal of Probability and Statistics* 27 (1): 31–53.

Mazucheli, J, AFB Menezes, LB Fernandes, RP De Oliveira, e ME Ghitany. 2020. "The unit-Weibull distribution as an alternative to the Kumaraswamy distribution for the modeling of quantiles conditional on covariates". *Journal of Applied Statistics* 47 (6): 954–74.

Mitnik, Pablo A. 2013. "New properties of the Kumaraswamy distribution". *Communications in Statistics-Theory and Methods* 42 (5): 741–55.

Mitnik, Pablo A, e Sunyoung Baek. 2013. "The Kumaraswamy distribution: median-dispersion re-parameterizations for regression modeling and simulation-based estimation". *Statistical Papers* 54 (1): 177–92.

Nadarajah, Saralees. 2008. "On the distribution of Kumaraswamy". *Journal of Hydrology* 348 (3): 568–69.

Pumi, Guilherme, Cristine Rauber, e Fábio M Bayer. 2020. "Kumaraswamy regression model with Aranda-Ordaz link function". *Test* 29 (4): 1051–71.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Sagrillo, Murilo, Renata Rojas Guerra, e Fábio M Bayer. 2021. "Modified Kumaraswamy distributions for double bounded hydro-environmental data". *Journal of Hydrology* 603: 127021.

Sundar, V, e K Subbiah. 1989. "Application of double bounded probability density function for analysis of ocean waves". *Ocean engineering* 16 (2): 193–200.