



Ajuste da distribuição Gumbel a dados de precipitação, baseado em valores de recordes

Larissa Bueno Fernandes¹ e Josmar Mazucheli¹

¹Programa de Pós-Graduação em Bioestatística/PBE - Universidade Estadual de Maringá/UEM

RESUMO

Este trabalho apresenta a caracterização frequentista da distribuição Gumbel baseada apenas nos valores de recorde, que pode ser de grande utilidade para a estimação dos parâmetros da distribuição quando apenas os valores de recorde são observados. A metodologia foi aplicada a um conjunto de dados de precipitações mensais da estação meteorológica de Maringá - PR, observados entre 1964 e 2016. Foi verificado que para alguns dos meses considerados, o ajuste aos dados apenas com os valores de recorde foi satisfatório quando comparado ao ajuste com os dados originais. Entretanto, para alguns meses o ajuste pelos recordes se mostra bastante distante da distribuição empírica dos dados, assim como do ajuste por meio dos dados originais. Desta forma, a adoção de outras abordagens ou distribuições para o ajuste ao conjunto de dados deve ser considerada.

Palavras chave: Recordes, Gumbel, precipitação.

1 INTRODUÇÃO

Um recorde é um registro em uma série temporal que é maior (ou menor) do que todos os registros anteriores. Nesse sentido, [1] aponta que um recorde é um valor extremo que é definido em relação a todos os valores anteriores na série de tempo. Especificada uma distribuição de probabilidade e caracterizando a mesma em termos de valores de recorde, é possível estimar os parâmetros da mesma em relação aos dados originais em situações em que apenas os recordes são observados.

Desta forma, considere os valores de recorde superior X_n , $n = 1, 2, \dots$, que se baseiam na sequência de variáveis aleatórias i.i.d. X_1, X_2, \dots , com uma função de distribuição acumulada contínua $F(x | \theta)$ e função de densidade de probabilidade $f(x | \theta)$. De acordo com [2] a função de verossimilhança baseada nos recordes é dada por:

$$L(\Theta | \mathbf{x}) = f(x_n | \Theta) \prod_{i=1}^{n-1} \frac{f(x_i | \Theta)}{1 - F(x_i | \Theta)}. \quad (1)$$

Neste contexto, o presente trabalho tem por objetivo caracterizar a distribuição Gumbel, com base nos valores de recorde superiores, aplicando a metodologia proposta para ajustar os recordes de precipitações mensais observadas na estação meteorológica convencional de Maringá - PR.

2 METODOLOGIA

2.1 Distribuição Gumbel

A distribuição Gumbel foi originalmente proposta por [3], em que os autores definiram três distribuições assintóticas de valores extremos, conhecidas como Gumbel, Fréchet e Weibull ou Valor Extremo tipo I, II e III, respectivamente. A função densidade de probabilidade da distribuição Gumbel é definida por:

$$f(x | \Theta) = \frac{1}{\sigma} \exp \left[-\frac{x - \mu}{\sigma} - \exp \left(-\frac{x - \mu}{\sigma} \right) \right], \quad (2)$$

em que $\Theta = (\mu, \sigma)$,

$$F(x | \Theta) = \exp \left[- \exp \left(- \frac{x - \mu}{\sigma} \right) \right]. \quad (3)$$

Substituindo a função densidade e a função de distribuição acumulada dadas em 2 e 3, respectivamente, na equação 1, a função de verossimilhança baseada nos recordes é dada por:

$$\begin{aligned} L(\Theta | X) &= \frac{1}{\sigma} \exp[-z_m - \exp(-z_m)] \prod_{i=1}^{m-1} \frac{\frac{1}{\sigma} \exp[-z_i - \exp(-z_i)]}{1 - \exp[-\exp(-z_i)]} \\ &= \frac{1}{\sigma^m} \prod_{i=1}^m \exp[-z_i - \exp(-z_i)] \prod_{i=1}^{m-1} (1 - \exp[-\exp(-z_i)])^{-1} \end{aligned} \quad (4)$$

em que $z_i = \frac{x_i - \mu}{\sigma}$ e m é o número de recordes observados. Assim, o logaritmo da função de verossimilhança é dado por:

$$\ell(\Theta | X) = m \log \sigma + \sum_{i=1}^m [-z_i - \exp(-z_i)] - \sum_{i=1}^{m-1} \log(1 - \exp[-\exp(-z_i)]) \quad (5)$$

Portanto, o estimador de máxima verossimilhança (EMV) do vetor de parâmetros Θ , denotado por $\hat{\Theta}$, pode ser obtido maximizando a equação 5, com respeito a Θ .

2.2 Dados

Considerou-se como variável de estudo o total de precipitação acumulada no mês da estação meteorológica convencional da cidade Maringá - PR, obtidas no Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP) do Instituto Nacional de Meteorologia (INMET), órgão responsável pela coleta e disponibilização informações meteorológicas oficiais.

O período considerado foi de 1964 a 2016, sendo que as informações referentes aos volumes de precipitação mensais não estavam disponíveis para todos os meses do mesmo. Assim, no total, foram utilizadas 527 observações, divididas em 12 séries mensais para o ajuste da distribuição.

3 RESULTADOS E DISCUSSÕES

Os resultados deste ajuste e do ajuste aos dados originais, assim como a distribuição empírica da precipitação mensal, são apresentados na Figura 1.

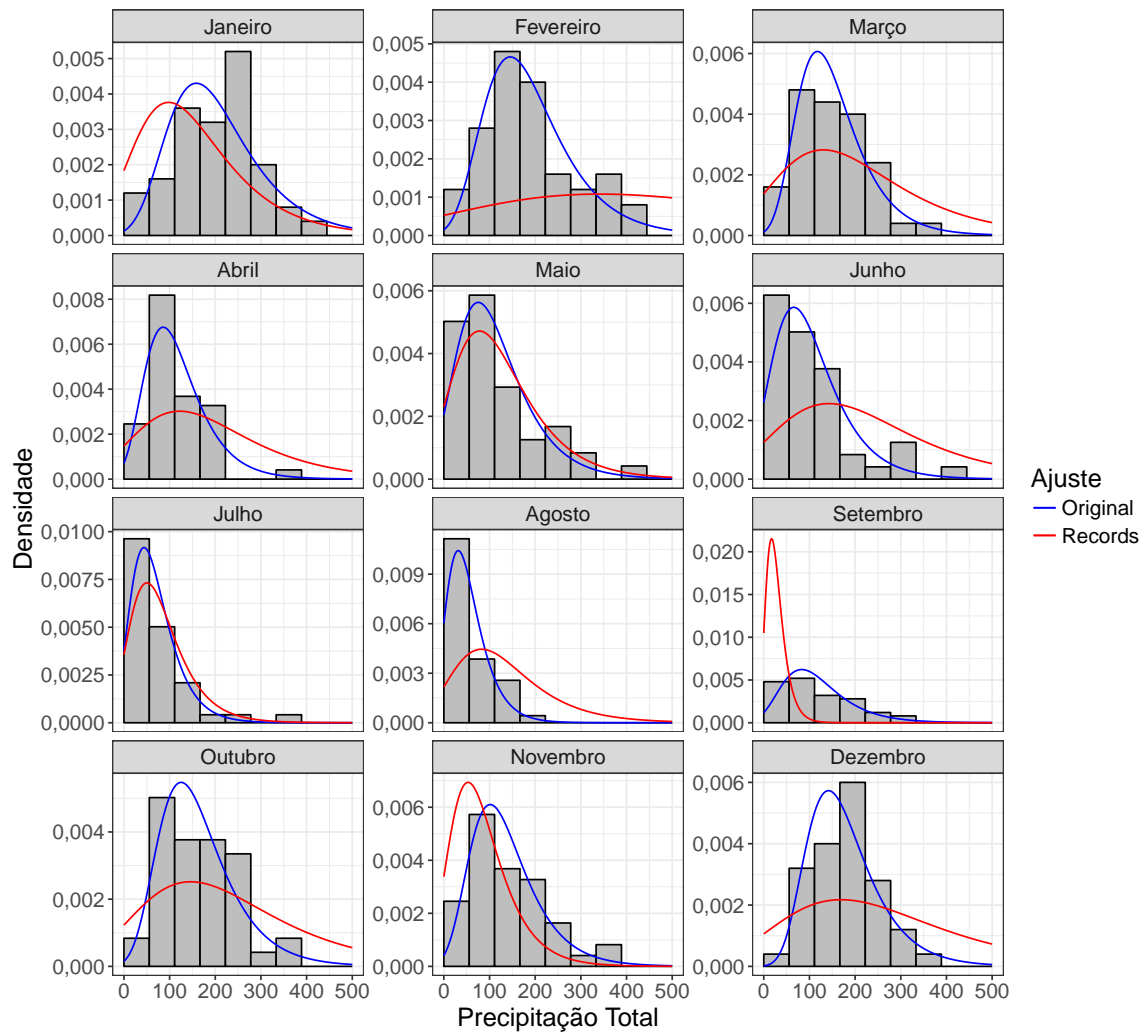


Figura 1: Distribuição e ajuste da distribuição Gumbel, por meio dos dados originais e dos valores de recorde, as séries históricas de precipitação pluviométrica agrupadas por mês, da estação meteorológica convencional de Maringá – PR, de 1964 a 2016.

Pela Figura 1, nota-se que os ajustes obtidos por meio dos valores de recorde são bastante próximos dos ajustes obtido por meio dos dados originais para os meses de Maio e Julho. Já para os meses de Fevereiro e Setembro, o ajuste pelos recordes se mostra bastante diferente da distribuição empírica dos volumes mensais de precipitação, assim como do ajuste por meio dos dados originais.

4 CONCLUSÃO

Com o presente trabalho foi avaliado a utilização do esquema de recordes para o ajuste de variáveis climatológicas, especificamente para o volume mensal de precipitação pluviométrica, considerando a distribuição Gumbel. Foi verificado que para alguns dos meses considerados, o ajuste aos dados apenas com os valores de recorde foi satisfatório quando comparado ao ajuste com os dados originais. Entretanto, para alguns meses o ajuste pelos recordes se mostra bastante diferente da distribuição empírica dos volumes mensais de precipitação, assim como do ajuste por meio dos dados originais.

Levanda-se a hipótese de que a abordagem Bayesiana para a estimação dos parâmetros da distribuição Gumbel ou até mesmo a consideração de outra distribuição, como a Nakagami, possa apresentar melhores resultados, sendo que tais propostas podem ser consideradas no desenvolvimento de trabalhos futuros.

5 AGRADECIMENTOS

Agradecemos ao Instituto Nacional de Meteorologia (INMET) por deixar disponíveis no portal (<http://www.inmet.gov.br>) os dados analisados neste trabalho.

Referências

- [1] WERGEN, G. Records statistics beyond the standard model-Theory and applications. Tese (Doutorado) — Universität zu Köln, 2012.
- [2] ARNOLD, B. C.; BALAKRISHNAN, N.; NAGARAJA, H. N. Records. John Wiley & Sons, 2011.
- [3] FISHER, R. A.; TIPPETT, L. H. C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In: CAMBRIDGE UNIV PRESS. Mathematical Proceedings of the Cambridge Philosophical Society. v. 24, n. 02, p. 180–190, 1928.