A new bivariate survival model: Modeling, Inference and Influence Analysis

Natan Hilário da Silva[1], *Adriano Kamimura Suzuki*[1]
[1] University of São Paulo (USP)

Abbreviated abstract: The study and modeling of lifetime data is a growing area of interest in problems involving healthcare and reliability systems. This work proposes a new model for the treatment of paired bivariate times based on a new one parameter distribution [1], considering right censored data. It is proposed mostly Bayesian inferential methods along with the models properties acquired from simulation studies. It is also proposed a way of *outlier* detection. This work presents applications on two different bivariate lifetime datasets.

Related publications:

[1] – Alshenawy, R. A new one parameter distribution: properties and estimation with applications to complete and type II censored data. Journal of Taibah University for Science, 14(1):11-18 (2020)
[2] – Peng & Dey. Bayesian analysis of outlier problems using divergence measures. Canadian Journal of Statistics, 23(2):199-213 (1995).

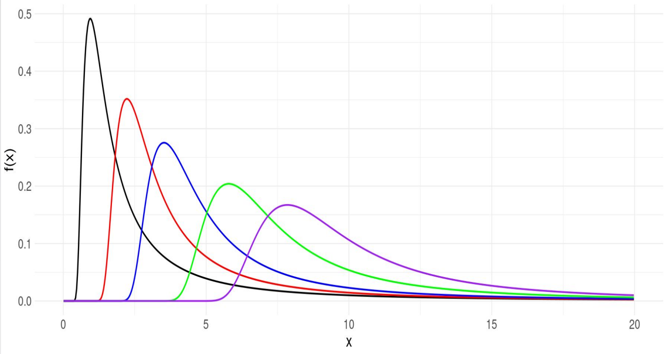
natan.hilario@usp.br

4th Conference on
Statistics and Data Science
Salvador, Brazil (online)
December 1-3, 2022

Previous Works and Main Idea

- The distribution A is a new one parameter distribution proposed in 2020 [1] with a heavy right tail, which shows promising properties for the study of lifetime data.
- New multivariate distributions can be directly obtained from the marginal distributions by the use of copula functions. A copula is a function C such that

$$C : I^n \to I,\ I = [0,1]$$



We propose the new bivariate model, described by marginals following the distribution A united by a copula from the family Farlie-Gumbel-Morgestern.

$$S_i(i) = P(i > u) = 1 - \exp\left(\frac{1}{\beta_i}\left[1 - \exp\left(\frac{\beta_i}{u}\right)\right]\right)$$
$$u > 0, \beta_i > 0, i = X, Y$$

$$C_\phi(u,v) = uv\left[1 + \phi(1-u)(1-v)\right]$$
$$-1 < \phi < 1$$

$$S_{X,Y}(x,y) = P[X > x, Y > y] = C_\phi(S_X(x), S_Y(y))$$

natan.hilario@usp.br

Methods

- Once the model is defined, we consider *priori* distributions for the unknown parameters $\boldsymbol{\theta} = (\phi, \beta_X, \beta_Y)$

$$\beta_X, \beta_Y \sim \text{Gamma}(0.0001, 0.0001)$$
$$\phi \sim U(-1, 1)$$

- The *posteriori* distributions for the parameters are obtained using the model's density function, defined as

$$g(x, y|\boldsymbol{\theta}) = \begin{cases} S_{X,Y}(x, y|\boldsymbol{\theta}) \text{ if } \delta_x = 0 \text{ and } \delta_y = 0, \\ -\dfrac{\partial S_{X,Y}(x, y|\boldsymbol{\theta})}{\partial x}, \text{ if } \delta_x = 1 \text{ and } \delta_y = 0, \\ -\dfrac{\partial S_{X,Y}(x, y|\boldsymbol{\theta})}{\partial y}, \text{ if } \delta_x = 0 \text{ and } \delta_y = 1, \\ \dfrac{\partial^2 S_{X,Y}(x, y|\boldsymbol{\theta})}{\partial x \partial y}, \text{ if } \underbrace{\delta_x = 1 \text{ and } \delta_y = 1}. \end{cases}$$

Censorship Index Function

- Besides the parameter estimations, it is proposed an *outlier* detection analysis using the ψ divergence, proposed originally in [2].

$$D_\psi(r) = \int \psi\left(\frac{\pi(\boldsymbol{\theta}|\boldsymbol{z}_{(r)})}{\pi(\boldsymbol{\theta}|\boldsymbol{z})}\right)\pi(\boldsymbol{\theta}|\boldsymbol{z})d\boldsymbol{\theta}$$

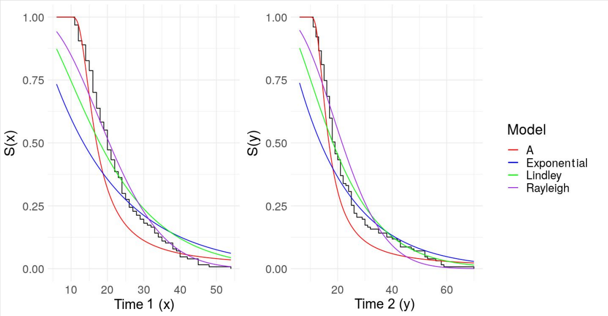
$$\widehat{D}_\psi(r) = \frac{1}{Q}\sum_{q=1}^{Q}\psi\left(\frac{\widehat{CPO}_r}{g(\boldsymbol{z}_r|\boldsymbol{\theta}^{(q)})}\right)$$

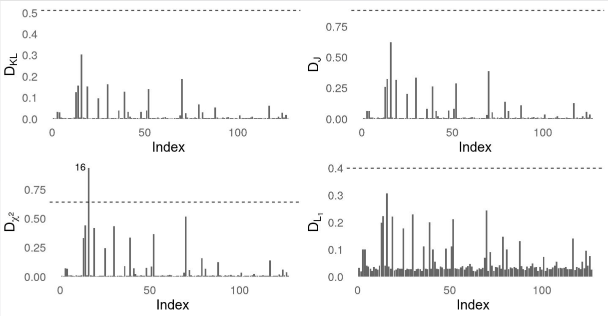
$$\widehat{CPO}_r = \left[\frac{1}{Q}\sum_{q=1}^{Q}\frac{1}{g(\boldsymbol{z}_r|\boldsymbol{\theta}^{(q)})}\right]^{-1}$$

natan.hilario@usp.br

Results and Conclusions

Australian Twins: Appendectomy Ages

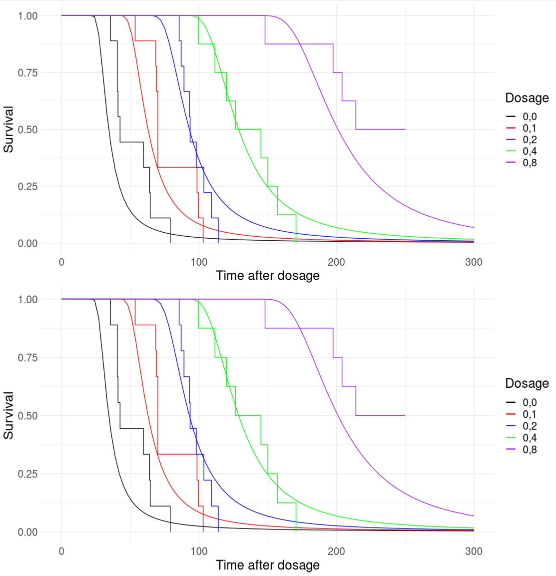




Regression Bivariate A Model: Lead Lab Rats

$$\beta_X = \exp\left\{\gamma_{0x} + \gamma_{2x}w_2 + \gamma_{3x}w_3\right\}$$

$$\beta_Y = \exp\left\{\gamma_{0y} + \gamma_{11y}w_1 + \gamma_{12y}w_1^2 + \gamma_{13y}w_1^3 + \gamma_{2y}w_2 + \gamma_{3y}w_3\right\}$$



natan.hilario@usp.br

4th Conference on Statistics and Data Science
Salvador, Brazil (online)
December 1-3, 2022