MDBED: An R package for simulating the Moran-Downton bivariate exponential distribution

Luis F. Duque[1]

1 School of Engineering, Newcastle University, Newcastle Upon Tyne NE1 7RU, UK

Summary

A bivariate exponential distribution (BED), the Moran-Downton bivariate exponential distribution (MDBED) to be specific, has been widely applied in several fields such as reliability, queueing systems, or hydrology. In the latter, it has been mainly used to simulate the bivariate relationship between the discharge exceedances and their durations (Choulakian, El-Jabi, & Moussi, 1990) or the intensity and duration of a storm of rainfall (Córdova & Rodríguez-Iturbe, 1985). Nagao & Kadoya (1971) further claimed that this bivariate distribution could be used to represent the joint behaviour of the exceedances of flow discharges at two points on a river or rainfall at two locations.

The bivariate modelling of the BED can be addressed by using a 'copula' (Nelsen, 2006), and the R package copula (Yan, 2007) provides a carefully designed and easily extensible platform for that purpose. However, this tool does not offer an option to describe/analyse the conditional properties of the bivariate distribution to be analysed. To the best of my knowledge, there is not a statistical free software that describes the BED fully. The `MDBED` package is an R package that meets this aim by using a parametric approach based on the equations of the Moran-Downton model (Downton, 1970). Through a set of functions, this package provides values of the joint and conditional probability density functions (PDFs) and cumulative distribution functions (CDFs), builds 3D plots of this bivariate distribution, and generates bivariate random values.

The `dBED` and `pBED` functions provide the values of $f(x, y)$ and $F(x, y)$ of the BED for positive pairs (x, y) respectively. The required inputs are the correlation coefficient between the marginal distributions of x and y (ρ_{xy}), the scale parameter of the marginal distribution of x (β_x), the scale parameter of the marginal distribution of y (β_y), and the pairs (x, y). The values of $f(x, y)$ and $F(x, y)$ are computed based on Eqs.18 and 10.54 described in Nagao & Kadoya (1971) and Balakrishna & Lai (2009) respectively.

The functions `PDF_3dPlot` and `CDF_3dPlot` build a 3D plot of $f(x, y)$ and $F(x, y)$ respectively. The main required inputs for these functions are ρ_{xy}, β_x, and β_y. Figure 1 shows an example of the output of the former function. This package also has the function `Countour_Plot` that builds the contour plot of the BED. The main required inputs are ρ_{xy}, β_x, β_y, and the values of $F(x, y)$ associated to the contours. This function also allows pairs of values to be plotted on the contour plot.

The `MDBED` package through the `CondBED` function also describes the conditional PDF and CDF of Y given $X = x$, i.e., $f(Y \mid X = x)$ and $F(Y \mid X = x)$ respectively. The required inputs are ρ_{xy}, β_x, β_y, and a value or vector of values of the marginal distribution of x. $f(Y \mid X = x)$ is defined by using the conditional moments, i.e., the conditional mean $E(Y \mid X = x)$ and variance $var(Y \mid X = x)$, of the Moran-Downton model described in Eqs. 21 and 22 in Nagao & Kadoya (1971). $F(Y \mid X = x)$
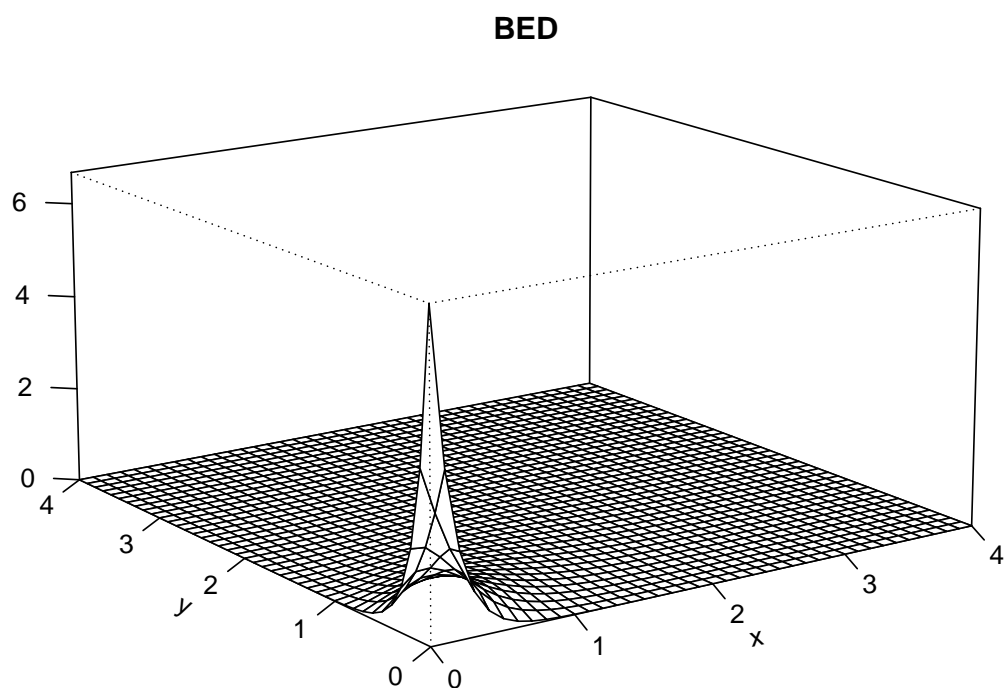
BED

Figure 1: Output of the `PDF_3dPlot` function of the `MDBED` R package. Adopted parameters: ρ_{xy}=0.85, β_x=1, β_y=1
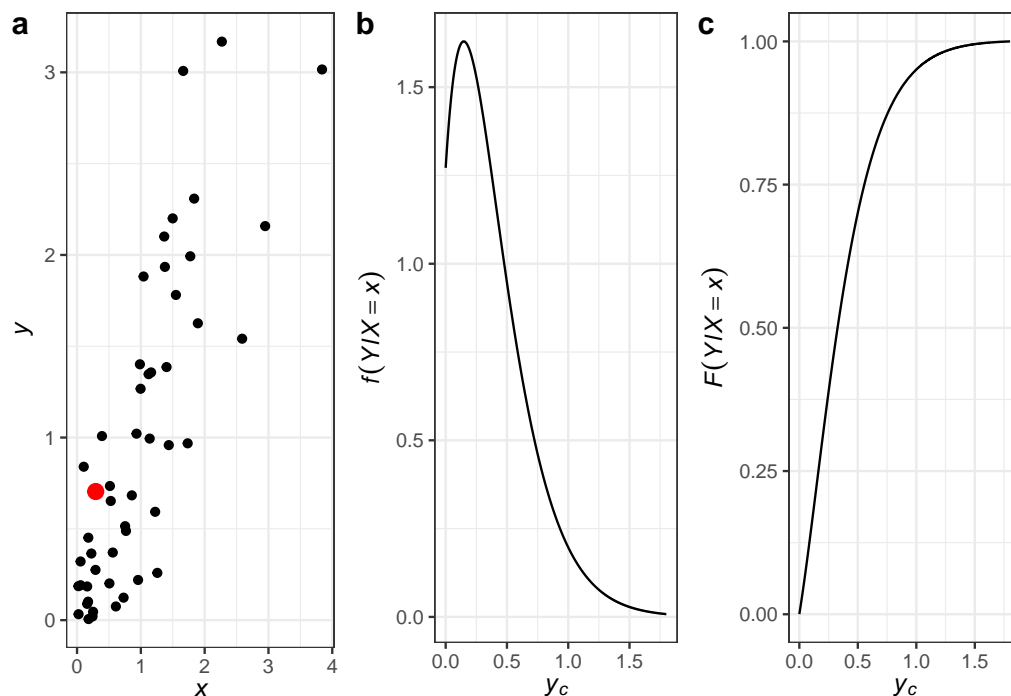

Figure 2: Figure to illustrate the description of the conditional properties of the BED through the MDBED package. a) n pairs (x, y) generated by using the `rBED` function, b) and c) conditional PDF and CDF of one the values of x shown in c)(red point). The values of the conditional PDF and CDF can be obtained through the `CondBED` function. Adopted parameters for this example: ρ_{xy}=0.85, β_x=1, β_y=1, n=50

Duque., (2020). MDBED: An R package for simulating the Moran-Downton bivariate exponential distribution. *Journal of Open Source Software*, 2(), . https://doi.org/

is computed via numerical integration of $f(Y \mid X = x)$ by using the Riemann sum method (Hughes-Hallett et al., 2008). This function returns a list with a dataframe named Conditional_Statistics and a sublist named Condyx. Conditional_Statistics contains the values of x with its conditional moments, and Condyx contains several dataframes with the information of the conditional distributions associated with each value of x. For each conditional distribution, the values of the conditional quantiles (y_c) with its respective values of $f(Y \mid X = x)$ and $F(Y \mid X = x)$ are provided. The first dataframe of Condyx corresponds to the first value of x shown in Conditional_Statistics; the second dataframe corresponds to the second value of x shown in Conditional_Statistics, and so on.

This package also has a bivariate generator; the `rBED` function. The required inputs are ρ_{xy}, β_x, β_y, and the n number of bivariate random values to be generated. This function generates jointly exponential random values based on the conditional approach (Lewis & Orav, 2018). Thus, It first generates n random values of X; then, the conditional moments $E(Y \mid X = x)$ and $var(Y \mid X = x)$ associated with each value of x are computed to define $F(Y \mid X = x)$. Finally, the random values of Y are obtained by drawing a random value from each $F(Y \mid X = x)$. This function returns a dataframe with the n pairs (x, y) generated.

Figure 2 illustrates how the `MDBED` package can be used to explore the conditional properties of the BED. Figure 2a shows n pairs (x, y) generated through the `rBED` function, and Figure 2b and 2c show the conditional PDF and CDF, obtained by using the `CondBED` function, of one of these x values (red point). This package is available on GitHub under the GNU General Public License at https://github.com/lfduquey/MDBED and on the Comprehensive R Archive Network (CRAN) at https://cran.r-project.org/web/packages/MDBED/index.html.

References

Balakrishna, N., & Lai, C. D. (2009). *Continuous bivariate distributions* (Second.). New York: Springer Science & Business Media. doi:10.1007/b101765

Choulakian, V., El-Jabi, N., & Moussi, J. (1990). On the distribution of flood volume in partial duration series analysis of flood phenomena. *Stochastic Hydrology and Hydraulics.* doi:10.1007/BF01543085

Córdova, J. R., & Rodríguez-Iturbe, I. (1985). On the probabilistic structure of storm surface runoff. *Water Resources Research.* doi:10.1029/WR021i005p00755

Downton, F. (1970). Bivariate Exponential Distributions in Reliability Theory. *Journal of the Royal Statistical Society: Series B (Methodological).* doi:10.1111/j.2517-6161.1970.tb00852.x

Hughes-Hallett, D., Gleason, A. M., McCallum, W. G., Lomen, D. O., Lovelock, D., Tecosky-Feldman, J., Tucker, T. W., et al. (2008). *Calculus: Single Variable* (p. 709). USA: Wiley.

Lewis, P. W. A., & Orav, E. J. (2018). *Simulation Methodology for Statisticians, Operations Analysts, and Engineers (1988).* New York: CRC Press.

Nagao, M., & Kadoya, M. (1971). Two-variate Exponential Distribution and Its Numerical Table for Engineering Application. *Bulletin of the Disaster Prevention Research Institute, 20*(3), 34.

Nelsen, R. B. (2006). *An introduction to copulas* (p. 265). New York: Springer.

Yan, J. (2007). Enjoy the joy of copulas: With a package copula. *Journal of Statistical Software.* doi:10.18637/jss.v021.i04