UDC 519.2

Bivariate Teissier Distribution

N. Kolev*, **N. Ngo***, **Y. T. Ju***

* *Department of Statistics, University of Sao Paulo, Brazil*

Abstract. We first give historical remarks about the forgotten univariate Teissier model. We introduce a bivarite version of the Teissier distribution and outline its basic properties. The corresponding copula is obtained and applications are discussed.

Keywords: Bivariate mean residual life time vector, Copula, Failure rate, Lambert W function, Teissier distribution, Proportional mean residual life model, Simulation and data analysis.

1. Introduction and historical remarks

The baseline model under consideration has been first introduced by the French biologist Teissier (1934) considering a mortality of several domestic animal species protected from accidents and disease, i.e., dying as a result of a "pure aging". Specifically, for a non-negative random variable X, the model is defined by the survival function

$$S_X(x) = P(X > x) = \exp\{x + 1 - e^x\}, \quad x \geq 0.$$

It is direct to check that the corresponding mean residual life function (MRLF hereafter) defined by $m_X(x) = E[X - x \,|\, X > x]$ is given by e^{-x}.

Teissier's distribution is motivated by the empirical fact that many vital functions are decaying exponentially. In reliability terms, relation $m_X(x) = e^{-x}$ means that one should consider a scenario where the breakdown of an used item deteriorate "trough wear alone" adopting unit exponential law of depletion.

By using a simple transformation, Laurent (1975) obtained an alternative distribution with a parameter $c \in (0, 1]$, written as

$$S_X(x; c) = P(X > x) = \exp\left[cx - \frac{e^{cx} - 1}{c}\right], \quad x \geq 0. \qquad (1)$$

Later on, Rinne (1985) used model (1) to estimate lifetime distribution (with lifetime expressed in kilometers) for a German data set based on prices of used cars. The Teissier's distribution and its location version (1) have been forgotten after that and we did not find any further reference in available literature.

Although the survival function, its hazard rate and MRLF are in one-to-one correspondence with each other, Muth (1977) justified the MRLF

E-mail: kolev.ime@gmail.com, ngochvn@gmail.com.

to be a superior concept than the failure (mortality) rate. Postulating MRLF of the form

$$m_X(x) = \exp\{-cx\} \quad \text{for} \quad c \in (0,1] \quad \text{and all} \quad x \geq 0,$$

Muth (1977) introduced via his Model 3 a continuous probability distribution with survival function given exactly by relation (1).

Let $c \to 0$ in (1) to get the unit exponential distribution, the fact reported by Leemis and McQueston (2008). Muth (1977) outlined few properties of model (1) observing that it displays heavier tail in comparison than commonly used unimodal right-skewed distributions (gamma, lognormal and Weibull). Recently, Jodra et al. (2015) rediscovered the distribution specified by (1) and named it "Muth distribution". One can find careful analysis and additional properties expressed by Lambert W function: its moment generating function, corresponding mode-median-mean inequality, moments of order statistics, its quantile function and parameter estimation. Most of these characteristics have been listed by Laurent (1975) already, but in terms of Gamma function.

Jodra et al. (2015) noted that $E(X) = 1$ and classified this fact as a strong restriction if one wishes to use model (1) for real data analysis. The authors applied a scaling transform $Y = bX$ for some $b > 0$ of (1) yielding

$$S_Y(x; c, b) = S_X(bx; c) = \exp\left[\frac{cx}{b} - \frac{e^{cx/b} - 1}{c}\right], \quad x \geq 0.$$

The last expression can be found in Laurent (1975) as well, see his Model II.

Let us introduce a parametrization $\frac{c}{b} = \alpha$ and $c = \alpha\theta$ in the last relation to get the *scaled Teissier distribution*

$$S_Y(x; \alpha, \theta) = \exp\left[\alpha x - \tfrac{1}{\alpha\theta}(e^{\alpha x} - 1)\right], \quad x \geq 0 \qquad (2)$$

with parameters $\alpha > 0$ and $\theta > 0$ such that $\alpha\theta \leq 1$. The expression of the inverse function of (2) is given by

$$S_Y^{-1}(u) = \tfrac{1}{\alpha} \log\left[-\alpha\theta W_{-1}\left(\tfrac{-u}{\alpha\theta e^{1/\alpha\theta}}\right)\right], \quad u \in [0,1], \qquad (3)$$

where $W_{-1}(.)$ is the real negative branch of Lambert W function solving equation $W(z)\exp(W(z)) = z$ and taking values in $(-\infty, -1]$ for $z \in [-\exp^{-1}, 0)$, consult Corollary 2 and Proposition 5 in Jodra et al. (2015).

In Section 2 we generate a bivariate version of scaled Teissier distribution corresponding to (2) using the bivariate proportional mean residual life approach following Sreeja and Sankaran (2007). Several basic properties are presented and we obtain the corresponding survival copula. We finish with a brief discussion regarding usefulness of the bivariate Teissier model and related simulation study.

2. Construction of bivariate Teissier distribution

We first introduce necessary notations and the bivariate proportional mean residual life model. Its particular version is a base and our motivation to generate a bivariate analog of the scaled Teissier distribution (2).

To proceed, let X_i be non-negative continuous random variables representing remaining lifetimes with survival functions $S_{X_i}(x_i) = P(X_i > x_i), i = 1, 2$, and denote their joint survival function by $S_{X_1,X_2}(x_1, x_2) = P(X_1 > x_1, X_2 > x_2)$.

Let $m_i(x_i) = E[X_i - x_i \,|\, X_i > x_i]$ be the MRLF of X_i, which can be interpreted as the expected remaining gap time of X_i, given that X_i is larger than x_i, $i = 1, 2$. For the recurrent events, the occurrence of the second event depends on the occurrence of the first one. Thus, one can consider conditional MRLF $m_i(x_1, x_2)$ of X_i given $\{X_j > x_j\}$, $i, j = 1, 2$, $i \neq j$. For example,

$$m_2(x_1, x_2) = E[X_2 - x_2 \,|\, X_1 > x_1, X_2 > x_2] = \frac{\int_{x_2}^{\infty} S_{X_1,X_2}(x_1, u)du}{S_{X_1,X_2}(x_1, x_2)} \qquad (4)$$

for all $x_1, x_2 \geq 0$. By analogy, $m_1(x_1, x_2) = E[X_1 - x_1 \,|\, X_1 > x_1, X_2 > x_2]$. The bivariate MRLF vector $(m_1(x_1, x_2), m_2(x_1, x_2))$ uniquely determine the joint distribution of X_1 and X_2. In fact,

$$S_{X_1,X_2}(x_1, x_2) = \frac{m_1(0)}{m_1(x_1)} \frac{m_2(x_1, 0)}{m_2(x_1, x_2)} \exp\left[-\int_0^{x_1} \frac{du}{m_1(u)} - \int_0^{x_2} \frac{du}{m_2(x_1, u)}\right]. \qquad (5)$$

Following Sreeja and Sankaran (2007) we define a bivariate proportional mean residual life model for the vector (X_1, X_2) as

$$m_1(x_1; \theta_1) = \theta_1 m_{10}(x_1) \quad \text{and} \quad m_2(x_1, x_2; \theta_2) = \theta_2 m_{20}(x_1, x_2), \qquad (6)$$

for some appropriate positive parameters θ_1 and θ_2. In model (6), $m_1(x_1; \theta_1)$ is the MRLF at time x_1 where θ_1 is a given constant and $m_{10}(x_1)$ is a baseline MRLF. The interpretation of $m_2(x_1, x_2; \theta_2)$ is similar.

Using (5) and (6), we obtain

$$S_{X_1,X_2}(x_1, x_2) = \frac{m_{10}(0)}{m_{10}(x_1)} \frac{m_{20}(x_1, 0)}{m_{20}(x_1, x_2)} \exp\left[-\int_0^{x_1} \frac{du}{m_1(u; \theta_1)} - \int_0^{x_2} \frac{du}{m_2(x_1, u; \theta_2)}\right].$$

Apply (6) in the last formula for $\theta_1 = \theta_2 = \theta > 0$ and assume that the baseline MRLFs are specified by $m_{10}(x_1) = \exp\{-\alpha x_1\}$ and $m_{20}(x_1, x_2) = \exp\{-\alpha(x_1 + x_2)\}$ with $\alpha \in (0, 1]$, yielding

$$S_{X_1,X_2}(x_1, x_2) = \exp\left\{\alpha(x_1 + x_2) - \frac{1}{\theta\alpha}\left[e^{\alpha(x_1+x_2)} - 1\right]\right\}, \; x_1, x_2 \geq 0, \quad (7)$$

i.e., we got our *bivariate Teissier distribution.* Note that its marginal survival functions are identical with univariate Teissier distribution (2).

In order $S_{X_1,X_2}(x_1, x_2)$ given by (7) to be a proper joint survival function, the condition $\frac{\partial^2}{\partial x_2 \partial x_2} S_{X_1,X_2}(x_1, x_2) \geq 0$ must be satisfied, implying that $\theta\alpha \leq \frac{2}{3+\sqrt{5}}$. Thus, the parameter space Ω_1 of the bivariate Teissier distribution (7) is

$$\Omega_1 = \Omega_1(\theta, \alpha) = \left\{ \theta > 0 \quad \text{and} \quad \alpha > 0, \quad \text{such that} \quad \theta\alpha \leq \frac{2}{3+\sqrt{5}} \right\}.$$

One might conclude that the bivariate Teissier distribution (7) is *exchangeable*, i.e., having the same marginal distributions represented by (2). They can serve as a marginal distribution of (7), only if their parameter space is given by Ω_1.

Now, let us consider the ratio $\frac{S_{X_1,X_2}(x_1,x_2)}{S_{X_1}(x_1)S_{X_2}(x_2)}$, where $S_{X_1,X_2}(x_1, x_2)$ is given by (7) and marginal survival functions are of the form (2). A simple substitution implies

$$\frac{S_{X_1,X_2}(x_1,x_2)}{S_{X_1}(x_1)S_{X_2}(x_2)} = \exp\left\{-(\alpha\theta)^{-1}[e^{\alpha x_1} - 1][e^{\alpha x_2} - 1]\right\}, \quad x_1, x_2 \geq 0.$$

The right hand side in the last equation is no larger than 1, i.e., the bivariate Teissier model (7) is *negative quadrant dependent.*

The expressions of the correlation coefficient of bivariate Teissier distribution and corresponding survival copula are given in the following two statements.

Lemma 1. *Let the random vector (X_1, X_2) follows the bivariate Teissier distribution (7). Its correlation coefficient can be expressed as*

$$Corr(X_1, X_2) = \frac{\exp\left(\frac{1}{\alpha\theta}\right) E_1\left(\frac{1}{\alpha\theta}\right) - \alpha\theta}{2\exp\left(\frac{1}{\alpha\theta}\right) E_1\left(\frac{1}{\alpha\theta}\right) - \alpha\theta},$$

where $E_1(.)$ is the exponential integral.

Lemma 2. *The Teissier's survuval copula function corresponding to (7) is given by*

$$C(u, v) = uv \exp\left\{-a^{-1}\left[aW_{-1}\left(\frac{-u}{ae^{1/a}}\right) + 1\right]\left[aW_{-1}\left(\frac{-v}{ae^{1/a}}\right) + 1\right]\right\}$$

for $u, v \in [0, 1]$, where $W_{-1}(.)$ is the real negative branch of Lambert W function and $a = \alpha\theta$.

3. Discussion

In many survival studies, each subject can potentially experience a series of events, which may be repetitions of essentially the same event or may be events of entirely different natures. Such outcomes have been termed recurrent events. To analyze recurrent event data, the focus can be placed on two types of time scale: the time since entering the study and the time since the last event (gap time). For the situation where the time since study entry is of interest, a variety of statistical methods have been proposed in literature.

The bivariate Teissier model (7) is a new contribution to the bivariate distribution theory and can serve as an alternative to the existing once for analysis of recurrent events which typically occur in Insurance, Finance, Reliability, Engineering, Medicine, etc. We will present a simulation study and will compare our analysis of real bivariate data set studied by other authors as well.

References

1. *Jodra P., Jimenez-Gamero M., Alba-Fernandez M.* On the Muth Distribution // Mathematical Modelling and Analysis. — 2015. — Vol. 20. — P. 291–310.
2. *Leemis L., McQueston, J.* Univariate distribution relationships // American Statistician. — 2008. — Vol. 62. — P. 45–53.
3. *Laurent A.* Failure and mortality from wear and aging. The Teissier model // Statistical Distributions in Scientific Work - Model Building and Model Selection. Vol. 2 (Patil, G., Kotz, S. and Ord, H., eds.). — 1975. — P. 301-320.
4. *Muth J.* Reliability models with positive memory derived from the mean residual life function // The Theory and Applications of Reliability. Vol. 2, (Tsokos, C. and Shimi, I., eds.). — Academic Press Inc., New York, 1977. — P. 401–435.
5. Rinne H. (1985). Estimating the lifetime distribution of private motorcars using prices of used cars: The Teissier model // Statistiks Zwischen Theorie und Praxis (Buttler, G. et al., eds.). — Gottingen (in German), 1985. — P. 172–184.
6. *Sankaran P., Sreeja V.* Proportional hazards model for multivariate failure time data // Communications in Statistics. Theory and Methods. — 2007. — Vol. 36. — P. 1627–1642.
7. *Teissier G.* Recherches sur le vieillissement et sur les lois de mortalite // Annales de Physiologie et de Physicochimie Biologique. — 1934. — Vol. 10. — P. 237–284.