

The Transmuted Weibull Regression Model: an Application to Type 2 Diabetes Mellitus Data

Daniele C. T. Granzotto¹, Carlos A. Santos² & Francisco Louzada³

^{1,2} Department of Statistics, State University of Maringá

³ Department of Applied Mathematics & Statistics, ICMC, University of São Paulo

Correspondence: Daniele C. T. Granzotto, Department of Statistics, State University of Maringá.

Received: November 19, 2017 Accepted: December 14, 2017 Online Published: January 8, 2018

doi:10.5539/ijsp.v7n2p1 URL: <https://doi.org/10.5539/ijsp.v7n2p1>

Abstract

In this paper, we are considering the transmuted Weibull distribution to investigate the role of oxidative stress markers as indicators of risk of injury of the lower extremities in patients with type 2 diabetes mellitus. A group of patients was studied until medical discharge assessing the following variables: glycated hemoglobin and three blood markers of oxidative status. In face of explicative covariates, we are considering the regression approach of the transmuted model to fit this real dataset. The inference was considered by using the method of maximum likelihood and the consistency of the estimators were verified by a Monte Carlo simulation study presented in this manuscript. Some properties as the moment generation function, median and the behaviors of the hazard and survival functions, are also included in this study.

Keywords: Regression approach, Weibull distribution, transmutation map and reliability analysis

1. Introduction

Generally transmutation maps are a convenient way of constructing new distributions, in particular, survival ones. According to Shaw and Buckley (2007) transmutation maps comprise the functional composition of the cumulative distribution function of one distribution with the inverse cumulative distribution (quantile) function of another one. After that, some studies involving quadratic rank transmutation map can be seen in other application areas such as survival analysis and reliability. Aryal and Tsokos (2009) proposed a generalization of the extreme value distribution by using the quadratic rank transmutation map and applied such new distribution to analyzing the snow fall data in Midway Airport in the state of Illinois, USA. Further, Aryal and Tsokos (2011) proposed the transmuted Weibull distribution as a generalization of the Weibull probability distribution and the usefulness was illustrated using two published data sets. Granzotto and Louzada, (2015); Louzada and Granzotto, (2016) proposed the transmuted log-logistic model and the regression one in the study of the time until the first calve of a polled Tabapua race. Interested readers can refer to Ghitany (2001), Marshall and Olkin (2007), Sarabia and Prieto (2009) and Lai (2013), and their references, whose discusses some common approaches for construct lifetime distributions.

Although several studies involving quadratic rank transmutation map can be seen in any areas such as survival analysis and reliability, limited are the numbers of papers that provide a regression approach. In this paper we proposed a regression approach for the transmuted Weibull model, called transmuted Weibull regression model, to investigate the role of oxidative stress markers as indicators of risk of injury of the lower extremities in patients with type 2 diabetes mellitus. For that, the sections are organized as follows. Section 2 we presented a background of the transmuted model including the way of the model was constructed, the main properties and the regression approach showing the reliability functions and their main properties. The likelihood function and estimation methods are be showed in Section 3. The numerical experiments are being showed in Section 4 where we present a simulation study by using the Monte Carlo method and the usefulness of the model is showed in the oxidative stress markers data, used as an indicator of risk of injury. Finally, some conclusions and remarks are presented in Section 6 following references and appendix sections.

2. Transmuted Weibull Regression Distribution

2.1 Background

By using the quadratic rank transmutation map, Aryal and Tsokos (2011) proposed a new generalization of the Weibull model called transmuted Weibull distribution. The author derive the Weibull distribution using the quadratic rank transmutation map studied by Shaw and Buckley (2007). For that, let F_1 and F_2 be the cumulative distribution functions, of two distributions with a common sample space. The general rank transmutation as given in Shaw and Buckley (2007) is defined as

$$G_{R12}(u) = F_2(F_1^{-1}(u)) \text{ and } G_{R21}(u) = F_1(F_2^{-1}(u)).$$

The functions $G_{R12}(u)$ and $G_{R21}(u)$ both map the unit interval $I = [0, 1]$ into itself, and under suitable assumptions are mutual inverses and they satisfy $G_{Rij}(0) = 0$ and $G_{Rij}(1) = 1$. A Quadratic Rank Transmutation Map is defined as

$$G_{R12}(u) = u + \lambda u(1 - u), |\lambda| \leq 1, \tag{1}$$

from which it follows that the cdf's satisfy the relationship

$$F_2(x) = (1 + \lambda)F_1(x) - \lambda F_1(x)^2, \tag{2}$$

which on differentiation yields,

$$f_2(x) = f_1(x) [(1 + \lambda) - 2\lambda F_1(x)], \tag{3}$$

where $f_1(x)$ and $f_2(x)$ are the corresponding pdfs associated with cdf $F_1(x)$ and $F_2(x)$ respectively (observe that at $\lambda = 0$ we have the distribution of the base random variable).

In equation (2), consider $F_1(x)$ the cumulative distribution of the baseline model Weibull and $F_2(x)$, which will be called just $F(x)$, the generalization of this distribution or transmuted model.

Hence, the cdf and pdf of a transmuted generalized Weibull distribution are given respectively by

$$F(x) = \left[1 - \exp \left[- \left(\frac{x}{\mu} \right)^\beta \right] \right] \left[1 + \lambda \exp \left[- \left(\frac{x}{\mu} \right)^\beta \right] \right]$$

and

$$f(x) = \frac{\beta}{\mu} \left(\frac{x}{\mu} \right)^{\beta-1} \exp \left[- \left(\frac{x}{\mu} \right)^\beta \right] \left[1 - \lambda + 2\lambda \exp \left[- \left(\frac{x}{\mu} \right)^\beta \right] \right]. \tag{4}$$

Note that the transmuted Weibull distribution is an extended model to analyze more complex data and it generalizes some of the widely used distributions. Figure 1 illustrates some of the possible shapes of the transmuted Weibull pdf and cdf for selected values of β and λ and for $\mu = 1$.

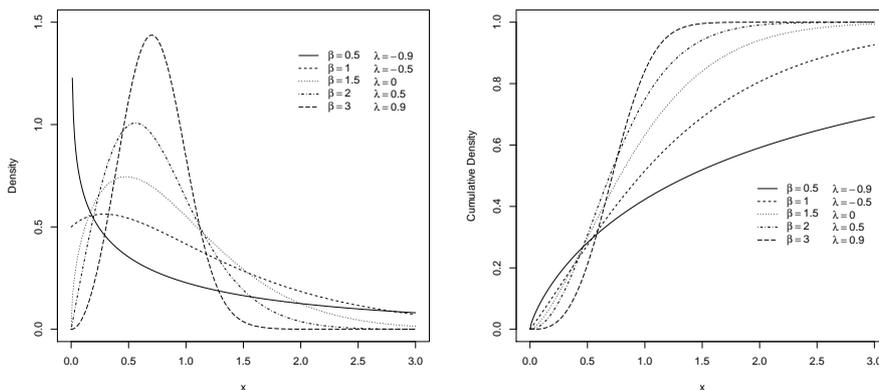


Figure 1. The pdf and cdf curves of transmuted Weibull distribution for $\mu = 1$ and different values of β and λ .

2.2 The Regression Model

Now, let T be a random variable denoting the lifetimes with a probability distribution and $\gamma(\mathbf{x})$ as a parameter depending on a covariate vector $\mathbf{x} = (1, x_1, \dots, x_p)'$ given by

$$\gamma(\mathbf{x}) = \gamma_0 + \gamma_1 x_1 + \dots + \gamma_p x_p$$

Then, the hazard function of the transmuted Weibull regression model $h(t | \gamma(\mathbf{x}), \beta, \lambda)$ which denote the hazard function at time t for an individual with covariate vector \mathbf{x} is given by

$$h(t | \gamma(\mathbf{x}), \beta, \lambda) = \frac{\beta t^{\beta-1}}{\gamma(\mathbf{x})^\beta} \left\{ \frac{1 - \lambda + 2\lambda \exp \left[- \left(\frac{t}{\gamma(\mathbf{x})} \right)^\beta \right]}{1 - \lambda + \lambda \exp \left[- \left(\frac{t}{\gamma(\mathbf{x})} \right)^\beta \right]} \right\} \tag{5}$$

where $\lambda > 0, \beta > 0$ and $\gamma(\mathbf{x})$ is a regression defined above. Consequently, the reliability function may be written as

$$R(t | \gamma(\mathbf{x}), \beta, \lambda) = \exp \left[- \left(\frac{t}{\gamma(\mathbf{x})} \right)^\beta \right] \left[1 - \lambda + \lambda \exp \left[- \left(\frac{t}{\gamma(\mathbf{x})} \right)^\beta \right] \right] \tag{6}$$

Some examples of behaviors of the hazard function are presented in Figure 2, right panel.

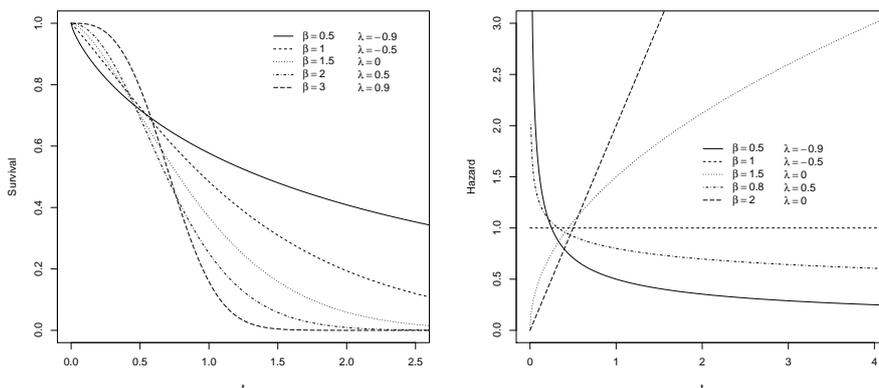


Figure 2. Survival and hazard plot of transmuted Weibull distribution for $\mu = 1$ and different values of β and λ .

2.3 Some Important Statistics

In this section we present some important statistics as moments and quantiles.

If X has the Transmuted Weibull distribution with $\mu, \beta > 0$ and $|\lambda| \leq 1$, then the the moment generating function of X is given as follows.

$$M_X(t) = \int_0^\infty e^{tx} f(x|\mu, \beta) dx,$$

then,

$$M_X(t) = \int_0^\infty e^{tx} \frac{\beta}{\mu} \left(\frac{x}{\mu} \right)^{\beta-1} \exp \left[- \left(\frac{x}{\mu} \right)^\beta \right] \left[1 - \lambda + 2\lambda \exp \left[- \left(\frac{x}{\mu} \right)^\beta \right] \right] dx$$

Substituting $\omega = \mu t^\beta$ and using Taylor expansion $e^x = \sum_{i=0}^\infty \frac{x^i}{i!}$, the final function is expressed as

$$M_X(t) = \sum_{i=0}^\infty \frac{t^i}{i!} \mu^{-i/\beta} \Gamma \left(1 + \frac{1}{\beta} \right) \left[1 - \lambda + \lambda 2^{-i/\beta} \right]. \tag{7}$$

By using the expression (7), the expected value and variance of the model are given, respectively, by

$$E(X) = \mu \Gamma \left(\frac{\beta + 1}{\beta} \right) \left[1 - \lambda + 2^{-(1/\beta)} \lambda \right] \tag{8}$$

and

$$V(X) = \mu^2 \left\{ \Gamma \left(\frac{\beta + 2}{\beta} \right) \left[1 - \lambda + 2^{-(2/\beta)} \lambda \right] - \Gamma^2 \left(\frac{\beta + 1}{\beta} \right) \left[1 - \lambda + 2^{-(1/\beta)} \lambda \right]^2 \right\}. \tag{9}$$

The quantile q^{th} quantile is given by

$$x_q = \mu \left\{ - \ln \left[1 - \left(\frac{1 + \lambda - \sqrt{(1 + \lambda)^2 - 4\lambda q}}{2\lambda} \right) \right] \right\}^{1/\beta} \tag{10}$$

3. Inference

The maximum likelihood estimates, MLEs, of the parameters that are inherent within the transmuted log-logistic probability distribution function is given by the following: Let y_1, y_2, \dots, y_n be a sample of size n from a transmuted log-logistic distribution; $\mathbf{x}_i = (1, x_{1i}, x_{2i}, \dots, x_{pi})$ be an i^{th} vector of covariates, $i = 1, \dots, n$. Also, consider the following relationship between the vector of covariates and the parameters $\gamma(\mathbf{x}) = \gamma_0 + \gamma_1\mathbf{x}_1 + \dots + \gamma_p\mathbf{x}_p$.

Then the likelihood function is given by

$$L(\gamma_0, \dots, \gamma_p, \beta, \lambda | \mathbf{y}, \mathbf{x}) = \exp \left[- \sum_{i=1}^n \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right)^\beta \right] \prod_{i=1}^n \left\{ \frac{\beta y_i^{\beta-1}}{\left(\sum_{j=1}^p \gamma_j x_{ij} \right)^\beta} \left[1 - \lambda + 2\lambda \exp \left(- \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right)^\beta \right) \right] \right\} \tag{11}$$

Hence, the log-likelihood function $l(\gamma_0, \dots, \gamma_p, \beta, \lambda | \mathbf{y}, \mathbf{x}) = \ln L(\gamma_0, \dots, \gamma_p, \beta, \lambda | \mathbf{y}, \mathbf{x})$ becomes

$$l(\gamma_0, \dots, \gamma_p, \beta, \lambda | \mathbf{y}, \mathbf{x}) = - \sum_{i=1}^n \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right)^\beta + n \log \beta + (\beta - 1) \sum_{i=1}^n \log y_i - \beta \sum_{i=1}^n \log \left(\sum_{j=1}^p \gamma_j x_{ij} \right) + \sum_{i=1}^n \log \left\{ 1 - \lambda + 2\lambda \exp \left[- \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right)^\beta \right] \right\} \tag{12}$$

Therefore, the maximum likelihood estimates of γ_j , λ and β which maximize (12) must satisfy the following normal equations

$$\begin{aligned} \frac{\partial l}{\partial \gamma_j} &= \sum_{i=1}^n \left\{ \frac{2\lambda \beta \left(\sum_{j=1}^p x_{ij} \right) \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right)^\beta \exp \left[- \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right)^\beta \right]}{\left(\sum_{j=1}^p \gamma_j x_{ij} \right) \left[1 - \lambda + 2\lambda \exp \left[- \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right)^\beta \right] \right]} \right\} \\ &+ \sum_{i=1}^n \left[\frac{\beta \left(\sum_{j=1}^p x_{ij} \right) \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right)^\beta}{\sum_{j=1}^p \gamma_j x_{ij}} \right] - \beta \sum_{i=1}^n \sum_{j=1}^p \frac{x_{ij}}{\gamma_j x_{ij}} = 0 \\ \frac{\partial l}{\partial \lambda} &= \sum_{i=1}^n \left[\frac{2 \exp \left[- \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right)^\beta \right] - 1}{1 - \lambda + 2\lambda \exp \left[- \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right)^\beta \right]} \right] = 0 \\ \frac{\partial l}{\partial \beta} &= \frac{n}{\beta} + \sum_{i=1}^n \log y_i - \sum_{i=1}^n \log \left(\sum_{j=1}^p \gamma_j x_{ij} \right) - \sum_{i=1}^n \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right)^\beta \log \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right) \\ &- \sum_{i=1}^n \left\{ \frac{2\lambda \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right)^\beta \log \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right) \exp \left[- \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right)^\beta \right]}{1 - \lambda + 2\lambda \exp \left[- \left(\frac{y_i}{\sum_{j=1}^p \gamma_j x_{ij}} \right)^\beta \right]} \right\} = 0 \end{aligned}$$

The maximum likelihood estimator $\hat{\theta} = (\hat{\gamma}_0, \dots, \hat{\gamma}_p, \hat{\beta}, \hat{\lambda})'$ is obtained by solving the above nonlinear system of equations, Lawless (2003); Marshall and Olkin (2007). It is usually more convenient to use nonlinear optimization algorithms such as quasi-Newton or Newton-Raphson algorithms to numerically maximize the log-likelihood function given in (12). In order to compute the standard error and asymptotic confidence interval we use the usual large sample approximation in which the maximum likelihood estimators of θ can be treated as being approximately $(p + 3)$ variate normal. Hence as

$n \rightarrow \infty$ the asymptotic distribution of the MLE $(\hat{\gamma}_0, \dots, \hat{\gamma}_p, \hat{\beta}, \hat{\lambda})$ is given by,

$$\begin{pmatrix} \hat{\gamma}_0 \\ \hat{\gamma}_1 \\ \vdots \\ \hat{\gamma}_p \\ \hat{\beta} \\ \hat{\lambda} \end{pmatrix} \sim N \left(\begin{pmatrix} \hat{\gamma}_0 \\ \hat{\gamma}_1 \\ \vdots \\ \hat{\gamma}_p \\ \hat{\beta} \\ \hat{\lambda} \end{pmatrix}, \begin{pmatrix} \hat{V}_{11} & \hat{V}_{12} & \dots & \hat{V}_{1(p+3)} \\ \hat{V}_{21} & \hat{V}_{22} & \dots & \hat{V}_{2(p+3)} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{V}_{(p+3)1} & \hat{V}_{(p+3)2} & \dots & \hat{V}_{(p+3)(p+3)} \end{pmatrix} \right) \tag{13}$$

where, $\hat{V}_{ij} = V_{ij} |_{\theta=\hat{\theta}}$ and it is determined by the inverse of Hessian matrix.

Thereby, an approximate $100(1 - \alpha)\%$ two sided confidence intervals for γ_j, β and λ are, respectively, given by

$$\hat{\gamma}_j \pm z_{\alpha/2} \sqrt{\hat{V}_{(j+1)(j+1)}}, \quad \hat{\beta} \pm z_{\alpha/2} \sqrt{\hat{V}_{(p+2)(p+2)}} \quad \text{and} \quad \hat{\lambda} \pm z_{\alpha/2} \sqrt{\hat{V}_{(p+3)(p+3)}}$$

where $j = 0, \dots, p, z_{\alpha}$ is the upper $\alpha - th$ percentiles of the standard normal distribution.

Different models can be compared by penalizing over-fitting by using the Akaike information criterion (Akaike, 1973), which intends to minimize the Kullback-Leibler divergence between the true distribution and the estimate from a candidate model. It is given by $AIC = -2l(\hat{\theta}) + 2 \text{size}(\theta)$, where $l(\hat{\theta})$ denotes the log likelihood function evaluated at the maximum and $\text{size}(\theta)$ is the number of model parameters. The model with the lowest value of each of this criterion (among all considered models) is regarded as the preferred model for describing the given dataset.

The Hessian matrix is showed in Appendix A.

4. Numerical Experiments

4.1 Simulation

This section presents the results of a Monte Carlo experiment on the finite sample behavior of the MLEs. All results were obtained from 1,000 Monte Carlo replications. The sample sizes n range from 30 to 500, generated by using the inversion method according to a transmuted Weibull regression distribution for each combination of the parameter values (β, λ) and fixed values of $\alpha_0 = -1, \alpha_1 = -2$. Table 4.1 shows the estimates of parameters and their variances and Table 4.1 shows that the mean square error (mse) become smaller when the sample size increases beyond the coverage probability to the confidence interval converge to 95% as it was generated.

Table 1. MLEs of the parameters and variance estimated for each combination of the parameters.

| Sample Size | Generated | | Parameters | | | | Variance | | | |
|-------------|-----------|-----------|------------------|------------------|---------------|-----------------|---------------------|---------------------|------------------|--------------------|
| | β | λ | $\hat{\alpha}_0$ | $\hat{\alpha}_1$ | $\hat{\beta}$ | $\hat{\lambda}$ | $V(\hat{\alpha}_0)$ | $V(\hat{\alpha}_1)$ | $V(\hat{\beta})$ | $V(\hat{\lambda})$ |
| 30 | 0.5 | -0.6 | -1.1352 | -2.1858 | 0.5140 | -0.5138 | 1.0175 | 1.1751 | 0.1310 | 0.2930 |
| | 1.0 | -0.5 | -1.0476 | -2.0800 | 1.0413 | -0.4378 | 0.4660 | 0.5607 | 0.2433 | 0.3102 |
| | 1.5 | -0.4 | -1.0387 | -2.0607 | 1.5770 | -0.3568 | 0.3331 | 0.3877 | 0.3346 | 0.3266 |
| | 0.5 | 0.6 | -1.0609 | -2.1653 | 0.5302 | 0.4889 | 1.1265 | 1.3358 | 0.0840 | 0.2987 |
| | 1.0 | 0.5 | -1.0423 | -2.0912 | 1.0610 | 0.4169 | 0.5558 | 0.6647 | 0.1670 | 0.3240 |
| | 1.5 | 0.4 | -1.0343 | -2.0653 | 1.5934 | 0.3390 | 0.3647 | 0.4374 | 0.2556 | 0.3425 |
| 50 | 0.5 | -0.6 | -1.0465 | -2.0803 | 0.5183 | -0.5568 | 0.5901 | 0.7392 | 0.0764 | 0.2375 |
| | 1.0 | -0.5 | -1.0300 | -2.0483 | 1.0382 | -0.4767 | 0.3248 | 0.3996 | 0.1493 | 0.2587 |
| | 1.5 | -0.4 | -1.0192 | -2.0315 | 1.5596 | -0.3884 | 0.2099 | 0.2561 | 0.2135 | 0.2732 |
| | 0.5 | 0.6 | -1.0494 | -2.0951 | 0.5166 | 0.5478 | 0.7905 | 0.9436 | 0.0607 | 0.2483 |
| | 1.0 | 0.5 | -1.0333 | -2.0545 | 1.0343 | 0.4650 | 0.3878 | 0.4688 | 0.1208 | 0.2694 |
| | 1.5 | 0.4 | -1.0261 | -2.0397 | 1.5532 | 0.3745 | 0.2532 | 0.3089 | 0.1817 | 0.2779 |
| 100 | 0.5 | -0.6 | -1.0122 | -2.0275 | 0.5115 | -0.5915 | 0.4126 | 0.4829 | 0.0453 | 0.1843 |
| | 1.0 | -0.5 | -1.0077 | -2.0159 | 1.0231 | -0.4976 | 0.2112 | 0.2490 | 0.0901 | 0.1957 |
| | 1.5 | -0.4 | -1.0066 | -2.0123 | 1.5340 | -0.4002 | 0.1448 | 0.1710 | 0.1346 | 0.2017 |
| | 0.5 | 0.6 | -1.0260 | -2.0445 | 0.5092 | 0.5907 | 0.5386 | 0.6336 | 0.0414 | 0.1874 |
| | 1.0 | 0.5 | -1.0164 | -2.0251 | 1.0191 | 0.4954 | 0.2660 | 0.3143 | 0.0832 | 0.2004 |
| | 1.5 | 0.4 | -1.0123 | -2.0181 | 1.5295 | 0.3945 | 0.1744 | 0.2069 | 0.1256 | 0.2055 |

| | | | | | | | | | | |
|-----|-----|------|---------|---------|--------|---------|--------|--------|--------|--------|
| 200 | 0.5 | -0.6 | -0.9987 | -2.0088 | 0.5066 | -0.5952 | 0.2823 | 0.3309 | 0.0301 | 0.1325 |
| | 1.0 | -0.5 | -0.9998 | -2.0053 | 1.0131 | -0.4952 | 0.1457 | 0.1708 | 0.0601 | 0.1382 |
| | 1.5 | -0.4 | -1.0002 | -2.0042 | 1.5194 | -0.3952 | 0.1002 | 0.1175 | 0.0899 | 0.1427 |
| | 0.5 | 0.6 | -1.0041 | -2.0154 | 0.5048 | 0.6046 | 0.3702 | 0.4351 | 0.0280 | 0.1346 |
| | 1.0 | 0.5 | -1.0034 | -2.0093 | 1.0101 | 0.5042 | 0.1837 | 0.2162 | 0.0558 | 0.1406 |
| | 1.5 | 0.4 | -1.0027 | -2.0069 | 1.5158 | 0.4035 | 0.1212 | 0.1427 | 0.0840 | 0.1445 |
| 500 | 0.5 | -0.6 | -0.9956 | -2.0024 | 0.5027 | -0.5973 | 0.1757 | 0.2058 | 0.0183 | 0.0817 |
| | 1.0 | -0.5 | -0.9979 | -2.0015 | 1.0053 | -0.4972 | 0.0908 | 0.1062 | 0.0368 | 0.0856 |
| | 1.5 | -0.4 | -0.9987 | -2.0013 | 1.5079 | -0.3972 | 0.0625 | 0.0730 | 0.0553 | 0.0885 |
| | 0.5 | 0.6 | -0.9988 | -2.0068 | 0.5019 | 0.6018 | 0.2322 | 0.2690 | 0.0176 | 0.0825 |
| | 1.0 | 0.5 | -0.9998 | -2.0038 | 1.0041 | 0.5016 | 0.1157 | 0.1337 | 0.0351 | 0.0864 |
| | 1.5 | 0.4 | -1.0000 | -2.0027 | 1.5064 | 0.4014 | 0.0763 | 0.0881 | 0.0528 | 0.0892 |

Table 2. Mean square error (mse) and coverage probability of the MLEs.

| Sample Size | Generated | | Mean Square Error | | | | Coverage Probability 95% | | | |
|-------------|-----------|-----------|-------------------|------------|---------|-----------|--------------------------|------------|---------|-----------|
| | β | λ | α_0 | α_1 | β | λ | α_0 | α_1 | β | λ |
| 30 | 0.5 | -0.6 | 1.036 | 1.210 | 0.131 | 0.300 | 0.861 | 0.872 | 0.874 | 0.884 |
| | 1.0 | -0.5 | 0.468 | 0.567 | 0.245 | 0.314 | 0.877 | 0.891 | 0.895 | 0.896 |
| | 1.5 | -0.4 | 0.335 | 0.391 | 0.341 | 0.328 | 0.884 | 0.896 | 0.911 | 0.903 |
| | 0.5 | 0.6 | 1.130 | 1.363 | 0.085 | 0.311 | 0.869 | 0.888 | 0.940 | 0.817 |
| | 1.0 | 0.5 | 0.558 | 0.673 | 0.171 | 0.331 | 0.880 | 0.889 | 0.943 | 0.852 |
| | 1.5 | 0.4 | 0.366 | 0.442 | 0.264 | 0.346 | 0.887 | 0.894 | 0.942 | 0.865 |
| 50 | 0.5 | -0.6 | 0.592 | 0.746 | 0.077 | 0.239 | 0.923 | 0.932 | 0.934 | 0.913 |
| | 1.0 | -0.5 | 0.326 | 0.402 | 0.151 | 0.259 | 0.924 | 0.932 | 0.934 | 0.909 |
| | 1.5 | -0.4 | 0.210 | 0.257 | 0.217 | 0.273 | 0.928 | 0.935 | 0.936 | 0.914 |
| | 0.5 | 0.6 | 0.793 | 0.953 | 0.061 | 0.251 | 0.908 | 0.921 | 0.945 | 0.879 |
| | 1.0 | 0.5 | 0.389 | 0.472 | 0.122 | 0.271 | 0.920 | 0.927 | 0.946 | 0.898 |
| | 1.5 | 0.4 | 0.254 | 0.310 | 0.185 | 0.279 | 0.920 | 0.928 | 0.945 | 0.907 |
| 100 | 0.5 | -0.6 | 0.413 | 0.484 | 0.045 | 0.184 | 0.922 | 0.944 | 0.929 | 0.927 |
| | 1.0 | -0.5 | 0.211 | 0.249 | 0.091 | 0.196 | 0.925 | 0.947 | 0.926 | 0.931 |
| | 1.5 | -0.4 | 0.145 | 0.171 | 0.136 | 0.202 | 0.927 | 0.948 | 0.926 | 0.942 |
| | 0.5 | 0.6 | 0.539 | 0.636 | 0.041 | 0.187 | 0.927 | 0.941 | 0.945 | 0.911 |
| | 1.0 | 0.5 | 0.266 | 0.315 | 0.084 | 0.200 | 0.929 | 0.944 | 0.939 | 0.917 |
| | 1.5 | 0.4 | 0.175 | 0.207 | 0.126 | 0.206 | 0.928 | 0.950 | 0.936 | 0.923 |
| 200 | 0.5 | -0.6 | 0.282 | 0.331 | 0.030 | 0.133 | 0.939 | 0.954 | 0.950 | 0.935 |
| | 1.0 | -0.5 | 0.146 | 0.171 | 0.060 | 0.138 | 0.939 | 0.953 | 0.950 | 0.938 |
| | 1.5 | -0.4 | 0.100 | 0.117 | 0.090 | 0.143 | 0.943 | 0.953 | 0.946 | 0.940 |
| | 0.5 | 0.6 | 0.370 | 0.435 | 0.028 | 0.135 | 0.944 | 0.954 | 0.957 | 0.923 |
| | 1.0 | 0.5 | 0.184 | 0.216 | 0.056 | 0.141 | 0.946 | 0.955 | 0.955 | 0.929 |
| | 1.5 | 0.4 | 0.121 | 0.143 | 0.084 | 0.144 | 0.949 | 0.956 | 0.950 | 0.931 |
| 500 | 0.5 | -0.6 | 0.176 | 0.206 | 0.018 | 0.082 | 0.936 | 0.952 | 0.942 | 0.948 |
| | 1.0 | -0.5 | 0.091 | 0.106 | 0.037 | 0.086 | 0.940 | 0.952 | 0.942 | 0.949 |
| | 1.5 | -0.4 | 0.062 | 0.073 | 0.055 | 0.088 | 0.937 | 0.953 | 0.943 | 0.946 |
| | 0.5 | 0.6 | 0.232 | 0.269 | 0.018 | 0.083 | 0.943 | 0.948 | 0.944 | 0.947 |
| | 1.0 | 0.5 | 0.116 | 0.134 | 0.035 | 0.086 | 0.943 | 0.950 | 0.941 | 0.950 |
| | 1.5 | 0.4 | 0.076 | 0.088 | 0.053 | 0.089 | 0.943 | 0.950 | 0.944 | 0.950 |

Also, in Figure 3 we can clearly see that the mean square error (mse) become smaller when the sample size increases for all parameters ($\alpha_0, \alpha_1, \beta, \lambda$).

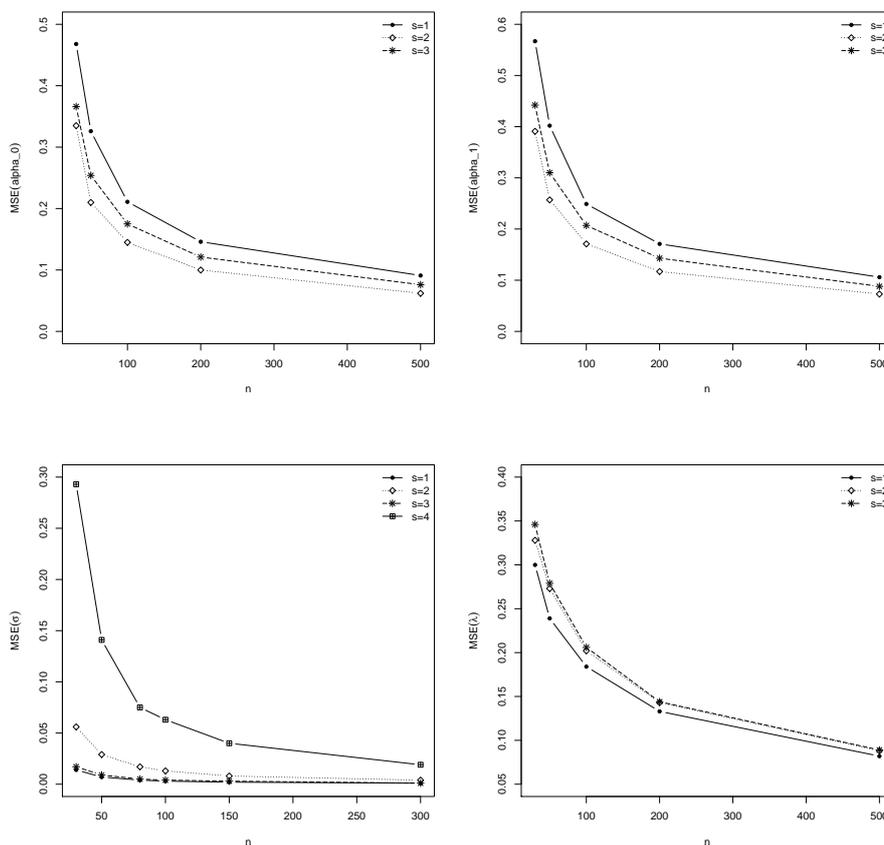


Figure 3. MSE from the estimators for different samples.

5. Application

Oxidative stress, caused by an imbalance between production of reactive oxygen species and endogenous antioxidant capacity is a major determinant of diabetic complications. In order to illustrate the usefulness of the transmuted Weibull regression model, we investigated the role of oxidative stress markers as indicators of risk of injury of the lower extremities in patients with type 2 diabetes mellitus (PPDMT2). For that, three groups of 29 patients was studied until medical discharge and we assessed the following parameters: glycated hemoglobin (HbA1c); three blood markers of oxidative status (total antioxidant capacity - TAC, reduced proteins thiol groups and reduced levels of protein carbonyls). Its important to emphasize that we are interested in study only the first group which one 28 patients actually initiated the treatment, i.e., one patient withdrew before the tests have been done.

Firstly, a brief descriptive analyze is made. The minimum observed time was about half month; the maximum observed was 40 months. The median time observed was 12.5 months and even the first and third quantiles are, 5 and 16.25 months, respectively. Figures 4, left and right panels, show the TTTPlot and the histogram of the observed times. In the first one we can observe that indicate increasing hazard, which can be fitted by a transmuted Weibull model.

After the initial analysis, the models transmuted Weibull, Weibull and exponential regression were fitted. The estimates of the parameters, standard errors and confidence intervals are be showed in Table 3. Note that, we considere $\gamma(\mathbf{x}) = \gamma_1\text{Protein} + \gamma_2\text{Albumin} + \gamma_3\text{GJ} + \gamma_4\text{HbA1C}$ (the albumin and protein are measure in mg per mol).

We observed the $-2 \log$ Likelihood estimated values for all fitted models. By using this criteria, we observe 52% in favor of transmuted model against 47% in favor of Weibull and 1% in favor of exponential model. Although there is no large difference between the transmuted and the Weibull model, we will consider the setting and description of transmuted Weibull. Also, the proximity between the parametric and nonparametric survival curves can be seem in the Figure 5, left panel (we are estimating the scale parameter by setting the covariates in the mean point).

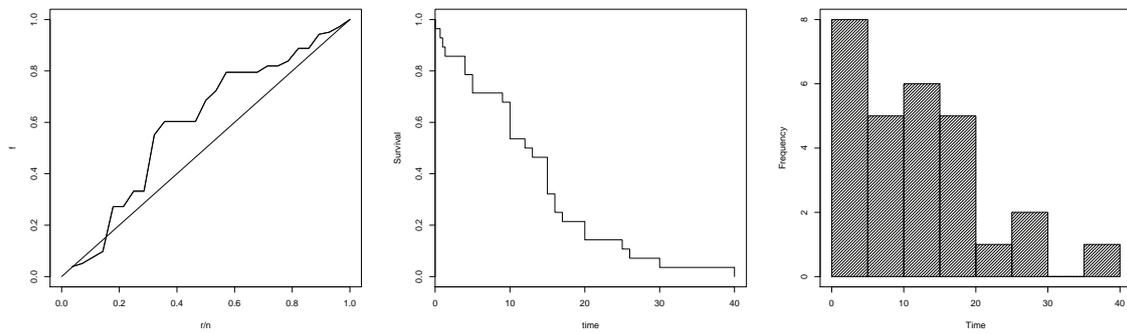


Figure 4. TTTPlot, survival curve and histogram of the observed times, respectively.

Table 3. Estimates of transmuted Weibull regression, Weibull regression and exponential regression parameters.

| Model | Parameter | Estimate | Standard Error | P-Value | Confidence Interval (95%) | |
|-------------|------------|----------|----------------|---------|---------------------------|--------|
| | | | | | Lower | Upper |
| Transmuted | γ_1 | -6.160 | 1.594 | 0.001 | -9.425 | -2.896 |
| | γ_2 | 9.272 | 2.992 | 0.004 | 3.144 | 15.400 |
| | γ_3 | -0.249 | 0.063 | 0.001 | -0.378 | -0.119 |
| | γ_4 | 7.258 | 1.894 | 0.001 | 3.378 | 11.139 |
| | λ | -0.254 | 0.243 | 0.044 | -1.367 | -0.085 |
| | β | 1.499 | 0.323 | < .0001 | 0.839 | 2.160 |
| Weibull | γ_1 | -6.707 | 1.151 | < .0001 | -9.065 | -4.350 |
| | γ_2 | 9.984 | 2.738 | 0.001 | 4.376 | 15.592 |
| | γ_3 | -0.270 | 0.046 | < .0001 | -0.364 | -0.175 |
| | γ_4 | 7.926 | 1.354 | < .0001 | 5.152 | 10.700 |
| | β | 1.584 | 0.251 | < .0001 | 1.071 | 2.098 |
| Exponential | γ_1 | -5.862 | 1.583 | 0.001 | -9.105 | -2.619 |
| | γ_2 | 8.149 | 3.849 | 0.043 | 0.265 | 16.033 |
| | γ_3 | -0.232 | 0.063 | 0.001 | -0.360 | -0.104 |
| | γ_4 | 7.119 | 1.947 | 0.001 | 3.131 | 11.108 |

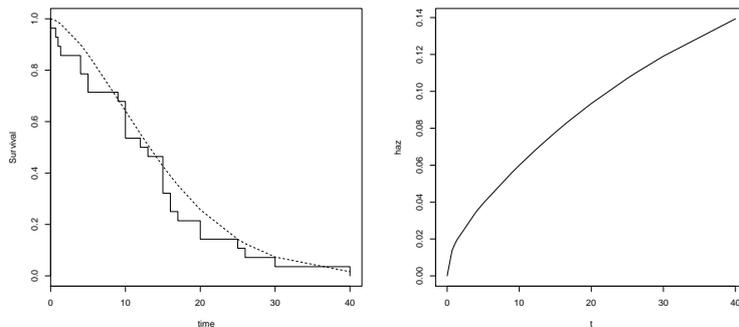


Figure 5. The transmuted Weibull, Weibull and exponential estimated survival curves versus the Kaplan-Meier, left panel, and the estimated hazard curve, right panel.

Also, the estimate of β is coherent with the TTT-Plot that indicates an increasing hazard (β is larger than 1). Figure 5, right panel, presents the estimated increasing hazard. Figure 6 presents the perspective curves of the estimated hazards for each covariates protein, albumin, GJ and HbA1C, ranging from the minimum to the maximum observed values, respectively, and the others covariates fixed at the mean observed value.

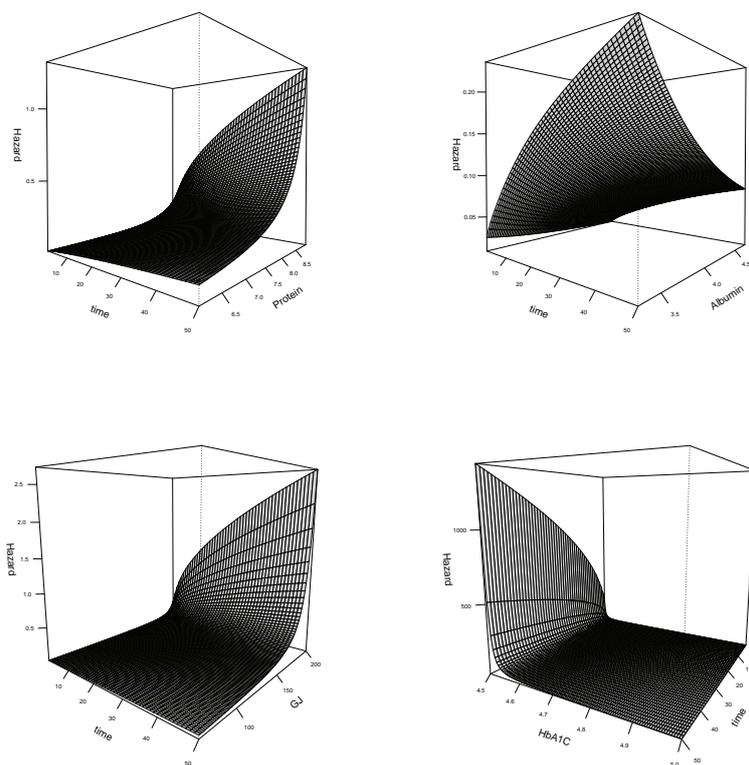


Figure 6. Perspective curves of the estimated hazards for each covariates protein, albumin, GJ and HbA1C, ranging from the minimum to the maximum observed values, respectively, and the others covariates fixed at the mean observed value.

The perspective plots are showing the behaviors of the hazard rate and, in all cases, we can clearly see the increasing behavior of the curves with respect all covariates, except for the HbA1C that present a decreasing hazard behavior. However, we can clearly see that if the covariate albumin increase in value, the hazard rate curves is being dominated by the others which is the opposite that occurs with the covariates protein and GJ.

Furthermore, some estimates of the fitted model were done. The estimated mean and respective confidence interval of 95% of the discharge time of the patients are, respectively, 16.40 and 9.20; 23.60 months. The estimated median and the first and third quantiles are given by: 14.62, 8.43 and 22.43 months, respectively.

6. Concluding Remarks

In this paper we presented the transmuted Weibull regression model in order to investigate the role of oxidative stress markers as indicators of risk of injury of the lower extremities in patients with type 2 diabetes mellitus. For that, we reviewed the transmuted Weibull model adding covariates on that model. The numerical experiment showed us the increasing behavior of the hazard curve and the shapes of the hazard for plenty of values of the covariates.

Also, some new properties as the moment generation function were presented for the transmuted Weibull model. The consistence of the model was showed in a simulation study by applying the Monte Carlos method for sample sizes range from 30 to 500 and different combinations of the parameters values.

References

Akaike, H. (1973). Information theory and the maximum likelihood principle. *International Symposium on Information Theory*, eds. V. Petrov and F.

Aryal, G. R., & Tsokos, C. P. (2009). On the transmuted extreme value distribution with applications. *Nonlinear Analysis*, 71, 1401C1407.

Aryal, G. R., & Tsokos, C. P. (2011). Transmuted Weibull distribution: a generalization of the Weibull probability distribution. *European Journal of Pure and Applied Mathematics*, 4(2), 89C102.

Ghitany, M. E. (2001). A compound rayleigh survival model and its application to randomly censored data. *Statistical Paper, Springer*, 42, 437C450.

Granzotto, D. C. T., & Louzada, F. (2015). The transmuted log-logistic distribution: Modeling, inference, and an application to a polled tabapua race time up to first calving data. *Communications in Statistics-Theory and Methods*, 44(16), 3387C3402.

Lai, C. D. (2013). Constructions and applications of lifetime distributions. *Applied Stochastic Models in Business and Industry*, pages 127C140.

Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data*. John Wiley and Sons, Hoboken.

Louzada, F., & Granzotto, D. C. (2016). The transmuted log-logistic regression model: a new model for time up to first calving of cows. *Statistical Papers*, 57(3), 623C640.

Marshall, A., & Olkin, I. (2007). *Life distributions: structure of nonparametric, semiparametric and parametric families*. Springer, New York.

Sarabia, J. M., & Prieto, F. (2009). The Pareto-positive stable distribution: a new descriptive method for city size data. *Physica A: Statistical Mechanics and its Applications*, 388(19), 4179C4191.

Shaw, W. T., & Buckley, I. R. C. (2007). The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *UCL Discovery Repository*, pages 1C16.

A. Hessian Matrix

The Hessian matrix is given by

$$A = \begin{pmatrix} A_{11} & \dots & A_{1(p+3)} \\ A_{21} & \dots & A_{2(p+3)} \\ \vdots & \ddots & \vdots \\ A_{(p+3)1} & \dots & A_{(p+3)(p+3)} \end{pmatrix}$$

where

$$\begin{pmatrix} \hat{V}_{11} & \dots & \hat{V}_{13} \\ \hat{V}_{21} & \dots & \hat{V}_{23} \\ \vdots & \ddots & \vdots \\ \hat{V}_{(p+3)1} & \dots & \hat{V}_{(p+3)(p+3)} \end{pmatrix} = \begin{pmatrix} A_{11} & \dots & A_{1(p+3)} \\ A_{21} & \dots & A_{2(p+3)} \\ \vdots & \ddots & \vdots \\ A_{(p+3)1} & \dots & A_{(p+3)(p+3)} \end{pmatrix}^{-1}$$

and

$$A_{(j+1)(j+1)} = -\frac{\partial^2 l}{\partial \gamma_j^2} = \sum_{i=1}^n \left[\frac{2\lambda \beta C_{i,j}^\beta \exp(-C_{i,j}^\beta)}{\left(\sum_{j=1}^p \gamma_j x_{i,j}\right)^2 (1 - \lambda + 2\lambda \exp(-C_{i,j}^\beta))} \right] \times$$

$$\left[C_{i,j}^\beta \left(1 - \frac{2\lambda}{1 - \lambda + 2\lambda \exp(-C_{i,j}^\beta)} \right) - (1 + \beta) \right] + \beta \sum_{i=1}^n \sum_{j=1}^p \left(\frac{x_{i,j}}{\gamma_j x_{i,j}} \right)^2$$

$$- \sum_{i=1}^n \frac{C_{i,j}^\beta \beta \left(\sum_{j=1}^p x_{i,j}\right)^2}{\left(\sum_{j=1}^p \gamma_j x_{i,j}\right)^2} (\beta + 1)$$

$$\begin{aligned}
 A_{(p+2)(p+2)} &= -\frac{\partial^2 l}{\partial \beta^2} = -\frac{n}{\beta^2} - \sum_{i=1}^n (C_{i,j}^\beta \log C_{i,j}^2) \\
 &\quad + \sum_{i=1}^n \frac{2\lambda C_{i,j}^\beta \log C_{i,j}^2 \exp(-C_{i,j}^\beta)}{1 - \lambda + 2\lambda \exp(-C_{i,j}^\beta)} \left[C_{i,j}^\beta - 1 - \frac{2\lambda C_{i,j}^\beta \exp(-C_{i,j}^\beta)}{1 - \lambda + 2\lambda \exp(-C_{i,j}^\beta)} \right] \\
 A_{(p+3)(p+3)} &= -\frac{\partial^2 l}{\partial \lambda^2} = -\sum_{i=1}^n \left[\frac{2 \exp(-C_{i,j}^\beta) - 1}{1 - \lambda + 2\lambda \exp(-C_{i,j}^\beta)} \right]^2
 \end{aligned}$$

where $C_{ij} = \frac{y_i}{\sum_{j=1}^p \gamma_j x_{i,j}}$ and $A_{l \times m}$ ($l \neq m$) are the partial derivatives.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).