

## The geometric power half-normal regression model with cure rate

Yolanda M. Gómez\*<sup>†</sup> and Heleno Bolfarine<sup>‡</sup>

### Abstract

In this paper we consider the geometric cure rate model defined in [16], using for  $S_0(\cdot)$ , the survival function of carcinogenic cells, an extension of the half-normal distribution based on the distribution of the maximum of a random sample. The implementation of maximum likelihood estimation for the model parameters is discussed and, finally, the model is fitted to a real database (Melanoma data set), and comparisons are performed with alternatives to the new  $S_0(\cdot)$ .

**Keywords:** Cure rate survival model, Half-normal distribution, Maximum likelihood, Power distribution.

*2000 AMS Classification:* 62N01, 62N02

*Received :* 01.04.2016 *Accepted :* 22.06.2016 *Doi :* 10.15672/HJMS.201613820026

### 1. Introduction

To generate alternative families of distributions one can use the power of an absolutely continuous distribution, as initiated by [9], where the famous Lehmann's alternatives are proposed. Given a cumulative distribution function,  $\mathcal{F}$ , this family of models can be written as

$$(1.1) \quad \mathcal{F}_F(z; \alpha) = \{F(z)\}^\alpha, \quad z \in \mathbb{R}.$$

Lehmann has considered  $\alpha$  as a rational number. In the case of  $\alpha$  integer, this distribution is the distribution of the maximum of a random sample. Durrans extends the distribution for the case of  $\alpha \in \mathbb{R}$  in a hydrological context (see [6]), calling it the fractional order distribution in which case the density function can be written as

---

\*Departamento de Matemática, Facultad de Ingeniería, Universidad de Atacama, Copiapó, Chile, Email: [yolanda.gomez@uda.cl](mailto:yolanda.gomez@uda.cl)

<sup>†</sup>Corresponding Author.

<sup>‡</sup>Departamento de Estatística, IME, Universidad de Sao Paulo, Sao Paulo, Brasil, Email: [hbolfar@ime.usp.br](mailto:hbolfar@ime.usp.br)

$$(1.2) \quad \varphi_F(z; \alpha) = \alpha f(z) \{F(z)\}^{\alpha-1}, \quad z \in \mathbb{R}^+.$$

where  $\alpha \in \mathbb{R}^+$  is called a resilience parameter. The case where  $F = \Phi$ , was originally studied in [6] with further results in [8] and [14], where its Fisher information matrix was derived and shown to be nonsingular. [14] call it the *Power distribution* and in the case of  $F = \Phi$ , the Power-Normal distribution. The notation  $Z \sim P_F(\alpha)$  means that a random variable  $Z$  has a power distribution. In the equation (1.2),  $\alpha$  is called a resilience parameter family (see [11]). Considering in 1.2 a Half-Normal(HN) distribution function, we obtain the Power Half-Normal(PHN) model, studied in [7].

Let  $T \sim PHN(\sigma, \alpha)$ . Then, the survival function of T is given by

$$(1.3) \quad S_{PHN}(t; \sigma, \alpha) = 1 - \left( 2\Phi\left(\frac{t}{\sigma}\right) - 1 \right)^\alpha, \quad t > 0,$$

where  $\sigma > 0$  is a scale parameter and  $\alpha > 0$  is a resilience parameter.

Cumulative damage such as chronic heart disease and several different types of cancer can be caused in individuals by several unknown causes or risk factors. This degradation leads to a fatigue process, the propagation life time of which may be suitably shaped by the PHN distribution. To model such phenomenon, [4] proposed the geometric cure rate Birnbaum-Saunders regression model. The main of the present article is to propose a new distribution, the Geometric Power Half-Normal cure rate model (GPHNcr), designed in a scenario of latent causes with fraction of healing and where there is no information about which causes were responsible for the individual's death or reappearance of the tumour. For instance, in clinical studies, part of the population can respond favourably to the treatment, being considered cured. The proportion of such fraction of the population which is not susceptible to the event of interest is termed *cure fraction*. Distributions which accommodate cured fraction have been widely developed. Perhaps, the most popular type of cure rate model is the mixture of distributions introduced by [3] and [2]. In this distribution, it is assumed that a certain proportion of the patients, say  $p$ , are cured, in the sense that they do not present the event of interest during a long period of the time and can be seen to be the immune or cured to the cause of death under study (see [12]). The key reference on the mixture distribution approach is [10]. To the best of our knowledge, there is no literature considering the mixture PHN (MPHN) as a distribution with survival function given by

$$(1.4) \quad S_{MPHN}(t) = p_0 + (1 - p_0)S_{PHN}(t), \quad t > 0,$$

where  $S_{PHN}(t)$  is given by (1.3) and  $p_0$  is the cured fraction.

The sections of this paper are organized in the following manner. In Section 2, we explain the model formulation and give some of its main properties. In Section 3, we develop parameters estimation for the model via maximum likelihood. In Section 4, we perform a simulation study to determine when the maximum likelihood estimators (MLE) perform well. In Section 5, we show a real data application where inference is made with the MLE. Finally, some conclusions are given in Section 6.

## 2. Model formulation

The distribution GPHNcr is derived as follows. For an individual picked at random in the population, let  $M$  be the unobserved number of causes (risk factors) causing the

event of interest. Suppose that  $M$  follows a geometric distribution with parameter  $\theta$  and probability function

$$(2.1) \quad P(M = m) = (1 - \theta)\theta^m, \quad m = 0, 1, \dots$$

Consider that the time for the  $j$ -th cause to produce the event of interest is indicated by  $Z_j$ ,  $j = 1, \dots, M$ . We assume that, conditional on  $M$ , the  $Z_j$ 's are independent and identically distributed (i.i.d.) with PHN distribution given in (1.3). Moreover, we assume that  $Z_1, Z_2, \dots$  are independent of  $M$ .

The observed time to event is defined by the random variables  $T = \min(Z_1, \dots, Z_M)$ , and  $T = \infty$  if  $M = 0$  with  $P(T = \infty | M = 0) = 1$ . Under this configuration, the population survivor function is given next.

**1. Proposition.** Under then above formulation, the population survivor function is given by

$$(2.2) \quad S_{GPHNcr}(t) = \frac{1 - \theta}{1 - \theta[1 - (2\Phi(\frac{t}{\sigma}) - 1)^\alpha]},$$

where  $\Phi(\cdot)$  is the standard normal distribution.

**Proof.** The proof follows directly from [16], and using the survivor function in (1.3).  $\square$

The cure fraction is given by  $S_{GPHNcr}(\theta) = 1 - \theta$ , indicating that this not a proper survivor function. The corresponding density function is given by

$$(2.3) \quad f_{GPHNcr}(t) = \theta(1 - \theta)f_{PHN}(t) \left\{ 1 - \theta \left[ 1 - \left( 2\Phi\left(\frac{t}{\sigma}\right) - 1 \right)^\alpha \right] \right\}^{-2},$$

where  $f_{PHN}(t)$  is the pdf of the PHN distribution. The population risk function is given by

$$(2.4) \quad h_{GPHNcr}(t) = \frac{\theta f_{PHN}(t)}{1 - \theta [1 - (2\Phi(\frac{t}{\sigma}) - 1)^\alpha]},$$

The distribution in (1.4) can be written as a mixture distribution (see [2]). Therefore,

$$(2.5) \quad S_{GPHNcr}(t) = (1 - \theta) + \theta \left\{ \frac{(1 - \theta) [1 - (2\Phi(\frac{t}{\sigma}) - 1)^\alpha]}{1 - \theta [1 - (2\Phi(\frac{t}{\sigma}) - 1)^\alpha]} \right\}.$$

The survivor function for the noncured population, for the GPHN model is given by

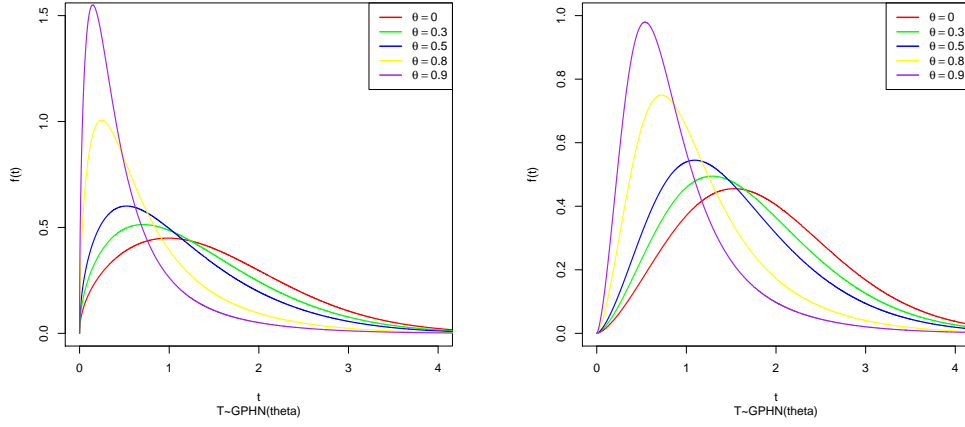
$$(2.6) \quad S_{GPHN}(t) = P(T > t | M \geq 1) = \frac{(1 - \theta) [1 - (2\Phi(\frac{t}{\sigma}) - 1)^\alpha]}{1 - \theta [1 - (2\Phi(\frac{t}{\sigma}) - 1)^\alpha]}, \quad t > 0.$$

Notice that  $S_{GPHN}(0) = 1$  and  $S_{GPHN}(\infty) = 0$ , so that (2.6) is a proper survivor function.

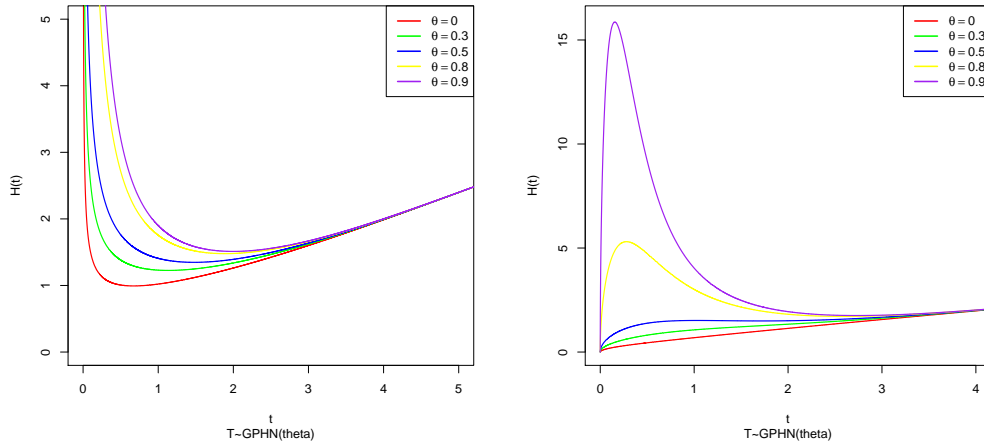
The pdf for the GPHN distribution is given by

$$(2.7) \quad f_{GPHN}(t) = \frac{(1 - \theta)2\alpha\phi(\frac{t}{\sigma})(2\Phi(\frac{t}{\sigma}) - 1)^{\alpha-1}}{\sigma(1 - \theta [1 - (2\Phi(\frac{t}{\sigma}) - 1)^\alpha])^2}, \quad t > 0.$$

Figure 1 shows the GPHN probability density functions for some fixed values of  $\theta$ . The plots in these figures show that the GPHN distribution is flexible and that the value of  $\theta$  has a substantial effect on its skewness and kurtosis, as we shall observe further in Figure (3).



**Figure 1.** Plots for the pdf of the GPHN distribution for  $\sigma = 1.5, \alpha = 1.5$ (left),  $\alpha = 2.5$ (right)



**Figure 2.** Plots for the risk function for the GPHN distribution for  $\sigma = 1.5, \alpha = 0.5$ (left),  $\alpha = 1.5$ (right)

Moreover, from (2.6) and (2.7) it is easy to verify that the risk function for the noncured population is given by

$$(2.8) \quad h_{GPHN}(t) = \frac{h_{PHN}(t)}{(1 - \theta [1 - (2\Phi(\frac{t}{\sigma}) - 1)^\alpha])^2}, \quad t > 0,$$

where  $h_{PHN}(t)$  is the risk function for the PHN distribution.

**2.1. Moments.** Moments of the GPHN model can be computed numerically using the routine “integrate” from the software R (see [15]). The following proposition presents the  $r$ -th moment of a random variable following the GPHN distribution.

**2. Proposition.** The  $r$ -th moment of the random variable  $T \sim GPHN(\sigma, \alpha, \theta)$ , is given by

$$\mu_r = E(T^r) = \alpha \sigma^r (1 - \theta) \kappa_r(\alpha, \theta), \quad r = 1, 2, \dots,$$

where  $\kappa_r(\alpha, \theta) = \int_0^1 \frac{(\Phi^{-1}(\frac{1+u}{2}))^r u^{\alpha-1}}{(1-\theta(1-u)^\alpha)^2} du$  is computed numerically.

**Proof.** The moment definition implies

$$\mu_r = E(T^r) = \int_0^\infty t^r \frac{(1 - \theta) 2\alpha \phi(\frac{t}{\sigma}) (2\Phi(\frac{t}{\sigma}) - 1)^{\alpha-1}}{\sigma (1 - \theta [1 - (2\Phi(\frac{t}{\sigma}) - 1)^\alpha])^2} dt.$$

The result follows after making the variable change  $u = 2\Phi(\frac{t}{\sigma}) - 1$ .  $\square$

Therefore, the first four moments are given by

- (1)  $\mu_1 = E(T) = \alpha(1 - \theta)\sigma\kappa_1(\alpha, \theta)$
- (2)  $\mu_2 = E(T^2) = \alpha(1 - \theta)\sigma^2\kappa_2(\alpha, \theta)$
- (3)  $\mu_3 = E(T^3) = \alpha(1 - \theta)\sigma^3\kappa_3(\alpha, \theta)$
- (4)  $\mu_4 = E(T^4) = \alpha(1 - \theta)\sigma^4\kappa_4(\alpha, \theta)$

**1. Corollary.** The asymmetry and kurtosis coefficients are given, respectively, by

$$\sqrt{\beta_1} = \frac{\kappa_3(\alpha, \theta) - 3\alpha(1 - \theta)\kappa_1(\alpha, \theta)\kappa_2(\alpha, \theta) + 2\alpha^2(1 - \theta)^2\kappa_1^3(\alpha, \theta)}{\sqrt{\alpha(1 - \theta)(\kappa_2(\alpha, \theta) - \alpha(1 - \theta)\kappa_1^2(\alpha, \theta))^{3/2}}},$$

and

$$\beta_2 = \frac{\kappa_4(\alpha, \theta) - 4\alpha(1 - \theta)\kappa_1(\alpha, \theta)\kappa_3(\alpha, \theta) + 6\alpha^2(1 - \theta)^2\kappa_1(\alpha, \theta)^2\kappa_2(\alpha, \theta) - 3\alpha^3(1 - \theta)^3\kappa_1^4(\alpha, \theta)}{\alpha(1 - \theta)(\kappa_2(\alpha, \theta) - \alpha(1 - \theta)\kappa_1^2(\alpha, \theta))^2}.$$

**Table 1.** Moments for some combinations of parameters of the GPHN distribution.

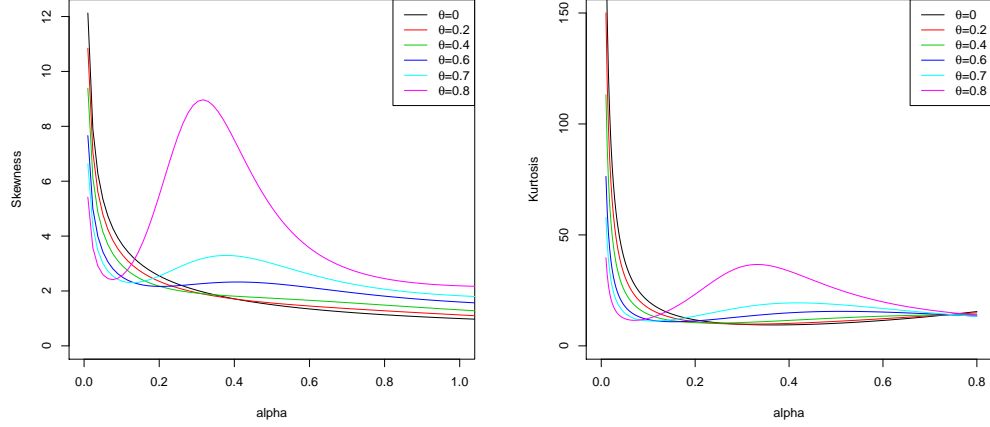
	$\theta = 0$				$\theta = 0.5$				$\theta = 0.8$			
	$\sigma = 1$		$\sigma = 2$		$\sigma = 1$		$\sigma = 2$		$\sigma = 1$		$\sigma = 2$	
	$\alpha = 1$	$\alpha = 2$	$\alpha = 1$	$\alpha = 2$	$\alpha = 1$	$\alpha = 2$	$\alpha = 1$	$\alpha = 2$	$\alpha = 1$	$\alpha = 2$	$\alpha = 1$	$\alpha = 2$
$\mu_1$	0.798	1.128	1.596	2.257	0.586	0.624	1.172	1.247	0.364	0.274	0.729	0.548
$\mu_2$	1	1.637	4	6.547	0.627	0.859	2.507	3.435	0.313	0.357	1.252	1.429
$\mu_3$	1.596	2.821	12.766	22.568	0.916	1.444	7.327	11.559	0.413	0.588	3.304	4.706
$\mu_4$	3	5.547	48	88.744	1.636	2.808	26.169	44.924	0.701	1.133	11.211	18.124
$\sqrt{\beta_1}$	0.995	0.704	0.995	0.704	1.436	1.004	1.436	1.004	2.188	2.242	2.188	2.242
$\beta_2$	3.869	3.435	3.869	3.435	5.311	3.414	5.311	3.414	9.079	7.943	9.079	7.943

From Table 1, we note that for  $\alpha = \sigma = 1$  and  $\theta = 0$  the coefficients of skewness and kurtosis coincide with the coefficients of the HN distribution,  $\sqrt{\beta_1} = 0.995$  and  $\beta_2 = 3.869$ .

### 3. Inference

Letting  $C_i$  be the  $i$ -th censoring time, we observe that  $Y_i = \min\{T_i, C_i\}$  and  $\delta_i = I(T_i \leq C_i)$ ,  $i = 1, \dots, n$ . Let  $\gamma$  the parameter vector for the distribution of time to event  $Z$  in (3). Given a sample of  $n$  pairs  $(y_1, \delta_1), \dots, (y_n, \delta_n)$  the corresponding likelihood function under uninformative censoring is given by

$$(3.1) \quad L(\gamma, \theta) \propto \prod_{i=1}^n \{f_{GPHN_{cr}}(y_i; \gamma, \theta)\}^{\delta_i} \{S_{GPHN_{cr}}(y_i; \gamma, \theta)\}^{1-\delta_i},$$



**Figure 3.** Plots for the asymmetry coefficient (left panel) and the kurtosis coefficient (right panel) of the GPHN model.

where  $S_{GPHNcr}(y_i; \gamma, \theta)$  and  $f_{GPHNcr}(y_i; \gamma, \theta)$  are given in (2.2) and (2.3), respectively. Hereafter, denote  $1 - \theta = p_0$ . To complete specification of the model, the cure fraction is related to the covariates  $\mathbf{x}_i$  by considering the link function  $\log(\frac{p_{0i}}{1-p_{0i}}) = \mathbf{x}_i^T \boldsymbol{\beta}$ , that is,

$$p_{0i} = \frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta})}{1 + \exp(\mathbf{x}_i^T \boldsymbol{\beta})}, \quad i = 1, \dots, n,$$

where  $\boldsymbol{\beta}$  represents the regression coefficient parameter vector.

We remember that covariates are traditionally used to model the expectation of the number of competing causes. For instance, in the proposed model, from (2.1), we have

$$E(M_i) = \frac{\theta_i}{(1 - \theta_i)} = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$$

and  $p_{0i} = 1 - \theta_i$ , so that

$$(3.2) \quad p_{0i} = \frac{1}{(1 + \exp(\mathbf{x}_i^T \boldsymbol{\beta}))}.$$

The connection between the cured fraction and the covariates is much more cumbersome in this expression than in the logistic link.

Thus, the improper functions given in (2.3) and (2.6) can be written as

$$(3.3) \quad S_{GPHNcr}(y_i; \gamma, \boldsymbol{\beta}) = \{1 + (p_{0i}^{-1} - 1)F_{PHN}(y_i; \gamma)\}^{-1}, \quad y_i > 0$$

and

$$(3.4) \quad f_{GPHNcr}(y_i; \gamma, \boldsymbol{\beta}) = \{1 + (p_{0i}^{-1} - 1)F_{PHN}(y_i; \gamma)\}^{-2} \{p_{0i}^{-1} - 1\} f_{PHN}(y_i; \gamma), \quad y_i > 0.$$

The likelihood function given in (3.1) can be expressed as

$$L(\theta, \mathfrak{D}) \propto \prod_{i=1}^n \{(p_{0i}^{-1} - 1) f_{PHN}(y_i; \gamma)\}^{\delta_i} \{1 + (p_{0i}^{-1} - 1) F_{PHN}(y_i; \gamma)\}^{-\delta_i - 1},$$

where  $\theta = (\boldsymbol{\beta}^T, \boldsymbol{\gamma}^T)^T$ ,  $\mathfrak{D} = (y, \delta, X)$ , and  $X = (x_1^T, \dots, x_n^T)$ .

The maximum likelihood estimation of the parameter vector  $\theta$  is carried out by direct numerical maximization of the log-likelihood function  $l(\theta; \mathfrak{D}) = \log(L(\theta; \mathfrak{D}))$ , which is accomplished by using existing software (see [15]). The computational program is available from the authors upon request. Under suitable regularity conditions, it can be shown that the asymptotic distribution of the maximum likelihood estimator  $\hat{\theta}$  is multivariate normal with mean vector  $\theta$  and covariance matrix  $\Sigma(\hat{\theta})$ , which can be estimated by

$$\hat{\Sigma}(\hat{\theta}) = \left\{ -\frac{\partial^2 l(\theta; \mathfrak{D})}{\partial \theta \partial \theta^T} \right\}^{-1}, \quad \text{evaluated at } \theta = \hat{\theta}.$$

The required second derivatives are computed numerically.

To compare model fits, we used the Akaike criterion (see [1]), namely

$$AIC = -2 \log [L(\hat{\theta})] + 2k,$$

where  $k$  is the dimension of  $\theta$  which is the vector of parameters of the model being considered. We considered also the BIC (see [17]), namely

$$BIC = -2 \ln [L(\hat{\theta})] + k [\log(n)].$$

The best model is the one with the smallest AIC (BIC).

#### 4. Simulation study

To evaluate the performance of the parameter estimation procedure for the proposed models, we conducted a simulation study. In this study we considered the proposed model with the PHN distribution for the event times ( $Z$ ) with parameter  $\alpha = 2$  and  $\sigma = 1$ . For the  $i$ -th individual, the number of causes of the event of interest, ( $M_i$ ), is generated from the Geometric distribution with parameter,

$$1 - \theta_i = p_{0i} = \exp(\beta_0 + \beta_1 x_i) / (1 + \exp(\beta_0 + \beta_1 x_i)), \quad i = 1, \dots, n.$$

In our simulations we consider a binary covariate  $x$  with values drawn from a Bernoulli distribution with parameter 0.5. We took  $\beta_0 = 0.5$  and  $\beta_1 = -1$  so that the cured fraction for the two levels of  $x$  are  $p_0^{(0)} = 0.62$  and  $p_0^{(1)} = 0.38$  respectively.

We took the sample sizes to be  $n = 50, 100, 200, 400$  and  $800$ . For each set up, we conducted 1000 simulations and calculated the average of the maximum likelihood estimates (MLEs) of the cured fraction ( $p_0^{(0)}$  and  $p_0^{(1)}$ ), standard deviation (SD) of MLEs and the square root of mean square errors (SRMSE) of the MLEs. The simulation results are shown in Table 2 for simulated data from the GPHN cure model. We can observe that the average of MLEs are closed to the true parameter values, with the SDs and SRMSEs decreasing as sample size increases, suggesting the consistency of the estimates, as expected.

**Table 2.** Averages of maximum likelihood estimates (MLEs), standard deviation (SD) and square root of mean square error (RMSE) of cure fraction  $p_0^{(0)}$  and  $p_0^{(1)}$  for simulated data from PHN cure rate model.

$n$	Average of MLE		SD of MLE		SRMSE	
	$p_0^{(0)}$	$p_0^{(1)}$	$p_0^{(0)}$	$p_0^{(1)}$	$p_0^{(0)}$	$p_0^{(1)}$
50	0.6169	0.3727	0.0974	0.0935	0.0088	0.0095
100	0.6230	0.3779	0.0673	0.0664	0.0045	0.0044
200	0.6233	0.3765	0.0483	0.0463	0.0024	0.0022
400	0.6210	0.3778	0.0334	0.0326	0.0011	0.0011
800	0.6223	0.3778	0.0245	0.0234	0.0006	0.0006

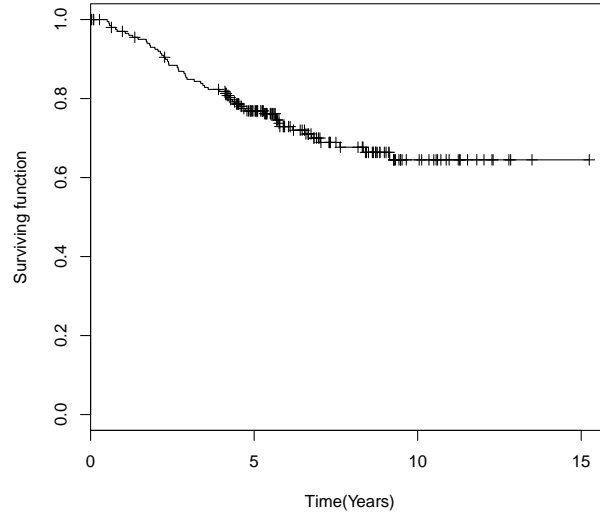


Figure 4. Kaplan-Meier estimate of the survival of function.

## 5. Illustration

In this section we work out an example employing the modelling presented in Section 2. The data set includes 205 patients observed after surgery for the removal of malignant melanoma in a follow-up period of 15 years. These data are available in the `timereg` package in R (Scheike, 2009). The observed time ( $T$ ) ranges from 10 to 5565 days (from 0.0274 to 15.25 years, with mean=5.9 and standard deviation=3.1 years) and refers to the time until the patient's death or the censoring time. Dead patients from other causes, as well as patients still alive at the end of the study are censored observations (72%). We take ulceration status ( $x_1$ ) (absent,  $n=115$ ; present,  $n=90$ ) and tumour thickness ( $x_2$ ) (in mm., mean=2.92 and standard deviation=2.96) as covariates. The Kaplan-Meier estimate of the survivor function is given in Figure (4). The presence of a plateau above 0.6 indicates that models that ignore the possibility of cure will not be suitable for these data.

We then fit the GPHNcr distribution with  $p_{0i}$  as in (3.2), as well as the MPH distribution (1.4) with  $p_{0i}$  as in (3.2). The maximum likelihood estimates (MLEs) of the model parameters are given in the Table 3. In the same Table also is presented the AIC and BIC selection criteria on the two candidate models. According to both criteria, the GPHNcr model stands out as the best one and then is chosen to be our working model. The QQ plot of the normalized randomized quantile residuals, in Figure (5), suggests that the GPHNcr model yields an acceptable fit.

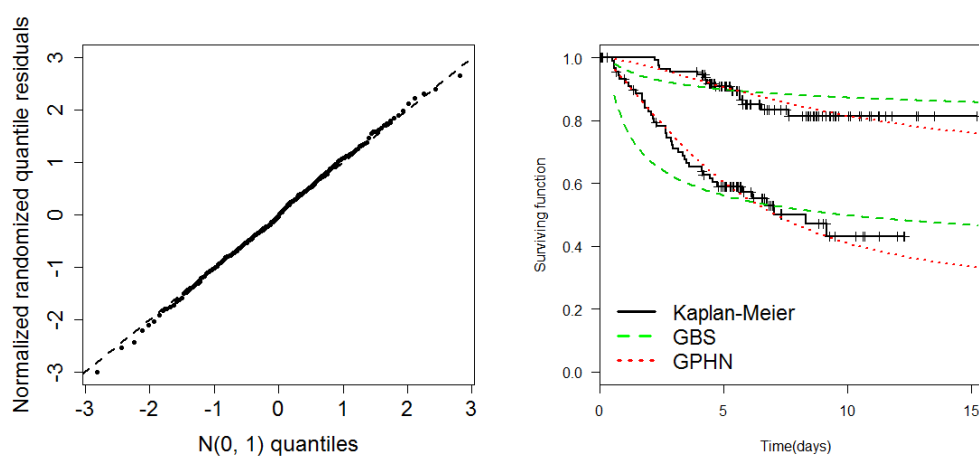
## 6. Concluding remarks

The paper focuses on proposing a new model for life-time data based on the GPHNcr distribution generated in a scenario of latent and competitive causes, which includes a cure fraction. The practical relevance of the new distribution was demonstrated in a real



**Table 3.** Maximum likelihood estimates for the parameters in models MPHN and GPHNcr.

Model	Estimates					Criteria	
	$\alpha$	$\sigma$	$\beta_{intercept}$	$\beta_{thickness}$	$\beta_{ulceration}$	AIC	BIC
GPHNcr	1.905 (0.293)	5.004 (1.079)	1.809 (0.346)	-0.179 (0.054)	-1.4804 (0.357)	420.86	437.472
MPHN	1.528 (0.172)	4.168 (0.155)	1.213 (0.336)	0.172 (0.076)	1.534 (0.411)	431.39	447.995
GBScr	1.357 (0.560)	10.091 (3.437)	1.257 (0.3455)	-0.168 (0.053)	-1.422 (0.353)	426.7	443.3
MBS	1.312 (0.201)	6.442 (0.565)	2.421 (0.763)	-0.956 (0.556)	-1.331 (0.821)	430.9	447.5



**Figure 5.** Residual plots for the model (left panel) and Kaplan-Meier estimates for ulceration patients (right panel)

data analysis as was indicated by computing Akaike and BIC scores to compare model fitting with MBS and GBScr models proposed by [4].

### Acknowledgments

We acknowledge two referees for comments and suggestions that substantially improved the presentation. The research of Yolanda M. Gómez was supported by Becas-Chile of the Chilean government. The research of Heleno Bolfarine was supported by CNPq and Fapesp funded by the Brazilian government.

## References

- [1] Akaike, H. (1974). *A new look at the statistical model identification*. IEEE Transactions of Automatic Control, **19**(6), 716-723.
- [2] Berkson, J., Gage, R. (1952). *Survival curve for cancer patients following treatment*. Journal of the American Statistical Association, **47**, 501-515.
- [3] Boag, J.W. (1949). *Maximum likelihood estimates of the proportion of patients cured by cancer therapy*. Journal of the Royal Statistical Society B, **11**,15-53.
- [4] Cancho, V.G., Louzada, F., Barriga, G.D.C. (2011) *The Geometric Birnbaum-Saunders regression model with cure rate*. Journal of Statistical Planning and Inference, **142**, 993-1000.
- [5] Cooray, K. and Ananda, M.M.A. (2008). *A Generalization of the Half-Normal Distribution with Applications to Lifetime Data*. Communications in Statistics - Theory and Methods, **37**, 1323-1337.
- [6] Durrans, S.R. (1992). *Distributions of fractional order statistics in hydrology*. Water Resources Research, **28**, 1649-1655.
- [7] Gómez, Y.M., Bolfarine, H. (2015). *Likelihood-based inference for power half-normal distribution*. Journal of Statistical Theory and Applications, **14**, 383-398.
- [8] Gupta, D. and Gupta, R.C. (2008). *Analyzing skewed data by power normal model*. Test, **17**, 197-210.
- [9] Lehmann, E.L. (1953). *The power of rank tests*. Annals of Mathematical Statistics, **24**, 23-43.
- [10] Maller, R.A., Zhou, X., (1996). *Survival Analysis with Long-Term Survivors*. Wiley, New York.
- [11] Marshall A.W. and Olkin I. (2007). *Life distributions: Structure of nonparametric, semi-parametric, and parametric families*. Springer Science+Business Media, LLC, New York.
- [12] Perdoná, G.S.C., Louzada-Neto, F. (2011). *A general hazard model for lifetime data in the presence of cure rate*. Journal of Applied Statistics **38**, 1395-1405.
- [13] Pescim, R.R., Demétrio, C.G.B., Cordeiro, G.M., Ortega, E.M.M and Urbano, M.R. (2010). *The beta generalized half-normal distribution*. Computational Statistics and Data Analysis, **54**, 945-957.
- [14] Pewsey, A., Gómez, H.W. and Bolfarine, H. (2012). *Likelihood-based inference for power distributions*. Test, **21**(4), 775-789.
- [15] R Development Core Team R. (2016). *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- [16] Rodrigues, J., Cancho, V.G., De Castro, M., Louzada-Neto, F. (2009). *On the unification of the long-term survival models*. Statistics and probability Letters, **79**, 753-759.
- [17] Schwarz, G. (1978). *Estimating the Dimension of a Model*. Annals of Statistics, **6**, 461-464.