International Journal of Sciences: Basic and Applied Research (IJSBAR)

ISSN 2307-4531
(Print & Online)

http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

Mixture Method to Estimate Baseline Hazard for Non-Arbitrary Function of the Cox Proportional Model

Farag Hamad[a*], Salem Abdulkarim[b], Ayman Hamad[c]

[a]*Department of Statistic, Faculty of Art & Science, University of Benghazi, Al Marj City, Libya*
[b,c]*Department of Physics, Faculty of Art& Science, University of Benghazi, Al Marj City, Libya*
[a]*Email: farag.hamad@uob.edu.ly*

Abstract

This study is extended to the work that was published by (Hamad & Kachouie, 2019). In this paper, a mixture method was used to estimate the Gompertz distribution parameters using hazard and cumulative hazard functions. This mixture method depends on the two different models (semi-parametric model, and nonparametric model). Cox proportional hazard model and Kaplan Meier used to estimate the Gompertz distribution parameters. The parameters of the logistic function (RHS) in the Cox proportional hazard can be estimated by the partial likelihood method. The hazard function (LHS) in the Cox model can be estimated by Kaplan Meier. The estimated parametric of the logistic function combined with the nonparametric estimate of the survival function by Kaplan Meier in order to get an estimate of the baseline hazard for Gompertz distribution. Improvement was archived in the estimated parameters of the baseline hazard using the mixture method compared to the use of the Cox method. The improvement of the mixture method was measured based on the estimated parameters for the baseline hazard as well as by the model goodness of fit. Different data types (simulation and real data) were used to measure the improvement of the mixture method. Monte Carlo simulations were carried out for evaluating the proposed method's performance. The results showed that the mixture method provided a better estimate value of the baseline and the model parameters compared to the estimated values using the Cox model.

Keywords: Kaplan Meier; Cox proportional hazard model; maximum likelihood; baseline hazard and partial likelihood; Gompertz distribution; Akaike information criterion; Bayesian information criterion.

---

* Corresponding author.

1. Introduction

Survival analysis is one of the statistical branches which deals with survival data. The goal of using survival analysis is to fit a model that illustrates the relationship between the event's time and some independent factors that might affect the survival time. Several approaches have been used for modeling survival data such as parametric model (linear regression), semi-parametric model (Cox proportional hazard), and nonparametric model (Kaplan Meier)[1]. Researchers now focusing on the modeling and fitting survival data. Survival data was modeled to estimate the failure distribution function, survival function, and hazard function [2].

The Cox proportional hazard model is the most popular semi-parametric regression model. A semi-parametric model makes some assumptions about a nonparametric baseline function to facilitate the estimation of the parameters of a function of independent variables [1]. The Cox proportional hazard model includes two components, baseline hazard and logistic regression function [3]. These two components can be estimated separately using the partial likelihood method. The coefficients of the logistic components were estimated first and then these estimated parameters can be used for estimating the baseline hazard [1]. There is no need to specify the baseline hazard to estimate the coefficients of the logistic component, and this is an important feature of the Cox proportional hazard model [4]. The partial likelihood were used to estimate the coefficients of the logistic components first [5].

The Kaplan Meier is one of the most popular forms of nonparametric models that can use to describe the survival data. In the survival analysis using Kaplan Meier, the product limit formula was used to estimate the probability of survival at a specific given time [1, 6]. The Kaplan Meier provides an estimate for the survival function and the hazard function. Estimated the survival function from the sample of censored data and using it to determine the survival curve [7]. The main goal of this study is using another estimator of the baseline hazard of the Gompertz distribution as well as provide an estimate for the parameters of the baseline hazard.

Since the calculation of the previous method depends on the partial likelihood method, which makes the estimate difficult. These limitations make biostatisticians sometimes need to develop methods for analyzing the survival data. Cox and Kaplan Meier provided a very thorough explanation of the hazard functions. The motivation of this study is to investigate the potential and useful baseline hazard of the Gompertz distribution. Using several mathematical properties of the hazard model, we estimated the baseline hazard for non-arbitrary of the Cox proportional model. This proposed method provides an estimate of the baseline hazard of the Cox model, which is not difficult to calculate.

2. Models

2.1 Semi-parametric model: Cox proportional hazard

Family of models were introduced by David Cox in 1972. These models are written in terms of hazard function. The Cox proportional hazard is a member of this family. The Cox model provides an expression for the hazard function at time t for an individual with a given set of explanatory variables [1]. The model is generally used in medical research to study the relationship between the survival time of patients and at least one indicator factor

[8]. It has also found a broad range of applications outside the medical research. For example, the Cox proportional hazard model is used in the financial sector for estimating the probability that a bank may survive until a specific time in the future [9]. The Cox hazard model:

$$h(t,x,\theta) = h_0(t)e^{\sum_{i=1}^{k}\theta_i X_i}, \qquad (2.1.1)$$

where $h_0(t)$, is an arbitrary baseline hazard which is a function of time t and X is a vector of explanatory variables. If the independent variable is a dichotomous variable, the hazard function is equal to the baseline hazard function $h_0(t)$ for $X = 0$ and equal to $h_0(t)e^{\sum_{i=1}^{k}\theta_i}$ for $X = 1$.

2.1.1 Estimating the parameters of the Cox proportional hazard

The Cox proportional hazard model contain two components, a baseline hazard $h_0(t)$ and a logistic regression function to model explanatory variables $x_1, x_2, \ldots\ldots, x_k$ [3]. These two components can be estimated independently. The coefficients of the logistic function θ are estimated first and then these estimated parameters can be used to estimate the baseline hazard [1]. This is an important feature of Cox proportional hazard model since it does not need to specify the baseline hazard for estimating the coefficients of the independent variables. The partial likelihood is used to estimate the coefficients(θ). For more details refer to [4], [7].

2.1.2 Estimation of the Cox model coefficients

Partial likelihood is used to estimate logistic coefficients $\theta_1, \theta_2, \ldots\ldots \theta_K$ [10]:

$$L(\theta) = \prod_{j=1}^{r} \frac{e^{\sum_{i=1}^{k}\theta_i X_i}}{\sum_{l\in w(t_j)} e^{\sum_{i=1}^{k}\theta_i X_i}} \qquad (2.1.2)$$

where $t_1, t_2, \ldots\ldots t_n$ are survival times and n is the sample size. The likelihood can be rewritten as:

$$L(\theta) = \prod_{j=1}^{r} \left[\frac{e^{\sum_{i=1}^{k}\theta_i X_i}}{\sum_{l\in w(t_j)} e^{\sum_{i=1}^{k}\theta_i X_l}} \right]^{c_i} \qquad (2.1.3)$$

where $w(t_j)$ is the risk at the time t_i and c_i is a binary variable which is zero if the observation i survives until time t_i. The coefficients are estimated by maximizing the partial likelihood function. From the computational perspective, estimating the values of coefficients that maximize the partial likelihood is equivalent to finding the values of the coefficients that maximize the log-likelihood function [8].

$$\log\big(L(\theta)\big) = \sum_{i=1}^{n} c_i \left[\sum_{i=1}^{k} \theta_i X_i - \log \sum_{l\in w(t_i)} e^{\sum_{i=1}^{k}\theta_i X_l} \right] \qquad (2.1.4)$$

The first derivative of log-likelihood function will be set to zero:

$$\frac{\partial \log(L(\theta_i))}{\partial \theta_i} = 0, \quad i = 1,2,\dots,k \tag{2.1.5}$$

and using Newton Raphson, the estimates $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$ are obtained. Furthermore, the variance of the estimated coefficients gets by taking the second derivative.

$$var(\hat{\theta}) = \left(\frac{d^2 \log(\theta)}{d\theta^2}\right)^{-1} \tag{2.1.6}$$

We can use the approximation of the variance to estimate a confidence interval of θ [1].

2.1.3 The baseline hazard

In 1973, Kalbfleisch and Prentice derived an estimate of baseline hazard function [1]. The estimate of baseline hazard was based on the partial likelihood estimates of the logistic model coefficients $\widehat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$. The estimate of baseline hazard function is given as:

$$\hat{S}_0(t_i) = 1 - \left(1 - \frac{e^{\sum_{i=1}^{k} \theta_i X_j}}{\sum_{l \in w(t_j)} e^{\sum_{i=1}^{k} \theta_i X_j}}\right)^{e^{\sum_{i=1}^{k} \theta_i X_j}} \tag{2.1.7}$$

where $w(t_j)$ is the set of all individuals n_j at risk at time t_i and x_j is the vector of independent covariates for the individual who experiences the event at time t_i [11, 12].

2.2 Non-parametric model: Kaplan Meier

The Kaplan Meier (KM) is a popular non-parametric survival analysis method that estimates the probability of survival at given time [1], [6]. To determine the Kaplan Meier curve, we estimate the survival function $S(t)$ from a sample of censored survival data. The estimate of the survival function is given as:

$$\hat{S}(t) = \prod_{i=1}^{k} \frac{n_i - d_i}{n_i} = \prod_{i=1}^{k} (1 - \frac{d_i}{n_i}) \tag{2.2.1}$$

For $t_k < t < t_{(k+1)}, k = 1,2, \dots., S(t) = 1$ for $t < t_1$, n_i denotes the number of individuals at risk at time t_i, d_i is the number of individuals that experienced the event at time $t_{(i)}$ [13], [14]. The stander error of $\hat{S}(t)$ can be estimated by:

$$S.e_{s(t)} \approx \hat{S}(t) \sqrt{\sum_{i}^{k} \left(1 - \frac{d_i}{n_i}\right)} \tag{2.2.2}$$

and the confidence interval of $S(t)$ is:

$$\left(\hat{S}(t) - z_{\frac{\alpha}{2}} S.e_{\hat{s}(t)}\ ,\qquad \hat{S}(t) + z_{\frac{\alpha}{2}} S.e_{\hat{s}(t)}\right) \tag{2.2.3}$$

2.3 Full parametric model with Gompertz baseline hazard

The baseline survival in Cox proportional hazard is an arbitrary function. Therefore, the corresponding survival function can be written as:

$$S(t,x,\theta) = e^{-\int_0^t h(u,x,\theta)du} \tag{2.3.1}$$

where t is survival time, x is set of factors, θ is set of the model parameters, and $h(t,x,\theta)$ is the hazard function. We assume hazard function has two components as follows:

$$h(t,x,\theta) = h_0(t,\theta_1)h_1(x,\theta_2) \tag{2.3.2}$$

where $h_0(t,\theta_1)$ is parametric base hazard and $h_1(x,\theta_2)$ is factor-specific hazard [15]. Hence, the full parametric survival model can be rewritten as:

$$S(t,x,\theta) = e^{-\int_0^t h(u,x)du} = e^{-\int_0^t h_0(t,\theta_1)h_1(x,\theta_2)du} = e^{-h_1(x,\theta_2)\int_0^t h_0(t,\theta_1)du} \tag{2.3.3}$$

and it can be simplified to:

$$S(t,x,\theta) = e^{-H_0(t,\theta_1)h_1(x,\theta_2)} = (S_0(t,\theta_1))^{h_1(x,\theta_2)} \tag{2.3.4}$$

where $H_0(t,\theta_1) = \int_0^t h_0(u,\theta_1)du$ is the total base hazard and $S_0(t,\theta_1)$ is base survival. Taking the natural log, we get:

$$\ln\big(S(t,x,\theta)\big) = -H_0(t,\theta_1)h_1(x,\theta_2) = h_1(x,\theta_2)\ln(S_0(t,\theta_1)) \tag{2.3.5}$$

We have discussed baseline survival function in the previous sections. The baseline survival function can have any relevant parametric form based on the application at hand. The next step after defining the baseline survival is to estimate the parameters of the model. We defined the baseline hazard as a Gompertz baseline. To estimate the baseline and its parameters, we assume $h_1(x,\theta_2)$ has a linear form and we estimate its parameter set $\theta_2$using least square method. Then we estimate the survival time $S(t,x)$ using Kaplan Meier method. By this way, the left-hand side of Equation (2.3.5) as well as factor-specific hazard $h_1(x,\theta_2)$ in the right-hand side of Equation (2.3.5) are estimated. The last step is estimating the baseline survival $S_0(t,\theta_1)$ using the estimates of θ_2 and $\ln\big(S(t,x,\theta)\big)$. Using the pervious equations to estimate the baseline hazard of the Cox proportional hazard in Equation (2.1.1).

2.4 Gompertz Distribution

The Gompertz distribution is commonly used in many applied research, particularly in lifetime data analysis

[16]. Gompertz distribution is a generalization of the exponential distribution and is related to Weibull distribution in which the log of the hazard is a linear function in of t [17]. A Gompertz distribution has been widely used in actuarial and biological applications and demography [18]. The Gompertz distribution imposes the following functional forms on the density, survival, hazard, and cumulative hazard function:

Probability density function	$f(t,a,b) = ae^{bt}e^{-\frac{a}{b}(e^{bt}-1)}$	(2.4.1)
Survival function	$S(t,a,b) = e^{-\frac{a}{b}(e^{bt}-1)}$	(2.4.2)
Hazard function	$h(t) = (t,a,b) = ae^{bt}$	(2.4.3)
Cumulative hazard function	$H(t) = \frac{a}{b}(e^{bt}-1)$	(2.4.4)

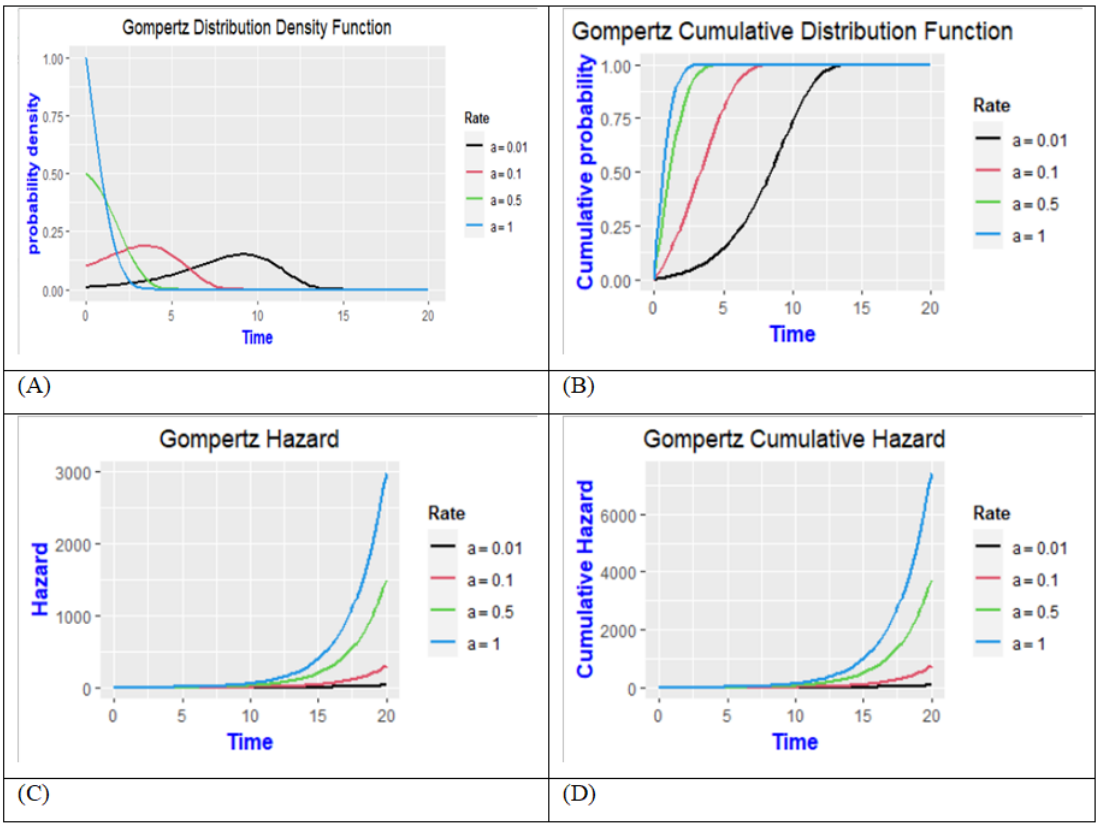

Figure 1: Gompertz hazard functions with different rate parameters: f(t), F(t), h(t), and H(t).

The Gompertz distribution can be characterized by using the hazard function

Figure **1**(C). The hazard function of the Gompertz distribution is increasing from a at time zero to ∞ at time ∞. Moreover, the natural log of the hazard is a linear function in time t.

3. Results and Discussion

This section evaluates the performance of the proposed method and compared it with the Cox Breslow method. The comparisons were done using two studies type (simulation and real application study). The results of the

estimated values of the parameters are compared based on four criteria including Akaike information criterion (AIC), Bayesian information criterion (BIC), coefficient of determination R^2, and residual standard error σ. Furthermore, the graphical comparison of these estimated baseline hazard, the estimated parameters of the Gompertz distribution and its bias are considered. The numerical evaluations were implemented using R language. The section 3.1 shows the implemented simulation study.

3.1 Simulation study

Simulation study carried out to evaluate the parametric baseline hazard estimated using the proposed method. The Gompertz baseline hazard was estimated for the survival data using the Cox model and proposed method. Three random samples were generated separately from a uniform distribution with different sizes 25, 50, and 100 observations. These observations were used to generate a random survival time using the Gompertz distribution baseline (rate =0.5 and shape=2). The independent variable of the logistic component was generated from a uniform distribution with parameters 2 and 100. RStudio was used to replicate the Monte Carlo simulation for N=5,000 with 10% additive white noise. Both methods have been applied to the simulated data and estimated the baseline hazard. Four criteria were considered to compare the performance of both methods including Akaike information criterion (AIC), Bayesian information criterion (BIC), coefficient of determination R^2, and residual standard error σ. The results are demonstrated in

Figure **1**Figure 4 and Table 1.

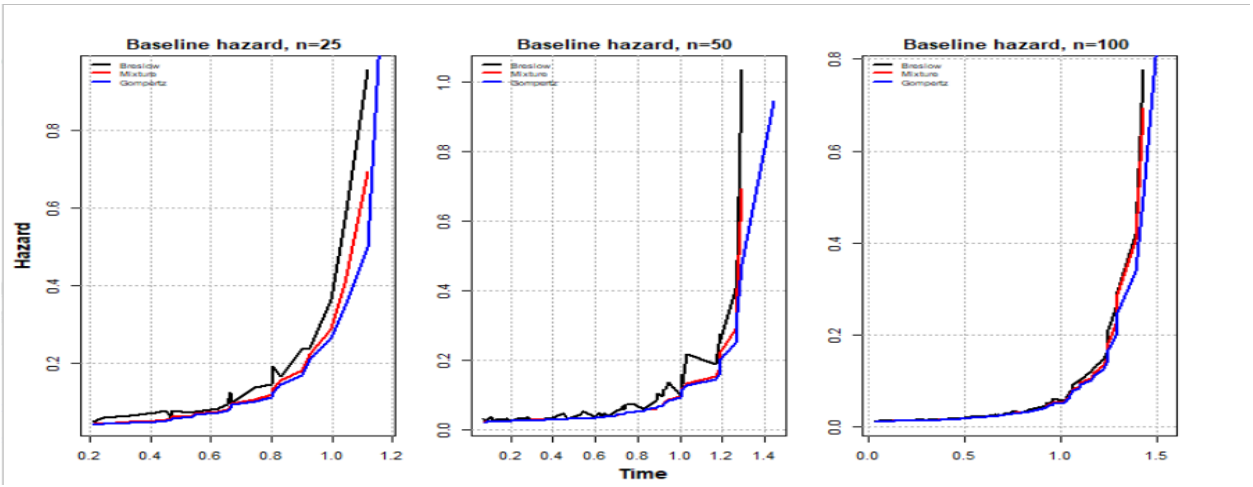

Figure 2: Cumulative baseline hazard for the Gompertz distribution using sample size of 25, 50, 100.

Figure 2 demonstrates the comparison between the estimated baseline hazard using the mixture method and Cox method for different sample sizes. The result shows that the estimated baseline hazard using the mixture method provides a better estimate of the Gompertz baseline hazard. By increasing the sample size to 100, the estimate of the baseline hazard using mixture method closer to the Gompertz baseline hazard. The model (2.3.5) residuals are shown in Figure 3. The estimated residuals using mixture and Cox Breslow methods were computed for different sample sizes. The scatter plots show that the estimated residuals using mixture method and Breslow method are comparable. The scatter plots of the estimated residuals using mixture method have less trend in

comparison with the scatter plots of the estimated residuals using Cox Breslow method. Moreover, the scatter plots of the model residuals versus the factor indicate have approximately Gaussian distribution.

Table 1: The Gompertz parameters estimation.

Method \ Parameters	n	$\hat{a}$	$S_{\hat{a}}$	$\hat{b}$	$S_{\hat{b}}$	AIC	BIC	R^2	$\hat{\sigma}$
Breslow	25	0.51	0.12	2.42	0.150	52.37	58.11	0.849	0.392
	50	0.46	0.08	2.46	0.106	54.98	61.93	0.880	0.339
	100	0.44	0.11	2.62	0.140	146.80	154.61	0.780	0.494
Mixture	25	0.54	0.10	2.34	0.120	49.94	55.56	0.856	0.390
	50	0.51	0.08	2.35	0.105	39.92	46.79	0.870	0.309
	100	0.52	0.09	2.44	0.120	112.68	120.43	0.798	0.421

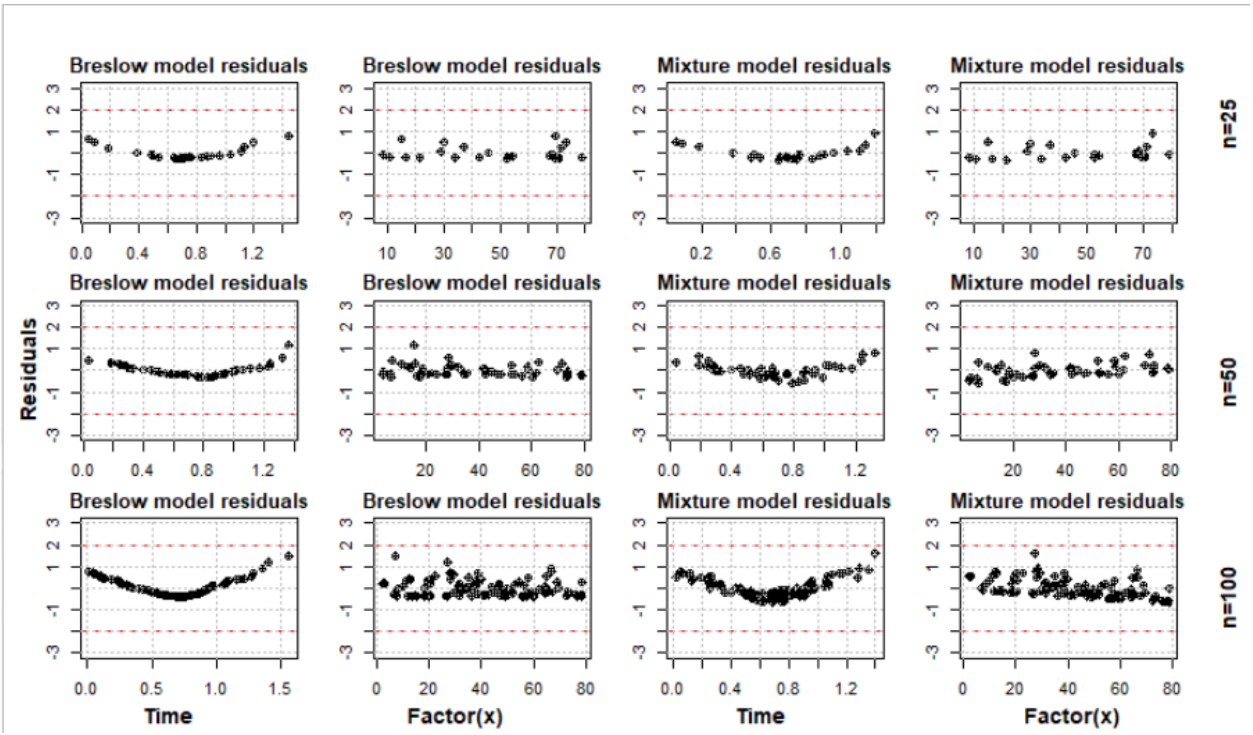

Figure 3: Model residuals using Gompertz baseline hazard for different samples size.

Mixture method was applied to estimate the parameters of the baseline survival in the model (2.3.5). Gompertz hazard was assumed for the arbitrary baseline hazard. The estimated parameters of the baseline survival using the mixture method were compared with those estimated parameters of the baseline survival using the Breslow method. Table 1 shows the estimated values for the Gompertz baseline survival parameters (rate and shape) and the model goodness of fit including AIC, BIC, R^2, and σ. The model results shows that the estimated parameters by the mixture method were closer estimates to the true parameters (rate=0.5 and shape=2) in comparison to those estimates using the Breslow method. In most cases, the proposed method gives a comparable performance

comparing with the Breslow method. Moreover, the estimated parameters using the proposed method for sample size n=25 closer to the true parameters compared with the estimated parameters using the Breslow method. It is clear from the results, there are an improvement using the proposed method for which the estimated value of the rate parameter using mixture method 0.54(0.10) compared with the estimated value of the same parameter using Breslow method 0.51(0.12). For the estimated shape of the baseline survival are 2.34(0.12) and 2.42(0.15) using the mixture and Breslow method respectively. By comparing the models goodness of fit using both methods, it is obvious that there are some improvements in the proposed model values in which the $\mathrm{AIC} = 49.94, \mathrm{BIC} = 55.56, \mathrm{R}^2 = 0.856, \text{and } \sigma = 0.390$ various $\mathrm{AIC} = 52.37, \mathrm{BIC} = 58.11, \mathrm{R}^2 = 0.849, \text{and } \sigma = 0.392.$

For n=100, we can see that the estimated value of the rate parameter is 0.52 with a stander error of 0.08 using the mixture method, while the same estimate of the rate parameter using the Breslow method is 0.44 with 0.11 stander error. Furthermore, the estimated value of the shape parameter using the mixture method is 2.44 with 0.120 stander error in comparison with 2.62 and 0.140 stander error. For the sample size n=100, we can see that the estimated values of the model coefficients using the mixture method are: $\mathrm{AIC} = 112.68, \mathrm{BIC} = 120.43, \mathrm{R}^2 = 0.798, \text{ and } \sigma = 0.421$ in comparison with the estimated values of the model coefficients using Breslow method $\mathrm{AIC} = 146.80, \mathrm{BIC} = 154.61, \mathrm{R}^2 = 0.780, \text{ and } \sigma = 0.494$. We can observe that the estimated coefficients using the mixture method were improved in comparison with the estimated coefficients using the Breslow method. More results about the estimated parameters of the baseline survival as well as the model goodness of fit can be found in Table 1.

Figure 3 demonstrates the scatterplots of the model residuals estimated using the mixture method and Breslow method. For both methos, we find the scatterplots for the residuals versus the survival time and the residuals versus the independent variable (X). From Figure 3, we can observe that the residuals are approximately normally distributed. Moreover, we can see that there is a slight trend in the residual scatterplots estimated using the Breslow method in comparison to scatterplots of the residual estimated using the mixture method.

The bias of the estimated parameters for the Gompertz baseline survival was computed using both methods. Figure 4, shows the histogram of the bias for the Gompertz parameters for different sample sizes. From the histogram, the of bias estimated using the mixture method is comparable with the bias that estimated using the Breslow method. Furthermore, the histogram of the bias indicts that the bias approximately normal distribution.

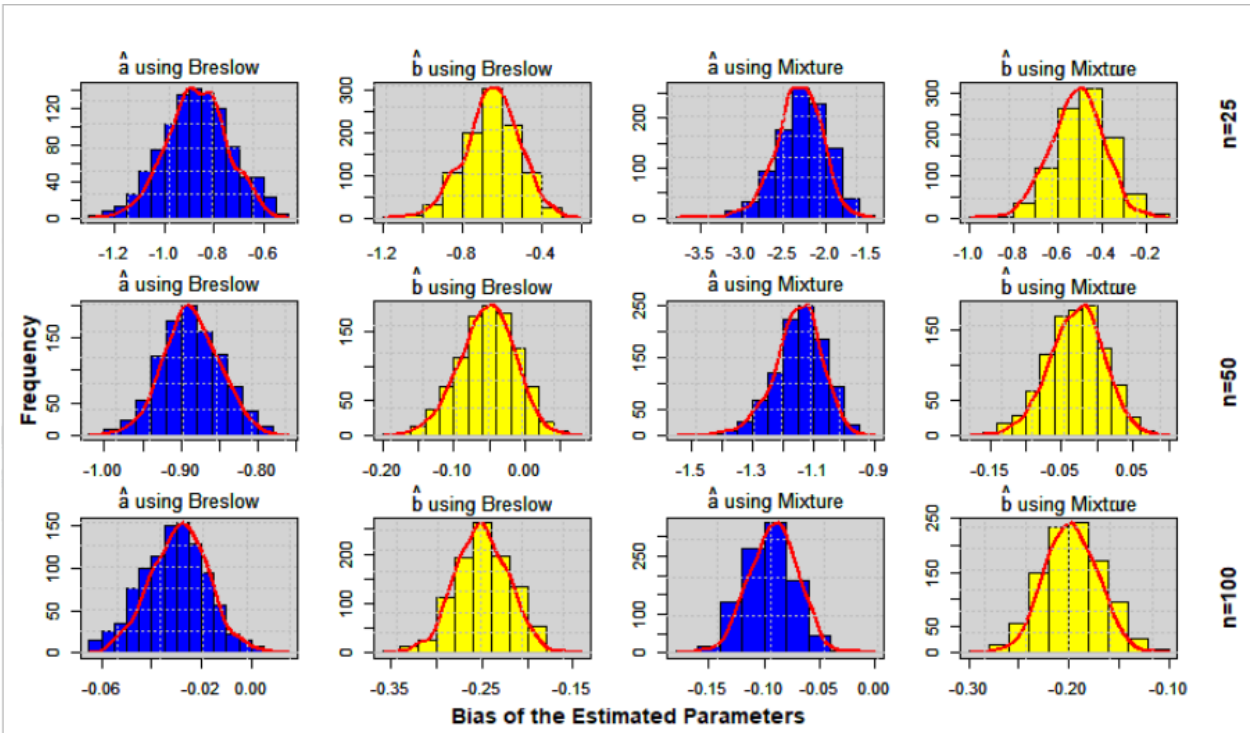

Figure 4: Estimated bias of Gompertz distribution parameters (a, b) for different samples size.

3.2 Real application study

In this step, the mixture model applied to estimate the baseline survival for the patients who have breast cancer. The dataset includes 272 breast cancer patients (as rows), and the data consists several factors that are related to the patients such as patient information, treatment, and survival time (data.World, 2016). The baseline hazard and baseline survival are estimated for the patients who have breast cancer using both methods. By considering the survival time of the patients who have breast cancer follows the Gompertz distribution, we estimated parameters of the Gompertz distribution. We compared the results visually and by evaluating the estimated models using the goodness of fit. The results are summarized in Table 2.

Table 2 : The estimated parameters of the baseline survival using both mixture and Breslow model.

Parameter / Method	n	â	b̂	AIC	BIC	R^2	σ̂
Breslow	272	0.055	0.291	146.144	154.105	0.851	0.476
Mixture		0.079	0.306	134.360	142.264	0.861	0.456

Table 2 shows the estimated baseline survival parameters using the mixture and Breslow method. From the results, it is clear that the estimated parameters of the baseline survival using both methods were comparable. Furthermore, we can see that the proposed performance was great in comparison with the Breslow method. By comparing those two estimated models visually, it observed that the estimated coefficients (goodness of fit coefficients) using the proposed method were improved in comparison with those estimated coefficients using

the Breslow method.

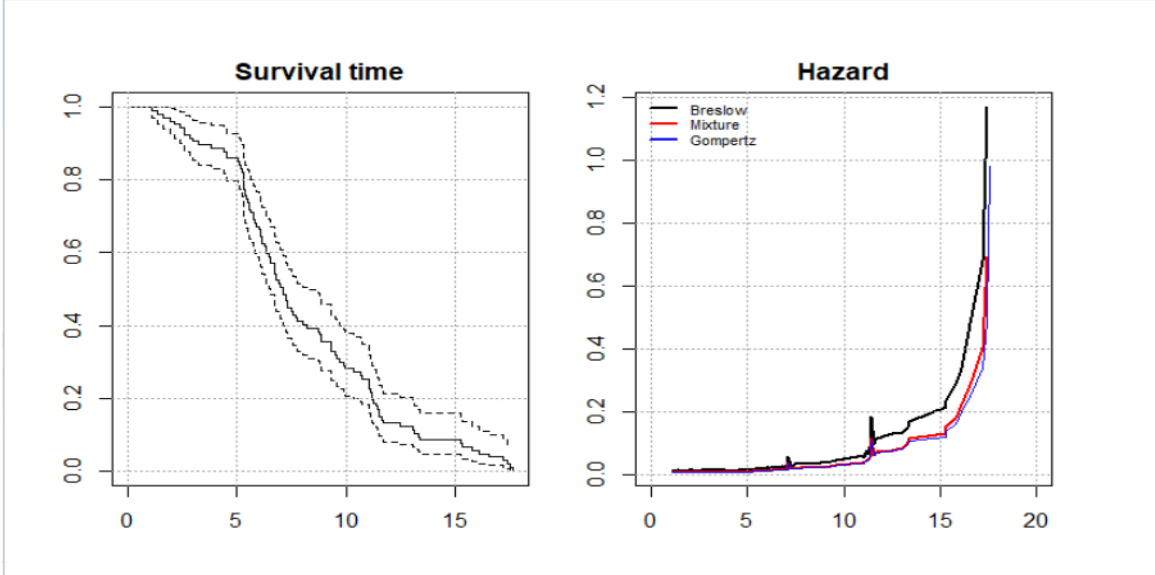

Figure 5: Survival time and hazard curves using cancer data.

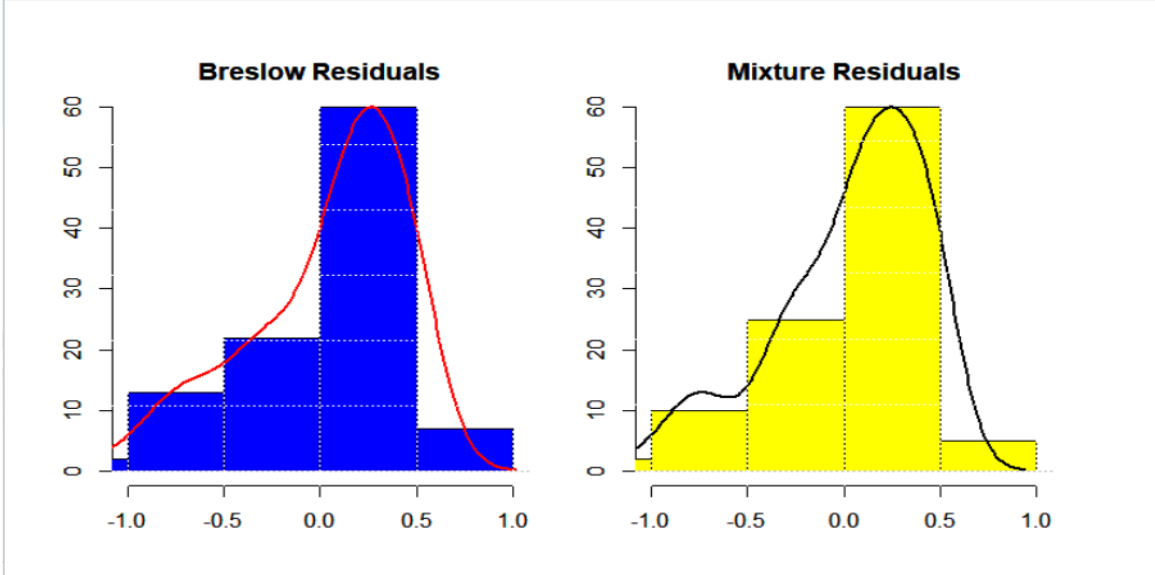

Figure 6: The histogram of the estimated residuals for the survival time of the patients.

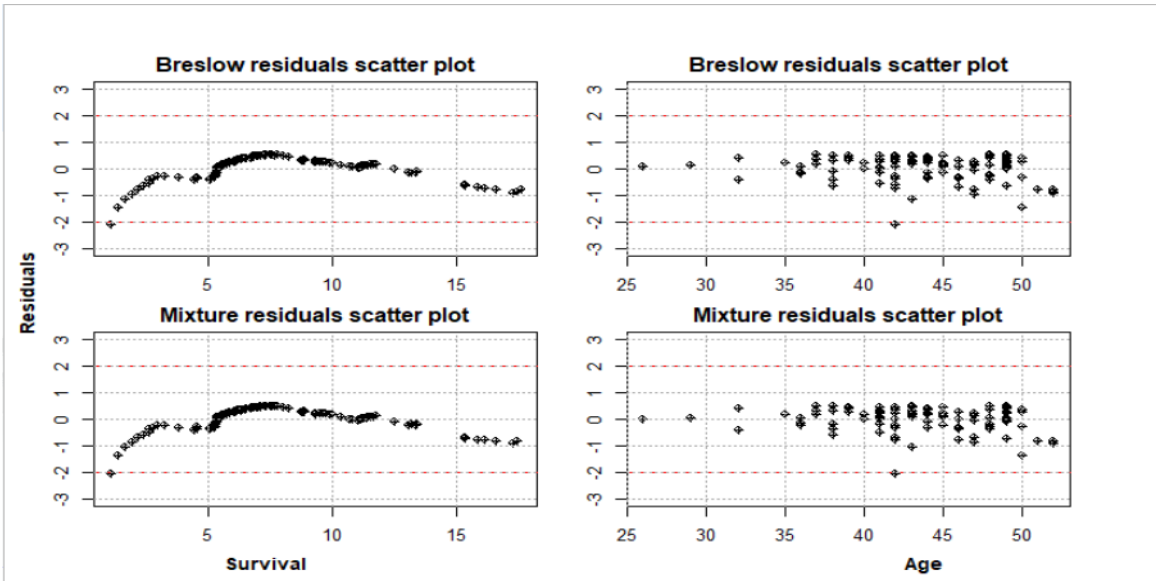

Figure 7: The scatter plots for the estimated residuals for the survival time model.

The curves in Figure 5 show the estimated baseline survival and hazard function using the survival time of the patients with cancer. Using the risk probability function, shows that the estimated baseline hazard using the mixture method and Breslow method were equivalent estimates. Moreover, we can see that the risk probability increases fast after 10 years of survive with breast cancer. The scatterplot in Figure 7 demonstrates the residuals vs. the survival time and residuals vs. the patient age for the proposed and Breslow method. The scatter plots of the residuals for both models are similar. Approximately normal residuals were obtained when the residuals were plotted vs. the patient's age. Moreover, there was a trend that appeared in the scatter plots of the residuals vs. the patient survival time.

4. Conclusion

The objective of this work was introducing a mixture method for estimating the baseline survival of the Cox model. The proposed method provides an estimate for the baseline survival based on the combination of the semiparametric and nonparametric models. The Cox proportional hazard model contains two components logistic component (RHS) and hazard component (LHS). Each component was estimated separately by partial likelihood (RHS) and by Kaplan Meier (LHS). Combine those estimations to get an estimate for the baseline hazard which was assumed as Gompertz baseline hazard. The results show that the estimated parameters of the baseline hazard using the mixture method were improved in comparison with the estimated parameters of the baseline hazard using the Cox Breslow method. The performance of the proposed method also appears on the goodness of fit coefficients. The goodness of fit coefficients shows that the proposed method provides a better estimating in comparison with the estimate of the model coefficients using the Breslow method. Despite the estimated baseline parameters using the mixture method and the Breslow method being comparable, the estimated values of the $\mathrm{AIC}, \mathrm{BIC}, \mathrm{R}^2$, and σ show that the fitted model using the mixture method provides a better estimate. For future research will be conducted a nonparametric baseline hazard to compare with the proposed method.

References

[1] D. Collett, *Modelling survival data in medical research*. CRC press, 2015.

[2] Z. Zhang, "Parametric regression model for survival data: Weibull regression model as an example," *Ann. Transl. Med.*, vol. 4, no. 24, 2016.

[3] J. Wieczorek, "Finite sample properties of minimum Kolmogorov-Smirnov estimator and maximum likelihood estimator for right-censored data," *Available ReseachGate*, 2009.

[4] F. E. Harrell, *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*, vol. 608. Springer, 2001.

[5] P. Hu, A. A. Tsiatis, and M. Davidian, "Estimating the parameters in the Cox model when covariate variables are measured with error," *Biometrics*, pp. 1407–1419, 1998.

[6] E. T. Lee and J. Wang, *Statistical methods for survival data analysis*, vol. 476. John Wiley & Sons, 2003.

[7] P. Wang, Y. Li, and C. K. Reddy, "Machine learning for survival analysis: A survey," *ACM Comput. Surv. CSUR*, vol. 51, no. 6, pp. 1–36, 2019.

[8] D. G. Kleinbaum and M. Klein, *Survival analysis*, vol. 3. Springer, 2010.

[9] G. Whalen, "A proportional hazards model of bank failure: an examination of its usefulness as an early warning tool," *Econ. Rev.*, vol. 27, no. 1, pp. 21–31, 1991.

[10] I. Langner, R. Bender, R. Lenz-Tönjes, H. Küchenhoff, and M. Blettner, "Bias of maximum-likelihood estimates in logistic and Cox regression models: a comparative simulation study," Discussion Paper, 2003.

[11] M. Franco, N. Balakrishnan, D. Kundu, and J.-M. Vivo, "Generalized mixtures of Weibull components," *Test*, vol. 23, no. 3, pp. 515–535, 2014.

[12] P. Royston, "Estimating a smooth baseline hazard function for the Cox model," *Lond. Dep. Stat. Sci. Univ. Coll. Lond.*, 2011.

[13] V. Bewick, L. Cheek, and J. Ball, "Statistics review 12: survival analysis," *Crit. Care*, vol. 8, no. 5, pp. 1–6, 2004.

[14] S. J. Wang, J. Kalpathy-Cramer, J. S. Kim, C. D. Fuller, and C. R. Thomas Jr, "Parametric survival models for predicting the benefit of adjuvant chemoradiotherapy in gallbladder cancer," 2010, vol. 2010, p. 847.

[15] F. Hamad and N. N. Kachouie, "A hybrid method to estimate the full parametric hazard model," *Commun. Stat.-Theory Methods*, vol. 48, no. 22, pp. 5477–5491, 2019.

[16] A. A. Jafari, S. Tahmasebi, and M. Alizadeh, "The beta-Gompertz distribution," *Rev. Colomb. Estad.*, vol. 37, no. 1, pp. 141–158, 2014.

[17] G. M. Cordeiro, M. Alizadeh, A. D. Nascimento, and M. Rasekhi, "The Exponentiated Gompertz Generated Family of Distributions: Properties and Applications.," *Chil. J. Stat. ChJS*, vol. 7, no. 2, 2016.

[18] A. Lenart, "The Gompertz distribution and maximum likelihood estimation of its parameters: a revision," *MPDIR Work Pap*, vol. 49, pp. 0–19, 2012.

[19] world data, "https://data.world/deviramanan2016/nki-breast-cancer-data." 2016.