International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 06, 2020
ISSN: 1475-7192

Modified Cox-Snell Residuals in Evaluating Gompertz Regression Model with Censored Data

[*1]Nur Niswah Naslina Azid @ Maarof,[2]Jayanthi Arasan,[3]Hani Syahida Zulkafli, [4]Mohd Bakri Adam

.

ABSTRACT--- *In this research, a two parameter Gompertz parametric survival model was extended to incorporate with covariate in the presence of right censored and uncensored data. The estimation procedure was studied at different sample sizes and censoring percentiles via simulation methodology. Statistically, the simulated data were assessed using the bias, standard error and root mean square error of the parameter estimates for the Gompertz regression model. Subsequently, various combinations of sample sizes and censoring levels were employed to evaluate the performance of the proposed modifications to the Cox-Snell residuals for both censored and uncensored observations. The results clearly indicate that the estimates perform well when the censoring degrees are lower, and the sample sizes are greater. The performance of the modified Cox-Snell residuals based on harmonic mean outperformed than the other approaches.*

Keywords--- *Gompertz Model, Right Censored Covariate, Cox-Snell Residuals, Simulation.*

I. INTRODUCTION

Survival modeling examines the relationship between survival time with one or more covariates. Nevertheless, a crucial analytical problem arises in analyzing survival regression model when censoring occurs due to incomplete data. This special type of missing data occurs in survival analyses when subjects do not experience the event of interest at the end of study or specified study time. One of the common censoring occurs in survival analyses is the right censoring. Right censoring ensues when the study end before the subject experienced the event of interest or observations that are lost to follow up or leave the study before an event occurs or withdraws. Standard statistical procedures are not amenable to handle with the censored observations. Undeniably, the use of diagnostics procedures for model checking is the vital part in the modeling process. Hence, Cox-Snell residuals is a widely used tool to assess the overall goodness of fit of the survival models (Cox & Snell, 1968).

In this paper, the performance of the Gompertz distribution with covariate in the presence of right censored data was studied extensively. A simulation study was carried out to evaluate the maximum likelihood estimation (MLE) procedure for the parameters of the Gompertz regression model at various censoring proportions and sample sizes by computing the values of bias, standard error (SE) and root mean square error (RMSE). Thereafter,

[1] Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, Malaysia, niswah@uitm.edu.my.*
[2] Assc. Statistics, Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, Malaysia.
[3] Statistics, Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, Malaysia.
[4] Statistics, Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, Malaysia

Received: 22 Sep 2019 | Revised: 13 Oct 2019 | Accepted: 15 Jan 2020

several modifications of the Cox-Snell residuals have been proposed. The performance of these methods was analyzed comprehensively at different level of censoring percentiles and sample sizes.

Historically, the Gompertz model was developed by a British actuary, (Gompertz, 1825) in describing human mortality curves and fitting actuarial tables. Numerous numbers of researchers have done studies on different features and statistical methodology of the Gompertz distribution, for instance, Garg et al. (1970) studied on the properties of the Gompertz model and compare the estimates by using the least-squares and maximum likelihood methods. Following that, Gordon (1990) had considered on the maximum likelihood estimates for the mixing proportions of two Gompertz distributions when censoring occurs. Makany (1991) conferred on theoretical justification of Gompertz model in the cases of accretionary growth. Later, Witten and Satzer (1992) discussed on the sensitivity of the parameter estimates using an alternative algorithm for estimating the model parameters of the Gompertz mortality rate model. Chen (1997) constructed an exact confidence interval and an exact joint confidence region for the parameters of the Gompertz distribution. While, Wu et al. (2004) explored on the unweighted and weighted least squares estimates for parameters of the Gompertz distribution under complete set of data and first failure censored data.

Lenart (2012) proved that maximum likelihood estimation gives a higher accuracy in parameter estimates for Gompertz model as compared to the method of moments. Kiani et al. (2012) studied on performance of the Gompertz model with time-dependent covariate as well as fixed covariate in the presence of right censored data and compared confidence interval estimations by Wald and Jackknife methods. Kiani and Arasan (2013) explored on the confidence intervals of the Gompertz model with the time-dependent covariate and fixed covariate in the presence of interval-, right-, left-censored and uncensored data based on the coverage probability. Dey (2018) deliberated several properties and methods in estimating the unknown parameters of Gompertz distribution. Recently, Ieren et al. (2019) extended the conventional Gompertz distribution to become a three parameters model known as power Gompertz distribution.

II. PURPOSES OF STUDY

This study intends to:

1. Evaluate the performance of the maximum likelihood estimation (MLE) for the parameters of the Gompertz regression model in the presence of right censored data via simulation study at various censoring proportions and sample sizes using bias, standard error (SE) and root mean square error (RMSE).

2. Assess the performance of several modifications of the Cox-Snell residuals for the model checking in Gompertz regression model in the presence of right censored data at different level of censoring percentiles and sample sizes.

1) Gompertz Distribution with Covariate and Right Censored Data

Let T be positive random variable representing the survival time. If T follows the Gompertz distribution, then the probability density function (PDF) is,

$$f(t;\gamma;\lambda) = \lambda\, exp(\gamma t) \times exp\left[\frac{\lambda}{\gamma}(1-e^{\gamma t})\right],\ t \geq 0, \lambda > 0, \gamma > 0, \tag{1}$$

where λ is known as baseline mortality, whereas γ is the senescent component. The corresponding survivor function and hazard function of the Gompertz distribution are given by,

$$S(t;\gamma;\lambda) = exp\left[\frac{\lambda}{\gamma}(1-e^{\gamma t})\right], \tag{2}$$

$$h(t;\gamma;\lambda) = \lambda\, exp(\gamma t), \tag{3}$$

respectively. The effect of covariate on the survival time of the i^{th} can be incorporated to the hazard function by letting parameter λ be a function of the covariate,

$$\lambda = exp(\beta' X). \tag{4}$$

For data set with a covariate x_i where $i = 1,2,\dots,n$, the hazard function for the i^{th}subject can be expressed as,

$$h(t_i;\gamma;\lambda) = \lambda_i\, exp(\gamma t_i), \tag{5}$$

where

$$\lambda_i = e^{\beta_0+\beta_1 x_i}. \tag{6}$$

Consequently, the hazard function is

$$h(t_i;x_i;\beta;\gamma) = e^{\beta_0+\beta_1 x_i+\gamma t_i}, \tag{7}$$

the probability density function is

$$f(t_i;x_i;\beta;\gamma) = e^{\beta_0+\beta_1 x_i+\gamma t_i+\frac{e^{\beta_0+\beta_1 x_i}}{\gamma}(1-e^{\gamma t_i})}, \tag{8}$$

with the corresponding survivor function given by,

$$S(t_i;x_i;\beta;\gamma) = e^{\left[\frac{e^{\beta_0+\beta_1 x_i}}{\gamma}(1-e^{\gamma t_i})\right]}. \tag{9}$$

2) Maximum Likelihood Estimation

In this study, the maximum likelihood estimation (MLE) was employed to obtain the parameter estimates of the model. In order to construct the likelihood function which is a function of the unknown parameters, following censoring indicators can be used to define for the i^{th}observations,

$$\delta_{E_i} = \begin{cases} 1, & \text{if the } i^{th} \text{ subject is uncensored} \\ 0, & \text{otherwise} \end{cases}$$

$$\delta_{R_i} = \begin{cases} 1, & \text{if the } i^{th} \text{ subject is right censored} \\ 0, & \text{otherwise} \end{cases}$$

The likelihood function for the full sample consisting of uncensored and right censored data for $i = 1,2,\dots,n$, is,

$$L(\beta;\gamma)=\prod_{i=1}^{n}[f(t_i,x_i,\beta,\gamma)]^{\delta_{E_i}}[S(r_i,x_i,\beta,\gamma)^{\delta_{R_i}}]$$

$$=\prod_{i=1}^{n}\left[exp\left[\beta_0+\beta_1x_i+\gamma t_i+\frac{e^{\beta_0+\beta_1x_i}}{\gamma}(1-e^{\gamma t_i})\right]\right]^{\delta_{E_i}}\left[exp\left[\frac{e^{\beta_0+\beta_1x_i}}{\gamma}(1-e^{\gamma t_i})\right]\right]^{\delta_{R_i}}. \quad (10)$$

and log-likelihood function is,

$$ln[L(\beta;\gamma)]=\sum_{i=1}^{n}\delta_{Ei}\left[\beta_0+\beta_1x_i+\gamma t_i+\frac{e^{\beta_0+\beta_1x_i}}{\gamma}(1-e^{\gamma t_i})\right]+\sum_{i=1}^{n}\delta_{Ri}\left[\frac{e^{\beta_0+\beta_1x_i}}{\gamma}(1-e^{\gamma t_i})\right] \quad (11)$$

The MLE method estimates the parameters of the model by maximizing the log-likelihood function. The Newton Raphson (NR) method also known as numerical iterative technique was employed to solve the non-linear equations simultaneously. The general formula for NR is,

$$\hat{\theta}=\theta_0+i^{-1}(\hat{\theta})u(\hat{\theta}) \quad 12)$$

where $u(\hat{\theta})$ is score vector as follows,

$$u(\hat{\theta})=\frac{\partial y}{\partial x}\Big|_{\theta=\hat{\theta}} \quad (13)$$

and $i(\hat{\theta})$ is the negative of the second partial derivatives of the log-likelihood function evaluated at $\boldsymbol{\hat{\theta}}=(\boldsymbol{\hat{\beta}_0},\boldsymbol{\hat{\beta}_1},\boldsymbol{\hat{\gamma}})$ as shown below,

$$i(\hat{\beta}_0,\hat{\beta}_1,\hat{\gamma})^{-1}=\left[\begin{bmatrix}\frac{\partial^2 l}{\partial\beta_0^2} & \frac{\partial^2 l}{\partial\beta_0\partial\beta_1} & \frac{\partial^2 l}{\partial\beta_0\partial\gamma}\\ \frac{\partial^2 l}{\partial\beta_1\partial\beta_0} & \frac{\partial^2 l}{\partial\beta_1^2} & \frac{\partial^2 l}{\partial\beta_1\partial\gamma}\\ \frac{\partial^2 l}{\partial\gamma\partial\beta_0} & \frac{\partial^2 l}{\partial\gamma\partial\beta_1} & \frac{\partial^2 l}{\partial\gamma^2}\end{bmatrix}\right]^{-1} \quad (14)$$

Equation (14) also known as the inverse of observed information matrix provides the estimates of the variance and covariance matrix**.**

III. SIMULATION STUDY AND RESULTS

1) Assessing Performance of the Parameter Estimates

A simulation study using 1000 samples each with $n=30,40,50,80$ and 100 was conducted for the Gompertz regression model with both censored and uncensored observations as well as fixed covariate, x_i. The covariate values were simulated independently from the standard normal distribution. The values of $-5, 0.3$ and 0.5 were chosen as the parameters of β_0,β_1 and γ to mimic real life survival data. A sequence of random numbers, u_i's from the standard uniform distribution on the interval $(0,1)$ was generated to produce lifetimes t_i for $i=1,2,\dots,n$ subjects. The censoring times, c_i were generated from the exponential distribution where the value μ could be

Received: 22 Sep 2019 | Revised: 13 Oct 2019 | Accepted: 15 Jan 2020

adjusted to obtain the desired approximate censoring proportion (cp) for the data with $cp = 0\%, 10\%, 20\%, 40\%$ and 50%. The simulated survival time is considered censored if $t_i > c_i$, and will be replaced by the corresponding censoring time. The survival time t_i was generated by,

$$t_i = \frac{1}{\gamma} log\left[1 - \frac{\gamma\ log(1-u_i)}{\gamma}\right] \quad (15)$$

A set of measures were used to evaluate the performance of parameters β_0, β_1 and γ. Bias, standard error and root mean square error has been calculated to evaluate the accuracy, precision and stability of estimator's performance.

Based on the Table 1, the results from the bias values show an inconsistent pattern as the sample size and the censoring proportion increase. However, in Table 2, the standard error values increase when the censoring proportions increase. Meanwhile, as the sample sizes increase, the standard error values will be decreased. Table 3 substantiates that the root mean square error values also increase as the censoring proportion increase. This indicates that poorer performance for the parameter estimates at smaller sample sizes and higher censoring proportions, whereas larger sample sizes and lower censoring proportions would have higher accuracy and efficiency of the parameter estimates.

Table 1: Bias of the parameters β_0, β_1 and γ

Estimates	*Sample Size, n*	*Censoring Proportion (%)*					
		0	*10*	*20*	*30*	*40*	*50*
$\hat{\beta}_0$	30	-0.2716	-0.2158	-0.1712	-0.0849	0.0058	0.0517
	40	-0.1900	-0.1293	-0.0630	0.0071	0.1069	0.1798
	50	-0.1598	-0.1128	-0.0089	0.0857	0.1680	0.2651
	80	-0.1069	0.0039	0.0644	0.1545	0.2916	0.3368
	100	-0.0659	0.0219	0.1028	0.1637	0.3399	0.3935
$\hat{\beta}_1$	30	0.0194	0.0247	0.0158	0.0018	0.0211	-0.0075
	40	0.0135	0.0129	0.0021	0.0016	0.0047	-0.0032
	50	0.0185	0.0172	0.0030	-0.0050	-0.0082	-0.0051
	80	0.0096	0.0060	-0.0062	-0.0114	-0.0143	-0.0178
	100	0.0025	0.0023	0.0029	-0.0109	-0.0202	-0.0154
$\hat{\gamma}$	30	0.0405	0.0373	0.0334	0.0304	0.0258	0.0237
	40	0.0293	0.0233	0.0199	0.0144	0.0084	0.0045
	50	0.0237	0.0211	0.0121	0.0045	-0.0002	-0.0096
	80	0.0160	0.0048	0.0013	-0.0058	-0.0170	-0.0196
	100	0.0107	0.0020	-0.0049	-0.0079	-0.0250	-0.0277

Table 2: Standard Error (SE) of the parameters β_0, β_1 and γ

Estimates	*Sample Size, n*	*Censoring Proportion (%)*					
		0	*10*	*20*	*30*	*40*	*50*
	30	0.7051	0.7704	0.8084	0.8872	0.8634	0.9877

$\hat{\beta}_0$	40	0.5842	0.6418	0.6334	0.7326	0.7998	0.7589
	50	0.5173	0.5490	0.5764	0.6084	0.6579	0.6554
	80	0.3869	0.4103	0.4462	0.4570	0.4855	0.5129
	100	0.3629	0.3659	0.3824	0.4050	0.4331	0.4683
$\hat{\beta}_1$	30	0.2243	0.2332	0.2549	0.2803	0.3281	0.3629
	40	0.1914	0.1920	0.2064	0.2339	0.2740	0.2747
	50	0.1580	0.1668	0.1770	0.2040	0.2097	0.2269
	80	0.1218	0.1342	0.1359	0.1447	0.1647	0.1728
	100	0.1130	0.1158	0.1286	0.1282	0.1434	0.1585
$\hat{\gamma}$	30	0.0887	0.0980	0.1029	0.1171	0.1145	0.1322
	40	0.0735	0.0805	0.0818	0.0944	0.1053	0.1022
	50	0.0655	0.0696	0.0752	0.0808	0.0870	0.0854
	80	0.0492	0.0524	0.0581	0.0602	0.0637	0.0689
	100	0.0455	0.0468	0.0495	0.0524	0.0573	0.0622

Table 3: Root Mean Square Error (RMSE) of the parameters β_0, β_1 and γ

Estimates	*Sample Size, n*	*Censoring Proportion (%)*					
		0	*10*	*20*	*30*	*40*	*50*
$\hat{\beta}_0$	30	0.7556	0.8001	0.8263	0.8913	0.8634	0.9890
	40	0.6144	0.6547	0.6365	0.7326	0.8069	0.7799
	50	0.5414	0.5605	0.5765	0.6144	0.6790	0.7070
	80	0.4014	0.4103	0.4508	0.4824	0.5663	0.6136
	100	0.3688	0.3666	0.3960	0.4368	0.5506	0.6117
$\hat{\beta}_1$	30	0.2251	0.2345	0.2554	0.2803	0.3288	0.3630
	40	0.1919	0.1925	0.2064	0.2339	0.2741	0.2747
	50	0.1590	0.1677	0.1770	0.2040	0.2099	0.2270
	80	0.1221	0.1343	0.1360	0.1451	0.1653	0.1737
	100	0.1130	0.1158	0.1286	0.1287	0.1449	0.1592
$\hat{\gamma}$	30	0.0975	0.1048	0.1081	0.1210	0.1174	0.1343
	40	0.0791	0.0838	0.0842	0.0955	0.1056	0.1023
	50	0.0696	0.0727	0.0761	0.0809	0.0870	0.0860
	80	0.0517	0.0526	0.0581	0.0604	0.0659	0.0717
	100	0.0467	0.0468	0.0497	0.0530	0.0625	0.0681

IV. ASSESSING MODEL FIT

1) Modification of Cox-Snell Residuals

Cox-Snell residuals, r_{Ci}, is commonly used in the analysis of survival data as a model adequacy procedure. In this study, model adequacy was evaluated via graphical plot of residuals. A log-cumulative hazard plot of residuals ican be obtained by plotting the Cox-Snell residual against the cumulative hazard function to assess the model's fit. A well fit model should have an intercept that approaches to zero and the slope as well as R-square approach to one. One criticism of Cox-Snell residuals is that they do not account for censored observations, therefore the adjusted Cox-Snell residuals were devised by Crowley & Hu (1977) whereby the standard Cox-Snell residual, r_{Ci} could be used for uncensored observations and $r_{Ci} + \Delta$ which $\Delta = \log(2) = 0.693$, is used to adjust the residual. The Cox-Snell residuals for the i^{th} individual, $i = 1,2,\ldots,n$ is given by,

$$r_{Ci} = -log\hat{S}_i(t_i) \tag{16}$$

The modified Cox-Snell residuals has been proposed to account for censored data. Crowley and Hu (1977) found that the addition of unity to a Cox-Snell residual for a censored observation inflated the residual to too great an extent. Hence, the median value was calculated for the excess residual. A second version of the modified Cox-Snell residual is,

$$r_{Ci}^1 = \begin{cases} r_{Ci}, & \text{for uncensored observations,} \\ r_{Ci} + 0.693, & \text{for censored observations.} \end{cases}$$

In this research, we had proposed two modifications to the Cox-Snell residuals as follows,

$$r_{Ci}^2 = \begin{cases} r_{Ci}, & \text{for uncensored observations,} \\ r_{Ci} + G, & \text{for censored observations.} \end{cases}$$

and

$$r_{Ci}^3 = \begin{cases} r_{Ci}, & \text{for uncensored observations,} \\ r_{Ci} + H, & \text{for censored observations.} \end{cases}$$

where G is the geometric mean of residuals and H is the harmonic mean of residuals.

2) Simulation Study

A simulation study by using 1000 samples each with different number of sample sizes, $n = 30, 40, 50, 80$ and 100 as well as the censoring proportions, $cp = 0\%, 10\%, 20\%, 40\%$ and 50% was conducted to compare the residual values. Plot of $ln\left[-ln\left(\hat{S}(r_{Ci})\right)\right]$ against $ln(r_{Ci})$ should exhibit a linear line through the origin with a unit gradient if the data fits the model well. Several modification of the Cox-Snell residuals were used and compare the performance for censored and uncensored data.

a) Cox-Snell Residuals, r_{Ci}.

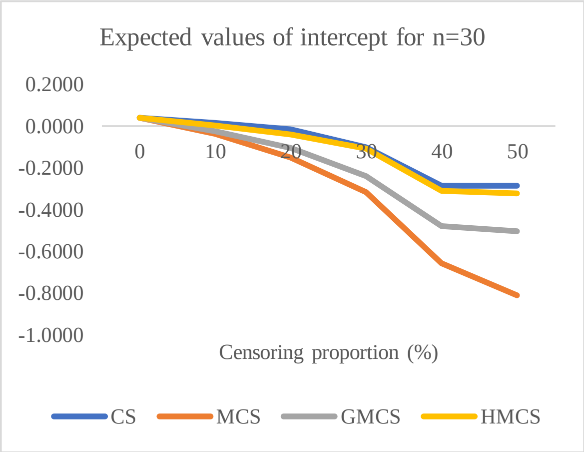
b) Modified Cox-Snell Residuals, r_{Ci}^1.

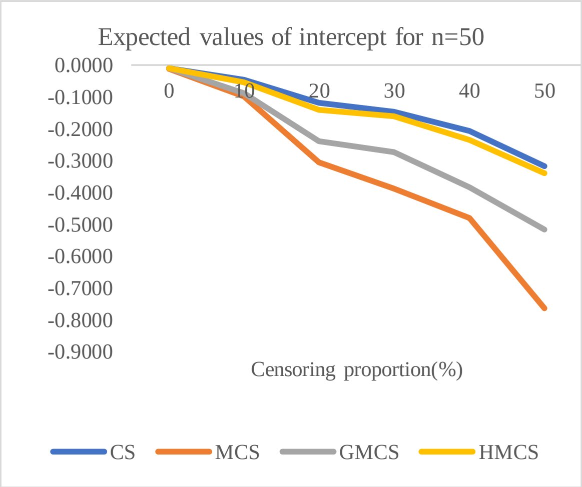
c) Replace the median with geometric mean of existing data, r_{Ci}^2.

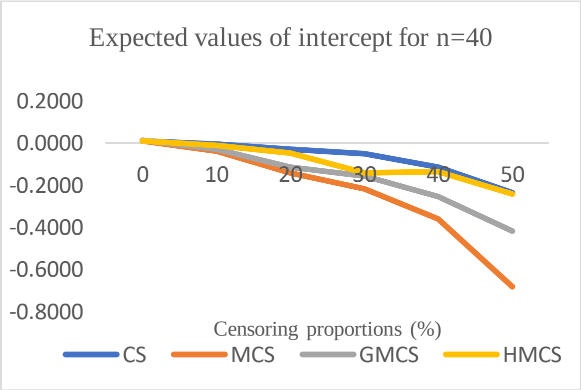
d) Replace the median with harmonic mean of existing data, r_{Ci}^3.

V. DISCUSSION

The selection criterions which are intercept, slope, r and R-square were used to compare the residuals performance. Figure 1 shows the comparison of residuals for estimated values of intercept. The modified Cox-Snell residuals using harmonic mean demonstrates an alike pattern as in Cox-Snell residuals. As the sample sizes increase, the slopes become closer to one. However, when the censoring proportions becomes higher, expected values for intercept and slopes go further than zero and one respectively. Further, the range for r and R^2 values also become wider as it across higher censoring proportions. The residuals perform well in model diagnosis when sample size is large and censoring proportion is low.

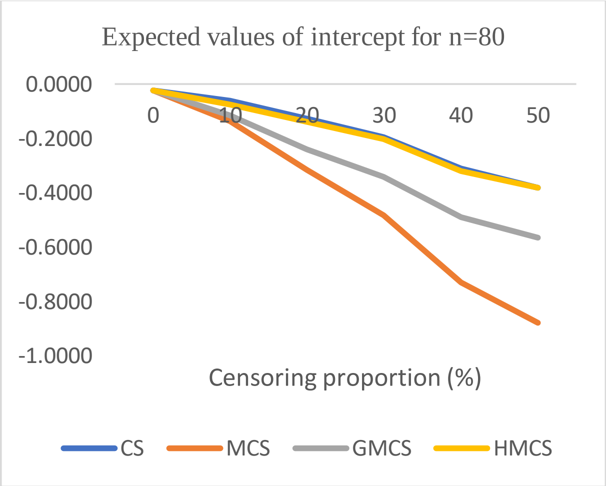






Received: 22 Sep 2019 | Revised: 13 Oct 2019 | Accepted: 15 Jan 2020

International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 06, 2020
ISSN: 1475-7192



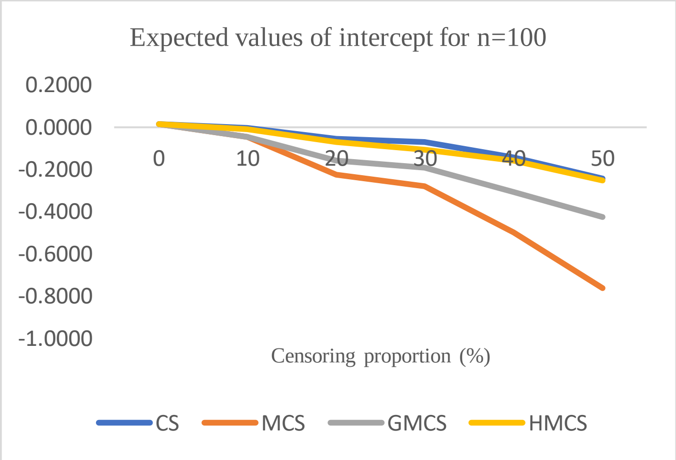
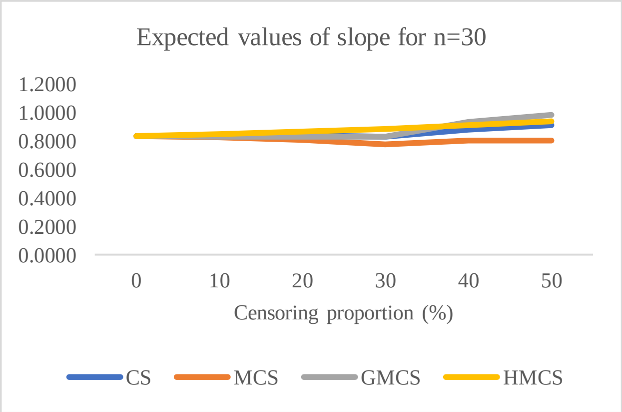


Figure 1: Comparison of residuals for estimated values of intercept



Received: 22 Sep 2019 | Revised: 13 Oct 2019 | Accepted: 15 Jan 2020

International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 06, 2020
ISSN: 1475-7192

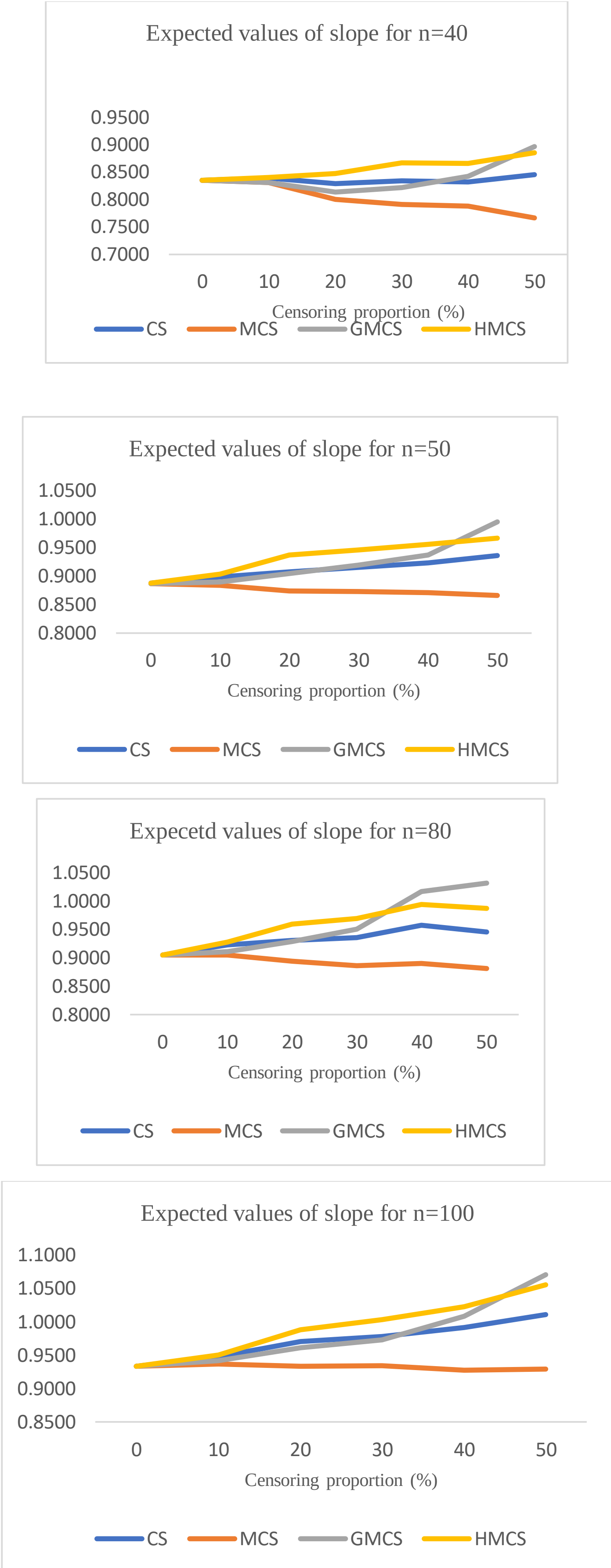


Figure 2: Comparison of residuals for estimated values of slope

Received: 22 Sep 2019 | Revised: 13 Oct 2019 | Accepted: 15 Jan 2020

Table 4: Comparison of intercept values for various residuals

Sample Size, n	*Censoring Proportion (%)*	*CS*	*MCS*	*GMCS*	*HMCS*
30	0	0.0405	0.0405	0.0405	0.0405
	10	0.0147	-0.0386	-0.0258	0.0036
	20	-0.0167	-0.1499	-0.1060	-0.0391
	30	-0.1030	-0.3179	-0.2405	-0.1091
	40	-0.2855	-0.6591	-0.4794	-0.3109
	50	-0.2863	-0.8114	-0.5041	-0.3221
40	0	0.0110	0.0107	0.0117	0.0109
	10	-0.0057	-0.0389	-0.0307	-0.0119
	20	-0.0310	-0.1428	-0.1147	-0.0493
	30	-0.0513	-0.2189	-0.1570	-0.1412
	40	-0.1157	-0.3607	-0.2546	-0.1376
	50	-0.2371	-0.6827	-0.4190	-0.2412
50	0	-0.0102	-0.0108	-0.0100	-0.0104
	10	-0.0464	-0.0991	-0.0883	-0.0548
	20	-0.1180	-0.3049	-0.2383	-0.1396
	30	-0.1458	-0.3873	-0.2730	-0.1611
	40	-0.2065	-0.4809	-0.3836	-0.2343
	50	-0.3182	-0.7643	-0.5159	-0.3400
80	0	-0.0228	-0.0231	-0.0227	-0.0229
	10	-0.0613	-0.1357	-0.1149	-0.0742
	20	-0.1259	-0.3144	-0.2399	-0.1390
	30	-0.1937	-0.4818	-0.3417	-0.2030
	40	-0.3101	-0.7299	-0.4900	-0.3194
	50	-0.3805	-0.8772	-0.5644	-0.3807
100	0	0.0154	0.0154	0.0154	0.0154
	10	-0.0031	-0.0452	-0.0474	-0.0114
	20	-0.0548	-0.2245	-0.1586	-0.0713
	30	-0.0714	-0.2811	-0.1923	-0.1077
	40	-0.1440	-0.4993	-0.3066	-0.1591
	50	-0.2447	-0.7627	-0.4259	-0.2526

Table 5: Comparison of slope values for various residuals

Sample Size, n	*Censoring Proportion (%)*	*CS*	*MCS*	*GMCS*	*HMCS*
30	0	0.8353	0.8353	0.8353	0.8353
	10	0.8355	0.8254	0.8310	0.8490
	20	0.8392	0.8075	0.8308	0.8659
	30	0.8314	0.7739	0.8288	0.8846
	40	0.8810	0.8044	0.9332	0.9107

Received: 22 Sep 2019 | Revised: 13 Oct 2019 | Accepted: 15 Jan 2020

	50	0.9090	0.8048	0.9808	0.9386
40	0	0.8354	0.8353	0.8354	0.8354
	10	0.8389	0.8307	0.8311	0.8405
	20	0.8286	0.8005	0.8131	0.8474
	30	0.8339	0.7911	0.8218	0.8670
	40	0.8322	0.7874	0.8420	0.8662
	50	0.8452	0.7662	0.8962	0.8853
50	0	0.8871	0.8866	0.8863	0.8869
	10	0.8983	0.8835	0.8894	0.9028
	20	0.9072	0.8734	0.9037	0.9362
	30	0.9145	0.8722	0.9186	0.9456
	40	0.9232	0.8701	0.9364	0.9552
	50	0.9354	0.8652	0.9952	0.9663
80	0	0.9045	0.9044	0.9044	0.9045
	10	0.9222	0.9050	0.9111	0.9274
	20	0.9301	0.8939	0.9284	0.9592
	30	0.9354	0.8855	0.9505	0.9689
	40	0.9573	0.8902	1.0167	0.9937
	50	0.9454	0.8809	1.0317	0.9871
100	0	0.9333	0.9333	0.9333	0.9333
	10	0.9476	0.9367	0.9416	0.9496
	20	0.9702	0.9332	0.9607	0.9876
	30	0.9776	0.9339	0.9728	1.0030
	40	0.9907	0.9273	1.0079	1.0221
	50	1.0105	0.9293	1.0701	1.0546

Table 6: Comparison of r values for various residuals

n	*Censoring Proportion (%)*	*CS*		*MCS*		*GMCS*		*HMCS*	
		min	*max*	*min*	*max*	*min*	*max*	*min*	*max*
30	0	0.7086	0.9879	0.7086	0.9879	0.7086	0.9879	0.7086	0.9879
	10	0.6316	0.9915	0.6290	0.9804	0.5920	0.9804	0.6427	0.9850
	20	0.6931	0.9855	0.5519	0.9831	0.5164	0.9795	0.6927	0.9871
	30	0.7488	0.9837	0.5844	0.9825	0.5017	0.9881	0.7299	0.9888
	40	0.7394	0.9795	0.6245	0.9848	0.5282	0.9705	0.7678	0.9825
	50	0.7453	0.9815	0.5411	0.9688	0.5011	0.9741	0.7496	0.9782
40	0	0.8802	0.9904	0.8802	0.9904	0.8804	0.9905	0.8802	0.9904
	10	0.8916	0.9925	0.8916	0.9929	0.8920	0.9929	0.8851	0.9918
	20	0.8837	0.9903	0.8627	0.9919	0.8548	0.9912	0.8698	0.9934
	30	0.8917	0.9875	0.8291	0.9857	0.8007	0.9873	0.8673	0.9897
	40	0.8816	0.9895	0.8304	0.9880	0.7910	0.9864	0.8659	0.9910

	50	0.8329	0.9846	0.7451	0.9885	0.7166	0.9888	0.8275	0.9858
50	0	0.7728	0.9913	0.7728	0.9913	0.7734	0.9913	0.7728	0.9913
	10	0.7687	0.9924	0.7638	0.9898	0.7548	0.9898	0.7690	0.9903
	20	0.8150	0.9886	0.7478	0.9898	0.6943	0.9914	0.8123	0.9901
	30	0.8059	0.9908	0.7019	0.9904	0.6810	0.9927	0.8216	0.9918
	40	0.8187	0.9935	0.6969	0.9896	0.6626	0.9850	0.8117	0.9899
	50	0.7999	0.9846	0.5915	0.9866	0.5938	0.9863	0.7966	0.9885
80	0	0.7955	0.9968	0.7955	0.9968	0.7955	0.9968	0.7955	0.9968
	10	0.7938	0.9966	0.7767	0.9960	0.7762	0.9960	0.7892	0.9962
	20	0.7958	0.9949	0.7446	0.9947	0.7405	0.9959	0.7952	0.9948
	30	0.7905	0.9955	0.7103	0.9948	0.7104	0.9936	0.7964	0.9948
	40	0.7558	0.9922	0.6776	0.9892	0.6319	0.9905	0.7748	0.9908
	50	0.7422	0.9895	0.6486	0.9871	0.6265	0.9878	0.7645	0.9904
100	0	0.9079	0.9951	0.9079	0.9951	0.9079	0.9951	0.9079	0.9951
	10	0.9010	0.9944	0.9026	0.9941	0.8896	0.9946	0.9010	0.9943
	20	0.9154	0.9948	0.8746	0.9945	0.8709	0.9943	0.9144	0.9949
	30	0.9127	0.9946	0.8654	0.9949	0.8610	0.9946	0.9095	0.9940
	40	0.8882	0.9941	0.8317	0.9938	0.8026	0.9933	0.9019	0.9937
	50	0.8781	0.9928	0.7545	0.9867	0.7723	0.9913	0.8804	0.9941

Table 7: Comparison of R-square values for various residuals

n	*Censoring Proportion (%)*	*CS*		*MCS*		*GMCS*		*HMCS*	
		min	*max*	*min*	*max*	*min*	*max*	*min*	*max*
30	0	0.6978	0.9874	0.6978	0.9874	0.6978	0.9874	0.6978	0.9874
	10	0.6179	0.9911	0.6153	0.9796	0.5769	0.9797	0.6294	0.9845
	20	0.6817	0.9849	0.5359	0.9825	0.4991	0.9787	0.6814	0.9866
	30	0.7395	0.9829	0.5690	0.9818	0.4832	0.9877	0.7199	0.9883
	40	0.7297	0.9787	0.6106	0.9842	0.5108	0.9694	0.7592	0.9819
	50	0.7359	0.9808	0.5241	0.9677	0.4826	0.9732	0.7404	0.9774
40	0	0.8770	0.9901	0.8770	0.9901	0.8772	0.9902	0.8770	0.9901
	10	0.8886	0.9923	0.8886	0.9927	0.8891	0.9927	0.8820	0.9915
	20	0.8804	0.9900	0.8590	0.9916	0.8509	0.9910	0.8663	0.9932
	30	0.8887	0.9871	0.8245	0.9853	0.7953	0.9869	0.8638	0.9894
	40	0.8771	0.9892	0.8258	0.9877	0.7854	0.9860	0.8623	0.9907
	50	0.8284	0.9841	0.7382	0.9882	0.7090	0.9885	0.8228	0.9854
50	0	0.7679	0.9912	0.7679	0.9912	0.7686	0.9912	0.7679	0.9912
	10	0.7637	0.9922	0.7588	0.9895	0.7496	0.9895	0.7641	0.9901
	20	0.8111	0.9883	0.7424	0.9896	0.6878	0.9912	0.8083	0.9899
	30	0.8018	0.9906	0.6956	0.9902	0.6742	0.9925	0.8178	0.9916

Received: 22 Sep 2019 | Revised: 13 Oct 2019 | Accepted: 15 Jan 2020

	40	0.8149	0.9933	0.6905	0.9894	0.6554	0.9847	0.8077	0.9896
	50	0.7956	0.9841	0.5830	0.9863	0.5852	0.9860	0.7923	0.9883
80	0	0.7929	0.9967	0.7929	0.9967	0.7929	0.9967	0.7929	0.9967
	10	0.7912	0.9965	0.7738	0.9959	0.7733	0.9960	0.7865	0.9962
	20	0.7931	0.9948	0.7413	0.9947	0.7371	0.9959	0.7925	0.9947
	30	0.7878	0.9954	0.7065	0.9947	0.7066	0.9935	0.7937	0.9947
	40	0.7526	0.9921	0.6734	0.9891	0.6272	0.9904	0.7719	0.9907
	50	0.7389	0.9894	0.6440	0.9869	0.6217	0.9876	0.7614	0.9902
100	0	0.9069	0.9951	0.9069	0.9951	0.9069	0.9951	0.9069	0.9951
	10	0.9000	0.9943	0.9016	0.9940	0.8885	0.9945	0.9000	0.9943
	20	0.9145	0.9947	0.8733	0.9945	0.8696	0.9943	0.9136	0.9949
	30	0.9118	0.9945	0.8640	0.9949	0.8595	0.9945	0.9086	0.9939
	40	0.8870	0.9940	0.8299	0.9938	0.8005	0.9933	0.9009	0.9936
	50	0.8768	0.9928	0.7520	0.9865	0.7700	0.9912	0.8792	0.9940

VI. CONCLUSION

The results of bias, standard error and root mean square error show that poorer performance for the parameter estimates at higher censoring proportions and smaller sample sizes. The findings from Table 4 until Table 7 demonstrates the range of intercept, slope, r and R-square values are decrease as sample size increase. However, there are having an opposite trend as censoring proportions increase. The residuals perform well in model diagnosis when sample size is large and censoring proportion is small. Overall, we can conclude that the proposed modification of the Cox Snell residual using harmonic mean outperform than the other residuals.

VII. ACKNOWLEDGEMENT

We would like to thank the Putra Grant, Vot 9595300, Universiti Putra Malaysia for supporting this research project.

REFERENCES

1. Chen, Z. (1997). Parameter Estimation of the Gompertz Population. Biometrical Journal, 39(1), 117-124.
2. Cox, D.R., & Snell, E.J. (1968). A General Definition of Residuals. Journal of the Royal Statistical Society Series B (Methodological), 30(2), 248-275.
3. Dey, S., Moala, F.A., & Kumar, D. (2018). Statistical properties and different methods of estimation of Gompertz distribution with application. Journal of Statistics and Management Systems, 21(5), 839–876.
4. Garg, M.L., Rao, B.R., & Redmond, C.K. (1970). Maximum-Likelihood Estimation of the Parameters of the Gompertz Survival Function. Journal of the Royal Statistical Society Series C (Applied Statistics), 19(2), 152-159.

5. Gompertz, B. (1825). On the Nature of the Function Expressive of the Law of Human Mortality and on a New Mode of Determining the Value of Life Contingencies. Philosophical Transactions of the Royal Society of London, 115, 513-583.

6. Gordon, N.H. (1990). Maximum likelihood estimation for mixtures of two gompertz distributions when censoring occurs. Communications in Statistics - Simulation and Computation, 19(2), 733–747.

7. Ieren, T.G., Kromtit, F.M., Agbor, B.U., Eraikhuemen, I.B., & Koleoso, P.O. (2019). A Power Gompertz Distribution: Model, Properties and Application to Bladder Cancer Data. Asian Research Journal of Mathematics, 15(2), 1-14.

8. Kiani, K., & Arasan, J. (2013). Gompertz model with time-dependent covariate in the presence of interval-, right- and left-censored data. Journal of Statistical Computation and Simulation, 83(8), 1472-1490.

9. Kiani, K., Arasan, J., & Midi, H. (2012). Interval estimations for parameters of gompertz model with time dependent covariate and right censored data. Sains Malaysiana, 41(4), 471-480.

10. Lenart, A. (2012). The moments of the Gompertz distribution and maximum likelihood estimation of its parameters. Scandinavian Actuarial Journal, 2014(3), 255–277.

11. Makany, R. (1991). A Theoretical Basis for Gompertz's Curve. Biometrical Journal, 33(1), 121–128.

12. Witten, W., & Satzer, W. (1992). Gompertz survival model parameters: Estimation and sensitivity. Applied Mathematics Letters, 5(1), 7-12.

13. Wu, J.-W., Hung, W.-L., & Tsai, C.-H. (2004). Estimation of the parameters of the Gompertz distribution using the least squares method. Applied Mathematics and Computation, 158(1), 133–147.