Constant-Linear and Constant-Quadratic Piecewise Survival Models

Agung Prabowo, Ibnu Ginanjar Susilo, Agus Sugandha
Department of Mathematics, Faculty of Mathematics and Natural Sciences,
Universitas Jenderal Soedirman, Indonesia
agung.prabowo@unpad.ac.id, agung_nghp@yahoo.com, ibnuginanjar19@gmail.com,
agussugandha74@gmail.com

Mustafa Mamat
Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin,
Gong Badak Campus, 21300, KualaTerengganu, Terengganu, Darul Iman, Malaysia
musmat567@gmail.com, must@unisza.edu.my

Sukono
Department of Mathematics, Faculty of Mathematics and Natural Sciences,
Universitas Padjadjaran, Indonesia
sukono@unpad.ac.id

Abdul Talib Bon
Department of Production and Operations,
University Tun Hussein Onn Malaysia, Malaysia
talibon@gmail.com

Abstract

One distribution that is often used in modeling of survival time is exponential. An example of its application is the problem of the length of time until all the light bulbs go out. In this example, the observed event is the time until all the light bulbs fail to ignite. Thus, the random variable states that the length of time until all the light bulbs is go out. By using an exponential distribution, one assumption that can be made is that the hazard function has a constant hazard rate. This assumption was felt unsatisfactory, so a new model was made. The model is called the piecewise exponential model. The purpose of this study is to build a cumulative hazard model in the form of a non-composite function (one step/piece) with the assumption that the damage rate is constant at each time interval. Thus, the hazard function is a function of time. The model of the hazard function can be in the form of piecewise-exponential constant-linear and constant-quadratic model in the form of a composite function (several steps). The cumulative hazard function in the form of a non-composite (one step) function is obtained from the composite hazard function with the help of an indicator function. The cumulative hazard function in the form of a step function can be seen as a regression function. If data on the length of light bulbs lifespan are available, then the parameters in the cumulative hazard function can be estimated using the least squares method. Chi-square test was used to test the null hypothesis which states that the long life span of a light bulb follows a constant-linear and constant-quadratic piecewise model. The use of Chi-square test requires a frequency distribution table. The table is made with the Sturgess rule to determine the number of interval classes. Based on secondary data, two cumulative hazard models in the form of non-composite functions are produced as a combination of constant-linear and quadratic functions.

Keywords:
Constant-linear, constant-quadratic, exponential, piecewise, survival.

1. Introduction

The survival of a lamp can be measured statistically using an analysis of the life time. The survival time data obtained from the results of the study are non-negative variables and form a function (Lawless, 1982: 1). In the analysis of survival time there is a function used to measure the level of risk of failure of a product called the hazard rate function (Bowers, et al., 1997: 52). In the industrial field, the product failure test aims to obtain the possibility that a product will suffer damage for the first time (mean time to failure). By knowing the risk of failure of a product, companies can measure quality through the products produced.

One distribution that is often used in modeling survival is an exponential distribution. Suppose the random variable T states the length of time until the occurrence of the observed event. In this case, the event observed is the length of time until all the lamps go out. For example, the length of time until all the lights observed go out are expressed by a random variable $T \sim \text{Eksp}(\lambda)$. Probability density function for random variable T is $f(t) = \lambda \cdot e^{-\lambda \cdot t}$, $t \geq 0$ and has a constant hazard rate λ. The use of exponential distribution with constant hazard rate is less satisfying, so a new model called the piecewise exponential model is generated.

The use of exponential distribution in various cases has been reported by several researchers. London (1997) and Riaman et al (2018) produce a constant-quadratic composite hazard rate model. In this article two composite hazard rate models will be constructed, namely linear-constant and constant-quadratic, and determine the best model for the data used. The model built using the distribution of the length of time the event occurs is exponential in the form of composite or piecewise.

The piecewise exponential model is created by dividing the time interval into sub-intervals with each sub-interval having its own hazard rate. Because each sub-interval has a constant rate of damage, the model obtained is called the piecewise exponential model. Piecewise exponential model has two types, namely constant-linear and constant-quadratic piecewise exponential. From the two piecewise exponential models, the hazard rate function and cumulative hazard rate function were obtained and applied it to secondary data to obtain the best cumulative hazard rate and hazard rate models.

Weiss (2017) develop a procedure to learn forests as combinations of piecewise-constant and parameterized distributions to compactly model survival distributions from data. Pottmann et al (2000) study piecewise linear approximation of quadratic functions defined on *R(n)*. Bouaziz and Nuel (2017) suggest a new method to estimate the piecewise constant hazard rate model. The initiated method is used to find the number and location of the cut points, as well as to estimate the hazard at each cutting interval. The piecewise exponential model can be used in injury risk modeling as a function of experience and has the added benefit of interpretability of other similar flexible models (Kubo, et al., 2013). Sari, Thamrin, and Lawi (2014) conducted Bayesian estimates to model frailty with exponential piecewise for multivariate survival data.

2. Research Methodology

In this article, we describe examples of applications of survival models, specifically light bulb survival. For example, a random variable T states the length of time until all of the observed lights go out. In this case, the event observed was that all lamps went out, no matter how long.

This study uses secondary data with samples $n = 49$ lamps whose it survival time is measured in hours. Of all the lamps studied, we note the length of time so that all extinguished. The smallest data value is 1,051 hours and the largest is 17,568 hours, meaning that the 49 light bulbs have a long life span between 1,051 - 17,568 hours. That is, before 1,051 hours all of the light bulbs remained on and after 17,568 hours all of the light bulbs had gone out. Data on the life span of light bulbs in hours are listed in Table 1 (London: 1997):

Table 1. Light Bulb Survival Time Data (in Hours)

1,051	4,006	5,905	8,108	10,205	11,608	14,110
1,337	4,012	5,956	8,546	10,396	11,745	14,496
1,389	4,063	6,068	8,666	10,861	11,762	15,395
1,921	4,921	6,121	8,831	11,026	11,895	16,179
1,942	5,445	6,473	9,106	11,214	12,044	17,092
2,322	5,620	7,501	9,711	11,362	13,520	17,568
3,629	5,817	7886	9,806	11,604	13,670	17,568

The purpose of this study is to determine the cumulative hazard rate and hazard rate functions in the form of non-composite functions. Next, apply the two models with secondary data in Table 1. The model obtained is compared with that of other researchers to determine the best model based on MAPE criteria. To achieve this goal, the following are the steps of the research. Steps 1 - 4 are theoretical work. The next step is solved using secondary data as a practical application of the theory developed in steps 1 - 4.

Stage 1: Hazard Rate Function
In this step, the hazard rate function is modeled in the form of a constant-linear and constant-quadratic piecewise function in the form of a two-fragment composite function.

Stage 2: Survival Function
The survival function is modeled based on Equation (1)

$$S(t) = exp\left[-\int_0^t h(x)\,dx\right] \tag{1}$$

where $h(x)$ are $h_1(t)$ or $h_2(t)$ on stage 1.

Stage 3: Cumulative Hazard Rate Function
Cumulative Hazard Rate Function is obtained by Equation (2)

$$H(t) = -\ln S(t) \tag{2}$$

Stage 4: Regression Equations of the Cumulative Hazard Rate Function
Compile cumulative hazard rate function equation obtained in step 3 as a regression equation using the indicator function so that cumulative hazard rate function is obtained in the form of non-composite functions.

Stage 5: Empirical Survival Model
Empirical survival function $\overline{S}^0(t)$ generated from the data in Table 1 using Equation (3)

$$\overline{S}^0(t) = \begin{cases} 1 & t \le t_1 \\ \dfrac{n-j}{n} & t_j < t < t_{j+1} \quad ; j = 1,2,\dots,n-1 \\ 0 & t \ge t_n \end{cases} \tag{3}$$

Function $\overline{S}^0(t)$ is an estimate of the survival function $S(t)$ in Equation (1).

Stage 6: Empirical Cumulative Hazard Rate Model
The cumulative hazard rate function in Equation (2) is estimated by the empirical cumulative hazard function. Equation of empirical cumulative hazard functions is determined by Equation (4):

$$\overline{\Lambda}^0(t) = -ln\,\overline{S}^0(t) \tag{4}$$

Stage 7: Empirical Hazard Rate Model
The empirical hazard function is determined by Equation (5):

$$\overline{h}^0(t) = \frac{-\frac{d}{dt}\overline{S}^0(t)}{\overline{S}^0(t)} \tag{5}$$

The empirical hazard rate function is an estimate for the hazard rate function.

Stage 8: Regression Model of Empirical Cumulative Hazard Rate
Arranging the empirical cumulative hazard rate function into a regression function with two independent variables in Equation (6):

$$c = \lambda x_1 + \beta x_2 \tag{6}$$

Stage 9: Estimating Parameters in the Regression Model
Estimating parameters λ and β using secondary data in the form of long life span of light bulbs. Parameters λ and β can be estimated by the least squares method with Equation (7):

$$\hat{\theta} = \left(X^T X\right)^{-1} X^T Y \tag{7}$$

where:

$$\hat{\theta} = \begin{bmatrix} \hat{\theta}_0 \\ \hat{\theta}_1 \\ \hat{\theta}_2 \\ . \\ . \\ \hat{\theta}_k \end{bmatrix}, \left(X^T X\right) = \begin{bmatrix} n & \sum x_1 & . & . & \sum x_k \\ \sum x_1 & \sum x_1^2 & . & . & \sum x_1 x_k \\ \sum x_2 & \sum x_2 x_1 & . & . & \sum x_2 x_k \\ . & . & . & . & . \\ \sum x_k & \sum x_k x_1 & . & . & \sum X_k^2 \end{bmatrix}, \text{ dan } \left(X^T Y\right) = \begin{bmatrix} \sum y \\ \sum x_1 y \\ \sum x_2 y \\ . \\ \sum x_n y \end{bmatrix}$$

Stage 10: Goodness of Fit Test
Test the null hypothesis which states the length of time follows the constant-linear and constant-quadratic piecewise model with the Chi-square test. Hypothesis testing is done twice with the formula:
H_0: The data follows a constant-linear piecewise exponential model
H_1 : The data does not follow a constant-linear piecewise exponential model
and
H_0 : The data follows a constant-quadratic piecewise exponential model
H_1 : The data does not follow a constant-quadratic piecewise exponential model

Chi-square test requires a frequency distribution table. The table is made with the Sturgess rule to determine the number of interval classes. In the Chi-square test, the calculated Chi-square value is obtained by Equation (8)

$$\chi^2 = \frac{(O_i - E_i)^2}{E_i} \tag{8}$$

In this test, a significance level of α = 0.05 and a degree of freedom $v = k - c$ where k states the number of cells that are not empty and c states the number of parameters estimated. For the case example worked on, the free degree is $v = 7 - 2 = 5$. Chi square table value is $\chi^2_{0,05;7} = 11.0705$. H_0 is rejected if $\chi^2 > 11.0705$.

Stage 11: The Best Cumulative Hazard Model
Determination of the best model is measured from the smallest mean absolute percentage error (MAPE) with Equation (9).

$$\text{MAPE} = \frac{1}{n}\sum_{j=1}^{n}\left|\frac{X_j - F_j}{X_j}\right| \times 100\% \tag{9}$$

where X_j and F_j states the value of the data and the prediction of the data j respectively. The models being compared are the models obtained from this study and the models from two researchers namely London (1997) and Riaman et al. (2018).

3. Discussion Results

In this section, the piecewise exponential model is derived which produces two types of hazard rate models, namely the constant-linear and constant-quadratic piecewise models. The model obtained is tested with secondary data to produce a cumulative hazard rate function. Next, a comparison is made between the cumulative hazard rate values obtained from the data and the cumulative hazard rate values obtained from the model. In this case three cumulative hazard rate models are used. The first model was created by London (1997) and Riaman et al (2018). Two other models were produced in this study.

The two models are exponential piecewise constant-linear and constant-quadratic hazard rate models. Both models are based on data used by London (1997) and Riaman et al (2018). The difference in the solution method is on the construction of the frequency distribution table so that the resulting chi-square value and parameter estimates are also different. A limitation of this study, as it did in London (1997) and Riaman et al (2018) is the use of assumptions in the regression equation $c = \alpha + \lambda x_1 + \beta x_2$ where $\alpha = 0$ on Equation (6).

3.1 Hazard Rate Function

The hazard rate graph of the piecewise exponential model is expressed in the form of a composite function, namely (1) a constant-linear piecewise exponential combination and (2) a constant-quadratic piecewise combination. Applications of the two composite hazard rate models are found in the life of the light bulb.

In the first (constant-linear) model, the hazard rate function is given by Equation (10) and the graph is given in Figure 1 to the left. In the second model (constant-quadratic), the hazard rate function is given by Equation (11) and the function graph is available in Figure 1 to the right.

$$h_1(t)=\begin{cases}\lambda & 0\le t<y\\ \lambda+\beta(t-y) & t\ge y\end{cases} \tag{10}$$

$$h_2(t)=\begin{cases}\lambda & 0\le t<y\\ \lambda+\beta(t-y)^2 & t\ge y\end{cases} \tag{11}$$

where $\lambda > 0$ and $\beta > 0$.

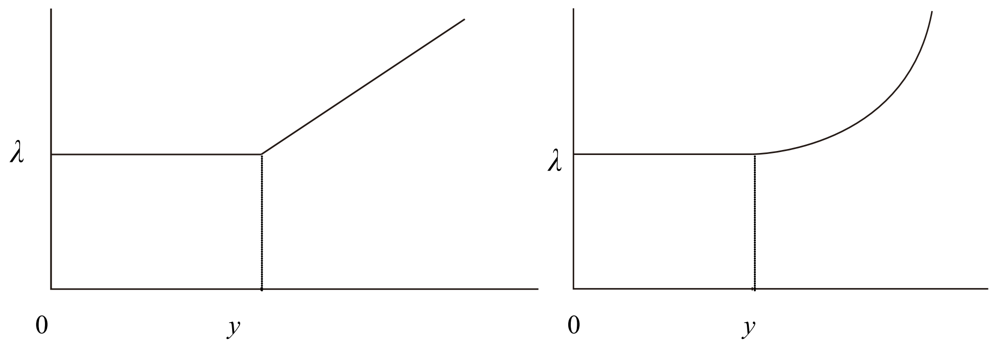

Figure 1. Graph of Hazard Rate Functions of the Constant-Linear (Left) and Constant-Quadratic Model

3.2 Survival Function

From the hazard rate function in Equations (10) and (11) a survival function can be built using Equation (1). The result is for the constant-linear piecewise exponential model (Equation (12)), and for the quadratic constant-quadratic piecewise model (Equation (13))

$$S_1(t)=\begin{cases}\exp(-\lambda t) & 0\le t<y\\ \exp\left[-\lambda t-\frac{\beta}{2}(t-y)^2\right] & t\ge y\end{cases} \tag{12}$$

$$S_2(t)=\begin{cases}\exp(-\lambda t) & 0\le t<y\\ \exp\left[-\lambda t-\frac{\beta}{3}(t-y)^3\right] & t\ge y\end{cases} \tag{13}$$

3.3 Cumulative Hazard Rate Function

The survival model in Equations (12) and (13) is a two-part model. Determining the values and the model is not easy to do. For this reason, the two equations are converted into non-composite functions.

Parameters λ and β estimated by the least squares method. However, estimation of parameters λ and β cannot be directly carried out from the composite survival function in Equations (12) and (13). The first difficulty is the survival function in the form of a composite or a two-piece function. The second difficulty, because it contains exponential so that the estimation using the least squares method becomes a complicated job. Besides λ and β, y is also a parameter that needs to be searched for. The optimum value y is determined by the minimum area criteria method.

Estimated λ and β carried out by freeing the survival function of the composite form and free of exponential content. For this reason it is necessary to formulate a cumulative hazard rate function obtained using Equation (2). The results are Equations (14) and (15) respectively for the constant-linear and constant-quadratic models.

$$H_1(t) = -\ln S_1(t) = \begin{cases} \lambda t & 0 \le t < y \\ \lambda t + \frac{\beta}{2}(t-y)^2 & t \ge y \end{cases} \tag{14}$$

$$H_2(t) = -\ln S_2(t) = \begin{cases} \lambda t & 0 \le t < y \\ \lambda t + \frac{\beta}{3}(t-y)^3 & t \ge y \end{cases} \tag{15}$$

3.4 Regression Equations of the Cumulative Hazard Rate Function

To obtain the non-composite form (one piece) of the cumulative hazard rate function in Equations (14) and (15) the indicator function in Equation (16) is used:

$$I(t) = \begin{cases} 0 & 0 \le t < y \\ 1 & t \ge y \end{cases} \tag{16}$$

The cumulative hazard rate function in the non-composite form is given in Equations (17) and (18), namely:

$$H_1(t) = \lambda t + \frac{\beta}{2}(t-y)^2 \cdot I(t) \tag{17}$$

$$H_2(t) = \lambda t + \frac{\beta}{3}(t-y)^3 \cdot I(t) \tag{18}$$

3.5 Empirical Survival Model

The data in Table 1 are used to calculate the empiric survival value, empiric cumulative hazard function and empiric hazard function. Values, and are calculated by Equations (3), (4), and (5) consecutively. The results are given in Table 2.

Table 2. Empirical Survival Values, Empirical Cumulative Hazard Levels and Empirical Hazard Levels

Long Live t (Hours)	Empirical Survival Value $\overline{S}^0(t)$	Empirical Cumulative Hazard Rate Value $\overline{\Lambda}^0(t)$	Empirical Hazard Rate Value $\overline{h}^0(t)$
1,000	49/49 = 1.000	0.000	0.00000
2,000	44/49 = 0.898	0.108	0.10200
3,000	43/49 = 0.878	0.131	0.02227
4,000	42/49 = 0.857	0.154	0.02392
5,000	38/49 = 0.776	0.254	0.09452
6,000	33/49 = 0.673	0.395	0.13273
7,000	30/49 = 0.612	0.491	0.09064
8,000	28/49 = 0.571	0.560	0.06699
9,000	24/49 = 0.490	0.714	0.14186
10,000	21/49 = 0.429	0.847	0.12449
11,000	18/49 = 0.367	1.001	0.14452
12,000	10/49 = 0.204	1.589	0.44414
13,000	9/49 = 0.184	1.695	0.09804
14,000	7/49 = 0.143	1.946	0.22283
15,000	5/49 = 0.102	2.282	0.28671
16,000	4/49 = 0.082	2.506	0.19608
17,000	3/49 = 0.061	2.793	0.25610
18,000	0/49 = 0.000	-	-

3.6 Estimating Parameters in Regression Equations

The cumulative hazard rate function (17) and (18) can be expressed as a regression function with two independent variables and in Equation (6), i.e.

$$c_1 = H_1(t), \quad x_1 = t, \text{ and } x_2 = \tfrac{1}{2}(t-y)^2 \cdot I(t) \tag{19}$$

$$c_2 = H_2(t), \quad x_1 = t, \text{ and } x_2 = \tfrac{1}{3}(t-y)^3 \cdot I(t) \tag{20}$$

Look at Equations (19) and (20) as regression equations. Parameters λ and β can be estimated by the least squares method with Equation (7).

3.7 The Constant-Linear Piecewise Exponential Model

In this section, the equation for the cumulative hazard rate function will be determined $c_1 = \lambda_1 x_1 + \beta_1 x_2$. The function can be expressed as a regression function with two independent variables x_1 and x_2 where $c_1 = H_1(t)$, $x_1 = t$ and $x_2 = \frac{1}{2}(t-y)^2 \cdot I(t)$ so that the regression model obtained in Equation (17). Parameters λ_1 and β_1 can be estimated by the least squares method in Equation (7) using the lamp survival life data in Table 3.

Table 3.Values of Regression Variables in the Constant-Linear Piecewise Exponential

$c_1 = H_1(t)$	$x_1 = t$	$x_2 = \frac{1}{2}(t-y)^2 \cdot I(t)$
0.000	1	0.000
0.108	2	0.000
0.131	3	0.000
0.154	4	0.180
0.254	5	1.280
0.395	6	3.380
0.491	7	6.480
0.560	8	10.580
0.714	9	15.680
0.847	10	21.780
1.001	11	28.880
1.589	12	36.980
1.695	13	46.080
1.946	14	56.180
2.282	15	67.280
2.506	16	79.380
2.793	17	92.480

With the least squares method in Equation (7),

$$\hat{\theta} = \left(X^T X\right)^{-1} X^T Y = \begin{bmatrix} 1785.0000 & 6514.7600 \\ 6514.7600 & 27748.7196 \end{bmatrix}^{-1} \begin{bmatrix} 228.8430 \\ 926.2955 \end{bmatrix} = \begin{bmatrix} 0.0445 \\ 0{,}0229 \end{bmatrix}$$

obtained $\hat{\lambda}_1$ = 0.0445 and $\hat{\beta}_1$ = 0,0229. The survival model is given in Equation (21):

$$S(t) = \begin{cases} \exp(-0.0445 \cdot t) & 0 \le t < y \\ \exp\left[-0.0445 \cdot t - \dfrac{0.0229}{2}(t-y)^2\right] & t \ge y \end{cases} \tag{21}$$

3.8 Goodness of Fit Test

Next, we will examine whether the data follows the survival function in Equation (21). From the data, we obtained the highest and lowest values respectively equal to 0 and 18,000 hours so the range of data is 18,000. The number of interval classes is determined by the Sturgess rule $k = 1 + 3.3 \log 49$ = 6.578 = 7. The width of the interval class is the ratio between the range and the number of interval classes, the width of each interval class is = 2,600. The test statistic used is the Chi-square test in Equation (8). The test requires data E_i that is calculated by Equations (22) - (25). Calculation results E_i and Chi-square calculations are given in Table 4.

$$E_1 = n \int_{bb_1}^{ba_1} \lambda \cdot exp(-\lambda) dt \tag{22}$$

$$E_2 = n \int_{bb_2}^{y} \lambda \cdot exp(-\lambda) dt + n \int_{y}^{ba_2} \left(\lambda + \beta(t-y) exp\left(-\lambda t - \frac{\beta}{2}(t-y)^2 \right) \right) dt \tag{23}$$

$$E_i = n \int_{bb_i}^{ba_i} \left(\lambda + \beta(t-y) exp\left(-\lambda t - \frac{\beta}{2}(t-y)^2 \right) \right) dt \ ; \ i = 3,4,\ldots,m-1 \tag{24}$$

$$E_m = \lim_{a \to \infty} n \int_{bb_m}^{a} \left(\lambda + \beta(t-y) exp\left(-\lambda t - \frac{\beta}{2}(t-y)^2 \right) \right) dt \tag{25}$$

where

n : sample number
ba : the upper limit of the class interval
bb : the lower limit of the class interval
m : the number of class interval
a : upper limit of the last class interval
y : the value calculated by minimum criteria area; and $y = 3.4$ (London, 1997)
λ, β : regression parameters

Estimated regression parameters are $\hat{\lambda}_1$ = 0.0445 and $\hat{\beta}_1$ = 0.0229. Here are some calculations E_i, with complete results available in Table 4.

$$E_1 = 49 \int_0^{2.6} 0.0445 \cdot exp(-0.445) dt = 5.3$$

$$E_2 = 49 \int_{2.6}^{3.4} 0.044 \cdot exp(-0.044) dt + 49 \int_{3.4}^{5.2} \left(0.044 + 0.0229(t-3.4) exp\left(-0.044t - \frac{0.0229}{2}(t-3.4)^2 \right) \right) dt = 6.1$$

$$E_3 = 49 \int_{5.2}^{7.8} \left(0.044 + 0.0229(t-3.4) exp\left(-0.044t - \frac{0.0229}{2}(t-3.4)^2 \right) \right) dt$$

Let $u = 0.044t + 0.0229(t-3.4)^2$, then $\frac{du}{dt} = 0.044t + \frac{0.0229}{2}(t-3.4)$ such that

$$E_3 = 49 \int_{5.2}^{7.8} \left(\frac{du}{dt} exp(-u) \right) dt = 49 \int_{5.2}^{7.8} (exp(-u)) du = 9{,}5$$

Table 4.Goodness of Fit Test for Constant-Linear Piecewise Exponential Combinations

Interval ($\times 1000$)	O_i	E_i	$\chi^2 = \frac{(O_i - E_i)^2}{E_i}$
0 – 2.6	6	5.3	0.092
2.6 – 5.2	5	6.1	0.198
5.2 – 7.8	9	9.5	0.026
7.8 – 10.4	10	10.0	0
10.4 – 13.0	10	8.0	0.5
13.0 – 15.6	5	5.2	0.008
> 15.6	4	4.9	0.165
Jumlah	49	49	0.989

From the calculations, we get a Chi-square value of 0.989. Therefore $\chi^2 = 0{,}989 < 11.0705$ then H_0 is accepted. Thus, the hazard rate function and cumulative hazard rate function for light bulb endurance can be modeled by Equations (26) and (27):

$$h_1(t) = \begin{cases} 0.0445 & 0 \le t < y \\ 0.0445 + 0.0229 \cdot (t - y) & t \ge y \end{cases} \quad (26)$$

$$H_1(t) = 0.0445 \cdot t + \frac{0.0229}{2}(t - y)^2 . I(t) \quad (27)$$

where $y = 3.4$

3.9 Constant-Quadratic Piecewise Exponential Model

In this section, the equation for the cumulative hazard rate function in the form will be determined $c_2 = \lambda_2 x_1 + \beta_2 x_2$. The function can be said as a regression function with two independent variables and by taking x_1 and x_2 where $c_2 = H_2(t)$, $x_1 = t$ and $x_2 = \frac{1}{3}(t - y)^3 \cdot I(t)$ so that the regression model obtained in Equation (9). Regression parameters λ_2 and β_2 can be estimated by the least squares method in Equation (7) using the lamp survival life data in Table 5.

Table 5.Values of Regression Variables in Constant-Quadratic Piecewise Exponential

$c_2 = H_2(t)$	$x_1 = t$	$x_2 = \frac{1}{3}(t - y)^3 \cdot I(t)$
0.000	1	0
0.108	2	0.000
0.131	3	0.000
0.154	4	0.072
0.254	5	1.365
0.395	6	5.859
0.491	7	15.552
0.560	8	32.445
0.714	9	58.539
0.847	10	95.832
1.001	11	146.325
1.589	12	212.019
1.695	13	294.912
1.946	14	397.005
2.282	15	520.299
2.506	16	666.792
2.793	17	838.485

With the least squares method in Equation (7)

$$\hat{\theta} = \left(X^T X\right)^{-1} X^T Y = \begin{bmatrix} 1{,}785.000000 & 48{,}168.99733 \\ 48{,}168.99733 & 1{,}743{,}270.574 \end{bmatrix}^{-1} \begin{bmatrix} 228.8428650 \\ 7107.469746 \end{bmatrix} = \begin{bmatrix} 0.071479931 \\ 0.002101999 \end{bmatrix}$$

obtained $\hat{\lambda}_2 = 0.0715$ and $\hat{\beta}_2 = 0.0021$. The survival model is expressed in Equation (28):

$$S(t)=\begin{cases} exp(-0.0715\cdot t) & 0\le t<y \\ exp\left[-0.0715\cdot t-\frac{0.0021}{3}(t-y)^3\right] & t\ge y \end{cases} \tag{28}$$

Next will be tested whether the data follows the survival function in Equation (23). The test statistic used is the Chi-square test in Equation (8). The test requires data E_i which is calculated by Equation (12) - (15). The calculation results E_i and the Chi-square count is given in Table 6.

Table 6.Goodness of Fit Test for Constant-Quadratic Piecewise Exponential Combinations

Interval	O_i	E_i	$\frac{(O_i-E_i)^2}{E_i}$
0 – 2.6	6	8.3	0.637
2.6 – 5.2	5	7.0	0.571
5.2 – 7.8	9	7.2	0.450
7.8 – 10.4	10	8.1	0.446
10.4 – 13.0	10	7.9	0.558
13.0 – 15.6	5	5.9	0.137
> 15.6	4	4.6	0.078
Jumlah	49	49	2.877

From the calculations, the Chi-square value of 2.877 was obtained. Therefore, $\chi^2=2{,}877<11.0705$ then H_0 is accepted. Thus, the hazard rate function and cumulative hazard rate function for light bulb endurance can be modeled by Equations (29) and (30):

$$h_2(t)=\begin{cases} 0.0715 & 0\le t<y \\ 0.0715+0.0021(t-y)^2 & t\ge y \end{cases} \tag{29}$$

$$H_2(t)=0.0715\cdot t+\frac{0.0021}{3}(t-y)^3\cdot I(t) \tag{30}$$

where $y=3.4$

3.10 Selection of the Best Cumulative Hazard Rate and Hazard Rate Functions

London (1997) and Riaman et al (2018) produced a constant-quadratic piecewise constant rate hazard model in Equations (31) and (32).

$$h_3(t)=\begin{cases} 0.049 & 0\le t<y \\ 0.049+0.0038(t-y)^2 & t\ge y \end{cases} \tag{31}$$

$$H_3(t)=0.049\cdot t+\frac{0.038}{3}(t-y)^3.I(t) \tag{32}$$

where $y=3.4$

In this section, we compare the function of hazard rate and cumulative hazard rate using MAPE in Equation (9). The purpose of the comparison is to choose the best comparison of the cumulative hazard rate and hazard rate functions. Table 7 lists the values for the empiric hazard rate function $\bar{h}_0(t)$ and the predictive hazard rate function, namely $h_1(t)$, $h_2(t)$ and $h_3(t)$, as well as the empiric cumulative hazard rate function $\bar{\Lambda}^0(t)$ and the predictive cumulative hazard rate function, namely $H_1(t)$, $H_2(t)$ and $H_3(t)$.

Table 7.Empirical and Predictive Hazard Rate and Hazard Rate Function Value

t	$\bar{h}_0(t)$	$h_1(t)$	$h_2(t)$	$h_3(t)$	$\bar{\Lambda}^0(t)$	$H_1(t)$	$H_2(t)$	$H_3(t)$
	Eq. (5)	Eq. (26)	Eq. (29)	Eq. (31)	Eq. (4)	Eq. (27)	Eq. (30)	Eq. (32)
1000	0.00000	0.071	0.044	0.049	0.000	0.071	0.044	0.049
2000	0.10200	0.071	0.044	0.049	0.108	0.142	0.088	0.098
3000	0.02227	0.071	0.044	0.049	0.131	0.213	0.132	0.137
4000	0.02392	0.176	0.07172	0.050368	0.154	0.2156	0.284151	0.196274
5000	0.09452	0.396	0.07612	0.058728	0.254	0.5016	0.357867	0.250190
6000	0.13273	0.616	0.08452	0.074688	0.395	1.0076	0.438303	0.316269
7000	0.09064	0.836	0.09692	0.098248	0.491	1.7336	0.529659	0.402113
8000	0.06699	1.056	0.11332	0.129408	0.560	2.6796	0.636135	0.515325
9000	0.14186	1.276	0.13372	0.168168	0.714	3.8456	0.761931	0.663505
10000	0.12449	1.496	0.15812	0.214528	0.847	5.2316	0.911247	0.854257
11000	0.14452	1.716	0.18652	0.268488	1.001	6.8376	1.088283	1.095183
12000	0.44414	1.936	0.21892	0.330048	1.589	8.6636	1.297239	1.393883
13000	0.09804	2.156	0.25532	0.399208	1.695	10.7096	1.542315	1.757961
14000	0.22283	2.376	0.29572	0.475968	1.946	12.9756	1.827711	2.195017
15000	0.28671	2.596	0.34012	0.560328	2.282	15.4616	2.157627	2.712655
16000	0.19608	2.816	0.38852	0.652288	2.506	18.1676	2.536263	3.318476
17000	0.25610	3.036	0.44092	0.751848	2.793	21.0936	2.967819	4.020083
18000	-	-	-	-	-	-	-	-

From these results, the best predictive hazard rate function $h_2(t)$ is the best and the predictive cumulative hazard rate function is $H_2(t)$

Table 8. MAPE Calculation Results

	$h_1(t)$	$h_2(t)$	$h_3(t)$	$H_1(t)$	$H_2(t)$	$H_3(t)$
Value of MAPE	882%	74%	108%	575%	9%	21%

4. Conclusions

The constant-linear and constant-quadratic piecewise hazard rate functions are given in equations (10) and (11). The form of the exponential survival function is piecewise a linear constant and a quadratic constant is given in equations (12) and (13). Furthermore, cumulative hazard rate function is obtained for the constant-linear and constant-quadratic piecewise exponential model in Equations (14) and (15).

Based on the data used the cumulative hazard rate and hazard rate function equations are obtained in Equations (26) and (27) and (29) and (30). Previously, London (1997) and Riaman (2018) had obtained results in the form of cumulative hazard rate and hazard rate functions in Equations (31) and (32). Based on the MAPE criteria the best hazard rate function is

$$h_2(t) = \begin{cases} 0.0715 & 0 \le t < y \\ 0.0715 + 0.0021(t-y)^2 & t \ge y \end{cases}$$

and the best cumulative hazard rate function is

$$H_2(t) = 0.0715 \cdot t + \frac{0.0021}{3}(t-y)^3 \cdot I(t)$$

where $y = 3.4$

References

Bouaziz, O. and Nuel, G (2017). L0 Regularisation for the Estimation of Piecewise Constant Hazard Rates in Survival Analysis. *Applied Mathematics, Scientific Research Publishing*. Vol. 8, No. 3, pp. 377-394.

Bowers, N. K., Gerber, H. U., Hickman, J. C., Jones, D. A., and Besbitt, C. J. (1997). *Actuarial Mathematics*. USA: The Society of Actuaries.

Chang, P., Wang, Y., and Liu, C. (2007). The Development of a Weight Evolving Fuzzy Neural Network for PCB Sales Forecasting. *Expert Systems with Applications*. Vol. 32, pp. 86-96.

Kubo, J., Cullen, M. R., Cantley, L., Slade, M., Tessier-Sherman, B., Taiwo, O., and Desai, M. (2013). Piecewise Exponential Models to Assess the Influence of Job-specific Experience on the Hazard of Acute Injury for Hourly Factory Workers. BMC Medical Research Methodology. Vol. 13, No. 89, pp. 1-10.

London, D. 1997. *Survival Models and Their Estimation*. Connecticut: Actex Publication.

Pottmann, H., Hamaan, B., Krasauskas, R., and Joy, K. (2000). On Piecewise Linear Approximation of Quadratic Functions. *Journal for Geometry and Graphics*. Vol. 4, No. 1, pp. 31-53.

Prabowo, A., Budiono, R., Atmanto, F.D., Tripena, A., Sukono. (2020). Ridge Regression Model Application in Case Human Development Index in Banten province. *International Journal of Advanced Science and Technology*. Vol. 29, No. 6, pp. 3221-3233.

Riaman, Lesmana, E., Subartini, B., and Supian, S. (2018). Time Prediction of Failure a Type of Lamps by Using General Composite Hazard Rate Model. *IOP Conf. Series: Material Science and Engineering*. 332 (2018) 012041.

Sari, I. M., Thamrin, S. A., and Lawi, A. (2014). Piecewise Exponential Frailty Model on Survival Data using Bayesian Approach. *Proceeding of The First International Conference on Science (ICOS-1)*. pp. 199-208.

Weiss, J. C. (2017). Piecewise-constant Parametric Approximations for Survival Learning. Proceedings of Machine Learning for Healthcare 2017.

Acknowledgements

Acknowledgments are conveyed to the Rector, Director of Directorate of Research and Community Service, and the Dean of Faculty of Mathematics and Natural Sciences, Universitas Jenderal Soedirman.

Biographies

Agung Prabowo is the staff of the Department of Mathematics, Universitas Jenderal Soedirman. His fields of research are financial mathematics, survival model analysis and ethno-mathematics.

Ibnu Ginanjar Susilo has been graduated from Department of Mathematics, Universitas Jenderal Soedirman. His field of research is survival model analysis.

Agus Sugandha is the staff of the Department of Mathematics, Universitas Jenderal Soedirman. His field of research is number theory.

Mustafa Mamat is a lecturerin the Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Malaysia. Currently serves as Dean of Graduate School Universiti Sultan Zainal Abidin, Terengganu, Malaysia. The field of applied mathematics, with a field of concentration of optimization.

Sukono is a lecturer in the Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran. Currently as Chair of the Research Collaboration Community (RCC), the field of applied mathematics, with a field of concentration of financial mathematics and actuarial sciences.

Abdul Talib Bon is a professor of Production and Operations Management in the Faculty of Technology Management and Business at the Universiti Tun Hussein Onn Malaysia since 1999. He has a PhD in Computer Science, which he obtained from the Universite de La Rochelle, France in the year 2008. His doctoral thesis was on topic Process Quality Improvement on Beltline Moulding Manufacturing. He studied Business Administration in the Universiti Kebangsaan Malaysia for which he was awarded the MBA in the year 1998. He's bachelor degree and diploma in Mechanical Engineering which his obtained from the Universiti Teknologi Malaysia. He received his postgraduate certificate in Mechatronics and Robotics from Carlisle, United Kingdom in 1997. He had published more 150 International Proceedings and International Journals and 8 books. He is a member of MSORSM, IIF, IEOM, IIE, INFORMS, TAM and MIM.