Indonesian Journal of Pure and Applied Mathematics
http://journal.uinjkt.ac.id/index.php/inprime
Vol. 2, No. 1 (2020), pp. 33–40
P-ISSN 2686-5335, E-ISSN 2716-2478
doi: 10.15408/inprime.v2i1.14556

Estimating the Cost of Car Warranty in Indonesia using the Gertsbakh-Kordonsky Method

Anggis Sagitarisman and Aceng Komarudin Mutaqin*
Statistics Study Program, Universitas Islam Bandung
Jl. Ranggagading No. 8 Bandung 40116, Indonesia
Email: anggissagita@gmail.com

Abstract

Car manufacturers in Indonesia need to determine reasonable warranty costs that do not burden companies or consumers. Several statistical approaches have been developed to analyze warranty costs. One of them is the Gertsbakh-Kordonsky method which reduces the two-dimensional warranty problem to one dimensional. In this research, we apply the Gertsbakh-Kordonsky method to estimate the warranty cost for car type A in XYZ company. The one-dimensional data will be tested using the Kolmogorov-Smirnov to determine its distribution and the parameter of distribution will be estimated using the maximum likelihood method. There are three approaches to estimate the parameter of the distribution. The difference between these three approaches is in the calculation of mileage for units that do not claim within the warranty period. In the application, we use claim data for the car type A. The data exploration indicates the failure of car type A is mostly due to the age of the vehicle. The Kolmogorov-Smirnov shows that the most appropriate distribution for the claim data is the three-parameter Weibull. Meanwhile, the estimated using the Gertsbakh-Kordonsky method shows that the warranty costs for car type A are around 3.54% from the selling price of this car unit without warranty i.e. around Rp. 4,248,000 per unit.

Keywords: warranty costs; the Gertsbakh-Kordonsky method; maximum likelihood estimation; Kolmogorov-Smirnov test.

Abstrak

Perusahaan produsen mobil di Indonesia perlu menentukan biaya garansi yang bersifat wajar tidak memberatkan perusahaan maupun konsumen. Beberapa pendekatan statistik telah dikembangkan untuk menganalisis biaya garansi. Salah satunya adalah metode Gertsbakh-Kordonsky yang mereduksi masalah garansi dua dimensi menjadi satu dimensi. Pada penelitian ini, metode Gertsbakh-Kordonsky akan digunakan untuk mengestimasi biaya garansi untuk mobil tipe A pada perusahaan XYZ. Data satu dimensi hasil reduksi diuji kecocokan distribusinya menggunakan uji kecocokan Kolmogorov-Smirnov dan taksiran parameter distribusinya menggunakan metode penaksir kemungkinan maksimum. Ada tiga pendekatan yang digunakan untuk menaksir parameter distribusi. Perbedaan dari ketiga pendekatan tersebut terletak pada perhitungan jarak tempuh untuk unit yang tidak melakukan klaim dalam periode garansi. Sebagai bahan aplikasi, kami menggunakan data klaim unit mobil tipe A. Hasil eksplorasi data menunjukkan bahwa kegagalan mobil tipe A lebih banyak disebabkan karena faktor usia kendaraan. Hasil uji kecocokan distribusi untuk data hasil reduksi menunjukkan bahwa distribusi yang cocok adalah distribusi Weibull 3-parameter. Sementara itu, hasil perhitungan taksiran biaya garansi menunjukan

bahwa taksiran biaya garansi untuk unit mobil tipe A sekitar 3,54% dari harga jual unit mobil tipe A tanpa garansi, atau sekitar Rp. 4.248.000,- per unit.
Kata Kunci: biaya garansi; metode Gertsbakh-Kordonsky; penaksiran kemungkinan maksimum; uji Kolmogorov-Smirnov.

1. INTRODUCTION

The almost type of manufactured goods are sold with a warranty to provide the consumer's protection. This compensation will be given when the failure occurred in the early time after purchase. It will increase the production cost that is related to warranty service. The warranty cost depends on characteristic, condition, and reliability of the product [1]. One of the manufactured goods that are sold with the warranty cost is the car. All car manufacturers offer a car warranty on their new to protect buyers from costly repairs due to manufacturer defects in materials or workmanship.

Generally, the car warranty depends on two dimensions (2D) i.e. age and usage. For example, car A is sold with a 5-years warranty or 150.000 km, whichever comes first. The buyers have the right to claim in the warranty period. To know the age and usage, we need some data for claim requirement: car identity, date of claim, and kilometer (km) reached.

Some statistical approaches have been developed to analyze the warranty cost based on the vehicle age (in years) and vehicle usage (in km). Generally, this approach divided into two types: two-dimensional (2D) and one-dimensional (1D) approaches. The two-dimensional approach does the bivariate analysis i.e. vehicle age and usage [2]. While the 1D approach only does the analysis on one variable that functions in age or usage [3].

One method in the 1D approach is developed by Gertsbakh and Kordonsky in 1998 (GK method in [4]). This method reduces the 2D-warranty problem into 1D. The GK method analyses one variable (for example V) which is a linear function from the age (X) and usage (Y), where X and Y are given a certain weight. Variable V is modeled by a specific distribution. This model is based on the censored data from variable V. Estimated warranty cost is obtained as a multiplication between the average of the total cost of claims per unit and the cumulative distribution function from variable V. In this research, we apply the GK method to estimate the warranty cost for a car in Indonesia. We use the claim data in the year 2012 – 2017 for car type A in XYZ company that is sold in 2012. The car type A sold in the 2D warranty i.e. 5 years and 150,000 km. The usage and age data for this car are reduced into one variable. To find the appropriate distribution, we test this 1D data using the Kolmogorov-Smirnov test and we use the maximum likelihood estimation method to estimate the parameter of the distribution. There are three approaches to estimate the parameter of the distribution. The difference between these three approaches is in the calculation of mileage for units that do not claim within the warranty period. The estimation of warranty cost will be calculated based on the most appropriate distribution.

2. METHOD

Basically, the warranty policy is categorized into two, namely one-dimensional policy (1D) and two-dimensional policy (2D) [5]. The 1D warranty policies characterized by a one-dimensional interval called warranty period (example years or km). The 2D warranty is an extension of the 1D warranty that is characterized by a rectangular area in the two-dimensional region with a horizontal axis represent time (or age) and the vertical axis represents vehicle usage (Figure 1). This figure describes the example of the 2D warranty region between time and vehicle usage. This figure shows that if the rate of vehicle usage is low, the warranty (first type) then the warranty can be finished until 5 years. However, if the rate of vehicle usage is high (second type) then the warranty can be finished before 150,000 km.

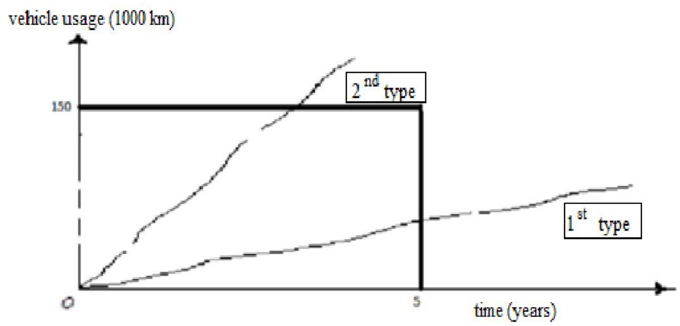

Figure 1. The illustration for the 2D warranty region.

In this research, we use secondary data from car manufacturers in Indonesia (called XYZ company) which are recorded in the year 2012 until 2017. This data describes the number of units sold for car type A in 2012 (Table 1). This type is sold with 2D warranty: 5 years and 150,000 km. We need two data to estimate the warranty cost i.e. claim and purchased data. Claim data in XYZ company consist of chassis number, engine number, the date of purchase or delivery (depends on the contract), the date of vehicle service, km reached, and vehicle's components claimed. In 2012, there are 868 units make claim and 6,141 units which don't claim.

Table 1. The number of units sold for car type A in 2012 in XYZ company.

Semester	The number of units sold
1	3,890
2	3,119
Total	7,009

Source: XYZ company (2018).

Gertsbakh-Kodonsky Method

GK method reduces the 2D problem into the 1D formulation. The 2D warranty provides coverage in $\Omega=[0,W)\times[0,U)$ where W is the limit of the warranty period in time (example: days, months, semesters, or years), U is the limit of vehicle's usage (example: km). Therefore, the failure will coverage in warranty if $(X,Y)\in\Omega$ where X is an age when the failure occurred and Y is the usage when the failure occurred. GK method reduces two variables (X and Y) into one variable (V). The realization i^{th} of this variable is

$$v_i = (1-\varepsilon)x_i + \varepsilon y_i; i = 1, 2, \cdots, K, \tag{1}$$

where $0 \leq \varepsilon \leq 1$ determined by minimizing the coefficient of sample variance,

$$CV = \frac{S_v}{\bar{v}}, \tag{2}$$

where $\overline{v} = \frac{\sum_{i=1}^{K} v_i}{K}$ and $S_v = \sqrt{\frac{\sum_{i=1}^{K}(v_i - \bar{v})^2}{K-1}}$ represent the average of sample and standard deviation for sample v in equation (2), respectively. Generally, Weibull distribution [6] [7] and lognormal [8] [9] are appropriate to data *v*. We will use data V to analyze the car warranty.

Parameter Estimation

In general, the warranty claim data is censored because not all buyers make claims within the warranty period. Only failure data are available when the units are claimed. Let *K* be the number of claims within the warranty period and *J* be the number of the unit which aren't claimed. Assume that $V_1, V_2, \ldots, V_K$ are the random variables that are independent of the failure time from the units within the warranty period on the V scale.

There are three approaches in determining the value of random variable V for the units are claimed in the warranty period. In the first approach, follow the next steps:

(1) Determine the unit age when censoring: w_j for $j = 1, 2, \cdots, J$. These w_js will have the same value with warranty period limit in the time scale for every unit which doesn't claim within the warranty period.

(2) Calculate the usage rate for claim data using:

$$r_i = \frac{y_i}{x_i}, i = 1,2, \ldots, K.$$

(3) Calculate the average of usage rate using:

$$\bar{r} = \frac{\sum_{i=1}^{K} r_i}{K}, i = 1,2, \ldots, K.$$

(4) Calculate the usage when censoring: u_j for $j = 1, 2, \cdots, J$ using:

$$u_j = \bar{r}w_j, j = 1, 2, \cdots, J.$$

These u_js will have the same value for every unit which doesn't claim within the warranty period.

(5) Calculate the value of V for units which doesn't claim within warranty period using:

$$\tilde{v}_j = (1-\varepsilon)u_j + \varepsilon w_j, j = 1, 2, \cdots, J. \tag{3}$$

Therefore, the data of defect unit which claim within the warranty period are $v_1, v_2, \ldots, v_K$ and unit data which doesn't make a claim (censored data) are $\tilde{v}_j$ for $j = 1,2, \ldots, J$. The warranty cost will be estimated based on these data.

The second approach has the same step with the previous but in step (3) we calculate the median usage rate. In the third approach, step (4) the u_j's value is determined by warranty period limit in time scale. Blischke and Murthy [4] apply all of these three approaches for motorcycle warranty cost. In that research, Blischke and Murthy [4] chose the most realistic to estimate the warranty cost.

Let $f_V(v;\theta)$ and $F_V(v;\theta)$ represent the density function and distribution function for random variable V, respectively. The likelihood function for $v_1, v_2, \ldots, v_K$, is $\prod_{i=1}^{K} f_V(v_i)$ and the likelihood function for $\tilde{v}_j$, $j = 1,2, \ldots, J$ is $\prod_{j=1}^{J}[1 - F_V(\tilde{v})]$. Then the likelihood function from these random variable is:

$$L(\theta) = \prod_{i=1}^{K} f_V(v_i) \prod_{j=1}^{J} [1 - F_V(\tilde{v})],$$

and the log-likelihood function is

$$log[L(\theta)] = \sum_{i=1}^{K} ln[f_V(v_i)] + \sum_{j=1}^{J} ln[1 - F_V(\tilde{v})]. \quad (4)$$

The θ maximum will be given by solving equation (4).

Warranty Cost Estimation

Warranty cost estimation needs the other costs i.e. cost related to defect item (c_s), cost related to the items that cannot be repaired (c_n), dan cost related to the items that can be repaired (c_r). The c_n is the total of the average costs required by the manufacturer to provide the substitute items including all indirect costs such as design, marketing, and others. The c_r is the average of repaired costs required by the manufacturer include parts and labor. Estimated cost per unit of warranty is multiplication between c_s and the unit proportion from the expected to fail within the warranty period $(F_V(\cdot))$, i.e.

$$E[C(\tilde{v})] = c_s F_V(\tilde{v}),$$

where $F_V(\cdot)$ is the cumulative distribution function for the random variable V. If the c_s is unknown then $F_V(\tilde{v})$ can be interpreted of the warranty cost estimation in percentage from the selling price without the warranty.

3. RESULTS AND DISCUSSIONS

Based on data in Table 1, we can plot age vs usage from the claim data for car type A in XYZ company (Figure 2). In this figure, we can see that the distribution of car type A that failed more was due to the age factor. To apply the GK method, we need to reduce the age and usage data of the car Type A using equation (1). The ε value is chosen such that minimizes the sample variance coefficient of the random variable V. We got $\varepsilon = 0.0107$ by using the Solver facility in Microsoft Excel 2016 with the minimum of the sample variance coefficient is 0.0952.

The next step is determining some values of random variable V for the unit of car unit type A so we can apply all of the three approaches above. The age of each car unit type A which doesn't claim within the warranty period is the limit of the warranty period in time scale (per 5years), i.e. $w_j = 1$, for $j = 1, 2, \cdots, 6{,}141$.

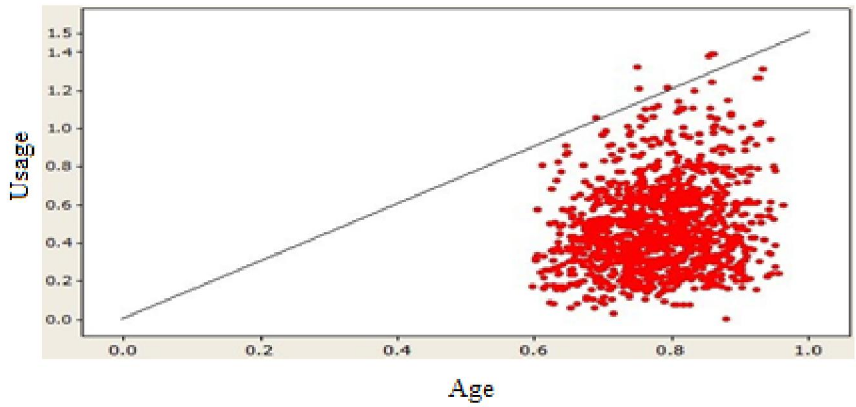

Figure 2. Plot age vs usage from the claim data for car type A.

In the first approach, the travel distance when censoring for each unit that doesn't claim within the warranty period in step (4) is $u_j = \bar{r}w_j = 0.6345 \times 1 = 0.6345$, for $j = 1, 2, \cdots, 6{,}141$. The value of random variable V for car unit type A which doesn't claim within the warranty period based on equation (3) is $\tilde{v}_j = (1-\varepsilon)u_j + \varepsilon w_j = (1 - 0.0107) \times 0.6345 + 0.0107 \times 1 = 0.6384$ for $j = 1, 2, \cdots, 6{,}141$. In the second approach, we get $u_j = mw_j = 0.5976 \times 1 = 0.5976$ for $j = 1, 2, \cdots, 6{,}141$. The value of random variable V using equation (3) is $\tilde{v}_j = (1-\varepsilon)u_j + \varepsilon w_j = (1 - 0.0107) \times 0.5976 + 0.0107 \times 1 = 0.6019$ for $j = 1, 2, \cdots, 6{,}141$. While in the last approach, the travel distance when censoring for each unit that doesn't claim within the warranty period is calculated based on the warranty period in usage scale (per 100,000 km) i.e. $u_j = 1.5$ for $j = 1, 2, \cdots, 6{,}141$. The value of random variable V using equation (3) is $\tilde{v}_j = (1-\varepsilon)u_j + \varepsilon w_j = (1 - 0.0107) \times 1.5 + 0.0107 \times 1 = 1.4947$ for $j = 1, 2, \cdots, 6{,}141$.

The result of Kolmogorov-Smirnov test, the most appropriate distribution for V is three-parameter Weibull i.e. almost all of the probability plots are between the confidence interval of the three-parameter Weibull distribution (Figure 3). The parameter estimation for this distribution based on joint data V (the units make claim and don't claim) for the three approaches are shown in Table 2. We use software Mathcad Prime 3.1 in these estimations. Based on this table, we can see that there is no solution to the third approach. Therefore, the estimation of warranty cost will be conducted using the first and second approaches.

Table 3 shows the value of the cumulative distribution function for the three-parameter Weibull for V on $v = \tilde{v}$. Based on this table, $F_V(\tilde{v}) = 0.035$. It means that the unit proportion from the expected to fail within the warranty period is 3.54% for the first approach and 0.3% for the second approach from the total of the number of the sold car. Refer to Blischke dan Murthy [4], the realistic result is the first approach. In Blischke dan Murthy [4], the unit proportion from the expected to fail within the warranty period is 4.1% from the total of the number of the sold car. Therefore, we use the first approach to estimate the warranty cost. The selling price for car type A from the XYZ manufacturer is Rp. 120 million so the warranty cost is Rp. $120 \text{ million} \times 3.54\% = \text{Rp.} 4.248$ million.

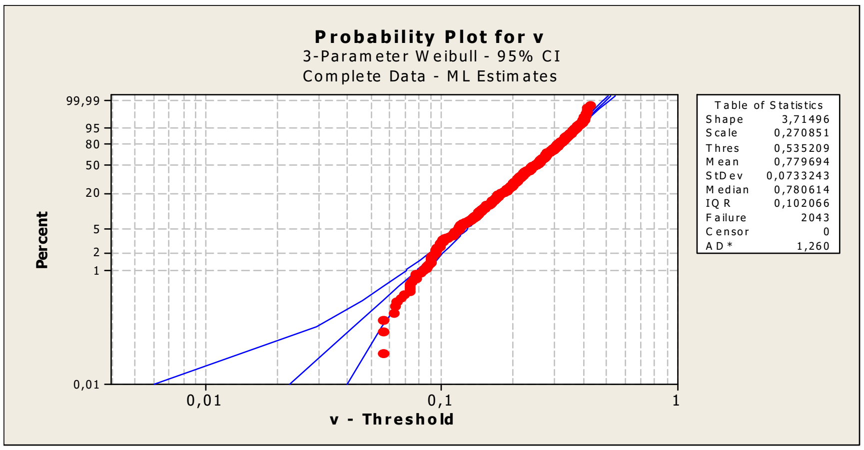

Figure 3. The probability plot for the three-parameter Weibull.

Table 2. The parameter estimation for the three-parameter Weibull distribution.

Parameter	Approach		
	I	II	III
$\hat{\alpha}$	2.295	3.219	-
$\hat{\beta}$	0.206	0.242	-
$\hat{\gamma}$	0.590	0.562	-

Table 3. The value of cumulative distribution function for the three-parameter Weibull

	First Approach	Second Approach
$\tilde{v}$	0.6384	0.6019
$F_V(\tilde{v})$	0.0354	0.0030

4. CONCLUSIONS

In this research, we apply the Gertsbakh-Kordonsky method to estimate the warranty estimate for car type A from the XYZ company. We use the claim data in 2012 – 2017 and selling data in 2012. There are three approaches to estimate the unit proportion from the expected to fail within the warranty period. However, only the first and second approaches obtain the solutions i.e. 3.54% for the first approach and 0.3% for the second approach from the total of the number of the sold car. Refer to the previous research, we choose the unit proportion from the expected to fail within the warranty period is 3.54% from the total of the number of the sold car. Because the selling price for the car type A is Rp. 120 million then the warranty cost for this car is Rp. 4,248,000.

REFERENCES

[1] A. K. Mutaqin and Komarudin, "Perhitungan pemi untuk asuransi kendaraan bermotor berdasarkan sejarah frekuensi klaim pemegang polis menggunakan analisis bayes," *PYTHAGORAS,* vol. 4, no. 1, pp. 47-55, 2008.

[2] B. P. Iskandar, Modelling and analysis of two-dimensional warranty policies, Queensland: School of Engineering, The University of Queensland, 1992.

[3] A. K. Mutaqin, Estimasi ongkos garansi sepeda motor yang dijual dengan garansi dua dimensi: studi kasus di PT. X, Bandung: Industrial Engineering and Management-FTI. Institut Teknologi Bandung, 2000.

[4] W. R. Blischke and D. N. Murthy, Case Studies in Reliability and Maintenance., New Jersey: John Wiley & Sons, Inc., 2003.

[5] W. R. Blischke and D. N. Murthy, Warranty Cost Analysis., New York: Marcel Dekker, Inc., 1994.

[6] D. Cousineau, "Fitting the Three-Parameter Weibull Distribution: Review and Evaluation of Existing and New Methods," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 16, no. 1, pp. 281-288, 2009.

[7] M. Ramakrishnan and N. Viswanathan, "Comparing the Methods of Estimation of Three-Parameter Weibull Distribution," *IOSR Journal of Mathematics,* vol. 13, no. 1, pp. 42-45, 2017.

[8] R. J. Aristizabal, Estimating the Parameters of Three-Parameter Lognormal Distribution, Florida International University: Florida: College of Art and Sciences, 2012.

[9] A. Goodarzi, A. Amiri and S. Asadzadeh, "Monitoring the Censored Lognormal Reliability Data in a Three-Stage Process using AFT Model," *Journal of Industrial Engineering International,* vol. 13, p. 67–80, 2017.