

# **Application of Product-Limit and Nelson-Aalen Methods in Health Insurance for Estimating Mean and Variance of Infected Duration of Dengue Fiver**

**Sukono, Dwi Susanti, Alit Kartiwa, Rani Q. F.**

Department of Mathematics, Faculty of Mathematics and Natural Sciences,  
Universitas Padjadjaran, Indonesia  
sukono@unpad.ac.id, dwi.susanti@unpad.ac.id, alit.kartiwa@unpad.ac.id

**Yuyun Hidayat**

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

**Abdul Talib Bon**

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

## **Abstract**

In health insurance knowing the length of the patient contracting a disease is important, due to the handling and maintenance costs. This paper analyzed the mean and variance estimator duration contracted dengue fever based on data on the number of cases of dengue fever in the Bandung city in January-April 2017. A person suffering from dengue fever symptoms is called Dengue Suspect. This Suspect has 2 opportunities, namely the chance of recovery or positive chance of contracting dengue fever. Opportunities suspect dengue fever to healt was analyzed by using survival model. The standard estimate in survival function was proposed by Kaplan and Meier in 1958, called Product-Limit Estimator. While the probability of suspected dengue fever to be positive contracted dengue is analyzed by using cumulative hazard function based on Nelson-Aalen estimator. Based on the survival function it is found that the time of suspect dengue fever to heal continues to decrease, while suspect dengue fever to become positive contracted dengue fever continues to increase based on cumulative hazard function. From the calculation results obtained average duration of dengue fever is for 8 days after suspect dengue fever, with variance 0.0758. Based on the average estimator and this variance can be used as consideration in the treatment planning and maintenance cost.

## **Keywords:**

Survival model, Kaplan-Meier, Product-Limit, Nelson-Aalen, dengue fever

## **1. Introduction**

Cases of Dengue Fever in Bandung increased. Until December 2016, Bandung Health Department recorded as many as 3,389 cases of dengue fever that occurred throughout Bandung. Efforts from the start of the precautionary measures to the treatment of dengue patients have been done by the government. Treatment of patients generally financing is done through health insurance program (Sukono et al., 2017.a; 2017.b). The problem that often arises is the budget reserved is not sufficient, because one of them is not known exactly how long the patient contracted dengue fever will be treated until healed (Camerron and Trivedi, 2005; Hanif et al., 2015).

To find the solution of the problem, Hanif et al. (2015), conducted an analysis of dengue patients associated with liver dysfunction using survival models. While Kenah (2011), performs endemic data analysis using nonparametric survival model. Meanwhile, Guibert & Planchet (2014), and also Mesike et al. (2012), undertook a nonparametric inference analysis of transition probabilities based on an integral estimator with the Aalen-Johansen model for semi-competing risk data applied to LCT insurance company. Furthermore, Christensen and Kallestrup-Lamb (2012), analyzed the relationship between health insurance coverage and survival of cervical cancer patients.

In 2016, dengue is an Extraordinary Events in the city of Bandung, because of the many positive patients contracted dengue fever since suspected of suffering from dengue fever symptoms. These problems can be analyzed and solved by applying survival model. Based on the above description, this paper discusses the nonparametric model approach of survival function and cumulative hazard function by using Product-Limit and Nelson-Aalen estimators to obtain the average and variance duration of dengue fever in Bandung in January-April 2017. The goal is to get an estimate of the duration of action services for patients with dengue fever.

## 2. Research Method

This section discusses nonparametric estimation methods for data on the number of dengue cases. The discussion covers from: estimation of survival function, estimation of cumulative hazard function, confidential interval for survival models, and estimation of residual life mean.

### 2.1 Estimation of Survival Function

The standard estimate of survival functions was proposed by Kaplan and Meier in 1958, called Product-Limit estimators. Defined by Klein & Moeschberger (1997) and Yen & Kassim (2013), suppose  $X$  random variable time dengue fever outbreak,  $x_i$  the duration of the contract dengue fever in individuals  $i$ , and  $n$  the number of data observations. Suppose that the incidence of dengue occurs at different times  $D$ , that is  $t_1 < t_2 < \dots < t_D$ , and at the time of the incident  $t_i$  there  $d_i$  events. Suppose, too  $Y_i$  states the number of individuals at risk of contracting dengue fever  $t_i$ . The estimator of survival function  $\hat{S}(x)$  is expressed as follows:

$$\hat{S}(x) = \begin{cases} 1 & , x < x_i \\ \prod_{x_i \leq x} \left[ 1 - \frac{d_i}{Y_i} \right] & , x_i \leq x \end{cases} \quad (1)$$

From equation (1), the estimator of variance of the Product-Limit Adjuster  $\hat{\sigma}_S^2(x)$  is expressed as follows:

$$\hat{\sigma}_S^2(x) = \hat{V}[\hat{S}(x)] = \hat{S}(x)^2 \sum_{x_i \leq x} \frac{d_i}{Y_i(Y_i - d_i)} \quad (2)$$

Thus, the standard error of the Product-Limit estimator  $\hat{\sigma}_S(x)$  is:

$$\hat{\sigma}_S(x) = \left\{ \hat{V}[\hat{S}(x)] \right\}^{1/2} \quad (3)$$

The Product-Limit estimate gives the right result in estimating survival function for right sensor data.

### 2.2 Estimation of Cumulative Hazard Function

Other estimates of average cumulative hazards that have small sample sizes are better than estimates based on Product-Limit estimators, first proposed by Nelson in 1972 in the context of reliability and then reiterated by Aalen in 1978b. The cumulative hazard function based on the Nelson-Aalen estimator is defined (Klein and Moeschberger, 1997; Rich et al., 2010). Using notations in Section 2.1, The cumulative hazard function based on the Nelson-Aalen estimator is defined as follows:

$$\hat{H}(x) = \begin{cases} 0 & , x \leq x_i \\ \sum_{x_i \leq x} \frac{d_i}{Y_i} & , x_i \leq x \end{cases} \quad (4)$$

While the variance of the Nelson-Aalen estimator is:

$$\sigma_{\hat{H}}^2(x) = \sum_{x_i \leq x} \frac{d_i}{Y_i^2} \quad (5)$$

### 2.3 Confidential Interval for Survival Models

For  $Z_{1-\alpha/2}$  percentile of standard normal distribution, with confidence level  $\alpha$  %, the confidence interval given for the survival function according to Borgan and Liestøl in 1990 was (Willem, 2014; Karnon et al., 2009):

– Linear

$$\hat{S}(x_0) - Z_{1-\alpha/2} \sigma_S(x_0) \hat{S}(x_0) \leq S(x_0) \leq \hat{S}(x_0) + Z_{1-\alpha/2} \sigma_S(x_0) \hat{S}(x_0) \quad (6)$$

– Transformation-Log

$$\left[ \hat{S}(x_0)^{1/\theta}, \hat{S}(x_0)^\theta \right],$$

where

$$\theta = \exp \left\{ \frac{Z_{1-\alpha/2} \sigma_S(x_0)}{\ln \left[ \hat{S}(x_0) \right]} \right\} \quad (7)$$

– The arcsin square root transformation

$$\sin^2 \left\{ \min \left[ \frac{\pi}{2}, \arcsin \left( \hat{S}(x_0)^{1/2} \right) - 0.5 Z_{1-\alpha/2} \sigma(x_0) \left( \frac{\hat{S}(x_0)}{1 - \hat{S}(x_0)} \right)^{1/2} \right] \right\} \leq S(x_0) \leq \sin^2 \left\{ \min \left[ \frac{\pi}{2}, \arcsin \left( \hat{S}(x_0)^{1/2} \right) + 0.5 Z_{1-\alpha/2} \sigma(x_0) \left( \frac{\hat{S}(x_0)}{1 - \hat{S}(x_0)} \right)^{1/2} \right] \right\} \quad (8)$$

### 2.4 Estimation of Residual Life Mean

Estimated mean values are limited to intervals  $[0, \tau]$ , with  $\tau$  is the longest observation time, given by (Nasir and Zaidi, 2009; Klein and Moeschberger, 1997):

$$\hat{\mu}_\tau = \int_0^\tau \hat{S}(x) dx \quad (9)$$

While the variance is:

$$\hat{\sigma}_{\mu_\tau}^2(x) = \hat{V}[\hat{\mu}_\tau] = \sum_{i=1}^D \left[ \int_{x_i}^\tau \hat{S}(x) dx \right]^2 \frac{d_i}{Y_i(Y_i - d_i)} \quad (10)$$

Thus, the standard error is:

$$\hat{\sigma}_{\mu_\tau}(x) = \left\{ \hat{V}[\hat{\mu}_\tau] \right\}^{1/2} \quad (11)$$

## 3. Result and Discussion

In the results and discussion section, a discussion of object data, survival estimation, and estimated hazard function. Therefore, the discussion begins with the object data problem as follows.

### 3.1 Object Data

In this section we will estimate the survival function and cumulative hazard function for the duration of dengue fever based on data on the number of dengue cases in Bandung in January-April 2017.

Suppose  $x_i$  is the duration of the patient contracted dengue fever,  $d_i$  is the number of patients who are positively infected with dengue, and  $Y_i$  is the number of patients at risk of contracting dengue fever. Data on the number of dmam cases in Bandung in January-April 2017 can be seen in Table 1 follows:

Table 1. Number of dengue cases in Bandung from January to April 2016

Times ( $x_i$ )	Number of Cases DB ( $d_i$ )	Number of Infection Risk DB ( $Y_i$ )
0	1	1000
1	19	999
2	83	872
3	67	735
4	90	618
5	76	505
6	66	405
7	53	325
8	44	263
9	46	204
10	25	153
11	24	125
12	19	99
13	16	78
14	8	61
15	11	51
16	5	40
17	3	35
18	3	32
19	1	29
20	3	26
21	2	23
23	2	21
30	4	19
31	1	15
32	4	13
33	3	9
34	1	6
36	1	5
42	2	4
45	1	2
49	1	1

Then we will calculate  $\hat{S}(x)$  using equation (1). The next step calculates the variance of  $\hat{S}(x)$  using equation (2). The calculation results  $\hat{S}(x)$  and  $\hat{V}[\hat{S}(x)]$  for  $x_i \leq x$  can be seen in Table 2 below:

Table 2. Calculation Results  $\hat{S}(x)$  and  $\hat{V}[\hat{S}(x)]$

Time Event	Estimator Product- Limit	Variance	Standard Error
$0 \leq x < 1$	0.9990	0.000001	0.0010
$1 \leq x < 2$	0.9800	0.000020	0.0044
$2 \leq x < 3$	0.8867	0.000111	0.0105
$3 \leq x < 4$	0.8059	0.000180	0.0134
$4 \leq x < 5$	0.6885	0.000262	0.0162
$5 \leq x < 6$	0.5849	0.000309	0.0176
$6 \leq x < 7$	0.4896	0.000332	0.0182
$7 \leq x < 8$	0.4097	0.000333	0.0183
$8 \leq x < 9$	0.3412	0.000320	0.0179
$9 \leq x < 10$	0.2643	0.000292	0.0171
$10 \leq x < 11$	0.2211	0.000266	0.0163
$11 \leq x < 12$	0.1786	0.000235	0.0153
$12 \leq x < 13$	0.1444	0.000203	0.0143
$13 \leq x < 14$	0.1147	0.000172	0.0131
$14 \leq x < 15$	0.0997	0.000154	0.0124

$15 \leq x < 16$	0.0782	0.000128	0.0113
$16 \leq x < 17$	0.0684	0.000115	0.0107
$17 \leq x < 18$	0.0626	0.000106	0.0103
$18 \leq x < 19$	0.0567	0.000098	0.0099
$19 \leq x < 20$	0.0547	0.000095	0.0097
$20 \leq x < 21$	0.0484	0.000086	0.0093
$21 \leq x < 23$	0.0442	0.000080	0.0089
$23 \leq x < 30$	0.0400	0.000073	0.0086
$30 \leq x < 31$	0.0316	0.000060	0.0077
$31 \leq x < 32$	0.0295	0.000056	0.0075
$32 \leq x < 33$	0.0204	0.000041	0.0064
$33 \leq x < 34$	0.0136	0.000029	0.0053
$34 \leq x < 36$	0.0113	0.000024	0.0049
$36 \leq x < 42$	0.0091	0.000020	0.0044
$42 \leq x < 45$	0.0045	0.000010	0.0032
$45 \leq x < 49$	0.0023	0.000005	0.0023

From Table 2 the plot of the survival function can be seen in Figure 1 below:

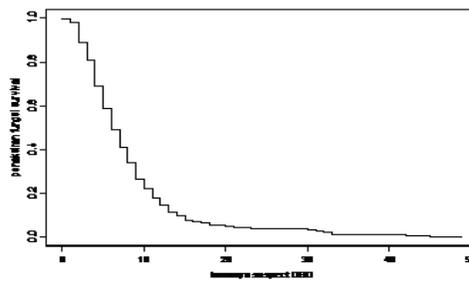


Figure 1. Plot of Survival Function

the time of suspecte patients from Table 2 and Figure 1, it can be seen that the survival function continues to decline, meaning that continues to decline until it becomes positive for dengue fever.

Then we will calculate  $\hat{H}(x)$  using equation (4). The next step calculates the variance of  $\hat{H}(x)$  using equation (5). The calculation results  $\hat{H}(x)$  and  $\sigma_H^2(x)$  for  $x_i \leq x$  can be seen in Table 3 below:

Table 3. Calculation Results  $\hat{H}(x)$  and  $\sigma_H^2(x)$

Time Event	Estimator <i>Nelson-Aalen</i>	Variance	Standard Error
$0 \leq x < 1$	0.0010	0.000001	0.0010
$1 \leq x < 2$	0.0200	0.000020	0.0045
$2 \leq x < 3$	0.1152	0.000129	0.0114
$3 \leq x < 4$	0.2064	0.000253	0.0159
$4 \leq x < 5$	0.3520	0.000489	0.0221
$5 \leq x < 6$	0.5025	0.000787	0.0281
$6 \leq x < 7$	0.6654	0.001189	0.0345
$7 \leq x < 8$	0.8285	0.001691	0.0411
$8 \leq x < 9$	0.9958	0.002327	0.0482
$9 \leq x < 10$	1.2213	0.003432	0.0586
$10 \leq x < 11$	1.3847	0.004500	0.0671
$11 \leq x < 12$	1.5767	0.006036	0.0777
$12 \leq x < 13$	1.7686	0.007975	0.0893
$13 \leq x < 14$	1.9738	0.010605	0.1030
$14 \leq x < 15$	2.1049	0.012755	0.1129
$15 \leq x < 16$	2.3206	0.016984	0.1303

$16 \leq x < 17$	2.4456	0.020109	0.1418
$17 \leq x < 18$	2.5313	0.022558	0.1502
$18 \leq x < 19$	2.6251	0.025488	0.1596
$19 \leq x < 20$	2.6595	0.026677	0.1633
$20 \leq x < 21$	2.7749	0.031115	0.1764
$21 \leq x < 23$	2.8619	0.034895	0.1868
$23 \leq x < 30$	2.9571	0.039430	0.1986
$30 \leq x < 31$	3.1676	0.050511	0.2247
$31 \leq x < 32$	3.2343	0.054955	0.2344
$32 \leq x < 33$	3.5420	0.078624	0.2804
$33 \leq x < 34$	3.8753	0.115661	0.3401
$34 \leq x < 36$	4.0420	0.143439	0.3787
$36 \leq x < 42$	4.2420	0.183439	0.4283
$42 \leq x < 45$	4.7420	0.308439	0.5554
$45 \leq x < 49$	5.2420	0.558439	0.7473

From Table 3 the plot of the cumulative hazard function can be seen in Figure 2:

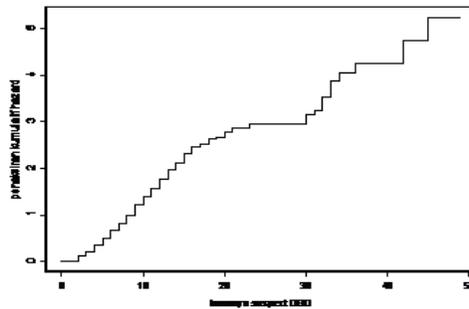


Figure 2. Plot Cumulative hazard function

From Table 3 and Figure 2, it can be seen that the cumulative hazard function continues to increase, meaning that the time of positive patients contracting dengue fever continues to increase.

Then we will calculate the confidence interval based on the value of survival estimation for the duration of the patient with dengue as shown in Table 2. If within 1 week (7 days), the survival function is worth  $S(7) = 0.4896$  and variance  $0.0182^2$ , then:

$$\sigma_S^2(7) = (0.0182/0.4896)^2 = 0.0372^2.$$

The linear confluence interval of 95% in the first week, based on (6) is:

$$0.4896 \pm 1.96 \times 0.0372 \times 0.4896 = (0.4539, 0.5253)$$

The next step is based on (7), the log conversion transform confidence interval in the first week we obtain:

$$\theta = \exp \left[ \frac{1.96 \times 0.0372}{\ln(0.4896)} \right] = 0.9029.$$

So the interval is:

$$\left( 0.4896^{1/0.9029}, 0.4896^{0.9029} \right) = (0.4534, 0.5248).$$

The confound interval of 95% arcsin square root transformation in the first week, based on (8), is  $(0.4890, 0.4902)$ .

Estimating the average length of contracting dengue fever based on data on the number of dengue cases in the city in January-April 2017 is obtained by calculating  $\hat{\mu}_\tau$  using equation (6), with  $\tau$  as the longest observation time, which is 49 days, then obtained:

$$\hat{\mu}_\tau = \int_0^{49} \hat{S}(x) dx = 8.1407$$

The next step is to calculate  $\hat{V}(\hat{\mu}_\tau)$  using equation (7), for  $\tau = 49$ , obtained  $\hat{V}(\hat{\mu}_{49}) = 0.0758$ . From the above calculation results obtained average duration of positive patients contracted dengue fever is after 8 days experiencing symptoms of dengue fever.

#### **4. Conclusions**

The Product-Limit estimator can be applied to estimate the survival function and cumulative hazard function in calculating the average duration of dengue fever in the city of Bandung from January to April 2016. In addition, the Nelson-Aalen Appraiser can also be applied to estimate the cumulative hazard function. Estimated survival function and cumulative hazard function is obtained by entering data on the number of dmam cases of bloody, so that the calculation results obtained the average duration of dengue fever in the city of Bandung in January to April 2016 is 8 days.

#### **Acknowledgements**

Further thanks to the Rector, Director of DRPMI, and Dean of FMIPA, Universitas Padjadjaran, which has a grant program of Competence Research of Lecturer of Unpad (*Riset Kompetensi Dosen Unpad/RKDU*) under the coordination of Dr. Sukono, which is a means to increase research activities and publications to researchers at the Universitas Padjadjaran.

#### **References**

- Cameron, A.C. and Trivedi, P.K. *Microeconometrics Methods and Applications*, Cambridge University Press, 2005.
- Christensen, B. J. and Kallestrup-Lamb, M. The Impact of Health Changes on Labor Supply: Evidence from Merged Data on Individual Objective Medical Diagnosis Codes and Early Retirement Behavior, *Health Econ. 21(Suppl. 1)*: 56–100, 2012. Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/hec.2811.
- Guibert, Q. and Planchet, F. Non-Parametric Inference of Transition Probabilities Based on Aalen-Johansen Integral Estimators for Semi-Competing Risks Data: Application to LTC Insurance, *Working Paper*, ISFA - Laboratories SAF, 2014.
- Hanif, A., Butt, A., Ahmed, A., Sajid, M. R., Ashraf, T. and Nawaz, A.A. Survival Analysis of Dengue Patients in Relation to Severity of Liver Dysfunction in Pakistan, *Advances in Biological Research* 9 (2): 91-94, 2015. ISSN 1992-0067, © IDOSI Publications. DOI: 10.5829/idosi.abr.2015.9.2.92193.
- Karnon, J., Mackay, M. and Mills, T.M., Mathematical Modelling in Health Care. *Working Paper*. 18th World IMACS/ MODSIM Congress, Cairns, Australia 13-17 July 2009. <http://mssanz.org.au/modsim09>, 2009
- Kenah, E. Nonparametric Survival Analysis of Epidemic Data, *Working Paper*, Department of Biostatistics, University of Washington, Seattle, USA, 2011.
- Klein, J.P. and Moeschberger, M.L. Survival Analysis. *Techniques for Censored and Truncated Data*. Newyork: Springer-Verlag, New York, Inc., 1997.
- Mesike, G.C., Adeleke, I.A. and Ibiwoye, A., Predictive Actuarial Modeling of Health Insurance Claim Costs. *International Journal of Mathematics and Computation*, ISSN 0974-570X (Online), ISSN 0974-5718 (Print), Vol. 14, Issue No. 1, Year 2012. Copyright © 2011-12 by CESER Publications. 2011
- Nasir, J.A. and Zaidi, S.A.A. Modelling Survival Data of Thalassaemia Patients in Pakistan, *J Ayub Med Coll Abbottabad*; 21 (1). 2009
- Rich, J.T., MD, Neely, J.G., MD, Paniello, R.C., MD, Voelker, C.C. J., MD, DPhil, Nussenbaum, B., MD, and Wang, E.W., MD, St. Louis, MO, A Practical Guide to Understanding Kaplan-Meier Curves, *Invited Article, Otolaryngology–Head and Neck Surgery*, 143, 331-336. 2010
- Sukono, Nahar, J., Putri, F. T., Subiyanto, Mamat, M., & Supian, S. Estimation of outstanding claims reserving based on inflation risk on car insurance companies by using the bootstrap method. *Far East Journal of Mathematical Sciences*, 102(4), 687-706. 2017.a. doi:10.17654/MS102040687
- Sukono, Suyudi, M., Islamiyati, F., & Supian, S. Estimation model of life insurance claims risk for cancer patients by using Bayesian method. Paper presented at the *IOP Conference Series: Materials Science and Engineering*, 166(1), 2017.b. doi:10.1088/1757-899X/166/1/012022
- Willems, S.J.W. Inverse Probability Censoring Weights for Routine Outcome Monitoring Data, *Master Thesis*, Department of Mathematics, University of Leiden The Netherlands, 2014.
- Yen, F.Y. and Kassim, S. A Non-parametric Survival Estimate After Elimination of a Cause of Failure, *Sains Malaysian* 42(5): 673–683. 2013

## **Biographies**

**Sukono** is a lecturer in the Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran. Currently serves as Head of Master's Program in Mathematics, the field of applied mathematics, with a field of concentration of financial mathematics and actuarial sciences.

**Alit Kartiwa** is a lecturer at the Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran. He currently serves as Head of Pure Mathematics Laboratory, a field of pure mathematics, with a field of concentration analysis.

**Dwi Susanti & Rani Q.F.** is a lecturer & Student at the Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, the field of applied mathematics, with a field of concentration of financial mathematics and actuarial sciences.

**Yuyun Hidayat** is a lecturer at the Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran. He currently serves as Deputy of the Head of Quality Assurance Unit, a field of Statistics, with a field of Quality Control Statistics.

**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.