

Weibull Regression Survival Analysis on the Rate of Recovery of Typhoid Fever Patients: Case Study of RSUD Haji Makassar

Nurul Maghfira⁽¹⁾, Bobby Poerwanto⁽²⁾, Hardianti Hafid⁽³⁾

^{1,2,3} Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Negeri Makassar

Kampus UNM Parangtambung, Jalan Daeng Tata Makassar

e-mail: nurulmaghfira90@gmail.com⁽¹⁾, bobby_poerwanto@unm.ac.id⁽²⁾,
hardiantihf@unm.ac.id⁽³⁾

ABSTRAK

Analisis *survival* adalah metode statistika yang digunakan dalam mempelajari ketahanan hidup yang berhubungan dengan waktu, mulai waktu awal penelitian sampai waktu akhir penelitian. Tujuan analisis *survival* adalah untuk mengetahui hubungan antara waktu *survival* dengan variabel bebas pada penelitian yang diduga mempengaruhi waktu *survival*-nya. Pada penelitian ini menggunakan analisis *survival* regresi weibull untuk melihat faktor-faktor yang berpengaruh secara signifikan terhadap laju kesembuhan pasien penderita demam tifoid di RSUD Haji Makassar Tahun 2022. weibull untuk melihat faktor-faktor yang berpengaruh secara signifikan terhadap laju kesembuhan pasien penderita demam tifoid di RSUD Haji Makassar Tahun 2022. Diperoleh hasil analisis faktor yang berpengaruh signifikan terhadap laju kesembuhan pasien penderita demam tifoid adalah nyeri ulu hati. Pasien dengan riwayat mengalami nyeri ulu hati memiliki nilai *hazard ratio* sebesar 1,779, yang berarti bahwa pasien penderita demam tifoid yang mengalami nyeri ulu hati memiliki laju kesembuhan lebih cepat sebesar 1,779 kali dibandingkan dengan pasien yang tidak mengalami nyeri ulu hati.

Kata kunci: Anslisis Survival, Regresi Weibull, Demam Tifoid

ABSTRACT

Survival analysis is a statistical method used in studying survival related to time, from the beginning of the study to the end of the study. The purpose of survival analysis is to determine the relationship between survival time and independent variables in research that are thought to affect survival time. This study uses weibull regression survival analysis to see the factors that significantly affect the recovery rate of patients with typhoid fever at the Makassar Hajj Hospital in 2022. The results of the analysis of factors that have a significant effect on the rate of recovery of patients with typhoid fever are heartburn. Patients with a history of having heartburn have a hazard ratio value of 1.779, which means that patients with typhoid fever who experience heartburn have a faster recovery rate of 1.779 times compared to patients who do not experience heartburn.

Keywords: Survival Analysis, Weibull Regerssion, Thyphoid Fever

INTRODUCTION

Survival analysis is one of the statistical methods that is often used to analyze data related to the timing of events [1]. Survival time is defined as the survival time of an object at the beginning of observation until the occurrence of an event [2] Events in survival analysis can be death, disease relapse, treatment or others [3].

The purpose of survival analysis is to determine the relationship between the time of the event and the independent variables measured at the time of the study. In addition, it is also used to see the factors that significantly affect an event or event [4].

One of the modeling of survival analysis is a parametric survival model, which is a model that has the assumption of following a certain known distribution on its dependent variable. In some circumstances, parametric models can provide more accurate estimates ([5]. Distributions that are often used in survival analysis include the Weibull, exponential, log normal, and log-logistic distributions.

The Weibull distribution is a generalization of the exponential distribution, it is known that in the exponential distribution the hazard value is constant, so it often does not match the actual situation. However, in the Weibull distribution, the hazard value is not constant so that it is closer to the actual situation. [6]. The Weibull distribution is flexible due to the shape parameter that determines the change in shape of the hazard curve [7].

Typhoid fever is an acute infectious disease of the digestive system caused by salmonella typhi or salmonella paratyphoid bacteria. Typhoid fever is transmitted through food or drink contaminated with salmonella typhi bacteria, in addition the disease can be transmitted through direct contact with feces, urine or secretions of people with typhoid fever. [8].

Based on data from the World Health Organization (WHO) an estimated 11-20 million people fall ill with typhoid and about 128,000 to 161,000 people die from it each year. The disease is mostly found in areas that lack clean water and environmental sanitation. It can also be caused by unhealthy food [9].

Typhoid fever is one of the endemic diseases in Indonesia that must be given serious attention because it can be a public health threat. In Indonesia alone, typhoid fever cases range from 350-810 per 100,000 population, the prevalence of this disease in Indonesia is 1.6% and ranks 5th in infectious diseases that occur at all ages in Indonesia, which is 6.0% and ranks 15th in the cause of death of all ages in Indonesia, which is 1.6%. [10], So it is necessary to reduce the mortality rate and one of the things that can be done is to analyze the factors that affect the recovery rate of patients with typhoid fever.

METHOD

1. Data Source

The data used in this study are secondary data and identification of variables set as criteria for typhoid fever patients obtained from the medical records of the RSUD Haji Makassar in 2022.

2. Variable Operational Definition

Tabel 1. Variable Operational Definition

Variable	Name Variable	Descriptive
Y	Survival Time	Time during which typhoid fever patients were hospitalized (days).
S	Status	0 = If the patient was involuntarily discharged, or changed treatment, died during the study period; 1 = If the patient recovers
X ₁	Gender	0 = Male; 1 = Female
X ₂	Age	-
X ₃	Heartburn	0 = Not heartburn; 1 = Heartburn
X ₄	Dirty Tongue	0 = Not dirty tongue; 1 = Dirty tongue
X ₅	Vomit	0 = Not Vomit; 1 = Vomit
X ₆	Diarrhea	0 = Not diarrhea; 1 = Diarrhea
X ₇	Dizzy	0 = Not Dizzy; 1 = Dizzy

3. Research Produce

The procedures carried out in this study are:

1. Collecting sources of information needed in the research.
2. Recapitulation of typhoid fever patient data at this stage the data is taken from the Makassar Hajj Hospital to be processed using the Weibull regression method.
3. Performing data processing using statistical software.
4. Compile a research report.
5. Make conclusions based on the problems that have been discussed

4. Data Analysis Technique

The data analysis techniques in this study are:

1. Describe the characteristics of data in patients with typhoid fever.
2. Testing the assumption of proportional hazard

There are three types of checking the Proportional Hazard assumption [11]:

1. The survival lines on the Kaplan-Meier curve do not intersect each other.
2. Survival lines on the ln-ln survival do not intersect each other.
3. Global test or Goodness of Fit (GOF) test
4. Weibull distribution test on dependent variable [12]

$$D = \sup |S(t) - F_0(t)|$$

where:

D = kolmogorov Smirnov

$S(t)$: empirical value of sample cumulative distribution

$F_0(t)$: cumulative distribution function

\sup : Supremum or upper limit

5. Weibull regression modeling [13]:

$$h(t|X) = \lambda \gamma t^{-1} \exp(\beta^t X)$$

where:

$h(t|X)$ = Cox model

β^t = regression coefficient

λ = scale parameter value

γ = shape parameter value

t = time variable

6. Parameter estimation with the method Maximum Likelihood Estimation [13], [14]:

$$L(\theta|t) = \frac{n!}{(n-r)!} S(t|X)^{n-r} \prod_{i=1}^r f(t|X)$$

Where:

$L(\theta|t)$ = likelihood θ function

X_i = a vector of variables of failed individuals at time t_i

r = the desired number of deaths or failures

n = the number of observed data.

7. Parameter significance testing [15], [16]:

- a. Simultaneous Test

$$G^2 = -2[\ln L_R - \ln L_f]$$

L_R = partial likelihood starting model.

L_f = partial likelihood final model.

- b. Partial Test

$$Z = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)}$$

$\hat{\beta}_i$ = the parameters of the regression model to be estimated,

$SE(\hat{\beta}_i)$ = standard error the parameters of the regression model to be estimated,

- c. Calculate the hazard ratio value [17]:

The hazard ratio for categorical data is as follows:

$$\widehat{HR} = \frac{h(t|x = 1)}{h(t|x = 0)} = \frac{h(t)e^{\hat{\beta}}}{h(t)} = e^{\hat{\beta}}$$

\widehat{HR} = hazard ratio estimation

$e^{\hat{\beta}}$ = exponential coefficient

- d. Interpretation of the Weibull regression results obtained
e. Conclusion.

RESULT AND DISCUSSION

1. Descriptive Analysis

Data are presented in the form of histograms, tables and diagrams with a total of 75 patients with typhoid fever at the Makassar Hajj Hospital in 2022. The longest typhoid fever patient was 8 days as many as 1 patient, while the fastest typhoid fever patient was 1 day as many as 4 patients and the most patients were 3 days, namely 19 patients. The percentage of typhoid fever patients

who recovered was 86.67% (65 people). The oldest age of the patient was 69 years, while the youngest age of the patient was 1 year. the average age of the patient was 25 years. Inpatients with typhoid fever were more female than male. More typhoid fever inpatients experienced heartburn, vomiting, dizziness and compared to those who did not experience heartburn, vomiting, and dizziness, and more typhoid fever patients did not experience dirty tongue and diarrhea than those who experienced dirty tongue and diarrhea.

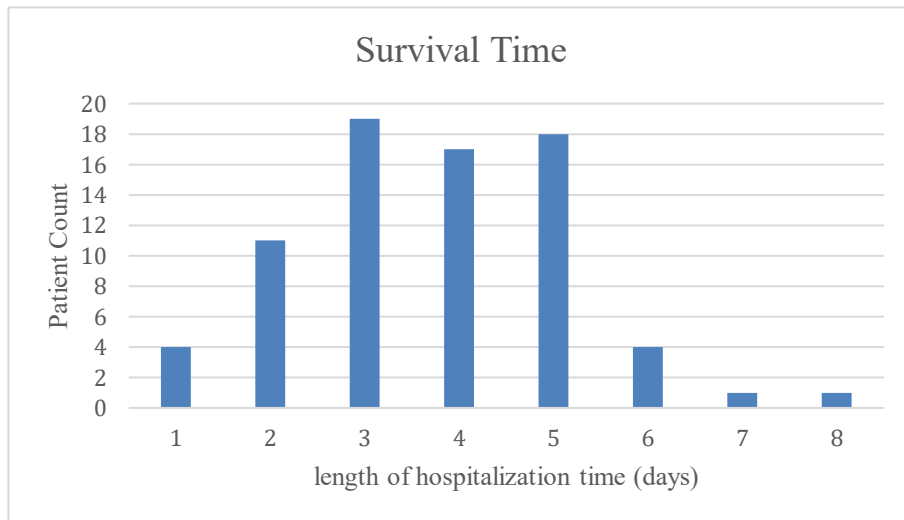
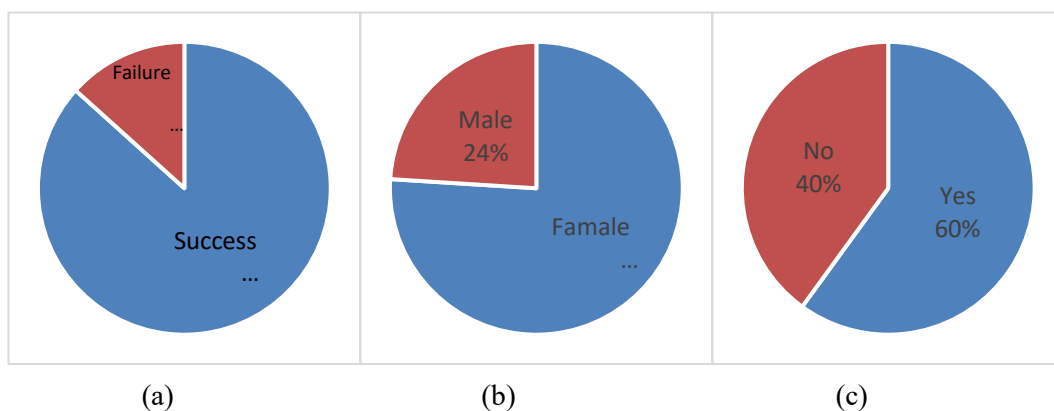


Figure 1. Survival Time Histogram

Table 2. Age of Typhoid Fever Patient

Variable	Average (Year)	Maximum (Year)	Minimum (Year)
Age	25	69	1



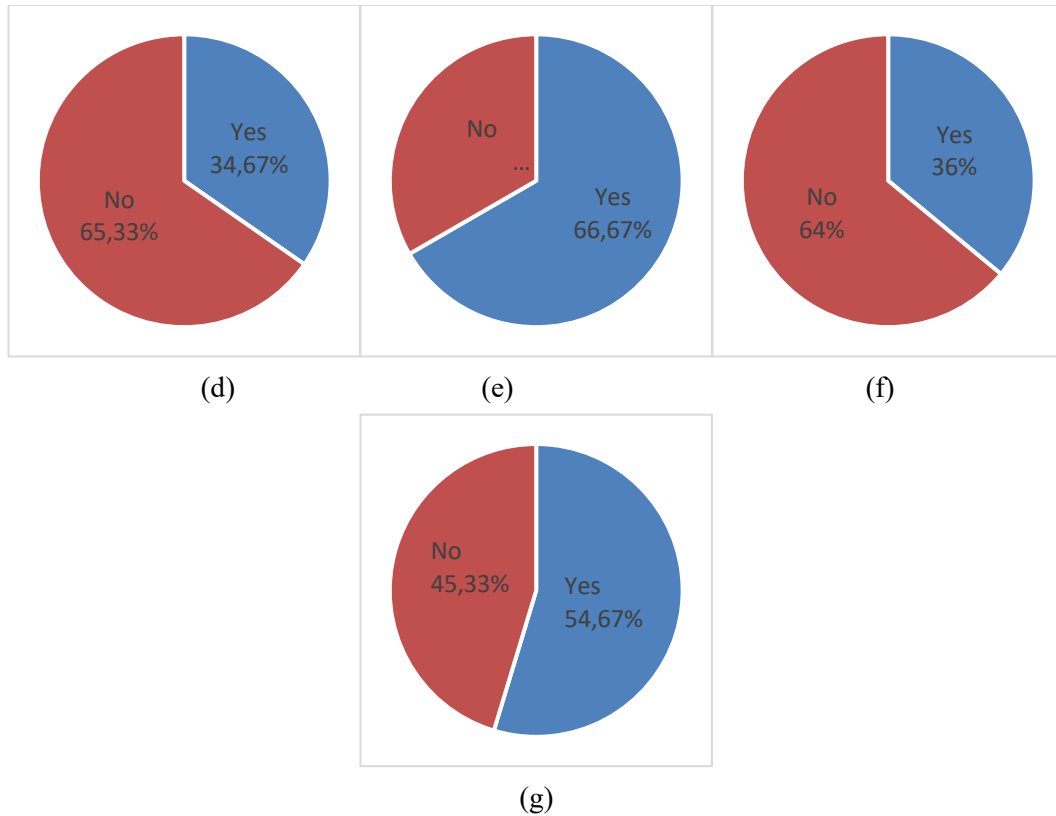


Figure 2. Pie Chart for (a) Event Censorship Status, (b) Gender, (c)Heartburn, (d)Dirty Tongue, (e)Vomit, (f)Diarrhea, (g)Dizzy.

2. Testing the Proportional Hazard Assumption

In testing the proportional hazard assumption, this study uses a global test or goodness of fit test. With hypothesis:

H_0 = the data meets the proportional hazard assumption

H_1 = the data does not fulfill the proportional hazard assumption

Reject H_0 if $p - value < \alpha(0,05)$

Table 3. Global Test p-value

Variable	p-value	Description
Gender (X_1)	0,52	Accept H_0
Age (X_2)	0,46	Accept H_0
Heartburn (X_3)	0,97	Accept H_0
Dirty Tongue (X_4)	0,89	Accept H_0
Vomit (X_5)	0,65	Accept H_0
Diarrhea (X_6)	0,70	Accept H_0
Dizzy (X_7)	0,96	Accept H_0

Based on table 3, it shows that the results of the global test for each variable each have a p-value $> \alpha = 0.05$, so it can be concluded that it Accept H_0 , which means that the proportional hazard assumption on typhoid fever patient data is met.

3. Testing Data Distribution

Data distribution testing is used to determine the distribution used is in accordance with the length of hospitalization time (Y) data. With Hypothesis:

H_0 = survival data follows the Weibull distribution

H_1 = survival data does not follow the Weibull distribution

Reject H_0 if $D_{hit} > D_{(\alpha,n-1)}$ or $p - value < \alpha(0,05)$

Table 4. Distribution Test Values

Distribution	D_{hit}	$p-value$	Description
Weibull	0,132	0,146	Accept H_0

Based on table 4, it shows that the data on the length of hospitalization of patients with typhoid fever obtained a p-value from the Kolmogorov Smirnov test of 0.146 where the p-value = 0.146 $> \alpha = 0$. and the value of $D_{hit} = 0,132 < D_{(0,05,74)} = 0,157$ so that it Accept H_0 , which means that the data on the length of hospitalization of patients with typhoid fever follows the Weibull distribution.

4. Parameter Estimation and Parameter Testing

Parameter estimates in the Weibull regression model can be estimated using the Maximum Likelihood Estimation (MLE) method. In parameter testing including simultaneous testing and partial testing, the results of parameter estimation and parameter testing are obtained as follows:

Table 5. Parameter Estimation and Partial Test

Variable	Z	Coefficient	P-value	Description
Gender (X_1)	1,860	0,625	0,063	Not significant
Age (X_2)	0,444	-0,004	0,645	Not significant
Heartburn (X_3)	2	0,576	0,045	Significant
Dirty Tongue (X_4)	1,063	-0,300	0,287	Not significant
Vomit (X_5)	0,056	-0,016	0,955	Not significant
Diarrhea (X_6)	0,418	-0,134	0,675	Not significant
Dizzy (X_7)	1,436	-0,405	0,151	Not significant

a. Simultaneous test

The simultaneous test is carried out to see the effect of the independent variables together on the dependent variable in a model. With hypothesis:

H_0 = The independent variable has no effect on the dependent variable

H_1 = At least one of the independent variables affects the variable.
 Reject H_0 if $p\text{-value} \leq \alpha(0,05)$ or $G_{hit} \geq x^2_{\alpha;db=p}$

Table 6. Simultaneous test for all variables

<i>Likelihood ratio test (G)</i>	$x^2_{(0,05;6)}$	<i>p-value</i>
16,4	12,59	0,021

Based on the results of the simultaneous test, the likelihood ratio test value is 16.4 and the p-value = 0.021 < $\alpha = 0.05$, because the value of $G_{hit} = 16,4 > x^2_{(0,05;6)} = 12,59$, it can be concluded that rejecting H_0 means that there are one or more variables that have a significant effect or it can be interpreted that the model as a whole can contribute to the recovery rate of patients with typhoid fever.

b. Partial Test

This test is a test conducted to determine the effect of each independent variable on the dependent variable. With hypothesis:

H_0 = Independent variable i has no effect on survival time

H_1 = Independent variable i affects survival time

Reject H_0 if $p\text{-value} \leq \alpha(0,05)$ or $|Z| \geq \frac{Z_{0,05}}{2}$

Based on the results in table 5, the test results for each variable are shown. Based on the H_0 rejection criteria, it can be concluded that only the heartburn variable has a p-value < $\alpha = 0.05$ and a value of $|Z| = 2 > \frac{Z_{0,05}}{2} = 1,96$ it can be concluded that reject H_0 , which means that only the heartburn variable has a significant effect on the recovery rate of patients with typhoid fever.

From the parameter estimation results in the table above, a Weibull regression model is obtained which has a significance value of less than $\alpha = 0.05$ so that a Weibull regression model is obtained as follows:

$$h(t|X) = \lambda \gamma t^{\gamma-1} \exp(\beta^t X)$$

$$h(t|x) = (4,190)(2,836)t^{2,836-1} \exp(0,576X_3)$$

$$h(t|x) = (4,190)(2,836)t^{1,836} \exp(0,576X_3)$$

5. Hazard Ratio

To find out the survival rate of patients with typhoid fever can be seen by looking for the hazard ratio value of significant variables. The results of the analysis for the hazard ratio value can be seen in the following table:

Table 7. Hazard Ratio values for significant variables

Variable	Hazard Ratio
Heartburn (X_3)	1,779

Based on table 7, the hazard ratio value for the heartburn variable is 1.779. This value is more than one, which means that patients with typhoid fever who experience heartburn have a faster recovery rate of 1.779 times compared to patients who do not experience heartburn.

CONCLUSION

Of the several factors that are considered to have a significant effect on the recovery rate of inpatients with typhoid fever at the Makassar Hajj Hospital, there is only one variable that has a significant effect, namely the variable of heartburn. The Weibull regression model for data on typhoid fever patients at the RSUD Haji Makassar in 2022 is:

$$h(t|x) = (4,190)(2,836)t^{1,836}exp(0,576X_3)$$

The hazard ratio value of 1.779 was obtained, which means that patients with typhoid fever who experience heartburn have a faster recovery rate of about 1.779 times compared to patients who do not experience heartburn.

For future researchers, the authors suggest using other methods in determining the factors that influence the recovery rate of typhoid fever patients. before designing a study, it is better to find out in advance whether or not the data to be used in the study so that it will be easier when doing a study and can conduct research with different case studies.

REFERENCE

- [1] Collet David, *Modelling Survival Data In Medical Research third Edition*. 2013.
- [2] R. Y. Fa'rifah and Purhadi, "Analisis Survival Faktor-Faktor yang Mempengaruhi Laju Kesembuhan Pasien Penderita Demam Berdarah Dengue (DBD) di RSUD Haji Surabaya dengan Regresi Cox," *Jurnal Sains dan Seni ITS*, vol. 1, no. 1, pp. 271–276, 2012.
- [3] M. K. Yusuf, "Analisis Survival Lama Waktu Sembuh Dengan Perawatan Standar Pada Pasien Rawat Inap Penyakit Tifus Dengan Menggunakan Metode Regresi Cox Proportional Hazard," 2018, [Online]. Available: <http://search.ebscohost.com/login.aspx?direct=true&db=sph&AN=119374333&site=ehost-live&scope=site%0Ahttps://doi.org/10.1016/j.neuron.2018.07.032%0Ahttp://dx.doi.org/10.1016/j.tics.2017.03.010%0Ahttps://doi.org/10.1016/j.neuron.2018.08.006>
- [4] A. R. Faisal, M. N. Bustan, and S. Annas, "ANALISIS SURVIVAL DENGAN PEMODELAN REGRESI COX PROPORTIONAL HAZARD MENGGUNAKAN PENDEKATAN BAYESIAN (Studi Kasus: Pasien Rawat Inap Penderita Demam Tifoid di RSUD Haji Makassar)," *VARIANSI: Journal of Statistics and Its application on Teaching and Research*, vol. 2, no. 2, p. 62, 2020, doi: 10.35580/variensium14629.
- [5] H. P. Zhu, X. Xia, C. H. Yu, A. Adnan, S. F. Liu, and Y. K. Du, "Application of Weibull model for survival of patients with gastric cancer," *BMC Gastroenterol*, vol. 11, pp. 1–6, 2011, doi: 10.1186/1471-230X-11-1.
- [6] L. E. Widyawati and R. D. Bkti, "Analisis Faktor yang Mempengaruhi Laju Kesembuhan Pasien Demam Berdarah Dengue Menggunakan Regresi Cox dan Regresi Weibull (Studi

- Kasus : Kabupaten Bantul, Yogyakarta),” *Jurnal Statistika Industri dan Komputasi*, vol. 05, no. 2, p. 26, 2020.
- [7] A. Solehah, “Analisis Ketahanan Hidup Pasien Kanker Paru Menggunakan Regresi Weibull,” 2018.
- [8] Y. Levani and A. D. Prastya, “Demam Tifoid: Manifestasi Klinis, Pilihan Terapi Dan Pandangan Dalam Islam,” *Al-Iqra Medical Journal : Jurnal Berkala Ilmiah Kedokteran*, vol. 1, no. 2, pp. 10–16, 2020, doi: 10.26618/aimj.v3i1.4038.
- [9] Juniah and M. Arianti, “Pentingnya Pengetahuan Kesehatan Demam Typhoid Anak,” *Jurnal Pengabdian Masyarakat*, vol. 2, no. 1, pp. 7–14, 2023, doi: 10.59030/jpmbd.v2i1.18.
- [10] S. Khairunnisa, E. M. Hidayat, and R. Herardi, “Hubungan Jumlah Leukosit dan Persentase Limfosit terhadap Tingkat Demam pada Pasien Anak dengan Demam Tifoid di RSUD Budhi Asih Tahun 2018 – Oktober 2019,” *Seminar Nasional Riset Kedokteran (SENSORIK)*, pp. 60–69, 2020.
- [11] M. Sopiudin Dahlan, *Analisis Survival Dasar-Dasar Teori Dan Aplikasi Program Stata*. Jakarta: Sagung Seto, 2013.
- [12] N. H. Hasa, M. N. Bustan, and Aswi, “ANALISIS BAYESIAN SURVIVAL WEIBULL UNTUK MENENTUKAN FAKTOR YANG MEMPENGARUHI LAJU KESEMBUHAN PASIEN RAWAT INAP KANKER SERVIKS DI RSUD KOTA MAKASSAR,” *VARIANSI: Journal of Statistics and Its Application on Teaching and Research ISSN 2684-7590 (Online) Vol. 4 No. 1 (2022), 1-8*, vol. 4, no. 1, pp. 1–8, 2022.
- [13] I. Sari, “Analisis Regresi Weibull (Kasus: Faktor-faktor yang mempengaruhi laju kesembuhan penderita penyakit demam berdarah dengue di RSAL JALA AMMARI Makassar),” 2017.
- [14] Yogi Adam Firdaus, N. Ngatini, and Sekarsari Utami Wijaya, “Pemodelan Regresi Data Panel Harga Beras di Wilayah Indonesia Bagian Barat,” *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, vol. 16, no. 2, pp. 486–498, Dec. 2023, doi: 10.36456/jstat.vol16.no2.a8061.
- [15] C. H. Imanina, “ANALISIS SURVIVAL TERHADAP PASIEN PENYAKIT GINJAL KRONIS DENGAN MENGGUNAKAN COX REGRESSION (Studi Kasus: Pasien Penyakit Ginjal Kronis di RSUD Arifin Achmad, Pekanbaru, Riau),” no. 21, pp. 1–9, 2018.
- [16] G. Anuraga, A. Indrasetianingsih, and M. Athoillah, “Pelatihan pengujian hipotesis statistika dasar dengan software r,” *BUDIMAS: Jurnal Pengabdian Masyarakat*, vol. 3, no. 2, pp. 327–334, 2021.
- [17] Riangkaryaman, “MODEL SEMIPARAMETRIK COX PH DAN PARAMETRIK (WEIBULL DAN LOGNORMAL) DALAM ANALISIS SURVIVAL (Kasus: Rawat Inap Penderita DBD di RSUD Haji Kota Makassar),” 2017.