

ShinySurv : A survival analysis and Shiny web application

Fatih Sağlam *

*Department of Statistics
Faculty of Art and Sciences
Ondokuz Mayıs University
Samsun 55270
Turkey*

Tuba Şanlı [†]

*Department of Finance and Banking
Görele School of Applied Sciences
Giresun University
Giresun 28800
Turkey*

Mehmet Ali Cengiz [§]

*Department of Statistics
Faculty of Art and Sciences
Ondokuz Mayıs University
Samsun 55270
Turkey*

Abstract

There are different programs available for survival analysis. However, these programs do not deal with survival problems with all censor types. The aim of this study is to form and explain “ShinySurv” application of survival analysis methods in R-Shiny for different types of censorship. Using ShinySurv, the user can upload his own data set to the system

* ORCID ID: 0000-0002-2084-2008

† ORCID ID: 0000-0001-6019-8037

§ ORCID ID: 0000-0002-1271-2588

* E-mail: fatih.saglam@omu.edu.tr (Corresponding Author)

† E-mail: tuba.sanli@giresun.edu.tr

§ E-mail: macengiz@omu.edu.tr

or benefit from sample data sets. ShinySurv provides visualization of analysis results with graphs and reporting of results for parametric and non-parametric survival methods. An example application of a survival analysis is given. The web application and the codes can be accessed at <https://scypster.shinyapps.io/ShinySurv/> and <https://github.com/fatihsglam/ShinySurv>.

Subject Classification: 62N02, 62N03, 62G07.

Keywords: Shiny, Web application, Survival analysis, Censored data.

1. Introduction

R program is a programming language in which statistical methods are applied and graphics are created. The R is constantly being developed, reported bugs, and can be modified. Shiny, a package of R, is a program that users can interact with over the web, making analysis and graphics fast and simple. Thanks to Shiny, users can access via the web even with tablets and smartphones without the R program. This makes the program more flexible and comfortable to use [28]. For this reason, Shiny applications for different statistical methods have been developed by researchers. Thus, analyzes made especially in the field of biostatistics have become easier for users. Wang and Zhou [35] developed parametric and semi-parametric models in right-censored data to deal with large data sets in biostatistics and bioinformatic. They used the SurvELM package and Shiny in R for this. Zhou et al. [41] developed the Shiny application named MEPHAS, which includes statistical analyzes for medicine and pharmacists in their study. Soutinho and Machado [29] developed a web application called MSM.App in Shiny in 2022. In this application, the Kaplan-Meier method, Cox regression and accelerated time to failure models, which are among the survival analysis methods, were explained. Yaşar et al. [38] applied statistical methods, hypothesis testing, correlation and regression analysis in the Shiny application they made using the R software. Arslan et al. [4, 5] used the Shiny to calculate sample size and power analysis for hypothesis testing, diagnostic tests, correlation and regression analysis. In another study they conducted in 2018, they applied the Kruskal Wallis H Test using the Shiny package. Holland and Schneider [17] developed the R Shiny application in dose-response experiments (beam designs) applied in optimal experimental design theory, which can be applied to the selection of optimal dose levels for each available observation and to both single dose-response relationships and interaction experiments in which two substances are given at the same time in several

different mixing ratios. Chopra et al. [9] taking full advantage of cloud computing technology requires redesigning applications to take full advantage of them. They proposed many strategies to improve the quality of web applications deployed in the cloud. Li et al. [23] They proposed a method to generate and verify authentication codes using cryptographic techniques that link a generated code to a specific user. Chu et al. [10] developed the “SMART” Shiny app that helps researchers perform demographic analysis of the original HIS datasets, Kaplan-Meier survival analysis and Cox proportional hazard ratio analysis. Dwivedi et al. [13] in their shiny application named “Survival Genie”, analyzed 53 datasets of 27 different malignancies from 11 different cancer programs related to adult and pediatric cancers with the Cox proportional hazard model. Wei et al. [36] used dual criteria and confidence interval-based approaches in drug development to enable quantitative decision making, and they made the application in shiny “GOahead”. Wissel et al. [37] developed “COVID-19 Watcher”, an interactive online dashboard Shiny app to monitor COVID-19 in US cities and states in real time. Yaşar et al. [39] developed the “DTROC” software, a web-based and user-friendly Shiny application, with which clinicians and researchers can perform two- and three-dimensional receiver operating characteristic (ROC) curve analyzes and calculate diagnostic test performance metrics. Özhan et al. [27] developed a Shiny application to perform multivariate survival analysis of TCGA data with “SmulTCan” gene sets. When the studies were examined, the applications of survival analysis methods for different variables were made in Shiny. We see application of survival analysis methods for different variables made in Shiny. However, the different types of censorship in survival analysis have not been examined in detail.

It is possible for an individual or object to leave the study for any reason. Often times, the failure process is completely unobservable. In this case, the only information obtained would be when the individual stopped functioning. If the individual is known to be alive at time t , time is called the censored survival time. Right censoring occurs only if the time of investigated event is known to be longer than the censoring time at the end of the study, loss of follow-up, or patient withdrawal. Left censoring exists when the individual's actual survival is less than the observed survival time. In intermittent censorship, it is known that the time interval in which the investigated event occurs is somewhere between a and b [12]. If there is both right and left censoring in the data set, this is called double censoring.

Packages such as `survival`, `KMsurv`, `npsurv`, `NADA` and `dblcens` available in R can perform survival analysis by writing code. Apart from R, it is also possible to perform survival analysis in package programs such as SAS, Stata, SPSS and Minitab.

It is difficult to encounter datasets with only left censorship. Left-censored datasets often contain both left- and right-censored samples [20]. Therefore, these datasets are considered doubly censored [33]. As far as we know, the options for users to perform a double censored survival analysis are only available in the `dblcens` package. However, the `dblcens` package does not provide detailed results and requires coding. We took this into account when writing the ShinySurv app. ShinySurv can handle right, left, double and interval censorship types both parametrically and non-parametrically.

This article introduces the ShinySurv application, its properties and how it works. In Chapter 2, the structure of ShinySurv and data entry are explained. Chapter 3 shows all phases of the program on a sample application. In Chapter 4, ShinySurv is compared with other Shiny survival analysis applications, R packages, and package programs. Chapter 5 discusses the results of the article and future work.

2. Framework and Data Entry

ShinySurv is an application that includes parametric and non-parametric methods in survival analysis. In Figure 1, the structure of the program is explained in general. Example data sets are available in the interface so that users can easily understand and practice different types of censorship. At the same time, the users can enter their own data.

In Figure 2, flowchart of data entry in data menu is shown. Example data sets for different censorship types are available in the “Load data” section of the data menu. There are right censored, double (right and left, that is, double) censored, interval censored data set examples.



Figure 1
ShinySurv framework

Example 1 (“right censored”) is a right censored data example. In the study conducted on 42 children with leukemia, the patients were randomly divided into 2 groups. The first group was treated with 6-MP medication, while the other group was determined as placebo. The study was followed until the leukemia relapsed or the end of the study period [20].

Example 2 (“left and right censored”) is an Example of Double Censored Data. The question “When did you first use cannabis” was asked to 191 male students studying at high school in the United States. Accordingly, those who said “I have never used it” among the answers were censored from the right. Those who answered “I used it but I don’t remember when I used it for the first time” were considered censored from the left. This data set was taken as double censored [20].

Example 3 (“interval censored”) is an Interval Censored Data Example. Women treated for breast cancer were divided into two groups. While only radiation therapy was applied to 46 women, 48 women received both radiation and chemotherapy treatment. Patients were observed in the range of 4-6 months and cancer cells were followed. Since observation was made in this time interval, the data was taken as an example for interval censorship [20].

Example 4 (“lung”) is the Lung Cancer Data Example. They analyzed how well lung cancer patients were able to perform their usual daily activities by measuring their performance scores with survival analysis methods. The data set is “NCCTG Lung Cancer Data”. We studied 167 patients free of missing data [24].

Example 5 (“ovarian”) is the Ovarian Cancer Data Example. 26 people were randomly selected from advanced ovarian cancer patients. The patients were treated every 3 weeks and the course of the tumor was

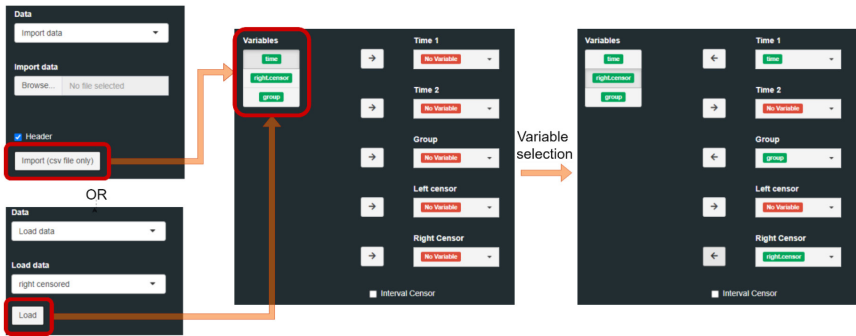


Figure 2

Data Entry

followed. The age of ovarian cancer patients and the effect of the treatment method on cancerous cells were compared [14].

Users can upload files in “csv” format in the “import data” section so that they can make applications on their own datasets. Rows that include missing observations in the loaded data set are automatically deleted.

Users should pay attention to the following points when entering data according to censorship types.

- “Time 1” is the observation start time and “Time 2” is the observation end time.
- For right, left censored or uncensored data, when only one of “Time 1” and “Time 2” is entered, the application will accept the start of tracking as 0 and will take the entered variable as the end time of tracking.
- Interval censorship requires both “Time 1” and “Time 2” variable. For left-censored cases, “Time 1” should be 0. For right-censored cases, the follow-up end time in “Time 2” should be left blank.
- “Group” is the group variable. All analyzes will be performed taking into account the group variable entered.
- “Right censor” and “Left censor” sections accept the censor variable. It should take the values 0, “censored” and 1, “uncensored” for both right and left censored. If no variables are entered, it means that there is no censorship.
- Censor information for interval censorship must be handled in “Time 1” and “Time 2”. So “Right censor” and “Left censor” is useless for interval censoring.

A summary preview of the selected dataset is given below the button before it is imported or loaded. Finally, the user who imports or loads the data can obtain the application results by choosing one of the parametric and non-parametric survival analysis methods.

3. Application and Results

The image of the results obtained when the “right censored” sample data set in the Load data menu of the ShinySurv is as in Figure 3. Summary information about the data set to be used is accessed from the “Data Summary” tab. In the “Dataset”, there is the table of the variables selected

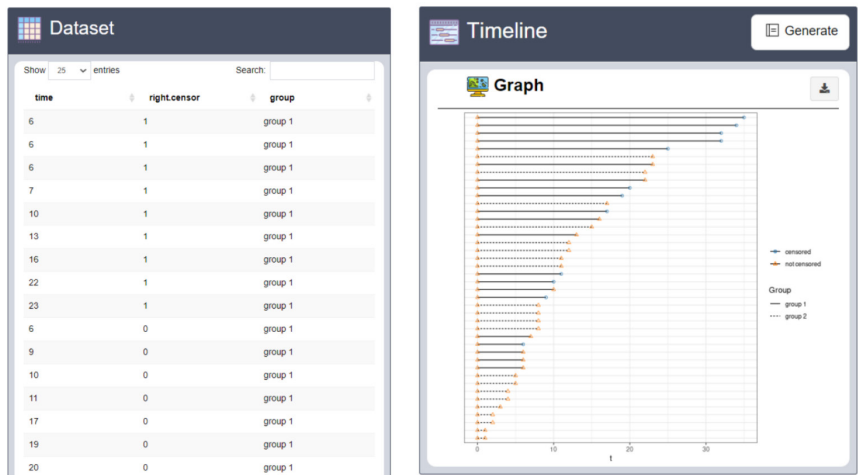


Figure 3
“Data Summary” tab

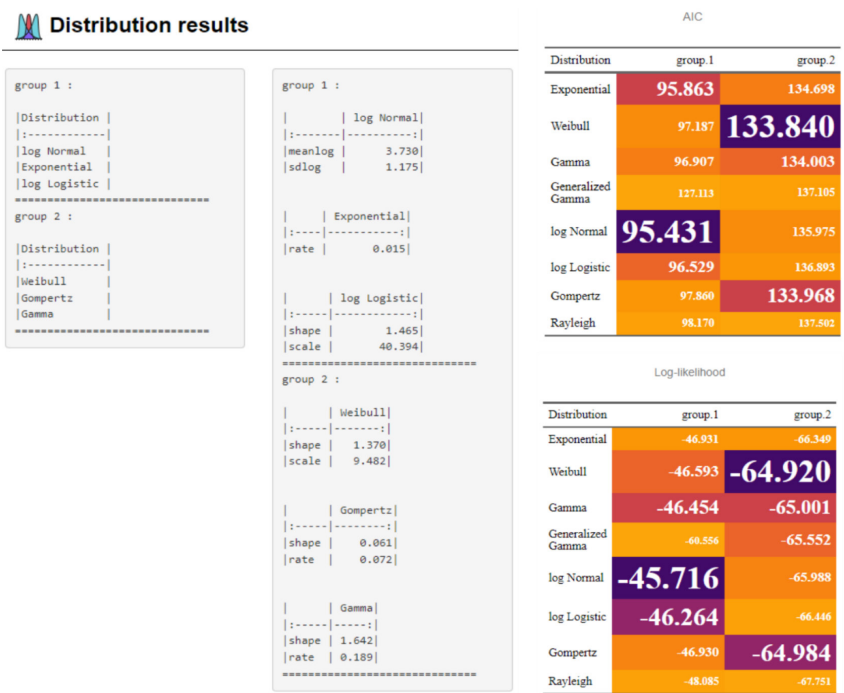


Figure 4
“Distribution Selection” tab

for analysis. The time interval in which individuals are observed from the beginning to the end is visualized in the graphic on the “Timeline” box. The censored cases in this time interval are separated according to the endpoint shapes, and the groups are according to the curve types. Using the download button, the graphic can be downloaded in different measurements and formats.

There can be many reasons that lead to the failure or death of a person at a given time. It is very difficult, if not impossible, to isolate the causes and explain them all mathematically. It is therefore difficult to choose a theoretical distribution to approximate survival data. ShinySurv offers 8 different theoretical distributions that are commonly used to define survival time: Exponential, Weibull, Gamma, Generalized Gamma, log-Normal, logLogistic, Gompertz, and Rayleigh.

In the “Distribution Selection” tab, the fit of the existing distributions to the survival data is compared. Comparison results are shown in Figure 4. AIC and log-likelihood values of distributions are tabulated for users. According to the AIC value, the 3 most suitable distributions and their corresponding parameters are printed on the screen.

After selecting the appropriate censorship type, the analysis is conducted in the “Parametric”, “Non-Parametric” tabs.

3.1 *Parametric Methods*

Survival time data include failure, death, response, relapse, development of a particular disease, etc. It refers to the time elapsed until an investigated event. This defined time is a random variable and has a distribution. The parameters of the distribution are calculated using the maximum likelihood method.

Censorships need to be handled carefully when constructing a likelihood function. For each observation, there is a different structure of information according to the censorship. In the likelihood function, the information that an uncensored event will add is the density of the event at that moment. For cases with right censorship, it is the probability that the event is greater than that time. For left censorship, it is the probability of the event being less than that time [20]. For interval censored data, the probability of the event occurring within the specified range, $[a, b]$, or the density of the midpoint of the range, i.e. $t = (a + b)/2$ should be taken into account. ShinySurv takes into account the density of the midpoint of this range. This point is coded as left censored if the starting point of the range is 0, and right censored if the end point is empty. It is then subject to the

procedure mentioned in the right and left censorship. In this case, the general formula of the likelihood function is

$$L \propto \prod_{i \in D} f(x_i) \prod_{i \in R} S(C_i) \prod_{i \in L} F(C_i).$$

Here, L is the likelihood function, C_r and C_l are the observation times of the right and left censored samples. D , R , and L are sets of uncensored, right-censored, and left-censored events, respectively. $S(t)$ is the survival function, which gives the probability of $P(T > t)$, and $F(t)$ is the failure function, which gives the probability of $P(T < t)$.

Parametric survival analysis is performed under the “Parametric” tab. Figure 5 shows the structure of this tab. In parametric survival analysis, survival, failure, cumulative hazard, hazard and density functions are obtained and presented graphically. There are 8 different distribution options to obtain these functions. Below the graph, information about the mean and median survival times is given. Below it, the distribution parameters are given. Finally, a life table containing 30 equally spaced time grid in the maximum and minimum survival time. In order to obtain the confidence intervals of all these results, the “Confidence Interval” should be marked in the “Curve Settings” box. Bootstrap methods are used to obtain confidence intervals. Available bootstrap methods in “bootstrap type” are Non-parametric bootstrap and Jackknife methods. In the “Graph Settings” box, the graphic design can be adjusted in detail. The “Apply” button applies the changes. Finally, the user can click the “Generate Report” button to generate the parametric survival

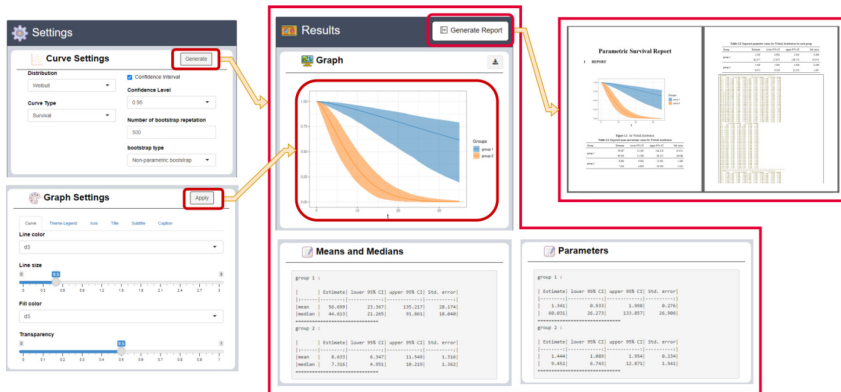


Figure 5
“Parametric” tab

analysis report in “.docx” format. At the same time, the resulting graphic can be downloaded separately in different measurement and file formats.

3.1.1. Survival

Let T be the survival time. Survival function denoted by $S(t)$; It is defined as the probability that an individual will survive longer than a certain time t .

$$S(t) = P(\text{“the probability of an individual surviving longer than } t\text{”}) = P(T > t)$$

The $S(t)$ function is also known as the cumulative survival rate. The $S(t)$ plot is called the survival curve.

3.1.2. Failure

“Failures” or “deaths” in survival analysis are the “detects” of environmental science—the observations with a single known value [16]. Failure function is denoted by $F(t)$ and equals to the probability $P(T < t)$, which is $1 - S(t)$.

3.1.3. Density

The probability density function for the time of death of an individual is expressed in

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P[\text{Individual's death in the interval } [t, t + \Delta t)]}{\Delta t}$$

with the limit definition of the probability that the individual's death will be a small $[t, t + \Delta t)$ of width Δt . The probability density function is also known as the “unconditional failure rate”.

If the probability density function $f(t)$ for a continuous random variable T defined in the interval $[0, \infty)$ is represented by $f(t)$, this function is the derivative of the cumulative distribution function $F(t)$, if any [18].

3.1.4. Hazard

The hazard function $h(t)$ of survival time T gives the conditional failure rate. The Hazard (hazard, risk) function T is a conditional distribution of survival time. Given that an asset lives to time t , the probability that it will die in a given time (Δt) is represented by the hazard function $h(t)$,

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P \left[\begin{array}{c} \text{Failure of the individual in the time interval } (t, t+\Delta t), \\ \text{given that they have survived until the time } t \end{array} \right]}{\Delta t}.$$

While the survival function focuses on the absence of an event, the hazard function, on the contrary, focuses on the probability of the event occurring [25].

The cumulative hazard function is

$$H(t) = \int_0^t h(x) dx.$$

At the same time, between the cumulative hazard function and the survival function is

$$H(t) = -\log S(t). \quad (1)$$

In Equation 1, for $t = 0$, $S(t) = 1$. If $H(t) = 0$, then $S(t) = 0$ and $H(t) = \infty$ for $t = \infty$. The cumulative hazard function can take any value between 0 and ∞ .

3.2 Non-Parametric Methods

The Kaplan-Meier estimator [19], also known as the product limit estimator, is a non-parametric statistic used to estimate survival function from survival data. In medical research, it is often used to measure the proportion of patients who live for a certain period of time after treatment.

Survival function of Kaplan-Meier estimator (probability of survival longer than t) is

$$\hat{S}(t) = \prod_{j=1}^k \left[\frac{n_j - d_j}{n_j} \right], \quad k \leq n; t_{(j)} \leq t \leq t_{(j+1)}.$$

Here, n_j is the number of individuals in the transaction at the j -th point, and d_j is the number of individuals who died at the j -th point. k is the number of observed observations and n is the total number of individuals in the research. The standard error of this estimate is

$$\sigma_{S(t)} = \hat{S}(t) \sqrt{\sum_{j=1}^k \frac{d_j}{n_j(n_j - d_j)}}, \quad k \leq n; t_{(j)} \leq t \leq t_{(j+1)}$$

Confidence interval can be calculated directly with $\hat{S}(t) \pm Z_{1-\alpha/2} \sigma_{S(t)}$. However, since this confidence interval can take values outside the $[0, 1]$ interval, alternative methods have been proposed. The types of confidence intervals that ShinySurv can calculate are linear, Peto, log, loglog, Rothman, Grunkemeier, Hall-Wellner, log Hall-Wellner, Equal Precision, and log

Equal Precision. These methods are calculated in R with the help of km.ci package [31]. Cumulative hazard function and standard error by Nelson-Aalen method is

$$\hat{H}(t) = \sum_{t_j \leq t} \frac{d_j}{n_j}, \sigma_{H(t)} = \sqrt{\sum_{t_j \leq t} \frac{d_j}{n_j^2}} \quad (2)$$

Turnbull's algorithm for left censorship consists of four steps. Table 1 shows Turnbull's algorithm steps. The final $\hat{d}$ obtained can be substituted in Equation 2 to obtain the cumulative hazard function for the left censor. Confidence intervals can also be obtained by replacing d with $\hat{d}$.

Table 1

Turnbull's survival function algorithm for the last censored data

Step 1: Calculate initial $\hat{S}_K(t)$.

Step 2: Calculate $p_{ij} = P[t_j - 1 < X \leq t_i | X \leq t_i]$ using $\frac{S_K(t_{j-1}) - S_K(t_j)}{1 - S_K(t_i)}, j \leq i$

Step 3: Calculate the number of events in t_i , $\hat{d}_i = d_i + \sum_{i=j}^m c_i p_{ij}$.

Step 4: Calculate $\hat{S}_{K+1}(t) = \prod_{j=1}^k \left[\frac{n_j - \hat{d}_j}{n_j} \right]$ using $\hat{d}_i$. If S_{K+1} is not close enough to $S_K(t)$, return to Step 2.

Turnbull's proposed method for interval censorship lists the start and end time times as a_i and b_i , $0 = \tau_0 < \tau_1 < \dots < \tau_m$. The i -th sample is given weight α_{ij} . α_{ij} takes the value 1 if $(\tau_{j-1}, \tau_j]$ contains the interval, $(a_i, b_i]$, and 0 if it does not. Then the five-step algorithm is passed. Table 2 shows the steps of the algorithm. The cumulative hazard function for the

Table 2

Turnbull's survival function algorithm for interval censored data

Step 1: Calculate initial $\hat{S}_K(t)$.

Step 2: Calculate $p_j = S(\tau_{j-1}) - S(\tau_j), j = 1, \dots, m$ for τ_j .

Step 3: Calculate $d_i = \sum_{i=1}^n \frac{\alpha_{ij} p_j}{\sum_k \alpha_{ik} p_k}$

Step 4: Calculate $Y = \sum_{k=j}^m d_k$.

Step 5: Calculate $\hat{S}_{K+1}(t) = \prod_{j=1}^k \left[\frac{n_j - d_j}{n_j} \right]$. If S_{K+1} is not close enough to $S_K(t)$, return to Step 2.

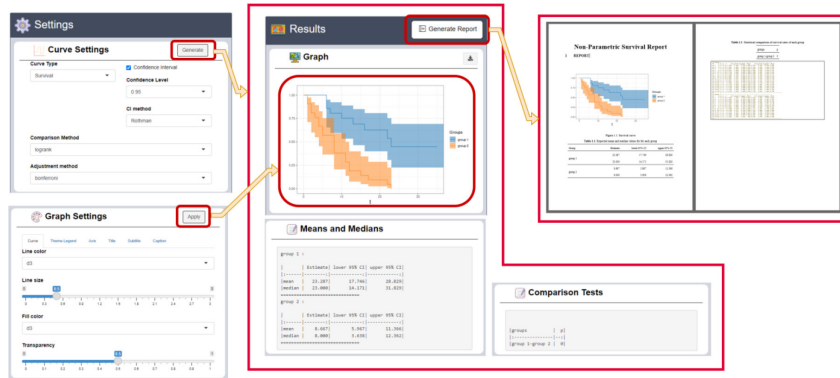


Figure 6
“Non-Parametric” tab

left censor can be obtained by substituting in the final d in Equation 2. Confidence intervals are obtained in the same way.

The “Non-Parametric” tab of the ShinySurv is shown in Figure 6. In the “Curve Settings” box, confidence interval and hypothesis test options are determined after the curve type. The survival graph, mean and median values and comparison test results are obtained by clicking the Generate button. In non-parametric methods, statistical tests are used to compare survival rates between groups. The absence hypothesis is based on the fact that there is no difference between the survival curves of the groups. Logrank and Wilcoxon tests are available to compare survival curves between groups. For multiple comparison methods, importance level corrections can be made under the “adjustment method” heading. The methods added for this are “bonferroni”, “holm”, “hochberg”, “hommel”, “BH”, “fdr” methods are taken from the stats package in R [7]. “Graph Settings” and “Generate Report” are the same as their counterparts in the “Parametric” tab.

ShinySurv allows the user to report and save their results. It also provides interactive charts that display values on graphs and can be configured for download by users.

4. Comparison with Similar Softwares

Programs such as SAS, Stata SPSS and Minitab, which are frequently preferred for statistical tests, make survival analysis easier for users. However, basic statistics must be known in order to perform survival analysis in these programs, and these programs are licensed, which makes

it difficult for users to access. In SAS, it is possible to analyze left, right and interval censored data with the LIFEREG procedure. Censorship types are determined by different MODEL definitions [3]. STATA survival analysis reports provides some information about left censorship, however the applications are insufficient. For example, it cannot handle the doubly censored data [11, 30]. SPSS is an easy-to-apply statistical program for social sciences. Nevertheless, it is frequently used in health sciences and natural sciences. The fact that the SPSS does not require code writing and has sufficient resources for output comments is an important reason for preference. SPSS can perform Kaplan-Meier and Cox Regression, one of the survival analysis methods, for right censoring [21, 6]. However, it cannot perform survival analysis for other types of censorship. In the Minitab program, right, left and interval censoring can be conducted. While applying left censorship, the data is converted to be right censored and the curve is obtained in this way [16]. However Minitab cannot consider double censor. ShinySurv can handle all types of censorship and report both parametric and non-parametric survival results and curves.

Survival analysis is conducted in package programs such as SAS, STATA, SPSS and Minitab programs. Different types of censorship are also included in them. However, “R” is always ahead of these programs due to its developing structure and ease of access. Studies for left censoring [1,8], interval censoring [2,15] and other types of censoring [26] is frequently encountered. The most commonly used packages are survival [32], KMsurv [20], npsurv [34], NADA [22], dblcens [40], and icenreg [2]. But, the application of different methods and censorship types in each package causes confusion for users. Also, these packages are weak on their own in terms of visualization and require the visualization skills of the user with different visualization packages. Having too many packages also makes it difficult to implement in R [25]. ShinySurv does not need any coding and contains all types of censorship. It also offers many options to the user in terms of visualization.

Apart from package programs and R packages, there are web applications made with Shiny. Examples of these web applications are MEPHAS, MSM.app, SMART, and Survival Genie. None of these web applications allow different types of censorship. In addition, it cannot perform parametric survival analysis. Table 3 compares ShinySurv and the mentioned software according to their different features. ShinySurv has the feature of being free of R packages and Shiny web applications. It is easy for users to access and requires no coding. As can be seen in Table 3, alternative software that addresses left censorship types are scarce.

Table 3
Comparison of ShinySurv with available programs

Application/ package/ software	Payment required	Coding required	Right censor	Left censor	Double-censor	Interval censor	Adjustable graph	Parametric	Non- parametric
ShinySurv	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SAS	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
STATA	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
SPSS	Yes	No	Yes	No	No	No	Yes	Yes	Yes
Minitab	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes
survival	No	Yes	Yes	Yes*	No	Yes*	No	Yes	Yes
KMSurv	No	Yes	Yes*	No	No	No	No	No	Yes
npsurv	No	Yes	No	No	No	Yes*	No	No	Yes
NADA	No	Yes	No	Yes*	No	No	No	No	Yes
dblcens	No	Yes	Yes*	Yes*	Yes*	No	No	No	Yes
icenreg	No	Yes	No	No	No	Yes	No	Yes	Yes
MEPHAS	No	No	Yes*	No	No	No	No	No	Yes
MSM.app	No	No	Yes*	No	No	No	No	No	Yes
SMART	No	No	Yes*	No	No	No	No	No	Yes
Survival Genie	No	No	Yes*	No	No	No	No	No	Yes

*only non-parametric

Other than ShinySurv, only dblcens can handle double censorship. ShinySurv can handle all types of censorship both parametrically and non-parametrically. Besides, ShinySurv can report analysis results in “docx” format.

5. Conclusion and Future Works

Different types of censorship are applied in survival analysis. These censorship types vary according to the structure of the data, and the resulting survival times are also affected. ShinySurv, the web-based application we have developed, allows the analysis of more censorship types compared to other programs. In addition, the analysis of parametric and non-parametric methods applied in survival analysis and the interpretation of their outputs have been facilitated. Analysis outputs can be easily obtained with the report file created by ShinySurv. ShinySurv complements and combines the shortcomings of other programs in generating survival curves.

ShinySurv does not yet include regression models. In future studies, the methods of survival analysis such as cox regression, accelerated failure times, and proportional hazards are open to be added. In addition, we are planning to add truncation data types to ShinySurv.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

We have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] J. P. Albaladejo, J. Antoni, and M. Fernández, “zCompositions — R package for multivariate imputation of left-censored data under a compositional approach,” *Chemometrics and Intelligent Laboratory Systems*, vol. 143, pp. 85-96 (2015). doi: <https://doi.org/10.1016/j.chemolab.2015.02.019>.
- [2] C. Anderson-Bergman, “icenReg: Regression Models for Interval Censored Data in R,” *Journal of Statistical Software*, vol. 81, pp. 1-23 (2017). doi: <https://doi.org/10.18637/jss.v081.i12>.
- [3] P. D. Allison, *Survival Analysis Using SAS: A Practical Guide*, USA: SAS Institute (2010).
- [4] A. K. Arslan, Ş. Yaşar, C. Çolak, and S. Yoloğlu, “WSSPAS: an interactive web application for sample size and power analysis with R using shiny,” *Türkiye Klinikleri Biyoistatistik*, vol. 10, no. 3, pp. 224-246 (2018). doi: <https://doi.org/10.5336/biostatic.2018-62787>.
- [5] A. K. Arslan, Ş. Yaşar, C. Çolak, and S. Yoloğlu, “R Shiny Paketi ile Kruskal Wallis H Testi için İnteraktif Bir Web Uygulaması,” *İnönü Üniversitesi Sağlık Bilimleri Dergisi*, vol. 7, no. 2, pp. 49-55 (2018).
- [6] B. Barton and J. Peat, *Medical Statistics: A Guide to SPSS, Data Analysis and Critical Appraisal*, John Wiley & Sons (2014).
- [7] A. Bunn and M. Korpela, “R: A language and environment for statistical,” [Online]. Available: <https://www.R-project.org/>.
- [8] R. A. Canales, A. M. Wilson, J. I. Pearce-Walker, M. P. Verhougstraete, and K. A. Reynolds, “Methods for Handling Left-Censored Data in Quantitative Microbial Risk Assessment,” *Public and Environmental Health Microbiology*, vol. 84, no. 20, e01203-18 (2018). doi: <https://doi.org/10.1128/AEM.01203-18>.
- [9] R. Chopra, K. Kapoor, S. Madan, and R. Sibal, “Improving quality of web applications deployed on cloud,” *Journal of Information and Optimization Sciences*, vol. 41, no. 6, pp. 1299-1308 (2020). doi: [10.1080/02522667.2020.1809088](https://doi.org/10.1080/02522667.2020.1809088).

- [10] Y.-C. Chu, W.-T. Kuo, Y.-R. Cheng, F. C. Lin, C.-Y. Lee, C.-Y. Shiau, and F. Lai, "SMART Survival Metadata Analysis Responsive Tool," in 2017 XXVI International Conference on Information, Communication and Automation Technologies (ICAT), pp. 1-6 (2017).
- [11] M. Cleves, W. Gould, R. G. Gutierrez, and Y. V. Marchenko, *An Introduction to Survival Analysis Using Stata*, Stata Press (2010).
- [12] D. Collett, *Modelling Survival Data in Medical Research*, CRC Press (2015).
- [13] B. Dwivedi, H. Mumme, S. Satpathy, S. S. Bhasin, and M. Bhasin, "Survival Genie, a web platform for survival analysis across pediatric and adult cancers," *Sci. Rep.* (2022). doi: <https://doi.org/10.1038/s41598-022-06841-0>.
- [14] J. Edmunson, T. Fleming, D. Decker, G. Malkasian, J. Jefferies, M. Webb, and L. Kvols, "Different Chemotherapeutic Sensitivities and Host Factors Affecting Prognosis in Advanced Ovarian Carcinoma vs. Minimal Residual Disease," *Cancer Treatment Reports*, vol. 63, pp. 241-247 (1979).
- [15] G. Gomez, M. L. Calle, and R. O. Langohr, "Tutorial on methods for interval-censored data and their implementation in R," *Stat. Modelling*, vol. 9, no. 4, pp. 259-297 (2009). doi: <https://doi.org/10.1177/1471082X0900900402>.
- [16] D. R. Helsel, *Statistics for Censored Environmental Data Using Minitab and R*, John Wiley & Sons (2011).
- [17] T. Holland-Letz and A. Kopp-Schneider, "An R-Shiny application to calculate optimal designs for single substance and interaction trials in dose response experiments," *Toxicol. Lett.* (2020). doi: <https://doi.org/10.1016/j.toxlet.2020.11.018>.
- [18] S. P. Jenkins, "Survival analysis," unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester, UK, pp. 42, 54-56 (2005).
- [19] E. L. Kaplan and P. Meier, "Nonparametric Estimation From Incomplete Observations," *J. Am. Stat. Assoc.*, pp. 457-481 (1958).
- [20] J. P. Klein and M. L. Moeschberger, *Survival Analysis Techniques for Censored and Truncated Data*, New York: Springer (2003).
- [21] S. Landau and B. S. Everitt, *A Handbook of Statistical Analyses Using SPSS*, Chapman and Hall/CRC (2004). doi: <https://doi.org/10.1201/9780203009765>.

- [22] L. Lee, "NADA: Nondetects and Data Analysis for Environmental Data," R package version 1.6-1.1 (2020). [Online]. Available: <https://CRAN.R-project.org/package=NADA>.
- [23] M. Li, W. Susilo, and J. Tonien, "The code for securing web applications," *J. Inf. Optim. Sci.*, vol. 40, no. 4, pp. 905-917 (2019). doi: <https://doi.org/10.1080/02522667.2018.1468609>.
- [24] C. L. Loprinzi, J. A. Laurie, H. S. Wieand, J. E. Krook, P. J. Novotny, J. W. Kugler, et al., "Prospective evaluation of prognostic variables from patient-completed questionnaires. North Central Cancer Treatment Group," *J. Clin. Oncol.*, vol. 12, no. 3, pp. 601-607 (1994). doi: <https://doi.org/10.1200/JCO.1994.12.3.601>.
- [25] D. F. Moore, *Applied Survival Analysis Using R*, Springer (2016).
- [26] R. Oller and K. Langohr, "FHtest: An R package for the comparison of survival curves with censored data," *J. Stat. Softw.*, vol. 81, pp. 1-25 (2017). doi: <https://doi.org/10.18637/jss.v081.i15>.
- [27] A. Özhan, M. Tombaz, and Ö. Konu, "SmulTCan: A Shiny application for multivariable survival analysis of TCGA data with gene sets," *Comput. Biol. Med.*, vol. 137, p. 104793 (2021). doi: <https://doi.org/10.1016/j.combiomed.2021.104793>.
- [28] N. Satyahadewi and H. Perdana, "Web application development for inferential statistics using R Shiny," *Adv. Soc. Sci. Educ. Humanities Res.*, vol. 550 (2020). doi: <https://doi.org/10.2991/assehr.k.210508.099>.
- [29] G. Soutinho and L. M. Machado, "Analysis of complex survival data: A tutorial using the Shiny MSM.app application," *arXiv*, pp. 1-26 (2022). doi: <https://doi.org/10.48550/arXiv.2202.09160>.
- [30] STATA, "Survival analysis features," April 16 (2022). [Online]. Available: <https://www.stata.com/features/survival-analysis/>.
- [31] R. Strobl, "km.ci: Confidence intervals for the Kaplan-Meier estimator," R package version 0.5-2 (2009). [Online]. Available: <https://CRAN.R-project.org/package=km.ci>.
- [32] T. M. Therneau and P. M. Grambsch, "The Cox model," in *Modeling Survival Data: Extending the Cox Model*, Springer, New York, NY, pp. 39-77 (2000). doi: https://doi.org/10.1007/978-1-4757-3294-8_3.
- [33] B. W. Turnbull, "Nonparametric estimation of a survivorship function with doubly censored data," *J. Am. Stat. Assoc.*, vol. 69, no. 345, pp. 169-173 (1974).

- [34] Y. Wang, “npsurv: Nonparametric Survival Analysis,” R package version 0.5-0 (2020). [Online]. Available: <https://CRAN.R-project.org/package=npsurv>.
- [35] H. Wang and L. Zhou, “SurvELM: An R package for high dimensional survival analysis with extreme,” *Knowl.-Based Syst.*, vol. 160, pp. 28-33 (2018). doi: <https://doi.org/10.1016/j.knosys.2018.07.009>.
- [36] B. Wei, W. Zhong, R. Liu, A. Wu, A. Chiang, M. Branson, and N. Ge, “A comprehensive toolbox to facilitate quantitative decision science in drug development: A web-based R shiny application GOahead,” *arXiv preprint arXiv:2201.11267* (2022). doi: <https://doi.org/10.48550/arXiv.2201.11267>.
- [37] B. D. Wissel, P. V. Camp, M. Kouril, C. Weis, T. A. Glauser, P. S. White, et al., “An interactive online dashboard for tracking COVID-19 in U.S. counties, cities, and states in real time,” *J. Am. Med. Inform. Assoc.*, vol. 27, no. 7, pp. 1121-1125 (2020). doi: <https://doi.org/10.1093/jamia/ocaa071>.
- [38] Ş. Yaşar, A. K. Arslan, C. Colak, and S. Yoloğlu, “A developed interactive web application for statistical analysis: Statistical Analysis Software,” *Middle Black Sea J. Health Sci.*, vol. 2, no. 6, pp. 227-239 (2020). doi: <https://doi.org/10.19127/mbsjohs.704456>.
- [39] Ş. Yaşar, F. H. Yağın, A. K. Arslan, C. Colak, and S. Yoloğlu, “Interactive web-based software for evaluating diagnostic tests and ROC curve analyses in health sciences,” *Ann. Med. Res.*, vol. 28, no. 11, pp. 2012-2018 (2021). doi: <https://doi.org/10.5455/annalsmedres.2021.01.103>.
- [40] M. Zhou, L. Lee, and K. Chen, “dblcens: Compute the NPMLE of distribution from doubly censored data,” R package version 1.1.7 (2012). [Online]. Available: <https://CRAN.R-project.org/package=dblcens>.
- [41] Y. Zhou, S.-W. Leung, S. Mizutani, and T. T.-S. Tian, “MEPHAS: An interactive graphical user interface for medical and pharmaceutical statistical analysis with R and Shiny,” *Bioinformatics*, vol. 183, no. 21, pp. 1-11 (2020). doi: <https://doi.org/10.1186/s12859-020-3494-x>.

Received June, 2022

Revised September, 2022