

How to Easily Perform a Log Rank Test in R

Authored by
stats writer

December 6, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Easily Perform a Log Rank Test in R*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=106373>

The Log-Rank Test is a fundamental and widely utilized non-parametric statistical procedure employed in survival analysis. Its primary objective is to rigorously compare the survival distributions--or the probability of an event occurring over time--across two or more distinct groups of subjects. This powerful technique is indispensable in fields like clinical trials, epidemiology, and reliability engineering where the timing of an event, such as death, recurrence of disease, or machine failure, is the key outcome of interest. In the context of the R programming environment, this test is executed using the `survdiff()` function, which is housed within the essential survival package. The process requires supplying the function with a well-defined survival object, which encapsulates the crucial elements of time-to-event data and the associated event status, alongside a vector defining the categorical group assignments (e.g., treatment A versus treatment B). The culmination of this process is the generation of a p-value, which serves as the quantitative basis for determining the statistical significance of any observed differences in survival profiles between the comparison groups.

Understanding the Log-Rank Test in Survival Analysis

The Log-Rank Test, also frequently referred to as the Mantel-Cox test, stands as the most common and robust method for comparing two or more survival curves, often visualized using Kaplan-Meier plots. Survival analysis, unlike standard regression analysis, handles two unique features of time-to-event data: the time scale itself and the issue of censoring. Censoring occurs when the exact time of the event is unknown for some subjects, perhaps because they dropped out of the study, were lost to follow-up, or the study concluded before the event occurred. The Log-Rank Test efficiently utilizes information from both uncensored and censored observations, enabling a reliable comparison of survival probabilities over the entire observation period.

The fundamental principle behind the Log-Rank Test involves comparing the observed number of events (deaths or failures) in each group to the number of events statistically expected if the survival curves for all groups were, in fact, identical. This comparison is calculated sequentially at every time point where an event occurs. The test statistic generated summarizes these differences across all event times, providing a comprehensive measure of disparity between the survival experiences of the groups under comparison. A larger difference between observed and expected events yields a larger test statistic, which subsequently corresponds to a smaller p-value, suggesting greater evidence against the null hypothesis of no difference.

In practice, utilizing the Log-Rank Test demands careful consideration of its underlying assumptions. While it is non-parametric, meaning it does not require assumptions about the shape of the underlying survival distribution (e.g., normal or exponential), it does assume that the ratio of the hazard rates between the groups is constant over time. This crucial assumption is known as the proportional hazards assumption. While violations of this assumption do not invalidate the test

entirely, they may suggest that alternative tests, such as the Peto-Prentice or the weighted Log-Rank tests, might offer better power or more appropriate interpretations, particularly if the treatment effect changes significantly late in the follow-up period.

The Null and Alternative Hypotheses

Like all rigorous statistical hypothesis tests, the Log-Rank Test is framed around two mutually exclusive statements regarding the population parameters. These hypotheses guide the interpretation of the test results and dictate the conclusions drawn about the survival data. Understanding the precise formulation of these hypotheses is critical for accurate reporting and scientific inference, especially when comparing the efficacy of different treatments or prognostic factors.

The **Null Hypothesis** (H_0) asserts that there is no meaningful or statistically significant difference in the underlying survival distributions between the groups being compared. This hypothesis implies that, regardless of treatment assignment or group classification, the probability of survival at any given time point is the same for all subjects across the groups. If the test yields a high p-value, we fail to reject this hypothesis, concluding that the data does not provide sufficient evidence to suggest different survival profiles.

H₀: There is no difference in survival distributions or survival experience between the two groups.

The **Alternative Hypothesis** (H_A) is the counterpoint, stating that a statistically significant difference in survival does exist between the groups. This means that the survival curves are not identical and that subjects in one group have a higher or lower probability of surviving compared to subjects in the other group. When the calculated p-value is sufficiently small--typically less than the predetermined significance level (alpha, often set at $\alpha = .05$)--we reject the null hypothesis in favor of the alternative. This rejection provides strong statistical evidence to conclude that the observed difference in survival is unlikely to be due merely to random chance.

H_A: There *is* a difference in survival distributions or survival experience between the two groups.

Crucially, if the p-value of the test is less than the chosen significance level (e.g., $\alpha = .05$), then we can confidently reject the null hypothesis. The practical implication is that we have sufficient evidence to assert that the intervention (e.g., the treatment type) has had a measurable and meaningful impact on the survival times of the subjects. This conclusion forms the basis for medical recommendations or policy changes derived from the survival study.

Prerequisites: The R Programming Environment and the Survival Package

To successfully execute a Log-Rank Test, or any survival analysis procedure, within the R programming environment, researchers must rely on specialized statistical packages developed for this purpose. The primary tool is the survival package, which is maintained by Terry Therneau and provides a comprehensive suite of functions for handling time-to-event data, including the creation of survival objects, fitting Cox proportional hazards models, and performing hypothesis tests like the Log-Rank Test.

Before any analysis can commence, the survival package must be installed and loaded into the current R session. Installation is typically a one-time process using the `install.packages()` function, while loading the package using the `library()` function is required at the beginning of every new session where the functions are needed. This step ensures that R recognizes commands such as `Surv()`, which creates the survival object, and `survdiff()`, which executes the test itself. Neglecting this crucial loading step will result in R returning an error, indicating that the function is not found.

The core of the data preparation involves the `Surv()` function. This function constructs the survival object required by subsequent analysis functions. It typically takes two primary arguments: the time variable (the duration until the event or censoring) and the event status indicator (a binary variable specifying whether the event occurred or if the observation was censored). Correctly defining the survival object is paramount, as it accurately captures the longitudinal nature and the status of the patient at the end of follow-up, which are the unique characteristics distinguishing survival data from standard outcome data.

Syntax and Functionality of `survdiff()`

The `survdiff()` function is the designated tool within the survival package for conducting the Log-Rank Test. It uses a standard formula syntax commonly employed across R's statistical modeling functions, making it highly intuitive for users familiar with the R environment. Understanding this syntax is key to correctly specifying the model for comparison.

The general syntax for the `survdiff()` function is defined as follows, where the structure mirrors that of other formula-based functions in R:

```
survdiff(Surv(time, status) ~ predictors, data)
```

In this structure, `Surv(time, status)` represents the survival object, which must be created beforehand using the time-to-event variable and the event indicator variable. The tilde symbol (`~`) separates the survival outcome from the grouping variable(s). The `predictors` element refers to the categorical variable(s) that define the groups whose survival curves are being compared (e.g., treatment group or disease stage). Finally, the `data` argument specifies the data frame containing all the necessary variables.

Upon execution, the `survdiff()` function returns a comprehensive output that includes the key results of the Log-Rank Test. Specifically, it provides the **Chi-Squared test statistic** and its associated **degrees of freedom**. The Chi-Squared statistic measures the overall discrepancy between the observed and expected number of events across all groups. Most importantly, the function reports the corresponding **p-value**, which is essential for hypothesis testing and making a final determination regarding the equality of the survival curves.

Practical Application: Analyzing the `ovarian` Dataset

To illustrate the practical application of the Log-Rank Test, we will utilize a built-in dataset commonly used in R survival examples. The `ovarian` dataset, which is readily available within the survival package, provides a concise yet realistic example of clinical trial data. This dataset contains crucial clinical and demographic information for 26 patients diagnosed with ovarian cancer who participated in a clinical study.

The `ovarian` dataset includes several variables pertinent to survival analysis, making it an excellent candidate for demonstrating the Log-Rank Test. The key variables we must focus on are:

Survival time (`futime`): Represents the duration, measured in months, following the initial diagnosis of ovarian cancer. This is the time component of our survival object.

Event Status (`fustat`): Represents the event status indicator. A value of 1 signifies that the event (death) occurred, while a value of 0 indicates censoring (the patient was still alive or lost to follow-up).

Treatment Type (`rx`): Denotes the type of treatment received (`rx = 1` or `rx = 2`). This is the primary categorical predictor variable we will use to define our comparison groups.

Before proceeding with the test, it is prudent to inspect the structure and initial rows of the data. This preliminary step helps confirm that the variables are correctly formatted and that the data frame is loaded properly, ensuring the subsequent statistical analysis is based on accurate inputs. The following code demonstrates how to load the necessary package and view the structure of the data:

```
library(survival)
```

```
#view first six rows of dataset
```

```
head(ovarian)
```

```
futime fustat age resid.ds rx ecog.ps
```

```
1 59 1 72.3315 2 1 1
```

```
2 115 1 74.4932 2 1 1
```

```
3 156 1 66.4658 2 1 2
```

```
4 421 0 53.3644 2 2 1
5 431 1 50.3397 2 1 1
6 448 0 56.4301 1 1 2
```

The visual inspection confirms that `futime` (time), `fustat` (status), and `rx` (group) are available and correctly represented in the data frame, setting the stage for the formal hypothesis test comparing the survival experiences of patients receiving treatment 1 versus treatment 2.

Executing the Test and Interpreting Output

The core objective of this analysis is to determine if the survival distributions differ significantly between the two treatment groups defined by the `rx` variable. We formalize this by performing the Log-Rank Test using the `survdiff()` function. The survival object is constructed using `futime` as the time variable and `fustat` as the event indicator, while `rx` serves as the grouping factor.

The following R code snippet executes the required statistical procedure. We are instructing R to compare the survival curve (defined by time and status) across the levels of the treatment variable, utilizing the ovarian dataset:

```
#perform log rank test
survdiff(Surv(futime, fustat) ~ rx, data=ovarian)
```

Call:

```
survdiff(formula = Surv(futime, fustat) ~ rx, data = ovarian)
```

```
N Observed Expected (O-E)^2/E (O-E)^2/V
```

```
rx=1 13 7 5.23 0.596 1.06
```

```
rx=2 13 5 6.77 0.461 1.06
```

```
Chisq= 1.1 on 1 degrees of freedom, p= 0.3
```

The output generated by `survdiff()` provides rich detail essential for interpretation. For each group (`rx=1` and `rx=2`), we see the total number of subjects (N), the total observed number of events (Observed), and the total expected number of events (Expected) if the null hypothesis were true. For group 1, 7 events were observed versus 5.23 expected; for group 2, 5 events were observed versus 6.77 expected.

The final line of the output summarizes the hypothesis test results. The calculated ****Chi-Squared test statistic is 1.1****, derived from the accumulated differences between observed and expected counts, with ****1 degree of freedom****. Crucially, the corresponding ****p-value is 0.3****. Using the conventional significance threshold of $\alpha = 0.05$, we compare the calculated p-value (0.3) to

this threshold. Since 0.3 is significantly greater than 0.05 , we **fail to reject the null hypothesis**. The statistical conclusion is that there is insufficient evidence to claim a statistically significant difference in the overall survival distributions between patients receiving treatment 1 and patients receiving treatment 2.

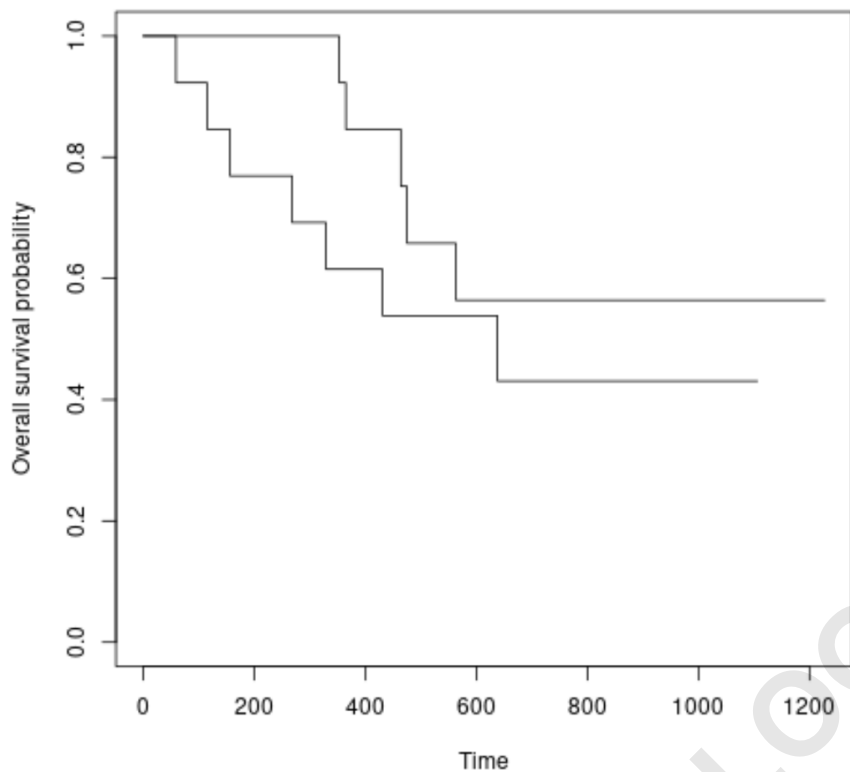
Visualizing Survival Curves for Comparison

While the Log-Rank Test provides a definitive statistical conclusion based on the calculated p-value, it is standard practice in survival analysis to visualize the survival experience through Kaplan-Meier curves. These plots offer a clear, intuitive representation of the survival probability over time for each group, complementing the numerical output of the test.

To generate these plots, we first use the `survfit()` function from the survival package to estimate the Kaplan-Meier survival curves for each treatment group. We then pass this fitted survival object to the `plot()` function. The syntax below specifies the creation of the survival fit based on the `futime` and `fustat` variables, stratified by the `rx` treatment group, and customizes the axis labels for clarity:

```
#plot survival curves for each treatment group  
plot(survfit(Surv(futime, fustat) ~ rx, data = ovarian),  
xlab = "Time",  
ylab = "Overall survival probability")
```

The resulting plot visually represents the survival probability (y-axis) as a function of time in months (x-axis) for both treatment regimens. Each step down on the curve corresponds to an observed event (death). The visual representation below allows for immediate qualitative comparison:



Upon examining the resulting Kaplan-Meier plot, we can qualitatively observe that the two survival curves appear slightly separated, suggesting a potential difference in outcomes between Treatment 1 and Treatment 2. Specifically, the survival curves are slightly different, but the Log-Rank Test mathematically confirms that this observed visual separation is within the range of expected random variation. Therefore, although there are minor visual differences in survival patterns, the Log-Rank Test concludes that the disparity is not substantial enough to be declared **statistically significant**, reinforcing the finding that neither treatment provides a demonstrably superior survival benefit over the other based on this dataset.