

Perform a Log Rank Test in R

Authored by
Mohammed looti

November 5, 2025

RECOMMENDED CITATION

Mohammed looti (2025). *Perform a Log Rank Test in R*. PSYCHOLOGICAL STATISTICS.
Retrieved from <https://statistics.arabpsychology.com/?p=10529>

Introduction to the Log Rank Test in Survival Analysis

In the specialized field of [survival analysis](#), a core methodological requirement is the ability to rigorously compare the survival experiences--or time-to-event outcomes--across two or more distinct cohorts. Researchers, particularly those involved in clinical trials and epidemiological studies, must determine whether differences observed in survival times between groups (e.g., treatment A vs. treatment B) are statistically meaningful or merely due to random chance. The [log rank test](#) emerges as the most widely accepted and robust non-parametric statistical method specifically designed for this purpose. It provides a powerful analytical framework for evaluating the equality of [survival curves](#) derived from independent groups.

The utility of this statistical tool extends across various disciplines where time until an event is the critical outcome variable. Whether assessing the efficacy of a new drug in prolonging patient life, analyzing the impact of an environmental exposure on disease recurrence, or determining how demographic factors influence equipment failure rates, the [log rank test](#) provides an essential, quantitative answer. It achieves this by contrasting the observed number of events (such as deaths or failures) in each group against the number of events that would be mathematically expected if all groups shared an identical underlying survival distribution. This careful comparison allows investigators to isolate the influence of the predictor variable--the treatment, exposure, or demographic factor--on the time-to-event outcome, ensuring accurate assessment in complex data environments.

Statistical Framework: Null and Alternative Hypotheses

The application of the [log rank test](#) is firmly rooted in the principles of [statistical hypothesis testing](#). Before any data analysis can commence, the researcher must clearly articulate the competing claims regarding the population parameters. This process involves establishing a specific framework that governs how the observed data will be used to make inferences about the broader population. The test is constructed to compare survival experiences based on the cumulative hazard function across the cohorts being analyzed.

The core hypotheses that guide the application and interpretation of the log rank test are formally defined to address the question of whether group membership affects the survival outcome. The test operates by assuming, initially, that no difference exists. The formal articulation of these competing claims is vital for framing the subsequent statistical interpretation:

H₀ (The Null Hypothesis): There is no underlying difference in the survival distributions among the groups being compared. Statistically, the survival curves are considered identical, implying that the grouping factor (e.g., treatment type) has no effect on the time until the event occurs.

H_A (The Alternative Hypothesis): There is a statistically significant difference in the survival

distributions between at least two of the groups under comparison. This suggests that the grouping factor does meaningfully influence the survival experience, causing the survival curves to diverge.

The final statistical decision to reject or retain the [null hypothesis](#) (H0) hinges entirely upon the computed [p-value](#). If this p-value falls below the pre-specified level of significance (commonly denoted as α , typically set at 0.05), we possess sufficient statistical evidence to reject H0. Rejecting the null hypothesis allows us to conclude that the differences observed in survival between the groups are highly unlikely to be random, thereby asserting a significant relationship between the grouping variable and the survival outcome.

Prerequisites and Setup for R Implementation

To effectively execute the [log rank test](#) within the dynamic [R statistical computing environment](#), researchers rely primarily on the robust functionality provided by the `survival` package. This package is foundational for handling time-to-event data in R, offering specialized functions optimized for survival modeling and comparison. The primary function employed for conducting the log rank test is `survdifftest()`, which is expertly designed to process survival objects and calculate the necessary test statistics.

The initial prerequisite for any survival analysis in R is ensuring that the `survival` library is loaded into the current session. Furthermore, the data must be structured appropriately, containing two key variables essential for defining the survival object: the time variable (duration of follow-up) and the status variable (indicating whether the event occurred or if the observation was [censored](#)). Understanding the data structure and loading the necessary library are non-negotiable preparatory steps before model execution.

The general syntax for the `survdifftest()` function adheres to the standard R formula structure, which is designed for clarity and flexibility in defining statistical relationships. This formula explicitly links the survival outcome to the predictor variables of interest, ensuring that the test accurately compares the groups defined by those predictors.

Detailed Syntax and Mechanics of the `survdifftest()` Function

The structure required to invoke the `survdifftest()` function is crucial for defining the scope of the log rank test. It dictates how R interprets the time-to-event data and which factor should be used for group comparison. The fundamental structure is defined as: `survdifftest(Surv(time, status) ~ predictors, data)`. This formula is divided into essential components, each playing a critical role in the computation of the test statistic.

The left side of the formula, `Surv(time, status)`, mandates the creation of a survival object. The `time` parameter represents the total duration from the start of observation until the event occurred

or until the observation was terminated. The `status` parameter is a binary indicator, typically coded as 0 for observations where the event did not occur (i.e., the observation was [censored](#)) and 1 for observations where the event did occur. Correctly defining this survival object is the cornerstone of accurate [survival analysis](#).

The right side of the formula, specified after the tilde (`~`), contains the `predictors`. This is the categorical grouping factor whose influence on survival we intend to test. The `survdiff()` function then systematically compares the observed events against the expected events at every distinct failure time across all levels of this predictor variable. This iterative comparison process yields the overall summary statistic, which is asymptotically distributed as a [Chi-Squared test statistic](#). The final output provides this statistic along with its degrees of freedom and the associated [p-value](#), enabling the formal statistical inference regarding the equality of survival curves.

Practical Application: Case Study Using the Ovarian Cancer Data

To move from theoretical understanding to practical application, we will employ a widely recognized dataset: the built-in `ovarian` dataset available within the R `survival` package. This dataset offers a concise yet comprehensive set of survival information for 26 patients diagnosed with ovarian cancer. Utilizing this real-world data allows us to demonstrate precisely how the log rank test is executed and how the input variables map onto the statistical requirements.

The key variables within the `ovarian` dataset that are essential for conducting the log rank test are clearly defined:

`futime`: The survival time, measured in months, representing the duration of follow-up after the cancer diagnosis. This serves as our time variable.

`fustat`: The event status, coded as 1 if the patient's event (e.g., death or recurrence) occurred, and 0 if the observation was [censored](#). This is our status variable.

`rx`: The type of treatment received by the patient, categorized as either Treatment 1 (`rx = 1`) or Treatment 2 (`rx = 2`). This serves as our grouping predictor.

As a fundamental step in any data analysis workflow, we must first load the necessary library and inspect the dataset structure to confirm variable names and formats. The code snippet below illustrates the command to load the `survival` package and displays the initial observations, ensuring proper identification of the survival components (`futime` and `fustat`) and the grouping factor (`rx`). Our primary analytical objective is to rigorously determine if the survival experience, defined by these time-to-event variables, differs significantly based on the assigned treatment type.

```
library(survival)
```

```
# View first six rows of dataset to confirm structure
```

```
head(ovarian)

fuptime 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
```

Executing and Interpreting the R Output

With the data prepared and the objective clearly defined--testing for a significant difference in survival between treatment groups $rx=1$ and $rx=2$ --we proceed to execute the **log rank test** using the established `survdifft()` syntax. This command integrates the survival object `Surv(fuptime, fustat)` with the grouping factor `rx`, directing the R environment to perform the necessary statistical comparison based on observed and expected events.

```
# Perform log rank test comparing survival by treatment group
survdifft(Surv(fuptime, fustat) ~ rx, data=ovarian)
```

Call:

```
survdifft(formula = Surv(fuptime, 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 resulting output provides a comprehensive summary of the test statistics for each group. Specifically, it details the number of subjects (N), the number of observed events (Observed), and the number of events that would have been statistically anticipated if the **null hypothesis** of equal survival were true (Expected). Analyzing these counts reveals the nature of the difference: for group $rx=1$, 7 events were observed compared to 5.23 expected, suggesting slightly worse survival than anticipated under the null; conversely, group $rx=2$ observed 5 events versus 6.77 expected, indicating slightly better survival than anticipated. The columns $(O-E)^2/E$ and $(O-E)^2/V$ contribute to the overall weighted sum that constitutes the test statistic.

The most critical information for drawing a statistical conclusion is presented in the final line of the

output. Here, the overall test result is summarized: the calculated **Chi-Squared test statistic** is **1.1**, based on 1 degree of freedom (which is calculated as the number of groups minus one). Crucially, the associated **p-value** is **0.3**. Given that 0.3 is substantially greater than the conventional significance threshold of $\alpha = 0.05$, we lack sufficient evidence to reject the **null hypothesis**. Therefore, the statistical conclusion dictates that the observed differences in survival outcomes between the two distinct treatment groups ($rx=1$ and $rx=2$) are not statistically significant, implying that neither treatment can be confidently deemed superior based on this dataset alone.

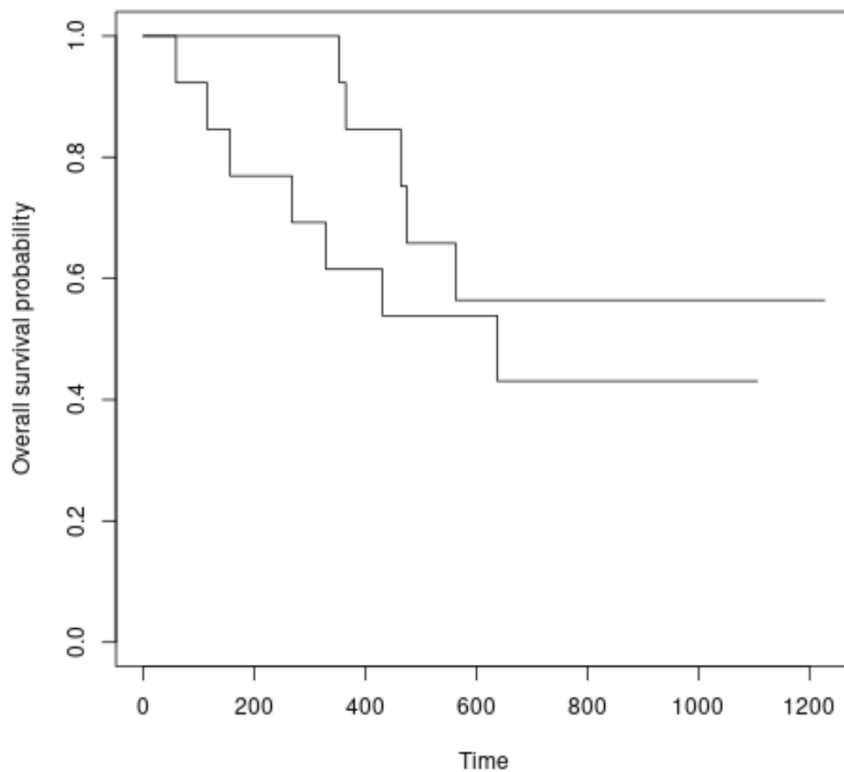
Visual Validation with Kaplan-Meier Survival Curves

While the **log rank test** provides a definitive quantitative measure, best practice in **survival analysis** mandates that statistical findings be complemented by graphical representation. Plotting the **Kaplan-Meier survival curves** allows researchers to visually inspect the survival experience over time for each group, providing essential context for the calculated p-value. This visualization helps confirm whether the magnitude of any observed separation between the curves aligns with the statistical inference.

To generate these curves, we first fit the survival model using the `survfit()` function, which estimates the survival probability at each time point. This estimated model is then passed to the `plot()` function, instructing R to render the curves, typically showing survival probability on the Y-axis and time on the X-axis. The following code executes the plotting command, facilitating visual confirmation of our statistical findings:

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

The execution of this code renders the visual output, enabling immediate comparison of the two treatment cohorts.



Upon examining the resulting plot, one can observe that the two survival curves exhibit a slight degree of separation over time. However, the curves remain relatively close throughout the observation period, and the visual divergence is clearly insufficient to suggest a major clinical or statistical difference. This visual evidence perfectly corroborates the results of the [log rank test](#): while there is a numerical disparity in outcomes, this separation does not cross the established threshold for statistical significance. Therefore, based on both the quantitative p-value (0.3) and the visual representation of the [Kaplan-Meier curves](#), the conclusion remains robust: we cannot confidently assert that one treatment regimen provides a statistically superior survival benefit over the other within this specific patient cohort.

Conclusion: Synthesizing Statistical and Visual Evidence

The process of comparing survival outcomes across groups requires a rigorous blend of statistical testing and graphical confirmation. The [log rank test](#) provides the quantitative foundation for this comparison, offering a non-parametric assessment of the equality of underlying survival distributions. By leveraging the `survdiff()` function in R, researchers can efficiently calculate the [Chi-Squared test statistic](#) and the associated [p-value](#), enabling clear decision-making regarding the [null hypothesis](#).

In our analysis of the `ovarian` dataset, the low Chi-Squared value (1.1) and the high p-value (0.3) led to the definitive conclusion that the treatment type (rx=1 vs. rx=2) did not produce a statistically

significant difference in patient survival. This statistical finding was then visually reinforced by the [Kaplan-Meier curves](#), which showed only a modest, non-significant separation between the two groups. Mastering the execution and interpretation of the log rank test in R is an indispensable skill for anyone conducting rigorous time-to-event research and interpreting the results of clinical and epidemiological studies.