

Calculating P-Values from Z-Scores with R: A Step-by-Step Guide

Authored by
Mohammed loot

November 7, 2025

RECOMMENDED CITATION

Mohammed loot (2025). *Calculating P-Values from Z-Scores with R: A Step-by-Step Guide*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=12549>

The Foundational Role of P-Values and Z-Scores in Statistical Inference

In the rigorous discipline of [statistical hypothesis testing](#), the relationship between the [Z-score](#) and the corresponding [P-value](#) is absolutely central. The **Z-score** serves as the standardized test statistic, quantifying the precise distance, measured in standard deviations, between an observed data point or sample mean and the population mean. This standardization assumes that the underlying data follows a [normal distribution](#). By converting raw data into a standardized Z-score, statisticians can compare results derived from entirely different scales and distributions, preparing the data for the crucial probabilistic assessment that follows.

The [P-value](#) is the probability measure derived from this Z-score, connecting the empirical evidence back to the theoretical distribution. Specifically, it represents the probability of observing a test statistic as extreme as, or more extreme than, the one calculated, assuming that the **null hypothesis** is true. This probability is the core metric used in statistical decision-making. If this P-value is very small, it suggests the observed result is unlikely under the null hypothesis, providing strong evidence to reject it.

The primary objective in this analytical chain is to compare the calculated [P-value](#) against a predetermined threshold known as the **significance level** (α). Typically set at 0.05 or 0.01, the significance level defines the tolerance for incorrectly rejecting a true null hypothesis. If the P-value is less than or equal to α , we declare the result **statistically significant** and reject the null hypothesis. Given the demand for precision and efficiency in modern data analysis, computational tools such as the [R](#) programming environment are indispensable for accurately deriving these probabilities, moving past the limitations of manual Z-table lookups.

Leveraging the `pnorm()` Function in R for Probability Calculation

The most efficient and reliable method for determining the [P-value](#) associated with a given [Z-score](#) in [R](#) is through the built-in function, [pnorm\(\)](#). This function is designed to calculate the **Cumulative Distribution Function (CDF)** for the [normal distribution](#). In simple terms, [pnorm\(\)](#) calculates the area under the standard normal curve up to a specified value (the Z-score), which is the fundamental definition of a cumulative probability. Mastering the correct application of this function is crucial for accurately translating a test statistic into a statistically meaningful probability.

When working with standardized [Z-scores](#), the [pnorm\(\)](#) function simplifies considerably. Since Z-scores adhere to the standard [normal distribution](#) (where the mean is 0 and the standard deviation is 1), we typically rely on the default settings for these two parameters. The critical element that requires explicit manipulation based on the type of [hypothesis test](#) being performed is the `lower.tail` argument. This argument dictates the direction of the probability calculation--whether we are calculating the area to the left or the area to the right of the Z-score.

The core syntax of the function is defined as follows:

`pnorm(q, mean = 0, sd = 1, lower.tail = TRUE)`

Understanding the function's parameters is paramount for accurate results in statistical computation:

q: This represents the quantile of interest, which is the calculated **Z-score** derived from the sample data.

mean: The mean (μ) of the distribution. For standard Z-scores, the default of **0** must be maintained.

sd: The standard deviation (σ) of the distribution. For standard Z-scores, the default of **1** must be maintained.

lower.tail: This boolean parameter determines the calculation direction. Setting it to **TRUE** (the default) computes the cumulative probability (area) from negative infinity up to **q** (the left tail). Setting it to **FALSE** computes the area from **q** up to positive infinity (the right tail). This is the key control for determining P-values in one-tailed tests.

Case Study 1: Calculating P-Values for a Left-Tailed Test

A left-tailed [hypothesis test](#) is established when the alternative hypothesis proposes that the population parameter is smaller than the hypothesized value. In this scenario, the critical region for rejecting the null hypothesis lies exclusively in the extreme left tail of the standard [normal distribution](#). Consequently, the P-value required is the probability of observing a Z-score as low as, or lower than, the calculated test statistic.

Because the P-value for a left-tailed test involves calculating the area to the left of a (typically negative) [Z-score](#), this application directly aligns with the default behavior of the [pnorm\(\)](#) function. Even though setting `lower.tail = TRUE` is redundant since it is the default, it is often good practice to include it explicitly for clarity and robustness in the code.

Let us analyze a statistical outcome that results in a [Z-score](#) of **-0.77**. To find the P-value for this left-tailed test, we use [R](#) to compute the cumulative probability up to -0.77:

#find p-value for a left-tailed test where Z = -0.77

`pnorm(q=-0.77, lower.tail=TRUE)`

0.2206499

The resulting [P-value](#) is approximately **0.2206**. When this is compared against a conventional significance level of $\alpha = 0.05$, the P-value is clearly much larger ($0.2206 > 0.05$). This high

probability suggests that observing a Z-score this low or lower is not a sufficiently rare occurrence to reject the null hypothesis. Therefore, we would conclude that there is insufficient statistical evidence to support the claim that the true population parameter is less than the hypothesized value.

Case Study 2: Calculating P-Values for a Right-Tailed Test

For a right-tailed [hypothesis test](#), the alternative hypothesis posits that the true population parameter is greater than the hypothesized value. Consequently, the region of interest--the critical rejection region--is located entirely within the extreme right tail of the [normal distribution](#). When calculating the P-value for a (typically positive) [Z-score](#) in this context, we must determine the area to the right of the test statistic, which represents the probability of observing a value as high as, or higher than, the calculated Z-score.

To correctly perform this calculation using the [pnorm\(\)](#) function in [R](#), we must explicitly set the `lower.tail` argument to `FALSE`. This command instructs the function to calculate the area from the specified quantile (the Z-score) up to positive infinity, thereby yielding the precise P-value necessary for a right-tailed test. Failing to set this argument to `FALSE` would incorrectly return the area of the left tail, leading to a flawed statistical conclusion.

Consider a scenario where the standardized test statistic is a positive [Z-score](#) of **1.87**. We input this value into [R](#), ensuring the tail parameter is correctly set:

```
#find p-value for a right-tailed test where Z = 1.87
```

```
pnorm(q=1.87, lower.tail=FALSE)
```

```
0.03074191
```

The resulting [P-value](#) is calculated as **0.0307**. Comparing this P-value to the standard $\alpha = 0.05$ significance level, we find that 0.0307 is less than 0.05. This small probability signifies that obtaining a Z-score of 1.87 or larger is a statistically rare event if the null hypothesis were true. Based on this compelling statistical evidence, we would formally reject the null hypothesis, concluding that the data strongly supports the alternative hypothesis that the true parameter is indeed greater than the hypothesized value.

Case Study 3: Calculating P-Values for a Two-Tailed Test

A two-tailed [hypothesis test](#) is required when the alternative hypothesis merely states that the true population parameter is different from the hypothesized value, without specifying whether it is greater or smaller. This approach necessitates dividing the region of rejection equally between both extremes of the standard [normal distribution](#). Consequently, the P-value must account for the

probability of observing a result as extreme as the calculated test statistic in either the positive or the negative direction.

To accurately compute the two-tailed [P-value](#), we first calculate the area (probability) corresponding to the tail defined by the sign of the observed [Z-score](#). Subsequently, we must multiply this single-tail probability by two. This multiplication is a vital step, as it symmetrically accounts for the probability residing in the opposite, unobserved tail. A common practice is to always calculate the probability of the right tail using the absolute value of the Z-score and then multiply by two, simplifying the logic flow in the code.

Imagine we obtain a [Z-score](#) of **1.24** in a two-tailed test. We must determine the area to the right of $Z = 1.24$ (using `lower.tail = FALSE`) and then double this result to capture both tails:

```
#find p-value for two-tailed test where Z = 1.24
```

```
2*pnorm(q=1.24, lower.tail=FALSE)
```

```
0.2149754
```

The resulting two-tailed [P-value](#) is **0.2149**. When interpreting this outcome using the standard $\alpha = 0.05$ significance level, we observe that 0.2149 is significantly greater than 0.05. This relatively high probability indicates that observing a deviation as large as 1.24 standard deviations (in either direction) is a common event under the assumption that the null hypothesis is true. Consequently, we would fail to reject the null hypothesis, concluding that there is insufficient evidence to claim a statistically significant difference from the hypothesized population parameter.

Conclusion: Interpreting P-Values and Making Statistical Decisions

The successful calculation of the [P-value](#) using the `pnorm()` function within [R](#) marks the culmination of the analytical phase in a standardized statistical test. However, the true value of this calculation lies in its accurate interpretation relative to the chosen **significance level** (α). This comparison dictates the final statistical conclusion regarding the null hypothesis.

The fundamental rule for statistical decision-making remains simple and absolute:

If the P-value $\leq \alpha$, we **reject the null hypothesis**. This crucial decision implies that the observed data is highly improbable if the null hypothesis were true, suggesting that the observed effect is statistically significant.

If the P-value $> \alpha$, we **fail to reject the null hypothesis**. This signifies that the observed variation is likely due to random sampling chance, and there is inadequate statistical evidence to support the alternative hypothesis.

By consistently applying the appropriate configuration of the [pnorm\(\)](#) function--paying close attention to the definition of the Z-score (q) and the directional control provided by the `lower.tail` argument--analysts can confidently translate their test statistics into robust, defensible statistical conclusions. This rigorous application ensures that the inferences drawn from data analysis using [R](#) are both accurate and aligned with established statistical principles.

Related: You can also use this online [Z-score to P-value calculator](#) to find p-values.