

How can a Two Proportion Z-Test be performed in SAS?

Authored by
stats writer

June 23, 2024

RECOMMENDED CITATION

stats writer (2024). *How can a Two Proportion Z-Test be performed in SAS?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=148367>

A Two Proportion Z-Test is a statistical method used to compare the proportions of two independent populations. This test can be performed in SAS by using the PROC PROB procedure. First, the data must be organized in a table format with the two populations listed in separate columns. Then, the PROC PROB statement should be followed by the TEST statement, specifying the two populations to be compared. The output will include the test statistic, confidence interval, and p-value, which can be used to determine the significance of the difference between the two proportions. Additionally, SAS provides options for adjusting for continuity correction and specifying the alternative hypothesis. This allows for a comprehensive and accurate analysis of the two proportions using SAS.

Perform a Two Proportion Z-Test in SAS

A two proportion z-test is used to determine if there is a statistically significant difference between two population proportions.

This test uses the following null hypothesis:

$H_0: \mu_1 = \mu_2$ (the two population proportions are equal)

The alternative hypothesis can be either two-tailed, left-tailed, or right-tailed:

H_1 (two-tailed): $\pi_1 \neq \pi_2$ (the two population proportions are not equal)

H_1 (left-tailed): $\pi_1 < \pi_2$ (population 1 proportion is less than population 2 proportion)

H_1 (right-tailed): $\pi_1 > \pi_2$ (population 1 proportion is greater than population 2 proportion)

We use the following formula to calculate the test statistic z:

$$z = (p_1 - p_2) / \sqrt{p(1-p)(1/n_1 + 1/n_2)}$$

where p_1 and p_2 are the sample proportions, n_1 and n_2 are the sample sizes, and where p is the total pooled proportion calculated as:

$$p = (p_1 n_1 + p_2 n_2) / (n_1 + n_2)$$

If the p-value that corresponds to the test statistic z is less than your chosen significance level (common choices are 0.10, 0.05, and 0.01) then you can reject the null hypothesis.

The following example shows how to perform a two proportion z-test in SAS.

Example: Two Proportion Z-Test in SAS

Suppose we want to know if there is a difference in the proportion of residents who support a certain law in county A compared to the proportion who support the law in county B.

To test this, we collect a random sample of 50 residents

from each county and count how many support the law.

The following code shows how to create a dataset that summarizes the number of residents that support the law from each county:

```
/*create dataset*/  
data my_data;  
input county $ status $ count;  
datalines;  
A Support 34  
A Reject 16  
B Support 29  
B Reject 21  
;  
run;  
  
/*view dataset*/  
proc printdata=my_data;
```

| Obs | county | status | count |
|-----|--------|---------|-------|
| 1 | A | Support | 34 |
| 2 | A | Reject | 16 |
| 3 | B | Support | 29 |
| 4 | B | Reject | 21 |

```
/*perform two proportion z-test*/  
proc freqdata=my_data;  
weight count;  
tables county * status / riskdiff(equal var = null);  
run;
```

ARABPSYCHOLOGY.COM

The FREQ Procedure

| Frequency Percent Row Pct Col Pct | Table of county by status | | | |
|--|---------------------------|--------|---------|-------|
| | county | status | | |
| | | Reject | Support | Total |
| | A | 16 | 34 | 50 |
| | | 16.00 | 34.00 | 50.00 |
| | | 32.00 | 68.00 | |
| | | 43.24 | 53.97 | |
| | B | 21 | 29 | 50 |
| | | 21.00 | 29.00 | 50.00 |
| | | 42.00 | 58.00 | |
| 56.76 | | 46.03 | | |
| Total | 37 | 63 | 100 | |
| | 37.00 | 63.00 | 100.00 | |

Statistics for Table of county by status

| Column 1 Risk Estimates | | | | | | |
|-------------------------------|---------|--------|-----------------------|--------|-----------------------------|--------|
| | Risk | ASE | 95% Confidence Limits | | Exact 95% Confidence Limits | |
| Row 1 | 0.3200 | 0.0660 | 0.1907 | 0.4493 | 0.1952 | 0.4670 |
| Row 2 | 0.4200 | 0.0698 | 0.2832 | 0.5568 | 0.2819 | 0.5679 |
| Total | 0.3700 | 0.0483 | 0.2754 | 0.4646 | 0.2756 | 0.4724 |
| Difference | -0.1000 | 0.0960 | -0.2882 | 0.0882 | | |
| Difference is (Row 1 - Row 2) | | | | | | |

| Risk Difference Test | |
|-----------------------------|---------|
| H0: P1 - P2 = 0 Wald Method | |
| Risk Difference | -0.1000 |
| ASE (H0) | 0.0966 |
| Z | -1.0356 |
| One-sided Pr < Z | 0.1502 |
| Two-sided Pr > Z | 0.3004 |
| Column 1 (status = Reject) | |

From the Risk Difference Test table in the output we can see the following information:

z-test statistic: -1.0356 two-sided p-value: 0.3004

This particular two proportion z-test used the following hypotheses:

**H0: $\pi_1 = \pi_2$ (the two population proportions are equal)
H1: $\pi_1 \neq \pi_2$ (the two population proportions are not equal)**

Since the p-value in the output is not less than 0.05, we fail to reject the null hypothesis.

This means we do not have sufficient evidence to say that the proportion of residents who support this law is different between the two counties.

The following tutorials explain how to perform other common statistical tests in SAS: