

# Learning How to Conduct a Two Proportion Z-Test in Excel

Authored by  
**Mohammed looti**

November 9, 2025

## RECOMMENDED CITATION

Mohammed looti (2025). *Learning How to Conduct a Two Proportion Z-Test in Excel*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=14040>

## Understanding the Two Proportion Z-Test

The [Two Proportion Z-Test](#) is an indispensable statistical method used to determine if a significant difference exists between two independent [population proportions](#). This test is vital when researchers need to compare categorical outcomes collected from two distinct groups. Practical applications range widely, from assessing the relative success rates of two alternative marketing campaigns to comparing voter preference rates between two demographic populations.

The core objective of this methodology is to evaluate whether the observed difference between the two sample proportions is large enough to confidently assert that the underlying population proportions are truly unequal. If the difference observed is small, it may simply be attributed to inherent **random sampling variation**. The Z-test provides the formal statistical framework needed to distinguish between these two possibilities. Crucially, the test relies on the fundamental assumption that the sample sizes are sufficiently large, allowing the sampling distribution to be accurately approximated by the standard normal distribution, which permits the use of the Z-statistic.

To grasp this concept, consider a clear scenario: A school superintendent claims that the percentage of students preferring chocolate milk is identical in School 1 and School 2. A determined researcher collects a [simple random sample](#) of 100 students from each school. The survey results indicate that 70% of students in School 1 and 68% in School 2 prefer chocolate milk. The observed difference is 2%. The [Two Proportion Z-Test](#) formalizes the assessment required to determine if this minor 2% sample difference constitutes statistically significant evidence to refute the superintendent's original assertion of equality.

## The Foundational Steps of Hypothesis Testing

Regardless of whether the calculations are performed manually or automated via software like Excel, every robust hypothesis test must follow a standardized, four-step process. Adhering to these steps ensures objectivity, statistical rigor, and a clear path toward a verifiable conclusion.

### Step 1: State the Hypotheses

The first step involves clearly defining the [Null hypothesis](#) ( $H_0$ ) and the Alternative hypothesis ( $H_a$ ). The null hypothesis always represents the baseline assumption, typically stating that there is no effect or that the parameters are equal. In the context of our milk preference example, we are testing for equality:

$H_0$ :  $P_1 = P_2$  (The true population proportion of chocolate milk preference is equivalent in both schools.)

$H_a$ :  $P_1 \neq P_2$  (The proportions are different between the two schools. This configuration

defines a **two-tailed test**.)

If the research question implies a direction (e.g., asserting that School 1's preference rate is higher), the alternative hypothesis would be  $P_1 > P_2$ , which would necessitate a **one-tailed test**.

### Step 2: Determine the **Significance Level** ( $\alpha$ )

The **significance level**, symbolized by  $\alpha$ , is a critical threshold. It quantifies the maximum acceptable probability of committing a Type I error--rejecting the null hypothesis when it is actually true. Standard choices for  $\alpha$  in social and natural sciences are 0.01, 0.05, and 0.10. Throughout this demonstration, we will consistently set the benchmark at  $\alpha = 0.05$ , meaning we are willing to accept a 5% chance of incorrectly rejecting  $H_0$ .

### Step 3: Calculate the **Test Statistic** and **P-value**

This phase is the computational heart of the test. We calculate the Z-statistic, which standardizes the observed difference by measuring how many standard errors the difference between the two sample proportions is from the hypothesized difference (zero). Following the calculation of the Z-statistic, the corresponding probability, or  $p$ -value, is determined using the standard normal distribution.

### Step 4: Make a Decision and Draw a Conclusion

The decision rule is the final gateway to inference: If the calculated  $p$ -value is less than or equal to the predetermined  $\alpha$ , we **reject the null hypothesis**. Conversely, if the  $p$ -value exceeds  $\alpha$ , we **fail to reject the null hypothesis**. The concluding statement must then translate this statistical decision back into clear, non-technical language relevant to the original problem scenario.

## Calculating the Test Statistics Manually

While statistical software provides immediate results, understanding the mechanics of the manual calculation is paramount for accurate interpretation. The calculation of the Z-statistic for two proportions requires two sequential intermediate steps: calculating the pooled sample proportion and then inserting this value into the complex Z-statistic formula.

### The Pooled Sample Proportion ( $\hat{p}$ )

The **pooled sample proportion** is essential because the null hypothesis ( $H_0: P_1 = P_2$ ) assumes that the two population proportions are identical. Therefore, we must use a single, combined estimate for the standard error component of the Z-statistic. This pooled proportion

combines the successes and sample sizes from both groups, yielding the most accurate overall estimate of the common population proportion under the assumption of  $H_0$ .

The formula for the pooled sample proportion ( $\hat{p}$ ) is derived by summing the total number of successes across both samples and dividing by the total number of observations:

$$\hat{p} = \frac{(p_1 \cdot n_1) + (p_2 \cdot n_2)}{n_1 + n_2}$$

Where:

$p_1$  and  $p_2$  represent the observed sample proportions (0.70 and 0.68, respectively).

$n_1$  and  $n_2$  represent the corresponding sample sizes (100 and 100).

Applying the values from our chocolate milk preference example:

$$\hat{p} = \frac{(0.70 \cdot 100) + (0.68 \cdot 100)}{100 + 100} = \frac{70 + 68}{200} = \frac{138}{200} = 0.69$$

## The Z-Test Statistic Calculation

Once the pooled proportion ( $\hat{p}$ ) is secured, we proceed to calculate the Z-statistic. This statistic quantifies the standardized distance between the two sample proportions. If the proportions are truly equal in the population, we expect the Z-statistic to be close to zero. The formula is structured by dividing the observed difference in sample proportions by the pooled standard error:

$$z = \frac{(p_1 - p_2)}{\sqrt{\hat{p} \cdot (1 - \hat{p}) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Substituting the specific values for our scenario (where  $p_1=0.70$ ,  $p_2=0.68$ , and  $\hat{p}=0.69$ ):

$$z = \frac{(0.70 - 0.68)}{\sqrt{0.69 \cdot (1 - 0.69) \cdot \left(\frac{1}{100} + \frac{1}{100}\right)}}$$

$$z = \frac{0.02}{\sqrt{0.69 \cdot 0.31 \cdot 0.02}} = \frac{0.02}{\sqrt{0.004278}} \approx \frac{0.02}{0.0654}$$

The resulting Z-statistic is approximately **0.306**. By referencing a standard Z-distribution table or utilizing statistical software for this Z-score in a two-tailed context, we find the resulting  $p$ -value is **0.759**. Since 0.759 is dramatically larger than our preset [significance level](#) ( $\alpha$ ) of 0.05, we **fail to reject the null hypothesis**. This absence of strong statistical evidence suggests that the observed 2% difference in chocolate milk preference is likely attributable to chance, not a true difference between the schools.

## Performing a Two Proportion Z-Test in Excel

Using Microsoft Excel is the most efficient and accurate way to perform these calculations, especially when dealing with multiple datasets. Excel automates the complex formula steps, allowing the user to focus on inputting the sample data and interpreting the final  $p$ -value. The following sections provide step-by-step instructions for setting up the spreadsheet for both two-tailed and one-tailed tests, ensuring replicable results.

### Implementing the Two-Tailed Z-Test in Excel (Example 1)

This first example directly addresses the Null hypothesis of equality ( $H_0: P_1 = P_2$ ) versus the Alternative hypothesis of difference ( $H_a: P_1 \neq P_2$ ).

**Scenario Review:** We are testing if the 70% preference rate ( $n=100$ ) in School 1 differs significantly from the 68% rate ( $n=100$ ) in School 2, using  $\alpha = 0.05$ .

The image below illustrates the required data inputs and the corresponding Excel formulas necessary to compute the pooled proportion, the [Z-statistic](#), and the final  $p$ -value for this two-tailed procedure:

	A	B	C	D	E	F
1	Sample 1 proportion	0.7				
2	Sample 1 size	100				
3	Sample 2 proportion	0.68				
4	Sample 2 size	100				
5						
6	Pooled sample proportion	0.69	$= (B1*B2 + B3*B4) / (B2+B4)$			
7	Test statistic	0.30578	$= (B1-B3) / \text{SQRT}(B6*(1-B6)*((1/B2)+(1/B4)))$			
8	P-value	0.759772	$= 2*(1-\text{NORM.S.DIST}(B7, \text{TRUE}))$			
9						
10						
11						
12						

To reproduce this statistical analysis, the raw sample data must first be entered into cells **B1 through B4**. The subsequent derived values in cells **B6 through B8** are automatically generated using the mathematical logic defined by the formulas visible in cells **C6 through C8**. This structured spreadsheet design standardizes the entire analysis, enabling rapid and reliable substitution of new sample data for immediate hypothesis testing.

The specific Excel functions applied are detailed below:

**Formula in Cell C6 (Pooled Proportion):**

This formula accurately calculates the [pooled sample proportion](#) ( $\hat{p}$ ), which is critical for estimating the common population variance under  $H_0$ .

$$\hat{p} = \frac{(p_1 \cdot n_1) + (p_2 \cdot n_2)}{n_1 + n_2}$$

**Formula in Cell C7 (Z-Test Statistic):**

This calculation determines the standardized Z-score by dividing the difference in sample proportions by the pooled standard error.

$$z = \frac{(p_1 - p_2)}{\sqrt{\hat{p} \cdot (1 - \hat{p}) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

**Formula in Cell C8 (P-value Calculation for Two-Tailed Test):**

For a two-tailed test, the p-value represents the combined probability in both extreme tails of the distribution. We utilize the [NORM.S.DIST](#) function, which provides the cumulative probability for the standard normal distribution (Z-score). Since the p-value must account for extremity in both the positive and negative directions, the resulting area in one tail ( $1 - \text{NORM.S.DIST}(\text{Z-score}, \text{TRUE})$ ) must be multiplied by two.

**Conclusion for Example 1:** Given that the calculated p-value (**0.759**) is substantially greater than the predetermined [significance level](#) ( $\alpha$ ) of **0.05**, we **fail to reject the null hypothesis**. There is insufficient statistical evidence to conclude that the chocolate milk preference rates truly differ between the two schools.

**Performing a One-Tailed Z-Test in Excel (Example 2)**

A one-tailed test is employed when the research question hypothesizes a specific direction for the difference (either greater than or less than). This directional hypothesis fundamentally alters how the final [P-value](#) is computed, although the pooled proportion and Z-statistic remain unchanged.

**Scenario Review:** We are now testing the superintendent's claim that preference in School 1 is **less than or equal to** that in School 2 ( $H_0: P_1 \leq P_2$ ). Consequently, the Alternative hypothesis becomes  $H_a: P_1 > P_2$ . Sample data remains constant (70% vs. 68% at  $n=100$  each), with  $\alpha = 0.05$ . The key statistical question is: Does the data provide sufficient evidence to support the claim that School 1 has a strictly higher preference rate?

The screenshot below displays the necessary Excel configuration for this one-tailed test, highlighting the modification required for the p-value formula:

	A	B	C	D	E	F	G
1	Sample 1 proportion	0.7					
2	Sample 1 size	100					
3	Sample 2 proportion	0.68					
4	Sample 2 size	100					
5							
6	Pooled sample proportion	0.69	$= (B1*B2 + B3*B4) / (B2+B4)$				
7	Test statistic	0.30578	$= (B1-B3) / \text{SQRT}(B6*(1-B6)*((1/B2)+(1/B4)))$				
8	P-value	0.379886	$= 1 - \text{NORM.S.DIST}(B7, \text{TRUE})$				
9							
10							
11							
12							

Just as in the two-tailed example, the sample inputs are in **B1:B4**, and the resulting statistics are automatically calculated in **B6:B8**. It is important to note that the calculations for the pooled proportion (C6) and the Z-statistic (C7) are identical to the previous test, as they are derived solely from the sample data and the assumption of  $H_0$ .

The primary difference resides in the p-value calculation (Cell C8):

#### Formula in Cell C8 (P-value Calculation for One-Tailed Test):

Since the alternative hypothesis ( $H_a: P_1 > P_2$ ) predicts a positive difference (School 1 has a higher proportion), we are only concerned with the probability located in the upper (right) tail of the standard normal distribution. To find this area, we calculate the cumulative probability up to the Z-score using the [NORM.S.DIST](#) function, and then subtract that result from 1.

$$P\text{-value} = 1 - \text{NORM.S.DIST}(\text{Z-score}, \text{TRUE})$$

**Conclusion for Example 2:** In this directional test, the resulting p-value is **0.379**. Comparing this to the predetermined [significance level](#) of **0.05**, we observe that  $0.379 > 0.05$ . Consequently, we must again **fail to reject the null hypothesis**. The data does not provide adequate statistical support to conclude that the percentage of students preferring chocolate milk in School 1 is strictly greater than the percentage in School 2.

## Interpreting the Results and Drawing Conclusions

The ultimate goal of applying the [Two Proportion Z-Test](#), whether through manual calculation or automated Excel sheets, is to reach a statistically sound, dichotomous decision regarding the [Null hypothesis](#). A thorough understanding of both the Z-statistic and the p-value is crucial for

translating numerical outputs into meaningful conclusions about the population parameters.

The calculated [Z-statistic](#) of 0.306 in our examples provides valuable context: it indicates that the observed difference between the sample proportions (0.02) is only 0.306 standard errors away from zero. This minimal standardized deviation strongly suggests that the difference observed in the sample data is highly likely to be a consequence of inherent **random variation** rather than a genuine, systematic difference in the underlying population preference rates.

The final inference rests entirely on the  $p$ -value. A large  $p$ -value (such as 0.759 or 0.379) signifies that if the null hypothesis were true (i.e., the population proportions were equal), we would frequently observe sample results as extreme as, or even more extreme than, the data we collected. Conversely, a small  $p$ -value (typically below the  $\alpha$  threshold of 0.05) implies that the observed data would be exceptionally rare if  $H_0$  were true, providing compelling statistical evidence to reject the null hypothesis and endorse the alternative claim. In both scenarios presented, the evidence was insufficient to overturn the established assumption of equality. This demonstrates the necessity of utilizing formal statistical inference over relying on mere numerical proximity (70% being slightly greater than 68%) when drawing conclusions about large population parameters.