An Exploratory Analysis of Mathematics Interest by Gender Among Instagram Poll Respondents

Giodanno Limin

Data collector and author

Statistical writing and LaTeX assistance: ChatGPT

June 2026

Abstract

This paper presents an exploratory statistical analysis of an Instagram poll investigating self-reported interest in mathematics by gender. The poll received 170 viewers and 67 total responses. Among male respondents, 22 out of 25 reported liking mathematics, while among female respondents, 29 out of 42 reported liking mathematics. These correspond to sample proportions of 88.0% and 69.0%, respectively.

A one-sided two-proportion z-test was conducted to test whether the male respondent proportion was higher than the female respondent proportion. The test produced a statistic of approximately $z = 1.76$ and a one-sided p-value of approximately 0.039. However, because one cell count in the contingency table was small, Fisher's exact test was also considered. Fisher's exact test produced a more conservative one-sided p-value of approximately 0.069.

Confidence intervals were also calculated for the proportion of respondents who reported liking mathematics within each gender group. Using the normal approximation, the 95% confidence intervals were approximately $[0.753, 1.007]$ for male respondents and $[0.551, 0.830]$ for female respondents. Since the male interval exceeds 1, this illustrates a limitation of the simple normal approximation for small samples. Using the Wilson method, the 95% confidence intervals were approximately $[0.700, 0.958]$ for male respondents and $[0.540, 0.809]$ for female respondents.

Overall, the results are suggestive but not conclusive. Due to the voluntary, nonrandom nature of the poll and the substantial nonresponse rate, the findings should

be interpreted as descriptive of the respondents only and should not be generalized to Indonesians or to the general population.

1 Introduction

Interest in mathematics is often discussed in relation to gender, education, culture, and social environment. However, claims about gender differences in mathematics interest require careful statistical interpretation. Informal online polls can provide useful exploratory information, but they do not have the same strength as carefully designed random surveys.

This paper analyzes an Instagram poll that asked viewers whether they like mathematics, with response options separated by self-reported gender. The purpose of the analysis is to examine whether, among respondents to this specific poll, male respondents were more likely than female respondents to report liking mathematics.

The research question is:

> Among respondents to this Instagram poll, is the proportion of male respondents who reported liking mathematics higher than the proportion of female respondents who reported liking mathematics?

This study is exploratory. The aim is not to make a population-level claim about Indonesians or about males and females in general. Instead, the aim is to apply formal statistical reasoning to a small, self-selected dataset.

2 Data and Methods

2.1 Data Collection

The data were collected through an Instagram poll. The poll asked:

> Do you like maths?

The four response options were:

1. Yes, I am male
2. Yes, I am female
3. No, I am male

4. No, I am female

The poll received 170 story viewers and 67 total responses. The observed responses are shown in Table 1.

Table 1: Observed Instagram poll responses

Gender	Likes mathematics	Does not like mathematics	Total
Male	22	3	25
Female	29	13	42
Total	51	16	67

The response rate was

$$\frac{67}{170} \approx 0.394.$$

Thus, approximately 39.4% of viewers responded. The number of non-respondents was

$$170 - 67 = 103.$$

Therefore, approximately 60.6% of viewers did not respond. The gender and mathematics interest of these non-respondents are unknown.

2.2 Notation

Let group 1 represent male respondents and group 2 represent female respondents.

$$n_1 = 25$$

$$n_2 = 42$$

where n_1 is the number of male respondents and n_2 is the number of female respondents. Let

$$x_1 = 22$$

$$x_2 = 29$$

where x_1 is the number of male respondents who answered "Yes," and x_2 is the number of female respondents who answered "Yes."

The sample proportion of male respondents who reported liking mathematics is

$$\hat{p}_1 = \frac{x_1}{n_1}.$$

The sample proportion of female respondents who reported liking mathematics is

$$\hat{p}_2 = \frac{x_2}{n_2}.$$

Substituting the observed values gives

$$\hat{p}_1 = \frac{22}{25} = 0.880$$

and

$$\hat{p}_2 = \frac{29}{42} \approx 0.690.$$

The observed difference in sample proportions is

$$\hat{p}_1 - \hat{p}_2 = 0.880 - 0.690 \approx 0.190.$$

Thus, the observed proportion of respondents who reported liking mathematics was approximately 19.0 percentage points higher among male respondents than among female respondents.

2.3 Hypotheses

Let

$$p_1 = \text{the true proportion of male poll respondents who like mathematics}$$

and

$$p_2 = \text{the true proportion of female poll respondents who like mathematics.}$$

The null hypothesis is

$$H_0 : p_1 = p_2.$$

The alternative hypothesis is

$$H_a : p_1 > p_2.$$

The null hypothesis states that male and female poll respondents have the same underlying probability of reporting that they like mathematics. The alternative hypothesis states that male poll respondents have a higher probability of reporting that they like mathematics.

Since the alternative hypothesis uses >, the analysis uses a one-sided test.

2.4 Two-Proportion z-Test

A two-proportion z-test was used to compare the two sample proportions. Under the null hypothesis, the two population proportions are assumed to be equal. Therefore, the common proportion is estimated using the pooled sample proportion:

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}.$$

Substituting the data gives

$$\hat{p} = \frac{22 + 29}{25 + 42} = \frac{51}{67} \approx 0.761.$$

The standard error under the null hypothesis is

$$SE = \sqrt{\hat{p}(1 - \hat{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}.$$

Substituting the values gives

$$SE = \sqrt{0.761(1 - 0.761) \left(\frac{1}{25} + \frac{1}{42} \right)} \approx 0.108.$$

The test statistic is

$$z = \frac{\hat{p}_1 - \hat{p}_2}{SE}.$$

Therefore,

$$z = \frac{0.880 - 0.690}{0.108} \approx 1.76.$$

The one-sided p-value is

$$P(Z \geq 1.76) \approx 0.039.$$

Using the two-proportion z-test, the result is statistically significant at the 5% level because

$$0.039 < 0.05.$$

2.5 Small Cell Count and Fisher's Exact Test

The contingency table contains one small cell count. Specifically, only 3 male respondents answered "No." In a contingency table, each entry is called a cell. Therefore, the value 3 in the male "No" category is a small cell count.

This matters because the two-proportion z-test relies on a normal approximation. Normal approximations tend to be more reliable when sample sizes are large and when the number of observations in each category is not too small. When a cell count is small, the approximation may be less accurate.

For this reason, Fisher's exact test was also considered. Fisher's exact test is commonly used for small contingency tables because it does not rely on a normal approximation. Instead, it calculates the probability of observing a table as extreme as, or more extreme than, the observed table under the null hypothesis, conditional on the fixed row and column totals.

For the one-sided alternative

$$H_a : p_1 > p_2,$$

Fisher's exact test gives approximately

$$p \approx 0.069.$$

Since

$$0.069 > 0.05,$$

Fisher's exact test is not statistically significant at the 5% level.

2.6 Confidence Intervals for Individual Proportions

In addition to hypothesis testing, confidence intervals were calculated for the proportion of respondents who reported liking mathematics within each gender group.

A confidence interval gives a range of plausible values for a population proportion, based on the observed sample. In this study, the intervals should be interpreted only for the respondent population of this Instagram poll, not for Indonesians or the general population.

2.6.1 Normal Approximation Confidence Interval

The normal approximation confidence interval for a single proportion is

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.$$

For a 95% confidence interval,

$$z_{\alpha/2} = 1.96.$$

For male respondents,

$$\hat{p}_1 = \frac{22}{25} = 0.880.$$

The standard error is

$$SE_1 = \sqrt{\frac{0.880(1-0.880)}{25}} \approx 0.065.$$

Therefore, the 95% normal approximation confidence interval is

$$0.880 \pm 1.96(0.065).$$

This gives

$$[0.753, 1.007].$$

As percentages, this is approximately

$$[75.3\%, 100.7\%].$$

Since a proportion cannot be greater than 100%, the upper endpoint above 100% illustrates a weakness of the normal approximation method in this small-sample setting.

For female respondents,

$$\hat{p}_2 = \frac{29}{42} \approx 0.690.$$

The standard error is

$$SE_2 = \sqrt{\frac{0.690(1-0.690)}{42}} \approx 0.071.$$

Therefore, the 95% normal approximation confidence interval is

$$0.690 \pm 1.96(0.071).$$

This gives

$$[0.551, 0.830].$$

As percentages, this is approximately

$$[55.1\%, 83.0\%].$$

2.6.2 Wilson Confidence Interval

Because the sample sizes are relatively small and one observed proportion is close to 1, the Wilson confidence interval is also considered. The Wilson interval is often more reliable than the simple normal approximation interval for small samples or when the sample proportion is near 0 or 1.

For a sample proportion $\hat{p}$, sample size n, and critical value $z_{\alpha/2}$, the Wilson interval is

$$\frac{\hat{p} + \frac{z_{\alpha/2}^2}{2n} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{1 + \frac{z_{\alpha/2}^2}{n}}.$$

Using $z_{\alpha/2} = 1.96$, the 95% Wilson confidence interval for male respondents is approximately

$$[0.700, 0.958].$$

As percentages, this is approximately

$$[70.0\%, 95.8\%].$$

For female respondents, the 95% Wilson confidence interval is approximately

$$[0.540, 0.809].$$

As percentages, this is approximately

$$[54.0\%, 80.9\%].$$

The Wilson intervals stay within the valid range from 0 to 1 and are therefore more appropriate for this small-sample setting.

3 Results

3.1 Descriptive Results

Among male respondents,

$$\frac{22}{25} = 0.880,$$

so 88.0% reported liking mathematics.

Among female respondents,

$$\frac{29}{42} \approx 0.690,$$

so approximately 69.0% reported liking mathematics.

The observed difference was

$$0.880 - 0.690 \approx 0.190.$$

Therefore, male respondents had an observed mathematics interest rate approximately 19.0 percentage points higher than female respondents.

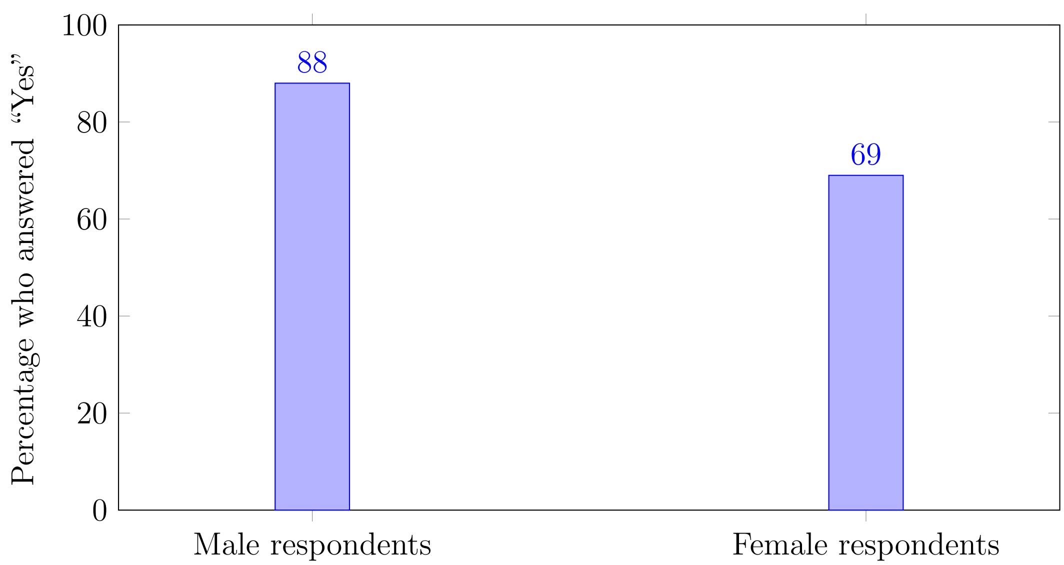

Figure 1: Percentage of respondents who reported liking mathematics by gender.

3.2 Confidence Interval Results

The confidence interval results are summarized in Table 2.

Table 2: 95% confidence intervals for the proportion who reported liking mathematics

Group	Estimate	Normal approximation CI	Wilson CI
Male respondents	0.880	$[0.753, 1.007]$	$[0.700, 0.958]$
Female respondents	0.690	$[0.551, 0.830]$	$[0.540, 0.809]$

The normal approximation confidence interval for male respondents has an upper endpoint above 1. Since proportions cannot exceed 1, this shows that the normal approximation interval is not ideal for this sample. The Wilson interval is more appropriate because it remains within the valid range and performs better for small samples.

The Wilson confidence intervals show that the estimated male proportion is higher than the estimated female proportion, but the intervals overlap. This overlap indicates that there is still substantial uncertainty in the estimates.

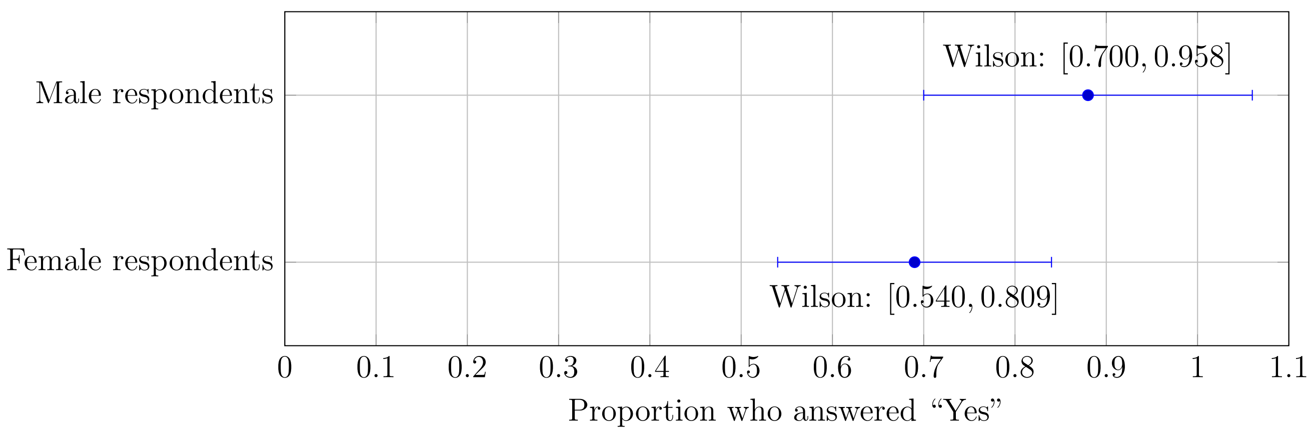

Figure 2: Wilson 95% confidence intervals for the proportion who reported liking mathematics.

3.3 Hypothesis Test Results

The one-sided two-proportion z-test produced

$$z \approx 1.76$$

with one-sided

$$p \approx 0.039.$$

This suggests statistical significance at the 5% level under the assumptions of the two-proportion z-test.

However, Fisher's exact test produced a more conservative one-sided p-value:

$$p \approx 0.069.$$

This is not statistically significant at the 5% level.

The results of the two tests are summarized in Table 3.

Table 3: Comparison of statistical test results

Test	One-sided p-value	Significant at 5% level?
Two-proportion z-test	≈ 0.039	Yes
Fisher's exact test	≈ 0.069	No

Figure 3 gives a conceptual visualization of the one-sided z-test. The shaded area represents the right-tail probability corresponding to the p-value.

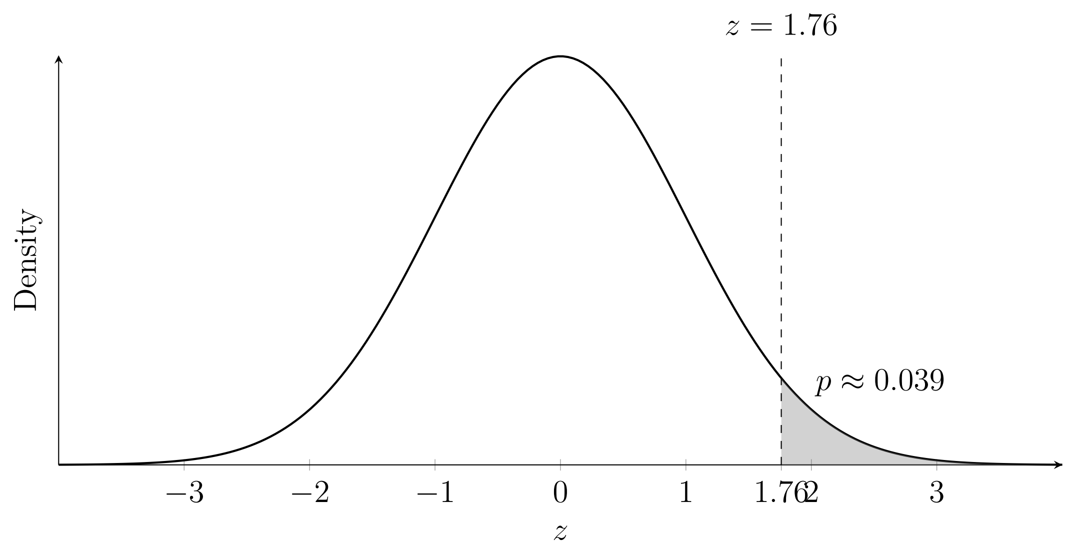

Figure 3: Conceptual visualization of the one-sided two-proportion z-test. The shaded right-tail area represents the approximate p-value.

4 Discussion

The observed data show that male respondents had a higher reported rate of mathematics interest than female respondents. Specifically, 88.0% of male respondents answered "Yes," compared with approximately 69.0% of female respondents.

The confidence intervals provide additional context. The Wilson 95% confidence interval for male respondents was approximately $[0.700, 0.958]$, while the Wilson 95% confidence interval for female respondents was approximately $[0.540, 0.809]$. These intervals are fairly wide because the sample sizes are small. They also overlap, which suggests that the difference should be interpreted cautiously.

The two-proportion z-test suggests that this difference is statistically significant at the 5% level. However, the evidence becomes weaker when Fisher's exact test is used. Since Fisher's exact test is more appropriate for small contingency tables, especially when one cell count is low, the result should be interpreted cautiously.

The difference between the two tests is important. The z-test relies on an approximation, while Fisher's exact test calculates probabilities exactly under the fixed-margin assumption. Because the male "No" cell contains only 3 observations, Fisher's exact test provides a useful check against overinterpreting the z-test result.

Therefore, the most appropriate interpretation is that the poll provides suggestive evidence of a higher observed mathematics interest rate among male respondents, but the evidence is not strong enough to support a firm conclusion.

5 Interpretation at the 10% Significance Level

The main analysis considered the 5% significance level. However, because this study is exploratory, it is also useful to consider the 10% significance level.

Let

$$\alpha = 0.10.$$

The decision rule is

$$\text{Reject } H_0 \quad \text{if} \quad p\text{-value} \le \alpha.$$

Equivalently,

$$\text{Reject } H_0 \quad \text{if} \quad p\text{-value} \le 0.10.$$

For the two-proportion z-test, the one-sided p-value was approximately

$$p_{\text{z-test}} \approx 0.039.$$

Since

$$0.039 < 0.10,$$

the two-proportion z-test rejects H_0 at the 10% significance level.

For Fisher's exact test, the one-sided p-value was approximately

$$p_{\text{Fisher}} \approx 0.069.$$

Since

$$0.069 < 0.10,$$

Fisher's exact test also rejects H_0 at the 10% significance level.

Therefore, both tests reject the null hypothesis at the 10% significance level. This provides suggestive evidence that the proportion of male respondents who reported liking mathematics was higher than the proportion of female respondents who reported liking mathematics.

The rigorous meaning of the 10% significance level is

$$P(\text{Reject } H_0 \mid H_0 \text{ is true}) \leq 0.10.$$

This probability is called the Type I error rate. It means that if the null hypothesis were actually true, then the testing procedure would falsely reject the null hypothesis at most 10% of the time in repeated sampling.

However, this does not mean that there is a 10% probability that the conclusion from this specific poll is wrong. In frequentist hypothesis testing, the significance level controls the long-run false rejection rate of the procedure, not the probability that the null hypothesis is true after observing the data.

Thus, rejecting H_0 at the 10% level should be interpreted as weaker evidence than rejecting at the 5% level. In this study, the result should be described as suggestive evidence rather than strong evidence.

6 Limitations

This analysis has several limitations.

First, the poll was voluntary. Viewers chose whether or not to respond, so the respondents may differ systematically from the non-respondents. This creates potential nonresponse bias.

Second, the sample was not randomly selected. The respondents were Instagram viewers of the author's story, not a random sample of Indonesians, students, or the general population.

Third, the response rate was only approximately 39.4%. Since 103 out of 170 viewers did not vote, the gender and mathematics interest of most viewers are unknown.

Fourth, the poll measured mathematics interest using a simple yes/no question. This does not capture the intensity of interest, educational background, mathematical confidence, school experience, major, age, or other relevant factors.

Fifth, the poll used only two gender categories. This simplifies the analysis but does not represent all possible gender identities.

Finally, the analysis is exploratory. If the research question was formed after observing the results, the hypothesis test should be interpreted more cautiously than if the hypothesis had been specified before data collection.

7 External Validity

The findings should not be generalized to Indonesians or to the general population. Even if the viewers were mostly Indonesian, the respondents were not randomly sampled from the Indonesian population. They were people who viewed a specific Instagram story and chose to respond.

Therefore, the appropriate target of interpretation is limited to the respondents of this specific Instagram poll. The results may provide insight into the responding portion of the author's Instagram audience, but they should not be treated as evidence of a general gender difference in mathematics interest.

8 Conclusion

This paper analyzed an Instagram poll about mathematics interest by gender. Among the 67 respondents, male respondents had a higher observed rate of reporting that they like mathematics than female respondents. The observed proportions were 88.0% for male respondents and approximately 69.0% for female respondents, giving an observed difference of approximately 19.0 percentage points.

Using the normal approximation, the 95% confidence interval for the male respondent proportion was approximately $[0.753, 1.007]$, while the corresponding interval for female respondents was approximately $[0.551, 0.830]$. Since the male interval exceeds 1, the normal approximation is not ideal in this setting. Using the Wilson method, the 95% confidence intervals were approximately $[0.700, 0.958]$ for male respondents and $[0.540, 0.809]$ for female respondents.

A one-sided two-proportion z-test gave a p-value of approximately 0.039, suggesting statistical significance at the 5% level. However, Fisher's exact test gave a more conservative one-sided p-value of approximately 0.069, which is not significant at the 5% level.

Because one cell count was small, the sample was voluntary and non-random, and a majority of viewers did not respond, the findings should be interpreted cautiously. The most appropriate conclusion is that male respondents showed a higher observed rate of mathematics interest in this specific Instagram poll, but the evidence is suggestive rather than conclusive and should not be generalized to a broader population.

AI Assistance Disclosure

The data were collected by Giodanno Limin through an Instagram poll. ChatGPT was used to assist with statistical explanation, report structure, wording, and LaTeX formatting. The author is responsible for the final interpretation and use of the report.