Performance Analysis of Ansari-Bradley and Flinger-Killeen Tests Under Varying Distribution Types and Ratios

November 15th, 2024

Muskaan Mahes, Sarah Khoja,
Tahira Quintos, Jameelah Ali

Introduction

This simulation aims to compare Type I Error and Power rates across 12 different combinations of tests, distributions, and dispersion ratios. Our group ran 12 tests of the Ansari-Bradley and Fligner-Killeen tests with sample sizes of (30,30) for two sample means. With each scenario, we tested both normal and exponential distribution and 3 potential power scenarios for each of these combinations (1.0, 1.5, 2.0) to see how these tests change with different distribution and power scenarios, both comparing the A-B to F-K test results and within each test group. From each of these 12 scenarios, we can discern trends about the effects of each test, distribution, and dispersion ratio has the largest effect on Type I Error and Power rates when we hold the sample size constant for test equality of variances. In comparing Type I Error, Power, and Histograms across these 12 scenarios, the simulation aims to provide insight into best practices for designing non-parametric tests and which factors are influential.

It's important to note that all of the test scenarios for the Ansari-Bradley (AB) and Fligner-Killeen (FK) have the same null hypothesis. The null hypothesis tells us that the two groups have no difference in the variance or rate parameter.

Methods

Sample Sizes: 30 observations per group

Dispersion Ratios: 1, 1.5. 2.0

Number of simulations: 1,000 iterations per scenario

Procedure:

- Generate two groups with specified dispersion ratios
- Apply the AB and FK tests for both normal and exponential distributions
- Record Type I and power values

1. Ansari-Bradley Tests
    a. Normal Distribution

        When applying normally distributed data to the Ansari-Bradley test, we can see a trend between the dispersion ratio and the power. As we can see in Table 1 as the dispersion ratio increases from 1.0 to 2.0, the power of the test also increases, rising from 0.059 to 0.301. This means that the test becomes more effective at detecting true differences between the groups as the variance disparity grows.

As for the Type I error the rate remains relatively stable across the different dispersion ratios, hovering around the nominal significance level of 0.05. This suggests that the test is well-controlled in terms of false positives, even as its power to detect true differences improves.

b. Exponential Distribution

The Ansari-Bradley test, when applied to exponentially distributed data demonstrates an intriguing relationship between dispersion ratio and power. We can see in Table 2 that as the dispersion ratio increases which indicates a greater difference in variance between the two groups, the test's power to detect a true difference also increases. This is evident in the rise of the power values from 0.287 to 0.535 as the dispersion ratio moves from 1.0 to 2.0.

However, it is crucial to note that this increase in power comes at a cost. The Type I error rate, which represents the probability of incorrectly rejecting the null hypothesis when it's actually true, also remains elevated across all three dispersion ratios. This means that the test becomes more sensitive to detecting differences, but it also becomes more prone to false positives as the variance between the groups grows.

2. Fligner-Killeen Tests
   a. Normal Distribution:

With normal distribution applied to the Fligner-Kileen Test, there are clear trends between dispersion ratios and the results. With smaller dispersion ratios, there is subsequently a lower power, and the power improved to 0.13 and 0.3 as the dispersion ratio increased to 1.5 and 2.0. This shows that with the differences between variances widening, the FK test can better reject a false null hypothesis.

Type I errors remained mostly stable as the dispersion ratio increased. This suggests that the FK test can control the rate of false positives and is reliable when it comes to error detection. This robustness in maintaining Type I error rates and better ability to detect differences as variance increases are notable.

b. Exponential Distribution:

For exponential distributions, the FK test assesses robustness under skewed conditions, particularly for higher dispersion ratios. As dispersion ratios the FK test demonstrates increasing power, rising from 0.063 at a dispersion ratio of 1.0 to 0.271 at 2.0. This trend indicates that the FK test becomes more effective at identifying true variance differences as these differences widen.

Type I error rates remained stable across all dispersion tattoos, hovering around 0.094. This stability indicates FK test maintains good control over false positives even when applied to skewed data. While its power remains lower compared to its performance under normal distributions, the FK tests still provide a reliable option for finding variance in exponential data.

For scenarios with unequal dispersions, the FK test demonstrated moderate power.

Results

Type I Error and Power

Ansari-Bradley Tests Normal

Dispersion Ratio	**1.0**	**1.5**	**2.0**
Type I Error	0.047	0.053	0.055
Power	0.059	0.143	0.301

This table summarizes the performance of the AB test for normal data. It demonstrates consistent control over Type I error rates under the null hypothesis and shows increasing power as dispersion ratios deviate from equality.

Ansari-Bradley Tests Exponential

Dispersion Ratio	**1.0**	**1.5**	**2.0**
Type I Error	0.274	0.297	0.292
Power	0.287	0.41	0.535

This table summarizes the performance of the AB test for normal data. It demonstrates consistent control over Type I error rates under the null hypothesis and shows increasing power as dispersion ratios deviate from equality.

Fligner-Killeen Tests Normal -

Dispersion Ratio	**1.0**	**1.5**	**2.0**
Type I Error	0.031	0.044	0.031
Power	0.036	0.13	0.354

The table above shows the performance of the Fligner-Killeen Tests with normal distribution for dispersion ratios of 1.0, 1.5, and 2.0. The Type I errors are under the null hypothesis and the powers are for each dispersion ratio. We can see a steady increase in Power and a stable Type I Error values

Fligner-Killeen Tests Exponential

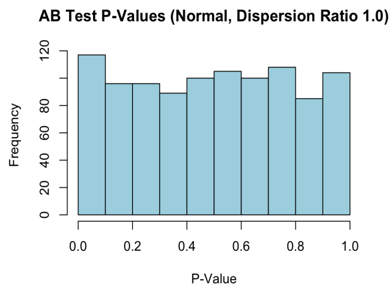
Dispersion Ratio	**1.0**	**1.5**	**2.0**
Type I Error	0.094	0.088	0.094
Power	0.063	0.157	0.271

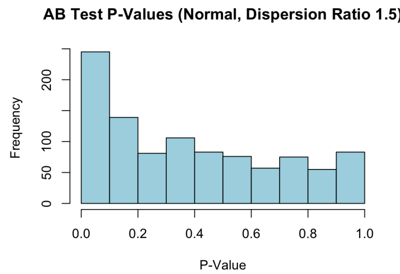
This table summarizes the performance of the FK test for exponential data. Type I error rates are provided for the null hypothesis scenario, and the power values are listed for scenarios with unequal dispersions.

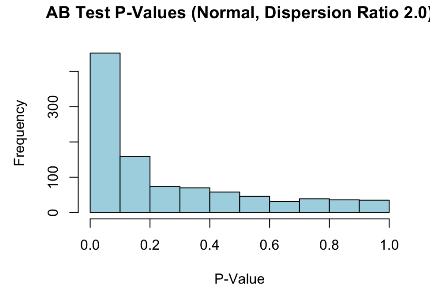
Histogram Analysis

Ansari-Bradley Tests:

The following histograms show the distribution of p-values for the AB test with normal distribution:

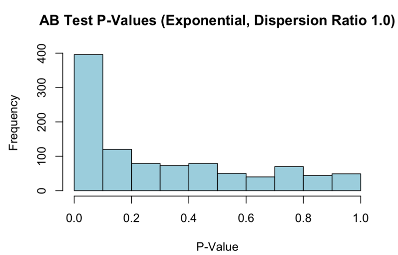


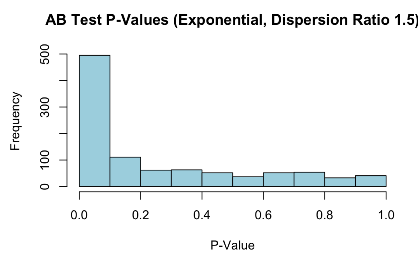


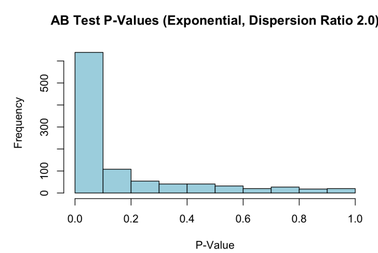


The histograms for the AB test under normal distribution show a well-behaved distribution of p-values, centered around the expected value under the null hypothesis, indicating good Type I error control.

The following histograms show the distribution of p-values for the AB test with normal distribution:



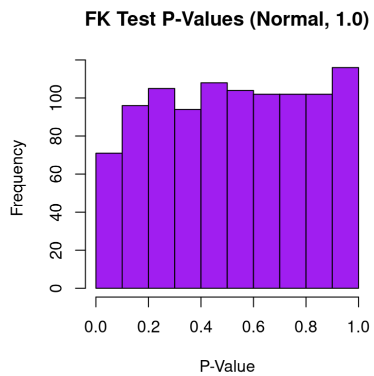


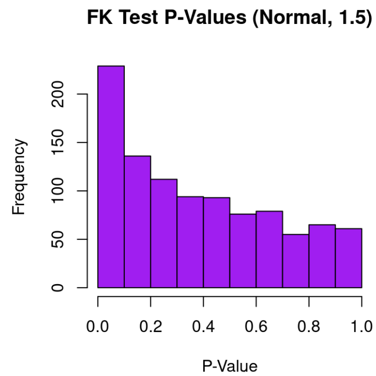


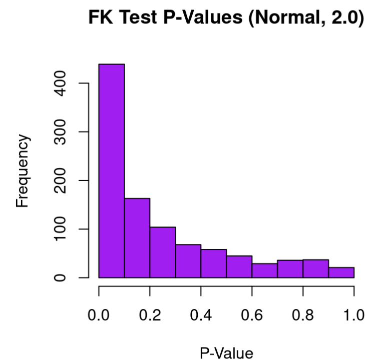
For exponential distributions, the histograms show a shift towards lower p-values, suggesting increased power but also a potential for inflated Type I error rates.

Fligner-Killeen Tests:

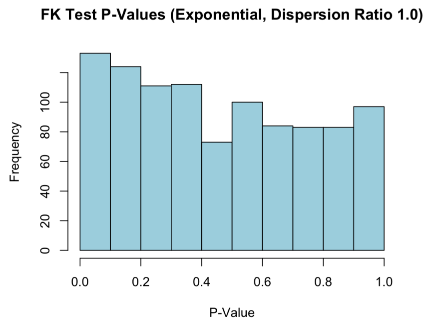
The following histograms show the distribution of p-values for the FK test with normal distribution:

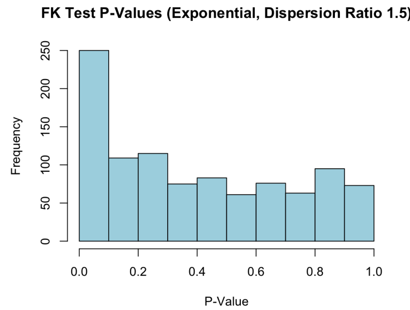


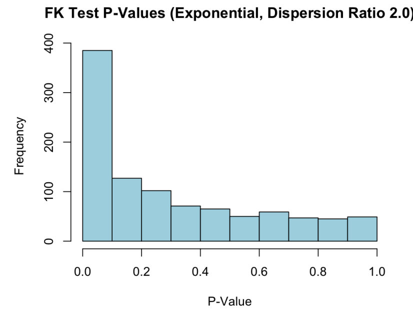




For all of the histograms, the x-axis is a range of p-values from the respective Fligner-Killeen Test on a scale of 0.0 to 1.0. The y-axis is the frequency of those p-values, however, each histogram has a different scale for the frequency values so comparing the actual height of the bar may not be indicative of the frequency. The first histogram is the Fligner-Killeen Test with normal distribution and a dispersion ratio of 1.0. The distribution is relatively uniform across a 0.0 to 1.0 range, as seen in the lack of curves or fluctuation in p-values. The uniformity suggests that the FK test is pretty robust under the null hypothesis when the dispersion ratio is 1.0 with a lack of sharp curves. The second histogram is for the Fligner-Killeen Test with normal distribution and a dispersion ratio of 1.5. There is a moderate left skew due to a higher frequency of smaller p-values. There is also more spread, especially compared to the third histogram showing the p-values of the test with a dispersion ratio of 2.0. This histogram has a balance of low and high p-values demonstrating reduced sensitivity compared to the final histogram. The final histogram is for the Fligner-Killeen Test with normal distribution and a dispersion ratio of 2.0. This histogram depicts most of the p-values near 0.0 with a steep decline, especially compared to the previous two histograms. This demonstrates that with the highest dispersion ratio, the test detects deviations from the null with many small p-values. However, with this comes increased sensitivity as there is a steep decline in frequency levels as the p-values approach 1.0.







The following histograms display the distribution of p-values for the FK test with exponential

The histogram shows a relatively uniform distribution of p-values, this suggests the test maintains reasonable control over Type I error under the null hypothesis

A shift in p-value frequencies reflects the increasing sensitivity of the FK test to variance difference as dispersion increases.

The histogram reveals a clustering of low values indicating stronger detection of variance differences at higher dispersion ratios.

Conclusion

This simulation study compared the performance of the AB and FK tests under normal and exponential distributions with varying dispersion ratios. Our key findings include:

1. Both tests perform well in controlling Type I error rates for normal data, though exponential data introduces variability that impacts both tests.
2. Power increases with dispersion for all scenarios. The FK test outperforms the AB test for normal data, while both tests struggle with power for exponential data.
3. The FK test demonstrates robustness across normal and non-normal data, making it the preferred choice for detecting unequal dispersions, especially for normal distributions.

By systematically analyzing Type I error and power across all 12 scenarios, this stimulation provides actionable insights into the strengths and limitations of these non-parametric tests.