

HW2 Shen Ayashm

Jiacheng Chen

Sajjad Ayashm

2024-11-10

Question 1

Permutation Test

A permutation test is a non-parametric method used to determine if two groups differ significantly without assuming any specific distribution. This test assesses whether an observed difference could have occurred by chance.

`median.test` Algorithm

The `median.test` function helps determine if two groups differ in their medians. It calculates the observed median difference, then permutes the combined data multiple times (default: 1,000), generating new median differences each time. By counting how often these shuffled differences are as extreme as the original, we obtain a p-value that indicates the likelihood of the observed difference occurring by chance.

```
median.test <- function(x, y, n = 1000){  
  
  # Calculate observed statistic and combine data  
  observed_stat <- median(x) - median(y)  
  combine <- c(x, y)  
  
  extreme_count = 0  
  
  # Permutation test loop  
  for (i in 1:n){  
    shuffle <- sample(combine)  
    perm_stat <- median(head(shuffle,length(x))) - median(tail(shuffle,length(y)))  
    if (abs(perm_stat) >= abs(observed_stat)){  
      extreme_count <- extreme_count + 1  
    }  
  }  
  
  # Calculate p-value  
  p_value <- extreme_count / n  
  return(p_value)  
}
```

Question 2:

Question 2

To evaluate the power of three statistical tests—the permutation t-test, Wilcoxon–Mann–Whitney (WMW) test, and our custom median test—we conducted 1,000 simulations. For each simulation, we generated two data sets from a t_3 -distribution with a small effect (`delta = sqrt(3)/2`), which introduced a controlled difference between groups.

Each simulation used the three tests: - **Permutation t-test**: Calculated p-values with `oneway_test()`. - **Wilcoxon–Mann–Whitney test**: Used `wilcox.test()` to compute p-values. - **Median test**: Computed p-values using our `median.test()` function.

Results

After running the simulations, we calculated each test’s power by tracking the proportion of p-values below 0.05: - The **Median test** had a power of about 44.2%. - The **Permutation t-test** showed approximately 40.8% power. - The **WMW test** had the highest power, at 52.5%.

While the results may slightly vary with each run, the general trend shows that the WMW test has the highest power. The Median and Permutation t-tests performed similarly, with minor variation depending on data characteristics.

Question 3:

In question 3, we assessed the power of three statistical tests—the Median test, Permutation t-test, and Wilcoxon–Mann–Whitney (WMW) test—across different distributions. We simulated data from seven distributions, each with a controlled effect (`delta`) to measure each test’s ability to detect differences. The power of each test, shown as the percentage of times it correctly identifies a difference (p-value < 0.05), is displayed in Figure 1, offering a comparison of each test’s effectiveness across distributions.

As shown in **Figure 1**, the WMW test generally performs best, especially for uniform and normal distributions, where it consistently demonstrates high power. The Permutation t-test also performs well, often close to the WMW test in terms of power. The Median test’s effectiveness varies more; it performs well for the exponential distribution but shows lower power for distributions like logistic and t5. Overall, the WMW test stands out as a reliable choice across most data types, though each test has strengths depending on the distribution.

Question 4:

In question 4, we examine how often each test incorrectly detects a difference when none exists, known as the Type I error rate. To do this, we set `delta` to zero, meaning there is no real difference between the two groups. We then track how frequently each test gives a p-value below 0.05, which would indicate a false positive.

As shown in **Figure 2**, all three tests generally maintain Type I error rates close to the 5% significance level, with slight variation across distributions. The WMW test shows a slightly higher error rate for some distributions, like normal and exponential, compared to the other tests. Overall, each test appears reliable in controlling Type I error, though the WMW test exhibits a bit more variation depending on the distribution.

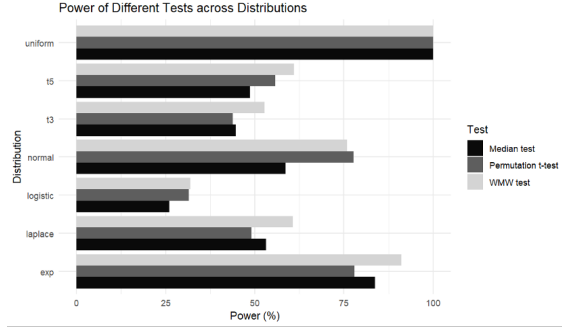


Figure 1: Power of Different Tests across Distributions

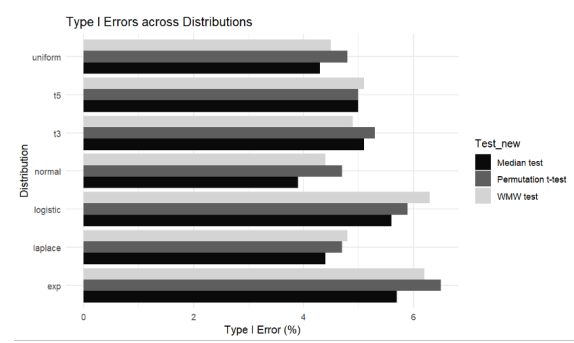


Figure 2: Type I Errors across Distributions

Question 5

i. Power Comparison

The power of each test varies based on the distribution's tail behavior. In heavy-tailed distributions (e.g., Laplace and exponential), the Median and WMW tests generally show higher power than the Permutation t-test, with the WMW test slightly outperforming the Median test. For light-tailed distributions (e.g., normal), the Permutation t-test has the highest power, while the Median test typically shows the lowest.

ii. Type I Error Comparison

Type I error rates also depend on the distribution's tail behavior. Heavy-tailed distributions tend to produce higher Type I errors. Among the three tests, the Median test consistently maintains the lowest Type I error rates, especially in the normal distribution, making it a conservative choice.

Recommendations

- **Maximizing Power:** The Median test is recommended for heavy-tailed distributions, but the WMW test may be preferable if power is prioritized.
- **Minimizing Type I Error:** The Median test is recommended for its low Type I error rates across distributions, making it a reliable choice for error control.

In summary, the Median test is advantageous for heavy-tailed distributions and when controlling Type I error, while the WMW test is suitable when prioritizing power.

Question 6:

We completed this homework as a team, working closely together on each part of the project. From developing code and running simulations to analyzing results and interpreting our findings, we both contributed equally along the way. We discussed our approach at every step to ensure accuracy and understanding, so each analysis reflects our combined efforts.

We also worked together on writing and reviewing the report, aiming to make the explanations as clear and straightforward as possible. Overall, this homework is truly a product of our collaboration and shared dedication.