

Why is the Mann-Whitney test significant even if the medians are equal?

Authored by
stats writer

June 30, 2024

RECOMMENDED CITATION

stats writer (2024). *Why is the Mann-Whitney test significant even if the medians are equal?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=161128>

The Mann-Whitney test is a statistical test used to compare two independent groups. It is significant even if the medians of the two groups are equal because it is based on the rank sums, not the actual values. This means that it takes into account the entire distribution of the data, not just the central tendency. Additionally, the test is robust to outliers and does not assume a normal distribution of the data, making it a reliable method for comparing groups with non-normal data. Therefore, even if the medians are equal, the Mann-Whitney test can still detect significant differences between the two groups.

FAQ: Why is the Mann-Whitney significant when the medians are equal?

A client has sent us the following question:

Q. I ran a Mann-Whitney test on two independent groups that have equal medians, the results were significant. I thought that the Mann-Whitney tested differences in medians. Why is the Mann-Whitney test significant when the medians are equal?

A. The answer is that the Mann-Whitney and the equivalent Wilcoxon test (hereafter called the Mann-Whitney-Wilcoxon test) are rank sum tests and not median tests. Basically, the Mann-Whitney-Wilcoxon test ranks all of the observations from both groups and then sums the ranks from one of the groups which is compared

with the expected rank sum. It is possible, although not very common, for groups to have different rank sums and yet have equal or nearly equal medians. An example is given below.

Consider the following example dataset of 120 observation (60 in each group) that has equal medians and a significant Mann-Whitney-Wilcoxon test.

```

+-----+
| y grp freq |
|-----|
| -2 1 20 |
| 0 1 20 |
| 5 1 20 |
|-----|
| -1 2 20 |
| 0 2 20 |
| 10 2 20 |
+-----+

```

The median for each of the two groups is zero, yet the Z-approximation for the Mann-Whitney-Wilcoxon is -2.16 with a p-value of .031.

On the other hand, for the median test the difference in medians is zero, since the two medians are equal, the t-value is 0 and has a p-value of 1.0 (some median tests report a chi-square value which in this case, will also be 0 with a p-value of 1.0).

The reason the Mann-Whitney-Wilcoxon is significant for the above data is the ranks for group 1 (other than those at the median) are lower than the ranks for group 2 (again, other than those values at the median). Here are the ranks for all the scores ignoring the frequencies to keep it simple.

```

+-----+
| y grp rank |
|-----|
| -2 1 1 |
| -1 2 2 |
| 0 2 3.5 |
| 0 1 3.5 |
| 5 1 5 |
| 10 2 6 |

```

+-----+

Note that the value of -2 has a rank of 1 and the value -1 has a rank of 2.

The values 5 and 10 have ranks of 5 and 6 respectively. The 0 scores all have a rank of 3.5.

Thus, other than the median (rank 3.5), all of the ranks for Grp1 are less than the ranks for group 2. With sufficient sample size the difference in ranks will be large enough to be significant even though the medians are equal.

The table below gives the sum of the ranks for each group for the full sample of 120 observations.

Grp	obs	rank sum
1	60	3230
2	60	4030

Group 2 clearly has a much larger sum of ranks than group 1. The difference in the sum of ranks is large enough to be statistically significant at the alpha equals .05 level.