IDM 404 Nonparametric Examples Data with 2 Levels Spring 2021

DATA:

|  |  |
| --- | --- |
| rating | feature |
| 1 | dur |
| 3 | dur |
| 4 | dur |
| 5 | dur |
| 1 | dur |
| 3 | dur |
| 4 | size |
| 4 | size |
| 4 | size |
| 4 | size |
| 3 | size |

 *We have two levels of feature and rating values that are not normally distributed. Note that there are 6 dur data points and 5 size data points.*

Minitab 17 Output

**Mood Median Test: rating versus feature**

Mood median test for rating

Chi-Square = 2.40 DF = 1 P = 0.122

 Individual 95.0% CIs

feature N< N≥ Median Q3-Q1 +---------+---------+---------+------

dur 4 2 3.00 3.25 (-------------------\*---------------)

size 1 4 4.00 0.50 (---------\*

 +---------+---------+---------+------

 1.0 2.0 3.0 4.0

Overall median = 4.00

\* NOTE \* Levels with < 6 observations have confidence < 95.0%

A 95.0% CI for median(dur) - median(size): (-3.00,1.04)

*I chose to conduct a Mood Median Test on this very small data set. The p-value suggests there is no significant difference between the medians. The CI suggests there is no significant difference between the medians. The small sample size is reason for caution.*

*Let's try a nonconsecutive list of features.*

*Let's make sure we have a sufficient number of observations.*

*Since we are not willing to assume the two data sets have similarly-shaped distributions, we will conduct a Mood Median Test instead of a Mann-Whitney.*

DATA:

rating1 pricelooks

1 looks

4 looks

1 looks

3 looks

4 looks

5 looks

1 price

3 price

5 price

3 price

5 price

4 price

5 price

4 price

5 price

4 price

3 price

3 price

3 price

3 price

1 price

5 price

4 price

4 price

5 price

5 price

5 price

5 price

5 price

5 price

5 price

5 price

3 looks

4 looks

2 looks

1 looks

2 looks

1 looks

2 looks

3 looks

2 looks

1 looks

2 looks

4 looks

**Mood Median Test: rating1 versus pricelooks**

Mood median test for rating1

Chi-Square = 9.68 DF = 1 P = 0.002

 Individual 95.0% CIs

pricelooks N≤ N> Median Q3-Q1 -----+---------+---------+---------+-

looks 17 1 2.00 3.00 (----\*--------------)

price 13 13 4.50 2.00 (--------\*----)

 -----+---------+---------+---------+-

 2.0 3.0 4.0 5.0

Overall median = 4.00

A 95.0% CI for median(looks) - median(price): (-3.00,-1.00)

*Mood's Median in Minitab reports a chi-square value.*

*Note that the p-value is small.*

*Note that the confidence intervals for the medians do not overlap.*

*Note that the confidence interval for the difference between the medians "goes from a minus to a minus".*

*Thus, the confidence interval does not contain zero.*

*The data suggest there is a significant difference between the medians.*

*Which median is larger?*

*We may conclude that price is more important than looks with this group of students.*

*By the way, the pull-down menu lists Mood's Median in the nonparametric statistics list.*

*The output for Minitab 17 shows Mood Median Test results.*

Additional Example using Mann Whitney (which requires similarly-shaped data sets)





 N Median

MW1 11 4.000

MW2 12 3.000

Point estimate for η1 - η2 is 0.500

95.5 Percent CI for η1 - η2 is (-1.000,2.000)

W = 145.0

Test of η1 = η2 vs η1 ≠ η2 is significant at 0.4417

The test is significant at 0.4296 (adjusted for ties)

*Note that the confidence interval for the difference between two medians contains zero. Note the phrasing “the test is significant at …” Since Minitab does not know the significance level (alpha) we are using, the phrasing is the same for high p-values and low p-values. We must remember that a test result that is above our assumed significance level of 0.05 indicates that the differences ARE NOT significant.*