Estimating Differences in Medians - Paired Samples

James M. Murray, Ph.D.
University of Wisconsin - La Crosse

Updated: September 13, 2017

PDF file location: http://www.murraylax.org/rtutorials/two_medians_paired.pdf

HTML file location: http://www.murraylax.org/rtutorials/two_medians_paired.html

---

Note on required packages: The following code required the package `psych` to perform statistics related to the median. If you have not already done so, download, install, and load the library with the following code:

```
install.packages("psych")
```

```
library("psych")
```

Here we investigate estimating the difference in the *medians* between two *paired samples*.

With **paired samples**, we have a single set of observations, but two measures or two variables for each observation. Obtaining two or more measures from each observation to pair can result from the following situations:

1. **Across-time measures:** The same dependent variable is measured for each individual but at two different time periods. For example, a sample of individuals may have their income measured once in 2013 and again in 2014, and a researcher could ask whether there was a change in average income from one year to the next.
2. **Different conditions measures:** The same dependent variable is measured for each individual, but under two different conditions, or before and after some treatment. For example, a sample of high school students may have test scores measured before introducing them to new curriculum and afterwards. A researcher could ask whether the curriculum affected test scores.
3. **Related topics measures:** Two slightly different variables are measured for each individual. For example, foreign language students may take separate exams for writing proficiency and speaking proficiency. A researcher could ask whether students are more proficient with writing in a foreign language versus speaking.

Testing differences in *medians* is appropriate when the variables are in the same units and at the ordinal, interval, or ratio scale.

Because ordinal data is categorical data, the mean is not an appropriate measure of center. However, since ordinal data can be sorted or ranked, it is possible to calculate the median.

While one can also measure the mean of interval or ratio data, it is often desirable to compute the median for populations that have a skewed distribution. That is, an asymmetric distribution where one end of the distribution extends farther from the median than another end. The extreme values of the long end of the distribution cause the mean to move towards that tail, away from the middle of the distribution.

An alternative measure for the median is the **interpolated median**. This is another measure of center which takes into account the percentage of the data that is strictly below versus strictly above the median.

1. Data

In this data set, students in fourth through sixth from three school districts in Michigan ranked their how important each of the following were to them: achieving good grades, having athletic ability, having popularity, and having money. A rank of 1 indicates highest importance and a rank of 4 indicates lowest importance.

The data set comes from Chase, M. A., and Dummer, G. M. (1992), "The Role of Sports as a Social Determinant for Children," *Research Quarterly for Exercise and Sport*, 63, 418-424. Available at: http://shapeamerica.tandfonline.com/doi/abs/10.1080/02701367.1992.10608764

The code below downloads the data and assigns the it to a data frame that we call `kidsdata`.

```
kidsdata <- read.csv(url("http://www.murraylax.org/datasets/gradeschool.csv"))
```

2. Compute Medians

The data set includes a variables `Grades` and `Sports`, which are each rankings on a scale of 1-4, with lower values indicating higher performance, for how much importance students place on good grades and on participating in sports, respectively. Let us begin by computing the **sample median** for each variable.

```
median(kidsdata$Grades)
```

```
## [1] 3
```

```
median(kidsdata$Sports)
```

```
## [1] 2
```

We see that the sample median response for the importance of grades is equal to 3, while the median response for sports is equal to 2. This means that *in our sample*, on average children place more importance on sports than on grades.

Similarly, we can compute the **sample interpolated median** for each variable:

```
interp.median(kidsdata$Grades)
```

```
## [1] 2.665414
```

```
interp.median(kidsdata$Sports)
```

```
## [1] 1.980519
```

Here we can see that the difference in the center of the distributions is somewhat smaller than implied by the simple median. The response for grades is centered somewhat below 3 (interpolated median = 2.67) and the response for sports is centered very close to 2 (interpolated median = 1.98).

3. Hypothesis Test on Differences in the Distributions

The **Wilcoxon Signed Rank Test** considers the hypothesis that the distributions for two paired samples are centered around the same value. Using the example above, let us test the hypothesis that in the population the distribution for *importance of grades* is centered around the same value as the distribution for the importance of sports.

Null hypothesis: The center of the distribution for grades is *equal* to the center of the distribution for sports
Alternative hypothesis: The center of the distribution for grades is *different* than the center of the distribution for sports

The following code calls the `wilcox.test()` function to test this hypothesis:

```r
wilcox.test(kidsdata$Grades, kidsdata$Sports, alternative="two.sided", paired=TRUE)
```

```
##
##  Wilcoxon signed rank test with continuity correction
##
## data:  kidsdata$Grades and kidsdata$Sports
## V = 77586, p-value = 4.085e-12
## alternative hypothesis: true location shift is not equal to 0
```

The first and second parameters `kidsdata$Grades` and `kidsdata$Sports` specify the two variables to compare. The third parameter, `alternative="two.sided"` specifies that this is a two-tailed test with an alternative hypothesis that the center of the two distributions are *different*, and the final parameter `paired=TRUE` tells the function to run a paired-samples test (versus an independent-samples test).

The p-value is equal to 4.085e-12 (or 4.085×10^{-12}). This is much less than a significance level of 0.05 or 5%. Therefore, we *find sufficient statistical evidence* that the median response for the importance for grades is different than the median response for the importance of sports.