The Wilcoxon test was developed to analyze data from studies with repeated-measures and matched-subjects designs. For a repeated-measures design, an individual is assessed on a measure on two occasions or under two conditions. Each individual is a case in the SPSS data file and has scores on two variables, the score obtained on the measure on one occasion (or under one condition) and the score obtained on the measure on a second occasion (or under a second condition). The goal of repeated-measures designs is to determine whether participants changed significantly across occasions (or conditions).

For a matched-subjects design, participants are paired, typically on one or more nuisance variables, and each participant in the pair is assessed once on a measure. Each pair of participants is a case in the SPSS data file and has scores on two variables, the score obtained on the measure by one participant under one condition and the score obtained on the measure by the other participant under the other condition. The purpose of matched-subjects designs is to evaluate whether the pairs of participants differ significantly under the two conditions.

For both types of studies, each case in the SPSS data file has scores on two variables (that is, paired scores). The Wilcoxon test evaluates differences between paired scores, either repeated or matched. The variables for the Wilcoxon test have multiple possible scores, with the focus on whether the median of the variables differs significantly.

The Wilcoxon test can be used to analyze data from a number of different designs:

1. Repeated-measures designs with an intervention
2. Repeated-measures designs without an intervention
3. Matched-subjects designs with an intervention
4. Matched-subjects designs without an intervention

Understanding the Wilcoxon Test

To help understand the Wilcoxon test, we will first describe what data are being analyzed. We will look at an example provided by Green and Salkind (2008) dealing with concerns for job security and job pay (Lesson 44 from Green & Salkind).

For this example: Security, is defined as concern for job security on a rating scale from 1, indicating no concern, to 10, indicating ultimate concern and Pay is defined as concern for job pay on a rating scale from 1, indicating no concern, to 10, indicating ultimate concern.

The Wilcoxon test involves ranking all nonzero difference scores disregarding sign, reattaching the sign to the rank, and then evaluating the mean of the positive and the mean of the negative ranks.
ASSUMPTIONS UNDERLYING A WILCOXON TEST

Assumption 1: Each pair of observations must represent a random sample from a population and must be independent of every other pair of observations.

If the data are from a repeated-measures design, the paired scores for each participant must be independent of the paired scores from any other participant. If the data are from a matched-subjects design, the paired scores from any matched set of participants must be independent of the paired scores of any other matched pair of participants. If the independence assumption is violated, the test is likely to yield inaccurate results. It should be noted that the analysis presumes dependency between scores within pairs.

Assumption 2: The z test yields relatively accurate results to the extent that the sample size is large.

The results for the test should be fairly accurate if the sample size is 16 or greater pairs.

Assumption 3: The Distribution of the differences scores is continuous and symmetrical in the population.

This assumption pertains to the difference scores, not to the ranked scores. If difference scores are continuously distributed, there should be no ties in the difference scores across pairs of scores.
EFFECT SIZE STATISTICS FOR THE WILCOXON TEST

SPSS does not report an effect size for the Wilcoxon test. However, simple indices can be reported to communicate the size of the effect. For the Wilcoxon test, the mean positive ranked difference score and the mean negative ranked difference score could be reported.

THE RESEARCH QUESTION

The research question used in this example is asked to reflect differences in medians between conditions.

Does median concern for job security differ from median concern for job pay?

CONDUCTING THE TWO RELATED-SAMPLES (WILCOXON) TEST IN SPSS

To conduct the Wilcoxon test in SPSS, use the following steps:

- Open the dataset in SPSS to be used for the Wilcoxon Test analysis
- Click Analyze, click (mouse over) Nonparametric Tests, and then click 2 Related Samples…
  - You should now be in the Two-Related Samples Tests dialog box
    - Holding down the Ctrl key, click Pay and Security, and click ▶ to move them to the Test Pair(s) List: box
    - Under Test Type – be sure that Wilcoxon is selected [✓] – this should be the default
    - Click Options
      - Under Statistics
        - Select [✓] Descriptive
      - Click Continue
        - Click OK
- You are now ready to analyze the output data…

SPSS OUTPUT

NPar Tests

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security</td>
<td>30</td>
<td>4.50</td>
<td>1.834</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Pay</td>
<td>30</td>
<td>5.67</td>
<td>1.493</td>
<td>3</td>
<td>9</td>
</tr>
</tbody>
</table>
Wilcoxon Signed Ranks Test

<table>
<thead>
<tr>
<th>Ranks</th>
<th>N</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pay - Security</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative Ranks</td>
<td>6a</td>
<td>12.25</td>
<td>73.50</td>
</tr>
<tr>
<td>Positive Ranks</td>
<td>20b</td>
<td>13.88</td>
<td>277.50</td>
</tr>
<tr>
<td>Ties</td>
<td>4c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. Pay < Security
b. Pay > Security
c. Pay = Security

Test Statistics

<table>
<thead>
<tr>
<th></th>
<th>Pay - Security</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z</td>
<td>-2.626a</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.009</td>
</tr>
</tbody>
</table>

a. Based on negative ranks.
b. Wilcoxon Signed Ranks Test

The Wilcoxon test, which evaluated difference between medians for job security and job pay concerns, is significant $z = -2.63$, $p < .01$. That is, the results indicate significant differential concern for pay versus security.

APA RESULTS

Based on the results produced from the above example, the APA results would be:

A Wilcoxon test was conducted to evaluate whether employees showed greater concern for pay or security. The results indicated a significant difference, $z = -2.63$, $p < .01$. The mean of the ranks in favor of pay was 13.88, while the mean of the ranks in favor of security was 12.25.

REFERENCE