

Mann-Whitney U Test using SPSS Statistics

Introduction The Mann-Whitney U test is used to compare differences between two independent groups when the dependent variable is either ordinal or continuous, but not normally distributed. For example, you could use the Mann-Whitney U test to understand whether attitudes towards pay discrimination, where attitudes are measured on an ordinal scale, differ based on gender (i.e., your dependent variable would be "attitudes towards pay discrimination" and your independent variable would be "gender", which has two groups: "male" and "female"). Alternately, you could use the Mann-Whitney U test to understand whether salaries, measured on a continuous scale, differed based on educational level (i.e., your dependent variable would be "salary" and your independent variable would be "educational level", which has two groups: "high school" and "university"). The Mann-Whitney U test is often considered the nonparametric alternative to the independent t-test although this is not always the case.

Unlike the independent-samples t-test, the Mann-Whitney U test allows you to draw different conclusions about your data depending on the assumptions you make about your data's distribution. These conclusions can range from simply stating whether the two populations differ through to determining if there are differences in medians between groups. These different conclusions hinge on the shape of the distributions of your data, which we explain more about later.

In our enhanced Mann-Whitney U test guide, we take you through all the steps required to understand when and how to use the Mann-Whitney U test, showing you the required procedures in SPSS Statistics, and how to interpret and report your output. You can access this enhanced Mann-Whitney U test guide by subscribing to the site [here](#). In this "quick start" guide, we show you the basics of the Mann-Whitney U test using one of SPSS Statistics's procedures when the critical assumption of this test is violated. Before we show you how to do this, we explain the different assumptions that your data must meet in order for a Mann-Whitney U test to give you a valid result. We discuss these assumptions next.

Assumptions When you choose to analyse your data using a Mann-Whitney U test, part of the process involves checking to make sure that the data you want to analyse can actually be analysed using a Mann-Whitney U test. You need to do this because it is only appropriate to use a Mann-Whitney U test if your data "passes" four assumptions that are required for a Mann-Whitney U test to give you a valid result. In practice, checking for these four assumptions just adds a little bit more time to your analysis, requiring you to click a few more buttons in SPSS Statistics when performing your analysis, as well as think a little bit more about your data, but it is not a difficult task.

Before we introduce you to these four assumptions, do not be surprised if, when analysing your own data using SPSS Statistics, one or more of these assumptions is violated (i.e., is not met). This is not uncommon when working with real-world data rather than textbook examples, which often only show you how to carry out a Mann-Whitney U test when everything goes well! However, don't worry. Even when your data fails certain assumptions, there is often a solution to overcome this. First, let's take a look at these four assumptions:

- **Assumption No1:** Your dependent variable should be measured at the ordinal or continuous level. Examples of ordinal variables include Likert items (e.g., a 7-point scale from "strongly agree" through to "strongly disagree"), amongst other ways of ranking categories (e.g., a 5-point scale explaining how much a customer liked a product, ranging from "Not very much" to "Yes, a lot"). Examples of continuous variables include revision time (measured in hours), intelligence (measured using IQ score), exam performance (measured from 0 to 100), weight (measured in kg), and so forth. You can learn more about ordinal and continuous variables in our article: [Types of Variable](#).
- **Assumption No 2:** Your independent variable should consist of two categorical, independent groups. Example independent variables that meet this criterion include gender (2 groups: male or female),

employment status (2 groups: employed or unemployed), smoker (2 groups: yes or no), and so forth.

- Assumption No3: You should have independence of observations, which means that there is no relationship between the observations in each group or between the groups themselves. For example, there must be different participants in each group with no participant being in more than one group. This is more of a study design issue than something you can test for, but it is an important assumption of the Mann-Whitney U test. If your study fails this assumption, you will need to use another statistical test instead of the Mann-Whitney U test (e.g., a Wilcoxon signed-rank test). If you are unsure whether your study meets this assumption, you can use our Statistical Test Selector, which is part of our enhanced content.
- Assumption No4: A Mann-Whitney U test can be used when your two variables are not normally distributed. However, in order to know how to interpret the results from a Mann-Whitney U test, you have to determine whether your two distributions (i.e., the distribution of scores for both groups of the independent variable; for example, 'males' and 'females' for the independent variable, 'gender') have the same shape. To understand what this means, take a look at the diagram below:

The Mann-Whitney U Identical Shaped Distributions

In the two diagrams above, the distribution of scores for 'males' and 'females' have the same shape. In the diagram on the left, you cannot see the distribution of scores for 'males' (illustrated in blue on the diagram on the right) because the two distributions are identical (i.e., both distributions are identical, so they are 'on top of each other' in the diagram, with the blue-coloured male distribution underneath the red-coloured female distribution). However, in the diagram on the right, even though both distributions have the same shape, they have a different location (i.e., the distribution of one of the groups of the independent variable has higher or lower values compared to the second distribution in our example, females have 'higher' values than males, overall).

When you analyse your own data, it is extremely unlikely that your two distributions will be identical, but they may have the same (or a 'similar') shape. If they do have the same shape, you can use SPSS Statistics to carry out a Mann-Whitney U test to compare the medians of your dependent variable (e.g., engagement score) for the two groups (e.g., males and females) of the independent variable (e.g., gender) you are interested in. However, if your two distributions have a different shape, you can only use the Mann-Whitney U test to compare mean ranks.

Therefore, when carrying out a Mann-Whitney U test, you must also use SPSS Statistics to determine whether your two distributions have the same shape or a different shape. This requires a few more procedures in SPSS Statistics, but it is an easy step-by-step process that we show you how to do in our enhanced Mann-Whitney U test guide. In this "quick start" guide, we show you how to carry out a Mann-Whitney U test assuming that your two distributions do not have a similar shape, such that you can only compare mean ranks and not medians.

You can check assumption No 4 using SPSS Statistics. Before doing this, you should make sure that your data meets the first three assumptions, although you don't need SPSS Statistics to do this. Just remember that if you do not check assumption No 4, you will not know whether you are correctly comparing mean ranks or medians, and the results you get when running a Mann-Whitney U test may not be valid. This is why we dedicate a number of sections of our enhanced Mann-Whitney U test guide to help you get this right. You can learn more about assumption No 4 and what you will need to interpret in the Assumptions section of our enhanced Mann-Whitney U test guide, which you can access by subscribing to the site here.

In the Test Procedure in SPSS Statistics section of this "quick start" guide, we illustrate the SPSS Statistics procedure to perform a Mann-Whitney U test assuming that your two distributions are not the

same shape and you have to interpret mean ranks rather than medians. First, we set out the example we use to explain the Mann-Whitney U test procedure in SPSS Statistics.

0.1 Example

The concentration of cholesterol (a type of fat) in the blood is associated with the risk of developing heart disease, such that higher concentrations of cholesterol indicate a higher level of risk, and lower concentrations indicate a lower level of risk. If you lower the concentration of cholesterol in the blood, your risk for developing heart disease can be reduced. Being overweight and/or physically inactive increases the concentration of cholesterol in your blood. Both exercise and weight loss can reduce cholesterol concentration. However, it is not known whether exercise or weight loss is best for lowering cholesterol concentration. Therefore, a researcher decided to investigate whether an exercise or weight loss intervention was more effective in lowering cholesterol levels. To this end, the researcher recruited a random sample of inactive males that were classified as overweight. This sample was then randomly split into two groups: Group 1 underwent a calorie-controlled diet (i.e., the 'diet' group) and Group 2 undertook an exercise-training programme (i.e., the 'exercise' group). In order to determine which treatment programme was more effective, cholesterol concentrations were compared between the two groups at the end of the treatment programmes.

- In SPSS Statistics, we entered the scores for cholesterol concentration, our dependent variable, under the variable name **Cholesterol**. Next, we created a grouping variable, called **Group**, which represented our independent variable. Since our independent variable had two groups - 'diet' and 'exercise' - we gave the diet group a value of "1" and the exercise group a value of "2". If you do not label your two groups, SPSS Statistics will not be able to distinguish between them and the Mann-Whitney U test will not run.

Note: There are two different procedures in SPSS Statistics to run a Mann-Whitney U test: a new and legacy procedure. How we have explained the data setup above relates to the legacy procedure (and the new procedure when your dependent variable is continuous), which is what we take you through in the Test Procedure in SPSS Statistics section next. We mention this because if you are using the new procedure, you have to make changes to your data setup if your dependent variable is ordinal (i.e., as opposed to being continuous). We explain how to do this in our enhanced Mann-Whitney U test guide, which you can access by subscribing to the site [here](#).

- If you read assumption No 4 earlier, you'll know that the SPSS Statistics procedure when analysing your data using a Mann-Whitney U test is different depending on the shape of the two distributions of your independent variable.
- In our example, where our dependent variable is cholesterol concentration, **Cholesterol**, we are referring to the two distributions of the independent variable, **Group** (i.e., the distribution of scores for Group 1 the 'diet' group and Group 2 the 'exercise' group).
- In the 10 steps below, we show you how to analyse your data using a Mann-Whitney U test in SPSS Statistics when these two distributions have a different shape, and therefore, you have to compare the mean ranks of your dependent variable rather than medians.

To use SPSS Statistics to determine whether your two distributions have the same or different shapes, or if you want to know how to use SPSS Statistics to carry out a Mann-Whitney U test when your two distributions have the same shape, such that you need to compare medians rather than mean ranks, you will need to access the Procedures section of our enhanced Mann-Whitney U test guide (N.B., you can do this by subscribing to the site here).

Furthermore, the 10 steps below also show you how to carry out a Mann-Whitney U test using the legacy procedure in SPSS Statistics. As we explained earlier, there are two different procedures in SPSS Statistics to run a Mann-Whitney U test: a new and legacy procedure. We recommend the new procedure if your two distributions have the same shape because it is a little easier to carry out, but the legacy procedure is fine if your two distributions have different shapes. We show you the new and legacy procedures in our enhanced Mann-Whitney U test guide.

At the end of the 9 steps below, we show you how to interpret the results from this test using mean ranks.

Click Analyze \downarrow Nonparametric Tests \downarrow Legacy Dialogs \downarrow 2 Independent Samples... on the top menu, as shown below:

The Mann-Whitney U Test Menu Published with written permission from SPSS Statistics, IBM Corporation. You will be presented with the Two-Independent-Samples Tests dialogue box, as shown below:

The Mann-Whitney U Test Dialogue Box Published with written permission from SPSS Statistics, IBM Corporation. Transfer the dependent variable, Cholesterol, into the Test Variable List: box and the independent variable, Group, into the Grouping Variable: box by using the SPSS Right Arrow Button button or by dragging-and-dropping the variables into the boxes.

The Mann-Whitney U Test Dialog Box Published with written permission from SPSS Statistics, IBM Corporation. Note: Make sure that the Mann-Whitney U checkbox is ticked in the Test Type area and the Grouping Variable: box is highlighted in yellow (as seen above). If it is not highlighted in yellow, simply click your cursor in the Grouping Variable: box to highlight it.

Click the Define Groups button. The button will not be clickable if you have not highlighted the Grouping Variable: box as instructed in Step 4. You will be presented with the following screen:

The Define Groups Dialogue Box within the Mann-Whitney U Test Published with written permission from SPSS Statistics, IBM Corporation. Click through to the next page for the remaining procedure and how to interpret the output.

Mann-Whitney U Test using SPSS Statistics (cont...)

Enter 1 into the Group 1: box and enter 2 into the Group 2: box. Remember that we labelled the Diet group as 1 and the Exercise group as 2.

The Define Groups Dialogue Box within the Mann-Whitney U Test Published with written permission from SPSS Statistics, IBM Corporation. Note: If you have more than two groups in your study (e.g., three groups: diet, exercise and drug groups), but only wanted to compare two (e.g., the diet and drug groups), you could type 1 into the Group 1: box and 3 into the Group 2: box (i.e., if you wished to compare the diet with drug group).

Click the Continue Button button.

If you wish to use this procedure to generate some descriptive statistics, click on the Options Button button and then tick Descriptive and Quartiles within the Statistics area. You will be presented with the dialogue box below:

The Options Dialogue Box within the Mann-Whitney U Test Published with written permission from SPSS Statistics, IBM Corporation. Click the Continue Button button, which will bring you back to the main dialogue box with the Grouping Variable: box now completed, as shown below:

The Mann-Whitney U Test Dialogue Box Published with written permission from SPSS Statistics, IBM Corporation. Click the button. This will generate the output for the Mann-Whitney U test.

0.2 Output and Interpretation

If you have been following this guide from page one, you will know that the following output and interpretation relates to the Mann-Whitney U test results when your two distributions have a different shape, such that you are comparing mean ranks rather than medians. This is what happens when your data has violated Assumption No 4 of the Mann-Whitney U test.

The output is also based on the use of the Legacy Procedure in SPSS Statistics. If you have used the New Procedure in SPSS Statistics or you need to know how to interpret medians because your data has met Assumption No 4 of the Mann-Whitney U test, we explain how to do this in our enhanced Mann-Whitney U test guide, which you can access by subscribing to the site [here](#).

In the SPSS Statistics output below, we show you how to report the Mann-Whitney U test using mean ranks. To do this, SPSS Statistics produces three tables of output:

Descriptives The Descriptive Statistics table looks as follows:

Descriptives Output from the Mann-Whitney U Test Published with written permission from SPSS Statistics, IBM Corporation.

Although we have decided to show you how you can get SPSS Statistics to provide descriptive statistics for the Mann-Whitney U test, they are not actually very useful. The reason for this is twofold. Firstly, in order to compare the groups, we need the individual group values, not the amalgamated ones. This table does not provide us with this vital information, so we cannot compare any possible differences between the exercise and diet groups. Secondly, we chose the Mann-Whitney U test because one of the individual groups (exercise group) was not normally distributed. However, we have not tested to see if the amalgamation of the two groups results in the larger group being normally distributed. Therefore, we do not know whether to use the mean and standard deviation or the median and interquartile range (IQR). The IQR is the 25th to 75th percentile. This will act as a surrogate to the standard deviation we would otherwise report if the data were normally distributed. For these reasons, we recommend that you ignore this table.

0.3 Ranks Table

The Ranks table is the first table that provides information regarding the output of the actual Mann-Whitney U test. It shows mean rank and sum of ranks for the two groups tested (i.e., the exercise and diet groups):

Descriptives Output from the Mann-Whitney U Test Published with written permission from SPSS Statistics, IBM Corporation. The table above is very useful because it indicates which group can be considered as having the higher cholesterol concentrations, overall; namely, the group with the highest mean rank. In this case, the diet group had the highest cholesterol concentrations.

Test Statistics Table This table shows us the actual significance value of the test. Specifically, the Test Statistics table provides the test statistic, U statistic, as well as the asymptotic significance (2-tailed) p-value.

Descriptives Output from the Mann-Whitney U Test Published with written permission from SPSS Statistics, IBM Corporation. From this data, it can be concluded that cholesterol concentration in the diet group was statistically significantly higher than the exercise group ($U = 110$, $p = .014$). Depending on the size of your groups, SPSS Statistics will produce both exact and asymptotic statistical significance levels. Understanding which one to use is explained in our enhanced guide.