

Chapter 430

Mann-Whitney U or Wilcoxon Rank-Sum Tests (Simulation)

Introduction

This procedure provides sample size and power calculations for one- or two-sided two-sample Mann-Whitney U or Wilcoxon Rank-Sum Tests based on simulation. This test is the nonparametric alternative to the traditional two-sample t-test. Other names for this test are the Mann-Whitney-Wilcoxon test or the Wilcoxon-Mann-Whitney test. The user can also conduct non-inferiority, superiority by a margin, and non-zero null tests

The design corresponding to this test procedure is sometimes referred to as a *parallel-groups* design. This design is used in situations such as the comparison of the income level of two regions, the nitrogen content of two lakes, or the effectiveness of two drugs.

There are several statistical tests available for the comparison of the center of two populations. You can examine the sections below to identify whether the assumptions and test statistic you intend to use in your study match those of this procedure, or if one of the other **PASS** procedures may be more suited to your situation.

Other PASS Procedures for Comparing Two Means

Procedures in PASS are primarily built upon the testing methods, test statistic, and test assumptions that will be used when the analysis of the data is performed. You should check to identify that the test procedure described below in the Test Procedure section matches your intended procedure. If your assumptions or testing method are different, you may wish to use one of the other two-sample procedures available in PASS. These procedures are Two-Sample T-Tests Assuming Equal Variance, Two-Sample T-Tests Allowing Unequal Variance, Two-Sample Z-Tests Assuming Equal Variance, and Two-Sample Z-Tests Allowing Unequal Variance. There is also a Mann-Whitney U or Wilcoxon Rank-Sum Tests procedure based on analytic results available. The methods, statistics, and assumptions for those procedures are described in the associated chapters.

If you wish to show that the mean of one population is larger (or smaller) than the mean of another population by a specified amount, you should use one of the clinical superiority procedures for comparing means. Non-inferiority, equivalence, and confidence interval procedures are also available.

Test Assumptions

When running a Mann-Whitney-Wilcoxon test, the basic assumptions are random sampling from each of the two populations and that the measurement scale is at least ordinal. These assumptions are sufficient for testing whether the two populations are different. If we can additionally assume that the two populations are identical except possible for a difference in location, then this test can be used as a test of equal means or medians.

Test Procedure

If we assume that the two populations differ only in location, with μ_1 and μ_2 representing the means of the two populations of interest, and that $\delta = \mu_1 - \mu_2$, the null hypothesis for comparing the two means (or medians) is $H_0 : \delta = 0$. The alternative hypothesis can be any one of

$$H_1 : \delta \neq 0$$

$$H_1 : \delta > 0$$

$$H_1 : \delta < 0$$

depending upon the desire of the researcher or the protocol instructions. A suitable Type I error probability (α) is chosen for the test, the data is collected, and the data from both groups are combined and then ranked.

The Mann-Whitney test statistic is defined as follows in Gibbons (1985).

$$z = \frac{W_1 - \frac{N_1(N_1 + N_2 + 1)}{2} + C}{s_W}$$

where

$$W_1 = \sum_{k=1}^{N_1} \text{Rank}(X_{1k})$$

The ranks are determined after combining the two samples. The standard deviation is calculated as

$$s_W = \sqrt{\frac{N_1 N_2 (N_1 + N_2 + 1)}{12} - \frac{N_1 N_2 \sum_{i=1} (t_i^3 - t_i)}{12(N_1 + N_2)(N_1 + N_2 - 1)}}$$

where t_i is the number of observations tied at value one, t_2 is the number of observations tied at some value two, and so forth.

The correction factor, C , is 0.5 if the rest of the numerator of z is negative or -0.5 otherwise. The value of z is then compared to the standard normal distribution.

The null hypothesis is rejected in favor of the alternative if,

for $H_1 : \delta \neq 0$,

$$z < z_{\alpha/2} \text{ or } z > z_{1-\alpha/2},$$

for $H_1 : \delta > 0$,

$$z > z_{1-\alpha},$$

or, for $H_1 : \delta < 0$,

$$z < z_{\alpha}.$$

Comparing the z -statistic to the cut-off z -value (as shown here) is equivalent to comparing the p -value to α .

Power Calculation

Simulation

Simulation allows us to estimate the power and significance level that is actually achieved by a test procedure in situations that are not mathematically tractable.

The steps to a simulation study are

1. Specify how the test is carried out. This includes indicating how the test statistic is calculated and how the significance level is specified.
2. Generate random samples from the distributions specified by the alternative hypothesis. Calculate the test statistics from the simulated data and determine if the null hypothesis is accepted or rejected. Tabulate the number of rejections and use this to calculate the test's power.
3. Generate random samples from the distributions specified by the null hypothesis. Calculate each test statistic from the simulated data and determine if the null hypothesis is accepted or rejected. Tabulate the number of rejections and use this to calculate the test's significance level.
4. Repeat steps 2 and 3 several thousand times, tabulating the number of times the simulated data leads to a rejection of the null hypothesis. The power is the proportion of simulated samples in step 2 that lead to rejection. The significance level is the proportion of simulated samples in step 3 that lead to rejection.

Generating Random Distributions

Two methods are available in *PASS* to simulate random samples. The first method generates the random variates directly, one value at a time. The second method generates a large pool (over 10,000) of random values and then draw the random numbers from this pool. This second method can cut the running time of the simulation by 70%.

As mentioned above, the second method begins by generating a large pool of random numbers from the specified distributions. Each of these pools is evaluated to determine if its mean is within a small relative tolerance (0.0001) of the target mean. If the actual mean is not within the tolerance of the target mean, individual members of the population are replaced with new random numbers if the new random number moves the mean towards its target. Only a few hundred such swaps are required to bring the actual mean to within tolerance of the target mean. This population is then sampled with replacement using the uniform distribution. We have found that this method works well as long as the size of the pool is the maximum of twice the number of simulated samples desired and 10,000.

Standard Deviations

Care must be used when either the null or alternative distribution is not normal. In these cases, the standard deviation is usually not specified directly. For example, you might use a gamma distribution with a shape parameter of 1.5 and a mean of 4 as the null distribution and a gamma distribution with the same shape parameter and a mean of 5 as the alternative distribution. This allows you to compare the two means. However, although the shape parameters are constant, the standard deviations, which are based on both the shape parameter and the mean, are not.

A Note on Specifying the Difference in Means/Medians

When the difference in means is specified in this procedure, it is the assumed difference in means for power or sample size calculations. It does not mean that the study will be powered to show that the mean difference is this amount, but rather that the design is powered to reject the null hypothesis of equal means if this were the true difference in means. If your purpose is to show that one mean is greater than another by a specific amount, you should use one of the clinical superiority procedures for comparing means.

Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design tab. For more information about the options of other tabs, go to the Procedure Window chapter.

Design Tab

The Design tab contains most of the parameters and options that you will be concerned with.

Solve For

Solve For

This option specifies whether you want to find *Power* or *Sample Size (NI)* from the simulation. Select *Power* when you want to estimate the power of a certain scenario. Select *Sample Size (NI)* when you want to determine the sample size needed to achieve a given power and alpha error level. Finding *Sample Size (NI)* is very computationally intensive, and so it may take a long time to complete.

Test and Simulations

Alternative Hypothesis

Specify the alternative hypothesis of the test. Since the null hypothesis is the opposite of the alternative, specifying the alternative is all that is needed. Usually, the two-tailed (\neq) option is selected. Note that $\delta = \mu_1 - \mu_2$.

The options containing only $<$ or $>$ are one-tailed tests. When you choose one of these options, you must be sure that the input parameters match this selection. Note that one-tailed tests could be used for a **non-inferiority test** or a **superiority by a margin test**.

Possible selections are:

- **Two-Sided (H1: $\delta \neq \delta_0$)**
This is the most common selection. It yields the *two-tailed t-test*. Use this option when you are testing whether the means are different, but you do not want to specify beforehand which mean is larger. Many scientific journals require two-tailed tests.
- **One-Sided (H1: $\delta < \delta_0$)**
This option yields a *one-tailed t-test*. Use it when you are only interested in the case in which δ is less than δ_0 .
- **One-Sided (H1: $\delta > \delta_0$)**
This option yields a *one-tailed t-test*. Use it when you are only interested in the case in which δ is greater than δ_0 .

Simulations

This option specifies the number of iterations, M , used in the simulation. As the number of iterations is increased, the accuracy and running time of the simulation will be increased also.

The precision of the simulated power estimates are calculated from the binomial distribution. Thus, confidence intervals may be constructed for various power values. The following table gives an estimate of the precision that is achieved for various simulation sizes when the power is either 0.50 or 0.95. The table values are interpreted as follows: a 95% confidence interval of the true power is given by the power reported by the simulation plus and minus the 'Precision' amount given in the table.

Mann-Whitney U or Wilcoxon Rank-Sum Tests (Simulation)

Simulation Size M	Precision when Power = 0.50	Precision when Power = 0.95
100	0.100	0.044
500	0.045	0.019
1000	0.032	0.014
2000	0.022	0.010
5000	0.014	0.006
10000	0.010	0.004
50000	0.004	0.002
100000	0.003	0.001

Notice that a simulation size of 1000 gives a precision of plus or minus 0.01 when the true power is 0.95. Also note that as the simulation size is increased beyond 5000, there is only a small amount of additional accuracy achieved.

Power and Alpha

Power

This option specifies one or more values for power. Power is the probability of rejecting a false null hypothesis, and is equal to one minus Beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected. In this procedure, a type-II error occurs when you fail to reject the null hypothesis of equal means when in fact the means are different.

Values must be between zero and one. Historically, the value of 0.80 (Beta = 0.20) was used for power. Now, 0.90 (Beta = 0.10) is also commonly used.

A single value may be entered here or a range of values such as *0.8 to 0.95 by 0.05* may be entered.

Alpha

This option specifies one or more values for the probability of a type-I error. A type-I error occurs when a true null hypothesis is rejected. In this procedure, a type-I error occurs when you reject the null hypothesis of equal means when in fact the means are equal.

Values must be between zero and one. Historically, the value of 0.05 has been used for alpha. This means that about one test in twenty will falsely reject the null hypothesis. You should pick a value for alpha that represents the risk of a type-I error you are willing to take in your experimental situation.

You may enter a range of values such as *0.01 0.05 0.10* or *0.01 to 0.10 by 0.01*.

Sample Size (When Solving for Sample Size)

Group Allocation

Select the option that describes the constraints on $N1$ or $N2$ or both.

The options are

- **Equal ($N1 = N2$)**

This selection is used when you wish to have equal sample sizes in each group. Since you are solving for both sample sizes at once, no additional sample size parameters need to be entered.

- **Enter $N2$, solve for $N1$**

Select this option when you wish to fix $N2$ at some value (or values), and then solve only for $N1$. Please note that for some values of $N2$, there may not be a value of $N1$ that is large enough to obtain the desired power.

Mann-Whitney U or Wilcoxon Rank-Sum Tests (Simulation)

- **Enter $R = N2/N1$, solve for $N1$ and $N2$**

For this choice, you set a value for the ratio of $N2$ to $N1$, and then PASS determines the needed $N1$ and $N2$, with this ratio, to obtain the desired power. An equivalent representation of the ratio, R , is

$$N2 = R * N1.$$

- **Enter percentage in Group 1, solve for $N1$ and $N2$**

For this choice, you set a value for the percentage of the total sample size that is in Group 1, and then PASS determines the needed $N1$ and $N2$ with this percentage to obtain the desired power.

$N2$ (Sample Size, Group 2)

This option is displayed if Group Allocation = "Enter $N2$, solve for $N1$ "

$N2$ is the number of items or individuals sampled from the Group 2 population.

$N2$ must be ≥ 2 . You can enter a single value or a series of values.

R (Group Sample Size Ratio)

This option is displayed only if Group Allocation = "Enter $R = N2/N1$, solve for $N1$ and $N2$."

R is the ratio of $N2$ to $N1$. That is,

$$R = N2 / N1.$$

Use this value to fix the ratio of $N2$ to $N1$ while solving for $N1$ and $N2$. Only sample size combinations with this ratio are considered.

$N2$ is related to $N1$ by the formula:

$$N2 = [R \times N1],$$

where the value $[Y]$ is the next integer $\geq Y$.

For example, setting $R = 2.0$ results in a Group 2 sample size that is double the sample size in Group 1 (e.g., $N1 = 10$ and $N2 = 20$, or $N1 = 50$ and $N2 = 100$).

R must be greater than 0. If $R < 1$, then $N2$ will be less than $N1$; if $R > 1$, then $N2$ will be greater than $N1$. You can enter a single or a series of values.

Percent in Group 1

This option is displayed only if Group Allocation = "Enter percentage in Group 1, solve for $N1$ and $N2$."

Use this value to fix the percentage of the total sample size allocated to Group 1 while solving for $N1$ and $N2$. Only sample size combinations with this Group 1 percentage are considered. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single or a series of values.

Sample Size (When Not Solving for Sample Size)

Group Allocation

Select the option that describes how individuals in the study will be allocated to Group 1 and to Group 2.

The options are

- **Equal ($N1 = N2$)**

This selection is used when you wish to have equal sample sizes in each group. A single per group sample size will be entered.

Mann-Whitney U or Wilcoxon Rank-Sum Tests (Simulation)

- **Enter N1 and N2 individually**

This choice permits you to enter different values for $N1$ and $N2$.

- **Enter N1 and R, where $N2 = R * N1$**

Choose this option to specify a value (or values) for $N1$, and obtain $N2$ as a ratio (multiple) of $N1$.

- **Enter total sample size and percentage in Group 1**

Choose this option to specify a value (or values) for the total sample size (N), obtain $N1$ as a percentage of N , and then $N2$ as $N - N1$.

Sample Size Per Group

This option is displayed only if Group Allocation = "Equal ($N1 = N2$)."

The Sample Size Per Group is the number of items or individuals sampled from each of the Group 1 and Group 2 populations. Since the sample sizes are the same in each group, this value is the value for $N1$, and also the value for $N2$.

The Sample Size Per Group must be ≥ 2 . You can enter a single value or a series of values.

N1 (Sample Size, Group 1)

*This option is displayed if Group Allocation = "Enter N1 and N2 individually" or "Enter N1 and R, where $N2 = R * N1$."*

$N1$ is the number of items or individuals sampled from the Group 1 population.

$N1$ must be ≥ 2 . You can enter a single value or a series of values.

N2 (Sample Size, Group 2)

This option is displayed only if Group Allocation = "Enter N1 and N2 individually."

$N2$ is the number of items or individuals sampled from the Group 2 population.

$N2$ must be ≥ 2 . You can enter a single value or a series of values.

R (Group Sample Size Ratio)

*This option is displayed only if Group Allocation = "Enter N1 and R, where $N2 = R * N1$."*

R is the ratio of $N2$ to $N1$. That is,

$$R = N2/N1$$

Use this value to obtain $N2$ as a multiple (or proportion) of $N1$.

$N2$ is calculated from $N1$ using the formula:

$$N2 = [R \times N1],$$

where the value $[Y]$ is the next integer $\geq Y$.

For example, setting $R = 2.0$ results in a Group 2 sample size that is double the sample size in Group 1.

R must be greater than 0. If $R < 1$, then $N2$ will be less than $N1$; if $R > 1$, then $N2$ will be greater than $N1$. You can enter a single value or a series of values.

Total Sample Size (N)

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This is the total sample size, or the sum of the two group sample sizes. This value, along with the percentage of the total sample size in Group 1, implicitly defines $N1$ and $N2$.

The total sample size must be greater than one, but practically, must be greater than 3, since each group sample size needs to be at least 2. You can enter a single value or a series of values.

Mann-Whitney U or Wilcoxon Rank-Sum Tests (Simulation)

Percent in Group 1

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This value fixes the percentage of the total sample size allocated to Group 1. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single value or a series of values.

Effect Size

Input Type

Select which type of input method to use to specify the simulated distributions and the effect size.

The choices are

- **Simple (Means)**
Provides the simplest possible input for specifying the distributions and effect size using means. This entry type is restricted to a set of distributions that can be specified using a mean and standard deviation. All simulated values come from one named distribution.
- **Simple (Differences)**
Provides the simplest possible input for specifying the distributions and effect size using differences. This entry type is restricted to a set of distributions that can be specified using a mean and standard deviation. All simulated values come from one named distribution.
- **General**
Provides all options for specifying the distributions and parameters. This allows for complete flexibility in specifying the simulated distributions but is more complicated.

Effect Size – Simulated Distribution

Displayed when Input Type = Simple

Distribution to Simulate

Specify the distribution that represents the type of data you expect from your study.

The possible distributions are Beta, Exponential, Gamma, Gumbel, Laplace, Logistic, Lognormal, Normal, Poisson, TukeyGH, Uniform, Weibull.

All these distributions can be specified with a mean and standard deviation. The Beta and TukeyGH distributions each require two additional parameters. The Exponential and Poisson distributions require only the mean to be specified since the standard deviation can be computed from the mean.

Normal vs. TukeyGH

Tukey's distribution can be used to generate values that are nearly Normal, with departure from normality controlled by entering skewness (G) and kurtosis (H) parameters. Tukey's distribution with $G = H = 0$ is equivalent to the Normal distribution.

Mann-Whitney U or Wilcoxon Rank-Sum Tests (Simulation)

Effect Size – Group 1 (Treatment)

Displayed when Input Type = Simple (Means)

 $\mu_{1.0}$ (Null Group 1 Mean)

Enter a value for the mean of group 1 under the null hypothesis. The null-hypothesized difference, δ_0 , is calculated using $\mu_{1.0}$ as

$$\delta_0 = \mu_{1.0} - \mu_2.$$

This value can be set to perform inequality, non-inferiority, and superiority by a margin tests. The value determines the null hypothesis of the test.

$\mu_{1.0}$ can be any value (positive, negative, or zero). You can enter a single value such as 0 or a series of values such as 0 1 2 or 0 to 2 by 1. When a series of values is entered, **PASS** will generate a separate calculation result for each value of the series.

 $\mu_{1.1}$ (Actual Group 1 Mean)

Enter a value for the actual mean of group 1 at which power and sample size are calculated. The actual difference, δ_1 , is calculated using $\mu_{1.1}$ as

$$\delta_1 = \mu_{1.1} - \mu_2.$$

$\mu_{1.1}$ can be any value (positive, negative, or zero). You can enter a single value such as 0 or a series of values such as 0 1 2 or 0 to 2 by 1. When a series of values is entered, **PASS** will generate a separate calculation result for each value of the series.

Effect Size – Differences ($\delta = \mu_1 - \mu_2$)

Displayed when Input Type = Simple (Differences)

 δ_0 (Null Difference)

Enter a value for the mean difference ($\mu_1 - \mu_2$) under the null hypothesis. δ_0 is calculated using $\mu_{1.0}$ and μ_2 as

$$\delta_0 = \mu_{1.0} - \mu_2.$$

This value can be set to perform inequality, non-inferiority, and superiority by a margin tests. The value determines the null hypothesis of the test.

δ_0 can be any value (positive, negative, or zero). You can enter a single value such as 0 or a series of values such as 0 1 2 or 0 to 2 by 1. When a series of values is entered, **PASS** will generate a separate calculation result for each value of the series.

 δ_1 (Actual Difference)

Enter a value for the actual mean difference ($\mu_1 - \mu_2$) at which power and sample size are to be calculated. δ_1 is calculated using $\mu_{1.1}$ and μ_2 as

$$\delta_1 = \mu_{1.1} - \mu_2.$$

δ_0 can be any value (positive, negative, or zero). You can enter a single value such as 0 or a series of values such as 0 1 2 or 0 to 2 by 1. When a series of values is entered, **PASS** will generate a separate calculation result for each value of the series.

Mann-Whitney U or Wilcoxon Rank-Sum Tests (Simulation)

Effect Size – Group 2 (Reference)

M2 (Group 2 Mean)

Enter a value for the mean of group 2, the reference group, under both the null and alternative hypotheses. The null and actual differences (δ_0 and δ_1) are calculated using μ_2 as

$$\delta_0 = \mu_{1.0} - \mu_2$$

$$\delta_1 = \mu_{1.1} - \mu_2.$$

δ_0 is the mean difference under the null hypothesis. δ_1 is the actual mean difference at which power and sample size are calculated.

μ_2 can be any value (positive, negative, or zero). You can enter a single value such as 10 or a series of values such as 10 20 30 or 5 to 50 by 5. When a series of values is entered, PASS will generate a separate calculation result for each value of the series.

Effect Size – Standard Deviations

Displayed when Input Type = Simple

Group Standard Deviations

Specify whether to simulate populations with group standard deviations that are equal or unequal.

σ (Standard Deviation)

Displayed when Group Standard Deviations = Equal

The standard deviation entered here is the assumed standard deviation for both the Group 1 population and the Group 2 population.

σ must be a positive number. You can enter a single value such as 5 or a series of values such as 1 3 5 7 9 or 1 to 10 by 1. When a series of values is entered, **PASS** will generate a separate calculation result for each value of the series.

σ_1 and σ_2 (Standard Deviation)

Displayed when Group Standard Deviations = Unequal

The standard deviations entered here are the assumed standard deviations for the Group 1 and Group 2 populations.

σ_1 and σ_2 must be positive numbers. You can enter a single value such as 5 or a series of values such as 1 3 5 7 9 or 1 to 10 by 1. When a series of values is entered, **PASS** will generate a separate calculation result for each value of the series.

Effect Size – Simulated Distributions

Displayed when Input Type = General

Group 1 & 2 Distributions|H0 (Null)

Specify the distributions of groups 1 and 2 to be used in the simulations for finding the actual significance level. These distributions also specify the value of the mean under the null hypothesis, H0. Usually, these two distributions will be identical. However, if you are planning a non-inferiority test, the means will be different.

The parameters of the distribution can be specified using numbers or letters. If letters are used, their values are specified in the boxes below. The value “M0” is reserved for the value of the mean under the null hypothesis.

Mann-Whitney U or Wilcoxon Rank-Sum Tests (Simulation)

Following is a list of the distributions that are available, and the syntax used to specify them. Each of the parameters should be replaced with a number or parameter name.

Distributions with Common Parameters

Beta(Shape1, Shape2, Min, Max)

Binomial(P, N)

Cauchy(Mean, Scale)

Constant(Value)

Exponential(Mean)

Gamma(Shape, Scale)

Gumbel(Location, Scale)

Laplace(Location, Scale)

Logistic(Location, Scale)

Lognormal(Mu, Sigma)

Multinomial(P1, P2, P3, ..., Pk)

Normal(Mean, Sigma)

Poisson(Mean)

TukeyGH(Mu, S, G, H)

Uniform(Min, Max)

Weibull(Shape, Scale)

Distributions with Mean and SD Parameters

BetaMS(Mean, SD, Min, Max)

BinomialMS(Mean, N)

GammaMS(Mean, SD)

GumbelMS(Mean, SD)

LaplaceMS(Mean, SD)

LogisticMS(Mean, SD)

LognormalMS(Mean, SD)

UniformMS(Mean, SD)

WeibullMS(Mean, SD)

Details of writing mixture distributions, combined distributions, and compound distributions are found in the chapter on *Data Simulation* and will not be repeated here.

Finding the Value of the Mean of a Specified Distribution

The mean of a distribution created as a linear combination of other distributions is found by applying the linear combination to the individual means. However, the mean of a distribution created by multiplying or dividing other distributions is not necessarily equal to applying the same function to the individual means. For example, the mean of $4 \text{ Normal}(4, 5) + 2 \text{ Normal}(5, 6)$ is $4*4 + 2*5 = 26$, but the mean of $4 \text{ Normal}(4, 5) * 2 \text{ Normal}(5, 6)$ is not exactly $4*4*2*5 = 160$ (although it is close).

Mann-Whitney U or Wilcoxon Rank-Sum Tests (Simulation)

Group 1 & 2 Distributions|H1 (Alternative)

Specify the distributions of groups 1 and 2 under the alternative hypothesis, H1. The difference between the means of these two distributions will be the actual difference specified under H1.

The parameters of the distribution can be specified using numbers or letters. If letters are used, their values are specified in the boxes below. The value "M1" is reserved for the value of the mean under the alternative hypothesis.

A list of the distributions that are available and the syntax used to specify them is given above.

Effect Size – Distribution Parameters

Displayed when Input Type = General

M0 (Mean|H0)

These values are substituted for *M0* in the distribution specifications given above. *M0* is intended to be the value of the mean hypothesized by the null hypothesis, H0.

You can enter a list of values using the syntax *0 1 2 3* or *0 to 3 by 1*.

M1 (Mean|H1)

These values are substituted for *M1* in the distribution specifications given above. Although it can be used wherever you want, *M1* is intended to be the value of the mean hypothesized by the alternative hypothesis, H1.

You can enter a list of values using the syntax *0 1 2 3* or *0 to 3 by 1*.

Parameter Values (S, A, B)

Enter the numeric value(s) of the parameters listed above. These values are substituted for the corresponding letter in all four distribution specifications.

You can enter a list of values for each letter using the syntax *0 1 2 3* or *0 to 3 by 1*.

You can also change the letter that is used as the name of this parameter using the pull-down menu to the side.

Options Tab

The Options tab contains simulation options.

Random Numbers

Random Number Pool Size

This is the size of the pool of random values from which the random samples will be drawn. Pools should be at least the maximum of 10,000 and twice the number of simulations. You can enter *Automatic* and an appropriate value will be calculated.

If you do not want to draw numbers from a pool, enter 0 here.

Example 1 – Finding the Sample Size

Researchers wish to compare two types of local anesthesia to determine whether there is a difference in time to loss of pain. Subjects will be randomized to treatment, the treatment will be administered, and the time to loss of pain measured. The researchers would like to generate a sample size for the study with 90% power to reject the null hypothesis of equal loss-of-pain time if the true difference is at least 3 minutes. How many participants are needed to achieve 90% power at significance levels of 0.01 and 0.05?

Past experiments of this type have had standard deviations in the range of 1 to 5 minutes. It is anticipated that the standard deviation of the two groups will be equal.

It is unknown which treatment has lower time to loss of pain, so a two-sided test will be used.

The researchers will be performing a Mann-Whitney-Wilcoxon test instead of the t-test, because it is anticipated that the distribution of the two populations is not Normal. The researchers assume that the distribution shape that most closely resembles the two populations is Tukey's Lambda distribution with Skewness value 0.12 and Elongation 0.07 (Tukey's Lambda distribution adjusts the Normal distribution by a specified Skewness factor and Kurtosis factor – you can examine the distribution with the Data Simulator tool, in the Tools menu).

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Mann-Whitney U or Wilcoxon Rank-Sum Tests (Simulation)** procedure window. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Sample Size
Alternative Hypothesis	Two-Sided (H1: $\mu_1 - \mu_2 \neq \delta_0$)
Simulations	2000
Power	0.90
Alpha	0.01 0.05
Group Allocation	Equal (N1 = N2)
Input Type	Simple (Differences)
Distribution to Simulate	TukeyGH
G (Skewness)	0.12
H (Kurtosis)	0.07
δ_0 (Null Difference)	0
δ_1 (Actual Difference)	3
μ_2 (Group 2 Mean)	0
Group Standard Deviations	Equal
σ (Standard Deviation)	1 to 5 by 1

Mann-Whitney U or Wilcoxon Rank-Sum Tests (Simulation)

Annotated Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Numeric Results

Hypotheses: $H_0: \delta = \delta_0$ vs. $H_1: \delta \neq \delta_0$

Test Statistic: Mann-Whitney U or Wilcoxon Rank-Sum Test

Simulated Distribution: TukeyGH

Power	N1	N2	N	$\mu_{1.0}$	$\mu_{1.1}$	μ_2	δ_0	δ_1	σ	Target Alpha	Actual Alpha	G H
0.9050 (0.0129) [0.8921 0.9179]	7	7	14	0.0	3.0	0.0	0.0	3.0	1.0	0.010	0.005 (0.003) [0.002 0.009]	0.12 0.07
0.9335 (0.0109) [0.9226 0.9444]	16	16	32	0.0	3.0	0.0	0.0	3.0	2.0	0.010	0.008 (0.004) [0.004 0.012]	0.12 0.07
0.9060 (0.0128) [0.8932 0.9188]	30	30	60	0.0	3.0	0.0	0.0	3.0	3.0	0.010	0.009 (0.004) [0.005 0.014]	0.12 0.07
0.9060 (0.0128) [0.8932 0.9188]	50	50	100	0.0	3.0	0.0	0.0	3.0	4.0	0.010	0.011 (0.005) [0.006 0.016]	0.12 0.07
0.9160 (0.0122) [0.9038 0.9282]	79	79	158	0.0	3.0	0.0	0.0	3.0	5.0	0.010	0.008 (0.004) [0.004 0.013]	0.12 0.07
0.9345 (0.0108) [0.9237 0.9453]	5	5	10	0.0	3.0	0.0	0.0	3.0	1.0	0.050	0.038 (0.008) [0.029 0.046]	0.12 0.07
0.9150 (0.0122) [0.9028 0.9272]	11	11	22	0.0	3.0	0.0	0.0	3.0	2.0	0.050	0.054 (0.010) [0.044 0.064]	0.12 0.07
0.9060 (0.0128) [0.8932 0.9188]	21	21	42	0.0	3.0	0.0	0.0	3.0	3.0	0.050	0.050 (0.010) [0.040 0.060]	0.12 0.07
0.9125 (0.0124) [0.9001 0.9249]	36	36	72	0.0	3.0	0.0	0.0	3.0	4.0	0.050	0.052 (0.010) [0.042 0.062]	0.12 0.07
0.9025 (0.0130) [0.8895 0.9155]	53	53	106	0.0	3.0	0.0	0.0	3.0	5.0	0.050	0.051 (0.010) [0.041 0.060]	0.12 0.07

Notes:

Pool Size: 10000. Simulations: 2000. Run Time: 26.69 seconds.

References

Chow, S.C., Shao, J., Wang, H., and Lohknygina, Y. 2018. Sample Size Calculations in Clinical Research, Third Edition. Taylor & Francis/CRC. Boca Raton, Florida.

Devroye, Luc. 1986. Non-Uniform Random Variate Generation. Springer-Verlag. New York.

Matsumoto, M. and Nishimura, T. 1998. 'Mersenne twister: A 623-dimensionally equidistributed uniform pseudorandom number generator.' ACM Trans. On Modeling and Computer Simulations.

Zar, Jerrold H. 1984. Biostatistical Analysis (Second Edition). Prentice-Hall. Englewood Cliffs, New Jersey.

Report Definitions

Power is the probability of rejecting a false null hypothesis. Second Row: (Power Inc.) [95% LCL and UCL Power].

N1 and N2 are the number of items sampled from each population.

$N = N1 + N2$ is the total sample size.

$\mu_{1.0}$ is the mean of group 1 under the null hypothesis.

$\mu_{1.1}$ is the actual mean of group 1 at which power and sample size are calculated.

μ_2 is the mean of group 2 under both the null and alternative hypotheses.

$\delta_0 = \mu_{1.0} - \mu_2$ is the mean difference under the null hypothesis.

$\delta_1 = \mu_{1.1} - \mu_2$ is the actual mean difference at which power and sample size are calculated.

σ is the assumed population standard deviation for both groups 1 and 2.

Target Alpha is the probability of rejecting a true null hypothesis. It is set by the user.

Actual Alpha is the alpha level that was actually achieved by the experiment. Second Row: (Alpha Inc.) [95% LCL and UCL Alpha].

Additional columns are the other parameters required to specify the distributions.

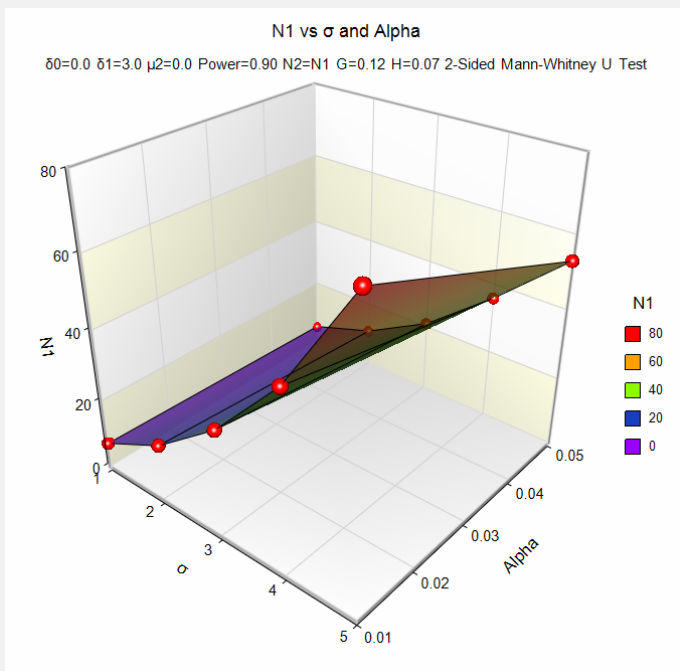
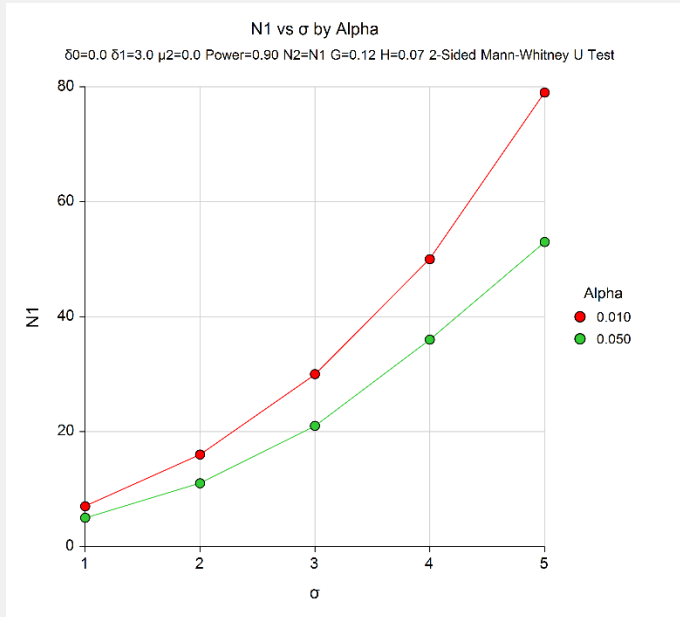
Mann-Whitney U or Wilcoxon Rank-Sum Tests (Simulation)

Summary Statements

Group sample sizes of 7 and 7 achieve 91% power to detect a mean difference of 3.0 when the null-hypothesized mean difference is 0.0 using a two-sided Mann-Whitney U or Wilcoxon Rank-Sum Test at a 0.010 significance level (alpha) with a standard deviation of 1.0 for both groups 1 and 2. These results are based on 2000 Monte Carlo samples from the TukeyGH distribution with $G = 0.12$, $H = 0.07$.

These reports show the values of each of the parameters, one scenario per row. Since these results are based on simulation, they will vary from one calculation to the next.

Plots Section



These plots show the relationship between the standard deviation and sample size for the two alpha levels.

Mann-Whitney U or Wilcoxon Rank-Sum Tests (Simulation)

Example 2 – Comparing the Power to the T-Test for Normal Data

Suppose a new corn fertilizer is to be compared to a current fertilizer. The researchers need only show that there is difference (increase) in yield with the new fertilizer. With 90 plots available, they would like to examine the power of the test if the improvement in yield is at least 10 lbs.

Researchers plan to use a one-sided test with alpha equal to 0.05. Previous studies indicate the standard deviation for plot yield to be 25 lbs. The distribution of plot yield values is unknown, so the researchers would like to see the loss in power if the distribution turns out to be Normal and the Mann-Whitney-Wilcoxon test is used rather than the standard t-test. The power for this scenario with the standard t-test is 0.594.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Mann-Whitney U or Wilcoxon Rank-Sum Tests (Simulation)** procedure window. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Solve For	Power
Alternative Hypothesis	One-Sided (H1: $\mu_1 - \mu_2 > \delta_0$)
Simulations	100000
Alpha	0.05
Group Allocation	Equal (N1 = N2)
Sample Size Per Group	45
Input Type	Simple (Differences)
Distribution to Simulate	Normal
δ_0 (Null Difference)	0
δ_1 (Actual Difference)	10
μ_2 (Group 2 Mean)	0
Group Standard Deviations	Equal
σ (Standard Deviation)	25

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Numeric Results											
Hypotheses: $H_0: \delta \leq \delta_0$ vs. $H_1: \delta > \delta_0$											
Test Statistic: Mann-Whitney U or Wilcoxon Rank-Sum Test											
Simulated Distribution: Normal											
Power	N1	N2	N	$\mu_{1.0}$	$\mu_{1.1}$	μ_2	δ_0	δ_1	σ	Target Alpha	Actual Alpha
0.5828 (0.0031) [0.5798 0.5859]	45	45	90	0.0	10.0	0.0	0.0	10.0	25.0	0.050	0.050 (0.001) [0.049 0.051]

The power of the Mann-Whitney test in this scenario is 0.5828 (0.5798 0.5859). This power is only slightly less than the power of the t-test (0.594) for the corresponding scenario.

Example 3 – Validation using Zhao, Rahardja, and Qu (2008)

Zhou, Rahardja, and Qu (2008) page 467 present a table of sample sizes and powers based on their formulas and simulation results. In the example, the data considered follow a multinomial distribution, which heavily increases the likelihood of ties. For the first set of scenarios, the power is 80%, the significance level is 0.05, and Group 2 is assumed to have 53% of the total number of individuals in the sample. The ninth line has the two different multinomial distributions defined as $M(0.66, 0.15, 0.19)$ for Group 1 and $M(0.55, 0.15, 0.30)$ for Group 2. The total sample size is estimated at 502.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Mann-Whitney U or Wilcoxon Rank-Sum Tests (Simulation)** procedure window. You may then make the appropriate entries as listed below, or open **Example 3** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Sample Size
Alternative Hypothesis	Two-Sided ($H_1: \mu_1 - \mu_2 \neq \delta_0$)
Simulations	10000
Power	0.80
Alpha	0.05
Group Allocation	Enter $R = N_2/N_1$, solve for N_1 and N_2
R	1.12766
Input Type	General
Group 1 Distribution H0	Multinomial(0.66 0.15 0.19)
Group 1 Distribution H1	Multinomial(0.66 0.15 0.19)
Group 2 Distribution H0	Multinomial(0.66 0.15 0.19)
Group 2 Distribution H1	Multinomial(0.55 0.15 0.30)

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Numeric Results							
Hypotheses: $H_0: \delta = \delta_0$ vs. $H_1: \delta \neq \delta_0$							
Test Statistic: Mann-Whitney U or Wilcoxon Rank-Sum Test							
Group 1 Distributions: H0: Multinomial(0.66 0.15 0.19) H1: Multinomial(0.66 0.15 0.19)							
Group 2 Distributions: H0: Multinomial(0.66 0.15 0.19) H1: Multinomial(0.55 0.15 0.30)							
Power	N1	N2	N	δ_0	δ_1	Target Alpha	Actual Alpha
0.8092 (0.0077)	236	266	502	0.0	-0.2	0.050	0.051 (0.004)
	[0.8015 0.8169]						[0.047 0.056]

The estimated sample size is 502, which matches the expected result exactly.