Chapter 414

One-Sample Z-Tests

Introduction

The one-sample *z*-test is used to test whether the mean of a population is greater than, less than, or not equal to a specific value. Because the standard normal distribution is used to calculate critical values for the test, this test is often called the one-sample *z*-test. The *z*-test assumes that the population standard deviation is known.

Other PASS Procedures for Testing One Mean or Median

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 one-sample procedures available in **PASS**–the One-Sample T-Tests and the nonparametric Wilcoxon Signed-Rank Test procedures. The methods, statistics, and assumptions for those procedures are described in the associated chapters.

If you wish to show that the mean of a population is larger (or smaller) than a reference value 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.

The Statistical Hypotheses

In the usual *z*-test setting, the null (H_0) and alternative (H_1) hypotheses for two-sided tests are defined as

$$H_0: \mu = \mu_0 \quad \text{versus} \quad H_1: \mu \neq \mu_0$$

Rejecting H_0 implies that the mean is not equal to the value μ_0. The hypotheses for one-sided upper-tail tests are

$$H_0: \mu \leq \mu_0 \quad \text{versus} \quad H_1: \mu > \mu_0$$

Rejecting H_0 implies that the mean is larger than the value μ_0. This test is called an *upper-tail test* because H_0 is rejected in samples in which the sample mean is larger than μ_0.

The *lower-tail test* is

$$H_0: \mu \geq \mu_0 \quad \text{versus} \quad H_1: \mu < \mu_0$$

It will be convenient to adopt the following specialize notation for the discussion of these tests.

Parameter	**PASS Input/Output**	**Interpretation**
μ	μ	*Population mean*. If the data are paired differences, this is the mean of those differences. This parameter will be estimated by the study.
μ_1	$\mu 1$	*Actual population mean at which power is calculated*. This is the assumed population mean used in all calculations.
μ_0	$\mu 0$	*Reference value*. Usually, this is the mean of a reference population. If the data are paired differences, this is the hypothesized value of the mean difference.
δ	δ	*Population difference*. This is the value of $\mu - \mu_0$, the difference between the population mean and the reference value. This parameter will be estimated by the study.
δ_1	$\delta 1$	*Actual difference at which power is calculated*. This is the value of $\mu_1 - \mu_0$, the assumed difference between the mean and the reference value for power calculations.

Assumptions for One-Sample Tests

This section describes the assumptions that are made when you use one of the one-sample tests. The key assumption relates to normality or non-normality of the data. One of the reasons for the popularity of the *t*-test is its robustness in the face of assumption violation. However, if an assumption is not met even approximately, the significance levels and the power of the *t*-test are invalidated. Unfortunately, in practice it often happens that several assumptions are not met. This makes matters even worse! Hence, take the steps to check the assumptions before you make important decisions based on these tests.

One-Sample Z-Test Assumptions

The assumptions of the one-sample *z*-test are:

1. The data are continuous (not discrete).
2. The data follow the normal probability distribution.
3. The sample is a simple random sample from its population. Each individual in the population has an equal probability of being selected in the sample.
4. The population standard deviation is known.

One-Sample T-Test Assumptions

The assumptions of the one-sample *t*-test are:

1. The data are continuous (not discrete).
2. The data follow the normal probability distribution.
3. The sample is a simple random sample from its population. Each individual in the population has an equal probability of being selected in the sample.

Paired T-Test Assumptions

The assumptions of the paired *t*-test are:

1. The data are continuous (not discrete).
2. The data, i.e., the differences for the matched pairs, follow a normal probability distribution.
3. The sample of pairs is a simple random sample from its population. Each individual in the population has an equal probability of being selected in the sample.

Wilcoxon Signed-Rank Test Assumptions

The assumptions of the Wilcoxon signed-rank test are as follows (note that the difference is between a data value and the hypothesized median or between the two data values of a pair):

1. The differences are continuous (not discrete).
2. The distribution of each difference is symmetric.
3. The differences are mutually independent.
4. The differences all have the same median.
5. The measurement scale is at least interval.

Limitations

There are few limitations when using these tests. Sample sizes may range from a few to several hundred. If your data are discrete with at least five unique values, you can often ignore the continuous variable assumption. Perhaps the greatest restriction is that your data come from a random sample of the population. If you do not have a random sample, your significance levels will probably be incorrect.

One-Sample Z-Test Statistic

The one-sample *z*-test assumes that the data are a simple random sample from a population of normally distributed values that all have the same mean and variance (known). This assumption implies that the data are continuous, and their distribution is symmetric. The calculation of the *z*-test proceeds as follows

$$z = \frac{\bar{X} - A}{\sigma / \sqrt{n}}$$

where

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

and A is the value of the mean hypothesized by the null hypothesis that incorporates both μ_0 and M_{NI}.

The significance of the test statistic is determined by computing the p-value. If this p-value is less than a specified level (usually 0.05), the hypothesis is rejected. Otherwise, no conclusion can be reached.

Population Size

This is the number of subjects in the population. Usually, you assume that samples are drawn from a very large (infinite) population. Occasionally, however, situations arise in which the population of interest is of limited size. In these cases, appropriate adjustments must be made.

When a finite population size is specified, the standard deviation is reduced according to the formula:

$$\sigma_1^2 = \left(1 - \frac{n}{N}\right)\sigma^2$$

where *n* is the sample size, *N* is the population size, σ is the original standard deviation, and σ_1 is the new standard deviation.

The quantity *n/N* is often called the sampling fraction. The quantity $\left(1 - \frac{n}{N}\right)$ is called the *finite population correction factor*.

Power Calculation for the One-Sample Z-Test

When the standard deviation is known, the power is calculated as follows for a directional alternative (one-tailed test) in which $\mu_1 > \mu_0$.

1. Find z_α such that $1 - \Phi(z_\alpha) = \alpha$, where $\Phi(x)$ is the area to the left of *x* under the standardized normal curve.
2. Calculate: $X_1 = \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$.
3. Calculate: $z_1 = \frac{X_1 - \mu_1}{\frac{\sigma}{\sqrt{n}}}$.
4. Power = $1 - \Phi(z_1)$.

Example 1 – Power after a Study

This example will cover the situation in which you are calculating the power of a *z*-test on data that have already been collected and analyzed. For example, you might be playing the role of a reviewer, looking at the power of a *z*-test from a study you are reviewing. In this case, you would not vary the means or sample size since they are given by the experiment. Instead, you investigate the power of the significance tests. You might look at the impact of different alpha values on the power.

Suppose an experiment involving 100 individuals yields the following summary statistics:

Hypothesized mean (μ0)	100.0
Sample mean (μ1)	110.0
Sample size	100

Given the above data, analyze the power of a *z*-test which tests the hypothesis that the population mean is 100 versus the alternative hypothesis that the population mean is 110. Consider the power at significance levels 0.01, 0.05, 0.10 and sample sizes 20 to 120 by 20. The standard deviation is known to be 40.

Note that we have set *μ1* equal to the sample mean. In this case, we are studying the power of the *z*-test for a mean difference the size of that found in the experimental data.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For .... **Power**
Alternative Hypothesis .... **Two-Sided (H1: μ ≠ μ0)**
Population Size .... **Infinite**
Alpha.. **0.01 0.05 0.10**
N (Sample Size)... **20 to 120 by 20**
μ0 (Null or Baseline Mean) **100**
μ1 (Actual Mean) **110**
σ (Standard Deviation).................................. **40**

Output

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

Numeric Reports

Numeric Results

Solve For: Power
Hypotheses: H0: μ = μ0 vs. H1: μ ≠ μ0

Power	Sample Size N	Mean Null μ0	Mean Actual μ1	Mean Difference μ1 - μ0	Standard Deviation σ	Effect Size	Alpha
0.07256	20	100	110	10	40	0.25	0.01
0.15996	40	100	110	10	40	0.25	0.01
0.26130	60	100	110	10	40	0.25	0.01
0.36702	80	100	110	10	40	0.25	0.01
0.46978	100	100	110	10	40	0.25	0.01
0.56466	120	100	110	10	40	0.25	0.01
0.20096	20	100	110	10	40	0.25	0.05
0.35261	40	100	110	10	40	0.25	0.05
0.49069	60	100	110	10	40	0.25	0.05
0.60878	80	100	110	10	40	0.25	0.05
0.70542	100	100	110	10	40	0.25	0.05
0.78191	120	100	110	10	40	0.25	0.05
0.30202	20	100	110	10	40	0.25	0.10
0.47523	40	100	110	10	40	0.25	0.10
0.61489	60	100	110	10	40	0.25	0.10
0.72286	80	100	110	10	40	0.25	0.10
0.80378	100	100	110	10	40	0.25	0.10
0.86298	120	100	110	10	40	0.25	0.10

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N The size of the sample drawn from the population.
μ0 The value of the population mean under the null hypothesis.
μ1 The actual value of the population mean at which power and sample size are calculated.
μ1 - μ0 The difference between the actual and null means.
σ The standard deviation of the population. It measures the variability in the population.
Effect Size The relative magnitude of the effect. Effect Size = |μ1 - μ0|/σ.
Alpha The probability of rejecting a true null hypothesis.

Summary Statements

A single-group design will be used to test whether the mean is different from 100 (H0: μ = 100 versus H1: μ ≠ 100). The comparison will be made using a two-sided, one-sample Z-test, with a Type I error rate (α) of 0.01. The (known) standard deviation is assumed to be 40. To detect a mean of 110 (corresponding to a difference of 10) with a sample size of 20, the power is 0.07256.

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	20	25	5
20%	40	50	10
20%	60	75	15
20%	80	100	20
20%	100	125	25
20%	120	150	30

Dropout Rate The percentage of subjects (or items) that are expected to be lost at random during the course of the study and for whom no response data will be collected (i.e., will be treated as "missing"). Abbreviated as DR.
N The evaluable sample size at which power is computed (as entered by the user). If N subjects are evaluated out of the N' subjects that are enrolled in the study, the design will achieve the stated power.
N' The total number of subjects that should be enrolled in the study in order to obtain N evaluable subjects, based on the assumed dropout rate. N' is calculated by inflating N using the formula N' = N / (1 - DR), with N' always rounded up. (See Julious, S.A. (2010) pages 52-53, or Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. (2018) pages 32-33.)
D The expected number of dropouts. D = N' - N.

Dropout Summary Statements

Anticipating a 20% dropout rate, 25 subjects should be enrolled to obtain a final sample size of 20 subjects.

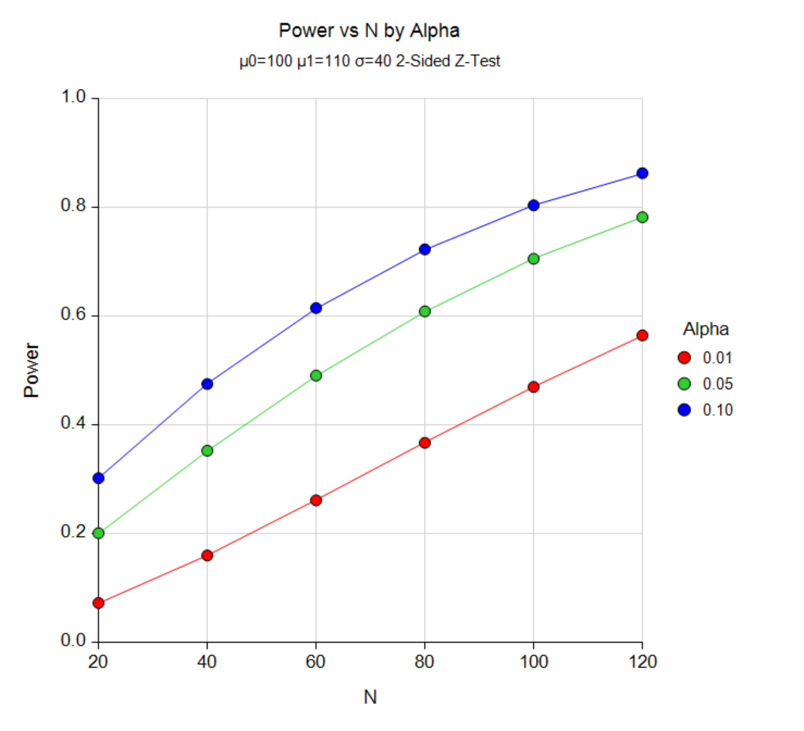
References

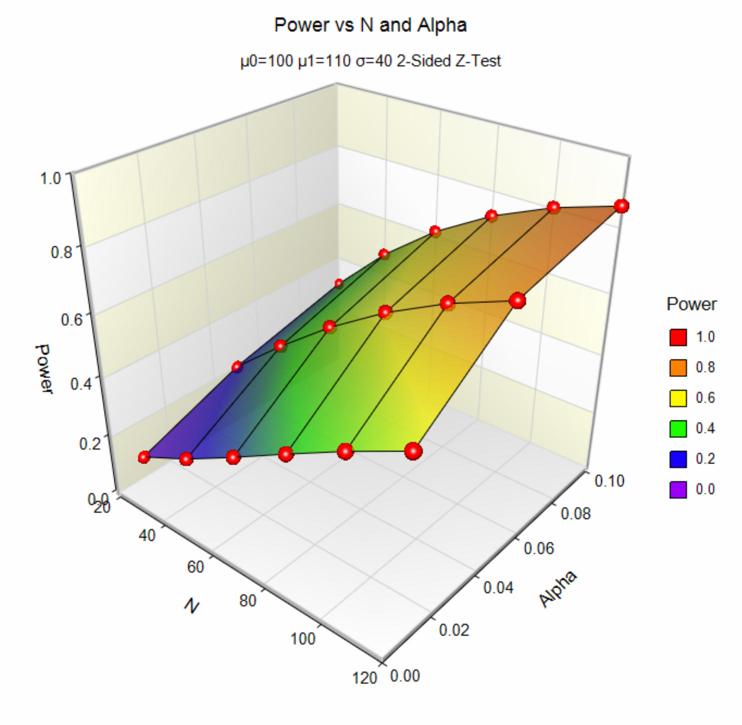
Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. 2018. Sample Size Calculations in Clinical Research, Third Edition. Taylor & Francis/CRC. Boca Raton, Florida.
Machin, D., Campbell, M., Fayers, P., and Pinol, A. 1997. Sample Size Tables for Clinical Studies, 2nd Edition. Blackwell Science. Malden, MA.
Zar, Jerrold H. 1984. Biostatistical Analysis (Second Edition). Prentice-Hall. Englewood Cliffs, New Jersey.

This report shows the values of each of the parameters, one scenario per row. The values of power and beta were calculated from the other parameters.

Plots Section

Plots





These plots show the relationship between sample size and power for various values of alpha.

Example 2 – Finding the Sample Size

This example will consider the situation in which you are planning a study that will use the one-sample *z*-test and want to determine an appropriate sample size. This example is more subjective than the first because you now have to obtain estimates of all the parameters. In the first example, these estimates were provided by the data.

In studying deaths from SIDS (Sudden Infant Death Syndrome), one hypothesis put forward is that infants dying of SIDS weigh less than normal at birth. Suppose the average birth weight of infants is 3300 grams with a known standard deviation of 663 grams. Use an alpha of 0.05 and power of both 0.80 and 0.90. How large a sample of SIDS infants will be needed to detect a drop in average weight of 25%? Of 10%? Of 5%? Note that applying these percentages to the average weight of 3300 yields 2475, 2970, and 3135.

Although a one-sided hypothesis is being considered, sample size estimates will assume a two-sided alternative to keep the research design in line with other studies.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For	**Sample Size**
Alternative Hypothesis	**Two-Sided (H1: μ ≠ μ0)**
Population Size	**Infinite**
Power	**0.80 0.90**
Alpha	**0.05**
μ0 (Null or Baseline Mean)	**3300**
μ1 (Actual Mean)	**2475 2970 3135**
σ (Standard Deviation)	**663**

Output

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

Numeric Results

Solve For: Sample Size
Hypotheses: H0: μ = μ0 vs. H1: μ ≠ μ0

		Mean					
Power	Sample Size N	Null μ0	Actual μ1	Difference μ1 - μ0	Standard Deviation σ	Effect Size	Alpha
0.86171	6	3300	2475	-825	663	1.244	0.05
0.90861	7	3300	2475	-825	663	1.244	0.05
0.80391	32	3300	2970	-330	663	0.498	0.05
0.90387	43	3300	2970	-330	663	0.498	0.05
0.80085	127	3300	3135	-165	663	0.249	0.05
0.90058	170	3300	3135	-165	663	0.249	0.05

This report shows the values of each of the parameters, one scenario per row. Since there were three values of μ1 and two values of power, there are a total of six rows in the report.

We were solving for the sample size, *N*. Notice that the increase in sample size seems to be most directly related to the difference between the two means. The difference in beta values does not seem to be as influential, especially at the smaller sample sizes.

Note that even though we set the power values at 0.8 and 0.9, these are not the power values that were achieved. This happens because *N* can only take on integer values. The program selects the first value of *N* that gives at least the values of alpha and power that were desired.

Example 3 – Finding the Minimum Detectable Difference

This example will consider the situation in which you want to determine how small of a difference between the two means can be detected by the *z*-test with specified values of the other parameters.

Continuing with the previous example, suppose about 50 SIDS deaths occur in a particular area per year. Using 50 as the sample size, 0.05 as alpha, and 0.80 as power, how large of a difference between the means is detectable?

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 3** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For .... **μ1 (Search < μ0)**
Alternative Hypothesis .... **Two-Sided (H1: μ ≠ μ0)**
Population Size .... **Infinite**
Power.. **0.80**
Alpha... **0.05**
N (Sample Size).. **50**
μ0 (Null or Baseline Mean) **3300**
σ (Standard Deviation).................................. **663**

Output

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

Numeric Results

Solve For: μ1 (Search < μ0)
Hypotheses: H0: μ = μ0 vs. H1: μ ≠ μ0

Power	Sample Size N	Mean: Null μ0	Mean: Actual μ1	Mean: Difference μ1 - μ0	Standard Deviation σ	Effect Size	Alpha
0.8	50	3300	3037.3	-262.7	663	0.396	0.05

With a sample of 50, a difference of 3037.3 - 3300 = -262.7 would be detectable.

Example 4 – Validation using Chow, Shao, Wang, and Lokhnygina (2018)

Chow, Shao, Wang, and Lokhnygina (2018) presents an example on pages 45 and 46 of a two-sided one-sample *z*-test sample size calculation in which μ0 = 1.5, μ1 = 2.0, σ = 1.0, alpha = 0.05, and power = 0.80. They obtain a sample size of 32.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 4** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For ....**Sample Size**
Alternative Hypothesis ....**Two-Sided (H1: μ ≠ μ0)**
Population Size..**Infinite**
Power..**0.80**
Alpha..**0.05**
μ0 (Null or Baseline Mean) **1.5**
μ1 (Actual Mean) **2**
σ (Standard Deviation).................................**1**

Output

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

Numeric Results

Solve For: Sample Size
Hypotheses: H0: μ = μ0 vs. H1: μ ≠ μ0

Power	Sample Size N	Mean Null μ0	Mean Actual μ1	Mean Difference μ1 - μ0	Standard Deviation σ	Effect Size	Alpha
0.80743	32	1.5	2	0.5	1	0.5	0.05

The sample size of 32 matches Chow, Shao, Wang, and Lokhnygina (2018) exactly.

Example 5 – Validation using Machin (1997)

Machin, Campbell, Fayers, and Pinol (1997) page 37 presents an example of a one-sample *t*-test in which μ0 = 0.0, μ1 = 0.2, σ = 1.0, alpha = 0.05, and beta = 0.20. They obtain a sample size of 199. The *z*-test should give a similar but slightly lower result because the normal distribution approximates the *t* distribution very well at this sample size.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 5** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For .... **Sample Size**
Alternative Hypothesis .... **Two-Sided (H1: μ ≠ μ0)**
Population Size... **Infinite**
Power.. **0.80**
Alpha.. **0.05**
μ0 (Null or Baseline Mean) **0**
μ1 (Actual Mean) **0.2**
σ (Standard Deviation).................................. **1**

Output

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

Numeric Results

Solve For: Sample Size
Hypotheses: H0: μ = μ0 vs. H1: μ ≠ μ0

Power	Sample Size N	Mean: Null μ0	Mean: Actual μ1	Mean: Difference μ1 - μ0	Standard Deviation σ	Effect Size	Alpha
0.80155	197	0	0.2	0.2	1	0.2	0.05

The sample size of 197 is very close to and just less than Machin's result for the *t*-test.