Excel **MASTER** Series

Nonparametric Testing in Excel

The Complete Guide

Excel Statistical Master

**Clear and Simple
yet THOROUGH
Statistical Instruction for the
Graduate Student and
Business Manager
with
LOTS of Worked-Out Problems
and Screen Shots**

Mark Harmon MBA

Nonparametric Testing in Excel

The Excel Statistical Master

By Mark Harmon

Copyright © 2011 Mark Harmon

No part of this publication may be reproduced
or distributed without the express permission
of the author.

mark@ExcelMasterSeries.com

www.ExcelMasterSeries.com

ISBN: 978-0-9833070-4-4

Table of Contents

(Click On Chapters To Go To Them)

Nonparametric Tests - When To Use Them

Statistical procedures are either parametric or nonparametric. Parametric statistical tests require assumptions about the population from which the samples are drawn. For example, many tests such as the t Test, Chi-Square tests, z Tests, and F tests, and many types of hypothesis tests require the underlying population to be normally distributed. Some tests require equal variances of both populations.

These assumptions cannot be always be assumed. Examples of this would be if the population is highly skewed or if the underlying distribution or variances were entirely unknown. Creating a quick Histogram of the data is fast way to check whether the distribution is skewed or not according any specific distribution. This is one brief step that would always be run on available population or sample data.

Nonparametric tests have no assumptions regarding distribution of underlying populations or variance. Most nonparametric tests are very easy to perform but they are not usually as precise as parametric tests and their Null Hypotheses usually requires more evidence to be rejected than would be the case in a parametric test.

Nonparametric tests are often used as shortcut replacements for more complicated parametric tests. You can usually get a quick answer that requires little calculation by running a nonparametric test.

Nonparametric tests are often used when the data is ranked but cannot be quantified. For example, how would you quantify consumer rankings such as very satisfied, moderately satisfied, just satisfied, less than satisfied, dissatisfied?

Nonparametric tests can be applied when there are a lot of outliers that might skew the results. Nonparametric tests often evaluate medians rather than means and therefore if the data have one or two outliers, the outcome of the analysis is not affected.

Nonparametric tests come in especially handy when dealing with non-numeric data, such as having customers rank products or attributes according to preference.

The Mann-Whitney U Test

Done in Excel

- The Mann-Whitney U Test is a great substitute for the two-sample t test when it can't be determined that both samples come from Normally-distributed populations.

- Is used to decide if whether or not there is a difference between two groups or, equivalently, whether or not both groups come from the same population.

Procedures

Objective: To determine if these two groups come from different populations

15	16
19	17
20	22
18	21
21	24
25	23
31	28

Step 1 - Combine All Samples Into 1 Group

Create a mechanism to **Keep Track of Each Sample's Group**. In this case, each sample from the 1st group was labeled with an "A" and each sample from the 2nd group was labeled with a "B."

15	A
19	A
20	A
18	A
21	A
25	A
31	A
16	B
17	B
22	B
21	B
24	B
23	B
28	B

Step 2 - Rank All Samples. – Sort the sample by the value of each sample and rank them in order of sample value.

1	15	A
2	16	B
3	17	B
4	18	A
5	19	A
6	20	A
7	21	A
8	21	B
9	22	B
10	23	B
11	24	B
12	25	A
13	28	B
14	31	A

Step 3 - Put Samples Back In Original Groups – Sort by the "A and B" column

1	15	A
4	18	A
5	19	A
6	20	A
7	21	A
12	25	A
14	31	A
2	16	B
3	17	B
8	21	B
9	22	B
10	23	B
11	24	B
13	28	B

Step 4 - Count the Rankings For Each Group

1	15	A
4	18	A
5	19	A
6	20	A
7	21	A
12	25	A
14	31	A

49   = R1

2	16	B
3	17	B
8	21	B
9	22	B
10	23	B
11	24	B
13	28	B

56   = R2

N1 = Count of Samples in 1st Group = 7

N2 = Count of Samples in 2nd Group = 7

Step 5 - Calculate the U Statistic by the following formula:

$U = N1*N2 + [(N1 * (N1 + 1)) / 2] - R1 = 28$

Step 6 - Calculate the Mean, μ_U , and Standard Deviation, σ_U, by the following formulas:

μ_U = N1 * N2 / 2 = 24.5

σ_U = SQRT [(N1 * N2 * (N1 + N2 + 1)) / 12] = 7.826

Step 7 - Calculate the Z Score With the Following Formula:

Z Score = (U - μ_U) / σ_U = 0.447

Step 8 - Compare this Z Score to Z Critical for the Given α

For a one-tailed test, Z Critical = NORMSINV(α)

For a two-tailed test, Z Critical = NORMSINV(α/2)

If We Require a 95% Degree of Certainty, Then α = 0.05

For a one-tailed test, Z Critical = NORMSINV(α) = NORMSINV(0.05) = - 1.645

For a two-tailed test, Z Critical = NORMSINV(α/2) =NORMSINV(0.025) - 1.960

Rule

The Result if Significant If | Z Score | > | Z Critical |

In This Case, | Z Score | (0.447) > | Z Critical | (1.96 for two-tailed and 1.645 for one-tailed)

So The Result Is Not Statistically Significant and We Cannot State With 95% Certainty That There is a Difference Between The Two Groups for Either a One-Tailed Test Or a Two-Tailed Test.

The Kruskal-Wallis Test - Done in Excel

- Is used to decide whether or not there is a difference between more than two groups or, equivalently, whether or not all groups come from the same population.

- Is a good substitute for Single-Factor (One-Way) ANOVA when it cannot be determine that the samples are drawn from Normally-distributed populations. It is almost identical to the one-way (single-factor) ANOVA except that the data is replaced by the ranks of each piece of data.

Procedures

Objective: To determine if these three groups come from different populations

15	16	14
19	17	2
20	22	7
18	21	9
21	24	11
25	23	34
31	28	36

Step 1 - Combine all samples into 1 group.

Create a mechanism to keep track of each sample's group. In this case, each sample in the 1st group was labeled "A," each sample in the 2nd group was labeled "B," and each sample in the 3rd group was labeled "C." This step is taken so that the data can be sorted by this column to return all records to their original groups.

15	A
19	A
20	A
18	A
21	A
25	A
31	A
16	B
17	B
22	B
21	B
24	B
23	B
28	B
14	C
2	C
7	C
9	C
11	C
34	C
36	C

Step 2 - Rank All Samples

1	2	C
2	7	C
3	9	C
4	11	C
5	14	C
6	15	A
7	16	B
8	17	B
9	18	A
10	19	A
11	20	A
12	21	A
13	21	B
14	22	B
15	23	B
16	24	B
17	25	A
18	28	B
19	31	A
20	34	C
21	36	C

Step 3 - Put samples back in original groups. Do this by sorting by the "A B C" column.

1	15	A
4	18	A
5	19	A
6	20	A
7	21	A
12	25	A
14	31	A
49	**= R1**	

2	16	B
3	17	B
8	21	B
9	22	B
10	23	B
11	24	B
13	28	B
56	**= R2**	

1	2	C
2	7	C
3	9	C
4	11	C
5	14	C
20	34	C
21	36	C
56	**= R3**	

Step 4 – Calculate the Following Metrics

N1 = Count of Samples in 1st Group = 7

N2 = Count of Samples in 2nd Group = 7

N3 = Count of Samples in 1st Group = 7

n = Total Number of Samples = 21

k = Number of Groups = 3

df = Degrees of Freedom = k - 1 = 2

Step 5 - Calculate the H Statistic by the following formula:

H Statistic =

[12 / { n*(n + 1) }] * [(R1^2)/N1 + (R2^2)/N2 + … + (RN^2)/NN] – 3*(n + 1)

H Statistic = -33.818

Step 6 - Calculate H Critical

Test Statistic H is very nearly distributed as the Chi-Square Distribution With k - 1 Degrees of Freedom as long as the number of samples in each group Is at least 5.

Therefore the Critical Value Can Be Found By The Following Excel Formula:

H Critical = CHIINV(α, df)

If we wish to be at least 95% of whether or not the result is statistically significant, then α = 0.05

H Critical = CHIINV(α, df) = 5.991

Rule

The Result if Significant If H Statistic > H Critical

In This Case, H Statistic > H Critical

So The Result Is Statistically Significant and We Can State Within 95% That There is Difference Between The Groups, Or, Equivalently, They Come From Different Populations.

The Wilcoxon Rank Sum Test
Done in Excel

The Wilcoxon Rank Sum Test has the same purpose as the t Test; to determine whether there is a difference between two groups. Specifically, we are trying to determine whether the two samples came from the same populations or not. The Wilcoxon Rank Sum Test can be used in place of the t Test when either of the sample groups does not appear to be normally distributed.

The Wilcoxon Rank Sum Test is a hypothesis test and has the following Null Hypothesis and an Alternate Hypothesis.

H0 - Null Hypothesis states that the two samples came from the same populations. In other words, there is no difference between the two groups.

H1 - Alternate Hypothesis states that the two samples do not come from the same populations.

Test Procedures in Excel

Large and Small sampling require different procedures for the Wilcoxon Rank Sum Test. The first couple steps of this test will be the same for both large and small samples:

We will collect two groups of samples and then calculate the following characteristics of each group:

n1 = Sample size of the 1st group - Number of samples in the group

n2 = Sample size of the 2nd group - Number of samples in the group

N1 = Sum of the ranks of all samples in the 1st group

N2 = Sum of the ranks of all samples in the 2nd group

R = either N1 or N2, whichever comes from the group with the smaller sample size

On the next pages we will show how to calculate N1 and N2 in Excel.

First, we have to determine whether we are dealing with small samples or large samples. The small sample procedure is used if n1 and n2 are both each less than 10. Use the large sample procedure is there are more than 20 samples total (n1 + n2 > 20).

The small sample procedure and the large sample procedure calculate n1, n2, N1, N2, and R the same way. Small samples and large samples use different procedures beyond this point.

Wilcoxon Rank Sum Test

for Small Samples In Excel

"Small samples" occur when the sample size of both groups is less than 10. In this case, n1 and n2 are both less than 10.

Problem: We have collected sales data from the last 6 Tuesdays and the last 7 Wednesdays. We want to determine if Tuesday sales have been different from Wednesday sales over the last month and a half. The data in either group does not appear to be normally distributed. In this case we will have to use a nonparametric test to determine whether the groups are different. The Wilcoxon Rank Sum Test for Small Samples is ideal for this situation.

 Here are the data samples for each of the two groups:

Tuesday Sales	Wednesday Sales
75	83
78	75
84	87
79	69
81	84
88	89
70	

Step 1 - Count the samples in Each Group

Sample Size =

n1 = 7

n2 = 6

Step 2 - Rank all samples according to sample size in Excel

Sample	Day
75	Tuesday
78	Tuesday
84	Tuesday
79	Tuesday
81	Tuesday
88	Tuesday
70	Tuesday
83	Wednesday
75	Wednesday
87	Wednesday
69	Wednesday
84	Wednesday
89	Wednesday

Step 3 - Combine All Samples and Rank Them. If any samples have the same

values, assign them the rank that would be the average of the two ranks if they weren't equal.

Rank	Sample	Day
1	69	Wednesday
2	70	Tuesday
3.5	75	Tuesday
3.5	75	Wednesday
4	78	Tuesday
5	79	Tuesday
6	81	Tuesday
7	83	Wednesday
8.5	84	Tuesday
8.5	84	Wednesday
9	87	Wednesday
10	88	Tuesday
11	89	Wednesday

Step 4 - Re-sort the Rows According To Day

Rank	Sample	Day
2	70	Tuesday
3.5	75	Tuesday
4	78	Tuesday
5	79	Tuesday
6	81	Tuesday
8.5	84	Tuesday
10	88	Tuesday
1	69	Wednesday
3.5	75	Wednesday
7	83	Wednesday
8.5	84	Wednesday
9	87	Wednesday
11	89	Wednesday

Step 5 - Calculate the Sum of the Ranks (N1 and N2) For Each Sample Group

Rank	Sample	Day
2	70	Tuesday
3.5	75	Tuesday
4	78	Tuesday
5	79	Tuesday
6	81	Tuesday
8.5	84	Tuesday
10	88	Tuesday

N1 = 39

Rank	Sample	Day
1	69	Wednesday
3.5	75	Wednesday
7	83	Wednesday
8.5	84	Wednesday
9	87	Wednesday
11	89	Wednesday

N2 = 40

Step 6 - Determine whether R = N1 or N2

R is taken from the sample with the smaller size.

n1 = 7 and n2 = 6 so sample group 2 is smaller

Therefore R = N2 = 40

Step 7 - Use the Wilcoxon Rank Sum Small Sample Chart to Evaluate R to Determine Whether R Falls Within or Outside the Lower and Upper Critical Values for the Given n1 and n2:

Below is the chart that is referenced. The green cells represent the n1, n2, and upper and lower limits for that n1 and n2.

Lower and Upper Critical Values for the Wilcoxon Rank Sum Test for small samples

Area in one outer tail = 0.025
α = 0.025 for one-tailed test (0.025 is only in one outer tail)
α = 0.05 for two-tailed test (0.025 is in each of both outer tails)

n1	3	4	5	6	7	8	9	10
n2								
3	5, 16	6, 18	6, 21	7, 23	7, 26	8, 28	8, 31	9, 33
4	6, 18	11, 25	12, 28	12, 32	13, 35	14, 38	15, 41	16, 44
5	6, 21	12, 28	18, 37	19, 41	20, 45	21, 49	22, 53	24, 56
6	7, 23	12, 32	19, 41	26, 52	28, 56	29, 61	31, 65	32, 70
7	7, 26	13, 35	20, 45	28, 56	37, 68	39, 73	41, 78	43, 83
8	8, 28	14, 38	21, 49	29, 61	39, 73	49, 87	51, 93	54, 98
9	8, 31	15, 41	22, 53	31, 65	41, 78	51, 93	63, 108	66, 114
10	9, 33	16, 41	24, 56	32, 70	43, 83	54, 98	66, 114	79, 131

Area in one outer tail = 0.05
α = 0.05 for one-tailed test (0.05 is only in one outer tail)
α = 0.10 for two-tailed test (0.05 is in each of both outer tails)

n1	3	4	5	6	7	8	9	10
n2								
3	6, 15	7, 17	7, 20	8, 22	9, 24	9, 27	10, 29	11, 31
4	7, 17	12, 24	13, 27	14, 30	15, 33	16, 36	17, 39	18, 42
5	7, 20	13, 27	19, 36	24, 40	22, 43	24, 46	25, 50	26, 54
6	8, 22	14, 30	20, 40	28, 50	30, 54	32, 58	33, 63	35, 67
7	9, 24	15, 33	22, 43	30, 54	39, 66	41, 71	43, 76	46, 80
8	9, 27	16, 36	24, 46	32, 58	41, 71	52, 84	54, 90	57, 95
9	10, 29	17, 39	25, 50	33, 63	43, 76	54, 90	66, 105	69, 111
10	11, 31	18, 42	26, 54	35, 67	46, 80	57, 95	69, 111	83, 127

The chart on the previous page is a standard Critical Value chart for the Wilcoxon Small Sample Rank Sum Test.

If R = 40, N1 = 7, and N2 = 6, then R is between the upper and lower critical values for both α = 0.025 (Critical Values = 28 and 56) and α = 0.05 (Critical Values = 30 and 54). That is shown on the above chart.

We therefore do not reject the Null Hypothesis, which states that there is no difference between the two sample groups.

Wilcoxon Rank Sum Test

for Large Samples in Excel

"Large samples" imply that both n1 and n2 add up to more than 20 and that each is greater than 10.

Problem: We have collected sales data from the last 14 Tuesdays and the last 12 Wednesdays. We want to determine if Tuesday sales have been different from Wednesday sales over the last month and a half. Calculate with at least 95% certainty whether there is a difference between the two groups.

The data in either group does not appear to be normally distributed. In this case we will have to use a nonparametric test to determine whether the groups are different. The Wilcoxon Rank Sum Test for Large Samples is ideal for this situation.

For large samples, you will calculate n1, n2, N1, N2, and R as before.

Here are the data samples for each of the two groups:

Tuesday Sales	Wednesday Sales
75	83
78	75
84	87
79	69
81	84
88	89
70	83
75	75
78	87
84	69
79	84
81	89
88	
70	

Step 1 - Count the Samples in Each Group

Sample Size =

n1 = 14

n2 = 12

Step 2 - Rank All Samples According To Sample Size in Excel

Sample	Day
75	Tuesday
78	Tuesday
84	Tuesday
79	Tuesday
81	Tuesday
88	Tuesday
70	Tuesday
75	Tuesday
78	Tuesday
84	Tuesday
79	Tuesday
81	Tuesday
88	Tuesday
70	Tuesday
83	Wednesday
75	Wednesday
87	Wednesday
69	Wednesday
84	Wednesday
89	Wednesday
83	Wednesday
75	Wednesday
87	Wednesday
69	Wednesday
84	Wednesday
89	Wednesday

Step 3 - Combine All Samples and Rank Them. If any samples have the same

values, assign them the rank that would be the average of the two ranks if they weren't equal.

Rank	Sample	Day
1.5	69	Wednesday
1.5	69	Wednesday
3.5	70	Tuesday
3.5	70	Tuesday
6.5	75	Tuesday
6.5	75	Tuesday
6.5	75	Wednesday
6.5	75	Wednesday
9.5	78	Tuesday
9.5	78	Tuesday
11.5	79	Tuesday
11.5	79	Tuesday
13.5	81	Tuesday
13.5	81	Tuesday
15.5	83	Wednesday
15.5	83	Wednesday
18.5	84	Tuesday
18.5	84	Tuesday
18.5	84	Wednesday
18.5	84	Wednesday
21.5	87	Wednesday
21.5	87	Wednesday
23.5	88	Tuesday
23.5	88	Tuesday
25.5	89	Wednesday
25.5	89	Wednesday

Step 4 - Re-sort the Rows According to the Day of the Week

Rank	Sample	Day
3.5	70	Tuesday
3.5	70	Tuesday
6.5	75	Tuesday
6.5	75	Tuesday
9.5	78	Tuesday
9.5	78	Tuesday
11.5	79	Tuesday
11.5	79	Tuesday
13.5	81	Tuesday
13.5	81	Tuesday
18.5	84	Tuesday
18.5	84	Tuesday
23.5	88	Tuesday
23.5	88	Tuesday
1.5	69	Wednesday
1.5	69	Wednesday
6.5	75	Wednesday
6.5	75	Wednesday
15.5	83	Wednesday
15.5	83	Wednesday
18.5	84	Wednesday
18.5	84	Wednesday
21.5	87	Wednesday
21.5	87	Wednesday
25.5	89	Wednesday
25.5	89	Wednesday

Step 5 - Calculate the Sum of the Ranks (N1 and N2) For Each Sample Group

Rank	Sample	Day
3.5	70	Tuesday
3.5	70	Tuesday
6.5	75	Tuesday
6.5	75	Tuesday
9.5	78	Tuesday
9.5	78	Tuesday
11.5	79	Tuesday
11.5	79	Tuesday
13.5	81	Tuesday
13.5	81	Tuesday
18.5	84	Tuesday
18.5	84	Tuesday
23.5	88	Tuesday
23.5	88	Tuesday

$N1 =$ 173

Rank	Sample	Day
1.5	69	Wednesday
1.5	69	Wednesday
6.5	75	Wednesday
6.5	75	Wednesday
15.5	83	Wednesday
15.5	83	Wednesday
18.5	84	Wednesday
18.5	84	Wednesday
21.5	87	Wednesday
21.5	87	Wednesday
25.5	89	Wednesday
25.5	89	Wednesday

$N2 =$ 178

Step 6 - Determine Whether R = N1 or N2

R is taken from the sample with the smaller size.

n1 = 14 and n2 = 12 so sample group 2 is smaller

Therefore R = N2 = 178

Step 7 - Calculate Z Score and Z Critical to Determine Whether the Sample Groups Come From Different Populations.

When n1 and n2 are both > 10, the Normal distribution can approximate the distribution of R.

If the Z Score of R is greater than Z Critical, we reject the Null Hypothesis and state that the two sample groups come from different populations. If the Z Score of R is less than Z Critical, we do not reject the Null Hypothesis and state that both sample groups came from the same populations.

Z Score Calculation

Calculate μr and σr

$$\mu r = [\, n1 * (\, n1 + n2 + 1\,)\,]\, /\, 2$$

$$\sigma r = SQRT\, [\, (\, n1 * n2 * (\, n1 + n2 + 1\,)\,)\, /\, 12\,]$$

$$\mu r = [\, 14 * (\, 14 + 12 + 1\,)\,]\, /\, 2 = 189$$

$$\sigma r = SQRT\, [\, (\, 14 * 12 * (\, 14 + 12 + 1\,)\,)\, /\, 12\,]\ = 19.44$$

Calculate the Z Score

$Z = (R - \mu_r) / \sigma_r = -0.5658$

Calculate Z Critical

Zcrit = NORMSINV (1 - α / 2) For example, for α = 0.05 and two-tailed test

Zcrit = NORMSINV (1 - 0.05 / 2) = NORMSINV (0.975) = 1.96

Rule

Reject H0 if | Z | > | Z Critical | for the given α and correct number of tails

In this case, |Z| is less than | Z Critical | so we do not reject the Null Hypothesis. We state that we are 95% certain that the two sample groups were taken from the same population.

The Spearman Correlation Coefficient Test

Done in Excel

- Used to determine if the strength and direction of a relationship between 2 sets of numbers.

- Is sometimes considered to be similar to the Pearson Correlation Coefficient between the ranks of numbers.

Procedures

Objective: To determine if there is a correlation between these two sets of numbers:

Sale Price	Time Spent On Website
26	27
19	18
25	21
16	21.5
24.5	23
24	24
28	29

Step 1 - Rank the Data.

Sale Price	Rank 1	Time Spent On Website	Rank2
26	6	27	6
19	2	18	1
25	5	21	2
16	1	21.5	3
24.5	4	23	4
24	3	24	5
28	7	29	7

Step 2 - Calculate the Difference Between Ranks 1 and 2, and Then Square That Difference.

Rank 1	Rank2	d = Difference Between Rankings	d^2
6	6	0	0
2	1	1	1
5	2	3	9
1	3	-2	4
4	4	0	0
3	5	-2	4
7	7	0	0

$\Sigma (d^2) = 18$

n = Number of Samples = 7

Step 3 – Calculate the Spearman Rank Correlation Coefficient

Spearman Rank Correlation Coefficient, r_s, = 1 – [6*Σ(d^2)] / [n*(n^2 – 1)]   = 0.6786

Rule:

r_s **is statistically significant if the following p Value is less than α:**

p Value = FDIST [(df * r_s^2) / (1 - r_s^2), 1, df]

df = degrees of freedom = n - 2 = 5

p Value = 0.001635228

If We Require a 95% Degree of Certainty, Then α = 0.05

In This Case, We Can Be At Least 95% Certain That There Is a Relationship Between The Variables.

Here Is a Quick Way To Assign Ranking To Samples In a Data Set Such As This:

Step 1 - Assign the Count to Each Sample According Sample Size

Step 2 - Sort Both Columns By the Original Data Column (Right Column) and Assign a Numerical Rank (New Left Yellow-Highlighted Column)

1	5	4
2	3	12
3	1	23
4	2	34
5	6	43
6	4	76

Step 3 - Sort All 3 Columns By the Middle Column

3	1	23
4	2	34
2	3	12
6	4	76
1	5	4
5	6	43

Step 4 – Delete the Middle Column, Leaving the Original Data Column (Right) and the Rank of Record (Left Column)

3	23
4	34
2	12
6	76
1	4
5	43

The Sign Test (Nonparametric)

Done in Excel

1) Used to determine whether a population median is equal to, less than, or greater than a specific number

2) Used to determine whether there is a real difference between pairs of data, such as before and after measurements

3) Has wide application but is not as statistically accurate as the parametric tests such as the paired samples t test.

Procedures

Objective: To determine whether "Before" and "After" data are different.

Small samples - less than 25 samples

For Case 1 above, in the event of testing to determine whether a population median is equal to, less than, or greater to a specific number, subtract the sample from the fixed number.

For Case 2 above, in the event of testing to determine whether there is a difference between before and after measurements, subtract the before-measurement from the after-measurement.

Small Sample Exercise: Determine Whether the Before and After Data is Different:

Here is the original Before and After Data:

Before	After
6	7
5	7
6	6
4	4
4	5
5	8
6	7
5	5
5	7
3	3
4	3
4	5
5	6

Step 1 - Take the Difference Between Each Before and After Data Point.

Step 2 - Record Whether That Difference Is Positive or Negative.

Step 3 - Record the Total Number of Positive and Negative Results.

Before	After	Difference	+ numbers	- numbers
6	7	1	1	
5	7	2	1	
6	6	0		
4	4	0		
4	5	1	1	
5	8	3	1	
6	7	1	1	
5	5	0		
5	7	2	1	
3	3	0		
4	3	-1		1
4	5	1	1	
5	6	1	1	

Σ (+ numbers) =	8
Σ (- numbers) =	1

X = Lesser of (+) or (-) Numbers = 1

N = Sum of (+) and (-) Numbers = 3 + 5 = 9

X = Lesser of (+) or (-) Numbers = 1

N = Sum of (+) and (-) Numbers = 3 + 5 = 9

p Value = -2*BINOMDIST(X, N, 0.5, 1) = 0.0390625

The p value is the total area under the outer 2 tails in a 2-tailed test

Rule

The Result is Significant if p Value < α

If 95% certainty is required, then α = 0.05

In This Case, the Result Is Significant and We Can State That There Is a Difference Between the Before and After Numbers.

Large Samples - 25 or More Samples

Plug the X and N into the Following Formula

Z Score = [(X + 0.5) – (N – 2)] / [SQRT(N) / 2]

Then Compare this Z Score to Z Critical for the Given α

For a one-tailed test, Z Critical = NORMSINV(α)

For a two-tailed test, Z Critical = NORMSINV(α/2)

Rule

The Result if Significant If | Z Score | > | Z Critical |

For Example:

X = 3

N = 9

Z Score = [(X + 0.5) – (N – 2)] / [SQRT(N) / 2] = -2.33

If We Require a 95% Degree of Certainty, Then α = 0.05

For a one-tailed test, Z Critical = NORMSINV(α) = -1.64

For a two-tailed test, Z Critical = NORMSINV(α/2) = -1.96

In this case, we can state with at least 95% certainty that the result is significant and that there is a difference between Before and After data for both the one-tailed and two-tailed tests.

Wilcoxon Signed-Rank Test in Excel

for Small Samples

The **Wilcoxon Signed-Rank Test** is used as an alternate to the **Paired Student's t Test** when the underlying population cannot be determined to be normally distributed. This test is used to determine whether there is a difference between paired samples.

Unlike another similar nonparametric test - the **Paired Sample Sign Test** - this test considers the magnitude of differences between the paired samples.

The Wilcoxon Signed-Rank Test is performed differently for small samples than it is for large samples. Small samples mean sample groups with less than 30 pairs of sample data.

The Null Hypothesis for the Wilcoxon Signed-Rank Test states that there is no difference between the paired samples. For example, this Null Hypothesis might state that there is no difference between before and after data, if each pair represents a before and after measurement of one sampled object.

The Alternate Hypothesis states that there is a difference between the paired samples.

Procedures For the Wilcoxon Signed-Rank Test

In this test, we are trying to determine if landing pages produce greater average sales when a bonus eBook is added to the offering. We wish to be at least 95% certain whether average sales do increase when an eBook is added to the offering on each landing page.

A landing page is that web page on a web site that a visitor is directed to (lands on) when the visitor clicks on an ad that is external to the web site. The ad can be any type of ad, such as a banner ad in another web site, a Google AdWords ad, or a link in a promotional email.

Bonus eBooks were added to the offering that appeared on 9 separate landing pages. Average daily sales from each landing page were recorded for one month before and one month after the bonus eBook was added to each landing page's offering. All 9 landing pages were very similar and were the destination URLs (landing

pages) of similar Google AdWords ads.

We do not know whether sales results from the landing pages are normally distributed so we need to use a nonparametric test. The most accurate nonparametric test for paired data is the **Wilcoxon Signed-Rank Test**. Here is how it's done in Excel:

Here is the Original Before and After Sales Data:

Landing Page	Avg. Daily Sales Before Bonus eBook	Avg. Daily Sales After Bonus eBook
A	28	31
B	30	29
C	31	34
D	29	31
E	27	32
F	29	35
G	31	36
H	30	32
I	28	31

Step 1 - Calculate Differences Between Before and After Data For Each Pair

Landing Page	Avg. Daily Sales Before Bonus eBook	Avg. Daily Sales After Bonus eBook	Diff.
A	28	31	3
B	30	29	-1
C	31	34	3
D	29	31	2
E	27	32	5
F	29	35	6
G	31	36	5
H	30	32	2
I	28	31	3

Step 2 - Take the Absolute Value of that Difference

Diff.	| Diff. |
3	3
-1	1
3	3
2	2
5	5
6	6
5	5
2	2
3	3

http://excelmasterseries.com/New_Manuals.php

Step 3 - Rank of the Differences and Record the Sign of Each Difference

When the differences are the same, assign to each the average rank between them.

Diff.	| Diff. |	Rank of | Diff. |	Sign of Diff.
3	3	5	+
-1	1	1	-
3	3	5	+
2	2	2.5	+
5	5	7.5	+
6	6	9	+
5	5	7.5	+
2	2	2.5	+
3	3	5	+

Step 4 - Combine the Rank of Each Difference With the Sign of the Difference

Rank of | Diff. |	Sign of Diff.	Sign Diff. and Rank | Diff.|
5	+	5
1	-	-1
5	+	5
2.5	+	2.5
7.5	+	7.5
9	+	9
7.5	+	7.5
2.5	+	2.5
5	+	5

Calculate the following two numbers:

a) The absolute value of the sum of all positive sign ranks.

b) The absolute value of the sum of all negative sign ranks.

The Test Statistics W equals the lesser of the two absolute values.

| Sum of all positive sign ranks | = | 5 + 5 + 2.5 + 7.5 + 9 + 7.5 + 2.5 + 5 | = | 44 | = 44

| Sum of all negative ranks | = | -1 | = 1

The Test Statistic W therefore equals 1.

Rule

Reject the Null Hypothesis if Test Statistic W is greater than or equal to W Critical

Below is a standard chart containing W Critical Values we need:

**W Critical Values for the
Wilcoxon Signed-Rank Test for small samples**

One-Tailed α		0.05	0.025	0.01	0.005
Two-Tailed α		0.1	0.05	0.02	0.01
n = number of sample pairs					
5	1				
6	2	1			
7	4	2	0		
8	6	4	2	0	
9	8	6	3	2	
10	11	8	5	3	
11	14	11	7	5	
12	17	14	10	7	

W Critical for n = 9 (9 pairs of sample data) for a one-tailed test in which alpha = 0.05 equals 8.

Since Test Statistics W (1) is less than W Critical (8), we reject the Null Hypothesis and state with at least 95% certainty that adding a bonus eBook to the offering on each landing page does increase the average sales from each landing page.

Wilcoxon Signed-Rank

Test in Excel for

Large Samples

The **Wilcoxon Signed-Rank Test** is used as an alternate to the **Paired Student's t Test** when the underlying population cannot be determined to be normally distributed. This test is used to determine whether there is a difference between paired samples. The most widely used application of the Wilcoxon Signed-Rank Test is to determine whether there really is a difference between before and after data. In this article, we will show how to use the Wilcoxon Signed-Rank Test in Excel to determine whether sales really did increase across a group of landing pages after a free bonus was added to the offer on each landing page.

Unlike another similar nonparametric test - the **Paired Sample Sign Test** - this test considers the magnitude of differences between the paired samples.

The Wilcoxon Signed-Rank Test is performed very differently for large samples than it is for small samples. Large samples mean sample groups with more than 30 pairs of sample data.

The Null Hypothesis for the Wilcoxon Signed-Rank Test states that there is no difference between the paired samples. For example, this Null Hypothesis might state that there is no difference between before and after data, if each pair represents a before and after measurement of one sampled object.

The Alternate Hypothesis states that there is a difference between the paired samples.

Procedures For the Wilcoxon Signed-Rank Test

In this test, we are trying to determine if landing pages produce greater average sales when a bonus eBook is added to the offering. We wish to be at least 95% certain whether average sales do increase when an eBook is added to the offering on each landing page.

A landing page is that web page on a web site that a visitor is directed to (lands on) when the visitor clicks on an ad that is external to the web site. The ad can be any type of ad, such as a banner ad in another web site, a Google AdWords ad that came up during a Google search, or a link in a promotional email.

Bonus eBooks were added to the offering that appeared on 35 separate landing pages. Average daily sales from each landing page were recorded for one month before and one month after the bonus eBook was added to each landing page's offering. All 35 landing pages were very similar and were the destination URLs (landing pages) of similar Google AdWords ads.

We do not know whether sales results from the landing pages are normally distributed so we need to use a nonparametric test. The most accurate nonparametric test for paired data is the **Wilcoxon Signed-Rank Test**. Here is how it's done in Excel for large samples (sample size is greater than 30):

1) Before and After Sales Data

Initial Data

Landing Page	Avg. Daily Sales Before Bonus eBook	Avg. Daily Sales After Bonus eBook
1	28	31
2	30	29
3	31	34
4	29	31
5	27	32
6	29	35
7	31	36
8	30	32
9	28	31
10	28	29
11	30	16
12	31	38
13	29	14
14	27	18
15	29	29
16	31	25
17	30	29
18	28	32
19	28	37
20	30	35
21	31	24
22	29	29
23	27	36
24	29	42
25	31	18
26	30	22
27	28	37
28	28	28
29	30	29
30	31	30
31	29	31
32	27	22
33	29	25
34	31	29
35	30	31

2) Calculate Differences Between Before and After Data For Each Pair On Next Page

Landing Page	Avg. Daily Sales Before Bonus eBook	Avg. Daily Sales After Bonus eBook	Diff.
1	28	31	3
2	30	29	-1
3	31	34	3
4	29	31	2
5	27	32	5
6	29	35	6
7	31	36	5
8	30	32	2
9	28	31	3
10	28	29	1
11	30	16	-14
12	31	38	7
13	29	14	-15
14	27	18	-9
15	29	29	0
16	31	25	-6
17	30	29	-1
18	28	32	4
19	28	37	9
20	30	35	5
21	31	24	-7
22	29	29	0
23	27	36	9
24	29	42	13
25	31	18	-13
26	30	22	-8
27	28	37	9
28	28	28	0
29	30	29	-1
30	31	30	-1
31	29	31	2
32	27	22	-5
33	29	25	-4
34	31	29	-2
35	30	31	1

3) Calculate the Absolute Value of the Difference and Then the Rank and Sign of that Difference When the differences are the same, assign to each the

average rank between them:

Diff.	| Diff. |	Rank of | Diff. |	Sign of Diff.
3	3	12	+
-1	1	2.5	-
3	3	12	+
2	2	8.5	+
5	5	17.5	+
6	6	20.5	+
5	5	17.5	+
2	2	8.5	+
3	3	12	+
1	1	2.5	+
-14	14	31	-
7	7	22.5	+
-15	15	32	-
-9	9	6.25	-
0	0		
-6	6	20.5	-
-1	1	2.5	-
4	4	14.5	+
9	9	6.25	+
5	5	17.5	+
-7	7	22.5	-
0	0		
9	9	6.25	+
13	13	29.5	+
-13	13	29.5	-
-8	8	24	-
9	9	6.25	+
0	0		
-1	1	2.5	-
-1	1	2.5	-
2	2	8.5	+
-5	5	17.5	-
-4	4	14.5	-
-2	2	8.5	-
1	1	2.5	+

4) Combine the Rank and Sign of Each Difference

Rank of | Diff. |	Sign of Diff.	Sign | Diff. | and Rank Diff.
12	+	12
2.5	-	-2.5
12	+	12
8.5	+	8.5
17.5	+	17.5
20.5	+	20.5
17.5	+	17.5
8.5	+	8.5
12	+	12
2.5	+	2.5
31	-	-31
22.5	+	22.5
32	-	-32
6.25	-	-6.25
		0
20.5	-	-20.5
2.5	-	-2.5
14.5	+	14.5
6.25	+	6.25
17.5	+	17.5
22.5	-	-22.5
		0
6.25	+	6.25
29.5	+	29.5
29.5	-	-29.5
24	-	-24
6.25	+	6.25
		0
2.5	-	-2.5
2.5	-	-2.5
8.5	+	8.5
17.5	-	-17.5
14.5	-	-14.5
8.5	-	-8.5
2.5	+	2.5

5) Calculate the Sums of All Ranks

Sign | Diff. | and Rank Diff.	All Positive Signed- Ranks	All Negative Signed- Ranks
12	12	
-2.5		-2.5
12	12	
8.5	8.5	
17.5	17.5	
20.5	20.5	
17.5	17.5	
8.5	8.5	
12	12	
2.5	2.5	
-31		-31
22.5	22.5	
-32		-32
-6.25		-6.25
0		
-20.5		-20.5
-2.5		-2.5
14.5	14.5	
6.25	6.25	
17.5	17.5	
-22.5		-22.5
0		
6.25	6.25	
29.5	29.5	
-29.5		-29.5
-24		-24
6.25	6.25	
0		
-2.5		-2.5
-2.5		-2.5
8.5	8.5	
-17.5		-17.5
-14.5		-14.5
-8.5		-8.5
2.5	2.5	
Sum of All Positive Signed- Ranks	**Sum of All Negative Signed- Ranks**	
224.75	**-216.25**	

Sum of All Ranks = 224.75 - 216.25 = 8.5

6) Calculate the Test Statistic W

Test Statistic W equals the sum of the signed ranks, which was calculated above:

W = 8.5

Since sample size is greater than 30, the distribution of Test Statistic T can be approximated by the Normal Distribution.

7) Calculate the Z Score for Test Statistic T

With large sample sizes (sample size, n, greater than 30), the distribution of Test Statistic W can approximated with the Normal Distribution. We can therefore calculate a Z Score for Test Statistic W.

The Z Score is the number of standard errors that Test Statistic T is from the mean of the Normal Distribution curve that was use to approximate the distribution of Test Statistic T.

There are 2 pieces of information that are needed in order to construct the specific Normal Distribution curve that is used to approximate the distribution of Test Statistic W. These 2 pieces of information are the 2 parameters necessary to construct any Normal Distribution curve. They are:

- Mean (of Test Statistic W) = μ_W

- Standard Deviation (of Test Statistic W) = σ_W

The Mean of the Normal Distribution curve used to estimate the distribution of Test Statistic W (μ_W) is determined by the Null Hypothesis. The Null Hypothesis states that there is no difference between the before and after sales. Put another way, the Null Hypothesis states that the difference between the before and after sales equals zero. Therefore the mean of Test Statistic W and the Normal Distribution curve that is used to approximate its distribution also equals zero. So;

$\mu_W = 0$

The standard deviation of Test Statistic W and therefore of the Normal Distribution curve that is used to approximate the distribution of Test Statistic W is found by the following formula:

σ_W = SQRT **[(** n * **(**n+1**)** * **(**2n + 1**))** / 6 **]**

n equals the number of before-and-after pairs of sample data whose difference does not equal zero. In this case n = 32.

The above formula for σ_W can be derived from the fact that for any particular value of n, it can be shown that the standard deviation of the sampling distribution of W is equal to:

σ_W = SQRT **[(** n * **(**n+1**)** * **(**2n + 1**))** / 6 **]**

A number of sources incorrectly state that the above formula is as follows:

σ_W = SQRT **[(** n * **(**n+1**)** * **(**2n + 1**))** / 24 **]**

Since n = 32

σ_W = 106.96 which can be rounded off to σ_W = 107

We can now map the Normal Distribution curve that is used to approximate the distribution of Test Statistic W as follows:
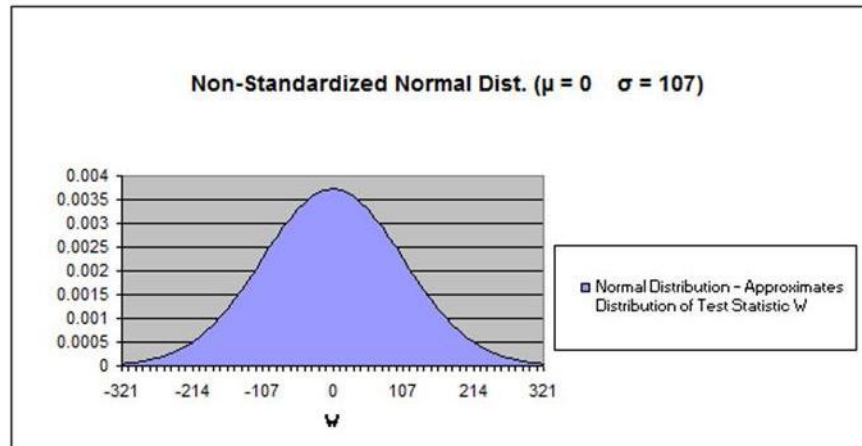
The Z Score of Test Statistic W can be determined by the following formula:

Z Score = (W - μ_W ± 0.5) **/** σ_W

The Z Score above includes a continuity correction, to account for the fact that we use a continuous probability function (the Normal Distribution) and our test value is based on discrete (ordinal, ranked) data. If the difference W - μ_W is negative the continuity correction is +0.5 (this is the case when W is less than; μ_W). If the difference W - μ_W is positive the correction is -0.5 (this is the case when W is greater than μ_W).

Z Score = (W - μ_W ± 0.5) **/** σ_W

Z Score = (8.5 - 0 - 0.5) / 107 = 0.0748

8) Calculate the p Value That Corresponds With That Z Score

In Excel, the formula is for calculating the p value when it is in the right outer tail is:

p Value = 1 - NORMDIST (X, μ_x, σ_x , Cumulative ?)

Which in this case is:

p Value = 1 - NORMDIST (Test Statistic W, μ_w, σ_w , Cumulative ?)

p Value = 1 - NORMDIST (8.5, 0, 107, TRUE) = 1 - 0.53 = 0.47

Rule:

Reject the Null Hypothesis if the p Value Test Statistic W is less than α, which is the Level of Significance.

Since we require 95% certainty in our test, α = 0.05.

We do not reject the Null Hypothesis in this case because the p Value of Test Statistic W is significantly greater than α.

Here is a graphical representation of the answer:
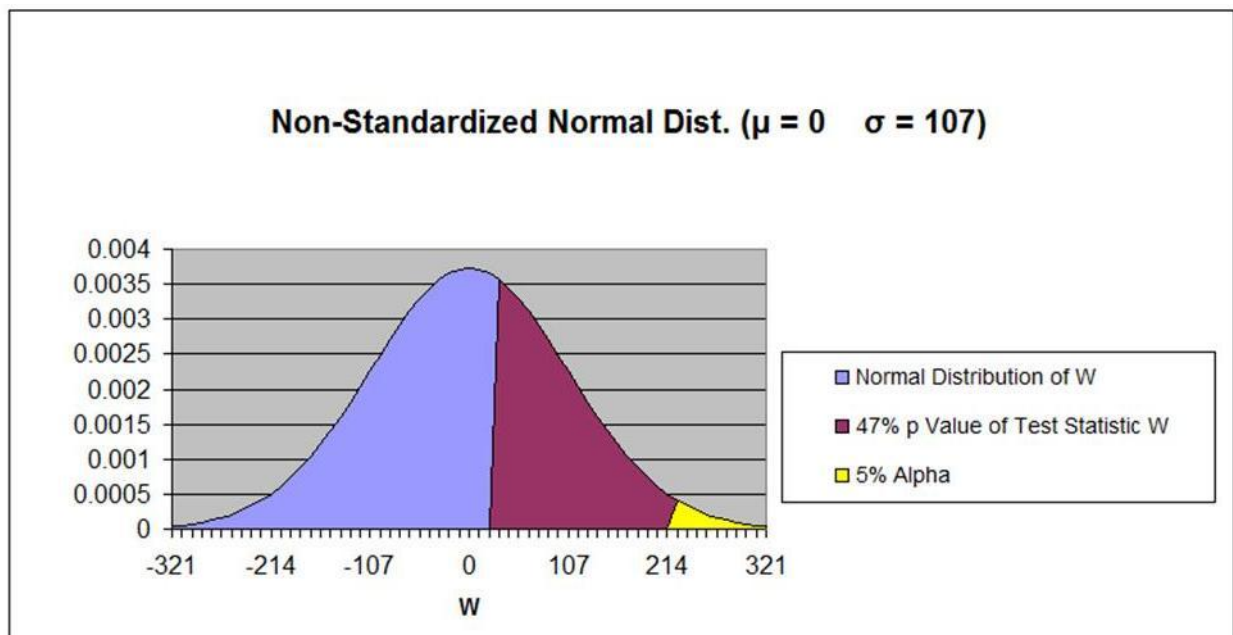
Statistical Mistakes You Don't Want To Make

1) Assuming that correlation equals causation

This is, of course, not true. However, if you find a correlation, you should look hard for links between the two. The correlation may be pure chance, but then again, it may not be. A correlation is a reason to look for underlying causes behind the behavior. Correlation is often a symptom of a larger issue.

2) Not graphing and eyeballing the data prior to performing regression analysis

Always graph the data before you do regression analysis. You'll know immediately whether you're dealing with linear regression, non-linear regression, or completely unrelated data that can't be regressed.

3) Not doing correlation analysis on all variables prior to performing regression

You'll save yourself a lot of time if you can remove any input variables that have a low correlation with the dependent
(output –> Y) variable or that have a high correlation with another input variable (this error is called multicollinearity). In the 2nd case, you would want to remove the input variable from the highly correlated pair of input variables that has the lowest correlation with the output variable.

4) Adding a large number of new input variables into a regression analysis all at once

You always want to add new input variables one at a time and run a separate regression each time a new input variable is added. The changes you observe to the output of the regression will tell you whether the new input variable adds to the predictive power of the regression equation. Adjusted r squared only increases when a new variable adds greater predictive power to the regression equation.

5) Applying input variables to a regression equation that are outside of the value of the original input variables that were used to create the regression equation

Here is an example to illustrate why this might produce totally invalid results. Suppose that you created a regression equation that predicted a child's weight based upon the child's age, and then you provided an adult age as an input. This regression equation would predict a completely incorrect weight for the adult, because adult data was not used to construct the original regression equation.

6) Not examining the residuals in regression

You should always at least eyeball the residuals. If the residuals show a pattern, your regression equation is not explaining all of the behavior of the data.

7) Only evaluating r square in a regression equation

In the output of regression performed in Excel, there are actually four very important components of the output that should be looked at. There is an article in my blog: http://blog.excelmasterseries.com/2010/03/how-to-quickly-read-output-of-excels.html that covers this topic in a lot more detail than could be done in this bullet point.

8) Not drawing a representative sample from a population

This is usually solved by taking a larger sample and using a random sampling technique such as nth-ing (sampling every nth object in the population).

9) Drawing a conclusion without applying the proper statistical analysis

This occurs quite often when people simply eyeball the results instead of performing a hypothesis test to determine if the observed change has at least an 80% chance (or whatever level of certainty you desire) of not being pure chance.

10) Drawing a conclusion before a statistically significant result has been reached

This is often caused by choosing a statistical test requiring a lot of samples but depending on a low sample rate. A common occurrence of this would be performing multivariate testing on a web site that does not have sufficient traffic. Such a test is likely to be concluded prematurely. A better solution might be to perform a number of successive A/B split-tests in place of multivariate analysis. You get a lot more testing done a lot faster, and correctly.

11) Analyzing non-normal data with the normal distribution

Data should always be eyeballed in a Histogram and then and analyzed for normality before using the normal distribution-based testing techniques. If the data is not normally distributed, you must use data fitting techniques to determine which statistical distribution most closely fits the data.

12) Not removing outliers prior to statistical analysis

A couple of outliers can skew results badly. Once again, eyeball the data and determine what belongs and what doesn't. Another solution is to test the data using nonparametric analysis. Many nonparametric tests are not significantly affected by a few outliers.

13) Not controlling or taking into account other variables besides the one(s) being testing when using the t test, ANOVA, or hypothesis tests.

Other variables that not part of the test need to be held as constant as possible during the above tests or your answer might be invalid without you realizing it.

14) Using the wrong t test

The t-test to be applied depends upon factors such as whether or samples have the same size and variance. It is important to pick the right t-test before starting. Read the t-Test chapters in this e-manual.

15) Attempting to apply the wrong type of hypothesis test

There are 4 ways that the data must be classified before the correct hypothesis test can be selected. Another article in this blog discusses this. Also, Chapters 8 and 9 of the Excel Statistical Master: http://excelmasterseries.com/ClickBank/Download_ESM_1234_Site_3.php provide clear, detailed instructions on how to analyze your data prior to hypothesis test selection. You probably wouldn't get far into a hypothesis test if you have incorrectly classified the data and selected the wrong hypothesis test.

16) Not using Excel

This point may sound a little self-serving, but knowing how to do this stuff in Excel is a real time-saver, particularly if you are in marketing, and especially if you're an Internet marketer. You'll never need to pick up another thick confusing statistics text book or figure how to work those confusing statistics tables ever again. I've actually thrown out all of my statistics text books (well, not quite, I sold them on eBay).

17) Always requiring 95% certainty

This could really slow you down. For example, if I'm A/B split-testing keywords or ads in an AdWords campaign, I will typically pick a winner when my split-tester tells me that it is 80% sure that one result is better than the other. Achieving 95% certainty would often take too long.

18) Thinking it is impossible to get a statistically significant sample if your target market is large

The sample size you need from a large population is probably quite a bit smaller than you think. Nationwide surveys are normally within a percentage point or two from real answer after only several thousand interviews have been conducted. That of course depends hugely on obtaining a representative sample to interview.

19) Not taking steps to ensure that your sample is normally distributed when analyzing with the normal distribution

One way to ensure that you have a normally distributed sample for analysis is to take a large number of samples (at 30) with each sample consisting of several random and simultaneously-chosen data points and then take the mean from each sample. Make that mean the sample. Your samples will now be Normally distributed. You can take as few as 2 data points per sample, but the more data points per sample, the fewer data points it will take for your samples (each sample is the mean of the data points collected for one sample) to appear to be Normally distributed. If you are taking only 3 data points per sample, you may have to collect over samples (that would be a total of 300 data points) for your samples to appear to be Normally distributed.

   http://excelmasterseries.com/New_Manuals.php

A statistical theory called the Central Limit Theory states that the means of samples (at least 30 samples and each sample having at least 2 data points that are averaged to get a mean, which will be the value of the sample) will be Normally distributed, no matter how the underlying population is distributed. You can then perform statistical analysis on that group using the normal distribution-based techniques.

20) Using covariance analysis instead of correlation analysis

The output of covariance analysis is dependent upon the scale used to measure the data. Different scales of measurement can produce completely different results on the same data if covariance analysis is used. Correlation analysis is completely independent of the scale used to measure the data. Different scales of measurement will produce the same results on a data set using correlation analysis, unlike covariance analysis.

21) Using a one-tailed test instead of a two-tailed test when accuracy is needed

If accuracy it what you are seeking, it might be better to use the two-tailed when performing, for example, a hypothesis test. The two-tailed test is more stringent than the one-tailed test because the outer regions (I refer to them the regions of uncertainty) are half the size in a two-tailed test than in a one-tailed test. The two-tailed test tells you merely that the means are different. The one-tailed test tells you that the means are different in one specific direction.

22) Not using nonparametric tests when analyzing small samples of unknown distribution

The t Distribution should only be used in small sample analysis if the population from which the samples were drawn was Normally distributed. Nonparametric tests are valid when the population distribution is not known, or is known not to be Normally distributed. Using the t distribution in either of these cases for small sample analysis is invalid.

Correctable Reasons Why Your

Data Is Not Normally Distributed

In the ideal world, all of your data samples are normally distributed. In this case you can usually apply the well-known parametric statistical tests such as ANOVA, the t Test, and regression to the sampled data.

What can you do if your data does not appear to be normally distributed?

You can either:

- Apply nonparametric tests to the data. Nonparametric tests do not rely on the underlying data to have any specific distribution

- Evaluate whether your "non-normal" data was really normally- distributed before it was affected by one of the seven correctable causes listed below:

The Biggest 7 Correctable Causes of Non-Normality in Data Samples

1) Outliers – Too many outliers can easily skew normally-distributed data. If you can identify and remove outliers that are caused by error in measurement or data entry, you might be able to obtain normally-distributed data from your skewed data set. Outliers should only be removed if a specific cause of their extreme value is identified. The nature of the normal distribution is that some outliers will occur. Outliers should be examined carefully if there are more than would be expected.

2) Data has been affected by more than one process – It is very important to understand all of the factors that can affect data sample measurement. Variations to process inputs might skew what would otherwise be normally-distributed output data. Input variation might be caused by factors such as shift changes, operator changes, or frequent changes in the underlying process. A common symptom that the output is being affected by more than one process is the occurrence of more than one mode (most commonly

occurring value) in the output. In such a situation, you must isolate each input variation that is affecting the output. You must then isolate the overall effect which that variation had on the output. Finally, you must remove that input variation's effect from output measurement. You may find that you now have normally-distributed data.

3) Not enough data – A normal process will not look normal at all until enough samples have been collected. It is often stated that 30 is the where a "large" sample starts. If you have collected 50 or fewer samples and do not have a normally-distributed sample, collect at least 100 samples before re-evaluating the normality of the population from which the samples are drawn.

4) Measuring devices that have poor resolution – Devices with poor resolution may round off incorrectly or make continuous data appear discrete. You can, of course, use a more accurate measuring device. A simpler solution is to use a much larger sample size to smooth out sharp edges.

5) A different distribution describes the data – Some forms of data inherently follow different distributions. For example, radioactive decay is described by the exponential distribution. The Poisson distribution describes events event that tend to occur at predictable intervals over time, such as calls over a switchboard, number of defects, or demand for services. The lengths of time between occurrences of Poisson-distributed processes are described by the exponential distribution. The uniform distribution describes events that have an equal probability of occurring. Application of the Gamma distribution often based on intervals between Poisson-distributed events, such as queuing models and the flow of items through a manufacturing process. The Beta distribution is often used for modeling planning and control systems such are PERT and CPM. The Weibull distribution is used extensively to model time between failure of manufactured items, finance, and climatology. It is important to become familiar with the applications of other distributions. If you know that the data is described by a different distribution than the normal distribution, you will have to apply the techniques of that distribution or use nonparametric analysis techniques.

6) Data approaching zero or a natural limit – If the data has a large number of value than are near zero or a natural limit, the data may appear to be skewed. In this case, you may have to adjust all data by adding a specific value to all data being analyzed. You need to make sure that all data being analyzed is "raised" to the same extent.

7) Only a subset of process' output is being analyzed – If you are sampling

only a specific subset of the total output of a process, you are likely not collecting a representative sample from the process and therefore will not have normally distributed samples. For example, if you are evaluating manufacturing samples that occur between 4 and 6AM and not an entire shift, you might not obtain the normally-distributed sample that a whole shift would provide. It is important to ensure that your sample is representative of an entire process.

If you are unable to obtain a normally-distributed data sample, you can usually apply non-parametric tests to the data.

Meet the Author

Mark Harmon is a master number cruncher. Creating overloaded Excel spreadsheets loaded with complicated statistical analysis is his idea of a good time. His profession as an Internet marketing manager provides him with the opportunity and the need to perform plenty of meaningful statistical analysis at his job.

Mark Harmon is also a natural teacher. As an adjunct professor, he spent five years teaching more than thirty semester-long courses in marketing and finance at the Anglo-American College in Prague and the International University in Vienna, Austria. During that five-year time period, he also worked as an independent marketing consultant in the Czech Republic and performed long-term assignments for more than one hundred clients. His years of teaching and consulting have honed his ability to present difficult subject matter in an easy-to-understand way.

Harmon received a degree in electrical engineering from Villanova University and MBA in marketing from the Wharton School.