Notes:

Nonparametric Tests

Key Learning Points

1. Describe the importance of using nonparametric tests.
2. Explain why to use nonparametric tests.
3. Utilize nonparametric tests in improvement projects.

What are Nonparametric Tests?

Many statistical tests used in Six Sigma projects require normally distributed data to ensure an accurate result. Unfortunately not all data follows the bell curve of the normal distribution.

A statistical method is nonparametric when it does not rely on computing the parameters of a specific probability distribution.

Nominal

Nonparametric methods can be used for data with a nominal (classes or categories) measurement scale (e.g. red, yellow, blue).

Ordinal

Nonparametric methods can be used for data with an ordinal measurement scale (ordered data). Only comparisons of "greater," "less," or "equal" between measurements are possible (e.g. low, medium, high).

Continuous

Nonparametric methods can be used for continuous data when the distribution is either unknown or is known to violate an assumption for one of the parametric methods (e.g. is not normally distributed).

Why Study Nonparametric Tests

Nonparametric tests should be used for the following reasons:

- Non-normal populations can lead to misleading results, biased predictions, and invalid confidence intervals.
- Some data only specify order or counts of events. There is not enough information to use traditional methods, and nonparametric methods are the only possibility.
- Medians may be more useful representations of central tendency than means in the presence of skewed populations and extreme outliers.

Test Usage and Comparison

Parametric Statistical Test	Nonparametric Equivalents	Nonparametric Distribution Assumptions
1 Sample-T	1 Sample Sign	Random and independent bivariate variables. Robust against outliers but need large sample size to be effective.
	1 Sample Wilcoxon	Random, independent sample is from a population with a symmetric distribution.
2 Sample-T	Wilcoxon Mann-Whitney	Independent random samples from two populations that have the same shape and equal variances.
Paired-T	Wolcoxon (Paired Samples)	Independent random samples with symmetric distributions.

Notes:

Notes:

Parametric Statistical Test	Nonparametric Equivalents	Nonparametric Distribution Assumptions
One-Way ANOVA	Kruskal-Wallis	Samples (at least 5 measures) are random and mutually independent from populations whose distribution functions have the same shape and equal variances. Kruskal-Wallis is more powerful than Mood's Median for data from many distributions, but less robust against outliers.
	Mood's Median	Independent random samples from population distributions that have the same shape. Mood's Median test is robust against outliers.

1 Sample Tests

1-Sample Sign test is useful for outliers and non-symmetrical distributions. It requires larger sample sizes for accuracy.

The 1-Sample Wilcoxon test is more robust, for either small or large sample sizes, and is the preferred test. This, however, is not a good test if there are many outliers.

1 Sample Sign Test

What is a 1 Sample Sign Test?

A 1-Sample Sign test is a statistical test used to compare a sample MEDIAN to a specified target. The target may be based on performance standards, process targets, or historical data.

The underlying distribution does not have to be normal or symmetrical for this test to work. For symmetrical distributions and small sample sizes, the Wilcoxon test is more powerful.

The 1-Sample Sign test is more useful for larger sample sizes. It requires large samples for more power. This test is especially robust when there are several outliers.

A 1-Sample Sign test is a simple rank test. It places the data (sample size n) in ascending order.
It then counts the number of data points that are above the tested median and the number below the tested median.

Steps in Hypothesis Testing

1. State the practical problem.
2. Establish the hypotheses.
3. Decide on appropriate statistical test.
4. Set the alpha level.
5. Determine initial sample size. Calculate the power.
6. Develop the sampling plan and collect the data.
7. Use the appropriate graphical tool to explore the data.
8. Check data assumptions (change choice of tool if appropriate).
    a. Not Normal
9. Run the statistical test and determine the statistical conclusion.
10. Translate statistical conclusion to a practical conclusion.

Example: Boston Suburb Housing Prices

Test the sample price of Boston Suburb housing against the median housing price across the USA to discover if the Boston Suburb housing has a comparable median housing price to the USA median of $206K.

Remember that for a 1 Sample Sign test, data do not have to be sorted and it does not matter if there are many outliers. This test requires a large sample to detect significant differences.

Step 1: State the Practical Problem

The Practical Problem:

Do Boston suburb houses have a comparable median price to the USA median of $206,000?

Step 2: Establish the Hypotheses

H_o: Median Price Boston = Median Price USA

H_a: Median Price Boston ≠ Median Price USA

Step 3: Decide on Appropriate Statistical Test

Since the data in the sample of Boston Suburb Housing Prices are not normal, use a nonparametric test comparing a sample to a target, a 1 Sample Sign Test.

Step 4: Set the Alpha Level

You want to be 95% confident that you are making the correct conclusion.
$\alpha = 0.05$

Notes:

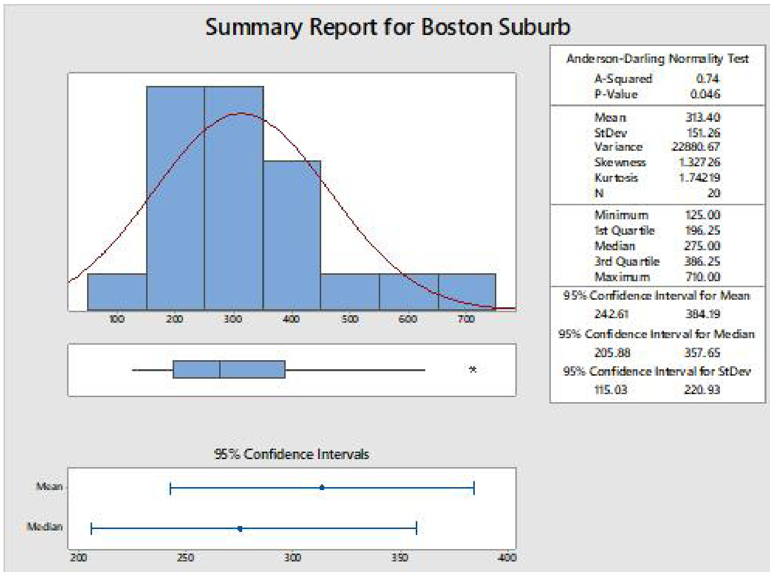
Steps 5&6: Set the Power and Sample Size, and Collect the Data

A sample of the sale prices of 20 houses was randomly collected over the course of a year in a Boston suburb.

Step 7: Use the Appropriate Graphical Tool To Explore the Data

The following graphical summary was created from the sample data.

Graphical Summary



Step 8: Check Data Assumptions

The data is indeed not normally distributed. Since a sample is being compared to a target, a 1 Sample Sign test should still be used.

Step 9: Run the Statistical Test

- **Minitab: Stat > Nonparametrics > 1 Sample Sign**
  - Variable: Boston Suburb
  - Test Median: 206
  - Alternative: Not Equal

Notes:

Results

Sign Test for Median: Boston Suburb

Method

η: median of Boston Suburb

Descriptive Statistics

Sample	N	Median
Boston Suburb	20	275

Test

Null hypothesis	H_0: η = 206
Alternative hypothesis	H_1: η ≠ 206

Sample	Number < 206	Number = 206	Number > 206	P-Value
Boston Suburb	6	0	14	0.115

Statistical Conclusion

The P-Value (0.115) for the Boston suburb median pricing is higher than alpha (0.05), therefore, fail to reject the null hypothesis.

Step 10: Translate the Statistical Conclusion Into a Practical Conclusion

Based on the 1 Sample Sign test you can conclude that Boston suburb median housing pricing is not significantly different than $206,000, the USA average.

Wilcoxon

What is a 1 Sample Wilcoxon Test?

A 1-Sample Wilcoxon test is a statistical test used to compare a sample median to a target value in order to identify if there is a significant difference between them.

This test can be used when the sample data are not normally distributed, but it should be from a population whose measurement scale is at least ordinal or continuous and the distribution is assumed to be symmetrical.

The 1-Sample Wilcoxon is a robust test for non-normal data, especially for smaller sample sizes, but is not a good test to use if there are many outliers.

An assumption for the 1-sample Wilcoxon test and confidence interval is that the data are a random sample from a continuous, symmetric population.

When the population is normally distributed, this test is slightly less powerful (the confidence interval is wider, on the average) than the t-Test. It may be considerably more powerful (the confidence interval is narrower, on the average) for other populations.

Theory: Wilcoxon

A 1-Sample Wilcoxon test is a rank test. It assigns ranks to the data (sample size n)

Notes:

based upon the difference each point is from the median being tested against.

Arrange differences in absolute value ascending order and assign ranks (1 for smallest, n for the largest).

The ranks are then given positive or negative signs corresponding to the difference and the sum of either positive or negative ranks. They are then compared to a Wilcoxon Quantile Table (with α values for different sample sizes (n)).

Expect the sum of positive and negative ranks to be similar if the data are drawn from a population that has a median which is not statistically different from the target median.

The estimated median is based upon the median value of Walsh averages of the sample data, and the CI around the median is based upon Walsh averages and Wilcoxon probabilities for 2-tailed tests.

Calculating Wilcoxon Statistic (W)

The Wilcoxon Statistic distribution can be approximated by the normal using the calculation: $z = W / \sqrt{[n(n+1)(2n+1)/6]}$

1. Subtract the test median from each sample data point.
2. Take absolute value of these differences in rank order.
3. Assign ranks to the absolute value and sum the ranks for both positive and negative values in separate columns.
4. Either of the summed values is W and can be compared with special Z tables for Wilcoxon Signed Rank (or MINITAB® will do it for you at your specified confidence level).
5. If the smaller W value is less than the value in the table, then reject the null hypothesis.

Using a Wilcoxon Table:

1. Locate the sample size n in the left column.
2. Locate a level in the top column (≠ would be 2 tail and > or < would be 1 tail).
3. Where they intersect is the lowest value you expect randomly. If your calculated
4. alue of W is less than the critical value from the table, the test is significant—that is, you reject the null.

Notes:

Wilcoxon Table

Table 10 Critical values of the Wilcoxon signed-rank statistic

One tail Two tail	10% 20%	5% 10%	2.5% 5%	1% 2%	0.5% 1%
n					
3	0				
4	1	0			
5	2	1	0		
6	4	2	1	0	
7	6	4	2	0	0
8	8	6	4	2	0
9	11	8	6	3	2
10	14	11	8	5	3
11	18	14	11	7	5
12	22	17	14	10	7
13	26	21	17	13	10
14	31	26	21	16	13
15	37	30	25	20	16
16	42	36	30	24	19
17	49	41	35	28	23
18	55	47	40	33	28
19	62	54	46	38	32
20	70	60	52	43	37

Theory:Walsh Averages

Walsh Averages are used to compute estimated medians and Confidence Intervals for sample data in a 1-Sample Wilcoxon test.

There are n(n+1)/2 Walsh Averages for any dataset.

Example: Scores from five golfers on one hole : (4, 5, 5, 6, 9).

With n=5, there are 15 Walsh Averages.

Notes:

Notes:

Data	Walsh Calculation	Walsh Average
4	(4+4)/2	4.0
5	(4+5)/2	4.5
5	(4+5)/2	4.5
6	(4+6)/2	5.0
9	(4+9)/2	6.5
	(5+5)/2	5.0
	(5+5)/2	5.0
	(5+6)/2	5.5
	(5+9)/2	7.0
	(5+5)/2	5.0
	(5+6)/2	5.5
	(5+9)/2	7.0
	(6+6)/2	6.0
	(6+9)/2	7.5
	(9+9)/2	9.0

The median of the dataset is 5.0 while the median of the Walsh Averages is 5.5

The Confidence Interval of the Walsh averages is based upon Quantiles of the Wilcoxon Statistic for a and n.

For example, with n = 5, the lower 80% CI (α/2 is 10) for the median would be the 3rd Walsh Average. The lower CI is = 4.5.

Quantiles of Wilcoxon Signed

	$W_{0.005}$	$W_{0.01}$	$W_{0.025}$	$W_{0.05}$	$W_{0.10}$
n = 4	0	0	0	0	1
5	0	0	0	1	3
6	0	0	1	3	4
7	0	1	3	4	6
8	1	2	4	6	9
9	2	4	6	9	11
10	4	6	9	11	15

Steps in Hypothesis Testing

1. State the practical problem.
2. Establish the hypotheses.
3. Decide on appropriate statistical test.
4. Set the alpha level.
5. Determine initial sample size. Calculate the power.
6. Develop the sampling plan and collect the data.
7. Use the appropriate graphical tool to explore the data.
8. Check data assumptions (change choice of tool if appropriate).

Notes:

a. Not Normal

9. Run the statistical test and determine the statistical conclusion.

10. Translate statistical conclusion to a practical conclusion.

Example: Boston Suburb Housing Prices

Return to the Boston suburb median housing pricing example.

Test the sample price of Boston Suburb housing against the median housing price across the USA to discover if the Boston Suburb housing has a comparable median housing price to the USA median of $206K.

Step 1: State the Practical Problem

The Practical Problem:
Do Boston suburb houses have a comparable median price to the USA median of $206,000?

Step 2: Establish the Hypotheses

H_o: Median Price $_{\text{Boston}}$ = Median Price $_{\text{USA}}$

H_a: Median Price $_{\text{Boston}}$ ≠ Median Price $_{\text{USA}}$

Step 3: Decide on Appropriate Statistical Test

Since the data in the sample of Boston Suburb Housing Prices are not normal, use a nonparametric test comparing a sample to a target, a 1 WilcoxonTest.

Step 4: Set the Alpha Level

You want to be 95% confident that you are making the correct conclusion.
$\alpha = 0.05$

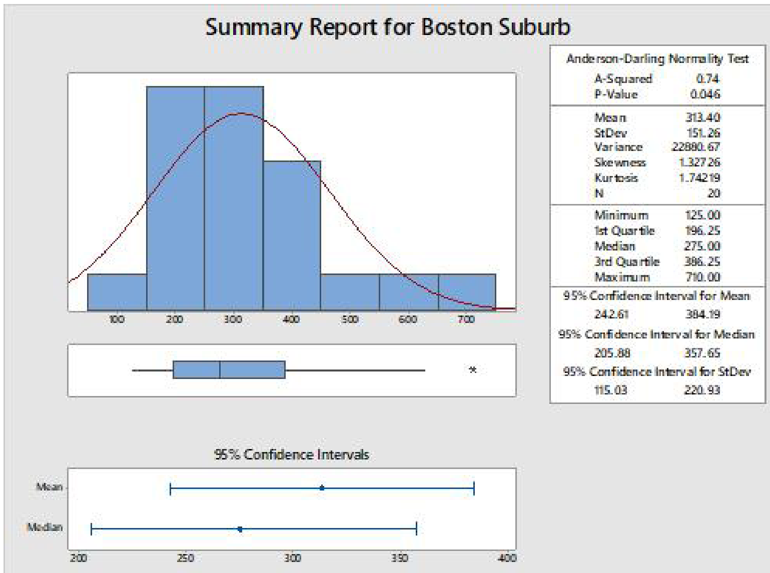
Steps 5&6: Set the Power and Sample Size, and Collect the Data

A sample of the sale prices of 20 houses was randomly collected over the course of a year in a Boston suburb.

Step 7: Use the Appropriate Graphical Tool To Explore the Data

The following graphical summary was created from the sample data.

Graphical Summary



Step 8: Check Data Assumptions

The data are indeed not normally distributed. Since a sample is being compared to a target, a 1 Sample Wilcoxon test should be used.

Step 9: Run the Statistical Test

- **Minitab: Stat > Nonparametrics > 1 Sample Wilcoxon**
  - Variable: Boston Suburb
  - Test Median: 206
  - Alternative: Not Equal

Results

Wilcoxon Signed Rank Test: Boston Suburb

Method

η: median of Boston Suburb

Descriptive Statistics

Sample	N	Median
Boston Suburb	20	292.5

Test

Null hypothesis	H_0: η = 206
Alternative hypothesis	H_1: η ≠ 206

Sample	N for Test	Wilcoxon Statistic	P-Value
Boston Suburb	20	180.00	0.005

Notes:

Statistical Conclusion

The P-Value (0.005) for the Boston suburb median pricing is lower than alpha (0.05), therefore, reject the null hypothesis.

Step 10: Translate the Statistical Conclusion Into a Practical Conclusion

Based on the 1 Sample Wilcoxon test we conclude that Boston suburb median housing pricing is significantly different than $206,000, the USA average.

2 Sample Tests

Feature	Mann-Whitney	Mood's Median
Use with. . .	Samples with similar distributions and few outliers	Any distribution, with several outliers, but requires large samples
Test Statistic	Wilcoxon W	Chi-squared based on expected equality of sum of ranks
Power	Relatively robust, preferred unless several outliers are present	Relatively weak
Confidence Interval	On median difference based on Mann-Whitney U Statistic	On actual median using binomial distribution
Median Estimate	Estimate of the difference between two medians with median of pairwise differences in sample	Actual median of sample

Mann-Whitney

What is a Mann-Whitney Test?

The Mann-Whitney test is a statistical test used to compare the medians from two random samples in order to determine if they are from the same population (i.e. the medians are equal) or from different populations (the medians are different).

It requires:

- The measurement scale is at least ordinal
- The two sample distributions are of similar shape
- The two variances are not significantly different

But...

- Does not assume normality
- Distributions do not need to be symmetrical

Notes:

- The two samples do not need to be the same size

MINITAB® uses the Wilcoxon formulation to analyze Mann-Whitney and is often called the Wilcoxon Mann-Whitney.

Theory: Mann-Whitney

MINITAB® uses the Wilcoxon formulation to analyze Mann-Whitney and is often called the Wilcoxon Mann-Whitney.

Wilcoxon Mann-Whitney is a rank test. It combines both samples (n and m) and assigns ranks to the data (sample size n + m). Samples can be different sizes.

Key Assumptions for Mann-Whitney Test

- Key assumptions for the Mann-Whitney test are that:
- The data are independent random samples
- The distributions for the two samples have the same shape
  - Their variances are not significantly different from each other
  - Both have a scale that is continuous or ordinal

The 2-sample Mann-Whitney rank test is slightly less powerful (the confidence interval is wider on the average) than the 2-sample t-Test with pooled sample variance when the populations are normal, but it is considerably more powerful (confidence interval is narrower, on the average) for many other distributions.

Example

The following table lists the number of errors discovered by location (Liverpool and Bournemouth). Discover if the error rate for B/Mouth is > Liverpool at a of 0.05.

L/Pool	164	212	132	140	116	104	167	
B/ Mouth	208	246	197	153	118	169	120	144

To solve this problem, first rank each measurement in order from lowest to highest.

104	116	118	120	132	140	144	153	164	167	169	197	208	212	146
1	2			5	6			9	10				14	
		3	4			7	8			11	12	13		15

Notes:

Next, separate out the two samples (m and n), and sum the ranks for each sample.

L/Pool	164	212	132	140	116	104	167	
S_m	9	14	5	6	2	1	10	
B/ Mouth	208	246	197	153	118	169	120	144
S_n	13	15	12	8	3	11	4	7

The Sum of the Ranks for Liverpool, S_m = 9+14+5+6+2+1+10 = 47
The Sum of the Ranks for Bournemouth, S_n = 13+15+12+8+3+11+4+7=73

Now compare the summed ranks to a Wilcoxon table with α values for different sample sizes (n+m).

TABLE O Critical Values of Smaller Rank Sum for the Wilcoxon-Mann-Whitney Test*

			n_1 (smaller sample)											
n_2	α for 2-sided test	α for 1-sided test	1	2	3	4	5	6	7	8	9	10	11	12
3	0.20	0.10		3	7									
	0.10	0.05			6									
	0.05	0.025												
	0.01	0.005												
4	0.20	0.10		3	7	13								
	0.10	0.05			6	11								
	0.05	0.025				10								
	0.01	0.005												
5	0.20	0.10		4	8	14	20							
	0.10	0.05		3	7	12	19							
	0.05	0.025			6	11	17							
	0.01	0.005					15							
6	0.20	0.10		4	9	15	22	30						
	0.10	0.05		3	8	13	20	28						
	0.05	0.025			7	12	18	26						
	0.01	0.005				10	16	23						
7	0.20	0.10		4	10	16	23	32	41					
	0.10	0.05		3	8	14	21	29	39					
	0.05	0.025			7	13	20	27	36					
	0.01	0.005				10	16	24	32					
8	0.20	0.10		5	11	17	25	34	44	55				
	0.10	0.05		4	9	15	23	31	41	51				
	0.05	0.025		3	8	14	21	29	38	49				
	0.01	0.005				11	17	25	34	43				
9	0.20	0.10	1	5	11	19	27	36	46	58	70			
	0.10	0.05		4	9	16	24	33	43	54	66			
	0.05	0.025		3	8	14	22	31	40	51	62			
	0.01	0.005			6	11	18	26	35	45	56			
10	0.20	0.10	1	6	12	20	28	38	49	60	73	87		
	0.10	0.05		4	10	17	26	35	45	56	69	82		
	0.05	0.025		3	9	15	23	32	42	53	65	78		
	0.01	0.005			6	12	19	27	37	47	58	71		
11	0.20	0.10	1	6	13	21	30	40	51	63	76	91	106	
	0.10	0.05		4	11	18	27	37	47	59	72	86	100	
	0.05	0.025		3	9	16	24	34	44	55	68	81	96	
	0.01	0.005			6	12	20	28	38	49	61	73	87	
12	0.20	0.10	1	7	14	22	32	42	54	66	80	94	110	127
	0.10	0.05		5	11	19	28	38	49	62	75	89	104	120
	0.05	0.025		4	10	17	26	35	46	58	71	84	99	115
	0.01	0.005			7	13	21	30	40	51	63	76	90	105

*Reproduced with permission from Tate, M. W. and Clelland, R. C. (1957). *Non-parametric and Shortcut Statistics*. The Interstate Printers & Publishers, Danville, IL.

Notes:

Notes:

The critical value for m=7, n=8 is 41 for the 1-tailed test with $\alpha = 0.05$. $47 > 41$ so there is no evidence to reject H_o.

Steps in Hypothesis Testing

1. State the practical problem.
2. Establish the hypotheses.
3. Decide on appropriate statistical test.
4. Set the alpha level.
5. Determine initial sample size. Calculate the power.
6. Develop the sampling plan and collect the data.
7. Use the appropriate graphical tool to explore the data.
8. Check data assumptions (change choice of tool if appropriate).
   a. Not Normal
9. Run the statistical test and determine the statistical conclusion.
10. Translate statistical conclusion to a practical conclusion.

Changing Processes

In this example you want to test whether changing from Process 1 to Process 2 reduces the amount of material lost per batch.

For a Mann-Whitney test you need 2 samples:

- Not necessarily equal
- Each with similar continuous distributions, but not necessarily normal or symmetric

The batches are the results before and after the change.

Note: There is no easy and exact way to determine whether the two distributions are similar. The easiest method is to plot both. If the two shapes are not substantially different, then it is okay to proceed. There is some judgment in this matter and exact statistical tests are both complex and unnecessary at this point.

Step 1: State the Practical Problem

The Practical Problem:

You want to compare the medians of two processes to see if one reduces loss.

Notes:

Step 2: Establish the Hypotheses

- H_o: Median 1 = Median 2
- H_a: Median 1 > Median 2

or

- H_a: Process 2 reduces loss

Step 3: Decide on Appropriate Statistical Test

Since the data is not normal, test the medians, not the means of the two sets of data. The statistical test to use is a Mann-Whitney Test.

Step 4: Set the Alpha Level

You want to be 95% confident that you are making the correct conclusion.
$\alpha = 0.05$

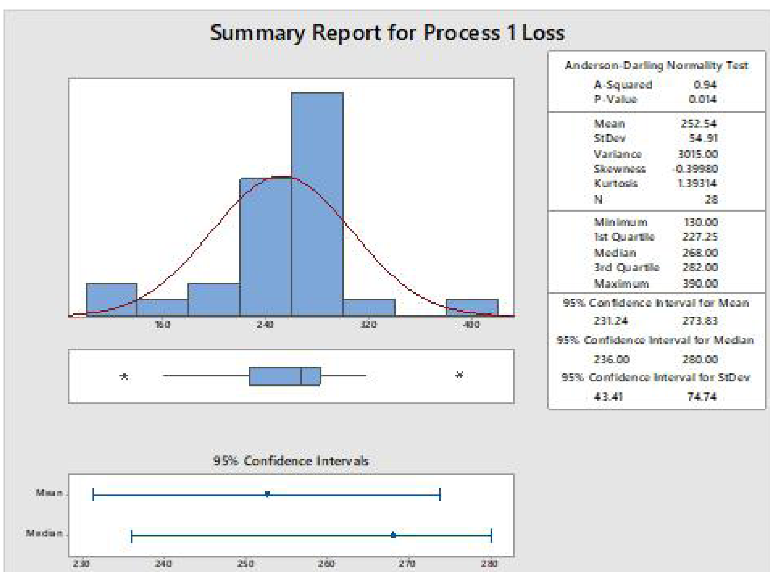
Steps 5&6: Set the Power and Sample Size, and Collect the Data

You have data from Process 1 and you need to collect data from Process 2.
You do not need the same number of data points.

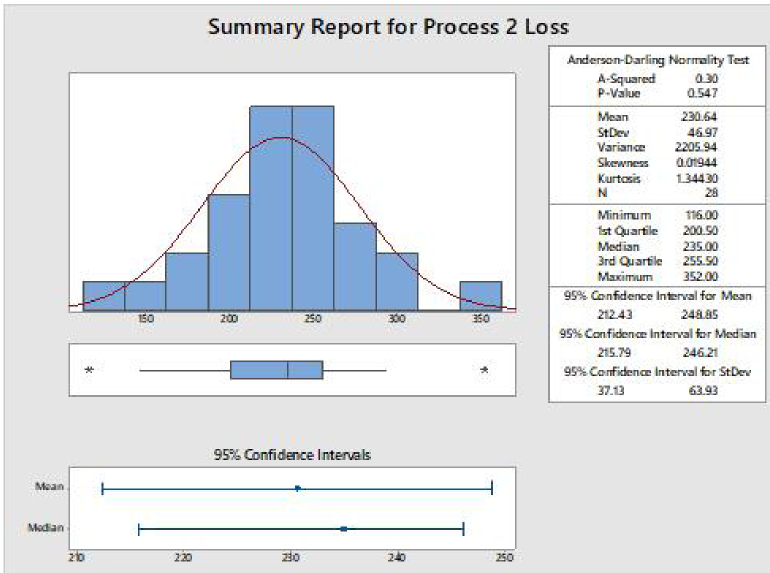
Step 7: Use the Appropriate Graphical Tool To Explore the Data

The following graphical summaries were created from the sample data. The box plots compare process 1 to process 2.

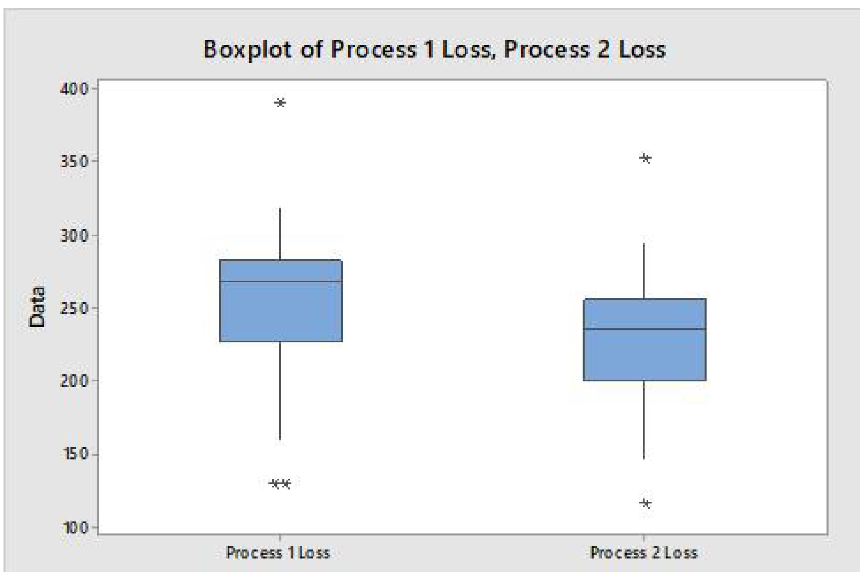
Graphical Summary Process 1



Graphical Summary Process 2



Box Plot of Process 1 Loss, vs. Process 2 Loss



Step 8: Check Data Assumptions

The assumption is that the data do not have to be normal, but the two processes must have similar distribution shapes.

The distribution for Process 1 is normal, but the distribution for Process 2 is not. The two distributions are shaped in a similar manner, so the Mann-Whitney test should be used.

Step 9: Run the Statistical Test

- **Minitab: Stat > Nonparametrics > Mann-Whitney**
  - Select First Sample “Process 1 Loss” and Second Sample “Process 2

Notes:

Loss”

- Confidence level at “95%”
- Select Alternative: “greater than”

Results

Mann-Whitney: Process 1 Loss, Process 2 Loss

Method

η_1: median of Process 1 Loss
η_2: median of Process 2 Loss
Difference: $\eta_1 - \eta_2$

Descriptive Statistics

Sample	N	Median
Process 1 Loss	28	268
Process 2 Loss	28	235

Estimation for Difference

Difference	Lower Bound for Difference	Achieved Confidence
28	4	95.02%

Test

Null hypothesis $H_0: \eta_1 - \eta_2 = 0$
Alternative hypothesis $H_1: \eta_1 - \eta_2 > 0$

Method	W-Value	P-Value
Not adjusted for ties	921.50	0.022
Adjusted for ties	921.50	0.022

Statistical Conclusion

The P-Value (0.0219) is lower than alpha (0.05), therefore, reject the null hypothesis. Process 1 > Process 2

Step 10: Translate the Statistical Conclusion Into a Practical Conclusion

The Sample 1 Median is greater than the Sample 2 Median. This means that Process 2 does reduce the loss.

Mood’s Median

What is a Mood’s Median Test?

The Mood’s Median test is a statistical test used to compare two or more sample medians in order to make inferences about the population medians (equal or not). It does not require the distributions to be normal or symmetrical, and is especially useful if there are several outliers. The measurement scale can be ordinal or continuous. This test does not work well with small samples and is not as robust as the Wilcoxon Mann-Whitney test or the Kruskal-Wallis test.

- MINITAB® uses the Chi-Square Tables and binomial probabilities to analyze the difference in medians from multiple samples.
- The median from all sample data is calculated.
- Mood’s Median is similar to the Sign test. It is a Simple Sign test that

Notes:

places data from all samples (n, m, …) in rank order.

- The count of the number of data points that are above and below the total sample median for each sample is placed in a Chi-Square Table.
- The Analysis is based upon proportions of events, that are above and below, against an hypothesized standard of 50%. The expected number above and below the overall median would be proportionally the same for all samples if no difference exists.
- The Confidence Interval around the median is calculated using binomial probabilities.

Steps in Hypothesis Testing

1. State the practical problem.
2. Establish the hypotheses.
3. Decide on appropriate statistical test.
4. Set the alpha level.
5. Determine initial sample size. Calculate the power.
6. Develop the sampling plan and collect the data.
7. Use the appropriate graphical tool to explore the data.
8. Check data assumptions (change choice of tool if appropriate).
   a. Not Normal
9. Run the statistical test and determine the statistical conclusion.
10. Translate statistical conclusion to a practical conclusion.

Changing Processes

In this example you want to test whether changing from Process 1 to Process 2 reduces the amount of material lost per batch. Mood's Median test will test if Process 2 and Process 1 have different medians and therefore come from different populations.

For a Mood's Median test you need 2 samples:

- Not necessarily equal
- The samples can come from any distribution and symmetry is not a requirement.

The batches are the results before and after the change.

Batches

Test that Process 2 has been improved and that its median loss is less than the

Notes:

losses from Process 1.

Process 1 Loss	Process 2 Loss
270	218
236	234
225	214
130	116
280	200
272	276
160	146
220	182
130	238
242	288
186	190
266	236
216	244
318	258
294	240
282	294
234	220
258	200
276	220
282	186
390	352
290	202
280	218
278	248
288	278
288	248
244	270
236	242

Step 1: State the Practical Problem

The Practical Problem:

We will compare the medians of two processes to see if they are the same.

Step 2: Establish the Hypotheses

H_o: Process medians are the same

H_a: Process medians are different

Step 3: Decide on Appropriate Statistical Test

Use the Mood's Median Test.

Notes:

Step 4: Set the Alpha Level

You want to be 95% confident that you are making the correct conclusion.
$\alpha = 0.05$

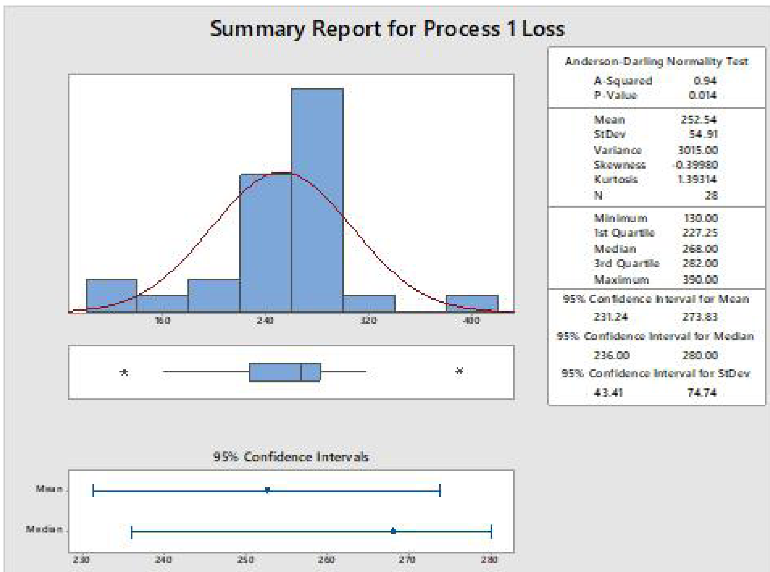
Steps 5&6: Set the Power and Sample Size, and Collect the Data

You have data from Process 1 and you need to collect data from Process 2.
You do not need the same number of data points.

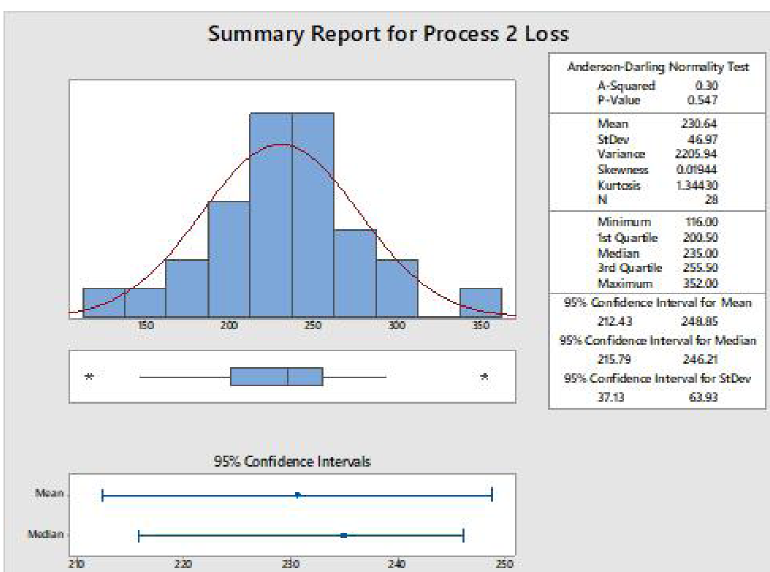
Step 7: Use the Appropriate Graphical Tool To Explore the Data

The following graphical summary was created from the sample data.

Graphical Summary Process 1

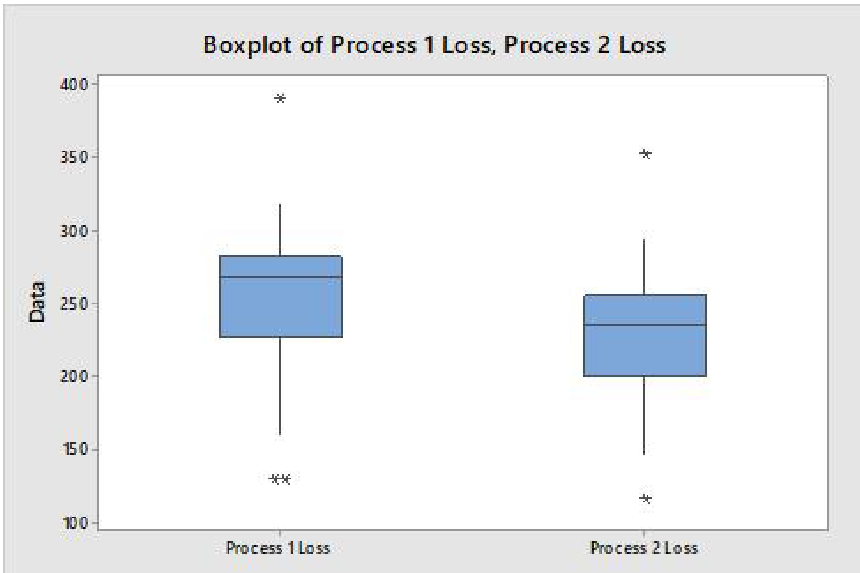


Graphical Summary Process 2



Notes:

Box Plot of Process 1 Loss, vs. Process 2 Loss



Step 8: Check Data Assumptions

The assumption is that the data does not have to be normal, but the two processes have similar distribution shapes.

The distribution for Process 1 is normal, but the distribution for Process 2 is not. Use the Mood's Median Test.

Step 9: Run the Statistical Test

To use Mood's Median test, you need to first stack the data.

- **Minitab: Stat > Nonparametrics > Mood's Median**
  - Response: Stacked Data
  - Factor: Subscripts

Notes:

Results

Mood's Median Test: Response versus Subscripts

Descriptive Statistics

Subscripts	Median	N <= Overall Median	N > Overall Median	Q3 – Q1	95% Median CI
Process 1 Loss	268	11	17	54.75	(236, 280)
Process 2 Loss	235	18	10	55.00	(215.793, 246.207)
Overall	242				

95.0% CI for median(Process 1 Loss) - median(Process 2 Loss): (-0.953865,60.2385)

Test

Null hypothesis	H_0: The population medians are all equal
Alternative hypothesis	H_1: The population medians are not all equal

DF	Chi-Square	P-Value
1	3.50	0.061

Statistical Conclusion

The P-Value (0.061) is higher than alpha (0.05), therefore, fail to reject the null hypothesis. The population medians are equal.

Step 10: Translate the Statistical Conclusion Into a Practical Conclusion

Based on the Mood's Median Test, The Process 1 Median is equal to the Process 2 Median so that the observed difference in medians is not statistically significant.

Two or More Samples: Kruskal-Wallis

What is a Kruskal-Wallis Test?

The Kruskal-Wallis test is a statistical test used to compare medians from two or more random samples to see if they come from the same or different populations. It is an extension of the Wilcoxon Mann-Whitney test.

It requires:

- The measurement scale is at least ordinal
- The distributions are of similar shape
- Distribution is any continuous one

But . . .

- Does not assume normality
- Distribution does not need to be symmetric
- The samples do not need to be the same size

The Kruskal-Wallis is the most robust test to use when you are testing medians from more than 2 samples.

An assumption for this test is that the samples from the different populations are independent random samples from continuous distributions, with the distributions

Notes:

having the same shape. The Kruskal-Wallis test is more powerful than Mood's median test for data from many distributions, including data from the normal distribution, but is less robust against outliers.

The Kruskal-Wallis test extends the Wilcoxon Mann-Whitney test from 2 samples only, to compare medians from two or more samples.

This is also a rank test. It combines all samples (n, m,…) and assigns ranks to the data (sample size n + m +…). Samples can be different sizes.

The samples are then separated out and the ranks are summed for each sample. The summed ranks are compared to a Wilcoxon table [with values for different sample sizes (n + m+…)].

Expect the sum of ranks to be proportionally similar for all samples if the medians are equal.

Steps in Hypothesis Testing

1. State the practical problem.
2. Establish the hypotheses.
3. Decide on appropriate statistical test.
4. Set the alpha level.
5. Determine initial sample size. Calculate the power.
6. Develop the sampling plan and collect the data.
7. Use the appropriate graphical tool to explore the data.
8. Check data assumptions (change choice of tool if appropriate).
    a. Not Normal
9. Run the statistical test and determine the statistical conclusion.
10. Translate statistical conclusion to a practical conclusion.

Long Cycle Times

A project team is testing theories on causes for long cycle times to resolve claims in an insurance company.

Three different departments process claims in the company.
Some of the theories relate to which of the three departments does the processing. The team wishes to test the theory that there is a difference among the three departments in claim processing time.

Step 1: State the Practical Problem

The Practical Problem:

Notes:

Do the three departments resolve claims at the same rate?

Step 2: Establish the Hypotheses

H_o: There is no difference among Department resolution time

H_a: There is a difference in Department resolution time

Step 3: Decide on Appropriate Statistical Test

Since the data from three populations is believed to be non-normal, a Kruskal-Wallis test will be used.

Step 4: Set the Alpha Level

You want to be 95% confident that you are making the correct conclusion.
$\alpha = 0.05$

Steps 5&6: Set the Power and Sample Size, and Collect the Data

For this test, ten samples were collected from each of the three departments.

The Data

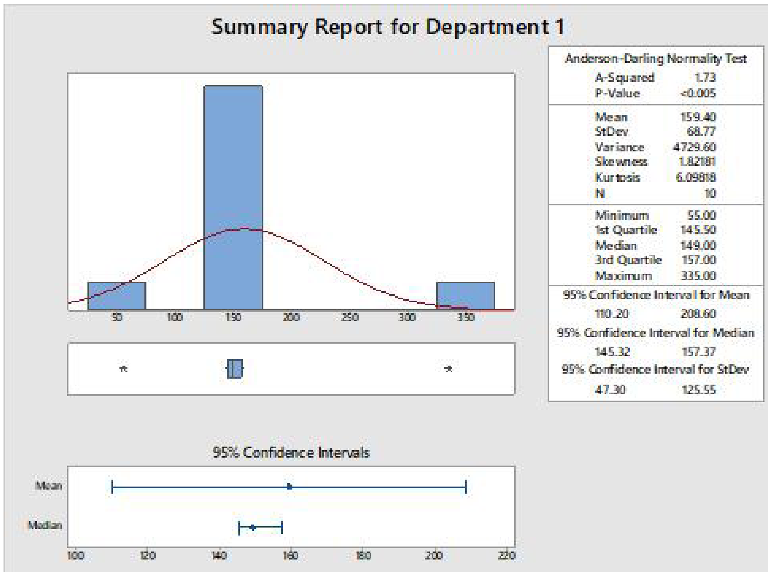
Department 1	Department 2	Department 3
150	130	125
335	315	310
148	128	123
146	126	121
160	140	135
156	136	131
152	132	127
148	128	123
144	124	119
55	35	30

Step 7: Use the Appropriate Graphical Tool To Explore the Data

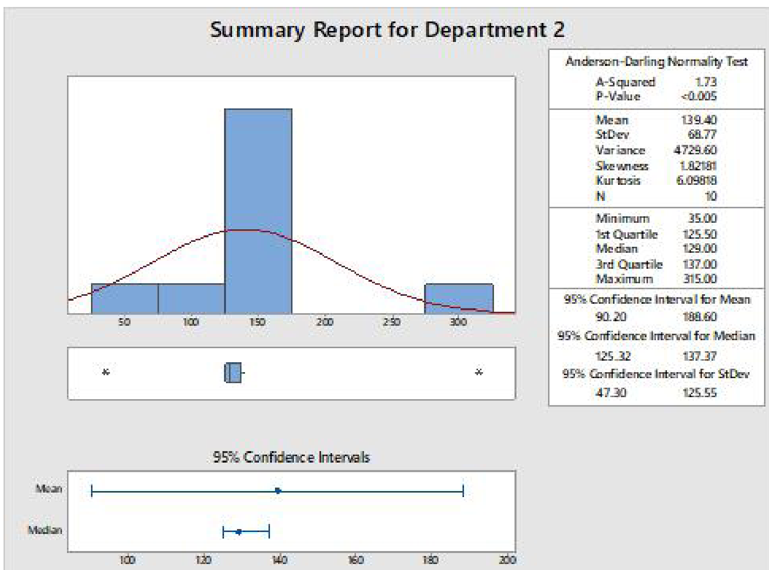
The following graphical summary was created from the sample data.

Notes:

Graphical Summary: Department 1

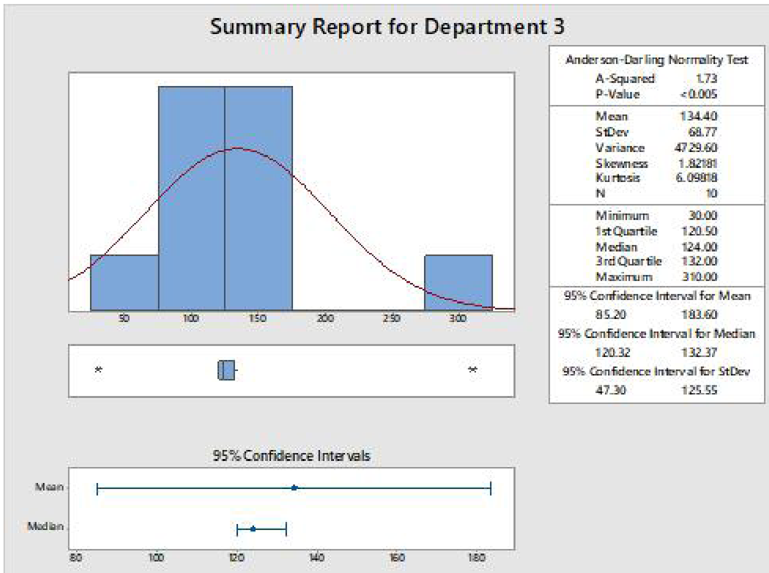


Graphical Summary: Department 2

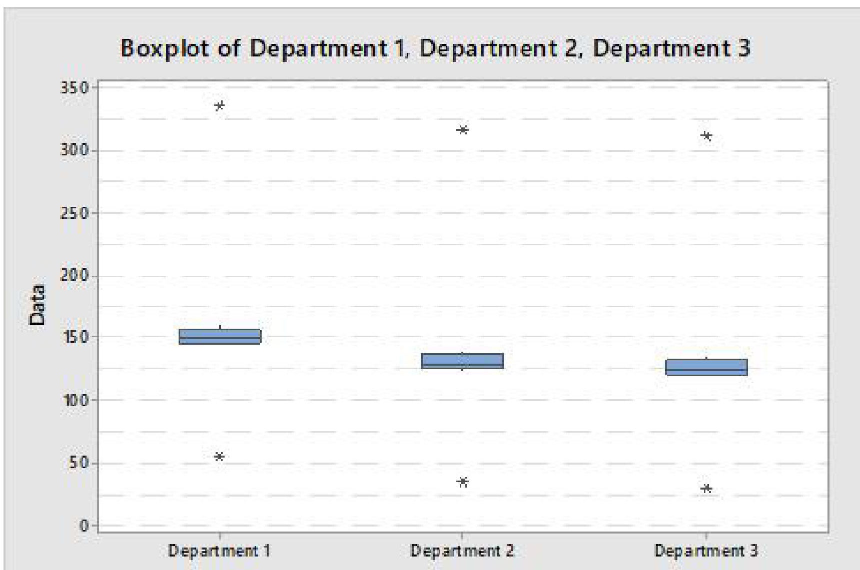


Notes:

Graphical Summary: Department 3



Boxplot of the Three Departments



Step 8: Check Data Assumptions

The data for all three departments are indeed not normal.

Step 9: Run the Statistical Test

To use Kruskal-Wallis test, you need to first stack the data.

- **Minitab: Stat > Nonparametrics > Kruskal-Wallis**
  - Response: Response Time
  - Factor: Department

Notes:

Results

Kruskal-Wallis Test: Response Time versus Department

Descriptive Statistics

Department	N	Median	Mean Rank	Z-Value
Department 1	10	149	22.1	2.90
Department 2	10	129	14.1	-0.62
Department 3	10	124	10.3	-2.29
Overall	30		15.5	

Test

Null hypothesis H_0: All medians are equal
Alternative hypothesis H_1: At least one median is different

Method	DF	H-Value	P-Value
Not adjusted for ties	2	9.36	0.009
Adjusted for ties	2	9.37	0.009

Statistical Conclusion

The P-Value (0.009) is lower than alpha (0.05), therefore, reject the null hypothesis. At least one of the medians is different.

Step 10: Translate the Statistical Conclusion Into a Practical Conclusion

There is a difference in Department claim resolution time.

When Should Nonparametric Tests Be Used?

Use nonparametric tests any time data fail to pass the normality test.

Pitfalls to Avoid

Nonparametric tests are less likely to yield inappropriate analysis to non-normal data. It is still a good idea to verify your conclusions with follow-up studies.

Notes: