4
Analyze Phase

Note: There is no change to Section 4 in the 2022 ASQ CSSGB BoK update.

A Exploratory Data Analysis

01 Multi-vari Studies

02 Correlation and linear regression

B Hypothesis Testing

01 Basics

02 Tests for means, variances, and proportions

4A-1 Multi-vari Studies

01 Three types of variation

02 Example–Bearing Measurements

03 Example–Call Centre (Minitab)

Multi-vari Chart

- A tool to visually show a variety of sources of variation.

3 Types of Variations

Variation

- Positional (within parts)
- Cyclical (between parts)
- Temporal (over a period of time)

3 Types of Variations

Variation

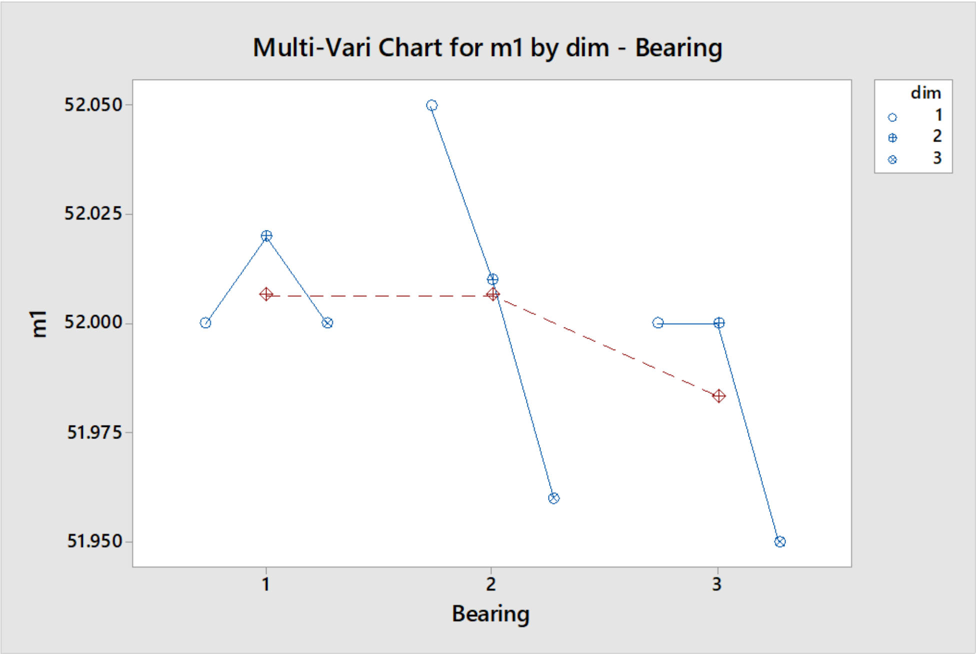
- ❖ Positional (within parts)
- ❖ Cyclical (between parts)

Bearing	dim	m1
1	1	52.00
1	2	52.02
1	3	52.00
2	1	52.05
2	2	52.01
2	3	51.96
3	1	52.00
3	2	52.00
3	3	51.95

3 Types of Variations

Variation

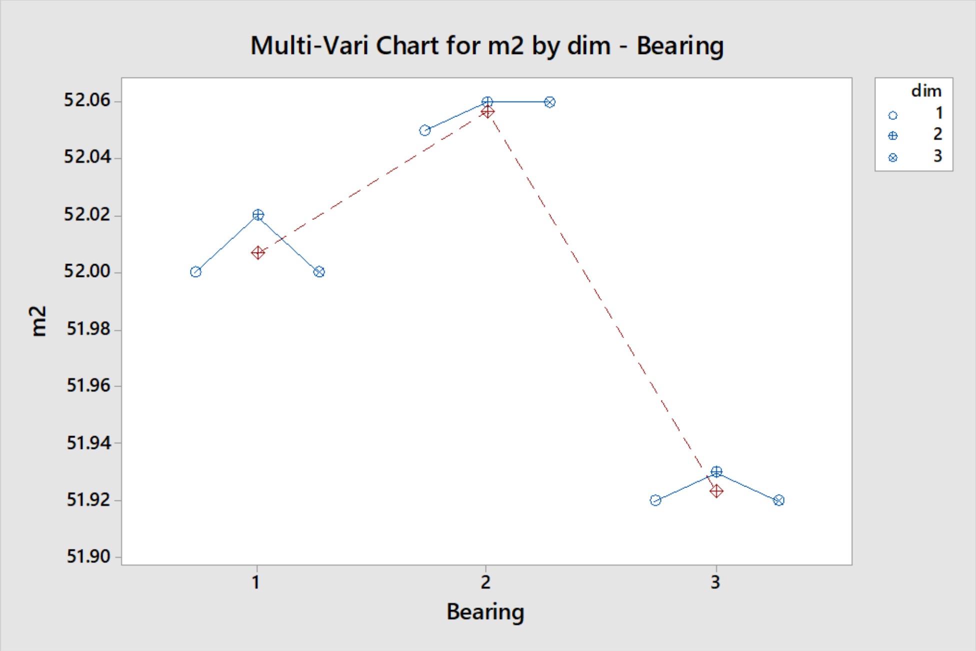
Bearing	dim	m1
1	1	52.00
1	2	52.02
1	3	52.00
2	1	52.05
2	2	52.01
2	3	51.96
3	1	52.00
3	2	52.00
3	3	51.95



3 Types of Variations

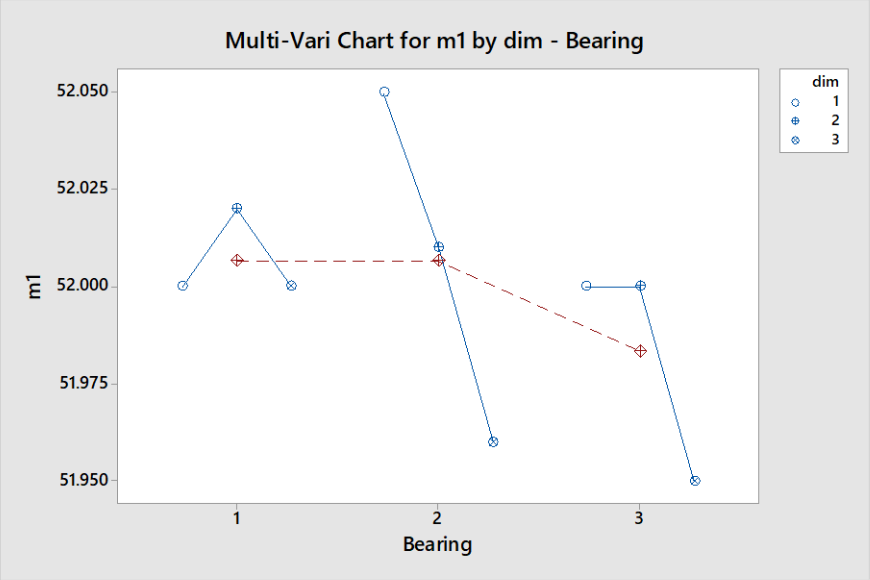
Variation

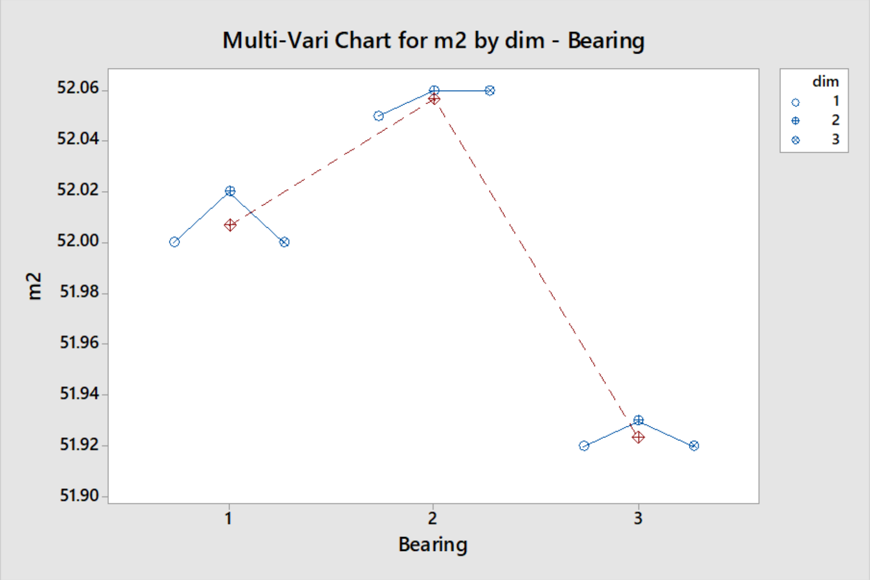
Bearing	dim	m2
1	1	52.00
1	2	52.02
1	3	52.00
2	1	52.05
2	2	52.06
2	3	52.06
3	1	51.92
3	2	51.93
3	3	51.92



3 Types of Variations

Variation





3 Types of Variations

Demo Using Minitab 18

Call Centres.MTW ***

↓	C1-T	C2-T	C3-D	C4	C5-T	C6	C7	C8-T
	Call Centre	Request	Dates	Durations	Customer Cat	Week	Hour	Day
85	Montpellier	Technical Support	03/09/2007	906	D	36	13	Mon
86	Montpellier	Individual accounts	03/09/2007	599	C	36	13	Mon
87	Saint-Quentin	New account	03/09/2007	613	E	36	13	Mon
88	Montpellier	Individual accounts	03/09/2007	765	A	36	14	Mon
89	Montpellier	Credit Cards	03/09/2007	779	A	36	14	Mon
90	Saint-Quentin	Queries	03/09/2007	607	C	36	14	Mon
91	Montpellier	Technical Support	03/09/2007	1077	A	36	14	Mon
92	Montpellier	Credit Cards	03/09/2007	487	A	36	14	Mon
93	Montpellier	Technical Support	03/09/2007	752	A	36	14	Mon
94	Saint-Quentin	Queries	03/09/2007	580	B	36	14	Mon

Multi-Vari Charts

4A-2 Correlation and Linear Regression

01 Calculate correlation coefficient

02 Correlation vs causation

03 Linear regression equation

04 Regression model for estimation and prediction

Correlation

- Y = f(X),
  - where Y is Dependent variable or the result (output)
  - X is Independent variable, input or the controllable variable

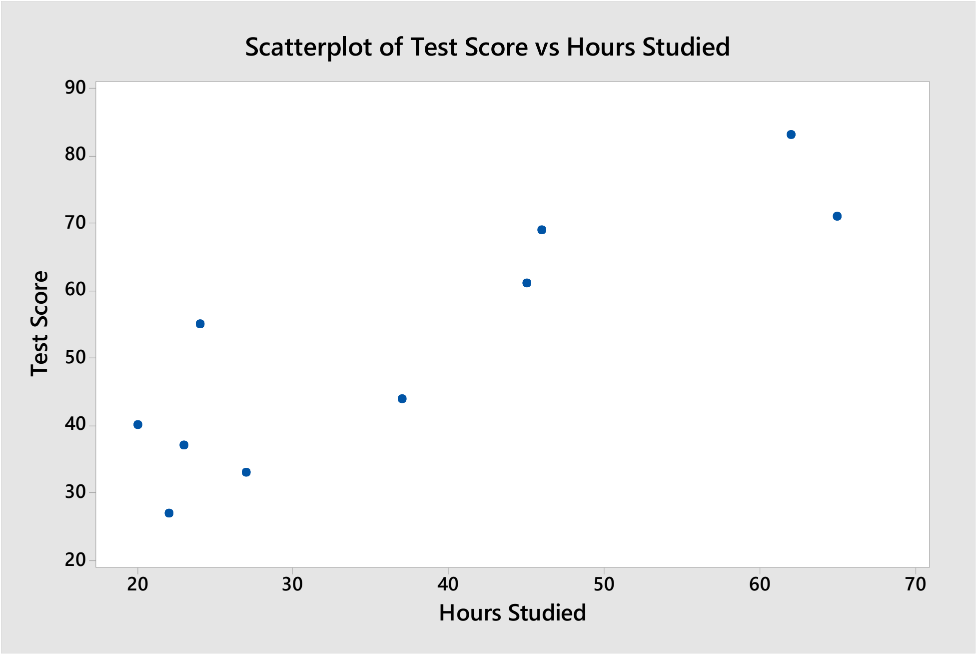
- For example in the study of marks obtained by students in a subject (Y) vs hours of study (X)

Correlation

Correlation

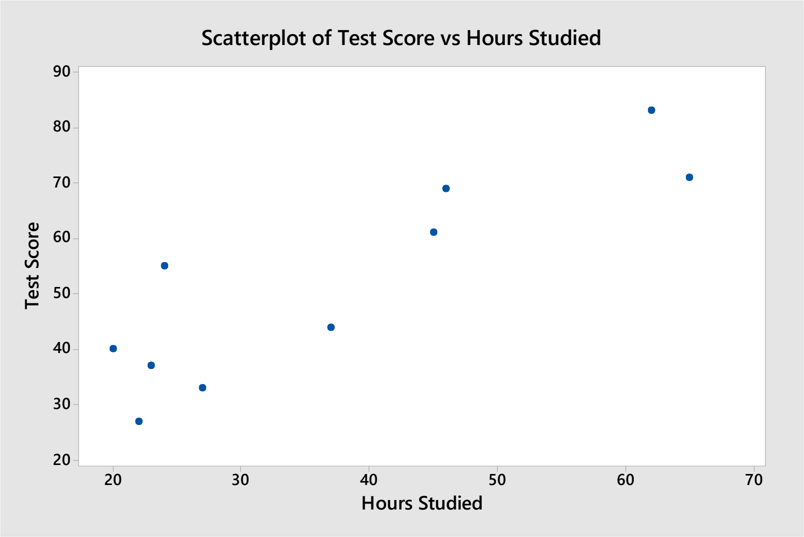
Hours Studied (X)	Test Score % (Y)
20	40
24	55
46	69
62	83
22	27
37	44
45	61
27	33
65	71
23	37

Correlation



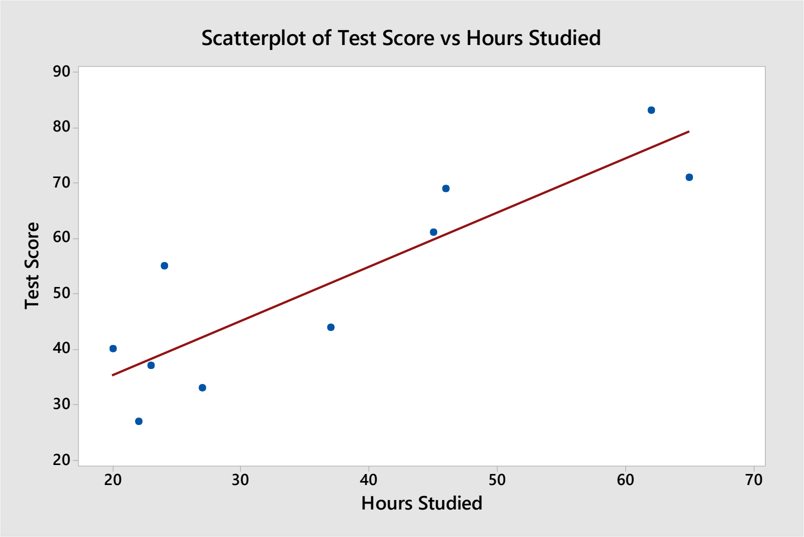
Hours Studied (X)	Test Score % (Y)
20	40
24	55
46	69
62	83
22	27
37	44
45	61
27	33
65	71
23	37

Scatter Plot



Correlations

Pearson correlation	0.879
P-value	0.001



Regression Equation

Test Score = 15.79 + 0.976 Hours Studied

Correlation / Regression

	Hours Studied (X)	Test Score % (Y)	XY	X2	Y2
	20	40	800	400	1600
	24	55	1320	576	3025
	46	69	3174	2116	4761
	62	83	5146	3844	6889
	22	27	594	484	729
	37	44	1628	1369	1936
	45	61	2745	2025	3721
	27	33	891	729	1089
	65	71	4615	4225	5041
	23	37	851	529	1369
SUM	**371**	**520**	**21764**	**16297**	**30160**

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2]\,[n(\sum y^2) - (\sum y)^2]}}$$

Pearson Correlation Coefficient

	Hours Studied (X)	Test Score % (Y)	XY	X2	Y2
	20	40	800	400	1600
	24	55	1320	576	3025
	46	69	3174	2116	4761
	62	83	5146	3844	6889
	22	27	594	484	729
	37	44	1628	1369	1936
	45	61	2745	2025	3721
	27	33	891	729	1089
	65	71	4615	4225	5041
	23	37	851	529	1369
SUM	**371**	**520**	**21764**	**16297**	**30160**

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2]\,[n(\sum y^2) - (\sum y)^2]}}$$

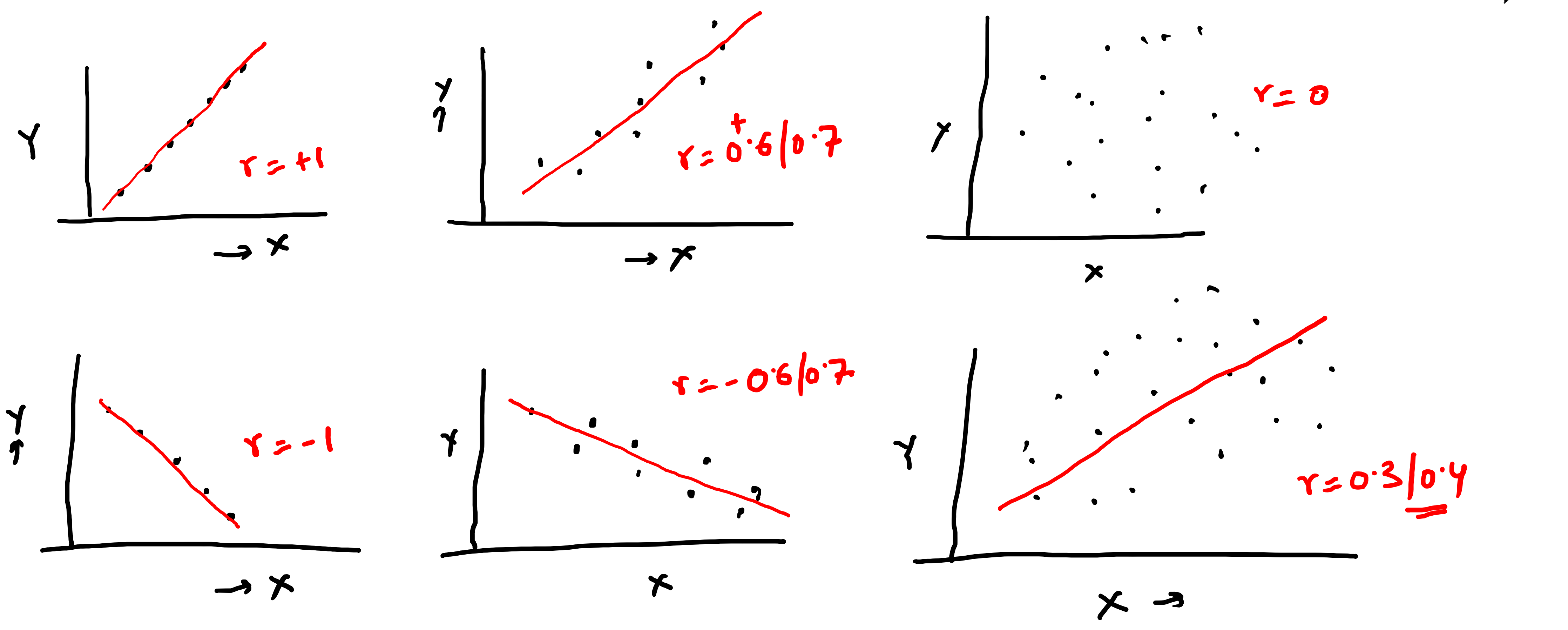
$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Pearson Correlation Coefficient

Pearson Correlation Coefficient

- Measures the strength of linear relationship between Y and X
- Pearson Correlation Coefficient, r (r varies between -1 and +1)
  - Perfect positive relationship: r = 1
  - No relationship: r = 0
  - Perfect negative relationship: r = -1

Correlation



Pearson Correlation Coefficient

Correlation

- Population correlation (ρ) – usually unknown
- Sample correlation (r)

Correlation

Correlation

- ❖ Coefficient of Determination, r^2
- ❖ Proportion of the variance in the dependent variable that is predictable from the independent variable
- ❖ (varies from 0.0 to 1.0 or zero to 100%)
  - ❖ None of the variation in Y is explained by X, $r^2 = 0.0$
  - ❖ All of the variation in Y is explained by X, $r^2 = 1.0$
  - ❖ $r = 0.88$, $r^2 = 0.77$

Correlation

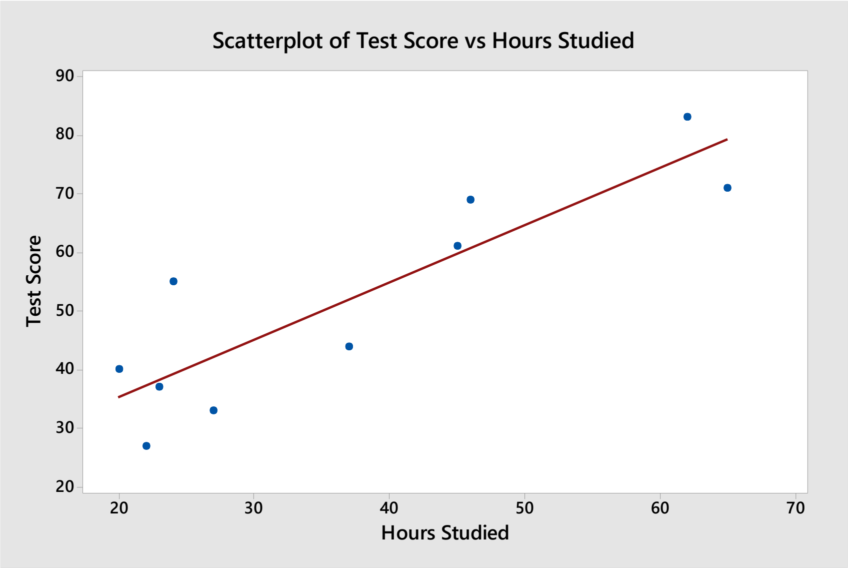
Correlation vs Causation

- **Correlation does not imply causation**
  - a correlation between two variables does not imply that one causes the other

Correlation

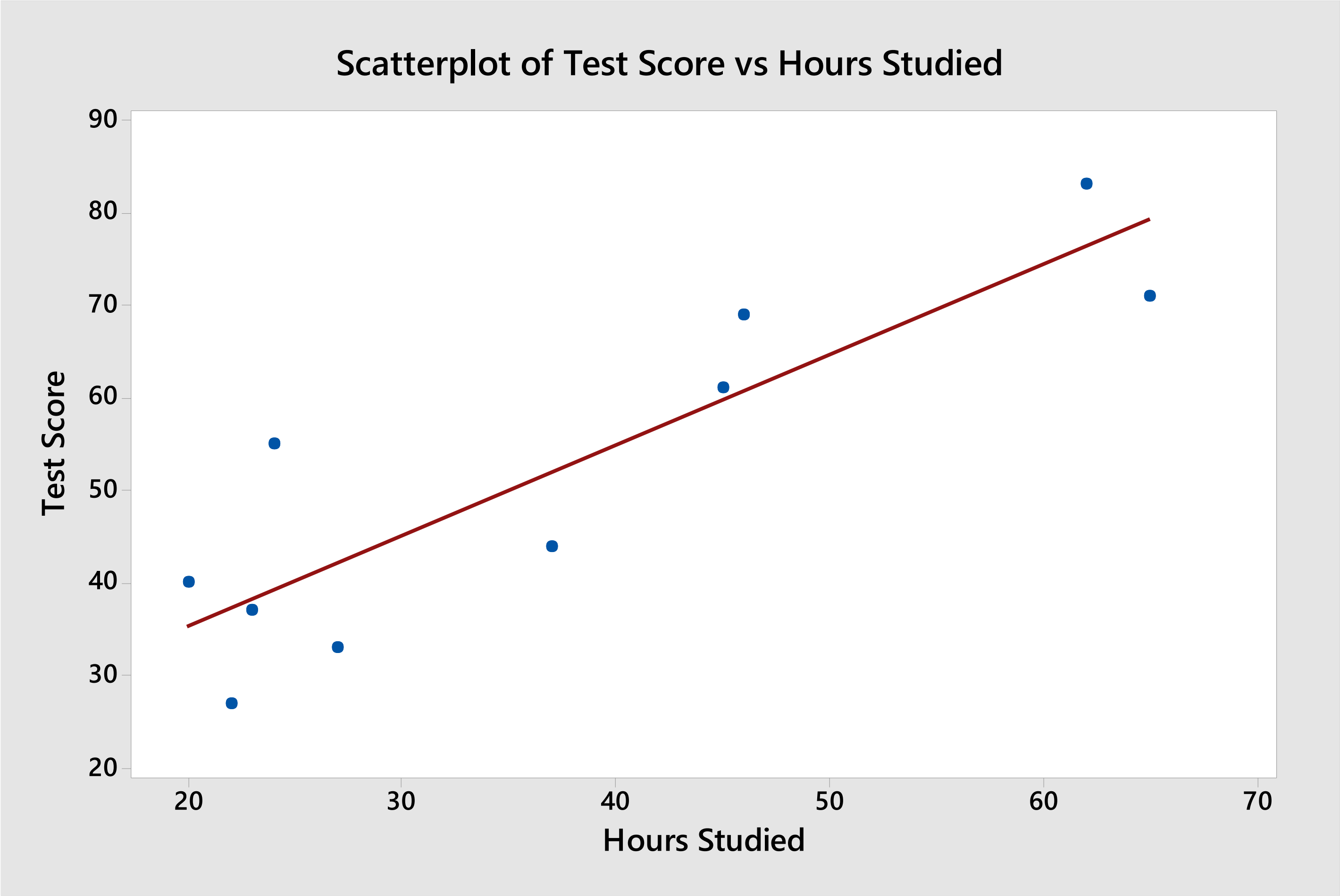
Regression Analysis

- Quantifies the relationship between Y and X ($Y = a + bX$)



Regression

Ordinary Least Square (OLS) Method



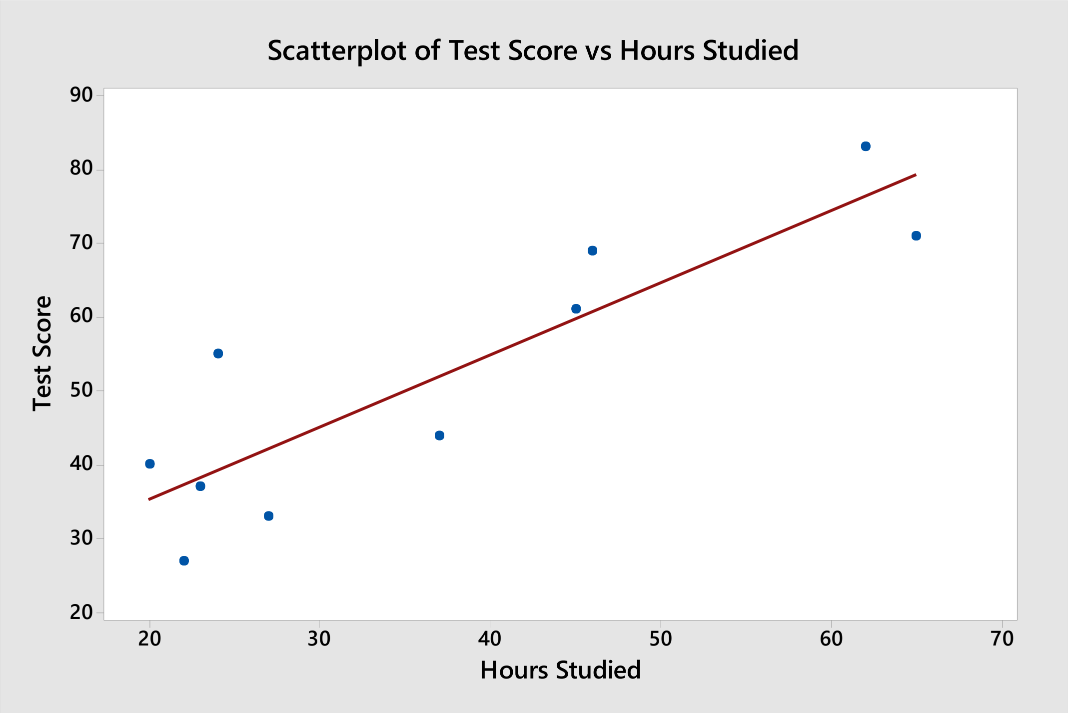
Regression

Ordinary Least Square (OLS) Method

$$Y = a + bX$$

$$b = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{N\Sigma X^2 - (\Sigma X)^2}$$

$$a = \frac{\sum Y - b\sum X}{N}$$



Regression

Ordinary Least Square (OLS) Method

$$Y = a + bX$$

$$b = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{N\Sigma X^2 - (\Sigma X)^2}$$

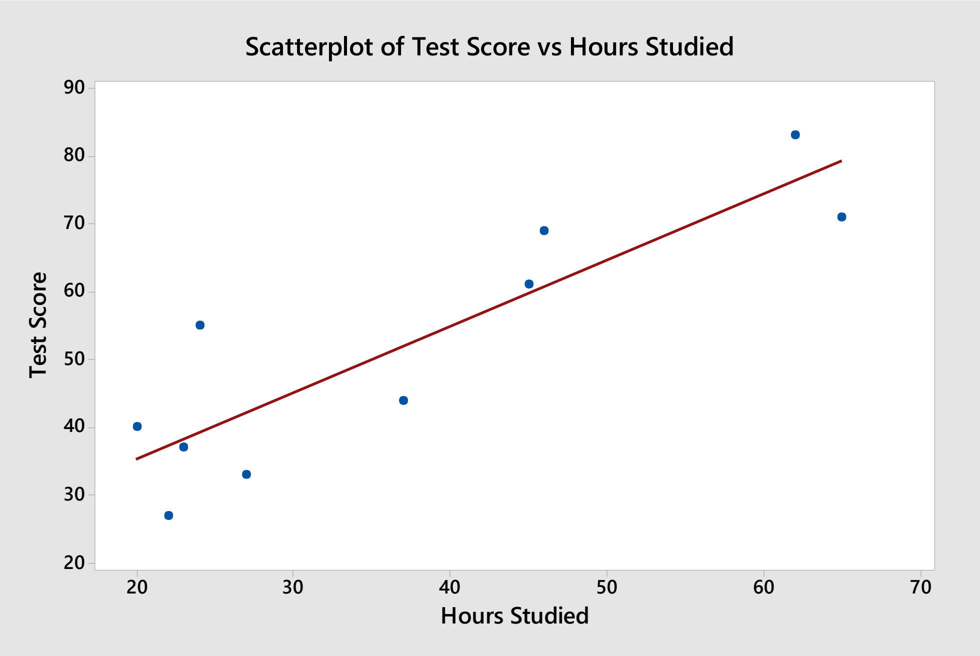
$$a = \frac{\sum Y - b\sum X}{N}$$

	Hours Studied (X)	Test Score % (Y)	XY	X2	Y2
	20	40	800	400	1600
	24	55	1320	576	3025
	46	69	3174	2116	4761
	62	83	5146	3844	6889
	22	27	594	484	729
	37	44	1628	1369	1936
	45	61	2745	2025	3721
	27	33	891	729	1089
	65	71	4615	4225	5041
	23	37	851	529	1369
SUM	371	520	21764	16297	30160

$$\boldsymbol{Y = 15.79 + 0.97\, X}$$

Regression

Ordinary Least Square (OLS) Method

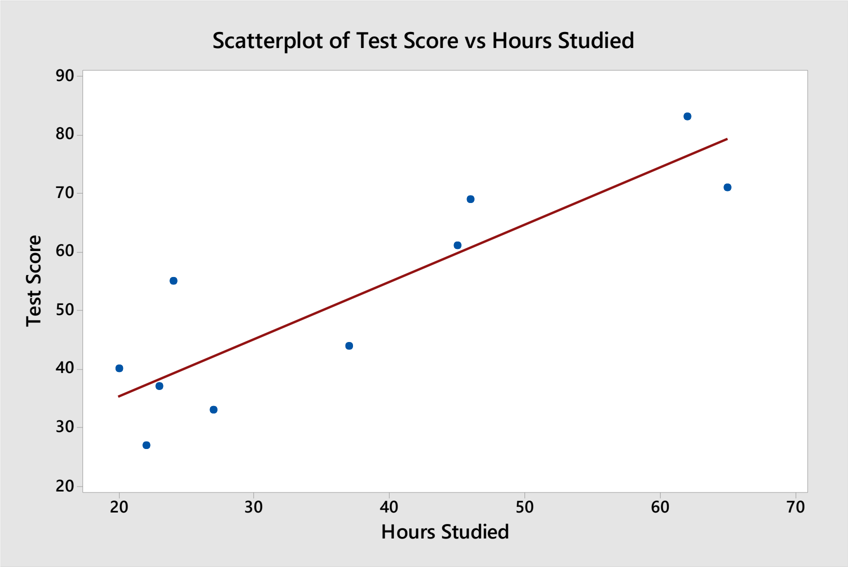


Hours Studied (X)	Test Score % (Y)	Y = 15.79 + 0.97.X	Residual
20	40	35.19	4.81
24	55	39.07	15.93
46	69	60.41	8.59
62	83	75.93	7.07
22	27	37.13	-10.13
37	44	51.68	-7.68
45	61	59.44	1.56
27	33	41.98	-8.98
65	71	78.84	-7.84
23	37	38.1	-1.1

$$\boldsymbol{Y = 15.79 + 0.97\,X}$$

Regression

- For a student studying 50 hrs what is the expected test score %?



$$Y = 15.79 + 0.97\,X$$

$$Y = 64.58$$

Regression

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	2412.6	2412.56	27.28	0.001
Hours Studied	1	2412.6	2412.56	27.28	0.001
Error	8	707.4	88.43		
Total	9	3120.0			

If P-value is less than 0.05, you can say with 95% confidence the significant relationship exists between variables.

Regression

4B Hypothesis Testing

4B-1 Hypothesis Testing Basics

01 Statistical vs practical significance

02 Hypothesis testing steps

03 Type 1 and 2 errors

04 The p Value

Hypothesis Testing

- A statistical hypothesis test is a method of statistical inference.
- Commonly used tests include:
  - Compare sample statistic with the population parameter
  - Compare two datasets

Hypothesis Testing

Statistical Significance

- Case of a perfume making company:
- Mean Volume 150 cc and sd=2 cc

Hypothesis Testing

Practical Significance

- Practical significance of an experiment tells us if there is any actionable information from the result.
- Large samples can find out statistical difference for very small difference. These small difference might not have practical significance.

Hypothesis Testing

Hypothesis Testing

1. State the Alternate Hypothesis.
2. State the Null Hypothesis.
3. Select a probability of error level (alpha level). Generally 0.05
4. Calculate the test statistic (e.g t or z score)
5. Critical test statistic
6. Interpret the results.

Hypothesis Testing

Hypothesis Testing

- Null Hypothesis: The person is innocent
- Alternate Hypothesis: The person is guilty. You need to provide proof of this.
- Court conclusion is: Guilty or Not Guilty (not the innocent)

Hypothesis Testing

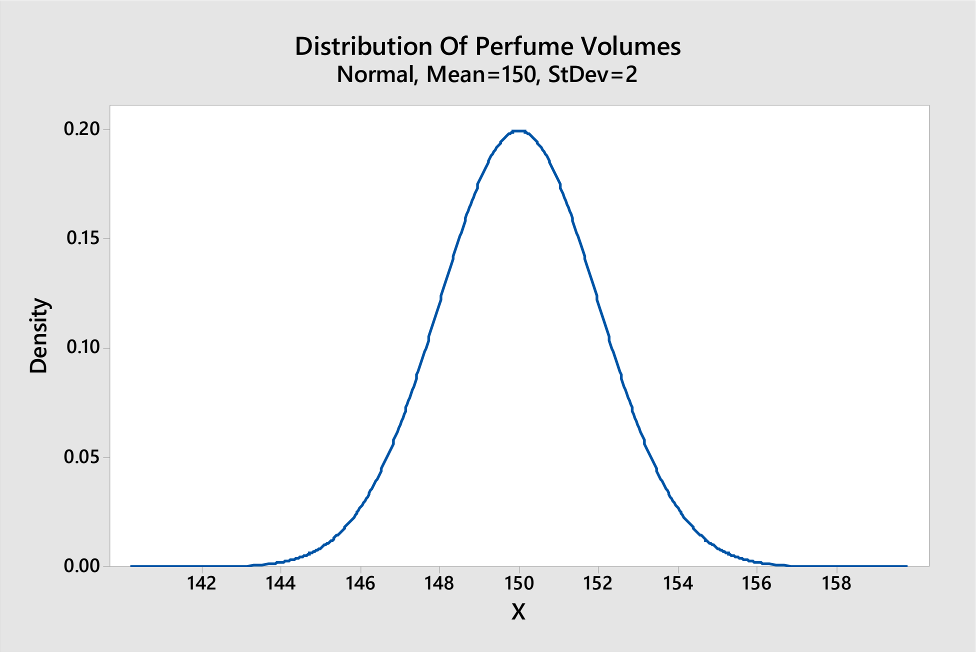
Hypothesis Testing

- Null Hypothesis: The person is innocent
- Alternate Hypothesis: The person is guilty. You need to provide proof of this.
- In statistical terms you:
  - Reject the Null Hypothesis, or
  - Fail to reject the Null Hypothesis (not accept the Null Hypothesis)

Hypothesis Testing

Hypothesis Testing

- Null Hypothesis: The machine is filling the bottles with 150 cc
- Alternate Hypothesis: The machine is "not" filling the bottles with 150 cc.



Hypothesis Testing

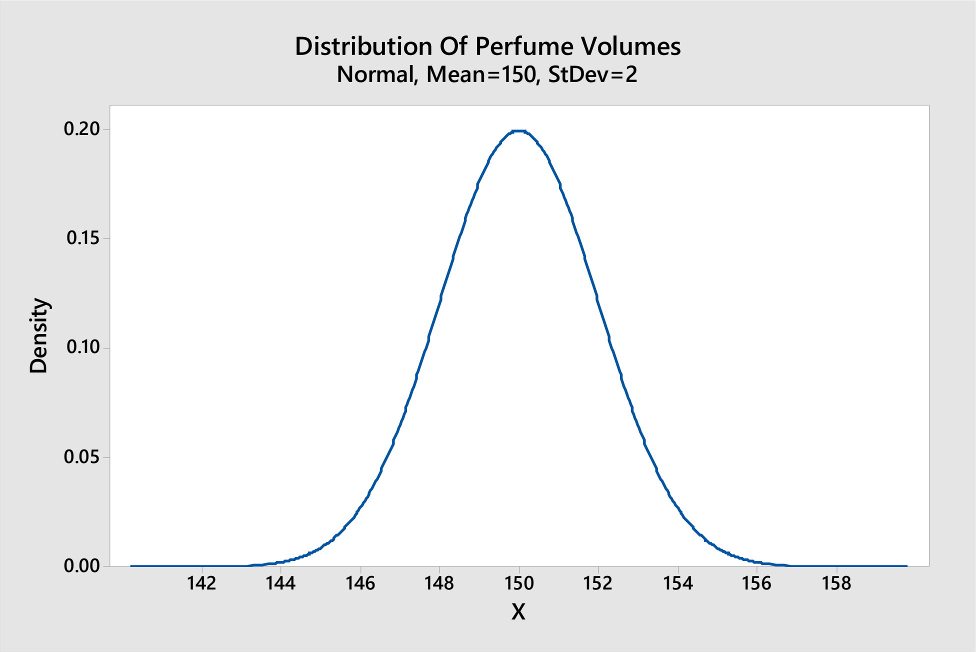
Hypothesis Testing

- Lower Tail Tests
  - H_0: $\mu \geq 150cc$
  - H_a: $\mu < 150cc$
- Upper Tail Tests
  - H_0: $\mu \leq 150cc$
  - H_a: $\mu > 150cc$
- Two Tail Tests
  - H_0: $\mu = 150cc$
  - H_a: $\mu \neq 150cc$

Hypothesis Testing

Hypothesis Testing

- What would you conclude if you pick one sample and find the volume as:
  - 147 cc ….. or
  - 156 cc



Hypothesis Testing

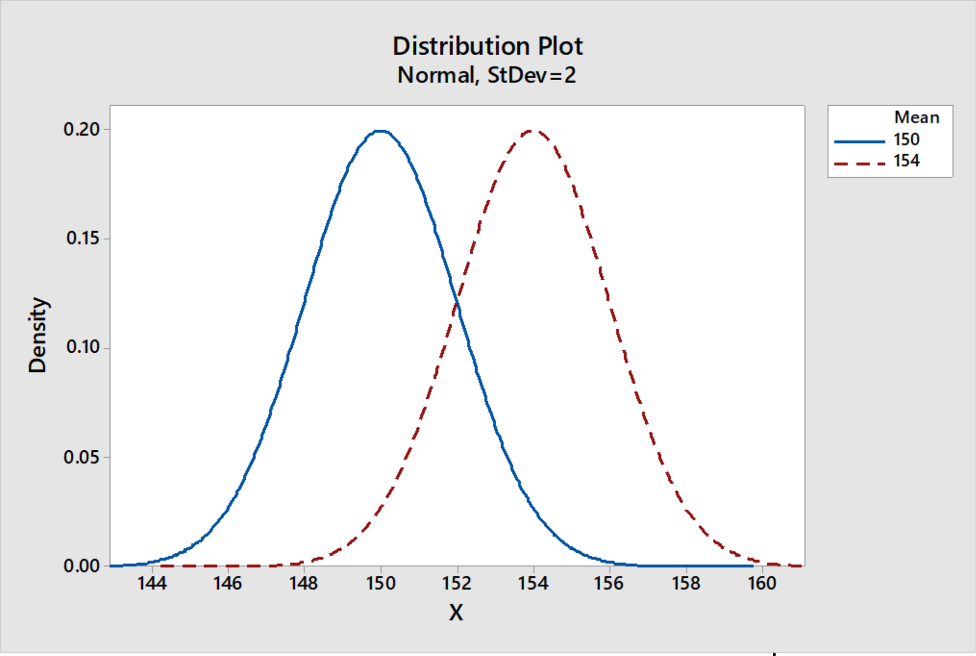
Hypothesis Testing

1. State the Alternate Hypothesis.
2. State the Null Hypothesis.
3. Select a probability of error level (alpha level). Generally 0.05
4. Calculate the test statistic (e.g t or z score)
5. Critical test statistic
6. Interpret the results.

Hypothesis Testing - Errors

		True State of Nature	
		H_0 Is true	H_a Is true
Conclusion	Support H_0 / Reject H_a	Correct Conclusion	Type II Error
	Support H_a / Reject H_0	Type I Error	Correct Conclusion (Power)

Hypothesis Testing – Errors



$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

		True State of Nature	
		H_0 Is true	H_a Is true
Conclusion	Support H_0 / Reject H_a	Correct Conclusion	Type II Error
	Support H_a / Reject H_0	Type I Error	Correct Conclusion (Power)

Hypothesis Testing – Errors

		True State of Nature	
		H_0 Is true	H_a Is true
Conclusion	Support H_0 / Reject H_a	Correct Conclusion	Type II Error
	Support H_a / Reject H_0	Type I Error	Correct Conclusion (Power)

	Type I error (alpha)	Type II error (beta)
Name	Producer's risk/ **Significance level**	Consumer's risk
1 minus error is called	**Confidence level**	Power of the test
Example of Fire Alarm	False fire alarm leading to inconvenience	Missed fire leading to disaster
Effects on process	Unnecessary cost increase due to frequent changes	Defects may be produced
Control method	Usually fixed at a pre-determined level, 1%, 5% or 10%	Usually controlled to < 10% by appropriate sample size
Simple definition	Innocent declared as guilty	Guilty declared as innocent

Hypothesis Testing – Errors

Hypothesis Testing – Errors

Confidence Level:

C = 0.90, 0.95, 0.99 (90%, 95%, 99%)

Level of Significance or Type I Error:

$\alpha = 1 - C$ (0.10, 0.05, 0.01)

	Type I error (alpha)	Type II error (beta)
Name	Producer's risk/ **Significance level**	Consumer's risk
1 minus error is called	**Confidence level**	Power of the test
Example of Fire Alarm	False fire alarm leading to inconvenience	Missed fire leading to disaster
Effects on process	Unnecessary cost increase due to frequent changes	Defects may be produced
Control method	Usually fixed at a pre-determined level, 1%, 5% or 10%	Usually controlled to < 10% by appropriate sample size
Simple definition	Innocent declared as guilty	Guilty declared as innocent

Power

- Power = 1 – β (or 1 - type II error)
- Type II Error: Failing to reject null hypothesis when null hypothesis is false.
- Power: Likelihood of rejecting null hypothesis when null hypothesis is false.
- Or: Power is the ability of a test to correctly reject the null hypothesis.

		True State of Nature	
		H_0 Is true	H_a Is true
Conclusion	Support H_0 / Reject H_a	Correct Conclusion	Type II Error
	Support H_a / Reject H_0	Type I Error	Correct Conclusion (Power)

Hypothesis Testing – Errors

Alpha vs Beta

- Researcher can not commit both Type I and II error. Only one can be committed.
- As the value of α increases (say 0.01 to 0.05) β goes down and the Power of test increases.
- To reduce both Type I and II errors increase sample size.

Hypothesis Testing - Errors

Hypothesis Testing

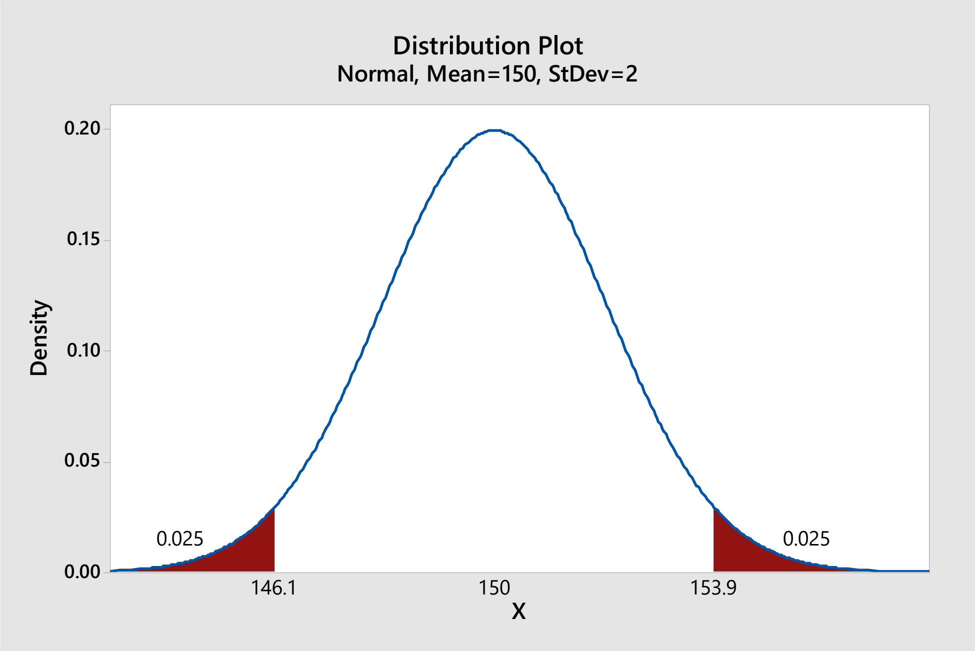
1. State the Alternate Hypothesis.
2. State the Null Hypothesis.
3. Select a probability of error level (alpha level). Generally 0.05
4. Calculate the test statistic (e.g t or z score)
5. Critical test statistic
6. Interpret the results.

Hypothesis Testing

Hypothesis Testing

- Let's pick 1 bottle from the production line and the volume is 153.8 cc
- With 95% confidence level we will fail to reject the null hypothesis.

H_0: $\mu = 150cc$
H_a: $\mu \neq 150cc$

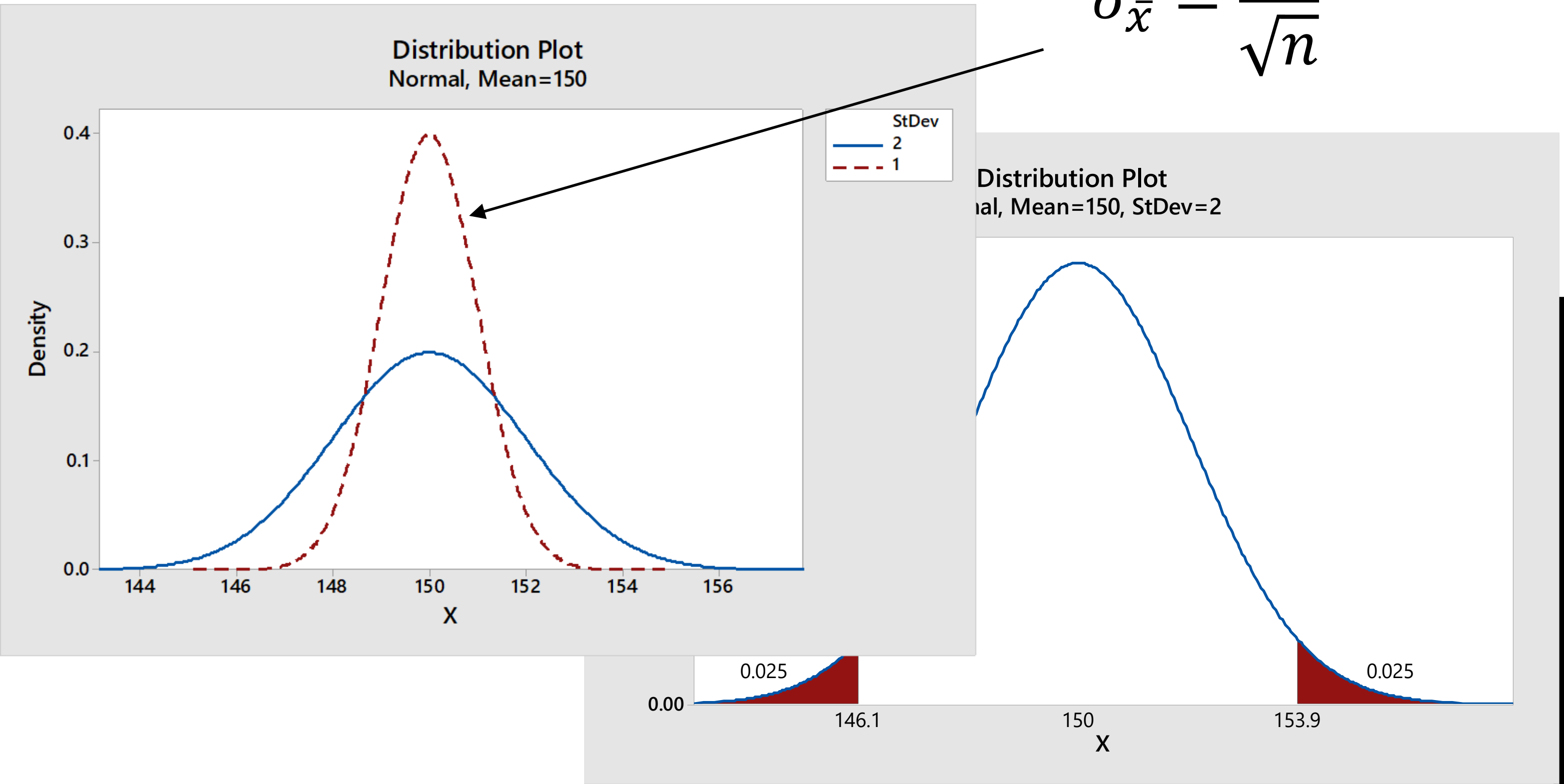


Hypothesis Testing

Hypothesis Testing

- Let’s pick 4 bottles from the production line and find the average volume as 153.8 cc.

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

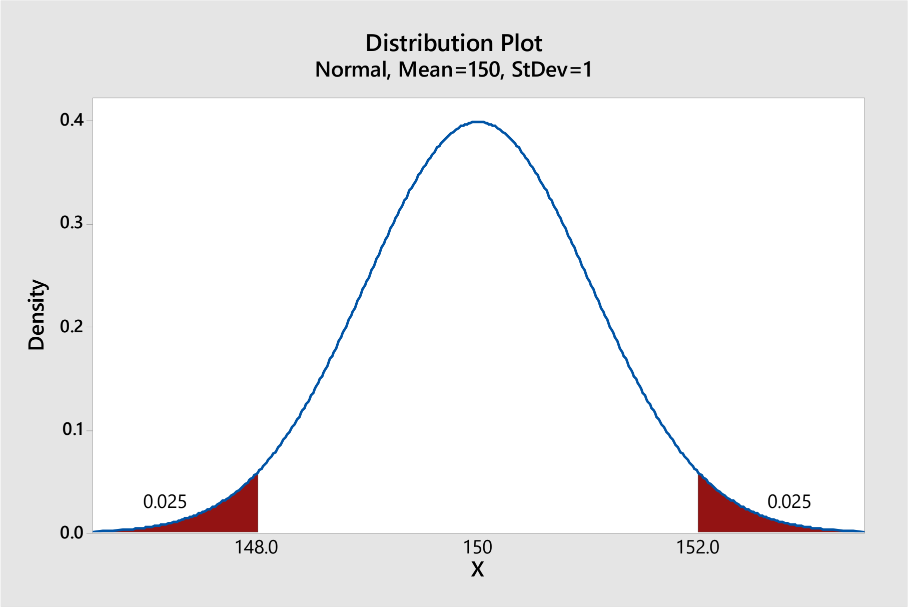


Hypothesis Testing

Hypothesis Testing

- Let's pick 4 bottles from the production line and find the average volume as 153.8 cc.
- With 95% confidence level we will reject the null hypothesis.

H_0: $\mu = 150cc$
H_a: $\mu \neq 150cc$



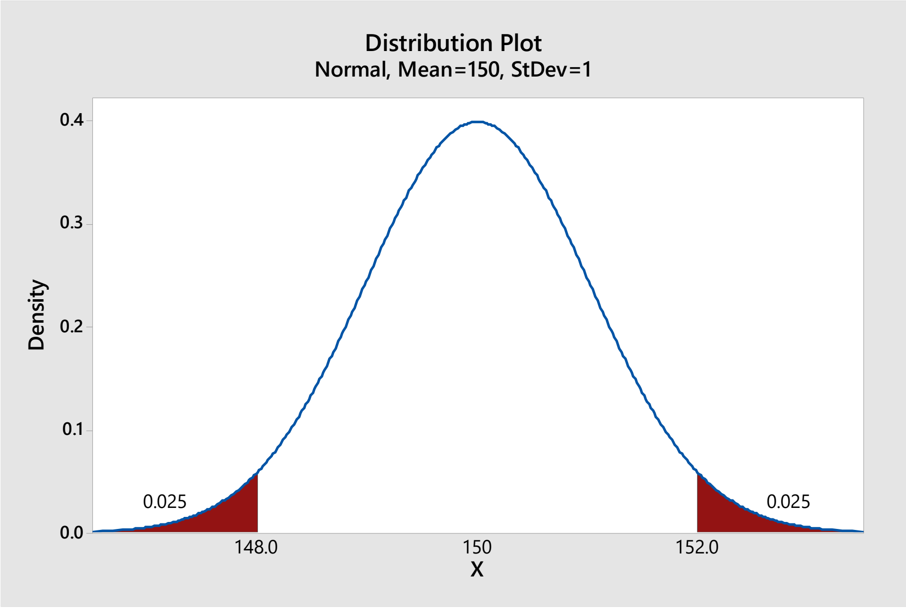
Hypothesis Testing

Hypothesis Testing

$$z_{cal} = \frac{(\bar{X} - \mu)}{\sigma_x/\sqrt{n}} \qquad z_{cal} = \frac{(153.8 - 150)}{2/\sqrt{4}} = 3 \cdot 8$$

For α = 0.05 Two Tails means 0.025 on both tails. Z Critical = 1.96

H_0: μ = 150cc
H_a: μ ≠ 150cc



Hypothesis Testing

Hypothesis Testing

1. State the Alternate Hypothesis.
2. State the Null Hypothesis.
3. Select a probability of error level (alpha level). Generally 0.05
4. Calculate the test statistic (e.g t or z score)
5. Critical test statistic
6. Interpret the results.

Hypothesis Testing

z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

Hypothesis Testing

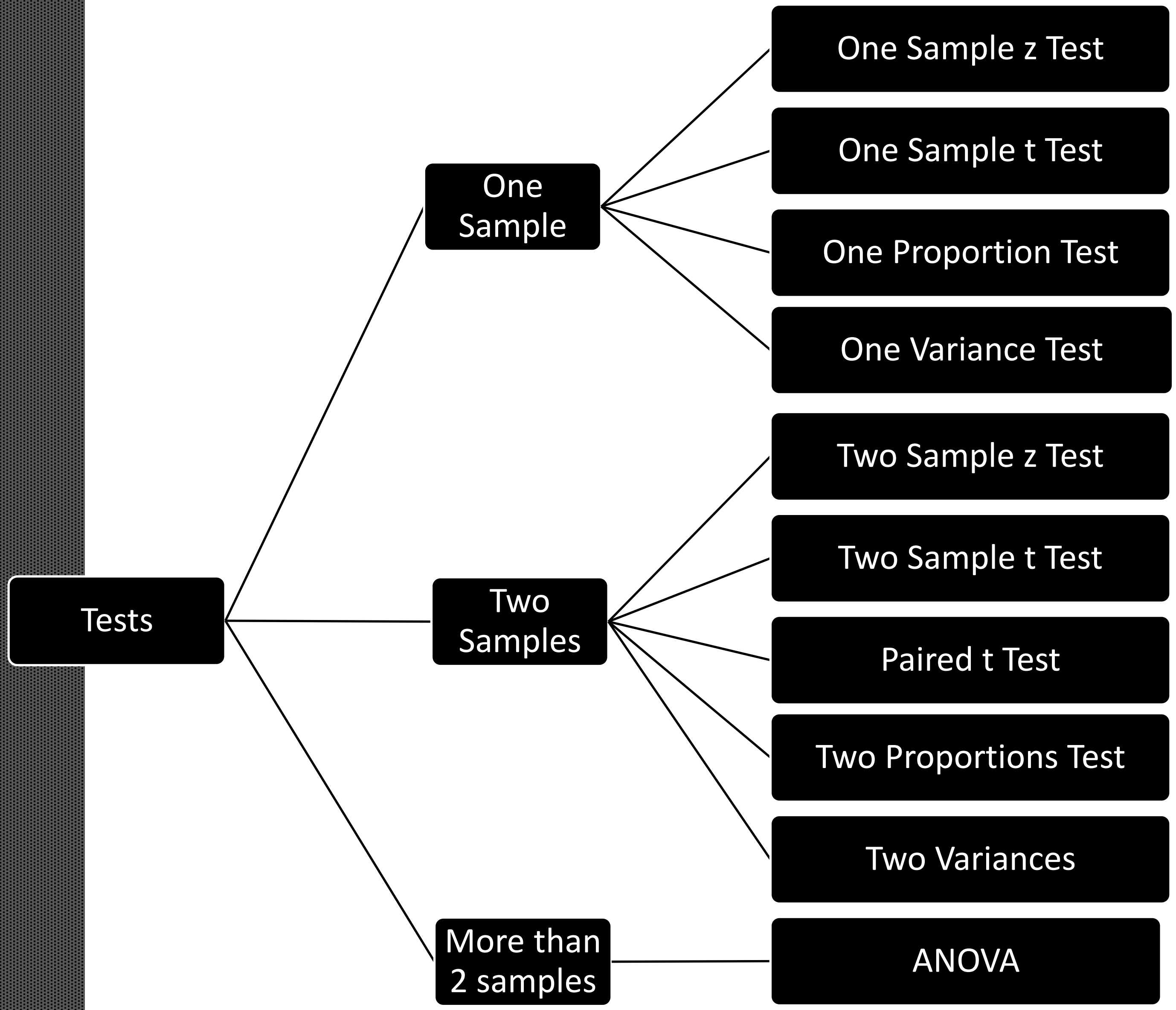
- α = 0.01 Two Tails means 0.005 on both tails. Z Critical = 2.575
- α = 0.05 Two Tails means 0.025 on both tails. Z Critical = 1.96
- α = 0.10 Two Tails means 0.05 on both tails. Z Critical = 1.645
- α = 0.05 Single Tails
  - Z Critical = 1.645

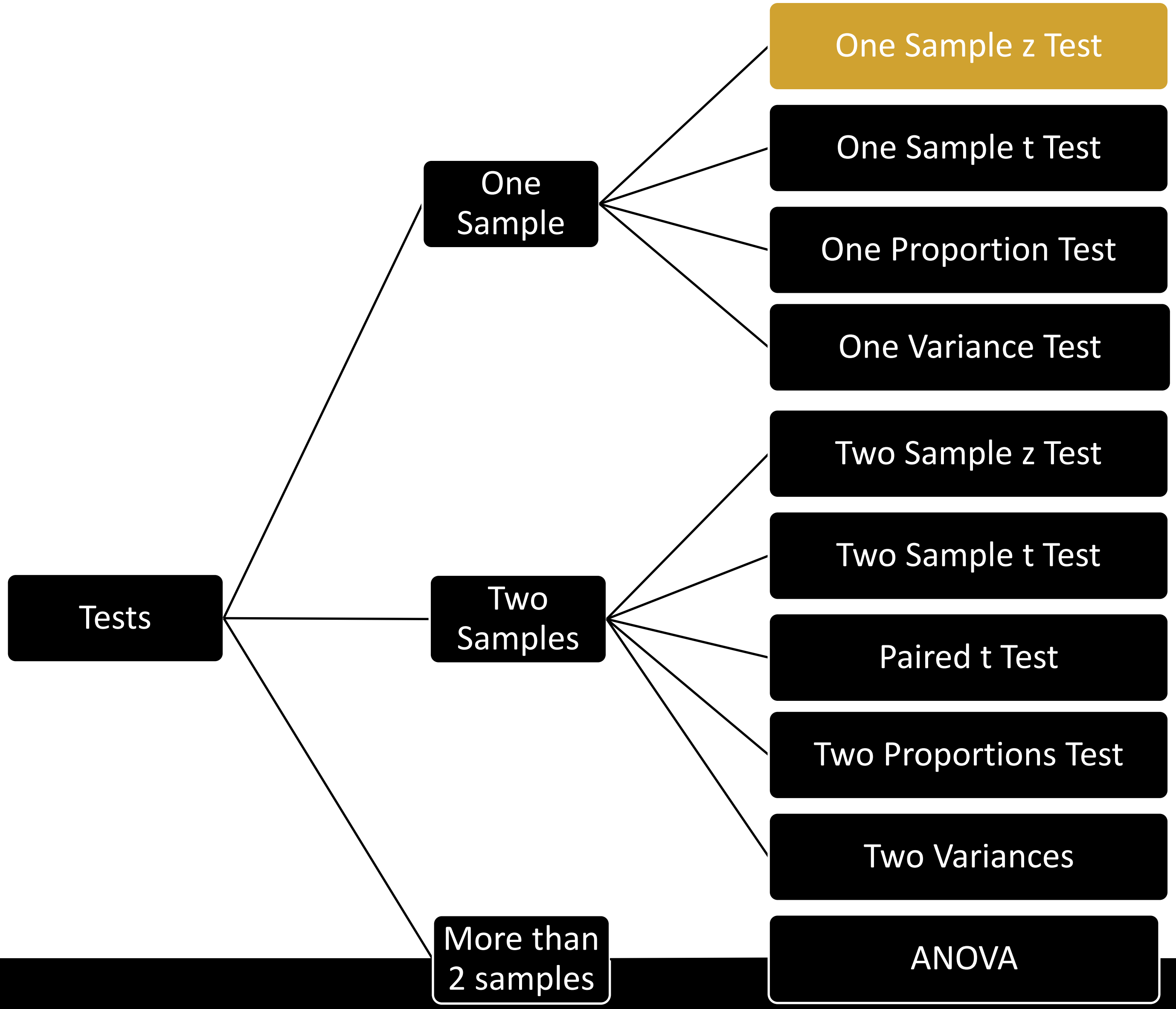
p Value

- p value is the lowest value of alpha for which the null hypothesis can be rejected. (Probability that the null hypothesis is correct)
- For example, if $p = 0.045$ you can reject the null hypothesis at $\alpha = 0.05$
- p is low the null must go / p is high the null fly.

Hypothesis Testing

4B-2
Tests for means, variances and proportions



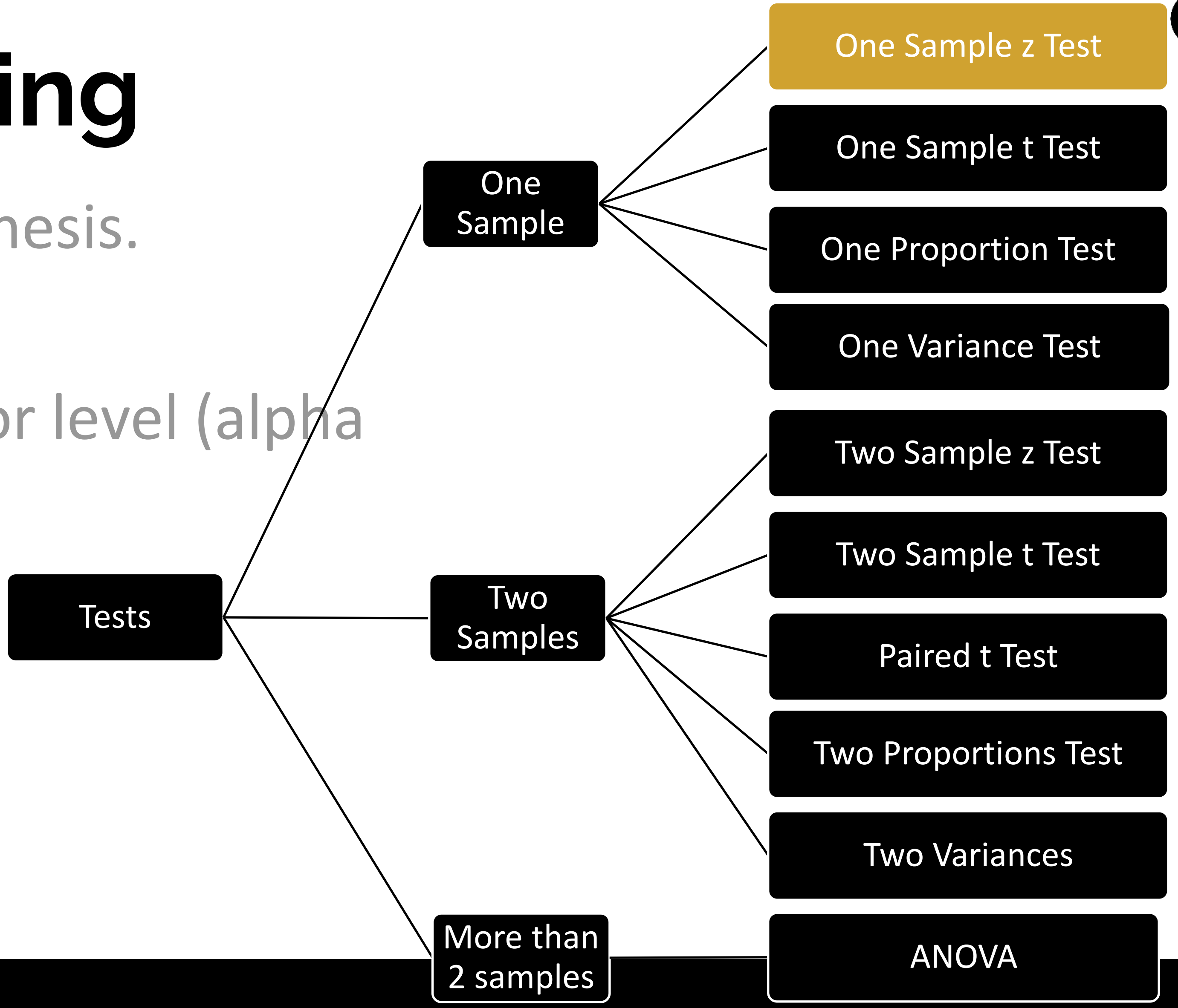


One Sample z Test

Hypothesis Testing

1. State the Alternate Hypothesis.
2. State the Null Hypothesis.
3. Select a probability of error level (alpha level). Generally 0.05
4. Calculate the test statistic

   (e.g t or z score)
5. Critical test statistic
6. Interpret the results.



One Sample z Test

Conditions for z Test

QG

- Random samples
- Each observation should be independent of other
  - Sampling with replacement
  - If sampling without replacement, the sample size should not be more than 10% of the population
- Sampling distribution approximates Normal Distribution
  - Population is Normally distributed and the population standard deviation is known *** OR ***
  - Sample size ≥ 30

Calculated Test Statistic

QG

H_0: μ = 150cc
H_a: μ ≠ 150cc

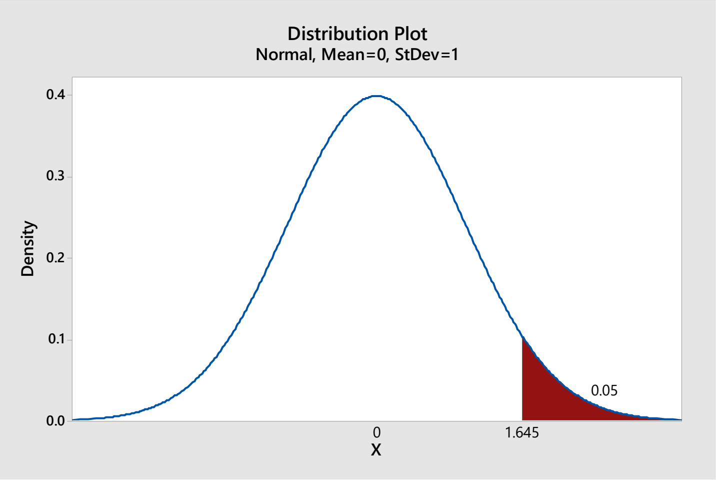
$$z_{cal} = \frac{(\bar{x} - \mu)}{\sigma/\sqrt{n}}$$

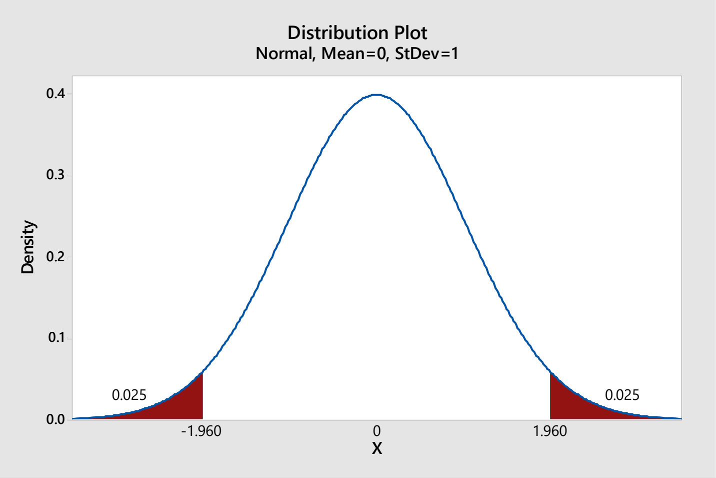
- Example: Perfume bottle producing 150cc with sd of 2 cc, 100 bottles are randomly picked and the average volume was found to be 150.2 cc. Has mean volume changed? (95% confidence)
- $z_{calculated}$ = (150.2-150)/[2 / sqrt(100)] = 0.2/0.2 = 1
- $z_{critical}$ = ?

One Sample z Test

Critical Test Statistic

z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000





- α = 0.05 One Tail
- Z Critical = 1.645

- α = 0.10 One Tail
- Z Critical = 1.282

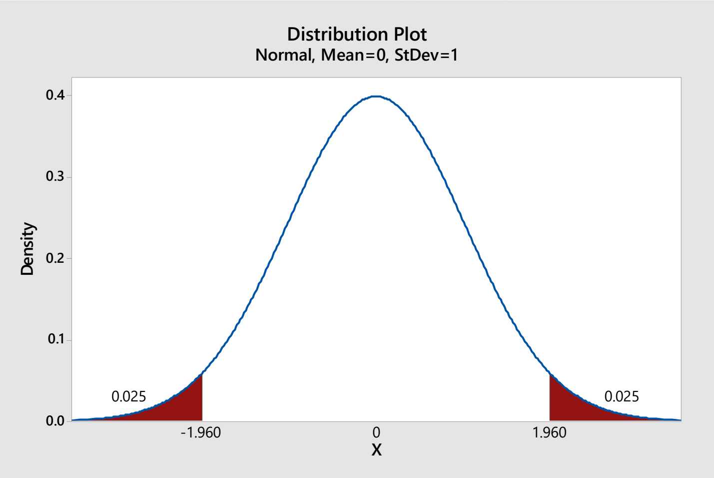
- α = 0.05 Two Tails
- Z Critical = 1.96

- α = 0.10 Two Tail
- Z Critical = 1.645

One Sample z Test

Critical Test Statistic

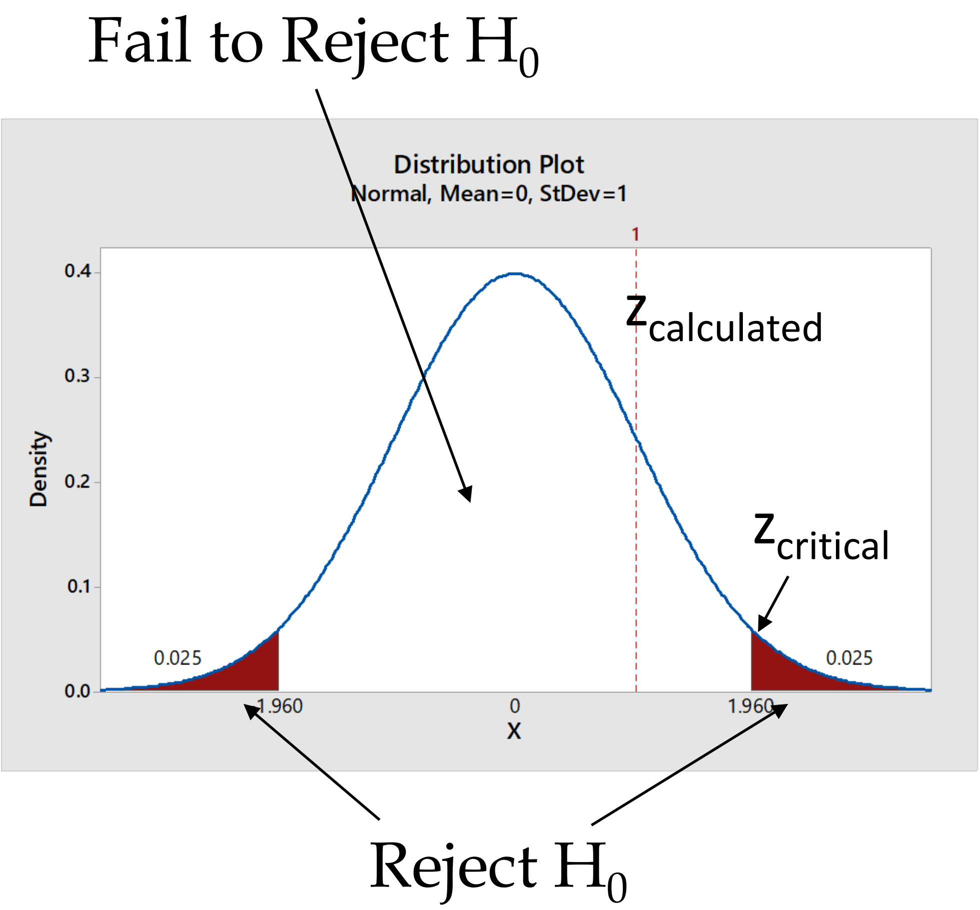
z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	[illegible]									.0183
2.1	[illegible]									.0143
2.2	[illegible]									.0110
2.3	[illegible]									.0084
2.4	[illegible]									.0064
2.5	[illegible]									.0048
2.6	[illegible]									.0036
2.7	[illegible]									.0026
2.8	[illegible]									.0019
2.9	[illegible]									.0014
3.0	[illegible]									.0010
3.1	[illegible]									.0007
3.2	[illegible]									.0005
3.3	[illegible]									.0003
3.4	[illegible]									.0002
3.5	[illegible]									.0002
3.6	[illegible]									.0001
3.7	[illegible]									.0001
3.8	[illegible]									.0001
3.9	[illegible]									.0000



- Example: Perfume bottle producing 150cc with sd of 2 cc, 100 bottles are randomly picked and the average volume was found to be 150.2 cc. Has mean volume changed? (95% confidence)
- $z_{calculated}$ = (150.2-150)/[2 / sqrt(100)] = 0.2/0.2 = 1
- $z_{critical}$ = 1.96

One Sample z Test

Interpret the Results



- Example: Perfume bottle producing 150cc with sd of 2 cc, 100 bottles are randomly picked and the average volume was found to be 150.2 cc. Has mean volume changed? (95% confidence)
- $z_{calculated}$ = (150.2-150)/[2 / sqrt(100)] = 0.2/0.2 = 1
- $z_{critical}$ = 1.96

One Sample z Test

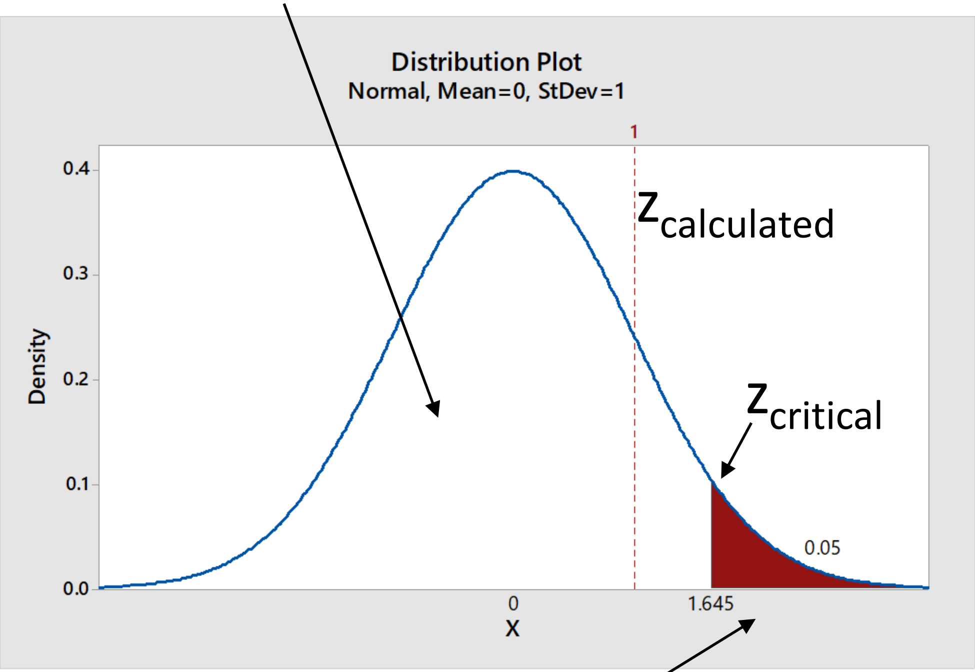
Interpret the Results

One Tail Tests

H_0: $\mu \le 150$cc

H_a: $\mu > 150$cc

Fail to Reject H_0



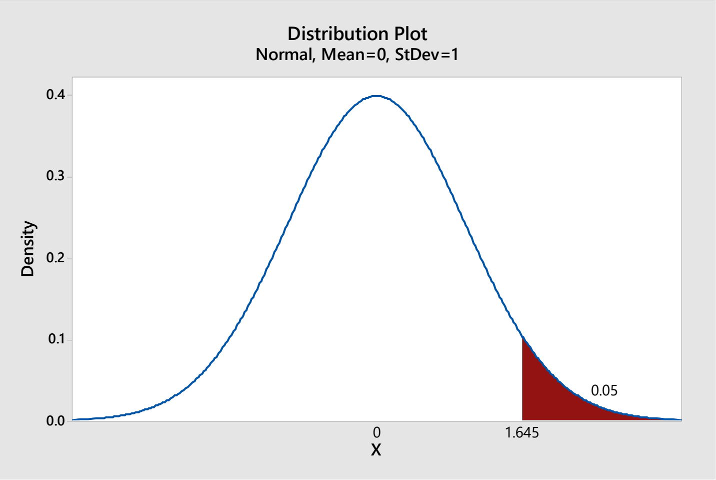
Reject H_0

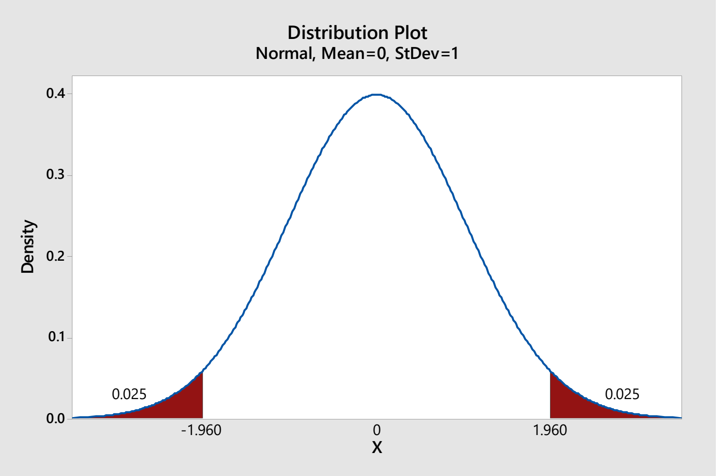
- Example: Perfume bottle producing 150cc with sd of 2 cc, 100 bottles are randomly picked and the average volume was found to be 150.2 cc. Has mean volume ~~changed~~ increased? (95% confidence)
- $z_{calculated}$ = (150.2-150)/[2 / sqrt(100)] = 0.2/0.2 = 1
- $z_{critical}$ = ~~1.96~~ 1.645

One Sample z Test

Critical Test Statistic

z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000





- α = 0.05 One Tail
- Z Critical = 1.645

- α = 0.10 One Tail
- Z Critical = 1.282

- α = 0.05 Two Tails
- Z Critical = 1.96

- α = 0.10 Two Tail
- Z Critical = 1.645

One Sample z Test

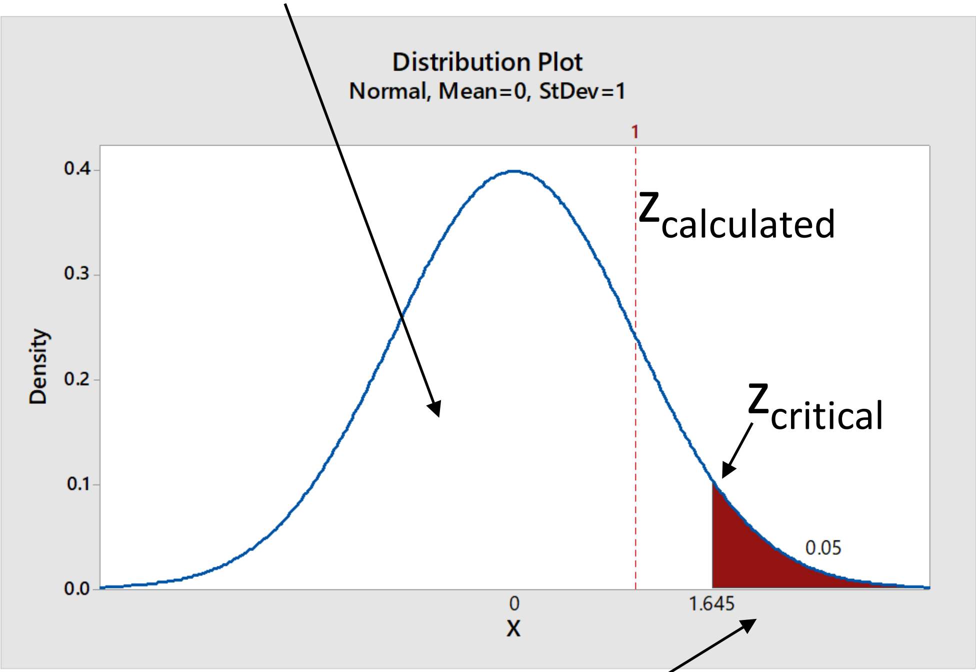
One Tail Tests

H_0: $\mu \leq 150$cc

H_a: $\mu > 150$cc

Interpret the Results

Fail to Reject H_0

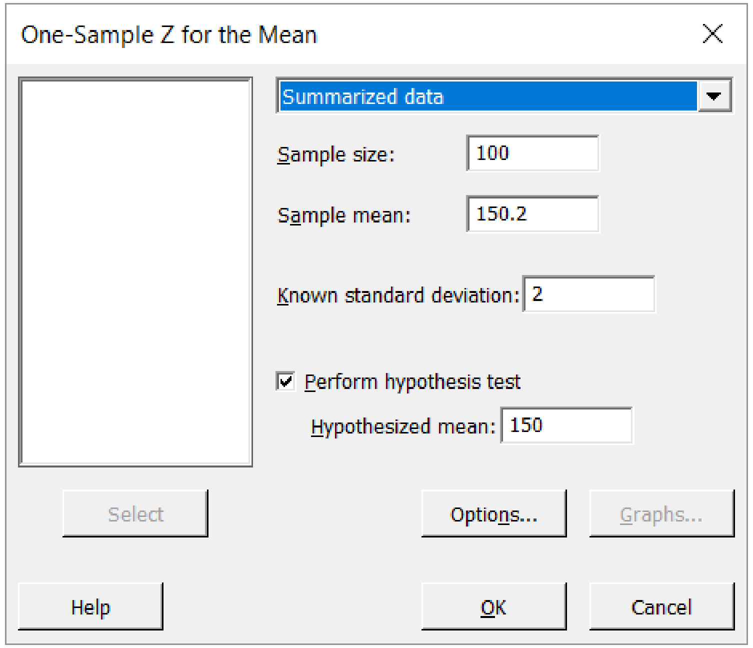


Reject H_0

- Example: Perfume bottle producing 150cc with sd of 2 cc, 100 bottles are randomly picked and the average volume was found to be 150.2 cc. Has mean volume ~~changed~~ increased? (95% confidence)
- $z_{calculated}$ = (150.2-150)/[2 / sqrt(100)] = 0.2/0.2 = 1
- $z_{critical}$ = ~~1.96~~ 1.645

One Sample z Test

One Sample z Test - Minitab



One-Sample Z

Descriptive Statistics

N	Mean	SE Mean	95% CI for μ
100	150.200	0.200	(149.808, 150.592)

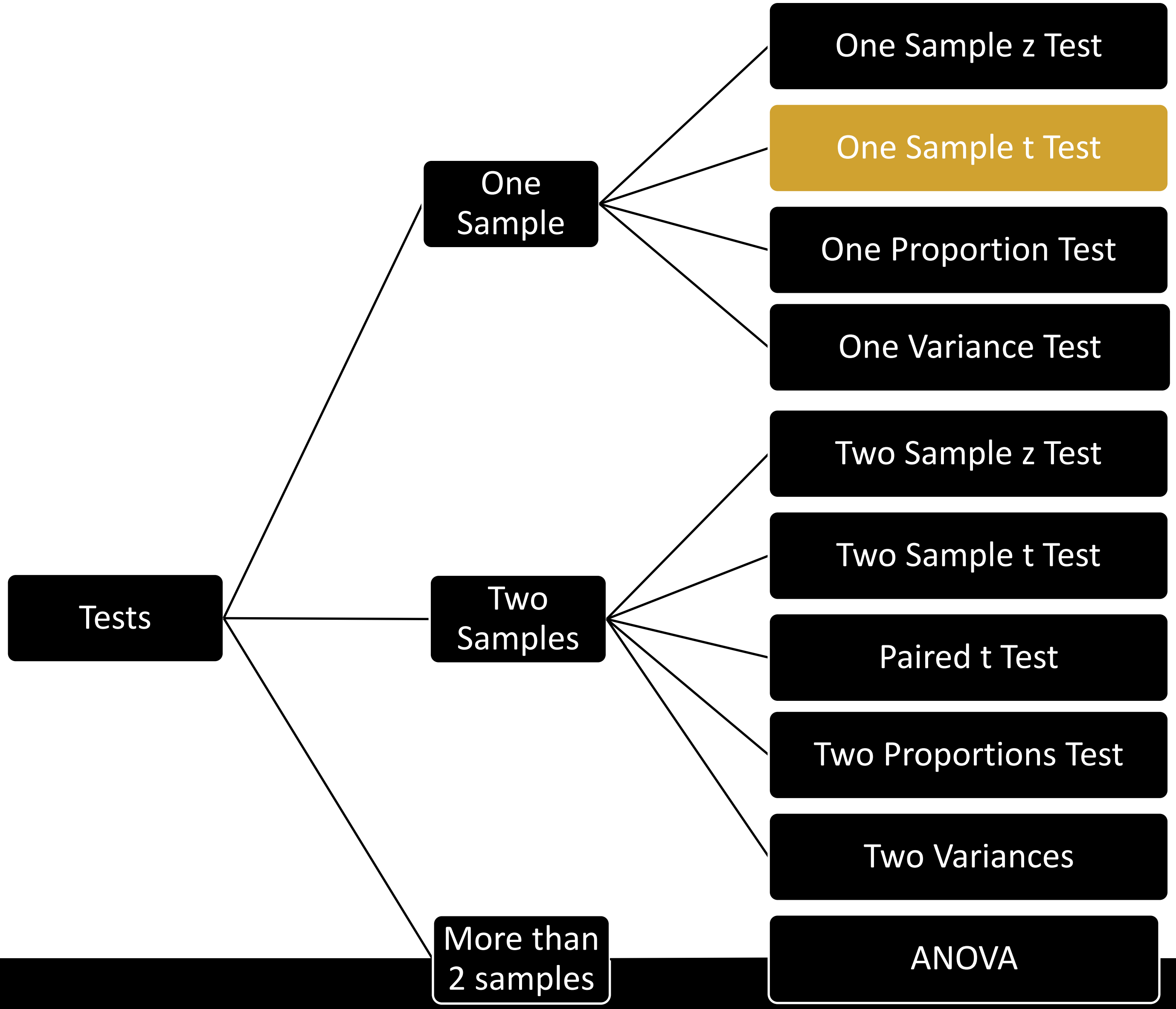
μ: mean of Sample
Known standard deviation = 2

Test

Null hypothesis	H_0: μ = 150
Alternative hypothesis	H_1: μ ≠ 150

Z-Value	P-Value
1.00	0.317

One Sample z Test



One Sample t Test

Conditions for t Test

- Random samples
- Each observation should be independent of other
  - Sampling with replacement
  - If sampling without replacement, the sample size should not be more than 10% of the population
- Sampling distribution approximates Normal Distribution
  - Population is Normally distributed and the standard deviation is **unknown** *** AND ***
  - Sample size < 30

One Sample t Test

Conditions for t Test

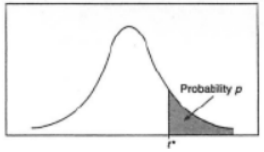
H_0: μ = 150cc
H_a: μ ≠ 150cc

$$t_{cal} = \frac{(\bar{x} - \mu)}{s/\sqrt{n}}$$

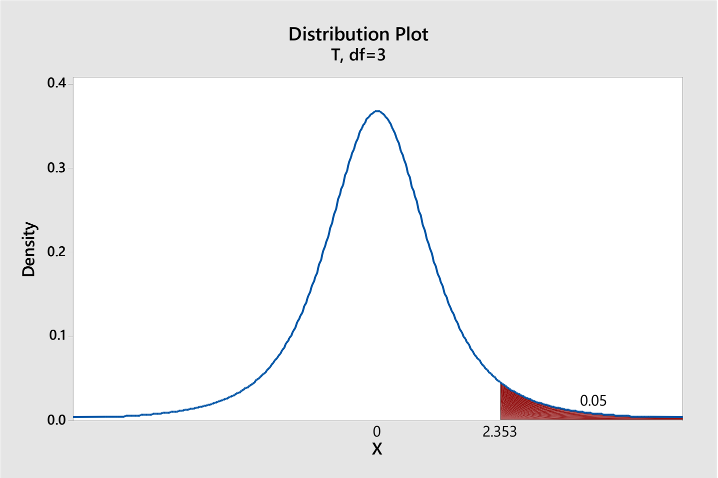
- Example: Perfume bottle producing 150cc, 4 bottles are randomly picked and the average volume was found to be 151cc and sd of the sample bottles was 2 cc. Has mean volume changed? (95% confidence)
- $t_{calculated} = \frac{(\bar{x}-\mu)}{s/\sqrt{n}} = \frac{(151-150)}{2/\sqrt{4}} = \frac{1}{1} = 1$
- $t_{critical}$ = ?

One Sample t Test

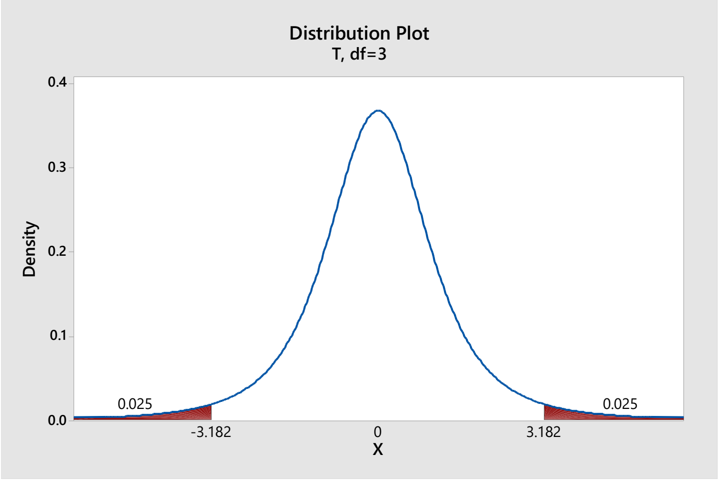
Critical Test Statistic



	TAIL PROBABILITY P											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646



- $\alpha = 0.05$ One Tails
- Df = 3
- t Critical = 2.353

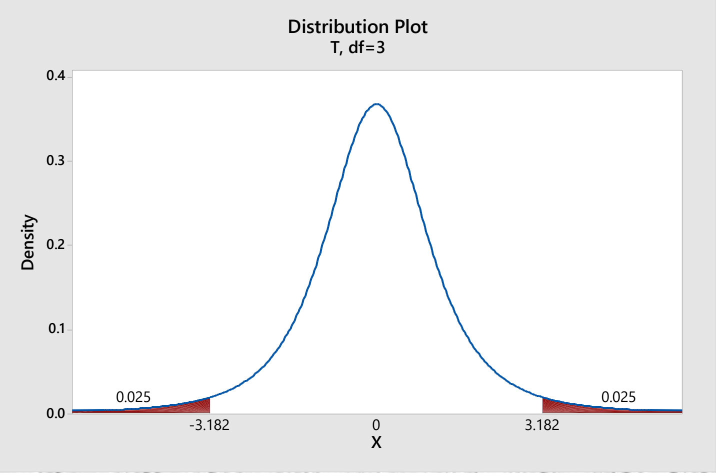


- $\alpha = 0.05$ Two Tails
- Df = 3
- t Critical = 3.182

One Sample t Test

Critical Test Statistic

df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
	TAIL PROBABILITY P											
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3.852	4.221
14	.692										3.787	4.140
15	.691										3.733	4.073
16	.690										3.686	4.015
17	.689										3.646	3.965
18	.688										3.611	3.922
19	.688										3.579	3.883
20	.687										3.552	3.850
21	.686										3.527	3.819
22	.686										3.505	3.792
23	.685										3.485	3.768
24	.685										3.467	3.745
25	.684										3.450	3.725
26	.684										3.435	3.707
27	.684										3.421	3.690
28	.683										3.408	3.674
29	.683										3.396	3.659
30	.683										3.385	3.646



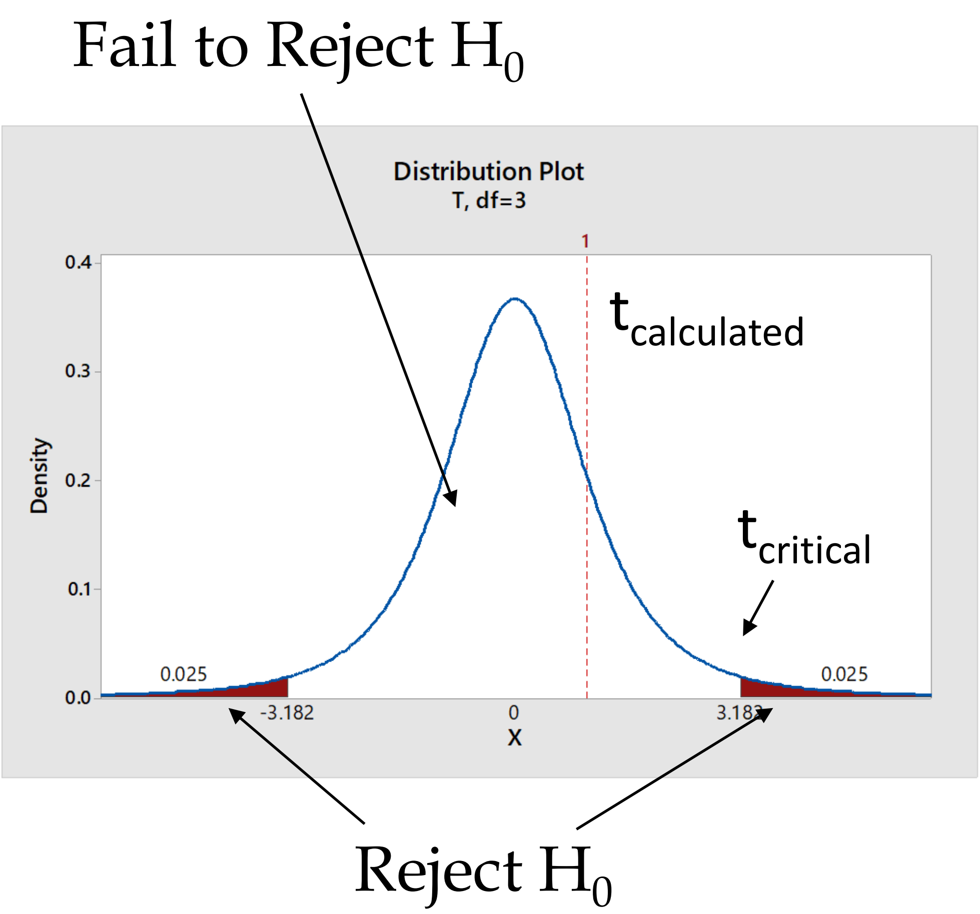
- Example: Perfume bottle producing 150cc, <u>4 bottles</u> are randomly picked and the average volume was found to be 151cc and sd of the <u>sample bottles</u> was 2 cc. Has mean volume changed? (95% confidence)

- $t_{calculated} = \frac{(\bar{x}-\mu)}{s/\sqrt{n}} = \frac{(151-150)}{2/\sqrt{4}} = \frac{1}{1} = 1$

- $t_{critical} = 3.182$

One Sample t Test

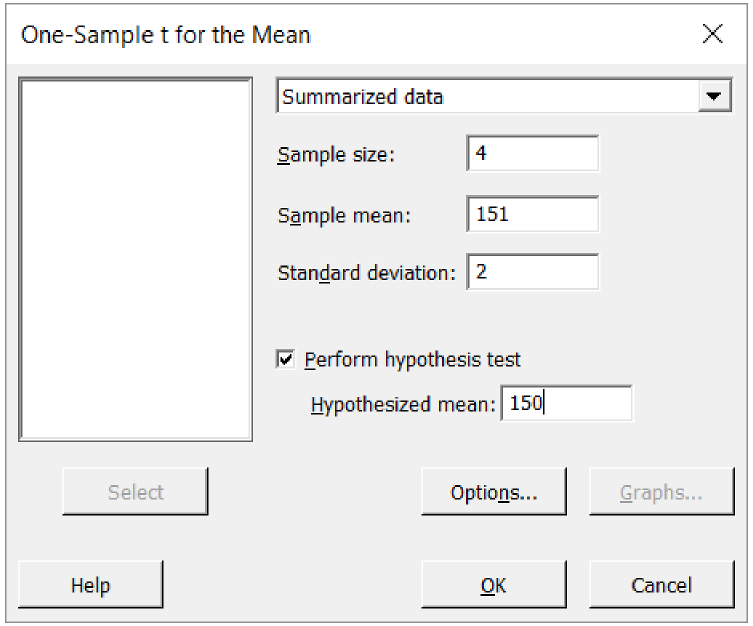
Interpret the Results



- Example: Perfume bottle producing 150cc, 4 bottles are randomly picked and the average volume was found to be 151cc and sd of the sample bottles was 2 cc. Has mean volume changed? (95% confidence)
- $t_{calculated} = \frac{(\bar{x}-\mu)}{s/\sqrt{n}} = \frac{(151-150)}{2/\sqrt{4}} = \frac{1}{1} = 1$
- $t_{critical} = 3.182$

One Sample t Test

One Sample t Test - Minitab



One-Sample T

Descriptive Statistics

N	Mean	StDev	SE Mean	95% CI for μ
4	151.00	2.00	1.00	(147.82, 154.18)

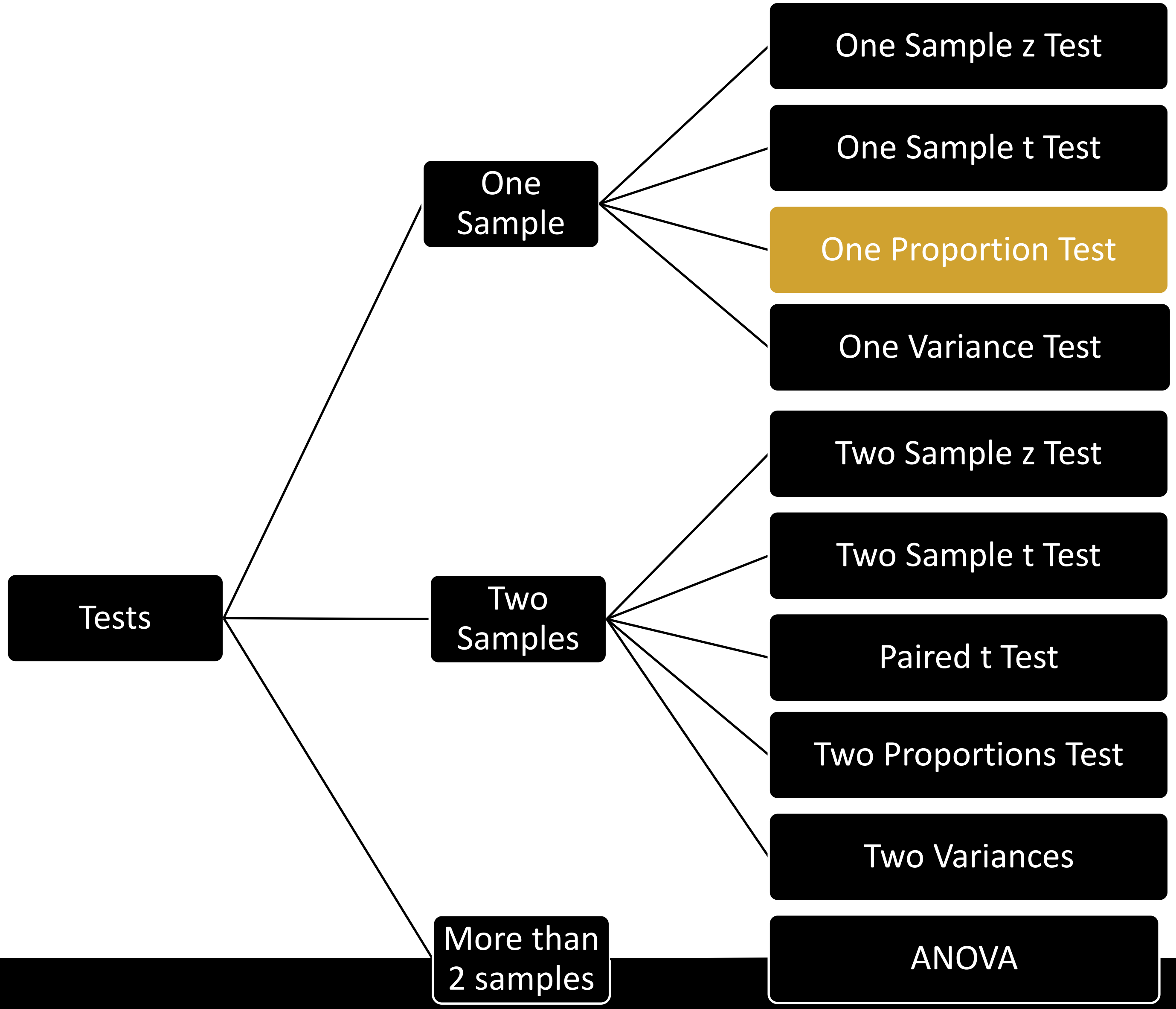
μ: mean of Sample

Test

Null hypothesis	H_0: μ = 150
Alternative hypothesis	H_1: μ ≠ 150

T-Value	P-Value
1.00	0.391

One Sample t Test



One Proportion Test

Conditions for One Proportion Test

- Random samples
- Each observation should be independent of other
  - Sampling with replacement
  - If sampling without replacement, the sample size should not be more than 10% of the population
- The data contains only two categories, such as pass/fail or yes/no
- For Normal approximation:
  - both $np \geq 10$ and $n(1-p) \geq 10$ (data should have at least 10 "successes" and at least 10 "failures") (in some books it is 5)

One Proportion Test

H_0: $p = p_0$
H_a: $p \neq p_0$

$$z = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

- Example: Smoking rate in a town in past was 21%, 100 samples were picked and found 14 smokers. Has smoking habit changed?
- Can Normality assumption be made?
  - $p_0 = 0.21$, $p=0.14$
  - $np_0 = 0.21x100 = 21$
  - $n(1-p_0) = 0.79x100 = 79$
  - >10 means sample size is sufficient.

One Proportion Test

One Proportion Test

H_0: $p = p_0$
H_a: $p \neq p_0$

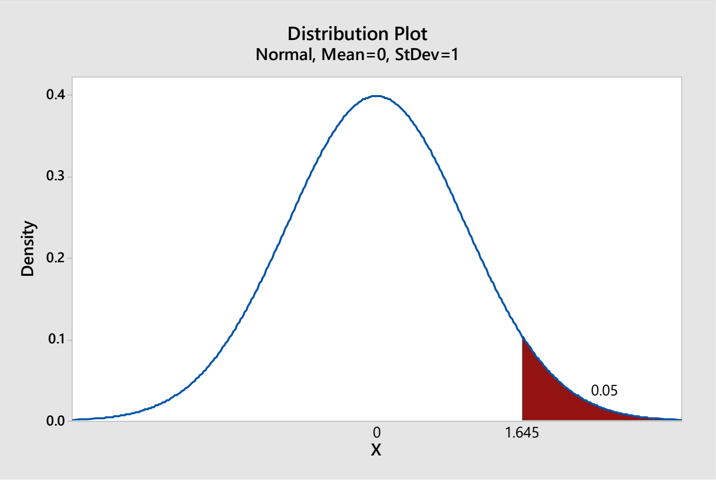
$$z = \frac{p - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

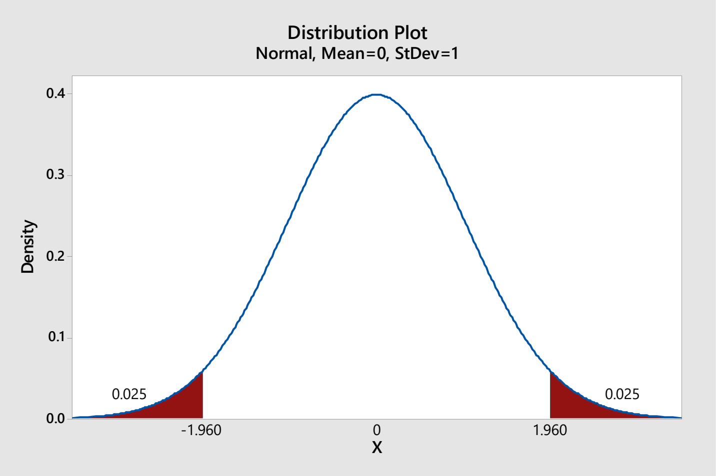
- Example: Smoking rate in a town in past was 21%, 100 samples were picked and found 14 smokers. Has smoking habit changed?
- $z_{calculated} = \frac{p-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\mathbf{0.14-0.21}}{\sqrt{\frac{\mathbf{0.21(1-0.21)}}{\mathbf{100}}}} = -1.719$
- $z_{critical} = ?$

One Proportion Test

Critical Test Statistic

z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000





- α = 0.05 One Tail
- Z Critical = 1.645

- α = 0.10 One Tail
- Z Critical = 1.282

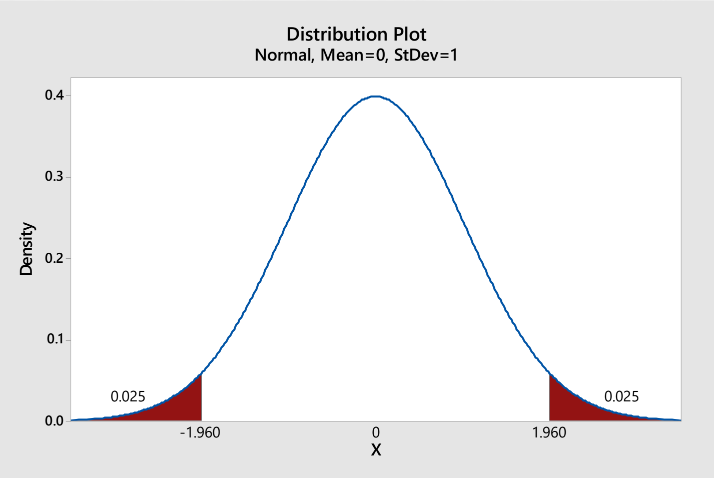
- α = 0.05 Two Tails
- Z Critical = 1.96

- α = 0.10 Two Tail
- Z Critical = 1.645

One Proportion Test

One Proportion Test

z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.022									.0183
2.1	.017									.0143
2.2	.013									.0110
2.3	.010									.0084
2.4	.008									.0064
2.5	.006									.0048
2.6	.004									.0036
2.7	.003									.0026
2.8	.002									.0019
2.9	.001									.0014
3.0	.001									.0010
3.1	.001									.0007
3.2	.000									.0005
3.3	.000									.0003
3.4	.000									.0002
3.5	.000									.0002
3.6	.000									.0001
3.7	.000									.0001
3.8	.000									.0001
3.9	.000									.0000

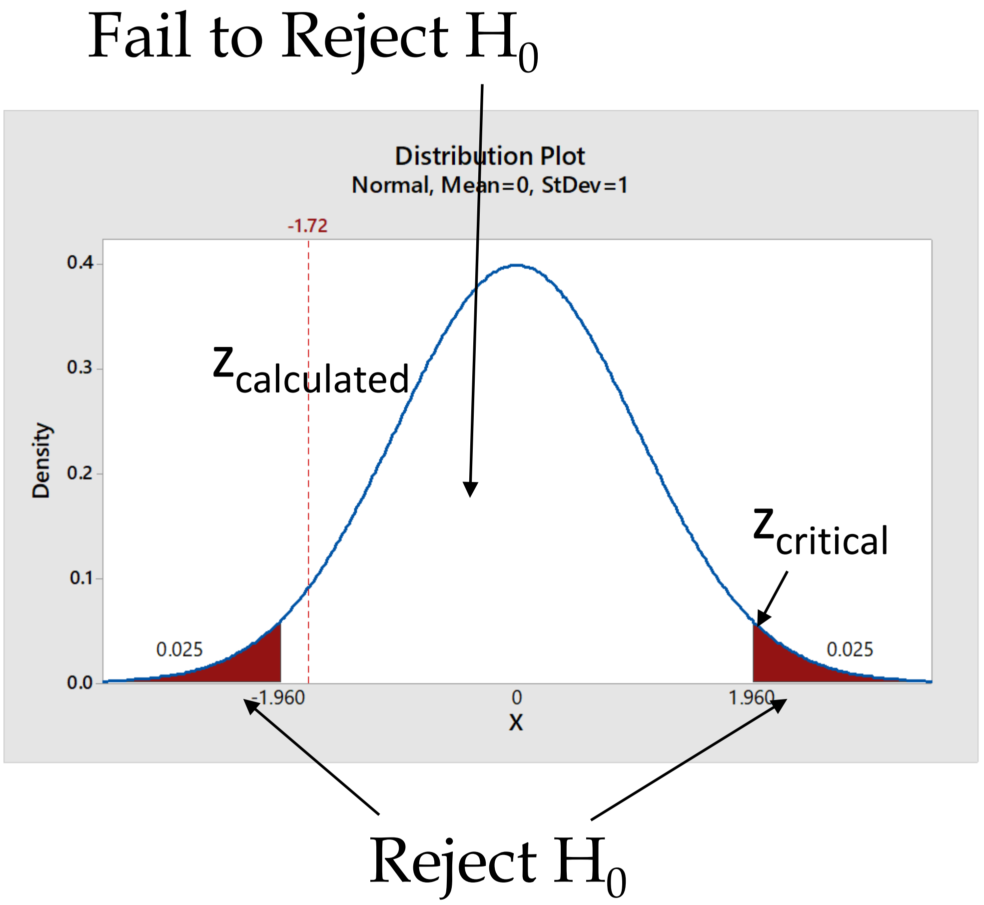


- Example: Smoking rate in a town in past was 21%, 100 samples were picked and found 14 smokers. Has smoking habit changed? (95% confidence)
- $z_{calculated} = \dfrac{p - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \dfrac{\mathbf{0.14 - 0.21}}{\sqrt{\dfrac{\mathbf{0.21(1-0.21)}}{\mathbf{100}}}} = -1.719$
- $z_{critical} = 1.96$

One Proportion Test

Interpret the Results

H_0: $p = p_0$
H_a: $p \neq p_0$



- Example: Smoking rate in a town in past was 21%, 100 samples were picked and found 14 smokers. Has smoking habit changed? (95% confidence)
- $z_{calculated} = \dfrac{p - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \dfrac{\mathbf{0.14 - 0.21}}{\sqrt{\dfrac{\mathbf{0.21(1-0.21)}}{\mathbf{100}}}} = -1.719$
- $z_{critical} = 1.96$

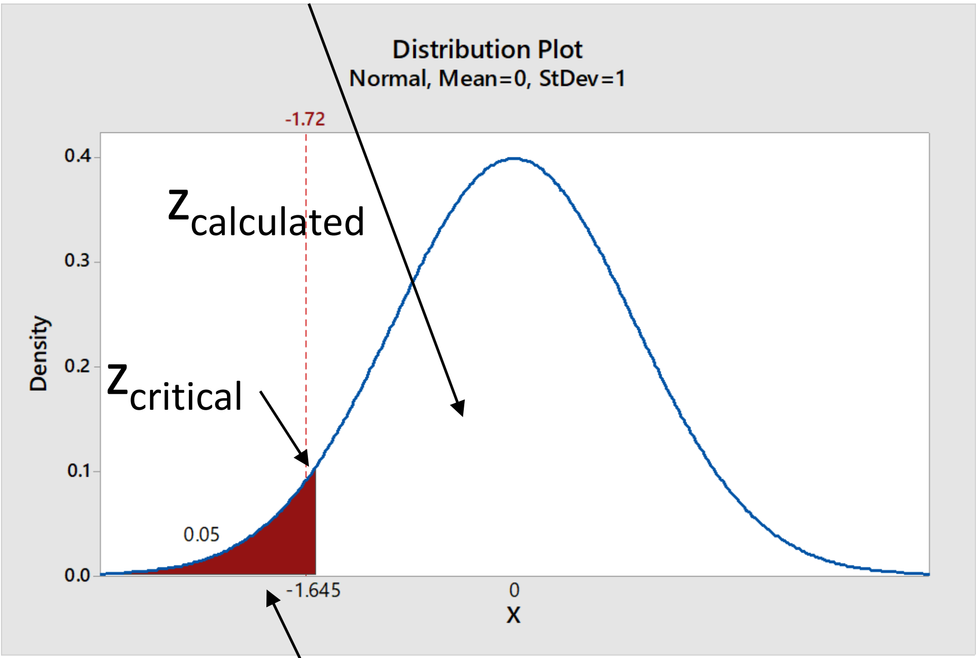
One Proportion Test

One Tail Tests

H_0: $p \geq p_0$

H_a: $p < p_0$

Fail to Reject H_0



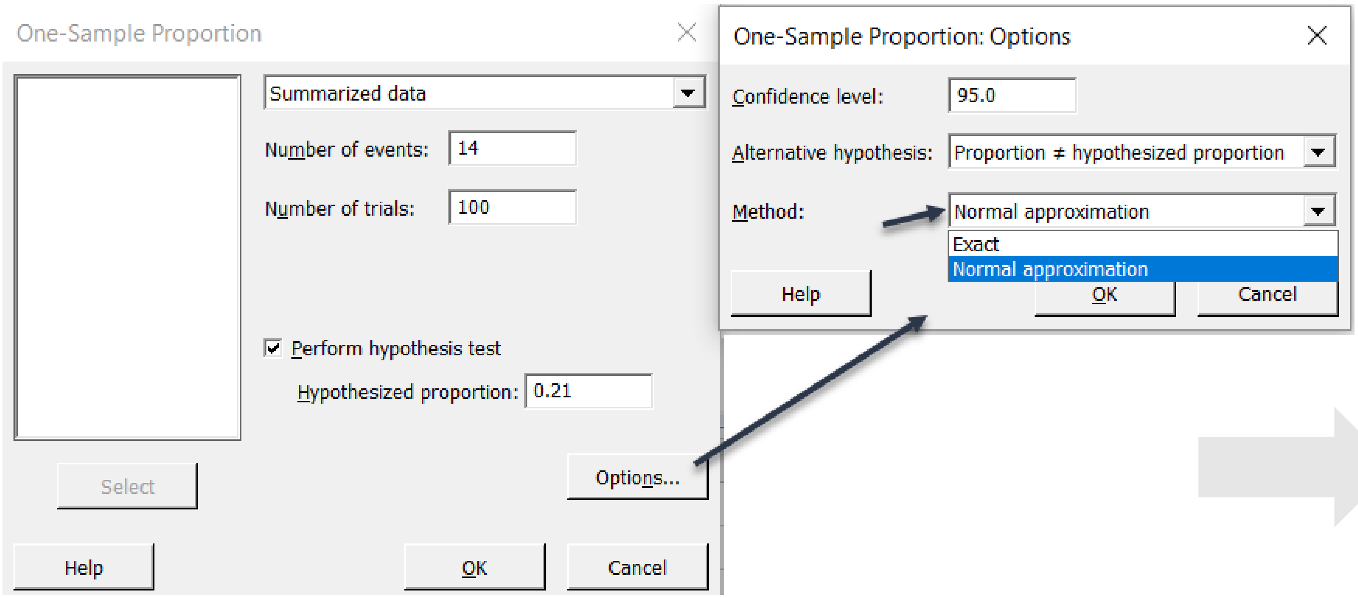
Reject H_0

Interpret the Results

- Example: Smoking rate in a town in past was 21%, 100 samples were picked and found 14 smokers. Has smoking habit <u>reduced</u> at 95% confidence? (one tail)
- $z_{calculated} = \dfrac{p - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \dfrac{\mathbf{0.14 - 0.21}}{\sqrt{\dfrac{\mathbf{0.21(1-0.21)}}{\mathbf{100}}}} = -1.719$
- $z_{critical}$ = ~~1.96~~ 1.645

One Proportion Test

One Proportion Test - Minitab



Test and CI for One Proportion

Method

p: event proportion
Normal approximation method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	95% CI for p
100	14	0.140000	(0.071992, 0.208008)

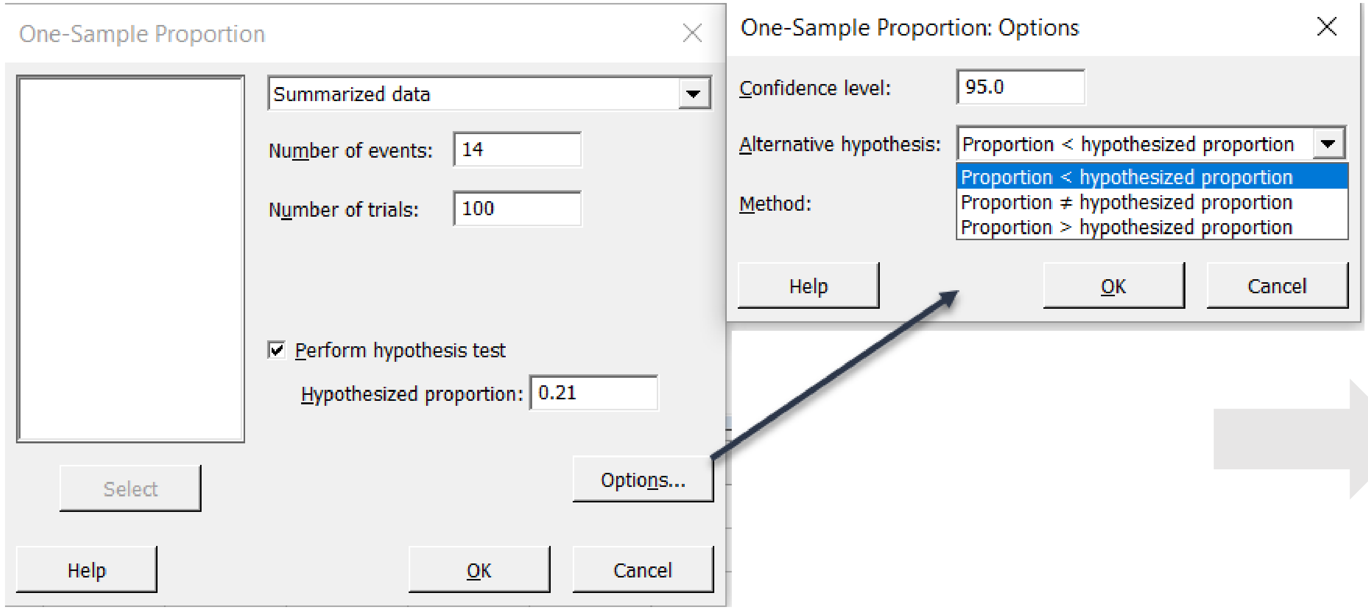
Test

Null hypothesis	H_0: p = 0.21
Alternative hypothesis	H_1: p ≠ 0.21

Z-Value	P-Value
-1.72	0.086

One Proportion Test

One Proportion Test - Minitab



Test and CI for One Proportion

Method

p: event proportion
Normal approximation method is used for this analysis.

Descriptive Statistics

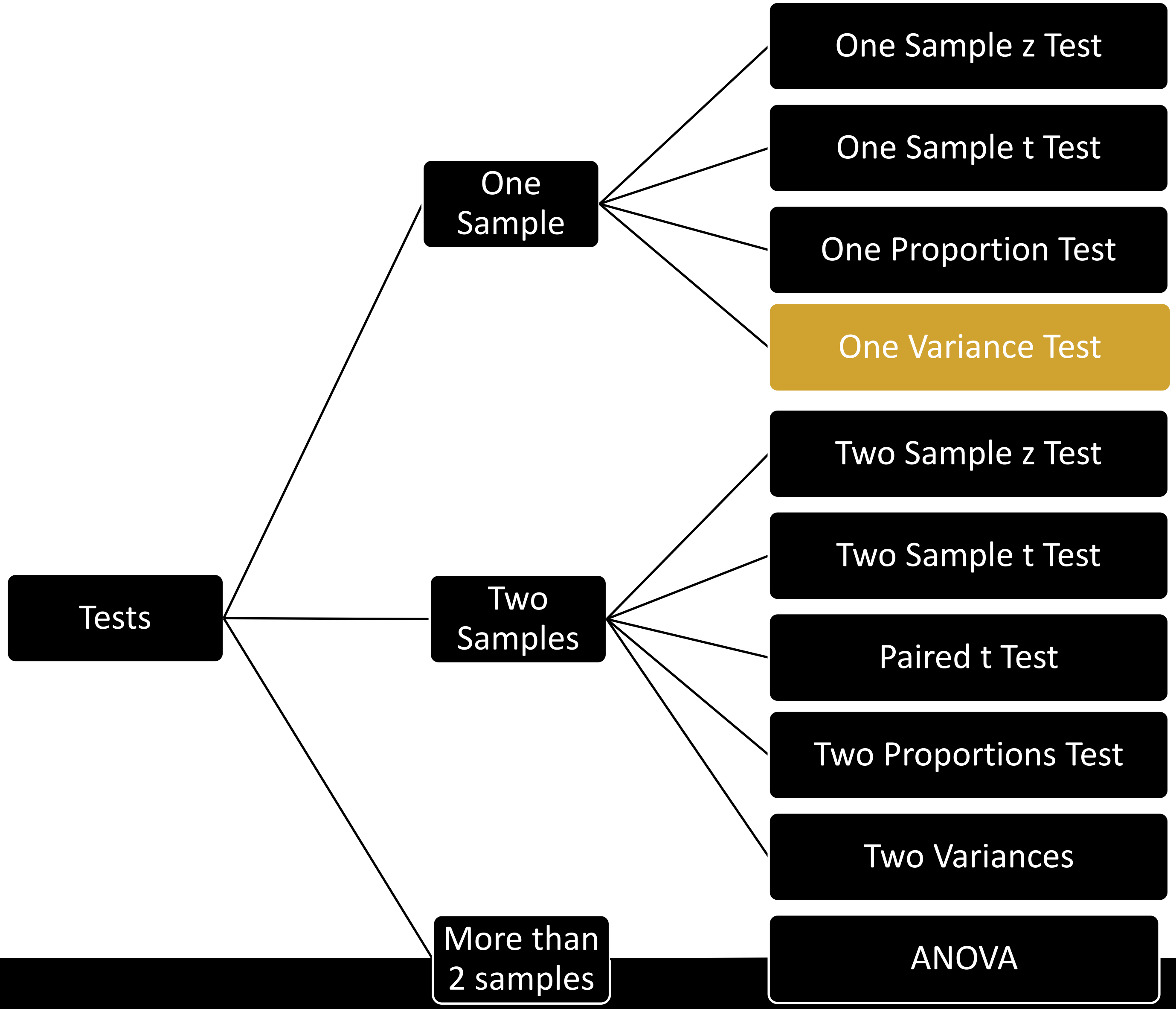
N	Event	Sample p	95% Upper Bound for p
100	14	0.140000	0.197074

Test

Null hypothesis	H_0: p = 0.21
Alternative hypothesis	H_1: p < 0.21

Z-Value	P-Value
-1.72	0.043

One Proportion Test



One Variance Test

Conditions for One Variance Test

- Random samples
- Each observation should be independent of other
  - Sampling with replacement
  - If sampling without replacement, the sample size should not be more than 10% of the population
- The data follows a Normal Distribution

Variance Tests

- Chi-square test
  - For testing the population variance against a specified value
  - testing goodness of fit of some probability distribution
  - testing for independence of two attributes (Contingency Tables)

- F-test
  - for testing equality of *two* variances from different populations
  - for testing equality of several means with technique of ANOVA.

One Variance Test

One Variance Test

Ho: $s^2 <= \sigma^2$
Ha: $s^2 > \sigma^2$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

- Example 1: A sample of 51 bottles was selected. The standard deviation of these 51 bottles was 2.35 cc. Has it **increased** from established 2 cc? 90% confidence level.

- $\chi^2(cal.) = \frac{(n-1)s^2}{\sigma^2} = \frac{\mathbf{(50)\ x\ 2.35^2}}{\mathbf{2^2}} = 69.03$

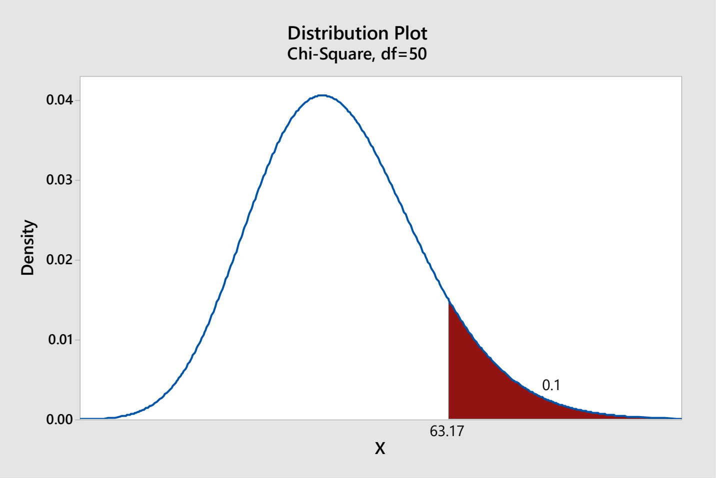
- What is critical value of Chi Square for 50 degrees of freedom?

One Variance Test

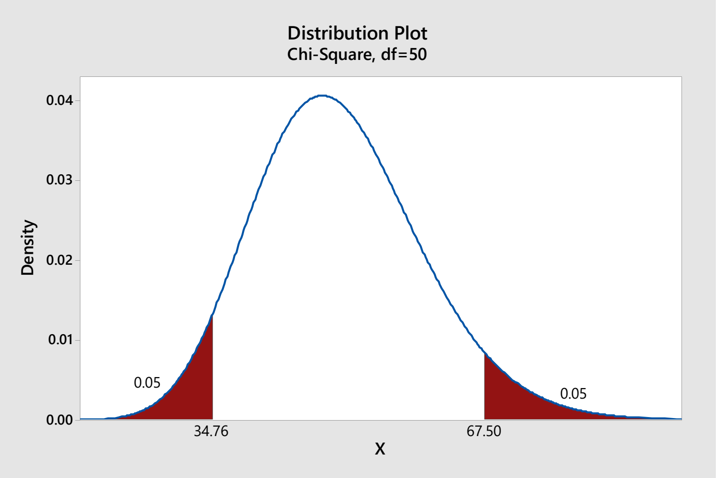
Critical Test Statistic

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41
18	7.015	9.390	10.865	13.675	17.338	21.60	25.99	28.87	34.80
19	7.633	10.117	11.651	14.562	18.338	22.72	27.20	30.14	36.19
20	8.260	10.851	12.443	15.452	19.337	23.83	28.41	31.41	37.57
22	9.542	12.338	14.041	17.240	21.337	26.04	30.81	33.92	40.29
24	10.856	13.848	15.659	19.037	23.337	28.24	33.20	36.42	42.98
26	12.198	15.379	17.292	20.843	25.336	30.43	35.56	38.89	45.64
28	13.565	16.928	18.939	22.657	27.336	32.62	37.92	41.34	48.28
30	14.953	18.493	20.599	24.478	29.336	34.80	40.26	43.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38



- α = 0.10 One Tail
- Df = 50
- χ^2 Critical = 63.17



- α = 0.10 Two Tail
- Df = 50
- χ^2Critical = 34.76 and 67.50

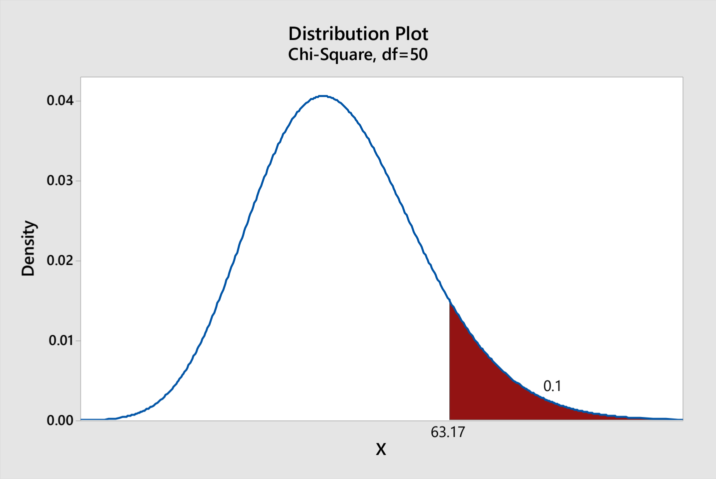
One Variance Test

Critical Test Statistic

QG

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11							.28	19.68	24.72
12							.55	21.03	26.22
13							.81	22.36	27.69
14							.06	23.68	29.14
15							.31	25.00	30.58
16							.54	26.30	32.00
17							.77	27.59	33.41
18							.99	28.87	34.80
19							.20	30.14	36.19
20							.41	31.41	37.57
22							.81	33.92	40.29
24							.20	36.42	42.98
26							.56	38.89	45.64
28							.92	41.34	48.28
30							.26	43.77	50.89
40							.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38

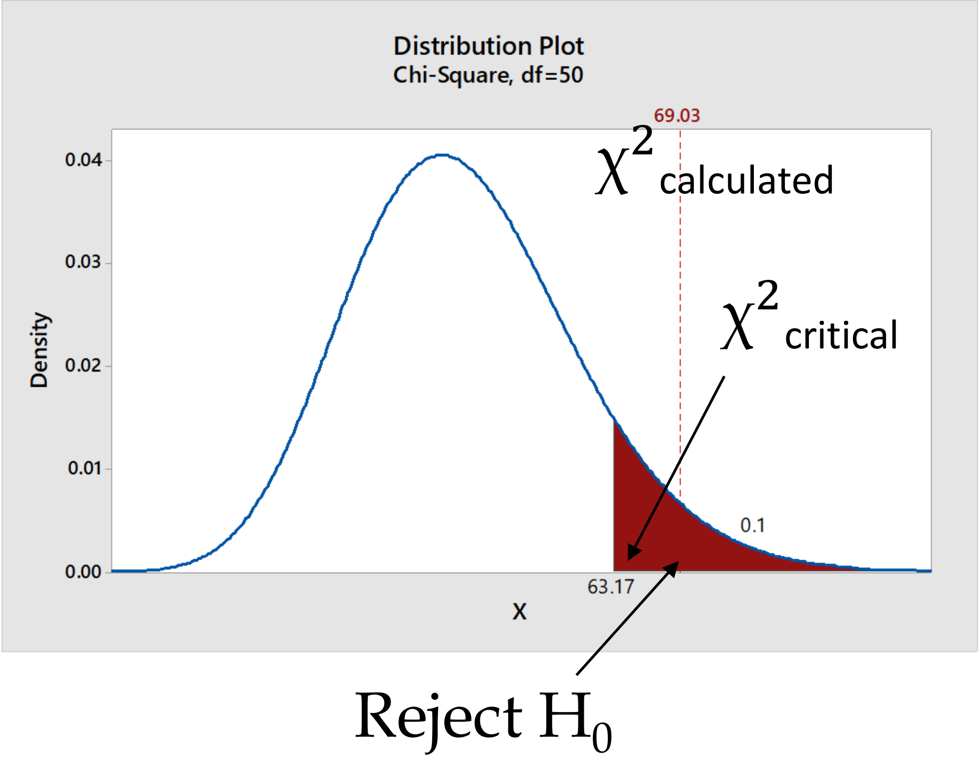


- Example 1: A sample of 51 bottles was selected. The standard deviation of these 51 bottles was 2.35 cc. Has it **increased** from established 2 cc? 90% confidence level.
- $\chi^2(cal.) = \frac{(n-1)s^2}{\sigma^2} = \frac{(50)\ x\ 2.35^2}{2^2} = 69.03$
- $\chi^2(\text{critical}) = 63.17$

One Variance Test

One Variance Test

Ho: $s^2 <= \sigma^2$
Ha: $s^2 > \sigma^2$



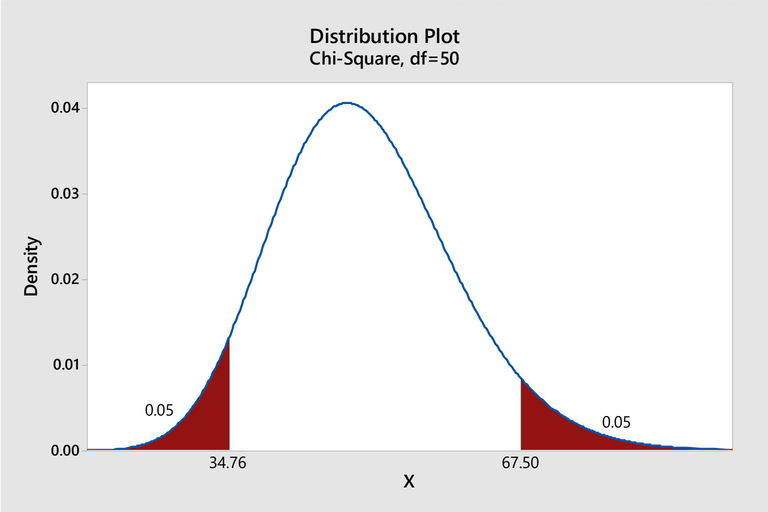
- Example 1: A sample of 51 bottles was selected. The standard deviation of these 51 bottles was 2.35 cc. Has it **increased** from established 2 cc? 90% confidence level.
- $\chi^2(cal.) = \frac{(n-1)s^2}{\sigma^2} = \frac{(50)\ x\ 2.35^2}{2^2} = 69.03$
- $\chi^2(\text{critical}) = 63.17$

One Variance Test

One Variance Test

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10								18.31	23.21
11								19.68	24.72
12								21.03	26.22
13								22.36	27.69
14								23.68	29.14
15								25.00	30.58
16								26.30	32.00
17								27.59	33.41
18								28.87	34.80
19								30.14	36.19
20								31.41	37.57
22								33.92	40.29
24								36.42	42.98
26								38.89	45.64
28								41.34	48.28
30								43.77	50.89
40								55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38



- **Example 2:** A sample of 51 bottles was selected. The standard deviation of these 51 bottles was 2.35 cc. Has it **changed** from established 2 cc? 90% confidence level.

- $\chi^2(cal.) = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{(50)\ x\ 2.35^2}{2^2} = 69.03$

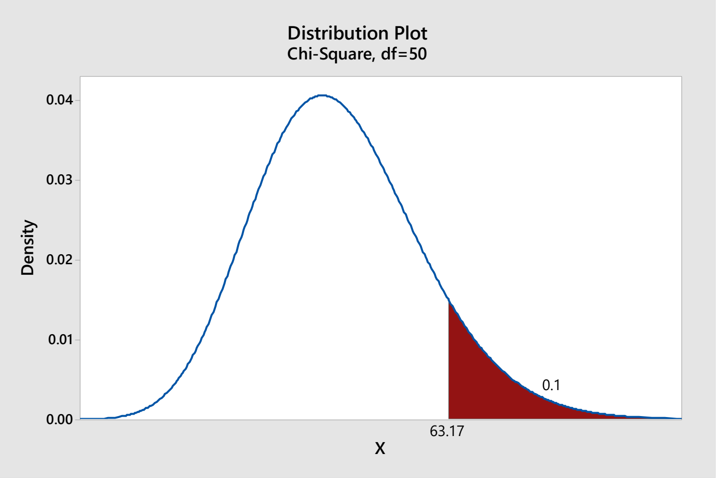
- What is critical value of Chi Square for 50 degrees of freedom? (two tails test)

One Variance Test

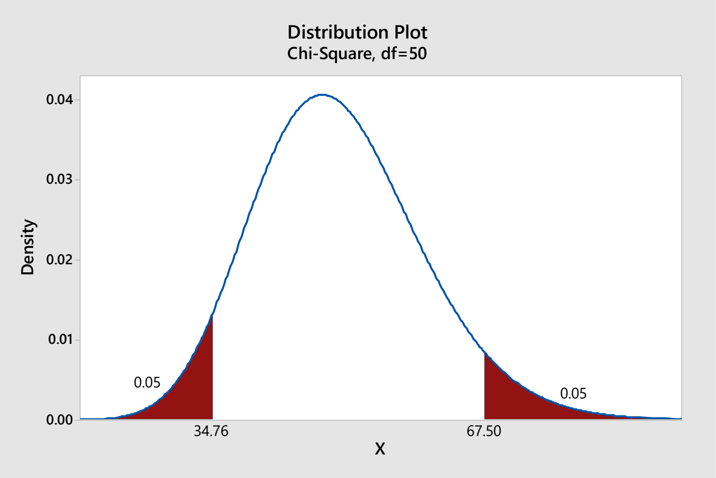
Critical Test Statistic

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	**0.95**	**0.90**	**0.75**	**0.50**	**0.25**	**0.10**	**0.05**	**0.01**
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41
18	7.015	9.390	10.865	13.675	17.338	21.60	25.99	28.87	34.80
19	7.633	10.117	11.651	14.562	18.338	22.72	27.20	30.14	36.19
20	8.260	10.851	12.443	15.452	19.337	23.83	28.41	31.41	37.57
22	9.542	12.338	14.041	17.240	21.337	26.04	30.81	33.92	40.29
24	10.856	13.848	15.659	19.037	23.337	28.24	33.20	36.42	42.98
26	12.198	15.379	17.292	20.843	25.336	30.43	35.56	38.89	45.64
28	13.565	16.928	18.939	22.657	27.336	32.62	37.92	41.34	48.28
30	14.953	18.493	20.599	24.478	29.336	34.80	40.26	43.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38



- α = 0.10 One Tail
- Df = 50
- χ^2 Critical = 63.17

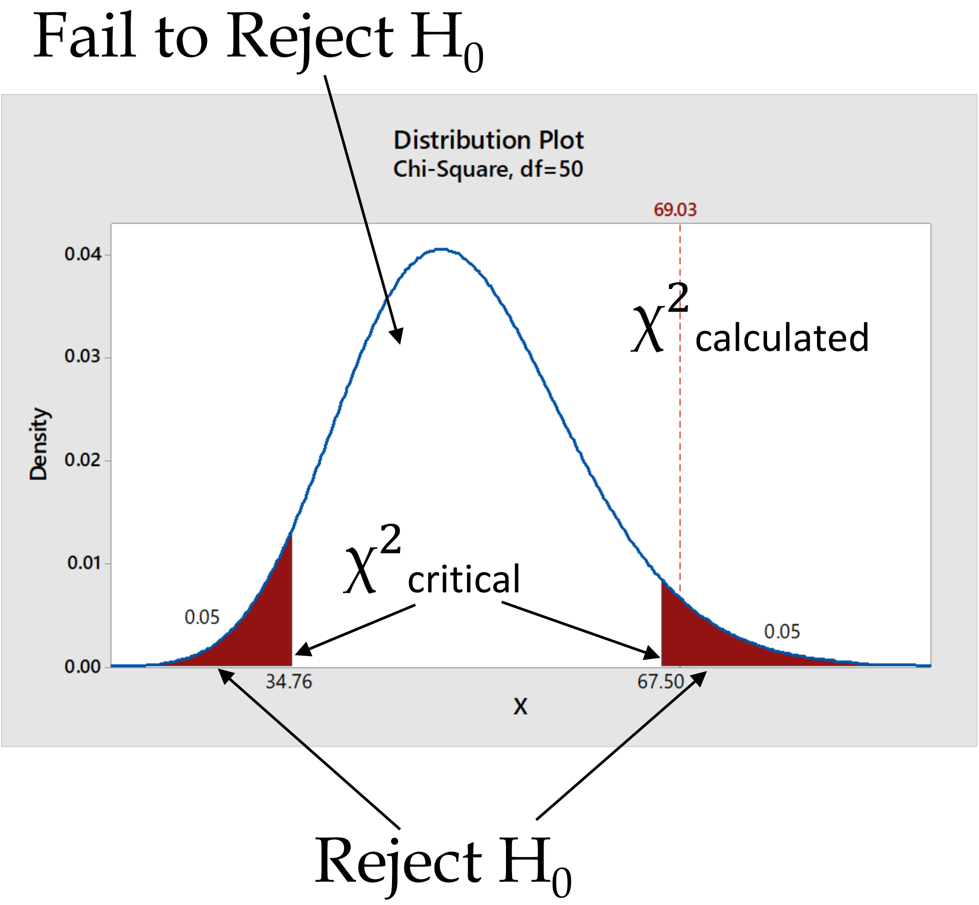


- α = 0.10 Two Tail
- Df = 50
- χ^2Critical = 34.76 and 67.50

One Variance Test

One Variance Test

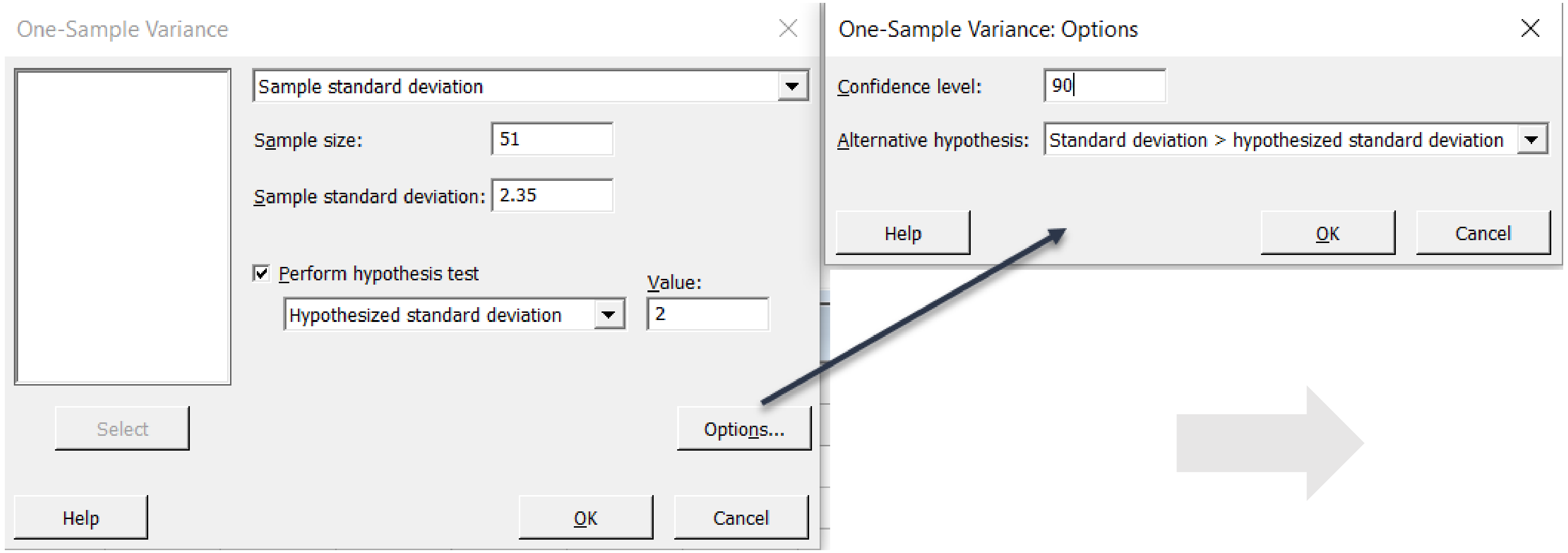
Ho: $s^2 = \sigma^2$
Ha: $s^2 \neq \sigma^2$



- **Example 2:** A sample of 51 bottles was selected. The standard deviation of these 51 bottles was 2.35 cc. Has it **changed** from established 2 cc? 90% confidence level.
- $\chi^2(cal.) = \frac{(n-1)s^2}{\sigma^2} = \frac{(50)\ x\ 2.35^2}{2^2} = 69.03$
- $\chi^2(\text{critical}) = 34.76$ and 67.50

One Variance Test

One Variance Test - Minitab



Test and CI for One Variance

Method

σ: standard deviation of Sample
The Bonett method cannot be calculated for summarized data.
The chi-square method is valid only for the normal distribution.

Descriptive Statistics

N	StDev	Variance	90% Lower Bound for σ using Chi-Square
51	2.35	5.52	2.09

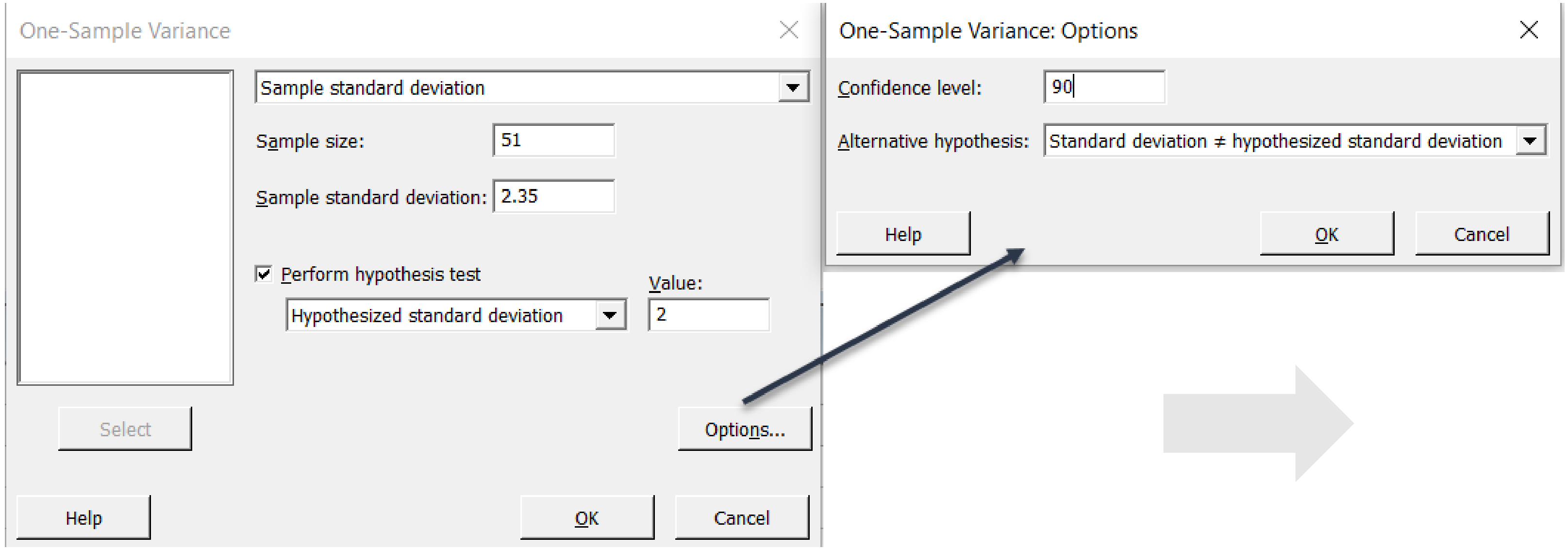
Test

Null hypothesis $H_0: \sigma = 2$
Alternative hypothesis $H_1: \sigma > 2$

Method	Test Statistic	DF	P-Value
Chi-Square	69.03	50	0.038

One Variance Test

One Variance Test - Minitab



Test and CI for One Variance

Method

σ: standard deviation of Sample
The Bonett method cannot be calculated for summarized data.
The chi-square method is valid only for the normal distribution.

Descriptive Statistics

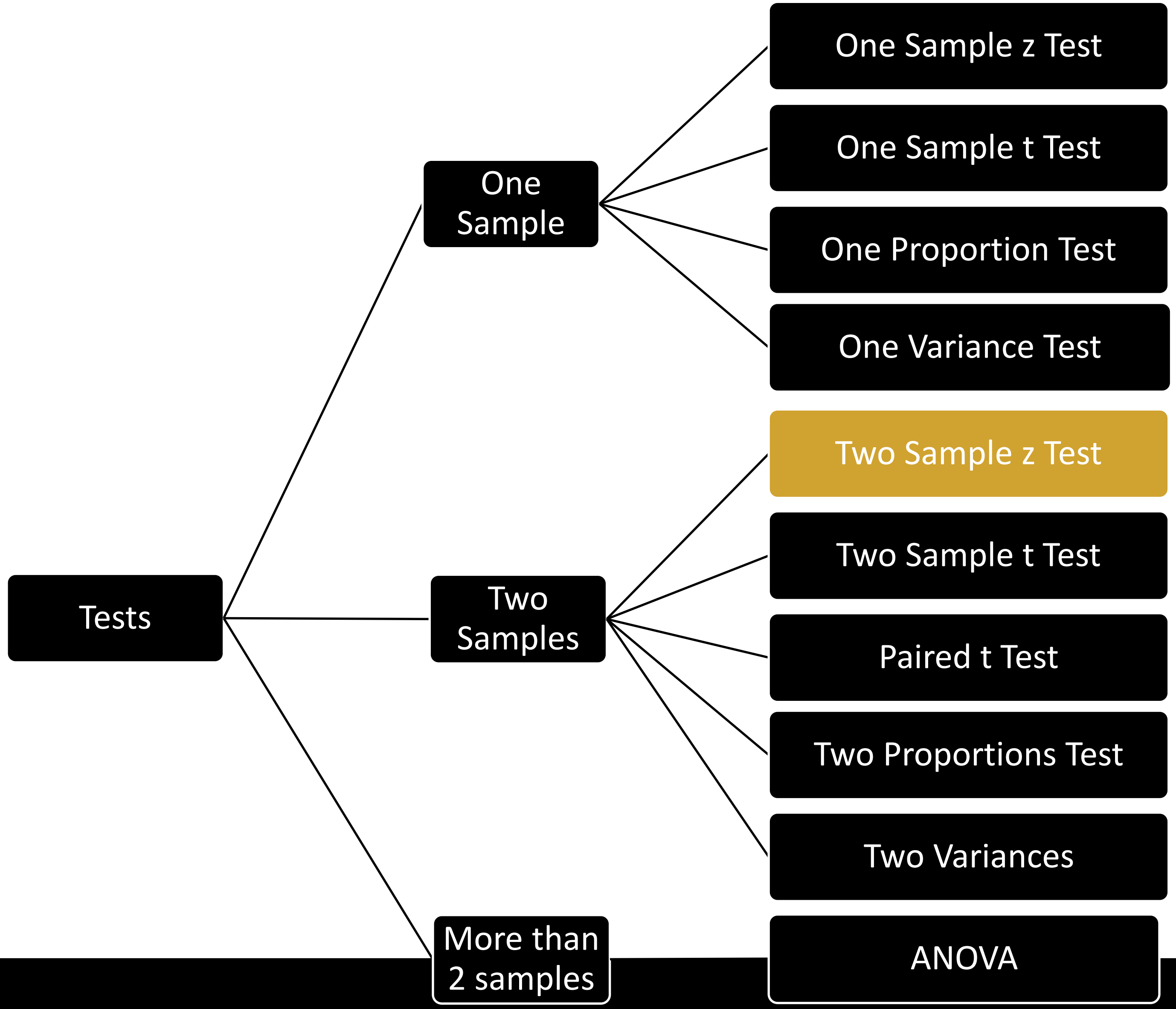
N	StDev	Variance	90% CI for σ using Chi-Square
51	2.35	5.52	(2.02, 2.82)

Test

Null hypothesis	H_0: σ = 2
Alternative hypothesis	H_1: σ ≠ 2

Method	Test Statistic	DF	P-Value
Chi-Square	69.03	50	0.077

One Variance Test



Two Sample Z Test

Conditions for z Test

- Random samples
- Each observation should be independent of other
  - Sampling with replacement
  - If sampling without replacement, the sample size should not be more than 10% of the population
- Sampling distribution approximates Normal Distribution
  - Population is Normally distributed and the population standard deviation is known *** OR ***
  - Sample size ≥ 30

Z Test

One Sample

H_0: μ = 150cc
H_a: $\mu \neq$ 150cc

$$z_{cal} = \frac{(\bar{x} - \mu)}{\sigma/\sqrt{n}}$$

Two Sample

- **Null hypothesis**: H_0: $\mu_1 = \mu_2$
  - or H_0: $\mu_1 - \mu_2 = 0$
- **Alternative hypothesis**: H_a : $\mu_1 \neq \mu_2$

$$z_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Two Sample Z Test

Calculated Test Statistic

H_0: $\mu_1 = \mu_2$

H_a: $\mu_1 \neq \mu_2$

$$z_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$z_{cal} = \frac{(151.2 - 151.9)}{\sqrt{\frac{2.1^2}{100} + \frac{2.2^2}{100}}}$$

- Example: From two machines 100 samples each were drawn.

  Machine 1: Mean = 151.2 / sd = 2.1

  Machine 2: Mean = 151.9 / sd = 2.2

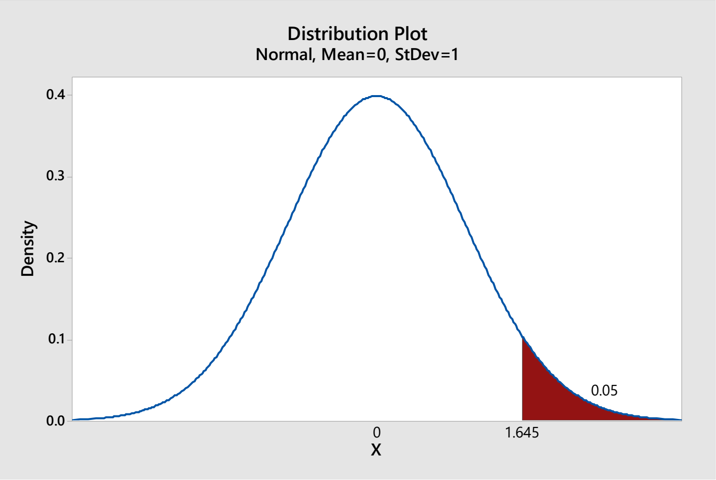
  Is there difference in these two machines. Check at 95% confidence level.

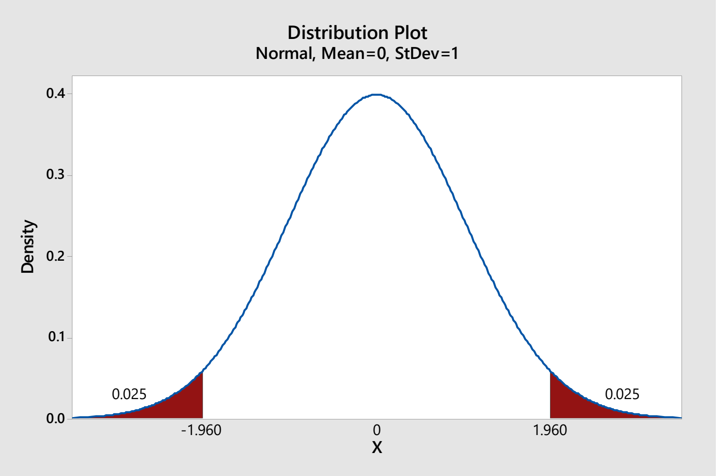
- Zcal = - 0.7 / 0.304 = -2.30
- $z_{critical}$ = ? (for alpha = 0.05, two tail test)

Two Sample Z Test

Critical Test Statistic

z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000





- α = 0.05 One Tail
- Z Critical = 1.645

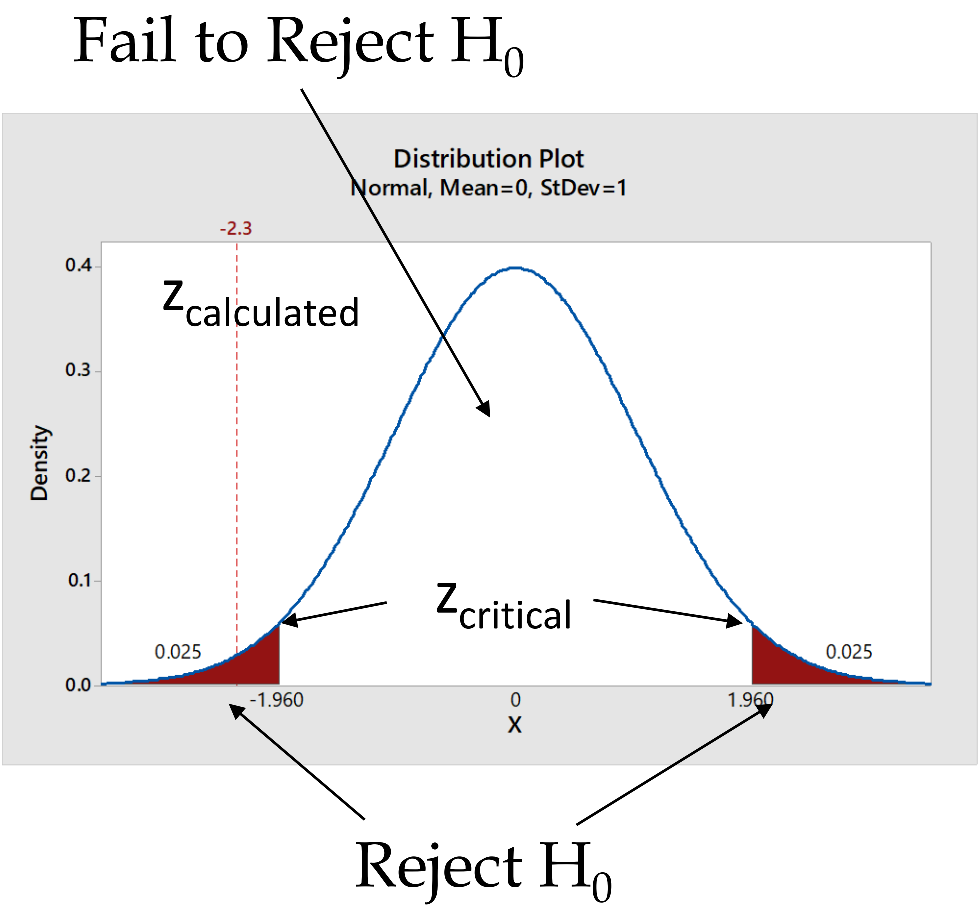
- α = 0.10 One Tail
- Z Critical = 1.282

- **α = 0.05 Two Tails**
- **Z Critical = 1.96**

- α = 0.10 Two Tail
- Z Critical = 1.645

Two Sample Z Test

Interpret the Results



- Example: From two machines 100 samples each were drawn.

  Machine 1: Mean = 151.2 / sd = 2.1

  Machine 2: Mean = 151.9 / sd = 2.2

  Is there difference in these two machines. Check at 95% confidence level.
- Zcal = - 0.7 / 0.304 = -2.30
- $z_{critical}$ = 1.96
- Conclusion: Reject Null Hypothesis
  - H_0: $\mu_1 = \mu_2$

Two Sample Z Test

Two Sample z Test - Minitab

QG

1. Download Macro:

   https://support.minitab.com/en-us/minitab/18/TwoZTest.mac

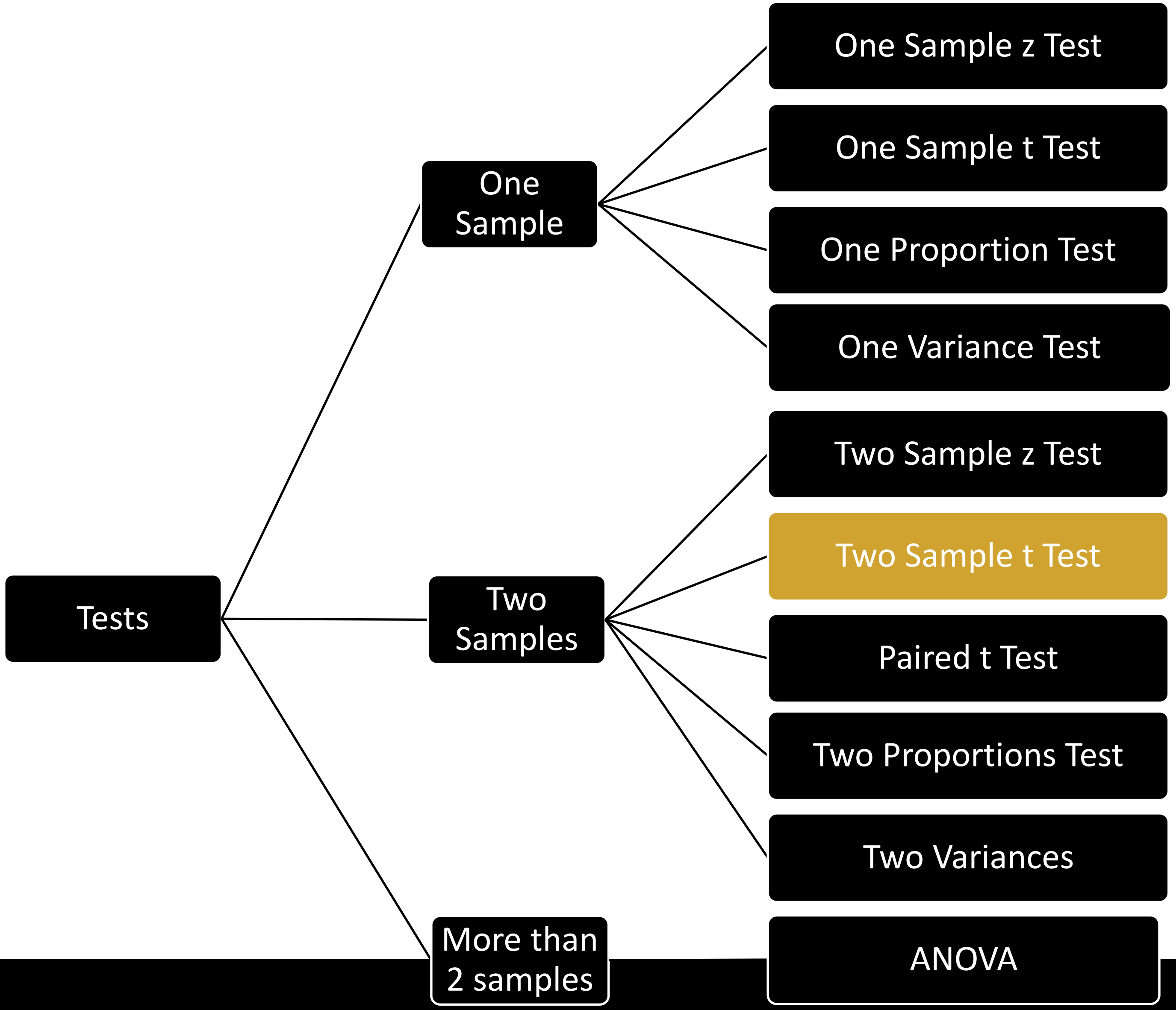
2. Provide the path of Macro:

   Choose Tools > Options > General. Under Macro location browse to the location where you save macro files.

3. Run command:

   Edit > Command Line Editor and type:
   %TWOZTEST C1 C2 sd1 sd2

Two Sample Z Test



Two Sample t Test

Conditions for t Test

- ❖ Random samples
- ❖ Each observation should be independent of other
  - ❖ Sampling with replacement
  - ❖ If sampling without replacement, the sample size should not be more than 10% of the population
- ❖ Sampling distribution approximates Normal Distribution
  - ❖ Population is Normally distributed and the standard deviation is **<u>unknown</u>** *** AND ***
  - ❖ Sample size < 30

Two Sample t Test

Types of t Tests

- If two set of data are independent or dependent.
  - If the values in one sample reveal no information about those of the other sample, then the samples are independent.
    - Example: Volume produced by two machines

  Two sample t test

  - If the values in one sample affect the values in the other sample, then the samples are dependent.
    - Example: Blood pressure before and after a specific medicine

  Paired t test

Two Sample t Test

One Sample t Test

$$t_{cal} = \frac{(\bar{x} - \mu)}{s/\sqrt{n}}$$

Two Sample t Tests

- Is variance for two samples equal?

Yes

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{s_p\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$df = n_1 + n_2 - 2$$

No

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

$$df = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^2}{\frac{(s_1^2/n_1)^2}{(n_1 - 1)} + \frac{(s_2^2/n_2)^2}{(n_2 - 1)}}$$

Two Sample t Test

Conditions for t Test

$H_0: \mu_A = \mu_B$
$H_a: \mu_A \neq \mu_B$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{s_p\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$df = n_1 + n_2 - 2$$

- Example: Samples from two machines A and B have the following volumes in bottles. **(assume equal variance)**
  - **Machine A:** 150, 152, 154, 152, 151
  - **Machine B:** 156, 155, 158, 155, 154

  Is the mean different? Calculate with 95% confidence.

  - $n_1 = 5$, $n_2 = 5$, $s_1 = 1.48$, $s_2 = 1.52$
  - $\bar{x}_1 = 151.8$, $\bar{x}_2 = 155.6$

Two Sample t Test

Conditions for t Test

$$H_0: \mu_A = \mu_B$$
$$H_a: \mu_A \neq \mu_B$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s_p = 1.50$$

- Example: Samples from two machines A and B have the following volumes in bottles. **(assume equal variance)**
  - **Machine A:** 150, 152, 154, 152, 151
  - **Machine B:** 156, 155, 158, 155, 154

  Is the mean **different**? Calculate with 95% confidence.

  - $n_1 = 5$, $n_2 = 5$, $s_1 = 1.48$, $s_2 = 1.52$
  - $\bar{x}_1 = 151.8$, $\bar{x}_2 = 155.6$

Two Sample t Test

Conditions for t Test

$H_0: \mu_A = \mu_B$
$H_a: \mu_A \neq \mu_B$

$s_p = 1.50$

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{s_p\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$t_{cal} = - 4.01$

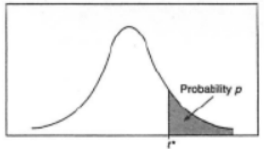
- Example: Samples from two machines A and B have the following volumes in bottles. **(assume equal variance)**
  - **Machine A:** 150, 152, 154, 152, 151
  - **Machine B:** 156, 155, 158, 155, 154

  Is the mean different? Calculate with 95% confidence.

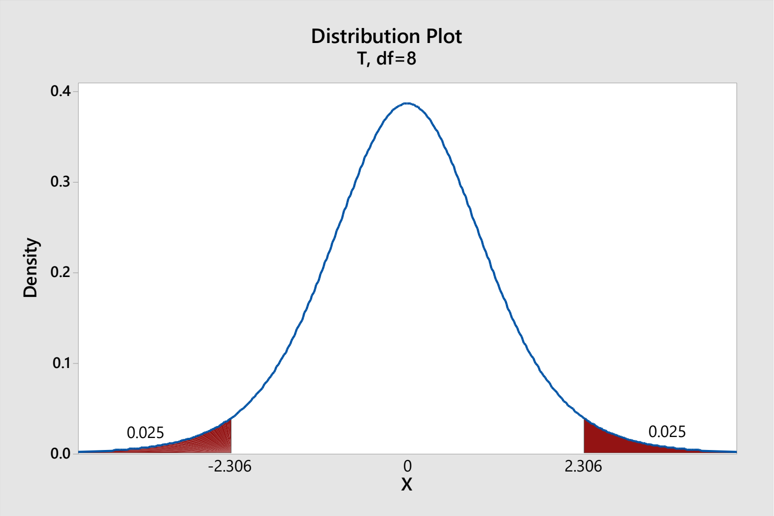
  - $n_1 = 5$, $n_2 = 5$, $s_1 = 1.48$, $s_2 = 1.52$
  - $\bar{x}_1 = 151.8$, $\bar{x}_2 = 155.6$

Two Sample t Test

Critical Test Statistic



df	TAIL PROBABILITY P											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646



$$df = n_1 + n_2 - 2$$

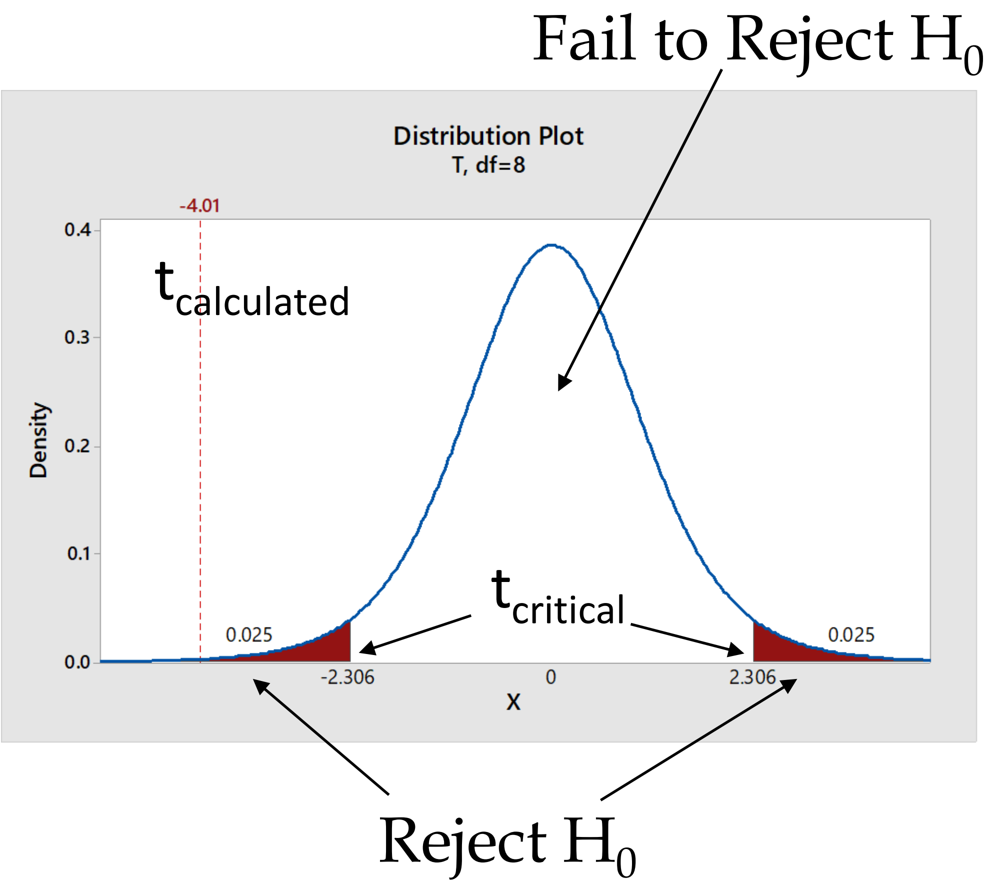
$$df = 8$$

- α = 0.05 Two Tails
- Df = 8
- t Critical = 2.306

Two Sample t Test

$H_0: \mu_A = \mu_B$
$H_a: \mu_A \neq \mu_B$

Interpret the Results



- Example: Samples from two machines A and B have the following volumes in bottles. **(assume equal variance)**
  - **Machine A:** 150, 152, 154, 152, 151
  - **Machine B:** 156, 155, 158, 155, 154

  Is the mean **different**? Calculate with 95% confidence.
- $t_{calculated} = -4.01$
- $t_{critical} = 2.306$

Two Sample t Test

One Sample t Test

$$t_{cal} = \frac{(\bar{x} - \mu)}{s/\sqrt{n}}$$

Two Sample t Tests

- Is variance for two samples equal?

Yes

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{s_p\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$df = n_1 + n_2 - 2$$

No

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

$$df = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^2}{\frac{(s_1^2/n_1)^2}{(n_1 - 1)} + \frac{(s_2^2/n_2)^2}{(n_2 - 1)}}$$

Two Sample t Test

Conditions for t Test

QG

H_0: $\mu_A = \mu_C$
H_a: $\mu_A \neq \mu_C$

$$\cancel{t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{s_p\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}}$$

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

- Example: Samples from two machines A and B have the following volumes in bottles. (**unequal variance**)
  - **Machine A:** 150, 152, 154, 152, 151
  - **Machine C:** 144, 162, 177, 150, 140

  Is the mean different? Calculate with 95% confidence.

  - n_1 = 5, n_2 = 5, s_1 = 1.48, s_2 = 15.0
  - $\bar{x}_1$ = 151.8, $\bar{x}_2$ = 154.6

Two Sample t Test

Conditions for t Test

H_0: $\mu_A = \mu_C$
H_a: $\mu_A \neq \mu_C$

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

t_{cal} = - 0.41

- Example: Samples from two machines A and B have the following volumes in bottles. (**unequal variance**)
  - **Machine A:** 150, 152, 154, 152, 151
  - **Machine C:** 144, 162, 177, 150, 140

  Is the mean different? Calculate with 95% confidence.

  - n_1 = 5, n_2 = 5, s_1 = 1.48, s_2 = 15.0
  - $\bar{x}_1$ = 151.8, $\bar{x}_2$ = 154.6

Two Sample t Test

Conditions for t Test

$$H_0: \mu_A = \mu_C$$
$$H_a: \mu_A \neq \mu_C$$

$$\cancel{df = n_1 + n_2 - 2}$$

$$df = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^2}{\frac{(s_1^2/n_1)^2}{(n_1 - 1)} + \frac{(s_2^2/n_2)^2}{(n_2 - 1)}}$$

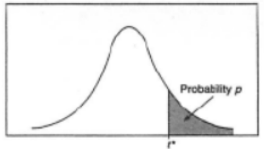
$df = 4.078$ or 4 (round down to nearest integer)

- Example: Samples from two machines A and B have the following volumes in bottles. (**unequal variance**)
  - **Machine A:** 150, 152, 154, 152, 151
  - **Machine C:** 144, 162, 177, 150, 140

  Is the mean different? Calculate with 95% confidence.

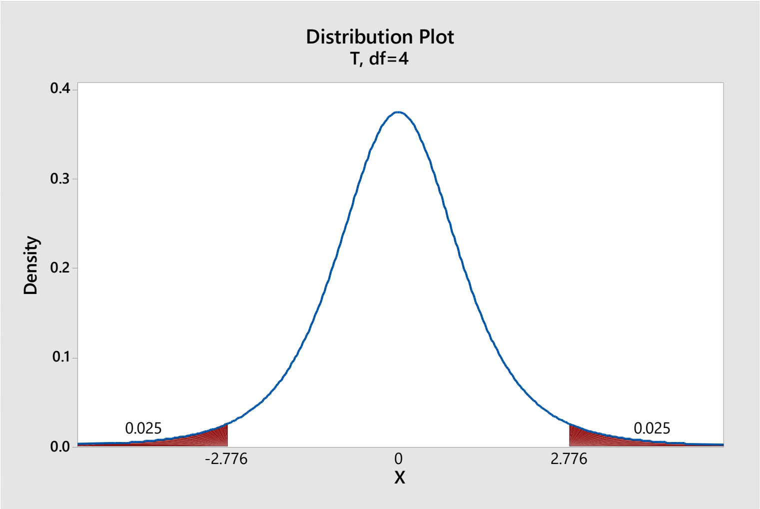
  - $n_1 = 5$, $n_2 = 5$, $s_1 = 1.48$, $s_2 = 15.0$
  - $\bar{x}_1 = 151.8$, $\bar{x}_2 = 154.6$

Two Sample t Test



Critical Test Statistic

df	TAIL PROBABILITY P											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646



$$df = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^2}{\frac{(s_1^2/n_1)^2}{(n_1-1)} + \frac{(s_2^2/n_2)^2}{(n_2-1)}}$$

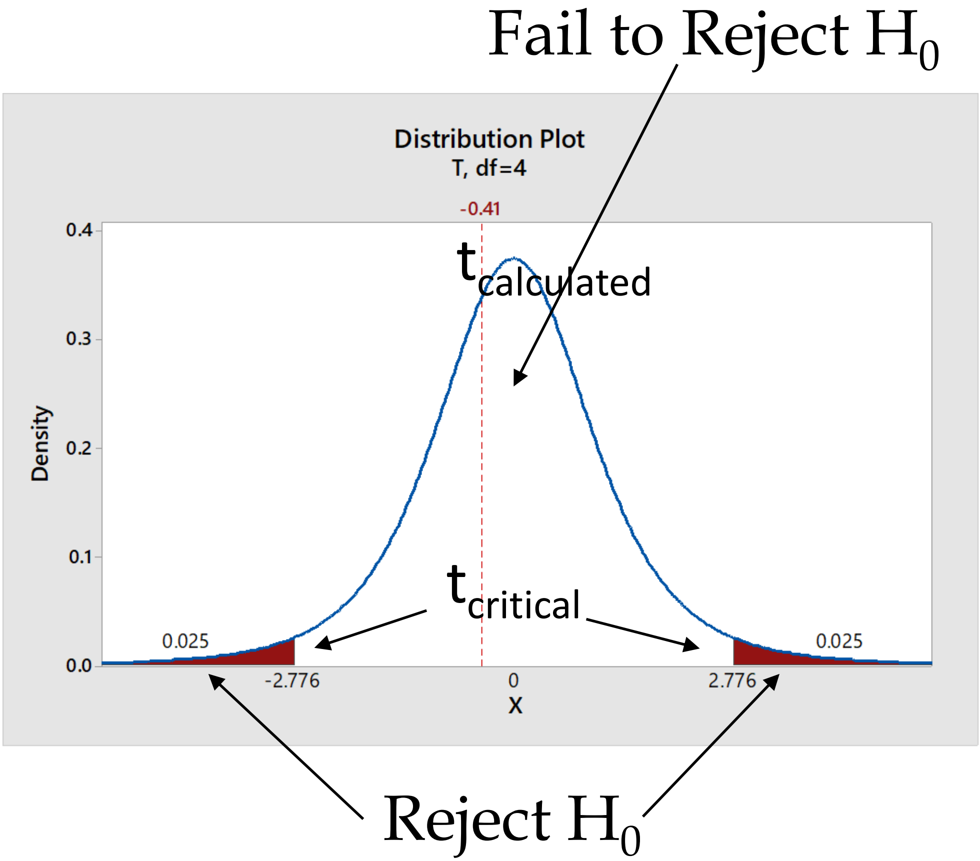
$$df = 4$$

- α = 0.05 Two Tails
- Df = 4
- t Critical = 2.776

Two Sample t Test

Interpret the Results

$H_0: \mu_A = \mu_C$
$H_a: \mu_A \neq \mu_C$



- Example: Samples from two machines A and B have the following volumes in bottles. (**unequal variance**)
  - **Machine A:** 150, 152, 154, 152, 151
  - **Machine C:** 144, 162, 177, 150, 140

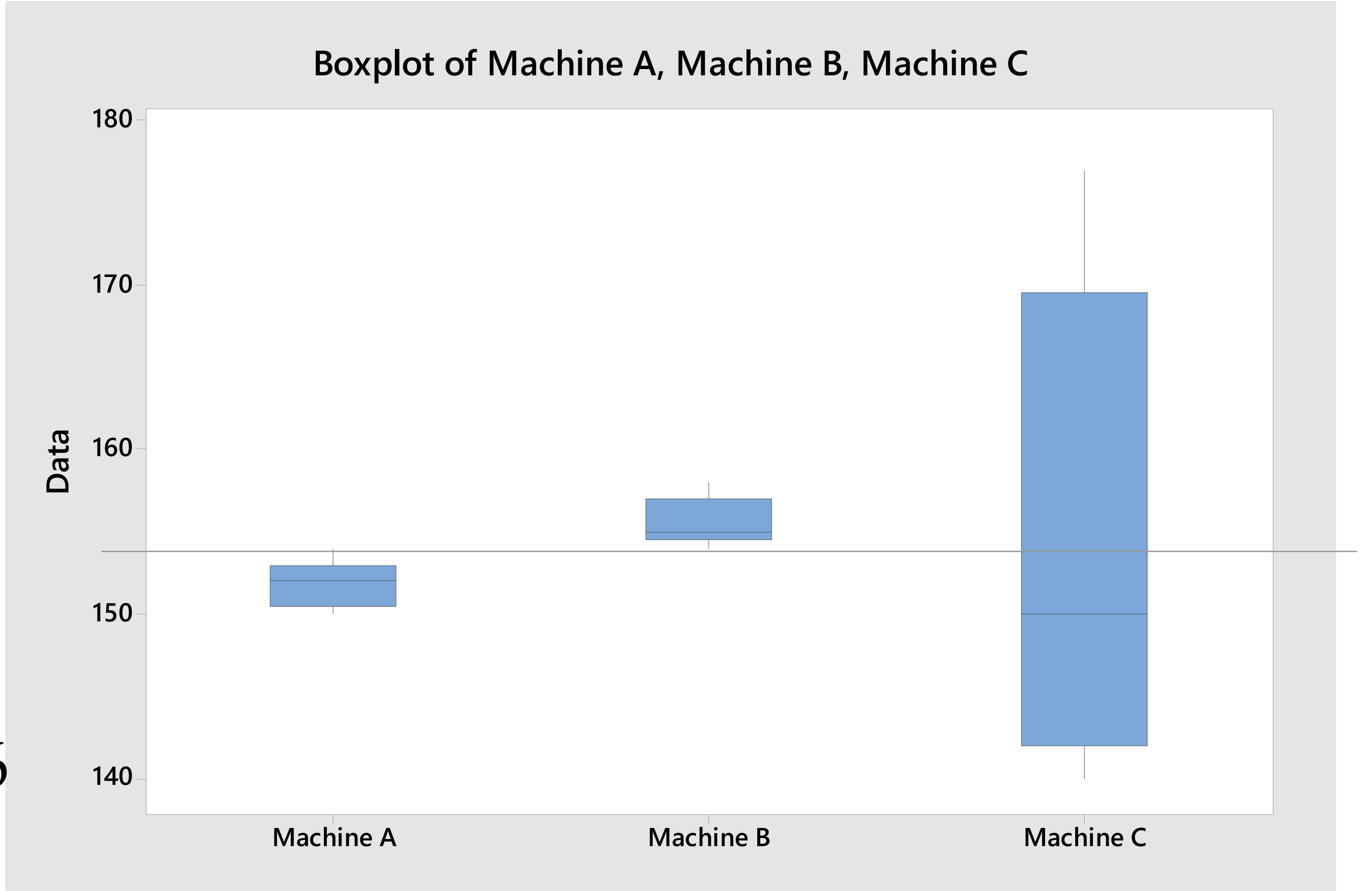
  Is the mean different? Calculate with 95% confidence.
- $t_{calculated} = -0.41$
- $t_{critical} = 2.776$

Two Sample t Test

Interpret the Results

Machine A	Machine B	Machine C
150	156	144
152	155	162
154	158	177
152	155	150
151	154	140

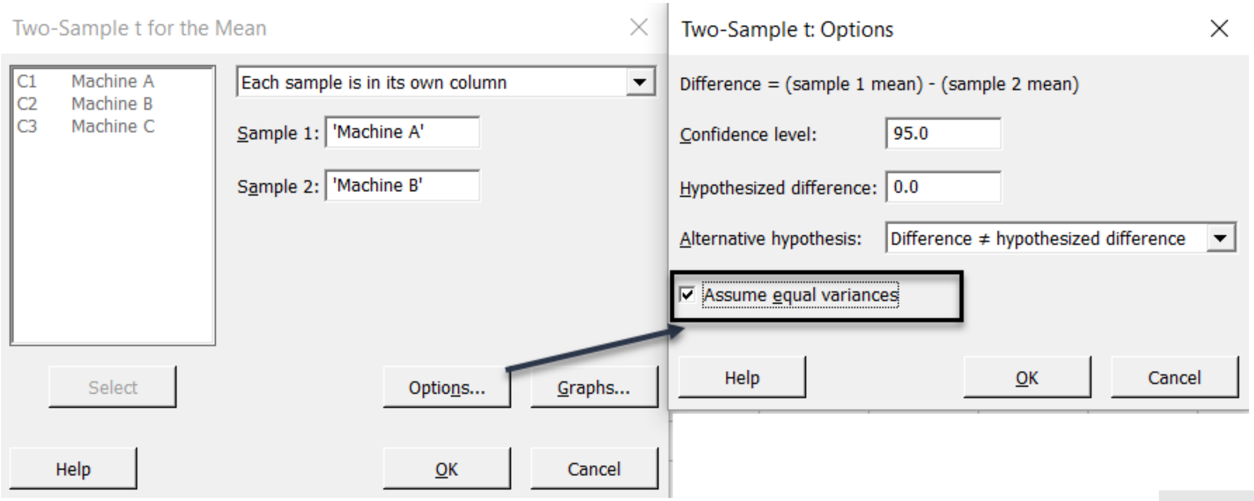
$\bar{x}_{\mathbf{A}} = 151.8$, $\bar{x}_{\mathrm{B}} = 155.6$ $\bar{x}_{\mathrm{C}} = 154.6$



Two Sample t Test

Two Sample t Test - Minitab

Equal Variance



Two-Sample T-Test and CI: Machine A, Machine B

Method

μ_1: mean of Machine A
μ_2: mean of Machine B
Difference: $\mu_1 - \mu_2$

Equal variances are assumed for this analysis.

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
Machine A	5	151.80	1.48	0.66
Machine B	5	155.60	1.52	0.68

Estimation for Difference

Difference	Pooled StDev	95% CI for Difference
-3.800	1.500	(-5.988, -1.612)

Test

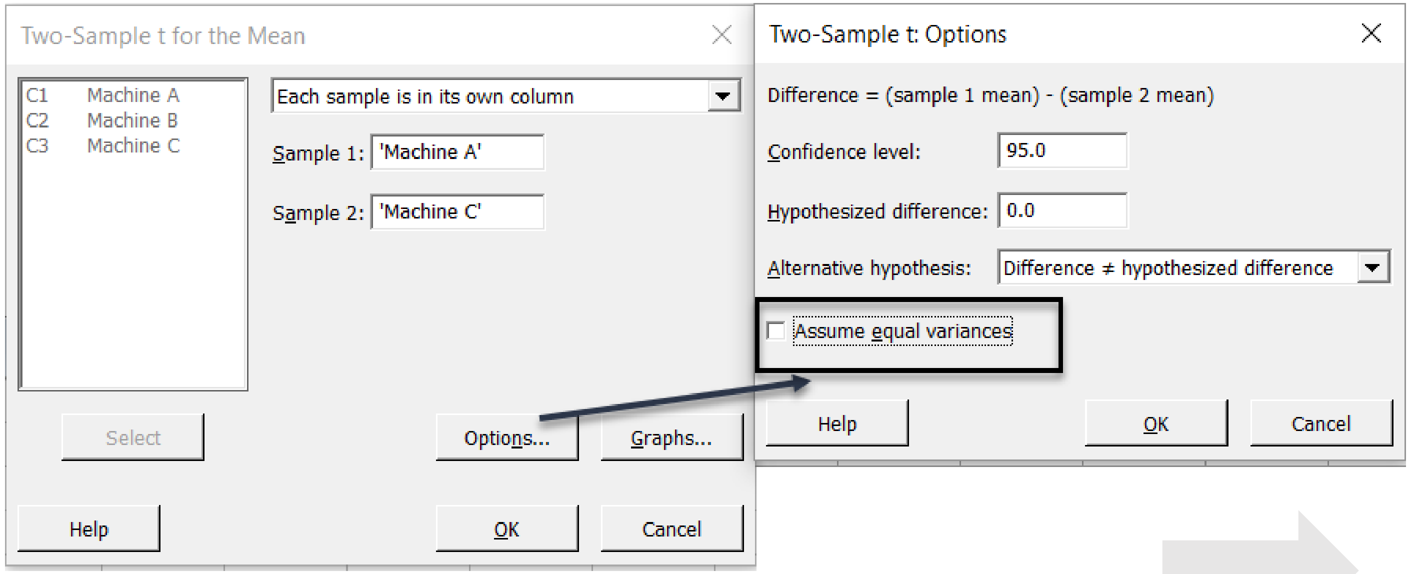
Null hypothesis $H_0: \mu_1 - \mu_2 = 0$
Alternative hypothesis $H_1: \mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
-4.01	8	0.004

Two Sample t Test

Two Sample t Test - Minitab

Unequal Variance



Two-Sample T-Test and CI: Machine A, Machine C

Method

μ_1: mean of Machine A
μ_2: mean of Machine C
Difference: $\mu_1 - \mu_2$

Equal variances are not assumed for this analysis.

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
Machine A	5	151.80	1.48	0.66
Machine C	5	154.6	15.0	6.7

Estimation for Difference

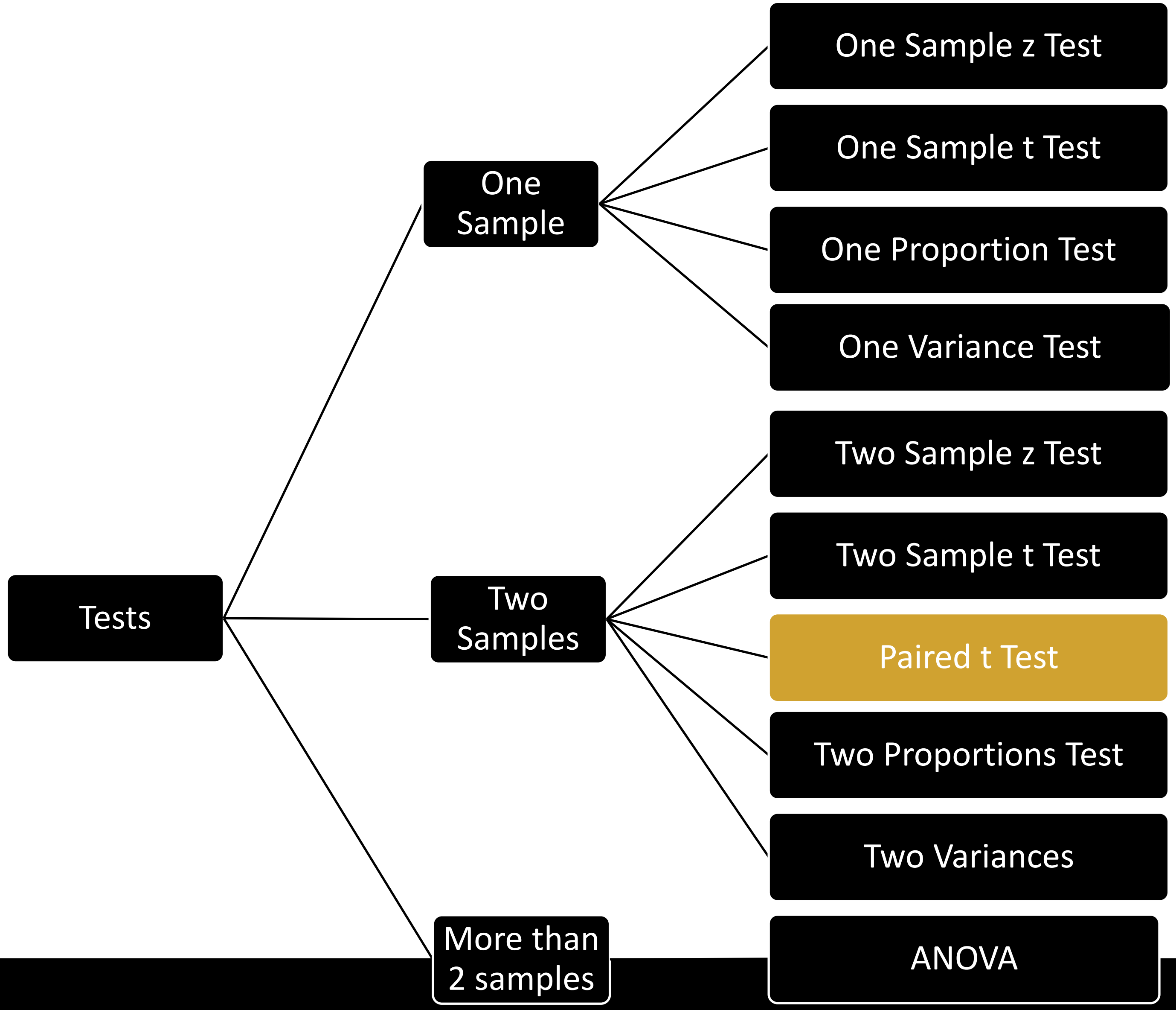
Difference	95% CI for Difference
-2.80	(-21.55, 15.95)

Test

Null hypothesis	$H_0: \mu_1 - \mu_2 = 0$
Alternative hypothesis	$H_1: \mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
-0.41	4	0.700

Two Sample t Test



Paired t Test

Types of t Tests

- If two set of data are independent or dependent.
  - If the values in one sample reveal no information about those of the other sample, then the samples are independent.
    - Example: Blood pressure of male/female

Two sample t test

  - If the values in one sample affect the values in the other sample, then the samples are dependent.
    - Example: Blood pressure before and after a specific medicine

Paired t test

Paired t Test

One Sample t Test

$$t_{cal} = \frac{(\bar{x} - \mu)}{s/\sqrt{n}}$$

Paired t Tests

- ❖ Find the difference between two set of readings as d1, d2 …. dn.
- ❖ Find the mean and standard deviation of these differences.

$$t = \frac{\bar{d}}{s/\sqrt{n}}$$

Paired t Test

Paired t Tests

H_0: $\mu_{before} = \mu_{after}$
H_a: $\mu_{before} \neq \mu_{after}$

- Example: Before and after medicine BP was measured. Is there a difference at 95% confidence level?

Patient	Before	After
1	120	122
2	122	120
3	143	141
4	100	109
5	109	109

Paired t Test

Paired t Tests

$H_0: \mu_{before} = \mu_{after}$
$H_a: \mu_{before} \neq \mu_{after}$

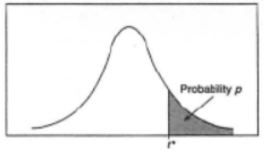
$$t = \frac{\bar{d}}{s/\sqrt{n}}$$

Patient	Before	After	difference
1	120	122	-2
2	122	120	2
3	143	141	2
4	100	109	-9
5	109	109	0

- Example: Before and after medicine BP was measured. Is there a **difference** at 95% confidence level?

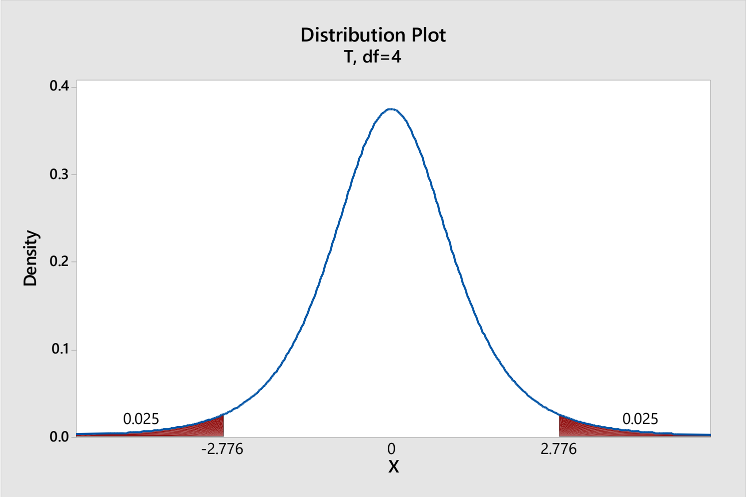
- $\bar{d}$ = -1.4 , s = 4.56 , n = 5
- $t_{cal.}$ = 1.4/2.04 = - 0.69

Paired t Test



Critical Test Statistic

df	TAIL PROBABILITY P											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646



$$df = n - 1$$

$$df = 4$$

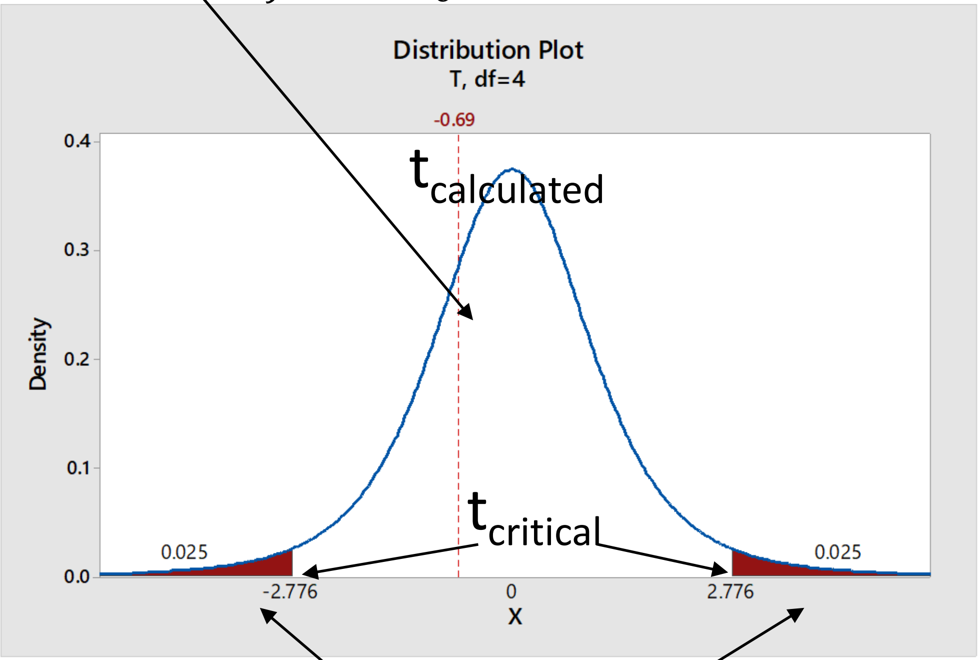
- α = 0.05 Two Tails
- df = 4
- t Critical = 2.776

Paired t Test

H_0: $\mu_{before} = \mu_{after}$
H_a: $\mu_{before} \neq \mu_{after}$

Interpret the Results

Fail to Reject H_0



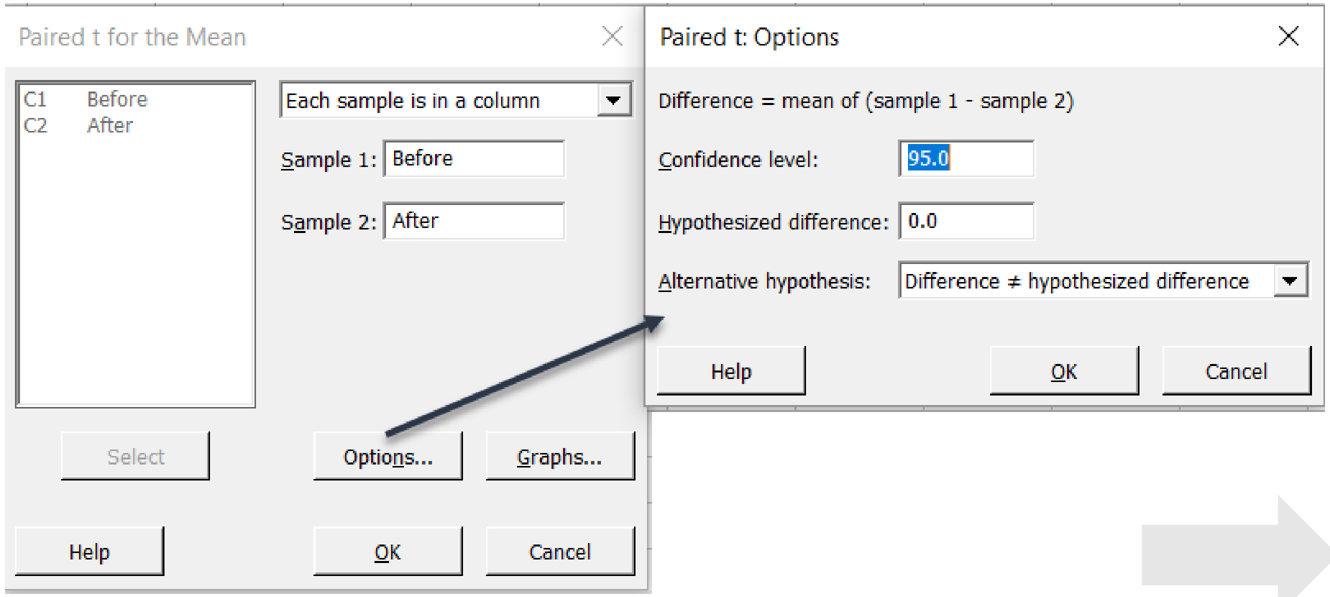
Reject H_0

- Example: Before and after medicine BP was measured. Is there a **difference** at 95% confidence level?

- $t_{calculated}$ = - 0.69
- $t_{critical}$ = 2.776

Paired t Test

Paired t Test - Minitab



Paired T-Test and CI: Before, After

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
Before	5	118.80	16.18	7.23
After	5	120.20	13.10	5.86

Estimation for Paired Difference

Mean	StDev	SE Mean	95% CI for μ_difference
-1.40	4.56	2.04	(-7.06, 4.26)

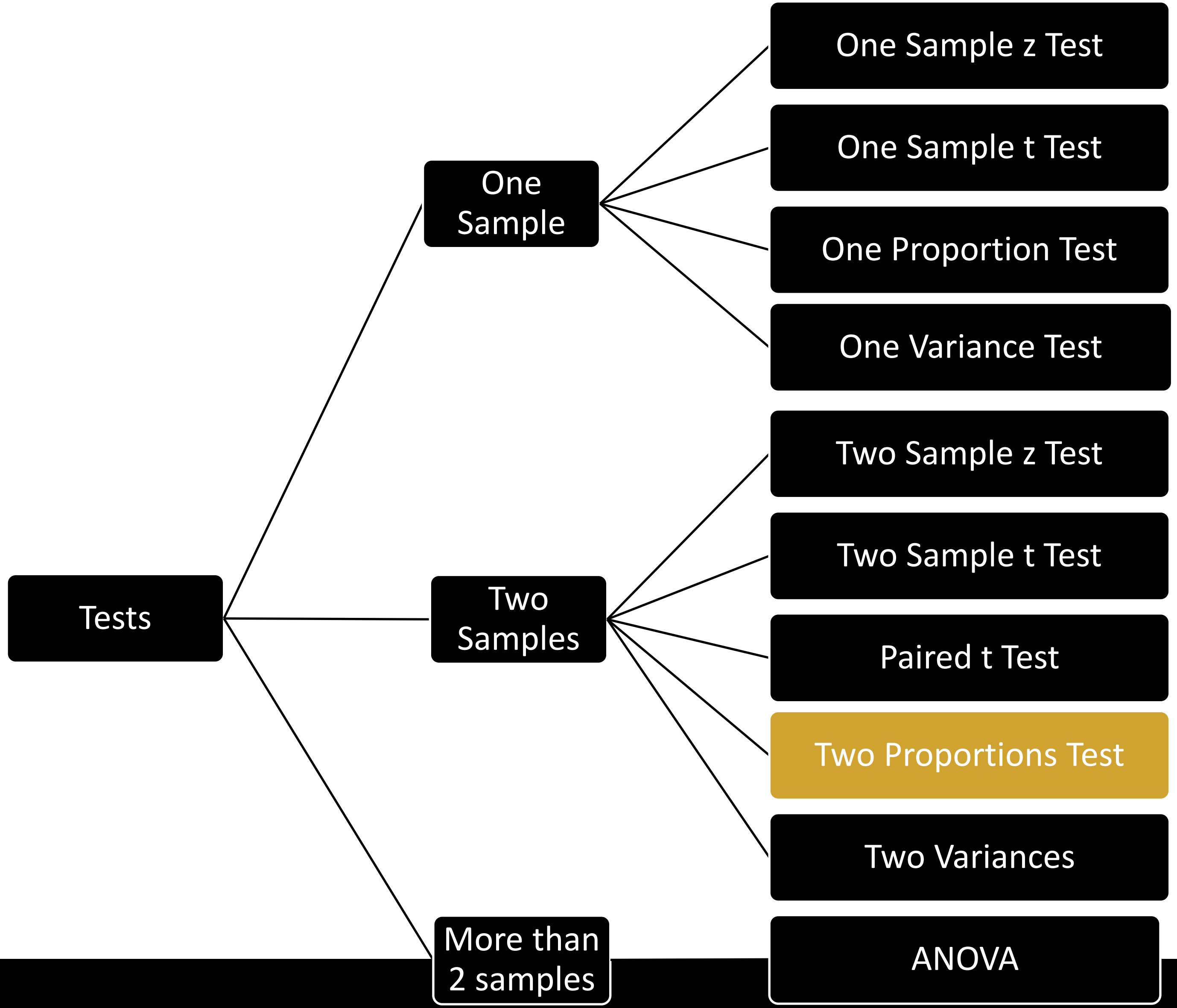
μ_difference: mean of (Before - After)

Test

Null hypothesis	H_0: μ_difference = 0
Alternative hypothesis	H_1: μ_difference ≠ 0

T-Value	P-Value
-0.69	0.530

Paired t Test



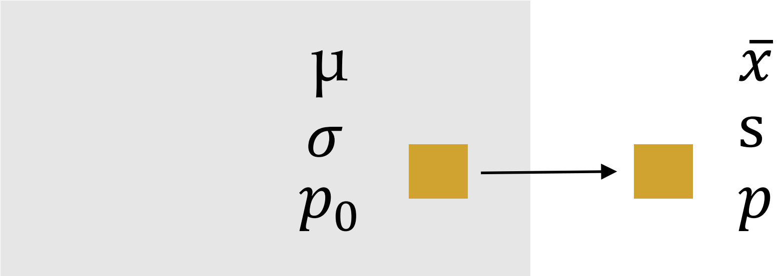
Two Proportions Test

Conditions for Proportions Test

- Random samples
- Each observation should be independent of other
  - Sampling with replacement
  - If sampling without replacement, the sample size should not be more than 10% of the population
- The data contains only two categories, such as pass/fail or yes/no
- For Normal approximation:
  - both $np \geq 10$ and $n(1-p) \geq 10$ (data should have at least 10 "successes" and at least 10 "failures") **for each sample** (in some books it is 5)

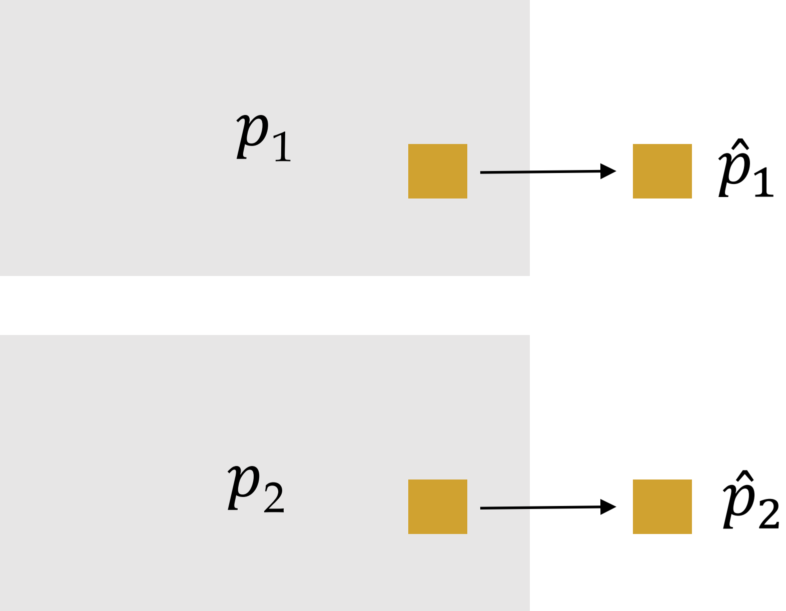
Proportions – Sample vs Population

How do we represent sample and population proportions?



One Proportion Test

$$z = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$



Two Proportions Test

One Proportion Test

$$z = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Two Proportions Tests

Test for **no** difference between proportions

Yes

Pooled

H_0: $p_1 - p_2 = 0$
H_a: $p_1 - p_2 \neq 0$

$$\bar{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{(n_1 + n_2)}$$

$$z_{cal} = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

No

Un-pooled

H_0: $p_1 - p_2 = d$
H_a: $p_1 - p_2 \neq d$

$$z_{cal} = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

Two Proportions Test

Two Proportions Tests

H_0: $p_1 - p_2 = 0$
H_a: $p_1 - p_2 \neq 0$

Test if Normality can be assumed?

$$\bar{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{(n_1 + n_2)}$$

$$z_{cal} = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

- Example: From vendor A we test 200 pieces and find 30 defectives. From vendor B we test 100 pieces and we find 10 defectives. Is there a significant difference in the quality of these two vendors? Use 95% confidence level.
- $z_{calculated}$ = ?
- $z_{critical}$ = ?

Two Proportions Test

Two Proportions Tests

H_0: $p_1 - p_2 = 0$
H_a: $p_1 - p_2 \neq 0$

Test if Normality can be assumed?

$$\bar{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{(n_1 + n_2)}$$

$$z_{cal} = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

- Example: From vendor A we test 200 pieces and find 30 defectives. From vendor B we test 100 pieces and we find 10 defectives. Is there a significant difference in the quality of these two vendors? Use 95% confidence level.
- $\hat{p}_1$ = 30/200 = 0.15
- $\hat{p}_2$ = 10/100 = 0.10
- n_1 = 200, n_2 = 100

Two Proportions Test

Two Proportions Tests

H_0: $p_1 - p_2 = 0$
H_a: $p_1 - p_2 \neq 0$

Test if Normality can be assumed?

$$n_1\hat{p}_1 \geq 10$$

$$n_1(1-\hat{p}_1) \geq 10$$

$$n_2\hat{p}_2 \geq 10$$

$$n_1(1-\hat{p}_2) \geq 10$$

- Example: From vendor A we test 200 pieces and find 30 defectives. From vendor B we test 100 pieces and we find 10 defectives. Is there a significant difference in the quality of these two vendors? Use 95% confidence level.
- $\hat{p}_1$= 30/200 = 0.15
- $\hat{p}_2$= 10/100 = 0.10
- n_1 = 200, n_2 = 100

Two Proportions Test

Two Proportions Tests

H_0: $p_1 - p_2 = 0$
H_a: $p_1 - p_2 \neq 0$

$$\bar{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{(n_1 + n_2)}$$

$$\bar{p} = 0.1333$$

- Example: From vendor A we test 200 pieces and find 30 defectives. From vendor B we test 100 pieces and we find 10 defectives. Is there a significant difference in the quality of these two vendors? Use 95% confidence level.
- $\hat{p}_1$= 30/200 = 0.15
- $\hat{p}_2$= 10/100 = 0.10
- n_1 = 200, n_2 = 100

Two Proportions Test

Two Proportions Tests

H_0: $p_1 - p_2 = 0$

H_a: $p_1 - p_2 \neq 0$

$$\bar{p} = 0.1333$$

$$z_{cal} = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$\mathrm{z_{cal}} = \frac{(0.15 - 0.10)}{\sqrt{0.133(1-0.133)\left(\frac{1}{200} + \frac{1}{100}\right)}}$$

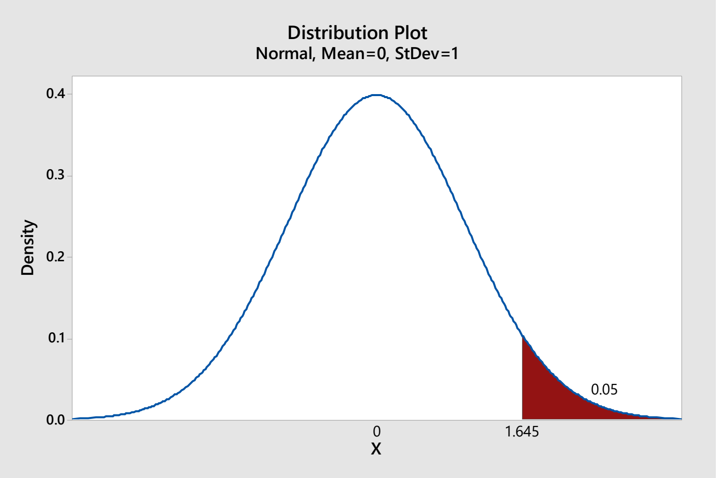
$$\mathrm{z_{cal}} = 1.20$$

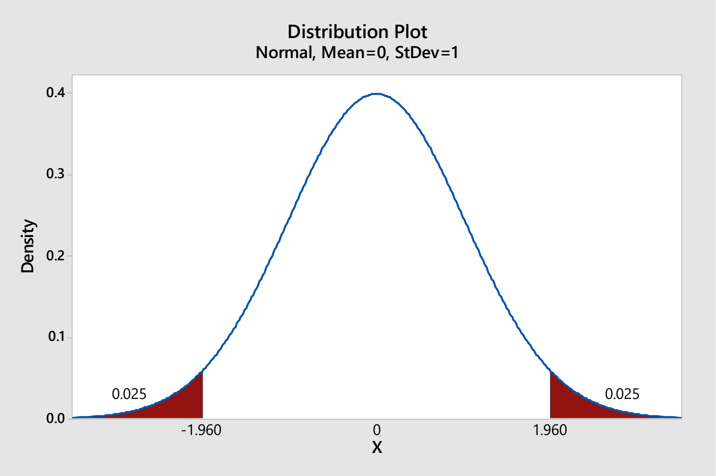
- Example: From vendor A we test 200 pieces and find 30 defectives. From vendor B we test 100 pieces and we find 10 defectives. Is there a significant difference in the quality of these two vendors? Use 95% confidence level.
- $\hat{p}_1$= 30/200 = 0.15
- $\hat{p}_2$= 10/100 = 0.10
- n_1 = 200, n_2 = 100

Two Proportions Test

Critical Test Statistic

z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000





- α = 0.05 One Tail
- Z Critical = 1.645

- α = 0.10 One Tail
- Z Critical = 1.282

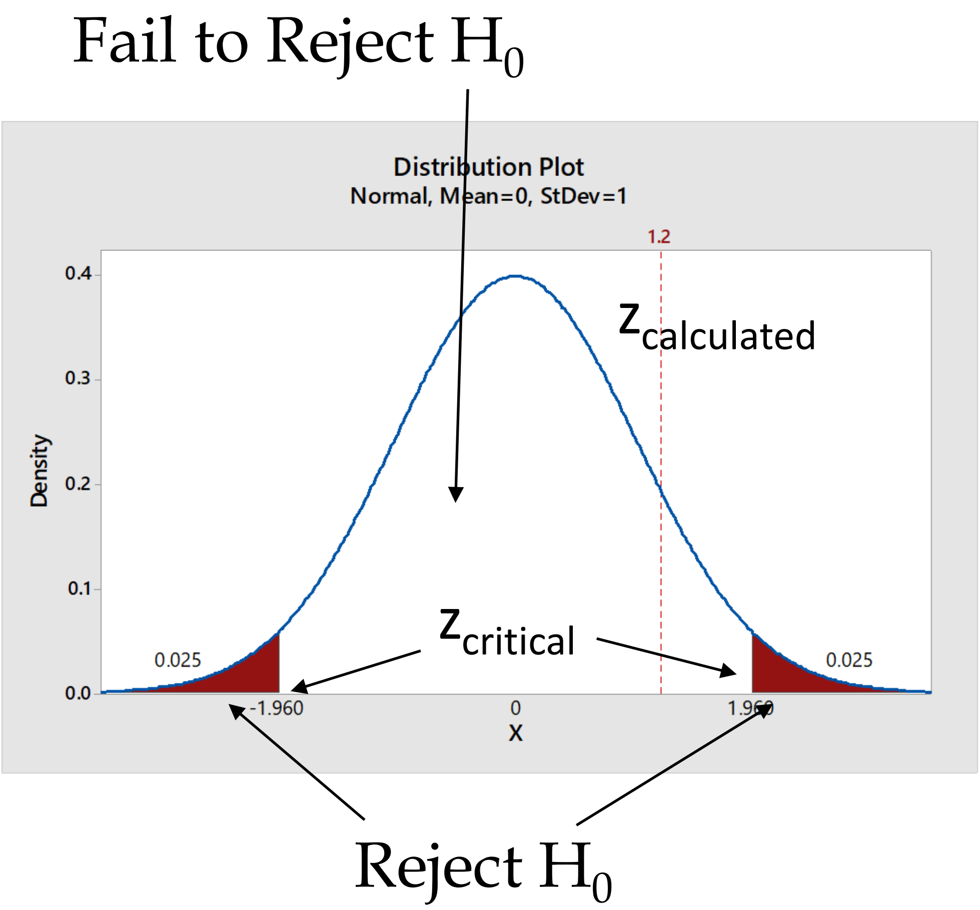
- α = 0.05 Two Tails
- Z Critical = 1.96

- α = 0.10 Two Tail
- Z Critical = 1.645

Two Proportions Test

H_0: $p = p_0$
H_a: $p \neq p_0$

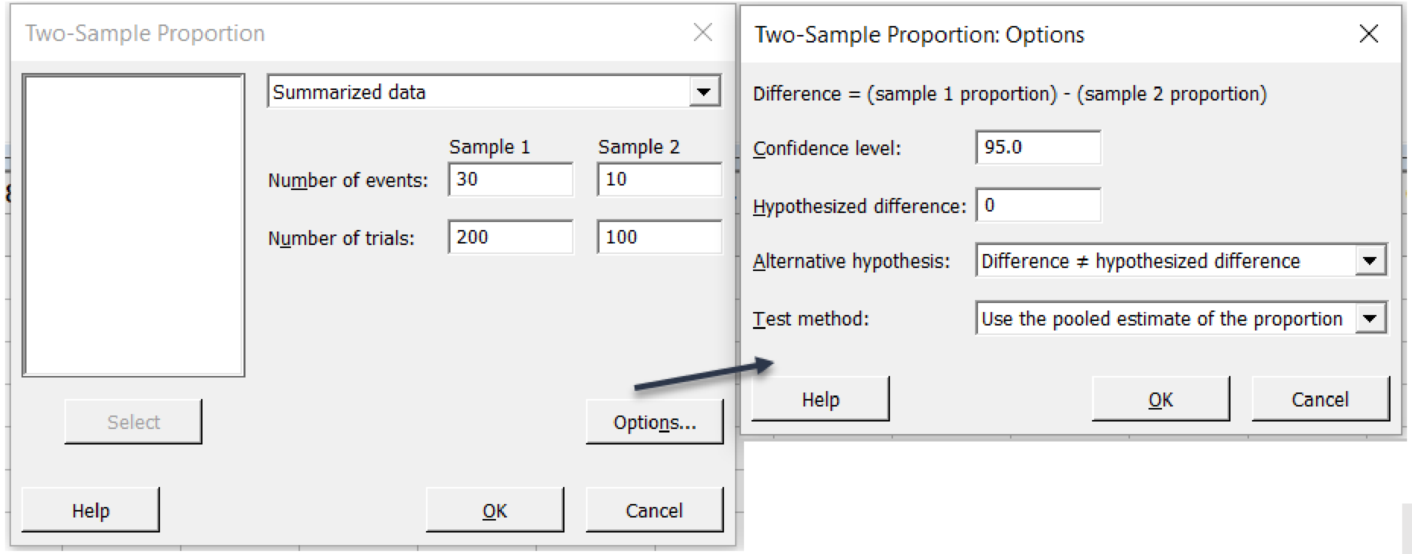
Interpret the Results



- Example: From vendor A we test 200 pieces and find 30 defectives. From vendor B we test 100 pieces and we find 10 defectives. Is there a significant difference in the quality of these two vendors? Use 95% confidence level.
- $z_{calculated} = 1.20$
- $z_{critical} = 1.96$

Two Proportions Test

Two Proportions Test - Minitab



Test and CI for Two Proportions

Method

p_1: proportion where Sample 1 = Event
p_2: proportion where Sample 2 = Event
Difference: $p_1 - p_2$

Descriptive Statistics

Sample	N	Event	Sample p
Sample 1	200	30	0.150000
Sample 2	100	10	0.100000

Estimation for Difference

Difference	95% CI for Difference
0.05	(-0.026852, 0.126852)

CI based on normal approximation

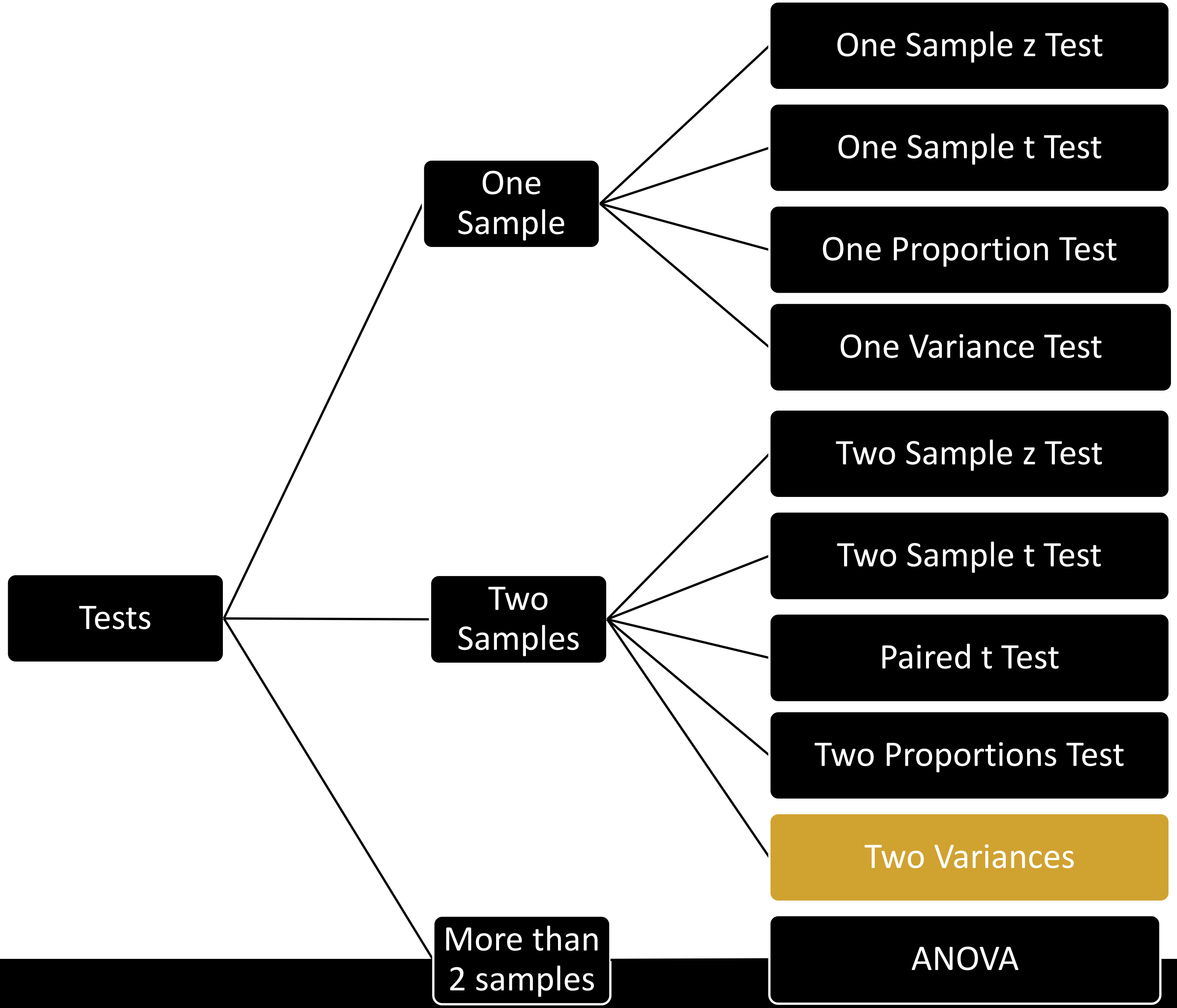
Test

Null hypothesis $H_0: p_1 - p_2 = 0$
Alternative hypothesis $H_1: p_1 - p_2 \neq 0$

Method	Z-Value	P-Value
Normal approximation	1.20	0.230
Fisher's exact		0.281

The pooled estimate of the proportion (0.133333) is used for the tests.

Two Proportions Test



Two Variances Test

Conditions for Variance Tests

- Random samples
- Each observation should be independent of other
  - Sampling with replacement
  - If sampling without replacement, the sample size should not be more than 10% of the population
- The data follows a Normal Distribution

Variance Tests

- Chi-square test
  - For testing the population variance against a specified value
  - testing goodness of fit of some probability distribution
  - testing for independence of two attributes (Contingency Tables)

- F-test
  - for testing equality of *two* variances from different populations
  - for testing equality of several means with technique of ANOVA.

Two Variances Test

Two Variances Test

$H_0: \sigma^2_1 = \sigma^2_2$
$Ha: \sigma^2_1 \neq \sigma^2_2$

$$F_{cal} = \frac{s_1^2}{s_2^2}$$

- Example: We took 8 samples from machine A and the **standard deviation** was 1.1. For machine B we took 5 samples and the **variance** was 11. Is there a difference in variance at 90% confidence level?
- $n1 = 5$, $s^2_1 = 11$, $df_1 = 4$ (numerator)
- $n2 = 8$, $s_2 = 1.1$, $s^2_2 = 1.21$, $df_2 = 7$ (denominator)
- $F_{calculated} = 11/1.21 = 9.09$ **(higher value at top)**

Two Variances Test

Critical Test Statistic

F - Distribution (α = 0.05 in the Right Tail)

Numerator Degrees of Freedom

df_2 \ df_1	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
4	7.7086	9.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	6.9988
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799

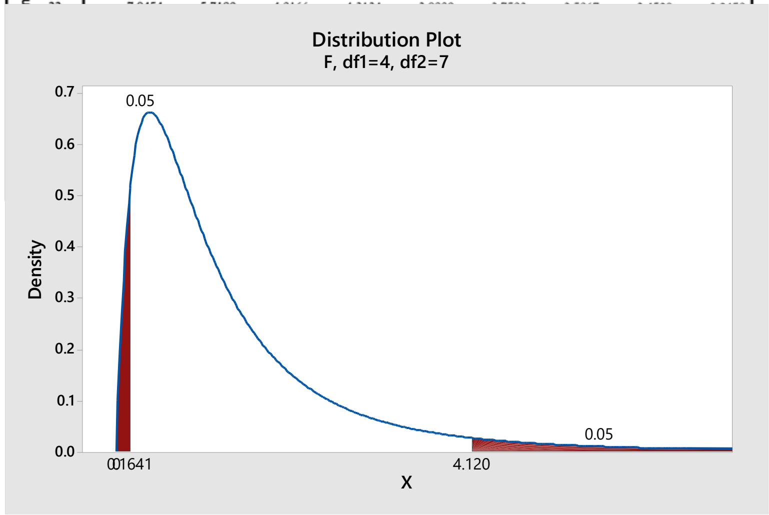
Denominator Degrees of Freedom

F - Distribution (α = 0.01 in the Right Tail)

Numerator Degrees of Freedom

df_2 \ df_1	1	2	3	4	5	6	7	8	9
1	4052.2	4999.5	5403.4	5624.6	5763.6	5859.0	5928.4	5981.1	6022.5
2	98.503	99.000	99.166	99.249	99.299	99.333	99.356	99.374	99.388
3	34.116	30.817	29.457	28.710	28.237	27.911	27.672	27.489	27.345
4	21.198	18.000	16.694	15.977	15.522	15.207	14.976	14.799	14.659
5	16.258	13.274	12.060	11.392	10.967	10.672	10.456	10.289	10.158
6	13.745	10.925	9.7795	9.1483	8.7459	8.4661	8.2600	8.1017	7.9761
7	12.246	9.5466	8.4513	7.8466	7.4604	7.1914	6.9928	6.8400	6.7188
8	11.259	8.6491	7.5910	7.0061	6.6318	6.3707	6.1776	6.0289	5.9106
9	10.561	8.0215	6.9919	6.4221	6.0569	5.8018	5.6129	5.4671	5.3511
10	10.044	7.5594	6.5523	5.9943	5.6363	5.3858	5.2001	5.0567	4.9424
11	9.6460	7.2057	6.2167	5.6683	5.3160	5.0692	4.8861	4.7445	4.6315
12	9.3302	6.9266	5.9525	5.4120	5.0643	4.8206	4.6395	4.4994	4.3875
13	9.0738	6.7010	5.7394	5.2053	4.8616	4.6204	4.4410	4.3021	4.1911
14	8.8616	6.5149	5.5639	5.0354	4.6950	4.4558	4.2779	4.1399	4.0297
15	8.6831	6.3589	5.4170	4.8932	4.5556	4.3183	4.1415	4.0045	3.8948
16	8.5310	6.2262	5.2922	4.7726	4.4374	4.2016	4.0259	3.8896	3.7804
17	8.3997	6.1121	5.1850	4.6690	4.3359	4.1015	3.9267	3.7910	3.6822
18	8.2854	6.0129	5.0919	4.5790	4.2479	4.0146	3.8406	3.7054	3.5971
19	8.1849	5.9259	5.0103	4.5003	4.1708	3.9386	3.7653	3.6305	3.5225
20	8.0960	5.8489	4.9382	4.4307	4.1027	3.8714	3.6987	3.5644	3.4567
21	8.0166	5.7804	4.8740	4.3688	4.0421	3.8117	3.6396	3.5056	3.3981

minator Degrees of Freedom



- Numerator df = 4
- Denominator df = 7

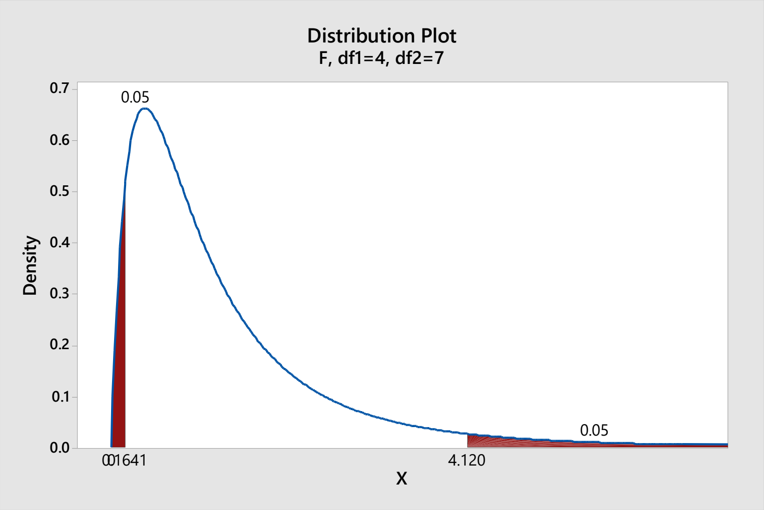
- α = 0.10 Two Tail
- $F_{0.05, 4, 7}$ = 4.1203

Two Variances Test

Critical Test Statistic

F - Distribution (α = 0.05 in the Right Tail)

df_2 \ df_1 (Denominator Degrees of Freedom \ Numerator Degrees of Freedom)	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
4	7.7086	9.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	6.9988
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799



- Numerator df = 4
- Denominator df = 7

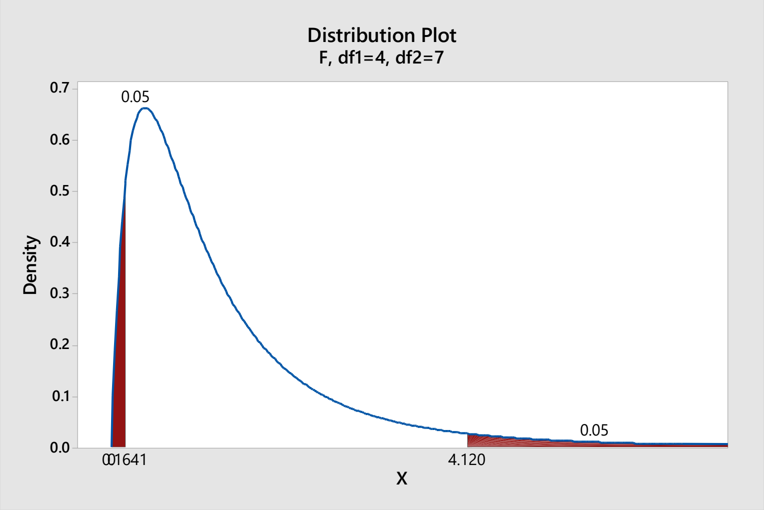
- α = 0.10 Two Tail
- $F_{0.05, 4, 7}$ = 4.1203
- $F_{0.95, 4, 7}$ = ?

Two Variances Test

Critical Test Statistic

F - Distribution ($\alpha = 0.05$ in the Right Tail)

df_2 \ df_1 (Numerator Degrees of Freedom →; Denominator Degrees of Freedom ↓)	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
4	7.7086	9.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	6.9988
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799



- Numerator df = 4
- Denominator df = 7

- $\alpha = 0.10$ Two Tail
- $F_{0.05, 4, 7} = 4.1203$

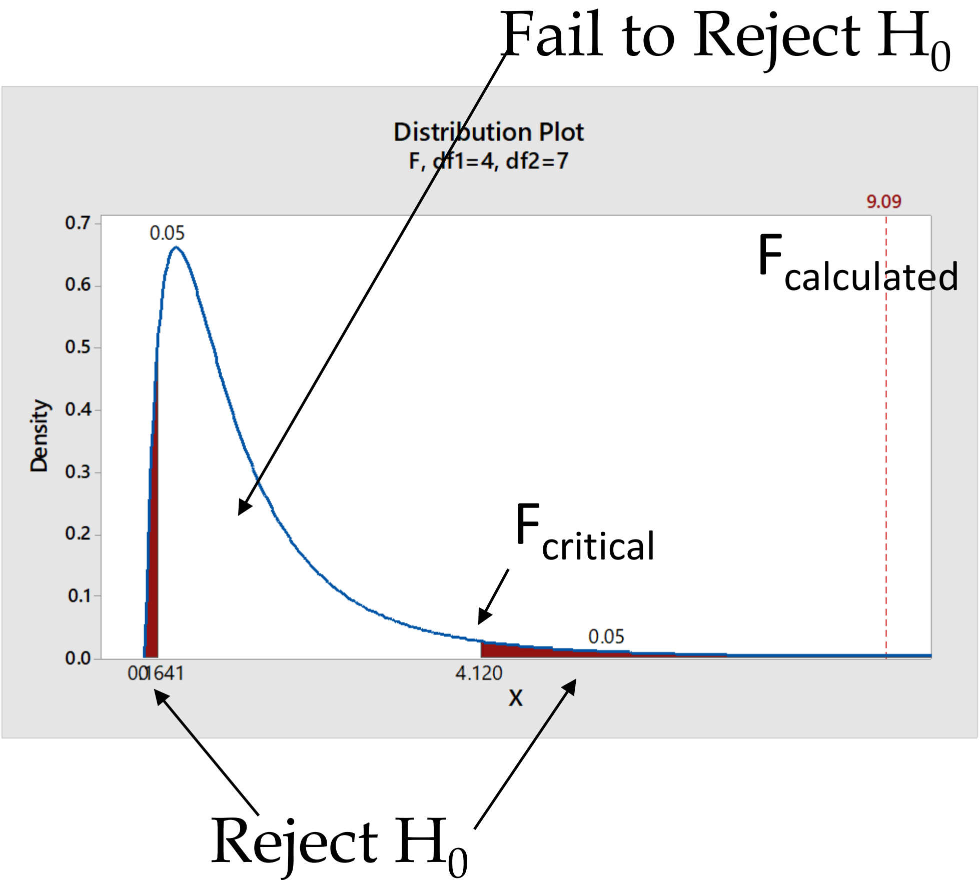
$F_{0.95, 4, 7} = 1 / F_{0.05, 7, 4}$

$F_{0.95, 4, 7} = 1 / 6.0942 = 0.164$

Two Variances Test

Two Variances Test

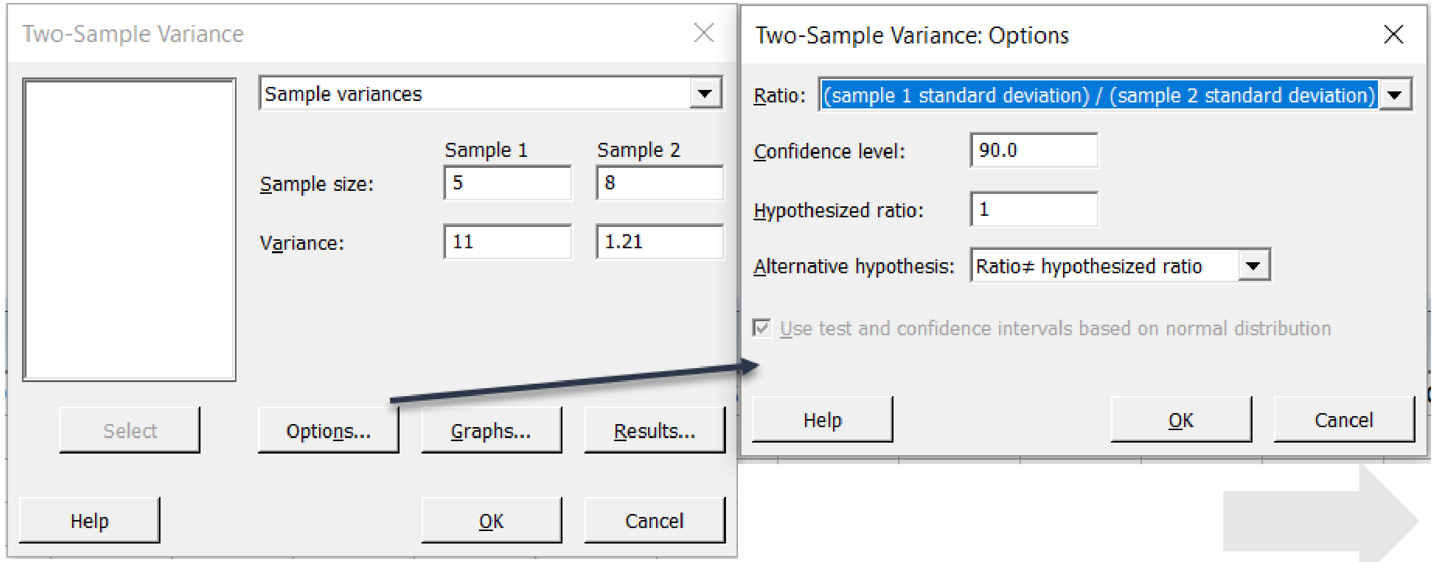
$H_0: \sigma^2_1 = \sigma^2_2$
$Ha: \sigma^2_1 \neq \sigma^2_2$



- Example: We took 8 samples from machine A and the **standard deviation** was 1.1. For machine B we took 5 samples and the **variance** was 11. Is there a difference in variance at 90% confidence level?
- $n1 = 5, s^2_1 = 11, df_1 = 4$ (numerator)
- $n2 = 8, s_2 = 1.1, s^2_2 = 1.21, df_2 = 7$ (denominator)
- $F_{calculated} = 11/1.21 = 9.09$ (higher value at top)
- $F_{critical} = 0.0164$ and 4.120

Two Variances Test

One Variance Test - Minitab



Test and CI for Two Variances

Method

σ_1^2: variance of Sample 1
σ_2^2: variance of Sample 2
Ratio: σ_1^2/σ_2^2
F method was used. This method is accurate for normal data only.

Descriptive Statistics

Sample	N	StDev	Variance	90% CI for σ
Sample 1	5	3.317	11.000	(2.154, 7.868)
Sample 2	8	1.100	1.210	(0.776, 1.977)

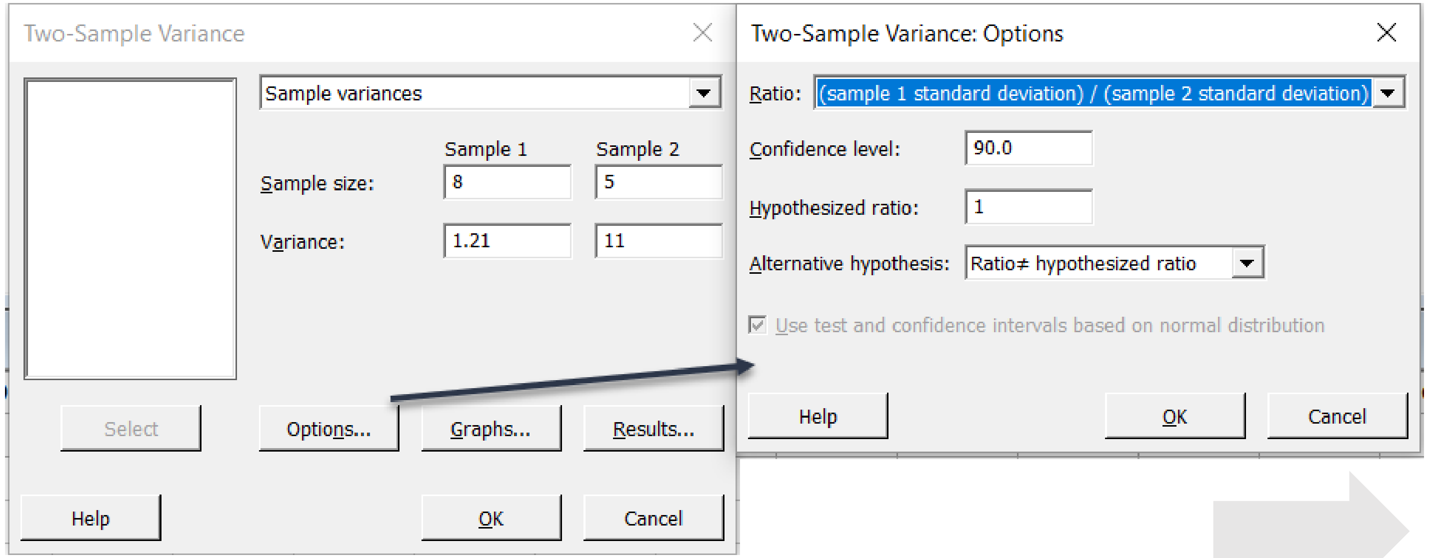
Test

Null hypothesis	H_0: $\sigma_1 / \sigma_2 = 1$
Alternative hypothesis	H_1: $\sigma_1 / \sigma_2 \neq 1$
Significance level	$\alpha = 0.1$

Method	Test Statistic	DF1	DF2	P-Value
F	9.09	4	7	0.013

Two Variances Test

Two Variances Test - Minitab



Test and CI for Two Variances

Method

σ_1^2: variance of Sample 1
σ_2^2: variance of Sample 2
Ratio: σ_1^2/σ_2^2
F method was used. This method is accurate for normal data only.

Descriptive Statistics

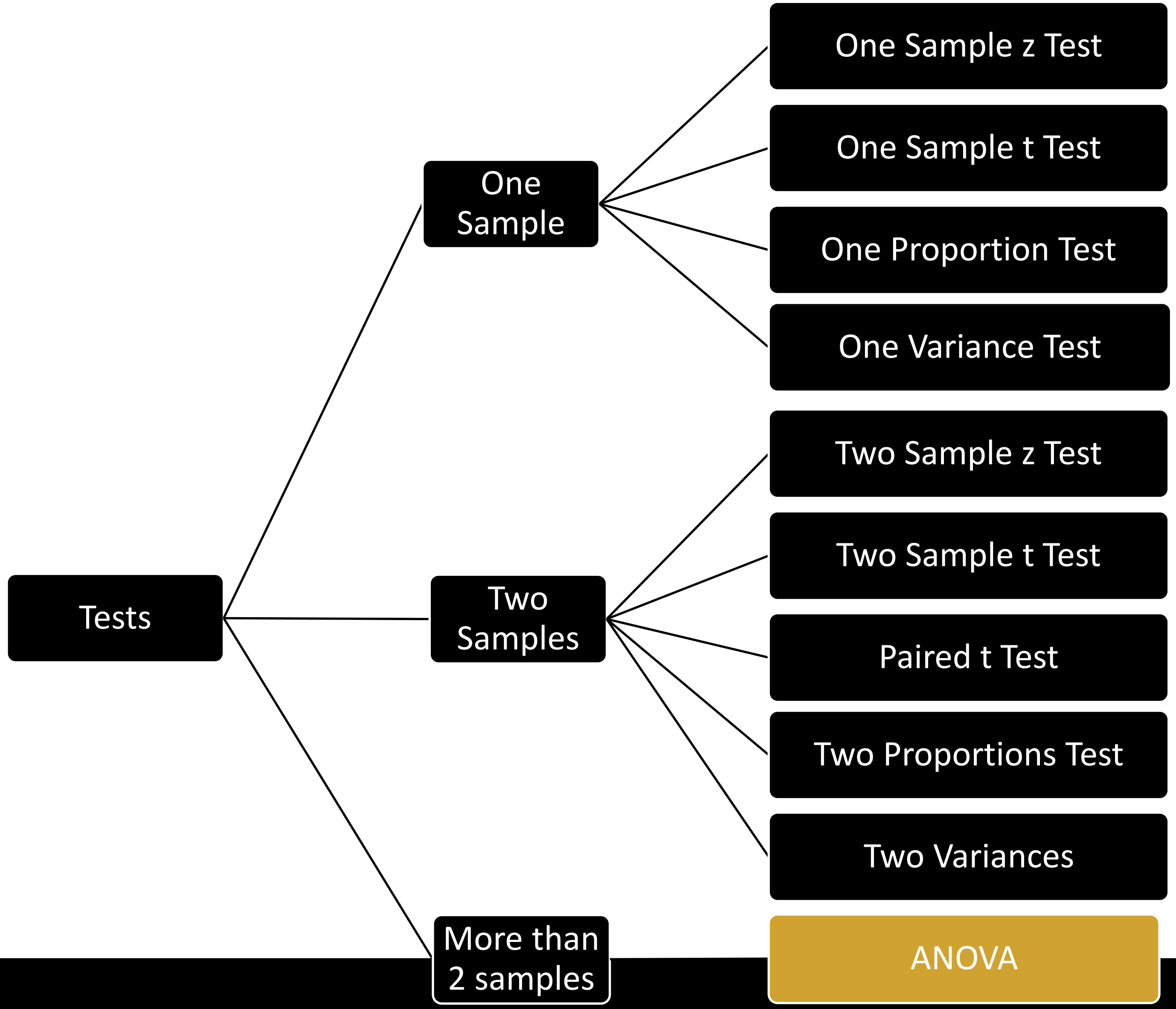
Sample	N	StDev	Variance	90% CI for σ
Sample 1	8	1.100	1.210	(0.776, 1.977)
Sample 2	5	3.317	11.000	(2.154, 7.868)

Test

Null hypothesis	H_0: $\sigma_1 / \sigma_2 = 1$
Alternative hypothesis	H_1: $\sigma_1 / \sigma_2 \neq 1$
Significance level	$\alpha = 0.1$

Method	Test Statistic	DF1	DF2	P-Value
F	0.11	7	4	0.013

Two Variances Test


Tests
One Sample
Two Samples
More than 2 samples
One Sample z Test
One Sample t Test
One Proportion Test
One Variance Test
Two Sample z Test
Two Sample t Test
Paired t Test
Two Proportions Test
Two Variances
ANOVA

ANOVA

Variance Tests

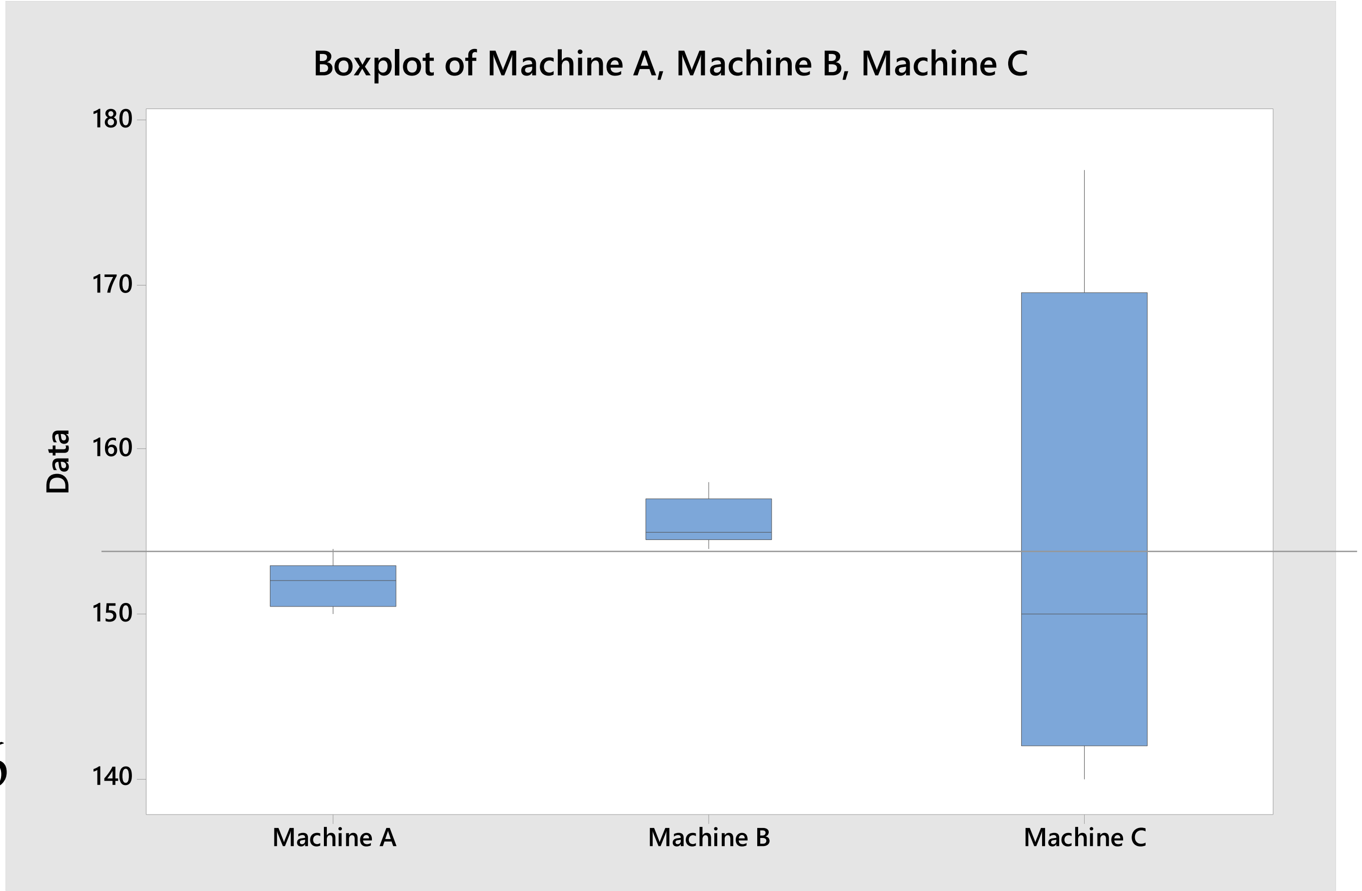
- ❖ Chi-square test
  - ❖ For testing the population variance against a specified value
  - ❖ testing goodness of fit of some probability distribution
  - ❖ testing for independence of two attributes (Contingency Tables)

- ❖ F-test
  - ❖ for testing equality of *two* variances from different populations
  - ❖ for testing equality of several means with technique of ANOVA.

Two Sample t Tests

Machine A	Machine B	Machine C
150	156	144
152	155	162
154	158	177
152	155	150
151	154	140

$\bar{x}_{\mathbf{A}} = 151.8$, $\bar{x}_{\mathrm{B}} = 155.6$ $\bar{x}_{\mathrm{C}} = 154.6$



ANOVA

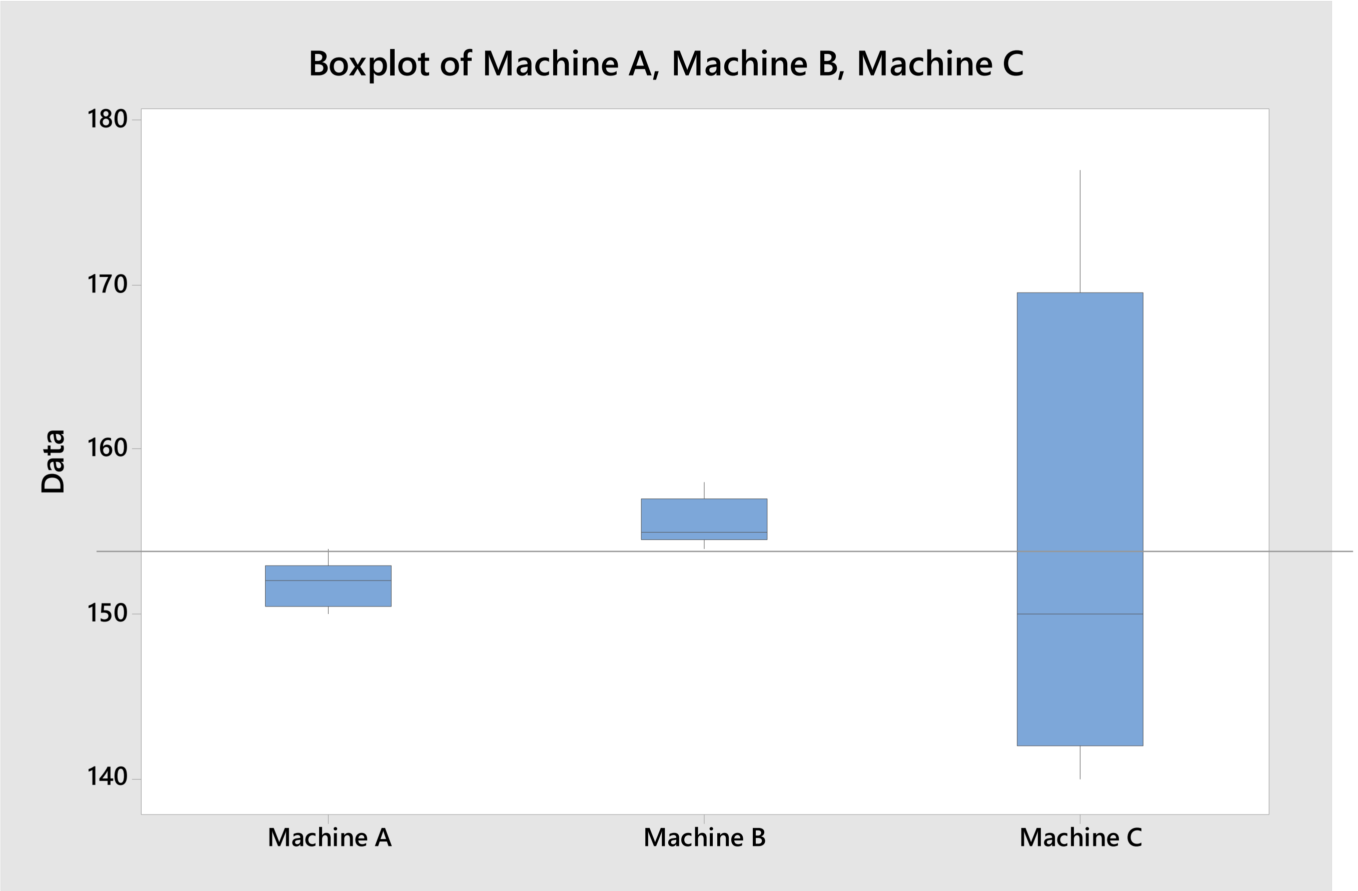
T Test vs ANOVA

2 Sample T Test

$H_0: \mu_A = \mu_B$
$H_a: \mu_A \neq \mu_B$

ANOVA

$H_0: \mu_A = \mu_B = \mu_C = \mu_D \ldots\ldots = \mu_k$
H_a: At least one of the means is different from others



ANOVA

T Test vs ANOVA

T Test

H_0: $\mu_A = \mu_B$
H_a: $\mu_A \neq \mu_B$

ANOVA

H_0: $\mu_A = \mu_B = \mu_C = \mu_{D\,.....} = \mu_k$
H_a: At least one of the means is different from others

- **Why ANOVA?**
  - We used t test to compare the means of two populations.
  - What if we need to compare more than two populations? With ANOVA we can find out if one or more populations have different mean or comes from a different population.
  - We could have conducted multiple t Test.
  - How many t Test we need to conduct if have to compare 4 sample means? … 6

T Test vs ANOVA

4				x
3			x	3 vs 4
2		x	2 vs 3	2 vs 4
1	x	1 vs 2	1 vs 3	1 vs 4
	1	2	3	4

- **Why ANOVA?**
  - We used t test to compare the means of two populations.
  - What if we need to compare more than two populations? With ANOVA we can find out if one or more populations have different mean or comes from a different population.
  - We could have conducted multiple t Test.
  - How many t Test we need to conduct if have to compare 4 sample means? … 6

ANOVA

T Test vs ANOVA

4				x
3			x	3 vs 4
2		x	2 vs 3	2 vs 4
1	x	1 vs 2	1 vs 3	1 vs 4
	1	2	3	4

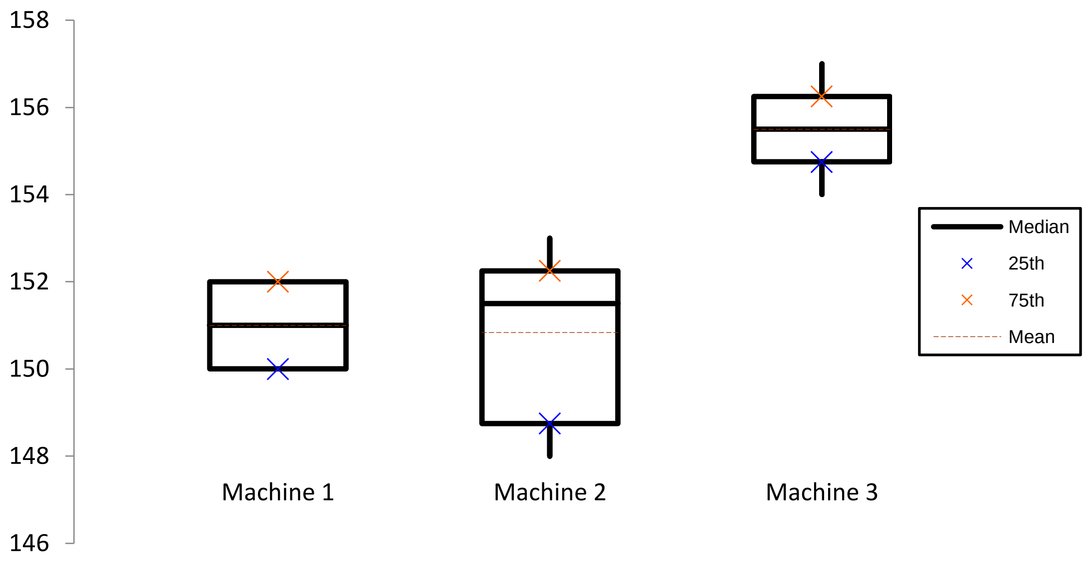
- Why ANOVA?
  - How many t Test we need to conduct if have to compare 4 samples? … 6
  - Each test is done with alpha = 0.05 or 95% confidence.
  - 6 tests will result in confidence level of 0.95x0.95x0.95x0.95x0.95x0.95 = 0.735

ANOVA

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	$\bar{x}_2 = 150.83$	$\bar{x}_3 = 155.50$

ANOVA

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	**$\bar{x}_2 = 150.83$**	**$\bar{x}_3 = 155.50$**

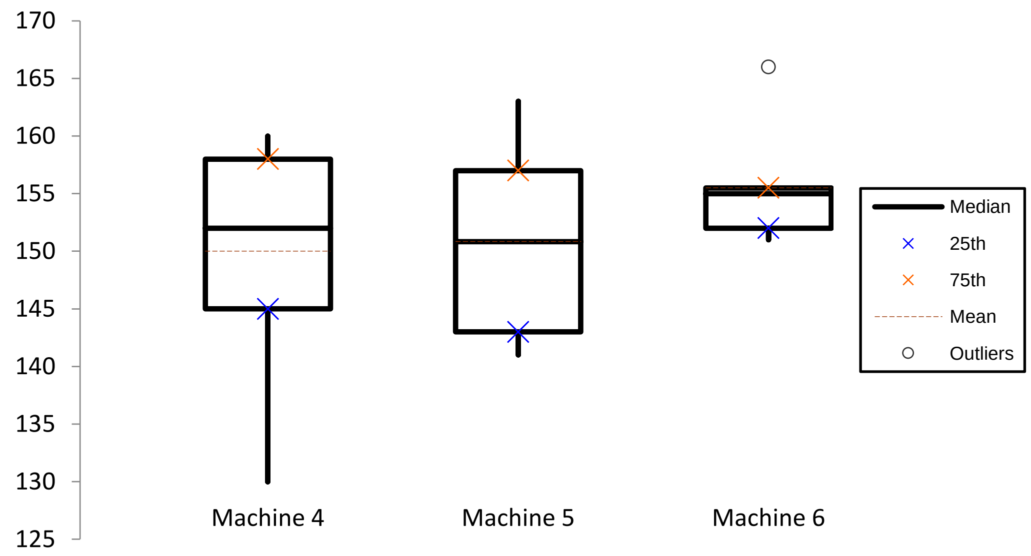


ANOVA

Machine 4	Machine 5	Machine 6
130	163	166
155	152	154
160	143	155
158	141	151
152	149	152
145	157	155
$\bar{x}_4$ = 151.00	**$\bar{x}_5$ = 150.83**	**$\bar{x}_6$ = 155.50**

ANOVA

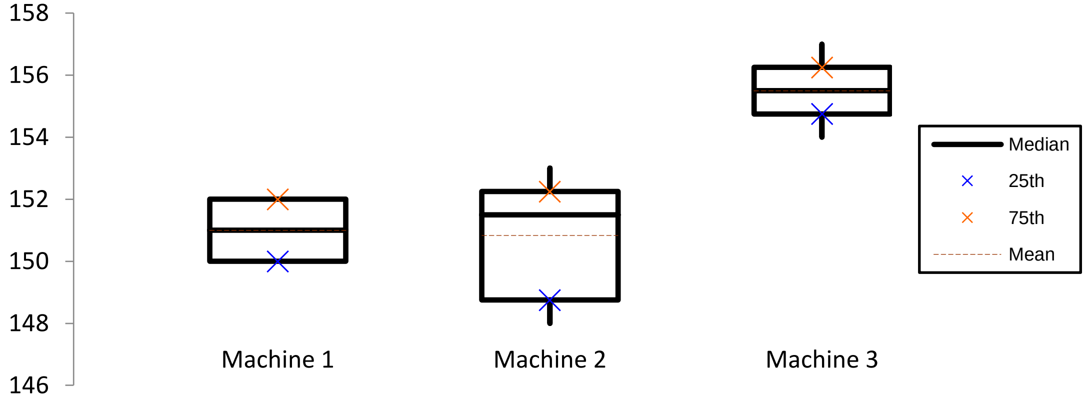
Machine 4	Machine 5	Machine 6
130	163	166
155	152	154
160	143	155
158	141	151
152	149	152
145	157	155
$\bar{x}_4 = 151.00$	$\bar{x}_5 = 150.83$	$\bar{x}_6 = 155.50$

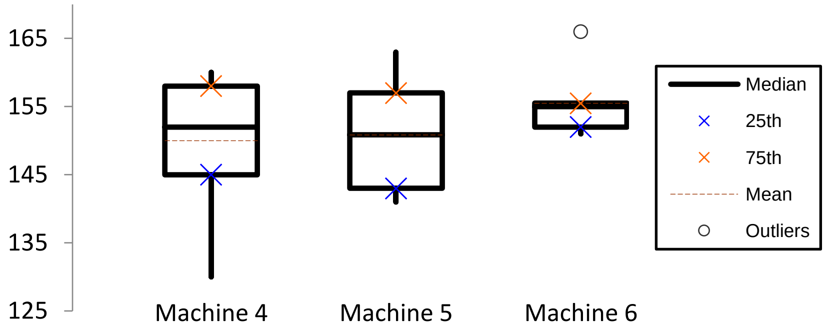


ANOVA

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151.00$	$\bar{x}_2 = 150.83$	$\bar{x}_3 = 155.50$

Machine 4	Machine 5	Machine 6
130	163	166
155	152	154
160	143	155
158	141	151
152	149	152
145	157	155
$\bar{x}_4 = 151.00$	$\bar{x}_5 = 150.83$	$\bar{x}_6 = 155.50$





ANOVA

- ANOVA is Analysis of Variance
- Variance

$$s^2 = \frac{\sum(x_i - \bar{X})^2}{n-1}$$

- Numerator of this formula is Sum of Squares, the denominator is the degrees of freedom for the sample.

ANOVA

ANOVA

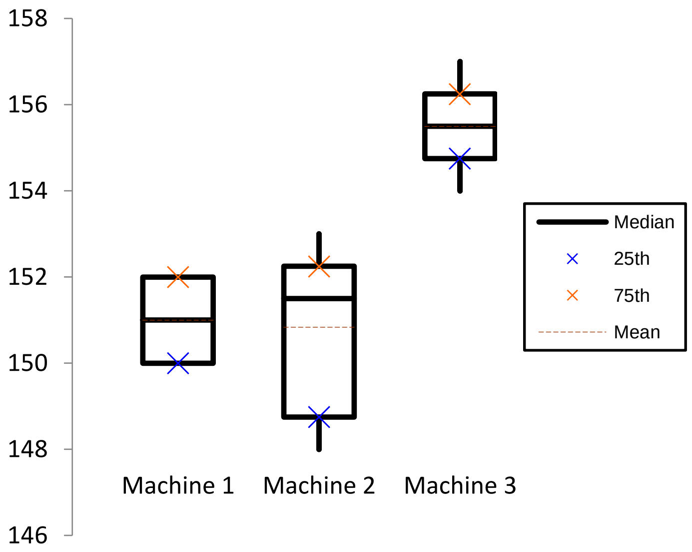
$$F = \frac{s_1^2}{s_2^2}$$

$$F = \frac{\frac{\sum(x-\bar{x}_1)^2}{n_1-1}}{\frac{\sum(x-\bar{x}_2)^2}{n_2-1}}$$

$$F = \frac{\frac{SS_1}{\mathrm{d}f_1}}{\frac{\mathrm{SS}_2}{\mathrm{d}f_2}} \qquad F = \frac{MSS_1}{MSS_2}$$

$$F = \frac{MSS_{between}}{MSS_{within}}$$

$$F = \frac{\frac{SS_{\mathrm{between}}}{\mathrm{d}f_{\mathrm{between}}}}{\frac{\mathrm{SS}_{within}}{\mathrm{d}f_{\mathrm{within}}}}$$

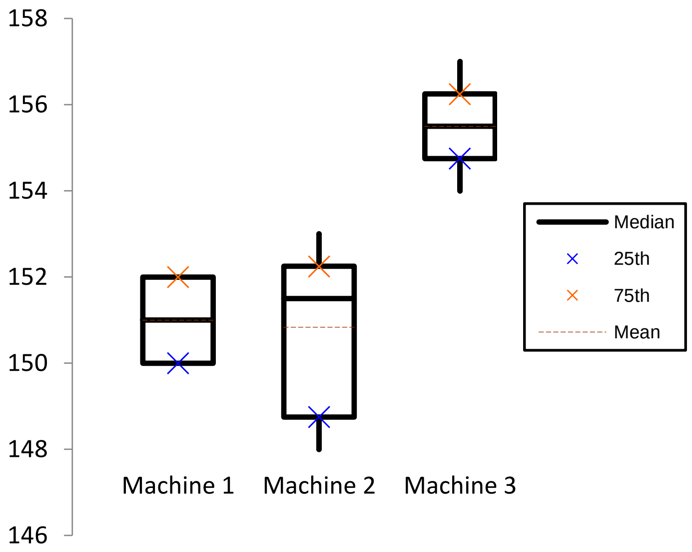


ANOVA

ANOVA

$$F = \frac{\frac{SS_{\text{between}}}{\mathrm{d}f_{\text{between}}}}{\frac{SS_{within}}{\mathrm{d}f_{\text{within}}}}$$

$$F = \frac{MSS_{between}}{MSS_{within}}$$



- $SST = SS_{\text{between(or treatment, or column)}} + SS_{\text{within(or error)}}$

ANOVA

ANOVA

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	**$\bar{x}_2 = 150.83$**	**$\bar{x}_3 = 155.50$**

- **Null hypothesis**: $H_0: \mu_1 = \mu_2 = \mu_3$
- **Alternative hypothesis**: H_a : Means are not all equal

Check at 95% confidence level.

- $SS_{\text{between(or treatment, or column)}}$
- $SS_{\text{within(or error)}}$

$$F = \frac{\frac{SS_{\text{between}}}{df_{\text{between}}}}{\frac{SS_{\text{within}}}{df_{\text{within}}}} \qquad F = \frac{MSS_{between}}{MSS_{within}}$$

ANOVA

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	$\bar{x}_2 = 150.83$	$\bar{x}_3 = 155.50$

- $SS_{within} = 4.00+18.83+5.50 = 28.33$

Machine 1	x1 - x̄1	Sqr(x1 - x̄1)	Machine 2	x2 - x̄2	Sqr(x2 - x̄2)	Machine 3	x3 - x̄3	Sqr(x3 - x̄3)	
150.00	-1.00	1.00	153.00	2.17	4.69	156.00	0.50	0.25	
151.00	0.00	0.00	152.00	1.17	1.36	154.00	-1.50	2.25	
152.00	1.00	1.00	148.00	-2.83	8.03	155.00	-0.50	0.25	
152.00	1.00	1.00	151.00	0.17	0.03	156.00	0.50	0.25	
151.00	0.00	0.00	149.00	-1.83	3.36	157.00	1.50	2.25	
150.00	-1.00	1.00	152.00	1.17	1.36	155.00	-0.50	0.25	
151.00			150.83			155.50			152.44
		4.00			18.83			5.50	

ANOVA

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	$\bar{x}_2 = 150.83$	$\bar{x}_3 = 155.50$

- $SS_{between} = (2.07+2.58+9.36)x6 = 84.06$

Machine 1	x1 - x̄1	Sqr(x1 - x̄1)	Machine 2	x2 - x̄2	Sqr(x2 - x̄2)	Machine 3	x3 - x̄3	Sqr(x3 - x̄3)	
150.00	-1.00	1.00	153.00	2.17	4.69	156.00	0.50	0.25	
151.00	0.00	0.00	152.00	1.17	1.36	154.00	-1.50	2.25	
152.00	1.00	1.00	148.00	-2.83	8.03	155.00	-0.50	0.25	
152.00	1.00	1.00	151.00	0.17	0.03	156.00	0.50	0.25	
151.00	0.00	0.00	149.00	-1.83	3.36	157.00	1.50	2.25	
150.00	-1.00	1.00	152.00	1.17	1.36	155.00	-0.50	0.25	
151.00			150.83			155.50			152.44
		4.00			18.83			5.50	
	-1.44	2.07		-1.61	2.58		3.06	9.36	

ANOVA

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	$\bar{x}_2 = 150.83$	$\bar{x}_3 = 155.50$

- $SS_{within} = 4.00+18.83+5.50 = 28.33$
- $SS_{between} = (2.07+2.58+9.36)x6 = 84.06$

Machine 1	x1 - x̄1	Sqr(x1 - x̄1)	Machine 2	x2 - x̄2	Sqr(x2 - x̄2)	Machine 3	x3 - x̄3	Sqr(x3 - x̄3)	
150.00	-1.00	1.00	153.00	2.17	4.69	156.00	0.50	0.25	
151.00	0.00	0.00	152.00	1.17	1.36	154.00	-1.50	2.25	
152.00	1.00	1.00	148.00	-2.83	8.03	155.00	-0.50	0.25	
152.00	1.00	1.00	151.00	0.17	0.03	156.00	0.50	0.25	
151.00	0.00	0.00	149.00	-1.83	3.36	157.00	1.50	2.25	
150.00	-1.00	1.00	152.00	1.17	1.36	155.00	-0.50	0.25	
151.00			150.83			155.50			152.44
		4.00			18.83			5.50	
	-1.44	2.07		-1.61	2.58		3.06	9.36	

ANOVA

ANOVA

- SST = SS $_{\text{between(or treatment)}}$ +SS $_{\text{within(or error)}}$
- SST = 84.06 +28.33 = 112.39

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1$ = 151	**$\bar{x}_2$ = 150.83**	**$\bar{x}_3$ = 155.50**

- **Degrees of freedom**
- Total df = df $_{\text{between(or treatment)}}$ + df $_{\text{within(or error)}}$
- (N-1) = (C-1) + (N-C)
- df $_{\text{between}}$ = 3-1=2,
- df $_{\text{total}}$ = 17 , df $_{\text{within}}$ = 17-2=15

- SS $_{\text{within}}$ = 28.33
- SS $_{\text{between}}$ = 84.06

QG

ANOVA

ANOVA

- **<u>Mean Sum of Square = SS / df</u>**
- $MSS_{between} = SS_{between} / df_{between}$
- $MSS_{between} = 84.06 / 2 = 42.03$

- $MSS_{within} = SS_{within} / df_{within}$
- $MSS_{within} = 28.33 / 15 = 1.89$

- $F = MSS_{between} / MSS_{within} = 42.03/1.89 = 22.24$

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	**$\bar{x}_2 = 150.83$**	**$\bar{x}_3 = 155.50$**

- $SS_{within} = 28.33$
- $SS_{between} = 84.06$
- $df_{within} = 15$
- $df_{between} = 2$

ANOVA

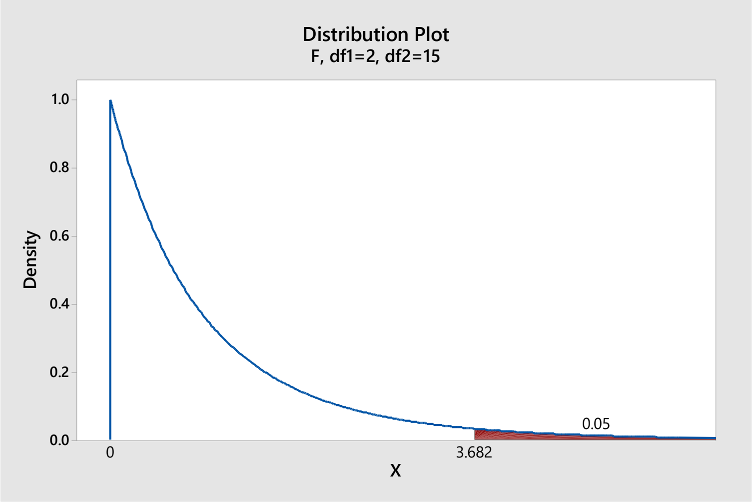
Critical Test Statistic

F - Distribution (α = 0.05 in the Right Tail)

df_2 \ df_1 (Numerator Degrees of Freedom) / Denominator Degrees of Freedom	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
4	7.7086	9.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	6.9988
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799

$$F = \frac{\dfrac{SS_{\text{between}}}{df_{\text{between}}}}{\dfrac{SS_{within}}{df_{\text{within}}}}$$

- Numerator (between) df = 2
- Denominator (within) df = 15



- α = 0.05 One Tail
- F 0.05, 2, 15 = 3.68

ANOVA

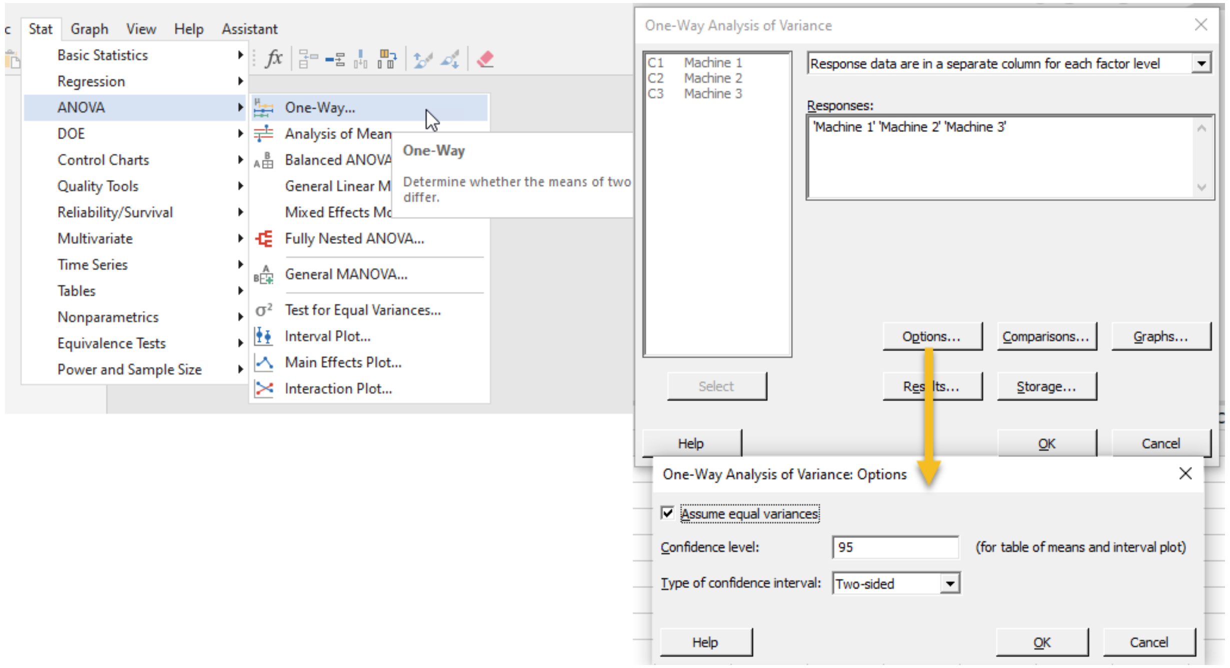
ANOVA

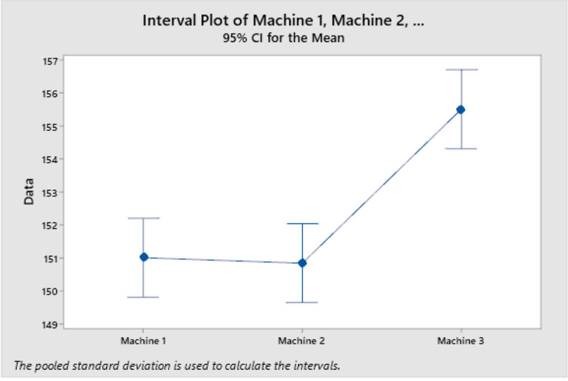
- $F = MSS_{between} / MSS_{within} = 42.03/1.89 = 22.24$

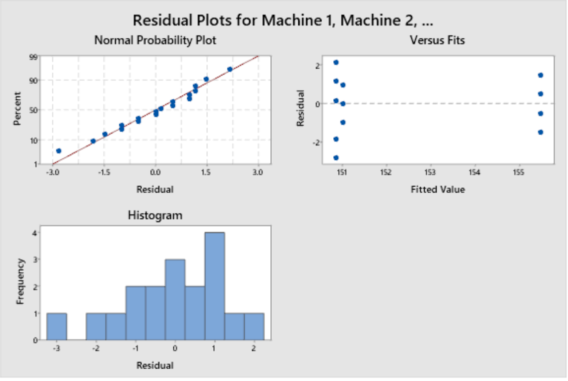
- Compare this with $F_{critical}$
- $F\ (0.05, 2, 15) = 3.68$

- Reject Null Hypothesis

ANOVA- Minitab







18 ONE WAY ANOVA.MTW

One-way ANOVA: Machine 1, Machine 2, Machine 3

Method

Null hypothesis	All means are equal
Alternative hypothesis	Not all means are equal
Significance level	α = 0.05

Equal variances were assumed for the analysis.

Factor Information

Factor	Levels	Values
Factor	3	Machine 1, Machine 2, Machine 3

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	2	84.11	42.056	22.26	0.000
Error	15	28.33	1.889		
Total	17	112.44			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1.37437	74.80%	71.44%	63.72%

Means

Factor	N	Mean	StDev	95% CI
Machine 1	6	151.000	0.894	(149.804, 152.196)
Machine 2	6	150.833	1.941	(149.637, 152.029)
Machine 3	6	155.500	1.049	(154.304, 156.696)

Pooled StDev = 1.37437

ANOVA

Variance Tests

- Chi-square test
  - For testing the population variance against a specified value
  - testing goodness of fit of some probability distribution
  - testing for independence of two attributes (Contingency Tables)

- F-test
  - for testing equality of *two* variances from different populations
  - for testing equality of several means with technique of ANOVA.

Variance Tests – Chi-Square

Goodness of Fit Test (Chi Square)

- ❖ To test if the sample is coming from a population with specific distribution.
- ❖ Other goodness-of-fit tests are
  - ❖ Anderson-Darling
  - ❖ Kolmogorov-Smirnov

Goodness of Fit Test

Goodness of Fit Test (Chi Square)

- H_0: The data follow a specified distribution.
- Ha: The data do not follow the specified distribution.
- Calculated Statistic: $$\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$$
- Critical Statistic: Chi square for k-1 degrees of freedom for specific alpha.

Goodness of Fit Test

Goodness of Fit Test (Chi Square)

- A coin is flipped 100 times. Number of heads and tails are noted. Is this coin biased? Check with 95% Confidence Level.

Head	40
Tail	60

Goodness of Fit Test

Ho: Coin is not biased.
Ha: Coin is biased.
Alpha = 0.05

- Example 1: A coin is flipped 100 times. Number of heads and tails are noted. Is this coin biased? Check with 95% Confidence Level.

Head	40
Tail	60

Goodness of Fit – Chi-Square

Variance Tests

- Chi-square test
  - For testing the population variance against a specified value
  - testing goodness of fit of some probability distribution
  - testing for independence of two attributes (Contingency Tables)

- F-test
  - for testing equality of *two* variances from different populations
  - for testing equality of several means with technique of ANOVA.

Variance Tests – Chi-Square

Contingency Tables

- To find relationship between two discrete variables.

	Smoker	Non Smoker	
Male	60	40	100
Female	35	40	75
	95	80	175

	Operator 1	Operator 2	Operator 3	
Shift 1	22	26	23	71
Shift 2	28	62	26	116
Shift 3	72	22	66	160
	122	110	115	347

Contingency Tables

	Smoker	Non Smoker	
Male	60	40	100
Female	35	40	75
	95	80	175

	Operator 1	Operator 2	Operator 3	
Shift 1	22	26	23	71
Shift 2	28	62	26	116
Shift 3	72	22	66	160
	122	110	115	347

- Null hypothesis is that there is no relationship between the row and column variables.
- Alternate hypothesis is that there is a relationship. Alternate hypothesis does not tell what type of relationship exists.

Contingency Tables

Calculating Test Statistic (Chi Square)

$$\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$$

	Operator 1	Operator 2	Operator 3	
Shift 1	22	26	23	71
Shift 2	28	62	26	116
Shift 3	72	22	66	160
	122	110	115	347

Calculating Test Statistic (Chi Square)

OBSERVED	Operator 1	Operator 2	Operator 3	
Shift 1	22	26	23	71
Shift 2	28	62	26	116
Shift 3	72	22	66	160
	122	110	115	347

EXPECTED	Operator 1	Operator 2	Operator 3	
Shift 1	122x71/347	110x71/347	115x71/347	71
Shift 2	122x116/347	110x116/347	115x116/347	116
Shift 3	122x160/347	110x160/347	115x160/347	160
	122	110	115	347

$$\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$$

Contingency Tables

Calculating Test Statistic (Chi Square)

OBSERVED	Operator 1	Operator 2	Operator 3	
Shift 1	22	26	23	71
Shift 2	28	62	26	116
Shift 3	72	22	66	160
	122	112	115	347

EXPECTED	Operator 1	Operator 2	Operator 3	
Shift 1	122x71/347	110x71/347	115x71/347	71
Shift 2	122x116/347	110x116/347	115x116/347	116
Shift 3	122x160/347	110x160/347	115x160/347	160
	122	110	115	347

EXPECTED	Operator 1	Operator 2	Operator 3	
Shift 1	24.96	22.51	23.53	71
Shift 2	40.78	36.77	38.44	116
Shift 3	56.25	50.72	53.02	160
	122	112	115	347

$$\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$$

Contingency Tables

Calculating Test Statistic (Chi Square)

OBSERVED	Operator 1	Operator 2	Operator 3	
Shift 1	22	26	23	71
Shift 2	28	62	26	116
Shift 3	72	22	66	160
	122	110	115	347

EXPECTED	Operator 1	Operator 2	Operator 3	
Shift 1	122x71/347	110x71/347	115x71/347	71
Shift 2	122x116/347	110x116/347	115x116/347	116
Shift 3	122x160/347	110x160/347	115x160/347	160
	122	110	115	347

$(O-E)^2/E$	Operator 1	Operator 2	Operator 3	
Shift 1	$(22-24.96)^2/24.96 = 0.35$	0.54	0.01	71
Shift 2	$(28-40.78)^2/40.78 = 4.00$	17.31	4.03	116
Shift 3	$(72-56.25)^2/56.25 = 4.41$	16.26	3.18	160
	122	112	115	347

EXPECTED	Operator 1	Operator 2	Operator 3	
Shift 1	24.96	22.51	23.53	71
Shift 2	40.78	36.77	38.44	116
Shift 3	56.25	50.72	53.02	160
	122	112	115	347

$$\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$$

Contingency Tables

Calculating Test Statistic (Chi Square)

$(O-E)^2/E$	Operator 1	Operator 2	Operator 3	
Shift 1	$(22-24.96)^2/24.96 = 0.35$	0.54	0.01	71
Shift 2	$(28-40.78)^2/40.78 = 4.00$	17.31	4.03	116
Shift 3	$(72-56.25)^2/56.25 = 4.41$	16.26	3.18	160
	122	112	115	347

$X^2 = 50.09$

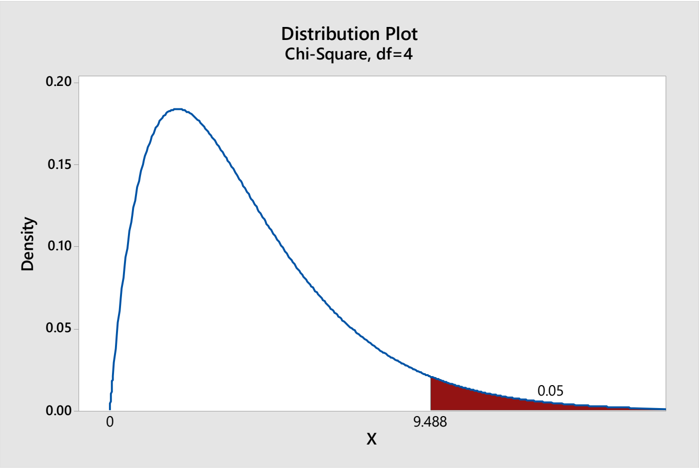
$$\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$$

Contingency Tables

Critical Test Statistic

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41
18	7.015	9.390	10.865	13.675	17.338	21.60	25.99	28.87	34.80
19	7.633	10.117	11.651	14.562	18.338	22.72	27.20	30.14	36.19
20	8.260	10.851	12.443	15.452	19.337	23.83	28.41	31.41	37.57
22	9.542	12.338	14.041	17.240	21.337	26.04	30.81	33.92	40.29
24	10.856	13.848	15.659	19.037	23.337	28.24	33.20	36.42	42.98
26	12.198	15.379	17.292	20.843	25.336	30.43	35.56	38.89	45.64
28	13.565	16.928	18.939	22.657	27.336	32.62	37.92	41.34	48.28
30	14.953	18.493	20.599	24.478	29.336	34.80	40.26	43.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38



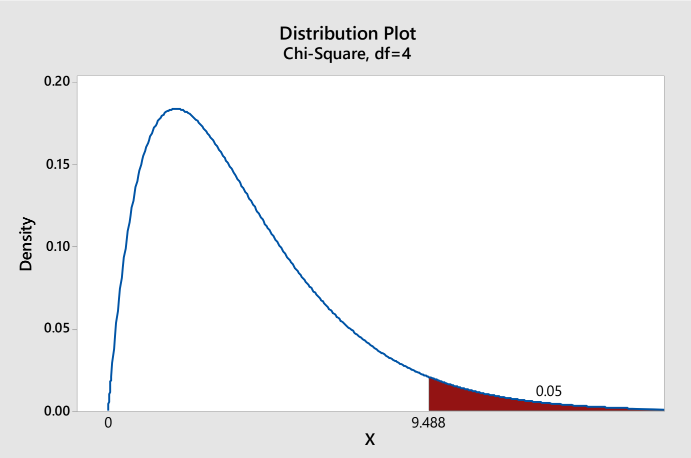
- Df = (r-1)(c-1) = 4
- α = 0.05 One Tail
- χ^2 Critical = 9.49

Contingency Tables

Critical Test Statistic

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.0							19.68	24.72
12	3.5							21.03	26.22
13	4.1							22.36	27.69
14	4.6							23.68	29.14
15	5.2							25.00	30.58
16	5.8							26.30	32.00
17	6.4							27.59	33.41
18	7.0							28.87	34.80
19	7.6							30.14	36.19
20	8.2							31.41	37.57
22	9.5							33.92	40.29
24	10.							36.42	42.98
26	12.							38.89	45.64
28	13.							41.34	48.28
30	14.							43.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38

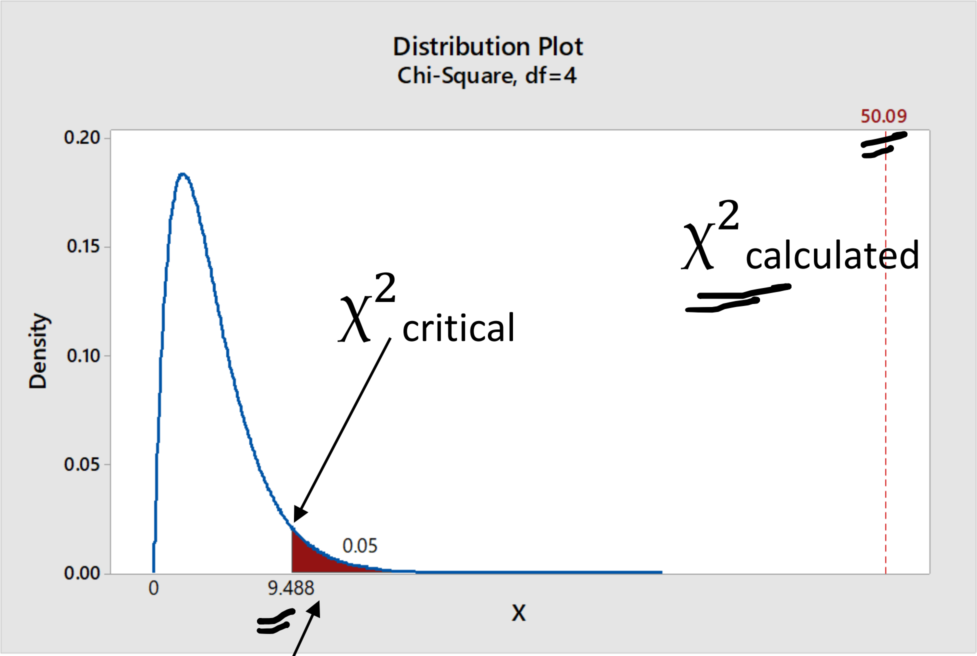


OBSERVED	Operator 1	Operator 2	Operator 3	
Shift 1	22	26	23	71
Shift 2	28	62	26	116
Shift 3	72	22	66	160
	122	112	115	347

- $\chi^2(\mathit{calculated}) = 50.09$
- $\chi^2(\text{critical}) = 9.49$

Contingency Tables

Null hypothesis is that there is no relationship between the row and column variables.
Alternate hypothesis is that there is a relationship. Alternate hypothesis does not tell what type of relationship exists.
Alpha = 0.05



Reject H_0

OBSERVED	Operator 1	Operator 2	Operator 3	
Shift 1	22	26	23	71
Shift 2	28	62	26	116
Shift 3	72	22	66	160
	122	112	115	347

- $\chi^2(calculated) = 50.09$
- $\chi^2(\mathrm{critical}) = 9.49$

Contingency Tables

- **Practice Exercise:**
- Calculate the Expected value for Non Smoker Male?
- What will be the degrees of freedom in this example?

	Smoker	Non Smoker	
Male	60	40	100
Female	35	40	75
	95	80	175

Contingency Tables

- **Practice Exercise: SOLUTION**
- Calculate the Expected value for Non Smoker Male? = 80x100/175 = 45.71
- What will be the degrees of freedom in this example? (2-1)(2-1)=1

	Smoker	Non Smoker	
Male	60	40	100
Female	35	40	75
	95	80	175

Contingency Tables

Ho: Coin is not biased.
Ha: Coin is biased.
Alpha = 0.05

- Example 1: A coin is flipped 100 times. Number of heads and tails are noted. Is this coin biased? Check with 95% Confidence Level.

$$\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$$

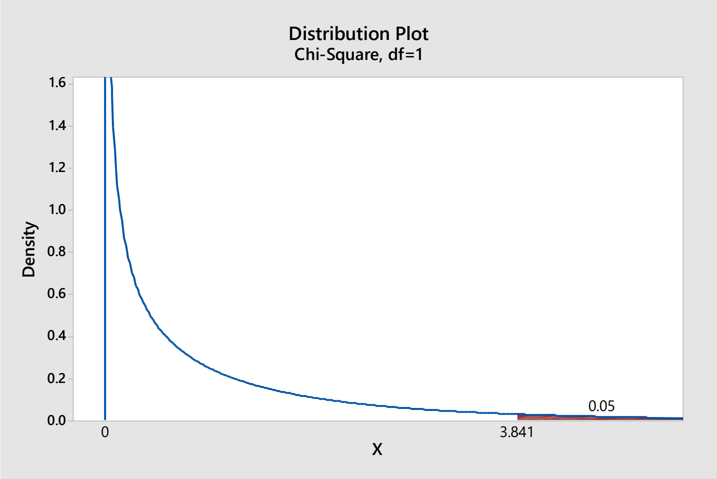
Flip	Expected	Observed	O-E	$(O-E)^2$	$(O-E)^2/E$
Head	50	40	-10	100	2
Tail	50	60	10	100	2
					$X^2 = 4$

Goodness of Fit – Chi-Square

Critical Test Statistic

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41
18	7.015	9.390	10.865	13.675	17.338	21.60	25.99	28.87	34.80
19	7.633	10.117	11.651	14.562	18.338	22.72	27.20	30.14	36.19
20	8.260	10.851	12.443	15.452	19.337	23.83	28.41	31.41	37.57
22	9.542	12.338	14.041	17.240	21.337	26.04	30.81	33.92	40.29
24	10.856	13.848	15.659	19.037	23.337	28.24	33.20	36.42	42.98
26	12.198	15.379	17.292	20.843	25.336	30.43	35.56	38.89	45.64
28	13.565	16.928	18.939	22.657	27.336	32.62	37.92	41.34	48.28
30	14.953	18.493	20.599	24.478	29.336	34.80	40.26	43.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38



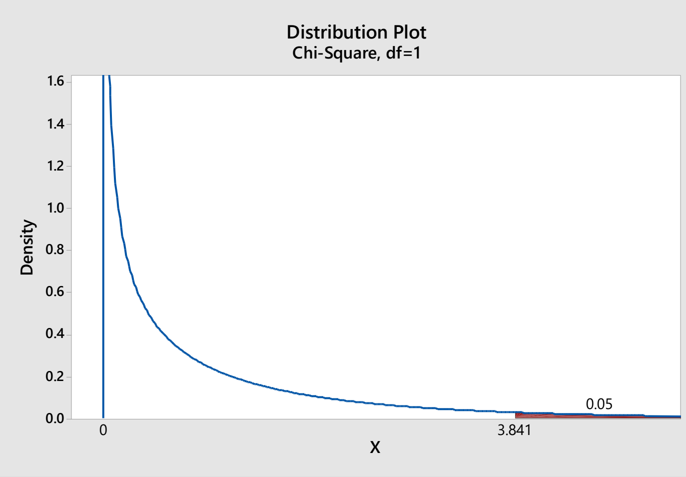
- $\alpha = 0.05$ One Tail
- Df = 1
- χ^2 Critical = 3.84

Goodness of Fit – Chi-Square

Critical Test Statistic

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	**0.95**	**0.90**	**0.75**	**0.50**	**0.25**	**0.10**	**0.05**	**0.01**
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11								19.68	24.72
12								21.03	26.22
13								22.36	27.69
14								23.68	29.14
15								25.00	30.58
16								26.30	32.00
17								27.59	33.41
18								28.87	34.80
19								30.14	36.19
20								31.41	37.57
22								33.92	40.29
24								36.42	42.98
26								38.89	45.64
28								41.34	48.28
30								43.77	50.89
40								55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38



❖ Example 1: A coin is flipped 100 times. Number of heads and tails are noted. Is this coin biased? Check with 95% Confidence Level.

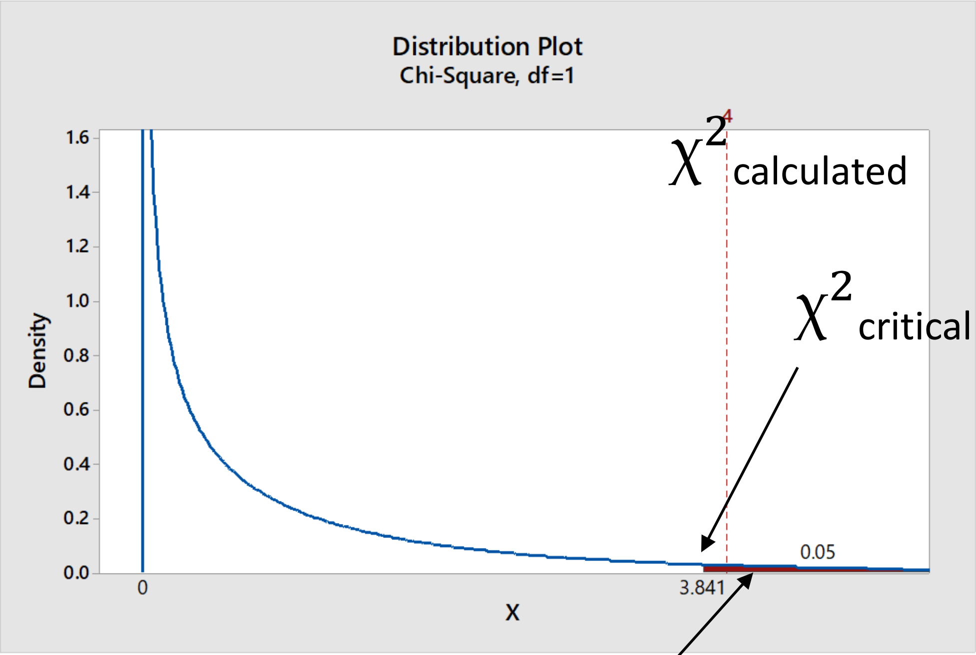
Head	40
Tail	60

❖ $\chi^2(calculated) = 4.0$

❖ $\chi^2(\text{critical}) = 3.84$

Goodness of Fit – Chi-Square

Ho: Coin is not biased.
Ha: Coin is biased.
Alpha = 0.05



Reject H_0

- Example 1: A coin is flipped 100 times. Number of heads and tails are noted. Is this coin biased? Check with 95% Confidence Level.

Head	40
Tail	60

- $\chi^2(calculated) = 4.0$

- $\chi^2(\text{critical}) = 3.84$

Goodness of Fit – Chi-Square

H0: The data follow a specified distribution.
Ha: The data do not follow the specified distribution.
Alpha = 0.05

- Example 2: A t-shirt manufacturer expects vs actual sale.

Size	Proportions	Counts
Small	0.1	25
Medium	0.2	41
Large	0.4	91
Extra Large	0.3	68

Goodness of Fit – Chi-Square

QG

- Example 2: A t-shirt manufacturer expects vs actual sale.

H0: The data follow a specified distribution.
Ha: The data do not follow the specified distribution.
Alpha = 0.05

$$\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$$

Size	Proportions	Expected	Observed
Small	0.1	22.5	25
Medium	0.2	45	41
Large	0.4	90	91
Extra Large	0.3	67.5	68

Goodness of Fit – Chi-Square