2-Sample t-Test to Compare Region 2 to Region 3

Stat>Basic Statistics>2-Sample t

Two Sample T-Test and Confidence Interval

Two sample T for Norfolk vs Phoenix

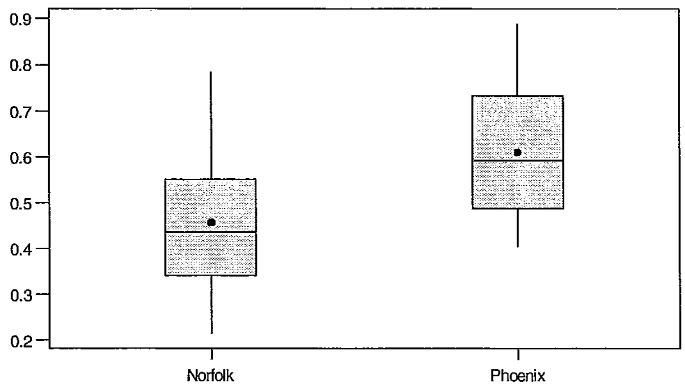
	N	Mean	StDev	SE Mean
Norfolk	24	0.457	0.143	0.029
Phoenix	12	0.610	0.147	0.042

95% CI for mu Norfolk - mu Phoenix: (-0.260, -0.046)

T-Test mu Norfolk = mu Phoenix (vs not =): T= -2.98 P=**0.0072** DF= 21

Boxplots of Norfolk and Phoenix

(means are indicated by solid circles)



reject H_o, a statistical difference does exist

Now try to see if there is a difference between subgroups

Unavailability of Call Takers (cont'd)

Note$_1$: You Must First Unstack the Data to Perform the 2-Sample t-Test

To Unstack the Data

Manip>Stack/Unstack>Unstack Block

Note$_2$: Name the new columns first: C15 - Memphis, C16 - Norfolk, C17 - Phoenix

The Data Window should now look like the one below:

MINITAB - Crprjda1.mtw - [Data]

File Edit Manip Calc Stat Graph Editor Window Help Six Sigma

↓	C9-T	C10	C11	C12	C13	C14	C15	C16	C17	C
		group	cls mtg	cls aux	cls un		Memphis	Norfolk	Phoenix	
1		1	0.024000	0.304000	0.328000		0.44760	0.50970	0.63810	
2		1	0.061000	0.290000	0.351000		0.54060	0.32200	0.65150	
3	C	1	0.077000	0.245000	0.322000		0.57780	0.26980	0.60750	
4	L	1	0.056000	0.243000	0.299000		0.53120	0.37820	0.47450	
5	O	1	0.037000	0.194000	0.231000		0.54270	0.35870	0.52980	
6	S	1	0.010000	0.222000	0.232000		0.81410	0.33250	0.75900	
7	E	1	0.061000	0.250000	0.311000		0.91770	0.29700	0.88860	
8		1	0.030000	0.234000	0.264000		0.97580	0.26860	0.80680	
9	D	1	0.036000	0.228000	0.264000		1.00000	0.21450	0.49920	
10	A	1	0.048000	0.249000	0.297000		0.29750	0.43420	0.48350	
11	T	1	0.011000	0.245000	0.256000		0.32660	0.55310	0.57580	
12	A	1	0.023000	0.256000	0.279000		0.25130	0.56090	0.40400	
13		1	0.009000	0.213000	0.222000		0.15370	0.51640		
14	T	2	0.043300	0.262600	0.305900		0.74740	0.54480		
15	O	2	0.038100	0.267600	0.305700		0.77760	0.40760		
16		2	0.036700	0.227400	0.264100		0.84330	0.42770		
17	T	2	0.055700	0.233800	0.289500		0.32690	0.42320		
18	H	2	0.031500	0.263300	0.294800		0.38830	0.44240		
19	E	2	0.029700	0.269500	0.299200		0.30690	0.52630		
20		2	0.036200	0.316300	0.352500		0.37620	0.78300		
21	R	2	0.083100	0.275000	0.358100		0.62370	0.69140		

11:09 AM

Démarrer | Boîte de récepti... | Martel, Bernard... | Microsoft Power... | MINITAB - C... | MINITAB Help | 11:10

3. Fixtures 1,2,4,6,7, & 8 to Fixture 3

What are the Hypotheses?

H_o:Fixture 1,2,4,6,7, & 8 = Fixture 3

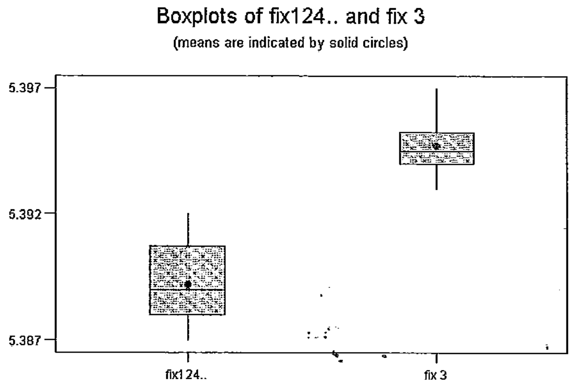
H_a:Fixture 1,2,3,4,6,7,& 8 ≠ Fixture 3

What is the rejection criteria?

p ≥ .05, Fail to reject H_o

p < .05, reject H_o

What conclusion can be drawn?



Two Sample T-Test and Confidence Interval

```
Two sample T for fix124.. vs fix 3
          N      Mean     StDev   SE Mean
fix124..  60  5.38920   0.00140   0.00018
fix 3     10  5.39470   0.00116   0.00037

95% CI for mu fix124.. - mu fix 3: ( -0.00643, -0.00457)
T-Test mu fix124.. = mu fix 3 (vs not =): T= -11.75  P=0.0000  DF=  68
Both use Pooled StDev = 0.00137
```

The means of the Fixtures are Statistically Different

4. Fixtures 1,2,4,6,7, & 8 to Fixture 5

What are the Hypotheses?

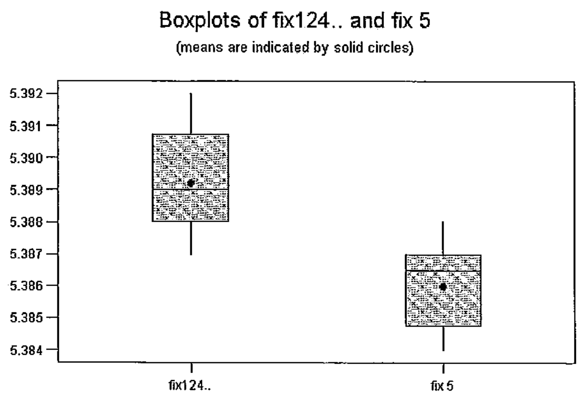
H_o:Fixture 1,2,4,6,7, & 8 = Fixture 5

H_a:Fixture 1,2,3,4,6,7,& 8 ≠ Fixture 5

What is the rejection criteria?

p ≧ .05, Fail to reject H_o

p< .05, reject H_o



What conclusion can be drawn?

Two Sample T-Test and Confidence Interval

```
Two sample T for fix124.. vs fix 5
          N      Mean     StDev   SE Mean
fix124..  60  5.38920   0.00140   0.00018
fix 5     10  5.38600   0.00141   0.00045

95% CI for mu fix124.. - mu fix 5: ( 0.00224, 0.00416)
T-Test mu fix124.. = mu fix 5 (vs not =): T= 6.68  P=0.0000  DF=  68
Both use Pooled StDev = 0.00140
```

The means of the Fixtures are Statistically Different

Class Exercise Answers: Page 55

Perform a 2-Sample t-Test to Compare Fixture-to-Fixture or Fixtures to Fixtures

1. Fixture 1 to Fixture 5

What are the Hypotheses?

H_o:

H_a:

What is the rejection criteria?

What conclusion can be drawn?

2. Fixture 3 to Fixture 5

What are the Hypotheses?

H_o:

H_a:

What is the rejection criteria?

What conclusion can be drawn?

3. Fixtures 1,2,4,6,7, & 8 to Fixture 3

What are the Hypotheses?

H_o:

H_a:

What is the rejection criteria?

What conclusion can be drawn?

4. Fixtures 1,2,4,6,7, & 8 to Fixture 5

What are the Hypotheses?

H_o:

H_a:

What is the rejection criteria?

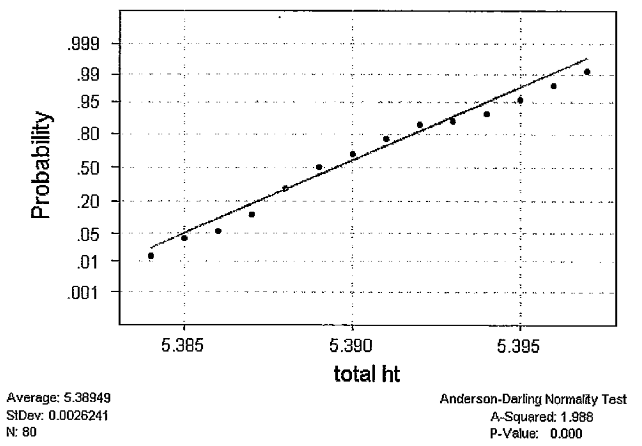
What conclusion can be drawn?

Normality Test for the Lower Transmission Housing Example - page 19

Stat>Basic Statistics>Normality Test

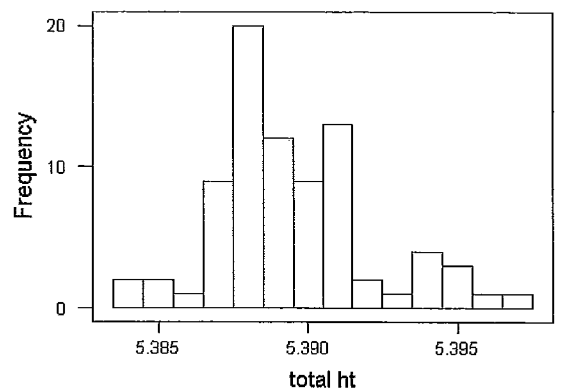
H_o: The data is normal
H_a: The data is non-normal

Normal Probability Plot



P value is < .05; reject H_o

Note: Section on Minitab (Phase 1) explains how to handle non-normal data

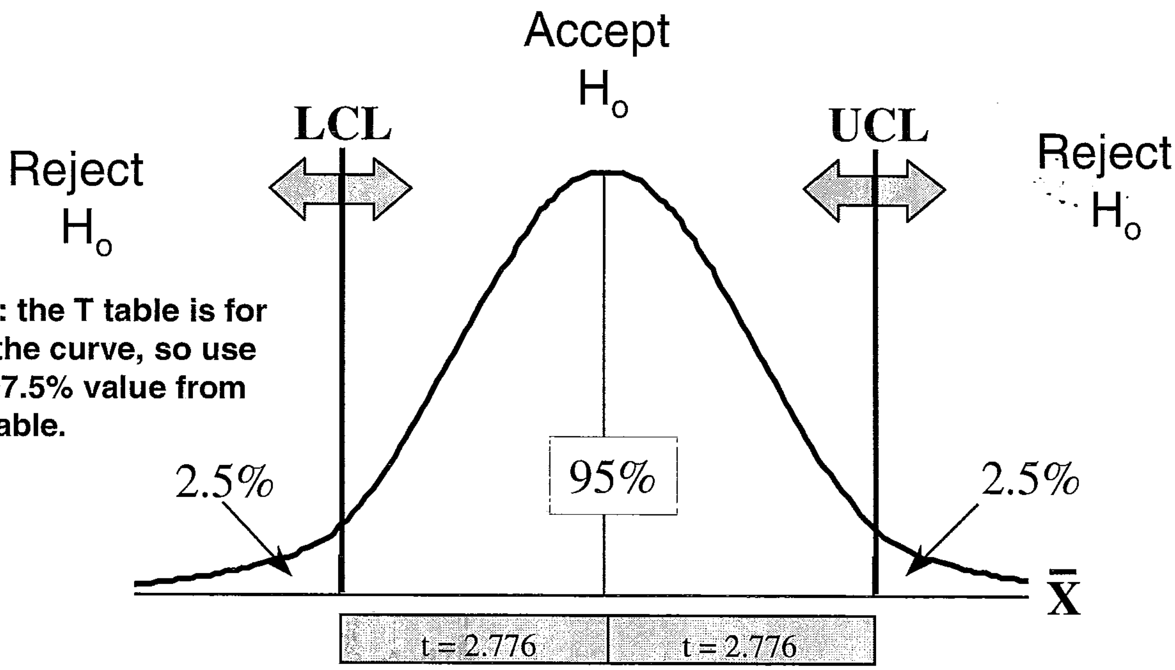


Distribution looks like it may be bi-modal. If so, it can't be transformed into normal data

The data is non-normal

Two-Sided Use of the t Distribution

Distribution of Sampling Averages



Note: the T table is for half the curve, so use the 97.5% value from the table.

95% Confidence Interval

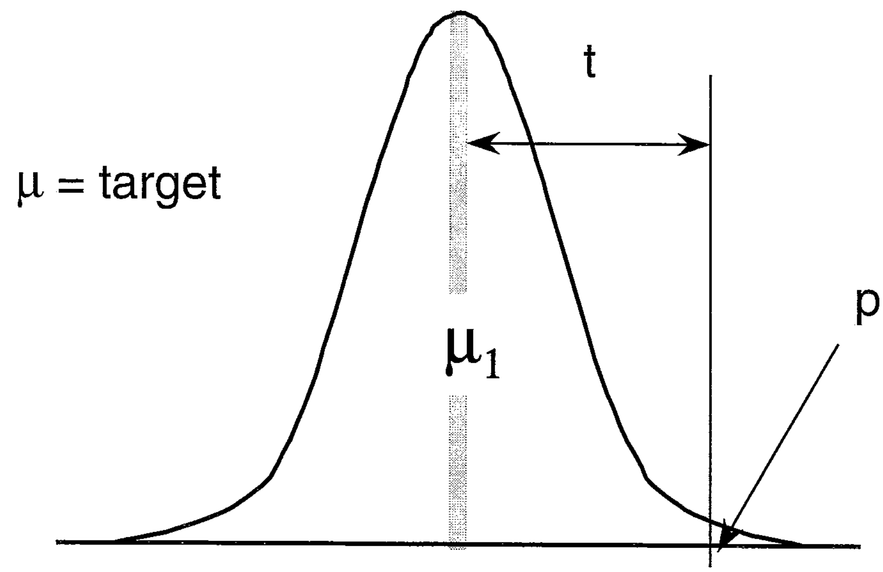
$$LCL = \bar{X} - t_{\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \qquad UCL = \bar{X} + t_{\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}$$

$$df = 4$$

$$LCL = \bar{X} - 2.776\frac{\hat{\sigma}}{\sqrt{5}} \qquad UCL = \bar{X} + 2.776\frac{\hat{\sigma}}{\sqrt{5}}$$

There is 95% certainty that the true universe mean will be contained within the given confidence interval. If we observe a sampling average greater than UCL or less than LCL, we may conclude that such an event could only occur 5 times out of 100 by random chance (sampling variations).

t Test: Testing a Single Mean



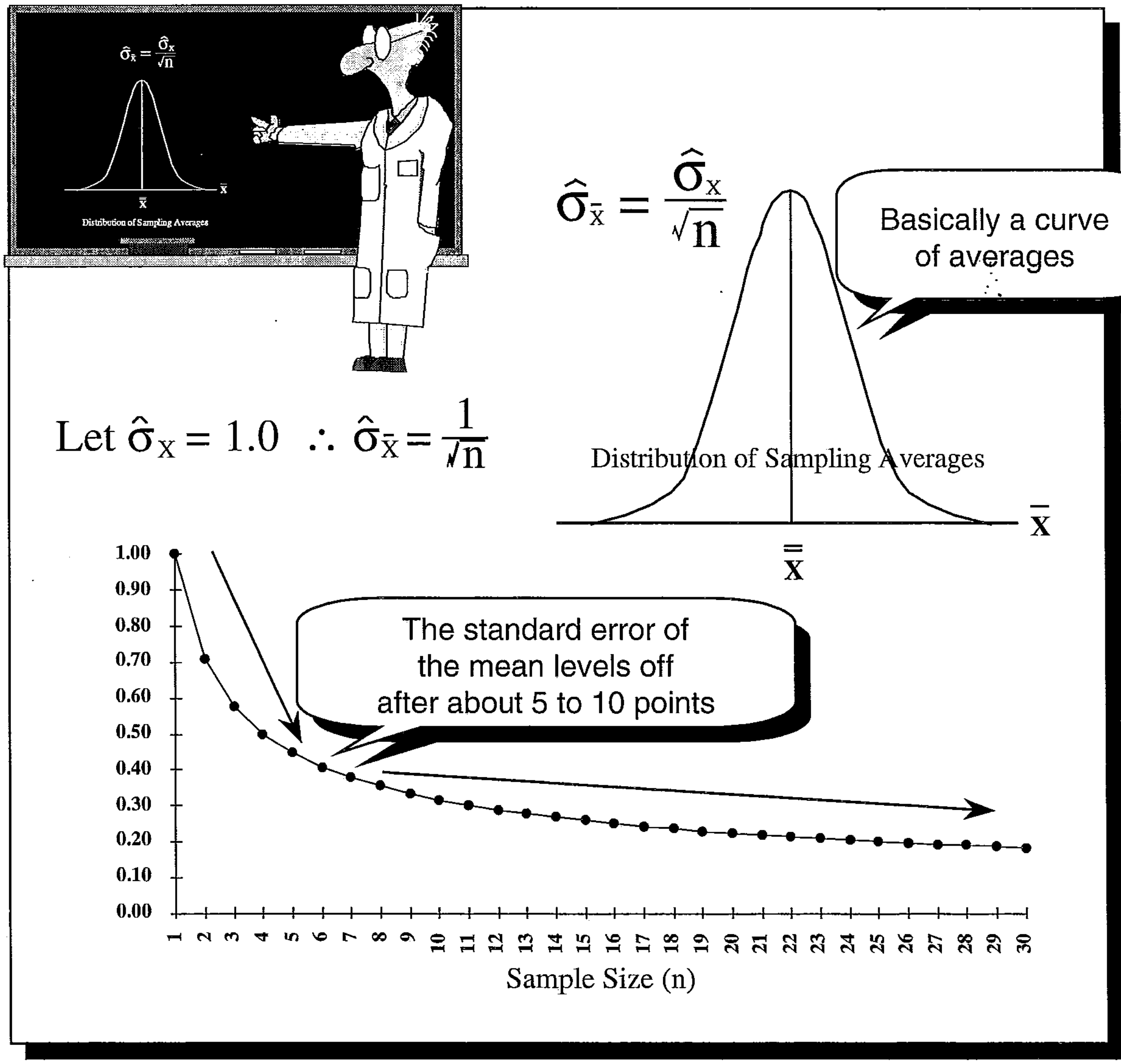
The larger the t, the smaller the area in the tail, p. The probability of being wrong when we claim there is a difference is less and less with smaller p values. [Smaller p-values = greater confidence that there IS a difference.]

Comparing Means: t-Test

Requirements

- Needs to be based on normal or close-to-normal data.
- If two groups are being compared, their variances (standard deviations) must be close and their quantity must be similar.
- Good for small sample sizes (less than 30) as it has a more forgiving distribution.
- Uses the Standard Error of the Mean.

The Standard Error of the Mean



The standard error of the mean is the spread of the population we are using for a particular test. Given a sample set of data, we can calculate a mean and a standard deviation. The standard deviation divided by the square root of sample size gives the standard error of the mean (this division adjusts the curve for sample size). This is the only data we need to draw the normal curve for this data.

Homogeneity of Variance

Interpreting the Results

There are two important issues to note:

- When using the F-test, a larger number indicates a difference because the F-value is used. In contrast, when using the Bartlett's or Levene's tests, a smaller number indicates difference because the probability statistic is used (p value).

- The test indicates only **difference**, not goodness or badness.

Homogeneity of Variance

Performing the Test

- The F test can be used to test **two** distributions
- Minitab will not perform the F test directly. It uses Bartlett's Test for Normal Data. However, the calculation is not difficult to perform on a calculator. The formula is:

$$F = \sigma_1^2/\sigma_2^2$$

where:

σ_1^2 = variance of one distribution (the larger of the 2)

σ_2^2 = variance of the other distribution (the smaller of the 2)

  - **<u>NOTE</u>**: "F" is a statistic, and it has its own distribution.

- Compare the calculated value to a tabled value - $F_{critical}$. The critical F value is selected from an F table by using n -1 (degrees of freedom) for the numerator and denominator. If the calculated value is greater than the tabled value (F_{crit}), then we reject the null hypothesis (H_o) and say that the two distributions <u>are</u> different.

(Reference pg. 15)

Comments on Hypotheses

Hypotheses represent the translation of a practical question into a statistical question. In this manner, the "real world" problem is represented in terms which are suitable for scientific examination and testing. In essence, hypotheses are statements related to the parameters of a given probability distribution; e.g., the mean and/or variance. In other words, hypotheses are statements which allow us to represent all possible outcomes prior to conducting an investigation. Following the statistical investigation, we simply accept or reject each hypothesis which, in turn, provides a solid foundation for making practical "real world" decisions.

When stated in the null form, the general meaning of the hypothesis is associated with the distribution of chance events. This particular hypothesis is most often referred to as the "null hypothesis" and designated as "H_0." Often, it is called the "straw man hypothesis." Its meaning is simple--the universe parameters under investigation are equal; i.e., there is no difference in the universe with respect to the parameter of interest (mean and/or variance.)

In direct contrast to the null hypothesis is the alternate hypothesis (H_a). In this form, such hypotheses are generally associated with distributions <u>other than</u> chance and, as such, are said to be "statistically significantly different" from the chance distribution. This means that any observed difference in the sample parameters under investigation could not have resulted by chance variations inherent to the sample. If the observed difference in the sample can not be explained by chance variations, then we conclude that sample membership is, in one or more respects, different from the qualifications necessary for membership in the universe of interest. Thus, we accept the alternate hypothesis of inequality and conclude the sample was drawn from a universe other than the one under investigation.

When accepting or rejecting the null and alternate hypotheses, we do so with a known degree of risk and confidence. To do this, we specify (in advance of the investigation) the magnitude of decision risk (α,β) and test sensitivity (δ/σ) which is acceptable. Once this has been accomplished, we have the information necessary to determine a "rational" sample size. Mathematical equations do exist for this purpose; however, we must balance such computations against the practical limitations of cost, time, and available resources in order to arrive at a "rational" sampling plan.

15 Steps in Testing for Differences

1. Define the problem

2. State the objectives

3. Establish the Hypotheses*
   - State the Null Hypothesis (H_o): The elements are the same **(Assume the defendant is innocent)**
   - State the Alternative Hypothesis (H_a): Something is **different (what needs to be proved)**

     (Must find evidence that the defendant is guilty)

4. Decide on appropriate statistical test (assumed probability distribution, t, F, or χ^2).

5. State the level of Alpha (α) risk (usually 5%)

Risk of finding an innocent person guilty; rejecting good parts

6. State the level of Beta (β) risk (usually 10-20%)

Risk of finding a guilty person innocent; accepting bad parts

These six elements <u>must</u> be defined before the test for differences is performed

Key Concepts:
Testing for Differences - Continuous Data

- H_o: Things are the same
  Status Quo

- H_a: Things are different

- 1-Tailed test is looking to prove whether one sample set is better than or worse than a second sample set.(1-tailed test: "less than" or "greater than")

- 2-Tailed test is looking to prove that the sample set is different from a target or a second sample set. (2-tailed test: "not equal to")

- F-test compares variances of different sample sets.
  In Minitab, the F-test is called "Homogeneity of Variance".

- Bartlett's Test compares variances of normal data.

- Levene's Test compares variances of non-normal data.

- t-test compares means of a sample set to a target or to other sample sets.

- 1 sample t-test compares a sample mean to a target mean.

- 2 Sample t-test compares a sample mean to a second sample mean.

- Paired t-test compares the difference of the mean of two sample sets of the same data through two operations.

Summarizing the Concepts . . . .

α error: claiming something is different when it's really the same (rejecting good parts at the assembly line)

α risk: the risk of making an α error

- The rule of thumb for the level of α risk you are willing to accept is usually 5% (or $\alpha = 0.05$)

p value*:* the probability of making an α error

☆ If the probability of making an α error ("p") is *less than* the risk I'm willing to take ("α"), then I accept Ha (reject H_o).

☆ If the probability of making an α error ("p") is *greater than* the risk I'm willing to take ("α"), then I reject Ha (fail to reject H_o).

Why? Because I can't PROVE there is a difference with 95% confidence! (I haven't "risen above the fog" of my initial assumption, H_o)

Class Exercise

Problem Statement

Consumer Relations only answers 61% of calls received. Surveys indicate that 34% of consumers find it difficult to get through. Call center performance varies widely and associate unavailability averages 54.5%, which contributes to low answer rate and consumer dissatisfaction.

CTQ: C.R. Availability
Defect: Time Unavailable > 35%
Unit: Each Day, M-F
Opportunity: Each Day, M-F
Customer: Consumer Calling

Measurements

Gage R&R:	14%	
DPMO:		749,594
Yield:		25.04%
Estimated Baseline σ :		**0.83**

Open worksheet file CRPRJDA1
C:\Mtbwin\Data\Ph_1data\CRPRJDA1.mtw

Use the data in columns C2 & C6 to compare the three regions to each other in regards to unavailability.

Apply what you have learned to this point to determine the potential Vital "X"s

Perform a 2-Sample t-Test to Compare the Means of the Calipers

Two Sample T-Test and Confidence Interval

```
Two sample T for calip1 vs calip2
         N      Mean     StDev   SE Mean
calip1  12   0.26625   0.00122   0.00035
calip2  12   0.26600   0.00176   0.00051

95% CI for mu calip1 - mu calip2: ( -0.00103,  0.00153)
T-Test mu calip1 = mu calip2 (vs not =): T= 0.41  P=0.69  DF=  22
Both use Pooled StDev = 0.00151
```

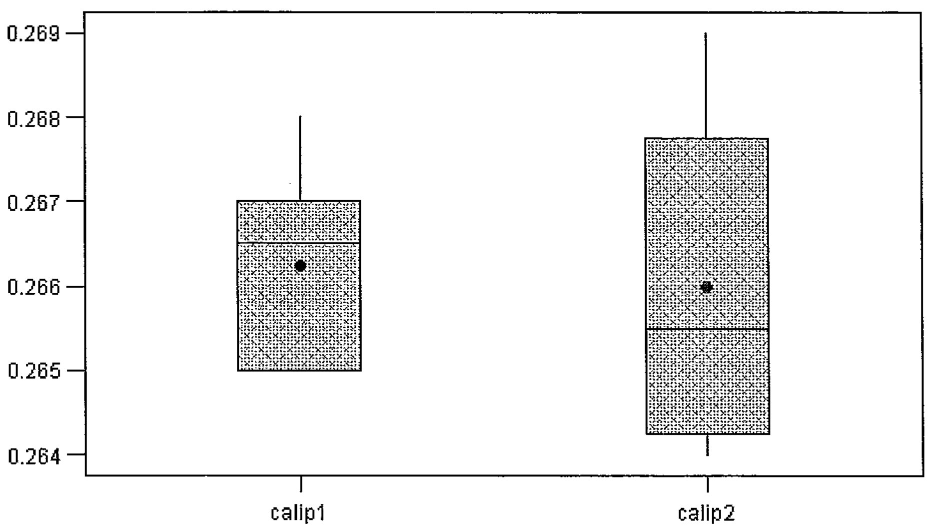
0 is within Confidence Interval; accept H_o

The Hypotheses for H_o and H_a

p value > .05; accept H_o

Boxplots of calip1 and calip2

(means are indicated by solid circles)



Same Result We Found with 1-Sample t-Test: We cannot say the means are different

Next, Compare the Means by using 1-Sample t-Test

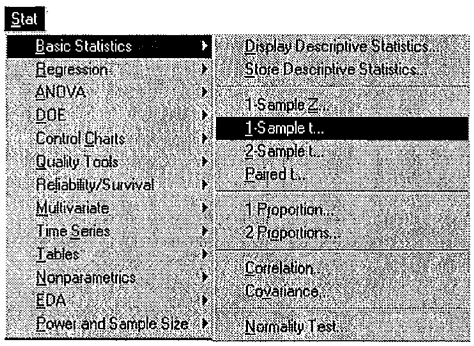
What are the Hypotheses?

H_o:

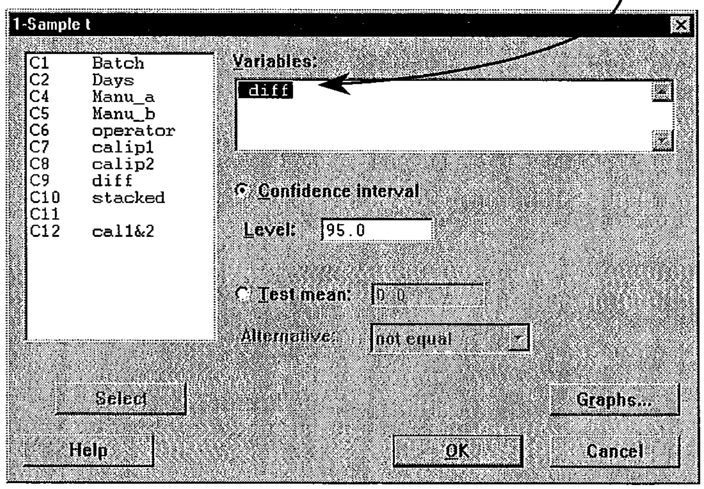
H_a:

What is the rejection criteria?

Stat>Basic Statistics>1-Sample t



Must use the difference between 2 measurements for 1 sample t



T Confidence Intervals

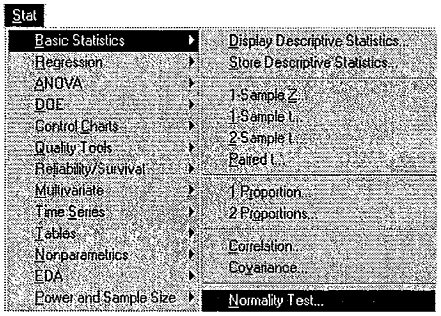
"0" does fall within the C.I.; accept H_o

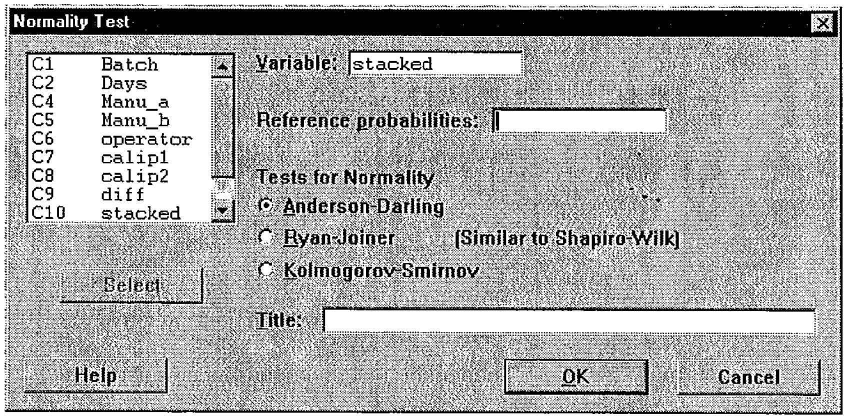
Variable	N	Mean	StDev	SE Mean	95.0 % CI
diff	12	0.00025	0.00201	0.00058	(-0.00102, 0.00152)

We Cannot Say the Means of the Calipers are Different

Check for Normality

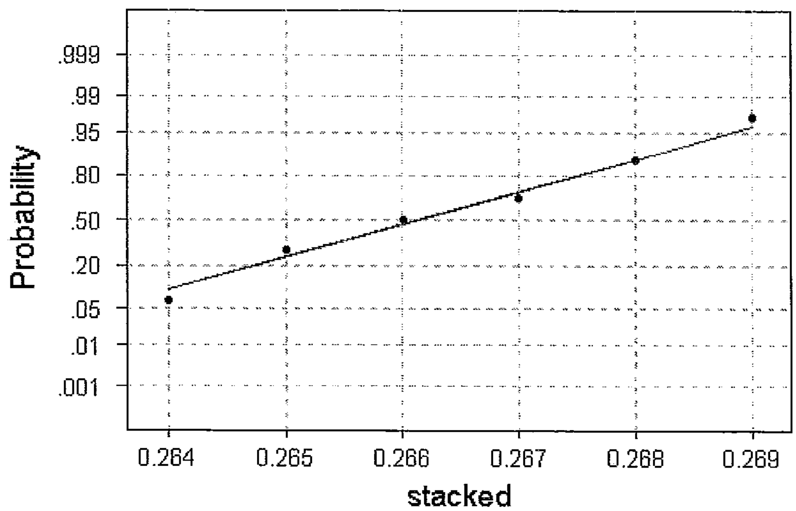
Stat>Basic Statistics>Normality Test





H_o: Data is normal
H_a: Data is non-normal

Normal Probability Plot



Average: 0.266125
StDev: 0.0014836
N: 24

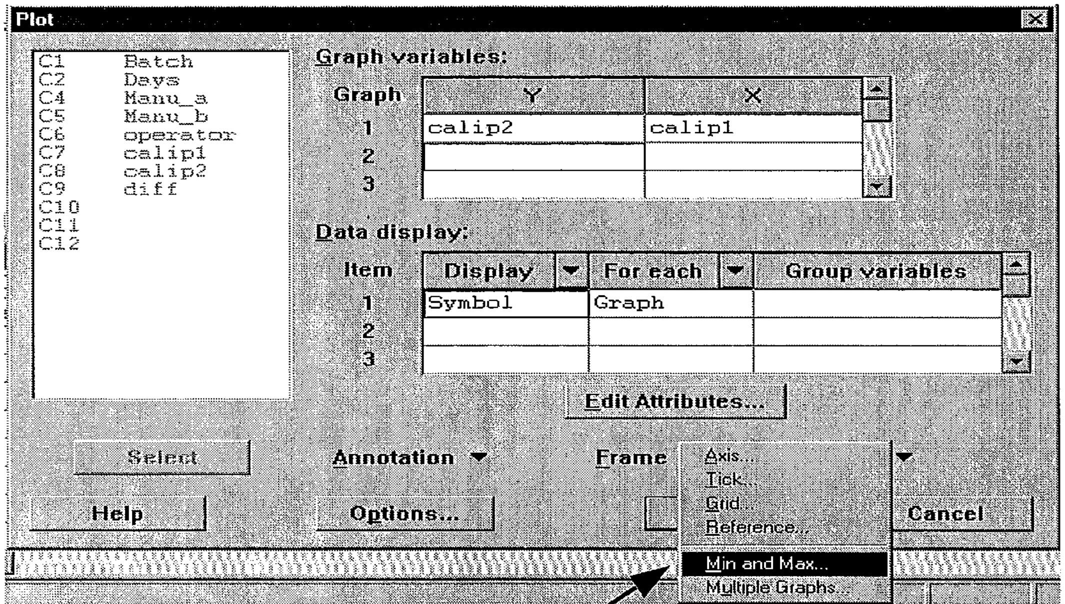
Anderson-Darling Normality Test
A-Squared: 1.013
P-Value: 0.009

p value < .05; reject H_o

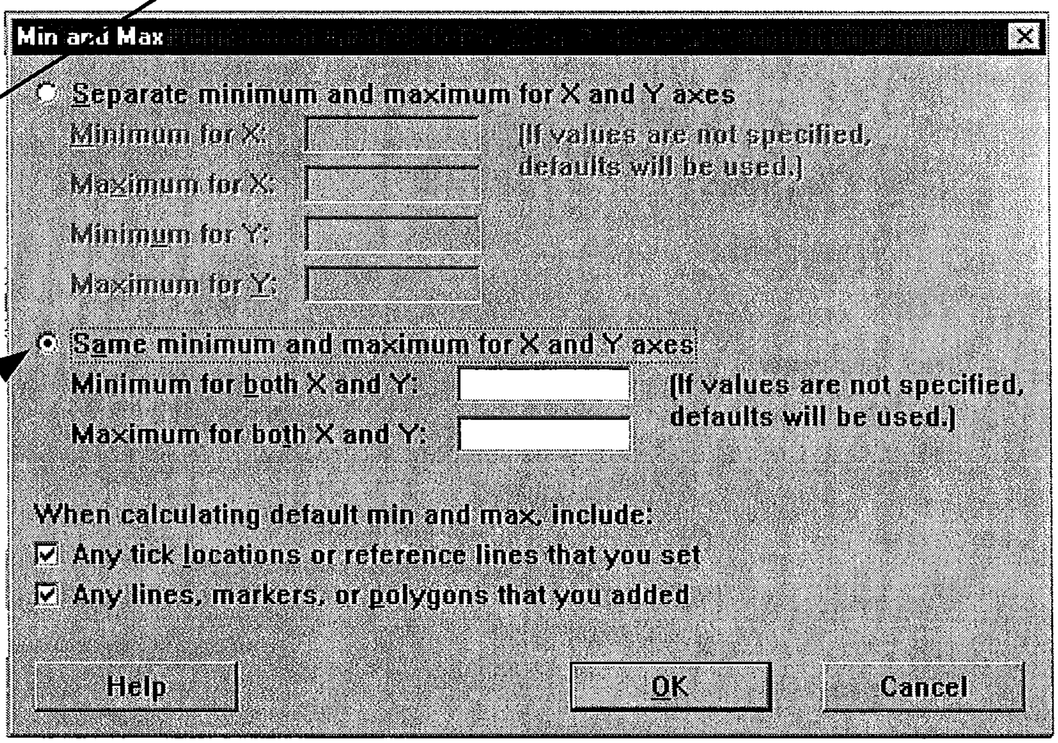
Conclusion: Data is Non-Normal

Graph the data first

Scatter plot of calip2 versus calip1



Click Frame and Min/Max to show identical X and Y axis ranges



Your Worksheet should look like the Worksheet below:

MINITAB - TTESTS.MTW - [Data]

File Edit Manip Calc Stat Graph Editor Window Help Six Sigma

	C6	C7	C8	C9	C10	C11	C12
↓	operator	calip1	calip2	diff	stacked		Cal1&2
1	1	0.265000	0.264000	0.0010000	0.265000	1	1
2	2	0.265000	0.265000	0.0000000	0.265000	2	1
3	3	0.266000	0.264000	0.0020000	0.266000	3	1
4	4	0.267000	0.266000	0.0010000	0.267000	4	1
5	5	0.267000	0.267000	0.0000000	0.267000	5	1
6	6	0.265000	0.268000	-0.0030000	0.265000	6	1
7	7	0.267000	0.264000	0.0030000	0.267000	7	1
8	8	0.267000	0.265000	0.0020000	0.267000	8	1
9	9	0.265000	0.265000	0.0000000	0.265000	9	1
10	10	0.268000	0.267000	0.0010000	0.268000	10	1
11	11	0.268000	0.268000	0.0000000	0.268000	11	1
12	12	0.265000	0.269000	-0.0040000	0.265000	12	1
13					0.264000	1	2
14					0.265000	2	2
15					0.264000	3	2
16					0.266000	4	2
17					0.267000	5	2
18					0.268000	6	2
19					0.264000	7	2

What test should be run first?

What are the Hypotheses?

H_o:

H_a:

What is the rejection criteria?

Taking t-Test One Step Further..... Testing Paired Samples (Paired t-Test)

Why use it?

- To block out excess variability
- To determine if multiple processes affect the response

When should it be used?

- When taking two measurements of the same response
- When each row of data has a mate

  Examples: - The flatness of a lid before and after paint

  - The performance of a call- taker before training and after training

Hypotheses

H_o: $\mu_1 = T$
H_a: $\mu_1 \neq T$
H_a: $\mu_1 < T$
H_a: $\mu_1 > T$

For a 1-sample t-test in Minitab, take the difference of the two measurements, enter the difference into a new column

H_o: $\mu_1 - \mu_2 = 0$
H_a: $\mu_1 - \mu_2 \neq 0$
H_a: $\mu_1 - \mu_2 < 0$
H_a: $\mu_1 - \mu_2 > 0$

For a 2- sample t-test in Minitab, enter the two measurements in separate columns.

Results of the 2-Sample t-Test to Compare Fixtures 1,2,4,6,7, & 8 to Fixtures 3 & 5

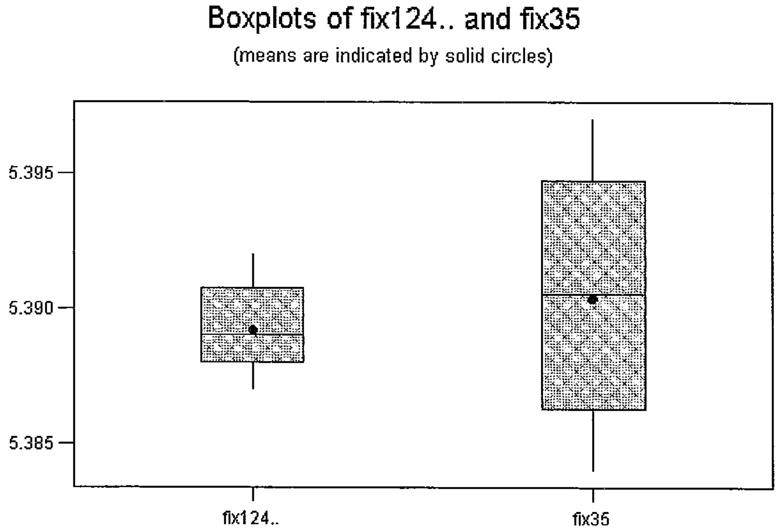
Two Sample T-Test and Confidence Interval

```
Two sample T for fix124.. vs fix35

            N      Mean     StDev   SE Mean
fix124..   60   5.38920   0.00140   0.00018
fix35      20   5.39035   0.00464    0.0010

95% CI for mu fix124.. - mu fix35: ( -0.00248,  0.0002)
T-Test mu fix124.. = mu fix35 (vs not =): T = -1.72  P = 0.090  DF = 78
Both use Pooled StDev = 0.00259
```

Since 0 falls inside the C.I.; accept H_o

The Hypotheses for H_o and H_a

Since p value is > .05; accept H_o



No statistical conclusion can be drawn!
We may need to take more data to see a difference

Stacked Data Appears in the Data Window

Fixtures 1,2,4,6,7,8 are all stacked in column C14
Fixtures 3 & 5 are all stacked in column C17

MINITAB - LTH.MTW - [Data]

File Edit Manip Calc Stat Graph Editor Window Help Six Sigma

	C14	C15	C16	C17	C18	C19	C20
↓	fix124..			fix35			
1	5.39000	1		5.39400	1		
2	5.38900	1		5.39400	1		
3	5.39000	1		5.39300	1		
4	5.38900	1		5.39400	1		
5	5.38800	1		5.39400	1		
6	5.39100	1		5.39500	1		
7	5.39100	1		5.39600	1		
8	5.39100	1		5.39700	1		
9	5.39100	1		5.39500	1		
10	5.38900	1		5.39500	1		
11	5.38700	2		5.38600	2		
12	5.38700	2		5.38400	2		
13	5.38700	2		5.38500	2		
14	5.38700	2		5.38500	2		
15	5.38800	2		5.38400	2		
16	5.38800	2		5.38700	2		
17	5.38900	2		5.38800	2		
18	5.38900	2		5.38700	2		
19	5.38800	2		5.38700	2		

What are the Hypotheses:

Ho:

Ha:

What is the rejection criteria?

Minitab Gives a Confidence Interval and a Hypothesis Test

Two Sample T-Test and Confidence Interval

```
Two sample T for fix 1 vs fix 3
        N      Mean     StDev   SE Mean
fix 1  10   5.38990   0.00110   0.00035
fix 3  10   5.39470   0.00116   0.00037

95% CI for mu fix 1 - mu fix 3: ( -0.00586,  -0.00374)
T-Test mu fix 1 = mu fix 3 (vs not =): T= -9.50  P=0.0000  DF=  18
Both use Pooled StDev = 0.00113
```

Confidence interval (does not include 0.0)

The Hypotheses for H_o and H_a

p value is < .05; reject H_o

Interpretation of the Hypothesis test

- The ***p value is less than .05***, so we reject Ho and conclude than the difference between the 2 fixtures is real, and not due to chance alone.
- In fact, we have ***greater than 99.9% confidence*** that the difference is real.
- The absolute value of the calculated t (|-9.50| = 9.50) is greater than the tabled value (2.101) with 18 degrees of freedom, indicating that the difference is statistically significant.

There is a Statistical Difference between the means of Fixture 1 and Fixture 3. Fixture is a critical X!

Minitab Output - Comparison of Fixture 1 to Fixture 3 (2-Sample t-Test)

Two Sample T-Test and Confidence Interval

Confidence interval for difference of the 2 means does not include 0.0

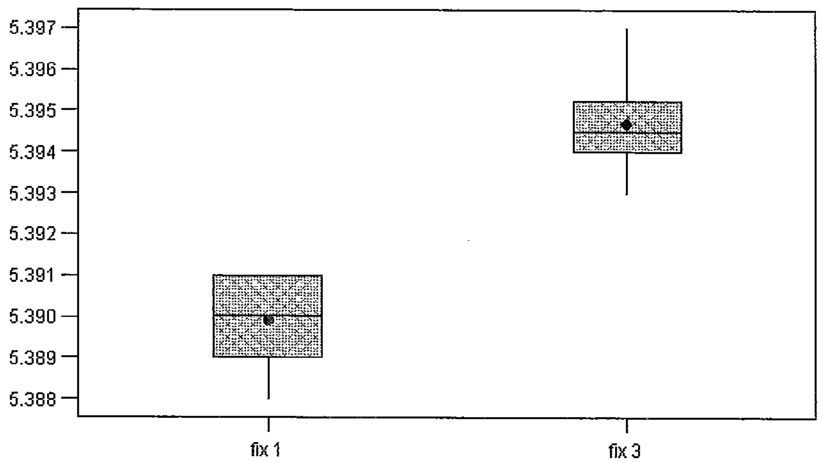
```
Two sample T for fix 1 vs fix 3

        N      Mean     StDev   SE Mean
fix 1  10   5.38990   0.00110   0.00035
fix 3  10   5.39470   0.00116   0.00037

95% CI for mu fix 1 - mu fix 3: ( -0.00587,  -0.00373)
T-Test mu fix 1 = mu fix 3 (vs not =): T = -9.50  P = 0.0000  DF = 17
```

p value is < .05; reject H_o

The Hypotheses for H_o and H_a

Boxplots of fix 1 and fix 3

(means are indicated by solid circles)



Class Exercise:

Perform a 1-Sample t-Test to Compare Fixtures 2 & 5 to the Target Mean

1. Fixture 2 compared to the target mean of 5.394"

   What are the Hypotheses?

   H_o:

   H_a:

   What is the rejection criteria?

   What conclusion can be drawn?

2. Fixture 5 compared to the target mean of 5.394"

   What are the Hypotheses?

   H_o:

   H_a:

   What is the rejection criteria?

   What conclusion can be drawn?

1-Sample t-Test for Fixture 1 to the Target Mean

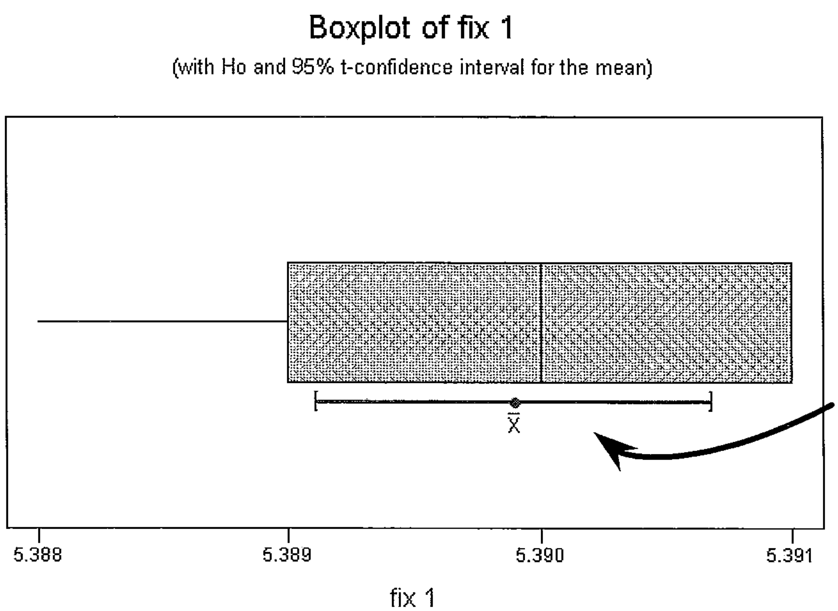
The Hypotheses for H_o and H_a

T-Test of the Mean

```
Test of mu = 5.39400 vs mu not = 5.39400
Variable        N      Mean     StDev    SE Mean        T        P
fix 1          10   5.38990   0.00110    0.00035   -11.78   0.0000
```

p value < .05; reject H_o

Graph



Confidence Interval indicates the range that the mean may fall within

There is a statistical difference between Fixture 1 and the Target Mean!

The Data Window will have unstacked fixture data in columns C5 - C12

MINITAB - LTH.MTW - [Data]

File Edit Manip Calc Stat Graph Editor Window Help Six Sigma

	C5	C6	C7	C8	C9	C10	C11	C
↓	fix 1	fix 2	fix 3	fix 4	fix 5	fix 6	fix 7	x 8
1	5.39000	5.38700	5.39400	5.38800	5.38600	5.38800	5.38800	5.
2	5.38900	5.38700	5.39400	5.38900	5.38400	5.38800	5.38900	5.
3	5.39000	5.38700	5.39300	5.38800	5.38500	5.38800	5.38800	5.
4	5.38900	5.38700	5.39400	5.39000	5.38500	5.38800	5.38800	5.
5	5.38800	5.38800	5.39400	5.38900	5.38400	5.38800	5.38800	5.
6	5.39100	5.38800	5.39500	5.39200	5.38700	5.39100	5.39100	5.
7	5.39100	5.38900	5.39600	5.39100	5.38800	5.39100	5.39200	5.
8	5.39100	5.38900	5.39700	5.39100	5.38700	5.39100	5.39100	5.
9	5.39100	5.38800	5.39500	5.39100	5.38700	5.39000	5.38900	5.
10	5.38900	5.38700	5.39500	5.39000	5.38700	5.39000	5.38900	5.
11								
12								
13								
14								
15								
16								
17								
18								
19								

Now, run a 1-Sample t-Test or 2-Sample t-Test . . .

How to Use the t-table to Draw a Conclusion (Calculated t-value vs Table Value)

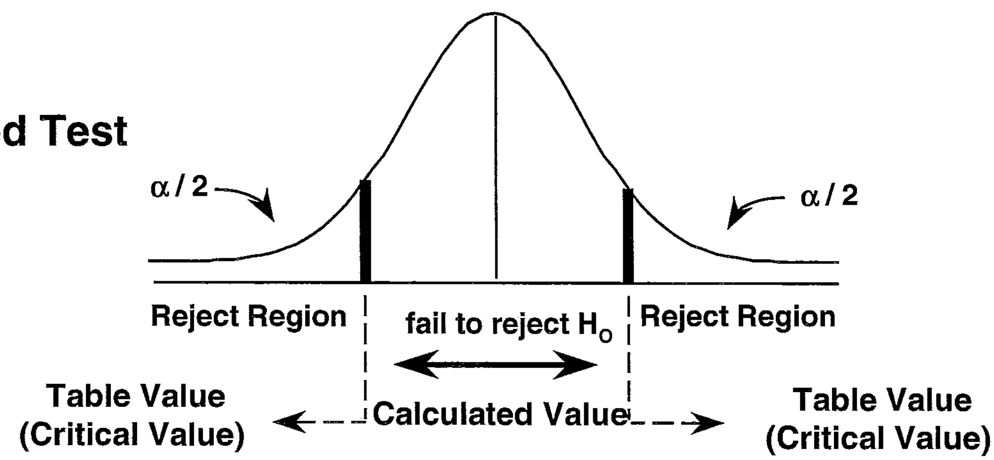
Using the previous example, use the Table Value to accept or to reject H_o. The sample has an $\bar{x}$ ($\hat{\mu}$) of 5.38990", a σ of .00110 and a sample size of 10. The Target μ is 5.3940". Is the sample mean equal to the target mean? Note: This is a 2-tailed test.

H_o: $\hat{\mu}_1 = T$
H_a: $\hat{\mu}_1 \neq T$
N = 10
Degrees of Freedom = 9; (10 - 1)
α = .05
The **Calculated Value = -11.78**

$$t = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

Rejection Region is the table value of +/- $\alpha/2$
(table value = +/ - 2.262)

2-Tailed Test



-11.78 < -2.262; reject H_o

There is a statistical difference between the Sample Mean and the Target Mean

Effect of α on the Confidence Interval

Calculate the 90% Confidence Interval for the mean of the height of the transmission housings made in fixture 3. Use n = 10. ($\bar{x}$ = 5.3947, s = 0.00116)

Calculate the 99% Confidence Interval for the mean of the height of the transmission housings made in fixture 3. Use n = 10. ($\bar{x}$ = 5.3947, s = 0.00116)

What effect does α have on the confidence interval?

Effect of Sample Size on the Confidence Interval

Let's take 20 more samples (total n=30) and see what happens to the 95% confidence interval. Suppose the average and standard deviation stay the same: $\bar{x} = 5.3947$ and s = 0.00116.

$$\text{Lower Confidence Limit} = \bar{x} - t_{(\alpha/2,\ df)} \frac{s}{\sqrt{n}}$$

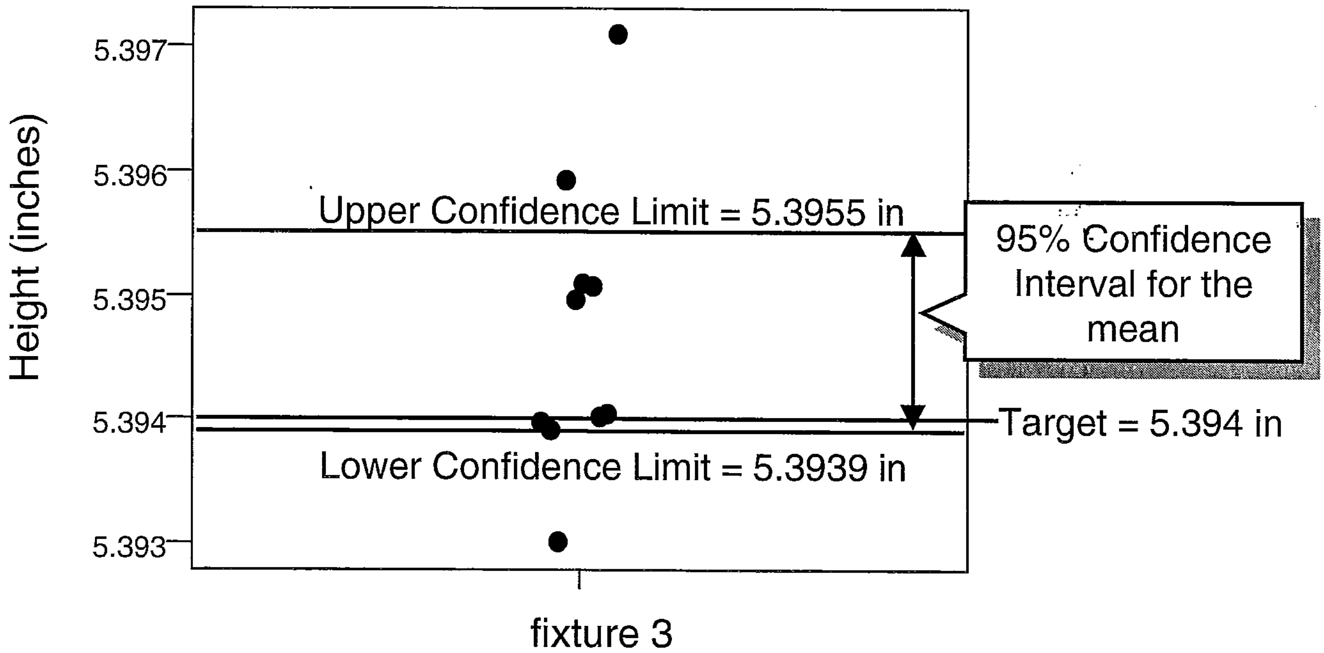
$$= \boxed{} - \boxed{} \left(\frac{\boxed{}}{\sqrt{\boxed{}}} \right) = \boxed{}$$

$$\text{Upper Confidence Limit} = \bar{x} + t_{(\alpha/2,\ df)} \frac{s}{\sqrt{n}}$$

$$= \boxed{} + \boxed{} \left(\frac{\boxed{}}{\sqrt{\boxed{}}} \right) = \boxed{}$$

Example - continued

Is the average of the parts made in fixture 3 on target?



The most-likely estimate for the population average of fixture 3 is 5.3947, but the true value may be somewhat larger or smaller than this.

The range of plausible values for the population average of fixture 3 is 5.3939 to 5.3955.

Confidence Interval

Calculate the Confidence Interval for the mean of the height of the transmission housings made in fixture 3. Use $\alpha = 0.05$ (95% Confidence Interval).

$\bar{x} = 5.3947$

$s = 0.00116$

$n = 10$

$df = n - 1 = 9$

$t_{(\alpha/2,df)}$ is obtained from the t table.

$t_{(0.025,9)} = 2.262$

$$\text{Lower Confidence Limit} = \bar{x} - t_{(\alpha/2,\ df)}\frac{s}{\sqrt{n}}$$

$$= \boxed{5.3947} - \boxed{2.262}\left(\frac{\boxed{0.00116}}{\sqrt{\boxed{10}}}\right) = \boxed{}$$

8.29x10^-4

5.3938

$$\text{Upper Confidence Limit} = \bar{x} + t_{(\alpha/2,\ df)}\frac{s}{\sqrt{n}}$$

5.3955

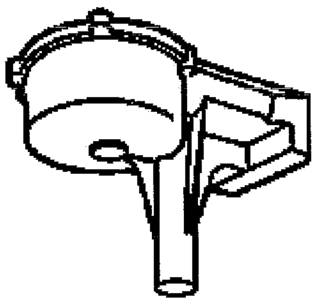
$$= \boxed{} + \boxed{}\left(\frac{\boxed{}}{\sqrt{\boxed{}}}\right) = \boxed{}$$

Example

The total height of the machined lower transmission housing on the washing machine affects brake performance. **The project Y is *Total Height***, with a target = 5.394". There are 8 fixtures that hold the part for machining.

What do you want to know?

Is the average of the parts made in fixture 3 on target?



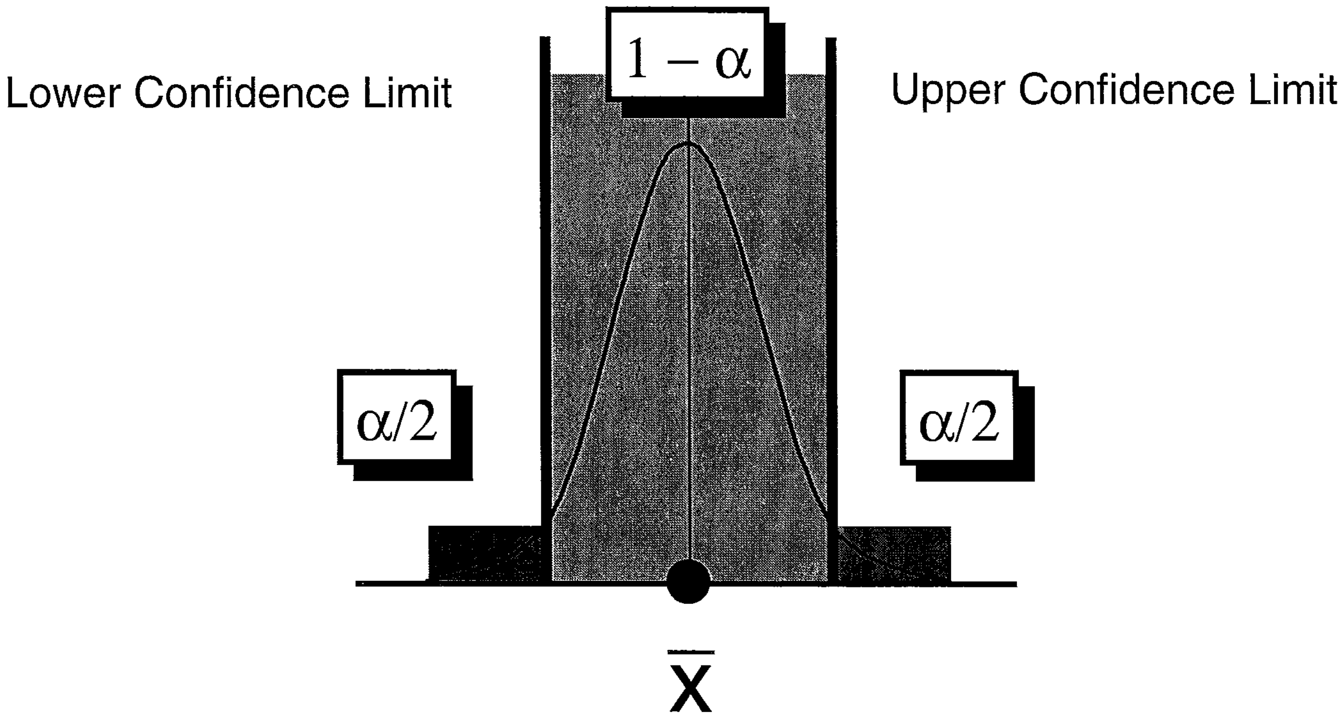
Steps in the analysis:

1. Graph the data.
2. Use hypothesis tests and confidence intervals to determine if the observed difference is real.
3. Draw conclusions.

Confidence Interval : 1-sample t

What is the range of plausible values for the mean of the population to be on target? Let's calculate the Confidence Interval to find out!

Confidence Interval for a Single Mean



There is (1-α)100% Confidence that the true population mean is contained within the Confidence Interval.

How to Perform the t-Test...

t- Test calculated by Hand

$$t_{calc} = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}\ ^{*}$$

where

$\bar{X}$ = the observed value
μ = the expected value
σ = the standard deviation of the observed
n = the sample size of the observed distribution

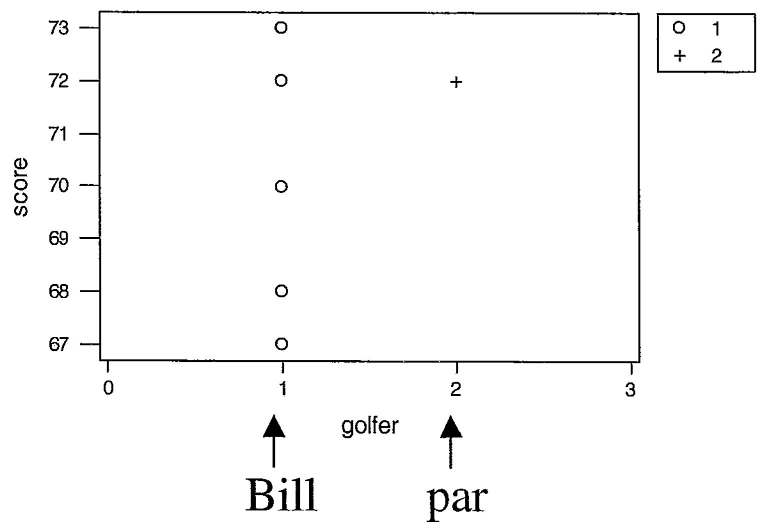
Minitab Method

Stat>Basic Statistics>

- **1-sample t** Compares an observed mean to a target
  Compares a paired distribution (Paired t - test)

Hypotheses

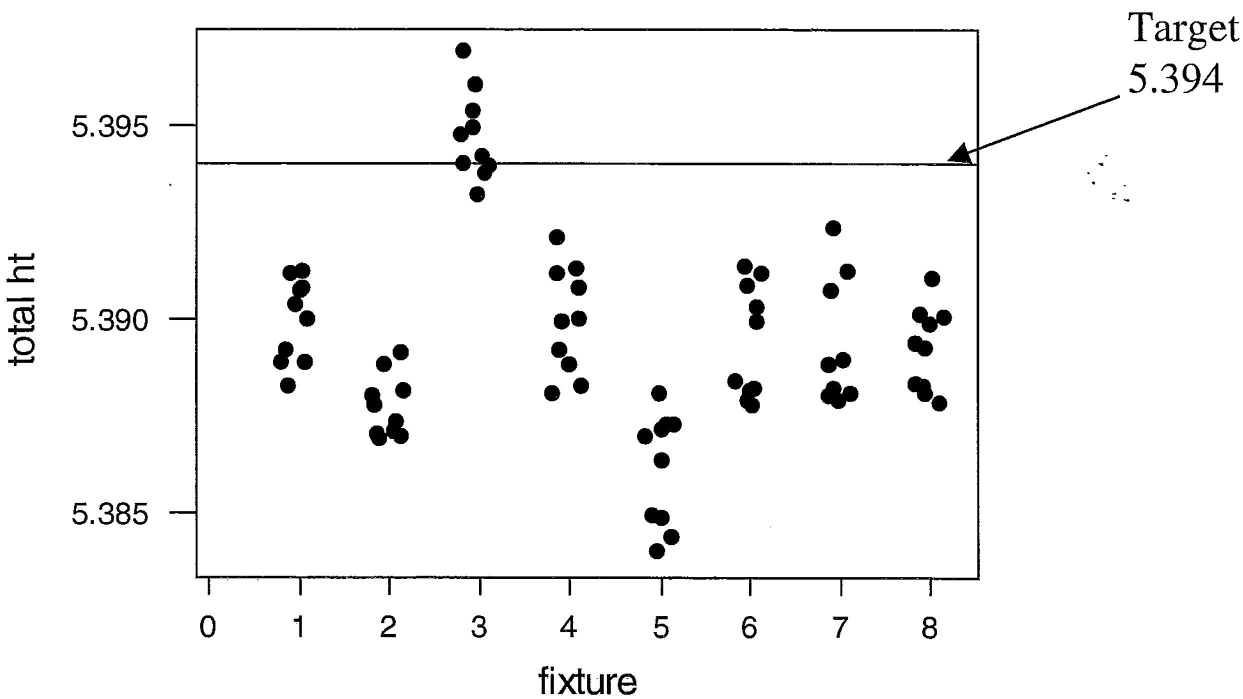
$H_o = \mu_1 = T$
$H_a = \mu_1 \neq T$
$H_a = \mu_1 < T$
$H_a = \mu_1 > T$



Fail to reject H_o when $p \geq .05$; Accept H_a when $p < .05$

*See Appendix

Now let's switch to Averages

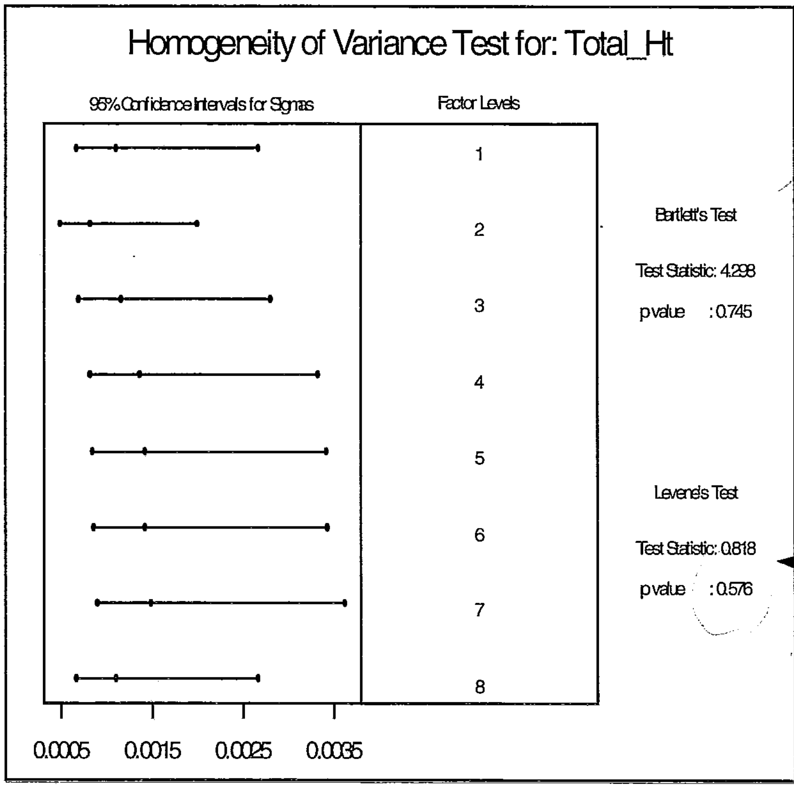


What do you want to know?

- **Statistical question** -- Are the apparent differences in averages between the fixtures real, or could they be due to chance alone?

- **Practical question** -- Should we put effort into making all fixtures like #3, in order to make them closer to the target?

Homogeneity of Variance Test



p value is > .05, fail to reject Ho.

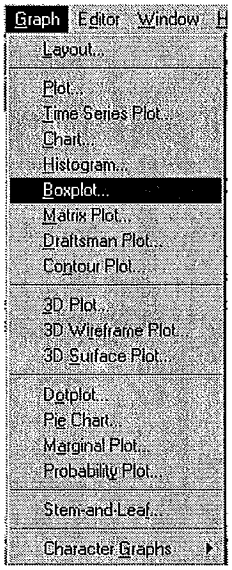
None of the fixtures can be proven different, since all the Confidence Intervals overlap

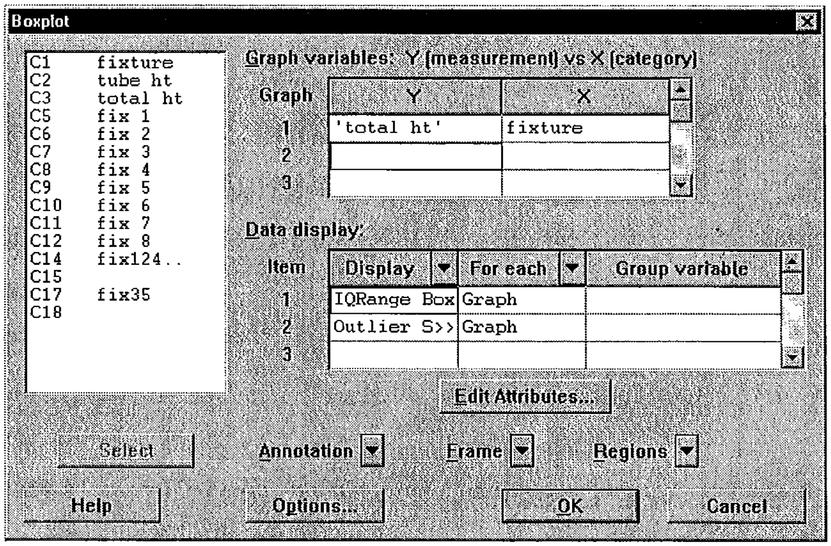
Use the Levene's Test since the data was not normal.

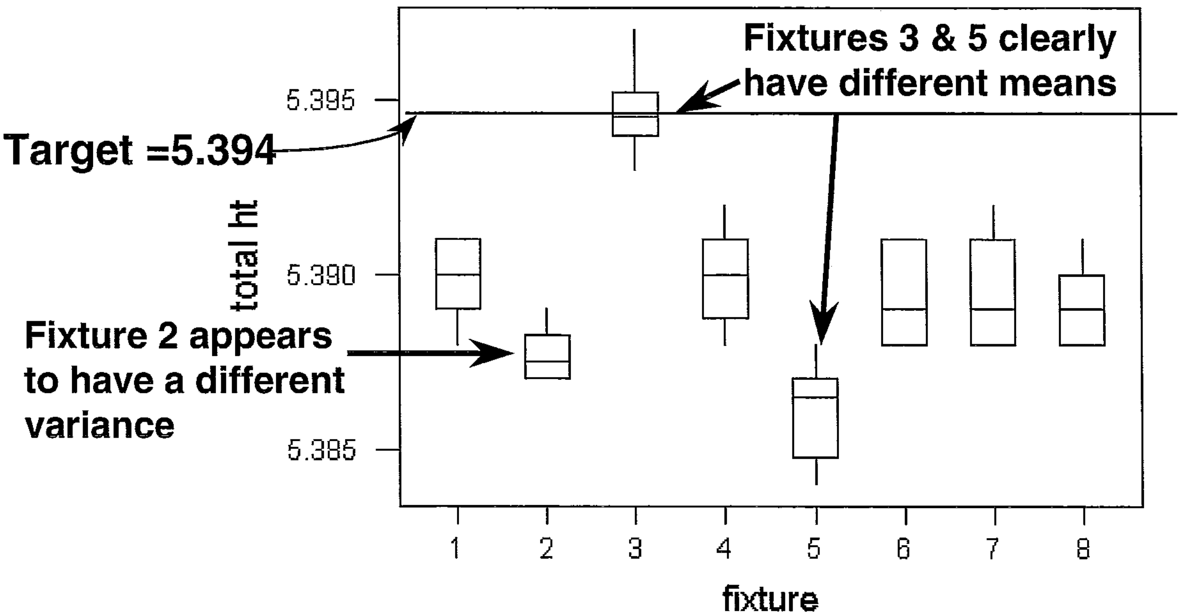
Statistical conclusion: We cannot conclude that the variation is different between the 8 fixtures. The differences we observed in the graph may be due to chance.
Practical conclusion: It would not be wise to try to make all fixtures like fixture 2 in order to reduce variation

Box Plot can give a picture of fixture to fixture difference

Graph>Boxplot







Graphs led to an answer, **Hypothesis Testing** gives **Statistical Confidence** to the answer

F Table

Denominator degrees of freedom

Numerator degrees of freedom

Numerator Degrees of Freedom

Denom	1	2	3	4	5	6	7	8	9	10
1	161.40	199.50	215.70	224.60	230.20	234.00	236.80	238.90	240.50	241.90
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83

How to Perform the F-Test...

F - Test calculated by Hand

$F = \sigma_1^2/\sigma_2^2$ Where σ_1^2 = the variance of one distribution (the larger of the two)

and σ_2^2 = the variance of the other distribution (the smaller of the two)

Minitab Method

Homogeneity of Variance

Stat>ANOVA>Homogeneity of Variance

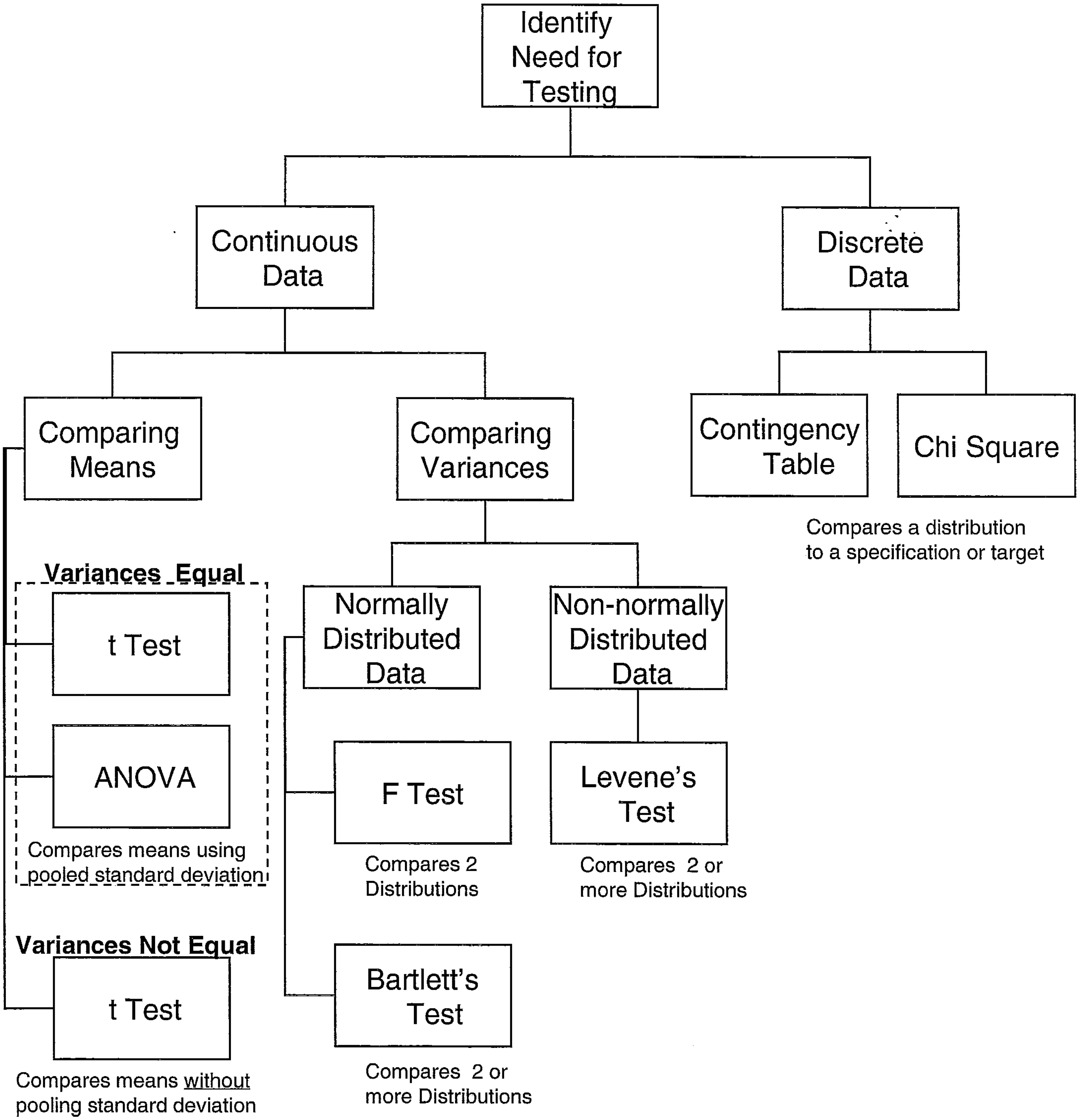
- Bartlett's Test - Normal Data
- Levene's Test - Non-Normal Data

Hypotheses

H_o : $\sigma_1^2 = \sigma_2^2$ Variances are equal when $p \geq \alpha$ (Fail to reject H_o)

H_a : $\sigma_1^2 \neq \sigma_2^2$ Variances not equal when $p < \alpha$ (reject Ho)

Summary- Statistical Testing Tools



Defining the Hypotheses: H_o and H_a

H_o

The starting point for a hypothesis test is the "null" hypothesis - H_o. H_o is the hypothesis of ***sameness***, or ***no difference***.

Example: The population mean equals the test mean.

H_a

The second hypothesis is H_a - the "alternative" hypothesis. It represents the hypothesis of ***difference***.

Example: The population mean does not equal the test mean.

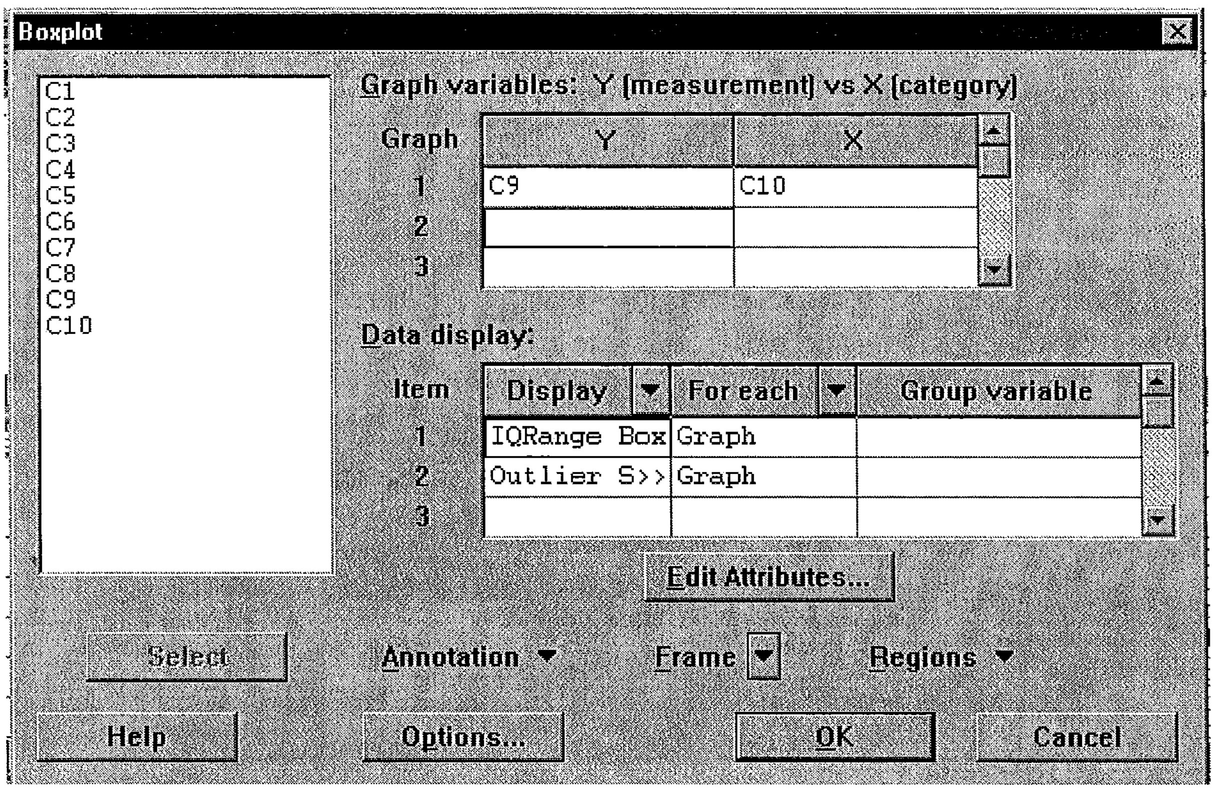
- **You usually want to show that there is a difference (H_a).**
- **Start by assuming equality (H_o).**
- **If the data show they are not equal, then they must be different (H_a).**

Why Do We Use Hypothesis Tests and Confidence Intervals?

3. Graph the data and look for differences.

>Graph >Boxplot

Graph c9 (Y) versus c10 (X, fixture).



Why Do We Use Hypothesis Tests and Confidence Intervals?

Since there is variation in all data, we expect to see small differences in the sample data, even if the populations are the same. Let's see what happens when we generate some random data, with no real patterns.

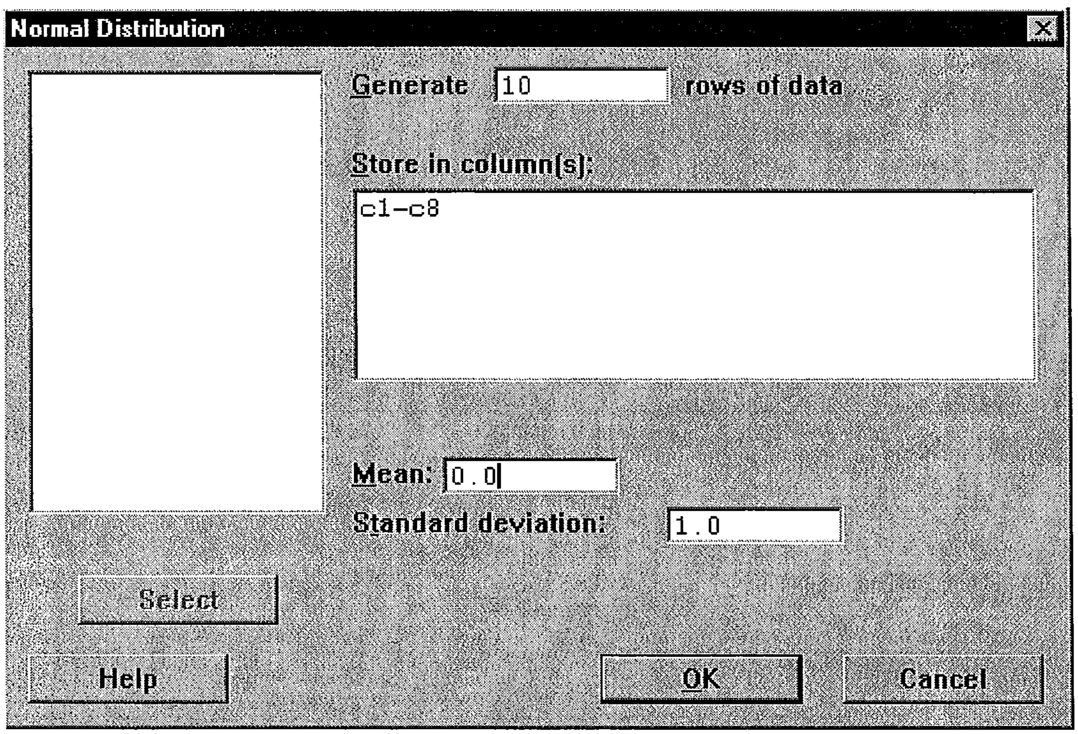
1. **Generate 8 sets of random data.**

   Open a new worksheet:

   > File > New . . . > Minitab Worksheet > OK

   Generate 10 rows of data. Store in columns c1-c8.

   >Calc >Random data > Normal



The Gap... (Observed - Expected)

How much of a difference (δ) is important to you?

- **Statistical Significance**
  - A difference that can be detected by use of statistical tools at a stated level of confidence

- **Practical Significance**
  - A difference that is important enough for you to take action

You must decide if a <u>statistically</u> significant difference is <u>practically</u> significant...Use the Hypothesis tests and Confidence intervals to determine if the difference is real!

Statistical Analysis

Observed - Expected

- In EVERY statistical test, you will be comparing some **observed** value to some **expected** value by comparing estimates (mean, standard deviation, variance)

- These estimates of the **TRUE** parameters are obtained by using sample data

- The ability to detect a **difference** between what is observed and what is expected depends on the **spread** of the sample data

- **Increasing** the **sample size improves** the **estimate** and your confidence in the statistical conclusions

but....

there ***is*** a difference between **statistical** and **practical** significance!

Testing for Differences - Continuous Data

PURPOSE:

The purpose of this tab is to introduce the confidence intervals and the two basic types of Hypothesis Tests for continuous data: **t-Test and F-Test**. Pre-requisites for application and rules of usage will be discussed.

OBJECTIVES:

- Understand the difference between practical and statistical significance
- Define the hypotheses and understand the three methods for acceptance/rejection of hypotheses
- Understand the rationale for hypothesis tests and confidence intervals
- Understand the concept of 1-tailed and 2-tailed hypothesis tests
- Apply F-test (comparison of variances) and t-Test (comparison of means) to real-life project examples

Let's learn how to tell if the differences we measure are statistically significant!