The Data Driven Manager

4

Two Sample Dependent Tests

Learning Objectives

- Discern between samples that are independent and dependent
- Perform a statistical test for differences in means (dependent groups) with repeated measures
- Perform a statistical test for differences in means (dependent groups) with matched pairs
- Perform a statistical test for differences in variances (dependent groups) when underlying distribution is normal

Learning Objectives

- Perform a statistical test for differences in variances (dependent groups) when underlying distribution is not normal
- Perform a statistical test for differences in proportions (dependent groups)
- Perform a statistical test for differences in counts (dependent groups)

Introduction to Dependent Tests

Basic Concepts

- Dependency between two sets (groups) of scores can be
  - Dependent by nature
  - Dependent by design

- The difference is important because it permits flexibility in the analyses if certain assumptions are not met

Dependent By Nature

- Repeated measures on the same items
  - May occur as a result of “experimental manipulations”
  - The same specimens may be measured at different points in time
  - The dependency in this case must be “honored” in the analysis of the data

Dependent By Design

- Results from “linking” the items in two independent groups
- Linking (the dependency) results from
  - Matching on some basis
    - Criterion measure matching
    - Covariate measure matching
  - Pairing (without matching), also called blocking
  - Subdividing or taking portions of a homogeneous unit (similar to repeated measures)

Two Sample Statistical Tests

Parameters / Condition	Means	Variances	Proportions	Counts (Poisson)
Independent	• Two Sample t-test for Independent Measures • Two Sample Approximate t-test	• F - test • Levene and ADM(n-1)	• Two Sample Exact Binomial Test (Fisher's Exact Test)	• Two Sample Poisson Test
Dependent	• Repeated Measures t-test • Matched Pairs t-test	• Matched Pairs t-test for Variances	• McNemar's Dependent Proportions Test	• Wilcoxon Signed Ranks Test

Two Dependent Sample Tests

- Repeated Measures t test for Means
- Paired t-test for Means
- Matched Pair t test for Variances
- McNemar's Test of Change
- Wilcoxon Signed Ranks

Two Sample Hypothesis Tests for **Means**

Dependent Conditions

Repeated Measures

t test for the Means

Repeated Measures t Test for Means

- Used to compare means of paired groups
- Test the following hypothesis:

$$H_0: \mu_1 = \mu_2$$

or

$$H_0: \mu_D = 0$$

Repeated Measures
t Test for Means

- Test Statistic

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{n} - 2r\frac{s_1}{\sqrt{n}}\frac{s_2}{\sqrt{n}}}} = \frac{\bar{D}}{s_d/\sqrt{n}}$$

Where n is the sample size for the **pairs** of scores

Repeated Measures
t Test for Means

- Underlying Assumptions
  - The n pairs of scores are independent of one another. This assumption is critical and is satisfied by randomly sampling the items
  - The population for difference scores is normally distributed, as are the populations for each group

Example - Repeated Measures

- A Black Belt is attempting to solve a noise level problem with a type of wiper motor.
- They suspect that a major source of noise problems may be traced to the material from which the bearing is made.
- In an effort to determine whether the average noise level of the motors can be decreased by changing to a new bearing material, they randomly select ten motors from the current production line.

Example - Repeated Measures

- A noise level reading is taken on each of the ten motors.
- Next, the bearings are replaced with bearings made from the new material, and retests the motors.
- The noise level for the each of the modified motors is recorded as was previously done.

Example - Repeated Measures

- All testing is conducted under the same essential conditions as for the first set of measurements. The data were collected and entered into the file **Noise.txt**.

- Test an appropriate hypothesis to determine whether it is reasonable to assume that the new bearing material will effectively reduce the initial noise level of the motors if implemented across the process. Use a significance level of 0.05 and set this up as a directional test.

Example - Repeated Measures

In RStudio

```
t.test.twosample.dependent( )
```

View the Full Hypothesis Testing Procedure here:

https://tinyurl.com/2smpdepttestrepmeas

ISO Plot

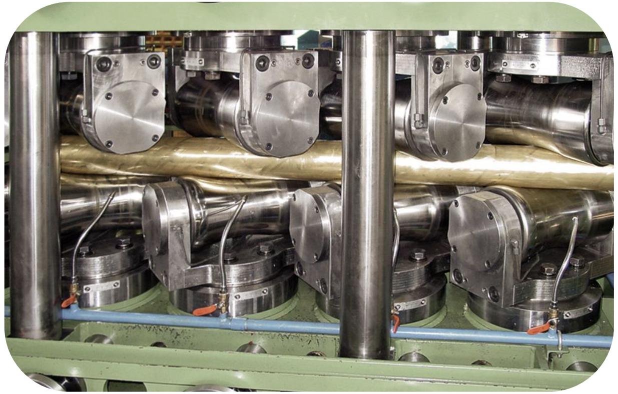
- Paired analysis should be accompanied by an iso-plot (Plot a Y=X line, Slope =1)

Iso Line Interpretation

- If the two variables are identical, the plotted points will fall exactly on the X = Y line.
- If one group has an average that is different from the other (but the variances are equal), the plotted points will fall on a straight line which is offset from, either above or below, the X = Y iso-line.
- If the two groups have different variances, the plotted points will have a slope different from the slope of 1.

Your Turn!

- A steel bar straightening machine has been in use for 15 years. When the machine was purchased (prior to the implementation of quality improvement policies), it was assumed that it would perform as advertised. That is, that it would significantly improve the straightness of the product.



Your Turn!

- The engineering manager just recently attended data driven management course on Coursera. They became anxious to statistically test whether the straightening operation was changing the process mean for bar straightness.

Your Turn!

- 20 bars were randomly selected from a production lot and identified in a serial fashion. It was then arranged to have an inspector measure the straightness of each bar. Subsequently, the bars were put through the straightening operation. Each of the bars was remeasured for straightness after the operation by the same inspector.
- Because the measurements were repeated on the same parts and the data were paired according to the bar identification, we have data that are dependent by nature.

Your Turn!

- The first column contains the straightness values measured prior to the operation and the second column contains the values measured after the operation. The values in each row represent the same specimen, in this case a bar. A flat bar has a value of zero, a non-flat bar has values above zero. The data file is named **Straight.txt**.

- Has the straightening operation significantly changed the process mean? Has the straightness of the population of bars been improved?

Repeated Measures
t Test for Means

1. The hypothesis is about:
   a. μ
   b. σ^2
   c. π
   d. λ
   e. ρ

2. What is the source of the dependency?
   a. Repeated measures
   b. Matching
   c. Pairing (natural or artificial)
   d. Subdividing a homogenous unit

Repeated Measures
t Test for Means

3. What is the proper test statistic?
    a. Z
    b. t
    c. F
    d. χ^2
    e. Exact

4. What is the **value** of the proper test statistic?
    a. 0.008
    b. -1.715
    c. -0.909
    d. -0.908

Repeated Measures
t Test for Means

5. Do you
   a. Reject the null hypothesis?
   b. Fail to reject the null hypothesis?

Matched Pairs

t test for the Means

Matched Pairs
t Test for Means

- Underlying assumptions
  - The specimens in the two samples are independent (by nature).
  - The population for difference scores is normally distributed, as are the populations for each group
  - Homogeneity of variance is assumed. (Not critical if sample sizes are equal.)
  - The units or specimens in the two samples are dependent by design.

Matched Pairs
t Test for Means

- Test Statistic

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{n} - 2r\frac{s_1}{\sqrt{n}}\frac{s_2}{\sqrt{n}}}}$$

Where n is the sample size for the **pairs of scores**

Example - Matched Pairs

- A production manager wishes to determine whether two secondary blanking presses are producing raw plates with equal average flatness, a smaller number is better.
- It is known that hardness of the material is a variable that can affect the flatness of the plates.
- Therefore, **60 plates** are randomly selected from the primary blanking operation and tested for hardness.

Example - Matched Pairs

- After rank ordering and pairing the plates according to their tested hardness (i.e., lowest with the next lowest, and so on up the measured hardness scale), they randomly assign the plates from each pair to one of the two groups (i.e., one plate from each pair to the group going to press 1 and one plate from each pair to the group going to press 2).

Example - Matched Pairs

- After the assignments are complete, the manager verifies that there are no significant differences between the two groups in terms of plate hardness on the basis of their variances and means. Why did they do this?

Example - Matched Pairs

- Then, the two groups are run concurrently as pairs through the presses, and the resultant flatness data are recorded. The summary statistics for the two groups of data are as follows.

  $mean_1 = 35.24$ $mean_2 = 38.02$

  $s_1 = 5.18$ $s_2 = 5.63$

  $r_{12} = 0.60$
- Test an appropriate hypothesis.

Example - Matched Pairs

In RStudio

```
cor.pearson.r.onesample.simple( )
t.test.twosample.dependent.simple.meandiff( )
```

View the Full Hypothesis Testing Procedure here:

https://tinyurl.com/2smpdepttestmatchpair

Your Turn!

- An engineering manager has been assigned the task of improving the joint strength of a clutch assembly.

- The assembly is comprised of two components: the clutch cup and the drive sprocket. The two components are joined together by means of a brazing operation.

Your Turn!

- The manager wanted to investigate whether changing the brazing material would affect average braze strength.
- They decided they would braze one group with pure copper wire and braze another group with the existing copper alloy.
- One of the two components, the sprocket, is manufactured from powdered metal.

Your Turn!

- They had observed from past studies that there was a strong relationship between the density of the sprocket and the joint strength of the assembly after brazing.
- The greater the porosity of the powdered metal sprocket, the more absorption of the braze material takes place, resulting in a weaker bond.

Your Turn!

- It was decided to use the known relationship between sprocket porosity and joint strength for setting up a matched-pairs design.
- The manager randomly selected 100 sprockets from inventory and had them identified and measured for density.
- They were then rank ordered from lowest to highest and paired according to their density. Each pair of sprockets was randomly assigned to one of two groups that were to receive the different brazing material.

Your Turn!

- Once the random assignments were completed, the engineering manager verified that there was no statistically significant difference between the two groups for density in terms of central tendency and dispersion.

- The two groups were then run in pairs through the braze furnace. The braze strength of the assemblies was tested to failure in the laboratory on an Instron machine.

Your Turn!

Test an appropriate hypothesis. Present your conclusions based upon the results of your analysis. Use an α = 0.05,

Copper Alloy	Pure Copper
$Mean_1$ = 3671 lbs	$Mean_2$ = 4228 lbs
s_1 = 246 lbs	s_2 = 182 lbs
$n_1 = n_2$ = 50 (pairs)	r_{12} = 0.78

Matched Pairs
t Test for Means

1. The hypothesis is about:
   a. μ
   b. σ^2
   c. π
   d. λ
   e. ρ

2. What is the source of the dependency?
   a. Repeated measures
   b. Matching
   c. Pairing (natural or artificial)
   d. Subdividing a homogenous unit

Matched Pairs
t Test for Means

3. What is the proper test statistic?
   a. Z
   b. t
   c. F
   d. χ^2
   e. Exact

4. What is the **value** of the proper test statistic?
   a. 4.839
   b. 3.387
   c. -36.108
   d. -25.532

Matched Pairs
t Test for Means

5. Do you
   a. Reject the null hypothesis?
   b. Fail to reject the null hypothesis?

Two Sample Hypothesis Tests for **Variances**

Dependent Conditions

The Dependent Sample
t Test for Variances

- This test is used to test for differences in variances between two dependent groups
- This test uses a different approach than the F test, which is used for independent groups
- This analysis should be accompanied by an iso-plot which provides a very easy to understand picture of the data used for this test

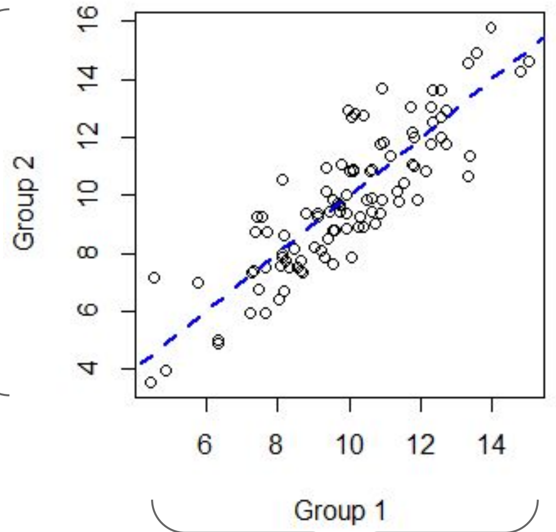
The Dependent Sample t Test for Variances

- Hypotheses: $$\mathrm{H}_0: \sigma_1^2 = \sigma_2^2$$

- Test Statistic $$t = \frac{s_1^2 - s_2^2}{2 s_1 s_2 \sqrt{\frac{1 - r^2}{n - 2}}}$$

The Dependent Sample t Test for Variances

Variation



Variation

Equal variation in both directions in the null case

The Dependent Sample t Test for Variances

- Underlying Assumptions
  - The pairs of scores are independent of one another (critical)
  - The sample data are either dependent by nature, or dependent by design (critical, therefore you may be required to test the correlation)
  - The underlying process distributions are normally distributed (not critical if n is large)

Example t Test for Variances Problem

- It is quite possible that the two blanking presses (see Sample Problem - Dependent by Design) could be producing product which have equivalent means, but are different in terms of dispersion. Assuming that you would like to know whether the processes are different in terms of piece-to-piece variability, test an appropriate hypothesis for that data.

Example t Test for Variances Problem

$\text{mean}_1 = 35.24$ $\quad$ $\text{mean}_2 = 38.02$

$s_1 = 5.18$ $\quad$ $s_2 = 5.63$

$r_{12} = 0.60$

Example t Test for Variances Problem

In RStudio

```
variance.test.twosample.dependent.simple( )
```

View the Full Hypothesis Testing Procedure here:

https://tinyurl.com/2smpdepttestvar

Your Turn!

- A process engineer arranged to measure the out-of-flatness characteristic after the heat and quench operation on 20 randomly selected units.

- The components were then stress relieved with a typical production lot. The same 20 units were re-measured for out of flatness after tempering.

Your Turn!

- The recorded data are stored in a data file called **Temper.txt**.
- Zero indicates no warp. A value greater than zero indicates the presence of warp.
- Use the appropriate hypothesis test to determine whether the fixture and the stress relief operation have effectively reduced out-of-flatness (on the average and variance).

The Dependent Sample t Test for Variances

1. The hypothesis is about:
   a. μ
   b. σ^2
   c. π
   d. λ
   e. ρ
2. What is the source of the dependency?
   a. Repeated measures
   b. Matching
   c. Pairing (natural or artificial)
   d. Subdividing a homogenous unit

The Dependent Sample *t* Test for Variances

3. What is the proper test statistic?
   a. Z
   b. t
   c. F
   d. χ^2
   e. Exact

4. What is the **value** of the proper test statistic?
   a. 6.042
   b. 1.577
   c. 10.541
   d. 2.291

The Dependent Sample t Test for Variances

5. Do you
   a. Reject the null hypothesis?
   b. Fail to reject the null hypothesis?

Two Sample Hypothesis Tests for **Proportions**

Dependent Conditions

McNemar's Test for Change: Dependent Proportions

- Suppose we test 100 randomly-selected units of product, and find that 20% are defective.
- Then, imagine that we apply some type of treatment to the units; and on a post-test, we find again that 20% are defective.

McNemar's Test for Change: Dependent Proportions

- We might be tempted to suppose that no hypothesis test is required under these conditions, in that the 'Before' and 'After' sample proportions are identical, and would surely result in a test statistic value of 0.00 (and a p-value of 1.0 indicating that this could have easily occurred by chance and chance alone).

McNemar's Test for Change: Dependent Proportions

- The problem with this thinking, however, is that the two sample proportions are Dependent because each unit was measured twice.

- Thus, it is possible that the 20 units that were defective originally were still defective. But, it is also possible that the 20 units that were defective on the second test were a completely different set of 20 units! It makes a difference.

McNemar's Test for Change: Dependent **Proportions**

- It is for precisely this type of situation that McNemar's Test for Change in Dependent-Sample Proportions is applicable.
- McNemar's Test employs two unique features for testing the difference between two dependent sample proportions:
  - A special fourfold (2x2) contingency table
  - A special-purpose chi-square ($\chi 2$) test statistic (the original approximate test).
  - The exact test is what we will use, which uses the binomial distribution and p = 0.5

McNemar's Test for Change: Dependent Proportions

- What McNemar discovered is that the test reduces to an assessment of the difference between the cells (either counts or proportions) that change and does not depend on those that stay the same.

McNemar's Test for Change: Dependent Proportions

- To understand how this test works, consider the following simplistic situation and data.
    - 13 items are assessed on one occasion and rated as Pass(0) or Fail(1)
    - The same items, after the application of a "fix", are assessed again as either Pass(0) or Fail(1)

- The data were collected and summarized as follows:

McNemar's Test for Change: Dependent Proportions

Item	M1	M2	**Diff**
1	1	1	0
2	1	0	1
3	1	0	1
4	1	1	0
5	0	1	-1
6	0	0	0
7	0	0	0
8	1	0	1
9	0	1	-1
10	1	0	1
11	0	1	-1
12	0	1	-1
13	1	1	0

		M2	
		0	1
M1	0	No Change	**- Change**
	1	**+ Change**	No Change

		M2	
		0	1
M1	0	A	**B**
	1	**C**	D

		M2	
		0	1
M1	0	2	**4**
	1	**4**	3

McNemar's Test for Change: Dependent Proportions

- Underlying Assumptions:
  - Each observation (cell entry) is independent of every other cell entry (this represents "pairs")
  - Each cell entry represents only one joint event
  - Expected cell frequencies must be 5 or use the Exact Probability option

McNemar’s Test for Change: Dependent **Proportions**

$$H_0: Pass_1 Fail_2 = Fail_1 Pass_2$$

After Condition

Before Condition		Pass	Fail
	Pass	a	**b**
	Fail	**c**	d

where

$(a+b) + (c+d) = (a+c) + (b+d) = n$ = number of pairs of units evaluated and where df = 1

McNemar's Test for Change: Dependent Proportions

- Hypotheses: $Pass_1 Fail_2 = Fail_1 Pass_2$

- Test Statistic $$P(r) = \frac{n!}{(n-r)!\,r!} p^r q^{n-r}$$

McNemar's Test for Change: Dependent Proportions

- An operations manager in a manufacturing plant wishes to determine whether a new maintenance procedure is likely to improve the repeatability of a particular test at a test station.
- They select a random sample of 120 electronically tuned radios, which contain nonconforming as well as conforming units at the same level as daily production levels.

McNemar's Test for Change: Dependent Proportions

- The entire sample is then tested.
- The maintenance procedure is then performed and the test is repeated on the same sample of 120 radios.
- In both tests, the radios are tested in a random order. They are also numbered with a unique identifier so the results of the two tests may be recorded for the proper units. Note that this is a repeated assessment on the same radios.

McNemar's Test for Change: Dependent Proportions

- The summary data resulting from the study appear as follows.

Number of Units	Status Before Maintenance	Status After Maintenance
4	Fail	Fail
4	Pass	Fail
56	Fail	Pass
56	Pass	Pass

McNemar's Test for Change: Dependent Proportions

- Place these data in the proper cells of the 2x2 contingency table before we demonstrate the test.

	Pass	Fail
Pass	56	**4**
Fail	**56**	4

McNemar's Test for Change: Dependent Proportions

Create a vector of the frequencies (counts)	`ct<-(a,c,b,d)`
Create a 2x2 contingency table	`matrix(ct,nrow = 2)`
Perform McNemar's Test	`proportion.test.mcnemar.simple( )`

McNemar’s Test for Change: Dependent Proportions

View the Full Hypothesis Testing Procedure here:

https://tinyurl.com/mcnemartestprop

Your Turn!

- The characteristic “shiny finish” on coated plastic cosmetic containers has been established as a critical-to-quality, CTQ, characteristic.
- The manufacturer of these products has incorporated an extra buffing operation to ensure itself of meeting this customer expectation.
- Some of the components have a metallic chrome finish. The company has assumed that the buffing operation, which follows the chroming operation, naturally improves the shiny appearance of the surface.

Your Turn!

- One employee wanted to statistically test this assumption and was given permission by management to do so.
- 150 cosmetic bases were randomly selected from production after the chroming operation and a visual inspection procedure for judging acceptable and unacceptable shininess was performed.
- Each part was carefully numbered before visually inspecting them in random order.

Your Turn!

- The employee then arranged to have the parts buffed with regular production units.
- The parts were retrieved and were visually inspected in the appropriate manner a second time.
- The results are shown in the table on the next page. Use the appropriate test and procedures to help report the appropriate findings to management.
- Use an α = 0.05.

Your Turn!

Number of Bases	Status *before* Buffing	Status *after* Buffing
10	Reject	Reject
4	Reject	Accept
102	Accept	Accept
34	Accept	Reject

McNemar's Test for Change

1. The hypothesis is about:
   a. μ
   b. σ^2
   c. π
   d. λ
   e. ρ

2. What is the source of the dependency?
   a. Repeated measures
   b. Matching
   c. Pairing (natural or artificial)
   d. Subdividing a homogenous unit

McNemar's Test for Change

3. What is the proper test statistic?
   a. Z
   b. t
   c. F
   d. χ^2
   e. Exact

4. What is the **p value** of the proper test statistic?
   a. 0.001
   b. 0.002
   c. 0.000
   d. 1.000

McNemar's Test for Change

5. Do you
    a. Reject the null hypothesis?
    b. Fail to reject the null hypothesis?

Two Sample Hypothesis Tests for Counts / Ordinal Data

Dependent Conditions

Wilcoxon Signed Ranks Test

- The Wilcoxon signed-rank test is a non-parametric statistical hypothesis test used to compare two related samples, matched samples, or repeated measurements on a single sample to assess whether their population mean ranks differ.
- Essentially a one-sample signed-ranks test conducted on the difference scores between the two groups

Wilcoxon Signed Ranks Test

- Underlying assumptions
  - The pairs of scores are independent of one another
  - The data is measured on at least an ordinal scale
  - The underlying characteristic is continuous
  - The absolute value of the difference scores is ordinal

Wilcoxon Signed Ranks Test Example

- An advertising researcher is interested in knowing whether a message communicated by way of a television commercial (video) is able to influence a consumer's assessment of the "freshness" of a loaf of bread.
- The freshness is to be assessed by squeezing the loaf of bread and then giving it a rating on a 10-point scale from 1 = "Not At All Fresh" to 10 = "Very Fresh."

Wilcoxon Signed Ranks Test Example

- The study was designed as follows. Forty-five people (representative of typical purchasers of the company's bread) were asked to squeeze three specimens and rate them for freshness.
- The specimens were actually foam rubber "loaves" in a typical plastic bread packaging bag.

Wilcoxon Signed Ranks Test Example

- Furthermore, the researcher was only interested in the ratings given to the one specimen whose softness (which has been shown to be related to freshness) was similar to actual loaves of bread as assessed by a quantitative gauge.

Wilcoxon Signed Ranks Test Example

- After the participants had rated the three specimens, they were shown a commercial that “taught them how to squeeze bread to determine its freshness”.
- This was done in the context of checking bread that is on the shelf at a market.

Wilcoxon Signed Ranks Test Example

- After viewing the commercial, the participants were asked to reassess and to rate the same three foam rubber “loaves of bread”. The data for this exercise are in the file named Fresh.txt. Test an appropriate hypothesis at the $\alpha = 0.05$ significance level.

Wilcoxon Signed Ranks Test

In RStudio

```
median.test.twosample.dependent.wilcoxon()
```

View the Full Hypothesis Testing Procedure here:

https://tinyurl.com/2smpdepwilcoxon

Your Turn!

- A professor wants to determine whether their new teaching method is more effective than the old one.
- They conducted a test before and after introducing the new method to the same group of 15 students.
- After each test, the students were asked to rank their understanding of the material on a scale from 1 (very poor understanding) to 5 (excellent understanding).

Your Turn!

- The data are in a file called teaching.txt.
- Test an appropriate hypothesis with $\alpha = 0.05$

Wilcoxon Signed Ranks Test

1. The hypothesis is about:
   a. μ
   b. M (distribution)
   c. π
   d. λ
   e. ρ

2. What is the source of the dependency?
   a. Repeated measures
   b. Matching
   c. Pairing (natural or artificial)
   d. Subdividing a homogenous unit

Wilcoxon Signed Ranks Test

3. What is the proper test statistic?
   a. Z
   b. t
   c. F
   d. χ^2
   e. T

4. What is the **p value** of the proper test statistic?
   a. 0.001
   b. 0.002
   c. 0.000
   d. 1.000

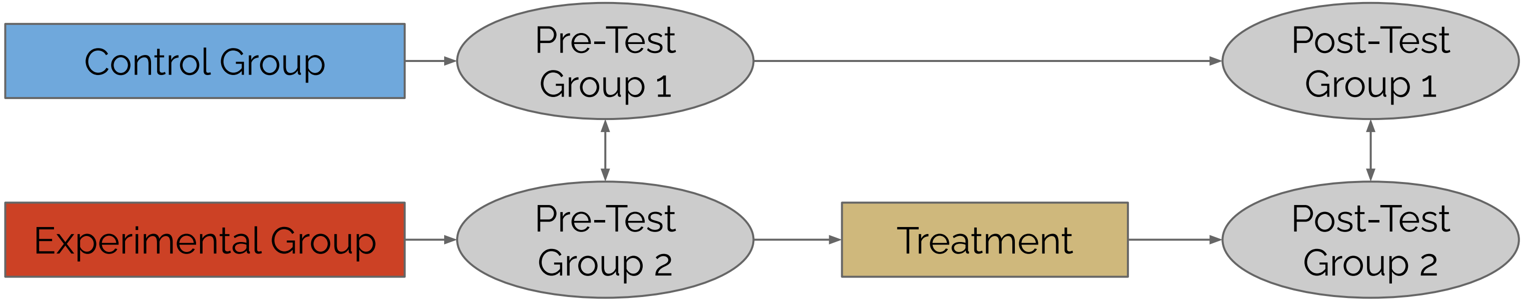
Wilcoxon Signed Ranks Test

5. Do you
   a. Reject the null hypothesis?
   b. Fail to reject the null hypothesis?

Pre-Test / Post-Test Design

Two Sample Tests

- One experimental group and one control group
- Randomized selection of subjects for both groups



Case Study

- A popular beverage company has been having a significant problem associated with the aluminum can label used for its flagship brand.

- There are two critical characteristics associated with the appearance of the brand label.

Case Study

- **Budweiser Red**: which is measured on a spectrophotometer, and which generates a continuous (ratio) measure. This quality characteristic has a target specification for it's **a*** value of 47; and many of the cans currently manufactured are below this Target Value, creating the appearance of a faded red color.

Case Study

- **Neck Scratches:** which is measured by a vision system, which assesses each can as either 'Good=0' or 'Bad=1'; based upon the visual inspection of all cans prior to the palletizer.

- As with the a* characteristic, the current level of unacceptable cans due to scratches produced, based on the visual camera inspection of the neck area, is too high and costly.

Case Study

- The Product Engineering group has come up with a solution that they believe will solve both these problems.
- They believe that by using a new overvarnish material, the red color will be closer to target and scratches will be significantly reduced.

Case Study

- In order to determine whether the new overvarnish process works, before they commit to the added expense for the new material and equipment, the engineers decide to run a Pre-Test / Post-Test Experimental Design.

Case Study

- To run this experiment, the engineers take a random sample of 1000 cans, produced under current and standard conditions.

- They then assign these cans at random to two groups (Group 1= Control & 2=Treatment) of 500 cans each.

Case Study

- Each can is marked with a unique identifier, and then tested in a randomized block sequence for color.

- The data are recorded and stored in a data file called **red.txt**.

Case Study

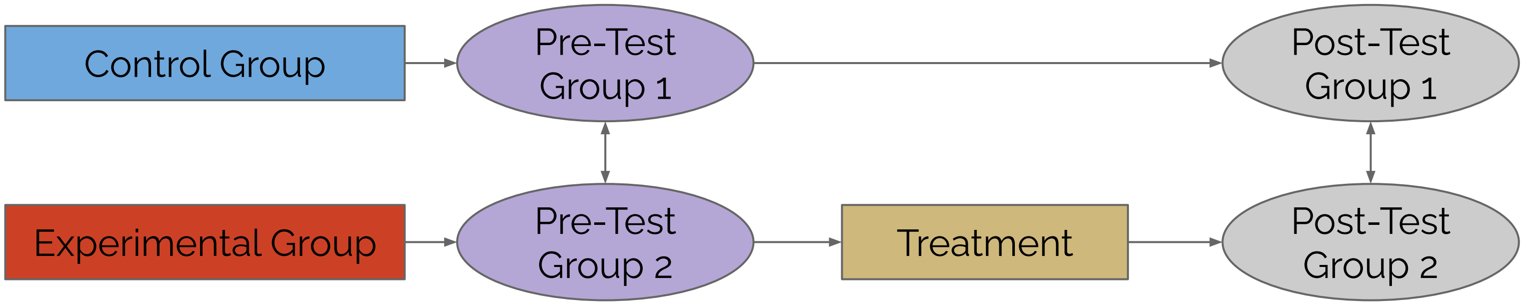
- After the color test has been completed, 200 cans are randomly identified within each of the two Groups (all of which were found to be acceptable based on the color test) and were subsequently evaluated for scratches.
- Each can is then assigned a value of 1 ('Fail') or '0' (Pass) depending on whether the can did or did not exhibit an excessive neck scratch as judged by the inspector. These data are in a file named **scratches.txt.**

Case Study

- Using a Type I Error level of 5% (maximum), and two-tailed tests throughout, perform the following assessments and answer the following questions:

Problem 1

- Are the two independent **Pre-Test** Groups equivalent? Test all appropriate hypotheses (for both color and scratches), and be sure to check the underlying assumptions for the hypothesis tests you perform.



Problem 2

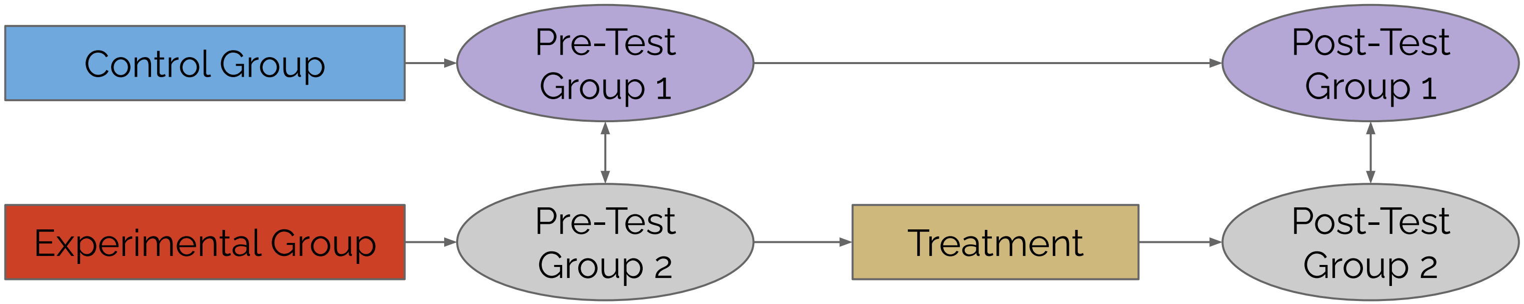
- Following the initial testing (the pre phase), the engineers ran the cans from Group 1, the Control Group, through Printer 1 using the old (original type of) material.
- Group 2, the Treatment Group, ran through Printer 1 using the new, proposed material.
- All tests were then repeated, (the post phase) and the newly collected data stored in the same data files.

Problem 2

- Because some of the data now should be compared on a dependent, repeated measures basis, this must be considered in your analysis.

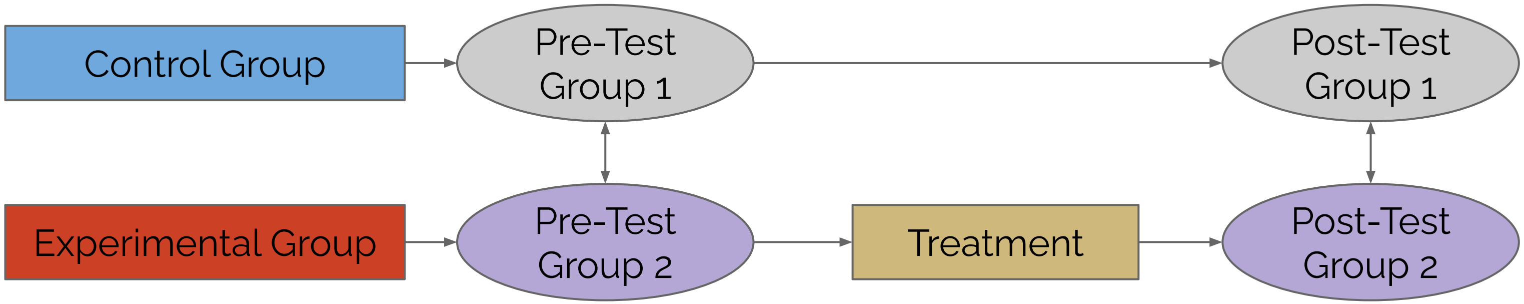
Problem 2

- Test all appropriate hypotheses to determine if the **Group 1** (Control Group) Cans changed from the Pre-Test to the Post-Test. Test all hypotheses required to check underlying assumptions, and show all your work and results.



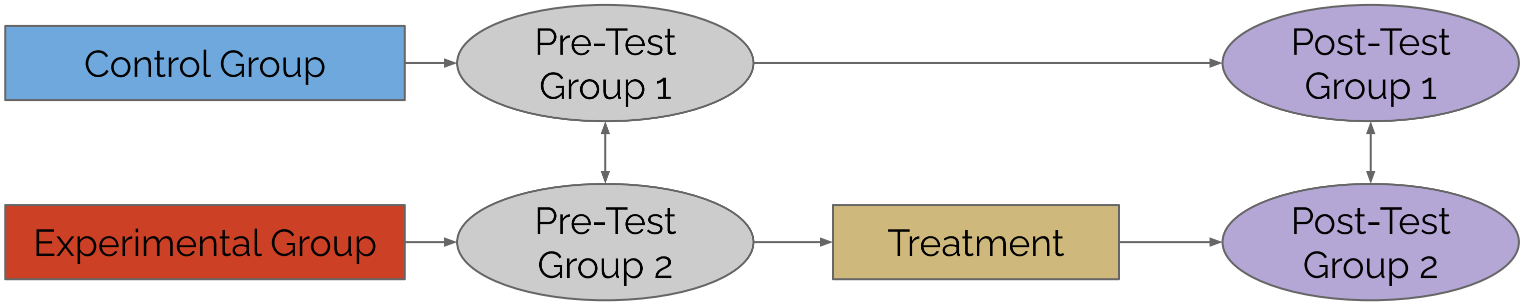
Problem 3

- Test all appropriate hypotheses to determine if the **Group 2** (Treatment Group) Cans changed from the Pre-Test to the Post-Test. Test all hypotheses required to check underlying assumptions, and show all your work and results.



Problem 4

- Test all appropriate hypotheses to determine if the Control Group 1 cans and the Treatment Group 2 cans are different for the **Post-Test** assessments. Test all hypotheses required to check underlying assumptions, and show all your work and results.



Problem 4

- Based on all of your analyses, was the Treatment (the change in the overvarnish material) effective?