Comparing Two Treatments

Designed Experiments - Why Use Them?

Designed experiments are used to:

- Determine which independent variables (X's) have the greatest impact on the response (Y).
- Quantify the effects of the independent variables (X's) on the response (Y).
- Prove that the independent variable (X's) you think are important really <u>do</u> affect the process.

The results of experiments may be used to:

1. **Shift the average of a process.** For example, using a higher cure temperature may increase hardness and move it to the center of the spec.
2. **Reduce the variation.** For example, a two week training program for all call takers may be more effective than a 1-week program, which may then reduce the variation in call-taking time from one call-taking center to another.
3. **Shift the average of a process AND reduce the variation.** For example, a higher setting of cleaning solution concentration may lead to better and more consistent cleanliness of parts in a cleaning system.

Observations Vary (cont'd)

We expect observations to vary, and would be concerned if they did not vary.

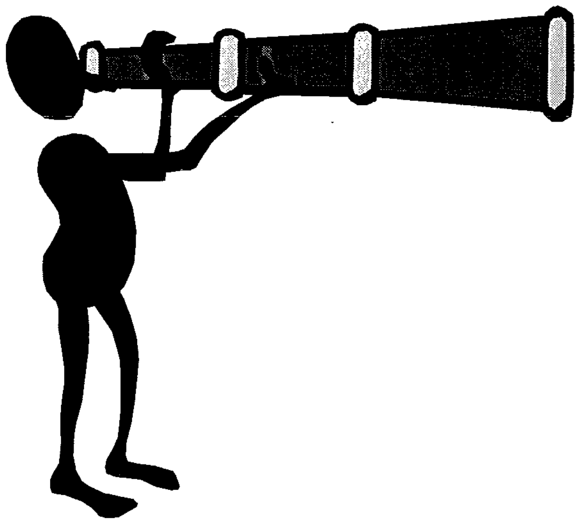
- ☆ If the sales volume of electric ranges was exactly the same for all regions, then we would suspect a corrupt database.
- ☆ If we measured 10 refrigerators and obtained the identical energy measurements, then we would question the quality of the measurements.

This variation makes our tasks challenging!

We generally cannot believe results from one data point. We often collect **multiple** data points, and take great care in how we collect these samples to observe the variation.

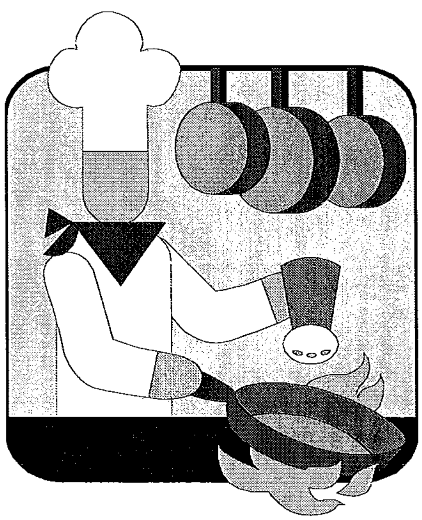
Variation is natural, expected, and the foundation of statistics

Passive Observation Vs. Experimental Design



Passive observation means to **watch** a process, without intentionally making changes (i.e., no “tweaking” the process during the collection of the baseline data).

Experimental Design means the **active manipulation** of independent variables, and observation of the effects on the dependent variables (responses).



Directed Experimentation

Directed experimentation is a method of learning through the combination of a critical event and an informed observer.

- Technology change is occurring far more rapidly than in the past.
- This is largely due to increases in the probability of learning through scientific methods.

We can improve this learning process by:

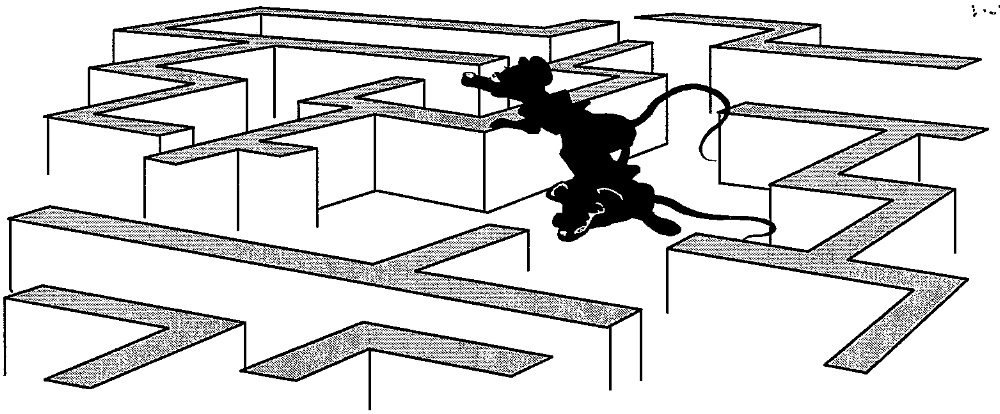
1. Making sure that naturally occurring informative events are brought to the attention of perceptive observers.
2. Inducing the occurrence of informative events through directed experimentation.

Through experimental design, one can demonstrate the ability to manipulate or control the dependent variables ("Y") by making changes to the independent variables ("X"s).

If you change things, maybe something exciting will happen.....

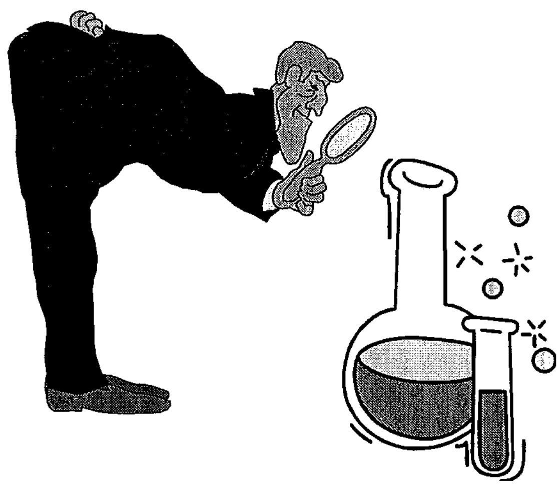
Why Do Experiments?

1. We want to demonstrate an ability to change and control a process.
2. It is a systematic approach to improvement.



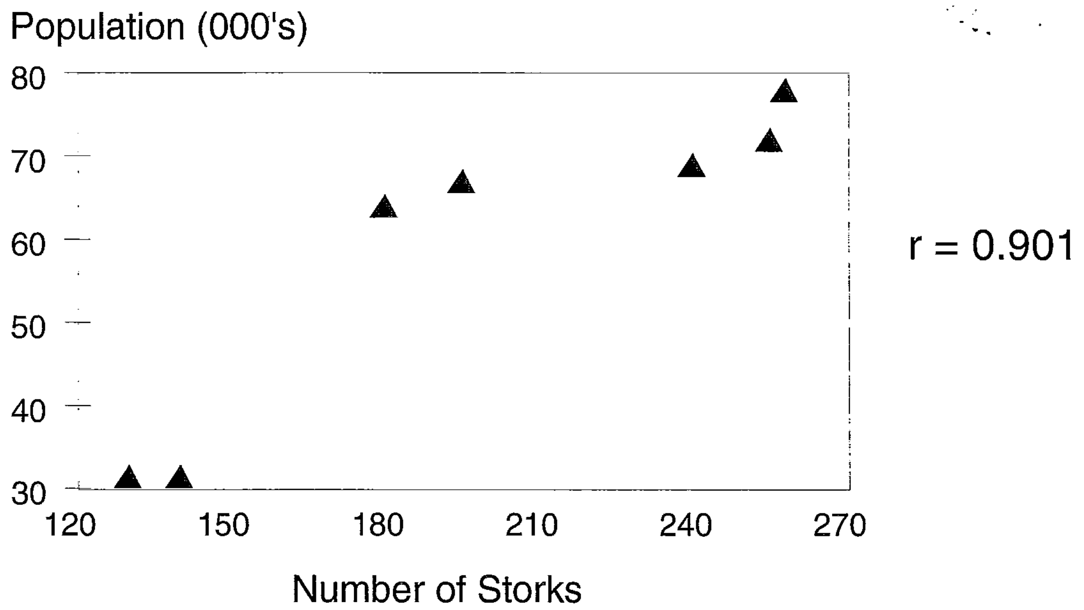
What is needed for Successful Learning

1. Make something exciting happen.
2. Someone needs to notice.



Caution!! Correlation Does Not Imply Causation

Note the graph of the population and storks below.
(Actual Data from Oldenberg, Germany, 1930-1936 published in Box, Hunter, Hunter, *Statistics for Experimenters*, page 8.)



Removing storks would not be a good method of birth control!

We may identify a relationship by observing a process:

- Two variables may tend to increase or decrease together
- However, this does **not** mean that we can adjust one variable through manipulation of the other.

Even Though "X" & "Y" are Correlated, We May Not be Able to Control the "Y" With That "X"! (CORRELATION ≠ CAUSATION)

What can we use to help us analyze this Problem?

We can see that the standard deviation for both wash times are the same. Are the two means statistically different?

Let's hypothesize that the two wash times produce the same level of brightness.

From Session 2, we remember:

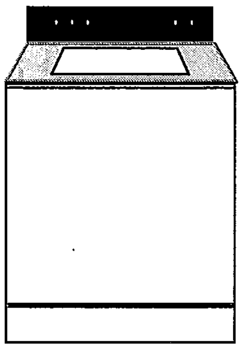
$$H_o : \mu_1 = \mu_2$$
$$H_a : \mu_1 \neq \mu_2$$

We can perform a 2-sample t test to determine this, but that will not help us quantify the difference between the two wash times if one exists.

However, Confidence Intervals will provide us with both answers.

Let's review a little about Confidence Intervals....

For the Horizontal Axis Washer example:



From our trials:

20 minute wash:
average = 18.56 standard deviation = 3.57

10 minute wash:
average = 15.85 standard deviation = 3.54

$$C.I. = (\bar{x}_1 - \bar{x}_2) +/- t_{\left(8+8-2, \frac{.05}{2}\right)} * s_p * \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Where:

$$s_p = \sqrt{\frac{\{(n_1-1)*s_1^2 + (n_2-1)*s_2^2\}}{(n_1+n_2-2)}} = \sqrt{\frac{\{(8-1)*3.57^2 + (8-1)*3.54^2\}}{(8+8-2)}}$$

$$\boxed{s_p = 3.56}$$

C.I. = (18.56 - 15.85) +/- 2.145 * 3.56 * √(1/8 + 1/8)

C.I. = 2.71 +/- 3.82 = **(-1.1, 6.5)**

Are the two means statistically different?

A 20 minute wash may improve reflectance over a 10 minute wash by as much as 6.5 units, or it may decrease reflectance by as much as 1.1 units.

The Effect of Sample Size on the Distribution of an Average

Suppose we sampled reflectance for a 20 minute wash every day for **30 days** (and the washing process is stable during this time).

Suppose it follows a normal distribution with an average of 19 and a standard deviation of 3.

We could :

- Measure 1 sample per day or
- 4 samples per day and record the average.

The following is a simulation of the results of measuring either 1 or 4 samples per day.

Daily average	**1 per day (30 samples)**	**Avg of 4 per day (120 total samples)**
24	X	
23	XX	
22	XX	
21	XXX	XX
20	XX	XXXXXX
19	XXXXXX	XXXXX
18	XXXXX	XXXXXXXX
17	X	XXXXXX
16	XXXX	XXX
15	XX	
14		
13	XX	
average:	**18.6**	**18.4**
standard deviation:	**2.84**	**1.45**

Using Minitab to Compare Two Treatments

Example: Using a horizontal axis washer, what is the effect of wash time on the ability to remove soot?

(i.e., if we wash the clothes longer will they look "cleaner"?)

We are using reflectance (brightness) as the measure of soot removal.

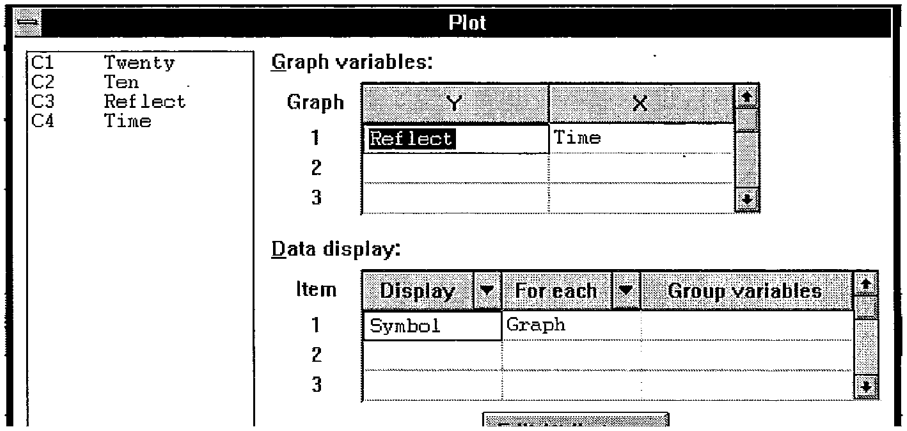
20 min	10 min
17.4	20.4
17.7	19.3
23.2	17.6
20.4	16.3
15.0	9.7
24.0	16.4
15.6	14.8
15.2	12.3

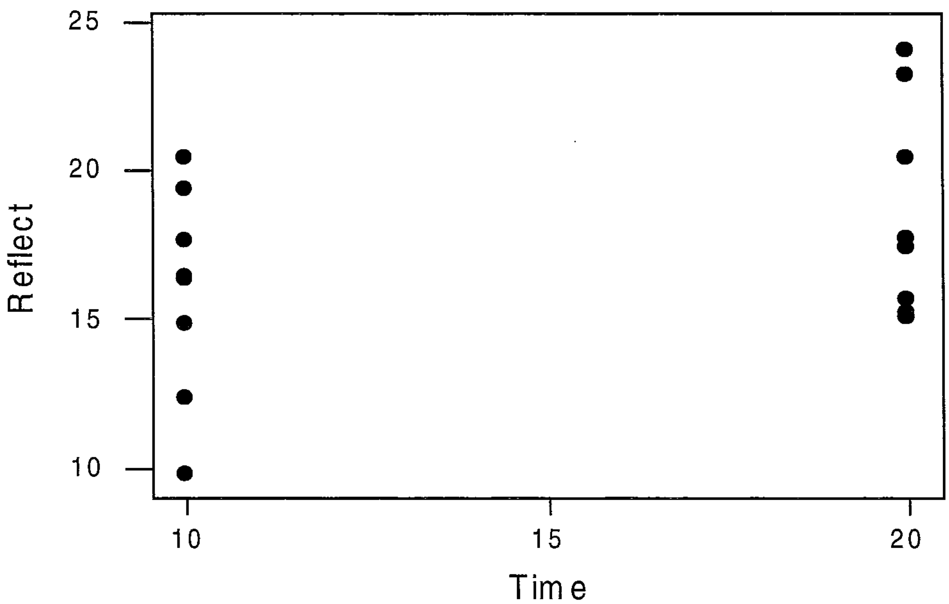
Simple graphs, such as scatterplots and dotplots, should be used to look at the data prior to any statistical analysis

Note: There are only 8 points in each group, so a histogram would not be very informative.

Using Minitab to Graphically Compare Two Treatments (cont'd)

Graph>Plot and / or Boxplot
Graph variables: Y = Reflect
X = Time



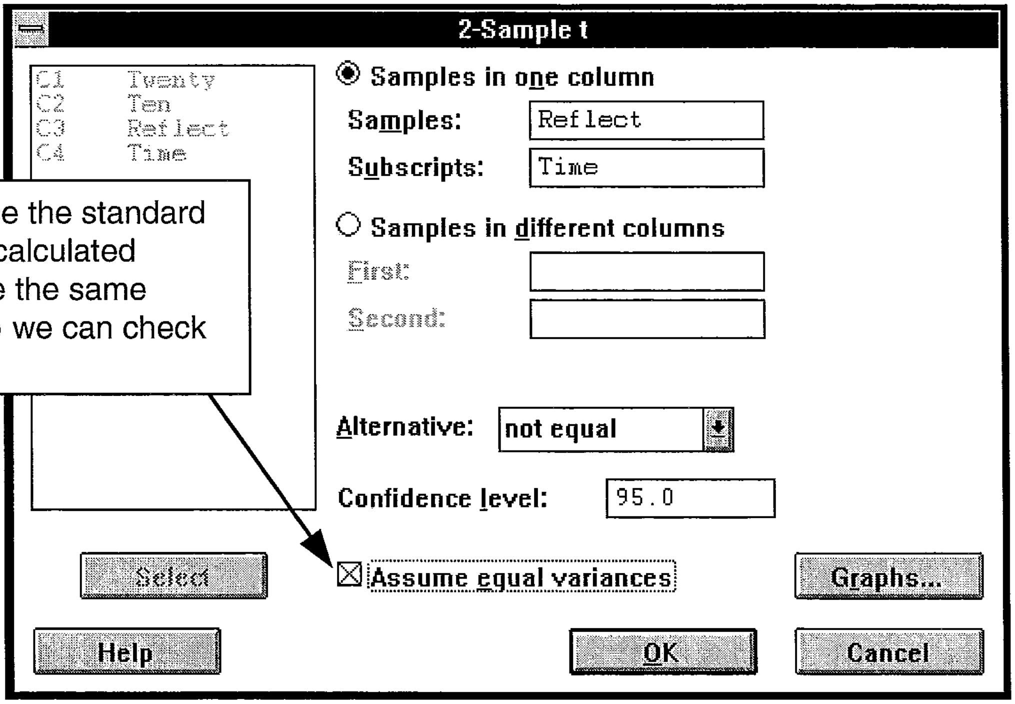


Using Minitab to Mathematically Compare Two Treatments

To generate the statistics and a Confidence Interval, use the 2-Sample t-Test:

Stat>Basic Stat>2-sample t

Fill in the dialog box as shown below. Then click 'OK'.



Note: Since the standard deviations calculated earlier were the same (3.57, 3.54) we can check this box.

We can obtain a point estimate and a confidence interval of the average difference between any two sets of data using a t-test.

If there is a statistical difference between the means of the two treatments, the confidence interval will not contain the value '0'.

Class Exercise - Helicopter Flight Time

Objective: Estimate the effect of wing width on the flight time of a paper helicopter.

1. What is the dependent variable (Y)? ______________

2. What is the independent variable (X)? ______________

A "treatment" is a level of an X variable.

3. How many treatments are there in this experiment? ____

Trial Number	Flight Time Narrow	Flight Time Wide
1		
2		
3		
4		
5		
6		
7		
8		

Key Concepts: Comparing Two Treatments

1. The objective of an experiment is to estimate the effects of the independent variables (X's) on the response (Y).
2. Two types of variation occur when collecting data: **systematic variation** (expected and predictable) and **random variation** (NOT predictable).
3. Statistics deals with variation through descriptive statistics, statistical inference, and design of experiments.
4. Multiple experiments will more than likely be required to find a solution.
5. Correlation does not imply causation.
6. A "treatment" is a level of an independent variable (X).
7. In an experiment comparing two treatments, you are estimating a difference between the averages and variation of each treatment.
8. The confidence interval gives the **range of plausible values** for the true difference between population averages. If the confidence interval contains "0", you cannot say there is a statistical difference.

Appendix

Reference Distributions and Controls

In an example similar to the helicopter experiment, we might be currently running a process with narrow wings, and we might be interested in determining if we can improve flight time by using wider wings. It might be tempting to run a test with wide wings only, and base our decision on that result alone.

But we are often interested in the **difference** attributable to a change, and some sort of a baseline or **control** is often of enormous benefit. Before starting any experiment it is a good idea to ask if a control should be included. More generally, it is a good idea to ask the following two questions before starting an experiment:

1. What are you going to do with the data after you collect it?
2. What decisions are to be made with the data?

These two questions might lead you to add a control, or to make other changes in the experimental procedure.