

Introductions

- NAME
- HOME TOWN
- FIELD OF RESEARCH AND/OR MAJOR
- SOMETHING FUN THAT YOU DID THIS SUMMER

1

Course Web Page

<http://www.public.iastate.edu/~dnett/S401/stat401.shtml>

The course web page contains

- Reading assignments including Wednesday's
- Homework assignments including the first
- Course handouts
- Lecture notes
- Practice exams
- Computer programs

2

Today's Handouts

Syllabus

Computing instructions

- You will need to use computers
- Find a friend to work with
- Consult former students in your lab
- Read the handout

3

Purpose of the Course

This section of Statistics 401 is designed for graduate students in agriculture and biological sciences.

The main goal is to help students understand basic statistical reasoning and learn methods of data analysis that will be helpful as students plan to conduct and analyze their own experiments.

4

Statistics

- Statistics is the science of drawing conclusions from data.
- Statistics is boring and frustrating if you do not understand it.
- Statistics is interesting and useful when you do understand it.

5

Basic Statistical Reasoning

- Suppose you have developed a treatment (drug, genetic alteration, feed, method of raising, etc.) that you believe will increase the lean percentage of hogs at slaughter.

(Lean percentage is often estimated as a function of carcass weight, back fat, and loin eye depth.)

6

- You have 100 pigs to use in an experiment to test your claim that the treatment will increase lean percentage.
- You randomly allocate 50 pigs to receive the treatment and the other 50 to be the control group.
- You treat all pigs in the same manner aside from giving the one group of 50 the treatment and not the other. (*Control group should get a placebo.*)

7

- You record the lean percentages of the 100 pigs at the time of slaughter.

Treatment	Control
49.2 52.4 50.0 54.4 51.0	49.5 53.5 46.7 52.1 49.2
54.3 52.5 52.9 56.0 54.4	48.2 51.5 51.7 50.4 52.5
54.8 53.8 53.0 49.7 50.4	52.5 49.8 50.5 51.2 49.3
53.4 52.2 56.1 52.4 48.9	48.5 52.8 51.7 53.1 47.1
50.9 52.5 51.9 56.5 49.4	52.6 49.7 47.9 54.6 53.0
53.6 52.1 50.4 49.1 53.4	52.9 52.3 52.3 50.1 51.1
47.8 54.6 53.2 54.4 51.6	49.8 53.8 51.1 52.2 50.9
54.3 47.6 50.6 51.0 50.5	50.0 50.1 50.3 50.5 49.7
50.0 53.6 50.2 51.1 53.3	50.0 49.7 48.1 50.0 51.8
53.0 52.9 54.4 53.7 54.0	52.3 49.1 48.7 51.8 51.8

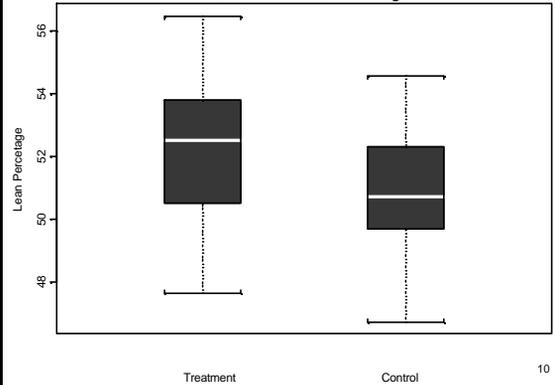
8

Summary Statistics

	Treatment	Control
Sample Size	50	50
Mean	52.3	50.8
Standard Deviation	2.1	1.8
Maximum	56.5	54.6
Q3	53.77	52.27
Median	52.5	50.7
Q1	50.52	49.7
Minimum	47.6	46.7

9

Side-by-Side Boxplots of the Lean Percentages of Control and Treated Hogs



10

Is there evidence that the treatment increased lean percentage?

- If you randomly split a group of 100 hogs into two groups of 50, you certainly wouldn't get the same mean lean percentage in both groups even if you treated them exactly alike. Maybe the difference that we see is just such a difference?

11

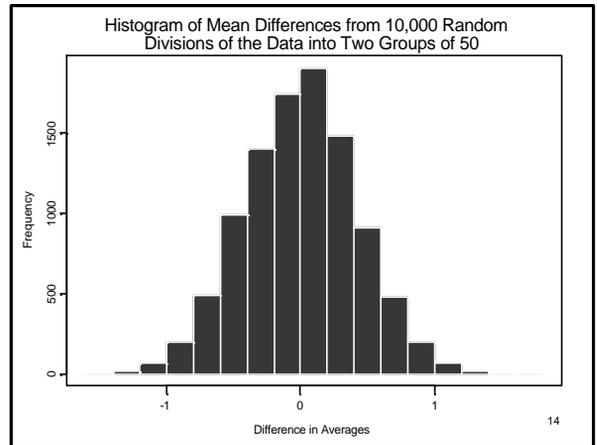
- Skeptic: "Your treatment didn't have anything to do with the lean percentages. When you split the hogs into two groups, you just happened to put the leaner hogs into the treatment group."
- Response: "Suppose the treatment had no effect, and the hogs developed their lean percentages independently of the treatment or placebo. What is the chance that the 50 hogs randomly assigned to the treatment group would have an average at least $52.3 - 50.8 = 1.5$ points higher than the average for the hogs randomly assigned to the control group?"

12

Randomization Test

1. Randomly divide the 100 lean percentages that we observed in this experiment into two groups of 50.
2. Note the difference in means between the two groups.
3. Repeat 1 and 2 a total of 10,000 times.
4. Note the proportion of times the difference was as from 0 as it was in the original experiment ($52.3-50.8=1.5$).

13



- The mean difference was as far from 0 as 1.5 for only 2 out of the 10,000 random divisions of the data into two groups of 50.
- Thus the difference between the mean lean percentages would almost always be less than the observed difference of $52.3-50.8=1.5$ if the treatment had no effect.
- It seems reasonable to believe that the treatment caused the difference in mean lean percentages.

15