

# Basic Data Analysis

Stephen Turnbull

Business Administration and Public Policy

Lecture 2: 15

**Abstract**



•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•



•

•

•

÷

•

•

•

•

1

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•



•

•

•

•

•

•

•

•

•

•

•

- control

- 

- 

-

•

•

•

•

•

•

•

•

•

•

•

•

•

# Homework 1: due April 19, 11:45am

**Submit** homework to `data-hw@turnbull.sk.tsukuba.ac.jp` by **email**. Note the due date is April 19, **11:45am**. Submission time is time of receipt by the server.

For this homework, please submit as *plain text* (no wordprocessor or PDF attachments). In other words, “just type” your answer in the email. In the *first line*, include your name, your student ID number, and the words “Homework 1”.

You may answer in English or Japanese. In answering, you *must* explain why the answer you give is correct *in your own words* to receive credit.

If you wish to ask me questions and get an answer, *write a separate email* to `data-help@turnbull.sk.tsukuba.ac.jp`.



# Reading

Read **Chapters 1 and 2** in Freedman. (They're short, and you don't need to study them so carefully.)

Especially concentrate on **sections 1.4 and 2.7** (the summaries).

Please read *all* of the **problems in sections 2.5 and 2.6** in the following order: starting with the problem number having the last digit in your student ID number, read every 5th problem. *E.g.*, if your student number is 201220666, start with problem 6, then 11, *etc.* When you run out of problems, keep counting from the beginning. (This is important, see Problems 1 and 2 below.)

# Problems

1. Of the problems in Ch. 2, which problem would you *most* like to see on the midterm? (The most popular problem *will* be on the

midterm!)

2. Of the problems in Ch. 2, which do you think are “impossible”? (The least popular problem will *not* be on the midterm!)
3. In student evaluations of a course, one question that might be asked is “was the pace of lectures (1) too slow; (2) about right; (3) too fast?” The students’ answers might be collected as data for a variable named **pace**. Is **pace** a qualitative, ordinal, or cardinal variable? If cardinal, is it discrete or continuous?
4. In Problem ??, suppose that 50% of the students say the lectures were “too slow”, and 50% say that it was “too fast”. Does it make sense to take the average, and say that pace was “about right”? Explain why you think so.

5. In the mass media (*e.g.*, newspapers or their Internet home pages, but *not* a statistical textbook or the Wikipedia), find an example of each type of variable:

(a) qualitative

(b) ordinal

(c) discrete cardinal

(d) continuous cardinal

Also give the URL or bibliographic information about where you found the variable. You may use one source for all the types, or different sources, as you find convenient.

6. In the mass media, find an example of statistics derived from each of

(a) a controlled experiment

(b) an observational study

Also give the URL or bibliographic information about where you found the example. You may use one source for both examples, or different sources, as convenient.

7. In the *controlled experiment* you gave as an example in your answer to Problem ??, was the treatment chosen randomly?

8. For the *observational study* you gave as an example in your answer to Problem ??, give an example of a possible *confounding variable*.