Statistics I, Fall 2012
Homework 01

Ling-Chieh Kung
Department of Information Management
National Taiwan University

1. (25 points; 5 points each) Classify each of the following as nominal, ordinal, interval, or ratio data:

   (a) The ranking of a company by *Fortune* 500.

   (b) The number of tickets sold at a movie theater on any given night.

   (c) The birthplace of a student.

   (d) Per capita income.

   (e) Profit/loss in dollars.

2. (25 points; 5 points each) Suppose you are deciding the color of a backpack and you want to choose among red, blue, and yellow. You make a website with three designs of your new backpack, each with a different color. Viewers of this website can vote for their favorite design.

   (a) What is the population?

   (b) Suppose 10000 people vote on your website. Do these votes form a sample? Why or why not?

   (c) What is the scale of measurement of these 10000 votes?

   (d) What is the parameter you want to know?

   (e) What kind of statistics would you design? Briefly explain why.

3. (50 points) The MS Excel file "StatFa12_hw01.xlsx" contains the monthly weight of cold storage holdings for green beans over a 14-year period. In column B, each number represents a weight of green beans (in 1000 lb).

   (a) (15 points) Prepare a frequency distribution of this set of data. Create nine classes. Let the lower bound of the first class be 30 million lb. Use an equal class interval of 40 million lb for all classes. Include relative frequency, cumulative frequency, and cumulative relative frequency in your table.

   (b) (10 points) Draw a histogram based on your answer in Part (a).

   (c) (5 points) According to the histogram, suggest a lower bound L and an upper bound U so that most values fall in $[L, U)$.

   (d) (5 points) Within $[L, U)$, there are probably multiple classes. If this is the case, do values spread roughly equally among these classes in $[L, U)$? Or there is a big difference among these classes in $[L, U)$? How do you define similarity among classes?

   (e) (15 points) Prepare another histogram for this set of data. This time use five classes with the first class starting at 30 million lb and class interval 70 million lb. Is the information revealed by these two histogram different? For example, if you are going to find the two bounds for the second histogram as you did in Part (c), would the bound be the same? What are other differences?

   Note. You are strongly suggested to use Excel (or Google Document or anything spreadsheet software) to do this problem. The following MS Excel functions may help: MAX(), MIN(), COUNT(), COUNTIF(), and FREQUENCY(). Knowing how to use "$" and "&" can also be helpful. Probably you cannot find a type of chart called "histogram" in MS Excel. In this case, try to make a "column chart" look like a histogram. Somehow find your own way to complete this problem. The TAs will give you a suggested way of doing it in the lab.

4. (0 point) Below are some questions regarding your feeling in the first lecture. The responses of all the students in this class will be useful for improving the quality of this course. They will be used as raw data for you to conduct statistical studies later in this semester.

In order to protect you, please detach this page, answer the following questions on this page, and DO NOT write down your name or anything that may reveal your identity. When submitting your homework, put TWO things into my mailbox:

- Your answers to Problems 1 to 3 with your student ID and name; and
- Your answers to those questions on this page WITHOUT your ID or name.

It is completely fine if you do not submit the second part.

(a) After the first lecture, do you think Statistics is important:

Yes   No

Why? _____

_____

(b) You think understanding the instructor's English was:

Easy   Not easy but OK   Hard

How may he improve (e.g., slower, faster, louder, etc.)?

_____

(c) You think the pace of the first lecture was:

Too slow   OK   Too fast

(d) You expect to spend how much time on Statistics per week (excluding lecture time)?

0 to 2 hour   2 to 4 hours   4 to 6 hours

Even more! How much? _____

(e) At the end of this semester, you expect to get:

A+   A   A−   B+   B   B−   C+   C   C−