Statistics — Class Project — Part 1 of 3

This is the first part of the three-part class project. The class project is designed to give you a longitudinal perspective of the topics covered in the course. The grading for the course project will be 80% correctness, 20% clarity of presentation. I encourage you to *neatly* write up your answers, as you will be using them in the future parts of the project.

1. DATA COLLECTION

(a) Flip a coin 120 times. Record the results of this experiment in 5 rows, with 24 flips per row. For example, the first row may be

H T H T T T H T T H T T H H H T T H T H H H H T

(b) For each *pair* of flips, write down the number of heads in that pair. For the 24 flips shown above, this would be done as follows:

$$\underbrace{\mathrm{H\ T}}_{1}\ \underbrace{\mathrm{H\ T}}_{1}\ \underbrace{\mathrm{T\ T}}_{0}\ \underbrace{\mathrm{H\ T}}_{1}\ \underbrace{\mathrm{T\ H}}_{1}\ \underbrace{\mathrm{T\ T}}_{0}\ \underbrace{\mathrm{H\ H}}_{2}\ \underbrace{\mathrm{H\ T}}_{1}\ \underbrace{\mathrm{T\ H}}_{1}\ \underbrace{\mathrm{T\ H}}_{1}\ \underbrace{\mathrm{H\ H}}_{2}\ \underbrace{\mathrm{H\ T}}_{1}$$

The numbers would then be separately recorded into a data set. You should *explicitly* show underbraces in your write-up. The result will be a data set with 60 numbers — each either 0, 1, or 2. We will refer to this as Data Set A.

(c) Repeat (b), except for each *triple* of flips. For the 24 flips shown above, the result would be

2 0 1 1 3 1 2 2

The full data set will have 40 numbers — each either 0, 1, 2, or 3. We will refer to this as Data Set B.

2. DATA PRESENTATION

(a) For Data Set A, make a *frequency table* displaying how often each pair had 0, 1, and 2 heads.

(b) For Data Set B, make a *frequency table* displaying how often each triple had 0, 1, 2, and 3 heads.

(c) Using these frequency tables, draw *relative-frequency histograms* for both Data Sets A and B. Use 3 classes for Data Set A and 4 classes for Data Set B. Important: Since the class width is 1 for both data sets, you don't need to worry about class limits and class boundaries for this problem — that is, the production of these histograms will be somewhat easier than the examples from the book and from class. For Data Set A, just draw rectangles of appropriate height and width 1 over 0, 1, and 2. For Data Set B, draw rectangles of width 1 over 0, 1, 2, and 3.

(d) *Without doing any calculations*, draw a rough picture of what you think the relative-frequency histogram for groups of *four* flips may look like. Briefly justify your picture.

3. DATA ANALYSIS

Compute the sample mean ($\overline{x}$) and sample standard deviation (s) for Data Sets A and B. You should select (and mention) the method you used to compute these descriptive statistics most efficiently.