

# Homework 1, due Wednesday, February 3

## Very important: Show All Working!!

1. In a certain year, the mean SAT score for all students is 1200 and the standard deviation is 300. Assume that the distribution is normal.
  - (a) What percentage of students scores above 1320? [3 points]
  - (b) A certain college decides to give automatic acceptance to all students who score in the top 12% of all SAT scores. What SAT score does that correspond to? [3 points]
2. The file SATscores.csv is on the Data page in sakai. Using R, answer the following questions:
  - (a) What are the sample mean  $\bar{y}$  and the sample standard deviation  $s_y$  of this dataset? [2 points]
  - (b) The *standard error* of a dataset is defined to be  $s_y/\sqrt{n}$ , where  $n$  is the sample size. For this dataset, what is the standard error? [2 points]
  - (c) Draw a histogram of the data [3 points]
  - (d) Draw a QQ-plot of the data [3 points]
  - (e) Based on the histogram and the QQ-plot, would you say the data are normally distributed? [2 points]
  - (f) A rough rule of thumb for when a sample mean is consistent with a hypothesized population mean (here, 1200) is that the difference between the two means should be less than 2 standard errors. Based on that, would you say the data are consistent with a population mean of 1200? [2 points]

## Hints for Using R

1. Use the R functions `pnorm`, `qnorm`. For explanations, type `?pnorm` or `?qnorm`.
2. To read a csv file into R, type something like

```
SAT=read.csv('SATscores.csv')
```

You may need to insert the directory path before the file name.

If you type `SAT` at the keyboard, you will get a listing of the data: the first few lines are

Observation	SAT
1	1 1190
2	2 1240
3	3 1120
4	4 1430

The values in the second column are the ones you need. If you prefer a short variable name, say `y`, you can enter `y=SAT$SAT`.

- (a) The sample mean and variance of the vector `y` are given by `mean(y)` and `var(y)`. For standard deviation, `sy=sqrt(var(y))`.
- (b) The sample size is `length(y)`.
- (c,d) Use the commands `hist` and `qqnorm`. You can find more information by typing in `?hist` and `?qqnorm`.