Test 1 will be Monday October 20th and will cover chapters 7, 8 and 9 in the book. (We did not cover, nor will you need to know, sections 8.9, 8.11, 8.12, 9.11 and the theory of 9.12.) What follows is a brief, and by no means complete, summary of what you should know.

- **Chapter 7** Chapter 7 was about displaying and summarizing data. You should be familiar with reading the graphs presented in this section in particular the boxplot and quantile plots which are likely new. As well, you should be able to calculate the mean, median and correlation for a given data set.
- **Chapter 8** Chapter 8 was much harder than chapter 7 as concepts from probability are used throughout.
 - random sampling: you should know what a random sample is and what the p.d.f. of a random sample is when you know the parent population
 - You should be able to compute a likelihood function or the negative log likelihood function for various assumptions on the randomness of the sample.
 - You should know what a statistic is and you should understand what sufficient means for a statistic.
 - You should understand the relationship between a sampling distribution of a statistic and the parent distribution where the sample is drawn from.
 - You should be able to appreciate the derivation of the sampling distribution of the median as shown in class.
 - You should be familiar with the language of sample proportion vs. population proportion.
 - You should know the mean and std. of the sample average in terms of the parent population.
 - You should know what the CLT is and how to interpret it. For example, what does it mean that *n* is large enough? How does it allow you to say $P(\bar{X} > x) = \dots$
 - You should be familiar with the fact that the sample mean for normal data is normally distributed and the distribution of $(n-1)S^2/\sigma^2$ is known.
- **Chapter 9** Chapter 9 begins to ask questions from a statistical viewpoint. We have some data, what does it say about the population parameters? We studied confidence intervals in this chapter. As well, we had a discussion of estimators and learned this language: an estimator, the m.s.e. of an estimator (compare to variance), the standard error of an estimator and a consistent estimator.

Can you describe confidence intervals in all of these situations?

- A sample proportion. Eg. a public opinion poll.
- Two sample proportions. Eg. the same poll taken on different weeks.
- A sample mean. When *n* is small, but the data is normal what pivotal quantity do you use?
- A confidence interval for the difference of means.
- A confidence interval for σ^2 based on the pivotal quantity $(n-1)S^2/\sigma^2$.

Can you explain how a confidence interval can have a 95% level of confidence when it either contains the population parameter or it doesn't.

Can you explain why the New York Times has this disclaimer

"In theory, in 19 cases out of 20, the results from such polls should differ by no more than plus or minus two percentage points from what would have been obtained by seeking to interview everyone who cast a ballot in the recall election."