Test 3 in 214 will cover the following sections in the book:
8.1,2,3; 9.1,2,3,4

Test 3 is basically the material on significance tests. The basic setup is important to understand. The pieces are

1. A pair of hypotheses H_0 and H_A.
2. A test statistic with known sampling distribution to judge the validity of an assumption that H_0 is true
3. A data set that yields an observed value of the test statistic
4. An analysis computing a p-value, or atleast comparing it to a significance level α.

We learned the following tests:

1. test for proportion
2. Test for mean
3. test of difference of propotions
4. test for difference of means (with different assumptions: equal variance, non-equal variance)

Each test has a different test statistic and a different set of assumptions. However all the test statistics (except for a single proportion) we met have this form

$$\frac{\text{observed} - \text{expected}}{\text{SE}}$$

What these means differs depending on the problem. For instance we have all of these possibilities for SE:

$$s/\sqrt{n}, \sqrt{\widehat{p}(1-\widehat{p})(1/n_1+1/n_2)}, s_p\sqrt{1/n_1+1/n_2}, \sqrt{s_1^2/n_1+s_2^2/n_2}, \ldots$$

For each significance test you should do all of the following:

1. Specify the null and alternative hypotheses. This is really important, as once you've written this down, the choice of a test statistic follows and after that the problem becomes mechanical.

2. Ask your self if the assumptions that allow you to "know" the sampling distribution of the test statistic are met. (I.e., normality of the population(s), variances are equivalent, large sample, ...)
3. Find the p-value (or atleast find bounds for it) and compare it to $\alpha = 0.05$. Answer if the differences are statistically significant.

Additionally, if you can, draw a nice picture of the sampling distribution of the test statistic, and shade in either the p-value or the critical values so that you can see the answer.

1. The English mathematican John Kerrich tossed a coin 10,000 times and obtained 5,06 heads. Is this significant evidence (at the 5% level) that the probability that Kerrich's coin comes up heads is different from 0.5?
2. Land's Beginning is considering buying a mailing list of 100,000 people, but only if they are reasonably confident that 5% (or more) will respond to direct mailing. They are given a random sample of 500 from the list and find that only 4% responded to a mailing. Is this difference from 5% statistically significant to make LB not want to buy these names?
3. Christmas tree growers want to know if there is a difference in preference for natural vs. artificial trees between city and suburb dwellers among those who buy christmas trees. They perform a small survey and find these preferences:

   ```
   population n prefer natural
   ---------------------------
   urban      160        64
   suburban   261        89
   ```

   Is the difference statistically significant?
4. Does removing a annual credit card fee change consumer purchases? A bank investigates by waiving the fee on a random sample of 500 of its customers and tracks differences in the difference in the amount they spent. The mean increase was \$565 with $s = \$267$. Is there significant evidence at the 5% level that the mean amount increased?
5. How accurate are home radon detectors? A company tested 6 in an environment with an *accurately* measured amount of 115. The 6 cheap detectors found

```
91.9  122.3 105.4  95.0 99.6  120.9
```

with summary

```
     xbar        sd         n
105.85000  13.02839   6.00000
```

Is there evidence to indicate that the home measurements are lower than they should be?

6. Do piano lessons improve spatio-temporal reasoning in children? To test, two independent groups were formed, one given piano lessons for one month the other not. Differences in reasoning were assessed numerically and are summarized by

group	n	xbar	s
piano	12	10.7	3.8
control	11	9.6	4.0

Is the difference statistically significant? (Did you assume equal variances? independent or paired samples?)

7. A study of iron deficiency between children who were bottle-fed versus nursed is summarized by

group	n	xbar	s
Nursed	23	13.3	1.7
Bottle	19	12.4	1.8

Is the difference statistically significant? (Did you assume equal variances? independent or paired samples?)

8. Does cocaine use by expectant mothers cause smaller birthweights? A restrospective study presents the following data

Group	n	xbar	s
Cocaine	134	2733	599
No Cocaine	5974	3118	672

Is the difference statistically significant? (Did you assume equal variances? independent or paired samples?)

9. A classic test on the effect of a home environment on IQ was performed in the 60s by Burt by using pairs of identical twins that were separated at birth. One twin raised by foster parents, one raised by the birth parents. For each pair, IQ scores were recorded. Suppose the data recorded was

```
              data                  n xbar    s
--------------------------------------------------
foster     120 101  95  97  97     5  102    10.2
biological 129 102  97  99 105     5  106.4 13
```

Let μ_f and μ_b be the population means for all such twins. Perform a two-sided significance test of $H_0 : \mu_f = \mu_b$.

As the two samples are not independent, this can not be done as a two saple test. Rather, do a one-sample test to see if the mean of the differences is 0.

(a) What is the p value?

(b) Why did the researcher used twins, and not two random samples?

(c) What assumptions did you make about the data?

10. What assumptions are needed to know the sampling distribution of

$$T = \frac{\bar{x} - \mu}{s/\sqrt{n}}?$$

11. What assumptions are needed to know the sampling distribution of

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}?$$