Opening day problems

Here are examples of the sorts of substantive problems we'll be addressing in this course. We'll go over what statistical techniques are and are not appropriate for such problems, and we'll show you how to use these techniques correctly

1. An advertiser wants to test (using $\alpha = .01$) a magazine publisher's claim that 25% of the magazine's readers are college students. A random sample of 200 subscribers is to be taken.

2. A police chief wants to test whether, on average, the incidence of crime is the same in four city districts. The number of police calls received from each district in each of the last 30 days is to be used as a sample. (Note that n = 4 * 30 = 120.)

3. A researcher is interested in whether severity of accident and location of accident are related. He plans to draw a sample of 100 accident records. In the accident records, severity is coded as either (1) property damage, (2) injury, or (3) fatality. The three possible locations for accidents are freeway, rural road, or city road.

4. A psychologist wants to test whether age influences IQ. A random sample of 60 45 year olds, whose IQ score at age 20 is known, is to be taken.

5. The mayor wants to know whether mean apartment rents are now higher than last year's $375 per month. A random sample of 20 of the town's apartments is to be taken. The desired significance level is $\alpha = .01$. It is assumed that the current population standard deviation is the same as last year, $\sigma = \$25$.

6. A pollster believes that voters are evenly divided in their support for 6 candidates. To test her idea, she plans to poll 200 people and ask them which of the 6 candidates they are most likely to vote for.

7. The Notre Dame administration wants to get a better understanding of alcohol consumption on campus. Five hundred students will be surveyed. The administration will see whether the amount of alcohol consumed (measured in ounces) differs by gender (male, female) and race (white, nonwhite).

8. A quality inspector wants to test the claim that the proportion of acceptable electronics components delivered by a foreign supplier (A) is lower than that coming from a domestic supplier (B). Random samples of $n_A = 100$ and $n_B = 150$ are to be taken from incoming shipments and tested; the desired significance level is $\alpha = .01$.

9. An economist wants to test the claim that the wages of hotel workers in New York (A) are different from those in Chicago (B). Random samples of workers, $n_A = 200$ and $n_B = 100$, are to be taken. The economist wants to use $\alpha = .01$. She assumes the populations are normally distributed and have equal variances.

We'll start off the semester by talking about probability theory. Probability theory provides the underpinnings for much of what we do in statistics. Here are a few sample problems. The first four are mostly for fun, but the fifth is similar to practical problems you might actually encounter.

1. In a family of 11 children, what is the probability that there will be more boys than girls? Assume that the biological odds of having a boy or a girl are equal.

2. There is a family with two children. You have been told this family has a daughter. What are the odds they also have a son, assuming the biological odds of having a male or female child are equal?

3. You are playing *Let's Make a Deal* with Monty Hall. You are offered your choice of door #1, door #2, or door #3. Monty tells you that goats are behind two of the doors. Behind the other door is a brand new car. You choose door #3. Monty then opens door #2, and shows you that it has one of the goats. He then offers you the choice of either keeping your own door, #3, or else switching to Monty's remaining door, #1. Should you switch? (Note: you can assume that Monty Hall knows what is behind each door. In addition, you can assume that Monty Hall follows this same procedure with every contestant, i.e. the contestant chooses a door, Monty then opens up one of the losing doors, and then offers the contestant the option to switch.)

4. You have to take a true-false test on a subject you know absolutely nothing about. You must get 2/3 or more of the answers right in order to pass. Would you rather take a 3 question test, where you had to get at least 2 answers right, or a 36 question test, where you had to get at least 24 answers right? Or would it not make any difference to you how long the test was? Explain your reasoning.

5. A researcher is doing a study of gender discrimination in the American labor force. She has come up with a 3-part classification of occupations (Occupation 1, Occupation 2, and Occupation 3) and a 2-part classification for wages ("good" and "bad"). In a study of 100 men and 100 women, she finds that, by gender, the distribution of occupation and wages is as follows:

	Women			Men		
Pay/Occ	Occ 1	Occ 2	Occ 3	Occ 1	Occ 2	Occ 3
Good Pay	20	7	10	7	10	60
Bad Pay	50	8	5	8	5	10

From the table, it is immediately apparent that 37% of all women receive good pay, compared to 77% of the men. At the same time, it is also very clear that the types of occupations are very different for men and women. For women, 70% are in occupation 1, which tends to pay poorly, while 70% of men are in occupation 3, which tends to pay very well. Therefore, the researcher wants to know whether differences in the types of occupations held by men and women account for the wage differential between them. How can she address this question?