Math 121 - Midterm 1 Review

Here are problems that are similar to the ones you might see on the exam. Be sure to also review old homework and lab questions too.

1. The following comes from an online news source:

   MONDAY, Feb. 18 (HealthDay News) Young adults and teens who smoke, are obese and have high blood sugar levels may be more likely to die before they reach their 55th birthday, new government research suggests. The average life expectancy in the United States is 78.7, according to the CDC. Dr. Saydah and her colleagues analyzed data on close to 9,250 people who took part in the third National Health and Nutrition Examination Survey. Participants were aged 12 to 39 when the study was conducted. Of these, more than 15 percent were obese, and 30 percent were smokers. Overall, 298 of the participants died before they turned 55. Those who smoked between the ages of 12 and 39 had an 86 percent greater risk of dying before 55, compared with those who did not, the data showed. Those who were obese when they were young had a 39 percent higher likeli- hood of dying before 55, compared with those not obese during these early years. In addition, the risk of dying before 55 tripled among those with high blood sugar levels between the ages of 12 to 39, the study showed. High blood pressure and high cholesterol levels, however, did not affect the risk of dying before age 55.

   Answer the following questions using the text above.

   (a) Who or what were the individuals in the study?

   (b) What is the sample size?

   (c) What were the variables collected?

   (d) Which variables were categorical and which were numerical?

2. A sample of households in a community is selected at random from the telephone directory. In this community 4% of the households have no telephone, 10% have only cell phones, and another 25% have unlisted telephone numbers. The sample will certainly suffer from

         A. nonresponse.

         B. undercoverage.

         C. false responses.

3. The average time for an emperor penguin egg to hatch is 70 days, with a standard deviation of 5 days. Assume the distribution of hatching times is approximately normally distributed.

   (a) Find the probability that an egg hatches between 65 and 75 days.

   (b) Find the probability that an egg hatches in 73 days or more.

   (c) By what day have 90% of the eggs hatched?

4. The following numbers are a sample of the monthly rental prices for apartments in Biloxi, Mississippi.

$$290 \ 180 \ 320 \ 260 \ 160 \ 980 \ 470 \ 150 \ 120 \ 850$$

   (a) Arrange the data in a stem and leaf diagram. The stems should be in the hundreds place.

   (b) Find the five number summary for the numbers.

   (c) Draw a box-and-whisker plot for this data.

   (d) Which is probably larger for this data, the mean or the median? Explain how you can know without calculating the mean.

5. The following chart from `salary.com` depicts the distribution of salaries for public school teachers in New York city.
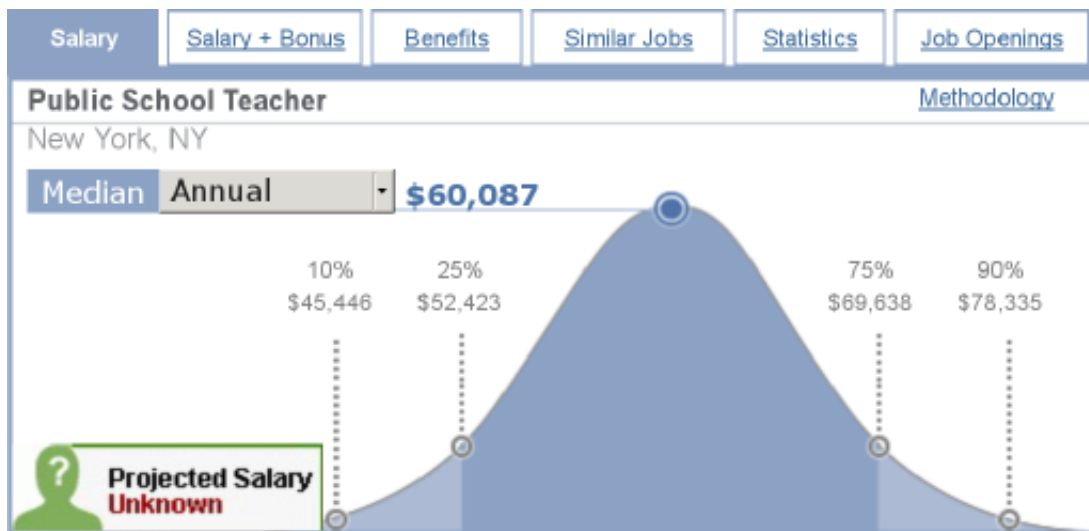
   (a) How can you tell that the distribution is not really normal, even though the picture makes it look like it is?

   (b) Is the distribution skewed left or right?

6. A sample of 200 frogs is collected near a nuclear power plant, and 16 frogs have visible mutations. What is the proportion of frogs that are mutants?

7. Is the answer to the previous question a parameter or a statistic? Explain why.

Read the following to answer the next two questions.

A state representative wants to know how voters in his district feel about enacting a statewide smoking ban in all enclosed public spaces, including bars and restaurants, as well as several other current statewide issues. He mails a questionnaire addressing these issues to a simple random sample of 800 voters in his district. Of the 800 questionairres mailed, 152 were returned.

8. The sample is

      A. the 800 voters receiving the questionnaire.

      B. the 152 voters returning the questionnaire.

      C. all voters in his district.

9. The population is

      A. the 800 voters receiving the questionnaire.

      B. the 152 voters returning the questionnaire.

      C. all voters in his district.

10. The Nurses' Health Study has interviewed a sample of more than 100,000 female registered nurses every two years since 1976. The study finds that "light-to-moderate drinkers had a significantly lower risk of death" than either nondrinkers or heavy drinkers. The Nurses' Health Study is

      A. an observational study.

      B. an experiement.

      C. Can't tell without more information.

11. Accident victims are sometimes taken by helicopter from the accident scene to a hospital. Helicopters save time. Do they also save lives? Consider the following hypothetical data.

	Helicopter	Road
Victim died	64	220
Victim survived	136	880
Total	200	1100

   (a) What are the column proportions for the two way table above?

   (b) Display the column proportions using a segmented bar graph. Does it look like helicopters save lives?

12. In the previous problem, we looked a whether helicopters save lives. Now suppose we separate the data above into two groups:

Serious Accidents	Helicopter	Road
Victim died	52	60
Victim survived	48	40
Total	100	100

Less Serious Accidents	Helicopter	Road
Victim died	12	160
Victim survived	88	840
Total	100	1000

From these two separate two way tables, it is clear that helicopters do improve survival odds for both serious accidents (48% versus 40%) and less serious accidents (88% versus 84%). This is an example of what important concept?

13. A study of king penguins looked for a relationship between how deep the penguins dive to seek food and how long they stay underwater. For all but the shallowest dives, there is a linear relationship that is different for different penguins. The study report gives a scatterplot for one penguin titled "The relation of dive duration (DD) to depth (D)." Duration DD is measured in minutes and depth D is in meters. The report then says, "The regression equation for this bird is: $DD = 2.69 + 0.0138D$".

   (a) What is the slope of the regression line? Explain in specific language what the slope says about this penguin's dives.

   (b) According to the regression line, how long does a typical dive to a depth of 100 meters last?

14. Which correlation coefficient is correct for the scatterplot below?

   A. $R = 0.3$ B. $R = 1$ C. $R = -0.5$ D. $R = 0.8$