1. (5 points) You use software to carry out a test of significance. The program tells you that the P-value is $p = 0.031$. You conclude

A. that the probability, computed assuming that H_0 is false, that the test statistic would take a value as extreme or more extreme than that actually observed is 0.031.

B. that the probability, computed assuming that H_0 is true, that the test statistic would take a value as extreme or more extreme than that actually observed is 0.031.

C. that the probability, computed assuming that H_0 is true, that the test statistic would take a value as extreme or less extreme than that actually observed is 0.031.

2. (5 points) You read an article about an experiment in which the researcher conducted a test of significance. The article tells you that the P-value is $p = 0.19$. This means that

A. the probability that the null hypothesis is true is 0.19.

B. neither of the above/below.

C. the value of the test statistic is not particularly large.

3. (5 points) Experiments on learning in animals sometimes measure how long it takes mice to find their way through a maze. The mean time is 18 seconds for one particular maze. A researcher thinks that a loud noise will cause the mice to complete the maze faster. She measures how long each of 10 mice takes with a noise as stimulus. The sample mean is $\bar{x} = 16.5$ seconds. The null hypothesis for the significance test is

A. $H_0 : \mu < 18$.

B. $H_0 : \mu = 16.5$.

C. $H_0 : \mu = 18$.

4. (5 points) A laboratory scale is known to have a standard deviation of $\sigma = 0.001$ gram in repeated weighings. Scale readings in repeated weighings are Normally distributed, with mean equal to the true weight of the specimen. Three weighings of a specimen on this scale give 3.412, 3.416, and 3.414 grams.

You want a 99% confidence interval for the true weight of this specimen. The margin of error for this interval will be

A. smaller than the margin of error for 95% confidence.

B. greater than the margin of error for 95% confidence.

C. about the same as the margin of error for 95% confidence.

5. (20 points) A class survey in a large class for first-year college students asked, "About how many minutes do you study on a typical weeknight?" The mean response of the 463 students was $\bar{x} = 118$ minutes. Suppose that we know that the study time follows a Normal distribution with standard deviation $\sigma = 65$ minutes in the population of all first-year students at this university.

The above survey excluded one response. There were actually 464 responses to the class survey. One student claimed to study 60,000 minutes per night. We know hes joking, so we left out this value. If we did a calculation without looking at the data, we would get $\bar{x} = 247$ minutes for all 464 students. Now what is the 99% confidence interval for the population mean? (Continue to use $\sigma = 65$.)

Compare the new interval with the one which leaves out the outlier. The message is clear: always look at your data, because outliers can greatly change your result.

6. Successful hotel managers must have personality characteristics often thought of as feminine (such as "compassionate") as well as those often thought of as masculine (such as "forceful"). The Bem Sex-Role Inventory (BSRI) is a personality test that gives separate ratings for female and male stereotypes, both on a scale of 1 to 7. A sample of 148 male general managers of three-star and four-star hotels had mean BSRI femininity score x $= 5.29$. The mean score for the general male population is $\mu = 5.19$. Do hotel managers, on the average, differ significantly in femininity score from men in general? Assume that the standard deviation of scores in the population of all male hotel managers is the same as the $\sigma = 0.78$ for the adult male population.

(a) State null and alternative hypotheses in terms of the mean femininity score μ for male hotel managers.

(b) Find the z test statistic.

(c) What is the P-value for your z? What do you conclude about male hotel managers?

7. (20 points) Shelias doctor is concerned that she may suffer from gestational diabetes (high blood glucose levels during pregnancy). There is variation both in the actual glucose level and in the blood test that measures the level. A patient is classified as having gestational diabetes if the glucose level is above 140 milligrams per deciliter (mg/dl) one hour after having a sugary drink. Shelias measured glucose level one hour after the sugary drink varies according to the Normal distribution with $\mu = 122$ mg/dl and $\sigma = 12$ mg/dl.

(a) If a single glucose measurement is made, what is the probability that Shelia is diagnosed as having gestational diabetes?

(b) If measurements are made on 4 separate days and the mean result is compared with the criterion 140 mg/dl, what is the probability that Shelia is diagnosed as having gestational diabetes?

8. **Getting teachers to come to school.** (20 points) Elementary schools in rural India are usually small, with a single teacher. The teachers often fail to show up for work. Here is an idea for improving attendance: give the teacher a digital camera with a tamper-proof time and date stamp and ask a student to take a photo of the teacher and class at the beginning and end of the day. Offer the teacher better pay for good attendanceverified by the photos. Will this work? A randomized comparative experiment started with 120 rural schools in Rajasthan and assigned 60 to this treatment and 60 to a control group. Random checks for teacher attendance showed that 21

(a) Outline the design of this experiment.

(b) Label the schools and choose the first 10 schools for the treatment group. Use Table B starting at line 108.