

Name: \_\_\_\_\_

Score: \_\_\_\_\_ / \_\_\_\_\_

### Homework 3

- 1 The deciles of any distribution are the nine values that divide the sorted data into ten equal parts, so that each part represents 10% of the sample or population. How many standard deviations from the mean are the deciles that mark off the highest and lowest 10% of a normal distribution? \_\_\_\_\_
- 2 The deciles of any distribution are the nine values that divide the sorted data into ten equal parts, so that each part represents 10% of the sample or population. How many standard deviations from the mean are the deciles that mark off the highest and lowest 10% of a t distribution with 10 degrees of freedom? \_\_\_\_\_
- 3 The deciles of any distribution are the nine values that divide the sorted data into ten equal parts, so that each part represents 10% of the sample or population. How many standard deviations from the mean are the deciles that mark off the highest and lowest 10% of a t distribution with 100 degrees of freedom? \_\_\_\_\_
- 4 The deciles of any distribution are the nine values that divide the sorted data into ten equal parts, so that each part represents 10% of the sample or population. How many standard deviations from the mean are the deciles that mark off the highest and lowest 2.5% of a normal distribution? \_\_\_\_\_
- 5 The deciles of any distribution are the nine values that divide the sorted data into ten equal parts, so that each part represents 10% of the sample or population. How many standard deviations from the mean are the deciles that mark off the highest and lowest 2.5% of a t distribution with 10 degrees of freedom? \_\_\_\_\_
- 6 The deciles of any distribution are the nine values that divide the sorted data into ten equal parts, so that each part represents 10% of the sample or population. How many standard deviations from the mean are the deciles that mark off the highest and lowest 2.5% of a t distribution with 100 degrees of freedom? \_\_\_\_\_
- 7 Using an online calculator, you find that your 3 year old daughter has a BMI of 13.9 and are concerned that she may be underweight. Assuming the distribution of BMI of 3 year old girls is normal with mean 15.4 and standard deviation 1.2, the probability that a three year old girl has BMI less than or equal to 13.9 is \_\_\_\_\_.
- 8 At a routine physical examination, you learn that your 3 year old son has a BMI of 18.5 and are shocked to hear the doctor warn you that he is overweight. Assuming the BMI of 3 year old boys is normal with mean 15.6 and standard deviation 1.3, explain why the doctor considers your son to be overweight.
- 9 The 2 hour oral glucose tolerance test is often used to classify individuals into three groups: "healthy" (glucose level  $< 140$  mg/dL), diabetic (glucose level 200 mg/dL or higher), or impaired glucose tolerance without diabetes (glucose levels in between). Suppose you are studying a population in which the distribution of glucose levels is normal with mean 100 mg/dL and standard deviation 30 mg/dL. The probability that a randomly sampled person from your population has a healthy glucose level is \_\_\_\_\_.
- 10 The 2 hour oral glucose tolerance test is often used to classify individuals into three groups: "healthy" (glucose level  $< 140$  mg/dL), diabetic (glucose level 200 mg/dL or higher), or impaired glucose tolerance without diabetes (glucose levels in between). Suppose you are

studying a population in which the distribution of glucose levels is normal with mean 100 mg/dL and standard deviation 30 mg/dL. The probability that a randomly sampled person from this population is diabetic is \_\_\_\_.

11 The 2 hour oral glucose tolerance test is often used to classify individuals into three groups: "healthy" (glucose level < 140 mg/dL), diabetic (glucose level 200 mg/dL or higher), or impaired glucose tolerance without diabetes (glucose levels in between). Suppose you are studying a population in which the distribution of glucose levels is normal with mean 100 mg/dL and standard deviation 30 mg/dL. The probability that a randomly sampled person from this population has impaired glucose tolerance but is not diabetic is \_\_\_\_.

12 An audit of over 5000 (flagged as suspicious) charges to Medicare estimated the average overcharge as \$2150 with 95% CI=(\$2013, \$2287). Mark ALL of the statements below that are correct.

- ☐ A. The probability is 0.95 that the true population mean lies between \$2013 and \$2287.
- ☐ B. In the long run, if we repeated the random sampling and calculated confidence intervals a large number of times, we would know that 95% of those intervals would contain the true population mean. However, for any one sample, we do not know whether the true population mean is included or not.
- ☐ C. If we took a very large number of samples of 5000 suspicious charges and calculated 95% confidence intervals for the mean each time, 95% of those intervals would include the true population mean, and 5% would not.
- ☐ D. There is a 95% chance that this interval contains the true population mean.

13 Using the dataset chinayoung.dta, which contains information on BMI of China Health and Nutrition Survey participants in their 20's, calculate a 95% confidence interval for the BMI of women in this age group using the t distribution. The lower limit of this 95% confidence interval is \_\_\_\_.

14 Using the dataset chinayoung.dta, which contains information on BMI of China Health and Nutrition Survey participants in their 20's, calculate a 95% confidence interval for the BMI of women in this age group using the t distribution. The upper limit of this 95% confidence interval is \_\_\_\_.

15 Using the dataset chinayoung.dta, which contains information on BMI of China Health and Nutrition Survey participants in their 20's, calculate a 99% confidence interval for the BMI of men in this age group using the t distribution. The lower limit of this 95% confidence interval is \_\_\_\_.

16 Using the dataset chinayoung.dta, which contains information on BMI of China Health and Nutrition Survey participants in their 20's, calculate a 99% confidence interval for the BMI of men in this age group using the t distribution. The upper limit of this 95% confidence interval is \_\_\_\_.

17 Using the dataset chinayoung.dta, which contains information on BMI of China Health and Nutrition Survey participants in their 20's, calculate and interpret the 90% confidence interval for the BMI of smokers in this age group using the t distribution.

18 Suppose that in Chapel Hill in October, the daily high temperature follows a normal distribution with mean 72 degrees F and standard deviation 3 degrees F. Suppose that in February, the daily low temperature follows a normal distribution with mean 31 degrees F and standard deviation 5 degrees F. What would be more unusual: having a high temperature of 85 degrees in Chapel Hill on a day in October, or having a low temperature of 10 degrees in Chapel Hill on a day in February? Explain how you obtained your solution based on the data provided.

- 19** Based on the previous data, the probability that the average of the daily high temperatures for Chapel Hill in October will be greater than 74 degrees F is \_\_\_\_.
- 20** Based on the previous data, the probability that the average of the daily high temperatures for Chapel Hill in October is between 71.5 and 72.5 degrees F is \_\_\_\_.
- 21** Suppose you are designing a study of total testosterone levels in men. Suppose that in the population as a whole, total testosterone levels are normally distributed with mean 750 ng/dL and standard deviation 175 ng/dL. You want to estimate the mean total testosterone level in your study sample. How many subjects are needed in your sample if you want to have a sample size for which 95% of the sample averages in similar samples would be within  $\pm 0.5$  standard deviations of the population mean? \_\_\_\_
- 22** Suppose you are designing a study of total testosterone levels in men. Suppose that in the population as a whole, total testosterone levels are normally distributed with mean 750 ng/dL and standard deviation 175 ng/dL. You want to estimate the mean total testosterone level in your study sample. How many subjects are needed in your sample if you want to have a sample size for which 95% of the sample averages in similar samples would be within  $\pm 0.25$  standard deviations of the population mean? \_\_\_\_
- 23** In this population, which event has the lowest probability?
- ☐ A. A group of 10 men with average total testosterone level  $< 600$  ng/dL
  - ☐ B. A group of 500 men with average total testosterone level  $> 775$  ng/dL
  - ☐ C. An individual man having a total testosterone level  $< 200$  ng/dL
  - ☐ D. A group of 1000 men with average total testosterone level  $< 725$  ng/dL