Name: _____

Lab Problems 5-8 (worth 15 pts.)

Statistics 1000 Dr. Nancy Pfenning

5. A student survey was completed by several hundred students in introductory statistics courses at Pitt. Students were asked to pick their favorite color from black, blue, green, orange, pink, purple, red, yellow.

   (a) If colors were equally popular, what proportion (to three decimal places) of students would choose each color?_____

   (b) Pick a color that you suspect will be *less* popular than others. _____Access the survey data and report the sample proportion who preferred the color you chose. _____Is it in fact lower than the proportion you calculated in (a)?

   (c) Use software to produce a 95% confidence interval for the proportion of all students who would choose that color._____

   (d) Does your confidence interval (i) contain the proportion you calculated in (a), or is it (ii) strictly above, or (iii) strictly below?

   (e) Use software to carry out a hypothesis test to see if the sample proportion choosing your color was low enough to assert that, overall, students picked that color less than if they were choosing at random from eight colors. Report the standardized sample proportion z _____and the p-value _____, and state your conclusions, using 0.05 as the cut-off for small p-values:

6. Several hundred students enrolled in introductory statistics courses at a large university were surveyed on various days of the week, and asked to report the number of minutes they'd spent exercising the day before.

   (a) Use software to access the survey data and construct a 95% confidence interval for the mean number of minutes of daily exercise for the population of students.

   _____

   (b) According to the American Time Use Survey published by the Bureau of Labor Statistics, the mean daily hours of exercise for Americans aged 15 and up is about 18 minutes. Use your confidence interval to choose from the following:

      i. 18 is a plausible value for population mean daily amount of exercise
      ii. 18 is not a plausible value for population mean daily amount of exercise: apparently students exercise more
      iii. 18 is not a plausible value for population mean daily amount of exercise: apparently students exercise less

   (c) Use mathematical notation to write null and alternative hypotheses to test if the population mean time for students could be 18 minutes:

7. Students were surveyed as to how many minutes they had exercised the day before. They were also asked whether or not they smoked.

   (a) If one group exercised more, would you expect it to be the (i) smokers or the (ii) non-smokers?

   (b) Use software to access the student survey data, and report the mean minutes exercised for smokers _____and for non-smokers _____. Which sample mean is higher? _____

   (c) Use software to carry out a (one-sided) test of the hypothesis that mean amount of exercise is the same for populations of smoking and non-smoking students. Which of these is relevant? (i) mean of differences (ii) difference between means. Report it. _____Report the p-value. _____
   Should the hypothesis of equality be rejected? _____

8. Several hundred students enrolled in introductory statistics courses at a large university were surveyed, and asked to pick a whole number at random from 1 to 20. Since the mean of the numbers from 1 to 20 is 10.5, if selections are truly random, they should average 10.5 in the long run.

   (a) Tell whether we would opt for a (i) z or (ii) t procedure if population standard deviation were unknown. Tell whether we would opt for a (i) z or (ii) t procedure if we take into account that the standard deviation of the numbers 1 through 20 is 5.766.

   (b) Use software to access the data and, with 5.766 as population standard standard deviation, construct a 95% confidence interval for mean selection by all students. _____

   (c) With 5.766 as population standard deviation, carry out a test to see if the students' random number selections were consistent with random selections from a population whose mean is 10.5: report the sample mean _____and p-value _____, and say whether or not selections could have been truly random: _____

   (d) Would the null hypothesis have been rejected against the one-sided alternative $H_a : \mu > 10.5$? _____Explain:

   (e) Would the null hypothesis have been rejected against the one-sided alternative $H_a : \mu < 10.5$? _____Explain:

   (f) Do people apparently perceive larger or smaller numbers to be more random? Explain:

   (g) Note that the sample standard deviation $s = 5.083$ is smaller than the assumed population standard deviation $\sigma = 5.766$. [This is partly due to the phenomenon that students tend to avoid the extremes 1 and 20 when making a "random" selection.] If $t = \frac{\bar{x}-\mu_0}{5.083/\sqrt{n}}$ had been used instead of $z = \frac{\bar{x}-\mu_0}{5.766/\sqrt{n}}$, would t have been (i) larger than z or (ii) smaller than z or (iii) the same size as z?

   (h) If $t = \frac{\bar{x}-\mu_0}{5.083/\sqrt{n}}$ had been used instead of $z = \frac{\bar{x}-\mu_0}{5.766/\sqrt{n}}$, would the p-value have been (i) larger or (ii) smaller than or (iii) the same size as the one obtained using z?