The University of Texas at San Antonio   Name: _____

Department of Management Science and Statistics

STA 3523, Mathematical Statistics

Instructor: Victor De Oliveira

March 25, 2010

EXAM 2: SOLUTION

Read carefully all the questions and don't spend too much time in any single question.
Show all your work to receive credit.

1. It is found in a random sample of 1000 homes in a city that 228 are heated by oil. Compute a 99%
   confidence interval for the proportion of homes in this city that are heated by oil.   (2 points)

   Answer

   99% CI for p is: $\hat{p} \pm z_{0.005}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.228 \pm 2.575\sqrt{\frac{0.228(0.772)}{1000}} = (0.194, 0.262)$.

2. Fifteen students who intended to major in engineering were compared with fifteen students who
   intended to major in language and literature. The table below shows summaries of the scores on
   the verbal and mathematics portion of the SAT for the two groups of students

	Verbal	**Mathematics**
Engineering	$\bar{x} = 446,\ s = 42$	$\bar{x} = 548,\ s = 57$
Language/literature	$\bar{x} = 534,\ s = 45$	$\bar{x} = 517,\ s = 52$

   Assume the populations are normal and the variances are equal.
   (a) Construct a 95% confidence interval for the difference in mean verbal scores of students majoring
   in engineering and those majoring in language/literature.   (2 points)
   (b) Construct a 95% confidence interval for the difference in mean math scores of students majoring
   in engineering and those majoring in language/literature.   (2 points)
   (c) Interpret the results in (a) and (b).   (1 point)

   Answer

   (a) $S_p^2 = \frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2} = \frac{(14)(1764)+(14)(2025)}{28} = 1894.5$, $S_p = 43.526$

   95% C.I.: $(\bar{x} - \bar{y}) \pm t_{0.025,28}S_p\sqrt{\frac{1}{m} + \frac{1}{n}} = -88 \pm (2.048)(15.893) = (-120.55, -55.45)$

   (b) $S_p^2 = \frac{(14)(3249)+(14)(2704)}{28} = 2976.5$, $S_p = 54.557$

   95% C.I.: $31 \pm (2.048)(19.921) = (-9.80, 71.80)$

   (c) The language/literature majors have better verbal skills than engineering majors, while both majors
   have comparable mathematics skills.

3. An experimenter is concerned that the variability of responses using two different experimental procedures may not be the same. Before conducting his study, he conducts a pre-study with random samples of size 10 and 8 responses and gets $s_1^2 = 7.14$ and $s_2^2 = 3.21$, respectively.

   (a) Find a 90% confidence interval for σ_1/σ_2. (2 points)

   (b) Decide based on your answer in (a) whether the data provide evidence to indicate that the population variances are not equal. (1 point)

   Answer.

   (a) The 90% C.I. for σ_1^2/σ_2^2 is $\left(\frac{s_1^2/s_2^2}{F_{\frac{0.1}{2},10-1,8-1}}, \frac{s_1^2/s_2^2}{F_{1-\frac{0.1}{2},10-1,8-1}} \right) = \left((\frac{7.14}{3.21})\frac{1}{3.68}, (\frac{7.14}{3.21})3.29 \right) = (0.6, 7.32)$.

   So the 90% C.I. for σ_1/σ_2 is $(\sqrt{0.6}, \sqrt{7.32}) = (0.77, 2.71)$.

   (b) Since $1 \in \text{CI}_{0.9}$, the data do not provide evidence indicating the population the variances are unequal.

4. A new additive has been developed for a certain type of cement that results in a comprehensive strength of 5000 kg per cm^2 and a standard deviation of 120. To test the hypothesis that $\mu = 5000$ against the alternative that $\mu < 5000$, a random sample of 50 pieces of cement is tested. The critical region is defined to be $\bar{x} < 4970$.

   (a) Find the probability of making a type I error by using this test. (1.5 points)

   (b) Find the probability of making a type II error by using this test when $\mu = 4960$. (1.5 points)

   (c) Find the test that has probability of type I error equal to 0.01. (1.5 points)

   (d) If the test in (c) is used, what is the necessary sample size to ensure $\beta(4960) = 0.1$. (1.5 points)

   Answer.

   (a) $\alpha = P(\bar{X} < 4970 \text{ when } \mu = 5000) = P(Z < \frac{4970-5000}{120/\sqrt{50}}) = \Phi(-1.77) = 0.038$.

   (b) $\beta(4960) = P(\bar{X} > 4970 \text{ when } \mu = 4960) = P(Z' > \frac{4970-4960}{120/\sqrt{50}}) = 1 - \Phi(0.59) = 0.277$.

   (c) $0.01 = P(\bar{X} < c \text{ when } \mu = 5000) = \Phi(Z < \frac{c-5000}{120/\sqrt{50}})$. Hence $\frac{c-5000}{120/\sqrt{50}} = z_{0.99} = -2.33$, so $c = 4960.46$. The test is to reject $\mu = 5000$ if $\bar{x} < 4960.46$.

   (d) $n = \left(\frac{\sigma(z_{0.1}+z_{0.01})}{\mu_o - \mu'} \right)^2 = \left(\frac{120(1.285+2.33)}{5000-4960} \right)^2 = 117.61$, so the necessary sample size is $n = 118$.

5. A new process for producing synthetic diamonds can be operated at a profitable level only if the average weight of the diamonds is grater than 0.5 karat. To evaluate the profitability of the process, six diamonds are produced with recorded weights $0.46, 0.61, 0.52, 0.48, 0.57$ and 0.54 karat.

   Do these data provide sufficient evidence to indicate that the average weight of the diamonds produced by the process is in excess of 0.5 karat ? Assume these data come from a normal distribution and use $\alpha = 0.1$. (3 points)

   Answer.

   We want to test the hypotheses $H_0 : \mu = 0.5$ vs $H_1 : \mu > 0.5$ at significance level 0.1, so the test should reject H_0 if $T > t_{0.1,5} = 1.476$.

   For these data $\bar{x} = 0.53$ and $s = 0.056$, and the test statistic is $t = \frac{0.53-0.5}{0.056/\sqrt{6}} = 1.312$.

   Since 1.312 is not in the rejection region, the conclusion is not to reject H_0, so the data do not provide sufficient evidence to indicate that the average weight exceeds 0.5 karat.

6. Labels on cans of paint usually indicate the area that can be covered in one coat. Most brands of paint indicate that, in one coat, a gallon will cover between 250 and 500 square feet, depending on the texture of the surface to be painted. One manufacturer, however, claims that a gallon of its paint will cover 400 square feet. To test this claim, a random sample of 10 one-gallon cans of white paint were used to paint ten identical areas using the same kind of equipment. The areas (in square feet) covered by these 10 gallons of paint were $310, 311, 412, 368, 447, 376, 303, 410, 365$ and 350.

   Do these data present sufficient evidence to indicate that the average coverage differs from 400 square feet ? Answer by computing the p-value of the test (more precisely by finding bounds for the p-value of the test). (3 points)

   Answer.

   We want to test the hypotheses $H_0 : \mu = 400$ vs $H_1 : \mu \neq 4000$. For these data $\bar{x} = 365.2$ and $s = 48.417$, and the test statistic is $t = \frac{365.2 - 400}{48.417/\sqrt{10}} = -2.27$.

   Then, p-value $= P(|T| > |-2.27|) = 2P(T > 2.27)$ where $T \sim t_{(9)}$. By looking in the t-table the row corresponding to 9 df, we find that $t_{0.025,9} = 2.262$ and $t_{0.01,9} = 2.821$.

   Hence $2(0.01) <$ p-value $< 2(0.025)$ or $0.02 <$ p-value < 0.05. If we use the popular significance level of 0.05 the conclusion is to reject H_0, so the manufacturer's claim is not supported by these data.

7. Decide whether the following statements are true or false:

   (a) Suppose X has mean 3 and variance 5, and Y has mean 2 and variance 3.
   Then $E(X - Y) = 1$ and $\text{var}(X - Y) = 2$.   **false**

   (b) Consider testing the hypotheses $H_0 : \mu = 1$ vs $H_1 : \mu = 3$. In this case $\beta = 1 - \alpha$.   **false**

   (c) A researcher always makes conclusions from test of hypotheses using a significance level of 0.025. She found that the p-value of a particular test of hypotheses is 0.037.
   Then her conclusion should be not to reject the null hypothesis.   **true**   (3 points)