Exam 3 - STAT 303 Session 201

Summer 2009

Name:

UIN:

Signature:

1. Do not open this test until told to do so.
2. Turn in your exam with your answers circled when you are done with the exam. You should not take the exam with you.
3. This is a closed book examination. You may use six both-sided sheet of formulas that you have brought with you. You should have no other printed or written material with you on the exam. Do not print out the handout as your cheating paper. (I will look everybody's cheating paper.)
4. You have 90 minutes to work on this exam. There are 7 multiple choice questions each worth 5 points and 4 work out questions, worth 25, 5, 15, 20 points.
5. You may use a calculator but not a phone during the exam.
6. If you are unsure of what a question is asking for, do not hesitate to ask the instructor or course assistant for clarification.
7. Do not sit directly next to another student.
8. Good Luck!!!

1. A sanitation department is interested in estimating the mean amount of garbage per bin for all bins in the city. In a random sample of 36 bins, the sample mean amount was 51.5 pounds and the population standard deviation was 4 pounds. (25 points)

   (a) Construct 95% and 99% confidence intervals. (10 Points)

   95% Confidence Interval

$$\bar{X} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 51.5 \pm 1.96\frac{4}{\sqrt{36}} = (50.193, 52.807)$$

   99% Confidence Interval

$$\bar{X} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 51.5 \pm 2.576\frac{4}{\sqrt{36}} = (49.783, 53.217)$$

   (b) Now, a sanitation supervisor is interested in testing to see if the mean amount of garbage per bin is different from 50. Conduct the appropriate hypothesis test using a 0.05 level of significance. (10 points)

   H_0: $\mu = 50$ vs H_1: $\mu \neq 50$. Then, test statistic is

$$z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{51.5 - 50}{4/\sqrt{36}} = 2.25$$

   Then, p-Value is

$$\begin{aligned} \text{p} - \text{Value} &= 2P(Z > |z|) = 2P(Z > 2.25) = 2(1 - P(Z < 2.25)) = 2(1 - .9878) \\ &= 0.0244 \end{aligned}$$

   We reject H_0 since p-Value $< \alpha$. Therefore, the mean amount of garbage per bin is different from 50.

   (c) Using the confidence intervals which calculates in (a) (95%, 99%), conduct the appropriate hypothesis test using 0.05 and 0.01 level of significance. Also, write down the decision would depend on whether we use a 95% or 99% confidence interval. (5 points)

   $\alpha = .05$. Since 50 is not in 95% confidence interval, we reject H_0.
   $\alpha = .01$. Since 50 is in 99% confidence interval, we fail to reject H_0.
   The decision depends on what confidence interval we use.

2. If we are very certain that the true standard deviation of the number of defects per screen is below 3, what sample size would be required so that the width of the 95% confidence interval for the mean number of defects per screen is at most 0.50? Make sure you enter a whole number below. (5 points)

   Calculate the margin of error: $m = .50/2 = 0.25$. Then, the proper sample size is

$$n = \left\lceil \frac{Z_{\alpha/2} \times \sigma}{m} \right\rceil = \left\lceil \frac{1.96 \times 3}{0.25} \right\rceil = 553.19 \approx 554$$

3. A hospital wishes to justify the benefits of nutrition programs for pregnant women using birth weight data from newborns. The hospital hopes to show that the mean birth weight for newborns from mothers who complete the program is higher than the birth weight for newborns from mothers who do not complete the program. A group of 20 pregnant women were randomly divided into two groups; the first group received the nutrition program and the second group did not receive the program. The sample mean and the standard deviation of group 1 are 2617 and 105.78 and the sample mean and the standard deviation of group 2 are 2512 and 80.52. (15 points)

   (a) Construct 95% confidence intervals for $\mu_1 - \mu_2$ without assumption of pooled variance. (5 points)

   Degree of Freedom: min(10 - 1, 10 - 1) = 9
   Then, 95% confidence interval is

$$\bar{X}_1 - \bar{X}_2 \pm t_{.025,9}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (2617 - 2512) \pm 2.262\sqrt{\frac{105.78^2}{10} + \frac{80.52^2}{10}}$$

   (b) Conduct the appropriate hypothesis test at the 0.05 level of significance with assumption of pooled variance. (10 points)

   H_0: $\mu_1 = \mu_2$ vs. H_a: $\mu_1 > \mu_2$. Then, the pooled variance is

$$S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{9 \cdot 105.78^2 + 9 \cdot 80.52^2}{10 + 10 - 2} = 8836.44$$

   Then, the test statistic is

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{(2617 - 2512) - 0}{94.002\sqrt{\frac{1}{10} + \frac{1}{10}}} = 2.497$$

   Then, the p-Value is

$$P(T > t) = P(T > 2.497) = \text{between 0.02 and 0.01}$$

Since p-Value $< \alpha$, we reject H_0. We can conclude that the mean birth weight for newborns from mothers who complete the program is higher than the birth weight for newborns from mothers who do not complete the program.

4. A web based software company is interested in estimating the proportion of individuals who use the Firefox browser. In a sample of 200 of individuals, 31 users stated that they used Firefox.

   (a) Using this data, construct a 95% confidence interval for the proportion of all individuals that use Firefox. (5 points)

   $\hat{p} = \frac{31}{200} = 0.155$. Then 95% confidence interval is

   $$\hat{p} \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.155 \pm 1.96\sqrt{\frac{0.155(1-0.155)}{200}} = (0.105, 0.205)$$

   (b) What sample size would be required so that the width of the 95% confidence interval would be at most 0.08 units wide? (5 points)

   Margin of Error: $0.08/2 = 0.04$. Then, the sample size is

   $$n = \left[\left(\frac{Z_{\alpha/2}}{m}\right)^2 \hat{p}(1-\hat{p})\right] = \left(\frac{1.96}{0.04}\right)^2 0.155(1-0.155) = 314.471 \approx 315$$

   (c) A research report claims that 20% of all individuals use Firefox to browse the web. A software company is trying to determine if the proportion of their users who use Firefox is significantly different from 0.2. Using this data, conduct the appropriate hypothesis test using a 0.05 level of significance. (10 points)

   H_0: $p = 0.2$ vs H_1: $p \neq 0.2$. Then, test statistic is

   $$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0}{n}}} = \frac{0.155 - 0.2}{\sqrt{\frac{.2\times.8}{200}}} = -1.59$$

   Then, p-Value is

   $$\begin{aligned} \text{p} - \text{Value} &= 2P(Z > |z|) = 2P(Z > |-1.59|) = 2(1 - P(Z < 1.59)) = 2(1 - .9441) \\ &= 0.1118 \end{aligned}$$

   We fail to reject H_0 since p-Value $> \alpha$. Therefore, the proportion of Firefox users is not significantly different from 0.2.

5. Select the FALSE statement. (Select Two Answers.)

   (a) To reduce margin of error, we need to reduce σ, increase n, decrease confidence level.

   (b) **Higher confidence levels yields narrower confidence intervals.**

   (c) The null hypothesis (H_0) is the statement of what we usually want to show evidence that the null hypothesis is NOT true.

   (d) The alternative hypothesis (H_a) is the statement of what we want to show is true instead of H_0.

   (e) Hypotheses are always about parameters of populations, never about statistics from samples.

   (f) **95% Confidence interval means that Approximately 95% of the intervals will contain the value μ.**

6. Select the FALSE definition. (Select TWO Answers.)

   (a) A test statistic measures the compatibility between the null hypothesis and the data.

   (b) **The p-value is the probability that we are going to reject H_0.**

   (c) These p-values are exact if the population distribution is normal and are approximately correct for large n in other cases.

   (d) Suppose H_0: $\mu = \mu_0$ and H_a: $\mu \neq \mu_0$. Since μ_0 falls outside $(1-\alpha)100\%$ confidence interval, we will reject the null hypothesis.

   (e) **Statistical inference is valid for all sets of data.**

7. Select the FALSE statement.

   (a) **Power: The probability, that a fixed level α significance test will fail to reject H_1 when a particular alternative value of the parameter is FALSE.**

   (b) To increase power, we need to increase α, increase sample size, or decrease σ.

   (c) **Type I Error**: When we reject H_0 (accept H_a) when in fact H_0 is true.

   (d) **Type II Error**: When we accept H_0 (fail to reject H_0) when in fact H_a is true.

   (e) The power of a fixed level test against a particular alternative is 1 minus the probability of a Type II Error for that alternative.

8. Suppose we want to test $H_0 : \mu = 60$ vs. $H_A : \mu > 60$ and the resulting p-value is 0.089. Which of the following is the correct conclusion?

   (a) We would reject at the 10% level and conclude that the true mean is more than 60.

   (b) We would fail to reject at the 5 and 1% levels and conclude that the true mean is not more than 60.

(c) We would fail to reject at the 5 and 1% levels and conclude that the true mean is no more than (less than) 60.

(d) **A and B are correct conclusions.**

(e) A and C are correct conclusions.

9. Suppose we are trying to test $H_0 : \mu = 3$ vs. $H_A : \mu \neq 3$, where μ is the average number of children a woman thinks a family should have. If we get a 95% confidence interval of (3.3,4.1). What conclusion is appropriate?

   (a) There is significant evidence that the average number of children a woman thinks a family should have is 3.

   (b) There is not significant evidence that the average number of children a woman thinks a family should have is 3.

   (c) **There is significant evidence that the average number of children a woman thinks a family should have is not 3.**

   (d) There is not significant evidence that the average number of children a woman thinks a family should have is not 3.

   (e) None of the above.

10. Suppose a test of $H_0 : \mu = 0$ vs. $H_A : \mu \neq 0$ is run with $\alpha = 0.05$. The p-value of the test is 0.069. If you were to calculate a 95% confidence interval for μ, would the resulting interval contain 0?

    (a) No, because based on the p-value for the hypothesis test we would fail to reject the null, which means that 0 is not a plausible value for μ.

    (b) **No, because based on the p-value for the hypothesis test we would reject the null, which means that 0 is not a plausible value for μ.**

    (c) Yes, because based on the p-value for the hypothesis test we would fail to reject the null, which means that 0 is a plausible value for μ.

    (d) Yes, because based on the p-value for the hypothesis test we would reject the null, which means that 0 is a plausible value for μ.

    (e) There is not enough information to answer this question.

11. An insurance company is conducting a study to justify higher rates for males. What type of hypothesis test should they use?

    (a) a 2-sample test of proportions comparing the proportion of males involved in accidents with that of females

    (b) a 2-sample test of means comparing the average number of accidents for males and females

(c) a 2-sample test of means comparing the average cost of accidents for males and females

(d) **All of the above would provide helpful information.**

(e) It doesn't matter, they're going to charge more anyway!