Math 3339

Homework 9 (7.7 - 7.9)

Name: James Richardson PeopleSoft ID: 1555520

Instructions:

- Homework will NOT be accepted through email or in person. Homework must be submitted through CourseWare BEFORE the deadline.
- Print out this file and complete the problems.
- Use blue or black ink or a dark pencil.
- Write your solutions in the space provided. You must show all work for full credit.
- Submit this assignment at http://www.casa.uh.edu under "Assignments" and choose **Homework 9.**
- Total possible points: **15.**

1. For each of the following scenarios state whether H_0 should be rejected or not. State any assumptions that you make beyond the information that is given.

(a) $H_0 : \mu = 4,\ H_1 : \mu \neq 4,\ n = 15,\ \overline{X} = 3.4,\ S = 1.5,\ \alpha = .05.$
(b) $H_0 : \mu = 21,\ H_1 : \mu < 21,\ n = 75,\ \overline{X} = 20.12,\ S = 2.1,\ \alpha = .10.$
(c) $H_0 : \mu = 10,\ H_1 : \mu \neq 10,\ n = 36,$ p-value $= 0.061.$

a) t = (x – m) / (s / sqrt(n)) = -1.549193
Pt(-1.549193, n-1) = 0.07182004 > .05: Fail to reject the null hypothesis

b) t = (x – m) / (s / sqrt(n)) = -3.629059
Pt(-3.629059, n-1) = 0.0002602994 < .1 : We will reject the null hypothesis

c) we assume that alpha is .05

the p-value of .061 is > .05, therefore we will fail to reject the null hypothesis

2. *Water samples are taken from water used for cooling as it is being discharged from a power plant into a river. It has been determined that as long as the mean temperature of the discharged water is at most 150 F, there will be no negative effects on the river's ecosystem. To investigate whether the plant is in compliance with regulations that prohibit a mean discharge-water temperature above 150 , 50 water sample will be tank at randomly selected times, and the temperature of each sample recorded.

a. Determine an appropriate null and alternative hypothesis for this test.

Null: Ho = If the average temperature of the water given by samples is less that 150 F then the ecosystem is safe.
Alternative: If the average temperature is above 150F then damage to the ecosystem will occur.

b. In the context of this situation, describe type I and type II errors.

Type 1 error would mean the average temperature of the water is above 150 F and we failed to reject it.
Type 2 error would mean that the average temperature is actually below 150F but we believed it to be above and thus rejected the null hypothesis.

c. Which type of error would you consider more serious? Explain.

Type 1 would be more serious, for the ecosystem at least. The temperature of the water will damage the ecosystem but we believed it to be safe.

∗ Problems came from Devore, Jay and Berk, Kenneth, *Modern Mathematical Statistics with Applications*, Thomson Brooks/Cole, 2007.

3. A sample of 12 radon detectors of a certain type was selected, and each was exposed to 100 pCI/L of radon. The resulting readings were as follows:

105.6 90.9 91.2 96.9 96.5 91.3 100.1 105.0 99.6 107.7 103.3 92.4

Does this data suggest that the population mean reading under these conditions differs from 100? Set up an appropriate hypothesis test to answer this question.

Ho: m = 100
Ha: m != 100

X = 98.375
S = 6.109475
N = 12
T = (X - M) / (S/sqrt(N)) = -0.9213828
Alpha = .05

pt(-0.9213828, N-1) = 0.1883081

0.1883081 > .05

Based on a confidence interval of .95 ,we will fail to reject the null hypothesis which states that the average reading is 100 pCI/L.

∗ Problems came from Devore, Jay and Berk, Kenneth, *Modern Mathematical Statistics with Applications*, Thomson Brooks/Cole, 2007.

4. In each of the following situations, for parts a – e write a short explanation of what is wrong and why.

a. A researcher wants to test $H = 30$ versus the two sided alternative $H_a: \neq 30$

The researcher's hypothesis should include the population mean, not the sample mean.

b. A two-sample t statistic gave a P-value of 0.97. From this we can reject the null hypothesis with 95% confidence.

We only reject a null hypothesis is the P-value is less than the alpha value (1 - confidence)

c. A manager wants to test the null hypothesis that average weekly demand is not equal to 100 units.

He needs to completely set up his experiment and include a alternative hypothesis

d. The z statistic had a value of 0.014, and the null hypothesis was rejected at the 5% level because 0.014 < 0.05.

The Z statistic is not what we use to reject or or fail to reject the null hypothesis. We have to calculate the p-value and then compare that to .05

e. Suppose the P-value is 0.03. Explain what is wrong with stating, "The probability that the null hypothesis is true is 0.03."

Because a p-value of .03 actually represents the chances of finding observed values given that the null hypothesis is true.

f. In terms of probability language, a P-value is a conditional probability. Define the event D as "observing a test statistic as extreme or more extreme than actually observed." Using the definition of a P-value, which of the two conditional probabilities represents a P-value:

i. P(H_0 is true | D)

ii. P(D | H_0 is true)

The second, ii, represents the definition of a P-value.

We want to find the probability of finding some value when our null hypothesis is true.

5. *For healthy individuals the level of prothrombin in the blood is approximately normally distributed with mean 20 mg/100 mL and standard deviation 4 mg/100 mL. Low levels indicate low clotting ability. In studying the effect of gallstones on prothrombin, the level of each patient in a sample is measured to see if there is a deficiency. Let μ be the true average level of prothrombin for gallstone patients.

   a. What are the appropriate null and alternative hypotheses?

      Ho: m = 20

      Hs: m < 20

   b. Let $\bar{x}$ denote the sample average level of prothrombin in a sample of $n = 20$ randomly selected gallstone patients. Consider the test procedure with test statistic $\bar{x}$ and rejection region $\bar{x} \leq 17.92$. What is the probability distribution of the test statistic when H_0 is true (i.e. determine center, spread, and shape of $\bar{x}$)? What is the probability of a type I error for this test procedure?

      > pnorm(17.92, 20, 4/sqrt(20)) = 0.01002233

      The probability that the sample average will be less than or equal to 17.92, given that our population average is actually 20, is 1 %.

      To make a type 1 error we must

   c. What is the probability distribution of the test statistic, $\bar{x}$ when $\mu = 16.7$? Using the test procedure of part (b), what is the probability that gallstone patients will be judged not deficient in prothrombin, when in fact $\mu = 16.7$ (a type II error)?

   d. How would you change the test procedure of part (b) to obtain a test with significance level 0.05? What impact would this change have on the error probability of part (c)?

6. *The IQ test scores for first-graders for the entire U.S. population has approximately a normal distribution with mean 100 and standard deviation 15. The following is a random sample of 10 scores from one school:

107	113	108	127	146	103	108	118	111	119

We are interested in seeing whether the population of first graders at this school is different from the national population. Assume that the normal distribution with standard deviation 15 is valid for the school, and test at the 0.05 level to see wither the school mean differs from the national mean. Summarize your conclusions in a sentence about these first graders.

Ho: m = 100
Ha: m != 100

data = c(107, 113, 108, 127, 146, 103, 108, 118, 111, 119)
m = 100
s = 15 # this is the population sd though
Alpha = .05
n = 10
z = (x - m) / (s / sqrt(n))
pnorm(z, m, s) = .321865e-08

Based on a p-value of .321865e-08 and a significance value of .05, we will reject the null hypothesis which stated, that the average IQ score for a 1st grader at this school is 100, in favor of the alternative.

∗ Problems came from Devore, Jay and Berk, Kenneth, *Modern Mathematical Statistics with Applications*, Thomson Brooks/Cole, 2007.

7. *In recent years major league baseball games have averaged three hours in duration. However, because games in Denver tend to be high-scoring, it might be expected that the games would be longer there. In 2001, the 81 games in Denver averaged 185.54 minutes with standard deviation 24.6. minutes. What would you conclude?

```
> #Ho: m = 180
> #Ha: m > 180
> x = 185.54; a = .05; s = 24.6 ; n = 81; m = 180
> t = (x - m) / (s / sqrt(n));
> pt(t, n-1)
[1] 0.9769944
```

Based on an alpha of .05 and a certainty of .95 we will fail to reject the null hypothesis which states that the average game length in Denver is 180 minutes.

* Problems came from Devore, Jay and Berk, Kenneth, *Modern Mathematical Statistics with Applications*, Thomson Brooks/Cole, 2007.

For problems 8 – 10 circle the best answer:

8. A one-sided significance test gives a p-value of .04. From this we can
   a. Reject the null hypothesis with 99% confidence.
   b. Reject the null hypothesis with 96% confidence.
   c. Say that the probability that the null hypothesis is false is .04.
   d. Say that the probability that the null hypothesis is true is .04.

9. Which of the following is not an assumption needed to perform a hypothesis test on a single mean using a z test statistic?
   a. An SRS of size n from the population.
   b. Known population standard deviation,
   c. Either a normal population or a large sample ($n \geq 30$).
   d. The population must be at least 10 times the size of the sample.
   e. All are needed.

10. Which statement about a hypothesis test is true?
   a. The alternative hypothesis says that the observed is due to chance.
   b. With p-value = 0.0001 there is significance.
   c. The test statistic is the probability of getting the observed value or extreme.
   d. None of these are true.

With a p-value that size we could reject most null hypothesizes

∗ Problems came from Devore, Jay and Berk, Kenneth, *Modern Mathematical Statistics with Applications*, Thomson Brooks/Cole, 2007.