Homework 4

Math 243

Due July 14th at 11:59 PM

Textbook Exercises

Chapter 16: 19, 20, 21, 22, 24

Chapter 17: 28, 29, 32, 35

Chapter 18: 28, 31, 33, 35

Chapter 20: 25, 27, 29, 31

Exercise 1: We've been talking about the relationship between confidence intervals and P-values. Let's make that a little more concrete. We can use a confidence interval for a population mean μ to test a null hypothesis H_0 for a subset of the population: we reject the null hypothesis if the subset mean is outside the confidence interval and fail to reject if it is inside. The alternative hypothesis, H_a, is always two-sided in this case, since the confidence interval extends in both directions from μ. For example, a confidence level of $C = .95$ leads to a significance test of $\alpha = .05$, since the interval is incorrect approximately 5% of the time.

a) In class, we had an example of blood pressures of male executives. In an SRS of 72 individuals, we found that $\overline{x} = 128.07$. The standard deviation for the population is $\sigma = 15$. Give a 90% confidence interval for the mean blood pressure μ of all executives in this age group.

We have that the confidence interval is $128.07 \pm 1.645 \dfrac{15}{\sqrt{72}} = 128.07 \pm 2.908$.

b) Let μ_0 be the mean blood pressure of the entire male population (in the appropriate age range — the study was conducted on those between 50 and 59). A hypothetical value for μ_0 is 130. Compare the null hypothesis that $\mu = 130$ to the alternative hypothesis, and show that the test is not statistically significant at the 10% level.

This can be done by the standard procedure, or by the confidence interval method mentioned at the start. To do the confidence interval method, the alternative hypothesis is forced to be two-sided — that is, that $\mu \neq 130$. Since 130 is in the confidence interval from part a), we fail to reject the null hypothesis.

c) Repeat part b), but now suppose $\mu_0 = 131$. Compare the null hypothesis that $\mu = 131$ to the alternative hypothesis, and show that the test **is** statistically significant at the 10% level.

By the same logic, since 131 is outside the confidence interval (by a hair), we reject the null hypothesis.

Exercise 2: A study was conducted on college students to determine if students identifying as male use a different number of words per day than those identifying as female. 20 men were surveyed, and their daily word counts were:

$$28408, 10084, 15931, 21688, 37786,$$

$$10575, 12880, 11071, 17799, 13182,$$

$$8918, 6495, 8153, 7015, 4429,$$

$$10054, 3998, 12639, 10974, 5255$$

Assuming these men are an SRS, can we use t procedures? (Hint: check for outliers and remove any). The mean number of words spoken by men daily is claimed to be approximately 7000. Does this data give statistically significant evidence that the mean number is different from 7000?

Computing the 5-number summary and using the IQR, we find that there is exactly one outlier: 37786. Removing that, we have 19 numbers left, and since there are no more outliers, t-scores can be used. The mean of the 19 numbers is $\overline{x} \approx 11555$, and the standard deviation is $s \approx 6095$. Now we want to run a hypothesis test again. Here, H_0 is that $\mu = 7000$ and H_a is that $\mu \neq 7000$. To find a p-value, we use the t-score table, using row 18 (since $DOF = n - 1 = 19 - 1$). Now the t-score is $t = \dfrac{11555 - 7000}{6095/\sqrt{19}} = 3.257$, and the closest entry on the table is 3.197, which has a two-sided p-value of .005. Since we weren't given a significance level α, we take the default of $\alpha = .05$. Since the p-value is smaller, we reject the null hypothesis.