PID:		
Last Name, First Name:		
a		
Section:		

Homework 6 Math 183, UCSD, Fall 2017 Due on Friday 1st, December 12:50pm Staple pages together

Exercise 1

For each of the following situations, state whether the parameter of interest is a mean or a proportion.

1. A poll shows that 64% of Americans personally worry a great deal about federal spending and the budget deficit.

2. A survey reports that local TV news has shown a 17% increase in revenue between 2009 and 2011 while newspaper revenues decreased by 6.4% during this time period.

3. In a survey, high school and college students are asked whether or not they use geolocation services on their smart phones.

4. In a survey, smart phone users are asked whether or not they use a web-based taxi service.

5. In a survey, smart phone users are asked how many times they used a web-based taxi service over the last year.

The San Diego County is currently experiencing a serious hepatitis A outbreak. Among 271 downtown residents who were tested for this disease in UCSD hospitals, 55 were found to carry it.

1. Create a 90% confidence interval for the percentage of Downtown residents that may carry the hepatitis A virus.

2. What concerns do you have about this sample?

3. If the county wants to cut the margin of error in half, how many residents must be tested?

According to a recent census, 12.4% of all houses in the U.S. are vacant. A county supervisor wonders if her county is different from this. She randomly selects 650 houses in her county and finds that 162 of the them are vacant. What should she conclude? Make sure to: 1) state hypotheses, 2) draw a picture of a sampling distribution, 3) shade a p-value, 4) find the p-value, and 5) state your decision based on the p-value.

In the U.S., the success rate of attaining the General Education Diploma (G.E.D.) is 75.7%.

1. In School A in 2013, there were 300 students that took the G.E.D., among which 258 succeeded. Do these data provide convincing evidence that the success rate of School A is different from the national success rate?

2. In School B in 2013, there were 200 students that took the G.E.D., among which 180 succeeded. Do these data provide convincing evidence that the success rate of School B is higher than the success rate of School A?

In a large university (around 30,000 students enrolled), we asked their personnal information (gender and residential area) to randomly picked students. Out of 860 students, 413 of them were women. Of the 600 students who said they lived on campus, 291 were women. Can it be argued that the difference in the proportion of men and women living on campus is statistically significant? Taking a level of confidence $\alpha = 0.05$, carry out an appropriate analysis to answer this question. Detail carefully your notation and all the steps of your reasoning.

We run a z-test on a mean of interest μ in a population. The z-test is computed using n independent random variables X_1, \ldots, X_n with distribution $N(\mu, \sigma)$. We let $Z = \frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}}$. We want to test the one-sided hypotheses:

$$H_0: \mu = \mu_0, \quad H_A: \mu > \mu_0.$$

We let μ_1 denote the *true* mean if H_A is true. We fix a confidence level α and we compute the critical value z^*_{α} such that $P(Z > z^*_{\alpha}|H_0 \text{ is true}) = \alpha$. (for $\alpha = 5\%$, $z^*_{\alpha} = 1.64$)

1. What is the distribution of Z under H_0 ? And under H_A ?

2. Show that the power $1 - \beta$ of this test can be written as

$$1 - \beta = P\left(Y > z_{\alpha}^* - \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}\right),$$

6

where Y has distribution N(0, 1).

3. Write $\Delta = \mu_1 - \mu_0$ for the difference of means. Study quantitatively the variations of the power of the test when:

(a) α varies;

(b) Δ varies;

(c) n varies.

4. For $\alpha = 5\%$, $\sigma = 1$ and n = 30, find the smallest difference of means Δ that we can differentiate in the test, while guaranteeing a power $1 - \beta \ge 80\%$.

5. For $\alpha = 5\%$ and $\sigma = 1$, what sample size *n* is necessary to differentiate a difference of means $\Delta = 0.1$ (and keep a power $\geq 80\%$)?

Exercise 7

Say you want to test if students with C.S. major, applied math major, environmental systems major, and other majors spend the same amount of time studying for Math 183. What type of test should you use? Explain.

Someone hands you a box of a dozen chocolate-covered candies, telling you that half are vanilla creams and the other half peanut butter. You pick candies at random and discover the first three you eat are all vanilla.

1. If there really were 6 vanilla and 6 peanut butter candies in the box, what is the probability that you would have picked three vanillas in a row?

2. Do you think there really might have been 6 of each? Explain.

3. Would you continue to believe that half are vanilla if the fourth one you try is also vanilla? Explain.

A researcher studies the effect of ten distinct fertilizers on the growth of a plant. He sprouts twelve seeds in each of 10 different Petri dishes containing the different fertilizers, with the exact same quantity of fertilizer in each dish. After three days, he measures the height of the 120 plants, and gets the following boxplots and ANOVA output:



1. What are H_0 and H_A ?

2. What does the ANOVA output say about this testing problem?

3. Do the assumptions for the test seem to be reasonable? Explain

4. His colleague analyses the same data by doing a *t*-test of every fertilizer against every other fertilizer. He finds that several of the fertilizers against every other fertilizer and finds that several of them have significantly higher mean heights. Does this match your finding of question 2? Give an explanation for the difference, if any, between the two results.