University of Texas at Austin

# Problem Set # 17

## Difference in two proportions.

**Problem 17.1.** A simple random sample of 200 students is selected from a large university. In this sample, there are 35 minority students. A simple random sample of 80 students is selected from the community college in the same town. In this sample, there are 28 minority students. What is the standard error of the difference in sample proportions of minority students?

**Problem 17.2.** Suppose that, in our usual notation, $\hat{p}_1 = 0.5, \hat{p}_2 = 0.4, n_1 = 30$ and $n_2 = 40$. What is the $p-$value for testing

$$H_0 : p_1 = p_2 \quad vs. \quad H_a : p_1 \neq p_2.$$

**Problem 17.3.** A simple random sample of 60 households in Whoville is taken. In the sample, there are 45 households that decorate their houses with lights for the holidays.

A simple random sample of 50 households is also taken from the neighboring Whoburgh. In the sample, there are 40 households that decorate their houses.

(i) What is a 95% confidence interval for the difference in population proportions of households that decorate their houses with lights for the holidays?

(ii) If you want to test the hypothesis whether one of the two cities has more festive inhabitants, i.e., whether one of the two cities has a higher proportion of decorated domiciles or not, what $p-$value would you obtain?

**Problem 17.4. Caveat tester.**

It is estimated that 780,000 surgical site infections (SSIs) occur each year. SSIs are the second most common type of healthcare-associated infections in U.S. hospitals and account for an extra $3.5 to $10 billion in healthcare costs per year. The national SSIs rate is 1.9%. A Georgetown medical office was interested in determining if their SSI rate were smaller than the national average. Out of a sample of 277 patients in their study, only one infection occurred.

(i) (1 point) What is the sample size $n$ in this study?

(ii) (1 point) What is the count of SSIs in this study?

(iii) (1 point) What is the observed sample proportion?

(iv) (2 points) What is the name of the approximate distribution of the the sample proportion (according to the CLT)?

(v) (3 points) What is the standard error for $\hat{P}$?

(vi) (3 points) What is the 95% confidence interval for the population proportion parameter?

(vii) (5 points) Test the hypothesis that the Georgetown medical office SSIs rate is less than the national average SSIs rate at the significence level 0.01.