

## STAT 252 – Handout 1 Review of Basic Terms, Ideas

Some very basic, but important, terms and ideas:

- **Variable:** any characteristic of a person or thing that can be assigned a number or a category
  - **Quantitative** variable: measures a numerical characteristic
  - **Categorical** variable: records a category designation
    - **Binary** variable: categorical with only two possible categories
- **Observational unit (case):** Person or thing on which the variable is measured or categorized
- **Population:** *entire* group of people or objects about which information is desired
- **Sample:** a (typically small) *part* of the population from which data are gathered
  - **Sample size:** number of cases (observational units) in the sample
- **Parameter:** numerical characteristic of a *population*
  - Population proportion  $\pi$  (for categorical variable)
  - Population mean  $\mu$  (for quantitative variable)
  - Population standard deviation  $\sigma$  (for quantitative variable)
- **Statistic:** numerical characteristic of a *sample*
  - Sample proportion  $\hat{p}$  (for categorical variable)
  - Sample mean  $\bar{y}$  (for quantitative variable)
  - Sample standard deviation  $s$  (for quantitative variable)

### Example 1: New Year's Eve

I'm interested in the proportion of Cal Poly students who were awake at midnight on New Year's Eve as the year 2009 became 2010. I suspect that this proportion is larger than three-fourths, and I'd like to estimate this proportion with 95% confidence. I'll investigate these questions by asking each of the students in this class whether he/she was awake at midnight on New Year's Eve this year.

- a) Identify the observational units in this study.
- b) Identify the variable in this study, and classify it as quantitative or categorical.
- c) Identify the population and sample in this study.
- d) Identify the parameter of interest, and indicate the symbol used to represent it.
- e) Calculate the relevant statistic, and also indicate the symbol used to represent it.

### Example 2: Holiday Movies

I'm also interested in knowing how many movies the average Cal Poly student has seen (in a theater) this holiday season, say between the weekend before Thanksgiving and this morning. Again I'll ask students in this class to provide their data on this question.

- a) Identify the observational units in this study.
  
- b) Identify the variable in this study, and classify it as quantitative or categorical.
  
- c) Identify the population and sample in this study.
  
- d) Identify the parameter of interest, and indicate the symbol used to represent it.

- **Statistical inference** draws a conclusion about a population parameter based on a sample statistic. The two major procedure types are **confidence intervals** and **hypothesis tests**.
  - A **confidence interval** estimates the value of a parameter with a range of values.
  - A **hypothesis test** assesses the plausibility of a particular claim about the parameter.

As pre-requisite for this course, you have studied four (one-sample) inference techniques:

- $z$ -interval for population proportion  $\pi$
- $z$ -test for population proportion  $\pi$
- $t$ -interval for population mean  $\mu$
- $t$ -test for population mean  $\mu$

### Example 3: New Year's Eve and Holiday Movies (cont.)

For each of the following research questions, identify which of these four inference procedures would be appropriate. Furthermore, if the research question calls for a hypothesis test, state the appropriate null and alternative hypotheses. If the research question calls for a confidence interval, clearly identify the parameter to be estimated.

- a) Were more than three-fourths of all Cal Poly students awake at midnight on New Year's Eve?
  
- b) What proportion of Cal Poly students were awake at midnight on New Year's Eve?

c) Have Cal Poly students seen an average of more than 1.5 movies in a theater this holiday season?

d) What is the average number of movies in a theater that Cal Poly students have seen this holiday season?

**Example 4: More Miscellany**

For each of the following research questions, identify which of these four inference procedures would be appropriate. Furthermore, if the research question calls for a hypothesis test, state the appropriate null and alternative hypotheses. If the research question calls for a confidence interval, clearly identify the parameter to be estimated.

a) What percentage of people wash their hands after using a public rest room?

b) How much, on average, is spent per transaction at *The Avenue*?

c) Are more than half of all college baseball games won by the home team?

d) What proportion of full-time Cal Poly students have at least one class on Friday?

e) Does a typical Cal Poly student study for 20 hours per week?

f) Do 70% of all new restaurants go out of business within the first year?