Intro to Social Science Data Analysis

Seminar 9: Overview of Statistical Inference (II)

Christopher Gandrud

October 15, 2012

Download Data from OpenIntro

```r
# Download data
download.file("http://bit.ly/PO9XsF",
              destfile = "ames.RData")

## Warning:  downloaded length 19027 != reported
length 146692


# Load Data
load("ames.RData")

## Error:  error reading from connection
```

Take a random sample

```
# Find number of observations
nrow(ames)

## Error:  object 'ames' not found


# Take a random sample of 100 observations
amesSamp <- ames[sample(1:nrow(ames), 100,
                        replace=FALSE),]

## Error:  object 'ames' not found


# Find number of observations in sample
nrow(amesSamp)

## Error:  object 'amesSamp' not found
```

Pick 3 continuous variables and formulate null and alternative hypotheses for testing whether the sample is representative of the population.

Test these hypotheses with 95% and 99% confidence intervals.

Hypothesis Tests

Test these hypotheses p-values using the 95% and 99% significance levels.

How do the results from the p-value tests correspond to the confidence interval tests?

Subset the one of your variables by the variable `Central.Air`.

It has the values `Y` and `N`.

Is there a significant difference in the mean value of the variable for houses with air conditioning and without air conditioning?