Week One

Help Center

Review of Basic Statistics and Introduction to Regression

During our first week together, we will review some basic statistical concepts including bias, confidence intervals and p-values. We will also begin our introduction into basic regression concepts and correlation. We will end our week with an introduction to STATA statistical software.

There is no homework for this first week, however, please be sure to check out the **Download STATA** page and check out some of the tutorial websites listed there to begin familiarizing yourself with STATA. You'll have the opportunity to install STATA later on during this week.

*Please be advised that you will be **unable** to install STATA without the software license code until after we email them on or around Thursday, March 26. They will be sent to the email address you use for logging into Coursera.*

Lectures

Please click on the links below to access the video lectures for this first week

- **Introduction and Course Highlights**
- **Central Tendency and Variability**
- **Sampling Distribution**
- **Bias**
- **Confidence Intervals**
- **p-Values**
- **Regression and Correlation**
- **Introduction to STATA**

Lecture Material

Please click on the link below to download the slides of the first week

Week One "Review of Basic Statistics and Introduction to Regression"

Conversations

Please join in the conversations around regression analysis in our **community forums** area. You can ask and answer questions and discover insights and help for yourself and others as we come together to encourage each other in our exploration.

Key Terms

Below are definitions of some important terms covered this week:

- **Degrees of Freedom**: The number of values in the calculation of a test statistic that are free to vary. For example, imagine a set of 5 test scores whose mean is 90. You are 'free' to pick the value of the first four test scores. However, once you know the first four values, the final value is fixed. In this example you have 4 degree of freedom, because you were 'free' to choose the first four scores, at which point the final score is fixed.
- **Central Limit Theorem**: A probability theorem which states that the mean of a large number of observations will be approximately normally distributed.
- **Total Sum of Squares**: The total sum of squares is a measure of the total variability of a set of scores around the mean of the dependent variable (outcome).
- **Variance**: A measure of the spread of data, or a measure of variability in data. You can calculate the variance by averaging the squared deviations of each observation in comparison to the mean.
- **Slope of Regression Line**: Coefficient of a predictor/independent variable in a regression model. The slope indicates how the outcome changes with increases in the predictor value. This slope can be interpreted as 'for every unit increase in predictor X, Y will increase or decrease by (slope value)'.

Homework

There is no homework this week.

Please visit the **Download STATA** page to learn more about the statistical software that will be demonstrated in our course homework starting in week two.

You can visit the **homework page** to also learn more.

Quiz

After you've gone through the materials for this week please be sure to visit the **quizzes area** to complete this week's quiz.

Created Mon 2 Feb 2015 11:48 AM PST

Last Modified Mon 30 Mar 2015 8:39 AM PDT