

Lecture 1

First Meeting & Introduction to Statistics and Data Analysis I

Lecture Information

Ceng272 *Statistical Computations* at February 15, 2012

First Meeting

Lecture Information

Overview

Text Book

Grading Criteria & Policies

Introduction to Statistics and Data Analysis

Overview

Dr. Cem Özdoğan
Computer Engineering Department
Çankaya University

Contents

1 First Meeting

Lecture Information

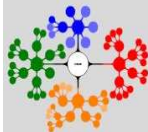
Overview

Text Book

Grading Criteria & Policies

2 Introduction to Statistics and Data Analysis

Overview



First Meeting

Lecture Information

Overview

Text Book

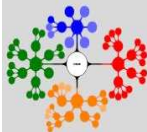
Grading Criteria & Policies

Introduction to Statistics and Data Analysis

Overview

First Meeting

- CENG 272 Statistical Computations Spring 2012
- WEDNESDAY 13:40-15:30 H-127
- THURSDAY 12:40-13:30 H-127
- Instructor: Cem Özdoğan Materials Science and Engineering Department, N-B08
- TA:
- WEB page:
<http://siber.cankaya.edu.tr/ozdogan/StatisticalComputations/spring2012/index.html>
- Announcements: Watch this space for the latest updates.
15.Subat.2012 10:27 In the first lecture, there will be first meeting and Introduction to Statistics and Data Analysis. Lecture notes for the other weeks are published, see Course Schedule section.
- All/Some the example Minitab files will be accessible via the link ???????.



First Meeting

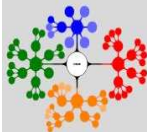
Lecture Information
Overview
Text Book
Grading Criteria & Policies

Introduction to Statistics and Data Analysis

Overview

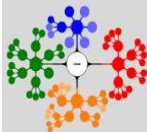
Lecture Information

- There are two groups for lecturing, you may attend any one of the lecture hours.
- We have a pop-up quiz policy.
- You MAY have quizzes (10-15 minutes, may be less; but not scheduled as before) for the previous lecture/chapter's subjects.
- There won't be any make-up for these quizzes.
- They will have weight for the final grading.
- Important announcements will be posted to the Announcements section of the web page, so please check this page frequently.
- You are responsible for all such announcements, as well as announcements made in lecture.



Overview

- Ceng 272 is intended as a calculus-based probability and statistics course and gives an introduction to probability theory and statistics that emphasises the mathematical foundations required to understand probability models and statistical methods.
- Among the topics covered will be;
 - To introduce various statistical methodologies and their applications in data analysis.
 - To provide students with a good understanding of the theory of probability and the basic methods of statistics.
 - Data, Graphical Descriptive Techniques. Use statistical concepts such as means, variances and various types of graphs to analyze datasets.
 - Understand basic concepts in probability including sample spaces, events, probability rules, conditional probability, independent events and Bayes rule.
 - Random variables and their distributions.
 - Sampling distributions - Sampling, law of large numbers, central limit theorem.
 - Understand confidence intervals and perform statistical inference such as hypothesis testing and regression??



Text Book

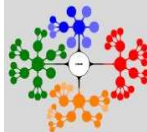
- **Required:** Readings will be assigned in Probability & Statistics for Engineers & Scientists, 8/E by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, Prentice Hall, 2006.
- **Recommended:** Statistics for Engineering and the Sciences, 5/E by William Mendenhall, Terry Sincich, Prentice Hall, 2006. This book is useful for reference, for an alternative point of view.



Required

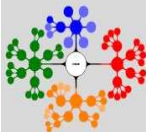


Recommended



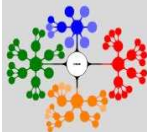
Grading Criteria & Policies

- There will be a midterm and a final exam, will count 25% and 45% of your grade, respectively.
- Quiz: 15% (worst of the quizzes will be discarded).
- Attendance is required and constitutes part of your course grade; 15%. Attendance is not compulsory, but you are responsible for everything said in class.
- I encourage you to ask questions in class. You are supposed to ask questions. Don't guess, ask a question!
- The code/homework you submit must be written completely by you. You can use anything from the textbook/notes with a clear understanding.



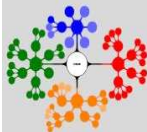
Computational Statistics vs Statistical Computation

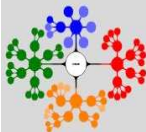
- Computational statistics is statistics that uses computation, often a lot of computation,
- While statistical computing is computational mathematics.
- Both fields are hybrids, but it seems that
 - in computational statistics the emphasis is on the statistics,
 - in statistical computation the emphasis is on the tools.
- Statistics means numerical descriptions to most people.
- Statistics does not just mean numerical fact, collection or summarization of the data.
- It also means that the whole body of systematic methods for collecting, summarizing, analyzing and interpreting data, and to make **decisions**.



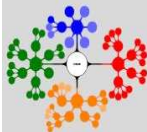
Success of Statistical Methods. Use of scientific data.

- In recent decades, the statistical methods have been successfully used in
 - Manufacturing,
 - Development of food products,
 - Computer software,
 - Pharmaceuticals,
 - Many other areas,
- Examples: Japanese industrial miracle; High-quality products. Improvement of quality in American industry
- The use of statistical methods involves the gathering of information of scientific data.
- Statistical methods are designed to make scientific judgements in the face of uncertainty and variation.
- Two types of statistics:
 - 1 Inferential statistics
 - 2 Descriptive statistics

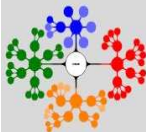




- *Inferential statistics* are the mathematical procedures hereby we convert information about the sample into intelligent guesses about the population.
- **Population.** A collection of all elements whose characteristics are being study. By countability, it can be either finite or infinite.
- **Sample.** A properly selected subset of the population, which is used to gain insight about the population.
 - (Simple) Random Sample: A sample drawn in such a way that each element of the population has an equal chance of being selected.
- **Parameter.** The characteristic(s) or attribute(s) of the population, which is generally unknown. e.g. μ (population mean).
- **Statistic.** The characteristic of the sample, which is a function of the data. It is usually numerical facts. e.g. \bar{x} (sample mean).



- Example: a manufacturer of computer boards wish to eliminate defects
 - 50 computer boards are sampled from the process to collect information.
 - Here, the population is all computer boards manufactured over a specific period of time.
- It utilizes sample data to make estimates, decisions, predictions, or other generalizations about larger set of data.
- This is also called *inductive reasoning* or *inductive statistics*.
 - *Deduction* - From the general facts to particular cases: Mathematics.
 - *Induction* - From a particular case to a general one: Statistics.



- *Descriptive statistics* consists of methods for organizing, displaying, and describing data (to convey information) by using tables, graphs, and summary measures.
 - Descriptive statistics are just descriptive. They do not involve generalizing beyond the data at hand.
 - It utilizes graphical and numerical methods to look for patterns, to summarize, and to present the information in a set of data.
 - Descriptive statistics are central to the professional sports and other areas.
- With descriptive statistics, we are simply describing what the data shows.
- With inferential statistics, we are trying to make inferences from our data to more general conditions.