

Statistics 528

- Introduction to the Course
- Class Survey
- Syllabus
- Course Website
- Overview
- Chapter 1: Introduction

Overview

- Chapter 1: Data Distributions
- Chapter 2: Looking at Relationships b/t Data
- Chapter 3: Experimental Design and Sample Surveys
- Chapter 4: Probability
- Chapter 5: Sampling Distributions
- Chapter 6: Introduction to Statistical Inference
- Chapter 7: More Statistical Inference

Chapter 1: Introduction

- Data sets are organized in terms of
 - Individuals** - objects described by a set of data. Individuals can be people, animals, things, etc.
 - Variables** - characteristics of individuals. A variable can take on different values depending on the individual it is describing.
- Questions to ask about data:
 1. What is the purpose of the data? Do we hope to answer some specific questions using the data?
 2. What individuals do the data describe? How many individuals appear in the data set.
 3. How many variables are there? What are the exact definitions of these variables and in what **units of measurement** is each variable recorded?

-
- **Categorical variables** place individuals into one or several groups or categories.
 - **Quantitative variables** take numerical values for which arithmetic operations such as adding and averaging make sense.
 - The **distribution** of a variable tells us what values it takes on and how often it take these values. Chapter 1 discusses graphical and numerical ways of describing distributions.

Example:

Data from a medical study contain values of many variables for each of the people who are subjects of the study. Which of the following variables are categorical and which are quantitative?

- a) Gender (female or male)
- b) Age (years)
- c) Race (Asian, black, white, or other)
- d) Smoker (yes or no)
- e) Systolic blood pressure (millimeters of mercury)
- f) Level of calcium in the blood (micrograms per millimeter)

Introduction to MINITAB

- Opening MINITAB: blank session and worksheet
- Two types of MINITAB files:
 - file.MPJ – MINITAB project (session, worksheet, and graphs)
 - data.MTW – MINITAB worksheet
- Data sets can be found in
 - C:\Program Files\MTBWIN\Data\
 - On IPS CD or the text's website
(<http://bcs.whfreeman.com/ips4e/default.asp>)
- Importing data: **Open > Open Worksheet**
Example: Poplar1.MTW

- Story: Researchers at Penn State Univ. planted Poplar Clone 252 at two different sites and measured the diameter (cm), height (m), and dry weight of the wood (kg) from a sample of 3-year-old trees. (Taken from Minitab tutorial.)
- Variables: Yield (C1), Height (C2), and Weight (C3)
Cases: 1-15
- Adding new data: go to row 16 and point the arrow to the right

1.52	[Enter]	2.9	[Enter]	.07	[Ctrl]+[Enter]
4.51	[Enter]	5.27	[Enter]	.79	[Ctrl]+[Enter]
1.18	[Enter]	2.2	[Enter]	.03	[Ctrl]+[Enter]
3.17	[Enter]	4.93	[Enter]	.44	[Ctrl]+[Enter]
3.33	[Enter]	4.89	[Enter]	.52	[Ctrl]+[Enter]

Statistics 528 - Lecture 1
Prof. Kate Calder

7

- Saving:
 - Save worksheet as work.MTW. **File > Save Current Worksheet As**
 - Save project as work.MPJ. (Recommended!!) **File > Save Project As**
- Adding new variable: Soil Type (store in C4)
 - 1 - well-drained soil (first 10 measurements)
 - 2 - sandy soil (last 10 measurement)

Type 1 into cell (1,4). Turn the cursor into a small cross in the lower right of cell (1,4) and drag it down 10 rows.
Similarly, add 2's to C4 rows 11-20.
- Using the calculator to create a new variable: D2H (store in C5)
 - **Calc > Calculator**
 - **Store Results in: C5**
 - **Expression: C1**2*C2**

Statistics 528 - Lecture 1
Prof. Kate Calder

8