Class Notes
for
Elementary Statistics
Letu Math–1423

Nicholas Capo
nicholas.capo@gmail.com

September 10, 2012
10:41am

Contents

This document comprises classroom notes from Statistics Class
at LeTourneau University, in the Fall of 2012.

Although the author will attempt to be complete and correct in these notes, it is the reader's responsibility to learn and understand the material. The author assumes no responsibility for the completeness or accuracy of this content.

If you have any suggestions or corrections feel free to email the author at
nicholas.capo@gmail.com

The latest version of this document is available at:
`https://bitbucket.org/nicholascapo/statisticsnotes/src/tip/StatisticsNotes.pdf`

Chapter 1

Introduction

Definition of Statistics
"Statistics is the science of collecting, organizing, analyzing, and interpreting data in order to make decisions."

1.1 Data

1.1.1 Data Sets

Population The collection of all outcomes, responses, measurements, or counts, that are of interest.

Sample A subset of the population.

Parameter A number that describes a population characteristic.

Statistic A number that describes a sample characteristic.

1.1.2 Types of Data

Qualitative Data Attributes, labels, or non-numerical entries.

Quantitative Data Numerical measurements or counts.

1.2 Sample Mean and Median

1.2.1 Definition

Sample Mean The average of the sample data points, however it may not be a data point.

$$\overline{x} = \sum_{i=1}^{n} \frac{x_i}{n} = \frac{x_1 + x_2 + x_3 \cdots x_n}{n}$$

Sample Median The middle value of the data.

$$\tilde{x} = \begin{cases} x_{(\frac{n+1}{2})} & \text{if } n \text{ is odd} \\ \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) & \text{if } n \text{ is even} \end{cases}$$

Trimmed Mean A trimmed mean is computed by trimming off the largest and smallest set of values. For example a 10% trimmed mean is found by eliminating the largest 10% and smallest 10% and computing the mean of the remaining values. This may be useful for data that contains possible outliers. Denoted by $x_{tr(\text{percent})}$

1.3 Measures of Variability

1.3.1 Standard Deviation

Sample Variance

$$s^2 = \sum_{i=1}^{n} \frac{(x_i - \overline{x})^2}{n-1}$$

Sample Standard Deviation

$$s = +\sqrt{s^2}$$

The standard deviation is 0 when all the data points are the same.

1.4 Descriptive Statistics

1.4.1 Quartiles

Quartiles approximately divide an ordered data set into four equal parts.

First Quartile, Q_1 About 25% of the data fall on or below Q_1

Second Quartile, Q_2 About 50% of the data fall on or below Q_2

Third Quartile, Q_3 About 75% of the data fall on or below Q_3

1.4.2 Range and Interquartile Range

Range

$$\text{range} = \text{max value} - \text{min value}$$

Interquartile Range

$$IQR = Q_3 - Q_1$$

To help find outliers, compute $1.5 \times IQR$, and any values that lie outside the interval $[Q_1 - 1.5 \times IQR, Q_3 + 1.5 \times IQR]$ is a possible (and probable) outlier.

1.4.3 Box and Whisker Plot

Exploratory Data Analysis Tool

- Requires
  - Min
  - Q_1
  - Median
  - Q_3
  - Max

Example

Example Data	[1, 2, 3, 4, 5, 6, 11]
Min	1
Median	4.0
Max	6
Outlier	11

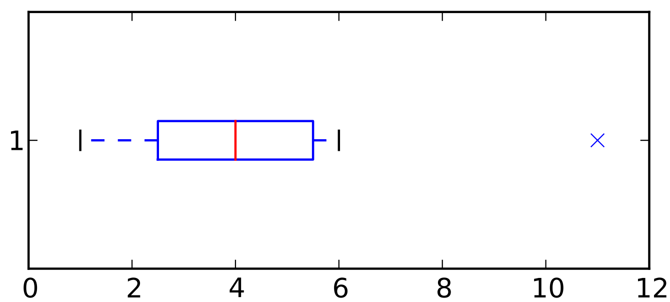

Figure 1.1: Example Box And Whisker Plot

1.5 Stem and Leaf Plots

These look like a sideways histogram
Data: [31, 21, 32, 33, 41, 42, 58, 25, 21]

Stem	Leaf	Key: $a\|b = ab$
2	1,1,5	
3	1,2,3	
4	1,2	
5	8	

1.5.1 Key Notation

Key: 4—5 = 45 Key: 4—5 = 4.5

1.5.2 Double Stem and Leaf

Separate the leaves into two groups, (0-4, and 5-9)

Data: [31, 21, 32, 33, 41, 42, 58, 25, 21]

Stem	Leaf Key: $a|b = ab$
2	1,1
2	5
3	1,2,3
4	1.2
4	
5	
5	8

1.6 Frequency Distribution

A table that shows classes or intervals of data with a count of the number of entries in each class.

1.6.1 Midpoint of a Class

Average of the class limits.

$$\frac{\text{(lower class limit)} + \text{(upper class limit)}}{2}$$

1.6.2 Relative Frequency

$$\frac{\text{class frequency}}{\text{sample size}} = \frac{f}{n}$$

1.7 Scatter Plots

Each entry in one data set corresponds to one entry in a second set, one-to-one mapping.

1.7.1 Example Scatter Plot

Data:	
X:	`[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]`
Y:	`[3, 1, 6, 11, 2, 7, 11, 8, 12, 12, 11, 1]`

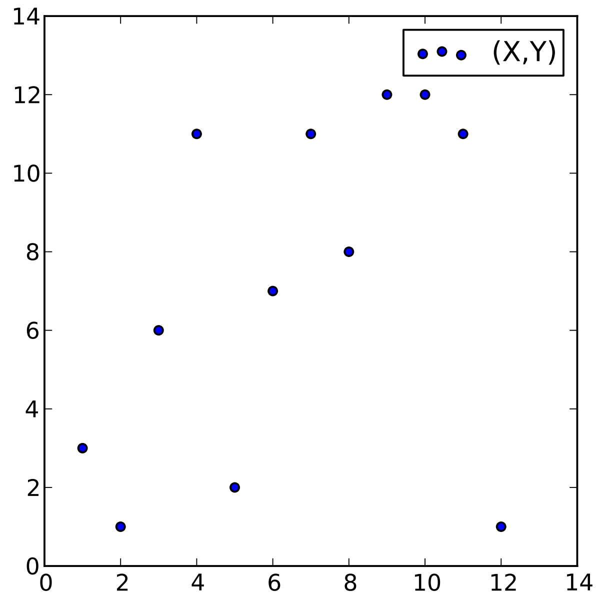

Figure 1.2: Example Scatter Plot

1.8 Homework

- Page 13 #'s 1.5, 1.6
- Page 17 #'s 1.11, 1.12
- Page 31 #'s 1.18, 1.19, 1.20, 1.29, 1.30

Chapter 2

Probability

2.1 Experiments

Any process that generates a set of data.

2.2 Sample Space

The set of all possible outcomes of a statistical experiment, denoted S. The sample space with no elements is the empty set or null set, denoted $\emptyset$

2.2.1 Example

$$S = \{3, 2, 1, 0\}$$

$$S = \{x|0 < x < 25\}$$

$$S = \{x^2|x \in \mathbb{R}\}$$

2.2.2 Tree Diagrams

A Tree Diagram can be used to list all possible outcomes

2.2.3 Events

An event is a subset of a sample space. The null set ($\emptyset$) and the sample space (S) are both subsets of the sample space S.

Intersection

The intersection of two events A and B, denoted $A \cap B$, is the event containing all elements that are common to A and B. If $A \cap B = \emptyset$ than A and B are called mutually exclusive or disjoint.

Union

The union of two events A and B, denoted $A \cup B$, is the event containing all elements that belong to A or B or both.

Compliment

The compliment of an event A with respect to S is a subset of all elements of S not in A, denoted A'

2.3 Counting Sample Points

2.3.1 Multiplication Rule

If an operation can be preformed in n_1 ways and if for each of the ways a second operation can be preformed in n_2 ways, then the two operations can be preformed together in $n_1 n_2$ ways. This principle can be extended to more than two operations. See Example 2.14 in Walpole et al. [1, p. 45]

2.3.2 Factorial

For any non-negative integer n, $n!$ called "n factorial", is defined as

$$n! = n(n-1)\cdots(2)(1)$$

with the special case $0! = 1$.

2.3.3 Permutation

A permutation is an arrangement of all or a part of a set objects. For permutations the order of objects matters. The number of permutations of n distinct objects is $n!$.

2.3.4 Permutations at a Time

The number of permutations of n distinct objects taken r at a time is

$$_nP_r = \frac{n!}{(n-r)!}$$

This is called "n Permute r"

2.3.5 Permutations in a Circle

The number of permutations of n objects arranged in a circle is $(n-1)!$.

2.3.6 Permutations of a Kind

The number of distinct permutations of n objects of which n_1 are of one kind, n_2 of a second kind, ..., n_k of a kth kind is

$$\frac{n!}{n_1!n_2!\cdots n_k!}$$

2.3.7 Partitioning

The number of way of partitioning a set of n objects into k cells with n_1 elements in the first cell, n_2 elements in the second cell, and so forth, is

$$\binom{n}{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1!n_2!\cdots n_k!}$$

Notice that this is the same as the last example.

2.4 Homework

- Page 42 #'s 2.3, 2.6, 2.10, 2.11, 2.14, 2.16, 2.18
- Page 51 #'s 2.24, 2.25, 2.28, 2.29, 2.30, 2.33, 2.34, 2.38, 2.39, 2.40, 2.44, 2.46, 2.47, 2.48

Homework Overview

Bibliography

[1] R.E. Walpole et al. *Probability and Statistics for Engineers and Scientists*. Pearson Education, 2010. ISBN: 9780321629111. URL: http://books.google.com/books?id=tzZxRQAACAAJ (cit. on p. 14).