Class Notes
for
Statistics
Letu Math–3403

Nicholas Capo
nicholas.capo@gmail.com

October 5, 2012
10:35am

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 document.

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

You may also view the LaTeX source code at
https://bitbucket.org/nicholascapo/statisticsnotes

Contents

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

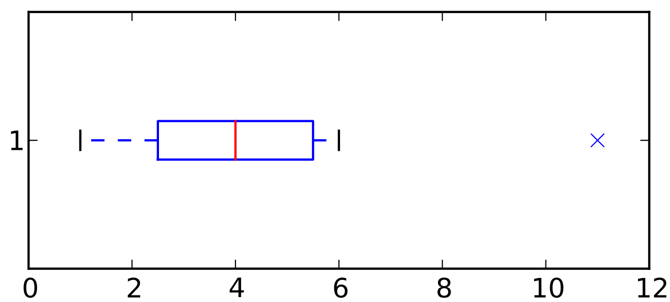

Figure 1.1: Example Box And Whisker Plot

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:	`[12, 7, 6, 9, 9, 3, 12, 6, 10, 12, 9, 7]`

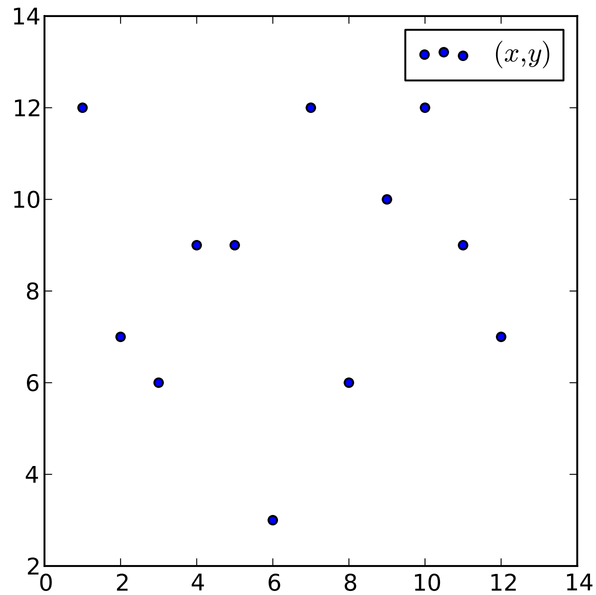

Figure 1.2: Example Scatter Plot

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

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

Note: 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

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

Where $n_1 + n_2 + \cdots + n_k = n$

Note: This is the same as the last example

2.3.8 Combinations

The number of combinations of n distinct object taken r at a time, is:

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

Note: This is partitioning with only two cells

2.3.9 Partitioning and Combinations

Note that:

$$\binom{10}{5,4,1} = \binom{10}{5}\binom{5}{4}\binom{1}{1}$$

2.4 Probability of an Event

The probability of an event A is the sum of the weifhts of all sample points on A. Therefore

$$0 \leq P(A) \leq 1$$

$$P(\emptyset) = 0$$

$$P(S) = 1$$

Furthermore, if $A_1, A_2, \cdots$ is a set of mutually exclusive events, then

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$$

2.4.1 Complimentary Probability

If A and A' are complimentary events, then

$$P(A) + P(A') = 1$$

2.4.2 Different Outcomes

If an experiment can result in any one of N different equally likely outcomes, and of exactly n of those outcomes correspond to event A, the probability of event A is

$$P(A) = \frac{n}{N}$$

2.4.3 Additive Rule of Probability

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

2.5 Conditional Probability

The conditional probability of B given A, denoted $P(B|A)$ is defined as

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \text{ where } P(A) \neq 0$$

Note: This is a filter

Alternate Notation

$$P(A|B) = \frac{n(A \cap B)}{n(B)}$$

Where $n(x)$ is defined as "the number of elements in x"

2.5.1 Multiplication Rule

If in an experiment the events A and B can both occur, then

$$P(A \cap B) = P(A)P(B|A), \text{ provided } P(A) > 0$$

2.5.2 Independent Events

Two events A and B are *independent* if and only if

$$P(B|A) = P(B) \text{ or } P(A|B) = P(A)$$

Assuming the existence of the conditional probabilities. Otherwise, A and B are dependent.

Note: Independence does not imply Mutual Exclusivity!

2.5.3 Corollary

Two events A and B are *independent* if and only if

$$P(A \cap B) = P(A)P(B)$$

Therefore, to obtain the probability that two independent events will occur, we simply find the product of their individual probabilities.

2.5.4 Compliments of Independent Events

If A and B are independent events so are the complements of these events, this means that:

- A' and B' are independent
- A' and B are independent
- A and B' are independent

Note: Also $P(A'|B) = 1 - P(A|B)$ (no need for independence)

2.5.5 Extended Multiplication for Conditional Probability

If, in an experiment, the events $A_1, A_2 \cdots A_k$ can occur, then

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1) \times P(A_2|A_1) \times P(A_3|A_1 \cap A_2) \times \cdots \times P(A_k|A_1 \cap A_2 \cap \cdots \cap A_{k-1})$$

2.6 Bayes' Rule

2.6.1 Partitioning the Sample Space

A sample space A can be partitioned into an arbitrary number of subsets. For partitions B_1, B_2, B_3 then

$$A = (B_1 \cap A) \cup (B_2 \cap A) \cup \cdots \cup (B_3 \cap A)$$

2.6.2 Theorem of Total Probability

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i)P(A|B_i)$$

2.6.3 Bayes Rule

$$P(B_i|A) = \frac{P(B_i \cap A)}{\sum_{i=0}^{k} P(B_i \cap A)} = \frac{P(B_i \cap A)}{\sum_{i=0}^{k} P(B_i)P(A|B)}$$

Chapter 3

Random Variables

3.1 Concept of a Random Variable

3.1.1 Definition

A random variable, denoted X, is a function that associates a real number with each element in the sample space.

3.1.2 Discrete Sample Space

If a sample space contains a finite number of possibilities or an unending sequence with as many elements as there are whole numbers, it is called a *discrete sample space.*

A random variable corresponding with a discrete sample space is called a *discrete random variable.*

3.1.3 Continuous Sample Space

If a sample space contains an infinite number of possibilities equal to the number of points on a line segment, it is called a *continuous sample space.*

A random variable corresponding with a continuous sample space is called a *continuous random variable.*

3.2 Discrete Probability Distribution

The set of ordered pairs $(x, f(x))$ is a *probability function*, *probability mass function*, or *probability distribution* of the discrete random variable X if, for each possible outcome x,

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$

3. $P(X = x) = f(x)$

Note: Probabilities are Functions!

3.2.1 Cumulative Distribution Function

The Cumulative Distribution Function $F(x)$ of a discrete random variable$X$ with probability distribution $f(x)$ is

$$F(x) = P(X \leq x) = \sum_{t \leq x} F(t), \text{ for } -\infty < x < \infty$$

Note: This keeps a running total of the cumulative probabilities up to a value

3.3 Continuous Random Variables

3.3.1 Probability Density Function

The function $f(x)$ is called a probability density function or pdf for the continuous random variable X, defined over the set of real numbers, if

1. $$F(x) \geq 0, \text{ for all } x \in R$$

2. $$\int_{-\infty}^{\infty} f(x)\mathrm{dx} = 1$$

3. $$P(a < X < b) = \int_{a}^{b} f(x)\mathrm{dx}$$

Note: $P(X = c) = 0$, where $a \leq c \leq b$

3.3.2 Cumulative Distribution Function

The Cumulative Distribution Function $F(x)$ of a continuous random variable X with density function $f(x)$ is

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\ \mathrm{dt}, \text{ for } -\infty < x < \infty$$

Note: $P(a < X < b) = F(a) - F(b)$

Note: $f(x) = F'(x)$

3.3.3 Joint Probability Mass Functions

The function $f(x, y)$ is a Joint Probability Function or Probability Mass Function of the discrete random variables X and Y if

1. $$f(x, y) \geq 0 \text{ for all } (x, y)$$

2.
$$\sum_x \sum_y f(x,y) = 1$$

3.
$$P(X = x, Y = y) = f(x,y)$$

4. For any region A in the xy plane,
$$P((X,Y) \in A) = \sum_A \sum f(x,y)$$

3.4 Joint Probability Distributions

he function $f(x,y)$ is a joint density function of the continuous random variables X and Y if

1.
$$f(x,y) \geq 0, \text{ for all}(x,y)$$

2.
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\text{dx dy} = 1$$

3.
$$P((X,Y) \in A) = \int_A \int f(x,y)\text{dx dy}$$

3.4.1 Marginal Distribution

The marginal distibution of X alone Y alone are

(for the discrete case)

$$g(x) = \sum_y f(x,y) \text{ and } h(y) = \sum_x f(x,y)$$

(for the continuous case)

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\text{dy and } h(y) = \int_{-\infty}^{\infty} f(x,y)\text{dx}$$

Note: Marginal distributions are probability distributions

3.5 Joint PDFs

Let X and Y be two random variables, sicrete or continuous. The conditional distribution of the random variable Y given that $X = x$ is

$$f(y|x) = \frac{f(x,y)}{g(x)}, \text{ provided } g(x) > 0$$

Similarly, the conditional distribution of the random variable X given that $Y = y$ is

$$f(y|x) = \frac{f(x,y)}{h(y)}, \text{ provided } h(y) > 0$$

Where $g(x)$ and $h(y)$ are the respective marginal distributions.

3.5.1 Statistical Independence

Let X and Y be two random variables, discrete or continuous, with joint probability distribution $f(x,y)$ and marginal distributions $g(x)$ and $h(y)$ respectively. The random variables X and Y are said to be statistically independent if and only if

$$f(x,y) = g(x)h(y)$$

for all (x,y) withing their range.

Statistical Independence (Extended Case)

Let $X_1, X_2, \cdots, X_n$ be n random variables, discrete or continuous, with joint probability distribution $f(x_1, x_2, \cdots, x_n)$ and marginal distributions $f_1(x_1), f_2(x_2), \cdots, f_n(x_n)$. The random variables $X_1, X_2, \cdots, X_n$ are said to be statistically independent if and only if

$$f(x_1, x_2, \cdots, x_n) = f_1(x_1), f_2(x_2), \cdots, f_n(x_n)$$

for all (x,y) withing their range.

Chapter 4

Mean of a Random Variable

4.1 Mean of a Random Variable

Let X be a random variable with probability distrobution $f(x)$. The mean, or expected value, of X is:

Discrete Case:

$$\mu = E(X) = \sum_x x f(x)$$

Continuous Case:

$$\mu = E(X) = \int_{=\infty}^{\infty} x f(x) \mathrm{dx}$$

Appendix A

Homework

A.1 Homework Set 1

Page	Problem Numbers
13	1.5, 1.6
17	1.11, 1.12
31	1.18, 1.19, 1.20, 1.29, 1.30

A.2 Homework Set 2

Page	Problem Numbers
42	2.3, 2.6, 2.10, 2.11, 2.14, 2.16, 2.18
51	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
59	2.52, 2.53, 2.56, 2.57, 2.58, 2.62, 2.67, 2.70

A.3 Homework Set 3

Page	Problem Numbers
69	2.74, 2.75, 2.79, 2.82, 2.84, 2.87, 2.90, 2.92

A.4 Homework Set 4

Page	Problem Numbers
77	2.96, 2.97, 2.99, 2.100
91	3.1, 3.2, 3.5, 3.10, 3.11, 3.12, 3.25

A.5 Homework Set 5

Page	Problem Numbers
91	3.17, 3.27, 3.30, 3.31, 3.33, 3.35
104	3.37, 3.38 3.44, 3.45, 3.46, 3.52, 3.57

A.6 Homework Set 6

Page	Problem Numbers
See Blackboard	

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. 12).