Class Notes
for
Statistics
Letu Math–3403

Nicholas Capo
nicholas.capo@gmail.com

November 9, 2012
10:09am

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

1 Introduction 7
1.1 Data 7
1.1.1 Data Sets 7
1.1.2 Types of Data 7
1.2 Sample Mean and Median 7
1.2.1 Definition 7
1.3 Measures of Variability 8
1.3.1 Standard Deviation 8
1.4 Descriptive Statistics 8
1.4.1 Quartiles 8
1.4.2 Range and Interquartile Range 8
1.4.3 Box and Whisker Plot 9
1.5 Stem and Leaf Plots 9
1.5.1 Key Notation 10
1.5.2 Double Stem and Leaf 10
1.6 Frequency Distribution 10
1.6.1 Midpoint of a Class 10
1.6.2 Relative Frequency 10
1.7 Scatter Plots 10
1.7.1 Example Scatter Plot 10

2 Probability 13
2.1 Experiments 13
2.2 Sample Space 13
2.2.1 Example 13
2.2.2 Tree Diagrams 13
2.2.3 Events 13
2.3 Counting Sample Points 14
2.3.1 Multiplication Rule 14
2.3.2 Factorial 14
2.3.3 Permutation 14
2.3.4 Permutations at a Time 14
2.3.5 Permutations in a Circle 14
2.3.6 Permutations of a Kind 15

2.3.7 Partitioning . . . 15
2.3.8 Combinations . . . 15
2.3.9 Partitioning and Combinations . . . 15
2.4 Probability of an Event . . . 15
2.4.1 Complimentary Probability . . . 16
2.4.2 Different Outcomes . . . 16
2.4.3 Additive Rule of Probability . . . 16
2.5 Conditional Probability . . . 16
2.5.1 Multiplication Rule . . . 16
2.5.2 Independent Events . . . 16
2.5.3 Corollary . . . 17
2.5.4 Compliments of Independent Events . . . 17
2.5.5 Extended Multiplication for Conditional Probability . . . 17
2.6 Bayes' Rule . . . 17
2.6.1 Partitioning the Sample Space . . . 17
2.6.2 Theorem of Total Probability . . . 17
2.6.3 Bayes Rule . . . 17

3 Random Variables 19
3.1 Concept of a Random Variable . . . 19
3.1.1 Definition . . . 19
3.1.2 Discrete Sample Space . . . 19
3.1.3 Continuous Sample Space . . . 19
3.2 Discrete Probability Distribution . . . 19
3.2.1 Cumulative Distribution Function . . . 20
3.3 Continuous Random Variables . . . 20
3.3.1 Probability Density Function . . . 20
3.3.2 Cumulative Distribution Function . . . 20
3.3.3 Joint Probability Mass Functions . . . 20
3.4 Joint Probability Distributions . . . 21
3.4.1 Marginal Distribution . . . 21
3.5 Joint PDFs . . . 21
3.5.1 Statistical Independence . . . 22

4 Mean of a Random Variable 23
4.1 Mean of a Random Variable . . . 23
4.1.1 Expected Values of Functions of Random Variables . . . 23
4.1.2 Expected Values of Joint distributions . . . 23
4.2 Variance and Covariance of Random Variables . . . 24
4.2.1 Variance/Standard Deviation . . . 24
4.2.2 Variance of a Random Variable . . . 24
4.2.3 Variance of a Random Variable . . . 24
4.2.4 The Empirical Rule . . . 25
4.2.5 Chebyshev's Theorem . . . 25
4.2.6 Variance of a Function of a Random Variable . . . 25
4.2.7 Alternative Form of Variance . . . 25

4.2.8 Covariance of X and Y 25
4.2.9 Alternate Form 26
4.2.10 Description of Covariance 26
4.2.11 Correlation Coefficient 26
4.3 Linear Combinations of Random Variables 26
4.3.1 Expected Value of a Sum or Difference of Functions . . . 26
4.3.2 Properties of Expected Values 26
4.3.3 26
4.3.4 27
4.3.5 Piston Example 27

5 Binomial and Multinomial Distributions 29
5.1 Introduction 29
5.2 Bernoulli Experiment 29
5.2.1 Probability Distribution 29
5.2.2 Multinomial Distribution 30
5.3 Hypergeometric Distribution 30
5.3.1 Hypergeometric Formula 30

6 Continuous Probability Distributions 31
6.1 Continuous Uniform Distribution 31
6.1.1 Mean and Variance 31
6.2 Normal Distribution 31
6.2.1 Properties of Normal Distribution 31
6.2.2 The Standard Normal Distribution 32
6.2.3 Conversion from Normal to Standard Normal Distributions 32
6.2.4 Probability and Normal Distributions 32
6.6 Gamma and Exponential Distributions 32
6.6.1 Gamma Function 32
6.6.2 Gamma Distribution 33
6.6.3 Mean and Variance 33

8 Sampling Distribution 35
8.3 Sampling Distribution 35
8.3.1 Properties of Sampling Distributions of Sample Means . . 35
8.3.2 Central Limit Theorem 35
8.3.3 36
8.4 Chi Squared Distribution 36
8.5 Sampling Distribution of S^2 36
8.6 t-Distribution 37
8.6.1 Alternative Form 37
8.6.2 Properties of t-distribution 37

9 One and Two Sample Estimation Problems **39**
9.1 Notes 39
9.2 Point Estimate for Population 39
9.3 Interval Estimate 39
9.4 Confidence Intervals 40
9.5 Margin of Error 40
9.5.1 Sample Size 40
9.6 9.8 Confidence Intervals for the Difference Between Means 40
9.6.1 40
9.7 Confidence Intervals for Proportions 40
9.7.1 Population proportion 40
9.7.2 Confidence Interval for Proportion 41
9.7.3 Sample Size 41
9.8 Estimating the Difference between Two Points 41

A Homework **43**
A.1 Homework Set 1 43
A.2 Homework Set 2 43
A.3 Homework Set 3 43
A.4 Homework Set 4 43
A.5 Homework Set 5 43
A.6 Homework Set 6 44
A.7 Homework Set 7 44

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

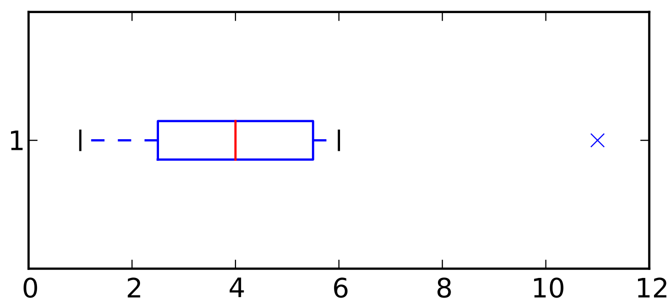

Figure 1.1: Example Box And Whisker Plot

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, 11, 10, 4, 7, 11, 6, 8, 2, 3, 12, 6]`

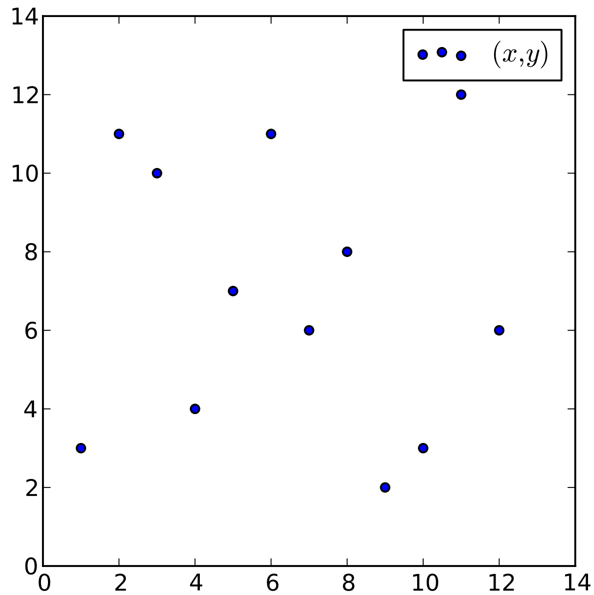

Figure 1.2: Example Scatter Plot

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.

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

3.

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

3.4.1 Marginal Distribution

The marginal distribution of X alone Y alone are

Discrete

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

Continuous

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

Note: Marginal distributions are probability distributions

3.5 Joint PDFs

Let X and Y be two random variables, discrete 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 distribution $f(x)$. The mean, or expected value, of X is:

Discrete:

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

Continuous:

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

4.1.1 Expected Values of Functions of Random Variables

Let X be a random variable with pdf $f(x)$. The expected value of the random variable $g(X)$ is

Discrete:

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

Continuous:

$$E(g(X)) = \int_{-\infty}^{\infty} g(x) f(x)$$

4.1.2 Expected Values of Joint distributions

Let X and Y be random variables with joint probability distribution $f(x, y)$. The mean, or expected value, of a random random variable $g(x, y)$ is

Discrete:

$$\mu_{g(x,y)} = E[g(x, y)] = \sum_X \sum_Y g(x, y) f(x, y)$$

Continuous:

$$\mu_{g(x,y)} = E[g(x,y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)$$

4.2 Variance and Covariance of Random Variables

4.2.1 Variance/Standard Deviation

Variance/Standard Deviation is a measure of the typical (average) amount the value deviates from the mean.

The further the values are from the mean, the grater the variance/standard variation.

4.2.2 Variance of a Random Variable

Let X be a random variable with probability distribution $f(x)$ and mean μ. The variance of X is

Discrete:

$$\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x)$$

$$\sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)$$

Standard Variation

$$\sigma = \sqrt{\sigma^2} = \sqrt{E((X-\mu)^2)}$$

- Variance — In terms of the mean's unit squared
- Standard Deviation — Terms of the mean's unit
- It does not makes sense to add or subtract the variance and the mean
- It does not make sense to add or subtract the standard deviation and the mean

4.2.3 Variance of a Random Variable

The variance of a Random Variable X is

$$\sigma^2 = E(X^2) - \mu^2$$

4.2.4 The Empirical Rule

For data with a (symmetric) bell-shaped distribution,

- $P[(\mu-\sigma) < X < (\mu+\sigma)] \approx 68\%$
- $P[(\mu-2\sigma) < X < (\mu+2\sigma)] \approx 95\%$
- $P[(\mu-3\sigma) < X < (\mu+3\sigma)] \approx 99.7\%$

4.2.5 Chebyshev's Theorem

The probability that any random variable X will assume a value within k standard deviations of the mean is at least $1-\frac{1}{k^2}$. That is,

$$P[(\mu-k\sigma) < X < (\mu+k\sigma)] \geq 1-\frac{1}{k^2}$$

4.2.6 Variance of a Function of a Random Variable

Let X be a random variable with probability distribution $f(x)$. The variance of the random variable $g(X)$ is

Discrete:

$$\sigma^2_{g(X)} = E\{[g(X)-\mu_{g(X)}]^2\} = \sum_x [g(x)-\mu_{g(X)}]^2 f(x)$$

Continuous:

$$\sigma^2_{g(X)} = E\{[g(X)-\mu_{g(X)}]^2\} = \int_{-\infty}^{\infty} [g(x)-\mu_{g(X)}]^2 f(x)\mathrm{dx}$$

4.2.7 Alternative Form of Variance

$$\sigma^2 = E[[g(x)]^2] - \mu^2_{g(x)}$$

4.2.8 Covariance of X and Y

Let X and Y be random variable with joint pdf $f(x,y)$. The covariance of X and Y is

Discrete:

$$\sigma_{XY} = E[(X-\mu x)(Y-\mu y)] = \sum_X \sum_Y (x-\mu x)(y-\mu y) f(x,y)$$

Continuous:

$$\sigma_{XY} = E[(X-\mu x)(Y-\mu y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu x)(y-\mu y) f(x,y)\mathrm{dx}\ \mathrm{dy}$$

4.2.9 Alternate Form

$$\sigma_{XY} = E(XY) - \mu_x \mu_y$$

4.2.10 Description of Covariance

[See Slide 4 − 26]

4.2.11 Correlation Coefficient

The correlation coefficient is a unit free measure of the strength of the linear relationship between two variables

Let X and Y be random variables with covariance σ_{XY}, and standard deviations σ_X and σ_Y respectively. The correlation coefficient of X and Y is

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

For $-1 \leq \rho \leq 1$

4.3 Linear Combinations of Random Variables

4.3.1 Expected Value of a Sum or Difference of Functions

The expected value of the sum or difference of two or more functions of a random variable X is the sum or difference of the expected values of the functions. That is,

$$E[g(X) \pm h(X)] = E[g(X)] \pm E[h(X)].$$

4.3.2 Properties of Expected Values

- Where k is a constant, and X and Y are Random Variables
- $E(k) = k$
- $E(kX) = kE(X)$
- Example: $E(3X + 5Y + 2) = 3E(X) + 5E(Y) + 2$

4.3.3

Let X and Y be independent random variables. Then

$$E(XY) = E(X)E(Y)$$

Note: The inverse does not hold

4.3.4

If X and Y are random variables with joint probability distribution $f(x, y)$ and a, b, c are real constants then

$$\sigma^2_{aX+bY+c} = a^2\sigma^2_X + b^2\sigma^2_Y + 2ab\sigma_{XY}$$

Independent Covariance

If X and Y are independent random variables

$$\sigma_{XY} = 0$$

4.3.5 Piston Example

[See Slide 4–43]

Chapter 5

Binomial and Multinomial Distributions

5.1 Introduction

[Not included in these Notes]

5.2 Bernoulli Experiment

1. The experiment is repeated for a fixed number of trials, where each trial is independent of other trials
2. There are only two possible outcomes of interest: success (S) or failure (F)
3. The probability of success p is the same for each trial
4. The random variable X counts the number of successful trials

5.2.1 Probability Distribution

The probability of exactly x successes in n trials is

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, 3, \cdots, n$$

Where:

$n =$ number of trials

$p =$ probability of success

$q =$ $(1 - p)$ probability of failure

$x =$ number of successes in n trials

$n - x =$ number of failures in n trials

Binomial PDF Function

Found in Calculator: Catalog → (F3) Flash Apps → `binomPdf(`

Binomial CDF Function

Found in Calculator: Catalog → (F3) Flash Apps → `binomCdf(`

Mean and Variance of Binomial Distribution

The mean and variance of the binomial distribution $b(x; n, p)$ are $\mu = np$ and $\sigma^2 = npq$.

5.2.2 Multinomial Distribution

Similar to binomial distribution, except the number of possible outcomes is more than two.

$$P(X_1 = x_1, X_2 = x_2, \cdots X_k = x_k) = \binom{n}{x_1, x_2, x_3 \cdots x_n} p_1^{x_1} p_2^{x_2} \cdots p_n^{x_n}$$

$n =$ number of trials

$p_i =$ probability of the i^{th} outcome

$x_i =$ number of occurrences of the i^{th} outcome in n trials

Note: $\binom{n}{x_1, x_2, x_3} = \frac{n!}{x_1! \times x_2! \times x_3!}$

5.3 Hypergeometric Distribution

1. A random sample of size n is selected without replacement from N items. (dependent events)
2. k items are classified as successes and $N - k$ items as failures
3. x represents the number of successes in a random sample of size n

5.3.1 Hypergeometric Formula

$$P(X = x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

Chapter 6

Continuous Probability Distributions

6.1 Continuous Uniform Distribution

Uniform: $f(x; A, B) = \frac{1}{B-A}$, $A \leq x \leq B$

6.1.1 Mean and Variance

the mean and average of the uniform distribution are

$$\mu = \frac{A+B}{2}$$

$$\sigma^2 = \frac{(B-A)^2}{12}$$

6.2 Normal Distribution

Many Random Variables have distributions that can be approximated by a bell shaped curve.

The Distribution Formula is

$$n(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \text{ to the power of } e^{\frac{-1}{2\sigma^2}(x-\mu)^2}$$

Where $-\infty < x < \infty$

6.2.1 Properties of Normal Distribution

1. The mean median and mode are equal
2. The normal curve approaches, but does not touch the x-axis. That is: Horizontal Asptote at $y = 0$

3. The line of symmetry gives the location of the mean.

4. The inflection points provides a way to estimate the standard deviation from the graph

6.2.2 The Standard Normal Distribution

The Standard Normal Distribution has a mean of 0 and a standard deviation of 1.

Calculator Help

Catalog → F3 (Flash Apps) → normCDF(x_{low}, x_{high}, [μ, σ])

6.2.3 Conversion from Normal to Standard Normal Distributions

Any x-value of a normal distribution can be transformed into a z-value in a standard normal distribution by

$$z = \frac{x - \mu}{\sigma}$$

6.2.4 Probability and Normal Distributions

If a random variable x is normally distributed, you can find the probability that x will occur in a given interval by calculating the area under the normal curve for that interval.

6.6 Gamma and Exponential Distributions

The gamma and exponential distributions are distributions that are often used to model time to failure, waiting, times and between arrival times.

6.6.1 Gamma Function

The gamma function is defined by

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \mathrm{dx}, \text{ for } \alpha > 0$$

If α is a positive integer then

$$\Gamma(n) = (n-1)!$$

6.6.2 Gamma Distribution

The continuous random variable X has a gamma distribution, with parameters α and β, if its density function is given by

$$f(x;\alpha,\beta) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & , \quad \text{for } x > 0 \\ 0 & , \quad \text{elsewhere} \end{cases}$$

When $\alpha = 1$ the gamma distribution becomes the exponential distribution

$$f(x,\beta) = \begin{cases} \frac{1}{\beta} x^{\alpha-1} e^{-x/\beta} & , \quad \text{for } x > 0 \\ 0 & , \quad \text{elsewhere} \end{cases}$$

Where $\alpha > 0$ and $\beta > 0$

6.6.3 Mean and Variance

Gamma

$$\mu = \alpha\beta$$

$$\sigma^2 = \alpha\beta^2$$

Exponential Distribution

$$\mu = \beta$$

$$\sigma^2 = \beta^2$$

$$\sigma = \beta$$

Chapter 8

Sampling Distribution

8.3 Sampling Distribution

The probability distribution of a statistic is called a sampling distribution. We are interested in the sampling distribution of the sample mean $\overline{X}$

The probability distribution of the sample means is called the sampling distribution of the sample mean.

8.3.1 Properties of Sampling Distributions of Sample Means

1. The mean of the sample means, $\mu_{\overline{x}}$ is equal to the population mean μ.

2. The standard deviation of the sample means $\sigma_{\overline{x}}$ is equal to the population standard deviation σ divided by the square root of the sample size n.

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

   This is called the standard error of the mean.

8.3.2 Central Limit Theorem

If $\overline{X}$ is the mean of a random sample of size n taken from a population with mean μ and finite variance σ^2, then the limiting form of the distribution

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \text{ as } n \to \infty$$

In other words for sufficiently large n, the mean of a sample of size n approximates the normal distribution with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$

Minimum Sample Size

What is the minimum sample size needed for a "good" approximation?

1. If the population is normally distributed, the sample size can be any number.
2. For non-normally distributed populations, a good rule of thumb is $n \geq 30$

8.3.3

If independent samples of size n_1 and n_2 are drawn from two populations with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2 respectively, then the sampling distribution of the differences of means $\overline{X}_1 - \overline{X}_2$, is approximately normally distributed with mean and variance given by,

$$\mu_{\overline{X}_1 - \overline{X}_2} = \mu_1 - \mu_2$$

and

$$\sigma^2_{\overline{X}_1 - \overline{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

Also note when adding as opposed to subtracting:

$$\mu_{\overline{X}_1 + \overline{X}_2} = \mu_1 + \mu_2$$

and

$$\sigma^2_{\overline{X}_1 + \overline{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

8.4 Chi Squared Distribution

A special case of the Gamma distribution when $\alpha = \frac{v}{2}$ and $\beta = 2$

$$f(x;v) = \begin{cases} \frac{1}{2^{v/2}\Gamma(v/2)} x^{(v/2)-1} e^{-x/2} & , \quad x > 0 \\ 0 & , \quad \text{else} \end{cases}$$

8.5 Sampling Distribution of S^2

If S^2 is the variance of the random sample of size n from a *normal* population having variance σ^2, the the statistic

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

Has a χ^2 distribution with $v = n - 1$ degrees of freedom.

Calculator Help

Catalog → F3 (Flash Apps) → Chi2CDF(x_{low}, x_{high}, [degrees of freedom])
Catalog → F3 (Flash Apps) → invChi2([1-area], [degrees of freedom])

8.6 t-Distribution

Let Z be a standard normal random variable and V be a chi-squared random variable with v degrees of freedom. If Z and V are independent, then the distribution of the random variable T, where

$$T = \frac{Z}{\sqrt{V/v}}$$

is called a t-distribution with v degrees of freedom.

8.6.1 Alternative Form

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

with $v = n - 1$ degrees of freedom

8.6.2 Properties of t-distribution

1. The total area under a t-curve is 1 or 100%
2. The mean, median, and mode are equal to zero
3. As the degrees of freedom increase, the t-distribution approaches the standard normal distribution. After 30 degrees of freedom the t-distribution is very close to the standard normal z-distribution.

Calculator Help

Catalog → F3 (Flash Apps) → tCDF(x_{low}, x_{high}, [degrees of freedom])
Catalog → F3 (Flash Apps) → inv_t(area, degrees of freedom)

Chapter 9

One and Two Sample Estimation Problems

A statistic $\hat{\Theta}$ is said to be an unbiased estimator of the parameter θ if

$$\mu_{\hat{\Theta}} = E(\hat{\Theta}) = \theta$$

If we consider all possible estimators of some parameter θ, the one with the smallest variance is called the most efficient estimator of θ.

9.1 Notes

1. The sample mean and median are both unbiased estimators of the population mean. The sample variance is an unbiased estimator of the population variance.

2. The sample standard deviation is a *biased* estimator of the population standard deviation.

3. For normal populations, the sample mean is the most efficient estimator of the population mean

9.2 Point Estimate for Population

A single value estimate for a population parameter Most unbiased point estimate of the population mean μ is the sample mean $\overline{x}$

9.3 Interval Estimate

An interval or range of values used to estimate a population parameter.

9.4 Confidence Intervals

A $(1-\alpha)$ confidence interval for the population mean μ

$$x - E < \mu < x + E \text{ where } E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

The probability that the confidence interval contains μ is $1-\alpha$

9.5 Margin of Error

The greatest possible distance between the point estimate and the value of the population parameter for a given level of confidence. Denoted by

$$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

This is sometimes called the maximum error of estimate or error tolerance.

9.5.1 Sample Size

Given a $(1-\alpha)$ confidence level and a margin of error E, the minimum sample size n needed to estimate the population mean μ is

$$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2$$

If σ is unknown you can estimate it using s provided you have a preliminary sample with at least 30 members.

9.6 9.8 Confidence Intervals for the Difference Between Means

9.6.1

If $\overline{x_1}$ and $\overline{x_2}$ are means of independent random samples of size n_1 and n_2 from populations with variances σ_1^2 and σ_2^2 respectively. Then a $(1-\alpha)$ confidence interval for $\mu_1 - \mu_2$ is given by

$$(\overline{x_1} - \overline{x_2}) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_1}}$$

9.7 Confidence Intervals for Proportions

9.7.1 Population proportion

The probability of success in a single trial of a binomial experiment

The proportion of successes in a sample is denoted by

$$\hat{p} = \frac{x}{n} = \frac{\text{successes}}{\text{sample size}}$$

9.7.2 Confidence Interval for Proportion

$$[\hat{p} - E, \hat{p} - E]$$

Where

$$E = z_c\sqrt{\frac{\hat{p}\hat{q}}{n}} \text{ and } \hat{q} = 1 - \hat{p}$$

If

$$n\hat{p} \geq 5 \text{ and } n\hat{q} \geq 5$$

9.7.3 Sample Size

Given a $(1-\alpha)$ confidence level and a margin or error E, the minimum sample size n needed to estimate p is

$$n = \hat{p}\hat{q}\left(\frac{z_{\alpha/2}}{E}\right)^2$$

This assumes you have an estimate for $\hat{p}$ and $\hat{q}$, if not set $\hat{p} = \hat{q} = 0.5$ this gives an upper bound for the error.

9.8 Estimating the Difference between Two Points

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
106	3.53, 3.55, 3.58, 3.59, 3.60, 4.1, 4.2, 4.5, 4.7
117	4.10, 4.11, 4.13, 4.18, 4.19, 4.24, 4.28, 4.30, 4.34, 4.35, 4.49
127	4.39, 4.41, 4.44,4.45,4.46, 4.75, 4.77

A.7 Homework Set 7

Page	Problem Numbers
127	4.51
138	4.58, 4.60, 4.61, 4.62, 4.63, 4.70, 4.89
150	5.4, 5.6, 5.10, 5.17, 5.18, 5.19, 5.21,
157	5.32, 5.39, 5.43, 5.44
	5.50, 5.52, 5.54, 5.57, 5.61, 5.66, 5.72

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).