Section 1 – Introduction to R

1.1 – Getting Started With R and RStudio 1
General introduction 1
Getting started 1
RStudio layout 1
R scripts and R markdown 3
1.2 – Coding in RStudio 3
Try out some code! 3
Variables 3
Vectors 4
1.3 – Analyzing Data Sets in RStudio 5
Loading data 5
Loading data into R markdown 5
Working with data 6
Getting help 7
1.4 – Additional Practice 7

Section 2 – Statistical Investigations

2.1 – Introduction 9
Designing an investigation 9
Probability and statistics 10
2.2 – Sampling and Replacement 10
Types of samples 10
Replacement 12
2.3 – Variables and Measurement 13
Types of variables 13
Measuring variables 15
2.4 – Summarizing Variables 16
Summarizing categorical variables 16
Summarizing quantitative variables: center 17
Summarizing quantitative variables: variation 18
2.5 – Graphing and Plotting Variables 20
Graphics for categorical variables 20
Graphics for quantitative variables 21
2.6 – Additional Practice 23

Section 3 - Probability and Random Variables

3.1 – Core Probability Concepts 25
Sample spaces 25
Properties of probabilities 26
Conditional probability 26
3.2 – Discrete Random Variables 28
Definition of random variable 28
Probability mass functions 28
Describing and summarizing discrete random variables 29
The binomial distribution 31
3.3 – Continuous Random Variables 34
Probability density functions 34
The normal distribution 35
3.4 – Additional Practice 37

Section 4 - Introduction to Statistical Inference

4.1 – Testing With Simulation 39
Modeling hypotheses 39
Helper or hinderer? 40
Simulating the coin experiment for the helper or hinderer study 40
Conducting the TinkerPlots simulation in R 42
4.2 – Formalizing the Hypothesis Test 43
Defining hypotheses 43
Measuring evidence 44
Using R to conduct a hypothesis test 45
4.3 – Additional Practice 47

Section 5 - Exploring Hypothesis Testing

5.1 – Two-sided Alternatives 49
5.2 – Errors and Power 50
Errors in hypothesis testing 50
Probabilities of errors 50
Practical significance 52
5.3 – Additional Practice 54

Section 6 - Estimating a Proportion

6.1 – Deriving a Confidence Interval 55
Constructing an interval 55
Finding an interval in R 58
6.2 – Additional Estimation Topics 59
Connecting confidence intervals and hypothesis testing 59
Bootstrap confidence intervals 59
6.3 – Additional Practice 62

Section 7 - Statistical Inference for One Mean

7.1 – The Central Limit Theorem 63
Normality through sampling 63
The central limit theorem 65
7.2 - Hypothesis Testing for One Mean 66
Theoretical background 66
R code for testing a mean 68
QQ plots 70
7.3 – Confidence Interval for a Mean 71
Theoretical background 71
R code for estimating a mean 72
Bootstrap confidence intervals 73
7.4 – Additional Practice 75

Section 8 - Statistical Inference for Two Proportions

8.1 – Validity of a Statistical Study 77
Introduction 77
Internal and external validity 77
8.2 – Statistical Inference for a Difference in Proportions 81
Theoretical background 81
Testing a difference in proportions by hand 82
R code for testing a difference in proportions 83
8.3 – Conducting Tests with Simulation 85
Experimental design 85
Observational studies 86
8.4 – Additional Practice 89

Section 9 - Statistical Inference for Two Means

9.1 – Comparing Dependent Outcomes 91
Introduction 91
Repeated observations 91
Related observations 93
Paired samples in R 94
9.2 – Comparing Independent Samples 94
Theoretical background 94
Testing a difference in means by hand 96
R code for testing a difference in means 97
Using simulation to conduct a difference in means test 98
9.3 – Additional Practice 100

Section 10 - Simple Linear Regression

10.1 – The Least Squares Criterion 101
Introduction 101
Finding the "best" line 101
Sum of squared errors 102
A model-based approach 102
10.2 – Simple Linear Regression 103
Using R for linear regression 103
Checking assumptions 104
10.3 – Assessing and Using Linear Models 105
Hypothesis testing and confidence intervals for regression parameters 105
Assessing the model 106
Prediction intervals 107
Dealing with heteroskedasticity 107
10.4 – Additional Practice 111

Section 11 - Multiple Linear Regression

11.1 – Simulation in Linear Regression 113
11.2 – Multiple Linear Regression 114
A linear model for multiple predictors 114
Using R for multiple linear regression 114
Polynomial Terms 117
11.3 – Additional Practice 118

Section 12 - Analysis of Variance

12.1 – Comparing Three or More Groups 119
Introduction 119
The one-way ANOVA model 120
Calculating an ANOVA test 120
12.2 – One-way ANOVA 122
The ANOVA table 122
Post-hoc analysis 123
Conducting one-way ANOVA in R 124
12.3 – Two-way ANOVA 125
Theoretical model 125
Building a two-way ANOVA model 125
12.4 – Additional Practice 127

Acknowledgements

This course is a summation of a great amount of work, not all of which is my own. For making this course possible, I would like to thank:

- Carly Roberts, for your work on making a series of tutorial videos for TinkerPlots.
- Luke Thorell, for creating various practice and homework problems for each section.
- All of my undergraduate course assistants, who put in immense work to prepare for each class, and are vital to creating a vibrant classroom community and establishing one-on-one connections with students.
- Dr. Jennifer Noll, for introducing me to the TinkerPlots and the original CATALST curriculum, and shaping my teaching philosophies to be student-centered and focused on active learning.
- Dr. Kelly Findley and Dr. Vimal Rao, for your constant encouragement and furthering the inspiration for this curriculum.