

A Guide to Statistical Methods for Researchers

Melinda Higgins

2024-10-14

Table of contents

Preface	3
1 Introduction	4
2 Descriptive Statistics	5
3 Foundation of Correlation/Association	6
4 Comparison of 2 Independent Groups	8
5 Comparison of 3 or more Independent Groups	9
6 One Outcome - Multiple Predictors and/or Covariates	10
7 One Outcome - Multiple Predictors and/or Covariates	11
8 One Outcome - Multiple Predictors and/or Covariates	12
9 Summary	13
References	14

Preface

This is a Quarto book.

To learn more about Quarto books visit <https://quarto.org/docs/books>.

1 + 1

[1] 2

1 Introduction

- Overview and intent of book
- explain the parts

This is a book created from markdown and executable code.

See Knuth (1984) for additional discussion of literate programming.

```
1 + 1
```

```
[1] 2
```

2 Descriptive Statistics

- understanding how to describe data
- exploring data issues
- checking distributions - understanding the nuances of parametrics and non-parametric
- why does the CLT and normality matter - or does it?
- exploring methods - means vs medians, sd vs IQR, CIs and bootstrapped percentiles (2.5th, 97.5th)
- parametric, non-parametric rank based methods, semi-parametric bootstrapping

This is a book created from markdown and executable code.

See Knuth (1984) for additional discussion of literate programming.

1 + 1

[1] 2

3 Foundation of Correlation/Association

- beyond 1 variable at the time
- looking at association - continuous, categorical, ordinal
- understanding the correlation matrix
 - directionality - dependent/independent - supervised (target) vs unsupervised - implied causality
 - cont - cont - correlation, regression, scatterplots
 - cont - 2grp - t-tests, mw tests, linear reg, pt-biserial corr
 - cont - 3+grp - ANOVAs and the General Linear Model
 - 2 grp - cont - logistic regression (flip side of a t-test) - generalized linear models
 - 2+ grp w/ 2+grp - logistic regression (or nominal, ordinal reg) - and chi-square tests
 - also show 2grp diff proportions - chisq goodness of fit - explain understanding
 - other types of responses - like count data Poisson or neg binomial regression
 - dependent data - repeated measures and other dependent types (like right/left, etc)
 - no specific outcome - unsupervised ML, PCA, FA, cluster analyses, etc
- thinking about the point of a given model
 - is it for hypothesis testing,
 - to estimate the size of an effect,
 - make predictions
 - create a functional form to understand mechanisms
- breaking down the analysis approaches:
 - NHST, frequentist stats
 - bootstrapping
 - simulation?
 - permutations
 - bayesian

This is a book created from markdown and executable code.

See Knuth (1984) for additional discussion of literate programming.

1 + 1

[1] 2

4 Comparison of 2 Independent Groups

The focus of this chapter is mainly comparing 2 independent groups

- t-tests
- Mann Whitney test
- bootstrapping
- permutations
- simulations - maybe
- bayesian
- using a regression approach
- using a GLM approach (e.g. F-test = t-test squared)
- add a discussion of effect sizes for each of these approaches

This is a book created from markdown and executable code.

See Knuth (1984) for additional discussion of literate programming.

1 + 1

[1] 2

5 Comparison of 3 or more Independent Groups

The focus of this chapter is methods for comparing 3 or more independent groups

- ANOVA - including post hoc tests
- Kruskal-Wallis test - including post hoc tests with adjusted p-values
- bootstrapping
- permutations
- simulations - maybe
- bayesian
- using a regression approach (dummy variables - thinking about multiple comparisons)
- using a GLM approach (e.g. F-test = t-test squared)
- add a discussion of effect sizes for each of these approaches

This is a book created from markdown and executable code.

See Knuth (1984) for additional discussion of literate programming.

1 + 1

[1] 2

6 One Outcome - Multiple Predictors and/or Covariates

This chapter (2 chapters) will focus the regression modeling approach with more than 1 predictor. Break into 2 chapters

- continuous outcome - usual linear regression
 - bootstrapping
 - permutations
 - simulations - maybe
 - bayesian
- categorical or other outcomes - generalized linear models
 - bootstrapping
 - permutations
 - simulations - maybe
 - bayesian
- add a discussion of effect sizes for each of these approaches

This is a book created from markdown and executable code.

See Knuth (1984) for additional discussion of literate programming.

```
1 + 1
```

```
[1] 2
```

7 One Outcome - Multiple Predictors and/or Covariates

This chapter will get into why dependence matters - usually time within subject but could be any nested structure data...

- continuous outcome - usual linear regression
 - bootstrapping
 - permutations
 - simulations - maybe
 - bayesian
- can also do this for categorical, binary, poisson, and other generalized outcomes...

This is a book created from markdown and executable code.

See Knuth (1984) for additional discussion of literate programming.

1 + 1

[1] 2

8 One Outcome - Multiple Predictors and/or Covariates

Up to this point we've focused on models with a specified target - what if we don't have one.

- PCA and FA
- Cluster Analyses
- other methods...

maybe add a chapter on MANOVA - links to logistic regression and discriminant analysis

This is a book created from markdown and executable code.

See Knuth (1984) for additional discussion of literate programming.

```
1 + 1
```

```
[1] 2
```

9 Summary

In summary, this book has no content whatsoever.

1 + 1

[1] 2

References

Knuth, Donald E. 1984. “Literate Programming.” *Comput. J.* 27 (2): 97–111. <https://doi.org/10.1093/comjnl/27.2.97>.