Learning Objectives

Chapter 1: Introduction to statistics

Theory:

What is statistics, what can it describe, and how can it be applied to the real world?

What is the difference between independent and dependent variables?

What is the difference between continuous and discrete variables?

How can data sets be organized/classified (one-way vs two-way classification)?

Computational:

How do we load data into R Statistical Software Program?

How do we set a working directory in R?

What is the difference between a vector, a data frame and a matrix in R?

How do we save and knit files in R?

Chapter 2: Descriptive Statistics

Theory:

What is the difference between population parameters and sampling statistics?

How do you measure central tendency of a data set?

How do you measure dispersion of a data set?

What is the five-number summary?

Computational:

How do we calculate the mean, median and mode in R?

How do we calculate the variance, the standard deviation and the range in R?

How do we calculate the five-number summary in R?

Chapter 3: Population Distributions

Theory:

What are the three population distributions observed in agriculture?

What are the characteristics of a binomial population?

When is it appropriate to use a poisson distribution?

What are the two parameters that characterize a normal distribution?

What are the unique parameter settings for a Standard Normal Distribution?

Computational:

Plot a histogram of a population data set

Create a box and whisker plot for a population dataset.

Label the upper and lower fences, upper and lower hinges and the median of the box plot.

Calculate the coefficient of variation for a population.

Chapter 4: Sampling Distributions

Theory:

What are the four sampling distributions observed in agriculture?

What is the central limit theorem?

When should a student T distribution be used over a normal distribution for samples?

What does a chi-square distribution describe?

What does a F distribution describe?

Computational:

Plot a normal distribution with sample size 20, sample size 200 and sample size 2000.

Plot a chi-square distribution and vary the degrees of freedom.

How do you calculate the difference between sample means? Between sample variances?

Chapter 5: Confidence Intervals and Hypothesis Testing for a Population

Theory:

What is the difference between the null hypothesis and alternative hypothesis?

What is the level of significance? What is the power of a test?

What values are needed to calculate a confidence interval?

What is the difference between a one-tailed and two-tailed test?

What are type I and type II errors?

Computational:

Compare the length of (two-tailed) confidence intervals with significance levels of 0.01, 0.05 and 0.10 in R.

Compare a one-tailed confidence interval and a two-tailed confidence interval with alpha levels of 0.05 in R.

Identify which calculated p-values allow you to reject/fail to reject the null hypothesis.

Chapter 6: Confidence Intervals and Hypothesis Testing for Two Populations

Theory:

What are the three main cases for population comparisons?

When are samples considered independent?

When can you pool the sample variances?

How do you determine if your samples are paired?

Computational:

How do we test if the sample variances are equal in R?

How do we test if the sample means are equal in R?

How do we perform a hypothesis test for paired samples?

Chapter 7: Analysis of Variance and Hypothesis Testing for Multiple Treatments

Theory:

What are the assumptions of ANOVA?

When is an ANOVA more appropriate than a T-test for hypothesis testing?

What are the six steps of running an ANOVA?

How does the calculation of experimental error change with and without subsamples?

Computational:

Run an ANOVA, report the calculated F-value, report the p-value and state your hypothesis decision.

Run a T-test and report the calculated t-value. Compare this value to the F-value; what is the relationship?

Compare the calculated experimental error with and without subsamples in R.

Chapter 8: Completely Randomized Design vs Randomized Complete Block Design

Theory:

What is a completely randomized design?

What is a randomized complete block design?

What are the reasons to incorporate blocking in an experiment?

What is an interaction?

Computational:

Run an ANOVA on a dataset as a CRD in R.

Run an ANOVA on the same dataset with blocks as a RCBD in R.

Compare the p-values for treatment for the CRD and RCBD.

Plot treatment means across blocks to identify interactions.

Chapter 9: Multiple Comparisons and Least Significant Difference (LSD)

Theory:

What is the purpose of LSD in pairwise comparisons?

What ANOVA results predicate performing pairwise comparisons?

What are the values needed to calculate the LSD?

What is an orthogonal comparison?

Computational:

Perform an ANOVA on your dataset in R.

Calculate the LSD of your dataset in R.

Identify groups that are significantly different from each other in your dataset.

Perform an F-test for an orthogonal comparison of your dataset.

Chapter 10: Simple Regression and Correlation

Theory:

What is the standard equation for a simple linear regression?

What do the coefficients of the simple linear regression represent?

What is the difference between Y, Y(hat) and Y(bar)?

What is the difference between the error due to the regression and the residual error?

How do you test the fitness of your regression line? What MS values do you use for your F-test?

Computational:

Calculate the coefficients (a and B) for your simple linear regression from the dataset.

Calculate the goodness of fit of your regression line by performing an ANOVA.

Calculate the coefficient of determination for your regression line.

Calculate the confidence intervals for B and a coefficients of the SLR.

Compare the confidence intervals around a predicted mean and a predicted individual observation.