Biology 2060: Statistics For Biologists

LECTURES: MON. & WED. 11:30am - 12:30 LOCATION: ACW 109

TUTORIALS: FRI. 2:30 - 5:30 LOCATION: VHA

COURSE DIRECTOR: Dr. Joel S. Shore, Professor of Biology

GRADING SCHEME:

Midterm 1 (20%) Friday Oct 4, 2013
Midterm 2 (30%) Friday Nov 8, 2013
Final exam (50%) tba in the Exam period

REQUIRED TEXT BOOK


(buy it and begin reading!). We will largely follow the text from the beginning up to and including at least chapter 18. But see below for explicit details of which chapters and sections.

TUTORIALS

The tutorials will be devoted to a number of tasks including.

1) Taking up problem sets assigned throughout the course.

2) Learning to use the statistical computing software SAS.

3) Midterm tests will be written during the tutorial period.

4) Dealing with individual problems/questions of students

LEARNING OUTCOMES
Statistics is an extremely important field for biology, and many other disciplines.

By the end of this course students should be able to:

1) Provide a summary of categorical and numeric data using graphical methods and statistics
2) Apply the most powerful hypothesis test(s) to data from a range of biological experiments involving categorical or numeric data.
3) Test the assumptions of various hypothesis tests
4) Interpret the results of the hypothesis tests carried out
5) Carry out the hypothesis tests both by “hand” and using the statistical computer program called SAS

COURSE OUTLINE

1) INTRODUCTION TO STATISTICS (read chapter 1)

What is statistics?
Populations and samples
Random sampling
Types of data - categorical versus numeric
Explanatory versus Response variables
Frequency and probability distributions

2) VISUALIZING DATA (read chapter 2)

Plotting frequency distributions
Bar graphs, histograms
Cumulative frequency distributions
Contingency tables
Scatterplots for two variables

3) DESCRIPTIVE STATISTICS (read chapter 3)

Sample mean and sample median
Variance and standard deviation
(computational formula)
Quartiles and box plots
Proportions

4) Estimation (read chapter 4)
Estimating a population parameter
Sampling distribution of the estimates
e.g. mean
Standard error
Confidence intervals

5) Probability (read chapter 5 sections 5.1 through 5.7 inclusive)
Probability of events
Mutually exclusive events
Probability distributions
Addition and multiplication rules
Independent events
Probability trees

6) Hypothesis testing (Chapter 6 & 7)
Null (Ho) versus Alternative (Ha) hypotheses
One versus two-sided tests
Examples
P-value
Type I and Type II errors
Analysing proportions and hypothesis tests using the Binomial distribution

OMIT THE MATERIAL ON CALCULATING CONFIDENCE INTERVALS FOR PROPORTIONS ON PAGES (162-163).

7) Goodness-of-Fit and Contingency tests (Chapters 8 & 9)

X² goodness-of-fit tests

assumptions

examples

Fitting data to probability distributions

Contingency tests

G-statistic

OMIT SECTION 9.2 ON ESTIMATING ASSOCIATION IN 2X2 TALBES: ODDS RATIOS

8) The normal distribution (Chapter 10)

Ignore normal approximation for the binomial distribution (section 10.7)

The standard normal distribution and probabilities

Normal distribution of sample means

Central limit theorem

9) Student's t-test (Chapter 11 & 12)

the t-distribution

confidence intervals

one- and two-sample t-tests

assumptions

paired t-test

F-test of equal variances
OMIT SECTION 11.5 ON CONFIDENCE LIMITS FOR VARIANCE AND STANDARD DEVIATION

10) Violations of assumptions, transformation and non-parametric tests (Chapter 13)

Detecting deviations from normality read section, but won't carry out Shapiro-Wilks test for normality.

Transforming data to meet normality assumption

Non-parametric alternatives to t-tests

Ignore chapter 14.

11) Analysis of Variance (Chapters 15)

Single factor ANOVA

Planned versus unplanned comparisons

Fixed versus random effects

(Unknown Nonparametric alternatives pg. 404, but you should know of their existence and when to use them).

12) Correlation and Regression (Chapter 16 and 17)

Chpt 16 only sections 16.1 (but ignore confidence interval material)

include 16.2, 16.3

Chpt 17 only sections 17.1 (ignore confidence interval)

include 17.3 (ignore ANOVA approach).

Correlation coefficient

Linear regression

Estimation of least squares line

Hypothesis test of slope

Assumptions and transformations
13) Two factor ANOVA and Randomized Block ANOVA

Chpt 18 Sections 18.2 – 18.4