FISH 560: Applied Multivariate Statistics for Ecologists
Autumn 2020

Instructor: Julian Olden
Office Location: Julian’s Living Room
Contact information: olden@uw.edu, 616-3112
Class location: Zoom (https://washington.zoom.us/j/97156093256, pw: fish560)
Prerequisite(s): QSCI 482 or equivalent or permission from instructor

E-mail: olden@uw.edu
Office hours: Wed 3:30-4:30 pm / appointment
Class hours: Wed 13:30-15:20, Fri 12:30-14:20

COURSE DESCRIPTION

With recent advances in data collection technology and ambitious field research, scientists are increasingly relying upon multivariate statistics to explore and test for patterns in their data. The goal of this course is to introduce graduate students in the ecological sciences to the multivariate statistical techniques necessary to carry out sophisticated analyses and to critically evaluate scientific papers using these approaches. This is a practical and hands-on course emphasizing the analysis and interpretation of multivariate analysis, covering a variety of approaches used by ecologists. The focus of the course is on the conceptual understanding and practical use of the methods (not the matrix algebra), with the singular hope of de-mystifying the “alphabet soup” of multivariate analysis.

The intent of this course is to provide you with the following: (1) an introduction to the use of multivariate statistics in ecology; (2) a conceptual organization of the various multivariate techniques, with respect to the types of research questions and data sets appropriate for each technique; and (3) a working understanding of how to use and interpret the results of each technique, including a conceptual overview, list of assumptions, diagnostics for assessing the assumptions, mechanics of performing the analysis using the R package, and how to interpret the statistical output of the analysis.

METHOD OF INSTRUCTION

Lectures/labs – Lectures will integrate both theoretical and applied aspects of multivariate statistics and provide solutions and interpretations of analyses. For each topic there will be a formal lecture followed by a computer-based lab where R will be used to demonstrate the particular multivariate technique.

Assignments – There will be a number of short assignments that are designed to enhance your conceptual understanding of multivariate approaches and increase your technical prowess in R.

Final report and peer review – The final paper will consist of a multivariate analysis, but the nature of the question, the source of the data, and the kinds of analysis employed are flexible. The requirement is that the data and analysis must address one or more specific biological hypotheses, which are to be tested using an appropriate method(s) of multivariate analysis. The primary goal is a coherent scientific paper, not excessive number crunching. You will also peer review other class members’ papers.

DATASETS

Personal dataset – A primary goal of this course is to provide you the opportunity to get better acquainted with your own data. The data set may be your own, one obtained from the literature or one
provided by the Instructor. Ideally you should use data that you have collected or are otherwise somewhat familiar with. The data set should be one or more matrices of entities × attributes (e.g., samples × species, species × characteristics of species, sites × environmental factors, etc.). The only data requirements are that it be adequate to test the hypotheses addressed in your final report. If you do not have access to a multivariate dataset, then I would be please to provide one.

Class dataset – Even if you do have a multivariate dataset, it is unlikely to be suitable for all the techniques covered in class. To address this issue, a common dataset containing stream fish community and environmental data is provided to all students at the beginning of the quarter. By using the class data you will be able to conduct all the statistical approaches covered in the class. Moreover, this dataset will serve as the basis for the R tutorials. You will be expected to work with both your own dataset and the class dataset. Data is sourced from the Environmental Protection Agency.

Assignment dataset – Analyses for all assignments will performed using a variety of datasets, including on containing information for single-malt whisky distilleries from different regions of Scotland that have been assigned an overall quality score and are attributed for different taste categories. Data is sourced from Whisky Advocate (http://whiskyadvocate.com/).

TEXTBOOK(S)

There is no required text for this course, however, here are some recommendations:

Conducting multivariate analyses in R (although the tutorials provided in class is all that you need!):

GRADING PLAN

The Standard UW Numerical Grading System will be used according to the breakdown provided below.

<table>
<thead>
<tr>
<th>Task</th>
<th>Due date</th>
<th>% of grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assignments (7)</td>
<td>Multiple</td>
<td>35%</td>
</tr>
<tr>
<td>One-page proposal</td>
<td>November 3rd</td>
<td>5%</td>
</tr>
<tr>
<td>Final paper</td>
<td>December 11th</td>
<td>40%</td>
</tr>
<tr>
<td>Peer-review reports</td>
<td>December 18th</td>
<td>20%</td>
</tr>
</tbody>
</table>
TOPICS (see Canvas website for class schedule)

Multivariate resemblance
- Modes of analysis, analytical spaces
- Similarity and distance coefficients (binary, categorical, quantitative)

Cluster analysis
- Hierarchical agglomerative clustering (e.g., linkage, UPGMA)
- Non-hierarchical divisive clustering (e.g., K-means, PAM)
- Model-based clustering

Direct Ordination
- Principal component analysis (PCA)
- Principal coordinate analysis (PCoA)
- Non-metric multidimensional scaling (NMDS)
- Correspondence analysis (CA)
- Detrended correspondence analysis (DCA)

Clustering and Ordination with Missing Data

Indirect Ordination
- Redundancy analysis (RDA)
- Canonical correspondence analysis (CCA)
- Canonical correlation analysis (CCorA)
- Canonical analysis on principal coordinates (CAP)
- Partial RDA and CCA
- Hierarchical RDA and CCA
- Principal Response Curves (PRC)

Fourth-corner and RLQ methods – Guest lecture (Dr. Pedro Peres-Neto)
  - Fourth-corner solution
  - RLQ methods
  - Model-based methods

Classification of Groups
- Discriminant Function Analysis
- Classification and Regression Trees

Testing for Similarities and Differences among Groups
- Analysis of similarity (ANOSIM)
- Multi-response Permutation Procedure (MRPP)
- Permutational MANOVA (perMANOVA)
- Permutation test of multivariate dispersion

Testing for Associations among Matrices
- Mantel tests
- Procrustes Analysis

Testing for Thresholds
- Gradient Forest Analysis
- Threshold Indicator Taxa ANalysis (TITAN)

ZOOM CODE OF CONDUCT

The UW Student Code of Conduct applies to online behavior as well as in-person or classroom behavior. You are expected to be professional and respectful when attending class on Zoom. The following are
class policies for our meetings with Zoom. Please read carefully. All students are expected to adhere to the policies.

General. Sign in with your full first name and last name as listed on the class roster. Do not use a nickname or other pseudonym when you log in. If you have changed your name to better reflect your gender identity, please send me an email so this can be noted on the roster and then you can use your current name on Zoom. If you do not have access to a computer or smartphone with internet access, call into class using a landline phone. This is not optimal; please try to locate an internet-enabled device to use for class. Please stay engaged in class activities. Close any apps on your device that are not relevant and turn off notifications.

Video. Turn on your video when possible. It is helpful to be able to see each other, just as in an in-person class. Exceptions include if you have limited internet bandwidth or no webcam or if you’re unable to find an environment without a lot of visual distractions. In the latter, please consider using a virtual background (instructions here).

Audio. Mute your microphone when you are not talking. This helps eliminate background noise. Use a headset when possible. If you own headphones with a microphone, please use them. This improves audio quality. Be in a quiet place when possible. Find a quiet, distraction-free spot to log in. Turn off any music, videos, etc. in the background.

Chat. Stay on topic. Use the chat window for questions and comments that are relevant to class. The chat window is not a place for socializing or posting comments that distract from the course activities. If you fill it up with random comments, I will be unable to sort through the information quickly to address students’ real questions/concerns about the course. No disrespect or hate speech. Just like in our in-person class, respectful behavior is expected. Consider Zoom a professional environment, and act like you’re at a job interview, even when you’re typing in the chat.

RELIGIOUS ACCOMMODATION
Washington state law requires that UW develop a policy for accommodation of student absences or significant hardship due to reasons of faith or conscience, or for organized religious activities. The UW’s policy, including more information about how to request an accommodation, is available at Religious Accommodations Policy (https://registrar.washington.edu/students/religious-accommodations-request/). Accommodations must be requested within the first two weeks of this course using the Religious Accommodations Request form (https://registrar.washington.edu/students/religious-accommodations-request/).

DISABILITY ACCOMMODATIONS
It is crucial that all students in this class have access to the full range of learning experiences. At the University of Washington, it is the policy and practice to create inclusive and accessible learning environments consistent with federal and state law. If you anticipate or experience barriers to your learning or full participation in this course based on a physical, learning, or mental health disability, please immediately contact the instructor to discuss possible accommodation(s). A more complete description of the disability policy of the College of the Environment can be found here. If you have, or think you have, a temporary or permanent disability that impacts your participation in any course, please
also contact Disability Resources for Students (DRS) at: 206-543-8924 V; 206-543-8925 TDD; uwdss@uw.edu; http://depts.washington.edu/uwdrs/.

ACADEMIC INTEGRITY

The University of Washington Student Conduct Code (WAC 478-121) defines prohibited academic and behavioral conduct and describes how the University holds students accountable. We expect that you will know and follow university policies on cheating and plagiarism.

Acts of academic misconduct include:
Cheating: unauthorized assistance in person and/or online for assignments, quizzes, tests or exams, using another student’s work without permission and instructor authorization, allowing anyone to take a course, assignment or exam for you without instructor authorization.
Falsification: intentional use of falsified data, information or records.
Plagiarism: representing the work of others as your own without giving appropriate credit to the original author(s).
Unauthorized collaboration: working with each other on assignments without permission.
Engaging in behavior prohibited by an instructor.
Multiple submissions of the same work in different courses without instructor permission.
Deliberately damaging or destroying student work to gain advantage.
Unauthorized recording, and/or subsequent dissemination of instructional content.

Any suspected cases of academic misconduct will be handled according to university regulations. For more information, see the College of the Environment’s Academic Misconduct Policy and the Community Standards and Student Conduct website.

SAFETY

If you feel unsafe or at-risk in any way while taking any course, contact SafeCampus, 206-685-7233 anytime—no matter where you work or study—to anonymously discuss safety and well-being concerns for yourself or others. SafeCampus can provide individualized support, discuss short- and long-term solutions, and connect you with additional resources when requested. For a broader range of resources and assistance see the Husky Health & Well-Being website.
FISH 560: Final Report and Peer Review
Autumn 2020

PURPOSE
1. Improve your ability to select appropriate multivariate methods for answering ecological questions.
2. Improve your understanding of how the choice of method may influence your results.
3. Advance your graduate research projects.
4. Broaden your experience with multivariate software.

DATASET
The data set may be your own, one obtained from the literature or one provided by the Instructor. Ideally you should use data that you have collected or are otherwise somewhat familiar with. The data set should be one or more matrices of entities \( \times \) attributes (e.g., samples \( \times \) species, species \( \times \) characteristics of species, sites \( \times \) environmental factors, etc.). The only data requirements are that it be adequate to test the hypotheses addressed in your final report. If you do not have access to a multivariate dataset, then I would be please to provide one.

ANALYSIS
Methods of analysis should be chosen to be compatible with hypotheses and data. Procedures can be either exploratory or inferential.

ONE-PAGE PROPOSAL
Your proposal should include the following items:
- Title
- Name
- Date
- Brief background
- Questions or objectives for the analysis
- Source of the data
- Data Structure
  - What is the sample unit?
  - What is each data matrix?
  - Describe any hierarchical structure in the data (e.g., subsampling with plots which are then averaged to a single plot value) and a statement of what level(s) in this structure you plan to do your analyses.
  - If aggregation of the raw data is required to construct a data matrix, describe how you will do that.

FINAL PAPER
The final paper should be written in standard scientific format, and it will be considered for web publication in the virtual journal, the *Electronic Journal of Applied Multivariate Statistics* (EJAMS). Dr. Julian Olden is the Editor-in-Chief of EJAMS and will post all accepted paper on the web for future class use, with your permission. The journal will be available only to students registered in this class. See [http://hdl.handle.net/1773/19723](http://hdl.handle.net/1773/19723).
Format:
- **Title**
- **Author**
- **Institutional affiliation** (Department or Program, University)
- **Email address**
- **Abstract**: 200 word summary (separate page)
- **Introduction**: Background to the problem with enough information to allow me to understand what question you’re asking, why you’re asking it, and why you find it interesting. Explicit statements of research hypotheses/objectives.
- **Methods**: Description and justification of methods of analysis chosen.
- **Results**: Tabular and text summaries of results, including descriptive statistics, test statistics, degrees of freedom, probabilities, and assessments of statistical significance.
- **Discussion**: Biological interpretation of the results, contribution to science and implications.
- **Acknowledgements**
- **References** (follow EJAMS guidelines)
- **Tables**
- **Figures**: Please include all figures **at the end of the document**. Select your figures sparingly. Excess figures should be included in an appendix **ONLY** when necessary.

Other details:
- Upload your final paper (Word format) on the Canvas website by the deadline. Name the Word document by your last name and first two initials, e.g. oldenjd.doc
- Follow the conventions of Standard Written English, which should include correct spelling, grammar, punctuation, capitalization, paragraph structure, and sentence construction
- 15-page maximum (not including figures), 10-page minimum
- Times Roman 10 point font
- Headings in 12 point bold font
- Single spacing
- 1 inch margins on all sides

**PEER-REVIEW OF FINAL PAPERS**

Each student will be randomly assigned to review two submitted papers. Reviewers are required to write a 1-page report and must decide to either accept (as is) or reject the paper. Please qualify your decision. For example, you might point out if a paper has a few misleading flaws that you don’t think should be published for others to follow, but the paper is otherwise in good shape. Note that the average acceptance rate for EJAMS is 20-30%. Please keep this in mind when you are rendering your decision. Be courteous and respectful in your reviews. A good practice for reviewing papers is to pretend that you are writing the review for a colleague who is a friend. You should provide constructive criticism, but it should be done in a friendly, gentle, non-hostile way. The Editor-in-Chief will make his final decision based on the reviewers’ recommendations, and his own review of the paper. **All peer-reviews must be uploaded to the Canvas website by the deadline.**