

# Stat 547N/FISH 506H - Statistics in Ecology and Marine Sciences

## Syllabus

Tuesday/Thursday 14:00-15:30 pm, September 3<sup>rd</sup> – October 17<sup>th</sup> 2019  
Room ESB 4192

Instructor: Marie Auger-Méthé, [auger-methe@stat.ubc.ca](mailto:auger-methe@stat.ubc.ca)

Office hours: Thursday 3:30-4:30 Room AERL 241

*Data in ecology and marine sciences are frequently associated with large challenges. Controlled experiments are often difficult and observational studies are often associated with missing data and measurement error. This class will introduce some of the challenges of using statistics to answer questions in ecology and marine sciences and the statistical tools developed to handle them. Topics covered in this class are: missing data, multiple imputation, censored and truncated data, GLMs, overdispersion, hidden Markov models, and state-space models. This course is a statistics class for graduate students in the Department of Statistics (STAT) and the Ocean and Fisheries Graduate program (OCF), but interested students from other departments should contact the instructor. This class is intended for students with good statistics background and some familiarity with R.*

### Tentative schedule:

Date	Lecture	Paper discussion	R tutorial	Assignment Due
Sep 3 -Tu	No lecture – Imagine UBC	--	--	--
Sep 5 -Th	Intro to class, challenges of ecological data, statistical inference approaches	--	--	--
Sep 10 -Tu	--	p-values	--	Paper summary & questions
Sep 12 -Th	Missing data	--	--	--
Sep 17 -Tu	--	Missing data	--	Paper summary & questions
Sep 19 -Th	--	--	Missing data	--
Sep 24 -Tu	Truncated & censored data	--	--	Final project outline
Sep 26 -Th	--	--	Censored data	--
Oct 1 -Tu	GLMs & Overdispersion	--	--	--
Oct 3 -Th	--	--	GLMs	--
Oct 8 -Tu	--	GLMs	--	Paper summary & questions
Oct 10 -Th	Hidden Markov model	--	--	--
Oct 15 -Tu	State-space models	--	--	--
Oct 17 -Th	Project presentations	--	--	Project report (due Friday 18)

## Assessments:

1. Paper discussions: 30%
2. R tutorials: 30%
3. Final project: 40%

### 1. Paper discussion guidelines and associated assignments:

The course will have 3 discussion sessions. The class will be assigned scientific papers to read in advance, will be asked to hand in an associated assignment prior to the in-class discussion. Each group will be asked to summarise *in their own words* the papers assigned that week. The summary of the papers should be only one paragraph and no longer than 350 words. If multiple papers were assigned for a session, make sure to highlight whether they have contrasting views on the topic. In addition, the group will be asked to write down 2 discussion questions that arose from the reading the papers. Find questions that should spark debate and that are specific to the topic in question. The questions should be inspired/informed by the papers. Please, hand in the written assignment at the beginning of the class.

You will be graded as a group. Make sure all members of the group participate. You will be assigned to a group in the second class.

Topic	Paper	Link <i>*see below for how to access off campus*</i>
<b>p-value</b>		
	Wasserstein RL & Lazar NA (2016). The ASA's statement on p-values: context, process, and purpose. <i>The American Statistician</i> 70:129-133	<a href="http://amstat.tandfonline.com/doi/pdf/10.1080/00031305.2016.1154108?needAccess=true">http://amstat.tandfonline.com/doi/pdf/10.1080/00031305.2016.1154108?needAccess=true</a>
	Ionides <i>et al.</i> (2017). Response to the ASA's statement on p-values: context, process, and purpose. <i>The American Statistician</i> 71:88-89	<a href="http://amstat.tandfonline.com/doi/pdf/10.1080/00031305.2016.1234977?needAccess=true">http://amstat.tandfonline.com/doi/pdf/10.1080/00031305.2016.1234977?needAccess=true</a>
	<i>Dushoff et al.</i> (2019) I can see clearly now: reinterpreting statistical significance. <i>Methods in Ecology and Evolution</i> 10:756-759	<a href="https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/2041-210X.13159">https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/2041-210X.13159</a>
<b>Missing data</b>		
	Ellington EH <i>et al.</i> (2015) Using multiple imputation to estimate missing data in meta-regression. <i>Methods in Ecology &amp; Evolution</i> 6:153-163	<a href="http://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12322/epdf">http://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12322/epdf</a>
<b>GLMs</b>		
	O'Hara RB & Kotze DJ (2010) Do not log-transform count data. <i>Methods in Ecology &amp; Evolution</i> 1:118-122	<a href="http://onlinelibrary.wiley.com/doi/10.1111/j.2041-210X.2010.00021.x/epdf">http://onlinelibrary.wiley.com/doi/10.1111/j.2041-210X.2010.00021.x/epdf</a>

\* Need to either be on campus to access the link or use the EZ-proxy tools from the library if off-campus (<https://services.library.ubc.ca/electronic-access/connect/ezproxy-toolkit/>)

## 2. R tutorials:

These tutorials will be described in a separate document handed on the appropriate day. However, make sure you have R and R Studio installed on your laptop. I am expecting that you are familiar with R. These tutorials are intended to be completed in class. **Bring your laptop to class!**

## 3. Final project:

The goal of the final project is to explore some of the analyses covered in class in more details, and, in particular, to learn to apply these analyses to real data. You can either use your own data or free data available online. This will be a group project and I am open to various project ideas as long as you explore some of the methods covered in class (e.g. GLMs, Multiple Imputation) and show how different ways to analyse the data affect the results. While it is important that at least one of the analyses of the project is one of the methods covered in class, you are welcome to explore other techniques in the final project.

Example 1: use data from the thesis of one of the group members or ecological data online (e.g. on dryad: <https://datadryad.org>) and show how different approaches affect the results (e.g. using transformation vs glms, and exploring methods to account for overdispersion).

Example 2: you could find an ecological paper that has an associated online dataset that you can use to try to reproduce the analyses used in the paper and show how analysing the data in a different ways would affect the results (e.g. compare imputation to doing list-wise deletion of missing data).

The project will be divided into 3 parts.

### 1. Outline (10%)

*Due: Sept 24 at beginning of class*

*1 paper copy per group*

One page summary of your project where you present the goal of your final projects and the analyses you will perform. Make sure to describe:

- data
  - if the data is taken online, make sure to provide a full citation of the data set and paper
- main questions to be answered
- analyses to be performed

### 2. Presentation (40%)

*Due: Oct 17*

*Group presentation in class*

15 minute presentation (12 min talk, 3 min for questions), where you will discuss:

- goal of the project
- dataset
- analyses performed
- comparisons of the analyses in terms of impact on results and conclusions
- recommendation

Note that part of the grade will be assessed by your peers.

Make sure all group members participate.

### *3. Written report (50%)*

*Due: Oct 18 5pm*

*Send me a pdf version by email. One report per group.*

Similar to the presentation, the main goals of this 10-15 pages (double spaced, including figures) written report is to explain the goal of the project and discuss the pros and cons of the methods explored. You should conclude with a recommendation with regards to the best analysis and the interpretation of the results. The format should include: Introduction, Methods, Results, and Discussion.

- Introduction should explain the main goals of paper
- Methods should explain the methods explored
- Results should present the results from all the methods explore and focus on the differences between the different methods.
- Discussion should focus on discussing the pros and cons the different methods, and should make a recommendation on which of the method explored is the most appropriate for the data and question. Make sure to give a clear interpretation of the results.

You will be graded on your understanding of the statistical analyses performed and the quality of report. For me to be able to assess whether you understand the analyses, you need to clearly describe all of the methods you used, including those covered in class. You need to emphasize why you are exploring specific analyses. Because the quality of the report is also assessed, pay attention to grammar, typos, and paragraph structure (e.g. include topic sentences). Verify that you are clear and concise and that your figures and tables are easy to read (e.g. make sure the axes are written with large enough font to be readable and that the axis titles are easily interpretable). Here you should write this report like if it was a scientific publication. Thus, describe your methods with words and equations, not with R code. Similarly, describe your results with words, tables, and figures, not with R outputs.

### **Missed classes, late assignments, and grade changes:**

*I do not provide extension for discussion documents and tutorials, as these are associated with in-class activities. If you do not hand-in your assignment on time or miss a class, the grade will*

*be weighted into your final project (i.e., you won't lose marks, but your final project will count for more). For the final project, if you return your assignment late, I will remove 10% for each day past the deadline. Many of the activities are group based. Please be respectful of others and participate equally. I do not discuss grades unless changing a grade will change more than 1% of the final grade (e.g. for the tutorial, this would mean that I am only willing to discuss a grade change that would rise the grade of this assignment by more than 10%).*

**Further reading:**

The early classes are based on the book *Ecological Statistics: contemporary theory and application* edited by G.A. Fox, S. Negrete-yankelevich, V.J. Sosa (available online: <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199672547.001.0001/acprof-9780199672547>)

You are not required to read the associated chapters, but if you find the material challenging I highly recommend that you do.