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

Syllabus

Tuesday/Thursday 14:00-15:30 pm, Sep 8th – Oct 20th 2022
Room: Neville Scarfe 1005

Instructor: Marie Auger-Méthé, auger-methe@stat.ubc.ca
Office hours: Thursday 15:30-16:30, AERL 241

Most information and assignments accessible via Canvas

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 are welcomed to enroll via FISH 506H. This class is intended for students with a good statistics background and some familiarity with R.

Tentative schedule:

Date	Lecture	Paper discussion	R tutorial	Assignment Due
Sep 6 -*Tu*	*No class – Imagine UBC https://students.ubc.ca/new-to-ubc/orientations/imagine-ubc*			
Sep 8 –*Th*	Intro to class + challenges of ecological data + review of statistical approaches	--	--	--
Sep 13 -*Tu*	Missing data	p-values	--	Paper summary & questions - *submit by noon*
Sep 15 -*Th*		--	Missing data	--
Sep 20 -*Tu*	Truncated + censored data part 1 + Setting groups	--	--	--
Sep 21 -*Th*	Truncated + censored data part 2	--	--	--
Sep 27 -*Tu*	--	--	Censored data	--
Sep 29 -*Th*	GLMs + Overdispersion part 1 + Time devoted to group project	--	--	--
Oct 4 -*Tu*	GLMs + Overdispersion part 2	GLMs	--	Paper summary & questions - *submit*

				by noon
Oct 6 -*Th*	--	--	GLMs	Final project outline
Oct 11 -*Tu*	Mixture models	--	--	Formative peer evaluations
Oct 13 -*Th*	Hidden Markov models	--	--	--
Oct 18 -*Tu*	State-space models	--	--	--
Oct 20 -*Th*	Project presentations	--	--	--
Oct 28 - Fr	(No class)			Project report + summative peer evaluations

Assessments:

1. Paper discussions: 30%
2. R tutorials: 30%
3. Final project: 40%
4. Peer evaluations (contribute to final project grade)

1. Paper discussion guidelines and associated assignments:

The course will have 2 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 student 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 300 words. If multiple papers were assigned for a session, make sure to highlight whether they have contrasting views on the topic. In addition, each student 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, submit at pdf file on Canvas before noon the day of the class discussion.

Topic	**Paper**	**Link** ***see below for how to access off campus****
p-value		
	Wasserstein RL & Lazar NA (2016). The ASA's statement on p-values: context, process, and purpose. The American Statistician 70:129-133	http://amstat.tandfonline.com/doi/pdf/10.1080/00031305.2016.1154108?needAccess=true
	Ionides *et al.* (2017). Response to the ASA's statement on p-values: context, process, and purpose. The American Statistician 71:88-89	http://amstat.tandfonline.com/doi/pdf/10.1080/00031305.2016.1234977?needAccess=true
	Dushoff et al. (2019) I can see clearly now: reinterpreting statistical significance. Methods in Ecology and Evolution 10:756-759	https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/2041-210X.13159
GLMs		
	O'Hara RB & Kotze DJ (2010) Do not log-	http://onlinelibrary.wiley.com/doi/

	transform count data. Methods in Ecology & Evolution 1:118-122	10.1111/j.2041-210X.2010.00021.x/epdf

* 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 available as an assignment on *Canvas* on the appropriate day. Make sure you have R and R Studio installed on your laptop before the first tutorial. I am expecting that you are familiar with R. These tutorials are intended to be completed in class. I will be in class to answer any questions you may have. They will be available as an assignment on *Canvas*.

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 and compare how different methods perform. 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.

You will be assigned to a group in the fourth class. There will be a peer-review process, where your participation will be assessed by the other members. Information will be provided in class.

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 way would affect the results (e.g. compare imputation to doing list-wise deletion of missing data).

The project will be divided into 3 parts.

i. Outline (10%)
Due: Oct 6 at beginning of class submit via Canvas. One outline 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

ii. Presentation (40%)
Due: Oct 20
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

While it is a group presentation, you will be graded individually based on the complete presentation and your own presentation skills. Make sure all group members participate and that you help each other create a good overall presentation.
Note that 10% of the grade will be assessed by your peers.

iii. Written report (50%)
Due: Oct 28 5pm submit via Canvas. 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 explored 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). 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. ***I will expect the students to make changes according to the written comments I have made after the presentation.***

4. Peer evaluations:
The final project, which is a core component of the class, is a group assignment and many in-class activities are done in groups. Throughout the term there will be peer evaluations made to received feedback from your peers and assess your contribution to the group work. In brief, you will be asked to write and sign a group contract (to be submitted via Canvas) and then you will be asked to fill two sets of peer evaluations (to be filled via ipeer, see Canvas).

i. Formative peer evaluations
The first set of peer evaluations will be not be anonymous and serve as a way for your team to give constructive feedback to other team members and help everyone improve.

ii. Summative peer evaluations
The second set of peer evaluations will be used to calculate your team skill grade on the final project report. 10% of the final report grade will be allocated based on the peer evaluations. The calculation will be: I will take the average of the % points the team members allocated you. If your average % points is higher or equal than to the % points that would be allocated for equal contribution, you will get the full 10 %. If your average is lower than the one that would be allocated for equal contribution, you will get the relative amount.
Example 1, you are in a team of 3 and thus in each evaluation you are compared to one other team member. If your average % points is 50%, you will get 50/50 = 1 (full) of the 10% (i.e. 10%).
Example 2, you are in a team of 3 and thus in each evaluation you are compared to one other team member. If your average % points is 40%, you will get 40/50 = 0.8 of the 10% (i.e. 8%).
If your allocation is really low, I may have to further discuss with you and the members of the group. You are expected to participate almost equally.
It is due at the same time as the final report.

Teams with students auditing
While auditing students are not expected to participate in the final project, you should all include auditing students in your evaluations. Keep in mind that it is ok to give the auditing students a 0% allocation, as they are not expected to actively participate. You do not need to ask them (nor should you expect them) to participate or improve. However, I ask auditing students to actively evaluate the participations of students taking the class for credits. So please cc them on all emails, so that they can see who is contributing the most to the group project. **Auditing students:** providing the complete peer evaluations is the only portion of the final project I expect you to complete and do well. Please be fair and objective in your evaluation of the contribution of other students. I will make sure to adjust the grading of group contribution to account for the fact that auditing students are not participating.

iii. Peer review of presentation.
In addition, I will ask student to peer review the final presentation of the other groups. More information is available on Canvas.

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 due to valid reasons (e.g. you are sick), the grade will be weighted into your final project (i.e. you won't lose marks, but your final project will count for more). If you have a health-related reason to be late for your final project, I will provide an extension. Otherwise, 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.

Further reading:

The first set of lectures 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 recommend that you do.

Statement about UBC's values and policies:

UBC provides resources to support student learning and to maintain healthy lifestyles but recognizes that sometimes crises arise and so there are additional resources to access including those for survivors of sexual violence. UBC values respect for the person and ideas of all members of the academic community. Harassment and discrimination are not tolerated nor is suppression of academic freedom. UBC provides appropriate accommodation for students with disabilities and for religious, spiritual and cultural observances. UBC values academic honesty and students are expected to acknowledge the ideas generated by others and to uphold the highest academic standards in all of their actions. Details of the policies and how to access support are available senate.ubc.ca/policies-resources-support-student-success.

We acknowledge that the UBC Vancouver campus is situated within the traditional, ancestral and unceded territory of the xʷməθkʷəy̓əm (Musqueam).