

Student's *T* Time

Professor Ranjini Grove:

Welcome back from break and a happy new year to all!

The second edition of Student's t-time, titled *Making Connections* profiles students who have been able to foster connections within the U.W. community as well as the statistics community at large. I commend these students for their perseverance and determination in “putting themselves out there” and taking risks.

In other news, the statistics capstone STAT 423 is off to a great start as you can see from the photos taken during snack break in between our Wednesday double-header session. Seniors will present their final projects at the joint CSSS/STAT poster session for the first time ever, so stay tuned for more details.

Ask Mee Ling:

Question: “I got an internship and would like to apply for Curricular Practical Training (CPT). What should I do?”

Answer: “Congratulations! First of all, you will need to send the job descriptions and the job offer letter to Prof Ranjini Grove (grover4@uw.edu) for approval. She will make sure the job is related to statistics. With her approval I will help you apply for the CPT training. Please do not start to work until the CPT training is granted by the International Student Services. You will also need to sign up for CPT related credits, that is, STAT 499. Please discuss with Ranjini for the number of credits for STAT 499 and the expectations for earning the credits.”

Annie Wan From the S.P.A Desk:



Happy New Year to the students, faculty and staff of the Department of Statistics! I hope you all had a great holiday

W

break and gained a good amount of weight to insulate yourselves from this chilly weather. SPA is kicking off 2019 by planning a few exciting events. The very first Women in STEM Panel will be held some evening between February 25th and March 1st. Additionally, SPA is planning on hosting a day of volleyball and quite possibly a movie night with the graduate student social group. Lastly, we are currently designing T-shirts, hats and socks to sell so the Stat Department can be represented proudly across campus.

All in all, there are quite a few things to look forward to. If you're interested in helping out SPA, getting to know other students in the stat major, and gaining communication skills, feel free to attend our weekly meetings on Wednesdays at 5 pm in PDL C301. Best of luck to you all in Winter Quarter 2019!

(Right) Benjamin Xie, a doctoral student in the Information School taught STAT 423 students about dplyr and ggplot. (Below) Scenes from snack break



“SAVE THE DATE!”
April 5 -- 7:
DataFest
2019!
More soon.



KUDOS!
New Phi Beta
Inductees:
Emily Flanagan,
Jiaqi Zhu
and Junming Shi



News from Afar

Sam Stockman:



Last year I was lucky enough to spend the whole academic year on exchange at UW from my home University in Bristol in the UK. It was so amazing to live and study in a

different country and I'll never forget the memories I made over the 9 months I spent there. It both opened my eyes up to world as well as put me out of my comfort zone and I would highly recommend it to anyone who has the opportunity.

I am a mathematics student specialising in statistics, so throughout my time at UW I took mainly stats classes. I therefore joined the stats majors for what was to be their final year. This was quite a different experience to what I was used to at my home university.

My statistics classes at home usually involve much more independent study throughout the quarter with a final exam at the end. I therefore found the constant assessment a little hard to adjust to at first but I actually found that by the end I preferred it.

I enjoyed being part of a small group of students that somehow managed to achieve the sense of community that helped me settle in, as well as the level of competitiveness that helped me push myself to get the most out of my time. It being the final classes for the stats majors, I got the chance to do the capstone project, a small research project as part of the linear models class. This was a great way to conclude my time amongst the rest of the class, although I'm sure for them it was a bigger deal since it was part of the final steps towards graduation.

STAT 499

Aaron Huang:

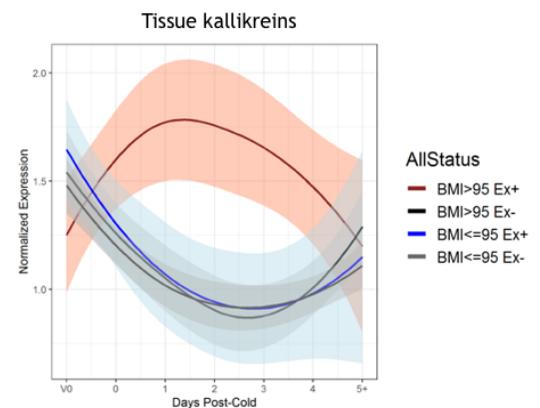
Faculty advisor: Dr. Matt Altman

For the past several months, I have been working with Dr. Matthew Altman, an allergy and immunology researcher at the University of Washington Medical Center. I found this research position by scrolling through the undergraduate research database and sifting through different positions. Research professors generally responded quickly, provided the listed positions were not older ones that simply had not been taken down. I contacted Dr. Altman, and was able to get started right away.



Dr. Altman is part of a NIH consortium conducting clinical studies investigating severe asthma in children. We analyzed a dataset containing longitudinal whole genome expression data collected at three time points: baseline when asthma is stable, at the start of a cold, and at the start of an asthma exacerbation. What makes this dataset exciting is that the data was summarized using network analysis approaches to determine clusters of highly correlated genes. Each cluster of genes could then be summarized to a single expression value, reducing the number of dimensions and the number of models we needed to construct. This allowed us to improve the statistical power and aid in the interpretation of results. This dimension reduction technique is part of a relatively new field of biostatistics known as systems biology.

I helped Dr. Altman work through secondary subgroup analyses of this dataset. We found that obesity was a statistically important and clinically relevant covariate (see fig). Our write-up of these results was selected for an oral presentation at the national meeting of the American College of Allergy, Asthma, and Immunology which I attended this past November.



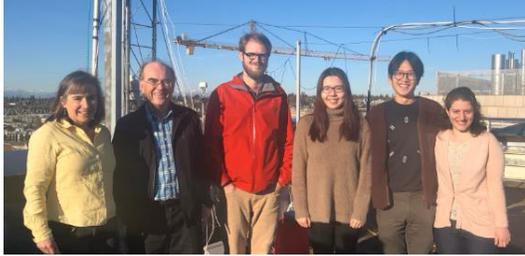


Zihui Zhang:

Faculty advisor: Adrian Raftery

Graduate Student Mentor: Hannah Director

My research project is about bias correction methods for predictions on sea ice extent. Sea ice is ice formed by the freezing of seawater. It is the masses of floating ice drifting in the ocean.



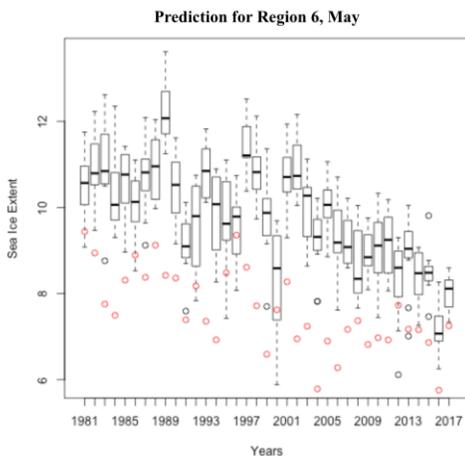
Sea ice extent is a measure of the surface area of the ocean covered by sea ice. Sea ice in the

Arctic is declining rapidly due to climate change and there are effects are on wildlife, ecosystems, and even human society. Although we are now able to predict sea ice reasonably well using deterministic models run on supercomputers, sometimes the predictions do not match the observations. My job is to find out where these differences are and how we can use various bias correction techniques to minimize the differences.

The first step I took is using graphs to find out where these differences in the sea ice extent appear. I made box-plots of this value for the 12 samples that are generated by the deterministic model and added the observations as points onto the plot. Now that I have found the pattern, I am going to use linear regression to shift the predictions and make them closer to the observations. I will then then use rank histograms to assess the calibration.

My research is a part of a large research group on sea ice and statistics led by Prof.

Cecilia Bitz, UW Atmospheric Sciences, and Prof. Adrian Raftery, UW Statistics and Sociology. People in the group are doing analyses on different aspects of sea ice, such as extent and thickness. I meet weekly with Hannah Director, my graduate student mentor, and we meet as a large group once a



month. During the meeting, we share our progress and problems we encountered and give suggestions to each other. I am glad to work with a group of intelligent people, and look forward to seeing what I can achieve in this research.

Research Project with Caren Marzban

Two students in STAT 390 - one our major, one not (yet!) - wanted to work on a research problem with Caren Marzban. They are Haoran Yu (front/left in the photo) and Jueyi Liu (front right). So, they signed up for STAT 499 and started working, in Spring and Autumn of 2018 respectively. Already they have produced very good work, which is probably even publishable in some good research journal. Coming out of STAT 390, they had to build on what they had learned there.

Haoran worked on (re)formulating the 1-Way ANOVA problem as a classification problem. The idea dates back to a little-advertised preprint by Jerome Friedman (who used to be a physicist) wherein it was shown that one can cast the standard 2-sample testing problem as a machine learning problem:

<https://www.slac.stanford.edu/econf/C030908/papers/THPD002.pdf>. Friedman addresses the multivariate situation as well,

but he left open plenty of issues and questions. Werner Stuetzle is acknowledged for bringing Friedman's paper to Caren's attention.



Jueyi wanted to learn more about machine learning. The main question she addresses is How does the variability of a performance measure (say, mean-squared-error) due to resampling (e.g., from cross-validation) compare with the variability due to local minima? Most machine learning approaches involve some sort of nonlinear optimization, and there is a chance that the machine gets stuck in a local minimum. Usually, this source of variability is ignored by practitioners. Jueyi used mixed effects models to estimate and compare the two variance components. Of course, there is no simple answer to the question asked, but some of the results are pleasantly surprising. Caren (who also used to be a physicist) himself is acknowledged (blamed) because that idea dates back to a preprint he wrote (when he was a physicist) - a preprint that was described as "novel and interesting," but rejected by NIPS.





Conferences

Vicente Velasco:

My experience going to the SAMSI Undergraduate workshop was one I will never forget. Being able to fly to North Carolina was amazing as I had never been before. SAMSI offers multiple workshops for undergraduates. The one that I attended was a 2-day workshop, but they also offer one that is a whole week. This workshop itself focuses on the applications of mathematics and statistics to precision medicine, but not all the seminars focused on that. The first workshop that I thought was helpful was an introductory workshop on the programming language R. While I have some experience in R already, it was helpful to revisit and learn from the beginning. One of the best parts of the program in my opinion was the panel they hosted with the post-docs that worked in SAMSI. It was helpful for me to hear about the experiences of people. They offered great advice on what grad school is like and how to find research opportunities to help prepare for graduate studies applications. While spending most of my time-off at the hotel, I was also able to go down to Raleigh and explore downtown a little bit with some friends I had made in the program. Coming back, I know I gained not only valuable skills and ideas for where I am heading towards in my career, but I also gained experiences that I will never forget.

Stephanie Wu:

Rarely can one find opportunities to attend fully funded statistics workshops as an undergraduate, so I was immediately intrigued when I read an email from Mee-Ling, our wonderful adviser, advertising such an opportunity at the Statistical and Applied Mathematical Sciences Institute (SAMSI). SAMSI is a partnership between the National Science Foundation, Duke University, North Carolina State University, and the University of North Carolina at Chapel Hill. It hosts various undergraduate workshops throughout the year—in late February, late May, and late October—that focus on current topics in statistics and applied mathematics.

Applying to one of these workshops is very straightforward: the application only requires you fill out some personal and academic information and submit a “statement of interest” describing why you want to attend the workshop (no minimum or maximum word count). If you are accepted, you are automatically granted full funding for travel, food, and hotel expenses.



The workshop I attended was a two-day event centered around the topic of Statistical, Mathematical, and Computational Methods for Precision Medicine. Precision medicine is a fascinating new field that uses individuals’ genetic information to tailor medical treatment and diagnosis. I was intrigued by the topic not only because it is a burgeoning field in biostatistics, but also because I am studying Public Health and have been learning about the ethical implications of these new genetic technologies. The workshop featured talks, led by students and faculty affiliated with SAMSI institutions, that provided an overview of current R software, mathematical models, and statistical methods used to analyze precision medicine data. We also heard sage advice from Lisa LaVange, president of the ASA, on how statisticians can be leaders, and we were given the chance to seek life and career advice from current graduate students in statistics. Though we listened to many presentations, there was also plenty of time to socialize and form surprisingly long-lasting connections with students from around the country who share a passion for statistics.