Goodsell Gazette

Carleton College *12 May 2017*

Northfield, MN 55057 *The newsletter for the Carleton mathematics and statistics community* *Vol. 35, No. 19*

Seventh Week Update from the Math & Stats Department

Lots is happening here in the Math & Stats Department; read on to learn about the group comps presentations happening next week, to hear exciting news on math and stats careers, and to stay updated on conference, job, and internship opportunities.

Group Comps Talks

Students who elected to pursue a Winter/Spring groups comps project will present their findings Tuesday, May 16 and Thursday, May 18 in OLIN 141. Take a look at what they'll be speaking about just below, then be sure to stop by and support them while they share what they've learned with the Math & Stats Department; you're likely to learn something new yourself as well! Then stay afterwards for dinner catered by Bon Appetit.

Tuesday, May 16

Title: Reducibility, Stability, and Polynomial Iterates
Speakers: David DeMark, Moses Misplon, Michael Stoneman
Time: 3:30 p.m.

We investigate the reducibility of monic quadratic polynomials with no linear term and rational coefficients. We conjecture that unless the constant term is 0 or -1, the polynomial is stable or eventually stable--i.e., the number of irreducible factors any iterate of the polynomial decomposes into is bounded. We provide proofs for the conjecture in a variety of special cases. Many of these are quite easily dispatched with a natural generalization of the Eisenstein criterion; however, other cases require deeper results from Galois theory and rely on properties of divisibility sequences or a parametrization for "near-Pythagorean" triples. Using this parametrization, we also present an algorithm to determine in finite time whether any given such polynomial is stable.

Title: Using Time Series Methods to Estimate Temperature Trend Uncertainty Due to Internal Variability in a Climate Model
Speakers: Melissa Bain, Caitlin Eichten, Il Shan Ng, Lianne Siegel
Time: 4:30 p.m.

In order to identify systematic changes in regional and global mean temperatures, it is important to account for uncertainties that arise from internal variability in the climate system. Physical processes that are temporally correlated may produce apparent trends even in the absence of any systematic changes. One way of estimating variation in temperature trends due to internal variability involves simulating an ensemble of climate model runs by using the same model and forcings for each run, but seeding the initial conditions differently. The variation in trends across ensemble members is a measure of trend uncertainty due to internal variability. These simulations, however, are costly and time-consuming. Statistical time series methods, which do not require large ensemble simulations, provide another way of estimating the variation in temperature trends. For example, Thompson et al. (2015) modeled internal variability of regional temperatures in a single equilibrated climate model simulation as an AR(1) process and demonstrated that this model can be used to derive good approximations of the trend variation across an initial condition ensemble of forced 21st-century climate model runs. In this paper, we show that global, rather than regional, temperatures in the same climate model studied by Thompson et al. (2015) are better described as an ARMA(3,1) process than as an AR(1) process. We compare the estimates of trend variation derived from this

ARMA(3,1) process, over a range of trend lengths, to those derived from an AR(1) process. We find that an AR(1) model overestimates trend variation for short trend lengths and underestimates it for longer trend lengths compared to the ARMA(3,1) model, which performs well for a wider range of trend lengths. We explain these differences in terms of how the two models capture quasi-periodic and long-timescale variability differently. Our findings highlight the importance of specifying a statistical model that captures the inter-temporal correlation structure when estimating temperature trend uncertainty.

Time: 5:30 p.m.
Dinner will be provided for all who attend the talks.

Title: Shift Spaces and S-limited Shifts
Speaker: Ben Matson
Time: 6:00 p.m.

When storing data on a CD or DVD, there is a specified minimum and maximum number of zeroes that can come between consecutive ones. This idea has been generalized into what are called S-gap shifts, where there is some set S so that there must be some number of zeroes in S between consecutive ones. In this talk, I will give an overview of S-gap shifts, and will also present a further generalization of this idea, called an S-limited shift, that can describe strings using more characters than only zeroes and ones.

Thursday, May 18

Title: Persistence is Key: Using Topology to Determine the Shape of Data
Speakers: Thomas Redding, Nathan Bern, Nate Osher
Time: 3:30 p.m.

Given a large number of data points, how can we determine its "shape?"Â In order to answer this question, we will employ techniques from topology-specifically the notion of "persistent homology".Â Over the course of our talk, we will explain how to compute persistent homology, how to interpret the results visually, and analyze a dataset regarding the orientations of Spinel crystals observed in New Caledonia by Carleton's Geology Department. We will conclude by discussing our results, and the direction we hope to take our research in the future.

Title: /r/comps
A Statistical Analysis of Network Data from Reddit
Speakers: Graham Earley, Nikita Fomichev, Willa Langworthy, Ruyi Shen
Time: 4:30 p.m.

Network structures are everywhere, from social networks to health epidemics. When making statistical models, it is important to be able to account for a network structure, since network data violates the assumption of independence that a regular linear model requires. In this talk, we explore descriptive statistics of networks, purely mathematical models of networks, and Exponential Random Graph Models (ERGMs) -- a statistical modeling method that accounts for network structures. Finally, we construct two ERGMs from network datasets that we gathered from two politically charged subreddits: the subreddit for Donald Trump supporters and the subreddit for Bernie Sanders supporters. Using sentiment and text analysis of Reddit comments in our model, we are able to quantify how likely a user is to comment positively in a comments section that consists largely of positive comments, and vice-versa. We find that commenting positively in an already negative discussion is less likely than commenting negatively in an already negative discussion. Additionally, we discuss structural and numerical characteristics of our networks, in the hope of better understanding how Redditors interact with each other in these political environments.

Time: 5:30 p.m.
Dinner will be provided for all who attend the talks.

Check out these new job rankings!

Careercast ranks best jobs of 2017:

#1 Statistician
#5 Data Scientist

#7 Mathematician
#11 Actuary

http://www.careercast.com/jobs-rated/2017-jobs-rated-report

Job & Internship Opportunities

Great Lakes Analytics in Sports Conference

Have you conducted research or accumulated experience in the field of sport analytics? Are you looking for a forum to present your findings and gather feedback? Hoping to connect with others in the field? Consider presenting at the Great Lakes Analytics in Sports Conference.

Scheduled for July 13 at the University of Wisconsin-Stevens Point, GLASC will feature an opportunity for students, faculty, professionals, and others affiliated with sports analytics to present poster or oral presentations on topics as diverse as mathematics, athletic training, business, psychology, biology, computer science, GIS, journalism, technology, and data visualization. Research papers resulting from these presentations will be submitted for a special issue of the Journal of Sports Analytics.

Deadline for abstract submission is Monday, May 15. Click here to submit a talk or poster. Contact GLASC director Scott Tappa with questions.

Capstone Program in Actuarial Science

Are you a soon-to-be graduate uncertain about your career options or desires? UW-Madison offers a Capstone certificate program in actuarial science that helps students prepare for an actuarial career. The certificate program is a great option for students who have completed a bachelor's degree, have excellent math and statistics skills, seek a well-compensated profession in a growing field, and want to launch an actuarial career quickly.

Applications for starting the program in Fall 2017 are due June 1, 2017 (there is some flexibility in this deadline). The program consists of five classes and is usually completed in 9 months by full-time students. Students can find more information on the Capstone certificate program here: https://advanceyourcareer.wisc.edu/capstone-certificates/actuarial-science/.

In addition to becoming prepared to pass the initial professional exams, Capstone students can take advantage of our risk and insurance career fair (60+ employers attend) and networking events to help connect them with prospective employers. Graduates of UW-Madison's actuarial science program are highly recruited.

If you have questions about the Capstone program in actuarial science, please contact Gordon Enderle at genderle@wisc.edu. Please note that this is a certificate program and not a degree program.

Marketing Research Intern

The Marketing Research Intern at Scholarship America, Minneapolis, MN is responsible for providing reporting and actionable analytical insights on market trends, partnerships and competitors to support key business decisions. They will have the opportunity to work year-round for the identified term (summer, fall, and winter).

Applications will be accepted until May 12. For more information and to apply, visit: https://carleton-csm.symplicity.com/students/index.phps=jobs&ss=jobs&mode=form&id=fd82d02729243cff875716727f1cba32

Full Time Teacher

Southern Teachers was founded in 1902 and is the oldest teacher-placement service in the country, working with about 600 private/independent, college-preparatory schools around the South: from Maryland to Miami to Midland, Texas.

Southern Teachers currently has over 150 math openings that they are trying to fill, and new openings are being listed with every day.

It's important to know that these schools can consider candidates who have not gone through an education program or student teaching. They typically can consider math majors and minors with strong academic records. Southern Teacher services are free for candidates; the schools support and pay for Southern Teacher's efforts because they trust their professional expertise they've developed over more than a century of work.

If you are seeking a full-time teaching job, you can find more information and learn how to apply at the Southern Teacher website.

Problems of the Fortnight

Having trouble seeing the problem of the fortnight? Try enabling images for the message.

PROBLEMS OF THE FORTNIGHT

To be acknowledged in the next *Gazette*, solutions to the problems below should reach me by noon on Tuesday, May 23.

1. In the "modern" table tennis scoring system introduced some years ago, each game is won by the first player to reach 11 points in that game, provided there is at least a two-point margin. If the score of 10-10 is reached, from then on the players alternate serving until one of them has gained a two-point advantage over the other (and thereby wins the game). Every serve results in one or the other player, but not necessarily the server, scoring one point.

Lucy and Linus have found that they are very closely matched at table tennis, so that they often reach 10-10. From that point on, how long (how many serves) can a game be expected to last:

a) if for each serve, the player serving has probability p of winning the point;

b) if, instead, for each serve, Lucy has probability p of winning the point, whether or not she is the one serving?

2. A few years ago, Sam Ihlenfeldt '14 told me that he had seen the amazing (but true) fact on the Internet that

$$\frac{1}{998001} = .\overline{000\,001\,002\,003\,004\,005\cdots 995\,996\,997\,999}\,;$$

that is, the decimal expansion of the fraction on the left has the repeating pattern shown on the right, going through *all* three-digit combinations in order through 997, then skipping 998 and repeating back to the beginning after 999. Explain this phenomenon, and show how to generalize it from "three-digit combinations" to "n-digit combinations". For example, for $n = 10$, what rational number would have a decimal expansion that starts off

$$.00000\,00000\,00000\,00001\,00000\,00002\cdots$$

and continues in this vein until it reaches

$$\cdots 99999\,99996\,99999\,99997\,99999\,99999\,,$$

then repeats back to the beginning? And how do you know that this will work for any n?

It's a pleasure to report that Yuki Segawa solved the second problem posed March 31; he should stop by CMC 217 at his convenience to select a B.B.O.P. item. As for more recent problems, there was a serious attempt (that can perhaps be salvaged) on the first problem posed April 28; no other student solutions have come in. Meanwhile, John Snyder in Oconomowoc solved both problems posed April 28.

- Mark Krusemeyer

Editors: ***Elaina Thomas, Steve Kennedy***
Problems of the Week: ***Mark Krusemeyer***
Web & Subscriptions: ***Sue Jandro***