

# OREGON CHAPTER American Statistical Association

## Fall 2017 Meeting

## **October 19<sup>th</sup>, 2017**

Oregon State University - Corvallis, OR

#### Minutes of the Spring 2017 Meeting (submitted by Lucas Beverlin) Thursday, May 18<sup>th</sup>, 2017

- 1. Location: Szechuan Chef Portland, OR
- 2. Participants: 23
- Treasurer's report by Lucas Beverlin. The starting balance at the Winter 2017 meeting was \$8,929.58 (February 8<sup>th</sup>, 2017). The ending balance at the Spring 2017 meeting was \$8,925.40 (May 18<sup>th</sup>, 2017).
- 4. Joseph Maurer (OSU) gave the presentation: "Can Bird and Bat Mortality at Wind Projects be Estimated Based of Road and Pad Searches?"

### <u>Agenda</u>

- 6:00 PM Dinner
- 6:45 7:15 PM Business Meeting
  - Welcome from the president Treasurer's report Update on upcoming workshops and other activities Voting for new secretary
- 7:15 7:45 PM Presentation by Katherine McLaughlin
- 7:45 8:15 PM Presentation by Chester Ismay

#### Titles and abstracts of the talks

Title: Models for the Respondent-Driven Sampling Recruitment Process

**Abstract**: Respondent-driven sampling (RDS) is a network sampling method used worldwide to access hidden or hard-to-reach populations that are not reachable using traditional samples. In RDS, study participants recruit their peers to enroll, resulting in a sampling mechanism that is partially unknown to researchers. Additionally, recruitment chains are often initiated by seeds chosen via a convenience sample, so the sampling mechanism is not ignorable. It is therefore necessary to model the RDS recruitment process to obtain inclusion probabilities used for design-based inference. In this talk I discuss several commonly used models for the RDS recruitment process, including the Markov chain sampling process, the successive sampling process, and the homophily configuration graph. I then propose a new model, the rational-choice preferential recruitment (RCPR) model, which incorporates preferential recruitment based on observed nodal or dyadic covariates. Inference is made about recruitment preferences by maximizing the likelihood of the observed recruitment chain given the covariates.

**Title**: Something old, something new, something borrowed, something blue: Ways to teach data science (and learn it too!)

**Abstract**: How can we effectively but gently launch new students into the world of data science? In this talk, I will discuss the ways that Albert Y. Kim and I have gone about approaching this by creating an open source, fully reproducible textbook using the bookdown package at ModernDive.com. The textbook uses the paradigm of books as versions instead of editions featuring an introductory "getting started" chapter with links to many videos and interactive content available on DataCamp.com to support new R users. I'll also discuss how we used #chalktalk (instead of slides) to slow down our instruction to help beginners grasp tidyverse and coding concepts. I will take a glimpse into the new infer package for statistical inference, which performs statistical inference using an expressive syntax that follows tidy design principles. Lastly, I'll demonstrate vignettes and R Markdown reports that our students created to further support the emerging tidyverse community ecosystem and I'll provide future goals for our ModernDive.com project.

#### About the speakers:

Dr. McLaughlin's research interests are in the area of survey sampling methodology, social network analysis, especially network and adaptive sampling techniques for hidden populations. She is also interested in social science applications of statistics. Before joining Oregon State University in 2016, she completed her PhD in Statistics at the University of California, Los Angeles. Her thesis focused on a rational-choice preferential recruitment model for respondent-driven sampling. Working collaboratively with the Hard-to-Reach Population Methods Research Group and the World Health Organization, she developed new statistical methodology geared toward improved estimation of hidden populations, including those at high risk for HIV/AIDS.

Chester builds (and helps instructors build) R and SQL courses for DataCamp. He obtained a PhD in Statistics from Arizona State University and has taught college/university courses and led workshops in mathematics, computer science, statistics, and sociology. Before joining DataCamp, Chester worked first as an actuary, then as a professor, and most recently as a statistical/data scientist consultant in academia. In addition, he has worked as an R consultant for actuarial firms, companies, and the Portland Trailblazers NBA team. He is co-author of the <u>fivethirtyeight</u> R package and author of the <u>thesisdown</u> R package. He is also a co-author of <u>ModernDive</u>, an open source textbook for introductory statistics and data science students using R.

#### **Treasurer's Report**

Account Checking (XXXX-8699) Saving (XXXX-9788)	
Previous Balance (as of 05/18/2017)	8925.40
Membership Dues	+533.00
Interest	+0.55
Expenses	-0
Balance (as of 10/19/2017)	9458.95

#### Fall 2017

#### **Oregon Chapter Officers**

Duo Jiang, PhD Assistant Professor Oregon State University fuentesc@stat.oregonstate.edu Secretary and Treasurer

Miguel Marino, PhD Assistant Professor Oregon Health & Science University marinom@ohsu.edu President Lucas Beverlin, PhD Intel Corporation Iucas.p.beverlin@intel.com Vice President

Jessica Minnier, PhD Assistant Professor Oregon Health & Science University minnier@ohsu.edu Chapter Representative