

Jill Lundell, PhD, PSTAT

Dana-Farber Cancer Institute
Department of Data Sciences
3 Blackfan Circle
Boston MA 02115
jlundell@gmail.com
www.jilllundell.com

Education

PhD, Mathematical Sciences , Utah State University, Logan, Utah Dissertation: Tuning Hyperparameters in Supervised Learning Models and Applications of Statistical Learning in Genome-Wide Association Studies with Emphasis on Heritability. Advisor: Dr. D. Richard Cutler	2019
M.S., Statistics , Utah State University, Logan, Utah Thesis: On the Model Selection in a Frailty Setting Advisor: Dr. Olcay Akman	1998
B.S., Mathematics , Utah State University, Logan, Utah	1996

Programming Languages and Computer Skills

Proficient	Machine learning, classification, prediction models, statistical genetics, bioinformatics, visualization, proteomics, immunology, Python, R, single cell and bulk data analysis, genomics, epigenomics, cluster computing, cloud computing, LaTeX, Linux/Unix, git, bash, GATK, samtools, plink, other bioinformatics software
Working Knowledge	Deep learning
Exposure	C, C++

Experience

Postdoctoral Research Fellow, Harvard T.H. Chan School of Public Health and Dana-Farber Cancer Institute Department of Data Science Boston, Massachusetts Mentor: Dr. Rafael Irizarry <ul style="list-style-type: none">Develop and implement data science technologies in collaborative research at Dana-Farber Cancer Institute.Develop new data science methods in omics and machine learning.Organize the Quantitative Issues in Cancer Research Working Seminar.Assist in grant writing for collaborative work and seek funding for own research.Write and promote Bioconductor packages for research projects.Publish articles.	2019-Present
---	--------------

Senior Statistician, Water and Environmental Technologies

2019-Present

Butte, Montana

Water and Environmental Technologies is a small geoscience, environmental, and engineering firm located in Montana.

- Provide statistical support for projects throughout the company including all of the statistical design and analysis for groundwater monitoring of five large coal powerplants.
- Write R packages to streamline analysis and ensure reproducibility for statistical projects that are ongoing at Department of Energy laboratories and other environmental sites.

Senior Statistician, North Wind, Inc.

2017–2019

Idaho Falls, Idaho

North Wind specializes in environmental management, scientific consulting, and construction.

- Managed all statistical activities at North Wind Group and affiliated North Wind companies.
- Created sampling designs, identified data needs, performed all statistical analyses, and wrote analytical reports.
- Wrote proposals to obtain external funding from private companies and government agencies.
- Provided statistical expertise on projects in many disciplines including nuclear energy, ecology, environmental remediation, and environmental monitoring.
- Managed all non-routine sampling activities for Idaho National Laboratories.

Senior Statistician, Portage, Inc.

2000–2017

Idaho Falls, Idaho

Portage, Inc. specialized in environmental management, engineering and scientific consulting, and information technology.

- Performed the same duties as at North Wind, Inc.
- Created and taught short courses to clients and employees on introductory statistics, including sampling design and hypothesis testing.
- Wrote proposals to obtain new statistical work for Portage and was a key contributor to proposals for large company contracts. Funding was secured from agencies such as the Environmental Protection Agency, U.S. Army Corps of Engineers, Bureau of Land Management, and private companies.
- Wrote business plans to propose and develop new company services and software related to statistics and data technologies.

Faculty, Brigham Young University Idaho

2011–2014

Rexburg, Idaho

- Taught undergraduate mathematics and statistics to classes of up to 50 students.
- Assisted in the creation of a statistics minor for the university with two other instructors.
- Developed a new introduction to statistics curriculum with a team of four instructors and curriculum designers.

- Created new upper division statistics courses.
- Created a training module for faculty and students to teach them how to use R.
- Developed new educational strategies for teaching collaboration skills and computational statistics.

Temporary Lecturer and Consultant, Utah State University

Logan, Utah

1998–2000

- Taught undergraduate and graduate level classes in mathematics and statistics.
- Ran the university statistical consulting center with Dr. Richard Cutler. The center provided statistical consulting to all faculty and graduate students doing research at the university.

Graduate Research Assistant, Los Alamos National Laboratories

1997

Los Alamos, New Mexico

- Provided statistical analysis for a Hepatitis C phylogenetic project in the Theoretical Biology group. I was the only statistician in the group so I had to research the methods that were needed for the project. Methods included bootstrapping genetic code, classifying phylogenies, and bootscanning to look for mosaicism in viral RNA.
- Discovered that a type of hepatitis C that had been identified as a separate species was a short mutation in the RNA for a small group of people in Indonesia. That group was reclassified as a subtype of a different species because of this discovery.

Published R Packages

cytofQC, a package for labeling CyTOF data for cleaning.

2022

cytofQC can be downloaded from Github and is currently under review with Bioconductor. The package can import an fcs file and identify events that are beads, debris, aggregates, or permeable using a statistical modeling approach. Events are labeled with the most likely event type and metrics are provided that indicate how much each event resembles the different event types. Plotting and reporting functions are also included.

VDJdive, a package for analyzing immune receptor repertoire data.

2022

VDJdive can be downloaded from Bioconductor. The package contains functions for reading immune receptor repertoire data, merging it with paired single cell data, quantifying clonotype abundances, calculating diversity metrics, and producing graphs. It implements an E-M algorithm to make use of ambiguous cells for improved quantification.

EZtune: a package for auto tuning supervised learning models.

2018

EZtune can be downloaded from CRAN. The package contains functions for automatically tuning and validating support vector machines, gradient boosting machines, adaboost, and elastic net. Functions are designed to be easy to use for novice R users and can find a well-tuned model with minimal computation time.

Selected Talks

- Street K, Movassagh M, **Lundell JF**. Single-cell TCR-seq data analysis. Dana-Farber Cancer Institute Neovax Scientific Meeting, Boston, Massachusetts. November 2021. Virtual. 2021
- Lundell JF**, Street K. Where did my cancer cells go? cytoQC for better CyTOF data cleaning. Bioconductor Conference, Seattle, Washington. July, 2022. 2022
- Lundell JF**. There has to be an easier way: a simple alternative for parameter tuning of supervised learning methods. Joint Statistical Meetings, Baltimore, Maryland. August 2017. 2017

Honors and Awards

- Presidential Doctoral Research Fellow at Utah State University 2015-2019
- Second place winner of the 2018 Data Expo at the Joint Statistical Meetings (JSM) in Vancouver, British Columbia. This competition is sponsored by the Section on Computational Statistics and Graphics of the ASA and is held approximately every 3 years. The competition requires innovative and comprehensive analysis of a provided data set with results presented at a special session of JSM. My entry was titled "Let's talk about the weather". 2018
- PhD Recipient of the Ellis R. Ott Scholarship for Applied Statistics and Quality sponsored by the ASQ Statistics Division. One scholarship is awarded to a Ph D student each year. 2018
- Exemplary Faculty Award at Brigham Young University Idaho 2013

Credentials

- Accredited Professional Statistician™ (PSTAT)** 2014-Present
American Statistical Association

Affiliations

- American Statistical Association 2005-Present
- Society of Industrial and Applied Mathematics 2016-2020
- Western North American Region of The International Biometric Society 2016-2020

Websites

- Github: <https://github.com/jillbo1000>
- Google Scholar: <https://scholar.google.com/citations?hl=en&user=ShrO3-AAAAAJ>
- LinkedIn: <https://www.linkedin.com/in/jill-lundell-00894256>
- Professional: www.jilllundell.com
-