

CHAD M. SCHAFER

Address: 5000 Forbes Avenue, Department of Statistics & Data Science, Pittsburgh, PA 15213

Phone: 412-268-3967

Email: cschafer at cmu.edu

URL: <http://www.stat.cmu.edu/~cschafer>

EDUCATION

Ph.D. in Statistics from University of California, Berkeley, 2004.

Master of Science in Statistics from University of Illinois, Urbana-Champaign, 1998.

Bachelor of Science from Western Michigan University, 1995.

POSITIONS

Professor, Department of Statistics & Data Science, Carnegie Mellon University, July 2020 to present

Associate Professor (with tenure), Department of Statistics, Carnegie Mellon University, September 2013 to June 2020

Associate Professor (without tenure), Department of Statistics, Carnegie Mellon University, September 2011 to September 2013

Assistant Professor, Department of Statistics, Carnegie Mellon University, September 2007 to September 2011.

Visiting Assistant Professor, Department of Statistics, Carnegie Mellon University, September 2004 to September 2007.

PUBLICATIONS

Bundy, K., et al. (25 authors). (2020) [“The Keck-FOBOS spectroscopic facility: conceptual design.”](#) Proceedings of *Ground-based and Airborne Instrumentation for Astronomy VIII*. Volume 11447.

Cisewski-Kehe, J., Weller, G., Schafer, C. (2019). [“A Preferential Attachment Model for the Stellar Initial Mass Function.”](#) *Electronic Journal of Statistics*. (13): 1580.

- Wang, K., Mao, Y., Zentner, A., van den Bosch, F., Lange, J., Schafer, C., Villareal, A., Hearin, A., Campbell, D. (2019). "[How to Optimally Constrain Galaxy Assembly Bias: Supplement Projected Correlation Functions with Count-in-cells Statistics.](#)" *Monthly Notices of the Royal Astronomical Society*. (488): 3541.
- Dalmasso, N., Dunn, R., LeRoy, B., Schafer C. (2019). "[A Flexible Pipeline for Prediction of Tropical Cyclone Paths.](#)" ICML Workshop (RESEARCH Track) "Climate Change: How can AI Help?."
- Eadie, G., et al. (15 authors). (2019) "[Realizing the potential of astrostatistics and astroinformatics.](#)" Submitted to the *Decadal Survey on Astronomy and Astrophysics*.
- Bauer, A. E., et al. (27 authors). (2019) "[Petabytes to Science.](#)" arXiv preprint arXiv:1905.05116.
- Smith, A.M., et al.. (20 authors). (2019) "[Astro2020 APC White Paper: Elevating the Role of Software as a Product of the Research Enterprise.](#)" arXiv preprint arXiv:1907.06981.
- Croft, R., Freeman, P., Schuster, T., Schafer, C. (2017). "[Prediction of galaxy ellipticities and reduction of shape noise in cosmic shear measurements.](#)" *Monthly Notices of the Royal Astronomical Society*. (469): 4422.
- Najita, J., et al. (51 authors). (2016). "[Maximizing Science in the Era of LSST: A Community-Based Study of Needed US Capabilities.](#)" arXiv preprint arXiv:1610.01661.
- Schafer, C. (2015). "[A framework for statistical inference in astrophysics.](#)" *Annual Review of Statistics and Its Application*. (2): 141.
- Izbicki, R. Lee, A., Schafer, C. (2014). "[High-dimensional density ratio estimation with extensions to approximate likelihood computation.](#)" *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics*. (33): 420.
- De Rubeis, S., et al. (90 authors). (2014) "[Synaptic, transcriptional and chromatin genes disrupted in autism.](#)" *Nature*. (515): 209.
- Weyant, A., Schafer, C., Wood-Vasey, W. (2013). "[Likelihood-free cosmological inference with type Ia supernovae: approximate Bayesian computation for a complete treatment of uncertainty.](#)" *The Astrophysical Journal*. (764): 116.
- Schafer, C., Campbell, N., Cai, G., Yu, F., Makarov, V., Yoon, S., Daly, M., Gibbs, R., Schellenberg, G., Devlin, B. (2013). "[Whole exome sequencing reveals minimal differences between cell line and whole blood derived DNA.](#)" *Genomics*. (102): 270.
- Neale, B., et al. (59 authors). (2012). "[Patterns and rates of exonic de novo mutations in autism spectrum disorders.](#)" *Nature*. (485): 242.
- Richards, J., Lee, A., Schafer, C., Freeman, P. (2012). "[Prototype selection for parameter estimation in complex models.](#)" *The Annals of Applied Statistics*. (6): 383.

- Schafer, C., Freeman, P. (2012). "[Likelihood-free inference in cosmology: Potential for the estimation of luminosity functions.](#)" *Statistical Challenges in Modern Astronomy V*.
- Freeman, P., Izbicki, R., Lee, A., Schafer, C., Slepcev, D., Newman, J., (2012). "[Detecting Galaxy Mergers at High Redshift.](#)" *Statistical Challenges in Modern Astronomy V*.
- Didier, J., Schafer, C., LeDuc, P. (2012). "Programmed Biologically Inspired Synthetic Templating of Multifunctional Nanoarchitectures for Small-Scale Reactions." *European Journal of Inorganic Chemistry*. (2012): 5405.
- Buchman, S., Lee, A., Schafer, C., (2011). "[High-dimensional density estimation via SCA: An example in the modelling of hurricane tracks.](#)" *Statistical Methodology*. (8): 18.
- Richards, J., Homrighausen, D., Freeman, P., Schafer, C., Poznanski, D. (2011). "[Semi-supervised learning for photometric supernova classification.](#)" *Monthly Notices of the Royal Astronomical Society*. (419): 1121.
- Richards, J., Freeman, P., Lee, A., Schafer, C. (2009). "[Exploiting low-dimensional structure in astronomical spectra.](#)" *The Astrophysical Journal*. (691): 32.
- Freeman, P., Newman, J., Lee, A., Richards, J., Schafer, C., (2009). "[Photometric redshift estimation using spectral connectivity analysis.](#)" *Monthly Notices of the Royal Astronomical Society*. (398): 2012.
- Richards, J., Freeman, P., Lee, A., Schafer, C. (2009). "[Accurate parameter estimation for star formation history in galaxies using SDSS spectra.](#)" *Monthly Notices of the Royal Astronomical Society*. (399): 1044.
- Schafer, C., Stark, P., (2009). "[Constructing confidence regions of optimal expected size.](#)" *Journal of the American Statistical Association*. (104): 1080.
- Schafer, C., Doksum, K. (2009). "[Selecting local models in multiple regression by maximizing power.](#)" *Metrika*. (69): 283.
- Freeman, P., Richards, J., Schafer, C., Lee, A. (2008). "[Astrostatistics: The final frontier.](#)" *Chance*. (21): 31.
- Schafer, C. (2007). "[A statistical method for estimating luminosity functions using truncated data.](#)" *The Astrophysical Journal*. (661): 703.
- Bryan, B., McMahan, H., Schafer, C., Schneider, J. (2007). "Efficiently computing minimax expected-size confidence regions." *Proceedings of the 24th international conference on Machine learning*. 97.
- Doksum, K., Schafer, C. (2006). "[Powerful choices: tuning parameter selection based on power.](#)" *Frontiers in Statistics*. 113.

- Schafer, C., Stark, P., Evans, S., Hansen, B. (2003). “Using what we know: Inference with physical constraints.” *Statistical Problems in Particle Physics, Astrophysics, and Cosmology*. 25.
- Jacob, R., Schafer, C., Foster, I., Tobis, M., Anderson, J. (2001). “Computational design and performance of the Fast Ocean Atmosphere Model, Version One.” *International Conference on Computational Science*. 175.
- Tobis, M., Schafer, C., Foster, I., Jacob, R., Anderson, J. (1997). “FOAM: Expanding the horizons of climate modeling.” *Proceedings of the 1997 ACM/IEEE conference on Supercomputing*.
- Hibbard, W., Anderson, J., Foster, I., Paul, B., Jacob, R., Schafer, C., Tyree, M. (1996). “Exploring coupled atmosphere-ocean models using Vis5D.” *The International Journal of Supercomputer Applications and High Performance Computing*. (10): 211.

PRESENTATIONS

- “Astrostatistics in the Era of LSST.” October 2020. *Fall Southeastern Sectional Meeting of AMS*.
- “Generating Realistic Galaxy Images.” July 2019. *Joint Statistical Meetings*. Denver, Colorado.
- “A Statistical Perspective on Deep Learning.” January 2019. *PhyStat-nu*. Geneva, Switzerland.
- “From Statistics to Data Science: What has Changed?” October 2018. *Corporate presentation to Morgan Stanley*. New York City.
- “Incorporating Spectral Features into Classifying Astronomical Light Curves.” April 2018. *Departmental Seminar*. Texas A&M University.
- “Maximizing Science Return from Large Surveys.” February 2018. *NOAO Community Needs for Science in the 2020s*. Tucson, Arizona.
- “Overview of Statistical Learning.” February 2018. *DM-Stat: Statistical Challenges in the Search for Dark Matter*. Banff, Canada.
- “The Potential of Deep Learning with Astronomical Data.” June 2017. *American Astronomical Society Meeting*. Austin, Texas.
- “Building Models for Emission Lines.” August 2016. *Joint Statistical Meetings*. Chicago, Illinois.
- “The Biggest of the Big Data: The Role of Statistics in Astronomy.” October 2015. *Departmental Seminar*. Western Michigan University. Kalamazoo, Michigan.
- “Big Science + Big Data = Big Opportunities: An Overview of Statistics in Astronomy.” May 2015. *Seminar*. Lawrence University. Appleton, Wisconsin.

- “Classifying Astronomical Sources Via Auxiliary Information.” September 2014. *Departmental Seminar*. University of Pennsylvania.
- “A Framework for Statistical Inference in Astronomy.” August 2014. *Introductory Overview Lecture at Joint Statistical Meetings*. Boston, Massachusetts.
- “Minimal Differences in Single Nucleotide Variation Calls between Blood and Cell Line Derived DNA from the same Individuals.” November 2012. *American Society for Human Genetics Annual Meeting*. San Francisco, California.
- “Constructing Exponential Family Approximations to Cosmological Models.” August 2012. *Joint Statistical Meetings*. San Diego, California.
- “Facing Heteroscedastic Measurement Error in Astronomical Surveys.” August 2011. *Joint Statistical Meetings*. Miami Beach, Florida.
- “Addressing the Challenges of Luminosity Function Estimation via Likelihood-Free Inference.” June 2011. *Statistical Challenges in Modern Astronomy V*. State College, Pennsylvania.
- “The Challenge and Potential of Likelihood-Free Inference in Cosmology.” March 2011. *Center for Time Domain Informatics Seminar Series*. Berkeley, California.
- “Facing the Supernova Challenge: Complex Theory and Complex Data.” April 2010. *New England Statistics Symposium*. Boston, Massachusetts.
- “Improved Astronomical Inferences via Nonparametric Density Estimation.” January 2010. *215th American Astronomical Society Meeting*. Washington, D.C.
- “Stochastic Models for High-Dimensional, Nonstandard Data.” October 2009. *Conference on Intelligent Data Understanding*. NASA Ames, California.
- “Testing Cosmological Theories: Methodology for the Inference Challenges.” October 2009. *Departmental Seminar*. University of California, Davis.
- “Issues and Directions in Parameter Estimation with Complex Models.” September 2009. *Statistical Frontiers of Astrophysics*. Institute for Physics and Mathematics of the Universe, Japan.
- “Improved Statistical Inference via Dimension Reduction.” July 2009. *COSMOSTATS09*. Ascona, Switzerland.
- “Semiparametric Bivariate Density Estimation with Irregularly Truncated Data.” July 2007. *Joint Statistical Meetings*. Salt Lake City, Utah.
- “Cosmological Inference via Measurements of the CMB: Nonparametric and Parametric Frequentist Approaches.” November 2005. *Probing the Distant Universe with Gravitational Waves*. Johnson City, Tennessee.

“Two Examples of Statistical Inference using Astronomical Data.” March 2005. *Departmental Seminar*. University of Wisconsin.

“Cosmological Parameter Estimation: A Frequentist Perspective.” July 2001. *Statistical Challenges in Modern Astronomy III*. State College, Pennsylvania.

PROFESSIONAL ACTIVITIES

Member of Steering Committee for *Master of Science in Computational Finance* program at CMU, August 2018 to present.

Co-Chair of *Statistics and Informatics Science Collaboration* for the *Large Synoptic Survey Telescope*, 2015 to present.

Co-Chair of *Statistical Challenges in Modern Astronomy VII*, June 2021.

Chair of the *Astrostatistics Interest Group* of the *American Statistical Association*, January 2019 to January 2020.

Instructor in *Pan-European Advanced Statistics School*, hosted at DESY in October 2019.

Director of *Summer Undergraduate Research Experience in Statistics* at CMU during Summers of 2015, 2016, 2017, and 2018.

Guest editor of special issue of *Chance* on the topic of Astrostatistics, September 2019.

Co-Chair of *Statistical Challenges in Modern Astronomy VI*, June 2016.

Organized session for AAAS Meeting in Austin, “A Universe of Discoveries: Progress in Astronomy, Statistics, and Machine Learning,” February 2018.

Moderator for Methodology Section of Statistics Area on arXiv, April 2007 to September 2018.

Joint Statistical Meetings Program Committee representative from the Nonparametrics Section for 2016.

President of the *Pittsburgh Chapter of the American Statistical Association*, August 2010 to August 2011.

Director of Graduate Studies for Department of Statistics at CMU, September 2011 to September 2016.

Member of Faculty Senate at CMU, September 2013 to September 2015.

Associate Editor for *Journal of the American Statistical Association - Theory and Methods*, 2014 to 2016.

Associate Editor for *Electronic Journal of Statistics*, 2007 to 2012.

Instructor for *Penn State Summer School on Astrostatistics*, June 2010, 2012, 2013, 2014, 2015, 2017, 2018, 2019, and 2021.

Member of Scientific Organizing Committee for *Statistical Challenges in Modern Astronomy V*, June 2011.

Co-taught “Tutorial on Nonparametric Inference, with R” for astronomers as a part of *SAMSI Opening Workshop for Astrostatistics Program*, January 2006.

Organized special session on astrostatistics for 2006 IMS Meetings.

FUNDED GRANTS

“Complexity to Clarity: Nonparametric Procedures that Exploit Structured Data and Models.” NSF Award #1521786. PI: A. Lee. Co-PIs: C. Schafer, S. Ho. \$450,000.00 for 9/2015 to 8/2019.

“Machine Learning Algorithms for Matching Theories, Simulations, and Observations in Cosmology.” DOE Award #DE-SC0011114. PI: B. Poczos. Co-PIs: S. Ho, J. Schneider, C. Genovese, R. Mandelbaum, C. Schafer, H. Trac. \$1,649,402.00 for 1/2014 to 12/2018.

“Statistical Challenges in Modern Astronomy VI.” Award Number: 1602592; PI: C. Schafer. Co-PI: S. Ho. \$20,000 for 6/2016.

“Nonparametric Inference for Complex Physical Models.” PI: C. Schafer, Co-PIs: A. Lee, C. Genovese, L. Wasserman, W. Wood-Vasey. NSF Proposal #1106956. \$100,000 for 8/2011 to 8/2013.

“Stochastic Models for High-Dimensional, Nonstandard Data.” PI: C. Schafer, Co-PIs: A. Lee and P. Freeman. NASA Proposal #08-AISR08-0112. \$250,000 for 9/2009 to 9/2011.

“MSPA - AST: Sparse Representation and Efficient Inference for Astronomical Spectra.” NSF Award #0707059. PI: A. Lee, Co-PIs: P. Freeman and C. Schafer. NSF Proposal #0707059. \$240,000 for 9/2007 to 9/2010.

PH.D. STUDENTS

Joseph Richards. May 2010. “Fast and Accurate Estimation for Astrophysical Problems in Large Databases.” Vice President of Data and Analytics at GE Digital.

Susan Buchman. May 2011. “High-Dimensional Adaptive Basis Density Estimation.” Special Faculty, Department of Statistics & Data Science, Carnegie Mellon University.

Beatriz Estefania Etchegaray. May 2014. "Classification Via Auxiliary Information." Head of Data Science, IBM Chief Analytics Office.

Michael Vespe. May 2016. "Constructing Approximately Sufficient ABC Summary Statistics." Capital One.