Ann B. Lee

Department of Statistics ⋄ Carnegie Mellon University ⋄ 5000 Forbes Avenue ⋄ Pittsburgh, PA 15213; annlee@stat.cmu.edu
Citizenship: U.S.

Education

Ph.D. in Physics, 2002, Brown University
Advisors: Professors David Mumford (Applied Mathematics) and Leon N Cooper (Physics)
Thesis: "Statistics, Models and Learning in BCM Theory of a Natural Visual Environment"

M.S./B.S. in Engineering Physics, 1995, Chalmers University of Technology, Sweden

Employment

Associate Professor, Department of Statistics, Carnegie Mellon University, 2010–present.
Full voting member, Department of Machine Learning, CMU, 2008–present.
Estella Loomis McCandless Chair, CMU, 2011-2013.

Assistant Professor, Department of Statistics, Carnegie Mellon University, 2005–2010.

J.W. Gibbs Assistant Professor, Department of Mathematics, Yale University, 2004–2005.

J.W. Gibbs Instructor, Department of Mathematics, Yale University, 2002–2004.

Visiting Research Associate, Division of Applied Mathematics, Brown University, 2001–2002.

Journal Publications

Richards, J. W., Lee, A. B., Schafer, C. M., and Freeman, P. E., "Prototype Selection for Parameter Estimation in Complex Models", arXiv:1105.6344. To appear in *Annals of Applied Statistics*, 2012.

Wang, W., Ozolek, J.A., Slepcev, D., Lee, A. B., Chen, C., and Rohde, G. K. "An optimal transportation approach for nuclear structure-based pathology". *IEEE Trans. Med. Imag.*, 30(3):621-631, 2011.

Buchman, S. M., Lee, A. B., and Schafer, C. M., "High-Dimensional Density Estimation via SCA: An Example in the Modelling of Hurricane Tracks", *Statistical Methodology*, 8(1):18-30, 2011.

Lee, A. B. and Wasserman, L., "Spectral Connectivity Analysis", *Journal of the American Statistical Association*, 105(491): 1241-1255, 2010.

Lee, A. B., Luca, D., and Roeder, K., "A Spectral Graph Approach to Discovering Genetic Ancestry". *Annals of Applied Statistics*, 4(1): 179-202, 2010.

Lee, A. B., Luca, D., Klei, L., Devlin, B., and Roeder, K., "Discovering Genetic Ancestry Using Spectral Graph Theory". *Genetic Epidemiology*, 34(1):51-59, 2010.

Richards, J. W., Freeman, P. E., Lee, A. B., and Schafer, C. M., "Accurate parameter estimation for star formation history in galaxies using SDSS spectra", *Monthly Notices of the Royal Astronomical Society*, 399: 1044-1057, 2009.

Freeman, P. E., Newman, J. A., Lee, A. B., Richards, J. W., and Schafer, C. M., "Photometric Redshift Estimation Using SCA", *Monthly Notices of the Royal Astronomical Society*, 398: 2012-2021, 2009

Richards, J. W., Freeman, P. E., Lee, A. B., and Schafer, C. M., "Exploiting Low-Dimensional Structure in Astronomical Spectra", *Astrophysical Journal* 691:32-42, 2009

Freeman, P. E., Richards, J. W., Schafer, C. M., and Lee, A. B., "Astrostatistics: The Final Frontier", *Chance*, 21(3):31-35, 2008.

Lee, A. B., Nadler, B., and Wasserman, L., "Treelets – An Adaptive Multi-Scale Basis for Sparse Unordered Data", *Annals of Applied Statistics*, 2(2):435-471, 2008. **Discussion paper**.

Lee, A. B., Nadler, B., and Wasserman, L., "Rejoinder of: Treelets", *Annals of Applied Statistics*, 2(2):494-500, 2008.

Luca, D., Ringquist, S., Klei, L., Lee, A.B., Gieger, C., Wichmann, H.-E., Schreiber, S., Krawczak, M., Liu, Y., Styche, A., Devlin, B., Roeder, K., and Trucco, M., "On the Use of General Control Samples for Genome-Wide Association Studies: Genetic Matching Highlights Causal Variants", *The American Journal of Human Genetics*, 82:1-11, 2008.

Lafon, S. and Lee, A. B., "Diffusion Maps and Coarse-Graining: A Unified Framework for Dimensionality Reduction, Graph Partitioning and Data Set Parameterization", *IEEE Trans. Pattern Analysis and Machine Intelligence*, 28(9): 1393-1403, 2006.

Coifman, R.R., Lafon, S., Lee, A. B., Maggioni, M., Nadler, B., Warner, F., and Zucker, S., "Geometric Diffusions as a Tool for Harmonic Analysis and Structure Definition of Data: Diffusion Maps", *Proc. Natl. Acad. Sci.* 102(21): 7426-7431, 2005.

Coifman, R.R., Lafon, S., Lee, A. B., Maggioni, M., Nadler, B., Warner, F., and Zucker, S., "Geometric Diffusions as a Tool for Harmonic Analysis and Structure Definition of Data: Multiscale Methods", *Proc. Natl. Acad. Sci.* 102(21): 7432-7437, 2005.

Lee, A. B., Pedersen, K. S., and Mumford D. , "The Nonlinear Statistics of High-Contrast Patches in Natural Images", *International Journal of Computer Vision* 54 (1-2): 83-103, 2003.

Srivastava, A., Lee, A. B., Simoncelli, and E. P., Zhu, S.-C., "On Advances in Statistical Modeling of Natural Images", *Journal of Mathematical Imaging and Vision* 18: 17-33, 2003.

Lee, A. B., Mumford, D., and Huang, J., "Occlusion Models for Natural Images", *International Journal of Computer Vision* 41(1/2): 35-59, 2001.

Lee, A. B., Blais, B. S., Shouval, H., and Cooper, L. N., "Statistics of Lateral Geniculate Nucleus (LGN) Activity Determine the Segregation of ON/OFF Subfields for Simple Cells in Visual Cortex", *Proc. Natl. Acad. Sci.* 97(23): 12875-12879, 2000.

Wahnström, G., Lee, A. B., and Strömquist, J., "Motion of 'hot' oxygen adatoms on corrugated metal surfaces", *J. Chem. Phys.* 105: 326-336, 1996.

Conference Proceedings

Lee, A. B., and Freeman, P. E., "Exploiting Non-Linear Structure in Astronomical Data for Improved Statistical Inference", In *Statistical Challenges in Modern Astronomy (SCMA) V*, Penn State University, June 2011 (invited talk); arXiv:1111.0911.

Lee, A.B., "Commentary on 'Data Compression Methods in Astrophysics' by Raul Jimenez", In *Statistical Challenges in Modern Astronomy (SCMA) V*, Penn State University, June 2011 (invited commentary).

Rohde, G. K., Wang, W., Slepcev, D., Lee, A. B., Chen, C., and Ozolek, J.A. "Detecting and classifying cancers from image data using optimal transportation". In *26th Southern Biomedical Engineering Conference*, University of Maryland, 2010.

Lee, A. B., and Nadler, B. "Treelets – A Tool for Dimensionality Reduction and Multi-Scale Analysis of Unstructured Data". In *Proc. of the 11th International Conf. on Artificial Intelligence and Statistics* (AISTATS*2007), San Juan, Puerto Rico, March 2007 (poster).

Pedersen, K. S., and Lee, A. B. "Towards a Full Probability Model of Edges in Natural Images". In *Proc. of European Conference of Computer Vision* (ECCV*2002), Copenhagen, May 2002 (poster).

Lee, A. B., Pedersen, K. S., and Mumford D. "The Complex Statistics of High-Contrast Patches in Natural Images". In *Proc. of IEEE Workshop on Statistical and Computational Theories of Vision* (with ICCV*2001), Vancouver, CA, July 2001 (talk).

Huang, J., Lee, A. B., and Mumford, D. "Statistics of Range Images". In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (CVPR*2000), Hilton Head Island, South Carolina, June 2000 (talk).

Lee, A. B., Blais, B. S., Shouval, H., and Cooper, L. N. "Statistics of LGN Activity Determine the Segregation of ON/OFF Subfields for Simple Cells in Cortex". In *Proc. of the Eighth CNS Conference* (CNS*99), Pittsburgh, PA, July 1999 (talk).

Lee, A. B., and Mumford D. "An Occlusion Model Generating Scale-Invariant Images", In *Proc. of IEEE Workshop on Statistical and Computational Theories of Vision* (with CVPR*99), Fort Collins, Co, June 1999 (talk).

Principal Grants and Awards

Active grants

NSF DMS-1106956 (**co-PI**, with PI C. Schafer and co-PIs C. Genovese, L. Wasserman, W. Wood-Vasey): "Nonparametric Inference for Complex Physical Models", 8/15/2011-7/31/2013.

Past grants and other awards

NASA NNX09AK59G (**co-PI**, with PI C. Schafer and co-PI P. Freeman): "Stochastic Models for High-Dimensional, Nonstandard Data", 9/1/2009-8/31/2011.

ONR #00424143 (**PI**): "Treelets and Multi-Scale Tools for Statistical Data Analysis", 3/1/2008-2/28/2011.

NSF DMS-0707059 (**PI**, with co-PI C. Schafer): "Sparse Representation and Efficient Inference for Astronomical Spectra", 9/15/2007 - 9/14/2010.

NSF CCF-0625879 (**co-PI**, with PI J. Lafferty and co-PI L. Wasserman): "Nonparametric Learning in High Dimensions", 9/1/2006-8/31/2009.

Estella Loomis McCandless Chair, 2011.
Galkin Foundation Fellowship, 2000.
Burroughs Wellcome Fellowship, 1999.
Professor R. Bruce Lindsay Fellowship, 1997.
Brown University Fellowship, 1995.

Course Instruction

Statistical Practice (STAT 36-726), Spring 2012.
Engineering Statistics and Quality Control (STAT 36-220), Fall 2010, 2011.
Machine Learning Journal Club (ML 10-915), Fall 2009.
Probability and Statistics for Business Applications (STAT 36-207), CMU, Fall 2009.
Probability and Statistics II (STAT 36-626), CMU, Spring 2006, 2007, 2008, 2010.
Probability and Statistics I (STAT 36-625), CMU, Fall 2005, 2006, 2007.
Applied Mathematics and Engineering I (AMTH 251), Yale University, Fall 2004, 2003.
Introduction to Calculus in Several Variables (MATH 118), Yale University, Spring 2004.
Pattern Theory and its Applications, 12th Jyväskylä Ph.D. Summer School, Aug 2002, Finland.

Advising of Ph.D. Students (Thesis Work)

Current
Rafael Izbicki, Department of Statistics, CMU
Di Liu, Department of Statistics, CMU
Andrew Crossett, Department of Statistics, CMU (co-advised with K. Roeder)

Former
Susan Buchman
– Ph.D. March 2011, Department of Statistics, CMU (co-advised with C. Schafer)
– Thesis title: "High-Dimensional Adaptive Basis Density Estimation"
Joseph W. Richards
– Ph.D. July 2010, Department of Statistics, CMU (co-advised with C. Schafer)
– Thesis title: "'Fast and Accurate Estimation for Astrophysical Problems in Large Databases"
Diana Luca
– Ph.D. Sept 2008, Department of Statistics, CMU (co-advised with K. Roeder)
– Thesis title: "Genetic Matching by Ancestry in Genome-Wide Association Studies"

Professional Activities

Associate Editor, Statistics and Computing, Springer, 2010-present.
ML Education Co-Director, Department of Machine Learning, CMU, 2009-present.
Graduate Admissions Co-Chair, Department of Machine Learning, CMU, 2009-present.
Faculty Senator, CMU, 2007-2009.
Graduate Admissions Co-Chair, Department of Statistics, CMU, 2005-2008.

Program and review committees
NIPS*2008, ICML*2010
European Conference on Computer Vision (ECCV*2004).
Intl. Workshop on Generative Model-Based Vision (2002, 2004).

Journal Reviewing
Annals of Applied Statistics
Electronic Journal of Statistics
IEEE Pattern Analysis and Machine Intelligence
Journal of the American Statistical Association
Journal of Classification
Journal of Computational and Graphical Statistics
Journal of Mathematical Imaging and Vision
Proceedings of the Royal Society A
SIAM Journal of Imaging Sciences
Signal Processing D
Statistica Sinica
Statistical Analysis and Data Mining
Transactions on Knowledge and Data Engineering

Memberships
ASA, IMS