

CURRICULUM VITA

DONALD GEMAN
Professor, Department of
Applied Mathematics and Statistics
The Johns Hopkins University
302A Clark Hall
Baltimore, Maryland 21218
geman@jhu.edu
410-516-7678

Affiliations

Institute for Computational Medicine, JHU
Center for Imaging Science, JHU
Mathematical Institute for Data Science (MINDS), JHU

Research Areas

Computational Vision

Scene interpretation
Turing tests for vision

Computational Medicine

Tumorigenesis
Single-cell transcriptomics
Predicting disease phenotypes

Machine Learning

Small-sample learning
Mechanistic classification

Education

Columbia University, New York, NY, 1961-1963
University of Illinois, Urbana, IL, 1963-1965
B.A. in English Literature
Northwestern University, Evanston, IL, 1966-1970
Ph.D. in Mathematics

Employment

Johns Hopkins University, Department of Applied Mathematics and Statistics
Professor, 2001-present
University of Massachusetts, Department of Mathematics and Statistics
Assistant and Associate Professor, 1970-1980
Professor and Distinguished Professor, 1981-2001

Visiting Positions

Department of Statistics, University of North Carolina, 1976-1977 ; Division of Applied Mathematics, Brown University, 1991-1992 ; ETH, Zurich, 1993 ; INRIA, Paris, France, periodic visits, 1990-present ; Department

of Applied Mathematics, Ecole Polytechnique, Palaiseau, France, Fall, 1997-99 ; Department of Statistics, University of Chicago, Spring, 2000 ; Centre de Mathematiques et Leurs Applications, ENS-Cachan, France, Spring, 2001-2013.

Honors

Member, National Academy of Sciences

Fellow, Society for Industrial and Applied Mathematics (SIAM)

Fellow, Institute of Mathematical Statistics (IMS)

Plenary or Keynote Speaker: ICIP, Lausanne, Switzerland, 1996; Annual NESS Meeting, Univ.of Connecticut, 1999; Annual French Statistical Society Meeting, Nantes, France, 2001; Biannual EMMCVPR, Sophia Antipolis, France, 2001; MAA Meeting, Baltimore, MD, 2003; ACIVS, Brussels, Belgium, 2004; Snowbird Learning Workshop, Snowbird, UT, 2006; International Symposium on Information Theory (ISIT06), Seattle, WA, 2006; Norwegian Society for Image Processing and Pattern Recognition, Oslo, Norway, 2006; Multimedia Image Retrieval (MIR'06), Santa Barbara, 2006; MCBIOS Annual Conference , College Station, 2011; German Pattern Recognition Society, Frankfurt, 2011; Medallion Lecture, 2012 Joint Statistical Meetings, San Diego, 2012.

Professional Societies

Institute of Mathematical Statistics, Institute of Electrical and Electronics Engineers, American Mathematical Society , Society for Industrial and Applied Mathematics

Doctoral Students

Carmen Acuna, *Parameter estimation for stochastic texture models*, Univ. of Mass., 1988; **Chengda Yang**, *Stochastic methods for image restoration*, Univ. of Mass., 1991; **Keith Hartt**, *Bayesian estimation of surface information from radar images*, Univ. of Mass., 1993; **Bruno Jedynak**, *Stochastic models and deterministic methods for finding roads in remotely-sensed images*, Univ. de Paris - Sud, 1995; **Decheng Wang**, *Stochastic modeling of magnetic resonance images with applications to tissue classification*, Univ. of Mass., 1996; **Kenneth Wilder**, *Decision tree algorithms for handwritten digit recognition*, Univ. of Mass., 1998; **Chunming Li**, *Classification by active testing with applications to imaging and change detection*, Univ. of Mass., 1998; **Francois Fleuret**, *Hierarchical face detection by statistical learning*, Univ. de Paris VI, 2000; **Alexey Koloydenko**, Univ. of Mass., *Modeling natural microimage statistics*, Univ. of Mass., 2000; **Franck Jung**, *Reconnaissance d'objets par focalisation et detection de changements*, Ecole Polytechnique, 2001; **Hichem Sahbi**, *Support vector machines for hierarchical face detection*, Universite de Versailles, 2003; **Christian d'Avignon**, *Applying machine learning to biomedical data: the small-sample and interpretability dilemmas*, Johns Hopkins University, 2004; **Xiaodong Fan**, *Learning a hierarchy of classifiers for multi-class shape detection*, Johns Hopkins University, 2006; **Sachin Gangaputra**, *Invariant coarse-to-fine object detection and tracking*, Johns Hopkins University, 2006; **Mary Lin**, *Rank-based methods for statistical analysis of gene expression microarray data*, Johns Hopkins University, 2008; **Erdem Yoruk**, *Learning graphical models with limited observations of high-dimensional data*, Johns Hopkins University, 2011; **Francisco Sanchez**, *Small sample learning of multivariate distributions using probabilistic graphical models*, Johns Hopkins University, 2012; **David Simcha**, *Statistical learning applied to transcriptional regulation in small N, large D domains*, Johns Hopkins University, 2012; **Bahman Afsari**, *Modeling cancer phenotypes with order statistics of transcript data*, Johns Hopkins University, 2013; **Asma Rejeb**, *Fine-grained object categorization: plant species identification*, ParisTech, 2014; **Ehsan Jahangiri**, Johns Hopkins University, 2018; **Kamel Lahouel**,

Coarse-to-fine multiple testing strategies, Johns Hopkins University, 2018; **Qian Ke**, Johns Hopkins University, in progress; **Yiran Xu**, Johns Hopkins University, in progress; **Lanlan Ji**, Johns Hopkins University, in progress.

Patents

K. Manbeck, C. Yang, D. Geman, and S. Geman. Cadence Editing. US 6,542,199, 2003; K. Manbeck, C. Yang, D. Geman, and S. Geman. Video Field Labeling. US 6,624,844, 2003; C. Yang, K. Manbeck, S. Geman, D. Geman. Format Conversion. US 7,064,792, 2006; K. Manbeck, J. Cassidy, S. Geman, D. Geman, and D. McClure. High Resolution Color Conforming. US 7,113,223, 2006; K. Manbeck. D. Geman, S. Geman, and M. Orton. Automated Color Control in Film-to-Digital Transfer. US 7,068,838, 2006.

Publications

Essay: D. Geman and S. Geman (2016), “Opinion: Science in the age of selfies,” *Proc. Nat’l. Acad. Sci.*, 113, 9384-9387.

Research in Computational Biology:

- Baloni, P., Dinalankara, W., Earls, J. C., Knijnenburg, T. A., Geman, D., Marchionni, L., & Price, N. D. (2021). “Identifying personalized metabolic signatures in breast cancer,” *Metabolites*, 11(1), 20.
- Lahouel, K., Younes, L., Danilova, L., Giardiello, F. M., Hruban, R. H., Groopman, J., Kinzler, K.W., Vogelstein, B., Geman, D. and Tomasetti, C. (2020). “Revisiting the tumorigenesis timeline with a data-driven generative model”. *Proceedings of the National Academy of Sciences*, 117(2), 857-864.
- Ke, Q., Dinalankara, W., Younes, L., Geman, D., & Marchionni, L. (2020). “Efficient Representations of Tumor Diversity with Paired DNA-RNA Aberrations”. bioRxiv.
- Afsari, B., Guo, T., Considine, M., Florea, L., Kagohara, L.T., Stein-O’Brien, G.L., Kelley, D., Flam, E., Zambo, K.D., Ha, P.K., Geman, D. and Fertig, E.J. (2018). “Splice Expression Variation Analysis (SEVA) for inter-tumor heterogeneity of gene isoform usage in cancer”. *Bioinformatics*, 34(11), pp.1859-186.
- Dinalankara, W., Ke, Q., Xu, Y., Ji, L., Pagane, N., Lien, A., Matam, T., Fertig, E.J., Price, N.D., Younes, L., Marchionni, L. and Geman, D. (2018). “Digitizing omics profiles by divergence from a baseline”. *Proceedings of the National Academy of Sciences*, 115(18), pp.4545-4552.
- Slama, P., Hoopmann, M. R., Moritz, R. L., & Geman, D. (2018). “Robust determination of differential abundance in shotgun proteomics using nonparametric statistics.” *Molecular omics*, 14(6), 424-436.
- Ament, S. A., Pearl, J. R., Grindeland, A., St Claire, J., Earls, J. C., Kovalenko, M., ... & Hood, L. E. (2017). “High resolution time-course mapping of early transcriptomic, molecular and cellular phenotypes in Huntington’s disease CAG knock-in mice across multiple genetic backgrounds”. *Human molecular genetics*, 26(5), 913-922.
- Lahouel, K., Geman, D., & Younes, L. (2015). “Coarse-to-fine multiple testing strategies”. arXiv preprint arXiv:1509.03737.
- Chang, L. B., & Geman, D. (2015). “Tracking cross-validated estimates of prediction error as studies accumulate”. *Journal of the American Statistical Association*, 110(511), 1239-1247.
- Geman, D., Ochs, M., Price, N. D., Tomasetti, C., & Younes, L. (2015). “An argument for mechanism-based statistical inference in cancer”. *Human genetics*, 134(5), 479-495.
- Ma, S., J. Sung, A. Magis, Y. Wang, D. Geman and N.D. Price (2014), “Measuring the effect of inter-study variability on estimating prediction error,” *PLOS ONE*, 9.
- Afsari, B., E. Fertig, D. Geman and L. Marchionni (2014), “switchbox: An R package for k-Top Scoring Pairs (kTSP) classifier development,” *Bioinformatics*, September.

- Afsari, B., D. Geman and E. Fertig (2014), "Learning dysregulated pathways in cancer from differential variability analysis," *Cancer Informatics*, 13, 62-67.
- Afsari, B., E. Fertig, L. Younes, D. Geman and L. Marchionni (2014), "Hardwiring mechanism into predicting cancer phenotypes by computational learning," *Cancer Research*, 74, 534.
- Afsari, B., U. Braga-Neto and D. Geman (2014), "Rank discriminants for predicting phenotypes from RNA expression," *Annals of Applied Statistics*, 8, 1469-1491.
- Simcha, D., L. Younes, M. Aryee and D. Geman (2013), "Identification of direction in gene networks from expression and methylation," *BMC Systems Biology*, 7:118
- Marchionni, L., B. Afsari, D. Geman and J.T. Leek (2013), "A simple and reproducible breast cancer prognostic test," *BMC Genomics* 14:336.
- Sung J., P-J Kim, C. Funk, S. Ma, A. Magis, Y. Wang, L. Hood, D. Geman and N.D. Price (2013), "Multi-study integration of brain cancer transcriptomes reveals organ-level diagnostic signatures," *PLOS Computational Biology*, 9, e1003148.
- Simcha, D., N.D. Price and D. Geman (2012), "The limits of de novo DNA motif discovery," *PLOS ONE*, 7(11).
- Winslow, R.L., N. Trayanova, D. Geman and M.I. Miller (2012), "The emerging discipline of computational medicine," *Science Translational Medicine*, 4, 31 October 2012.
- Yoruk, E., M. Ochs, D. Geman and L. Younes (2011), "A comprehensive statistical model for cell signaling," *IEEE Trans. Computational Biology and Bioinformatics*, 592-606.
- Slama, P. and D. Geman (2010), "Identification of family-determining residues in PHD fingers," *Nucleic Acids Research*, 1-14.
- Leek, J.T., R. Scharpf, H. Corrada Brav, D. Simcha, B. Langmead, E.W. Johnson, D. Geman, K. Baggerly and R.A. Irizarry (2010), "Tackling the widespread and critical impact of batch effects in high-throughput data," *Nature Reviews Genetics*, 11, 733-739.
- Lin, X., B. Afsari, L. Marchionni, G. Parmigiani, L. Cope, D. Naiman and D. Geman (2009), "The ordering of expression among a few genes can provide simple cancer biomarkers and signal BRCA1 mutations," *BMC Bioinformatics*, 10:256.
- Eddy, J.A., N.D. Price, L. Hood and D. Geman (2009), "Identifying tightly regulated and variably expressed networks by differential rank conservation (DIRAC)," *PLoS Computational Biology*, 6.
- Eddy, J.A., J. Sung, D. Geman and N.D. Price (2009), "Relative expression analysis for molecular diagnosis and prognosis," *Technology in Cancer Research and Treatment*, 9, 149-159.
- Edelman, L.B., G. Goia, D. Geman, W. Zhang and N.D. Price (2009), "Two-transcript gene expression classifiers in the diagnosis and prognosis of human diseases," *BMC Genomics*, 10:583.
- Geman, D., B. Afsari, A.C. Tan and D. Naiman (2008), "Microarray classification from several two-gene expression comparisons," *Proceedings ICMLA 2008*, (Winning entry, ICMLA Microarray Classification Algorithm Competition).
- Xu, L., A.C. Tan, R.L. Winslow and D. Geman (2008), "Merging microarray data from separate breast cancer studies provides a robust prognostic signature," *BMC Bioinformatics* 9:125.
- Xu, L., D. Geman and R. Winslow (2007), "Large-scale integration of cancer microarray data identifies a robust common cancer signature," *BMC Bioinformatics* 8:275.
- Anderson, T. J., I. Tchernyshyov, R. Diez, R.N. Cole, D. Geman, C. V. Dang and R. L. Winslow (2007). "Discovering robust protein biomarkers for disease from relative expression reversals in 2D DIGE data," *Proteomics* 7:1197-1207.
- Tan, A. C., D. Q. Naiman, L. Xu, R. L. Winslow and D. Geman (2005). "Simple decision rules for classifying human cancers from gene expression profiles." *Bioinformatics* 21(20): 3896-3904.
- Xu, L., A. C. Tan, D. Q. Naiman, D. Geman and R. L. Winslow (2005). "Robust prostate cancer marker genes emerge from direct integration of inter-study microarray data." *Bioinformatics* 21(20):3905-3911.
- Geman, D., C. D'Avignon, D. Q. Naiman and R. L. Winslow (2004). "Classifying gene expression profiles from pairwise mRNA comparisons." *Stat. Appl. Genet. Mol. Biol.* 3(1): Article 19.

- D'Avignon, C. and D. Geman (2003). "Tree-structured neural decoding." *Journal of Machine Learning Research* 4: 743-754.

Research in Machine Learning:

- Chang, L-B and D. Geman (2015), "Tracking cross-validated estimates of prediction error as studies accumulate," *Journal of the American Statistical Association*, 110, 1239-1247.
- Lahouel, K., D. Geman and L. Younes (2015), "Coarse-to-fine multiple testing strategies," arXiv:1509.03737.
- Geman, D., H. Geman and N.N. Taleb (2015), "Tail risk constraints and maximum entropy," *Entropy*, 17, 3724-3737.
- Sanchez-Vega, F., J. Eisner, L. Younes and D. Geman (2013), "Learning multivariate distributions by competitive assembly of marginals," *IEEE Trans. Pattern Analysis and Machine Intell*, 35, 398-410.
- Fleuret, F., T. Li, C. Dubout, E.K. Wampler, S. Yantis and D. Geman (2011), "Comparing machines and humans on a visual categorization test," *Proc. Nat. Acad. Sci.*, 108: 17621-17625.
- Sahbi, H. and D. Geman (2006). "A hierarchy of support vector machines for pattern detection." *Journal of Machine Learning Research* 7: 2087-2123.
- Blanchard, G. and D. Geman (2005). "Sequential testing designs for pattern recognition." *Annals of Statistics* 33(3): 1155-1202.
- Geman, D. and B. Jedynek (2001). "Model-based classification trees." *IEEE Trans. Information Theory* 47(3): 1075-1082.

Research in Computer Vision:

- Geman, D., S. Geman, N. Hallonquist and L. Younes (2015), "Visual Turing test for computer vision systems," *Proc. Nat. Acad. Sci.*, 112, 3618-3623.
- Rejeb-Sfar, A., N. Boujemaa and D. Geman (2015), "Confidence sets for fine-grained classification and plant identification," *International Journal of Computer Vision*, 111, 255-275.
- Rejeb-Sfar, A, N. Boujemaa and D. Geman (2013), "Vantage feature frames for fine-grained categorization," *Proceedings Computer Vision and Pattern Recognition (CVPR 2013)*.
- Ferecatu, M. and D. Geman (2009), "A statistical framework for image category search from a mental picture," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 31(6): 1087-1101.
- Fleuret, F. and D. Geman (2008), "Stationary features and cat detection," *Journal of Machine Learning Research*, 9:2547-2578.
- Gangaputra, S. and D. Geman (2006). "The trace model for object detection and tracking." Toward Category-Level Object Recognition, *Lecture Notes in Computer Science*, 4170. J. Ponce et al. Berlin, Springer-Verlag: 401-420.
- Amit, Y., D. Geman and X. Fan (2004). "A coarse-to-fine strategy for multi-class shape detection." *IEEE Trans. Pattern Analysis and Machine Intelligence* 26(12): 1606-1621.
- Geman, D. (2003). "Coarse-to-fine classification and scene labeling". Nonlinear Estimation and Classification. *Lecture Notes in Statistics*, 171. D. D. Denison, M. Hansen, C. C. Holmes, B. Mallick and B. Yu. New York, Springer-Verlag: 31-48.
- Fleuret, F. and D. Geman (2001). "Coarse-to-fine face detection." *International Journal of Computer Vision* 41(1-2): 85-107.
- Amit, Y. and D. Geman (1999). "A computational model for visual selection." *Neural Computation* 11: 1691-1715.
- Amit, Y. and D. Geman (1997). "Shape quantization and recognition with randomized trees," *Neural Computation* 9:1545-1588.
- Amit, Y., D. Geman and K. Wilder (1997). "Joint induction of shape features and tree classifiers," *IEEE*

Trans. Pattern Analysis and Machine Intelligence 19(11): 1300-1306.

- Geman, D. and B. Jedynak (1996). "An active testing model for tracking roads from satellite images." *IEEE Trans. Pattern Analysis and Machine Intelligence*. 18(1): 1-14.
- Geman, D. and C. G. Yang (1995). "Nonlinear image recovery with half-quadratic regularization." *IEEE Trans. Image Processing* 4(7): 932-946.
- Geman, D. and G. Reynolds (1992). "Constrained restoration and the recovery of discontinuities." *IEEE Trans. Pattern Analysis and Machine Intelligence* 14(3): 367-383.
- Geman, S., D. E. McClure and D. Geman (1992). "A nonlinear filter for film restoration and other problems in image processing." *Computer Vision, Graphics, and Image Processing* 54(4): 281-289.
- Geman, D., S. Geman, C. Graffigne and P. Dong (1990). "Boundary detection by constrained optimization." *IEEE Trans. Pattern Analysis and Machine Intelligence* 12(7): 609-628.
- Geman, D. (1990). "Random Fields and Inverse Problems in Imaging." *Lecture Notes in Mathematics*, Springer-Verlag. 1427: 113-193.
- Geman, D., S. Geman and C. Graffigne (1987). "Locating object and texture boundaries." *Pattern Recognition Theory and Applications*. P. Devijver and J. Kittler, Springer-Verlag.
- Geman, D. (1987). "A stochastic model for boundary detection." *Image and Vision Computing* 5: 61-65.
- Geman, S. and D. Geman (1984). "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images." *IEEE Trans. Pattern Analysis and Machine Intelligence* 6: 721-741.

Research in Stochastic Processes:

- Geman, D., J. Horowitz and J. Rosen (1984). "A local time analysis of the intersections of Brownian paths in the plane." *Annals of Probability* 12: 86-107.
- Geman, D. and J. Horowitz (1981). "Smooth perturbations of a function with a smooth local time." *Trans. Amer. Math. Soc.* 267: 517-530.
- Geman, D. and J. Horowitz (1980). "Occupation densities." *Annals of Probability* 8: 1-67.
- Geman, D. (1979). "Dispersion points for linear sets and approximate moduli for some stochastic processes." *Trans. Amer. Math. Soc.* 253: 257-272.
- Geman, D. (1977). "On the approximate local growth of multi-dimensional random fields." *Z. Wahrscheinlichkeitstheorie verw. Geb.* 38: 237-251.
- Geman, D. (1976). "A note on the continuity of local times." *Proc. Amer. Math. Soc.* 57: 321-326.
- Geman, D., J. Horowitz and J. Zinn (1976). "Recurrence of stationary sequences." *Annals of Probability* 4: 372-381.
- Geman, D. and J. Horowitz (1975). "Polar sets and Palm measures in the theory of flows." *Trans. Amer. Math. Soc.* 208: 141-159.
- Geman, D. and J. Horowitz (1975). "Random shifts which preserve measure." *Proc. Amer. Math. Soc.* 49: 143-150.
- Geman, D. and J. Horowitz (1973). "Remarks on Palm measures," *Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques*, 9, 215-232.
- Geman, D. (1973). "A note on the distribution of hitting times." *Annals of Probability* 1: 854-856.
- Geman, D. and J. Horowitz (1973). "Occupation times for smooth stationary processes." *Annals of Probability* 1: 131-137.
- Geman, D. (1972). "On the variance of the number of zeros of a stationary Gaussian process." *Annals Math. Stat.* 43: 977-982.

