

Curriculum Vitae

James G. Booth

July 2008

Contact information

Department of Biological Statistics and Computational Biology, 1178 Comstock Hall, Cornell University, Ithaca, NY 14853. Phone: 607-254-6505. Email: jim.booth@cornell.edu

Education

| | | | |
|-----|-------------|------|------------------------------|
| PhD | Statistics | 1987 | University of Kentucky |
| MS | Statistics | 1984 | University of Kentucky |
| MSc | Statistics | 1982 | University of Leeds, England |
| BSc | Mathematics | 1981 | University of Leeds, England |

Positions

| | | |
|---------------------|-------------|--------------------------------|
| Professor | 2004 – | Cornell University |
| Visiting Professor | 2003 – 2004 | Cornell University |
| Professor | 2000 – 2004 | University of Florida |
| Visiting Professor | 1996 – 1997 | Colorado State University |
| Associate Professor | 1993 – 2000 | University of Florida |
| Research Fellow | 1990 – 1992 | Australian National University |
| Assistant Professor | 1987 – 1993 | University of Florida |
| Instructor | 1986 – 1987 | Transylvania University, KY |

Awards

- *ASA Fellow*: 2004
- *Best Contributed Presentation*: Joint Statistical Meetings, New York, 2003
- *Teaching Improvement Program Award*: University of Florida, 1995
- *R.L. Anderson Outstanding Graduate Student*: University of Kentucky, 1985

Teaching

- College courses including: Categorical Data Analysis, Generalized Linear Models, Theory of Linear Models, Matrix Algebra, Multivariate Analysis, Sampling, Probability Theory, Stochastic Processes, and an assortment of undergraduate and graduate level service courses in both statistics and mathematics.
- Shortcourses for industry on Generalized Linear Models, Logistic Regression, and Categorical Data Analysis

Departmental Service

- *Chair*, Department of Biological Statistics and Computational Biology, College of Agriculture and Life Sciences, Cornell University, 2007–present
- *College Curriculum Committee*, College of Agriculture and Life Sciences, Cornell University, 2004–8
- *Acting Chair*, Department of Statistics, University of Florida, Fall 2002
- *Associate Chair*, Department of Statistics, University of Florida, 2001–2003
- *Graduate Coordinator*, Department of Statistics, University of Florida, 1995–1996, 1997–2001

Professional Service

- *Executive Committee*: Statistical Modelling Society
- *Associate Editor*: Journal of the American Statistical Association, 2006–, Statistical Modelling: 2001–, ANZ Journal of Statistics: 2001–8, Journal of the Royal Statistical Society, Series B: 2001–5, Computational Statistics: 1996–2002
- *Book Review Board*: Journal of the American Statistical Association: 1999–2002
- *NSF Review Panel*: Program in Statistics: December 10-12, 2001

Grants

- PI, NSF Grant DMS-0805865, “Applications and computational issues involving generalized linear and mixed models”, 2008-2010. Award amount \$150,000.
- Co-PI, NSF Grant DMS-0405543, “Cluster analysis, predictive distributions, and stochastic search algorithms,” 2004-2007. PI George Casella, Department of Statistics, University of Florida. Award amount: \$450,000.
- Co-PI, NSF Grant DMS-0072827, “Combining EM and Monte Carlo to maximize intractable likelihood functions,” 2000-2004, awarded \$346,000.
- PI, NSF Grant DMS-9813374, “NSF/CBMS Regional Conference in the Mathematical Sciences - Generalized Linear Mixed Models and Related Topics - June 8-12, 1999,” awarded \$22,323.
- PI, NSF Grant DMS-93010836, “Some new bootstrap methods for sample surveys,” 1993-1996, awarded \$34,658.

Invited Talks at Meetings

- Joint Statistical Meetings, August 6-10, 2006, Seattle, WA.
- Euroworkshop on Statistical Modelling - Mixed Models, Nov 2-5, 2000, Schloß Höhenried, Bernried, near Munich Germany
- Workshop on Inference and Asymptotics, July 9-14, 2000, Ascona, Switzerland
- ISI Meetings, Aug.9-18, 1999, Helsinki, Finland
- Joint Statistical Meetings, Aug 9-13, 1998, Dallas, TX
- Southern Regional Conference on Statistics, June 7-11, 1998, Navarre Beach, FL.
- Interface 98: 30th Symposium on the Interface of Computing Science and Statistics, May 13-16, 1998, Minneapolis, MN
- 7th CRM Summer School, Likelihood and Asymptotics, Aug 1-10, 1997, Banff, Canada
- 14th Statistical Methods Symposium, May 5-7, 1997, San Antonio, TX
- Miniconference on Small Area Estimation, May 1996, Southampton, England.

- International Workshop on Statistical Modelling, July 1-14, 1995, Innsbruck, Austria

Invited Talks in Other Departments

Australian National, Bath UK, Berkeley (MSRI), Buffalo, Cambridge UK, Chicago, Colorado State, Cornell, Denver, EPFL CH, Florida, Georgia, Technical U. Graz, Iowa, Iowa State, Johns-Hopkins, Kentucky, Lancaster UK, Maryland, M.D. Anderson, Munich D, Northern Illinois, Ohio State, Otago NZ, Penn State, Rice, Southampton UK, Southern Methodist, Swinbourne AUS, Sydney AUS, Texas A&M, Toronto, Virginia, Wyoming.

Graduate Students Supervised

- Haim Bar, PhD, Cornell University. Expected graduation August 2009.
- Matthias Kormacksson, PhD, Cornell University. Expected graduation August 2009.
- Vadim Zipunikov, PhD, Cornell University. Expected graduation December 2008.
- Serena Chan, MS, Cornell University. May 2006.
- David Hitchcock, PhD, August 2004. Assistant Professor at University of South Carolina. (Co-advisor George Casella)
- Bernhard Klingenberg, PhD, August 2004. Assistant Professor at Williams College. (Co-advisor Alan Agresti)
- Brian Caffo, PhD, August 2001. Associate Professor at Johns-Hopkins University.
- Wolfgang Jank, PhD, August 2001. Associate Professor at University of Maryland.
- Glen Hartless, PhD, August 2000. CIA Analyst in Washington D.C. (Co-advisor: Ramon Littell)
- Somnath Sarkar, PhD, August 1996. Statistician at Eli-Lilly, Indianapolis. (Co-advisor: M. Ghosh)

Refereed Publications in Statistics

- [1] James G Booth, George Casella, and James P Hobert. Clustering using objective functions and stochastic search. *Journal of the Royal Statistical Society*, B 70(1):119–139, 2008.

- [2] Yongsung Joo, James G. Booth, and George Casella. Model-based Bayesian clustering (MBBC). *Bioinformatics*, 2008. To appear.
- [3] Ronald W. Butler, Richard K. Sutton, James G. Booth, and Pamela Ohman Strickland. Simulation-assisted saddlepoint approximation. *Journal of Statistical Computation and Simulation*, 2008. To appear.
- [4] David B. Hitchcock, James G. Booth, and George Casella. The effect of pre-smoothing functional data on cluster analysis. *Journal of Statistical Computation and Simulation*, 77(12):1043–1055, 2007.
- [5] David B. Hitchcock, George Casella, and James G. Booth. Improved estimation of dissimilarities by smoothing functional data. *Journal of the American Statistical Association*, 101(473):211–222, 2006.
- [6] James G. Booth, Marinela Capanu, and Ludwig Heigenhauser. Exact conditional p-value calculation for the quasi-symmetry model. *Journal of Computational and Graphical Statistics*, 14(3):716–725, 2005.
- [7] James G. Booth, George Casella, Herwig Friedl, and James P. Hobert. Negative binomial loglinear mixed models. *Statistical Modelling*, 3(3):179–191, 2003.
- [8] Glen L. Hartless, James G. Booth, and Ramon C. Littell. Local influence of predictors in multiple linear regression. *Technometrics*, 45(4):326–332, 2003.
- [9] Wolfgang Jank and James G. Booth. Efficiency of Monte Carlo EM and simulated maximum likelihood in generalized linear mixed models. *Journal of Computational and Graphical Statistics*, 12(1):214–229, 2003.
- [10] Brian S. Caffo and James G. Booth. Monte Carlo conditional tests for log-linear and logistic models: a survey of current methodology. *Statistical Methods in Medical Research*, 12:1–15, 2003.
- [11] Brian S. Caffo, James G. Booth, and Anthony C. Davison. Empirical sup rejection sampling. *Biometrika*, 89:745–754, 2002.
- [12] James G. Booth and Brian S. Caffo. Unequal sampling for Monte Carlo EM algorithms. *Computational Statistics and Data Analysis*, 39:261–270, 2002.

- [13] James G. Booth, James P. Hobert, and Wolfgang Jank. A survey of Monte Carlo algorithms for maximizing the likelihood of a two-stage hierarchical model. *Statistical Modelling*, 1:333–349, 2001.
- [14] Brian S. Caffo and James G. Booth. A Markov chain Monte Carlo algorithm for approximating exact conditional tests. *Journal of Computational and Graphical Statistics*, 10:730–745, 2001.
- [15] Alan Agresti, James G. Booth, James P. Hobert, and Brian Caffo. Random effects modeling of categorical response data. *Sociological Methodology*, 30:27–80, 2000.
- [16] James G. Booth and James P. Hobert. Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo EM algorithm. *Journal of the Royal Statistical Society*, B 61:265–285, 1999.
- [17] James G. Booth, James P. Hobert, and Pamela A. Ohman. On the probable error of the ratio of two gamma means. *Biometrika*, 86:439–452, 1999.
- [18] James G. Booth and Ronald W. Butler. An importance sampling algorithm for exact conditional tests in log-linear models. *Biometrika*, 86:321–332, 1999.
- [19] James G. Booth and James P. Hobert. Standard errors of prediction in generalized linear mixed models. *Journal of the American Statistical Association*, 93:262–272, 1998.
- [20] James G. Booth and Somnath Sarkar. Monte Carlo approximation of bootstrap variances. *American Statistician*, 52:354–357, 1998.
- [21] James G. Booth and Brett Presnell. Allocation of Monte Carlo resources for the iterated bootstrap. *Journal of Computational and Graphical Statistics*, 7:92–112, 1998.
- [22] James G. Booth, Ronald W. Butler, Snehelata Huzurbazar, and Andrew T. A. Wood. Saddlepoint approximations for p-values of some tests of covariance matrices. *Journal of Statistical Computation and Simulation*, 53:165–180, 1995.

- [23] James G. Booth and Andrew T. A. Wood. An example in which the lugannani-rice saddlepoint formula fails. *Statistics and Probability Letters*, 23:53–61, 1995.
- [24] James G. Booth, Ronald W. Butler, and Peter Hall. Bootstrap methods for finite populations. *Journal of the American Statistical Association*, 89:1282–1289, 1994.
- [25] James G. Booth, Peter Hall, and Andrew T. A. Wood. On the validity of edgeworth and saddlepoint expansions. *Journal of Multivariate Analysis*, 51:121–138, 1994.
- [26] James G. Booth and Peter Hall. Monte Carlo approximation and the iterated bootstrap. *Biometrika*, 81(2):331–340, 1994.
- [27] James G. Booth and Peter Hall. An improvement of the jackknife distribution function estimator. *Annals of Statistics*, 21(3):1476–1485, 1993.
- [28] James G. Booth and Peter Hall. Bootstrap confidence regions for functional relationships in errors-in-variables models. *Annals of Statistics*, 21(4):1780–1791, 1993.
- [29] James G. Booth and Kim-Anh Do. Simple and efficient resampling methods for constructing bootstrap confidence intervals. *Computational Statistics*, 8:333–346, 1993.
- [30] Andrew T. A. Wood, James G. Booth, and Ronald W. Butler. Saddlepoint approximations to the cdf of some statistics with nonnormal limit distributions. *Journal of the American Statistical Association*, 88(422):680–686, 1993.
- [31] James G. Booth, Peter Hall, and Andrew T. A. Wood. Balanced importance resampling for the bootstrap. *Annals of Statistics*, 21(1):286–298, 1993.
- [32] Ronald W. Butler, James G. Booth, and Snehelata Huzurbazar. Saddlepoint approximations for tests of block independence, sphericity, and equal variances and covariances. *Journal of the Royal Statistical Society, B* 55:171–184, 1993.

- [33] Ronald W. Butler, Snehelata Huzurbazar, and James G. Booth. Saddlepoint approximations for the bartlett–nanda–pillai trace statistic in multivariate analysis. *Biometrika*, 79:705–715, 1992.
- [34] James G. Booth. A note on a one compartment model with clustering. *Journal of Applied Probability*, 29:535–542, 1992.
- [35] James G. Booth, Peter Hall, and Andrew T. A. Wood. Bootstrap estimation of conditional distributions. *Annals of Statistics*, 20(3):1594–1610, 1992.
- [36] Ronald W. Butler, Snehelata Huzurbazar, and James G. Booth. Saddlepoint approximations to the generalized variance and wilk’s statistic. *Biometrika*, 79:157–169, 1992.
- [37] James G. Booth. A note on the estimation of rate parameters in stochastic compartmental models. *Communications in Statistics*, B 20(1):391–397, 1991.
- [38] James G. Booth and Ronald W. Butler. Randomization distributions and saddlepoint approximations in generalized linear models. *Biometrika*, 77:787–796, 1990.
- [39] James G. Booth. On the limiting behavior of downtown’s carrier epidemic in the case of a general infection mechanism. *Journal of Applied Probability*, 26:625–630, 1989.
- [40] James G. Booth, Joseph Gani, Marie-Pierre Malice, H. Mansouri, and Gaby Maravankin. A general solution for the epidemic with carriers. *Statistics and Probability Letters*, 4:9–15, 1986.

Collaborative Publications

- Davis, Lewenstein, Simon, Booth, Connolly, ”Open access publishing, article downloads, and citations: randomised controlled trial”, *British Medical Journal* 2008;337:a568
- Yongsung Joo, James G. Booth, Younghwan Namkoong, and George Casella “Model-based Bayesian clustering (MBBC)”, *Bioinformatics* March 2008; 24:874-875.

- Brooks, MacNguyen, Hall, Gupta and Booth, “Indirect Carrier Detection of Canine Hemophilia A Using Factor VIII Microsatellite Markers”, *Animal Genetics*, June 2008, 39:306-309.
- Kim, Booth, Gauch, Sun, Park, Lee and Lee, “Simple sequence repeats in *Neurospora crassa*: distribution, polymorphism and evolutionary inference”, *BMC Genomics* 2008, 9:31.
- Michael, Park, Kim, Booth, Byer, Sun, Chory and Lee (2007), “Simple sequence repeats provide a substrate for phenotypic variation in the *Neurospora crassa* circadian clock”, *PLoS ONE*, 2(8):e795.
- Joo, Casella, Booth, Lee and Enkeman (2007), “Normalization of Dye Bias in Microarrays Using the Mixture of Splines Model”, *Statistical Applications in Genetics and Molecular Biology*,6(1):Art.2.
- Booth and Willett (1996), “The statistical analysis of standard costing variances”, *IMA Journal of Mathematical Applications in Business and Industry*, 8:167-179.

Non-refereed Publications

- James G. Booth, Review of “Generalized Linear Models with Random Effects: Unified Analysis via H-likelihood”, by Lee, Nelder and Pawitan. *Biometrics* 63:1296
- James G. Booth, Review of “Exact Analysis of Discrete Data”, by Karim F. Hirji. *J. Amer. Statist. Assoc.* 102(478):764, 2007.
- James G. Booth, Review of “Analysis of Variance for Random Models: Theory, Methods, Applications, and Data Analysis. Volume 1: Balanced Data, and Volume 2: Unbalanced Data”, by Hardeo Sahai and Mario Miguel Ojeda. *J. Amer. Statist. Assoc.* 101, No.475, 2006.
- James G. Booth and Brett Presnell, “Resampling Methods for Sample Surveys,” *Bulletin of the International Statistical Institute*, 52nd Session Proceedings, 1999.
- James G. Booth, Review of *Odds Ratios in the Analysis of Contingency Tables*, by Tamas Rudas. #119 in Series: Quantitative Applications in the Social Sciences (1998), in *J. Amer. Statist. Assoc.* 94, 657, 1999.
- James G. Booth, Review of *Series Approximation Methods in Statistics*, 2nd Edition, by J. E. Kolassa. Lecture Notes in Statistics 88: Springer (1996), in *J. Amer. Statist. Assoc.* 94, 343, 1999

- James G. Booth and Ronald W. Butler, Contribution to the discussion of “Approximately exact inference for the common odds-ratio in several 2×2 tables,” by Strawderman, R.L. and Wells, M.T. *J. Amer. Statist. Assoc.* **93**, 1310-1313, 1998.
- James G. Booth, “Bootstrap methods for generalized linear mixed models with applications to small area estimation,” Proceedings of the 10th International Workshop on Statistical Modelling, Lecture Notes in Statistics Series: Springer, 1995.
- James G. Booth, “A note of the accuracy of two saddlepoint tail probability approximations,” Proceedings of the American Statistical Association – Section on Physical and Engineering Sciences, 1994.
- James G. Booth, Review of *Resampling-Based Multiple Testing* by P.H. Westfall and S.S. Young. New York: John Wiley (1993), in *J. Amer. Statist. Assoc.* **89**, 354-55, 1994.
- James G. Booth and Andrew T.A. Wood, Contribution to the discussion of “Practical use of higher order asymptotics for multiparameter exponential families,” by Pierce, D.A. & Peters, D. *J. Roy. Statist. Soc.*, **B 54**, 701-38, 1992.

Submitted Articles

- Booth, Federer, Wells and Wolfinger, “A multivariate variance components model for the analysis of covariance in designed experiments”, Submitted to *Statistical Science*
- Zipunnikov, Booth and Yoshida, “Table counting using the saddlepoint approximation”, Submitted to *Journal of Computational and Graphical Statistics*.
- Hirschl, Booth and Glenna, “The link between biblical literalism and voter choice in the U.S. presidential elections, 1980-2000”, Submitted to *Journal for the Scientific Study of Religion*.