

Research summary:

My research involves computationally intensive statistical methods. I apply these methods to models for clustered and longitudinal categorical data and medical image processing and analysis. My current research agenda focuses on the reconstruction and validation of reconstruction methods for nuclear medical images and the analysis of experimental medical imaging data.

A specific focus of my research is the computation and application of statistical tools called multilevel models. This flexible class of statistical techniques, which includes linear mixed effect models, generalized- and non-linear mixed effect models and Bayesian models, are applicable to a nearly endless variety of scientific settings. In particular, I have applied these methods to social science data, general contingency table data, pharmacological data, brain imaging region of interest studies, voxel-by-voxel brain imaging studies, electro encephalogram data and sleep transition data.

Due to a successful K25 application, I have endeavored to focus my research on the application and computation of multilevel models to medical imaging data. I have focused on nuclear medicine and brain imaging studies from modalities such as: positron emission tomography, single photon emission computed tomography, magnetic resonance imaging, and functional magnetic resonance imaging.

My more theoretical work focuses on the computational aspects of these models and general computing algorithms. As such, I have worked on likelihood based marginalized models for categorical data, the expectation/maximization algorithm and Markov chain Monte Carlo.

To give examples of my research, in Booth and Caffo (2002) and Caffo et al. (2004) we explored novel adaptations of the Monte Carlo EM algorithm. A new Monte Carlo sampling procedure was presented in Caffo et al. (2002). Monte Carlo applications to the analysis of contingency tables were developed in Caffo and Booth (2001) and later modified into software listed at the Comprehensive R-Archive Network in Caffo (2006). Further work on the analysis and computation for longitudinal categorical data can be found in Caffo et al. (2007) and Caffo and Griswold (2006). My work imaging started with a high-profile manuscript on fMRI and Alzheimer's disease (Bassett et al. 2006) and was further explored in MRI and environmental toxicology in Schwartz et al. (2007) and Caffo et al. (2007b). Finally, I have explored algorithms for three dimensional curve fitting in Caffo et al. (2008).

Bassett S, Yousem D, Cristinzio C, Kusevic I, Yassa M, Caffo B, Zeger S (2006). Familial Risk for Alzheimer's Disease Alters fMRI Activation Patterns: A Study of Asymptomatic Offspring. *Brain* 129: 1229-1239.

Booth J, Caffo B (2002). Unequal Sampling for Monte Carlo EM Algorithms. *Computational Statistics and Data Analysis* 39: 261-270.

Bowman D, Caffo B, Bassett S, Kilts C (2007). A Bayesian Hierarchical Framework for

- Spatial Modeling of fMRI Data. *Neuroimage*. Published online, print version in press.
- Caffo B (2006). Exact Hypothesis Tests for Log-linear Models with exactLoglinTest. *The Journal of Statistical Software* 17 (7): 1-17.
- Caffo B, An M, Rohde C (2007a). Flexible Random Intercept Models for Binary Outcomes Using Mixtures of Normals. *Computational Statistics and Data Analysis, Special Issue on Advances in Mixture Models* 51: 5220-5235.
- Caffo B, Booth J (2001). A Markov Chain Monte Carlo Algorithm for Approximating Exact Conditional Probabilities. *The Journal of Computational and Graphical Statistics* 10: 730-745.
- Caffo B, Booth J, Davison A (2002). Empirical Sup Rejection Sampling. *Biometrika* 89: 745-754.
- Caffo B, Chen S, Bolla K, Davatzikos C, Youssef D, Schwartz B (2007b). Is Lead's Effect On Cognition Mediated Through Structural Changes Detectable by MRI? *American Journal of Epidemiology*. Published online, print version in press.
- Caffo, B, Crainiceanu, C, Deng, L, Hendrix, C (2008). A Case Study in Pharmacologic Imaging Using Principal Curves in Single Photon Emission Computed Tomography. Accepted in *Journal of the American Statistical Association*
- Caffo B, Jank W, Jones G (2004). Ascent-based MCEM. *The Journal of the Royal Statistical Society Series B* 67: 235-251.
- Caffo B, Griswold M (2006) A User-friendly Tutorial on Link-probit-normal Models. *The American Statistician* 60: 139-145.
- Schwartz B, Chen S, Caffo B, Stewart F, Bolla K, Youssef D, Liu D, Davatzikos K (2007). Relations of Brain Volumes With Cognitive Function in Males 45 Years and Older With Past Lead Exposure. *Neuroimage* 37: 633-641