

The Final Oral Examination  
for the Degree of

DOCTOR OF PHILOSOPHY  
(Department of Mathematics and Statistics)

**RABIH SAAB**

|        |                   |     |
|--------|-------------------|-----|
| [2006] | [York University] | MSc |
| [2005] | [York University] | BSc |

“Nonparametric estimation of the mixing distribution in mixed models  
with random intercepts and slopes”

April 9<sup>th</sup>, 2013, at 10:00 am  
Social Sciences and Mathematics Building, Room A144

Supervisory Committee:

Dr. Mary Lesperance, Department of Mathematics and Statistics, UVic  
(Supervisor)

Dr. Farouk Nathoo, Department of Mathematics and Statistics, UVic  
(Academic Unit Member)

Dr. David E. Giles, Department of Economics, UVic  
(Non Unit Member)

External Examiner:

Dr. Peter D.M. Macdonald, McMaster University

Chair of Oral Examination:

Dr. Francis Lau, School of Health Information Science, UVic

## Abstract

Generalized linear mixture models (GLMM) are widely used in statistical applications to model count and binary data. We consider the problem of nonparametric likelihood estimation of mixing distributions in GLMM's with multiple random effects. The log-likelihood to be maximized has the general form

$$l(G) = \sum_i \log \int f(y_i, \gamma) dG(\gamma)$$

where  $f(\cdot, \gamma)$  is a parametric family of component densities,  $y_i$  is the  $i^{th}$  observed response dependent variable, and  $G$  is a mixing distribution function of the random effects vector  $\gamma$  defined on  $\Omega$ .

The literature presents many algorithms for maximum likelihood estimation (MLE) of  $G$  in the univariate random effect case such as the EM algorithm (Laird, 1978), the intra-simplex direction method, ISDM (Lesperance and Kalbfleish, 1992), and vertex exchange method, VEM (Böhning, 1985). In this dissertation, the constrained Newton method (CNM) in Wang (2007) is extended to fit clustered datasets with multiple random effects. Owing to the general equivalence theorem from the geometry of mixture likelihoods (see Lindsay, 1995), many NPMLE algorithms including CNM and ISDM maximize the directional derivative of the log-likelihood to add potential support points to the mixing distribution  $G$ . Our method, *Direct Search Directional Derivative* (DSDD), uses a directional search method to find local maxima of the multi-dimensional directional derivative function. The DSDD's performance is investigated in GLMM where  $f$  is a Bernoulli or Poisson distribution function. The algorithm is also extended to cover GLMM's with zero-inflated data.

Goodness-of-fit (GOF) and selection methods for mixed models have been developed in the literature, however their application in models with nonparametric random effects distributions is vague and ad-hoc. Some popular measures such as the Deviance Information Criteria (DIC), conditional Akaike Information Criteria (cAIC) and  $R^2$  statistics are potentially useful in this context. Additionally, some cross-validation goodness-of-fit methods popular in Bayesian applications, such as the conditional predictive ordinate (CPO) and numerical posterior predictive checks, can be applied with some minor modifications to suit the non-Bayesian approach.

## Published Papers and Conference Proceedings:

- Lu, L., Mann, R. K., Saab, R. and Stone, R. (2011). Disease status determination: Exploring imputation and selection techniques. *Canadian Journal of Statistics*, **39** (2), 197-201.
- Lesperance, M., Saab, R., and Neuhaus, J. (2013). Non-parametric estimation of the mixing distribution in logistic regression mixed models with random intercepts and slopes. Unpublished manuscript.

## Presentations and Poster Sessions:

- Saab, R. and Lesperance, M. Nonparametric maximum likelihood estimation for generalized linear mixture models (GLMM) with random intercepts and slopes. *Kalbfleish Symposium (Ann Arbor, Michigan, Sep. 28th-30th, 2012)*, poster.
- Lu, L., Mann, R., Saab, R. and Stone, R. Proteomic Biomarkers for Disease Status. *Statistical Society of Canada Conference (Vancouver, BC, May 31-June 3, 2009)*, poster.

## Awards, Grants and Fellowships:

- Best Junior award in Mathematics, *University of Tampa, 2003*
- York University graduate fellowship, *York University, 2005*
- University of Victoria graduate fellowship, *University of Victoria, 2007*