

# Accurate Frequentist Generalised Linear Mixed Model Analysis via Expectation Propagation

by **James Yu**

Thesis submitted in fulfilment of the requirements for the degree of

**Doctor of Philosophy in Mathematics**

under the supervision of **Prof. Matt Wand** and **Dr. Shev MacNamara**

**University of Technology Sydney**

Faculty of Science

**April 22, 2021**

# Certificate of original authorship

I, **James Yu** declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy in Mathematics, in the School of Mathematical and Physical Sciences at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree at any other academic institution except as fully acknowledged within the text. This thesis is the result of a Collaborative Doctoral Research Degree program with the Australian Red Cross Blood Service.

This research is supported by an Australian Government Research Training Program.

Production Note:

**Signature:** Signature removed prior to publication.

**Date:** April 22, 2021

# Acknowledgements

This thesis would not have been possible without the support of a number of individuals.

Many thanks to my mentor and supervisor Matt Wand. Thank you for your patience and kindness. This achievement would not have been possible without you. I could not have asked for a better supervisor.

I would also like to thank Stephen Wright. Thank you for your care and for the time you spent with me. I really appreciate it.

To my colleagues and friends at UTS, thank you for your help, constant support, for making me laugh and for listening to me.

Finally to Mum, Dad and Mario, I am eternally grateful to you for giving me the opportunities and experiences that have made me who I am. I dedicate this milestone to you.

# List of papers/publications

The following list of paper and publications awards relate to the work presented in this thesis:

- Hall, P., Johnstone, I. M., Ormerod, J. T., Wand, M. P., & Yu, J. C. F. (2020). Fast and Accurate Binary Response Mixed Model Analysis via Expectation Propagation, *Journal of the American Statistical Association*, 115:532, 1902–1916, DOI: 10.1080/01621459.2019.1665529.
- Wand, M. P. & Yu, J. C. F. (2020). glmmEP: Generalized Linear Mixed Model Analysis via Expectation Propagation (Version 1.0-3.1). Comprehensive R Archive Network. Retrieved from <https://cran.r-project.org/web/packages/glmmEP/index.html>
- Yu, J. C. F. (2019), “Fast and Accurate Frequentist Generalised Linear Mixed Model Analysis via Expectation Propagation”, (2019) Enabling Algorithms Theme Symposium, University of Technology Sydney, Sydney, 13-14 June.
- Wand, M. P., & Yu, J. C. F. (2021) Density Estimation via Bayesian Inference Engines.
- First Place, 20th Annual J.B. Douglass Award, New South Wales Branch of Statistical Society of Australia, (2019).

# Notation

In this chapter, important notations are introduced that we refer to throughout this thesis.

## Acronyms

Table 1: Table with acronyms used in the thesis with their meanings.

<b>Acronym</b>	<b>Meaning</b>
AGHQ	Adaptive Gauss-Hermite quadrature
BFGS	Broyden-Fletcher-Goldfarb-Shanno
BP	Best predictor
CDF	Cummulative density function
DAG	Directed acyclic graph
DC	Data cloning
EP	Expectation propagation
GHQ	Gauss-Hermite quadrature
GLMM	Generalised linear mixed models
KL	Kullback-Leiber
MCMC	Markov chain Monte Carlo
NM	Nelder-Mead
PDF	Probability density function
PQL	Penalised quasi-likelihood

# Contents

<b>1 Introduction and background</b>	<b>2</b>
1.1 Introduction	2
1.2 Thesis aim	4
1.3 Matrix storage and notation	5
1.4 Notation for spaces	7
1.5 Exponential family theory and distributions	7
1.5.1 Exponential family theory	7
1.5.2 Probability distribution	9
1.5.2.1 Univariate normal distribution	9
1.5.2.2 Multivariate normal distribution	10
1.5.2.3 Bernoulli distribution	11
1.5.2.4 Poisson distribution	12
1.5.2.5 Negative binomial distribution	13
1.6 Graph theory	13
1.6.1 Directed acyclic graphs	14
1.6.2 Factor graphs	15
1.7 Multilevel datasets	15
1.8 Generalised linear mixed models	19
1.8.1 Binary response models	22
1.8.2 Count response models	22
1.9 Maximum likelihood	23
1.9.1 Likelihood functions	23
1.9.2 Maximum likelihood	25
1.10 Best prediction	25
1.11 Current approximation methods	26
1.11.1 Laplace approximation	27
1.11.2 Gauss Hermite quadrature	28
1.11.3 Other methods	29
1.12 Expectation propagation	30
1.12.1 Kullback Leiber divergence and projection	31
1.12.2 Mean field approximations	33
1.12.3 Expectation propagation	34
1.12.4 Message passing	37
1.13 Thesis structure	39

<b>2</b>	<b>Expectation propagation for the simplest one level probit mixed model</b>	<b>40</b>
2.1	Traditional quadrature likelihood approximation	41
2.2	Expectation propagation likelihood approximation	42
2.2.1	Projection onto the unnormalised normal family	45
2.2.2	Message passing formulation	47
2.2.3	Starting values for Algorithm 3	54
2.3	Evaluation of the estimates	55
2.4	Computing point estimates and confidence intervals	56
2.4.1	Confidence interval estimation	57
2.4.2	Derivative approximation	58
2.4.2.1	Expectation propagation analytical approach I	61
2.4.2.2	Expectation propagation analytical approach II	63
2.5	Best predictor	70
2.6	Simulation Study	71
2.7	Appendix	75
2.7.1	Proof of Result 3	75
2.7.2	Proof of Result 5	76
2.7.2.1	Proof of Lemma 1	77
2.7.3	Proof of Result 7	80
2.7.4	Details on finding the inverse function gradient map log-partition function for each $s$	81
2.7.4.1	The $s = 0$ case	81
2.7.4.2	The $s = 2$ case	82
2.7.4.3	The $s = 4$ case	82
<b>3</b>	<b>Expectation propagation for general one level probit mixed models</b>	<b>83</b>
3.1	Expectation propagation likelihood approximation	84
3.1.1	Projection onto the unnormalised multivariate normal family	88
3.1.2	Message passing formulation	89
3.1.3	Starting values for Algorithm 6	95
3.2	Computation of point estimates and confidence intervals	96
3.3	Best predictor	99
3.4	Simulation study	100
3.4.1	Comparison of maximum likelihood estimates for univariate random effects	100
3.4.2	Maximum likelihood estimates for bivariate random effects	104
3.4.3	Maximum likelihood estimates for trivariate random effects	106
3.5	Appendix	108
3.5.1	Proof of Definition 13	108
3.5.1.1	Proof of Lemma 2	108
3.5.1.2	Proof of Lemma 3	109
3.5.2	Proof of Result 12	111
<b>4</b>	<b>Expectation propagation for one level logistic mixed models</b>	<b>113</b>
4.1	The simplest logistic mixed model	114
4.1.1	Traditional quadrature likelihood approximation	115
4.1.2	Expectation propagation likelihood approximation	115

4.1.2.1	Projection onto the unnormalised normal family	118
4.1.2.2	Message passing formulation	119
4.1.2.3	Starting values for the univariate logistic case	120
4.1.3	Evaluation of the estimates	121
4.1.4	Best predictor	123
4.2	General logistic mixed models	123
4.2.1	Expectation propagation likelihood approximation	124
4.2.1.1	Projection onto the unnormalised multivariate normal family	127
4.2.1.2	Message passing formulation	130
4.2.1.3	Starting values for the multivariate logistic case	131
4.2.2	Simulation study	132
4.2.2.1	Comparison of maximum likelihood estimates for univariate random effects	132
4.2.2.2	Maximum likelihood estimates for bivariate random effects	134
4.3	Appendix	137
4.3.1	Proof of Result 16	137
4.3.2	Proof of Definition 17	138
4.3.3	Proof of Definition 18	140
<b>5</b>	<b>Expectation propagation for one level count response mixed models</b>	<b>142</b>
5.1	The simplest Poisson mixed model	143
5.1.1	Traditional quadrature likelihood approximation	144
5.1.2	Expectation propagation likelihood approximation	144
5.1.2.1	Projection onto the unnormalised normal family	146
5.1.2.2	Message passing formulation	147
5.1.2.3	Starting values for the Poisson case	148
5.1.3	Evaluation of the estimates	150
5.1.4	Best predictor	151
5.2	General Poisson mixed models	151
5.2.1	Expectation propagation likelihood approximation	152
5.2.1.1	Projection onto the unnormalised multivariate normal family	154
5.2.1.2	Message passing formulation	155
5.2.1.3	Starting values for the Poisson case	156
5.2.2	Simulation study	157
5.3	The simplest negative binomial models	160
5.3.1	Traditional quadrature likelihood approximation	161
5.3.2	Expectation propagation likelihood approximation	161
5.3.2.1	Projection onto the unnormalised normal family	163
5.3.2.2	Message passing formulation	165
5.3.2.3	Starting values for the negative binomial case	166
5.3.3	Evaluation of the estimates	167
5.3.4	Computation of point estimates and confidence intervals	168
5.3.4.1	Derivative approximation	169
5.4	General negative binomial model	169
5.4.1	Expectation propagation likelihood approximation	170



5.4.1.1	Projection onto the unnormalised multivariate normal family	172
5.4.1.2	Message passing formulation	174
5.4.1.3	Starting values for the negative binomial case	175
5.4.2	Computation of point estimates and confidence intervals	176
5.4.3	Simulation study	177
5.5	Varying dispersion negative binomial model	179
5.6	Appendix	181
5.6.1	Proof of Definition 21	181
5.6.2	Proof of Definition 23	182
5.6.3	Proof of Definition 25	184
5.6.4	Proof of Definition 27	185
<b>6</b>	<b>Expectation propagation for two level and crossed random effects probit models</b>	<b>187</b>
6.1	The general probit crossed mixed model	187
6.1.1	Expectation propagation likelihood approximation	188
6.1.1.1	Message passing formulation	189
6.1.2	Computation of point estimates and confidence intervals	196
6.1.3	Simulation study	199
6.1.3.1	Comparison with MCMC and Laplace approximation	199
	maximum likelihood for crossed random effects	199
6.2	The general probit two level mixed model	200
6.2.1	Expectation propagation likelihood approximation	202
6.2.1.1	Message passing formulation	202
<b>7</b>	<b>Applications of expectation propagation for one level probit mixed models</b>	<b>210</b>
7.1	R package “glmmEP”	210
7.2	Modelling immunisation of Guatemalan children	211
7.3	Modelling donor attendance of the Australian Red Cross Blood Service	216
7.3.1	Data cleaning	220
7.3.2	Modelling continuous variables	222
7.3.3	Initial model	223
7.3.4	Second model	226
7.3.4.1	Results of model fit	227
7.4	Appendix	233
<b>8</b>	<b>Discussion and conclusion</b>	<b>235</b>
<b>9</b>	<b>References</b>	<b>239</b>

# Abstract

Generalised linear mixed models are a particularly powerful and well established statistical tool. Unlike linear mixed models, where the integrals arising in likelihood functions can be expressed in closed form, the likelihood functions expressed in generalised linear mixed models do not follow tractable solutions. Methods such as Gauss-Hermite quadrature and Laplace approximation are the standard approaches to overcome these integrals. Although Gauss-Hermite quadrature is accurate it is also slow, rendering it unsuitable for analyses with more than two or three random effects. Laplace approximations are the most feasible solution, however the approximate inference they provide in binary models is well known to be inaccurate. A less common approach is to use Bayesian ideas such as data cloning, however they involve a number of technicalities and as such are difficult to implement. Although expectation propagation is generally used in Bayesian settings, in this thesis we introduce a novel approach where we use it as frequentists to achieve high accuracy results with minimal computational cost for inference on generalised linear mixed models. We show our methodology can be used to solve one level probit models without the need for quadrature, providing consistent and accurate results. We explain how using quadrature we can also extend our method to logistic, Poisson and negative-binomial models. Additionally we show how these models can be extended to two level models and crossed random effects models for the probit case. Finally we present applications of our methodology on two real datasets, both with different technical challenges.