
Contents

PREFACE	V
ACKNOWLEDGEMENTS.....	V
DATASETS AND R CODE USED IN THIS BOOK	VI
CHAPTER 1 OF ZUUR ET AL. (2012A) AND ZUUR (2012B)	VI
COVER ART	VI
CONTRIBUTORS	XII

1 INTRODUCTION TO GENERALIZED LINEAR MODELS1

1.1 LINEAR REGRESSION APPLIED ON FISHERIES DATA	1
1.2 POISSON GLM.....	7
1.2.1 Poisson distribution for count data	7
1.2.2 Predictor function	8
1.2.3 Linking the mean and the predictor function	9
1.2.4 Maximum likelihood to estimate the parameters	9
1.2.5 Application of Poisson GLM on the fisheries data	11
1.2.6 Overdispersion	19
1.2.7 Adding covariates	23
1.2.8 Using the offset.....	24
1.3 NEGATIVE BINOMIAL GLM	26
1.3.1 Negative binomial distribution for count data	26
1.3.2 Example of Negative binomial regression	28
1.3.3 Heterogeneous Negative binomial regression	34
1.3.4 A note on modelling under-dispersion	36
1.4 BINOMIAL GLM FOR BINARY DATA	36
1.4.1 Parasites in honeybee larvae	36
1.4.2 Visualizing the data	36
1.4.3 Defining the three steps of a binomial GLM	38
1.4.4 Results for the bee data	40
1.4.5 Likelihood function for a binomial GLM	42
1.4.6 Other link functions	42
1.5 BINOMIAL GLM FOR PROPORTIONAL DATA	43
1.5.1 Binomial distribution.....	43
1.5.2 Predictor function	45
1.5.3 Link function.....	45
1.5.4 Fitting the model in R.....	45
1.6 OTHER DISTRIBUTIONS	47

2 GENERALIZED LINEAR MODELLING APPLIED TO RED

SQUIRREL DATA	49
2.1 RED SQUIRRELS	49
2.2 IMPORTING THE DATA	50
2.3 DATA EXPLORATION	51
2.3.1 Outliers	51
2.3.2 Collinearity.....	52

2.3.3 Relationships	54
2.4 FITTING THE POISSON GLM IN R	55
2.4.1 Specifying the model.....	55
2.4.2 Execute the Poisson GLM in R.....	55
2.4.3 Model validation	57
2.5 FITTING THE NEGATIVE BINOMIAL GLM IN R	60
2.5.1 Using the <i>glm.nb</i> function	60
2.5.2 Heterogeneous negative binomial GLM	63
2.6 BAYESIAN APPROACH – RUNNING THE POISSON GLM	66
2.6.1 Obtaining and installing JAGS	66
2.6.2 Specifying the data for JAGS	67
2.6.3 Specifying the model for JAGS.....	68
2.6.4 Specifying the initial values.....	69
2.6.5 Parameters to store	69
2.6.6 Running JAGS via R.....	69
2.6.7 Generalizing the JAGS modelling code	72
2.7 ASSESSING MIXING OF CHAINS	74
2.7.1 Assess mixing of chains if <i>R2jags</i> is used.....	74
2.8 MODEL VALIDATION	76
2.8.1 Checking for overdispersion	76
2.8.2 Obtaining Pearson residuals.....	77
2.9 APPLYING A NEGATIVE BINOMIAL GLM IN JAGS.....	79
2.10 MIXING OF CHAINS	82
2.11 MODEL VALIDATION	83
2.12 MODEL INTERPRETATION	84
2.13 DISCUSSION.....	87
2.14 WHAT TO PRESENT IN A PAPER	87

3 GLM APPLIED TO PRESENCE-ABSENCE POLYCHAETA

DATA	89
3.1 MARINE BENTHIC DATA	89
3.2 IMPORTING THE DATA AND HOUSEKEEPING	90
3.3 DATA EXPLORATION	91
3.4 BINARY GLM; A FREQUENTIST APPROACH	94
3.4.1 Specifying the distribution and link function.....	94
3.4.2 Specifying the predictor function	95
3.4.3 Running the <i>glm</i> function	96
3.4.4 Results of the <i>glm</i> function	96
3.4.5 Model selection	97
3.4.6 Results of the optimal model	100
3.4.7 Model validation	101
3.4.8 Visualizing the model	102
3.5 FITTING A BERNOULLI GLM IN JAGS	103
3.5.1 Specifying the data for JAGS	103
3.5.2 JAGS modelling code	104
3.5.3 Initial values and parameters to save	105
3.5.4 Running JAGS from R	105

3.5.5 JAGS results presented within R	106
3.6 MODEL SELECTION USING AIC, DIC AND BIC IN JAGS	107
3.7 MODEL INTERPRETATION	110
3.8 DISCUSSION	113
3.9 WHAT TO PRESENT IN A PAPER	114
4 INTRODUCTION TO MIXED EFFECTS MODELS.....	115
4.1 SPIDERS	115
4.2 LINEAR REGRESSION APPLIED ON THE SPIDER DATA	115
4.3 LINEAR MIXED EFFECTS MODELS	118
4.3.1 Model formulation and interpretation.....	118
4.3.2 Fitting a linear mixed effects model using lmer	119
4.3.3 Analysis using lmer	122
4.4 FITTING A LINEAR MIXED EFFECTS MODEL IN JAGS	128
4.5 USING A VARIABLE AS A FIXED OR RANDOM TERM?.....	131
4.6 RANDOM INTERCEPT AND SLOPE MODEL	131
4.7 GENERALIZED LINEAR MIXED EFFECTS MODELS	132
5 GLMM APPLIED ON HONEYBEE POLLINATION DATA	133
5.1 HONEYBEES AND DANDELION POLLEN	133
5.2 DATA DESCRIPTION AND IMPORTING THE DATA	134
5.3 DATA EXPLORATION	135
5.4 BUILDING UP A MODEL	136
5.5 POISSON GLMM USING GLMER	137
5.6 POISSON GLMM USING JAGS	140
5.6.1 Data for JAGS	140
5.6.2 JAGS modelling code	141
5.6.3 Likelihood	142
5.6.4 Priors	142
5.6.5 Initial values	144
5.6.6 Parameters to save	144
5.6.7 Executing JAGS and obtaining results	145
5.7 NEGATIVE BINOMIAL GLMM USING GLMMADMB	146
5.8 NEGATIVE BINOMIAL GLMM USING JAGS	147
5.8.1 Data for JAGS	147
5.8.2 JAGS modelling code	147
5.8.3 Initial values	148
5.8.4 Parameters to save	149
5.8.5 Executing JAGS and obtaining results	149
5.8.6 Mixing of chains	150
5.8.7 Model validation	150
5.8.8 Model interpretation	152
5.9 GLMM WITH AUTO-REGRESSIVE CORRELATION.....	154
5.9.1 Simulate temporal correlated counts	155
5.9.2 JAGS to estimate the Poisson GLM with AR correlation...158	
5.9.3 Multiple Poisson time series	161
5.9.4 Poisson GLMM with AR correlation	161
5.10 WHAT TO PRESENT IN A PAPER	164

6 GLMM FOR STRICTLY POSITIVE DATA: BIOMASS OF RAINFOREST TREES	165
6.1 RAINFOREST TREE SPECIES	165
6.2 IMPORTING THE DATA AND HOUSEKEEPING	167
6.3 DATA EXPLORATION	168
6.3.1 <i>Outliers</i>	168
6.3.2 <i>Collinearity</i>	169
6.3.3 <i>Relationships</i>	170
6.4 MULTIPLE LINEAR REGRESSION: A FREQUENTIST APPROACH	173
6.5 GAMMA GLM USING A FREQUENTIST APPROACH	175
6.5.1 <i>Formulating the gamma GLM</i>	175
6.5.2 <i>Scale and shape</i>	176
6.5.3 <i>Visualizing the gamma distribution</i>	176
6.5.4 <i>Different link functions</i>	178
6.5.5 <i>Running the Gamma GLM using the glm function</i>	179
6.5.6 <i>Scale confusion</i>	179
6.5.7 <i>Identity link and inverse link function</i>	182
6.6 FITTING A GAMMA GLM USING JAGS	183
6.6.1 <i>Specifying the data for JAGS</i>	183
6.6.2 <i>JAGS modelling code</i>	185
6.6.3 <i>Priors</i>	185
6.6.4 <i>Likelihood function</i>	185
6.6.5 <i>Initial values and parameters to save</i>	186
6.6.6 <i>Running JAGS from R</i>	186
6.6.7 <i>JAGS results presented within R</i>	187
6.6.8 <i>Model interpretation</i>	190
6.6.9 <i>Model validation</i>	193
6.7 ADDING MORE COVARIATES TO THE GAMMA GLM IN JAGS	195
6.8 GAMMA GLMM	195
6.8.1 <i>R code for a gamma GLMM in JAGS</i>	196
6.8.2 <i>Results from JAGS for the gamma GLMM</i>	198
6.9 TRUNCATED GAUSSIAN LINEAR REGRESSION	199
6.9.1 <i>Zero trick to fit any statistical distribution in JAGS</i>	199
6.9.2 <i>Multiple linear regression in JAGS with the zero trick</i>	200
6.9.3 <i>Tobit model in JAGS</i>	202
6.9.4 <i>Tobit model with random effects in JAGS</i>	205
6.10 DISCUSSION	205
6.11 WHAT TO PRESENT IN A PAPER	206
7 BINOMIAL, BETA-BINOMIAL, AND BETA GLMM APPLIED TO CHEETAH DATA	207
7.1 STEREOTYPIC BEHAVIOURS IN CAPTIVE CHEETAHS	207
7.2 IMPORTING THE DATA	209
7.3 DATA EXPLORATION	209
7.3.1 <i>Outliers</i>	209
7.3.2 <i>Collinearity</i>	210
7.4 BINOMIAL GLMM USING A FREQUENTIST APPROACH	212

7.4.1 Standardizing covariates	212
7.4.2 Binomial GLMM with random intercept zoo	213
7.4.3 Executing the GLMM using the <i>glmer</i> function	213
7.4.4 Overdispersion	215
7.4.5 Binomial GLMM with observation level random intercept	216
7.4.6 Visualization of results	220
7.5 BINOMIAL GLMM WITH RANDOM INTERCEPT ZOO IN JAGS	223
7.5.1 Data for JAGS	223
7.5.2 JAGS modelling code for a binomial GLMM.....	224
7.5.3 Results for the binomial GLMM.....	226
7.5.4 Overdispersion	226
7.6 BETA-BINOMIAL GLMM IN JAGS	228
7.6.1 The Beta distribution	228
7.6.2 From beta to beta-binomial distribution	229
7.6.3 JAGS code for beta-binomial GLMM	230
7.6.4 Beta-binomial GLMM results.....	231
7.6.5 Model validation of the beta-binomial GLMM	232
7.7 USING A BETA GLMM FOR PROPORTIONS	234
7.8 COMPARING ESTIMATED PARAMETERS FROM ALL MODELS.....	237
7.9 MODEL SELECTION FROM A FREQUENTIST POINT OF VIEW	239
7.10 MODEL SELECTION FROM A BAYESIAN POINT OF VIEW	241
7.10.1 Using the <i>DIC</i> , <i>AIC</i> and <i>BIC</i>	242
7.10.2 Inclusion probabilities.....	246
7.11 WHAT TO PRESENT IN A PAPER	246
REFERENCES	247
INDEX	251
BOOKS BY HIGHLAND STATISTICS	255
UPCOMING BOOKS IN 2013 AND 2014	256