Methods for analyzing longitudinal binary data with missing
responses

by

Sangook Kim

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements
for the degree of

Master of Science

in

Probability and Statistics

Carleton University

Ottawa, ON

© 2015

Sangook Kim

Abstract

In this thesis, we explore different estimation methods for longitudinal data with binary responses and drop-outs. We also study the effect of incorrectly specifying the dependence structure or the drop-out mechanism in longitudinal data. Although highly efficient, the traditional maximum likelihood (ML) method becomes complex when the number of responses increases, requiring intensive computation. Alternative methods such as generalized estimating equations (GEE) and weighted GEE had been proposed in the literature to overcome the limitation of the ML method. However, both estimators are known to be biased under non-ignorable drop-out mechanisms. The bivariate maximum pseudo likelihood is a pseudo likelihood method that takes into account the correlation between the current and baseline responses. Originally developed for non-monotone missing data, it was modified to be adapted for monotone drop-outs. We conduct a simulation study to assess the sensitivity of each method to model misspecifications in which a non-ignorable drop-out mechanism is our primary interest.

Acknowledgements

I thank Pr. Sinha for his guidance, encouragements and patience. He had exposed me to the wonderful world of Biostatistics and taught me the right way to conduct research, for which I am forever indebted. I am grateful for the support from Public Health Agency of Canada. In particular, I thank Dr. Rosamund Lewis, Dr. Barbara Law, Dr. Nooshin Ahmadipour and Julie Laflèche for their understanding and support. I also thank both Pr. Shirley Mills and Pr. Chul Gyu Park for their wonderful introduction to Statistics and for continuing to inspire me. I also thank Pr. Chul Gyu Park, Pr. Pierre-Jérôme Bergeron and Pr. Jiankang Zhang who accepted to be the committee members. Lastly, I thank my lovely wife Chin Myong, my daughter Youmi and our parents, my motivation and support, without whom I would have never made so far.

Dedicated to my wife Chin Myong with love
for our last two unforgettable years

Contents

List of Tables	vi
1 Introduction	1
2 Generalized Linear Model	5
2.1 Introduction	5
2.2 Definition and Properties of Generalized Linear Model	6
2.2.1 Definition of Generalized Linear Model	6
2.2.2 Some Properties of Generalized Linear Model	8
2.3 Estimation Methods	9
2.3.1 Maximum likelihood estimation	9
2.3.2 Maximum quasi-likelihood estimation	12
3 Longitudinal Data Analysis	16
3.1 Introduction	16
3.2 Regression models for longitudinal data	17
3.2.1 Notation	17
3.2.2 Regression models	17
3.2.3 Model for multivariate binary responses	19
3.3 Estimation methods	21
3.3.1 Independent Estimating Equations	22
3.3.2 Generalized Estimating Equations	24
3.3.3 Modified Generalized Estimating Equations	26

4	Missing Data and Estimation Methods	29
4.1	Introduction	29
4.2	Overview of missing data	30
4.2.1	Missing-data pattern	30
4.2.2	Missing-data mechanism	30
4.2.3	Class of methods for missing data	32
4.3	Estimation Methods for Binary Longitudinal Data with Drop-outs .	33
4.3.1	Maximum Likelihood	33
4.3.2	Generalized Estimating Equations	38
4.3.3	Weighted Generalized Estimating Equations	40
4.3.4	Pseudo-Maximum Likelihood Estimation	44
5	Simulation Study	50
5.1	Simulation settings	50
5.1.1	Responses model	50
5.1.2	Drop-out model	51
5.1.3	Estimation methods	54
5.1.4	Model misspecifications	57
5.2	Results	58
5.2.1	Results for Exchangeable Correlation	59
5.2.2	Results for AR(1) Correlation	93
5.3	Discussion	124
6	Conclusion and Future Study	127
	Bibliography	129

List of Tables

5.1	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), exchangeable correlation ($\alpha = 0.2$) and MCAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	61
5.2	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.2$) and MCAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	62
5.3	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), exchangeable correlation ($\alpha = 0.4$) and MCAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	63
5.4	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.4$) and MCAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	64

5.5	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), exchangeable correlation ($\alpha = 0.2$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 2, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	67
5.6	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.2$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 2, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	68
5.7	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), exchangeable correlation ($\alpha = 0.4$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 2, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	69
5.8	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.4$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 2, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	70
5.9	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), exchangeable correlation ($\alpha = 0.2$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	71

5.10	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.2$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	72
5.11	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), exchangeable correlation ($\alpha = 0.4$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	73
5.12	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.4$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	74
5.13	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), exchangeable correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	77
5.14	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	78

5.15	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), exchangeable correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	79
5.16	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	80
5.17	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), exchangeable correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	81
5.18	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	82
5.19	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), exchangeable correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	83

5.20	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	84
5.21	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), exchangeable correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	85
5.22	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	86
5.23	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), exchangeable correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	87
5.24	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	88

5.25	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), exchangeable correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	89
5.26	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	90
5.27	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), exchangeable correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	91
5.28	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	92
5.29	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), AR(1) correlation ($\alpha = 0.2$) and MCAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	94

5.30	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), AR(1) correlation ($\alpha = 0.2$) and MCAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	95
5.31	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), AR(1) correlation ($\alpha = 0.4$) and MCAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	96
5.32	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), AR(1) correlation ($\alpha = 0.4$) and MCAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	97
5.33	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), AR(1) correlation ($\alpha = 0.2$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 2, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	99
5.34	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), AR(1) correlation ($\alpha = 0.2$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 2, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	100

5.35	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), AR(1) correlation ($\alpha = 0.4$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 2, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	101
5.36	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), AR(1) correlation ($\alpha = 0.4$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 2, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	102
5.37	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), AR(1) correlation ($\alpha = 0.2$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	103
5.38	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), AR(1) correlation ($\alpha = 0.2$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	104
5.39	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), AR(1) correlation ($\alpha = 0.4$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	105

5.40	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), AR(1) correlation ($\alpha = 0.2$) and MAR drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	106
5.41	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), AR(1) correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	108
5.42	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), AR(1) correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	109
5.43	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), AR(1) correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	110
5.44	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), AR(1) correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	111

5.45	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), AR(1) correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	112
5.46	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), AR(1) correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	113
5.47	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), AR(1) correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	114
5.48	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), AR(1) correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	115
5.49	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), AR(1) correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	116

5.50	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), AR(1) correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	117
5.51	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), AR(1) correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	118
5.52	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), AR(1) correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	119
5.53	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), AR(1) correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	120
5.54	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), AR(1) correlation ($\alpha = 0.2$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	121

5.55	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 3 responses ($T = 3$), AR(1) correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	122
5.56	Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), AR(1) correlation ($\alpha = 0.4$) and NI drop-outs $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$	123

*

Chapter 1

Introduction

In biological experiments, often measurements are collected from a random sample of individuals at a single time point. Such experiments are called cross-sectional because the measurements taken at a single time point can be considered as a snapshot of the state of the random sample. Naturally, rather than a snapshot, one might be more interested in incorporating the time into the analysis by taking several measurements at different time points. Longitudinal studies consist of taking measurements repeatedly from an individual by following the participants at different times. As a practical example, consider the Indonesian Children's Health Study (ICHS), in which the presence of some respiratory infectious diseases and the vitamin A deficiency are collected from participants over time, to investigate the relation between the infection and the vitamin intake (Diggle et al., 2002).

For classical statistical models, incorporating the time component is not a trivial step because the measurements for an individual at different time points can be associated in a complex manner. An obvious consequence of ignoring the inherent association between the measurements within individuals and carrying on with a standard statistical analysis is incorrect inference in the form of biased and inefficient estimators of model parameters. It is, therefore, necessary for a model suitable for longitudinal data to take into account the correlation structure among the measurements within individuals. The problem is further complicated by the

fact that the parameters for the association between responses within individuals quickly proliferate with the increase in the number of time points. In general, the primary interest of an experiment is in the relation between covariates and responses and not in the association between the responses within individuals, which can be considered as ‘nuisance’ parameters. In the presence of a large number of parameters, traditional estimation methods such as the maximum likelihood becomes unusable because the method requires first and second-order derivatives which become intractable as the number of parameters increases. Therefore, in addition to considering the association structure of measurements within individuals, a good estimation method for longitudinal data requires a way to treat nuisance parameters differently from the parameters of main interests in order to avoid an overwhelming number of parameters to estimate simultaneously.

The generalized estimating equations (GEE) method proposed by Liang and Zeger (1986) is commonly used to estimate regression parameters for longitudinal data. The main feature of the GEE method is that one can specify the assumption for the correlation structure between responses within individuals in the form of ‘working’ correlation and with large sample size, the estimator is unbiased even when the correlation structure is misspecified. Its main advantage over the maximum likelihood method is that the estimation procedure for the GEE method is unaffected by the number of correlation parameters as it estimates the association between measurements within individuals separately from regression parameters. Due to the advantage, it is free of the computational burden and does not require complicated numerical optimization unlike the maximum likelihood method when there is a large number of parameters to estimate.

A greater problem not limited to longitudinal studies is the missing-data problem. In practice, measurements are often missing, complicating the estimating procedures. In longitudinal studies, it is common to have participants dropping out from the study and never returning back. To discern the impact of the missing data on estimation, suppose the drop-outs are random. The sample obtained

after removing the participants who dropped out from the study is still a random sub-sample of the original sample because the drop-outs occurred randomly and so carrying on with the statistical analysis after removing the drop-outs is still a valid approach. However, such a missing-data pattern is unlikely to occur in practice and it is not difficult to imagine other scenarios in which the participants' drop-outs are caused by their state prior to the drop-out or the unobserved outcomes after the drop-out. For example, in a clinical trial, a participant might drop out from the study due to a sudden deterioration in health at the time of drop-out caused by the treatment, which cannot be recorded due to the participant's drop-out. For such drop-outs, ignoring the individuals who dropped out from the study is a simplistic approach resulting in incorrect inference because the drop-outs mask the effect of the treatment on the participants' health. A sensible estimation method for longitudinal data with drop-outs should take into consideration the underlying drop-out mechanism so that one does not overestimate or underestimate the treatment effect.

The missing-data problem had been formerly treated by Rubin (1976). However, Rubin's primary focus was on continuous measurements. In biological experiments or social studies, responses are often discrete, in which case statistical analyses at multivariate level become difficult due to the lack of a proper multivariate distribution for discrete data. In this thesis, we explore different estimation methods for longitudinal data with binary responses and drop-outs. In particular, we are interested in drop-outs which are not ignorable, that is the underlying drop-out mechanism depends on unobserved outcomes following the participants' drop-outs. We examine GEE, weighted GEE, maximum likelihood and bivariate maximum pseudo likelihood and apply the methods in a simulation study to evaluate the empirical behaviours such as the empirical bias and efficiency of an estimator and to outline each method's sensitivity to different types of drop-out mechanisms and to misspecifications of the correlation structure among responses within individuals.

We start by reviewing the generalized linear model in Chapter 2, the building block of models for longitudinal data, and the estimation methods for them. In Chapter 3, we review models for analyzing longitudinal data including commonly used models for multivariate binary responses and the estimation methods including GEE. Missing-data problems are reviewed in Chapter 4 in which we examine the estimation methods such as weighted GEE, maximum likelihood and bivariate maximum pseudo likelihood. Results from the simulation study are summarized in Chapter 5 in which we elaborate on the sensitivity of each estimation method to model misspecifications. We finally state our conclusions in Chapter 6.

Chapter 2

Generalized Linear Model

2.1 Introduction

In biological science, the outcome of an experiment is often collected as a binary variable. A typical example is a clinical trial in which the effectiveness of a novel drug over a placebo is measured by patients' successful improvement following the treatments. Such data cannot be modeled by classical regression as the model requires the response to be continuous. Given its limitation, several models for binary responses were developed. An earlier model is probit regression model, in which the mean of a binary response is projected onto a real space using a log-normal transformation.

Nelder and Wedderburn (1972) recognized that the same principle could be applied to other non-linear models and proposed the Generalized Linear Model (GLM). As shown later in the chapter, GLM is, in fact, a generalization of classical linear regression and unifies many non-linear models with special cases such as logit and probit regression for binary responses, loglinear and multinomial regression for count data and other regression models for survival analysis. GLM is also the building block of models for longitudinal data, which are the centre of our interests and will be reviewed in Chapter 3.

In this chapter, an overview of GLM is provided, in order to build the context

for longitudinal data analysis in Chapter 3. The model will first be defined in section 2.2 along with some interesting properties. Then the estimation methods for GLM are reviewed in section 2.3 which include the maximum likelihood (ML) and maximum quasi-likelihood methods.

2.2 Definition and Properties of Generalized Linear Model

2.2.1 Definition of Generalized Linear Model

For notation, suppose n individuals constitute a random sample. The design matrix $\mathbf{X}$ with p covariates and response vector $\mathbf{Y}$ can be formed as follows:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \quad \text{and} \quad \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Additionally, the vector of p regression parameters is denoted by $\boldsymbol{\beta}$. For the i^{th} observation, the response and covariates are denoted by y_i and $\mathbf{x}_i$, respectively.

As in classical regression, GLM works under the assumption that the observations in a sample are independent from each other and the response Y_i has a constant variance σ^2 . GLM consists of the following three components:

- Random component $\mathbf{Y}$: in addition to the independence of the response variable with constant variance, GLM requires the random components Y_i 's to follow a density function in the exponential family defined by:

$$f_{Y_i}(y_i|\theta_i, \phi) = \exp \left\{ \frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi) \right\}, \quad (2.1)$$

where $a(\cdot)$, $b(\cdot)$ and $c(\cdot)$ are some known functions and θ_i and ϕ are some parameters.

-
- Systematic component $\boldsymbol{\eta}$: the linear function of the covariates is the systematic component in the regression model

$$\eta_i = \sum_{j=1}^p x_{ij} \beta_{ij}.$$

- Link function $g(\cdot)$: the link function is a function which transforms the mean response μ_i and links it to the systematic component:

$$g(\mu_i) = \eta_i.$$

To illustrate the model, suppose Y_i is a binary response with a Bernoulli distribution, that is

$$Pr(Y_i = y_i) = \mu_i^{y_i} (1 - \mu_i)^{1-y_i} \quad (2.2)$$

Bernoulli density can be rearranged to take the form (2.1) as

$$\begin{aligned} Pr(Y_i = y_i) &= \mu_i^{y_i} (1 - \mu_i)^{1-y_i} \\ &= \left(\frac{\mu_i}{1 - \mu_i} \right)^{y_i} (1 - \mu_i) \\ &= \exp \left\{ y_i \log \left(\frac{\mu_i}{1 - \mu_i} \right) + \log(1 - \mu_i) \right\}, \end{aligned}$$

which is in the form (2.1) with

$$a(\phi) = 1, \theta_i = \log \left(\frac{\mu_i}{1 - \mu_i} \right), b(\theta_i) = -\log(1 - \mu_i) \text{ and } c(y_i, \phi) = 0.$$

Here, the link function is chosen to be the logit function

$$\text{logit}(\mu_i) = \log \left(\frac{\mu_i}{1 - \mu_i} \right).$$

in which case, the GLM becomes a logistic regression model

$$\mu_i = E(Y_i|\mathbf{x}_i, \boldsymbol{\beta}) = Pr(Y_i = 1|\mathbf{x}_i, \boldsymbol{\beta}) = \frac{\exp\left(\sum_{j=1}^p x_{ij}\beta_{ij}\right)}{1 + \exp\left(\sum_{j=1}^p x_{ij}\beta_{ij}\right)}.$$

Note other common link functions for binary response are the log-normal function leading to probit regression model and complementary log-log function (Nelder and Wedderburn, 1972).

2.2.2 Some Properties of Generalized Linear Model

Consider the log-likelihood of Y_i following a GLM:

$$l_{Y_i}(\theta_i, \phi|y_i) = \log f_{Y_i}(y_i|\theta_i, \phi) = \frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi).$$

The mean and the variance of Y_i can be derived using the following equations:

$$E\left(\frac{\partial l_{Y_i}}{\partial \theta_i}\right) = 0, \tag{2.3}$$

$$E\left(\frac{\partial^2 l_{Y_i}}{\partial \theta_i^2}\right) + E\left(\frac{\partial l_{Y_i}}{\partial \theta_i}\right)^2 = 0. \tag{2.4}$$

Solving (2.3), the mean of Y_i is obtained

$$\begin{aligned} E\left(\frac{\partial l_{Y_i}}{\partial \theta_i}\right) &= E\left(\frac{y_i - b'(\theta_i)}{a(\phi)}\right) \\ &= \frac{\mu_i - b'(\theta_i)}{a(\phi)} = 0 \\ \Leftrightarrow E(Y_i) &= \mu_i = b'(\theta_i). \end{aligned} \tag{2.5}$$

And the variance of Y_i is derived by solving (2.4) as

$$\begin{aligned}
E\left(\frac{\partial^2 l_{Y_i}}{\partial \theta_i^2}\right) + E\left(\frac{\partial l_{Y_i}}{\partial \theta_i}\right)^2 &= -E\left(\frac{b''(\theta_i)}{a(\phi)}\right) + E\left(\frac{y_i - b'(\theta_i)}{a(\phi)}\right)^2 \\
&= -\frac{b''(\theta_i)}{a(\phi)} + \frac{\text{var}(Y_i)}{a(\phi)^2} = 0 \\
\Leftrightarrow \text{var}(Y_i) &= b''(\theta_i)a(\phi).
\end{aligned} \tag{2.6}$$

Equations (2.5) and (2.6) show that the mean and variance of a GLM depend respectively on the first and second-order derivatives of $b(\theta_i)$ with respect to θ_i . The variance is, therefore, a function of the mean, an important property, which is later used to build the maximum quasi-likelihood estimator in Section 2.3.2. Note the variance also depends on the function $a(\phi)$ in order to deal with the problem of overdispersion. ϕ is called the dispersion parameter.

2.3 Estimation Methods

2.3.1 Maximum likelihood estimation

From (2.1), the log-likelihood for n independent observations is

$$l = \sum_{i=1}^n \left[\frac{y_i \theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi) \right]$$

The maximum likelihood estimator (MLE) of the regression parameter β is obtained by maximizing the log-likelihood l , which is equivalent to finding the solution to the score equations

$$\frac{\partial l}{\partial \beta} = \mathbf{0}.$$

The score function is

$$\begin{aligned}
\frac{\partial l}{\partial \boldsymbol{\beta}} &= \sum_{i=1}^n \frac{\partial \left[\frac{y_i \theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi) \right]}{\partial \boldsymbol{\beta}} \\
&= \frac{1}{a(\phi)} \sum_{i=1}^n \left[y_i \frac{\partial \theta_i}{\partial \boldsymbol{\beta}} - \frac{\partial b(\theta_i)}{\partial \boldsymbol{\beta}} \right] \\
&= \frac{1}{a(\phi)} \sum_{i=1}^n \left[y_i \frac{\partial \theta_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \boldsymbol{\beta}} - \frac{\partial b(\theta_i)}{\partial \theta_i} \frac{\partial \theta_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \boldsymbol{\beta}} \right] && \text{(by chain rule)} \\
&= \frac{1}{a(\phi)} \sum_{i=1}^n \left[\left(y_i - \frac{\partial b(\theta_i)}{\partial \theta_i} \right) \frac{\partial \theta_i}{\partial \eta_i} \mathbf{x}_i \right] \\
&= \frac{1}{a(\phi)} \sum_{i=1}^n (y_i - \mu_i) \frac{\partial \theta_i}{\partial \eta_i} \mathbf{x}_i. && \text{(using (2.5))}
\end{aligned}$$

The score function can be written in the matrix form as

$$\frac{\partial l}{\partial \boldsymbol{\beta}} = \frac{1}{a(\phi)} \mathbf{X}' \boldsymbol{\Delta} (\mathbf{y} - \boldsymbol{\mu}) \quad (2.7)$$

where $\boldsymbol{\Delta}$ is a diagonal matrix, given by

$$\boldsymbol{\Delta} = \begin{bmatrix} \frac{\partial \theta_1}{\partial \eta_1} & 0 & \cdots & 0 \\ 0 & \frac{\partial \theta_2}{\partial \eta_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\partial \theta_n}{\partial \eta_n} \end{bmatrix}.$$

Under some regularity conditions (Casella and Berger, 1990), the MLE $\hat{\boldsymbol{\beta}}$ is such that $\hat{\boldsymbol{\beta}} \xrightarrow{P} \boldsymbol{\beta}$ and satisfies the asymptotic property

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \sim N_p(0, \mathbf{I}^{-1}(\boldsymbol{\beta})),$$

where $\mathbf{I}(\boldsymbol{\beta})$ is the Fisher information for $\boldsymbol{\beta}$.

The MLEs of the regression parameters $\boldsymbol{\beta}$ are generally computed iteratively

using numerical methods such as the Fisher's scoring method, given by

$$\boldsymbol{\beta}^{(m+1)} = \boldsymbol{\beta}^{(m)} + \mathbf{I} \left(\boldsymbol{\beta}^{(m)} \right)^{-1} \frac{\partial l}{\partial \boldsymbol{\beta}} \Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}^{(m)}} \quad (2.8)$$

for $m = 0, 1, 2, \dots$, where $\boldsymbol{\beta}^{(m)}$ are the estimators of $\boldsymbol{\beta}$ at m^{th} iteration and the right-hand side of the equation is evaluated at $\boldsymbol{\beta} = \boldsymbol{\beta}^{(m)}$. The iterations continue until convergence is reached. The converged values are the numerical approximation of the ML estimates.

It can be shown that the Fisher information matrix $\mathbf{I}$ is

$$\mathbf{I} = -E \left(\frac{\partial^2 l}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} \right) = \frac{1}{a(\phi)} \mathbf{X}' \boldsymbol{\Delta} \mathbf{A} \boldsymbol{\Delta} \mathbf{X} \quad (2.9)$$

where $\mathbf{A}$ is a diagonal matrix, given by

$$\mathbf{A} = \begin{bmatrix} b''(\theta_1) & 0 & \dots & 0 \\ 0 & b''(\theta_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & b''(\theta_n) \end{bmatrix}.$$

Using (2.7) and (2.9), the explicit Fisher's scoring algorithm is:

$$\boldsymbol{\beta}^{(m+1)} = \boldsymbol{\beta}^{(m)} + \{ \mathbf{X}' \boldsymbol{\Delta} \mathbf{A} \boldsymbol{\Delta} \mathbf{X} \}^{-1} \{ \mathbf{X}' \boldsymbol{\Delta} (\mathbf{y} - \boldsymbol{\mu}) \} \Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}^{(m)}}, \quad (2.10)$$

for $m = 0, 1, 2, \dots$.

The variance-covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$ of $\hat{\boldsymbol{\beta}}$ can be approximated by

$$\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}} = \mathbf{I}^{-1} \left(\hat{\boldsymbol{\beta}} \right) = \left(\frac{1}{a(\phi)} \mathbf{X}' \boldsymbol{\Delta} \mathbf{A} \boldsymbol{\Delta} \mathbf{X} \right)^{-1} \Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}}.$$

For binary responses in (2.2), let δ_{ii} be the i^{th} diagonal element of the matrix $\boldsymbol{\Delta}$.

Then δ_{ii} can be obtained using

$$\begin{aligned}
\delta_{ii} &= \frac{\partial \theta_i}{\partial \eta_i} \\
&= \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} && \text{(by chain rule)} \\
&= \left(\frac{\partial \mu_i}{\partial \theta_i} \right)^{-1} \left(\frac{\partial \eta_i}{\partial \mu_i} \right)^{-1} \\
&= \frac{1}{a(\phi)b''(\theta_i)} \left(\frac{\partial \log \left(\frac{\mu_i}{1-\mu_i} \right)}{\partial \mu_i} \right)^{-1} && \text{(using (2.6))} \\
&= \frac{\mu_i(1-\mu_i)}{a(\phi)b''(\theta_i)}.
\end{aligned}$$

The variance of Y_i under Bernoulli distribution is

$$b''(\theta_i) = \mu_i(1 - \mu_i),$$

so $\mathbf{A}$ can also be calculated using $b''(\theta_i)$. The matrices $\mathbf{\Delta}$ and $\mathbf{A}$ can then be substituted in (2.10) to compute the ML estimates.

2.3.2 Maximum quasi-likelihood estimation

The maximum quasi-likelihood method is an alternative to the ML and its advantage over the ML is that one does not need to assume any distribution for the response Y_i (Wedderburn, 1974). The method only requires that the mean and variance functions of Y_i are known *a priori*.

Consider a random variable Y_i following a GLM such that $E(Y_i) = \mu_i$ and $\text{var}(Y_i) = a(\phi)v(\mu_i)$ where $v(\mu_i)$ is the first-derivative of μ_i with respect to θ as in

(2.6). First, note

$$\begin{aligned}
\text{var} \left[\frac{\partial \log f_{Y_i}(y_i)}{\partial \mu_i} \right] &= \text{var} \left[\frac{\partial \log f_{Y_i}(y_i)}{\partial \theta_i} \frac{\partial \theta_i}{\partial \mu_i} \right] && \text{(by chain rule)} \\
&= \text{var} \left[\frac{\partial \log f_{Y_i}(y_i)}{\partial \theta_i} \right] \left(\frac{\partial \theta_i}{\partial \mu_i} \right)^2 \\
&= -E \left[\frac{\partial^2 \log f_{Y_i}(y_i)}{\partial \theta_i^2} \right] \left(\frac{\partial \theta_i}{\partial \mu_i} \right)^2 \\
&= \frac{v(\mu_i)}{a(\phi)} \left(\frac{\partial \mu_i}{\partial \theta_i} \right)^{-2} \\
&= \frac{v(\mu_i)}{a(\phi)} \frac{1}{v(\mu_i)^2} && \text{(using (2.5))} \\
&= \frac{1}{a(\phi)v(\mu_i)}. && (2.11)
\end{aligned}$$

Now define a quantity q_i such that

$$q_i = \frac{y_i - \mu_i}{a(\phi)v(\mu_i)} \quad (2.12)$$

and Q_i , its integral with respect to the mean μ_i , such that

$$Q_i = \int_{y_i}^t \frac{y_i - t}{a(\phi)v(t)} dt. \quad (2.13)$$

Clearly,

$$E \left(\frac{\partial Q_i}{\partial \mu_i} \right) = E(q_i) = 0 \quad (2.14)$$

and

$$\begin{aligned}
-E \left(\frac{\partial^2 Q_i}{\partial \mu_i^2} \right) &= -E \left(\frac{\partial q_i}{\partial \mu_i} \right) \\
&= E \left[\frac{v(\mu_i) + (y_i - \mu_i)v'(\mu_i)}{a(\phi)v(\mu_i)^2} \right] \\
&= \frac{1}{a(\phi)v(\mu_i)}. && (2.15)
\end{aligned}$$

Note (2.14) is equivalent to (2.11), therefore, by (2.14) and (2.15), q_i satisfies the definition of a score function. The quantity (2.12) is called the log quasi-likelihood

for y_i .

The log quasi-likelihood for $\mathbf{y}$ is $\sum_{i=1}^n Q_i$ and analogous to ML estimation, we derive the maximum quasi-likelihood estimator, which maximizes the quasi-likelihood. To achieve the maximization, we first derive the score function called quasi-score function

$$\begin{aligned}
\sum_{i=1}^n \frac{\partial Q_i}{\partial \boldsymbol{\beta}} &= \sum_{i=1}^n q_i \frac{\partial \mu_i}{\partial \boldsymbol{\beta}} \\
&= \sum_{i=1}^n \frac{y_i - \mu_i}{a(\phi)v(\mu_i)} \frac{\partial \mu_i}{\partial \boldsymbol{\beta}} \\
&= \frac{1}{a(\phi)} \sum_{i=1}^n \frac{y_i - \mu_i}{v(\mu_i)} \frac{\partial \mu_i}{\partial \theta_i} \frac{\partial \theta_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \boldsymbol{\beta}} \quad (\text{by chain rule}) \\
&= \frac{1}{a(\phi)} \sum_{i=1}^n (y_i - \mu_i) \frac{\partial \theta_i}{\partial \eta_i} \mathbf{x}_i' \quad (\text{using (2.5) and (2.6)})
\end{aligned}$$

or in matrix notation

$$\frac{1}{a(\phi)} \mathbf{X}' \boldsymbol{\Delta} (\mathbf{y} - \boldsymbol{\mu}) = \mathbf{0} \quad (2.16)$$

where

$$\boldsymbol{\Delta} = \begin{bmatrix} \frac{\partial \theta_1}{\partial \eta_1} & 0 & \cdots & 0 \\ 0 & \frac{\partial \theta_2}{\partial \eta_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\partial \theta_n}{\partial \eta_n} \end{bmatrix}.$$

which is surprisingly identical to (2.7), the score function of the ML, therefore under the current settings, the maximum quasi-likelihood and the ML produce the same estimates.

Analogous to the ML, (2.16) is maximized using Fisher's scoring algorithm as highlighted in the previous section. It can be shown that Fisher's information matrix is

$$\mathbf{I} = -E \left(\frac{\partial^2 l}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} \right) = \frac{1}{a(\phi)} \mathbf{X}' \boldsymbol{\Delta} \mathbf{A} \boldsymbol{\Delta} \mathbf{X}$$

where

$$\mathbf{A} = \begin{bmatrix} b''(\theta_1) & 0 & \cdots & 0 \\ 0 & b''(\theta_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b''(\theta_n) \end{bmatrix}.$$

Then the maximum quasi-likelihood estimates are computed using

$$\boldsymbol{\beta}^{(m+1)} = \boldsymbol{\beta}^{(m)} + \{\mathbf{X}'\boldsymbol{\Delta}\mathbf{A}\boldsymbol{\Delta}\mathbf{X}\}^{-1} \{\mathbf{X}'\boldsymbol{\Delta}(\mathbf{y} - \boldsymbol{\mu})\} \Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}^{(m)}}$$

The variance-covariance matrix of the estimates are calculated at the convergence using

$$\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}} = I^{-1}(\hat{\boldsymbol{\beta}}) = \left(\frac{1}{a(\phi)} \mathbf{X}'\boldsymbol{\Delta}\mathbf{A}\boldsymbol{\Delta}\mathbf{X} \right)^{-1} \Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}}.$$

In Chapter 3, we illustrate the use of GLM to build regression models for longitudinal data and present some estimation methods.

Chapter 3

Longitudinal Data Analysis

3.1 Introduction

Unlike cross-sectional studies, in which a single time point is considered, longitudinal studies, also referred to as panel studies by sociologists and economists, are studies in which multiple responses from participants are collected over time. One obvious advantage of such studies over cross-sectional studies is its ability to capture the change in individuals' responses over time, defined as the age effect. An additional advantage is to compare groups of individuals with different baseline levels, called cohort effects.

When observations are independent by the study design, classical linear or generalized regression models can be used to analyze cross-sectional data. However, in longitudinal data, the responses within individuals tend to be inter-correlated. Therefore, classical regression can lead to an incorrect inference, especially when the responses within individuals are highly correlated. So it is important to use a model which takes into account the correlation within individuals. In this chapter, we will review regression models for longitudinal data and models for joint responses in Section 3.2 followed by estimation methods in Section 3.3.

3.2 Regression models for longitudinal data

3.2.1 Notation

Suppose T responses were collected from a random sample of N individuals. Let $\mathbf{x}_{it} = (x_{1it}, \dots, x_{pit})'$ denote p -covariates for the i^{th} individual at time t and Y_{it} the corresponding response taken at time t with $t = 1, \dots, T$. Additionally, suppose $\mu_{it} = E(Y_{it})$ denotes the mean response, $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{iT})'$ denote a vector of T responses and $\boldsymbol{\mu}_i = (\mu_{i1}, \dots, \mu_{iT})'$, its corresponding mean vector. Although the covariates can change over time, for simplicity, we assume throughout the review that the covariates are taken prior to the beginning of the longitudinal study and do not change in the course of the study *i.e.* $\mathbf{x}_i = \mathbf{x}_{it}$ for all $t = 1, \dots, T$. For covariates which change over time, one can replace $\mathbf{x}_i$ by $\mathbf{x}_{it}$.

3.2.2 Regression models

When the association between responses within an individual is non-negligible, classical regression is an inappropriate method to draw an inference because the model is based on the assumption that the observations are independent. In practice, the responses within individuals are often inter-correlated and in such a case, classical regression would lead to not only an incorrect inference, but also inefficient estimates (Diggle et al, 2002). An appropriate regression model for longitudinal data must take into account the correlation within individuals. In this section, three regression models for longitudinal data such as marginal model, random effects model and transition model are reviewed. Marginal regression models are the primary focus here and will be reviewed in details. The details for other methods can be found in Diggle et al. (2002).

Marginal regression model makes assumptions about the correlation between responses, and models the expectation $E(Y_{ij})$ as a function of the covariates $\mathbf{x}_i$. The expectation $E(Y_{ij}|\mathbf{x}_i)$ is called marginal expectation because it is the average over a group sharing the same value of $\mathbf{x}_i$. Such a model is suited for studies in

which the population average is the primary interest. The following assumptions are required for marginal regression models:

1. The expectation $E(Y_{ij}) = \mu_{ij}$ is a function of covariates $\mathbf{x}_i$:

$$g(\mu_{ij}) = \mathbf{x}_i' \boldsymbol{\beta},$$

where $g(\cdot)$ is a known link function.

2. The variance $\text{var}(Y_{ij})$ is a function of the marginal mean $E(Y_{ij}|\mathbf{x}_i, \boldsymbol{\beta}) = \mu_{ij}$:

$$\text{var}(Y_{ij}) = v(\mu_{ij})a(\phi),$$

where a and v are some known functions and ϕ a dispersion parameter.

3. Responses Y_{ij} and Y_{ik} are correlated by a given correlation parameter $\alpha_{i_{jk}}$ such that:

$$\text{Corr}(Y_{ij}, Y_{ik}) = \rho(\mu_{ij}, \mu_{ik}|\alpha_{i_{jk}}),$$

where ρ is some function.

The GLM is naturally suited for the marginal regression model as it satisfies the first two assumptions by definition. The correlation between responses is a new concept and is elaborated in Section 3.3. It is a source of problem for estimation methods.

Random effects model for longitudinal data assumes each observation is heterogeneous and has its specific regression coefficients $\boldsymbol{\beta}_i$, which are considered as random variables with a density. Furthermore, the model assumes that correlation between responses within individuals arises due to unobserved latent variables $\mathbf{U}_i$ specific to each individual, which are independent and has a common multivariate distribution. Conditioning on the random effects $\mathbf{U}_i$, $Y_{i1}, \dots, Y_{iT}$ are independent and follow a density $f(Y_{it}|\mathbf{U}_i)$ in the exponential family. The random effects model

is:

$$h\{E(Y_{ij}|\mathbf{U}_i)\} = \mathbf{x}_i'\boldsymbol{\beta} + \mathbf{d}_i'\mathbf{U}_i$$

where h is the link function and $\mathbf{d}_i$ is a subset of the covariates $\mathbf{x}_i$.

Random effects model is useful when the interest is in making an inference about individuals rather than the general populations.

Transition model assumes that the response Y_{ij} depends on the previous responses $Y_{i1}, \dots, Y_{i,j-1}$ and the covariates $\mathbf{x}_i$, and models the conditional expectation

$$E(Y_{ij}|Y_{i1}, \dots, Y_{i,j-1}, \mathbf{x}_i)$$

3.2.3 Model for multivariate binary responses

The log-linear model is widely used for multivariate binary variables (Bishop et al, 1975), and is given by

$$P(\mathbf{Y}_i = \mathbf{y}_i) = c(\boldsymbol{\theta}) \exp \left[\sum_{t=1}^T \theta_t^{(1)} y_{it} + \sum_{t_1 < t_2} \theta_{t_1 t_2}^{(2)} y_{it_1} y_{it_2} + \dots + \theta_{1, \dots, T}^{(T)} y_{i1} \dots y_{iT} \right], \quad (3.1)$$

where $c(\boldsymbol{\theta})$ is a normalizing constant and $\boldsymbol{\theta}$ is a $(2^n - 1)$ vector of canonical parameters. Model (3.1) is, in fact, a saturated log-linear model.

Bahadur model is a log-linear model, where the joint distribution of responses is expressed by marginal means, pairwise correlations and higher moments of standardized responses (Bahadur, 1961). The model is given by

$$P(Y_{i1} = y_{i1}, \dots, Y_{iT} = y_{iT}) = \prod_{t=1}^T \mu_{it}^{y_{it}} (1 - \mu_{it})^{(1-y_{it})} \left[1 + \sum_{t_1 < t_2} \alpha_{it_1 t_2} z_{it_1} z_{it_2} + \sum_{t_1 < t_2 < t_3} \alpha_{it_1 t_2 t_3} z_{it_1} z_{it_2} z_{it_3} + \dots + \alpha_{i1, \dots, T} z_{i1} \dots z_{iT} \right],$$

where

$$\begin{aligned}
\mu_{it} &= E(Y_{it}|\mathbf{x}_i, \boldsymbol{\beta}) = Pr(Y_{it} = 1|\mathbf{x}_i, \boldsymbol{\beta}), \\
Z_{it} &= \frac{Y_{it} - \mu_{it}}{\sqrt{\mu_{it}(1 - \mu_{it})}}, \\
\alpha_{i_{t_1 t_2}} &= E(z_{i1}, z_{i2}) = Corr(y_{i1}, y_{i2}), \\
&\dots \\
\alpha_{i_{1, \dots, T}} &= E(z_{i1}, \dots, z_{iT}).
\end{aligned}$$

Bahadur's representation is appealing because it uses the familiar marginal probabilities and correlations expressing the association between responses. Although an attractive feature, the use of correlation complicates the constraints on the marginal mean (Prentice, 1988). Consider a simple bivariate case with responses Y_i and Y_j . By definition, the correlation between Y_i and Y_j , denoted by α , must satisfy the inequalities

$$-1 \leq \alpha = \frac{Pr(Y_i = 1, Y_j = 1) - \mu_i \mu_j}{\sqrt{\mu_i(1 - \mu_i)\mu_j(1 - \mu_j)}} \leq 1 \quad (3.2)$$

As the joint probability $Pr(Y_i = 1, Y_j = 1)$ must be a non-negative number less than 1, to satisfy the constraint (3.2), the joint probability is constrained by the marginal means

$$\max(0, \mu_1 + \mu_2 - 1) < Pr(Y_1 = 1, Y_2 = 1) < \min(\mu_1, \mu_2)$$

As we will review in the following section, the constraints on the joint probability becomes problematic in the likelihood-based estimation procedures because some values of $\boldsymbol{\alpha}$ may lead to a negative probability or probability exceeding 1. The details are discussed later in the simulation study in Chapter 5.

An alternative model to remove the constraints on the association parameters is to use marginal odds ratio as measures of association between responses (Molen-

berghs and Lesaffre, 1994). For bivariate responses, the pairwise marginal odds ratio is

$$OR(Y_1, Y_2) = \frac{Pr(Y_i = 1|Y_j = 1)Pr(Y_i = 0|Y_j = 0)}{Pr(Y_i = 1|Y_j = 0)Pr(Y_i = 0|Y_j = 1)}$$

Odds ratio takes values in $(0, \infty)$ and is not constrained on a logarithmic scale. The joint distribution of responses can be formulated using marginal means, marginal odds ratio and contrasts of odds ratio in the form

$$P(Y_{i1} = y_{i1}, \dots, Y_{iT} = y_{iT}) = \sum_{y_{i3}, \dots, y_{iT}=0,1} (-1)^{b(\mathbf{y}_i)} \log OR(y_{i1}, y_{i2}|y_{i3}, \dots, y_{iT}),$$

where

$$b(\mathbf{y}_i) = \sum_{t=3}^T y_{it} + T - 2$$

A disadvantage of using marginal odds ratio is that the closed form of likelihood cannot be derived explicitly unless the number of responses is small, so the likelihood-based methods cannot be applied easily unlike Bahadur model (Liang et al., 1992).

3.3 Estimation methods

For marginal models, when the responses are continuous and follow multivariate Gaussian, the first two moments fully determines the likelihood. Therefore, the estimation can be performed using the usual ML method. However, for binary responses, as reviewed in Section 3.2, higher-order moments are required for the likelihood and no sensible, simple method exists (Diggle et al., 2002). Even with assumptions about higher-order moments, another problem of estimation is the number of association parameters: as the number of responses gets larger, the number of association parameters to estimate can quickly proliferate. For a marginal odds ratio model with $\frac{T(T-1)}{2}$ pairwise odds ratios, the number of odds ratios to estimate clearly grows exponentially when T increases. In general, the goal of marginal regression analysis is to estimate the regression parameters, leav-

ing the association parameters as nuisance. In such cases, estimation methods such as generalized estimating equations method can be a better approach than the ML.

The generalized estimating equations method is, in fact, a multivariate analog of the maximum quasi-likelihood method described in Chapter 2. As shown later in the section, the generalized estimating equations method differs from the maximum quasi-likelihood method in terms of assumptions about pairwise correlations between responses. Unlike the ML estimation, the generalized estimating equations method does not model the joint distribution, freeing from the computational burden, but still gives consistent estimates even when the correlation is misspecified (Liang and Zeger, 1986). To understand the concept, we begin by illustrating the independent estimating equations method.

3.3.1 Independent Estimating Equations

Under the naive assumption that the association between responses within individuals are independent, the association parameters are eliminated in independent estimating equations (IEE). Such assumption simply leads to T -variate log-likelihood equations, as described in section 2.3.1. Evaluating the first-order derivatives of the log-likelihood with respect to $\boldsymbol{\beta}$, IEE are given by the following score equations

$$U(\boldsymbol{\beta}) = \sum_{i=1}^N \mathbf{X}_i' \boldsymbol{\Delta}_i (\mathbf{Y}_i - \boldsymbol{\mu}_i) = \mathbf{0},$$

where

$$\boldsymbol{\Delta}_i = \begin{bmatrix} \frac{\partial \theta_1}{\partial \eta_1} & 0 & \dots & 0 \\ 0 & \frac{\partial \theta_2}{\partial \eta_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\partial \theta_T}{\partial \eta_T} \end{bmatrix}.$$

Using (2.7), Fisher information matrix is

$$\mathbf{I} = \sum_{i=1}^N \mathbf{X}_i' \Delta_i \mathbf{A}_i \Delta_i \mathbf{X}_i,$$

where

$$\mathbf{A}_i = \begin{bmatrix} b''(\theta_1) & 0 & \cdots & 0 \\ 0 & b''(\theta_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b''(\theta_T) \end{bmatrix}.$$

Under some regularity conditions, the solutions to IEE, $\hat{\boldsymbol{\beta}}$, are consistent and follow an asymptotic normal distribution:

$$\sqrt{N}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{P} \mathbf{N}(\mathbf{0}, \mathbf{I}^{-1}(\boldsymbol{\beta}))$$

However, the assumption of independent responses might not be true and consequently, might affect the estimates. To adjust for the possible misspecification of association between responses, Liang and Zeger (1986) applied a robust sandwich estimator proposed by Huber (1967) and White (1980) and obtained the variance-covariance matrix

$$\mathbf{V}_{\boldsymbol{\beta}} = \lim_{N \rightarrow \infty} N \mathbf{B}_{\boldsymbol{\beta}}^{-1} \mathbf{H}_{\boldsymbol{\beta}} \mathbf{B}_{\boldsymbol{\beta}}^{-1},$$

where

$$\begin{aligned} \mathbf{B}_{\boldsymbol{\beta}} &= \mathbf{I}(\boldsymbol{\beta}) = \sum_{i=1}^N \mathbf{X}_i' \Delta_i \mathbf{A}_i \Delta_i \mathbf{X}_i, \\ \mathbf{H}_{\boldsymbol{\beta}} &= \sum_{i=1}^N \mathbf{X}_i' \Delta_i \mathbf{A}_i (\mathbf{Y}_i - \boldsymbol{\mu}_i) (\mathbf{Y}_i - \boldsymbol{\mu}_i)' \mathbf{A}_i \Delta_i \mathbf{X}_i. \end{aligned}$$

The estimates are computed iteratively using the Fisher's scoring algorithm:

$$\boldsymbol{\beta}^{(m+1)} = \boldsymbol{\beta}^{(m)} + \left(\sum_{i=1}^N \mathbf{X}_i' \Delta_i \mathbf{A}_i \Delta_i \mathbf{X}_i \right)^{-1} \left(\sum_{i=1}^N \mathbf{X}_i' \Delta_i \mathbf{A}_i (\mathbf{Y}_i - \boldsymbol{\mu}_i) \right) \Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}^{(m)}}$$

for $m = 0, 1, 2, \dots$.

The terms at the right-hand side of the equations are evaluated at $\beta = \beta^{(m)}$ obtained at the m^{th} iteration, and the iterations stop once a convergence is reached. Sample variance-covariance matrix can be calculated with the values at convergence $\hat{\beta}$, using

$$\hat{\Sigma}_{\hat{\beta}} = B_{\hat{\beta}}^{-1} H_{\hat{\beta}} B_{\hat{\beta}}^{-1}.$$

The assumption of independent responses is unrealistic in practice, making IEE unfavourable to use. The next methods take into consideration the association between responses by making assumptions about the correlation between responses.

3.3.2 Generalized Estimating Equations

Generalized estimating equations method takes into account the correlation between responses within observations by incorporating a ‘working’ correlation matrix $R(\alpha)$ into the estimating equations. The ‘working’ correlation matrix $R(\alpha)$ allows one to specify the assumption about the association between responses.

Then generalized estimating equations (GEE) are

$$U_{\beta}(\beta, \alpha) = \sum_{i=1}^N D_i' V_i^{-1} (y_i - \mu_i) = 0, \quad (3.3)$$

where

$$D_i = \frac{\partial \mu_i}{\partial \beta} = A_i \Delta_i X_i \text{ and } V_i = A_i^{\frac{1}{2}} R(\alpha) A_i^{\frac{1}{2}}.$$

GEE resembles the quasi-likelihood equations (2.16) except that the ‘working’ variance matrix V_i is a function not only of β but also α . Some typical ‘working’ correlations are given below.

- When an independent ‘working’ correlation is assumed, *i.e.* $R(\alpha) = I$, GEE are reduced to IEE.
- A correlation is exchangeable if $Corr(Y_{it_1}, Y_{it_2}) = \alpha$ for all $t_1 \neq t_2$. The

‘working’ correlation matrix then becomes

$$\mathbf{R} = \begin{bmatrix} 1 & \alpha & \alpha & \cdots & \alpha \\ \alpha & 1 & \alpha & \cdots & \alpha \\ \vdots & & & \ddots & \\ \alpha & \alpha & \alpha & \cdots & 1 \end{bmatrix}.$$

Following Liang and Zeger (1986), the correlation parameter α may be estimated using Pearson residuals:

$$\hat{\alpha} = \sum_{i=1}^N \sum_{t_1 < t_2} \frac{\hat{r}_{it_1} \hat{r}_{it_2}}{\sum_{i=1}^N \frac{1}{2} T(T-1) - p},$$

where p is the number of covariates and $\hat{r}_{it}$ is the Pearson residual

$$\hat{r}_{it} = \frac{y_{it} - \hat{\mu}_{it}}{\sqrt{\hat{\mu}_{it}(1 - \hat{\mu}_{it})}}$$

- A first-order autoregressive correlation, $AR(1)$, is such that $Corr(Y_{it_1}, Y_{it_2}) = \alpha^{|t_2 - t_1|}$. The corresponding ‘working’ correlation matrix takes the form

$$\mathbf{R} = \begin{bmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^n \\ \alpha & 1 & \alpha & \cdots & \alpha^{(n-1)} \\ & \vdots & & \ddots & \\ \alpha & \alpha^2 & \alpha^3 & \cdots & 1 \end{bmatrix}.$$

Under such assumption,

$$E(r_{it_1} r_{it_2}) = Corr(y_{it_1}, y_{it_2}) = \alpha^{|t_2 - t_1|}.$$

The coefficient α may be estimated by the slope of the regression of $\log(\hat{r}_{it_1} \hat{r}_{it_2})$ on $|t_2 - t_1|$.

Under some regularity conditions, the solutions to GEE, $\hat{\beta}$, are consistent provided that the estimates of α are consistent, and are asymptotically normal. Similar to the IEE method, a robust sandwich estimator is used to derive the asymptotic variance-covariance matrix of $\hat{\beta}$, which is given by

$$\Sigma_{\beta} = \lim_{N \rightarrow \infty} N \mathbf{B}_{\beta}^{-1} \mathbf{H}_{\beta} \mathbf{B}_{\beta}^{-1},$$

where

$$\mathbf{B}_{\beta} = \sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1} \mathbf{D}_i, \quad (3.4)$$

$$\mathbf{H}_{\beta} = \sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1} (\mathbf{Y}_i - \boldsymbol{\mu}_i) (\mathbf{Y}_i - \boldsymbol{\mu}_i)' \mathbf{V}_i^{-1} \mathbf{D}_i. \quad (3.5)$$

The estimates of the regression parameters are calculated iteratively using the Fisher's scoring algorithm:

$$\boldsymbol{\beta}^{(m+1)} = \boldsymbol{\beta}^{(m)} + \left(\sum_{i=1}^K \mathbf{D}_i' \mathbf{V}_i^{-1} \mathbf{D}_i \right)^{-1} \left(\sum_{i=1}^K \mathbf{D}_i' \mathbf{V}_i^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i) \right) \Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}^{(m)}}$$

The estimates of the variance-covariance matrix are

$$\hat{\Sigma}_{\hat{\beta}} = \mathbf{B}_{\hat{\beta}}^{-1} \mathbf{H}_{\hat{\beta}} \mathbf{B}_{\hat{\beta}}^{-1}$$

evaluated at $\hat{\beta}$.

3.3.3 Modified Generalized Estimating Equations

Prentice (1988) extended the GEE method described in 3.3.2 by including a second set of GEE to estimate α instead of using Pearson statistic. The idea is to treat the “sample correlation” Z_{ij}

$$Z_{ij} = \frac{(Y_i - \mu_i)(Y_j - \mu_j)}{\sqrt{\mu_i(1 - \mu_i)\mu_j(1 - \mu_j)}}$$

as random variables and estimate $\boldsymbol{\alpha}$ using a second set of GEE. Let δ_{ij} and w_{ij} denote respectively the mean and variance of Z_{ij} such as

$$\delta_{ij} = \alpha_{ij}$$

and

$$w_{ij} = 1 + \frac{(1 - 2\mu_i)(1 - 2\mu_j)}{\sqrt{\mu_i(1 - \mu_i)\mu_j(1 - \mu_j)}}\delta_{ij} - \delta_{ij}^2.$$

Then the GEE for $\boldsymbol{\alpha}$ are

$$\mathbf{U}_{\boldsymbol{\alpha}}(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \sum_{i=1}^N \mathbf{E}_i' \mathbf{W}_i^{-1} (\mathbf{Z}_i - \boldsymbol{\delta}_i) = \mathbf{0}, \quad (3.6)$$

where

$$\boldsymbol{\delta}_i = \begin{bmatrix} \delta_{i12} \\ \delta_{i13} \\ \vdots \\ \delta_{i_{n-1,n}} \end{bmatrix}, \mathbf{E}_i = \begin{bmatrix} \frac{\delta_{i12}}{\alpha_{12}} \\ \frac{\delta_{i13}}{\alpha_{13}} \\ \vdots \\ \frac{\delta_{i_{n-1,n}}}{\alpha_{12}} \end{bmatrix}, \mathbf{W}_i = \begin{bmatrix} w_{i12} & 0 & \cdots & 0 \\ 0 & w_{i13} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & w_{i_{n-1,n}} \end{bmatrix}.$$

Under some regularity conditions, it can be shown that the solutions to (3.3) and (3.6) are consistent and the resulting estimators $(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}})$ have following asymptotic variance-covariance matrix

$$\Sigma_{\boldsymbol{\beta}, \boldsymbol{\alpha}} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{B} & \mathbf{C} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Lambda}_{11} & \boldsymbol{\Lambda}_{12} \\ \boldsymbol{\Lambda}_{21} & \boldsymbol{\Lambda}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B}' \\ \mathbf{0} & \mathbf{C} \end{bmatrix},$$

where

$$\begin{aligned}
\mathbf{A} &= \left(\sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1} \mathbf{D}_i \right)^{-1}, \\
\mathbf{B} &= \left(\sum_{i=1}^N \mathbf{E}_i' \mathbf{W}_i^{-1} \mathbf{E}_i \right)^{-1} \left(\sum_{i=1}^N \mathbf{E}_i' \mathbf{W}_i \frac{\partial \mathbf{Z}_i}{\partial \boldsymbol{\beta}} \right)^{-1} \left(\sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1} \mathbf{D}_i \right)^{-1}, \\
\mathbf{C} &= \left(\sum_{i=1}^N \mathbf{E}_i' \mathbf{W}_i^{-1} \mathbf{E}_i \right)^{-1}, \\
\Lambda_{11} &= \sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i) (\mathbf{y}_i - \boldsymbol{\mu}_i)' \mathbf{V}_i^{-1} \mathbf{D}_i, \\
\Lambda_{12} &= \sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i) (\mathbf{Z}_i - \boldsymbol{\delta}_i)' \mathbf{W}_i^{-1} \mathbf{E}_i, \\
\Lambda_{22} &= \sum_{i=1}^N \mathbf{E}_i' \mathbf{W}_i^{-1} (\mathbf{Z}_i - \boldsymbol{\delta}_i) (\mathbf{Z}_i - \boldsymbol{\delta}_i)' \mathbf{W}_i^{-1} \mathbf{E}_i, \\
\Lambda_{21} &= \Lambda_{12}'.
\end{aligned}$$

The estimates are computed iteratively using the Fisher's scoring algorithms. The strategy for the numerical method is to solve the first set of GEE for $\boldsymbol{\beta}$ and then to insert the solutions $\boldsymbol{\beta}^{(m)}$ at iteration m in the second set of GEE for $\boldsymbol{\alpha}$. The iterations are repeated until a convergence is reached. The following are the two sets of iterative equations for $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$:

$$\begin{aligned}
\boldsymbol{\beta}^{(m+1)} &= \boldsymbol{\beta}^{(m)} + \left(\sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1} \mathbf{D}_i \right)^{-1} \left(\sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i) \right) \Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}^{(m)}}, \\
\boldsymbol{\alpha}^{(m+1)} &= \boldsymbol{\alpha}^{(m)} + \left(\sum_{i=1}^N \mathbf{E}_i' \mathbf{W}_i^{-1} \mathbf{E}_i \right)^{-1} \left(\sum_{i=1}^N \mathbf{E}_i' \mathbf{W}_i^{-1} (\mathbf{Z}_i - \boldsymbol{\delta}_i) \right) \Big|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^{(m)}}.
\end{aligned}$$

Longitudinal data are characterized by correlated responses within individuals, which makes the estimation difficult. By avoiding the specification of the joint density, GEE is a better alternative than the traditional ML. However, GEE fails to be consistent in some missing-data scenarios, which we will examine in Chapter 4. We will also review other estimation methods such as weighted GEE and maximum pseudo-likelihood, to cope with the missing-data problem.

Chapter 4

Missing Data and Estimation

Methods

4.1 Introduction

The methods described in Chapter 3 work well for data with no missing values. In practice, missing-data problem is common and the methods in Chapter 3 might no longer be valid. In particular, in longitudinal studies, it is often common to have subjects dropping out during the study period and never returning back, forming a monotone missing data pattern. Sometimes, missing data are not treated appropriately and analyzed ignoring the missing data completely. Such an approach is considered as complete case analysis and despite its convenience, it does not always result in a correct inference as shown in Section 4.2. In general, missing data had not been properly studied until Rubin (1976) formalized theories on missing data.

In this chapter, a brief overview of types of missing data can be found in section 4.2, followed by the estimation methods for longitudinal data with drop-outs in section 4.3.

4.2 Overview of missing data

4.2.1 Missing-data pattern

Missing data can be classified based on the generated pattern. The simplest pattern occurs when a single variable is partially observed among other fully-observed variables creating a univariate missing-data pattern. A similar example is when, instead of a single variable, several variables are missing in a haphazard manner. Such pattern, commonly observed in surveys, are termed item non-response by survey practitioners.

Another missing-data pattern is monotone pattern in which the individuals in a sample provides n_i first responses out of T ordered questions, $Y_1, \dots, Y_{n_i}$, but not the rest, $Y_{n_i+1}, \dots, Y_T$. It is easily recognizable by its rectangle-trapezoidal shape when the data are rearranged in a way that a group of individuals in the sample have all $Y_1, \dots, Y_{n_i}$ observed and $Y_{n_i+1}, \dots, Y_T$ missing for all $n_i = 1, \dots, T$. The monotone missing-data pattern is of our interest because, in longitudinal studies, participants often drop out during the study period and never returning back, creating a monotone missing-data pattern. In practice, the missing-data pattern is nearly monotone and in such situations, it is recommended to remove small amount of data as a sacrifice to artificially create a monotone pattern (Little and Rubin, 2002).

4.2.2 Missing-data mechanism

Missing data may arise from an underlying missing-data mechanism. By masking the true values, missing data have an impact on inferential analysis, and its degree may vary depending on the type of missing-data mechanisms, which will be reviewed shortly. Let $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{iT})$ denote a vector of responses and $\mathbf{R}_i = (R_{i1}, \dots, R_{iT})$, its corresponding missing-data indicator vector with $R_{it} = 0$ if Y_{it} is missing and $R_{it} = 1$, otherwise. Then $\mathbf{R}_i$ is a vector of binary random variables, which can be assigned a probability distribution. A reasonable assumption

is that the missingness depends on the responses and a missingness parameter γ , resulting in the probability distribution of missing data,

$$f(\mathbf{R}_i|\mathbf{Y}_i, \gamma).$$

Partitioning the responses $\mathbf{Y}_i$ into $\mathbf{Y}_i^o$, observed components, and $\mathbf{Y}_i^m$, missing components, the joint probability distribution of $\mathbf{Y}_i$ is

$$f(\mathbf{Y}_i|\boldsymbol{\alpha}, \boldsymbol{\beta}) = f(\mathbf{Y}_i^o, \mathbf{Y}_i^m|\boldsymbol{\alpha}, \boldsymbol{\beta}), \quad (4.1)$$

where $\boldsymbol{\beta}$ are some regression parameters and $\boldsymbol{\alpha}$ are parameters of association between the responses.

Then the joint distribution of $\mathbf{Y}_i$ and $\mathbf{R}_i$ is

$$f(\mathbf{Y}_i, \mathbf{R}_i|\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) = f(\mathbf{Y}_i|\boldsymbol{\alpha}, \boldsymbol{\beta})f(\mathbf{R}_i|\mathbf{Y}_i, \gamma). \quad (4.2)$$

Here three scenarios of missing-data can be considered. First, when the missingness does not depend on $\mathbf{Y}_i$, the distribution of the missing data becomes

$$f(\mathbf{R}_i|\mathbf{Y}_i, \gamma) = f(\mathbf{R}_i|\gamma).$$

Such missing-data mechanism is called missing completely at random (MCAR). Clearly, the observed sub-sample resulting from such missing-data mechanism is still a random sub-sample of the original sample, so a complete-case analysis is still a valid method under MCAR. An obvious consequence is the increase in the sample variance due to the sample size reduction.

A second scenario is when the missingness depends only on the observed components $\mathbf{Y}^o$. In this case,

$$f(\mathbf{R}_i|\mathbf{Y}_i, \gamma) = f(\mathbf{R}_i|\mathbf{Y}_i^o, \gamma). \quad (4.3)$$

This missing-data mechanism is called missing at random (MAR). Notice that MCAR is a special case of MAR. Complete-case analysis might not be the right method under MAR, because the missingness is no longer random and depends on the observed values. So the sub-sample resulting from MAR might not be representative of the sample without missing data.

A third scenario is when the missingness depends on both observed and missing data, which is referred to as missing not at random (MNAR). Like MAR, the missingness is not random and depends on either observed or unobserved values.

4.2.3 Class of methods for missing data

Little and Rubin (2002) enumerated four classes of methods to deal with missing data:

1. Complete-case analysis: when some variables are partially observed in an observation, the method consists in removing the observation entirely and to carry on with the standard statistical analysis. It is the simplest method, but as reviewed in the previous section, it can result in biased estimates if the missing-data mechanism is not MCAR.
2. Weighted complete-case analysis: in survey designs, a common practice is to weight a sample unit by the inverse of its probability of selection. To deal with the missing-data problem, weighted complete-case analysis consists in modifying the weights as an adjustment for non-response. For longitudinal data with drop-outs, the weighted GEE method illustrated in Section 4.3.3 is an example of this class of methods.
3. Imputation analysis: the strategy for imputation-based methods is to fill in the missing response from a sample unit using the observed values from other sample units and to carry on with the standard statistical analysis. The common imputation methods include the mean imputation, which consists in substituting the missing values by the mean of the observed values, and

regression imputation, in which missing values are estimated by applying regression on the sets of observed values.

4. Model-based analysis: model-based methods consist in modeling missing-data mechanism using a probability function and drawing inference using procedures such as the ML. It is more beneficial than the other three classes of methods, because one can use statistical procedures, such as statistical tests for underlying assumptions, and diagnostic methods considering missing-data mechanism. Additionally, it allows one to adjust the variance estimation considering the missing-data mechanism in the data.

4.3 Estimation Methods for Binary Longitudinal Data with Drop-outs

4.3.1 Maximum Likelihood

Consider the joint density of $\mathbf{Y}_i$ and $\mathbf{R}_i$ given by (4.2). The marginal joint density of observed responses $\mathbf{Y}_i^o$ and the missing-data indicator $\mathbf{R}_i$ can be obtained by summing (4.2) over the missing responses $\mathbf{Y}_i^m$

$$f(\mathbf{Y}^o, \mathbf{R} | \alpha, \beta, \gamma) = \sum_{\mathbf{Y}_i^m} f(\mathbf{Y}_i^o, \mathbf{Y}_i^m | \alpha, \beta) f(\mathbf{R}_i | \mathbf{Y}_i^o, \mathbf{Y}_i^m, \gamma). \quad (4.4)$$

The marginal likelihood of α , β and γ is

$$\begin{aligned} L(\alpha, \beta, \gamma | \mathbf{Y}^o, \mathbf{R}) &= \prod_{i=1}^N f(\mathbf{Y}_i^o, \mathbf{R}_i | \alpha, \beta, \gamma) \\ &= \prod_{i=1}^N \left\{ \sum_{\mathbf{Y}_i^m} f(\mathbf{Y}_i^o, \mathbf{Y}_i^m | \alpha, \beta) f(\mathbf{R}_i | \mathbf{Y}_i^o, \mathbf{Y}_i^m, \gamma) \right\}. \end{aligned} \quad (4.5)$$

Taking the log of (4.5), the marginal log-likelihood is obtained as

$$l(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} | \mathbf{Y}^o, \mathbf{R}) = \sum_{i=1}^N \log \left\{ \sum_{\mathbf{Y}_i^m} f(\mathbf{Y}_i^o, \mathbf{Y}_i^m | \boldsymbol{\beta}) f(\mathbf{R}_i | \mathbf{Y}_i^o, \mathbf{Y}_i^m, \boldsymbol{\gamma}) \right\}. \quad (4.6)$$

As a special case, when missing-data mechanism is MAR, the likelihood can be simplified further as

$$\begin{aligned} L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} | \mathbf{Y}^o, \mathbf{R}) &= \prod_{i=1}^N \left\{ \sum_{\mathbf{Y}_i^m} f(\mathbf{Y}_i^o, \mathbf{Y}_i^m | \boldsymbol{\beta}) f(\mathbf{R}_i | \mathbf{Y}_i^o, \boldsymbol{\gamma}) \right\} \quad (\text{using (4.3)}) \\ &= \prod_{i=1}^N f(\mathbf{Y}_i^o | \boldsymbol{\beta}) f(\mathbf{R}_i | \mathbf{Y}_i^o, \boldsymbol{\gamma}). \end{aligned} \quad (4.7)$$

Note that the second term in (4.7) does not depend on the parameters of interests $\boldsymbol{\beta}$. In such cases, the missing-data mechanism can be “ignored” so MAR is also called an ignorable missing-data mechanism. If the missing-data mechanism also depend on missing responses $\mathbf{Y}_i^m$, a full specification of the density function of the missing data is required in the ML method, so such a missing-data mechanism is termed non-ignorable (NI).

The model parameters $\boldsymbol{\psi} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ are estimated by maximizing the log-likelihood (4.6), which is equivalent to solving the score equations:

$$\mathbf{S} = \frac{\partial l}{\partial \boldsymbol{\psi}} = \mathbf{0} \quad (4.8)$$

with respect to $\boldsymbol{\psi}$.

Under some regularity conditions, the MLE $\hat{\boldsymbol{\psi}}$ is asymptotically normal with the mean vector $\boldsymbol{\psi}$ and with an asymptotical variance-covariance matrix equal to the inverse of the Fisher information $\mathbf{I} = -E \left(\frac{\partial^2 l}{\partial \boldsymbol{\psi}^2} \right)$.

The solutions to (4.8) can be obtained numerically using Fisher’s scoring algorithm, which requires the derivation of the Fisher information matrix. However, deriving Fisher information matrix becomes cumbersome and complicated when the number of parameters increases. In particular, the burden of estimating

the nuisance parameters $\boldsymbol{\alpha}$ was already raised in Chapter 3 for multivariate binary responses. As the number of responses increases, the number of association parameters quickly proliferates, making the likelihood equations intractable. In general, with a large number of parameters, numerical optimization methods are used to obtain the ML estimates, such as derivative-free simplex method (Nelder and Mead, 1965). However, numerical method may fail to converge, converge too slowly or may not even be reliable, as with a large number of parameters, the likelihood surface may become multi-modal and in such a case, the initial values determine the values at convergence, so different initial values can produce different results, making the numerical method unreliable. Clearly, a need for a method alleviating the burden of estimating the nuisance parameters is identified. Non-likelihood-based methods, such as the GEE and weighted GEE, which do not require the specification of the joint distribution of the responses, can be alternatives to the ML and are reviewed in the next sections.

As an example, the ML method is demonstrated using tri-variate binary longitudinal responses with monotone drop-outs based on a Bahadur model for the responses. Recall from Section 3.2.3 that the main advantage of using a Bahadur model for multivariate binary responses is its ability to specify the likelihood function explicitly even when the number of responses gets large, unlike log-linear models based on marginal odds ratios. For $T = 3$, the joint density of (y_{i1}, y_{i2}, y_{i3}) under a Bahadur model is given by

$$f(y_{i1}, y_{i2}, y_{i3} | \boldsymbol{\beta}) = \prod_{t=1}^3 p_{it}^{y_{it}} (1 - p_{it})^{1-y_{it}} \\ (1 + \alpha_{12} z_{i1} z_{i2} + \alpha_{13} z_{i1} z_{i3} + \alpha_{23} z_{i2} z_{i3} + \alpha_{123} z_{i1} z_{i2} z_{i3}), \quad (4.9)$$

where

$$p_{it} = E(Y_{it} | \boldsymbol{\beta}, \mathbf{x}_i) = \text{Pr}(Y_{it} = 1 | \boldsymbol{\beta}, \mathbf{x}_i) = \frac{\exp(\mathbf{x}_i' \boldsymbol{\beta})}{1 + \exp(\mathbf{x}_i' \boldsymbol{\beta})},$$

and

$$z_{it} = \frac{y_{it} - p_{it}}{\sqrt{p_{it}(1 - p_{it})}}.$$

For simplicity, the third-order moment α_{123} is assumed to be zero.

The joint density (4.9) can also be written, equivalently, as a product of conditional density functions:

$$f(y_{i1}, y_{i2}, y_{i3} | \boldsymbol{\beta}) = f(y_{i1} | \boldsymbol{\beta}) f(y_{i2} | y_{i1}, \boldsymbol{\beta}) f(y_{i3} | y_{i1}, y_{i2}, \boldsymbol{\beta}),$$

where

$$\begin{aligned} f(y_{i1} | \boldsymbol{\beta}) &= p_{i1}^{y_{i1}} (1 - p_{i1})^{1-y_{i1}}, \\ f(y_{i2} | y_{i1}, \boldsymbol{\beta}, \alpha_{12}) &= \frac{f(y_{i1}, y_{i2} | \alpha_{12}, \boldsymbol{\beta})}{f(y_{i1})} \\ &= p_{i2}^{y_{i2}} (1 - p_{i2})^{1-y_{i2}} (1 + \alpha_{12} z_{i1} z_{i2}), \\ f(y_{i3} | y_{i1}, y_{i2}, \boldsymbol{\beta}, \boldsymbol{\alpha}) &= \frac{f(y_{i1}, y_{i2}, y_{i3} | \boldsymbol{\alpha}, \boldsymbol{\beta})}{f(y_{i1}, y_{i2})} \\ &= p_{i3}^{y_{i3}} (1 - p_{i3})^{1-y_{i3}} \left(\frac{1 + \alpha_{12} z_{i1} z_{i2} + \alpha_{13} z_{i1} z_{i3} + \alpha_{23} z_{i2} z_{i3}}{1 + \alpha_{12} z_{i1} z_{i2}} \right). \end{aligned}$$

Suppose the probability of a missing response at time t is a logistic function of y_{it} , the unobserved response at time of dropout, $y_{i,t-1}$, the response taken right before the drop-out and some missingness parameters $\boldsymbol{\gamma}$. That is,

$$Pr(R_{it} = 0 | R_{i1} = \dots = R_{i,t-1} = 1, y_{it}, y_{i,t-1}, \boldsymbol{\gamma}) = \frac{\exp(\gamma_0 + \gamma_1 y_{i,t-1} + \gamma_2 y_{it})}{1 + \exp(\gamma_0 + \gamma_1 y_{i,t-1} + \gamma_2 y_{it})}.$$

Then the probability that an individual drops out from the study at time m is

$$\begin{aligned} &Pr(R_{i1} = 1, \dots, R_{i,m-1} = 1, R_{im} = 0 | y_{i1}, \dots, y_{im}, \boldsymbol{\gamma}) \\ &= Pr(R_{i1} = 1) Pr(R_{i2} = 1 | y_{i1}, y_{i2}, \boldsymbol{\gamma}) Pr(R_{i3} = 1 | R_{i2} = 1, y_{i2}, y_{i3}, \boldsymbol{\gamma}) \dots \\ &Pr(R_{im} = 0 | R_{i2} = \dots = R_{i,m-1} = 1, y_{i1}, \dots, y_{im}, \boldsymbol{\gamma}) \\ &= \prod_{i=2}^{m-1} \left\{ \frac{1}{1 + \exp(\gamma_0 + \gamma_1 y_{i,t-1} + \gamma_2 y_{it})} \right\} \left\{ \frac{\exp(\gamma_0 + \gamma_1 y_{i,m-1} + \gamma_2 y_{im})}{1 + \exp(\gamma_0 + \gamma_1 y_{i,m-1} + \gamma_2 y_{im})} \right\}^{I(m \leq T)}, \end{aligned}$$

where $T = 3$ in our example.

Using (4.5), the marginal density of observed response when the drop-out occurs at $m = 2$ is

$$\begin{aligned}
& \sum_{Y_{i2}=0}^1 \sum_{Y_{i3}=0}^1 f(y_{i1}, y_{i2}, y_{i3} | \beta) \left\{ \frac{\exp(\gamma_0 + \gamma_1 y_{i1} + \gamma_2 y_{i2})}{1 + \exp(\gamma_0 + \gamma_1 y_{i1} + \gamma_2 y_{i2})} \right\} \\
&= f(y_{i1} | \beta) \sum_{Y_{i2}=0}^1 \left[f(y_{i2} | y_{i1}, \beta) \left\{ \frac{\exp(\gamma_0 + \gamma_1 y_{i1} + \gamma_2 y_{i2})}{1 + \exp(\gamma_0 + \gamma_1 y_{i1} + \gamma_2 y_{i2})} \right\} \sum_{Y_{i3}=0}^1 f(y_{i3} | y_{i1}, y_{i2}, \beta) \right] \\
&= f(y_{i1} | \beta) \sum_{Y_{i2}=0}^1 f(y_{i2} | y_{i1}, \beta) \left\{ \frac{\exp(\gamma_0 + \gamma_1 y_{i1} + \gamma_2 y_{i2})}{1 + \exp(\gamma_0 + \gamma_1 y_{i1} + \gamma_2 y_{i2})} \right\}. \tag{4.10}
\end{aligned}$$

When the dropout occurs at $m = 3$, the marginal density is

$$\begin{aligned}
& \sum_{Y_{i3}=0}^1 f(y_{i1}, y_{i2}, y_{i3} | \beta) \left\{ \frac{1}{1 + \exp(\gamma_0 + \gamma_1 y_{i1} + \gamma_2 y_{i2})} \right\} \left\{ \frac{\exp(\gamma_0 + \gamma_1 y_{i2} + \gamma_2 y_{i3})}{1 + \exp(\gamma_0 + \gamma_1 y_{i2} + \gamma_2 y_{i3})} \right\} \\
&= f(y_{i1} | \beta) f(y_{i2} | y_{i1}, \beta) \left\{ \frac{1}{1 + \exp(\gamma_0 + \gamma_1 y_{i1} + \gamma_2 y_{i2})} \right\} \\
& \sum_{Y_{i3}=0}^1 f(y_{i3} | y_{i1}, y_{i2}, \beta) \left\{ \frac{\exp(\gamma_0 + \gamma_1 y_{i2} + \gamma_2 y_{i3})}{1 + \exp(\gamma_0 + \gamma_1 y_{i2} + \gamma_2 y_{i3})} \right\}. \tag{4.11}
\end{aligned}$$

When all three responses are fully observed, that is $m = 4$, the marginal density is

$$\begin{aligned}
& f(y_{i1}, y_{i2}, y_{i3} | \beta) \left\{ \frac{1}{1 + \exp(\gamma_0 + \gamma_1 y_{i1} + \gamma_2 y_{i2})} \right\} \left\{ \frac{1}{1 + \exp(\gamma_0 + \gamma_1 y_{i2} + \gamma_2 y_{i3})} \right\} \\
&= f(y_{i1} | \beta) f(y_{i2} | y_{i1}, \beta) f(y_{i3} | y_{i1}, y_{i2}, \beta) \\
& \left\{ \frac{1}{1 + \exp(\gamma_0 + \gamma_1 y_{i1} + \gamma_2 y_{i2})} \right\} \left\{ \frac{1}{1 + \exp(\gamma_0 + \gamma_1 y_{i2} + \gamma_2 y_{i3})} \right\}. \tag{4.12}
\end{aligned}$$

The following likelihood can then be derived using the density functions (4.7)

- (4.9), :

$$L = \prod_{i=1}^N \left\{ \sum_{Y_{im}=0}^1 f(\mathbf{y}_i | \alpha, \beta) f(R_{i1} = 1, \dots, R_{i,m-1} = 1, R_{im} = 0 | y_{i1}, \dots, y_{im}, \gamma) \right\}$$

or the log-likelihood function

$$l = \sum_{i=1}^N \log \left[\sum_{Y_{im}=0}^1 f(\mathbf{y}_i | \boldsymbol{\alpha}, \boldsymbol{\beta}) f(R_{i1} = 1, \dots, R_{i,m-1} = 1, R_{im} = 0 | y_{i1}, \dots, y_{im}, \boldsymbol{\gamma}) \right].$$

The log-likelihood is maximized by solving the score equations

$$\mathbf{S} = \frac{\partial l}{\partial \boldsymbol{\psi}} = \mathbf{0}$$

with respect to $\boldsymbol{\psi} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$.

With only three responses, nine parameters are left to be estimated with the assumption that the third-order moment is null, which already makes it cumbersome to derive the Fisher information matrix. Due to the difficulties highlighted earlier in the section, one resorts to numerical optimization methods, which do not require the specification of the second-order derivatives to solve the score equations. In addition, numerical optimization packages in statistical software generally provide the Hessian matrix, which is a good approximation of the negative of the Fisher information matrix, and can be used to approximate the variance-covariance matrix of the MLE.

4.3.2 Generalized Estimating Equations

Recall from Chapter 3 that for multivariate binary responses without missing data, the GEE is favoured over the ML method because it does not require a specification of the joint distribution, therefore, facilitating the computational process to calculate the estimates. With drop-outs in the data, the GEE can still be applied provided that the missing-data mechanism is MCAR. Using only the

observed responses $\mathbf{y}_i^o$, Prentice (1988)'s modified GEEs are

$$U_{\beta}(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1} (\mathbf{y}_i^o - \boldsymbol{\mu}_i^o) = \mathbf{0}, \quad (4.13)$$

$$U_{\alpha}(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \sum_{i=1}^N \mathbf{E}_i' \mathbf{W}_i^{-1} (\mathbf{z}_i^o - \boldsymbol{\delta}_i^o) = \mathbf{0}. \quad (4.14)$$

The estimates are calculated iteratively using Fisher's scoring method. The following two sets of equations are alternated at each iteration:

$$\begin{aligned} \boldsymbol{\beta}^{(m+1)} &= \boldsymbol{\beta}^{(m)} + \left(\sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1} \mathbf{D}_i \right)^{-1} \left(\sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1} (\mathbf{y}_i^o - \boldsymbol{\mu}_i^o) \right) \\ \boldsymbol{\alpha}^{(m+1)} &= \boldsymbol{\alpha}^{(m)} + \left(\sum_{i=1}^N \mathbf{E}_i' \mathbf{W}_i^{-1} \mathbf{E}_i \right)^{-1} \left(\sum_{i=1}^N \mathbf{E}_i' \mathbf{W}_i^{-1} (\mathbf{z}_i^o - \boldsymbol{\delta}_i^o) \right) \end{aligned}$$

The solutions to (4.13) and (4.14) can be shown to be consistent and are asymptotically normal with the asymptotic variance-covariance matrix

$$\Sigma_{\beta, \alpha} = \begin{bmatrix} \mathbf{A}^o & \mathbf{0} \\ \mathbf{B}^o & \mathbf{C}^o \end{bmatrix} \begin{bmatrix} \Lambda_{11}^o & \Lambda_{12}^o \\ \Lambda_{21}^o & \Lambda_{22}^o \end{bmatrix} \begin{bmatrix} \mathbf{A}^o & \mathbf{B}^{o'} \\ \mathbf{0} & \mathbf{C}^o \end{bmatrix},$$

where

$$\begin{aligned}
\mathbf{A}^o &= \left(\sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1} \mathbf{D}_i \right)^{-1}, \\
\mathbf{B}^o &= \left(\sum_{i=1}^N \mathbf{E}_i' \mathbf{W}_i^{-1} \mathbf{E}_i \right)^{-1} \left(\sum_{i=1}^N \mathbf{E}_i' \mathbf{W}_i \frac{\partial \mathbf{Z}_i^o}{\partial \boldsymbol{\beta}} \right)^{-1} \left(\sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1} \mathbf{D}_i \right)^{-1}, \\
\mathbf{C}^o &= \left(\sum_{i=1}^N \mathbf{E}_i' \mathbf{W}_i^{-1} \mathbf{E}_i \right)^{-1}, \\
\Lambda_{11}^o &= \sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1} (\mathbf{y}_i^o - \boldsymbol{\mu}_i^o) (\mathbf{y}_i^o - \boldsymbol{\mu}_i^o)' \mathbf{V}_i^{-1} \mathbf{D}_i, \\
\Lambda_{12}^o &= \sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1} (\mathbf{y}_i^o - \boldsymbol{\mu}_i^o) (\mathbf{Z}_i^o - \boldsymbol{\delta}_i^o)' \mathbf{W}_i^{-1} \mathbf{E}_i, \\
\Lambda_{22}^o &= \sum_{i=1}^N \mathbf{E}_i' \mathbf{W}_i^{-1} (\mathbf{Z}_i^o - \boldsymbol{\delta}_i^o) (\mathbf{Z}_i^o - \boldsymbol{\delta}_i^o)' \mathbf{W}_i^{-1} \mathbf{E}_i, \\
\Lambda_{21}^o &= \Lambda_{12}^{o'}.
\end{aligned}$$

The GEE under a missing-data mechanism is similar to a complete-case analysis except that the observations with drop-outs are not entirely removed from the analysis but observed responses up to the time of drop-out m are included in the analysis. Under MCAR, as the observed responses are random sub-samples of fully observed responses, GEE yields consistent estimates. However, if the missingness is MAR or NI, the GEE might fail to be consistent, because MAR and NI mechanisms no longer mask data randomly but selectively, in which case:

$$E(\mathbf{Y}_i^o | \mathbf{x}_i, \boldsymbol{\beta}) \neq \boldsymbol{\mu}_i^o.$$

To cope with the missing-data mechanism, Robins (1995) proposed a weighting approach to the existing GEE, which is reviewed in the following section.

4.3.3 Weighted Generalized Estimating Equations

As the GEE method can produce biased estimates under a missing-data mechanism other than MCAR, an alternative method has been proposed by Robins

(1995), which is referred to as the weighted GEE (wGEE) method. The wGEE method is an extension of the GEE method, in which the inverse of an individual's probability of response at time t is used as weights in the GEE. Fitzmaurice (1995) modified the original weighted GEE for longitudinal binary responses with drop-outs and uses the inverse of an individual's probability of drop-out at time m rather than the probability of response at all time points.

Consider a longitudinal study with T time points. Let $\mathbf{R}_i$ be the vector of missing-data indicators for i^{th} participant with $R_{it} = 1$ if the response y_{it} is observed and $R_{it} = 0$, otherwise. The drop-out mechanism can be summarized by a single random variable M_i rather than a vector of missing-data indicators $\mathbf{R}_i$ such that

$$M_i = 1 + \sum_{t=1}^T R_{it}.$$

We assume that data are always observed at the first time point. Note that M_i is, in fact, the time of drop-out for the i^{th} individual with $M_i = 2, \dots, T+1$. When $M_i = T+1$, the responses are fully observed for that individual.

The probability of drop-out can then be expressed as

$$\nu_{im} = Pr(M_i = m_i | \mathbf{y}_i, \mathbf{x}_i, \boldsymbol{\gamma}).$$

The wGEE method consists in solving GEE weighted by the inverse of ν_{im} . Here, extended GEE model by Prentice (1988) is illustrated with weights. The two sets of wGEEs are

$$U_{\boldsymbol{\beta}}(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \sum_{i=1}^N \frac{1}{\nu_{im}} \mathbf{D}_i' \mathbf{V}_i^{-1} (\mathbf{Y}_i^o - \boldsymbol{\mu}_i^o) = \mathbf{0}, \quad (4.15)$$

$$U_{\boldsymbol{\alpha}}(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \sum_{i=1}^N \frac{1}{\nu_{im}} \mathbf{E}_i' \mathbf{W}_i^{-1} (\mathbf{Z}_i^o - \boldsymbol{\delta}_i^o) = \mathbf{0}. \quad (4.16)$$

The above equations (4.12) and (4.13) can be written as

$$U_{\beta}(\beta, \alpha) = \sum_{i=1}^N \sum_{m=2}^{T+1} \frac{I(M_i = m)}{\nu_{im}} \mathbf{D}_i(m)' \mathbf{V}_i(m)^{-1} (\mathbf{Y}_i(m) - \boldsymbol{\mu}_i(m)) = \mathbf{0}, \quad (4.17)$$

$$U_{\alpha}(\beta, \alpha) = \sum_{i=1}^N \sum_{m=2}^{T+1} \frac{I(M_i = m)}{\nu_{im}} \mathbf{E}_i(m)' \mathbf{W}_i(m)^{-1} (\mathbf{Z}_i(m) - \boldsymbol{\delta}_i(m)) = \mathbf{0}, \quad (4.18)$$

where $\mathbf{Y}_i(m)$ are the first $m - 1$ elements of $\mathbf{Y}_i$. Analogously, $\mathbf{D}_i(m)$, $\mathbf{V}_i(m)$, $\boldsymbol{\mu}_i(m)$, $\mathbf{E}_i(m)$, $\mathbf{W}_i(m)$, $\mathbf{Z}_i(m)$ and $\boldsymbol{\delta}_i(m)$ are defined. Then the solutions to the equations (4.17) are consistent provided that the drop-out probability is correctly specified, since

$$\begin{aligned} & E \left[\frac{I(M_i = m)}{\nu_{im}} \mathbf{D}_i(m)' \mathbf{V}_i(m)^{-1} (\mathbf{Y}_i^o(m) - \boldsymbol{\mu}_i^o(m)) \right] \\ &= E \left[E_{M_i|\mathbf{Y}_i} \left(\frac{I(M_i = m)}{\nu_{im}} \right) \mathbf{D}_i(m)' \mathbf{V}_i(m)^{-1} (\mathbf{Y}_i^o(m) - \boldsymbol{\mu}_i^o(m)) \right] \\ &= E \left[\frac{E_{M_i|\mathbf{Y}_i}(I(M_i = m))}{\nu_{im}} \mathbf{D}_i(m)' \mathbf{V}_i(m)^{-1} (\mathbf{Y}_i^o(m) - \boldsymbol{\mu}_i^o(m)) \right] \\ &= E \left[\frac{E_{M_i|\mathbf{Y}_i}(Pr(M_i = m_i | \mathbf{y}_i, \mathbf{x}_i, \gamma))}{\nu_{im}} \mathbf{D}_i(m)' \mathbf{V}_i(m)^{-1} (\mathbf{Y}_i^o(m) - \boldsymbol{\mu}_i^o(m)) \right] \\ &= E [\mathbf{D}_i(m)' \mathbf{V}_i(m)^{-1} (\mathbf{Y}_i^o(m) - \boldsymbol{\mu}_i^o(m))] \quad (\text{under MAR}) \\ &= \mathbf{0}. \end{aligned}$$

Similarly, it can be shown that the solutions to the equations (4.18) are consistent provided that the drop-out probability is correctly specified.

If the missing-data mechanism is MCAR or MAR, the missingness parameters γ can be consistently estimated from the observed responses using ML method, separately from β and α . However, if the missing-data mechanism is NI, the estimators of γ may no longer be consistent, affecting consequently the consistency of $\hat{\beta}$.

To find the solutions to wGEE, γ are first estimated using the ML and are used in

the following two Fisher's scoring algorithms which are alternated at each iteration:

$$\begin{aligned}\boldsymbol{\beta}^{(m'+1)} &= \boldsymbol{\beta}^{(m')} + \left(\sum_{i=1}^N \frac{1}{\nu_{im}} \mathbf{D}_i' \mathbf{V}_i^{-1} \mathbf{D}_i \right)^{-1} \left(\sum_{i=1}^N \frac{1}{\nu_{im}} \mathbf{D}_i' \mathbf{V}_i^{-1} (\mathbf{y}_i^o - \boldsymbol{\mu}_i^o) \right), \\ \boldsymbol{\alpha}^{(m'+1)} &= \boldsymbol{\alpha}^{(m')} + \left(\sum_{i=1}^N \frac{1}{\nu_{im}} \mathbf{E}_i' \mathbf{W}_i^{-1} \mathbf{E}_i \right)^{-1} \left(\sum_{i=1}^N \frac{1}{\nu_{im}} \mathbf{E}_i' \mathbf{W}_i^{-1} (\mathbf{Z}_i^o - \boldsymbol{\delta}_i^o) \right).\end{aligned}$$

The asymptotic covariance matrix can be obtained using the sandwich type estimator:

$$\Sigma_{\boldsymbol{\beta}, \boldsymbol{\alpha}} = \begin{bmatrix} \mathbf{A}^o & \mathbf{0} \\ \mathbf{B}^o & \mathbf{C}^o \end{bmatrix} \begin{bmatrix} \Lambda_{11}^o & \Lambda_{12}^o \\ \Lambda_{21}^o & \Lambda_{22}^o \end{bmatrix} \begin{bmatrix} \mathbf{A}^o & \mathbf{B}^{o'} \\ \mathbf{0} & \mathbf{C}^o \end{bmatrix},$$

where

$$\begin{aligned}\mathbf{A} &= \left(\sum_{i=1}^N \frac{1}{\nu_{im}} \mathbf{D}_i' \mathbf{V}_i^{-1} \mathbf{D}_i \right)^{-1}, \\ \mathbf{B}^o &= \left(\sum_{i=1}^N \frac{1}{\nu_{im}} \mathbf{E}_i' \mathbf{W}_i^{-1} \mathbf{E}_i \right)^{-1} \left(\sum_{i=1}^N \frac{1}{\nu_{im}} \mathbf{E}_i' \mathbf{W}_i \frac{\partial \mathbf{Z}_i^o}{\partial \boldsymbol{\beta}} \right)^{-1} \left(\sum_{i=1}^N \frac{1}{\nu_{im}} \mathbf{D}_i' \mathbf{V}_i^{-1} \mathbf{D}_i \right)^{-1}, \\ \mathbf{C} &= \left(\sum_{i=1}^N \frac{1}{\nu_{im}} \mathbf{E}_i' \mathbf{W}_i^{-1} \mathbf{E}_i \right)^{-1}, \\ \Lambda_{11}^o &= \sum_{i=1}^N \frac{1}{\nu_{im}^2} \mathbf{D}_i' \mathbf{V}_i^{-1} (\mathbf{y}_i^o - \boldsymbol{\mu}_i^o) (\mathbf{y}_i^o - \boldsymbol{\mu}_i^o)' \mathbf{V}_i^{-1} \mathbf{D}_i, \\ \Lambda_{12}^o &= \sum_{i=1}^N \frac{1}{\nu_{im}^2} \mathbf{D}_i' \mathbf{V}_i^{-1} (\mathbf{y}_i^o - \boldsymbol{\mu}_i^o) (\mathbf{Z}_i^o - \boldsymbol{\delta}_i^o)' \mathbf{W}_i^{-1} \mathbf{E}_i, \\ \Lambda_{22}^o &= \sum_{i=1}^N \frac{1}{\nu_{im}^2} \mathbf{E}_i' \mathbf{W}_i^{-1} (\mathbf{Z}_i^o - \boldsymbol{\delta}_i^o) (\mathbf{Z}_i^o - \boldsymbol{\delta}_i^o)' \mathbf{W}_i^{-1} \mathbf{E}_i, \\ \Lambda_{21}^o &= \Lambda_{12}^{o'}.\end{aligned}$$

The calculation of weights for the wGEE method is illustrated here for $T = 3$ time points. Suppose the first response is always observed and the missing-data mechanism is MAR such that the probability of response missing at time m_i depends on its immediately preceding response y_{i,m_i-1} and some drop-out parameter

γ through a logistic model

$$Pr(R_{it} = 0 | R_{i1} = \dots = R_{i,t-1} = 1, y_{i,t-1}, \gamma) = \frac{\exp(\gamma_0 + \gamma_1 y_{i,t-1})}{1 + \exp(\gamma_0 + \gamma_1 y_{i,t-1})}.$$

Then the probability that the i^{th} participant drops out from the study at time t ($t = 2, 3, 4$) is

$$\begin{aligned} \nu_{i2} &= \frac{\exp(\gamma_0 + \gamma_1 y_{i1})}{1 + \exp(\gamma_0 + \gamma_1 y_{i1})} && (\text{for } t = 2) \\ \nu_{i3} &= \left\{ \frac{1}{1 + \exp(\gamma_0 + \gamma_1 y_{i1})} \right\} \left\{ \frac{\exp(\gamma_0 + \gamma_1 y_{i2})}{1 + \exp(\gamma_0 + \gamma_1 y_{i2})} \right\} && (\text{for } t = 3) \\ \nu_{i4} &= \left\{ \frac{1}{1 + \exp(\gamma_0 + \gamma_1 y_{i1})} \right\} \left\{ \frac{1}{1 + \exp(\gamma_0 + \gamma_1 y_{i2})} \right\} && (\text{for } t = 4). \end{aligned}$$

Using the identities above, the log-likelihood can be calculated as

$$l = \sum_{i=1}^N \log(\nu_{im_i})$$

The likelihood score equations for γ is:

$$U = \frac{\partial l}{\partial \gamma} = \mathbf{0},$$

which is solved for the ML estimator $\hat{\gamma}$.

The ML estimates are computed iteratively using Fisher's scoring algorithm

$$\gamma^{(m'+1)} = \gamma^{(m')} - \left[\frac{\partial^2 l}{\partial \gamma \partial \gamma'} \right]^{-1} \left[\frac{\partial l}{\partial \gamma} \right].$$

4.3.4 Pseudo-Maximum Likelihood Estimation

Although the non-likelihood based methods, such as the GEE and wGEE, can produce consistent estimates under a correctly specified missing-data mechanism, both are sensitive to misspecification of missing-data mechanism (Liang and Zeger,

1986 and Robins, 1995). Under non-ignorable missingness, estimators for both GEE and wGEE methods may become biased and, consequently, the ML method could be more reliable. In section 4.2, the likelihood for multivariate binary responses was derived using the Bahadur model. Although the MLE is a highly efficient method, as the number of parameters increases, the estimating equations may quickly become intractable. Numerical methods have been developed to maximize the likelihood, however, when the number of parameters increases, the convergence might become slow or even unreliable, as the likelihood surface becomes highly multi-modal. In practice, the ML method is recommended when the number of responses is less than or equal to five (Sinha et al., 2011).

For marginal regression, the primary interest is often in the marginal regression coefficients and not in the association between responses within observations. To alleviate the burden of estimating a large number of parameters, the association parameters can be treated as a nuisance. In the presence of nuisance parameters, several methods including maximum partial likelihood (Cox, 1975) and pseudo-maximum likelihood (Gong and Samaniego, 1981) have been developed. The idea of pseudo-likelihood is to replace all nuisance parameters by their estimates using estimation methods other than the ML and to simplify the computation of the likelihood equations. Such method is attractive for binary multivariate models, *e.g.* Bahadur model, because the number of second and higher-order association parameters can quickly proliferate when the number of responses increase. To illustrate this approach, for the binary responses $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{iT})$, the Bahadur model is defined as a function of the regression parameters $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$. By replacing the true parameter $\boldsymbol{\alpha}$ by its estimate $\hat{\boldsymbol{\alpha}}$, the joint density can be written as a function of $\boldsymbol{\beta}$ as

$$f(\mathbf{Y}_i | \boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}).$$

The resulting likelihood is termed as the pseudo-likelihood, which takes the form

$$L = \prod_{i=1}^N f(\mathbf{Y}_i | \boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}).$$

By maximizing the pseudo-likelihood, the pseudo-maximum likelihood estimator $\hat{\boldsymbol{\beta}}_{PL}$ is obtained. Under regularity conditions, pseudo-maximum likelihood estimator is expected to be consistent provided $\hat{\boldsymbol{\alpha}}$ is consistent. Its efficiency also depends on the relative efficiency of $\hat{\boldsymbol{\alpha}}$.

A simple application is to treat all pairwise correlations $\alpha_{st} = \text{Corr}(Y_s, Y_t)$ in Bahadur model as nuisance parameters and fix them at zero. Under the naive assumption of independence, Troxel et al (1998) developed the independent pseudo-likelihood (IPL)

$$L(\boldsymbol{\beta}, \boldsymbol{\gamma} | \hat{\boldsymbol{\alpha}} = 0) = \prod_{i=1}^N f(\mathbf{Y}_i^o, R_i)$$

Summing over the missing responses, the IPL is obtained as a product of the marginal distribution of $\mathbf{Y}_i^o$:

$$L(\boldsymbol{\beta}, \boldsymbol{\gamma} | \hat{\boldsymbol{\alpha}} = 0) = \prod_{i=1}^N \left(\sum_{\mathbf{Y}_i^m} f(\mathbf{Y}_i^o, R_i) \right).$$

The resulting pseudo-score function is

$$S(\boldsymbol{\psi}) = \frac{\partial \log L(\boldsymbol{\psi})}{\partial(\boldsymbol{\psi})},$$

where $\boldsymbol{\psi} = (\boldsymbol{\beta}, \boldsymbol{\gamma})$.

It can be shown that $E[S(\boldsymbol{\psi})] = 0$ at the true values of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. The estimators $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ are consistent for $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, and their joint asymptotic variance Σ can be obtained using the sandwich estimator:

$$\Sigma = NE \left\{ \frac{\partial S(\boldsymbol{\psi})}{\partial(\boldsymbol{\psi})} \right\}^{-1} \sum_{i=1}^N E[S(\boldsymbol{\psi})S(\boldsymbol{\psi})'] E \left\{ \frac{\partial S(\boldsymbol{\psi})}{\partial(\boldsymbol{\psi})} \right\}^{-1}.$$

The maximization of the IPL may be performed using numerical methods,

such as derivative-free simplex algorithm (Nelder and Mead, 1965). The pseudo-maximum likelihood estimators $\hat{\beta}$ and $\hat{\gamma}$ are consistent provided that the joint density of the responses and the missing-data mechanism are correctly specified.

When the correlations between responses within individuals is close to zero, the IPL can produce satisfactory estimates. However, the estimator becomes highly inefficient when the association becomes strong. As an alternative approach, Sinha et al. (2011) proposed a bivariate pseudo-likelihood method, which takes into account the second-order correlation with respect to the first observation. The method is based on the assumption that the first response is always observed, which is reasonable in longitudinal studies, where a baseline measurement is collected from each individual prior to the enrollment. Assuming the baseline response provides a meaningful information on the following responses, intuitively, bivariate maximum pseudo-likelihood estimator should be more efficient than independent maximum pseudo-likelihood estimator, which naively treats the responses to be independent.

To illustrate the approach, consider the baseline response with the density

$$f(y_{i1}|\mathbf{x}_i, \beta) = p_{i1}^{y_{i1}}(1 - p_{i1})^{1-y_{i1}},$$

where p_{it} is modeled by a logistic regression

$$p_{it} = E(Y_{it}|\mathbf{x}_i, \beta) = Pr(Y_{it} = 1|\mathbf{x}_i, \beta) = \frac{\exp(\mathbf{x}_i'\beta)}{1 + \exp(\mathbf{x}_i'\beta)}.$$

Assuming the first response is always observed, Y_{it} can be written in the form of conditional distribution

$$f(y_{it}, r_{i1}|y_{i1}, \mathbf{x}_i, \beta, \alpha, \gamma) = f(y_{it}|y_{i1}, \mathbf{x}_i, \beta, \alpha)f(r_{it}|y_{i1}, y_{it}, \mathbf{x}_i, \gamma). \quad (4.19)$$

To obtain $f(y_{it}|y_{i1}, \mathbf{x}_i, \beta, \alpha)$, the bivariate distribution of Y_{i1} and Y_{it} is needed as

the definition of conditional probability is

$$f(y_{it}|y_{i1}, \mathbf{x}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}) = f(y_{i1}, y_{it}|\mathbf{x}_i, \boldsymbol{\beta}, \boldsymbol{\alpha})f(y_{i1}|\mathbf{x}_i, \boldsymbol{\beta}) \quad (4.20)$$

Adopting the Bahadur model, the bivariate density is obtained as

$$f(y_{it}, y_{i1}|\mathbf{x}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}) = p_{i1}^{y_{i1}}(1 - p_{i1})^{1-y_{i1}}p_{it}^{y_{it}}(1 - p_{it})^{1-y_{it}} \left\{ 1 + \alpha_{it} \frac{(y_{i1} - p_{i1})(y_{it} - p_{it})}{\sqrt{p_{i1}(1 - p_{i1})p_{it}(1 - p_{it})}} \right\}.$$

Then using (4.20),

$$f(y_{it}|y_{i1}, \mathbf{x}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}) = p_{it}^{y_{it}}(1 - p_{it})^{1-y_{it}} \left\{ 1 + \alpha_{it} \frac{(y_{i1} - p_{i1})(y_{it} - p_{it})}{\sqrt{p_{i1}(1 - p_{i1})p_{it}(1 - p_{it})}} \right\}.$$

Suppose the missing-data mechanism only depends on the first response y_{i1} , the current responses y_{it} and missingness parameters $\boldsymbol{\gamma}$ through a logistic function

$$f(R_{it} = 0|y_{i1}, y_{it}, \boldsymbol{\gamma}) = \frac{\exp(\gamma_0 + \gamma_1 y_{i1} + \gamma_2 y_{it})}{1 + \exp(\gamma_0 + \gamma_1 y_{i1} + \gamma_2 y_{it})}.$$

The missing-data mechanism may also depend on the covariates, however, they are ignored here for simplicity. Given y_{it} and treating it as independent, the bivariate pseudo-likelihood can be obtained as

$$L(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\gamma}|\mathbf{Y}) = \prod_{i=1}^N f(y_{i1}|\mathbf{x}_i, \boldsymbol{\beta}) \prod_{t=2}^T \{f(y_{it}, r_{it}|y_{i1}, \mathbf{x}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\gamma})\}^{r_{it}} \left\{ \sum_{y_{it}=0}^1 f(y_{it}, r_{it}|y_{i1}, \mathbf{x}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\gamma}) \right\}^{1-r_{it}}. \quad (4.21)$$

The bivariate maximum pseudo-likelihood estimator $\hat{\boldsymbol{\psi}} = (\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\gamma}})$ is obtained by maximizing (4.21), which is equivalent to solving the pseudo-likelihood equations

$$\mathbf{S}(\boldsymbol{\psi}) = \frac{\partial \log L(\boldsymbol{\psi})}{\partial \boldsymbol{\psi}}$$

with respect to $\boldsymbol{\psi} = (\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\gamma})$.

It can be shown that $E[S(\boldsymbol{\psi})] = 0$ at the true value $\boldsymbol{\psi}$. The bivariate maximum pseudo-likelihood estimator $\hat{\boldsymbol{\psi}}$ is consistent provided that the joint distribution and the missing-data mechanism in (4.20) are correctly specified, and $\hat{\boldsymbol{\psi}}$ is asymptotically normal with the variance-covariance matrix

$$\boldsymbol{\Sigma} = NE \left\{ \frac{\partial S(\boldsymbol{\psi})}{\partial(\boldsymbol{\psi})} \right\}^{-1} \left\{ \sum_{i=1}^N E[S(\boldsymbol{\psi})S(\boldsymbol{\psi})'] \right\} E \left\{ \frac{\partial S(\boldsymbol{\psi})}{\partial(\boldsymbol{\psi})} \right\}^{-1}. \quad (4.22)$$

The review of the estimation methods for longitudinal data with missing responses is provided in this chapter. Given the limitation of each method, a simulation was conducted to assess the sensitivity of each estimation method to misspecifications. In Chapter 5, the simulation settings will be presented followed by the results and discussion.

Chapter 5

Simulation Study

5.1 Simulation settings

5.1.1 Responses model

For our simulation study, we chose a single dichotomous covariate x_i , which is a group indicator for the i^{th} subject and is assumed to be fully observed. The corresponding measurement or response at time t , denoted by Y_{it} , is binary and its marginal distribution is assumed to be Bernoulli with the probability of success p_{it} such that:

$$p_{it} = E(Y_{it}|x_i, \boldsymbol{\beta}) = Pr(Y_{it} = 1|x_i, \boldsymbol{\beta}) = \frac{\exp(\beta_0 + \beta_G x_i + \beta_T(t-1))}{1 + \exp(\beta_0 + \beta_G x_i + \beta_T(t-1))}$$

where $t = 1, \dots, T$.

The joint distribution of the multivariate binary variable $\mathbf{Y}_i$ is reviewed in Chapter 3. The Bahadur model was chosen for the study because the likelihood can be derived explicitly regardless of the number of association parameters. Recall that

for T responses, the Bahadur model is given by

$$P(Y_{i1} = y_{i1}, \dots, Y_{iT} = y_{iT}) = \prod_{t=1}^T p_{it}^{y_{it}} (1 - p_{it})^{(1-y_{it})} \\ \times \left[1 + \sum_{t_1 < t_2} \alpha_{i_{t_1 t_2}} z_{it_1} z_{it_2} + \sum_{t_1 < t_2 < t_3} \alpha_{i_{t_1 t_2 t_3}} z_{it_1} z_{it_2} z_{it_3} + \dots + \alpha_{i_{1, \dots, T}} z_{i1} \dots z_{iT} \right],$$

where

$$z_{it} = \frac{y_{it} - p_{it}}{\sqrt{\text{var}(y_{it})}} = \frac{y_{it} - p_{it}}{\sqrt{p_{it}(1 - p_{it})}},$$

and

$$\begin{aligned} \alpha_{i_{t_1 t_2}} &= E(z_{it_1} z_{it_2}), \\ \alpha_{i_{t_1 t_2 t_3}} &= E(z_{it_1} z_{it_2} z_{it_3}), \\ &\dots \\ \alpha_{i_{1, \dots, T}} &= E(z_{i1} \dots z_{iT}). \end{aligned}$$

For simplicity, we only considered pairwise correlations in our simulation study and ignored all moments of order higher than 3, *i.e.* $\alpha_{123} = \dots = \alpha_{1, \dots, T} = 0$.

The number of responses in our simulation study was varied from 3 to 4 to find the impact of the increase in the number of responses on the estimates. For the true correlation, we chose exchangeable correlation and the first order auto-regressive correlation $AR(1)$. In each type of correlations, we increased the value of the correlation parameter α from 0.2 to 0.4 in order to study the properties of the regression estimators under both small and large values of α .

5.1.2 Drop-out model

The missing-data pattern in our simulation study was considered monotone, resulting from participants' drop-outs: when a response is missing at time m for the first time, all following responses are missing, *i.e.* $Y_{im}, \dots, Y_{iT}$ are missing. An additional assumption is that the baseline measurement Y_{i1} is always observed,

which is reasonable here, as the participants' baseline measurements are usually taken prior to the start of study.

Let $\mathbf{R}_i$ denote the vector of missing-value indicators such that $R_{it} = 0$ when the response from the i^{th} subject at time t is missing and $R_{it} = 1$ otherwise. Then the drop-out index for the i^{th} subject, M_i , can be formulated as

$$M_i = 1 + \sum_{t=1}^T R_{it}.$$

Assuming that the baseline measurement is always observed, *i.e.* $R_{i1} = 1$, M_i ranges from 2 to $T + 1$. When $M_i = T + 1$, all T responses are observed.

For the missing-data mechanism, the probability of a drop-out at time m was modeled by a logistic regression which takes into account the unobserved response y_{im} at the time of drop-out, and the observed response immediately preceding the drop-out, $y_{i,m-1}$, such that

$$\begin{aligned} Pr(R_{im} = 0 | R_{i2} = \dots = R_{i,m-1} = 1, y_{i,m-1}, y_{im}, \gamma) \\ = \frac{\exp(\gamma_0 + \gamma_P y_{i,m-1} + \gamma_C y_{im})}{1 + \exp(\gamma_0 + \gamma_P y_{i,m-1} + \gamma_C y_{im})}. \end{aligned} \quad (5.1)$$

It is possible to simulate different missing-data mechanisms such as MCAR, MAR and NI using (5.1). Setting both γ_P and γ_C to 0 in (5.1) results in MCAR as the drop-out mechanism only depends on γ_0 . In this case,

$$Pr(M_i = m | y_{i1}, \dots, y_{im}, \gamma) = Pr(M_i = m | \gamma_0) = c,$$

where c is a constant.

Assigning γ_P a non-zero value and $\gamma_C = 0$, the drop-out mechanism becomes MAR:

$$Pr(M_i = m | y_{i1}, \dots, y_{iT}, \gamma) = Pr(M_i = m | y_{i1}, \dots, y_{i,m-1}, \gamma_0, \gamma_P).$$

Assigning γ_C a non-zero value, the drop-out mechanism becomes informative or non-ignorable (NI).

Note that although the group effect might be present in the missing-data mechanism, it was neglected in our simulation study by setting (5.1) to be functionally independent of the group membership x_i .

In addition, if the drop-out mechanism is NI, *i.e.* $\gamma_C \neq 0$, we assumed the drop-out mechanism depends only on the response at the time of drop-out Y_{im} , which is not observed, but not on its following responses, $Y_{i,m+1}, \dots, Y_{iT}$ such that

$$Pr(M_i = m | y_{i1}, \dots, y_{iT}, \gamma) = Pr(M_i = m | y_{i1}, \dots, y_{im}, \gamma). \quad (5.2)$$

Then the probability of a drop-out at time m can be formulated as

$$\begin{aligned} \nu_{im} &= Pr(M_i = m | y_{i1}, \dots, y_{iT}, \gamma) \\ &= Pr(M_i = m | y_{i1}, \dots, y_{im}, \gamma) \quad (\text{using (5.2)}) \\ &= Pr(R_{im} = 0, R_{i2} = \dots = R_{i,m-1} = 1 | y_{i1}, \dots, y_{im}, \gamma) \\ &= Pr(R_{i2} = 1 | y_{i1}, \gamma) Pr(R_{i3} = 1 | R_{i2} = 1, y_{i1}, y_{i2}, \gamma) \dots \\ &\quad Pr(R_{im} = 0 | R_{i2} = \dots = R_{i,m-1} = 1, y_{i1}, \dots, y_{im}, \gamma) \\ &= \left\{ \prod_{t=2}^{m-1} Pr(R_{it} = 1 | R_{i1} = \dots = R_{i,t-1} = 1, y_{i1}, \dots, y_{it}, \gamma) \right\} \\ &\quad \{1 - Pr(R_{im} = 1 | R_{i1} = \dots = R_{im} = 1, y_{i1}, \dots, y_{im}, \gamma)\}^{I(m \leq T)} \\ &= \left\{ \prod_{t=2}^{m-1} \frac{\exp(\gamma_0 + \gamma_P y_{i,t-1} + \gamma_C y_{it})}{1 + \exp(\gamma_0 + \gamma_P y_{i,t-1} + \gamma_C y_{it})} \right\} \\ &\quad \left\{ \frac{\exp(\gamma_0 + \gamma_P y_{i,m-1} + \gamma_C y_{im})}{1 + \exp(\gamma_0 + \gamma_P y_{i,m-1} + \gamma_C y_{im})} \right\}^{I(m \leq T)}, \end{aligned} \quad (5.3)$$

where $I()$ is an indicator function.

For the simulation study, we varied γ to produce different MCAR, MAR and NI missing-data mechanisms.

5.1.3 Estimation methods

Four methods reviewed in Chapter 4 such as GEE, wGEE, ML and bivariate maximum pseudo-likelihood were used to estimate α and β in our simulation study.

For GEE, we solved the unweighted equations (4.14) and (4.15) using Fisher's scoring method. The sample variance-covariance matrix of the estimates for $\hat{\alpha}$ and $\hat{\beta}$ were approximated respectively by

$$\begin{aligned}\hat{\Sigma}_{\hat{\beta}} &= B_{\hat{\beta}}^{-1} H_{\hat{\beta}} B_{\hat{\beta}}^{-1}, \\ \hat{\Sigma}_{\hat{\alpha}} &= B_{\hat{\alpha}}^{-1} H_{\hat{\alpha}} B_{\hat{\alpha}}^{-1},\end{aligned}$$

where

$$\begin{aligned}B_{\hat{\beta}} &= \sum_{i=1}^N D_i' V_i^{-1} D_i, \\ H_{\hat{\beta}} &= \sum_{i=1}^N D_i' V_i^{-1} (Y_i^o - \mu_i^o) (Y_i^o - \mu_i^o)' V_i^{-1} D_i, \\ B_{\hat{\alpha}} &= \sum_{i=1}^N E_i' W_i^{-1} E_i, \\ H_{\hat{\alpha}} &= \sum_{i=1}^N E_i' W_i^{-1} (Z_i^o - \delta_i^o) (Z_i^o - \delta_i^o)' W_i^{-1} E_i.\end{aligned}$$

We also varied the 'working' correlation, which included independent, exchangeable and $AR(1)$ to assess the effects of 'working' correlation on the estimates.

For wGEE, we first computed the ML estimates of γ_0 and γ_P using the Fisher's scoring method, which were then used to calculate the weights ν_{im} in equations (4.15) and (4.16). The solutions were found iteratively using Fisher's scoring methods. The variance-covariance matrix of $\hat{\alpha}$ and $\hat{\beta}$ was approximated respectively

by

$$\begin{aligned}\hat{\Sigma}_{\hat{\beta}} &= B_{\hat{\beta}}^{-1} H_{\hat{\beta}} B_{\hat{\beta}}^{-1}, \\ \hat{\Sigma}_{\hat{\alpha}} &= B_{\hat{\alpha}}^{-1} H_{\hat{\alpha}} B_{\hat{\alpha}}^{-1},\end{aligned}$$

where

$$\begin{aligned}B_{\hat{\beta}} &= \sum_{i=1}^N \frac{1}{\nu_{im}} D_i' V_i^{-1} D_i, \\ H_{\hat{\beta}} &= \sum_{i=1}^N \frac{1}{\nu_{im}} D_i' V_i^{-1} (Y_i - \mu_i) (Y_i^o - \mu_i^o)' V_i^{-1} D_i, \\ B_{\hat{\alpha}} &= \sum_{i=1}^N \frac{1}{\nu_{im}} E_i' W_i^{-1} E_i, \\ H_{\hat{\alpha}} &= \sum_{i=1}^N \frac{1}{\nu_{im}} E_i' W_i^{-1} (Z_i^o - \delta_i^o) (Z_i^o - \delta_i^o)' W_i^{-1} E_i.\end{aligned}$$

We varied the ‘working’ correlation to study the effects of the correlation structure on the estimates.

The bivariate maximum pseudo-likelihood method was modified to adjust for the missing-data mechanism that differs from the original mechanism used in Sinha et al. (2011). In our simulation study, the missing-data mechanism is a drop-out, which depends on the response immediately preceding the drop-out or the response at the time of drop-out, whereas in Sinha et al. (2011), the missing data are non-monotone and the mechanism depends only on the first response, which is assumed to be fully observed, and the response at the time of drop-out. To account for the possible association between the the response immediately preceding the drop-out and the response at the time of drop-out, we considered the bivariate distribution between two consecutive responses rather than a bivariate relation with respect to the first response as in Sinha et al (2011).

The bivariate Bahadur model is

$$f(y_{i,t-1}, y_{it} | \mathbf{x}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}) = p_{i,t-1}^{y_{i,t-1}} (1 - p_{i,t-1})^{1-y_{i,t-1}} p_{it}^{y_{it}} (1 - p_{it})^{1-y_{it}} \left\{ 1 + \alpha_{i_{t-1},t} \frac{(y_{i,t-1} - p_{i,t-1})(y_{it} - p_{it})}{\sqrt{p_{i,t-1}(1 - p_{i,t-1})p_{it}(1 - p_{it})}} \right\}$$

from which, the conditional probability of y_{it} given the previous response $y_{i,t-1}$ is obtained as

$$f(y_{it} | y_{i,t-1}, \mathbf{x}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}) = p_{it}^{y_{it}} (1 - p_{it})^{1-y_{it}} \left\{ 1 + \alpha_{i_{t-1},t} \frac{(y_{i,t-1} - p_{i,t-1})(y_{it} - p_{it})}{\sqrt{p_{i,t-1}(1 - p_{i,t-1})p_{it}(1 - p_{it})}} \right\}. \quad (5.4)$$

Given the conditional probability (5.2) and the missing-data mechanism (5.3), the following pseudo-likelihood can be derived

$$L(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\gamma} | \mathbf{Y}) = \prod_{i=1}^N f(y_{i1} | \mathbf{x}_{i1}, \boldsymbol{\beta}) \prod_{t=2}^T \{f(y_{it}, r_{it} | y_{i,t-1}, \mathbf{x}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\gamma})\}^{r_{it}} \left\{ \sum_{y_{it}=0}^1 f(y_{it}, r_{it} | y_{i,t-1}, \mathbf{x}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\gamma}) \right\}^{1-r_{it}}.$$

The pseudo-maximum likelihood estimator is the solution to the following pseudo-score equations

$$\mathbf{S}(\boldsymbol{\psi}) = \frac{\partial \log \mathbf{L}(\boldsymbol{\psi})}{\partial \boldsymbol{\psi}} = \mathbf{0},$$

where $\boldsymbol{\psi} = (\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\gamma})$. For our simulation study, the estimates of $\boldsymbol{\psi}$ were calculated using numerical optimization method. We used the method by Byrd et al. (1995), in which boundary constraints on the parameters can be specified. The constraint specification for the correlation parameter α is necessary in the computational procedures because exceeding the boundary $[-1, 1]$ will result in the conditional probability (5.4) outside the probability space. Consequently, the likelihood may become negative or the log-likelihood may become undefined, in which case the numerical method will fail to converge. We used the optim package in R, which

implements the optimization method. The sample variance-covariance matrix was calculated using (4.22). The `optim` package conveniently provides the Hessian matrix at the convergence, which is a good approximation of the negative of Fisher's information matrix, so we used the Hessian matrix to approximate the Fisher's information matrix in the estimation of the variance-covariance matrix.

For the ML estimation, the likelihood for $T = 3$ time points was explicitly derived in section 4.3.1. Similarly, the likelihood for $T = 4$ can be derived. We modified the ML to include different correlation structures in order to make it comparable to GEE and wGEE such as $\alpha = 0$ for no correlations, $\alpha_{t_1, t_2} = \alpha$ for all $t_1 \neq t_2$ for exchangeable correlations and $\alpha_{t_1, t_2} = \alpha^{|t_2 - t_1|}$ for $AR(1)$ correlations. The ML estimates for our simulated data were computed using the numerical optimization method by Byrd et al (1995) by specifying the boundary constraints on α . Using the Hessian matrix, we estimated the sample variance-covariance matrix of the estimates.

5.1.4 Model misspecifications

The sensitivity of the estimation methods to model misspecifications is the primary interest of our empirical study. There are two possible misspecifications in the estimation procedures: first, the joint distribution of the responses can be misspecified in the estimation method by adopting an incorrect 'working' correlation. We used exchangeable and $AR(1)$ correlations for the true correlation structure and varied the 'working' correlation for the estimation methods including independent, exchangeable and $AR(1)$ processes to explore the sensitivity of the methods to the misspecification of the joint distribution of the responses. For each scenario, we varied α from 0.2 to 0.4 to assess how the correlation affect the regression estimators.

A second possible misspecification is for the drop-out mechanism. To study the effect of the misspecification in the drop-out mechanism, we varied γ in (5.3) to generate different types of drop-out mechanisms, which included MCAR, MAR

and NI. For MAR and NI, we varied γ_P and γ_C to investigate how each parameter affects the estimation method.

5.2 Results

In our simulation study, the bias of the estimator $\hat{\beta}$ is estimated by the empirical bias

$$\sum_{s=1}^S \frac{\hat{\beta}_s - \beta}{S},$$

where $\hat{\beta}_s$ is the estimate of β for the s^{th} simulated dataset and S is the total number of simulations. In our study, we used a total of 1000 simulated datasets for each misspecification scenario.

For the results, we calculated:

1. the percentage relative bias,

$$\sum_{s=1}^S \frac{\hat{\beta}_s - \beta}{\beta} \times 100,$$

2. the empirical mean squared error (MSE),

$$\sum_{s=1}^S \frac{(\hat{\beta}_s - \beta)^2}{S},$$

3. the 95% coverage probability of the estimator $\hat{\beta}$,

$$\sum_{s=1}^S \frac{I(\hat{\beta}_s - 1.96 \times SE(\beta_s) \leq \beta \leq \hat{\beta}_s + 1.96 \times SE(\beta_s))}{S},$$

where SE is the standard error of $\hat{\beta}_s$ and I is an indicator function.

5.2.1 Results for Exchangeable Correlation

Drop-out mechanism: MCAR

For MCAR drop-out mechanism, we chose $(\gamma_0, \gamma_P, \gamma_G) = (-1, 0, 0)$. Table 5.1 presents results from the simulations with 3 responses and $\alpha = 0.2$. Keeping the same setting, we varied the number of responses to 4. The results are summarized in Table 5.2. We then replicated the previous simulations but with $\alpha = 0.4$ instead. The results for $T = 3$ and $T = 4$ are shown in Tables 5.3 and 5.4 respectively.

For non-likelihood based methods such as GEE and wGEE, the relative bias of $\hat{\beta}$ was close to zero regardless of the ‘working’ correlation and the strength of α . The GEE was remarkably efficient for β_T with the lowest MSE. The wGEE showed slightly higher MSE. The coverage probabilities for both methods were close to the nominal confidence level 0.95. Increase in the sample size did not affect neither methods, except for the MSE.

The bivariate maximum pseudo-likelihood showed a relative bias of $\hat{\beta}$ close to zero. The increase in α did not affect the relative bias. Although the MSEs for the estimators of β_0 and β_G were low compared to the non-likelihood methods, the MSE was quite large for the time coefficient β_T and was slightly higher than that of the ML estimator with the correct ‘working’ correlation, which was the exchangeable correlation.

For the ML method, estimation with the independence ‘working’ correlation showed some small bias for the regression parameters. Tables 5.2 and 5.4 show that an increase in α resulted in an increase in the bias for both β_0 and β_G . More importantly, the ML with independence assumption was highly inefficient for β_T with the largest MSE and has noticeably low coverage probabilities, especially for β_T . From Tables 5.2 and 5.4, we also noticed that an increase in α leads to a higher MSE and considerably lower coverage probabilities.

The ML with exchangeable ‘working’ correlation showed an empirical bias close to zero for all the regression parameters when the sample size is $N = 500$, whereas when $N = 200$, some small bias was observed for β_T . It has the lowest MSE for

β_0 and β_G but is much less efficient than GEE and wGEE for β_T . The coverage probabilities for β_T were close to 0.95.

For the ML with $AR(1)$ ‘working’ correlation structure, the empirical relative bias was close to 0% for $\alpha = 0.2$. However, as evident from Tables 5.2 and 5.4, an increase in α resulted in a small bias, which was more pronounced for β_T . The MSE was slightly lower than that obtained from the method with exchangeable ‘working’ correlation, and the coverage probabilities were close to 0.95.

Table 5.1: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **exchangeable correlation** ($\alpha = 0.2$) and **MCAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		1.4	1.5	0.3		0.0431	0.0577	0.0162		96.3	95.4	94.6		
		Exchangeable	-0.4	1.4	1.6	0.1	0.0041	0.0437	0.0614	0.0156	94.5	96.7	95.1	94.4		
	GEE	AR1	10.0	1.4	1.7	0.1	0.0068	0.0439	0.0599	0.0169	90.5	96.2	95.4	94.8		
		Independent		1.1	1.1	0.4		0.0373	0.0524	0.0139		95.3	94.9	94.6		
	Exchangeable	-1.1	1.1	1.2	0.3	0.0037	0.0369	0.0516	0.0136	94.6	95.4	95.1	94.6			
		AR1	12.3	1.2	1.3	0.3	0.0065	0.0373	0.0520	0.0137	89.4	95.4	94.8	94.9		
	BPL	Exchangeable/AR1	-7.5	0.8	0.8	-0.2	0.0045	0.0380	0.0518	0.0418	92.8	95.1	93.9	91.6		
		ML	Independent		-1.1	-2.5	-4.6		0.0422	0.0557	0.0627		92.2	89.1	69.1	
	Exchangeable	-5.3	0.6	0.9	-2.9	0.0035	0.0366	0.0515	0.0348	94.1	95.5	95.6	93.1			
		AR1	-5.2	1.1	1.2	-1.4	0.0046	0.0375	0.0517	0.0401	93.2	95.4	94.6	92.8		
	WGEE	Independent		0.4	0.2	1.9		0.0168	0.0239	0.0060		96.1	94.7	95.9		
		Exchangeable	-0.0	0.4	0.2	1.9	0.0016	0.0170	0.0255	0.0058	95.5	96.1	94.2	95.8		
	GEE	AR1	11.0	0.4	0.1	2.0	0.0031	0.0171	0.0248	0.0063	89.0	95.9	94.2	95.5		
		Independent		0.4	0.1	2.0		0.0144	0.0204	0.0051		95.7	95.0	95.8		
	Exchangeable	-0.1	0.4	0.1	1.7	0.0015	0.0143	0.0202	0.0049	95.0	95.7	94.6	96.0			
		AR1	14.2	0.3	0.1	1.9	0.0032	0.0145	0.0204	0.0050	85.8	95.7	94.7	96.0		
	BPL	Exchangeable/AR1	-2.9	0.2	-0.1	1.3	0.0019	0.0146	0.0205	0.0182	92.3	95.7	94.0	93.3		
		ML	Independent		-1.2	-3.1	-6.6		0.0175	0.0248	0.0396		91.9	87.3	68.5	
	Exchangeable	-2.9	0.1	-0.1	1.2	0.0014	0.0143	0.0199	0.0142	95.0	95.6	94.7	94.6			
		AR1	-1.2	0.4	-0.0	1.5	0.0019	0.0146	0.0203	0.0172	93.9	95.4	93.7	94.0		

N= 500

Table 5.2: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **exchangeable correlation** ($\alpha = 0.2$) and **MCAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C = (-1, 0, 0)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability							
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T				
N= 200																		
	WGEE	Independent		-1.2	-0.5	-1.7		0.0373	0.0515	0.0076					95.3	95.9	94.5	
		Exchangeable	-1.7	-1.1	-0.4	-1.1	0.0029	0.0370	0.0531	0.0073		94.2	95.1	95.5	94.7			
	GEE	AR1	16.8	-1.1	-0.4	-1.3	0.0072	0.0371	0.0520	0.0078		86.3	95.1	95.8	94.9			
		Independent		-0.9	-0.2	-1.9		0.0341	0.0452	0.0066			95.1	95.9	94.7			
		Exchangeable	-1.8	-0.8	-0.2	-1.4	0.0027	0.0338	0.0443	0.0063		94.0	95.1	95.4	94.7			
		AR1	19.9	-0.8	-0.2	-1.6	0.0073	0.0341	0.0447	0.0065		82.6	95.0	95.4	94.4			
	BPL	Exchangeable/AR1	-6.4	-0.8	-0.6	-4.6	0.0038	0.0355	0.0453	0.0174		92.9	94.3	94.7	92.7			
		Independent		-2.1	-3.9	-11.7		0.0415	0.0496	0.0255			90.7	89.0	73.4			
	ML	Exchangeable	-3.4	-0.9	-0.5	-1.6	0.0026	0.0343	0.0442	0.0146		93.1	94.2	94.5	92.7			
		AR1	-3.2	-0.7	-0.2	-3.9	0.0039	0.0353	0.0450	0.0162		91.2	94.2	94.2	93.3			
			WGEE	Independent		0.6	0.8	0.1		0.0152	0.0221	0.0031				96.7	94.6	94.8
				Exchangeable	0.2	0.7	0.9	0.0	0.0012	0.0149	0.0224	0.0031		94.4	96.1	94.6	94.3	
GEE	AR1		20.2	0.7	0.8	0.1	0.0041	0.0149	0.0220	0.0033		79.5	96.5	94.6	94.4			
	Independent			0.4	0.3	-0.1		0.0128	0.0196	0.0027			96.1	94.4	94.6			
	Exchangeable		-0.1	0.4	0.4	-0.0	0.0010	0.0126	0.0189	0.0026		95.4	96.5	94.6	94.4			
	AR1		23.2	0.4	0.3	-0.1	0.0044	0.0127	0.0192	0.0027		74.0	96.6	94.1	93.8			
	BPL	Exchangeable/AR1	-1.7	0.3	0.1	-0.8	0.0014	0.0132	0.0192	0.0072		92.9	96.2	92.6	94.1			
		Independent		-0.2	-2.0	-8.2		0.0156	0.0219	0.0144			91.9	87.5	76.2			
	ML	Exchangeable	-0.7	0.4	0.5	-0.1	0.0009	0.0127	0.0188	0.0054		94.2	95.9	94.2	94.1			
		AR1	1.3	0.5	0.5	-1.0	0.0014	0.0129	0.0191	0.0066		91.6	95.9	92.3	94.8			
	N= 500																	

Table 5.3: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **exchangeable correlation** ($\alpha = 0.4$) and **MCAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias			MSE			Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T	
N= 200															
	WGEE	Independent		-0.1	0.1	0.0		0.0479	0.0768	0.0132		95.1	94.1	96.2	
		Exchangeable	-0.1	0.1	0.4	-0.2	0.0041	0.0486	0.0826	0.0121	95.3	94.7	93.5	94.9	
	GEE	AR1	12.9	0.0	0.4	-0.5	0.0077	0.0492	0.0816	0.0144	83.7	95.0	93.0	94.9	
		Independent		-0.6	-0.9	1.0		0.0428	0.0705	0.0115		94.1	94.6	94.8	
		Exchangeable	-0.9	-0.5	-0.7	0.5	0.0039	0.0416	0.0673	0.0103	95.0	94.8	94.5	95.3	
		AR1	15.0	-0.5	-0.7	0.2	0.0083	0.0427	0.0694	0.0106	81.4	94.6	94.3	94.6	
	BPL	Exchangeable/AR1	-8.4	-0.9	-1.4	-0.5	0.0057	0.0430	0.0693	0.0532	92.6	94.3	92.6	89.2	
	ML	Independent		-8.4	-11.7	11.3		0.0544	0.0761	0.1102		83.3	80.5	22.1	
		Exchangeable	-8.2	-1.0	-1.8	-6.1	0.0038	0.0422	0.0712	0.0236	94.0	91.9	92.1	95.8	
		AR1	-5.8	1.9	2.3	-10.9	0.0043	0.0415	0.0661	0.0377	94.6	95.1	93.3	93.0	
N= 500															
	WGEE	Independent		-0.1	0.7	0.9		0.0178	0.0261	0.0052		95.6	95.3	96.1	
		Exchangeable	-0.4	0.0	0.9	0.9	0.0015	0.0183	0.0285	0.0049	95.7	95.4	95.9	95.0	
	GEE	AR1	12.9	0.1	0.9	1.1	0.0046	0.0185	0.0280	0.0057	73.3	95.4	95.5	95.0	
		Independent		-0.1	0.7	0.5		0.0165	0.0239	0.0046		94.9	95.9	95.0	
		Exchangeable	-0.6	-0.1	0.8	0.4	0.0015	0.0164	0.0233	0.0043	95.2	94.6	95.2	94.8	
		AR1	15.6	-0.0	0.9	0.5	0.0058	0.0165	0.0237	0.0044	64.1	94.5	95.4	95.2	
	BPL	Exchangeable/AR1	-4.1	-0.2	0.4	-6.2	0.0022	0.0165	0.0236	0.0234	93.4	93.9	95.0	93.3	
	ML	Independent		-8.2	-11.3	0.6		0.0319	0.0339	0.0992		73.7	76.2	2.2	
		Exchangeable	-2.5	0.2	1.6	-2.1	0.0012	0.0154	0.0212	0.0102	96.1	93.9	95.3	93.6	
		AR1	1.7	2.5	3.9	-7.3	0.0016	0.0164	0.0236	0.0141	91.8	94.6	94.5	96.1	

Table 5.4: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **exchangeable correlation** ($\alpha = 0.4$) and **MCAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		0.2	1.2	0.0		0.0479	0.0759	0.0066		94.5	93.8	95.3	95.3	
		Exchangeable	0.2	0.2	1.4	-0.0	0.0032	0.0478	0.0771	0.0061	95.1	93.8	93.8	95.1		
	GEE	AR1	24.1	0.3	1.4	-0.0	0.0139	0.0483	0.0771	0.0072	57.0	93.8	92.7	94.6		
		Independent		0.2	0.8	1.0		0.0420	0.0680	0.0059		94.3	93.6	94.4		
	Exchangeable	0.1	0.2	0.9	0.6	0.0028	0.0419	0.0650	0.0051	95.0	93.8	93.2	94.3			
		AR1	27.7	0.3	1.0	0.6	0.0162	0.0427	0.0660	0.0056	45.8	93.7	93.6	94.0		
	BPL	Exchangeable/AR1	-5.6	0.1	0.5	-1.3	0.0042	0.0433	0.0662	0.0238	92.6	92.9	91.7	92.3		
	ML	Independent		-8.8	-11.0	-0.7		0.0687	0.0730	0.0480	75.4	78.9	26.7			
Exchangeable		-3.9	3.3	1.5	7.4	0.0022	0.0358	0.0456	0.0097	94.3	92.5	93.4	92.1			
	WGEE	AR1	-2.8	3.5	4.5	-6.0	0.0025	0.0407	0.0628	0.0163	90.3	93.5	89.7	94.0		
		Independent		-0.9	-1.1	-1.2		0.0171	0.0271	0.0024		96.0	94.8	96.3		
	Exchangeable	0.3	-0.8	-0.9	-0.9	0.0011	0.0166	0.0270	0.0022	96.4	95.5	95.0	96.5			
		AR1	24.5	-0.8	-0.9	-1.1	0.0112	0.0169	0.0269	0.0025	26.3	95.4	95.2	96.5		
	GEE	Independent		-0.6	-0.8	-0.7		0.0145	0.0239	0.0022		95.6	94.9	96.2		
		Exchangeable	0.1	-0.4	-0.6	-0.6	0.0010	0.0140	0.0222	0.0019	95.6	95.8	94.9	95.5		
	BPL	AR1	28.0	-0.4	-0.6	-0.7	0.0139	0.0144	0.0228	0.0020	13.3	96.0	95.1	95.9		
		Exchangeable/AR1	-2.4	-0.5	-0.8	-1.3	0.0014	0.0148	0.0227	0.0097	94.7	95.5	93.4	93.4		
ML	Independent		-10.6	-13.8	-0.9		0.0501	0.0393	0.0417	56.5	67.5	3.0				
	Exchangeable	-4.7	1.4	1.0	4.6	0.0009	0.0105	0.0138	0.0032	94.3	93.3	95.3	92.7			
	AR1		-2.6	3.7	4.8	-7.2	0.0007	0.0148	0.0230	0.0061	93.2	94.2	91.9	95.5		

Drop-out mechanism: MAR

A MAR drop-out mechanism was first considered for $(\gamma_0, \gamma_P, \gamma_C) = (-1, 2, 0)$. Table 5.5 shows the results from simulations with 3 responses and $\alpha = 0.2$. Table 5.6 repeats them for 4 responses. The results with $\alpha = 0.4$, and with $T = 3$ and $T = 4$ are shown in Tables 5.7 and 5.8, respectively. We then used $\gamma_P = 1$ and replicated the previous simulations. The results are shown in Tables 5.9 - 5.12.

Although the GEE showed high efficiency and low relative bias for β_0 and β_G , the estimates showed some non-negligible bias, which was noticeably more pronounced for the GEE with independence assumption. The bias of $\hat{\beta}_T$ was also affected by an increase in α and γ_P . For example, in Tables 5.5 and 5.7, the increase in α from 0.2 to 0.4 for $N = 500$ resulted in percentage relative bias jumping from -82.1% to -164.8% for the GEE with independence assumption; from -12.4% to -26.0% for exchangeable correlation assumption; and from -31.9% to -46.1% for $AR(1)$ correlation assumption. In Tables 5.5 and 5.9, an increase in γ_P from 1 to 2 also increased the relative bias.

For all ‘working’ correlation, the wGEE showed some small bias when $N = 200$ and $\gamma_P = 2$, which decreased with an increase in the sample size. When $\gamma_P = 1$, the relative bias was close to 0% regardless of the sample size. Additionally, the relative bias was unaffected by α . The MSEs for different ‘working’ correlations differed slightly, however, were much higher than the GEE with the correct ‘working’ correlation.

The bivariate maximum pseudo-likelihood showed high relative bias for β_T , which increases with α and γ_P . Tables 5.5 and 5.7 show that an increase in α from 0.2 to 0.4 resulted in an increase in relative bias from -47.6% to -112.5% for β_T , while Tables 5.5 and 5.9 show that an increase in γ_P from 1 to 2 resulted in an increase in relative bias from -21.4% to -47.6%. The MSE was higher than that obtained from the ML with exchangeable or $AR(1)$ ‘working’ correlations.

The ML with independent assumption showed considerably high relative bias for the time coefficient exceeding 100% in all scenarios considered. Table 5.8 shows

the worst results with -16.3% for the percentage relative bias of $\hat{\beta}_G$ and -279.4% for the percentage relative bias of $\hat{\beta}_T$. When $\gamma_P = 2$, the increase in α from 0.2 to 0.4 nearly doubles the relative bias for β_T whereas when $\gamma_P = 1$, the relative bias for β_T moderately decreases with the increase in α . MSE for β_T was nearly 40 times the MSE found using the ML with an exchangeable correlation assumption and the coverage probability was 0.0.

The ML with exchangeable correlation assumption showed some small bias when $N = 200$ but decreased to 0% with $N = 500$. Similarly, the coverage probability for β_T was close to 0.9 for $N = 200$ but increases to 0.95 with $N = 500$. MSE for β_0 and β_G was the smallest with bivariate maximum pseudo-likelihood but the MSE for β_T was almost 3 times larger than the MSE found from the wGEE method.

The ML with $AR(1)$ correlation assumption showed a pattern similar to the ML with independence assumption, but in a moderate degree. Table 5.8 shows the worst performance with the percentage relative bias of -84.9% for β_T .

Table 5.5: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **exchangeable correlation** ($\alpha = 0.2$) and **MAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 2, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE			Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		1.6	0.6	-5.6		0.0726	0.1251	0.0364		94.2	93.7	93.1		
		Exchangeable	-7.4	1.6	0.7	-5.5	0.0125	0.0689	0.1137	0.0369	88.6	94.2	94.1	92.5		
	GEE	AR1	1.8	1.7	0.7	-4.8	0.0173	0.0695	0.1160	0.0372	89.0	94.4	93.8	92.7		
		Independent		0.4	-0.6	-82.1		0.0377	0.0559	0.0493		94.5	94.5	80.1		
	Exchangeable		-16.1	1.6	1.7	-12.4	0.0075	0.0388	0.0578	0.0243	85.8	94.4	95.0	92.5		
		AR1	-14.2	0.7	1.3	-31.9	0.0099	0.0384	0.0572	0.0282	84.9	94.5	95.0	90.4		
	BPL	Exchangeable/AR1	-19.8	0.8	0.6	-44.6	0.0195	0.0388	0.0589	0.1135	86.7	94.0	94.7	86.4		
	ML	Independent		0.1	-6.4	-182.2		0.0368	0.0610	0.1781		94.5	92.1	49.7		
Exchangeable		-5.4	1.7	1.6	2.5	0.0163	0.0407	0.0605	0.1025	85.6	94.3	94.3	86.1			
	AR1		-17.0	1.0	0.8	-41.0	0.0203	0.0387	0.0585	0.1091	85.5	94.1	94.7	85.9		
	WGEE	Independent		1.5	2.2	-2.0		0.0302	0.0504	0.0126		93.8	93.3	95.3		
		Exchangeable	-4.3	1.4	1.9	-2.0	0.0044	0.0286	0.0463	0.0128	92.9	93.9	93.6	95.1		
	AR1	6.5	1.5	2.0	-1.8	0.0062	0.0288	0.0473	0.0129	92.1	93.9	93.2	95.3			
		Independent		-0.3	-0.8	-80.7		0.0147	0.0235	0.0334		95.4	94.1	57.3		
	Exchangeable		-15.6	0.7	1.1	-11.2	0.0034	0.0150	0.0244	0.0085	83.3	95.9	93.9	93.7		
		AR1	-13.2	-0.3	0.7	-31.3	0.0041	0.0149	0.0242	0.0120	85.3	95.8	94.0	89.0		
	BPL	Exchangeable/AR1	-14.3	-0.2	0.2	-47.6	0.0083	0.0149	0.0252	0.0530	91.2	95.6	94.2	88.0		
		ML	Independent		-0.6	-7.2	-187.3		0.0142	0.0290	0.1571		94.8	88.9	16.6	
	Exchangeable		-4.0	0.7	0.9	-1.2	0.0059	0.0149	0.0245	0.0378	91.4	95.8	94.0	92.2		
		AR1	-11.8	-0.1	0.2	-43.8	0.0080	0.0148	0.0247	0.0494	91.2	95.7	94.4	88.6		
N= 500																

Table 5.6: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.2$) and MAR drop-outs ($\gamma_0, \gamma_P, \gamma_C = (-1, 2, 0)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE			Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		5.6	2.5	-0.6		0.1081	0.1742	0.0269		90.8	92.9	92.1		
		Exchangeable	-13.9	5.3	2.1	1.8	0.0130	0.1007	0.1488	0.0275	81.1	91.8	93.2	91.5		
	GEE	AR1	-0.0	5.4	2.2	0.5	0.0223	0.1011	0.1552	0.0262	81.4	90.8	93.0	92.5		
		Independent		-0.3	-1.8	-76.9		0.0327	0.0522	0.0346		96.1	94.7	71.0		
		Exchangeable	-8.7	1.0	1.0	-4.4	0.0067	0.0338	0.0553	0.0133	83.8	96.4	95.2	93.7		
		AR1	-10.3	-0.8	0.2	-36.6	0.0097	0.0335	0.0540	0.0186	81.9	96.3	95.1	87.5		
	BPL	Exchangeable/AR1	-20.3	-0.4	-0.7	-52.8	0.0142	0.0341	0.0551	0.0577	90.6	95.5	95.0	85.9		
	ML	Independent		1.3	-7.6	-143.8		0.0320	0.0585	0.1035		95.2	92.1	42.9		
		Exchangeable	-2.7	0.9	0.4	1.5	0.0108	0.0346	0.0558	0.0453	89.8	96.0	95.1	89.5		
		AR1	-17.8	-0.4	-0.9	-52.5	0.0148	0.0337	0.0550	0.0527	89.1	95.9	94.4	86.0		
		WGEE	Independent		2.6	2.5	-1.0		0.0449	0.0765	0.0109		93.4	93.4	93.2	
	Exchangeable		-6.6	2.5	2.5	-1.0	0.0058	0.0419	0.0647	0.0107	87.2	93.3	94.2	93.0		
	GEE	AR1	11.0	2.5	2.5	-1.0	0.0104	0.0420	0.0687	0.0106	86.6	93.4	93.5	93.2		
		Independent		-0.0	-1.4	-77.2		0.0142	0.0211	0.0287		95.6	95.1	37.5		
		Exchangeable	-9.3	1.3	1.4	-5.5	0.0030	0.0146	0.0215	0.0053	82.6	95.9	95.3	91.9		
		AR1	-10.7	-0.7	0.5	-38.2	0.0045	0.0144	0.0210	0.0113	79.4	95.5	95.2	78.1		
	BPL	Exchangeable/AR1	-18.2	-0.3	-0.4	-57.7	0.0068	0.0144	0.0214	0.0310	90.5	94.9	94.7	80.2		
	ML	Independent		1.7	-7.4	-146.2		0.0139	0.0269	0.0948		94.9	88.7	11.2		
Exchangeable		-3.9	1.0	0.7	-2.7	0.0037	0.0145	0.0213	0.0158	92.6	95.7	95.9	92.6			
		AR1	-16.2	-0.3	-0.6	-56.6	0.0064	0.0143	0.0213	0.0294	90.7	95.2	94.4	81.9		
		N= 500														

Table 5.7: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **exchangeable correlation** ($\alpha = 0.4$) and **MAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 2, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE			Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		2.1	0.9	-6.7		0.0852	0.1518	0.0301		94.1	93.7	94.0		
		Exchangeable	-4.8	1.9	0.6	-4.0	0.0127	0.0763	0.1259	0.0297	90.5	94.4	94.1	92.3		
	GEE	AR1	7.1	2.0	0.7	-4.1	0.0161	0.0765	0.1277	0.0309	87.8	94.6	94.3	92.0		
		Independent		-0.2	-2.0	-164.8		0.0422	0.0688	0.1289		95.0	94.7	35.7		
	Exchangeable	Exchangeable	-15.9	1.4	1.2	-26.0	0.0115	0.0440	0.0709	0.0203	75.9	94.6	94.7	90.9		
		AR1	-7.2	0.4	1.2	-46.1	0.0113	0.0435	0.0707	0.0290	82.0	94.6	94.9	86.2		
	BPL	Exchangeable/AR1	-21.9	0.2	-0.5	-100.7	0.0298	0.0426	0.0718	0.1811	84.3	94.9	93.6	80.1		
		ML	Independent		-1.6	-16.1	-355.0		0.0384	0.0870	0.5353		94.0	84.8	3.8	
Exchangeable	Exchangeable		-6.1	2.6	0.8	12.6	0.0165	0.0478	0.0695	0.0966	86.3	93.3	94.9	86.3		
	AR1	-13.3	2.0	0.6	-81.1	0.0263	0.0423	0.0700	0.1315	82.5	95.1	93.8	85.6			
N= 500																
	WGEE	Independent		1.3	0.8	-0.7		0.0333	0.0597	0.0104		95.3	94.8	96.0		
		Exchangeable	-1.2	1.2	0.6	0.8	0.0050	0.0295	0.0492	0.0103	93.4	95.5	94.6	94.2		
	GEE	AR1	11.8	1.3	0.7	0.8	0.0083	0.0296	0.0500	0.0106	86.4	95.8	94.8	94.3		
		Independent		-1.0	-2.7	-166.0		0.0171	0.0276	0.1183		93.9	93.4	3.0		
	Exchangeable	Exchangeable	-15.3	0.4	0.2	-26.1	0.0069	0.0171	0.0270	0.0096	62.2	94.7	94.3	87.4		
		AR1	-6.2	-0.6	0.2	-47.1	0.0049	0.0170	0.0270	0.0168	79.5	95.1	94.1	75.7		
	BPL	Exchangeable/AR1	-17.6	-0.9	-1.5	-112.5	0.0155	0.0170	0.0282	0.1121	85.7	94.5	92.8	76.6		
		ML	Independent		-2.2	-17.3	-362.4		0.0159	0.0540	0.5372		93.2	72.3	0.0	
Exchangeable	Exchangeable		-1.4	0.4	-0.4	-1.6	0.0054	0.0171	0.0254	0.0306	90.9	95.2	94.3	91.0		
	AR1	-7.7	0.2	-0.2	-76.7	0.0100	0.0169	0.0268	0.0700	90.0	94.8	93.2	87.4			

Table 5.8: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.4$) and MAR drop-outs ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 2, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		6.4	1.1	-7.8		0.1510	0.2837	0.0259		90.1	89.0	90.2		
		Exchangeable	-8.6	6.5	1.5	0.3	0.0148	0.1282	0.2083	0.0240	83.8	89.8	91.0	91.4		
	GEE	AR1	11.6	5.9	1.1	-2.9	0.0223	0.1285	0.2238	0.0237	76.9	89.1	90.5	91.8		
		Independent		1.4	-1.6	-154.5		0.0448	0.0650	0.1075		94.5	94.7	19.3		
	Exchangeable	AR1	-13.6	3.1	2.9	-20.0	0.0109	0.0460	0.0664	0.0126	74.3	94.1	95.3	89.9		
		Exchangeable/AR1	-1.1	1.0	2.8	-51.3	0.0136	0.0452	0.0669	0.0267	74.3	94.5	95.0	76.3		
	BPL	Exchangeable/AR1	-20.8	1.1	0.8	-104.6	0.0218	0.0455	0.0670	0.1007	88.0	93.4	93.0	76.6		
	ML	Independent		4.4	-14.7	-275.6		0.0422	0.0808	0.3215		92.2	86.2	2.4		
		Exchangeable	0.1	3.1	1.5	5.8	0.0097	0.0464	0.0612	0.0299	89.5	93.5	94.4	92.0		
	AR1	Exchangeable	-12.1	2.3	1.5	-84.9	0.0195	0.0465	0.0638	0.0749	86.7	93.5	94.2	82.7		
	N= 500															
	WGEE	Independent		1.9	1.1	-0.7		0.0505	0.0994	0.0095		94.5	94.0	93.0		
		Exchangeable	-2.5	1.8	0.8	2.2	0.0066	0.0426	0.0725	0.0087	88.9	95.1	94.4	93.0		
	GEE	AR1	20.6	1.8	0.8	1.3	0.0158	0.0425	0.0782	0.0086	70.9	94.6	94.3	92.6		
		Independent		0.3	-1.9	-153.5		0.0156	0.0248	0.0985		95.7	95.0	0.3		
	Exchangeable	AR1	-13.0	1.7	2.0	-18.1	0.0062	0.0166	0.0268	0.0056	63.8	95.6	95.0	86.9		
		Exchangeable/AR1	0.0	-0.4	1.9	-50.9	0.0060	0.0164	0.0270	0.0167	71.6	95.8	94.8	60.7		
	BPL	Exchangeable/AR1	-18.1	-0.1	-0.1	-112.1	0.0112	0.0160	0.0253	0.0690	85.2	95.4	94.3	62.9		
	ML	Independent		3.4	-16.3	-279.4		0.0152	0.0482	0.3185		93.7	73.1	0.0		
		Exchangeable	0.2	0.8	0.3	2.9	0.0031	0.0148	0.0222	0.0086	91.1	95.1	94.7	93.2		
	AR1	Exchangeable	-9.6	0.5	1.1	-82.1	0.0065	0.0157	0.0238	0.0420	88.2	95.4	95.2	72.7		

N= 500

Table 5.9: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **exchangeable correlation** ($\alpha = 0.2$) and **MAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 1, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE			Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		0.1	0.1	1.0		0.0426	0.0613	0.0185		95.6	96.3	95.4		
		Exchangeable	-1.2	-0.0	-0.2	1.1	0.0054	0.0416	0.0606	0.0181	95.0	95.6	96.2	94.6		
	GEE	AR1	8.7	0.1	-0.1	1.6	0.0086	0.0417	0.0600	0.0189	91.2	95.6	96.5	95.1		
		Independent		-0.2	-0.7	-34.2		0.0354	0.0531	0.0211		96.0	95.8	91.4		
		Exchangeable	-6.7	0.1	-0.3	-0.6	0.0045	0.0354	0.0527	0.0157	93.3	96.0	95.9	95.4		
		AR1	3.1	-0.4	-0.3	-8.7	0.0068	0.0354	0.0528	0.0163	89.6	95.9	96.1	94.7		
	BPL	Exchangeable/AR1	-11.1	-0.5	-1.0	-19.0	0.0083	0.0353	0.0543	0.0730	92.1	96.1	94.7	90.1		
	ML	Independent		-0.3	-9.5	-156.1		0.0333	0.0635	0.1436		95.6	89.0	50.5		
Exchangeable		-5.1	-0.2	-0.8	1.3	0.0059	0.0359	0.0546	0.0556	92.5	96.1	95.2	92.8			
	AR1	-8.7	-0.2	-0.7	-17.6	0.0085	0.0352	0.0540	0.0678	91.7	96.2	94.7	91.1			
N= 500																
	WGEE	Independent		0.1	-0.1	-0.7		0.0185	0.0255	0.0076		94.7	94.3	94.4		
		Exchangeable	0.1	0.1	-0.1	-0.5	0.0021	0.0183	0.0253	0.0076	95.6	94.6	94.6	94.6		
	GEE	AR1	11.2	0.1	-0.1	-0.4	0.0036	0.0183	0.0250	0.0079	90.7	94.7	94.7	94.3		
		Independent		-0.3	-0.8	-36.2		0.0158	0.0215	0.0121		93.9	95.1	84.2		
		Exchangeable	-5.9	0.1	-0.1	-2.6	0.0019	0.0158	0.0213	0.0065	92.4	94.1	95.4	93.8		
		AR1	4.9	-0.5	-0.2	-10.9	0.0027	0.0158	0.0213	0.0071	90.8	93.8	95.7	93.3		
	BPL	Exchangeable/AR1	-5.2	-0.4	-0.7	-21.4	0.0035	0.0160	0.0219	0.0285	93.0	93.8	94.6	92.8		
	ML	Independent		-0.6	-10.9	-171.9		0.0149	0.0359	0.1375		93.4	82.6	25.9		
Exchangeable		-1.5	0.0	-0.3	-0.5	0.0024	0.0160	0.0212	0.0199	93.4	93.6	95.4	94.1			
	AR1	-2.8	-0.3	-0.6	-19.1	0.0034	0.0160	0.0216	0.0260	93.0	93.9	94.9	93.3			

Table 5.10: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.2$) and MAR drop-outs ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 1, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE			Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
WGEE	Independent Exchangeable AR1	Independent Exchangeable AR1		1.8	1.6	0.2		0.0486	0.0687	0.0103		95.0	95.2	94.4		
			-3.5	1.7	1.5	0.1	0.0044	0.0464	0.0645	0.0101	93.4	95.1	95.2	94.5		
			14.2	1.8	1.6	0.6	0.0094	0.0469	0.0661	0.0104	87.6	95.1	95.0	94.5		
				1.1	0.8	-33.8		0.0365	0.0508	0.0133		94.2	95.2	89.0		
GEE	Independent Exchangeable AR1	Independent Exchangeable AR1	-7.0	1.6	1.7	-0.9	0.0034	0.0367	0.0507	0.0085	92.2	94.8	95.5	95.0		
			8.1	0.6	1.6	-13.4	0.0067	0.0366	0.0510	0.0097	88.7	94.5	95.0	93.1		
			-10.8	0.9	0.9	-23.3	0.0065	0.0374	0.0514	0.0297	92.5	94.4	94.2	91.9		
				3.5	-8.2	-113.4		0.0351	0.0628	0.0731		93.2	87.8	49.0		
BPL ML	Independent Exchangeable AR1	Independent Exchangeable AR1	-5.1	1.5	1.1	0.4	0.0042	0.0372	0.0510	0.0224	93.3	93.9	95.3	92.9		
			-8.5	1.0	1.0	-22.9	0.0061	0.0370	0.0507	0.0274	92.1	94.3	94.0	92.1		
				0.9	0.7	0.8		0.0191	0.0279	0.0044		94.8	95.1	94.0		
			-0.5	0.9	0.7	0.9	0.0017	0.0181	0.0261	0.0043	94.6	95.3	95.2	94.3		
WGEE	Independent Exchangeable AR1	Independent Exchangeable AR1	18.7	0.8	0.7	0.7	0.0047	0.0182	0.0267	0.0044	82.9	95.2	95.2	94.2		
				-0.1	-0.4	-35.1		0.0137	0.0213	0.0085		94.8	93.8	77.2		
			-4.9	0.4	0.6	-1.4	0.0015	0.0135	0.0210	0.0035	93.2	95.9	95.1	93.7		
			11.0	-0.8	0.3	-14.8	0.0032	0.0136	0.0212	0.0045	85.3	95.5	94.8	90.6		
GEE	Independent Exchangeable AR1	Independent Exchangeable AR1	-4.7	-0.6	-0.0	-22.6	0.0027	0.0138	0.0217	0.0133	92.7	95.1	93.5	91.0		
				2.6	-10.0	-120.3		0.0134	0.0353	0.0670		93.7	80.9	21.9		
			-0.2	0.3	0.6	3.1	0.0015	0.0135	0.0209	0.0081	94.6	95.8	94.4	93.5		
			-1.8	-0.4	0.2	-21.2	0.0025	0.0135	0.0215	0.0119	91.9	95.3	93.4	92.1		
N= 500																

Table 5.11: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **exchangeable correlation** ($\alpha = 0.4$) and **MAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 1, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE			Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		1.5	1.4	0.9		0.0503	0.0834	0.0134		95.1	94.9	96.6		
		Exchangeable	-1.1	1.4	1.3	1.9	0.0052	0.0484	0.0808	0.0125	95.4	95.9	93.7	96.2		
	GEE	AR1	11.7	1.4	1.4	1.8	0.0086	0.0483	0.0806	0.0138	88.2	95.7	94.5	95.7		
		Independent		0.9	0.3	-68.6		0.0417	0.0695	0.0318		95.6	94.8	82.0		
		Exchangeable	-6.7	1.4	1.3	-2.2	0.0052	0.0415	0.0674	0.0108	92.3	95.9	94.5	96.2		
		AR1	7.0	0.9	1.4	-10.1	0.0066	0.0415	0.0681	0.0117	88.8	96.0	94.2	96.0		
BPL	Exchangeable/AR1	-13.8	0.5	0.0	-45.6	0.0135	0.0413	0.0681	0.1038	90.0	95.6	93.4	84.5			
	ML	Independent		-0.3	-16.2	-260.3		0.0371	0.0850	0.2910		94.4	80.1	3.5		
		Exchangeable	-7.9	2.9	1.0	1.6	0.0075	0.0424	0.0660	0.0482	92.3	94.9	93.9	90.9		
		WGEE	AR1	-8.4	3.4	2.5	-37.8	0.0106	0.0406	0.0660	0.0729	91.1	95.9	93.5	89.4	
WGEE		Independent		0.0	-0.3	-1.0		0.0193	0.0304	0.0058		95.7	95.2	95.6		
		Exchangeable	-0.8	0.0	-0.3	-0.4	0.0022	0.0183	0.0289	0.0054	95.1	96.1	95.5	96.1		
GEE		AR1	12.4	0.1	-0.2	-0.6	0.0051	0.0185	0.0289	0.0061	79.5	95.9	95.3	95.9		
		Independent		-0.3	-1.1	-72.2		0.0163	0.0257	0.0266		95.0	95.1	53.1		
	Exchangeable	-6.6	0.3	0.0	-4.8	0.0025	0.0161	0.0245	0.0049	87.8	95.7	94.8	94.2			
	AR1	7.5	-0.2	0.1	-12.9	0.0033	0.0163	0.0250	0.0058	85.8	94.5	95.2	93.4			
BPL	Exchangeable/AR1	-8.2	-0.3	-0.8	-48.7	0.0058	0.0162	0.0257	0.0532	92.1	94.2	94.3	86.7			
	ML	Independent		-1.5	-18.6	-268.4		0.0144	0.0547	0.2941		93.5	64.7	0.1		
		Exchangeable	-2.1	0.9	0.1	0.2	0.0022	0.0152	0.0227	0.0154	94.3	95.8	94.8	93.3		
			AR1	-0.8	1.8	1.8	-31.5	0.0035	0.0159	0.0243	0.0306	90.4	95.0	94.1	93.5	

Table 5.12: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), exchangeable correlation ($\alpha = 0.4$) and MAR drop-outs ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 1, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		1.4	1.9	0.5		0.0561	0.0994	0.0093		94.7	94.1	94.9	94.9	
		Exchangeable	-1.8	1.2	1.6	1.1	0.0043	0.0501	0.0855	0.0084	95.1	94.5	94.2	94.9		
	GEE	AR1	21.5	1.4	1.7	1.3	0.0136	0.0508	0.0882	0.0093	71.7	95.5	94.2	94.2		
		Independent		0.5	0.4	-69.3		0.0404	0.0714	0.0276		94.9	93.9	65.4		
	Exchangeable		-6.1	1.3	1.9	-2.7	0.0041	0.0397	0.0674	0.0067	91.7	94.8	94.4	95.4		
		AR1	16.4	0.2	1.9	-14.9	0.0098	0.0400	0.0691	0.0089	74.7	95.5	93.9	93.2		
	BPL	Exchangeable/AR1	-11.6	0.4	0.5	-54.2	0.0097	0.0393	0.0680	0.0491	92.6	94.8	92.1	86.9		
		ML	Independent		4.1	-18.8	-202.5		0.0357	0.0931	0.1734		92.8	75.7	2.3	
Exchangeable			-0.2	3.2	1.2	11.5	0.0039	0.0373	0.0584	0.0162	92.8	94.2	93.7	90.8		
	AR1	-5.6	2.3	3.1	-42.0	0.0060	0.0397	0.0668	0.0303	91.2	94.1	90.6	90.0			
N= 500																
	WGEE	Independent		1.5	1.0	0.3		0.0213	0.0336	0.0032		94.9	95.6	96.4		
		Exchangeable	-1.0	1.4	0.9	0.7	0.0018	0.0193	0.0293	0.0032	95.2	95.3	95.4	95.5		
	GEE	AR1	23.0	1.3	0.8	0.6	0.0110	0.0196	0.0304	0.0033	44.9	95.7	95.1	95.5		
		Independent		0.6	-0.4	-69.4		0.0156	0.0240	0.0224		95.2	94.9	32.4		
	Exchangeable		-6.1	1.3	1.1	-3.5	0.0022	0.0155	0.0233	0.0026	86.9	95.6	95.6	95.2		
		AR1	17.2	0.0	1.0	-15.9	0.0072	0.0156	0.0240	0.0038	54.3	94.9	94.6	91.0		
	BPL	Exchangeable/AR1	-7.6	0.4	0.1	-52.0	0.0042	0.0158	0.0237	0.0278	91.4	94.4	93.5	82.3		
		ML	Independent		4.6	-20.1	-203.8		0.0153	0.0582	0.1692		91.4	57.8	0.0	
Exchangeable			-0.7	1.6	-0.0	6.7	0.0012	0.0123	0.0176	0.0051	94.5	94.9	94.8	93.4		
	AR1	-4.6	3.0	3.5	-42.4	0.0017	0.0150	0.0222	0.0165	93.1	95.2	93.9	85.0			

Drop-out mechanism: NI

For NI drop-out mechanisms, we first considered two drop-out models which did not depend on the responses immediately preceding the drop-out by setting $\gamma_P = 0$. For the first model, we chose $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 0.5)$. We then increased γ_C for the second drop-out model so that $(\gamma_0, \gamma_P, \gamma_C) = (-1, 0, 1)$. The results for the first model are summarized in Tables 5.13 - 5.16 with the first two tables showing the results for $\alpha = 0.2$ with $T = 3$ and $T = 4$ respectively, and the last two tables showing the results for $\alpha = 0.4$ with $T = 3$ and $T = 4$ respectively. In the same way, Tables 5.17 - 5.20 show the results for the second drop-out model.

We also considered two drop-out models, which depended on both the responses immediately preceding the drop-out and the responses at the time of drop-out by setting $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 0.5)$ for the first drop-out model. The results are shown in Tables 5.21 - 5.24. We increased γ_C for the second drop-out model so that $(\gamma_0, \gamma_P, \gamma_C) = (-1, 1, 1)$. The results are summarized in Tables 5.25 - 5.28.

For NI drop-out, both the GEE and wGEE methods grossly underestimated β_T . When $\gamma_P \neq 0$ and $\gamma_C \neq 0$, estimators of both β_0 and β_G even showed a slight bias. For the wGEE, the differences in the relative bias for different ‘working’ correlations were negligible. For the GEE, the model with independence assumption gave the relative bias noticeably higher than that obtained from models with other correlation assumption. The dramatic contrast can be observed in Table 5.28, in which the GEE with independence assumption for $N = 500$ gives a percentage relative bias of -221.9% for β_T , almost twice as large as the relative bias found with GEE with the correct ‘working’ correlation. For both GEE and wGEE, the increase in α from 0.2 to 0.4 did not affect the bias. The increase in γ_C from 0.5 to 1, however, resulted in more than doubling the bias for β_T in most cases. The increase in the number of responses reduced the bias. The sample size did not affect the bias, but the coverage probability for β_T was dramatically reduced with the increased sample size. In Table 5.19, for example, the coverage probability for β_T from the GEE with independence assumption decreases from 53.0% to

13.2% when the sample size increased from 200 to 500. As expected, the ML and bivariate maximum pseudo-likelihood methods provided smaller bias and better coverages, as compared to the GEE and wGEE methods.

For the drop-out mechanisms with $\gamma_P = 0$, the bivariate maximum pseudo-likelihood for $T = 3$ showed some small bias for the time coefficient with $N = 200$, which then decreased close to 0% with $N = 500$. However, with an increase in the number of responses to $T = 4$, we observed a small relative bias for β_T , which increases with α . For example, Tables 5.18 and 5.20 show that an increase in α from 0.2 to 0.4 resulted in an increase in the relative bias for β_T from -11.7% to -22.2%. When $\gamma_P \neq 0$, the relative bias for β_T became noticeable and was observed to vary positively with α whereas the increase in γ_C had negligible impact on the relative bias. For example in Table 5.26, the percentage relative bias for β_T with $\alpha = 0.2$ was -39.8%, whereas in Table 5.28 with $\alpha = 0.4$, it increases to -86.2%. The bivariate maximum pseudo-likelihood was slightly less efficient than the ML with exchangeable ‘working’ correlation for β_0 and β_G but was much less efficient for β_T .

The ML with independence assumption was highly sensitive to the increase in α or γ_C . The increase in α from 0.2 to 0.4 nearly doubled the percentage relative bias for β_T . The increase in the number of responses decreased the relative bias. The ML with exchangeable ‘working’ correlation showed some small relative bias with $N = 200$, which decreased to near 0% with an increase in the sample size and was the most efficient method.

The ML with $AR(1)$ ‘working’ correlation behaved much like the bivariate maximum pseudo-likelihood under NI drop-out mechanisms but was slightly more efficient for β_T .

Table 5.13: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **exchangeable correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T			
N= 200																	
	WGEE	Independent		0.8	0.9	-55.1		0.0385	0.0587	0.0296		96.9	94.8	85.5			
		Exchangeable	-2.6	0.8	0.9	-55.4	0.0044	0.0388	0.0605	0.0291	94.7	97.1	94.8	84.5			
	GEE	AR1	7.4	0.5	0.9	-56.7	0.0070	0.0390	0.0593	0.0309	91.4	96.9	94.5	84.5			
		Independent		1.8	1.3	-53.1		0.0349	0.0546	0.0270		96.9	95.2	85.4			
		Exchangeable	-3.2	1.9	1.3	-49.7	0.0042	0.0349	0.0543	0.0251	93.9	96.6	95.2	86.0			
		AR1	9.3	1.6	1.3	-51.4	0.0068	0.0351	0.0542	0.0261	90.5	96.9	95.3	85.0			
	BPL	Exchangeable/AR1	-8.1	0.2	1.0	-1.5	0.0048	0.0356	0.0543	0.0523	93.9	96.7	94.3	91.7			
	ML	Independent		-1.5	-2.0	-22.9		0.0388	0.0584	0.0815		94.1	89.5	71.9			
		Exchangeable	-6.2	0.2	1.3	-4.5	0.0038	0.0356	0.0549	0.0441	94.5	96.5	95.0	92.8			
		AR1	-5.5	0.5	1.3	-2.9	0.0049	0.0355	0.0543	0.0505	94.3	96.7	94.3	92.0			
	N= 500																
		WGEE	Independent		1.3	1.0	-54.1		0.0162	0.0236	0.0180		95.3	94.5	72.8		
Exchangeable			-0.8	1.3	1.0	-54.2	0.0016	0.0163	0.0239	0.0179	96.0	95.6	94.7	71.4			
		AR1	9.6	1.0	1.0	-55.6	0.0029	0.0162	0.0235	0.0189	91.3	95.5	94.6	72.5			
		Independent		2.3	1.1	-52.5		0.0151	0.0222	0.0169		95.8	94.5	71.6			
	GEE	Exchangeable	-1.5	2.3	1.2	-48.4	0.0015	0.0150	0.0216	0.0150	94.7	95.4	94.9	73.9			
		AR1	11.4	2.0	1.2	-50.3	0.0030	0.0150	0.0218	0.0158	88.8	95.3	94.6	73.3			
	BPL	Exchangeable/AR1	-2.9	0.9	1.0	-5.2	0.0019	0.0149	0.0217	0.0197	94.1	95.8	94.5	94.6			
	ML	Independent		0.1	-1.3	-35.7		0.0164	0.0251	0.0501		94.1	88.3	77.0			
		Exchangeable	-1.7	0.9	1.1	-3.6	0.0014	0.0148	0.0214	0.0151	96.1	95.7	94.9	95.0			
			AR1	-0.9	1.1	1.1	-6.0	0.0019	0.0148	0.0217	0.0188	93.8	95.5	94.6	95.2		

Table 5.14: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **exchangeable correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C = (-1, 0, 0.5)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability							
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T				
N= 200																		
	WGEE	Independent		2.3	0.1	-38.2		0.0403	0.0570	0.0141		94.9	94.3	86.3				
		Exchangeable	-3.9	2.3	0.0	-37.9	0.0032	0.0394	0.0568	0.0137	93.4	95.1	94.5	87.4				
	GEE	AR1	13.3	1.8	0.1	-40.0	0.0071	0.0394	0.0558	0.0147	87.8	94.8	94.0	86.0				
		Independent		3.5	0.9	-38.1		0.0351	0.0515	0.0133		95.3	94.6	85.2				
		Exchangeable	-3.4	3.5	0.9	-34.1	0.0030	0.0345	0.0505	0.0119	93.4	95.4	94.4	86.3				
		AR1	16.4	2.9	1.0	-37.6	0.0072	0.0343	0.0505	0.0130	86.9	95.6	94.3	85.9				
	BPL	Exchangeable/AR1	-6.6	0.8	0.9	-5.9	0.0040	0.0356	0.0513	0.0206	94.4	94.5	93.0	93.7				
		Independent		0.0	-1.8	-24.1		0.0401	0.0555	0.0329		91.4	88.3	78.8				
	ML	Exchangeable	-3.3	0.9	0.8	0.4	0.0028	0.0339	0.0497	0.0166	93.2	95.5	93.9	93.6				
		AR1	-3.0	1.2	1.2	-6.8	0.0041	0.0349	0.0503	0.0195	92.4	94.8	93.2	94.0				
			WGEE	Independent		1.1	-1.3	-39.7		0.0168	0.0228	0.0095		94.2	95.4	71.6		
				Exchangeable	-0.9	1.1	-1.3	-39.1	0.0012	0.0164	0.0223	0.0092	94.7	93.8	95.0	71.8		
GEE	AR1		18.5	0.6	-1.3	-41.6	0.0038	0.0164	0.0222	0.0102	83.3	94.0	95.7	70.4				
	Independent			2.3	-0.7	-40.1		0.0150	0.0199	0.0094		95.0	95.4	69.2				
	ML		Exchangeable	-1.6	2.4	-0.6	-35.4	0.0012	0.0149	0.0192	0.0079	95.2	94.7	95.8	73.7			
			AR1	19.8	1.7	-0.7	-39.3	0.0039	0.0147	0.0194	0.0092	79.0	94.5	95.3	69.9			
		BPL	Exchangeable/AR1	-1.5	-0.4	-0.6	-4.8	0.0015	0.0152	0.0196	0.0078	93.3	94.8	93.7	95.4			
			Independent		-0.5	-2.0	-22.0		0.0170	0.0222	0.0157		92.0	88.3	85.8			
	ML	Exchangeable	-0.7	-0.2	-0.5	0.8	0.0011	0.0146	0.0190	0.0064	93.9	94.4	94.9	93.7				
		AR1	1.3	-0.1	-0.3	-5.7	0.0016	0.0148	0.0193	0.0073	92.5	94.7	94.0	95.7				
		N= 500																

Table 5.15: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **exchangeable correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		-0.3	0.3	-51.3		0.0448	0.0714	0.0233		95.8	94.6	87.9		
		Exchangeable	-0.4	-0.2	0.2	-50.5	0.0044	0.0451	0.0747	0.0223	95.3	95.8	94.6	85.4		
		AR1	12.5	-0.8	0.2	-54.8	0.0080	0.0451	0.0733	0.0259	86.2	95.9	94.4	86.0		
		Independent		1.6	1.1	-58.4		0.0419	0.0690	0.0254	96.2	95.1	82.4			
	GEE	Exchangeable	-1.5	1.7	1.2	-43.4	0.0043	0.0416	0.0673	0.0184	94.6	96.1	94.8	86.5		
		AR1	13.7	1.1	1.3	-47.7	0.0083	0.0415	0.0678	0.0202	84.9	96.4	95.0	85.2		
		Exchangeable/AR1	-9.1	0.2	0.7	-3.5	0.0066	0.0421	0.0675	0.0719	91.7	95.5	94.4	88.8		
		Independent		-4.2	-9.1	-69.0		0.0498	0.0689	0.1653	87.3	84.5	18.7			
	BPL	Exchangeable	-8.5	0.4	0.4	-5.5	0.0043	0.0422	0.0684	0.0283	94.5	94.3	92.1	93.5		
		AR1	-6.3	2.7	3.8	-14.3	0.0050	0.0425	0.0713	0.0499	94.1	95.2	92.5	92.1		
		Independent		-0.2	-0.7	-51.3		0.0180	0.0290	0.0155		94.8	95.1	73.4		
		Exchangeable	-0.2	-0.2	-0.9	-50.4	0.0016	0.0179	0.0297	0.0148	96.0	94.4	94.6	69.9		
	GEE	AR1	13.1	-0.8	-0.9	-54.9	0.0047	0.0180	0.0290	0.0175	75.4	94.5	95.1	70.0		
		Independent		1.5	0.0	-58.4		0.0171	0.0280	0.0181		94.8	94.8	63.6		
		Exchangeable	-1.1	1.6	-0.0	-43.4	0.0015	0.0168	0.0268	0.0117	96.1	95.4	95.2	73.3		
		AR1	14.5	0.9	-0.1	-47.9	0.0052	0.0169	0.0269	0.0135	70.6	95.1	94.7	70.4		
BPL	Exchangeable/AR1	-3.4	0.3	-0.2	-1.7	0.0021	0.0170	0.0269	0.0281	94.2	94.5	94.1	92.4			
	Independent		-3.8	-12.2	-101.6		0.0255	0.0405	0.1674		81.1	74.4	4.6			
	Exchangeable	-3.1	0.8	0.9	2.2	0.0012	0.0152	0.0247	0.0106	96.9	95.3	94.6	94.8			
	AR1	2.5	3.2	3.1	-11.4	0.0018	0.0168	0.0265	0.0172	89.8	94.8	93.8	96.1			
N= 500																

Table 5.16: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **exchangeable correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias			MSE			Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T	
N= 200															
	WGEE	Independent		1.0	-0.2	-39.6		0.0476	0.0712	0.0139		94.1	95.2	84.9	
		Exchangeable	-0.7	1.3	-0.1	-37.4	0.0033	0.0446	0.0675	0.0123	94.1	94.6	96.2	85.3	
		AR1	22.9	0.5	0.1	-42.7	0.0131	0.0452	0.0685	0.0150	64.8	94.4	95.6	83.6	
	GEE	Independent		2.7	0.2	-49.2		0.0419	0.0637	0.0173		95.3	96.1	78.5	
		Exchangeable	-1.6	3.1	0.6	-33.4	0.0032	0.0403	0.0593	0.0105	94.3	95.2	96.3	84.8	
		AR1	24.3	2.0	0.7	-40.8	0.0139	0.0407	0.0615	0.0134	58.3	95.9	96.2	82.3	
BPL ML	Exchangeable/AR1	-5.4	1.0	0.5	-12.1	0.0044	0.0418	0.0611	0.0317	93.7	95.1	93.7	92.3		
	Independent		-3.7	-11.4	-61.8		0.0615	0.0728	0.0832		80.3	79.3	26.2		
	Exchangeable	-2.6	4.3	0.5	12.3	0.0025	0.0358	0.0507	0.0107	95.4	93.6	93.4	92.3		
	AR1	-2.0	3.6	3.7	-14.3	0.0029	0.0388	0.0555	0.0205	90.1	94.9	92.7	92.5		
N= 500															
	WGEE	Independent		-0.0	-0.2	-42.4		0.0182	0.0280	0.0101		95.5	94.5	64.6	
		Exchangeable	-0.4	0.1	-0.4	-39.8	0.0014	0.0168	0.0264	0.0089	94.7	95.9	94.5	65.4	
		AR1	23.8	-0.9	-0.3	-45.7	0.0110	0.0168	0.0262	0.0114	32.7	95.5	94.5	60.9	
	GEE	Independent		1.7	0.3	-52.5		0.0155	0.0248	0.0138		95.8	94.9	48.2	
		Exchangeable	-1.4	1.9	0.4	-35.4	0.0013	0.0148	0.0229	0.0073	95.2	95.9	95.6	68.5	
		AR1	25.0	0.7	0.4	-43.5	0.0117	0.0146	0.0232	0.0101	26.0	96.9	96.0	58.5	
BPL ML	Exchangeable/AR1	-1.9	-0.2	0.4	-15.4	0.0015	0.0151	0.0235	0.0132	94.1	96.1	94.1	94.2		
	Independent		-2.6	-13.3	-88.9		0.0366	0.0413	0.0871		71.5	70.7	9.5		
	Exchangeable	-5.7	1.3	2.1	6.5	0.0012	0.0114	0.0174	0.0039	90.6	98.0	95.0	93.6		
	AR1	-1.9	3.9	5.4	-20.0	0.0008	0.0149	0.0237	0.0094	93.5	95.8	92.6	92.6		

Table 5.17: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **exchangeable correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability			
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T
N= 200	WGEE	Independent		2.2	0.0	-123.8		0.0426	0.0585	0.0806		95.2	96.3	53.7
		Exchangeable	-4.2	2.1	-0.2	-124.0	0.0052	0.0424	0.0589	0.0803	93.5	95.0	96.2	51.6
		AR1	4.7	1.5	-0.2	-127.2	0.0080	0.0422	0.0582	0.0843	90.6	95.5	96.2	51.2
	GEE	Independent		4.1	1.0	-123.5		0.0414	0.0560	0.0795		94.7	96.1	53.0
		Exchangeable	-5.7	4.1	1.1	-114.3	0.0050	0.0416	0.0559	0.0705	92.5	94.6	96.5	55.9
		AR1	5.0	3.5	1.0	-118.8	0.0077	0.0411	0.0560	0.0748	90.4	95.2	96.2	53.2
	BPL	Exchangeable/AR1	-7.3	1.3	1.0	-10.1	0.0059	0.0412	0.0574	0.0668	93.6	94.9	95.0	91.7
		Independent		0.4	-1.1	-62.6		0.0431	0.0611	0.1175		93.2	90.3	76.2
		Exchangeable	-5.8	1.0	0.9	-10.1	0.0044	0.0401	0.0562	0.0566	94.2	95.9	95.9	92.2
	ML	AR1	-5.3	1.5	1.2	-13.7	0.0059	0.0408	0.0563	0.0658	93.3	94.9	95.4	92.0
N= 500	WGEE	Independent		1.7	-0.4	-125.1		0.0178	0.0264	0.0694		94.8	95.0	14.2
		Exchangeable	-4.4	1.7	-0.5	-125.1	0.0019	0.0178	0.0266	0.0692	94.3	94.3	94.3	13.4
		AR1	5.6	1.1	-0.5	-128.5	0.0030	0.0176	0.0262	0.0729	92.0	94.7	94.5	12.9
	GEE	Independent		3.8	0.7	-124.5		0.0178	0.0247	0.0686		94.3	94.7	13.2
		Exchangeable	-5.8	3.9	0.8	-115.0	0.0018	0.0180	0.0246	0.0593	93.1	94.3	94.7	17.8
		AR1	6.1	3.2	0.8	-119.6	0.0028	0.0174	0.0245	0.0637	92.5	94.1	94.6	14.7
	BPL	Exchangeable/AR1	-2.6	0.9	0.7	-4.0	0.0021	0.0169	0.0241	0.0290	94.4	94.8	93.8	92.5
		Independent		1.0	-0.1	-56.2		0.0180	0.0253	0.0630		93.0	91.2	81.3
		Exchangeable	-2.5	1.0	0.9	-1.4	0.0016	0.0168	0.0236	0.0232	95.1	94.6	94.4	92.6
	ML	AR1	-0.5	1.2	1.0	-6.5	0.0022	0.0169	0.0239	0.0279	94.1	94.7	93.9	92.5

Table 5.18: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **exchangeable correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability							
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T				
N= 200																		
	WGEE	Independent		5.6	-0.7	-92.7		0.0457	0.0645	0.0447		94.5	95.1	54.1				
		Exchangeable	-7.4	5.6	-0.8	-91.4	0.0041	0.0441	0.0619	0.0434		92.0	94.7	94.8	54.4			
	GEE	AR1	8.6	4.4	-0.8	-96.4	0.0080	0.0433	0.0628	0.0475		88.9	95.3	94.8	50.6			
		Independent		6.6	0.3	-98.0		0.0408	0.0564	0.0482		94.5	95.1	47.1				
	Exchangeable		-8.9	6.7	0.5	-88.2	0.0038	0.0405	0.0554	0.0404		91.3	94.4	95.1	53.0			
		AR1	7.8	5.3	0.4	-96.4	0.0072	0.0390	0.0561	0.0468		87.5	94.5	94.9	47.6			
	BPL	Exchangeable/AR1	-7.9	1.1	0.7	-12.4	0.0049	0.0373	0.0560	0.0284		93.2	94.6	93.3	93.3			
		Independent		0.9	-0.8	-44.5		0.0415	0.0596	0.0441		92.8	89.8	83.1				
	ML	Exchangeable	-4.4	1.2	0.7	1.1	0.0033	0.0366	0.0531	0.0232		93.6	95.2	95.0	92.5			
		AR1	-4.5	1.5	1.1	-14.3	0.0051	0.0369	0.0551	0.0272		92.0	95.2	93.3	94.0			
		WGEE	Independent		4.7	-1.6	-91.8		0.0191	0.0251	0.0378		93.4	94.8	16.5			
			Exchangeable	-5.0	4.7	-1.7	-90.4	0.0017	0.0184	0.0238	0.0366		93.2	93.3	95.0	16.1		
GEE		AR1	12.8	3.5	-1.7	-95.4	0.0037	0.0177	0.0242	0.0405		86.6	93.6	94.7	14.4			
		Independent		6.1	-0.3	-97.0		0.0189	0.0225	0.0415		92.1	94.5	12.0				
Exchangeable			-6.7	6.3	-0.1	-86.7	0.0016	0.0187	0.0216	0.0337		92.2	92.4	94.9	16.6			
		AR1	11.9	4.8	-0.2	-95.0	0.0033	0.0173	0.0219	0.0399		87.3	92.8	94.3	11.8			
	BPL	Exchangeable/AR1	-2.3	0.9	0.4	-11.7	0.0018	0.0158	0.0216	0.0096		93.3	93.8	93.4	94.8			
		Independent		1.5	0.1	-43.2		0.0173	0.0229	0.0200		92.2	89.9	82.7				
	ML	Exchangeable	-0.8	1.2	0.3	-0.2	0.0012	0.0152	0.0207	0.0072		94.6	94.4	94.6	94.7			
		AR1	0.8	1.4	0.7	-13.8	0.0018	0.0156	0.0214	0.0093		93.7	93.8	93.6	95.2			

Table 5.19: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **exchangeable correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C = (-1, 0, 1)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
WGEE	Independent	Independent	-0.3	-0.7	-123.9	0.0464	0.0728	0.0775			95.8	95.7	51.2			
		Exchangeable	-3.6	-0.3	-121.3	0.0056	0.0443	0.0699	0.0740	92.8	95.9	95.9	48.4			
	AR1	Independent	9.0	-1.7	-130.8	0.0081	0.0448	0.0697	0.0851	89.6	95.9	95.9	47.4			
		Exchangeable	-5.7	3.1	-143.6		0.0425	0.0672	0.0988		95.6	95.9	37.4			
GEE	Independent	Independent	8.6	1.9	-119.8	0.0076	0.0407	0.0643	0.0717	88.6	96.1	95.3	46.7			
		Exchangeable/AR1	-8.7	0.7	-8.0	0.0070	0.0407	0.0637	0.0851	92.1	95.4	94.8	88.9			
	ML	Independent	-1.6	-9.1	-177.0		0.0448	0.0734	0.3199		91.4	85.2	26.0			
		Exchangeable	-8.8	1.2	3.2	0.0047	0.0404	0.0548	0.0364	94.3	94.7	95.7	93.4			
WGEE	Independent	Independent	-5.2	3.4	-27.1	0.0058	0.0401	0.0617	0.0643	92.9	96.0	95.3	91.7			
		Exchangeable														
	AR1	Independent	-0.3	-1.8	-118.3	0.0208	0.0327	0.0629			92.9	93.5	15.8			
		Exchangeable	-3.6	-0.3	-115.7	0.0023	0.0203	0.0319	0.0599	94.9	92.9	92.6	13.8			
GEE	Independent	Independent	9.4	-1.6	-125.3	0.0039	0.0205	0.0313	0.0698	86.6	92.5	93.1	11.6			
		Exchangeable	-5.9	3.1	-138.7		0.0202	0.0306	0.0836		92.6	93.9	5.8			
	AR1	Independent	8.8	2.0	-115.0	0.0025	0.0201	0.0295	0.0497	90.7	92.9	93.8	18.8			
		Exchangeable	-4.2	0.9	-3.6	0.0037	0.0196	0.0294	0.0590	85.5	93.1	94.0	12.0			
BPL	Independent	Independent	-4.9	2.2	-240.5	0.0024	0.0198	0.0301	0.0374	94.3	92.8	92.5	92.0			
		Exchangeable/AR1	1.9	3.8	-18.6		0.0214	0.0459	0.3571		88.0	74.8	13.3			
	ML	Independent	-4.9	2.2	2.5	0.0016	0.0173	0.0279	0.0130	95.3	94.1	93.1	93.8			
		Exchangeable				0.0021	0.0195	0.0295	0.0243	90.1	92.8	92.3	95.4			
N= 500																

Table 5.20: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **exchangeable correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		2.2	-1.8	-93.5		0.0558	0.0889	0.0439		93.4	95.6	48.2		
		Exchangeable	-3.2	2.4	-2.4	-86.3	0.0047	0.0509	0.0800	0.0380	92.7	94.4	94.8	49.6		
	GEE	AR1	19.9	0.1	-2.3	-100.1	0.0128	0.0505	0.0817	0.0488	75.5	94.2	94.4	39.9		
		Independent		4.6	0.1	-119.9		0.0480	0.0750	0.0668		94.7	95.0	27.6		
	Exchangeable		-6.5	5.0	0.2	-82.8	0.0048	0.0463	0.0710	0.0349	90.0	93.9	94.2	47.4		
		AR1	17.1	2.3	0.1	-101.7	0.0106	0.0443	0.0718	0.0495	76.8	95.4	94.6	34.4		
	BPL	Exchangeable/AR1	-5.7	0.7	0.7	-29.3	0.0052	0.0454	0.0715	0.0438	92.6	94.2	92.5	91.5		
	ML	Independent		2.1	-9.9	-149.5		0.0574	0.0869	0.1745		85.5	77.6	34.7		
Exchangeable		-2.4	4.2	1.3	11.4	0.0031	0.0421	0.0592	0.0148	93.9	93.5	91.9	92.5			
	AR1		-0.8	3.4	3.2	-29.1	0.0042	0.0444	0.0675	0.0286	88.5	93.5	91.9	92.5		
	WGEE	Independent		0.9	-3.4	-91.8		0.0179	0.0314	0.0371		95.9	95.1	11.8		
		Exchangeable	-2.5	1.2	-3.6	-85.9	0.0017	0.0164	0.0280	0.0326	94.9	96.8	95.5	13.2		
	AR1	21.5	-1.0	-3.6	-98.9	0.0096	0.0166	0.0284	0.0426	48.9	96.1	95.2	6.9			
		Independent		3.8	-0.5	-119.1		0.0167	0.0252	0.0603		95.8	96.4	1.8		
	Exchangeable		-6.1	4.2	-0.0	-82.2	0.0020	0.0166	0.0234	0.0299	90.5	96.4	96.2	11.8		
		AR1	18.4	1.6	-0.2	-100.5	0.0074	0.0152	0.0236	0.0435	54.7	97.0	96.6	4.8		
BPL	Exchangeable/AR1	-2.2	-0.1	0.5	-22.2	0.0018	0.0155	0.0235	0.0166	94.5	96.0	95.0	94.3			
ML	Independent		3.7	-12.8	-181.8		0.0261	0.0472	0.1881		82.8	69.8	20.7			
	Exchangeable	-3.0	2.2	0.5	9.4	0.0011	0.0123	0.0183	0.0044	92.3	94.0	96.2	94.5			
	AR1	-0.8	3.6	4.6	-28.2	0.0011	0.0162	0.0245	0.0116	91.0	95.1	93.6	92.6			

Table 5.21: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **exchangeable correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C = (-1, 1, 0.5)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		0.3	-3.6	-66.7		0.0455	0.0697	0.0382		95.6	94.8	87.6		
		Exchangeable	-15.6	0.2	-3.6	-66.6	0.0076	0.0446	0.0679	0.0377	89.0	95.6	94.5	87.2		
		AR1	-10.7	-0.1	-3.7	-67.7	0.0105	0.0450	0.0685	0.0388	90.4	95.4	94.5	87.1		
	GEE	Independent		1.3	-1.9	-102.3		0.0389	0.0579	0.0596		95.6	95.3	68.8		
		Exchangeable	-20.6	1.8	-1.0	-66.8	0.0067	0.0395	0.0581	0.0356	86.4	95.4	95.6	83.0		
		AR1	-15.8	1.0	-1.2	-77.6	0.0086	0.0394	0.0583	0.0420	88.0	95.4	95.2	79.0		
BPL	Exchangeable/AR1	-10.4	0.3	-0.4	-17.3	0.0093	0.0400	0.0596	0.0824	91.2	95.5	94.6	90.2			
	ML	Independent		1.2	-6.1	-168.0		0.0379	0.0672	0.1761		94.8	90.3	59.6		
		Exchangeable	-4.1	0.4	-0.1	3.2	0.0072	0.0405	0.0599	0.0644	92.2	95.0	94.6	90.9		
			AR1	-8.0	0.5	-0.2	-17.4	0.0098	0.0403	0.0597	0.0791	90.9	95.1	94.5	91.1	
N= 500																
	WGEE	Independent		-0.2	-2.8	-71.0		0.0181	0.0293	0.0288		95.1	94.0	66.8		
		Exchangeable	-15.0	-0.2	-2.8	-70.9	0.0036	0.0177	0.0287	0.0284	87.8	95.3	93.7	66.9		
		AR1	-8.9	-0.5	-2.7	-72.4	0.0044	0.0179	0.0287	0.0296	89.2	95.3	93.7	66.3		
	GEE	Independent		1.4	-0.7	-105.9		0.0148	0.0220	0.0523		95.4	94.8	31.4		
		Exchangeable	-20.7	1.8	0.1	-69.7	0.0037	0.0151	0.0221	0.0267	80.5	95.2	95.1	60.5		
		AR1	-15.1	1.0	-0.1	-80.9	0.0040	0.0149	0.0221	0.0336	84.7	95.3	94.7	50.8		
BPL	Exchangeable/AR1	-5.7	0.4	0.9	-23.6	0.0040	0.0152	0.0225	0.0378	92.5	95.1	94.3	91.7			
	ML	Independent		1.3	-5.3	-178.4		0.0145	0.0283	0.1578		94.6	90.4	36.6		
		Exchangeable	-2.0	0.8	1.0	-0.3	0.0029	0.0150	0.0221	0.0277	93.3	95.1	95.3	92.1		
			AR1	-3.5	0.6	0.9	-22.2	0.0040	0.0151	0.0223	0.0353	92.7	94.9	94.2	92.3	

Table 5.22: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), **exchangeable correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C = (-1, 1, 0.5)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE			Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
WGEE	Independent	Independent		2.1	-3.4	-53.1		0.0562	0.0888		93.6	93.3	84.6			
		Exchangeable	-20.8	2.1	-3.4	-51.9	0.0072	0.0532	0.0814	0.0241	84.9	93.4	93.6	85.0		
	AR1	Independent	-13.3	1.6	-3.4	-54.2	0.0111	0.0538	0.0842	0.0253	87.3	93.5	93.8	84.2		
		Exchangeable	-23.8	2.4	-1.1	-89.5		0.0354	0.0547	0.0426		95.3	94.7	56.0		
GEE	Independent	Independent		3.0	0.1	-54.6		0.0353	0.0534	0.0225	81.4	95.1	94.3	78.6		
		Exchangeable	-18.5	1.5	-0.3	-71.8	0.0085	0.0352	0.0540	0.0315	82.0	95.0	94.2	69.2		
	AR1	Independent	-16.7	0.3	0.6	-36.0	0.0080	0.0361	0.0564	0.0401	90.9	94.4	93.8	90.1		
		Exchangeable	-5.4	1.1	1.2	-122.9		0.0352	0.0621	0.0897		94.5	89.9	58.1		
BPL	Independent	Independent		1.1	1.2	1.8		0.0054	0.0357	0.0308	92.1	95.1	93.7	91.7		
		Exchangeable	-13.0	0.6	0.9	-34.3	0.0079	0.0356	0.0560	0.0373	91.5	94.6	93.7	90.4		
	AR1	Independent		0.8	-5.4	-51.8		0.0233	0.0376	0.0161		93.4	92.2	68.7		
		Exchangeable	-14.8	0.9	-5.1	-51.0	0.0033	0.0222	0.0346	0.0157	85.4	93.8	92.3	69.6		
GEE	Independent	Independent		0.2	-5.2	-53.5	0.0049	0.0224	0.0355	0.0168	86.6	93.4	92.6	67.3		
		Exchangeable	-4.0	1.4	-2.9	-90.0		0.0146	0.0232	0.0364		95.1	93.6	19.1		
	AR1	Independent	-19.8	2.0	-1.6	-53.9	0.0032	0.0149	0.0226	0.0156	74.8	95.3	94.2	55.6		
		Exchangeable	-12.5	0.4	-2.1	-71.7	0.0036	0.0145	0.0227	0.0246	83.4	95.4	93.6	37.0		
BPL	Independent	Independent		-0.8	-1.1	-36.0	0.0032	0.0151	0.0225	0.0176	92.4	95.3	93.1	90.1		
		Exchangeable	-8.5	2.7	-6.3	-128.8		0.0149	0.0289	0.0784		94.5	87.9	26.7		
	AR1	Independent	-1.0	0.0	-0.3	0.2	0.0018	0.0148	0.0226	0.0093	93.5	95.7	94.0	95.0		
		Exchangeable	-5.6	-0.6	-0.8	-34.6	0.0030	0.0149	0.0225	0.0159	92.2	95.3	93.5	91.2		
N= 500																
WGEE	Independent	Independent		0.8	-5.4	-51.8		0.0233	0.0376	0.0161		93.4	92.2	68.7		
		Exchangeable	-14.8	0.9	-5.1	-51.0	0.0033	0.0222	0.0346	0.0157	85.4	93.8	92.3	69.6		
	AR1	Independent	-4.0	0.2	-5.2	-53.5	0.0049	0.0224	0.0355	0.0168	86.6	93.4	92.6	67.3		
		Exchangeable	-19.8	1.4	-2.9	-90.0		0.0146	0.0232	0.0364		95.1	93.6	19.1		
GEE	Independent	Independent		2.0	-1.6	-53.9	0.0032	0.0149	0.0226	0.0156	74.8	95.3	94.2	55.6		
		Exchangeable	-12.5	0.4	-2.1	-71.7	0.0036	0.0145	0.0227	0.0246	83.4	95.4	93.6	37.0		
	AR1	Independent	-8.5	-0.8	-1.1	-36.0	0.0032	0.0151	0.0225	0.0176	92.4	95.3	93.1	90.1		
		Exchangeable	-1.0	2.7	-6.3	-128.8		0.0149	0.0289	0.0784		94.5	87.9	26.7		
BPL	Independent	Independent		0.0	-0.3	0.2	0.0018	0.0148	0.0226	0.0093	93.5	95.7	94.0	95.0		
		Exchangeable	-5.6	-0.6	-0.8	-34.6	0.0030	0.0149	0.0225	0.0159	92.2	95.3	93.5	91.2		

Table 5.23: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($\mathbf{T} = \mathbf{3}$), **exchangeable correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 1, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		0.0	-2.0	-69.2		0.0566	0.0915	0.0380		94.9	94.6	86.0		
		Exchangeable	-9.9	0.0	-2.1	-67.0	0.0081	0.0533	0.0836	0.0361	91.9	95.0	94.4	86.1		
		AR1	1.4	-0.5	-2.1	-71.1	0.0087	0.0541	0.0830	0.0397	93.3	95.1	94.9	85.9		
	GEE	Independent		1.3	-0.6	-154.0		0.0450	0.0668	0.1118		94.1	94.6	34.7		
		Exchangeable	-16.6	2.0	0.8	-70.5	0.0094	0.0453	0.0650	0.0347	79.6	94.4	95.6	79.1		
		AR1	-5.5	0.9	0.6	-85.1	0.0074	0.0454	0.0648	0.0446	89.7	94.6	95.8	71.4		
	BPL	Exchangeable/AR1	-15.2	0.6	0.4	-62.8	0.0148	0.0458	0.0650	0.1311	90.9	94.5	94.7	85.3		
		Independent		0.3	-16.0	-350.0		0.0409	0.0862	0.5200		93.3	83.5	4.9		
	ML	Exchangeable	-8.5	2.0	1.0	-7.3	0.0093	0.0498	0.0685	0.0725	91.5	92.9	94.2	88.7		
		AR1	-10.6	3.1	2.3	-63.6	0.0130	0.0464	0.0659	0.1030	90.8	94.6	94.3	88.9		
		WGEE	Independent		-1.0	-3.6	-66.0		0.0221	0.0381	0.0251		95.1	94.4	67.9	
			Exchangeable	-7.7	-1.0	-3.6	-64.9	0.0036	0.0206	0.0346	0.0241	90.9	94.7	94.5	67.5	
	GEE		AR1	4.3	-1.7	-3.6	-69.1	0.0037	0.0210	0.0345	0.0269	92.8	95.0	94.2	66.5	
		Independent		0.8	-1.6	-152.7		0.0173	0.0277	0.1002		94.1	94.7	3.4		
		Exchangeable	-15.1	1.7	-0.0	-68.1	0.0057	0.0174	0.0266	0.0246	70.3	94.8	94.9	54.7		
	BPL	AR1	-3.3	0.4	-0.2	-82.5	0.0029	0.0171	0.0266	0.0337	89.6	94.5	94.4	43.6		
		Exchangeable/AR1	-7.5	0.4	0.0	-48.9	0.0051	0.0174	0.0272	0.0601	94.3	94.0	93.9	88.6		
		Independent		-0.6	-19.4	-363.3		0.0154	0.0601	0.5373		93.3	66.1	0.3		
ML		Exchangeable	-2.3	2.2	0.0	1.9	0.0021	0.0163	0.0239	0.0177	95.1	94.6	95.2	92.7		
	AR1	-1.3	2.8	1.5	-39.1	0.0034	0.0172	0.0253	0.0358	92.2	94.1	95.1	94.1			
N= 500																

Table 5.24: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for 4 responses ($T = 4$), **exchangeable correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 1, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		1.7	-3.3	-51.6		0.0718	0.1313	0.0248		94.7	93.1	83.4		
		Exchangeable	-9.5	1.7	-3.4	-47.1	0.0083	0.0630	0.1081	0.0220	88.3	94.8	93.6	84.6		
	GEE	AR1	11.6	0.5	-3.4	-53.8	0.0127	0.0638	0.1144	0.0251	83.0	94.4	93.9	83.6		
		Independent		2.4	-0.9	-136.3		0.0427	0.0669	0.0860		94.3	95.5	24.5		
	Exchangeable	AR1	-16.2	3.4	1.3	-54.0	0.0092	0.0431	0.0653	0.0214	74.1	94.1	95.6	74.4		
		Independent	1.9	1.0	1.0	-77.6	0.0081	0.0426	0.0666	0.0353	84.5	94.2	95.5	60.0		
	BPL	Exchangeable/AR1	-11.7	0.7	1.1	-66.6	0.0108	0.0430	0.0657	0.0643	90.7	93.3	94.7	86.9		
	ML	Independent		5.9	-17.4	-264.2		0.0419	0.0925	0.2989		91.4	79.0	5.5		
Exchangeable		0.2	3.3	1.8	10.9	0.0052	0.0407	0.0554	0.0228	90.6	93.9	94.1	89.9			
	AR1	Independent	-6.1	2.7	3.2	-55.9	0.0076	0.0413	0.0594	0.0441	89.5	94.9	94.2	89.9		
		Exchangeable														
	WGEE	Independent		0.8	-4.0	-51.4		0.0260	0.0492	0.0153		96.1	94.3	68.0		
		Exchangeable	-7.4	1.0	-3.9	-47.0	0.0036	0.0230	0.0402	0.0131	88.6	95.8	94.5	71.0		
	AR1	Independent	14.7	-0.2	-3.8	-54.3	0.0076	0.0235	0.0427	0.0163	74.7	95.4	94.1	64.6		
		Exchangeable		1.4	-1.6	-140.1		0.0173	0.0266	0.0828		95.3	95.6	0.7		
	GEE	AR1	-15.5	2.4	0.4	-56.3	0.0058	0.0177	0.0257	0.0161	60.4	94.8	96.1	46.7		
		Independent	3.0	-0.1	0.2	-81.2	0.0034	0.0173	0.0261	0.0305	83.0	95.2	95.6	23.5		
	BPL	Exchangeable/AR1	-8.4	-0.2	0.5	-72.9	0.0045	0.0172	0.0262	0.0398	90.4	94.8	93.9	79.0		
		Independent		5.5	-19.8	-278.4		0.0181	0.0624	0.3160		89.3	64.5	0.0		
	ML	Exchangeable	0.6	0.9	0.7	4.0	0.0015	0.0150	0.0222	0.0060	92.2	94.4	94.2	93.3		
		AR1	-5.1	2.0	3.3	-60.2	0.0023	0.0167	0.0262	0.0262	92.5	95.0	93.0	80.8		
N= 500																

Table 5.25: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **exchangeable correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C = (-1, 1, 1)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		-1.8	-8.0	-164.2		0.0491	0.0831	0.1351		96.3	94.2	50.3		
		Exchangeable	-35.0	-1.9	-8.1	-164.6	0.0127	0.0479	0.0798	0.1355	82.2	96.3	94.7	48.3		
	GEE	AR1	-37.2	-2.3	-8.1	-166.1	0.0183	0.0483	0.0806	0.1375	83.3	95.9	94.5	49.0		
		Independent		1.0	-2.8	-197.2		0.0385	0.0568	0.1797		95.6	95.3	25.2		
	Exchangeable	AR1	-38.1	1.4	-2.0	-163.2	0.0109	0.0389	0.0566	0.1302	74.2	95.2	95.1	38.0		
		Independent	-38.7	0.6	-2.3	-175.8	0.0140	0.0387	0.0566	0.1472	77.2	95.3	95.1	33.5		
	BPL	Exchangeable/AR1	-14.8	-0.8	-0.2	-41.9	0.0113	0.0404	0.0581	0.1105	92.0	94.9	95.2	90.6		
	ML	Independent		0.6	-3.1	-187.9		0.0381	0.0617	0.2295		95.1	92.8	66.2		
Exchangeable		-8.1	-0.8	-0.3	-16.9	0.0079	0.0404	0.0576	0.0950	93.3	95.1	95.2	90.5			
	AR1	-13.2	-0.6	-0.1	-42.6	0.0114	0.0397	0.0572	0.1054	91.4	95.0	95.4	91.1			
N= 500																
	WGEE	Independent		-1.4	-8.2	-158.4		0.0191	0.0377	0.1110		96.7	92.7	12.5		
		Exchangeable	-30.8	-1.4	-8.1	-158.6	0.0070	0.0185	0.0362	0.1110	77.6	96.9	92.5	11.8		
	GEE	AR1	-31.0	-2.0	-8.1	-161.1	0.0090	0.0189	0.0365	0.1142	77.5	96.6	92.5	11.2		
		Independent		1.8	-2.7	-192.1		0.0148	0.0233	0.1566		96.1	95.2	1.5		
	Exchangeable	AR1	-35.1	2.3	-1.8	-156.9	0.0071	0.0150	0.0227	0.1075	59.1	95.7	95.5	6.6		
		Independent	-34.1	1.3	-2.1	-170.1	0.0081	0.0147	0.0229	0.1248	66.5	95.8	95.5	3.9		
	BPL	Exchangeable/AR1	-5.3	-0.1	0.1	-24.1	0.0044	0.0149	0.0230	0.0414	92.0	95.8	95.5	92.6		
	ML	Independent		1.6	-2.6	-180.4		0.0147	0.0253	0.1677		95.2	93.0	43.3		
Exchangeable		-1.4	0.1	0.1	0.0	0.0028	0.0148	0.0226	0.0300	93.8	96.1	95.7	93.7			
	AR1	-3.5	0.2	0.2	-25.5	0.0042	0.0148	0.0229	0.0387	92.6	96.1	95.3	93.2			

Table 5.26: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **exchangeable correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C = (-1, 1, 1)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE			Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		3.4	-9.4	-117.7		0.0681	0.1155		0.0745		94.6	93.0	57.1	
		Exchangeable	-32.9	3.3	-9.3	-117.2	0.0122	0.0646	0.1045	0.0737	75.1	94.7	93.3	56.7		
	GEE	AR1	-31.7	2.6	-9.2	-120.4	0.0194	0.0654	0.1083	0.0766	78.3	94.9	93.1	55.4		
		Independent		4.5	-2.6	-158.1		0.0407	0.0555	0.1135		94.5	94.8	20.5		
		Exchangeable	-36.4	5.1	-1.3	-124.1	0.0097	0.0412	0.0553	0.0751	70.6	94.7	95.3	37.1		
		AR1	-34.5	3.5	-1.8	-142.5	0.0123	0.0401	0.0552	0.0946	74.8	94.5	94.9	27.1		
	BPL	Exchangeable/AR1	-12.8	0.7	1.0	-43.5	0.0089	0.0410	0.0559	0.0496	92.5	93.9	94.4	90.8		
		Independent		3.6	-1.9	-137.6		0.0401	0.0591	0.1139		93.6	92.5	64.0		
	ML	Exchangeable	-2.5	1.2	1.2	-2.2	0.0063	0.0408	0.0563	0.0400	92.5	94.4	95.0	92.2		
		AR1	-9.7	0.9	1.0	-43.4	0.0091	0.0406	0.0559	0.0470	91.4	93.8	94.2	91.0		
		WGEE	Independent		2.8	-10.1	-114.8		0.0244	0.0515		0.0596		94.3	90.1	20.0
			Exchangeable	-31.0	2.8	-9.9	-113.7	0.0071	0.0231	0.0474	0.0584	69.6	94.8	90.1	19.4	
	GEE	AR1	-30.4	1.8	-10.0	-117.6	0.0102	0.0233	0.0493	0.0620	74.6	95.2	90.1	17.2		
		Independent		4.0	-3.0	-157.4		0.0166	0.0244	0.1043		93.9	94.4	0.7		
	ML	Exchangeable	-36.8	4.6	-1.8	-123.2	0.0071	0.0170	0.0236	0.0659	47.8	93.7	94.8	5.1		
		AR1	-35.5	2.9	-2.3	-142.7	0.0080	0.0159	0.0238	0.0866	59.1	94.0	94.5	1.8		
		BPL	Exchangeable/AR1	-9.9	-0.2	0.4	-39.8	0.0037	0.0159	0.0239	0.0210	92.6	94.1	93.5	90.8	
		Independent		3.0	-1.9	-131.6		0.0161	0.0255	0.0836		93.3	91.3	29.8		
	ML	Exchangeable	-1.4	0.5	0.8	2.9	0.0021	0.0156	0.0231	0.0125	95.5	95.0	94.9	94.8		
		AR1	-7.3	0.1	0.5	-39.4	0.0034	0.0157	0.0236	0.0196	93.2	94.2	94.1	91.1		
N= 500																

Table 5.27: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **exchangeable correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C = (-1, 1, 1)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE			Coverage Probability							
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T			
N= 200																	
	WGEE	Independent		0.0	-5.8	-151.4		0.0645	0.1070		0.1190		95.5	94.6	55.1		
		Exchangeable	-16.9	-0.1	-6.0	-148.6	0.0154	0.0591	0.0945	84.4	95.4	95.0	53.5				
		AR1	-7.0	-1.3	-5.9	-157.3	0.0152	0.0595	0.0952	88.7	95.5	95.1	51.9				
	GEE	Independent		2.8	-1.3	-253.7		0.0426	0.0656		0.2797		95.2	95.8	6.8		
		Exchangeable	-25.0	3.6	0.2	-155.2	0.0167	0.0432	0.0653	64.0	95.6	95.5	32.3				
		AR1	-15.8	1.9	0.0	-176.8	0.0133	0.0424	0.0651	77.8	95.6	95.5	24.6				
	BPL	Exchangeable/AR1	-11.0	1.4	1.2	-64.6	0.0130	0.0425	0.0659	93.1	94.9	95.0	88.2				
		Independent		2.1	-12.2	-421.6		0.0406	0.0830		0.7824		93.7	86.9	13.7		
	ML	Exchangeable	-8.5	2.3	1.0	-5.4	0.0094	0.0489	0.0717	92.2	92.0	93.6	88.0				
		AR1	-7.6	4.3	2.8	-70.7	0.0121	0.0432	0.0651	92.0	94.7	94.9	91.2				
		WGEE	Independent		-2.8	-8.7	-151.4		0.0246	0.0528		0.1023		95.5	93.1	14.6	
			Exchangeable	-17.2	-2.7	-8.8	-148.0	0.0086	0.0227	0.0478	80.3	95.6	93.0	13.7			
AR1			-7.0	-4.2	-8.9	-158.1	0.0060	0.0238	0.0486	90.1	94.6	92.8	11.4				
GEE		Independent		1.4	-2.6	-252.5		0.0163	0.0289		0.2639		95.8	95.5	0.0		
		Exchangeable	-26.1	2.3	-1.1	-155.2	0.0133	0.0165	0.0279	36.9	95.7	94.6	3.7				
		AR1	-16.9	0.5	-1.3	-178.0	0.0081	0.0160	0.0281	64.6	96.2	94.6	1.3				
	BPL	Exchangeable/AR1	-7.2	0.3	0.3	-57.6	0.0057	0.0161	0.0278	92.5	96.5	94.1	89.3				
		Independent		0.4	-15.6	-439.9		0.0155	0.0542		0.8015		94.8	73.8	1.2		
	ML	Exchangeable	-1.8	1.2	0.1	3.5	0.0024	0.0150	0.0259	94.2	97.1	95.1	94.8				
		AR1	-1.8	2.5	1.6	-49.7	0.0041	0.0159	0.0268	92.3	96.6	94.5	92.8				
	N= 500																

Table 5.28: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **exchangeable correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 1, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE			Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		3.6	-10.9	-120.4		0.1015	0.1880		91.3	91.9	57.2			
		Exchangeable	-20.1	4.0	-10.7	-110.3	0.0173	0.0881	0.1477		75.7	92.6	92.4	56.9		
	GEE	AR1	-2.8	1.7	-10.6	-124.4	0.0185	0.0873	0.1554	0.0808	83.7	92.6	92.3	51.8		
		Independent		4.6	-2.4	-222.0		0.0455	0.0645	0.2111		95.1	95.3	2.1		
	Exchangeable	AR1	-28.8	5.8	-0.1	-128.9	0.0192	0.0458	0.0628	0.0790	50.0	94.4	95.9	26.0		
		Exchangeable	-15.3	2.5	-0.5	-164.0	0.0131	0.0436	0.0629	0.1216	75.0	95.4	95.6	13.7		
	BPL	Exchangeable/AR1	-12.0	1.5	1.5	-84.0	0.0111	0.0441	0.0647	0.0886	92.6	94.8	94.0	86.2		
	ML	Independent		6.8	-11.5	-313.0		0.0462	0.0838	0.4356		91.5	86.4	11.0		
Exchangeable		0.7	2.7	1.2	8.0	0.0057	0.0400	0.0578	0.0252	91.4	95.3	95.1	92.1			
	AR1	Exchangeable	-5.8	2.8	2.7	-69.8	0.0094	0.0408	0.0613	0.0590	89.7	95.7	95.0	88.4		
		AR1														
	WGEE	Independent		1.9	-10.5	-115.6		0.0357	0.0825	0.0613		94.9	91.2	23.8		
		Exchangeable	-17.3	2.1	-10.6	-107.5	0.0089	0.0304	0.0663	0.0533	77.1	95.3	90.6	22.3		
	AR1	Exchangeable	2.2	-0.3	-10.4	-121.9	0.0069	0.0306	0.0700	0.0667	88.3	95.1	90.8	15.0		
		Independent		3.3	-2.8	-221.9		0.0172	0.0282	0.2021		94.6	95.5	0.0		
	Exchangeable	AR1	-28.6	4.2	-1.0	-128.1	0.0153	0.0176	0.0266	0.0703	20.2	94.5	95.0	2.1		
		Exchangeable	-14.7	0.8	-1.3	-164.4	0.0070	0.0162	0.0270	0.1133	64.6	95.6	94.9	0.1		
	BPL	Exchangeable/AR1	-8.5	0.1	0.8	-86.2	0.0046	0.0165	0.0271	0.0520	92.0	95.9	94.0	78.6		
		Independent		5.9	-13.6	-328.3		0.0186	0.0495	0.4479		91.1	78.0	0.2		
	ML	Exchangeable	0.6	0.8	0.4	2.2	0.0016	0.0141	0.0227	0.0077	92.9	94.7	95.5	93.7		
		AR1	-5.5	1.9	3.0	-74.2	0.0030	0.0165	0.0269	0.0353	91.9	94.5	93.9	79.1		
N= 500																

5.2.2 Results for AR(1) Correlation

In this section, we present results from the simulations, where data were generated using the $AR(1)$ correlation structure. The simulation settings were kept the same as in the previous section, and the tables in this section are also presented in the same order as in the previous section.

Drop-out mechanism: MCAR

With an $AR(1)$ correlation structure, under MCAR, the GEE and wGEE estimators behaved similarly to those obtained under the exchangeable correlation. The same is also true for the bivariate maximum pseudo-likelihood estimators.

For the ML, estimation with the independence ‘working’ correlation showed some bias for regression parameters, especially for β_T . Tables 5.31 and 5.32 show that an increase in α results in an increase in the bias for β_0 and β_G . More importantly, the ML with independence assumption was highly inefficient for β_T with the largest MSE and has noticeably low coverage probabilities, especially for β_T . From Tables 5.31 and 5.32, we noticed an increase in α led to a higher MSE and lower coverage probabilities.

The ML with exchangeable ‘working’ correlation showed the same behaviour as the ML with independent ‘working’ correlation except that the coverage probabilities were not as dramatically low as we observed under independence. In addition, the increase in α had a considerable impact on the percentage relative bias of $\hat{\beta}_T$, resulting in a larger relative bias of close to 20%. The increase in sample size also resulted in an increase in the bias.

The ML with $AR(1)$ ‘working’ correlation showed an empirical bias close to zero and has the lowest MSE for β_0 and β_G but is much less efficient than GEE and wGEE for β_T . In addition, Tables 5.29 and 5.31 show that the coverage probability for β_T is nearly 0.90 when $N = 200$ and required a larger sample size to obtain better coverages.

Table 5.29: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **AR(1) correlation** ($\alpha = 0.2$) and **MCAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		-0.2	-0.8	-0.1		0.0428	0.0586	0.0185		95.0	95.1	94.8		
		Exchangeable	-22.0	-0.2	-0.7	-0.5	0.0063	0.0430	0.0614	0.0182	87.2	94.7	95.0	94.4		
	GEE	AR1	-2.1	-0.2	-0.7	-0.3	0.0054	0.0428	0.0598	0.0191	93.9	94.8	94.9	94.6		
		Independent		-0.7	-0.7	-1.0		0.0383	0.0501	0.0173		94.6	94.9	93.1		
	Exchangeable	-25.9	-0.7	-0.7	-1.6	0.0067	0.0378	0.0493	0.0173	80.8	94.4	95.0	92.7			
	BPL	AR1	-2.1	-0.6	-0.7	-1.5	0.0047	0.0376	0.0489	0.0172	94.2	94.5	94.7	92.8		
		Exchangeable/AR1	-8.7	-0.9	-1.2	-0.7	0.0042	0.0381	0.0502	0.0475	93.7	94.4	93.8	90.4		
	ML	Independent		-3.1	-4.6	0.1		0.0432	0.0558	0.0654		91.0	88.9	70.7		
Exchangeable		-31.1	-1.4	-1.9	-5.2	0.0075	0.0383	0.0507	0.0512	78.9	93.3	94.0	85.3			
	WGEE	AR1	-8.3	-0.9	-1.2	-0.3	0.0042	0.0381	0.0501	0.0474	93.1	94.2	94.0	89.8		
		Independent		-0.1	0.2	-1.5		0.0173	0.0233	0.0068		95.1	95.5	94.7		
	GEE	Exchangeable	-19.6	-0.1	0.2	-1.4	0.0031	0.0175	0.0245	0.0067	84.4	95.0	95.6	94.5		
		AR1	0.4	-0.1	0.2	-1.4	0.0020	0.0175	0.0239	0.0070	94.9	95.0	95.7	94.6		
	BPL	Independent		0.1	0.3	-1.1		0.0147	0.0201	0.0063		95.1	95.0	93.6		
		Exchangeable	-23.6	0.1	0.3	-0.9	0.0036	0.0147	0.0199	0.0062	74.9	94.9	95.1	93.8		
	ML	AR1	0.1	0.1	0.3	-1.0	0.0017	0.0147	0.0199	0.0062	95.1	95.1	95.0	93.6		
		Exchangeable/AR1	-2.5	0.0	0.1	-2.8	0.0016	0.0150	0.0201	0.0181	95.6	95.4	94.9	93.4		
	WGEE	Independent		-1.1	-2.7	-12.7		0.0172	0.0234	0.0394		91.7	88.6	69.6		
		Exchangeable	-25.4	-0.1	-0.3	-7.3	0.0039	0.0150	0.0202	0.0204	73.8	94.9	95.2	90.6		
	GEE	AR1	-2.4	0.1	0.1	-2.9	0.0016	0.0149	0.0200	0.0182	95.4	95.3	95.2	93.5		
		Independent		-0.1	0.2	-1.5		0.0173	0.0233	0.0068		95.1	95.5	94.7		
	BPL	Exchangeable	-19.6	-0.1	0.2	-1.4	0.0031	0.0175	0.0245	0.0067	84.4	95.0	95.6	94.5		
		AR1	0.4	-0.1	0.2	-1.4	0.0020	0.0175	0.0239	0.0070	94.9	95.0	95.7	94.6		
	ML	Independent		0.1	0.3	-1.1		0.0147	0.0201	0.0063		95.1	95.0	93.6		
		Exchangeable	-23.6	0.1	0.3	-0.9	0.0036	0.0147	0.0199	0.0062	74.9	94.9	95.1	93.8		

Table 5.30: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **AR(1) correlation** ($\alpha = 0.2$) and **MCAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		1.0	0.7	1.1		0.0408	0.0501	0.0090		94.5	94.7	95.5		95.5
		Exchangeable	-34.9	1.0	0.7	1.0	0.0076	0.0407	0.0515	0.0089	70.2	94.3	95.2	95.7		95.7
		AR1	-1.6	1.0	0.7	0.9	0.0041	0.0407	0.0500	0.0091	94.3	94.4	94.8	95.3		95.3
	GEE	Independent		1.1	0.8	1.3		0.0350	0.0441	0.0082		95.0	94.0	94.1		94.1
		Exchangeable	-38.9	1.1	0.9	1.4	0.0084	0.0348	0.0437	0.0082	59.7	95.2	94.1	94.5		94.5
		AR1	-1.6	1.1	0.9	1.2	0.0034	0.0350	0.0436	0.0081	93.8	94.8	94.0	94.8		94.8
	BPL	Exchangeable/AR1	-5.9	0.9	0.5	0.5	0.0032	0.0362	0.0434	0.0181	94.5	94.3	94.4	93.0		93.0
		Independent		-0.4	-2.6	-5.5		0.0422	0.0474	0.0245		91.0	90.5	76.1		76.1
	ML	Exchangeable	-40.9	0.4	-0.5	-3.5	0.0089	0.0376	0.0440	0.0206	56.0	93.6	94.5	87.9		87.9
		AR1	-5.3	0.9	0.5	0.9	0.0031	0.0359	0.0434	0.0178	93.9	94.6	94.4	93.0		93.0
	N= 500															
	WGEE	Independent		0.2	0.1	-0.8		0.0153	0.0199	0.0032		94.3	95.3	96.0		96.0
		Exchangeable	-33.9	0.2	0.1	-0.9	0.0056	0.0154	0.0206	0.0032	43.8	94.2	95.2	96.1		96.1
		AR1	0.2	0.2	0.1	-0.7	0.0015	0.0154	0.0200	0.0033	95.1	93.8	95.3	95.8		95.8
	GEE	Independent		-0.2	-0.2	-0.8		0.0134	0.0165	0.0028		94.7	95.5	96.1		96.1
		Exchangeable	-37.8	-0.2	-0.2	-0.9	0.0066	0.0135	0.0166	0.0028	30.8	94.8	94.9	96.1		96.1
		AR1	0.2	-0.1	-0.2	-0.9	0.0013	0.0135	0.0165	0.0028	95.2	94.6	95.2	96.0		96.0
	BPL	Exchangeable/AR1	-1.3	-0.1	-0.3	-2.5	0.0012	0.0140	0.0165	0.0068	96.0	94.7	95.1	95.2		95.2
		Independent		-0.7	-2.6	-9.1		0.0174	0.0200	0.0131		90.3	89.1	79.5		79.5
	ML	Exchangeable	-38.4	-0.2	-0.9	-7.7	0.0068	0.0145	0.0175	0.0086	27.5	94.0	94.8	90.7		90.7
		AR1	-1.3	-0.1	-0.4	-2.6	0.0011	0.0140	0.0166	0.0067	94.9	95.0	95.4	95.2		95.2

Table 5.31: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **AR(1) correlation** ($\alpha = 0.4$) and **MCAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability			
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T
N= 200														
	WGEE	Independent		0.9	1.2	0.4		0.0502	0.0694	0.0169		94.5	95.3	94.6
		Exchangeable	-15.4	1.2	1.7	0.5	0.0079	0.0508	0.0754	0.0163	84.3	93.9	95.5	93.5
	GEE	AR1	-0.5	1.2	1.6	0.4	0.0044	0.0501	0.0727	0.0179	95.3	94.1	95.4	93.6
		Independent		0.4	0.0	0.4		0.0450	0.0615	0.0156		94.2	95.1	93.9
	Exchangeable		-18.5	0.6	0.4	0.5	0.0093	0.0443	0.0601	0.0153	77.7	94.0	95.5	93.8
		AR1	-0.6	0.6	0.4	0.4	0.0041	0.0435	0.0591	0.0152	95.1	94.3	95.7	93.8
	BPL	Exchangeable/AR1	-8.1	0.3	-0.2	-2.3	0.0050	0.0435	0.0596	0.0579	93.8	94.1	95.5	87.8
	ML	Independent		-6.9	-10.4	6.9		0.0549	0.0645	0.1107		85.3	84.9	27.9
Exchangeable		-25.5	-0.2	-1.7	-24.1	0.0146	0.0432	0.0601	0.0541	63.2	93.4	94.0	77.3	
	AR1		-6.0	0.7	0.1	-2.0	0.0046	0.0425	0.0584	0.0508	92.7	94.6	94.5	88.6
	WGEE	Independent		0.4	1.2	-0.8		0.0198	0.0272	0.0060		93.8	95.0	94.8
		Exchangeable	-15.3	0.4	1.2	-0.7	0.0053	0.0200	0.0291	0.0057	66.6	94.5	95.1	94.2
	AR1	-0.2	0.3	1.2	-1.0	0.0016	0.0201	0.0285	0.0063	96.0	94.3	95.0	95.3	
		Independent		0.5	1.1	0.2		0.0174	0.0235	0.0057		93.6	94.9	94.2
	Exchangeable	-18.3	0.5	1.1	0.1	0.0068	0.0170	0.0229	0.0055	52.1	94.2	95.3	94.4	
		AR1	-0.3	0.4	1.1	-0.1	0.0015	0.0171	0.0232	0.0054	95.2	94.1	94.7	94.2
	BPL	Exchangeable/AR1	-3.7	0.3	0.8	-2.6	0.0017	0.0174	0.0234	0.0252	96.4	93.4	94.9	92.5
		Independent		-8.1	-10.9	10.3		0.0321	0.0321	0.1001		73.8	78.9	2.5
	ML	Exchangeable	-22.8	-0.1	-0.4	-26.0	0.0101	0.0170	0.0238	0.0301	39.3	93.0	93.9	81.7
		AR1	-2.4	0.6	1.2	-1.2	0.0014	0.0168	0.0230	0.0202	95.5	93.2	94.9	92.7
N= 500														

Table 5.32: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **AR(1) correlation** ($\alpha = 0.4$) and **MCAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability			
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T
N= 200														
WGEE	Independent	Independent		0.9	1.0	2.1		0.0443	0.0658	0.0090		95.3	94.5	94.3
		Exchangeable	-28.1	0.8	0.8	2.0	0.0156	0.0432	0.0670	0.0088	47.3	95.6	94.1	94.3
	AR1	Independent	-1.7	0.9	0.9	1.9	0.0035	0.0422	0.0641	0.0091	95.2	95.8	93.7	94.4
		Exchangeable	-31.5	1.2	1.1	2.8		0.0380	0.0550	0.0085		95.3	95.0	94.4
GEE	Independent	Independent		1.1	1.0	2.8	0.0185	0.0371	0.0533	0.0086	33.1	95.3	94.7	94.0
		Exchangeable	-1.6	1.1	1.0	2.7	0.0030	0.0361	0.0521	0.0081	94.6	95.2	94.7	94.2
	AR1	Exchangeable/AR1	-6.0	0.7	0.4	4.5	0.0033	0.0362	0.0526	0.0266	95.6	95.8	94.0	89.7
		Independent	-8.4	-11.5	2.4		0.0609	0.0628	0.0483		79.0	81.5	31.8	
BPL	Independent	Independent		-0.8	-2.6	-19.4	0.0259	0.0374	0.0532	0.0340	18.1	93.3	92.2	67.5
		Exchangeable	-37.9	-0.8	-2.6	-19.4		0.0374	0.0532	0.0340		93.3	92.2	67.5
	AR1	Independent	-1.2	0.3	-0.3	4.4	0.0026	0.0337	0.0472	0.0191	94.4	95.8	94.7	91.3
		Exchangeable												
N= 500														
WGEE	Independent	Independent		0.5	0.3	0.5		0.0177	0.0240	0.0037		95.9	95.0	94.5
		Exchangeable	-26.5	0.6	0.5	0.6	0.0124	0.0175	0.0247	0.0036	16.1	95.6	95.3	95.1
	AR1	Independent	0.2	0.5	0.4	0.1	0.0014	0.0171	0.0234	0.0038	94.6	95.3	95.7	94.6
		Exchangeable	-29.9	0.7	0.2	1.1		0.0150	0.0218	0.0034		95.2	94.8	94.6
GEE	Independent	Independent		0.7	0.4	1.1	0.0154	0.0149	0.0210	0.0034	6.2	95.0	95.0	94.2
		Exchangeable	-29.9	0.7	0.4	1.1		0.0149	0.0210	0.0034		95.0	95.0	94.2
	AR1	Independent	0.1	0.7	0.4	0.7	0.0013	0.0145	0.0205	0.0033	93.8	95.6	95.1	94.1
		Exchangeable/AR1	-2.4	0.6	0.1	-0.0	0.0012	0.0146	0.0205	0.0111	95.3	95.5	94.7	92.5
BPL	Independent	Independent		-9.6	-12.8	1.3	0.0484	0.0355	0.0415		58.1	70.3	5.3	
		Exchangeable	-34.0	-0.2	-2.5	-23.9	0.0200	0.0151	0.0225	0.0186	1.5	93.9	92.1	73.5
	AR1	Independent	-0.1	0.4	-0.2	2.2	0.0008	0.0132	0.0183	0.0074	93.1	96.1	94.9	93.1
		Exchangeable												

Drop-out mechanism: MAR

With the $AR(1)$ correlation, the same patterns were observed in the GEE and wGEE estimators as we have seen among estimators obtained under the exchangeable correlation.

The bivariate maximum pseudo-likelihood method was highly efficient, comparable to the ML with $AR(1)$ assumption, but showed some moderate bias for the time coefficient, which was affected by an increase in α or γ_P . For example, Tables 5.33 and 5.35 show that the increase in α from 0.2 to 0.4, and for $T = 3$ and $N = 500$, the relative bias for $\hat{\beta}_T$ increased from 0.9% to -10.3%. Tables 5.35 and 5.39 show that an increase in γ_P from 1 to 2, and for $T = 3$ and $N = 500$, the relative bias increases from -3.5% to -10.3%. The coverage probabilities for β_0 and β_G were close to 0.95 for both $N = 200$ and $N = 500$, but for β_T , the coverage probability was close to 0.90 for $N = 200$ and ranged from 0.91 to 0.94 for $N = 500$.

The ML with independent assumption showed considerably high relative bias and large MSE for the time coefficient, and also low coverage probabilities for the same coefficient. Table 5.35 showed the worst results with -17.3% for the percentage relative bias of $\hat{\beta}_G$ and -317.0% for the percentage relative bias of $\hat{\beta}_T$. The MSE for β_T was nearly 10 times the MSE found from the ML with $AR(1)$ assumption and the coverage probability was 0.0.

The ML with exchangeable correlation assumption showed a pattern similar to the ML with independence assumption, but in a moderate degree. Table 5.39 shows the worst performance with the percentage relative bias of -58.7%

The ML with $AR(1)$ correlation assumption showed small bias when $N = 200$, but decreased to 0% with $N = 500$. Similarly, the coverage probability for β_T was below 0.9 for $N = 200$, but increases to 0.95 with $N = 500$. The MSE for β_0 and β_G were the smallest with bivariate maximum pseudo-likelihood, but the MSE for β_T was 3 times larger than MSE found from the wGEE.

Table 5.33: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **AR(1) correlation** ($\alpha = 0.2$) and **MAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C = (-1, 2, 0)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE			Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		1.6	0.3	-0.2		0.0678	0.1055	0.0401		94.7	94.3	93.7		
		Exchangeable	-27.8	1.5	0.3	-0.6	0.0149	0.0653	0.0989	0.0401	82.0	95.3	94.3	93.3		
	AR1	Independent	-7.9	1.6	0.5	0.4	0.0140	0.0653	0.0993	0.0401	90.5	95.2	94.3	93.2		
		Exchangeable	-33.1	2.0	-1.1	-52.2		0.0379	0.0537	0.0334		95.8	94.9	89.4		
	GEE	Independent		1.0	-1.1	-52.2		0.0379	0.0537	0.0334		95.8	94.9	89.4		
		Exchangeable	-33.1	2.0	0.7	2.8	0.0113	0.0388	0.0561	0.0258	74.5	95.7	95.1	92.3		
	BPL	Independent	-19.5	1.0	0.7	-6.8	0.0091	0.0386	0.0559	0.0248	85.3	96.0	95.2	93.5		
		Exchangeable/AR1	-7.8	1.1	1.0	10.7	0.0167	0.0391	0.0589	0.1088	89.0	95.3	94.6	89.2		
	ML	Independent		0.8	-6.4	-146.6		0.0364	0.0593	0.1315	95.3	92.2	63.1			
		Exchangeable	-33.3	1.5	-0.7	-13.8	0.0226	0.0396	0.0608	0.1173	78.9	95.2	93.2	82.2		
	AR1	Independent	-7.5	1.0	0.8	12.2	0.0173	0.0391	0.0589	0.1106	87.2	95.5	94.5	87.0		
		Exchangeable														
N= 500																
	WGEE	Independent		0.2	0.4	-0.8		0.0279	0.0437	0.0133		95.8	94.4	95.5		
		Exchangeable	-23.8	0.2	0.5	-0.5	0.0067	0.0268	0.0405	0.0137	84.9	95.9	95.2	95.8		
	AR1	Independent	-3.6	0.3	0.5	-0.0	0.0050	0.0267	0.0406	0.0136	94.0	95.8	95.1	95.8		
		Exchangeable	-31.9	1.1	-1.0	-55.9		0.0156	0.0210	0.0209		94.8	95.4	76.5		
	GEE	Independent		0.1	-1.0	-55.9		0.0156	0.0210	0.0209		94.8	95.4	76.5		
		Exchangeable	-31.9	1.1	0.8	0.8	0.0066	0.0160	0.0215	0.0094	64.3	95.2	96.3	94.0		
	BPL	Independent	-18.6	0.0	0.7	-10.3	0.0041	0.0159	0.0214	0.0094	81.8	95.0	96.1	93.8		
		Exchangeable/AR1	-4.3	0.1	0.9	-0.6	0.0069	0.0159	0.0219	0.0393	92.5	94.8	96.1	93.8		
	ML	Independent		-0.1	-6.9	-156.3		0.0151	0.0261	0.1142	93.6	91.5	30.2			
		Exchangeable	-32.1	0.9	-0.3	-22.7	0.0112	0.0158	0.0230	0.0449	83.0	94.3	95.6	90.9		
	AR1	Independent	-4.3	-0.0	0.7	0.6	0.0068	0.0159	0.0218	0.0393	92.5	94.4	95.7	92.7		
		Exchangeable														

Table 5.34: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **AR(1) correlation** ($\alpha = 0.2$) and **MAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C = (-1, 2, 0)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability							
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T				
N= 200																		
WGEE	Independent Exchangeable AR1	Independent		4.2	2.9	1.1		0.1071	0.1642	0.0330		0.1071	0.1642	0.0330		93.7	92.4	92.6
		Exchangeable	-43.1	3.8	2.5	0.8	0.0192	0.1018	0.1455	0.0349		0.1018	0.1455	0.0349	68.0	93.6	92.5	92.2
	GEE	Independent	-8.7	4.0	2.8	1.8	0.0148	0.0996	0.1480	0.0334		0.0996	0.1480	0.0334	88.1	93.6	92.0	92.3
		Exchangeable		1.9	-0.6	-40.8		0.0353	0.0472	0.0193		0.0353	0.0472	0.0193		96.1	96.3	88.0
BPL ML	Exchangeable/AR1	Independent	-33.9	2.9	1.5	11.9	0.0111	0.0366	0.0498	0.0160		0.0366	0.0498	0.0160	71.1	96.1	96.5	92.0
		Exchangeable	-14.8	1.0	1.3	-3.6	0.0080	0.0357	0.0494	0.0137		0.0357	0.0494	0.0137	86.1	96.2	96.5	93.4
	ML	Independent	-4.9	1.1	1.6	5.7	0.0134	0.0366	0.0519	0.0456		0.0366	0.0519	0.0456	89.4	95.8	95.5	90.9
		Exchangeable		3.3	-6.4	-106.5		0.0347	0.0527	0.0674		0.0347	0.0527	0.0674		95.3	92.8	62.3
N= 500																		
WGEE	Independent Exchangeable AR1	Independent	-37.9	2.8	-0.7	-9.5	0.0197	0.0366	0.0532	0.0544		0.0366	0.0532	0.0544	76.4	95.1	94.2	86.5
		Exchangeable	-2.9	0.7	0.9	6.6	0.0138	0.0366	0.0511	0.0444		0.0366	0.0511	0.0444	87.0	95.8	95.9	89.4
	GEE	Independent		1.8	0.7	0.7		0.0448	0.0662	0.0128		0.0448	0.0662	0.0128		93.6	92.5	93.6
		Exchangeable	-38.5	1.8	0.8	0.9	0.0107	0.0427	0.0587	0.0131		0.0427	0.0587	0.0131	65.9	93.3	92.8	92.7
BPL ML	Exchangeable/AR1	Independent	-4.6	1.9	0.8	1.3	0.0060	0.0418	0.0596	0.0127		0.0418	0.0596	0.0127	91.0	93.3	92.5	93.2
		Exchangeable		1.9	-0.9	-41.5		0.0148	0.0195	0.0123		0.0148	0.0195	0.0123		95.2	95.6	78.0
	ML	Independent	-34.0	3.1	1.5	11.4	0.0069	0.0158	0.0207	0.0069		0.0158	0.0207	0.0069	56.3	94.6	95.6	90.7
		Exchangeable	-15.1	1.1	1.1	-5.6	0.0032	0.0148	0.0202	0.0059		0.0148	0.0202	0.0059	85.5	95.8	95.8	92.8
N= 500																		
BPL ML	Exchangeable/AR1	Independent	-1.2	1.0	1.5	2.8	0.0045	0.0148	0.0208	0.0186		0.0148	0.0208	0.0186	93.9	95.6	95.2	92.3
		Exchangeable		3.5	-6.6	-108.3		0.0150	0.0244	0.0569		0.0150	0.0244	0.0569		93.6	90.3	31.6
	ML	Independent	-38.7	3.1	-0.3	-12.9	0.0106	0.0154	0.0212	0.0229		0.0154	0.0212	0.0229	73.1	94.6	94.7	88.7
		Exchangeable	-1.6	0.7	1.0	1.4	0.0043	0.0147	0.0203	0.0174		0.0147	0.0203	0.0174	93.9	96.0	95.5	92.5

Table 5.35: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **AR(1) correlation** ($\alpha = 0.4$) and **MAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C = (-1, 2, 0)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE			Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		2.4	1.7	-1.8		0.0864	0.1555		0.0337		93.7	92.3	93.2	
		Exchangeable	-17.0	2.4	1.8	0.6	0.0172	0.0782	0.1327		0.0364	84.0	94.4	92.6	92.8	
		AR1	-1.9	2.4	1.8	1.2	0.0130	0.0765	0.1314	0.0350		89.9	94.3	92.6	92.2	
		Independent		1.1	-1.8	-125.9		0.0409	0.0665	0.0853		95.6	94.5	60.1		
	GEE	Exchangeable	-27.1	2.8	1.7	-5.7	0.0193	0.0428	0.0681	0.0219		60.9	95.3	95.1	93.7	
		AR1	-15.3	1.2	1.5	-22.1	0.0122	0.0419	0.0675	0.0245		78.4	95.7	95.3	91.7	
		Exchangeable/AR1	-9.9	1.3	1.5	-10.9	0.0199	0.0423	0.0673	0.1328		88.5	94.7	95.6	87.7	
		Independent		-0.2	-15.9	-314.9		0.0371	0.0849	0.4262		94.5	84.5	5.4		
	BPL	Exchangeable	-27.8	2.2	-0.7	-39.2	0.0362	0.0441	0.0689	0.1316		78.3	93.4	93.8	82.6	
		AR1	-8.1	1.0	0.9	-11.8	0.0191	0.0410	0.0628	0.1048		87.0	94.7	95.8	87.5	
		Independent		0.9	0.5	-2.8		0.0327	0.0574	0.0131		95.4	94.3	95.2		
		Exchangeable	-17.5	0.8	0.4	-1.8	0.0097	0.0298	0.0491	0.0136		81.0	95.6	93.9	95.4	
	GEE	AR1	-2.4	0.9	0.5	-1.5	0.0051	0.0294	0.0491	0.0133		93.6	95.7	94.2	95.2	
		Independent		0.6	-2.6	-125.1		0.0160	0.0258	0.0708		95.2	94.9	22.7		
		Exchangeable	-29.2	2.0	0.4	-8.3	0.0164	0.0165	0.0253	0.0083		27.8	95.0	95.3	93.2	
		AR1	-17.3	0.4	0.3	-25.6	0.0079	0.0161	0.0252	0.0108		63.5	95.2	95.6	89.9	
BPL	Exchangeable/AR1	-5.4	0.6	0.5	-10.3	0.0083	0.0163	0.0252	0.0576		90.5	94.6	96.1	91.1		
	Independent		-0.7	-17.3	-317.0		0.0146	0.0524	0.4137		94.3	73.7	0.0			
	Exchangeable	-27.0	1.4	-1.4	-41.6	0.0217	0.0167	0.0256	0.0580		78.8	94.5	95.4	87.6		
		AR1	-4.3	0.2	-0.1	-3.8	0.0069	0.0160	0.0244	0.0450		90.9	95.0	95.7	91.6	
N= 500																

Table 5.36: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **AR(1) correlation** ($\alpha = 0.4$) and **MAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 2, 0)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		6.1	3.9	-4.2	0.1334	0.2298	0.0340		91.9	90.0	92.2			
		Exchangeable AR1	-32.5	5.7	3.2	0.9	0.0297	0.1175	0.1836	0.0355	62.6	92.0	91.7	92.2		
	GEE	Independent	-6.4	5.4	3.3	-0.3	0.0144	0.1114	0.1840	0.0327	84.7	92.0	91.3	91.8		
		Exchangeable AR1	-31.5	5.2	-2.2	-93.4	0.0400	0.0581	0.0482		94.2	95.0	62.6			
	BPL ML	Exchangeable/AR1	-13.6	2.3	1.4	-11.1	0.0117	0.0392	0.0599	0.0152	72.9	94.9	95.9	91.4		
		Independent	-6.1	2.3	1.6	2.8	0.0131	0.0396	0.0623	0.0583	91.9	94.6	95.3	90.9		
		Exchangeable AR1		6.3	-16.8	-219.1	0.0383	0.0801	0.2099		92.8	82.2	7.3			
			-32.6	3.6	-3.3	-20.8	0.0380	0.0418	0.0645	0.0629	71.0	94.0	93.2	80.0		
N= 500																
	WGEE	Independent		3.6	3.5	-1.4	0.0588	0.0980	0.0132		91.6	91.1	92.8			
		Exchangeable AR1	-29.3	3.3	3.0	0.3	0.0196	0.0524	0.0775	0.0140	53.5	92.5	92.0	91.9		
	GEE	Independent	-2.6	3.3	3.2	-0.5	0.0063	0.0500	0.0794	0.0128	90.3	92.1	91.6	92.4		
		Exchangeable AR1		3.5	-1.5	-97.1	0.0167	0.0241	0.0428		94.4	95.2	21.8			
	BPL ML	Exchangeable/AR1	-31.8	4.8	2.3	8.5	0.0194	0.0182	0.0255	0.0058	21.3	94.9	95.0	92.0		
		Independent	-13.9	1.7	2.1	-17.1	0.0068	0.0158	0.0252	0.0068	62.7	95.9	94.7	90.7		
		Exchangeable AR1	-2.7	1.8	2.6	-3.8	0.0048	0.0158	0.0258	0.0227	94.2	95.9	94.9	92.8		
			-29.7	2.7	-1.2	-15.4	0.0211	0.0171	0.0251	0.0212	56.5	93.7	93.5	89.8		
N= 500																
	WGEE	Independent		3.6	3.5	-1.4	0.0588	0.0980	0.0132		91.6	91.1	92.8			
		Exchangeable AR1	-29.3	3.3	3.0	0.3	0.0196	0.0524	0.0775	0.0140	53.5	92.5	92.0	91.9		
	GEE	Independent	-2.6	3.3	3.2	-0.5	0.0063	0.0500	0.0794	0.0128	90.3	92.1	91.6	92.4		
		Exchangeable AR1		3.5	-1.5	-97.1	0.0167	0.0241	0.0428		94.4	95.2	21.8			
	BPL ML	Exchangeable/AR1	-31.8	4.8	2.3	8.5	0.0194	0.0182	0.0255	0.0058	21.3	94.9	95.0	92.0		
		Independent	-13.9	1.7	2.1	-17.1	0.0068	0.0158	0.0252	0.0068	62.7	95.9	94.7	90.7		
		Exchangeable AR1	-2.7	1.8	2.6	-3.8	0.0048	0.0158	0.0258	0.0227	94.2	95.9	94.9	92.8		
			-29.7	2.7	-1.2	-15.4	0.0211	0.0171	0.0251	0.0212	56.5	93.7	93.5	89.8		
N= 500																
	WGEE	Independent		3.6	3.5	-1.4	0.0588	0.0980	0.0132		91.6	91.1	92.8			
		Exchangeable AR1	-29.3	3.3	3.0	0.3	0.0196	0.0524	0.0775	0.0140	53.5	92.5	92.0	91.9		
	GEE	Independent	-2.6	3.3	3.2	-0.5	0.0063	0.0500	0.0794	0.0128	90.3	92.1	91.6	92.4		
		Exchangeable AR1		3.5	-1.5	-97.1	0.0167	0.0241	0.0428		94.4	95.2	21.8			
	BPL ML	Exchangeable/AR1	-31.8	4.8	2.3	8.5	0.0194	0.0182	0.0255	0.0058	21.3	94.9	95.0	92.0		
		Independent	-13.9	1.7	2.1	-17.1	0.0068	0.0158	0.0252	0.0068	62.7	95.9	94.7	90.7		
		Exchangeable AR1	-2.7	1.8	2.6	-3.8	0.0048	0.0158	0.0258	0.0227	94.2	95.9	94.9	92.8		
			-29.7	2.7	-1.2	-15.4	0.0211	0.0171	0.0251	0.0212	56.5	93.7	93.5	89.8		
N= 500																
	WGEE	Independent		3.6	3.5	-1.4	0.0588	0.0980	0.0132		91.6	91.1	92.8			
		Exchangeable AR1	-29.3	3.3	3.0	0.3	0.0196	0.0524	0.0775	0.0140	53.5	92.5	92.0	91.9		
	GEE	Independent	-2.6	3.3	3.2	-0.5	0.0063	0.0500	0.0794	0.0128	90.3	92.1	91.6	92.4		
		Exchangeable AR1		3.5	-1.5	-97.1	0.0167	0.0241	0.0428		94.4	95.2	21.8			
	BPL ML	Exchangeable/AR1	-31.8	4.8	2.3	8.5	0.0194	0.0182	0.0255	0.0058	21.3	94.9	95.0	92.0		
		Independent	-13.9	1.7	2.1	-17.1	0.0068	0.0158	0.0252	0.0068	62.7	95.9	94.7	90.7		
		Exchangeable AR1	-2.7	1.8	2.6	-3.8	0.0048	0.0158	0.0258	0.0227	94.2	95.9	94.9	92.8		
			-29.7	2.7	-1.2	-15.4	0.0211	0.0171	0.0251	0.0212	56.5	93.7	93.5	89.8		
N= 500																
	WGEE	Independent		3.6	3.5	-1.4	0.0588	0.0980	0.0132		91.6	91.1	92.8			
		Exchangeable AR1	-29.3	3.3	3.0	0.3	0.0196	0.0524	0.0775	0.0140	53.5	92.5	92.0	91.9		
	GEE	Independent	-2.6	3.3	3.2	-0.5	0.0063	0.0500	0.0794	0.0128	90.3	92.1	91.6	92.4		
		Exchangeable AR1		3.5	-1.5	-97.1	0.0167	0.0241	0.0428		94.4	95.2	21.8			
	BPL ML	Exchangeable/AR1	-31.8	4.8	2.3	8.5	0.0194	0.0182	0.0255	0.0058	21.3	94.9	95.0	92.0		
		Independent	-13.9	1.7	2.1	-17.1	0.0068	0.0158	0.0252	0.0068	62.7	95.9	94.7	90.7		
		Exchangeable AR1	-2.7	1.8	2.6	-3.8	0.0048	0.0158	0.0258	0.0227	94.2	95.9	94.9	92.8		
			-29.7	2.7	-1.2	-15.4	0.0211	0.0171	0.0251	0.0212	56.5	93.7	93.5	89.8		
N= 500																
	WGEE	Independent		3.6	3.5	-1.4	0.0588	0.0980	0.0132		91.6	91.1	92.8			
		Exchangeable AR1	-29.3	3.3	3.0	0.3	0.0196	0.0524	0.0775	0.0140	53.5	92.5	92.0	91.9		
	GEE	Independent	-2.6	3.3	3.2	-0.5	0.0063	0.0500	0.0794	0.0128	90.3	92.1	91.6	92.4		
		Exchangeable AR1		3.5	-1.5	-97.1	0.0167	0.0241	0.0428		94.4	95.2	21.8			
	BPL ML	Exchangeable/AR1	-31.8	4.8	2.3	8.5	0.0194	0.0182	0.0255	0.0058	21.3	94.9	95.0	92.0		
		Independent	-13.9	1.7	2.1	-17.1	0.0068	0.0158	0.0252	0.0068	62.7	95.9	94.7	90.7		
		Exchangeable AR1	-2.7	1.8	2.6	-3.8	0.0048	0.0158	0.0258	0.0227	94.2	95.9	94.9	92.8		
			-29.7	2.7	-1.2	-15.4	0.0211	0.0171	0.0251	0.0212	56.5	93.7	93.5	89.8		
N= 500																
	WGEE	Independent		3.6	3.5	-1.4	0.0588	0.0980	0.0132		91.6	91.1	92.8			
		Exchangeable AR1	-29.3	3.3	3.0	0.3	0.0196	0.0524	0.0775	0.0140	53.5	92.5	92.0	91.9		
	GEE	Independent	-2.6	3.3	3.2	-0.5	0.0063	0.0500	0.0794	0.0128	90.3	92.1	91.6	92.4		
		Exchangeable AR1		3.5	-1.5	-97.1	0.0167	0.0241	0.0428		94.4	95.2	21.8			
	BPL ML	Exchangeable/AR1	-31.8	4.8	2.3	8.5	0.0194	0.0182	0.0255	0.0058	21.3	94.9	95.0	92.0		
		Independent	-13.9	1.7	2.1	-17.1	0.0068	0.0158	0.0252	0.0068	62.7	95.9	94.7	90.7		
		Exchangeable AR1	-2.7	1.8	2.6	-3.8	0.0048	0.0158	0.0258	0.0227	94.2	95.9	94.9	92.8		
			-29.7	2.7	-1.2	-15.4	0.0211	0.0171	0.0251	0.0212	56.5	93.7	93.5	89.8		
N= 500																
	WGEE	Independent		3.6	3.5	-1.4	0.0588	0.0980	0.0132		91.6	91.1	92.8			
		Exchangeable AR1	-29.3	3.3	3.0	0.3	0.0196	0.0524	0.0775	0.0140	53.5	92.5	92.0	91.9		
	GEE	Independent	-2.6	3.3	3.2	-0.5	0.0063	0.0500	0.0794	0.0128	90.3	92.1	91.6	92.4		
		Exchangeable AR1		3.5	-1.5	-97.1	0.0167	0.0241	0.0428		94.4	95.2	21.8			
	BPL ML	Exchangeable/AR1	-31.8	4.8	2.3	8.5	0.0194	0.0182	0.0255	0.0058	21.3	94.9	95.0	92.0		
		Independent	-13.9	1.7	2.1	-17.1	0.0068	0.0158	0.0252	0.0068	62.7	95.9	94.7	90.7		
		Exchangeable AR1	-2.7	1.8	2.6	-3.8	0.0048	0.0158	0.0258	0.0227	94.2	95.9	94.9	92.8		
			-29.7	2.7	-1.2	-15.4	0.0211	0.0171	0.0251	0.0212	56.5	93.7	93.5	89.8		
N= 500																
	WGEE	Independent		3.6	3.5	-1.4	0.0588	0.0980	0.0132		91.6	91.1	92.8			
		Exchangeable AR1	-29.3	3.3	3.0	0.3	0.0196	0.0524	0.0775	0.0140	53.5	92.5	92.0	91.9		
	GEE	Independent	-2.6	3.3	3.2	-0.5	0.0063	0.0500	0.0794	0.0128	90.3	92.1	91.6	92.4		
		Exchangeable AR1		3.5	-1.5	-97.1	0.0167	0.0241	0.0428		94.4	95.2	21.8			
	BPL ML	Exchangeable/AR1	-31.8	4.8	2.3	8.5	0.0194	0.0182	0.0255	0.0058	21.3	94.9	95.0	92.0		
		Independent	-13.9	1.7	2.1	-17.1	0.0068	0.0158	0.0252	0.0068	62.7	95.9	94.7	90.7		
		Exchangeable AR1	-2.7	1.8	2.6	-3.8	0.0048	0.0158	0.0258	0.0227	94.2	95.9	94.9	92.8		
			-29.7	2.7	-1.2	-15.4	0.0211	0.0171	0.0251	0.0212	56.5	93.7	93.5	89.8		
N= 500																
	WGEE	Independent		3.6	3.5	-1.4	0.0588	0.0980	0.0132		91.6	91.1	92.8			
		Exchangeable AR1	-29.3	3.3	3.0	0.3	0.0196	0.0524	0.0775	0.0140	53.5	92.5	92.0	91.9		
	GEE	Independent	-2.6	3.3	3.2	-0.5	0.0063	0.0500	0.0794	0.0128	90.3	92.1	91.6	92.4		
		Exchangeable AR1		3.5	-1.5	-97.1	0.0167	0.0241	0.0428		94.4	95.2	21.8			
	BPL ML	Exchangeable/AR1	-31.8	4.8	2.3	8.5	0.0194	0.0182	0.0255	0.0058	21.3	94.9	95.0	92.0		
		Independent	-13.9	1.7	2.1	-17.1	0.0068	0.0158	0.0252	0.0068	62.7	95.9	94.7	90.7		
		Exchangeable AR1	-2.7	1.8	2.6	-3.8	0.0048	0.0158	0.0258	0.0227	94.2	95.9	94.9	92.8		
			-29.7	2.7	-1.2	-15.4	0.0211	0.0171	0.0251	0.0212	56.5	93.7	93.5	89.8		
N= 500																
	WGEE	Independent		3.6	3.5	-1.4	0.0588	0.0980	0.0132		91.6	91.1				

Table 5.37: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **AR(1) correlation** ($\alpha = 0.2$) and **MAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C = (-1, 1, 0)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
WGEE	Independent Exchangeable AR1	Independent		2.1	1.9	1.6		0.0451	0.0629	0.0204		94.7	94.0	94.4		
		Exchangeable	-23.7	2.1	2.0	1.5	0.0077	0.0446	0.0626	0.0203	86.5	95.2	93.9	94.3		
	GEE	Independent	-3.7	2.1	2.0	1.9	0.0066	0.0446	0.0617	0.0206	94.1	95.0	93.7	94.6		
		Exchangeable		2.2	1.5	-20.8		0.0391	0.0542	0.0202		94.8	94.1	94.4		
BPL ML	Exchangeable/AR1	Independent	-30.2	2.5	2.1	4.6	0.0081	0.0397	0.0544	0.0188	78.8	94.7	94.7	94.9		
		Exchangeable	-9.8	2.0	2.1	0.8	0.0057	0.0394	0.0542	0.0185	91.6	94.8	94.8	94.3		
	ML	Independent	-10.4	1.8	1.8	5.4	0.0083	0.0388	0.0553	0.0732	92.1	95.0	94.4	89.5		
		Exchangeable		1.9	-6.5	-133.5		0.0362	0.0621	0.1209		93.8	89.2	59.0		
WGEE	Independent Exchangeable AR1	Independent	-36.9	2.3	0.2	-25.7	0.0130	0.0389	0.0564	0.0763	79.7	94.0	94.0	84.3		
		Exchangeable	-10.1	1.8	1.8	5.3	0.0083	0.0390	0.0551	0.0726	92.0	95.0	94.3	89.7		
	GEE	Independent Exchangeable AR1	Independent		0.5	0.2	-0.3		0.0177	0.0245	0.0082		95.4	94.7	94.7	
			Exchangeable	-21.9	0.6	0.2	-0.4	0.0038	0.0174	0.0244	0.0080	83.9	95.5	94.5	94.5	
BPL ML		Independent	-2.1	0.6	0.2	-0.2	0.0024	0.0174	0.0241	0.0082	95.6	95.9	94.8	94.8		
		Exchangeable		0.8	-0.3	-23.3		0.0152	0.0207	0.0094		94.7	95.4	91.8		
WGEE	Independent Exchangeable AR1	Independent	-28.8	1.1	0.2	2.4	0.0049	0.0153	0.0206	0.0072	68.2	95.0	95.5	94.6		
		Exchangeable	-8.5	0.5	0.2	-1.8	0.0022	0.0151	0.0206	0.0072	93.1	95.0	95.2	94.1		
	BPL ML	Independent	-4.9	0.4	-0.0	-1.3	0.0030	0.0153	0.0210	0.0273	93.4	94.6	95.2	92.7		
		Exchangeable		0.5	-9.7	-152.4		0.0144	0.0328	0.1129		93.6	83.1	33.1		
WGEE	Independent Exchangeable AR1	Independent	-32.8	1.2	-1.1	-30.2	0.0070	0.0152	0.0217	0.0306	76.9	94.5	95.2	90.2		
		Exchangeable	-4.8	0.4	-0.0	-0.3	0.0030	0.0153	0.0210	0.0274	93.0	94.7	95.5	93.0		

Table 5.38: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **AR(1) correlation** ($\alpha = 0.2$) and **MAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C = (-1, 1, 0)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE			Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		1.1	1.0	0.5		0.0484	0.0620	0.0122		94.9	95.4	94.9		
		Exchangeable	-37.7	1.1	0.9	0.7	0.0096	0.0471	0.0598	0.0125	70.2	95.2	94.8	95.1		
	GEE	AR1	-5.2	1.1	0.9	1.0	0.0056	0.0467	0.0590	0.0123	92.6	95.1	95.1	95.1		
		Independent		1.5	0.0	-15.8		0.0354	0.0440	0.0113		95.2	95.5	92.9		
	Exchangeable		-39.5	1.9	0.7	6.0	0.0094	0.0357	0.0443	0.0108	66.2	94.8	95.4	94.3		
		AR1	-8.6	0.9	0.7	0.4	0.0045	0.0351	0.0437	0.0103	91.3	94.6	95.7	94.4		
		Exchangeable/AR1	-8.4	0.8	0.5	1.8	0.0061	0.0352	0.0448	0.0300	92.8	94.8	95.3	91.6		
		Independent		3.6	-8.5	-92.4		0.0337	0.0561	0.0576		93.0	90.1	61.2		
BPL	Exchangeable	-45.3	2.6	-1.4	-21.8	0.0130	0.0353	0.0470	0.0334	68.4	93.8	95.1	83.9			
	AR1	-7.4	0.8	0.5	2.5	0.0060	0.0354	0.0445	0.0296	91.6	94.6	95.5	91.4			
N= 500																
	WGEE	Independent		1.1	0.8	1.5		0.0207	0.0268	0.0048		93.8	94.7	95.3		
		Exchangeable	-34.2	1.0	0.7	1.5	0.0062	0.0200	0.0256	0.0047	57.2	94.1	95.0	95.2		
	GEE	AR1	-0.5	1.1	0.8	1.7	0.0021	0.0199	0.0254	0.0047	95.7	93.8	95.3	94.6		
		Independent		1.7	-0.1	-16.0		0.0151	0.0207	0.0046		94.3	94.3	92.8		
	Exchangeable		-37.1	2.1	0.6	6.7	0.0067	0.0152	0.0206	0.0039	40.7	94.1	94.5	95.0		
		AR1	-5.3	1.0	0.6	0.7	0.0017	0.0148	0.0205	0.0036	93.2	94.6	94.5	95.6		
		Exchangeable/AR1	-2.7	1.1	0.4	-1.0	0.0024	0.0151	0.0205	0.0110	93.8	94.8	94.7	93.7		
		Independent		4.3	-9.9	-103.6		0.0155	0.0322	0.0519		92.0	81.5	32.4		
BPL	Exchangeable	-42.8	3.1	-1.5	-26.6	0.0093	0.0155	0.0215	0.0153	45.6	93.1	94.0	87.5			
	AR1	-2.6	1.0	0.3	-0.8	0.0023	0.0151	0.0203	0.0107	92.9	94.6	94.3	93.5			

Table 5.39: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **AR(1) correlation** ($\alpha = 0.4$) and **MAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C = (-1, 1, 0)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability			
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T
N= 200														
	WGEE	Independent		0.9	1.7	-0.5		0.0501	0.0738	0.0175		95.0	95.4	94.5
		Exchangeable	-15.3	1.0	1.9	0.2	0.0096	0.0484	0.0723	0.0171	84.8	95.0	95.9	94.7
	GEE	AR1	-0.3	1.0	2.0	0.5	0.0062	0.0478	0.0708	0.0177	93.3	95.1	95.7	94.4
		Independent		1.3	0.5	-52.1		0.0437	0.0613	0.0277		93.8	95.8	87.8
		Exchangeable	-21.5	2.1	1.8	4.6	0.0124	0.0439	0.0609	0.0157	70.9	94.0	96.0	94.4
		AR1	-5.4	1.3	1.9	-2.2	0.0059	0.0433	0.0604	0.0157	91.7	93.5	96.1	94.0
BPL	Exchangeable/AR1	-8.4	0.8	0.9	-4.1	0.0101	0.0439	0.0619	0.0893	93.1	93.7	95.7	88.5	
	ML	Independent		-0.2	-16.9	-246.8		0.0382	0.0780	0.2674		92.9	82.1	4.4
		Exchangeable	-29.7	1.4	-1.1	-56.1	0.0254	0.0446	0.0616	0.0884	72.1	93.4	95.4	80.6
		WGEE	AR1	-7.0	0.9	0.9	-2.6	0.0095	0.0435	0.0605	0.0821	91.0	94.4	95.7
Independent				0.4	0.3	0.2		0.0192	0.0310	0.0069		96.3	94.6	95.2
GEE		Exchangeable	-15.7	0.4	0.3	0.5	0.0060	0.0187	0.0310	0.0067	75.0	95.7	94.3	95.0
		AR1	-0.5	0.5	0.3	0.4	0.0022	0.0184	0.0304	0.0070	94.9	95.5	94.2	95.6
		Independent		0.8	-0.6	-52.4		0.0166	0.0261	0.0175		94.5	95.5	75.3
		Exchangeable	-22.0	1.4	0.5	3.9	0.0094	0.0168	0.0263	0.0060	42.6	94.8	94.9	95.3
BPL	AR1	-5.9	0.5	0.4	-3.5	0.0024	0.0165	0.0260	0.0061	91.2	94.9	94.9	95.7	
	ML	Exchangeable/AR1	-3.8	0.4	0.2	-2.4	0.0036	0.0164	0.0265	0.0367	95.5	95.0	94.4	92.3
		Independent		-0.7	-18.9	-251.7		0.0144	0.0556	0.2599		93.5	64.4	0.0
		ML	Exchangeable	-28.0	1.2	-1.9	-58.7	0.0176	0.0163	0.0274	0.0466	58.0	94.1	93.1
AR1			-3.3	0.4	0.1	-2.0	0.0031	0.0161	0.0254	0.0283	94.3	93.9	94.5	93.2
N= 500														

Table 5.40: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **AR(1) correlation** ($\alpha = 0.2$) and **MAR drop-outs** ($\gamma_0, \gamma_P, \gamma_C = (-1, 1, 0)$) with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE			Coverage Probability							
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T			
N= 200																	
	WGEE	Independent		2.8	2.8	0.8		0.0591	0.0798	0.0128		95.1	96.1	94.4			
		Exchangeable	-27.6	2.7	2.7	1.3	0.0167	0.0562	0.0738	0.0128	60.1	95.4	96.0	94.7			
	GEE	AR1	-1.0	2.7	2.7	1.4	0.0050	0.0549	0.0734	0.0126	94.2	95.0	96.2	94.8			
		Independent		3.2	0.5	-40.7		0.0445	0.0586	0.0174		94.9	96.3	86.5			
	GEE	Exchangeable	-31.2	3.8	1.8	9.1	0.0194	0.0457	0.0601	0.0109	42.4	94.6	95.4	93.4			
		AR1	-5.0	2.0	1.8	-1.1	0.0048	0.0436	0.0596	0.0105	91.2	94.2	95.5	93.0			
	BPL	Exchangeable/AR1	-7.1	2.0	1.1	-7.4	0.0070	0.0438	0.0594	0.0443	95.8	94.4	95.2	88.5			
		ML	Independent		6.5	-19.4	-179.0		0.0408	0.0819	0.1399		91.0	76.7	4.6		
	BPL	Exchangeable	-41.1	3.6	-3.8	-53.9	0.0364	0.0435	0.0596	0.0513	43.9	92.6	92.8	72.6			
		AR1	-2.1	0.6	-0.7	-0.1	0.0059	0.0415	0.0544	0.0335	91.3	94.1	95.0	90.0			
		WGEE	Independent		0.2	0.1	-0.8		0.0200	0.0309	0.0048		96.0	95.2	93.9		
			Exchangeable	-27.6	0.2	0.1	-0.9	0.0139	0.0187	0.0283	0.0048	29.0	96.1	95.1	94.4		
GEE		AR1	-0.9	0.2	0.1	-0.8	0.0020	0.0181	0.0277	0.0047	95.4	96.1	95.3	94.0			
		Independent		1.4	-1.0	-42.3		0.0152	0.0235	0.0111		95.2	94.4	72.5			
	GEE	Exchangeable	-31.2	2.1	0.5	7.4	0.0170	0.0155	0.0231	0.0041	10.9	95.1	95.2	93.9			
		AR1	-5.0	0.2	0.3	-3.4	0.0021	0.0147	0.0226	0.0039	90.5	95.1	95.1	94.5			
	BPL	Exchangeable/AR1	-2.8	0.3	0.2	-5.1	0.0023	0.0147	0.0232	0.0157	96.0	95.2	94.8	93.2			
		ML	Independent		4.9	-21.9	-182.5		0.0149	0.0646	0.1366		91.4	52.7	0.0		
	BPL	Exchangeable	-39.2	2.4	-4.5	-55.7	0.0292	0.0143	0.0265	0.0318	10.9	95.2	90.7	74.0			
		AR1	-0.6	-0.4	-0.9	-0.2	0.0016	0.0138	0.0213	0.0108	93.6	96.1	94.9	93.3			
	N= 500																

Drop-out mechanism: NI

For NI drop-out and $AR(1)$ correlation, the GEE and wGEE methods behaved similarly to those studied under exchangeable correlation.

The bivariate maximum pseudo-likelihood with $\gamma_P = 0$ showed some small bias for the time coefficient when $N = 200$. But the bias decreased to 0% when $N = 500$. When $\gamma_P \neq 0$, a small bias in $\hat{\beta}_T$ was observed, which did not disappear even when the sample size increased. For example in Table 5.55, the percentage relative bias for β_T was 10.1%. The increase in α slightly increased the relative bias. The increase in γ_C had a moderate effect on the bias. The bivariate maximum pseudo-likelihood was slightly less efficient than the ML with $AR(1)$ and the coverage probabilities required a sample size of $N = 500$ to be close to the nominal value 0.95. A sample size of $N = 200$ resulted in coverage probabilities close to 0.90. The ML with independence assumption was highly sensitive to the increase in α or γ_C . The increase in α from 0.2 to 0.4 nearly tripled the percentage relative bias for β_T . The increase in the number of responses decreased the relative bias.

Table 5.41: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **AR(1) correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability			
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T
N= 200														
	WGEE	Independent		2.9	1.1	-51.8		0.0450	0.0569	0.0290		94.1	94.5	87.0
		Exchangeable	-21.2	2.9	1.2	-53.0	0.0061	0.0453	0.0589	0.0292	87.7	93.8	94.7	86.1
	GEE	AR1	-2.1	2.8	1.2	-52.8	0.0055	0.0449	0.0577	0.0298	94.2	93.7	94.6	87.0
		Independent		2.9	1.3	-52.4		0.0403	0.0516	0.0285		94.4	95.0	86.4
	Exchangeable		-25.5	2.9	1.4	-50.4	0.0066	0.0403	0.0513	0.0276	83.3	94.4	95.2	86.4
		AR1	-3.3	2.7	1.4	-50.6	0.0050	0.0399	0.0513	0.0276	93.6	94.0	95.4	86.3
	BPL	Exchangeable/AR1	-8.2	1.4	1.0	-5.8	0.0046	0.0399	0.0510	0.0583	93.9	94.1	95.3	89.8
	ML	Independent		-0.1	-2.2	-30.2		0.0439	0.0553	0.0869		91.7	91.2	72.0
	Exchangeable		-29.2	1.3	0.7	-15.7	0.0073	0.0408	0.0519	0.0595	80.5	94.6	94.8	88.7
		AR1	-7.5	1.5	1.1	-6.7	0.0048	0.0401	0.0512	0.0579	93.3	94.4	95.2	89.7
	N= 500													
		WGEE	Independent		0.9	-0.1	-54.2		0.0151	0.0218	0.0190		96.5	95.6
Exchangeable			-21.0	0.9	-0.2	-55.1	0.0035	0.0152	0.0226	0.0193	80.4	96.2	95.1	73.6
GEE		AR1	-1.5	0.7	-0.2	-55.6	0.0022	0.0151	0.0221	0.0197	94.1	96.4	95.3	73.9
		Independent		1.2	0.2	-54.2		0.0140	0.0198	0.0186		96.0	95.4	73.8
	Exchangeable		-24.5	1.2	0.2	-51.6	0.0040	0.0140	0.0198	0.0174	75.1	95.8	95.8	74.9
		AR1	-1.9	0.9	0.2	-52.3	0.0020	0.0139	0.0197	0.0177	94.0	95.9	95.8	74.6
	BPL	Exchangeable/AR1	-2.9	-0.3	0.2	-3.8	0.0019	0.0140	0.0195	0.0234	94.0	95.4	95.8	93.3
	ML	Independent		-1.0	-2.0	-32.7		0.0158	0.0224	0.0510		92.8	91.2	78.1
	Exchangeable		-25.3	-0.4	-0.1	-11.6	0.0041	0.0141	0.0199	0.0244	74.0	95.0	95.3	92.1
		AR1	-2.9	-0.2	0.2	-4.1	0.0018	0.0140	0.0195	0.0233	93.7	95.2	95.7	93.4

Table 5.42: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **AR(1) correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		2.9	-0.4	-38.9		0.0400	0.0561	0.0152		94.6	93.6	87.2		
		Exchangeable AR1	-35.5	2.7	-0.5	-39.9	0.0078	0.0396	0.0566	0.0154	71.5	94.7	93.5	87.3		
	GEE	Independent	-2.8	2.4	-0.4	-40.3	0.0041	0.0388	0.0553	0.0156	95.5	94.9	93.7	87.0		
		Exchangeable AR1	-38.6	3.0	-0.2	-39.1		0.0353	0.0494	0.0149		95.3	93.5	85.6		
	BPL	Exchangeable/AR1	-3.1	2.5	-0.2	-38.6	0.0085	0.0353	0.0496	0.0144	65.2	95.1	93.1	86.3		
		Independent	-5.3	0.2	-0.3	-4.4	0.0036	0.0346	0.0492	0.0145	93.9	95.4	93.3	86.1		
	ML	Independent	-0.9	-2.9	-20.9		0.0402	0.0530	0.0334		94.6	95.7	93.4	91.5		
		Exchangeable AR1	-39.3	-0.2	-1.1	-10.0	0.0034	0.0348	0.0491	0.0235	92.2	88.9	78.7			
	WGEE	Independent	-4.7	0.2	-0.3	-4.2	0.0087	0.0369	0.0499	0.0270	64.0	95.4	93.2	87.2		
		Exchangeable					0.0034	0.0347	0.0490	0.0234	94.8	95.6	93.5	90.9		
	GEE	Independent		2.4	-0.2	-38.7		0.0163	0.0206	0.0099		95.3	95.6	75.7		
		Exchangeable AR1	-35.3	2.3	-0.4	-39.5	0.0061	0.0160	0.0205	0.0101	44.0	95.3	95.7	75.0		
	BPL	Independent	-2.2	1.9	-0.4	-40.1	0.0017	0.0157	0.0201	0.0103	95.0	95.5	95.7	74.5		
		Exchangeable AR1	-38.8	2.8	-0.1	-38.2		0.0146	0.0181	0.0093		95.6	95.9	75.3		
	ML	Independent	-3.1	2.7	-0.2	-36.1	0.0071	0.0144	0.0179	0.0087	33.1	95.6	95.7	76.9		
		Exchangeable/AR1	-2.5	2.2	-0.2	-37.4	0.0015	0.0141	0.0178	0.0091	94.2	95.8	95.5	75.3		
	WGEE	Independent	-0.0	-0.2	-1.6		0.0014	0.0142	0.0179	0.0080	95.6	95.5	95.5	95.0		
		Exchangeable	0.1	-1.2	-19.8		0.0163	0.0196	0.0144		93.3	92.0	86.4			
	GEE	Independent	-38.0	-0.0	-0.3	-6.9	0.0067	0.0150	0.0180	0.0092	31.6	94.8	95.5	92.2		
		Exchangeable AR1	-2.5	-0.0	-0.2	-1.6	0.0013	0.0141	0.0177	0.0080	95.9	95.5	95.6	95.2		
N= 500																

Table 5.43: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **AR(1) correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$.

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability				
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T	
N= 200															
	WGEE	Independent		1.6	1.0	-52.6		0.0436	0.0677	0.0276		94.9	95.9	87.5	
		Exchangeable	-16.9	1.6	1.2	-53.5	0.0089	0.0442	0.0717	0.0275	81.9	95.1	95.4	85.3	
	GEE	AR1	-2.1	1.2	1.1	-54.6	0.0048	0.0435	0.0693	0.0292	95.4	95.2	95.4	85.8	
		Independent		1.9	1.0	-60.1		0.0396	0.0597	0.0302		95.1	95.4	83.4	
		Exchangeable	-20.5	2.0	1.3	-49.2	0.0109	0.0396	0.0592	0.0249	74.7	95.0	96.0	86.5	
		AR1	-3.4	1.5	1.2	-51.3	0.0047	0.0388	0.0584	0.0256	94.5	94.8	95.7	85.5	
	BPL	Exchangeable/AR1		0.6	0.6	-12.7	0.0058	0.0396	0.0596	0.0732	93.1	95.1	95.3	88.4	
		Independent	-9.6	-3.4	-9.6	-86.4		0.0468	0.0640	0.1760		88.6	86.3	19.2	
	ML	Exchangeable	-27.2	-0.1	-1.0	-36.5	0.0167	0.0397	0.0606	0.0766	63.5	94.3	94.1	79.9	
		AR1	-7.8	0.9	0.7	-12.5	0.0056	0.0384	0.0585	0.0673	92.4	95.5	95.3	88.5	
		WGEE	Independent		1.0	-0.0	-51.1		0.0184	0.0263	0.0163		94.3	95.1	76.0
			Exchangeable	-15.3	1.0	-0.0	-52.4	0.0055	0.0183	0.0270	0.0167	71.2	94.2	94.9	73.5
GEE		AR1	-0.4	0.7	-0.0	-53.0	0.0019	0.0180	0.0260	0.0174	95.6	94.6	95.1	74.6	
		Independent		1.5	0.4	-59.8		0.0167	0.0242	0.0202		95.8	95.5	66.1	
		Exchangeable	-18.8	1.6	0.6	-48.9	0.0074	0.0164	0.0234	0.0152	58.0	95.3	95.3	75.0	
		AR1	-1.6	1.2	0.6	-50.6	0.0018	0.0160	0.0229	0.0158	94.6	95.6	95.4	73.7	
	BPL	Exchangeable/AR1		0.4	0.4	-5.9	0.0019	0.0160	0.0228	0.0312	96.5	95.6	95.8	93.2	
		Independent	-3.2	-4.0	-12.1	-103.3		0.0261	0.0375	0.1732		81.9	75.2	4.6	
	ML	Exchangeable	-21.8	-0.3	-0.4	-35.0	0.0097	0.0166	0.0233	0.0376	50.1	94.6	94.7	86.2	
		AR1	-2.4	0.5	0.5	-4.2	0.0017	0.0156	0.0222	0.0265	94.4	95.8	95.7	93.6	

Table 5.44: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **AR(1) correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability				
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T	
N= 200															
WGEE	Independent	Independent		2.8	0.7	-38.1		0.0496	0.0669	0.0162		94.9	94.9	86.5	
		Exchangeable	-28.2	2.6	0.5	-38.6	0.0160	0.0473	0.0647	0.0162	50.6	94.2	94.8	86.1	
	AR1	AR1	-2.0	2.1	0.6	-40.0	0.0038	0.0464	0.0628	0.0166	95.4	94.4	94.5	86.6	
		Independent		3.9	1.6	-44.4		0.0440	0.0583	0.0181	95.0	95.7	83.1		
GEE	Independent	Exchangeable	-31.7	3.8	1.5	-34.9	0.0190	0.0433	0.0565	0.0151	39.4	94.6	94.8	86.1	
		AR1	-3.3	2.9	1.6	-38.1	0.0036	0.0422	0.0560	0.0156	95.0	94.0	94.4	84.8	
	Exchangeable/AR1	Exchangeable	-5.5	1.5	1.4	-2.1	0.0036	0.0424	0.0575	0.0309	96.4	95.0	93.6	92.5	
		Independent		-1.9	-10.4	-64.6		0.0613	0.0656	0.0815	81.2	81.7	28.5		
BPL	Independent	Exchangeable	-35.6	0.2	-1.1	-25.9	0.0238	0.0441	0.0569	0.0391	27.4	91.5	92.6	75.8	
		AR1	-0.8	0.9	0.4	3.4	0.0029	0.0391	0.0519	0.0220	94.5	95.2	94.8	92.0	
	WGEE	Independent	Independent		2.4	-0.0	-37.7		0.0183	0.0263	0.0095		94.7	94.1	75.1
			Exchangeable	-26.9	2.3	-0.1	-38.6	0.0130	0.0177	0.0258	0.0098	18.9	94.9	94.7	74.6
AR1		AR1	-0.5	1.7	-0.2	-39.7	0.0016	0.0173	0.0249	0.0101	94.6	95.1	94.1	73.7	
		Independent		3.0	0.1	-43.8		0.0167	0.0227	0.0115	94.9	94.9	67.3		
GEE	Independent	Exchangeable	-30.3	3.0	0.2	-34.4	0.0160	0.0166	0.0221	0.0085	9.7	95.1	94.5	76.7	
		AR1	-1.8	2.1	0.1	-37.5	0.0015	0.0160	0.0218	0.0092	93.5	95.3	94.2	73.6	
	Exchangeable/AR1	Exchangeable	-2.0	0.8	0.1	-1.5	0.0012	0.0160	0.0218	0.0120	96.6	95.6	94.3	93.9	
		Independent		-0.9	-13.4	-84.5		0.0354	0.0409	0.0807	71.2	69.6	10.9		
BPL	Independent	Exchangeable	-32.3	-0.0	-1.3	-27.8	0.0183	0.0163	0.0221	0.0203	2.8	93.4	92.8	80.8	
		AR1	-0.2	0.5	-0.3	2.4	0.0010	0.0144	0.0195	0.0086	93.5	95.7	93.9	94.4	
	N= 500														

Table 5.45: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **AR(1) correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T			
N= 200																	
	WGEE	Independent		3.3	-0.0	-124.8		0.0446	0.0633	0.0828		95.5	94.9	57.0			
		Exchangeable	-27.4	3.1	-0.1	-127.1	0.0080	0.0439	0.0631	0.0848	84.6	95.3	94.6	55.3			
	GEE	AR1	-8.8	2.7	-0.1	-127.6	0.0065	0.0435	0.0622	0.0857	92.2	95.1	94.8	56.0			
		Independent		3.9	0.9	-125.7		0.0418	0.0586	0.0829		94.9	95.1	54.4			
	Exchangeable		-31.5	3.8	1.0	-121.0	0.0087	0.0413	0.0577	0.0782	79.3	94.7	94.9	55.5			
		AR1	-10.8	3.3	1.0	-122.1	0.0062	0.0409	0.0574	0.0792	91.9	94.5	95.0	55.6			
	BPL	Exchangeable/AR1	-10.2	1.0	1.1	-7.3	0.0054	0.0401	0.0571	0.0636	94.1	95.1	94.8	92.7			
	ML	Independent		0.3	-0.6	-53.9		0.0421	0.0605	0.1009	93.5	92.0	79.8				
	Exchangeable		-31.3	0.7	0.7	-22.1	0.0083	0.0404	0.0574	0.0680	82.0	95.2	95.3	91.3			
		AR1	-10.0	1.0	1.1	-8.1	0.0054	0.0401	0.0569	0.0638	94.0	95.0	95.0	91.9			
	N= 500																
		WGEE	Independent		2.7	-0.5	-123.0		0.0174	0.0236	0.0683		95.1	95.1	20.4		
Exchangeable			-24.5	2.6	-0.6	-124.8	0.0044	0.0173	0.0237	0.0701	76.2	95.0	94.9	19.5			
AR1		-5.8	2.2	-0.6	-125.8	0.0026	0.0170	0.0233	0.0712	92.7	95.0	95.2	18.9				
		Independent		3.7	0.7	-123.5		0.0174	0.0222	0.0687		93.8	94.7	19.5			
	GEE	Exchangeable	-28.3	3.7	0.8	-118.1	0.0050	0.0174	0.0218	0.0636	69.2	93.7	95.1	23.8			
		AR1	-7.4	3.1	0.8	-119.6	0.0025	0.0169	0.0218	0.0650	92.3	93.8	95.0	22.6			
	BPL	Exchangeable/AR1	-4.0	0.8	1.0	-0.9	0.0022	0.0162	0.0216	0.0240	93.8	94.7	94.6	94.4			
	ML	Independent		0.6	0.0	-47.2		0.0180	0.0237	0.0535	92.6	90.9	83.3				
	Exchangeable		-26.0	0.6	0.9	-12.5	0.0043	0.0163	0.0217	0.0255	74.1	94.8	94.7	93.5			
		AR1	-4.0	0.8	1.0	-0.8	0.0021	0.0162	0.0216	0.0240	93.5	95.1	94.6	94.4			

Table 5.46: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($\mathbf{T} = 4$), **AR(1) correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		4.9	-1.4	-92.3		0.0413	0.0547	0.0451			95.5	95.4	60.4	
		Exchangeable	-38.6	4.7	-1.4	-93.7	0.0097	0.0404	0.0537	0.0461		69.4	95.4	95.6	58.8	
	GEE	AR1	-7.1	3.9	-1.5	-95.1	0.0054	0.0393	0.0531	0.0472		92.5	95.7	95.7	58.0	
		Independent		5.6	-0.2	-93.2		0.0382	0.0485	0.0452			94.7	95.2	56.9	
	Exchangeable	AR1	-40.9	5.6	0.0	-89.0	0.0101	0.0384	0.0487	0.0420		64.9	94.8	95.4	59.2	
		Independent		4.6	-0.1	-91.8	0.0050	0.0371	0.0484	0.0439		91.6	95.3	95.4	57.1	
	BPL	Exchangeable/AR1	-7.2	0.2	0.5	-5.0	0.0043	0.0365	0.0478	0.0289		94.6	95.8	95.4	92.5	
	ML	Independent		-0.1	-1.0	-38.2		0.0401	0.0495	0.0433			93.2	92.4	83.3	
Exchangeable		-39.2	-0.1	0.0	-11.5	0.0095	0.0379	0.0489	0.0317		66.6	94.9	94.9	89.8		
	AR1		-6.9	0.1	0.5	-5.1	0.0042	0.0366	0.0480	0.0283		94.1	95.8	95.4	92.7	
		Independent		5.6	-0.6	-89.5		0.0197	0.0222	0.0367			93.4	95.1	24.6	
	WGEE	Exchangeable	-36.0	5.5	-0.7	-91.2	0.0066	0.0192	0.0215	0.0379		51.2	93.1	95.5	24.1	
		AR1	-4.1	4.6	-0.7	-92.6	0.0019	0.0182	0.0212	0.0389		94.6	93.7	95.2	22.7	
	GEE	Independent		6.3	0.3	-91.5		0.0189	0.0190	0.0379			92.5	95.4	20.5	
		Exchangeable	-38.8	6.2	0.5	-87.1	0.0073	0.0188	0.0188	0.0348		40.1	92.5	95.5	24.2	
	AR1		-6.0	5.1	0.4	-90.0	0.0018	0.0175	0.0187	0.0368		93.1	92.8	96.0	21.7	
		BPL	Exchangeable/AR1	-1.8	0.7	0.9	-0.8	0.0015	0.0153	0.0185	0.0110		95.5	94.4	95.6	93.3
	ML	Independent		1.2	0.5	-33.4		0.0167	0.0197	0.0190			92.8	92.5	84.4	
		Exchangeable	-35.2	0.7	0.8	-7.8	0.0061	0.0159	0.0188	0.0123		46.8	93.9	95.5	89.8	
	AR1		-1.7	0.7	0.9	-0.3	0.0014	0.0153	0.0184	0.0106		95.6	94.5	95.4	93.9	

Table 5.47: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **AR(1) correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T			
N= 200																	
	WGEE	Independent		1.8	-1.2	-119.3		0.0487	0.0703	0.0759		94.6	94.9	58.6			
		Exchangeable	-16.4	1.6	-1.3	-121.7	0.0097	0.0472	0.0683	0.0778	83.8	94.1	95.2	55.8			
	GEE	AR1	-1.8	1.0	-1.3	-123.5	0.0059	0.0468	0.0670	0.0801	94.6	94.4	95.2	55.2			
		Independent		3.2	0.5	-139.8		0.0456	0.0660	0.0970		95.2	94.8	45.2			
	Exchangeable		-21.3	3.4	0.8	-114.9	0.0124	0.0450	0.0637	0.0712	73.6	94.5	95.2	58.0			
		AR1	-5.1	2.3	0.7	-119.2	0.0059	0.0443	0.0632	0.0750	91.7	95.0	94.9	54.4			
	BPL	Exchangeable/AR1	-7.4	1.1	0.6	-15.2	0.0064	0.0443	0.0651	0.0894	93.8	94.5	95.1	90.0			
		ML	Independent		-0.8	-9.9	-192.6		0.0479	0.0752	0.3252		89.6	84.4	24.0		
	Exchangeable		-24.9	0.1	-0.1	-44.3	0.0157	0.0442	0.0632	0.0957	72.2	94.0	94.3	82.6			
		AR1	-5.9	1.1	0.8	-10.7	0.0060	0.0429	0.0638	0.0807	91.8	94.6	94.5	89.3			
	N= 500																
		WGEE	Independent		1.0	-1.9	-122.0		0.0204	0.0289	0.0665		94.2	94.4	18.7		
Exchangeable			-16.6	0.7	-2.3	-124.6	0.0066	0.0198	0.0285	0.0688	69.8	94.0	93.9	15.6			
GEE		AR1	-2.0	0.0	-2.3	-126.1	0.0024	0.0196	0.0278	0.0708	94.4	94.0	94.1	16.5			
		Independent		2.6	-0.1	-142.2		0.0200	0.0275	0.0878		93.2	94.6	7.6			
Exchangeable			-21.5	2.5	-0.2	-117.4	0.0095	0.0197	0.0266	0.0617	51.6	93.6	95.1	18.6			
		AR1	-5.3	1.5	-0.2	-121.5	0.0027	0.0191	0.0262	0.0657	91.3	93.7	95.0	15.3			
	BPL	Exchangeable/AR1	-3.3	0.3	-0.2	-4.5	0.0021	0.0192	0.0266	0.0370	94.5	92.5	94.2	92.1			
		ML	Independent		-0.1	-12.2	-243.0		0.0219	0.0446	0.3641		86.8	75.2	12.3		
	Exchangeable		-20.8	-0.7	-0.4	-37.6	0.0092	0.0196	0.0264	0.0420	53.7	92.2	94.1	87.2			
		AR1	-2.5	0.4	-0.0	-3.3	0.0020	0.0188	0.0258	0.0313	94.1	93.3	94.7	92.5			

Table 5.48: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **AR(1) correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 0, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability				
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T	
N= 200															
	WGEE	Independent		4.8	-0.6	-98.0		0.0483	0.0728	0.0504		96.4	95.9	56.1	
		Exchangeable	-29.5	4.3	-1.0	-100.2	0.0183	0.0450	0.0669	0.0521	57.5	96.3	95.8	54.8	
	GEE	AR1	-3.8	3.0	-1.0	-102.4	0.0051	0.0427	0.0644	0.0536	93.3	96.5	95.9	52.7	
		Independent		6.3	1.2	-113.1		0.0435	0.0591	0.0628		95.0	95.3	42.7	
		Exchangeable	-34.2	6.2	1.4	-92.5	0.0226	0.0428	0.0567	0.0458	38.0	95.2	95.8	56.0	
		AR1	-7.8	4.2	1.2	-99.3	0.0054	0.0397	0.0552	0.0505	89.7	95.8	95.3	50.2	
BPL	Exchangeable/AR1	-5.9	1.6	1.4	-10.0	0.0041	0.0388	0.0549	0.0424	95.4	95.9	95.0	89.7		
	ML	Independent		3.5	-9.0	-142.9		0.0523	0.0690	0.1681		86.7	85.2	34.1	
		Exchangeable	-34.0	-0.6	-0.9	-33.2	0.0230	0.0425	0.0544	0.0524	39.5	93.3	94.3	78.5	
			AR1	-0.9	0.6	0.2	-3.5	0.0039	0.0381	0.0534	0.0297	92.7	94.7	95.1	91.5
N= 500															
	WGEE	Independent		3.5	-1.5	-96.0		0.0207	0.0297	0.0413		95.7	95.0	18.8	
		Exchangeable	-27.7	3.0	-1.9	-98.2	0.0138	0.0193	0.0278	0.0429	27.6	95.6	94.7	17.7	
		AR1	-1.8	1.7	-1.8	-100.7	0.0018	0.0183	0.0273	0.0447	96.2	95.5	95.2	13.9	
	GEE	Independent		5.2	0.2	-111.2		0.0192	0.0247	0.0538		93.5	95.7	8.0	
		Exchangeable	-32.4	5.1	0.3	-91.0	0.0182	0.0187	0.0239	0.0372	8.2	94.4	95.6	19.6	
		AR1	-5.8	3.1	0.1	-97.9	0.0021	0.0169	0.0236	0.0423	91.6	94.7	95.5	12.2	
	BPL	Exchangeable/AR1	-2.1	0.6	0.7	-0.4	0.0015	0.0164	0.0235	0.0151	95.6	94.8	95.2	93.0	
		ML	Independent		4.1	-11.7	-166.1		0.0270	0.0444	0.1736		81.2	70.9	23.2
			Exchangeable	-30.1	-1.7	-0.5	-22.4	0.0159	0.0165	0.0223	0.0182	6.4	93.5	94.2	82.9
			AR1	0.0	0.3	0.0	1.6	0.0011	0.0147	0.0218	0.0108	94.1	94.9	94.9	93.5

Table 5.49: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **AR(1) correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 1, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability			
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T
N= 200														
	WGEE	Independent		0.8	-2.1	-70.3		0.0464	0.0673	0.0411		96.0	95.1	86.3
		Exchangeable	-35.2	0.7	-2.0	-71.4	0.0114	0.0454	0.0660	0.0415	79.6	96.1	95.2	86.4
	GEE	AR1	-17.6	0.6	-2.0	-71.1	0.0091	0.0452	0.0656	0.0415	90.3	96.0	95.1	86.4
		Independent		1.8	-0.8	-92.0		0.0390	0.0541	0.0524		95.0	95.0	74.3
	Exchangeable	AR1	-41.4	2.2	-0.1	-65.7	0.0119	0.0392	0.0542	0.0362	70.1	95.5	94.8	84.1
		Independent	-24.3	1.5	-0.1	-69.8	0.0084	0.0389	0.0539	0.0379	85.7	95.6	95.0	83.4
	BPL	Exchangeable/AR1	-7.7	0.6	0.5	-1.4	0.0092	0.0393	0.0556	0.0832	92.0	95.9	95.4	90.3
		ML	Independent		1.7	-5.2	-159.2		0.0380	0.0638	0.1662		94.8	91.3
	Exchangeable	AR1	-33.9	1.3	-0.2	-35.5	0.0128	0.0397	0.0571	0.0867	84.0	95.3	94.5	88.7
		Independent	-7.7	0.6	0.4	-1.8	0.0092	0.0395	0.0559	0.0833	91.0	95.7	95.4	90.5
	N= 500													
		WGEE	Independent		-0.2	-3.2	-69.6		0.0182	0.0288	0.0280		96.0	94.5
Exchangeable			-34.2	-0.3	-3.2	-70.2	0.0072	0.0177	0.0282	0.0282	71.7	96.7	94.8	70.8
GEE		AR1	-16.9	-0.5	-3.2	-70.7	0.0042	0.0177	0.0279	0.0286	90.1	96.5	94.3	70.6
		Independent		1.5	-1.4	-91.2		0.0150	0.0224	0.0410		95.6	95.0	45.7
	Exchangeable	AR1	-40.7	1.8	-0.8	-64.3	0.0085	0.0151	0.0222	0.0243	48.5	95.7	95.2	69.7
		Independent	-23.5	1.1	-0.8	-69.2	0.0045	0.0149	0.0221	0.0267	78.9	95.4	95.1	66.9
	BPL	Exchangeable/AR1	-2.8	0.3	0.3	-0.1	0.0036	0.0151	0.0223	0.0329	92.0	95.6	95.7	92.7
		ML	Independent		1.5	-5.6	-160.2		0.0148	0.0282	0.1328		94.6	90.1
	Exchangeable	AR1	-29.3	1.1	-0.1	-26.8	0.0064	0.0152	0.0223	0.0342	79.5	94.9	95.4	91.5
		Independent	-3.1	0.2	0.3	-0.2	0.0036	0.0152	0.0223	0.0326	91.7	95.5	95.5	92.6

Table 5.50: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **AR(1) correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 1, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE			Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		2.1	-3.7	-53.1		0.0526	0.0789		94.7	94.1	85.1			
		Exchangeable	-47.4	2.0	-3.7	-54.3	0.0145	0.0510	0.0749		64.5	94.8	94.1	84.8		
	GEE	Independent	-18.6	1.7	-3.5	-54.3	0.0088	0.0504	0.0744		87.2	95.1	94.0	85.8		
		Exchangeable	-50.2	4.0	-1.3	-69.2		0.0353	0.0493	0.0300		94.2	95.1	75.1		
	BPL	Independent	-24.1	4.4	-0.5	-47.2	0.0139	0.0358	0.0494		52.0	94.3	95.1	83.7		
		Exchangeable	-5.5	3.1	-0.5	-53.2	0.0072	0.0348	0.0490		83.0	94.3	94.9	82.2		
	ML	Independent	-41.3	1.2	0.8	2.1	0.0068	0.0359	0.0506		92.6	95.4	95.1	92.9		
		Exchangeable	-4.7	4.9	-4.7	-101.4		0.0360	0.0563	0.0689		93.1	91.3	68.0		
	WGEE	Independent	-47.4	3.1	-0.2	-21.8	0.0134	0.0366	0.0520		71.5	94.3	94.4	89.0		
		Exchangeable	-4.7	1.1	0.7	2.4	0.0067	0.0360	0.0508		91.6	95.4	94.7	92.8		
		WGEE	Independent		0.6	-3.9	-53.1		0.0218	0.0302		94.5	93.9	72.6		
			Exchangeable	-45.0	0.4	-3.9	-54.1	0.0102	0.0211	0.0289		47.9	94.6	93.6	72.1	
GEE		Independent	-15.0	0.0	-3.8	-54.7	0.0037	0.0209	0.0287		87.1	94.4	93.9	70.7		
		Exchangeable	-48.7	2.7	-1.9	-69.5		0.0157	0.0201	0.0238		94.1	94.6	45.0		
	BPL	Independent	-21.8	3.0	-1.1	-47.1	0.0109	0.0161	0.0203		25.2	94.2	94.8	70.2		
		Exchangeable	-0.2	1.6	-1.1	-53.8	0.0038	0.0154	0.0201		77.8	94.5	94.8	62.4		
	ML	Independent	-38.5	-0.3	0.5	1.7	0.0025	0.0156	0.0207		94.3	94.0	95.2	94.4		
		Exchangeable	-0.7	3.7	-4.8	-101.9	0.0081	0.0162	0.0254		92.2	90.1	46.6			
N= 500																
	WGEE	Independent	-38.5	1.8	-0.3	-20.5	0.0081	0.0161	0.0212		57.8	93.7	94.4	91.0		
		Exchangeable	-0.7	-0.4	0.3	1.1	0.0024	0.0158	0.0205		94.2	93.8	95.3	94.6		
		WGEE	Independent		0.6	-3.9	-53.1		0.0218	0.0302		94.5	93.9	72.6		
			Exchangeable	-45.0	0.4	-3.9	-54.1	0.0102	0.0211	0.0289		47.9	94.6	93.6	72.1	
GEE		Independent	-15.0	0.0	-3.8	-54.7	0.0037	0.0209	0.0287		87.1	94.4	93.9	70.7		
		Exchangeable	-48.7	2.7	-1.9	-69.5		0.0157	0.0201	0.0238		94.1	94.6	45.0		
	BPL	Independent	-21.8	3.0	-1.1	-47.1	0.0109	0.0161	0.0203		25.2	94.2	94.8	70.2		
		Exchangeable	-0.2	1.6	-1.1	-53.8	0.0038	0.0154	0.0201		77.8	94.5	94.8	62.4		
	ML	Independent	-38.5	-0.3	0.5	1.7	0.0025	0.0156	0.0207		94.3	94.0	95.2	94.4		
		Exchangeable	-0.7	3.7	-4.8	-101.9	0.0081	0.0161	0.0212		92.2	90.1	46.6			

Table 5.51: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **AR(1) correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 1, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$.

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability			
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T
N= 200														
	WGEE	Independent		1.3	-1.2	-63.2		0.0518	0.0829	0.0377		95.9	95.1	87.3
		Exchangeable	-22.2	1.2	-1.1	-64.8	0.0149	0.0482	0.0759	0.0379	79.2	96.4	95.0	87.9
	GEE	AR1	-8.0	0.8	-1.1	-65.4	0.0087	0.0480	0.0756	0.0388	92.8	96.4	95.2	87.8
		Independent		3.0	-0.1	-127.6		0.0405	0.0613	0.0839		95.2	95.6	55.5
	Exchangeable	-30.1	3.8	1.6	-58.8	0.0198	0.0411	0.0611	0.0315	54.9	95.6	95.3	85.7	
		AR1	-15.8	2.4	1.4	-68.7	0.0101	0.0403	0.0609	0.0367	83.1	95.2	95.4	81.8
	BPL	Exchangeable/AR1	-9.1	1.9	1.6	-6.8	0.0102	0.0416	0.0627	0.1065	93.8	95.3	95.3	89.2
	ML	Independent		1.6	-16.1	-321.6		0.0349	0.0789	0.4507		94.5	84.9	7.8
	Exchangeable	-28.3	2.3	0.3	-53.5	0.0237	0.0412	0.0604	0.1040	76.1	95.0	96.0	84.9	
		AR1	-8.1	2.0	1.6	-7.0	0.0096	0.0405	0.0596	0.0943	92.8	94.9	95.7	89.4
	WGEE	Independent		-1.6	-4.5	-66.7		0.0225	0.0375	0.0268		94.9	93.3	70.9
		Exchangeable	-20.7	-1.7	-4.4	-68.0	0.0095	0.0217	0.0353	0.0273	65.6	94.6	93.7	68.5
	GEE	AR1	-6.1	-2.0	-4.3	-68.6	0.0034	0.0218	0.0351	0.0277	93.5	94.6	93.6	68.5
		Independent		0.5	-2.7	-132.2		0.0178	0.0276	0.0778		94.5	94.4	15.7
	Exchangeable	-29.3	1.2	-1.3	-63.0	0.0158	0.0181	0.0272	0.0235	23.8	93.8	94.6	65.7	
		AR1	-14.7	-0.2	-1.4	-72.9	0.0057	0.0180	0.0272	0.0289	73.1	93.5	94.3	55.5
	BPL	Exchangeable/AR1	-2.7	-0.4	-0.6	-2.6	0.0031	0.0184	0.0276	0.0434	96.1	93.4	94.3	92.6
	ML	Independent		-1.0	-20.8	-346.1		0.0158	0.0653	0.4882		92.7	63.9	0.1
	Exchangeable	-24.3	0.3	-1.4	-53.5	0.0140	0.0184	0.0277	0.0474	66.0	93.6	93.7	89.9	
		AR1	-2.3	-0.4	-0.6	-0.2	0.0030	0.0179	0.0271	0.0359	94.9	93.5	93.9	93.1

Table 5.52: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **AR(1) correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 1, 0.5)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$.

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		2.4	-2.1	-55.4		0.0715	0.1153	0.0277		94.1	92.4	87.4		
		Exchangeable	-32.8	2.2	-2.3	-54.9	0.0234	0.0663	0.1016	0.0281	58.3	94.0	92.7	88.0		
	GEE	Independent	-7.3	1.3	-2.3	-56.6	0.0077	0.0647	0.1011	0.0277	90.6	93.9	92.5	86.8		
		Exchangeable	-38.3	4.5	-1.5	-105.4		0.0445	0.0658	0.0558		94.2	93.3	49.3		
	BPL	Independent	-15.4	5.2	0.3	-44.9	0.0279	0.0459	0.0668	0.0198	30.7	93.6	93.5	86.0		
		Exchangeable	-4.5	2.5	0.1	-60.0	0.0087	0.0430	0.0659	0.0254	81.1	93.8	93.3	77.7		
	ML	Independent	-34.3	7.7	-18.8	-240.1	0.0063	0.0430	0.0664	0.0465	96.0	94.0	93.7	92.0		
		Exchangeable	0.2	2.4	-2.7	-40.4	0.0293	0.0453	0.0680	0.0607	59.4	91.5	91.7	79.3		
	WGEE	Independent		0.8	-0.3	5.3	0.0064	0.0412	0.0618	0.0358	90.9	94.9	94.0	92.1		
		Exchangeable	-31.4	-1.2	-4.6	-56.5	0.0184	0.0241	0.0392	0.0192	33.3	95.3	93.9	71.4		
	GEE	Independent	-5.9	-1.9	-4.5	-57.7	0.0035	0.0237	0.0391	0.0196	91.0	95.8	93.7	68.9		
		Exchangeable	-37.8	2.1	-2.7	-106.6		0.0172	0.0252	0.0504		93.9	94.5	13.4		
	BPL	Independent	-14.8	2.8	-0.8	-46.8	0.0247	0.0176	0.0243	0.0136	5.6	93.8	94.7	66.6		
		Exchangeable	-2.0	0.1	-1.0	-62.0	0.0056	0.0166	0.0242	0.0202	66.7	94.6	94.3	51.1		
	ML	Independent	-31.7	5.6	-21.9	-249.2	0.0023	0.0169	0.0244	0.0199	96.0	94.4	94.4	92.3		
		Exchangeable	0.2	0.3	-2.2	-39.0	0.0198	0.0170	0.0245	0.0248	22.0	93.0	93.1	86.3		
N= 500																
	WGEE	Independent		-0.9	-4.6	-55.2		0.0260	0.0445	0.0188		95.7	93.5	70.3		
		Exchangeable	-31.4	-1.2	-4.6	-56.5	0.0184	0.0241	0.0392	0.0192	33.3	95.3	93.9	71.4		
	GEE	Independent	-5.9	-1.9	-4.5	-57.7	0.0035	0.0237	0.0391	0.0196	91.0	95.8	93.7	68.9		
		Exchangeable	-37.8	2.1	-2.7	-106.6		0.0172	0.0252	0.0504		93.9	94.5	13.4		
	BPL	Independent	-14.8	2.8	-0.8	-46.8	0.0247	0.0176	0.0243	0.0136	5.6	93.8	94.7	66.6		
		Exchangeable	-2.0	0.1	-1.0	-62.0	0.0056	0.0166	0.0242	0.0202	66.7	94.6	94.3	51.1		
	ML	Independent	-31.7	5.6	-21.9	-249.2	0.0023	0.0169	0.0244	0.0199	96.0	94.4	94.4	92.3		
		Exchangeable	0.2	0.3	-2.2	-39.0	0.0198	0.0170	0.0245	0.0248	22.0	93.0	93.1	86.3		
N= 500																
	WGEE	Independent		0.8	-0.3	5.3	0.0064	0.0412	0.0618	0.0358	90.9	94.9	94.0	92.1		
		Exchangeable	-31.4	-1.2	-4.6	-56.5	0.0184	0.0241	0.0392	0.0192	33.3	95.3	93.9	71.4		
	GEE	Independent	-5.9	-1.9	-4.5	-57.7	0.0035	0.0237	0.0391	0.0196	91.0	95.8	93.7	68.9		
		Exchangeable	-37.8	2.1	-2.7	-106.6		0.0172	0.0252	0.0504		93.9	94.5	13.4		
	BPL	Independent	-14.8	2.8	-0.8	-46.8	0.0247	0.0176	0.0243	0.0136	5.6	93.8	94.7	66.6		
		Exchangeable	-2.0	0.1	-1.0	-62.0	0.0056	0.0166	0.0242	0.0202	66.7	94.6	94.3	51.1		
	ML	Independent	-31.7	5.6	-21.9	-249.2	0.0023	0.0169	0.0244	0.0199	96.0	94.4	94.4	92.3		
		Exchangeable	0.2	0.3	-2.2	-39.0	0.0198	0.0170	0.0245	0.0248	22.0	93.0	93.1	86.3		
N= 500																
	WGEE	Independent		0.8	-0.3	5.3	0.0064	0.0412	0.0618	0.0358	90.9	94.9	94.0	92.1		
		Exchangeable	-31.4	-1.2	-4.6	-56.5	0.0184	0.0241	0.0392	0.0192	33.3	95.3	93.9	71.4		
	GEE	Independent	-5.9	-1.9	-4.5	-57.7	0.0035	0.0237	0.0391	0.0196	91.0	95.8	93.7	68.9		
		Exchangeable	-37.8	2.1	-2.7	-106.6		0.0172	0.0252	0.0504		93.9	94.5	13.4		
	BPL	Independent	-14.8	2.8	-0.8	-46.8	0.0247	0.0176	0.0243	0.0136	5.6	93.8	94.7	66.6		
		Exchangeable	-2.0	0.1	-1.0	-62.0	0.0056	0.0166	0.0242	0.0202	66.7	94.6	94.3	51.1		
	ML	Independent	-31.7	5.6	-21.9	-249.2	0.0023	0.0169	0.0244	0.0199	96.0	94.4	94.4	92.3		
		Exchangeable	0.2	0.3	-2.2	-39.0	0.0198	0.0170	0.0245	0.0248	22.0	93.0	93.1	86.3		
N= 500																
	WGEE	Independent		0.8	-0.3	5.3	0.0064	0.0412	0.0618	0.0358	90.9	94.9	94.0	92.1		
		Exchangeable	-31.4	-1.2	-4.6	-56.5	0.0184	0.0241	0.0392	0.0192	33.3	95.3	93.9	71.4		
	GEE	Independent	-5.9	-1.9	-4.5	-57.7	0.0035	0.0237	0.0391	0.0196	91.0	95.8	93.7	68.9		
		Exchangeable	-37.8	2.1	-2.7	-106.6		0.0172	0.0252	0.0504		93.9	94.5	13.4		
	BPL	Independent	-14.8	2.8	-0.8	-46.8	0.0247	0.0176	0.0243	0.0136	5.6	93.8	94.7	66.6		
		Exchangeable	-2.0	0.1	-1.0	-62.0	0.0056	0.0166	0.0242	0.0202	66.7	94.6	94.3	51.1		
	ML	Independent	-31.7	5.6	-21.9	-249.2	0.0023	0.0169	0.0244	0.0199	96.0	94.4	94.4	92.3		
		Exchangeable	0.2	0.3	-2.2	-39.0	0.0198	0.0170	0.0245	0.0248	22.0	93.0	93.1	86.3		
N= 500																
	WGEE	Independent		0.8	-0.3	5.3	0.0064	0.0412	0.0618	0.0358	90.9	94.9	94.0	92.1		
		Exchangeable	-31.4	-1.2	-4.6	-56.5	0.0184	0.0241	0.0392	0.0192	33.3	95.3	93.9	71.4		
	GEE	Independent	-5.9	-1.9	-4.5	-57.7	0.0035	0.0237	0.0391	0.0196	91.0	95.8	93.7	68.9		
		Exchangeable	-37.8	2.1	-2.7	-106.6		0.0172	0.0252	0.0504		93.9	94.5	13.4		
	BPL	Independent	-14.8	2.8	-0.8	-46.8	0.0247	0.0176	0.0243	0.0136	5.6	93.8	94.7	66.6		
		Exchangeable	-2.0	0.1	-1.0	-62.0	0.0056	0.0166	0.0242	0.0202	66.7	94.6	94.3	51.1		
	ML	Independent	-31.7	5.6	-21.9	-249.2	0.0023	0.0169	0.0244	0.0199	96.0	94.4	94.4	92.3		
		Exchangeable	0.2	0.3	-2.2	-39.0	0.0198	0.0170	0.0245	0.0248	22.0	93.0	93.1	86.3		
N= 500																
	WGEE	Independent		0.8	-0.3	5.3	0.0064	0.0412	0.0618	0.0358	90.9	94.9	94.0	92.1		
		Exchangeable	-31.4	-1.2	-4.6	-56.5	0.0184	0.0241	0.0392	0.0192	33.3	95.3	93.9	71.4		
	GEE	Independent	-5.9	-1.9	-4.5	-57.7	0.0035	0.0237	0.0391	0.0196	91.0	95.8	93.7	68.9		
		Exchangeable	-37.8	2.1	-2.7	-106.6		0.0172	0.0252	0.0504		93.9	94.5	13.4		
	BPL	Independent	-14.8	2.8	-0.8	-46.8	0.0247	0.0176	0.0243	0.0136	5.6	93.8	94.7	66.6		
		Exchangeable	-2.0	0.1	-1.0	-62.0	0.0056	0.0166	0.0242	0.0202	66.7	94.6	94.3	51.1		
	ML	Independent	-31.7	5.6	-21.9	-249.2	0.0023	0.0169	0.0244	0.0199	96.0	94.4	94.4	92.3		
		Exchangeable	0.2	0.3	-2.2	-39.0	0.0198	0.0170	0.0245	0.0248	22.0	93.0	93.1	86.3		
N= 500																
	WGEE	Independent		0.8	-0.3	5.3	0.0064	0.0412	0.0618	0.0358	90.9	94.9	94.0	92.1		
		Exchangeable	-31.4	-1.2	-4.6	-56.5	0.0184	0.0241	0.0392	0.0192	33.3	95.3	93.9	71.4		
	GEE	Independent	-5.9	-1.9	-4.5	-57.7	0.0035	0.0237	0.0391	0.0196	91.0	95.8	93.7	68.9		
		Exchangeable	-37.8	2.1	-2.7	-106.6		0.0172	0.0252	0.0504		93.9	94.5	13.4		
	BPL	Independent	-14.8	2.8	-0.8	-46.8	0.0247	0.0176	0.0243	0.0136	5.6	93.8	94.7	66.6		
		Exchangeable	-2.0	0.1	-1.0	-62.0	0.0056	0.0166	0.0242	0.0202	66.7	94.6	94.3	51.1		
	ML	Independent	-31.7	5.6	-21.9	-249.2	0.0023	0.0169	0.0244	0.0199	96.0	94.4	94.4	92.3		
		Exchangeable	0.2	0.3	-2.2	-39.0	0.0198	0.0170	0.0245	0.0248	22.0	93.0	93.1	86.3		
N= 500																
	WGEE	Independent		0.8	-0.3	5.3	0.0064	0.0412	0.0618	0.0358	90.9	94.9	94.0	92.1		
		Exchangeable	-31.4	-1.2	-4.6	-56.5	0.0184	0.0241	0.0392	0.0192	33.3	95.3	93.9	71.4		
	GEE	Independent	-5.9	-1.9	-4.5	-57.7	0.0035	0.0237	0.0391	0.0196	91.0	95.8	93.7	68.9		
		Exchangeable	-37.8	2.1	-2.7	-106.6		0.0172	0.0252	0.0504		93.9	94.5	13.4		
	BPL	Independent	-14.8	2.8	-0.8	-46.8	0.0247	0.0176	0.0243	0.0136	5.6	93.8	94.7	66.6		
		Exchangeable	-2.0	0.1	-1.0	-62.0	0.0056	0.0166	0.0242	0.0202	66.7	94.6	94.3	51.1		
	ML	Independent	-31.7	5.6	-21.9	-249.2	0.0023	0.0169	0.0244	0.0199	96.0	94.4	94.4	92.3		
		Exchangeable	0.2	0.3	-2.2	-39.0	0.0198	0.0170	0.0245	0.0248	22.0	93.0	93.1	86.3		
N= 500																
	WGEE	Independent		0.8	-0.3	5.3	0.0064	0.0412	0.0618	0.0358	90.9	94.9	94.0	92.1		
		Exchangeable	-31.4	-1.2	-4.6	-56.5	0.0184	0.0241	0.0392	0.0192	33.3	95.3	93.9	71.4		
	GEE	Independent	-5.9	-1.9	-4.5	-57.7	0.0035	0.0237	0.0391	0.0196	91.0	95.8	93.7	68.9		
		Exchangeable	-37.8	2.1	-2.7	-106.6		0.0172	0.0252	0.0504		93.9	94.5	13.4		
	BPL	Independent	-14.8	2.8	-0.8	-46.8	0.0247	0.0176	0.0243	0.0136	5.6	93.8	94.7	66.6		
		Exchangeable	-2.0	0.1	-1.0	-62.0	0.0056	0.0166	0.0242	0.0202	66.7	94.6	94.3	51.1		
	ML	Independent	-31.7	5.6	-21.9	-249.2	0.0023	0.0169	0.0244	0.0199	96.0	94.4	94.4	92.3		
		Exchangeable	0.2	0.3	-2.2	-39.0	0.0198	0.0170	0.0245	0.0248	22.0	93.0	93.1			

Table 5.53: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **AR(1) correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 1, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		-0.0	-5.7	-167.2		0.0505	0.0846	0.1444		95.6	93.5	51.0		
		Exchangeable	-50.3	-0.2	-5.7	-169.2	0.0181	0.0494	0.0815	0.1470	69.5	95.9	93.6	51.0		
	GEE	Independent	-36.8	-0.5	-5.7	-169.4	0.0154	0.0495	0.0812	0.1471	82.4	95.8	93.8	49.7		
		Exchangeable	-55.9	2.9	-1.3	-183.5		0.0394	0.0565	0.1622		96.1	95.3	34.9		
	BPL	Independent	-42.0	3.1	-0.8	-160.3	0.0178	0.0397	0.0564	0.1310	57.7	95.7	95.3	44.4		
		Exchangeable	-7.4	2.4	-0.8	-164.0	0.0135	0.0394	0.0563	0.1354	74.0	95.8	95.1	41.8		
	ML	Independent	-32.6	0.8	1.5	-3.8	0.0099	0.0406	0.0579	0.0985	90.9	94.9	94.9	92.0		
		Exchangeable	-7.5	2.6	-1.3	-167.9		0.0394	0.0604	0.1941		94.7	93.3	72.4		
	WGEE	Independent	-32.6	1.4	1.1	-36.9	0.0132	0.0412	0.0587	0.1010	85.8	94.8	94.7	90.1		
		Exchangeable	-7.5	0.8	1.3	-2.7	0.0099	0.0405	0.0579	0.0972	90.5	94.8	94.9	92.0		
	GEE	Independent	-48.2	-0.2	-6.5	-160.9		0.0194	0.0336	0.1152		95.7	93.4	13.9		
		Exchangeable	-34.1	-0.3	-6.5	-162.4	0.0122	0.0191	0.0328	0.1170	56.5	95.9	93.0	12.9		
	BPL	Independent	-55.1	-0.7	-6.5	-163.1	0.0082	0.0191	0.0327	0.1179	77.6	95.5	93.1	12.8		
		Exchangeable	-40.9	2.6	-2.0	-180.5		0.0168	0.0238	0.1397		94.5	94.2	3.1		
	ML	Independent	-55.1	2.8	-1.5	-156.7	0.0141	0.0170	0.0237	0.1081	26.9	94.1	94.4	8.6		
		Exchangeable	-2.5	2.0	-1.5	-161.0	0.0090	0.0167	0.0238	0.1132	54.8	94.1	94.4	7.0		
	WGEE	Independent	-28.4	0.4	0.8	-1.5	0.0035	0.0166	0.0239	0.0378	94.5	93.8	94.3	93.2		
		Exchangeable	-2.7	2.3	-1.5	-161.2		0.0166	0.0258	0.1395		93.9	91.5	52.8		
	GEE	Independent	-2.7	1.2	0.8	-30.5	0.0063	0.0167	0.0241	0.0370	82.5	93.9	94.3	93.5		
		Exchangeable	-2.7	0.3	0.7	-1.4	0.0035	0.0167	0.0240	0.0373	94.5	93.9	94.4	93.3		

Table 5.54: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **AR(1) correlation** ($\alpha = 0.2$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 1, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability					
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T		
N= 200																
	WGEE	Independent		4.6	-6.7	-120.8		0.0647	0.1048	0.0796		93.5	91.2	63.4		
		Exchangeable	-61.9	4.4	-6.8	-122.0	0.0222	0.0622	0.0995	0.0807	52.9	93.8	91.4	62.8		
	GEE	AR1	-39.3	3.8	-6.8	-122.9	0.0154	0.0617	0.0996	0.0812	76.9	93.6	91.5	61.8		
		Independent		6.7	-2.5	-134.3		0.0423	0.0546	0.0860		94.5	93.3	36.3		
	Exchangeable		-63.1	6.9	-1.9	-115.2	0.0200	0.0430	0.0549	0.0682	39.8	93.9	93.9	48.5		
		AR1	-41.9	5.5	-1.9	-120.8	0.0123	0.0413	0.0548	0.0725	69.8	94.3	93.5	46.3		
	BPL	Exchangeable/AR1	-6.1	1.8	0.9	-0.8	0.0075	0.0404	0.0567	0.0424	93.7	94.8	93.5	92.4		
	ML	Independent		5.5	-1.5	-105.0		0.0412	0.0578	0.0811		93.4	91.2	74.9		
Exchangeable		-37.8	3.3	0.5	-17.1	0.0137	0.0411	0.0575	0.0513	76.4	94.6	92.7	90.9			
	AR1		-5.9	1.7	0.7	-1.6	0.0077	0.0403	0.0561	0.0414	91.8	94.9	93.4	92.2		
	WGEE	Independent		2.4	-8.8	-116.5		0.0261	0.0457	0.0627		94.8	91.6	24.3		
		Exchangeable	-58.1	2.2	-8.7	-117.9	0.0161	0.0255	0.0435	0.0640	37.1	94.7	91.5	24.0		
	AR1		-34.4	1.5	-8.7	-119.1	0.0083	0.0250	0.0434	0.0649	74.4	95.1	91.4	22.0		
		Independent		5.2	-3.0	-134.0		0.0180	0.0229	0.0778		93.3	94.3	4.8		
	Exchangeable		-61.7	5.5	-2.2	-114.5	0.0167	0.0184	0.0228	0.0587	10.6	93.0	94.9	13.2		
		AR1	-40.0	4.0	-2.3	-120.8	0.0083	0.0170	0.0227	0.0643	51.4	93.3	94.7	9.9		
	BPL	Exchangeable/AR1	-1.4	0.1	0.5	2.3	0.0029	0.0161	0.0226	0.0160	94.9	94.6	94.7	94.2		
		ML		4.0	-1.5	-101.0		0.0168	0.0233	0.0548		93.0	92.1	55.1		
	Exchangeable		-35.6	1.9	0.3	-13.8	0.0078	0.0162	0.0228	0.0185	67.7	94.3	94.7	91.3		
		AR1	-1.4	0.0	0.3	2.2	0.0028	0.0161	0.0225	0.0151	94.7	94.1	94.7	94.1		
N= 500																

Table 5.55: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **3 responses** ($T = 3$), **AR(1) correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 1, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T			
N= 200																	
	WGEE	Independent		-1.5	-7.4	-165.1		0.0633	0.1132	0.1437		95.1	93.5	53.7			
		Exchangeable	-31.2	-1.7	-7.5	-167.2	0.0264	0.0593	0.1025	0.1455	70.8	95.0	93.6	52.0			
		AR1	-17.7	-2.4	-7.5	-168.5	0.0167	0.0587	0.1014	0.1473	83.8	95.5	93.5	51.4			
	GEE	Independent		1.9	-2.7	-237.1		0.0410	0.0641	0.2510		96.0	96.0	13.5			
		Exchangeable	-40.5	2.6	-1.2	-161.3	0.0332	0.0416	0.0637	0.1299	37.2	96.0	96.0	39.8			
		AR1	-27.8	1.0	-1.3	-173.5	0.0201	0.0406	0.0633	0.1457	62.2	96.1	95.9	34.6			
	BPL	Exchangeable/AR1	-9.0	0.3	0.3	-18.4	0.0111	0.0410	0.0639	0.1383	94.0	96.3	95.9	88.5			
		Independent		1.0	-13.7	-397.4		0.0389	0.0842	0.7097		95.0	86.9	17.9			
	ML	Exchangeable	-26.3	0.3	-0.4	-61.2	0.0238	0.0448	0.0637	0.1425	78.0	95.1	95.2	84.8			
		AR1	-7.6	0.6	0.3	-16.4	0.0105	0.0404	0.0612	0.1191	92.1	96.5	96.2	89.5			
	N= 500																
		WGEE	Independent		-1.7	-7.0	-162.4		0.0249	0.0467	0.1183		96.4	93.6	14.4		
Exchangeable			-29.3	-1.9	-7.1	-164.9	0.0178	0.0233	0.0424	0.1212	54.7	96.4	93.3	12.6			
AR1			-15.5	-2.6	-7.1	-166.5	0.0082	0.0230	0.0418	0.1234	84.2	96.2	93.3	12.7			
GEE		Independent		2.3	-2.4	-235.9		0.0170	0.0262	0.2325		96.1	94.4	0.1			
		Exchangeable	-40.3	2.9	-1.0	-160.2	0.0285	0.0173	0.0254	0.1123	8.3	96.1	94.9	6.6			
		AR1	-27.3	1.2	-1.2	-172.9	0.0148	0.0164	0.0251	0.1291	37.0	96.3	94.8	3.8			
	BPL	Exchangeable/AR1	-3.0	0.6	0.6	-10.1	0.0036	0.0164	0.0255	0.0542	95.2	96.3	94.8	92.1			
		Independent		1.1	-15.8	-426.5		0.0158	0.0534	0.7560		94.7	74.4	1.9			
	ML	Exchangeable	-23.4	1.1	0.5	-58.4	0.0128	0.0169	0.0256	0.0524	67.3	96.2	94.6	90.8			
		AR1	-2.6	0.6	0.5	-7.6	0.0033	0.0162	0.0251	0.0439	93.9	95.8	94.9	92.6			

Table 5.56: Percentage relative bias, mean squared error and coverage probability of GEE, wGEE, BPL and ML with different ‘working’ correlations for **4 responses** ($T = 4$), **AR(1) correlation** ($\alpha = 0.4$) and **NI drop-outs** ($\gamma_0, \gamma_P, \gamma_C$) = $(-1, 1, 1)$ with regression parameters $(\beta_0, \beta_G, \beta_T) = (-1, 1, 0.2)$

N	Method	Working Correlation	Relative bias				MSE				Coverage Probability						
			α	β_0	β_G	β_T	α	β_0	β_G	β_T	α	β_0	β_G	β_T			
N= 200																	
	WGEE	Independent		3.3	-7.4	-131.8		0.0873	0.1605	0.0971		94.0	92.1	59.6			
		Exchangeable	-43.5	3.0	-7.5	-131.0	0.0400	0.0788	0.1369	0.0954	47.7	94.5	91.9	60.3			
	GEE	AR1	-20.0	1.6	-7.4	-134.4	0.0176	0.0749	0.1358	0.0975	77.4	94.6	92.2	57.8			
		Independent		6.2	-2.4	-184.5		0.0443	0.0665	0.1525		94.9	93.7	15.0			
	Exchangeable		-49.4	6.8	-0.6	-121.9	0.0442	0.0451	0.0666	0.0758	15.8	94.8	93.9	41.6			
		AR1	-29.5	3.7	-0.9	-139.4	0.0198	0.0418	0.0655	0.0934	54.5	95.7	94.1	31.9			
	BPL	Exchangeable/AR1	-7.0	1.9	1.4	-14.9	0.0075	0.0426	0.0678	0.0647	96.6	95.7	94.1	89.6			
	ML	Independent		8.0	-12.0	-278.8		0.0448	0.0856	0.3566		93.4	84.7	18.8			
	Exchangeable		-31.0	2.0	-1.0	-35.3	0.0271	0.0451	0.0671	0.0725	64.7	93.5	93.0	83.0			
		AR1	-1.0	1.1	0.1	0.0	0.0076	0.0406	0.0629	0.0456	90.9	95.7	93.9	91.3			
		WGEE	Independent		0.2	-8.8	-124.9		0.0365	0.0685	0.0715		95.1	92.2	26.2		
			Exchangeable	-39.0	-0.2	-8.9	-126.9	0.0288	0.0333	0.0590	0.0735	36.6	95.3	92.4	24.7		
AR1		-14.9	-1.8	-8.8	-129.9	0.0086	0.0321	0.0584	0.0760	79.0	95.5	92.5	19.4				
		Independent		5.0	-3.1	-182.5		0.0204	0.0261	0.1388		92.7	94.2	0.4			
	GEE	Exchangeable	-48.6	5.5	-1.5	-120.2	0.0402	0.0211	0.0255	0.0637	1.2	92.4	94.2	8.9			
		AR1	-28.4	2.2	-1.8	-138.6	0.0157	0.0184	0.0253	0.0823	27.9	94.1	94.2	2.8			
	BPL	Exchangeable/AR1	-2.2	0.3	0.8	-6.9	0.0028	0.0181	0.0259	0.0230	95.0	94.5	94.6	93.0			
	ML	Independent		7.2	-13.7	-284.7		0.0219	0.0483	0.3450		87.8	76.2	1.7			
	Exchangeable		-28.5	0.0	-0.9	-26.4	0.0167	0.0185	0.0248	0.0228	32.9	92.8	94.2	87.7			
		AR1	0.3	-0.4	-0.1	2.2	0.0023	0.0178	0.0247	0.0152	92.2	93.6	93.8	94.0			
	N= 500																

5.3 Discussion

In our study, we found that the relative biases for the estimators of the group effect β_G were relatively small, even with the misspecification of the drop-out mechanisms or the association between responses. Such results were not unexpected because the drop-out mechanism was assumed to be independent of the group $\mathbf{x}$. Therefore, the focus of our discussion is on the time coefficient β_T .

Under MCAR, it was expected to obtain asymptotically unbiased estimates for GEE and wGEE as the missing-data mechanism was correctly specified and the effect of the ‘working’ correlation becomes negligible when the sample size is large. The GEE and wGEE were also highly efficient, especially for $\hat{\beta}_T$ requiring a relatively small sample size. In our empirical study, a sample size of $N = 200$ was sufficient to obtain near 0% relative bias and close to 0.95 for coverage probabilities.

For the GEE, the misspecification occurs when the drop-out mechanism is MAR or NI, under which the estimates were expected to be biased. The GEE with independence assumption was highly sensitive to the misspecification of the missing-data mechanism, especially when the correlations between responses were strong and the weight on the response at the time of drop-out for the missing-data mechanism was large. Under MAR, the GEE with exchangeable or $AR(1)$ correlation assumption was quite robust to the misspecification but not under NI, in which case the method became unreliable.

For the wGEE, the drop-out mechanism was correctly specified under MAR and it was expected to have asymptotically unbiased estimates. A sample size of $N = 200$ was sufficient to obtain near 0% relative bias. Under NI drop-out, the drop-out model was misspecified for wGEE and the estimates were expected to be asymptotically biased. Unlike the GEE, the differences in the relative biases for different ‘working’ correlations were negligible. In other words, ‘working’ correlation did not have much impact on the relative bias of the wGEE estimator.

For both the GEE and wGEE, the strength in the association between re-

sponses had a negligible impact on the relative bias of the regression estimators. In addition, an increase in the number of responses decreased the relative bias moderately. As the robust sandwich estimator was used to estimate the variance of $\hat{\beta}_T$, the coverage probability was good as compared to the ML method, even with the model misspecification.

For the ML method, the drop-out mechanism was correctly specified for all three types of drop-outs. However, the ML method was highly sensitive to the misspecification of the correlation structure. When the association between responses was misspecified, the relative bias and MSE of the ML estimator became large. In particular, the ML with independence assumption led to relative bias much higher than those obtained from other methods, making the method inappropriate under a misspecified correlation structure. Additionally, MSEs were the largest for the ML method with independence assumption and the coverage probabilities were the lowest, because unlike other methods, the variance of $\hat{\beta}_T$ was not estimated using the sandwich estimator. Also, the relative bias was considerably affected by the strength in the association between responses and by the parameters in the drop-out mechanism. When the association was correctly specified, the ML method required a large sample of size $N = 500$ to obtain a relative bias close to 0% and coverage probability close to 0.95, whereas with a smaller sample size $N = 200$, some small bias was present and the coverage probability was close to 0.90.

For bivariate maximum pseudo-likelihood method, the asymptotic relative bias was small only when $\gamma_P = 0$ for exchangeable correlation and requires a large sample size to obtain a relative bias close to 0%. In our case, $N = 500$ was required to obtain a relative bias near 0%. When $\gamma_P \neq 0$, the asymptotic relative bias became noticeably large and varied positively with the strength in the association between responses and the weights for γ_P in the drop-out mechanism. With $AR(1)$ correlation, the bivariate maximum pseudo-likelihood behaved much like the ML with the correct specification of the correlation. Our bivariate maximum

pseudo-likelihood method is appropriate only for cases in which the correlation between two time points decreases as the distance between two time points increases. $AR(1)$ is a good example of such a correlation structure. When the correlation between distant time points increases or stays constant such as the exchangeable correlation, bivariate maximum pseudo-likelihood estimator is unbiased provided the drop-out mechanism does not depend on the responses immediately preceding the drop-outs *i.e.* $\gamma_P = 0$.

Chapter 6

Conclusion and Future Study

The problem of missing data in longitudinal studies is common. In particular, for longitudinal data, missingness is often monotone resulting from participants' drop-outs. We have reviewed non-likelihood methods, such as the GEE and wGEE, and likelihood methods, such as bivariate maximum pseudo-likelihood and ML. We have conducted a simulation study in which the estimation methods were studied under several drop-out mechanisms including MCAR, MAR and NI drop-outs. We assumed that the drop-out mechanism was independent of the covariate x_i , which is a binary variable indicating a group membership. As a result, for all the reviewed methods, the bias for the group effect was negligible. However, such assumption might not always be true in practical situations. For example, in case-control studies, the missing-data pattern for the group with a disease might be different from the group without the disease. In practice, it is recommended to verify that the missingness in the two groups do not differ considerably. Otherwise, the drop-out mechanism must integrate the covariate.

Although the bias for the group effect was negligible, the bias for the time effect changed considerably from one estimation method to another. Under MCAR, estimates obtained by both GEE and wGEE were asymptotically unbiased. However, under MAR, the GEE estimates were moderately biased, whereas, the wGEE estimates were unbiased. Under NI, both methods gave comparable relative bias,

which were, in some cases, too large making the estimation methods unsuitable in practice. For the bivariate maximum pseudo-likelihood, the bias was negligible for $AR(1)$ correlation. The method becomes unreliable when the correlation between two time points stays constant or increases with the distance between the time points, such as in exchangeable correlation. The bias for estimating the time effect can be high when the drop-out mechanism is a function of the responses immediately preceding the drop-out and those at the time of drop-out. When the correlation between distant time points are weak, the bivariate maximum pseudo-likelihood behaves much like the ML with a correct specification of the association between responses. Finally, although ML estimator is unbiased under correctly specified models, it was highly sensitive to the misspecification of the association between responses.

To conclude, for NI drop-outs in longitudinal data, when the correlation between distant time points are weak, we recommend the bivariate maximum pseudo-likelihood, as the GEE and wGEE estimators are biased and the ML estimators are highly sensitive to model misspecification.

For our future research, consider the longitudinal clinical trial of contracepting women, in which different dosage amounts of depot-medroxyprogesterone acetate (DMPA) were administered to contracepting women and amenorrhea, the absence of menstrual bleeding for some number of days, was self-recorded each month (Machin et al., 1988). The clinical trial was affected by drop-outs and it is likely that the drop-outs depended on the missing response at the time of the drop-out. And one cannot exclude the possibility that the responses within the participants were strongly correlated. Such settings constitute a challenge for the estimation of the effect of DMPA on the amenorrhea. We would like to investigate novel methods based on the pseudo-likelihood approach, which are generally more efficient than the existing methods for analysing incomplete longitudinal data, and are not much sensitive to misspecified missing data models.

Bibliography

- [1] Bahadur R. T. (1961). A Representation of the Joint Distribution of Response to n Dichotomous items. *Studies in Item Analysis and Prediction*, Stanford University Press, 158-168.
- [2] Bishop, S.H. and Jones, B. (1984). A review of higher-order cross-over designs. *Journal of Applied Statistics*, **11**, 29-50.
- [3] Byrd, R. H., Lu, P., Nocedal, J. and Zhu, C. (1995) A limited memory algorithm for bound constrained optimization. *SIAM J. Scientific Computing*, **16**, 1190–1208.
- [4] Casella, G. and Berger, R.L. (1990). *Statistical Inference*. Wardsworth and Brooks/Cole, Pacific Grove
- [5] Cox D. R. (1975). Partial likelihood. *Biometrika*, **62**, 269-276.
- [6] Diggle, P. J., Heagerty P., Linang K. Y. and Zeger, S. L. (2002). *Analysis of Longitudinal Data*. Oxford University Press Inc., New York.
- [7] Fitzmaurice, G. M., Molenberghs, G., and Lipsitz, S. R. (1995). Regression Models for Longitudinal Binary Responses With Informative Drop-Outs. *Journal of the Royal Statistical Society, Ser.B*, **57**, 691-704.
- [8] Gong, G. and Samaniego, F. (1981). Pseudo maximum likelihood estimation: Theory and applications. *Annals of Statistics*, **9**, 861–869.

-
- [9] Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions, *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 221–233.
- [10] Liang, K. Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, **73**, 13-22.
- [11] Liang, K. Y., Zeger, S. L. and Qaqish, B. (1992). Multivariate regression analyses for categorical data. *Journal of the Royal Statistical Society, B*, **54**, 3-40.
- [12] Little, R. J. A. and Rubin, D. B. (2002). *Statistical Analysis with Missing Data*. Wiley, New York.
- [13] Machin, D., Farley T., Busca, B., Campbell, M. and d’Arcangues, C. (1988). Assessing changes in vaginal bleeding patterns in contracepting women. *Contraception*, **38**, 165-179.
- [14] McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models (2nd edition)*. Chapman and Hall/CRC, Boca Raton.
- [15] McCulloch, C. E., Searle, S. R. and Neuhaus, J. M. (2008). *Generalized, Linear and Mixed Models (2nd edition)*. Wiley & Sons, Inc, New Jersey.
- [16] Molenberghs, G. and Lesaffre, E. (1994). Marginal modeling of correlated ordinal data using a multivariate Plackett distribution. *Journal of the American Statistical Association*, **89**, 633-44.
- [17] Nelder, J. A. and Mead, R. (1965) A simplex algorithm for function minimization. *Computer Journal*, **7**, 308–313.
- [18] Nelder, J. A. and Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society, A*, **135**, 370–384.
- [19] Prentice, R. L. (1988). Correlated Binary Regression With Covariates Specific to Each Binary Observation. *Biometrics*, **44**, 1033-1048.

-
- [20] Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1995). Analysis of Semiparametric Regression Models for Repeated Outcomes in the Presence of Missing Data. *Journal of the American Statistical Association*, **90**, 106-121.
- [21] Rubin, D. B. (1976). Inference and Missing Data. *Biometrika*, **63**, 581-592.
- [22] Sinha, S. K., Troxel, A. B., Lipsitz, S. R., Sinha, D., Fitzmaurice, G. M., Molenberghs, G. and Ibrahim, J. G. (2011). A bivariate pseudo-likelihood for incomplete longitudinal binary data with nonignorable non-monotone missingness. *Biometrics*, **67**, 1119-1126.
- [23] Troxel, A. B., Lipsitz, S. R., and Harrington, D. P. (1998). Marginal models for the analysis of longitudinal measurements subject to nonignorable non-monotone missing data. *Biometrika*, **85**, 661– 672.
- [24] Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models and the Gaussian method. *Biometrika*, **61**, 439-47.
- [25] White, H. (1982). Maximum Likelihood Estimation of Misspecified Models. *Econometrica*, **50**, 1-26.