

REPUBLIC OF TURKEY
HACETTEPE UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

AN ADAPTED WALD TEST STATISTIC TO
DETERMINE THE VARIABLES WHICH DO NOT
SATISFY THE PROPORTIONALITY ASSUMPTION IN
THE ADJACENT CATEGORY LOGISTIC
REGRESSION MODEL

Nimet Anıl DOLGUN

Program of Biostatistics
DOCTOR OF PHILOSOPHY THESIS

ANKARA
2012

REPUBLIC OF TURKEY
HACETTEPE UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

AN ADAPTED WALD TEST STATISTIC TO
DETERMINE THE VARIABLES WHICH DO NOT
SATISFY THE PROPORTIONALITY ASSUMPTION IN
THE ADJACENT CATEGORY LOGISTIC
REGRESSION MODEL

Nimet Anıl DOLGUN

Program of Biostatistics
DOCTOR OF PHILOSOPHY THESIS

THESIS SUPERVISOR
Prof. Osman Saraçbaşı

ANKARA
2012

Department: Department of Biostatistics

Program: Biostatistics

Thesis Title: An Adapted Wald Test Statistic to Determine the Variables Which Do Not Satisfy the Proportionality Assumption in the Adjacent Category Logistic Regression Model

Student's Name - Surname: Nimet Anıl Dolgun

Date of Thesis Defense: 23.01.2012

This study has been accepted and approved as a Ph.D. thesis by the examining committee whose members listed below.

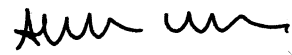
Chairperson of the jury: Prof. Ergun Karaağaoğlu
Hacettepe University



Supervisor: Prof. Osman Saraçbaşı
Hacettepe University



Member: Prof. Atilla H. Elhan
Ankara University



Member: Assoc. Prof. Mehtap Akçıl
Başkent University



Member: Assist. Prof. S. Kenan Köse
Ankara University



I hereby certify that this thesis has been accepted and approved by the committee above in conformity to the regulations and bylaws of the Hacettepe University Institute of Health Sciences.



Prof. Hakan S. Orer, MD, PhD

Director

ACKNOWLEDGEMENTS

I am heartily thankful to my supervisor, Professor Osman Saraçbaşı, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject. I am also grateful to the members of the thesis examining committee, whose suggestions and questions improved the clarity and quality of my thesis.

Lastly, I offer my thanks and regards to my family, colleagues in my department, and all of those who supported me in any respect during the completion of my thesis.

This thesis is supported by H.Ü.B.A.B, Ph. D. Thesis Grant (09.T01.101.001).

ABSTRACT

Dolgun, N. A. An Adapted Wald Test Statistic to Determine the Variables Which Do Not Satisfy the Proportionality Assumption in the Adjacent Category Logistic Regression Model. Hacettepe University Institute of Health Sciences, Ph. D. Thesis in Biostatistics, Ankara, 2012. Ordinal logistic regression models are classified as either proportional odds models, continuation ratio models or adjacent category models. The common model assumption of these models is that the log odds do not depend on the outcome category. This assumption is also known as the “proportionality” or “parallel logits” assumption. Non-proportional and partial proportional models are proposed for the proportional odds and continuation ratio model. The non-proportional and the partial proportional versions of the adjacent category model are also feasible. Prior to fitting any of ordinal logistic regression models, it is important to check whether the assumption of proportionality is satisfied by each independent variable. In the proportional odds model, the proportional odds assumption is checked by Brant’s Wald test statistic, and the standard Wald test statistic can be used in the continuation ratio model. However there is no valid approach to test whether the proportionality assumption is satisfied by each independent variable in the adjacent category model. The aim of the study is to determine the variables in the adjacent category model that violate the proportionality assumption. For this purpose, a Wald test is proposed for testing the proportionality assumption in the adjacent category model. Using the proposed test statistic, the variables in the adjacent category model that do not obey the proportionality assumption can be determined, and they can be modelled as non-proportional in the partial proportional adjacent category model.

Key Words: Ordinal logistic regression, odds ratio, proportionality assumption, Wald test statistics, likelihood ratio test statistics.

Supported by H.Ü.B.A.B, Ph. D. Thesis Grant (09.T01.101.001).

ÖZET

Dolgun, N. A. Ardışık Kategori Lojistik Regresyon Modelinde Orantısallığı Bozan Değişkenlerin Belirlenmesi İçin Uyarlanmış Wald Test İstatistiği. Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Programı Doktora Tezi, Ankara, 2012. Ordinal lojistik regresyon modelleri, orantısal odds modeli, sürekli oran modeli ve ardışık kategori lojistik regresyon modeli olarak üçe ayrılmaktadır. Bu modellerin tümünde, modelde oluşturulan logit'lerin "paralellliği" diğer adıyla "orantısallık" varsayımı bulunmaktadır. Orantısal odds modeli ile sürekli oran lojistik regresyon modelinde orantısallık varsayımının sağlanmadığı durumlar için orantısal olmayan ve kısmi orantısal modeller geliştirilmiştir. Ayrıca ardışık kategori lojistik regresyon modeli için de orantısal olmayan ve kısmi orantısal modeller uygulanabilmektedir. Herhangi bir ordinal lojistik regresyon modelini kurmadan önce, orantısallık varsayımının modeldeki değişkenler bazında sağlanıp sağlanmadığının kontrolünün yapılması oldukça önemli bir noktadır. Orantısal odds modelinde bu varsayımın testi Brant'ın Wald test istatistiği ile, sürekli oran modelinde ise standart Wald test istatistiği ile yapılabilmektedir. Ancak ardışık kategori lojistik regresyon modelinde bu varsayımın değişken bazında kontrolü için herhangi geçerli bir test istatistiği bulunmamaktadır. Bu nedenden ötürü, bu çalışmanın amacı, ardışık kategori lojistik regresyon modelinde orantısallık varsayımını bozan değişkenlerin saptanmasıdır. Bu amaçla, ardışık kategori lojistik regresyon modelinde orantısallığı değişken bazında test edebilecek bir Wald test istatistiği geliştirilmiştir. Önerilen test istatistiği ile, ardışık kategori lojistik regresyon modelinde orantısallık değişken bazında test edilip belirlenecek ve orantısallığı sağlamayan değişkenler kısmi orantısal ardışık kategori lojistik regresyon modeli kullanılarak modellenebilecektir.

Anahtar Kelimeler: Ordinal lojistik regresyon, odds oranı, orantısallık varsayımı, Wald test istatistiği, olabilirlik oranı test istatistiği.

Destekleyen Kurumlar: H.Ü.B.A.B, Tez Destekleme (09.T01.101.001).

TABLE OF CONTENTS

	Page
APPROVAL PAGE	iii
ACKNOWLEDGEMENTS	iv
ABSTRACT	v
ÖZET	vi
TABLE OF CONTENTS	vii
LIST OF ABBREVIATIONS	ix
LIST OF FIGURES	x
LIST OF TABLES	xi
1 INTRODUCTION	1
2 REVIEW OF LITERATURE	5
2.1 The Multinomial Logistic Regression Model	5
2.2 Ordinal Logistic Regression Models	7
2.2.1 The Proportional Odds Model	8
2.2.2 The Non-Proportional Odds Model	11
2.2.3 The Partial Proportional Odds Model	11
2.2.4 The Continuation Ratio Model	12
2.2.5 The Non-Proportional Continuation Ratio Model	14
2.2.6 The Partial Proportional Continuation Ratio Model	14
2.2.7 The Adjacent Category Model	14
2.2.8 The Non-Proportional Adjacent Category Model	17
2.3 Methods Used for Checking the Proportional Odds Assumption .	18
2.3.1 The Likelihood Ratio Test	18
2.3.2 Brant's Wald Test	19
3 MATERIAL AND METHODS	22
3.1 The Proposed Wald Test Statistic: Adaptation of Brant's Test for Adjacent Category Model	22
3.2 The Partial Proportional Adjacent Category Model	24
3.2.1 Model Specification	24
3.2.2 Parameter Estimation	24
3.3 Simulation Study	25

3.3.1	Simulation Study 1: Simulation of the Adapted Test Statistic in 2×3 Tables	25
3.3.2	Simulation Study 2: Simulation of the Adapted Test Statistic in $R \times C$ Tables	27
4	RESULTS	31
4.1	Results of Simulation Study 1	31
4.2	Results of Simulation Study 2	33
5	AN EXAMPLE	40
5.1	Data	40
5.2	An Example of the Proposed Test of Proportionality in the Ad- jacent Category Model	41
5.3	An Example of Parameter Estimation in the Partial Proportional Adjacent Category Model	42
6	DISCUSSION AND FUTURE WORK	47
	BIBLIOGRAPHY	49
A	R CODES FOR THE SIMULATION STUDY 1	52
B	R CODES FOR THE SIMULATION STUDY 2	63
C	R CODES FOR CALCULATING PROPOSED WALD TEST AND FITTING PARTIAL PROPORTIONAL ADJACENT CATEGORY MODEL	78

LIST OF ABBREVIATIONS

ACM	Adjacent category model
$a_k(\cdot)$	Logit representation of the adjacent category model
$a_k^{NP}(\cdot)$	Logit representation of the non-proportional adjacent category model
$a_k^{PP}(\cdot)$	Logit representation of the partial proportional adjacent category model
cdf	Cumulative distribution function
$c_k(\cdot)$	Logit representation of the proportional odds model
$c_k^{NP}(\cdot)$	Logit representation of the non-proportional odds model
$c_k^{PP}(\cdot)$	Logit representation of the partial proportional odds model
CRM	Continuation ratio model
$g_k(\cdot)$	Logit representation of the multinomial logistic regression model
$L(\cdot)$	Log-likelihood
L_{ACM}	Log-likelihood of the adjacent category model
L_{NPACM}	Log-likelihood of the non-proportional adjacent category model
LR	Likelihood ratio
MLRM	Multinomial logistic regression model
NPACM	Non-proportional adjacent category model
NPCRM	Non-proportional continuation ratio model
NPOM	Non-proportional odds model
OLRM	Ordinal logistic regression models
OR	Odds ratio
pdf	Probability density function
POM	Proportional odds model
PPACM	Partial proportional adjacent category model
PPCRM	Partial proportional continuation ratio model
PPOM	Partial proportional odds model
$r_k(\cdot)$	Logit representation of the continuation ratio model
$r_k^{NP}(\cdot)$	Logit representation of the non-proportional continuation ratio model
$r_k^{PP}(\cdot)$	Logit representation of the partial proportional continuation ratio model
Q-Q plot	Quantile-quantile plot

LIST OF FIGURES

	Page
1.1 The diagram of the ordinal logistic regression models	4
2.1 Constructing baseline category logits in the multinomial logistic regression model	6
2.2 Constructing cumulative logits in the proportional odds model . .	9
2.3 Illustration of the parallel regression assumption	10
2.4 Constructing logits in the continuation ratio model	12
2.5 Constructing adjacent category logits in the adjacent category model	15
4.1 The pdf, cdf, and Q-Q plots of the (a) proposed test statistic vs. χ^2_1 distribution and (b) LR test vs. χ^2_1 distribution for the scenario number 10	32
4.2 The pdf, cdf, and Q-Q plots of the (a) omnibus version of the proposed test statistic vs. $\chi^2_{J(K-2)}$ distribution and (b) LR test vs. $\chi^2_{J(K-2)}$ distribution for the scenario number 18	35
4.3 The pdf, cdf, and Q-Q plots of the (a) omnibus version of the proposed test statistic vs. $\chi^2_{J(K-2)}$ distribution and (b) LR test vs. $\chi^2_{J(K-2)}$ distribution for the scenario number 20	36
4.4 The pdf, cdf, and Q-Q plots of the individual version of the pro- posed test statistic vs. χ^2_{K-2} distribution for the scenario number 18	38
4.5 The pdf, cdf, and Q-Q plots of the individual version of the pro- posed test statistic vs. χ^2_{K-2} distribution for the scenario number 20	39

LIST OF TABLES

	Page
3.1 The frequency and the probability notations in the adjacent category model with three outcomes and one binary independent variable.	25
3.2 Multinomial probabilities that are used for simulation study 1. . .	26
3.3 Scenarios that are used for simulation study 1.	26
3.4 Frequencies in a $R \times C$ table.	27
3.5 Mean, standard deviation of the linear-by-linear model parameters, common adjacent odds ratios, $R \times C$ values, and sample sizes that are used in the simulation study 2.	29
4.1 Results of the simulation study 1.	31
4.2 Results of the simulation study 2.	34
5.1 Data code sheet for the variables in the attitudes toward working mothers general survey data.	40
5.2 Results of the LR test and the proposed test of proportionality for the attitudes toward working mothers data.	41
5.3 Estimation results of the multinomial logistic regression model. . .	42
5.4 Estimation results of the non-proportional adjacent category model.	43
5.5 Estimation results of the constrained multinomial logistic regression model with constraint $\beta_k = k\beta$	44
5.6 Estimation results of the adjacent category model.	45
5.7 Estimation results of the partial-proportional adjacent category model.	46

1 INTRODUCTION

The logistic regression analysis is a useful way of describing the relationship between one or more independent variables and a categorical response variable. Logistic regression models are categorized according to the type of categorical response variable as follows: binary logistic regression model, multinomial logistic regression model, and ordinal logistic regression models (1). The binary logistic regression model is used to model the binary response variable, whereas the multinomial logistic regression is a simple extension of the binary logistic regression model where the response variable has more than two unordered categories.

Response variables can also be measured on an ordinal scale such as extent of disease (none, some, severe), cardiac risk levels (low, medium and high risk) and grade of a tumour (grade I-II-III). Outcome data collected using an ordinal scale are sometimes analyzed with a binary logistic regression by creating dichotomies among the level of the outcome variable (2, 3). Although such approaches are not incorrect, they often result in a loss of information due to collapsing (or ignoring) some categories of the response, typically resulting in a considerable loss of statistical power (4). Additionally, the multinomial logistic regression model is often used for modelling multiple category ordinal outcomes. This analysis, however, would not take into account the ordinal nature of the outcome (1). Therefore, ordinal logistic regression models are used to model the relationship between independent variables and an ordinal response variable when the response variable has a natural ordering.

A number of ordinal logistic regression models for analyzing ordinal outcome data have been proposed (5–14), and the application of these methods are widely used, especially in epidemiological and biomedical research (15–20). Figure 1.1 gives the diagram of the ordinal logistic regression models and describes when to use them.

Probably the most frequently used ordinal logistic regression model is the proportional odds model of McCullagh because of the simplicity of its interpretation (6). The continuation ratio model may be best suited in situations where the individual categories of the response variable are of intrinsic interest and are not merely arbitrary groupings of an underlying continuous variable (15). Adjacent category models are often used due to their close connection to log-linear models (9).

Although these ordinal logistic regression models have different structures and strategies for constructing logits, they share the same model assumption. The common model assumption of these models is that the log odds do not depend on

the outcome category. This assumption is also known as the “proportionality” or “parallel logits” assumption (1). Unfortunately, this basic and strict assumption frequently does not hold (21).

When the assumption of parallel logits is rejected for any of the ordinal logistic regression models, alternative models should be considered that do not impose the constraint of parallel logits. As seen in Figure 1.1, for every ordinal logistic regression model, there are two alternative models to use in cases of non-proportional odds. Non-proportional (i.e., relaxing the proportionality assumption for all of the variables in a model) and partial proportional models (i.e., relaxing the proportionality assumption only for some) are proposed for the proportional odds model (14, 22, 23) and the continuation ratio model (24, 25). The non-proportional and the partial proportional versions of the adjacent category model are also feasible.

Prior to fitting any of ordinal logistic regression models, it is important to check whether the assumption of proportionality is satisfied by each independent variable. In the proportional odds model, the proportional odds assumption is checked by Brant’s Wald test statistic (26), and the standard Wald test statistic (1) can be used in the continuation ratio model. **However there is no valid approach to test whether the proportionality assumption is satisfied by each independent variable in the adjacent category model.**

The aim of this study is to determine the variables in the adjacent category model that violate the proportionality assumption. For this purpose, a Wald test is proposed for testing the proportionality assumption in the adjacent category model. With use of the proposed test statistic, the variables in the adjacent category model that do not obey the proportionality assumption can be determined, and they can be modelled as non-proportional in the partial proportional adjacent category model.

The main hypotheses in this study are as follows:

i) The proposed “Wald test statistic” is asymptotically distributed as a chi-square distribution under the null hypothesis of proportionality. Additionally, this test statistic is an effective way of testing proportionality in the adjacent category model, and it is superior to the likelihood ratio test statistic because it allows both an overall and individual test of proportionality in the adjacent category model.

ii) The partial proportional adjacent category model is superior to the adjacent category model in situations where the assumption of proportionality is not fulfilled for some variables in the model. The parameter estimates and the standard errors in the partial proportional model are feasible, and they can be

obtained by fitting a constrained version of the multinomial logistic regression model using the standard maximum likelihood estimation method.

Chapter 2 presents a comprehensive literature review. Through Sections 2.1 and 2.2, the multinomial and ordinal logistic regression models are presented. Methods used for checking the proportional odds assumption are also given in Section 2.3. Chapter 3 presents the materials and methods used in the study. In Section 3.1, we adapt the Wald test of Brant in order to test the proportionality assumption in the adjacent category model. In Section 3.2, we describe how the partial proportional adjacent category model can be fitted using the maximum likelihood estimation method. In Section 3.3, we examine the asymptotic distribution of the adapted test statistic. In Chapter 4, we give the results of our simulation study. An example of the calculations of the proposed test statistic and the parameter estimations in the partial proportional adjacent category model are given in Chapter 5. Finally, in Chapter 6, we discuss our findings.

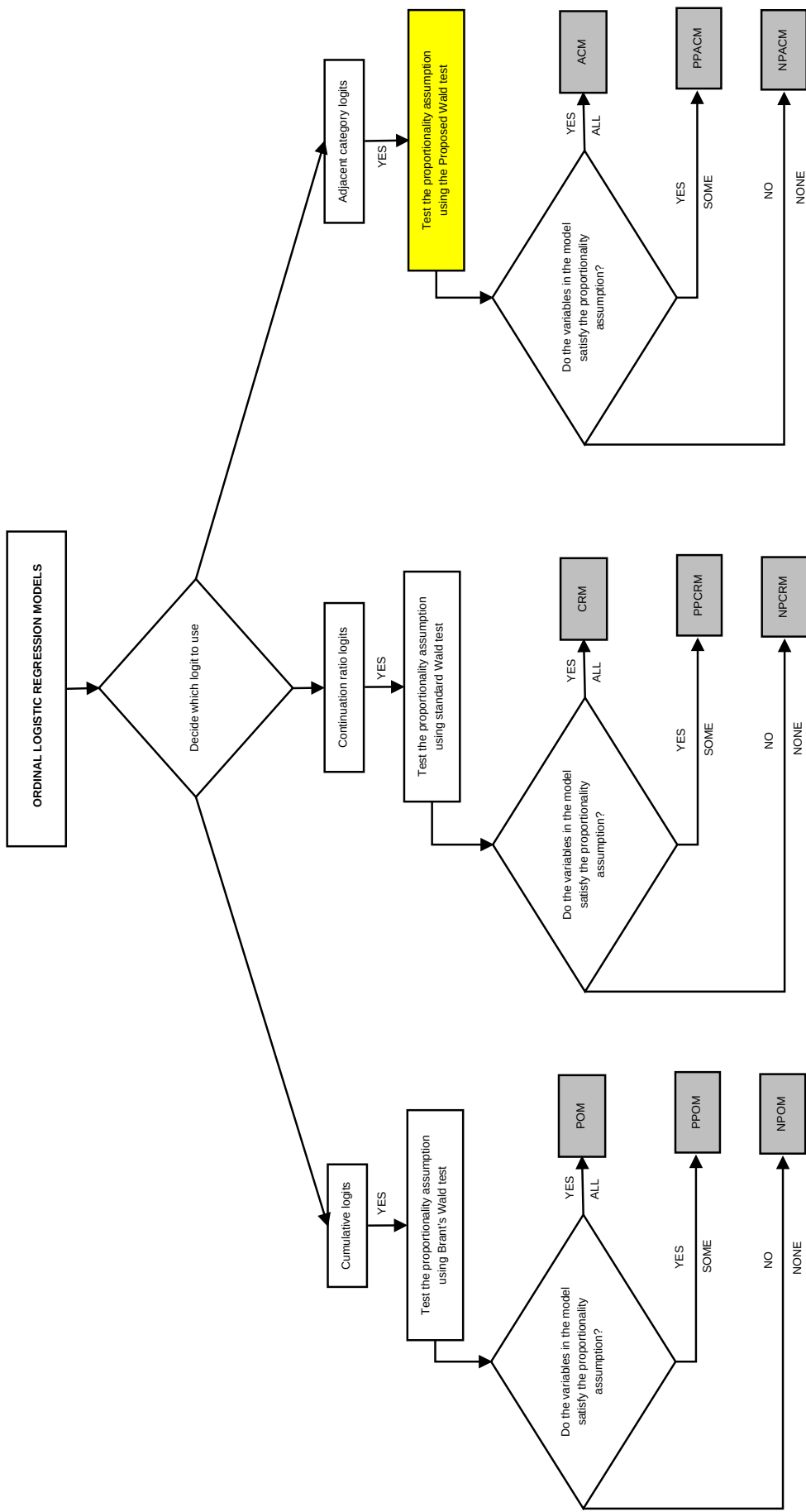


Figure 1.1: The diagram of the ordinal logistic regression models

POM: Proportional odds model, **PPOM:** Partial proportional odds model, **NPOM:** Non-proportional odds model,
CRM: Continuation ratio model; **PPCRM:** Partial proportional continuation ratio model, **NPCRM:** Non-proportional continuation ratio model,
ACM: Adjacent category model, **PPACM:** Partial proportional adjacent category model, **NPACM:** Non-proportional adjacent category model.

2 REVIEW OF LITERATURE

In this section, we undertake a review of the literature of the key models.

2.1 The Multinomial Logistic Regression Model

The multinomial logistic regression model (MLRM) is a simple extension of the binary logistic regression model in which the response variable has more than two unordered categories. This model can be fitted by simultaneously estimating binary logistic regression models for all possible comparisons of the outcome category with a baseline category. This is why MLRM is often called the “baseline category logistic regression model”.

Let Y denote the outcome variable and k denote the outcome variable category where $k = 0, 1, \dots, K$. The probability that the outcome is equal to k , conditional on a vector $\mathbf{x}$ of j covariates ($j = 1, 2, \dots, J$) is denoted by $P(Y = k|\mathbf{x}) = \phi_k(\mathbf{x})$. In the multinomial logistic regression model, generally the first category of the outcome ($k = 0$) is set to be the baseline, and the probability of the baseline, conditional on $\mathbf{x}$, is denoted by $P(Y = 0|\mathbf{x}) = \phi_0(\mathbf{x})$. Using the multinomial logit link, the MLRM can be expressed in terms of logit (log odds) as:

$$g_k(\mathbf{x}) = \ln \left[\frac{P(Y = k|\mathbf{x})}{P(Y = 0|\mathbf{x})} \right] = \beta_{0k} + \mathbf{x}'\boldsymbol{\beta}_k \quad k = 1, 2, \dots, K. \quad (2.1)$$

Additionally, the probabilities in Equation 2.1 are equal to:

$$P(Y = 0|\mathbf{x}) = \frac{1}{1 + \sum_{i=1}^K e^{\beta_{0i} + \mathbf{x}'\boldsymbol{\beta}_i}}, \quad (2.2)$$

$$P(Y = k|\mathbf{x}) = \frac{e^{\beta_{0k} + \mathbf{x}'\boldsymbol{\beta}_k}}{1 + \sum_{i=1}^K e^{\beta_{0i} + \mathbf{x}'\boldsymbol{\beta}_i}}, \quad (2.3)$$

and

$$\sum_{k=0}^K P(Y = k|\mathbf{x}) = 1. \quad (2.4)$$

In Equation 2.1, β_{0k} is the constant term, and $\boldsymbol{\beta}_k$ is the regression coefficient vector in the k th logit. The benefit of using MLRM is that it models the odds of each category relative to a baseline category as a function of covariates. As an example, let Y have four categories ($k = 0, 1, 2, 3$). Then we can construct

3 logits as given in Figure 2.1.

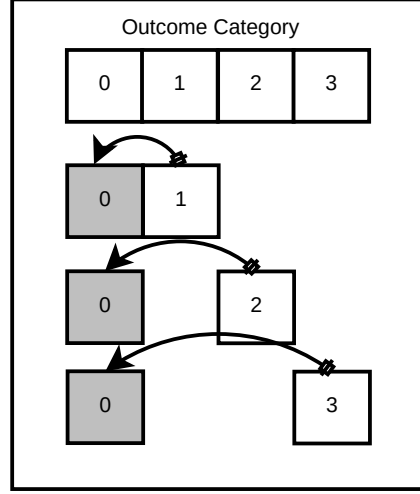


Figure 2.1: Constructing baseline category logits in the multinomial logistic regression model

In Figure 2.1, the outcome category $Y = 0$ (which is shaded with gray) is taken as the reference category, and the log-odds of other categories conditional on $\mathbf{x}$ are simultaneously modelled in the multinomial logistic regression model. The parameter and the standard error estimates, simultaneously satisfying K multinomial logits can be estimated using the maximum likelihood estimation method.

For $i = 1, 2, \dots, n$, let $\mathbf{y}_i = (y_{i0}, y_{i1}, \dots, y_{iK})$ represent the multinomial trial for subject i , where $y_{ik} = 1$ if the response is in category k and $y_{ik} = 0$ otherwise. Thus, $\sum_{k=0}^K (y_{ik}) = 1$. Let $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})'$ denote p independent variable values for subject i , and let $\boldsymbol{\beta}_k = (\beta_{k1}, \beta_{k2}, \dots, \beta_{kp})$ denote the regression coefficients for the k th logit.

Because $\phi_0(\mathbf{x}) = 1 - (\phi_1(\mathbf{x}) + \dots + \phi_K(\mathbf{x}))$ and $y_{i0} = 1 - (y_{i1} + y_{i2} + \dots + y_{iK})$, the contribution to the log-likelihood by subject i is:

$$\begin{aligned} \log \left[\prod_{k=0}^K \phi_k(\mathbf{x}_i)^{y_{ik}} \right] &= \sum_{k=1}^K y_{ik} \log \phi_k(\mathbf{x}_i) + \left(1 - \sum_{k=1}^K y_{ik} \right) \log \left[1 - \sum_{k=1}^K \phi_k(\mathbf{x}_i) \right] \\ &= \sum_{k=1}^K y_{ik} \log \frac{\phi_k(\mathbf{x}_i)}{1 - \sum_{k=1}^K \phi_k(\mathbf{x}_i)} + \log \left[1 - \sum_{k=1}^K \phi_k(\mathbf{x}_i) \right]. \end{aligned} \quad (2.5)$$

Now assume n independent observations. In the last expression above, substituting $\beta_{0k} + \mathbf{x}_i \boldsymbol{\beta}'_k$ for the logit in the first term and $\phi_0(\mathbf{x}_i) = 1/[1 +$

$\sum_{k=1}^K e^{\beta_{0i} + \mathbf{x}_i \boldsymbol{\beta}'_k}$] in the second term, the log-likelihood is

$$\begin{aligned} \mathbf{L}(\boldsymbol{\beta}) &= \log \prod_{i=1}^n \left[\prod_{k=0}^K \phi_k(\mathbf{x}_i)^{y_{ik}} \right] = \sum_{i=1}^n \left\{ \sum_{k=1}^K y_{ik} (\beta_{0k} + \mathbf{x}_i \boldsymbol{\beta}'_k) - \log \left[1 + \sum_{k=1}^K e^{\beta_{0k} + \mathbf{x}_i \boldsymbol{\beta}'_k} \right] \right\} \\ &= \sum_{k=1}^K \left[\beta_{0k} \left(\sum_{i=1}^n y_{ik} \right) + \sum_{j=1}^p \beta_{jk} \left(\sum_{i=1}^n x_{ij} y_{ik} \right) \right] \\ &\quad - \sum_{i=1}^n \log \left[1 + \sum_{k=1}^K e^{\beta_{0k} + \mathbf{x}_i \boldsymbol{\beta}'_k} \right]. \end{aligned} \tag{2.6}$$

The likelihood equations are found by taking the first partial derivatives of $\mathbf{L}(\boldsymbol{\beta})$ with respect to each of the unknown parameters. The maximum likelihood estimator, $\hat{\boldsymbol{\beta}}$, is obtained by setting these equations equal to zero and solving for $\boldsymbol{\beta}$. The solution requires an iterative computation (the Newton-Raphson method). The matrix of second partial derivatives is required to obtain the information matrix and the estimator of the covariance matrix of the maximum likelihood estimator. Further details can be found in McCullagh and Nelder's book (11, p.171–174).

MLRM is often used when the dependent variable is ordinal. If the dependent variable is ordinal and a model for nominal variables is used, there is a loss of efficiency because this information (ordinality) is ignored (21). In such cases, one should consider using ordinal logistic regression models (OLRM) instead. Additionally, before considering methods for ordinal outcomes, it is important to note that the fact that the variables can be ordered does not imply that the variable should be analyzed as ordinal. A variable might be ordered when considered for one purpose but can be unordered when used for another purpose. McCullagh and Nelder (11, p.151) make this point with the example of colours. While colours can be arranged according to the electromagnetic spectrum, this doesn't imply that this ordering is appropriate for all purposes. For example, when selecting a car, a customer doesn't prefer colours in an order that moves around the colour wheel from red, to orange, to yellow, and so on.

2.2 Ordinal Logistic Regression Models

When we move to an ordinal model, there are many alternative models. We must decide what outcomes to compare and what the most reasonable model is

for the logit. Ordinal logistic regression models have basically three members, the proportional odds model (POM), continuation ratio model (CRM), and adjacent category model (ACM). These models assume that the log odds do not depend on the outcome category. Other models are derivations of them, relaxing the proportionality assumption for some or for all of the independent variables in a model.

2.2.1 The Proportional Odds Model

Probably the most frequently used ordinal logistic regression model is the proportional odds model (POM) because of the simplicity of its interpretation (6). The POM uses cumulative logits while building the model. The cumulative logits are defined as

$$\text{logit}[Y \leq k|\mathbf{x}] = \ln \left[\frac{P(Y \leq k|\mathbf{x})}{1 - P(Y \leq k|\mathbf{x})} \right]. \quad (2.7)$$

A model that simultaneously uses all cumulative logits for all $k = 0, 1, \dots, K - 1$ is called the POM and is given as follows:

$$c_k(\mathbf{x}) = \ln \left[\frac{P(Y \leq k|\mathbf{x})}{1 - P(Y \leq k|\mathbf{x})} \right] = \tau_k - \mathbf{x}'\boldsymbol{\beta} \quad k = 0, 1, \dots, K - 1. \quad (2.8)$$

The cumulative probability in Equation 2.8 is defined as

$$P(Y \leq k|\mathbf{x}) = \frac{e^{\tau_k - \mathbf{x}'\boldsymbol{\beta}}}{1 + e^{\tau_k - \mathbf{x}'\boldsymbol{\beta}}}, \quad (2.9)$$

and the actual category probability is defined as

$$\begin{aligned} P(Y = k|\mathbf{x}) &= P(Y \leq k|\mathbf{x}) - P(Y \leq k - 1|\mathbf{x}) \\ &= \frac{e^{\tau_k - \mathbf{x}'\boldsymbol{\beta}}}{1 + e^{\tau_k - \mathbf{x}'\boldsymbol{\beta}}} - \frac{e^{\tau_{k-1} - \mathbf{x}'\boldsymbol{\beta}}}{1 + e^{\tau_{k-1} - \mathbf{x}'\boldsymbol{\beta}}}. \end{aligned} \quad (2.10)$$

As an example to how to construct the cumulative logits in POM, let Y have four ordinal categories ($k = 0, 1, 2, 3$). Then, in the POM, we can construct 3 cumulative logits as given in Figure 2.2.

Note that in the POM, each logit has its own intercept (τ_k), but the regression coefficient ($\boldsymbol{\beta}$) does not depend on k , implying that the model assumes that the relationship between x_i and Y is independent of k . McCullagh calls this assumption “the proportional odds” or “the parallel regression” assumption. The

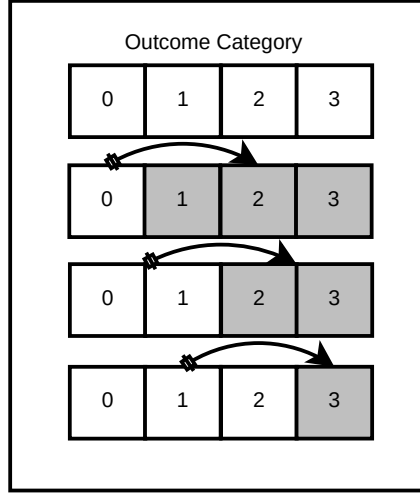


Figure 2.2: Constructing cumulative logits in the proportional odds model

idea of “parallel regression assumption” can be seen by rewriting the model in terms of the cumulative probability that an outcome is less than or equal to k . For $k = 0, 1, 2, 3$ and from Equation 2.9,

$$\begin{aligned}
 P(Y \leq 0|\mathbf{x}) &= \frac{e^{\tau_0 - \mathbf{x}'\boldsymbol{\beta}}}{1 + e^{\tau_0 - \mathbf{x}'\boldsymbol{\beta}}}, \\
 P(Y \leq 1|\mathbf{x}) &= \frac{e^{\tau_1 - \mathbf{x}'\boldsymbol{\beta}}}{1 + e^{\tau_1 - \mathbf{x}'\boldsymbol{\beta}}}, \\
 P(Y \leq 2|\mathbf{x}) &= \frac{e^{\tau_2 - \mathbf{x}'\boldsymbol{\beta}}}{1 + e^{\tau_2 - \mathbf{x}'\boldsymbol{\beta}}}.
 \end{aligned} \tag{2.11}$$

Figure 2.3 plots the three cumulative probability curves in Equation 2.11. As shown in Figure 2.3, changing the intercept shifts the cumulative probability curve to the right or to the left but does not change the slope.

When we examine the slope of these three curves at any point of the outcome probability, we find that

$$\frac{\partial P(Y \leq 0|\mathbf{x})}{\partial x} = \frac{\partial P(Y \leq 1|\mathbf{x})}{\partial x} = \frac{\partial P(Y \leq 2|\mathbf{x})}{\partial x}. \tag{2.12}$$

The POM in Equation 2.8 has some attractive features. Because the regression parameters are invariant to the outcome category, the odds ratios are the same over the K logits, and the common log odds ratio provides a single estimate of the log odds ratio over the cut-off points. This estimate is not a weighted average of the cut-off point-specific log odds ratios (27), but it is the optimum

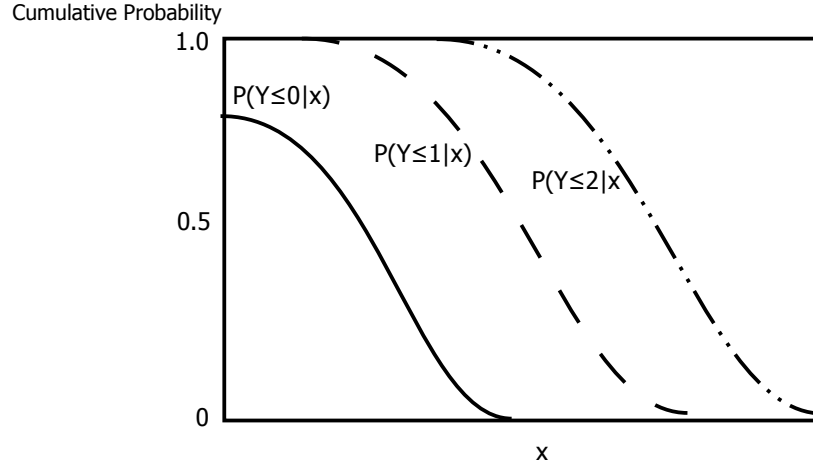


Figure 2.3: Illustration of the parallel regression assumption

estimate obtained using the maximum likelihood methods.

In the POM, cumulative logits are simultaneously modelled using the maximum likelihood estimation method. The log-likelihood of the POM is an adaptation of the multinomial log-likelihood given in Equation 2.6.

$$\begin{aligned}
 \mathbf{L}(\boldsymbol{\beta}) &= \log \prod_{i=1}^n \left[\prod_{k=0}^K \phi_k(\mathbf{x}_i)^{y_{ik}} \right] = \sum_{i=1}^n \sum_{k=0}^K y_{ik} \log \{P(Y \leq k|\mathbf{x}_i) - P(Y \leq k-1|\mathbf{x}_i)\} \\
 &= \sum_{i=1}^n \sum_{k=0}^K y_{ik} \log \left(\frac{e^{\tau_k - \mathbf{x}_i \boldsymbol{\beta}'}}{1 + e^{\tau_k - \mathbf{x}_i \boldsymbol{\beta}'}} - \frac{e^{\tau_{k-1} - \mathbf{x}_i \boldsymbol{\beta}'}}{1 + e^{\tau_{k-1} - \mathbf{x}_i \boldsymbol{\beta}'}} \right)
 \end{aligned} \tag{2.13}$$

The maximum likelihood estimator of $\boldsymbol{\beta}$ is found by maximizing the log-likelihood in Equation 2.13 using the Newton-Raphson method.

Prior to fitting a POM, it is important to check whether the assumption of proportionality is satisfied by each variable in the model. Methods used for checking the proportionality assumption are given in Section 2.3. In cases where the proportional odds assumption is inadequate, alternative strategies to improve the fit include (i) trying different link functions such as probit or log-log, (ii) adding additional terms such as interactions to the model, (iii) generalizing the model by adding dispersion parameters (11), (iv) using a multinomial logistic regression instead, and (v) permitting separate effects for each logit for some or all of the covariates in a model. Through sections 2.2.2 and 2.2.3, *non-proportional* and *partial proportional odds models* that allow separate effects for each logit for

all or some of the variables are presented.

2.2.2 The Non-Proportional Odds Model

The non-proportional odds model (NPOM) relaxes the assumption of proportional odds by allowing the effect of the covariates to vary with the point where the categories of the response variable are dichotomized. The NPOM is given as follows (22, 23):

$$c_k^{NP}(\mathbf{x}) = \ln \left[\frac{P(Y \leq k|\mathbf{x})}{1 - P(Y \leq k|\mathbf{x})} \right] = \tau_k - \mathbf{x}'\boldsymbol{\beta}_k \quad k = 0, 1, \dots, K - 1. \quad (2.14)$$

Note that in this model, each logit has its own intercept (τ_k) and regression coefficient ($\boldsymbol{\beta}_k$). Thus, it relaxes the proportional assumption for all covariates in the model. If only some of the covariates violate this assumption, then a more parsimonious model should be considered.

2.2.3 The Partial Proportional Odds Model

Under the partial proportional odds model (PPOM) (14), two sets of regression coefficients are estimated: one set having proportional odds and the second set having non-proportional odds. This model has the form:

$$c_k^{PP}(\mathbf{x}) = \ln \left[\frac{P(Y \leq k|\mathbf{x})}{1 - P(Y \leq k|\mathbf{x})} \right] = \tau_k - \mathbf{x}'\boldsymbol{\beta} - \mathbf{T}'\boldsymbol{\gamma}_k \quad k = 0, 1, \dots, K - 1. \quad (2.15)$$

In Equation 2.15, $\mathbf{x}$ is a $p \times 1$ vector of covariates that maintain the proportional odds property, and $\mathbf{T}$ is a $q \times 1$ vector of covariates that are non-proportional. If each covariate violates the proportionality, then the model in Equation 2.15 is equivalent to the NPOM in Equation 2.14. The independent variables (and thus the corresponding regression coefficients) which meet or do not meet the proportionality assumption in the PPOM, can be determined with the Brant's Wald test statistic given in Section 2.3.2.

The maximum likelihood estimation in the NPOM and the PPOM is determined by replacing the terms " $\tau_k - \mathbf{x}_i\boldsymbol{\beta}'$ " and " $\tau_{k-1} - \mathbf{x}_i\boldsymbol{\beta}'$ " in Equation 2.13 with " $\tau_k - \mathbf{x}_i\boldsymbol{\beta}'_k$ " and " $\tau_{k-1} - \mathbf{x}_i\boldsymbol{\beta}'_{k-1}$ " for the NPOM and with " $\tau_k - \mathbf{x}_i\boldsymbol{\beta}' - \mathbf{T}_i\boldsymbol{\gamma}'_k$ " and " $\tau_{k-1} - \mathbf{x}_i\boldsymbol{\beta}' - \mathbf{T}_i\boldsymbol{\gamma}'_{k-1}$ " for the PPOM, respectively.

2.2.4 The Continuation Ratio Model

Feinberg (10) proposed an alternative model for the analysis of ordinal categorical data. The continuation ratio model (CRM) is based on the conditional probabilities of being in a category ($Y = k$) among all subjects who are in higher categories ($Y > k$) and is given as:

$$r_k(\mathbf{x}) = \ln \left[\frac{P(Y = k|\mathbf{x})}{P(Y > k|\mathbf{x})} \right] = \theta_k - \mathbf{x}'\boldsymbol{\beta} \quad k = 0, 1, \dots, K - 1. \quad (2.16)$$

Note that in this model, each logit has its own intercept (θ_k), but the regression coefficient ($\boldsymbol{\beta}$) does not depend on k , implying that the model again assumes that the relationship between x_i and Y is independent of k . As an example, let Y have four ordinal categories ($k = 0, 1, 2, 3$); then, in the CRM, we can construct 3 logits as given in Figure 2.4.

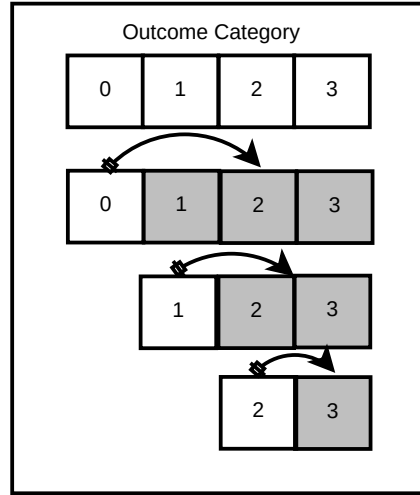


Figure 2.4: Constructing logits in the continuation ratio model

The continuation logits in Figure 2.4 possess the property of conditional independence between partitions. Because the joint probability of $P(Y = 0, Y = 1, Y = 2, Y = 3) = P(Y = 0|Y = 1, Y = 2, Y = 3)P(Y = 1|Y = 2, Y = 3)P(Y = 2|Y = 3)$, the CRM is a simple decomposition of a multinomial distribution (28). Using this property, multinomial likelihood for the model in Figure 2.4 can be decomposed into a series of conditionally independent binary logistic regressions as follows (29):

$$\begin{aligned}
L_{(1)} &= \prod_{i=1}^n \phi_0(x_i)^{y_{i0}} (1 - \phi_0(x_i))^{y_{i1}+y_{i2}+y_{i3}} \\
L_{(2)} &= \prod_{i=1}^n \left(\frac{\phi_1(x_i)}{(1 - \phi_0(x_i))} \right)^{y_{i1}} \left(1 - \frac{\phi_1(x_i)}{(1 - \phi_0(x_i))} \right)^{y_{i2}+y_{i3}} \\
L_{(3)} &= \prod_{i=1}^n \left(\frac{\phi_2(x_i)}{(1 - \phi_0(x_i) - \phi_1(x_i))} \right)^{y_{i2}} \left(1 - \frac{\phi_2(x_i)}{(1 - \phi_0(x_i) - \phi_1(x_i))} \right)^{y_{i3}}.
\end{aligned} \tag{2.17}$$

If we take the product of $L_{(1)}$, $L_{(2)}$ and $L_{(3)}$, the resulting likelihood will be the multinomial likelihood as follows:

$$\begin{aligned}
\prod_{i=1}^n \left[\prod_{k=0}^3 \phi_k(x_i)^{y_{ik}} \right] &= L_{(1)} L_{(2)} L_{(3)} \\
\prod_{i=1}^n \left[\prod_{k=0}^3 \phi_k(x_i)^{y_{ik}} \right] &= \prod_{i=1}^n \phi_0(x_i)^{y_{i0}} \phi_1(x_i)^{y_{i1}} \phi_2(x_i)^{y_{i2}} \phi_3(x_i)^{y_{i3}}.
\end{aligned} \tag{2.18}$$

Therefore, a natural way of modelling the continuation ratio logits would be to consider each ordinal response as a series of conditionally independent binary responses, each of which may be modelled via binary logistic regression. Each fit is based on a binary outcome y_k^* , defined as

$$y_k^* = \begin{cases} 1 & \text{if } Y = k \\ 0 & \text{if } Y > k \\ \text{Missing} & \text{else} \end{cases} \tag{2.19}$$

for $k = 0, 1, 2$. Wolfe (29) developed a command for fitting the continuation ratio model in STATA software. It is also possible to define the CRM in terms of $Y = k$ versus $Y < k$ for $k = 1, 2, \dots, K$. Unfortunately, the results obtained from these two parameterizations are not equivalent. In other words, unlike the POM, the CRM is neither preserved by a reversal nor the collapsibility of categories (4).

The test of the proportional odds assumption in the CRM can be performed using the likelihood ratio or the Wald test statistic. Details and an example can be found in Hosmer and Lemeshow (1, p.297). In cases where the proportional odds assumption does not hold for some or all of the covariates in the CRM, then alternative models should be considered. Through sections 2.2.5

and 2.2.6, alternative models which accommodate this assumption are considered.

2.2.5 The Non-Proportional Continuation Ratio Model

The non-proportional continuation ratio model (NPCRM) relaxes the assumption of proportional odds for all covariates in the model as follows (1, 24):

$$r_k^{NP}(\mathbf{x}) = \ln \left[\frac{P(Y = k|\mathbf{x})}{P(Y > k|\mathbf{x})} \right] = \theta_k - \mathbf{x}'\beta_k \quad k = 0, 1, \dots, K - 1. \quad (2.20)$$

Additionally, the NPCRM can be fit via K ordinary binary logistic regression models. Details and an example can be found in Hosmer and Lemeshow (1, p.296).

2.2.6 The Partial Proportional Continuation Ratio Model

If only some of the covariates violate the proportionality assumption in the CRM, then a more parsimonious model than the NPCRM should be considered. Under the partial proportional continuation ratio model (PPCRM) (25), two sets of regression coefficients are estimated: one set has proportional odds and the other set has non-proportional odds. This model has the form:

$$r_k^{PP}(\mathbf{x}) = \ln \left[\frac{P(Y = k|\mathbf{x})}{P(Y > k|\mathbf{x})} \right] = \theta_k - \mathbf{x}'\beta - \mathbf{T}'\gamma_k \quad k = 0, 1, \dots, K - 1. \quad (2.21)$$

Cole and Ananth (25) demonstrate a new method for fitting the PPCRM using a “person-threshold” data set. A separate binary logistic regression modelling approach is also valid because the PPCRM possesses the property of conditional independence.

2.2.7 The Adjacent Category Model

The adjacent category model (ACM) proceeds by forming K logits for all pairs of adjacent categories (9). In this model, each response is compared to the next lower response. The model can be written as:

$$a_k(\mathbf{x}) = \ln \left[\frac{P(Y = k|\mathbf{x})}{P(Y = k - 1|\mathbf{x})} \right] = \alpha_k + \mathbf{x}'\beta \quad k = 1, 2, \dots, K. \quad (2.22)$$

This model assumes that the log odds do not depend on the response

category k . As an example, let Y have four ordinal categories ($k = 0, 1, 2, 3$). Then, in the ACM, we can construct 3 logits as given in Figure 2.5.

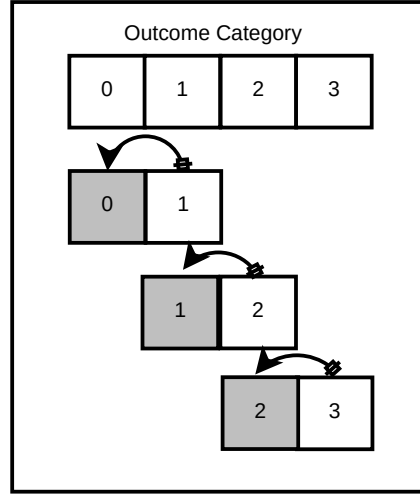


Figure 2.5: Constructing adjacent category logits in the adjacent category model

As seen in Equation 2.22, the adjacent category logistic regression model also assumes that the assumption of proportionality is satisfied by each variable in the model. The likelihood ratio test can be used to check the proportionality assumption in the adjacent category model in a global sense, but there is no valid approach to test if this assumption is satisfied individually.

The ACM is a constrained version of the MLRM. Thus, model parameters and standard error estimates can be directly obtained from the MLRM. The MLRM can be expressed in terms of the ACM as follows (1, 9):

$$\begin{aligned}
 \ln \left[\frac{P(Y = k|\mathbf{x})}{P(Y = 0|\mathbf{x})} \right] &= \ln \left[\frac{P(Y = 1|\mathbf{x})}{P(Y = 0|\mathbf{x})} \right] + \ln \left[\frac{P(Y = 2|\mathbf{x})}{P(Y = 1|\mathbf{x})} \right] + \dots + \ln \left[\frac{P(Y = k|\mathbf{x})}{P(Y = k-1|\mathbf{x})} \right] \\
 &= a_1(\mathbf{x}) + a_2(\mathbf{x}) + \dots + a_k(\mathbf{x}) \\
 &= (\alpha_1 + \mathbf{x}'\beta) + (\alpha_2 + \mathbf{x}'\beta) + \dots + (\alpha_k + \mathbf{x}'\beta) \\
 &= (\alpha_1 + \alpha_2 + \dots + \alpha_k) + k\mathbf{x}'\beta.
 \end{aligned}
 \tag{2.23}$$

Thus, we see that the model in Equation 2.23 is a version of the MLRM in Equation 2.1 with intercept $\beta_{0k} = (\alpha_1 + \alpha_2 + \dots + \alpha_k)$ and regression coefficient $\beta_k = k\beta$. In order to obtain the ACM estimates of α_k 's and β 's from the MLRM, we fit a MLRM using the two constraints given below:

$$\beta_{0k} = \sum_{i=1}^k \alpha_i \quad \text{and} \quad \beta_k = k\beta. \quad (2.24)$$

This approach for fitting ACM via MLRM is explained for STATA software (30), and a real data example can be found in Hosmer and Lemeshow's book (1, p. 294). Note that with this formalization, ACM gives rise to an ordinal interpretation in terms of adjacent odds where the MLRM does not.

The ACM became popular because of connections with the ordinal log-linear models (9). The corresponding ordinal log-linear model for the adjacent category logit model is the linear-by-linear association model. For two-way tables, let the assigned row scores and column scores of ordinal variables be $u_1 < u_2 < \dots < u_I$ and $v_1 < v_2 < \dots < v_J$, respectively. Then, the linear-by-linear association model is

$$\log \mu_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \beta u_i v_j, \quad (2.25)$$

with constraints such as $\lambda_I^X = \lambda_J^Y = 0$. The term $\beta u_i v_j$ in Equation 2.25 represents the deviation of $\log \mu_{ij}$ from independence. The deviation is linear in the Y scores at a fixed level of X and linear in the X scores at a fixed level of Y . Because of this property, this model is called the *linear-by-linear association model*.

A logit formulation of the linear-by-linear association model treats Y as a response and X as explanatory. Using logits for adjacent response categories,

$$\log \frac{\mu_{i,j+1}}{\mu_{i,j}} = (\lambda_{j+1}^Y - \lambda_j^Y) + \beta(v_{j+1} - v_j)u_i, \quad (2.26)$$

and for unit spaced v_j , the linear-by-linear association model simplifies to the ACM as

$$\log \frac{\pi_{j+1|i}}{\pi_{j|i}} = \alpha_j + \beta u_i, \quad (2.27)$$

where $\alpha_j = \lambda_{j+1}^Y - \lambda_j^Y$. In the linear-by-linear association model, local odds ratios for adjacent rows and columns have a common value e^β .

Additionally, using the connection between the ACM and the linear-by-linear association model, the parameter estimates in the ACM can be obtained from the Poisson log-likelihood. The log-likelihood in the linear-by-linear association model is

$$\begin{aligned} \mathbf{L}(\mu) = & n\lambda + \sum_i n_{i+} \lambda_i^X + \sum_j n_{+j} \lambda_j^Y + \beta \sum_i \sum_j u_i v_j n_{ij} \\ & - \sum_i \sum_j \exp(\lambda + \lambda_i^X + \lambda_j^Y + \beta u_i v_j). \end{aligned} \quad (2.28)$$

Differentiating $\mathbf{L}(\mu)$ with respect to $(\lambda_i^X, \lambda_j^Y, \beta)$ and using iterative methods such as Newton-Raphson method yields the maximum likelihood fit.

2.2.8 The Non-Proportional Adjacent Category Model

Like the proportional odds model and the continuation ratio model, the adjacent category logistic regression model assumes that the assumption of proportionality is satisfied by each variable in the model. The non-proportional adjacent category model (NPACM) relaxes the assumption of proportional odds for all variables in the model and is given as follows (31):

$$a_k^{NP}(\mathbf{x}) = \ln \left[\frac{P(Y = k|\mathbf{x})}{P(Y = k-1|\mathbf{x})} \right] = \alpha_k + \mathbf{x}' \beta_k \quad k = 1, 2, \dots, K. \quad (2.29)$$

The NPACM is also a constrained version of the MLRM. Thus, model parameters and standard error estimates can be directly obtained from the MLRM as follows:

$$\begin{aligned} \ln \left[\frac{P(Y = k|\mathbf{x})}{P(Y = 0|\mathbf{x})} \right] &= \ln \left[\frac{P(Y = 1|\mathbf{x})}{P(Y = 0|\mathbf{x})} \right] + \ln \left[\frac{P(Y = 2|\mathbf{x})}{P(Y = 1|\mathbf{x})} \right] + \dots + \ln \left[\frac{P(Y = k|\mathbf{x})}{P(Y = k-1|\mathbf{x})} \right] \\ &= a_1(\mathbf{x}) + a_2(\mathbf{x}) + \dots + a_k(\mathbf{x}) \\ &= (\alpha_1 + \mathbf{x}' \beta_1) + (\alpha_2 + \mathbf{x}' \beta_2) + \dots + (\alpha_k + \mathbf{x}' \beta_k) \\ &= (\alpha_1 + \alpha_2 + \dots + \alpha_k) + \mathbf{x}' (\beta_1 + \beta_2 + \dots + \beta_k). \end{aligned} \quad (2.30)$$

Thus, we see that the model in Equation 2.29 is a version of the MLRM in Equation 2.1 with intercept $\beta_{0k} = (\alpha_1 + \alpha_2 + \dots + \alpha_k)$ and regression coefficient $\beta_k = (\beta_1 + \beta_2 + \dots + \beta_k)$. In order to obtain the NPACM estimates of α_k 's and β_k 's from the MLRM, we fit a MLRM using the two constraints given below:

$$\beta_{0k} = \sum_{i=1}^k \alpha_i \quad \text{and} \quad \beta_k = \sum_{i=1}^k \beta_i. \quad (2.31)$$

Because the two constraints given in Equation 2.31 are additive rather than multiplicative, the NPACM and MLRM models have equal log-likelihoods. For this reason, the NPACM can be seen as an MLRM with a different parameterization (i.e., using adjacent logits rather than baseline logits).

2.3 Methods Used for Checking the Proportional Odds Assumption

Proportional odds, continuation ratio, and adjacent category models employ a strict assumption that the odds ratio (or log odds) does not depend on k . Prior to fitting any of them, it is important to determine whether the assumption of proportionality is satisfied by each independent variable j . In the next two sections, we will introduce methods used for checking the proportionality assumption in the ordinal logistic regression models.

2.3.1 The Likelihood Ratio Test

Let ordinal dependent variable Y take values of $k = 1, 2, \dots, K$ (note that, this display is different from that given as $k = 0, 1, \dots, K$ previously), and the proportional odds model has j independent variables ($j = 1, 2, \dots, J$). When the proportional odds assumption holds, the regression coefficients corresponding to the j th independent variable (β_{jk}) are equal to each other. This hypothesis can be formulated as follows (21):

$$H_0 : \beta_{1j} = \beta_{2j} = \dots = \beta_{(K-1)j} \quad j = 1, 2, \dots, J. \quad (2.32)$$

As the proportional odds model is nested within the non-proportional odds model, the likelihood ratio (LR) test, which evaluates differences in deviances, provides an omnibus test of proportionality. Under H_0 , the LR test statistic is distributed as chi-square with degrees of freedom equal to the difference in the number of parameters between the nested models (which is $J(K - 2)$)(25). The LR test is a very useful tool, and it can be used to check the proportionality assumption in the continuation and adjacent category models in a global sense.

Another method, an approximate likelihood ratio test (32), is also an omnibus test of proportionality that tests whether the coefficients for all variables are simultaneously equal. However, it compares the log-likelihood of the proportional odds model to the log-likelihood obtained from pooling $K - 1$ binary logistic regression models, adjusting for the correlation between $K - 1$ binary logistic regression models. This approach is an approximation to the LR test

because the log-likelihood obtained from pooling $K - 1$ binary logistic regression models is only an approximation to the likelihood of the non-proportional odds model.

2.3.2 Brant's Wald Test

The LR test is an omnibus test that does not show whether the proportionality assumption is violated for all independent variables or only for some. Brant (26) proposed a Wald test that tests the proportionality assumption for each variable individually and together. The approach is based on viewing the model in Equation 2.14 as a combination of $K - 1$ correlated binary logistic regression models. For $k = 1, 2, \dots, K$, $j = 1, 2, \dots, J$ and $i = 1, 2, \dots, N$, this test is constructed as follows (21):

1. Estimate $\hat{\beta}_k$'s and $Var(\hat{\beta}_k)$'s obtained from the $K - 1$ binary logistic regressions with outcomes defined by

$$z_k = \begin{cases} 1 & \text{if } Y > k \\ 0 & \text{if } Y \leq k \end{cases} \quad (2.33)$$

2. Estimate the probability that $z_{ik} = 1$ given x_i , by using the below formula:

$$\hat{\pi}_k(x_i) = \frac{e^{\hat{\tau}_k + \mathbf{x}' \hat{\beta}_k}}{1 + e^{\hat{\tau}_k + \mathbf{x}' \hat{\beta}_k}} \quad (2.34)$$

3. Estimate the covariance between $\hat{\beta}_k$ and $\hat{\beta}_l$ ($k \leq l$). Define:

$$w_{ikl} = \hat{\pi}_l(x_i) - \hat{\pi}_k(x_i)\hat{\pi}_l(x_i) \quad (2.35)$$

and let W_{ikl} be an $N \times N$ diagonal matrix whose i th element is w_{ikl} . Let X be the $N \times (J + 1)$ matrix with 1's in the first column and the independent variables in the remaining columns. Brant (26) showed that the covariances among the $\hat{\beta}$'s from different binary equations $\widehat{Var}(\hat{\beta}_k, \hat{\beta}_l)$ are estimated by deleting the first row and column of

$$(X'W_{kk}X)^{-1}(X'W_{kl}X)(X'W_{ll}X)^{-1} \quad (2.36)$$

4. Combine all estimates. Define $\hat{\beta}_k^* = (\hat{\beta}_1' \hat{\beta}_2' \dots \hat{\beta}_{K-1}')'$ and

$$\widehat{Var}(\hat{\beta}_k^*) = \begin{pmatrix} \widehat{Var}(\hat{\beta}_1) & \dots & \widehat{Var}(\hat{\beta}_1, \hat{\beta}_{K-1}) \\ \vdots & \ddots & \vdots \\ \widehat{Var}(\hat{\beta}_{K-1}, \hat{\beta}_1) & \dots & \widehat{Var}(\hat{\beta}_1) \end{pmatrix} \quad (2.37)$$

The diagonal elements $\widehat{Var}(\hat{\beta}_k)$ are the covariance matrices from each binary regression. The off-diagonal elements were defined at step 3.

5. Construct the Wald test of $H_0 : \beta_1 = \beta_2 = \dots = \beta_{(K-1)}$. This hypothesis corresponds to $H_0 : \mathbf{D}\beta^* = 0$ where

$$\mathbf{D} = \begin{pmatrix} \mathbf{I} & -\mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & -\mathbf{I} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \dots & -\mathbf{I} \end{pmatrix} \quad (2.38)$$

$\mathbf{I}$ is a $(J+1) \times (J+1)$ identity matrix, and $\mathbf{0}$ is a $(J+1) \times (J+1)$ matrix of 0's. The Wald test statistic

$$W = (\mathbf{D}\hat{\beta}^*)' [\mathbf{D}\widehat{Var}(\hat{\beta}^*)\mathbf{D}']^{-1} (\mathbf{D}\hat{\beta}^*) \quad (2.39)$$

distributes as chi-square with $J(K-2)$ degrees of freedom.

6. Construct tests for individual variables. The hypothesis $H_0 : \beta_{1j} = \beta_{2j} = \dots = \beta_{(K-1)j}$ can be tested by selecting only those rows and columns of $\mathbf{D}$, $\hat{\beta}^*$ and $\widehat{Var}(\hat{\beta}^*)$ that correspond to the coefficients being tested. The resulting test statistic has $K-2$ degrees of freedom. For the overall and individual tests, if the Wald test statistic is big enough, the proportionality assumption can be rejected.

Brant (26) proved that under the null hypothesis, the separate maximum likelihood estimates $\hat{\beta}_k^* = (\hat{\beta}'_1 \hat{\beta}'_2 \dots \hat{\beta}'_{K-1})'$ have a distribution that is asymptotically multivariate normal with $E(\hat{\beta}_k^*) = \beta_k$ and the asymptotic variance-covariance $\widehat{Var}(\hat{\beta}_k^*)$. Thus, the Wald test statistic, which is based on the correlated separate fits, provides an assessment of the proportionality assumption in the POM.

It is an important task to check the proportionality assumption in any ordinal logistic regression model. Bender and Grouven (2) showed that use of the

POM can lead to invalid results if the main model assumption is violated, and they make a point for POM as follows:

“With increasing use of the POM, the misuse of it will also increase, especially due to application of this method to data with non-proportional odds. Unfortunately, the results of logistic regression models are poorly presented in medical research papers. Hence, an invalid application of the POM is seldom detectable for readers of medical journals.”

Therefore, applying a proportionality test before proceeding with any of the ordinal logistic regression models is crucial. However, the existing methods for testing proportionality assumptions are restricted. An LR test is applicable for all of the ordinal logistic models, but it is only a way of testing proportionality in a global sense, and the Brant’s test of Wald must be generalized for other ordinal logistic regression models such as the adjacent category model.

3 MATERIAL AND METHODS

3.1 The Proposed Wald Test Statistic: Adaptation of Brant's Test for Adjacent Category Model

Brant (26) showed that the separate binary logistic regression estimates of $\hat{\beta}_k$'s are asymptotically unbiased and follow a multivariate normal distribution with a mean of β_k and covariance given in Equation 2.36. Following the idea of Brant, we adapted a test for checking the proportionality assumption in the adjacent category model as follows:

Proposition:

1. The ACM uses adjacent category logits, so the first step of the proposed test statistic is to estimate $\hat{\beta}_k$'s and $Var(\hat{\beta}_k)$'s from the $K - 1$ binary logistic regressions with outcomes defined by

$$z_k = \begin{cases} 1 & \text{if } Y = k \\ 0 & \text{if } Y = k - 1 \\ Missing & \text{else.} \end{cases} \quad (3.1)$$

2. Estimate the probability that $z_{ik} = 1$ given x_i by using the below formula:

$$\hat{\pi}_k(x_i) = \frac{e^{\hat{\tau}_k + \mathbf{x}' \hat{\beta}_k}}{1 + e^{\hat{\tau}_k + \mathbf{x}' \hat{\beta}_k}} \quad (3.2)$$

Note that in this step, the probabilities ($\hat{\pi}_k(x_i)$) that correspond to the missing categories of z_k 's should remain as missing because there will not be any positive outcome probability that corresponds to the missing outcome category in the binary logistic regression model.

3. Estimate the covariance between $\hat{\beta}_k$ and $\hat{\beta}_l$ ($k \leq l$). Define:

$$w_{ikl} = \hat{\pi}_l(x_i) - \hat{\pi}_k(x_i)\hat{\pi}_l(x_i) \quad (3.3)$$

and let W_{ikl} be a $N^* \times N^*$ diagonal matrix whose i th element is w_{ikl} , and N^* is the number of cases when we exclude the missing cases listwise. Let X be the $N^* \times (J + 1)$ matrix with 1's in the first column and the independent variables in the remaining columns. Estimate the $\widehat{Var}(\hat{\beta}_k, \hat{\beta}_l)$ by deleting the first row and column of

$$(X'W_{kk}X)^{-1}(X'W_{kl}X)(X'W_{ll}X)^{-1} \quad (3.4)$$

4. Combine all estimates. Define $\hat{\beta}_k^* = (\hat{\beta}'_1 \hat{\beta}'_2 \dots \hat{\beta}'_{K-1})'$ and

$$\widehat{Var}(\hat{\beta}_k^*) = \begin{pmatrix} \widehat{Var}(\hat{\beta}_1) & \dots & \widehat{Var}(\hat{\beta}_1, \hat{\beta}_{K-1}) \\ \vdots & \ddots & \vdots \\ \widehat{Var}(\hat{\beta}_{K-1}, \hat{\beta}_1) & \dots & \widehat{Var}(\hat{\beta}_1) \end{pmatrix} \quad (3.5)$$

The diagonal elements $\widehat{Var}(\hat{\beta}_k)$ are the covariance matrices from each binary regression. The off-diagonal elements were defined at step 3.

5. Construct the Wald test of $H_0 : \beta_1 = \beta_2 = \dots = \beta_{(K-1)}$. This hypothesis corresponds to $H_0 : \mathbf{D}\beta^* = 0$ where

$$\mathbf{D} = \begin{pmatrix} \mathbf{I} & -\mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & -\mathbf{I} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \dots & -\mathbf{I} \end{pmatrix} \quad (3.6)$$

$\mathbf{I}$ is a $(J+1) \times (J+1)$ identity matrix, and $\mathbf{0}$ is a $(J+1) \times (J+1)$ matrix of 0's. The Wald test statistic

$$W = (\mathbf{D}\hat{\beta}^*)' [\mathbf{D}\widehat{Var}(\hat{\beta}^*)\mathbf{D}']^{-1} (\mathbf{D}\hat{\beta}^*) \quad (3.7)$$

distributes as chi-square with $J(K-2)$ degrees of freedom.

6. Construct tests for individual variables. The hypothesis $H_0 : \beta_{1j} = \beta_{2j} = \dots = \beta_{(K-1)j}$ can be tested by selecting only those rows and columns of $\mathbf{D}$, $\hat{\beta}^*$ and $\widehat{Var}(\hat{\beta}^*)$ that correspond to the coefficients being tested. The resulting test statistic has $K-2$ degrees of freedom. For the overall and individual tests, if the Wald test statistic is big enough, the proportionality assumption can be rejected.

The validity of the proposed test is examined under H_0 with a Monte Carlo simulation study in Section 3.3, and the simulation results are given in Chapter 4. Additionally, an example of the proposed test of proportionality calculations in the adjacent category model is given in Chapter 5.

3.2 The Partial Proportional Adjacent Category Model

3.2.1 Model Specification

When some of the independent variables in the ACM do not meet the proportionality assumption, a need for a more parsimonious model than the NPACM naturally arises. Following the idea that was done before for the POM and the CRM, a *partial proportional model* for the ACM can be constructed as follows:

$$a_k^{PP}(\mathbf{x}) = \ln \left[\frac{P(Y = k|\mathbf{x})}{P(Y = k-1|\mathbf{x})} \right] = \alpha_k + \mathbf{x}'\beta + \mathbf{T}'\gamma_k \quad k = 1, 2, \dots, K. \quad (3.8)$$

Under the partial proportional adjacent category model (PPACM) in Equation (3.8), two sets of regression coefficients should be estimated: one set (denoted by β) meets the proportional odds assumption, and a second set (denoted by γ_k) violates the proportional odds assumption.

At this point, the independent variables (and thus the corresponding coefficients) that meet or do not meet the proportionality assumption in the PPACM can be determined with the proposed Wald test statistic given in Section 3.1.

3.2.2 Parameter Estimation

In order to estimate the β and γ_k 's in the PPACM, we can utilize the relationships between the adjacent category model, the non-proportional adjacent category model, and the multinomial logistic regression model which are given in sections 2.2.7 and 2.2.8.

Fitting (with the maximum likelihood estimation method) a multinomial logistic regression model using the two constraints given in Equation 2.31 gives the parameter estimates for a fully non-proportional adjacent category model (NPACM). On the other hand, fitting a multinomial logistic regression model using the two constraints given in Equation 2.24 gives the parameter estimates of a fully proportional adjacent category model (ACM). Then, fitting a multinomial logistic regression model using constraints given in Equation 2.24 for only some of the parameters in the model while leaving the rest of them without any constraints will give the parameter estimates of a partial version of the adjacent category model (PPACM). Yee (33) recently published the R codes (VGAM) for fitting the partial proportional adjacent category model. An example of the strategy used for parameter estimation in the partial proportional adjacent category model is given in Chapter 5.

3.3 Simulation Study

To determine whether the proposed test in Equation 3.7 is distributed as chi-square or not, we conducted a Monte Carlo simulation. In the simulation study, we simulated the distribution of the proposed test statistic under H_0 , and then we examined the appropriateness of chi-square approximation using some graphical methods such as Q-Q plots, probability distribution and cumulative distribution plots.

3.3.1 Simulation Study 1: Simulation of the Adapted Test Statistic in 2×3 Tables

For the first simulation study, the simplest situation in the adjacent category model with three ordinal outcomes ($Y = 3$ and $k = 1, 2, 3$) and one binary independent variable ($X = 0, 1$) is considered. In this case, frequencies in the ACM are given in Table 3.1.

Table 3.1: The frequency and the probability notations in the adjacent category model with three outcomes and one binary independent variable.

Independent Variable X	Outcome Variable Y		
	1	2	3
0	$n_{01}(p_{01})$	$n_{02}(p_{02})$	$n_{03}(p_{03})$
1	$n_{11}(p_{11})$	$n_{12}(p_{12})$	$n_{13}(p_{13})$

In order to simulate the distribution of the test statistic under H_0 , we rewrite the null hypothesis of proportionality in the ACM with the frequencies given in Table 3.1 as follows:

$$H_0 : \frac{n_{01} \cdot n_{12}}{n_{11} \cdot n_{02}} = \frac{n_{02} \cdot n_{13}}{n_{12} \cdot n_{03}} = OR \quad (3.9)$$

Equation 3.9 is another way of expressing the proportional odds assumption in the ACM, and it provides that the adjacent category odds ratios across the outcome categories are equal to each other. In order to achieve different values for the common adjacent odds ratio (i.e., $OR = 1, 2, \dots, 5$) from Equation 3.9, we generate frequencies from the multinomial distribution with probabilities given in Table 3.2.

Also, to consider the sample size effect, we used 4 sample size ($n = 200, 300, 500, 1000$) combinations for the multinomial distribution. In total, we

Table 3.2: Multinomial probabilities that are used for simulation study 1.

Common Adjacent OR	Multinomial Probabilities in the 2×3 Table					
	p_{01}	p_{02}	p_{03}	p_{11}	p_{12}	p_{13}
$OR = 1$	0.095	0.382	0.190	0.048	0.190	0.095
$OR = 2$	0.240	0.240	0.120	0.080	0.160	0.160
$OR = 3$	0.088	0.147	0.147	0.029	0.147	0.442
$OR = 4$	0.067	0.100	0.067	0.033	0.200	0.533
$OR = 5$	0.147	0.147	0.088	0.029	0.147	0.442

used 20 different scenarios, which are given in Table 3.3. We repeated each common adjacent odds ratio and sample size combination 1,000 times, and in each run, we computed the proposed test statistic given in Equation 3.7. We also calculated the likelihood ratio test statistics using the log-likelihood of the non-proportional adjacent category model (L_{NPACM}) and the adjacent category model (L_{ACM}), and we compared the proposed test statistic with the likelihood ratio test results.

Table 3.3: Scenarios that are used for simulation study 1.

Scenario Number	Common Adjacent OR	Frequencies in the 2×3 Table						Total Sample Size
		n_{01}	n_{02}	n_{03}	n_{11}	n_{12}	n_{13}	
1	$OR = 1$	19	76	38	10	38	19	200
2	$OR = 1$	28	115	57	15	57	28	300
3	$OR = 1$	47	191	95	24	95	48	500
4	$OR = 1$	95	382	190	48	190	95	1000
5	$OR = 2$	48	48	24	16	32	32	200
6	$OR = 2$	72	72	36	24	48	48	300
7	$OR = 2$	120	120	60	40	80	80	500
8	$OR = 2$	240	240	120	80	160	160	1000
9 [†]	$OR = 3$	18	29	29	6	29	89	200
10 [†]	$OR = 3$	27	44	44	9	44	132	300
11	$OR = 3$	44	73	73	15	74	221	500
12	$OR = 3$	88	147	147	29	147	442	1000
13 [†]	$OR = 4$	13	20	13	7	40	107	200
14	$OR = 4$	20	30	20	10	60	160	300
15	$OR = 4$	33	50	33	17	100	267	500
16	$OR = 4$	67	100	67	33	200	533	1000
17 [†]	$OR = 5$	29	29	18	6	29	89	200
18	$OR = 5$	44	44	27	9	44	132	300
19	$OR = 5$	74	74	44	14	73	221	500
20	$OR = 5$	147	147	88	29	147	442	1000

[†]The scenarios which the ad hoc adjustment is applied.

Note that, when one of the cell frequencies in the 2×3 tables equals 0, the resulting standard error estimate of adjacent odds ratio is infinity. In such scenarios (they are marked with [†] in Table 3.3) where the generated 2×3 table

contains a cell with 0 frequency, we applied the commonly-used ad hoc adjustment procedure of adding 0.50 to each cell frequency in that particular table.

The appropriateness of the chi-square approximation of the proposed test statistic was investigated using some graphical methods such as the probability density function (pdf), cumulative distribution function (cdf) and Q-Q plots as performed by Brant (26). Additionally, we obtained the estimated type I error rates for the proposed test and the LR tests as the percentage of times the test rejects the null hypothesis when the null hypothesis is true.

The simulation codes were developed in R (Version 2.11.1), and the simulation routine can be found in Appendix A. Results of Simulation Study 1 are given in Chapter 4.

3.3.2 Simulation Study 2: Simulation of the Adapted Test Statistic in $R \times C$ Tables

In Simulation Study 1, we investigated the distribution of the proposed test statistic under H_0 for the simplest case in the ACM. However, the results of Simulation Study 1 needed to be generalized for the $R \times C$ tables. For this purpose we re-conducted the simulation study for $R \times C$ tables.

Suppose that the ordinal dependent variable has C levels ($j = 1, 2, \dots, C$) and the independent variable has R levels ($i = 1, 2, \dots, R$), and let the frequencies in the $R \times C$ table be given as Table 3.4.

Table 3.4: Frequencies in a $R \times C$ table.

Independent Variable X	Outcome Variable Y			
	1	2	...	C
1	n_{11}	n_{12}	...	n_{1C}
2	n_{21}	n_{22}	...	n_{2C}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
R	n_{R1}	n_{R2}	...	n_{RC}

In the $R \times C$ table, the null hypothesis of proportionality or in other words, the common odds ratio across the adjacent categories, can be expressed as

$$H_0 : \frac{n_{ij} \cdot n_{(i+1)(j+1)}}{n_{i(j+1)} \cdot n_{(i+1)j}} = OR \quad i = 1, 2, \dots, R \quad j = 1, 2, \dots, C. \quad (3.10)$$

To generate frequencies from the multinomial distribution under Equation 3.10, one must calculate the probabilities of the adjacent categories that gives

the same odds ratio across every category. This calculation is simple for a 2×3 table; however, for an $R \times C$ table, it can be tedious. Therefore, upon generating frequencies in an $R \times C$ table under H_0 , we use the connection between the adjacent category model and the linear-by-linear model that is given in Section 2.2.7.

Data for Simulation Study 2 is generated under the following linear-by-linear model:

$$\log \mu_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \beta u_i v_j, \quad (3.11)$$

where λ is the general effect parameter, and λ_i^X and λ_j^Y are the main effects for variables X and Y, with constraints $\lambda_I^X = \lambda_J^Y = 0$. Under the linear-by-linear model, we guaranteed that the generated frequencies in the $R \times C$ table will satisfy Equation 3.10 and the common adjacent odds ratio as $\exp(\beta) = OR$.

In large samples, the parameters (λ , λ_i^X and λ_j^Y) in the linear-by-linear model are approximately normally distributed (9), and the formal significance testing of the parameters can be performed using a Wald test statistic. Therefore, we used the normal distribution for generating the parameters of the linear-by-linear model. Mean values of β were taken as $\beta = 0, 0.182, 0.262$ to obtain the common odds ratio values of $OR = 1, 1.2, 1.3$. We did not consider using any OR values greater than 1.3, because when we generate cell frequencies under H_0 , for let's say $OR=1.5$ in the 4×5 table, six cells ($6/20=30\%$ of the cells) have expected value less than five in the table, even if $N=1000$. Additionally when we increase the OR, for example when $OR=2$, fourteen cells ($14/20=70\%$ of the cells) have expected value less than five, which may lead us to invalid results for the simulation study.

The sample size was taken as $N = 300, 500, 1000$, and we applied the ad hoc adjustment procedure of adding 0.50 to each cell if the generated $R \times C$ table contained cells with 0 frequencies. Mean, standard deviation of the linear-by-linear model parameters, common adjacent odds ratios, R and C values, and sample sizes that were used in the simulation study are given in Table 3.5.

Table 3.5: Mean, standard deviation of the linear-by-linear model parameters, common adjacent odds ratios, $R \times C$ values, and sample sizes that are used in the simulation study 2.

Scenario Number	Distribution of the parameters			Common Adjacent OR	$R \times C$	Total Sample Size
	λ	λ_i^X	λ_j^Y			
1	N(1,1)	N(0,1)	N(0,1)	$OR = 1$	3×4	300
2	N(1,1)	N(0,1)	N(0,1)	$OR = 1$	3×4	500
3	N(1,1)	N(0,1)	N(0,1)	$OR = 1$	4×5	300
4	N(1,1)	N(0,1)	N(0,1)	$OR = 1$	4×5	500
5	N(1,1)	N(0,1)	N(0,1)	$OR = 1.2$	3×4	500
6	N(1,1)	N(0,1)	N(0,1)	$OR = 1.2$	3×4	1000
7	N(1,1)	N(0,1)	N(0,1)	$OR = 1.2$	4×5	500
8	N(1,1)	N(0,1)	N(0,1)	$OR = 1.2$	4×5	1000
9	N(1,1)	N(0,1)	N(0,1)	$OR = 1.3$	3×4	500
10	N(1,1)	N(0,1)	N(0,1)	$OR = 1.3$	3×4	1000
11	N(1,1)	N(0,1)	N(0,1)	$OR = 1.3$	4×5	500
12	N(1,1)	N(0,1)	N(0,1)	$OR = 1.3$	4×5	1000
13	N(1,1)	N(1,0.5)	N(1,0.5)	$OR = 1$	3×4	300
14	N(1,1)	N(1,0.5)	N(1,0.5)	$OR = 1$	3×4	500
15	N(1,1)	N(1,0.5)	N(1,0.5)	$OR = 1$	4×5	300
16	N(1,1)	N(1,0.5)	N(1,0.5)	$OR = 1$	4×5	500
17	N(1,1)	N(1,0.5)	N(1,0.5)	$OR = 1.2$	3×4	500
18	N(1,1)	N(1,0.5)	N(1,0.5)	$OR = 1.2$	3×4	1000
19	N(1,1)	N(1,0.5)	N(1,0.5)	$OR = 1.2$	4×5	500
20	N(1,1)	N(1,0.5)	N(1,0.5)	$OR = 1.2$	4×5	1000
21	N(1,1)	N(1,0.5)	N(1,0.5)	$OR = 1.3$	3×4	500
22	N(1,1)	N(1,0.5)	N(1,0.5)	$OR = 1.3$	3×4	1000
23	N(1,1)	N(1,0.5)	N(1,0.5)	$OR = 1.3$	4×5	500
24	N(1,1)	N(1,0.5)	N(1,0.5)	$OR = 1.3$	4×5	1000
25	N(1,1)	N(1,1)	N(1,1)	$OR = 1$	3×4	300
26	N(1,1)	N(1,1)	N(1,1)	$OR = 1$	3×4	500
27	N(1,1)	N(1,1)	N(1,1)	$OR = 1$	4×5	300
28	N(1,1)	N(1,1)	N(1,1)	$OR = 1$	4×5	500
29	N(1,1)	N(1,1)	N(1,1)	$OR = 1.2$	3×4	500
30	N(1,1)	N(1,1)	N(1,1)	$OR = 1.2$	3×4	1000
31	N(1,1)	N(1,1)	N(1,1)	$OR = 1.2$	4×5	500
32	N(1,1)	N(1,1)	N(1,1)	$OR = 1.2$	4×5	1000
33	N(1,1)	N(1,1)	N(1,1)	$OR = 1.3$	3×4	500
34	N(1,1)	N(1,1)	N(1,1)	$OR = 1.3$	3×4	1000
35	N(1,1)	N(1,1)	N(1,1)	$OR = 1.3$	4×5	500
36	N(1,1)	N(1,1)	N(1,1)	$OR = 1.3$	4×5	1000

$R \times C$ tables were generated 1,000 times from the given R , C , β , λ , λ_i^X and λ_j^Y combination, and in each run, the proposed Wald test statistic, which is given in Equation 3.7, was computed. We also computed the proportionality tests for the individual variables in the model. The appropriateness of the chi-square approximation of the omnibus and the individual test statistics were investigated using graphical methods such as the probability density function, cumulative distribution function and Q-Q plots. We also calculated the likelihood ratio test statistics using the log-likelihood of the non-proportional adjacent category model (L_{NPACM}) and the adjacent category model (L_{ACM}), and we compared the proposed test statistic with the likelihood ratio test results.

We obtained the estimated type I error rates for the proposed Wald and the LR tests as the percentage of times the test rejects the null hypothesis when the null hypothesis is true. The simulation codes were developed in R (Version 2.11.1), and the simulation routine can be found in Appendix B. Results of Simulation Study 2 are summarized in Chapter 4.

4 RESULTS

4.1 Results of Simulation Study 1

Because the simplest adjacent category model has three outcomes and a binary independent variable, the adapted Wald test statistic should be distributed as chi-square with $J(K - 2) = 1 \times (3 - 2) = 1$ (note that, for a 2×3 table, the number of independent variables is $J = C - 1 = 2 - 1 = 1$ and $K = 3$) degrees of freedom.

In order to check this, we calculated two measures and we used some graphical methods. First, we calculated the mean values of the proposed test and the likelihood ratio test statistics from 1,000 replications. Secondly, we calculated the type I error rates for both test statistics under the χ_1^2 distribution. All of the results from 20 combinations are given in Table 4.1.

Table 4.1: Results of the simulation study 1.

Scenario Number	Common	Total	Mean Values of		Type I error under χ_1^2 distr.	
	Adjacent OR	Sample Size	LR test	Prop. test	LR test	Prop. test
1	$OR = 1$	200	0.949	0.933	0.045	0.042
2	$OR = 1$	300	1.032	1.020	0.050	0.050
3	$OR = 1$	500	1.110	1.103	0.063	0.062
4	$OR = 1$	1000	0.999	0.997	0.052	0.050
5	$OR = 2$	200	0.985	0.972	0.050	0.046
6	$OR = 2$	300	1.030	1.022	0.051	0.050
7	$OR = 2$	500	0.989	0.985	0.038	0.038
8	$OR = 2$	1000	1.018	1.016	0.047	0.047
9 [†]	$OR = 3$	200	0.956	0.920	0.041	0.036
10 [†]	$OR = 3$	300	1.067	1.039	0.056	0.048
11	$OR = 3$	500	0.970	0.954	0.049	0.047
12	$OR = 3$	1000	1.050	1.043	0.062	0.060
13 [†]	$OR = 4$	200	1.056	1.021	0.052	0.042
14	$OR = 4$	300	1.050	1.025	0.070	0.062
15	$OR = 4$	500	0.998	0.987	0.051	0.049
16	$OR = 4$	1000	1.036	1.028	0.051	0.050
17 [†]	$OR = 5$	200	1.070	1.022	0.056	0.047
18	$OR = 5$	300	0.974	0.947	0.052	0.045
19	$OR = 5$	500	1.049	1.029	0.058	0.054
20	$OR = 5$	1000	0.996	0.987	0.053	0.052

[†]The scenarios which the ad hoc adjustment is applied.

LR test: The Likelihood Ratio Test Statistic

Prop. test: The Proposed Wald Test Statistic

dist. : Distribution

In Table 4.1, the mean values of the test statistics that were computed from 1,000 random samples are given in columns 4 and 5. Because the expected

value of the chi-square distribution is equal to its own degrees of freedom, the mean values of the adapted Wald test statistic (and also the likelihood ratio test statistic) should be around 1. As seen in columns 4 and 5, the values of the likelihood ratio test statistics and the proposed test statistics are close to each other and very close to 1.

The sixth and seventh columns of Table 4.1 give the percentage of rejections of the proposed test statistic and the LR test statistic under the χ_1^2 distribution. As seen in these columns, type I error rates of both test statistics are around the 0.05 level. Moreover, the type I error rates of the proposed test statistic are always smaller than or equal to the likelihood ratio test statistics.

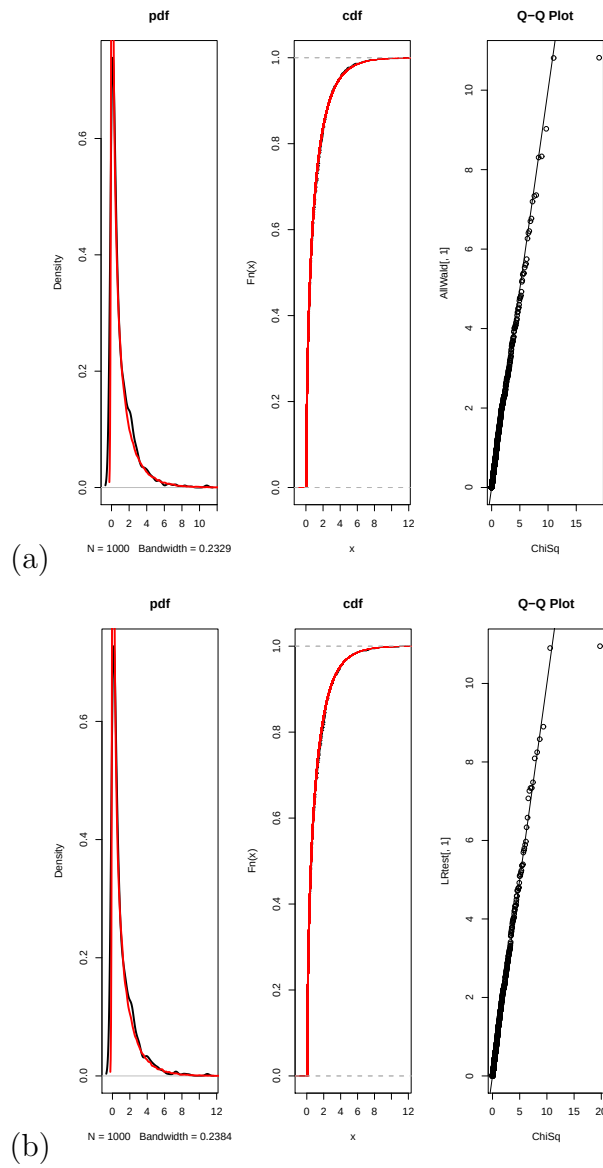


Figure 4.1: The pdf, cdf, and Q-Q plots of the (a) proposed test statistic vs. χ_1^2 distribution and (b) LR test vs. χ_1^2 distribution for the scenario number 10

Additionally, the appropriateness of the chi-square approximation of the proposed test statistic was investigated using some graphical methods such as the pdf, cdf and Q-Q plots. All of the graphs are not given here to save space; only the graphical results related to scenario number 10 are given in Figure 4.1.

In Figure 4.1, the red lines represent the χ_1^2 distribution and the black lines represent the proposed Wald test statistic (in Figure 4.1 part (a)) and LR test statistic values (in Figure 4.1 part (b)) which are calculated from 1,000 replications. As seen from that figure, the red and the black lines overlap almost perfectly. Also in the Q-Q plot, the values of the proposed Wald test statistic lie on the 45° line. Therefore, we conclude that the proposed Wald test statistic has a χ_1^2 distribution under the H_0 for the 2×3 table.

4.2 Results of Simulation Study 2

In Simulation Study 1 under H_0 , the distribution of the proposed Wald test statistic was found to be the χ_1^2 distribution. In order to generalize the distributional results, we re-conducted the simulation study for $R \times C$ tables. The results of the simulation study for 36 scenarios are given in Table 4.2.

In Table 4.2, the mean values of the omnibus LR and the proposed test statistics that were computed from 1,000 random samples are given in columns 3 and 4. The proposed Wald test statistic and the LR test statistic should be distributed as chi-square with a mean of $J(K - 2)$ (note that, for a $R \times C$ table, the number of independent variables is $J = R - 1$ and $K = C$) degrees of freedom. As seen in these columns, the mean values of the omnibus version of the proposed test statistic are approximately $J(K - 2)$. Again, the mean values of the omnibus version of the proposed test statistic are smaller than the likelihood ratio test statistics.

The fifth and sixth columns of Table 4.2 give the percentage of rejections of the proposed test statistic and the LR test statistic under the $\chi_{J(K-2)}^2$ distribution. As seen in these columns, type I error rates of both test statistics are around the 0.05 level, while the proposed test tends to be more conservative than the likelihood ratio test statistic.

Table 4.2: Results of the simulation study 2.

Scenario Number	$R \times C$	Mean Values of Omn.		Type I error of Omn.		Mean Values of			Type I error of		
		LR test	Prop. test	LR test	Prop. test	Prop. test (Ind.)			Prop. test (Ind.)		
1	3×4	3.968	3.863	0.043	0.035	1.904	1.954		0.039	0.051	
2	3×4	3.986	3.929	0.056	0.049	2.005	1.909		0.064	0.037	
3	4×5	9.270	8.725	0.059	0.031	1.944	1.961	1.942	0.017	0.014	0.007
4	4×5	9.101	8.796	0.047	0.037	1.959	1.975	1.993	0.013	0.020	0.017
5	3×4	4.083	4.004	0.054	0.042	1.948	1.965		0.044	0.041	
6	3×4	4.068	4.032	0.057	0.055	1.999	2.035		0.045	0.059	
7	4×5	9.195	8.634	0.052	0.031	1.920	1.902	1.886	0.014	0.009	0.009
8	4×5	9.291	9.044	0.073	0.057	2.044	2.057	1.947	0.020	0.022	0.016
9	3×4	4.045	3.935	0.063	0.050	1.943	1.966		0.046	0.041	
10	3×4	4.170	4.116	0.058	0.055	2.043	2.054		0.052	0.054	
11	4×5	8.568	7.820	0.032	0.012	1.651	1.656	1.658	0.007	0.003	0.010
12	4×5	9.226	8.636	0.059	0.036	1.943	1.887	1.907	0.012	0.005	0.014
13	3×4	3.921	3.730	0.042	0.031	1.975	1.788		0.048	0.027	
14	3×4	4.113	3.964	0.052	0.045	2.049	1.932		0.053	0.042	
15	4×5	9.044	8.424	0.053	0.032	1.973	1.940	2.026	0.017	0.007	0.018
16	4×5	8.958	8.568	0.044	0.030	1.959	1.876	1.946	0.017	0.015	0.018
17	3×4	3.953	3.891	0.042	0.037	1.879	1.995		0.035	0.047	
18	3×4	4.066	4.028	0.051	0.044	2.074	2.044		0.060	0.056	
19	4×5	8.875	8.543	0.041	0.026	1.974	1.891	1.815	0.015	0.012	0.010
20	4×5	8.999	8.837	0.050	0.041	1.991	2.000	1.985	0.019	0.020	0.015
21	3×4	4.051	3.965	0.060	0.049	1.936	2.018		0.039	0.052	
22	3×4	3.944	3.906	0.045	0.038	1.957	1.924		0.049	0.042	
23	4×5	8.980	8.403	0.051	0.032	1.801	1.840	1.828	0.006	0.005	0.010
24	4×5	9.094	8.790	0.046	0.032	2.017	1.911	1.962	0.016	0.017	0.014
25	3×4	3.894	3.677	0.048	0.036	1.890	1.782		0.032	0.035	
26	3×4	4.106	3.940	0.049	0.037	2.083	1.929		0.069	0.037	
27	4×5	8.828	8.080	0.036	0.015	1.908	1.753	1.896	0.013	0.005	0.014
28	4×5	9.065	8.562	0.054	0.032	1.932	1.905	1.924	0.018	0.011	0.017
29	3×4	4.156	4.059	0.061	0.055	2.095	2.010		0.058	0.052	
30	3×4	4.026	3.983	0.051	0.049	2.003	1.976		0.057	0.047	
31	4×5	9.447	9.052	0.059	0.042	2.020	2.007	1.989	0.012	0.012	0.019
32	4×5	9.046	8.850	0.049	0.040	1.966	1.938	2.056	0.014	0.019	0.024
33	3×4	4.215	4.113	0.063	0.058	2.110	2.090		0.055	0.057	
34	3×4	4.039	3.994	0.053	0.049	2.067	1.938		0.053	0.049	
35	4×5	9.264	8.680	0.058	0.033	2.015	1.898	1.857	0.010	0.007	0.012
36	4×5	8.993	8.701	0.046	0.031	2.012	2.010	1.932	0.020	0.017	0.009

LR test: The Likelihood Ratio Test Statistic

Prop. test: The Proposed Wald Test Statistic

Omn. : Omnibus

Ind. : Individual

The appropriateness of the chi-square approximation of the omnibus version of the proposed test statistic was investigated using pdf, cdf and Q-Q plots. All of the graphs are not given here to save space; only the graphical results related to scenario number 18 and 20 are given in Figure 4.2 and Figure 4.3, respectively.

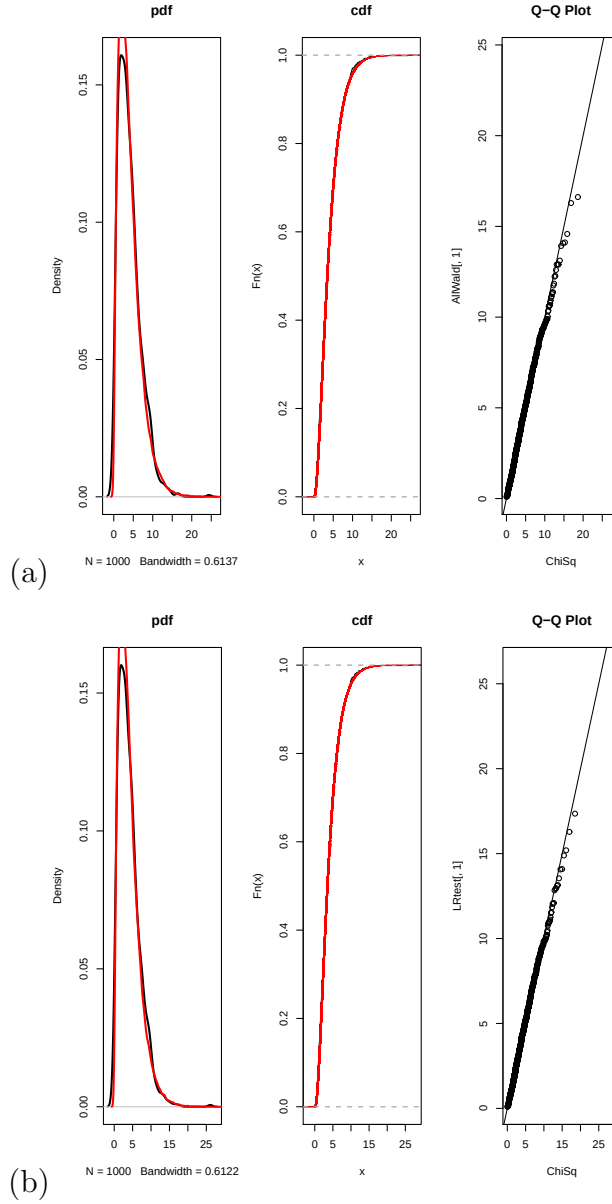


Figure 4.2: The pdf, cdf, and Q-Q plots of the (a) omnibus version of the proposed test statistic vs. $\chi^2_{J(K-2)}$ distribution and (b) LR test vs. $\chi^2_{J(K-2)}$ distribution for the scenario number 18

In Figure 4.2, the red lines represent the χ^2_4 distribution (note that for the 18th scenario the degrees of freedom is equal to 4) and the black lines represent the omnibus version of the proposed Wald test statistic (in Figure 4.2 part (a)) and LR test statistic values (in Figure 4.2 part (b)) which are calculated from 1,000 replications. As seen from that figure, the red and the black lines overlap. Also in the Q-Q plot, the values of the omnibus version of the proposed Wald test statistic lie on the 45° line.

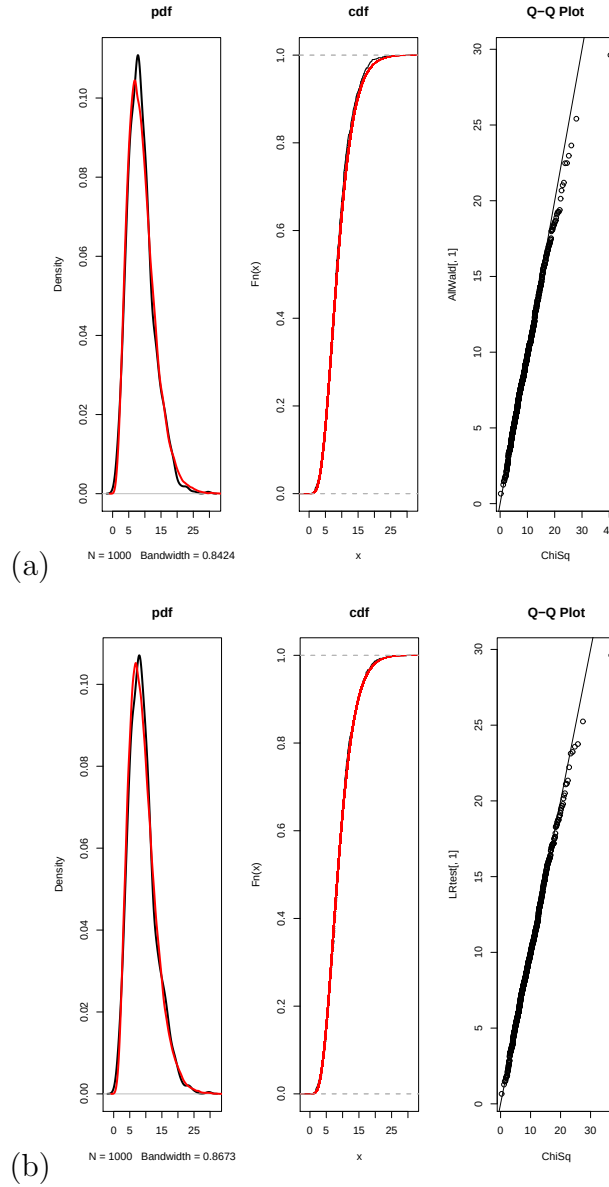


Figure 4.3: The pdf, cdf, and Q-Q plots of the (a) omnibus version of the proposed test statistic vs. $\chi^2_{J(K-2)}$ distribution and (b) LR test vs. $\chi^2_{J(K-2)}$ distribution for the scenario number 20

In Figure 4.3, the red lines represent the χ^2_9 distribution (note that for the 20th scenario the degrees of freedom is equal to 9) and the black lines represent the omnibus version of the proposed Wald test statistic (in Figure 4.3 part (a)) and LR test statistic values (in Figure 4.3 part (b)) which are calculated from 1,000 replications. As seen from that figure, the red and the black lines overlap. In the Q-Q plot, the values of the omnibus version of the proposed Wald test statistic also lie on the 45° line.

Therefore, after examining the results, we concluded that the omnibus version of the proposed Wald test statistic has a chi-square distribution under the H_0 for the $R \times C$ table.

In Table 4.2, the mean values of the individual version of the proposed test statistic that were computed from 1,000 random samples are given through columns 7–9. The Wald test statistic for the individual parameters should be distributed as chi-square with a mean of $K - 2$. As seen in these columns, the mean values of the individual test statistic are approximately $K - 2$ for small degrees of freedom (i.e., for a 3×4 table), whereas they are smaller than $K - 2$ in larger tables (i.e., for a 4×5 table).

Through columns 10–12, the percentage of rejections of proposed individual test statistics under the χ^2_{K-2} distribution are given. As seen in these columns, when the number of degrees of freedom is small (i.e., for a 3×4 table), type I error rates of the proposed test statistic are around the 0.05 level. However, the type I error rates of the individual test statistic are very conservative in the situations where the number of degrees of freedom is large (i.e., for a 4×5 table).

The appropriateness of the chi-square approximation of the individual version of the proposed test statistic was investigated using pdf, cdf and Q-Q plots. All of the graphs are not given here to save space; only the graphical results related to scenario number 18 and 20 are given in Figure 4.4 and Figure 4.5, respectively.

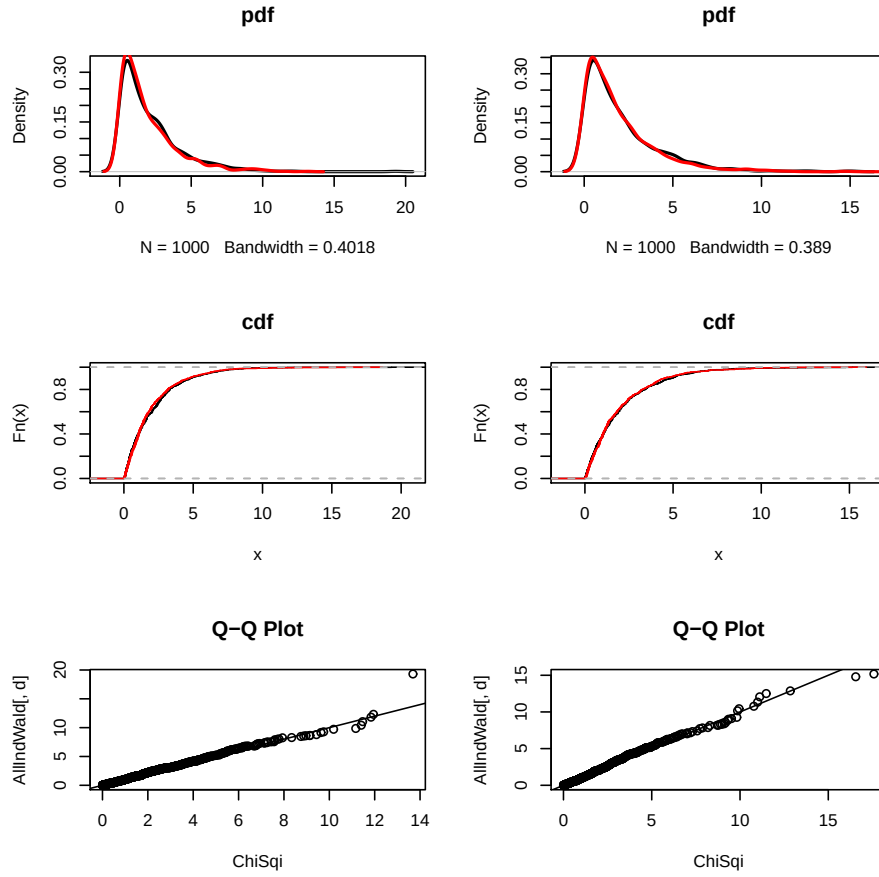


Figure 4.4: The pdf, cdf, and Q-Q plots of the individual version of the proposed test statistic vs. χ^2_{K-2} distribution for the scenario number 18

In Figure 4.4, the red lines represent the χ^2_2 distribution and the black lines represent the individual version of the proposed Wald test statistic which are calculated from 1,000 replications. Note that for the 18th scenario, there are two individual tests (i.e. for variables X_1 and X_2 , respectively) and the degrees of freedom is equal to 2 for the individual Wald tests. As seen from that figure, for both variables, the red and the black lines overlap, and in the Q-Q plots, the values of the individual version of the proposed Wald test statistic lie on the 45° line.

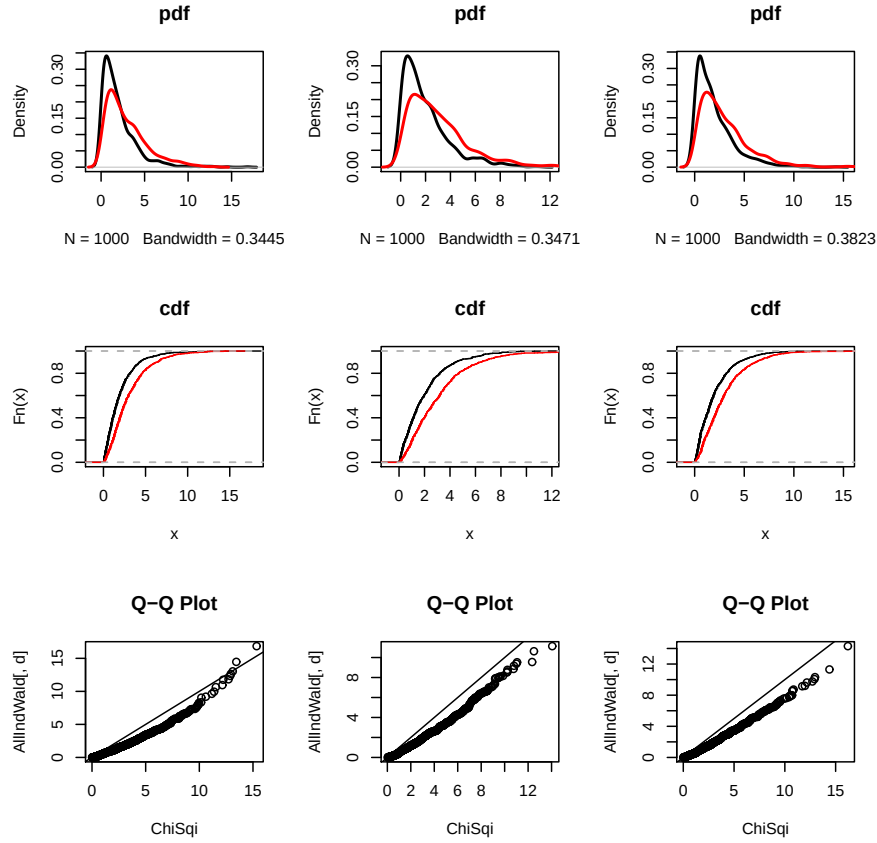


Figure 4.5: The pdf, cdf, and Q-Q plots of the individual version of the proposed test statistic vs. χ^2_{K-2} distribution for the scenario number 20

In Figure 4.5, the red lines represent the χ^2_3 distribution and the black lines represent the individual version of the proposed Wald test statistic which are calculated from 1,000 replications. Note that for the 20th scenario, there are three individual tests (i.e. for variables X_1 , X_2 and X_3 , respectively) and the degrees of freedom is equal to 3 for the individual Wald tests. As seen from that figure, for the individual tests, the red and the black lines do not overlap. Also in the Q-Q plots, the values of the individual version of the proposed Wald test statistic do not lie on the 45° line. Therefore, we can say that the chi-square approximation is labile when the number of independent variables (so the degrees of freedom) increases.

After examining the probability density function, cumulative distribution function and Q-Q plots of the individual version of the proposed test statistic, we concluded that under H_0 , the chi-square approximation is feasible for the individual version of the proposed Wald test statistic, provided that the number of degrees of freedom is not so large for the $R \times C$ table.

5 AN EXAMPLE

An example of the strategy used for parameter estimation in the partial proportional adjacent category model and the proposed test of proportionality in the adjacent category model are given in the next sections.

5.1 Data

The “Attitudes Toward Working Mothers General Survey” (from the Long’s book (21, p.125)) data are used to illustrate the methods used for this study. This data contains information from the 2,293 respondents on the attitudes toward working mothers. The survey is conducted in 1977 and 1989, and the respondents are asked to evaluate the following statement: “A working mother can establish just as warm and secure relationship with her children as a mother who does not work”. Responses are coded in the variable WARM as: 1 = Strongly Disagree (SD); 2 = Disagree (D); 3 = Agree (A); and 4 = Strongly Agree (SA). The goal of this study was to determine whether the survey year (YR89), sex of the respondent (MALE), race of the respondent (WHITE), age of the respondent (AGE) and education year of the respondent (EDU) are associated with the attitudes toward working mothers. The data code sheet is given in Table 5.1.

Table 5.1: Data code sheet for the variables in the attitudes toward working mothers general survey data.

Variable	Description	Codes/Values	Name
1	Response to: "A working mother can establish just as warm and secure relationship with her children as a mother who does not work"	1: Strongly Disagree (SD) 2: Disagree (D) 3: Agree (A) 4: Strongly Agree (SA)	WARM
2	Survey Year	1: 1989 0: 1977 [†]	YR89
3	Sex of the respondent	1: Male 0: Female [†]	MALE
4	Race of the respondent	1: White 0: Non-white [†]	WHITE
5	Age of the respondent	Age in years	AGE
6	Education year of the respondent	Education in years	EDU

[†] Reference category

5.2 An Example of the Proposed Test of Proportionality in the Adjacent Category Model

Suppose that we want to determine the effects of the survey year (YR89), sex of the respondent (MALE), race of the respondent (WHITE), age of the respondent (AGE) and education year of the respondent (EDU) on the attitudes toward working mothers (WARM) using the adjacent category model. Before proceeding with any modelling, we must check the proportionality assumption using the proposed method given in Section 3.1.

The R codes for calculating the proposed Wald test statistic for the attitudes toward working mothers data are given in Appendix C. Table 5.2 gives the results of the LR test and the proposed test of proportionality for this data.

Table 5.2: Results of the LR test and the proposed test of proportionality for the attitudes toward working mothers data.

	Test Statistic	df	p value
Likelihood ratio test	47.969	10	$p < 0.001$
Proposed Wald test (Omnibus)	46.930	10	$p < 0.001$
Proposed Wald test for YR89	10.811	2	0.0045
Proposed Wald test for MALE	24.689	2	$p < 0.001$
Proposed Wald test for WHITE	1.188	2	0.5522
Proposed Wald test for AGE	8.708	2	0.0129
Proposed Wald test for EDU	0.953	2	0.6211

According to both LR and the omnibus version of the proposed test results, at least one factor in the ACM is not satisfying the proportionality assumption ($p < 0.001$ and $p < 0.001$, respectively). When we examine the individual Wald test statistics of these factors, we observe that the survey year (YR89), sex of the respondent (MALE), and age of the respondent (AGE) violate the proportionality assumption. Therefore, we must construct a *partial proportional adjacent category model* for this data.

5.3 An Example of Parameter Estimation in the Partial Proportional Adjacent Category Model

In order to determine the parameter estimations in the PPACM, we utilize the relationships between the adjacent category model, the non-proportional adjacent category model, and the multinomial logistic regression model, which are given in sections 2.2.7 and 2.2.8. Because the adjacent category model and the non-proportional adjacent category model is a constrained version of the multinomial logistic regression model (See Equations 2.23 and 2.30), we start by fitting a multinomial logistic regression in order to get the NPACM, ACM and PPACM results. The R codes for fitting these models are given in Appendix C.

Table 5.3: Estimation results of the multinomial logistic regression model.

Variable	Coefficient	Std. Error	z value	p value
Logit 1 : Disagree				
Intercept	0.4230	0.4287	0.99	0.324
YR89	0.7348	0.1656	4.44	$p < 0.001$
MALE	0.0829	0.1403	0.59	0.554
WHITE	-0.4456	0.2464	-1.81	0.071
AGE	-0.0037	0.0043	-0.86	0.393
EDU	0.0703	0.0230	3.05	0.002
Logit 2 : Agree				
Intercept	1.0891	0.4311	2.53	0.012
YR89	1.0980	0.1637	6.71	$p < 0.001$
MALE	-0.3623	0.1406	-2.58	0.010
WHITE	-0.5321	0.2456	-2.17	0.030
AGE	-0.0246	0.0044	-5.61	$p < 0.001$
EDU	0.1192	0.0236	5.06	$p < 0.001$
Logit 3 : Strongly Agree				
Intercept	0.6679	0.4929	1.36	0.175
YR89	1.1622	0.1811	6.42	$p < 0.001$
MALE	-1.2308	0.1673	-7.36	$p < 0.001$
WHITE	-0.8283	0.2633	-3.15	0.002
AGE	-0.0311	0.0051	-6.08	$p < 0.001$
EDU	0.1590	0.0279	5.71	$p < 0.001$
WARM=Strongly Disagree is the baseline outcome				
Log-likelihood=-2825.204				

Table 5.3 gives the results of the multinomial logistic regression model for the attitudes toward working mothers data. The non-proportional adjacent category model is the constrained version of the MLRM.

Table 5.4 gives the parameter and standard error estimates of the non-proportional adjacent category model, which is directly obtained from the MLRM using the constraints in Equation 2.31. As seen in Table 5.3 and 5.4, these models have equal log-likelihoods.

Table 5.4: Estimation results of the non-proportional adjacent category model.

Variable	Coefficient	Std. Error	z value	p value
Logit 1 : Disagree				
Intercept	0.4230	0.4287	0.99	0.324
YR89	0.7348	0.1656	4.44	$p < 0.001$
MALE	0.0829	0.1403	0.59	0.554
WHITE	-0.4456	0.2464	-1.81	0.071
AGE	-0.0037	0.0043	-0.85	0.393
EDU	0.0703	0.0230	3.05	0.002
Logit 2 : Agree				
Intercept	0.6662	0.3116	2.14	0.033
YR89	0.3632	0.1067	3.40	$p < 0.001$
MALE	-0.4453	0.1041	-4.28	$p < 0.001$
WHITE	-0.0865	0.1616	-0.54	0.593
AGE	-0.0209	0.0033	-6.43	$p < 0.001$
EDU	0.0489	0.0175	2.79	0.005
Logit 3 : Strongly Agree				
Intercept	-0.4212	0.3732	-1.13	0.259
YR89	0.0642	0.1229	0.52	0.601
MALE	-0.8684	0.1312	-6.62	$p < 0.001$
WHITE	-0.2962	0.1707	-1.73	0.083
AGE	-0.0065	0.0040	-1.61	0.108
EDU	0.0398	0.0220	1.81	0.070
Log-likelihood=-2825.204				

In the non-proportional adjacent category model, the first logit estimation results of the coefficients should be the same as the multinomial logistic regression results. From Equation 2.31, the second logit estimates of the NPACM results are found by taking the difference of the logit 2 of the MLRM results from the logit 1, and the third logit results are found by taking the difference of logit 3 of the MLRM results from the logit 2. For example in Table 5.4, the coefficient estimate of the WHITE in the third logit can be found as $-0.8283 - (-0.5321) = -0.2962$.

The adjacent category model is also a constrained version of the multinomial logistic regression model. In order to get the parameter estimates in the adjacent category model, first we fit a constrained multinomial logistic regression model (using MLE methods) with constraint $\beta_k = k\beta$, which is given in Equation 2.24. Table 5.5 gives the maximum likelihood estimation results for the constrained multinomial logistic regression model.

Table 5.5: Estimation results of the constrained multinomial logistic regression model with constraint $\beta_k = k\beta$.

Variable	Coefficient	Std. Error	z value	p value
Logit 1 : Disagree				
Intercept	1.2817	0.1619	7.91	$p < 0.001$
YR89	0.3463	0.0506	6.84	$p < 0.001$
MALE	-0.4354	0.0493	-8.82	$p < 0.001$
WHITE	-0.2307	0.0752	-3.07	0.002
AGE	-0.0127	0.0015	-8.25	$p < 0.001$
EDU	0.0508	0.0083	6.15	$p < 0.001$
Logit 2 : Agree				
Intercept	1.6779	0.2976	5.64	$p < 0.001$
YR89	0.6926	0.1012	6.84	$p < 0.001$
MALE	-0.8708	0.0988	-8.82	$p < 0.001$
WHITE	-0.4615	0.1503	-3.07	0.002
AGE	-0.0253	0.0031	-8.25	$p < 0.001$
EDU	0.1016	0.0165	6.15	$p < 0.001$
Logit 3 : Strongly Agree				
Intercept	1.0325	0.4374	2.36	0.018
YR89	1.0388	0.1519	6.84	$p < 0.001$
MALE	-1.3063	0.1481	-8.82	$p < 0.001$
WHITE	-0.6922	0.2255	-3.07	0.002
AGE	-0.0379	0.0046	-8.25	$p < 0.001$
EDU	0.1524	0.0248	6.15	$p < 0.001$
Log-likelihood=-2849.189				

Note that in Table 5.5, since we applied the constraint $\beta_k = k\beta$, the coefficient estimates in the second logit are twice of the coefficient estimates in the first logit, and the coefficient estimates in the third logit are treble of the coefficient estimates in the first logit. For example, the coefficient estimate for the race of the respondent (WHITE) in the third logit will be equal to $3 \times -0.2307 = -0.6922$. Because the adjacent category model possesses the property that the independent variables have equal effects on the outcome across the adjacent categories, the effects of the independent variables do not vary across logits and they are equal to the coefficients which are given in the first logit of Table 5.5. Such as the co-

efficient estimate of survey year (YR89) and the coefficient estimate of education year of the respondent (EDU) for the adjacent category model will be equal to 0.3463 and 0.0508, respectively.

Table 5.6: Estimation results of the adjacent category model.

Variable	Coefficient	Std. Error	z value	p value
Intercept (Logit 1)	1.2817	0.1619	7.9144	$p < 0.001$
Intercept (Logit 2)	0.3962	0.1526	2.5965	0.009
Intercept (Logit 3)	-0.6454	0.1545	-4.1765	$p < 0.001$
YR89	0.3463	0.0506	6.8420	$p < 0.001$
MALE	-0.4354	0.0494	-8.8203	$p < 0.001$
WHITE	-0.2307	0.0752	-3.0699	0.002
AGE	-0.0127	0.0015	-8.2488	$p < 0.001$
EDU	0.0508	0.0083	6.1529	$p < 0.001$
Log-likelihood=-2849.189				

After fitting the multinomial logistic regression with constraint $\beta_k = k\beta$, and getting the regression coefficient estimates, now we can get the intercept terms for the adjacent category model using the property $\beta_{0k} = \sum_{i=1}^k \alpha_i$. For example, the intercept term for the first logit of adjacent category model (which is 1.2817) will be equal to the intercept of the first logit of the constrained multinomial logistic regression model in Table 5.6. The second intercept term of the adjacent category model is found by taking the difference of intercept of the second logit and the first logit, which is $1.6779 - 1.2817 = 0.3962$. Similarly, the intercept term for the third logit will be $1.0325 - 1.6779 = -0.6454$. Therefore, we get the parameter and standard error estimates of the adjacent category model as given in Table 5.6.

We have found that the survey year (YR89), sex of the respondent (MALE), and age of the respondent (AGE) violate the proportionality assumption. Thus, when we apply the constraints in Equation 2.24 for the race (WHITE) and education year of the respondent while not applying any constraints to the rest of the variables (YR89, MALE, AGE), we get the parameter estimates of partial version of the adjacent category model. The partial proportional adjacent category model results are given in Table 5.7.

Table 5.7: Estimation results of the partial-proportional adjacent category model.

Variable	Coefficient	Std. Error	z value	p value
Intercept (Logit 1)	0.4964	0.2672	1.8580	0.063
Intercept (Logit 2)	0.7454	0.2090	3.5663	$p < 0.001$
Intercept (Logit 3)	-0.6412	0.2242	-2.8597	0.004
WHITE	-0.2304	0.0751	-3.0661	0.002
EDU	0.0518	0.0084	6.1804	$p < 0.001$
Logit 1 : Disagree				
YR89	0.7545	0.1639	4.6025	$p < 0.001$
MALE	0.0786	0.1400	0.5618	0.574
AGE	-0.0050	0.0041	-1.2104	0.226
Logit 2 : Agree				
YR89	0.3585	0.1057	3.3909	0.001
MALE	-0.4424	0.1040	-4.2532	$p < 0.001$
AGE	-0.0206	0.0032	-6.5237	$p < 0.001$
Logit 3 : Strongly Agree				
YR89	0.0551	0.1216	0.4529	0.651
MALE	-0.8776	0.1308	-6.7114	$p < 0.001$
AGE	-0.0061	0.0039	-1.5421	0.123
Log-likelihood=-2826.187				

From Table 5.7 we can see that the race and education year of the respondents have the same effect across the adjacent logits. White respondents tend to agree more by $\exp(-0.2304) = 0.7942$ times (which means they do not agree) when compared to the non-white respondents, while with an increase of 1 year in education, respondents tend to agree more with odds of $\exp(0.0518) = 1.05$. The survey year, sex of the respondents and age of respondents have different effects across the logits. The coefficients of the variables can be interpreted individually; for example, when compared to females, males tend to agree $\exp(0.0786) = 1.08$ times for the first logit, $\exp(-0.4424) = 0.64$ times in the second logit, and $\exp(-0.8776) = 0.41$ times for the third logit. In other words, the degree of agreement with the statement decreases for males.

6 DISCUSSION AND FUTURE WORK

Prior to fitting any of ordinal logistic regression models, it is important to check whether the assumption of proportionality is satisfied by each independent variable. In the proportional odds model, the proportional odds assumption is checked by the Brant's Wald test statistic (26), and it is performed by the standard Wald test statistic (1) for the continuation ratio model. In the adjacent category logistic regression model, the test of proportionality assumption can be performed with the LR test. However, the LR test is an omnibus test that does not show whether the proportionality assumption is violated for all independent variables or only for some. Therefore, before using the ACM, one needs a valid method to test the proportionality assumption both in an omnibus and individual manner.

Following the idea of Brant (26), we adapted a Wald test that can test the proportionality assumption both in an omnibus and individual manner for checking the proportionality assumption in the ACM. The distribution of the proposed test statistic is examined under H_0 with a Monte Carlo simulation study. We examined the appropriateness of chi-square approximation of the proposed test statistic using some graphical methods, and **we showed that the proposed test statistic is asymptotically distributed as a chi-square distribution under the null hypothesis.**

The likelihood ratio test statistic evaluates the difference in deviances, and it provides an omnibus test of proportionality in the adjacent category model. Therefore, we also compared the results of the omnibus version of the proposed test statistic with the LR test results. According to the simulation results, the mean values of the omnibus version of the proposed test statistic are very close to the LR test statistics, and the proposed test tends to be more conservative than the LR test statistic. Agresti (9) states that as $n \rightarrow \infty$, the standard Wald test and the LR test become equivalent. We also observed that as sample size increases, the omnibus version of the proposed test and the LR test become equivalent.

Brant (26) points out that if the number of degrees of freedom is large, then the Wald test becomes conservative and suffers from the power for the proportional odds model. We also observed that the omnibus and the individual versions of the proposed test statistics tend to be more conservative when the

number of degrees of freedom is large for an $R \times C$ table.

The proposed test statistic is an effective way to test the proportionality assumption in the adjacent category model, and it is superior to the likelihood ratio test statistic because it allows both an overall and individual test of proportionality in the ACM. With use of the proposed test statistic, the variables in the ACM that do not obey the proportionality assumption can be determined, and they can be modelled as non-proportional in the partial proportional adjacent category model. Therefore, in this study, we suggest using the proposed test statistic for testing the proportionality assumption in the ACM. The test statistic suggested here has the advantage of requiring only the software for fitting binary logistic regression model plus some basic matrix manipulations. We also provide the R routine for the proposed test statistic in Appendix C.

The partial proportional adjacent category model is superior to the adjacent category model for situations where the assumption of proportionality is not fulfilled for some variables in the model. Additionally, this model is more parsimonious than the non-proportional adjacent category model when some variables violate the proportionality assumption. The parameter estimates and the standard errors in the partial proportional model are feasible, and they can be obtained by fitting a constrained version of the multinomial logistic regression model using the standard maximum likelihood estimation method. The maximum likelihood estimation in the partial proportional adjacent category model can be determined using the VGAM package in R software (33).

The simulation study scenarios are restricted for 2×3 and $R \times C$ tables. We simulated the distribution of the proposed test statistic under H_0 with 20 and 36 scenarios for a 2×3 and a $R \times C$ table, respectively. During data generation, we did not consider including any continuous independent variables in the ACM. Thus, the distributional results that are derived from the simulation study are restricted to only categorical independent variables. Possible violations in the chi-square approximation may occur when we add some continuous independent variables in the ACM. Another study to determine the effect of continuous independent variables on the null distribution of the proposed test statistic will be conducted as possible future work.

BIBLIOGRAPHY

1. Hosmer, D. W. and Lemeshow, S. (2000) *Applied Logistic Regression*, Second Edition, Wiley, New York.
2. Bender, R. and Grouven U. (1998) Using Binary Logistic Regression Models for Ordinal Data with Non-proportional Odds. *Journal of Clinical Epidemiology* 51: 809–816.
3. Norris C. M. et al. (2006) Ordinal regression model and the linear regression model were superior to the logistic regression models. *Journal of Clinical Epidemiology* 59: 448–456.
4. Ananth, C. V. and Kleinbaum, D. G. (1997). Regression models for ordinal responses: A review of methods and applications. *International Journal of Epidemiology*, 26, 1323–1333.
5. Walker S. H. and Duncan D. B. (1967) Estimation of the probability of an event as a function of several independent variables. *Biometrika* 54: 167–79.
6. McCullagh, P. (1980) Regression Models for Ordinal Data (with Discussion). *Journal of the Royal Statistical Society Series B* 42: 109–142.
7. Simon, G. (1974) Alternative analyses for the singly-ordered contingency table, *Journal of the American Statistical Association* 69: 971–976.
8. Goodman, L. A. (1983) The analysis of dependence in cross-classifications having ordered categories, using log-linear models for frequencies and log-linear models for odds. *Biometrika* 39: 149–160.
9. Agresti, A. (2002) *Categorical Data Analysis*, Second Edition, Wiley, New Jersey.
10. Fienberg S. E. (1977) *The Analysis of Cross-classified Categorical Data*. MIT Press: Cambridge, Mass.
11. McCullagh, P. and Nelder, J. A. (1989) *Generalized Linear Models*, Second Edition, Chapman and Hall, London.
12. Anderson J. A. (1984) Regression and ordered categorical variables (with discussion). *Journal of the Royal Statistical Society Series B* 46: 1–30.
13. Engel J. (1988) Polytomous logistic regression *Stat Neerlandica* 42: 233–52.
14. Peterson B. and Harrell F. E. (1990) Partial Proportional Odds Model for Ordinal Response Variables. *Applied Statistics* 2: 205–217.

15. Greenwood C. and Farewell V. (1988) A comparison of regression models for ordinal data in an analysis of transplanted-kidney function. *The Canadian Journal of Statistics*. 16: 325–335.
16. Holtbrugge W. and Schumacher M. (1991) A comparison of regression models for the analysis of ordered categorical data. *Applied Statistics*. 40: 249–259.
17. Harrell F. E. et al. (1998) Development of a clinical prediction model for an ordinal outcome: The World Health Organization Multicentre Study of Clinical Signs and Etiological Agents of Pneumonia, Sepsis and Meningitis in Young Infants. *Statistics in Medicine*. 17: 909–944.
18. Lall R. et al. (2002) A review of ordinal regression models applied on health-related quality of life assessments. *Statistical Methods in Medical Research*. 11: 49–67.
19. Imrey P. B. and Kingman A. (2004) Analysis of Clinical Trials Involving Non-cavitated Caries Lesions. *Journal of Dental Research*. 83: 103–108.
20. Terzi A. et al. (2008) The significance of immunohistochemical expression of Ki-67, p53, p21, and p16 in meningiomas tissue arrays. *Pathology Research and Practice* 204: 305–314.
21. Long S. J. (1997) *Regression Models for Categorical and Limited Dependent Variables*, Sage Publications, Thousand Oaks, CA.
22. Fu, V. K. (1998). Estimating Generalized Ordered Logit Models. *Stata Technical Bulletin*, 44, 27–30. Reprinted in *Stata Technical Bulletin Reprints*, 8: 160–164.
23. Williams, R. (2006) Generalized Ordered Logit/ Partial Proportional Odds Models for Ordinal Dependent Variables. *The Stata Journal* 6(1): 58–82.
24. Armstrong B. G. and Sloan M. (1989) Ordinal Regression Models for Epidemiologic Data. *American Journal of Epidemiology*, 129: 191–204.
25. Cole S. R. and Ananth C. V. (2001) Regression models for unconstrained, partially or fully constrained continuation odds ratios, *International Journal of Epidemiology* 30: 1379–1382.
26. Brant, R. (1990) Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression. *Biometrics* 46: 1771–1178.
27. Dale, J. R. (1986). Global cross-ratio models for bivariate, discrete and ordered responses. *Biometrics*, 42: 909–917.

28. Dos Santos, D. M. and Berridge, D. M. (2000) A continuation ratio random effects model for repeated ordinal responses. *Statistics in Medicine* 19: 3377–3388.
29. Wolfe, R. (1998) Continuation-ratio models for epidemiological data. *Stata Technical Bulletin*, 44: 18–21.
30. Stata Textbook Examples. Applied Logistic Regression, Second Edition, by Hosmer and Lemeshow, Chapter 8: Special Topics. UCLA: Academic Technology Services, Statistical Consulting Group. From <http://www.ats.ucla.edu/stat/stata/examples/alr2/alr2stata8.htm> (accessed November 23, 2010).
31. Jones, B. and Westerland C. (2006). Alternatives to Conventional Practices for Ordinal, Categorical Response Variables. Paper presented at the annual meeting of Midwest Political Science Association, Palmer House Hilton, Chicago, Illinois. From <http://psfaculty.ucdavis.edu/bsjjones/joneswesterland06.pdf> (accessed November 23, 2010).
32. Wolfe, R. and Gould, W. (1998). An Approximate Likelihood-ratio Test for Ordinal Response Models. *Stata Technical Bulletin*, 42: 24–27.
33. Yee, T. W. (2010). The VGAM package for categorical data analysis. *Journal of Statistical Software* 32: 1–34.

A R CODES FOR THE SIMULATION STUDY 1

```
#####
##### PART I- DATA GENERATION #####
#####

library(xlsReadWrite)
library(dummies)
library(VGAM)

nsim<-1000 # Number of iterations

N<-200 # Sample Size in the Table

R<-2 # Number of row levels
C<-3 # Number of column levels

Pr<-matrix(NA,R*C,1)
RC<-matrix(NA,R,C)

AllWald<-matrix(NA,nsim,2)
rejAllWald<-matrix(NA,nsim,1)
AllIndWald<-matrix(NA,nsim,(C-2))

DevianceNPACM<-matrix(NA,nsim,1)
DevianceACM<-matrix(NA,nsim,1)
LRtest<-matrix(NA,nsim,1)

set.seed(987654321)

#Probabilities under H0:

Prprime<-c(0.095,0.382,0.190,0.048,0.190,0.095) # For Odds Ratio=1
#Prprime<-c(0.240,0.240,0.120,0.080,0.160,0.160) # For Odds Ratio=2
#Prprime<-c(0.088,0.147,0.147,0.029,0.147,0.442) # For Odds Ratio=3
#Prprime<-c(0.067,0.100,0.067,0.033,0.200,0.533) # For Odds Ratio=4
#Prprime<-c(0.147,0.147,0.088,0.029,0.147,0.442) # For Odds Ratio=5

RCprime<-Prprime*N # Table under H0

for (t in 1:nsim)
```

```

{

RC<-rmultinom(1, N, Prprime)

RC<-t(matrix(RC,C,R))
Pr<-RC/N
print(RC)

data<-matrix(NA,N,2)
z<-matrix(NA,N,C-1)

#####
# Converting RXC table to long-format data for logistic regression #
#####

row<-0

for (i in 1:R) #Taking values for row effect
{
for (j in 1:C) #Taking values for column effect
{
for (k in 1:RC[i,j]) #Rows
{
row<-row+1

data[row,1]<-i
data[row,2]<-j

#####
# Generating Zk values for the data #
#####

for (m in 2:C)
{
if (data[row,2]==m)
{
z[row,m-1]=1
}
else {
if (data[row,2]==m-1)
{
z[row,m-1]=0
}
else z[row,m-1]=NA
}
}
}

```

```

}

} #End for Rows
} #End for column effect

} #End for row effect

#print(data)
#print(z)

#####
# Binding All Data #
#####

fulldata<-cbind(data,z)
fulldata<-as.data.frame(fulldata)
names(fulldata) <- NULL
b<-c("X","Y")
z <- 1:(C-1)
names(fulldata)<-c(b,z)
print(fulldata)

#write.xls(fulldata,"FullData.xls")

#####
# Likelihood Ratio Test Statistic Calculation #
#####

# Log-likelihood under Non-Proportional NPACM

NPACM<-vglm(Y~factor(X),acat(link="log",parallel=FALSE),fulldata)
DevNPACM<-deviance(NPACM)
DevianceNPACM[t,1]<-DevNPACM

# Verifying BetaHat and VarBetaHat Matrix from the NPACM results

BetasNPACM<-NPACM@coefficients[-(1:C-1)]
BetaVarsNPACM<-vcov(NPACM)
BetaVarsNPACM<-BetaVarsNPACM[-(1:C-1),-(1:C-1)]

# Log-likelihood under ACM

ACM<-vglm(Y~factor(X),acat(link="log",parallel=TRUE),fulldata)
DevACM<-deviance(ACM)
DevianceACM[t,1]<-DevACM

```

```
# Verifying BetaHat and VarBetaHat Matrix from the ACM results
```

```
BetasACM<-ACM@coefficients[-(1:C-1)]
BetaVarsACM<-vcov(ACM)
BetaVarsACM<-BetaVarsACM[-(1:C-1),-(1:C-1)]
```

```
#write.xls(BetasNPACM,"BetasNPACM.xls")
#write.xls(BetaVarsNPACM,"BetaVarsNPACM.xls")
```

```
# CALCULATING LR TEST STATISTICS
```

```
LRtest[t,1]<-DevianceACM[t,1]-DevianceNPACM[t,1]
```

```
#pchisq(deviance(ACM)-deviance(NPACM),df=df.residual(ACM)
#-df.residual(NPACM),lower.tail=FALSE)
```

```
#####
##### PART II- WALD TEST STATS #####
#####
```

```
#Matrix Declarations
```

```
BetaHat<-matrix(NA,(C-1),R)
Pik<-matrix(NA,length(fulldata$X),(C-1))
Combined.Cov.Matrix<-matrix(NA,(R-1)*(C-1),(R-1)*(C-1))
BetaStar<-matrix(NA,(R-1)*(C-1),1)
CovBetaHats<-matrix(NA,(R-1),(R-1)*factorial(C-1)/
(factorial(2)*factorial((C-1)-2)))
```

```
#####
# (C-1) Logistic Regression Results #
#####
```

```
n<-0
s<-1
```

```
ns<-0
for (m in 3:(2+C-1)) # for (C-1) Binary Logistic Regression
{
GLMfits<-glm(fulldata[col(fulldata)==m]~factor(fulldata$X),
data=fulldata,family=(binomial(link="logit")))
GLM<-summary(GLMfits)
print(GLM)
```

```

n<-n+1

BetaHat[n,<-GLM$coefficients[,1] #Beta Hats
#print(BetaHat)
VarBetaHat<-GLM$cov.unscaled[-1,-1] #Variance-Covariance of Beta Hats
#print(VarBetaHat)
Combined.Cov.Matrix[s:(s+(R-2)),s:(s+(R-2))]<-VarBetaHat
#Constructing the diagonal first
s<-s+(R-1)

XB0<-rep(1,length(fulldata$X))
XBetas<-dummy(fulldata$X)
XBetas<-XBetas[,-1]
XBetas<-cbind(XB0,XBetas) #X Matrix

# Positive probabilities for (C-1) logits

Pik[,n]<-exp(BetaHat[n,]%*%t(XBetas))/(1+exp(BetaHat[n,]%*%t(XBetas)))

}

print(Pik)

#####
# Reshaping Pik values for the data #
#####

for (row in 1:N)
{
  for (m in 2:C)
  {
    if (data[row,2]==m)
    {
      Pik[row,m-1]=Pik[row,m-1]
    }
    else {
      if (data[row,2]==m-1)
      {
        Pik[row,m-1]=Pik[row,m-1]
      }
      else Pik[row,m-1]=NA
    }
  }
}

```

```

print(Pik)

#####
# Covariance matrix of Beta's (k<l) from (C-1) logits #
#####

rows<-seq(1,((R-1)*(C-1)),by=(R-1))
cols<-seq(1,((R-1)*(C-1)),by=(R-1))

s<-1
for (k in 1:(C-2))
{
  for (l in (k+1):(C-1))
  {

A<-Pik[,k]-(Pik[,k]*Pik[,k])
B<-Pik[,l]-(Pik[,k]*Pik[,l])
G<-Pik[,l]-(Pik[,l]*Pik[,l])

CovBetaHat1<-solve(t(XBetas[-which(is.na(A)),])%%diag(A[!is.na(A)])
%%XBetas[-which(is.na(A)),])%%(t(XBetas[-which(is.na(B)),])
%%diag(B[!is.na(B)])
%%XBetas[-which(is.na(B)),])%%solve(t(XBetas[-which(is.na(G)),])
%%diag(G[!is.na(G)])
%%XBetas[-which(is.na(G)),])
CovBetaHat<- -CovBetaHat1[-1,-1]
#Eliminating the Beta zeros and giving minus signs
print(CovBetaHat)
CovBetaHats[,s:(s+(R-2))]<-CovBetaHat
s<-s+(R-1)
}
}
#print(CovBetaHats)

#write.xls(CovBetaHats,"Cov1.xls")

#####
# Gathering all variance and covariances in a Combined Matrix #
#####

s<-1
for (m in rows)
{
  for (n in cols[(rows[]>m)])

```

```

{
#print(m)
#print(n)

Combined.Cov.Matrix[m:(m+(R-2)),n:(n+(R-2))]<-CovBetaHats[,s:(s+(R-2))]
Combined.Cov.Matrix[n:(n+(R-2)),m:(m+(R-2))]<-CovBetaHats[,s:(s+(R-2))]
s<-s+(R-1)
}
}
print(Combined.Cov.Matrix)

#write.xls(Combined.Cov.Matrix,"Cov.xls")

#####
# Reshaping the Beta Hat Matrix to suit the matrix calculation (BetaStar) #
#####

BetaHat<-BetaHat[,-1]

if (R==2) BetaStar<-BetaHat else
{
BetaStar<-matrix(NA,1,(C-1)*(R-1))
s<-1
for (i in 1:(C-1))
{
BetaStar[1,s:(s+(R-2))]<-BetaHat[i,]
s<-s+(R-1)
}
BetaStar<-t(BetaStar)
}
print(BetaStar)

#####
# Constructing the D Matrix #
#####

ones<-diag(1,(R-1),(R-1))
minus<-diag(-1,(R-1),(R-1))
zeros<-matrix(0,(R-1),(R-1))

D<-matrix(NA,((C-2)*(R-1)),(C-1)*(R-1))
#print(D)

```

```

#First replacing the minus

r<-1
c<-r+(R-1)
for (m in 1:(C-2))
{

D[r:(r+(R-2)),c:(c+(R-2))]<-minus
r<-r+(R-1)
c<-c+(R-1)
}

#print(D)

# Replacing ones

r<-1
c<-1
for (m in 1:(C-2))
{
D[r:(r+(R-2)),c:(c+(R-2))]<-ones
r<-r+(R-1)
}

#print(D)

# Replacing zeros

D[is.na(D)]<-0
print(D)

#####
# Omnibus Wald test statistic #
#####

Wald<-t(D%*%BetaStar)%*%solve(D%*%Combined.Cov.Matrix%*%t(D))%*%(D%*%BetaStar)
AllWald[t,1]<-Wald

# If negative Wald values occur
#if (Wald>=0) AllWald[t,1]<-Wald else AllWald[t,1]<-NA

dfWald<-((C-2)*(R-1))
AllWald[t,2]<-dfWald

```

```

print(AllWald)

} # End for number of simulations

MeanLRtest<-colMeans(LRtest)
MeanWald<-colMeans(AllWald)

#####
# Finding the parameters of the Chi-Square Distribution #
#####

#####
# For Omnibus Wald #
#####

M<-mean(AllWald[,1])
V<-var(AllWald[,1])

critical<-qchisq(0.95,dfWald)
#Critical value for the chi-square distribution

#####
# Critical Value and the number of rejections #
#   for the chi square distribution           #
#####

#####
# For Omnibus Wald #
#####

Reject<-matrix(NA,nsim,1)
RejectLR<-matrix(NA,nsim,1)
Sorted<-sort(AllWald[,1])
#print(Sorted[nsim*0.95])
#print(critical)

for (t in 1:nsim)
{
  if (is.na(AllWald[t,1])) Reject[t,1]=NA
  else if (AllWald[t,1]>=critical) Reject[t,1]<-1
  else Reject[t,1]=0
  if (LRtest[t,1]>=critical) RejectLR[t,1]<-1 else RejectLR[t,1]=0
}

RejectLRtest<-mean(RejectLR[,1])

```

```

RejectWald<-mean(Reject[,1])

MeanLRtest
MeanWald

RejectLRtest
RejectWald

Sim.Results<-c(MeanLRtest,MeanWald[1],RejectLRtest,RejectWald)
print(Sim.Results,digits =5)

#####
# Graphical check of the chi-squared distribution assumption #
#####

#####
# For Omnibus Wald #
#####

par(mfrow=c(1,3))

#####
# Density check #
#####

ChiSq<-rchisq(nsim*100,dfWald)
plot(density(AllWald[,1]),main="pdf", col = "black",lwd=2)
lines(density(ChiSq), col = "red",lwd=2)

#####
# CDF check #
#####

plot(ecdf(AllWald[,1]),main="cdf",col = "black",lwd=1)
lines(ecdf(ChiSq),col = "red")

#####
# QQ plot #
#####

qqplot(ChiSq,AllWald[,1],main="Q-Q Plot")
abline(0,1)

#####

```

```

# Graphical check of chi-squared distribution assumption of the LR test #
#####

#####
# For Omnibus Wald #
#####

par(mfrow=c(1,3))

#####
# Density check #
#####

ChiSq<-rchisq(nsim*100,dfWald)
plot(density(LRtest[,1]),main="pdf", col = "black",lwd=2)
lines(density(ChiSq), col = "red",lwd=2)

#####
# CDF check #
#####

plot(ecdf(LRtest[,1]),main="cdf",col = "black",lwd=1)
lines(ecdf(ChiSq),col = "red")

#####
# QQ plot #
#####

qqplot(ChiSq,LRtest[,1],main="Q-Q Plot")
abline(0,1)

```

B R CODES FOR THE SIMULATION STUDY 2

```
#####
##### PART I- DATA GENERATION #####
#####

library(xlsReadWrite)
library(dummies)
library(VGAM)

nsim<-1000 # Number of iterations

Ns<-c(500,1000) # Sample Size in the Table

beta<-log(1.3) # Common Odds Ratio in the Table

R<-4 # Number of row levels
C<-5 # Number of column levels

mu<-1 # General effect
sigma<-1 # Deviation for the general effect

muX<-1 # Row (X) effect
sigmaX<-1 # Deviation for the row (X) effect

muY<-1 # Column (Y) effect
sigmaY<-1 # Deviation for the column (Y) effect

iv<-seq(from = 1, to = R, by =1) #Row levels
jv<-seq(from = 1, to = C, by =1) #Column levels

logMu<-matrix(NA,R,C)
RCprime<-matrix(NA,R,C)
Prprime<-matrix(NA,R,C)
Pr<-matrix(NA,R*C,1)
RC<-matrix(NA,R,C)
lambdaX<-matrix(NA,1,R)
lambdaY<-matrix(NA,1,C)

AllWald<-matrix(NA,nsim,2)
```

```

rejAllWald<-matrix(NA,nsim,1)
AllIndWald<-matrix(NA,nsim,(C-2))

DevianceNPACM<-matrix(NA,nsim,1)
DevianceACM<-matrix(NA,nsim,1)
LRtest<-matrix(NA,nsim,1)

Sim.Results<-matrix(NA,length(Ns),(4+2*(R-1)) )
set.seed(987654321)

lambda<-rnorm(1,mu,sigma) #Generating general effect Lambda
print(lambda)
lambdaX<-rnorm(R,muX,sigmaX) #Generating row effects
lambdaX[R]<-0
print(lambdaX)
lambdaY<-rnorm(C,muY,sigmaY) #Generating column effects
lambdaY[C]<-0
print(lambdaY)

for (i in 1:R) #Taking values for row effect
{
  for (j in 1:C) #Taking values for column effect
  {

logMu[i,j]<-lambda+lambdaX[i]+lambdaY[j]+(beta*i*j)
#Creating log-frequencies in RxC tables

RCprime[i,j]<-exp(logMu[i,j]) #Creating frequencies of RxC table

  } #End for column effect

} #End for row effect

print(RCprime)
Prprime<-RCprime/sum(RCprime)
print(Prprime)

m<-1
for (i in 1:R) #Taking values for row effect
{
  for (j in 1:C) #Taking values for column effect
  {

Pr[m,1]<-Prprime[i,j] #Multinomial Probabilities

```

```

m<-m+1
}
}

for (q in 1:length(Ns))

{

N<-Ns[q]

for (t in 1:nsim)
{

v<-rmultinom(1, N, Pr)
RC<-t(matrix(v,C,R))
print(RC)

if (any(RC==0)) RC<-ceiling(RC+0.5)

data<-matrix(NA,sum(RC),2)
z<-matrix(NA,sum(RC),C-1)

#####
# Converting RXC table to long-format data for logistic regression #
#####

row<-0

for (i in 1:R) #Taking values for row effect
{
for (j in 1:C) #Taking values for column effect
{
for (k in 1:RC[i,j]) #Rows
{
row<-row+1

data[row,1]<-i
data[row,2]<-j

#####
# Generating Zk values for the data #
#####

for (m in 2:C)
{

```

```

if (data[row,2]==m)
{
z[row,m-1]=1
}
else {
if (data[row,2]==m-1)
{
z[row,m-1]=0
}
else z[row,m-1]=NA
}
}

} #End for Rows
} #End for column effect

} #End for row effect

#print(data)
#print(z)

#####
# Binding All Data #
#####

fulldata<-cbind(data,z)
fulldata<-as.data.frame(fulldata)
names(fulldata) <- NULL
b<-c("X","Y")
z <- 1:(C-1)
names(fulldata)<-c(b,z)
print(fulldata)

#write.xls(fulldata,"FullData.xls")

#####
# Likelihood Ratio Test Statistic Calculation #
#####

# Log-likelihood under Non-Proportional NPACM

NPACM<-vglm(fulldata$Y~factor(X),acat(link="log",parallel=FALSE),
data=fulldata)
DevNPACM<-deviance(NPACM)

```

```

DevianceNPACM[t,1]<-DevNPACM

# Verifying BetaHat and VarBetaHat Matrix from the NPACM results

BetasNPACM<-NPACM@coefficients[-(1:C-1)]
BetaVarsNPACM<-vcov(NPACM)
BetaVarsNPACM<-BetaVarsNPACM[-(1:C-1),-(1:C-1)]

# Log-likelihood under ACM

ACM<-vglm(Y~factor(X),acat(link="log",parallel=TRUE),
data=fulldata)
DevACM<-deviance(ACM)
DevianceACM[t,1]<-DevACM

# Verifying BetaHat and VarBetaHat Matrix from the ACM results

BetasACM<-ACM@coefficients[-(1:C-1)]
BetaVarsACM<-vcov(ACM)
BetaVarsACM<-BetaVarsACM[-(1:C-1),-(1:C-1)]

#write.xls(BetasNPACM,"BetasNPACM.xls")
#write.xls(BetaVarsNPACM,"BetaVarsNPACM.xls")

# CALCULATING LR TEST STATISTICS

LRtest[t,1]<-DevianceACM[t,1]-DevianceNPACM[t,1]

#pchisq(deviance(ACM)-deviance(NPACM),df=df.residual(ACM)
#-df.residual(NPACM),lower.tail=FALSE)

#####
##### PART II- WALD TEST STATS #####
#####

#Matrix Declarations

BetaHat<-matrix(NA,(C-1),R)
Pik<-matrix(NA,length(fulldata$X),(C-1))
XB<-matrix(NA,length(fulldata$X),(C-1))
Combined.Cov.Matrix<-matrix(NA,(R-1)*(C-1),(R-1)*(C-1))
BetaStar<-matrix(NA,(R-1)*(C-1),1)
CovBetaHats<-matrix(NA,(R-1),(R-1)*factorial(C-1)/
(factorial(2)*factorial((C-1)-2)))

```

```
#####
# (C-1) Logistic Regression Results #
#####

n<-0
s<-1

ns<-0
for (m in 3:(2+C-1)) # for (C-1) Binary Logistic Regression
{
  GLMfits<-glm(fullldata[col(fullldata)==m]~factor(fullldata$X),
  data=fullldata,family=(binomial(link="logit")), na.action="na.omit")
  GLM<-summary(GLMfits)
  print(GLM)
  n<-n+1

  BetaHat[n,<-GLM$coefficients[,1] #Beta Hats
  #print(BetaHat)
  VarBetaHat<-GLM$cov.unscaled[-1,-1] #Variance-Covariance of Beta Hats
  #print(VarBetaHat)
  Combined.Cov.Matrix[s:(s+(R-2)),s:(s+(R-2))]<-VarBetaHat
  #Constructing the diagonal first
  s<-s+(R-1)

  XB0<-rep(1,length(fullldata$X))
  XBetas<-dummy(fullldata$X)
  XBetas<-XBetas[,-1]
  XBetas<-cbind(XB0,XBetas) # The X Matrix

#####
# Positive probabilities Pik's for (C-1) logits #
#####

Pik[,n]<-exp(BetaHat[n,]*%t(XBetas))/(1+exp(BetaHat[n,]*%t(XBetas)))

}
print(Pik)

#####
# Reshaping Pik values for the data #
#####

for (row in 1:N)
{
```

```

for (m in 2:C)
{
  if (data[row,2]==m)
  {
    Pik[row,m-1]=Pik[row,m-1]
  }
  else {
    if (data[row,2]==m-1)
    {
      Pik[row,m-1]=Pik[row,m-1]
    }
    else Pik[row,m-1]=NA
  }
}
}

print(Pik)

#####
# Covariance matrix of Beta's (k<l) from (C-1) logits #
#####

rows<-seq(1,((R-1)*(C-1)),by=(R-1))
cols<-seq(1,((R-1)*(C-1)),by=(R-1))

s<-1
for (k in 1:(C-2))
{
  for (l in (k+1):(C-1))
  {

    A<-Pik[,k]-(Pik[,k]*Pik[,k])
    B<-Pik[,l]-(Pik[,k]*Pik[,l])
    G<-Pik[,l]-(Pik[,l]*Pik[,l])

    CovBetaHat1<-solve(t(XBetas[-which(is.na(A)),])%*%
      diag(A[!is.na(A)]))
    %*%XBetas[-which(is.na(A)),])%*%(t(XBetas[-which(is.na(B)),])
    %*%diag(B[!is.na(B)])%*%XBetas[-which(is.na(B)),])%*%
    solve(t(XBetas[-which(is.na(G)),])
    %*%diag(G[!is.na(G)])%*%XBetas[-which(is.na(G)),])
    CovBetaHat<- -CovBetaHat1[-1,-1]
    #Eliminating the Beta zeros and giving minus signs
    print(CovBetaHat)
  }
}

```

```

CovBetaHats[,s:(s+(R-2))]<-CovBetaHat
s<-s+(R-1)
}
}
#print(CovBetaHats)

#write.xls(CovBetaHats,"Cov1.xls")

#####
# Gathering all variance and covariances in a Combined Matrix #
#####

s<-1
for (m in rows)
{
  for (n in cols[(rows[]>m)])
  {
    #print(m)
    #print(n)

    Combined.Cov.Matrix[m:(m+(R-2)),n:(n+(R-2))]<-CovBetaHats[,s:(s+(R-2))]
    Combined.Cov.Matrix[n:(n+(R-2)),m:(m+(R-2))<-CovBetaHats[,s:(s+(R-2))]
    s<-s+(R-1)
  }
}
print(Combined.Cov.Matrix)

#write.xls(Combined.Cov.Matrix,"Combined.Cov.Matrix.xls")

#####
# Reshaping the Beta Hat Matrix to suit the matrix calculation (BetaStar) #
#####

BetaHat<-BetaHat[, -1]

if (R==2) BetaStar<-BetaHat else
{
  BetaStar<-matrix(NA,1,(C-1)*(R-1))
  s<-1
  for (i in 1:(C-1))
  {
    BetaStar[1,s:(s+(R-2))]<-BetaHat[i,]
    s<-s+(R-1)
  }
}

```

```

}
BetaStar<-t(BetaStar)
}
print(BetaStar)

#####
# Constructing the D Matrix #
#####

ones<-diag(1,(R-1),(R-1))
minus<-diag(-1,(R-1),(R-1))
zeros<-matrix(0,(R-1),(R-1))

D<-matrix(NA,((C-2)*(R-1)),(C-1)*(R-1))
#print(D)

#First replacing the minus

r<-1
c<-r+(R-1)
for (m in 1:(C-2))
{

D[r:(r+(R-2)),c:(c+(R-2))]<-minus
r<-r+(R-1)
c<-c+(R-1)
}

#print(D)

# Replacing ones

r<-1
c<-1
for (m in 1:(C-2))
{
D[r:(r+(R-2)),c:(c+(R-2))]<-ones
r<-r+(R-1)
}

#print(D)

# Replacing zeros

D[is.na(D)]<-0

```

```

print(D)

#####
# Omnibus Wald test statistic #
#####

Wald<-t(D%*%BetaStar)%*%solve(D%*%Combined.Cov.Matrix%*%t(D))
%*%(D%*%BetaStar)
AllWald[t,1]<-Wald

# If negative Wald values occur
#if (Wald>=0) AllWald[t,1]<-Wald else AllWald[t,1]<-NA

dfWald<-((C-2)*(R-1))
AllWald[t,2]<-dfWald
print(AllWald)

#####
# Individual Wald test statistics #
#####

IndWald<-matrix(NA,1,(C-2))
s<-1
p<-0
for (d in 1:(C-2))
{
p<-p+1
newD<-rbind(D[s,],D[s+(R-1),])
IndWald[1,p]<-t(newD%*%BetaStar)%*%solve(newD%*%Combined.Cov.Matrix
%*%t(newD))
%*%(newD%*%BetaStar)
AllIndWald[t,]<-IndWald

#If any negative individual wald values occur
#AllIndWald[t,any(AllIndWald<0)]<-NA
s<-s+(R-2)
dfIndWald<-(R-1)
}

#print(IndWald)
#print(dfIndWald)

```

```

} # End for number of simulations

MeanLRtest<-colMeans(LRtest)
MeanWald<-colMeans(AllWald)
MeanIndWald<-colMeans(AllIndWald)

#####
# Finding the parameters of the Chi-Square Distribution #
#####

#####
# For Omnibus Wald #
#####

M<-mean(AllWald[,1])
V<-var(AllWald[,1])

#If any negative wald values occur
#M<-mean(AllWald[,1],na.rm=TRUE)
#V<-var(AllWald[,1],na.rm=TRUE)

critical<-qchisq(0.95,dfWald) #Critical value for the chi-square distribution

#####
# For Individual Walds #
#####

Mi<-colMeans(AllIndWald)
Vi<- apply(AllIndWald, 2,var)

#If any negative wald values occur
Mi<-colMeans(AllIndWald,na.rm=TRUE)
Vi<- apply(AllIndWald, 2,var,na.rm=TRUE)

criticali<-qchisq(0.95,dfIndWald)
#Critical value for the chi-square distribution

#####
# Critical Value and the number of rejections #
# for the chi square distribution #

```

```
#####

#####
# For Omnibus Wald #
#####

Reject<-matrix(NA,nsim,1)
RejectLR<-matrix(NA,nsim,1)
Sorted<-sort(AllWald[,1])
#print(Sorted[nsim*0.95])
#print(critical)

for (t in 1:nsim)
{
  if (is.na(AllWald[t,1])) Reject[t,1]=NA
  else if (AllWald[t,1]>=critical) Reject[t,1]<-1
  else Reject[t,1]=0
  if (LRtest[t,1]>=critical) RejectLR[t,1]<-1 else RejectLR[t,1]=0
}

RejectLRtest<-mean(RejectLR[,1])
RejectWald<-mean(Reject[,1])

#RejectWald<-mean(Reject[,1],na.rm=TRUE)

#####
# For Individual Walds #
#####

Rejecti<-matrix(NA,nsim,(C-2))
print(criticali)

for (d in 1:(C-2))
{
  for (t in 1:nsim)
  {
    if (is.na(AllIndWald[t,d])) Rejecti[t,d]<-NA
    else if (AllIndWald[t,d]>=criticali) Rejecti[t,d]<-1
    else Rejecti[t,d]<-0
  }
}

RejectIndWald<-colMeans(Rejecti)
#RejectIndWald<-colMeans(Rejecti,na.rm=TRUE)
```

```

MeanLRtest
MeanWald
MeanIndWald

RejectLRtest
RejectWald
RejectIndWald

Sim.Results[q,]<-c(MeanLRtest,MeanWald[1],RejectLRtest,RejectWald,
MeanIndWald,RejectIndWald)
}
print(Sim.Results,digits =5)

#####
# Graphical check of the chi-squared distribution assumption #
#####

#####
# For Omnibus Wald #
#####

par(mfrow=c(1,3))

#####
# Density check #
#####

ChiSq<-rchisq(nsim*100,dfWald)
plot(density(AllWald[,1]),main="pdf",col = "black",lwd=2)
lines(density(ChiSq), col = "red",lwd=2)

#####
# CDF check #
#####

plot(ecdf(AllWald[,1]),main="cdf",lwd=1)
lines(ecdf(ChiSq),col = "red")

#####
# QQ plot #
#####

qqplot(ChiSq,AllWald[,1],main="Q-Q Plot")

```

```

abline(0,1)

#####
# For Individual Walds #
#####

#####
# Density check #
#####

par(mfrow=c(3,(C-2)))

for (d in 1:(C-2))
{
  ChiSqi<-rchisq(nsim,dfIndWald)
  plot(density(AllIndWald[,d]),main="pdf", col = "black",lwd=2)
  lines(density(ChiSqi), col = "red",lwd=2)
}

#####
# CDF check #
#####

for (d in 1:(C-2))
{
  ChiSqi<-rchisq(nsim,dfIndWald)
  plot(ecdf(AllIndWald[,d]),main="cdf",col = "black",lwd=1)
  lines(ecdf(ChiSqi),col = "red")
}

#####
# QQ plot #
#####

for (d in 1:(C-2))
{
  ChiSqi<-rchisq(nsim,dfIndWald)
  qqplot(ChiSqi,AllIndWald[,d],main="Q-Q Plot")
  abline(0,1)
}

#####
# Graphical check of chi-squared distribution assumption of the LR test #
#####

```

```
#####
# For Omnibus Wald #
#####

par(mfrow=c(1,3))

#####
# Density check #
#####

ChiSq<-rchisq(nsim*100,dfWald)
plot(density(LRtest[,1]),main="pdf", col = "black",lwd=2)
lines(density(ChiSq), col = "red",lwd=2)

#####
# CDF check #
#####

plot(ecdf(LRtest[,1]),main="cdf",col = "black",lwd=1)
lines(ecdf(ChiSq),col = "red")

#####
# QQ plot #
#####

qqplot(ChiSq,LRtest[,1],main="Q-Q Plot")
abline(0,1)
```

C R CODES FOR CALCULATING PROPOSED WALD TEST AND FITTING PARTIAL PROPORTIONAL ADJACENT CAT- EGORY MODEL

```
#####
# AN EXAMPLE USING ATTITUDES TOWARD WORKING MOTHERS DATA #
#####

library(VGAM)
library(xtable)
library(xlsReadWrite)

#####
# INPUT DATA #
#####

warm.data<-read.xls("C:\\Users\\Anil\\Documents\\TEZ\\DOKTORA\\THESIS
\\example\\ordwarm.xls", colNames = TRUE, sheet=1, type = "data.frame")
is.data.frame(warm.data)

attach(warm.data)
as.factor(warm)
as.factor(yr89)
as.factor(male)
as.factor(white)

#####
# Maximum Likelihood Estimation for the MLRM #
#####

MLRM<-vglm(warm~yr89+male+white+age+ed,multinomial(parallel=FALSE,
refLevel=1),warm.data)
summary(MLRM)
DevMLRM<-deviance(MLRM)
DevMLRM
BetasMLRM<-MLRM@coefficients[-(1:3)]
BetasMLRM
#xtable(BetasMLRM)
BetaVarsMLRM<-vcov(MLRM)
BetaVarsMLRM
BetaVarsMLRM<-BetaVarsMLRM[-(1:3),-(1:3)]
BetaVarsMLRM
#xtable(BetaVarsMLRM)
```

```
#####
# Maximum Likelihood Estimation for the Non-Proportional ACM #
#####

NPACM<-vglm(warm~yr89+male+white+age+ed,acat(link="loge",
parallel=FALSE),warm.data)
summary(NPACM)
DevNPACM<-deviance(NPACM)
DevNPACM
BetasNPACM<-NPACM@coefficients[-(1:3)]
BetasNPACM
#xtable(BetasNPACM)
BetaVarsNPACM<-vcov(NPACM)
BetaVarsNPACM
BetaVarsNPACM<-BetaVarsNPACM[-(1:3),-(1:3)]
BetaVarsNPACM
#xtable(BetaVarsNPACM)

#####
# Maximum Likelihood Estimation for the ACM #
#####

ACM<-vglm(warm~yr89+male+white+age+ed,acat(link="loge",
parallel=TRUE),warm.data)
summary(ACM)
DevACM<-deviance(ACM)
DevACM
BetasACM<-ACM@coefficients[-(1:3)]
BetasACM
#xtable(BetasACM)
BetaVarsACM<-vcov(ACM)
BetaVarsACM
BetaVarsACM<-BetaVarsACM[-(1:3),-(1:3)]
BetaVarsACM
#xtable(BetaVarsACM)

#####
# LR test for the Models (ACM, NPACM)#
#####

LRtestACM_NPACM<-DevACM-DevNPACM
LRtestACM_NPACM
pACM_NPACM<-pchisq(deviance(ACM)-deviance(NPACM),df=df.residual(ACM)
-df.residual(NPACM),lower.tail=FALSE)
pACM_NPACM
```

```
LRtest<-c(LRtestACM_NPACM,df.residual(ACM)-df.residual(NPACM),pACM_NPACM)
```

```
#####
```

```
# PROPOSED WALD TEST STATISTIC CALCULATIONS #
```

```
#####
```

```
#####
```

```
# Constructing the D Matrix #
```

```
#####
```

```
D<-matrix(NA,10,15)
```

```
yr891<-c(1,-1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
```

```
yr892<-c(1, 0,-1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
```

```
#
```

```
male1<-c(0, 0, 0, 1,-1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
```

```
male2<-c(0, 0, 0, 1, 0,-1, 0, 0, 0, 0, 0, 0, 0, 0, 0)
```

```
#
```

```
white1<-c(0, 0, 0, 0, 0, 0, 0, 1,-1, 0, 0, 0, 0, 0, 0)
```

```
white2<-c(0, 0, 0, 0, 0, 0, 0, 1, 0,-1, 0, 0, 0, 0, 0)
```

```
#
```

```
age1<-c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1,-1, 0, 0, 0)
```

```
age2<-c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0,-1, 0, 0)
```

```
#
```

```
ed1<-c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1,-1, 0)
```

```
ed2<-c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0,-1)
```

```
#
```

```
#prst1<-c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1,-1, 0)
```

```
#prst2<-c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0,-1)
```

```
#
```

```
Dprime<-rbind(yr891,yr892,male1,male2,white1,white2,age1,age2,ed1,ed2)
```

```
D<-as.matrix(Dprime)
```

```
#####
```

```
# Omnibus Wald test statistic #
```

```
#####
```

```
BetaStar<-BetasNPACM
```

```
Combined.Cov.Matrix<-BetaVarsNPACM
```

```
Wald<-t(D%*%BetaStar)%*%solve(D%*%Combined.Cov.Matrix%*%t(D))
```

```
%*%(D%*%BetaStar)
```

```
Waldp<-pchisq(Wald,df=df.residual(ACM)-df.residual(NPACM)
```

```
,lower.tail=FALSE)
```

```
OmnibusWald<-c(Wald,df.residual(ACM)-df.residual(NPACM),Waldp)
```

```
#####
```

```

# Individual Wald Test Statistics #
#####

# Testing yr89

yr891<-c(1,-1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
yr892<-c(1, 0,-1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
D_yr89<-as.matrix(rbind(yr891,yr892))
Wald_yr89<-t(D_yr89%%BetaStar)%%solve(D_yr89%%Combined.Cov.Matrix
%%t(D_yr89))%%(D_yr89%%BetaStar)
Wald_yr89p<-pchisq(Wald_yr89,df=2,lower.tail=FALSE)
Wald_yr89<-c(Wald_yr89,2,Wald_yr89p)

# Testing male

male1<-c(0, 0, 0, 1,-1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
male2<-c(0, 0, 0, 1, 0,-1, 0, 0, 0, 0, 0, 0, 0, 0, 0)
D_male<-as.matrix(rbind(male1,male2))
Wald_male<-t(D_male%%BetaStar)%%solve(D_male%%Combined.Cov.Matrix
%%t(D_male))%%(D_male%%BetaStar)
Wald_malep<-pchisq(Wald_male,df=2,lower.tail=FALSE)
Wald_male<-c(Wald_male,2,Wald_malep)

# Testing white

white1<-c(0, 0, 0, 0, 0, 0, 0, 1,-1, 0, 0, 0, 0, 0, 0)
white2<-c(0, 0, 0, 0, 0, 0, 0, 1, 0,-1, 0, 0, 0, 0, 0)
D_white<-as.matrix(rbind(white1,white2))
Wald_white<-t(D_white%%BetaStar)%%solve(D_white%%Combined.Cov.Matrix
%%t(D_white))%%(D_white%%BetaStar)
Wald_whitep<-pchisq(Wald_white,df=2,lower.tail=FALSE)
Wald_white<-c(Wald_white,2,Wald_whitep)

# Testing age

age1<-c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1,-1, 0, 0, 0)
age2<-c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0,-1, 0, 0)
D_age<-as.matrix(rbind(age1,age2))
Wald_age<-t(D_age%%BetaStar)%%solve(D_age%%Combined.Cov.Matrix%%
t(D_age))%%(D_age%%BetaStar)
Wald_agep<-pchisq(Wald_age,df=2,lower.tail=FALSE)
Wald_age<-c(Wald_age,2,Wald_agep)

# Testing ed

```

```

ed1<-c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1,-1, 0)
ed2<-c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0,-1)
D_ed<-as.matrix(rbind(ed1,ed2))
Wald_ed<-t(D_ed%*%BetaStar)%*%solve(D_ed%*%Combined.Cov.Matrix%*%
t(D_ed))%*%(D_ed%*%BetaStar)
Wald_edp<-pchisq(Wald_ed,df=2,lower.tail=FALSE)
Wald_ed<-c(Wald_ed,2,Wald_edp)

TEST_RESULTS<-rbind(LRtest,OmnibusWald,Wald_yr89,Wald_male,
Wald_white,Wald_age,Wald_ed)
TEST_RESULTS
xtable(TEST_RESULTS)

#####
# Maximum Likelihood Estimation for the PPACM #
#####

# In the individual Wald tests, we have found that
# YR89, MALE and AGE are not satisfying the
# proportionality assumption. So we proceed with
# the PPACM as follows:

PPACM<-vglm(warm~yr89+male+white+age+ed,acat(link="loge",
parallel=FALSE~yr89+male+age),warm.data)
summary(PPACM)
DevPPACM<-deviance(PPACM)
DevPPACM
BetasPPACM<-PPACM@coefficients[-(1:3)]
BetasPPACM
#xtable(BetasPPACM)
BetaVarsPPACM<-vcov(PPACM)
BetaVarsPPACM
BetaVarsPPACM<-BetaVarsPPACM[-(1:3),-(1:3)]
BetaVarsPPACM
#xtable(BetaVarsPPACM)

```