1994

Parametric and semi-parametric models for the analysis of proportions in presence of over/under dispersion.

Ali Salim. Islam
University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation
https://scholar.uwindsor.ca/etd/2994

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.
NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.
PARAMETRIC AND SEMI-PARAMETRIC MODELS FOR THE ANALYSIS OF PROPORTIONS IN PRESENCE OF OVER/UNDER DISPERSION

BY

ALI SALIM ISLAM

A Dissertation submitted to the Faculty of Graduate Studies and Research through the Department of Mathematics and Statistics in Partial Fulfillment of the requirements for the Degree of Doctor of Philosophy at the University of Windsor

Windsor, Ontario, Canada
1994
The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L’auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L’auteur conserve la propriété du droit d’auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-93276-7
<table>
<thead>
<tr>
<th>SUBJECT TERMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject Categories</td>
</tr>
<tr>
<td>THE HUMANITIES AND SOCIAL SCIENCES</td>
</tr>
<tr>
<td>COMMUNICATIONS AND THE ARTS</td>
</tr>
<tr>
<td>Architecture</td>
</tr>
<tr>
<td>Art History</td>
</tr>
<tr>
<td>Cinema</td>
</tr>
<tr>
<td>Dance</td>
</tr>
<tr>
<td>Fine Arts</td>
</tr>
<tr>
<td>Information Science</td>
</tr>
<tr>
<td>Journalism</td>
</tr>
<tr>
<td>Library Science</td>
</tr>
<tr>
<td>Mass Communications</td>
</tr>
<tr>
<td>Music</td>
</tr>
<tr>
<td>Speech Communication</td>
</tr>
<tr>
<td>Theater</td>
</tr>
<tr>
<td>EDUCATION</td>
</tr>
<tr>
<td>General</td>
</tr>
<tr>
<td>Administration</td>
</tr>
<tr>
<td>Adult and Continuing</td>
</tr>
<tr>
<td>Agricultural</td>
</tr>
<tr>
<td>Art</td>
</tr>
<tr>
<td>Bilingual and Multicultural</td>
</tr>
<tr>
<td>Business</td>
</tr>
<tr>
<td>Community College</td>
</tr>
<tr>
<td>Curriculum and Instruction</td>
</tr>
<tr>
<td>Early Childhood</td>
</tr>
<tr>
<td>Elementary</td>
</tr>
<tr>
<td>Finance</td>
</tr>
<tr>
<td>Guidance and Counseling</td>
</tr>
<tr>
<td>Health</td>
</tr>
<tr>
<td>Higher Education</td>
</tr>
<tr>
<td>History</td>
</tr>
<tr>
<td>Home Economics</td>
</tr>
<tr>
<td>Industrial Arts</td>
</tr>
<tr>
<td>Language and Literature</td>
</tr>
<tr>
<td>Mathematics</td>
</tr>
<tr>
<td>Music</td>
</tr>
<tr>
<td>Philosophy</td>
</tr>
<tr>
<td>Physical</td>
</tr>
<tr>
<td>Psychology</td>
</tr>
<tr>
<td>Religion</td>
</tr>
<tr>
<td>Religious Studies</td>
</tr>
<tr>
<td>Sciences</td>
</tr>
<tr>
<td>Sociology</td>
</tr>
<tr>
<td>Teacher Education</td>
</tr>
<tr>
<td>Technology</td>
</tr>
<tr>
<td>Tests and Measurements</td>
</tr>
<tr>
<td>Vocational</td>
</tr>
<tr>
<td>LANGUAGE, LITERATURE AND LINGUISTICS</td>
</tr>
<tr>
<td>Language</td>
</tr>
<tr>
<td>Literature</td>
</tr>
<tr>
<td>Linguistics</td>
</tr>
<tr>
<td>Modern</td>
</tr>
<tr>
<td>Medieval</td>
</tr>
<tr>
<td>African</td>
</tr>
<tr>
<td>American</td>
</tr>
<tr>
<td>Asian</td>
</tr>
<tr>
<td>Canadian (English)</td>
</tr>
<tr>
<td>Canadian (French)</td>
</tr>
<tr>
<td>English</td>
</tr>
<tr>
<td>German</td>
</tr>
<tr>
<td>Latin American</td>
</tr>
<tr>
<td>Middle Eastern</td>
</tr>
<tr>
<td>Romance</td>
</tr>
<tr>
<td>Slavic and East European</td>
</tr>
<tr>
<td>SOCIETY</td>
</tr>
<tr>
<td>American Studies</td>
</tr>
<tr>
<td>Anthropology</td>
</tr>
<tr>
<td>Center for Public Affairs</td>
</tr>
<tr>
<td>Business Administration</td>
</tr>
<tr>
<td>Economics</td>
</tr>
<tr>
<td>Sociology</td>
</tr>
<tr>
<td>American</td>
</tr>
<tr>
<td>Business</td>
</tr>
<tr>
<td>Economics</td>
</tr>
<tr>
<td>Finance</td>
</tr>
<tr>
<td>History</td>
</tr>
<tr>
<td>Theory</td>
</tr>
<tr>
<td>Public Policy</td>
</tr>
<tr>
<td>Geography</td>
</tr>
<tr>
<td>Geology</td>
</tr>
<tr>
<td>History</td>
</tr>
<tr>
<td>THE SCIENCES AND ENGINEERING</td>
</tr>
<tr>
<td>BIOPHYSICAL SCIENCES</td>
</tr>
<tr>
<td>Agricultural</td>
</tr>
<tr>
<td>Agronomy</td>
</tr>
<tr>
<td>Animal Physiology</td>
</tr>
<tr>
<td>Animal Behavior</td>
</tr>
<tr>
<td>Animal Nutrition</td>
</tr>
<tr>
<td>Animal Pathology</td>
</tr>
<tr>
<td>Animal Temperature</td>
</tr>
<tr>
<td>Anatomy</td>
</tr>
<tr>
<td>Plant Pathology</td>
</tr>
<tr>
<td>Plant Physiology</td>
</tr>
<tr>
<td>Soil</td>
</tr>
<tr>
<td>ecology</td>
</tr>
<tr>
<td>Geology</td>
</tr>
<tr>
<td>Geophysics</td>
</tr>
<tr>
<td>Hydrology</td>
</tr>
<tr>
<td>Mineralogy</td>
</tr>
<tr>
<td>Paleontology</td>
</tr>
<tr>
<td>Paleobotany</td>
</tr>
<tr>
<td>Paleoclimates</td>
</tr>
<tr>
<td>Paleomagnetism</td>
</tr>
<tr>
<td>Paleontology</td>
</tr>
<tr>
<td>Paleontological</td>
</tr>
<tr>
<td>Paleontology</td>
</tr>
<tr>
<td>Physical Geography</td>
</tr>
<tr>
<td>Physical Oceanography</td>
</tr>
<tr>
<td>HEALTH AND ENVIRONMENTAL SCIENCES</td>
</tr>
<tr>
<td>Environmental Sciences</td>
</tr>
<tr>
<td>Health Sciences</td>
</tr>
<tr>
<td>Audiology</td>
</tr>
<tr>
<td>Chemistry</td>
</tr>
<tr>
<td>Dentistry</td>
</tr>
<tr>
<td>Education</td>
</tr>
<tr>
<td>Hospital Administration</td>
</tr>
<tr>
<td>Human Development</td>
</tr>
<tr>
<td>Immunology</td>
</tr>
<tr>
<td>Medicine and Surgery</td>
</tr>
<tr>
<td>Mental Health</td>
</tr>
<tr>
<td>Nursing</td>
</tr>
<tr>
<td>Nutrition</td>
</tr>
<tr>
<td>Obstetrics and Gynecology</td>
</tr>
<tr>
<td>Occupational Health &amp; Therapy</td>
</tr>
<tr>
<td>Ophthalmology</td>
</tr>
<tr>
<td>Pathology</td>
</tr>
<tr>
<td>Pharmacology</td>
</tr>
<tr>
<td>Physics</td>
</tr>
<tr>
<td>Physical Therapy</td>
</tr>
<tr>
<td>Public Health</td>
</tr>
<tr>
<td>Radiology</td>
</tr>
<tr>
<td>Speech Pathology</td>
</tr>
<tr>
<td>Engineering</td>
</tr>
<tr>
<td>Aerospace</td>
</tr>
<tr>
<td>Agricultural</td>
</tr>
<tr>
<td>Animal</td>
</tr>
<tr>
<td>Biomedical</td>
</tr>
<tr>
<td>Chemical</td>
</tr>
<tr>
<td>Clinical</td>
</tr>
<tr>
<td>Electronics and Electrical</td>
</tr>
<tr>
<td>Heat and Thermodynamics</td>
</tr>
<tr>
<td>Hydrology</td>
</tr>
<tr>
<td>Industrial</td>
</tr>
<tr>
<td>Material Science</td>
</tr>
<tr>
<td>Mechanical</td>
</tr>
<tr>
<td>Metallurgy</td>
</tr>
<tr>
<td>Mining</td>
</tr>
<tr>
<td>Nuclear</td>
</tr>
<tr>
<td>Parkway</td>
</tr>
<tr>
<td>Petroleum</td>
</tr>
<tr>
<td>Pharmacy</td>
</tr>
<tr>
<td>System</td>
</tr>
<tr>
<td>Geology</td>
</tr>
<tr>
<td>Operations Research</td>
</tr>
<tr>
<td>Plastics Technology</td>
</tr>
<tr>
<td>Textile Technology</td>
</tr>
<tr>
<td>PSYCHOLOGY</td>
</tr>
<tr>
<td>General</td>
</tr>
<tr>
<td>Behavioral</td>
</tr>
<tr>
<td>Clinical</td>
</tr>
<tr>
<td>Developmental</td>
</tr>
<tr>
<td>Experimental</td>
</tr>
<tr>
<td>Industrial</td>
</tr>
<tr>
<td>Personality</td>
</tr>
<tr>
<td>Pharmacological</td>
</tr>
<tr>
<td>Psychobiology</td>
</tr>
<tr>
<td>Psychometrics</td>
</tr>
<tr>
<td>Social</td>
</tr>
</tbody>
</table>
ABSTRACT

Data in the form of proportions arise in toxicology (Weil, 1970; Williams, 1975) and other similar fields (Crowder, 1978; Otake and Prentice, 1984). These proportions often exhibit variation greater than predicted by a simple binomial model. Several parametric models such as the beta-binomial (BB) (Skellam, 1948), the correlated binomial (Kuper and Haseman, 1978) and the additive and multiplicative binomial models (Altham, 1974) are available for analysing binomial data with over dispersion. Of these the correlated binomial and the additive binomial models are identical. The superiority of the beta-binomial model for the analysis of proportions has been shown by many authors (Paul, 1982; Pack, 1986).

The joint estimation of the mean and the dispersion or the intraclass correlation parameters is important in the over/under dispersed binomial data. The computation of the maximum likelihood estimates is quite intensive and not robust to variance misspecification. We consider several semi-parametric models as an alternative approach recently developed in the context of correlated binary data, which require assumption on the form of only the mean and variance. We study large and small sample efficiency of the mean and the intraclass correlation parameters.

An important problem is to compare proportions of a certain characteristic in several groups. A common test in these type of studies is to compare the proportion in a control group with that in a treatment group. A number of parametric and non-parametric procedures are available for testing homogeneity of proportions in the presence of over dispersion. Of these, the likelihood ratio test based on the beta-binomial model has found prominence in the literature (Pack, 1986(a)). We
consider procedures for testing the homogeneity of proportions in the presence of a common dispersion parameter. We develop $C(\alpha)$ (Neyman, 1959) or score type tests (Rao, 1947) based on a parametric model; namely, the extended beta-binomial model (Prentice, 1986) and two semi-parametric models using the quasi-likelihood (Wedderburn, 1974) and the extended quasi-likelihood (Nelder and Pregibon, 1987). We also derive a $C(\alpha)$ test using empirical variance based on quasi-likelihood. These procedures and a recent procedure by Rao and Scott (1992), based on the concept of design effect and effective sample size, are compared, through simulations in terms of size, power and robustness for departure from data distribution and dispersion homogeneity. To study robustness in terms of departure from data distribution, i.e., departure from the beta-binomial distribution, we simulate data from the beta-binomial distribution, the probit normal binomial distribution and the logit normal binomial distribution.

Further, we develop $C(\alpha)$ tests for testing the assumption of a common dispersion parameter based on semi-parametric models. In some cases the assumption of a common dispersion parameter might not be tenable. A $C(\alpha)$ test is derived for testing the homogeneity of proportions with unequal dispersion parameters based on semi-parametric models.
Respectfully Dedicated to the Loving Memories of my Father and Mother
ACKNOWLEDGEMENTS

I wish to express my deep appreciation to my supervisor professor S.R. Paul for his guidance and encouragement in preparation of this dissertation. I would like to thank Dr. M.P. Becker, Dr. K.Y. Fung, Dr. M. Hlynka, Dr. J.H. Ciborowski and Dr. D.S. Tracy for their useful comments and suggestions which greatly improved this dissertation.

I would like to thank the Department of Mathematics and Statistics for giving me financial support in terms of graduate teaching assistantship and Dr. S.R. Paul for research assistantship through his NSERC grant throughout my graduate studies.

I would like to thank the Governments of Kenya and Canada for providing me with the CIDA scholarship during the duration of my graduate studies. Without this financial support it would not have been possible to complete my studies. May God Bless you all.

I would like to thank my late parents for their guidance and support throughout my life. May Allah rest their souls in peace.

I would like to thank my wife, son, daughter and all the family members for their support, understanding and prayers during my studies.

I would like to thank my best friends Dr. and Mrs. Bell for their support and affection. I would also like to thank Ms. Rosemarie Gignac for her assistance in typing this thesis.
# TABLE OF CONTENTS

**ABSTRACT**

iv

**DEDICATION**

vi

**ACKNOWLEDGMENTS**

vii

**CHAPTER 1: INTRODUCTION AND SOME PRELIMINARIES**

1.1 Introduction  
1

1.2 Preliminaries  
6

1.2.1 $\sqrt{m}$ consistent estimators  
6

1.2.2 Likelihood ratio test  
6

1.2.3 $C(\alpha)$ test  
7

1.2.4 Lindeberg theorem  
8

**CHAPTER 2: PARAMETRIC AND SEMI-PARAMETRIC MODELS AND ESTIMATION**

2.1 Introduction  
10

2.2 Estimation  
11

2.2.1 The extended beta-binomial likelihood  
11

2.2.2 The quasi-likelihood  
13

2.2.3 The extended likelihood  
14

2.2.4 The Gaussian likelihood  
15

2.2.5 Quadratic estimating equations  
16

2.2.6 Moment methods  
17

2.3 A result of Inagaki  
20

2.4 Asymptotic variance covariance  
20
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4.1 Maximum likelihood estimates</td>
<td>20</td>
</tr>
<tr>
<td>2.4.2 Quasi-likelihood estimates</td>
<td>22</td>
</tr>
<tr>
<td>2.4.3 Extended quasi-likelihood estimates</td>
<td>23</td>
</tr>
<tr>
<td>2.4.4 Gaussian likelihood estimates</td>
<td>25</td>
</tr>
<tr>
<td>2.4.5 Quadratic estimates</td>
<td>26</td>
</tr>
<tr>
<td>2.4.6 Method of moments estimates</td>
<td>28</td>
</tr>
<tr>
<td>2.5 Efficiency</td>
<td>31</td>
</tr>
<tr>
<td>2.5.1 Asymptotic relative efficiency</td>
<td>32</td>
</tr>
<tr>
<td>2.5.2 Small sample relative efficiency</td>
<td>46</td>
</tr>
<tr>
<td>2.6 Examples</td>
<td>68</td>
</tr>
<tr>
<td>2.7 Discussion</td>
<td>73</td>
</tr>
</tbody>
</table>

CHAPTER 3: TESTING FOR HOMOGENEITY OF PROPORTIONS

IN THE PRESENCE OF A COMMON

DISPERSION PARAMETER

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Introduction</td>
<td>74</td>
</tr>
<tr>
<td>3.2 Parametric methods</td>
<td>77</td>
</tr>
<tr>
<td>3.2.1 Likelihood ratio test</td>
<td>77</td>
</tr>
<tr>
<td>3.2.2 $C(\alpha)$ statistics based on beta-binomial model</td>
<td>78</td>
</tr>
<tr>
<td>3.2.3 Simulations</td>
<td>81</td>
</tr>
<tr>
<td>3.3 Semi-parametric methods</td>
<td>85</td>
</tr>
<tr>
<td>3.3.1 $C(\alpha)$ statistics based on the quasi-likelihood</td>
<td>85</td>
</tr>
<tr>
<td>3.3.2 $C(\alpha)$ statistics based on the extended quasi-likelihood</td>
<td>89</td>
</tr>
<tr>
<td>3.3.3 Rao-Scott statistics</td>
<td>92</td>
</tr>
<tr>
<td>3.3.4 The empirical score test</td>
<td>93</td>
</tr>
<tr>
<td>3.3.4.1 Introduction</td>
<td>93</td>
</tr>
<tr>
<td>3.3.4.2 The empirical score test</td>
<td>94</td>
</tr>
</tbody>
</table>
3.4.5 Simulations

3.4.5.1 Data generation  
3.4.5.2 Significance level  
3.4.5.3 Power  
3.4.5.4 Robustness under dispersion heterogeneity  

3.4.6 Examples and discussion  

CHAPTER 4: TESTING THE ASSUMPTION OF A COMMON 
DISPERSION PARAMETER IN THE ANALYSIS 
OF PROPORTIONS  

4.1 Introduction  

4.2 The $C(\alpha)$ statistics based on the quasi-likelihood 
and method of moments  

4.3 The $C(\alpha)$ statistics based on the gaussian likelihood  

CHAPTER 5: TESTING OF HOMOGENEITY OF PROPORTIONS 
IN PRESENCE OF POSSIBLE UNEQUAL 
DISPERSION PARAMETERS  

5.1 Introduction  

5.2 The $C(\alpha)$ statistics based on the quasi-likelihood 
and method of moments (Breslow's)  

5.3 The $C(\alpha)$ statistics based on the quasi-likelihood 
and method of moments (Srivastava's)  

5.4 Simulations  

CHAPTER 6: SUMMARY AND RECOMMENDATIONS FOR 
FUTURE RESEARCH  

6.1 Summary  

6.2 Recommendations for future research
CHAPTER I

INTRODUCTION AND SOME PRELIMINARIES

1.1 Introduction

Data in the form of proportions arise in toxicology, teratology (Paul, 1982; Weil, 1970; Williams, 1975) and other similar fields (Crowder, 1978; Otake and Prentice, 1984). These proportions often exhibit variation greater than predicted by a simple binomial model. In studies where the experimental unit is a litter, it has been observed (Mantel, 1969; Weil, 1970) that an inherent characteristic of data from these types of studies is the ‘litter effect’, i.e., there is a tendency of littermates to respond more alike than animals from different litters. This litter effect is also known as intra-litter correlation. There is considerable current interest in modelling correlated binomial data (Crowder, 1978; 1985; Haseman and Kupper, 1979; Prentice, 1986; Srivastava and Wu, 1993; Williams, 1975). A number of parametric models have been developed to incorporate the extra-binomial variation due to the litter effect.

Several parametric models such as the beta-binomial (BB) (Skellam, 1948), the correlated binomial model (Kupper and Haseman, 1978) and the additive and multiplicative binomial models (Altham, 1978) are available for analysing binomial data with overdispersion. Of these, the correlated binomial and the additive binomial models are identical and give a first order approximation of the beta-binomial model (Srivastava and Wu, 1993). The correlated and multiplicative binomial models allow the possibility of negative correlation (or under dispersion) while the beta-binomial permits only positive correlation (over dispersion). The beta-binomial is a commonly used parametric model because it is easy to use, flexible and extends readily...
to more complex models of the extraneous variance as a function of covariates (Chen and Kodell, 1989). The superiority of the beta-binomial (BB) model for analysing proportions and model fitting has been shown by many authors (Paul, 1982; Pack, 1986(a), 1986(b)). In model fitting the superiority of the BB model over the others is that the BB model is more sensitive for departure from the binomial model and the likelihood under BB is the easiest to maximize. In analysing data sets the BB model is preferred to the others in that it is easier to use. This model has been employed in modelling consumer purchasing behaviour (Chatfield and Goodhart, 1970), in studies of dental caries in children (Weil, 1970), and in toxicological data (Williams, 1975). The beta-binomial model, however, takes into account only the positive correlation between littermates; although in practice the littermates may compete with each other and as a result negative correlations among the littermates may occur. To correct this drawback of the beta-binomial model Prentice (1986) developed an extension of the beta-binomial model to allow negative correlations among the binary variates corresponding to the littermates, which is known as extended beta-binomial model. The extended beta-binomial model is a fully parametric model. It may not fit the data if the variance structure is misspecified. Therefore it is important in practice to model the parametric assumption accurately.

An alternative approach to incorporate the extra-binomial variation is the use of semi-parametric models and the method of moments, which are known to be robust to misspecification of the variance-structure. Several semi-parametric models have been proposed that require the assumption of only the first two moments. For binomial data with over/under dispersion several semi-parametric models are possible. These include models based on the Quasi-likelihood (McCullagh, 1983; Wedderburn 1974; Williams, 1982), the extended quasi-likelihood (Nelder and Pregibon, 1987), the pseudo-likelihood (Davidian and Carrol, 1987) and those based on optimal qua-
dratic estimating equations (Crowder, 1987; Godambe and Thompson, 1989). For
the estimation of mean and dispersion parameters, several estimators based on the
method of moments have been proposed by Breslow (1990(a)), Kleinman (1973),

The maximum likelihood estimates of the mean and dispersion parameters based
on a beta-binomial distribution are known from the classical theory to be asympto-
totically consistent and efficient if the parametric assumption is correct. However,
these estimates are often computationally intensive and are found to be biased when
the assumed variance structure is incorrect (Williams, 1982). The semi-parametric
and method of moments procedures of estimation have many desirable properties
such as computational ease, high efficiencies for estimates of the mean (regression)
parameters and robustness of the estimates to variance structure misspecifications.
Some comparisons of efficiency of method of moments estimates of the parameters
of the two parameter model have been conducted (Kleinman, 1971; Crowder, 1985;
Moore, 1985). However, a comparison of the methods based on semi-parametric
models and method of moment in current use is lacking.

An important problem in toxicology and other similar fields is to compare propor-
tions of a certain characteristic in several groups. A special case is to compare the
proportion in a control group with that in a treatment group. The likelihood ratio
test and some non-parametric procedures have been developed for testing homo-
geunity of the proportions. Of these, the likelihood ratio test has found prominence
in literature. Williams (1975) tested treatment differences using the likelihood ratio
test based on the beta-binomial distribution. Pack (1986(a)) studied, by simulation,
the performance in terms of size and power of the likelihood ratio test (LR) using
beta-binomial distribution and some simpler t-tests for testing the homogeneity of
proportions. He concluded that the LR test is more powerful than the simple t-tests
in a wide range of situations. One of the drawbacks of the LR test is that estimates of the parameters under both the null and the alternative hypotheses are required.

An alternative is to use the score or $C(\alpha)$ tests which avoid having to use mle's of the parameters under the alternative hypotheses, but which have desirable optimality properties (See Chapter 3).

The score or $C(\alpha)$ statistics have been found to be useful for testing over-dispersion in binomial and Poisson data (Paul, Liang and Self, 1989: Dean and Lawless, 1990). Barnwal and Paul (1988) derived $C(\alpha)$ statistics for testing the equality of Poisson means in the presence of negative binomial over-dispersion. Breslow (1990(a)) developed a robust score test for testing the significance of added variables in the over-dispersed Poisson regression using quasi-likelihood models.

The main purpose of this thesis is to derive procedures for testing the homogeneity of proportions where the data are correlated binomial using parametric and semi-parametric models. This thesis is also concerned with the comparison of efficiencies of several estimation procedures for estimating the mean (regression) and dispersion parameters in over/under dispersed binomial data.

In Chapter 2, we investigate several procedures for the joint estimation of mean (regression) and dispersion (intraclass correlation) parameters in over/under dispersed binomial data with or without covariates. We consider joint estimates by the extended beta-binomial likelihood, a combination of the quasi-likelihood estimating equation for the mean and the moment estimating equation for the intraclass correlation (Breslow, 1990(a); Moore and Tsiatis, 1991), extended quasi-likelihood (Nelder and Pregibon, 1987), the Gaussian likelihood (Whittle, 1961; Crowder, 1985), quadratic estimating equations (Crowder, 1987; Godambe and Thompson, 1989) and several moment methods (Kleinman, 1973; Srivastava and Wu, 1993). We compare large and small sample efficiencies of the different estimates of the
mean (regression) and the intraclass correlation parameters relative to the maximum likelihood estimates using the extended beta-binomial model.

In Chapter 3, we derive procedures for testing the homogeneity of proportions in the presence of a common dispersion parameter for over/under dispersed binomial data. We derive $C(\alpha)$ or score tests based on a parametric model, namely the extended beta-binomial model. $C(\alpha)$ or scores tests are also derived based on the quasi-likelihood and the extended quasi-likelihood. These procedures and a recent procedure by Rao and Scott (1992), based on the concept of design effect and effective sample size are compared, through simulation, in terms of size, power and robustness for departures from the beta-binomial distribution and dispersion heterogeneity.

In Chapter 4, we derive procedures for testing the assumption of a common dispersion parameter in the over/under dispersed binomial data. Two $C(\alpha)$ or score tests have been derived based on the quasi-likelihood/method of moments and the Gaussian-likelihood. The performance study of these procedures has been left for future study.

In Chapter 5, we derive procedures for testing the homogeneity of proportions with unequal dispersion parameters. $C(\alpha)$ or score statistics are derived based on the quasi-likelihood/method of moments and method of moments proposed by Srivastava and Wu (1993). A small scale simulation study is conducted to compare the empirical size and robustness for departures from data distribution of $C(\alpha)$ statistics. The performance, in terms of size and robustness for departures from data distribution and dispersion heterogeneity, of the $C(\alpha)$ statistics based on empirical variances (Breslow, 1989;1990(a)) derived in section (3.4), needs to be investigated further. This has been left as a topic for future study.
1.2 Preliminaries

1.2.1 $\sqrt{m}$ consistent estimators

Definition:
Let $\{\hat{\theta}_m\}$, $m = 1, 2, \ldots$, be a sequence of estimators. If the quantity $|\hat{\theta}_m - \theta|/\sqrt{m}$ remains bounded in probability as $m \to \infty$, then the sequence of estimates $\hat{\theta}_m$ is called $\sqrt{m}$ consistent estimator.

Theorem:
Let $\hat{\theta}_m$ be a sequence of estimates of $\theta$, and $\var(\hat{\theta}_m) = O\left(\frac{1}{m}\right)$, then this sequence of estimates is $\sqrt{m}$-consistent.

Proof:
By using Chebyshev's inequality, for a given $\epsilon > 0$,

$$
P\left( |\hat{\theta}_m - \theta|/\sqrt{m} \leq \epsilon \right) \geq 1 - \frac{\var(\hat{\theta}_m)m}{\epsilon^2}
$$

Let $\hat{\theta}_m$ be the sequence of maximum likelihood estimates, then by the asymptotic properties of MLE, $\var(\hat{\theta}_m)$ tends to zero as $m$ tends to infinity, i.e. $\var(\hat{\theta}_m)$ is $O(m^{-1})$. Thus, MLE is root $m$-consistent. The method of moments estimators are also $\sqrt{m}$-consistent estimates, (Moore, 1986).

1.2.2 Likelihood Ratio Test

Suppose $Y = (Y_1, \ldots, Y_m)'$ is a random sample of size $m$ taken from a distribution with pdf $f(y; \lambda)$, where $\lambda = (\theta, \phi)' = (\theta_1, \ldots, \theta_k, \phi_1, \ldots, \phi_s)'$ is a $k + s$ component vector. Then the likelihood can be given as $L(Y_1, \ldots, Y_m, \lambda)$. It is of interest to test the null hypothesis $H_0 : \theta = \theta_o = (\theta_{1o}, \ldots, \theta_{ko})'$ treating $\phi = (\phi_1, \ldots, \phi_s)'$ as a nuisance parameter.

The likelihood ratio for testing $H_0$ is defined as

$$
\Lambda = \frac{L\left(Y_1, \ldots, Y_m, \theta_o, \phi\right)}{L\left(Y_1, \ldots, Y_m, \bar{\theta}, \bar{\phi}\right)}
$$
Then the log likelihood ratio is given by

\[ LR = -2 \ln \Lambda = 2(l_1 - l_0) \]

where \( l_o \) is the maximum log-likelihood function under \( H_o \) and \( l_1 \) is the maximum log-likelihood function under alternative hypotheses. Under the null hypotheses \( H_o \), for a large \( m \), the statistic \( LR \) is distributed approximately as a chi-square with \( k \) degrees of freedom.

1.2.3 \( C(\alpha) \) Test

Let \( l \) be the log-likelihood of the data. The partial derivatives evaluated at \( \theta = \theta_o = (\theta_{1o}, \ldots, \theta_{ko})' \) are

\[ \psi = \left. \frac{\partial l}{\partial \theta} \right|_{\theta = \theta_o} = \left[ \frac{\partial l}{\partial \theta_1}, \ldots, \frac{\partial l}{\partial \theta_k} \right]' \bigg|_{\theta = \theta_o} \]

and

\[ \gamma = \left. \frac{\partial l}{\partial \phi} \right|_{\theta = \theta_o} = \left[ \frac{\partial l}{\partial \phi_1}, \ldots, \frac{\partial l}{\partial \phi_k} \right]' \bigg|_{\theta = \theta_o} \]

Cramer (1946) has shown that under the null hypothesis and mild regularity conditions, \( \left( \frac{\partial l}{\partial \theta}, \frac{\partial l}{\partial \phi} \right) \) follows a multivariate normal distribution with mean vector 0 and variance covariance matrix \( I^{-1} \), where

\[ I = \begin{bmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{bmatrix} \]

is the Fisher information with elements

\[ I_{11} = \mathbb{E} \left( \frac{-\partial^2 l}{\partial \theta \partial \theta'} \bigg|_{\theta = \theta_o} \right) \]

\[ I_{12} = \mathbb{E} \left( \frac{-\partial^2 l}{\partial \theta \partial \phi'} \bigg|_{\theta = \theta_o} \right) \]

and

\[ I_{22} = \mathbb{E} \left( \frac{-\partial^2 l}{\partial \phi \partial \phi'} \bigg|_{\phi = \phi_o} \right) \]
Define $S = \frac{\partial l}{\partial \theta} - B \frac{\partial l}{\partial \phi}$, where $B$ is the partial regression coefficient matrix obtained by regressing $\frac{\partial l}{\partial \theta}$ on $\frac{\partial l}{\partial \phi}$. From Bartlett (1953), $B = I_{12} I_{22}^{-1}$ and the Dispersion matrix of $S$ is $I_{11.2}$ where $I_{11.2} = I_{11} - I_{12} I_{22}^{-1} I_{21}$.

Thus $S$ is multivariate normal with mean vector 0 and variance-covariance matrix $I_{11.2}$, i.e.

$$S \sim MN(0, I_{11.2}).$$

Hence following Neyman (1959) $S' I_{11.2}^{-1} S \sim \chi^2_{(k)}$. But the above expression depends on the nuisance parameters $\phi = (\phi_1, \ldots, \phi_s)'$, which makes the statistic unsuitable to use for testing the null hypothesis. Moran (1970) suggested that we replace the unknown nuisance parameter $\phi$ by its $\sqrt{m}$-consistent estimator, obtained from the data. Let $\hat{\phi} = (\hat{\phi}_1, \ldots, \hat{\phi}_s)'$ be some $\sqrt{m}$ consistent estimators of the parameters $\phi = (\phi_1, \ldots, \phi_s)'$. Hence following the Neyman (1959) procedure.

$$\chi^2_{(\alpha)} = S' I_{11.2}^{-1} \hat{S}$$

which is asymptotically distributed as chi-square with $k$ degrees of freedom. Note that if we replace the nuisance parameter $\phi$ by its mle $\hat{\phi}$, then the score function $S_i$ reduces to $\psi_i, i = 1, \ldots, k - 1$. The $O(\alpha)$ statistic then reduces to $\hat{\psi}' I_{11.2}^{-1} \hat{\psi}$ which is referred to as a score test (Rao, 1947). The score test is asymptotically equivalent to the likelihood ratio test (Moran, 1970; Cox and Hinkley, 1974).

1.2.4 Lindeberg theorem

The central limit theorem holds whenever for every $\epsilon > 0$ the truncated variables $U_j$ defined by

$$U_j = Z_j - \mu_j \quad \text{if } |Z_j - \mu_j| \leq \epsilon s_m$$

$$U_j = 0 \quad \text{if } |Z_j - \mu_j| > \epsilon s_m$$

satisfy the condition $s_m \to \infty$ and

$$\frac{1}{s_m^2} \sum_{j=1}^{m} E(U_j^2) \to 1.$$
where $\mu_j = E(Z_j)$, $\sigma_j^2 = \text{var}(Z_j)$ and $s_m^2 = \sum_{j=1}^{m} \sigma_j^2$.

This theorem implies that every uniformly bounded sequence of $\{Z_j\}$ of mutually independent random variables obeys the central limit theorem, provided that $s_m \rightarrow \infty$.

**Theorem**

Let $U_j(\beta, \phi), j = 1, \ldots, k$, represent an unbiased estimating function for $\beta_j$ and let $U_{k+1}(\beta, \phi)$ represent an unbiased estimating function for $\phi$. Then by the Lindeberg central limit theorem

\[
\frac{1}{\sqrt{m}} \begin{pmatrix} U_j \\ U_2 \end{pmatrix} \rightarrow_d N_{k+1} \left( 0, \sum_o \right)
\]

where $\sum_o$ denotes the limiting covariance matrix of $U_j$ and $U_2$. The details of the proof of this theorem are given by Moore (1985).
CHAPTER II
PARAMETRIC AND SEMI-PARAMETRIC MODELS
AND ESTIMATION

2.1 Introduction

There is considerable current interest in modelling correlated binomial data (Haseman and Kupper, 1979; Crowder, 1978, 1985; Prentice, 1986; Srivastava and Wu, 1993). In situations where interest is on the mean (regression) parameter, the dispersion parameter or the intraclass correlation parameter can be treated as a nuisance parameter. However, in many instances the dispersion parameter or the intraclass correlation parameter is important in its own right. In some binary data situations, it is interpreted as the ‘heritability of a dichotomous trait’ (see Crowder, 1982; Elston, 1977). Therefore joint estimation of the mean (regression) and the dispersion or intraclass correlation is important. Nelder and Lee (1992) emphasize the importance of joint estimation of the mean (regression) and dispersion parameters and indicate that for the analysis of quality improvement experiments, estimation of the dispersion parameter is itself of interest.

A classical approach to model correlated binomial data is to assume that the data come from a specified parametric model. A commonly used model to account for over dispersion is the Beta-binomial model (Williams, 1975; Crowder, 1978). More recently, Prentice (1986) extended the beta-binomial model to include over-dispersion as well as under-dispersion. Maximum likelihood estimates of the mean (regression) and the dispersion parameters can be obtained by using the extended beta-binomial likelihood. However, the MLEs are much more computation-
ally intensive than estimates by the method of moments or other semi-parametric methods. Williams (1988) reported serious biases resulting in the estimation of the dispersion parameter when the beta-binomial model is fitted to data simulated from the logistic-normal-binomial (LNB) distribution. This shows that ML estimation is sensitive to misspecification of the over-dispersion model. Further the beta-binomial distribution is not a member of the exponential family of models or exponential dispersion family of models (Jorgensen, 1987). Due to these unfavourable properties, it is suggested that the method of moments or other semi-parametric methods might be superior to the maximum likelihood method for over/under dispersed binomial data.

In this Chapter, we investigate several procedures for the joint estimation of the mean (regression) and intraclass correlation parameters in over (or under) dispersed binomial data with or without covariates. We consider a parametric procedure based on the extended beta-binomial model, several semi-parametric methods, namely, the quasi-likelihood (Breslow, 1990(a); Moore and Tsiatis, 1991), the extended quasi-likelihood (Nelder and Pregibon, 1987) and the quadratic estimating equations (Crowder, 1987; Godambe and Thompson, 1989), based on the assumption that the data come from any distribution that is specified by only the first two moments of the binomial response with over (or under) dispersion, and some moment methods (Kleinman, 1973; Srivastava and Wu, 1993). A method by Whittle (1961) is also used for the joint estimation of the mean (regression) and dispersion parameters. The estimates are developed in section 2.2.

2.2 Estimation

2.2.1 The Extended Beta-binomial Likelihood

A plausible model for over-dispersed binomial data is the beta-binomial model (BB) (Williams, 1975), and may be obtained in the following way. We assume that
$Y_i | p_i \sim$ binomial $(n_i, p_i)$ for $i = 1, \ldots, m$. We assume further that the binomial probability $p_i$ is distributed as a beta distribution with mean $\pi_i$ and variance $\pi_i (1 - \pi_i) \theta_i / (1 + \theta_i)$.

Thus

$$P(Y_i = y_i | p_i) = \binom{n_i}{y_i} p_i^{y_i} (1 - p_i)^{n_i - y_i}$$

and

$$f(p_i) = \frac{p_i^{\alpha_i - 1} (1 - p_i)^{\delta_i - 1}}{B(\alpha_i, \delta_i)} \quad 0 < p_i < 1,$$

$$\alpha_i > 0,$$

$$\delta_i > 0.$$  

The marginal distribution of $Y_i$ is then Beta-binomial with

$$P(Y_i = y_i) = \binom{n_i}{y_i} \frac{B(\alpha_i + y_i, n_i + \delta_i - y_i)}{B(\alpha_i, \delta_i)} \quad \text{for} \quad 0 \leq y_i \leq n_i.$$  

We define $\pi_i = \frac{\alpha_i}{\alpha_i + \delta_i}$, $\theta_i = \frac{1}{\alpha_i + \delta_i}$. The extended BB probability function is given by

$$P(Y_i = y_i | n_i) = \frac{\binom{n_i}{y_i} \prod_{r=0}^{\delta_i - 1} (\pi_i + r \theta_i) \prod_{r=0}^{\delta_i - 1 - y_i} (1 - \pi_i + r \theta_i)}{\prod_{r=0}^{n_i - 1} (1 + r \theta_i)},$$

$$i = 1, \ldots, m, \quad 0 \leq \pi_i \leq 1 \quad , \quad 0 \leq y_i \leq n_i,$$

and

$$\theta_i \geq \max \left\{ \frac{-\pi_i}{(n_i - 1)}, \frac{(1 - \pi_i)}{(n_i - 1)} \right\} \quad (\text{Prentice}, 1986).$$

Skellam (1948) introduced the beta-binomial model in which $\theta$ is strictly positive.

As discussed earlier, under dispersion can also occur in practice. To incorporate this Prentice (1986) extended the BB to extended BB model. The mean structure $\pi_i$ is assumed to follow the logistic model

$$\pi_i \sim \text{Logistic}(X_i; \beta) = e^{X_i \beta} / (1 + e^{X_i \beta})$$

where $X_i \beta = X_{i1} \beta_1 + \ldots + X_{ik} \beta_k$, $X_1, \ldots, X_k$ are $k$ explanatory variables, and $\beta_1, \ldots, \beta_k$ are the $k$ regression parameters. The mean and variance of the extended beta-binomial variate $Y_i$ are $n_i \pi_i$ and $n_i \pi_i (1 - \pi_i) \{1 + (n_i - 1) \phi\}$, where
\( \phi = \theta / (1 + \theta) \). The extended beta-binomial log-likelihood is
\[
l = \sum_{i=1}^{m} \left\{ \sum_{r=0}^{y_i-1} \log (\pi_i + r\theta) + \sum_{r=0}^{n_i-y_i-1} \log (1 - \pi_i + r\theta) - \sum_{r=0}^{n_i-1} \log (1 + r\theta) \right\}.
\]
The Maximum likelihood estimates (MLE) of \( \beta_1, \ldots, \beta_k \) and \( \theta \) can be obtained by maximizing \( l \) or alternatively, by solving the ml estimating equations:
\[
\frac{\partial l}{\partial \beta_j} = \sum_{i=1}^{m} \left\{ \sum_{r=0}^{y_i-1} \frac{1}{(\pi_i + r\theta)} - \sum_{r=0}^{n_i-y_i-1} \frac{1}{(1 - \pi_i + r\theta)} \right\} \pi_i (1 - \pi_i) X_{ij} = 0 \quad j = 1, \ldots, k
\]
and
\[
\frac{\partial l}{\partial \theta} = \sum_{i=1}^{m} \left\{ \sum_{r=0}^{y_i-1} \frac{r}{(\pi_i + r\theta)} + \sum_{r=0}^{n_i-y_i-1} \frac{r}{(1 - \pi_i + r\theta)} - \sum_{r=0}^{n_i-1} \frac{r}{(1 + r\theta)} \right\} = 0
\]
simultaneously. The MLE of \( \phi \) is obtained as \( \hat{\phi} = \hat{\theta} / (1 + \hat{\theta}) \), where \( \hat{\theta} \) is the mle of \( \theta \). Now, denote the parameter vector \( (\beta_1, \ldots, \beta_k, \phi) \) by \( \lambda \). Then, denote the maximum likelihood estimate of \( \lambda \) by \( \hat{\lambda}_{ML} \).

2.2.2 The Quasi-Likelihood

The quasi-likelihood (Wedderburn, 1974) is based on the knowledge of the form of the first two moments of the random variable \( Z_i = Y_i / n_i \)
\[
E(Z_i) = \pi_i, \quad \text{var}(Z_i) = \frac{\pi_i (1 - \pi_i)}{n_i} \left\{ 1 + (n_i - 1) \phi \right\}.
\]
\( i = 1, \ldots, m, 0 \leq \pi_i \leq 1 \) and \( \int_{\pi_i}^{1} \left( \frac{-1}{n_i - 1} \right) < \phi < 1 \). This specification of the mean and variance coincides with those based on the extended beta-binomial model. The quasi-likelihood for an observation \( Z_i \) with the above mean and variance is given by
\[
Q(z_i, \pi_i, \phi) = \int_{z_i}^{\pi_i} \frac{(z_i - t) n_i}{t (1 - t) \left\{ 1 + (n_i - 1) \phi \right\}} dt,
\]
which for the data under consideration becomes
\[
Q = \sum_{i=1}^{m} \frac{1}{\left\{ 1 + (n_i - 1) \phi \right\}} \left[ y_i \log (\pi_i / z_i) + (n_i - y_i) \log \left\{ (1 - \pi_i) / (1 - z_i) \right\} \right].
\]
Given $\phi$, the unbiased estimating equations for $\beta$ are

$$U_j(\beta, \phi) = \frac{\partial Q}{\partial \beta_j} = \sum_{i=1}^{m} \frac{(z_i - \pi_i) n_i X_{ij}}{1 + (n_i - 1) \phi} = 0. \quad j = 1, \ldots, k. \quad (2.2.2.1)$$

No such estimating equation exists for $\phi$. However, an unbiased estimating equation for $\phi$ can be obtained by using the moment method (Breslow, 1990(a); Moore and Tsiatis, 1991).

$$U_{k+1}(\beta, \phi) = \sum_{i=1}^{m} \frac{(z_i - \pi_i)^2 n_i}{\pi_i (1 - \pi_i) \{1 + (n_i - 1) \phi\}} - (m - k) = 0. \quad (2.2.2.2)$$

Maximum quasi-likelihood estimates of $\beta$ and $\phi$ are obtained by solving equations (2.2.2.1) and (2.2.2.2) simultaneously. Denote the estimates by $\hat{\lambda}_{QM}$.

2.2.3 The Extended quasi-likelihood

The quasi-likelihood $Q$ does not behave like a log-likelihood with respect to the $\phi$ derivative. The Extended quasi-likelihood of Nelder and Pregibon (1987) can be used for the simultaneous estimation of the $\beta_j$ and $\phi$. The extended quasi-log-likelihood for an observation $z_i$ with mean and variance specified above is

$$Q^+ (z_i, \pi_i, \phi) = -\frac{1}{2} \ln(2k) - \frac{1}{2} \ln \left[ \frac{z_i (1 - z_i) \{1 + (n_i - 1) \phi\}}{n_i} \right]$$

$$+ \int_{z_i}^{\pi_i} \frac{n_i (z_i - t)}{t (1 - t) \{1 + (n_i - 1) \phi\}} dt$$

The extended quasi-log-likelihood for the data under consideration, then, is

$$Q^+ = C - \frac{1}{2} \sum_{i=1}^{m} \left[ \ln \{1 + (n_i - 1) \phi\} + \frac{2 y_i \ln(\pi_i/z_i)}{1 + (n_i - 1) \phi} \right]$$

$$+ (n_i - y_i) \ln \{(1 - \pi_i)/(1 - z_i)\}$$

where $C$ is a term not involving the parameters. The unbiased estimating equations for $\beta_j$ and $\phi$ obtained from $Q^+$ are

$$U_j(\beta, \phi) = \frac{\partial Q^+}{\partial \beta_j} = \sum_{i=1}^{m} \frac{(z_i - \pi_i) n_i X_{ij}}{1 + (n_i - 1) \phi} = 0. \quad j = 1, \ldots, k \quad (2.2.3.1)$$
\[ U_{s+1}(\beta, \phi) = \frac{\partial Q^+}{\partial \phi} = \sum_{i=1}^{m} \frac{(n_i - 1)}{1 + (n_i - 1) \phi} \left[ y_i \ln \left( \frac{z_i}{\pi_i} \right) \right] \\
+ (n_i - y_i) \ln \left\{ \frac{(1 - z_i)}{(1 - \pi_i)} \right\} - \left\{ 1 + (n_i - 1) \phi \right\} / 2 \right] = 0. \]

Maximum extended quasi-likelihood estimates of \( \beta \) and \( \phi \) are obtained by solving (2.2.3.1) and (2.2.3.2) simultaneously. Denote the estimate by \( \hat{\lambda}_{EQE} \).

### 2.2.4 The Gaussian Likelihood

Whittle (1961) introduced the Gaussian estimation procedure which uses the normal log-likelihood, without assuming that the data are normally distributed. This procedure produces unbiased estimating equations for the \( \beta \) and \( \phi \) parameters.

The Gaussian likelihood is given by

\[ L = -\frac{1}{2} \sum_{i=1}^{m} \left\{ \ln \det \left( 2\pi \sum_i \right) + (y_i - M_i)^T (\sum_i)^{-1} (y_i - M_i) \right\} \]

where \( M_i = \pi_i 1_i \), \( 1_i = (1, \ldots, 1)^T \)

and \( \sum_i = \pi_i (1 - \pi_i) \{ (1 - \phi) I_i + \phi 1_i 1_i^T \} \)

where \( I_i \) is the unit matrix of order \( m \). The Gaussian log-likelihood, in our case, after simplifications, can be written as

\[ l = -\frac{1}{2} \sum_{i=1}^{m} \left\{ n_i \ln \{ \pi_i (1 - \pi_i) \} + (n_i - 1) \ln (1 - \phi) + \ln \{ 1 + (n_i - 1) \phi \} \right. \\
\left. + \{ \pi_i (1 - \pi_i) (1 - \phi) \}^{-1} \right\} n_i \pi_i (1 - \pi_i) + (1 - 2\pi_i) n_i (z_i - \pi_i) \\
- \phi \{ 1 + (n_i - 1) \phi \}^{-1} n_i^2 (z_i - \pi_i)^2 \right\} \].

Estimates of \( \beta_1, \ldots, \beta_k \) and \( \phi \) by the Gaussian estimation procedure are obtained by solving the following estimating equations

\[ U_j(\beta, \phi) = \frac{\partial l}{\partial \beta_j} = \sum_{i=1}^{m} \left\{ \frac{(1 - \phi)}{\{ 1 + (n_i - 1) \phi \}} + \frac{(1 - 2\pi_i)^2}{2\pi_i (1 - \pi_i) (1 - \phi)} \right\} \frac{n_i (z_i - \pi_i)}{\pi_i (1 - \pi_i) (1 - \phi)} \\
- \left\{ \frac{(1 - 2\pi_i) n_i \phi}{2\pi_i (1 - \pi_i) (1 - \phi)} \right\} \left\{ \frac{n_i (z_i - \pi_i)^2}{\pi_i (1 - \pi_i) (1 + (n_i - 1) \phi)} - 1 \right\} \pi_i (1 - \pi_i) X_{ij} = 0 \]

for \( j = 1, \ldots, k \).
and
\[ U_{k+1}(\beta, \phi) = \frac{\partial l}{\partial \phi} = \sum_{i=1}^{m} \left[ \frac{n_i}{2(1-\phi)^2} \left\{ \frac{1 + (n_i - 1)\phi^2}{1 + (n_i - 1)\phi} \right\} \left\{ \frac{n_i(z_i - \pi_i)^2}{\pi_i(1 - \pi_i)(1 + (n_i - 1)\phi)} - 1 \right\} \right. 

\left. - \left\{ \frac{(1 - 2\pi_i)}{2\pi_i(1 - \pi_i)(1 - \phi)^2} \right\} n_i(z_i - \pi_i) \right] = 0. \]

Denote the estimates by \( \hat{\lambda}_{GL} \).

2.2.5 Quadratic Estimating Equations

Crowder (1987) and Godambe and Thompson (1989) developed a general class of optimal quadratic estimating equations (QEE) for both the mean and the dispersion parameters. The estimators obtained by solving the QEEs has minimal asymptotic variance in the class of unbiased QEEs. The QEEs require the knowledge of the skewness and kurtosis of the data distribution which in practical situations is unlikely to be known exactly. Estimation of these requires much more data than is usually available in practice. For more discussions on the drawbacks of the QEEs, see Dean and Lawless (1989) and Nelder (1991). In our case, when the skewness and kurtosis are set to zero the QEEs convert to the pseudo-likelihood estimating equations of Davidian and Carrol (1987) and to a form of Gaussian estimating equations.

For the mean or regression parameters Gaussian and quasi-likelihood estimates are efficient (efficiency close to one) compared to the ML estimates based on the beta-binomial model. Dean and Lawless (1989) have similar findings for count data. Dean and Lawless (1989) suggested combining the quasi-likelihood estimating equations for the mean parameters and the optimal estimating equation of Crowder (1987) for dispersion parameter after setting the skewness and kurtosis to zero. In our case we follow the same suggestion.
The Quadratic Estimating Equation for $\phi$ using Crowder's and Godambe and Thompson's procedures is

$$U_{k+1}(\beta, \phi) = \sum_{i=1}^{m} \left\{ \frac{(z_{i} - \pi_{i})^{2} n_{i}}{\pi_{i} (1 - \pi_{i}) \{1 + (n_{i} - 1) \phi\}} - 1 - \frac{\gamma_{1i} (z_{i} - \pi_{i}) n_{i}^{1/2}}{[\pi_{i} (1 - \pi_{i}) \{1 + (n_{i} - 1) \phi\}]^{1/2}} \right\} \frac{(n_{i} - 1)}{[1 + (n_{i} - 1) \phi] \gamma_{i}}$$

(2.2.5.1)

where $\gamma_{1i}$ is the skewness and $\gamma_{2i}$ is the kurtosis and $\gamma_{i} = (\gamma_{2i} + 2 - \gamma_{1i}^{2})$.

The estimating equations so constructed are

$$U_{j}(\beta, \phi) = \sum_{i=1}^{m} \frac{(z_{i} - \pi_{i}) n_{i} X_{ij}}{\{1 + (n_{i} - 1) \phi\}} = 0, \quad j = 1, \ldots, k. \tag{2.2.5.2}$$

$$U_{k+1}(\beta, \phi) = \sum_{i=1}^{m} \frac{n_{i} (n_{i} - 1)}{\pi_{i} (1 - \pi_{i}) \{1 + (n_{i} - 1) \phi\}^{2}} \left[ (z_{i} - \pi_{i})^{2} - \frac{\pi_{i} (1 - \pi_{i}) \{1 + (n_{i} - 1) \phi\}}{n_{i}} \right] = 0. \tag{2.2.5.3}$$

The estimating equation for $\phi$ is obtained by setting $\gamma_{1i} = \gamma_{2i} = 0$ in equation (2.2.5.1). Denote the estimates, obtained by solving equations (2.2.5.2) and (2.2.5.3) simultaneously, by $\hat{\lambda}_{QEE}$.

2.2.6 Moment Methods

For the two parameter model, one can obtain moment estimators for $\pi$ and $\phi$ by equating the estimates of the mean and variance of the data $Z_{i}$ to their corresponding expectations. The estimating equations obtained for $\pi$ and $\phi$ can be extended to the regression situation for $\beta_{j}, j = 1, \ldots, k$ and $\phi$. Kleinman (1973) obtained moment estimates of $\pi$ and $\phi$ by equating the weighted mean

$$\bar{z} = \sum_{i=1}^{k} w_{i} z_{i} \sum_{i=1}^{k} w_{i}$$

and the weighted variance $S = \sum w_{i} (z_{i} - \bar{z})^{2}$, where
\( w_i = n_i / [\pi (1 - \pi) \{1 + (n_i - 1) \phi\}] \), to their corresponding expectations, that is, estimates of \( \pi \) and \( \phi \) are obtained by solving

\[
\hat{\pi} = \pi \tag{2.2.6.1}
\]

and

\[
S = E(S) \tag{2.2.6.2}
\]

simultaneously. Equations (2.2.6.1) can be written as an unbiased estimating equation for \( \pi \) by

\[
\sum_{i=1}^{m} w_i (z_i - \pi) = 0,
\]

which can be generalized to the regression situation. Thus, the unbiased estimating equation for \( \beta_j; j = 1, \ldots, k \), is

\[
U_j(\beta, \phi) = \sum_{i=1}^{m} \frac{(z_i - \pi_i) n_i X_{ij}}{\{1 + (n_i - 1) \phi\}} = 0 \tag{2.2.6.3}
\]

where \( \pi_i = \pi_i(X_i, \beta) = e^{X_i \beta} / (1 + e^{X_i \beta}) \).

Further, after obtaining \( E(S) \) in (2.2.6.2), the unbiased estimating equation for \( \phi \) can be written as

\[
U_{k+1}(\beta, \phi) = \sum_{i=1}^{m} \frac{(z_i - \pi_i)^2 n_i}{\pi_i (1 - \pi_i) \{1 + (n_i - 1) \phi\}} - \phi \left[ \sum_{i=1}^{m} \left\{ \frac{(n_i - 1)}{(1 + (n_i - 1) \phi)} - \frac{n_i (n_i - 1)}{\pi_i (1 - \pi_i) \{1 + (n_i - 1) \phi\}^2} V^{-1} \right\} \right]

- \sum_{i=1}^{m} \left\{ \frac{1}{(1 + (n_i - 1) \phi)} - \frac{n_i}{\pi_i (1 - \pi_i) \{1 + (n_i - 1) \phi\}^2} V^{-1} \right\} = 0 \tag{2.2.6.4}
\]

where

\[
V = \sum_{i=1}^{m} \frac{n_i}{\pi_i (1 - \pi_i) \{1 + (n_i - 1) \phi\}}.
\]
Thus, moment estimates by Kleinman's procedure are obtained by solving equations (2.2.6.3) and (2.2.6.4) simultaneously. Note that the moment estimating equation for $\beta_j$ is identical to the corresponding quasi-likelihood estimating equation. Denote these estimates by $\hat{\lambda}_{M1}$.

Other moment estimators of $\beta_j$ and $\phi$ are possible. For example, in the case of the two parameter model. Srivastava and Wu (1990) suggested the possibility of using two different estimates of $\tau$ and three different estimating equations for $\phi$, all based on method of moments. For estimating $\tau$, they preferred $\hat{\tau}_2 = \sum_{i=1}^{m} y_i / \sum_{i=1}^{m} n_i$, over another simple estimate, based on a smaller asymptotic variance. Now, $E(\hat{\tau}_2) = \tau$, so by equating $\hat{\tau}_2$ to $\tau$ we obtain the unbiased estimating equation

$$\sum_{i=1}^{m} n_i (z_i - \tau) = 0$$

This equation is identical in form to the quasi-likelihood estimating equation with $w_i = n_i$. From this the unbiased estimating equation for $\beta_j$ is obtained as

$$U_j (\beta, \phi) = \sum_{i=1}^{m} (z_i - \pi_i) n_i \pi_i (1 - \pi_i) X_{ij} = 0, \quad j = 1, \ldots, k \quad (2.2.6.5)$$

where $\pi_i = \pi_i (X_i, \beta) = e^{X_i \beta} / (1 + e^{X_i \beta})$. Of the three moment estimating equations for $\phi$ suggested by Srivastava and Wu (1990), they examined two of them for efficiency by an example and they found no reason to prefer one over the other. We consider both of these equations, which in our notation can be written as

$$U_{k+1} (\beta, \phi) = \sum_{i=1}^{m} \left\{ \frac{n_i (z_i - \pi_i)^2}{\pi_i (1 - \pi_i)} - \{1 + (n_i - 1) \phi\} \right\} = 0 \quad (2.2.6.6)$$

$$U_{k+1} (\beta, \phi) = \sum_{i=1}^{m} \left\{ \frac{n_i^2 (z_i - \pi_i)^2}{\pi_i (1 - \pi_i)} - n_i \{1 + (n_i - 1) \phi\} \right\} = 0. \quad (2.2.6.7)$$

In our study of efficiency we consider four moment estimates of $\hat{\lambda}$ denoted by $\hat{\lambda}_{M1}$, $\hat{\lambda}_{M2}$, $\hat{\lambda}_{M3}$ and $\hat{\lambda}_{M4}$. The estimates $\hat{\lambda}_{M1}$, as discussed earlier, are those obtained
by Kleinman’s moment procedure. The estimates \( \hat{\lambda}_{M2} \) are obtained by solving equations (2.2.6.5) and (2.2.6.6) simultaneously. Finally, the estimates \( \hat{\lambda}_{M3} \) and \( \hat{\lambda}_{M4} \) are obtained by combining the quasi-likelihood estimating equation (2.2.6.3) for \( \beta_j \) with the estimating equations for \( \phi \), (2.2.6.6) and (2.2.6.7) respectively.

2.3 A Result by Inagaki (1973)

The unbiased estimating functions obtained using the moment or semi-parametric procedures described in section (2.2) for convenience, are all denoted by the notations \( U_1, \ldots, U_k \) and \( U_{k+1} \), where \( U_j, j = 1, \ldots, k \), represents an unbiased estimating function for \( \beta_j \) and \( U_{k+1} \) represent an unbiased estimating function for \( \phi \). Let \( \hat{\lambda} \) be an estimate of \( \lambda \) using a moment or a semi-parametric method. Then, by Inagaki (1973) and under the usual regularity conditions such as: the parameter space has finite dimension, the expected values of \( U_j, j = 1, \ldots, k+1 \), exist for all the parameter space, the expected values of \( U_j \) are continuously differentiable, etc.,

\[
\text{var} \left( \hat{\lambda} \right) = \left[ A \left( \hat{\lambda} \right) \right]^{-1} \beta \left( \hat{\lambda} \right) \left\{ \left[ A \left( \hat{\lambda} \right) \right]^{-1} \right\}^T
\]

where \( A \left( \hat{\lambda} \right) \) and \( \beta \left( \hat{\lambda} \right) \) are \((k + 1) \times (k + 1)\) matrices with

\[
A_{js} = E \left( \frac{-\partial U_j}{\partial \beta_s} \right), \quad A_{j,k+1} = E \left( \frac{-\partial U_j}{\partial \phi} \right)
\]

\[
A_{k+1,j} = E \left( \frac{-\partial U_{k+1}}{\partial \beta_j} \right), \quad A_{k+1,k+1} = E \left( \frac{-\partial U_{k+1}}{\partial \phi} \right)
\]

\[
B_{js} = E \left( U_j U_s \right), \quad B_{k+1,k+1} = E \left( U_{k+1}^2 \right)
\]

\[
B_{j,k+1} = E \left( U_j U_{k+1} \right) = B_{k+1,j}
\]

2.4 Asymptotic variance covariance

2.4.1 Variance of Maximum Likelihood Estimator \( \hat{\lambda}_{ML} \)

The asymptotic variance-covariance matrix of the MLEs \( \hat{\beta}_1, \ldots, \hat{\beta}_k, \hat{\phi} \) is obtained by inverting the expected Fisher information matrix

\[
I = \begin{bmatrix}
I_{\beta \beta} & I_{\beta \theta} \\
I_{\beta \theta} & I_{\theta \theta}
\end{bmatrix},
\]

(2.4.1.1)
where
\[ I_{\beta\beta} = E \left( \frac{-\partial^2 l}{\partial \beta_j \partial \beta_s} \right) \]

is a \( k \times k \) matrix.

\[ I_{\beta\theta} = E \left( \frac{-\partial^2 l}{\partial \beta_j \partial \theta} \right) \]

is a \( k \times 1 \) matrix.

and
\[ I_{\theta\theta} = E \left( \frac{-\partial^2 l}{\partial \theta^2} \right). \]

The second derivatives of the beta-binomial loglikelihood \( l \) given in section (2.2.1) are obtained as
\[
\frac{\partial^2 l}{\partial \beta_j \partial \beta_s} = \left\{ \sum_{r=0}^{y_i-1} \frac{1}{(\pi_i + r\theta)^2} + \sum_{r=0}^{n_i-y_i-1} \frac{1}{(1 - \pi_i + r\theta)^2} \right\} \pi_i^2 (1 - \pi_i)^2 X_{ij} X_{is}, \quad j, s = 1, \ldots, k
\]
\[
\frac{\partial^2 l}{\partial \beta_j \partial \theta} = \left\{ -\sum_{r=0}^{y_i-1} \frac{r}{(\pi_i + r\theta)^2} + \sum_{r=0}^{n_i-y_i-1} \frac{r}{(1 - \pi_i + r\theta)^2} \right\} \pi_i (1 - \pi_i) X_{ij}, \quad j = 1, \ldots, k
\]
\[
\frac{\partial^2 l}{\partial \theta^2} = \left\{ -\sum_{r=0}^{y_i-1} \frac{r^2}{(\pi_i + r\theta)^2} + \sum_{r=0}^{n_i-y_i-1} \frac{r^2}{(1 - \pi_i + r\theta)^2} - \sum_{r=0}^{n_i-1} \frac{r^2}{(1 + r\theta)^2} \right\}. \]

The expectations of the negative of the second derivatives yield the Fisher information matrix with entries
\[
E \left( \frac{-\partial^2 l}{\partial \beta_j \partial \beta_s} \right) = \sum_{i=1}^{m} \left\{ \sum_{r=0}^{n_i-1} \frac{P(Y_i > r)}{(\pi_i + r\theta)^2} + \sum_{r=0}^{n_i-y_i-1} \frac{P(Y_i < n_i - r)}{(1 - \pi_i + r\theta)^2} \right\} \pi_i^2 (1 - \pi_i)^2 X_{ij} X_{is}, \quad j, s = 1, \ldots, k
\]
\[
E \left( \frac{-\partial^2 l}{\partial \beta_j \partial \theta} \right) = \sum_{i=1}^{m} \frac{1}{\theta^2} \left\{ -\pi_i \sum_{r=0}^{n_i-1} \frac{P(Y_i > r)}{(\pi_i + r\theta)^2} + (1 - \pi_i) \sum_{r=0}^{n_i-y_i-1} \frac{P(Y_i < n_i - r)}{(1 - \pi_i + r\theta)^2} \right\} \pi_i (1 - \pi_i) X_{ij}, \quad j = 1, \ldots, k
\]

and
\[
E \left( \frac{-\partial^2 l}{\partial \theta^2} \right) = \sum_{i=1}^{m} \frac{1}{\theta^2} \left\{ \sum_{r=0}^{n_i-1} \frac{\pi_i^2 P(Y_i > r)}{(\pi_i + r\theta)^2} + \sum_{r=0}^{n_i-y_i-1} \frac{(1 - \pi_i)^2 P(Y_i < n_i - r)}{(1 - \pi_i + r\theta)^2} \right\} - \sum_{r=0}^{n_i-1} \frac{1}{(1 + r\theta)^2}\right\}. \]
The detailed derivation and simplification of $I_{\beta\beta}, I_{\beta\theta}$ and $I_{\theta\theta}$ are given in the appendix A.

By using the delta method, the asymptotic variance of \( \hat{\phi} \) is then obtained as 
\[
\text{var}(\hat{\phi}) = \text{var}(\hat{\theta}) / (1 + \hat{\theta})^4.
\]

2.4.2 Variance of Quasi-likelihood Estimator $\hat{\lambda}_{QM}$

The asymptotic variance-covariance matrix of the estimators $\hat{\lambda}_{QM}$ is obtained using a result by Inagaki (1973) as described in section (2.2). The entries of the matrices $A$ and $B$ are:

\[
A_{js} = E \left( \frac{-\partial U_j}{\partial \beta_s} \right) = \sum_{i=1}^{m} \frac{n_i \pi_i (1 - \pi_i) X_{ij} X_{is}}{\{1 + (n_i - 1) \phi\}}, \quad j, s = 1, \ldots, k
\]

\[
A_{jk+1} = E \left( \frac{-\partial U_j}{\partial \phi} \right) = 0, \quad j = 1, \ldots, k
\]

\[
a_j = A_{k+1,j} = E \left( \frac{\partial U_{k+1}}{\partial \beta_j} \right) = \sum_{i=1}^{m} (1 - 2\pi_i) X_{ij}, \quad j = 1, \ldots, k
\]

\[
a_{k+1} = A_{k+1,k+1} = E \left( \frac{-\partial U_{k+1}}{\partial \phi} \right) = \sum_{i=1}^{m} \frac{(n_i - 1)}{\{1 + (n_i - 1) \phi\}}.
\]

\[
B_{js} = E(U_j U_s) = \sum_{i=1}^{m} \frac{n_i \pi_i (1 - \pi_i) X_{ij} X_{is}}{\{1 + (n_i - 1) \phi\}} = A_{js}, \quad j, s = 1, \ldots, k
\]

\[
b_j = B_{j,k+1} = E(U_j U_{k+1}) = \sum_{i=1}^{m} \frac{(1 - 2\pi_i)(1 + (2n_i - 1) \phi) X_{ij}}{(1 + \phi)\{1 + (n_i - 1) \phi\}}, \quad j = 1, \ldots, k
\]

\[
B_{k+1,j} = B_{j,k+1}
\]

\[
B_{k+1,k+1} = E(U_j U_{k+1}) = \sum_{i=1}^{m} \left\{ \frac{n_i^2 (n_i - 1)^2 E(z_i - \pi_i)^4}{\{n_i \pi_i (1 - \pi_i)\}^2 \{1 + (n_i - 1) \phi\}^4} - \frac{(n_i - 1)^2}{\{1 + (n_i - 1) \phi\}^2} \right\},
\]

\[
B_{k+1,k+1} = b_{k+1}.
\]
where

\[ E(z_i - \pi_i)^2 = \frac{\pi_i (1 - \pi_i) (1 - 2\pi_i) \{1 + (n_i - 1) \phi \} \{1 + (2n_i - 1) \phi \}}{n_i^2 (1 + \phi)} \]

\[ E(z_i - \pi_i)^4 = \left\{ n_i \pi_i (1 - \pi_i) \{1 + (n_i - 1) \phi \} \times \left[ (1 + (2n_i - 1) \phi) \frac{1 + (3n_i - 1) \phi}{(1 - \phi)} (1 - 3\pi_i (1 - \pi_i)) + \right. \right. \]
\[ \left. \left. \{n_i - 1\} (\phi + 3\pi_i (1 - \pi_i) \{1 + (n_i - 1) \phi \}) \right\} \frac{(1 - \phi)}{(1 + \phi)(1 + 2\phi) n_i^4} \right. \]

The elements of the matrices \( A(\tilde{\lambda}) \) and \( B(\tilde{\lambda}) \) are

\[ A = \begin{bmatrix} A & 0 \\ a' & a_{k+1} \end{bmatrix} \]

and

\[ B = \begin{bmatrix} A & b \\ b' & b_{k+1} \end{bmatrix} \]

where \( a = (a_1, \ldots, a_k) \) and \( b = (b_1, \ldots, b_k) \).

Therefore,

\[ \text{Var}(\lambda_{QM}) = \begin{bmatrix} A & 0 \\ a' & a_{k+1} \end{bmatrix}^{-1} \begin{bmatrix} A & b \\ b' & b_{k+1} \end{bmatrix} \begin{bmatrix} A & 0 \\ a' & a_{k+1} \end{bmatrix}^{-1} \]

\[ = \begin{bmatrix} \frac{1}{a_{k+1}} (b - a)' A^{-1} \frac{1}{a_{k+1}^2} \{b_{k+1} - 2b' A^{-1} a + a' A^{-1} a\} \end{bmatrix} \]

2.4.3 Variance of Extended Quasi-likelihood Estimator \( \hat{\lambda}_{EQE} \)

The asymptotic variance-covariance of the estimators \( \hat{\lambda}_{EQE} \) is obtained using a result by Inagaki (1973) which was described in section (2.2). The entries of the
matrices $A$ and $B$ are:

$$A_{j,s} = E\left( \frac{-\partial U_j}{\partial \beta_s} \right) = \sum_{i=1}^{m} \frac{n_i \pi_i (1 - \pi_i) X_{ij} X_{is}}{1 + (n_i - 1) \phi}, \quad j, s = 1, \ldots, k$$

$$A_{j,k+1} = E\left( \frac{-\partial U_j}{\partial \phi} \right) = 0, \quad j = 1, \ldots, k$$

$$A_{k+1,j} = E\left( \frac{-\partial U_{k+1}}{\partial \beta_j} \right) = 0, \quad j = 1, \ldots, k$$

$$A_{k+1,k+1} = E\left( \frac{-\partial U_{k+1}}{\partial \phi} \right) \approx \sum_{i=1}^{m} \left\{ \frac{(n_i - 1)}{(1 + (n_i - 1) \phi)} \right\}^2 = c_{k+1}$$

$$B_{j,s} = E(U_j U_s) = \sum_{i=1}^{m} \frac{n_i \pi_i (1 - \pi_i) X_{ij} X_{is}}{1 + (n_i - 1) \phi} = A_{j,s},$$

$$B_{j,k+1} = E(U_j U_{k+1}) = \sum_{i=1}^{m} \frac{(n_i - 1)(1 - 2\pi_i)(1 + (2n_i - 1) \phi) X_{ij}}{(1 + \phi) \pi_i (1 - \pi_i)(1 + (n_i - 1) \phi)^2}, \quad j = 1, \ldots, k$$

$$B_{k+1,j} = B_{j,k+1}.$$

$$B_{k+1,k+1} = E(U_{k+1}^2)$$

$$= \sum_{i=1}^{m} \left\{ \frac{n_i^2 (n_i - 1)^2 E(z_i - \pi_i)^4}{[n_i \pi_i (1 - \pi_i)]^2 (1 + (n_i - 1) \phi)^2} - \frac{(n_i - 1)^2}{(1 + (n_i - 1) \phi)^2} \right\}.$$

The elements of the matrices $A \left( \hat{\lambda} \right)$ and $B \left( \hat{\lambda} \right)$ are

$$A = \begin{bmatrix} A & 0 \\ 0 & a_{k+1} \end{bmatrix}$$

and

$$B = \begin{bmatrix} A & b \\ b & b_{k+1} \end{bmatrix}.$$ 

Therefore

$$\text{Var} \left( \hat{\lambda}_{E|E} \right) = \begin{bmatrix} A & 0 \\ 0 & a_{k+1} \end{bmatrix}^{-1} \begin{bmatrix} A & b \\ b & b_{k+1} \end{bmatrix} \left\{ \begin{bmatrix} A & 0 \\ 0 & a_{k+1} \end{bmatrix}^{-1} \right\}^T$$

$$= \begin{bmatrix} A^{-1} & \frac{1}{a_{k+1}} b A^{-1} \\ \frac{1}{a_{k+1}} b A^{-1} & \frac{b_{k+1}}{a_{k+1}^2} A^{-1} \end{bmatrix}.$$
2.4.4 Variance of Gaussian Likelihood Estimator $\hat{\lambda}_{GL}$

The asymptotic variance-covariance of the estimators $\hat{\lambda}_{GL}$ is obtained using a result of Inagaki (1973) as described in section (2.2). The entries of the matrices $A$ and $B$ are:

$$A_{js} = E \left( \frac{-\partial U_j}{\partial \beta_s} \right) = \sum_{i=1}^{m} \left\{ n_i \left[ \frac{1}{1 + (n_i - 1) \phi} + \frac{(1 - 2\pi_i)^2}{2\pi_i(1 - \pi_i)} \right] \pi_i (1 - \pi_i) X_{ij} X_{is} \right\},$$

$s, j = 1, \ldots, k$

$$A_{j,k+1} = E \left( \frac{-\partial U_j}{\partial \phi} \right) = \sum_{i=1}^{m} \frac{-n_i(n_i - 1)(1 - 2\pi_i)\phi X_{ij}}{2(1 - \phi)\{1 + (n_i - 1) \phi\}} = a_j, \quad j = 1, \ldots, k$$

$$A_{k+1,j} = E \left( \frac{-\partial U_{k+1}}{\partial \beta_j} \right) = \sum_{i=1}^{m} \frac{-n_i(n_i - 1)(1 - 2\pi_i)\phi X_{ij}}{2(1 - \phi)\{1 + (n_i - 1) \phi\}} = a_j, \quad j = 1, \ldots, k$$

$$A_{k+1,k+1} = E \left( \frac{-\partial U_{k+1}}{\partial \phi} \right) = \sum_{i=1}^{m} \frac{n_i(n_i - 1)(1 + (n_i - 1) \phi^2)}{2[(1 - \phi)\{1 + (n_i - 1) \phi\}]^2},$$

$$B_{js} = E (U_j U_s) = \sum_{i=1}^{m} \left\{ \left[ \frac{(1 - \phi)}{[1 + (n_i - 1) \phi]} + \frac{(1 - 2\pi_i)^2}{2\pi_i(1 - \pi_i)} \right] \left[ \frac{n_i}{\pi_i (1 - \pi_i)(1 - \phi)} + \frac{n_i(1 - 2\pi_i)^2 \{1 + (n_i - 1) \phi\}}{2 [\pi_i(1 - \pi_i)(1 - \phi)]^2} \right] \right. \right.$$

$$\left. \left. \left. - \frac{n_i(1 - 2\pi_i)^2 \phi (1 + (2n_i - 1) \phi)}{[\pi_i(1 - \pi_i)(1 - \phi)]^2 (1 + \phi)} \right) \right\}$$

$$+ \left[ \frac{(1 - 2\pi_i) \phi n_i}{2\pi_i(1 - \pi_i)(1 - \phi)} \right]^2 \left[ \frac{n_i^2 E (z_i - \pi_i)^4}{[\pi_i(1 - \pi_i)\{1 + (n_i - 1) \phi\}]^2} \right] - 1 \right\} \pi_i^2 (1 - \pi_i)^2 X_{ij} X_{is},$$

$j, s = 1, \ldots, k$
\[ b_j = B_{j,k+1} = E(U_j U_{k+1}) \]
\[ = \sum_{i=1}^{m} \left\{ \frac{(1 - 2 \pi_i)(1 + (2n_i - 1) \phi)}{(1 - \phi^2)} G \times D \frac{n_i}{2(1 - \phi)^2} \left[ \frac{n_i^2 E(z_i - \pi_i)^4}{[\pi_i(1 - \pi_i)(1 + (n_i - 1) \phi)]^2} \right] - \frac{n_i(1 - 2 \pi_i)(1 + (n_i - 1) \phi)}{2(1 - \phi)^3} \times G \right. \]
\[ - D \times \frac{n_i^2(1 - 2 \pi_i)}{4(1 - \phi)^3} \left[ \frac{n_i^2 E(z_i - \pi_i)^4}{[\pi_i(1 - \pi_i)(1 + (n_i - 1) \phi)]^2} - 1 \right] \left[ \frac{n_i(1 - 2 \pi_i)(1 + (n_i - 1) \phi)}{4\pi_i(1 - \pi_i)(1 - \phi)^3 (1 + \phi)} \right] \left\} \times X_{ij}, \quad j = 1, \ldots, k \]

\[ b_{k+1} = B_{k+1,k+1} = E(U_{k+1}^2) \]
\[ = \sum_{i=1}^{m} \left\{ \frac{n_i^2}{4(1 - \phi)^8} \times D^2 \left[ \frac{n_i^2 E(z_i - \pi_i)^4}{[\pi_i(1 - \pi_i)(1 + (n_i - 1) \phi)]^2} - 1 \right] - \frac{n_i(1 - 2 \pi_i)^2(1 + (2n_i - 1) \phi)}{2\pi_i(1 - \pi_i)(1 - \phi)^4(1 + \phi)} D \right. \]
\[ + \frac{n_i(1 - 2 \pi_i)^2(1 + (n_i - 1) \phi)}{4\pi_i(1 - \pi_i)(1 - \phi)^3 (1 + \phi)} \right\}, \]

where

\[ G = \left[ \frac{(1 - \phi)}{(1 + (n_i - 1) \phi)} + \frac{(1 - 2 \pi_i)^2}{2\pi_i(1 - \pi_i)} \right] \]
and
\[ D = \frac{(1 + (n_i - 1) \phi^2)}{(1 + (n_i - 1) \phi)}. \]

The elements of the matrices \( A(\lambda) \) and \( B(\lambda) \) are

\[ A = \begin{bmatrix} A & a \\ a' & a_{k+1} \end{bmatrix} \]
\[ B = \begin{bmatrix} B & b \\ b' & b_{k+1} \end{bmatrix}. \]

Therefore

\[ \text{Var}(\hat{\lambda}_{GL}) = \begin{bmatrix} A & a \\ a' & a_{k+1} \end{bmatrix}^{-1} \begin{bmatrix} B & b \\ b' & b_{k+1} \end{bmatrix} \left\{ \begin{bmatrix} A & a \\ a' & a_{k+1} \end{bmatrix} \right\}^T. \]

2.4.5 Variance of Quadratic Estimator \( \hat{\lambda}_{QEE} \)
Similarly, the asymptotic variance-covariance matrix of the estimator \( \hat{\Lambda}_{QEE} \) is obtained using a Inagaki (1973) result. The entries of the matrices \( A \) and \( B \) are:

\[
A_{js} = \left( \frac{-\partial \hat{u}_j}{\partial \beta_s} \right) = \sum_{i=1}^{m} \frac{n_i \pi_i (1 - \pi_i) X_{ij} \hat{x}_{is}}{1 + (n_i - 1) \hat{\phi}}. \quad j, s = 1, \ldots, k
\]

\[
A_{j,k+1} = E \left( \frac{-\partial \hat{u}_j}{\partial \hat{\phi}} \right) = 0. \quad j = 1, \ldots, k
\]

\[
a_j = A_{k+1,j} = E \left( \frac{-\partial \hat{u}_{k+1}}{\partial \beta_j} \right) = \sum_{i=1}^{m} \frac{(1 - 2 \pi_i) (n_i - 1) X_{ij}}{1 + (n_i - 1) \hat{\phi}}. \quad j = 1, \ldots, k
\]

\[
a_{k+1} = A_{k+1,k+1} = E \left( \frac{-\partial \hat{u}_{k+1}}{\partial \hat{\phi}} \right) = \sum_{i=1}^{m} \frac{(n_i - 1)^2}{1 + (n_i - 1) \hat{\phi}}.
\]

\[
B_{js} = E (U_j U_s) = \sum_{i=1}^{m} \frac{n_i \pi_i (1 - \pi_i) X_{ij} \hat{x}_{is}}{1 + (n_i - 1) \hat{\phi}} = A_{js}. \quad j, s = 1, \ldots, k
\]

\[
b_j = B_{j,k+1} = E(U_j U_{k+1}) = \sum_{i=1}^{m} \frac{(n_i - 1) (1 - 2 \pi_i) (1 + (2n_i - 1) \hat{\phi}) X_{ij}}{(1 + \hat{\phi}) (1 + (n_i - 1) \hat{\phi})} = A_{k+1,j}
\]

\[
B_{k+1,k+1} = E(U_{k+1}) = \sum_{i=1}^{m} \frac{ \left( \frac{n_i (n_i - 1)}{\pi_i (1 - \pi_i)} \right)^2 \frac{E(z_i - \pi_i)^4}{\{1 + (n_i - 1) \hat{\phi}\}^4} - \left( \frac{n_i - 1}{1 + (n_i - 1) \hat{\phi}} \right)^2}{\{1 + (n_i - 1) \hat{\phi}\}^2}.
\]

The elements of the matrices \( A(\hat{\Lambda}) \) and \( B(\hat{\Lambda}) \) are

\[
A = \begin{bmatrix} A & 0 \\ a' & a_{k+1} \end{bmatrix}
\]

and

\[
B = \begin{bmatrix} A & b \\ b' & b_{k+1} \end{bmatrix}
\]

Therefore,

\[
\text{Var} (\hat{\Lambda}_{QEE}) = \begin{bmatrix} A & 0 \\ a' & a_{k+1} \end{bmatrix}^{-1} \begin{bmatrix} A & b \\ b' & b_{k+1} \end{bmatrix} \left\{ \begin{bmatrix} A & 0 \\ a' & a_{k+1} \end{bmatrix}^{-1} \right\}^T
\]

\[
= \frac{1}{a_{k+1}} \begin{bmatrix} (b - a)' A^{-1} \\ a' \end{bmatrix} \frac{1}{a_{k+1}} \begin{bmatrix} b_{k+1} - 2b' A^{-1} a + a' A^{-1} a \end{bmatrix}
\]
2.4.6 Variance of Moment Estimators

(i) \( \text{Var}(\hat{\lambda}_{M_1}) \)

The asymptotic variance covariance matrix of the estimators \( \hat{\lambda}_{M_1} \) is obtained using a result by Inagaki (1973).

The entries of the matrices A and B are

\[
A_{js} = E \left( \frac{-\partial U_j}{\partial \beta_s} \right) = \sum_{i=1}^{m} \frac{n_i \pi_i (1 - \pi_i) X_{ij} X_{is}}{1 + (n_i - 1) \phi}, \quad s, j = 1, \ldots, k
\]

\[
A_{jk+1} = E \left( \frac{-\partial U_j}{\partial \phi} \right) = 0, \quad j = 1, \ldots, k
\]

\[
a_j = A_{k+1,j} = E \left( \frac{-\partial U_{k+1}}{\partial \beta_j} \right) = \sum_{i=1}^{m} (1 - 2\pi_i) X_{ij}, \quad j = 1, \ldots, k
\]

\[
a_{k+1} = A_{k+1,k+1} = E \left( \frac{-\partial U_{k+1}}{\partial \phi} \right) = \sum_{i=1}^{m} \frac{(1 - 2\pi_i)}{\pi_i (1 - \pi_i)}
\]

\[
B_{js} = E(U_j U_s) = \sum_{i=1}^{m} \frac{n_i \pi_i (1 - \pi_i) X_{ij} X_{is}}{1 + (n_i - 1) \phi}, \quad j, s = 1, \ldots, k
\]

\[
b_j = B_{j,k+1} = E(U_j U_{k+1}) = \sum_{i=1}^{m} \frac{(1 - 2\pi_i)(1 + (2n_i - 1) \phi) X_{ij}}{(1 + \phi) \{1 + (n_i - 1) \phi\}}, \quad j = 1, \ldots, k.
\]

To find \( E(U^2_{k+1}) \), we rewrite

\[
U_{k+1}(\beta, \phi) = \sum_{i=1}^{m} \left\{ \frac{n_i (\bar{z}_i - \pi_i)^2}{\pi_i (1 - \pi_i) \{1 + (n_i - 1) \phi\}} - C_i \right\}.
\]

where

\[
C_i = \phi \left\{ \frac{(n_i - 1)}{\{1 + (n_i - 1) \phi\}} \right\} - \frac{n_i (n_i - 1)}{\pi_i (1 - \pi_i) \{1 + (n_i - 1) \phi\}} V^{-1}
\]

\[
+ \left\{ \frac{1}{\{1 + (n_i - 1) \phi\}} - \frac{n_i}{\pi_i (1 - \pi_i) \{1 + (n_i - 1) \phi\}} \right\} V^{-1}
\]

and

\[
V = \sum_{i=1}^{m} \frac{n_i}{\pi_i (1 - \pi_i) \{1 + (n_i - 1) \phi\}}.
\]
Therefore
\[ B_{k+1,k+1} = E(U_{k+1}^2) = \sum_{i=1}^{m} \left\{ \frac{n_i^2 E(z_i - \pi_i)^4}{[\pi_i(1 - \pi_i)(1 + (n_i - 1)\phi)]^2} - 2C_i + C_i^2 \right\} \]
\[ = b_{k+1} \]

and
\[ \text{Var}(\hat{\lambda}_{M1}) = \begin{bmatrix} A^{-1} & \frac{1}{a_{k+1}} A^{-1} (b - a) \\ \frac{1}{a_{k+1}} (b - a)' A^{-1} & \frac{1}{a_{k+1}^2} \{ b_{k+1} - 2 \frac{b'}{a} A^{-1} a + a' A^{-1} a \} \end{bmatrix} \].

(ii) Var(\hat{\lambda}_{M2})

The asymptotic variance-covariance matrix of the estimator \( \hat{\lambda}_{M2} \) is obtained using a result by Inagaki (1973). The entries of the matrices A and B are:

\[ A_{js} = E\left(-\frac{\partial U_j}{\partial \beta_s}\right) = \sum_{i=1}^{m} n_i \pi_i^3 (1 - \pi_i)^2 X_{ij} X_{is}, \quad j, s = 1, \ldots, k \]

\[ A_{jk+1} = E\left(-\frac{\partial U_j}{\partial \phi}\right) = 0, \quad j = 1, \ldots, k \]

\[ A_{k+1,j} = E\left(-\frac{\partial U_{k+1}}{\partial \beta_j}\right) = \sum_{i=1}^{m} (1 - 2\pi_i) (1 + (n_i - 1)\phi) X_{ij}, \quad j = 1, \ldots, k \]

\[ = a_{k+1, j} \]

\[ A_{k+1,k+1} = E\left(-\frac{\partial U_{k+1}}{\partial \phi}\right) = \sum_{i=1}^{m} (n_i - 1), \]

\[ = a_{k+1} \]

\[ B_{js} = E(U_j U_s) = \sum_{i=1}^{m} n_i \pi_i^3 (1 - \pi_i)^3 \{ 1 + (n_i - 1)\phi \} X_{ij} X_{is}, \quad j, s = 1, \ldots, k \]

\[ B_{j,k+1} = E(U_j U_{k+1}) = \sum_{i=1}^{m} \frac{\pi_i (1 - \pi_i)(1 - 2\pi_i) (1 + (n_i - 1)\phi) \{ 1 + (2n_i - 1)\phi \} X_{ij}}{(1 + \phi)} \]

\[ = b_j = B_{j,k+1}, \quad j = 1, \ldots, k \]

and

\[ B_{k+1,k+1} = E(U_{k+1}^2) = \sum_{i=1}^{k} \left\{ \frac{n_i^2 E(z_i - \pi_i)^4}{\pi_i^2 (1 - \pi_i)^2} - \{ 1 + (n_i - 1)\phi \}^2 \right\} = b_{k+1} \].
The elements of the matrices \( A(\hat{\lambda}) \) and \( B(\hat{\lambda}) \) are

\[
A = \begin{bmatrix}
A & 0 \\
a' & a_{k+1}
\end{bmatrix}
\]

and

\[
B = \begin{bmatrix}
B & b \\
b & b_{k+1}
\end{bmatrix}.
\]

\[
\text{Var}(\hat{\lambda}_{M2}) = \begin{bmatrix}
A^{-1}B\{A^{-1}\}' & \frac{1}{a_{k+1}} (A^{-1}b - aA^{-1}B\{A^{-1}\}') \\
\frac{1}{a_{k+1}} (A^{-1}b - aA^{-1}B\{A^{-1}\})' & \frac{1}{a_{k+1}^2} \{b_{k+1} - 2b' A^{-1}a + a'A^{-1}BA^{-1}a\}
\end{bmatrix}.
\]

(iii) \( \text{Var}(\hat{\lambda}_{M3}) \)

The asymptotic variance-covariance matrix of the estimator \( \hat{\lambda}_{M3} \) is obtained using a result by Inagaki (1973). The entries of the matrices \( A \) and \( B \) are:

\[
A_{js} = E\left( \frac{-\partial U_j}{\partial \beta_s} \right) = \sum_{i=1}^{m} \frac{n_i \pi_i (1 - \pi_i) X_{ij} X_{is}}{1 + (n_i - 1) \phi}, \quad j, s = 1, \ldots, k
\]

\[
A_{j,k+1} = E\left( \frac{-\partial U_j}{\partial \phi} \right) = 0, \quad j = 1, \ldots, k
\]

\[
A_{k+1,j} = E\left( \frac{-\partial U_{k+1}}{\partial \beta_j} \right) = \sum_{i=1}^{m} (1 - 2n_i) \{1 + (n_i - 1) \phi\} X_{ij},
\]

\[
= a_j, \quad j = 1, \ldots, k
\]

\[
A_{k+1,k+1} = E\left( \frac{-\partial U_{k+1}}{\partial \phi} \right) = \sum_{i=1}^{m} (n_i - 1),
\]

\[
= a_{k+1},
\]

\[
B_{js} = E(U_j U_s) = \sum_{i=1}^{m} \frac{n_i \pi_i (1 - \pi_i) X_{ij} X_{is}}{1 + (n_i - 1) \phi} = A_{js}, \quad j, s = 1, \ldots, k
\]

\[
b_j = B_{j,k+1} = E(U_j U_{k+1}) = \sum_{i=1}^{m} \frac{(1 + (2n_i - 1) \phi) X_{ij}}{(1 + \phi)}, \quad j = 1, \ldots, k
\]

\[
b_{k+1} = B_{k+1,k+1} = E(U_{k+1}^2) = \sum_{i=1}^{m} \left\{ \frac{n_i^2 E(\pi_i - \pi_i)^4}{\pi_i^2 (1 - \pi_i)^2} - (1 + (n_i - 1) \phi)^2 \right\}.
\]

Therefore

\[
\text{Var}(\hat{\lambda}_{M3}) = \begin{bmatrix}
A^{-1} & \frac{1}{a_{k+1}} A^{-1} (b - a) \\
\frac{1}{a_{k+1}} A^{-1} (b - a) & \frac{1}{a_{k+1}^2} \{b_{k+1} - 2b' A^{-1}a + a'A^{-1}a\}
\end{bmatrix}.
\]
(iv) \( \text{Var}(\widehat{\Lambda}_{M4}) \)

The asymptotic variance covariance matrix of the estimator \( \widehat{\Lambda}_{M4} \) is obtained using a result by Inagaki (1973). The entries of the matrices \( A \) and \( B \) are:

\[
A_{j,s} = E \left( \frac{-\partial U_j}{\partial \beta_s} \right) = \sum_{i=1}^{m} \frac{n_i \pi_i (1 - \pi_i) X_{ij} X_{is}}{1 + (n_i - 1) \phi}, \quad j, s = 1, \ldots, k
\]

\[
A_{j,k+1} = E \left( \frac{-\partial U_j}{\partial \phi} \right) = 0, \quad j = 1, \ldots, k
\]

\[
A_{k+1,j} = E \left( \frac{-\partial U_{k+1}}{\partial \beta_j} \right) = \sum_{i=1}^{m} n_i (1 - 2\pi_i) \{1 + (n_i - 1) \phi\} X_{ij}
\]

\[
= a_j, \quad j = 1, \ldots, k
\]

\[
A_{k+1,k+1} = E \left( \frac{-\partial U_{k+1}}{\partial \phi} \right) = \sum_{i=1}^{m} n_i (n_i - 1)
\]

\[
= a_{k+1}.
\]

\[
B_{j,s} = E(U_j U_s) = \sum_{i=1}^{m} \frac{n_i \pi_i (1 - \pi_i) X_{ij} X_{is}}{1 + (n_i - 1) \phi}, \quad j, s = 1, \ldots, k
\]

\[
B_{j,k+1} = E(U_j U_{k+1}) = \sum_{i=1}^{m} \frac{n_i (1 - 2\pi_i) \{1 + (2n_i - 1) \phi\}}{(1 + \phi)} X_{ij}
\]

\[
= b_j, \quad j = 1, \ldots, k
\]

\[
B_{k+1,k+1} = E(U_{k+1}^2) = \sum_{i=1}^{m} \left\{ \frac{n_i^2 E(z_i - \pi_i)^4}{\pi_i^2 (1 - \pi_i)^2} - n_i^2 \{1 + (n_i - 1) \phi\}^2 \right\}
\]

\[
= b_{k+1},
\]

and therefore

\[
\text{Var}(\widehat{\Lambda}_{M4}) = \begin{bmatrix}
A^{-1} & \frac{1}{a_{k+1}} A^{-1} (b - a) \\
\frac{1}{a_{k+1}} (b - a)' A^{-1} & \frac{1}{a_{k+1}^2} \left\{ b_{k+1} - 2b' A^{-1} a + a' A^{-1} a \right\}
\end{bmatrix}
\]

2.5 Efficiency

In this section we compare asymptotic and small sample efficiencies of the moment and other estimates based on semi-parametric models relative to that of maximum likelihood estimates based on the extended beta-binomial model. In what follows we study large and small sample efficiency.
2.5.1 Asymptotic Relative Efficiency

In section 2.4 we derived the asymptotic variance-covariance matrix of various estimates for the general model involving regression parameters. However, for simplicity, numerical efficiency calculations are performed for the two parameter model. Note that $\text{Var}(\widehat{\pi}_{QM})$ is the same as that of $\text{Var}(\widehat{\pi}_{EQM})$, $\text{Var}(\widehat{\pi}_{QEE})$, $\text{Var}(\widehat{\pi}_{M1})$, $\text{Var}(\widehat{\pi}_{M3})$ and $\text{Var}(\widehat{\pi}_{M4})$. Thus, for $\pi$ we have three distinct relative efficiency comparisons to make. Let $\text{Var}(\widehat{\pi}_{ML})$ be the asymptotic variance of the maximum likelihood estimate $\widehat{\pi}$ obtained from (2.4.1) and let $\text{Var}(\widehat{\pi}^{(t)})$ be the asymptotic variance of $\widehat{\pi}^{(t)}$ obtained by using Inagaki (1973) results, where $t$ represents the three methods $QM$, GL and M2. Then the asymptotic relative efficiency of $\widehat{\pi}^{(t)}$ is given by $\text{var}(\widehat{\pi}_{ML})/\text{var}(\widehat{\pi}^{(t)})$. The asymptotic relative efficiency of $\widehat{\phi}^{(t)}$ is given by $\text{var}(\widehat{\phi}_{ML})/\text{var}(\widehat{\phi}^{(t)})$, where $t$ represents the methods $QM$, EQE, GL, QEE, M1, M2, M3, and M4. Because of the asymptotic nature of the efficiency results which are valid when $m \rightarrow \infty$, we consider litter sizes obtained from a real life experiment (Potthoff and Whittinghill, 1966) in which $m = 36$ is reasonably large. Thus, we consider the litter sizes $n_i; \{11, 1, 6, 7, 8, 6, 2, 19, 4, 2, 15, 6, 6, 10, 8, 4, 5, 6, 6, 4, 12, 8, 4, 5, 6, 4, 10, 8, 11, 4, 4, 4, 2, 2, 3\}$. In order to investigate the effect of $\pi$ and $\phi$ on the relative efficiency of $\widehat{\pi}$ and $\widehat{\phi}$, three sets of a combinations of the parameters $\pi$ and $\phi$ were chosen: $\phi = 0.10, 0.02, 0.06, 0.10, 0.15, 0.20, 0.30, 0.40, 0.50; \pi = 0.10, 0.02, 0.06, 0.10, 0.15, 0.20, 0.30, 0.40, 0.50; \pi = 0.40, 0.02, 0.06, 0.10, 0.15, 0.20, 0.30, 0.40, 0.50$. The efficiency results for $\widehat{\pi}$ are given in table 2.1 and plotted in figure 2.5.1 and those for $\widehat{\phi}$ are given in table 2.2 and plotted in figure 2.5.2.
We first discuss the efficiency of $\hat{\pi}^{(i)}$. $t = QM, GL, M2$ from figure 2.5.1 We observe the following:

(a) The efficiency of $\hat{\pi}_{QM}$ (as well as the efficiencies of $\hat{\pi}_{EQE}$, $\hat{\pi}_{QEE}$, $\hat{\pi}_{M1}$, $\hat{\pi}_{M3}$ and $\hat{\pi}_{M4}$) is the best (highest) and close to one.

(b) the Gaussian estimate $\hat{\pi}_{GL}$ also is highly efficient, although somewhat less efficient than $\hat{\pi}_{QM}$. This loses some efficiency when $\pi$ becomes small. The value of $\phi$ does not seem to have much effect on the efficiency.

(c) The estimate $\hat{\pi}_{M2}$ is the least efficient. The efficiency decreases as $\phi$ increases. However, the value of $\pi$ does not seem to have any effect on the efficiency.

Next, we discuss efficiency results for $\hat{\phi}^{(i)}$. $t = QM, EQM, GL, QEE, M1, M2, M3, M4$ from figure 2.5.2 We observe the following:

(a) The Gaussian estimator $\hat{\phi}_{GL}$ performs best overall and has high efficiency. For the cases studied the lowest efficiency found was 84%, which occurred when $\pi$ was very small.

(b) The second best estimate seems to be $\hat{\phi}_{QEE}$, obtained using the quadratic estimating equations. For large $\pi$ and $\phi$ this estimate performs best.

(c) Except when $\phi$ is small (less than 0.10) the estimator $\hat{\phi}_{M2}$ does quite well. In some situations, particularly, for $\pi = 0.10$ and $\phi > 0.20$, the efficiency marginally exceeds those of other estimates.

(d) In general, the remaining three estimators perform poorly.

(e) In general, there is some evidence that the efficiency of the estimators increases as $\pi$ increases.
Table 2.1(a): Asymptotic Relative Efficiencies of $\pi_{QM}, \pi_{M2}$ and $\pi_{GL}$ for $\phi = 0.10$ and $m = 36$ under Beta-binomial distribution.

<table>
<thead>
<tr>
<th>$\pi$</th>
<th>Method</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quasi</td>
<td>Moments 2</td>
<td>Gaussian</td>
<td></td>
</tr>
<tr>
<td>0.02</td>
<td>0.9960</td>
<td>0.9443</td>
<td>0.9580</td>
<td></td>
</tr>
<tr>
<td>0.06</td>
<td>0.9959</td>
<td>0.9443</td>
<td>0.9622</td>
<td></td>
</tr>
<tr>
<td>0.10</td>
<td>0.9959</td>
<td>0.9442</td>
<td>0.9668</td>
<td></td>
</tr>
<tr>
<td>0.15</td>
<td>0.9958</td>
<td>0.9442</td>
<td>0.9728</td>
<td></td>
</tr>
<tr>
<td>0.20</td>
<td>0.9958</td>
<td>0.9442</td>
<td>0.9790</td>
<td></td>
</tr>
<tr>
<td>0.30</td>
<td>0.9958</td>
<td>0.9441</td>
<td>0.9897</td>
<td></td>
</tr>
<tr>
<td>0.40</td>
<td>0.9958</td>
<td>0.9441</td>
<td>0.9950</td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td>0.9957</td>
<td>0.9441</td>
<td>0.9957</td>
<td></td>
</tr>
</tbody>
</table>
Table 2.1(b): Asymptotic Relative Efficiency of $\pi_{QM}$, $\pi_{M2}$ and $\pi_{GL}$ for $\pi = 0.10$ and $m = 36$ under Beta-binomial distribution.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>Quasi</th>
<th>Moments 2</th>
<th>Gaussian</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.02</td>
<td>0.9957</td>
<td>0.9901</td>
<td>0.9923</td>
</tr>
<tr>
<td>0.06</td>
<td>0.9959</td>
<td>0.9674</td>
<td>0.9788</td>
</tr>
<tr>
<td>0.10</td>
<td>0.9959</td>
<td>0.9442</td>
<td>0.9668</td>
</tr>
<tr>
<td>0.15</td>
<td>0.9954</td>
<td>0.9195</td>
<td>0.9561</td>
</tr>
<tr>
<td>0.20</td>
<td>0.9944</td>
<td>0.8991</td>
<td>0.9493</td>
</tr>
<tr>
<td>0.30</td>
<td>0.9916</td>
<td>0.8678</td>
<td>0.9422</td>
</tr>
<tr>
<td>0.40</td>
<td>0.9880</td>
<td>0.8446</td>
<td>0.9391</td>
</tr>
<tr>
<td>0.50</td>
<td>0.9843</td>
<td>0.8267</td>
<td>0.9377</td>
</tr>
</tbody>
</table>
Table 2.1(c): Asymptotic Relative Efficiency of $\pi_{QM}$, $\pi_{M2}$ and $\pi_{GL}$ for $\pi = 0.40$ and $m = 36$ under Beta-binomial distribution.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>Quasi</th>
<th>Moments 2</th>
<th>Gaussian</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.02</td>
<td>0.9957</td>
<td>0.9901</td>
<td>0.9957</td>
</tr>
<tr>
<td>0.06</td>
<td>0.9959</td>
<td>0.9673</td>
<td>0.9956</td>
</tr>
<tr>
<td>0.10</td>
<td>0.9958</td>
<td>0.9441</td>
<td>0.9950</td>
</tr>
<tr>
<td>0.15</td>
<td>0.9951</td>
<td>0.9192</td>
<td>0.9940</td>
</tr>
<tr>
<td>0.20</td>
<td>0.9939</td>
<td>0.8986</td>
<td>0.9918</td>
</tr>
<tr>
<td>0.30</td>
<td>0.9905</td>
<td>0.8668</td>
<td>0.9873</td>
</tr>
<tr>
<td>0.40</td>
<td>0.9864</td>
<td>0.8432</td>
<td>0.9826</td>
</tr>
<tr>
<td>0.50</td>
<td>0.9823</td>
<td>0.8250</td>
<td>0.9782</td>
</tr>
</tbody>
</table>
ASYMPTOTIC RELATIVE EFFICIENCY OF $\widehat{\pi}$ VS $\pi$
FOR $\phi = 0.10$ AND $M = 36$.

Figure 2.5.1(a)
ASYMPTOTIC RELATIVE EFFICIENCY OF $\hat{r}$ VS $\phi$
FOR $\pi = 0.10$ AND $M = 36$.

Figure 2.5.1(b)
ASYMPTOTIC RELATIVE EFFICIENCY OF $\hat{\pi}$ VS $\phi$
FOR $\pi = 0.40$ AND $M = 36$.

Figure 2.5.1(c)
Table 2.2(a): Asymptotic Relative Efficiencies of $\phi_{GL}$, $\phi_{QEE}$, $\phi_{QM}$, $\phi_{EQE}$, $\phi_{M1}$, $\phi_{M2}$, $\phi_{M3}$ and $\phi_{M4}$ for $\phi=0.10$ and $m = 36$ under Beta-binomial distribution.

<table>
<thead>
<tr>
<th>$\pi$</th>
<th>GL</th>
<th>QEE</th>
<th>QM</th>
<th>EQE</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.02</td>
<td>.8403</td>
<td>.7718</td>
<td>.6052</td>
<td>.2639</td>
<td>.5054</td>
<td>.7386</td>
<td>.8060</td>
<td>.4533</td>
</tr>
<tr>
<td>0.06</td>
<td>.8736</td>
<td>.8231</td>
<td>.6733</td>
<td>.3498</td>
<td>.6737</td>
<td>.7965</td>
<td>.8468</td>
<td>.5286</td>
</tr>
<tr>
<td>0.10</td>
<td>.9005</td>
<td>.8626</td>
<td>.7263</td>
<td>.4371</td>
<td>.7270</td>
<td>.8408</td>
<td>.8787</td>
<td>.5906</td>
</tr>
<tr>
<td>0.15</td>
<td>.9266</td>
<td>.8999</td>
<td>.7771</td>
<td>.5451</td>
<td>.7781</td>
<td>.8827</td>
<td>.9090</td>
<td>.6528</td>
</tr>
<tr>
<td>0.20</td>
<td>.9462</td>
<td>.9273</td>
<td>.8151</td>
<td>.6480</td>
<td>.8162</td>
<td>.9135</td>
<td>.9315</td>
<td>.7011</td>
</tr>
<tr>
<td>0.30</td>
<td>.9724</td>
<td>.9627</td>
<td>.8647</td>
<td>.8254</td>
<td>.8661</td>
<td>.9533</td>
<td>.9605</td>
<td>.7669</td>
</tr>
<tr>
<td>0.40</td>
<td>.9866</td>
<td>.9807</td>
<td>.8902</td>
<td>.9443</td>
<td>.8918</td>
<td>.9736</td>
<td>.9753</td>
<td>.8020</td>
</tr>
<tr>
<td>0.50</td>
<td>.9912</td>
<td>.9862</td>
<td>.8981</td>
<td>.9862</td>
<td>.8997</td>
<td>.9798</td>
<td>.9798</td>
<td>.8130</td>
</tr>
</tbody>
</table>
Table 2.2(b): Asymptotic Relative Efficiencies of $\phi_{GL}$, $\phi_{QEE}$, $\phi_{QM}$, $\phi_{EQE}$, $\phi_{M}$, $\phi_{M2}$, $\phi_{M3}$ and $\phi_{M4}$ for $\pi=0.10$ and $m = 36$ under Beta-binomial distribution.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>GL</th>
<th>QEE</th>
<th>QM</th>
<th>EQE</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.02</td>
<td>.9781</td>
<td>.9563</td>
<td>.6350</td>
<td>.6045</td>
<td>.6356</td>
<td>.6867</td>
<td>.7126</td>
<td>.9306</td>
</tr>
<tr>
<td>0.06</td>
<td>.9304</td>
<td>.8964</td>
<td>.6938</td>
<td>.4961</td>
<td>.6945</td>
<td>.7966</td>
<td>.8445</td>
<td>.7298</td>
</tr>
<tr>
<td>0.10</td>
<td>.9005</td>
<td>.8626</td>
<td>.7263</td>
<td>.4371</td>
<td>.7270</td>
<td>.8408</td>
<td>.8787</td>
<td>.5906</td>
</tr>
<tr>
<td>0.15</td>
<td>.8797</td>
<td>.8412</td>
<td>.7551</td>
<td>.3917</td>
<td>.7558</td>
<td>.8623</td>
<td>.8695</td>
<td>.4825</td>
</tr>
<tr>
<td>0.20</td>
<td>.8688</td>
<td>.8309</td>
<td>.7777</td>
<td>.3912</td>
<td>.7784</td>
<td>.8696</td>
<td>.8417</td>
<td>.4136</td>
</tr>
<tr>
<td>0.30</td>
<td>.8598</td>
<td>.8229</td>
<td>.8127</td>
<td>.3172</td>
<td>.8135</td>
<td>.8715</td>
<td>.7774</td>
<td>.3299</td>
</tr>
<tr>
<td>0.40</td>
<td>.8568</td>
<td>.8191</td>
<td>.8382</td>
<td>.2855</td>
<td>.8391</td>
<td>.8685</td>
<td>.7181</td>
<td>.2790</td>
</tr>
<tr>
<td>0.50</td>
<td>.8552</td>
<td>.8147</td>
<td>.8571</td>
<td>.2599</td>
<td>.8581</td>
<td>.8642</td>
<td>.6662</td>
<td>.2435</td>
</tr>
</tbody>
</table>

41
Table 2.2(c): Asymptotic Relative Efficiencies of $\phi_{GL}$, $\phi_{QEE}$, $\phi_{QM}$, $\phi_{EQE}$, $\phi_{M1}$, $\phi_{M2}$, $\phi_{M3}$ and $\phi_{M4}$ for $\pi = 0.40$ and $m = 36$ under Beta-binomial distribution.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>GL</th>
<th>QEE</th>
<th>QM</th>
<th>EQE</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.02</td>
<td>0.9948</td>
<td>0.9892</td>
<td>0.8117</td>
<td>0.9739</td>
<td>0.8129</td>
<td>0.8609</td>
<td>0.8619</td>
<td>0.9700</td>
</tr>
<tr>
<td>0.06</td>
<td>0.9909</td>
<td>0.9846</td>
<td>0.8643</td>
<td>0.9584</td>
<td>0.8657</td>
<td>0.9477</td>
<td>0.9497</td>
<td>0.8791</td>
</tr>
<tr>
<td>0.10</td>
<td>0.9866</td>
<td>0.9807</td>
<td>0.8902</td>
<td>0.9443</td>
<td>0.8918</td>
<td>0.9736</td>
<td>0.9753</td>
<td>0.8020</td>
</tr>
<tr>
<td>0.15</td>
<td>0.9797</td>
<td>0.9760</td>
<td>0.9075</td>
<td>0.9274</td>
<td>0.9093</td>
<td>0.9761</td>
<td>0.9765</td>
<td>0.7306</td>
</tr>
<tr>
<td>0.20</td>
<td>0.9714</td>
<td>0.9716</td>
<td>0.9168</td>
<td>0.9114</td>
<td>0.9187</td>
<td>0.9674</td>
<td>0.9558</td>
<td>0.6783</td>
</tr>
<tr>
<td>0.30</td>
<td>0.9526</td>
<td>0.9535</td>
<td>0.9251</td>
<td>0.8814</td>
<td>0.9274</td>
<td>0.9429</td>
<td>0.9368</td>
<td>0.6070</td>
</tr>
<tr>
<td>0.40</td>
<td>0.9334</td>
<td>0.9565</td>
<td>0.9275</td>
<td>0.8542</td>
<td>0.9301</td>
<td>0.9200</td>
<td>0.9089</td>
<td>0.5600</td>
</tr>
<tr>
<td>0.50</td>
<td>0.9157</td>
<td>0.9504</td>
<td>0.9277</td>
<td>0.8293</td>
<td>0.9306</td>
<td>0.9008</td>
<td>0.8848</td>
<td>0.5260</td>
</tr>
</tbody>
</table>
ASYMPTOTIC RELATIVE EFFICIENCY OF $\hat{\phi}$ VS $\pi$
FOR $\phi = 0.10$ AND $M = 36$.

Figure 2.5.2(a)
ASYMPTOTIC RELATIVE EFFICIENCY OF $\hat{\phi}$ VS $\phi$
FOR $\pi = 0.10$ AND $M = 36$. 

Figure 2.5.2(b)
ASYMPTOTIC RELATIVE EFFICIENCY OF $\hat{\sigma}$ VS $\phi$
FOR $\pi = 0.40$ AND $M = 36$.

Figure 2.5.2(c)
2.5.2 Small Sample Relative Efficiency

The asymptotic relative efficiency may not be very useful when comparing different estimators in small samples so we conducted a simulation study taking $m$ reasonably small. The litter sizes considered are those of the low dose group of Paul (1982). Thus, we take $m = 19$ and $n_i$: 5, 11, 7, 9, 12, 8, 6, 7, 6, 4, 6, 9, 6, 7, 5, 9, 1, 6, 9. For the simulation described below the combination of the parameters $\pi$ and $\phi$ used are $\phi = 0.1, \pi = 0.10, 0.15, 0.20, 0.30, 0.40, 0.50; \pi = 0.10, \phi = 0.10, 0.15, 0.20, 0.25, 0.30, 0.40, 0.50$ and $\pi = 0.40, \phi = 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.40, 0.50$. For each combination of $(\pi, \phi)$ we take 1000 samples from the beta-binomial distribution using the litter sizes given above. Beta-binomial observations were generated by first generating a beta $(\pi, \phi)$ random variate $p$ using the IMSL subroutine RNBET and then a binomial variable with expectation $n_i p$ ($i = 1, \ldots, 19$) using IMSL subroutine RNBIN. For each sample $\pi$ and $\phi$ are estimated by the nine procedures including the maximum likelihood procedure and the small sample efficiency are calculated as follows.

$$\text{Relative Efficiency } (\bar{\pi}^{(t)}) = \frac{MSE(\bar{\pi})}{MSE(\bar{\pi}^{(t)})} = \frac{1}{1000} \sum_{i=1}^{1000} \left( \bar{\pi}_i - \pi \right)^2, \quad \sum_{i=1}^{1000} \left( \bar{\pi}_i^{(t)} - \pi \right)^2.$$

Similarly,

$$\text{Relative Efficiency } (\hat{\phi}^{(t)}) = \frac{MSE(\hat{\phi})}{MSE(\hat{\phi}^{(t)})} = \frac{1}{1000} \sum_{i=1}^{1000} \left( \hat{\phi}_i - \phi \right)^2, \quad \sum_{i=1}^{1000} \left( \hat{\phi}_i^{(t)} - \phi \right)^2.$$

Estimated relative efficiency results for $\bar{\pi}^{(t)}$ are given in table 2.3 and plotted in figure 2.5.3. Those for $\hat{\phi}^{(t)}$ are given in table 2.4 and plotted in figure 2.5.4.

We first discuss the efficiency of $\bar{\pi}^{(t)}$ from figure 2.5.3. We observe the following:
(a) The Gaussian estimator \( \hat{\pi}_{GL} \) is best overall with efficiency greater than one for almost all combinations of the parameters considered. This shows that it is even superior to maximum likelihood estimate \( \hat{\pi}_{ML} \).

(b) The efficiencies of the estimators \( \hat{\pi}_{QM} \), \( \hat{\pi}_{QEE} \), \( \hat{\pi}_{EQE} \), \( \hat{\pi}_{M1} \), \( \hat{\pi}_{M3} \) and \( \hat{\pi}_{M4} \) are similar and they show high efficiency. Note that these estimators have the same estimating equation for \( \pi(\pi_{QM}) \). The efficiencies differ slightly because the estimates differ.

(c) The moment estimator \( \hat{\pi}_{M2} \) is also highly efficient, although less efficient than the estimators mentioned in part (a) and (b).

(d) All the estimators retain high efficiency.

Next, we discuss the efficiency results for \( \hat{\phi}^{(t)} \) from figure 2.5.4. We observe the following.

The Gaussian estimator \( \hat{\phi}_{GL} \) and the moment estimators \( \hat{\phi}_{M2} \) and \( \hat{\phi}_{M3} \) show consistently high efficiency. No such consistent pattern emerges for the remaining estimators. For example, the efficiency of the extended quasi-likelihood estimator \( \hat{\phi}_{EQE} \) is highest when \( \pi \) and \( \phi \) are small, but lowest for large \( \pi \) (see figure 2.5.4(b)).

To study if the same kind of conclusions hold for the efficiency of the estimators when data come from other over-dispersed binomial model we repeated the simulation study by taking samples from the probit normal binomial (PNB) distribution. The procedure of generating data from the PNB distribution is given below.

Data are generated from the probit normal binomial distribution using the method proposed by Ochi and Prentice (1984). First, we generate observations from the correlated probit normal using the following steps, (i) Let \( \omega^T = (\omega_1, \ldots, \omega_n) \) be a normally distributed variate with common mean \( \mu \), variance \( \sigma^2 \) and correlation \( \rho \) with density

\[
\rho_n (\omega; \mu, \sigma, \rho) = (2\pi)^{-n/2} |\Omega_n|^{-1/2} \exp \left\{ -\frac{1}{2} (\omega - \mu 1_n)^T \Omega_n^{-1} (\omega - \mu 1_n) \right\},
\]
where \(1_n\) is an \(n\)-vector of ones, \(\Omega_n = \sigma^2 \{(1 - \rho) I_n + \rho \Omega_n \Omega_n^T\}\). \(I_n\) is an \(n \times n\) identity matrix and \(-(n - 1)^{-1} < \rho < 1\). A binary variate is defined according to whether or not the components of \(\omega\) exceed a common threshold, which can be taken to have value zero without loss of generality. We define

\[
y_i \text{ as } 1 \quad \text{if } \omega_i > 0
\]

\[
as \quad \text{0 if } \omega_i \leq 0.
\]

Let \(y = \sum y_i\). Then the probit normal density is

\[
P(Y = y) = \binom{n}{y} \int_A \rho_n(\omega, \mu, \sigma, \rho) \, d\omega,
\]

where \(A = \{W|\omega_i > 0 \text{ if } i \leq y; \omega_i \leq 0; i > y\}\). Equation (3.4.4.1) can be simplified further if we let \(\gamma = \mu \sigma^{-1}\) and subtract \(\mu\). Then

\[
P(Y = y) = \binom{n}{y} \int_B \rho_n(\omega; 0, 1, \rho) \, d\omega.
\]

where

\(B = \{W|\gamma < \omega_i, i \leq y; \omega_i \leq -\gamma, i > y\}\).

The mean and variance of \(Y\) have the standard form of \(n \pi\) and \(n \pi(1 - \pi)(1 + (n - 1)\phi)\). Now for given \(\pi\) and \(\phi\), we obtain \(\rho\) by solving

\[
\phi = \left\{ \int_{-\infty}^{\gamma} \int_{-\infty}^{\gamma} \rho_2(\omega; 0, 1, \rho) \, d\omega - (\Phi(\gamma))^2 \right\} / \{\Phi(\gamma)\{1 - \Phi(\gamma)\}\},
\]

which can be written as

\[
\left\{ \int_{-\infty}^{\gamma} \int_{-\infty}^{\gamma} \rho_2(\omega; 0, 1, \rho) \, d\omega - (\Phi(\gamma))^2 \right\} - \phi \{\Phi(\gamma)\{1 - \Phi(\gamma)\}\} = 0,
\]

using IMSL subroutine ZBREN, where \(\Phi\) is the standard normal distribution function and \(\gamma = \Phi^{-1}(\pi)\).

(ii) \(n\) correlated probit normal observations are then generated using IMSL subroutine CHFAC and RNMVN with common variance 1 and correlation \(\rho\) obtained in (i).
(iii) We then dichotomize the multivariate observations obtained in (ii), i.e., we count the number of observations greater than zero. These counts are then distributed as probit normal binomial.

The conclusion of the results essentially remained the same, and the results are given in tables 2.5 and 2.6.
Table 2.3(a): Estimated Relative Efficiencies of $\pi_{GL}$, $\pi_{QEE}$, $\pi_{QM}$, $\pi_{EQE}$, $\pi_{M1}$, $\pi_{M2}$, $\pi_{M3}$ and $\pi_{M4}$ for $\phi=0.10$ and $m = 19$ under Beta-binomial distribution.

<table>
<thead>
<tr>
<th>$\pi$</th>
<th>GL</th>
<th>QEE</th>
<th>QM</th>
<th>EQE</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>1.1775</td>
<td>1.0739</td>
<td>1.0627</td>
<td>1.0840</td>
<td>1.0840</td>
<td>1.0515</td>
<td>1.0792</td>
<td>1.0840</td>
</tr>
<tr>
<td>0.15</td>
<td>1.1257</td>
<td>1.0679</td>
<td>1.0295</td>
<td>1.0108</td>
<td>1.0388</td>
<td>1.0262</td>
<td>1.0510</td>
<td>1.0468</td>
</tr>
<tr>
<td>0.20</td>
<td>1.0634</td>
<td>1.0687</td>
<td>1.0156</td>
<td>1.0068</td>
<td>1.0165</td>
<td>1.0055</td>
<td>1.0108</td>
<td>1.0181</td>
</tr>
<tr>
<td>0.30</td>
<td>1.1961</td>
<td>1.1575</td>
<td>1.1966</td>
<td>1.1711</td>
<td>1.1870</td>
<td>1.1818</td>
<td>1.1862</td>
<td>1.1882</td>
</tr>
<tr>
<td>0.40</td>
<td>1.0231</td>
<td>1.0068</td>
<td>1.0153</td>
<td>1.0100</td>
<td>1.0153</td>
<td>0.9875</td>
<td>1.0102</td>
<td>1.0044</td>
</tr>
<tr>
<td>0.50</td>
<td>1.0207</td>
<td>1.0241</td>
<td>1.0148</td>
<td>1.0272</td>
<td>1.0162</td>
<td>0.9996</td>
<td>1.0127</td>
<td>1.0155</td>
</tr>
</tbody>
</table>
Table 2.3(b): Estimated Relative Efficiencies of $\pi_{GL}$, $\pi_{QEE}$, $\pi_{QM}$, $\pi_{EQE}$, $\pi_{M1}$, $\pi_{M2}$, $\pi_{M3}$ and $\pi_{M4}$ for $\pi=0.10$ and $m=19$ under Beta-binomial distribution.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>GL</th>
<th>QEE</th>
<th>QM</th>
<th>EQE</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>1.0857</td>
<td>1.0036</td>
<td>1.0545</td>
<td>1.0253</td>
<td>1.0253</td>
<td>0.9934</td>
<td>1.0017</td>
<td>1.0156</td>
</tr>
<tr>
<td>0.10</td>
<td>1.1775</td>
<td>1.0739</td>
<td>1.0627</td>
<td>1.0840</td>
<td>1.0840</td>
<td>1.0515</td>
<td>1.0792</td>
<td>1.0840</td>
</tr>
<tr>
<td>0.15</td>
<td>1.1629</td>
<td>1.0960</td>
<td>1.0695</td>
<td>1.0740</td>
<td>1.0740</td>
<td>1.0236</td>
<td>1.0605</td>
<td>1.0732</td>
</tr>
<tr>
<td>0.20</td>
<td>1.1634</td>
<td>1.1364</td>
<td>1.0437</td>
<td>1.0629</td>
<td>1.0629</td>
<td>1.0104</td>
<td>1.0719</td>
<td>1.0881</td>
</tr>
<tr>
<td>0.30</td>
<td>1.1695</td>
<td>1.1546</td>
<td>1.0330</td>
<td>1.1144</td>
<td>1.0319</td>
<td>1.0409</td>
<td>1.0581</td>
<td>1.0792</td>
</tr>
<tr>
<td>0.40</td>
<td>1.1027</td>
<td>1.1223</td>
<td>1.0443</td>
<td>1.1153</td>
<td>1.0528</td>
<td>0.9919</td>
<td>1.0789</td>
<td>1.0850</td>
</tr>
<tr>
<td>0.50</td>
<td>1.0577</td>
<td>1.0364</td>
<td>0.9984</td>
<td>1.0798</td>
<td>1.0162</td>
<td>0.9579</td>
<td>1.0737</td>
<td>1.0980</td>
</tr>
</tbody>
</table>

51

Typeset by \texttt{AMS-\LaTeX}
Table 2.3(c): Estimated Relative Efficiencies of $\pi_{GL}$, $\pi_{QEE}$, $\pi_{QM}$, $\pi_{EQE}$, $\pi_{M1}$, $\pi_{M2}$, $\pi_{M3}$ and $\pi_{M4}$ for $\pi=0.40$ and $m = 19$ under Beta-binomial distribution.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>GL</th>
<th>QEE</th>
<th>QM</th>
<th>EQE</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>1.0116</td>
<td>1.0158</td>
<td>1.0048</td>
<td>1.0089</td>
<td>1.0177</td>
<td>1.0153</td>
<td>1.0163</td>
<td>1.0222</td>
</tr>
<tr>
<td>0.10</td>
<td>1.0231</td>
<td>1.0068</td>
<td>1.0153</td>
<td>1.0100</td>
<td>1.0153</td>
<td>0.9875</td>
<td>1.0102</td>
<td>1.0044</td>
</tr>
<tr>
<td>0.15</td>
<td>1.0377</td>
<td>1.0104</td>
<td>1.0093</td>
<td>1.0067</td>
<td>1.0106</td>
<td>1.0002</td>
<td>1.0127</td>
<td>1.0057</td>
</tr>
<tr>
<td>0.20</td>
<td>1.0224</td>
<td>1.0033</td>
<td>1.0093</td>
<td>1.0075</td>
<td>1.0095</td>
<td>0.9806</td>
<td>1.0018</td>
<td>0.9956</td>
</tr>
<tr>
<td>0.25</td>
<td>1.0148</td>
<td>1.0040</td>
<td>1.0045</td>
<td>1.0028</td>
<td>1.0044</td>
<td>0.9653</td>
<td>0.9955</td>
<td>0.9995</td>
</tr>
<tr>
<td>0.30</td>
<td>1.0146</td>
<td>0.9764</td>
<td>0.9965</td>
<td>0.9915</td>
<td>0.9982</td>
<td>0.9758</td>
<td>0.9944</td>
<td>0.9829</td>
</tr>
<tr>
<td>0.40</td>
<td>1.0237</td>
<td>1.0403</td>
<td>1.0008</td>
<td>0.9944</td>
<td>1.0020</td>
<td>0.9824</td>
<td>0.9951</td>
<td>0.9974</td>
</tr>
<tr>
<td>0.50</td>
<td>1.0004</td>
<td>0.9749</td>
<td>1.0036</td>
<td>0.9981</td>
<td>1.0038</td>
<td>0.9379</td>
<td>0.9799</td>
<td>1.0002</td>
</tr>
</tbody>
</table>
Table 2.4(a): Estimated Relative Efficiencies of $\phi_{GL}$, $\phi_{QEE}$, $\phi_{QM}$, $\phi_{EQE}$, $\phi_{M1}$, $\phi_{M2}$, $\phi_{M3}$ and $\phi_{M4}$ for $\phi = 0.10$ and $m = 19$ under Beta-binomial distribution.

<table>
<thead>
<tr>
<th>$\pi$</th>
<th>GL</th>
<th>QEE</th>
<th>QM</th>
<th>EQE</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.9922</td>
<td>0.8654</td>
<td>0.5361</td>
<td>1.6755</td>
<td>0.6601</td>
<td>1.0085</td>
<td>0.9553</td>
<td>0.7617</td>
</tr>
<tr>
<td>0.15</td>
<td>1.0536</td>
<td>0.9609</td>
<td>0.6139</td>
<td>1.1849</td>
<td>0.7888</td>
<td>1.0843</td>
<td>1.0547</td>
<td>0.8546</td>
</tr>
<tr>
<td>0.20</td>
<td>1.0195</td>
<td>1.1326</td>
<td>0.6273</td>
<td>0.8318</td>
<td>0.8005</td>
<td>1.0612</td>
<td>1.0784</td>
<td>1.0223</td>
</tr>
<tr>
<td>0.30</td>
<td>1.0107</td>
<td>1.0452</td>
<td>0.6502</td>
<td>0.6103</td>
<td>0.8260</td>
<td>1.0380</td>
<td>1.0221</td>
<td>0.9627</td>
</tr>
<tr>
<td>0.40</td>
<td>0.9520</td>
<td>0.8692</td>
<td>0.6835</td>
<td>0.5472</td>
<td>0.8449</td>
<td>0.9711</td>
<td>0.9716</td>
<td>0.8640</td>
</tr>
<tr>
<td>0.50</td>
<td>0.9470</td>
<td>0.8612</td>
<td>0.6586</td>
<td>0.4884</td>
<td>0.8200</td>
<td>0.9555</td>
<td>0.9548</td>
<td>0.8840</td>
</tr>
</tbody>
</table>
Table 2.4(b): Estimated Relative Efficiencies of $\phi_{GL}$, $\phi_{QEE}$, $\phi_{QM}$, $\phi_{EQE}$, $\phi_{M1}$, $\phi_{M2}$, $\phi_{M3}$ and $\phi_{M4}$ for $\pi=0.10$ and $m=19$ under Beta-binomial distribution.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>GL</th>
<th>QEE</th>
<th>QM</th>
<th>EQE</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.9770</td>
<td>0.9965</td>
<td>0.3982</td>
<td>1.5537</td>
<td>0.5309</td>
<td>0.9608</td>
<td>1.0017</td>
<td>0.8880</td>
</tr>
<tr>
<td>0.10</td>
<td>0.9922</td>
<td>0.8654</td>
<td>0.5361</td>
<td>1.6755</td>
<td>0.6601</td>
<td>1.0085</td>
<td>0.9553</td>
<td>0.7617</td>
</tr>
<tr>
<td>0.15</td>
<td>1.1168</td>
<td>0.8802</td>
<td>0.6284</td>
<td>1.4566</td>
<td>0.7822</td>
<td>1.1214</td>
<td>1.0829</td>
<td>0.8379</td>
</tr>
<tr>
<td>0.20</td>
<td>1.1427</td>
<td>0.8386</td>
<td>0.6583</td>
<td>1.4013</td>
<td>0.8030</td>
<td>1.0320</td>
<td>1.0321</td>
<td>0.8234</td>
</tr>
<tr>
<td>0.30</td>
<td>1.0610</td>
<td>0.8084</td>
<td>0.6813</td>
<td>1.1561</td>
<td>0.7534</td>
<td>0.9775</td>
<td>1.0038</td>
<td>0.7682</td>
</tr>
<tr>
<td>0.40</td>
<td>0.9482</td>
<td>0.6758</td>
<td>0.6314</td>
<td>0.8718</td>
<td>0.7231</td>
<td>0.8603</td>
<td>0.8771</td>
<td>0.7039</td>
</tr>
<tr>
<td>0.50</td>
<td>0.8535</td>
<td>0.5917</td>
<td>0.6127</td>
<td>0.7207</td>
<td>0.6982</td>
<td>0.7880</td>
<td>0.7607</td>
<td>0.6372</td>
</tr>
</tbody>
</table>
Table 2.4(c): Estimated Relative Efficiencies of $\phi_{GL}$, $\phi_{QEE}$, $\phi_{QM}$, $\phi_{EQE}$, $\phi_{M1}$, $\phi_{M2}$, $\phi_{M3}$ and $\phi_{M4}$ for $\pi=0.40$ and $m=19$ under Beta-binomial distribution.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>GL</th>
<th>QEE</th>
<th>QM</th>
<th>EQE</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.7983</td>
<td>0.8005</td>
<td>0.5186</td>
<td>0.4365</td>
<td>0.6572</td>
<td>0.8370</td>
<td>0.8421</td>
<td>0.7940</td>
</tr>
<tr>
<td>0.10</td>
<td>0.9520</td>
<td>0.8692</td>
<td>0.6835</td>
<td>0.5472</td>
<td>0.8449</td>
<td>0.9711</td>
<td>0.9716</td>
<td>0.8640</td>
</tr>
<tr>
<td>0.15</td>
<td>0.9784</td>
<td>0.8293</td>
<td>0.7199</td>
<td>0.5503</td>
<td>0.8716</td>
<td>0.9983</td>
<td>0.9848</td>
<td>0.8665</td>
</tr>
<tr>
<td>0.20</td>
<td>0.9795</td>
<td>0.8386</td>
<td>0.7346</td>
<td>0.5147</td>
<td>0.8965</td>
<td>1.0156</td>
<td>1.0110</td>
<td>0.8898</td>
</tr>
<tr>
<td>0.25</td>
<td>0.9768</td>
<td>0.8408</td>
<td>0.6889</td>
<td>0.4711</td>
<td>0.8545</td>
<td>0.9905</td>
<td>0.9809</td>
<td>0.8960</td>
</tr>
<tr>
<td>0.30</td>
<td>0.9806</td>
<td>0.8318</td>
<td>0.7236</td>
<td>0.4942</td>
<td>0.8964</td>
<td>0.9925</td>
<td>0.9799</td>
<td>0.9188</td>
</tr>
<tr>
<td>0.40</td>
<td>0.9406</td>
<td>0.8199</td>
<td>0.6877</td>
<td>0.4090</td>
<td>0.8603</td>
<td>0.9435</td>
<td>0.9509</td>
<td>0.8783</td>
</tr>
<tr>
<td>0.50</td>
<td>0.9174</td>
<td>0.7924</td>
<td>0.6434</td>
<td>0.3556</td>
<td>0.8198</td>
<td>0.9117</td>
<td>0.9015</td>
<td>0.8433</td>
</tr>
</tbody>
</table>
Table 2.5(a): Estimated Relative Efficiencies of $\pi_{GL}$, $\pi_{QEE}$, $\pi_{QM}$, $\pi_{EQE}$, $\pi_{M1}$, $\pi_{M2}$, $\pi_{M3}$ and $\pi_{M4}$ for $\phi=0.10$ and $m=19$ using data from PNB distribution.

<table>
<thead>
<tr>
<th>$\pi$</th>
<th>GL</th>
<th>QEE</th>
<th>QM</th>
<th>EQE</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>1.0793</td>
<td>1.0432</td>
<td>0.9989</td>
<td>1.0592</td>
<td>0.9992</td>
<td>0.9861</td>
<td>1.0013</td>
<td>1.0250</td>
</tr>
<tr>
<td>0.15</td>
<td>1.0918</td>
<td>1.0162</td>
<td>1.0035</td>
<td>0.9943</td>
<td>1.0102</td>
<td>1.0014</td>
<td>1.0108</td>
<td>1.0152</td>
</tr>
<tr>
<td>0.20</td>
<td>1.0460</td>
<td>1.0258</td>
<td>1.0134</td>
<td>1.0027</td>
<td>1.0345</td>
<td>1.0246</td>
<td>1.0371</td>
<td>1.0376</td>
</tr>
<tr>
<td>0.30</td>
<td>1.0351</td>
<td>1.0080</td>
<td>1.0072</td>
<td>0.9970</td>
<td>1.0194</td>
<td>1.0155</td>
<td>1.0193</td>
<td>1.0082</td>
</tr>
<tr>
<td>0.40</td>
<td>0.9924</td>
<td>0.9564</td>
<td>0.9853</td>
<td>0.9766</td>
<td>0.9963</td>
<td>0.9948</td>
<td>0.9972</td>
<td>0.9726</td>
</tr>
<tr>
<td>0.50</td>
<td>1.0167</td>
<td>1.0085</td>
<td>1.0001</td>
<td>1.0030</td>
<td>1.0068</td>
<td>1.0002</td>
<td>1.0029</td>
<td>1.0052</td>
</tr>
</tbody>
</table>
Table 2.5(b): Estimated Relative Efficiencies of $\pi_{GL}$, $\pi_{QEE}$, $\pi_{QM}$, $\pi_{EQE}$, $\pi_{M1}$, $\pi_{M2}$, $\pi_{M3}$, and $\pi_{M4}$ for $\pi=0.10$ and $m=19$ using data from PNB distribution.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>GL</th>
<th>QEE</th>
<th>QM</th>
<th>EQE</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>1.0808</td>
<td>0.9958</td>
<td>0.2993</td>
<td>1.0528</td>
<td>1.0063</td>
<td>1.0034</td>
<td>0.9917</td>
<td>0.9950</td>
</tr>
<tr>
<td>0.10</td>
<td>1.0793</td>
<td>1.0432</td>
<td>0.9989</td>
<td>1.0592</td>
<td>0.9992</td>
<td>0.9861</td>
<td>1.0013</td>
<td>1.0250</td>
</tr>
<tr>
<td>0.15</td>
<td>1.1275</td>
<td>1.1486</td>
<td>1.0319</td>
<td>1.0855</td>
<td>1.0418</td>
<td>1.0317</td>
<td>1.0404</td>
<td>1.0568</td>
</tr>
<tr>
<td>0.20</td>
<td>1.1471</td>
<td>1.1184</td>
<td>1.0372</td>
<td>1.1018</td>
<td>1.0473</td>
<td>1.0254</td>
<td>1.0466</td>
<td>1.0640</td>
</tr>
<tr>
<td>0.30</td>
<td>1.1305</td>
<td>1.0804</td>
<td>1.0239</td>
<td>1.0937</td>
<td>1.0436</td>
<td>1.0223</td>
<td>1.0524</td>
<td>1.0648</td>
</tr>
<tr>
<td>0.40</td>
<td>1.1073</td>
<td>1.0833</td>
<td>1.0451</td>
<td>1.0978</td>
<td>1.0530</td>
<td>1.0121</td>
<td>1.0953</td>
<td>1.1313</td>
</tr>
<tr>
<td>0.50</td>
<td>1.0846</td>
<td>1.0796</td>
<td>1.0283</td>
<td>1.0880</td>
<td>1.0386</td>
<td>0.9975</td>
<td>1.0602</td>
<td>1.0831</td>
</tr>
</tbody>
</table>
Table 2.5(c): Estimated Relative Efficiencies of $\pi_{GL}$, $\pi_{QEE}$, $\pi_{QM}$, $\pi_{EQE}$, $\pi_{M1}$, $\pi_{M2}$, $\pi_{M3}$ and $\pi_{M4}$ for $\pi=0.40$ and $m = 19$ using data from PNB distribution.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>$\pi_{GL}$</th>
<th>$\pi_{QEE}$</th>
<th>$\pi_{QM}$</th>
<th>$\pi_{EQE}$</th>
<th>$\pi_{M1}$</th>
<th>$\pi_{M2}$</th>
<th>$\pi_{M3}$</th>
<th>$\pi_{M4}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>1.0081</td>
<td>0.9896</td>
<td>0.9933</td>
<td>0.9870</td>
<td>1.0192</td>
<td>1.0269</td>
<td>1.0182</td>
<td>0.9975</td>
</tr>
<tr>
<td>0.10</td>
<td>0.9924</td>
<td>0.9564</td>
<td>0.9853</td>
<td>0.9766</td>
<td>0.9963</td>
<td>0.9943</td>
<td>0.9972</td>
<td>0.9726</td>
</tr>
<tr>
<td>0.15</td>
<td>1.0065</td>
<td>0.9761</td>
<td>1.0078</td>
<td>0.9749</td>
<td>0.9857</td>
<td>0.9779</td>
<td>0.9840</td>
<td>0.9774</td>
</tr>
<tr>
<td>0.20</td>
<td>1.0083</td>
<td>0.9997</td>
<td>1.0006</td>
<td>0.9823</td>
<td>0.9994</td>
<td>0.9868</td>
<td>0.9928</td>
<td>0.9792</td>
</tr>
<tr>
<td>0.25</td>
<td>1.0140</td>
<td>0.9978</td>
<td>0.9991</td>
<td>0.9866</td>
<td>1.0006</td>
<td>0.9815</td>
<td>0.9961</td>
<td>0.9828</td>
</tr>
<tr>
<td>0.30</td>
<td>1.0095</td>
<td>1.0045</td>
<td>0.9976</td>
<td>0.9893</td>
<td>0.9995</td>
<td>0.9800</td>
<td>0.9926</td>
<td>0.9952</td>
</tr>
<tr>
<td>0.40</td>
<td>1.0019</td>
<td>1.0189</td>
<td>0.9987</td>
<td>0.9925</td>
<td>1.0006</td>
<td>0.9715</td>
<td>0.9897</td>
<td>0.9862</td>
</tr>
<tr>
<td>0.50</td>
<td>1.0010</td>
<td>1.0258</td>
<td>0.9999</td>
<td>0.9934</td>
<td>1.0005</td>
<td>0.9617</td>
<td>0.9909</td>
<td>0.9820</td>
</tr>
</tbody>
</table>
ESTIMATED RELATIVE EFFICIENCY OF $\hat{p}$ VS $p$

FOR $\phi = 0.10$ AND $M = 19$.

Figure 2.5.3(a)
ESTIMATED RELATIVE EFFICIENCY OF $\bar{\pi}$ VS $\phi$
FOR $\pi = 0.10$ AND $M = 19.$

Figure 2.5.3(b)
ESTIMATED RELATIVE EFFICIENCY OF $\tilde{\pi}$ VS $\phi$
FOR $\pi = 0.40$ AND $M = 19.$

Figure 2.5.3(c)
ESTIMATED RELATIVE EFFICIENCY OF $\tilde{\phi}$ VS $\pi$
FOR $\phi = 0.10$ AND $M = 19$.

Figure 2.5.4(a)
ESTIMATED RELATIVE EFFICIENCY OF $\tilde{\phi}$ VS $\phi$
FOR $\pi = 0.10$ AND $M = 19$. 

Figure 2.5.4(b)
ESTIMATED RELATIVE EFFICIENCY OF $\hat{\phi}$ VS $\phi$
FOR $\pi = 0.40$ AND $M = 19$. 

Figure 2.5.4(c)
Table 2.6(a): Estimated Relative Efficiencies of $\phi_{GL}$, $\phi_{QEE}$, $\phi_{QM}$, $\phi_{EQE}$, $\phi_{M1}$, $\phi_{M2}$, $\phi_{M3}$ and $\phi_{M4}$ for $\phi=0.10$ and $m=19$ using data from PNB distribution.

<table>
<thead>
<tr>
<th>$\pi$</th>
<th>GL</th>
<th>QEE</th>
<th>QM</th>
<th>EQE</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>1.0392</td>
<td>0.9964</td>
<td>0.5054</td>
<td>1.6701</td>
<td>0.8093</td>
<td>1.0696</td>
<td>1.1210</td>
<td>0.8987</td>
</tr>
<tr>
<td>0.15</td>
<td>1.0290</td>
<td>1.0843</td>
<td>0.5992</td>
<td>1.1717</td>
<td>0.9228</td>
<td>1.0680</td>
<td>1.1176</td>
<td>0.9795</td>
</tr>
<tr>
<td>0.20</td>
<td>1.0198</td>
<td>1.1293</td>
<td>0.6449</td>
<td>0.8744</td>
<td>0.9825</td>
<td>1.0940</td>
<td>1.1195</td>
<td>1.0288</td>
</tr>
<tr>
<td>0.30</td>
<td>0.9934</td>
<td>1.0169</td>
<td>0.6389</td>
<td>0.5980</td>
<td>0.9590</td>
<td>1.0138</td>
<td>1.0193</td>
<td>0.9735</td>
</tr>
<tr>
<td>0.40</td>
<td>0.9853</td>
<td>0.9472</td>
<td>0.6411</td>
<td>0.5123</td>
<td>0.9861</td>
<td>1.0236</td>
<td>1.0147</td>
<td>0.9373</td>
</tr>
<tr>
<td>0.50</td>
<td>0.9421</td>
<td>0.8805</td>
<td>0.6319</td>
<td>0.5025</td>
<td>0.9371</td>
<td>0.9726</td>
<td>0.9671</td>
<td>0.9020</td>
</tr>
</tbody>
</table>
Table 2.6(b): Estimated Relative Efficiencies of $\phi_{GL}$, $\phi_{QEE}$, $\phi_{QM}$, $\phi_{EQE}$, $\phi_{M1}$, $\phi_{M2}$, $\phi_{M3}$ and $\phi_{M4}$ for $\pi=0.10$ and $m=19$ using data from PNB distribution.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>GL</th>
<th>QEE</th>
<th>QM</th>
<th>EQE</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.8540</td>
<td>0.9700</td>
<td>0.3567</td>
<td>1.4651</td>
<td>0.5798</td>
<td>0.8327</td>
<td>0.9147</td>
<td>0.8216</td>
</tr>
<tr>
<td>0.10</td>
<td>1.0392</td>
<td>0.9964</td>
<td>0.5054</td>
<td>1.6701</td>
<td>0.8093</td>
<td>1.0696</td>
<td>1.1210</td>
<td>0.8987</td>
</tr>
<tr>
<td>0.15</td>
<td>0.9929</td>
<td>0.9245</td>
<td>0.5540</td>
<td>1.5341</td>
<td>0.8224</td>
<td>1.0479</td>
<td>1.1495</td>
<td>0.8840</td>
</tr>
<tr>
<td>0.20</td>
<td>1.0280</td>
<td>0.8265</td>
<td>0.6137</td>
<td>1.3358</td>
<td>0.8834</td>
<td>1.0531</td>
<td>1.1031</td>
<td>0.8446</td>
</tr>
<tr>
<td>0.30</td>
<td>0.9805</td>
<td>0.7702</td>
<td>0.6473</td>
<td>1.0706</td>
<td>0.8738</td>
<td>0.9739</td>
<td>0.9872</td>
<td>0.7761</td>
</tr>
<tr>
<td>0.40</td>
<td>0.9111</td>
<td>0.6666</td>
<td>0.6228</td>
<td>0.8849</td>
<td>0.8338</td>
<td>0.9029</td>
<td>0.8598</td>
<td>0.7239</td>
</tr>
<tr>
<td>0.50</td>
<td>0.8594</td>
<td>0.6163</td>
<td>0.5993</td>
<td>0.7399</td>
<td>0.7935</td>
<td>0.8485</td>
<td>0.8100</td>
<td>0.6527</td>
</tr>
</tbody>
</table>
Table 2.6(c): Estimated Relative Efficiencies of $\phi_{GL}$, $\phi_{QEE}$, $\phi_{QM}$, $\phi_{EQE}$, $\phi_{M1}$, $\phi_{M2}$, $\phi_{M3}$ and $\phi_{M4}$ for $\pi=0.40$ and $m = 19$ using data from PNB distribution.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>GL</th>
<th>QEE</th>
<th>QM</th>
<th>EQE</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.8810</td>
<td>0.9210</td>
<td>0.5206</td>
<td>0.4182</td>
<td>0.7512</td>
<td>0.9048</td>
<td>0.3982</td>
<td>0.8761</td>
</tr>
<tr>
<td>0.10</td>
<td>0.9883</td>
<td>0.9472</td>
<td>0.6411</td>
<td>0.5123</td>
<td>0.9861</td>
<td>1.0236</td>
<td>1.0147</td>
<td>0.9373</td>
</tr>
<tr>
<td>0.15</td>
<td>1.0151</td>
<td>0.9205</td>
<td>0.6825</td>
<td>0.5138</td>
<td>0.9900</td>
<td>1.0187</td>
<td>1.0039</td>
<td>0.9405</td>
</tr>
<tr>
<td>0.20</td>
<td>0.9925</td>
<td>0.9126</td>
<td>0.6950</td>
<td>0.4925</td>
<td>0.9819</td>
<td>1.0052</td>
<td>1.0012</td>
<td>0.9086</td>
</tr>
<tr>
<td>0.25</td>
<td>0.9823</td>
<td>0.8533</td>
<td>0.7040</td>
<td>0.4692</td>
<td>0.9861</td>
<td>0.9895</td>
<td>0.9681</td>
<td>0.9050</td>
</tr>
<tr>
<td>0.30</td>
<td>0.9849</td>
<td>0.8347</td>
<td>0.6945</td>
<td>0.4445</td>
<td>0.9855</td>
<td>0.9883</td>
<td>0.9762</td>
<td>0.8991</td>
</tr>
<tr>
<td>0.40</td>
<td>0.9292</td>
<td>0.8128</td>
<td>0.6522</td>
<td>0.3912</td>
<td>0.9287</td>
<td>0.9294</td>
<td>0.9291</td>
<td>0.8698</td>
</tr>
<tr>
<td>0.50</td>
<td>0.9078</td>
<td>0.7663</td>
<td>0.6256</td>
<td>0.3432</td>
<td>0.9098</td>
<td>0.8988</td>
<td>0.8941</td>
<td>0.8432</td>
</tr>
</tbody>
</table>
2.6 Examples

In this section two data sets are analysed using the various estimation procedures presented in section 2.3. The first data set for the number of complementary crossover offspring and ++ in each of 36 Drosophila melanogaster families, given in table 2.6.1 is from Potthoff and Whittinghill (1966). The second data set from shell toxicology laboratory given in table 2.6.2 is that of the number of live foetuses in a litter affected by treatment and the number of live foetuses of low dose group in Paul (1982). Estimates of the parameters $\pi$ and $\phi$ by the eight procedures and the maximum likelihood estimates under beta-binomial distribution are presented. The estimated asymptotic variance-covariance for $\hat{\lambda}$ is obtained by replacing the unknown parameters by their estimates. Estimated efficiencies of $\pi$ and $\phi$ relative to the MLE are also presented. Note that the moment method of estimation denoted by $\lambda_{M_4}$ is similar to the one proposed by Liang (1992).

The results presented in table 2.6.3 show that for the estimation of the mean (regression) parameter all the procedures were highly efficient, except for the extended quasi-likelihood estimate with an efficiency of 0.87. For the joint estimation of $\pi$ and $\phi$, the quadratic ($\pi_{QEE}, \phi_{QEE}$), the Gaussian ($\pi_{GL}, \phi_{GL}$) and method of moments ($\pi_{M2}, \phi_{M2}$) and ($\pi_{M3}, \phi_{M3}$) estimating equations produce estimates which were highly efficient. These results are in agreement with the simulation results presented in section 2.5.2.

The results presented in table 2.6.4 show that for the estimation of the mean parameter, all the procedures were highly efficient with efficiency greater than 0.94 except for the moment method M4 with efficiency 0.89. For the joint estimation of parameters $\pi$ and $\phi$ we see that the moment estimators proposed by Srivastava and Wu (1993) ($\pi_{M2}, \phi_{M2}$) and ($\pi_{M4}, \phi_{M4}$) and the Gaussian estimator ($\pi_{GL}, \phi_{GL}$) were highly efficient. The second best estimates were the Kleinman estimator ($\pi_{M1}, \phi_{M1}$)
and the moment estimator \((π_M, φ_M)\). The maximum extended quasi-likelihood estimate for \(φ\) performs poorly. These results are in agreement with the simulation results given in section 2.5.2.
Table 2.6.1: Numbers of the cross-over offspring in \( m=36 \) families from Potthoff and Whittinghill (1966). \( y= \) number of ++ offspring. \( n= \) total cross-over offspring.

\[
\begin{array}{cccccccccccccccccccc}
y: & 7 & 1 & 4 & 3 & 5 & 3 & 0 & 1 & 1 & 3 & 0 & 1 & 0 & 3 & 0 & 4 & 2 & 2 & 3 & 5 & 2 & 1 & 2 & 3 & 1 & 1 & 4 & 5 & 3 & 3 & 5 & 1 & 1 & 3 & 4 & 0 & 1 & 2 \\
\end{array}
\]

Table 2.6.2: The Toxicological data of Low dose group from Paul (1982). \( m=19 \) litters. \( y= \) number of live foetuses affected by treatment. \( n= \) total number of live foetuses.

\[
\begin{array}{cccccccccccccccccccc}
y: & 0 & 1 & 1 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 & 3 & 0 & 0 & 1 & 5 & 0 & 0 & 3 \\
n: & 5 & 1 & 1 & 7 & 9 & 1 & 2 & 8 & 6 & 7 & 6 & 4 & 6 & 9 & 6 & 7 & 5 & 9 & 1 & 5 & 9 \\
\end{array}
\]
Table 2.6.3: The Estimates $\hat{\pi}$ and $\hat{\phi}$ of the parameters $\pi$ and $\phi$ and the Estimated Relative Efficiency of $\hat{\pi}$ and $\hat{\phi}$ for $t=ML, GL, QEE, QM, EQE, M1, M2, M3$ and $M4$ for Cross-over offspring data in Table 1.

<table>
<thead>
<tr>
<th>Method</th>
<th>Estimates of $\pi$</th>
<th>$\phi$</th>
<th>Estimated Relative Efficiency for $\hat{\pi}$</th>
<th>$\hat{\phi}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML</td>
<td>0.4726</td>
<td>0.0954</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>Gaussian</td>
<td>0.4741</td>
<td>0.0954</td>
<td>0.9991</td>
<td>0.9969</td>
</tr>
<tr>
<td>Quadratic</td>
<td>0.4741</td>
<td>0.0954</td>
<td>1.0025</td>
<td>0.9975</td>
</tr>
<tr>
<td>Quasi/Moments</td>
<td>0.4743</td>
<td>0.1195</td>
<td>0.9178</td>
<td>0.9045</td>
</tr>
<tr>
<td>Ext.Quasi</td>
<td>0.4743</td>
<td>0.1352</td>
<td>0.8722</td>
<td>0.7994</td>
</tr>
<tr>
<td>Moments 1</td>
<td>0.4742</td>
<td>0.1038</td>
<td>0.9688</td>
<td>0.8678</td>
</tr>
<tr>
<td>Moments 2</td>
<td>0.4737</td>
<td>0.0970</td>
<td>0.9379</td>
<td>0.9779</td>
</tr>
<tr>
<td>Moments 3</td>
<td>0.4741</td>
<td>0.0969</td>
<td>0.9933</td>
<td>0.9783</td>
</tr>
<tr>
<td>Moments 4</td>
<td>0.4741</td>
<td>0.0897</td>
<td>1.0207</td>
<td>0.8502</td>
</tr>
</tbody>
</table>
Table 2.6.4: The Estimates $\hat{\pi}_t$ and $\hat{\phi}_t$ of the parameters $\pi$ and $\phi$ and the Estimated Relative Efficiency of $\hat{\pi}_t$ and $\hat{\phi}_t$ for $t =$ ML, GL, QEE, QM, EQE, M1, M2, M3 and M4 for the Toxicological data in Table 2.

<table>
<thead>
<tr>
<th>Method</th>
<th>Estimates of $\pi$</th>
<th>Estimates of $\phi$</th>
<th>Estimated Relative Efficiency for $\pi$</th>
<th>Estimated Relative Efficiency for $\phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML</td>
<td>0.1273</td>
<td>0.1058</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>Gaussian</td>
<td>0.1331</td>
<td>0.1006</td>
<td>0.9731</td>
<td>0.9908</td>
</tr>
<tr>
<td>Quadratic</td>
<td>0.1265</td>
<td>0.1225</td>
<td>0.9460</td>
<td>0.7915</td>
</tr>
<tr>
<td>Quasi/Moments</td>
<td>0.1267</td>
<td>0.1162</td>
<td>0.9653</td>
<td>0.7734</td>
</tr>
<tr>
<td>Ext. Quasi</td>
<td>0.1267</td>
<td>0.1177</td>
<td>0.9607</td>
<td>0.4280</td>
</tr>
<tr>
<td>Moments 1</td>
<td>0.1280</td>
<td>0.0917</td>
<td>1.0508</td>
<td>0.9481</td>
</tr>
<tr>
<td>Moments 2</td>
<td>0.1353</td>
<td>0.0914</td>
<td>0.9874</td>
<td>1.0519</td>
</tr>
<tr>
<td>Moments 3</td>
<td>0.1272</td>
<td>0.1059</td>
<td>0.9994</td>
<td>0.9159</td>
</tr>
<tr>
<td>Moments 4</td>
<td>0.1257</td>
<td>0.1415</td>
<td>0.8916</td>
<td>1.0629</td>
</tr>
</tbody>
</table>
2.7 Discussion

By examining the efficiency results presented in section 2.5 and the results from the two examples given in section 2.6 we conclude that for the estimation of the mean (regression) parameters either of the Gaussian likelihood or quasi-likelihood procedure is a good choice. The Gaussian likelihood estimates, in general, have higher efficiencies, whereas the quasi-likelihood estimates are easier to calculate. However, for the estimation of \( \phi \) or the joint estimation of the mean parameter and the dispersion parameter \( \phi \) we recommend the Gaussian likelihood procedure and the moment procedure M2 proposed by Srivastava and Wu (1993). The estimates by these procedures, in general, have high efficiencies and show consistent behaviour throughout the parameter space examined. The moment estimating equation M2 for \( \phi \) is very simple, it does not depend on the skewness or the kurtosis and computationally it is less intensive than the maximum likelihood estimate. The Gaussian estimates performed very well, although the binary data is highly non-Gaussian in distribution. These estimates do not depend on the skewness or the kurtosis. Whittle (1961), showed that the Gaussian estimates are asymptotically efficient in the class of estimates based on the second order function of the observations even if the observations are not normally distributed. For the estimation of the mean parameter, the maximum likelihood estimate using the beta-binomial distribution is sensitive to misspecification of the variance structure, but the Gaussian, the moment M2 and the quasi-likelihood estimates are robust. For the estimation of the dispersion parameter, the Gaussian and the moment estimate M2 are robust to misspecification of the variance structure.
CHAPTER III

TESTING FOR HOMOGENEITY OF PROPORTIONS IN THE PRESENCE OF A COMMON DISPERSION PARAMETER

3.1 Introduction

An important problem that arises in toxicology and other similar fields is to compare proportions of a certain characteristic in several groups. Data that arise in practice can be described as follows. Suppose that there are $T$ treatment groups and that the $i$-th group has $m_i$ litters, $i = 1, \ldots, T$. The proportion responding in the $j$-th litter of $i$-th group is $y_{ij}/n_{ij}$, $j = 1, \ldots, m_i$; $i = 1, \ldots, T$. A typical data set with $T$ groups can be presented as

<table>
<thead>
<tr>
<th>Groups</th>
<th>Proportions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$y_{11}/n_{11}$ $y_{12}/n_{12}$ $\cdots$ $y_{1j}/n_{1j}$ $\cdots$ $y_{1m_1}/n_{1m_1}$</td>
</tr>
<tr>
<td>2</td>
<td>$y_{21}/n_{21}$ $y_{22}/n_{22}$ $\cdots$ $y_{2j}/n_{2j}$ $\cdots$ $y_{2m_2}/n_{2m_2}$</td>
</tr>
<tr>
<td></td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$i$</td>
<td>$y_{i1}/n_{i1}$ $y_{i2}/n_{i2}$ $\cdots$ $y_{ij}/n_{ij}$ $\cdots$ $y_{im_i}/n_{im_i}$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$T$</td>
<td>$y_{T1}/n_{T1}$ $y_{T2}/n_{T2}$ $\cdots$ $y_{Tj}/n_{Tj}$ $\cdots$ $y_{Tm_T}/n_{Tm_T}$</td>
</tr>
</tbody>
</table>

|
where

\[ n_{ij} = j\text{-th litter size in the } i\text{-th group} \]

\[ y_{ij} = \text{number responding within the } j\text{-th litter of the } i\text{-th group}. \]

A number of parametric and non-parametric procedures are available for testing homogeneity of proportions in presence of over-dispersion. Of these, the likelihood ratio test based on the beta-binomial model has found prominence in the literature.

The purpose of this Chapter is to develop \( C(\alpha) \) or score tests for testing the homogeneity of proportions in the presence of common over/under dispersion. The \( C(\alpha) \) test is based on the residual of a regression of the score function for the parameter(s) of interest on the score function for the nuisance parameters. The \( C(\alpha) \) or score test has been shown by many authors to be asymptotically equivalent to the likelihood ratio test and to tests using the maximum likelihood estimates (i.e. Wald tests). See, for example, Moran (1970) and Cox and Hinkley (1974). Potential drawbacks to the likelihood ratio and Wald tests include the fact that the likelihood ratio test require estimates of the parameter under both the null and alternative hypothesis and the Wald test require estimates under the alternative hypothesis. Several advantages of the \( C(\alpha) \) or score class of tests are (i) it often maintains, at least approximately, a preassigned level of significance, say \( \alpha \) (Bartoo and Puri, 1967), (ii) it is locally asymptotically most powerful (Bühler and Puri, 1966 and Moran, 1970), (iii) it requires estimates of the parameters only under the null hypothesis, and (iv) it often produces a statistic, which is simple to calculate.

The score or \( C(\alpha) \) statistic has been found to be useful for detecting over-dispersion in binomial and Poisson data (Paul, Liang and Self, 1989; Dean and Lawless, 1990). As homogeneity tests, the \( C(\alpha) \) class of tests have been widely used (see Neyman and Scott, 1966; Moran, 1973; Tarone, 1985; Barnwal and Paul, 1988; Paul, 1989). Breslow (1990(a)) developed two versions of the score test for testing
the significance of added variables in over-dispersed Poisson regression and quasi-likelihood models. One version is calculated from the usual model based covariance matrix and another using the empirical covariance matrix that has asymptotic justification. Breslow’s main objective was to develop score tests which are applicable to more general semi-parametric models that are robust to misspecification of the mean/variance relation. He noted that the empirical score test performed reasonably well when the sample size is sufficiently large. In the case of small samples, performance was not as good as the model-based score tests. Boos (1992) derived a score test known as a genaralized score test using empirical variance estimates.

In section (3.2) we consider parametric procedures, based on the extended beta-binomial model, for testing homogeneity of proportions in the presence of over/under dispersion. In particular, we derive a likelihood ratio statistic (LR) and two $C(\alpha)$ (Neyman, 1959) statistics. Of these, one is based on the maximum likelihood estimators (MLEs) and the other is based on the method of moments estimators (MMEs) of the nuisance parameters. The MLEs and the MMEs are $\sqrt{m}$ consistent. We then conduct a simulation study, for comparing the size and power of the $C(\alpha)$ statistics with those of the LR statistics.

The $C(\alpha)$ statistics derived in section (3.2) based on the extended beta-binomial model hold nominal level well, but do not produce simple forms. Also, they may not be robust against data departure. So in section (3.3) we consider semi-parametric procedures. The semi-parametric procedure requires assumptions only of the first two moments of the binomial response with some unknown common dispersion. Under this assumption we derive $C(\alpha)$ or score test based on the quasi-likelihood and the extended quasi-likelihood.

In a recent paper, Rao and Scott (1992) propose a procedure for testing homogeneity of proportions in several clusters based on the concept of design effect
and effective sample size. In this section, we also derive two versions of the Rao-Scott statistic. In section (3.4), we also derive a score test based on the empirical variance. A further simulation study was conducted to compare the behaviours in terms of size, power and robustness, of the $C(\alpha)$ statistics based on the quasi-likelihood and the extended quasi-likelihood, of the two $C(\alpha)$ statistics based on the extended beta-binomial distribution, and of the $C(\alpha)$ statistics based on the empirical variance and two versions of the Rao-Scott statistic. In studying robustness of the procedures we generate data from the beta-binomial (BB), the probit normal binomial (PNB), and the logit normal binomial (LNB) models. Robustness in terms of dispersion heterogeneity is also considered. Some real life examples are given.

3.2 Parametric Methods

3.2.1. Likelihood Ratio test (LR)

We consider that $Y_{ij} | n_{ij} \sim \text{extended BB}(\pi_i, \theta_i)$ for $j = 1, \ldots, m_i$ and $i = 1, \ldots, T$. Our interest is to test

$$H_0 : \pi_1 = \ldots = \pi_T$$

against

$$H_A : \text{not all } \pi_i \text{'s are the same},$$

assuming that $\theta_1 = \ldots = \theta_T = \theta$.

The extended Beta-binomial log-likelihood, apart from some constant, is

$$l = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \left\{ \sum_{r=0}^{y_{ij}-1} \log (\pi_i + r\theta_i) + \sum_{r=0}^{n_{ij}-y_{ij}-1} \log (1 - \pi_i + r\theta_i) - \sum_{r=0}^{n_{ij}-1} \log (1 + r\theta_i) \right\}$$

(3.2.1.1.)

The likelihood ratio test is defined as

$$LR = 2(l_1 - l_0)$$
where \( l_0 \) is the maximum log-likelihood function under the null hypothesis and \( l_1 \) is the maximum log-likelihood function under the alternative hypothesis. The \( LR \) statistic under the null hypothesis is distributed asymptotically chi-square with \((T - 1)\) degrees of freedom.

3.2.2. \( C(\alpha) \) statistics for testing equality of means in presence of a common dispersion parameter

For the derivation of the \( C(\alpha) \) statistics for testing the homogeneity of the proportions under common dispersion parameter it is convenient to reparametrize \( \pi_i \) under \( H_A \), by \( \pi_i = \pi + \tau_i \) with \( \tau_T = 0 \). Then testing \( H_0 \) is equivalent to testing \( \tau_i = 0, \ i = 1, \ldots, T - 1 \) with \( \pi \) and \( \theta \) treated as nuisance parameters. Tarone (1985) applied this technique to derive \( C(\alpha) \) statistics to test the equality of several odds ratios. Barnwal and Paul (1988) used this technique to obtain \( C(\alpha) \) statistics to test the homogeneity of Poisson means in the presence of negative binomial overdispersion. Using the reparametrization of \( \pi_i \), the log-likelihood, under \( H_A \) can be written as

\[
l = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \left\{ \sum_{r=0}^{\pi_{ij} - 1} \log(\pi + \tau_i + r\theta) + \sum_{r=0}^{\pi_{ij} - 1} \log(1 - \pi - \tau_i + r\theta) \right. \\
- \sum_{r=0}^{n_{ij} - 1} \log(1 + r\theta) \right\}.
\]

Define \( \tau = (\tau_1, \ldots, \tau_{T-1}) \) and \( \lambda = (\lambda_1, \lambda_2) = (\pi, \theta) \).

Then let

\[
\psi_i = \left. \frac{\partial l}{\partial \tau_i} \right|_{\tau = 0}, \quad i = 1, \ldots, T - 1
\]

and

\[
\gamma_k = \left. \frac{\partial l}{\partial \lambda_k} \right|_{\tau = 0}, \quad k = 1, 2.
\]

Now, let \( \hat{\lambda} \) be some \( \sqrt{n} \) consistent estimator of \( \lambda \) under the null hypothesis. Then the \( C(\alpha) \) test is based on

\[
S_i(\hat{\lambda}) = \psi_i(\hat{\lambda}) - \beta_1 \gamma_1(\hat{\lambda}) - \beta_2 \gamma_2(\hat{\lambda}), \quad i = 1, \ldots, T - 1
\]
where $\beta_{1i}$ and $\beta_{2i}$ are the partial regression coefficients of $\psi_i$ on $\gamma_1$ and $\psi_i$ on $\gamma_2$ respectively. The variance-covariance matrix of $S(\lambda) = \{S_1(\lambda), \ldots, S_{T-1}(\lambda)\}'$ is $D - AB^{-1}A'$ and the regression coefficients $\beta = (\beta_1, \beta_2) = AB^{-1}$, where $\beta_1 = (\beta_{11}, \ldots, \beta_{1(T-1)})$, $\beta_2 = (\beta_{21}, \ldots, \beta_{2(T-1)})$.

$$D_{it} = E \left[ \frac{-\partial^2 l}{\partial \tau_i \partial \tau_t} \right]_{\tau = 0}, \quad i, t = 1, \ldots, T - 1.$$ 

$$A_{ik} = E \left[ \frac{-\partial^2 l}{\partial \tau_i \partial \lambda_k} \right]_{\tau = 0}, \quad i = 1, \ldots, T - 1 \quad k = 1, 2$$

and

$$B_{ks} = E \left[ \frac{-\partial^2 l}{\partial \lambda_k \partial \lambda_s} \right]_{\tau = 0}, \quad k, s = 1, 2.$$

Using $\hat{\lambda}$ in S.A.B and D, the $C(\alpha)$ statistic is given by $S'(D - AB^{-1}A')^{-1}S$, which is approximately distributed as chi-square with $T - 1$ degrees of freedom (see Neyman, 1959; Neyman and Scott, 1966; Moran, 1970). If we use the maximum likelihood estimate of $\lambda$, then $S_i(\hat{\lambda}) = \psi_i(\hat{\lambda})$, reducing the $C(\alpha)$ statistic to a score statistic (Rao, 1948).

Using the log-likelihood (3.2.1.1) we obtain for $i = 1, \ldots, T - 1$, $t = 1, \ldots, T - 1$,

$$\psi_i = \left. \frac{\partial l}{\partial \tau_i} \right|_{\tau = 0} = \sum_{j=1}^{m_i} \left\{ \sum_{r=0}^{\gamma_{ij} - 1} \frac{1}{(\pi + r\theta)} - \sum_{r=0}^{n_{ij} - y_{ij} - 1} \frac{1}{(1 - \pi + r\theta)} \right\},$$

$$\gamma_1 = \left. \frac{\partial l}{\partial \lambda_1} \right|_{\tau = 0} = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \left\{ \sum_{r=0}^{\gamma_{ij} - 1} \frac{1}{(\pi + r\theta)} - \sum_{r=0}^{n_{ij} - y_{ij} - 1} \frac{1}{(1 - \pi + r\theta)} \right\},$$

$$\gamma_2 = \left. \frac{\partial l}{\partial \lambda_2} \right|_{\tau = 0} = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \left\{ \sum_{r=0}^{\gamma_{ij} - 1} \frac{r}{(\pi + r\theta)^2} + \sum_{r=0}^{n_{ij} - y_{ij} - 1} \frac{r}{(1 - \pi + r\theta)^2} \right\},$$

$$\left( \frac{-\partial^2 l}{\partial \tau_i^2} \right)_{\tau = 0} = \sum_{j=1}^{m_i} \left\{ \sum_{r=0}^{\gamma_{ij} - 1} \frac{1}{(\pi + r\theta)^2} + \sum_{r=0}^{n_{ij} - y_{ij} - 1} \frac{1}{(1 - \pi + r\theta)^2} \right\},$$

$$\left( \frac{-\partial^2 l}{\partial \tau_i \partial \tau_t} \right)_{\tau = 0} = 0, \quad \text{for } i \neq t,$$

$$\left( \frac{-\partial^2 l}{\partial \tau_i \partial \lambda_1} \right)_{\tau = 0} = \left( \frac{-\partial^2 l}{\partial \tau_i^2} \right)_{\tau = 0}.$$
\[
\left( \frac{-\partial^2 l}{\partial \tau \partial \lambda_2} \right)_{r=0} = \sum_{j=1}^{n_2} \left\{ \sum_{r=0}^{n_2-1} \frac{r}{(\pi + r\theta)^2} - \sum_{r=0}^{n_2-1} \frac{r}{(1 - \pi + r\theta)^2} \right\}.
\]

\[
\left( \frac{-\partial^2 l}{\partial \lambda_1^2} \right)_{r=0} = \sum_{i=1}^{T} \left\{ \sum_{j=1}^{m_i} \left\{ \sum_{r=0}^{r_{ij}-1} \frac{1}{(\pi + r\theta)^2} + \sum_{r=0}^{r_{ij}-1} \frac{1}{(1 - \pi + r\theta)^2} \right\} \right\}.
\]

\[
\left( \frac{-\partial^2 l}{\partial \lambda_1 \partial \lambda_2} \right)_{r=0} = \sum_{i=1}^{T} \left\{ \sum_{j=1}^{m_i} \left\{ \sum_{r=0}^{r_{ij}-1} \frac{r}{(\pi + r\theta)^2} - \sum_{r=0}^{r_{ij}-1} \frac{r}{(1 - \pi + r\theta)^2} \right\} \right\}.
\]

and

\[
\left( \frac{-\partial^2 l}{\partial \lambda_2^2} \right)_{r=0} = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \left\{ \sum_{r=0}^{r_{ij}-1} \frac{r^2}{(\pi + r\theta)^2} + \sum_{r=0}^{r_{ij}-1} \frac{r^2}{(1 - \pi + r\theta)^2} - \sum_{r=0}^{r_{ij}-1} \frac{r^2}{(1 + r\theta)^2} \right\}.
\]

Expectations of the negative of the second derivatives are given below, for \( i = 1, \ldots, T - 1; \ t = 1, \ldots, T - 1. \)

\[
D_{ii} = E \left( \left. \frac{-\partial^2 l}{\partial \tau_i^2} \right|_{r=0} \right) = \sum_{j=1}^{m_i} \left\{ \sum_{r=0}^{r_{ij}-1} \frac{P(Y_{ij} > r)}{(\pi + r\theta)^2} + \sum_{r=0}^{r_{ij}-1} \frac{P(Y_{ij} < n_{ij} - r)}{(1 - \pi + r\theta)^2} \right\},
\]

\[
D_{it} = E \left( \left. \frac{-\partial^2 l}{\partial \tau_i \partial \tau_t} \right|_{r=0} \right) = 0, \quad i \neq t,
\]

\[
A_{ii} = E \left( \left. \frac{-\partial^2 l}{\partial \tau_i \partial \lambda_1} \right|_{r=0} \right) = D_{ii},
\]

\[
A_{i2} = E \left( \left. \frac{-\partial^2 l}{\partial \tau_i \partial \lambda_2} \right|_{r=0} \right) = \sum_{j=1}^{m_i} \frac{1}{\theta} \left\{ \pi \sum_{r=0}^{r_{ij}-1} \frac{P(Y_{ij} > r)}{(\pi + r\theta)^2} - (1 - \pi) \sum_{r=0}^{r_{ij}-1} \frac{P(Y_{ij} < n_{ij} - r)}{(1 - \pi + r\theta)^2} \right\},
\]

\[
B_{11} = E \left( \left. \frac{-\partial^2 l}{\partial \lambda_1^2} \right|_{r=0} \right) = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \left\{ \sum_{r=0}^{r_{ij}-1} \frac{P(Y_{ij} > r)}{(\pi + r\theta)^2} + \sum_{r=0}^{r_{ij}-1} \frac{P(Y_{ij} < n_{ij} - r)}{(1 - \pi + r\theta)^2} \right\},
\]

\[
B_{12} = E \left( \left. \frac{-\partial^2 l}{\partial \lambda_1 \partial \lambda_2} \right|_{r=0} \right) = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{1}{\theta} \left[ -\pi \sum_{r=0}^{r_{ij}-1} \frac{P(Y_{ij} > r)}{(\pi + r\theta)^2} + (1 - \pi) \sum_{r=0}^{r_{ij}-1} \frac{P(Y_{ij} < n_{ij} - r)}{(1 - \pi + r\theta)^2} \right],
\]

and

\[
B_{22} = E \left( \left. \frac{-\partial^2 l}{\partial \lambda_2^2} \right|_{r=0} \right) = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{1}{\theta^2} \left\{ \sum_{r=0}^{r_{ij}-1} \frac{\pi^2 P(Y_{ij} > r)}{(\pi + r\theta)^2} + \sum_{r=0}^{r_{ij}-1} \frac{\pi^2 P(Y_{ij} < n_{ij} - r)}{(1 - \pi + r\theta)^2} - \sum_{r=0}^{r_{ij}-1} \frac{1}{(1 + r\theta)^2} \right\}.
\]
Simplifications of the expectations of the above expressions are given in Appendix A. Note that under the null hypothesis the parameters \( \pi \) and \( \theta \) are common across groups. Therefore, the estimation of \( \pi \) and \( \theta \) from the \( T \) groups can be considered to be estimation from a single group consisting of the combined \( \lambda \) in the \( T \) groups. The maximum likelihood estimates (MLEs) and the method of moments estimates (MMEs) of \( \pi \) and \( \theta \) have been shown to be \( \sqrt{m} \) consistent (Moore, 1986). Denote the MLEs of \( \lambda = (\pi, \theta) \) by \( \lambda_m \) and the MMEs by \( \lambda_{mm} \). If we use \( \lambda_{mm} \) in S, D, A and B, the \( C(\alpha) \) statistic based on the method of moment estimators is

\[
C_{mm} = S'(D - AB^{-1}A')^{-1} S.
\]

If \( \lambda_m \) is used in S,D,A and B, then \( S(\lambda) = \psi(\lambda_m) = (\psi_1(\lambda_m), \ldots, \psi_{T-1}(\lambda_m)) \) and the \( C(\alpha) \) or the score statistic is

\[
C_m = \psi'(D - AB^{-1}A')^{-1} \psi.
\]

Estimation of the parameters \( \pi \) and \( \phi \) using the maximum likelihood method and Kleinman's method of moments is described in detail in Chapter 2.

3.2.3 Simulations

A simulation study was conducted to compare the performance, in terms of size and power, of the likelihood ratio statistic LR and the \( C(\alpha) \) statistics \( C_m \) and \( C_{mm} \). In this study we consider \( T = 2 \) groups. The litter sizes and number of litters were chosen as those of the control group \( (m_1 = 27) \) and low dose treatment group \( (m_2 = 19) \) of Paul (1982). Thus, for group 1, the litter sizes are 12, 7, 6, 6, 7, 8, 10, 7, 8, 6, 6, 11, 7, 8, 9, 2, 7, 9, 7, 11, 10, 4, 8, 10, 12, 8, 7, 8 and those for group 2 are: 5, 11, 7, 9, 12, 8, 6, 7, 6, 4, 6, 9, 6, 7, 5, 9, 1, 6, 9. Beta-binomial observations were generated by first generating a beta \((\pi, \theta)\) random variate \( p \) using the IMSL subroutine RNBET and then a binomial random variable with expectation
\[ n_{ijp}(i = 1, 2; j = 1, \ldots, m_i) \] using IMSL subroutine RNBIN. Empirical levels were calculated based on 1000 replications for each combination of \( \pi = 0.06, 0.08, 0.10, 0.12, 0.20, 0.30 \) and \( \theta = 0.02, 0.06, 0.10, 0.20 \). The results are given in Table 3.1. For small \( \theta (\theta = 0.02) \) the \( C(\alpha) \) statistic \( C_{mm} \) shows some conservative behavior, otherwise all the statistics produce empirical levels close to the nominal. For power comparison, we considered \( \theta = 0.02, 0.06, 0.09, 0.12, 0.16 \). For each value of \( \theta \) empirical powers were calculated for \( (\pi_1, \pi_2) : (0.08, 0.10), (0.08, .15), (0.08, .20), (0.08, 0.30) \). The results are given in Table 3.2. Again, except for very small \( \theta \) the power of the statistics do not show marked disagreement. For small \( \theta (\theta = 0.02) \), the power of the statistic \( C_{mm} \) is to some extent smaller than those of the other two statistics LR and \( C_m \). This is because the level of the \( C_{mm} \) for the value of \( \theta \) is smaller than that of the other two. Either of the \( C(\alpha) \) statistics is preferable, as each requires estimates of the parameters only under the null hypothesis. The moment estimates are easier to obtain than the mle's, although for obtaining mle's under \( H_0 \) easily implementable subroutines are available, such as the IMSL subroutines, BCONF or NEQNF, or the subroutine by Smith (1983).

Pack (1986(a)) indicated that it would be of interest to examine what effect allowing \( \theta \) to be negative has on the properties of the hypothesis tests. We investigated this and found this effect to be negligible. The number of samples producing negative estimates of \( \theta \) is small compared to the number of samples used in the simulation study. This number becomes even smaller when we take a larger \( \theta \).
Table 3.1: Empirical levels (\%) : $\alpha = 0.05$; based on 1000 replications. $T=2$ groups. $m_1 = 27$, $m_2 = 19.$

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>Test</th>
<th>Statistics</th>
<th>$\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.06</td>
</tr>
<tr>
<td>0.02</td>
<td>LR</td>
<td>4.3</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td>$C_m$</td>
<td>3.7</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>$C_{mm}$</td>
<td>2.7</td>
<td>3.5</td>
</tr>
<tr>
<td>0.06</td>
<td>LR</td>
<td>6.0</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td>$C_m$</td>
<td>5.5</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td>$C_{mm}$</td>
<td>5.1</td>
<td>6.1</td>
</tr>
<tr>
<td>0.19</td>
<td>LR</td>
<td>5.4</td>
<td>5.6</td>
</tr>
<tr>
<td></td>
<td>$C_m$</td>
<td>4.8</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>$C_{mm}$</td>
<td>5.1</td>
<td>4.7</td>
</tr>
<tr>
<td>0.20</td>
<td>LR</td>
<td>4.8</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td>$C_m$</td>
<td>4.8</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td>$C_{mm}$</td>
<td>4.8</td>
<td>6.1</td>
</tr>
</tbody>
</table>
Table 3.2: Empirical power (%): $\alpha = 0.05$; based on 1000 replications. $T=2$
groups. $m_1 = 27$, $m_2 = 19$.

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>Test</th>
<th>$\pi_1, \pi_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistics</td>
<td>(0.08,0.10)</td>
</tr>
<tr>
<td>0.02</td>
<td>LR</td>
<td>8.7</td>
</tr>
<tr>
<td></td>
<td>$C_m$</td>
<td>8.8</td>
</tr>
<tr>
<td></td>
<td>$C_{mm}$</td>
<td>7.0</td>
</tr>
<tr>
<td>0.06</td>
<td>LR</td>
<td>8.9</td>
</tr>
<tr>
<td></td>
<td>$C_m$</td>
<td>9.0</td>
</tr>
<tr>
<td></td>
<td>$C_{mm}$</td>
<td>7.0</td>
</tr>
<tr>
<td>0.09</td>
<td>LR</td>
<td>9.1</td>
</tr>
<tr>
<td></td>
<td>$C_m$</td>
<td>10.1</td>
</tr>
<tr>
<td></td>
<td>$C_{mm}$</td>
<td>9.9</td>
</tr>
<tr>
<td>0.12</td>
<td>LR</td>
<td>9.2</td>
</tr>
<tr>
<td></td>
<td>$C_m$</td>
<td>8.4</td>
</tr>
<tr>
<td></td>
<td>$C_{mm}$</td>
<td>8.4</td>
</tr>
<tr>
<td>0.16</td>
<td>LR</td>
<td>9.0</td>
</tr>
<tr>
<td></td>
<td>$C_m$</td>
<td>8.4</td>
</tr>
<tr>
<td></td>
<td>$C_{mm}$</td>
<td>9.0</td>
</tr>
</tbody>
</table>
3.3 Semi-parametric methods

3.3.1 The $C(\alpha)$ statistics based on the quasi-likelihood

Following the parametrization used in section 3.2.2. the quasi log-likelihood (derived in Chapter 2) can be written as

$$Q = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{1}{1 + (n_{ij} - 1) \phi} \left[ y_{ij} \log \left( \frac{(\pi + \tau_i)}{z_{ij}} \right) + (n_{ij} - y_{ij}) \log \left( \frac{(1 - \pi - \tau_i)}{(1 - z_{ij})} \right) \right]$$

The likelihood score analogs corresponding to the parameters $\tau_1, \ldots, \tau_{T-1}, \pi$ are the quasi likelihood scores

$$g_{1i} = \frac{\partial Q}{\partial \tau_i} = \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi - \tau_i) n_{ij}}{(\pi + \tau_i)(1 - \pi - \tau_i)(1 + (n_{ij} - 1) \phi)}.$$  

and

$$g_{21} = \frac{\partial Q}{\partial \pi} = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi - \tau_i) n_{ij}}{(\pi + \tau_i)(1 - \pi - \tau_i)(1 + (n_{ij} - 1) \phi)}.$$

Note. for given $\phi$, $g_{1i}$, $i = 1, \ldots, T - 1$ and $g_{21}$ are unbiased estimating functions for the parameters $\tau_1, \ldots, \tau_{T-1}, \pi$. No such estimating function for $\phi$ can be obtained from $Q$. However, an unbiased estimating function for $\phi$, given $\tau_1, \ldots, \tau_{T-1}, \pi$ can be obtained by using the moment method (Breslow, 1990(a); Moore and Tsiatis, 1991).

$$g_{22} = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi - \tau_i)^2 n_{ij}}{(\pi + \tau_i)(1 - \pi - \tau_i)(1 + (n_{ij} - 1) \phi)} - (n - T).$$

where $n = \sum_{i=1}^{T} \sum_{j=1}^{m_i} n_{ij}$. Now, define $\tau = (\tau_1, \ldots, \tau_{T-1})$ and $\lambda = (\lambda_1, \lambda_2) = (\pi, \phi)$. Then we let

$$\psi_i = g_{1i} \bigg|_{\tau = 0}, \quad i = 1, \ldots, T - 1$$

and

$$\gamma_k = g_{2k} \bigg|_{\tau = 0}, \quad k = 1, 2.$$
Further define $\psi = (\psi_1, \ldots, \psi_{T-1})'$. Then asymptotically, as $m_i \to \infty$, using the Lindeberg Central-limit theorem, $\psi(\lambda) \sim N(0, \Delta_{11} - \Delta_{12} \Delta_{22}^{-1} \Delta_{21})$ where

$$
\Delta_{11} \equiv E \left( \frac{-\partial g_{1i}}{\partial \tau_i} \right)_{\tau=0}, \quad i, t = 1, \ldots, T - 1.
$$

$$
\Delta_{12} \equiv E \left( \frac{-\partial g_{1i}}{\partial \lambda_k} \right)_{\tau=0}, \quad i = 1, \ldots, T - 1, \quad k = 1, 2
$$

$$
\Delta_{21} \equiv E \left( \frac{-\partial g_{2i}}{\partial \tau_i} \right)_{\tau=0}, \quad i = 1, \ldots, T - 1, \quad k = 1, 2.
$$

and

$$
\Delta_{22} \equiv E \left( \frac{-\partial g_{2i}}{\partial \lambda_s} \right)_{\tau=0}, \quad k, s = 1, 2.
$$

The dimensions of $\Delta_{11}$, $\Delta_{12}$ and $\Delta_{22}$ are $(T-1) \times (T-1)$, $(T-1) \times 2$ and $2 \times 2$ respectively. If $\lambda$ in $\psi$, $\Delta_{11}$, $\Delta_{12}$, $\Delta_{21}$ and $\Delta_{22}$ is replaced by some $\sqrt{m}$ consistent estimate of $\lambda$ under the null hypothesis, then the quasi-likelihood score or $C(\alpha)$ statistic is

$$
C_Q = \psi' (\Delta_{11} - \Delta_{12} \Delta_{22}^{-1} \Delta_{21})^{-1} \psi. \quad (3.3.1.1)
$$

which has asymptotically a chi-square distribution with $(T-1)$ degrees of freedom.

Now, under $H_0$, the unbiased estimating equations for $\pi$ and $\phi$ are

$$
\sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi) n_{ij}}{\pi (1 - \pi) \{1 + (n_{ij} - 1) \phi\}} = 0 \quad (3.3.1.2)
$$

and

$$
\sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi)^2 n_{ij}}{\pi (1 - \pi) \{1 + (n_{ij} - 1) \phi\}} - (n - 2) = 0. \quad (3.3.1.3)
$$

The moment estimate $\hat{\lambda} = (\hat{\pi}, \hat{\phi})$ is the simultaneous solution, if it exists, of the equations (3.3.1.2) and (3.3.1.3) which is $\sqrt{m}$ consistent (Moore, 1986). The elements of $\psi$, $\Delta_{11}$, $\Delta_{12}$, $\Delta_{21}$, and $\Delta_{22}$ for the $C_Q$ statistic are

$$
\psi_i = g_{1i} \bigg|_{\tau=0} = \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi) n_{ij}}{\pi (1 - \pi) \{1 + (n_{ij} - 1) \phi\}}, \quad i = 1, \ldots, T.
$$
It is easily verified that for \( i, t = 1, \ldots, T - 1 \).

\[
\begin{align*}
\Delta_{11tt} &= E\left( \frac{-\partial g_{1i}}{\partial r_i} \bigg|_{r=0} \right) = \sum_{j=1}^{m_i} \frac{n_{ij}}{\pi (1 - \pi)} \left\{ 1 + (n_{ij} - 1) \frac{\phi}{(n_{ij} - 1) \phi} \right\} = d_i,
\Delta_{11it} &= 0 \quad i \neq t \\
\Delta_{21i1} &= E\left( \frac{-\partial g_{2i}}{\partial \lambda_1} \bigg|_{r=0} \right) = \sum_{j=1}^{m_i} \frac{n_{ij}}{\pi (1 - \pi)} \left\{ 1 + (n_{ij} - 1) \frac{\phi}{(n_{ij} - 1) \phi} \right\} = d_i.
\Delta_{21i2} &= E\left( \frac{-\partial g_{2i}}{\partial \lambda_2} \bigg|_{r=0} \right) = 0.
\Delta_{221i} &= E\left( \frac{-\partial g_{22}}{\partial r_i} \bigg|_{r=0} \right) = \sum_{j=1}^{m_i} \frac{n_{ij}}{\pi (1 - \pi)} \left\{ 1 + (n_{ij} - 1) \frac{\phi}{(n_{ij} - 1) \phi} \right\} = d_i.
\Delta_{222i} &= E\left( \frac{-\partial g_{22}}{\partial \lambda_1} \bigg|_{r=0} \right) = \sum_{j=1}^{m_i} \frac{(1 - 2\pi)}{\pi (1 - \pi)} = cm_i.
\end{align*}
\]

where \( C = \frac{(1 - 2\pi)}{\pi (1 - \pi)} \).

\[
\begin{align*}
\Delta_{2211} &= E\left( \frac{-\partial g_{21}}{\partial \lambda_1} \bigg|_{r=0} \right) = \sum_{i=1}^{T} d_i = d.
\Delta_{2212} &= E\left( \frac{-\partial g_{21}}{\partial \lambda_2} \bigg|_{r=0} \right) = 0.
\Delta_{2221} &= E\left( \frac{-\partial g_{22}}{\partial \lambda_1} \bigg|_{r=0} \right) = \sum_{i=1}^{T} cm_i = nc.
\Delta_{2222} &= E\left( \frac{-\partial g_{22}}{\partial \lambda_2} \bigg|_{r=0} \right) = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{(n_{ij} - 1)}{\{ 1 + (n_{ij} - 1) \phi \}}.
\end{align*}
\]

Using these results, we see that

\[
\Delta_{12} \Delta_{22}^{-1} \Delta_{21} = \Delta_{12} \Delta_{22}^{-1} \Delta_{12}^{t}.
\]

and

\[
\begin{bmatrix}
d_1 & 0 \\
\cdot & \cdot \\
\cdot & \cdot \\
d_{T-1} & 0
\end{bmatrix}
= \begin{bmatrix}
d & 0 \\
\cdot & \cdot \\
nc & e \\
\cdot & \cdot \\
d_{T-1} & 0
\end{bmatrix}^{-1}
= \begin{bmatrix}
d_1 & \cdot & \cdot & d_{T-1}
\end{bmatrix}
\]

87
\[
\begin{bmatrix}
\frac{d_1^2}{d_1} & d_1 d_2 & \ldots & d_1 d_{T-1} & d_1 d_T
\\
d_1 d_2 & d_2^2 & \ldots & d_2 d_{T-1} & d_2 d_T
\\
\vdots & \vdots & \ddots & \vdots & \vdots
\\
\vdots & \vdots & \ddots & \vdots & \vdots
\\
\vdots & \vdots & \ddots & \vdots & \vdots
\\
d_1 d_{T-1} & d_2 d_{T-1} & \ldots & d_{T-1}^2 & d_{T-1} d_T
\end{bmatrix}
\]

which can be expressed as \(\alpha ab'\), where \(\alpha = d^{-1}\) and \(a = b = (d_1, \ldots, d_{T-1})'\). Again, using the inverse of a partitioned matrix, we obtain.

\[
(\Delta_{11} - \Delta_{12} \Delta_{22}^{-1} \Delta_{21})^{-1} = \Delta_{11}^{-1} + \mathbb{1}' / d_T
\]

where \(\mathbb{1}\) is a (T-1) column vectors of 1s. Thus,

\[
\psi' (\Delta_{11} - \Delta_{12} \Delta_{22}^{-1} \Delta_{21})^{-1} \psi = \sum_{i=1}^{T} \frac{\psi_i^2}{d_i} + \left( \sum_{i=1}^{T-1} \psi_i \right)^2 / d_T.
\]

Now, from equation (3.3.1.2) we have

\[
\psi_T = - \sum_{i=1}^{T-1} \psi_i.
\]

Thus,

\[
\psi' (\Delta_{11} - \Delta_{12} \Delta_{22}^{-1} \Delta_{21})^{-1} \psi = \sum_{i=1}^{T-1} \frac{\psi_i^2}{d_i} + \frac{\psi_T^2}{d_T} = \sum_{i=1}^{T} \frac{\psi_i^2}{d_i}.
\]

Then, after simplifications we find that the \(C(\alpha)\) statistic (3.3.1.1) reduces to

\[
C_Q = \frac{1}{\bar{y}(1 - \bar{y})} \sum_{i=1}^{T} \left\{ \frac{\sum_{j=1}^{m_i} (z_{ij} - \bar{y}) n_{ij} \left\{ 1 + (n_{ij} - 1) \hat{\phi} \right\}^{-1} \right\}^2 \sum_{j=1}^{m_i} n_{ij} \left\{ 1 + (n_{ij} - 1) \hat{\phi} \right\}^{-1} \}
\]

88
3.3.2 The C(α) statistic based on the extended Quasi-likelihood

We reparameterize \( \pi_i \) as in section 3.2.2. The extended quasi log-likelihood (as derived in chapter 2) excluding a constant term, in terms of the parameters \( \tau_1, \ldots, \tau_{T-1}, \pi \) and \( \phi \) can be written as

\[
Q^+ = -\frac{1}{2} \sum_{i=1}^{T} \sum_{j=1}^{m_i} \left[ \log \{1 + (n_{ij} - 1) \phi\} + \frac{2}{\{1 + (n_{ij} - 1) \phi\}} y_{ij} \log \{(\pi + \tau_i)/z_{ij}\}
\]

\[
+ (n_{ij} - y_{ij}) \log \{(1 - \pi - \tau_i)/(1 - z_{ij})\} \right].
\]  

(3.3.2.1)

Now, define \( \tau = (\tau_1, \ldots, \tau_{T-1}) \) and \( \lambda = (\lambda_1, \lambda_2) = (\pi, \phi) \) and let

\[
\psi_i = \frac{\partial Q^+}{\partial \tau_i} \bigg|_{r=0}, \quad i = 1, \ldots, T - 1
\]

and

\[
\gamma_k = \frac{\partial Q^+}{\partial \lambda_k} \bigg|_{r=0}, \quad k = 1, 2.
\]

Again, as in section (3.3.1), define \( \psi = (\psi_1, \ldots, \psi_{T-1})' \). Then asymptotically the variance covariance of \( \psi(\lambda) \) is \( \Delta_{11} = \Delta_{12} \Delta_{22}^{-1} \Delta_{21} \), where

\[
\Delta_{11it} = E \left( \frac{-\partial^2 Q^+}{\partial \tau_i \partial \tau_t} \bigg|_{r=0} \right), \quad i, t = 1, \ldots, T - 1.
\]

\[
\Delta_{12ik} = E \left( \frac{-\partial^2 Q^+}{\partial \tau_i \partial \lambda_k} \bigg|_{r=0} \right), \quad i = 1, \ldots, T - 1, \quad k = 1, 2
\]

\[
\Delta_{21ki} = E \left( \frac{-\partial^2 Q^+}{\partial \lambda_k \partial \tau_i} \bigg|_{r=0} \right), \quad i = 1, \ldots, T - 1, \quad k = 1, 2
\]

and

\[
\Delta_{22ks} = E \left( \frac{-\partial^2 Q^+}{\partial \lambda_k \partial \lambda_s} \bigg|_{r=0} \right), \quad k, s = 1, 2.
\]

The dimensions of \( \Delta_{11}, \Delta_{12}, \Delta_{21}, \) and \( \Delta_{22} \) are \((T - 1) \times (T - 1)\), \((T - 1) \times 2\), \(2 \times (T - 1)\) and \(2 \times 2\) respectively. Now, if \( \lambda \) in \( \psi \), \( \Delta_{11}, \Delta_{12}, \Delta_{21} \) and \( \Delta_{22} \) is replaced by the maximum extended quasi-likelihood estimate \( \widetilde{\lambda} = (\widetilde{\pi}, \widetilde{\phi}) \), obtained by solving

\[
\sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi) n_{ij}}{\pi (1 - \pi) \{1 + (n_{ij} - 1) \phi\}} = 0
\]  

(3.3.2.2)
and

\[
\sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{(n_{ij} - 1)}{\{1 + (n_{ij} - 1) \phi\}^2} \left[ y_{ij} \log \left( \frac{z_{ij}}{\pi} \right) + (n_{ij} - y_{ij}) \log \left( \frac{1 - z_{ij}}{1 - \pi} \right) \right] \\
- \left\{1 + (n_{ij} - 1) \phi \right\} / 2 = 0
\]  

(3.3.2.3)

simultaneously, which is \(\sqrt{m}\) consistent, then the \(C(\alpha)\) or score statistic based on the extended quasi-likelihood is

\[
C_{Q^+} = \psi' \left( \Delta_{11} - \Delta_{12} \Delta_{22}^{-1} \Delta_{21} \right)^{-1} \psi.
\]

which has asymptotically a chi-square distribution with \((T - 1)\) degrees of freedom.

Now, for \(i = 1, \ldots, T\),

\[
\psi_i = \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi) n_{ij}}{\pi (1 - \pi) \{1 + (n_{ij} - 1) \phi\}}, \quad i = 1, \ldots, T.
\]

Further, it can be easily verified that for \(i, t = 1, \ldots, T - 1\),

\[
\left. \frac{-\partial^2 Q^+}{\partial \tau_i^2} \right|_{\tau = 0} = \sum_{j=1}^{m_i} \frac{1}{\{1 + (n_{ij} - 1) \phi\}^2} \left\{ \frac{y_{ij}}{\pi^2} + \frac{(n_{ij} - y_{ij})}{(1 - \pi)^2} \right\}.
\]

\[
\left. \frac{-\partial^2 Q^+}{\partial \tau_i \partial \lambda_1} \right|_{\tau = 0} = \left. \frac{-\partial^2 Q^+}{\partial \tau_i^2} \right|_{\tau = 0}.
\]

\[
\left. \frac{-\partial^2 Q^+}{\partial \lambda_1^2} \right|_{\tau = 0} = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{1}{\{1 + (n_{ij} - 1) \phi\}^2} \left\{ \frac{y_{ij}}{\pi^2} + \frac{(n_{ij} - y_{ij})}{(1 - \pi)^2} \right\}.
\]

\[
\left. \frac{-\partial^2 Q^+}{\partial \tau_i \partial \lambda_2} \right|_{\tau = 0} = \sum_{j=1}^{m_i} \frac{(n_{ij} - 1)}{\{1 + (n_{ij} - 1) \phi\}^2} \left\{ \frac{y_{ij}}{\pi} - \frac{(n_{ij} - y_{ij})}{(1 - \pi)} \right\}.
\]

\[
\left. \frac{-\partial^2 Q^+}{\partial \lambda_2^2} \right|_{\tau = 0} = 2 \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{(n_{ij} - 1)^2}{\{1 + (n_{ij} - 1) \phi\}^3} \left[ \frac{y_{ij} \log \frac{z_{ij}}{\pi}}{\pi} \\
+ (n_{ij} - y_{ij}) \log \left( \frac{1 - z_{ij}}{1 - \pi} \right) \right] - \frac{1}{2} \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{(n_{ij} - 1)^2}{\{1 + (n_{ij} - 1) \phi\}^2},
\]

\[
\left. \frac{-\partial^2 Q^+}{\partial \lambda_1 \partial \lambda_2} \right|_{\tau = 0} = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{(n_{ij} - 1)}{\{1 + (n_{ij} - 1) \phi\}^2} \left[ \frac{y_{ij}}{\pi} - \frac{(n_{ij} - y_{ij})}{(1 - \pi)} \right].
\]
Then, under the null hypothesis, the expected values of the negative of the second derivatives for $i, t = 1, \ldots, T - 1$ are:

$$\Delta_{11ii} = E \left( \frac{-\partial^2 Q^+}{\partial \tau_i^2} \bigg|_{\tau=0} \right) = \sum_{j=1}^{m_i} \frac{n_{ij}}{\pi (1 - \pi) \{1 + (n_{ij} - 1) \phi\}} = d_i.$$  

$\Delta_{11it} = 0.$ for $i \neq t.$

$$\Delta_{12i1} = E \left( \frac{-\partial^2 Q^+}{\partial \tau_i \partial \lambda_1} \bigg|_{\tau=0} \right) = \sum_{j=1}^{m_i} \frac{n_{ij}}{\pi (1 - \pi) \{1 + (n_{ij} - 1) \phi\}} = d_i.$$  

$$\Delta_{12i2} = E \left( \frac{-\partial^2 Q^+}{\partial \tau_i \partial \lambda_2} \bigg|_{\tau=0} \right) = 0.$$  

$$\Delta_{21ii} = \Delta_{12i1}, \quad \Delta_{21i2} = \Delta_{21i2}.$$  

$$\Delta_{2211} = E \left( \frac{-\partial^2 Q^+}{\partial \lambda_1^2} \bigg|_{\tau=0} \right) = \sum_{i=1}^{T} d_i = d.$$  

$$\Delta_{2212} = E \left( \frac{-\partial^2 Q^+}{\partial \lambda_1 \partial \lambda_2} \bigg|_{\tau=0} \right) = 0.$$  

$$\Delta_{2222} = E \left( \frac{-\partial^2 Q^+}{\partial \lambda_2^2} \bigg|_{\tau=0} \right) = \frac{1}{2} \sum_{i=1}^{T} \sum_{j=1}^{m_i} \left( \frac{n_{ij} - 1}{1 + (n_{ij} - 1) \phi} \right)^2 = e.$$  

Therefore,

$$\Delta_{12} \Delta_{22}^{-1} \Delta_{21} = \begin{bmatrix} d_1 & 0 & & & \\ \cdot & \cdot & \cdot & & \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ d_{T-1} & 0 & & & \\ \end{bmatrix} \begin{bmatrix} d & 0 \\ 0 & e \end{bmatrix}^{-1} \begin{bmatrix} d_1 & \cdots & d_{T-1} \\ 0 & \cdots & 0 \end{bmatrix}$$

$$= \begin{bmatrix} d_1^2/d & d_1 d_2/d & \cdots & d_1 d_{T-1}/d \\ d_1 d_2/d & d_2^2/d & \cdots & d_2 d_{T-1}/d \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ d_{1T-1}/d & d_2 d_{T-1}/d & \cdots & d_{T-1}^2/d \end{bmatrix}. $$
As in section (3.4.1), using the inverse of a partitioned matrix, we obtain

\[
(\Delta_{11} - \Delta_{12} \Delta_{22}^{-1} \Delta_{21})^{-1} = \Delta_{11}^{-1} + 1 \, \mathbf{1} \cdot d_T.
\]

After simplifications, the statistic \(C_Q^+\) reduces to

\[
C_Q^+ = \frac{1}{\bar{\pi} (1 - \bar{\pi})} \sum_{i=1}^{T} \left\{ \frac{\left[ \sum_{j=1}^{m_i} (z_{ij} - \bar{\pi}) n_{ij} \left\{ 1 + (n_{ij} - 1) \hat{\phi} \right\}^{-1} \right]^2}{\sum_{j=1}^{m_i} n_{ij} \left\{ 1 + (n_{ij} - 1) \hat{\phi} \right\}^{-1}} \right\}.
\]

Note that \(C_Q\) and \(C_Q^+\) have exactly the same simplified form. They only differ in the use of the estimates. Note, it can be shown that the asymptotic \(\text{var}(\tilde{\lambda}) \to 0\) as \(m \to \infty\). So, \(\tilde{\lambda}\) is \(\sqrt{m}\) consistent. However, this asymptotic result is dependent on large \(n_{ij}\) (see Appendix B). Thus, \(C_Q^+\) will approximate \(\chi^2 (T - 1)\) if \(m\) and \(n_{ij}\) are large.

3.3.3 The Rao-Scott (RS) statistics

Rao and Scott (1992) proposed a statistic for comparing homogeneity of proportions in several clusters based on the concept of design effect and effective sample size.

Define

\[
y_i = \sum_j y_{ij}, \quad n_i = \sum_j n_{ij}, \quad \hat{p}_i = y_i / n_i,
\]

\[
\forall_i = m_i (m_i - 1)^{-1} n_i^{-2} \sum_j (y_{ij} - n_{ij} \hat{p}_i)^2,
\]

\[
d_i = n_i \forall_i / \left\{ \hat{p}_i (1 - \hat{p}_i) \right\},
\]

\[
\bar{y}_i = y_i / d_i, \quad \text{and} \quad \bar{n}_i = n_i / d_i.
\]

The quantity \(d_i\) is called the design effect (variance inflation factor) and the quantity \(\bar{n}_i\) is called the effective sample size in the context of survey sampling (Kish, 1965).

The RS statistic for testing homogeneity of proportions is

\[
RS = \sum_{i=1}^{T} \left\{ (\bar{y}_i - \bar{n}_i \bar{p})^2 / \{\bar{n}_i \bar{p}(1 - \bar{p})\} \right\},
\]

92
where $\tilde{p} = \sum \tilde{y}_i / \sum \tilde{n}_i$. It is clear that the RS procedure does not assume any specific model for the intracluster correlations. It essentially uses the binomial model for the overall response $y_i$ in the $i$-th cluster after adjusting for the variance inflation due to clustering or the design effect. Thus, RS statistic can also be considered to be based on semi-parametric assumptions. Under the hypothesis of homogeneity of proportions, the statistic RS is distributed asymptotically as $\chi^2(T - 1)$. In the special case in which the population inflation factors $d_i$ are equal, say $d_i = d$ for $i = 1, \ldots, T$. Rao and Scott suggest using $\tilde{u}_i = y_i / d$ and $\tilde{n}_i = n_i / d$, where $d$ is a pooled estimate given by

$$d = \left[ \sum_{i=1}^{T} (1 - f_i) \frac{\tilde{p}_i (1 - \tilde{p}_i)}{\hat{p} (1 - \hat{p})} d_i \right] / (T - 1)$$

with $f_i = n_i / n$ and $n = \sum n_i$. In this case RS reduces to $ARS = \chi^2 / d$, where

$$\chi^2 = \sum_{i=1}^{T} \frac{(y_i - n_i \hat{p})^2}{n_i \hat{p} (1 - \hat{p})}.$$

with $\hat{p} = \sum y_i / \sum n_i$. We call ARS the adjusted Rao-Scott statistic. Again, under the hypothesis of homogeneity of proportions, ARS is asymptotically distributed as $\chi^2(T - 1)$. Both statistics RS and ARS will be included in the simulation study in section (3.4.5).

3.3.4 The Empirical Score Test

3.3.4.1 Introduction

Breslow (1989; 1990(a)) suggests using the empirical variance covariance matrix to attain robustness against variance misspecifications. The score test based on the empirical variances is known as empirical score test $C_{EQ}$ or it is also known as the generalized score test (Boos, 1992). It also retains the invariance property. The likelihood ratio statistic under misspecification usually has a non-standard null
asymptotic distribution (Kent, 1982; White, 1982). The empirical score test is preferred to the likelihood ratio test because of a simpler asymptotic distribution. The empirical score statistic arises from Taylor’s expansion of the estimating equations. This statistic is not as simple as the model-based score statistic $C_Q$, but due to its asymptotically type I error robustness, it is attractive for general use. It also maintains the computational appeal of the score statistics in that only null hypothesis estimation of the parameters is required. We follow the same procedure as described by Breslow (1989; 1990(a)) to derive the empirical score test.

3.3.4.2 The Empirical Score Test

We reparametrize $\pi_i$, under $H_A$ by $\pi_i = \pi + \tau_i$ with $\tau_T = 0$. Then testing $H_0$ is equivalent to testing

$$H_0 : \quad \tau_i = 0 \quad i = 1, \ldots, T - 1$$

$$H_A : \quad \tau_i \neq 0.$$ 

with $\pi$ treated as nuisance parameter. Define $\tau = (\tau_1, \ldots, \tau_{T-1})$ and $\lambda = (\lambda_1, \lambda_2) = (\tau, \pi)$. The unbiased estimating equations for $\tau_1, \ldots, \tau_{T-1}$ and $\pi$ are

$$g_{1i}(\tau, \phi) = \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi - \tau_i)n_{ij}}{(\pi + \tau_i)(1 - \pi + \tau_i) \{1 + (n_{ij} - 1)\phi\}}, \quad i = 1, \ldots, T - 1$$

and

$$g_{21}(\tau, \phi) = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi - \tau_i)n_{ij}}{(\pi + \tau_i)(1 - \pi - \tau_i) \{1 + (n_{ij} - 1)\phi\}}.$$

$E\left(\frac{\partial g_{1i}}{\partial \phi}\right) = 0$, together with sufficient regularity conditions imply that for any sequence of estimates $\hat{\tau}_i, i = 1, \ldots, T - 1$ converging to $\tau^0$, the statistics $g_{1i}\left(\hat{\tau}, \hat{\phi}\right)$ and $g_{1i}(\tau, \phi^*)$ are asymptotically equivalent where $\phi^*$ is the limiting value. Therefore, substituting $\hat{\phi}$ for $\phi^*$ does not affect the asymptotic distribution of the score test statistics that we derive from $g_{1i}(\tau, \phi)$. Under $H_0$, we expand $g_{1i}$ and $g_{21}$ with
respect to \( \lambda_2 \)

\[
g_{1i} (\tilde{\lambda}) = g_{1i} (\lambda^*) + \frac{\partial g_{1i}}{\partial \pi} (\tilde{\lambda}_2 - \lambda_2^*) + O_p(1) \tag{3.3.4.2.1}
\]

\[
0 = g_{21} (\tilde{\lambda}) = g_{21} (\lambda^*) + \frac{\partial g_{21}}{\partial \pi} (\tilde{\lambda}_2 - \lambda_2^*) + O_p(1). \tag{3.3.4.2.2}
\]

From equation (3.3.4.2.2) we obtain

\[
(\tilde{\lambda}_2 - \lambda_2^*) = -\left( \frac{\partial g_{21}}{\partial \pi} \right)^{-1} g_{21} (\lambda^*). \tag{3.3.4.2.3}
\]

Substituting equation (3.3.4.2.3) into equation (3.4.4.2.1) we obtain

\[
g_{1i} (\tilde{\lambda}) = g_{1i} (\lambda^*) - \left( \frac{\partial g_{1i}}{\partial \pi} \right) \left( \frac{\partial g_{21}}{\partial \pi} \right)^{-1} g_{21} (\lambda^*). \tag{3.3.4.2.4}
\]

Let

\[
g_1 (\tilde{\lambda}) = (g_{11} (\tilde{\lambda}), \ldots, g_{1T-1} (\tilde{\lambda}))
\]

and

\[
g_2 (\tilde{\lambda}) = g_{21} (\tilde{\lambda}).
\]

Then, replacing \( \left( \frac{\partial g_{1i}}{\partial \pi} \right) \) and \( \left( \frac{\partial g_{21}}{\partial \pi} \right)^{-1} \) by their asymptotically equivalent versions, equation (3.3.4.2.4) can be written as

\[
g_{1i} (\tilde{\lambda}) = (I_{T-1} - A_{12} A_{22}^{-1}) \left( g_{1} (\lambda^*) \right),
\]

where the dimension of \( A_{12} \) and \( A_{22} \) are \((T-1) \times 1\) and \(1 \times 1\) with elements

\[
A_{12} = E \left( \frac{-\partial g_{1i}}{\partial \pi} \right), \quad i = 1, \ldots, T - 1
\]

\[
A_{22} = E \left( \frac{-\partial g_{21}}{\partial \pi} \right),
\]

where \( I_{T-1} \) is the \((T - 1) \times (T - 1)\) identity matrix. It follows that the asymptotic distribution of the score statistic \( g_{1i} \) is multivariate normal with zero mean and
covariance matrix given by

\[
\text{var} \left( g_{1i} \left( \lambda \right) \right) = (I - A_{12}A_{22}^{-1}) \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} (I - A_{12}A_{22}^{-1})
\]

\[
= G_{11} - A_{12}A_{22}^{-1}G_{12}' - G_{12}A_{22}^{-1}A_{12}' + A_{12}A_{22}^{-1}G_{22}A_{22}^{-1}A_{12}'
\]

where \( G_{11}, G_{22}, \) and \( G_{12} \) are the empirical covariance matrix of the score \( g_{1i} \) and \( g_{21} \) with dimensions \((T - 1) \times (T - 1)\), \((T - 1) \times 1\) and \(1 \times 1\) respectively. We find

\[
G_{11} = \sum_{i=1}^{T-1} g_{1i}g_{1i}.
\]

\[
G_{12} = \sum_{i=1}^{T-1} g_{1i}g_{21}.
\]

\[
G_{22} = g_{21}g_{21}.
\]

The empirical score statistic for testing the null hypothesis \( H_0 \) is

\[
C_{EQ} = g_1^T \left[ G_{11} - A_{12}A_{22}^{-1}G_{12} - G_{12}A_{22}^{-1}A_{12}' + A_{12}A_{22}^{-1}G_{22}A_{22}^{-1}A_{12}' \right]^{-1} g_1.
\]

where \( g_1 = g_{1i} \), \( i = 1, \ldots, T-1 \). Under \( H_0 \) and suitable regularity conditions, \( C_{EQ} \) is asymptotically a \( \chi^2 \) distribution with \((T-1)\) degrees of freedom. Note, if the variance function has been correctly specified so that \( A = G \), the asymptotic variance reduces to the model-based variance form

\[
\text{Var} \left( g_{1i} \right) = A_{11} - A_{12}A_{22}^{-1}A_{21}.
\]

The derivation of the elements \( G_{11}, G_{12}, G_{22}, A_{11}, \) and \( A_{22} \) are given below:

\[
G_{11} = \sum_{j=1}^{m_i} g_{1j}g_{1j}' = \sum_{j=1}^{m_i} \frac{(zi_{ij} - \bar{\pi})^2 n_{ij}^2}{\bar{\pi}(1 - \bar{\pi}) \left( 1 + (n_{ij} - 1) \bar{\phi} \right)^2}, \quad i = 1, \ldots, T - 1
\]

\[
G_{12} = \sum_{j=1}^{m_i} g_{1j}g_{21}' = \sum_{j=1}^{m_i} \frac{(zi_{ij} - \bar{\pi})^2 n_{ij}^2}{\bar{\pi}(1 - \bar{\pi}) \left( 1 + (n_{ij} - 1) \bar{\phi} \right)^2}, \quad i = 1, \ldots, T - 1
\]

\[
G_{22} = \sqrt{\frac{T \sum_{i=1}^{m_i} \left( z_{ij} - \bar{\pi} \right)^2 n_{ij}^2}{\bar{\pi}(1 - \bar{\pi}) \left( 1 + (n_{ij} - 1) \bar{\phi} \right)^2}},
\]
\[ A_{12} = E \left( \frac{-\partial g_{1i}}{\partial \lambda_2} \bigg|_{\tau = 0} \right) = \sum_{j=1}^{m_i} \frac{n_{ij}}{\tilde{\pi}(1 - \tilde{\pi}) \left\{ 1 + (n_{ij} - 1) \phi \right\}}. \quad i = 1, \ldots, T - 1 \]

\[ A_{22} = E \left( \frac{-\partial g_{2i}}{\partial \lambda_2} \bigg|_{\tau = 0} \right) = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{n_{ij}}{\tilde{\pi}(1 - \tilde{\pi}) \left\{ 1 + (n_{ij} - 1) \phi \right\}}. \quad i = 1, \ldots, T - 1. \]

### 3.4.5 Simulations

We are concerned here with a study of the behaviour of the statistics \( C_m, C_{mm}, C_Q, C_{Q+}, C_{EQ}, RS \) and \( ARS \) in terms of size, power, and robustness. For this we conduct a simulation study. For studying robustness we generate data, as in examples given in section (3.4.4.1), from the beta-binomial (BB), the probit normal binomial (PNB), and the logit normal binomial (LNB) distribution. In all examples, we consider \( T = 2 \) groups. The litter sizes and number of litters are chosen as those of the control group \( (m_1 = 27) \) and the low dose treatment group \( (m_2 = 19) \) of Paul (1982). The litter sizes are given in section (3.3.3). The litter sizes and the number of litters here can be representative of those that occur in toxicological experiments.

For the comparison of levels, experiments were conducted for several combinations of common \( \pi \) and \( \phi \) (or common \( \psi \)). Note, for convenience of comparison, results from simulations are presented in terms of \( \theta \). Each experiment was based on 1000 replications. Monte Carlo standard error for size estimate of nominal size 5% corresponds to 0.7%. For power comparison we consider \( \pi_1 = \pi \) and \( \pi_2 = \pi + c \) with \( \pi = 0.05 \) and \( c = 0.02, 0.07, 0.12 \). Simulation experiments for power comparison were run for \( \theta = 0.06, 0.12, 0.16 \). Monte Carlo standard errors for point estimates of powers of magnitude 80%, 50% and 20% are respectively 1.3%, 1.6% and 1.3%.

#### 3.4.5.1 Data generation

**Example 1: The BB Distribution**

Data are generated from the beta-binomial distribution using the IMSL random number generators RNBET and RNBIN as described in section 2.5.2.

**Example 2: The PNB Distribution**
The procedure of generating data from the probit normal distribution is given in section 2.5.2.

Example 3: The LNB Distribution

Data are generated from the logit normal binomial distribution following the three step procedure used to generate data, in section 2.5.2, from the PNB distribution with $\gamma = \text{logit}(\pi)$.

3.4.5.2 Significance Level

Empirical levels corresponding to the nominal level $\alpha = 0.05$ are given in table 3.3 for all the statistics using data from the above three distributions. We note several points from table 3.3.

(i) The statistic RS often shows liberal behaviour.

(ii) The statistic $C_m$ performs well only when the data come from the beta-binomial distribution, otherwise it shows extreme liberal behaviour in most situations.

(iii) The performances of the statistics $C_{mm}, C_{Q+}$ and ARS are similar. They hold nominal level reasonably well except in some situations in which they show some nonconservative behaviour.

(iv) The statistic $C_{EQ}$ holds nominal level quite well in all data distribution situations.

(v) Data departures, i.e., data coming from distribution other than BB model, have little effect on the performance of the statistics $C_{mm}, C_Q, C_{Q+}, C_{EQ}, RS$ and ARS.

(vi) The performance of the statistic $C_Q$ is best in that it holds nominal level quite well in all data distribution situations with no apparent nonconservative behaviour. It also has a simple form.

3.4.5.3 Power

The empirical power corresponding to the nominal significance level $\alpha = 0.05$
has been computed for all the statistics. Power estimates for the statistics $C_m$ and RS were highest. These estimates are erroneous because these statistics are liberal. Power estimates for the statistics $C_m$, $C_{mm}$, $C_Q$, $C_{Q+}$ RS and ARS are given in table 3.4. For data generated from the BB and the PNB distributions the power performances of the statistics $C_{mm}$, $C_Q$ and $C_{Q+}$ are similar. The power of the statistic ARS is somewhat lower than that of the other three. For data coming from the LNB distribution, the statistic $C_Q$ has some advantage over the other statistics. The difference of the power estimates of the statistics $C_Q$ and ARS is often more than twice the estimated standard errors.

3.4.5.4 Robustness Under Dispersion Heterogeneity

We have seen in section (3.4.5.2) and (3.4.5.3) that the statistic $C_Q$ performs best in terms of level, power and also robustness for departure of data from the assumed distribution. It also has a simple form. The performance, in terms of size and robustness, of the empirical statistic $C_{EQ}$ (derived in section 3.4.4.1), is similar to the model-based score statistic $C_Q$. It does not have as simple a form as $C_Q$. From table 3.3, we see that the statistics RS and ARS show mild to moderate liberal behaviour in most situations. However, these statistics are very simple to use. In this section we present results of a limited simulation study to determine robustness of the statistics $C_Q$, $C_{EQ}$, RS and ARS under heterogeneous dispersions. Estimates of levels based on a nominal level of $\alpha = 0.05$ have been computed for the common $\pi = 0.06, 0.10$ and 0.20 and for the combinations of $(\theta_1, \theta_2) = (0.06..08), (0.06..10), (0.06..12), (0.06..20), (0.06..40), (0.06..80)$ using data simulated from the BB, PNB and the LNB distributions as described in section 3.4.4.1. The results are given in table 3.5. The results show evidence that under moderate departure from homogeneity of dispersion the statistics $C_Q$ and $C_{EQ}$ hold nominal level extremely well. The liberal behaviour of the statistics RS and ARS becomes more evident as the departure from
homogeneity of dispersion increases. Further investigation with other values of \( \pi \) and \((\theta_1, \theta_2)\) as in table 3.5 indicated similar performance. However, the departure of \( \theta_2 \) from \( \theta_1 \) further than that given in table 3.5 produces level estimates that show liberal behaviour. As a general investigation we considered different values of \( \pi \) and \( \theta_1 \) and \( \theta_2 = \theta_1(1 + c) \), with \( c = n^{-1/4}, n^{-1/8} \). 1 where \( n = n_1 + n_2 \). Simulation indicated the kind of robustness that is evident in table 3.5.

3.4.6 Examples and Discussion

We consider the toxicological data in Paul (1982). The data consist of live fociuses in a litter affected by treatment, and the number of live fociuses, for each of \( T = 4 \) dose groups: control (C), low dose (L), medium dose (M) and high dose (H). The data are given in table 3.6. Analysis of the data using the likelihood ratio procedure, based on the beta-binomial model, indicates evidence of homogeneity of the dispersion parameters, either across groups or between any pair of groups (see. Paul, 1982, Table 5a). For testing homogeneity of proportions across groups or between any pair of groups we apply the recommended statistic \( C_Q \). Under the null hypothesis of homogeneity of proportions with the assumption of a constant dispersion parameter we need estimates of \( \pi \) and \( \theta \) using the quasi-likelihood estimating equations. The estimates of the parameters and the value of the \( C_Q \) statistic are given in table 3.7. A comparison of the results in table 3.7 and those in table 5a of Paul (1982) indicate that for the toxicological data in table 3.6 the quasi-likelihood based score procedure results in similar conclusions, regarding treatment differences, as those of the likelihood ratio procedure based on the beta-binomial model. The RS and ARS show liberal behaviour.

As discussed earlier, the likelihood ratio procedure needs the assumption of a beta-binomial model for the data and it requires estimates of the parameters both under the null and the alternative hypothesis. The Rao-Scott statistics are very
simple to use. However, in some instances, these show liberal behaviour. See table 3.3 and 3.7 for specific examples. Under homogeneous dispersion the statistic ARS behaves slightly better than the statistic RS in maintaining γ level. However, its power performance is somewhat less satisfactory compared to other statistics, particularly compared to the $C_Q$ statistic for data coming from the LNB distribution. The somewhat poor performance of RS statistics may be attributed to the fact that these are derived using the central limit theorem on the assumption that the number of litters $m_i; i = 1, \ldots, T$, is large. In practice, in some instances, $m_i$ is not large enough for the central limit theorem to hold. The $C_{Q+}$ statistic has a simple form and needs fewer moment assumptions. Its performance is also unsatisfactory. The poor performance of this statistic may be attributed to the fact that we need both $m$ and $n_{ij}$ large. In practice the $n_{ij}$ are rarely large enough for the Taylor’s approximation to hold.

The $C_Q$ statistic needs fewer assumptions compared to the beta-binomial model. The maximum quasi-likelihood estimates of the parameters under the null hypothesis required for the $C_Q$ statistic are more simply obtained than the corresponding maximum likelihood estimates under the beta-binomial model. The $C_Q$ statistic has a simple form, holds the nominal level most effectively and has some advantage in power over the statistics in some data distribution situations. The $C_Q$ statistic also shows evidence of robustness in the presence of some dispersion heterogeneity.

The performance in terms of size and robustness of the empirical score statistic $C_{EQ}$ under dispersion homogeneity was similar to that of the model based $C_Q$ statistic. However, the drawback of the empirical score statistic $C_{EQ}$ is that it does not provide as simple a form as the model-based $C_Q$ statistic. The empirical score statistic $C_{EQ}$ can only be recommended in situations where there is a lack of knowledge of the variance structure or for other model inadequacies.
Table 3.3: Empirical level(%) $\alpha = 0.05; T=2$ groups based on 1000 replications

<table>
<thead>
<tr>
<th>Test statistic</th>
<th>$\pi$</th>
<th>(a) BB distribution</th>
<th>(b) PNB distribution</th>
<th>(c) LNB distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\theta$</td>
<td>0.06</td>
<td>0.10</td>
<td>0.20</td>
</tr>
<tr>
<td>$C_m$</td>
<td>0.06</td>
<td>5.5</td>
<td>5.9</td>
<td>5.5</td>
</tr>
<tr>
<td>$C_{mm}$</td>
<td></td>
<td>5.1</td>
<td>5.2</td>
<td>5.2</td>
</tr>
<tr>
<td>$C_Q$</td>
<td></td>
<td>4.3</td>
<td>4.6</td>
<td>5.1</td>
</tr>
<tr>
<td>$C_{Q+}$</td>
<td></td>
<td>6.8</td>
<td>5.1</td>
<td>4.3</td>
</tr>
<tr>
<td>$C_{EQ}$</td>
<td></td>
<td>4.9</td>
<td>5.3</td>
<td>5.8</td>
</tr>
<tr>
<td>RS</td>
<td></td>
<td>6.7</td>
<td>6.7</td>
<td>5.9</td>
</tr>
<tr>
<td>ARS</td>
<td></td>
<td>5.9</td>
<td>6.1</td>
<td>6.4</td>
</tr>
<tr>
<td>$C_m$</td>
<td>0.10</td>
<td>4.8</td>
<td>5.4</td>
<td>5.0</td>
</tr>
<tr>
<td>$C_{mm}$</td>
<td></td>
<td>5.1</td>
<td>5.3</td>
<td>4.9</td>
</tr>
<tr>
<td>$C_Q$</td>
<td></td>
<td>4.6</td>
<td>4.5</td>
<td>4.7</td>
</tr>
<tr>
<td>$C_{Q+}$</td>
<td></td>
<td>8.0</td>
<td>5.5</td>
<td>3.4</td>
</tr>
<tr>
<td>$C_{EQ}$</td>
<td></td>
<td>5.6</td>
<td>5.5</td>
<td>5.0</td>
</tr>
<tr>
<td>RS</td>
<td></td>
<td>6.7</td>
<td>6.6</td>
<td>6.0</td>
</tr>
<tr>
<td>ARS</td>
<td></td>
<td>5.2</td>
<td>5.7</td>
<td>5.0</td>
</tr>
<tr>
<td>$C_m$</td>
<td>0.20</td>
<td>4.8</td>
<td>7.5</td>
<td>5.9</td>
</tr>
<tr>
<td>$C_{mm}$</td>
<td></td>
<td>4.8</td>
<td>7.5</td>
<td>5.4</td>
</tr>
<tr>
<td>$C_Q$</td>
<td></td>
<td>3.6</td>
<td>5.7</td>
<td>5.2</td>
</tr>
<tr>
<td>$C_{Q+}$</td>
<td></td>
<td>8.3</td>
<td>8.0</td>
<td>4.8</td>
</tr>
<tr>
<td>$C_{EQ}$</td>
<td></td>
<td>4.8</td>
<td>7.0</td>
<td>5.9</td>
</tr>
<tr>
<td>RS</td>
<td></td>
<td>8.1</td>
<td>8.8</td>
<td>7.0</td>
</tr>
<tr>
<td>ARS</td>
<td></td>
<td>5.2</td>
<td>7.7</td>
<td>6.5</td>
</tr>
</tbody>
</table>
Table 3.4: Empirical power(%) \( \alpha = 0.05 \): \( T = 2 \) groups: based on 1000 replications:

\[ \pi_1 = \pi = 0.08; \pi_2 = \pi + c. \]

<table>
<thead>
<tr>
<th>Test statistic</th>
<th>( c )</th>
<th>(a) BB distribution</th>
<th>(b) PNB distribution</th>
<th>(c) LNB distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \theta )</td>
<td>0.02 0.07 0.12</td>
<td>0.02 0.07 0.12</td>
<td>0.02 0.07 0.12</td>
</tr>
<tr>
<td>( C_m )</td>
<td>0.06</td>
<td>10.7 40.5 77.6</td>
<td>22.3 54.1 87.6</td>
<td>66.4 84.0 91.0</td>
</tr>
<tr>
<td>( C_{mm} )</td>
<td></td>
<td>9.7   39.8 76.4</td>
<td>7.7   37.3 71.0</td>
<td>13.4 72.4 95.0</td>
</tr>
<tr>
<td>( C_Q )</td>
<td></td>
<td>8.9   40.0 75.0</td>
<td>8.1   39.1 72.2</td>
<td>10.7 69.9 96.2</td>
</tr>
<tr>
<td>( C_{Q+} )</td>
<td></td>
<td>10.3  39.7 74.8</td>
<td>9.4   40.6 71.9</td>
<td>11.2 60.3 92.2</td>
</tr>
<tr>
<td>RS</td>
<td></td>
<td>10.4  43.0 77.8</td>
<td>9.9   42.2 74.5</td>
<td>13.2 62.6 93.5</td>
</tr>
<tr>
<td>ARS</td>
<td></td>
<td>8.7   37.7 75.3</td>
<td>7.7   36.0 69.6</td>
<td>11.6 59.2 92.1</td>
</tr>
<tr>
<td>( C_m )</td>
<td>0.12</td>
<td>7.9   38.3 69.4</td>
<td>18.7  47.5 84.1</td>
<td>53.1 74.3 88.0</td>
</tr>
<tr>
<td>( C_{mm} )</td>
<td></td>
<td>8.1   38.6 69.7</td>
<td>8.1   32.4 63.6</td>
<td>11.6 49.6 82.6</td>
</tr>
<tr>
<td>( C_Q )</td>
<td></td>
<td>7.4   33.7 65.7</td>
<td>7.2   32.6 65.2</td>
<td>10.7 52.9 85.9</td>
</tr>
<tr>
<td>( C_{Q+} )</td>
<td></td>
<td>9.4   36.6 66.3</td>
<td>9.4   35.3 66.1</td>
<td>8.8  45.9 80.3</td>
</tr>
<tr>
<td>RS</td>
<td></td>
<td>10.6  37.9 69.4</td>
<td>10.4  35.1 67.1</td>
<td>11.2 50.2 83.3</td>
</tr>
<tr>
<td>ARS</td>
<td></td>
<td>7.7   31.6 63.6</td>
<td>7.4   29.1 61.5</td>
<td>9.8  45.9 81.0</td>
</tr>
<tr>
<td>( C_m )</td>
<td>0.16</td>
<td>6.9   35.4 67.5</td>
<td>17.9  45.7 81.8</td>
<td>46.4 70.4 83.6</td>
</tr>
<tr>
<td>( C_{mm} )</td>
<td></td>
<td>6.9   35.8 67.9</td>
<td>6.9   30.9 61.6</td>
<td>10.1 44.7 75.6</td>
</tr>
<tr>
<td>( C_Q )</td>
<td></td>
<td>5.9   31.2 63.7</td>
<td>6.6   30.0 60.5</td>
<td>10.0 47.3 81.1</td>
</tr>
<tr>
<td>( C_{Q+} )</td>
<td></td>
<td>7.8   33.2 66.3</td>
<td>9.2   33.1 63.0</td>
<td>8.9  42.8 73.7</td>
</tr>
<tr>
<td>RS</td>
<td></td>
<td>7.4   33.7 64.8</td>
<td>9.3   33.1 62.4</td>
<td>10.5 46.9 78.3</td>
</tr>
<tr>
<td>ARS</td>
<td></td>
<td>5.4   28.6 62.0</td>
<td>6.6   26.4 57.1</td>
<td>9.0  42.4 74.7</td>
</tr>
</tbody>
</table>
Table 3.5: Empirical level (%) of the statistics $C_Q$, $C_{EQ}$, RS and ARS under departure of homogeneity of dispersion $\alpha = 0.05$: T=2 groups: based on 1000 replications.

(I) $C_Q$

<table>
<thead>
<tr>
<th>$\pi$</th>
<th>(a)BB distribution</th>
<th>(b)PNB distribution</th>
<th>(c)LNB distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\theta_1, \theta_2)$</td>
<td>0.06 0.10 0.20 0.30</td>
<td>0.06 0.10 0.20 0.30</td>
<td>0.06 0.10 0.20 0.30</td>
</tr>
<tr>
<td>(0.06,0.08)</td>
<td>5.8 5.0 5.1 3.5</td>
<td>4.3 5.0 4.5 4.4</td>
<td>4.5 3.9 4.0 3.7</td>
</tr>
<tr>
<td>(0.06,0.10)</td>
<td>4.9 3.9 4.9 4.2</td>
<td>4.8 4.6 4.1 4.1</td>
<td>5.2 4.2 4.1 4.9</td>
</tr>
<tr>
<td>(0.06,0.12)</td>
<td>5.2 4.8 4.3 5.6</td>
<td>4.9 4.6 4.7 4.1</td>
<td>5.2 4.9 4.8 4.7</td>
</tr>
<tr>
<td>(0.06,0.20)</td>
<td>4.9 4.8 5.7 5.1</td>
<td>4.5 5.0 5.8 4.7</td>
<td>5.3 4.7 6.8 6.0</td>
</tr>
<tr>
<td>(0.06,0.40)</td>
<td>6.2 8.4 6.1 5.2</td>
<td>6.0 6.8 5.8 6.3</td>
<td>6.1 5.6 6.2 6.2</td>
</tr>
<tr>
<td>(0.06,0.80)</td>
<td>6.8 7.4 9.4 7.8</td>
<td>7.7 8.2 7.2 6.2</td>
<td>7.6 7.1 7.0 6.9</td>
</tr>
</tbody>
</table>

(II) $C_{EQ}$

<table>
<thead>
<tr>
<th>$\pi$</th>
<th>(a)BB distribution</th>
<th>(b)PNB distribution</th>
<th>(c)LNB distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\theta_1, \theta_2)$</td>
<td>0.06 0.10 0.20 0.30</td>
<td>0.06 0.10 0.20 0.30</td>
<td>0.06 0.10 0.20 0.30</td>
</tr>
<tr>
<td>(0.06,0.08)</td>
<td>5.1 6.0 5.6 4.3</td>
<td>6.1 5.5 4.8 4.7</td>
<td>5.4 5.6 6.4 4.7</td>
</tr>
<tr>
<td>(0.06,0.10)</td>
<td>5.3 4.4 5.0 4.4</td>
<td>6.7 5.6 4.7 4.3</td>
<td>5.1 5.5 6.5 6.1</td>
</tr>
<tr>
<td>(0.06,0.12)</td>
<td>6.7 5.5 5.4 6.0</td>
<td>6.8 5.2 5.2 4.4</td>
<td>4.9 5.6 5.3 5.4</td>
</tr>
<tr>
<td>(0.06,0.20)</td>
<td>6.8 5.1 5.6 5.3</td>
<td>6.6 5.8 5.1 4.3</td>
<td>5.2 3.9 4.5 4.8</td>
</tr>
<tr>
<td>(0.06,0.40)</td>
<td>8.6 8.9 6.0 6.2</td>
<td>7.9 7.2 5.5 5.2</td>
<td>7.4 4.6 4.9 4.2</td>
</tr>
<tr>
<td>(0.06,0.80)</td>
<td>10.8 8.0 8.1 6.6</td>
<td>11.4 9.2 5.8 5.1</td>
<td>9.0 7.3 4.9 5.7</td>
</tr>
</tbody>
</table>

104
### (III) RS

<table>
<thead>
<tr>
<th>$\pi$</th>
<th>(a)BB distribution</th>
<th>(b)PNB distribution</th>
<th>(c)LNB distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\theta_1, \theta_2)$</td>
<td>0.06 0.10 0.20 0.30</td>
<td>0.06 0.10 0.20 0.30</td>
<td>0.06 0.10 0.20 0.30</td>
</tr>
<tr>
<td>(0.06,0.08)</td>
<td>7.9 7.2 5.8 5.5</td>
<td>7.1 7.5 5.9 5.7</td>
<td>6.0 6.5 5.5 5.1</td>
</tr>
<tr>
<td>(0.06,0.10)</td>
<td>7.6 5.6 5.2 5.5</td>
<td>7.8 7.2 5.7 5.4</td>
<td>7.2 6.2 5.2 4.5</td>
</tr>
<tr>
<td>(0.06,0.12)</td>
<td>9.8 7.3 5.6 6.9</td>
<td>8.6 7.0 6.5 5.3</td>
<td>6.6 5.7 5.2 5.1</td>
</tr>
<tr>
<td>(0.06,0.20)</td>
<td>9.2 7.0 6.4 5.9</td>
<td>8.4 6.9 6.0 5.3</td>
<td>6.5 5.4 5.7 6.0</td>
</tr>
<tr>
<td>(0.06,0.40)</td>
<td>12.1 9.6 6.6 6.0</td>
<td>10.8 8.0 6.4 6.1</td>
<td>7.7 5.8 6.7 4.8</td>
</tr>
<tr>
<td>(0.06,0.80)</td>
<td>13.9 9.3 7.4 6.4</td>
<td>15.1 9.4 6.1 5.3</td>
<td>8.5 7.0 6.0 5.3</td>
</tr>
</tbody>
</table>

### (IV) ARS

<table>
<thead>
<tr>
<th>$\pi$</th>
<th>(a)BB distribution</th>
<th>(b)PNB distribution</th>
<th>(c)LNB distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\theta_1, \theta_2)$</td>
<td>0.06 0.10 0.20 0.30</td>
<td>0.06 0.10 0.20 0.30</td>
<td>0.06 0.10 0.20 0.30</td>
</tr>
<tr>
<td>(0.06,0.08)</td>
<td>6.4 6.5 6.5 5.1</td>
<td>7.2 6.8 6.0 5.5</td>
<td>6.4 6.4 5.8 5.4</td>
</tr>
<tr>
<td>(0.06,0.10)</td>
<td>6.3 4.9 5.6 5.9</td>
<td>7.3 6.5 5.6 5.5</td>
<td>6.9 6.6 5.1 5.0</td>
</tr>
<tr>
<td>(0.06,0.12)</td>
<td>8.1 6.8 6.0 7.0</td>
<td>7.3 6.0 6.3 5.3</td>
<td>6.2 6.3 5.7 5.2</td>
</tr>
<tr>
<td>(0.06,0.20)</td>
<td>7.8 6.3 6.5 6.2</td>
<td>7.1 6.4 6.4 5.4</td>
<td>7.3 6.3 6.2 5.7</td>
</tr>
<tr>
<td>(0.06,0.40)</td>
<td>10.9 9.5 6.5 6.2</td>
<td>9.2 8.4 6.4 6.5</td>
<td>9.6 6.8 6.4 5.4</td>
</tr>
<tr>
<td>(0.06,0.80)</td>
<td>13.4 9.8 8.4 7.3</td>
<td>13.7 10.7 7.2 5.9</td>
<td>10.8 9.6 7.1 6.0</td>
</tr>
</tbody>
</table>
Table 3.6: The Toxicological data from Paul (1982). (i) Number of live foetuses affected by treatment. (ii) Total number of live foetuses.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>(i)</th>
<th>(ii)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control. C</td>
<td>1 1 4 0 0 0 0 0 1 0 2 0 5 2 1 2 0 0</td>
<td>1 0 0 0 0 3 2 4 0</td>
</tr>
<tr>
<td>Low dose. L</td>
<td>0 1 1 0 2 0 1 0 1 0 0 3 0 0 1 5 0 0 3</td>
<td>5 1 1 7 9 1 2 8 6 7 6 4 6 9 6 7 5 9 1 6 9</td>
</tr>
<tr>
<td>Medium dose. M</td>
<td>2 3 2 1 2 3 0 4 0 0 4 0 0 6 6 5 4 1 0 3 6</td>
<td>4 4 9 8 9 7 8 9 6 4 6 7 3 1 3 6 8 1 1 7 6 1 0 6</td>
</tr>
<tr>
<td>High dose. H</td>
<td>1 0 1 0 1 0 1 2 0 4 1 1 4 2 3 1</td>
<td>9 1 0 7 5 4 6 3 8 5 4 4 5 3 8 6 8 6</td>
</tr>
</tbody>
</table>
Table 3.7: The quasi-likelihood estimates of the parameters $\pi$ and $\theta$ and the quasi-likelihood score statistic $C_Q$, the likelihood ratio statistic $G^2$ and the Rao Scott statistics RS and ARS for the Toxicological data in Table 3.5.

<table>
<thead>
<tr>
<th></th>
<th>Treatment Combinations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CLMH</td>
</tr>
<tr>
<td>$\hat{\pi}$</td>
<td>.208</td>
</tr>
<tr>
<td>$\hat{\theta}$</td>
<td>.283</td>
</tr>
<tr>
<td>$G^2$</td>
<td>11.30</td>
</tr>
<tr>
<td>P value</td>
<td>(.011)</td>
</tr>
<tr>
<td>$C_Q$</td>
<td>10.52</td>
</tr>
<tr>
<td>P value</td>
<td>(.015)</td>
</tr>
<tr>
<td>RS</td>
<td>12.60</td>
</tr>
<tr>
<td>P value</td>
<td>(.006)</td>
</tr>
<tr>
<td>ARS</td>
<td>13.92</td>
</tr>
<tr>
<td>P value</td>
<td>(.003)</td>
</tr>
<tr>
<td>d.f.</td>
<td>3</td>
</tr>
</tbody>
</table>
CHAPTER IV

TESTING THE ASSUMPTION OF A COMMON DISPERSION PARAMETER IN THE ANALYSIS OF PROPORTIONS

4.1 Introduction

When testing the homogeneity of proportions as discussed in Chapter three we assumed that the dispersion parameters of the T groups are equal. However, this assumption should be checked before performing a test of the homogeneity of proportions. As in the normal theory one-way analysis of variance situation, we consider the testing of homogeneity of the dispersion parameters as a preliminary test. If this hypothesis is not rejected we then test the hypothesis of the homogeneity of proportions. The commonly used parametric procedure for testing the assumption is the likelihood ratio test based on the beta-binomial distribution (Aeschbacher et al., 1977; Williams, 1975 and Pack, 1986(a)). The potential drawback of the likelihood ratio test is that it requires estimates of the parameters under both the null and the alternative hypothesis. One of the drawbacks of the $C(\alpha)$ test based on a parametric model is that the test is not robust under model misspecification.

As we have seen in Chapter three, the $C(\alpha)$ statistics based on semi-parametric models were robust for departure from the data distribution and dispersion heterogeneity. Therefore, in this chapter our main interest is to develop $C(\alpha)$ test procedures based on semi-parametric models for testing the homogeneity of dispersion. We develop a $C(\alpha)$ test based on the quasi-likelihood/method of moments proposed by Srivastava and Wu (1993). We also develop a $C(\alpha)$ test based on the Gaussian likelihood. The reason for selecting these two procedures, as has been
discussed in chapter two, is that for the estimation of dispersion parameter these
two procedures perform best.

We are interested in testing the homogeneity of the dispersion parameters i.e.
testing

$$H_0 : \theta_1 = \ldots = \theta_T$$

against

$$H_A : \text{not all } \theta_i \text{'s are the same.}$$

with the parameters $\pi_1, \ldots, \pi_T$ remaining unspecified. This hypothesis can also be written in terms of $\phi$:

$$H_0 : \phi_1 = \ldots = \phi_T$$

against

$$H_A : \text{not all } \phi_i \text{'s are the same.}$$

4.2 The C($\alpha$) statistics based on the quasi-likelihood and method of mo-
ments proposed by Srivastava and Wu (1993)

We reparametrize $\phi$ under $H_A$, by $\phi_i = \phi + \tau_i$ with $\tau_T = 0$. Then testing $H_0$ is equivalent to testing

$$H_0 : \tau_i = 0 \quad \text{against}$$

$$H_A : \tau_i \neq 0 \quad i = 1, \ldots, T - 1$$

with $\pi_1, \ldots, \pi_T$ and $\phi$ treated as nuisance parameters. Define

$$\tau = (\tau_1, \ldots, \tau_{T-1}) \quad \text{and} \quad \lambda = (\lambda_1, \ldots, \lambda_T, \lambda_{T+1}) = (\pi_1, \ldots, \pi_T, \phi).$$

Thus $\lambda_k = \pi_k$ for $k = 1, \ldots, T$ and $\lambda_{T+1} = \phi$. For given $\pi_k$, $k = 1, \ldots, T$ and $\phi$, the unbiased estimating function for the parameter $\tau_1, \ldots, \tau_{T-1}$ is obtained by the method of moments proposed by Srivastava and Wu (1993) as given below

$$g_{1i} = \sum_{j=1}^{m_i} \left\{ \frac{n_{ij}(z_{ij} - \pi_i)^2}{\pi_i(1 - \pi_i)} - 1 + (n_{ij} - 1)(\phi + \tau_i) \right\}. $$
For given \( \tau_1 \ldots \tau_{T-1}, \phi \), the unbiased estimating function for \( \pi_k, k = 1 \ldots , T \), is given by the quasi-likelihood score function

\[
g_{2k} = \sum_{j=1}^{m_k} \frac{n_{kj}(z_{kj} - \pi_k)}{\pi_k(1 - \pi_k)(1 + (n_{kj} - 1)(\phi + \tau_k))}, \quad k = 1, \ldots , T.
\]

Similarly, an unbiased estimating function for \( \phi \), given \( \tau_1 \ldots \tau_{T-1}, \pi_1 \ldots \pi_T \) can be obtained by using the method of moments proposed by Srivastava and Wu (1993), as given below

\[
g_{2, T+1} = \sum_{i=1}^T \sum_{j=1}^{m_i} \left\{ \frac{n_{ij}(z_{ij} - \pi_i)^2}{\pi_i(1 - \pi_i)} - \{1 + (n_{ij} - 1)(\phi + \tau_i)\} \right\}.
\]

Let

\[
\psi_i = g_{1i} \bigg|_{\tau=0}, \quad i = 1, \ldots , T - 1
\]

and

\[
\gamma_k = g_{2k} \bigg|_{\tau=0}, \quad k = 1, \ldots , T + 1.
\]

Further define

\[
\psi = (\psi_1, \ldots , \psi_{T-1})'.
\]

Then, if \( W \) is the asymptotic variance of \( \psi \), which will be derived later, the score statistic for testing the homogeneity of dispersion parameters is

\[
C_\phi = \psi'W^{-1}\psi
\]

which will have asymptotically a chi-square distribution with \((T-1)\) degrees of freedom.

Note that \( \psi \) and \( W \) involve \( \lambda \). According to the procedure discussed earlier we replace \( \lambda = (\pi_1, \ldots , \pi_T, \phi) \) in \( C_\phi \) by \( \sqrt{m_i} \) consistent estimates which are obtained by solving

\[
\sum_{j=1}^{m_k} \frac{n_{kj}(z_{kj} - \pi_k)}{\pi_k(1 - \pi_k)(1 + (n_{kj} - 1)\phi)} = 0, \quad k = 1, \ldots , T
\]

110
and
\[
\sum_{i=1}^{T} \sum_{j=1}^{m_i} \left\{ \frac{n_{ij}(z_{ij} - \mu_i)^2}{\mu_i(1 - \mu_i)} \right\} \left(1 + (n_{ij} - 1)\phi \right) = 0
\]
simultaneously. Denote the estimate of \( \lambda \) obtained above by \( \lambda^* \).

Now we obtain \( W \) following Inagaki (1973).

Let
\[
\psi(\lambda^*) = \left( \psi_1(\lambda^*), \ldots, \psi_{T-1}(\lambda^*) \right)^t
\]
and
\[
\gamma(\lambda^*) = \left( \gamma_1(\lambda^*), \ldots, \gamma_{T+1}(\lambda^*) \right)^t.
\]

We expand \( \psi \) and \( \gamma \) with respect to \( \lambda \) and obtain
\[
\psi(\lambda^*) = \psi(\lambda^0) + \frac{\partial \psi}{\partial \lambda} (\lambda^* - \lambda^0) + O_p(1) \quad (4.2.1)
\]
\[
\gamma(\lambda^*) = \gamma(\lambda^0) + \frac{\partial \gamma}{\partial \lambda} (\lambda^* - \lambda^0) + O_p(1). \quad (4.2.2)
\]
where \( \lambda^0 \) is the true value of \( \lambda \). From equation \( 4.2.2 \) we get
\[
(\lambda^* - \lambda^0) \approx \gamma(\lambda^0) \left( -\frac{\partial \gamma}{\partial \lambda} \right)^{-1} \quad (4.2.3)
\]
Substituting equation \( 4.2.3 \) into equation \( 4.2.1 \) we get
\[
\psi(\lambda^*) = \psi(\lambda^0) - \left( -\frac{\partial \psi}{\partial \lambda} \right) \left( -\frac{\partial \gamma}{\partial \lambda} \right)^{-1} \gamma(\lambda^0) + O_p(1).
\]
Then replacing \( -\frac{\partial \psi}{\partial \lambda} \) and \( -\frac{\partial \gamma}{\partial \lambda} \) by their asymptotically equivalent versions namely their expectations, \( C = E\left(-\frac{\partial \psi}{\partial \lambda}\right) \) and \( A = E\left(-\frac{\partial \gamma}{\partial \lambda}\right) \), we obtain
\[
\psi(\lambda^*) = \psi(\lambda^0) - E\left(-\frac{\partial \psi}{\partial \lambda}\right) E\left(-\frac{\partial \gamma}{\partial \lambda}\right)^{-1} \gamma(\lambda^0) = \psi(\lambda^0) - CA^{-1} \gamma(\lambda^0).
\]
Then, the asymptotic variance of \( \psi(\lambda^*) \) is
\[
\text{Var} \left( \psi(\lambda^*) \right) = \text{Var} \left( \psi(\lambda^0) - CA^{-1} \gamma(\lambda^0) \right).
\]
Now, the actual variance can be estimated by replacing $\lambda^0$ by $\hat{\lambda}$ in

$$\text{Var}(\psi(\lambda^0) - CA^{-1}\gamma(\lambda^0)) = E(\psi(\lambda^0) - CA^{-1}\gamma(\lambda^0))(\psi(\lambda^0) - CA^{-1}\gamma(\lambda^0))^t$$

$$= E(\psi(\lambda^0)\psi^t(\lambda^0) - \psi(\lambda^0)\gamma^t(\lambda^0)CA^{-1}C^t$$

$$- CA^{-1}\gamma(\lambda^0)\psi^t(\lambda^0) + CA^{-1}\gamma(\lambda^0)\gamma^t(\lambda^0)CA^{-1}C^t)$$

$$= E(\psi(\lambda^0)\psi^t(\lambda^0)) - E(\psi(\lambda^0)\gamma^t(\lambda^0))CA^{-1}C^t$$

$$- CA^{-1}E(\gamma(\lambda^0)\psi^t(\lambda^0)) + CA^{-1}E(\gamma(\lambda^0)\gamma^t(\lambda^0))CA^{-1}C^t.$$

Denote $E(\psi(\lambda^0)\psi^t(\lambda^0))$ by J, $E(\psi(\lambda^0)\gamma^t(\lambda^0))$ by H and $E(\gamma(\lambda^0)\gamma^t(\lambda^0))$ by G.

Then the score statistic using the estimate of the actual variance of $\psi(\hat{\lambda})$ is

$$C_{oA} = \psi^t(J - HA^{-1}C^t - CA^{-1}H - CA^{-1}GA^{-1}C^t)^{-1}\psi$$

where the matrices J, H, A, C and G are of the order $(T-1)\times(T-1)$, $(T-1)\times(T+1)$, $(T+1)\times(T+1)$, $(T-1)\times(T+1)$ and $(T+1)\times(T+1)$ respectively.

Breslow (1990(a), 1990(b)) suggests using the empirical variance of $\psi(\hat{\lambda})$ which is

$$\text{Empirical Var}(\psi(\hat{\lambda})) = \left(\psi(\lambda^0) - CA^{-1}\gamma(\lambda^0)\right)\left(\psi(\lambda^0) - CA^{-1}\gamma(\lambda^0)\right)^t$$

$$= \psi(\lambda^0)\psi^t(\lambda^0) - \psi(\lambda^0)\gamma^t(\lambda^0)CA^{-1}C$$

$$- CA^{-1}\gamma(\lambda^0)\psi^t(\lambda^0) + CA^{-1}\gamma(\lambda^0)\gamma^t(\lambda^0)CA^{-1}C^t.$$

Denote $\left(\psi(\lambda^0)\psi^t(\lambda^0)\right)$ by K, $\left(\psi(\lambda^0)\gamma^t(\lambda^0)\right)$ by M and $\left(\gamma(\lambda^0)\gamma^t(\lambda^0)\right)$ by N. Then the score statistic using the empirical variance of $\psi(\hat{\lambda})$ is

$$C_{oE} = \psi^t(K - MA^{-1}C^t - CA^{-1}M^t + CA^{-1}NA^{-1}C^t)^{-1}\psi.$$

Now we derive the elements of the matrices J, H, A, C and G in $C_{oA}$.

$$J_{it} = E(\psi_i\psi_t)$$

$$= \sum_{j=1}^{m_i} \left\{ \frac{n_{ij}^2E(z_{ij} - \pi_i)^4}{[\pi_i(1 - \pi_i)]^2} - \{1 + (n_{ij} - 1)\phi\} \right\} = d_i, \quad i = t = 1, \ldots, T - 1$$

$$= 0, \quad \text{for } i \neq t = 1, \ldots, T - 1.$$
\[ H_{ik} = E\left(\psi_i\gamma_k\right) \]

\[ = \sum_{j=1}^{m_i} \left\{ \frac{(1-2\pi_i)(1+(2n_{ij}-1)\phi)}{\pi_i(1-\pi_i)(1+\phi)} \right\} = f_i, \quad i = k = 1, \ldots, T-1 \]

\[ = 0, \quad i \neq k = 1, \ldots, T-1 \]

\[ = 0, \quad i = 1, \ldots, T-1; k = T \]

\[ = \sum_{j=1}^{m_i} \left\{ \frac{n_{ij}E(z_{ij}-\pi_i)^4}{\pi_i(1-\pi_i)} - \frac{1+(n_{ij}-1)\phi}{\pi_i(1-\pi_i)} \right\} = d_i, \quad i = 1, \ldots, T-1; k = T+1 \]

\[ A_{ks} = E\left(\frac{-\partial \tau_k}{\partial \lambda_s}\right), \]

\[ = \sum_{j=1}^{m_i} \frac{n_{ij}}{\pi_i(1-\pi_i)(1+(n_{ij}-1)\phi)} = g_i, \quad k = s = 1, \ldots, T \]

\[ = 0, \quad k \neq s = 1, \ldots, T \]

\[ = \sum_{j=1}^{m_i} \left\{ \frac{(1-2\pi_i)(1+(n_{ij}-1)\phi)}{\pi_i(1-\pi_i)} \right\} = e_i, \quad i = s = 1, \ldots, T; k = T+1 \]

\[ = \sum_{i=1}^{T} \sum_{j=1}^{m_i} (n_{ij} - 1) = \sum_{i=1}^{T} c_i = c, \quad k = s = T+1 \]

\[ C_{is} = E\left(\frac{-\partial \psi_i}{\partial \lambda_s}\right) \]

\[ = \sum_{j=1}^{m_i} \left\{ \frac{(1-2\pi_i)(1+(n_{ij}-1)\phi)}{\pi_i(1-\pi_i)} \right\} = e_i, \quad i = s = 1, \ldots, T-1 \]

\[ = 0, \quad i \neq s = 1, \ldots, T-1 \]

\[ = 0, \quad i = 1, \ldots, T-1; s = T \]

\[ = \sum_{j=1}^{m_i} (n_{ij} - 1) = c_i, \quad i = 1, \ldots, T-1; s = T+1 \]
\[ G_{ks} = E(\gamma_s \gamma_k) \]
\[ = \sum_{j=1}^{m_k} \frac{n_{ij}}{\pi_i(1 - \pi_i)(1 + (n_{ij} - 1)\phi)} = g_i. \quad k = s = 1, \ldots, T \]
\[ i = 1, \ldots, T \]
\[ = 0. \quad k \neq s = 1, \ldots, T \]
\[ = \sum_{j=1}^{m_i} \left\{ \frac{(1 - 2\pi_i)(1 + (2n_{ij} - 1)\phi)}{\pi_i(1 - \pi_i)(1 + \phi)} \right\} = f_i. \quad k = 1, \ldots, T; s = T + 1 \]
\[ i = 1, \ldots, T \]
\[ = \sum_{j=1}^{m_i} \left\{ \frac{n_{ij}^2 E(z_{ij} - \pi_i)^4}{[\pi_i(1 - \pi_i)]^2} - [1 + (n_{ij} - 1)\phi]^2 \right\} = \sum_{i=1}^{T} d_i = d. \quad s = k = T + 1 \]

Thus, the matrices J, H, A, C and G can be written in terms of \( d_i, f_i, g_i, e_i \) and \( c_i \) as

\[
J = \begin{pmatrix}
    d_1 & 0 & \cdots & 0 \\
    0 & d_2 & \cdots & 0 \\
    \vdots & \ddots & \ddots & \vdots \\
    0 & 0 & \cdots & d_{T-1}
\end{pmatrix},
\]

\[
H = \begin{pmatrix}
    f_1 & 0 & \cdots & 0 & 0 & d_1 \\
    0 & f_2 & \cdots & 0 & 0 & d_2 \\
    \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\
    0 & 0 & \cdots & f_{T-1} & 0 & d_{T-1}
\end{pmatrix},
\]

\[
A = \begin{pmatrix}
    g_1 & 0 & \cdots & 0 & 0 \\
    0 & g_2 & \cdots & 0 & 0 \\
    \vdots & \ddots & \ddots & \ddots & \vdots \\
    0 & 0 & \cdots & g_T & 0 \\
    e_1 & e_2 & \cdots & e_T & c
\end{pmatrix},
\]

114
\[ C = \begin{pmatrix}
c_1 & 0 & \cdots & 0 & 0 & c_1 \\
0 & c_2 & \cdots & 0 & 0 & c_2 \\
\vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & e_{T-1} & 0 & e_{T-1}
\end{pmatrix}. \]

\[ G = \begin{pmatrix}
g_1 & 0 & \cdots & 0 & f_1 \\
0 & g_2 & \cdots & 0 & f_2 \\
\vdots & \ddots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & g_T & f_T \\
f_1 & f_2 & \cdots & f_T & d
\end{pmatrix}. \]

Note that the quantities \( d_i \) and \( f_i \) in \( \mathbf{J}, \mathbf{H} \) and \( \mathbf{G} \) involve third and fourth moments, which is beyond the scope of the quasi-likelihood procedure. So we replace them by the third and fourth moments of the beta-binomial distribution. As such, the score statistic developed here uses the model based variance-covariance matrix.

For the statistic \( C_{\phi E} \) the elements of the matrices \( \mathbf{C} \) and \( \mathbf{A} \) are given above and the elements of the matrices \( \mathbf{K}, \mathbf{M} \) and \( \mathbf{N} \) are

\[
K_{it} = (\psi_i \psi_t)
\]

\[
= \sum_{j=1}^{m_i} \left\{ \frac{n_{ij}(z_{ij} - \pi_i)^4}{\pi_i(1 - \pi_i)^2} - \frac{2n_{ij}(z_{ij} - \pi_i)^2\{1 + (n_{ij} - 1)\phi\}}{\pi_i(1 - \pi_i)} \right\} + \{1 + (n_{ij} - 1)\phi\}^2 \]

\[
= d_{ii}, \quad i = 1, \ldots, T - 1; \quad i = t
\]

\[
K_{it} = \sum_{j=1}^{m_i} \left\{ \frac{n_{ij}(z_{ij} - \pi_i)^2}{\pi_i(1 - \pi_i)} - \{1 + (n_{ij} - 1)\phi\} \right\} \sum_{j=1}^{m_t} \left\{ \frac{n_{tj}(z_{tj} - \pi_t)^2}{\pi_t(1 - \pi_t)} - \{1 + (n_{tj} - 1)\phi\} \right\}
\]

\[
= d_{it}, \quad i = 1, \ldots, T - 1; \quad i \neq t
\]
\[ M_{ik} = (\psi_i \gamma_k) \]
\[ = \sum_{j=1}^{m_i} \left\{ \frac{n_{ij}^2(z_{ij} - \pi_i)^3}{[\pi_i(1 - \pi_i)]^2(1 + (n_{ij} - 1)\phi)} - \frac{n_{ij}(z_{ij} - \pi_i)}{\pi_i(1 - \pi_i)} \right\} = f_{ii}, \quad i = 1, \ldots, T - 1 \]
\[ i = k. \]
\[ M_{ik} = 0, \quad i = 1, \ldots, T - 1; \quad i \neq k. \]
\[ M_{ik} = \sum_{j=1}^{m_i} \left\{ \frac{n_{ij}^2(z_{ij} - \pi_i)^4}{[\pi_i(1 - \pi_i)]^2} - \frac{2n_{ij}(z_{ij} - \pi_i)^2(1 + (n_{ij} - 1)\phi)}{\pi_i(1 - \pi_i)} + \{1 + (n_{ij} - 1)\phi\}^2 \right\} \]
\[ = d_i, \quad i = 1, \ldots, T - 1; \quad k = T + 1. \]
\[ N_{ks} = (\gamma_k \gamma_s) \]
\[ = \sum_{j=1}^{m_s} \left\{ \frac{n_{kj}^2(z_{kj} - \pi_k)^2}{[\pi_k(1 - \pi_k)(1 + (n_{kj} - 1)\phi)]^2} \right\} = g_{kk}, \quad k = 1, \ldots, T; \quad k = s. \]
\[ N_{ks} = 0, \quad k = 1, \ldots, T; \quad k \neq s. \]
\[ N_{ks} = \sum_{j=1}^{m_i} \left\{ \frac{n_{ij}^2(z_{ij} - \pi_i)^3}{[\pi_i(1 - \pi_i)]^2(1 + (n_{ij} - 1)\phi)} - \frac{n_{ij}(z_{ij} - \pi_i)}{\pi_i(1 - \pi_i)} \right\} = f_{i}, \quad k = 1, \ldots, T \]
\[ s = T + 1. \quad i = 1, \ldots, T. \]
\[ N_{ks} = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \left\{ \frac{n_{ij}^2(z_{ij} - \pi_i)^4}{[\pi_i(1 - \pi_i)]^2} - \frac{2n_{ij}(z_{ij} - \pi_i)^2(1 + (n_{ij} - 1)\phi)}{\pi_i(1 - \pi_i)} + \{1 + (n_{ij} - 1)\phi\}^2 \right\} \]
\[ = \sum_{i=1}^{T} d_i = d, \quad s = k = T + 1. \]

Thus the matrices \(K, M\) and \(N\) can be written in terms of \(d_i, f_i\) and \(g_i\) as

\[
K = \begin{pmatrix}
  d_{11} & d_{12} & \cdots & d_{1T-1} \\
  d_{21} & d_{22} & \cdots & d_{2T-1} \\
  \vdots & \vdots & \ddots & \vdots \\
  d_{T-11} & d_{T-12} & \cdots & d_{T-1T-1}
\end{pmatrix},
\]

\[
M = \begin{pmatrix}
  f_{11} & 0 & \cdots & 0 & 0 & d_1 \\
  0 & f_{22} & \cdots & 0 & 0 & d_2 \\
  \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
  0 & 0 & \cdots & f_{T-1T-1} & 0 & d_{T-1}
\end{pmatrix},
\]

116
The statistic $C_{\phi E}$ obviously does not require any moments higher than the second moment.

4.3 The $C(\alpha)$ statistics based on the Gaussian likelihood

We reparametrize $\phi_i$ under $H_A$, by $\phi_i = \phi + \tau_i$ with $\tau_T = 0$.

Here $\pi_1, \ldots, \pi_T$ and $\phi$ are treated as nuisance parameters. Define $\tau = (\tau_1, \ldots, \tau_{T-1})$
and $\lambda = (\lambda_1, \ldots, \lambda_T, \lambda_{T+1}) = (\pi_1, \ldots, \pi_T, \phi)$.

The log-likelihood for the normal distribution, in terms of the parameters $\tau_i$, 
$(i = 1, \ldots, T - 1), \pi_k, (k = 1, \ldots, T)$ and $\phi$ can be written as

$$l = -\frac{1}{2} \sum_{i=1}^{T} \sum_{j=1}^{m_i} \left\{ n_{ij} \ln \{ \pi_i (1 - \pi_i) \} + (n_{ij} - 1) \ln (1 - \phi - \tau_i) 
+ \ln \{ 1 + (n_{ij} - 1)(\phi + \tau_i) \} + \{ \pi_i (1 - \pi_i)(1 - \phi - \tau_i) \}^{-1} \left[ n_{ij} \pi_i (1 - \pi_i) 
(1 - 2\pi_i) n_{ij} (z_{ij} - \pi_i) - (\phi + \tau_i) [1 + (n_{ij} - 1)(\phi + \tau_i)]^{-1} n_{ij}^2 (z_{ij} - \pi_i)^2 \right] \right\}.$$ 

Let

$$\psi_i = g_{1i} = \left( \frac{\partial l}{\partial \tau_i} \right)_{\tau = 0}, \quad i = 1, \ldots, T - 1$$

and

$$\gamma_k = \left( \frac{\partial l}{\partial \lambda_k} \right)_{\tau = 0}, \quad k = 1, \ldots, T + 1.$$ 

Further define

$$\psi = (\psi_1, \ldots, \psi_{T-1})^T.$$
Note that these score functions are treated as estimating equations because the data does not come from a normal distribution. Therefore the asymptotic variance of the score $\psi$ is obtained by using the Inagaki results as given in section 4.2.

Let $W$ be the asymptotic variance of $\psi$ which was derived in section 4.2. The score statistic for testing the homogeneity of the dispersion parameters is

$$C_{GL} = \psi' W^{-1} \psi$$

which has asymptotically a chi-square distribution with $(T-1)$ degrees of freedom. Note that $\psi$ and $W$ involve $\lambda$. Following the same procedure as in the preceding section we replace $\lambda = (\pi_1, \ldots, \pi_T, \phi)$ in $C_{GL}$ by $\sqrt{m_i}$ consistent estimates which are obtained by solving

$$\frac{\partial l}{\partial \pi_i} = \sum_{j=1}^{m_i} \left\{ \left[ \frac{1 - \phi}{1 + (n_{ij} - 1) \phi} + \frac{(1 - 2 \pi_i)^2}{2 \pi_i (1 - \pi_i)} \right] \frac{n_{ij} (z_{ij} - \pi_i)}{\pi_i (1 - \pi_i) (1 - \phi)} 

- \left[ \frac{(1 - 2 \pi_i) n_{ij} \phi}{2 \pi_i (1 - \pi_i) (1 - \phi)} \right] \frac{n_{ij} (z_{ij} - \pi_i)^2}{\pi_i (1 - \pi_i) (1 + (n_{ij} - 1) \phi)} - 1 \right\} = 0$$

and

$$\frac{\partial l}{\partial \phi} = \sum_{t=1}^{T} \sum_{j=1}^{m_i} \left\{ \frac{n_{ij}}{2(1 - \phi)^2} \left( \frac{1 + (n_{ij} - 1) \phi}{1 + (n_{ij} - 1) \phi} \right) \frac{n_{ij} (z_{ij} - \pi_i)^2}{\pi_i (1 - \pi_i) (1 + (n_{ij} - 1) \phi)} - 1 \right\} 

- \frac{(1 - 2 \pi_i) n_{ij}}{2 \pi_i (1 - \pi_i) (1 - \phi)^2} (z_{ij} - \pi_i) = 0$$

simultaneously. Denote the estimate of $\lambda$ obtained above by $\hat{\lambda}$. Then the score statistic using the estimate of the actual variance of $\psi(\lambda)$ is

$$W = J - HA^{-1}C' - CA^{-1}H' + CA^{-1}GA^{-1}C'$$

where the matrices $J$, $H$, $A$, $C$ and $G$ are of the order $(T-1) \times (T-1)$, $(T-1) \times (T+1)$, $(T+1) \times (T+1)$, $(T-1) \times (T+1)$, and $(T+1) \times (T+1)$ respectively.
Now we derive the elements of the matrices $J$, $H$, $A$, and $G$ in $C_{GL}$.

$$J_{it} = E(\psi_i \psi_t)$$

$$\sum_{j=1}^{m_i} \left\{ \begin{align*}
\frac{n_{ij}^2}{4(1 - \phi)^4} \left[ \frac{n_{ij}^2 E(z_{ij} - \pi_i)^4}{[\pi_i(1 - \pi_i)(1 + (n_{ij} - 1)\phi)]^2} - 1 \right] \times D^2
\end{align*} \right. - \frac{n_{ij}(1 - 2\pi_i)^2(1 + (2n_{ij} - 1)\phi)}{2\pi_i(1 - \pi_i)(1 + \phi)(1 - \phi)^4} \times D
+ \frac{n_{ij}(1 - 2\pi_i)^2(1 + (n_{ij} - 1)\phi)}{4\pi_i(1 - \pi_i)(1 - \phi)^4} \right\} = d_i, \quad i = 1, \ldots, T - 1; \quad i = t.
$$

$$J_{it} = 0, \quad i = 1, \ldots, T - 1; \quad i \neq t = 1, \ldots, T - 1.$$

$$H_{ik} = E(\psi_i \gamma_k)$$

$$= \sum_{j=1}^{m_i} \left\{ \begin{align*}
\frac{n_{ij}(1 - 2\pi_i)^2(1 + (2n_{ij} - 1)\phi)}{2\pi_i(1 - \pi_i)(1 + \phi)(1 - \phi)^3} \times D \times B
\end{align*} \right. - \frac{n_{ij}(1 - 2\pi_i)^2}{4\pi_i(1 - \pi_i)(1 - \phi)^3} \times D - \frac{n_{ij}(1 - 2\pi_i)(1 + (n_{ij} - 1)\phi)}{2\pi_i(1 - \pi_i)(1 - \phi)^3}
+ \frac{n_{ij}\phi(1 - 2\pi_i)^2(1 + (2n_{ij} - 1)\phi)}{4\pi_i^2(1 - \pi_i)^2(1 + \phi)(1 - \phi)^3} \right\} = f_i, \quad i = 1, \ldots, T - 1; \quad i = k.
$$

$$H_{ik} = 0, \quad i = 1, \ldots, T - 1; \quad i \neq k.$$

$$H_{ik} = d_i, \quad i = 1, \ldots, T - 1; \quad k = T + 1.$$

$$A_{ks} = E \left( \frac{-\partial \gamma_k}{\partial \lambda_s} \right)$$

$$= \sum_{j=1}^{m_i} \left\{ \begin{align*}
\frac{n_{ij}}{\pi_i(1 - \pi_i)(1 + (n_{ij} - 1)\phi)} + \frac{n_{ij}(1 - 2\pi_i)^2}{2\pi_i(1 - \pi_i)^2} \right\} = g_i, \quad k = s = 1, \ldots, T.
$$

$$A_{ks} = 0, \quad s = 1, \ldots, T; \quad k \neq s; \quad i = 1, \ldots, T,$$

$$A_{ks} = \sum_{j=1}^{m_i} \left\{ \begin{align*}
\frac{-n_{ij}(n_{ij} - 1)(1 - 2\pi_i)\phi}{2\pi_i(1 - \pi_i)(1 - \phi)(1 + (n_{ij} - 1)\phi)} \right\} = e_i, \quad k = 1, \ldots, T; \quad s = T + 1.$$

$$A_{ks} = e_i, \quad i = s = 1, \ldots, T; \quad k = T + 1,$$

$$A_{ks} = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \left\{ \begin{align*}
\frac{n_{ij}(n_{ij} - 1)(1 + (n_{ij} - 1)\phi)^2}{2[(1 - \phi)(1 + (n_{ij} - 1)\phi)]^2} \right\} = \sum_{i=1}^{T} c_i = c, \quad k = s = T + 1.$$
\[ C_{is} = E \left( \frac{-\partial \psi_i}{\partial \lambda_s} \right) \]
\[ = \sum_{j=1}^{m_i} \left\{ \frac{-n_{ij}(n_{ij} - 1)(1 - 2\pi_i)\phi}{2\pi_i(1 - \pi_i)(1 - \phi)[1 + (n_{ij} - 1)\phi]} \right\} = c_i, \quad i = s = 1, \ldots, T - 1. \]
\[ C_{is} = 0, \quad i \neq s = 1, \ldots, T - 1. \]
\[ C_{is} = 0, \quad i = 1, \ldots, T - 1; \quad s = T. \]
\[ C_{is} = \sum_{j=1}^{m_i} \left\{ \frac{n_{ij}(n_{ij} - 1)[1 + (n_{ij} - 1)\phi^2]}{2[(1 - \phi)[1 + (n_{ij} - 1)\phi]^2]} \right\} = c_i, \quad i = 1, \ldots, T - 1; \quad s = T + 1. \]
\[ G_{ks} = E(\gamma_k \gamma_s) \]
\[ = \sum_{j=1}^{m_i} \left\{ \left[ \frac{n_{ij}^2 E(z_{ij} - \pi_i)^2}{[\pi_i(1 - \pi_i)[1 + (n_{ij} - 1)\phi]^2]} - 1 \right] \times R \right. \]
\[ + \frac{n_{ij}[1 + (n_{ij} - 1)\phi]}{(1 - \phi)} \times B \]
\[ - \frac{(1 - 2\pi_i)[1 + (2n_{ij} - 1)\phi]}{\pi_i(1 - \pi_i)(1 - \phi^2)} \times B \times R \right\} = b_i, \quad k = s = 1, \ldots, T \]
\[ i = 1, \ldots, T, \]
\[ G_{ks} = 0, \quad k = 1, \ldots, T; \quad k \neq s. \]
\[ G_{ks} = f_i, \quad k = 1, \ldots, T; \quad s = T + 1. \]
\[ G_{ks} = \sum_{i=1}^{T} d_i = d, \quad s = k = T + 1, \]

where
\[ B = \sum_{j=1}^{m_i} \left\{ \frac{(1 - \phi)}{[1 + (n_{ij} - 1)\phi]} + \frac{(1 - 2\pi_i)^2}{2\pi_i(1 - \pi_i)} \right\}, \]
\[ D = \left[ \frac{(1 + (n_{ij} - 1)\phi^2)}{(1 + (n_{ij} - 1)\phi)} \right], \]
and
\[ R = \left[ \frac{n_{ij}\phi(1 - 2\pi_i)}{2\pi_i(1 - \pi_i)(1 - \phi)} \right]. \]

Thus, the matrices \( J, H, A, C \) and \( G \) can be written in terms of \( d_i, f_i, g_i, e_i \) and \( c_i \) as
\[ J = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{T-1} \end{pmatrix} \]

\[ H = \begin{pmatrix} f_1 & 0 & \cdots & 0 & 0 & d_1 \\ 0 & f_2 & \cdots & 0 & 0 & d_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & f_{T-1} & 0 & d_{T-1} \end{pmatrix} \]

\[ A = \begin{pmatrix} g_1 & 0 & \cdots & 0 & c_1 \\ 0 & g_2 & \cdots & 0 & e_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & g_T & e_T \\ c_1 & e_2 & \cdots & e_T & c \end{pmatrix} \]

\[ C = \begin{pmatrix} e_1 & 0 & \cdots & 0 & 0 & c_1 \\ 0 & e_2 & \cdots & 0 & 0 & c_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & e_{T-1} & 0 & c_{T-1} \end{pmatrix} \]

\[ G = \begin{pmatrix} g_1 & 0 & \cdots & 0 & f_1 \\ 0 & g_2 & \cdots & 0 & f_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & g_T & f_T \\ f_1 & f_2 & \cdots & f_T & d \end{pmatrix} \]
CHAPTER V

TESTING OF HOMOGENEITY

OF PROPORTIONS IN PRESENCE

OF POSSIBLE UNEQUAL DISPERSION PARAMETERS

5.1 Introduction

In this chapter, we develop methods for testing the homogeneity of proportions in the presence of unequal dispersion parameters. As in the normal theory one-way analysis of variance situation, we consider testing of the homogeneity of dispersion parameters $\theta_i$ (derived in chapter 4) as a preliminary test. If this preliminary test is not rejected we test the homogeneity of proportions with common dispersion parameter $\theta$ as described in chapter 3. Otherwise we test the homogeneity of proportions with unequal dispersion parameters $\theta_i$. The most commonly used parametric method for testing the homogeneity of proportion in the presence of unequal dispersion parameters $\theta_i$ is the likelihood ratio test based on the beta-binomial likelihood (Aebisher et al., 1977; William, 1975; Pack, 1986a)). Because of the complexity and of poor performance of the likelihood ratio test we omit this from further consideration.

Our main interest is to develop a $C(\alpha)$ or score test based on semi-parametric models. We derive a $C(\alpha)$ or score test based on the quasi-likelihood and method of moments (Breslow, 1990a) in section 5.2. As a matter of interest we derive a $C(\alpha)$ test based on the quasi-likelihood and method of moments proposed by Srivastava and Wu (1993) in section 5.3 in order to compare the performance in terms of size with the $C(\alpha)$ test derived in section 5.2. A small simulation study was conducted
to compare the behaviours in terms of size and robustness for departure from data
distribution of the two \( C(\alpha) \) statistics \( C_{Q^*} \) and \( C_{M_2} \).

We assume \( Y_{ij} \sim BB(\pi_i, \theta_i) \). We are interested in testing the homogeneity of
proportions when the dispersion parameters are unequal.

\[
H_0 : \pi_1 = \ldots = \pi_T \quad \text{against} \quad H_A : \text{not all } \pi_i \text{'s are the same},
\]

with the parameters \( \theta_1, \ldots, \theta_T \) remaining unspecified. This hypothesis is anal-
ogous to the famous Behrens-Fisher problem for \( T = 2 \) and its extension for \( T > 2 \).

\[5.2 \] The \( C(\alpha) \) Statistics based on the quasi-likelihood and method of mo-
ments (Breslow's)

We reparametrize \( \pi_i \) under \( H_A \) by \( \pi_i = \pi + \tau_i \) with \( \tau_T = 0 \). Then testing \( H_0 \)
is equivalent to testing \( \tau_i = 0, i = 1, \ldots, T-1 \) with \( \pi \) and \( \phi_1, \ldots, \phi_T \) are treated
as nuisance parameters. Define

\[
\tau = (\tau_1, \ldots, \tau_T)
\]

and

\[
\lambda = (\lambda_1, \ldots, \lambda_{T+1}) = (\pi, \phi_1, \ldots, \phi_T).
\]

The unbiased estimating functions for parameters \( \tau_1, \ldots, \tau_{T-1} \) and \( \pi \) are obtained
using the quasi-likelihood scores. We find

\[
\psi_i = g_{1i} = \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi - \tau_i)n_{ij}}{(\pi + \tau_i)(1 - \pi - \tau_i)\{1 + (n_{ij} - 1)\phi_i\}}, \quad i = 1, \ldots, T - 1
\]

and

\[
g_{21} = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi - \tau_i)n_{ij}}{(\pi + \tau_i)(1 - \pi - \tau_i)\{1 + (n_{ij} - 1)\phi_i\}}.
\]
The unbiased estimating function for $\phi_i$, given $\tau_1, \ldots, \tau_{T-1}$ and $\pi$ can be obtained by using moment method (Breslow, 1990(a); Moore and Tsiatis, 1991), given by

$$g_{2.k-1} = \sum_{j=1}^{m} \frac{(z_{kj} - \pi - \tau_k)^2 n_{kj}}{(\pi + \tau_k)(1 - \pi - \tau_k)(1 + (n_{kj} - 1) \phi_k)} - (m_k - 2), \quad k = 1, \ldots, T.$$  

Then we let

$$\psi_i = g_{1i} \bigg|_{\tau=0}, \quad i = 1, \ldots, T - 1$$

and

$$\gamma_k = g_{2k} \bigg|_{\tau=0}, \quad k = 1, \ldots, T + 1.$$  

We define $\psi = (\psi_1, \ldots, \psi_{T-1})'$. Then asymptotically, as $m_i \to \infty$, using Lindeberg central limit theorem,

$$\psi(\lambda) \sim N(0, \Delta_{11} - \Delta_{12} \Delta_{22}^{-1} \Delta_{21}).$$

where

$$\Delta_{11} = E \left( \left. \frac{-\partial g_{1i}}{\partial \tau_i} \right|_{\tau=0} \right), \quad i, t = 1, \ldots, T - 1$$

$$\Delta_{12} = E \left( \left. \frac{-\partial g_{1i}}{\partial \lambda_k} \right|_{\tau=0} \right), \quad i = 1, \ldots, T - 1, \quad k = 1, \ldots, T + 1$$

$$\Delta_{21} = E \left( \left. \frac{-\partial g_{2k}}{\partial \tau_i} \right|_{\tau=0} \right), \quad i = 1, \ldots, T - 1, \quad k = 1, \ldots, T + 1$$

and

$$\Delta_{22} = E \left( \left. \frac{-\partial g_{2k}}{\partial \lambda_s} \right|_{\tau=0} \right), \quad k, s = 1, \ldots, T + 1.$$  

If $\lambda$ in $\psi$, $\Delta_{11}$, $\Delta_{12}$, $\Delta_{21}$ and $\Delta_{22}$ is replaced by some $\sqrt{m_i}$ consistent estimate of $\lambda$ under the null hypothesis then the $C(\alpha)$ statistic is

$$C_{Q^*} = \psi' (\Delta_{11} - \Delta_{12} \Delta_{22}^{-1} \Delta_{21})^{-1} \psi,$$
which has asymptotically a chi-square distribution with \((T-1)\) degrees of freedom.

The unbiased estimating equations for \(\pi\) and \(\phi_1, \ldots, \phi_T\) are

\[
\sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi) n_{ij}}{\pi(1 - \pi)\{1 + (n_{ij} - 1)\phi_i\}} = 0
\]

and

\[
\sum_{j=1}^{m_i} \frac{(z_{ij} - \pi)^2 n_{ij}}{\pi(1 - \pi)\{1 + (n_{ij} - 1)\phi_i\}} - (m_i - 2) = 0. \quad i = 1, \ldots, T
\]

The solution, if it exists, of the above \(T + 1\) equations provide the \(\sqrt{m_i}\) consistent estimate for \(\lambda\). Now, we obtain

\[
\Delta_{11i_t} = E\left(\left.\frac{-\partial g_{1i}}{\partial \tau_t}\right|_{\tau=0}\right) = \sum_{j=1}^{m_i} \frac{n_{ij}}{\pi(1 - \pi)\{1 + (n_{ij} - 1)\phi_i\}} = d_i, \quad i = 1, \ldots, T - 1.
\]

\[
\Delta_{11i_t} = 0. \quad i \neq t.
\]

\[
\Delta_{12i1} = E\left(\left.\frac{-\partial g_{1i}}{\partial \lambda_1}\right|_{\tau=0}\right) = d_i. \quad i = 1, \ldots, T - 1.
\]

\[
\Delta_{12i1+k} = E\left(\left.\frac{-\partial g_{1i}}{\partial \lambda_{1+k}}\right|_{\tau=0}\right) = 0. \quad k = 1, \ldots, T
\]

\[
\Delta_{211} = E\left(\left.\frac{-\partial g_{21}}{\partial \tau_i}\right|_{\tau=0}\right) = d_i. \quad i = 1, \ldots, T - 1.
\]

\[
\Delta_{211+k,i} = E\left(\left.\frac{-\partial g_{21+k}}{\partial \tau_i}\right|_{\tau=0}\right) = \sum_{j=1}^{m_i} \frac{(1 - 2\pi)}{\pi(1 - \pi)} = cm_i. \quad i, k = 1, \ldots, T - 1,
\]

\[
\Delta_{211+T,i} = 0. \quad i = 1, \ldots, T - 1.
\]

\[
\Delta_{2211} = E\left(\left.\frac{-\partial g_{21}}{\partial \lambda_1}\right|_{\tau=0}\right) = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{n_{ij}}{\pi(1 - \pi)\{1 + (n_{ij} - 1)\phi_i\}}.
\]

125
\[ \sum_{i=1}^{T} d_i = d, \]

\[ \Delta_{21}^{21.k+k} = E \left( \frac{\partial g_{21}}{\partial \lambda_{1+k}} \bigg|_{r=0} \right) = 0. \quad k = 1, \ldots, T. \]

\[ \Delta_{22.1+k,1} = E \left( \frac{\partial g_{21+k}}{\partial \lambda_{1+k}} \bigg|_{r=0} \right) = cm_i. \quad i, k = 1, \ldots, T. \]

\[ \Delta_{22,1+k+1+k} = E \left( \frac{\partial g_{21+k}}{\partial \lambda_{1+k}} \bigg|_{r=0} \right) \]

\[ = \sum_{j=1}^{n_i} \frac{(n_{ij} - 1)}{(1 + (n_{ij} - 1)\phi_i)} = c_i. \quad i, k = 1, \ldots, T \]

Then, after simplification (as shown in chapter 3)

\[ \Delta_{11} - \Delta_{12} \Delta_{22}^{-1} \Delta_{12} = \]

\[ \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{T-1} \end{pmatrix} - \begin{pmatrix} d_1^2/d & d_1 d_2/d & \cdots & d_1 d_{T-1}/d \\ d_1 d_2/d & d_2^2/d & \cdots & d_2 d_{T-1}/d \\ \vdots & \vdots & \ddots & \vdots \\ d_1 d_{T-1}/d & d_2 d_{T-1}/d & \cdots & d_{T-1}^2/d \end{pmatrix}. \]

Using the inverse of a partitioned matrix, (details are shown in chapter 3) we obtain

\[ C_{Q^*} = \psi'(\Delta_{11} - \Delta_{12} \Delta_{22}^{-1} \Delta_{12})^{-1} \psi \]

\[ = \sum_{i=1}^{T} \frac{\psi_i^2}{d_i}. \]

5.3 The C(α) Statistics based on the quasi-likelihood and method of moments (Srivastava’s)

We follow the same procedure as in the preceding section. The unbiased estimating functions for parameters \( \tau_1, \ldots, \tau_{T-1} \) and \( \pi \) are obtained using the quasi-likelihood scores

\[ \psi_i = g_{2i} = \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi - \tau_i)n_{ij}}{(\pi + \tau_i)(1 - \pi - \tau_i)(1 + (n_{ij} - 1)\phi_i)}, \quad i = 1, \ldots, T - 1 \]
and

$$g_{21} = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi - \tau_i)n_{ij}}{(\pi + \tau_i)(1 - \pi - \tau_i)(1 + (n_{ij} - 1)\phi_i)}.$$ 

The unbiased estimating function for $\phi_i$; given $\tau_1 \ldots \tau_{T-1}$ and $\pi$ can be obtained by using moment method proposed by Srivastava and Wu (1990), which is given below

$$g_{2,k+1} = \sum_{j=1}^{m_i} \left\{ \frac{n_{kj}(z_{kj} - \pi - \tau_k)^2}{(\pi + \tau_k)(1 - \pi - \tau_k)} - (1 + (n_{kj} - 1)\phi_k) \right\}, \quad k = 1 \ldots T.$$ 

Under the null hypothesis using the same notation as in section 5.2. the $C(\alpha)$ statistic is given by

$$C_{M2} = \psi'(\Delta_{11} - \Delta_{12} \Delta_{22}^{-1} \Delta_{12})^{-1} \psi,$$

which has asymptotically a chi-square distribution with $(T-1)$ degrees of freedom.

The $\sqrt{m_i}$ consistent estimates, in this case, are obtained by solving the unbiased estimating equations for $\pi$ and $\phi_1 \ldots \phi_T$:

$$\sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{(z_{ij} - \pi)n_{ij}}{\pi(1 - \pi)(1 + (n_{ij} - 1)\phi_i)} = 0$$

and

$$\sum_{j=1}^{m_i} \left\{ \frac{n_{ij}(z_{ij} - \pi)}{\pi(1 - \pi)} - (1 + (n_{ij} - 1)\phi_i) \right\} = 0, \quad i = 1, \ldots, T$$

simultaneously. Also in this case the elements of the matrices $\Delta_{11}$, $\Delta_{12}$, $\Delta_{21}$ and $\Delta_{22}$ are

$$\Delta_{11it} = E\left( \frac{-\partial g_{ii}}{\partial \tau_t} \right)_{\tau=0} = \sum_{j=1}^{m_i} \frac{n_{ij}}{\pi(1 - \pi)(1 + (n_{ij} - 1)\phi_i)} = d_i, \quad i, t = 1, \ldots, T-1$$

$$\Delta_{11it} = 0, \quad i \neq t$$

$$\Delta_{12it} = E\left( \frac{-\partial g_{ii}}{\partial \lambda_t} \right)_{\tau=0} = d_i, \quad i = 1, \ldots, T-1$$

127
\[ \Delta_{12,i+1} = E \left( \frac{-\partial g_{21}}{\partial \lambda_{1+k}} \bigg|_{\tau=0} \right) = 0, \quad k = 1, \ldots, T \]

\[ \Delta_{21,i} = E \left( \frac{-\partial g_{21}}{\partial \tau_i} \bigg|_{\tau=0} \right) = d_i, \quad i = 1, \ldots, T - 1 \]

\[ \Delta_{211+k} = E \left( \frac{-\partial g_{21}}{\partial \tau_i} \bigg|_{\tau=0} \right) = \sum_{j=1}^{m_i} \frac{(1 - 2\pi)(1 + (n_{ij} - 1)\phi_i)}{\pi(1 - \pi)} = c_i, \quad i, k = 1, \ldots, T - 1 \]

\[ \Delta_{21,1+T,i} = 0, \quad i = 1, \ldots, T - 1 \]

\[ \Delta_{2211} = E \left( \frac{-\partial g_{21}}{\partial \lambda_1} \bigg|_{\tau=0} \right) = \sum_{i=1}^{T} \sum_{j=1}^{m_i} \frac{n_{ij}}{\pi(1 - \pi)(1 + (n_{ij} - 1)\phi_i)} = \sum_{i=1}^{T} d_i = d \]

\[ \Delta_{221,k} = E \left( \frac{-\partial g_{21}}{\partial \lambda_{1+k}} \bigg|_{\tau=0} \right) = 0, \quad k = 1, \ldots, T \]

\[ \Delta_{22,1+k,1} = E \left( \frac{-\partial g_{21}}{\partial \lambda_1} \bigg|_{\tau=0} \right) = \sum_{j=1}^{m_i} \frac{(1 - 2\pi)(1 + (n_{ij} - 1)\phi_i)}{\pi(1 - \pi)} = c_i, \quad i, k = 1, \ldots, T \]

\[ \Delta_{22,1+k,1+k} = E \left( \frac{-\partial g_{21}}{\partial \lambda_{1+k}} \bigg|_{\tau=0} \right) = \sum_{j=1}^{m_i} (n_{ij} - 1) = c_i, \quad i, k = 1, \ldots, T \]

The entries of matrices are

\[
\Delta_{11} = \begin{pmatrix}
d_1 & 0 & \cdots & 0 \\
0 & d_2 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & 0 & \cdots & d_{T-1}
\end{pmatrix},
\]

\[
\Delta_{12} = \begin{pmatrix}
d_1 & 0 & \cdots & 0 \\
d_2 & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
d_{T-1} & 0 & \cdots & 0
\end{pmatrix},
\]

128
\[
\Delta_{21} = \begin{pmatrix}
  d_1 & d_2 & \cdots & d_{T-1} \\
  c_1 & 0 & \cdots & 0 \\
  0 & c_2 & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & c_{T-1} \\
  0 & 0 & \cdots & 0
\end{pmatrix}
\]

and

\[
\Delta_{22} = \begin{pmatrix}
  d & 0 & 0 & \cdots & 0 \\
  c_1 & c_1 & 0 & \cdots & 0 \\
  c_2 & 0 & c_2 & \cdots & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  c_T & 0 & 0 & \cdots & c_T
\end{pmatrix}
\]

After simplification and using the inverse of a partitioned matrix, we obtain

\[
C_{M3} = \psi' (\Delta_{11} - \Delta_{12} \Delta_{22}^{-1} \Delta_{12})^{-1} \psi
\]

\[
= \sum_{i=1}^{T} \frac{\psi_i^2}{d_i}.
\]

### 5.4 Simulations

A small simulation study was conducted to compare the performance, in terms of size and robustness of the \(C(\alpha)\) statistics \(C_Q\) and \(C_{M3}\). In this study we consider \(T = 2\) groups. The litter sizes and number of litters were chosen as those of the control group \((m_1 = 27)\) and the low dose treatment group \((m_2 = 19)\) of Paul (1982). The litter sizes of both groups are given in chapter 3. Beta-binomial observations were generated using IMSL subroutines RNBET and RNBIN. Empirical levels were calculated based on 1000 replications for each combination of \(\pi = 0.06, 0.08, 0.10, 0.15, 0.20, 0.30\) and \((\theta_1, \theta_2) = (0.06, 0.08), (0.06, 0.10), (0.06, 0.12), (0.06, 0.16), (0.06, 0.20), (0.06, 0.40), (0.06, 0.80)\). For studying robustness we generate data from the probit normal binomial (PNB) distribution.
The empirical levels corresponding to nominal level $\alpha=0.05$ are given in table 5.1(a) for data generated from the beta-binomial distribution and in table 5.1(b) for data generated from the PNB distribution. We note that both statistics $C_{Q^*}$ and $C_{M3}$ hold nominal level well for almost all of the parameter space examined except when $(\theta_1-\theta_2)$ is largest (i.e. $(\theta_1, \theta_2) = (0.06, 0.80)$) and $\pi$ is low ($\pi =0.06, 0.08$) in which case they show liberal behaviour. The data departures do not seem to have much effect on the performance of the statistics $C_{Q^*}$ and $C_{M3}$.

We compare the empirical levels of the statistic $C_{Q^*}$ given in table 5.1(a) and the statistic $C_Q$ (based on testing of homogeneity of proportions under departure of homogeneity of dispersion discussed in chapter 3) given in table 3.7(I). We observe that for $(\theta_1, \theta_2) = (0.06, 0.08), (0.06, 0.10), (0.06, 0.12), (0.06, 0.20)$, the two statistics $C_{Q^*}$ and $C_Q$ maintain the significance level well. However, when $(\theta_1, \theta_2) = (0.06, 0.40), (0.06, 0.80)$ and $\pi=0.20, 0.30$, we see that the statistic $C_Q$ shows liberal behaviour, whereas the statistic $C_{Q^*}$ maintains the level well. We have shown that the statistic $C_{Q^*}$ performs well for almost all of the parameter space examined. Therefore we recommend the use of this statistic when testing for the homogeneity of proportions in the presence of possibly widely varying unequal dispersion parameters.
Table 5.1(a): Empirical level (%) of the statistics $C_Q$, and $C_{M3}$: $\alpha = 0.05$ : T=2 groups based on 1000 replications, using BB distribution.

(I) $C_Q$

<table>
<thead>
<tr>
<th>$\theta_1, \theta_2$</th>
<th>$\pi$</th>
<th>0.06</th>
<th>0.08</th>
<th>0.10</th>
<th>0.15</th>
<th>0.20</th>
<th>0.30</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.06, 0.08)</td>
<td></td>
<td>3.1</td>
<td>3.4</td>
<td>3.0</td>
<td>4.4</td>
<td>4.5</td>
<td>4.7</td>
</tr>
<tr>
<td>(0.06, 0.10)</td>
<td></td>
<td>3.8</td>
<td>5.2</td>
<td>4.5</td>
<td>3.6</td>
<td>4.1</td>
<td>4.3</td>
</tr>
<tr>
<td>(0.06, 0.12)</td>
<td></td>
<td>4.3</td>
<td>4.1</td>
<td>4.7</td>
<td>4.7</td>
<td>4.0</td>
<td>4.7</td>
</tr>
<tr>
<td>(0.06, 0.16)</td>
<td></td>
<td>4.1</td>
<td>4.2</td>
<td>5.2</td>
<td>4.1</td>
<td>4.2</td>
<td>4.1</td>
</tr>
<tr>
<td>(0.06, 0.20)</td>
<td></td>
<td>5.7</td>
<td>5.9</td>
<td>4.6</td>
<td>5.0</td>
<td>4.7</td>
<td>4.4</td>
</tr>
<tr>
<td>(0.06, 0.40)</td>
<td></td>
<td>5.5</td>
<td>6.7</td>
<td>5.8</td>
<td>6.0</td>
<td>3.6</td>
<td>3.1</td>
</tr>
<tr>
<td>(0.06, 0.80)</td>
<td></td>
<td>7.5</td>
<td>8.7</td>
<td>5.8</td>
<td>5.9</td>
<td>5.4</td>
<td>4.6</td>
</tr>
</tbody>
</table>

(II) $C_{M3}$

<table>
<thead>
<tr>
<th>$\theta_1, \theta_2$</th>
<th>$\pi$</th>
<th>0.06</th>
<th>0.08</th>
<th>0.10</th>
<th>0.15</th>
<th>0.20</th>
<th>0.30</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.06, 0.08)</td>
<td></td>
<td>5.2</td>
<td>4.8</td>
<td>6.1</td>
<td>5.9</td>
<td>5.8</td>
<td>4.4</td>
</tr>
<tr>
<td>(0.06, 0.10)</td>
<td></td>
<td>5.6</td>
<td>5.6</td>
<td>6.3</td>
<td>5.9</td>
<td>6.2</td>
<td>4.6</td>
</tr>
<tr>
<td>(0.06, 0.12)</td>
<td></td>
<td>5.1</td>
<td>5.8</td>
<td>6.4</td>
<td>6.0</td>
<td>6.2</td>
<td>4.5</td>
</tr>
<tr>
<td>(0.06, 0.16)</td>
<td></td>
<td>5.6</td>
<td>5.4</td>
<td>6.6</td>
<td>6.1</td>
<td>6.4</td>
<td>4.6</td>
</tr>
<tr>
<td>(0.06, 0.20)</td>
<td></td>
<td>4.8</td>
<td>5.5</td>
<td>7.1</td>
<td>6.9</td>
<td>5.7</td>
<td>4.9</td>
</tr>
<tr>
<td>(0.06, 0.40)</td>
<td></td>
<td>6.5</td>
<td>6.8</td>
<td>8.5</td>
<td>8.6</td>
<td>6.2</td>
<td>5.5</td>
</tr>
<tr>
<td>(0.06, 0.80)</td>
<td></td>
<td>8.0</td>
<td>9.2</td>
<td>9.5</td>
<td>8.0</td>
<td>7.4</td>
<td>6.4</td>
</tr>
</tbody>
</table>
Table 5.1(b): Empirical level (%) of the statistics $C_Q$ and $C_{M3}$: $\alpha = 0.05 : T=2$
groups based on 1000 replications. using PNB distribution.

(I) $C_Q$

<table>
<thead>
<tr>
<th>$(\theta_1, \theta_2)$</th>
<th>$\pi$</th>
<th>0.06</th>
<th>0.08</th>
<th>0.10</th>
<th>0.15</th>
<th>0.20</th>
<th>0.30</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.06,0.08)</td>
<td>4.0</td>
<td>4.7</td>
<td>4.6</td>
<td>4.1</td>
<td>4.0</td>
<td>2.9</td>
<td></td>
</tr>
<tr>
<td>(0.06,0.10)</td>
<td>3.9</td>
<td>4.9</td>
<td>5.1</td>
<td>4.2</td>
<td>3.7</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>(0.06,0.12)</td>
<td>4.2</td>
<td>5.5</td>
<td>5.2</td>
<td>4.6</td>
<td>4.2</td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td>(0.06,0.16)</td>
<td>4.7</td>
<td>4.7</td>
<td>5.8</td>
<td>4.4</td>
<td>4.6</td>
<td>3.6</td>
<td></td>
</tr>
<tr>
<td>(0.06,0.20)</td>
<td>5.1</td>
<td>4.2</td>
<td>5.2</td>
<td>5.2</td>
<td>4.5</td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td>(0.06,0.40)</td>
<td>7.5</td>
<td>6.7</td>
<td>6.3</td>
<td>6.8</td>
<td>4.6</td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td>(0.06,0.80)</td>
<td>7.3</td>
<td>9.1</td>
<td>8.6</td>
<td>6.6</td>
<td>5.7</td>
<td>5.1</td>
<td></td>
</tr>
</tbody>
</table>

(II) $C_{M3}$

<table>
<thead>
<tr>
<th>$(\theta_1, \theta_2)$</th>
<th>$\pi$</th>
<th>0.06</th>
<th>0.08</th>
<th>0.10</th>
<th>0.15</th>
<th>0.20</th>
<th>0.30</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.06,0.08)</td>
<td>3.5</td>
<td>4.2</td>
<td>4.7</td>
<td>5.7</td>
<td>5.8</td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td>(0.06,0.10)</td>
<td>4.0</td>
<td>5.6</td>
<td>5.5</td>
<td>4.4</td>
<td>5.2</td>
<td>5.7</td>
<td></td>
</tr>
<tr>
<td>(0.06,0.12)</td>
<td>4.2</td>
<td>5.3</td>
<td>5.9</td>
<td>5.9</td>
<td>5.3</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>(0.06,0.16)</td>
<td>4.6</td>
<td>5.0</td>
<td>6.2</td>
<td>4.8</td>
<td>5.4</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>(0.06,0.20)</td>
<td>5.5</td>
<td>6.3</td>
<td>5.4</td>
<td>6.4</td>
<td>6.2</td>
<td>6.2</td>
<td></td>
</tr>
<tr>
<td>(0.06,0.40)</td>
<td>5.4</td>
<td>7.0</td>
<td>6.9</td>
<td>6.6</td>
<td>4.9</td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td>(0.06,0.80)</td>
<td>6.1</td>
<td>9.8</td>
<td>6.7</td>
<td>7.3</td>
<td>7.2</td>
<td>5.7</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER VI

SUMMARY AND RECOMMENDATIONS

FOR FUTURE RESEARCH

This Chapter summarizes the conclusions of this thesis and recommends some problems for future research.

6.1 Summary

If the data of proportions or counts show over dispersion, it is essential to account for it by fitting a parametric model such as beta-binomial or a semi-parametric model.

In Chapter 2, we investigated several procedures for the estimation of the mean (regression) and intraclass correlation or dispersion parameters in over/under dispersed binomial data without covariates. The following methods were considered for the joint estimation of the mean and the dispersion parameters: ML, QE, EQE, QEE, GL, M1, M2, M3, M4. We studied large and small sample efficiencies of the estimates of the mean and intraclass correlation parameters.

For the estimation of the mean(regression) parameters then either the Gaussian likelihood or the quasi-likelihood procedure might be chosen. The Gaussian likelihood estimate (GL), in general has higher efficiency, whereas the quasi-likelihood estimate is easier to compute. Both of these methods are robust to misspecification of the variance structure.

For the estimation of \( \phi \) or the joint estimation of the mean and the dispersion parameter \( \phi \), we recommend the Gaussian likelihood (GL) and the method of moments estimates M2 and M3. In general, the estimates by these procedures have high effi-
ciences and show consistent behaviour throughout the parameter space examined. The method of moments estimates M2 and M3 are computationally less intensive than the maximum likelihood (ML) estimates based on beta-binomial model. The method of moments estimates are also robust to misspecification of the variance structure.

In Chapter 3, we derived two $C(\alpha)$ statistics $C_m$ and $C_{mm}$ for testing homogeneity of proportions in toxicology in the presence of beta-binomial over dispersion. The performance of these statistics was then compared with the likelihood ratio statistic (LR). The two statistics $C_m$ and $C_{mm}$ hold the nominal level well and are similar to those corresponding to the likelihood ratio statistics. Either of the $C(\alpha)$ statistics can be used, as they require estimates of the parameter only under the null hypotheses. However, the $C(\alpha)$ statistic based on the moment estimates is computationally less intensive than the $C(\alpha)$ statistic based on the MLEs. The two $C(\alpha)$ statistics based on the extended beta-binomial, although they hold nominal level well, do not produce simple forms. Therefore we derived two $C(\alpha)$ statistics for testing homogeneity of proportions based on a semi-parametric model using the quasi-likelihood, $C_Q$ and extended quasi-likelihood, $C_{Q^+}$. We also derived the $C(\alpha)$ statistic, $C_{EQ}$ based on empirical variances using quasi-likelihood. We compared procedures $C_m$, $C_{mm}$, $C_Q$, $C_{Q^+}$, $C_{EQ}$ and two statistics RS and ARS developed by Rao and Scott (1992), through simulation, in terms of size, power and robustness for data distribution and dispersion heterogeneity. The score test, $C_Q$, based on the quasi-likelihood performs best in that it holds the nominal level well in all data distribution situations considered and also shows robustness in the presence of moderate dispersion heterogeneity. This statistic has a very simple form and requires estimates of the parameters only under the null hypotheses. The maximum quasi-likelihood estimates of the parameters under the null hypotheses required for
the $C_Q$ statistic are much more simply obtained than the corresponding maximum likelihood estimates under the beta-binomial model. The score statistic $C_{EQ}$ based on empirical variance shows similar performance to that of the model based $C_Q$ statistic. However, the $C_{EQ}$ statistic does not provide as simple a form as the model based $C_Q$.

When testing for homogeneity of proportions in Chapter 3 we assumed a common dispersion. However, in some situations this condition might not be present. In Chapter 4, we derived two $C(\alpha)$ statistics $C_\alpha$ and $C_{GL}$ for testing this assumption based on quasi-likelihood/method of moments and Gaussian likelihood.

In Chapter 5, we derived the two $C(\alpha)$ statistics $C_{Q^*}$ and $C_{M_3}$ to test the homogeneity of proportion in the presence of possible unequal dispersion parameter. These statistics hold the nominal level well and also show robustness for departure from the data distribution. The statistic $C_{Q^*}$ can be recommended for use when testing for homogeneity of proportions in the presence of possibly widely varying unequal dispersion parameters.

6.2 Recommendation for future research

(1) The score test for testing the significance of added variables for over/under dispersed binomial regression model, using quasi-likelihood/method of moments.

The quasi-likelihood (as derived in Chapter 2) for an observation $z_i = y_i/n_i$, $i = 1, \ldots, m$ is obtained as

$$Q = \sum_{i=1}^{m} \frac{1}{1 + (n_i - 1)\phi} \left[ y_i \log(\frac{\pi_i}{z_i}) + (n_i - y_i) \log\left(\frac{1 - \pi_i}{1 - z_i}\right)\right].$$

The mean structure $\pi_i$ is given by the logistic model

$$\pi_i(X_i; \beta) = \frac{e^{X_i \beta}}{1 + e^{X_i \beta}},$$

where $X_i \beta = X_1 \beta_1 + \ldots + X_k \beta_k$ and $X_1, \ldots, X_k$ are $k$ explanatory variables and $\beta_1, \ldots, \beta_k$ are the $k$ regression parameters.
Given \( \phi \), the unbiased estimating equations for \( \beta \) are

\[
U_j(\beta, \phi) = \frac{\partial Q}{\partial \beta_j} = \sum_{i=1}^{m} \frac{(z_i - \pi_i)u_i}{\pi_i(1 - \pi_i)(1 + (n_{ij} - 1)\phi)} \frac{\partial \pi_i}{\partial \beta_j} = 0
\]

\[
= \sum_{i=1}^{m} \frac{(z_i - \pi_i)u_iX_{ij}}{1 + (n_{ij} - 1)\phi}, \quad j = 1, \ldots, k.
\]

The unbiased estimating equation for \( \phi \) can be obtained by using the moment method (Breslow, 1990(a); Moore and Tsiatis, 1991)

\[
U_{k+1}(\beta, \phi) = \sum_{i=1}^{m} \frac{n_i(z_i - \pi_i)^2}{\pi_i(1 - \pi_i)(1 + (n_{ij} - 1)\phi)} - (m - k) = 0.
\]

Let

\[
\beta = (\beta_1, \ldots, \beta_k)^t
\]

and

\[
U = (U_1, \ldots, U_k, U_{k+1})^t.
\]

Now, let \( X_i, \beta \) and \( U \) be partitioned as

\[
X_i = (X_{1i}, X_{2i}), \quad \beta = (\beta_1, \beta_2)^t \quad \text{and} \quad U = (U_1, U_2)^t,
\]

where

\[
X_{1i} = (X_{11}, \ldots, X_{1k_1})^t,
\]

\[
X_{2i} = (X_{2k_1+1}, \ldots, X_{2k})^t,
\]

\[
\beta_1 = (\beta_1, \ldots, \beta_{k_1})^t,
\]

\[
\beta_2 = (\beta_{k_1+1}, \ldots, \beta_k)^t,
\]

\[
U_1 = (U_1, \ldots, U_{k_1}, U_{k+1})^t,
\]

and

\[
U_2 = (U_{k_1+1}, \ldots, U_k)^t.
\]
Thus, $\beta_1$ is of dimension $(k_1 \times 1)$ and $\beta_2$ is of dimension $(k - k_1) \times 1$. Therefore $X_i\beta$ can be written as

$$X_i\beta = X_1i\beta_1 + X_2i\beta_2.$$ 

We are interested in testing

$$H_0 : \beta_2 = 0 \quad \text{against} \quad H_A : \beta_2 \neq 0.$$ 

treating $\lambda = (\beta_1, \ldots, \beta_{k_1}, \phi)$ as nuisance parameters. The score statistic for testing the null hypothesis is based on $U_{2j}(\beta, \phi)$ where

$$U_{2j}(\beta, \phi) = \sum_{i=1}^{m} \frac{(z_i - \pi_i)n_iX_{2ij}}{1 + (n_{ij} - 1)\phi}, \quad j = k_1 + 1, \ldots, k.$$ 

Denote the estimates of $\lambda$ by $\tilde{\lambda}$.

Using a Taylor series expansion of $U_2(\tilde{\lambda})$ and $U_1(\tilde{\lambda})$ we obtain

$$U_2(\tilde{\lambda}) = U_2(\lambda^0) - \frac{\partial U_2}{\partial \lambda} (\tilde{\lambda} - \lambda^0) + o_p(\sqrt{m}) \quad (6.2.1)$$ 

and

$$U_1(\tilde{\lambda}) = U_1(\lambda^0) - \frac{\partial U_1}{\partial \lambda} (\tilde{\lambda} - \lambda^0) + o_p(\sqrt{m}). \quad (6.2.2)$$ 

From equation (6.2.2) we obtain

$$(\tilde{\lambda} - \lambda^0) \approx \left( \frac{-\partial U_1}{\partial \lambda} \right) (\tilde{\lambda} - \lambda^0). \quad (6.2.3)$$ 

Substituting equation (6.2.3) into equation (6.2.1) we get

$$U_2(\tilde{\lambda}) = U_2(\lambda^0) - \left( \frac{-\partial U_2}{\partial \lambda} \right) \left( \frac{-\partial U_1}{\partial \lambda} \right)^{-1} U_1(\lambda^0).$$ 

We replace $\left( \frac{-\partial U_2}{\partial \lambda} \right)$ and $\left( \frac{-\partial U_1}{\partial \lambda} \right)$ by their asymptotically equivalent versions, namely their expectations $A_{12} = E\left( \frac{-\partial U_2}{\partial \lambda} \right)$ and $A_{11} = E\left( \frac{-\partial U_1}{\partial \lambda} \right)$, we obtain

$$U_2(\tilde{\lambda}) = U_2(\lambda^0) - A_{21}A_{11}^{-1} U_1(\lambda^0).$$
Then, the asymptotic variance of $U_2(\hat{\lambda})$ is

$$Var\left(U_2(\hat{\lambda})\right) = Var\left(U_2(\lambda^n) - A_{21}A_{11}^{-1}U_1(\lambda^n)\right).$$

Now, the actual variance can be estimated by replacing $\lambda^n$ by $\hat{\lambda}$ in

$$Var\left(U_2(\hat{\lambda})\right) = Var\left(U_2(\lambda^0) - A_{21}A_{11}^{-1}U_1(\lambda^0)\right)$$

$$= E\left(U_2(\lambda^0) - A_{21}A_{11}^{-1}U_1(\lambda^0)\right)^T\left(U_2(\lambda^0) - A_{21}A_{11}^{-1}U_1(\lambda^0)\right)$$

$$= E\left(U_2(\lambda^0)U_2^T(\lambda^0)\right) - E\left(U_2(\lambda^0)U_1^T(\lambda^0)\right)A_{11}^{-1}A_{12}$$

$$- A_{21}A_{11}^{-1}E\left(U_1(\lambda^0)U_2^T(\lambda^0)\right) + A_{21}A_{11}^{-1}E\left(U_1(\lambda^0)U_1^T(\lambda^0)\right)A_{11}^{-1}A_{12}.$$

If the variance has been correctly specified, then

$$E\left(U_2(\lambda^0)U_2^T(\lambda^0)\right) = A_{22},$$

$$E\left(U_2(\lambda^0)U_1^T(\lambda^0)\right) = A_{21},$$

$$E\left(U_1(\lambda^0)U_2^T(\lambda^0)\right) = A_{12},$$

and

$$E\left(U_1(\lambda^0)U_1^T(\lambda^0)\right) = A_{11}.$$

Then the variance of $U_2(\hat{\lambda})$ becomes

$$Var\left(U_2(\hat{\lambda})\right) = A_{22} - A_{21}A_{11}^{-1}A_{12}.$$

Therefore

$$U_2(\hat{\lambda}) \sim MN\left(0, \ A_{22} - A_{21}A_{11}^{-1}A_{12}\right).$$

where the dimensions of matrices $A_{11}$, $A_{12}$, $A_{21}$ and $A_{22}$ are $(k_1 + 1) \times (k_1 + 1)$, $(k_1 + 1) \times (k - k_1)$, $(k - k_1) \times (k_1 + 1)$ and $(k - k_1) \times (k - k_1)$ respectively with

$$A_{11} = E\left(\frac{-\partial U_1}{\partial \lambda}\right).$$

138
\[ A_{12} = E \left( \frac{-\partial U_1}{\partial \beta_2} \right). \]
\[ A_{21} = E \left( -\frac{\partial U_2}{\partial \lambda} \right). \]

and
\[ A_{22} = E \left( -\frac{\partial U_2}{\partial \beta_2} \right). \]

If we replace \( \lambda = (\beta_1, \ldots, \beta_{k_1}, \phi) \) by \( \sqrt{m} \) consistent estimates then the quasi-likelihood score statistic is
\[ C_{\delta} = U_2' (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} U_2. \]

which has asymptotically a chi-square distribution with \((k - k_1)\) degrees of freedom.

The score statistic \( C_{\delta} \) uses model based variance-covariance.

Breslow (1989: 1990(a), 1990(b)) suggests using the empirical variances of \( U_2(\hat{\lambda}) \)

which is

\[
\text{Empirical Var}\left(U_2(\hat{\lambda})\right) = \left(U_2(\lambda^0) - A_{21} A_{11}^{-1} U_1(\lambda^0)\right)\left(U_2(\lambda^0) - A_{21} A_{11}^{-1} U_1(\lambda^0)\right)^t
= U_2(\lambda^0)U_2^t(\lambda^0) - U_2(\lambda^0)U_2^t(\lambda^0)A_{11}^{-1}A_{12}
- A_{21} A_{11}^{-1} U_1(\lambda^0)U_2^t(\lambda^0) + A_{21} A_{11}^{-1} U_1(\lambda^0)U_1^t(\lambda^0)A_{11}^{-1}A_{12}
= G_{22} - G_{21} A_{11}^{-1} A_{12} - A_{21} A_{11}^{-1} G_{12} + A_{21} A_{11}^{-1} G_{11} A_{11}^{-1} A_{12}
\]

where \( G_{22} = \left(U_2(\lambda^0)U_2^t(\lambda^0)\right), \quad G_{21} = \left(U_2(\lambda^0)U_1^t(\lambda^0)\right), \quad G_{12} = \left(U_1(\lambda^0)U_2^t(\lambda^0)\right), \quad \)

and \( G_{11} = \left(U_1(\lambda^0)U_1^t(\lambda^0)\right). \) Then the score statistic using the empirical variance of \( U_2(\hat{\lambda}) \) is

\[ C_{\delta^*} = U_2' \left(G_{22} - G_{21} A_{11}^{-1} A_{12} - A_{21} A_{11}^{-1} G_{12} + A_{21} A_{11}^{-1} G_{11} A_{11}^{-1} A_{12}\right)^{-1} U_2. \]

Further work is required to investigate and compare the performance of these two \( C(\alpha) \) statistics \( C_{\delta} \) and \( C_{\delta^*} \) in terms of size, power and robustness.
(2) Another possible area of research is to derive various procedures for constructing the confidence intervals of the mean (regression) and dispersion parameters.

(3) From the literature of all the work done in this area, different litter numbers and litter sizes have been used when testing the homogeneity of proportions. A comprehensive study is suggested to investigate the effect of number of litters and litter sizes on the inference concerning the homogeneity of proportions in the presence of equal and unequal dispersion.

(4) The performances of the $C(\alpha)$ statistics for testing the homogeneity of proportions under equal and unequal dispersion parameters were investigated using a simulation study for two groups. A simulation study is suggested to investigate the effect of equal or unequal dispersion on inferences concerning the proportions for more than two groups.
APPENDIX

Appendix A

(i) Derivation of the first component of $E\left(\frac{\partial^2}{\partial \theta_1 \partial \theta_2}\right)$.

$$E\left(\sum_{r=0}^{y_i-1} \frac{1}{(\pi_i + r\theta)^2}\right) = \sum_{y_i=0}^{n_i} \sum_{r=0}^{y_i-1} \frac{P(Y_i)}{(\pi_i + r\theta)^2}$$

$$= \frac{P(1)}{(\pi_i + 0\theta)^2} + \frac{P(2)}{(\pi_i + 0\theta)^2} + \frac{P(2)}{(\pi_i + \theta)^2} + \frac{P(3)}{(\pi_i + \theta)^2} + \frac{P(3)}{(\pi_i + 2\theta)^2} + \cdots$$

$$+ \frac{P(n_i)}{(\pi_i + 0\theta)^2} + \frac{P(n_i)}{(\pi_i + \theta)^2} + \cdots + \frac{P(n_i)}{(\pi_i + (n_i-1)\theta)^2}$$

$$= \frac{P(Y_i > 0)}{(\pi_i + 0\theta)^2} + \frac{P(Y_i > 1)}{(\pi_i + \theta)^2} + \cdots + \frac{P(Y_i > n_i - 1)}{(\pi_i + (n_i-1)\theta)^2}$$

$$= \sum_{r=0}^{n_i-1} \frac{P(Y_i > r)}{(\pi_i + r\theta)^2}.$$
\[
= \sum_{y_i=0}^{n_i} \sum_{r=0}^{y_i-1} \frac{P(Y_i)}{(1 - \pi_i + r\theta)^2} \\
= \left\{ \frac{P(0)}{(1 - \pi_i + 0\theta)^2} + \frac{P(0)}{(1 - \pi_i + \theta)^2} + \cdots + \frac{P(0)}{(1 - \pi_i + (n_i - 1)\theta)^2} \right\} \\
+ \frac{P(n_i - 2)}{(1 - \pi_i + 2\theta)^2} + \frac{P(n_i - 2)}{(1 - \pi_i + \theta)^2} \\
+ \cdots \\
+ \frac{P(n_i - 1)}{(1 - \pi_i + n_i\theta)^2}
\]

= \frac{P(Y_i < 0)}{(1 - \pi_i + 0\theta)^2} + \frac{P(Y_i < n_i - 1)}{(1 - \pi_i + \theta)^2} + \cdots + \frac{P(Y_i < 1)}{(1 - \pi_i + (n_i - 1)\theta)^2} \\
= \sum_{r=0}^{n_i-1} \frac{P(Y_i < n_i - r)}{(1 - \pi_i + r\theta)^2}.
\]

(ii) Derivation of the first component of \(E \left( \frac{\partial^2 i}{\partial \theta_1 \partial \theta} \right) \).

\[
E \left( \sum_{r=0}^{y_i-1} \frac{r}{(\pi_i + r\theta)^2} \right) = \sum_{y_i=0}^{n_i} \sum_{r=0}^{y_i-1} \frac{rP(Y_i)}{(\pi_i + r\theta)^2}.
\]

We multiply the numerator and the denominator by \(\theta\) of the right hand side of the above equation, and obtain

\[
\sum_{y_i=0}^{n_i} \sum_{r=0}^{y_i-1} \frac{r\theta P(Y_i)}{(\pi_i + r\theta)^2} = \sum_{y_i=0}^{n_i} \sum_{r=0}^{y_i-1} \frac{(\pi_i + r\theta - \pi_i)P(Y_i)}{\theta(\pi_i + r\theta)^2} \\
= \frac{1}{\theta} \sum_{y_i=0}^{n_i} \sum_{r=0}^{y_i-1} \frac{P(Y_i)}{(\pi_i + r\theta)^2} - \frac{1}{\theta} \sum_{y_i=0}^{n_i} \sum_{r=0}^{y_i-1} \frac{\pi_i P(Y_i)}{(\pi_i + r\theta)^2}.
\]

We know that \(E \left( \frac{\partial^1 i}{\partial \theta_1} \right) = 0\), thus

\[
\sum_{y_i=0}^{n_i} \sum_{r=0}^{y_i-1} \frac{P(Y_i)}{(\pi_i + r\theta)^2} = 0,
\]

142
therefore:

\[
\sum_{y_i=0}^{n_i} \sum_{r=0}^{y_i-1} \frac{rP(Y_i)}{(\pi_i + r\theta)^2} = -\frac{1}{\theta} \sum_{y_i=0}^{n_i} \sum_{r=0}^{y_i-1} \frac{\pi_i P(Y_i)}{(\pi_i + r\theta)^2} = -\frac{1}{\theta} \sum_{r=0}^{n_i-1} \frac{\pi_i P(Y_i > r)}{(\pi_i + r\theta)^2}.
\]

We follow similar steps, for the second component of \( E\left( \frac{\partial^2 l}{\partial \theta^2} \right) \).

\[
E\left( \sum_{r=0}^{n_i-y_i-1} \frac{r}{(1-\pi_i + r\theta)^2} \right) = \sum_{y_i=0}^{n_i} \sum_{r=0}^{n_i-y_i-1} \frac{rP(Y_i)}{(1-\pi_i + r\theta)^2}
\]

\[
= \sum_{y_i=0}^{n_i} \sum_{r=0}^{n_i-y_i-1} \frac{(1-\pi_i + r\theta) - (1-\pi_i)P(Y_i)}{\theta(1-\pi_i + r\theta)^2}
\]

\[
= \frac{1}{\theta} \sum_{y_i=0}^{n_i} \sum_{r=0}^{n_i-y_i-1} \frac{P(Y_i)}{(1-\pi_i + r\theta)^2} - \frac{1}{\theta} \sum_{y_i=0}^{n_i} \sum_{r=0}^{n_i-y_i-1} \frac{(1-\pi_i)P(Y_i)}{(1-\pi_i + r\theta)^2}
\]

\[
= 0 - \frac{1}{\theta} \sum_{y_i=0}^{n_i} \sum_{r=0}^{n_i-y_i-1} \frac{(1-\pi_i)P(Y_i)}{(1-\pi_i + r\theta)^2}
\]

\[
= -\frac{1}{\theta} \sum_{r=0}^{n_i-1} \frac{(1-\pi_i)P(Y_i < n_i - r)}{(1-\pi_i + r\theta)^2}.
\]

(iii) Derivation of \( E\left( \frac{\partial^2 l}{\partial \theta^2} \right) \).

The expectation of the quantity \( E\left( \frac{\partial^2 l}{\partial \theta^2} \right) \) is easily obtained by rewriting the log-likelihood \( l \) in terms of \( \pi \) and \( c = \theta^{-1} \).

\[
l = \sum_{r=0}^{y_i-1} \log(\pi_i + r\theta) + \sum_{r=0}^{n_i-y_i-1} \log(1-\pi_i + r\theta) - \sum_{r=0}^{n_i-1} \log(1 + r\theta).
\]

Let \( c = \frac{1}{\theta} \), then

\[
l = \sum_{r=0}^{y_i-1} \log(\pi_i c + r) + \sum_{r=0}^{n_i-y_i-1} \log((1-\pi_i)c + r) - \sum_{r=0}^{n_i-1} \log(c + r)
\]

\[
-\frac{\partial^2 l}{\partial c^2} = \sum_{r=0}^{y_i-1} \frac{\pi_i^2}{(\pi_i c + r)^2} + \sum_{r=0}^{n_i-y_i-1} \frac{(1-\pi_i)^2}{((1-\pi_i)c + r)^2} - \sum_{r=0}^{n_i-1} \frac{1}{(c + r)^2}
\]

143
\[ E\left(-\frac{\partial^2 l}{\partial \phi_i^2}\right) = \sum_{y_i = 0}^{n_i} \left\{ \sum_{r=0}^{y_i - 1} \frac{\pi_i^2}{(\pi_i c + r)^2} + \sum_{r=0}^{n_i - y_i - 1} \frac{(1 - \pi_i)^2}{((1 - \pi_i)c + r)^2} - \sum_{r=0}^{n_i - 1} \frac{1}{(c + r)^2} \right\} P(Y_i). \]

The above equation is similar to the equation obtained in part (i), therefore:

\[ E\left(-\frac{\partial^2 l}{\partial \phi_i^2}\right) = \theta^{-4} E\left(-\frac{\partial^2 l}{\partial \phi_i^2}\right) \] and substituting \( e = \theta^{-1} \), we obtain:

\[ E\left(-\frac{\partial^2 l}{\partial \phi_i^2}\right) = \frac{1}{\delta} \left\{ \sum_{r=0}^{n_i - 1} \frac{\pi_i^2}{(\pi_i + r\theta)^2} P(Y_i > r) + \sum_{r=0}^{n_i - 1} \frac{(1 - \pi_i)^2}{(1 - \pi_i + r\theta)^2} P(Y_i < n_i - r) - \sum_{r=0}^{n_i - 1} \frac{1}{(1 + r\theta)^2} \right\}. \]

Appendix B

To show \( E \left( \frac{\partial Q^+}{\partial \phi_i} \right) \approx 0 \). From equation (3.4.2.3),

\[ \frac{\partial Q^+}{\partial \phi_i} = -\frac{1}{2} \sum_{j=1}^{m_i} \frac{(n_{ij} - 1)}{1 + (n_{ij} - 1)\phi_i} + \sum_{j=1}^{m_i} \frac{(n_{ij} - 1)}{1 + (n_{ij} - 1)\phi_i} \left[ y_{ij}\log(y_{ij}) 

- y_{ij}\log(n_{ij}) - (n_{ij} - y_{ij})\log(n_{ij}(1 - \pi_i)) + (n_{ij} - y_{ij})\log(n_{ij} - y_{ij}) \right]. \]

Now,

\[ E\left(y_{ij}\log(n_{ij})\right) = n_{ij}\pi_i log(n_{ij}) \pi_i \]

and

\[ E\left[(n_{ij} - y_{ij})\log(n_{ij}(1 - \pi_i))\right] = n_{ij}(1 - \pi_i) log(n_{ij}(1 - \pi_i)). \]

To find expectations of \( y_{ij}\log(y_{ij}) \) and \( (n_{ij} - y_{ij})\log(n_{ij} - y_{ij}) \) we use the Taylor series expansion

\[ f(Y) \approx f(a) + (Y - a)f'(a) + \frac{(Y - a)^2}{2} f''(a), \]

where \( a = E(Y) \). Then taking expectations we obtain

\[ E\left(f(Y)\right) \approx f(a) + \frac{\text{Var}(Y)}{2} f''(a). \]

144
where
\[
a = E(y_{ij}) = n_{ij} \pi_i, \quad \text{Var}(y_{ij}) = n_{ij} \pi_i(1 - \pi_i)(1 + (n_{ij} - 1)\phi_i)
\]
\[
f(Y) = y_{ij} \log(y_{ij}), \quad f'(Y) = \log(y_{ij}) + 1, \quad f''(Y) = \frac{1}{y_{ij}}.
\]
\[
f''(a) = \frac{1}{n_{ij} \pi_i}.
\]

Therefore,
\[
E\left(y_{ij} \log(y_{ij})\right) \simeq n_{ij} \pi_i \log(n_{ij} \pi_i) + \frac{n_{ij} \pi_i(1 - \pi_i)(1 + (n_{ij} - 1)\phi_i)}{2n_{ij} \pi_i}
\]
\[
\simeq n_{ij} \pi_i \log(n_{ij} \pi_i) + \frac{(1 - \pi_i)(1 + (n_{ij} - 1)\phi_i)}{2}.
\]

Also
\[
f(Y) = (n_{ij} - y_{ij}) \log(n_{ij} - y_{ij}), \quad f'(Y) = \log(n_{ij} - y_{ij}) + 1
\]
\[
f''(Y) = \frac{1}{(n_{ij} - y_{ij})}, \quad f''(a) = \frac{1}{n_{ij}(1 - \pi_i)}
\]
\[
E\left((n_{ij} - y_{ij}) \log(n_{ij} - y_{ij})\right) \simeq n_{ij}(1 - \pi_i) \log(n_{ij}(1 - \pi_i))
\]
\[
+ \frac{n_{ij} \pi_i(1 - \pi_i)(1 + (n_{ij} - 1)\phi_i)}{2n_{ij}(1 - \pi_i)}
\]
\[
\simeq n_{ij}(1 - \pi_i) \log(n_{ij}(1 - \pi_i)) + \frac{\pi_i(1 + (n_{ij} - 1)\phi_i)}{2}.
\]

\[
E\left(\frac{\partial Q^+}{\partial \phi_i}\right) = -\frac{1}{2} \sum_{j=1}^{m_i} \frac{(n_{ij} - 1)}{1 + (n_{ij} - 1)\phi_i} + \sum_{j=1}^{m_i} \frac{(n_{ij} - 1)}{1 + (n_{ij} - 1)\phi_i}^2 \left[ n_{ij} \pi_i \log(n_{ij} \pi_i)
\right.
\]
\[
\left. + \frac{(1 - \pi_i)(1 + (n_{ij} - 1)\phi_i)}{2} - n_{ij} \pi_i \log(n_{ij} \pi_i) + n_{ij}(1 - \pi_i) \log(n_{ij}(1 - \pi_i)) \right]
\]
\[
+ \frac{\pi_i(1 + (n_{ij} - 1)\phi_i)}{2} - n_{ij}(1 - \pi_i) \log(n_{ij}(1 - \pi_i)) \right] .
\]

After simplification,
\[
E\left(\frac{\partial Q^+}{\partial \phi_i}\right) = -\frac{1}{2} \sum_{j=1}^{m_i} \frac{(n_{ij} - 1)}{1 + (n_{ij} - 1)\phi_i} + \frac{1}{2} \sum_{j=1}^{m_i} \frac{(n_{ij} - 1)}{1 + (n_{ij} - 1)\phi_i} = 0.
\]
Appendix C

To show that

\[
E\left(-\frac{\partial^2 Q^+}{\partial \lambda_2^2}\bigg|_{r_0}\right) = \frac{1}{2} \sum_i \sum_j \frac{(n_{ij} - 1)^2}{\{1 + (n_{ij} - 1)\phi_i\}^2}.
\]

From section 3.4.2, we have

\[
\left(-\frac{\partial^2 Q^+}{\partial \lambda_2^2}\bigg|_{r_0}\right) = \frac{1}{2} \sum_i \sum_j \frac{(n_{ij} - 1)^2}{\{1 + (n_{ij} - 1)\phi_i\}^2} \left[ y_{ij} \log(y_{ij}) - y_{ij} \log(n_{ij}) \right] + (n_{ij} - y_{ij}) \log(n_{ij} - y_{ij}) - (n_{ij} - y_{ij}) \log[n_{ij}(1 - \pi)]
\]

Using the Taylor series expansion as in Appendix B, then,

\[
E\left(-\frac{\partial^2 Q^+}{\partial \lambda^2_2}\bigg|_{r_0}\right) = \frac{1}{2} \sum_i \sum_j \frac{(n_{ij} - 1)^2}{\{1 + (n_{ij} - 1)\phi_i\}^2} \left[ n_{ij} \pi \log(n_{ij}) - n_{ij} \pi \log(n_{ij}) + \frac{(1 - \pi)\{1 + (n_{ij} - 1)\phi\}}{2} + n_{ij}(1 - \pi) \log[n_{ij}(1 - \pi)] \right] - n_{ij}(1 - \pi) \log[n_{ij}(1 - \pi)] + \frac{\pi\{1 + (n_{ij} - 1)\phi\}}{2}
\]

After simplification we get

\[
E\left(-\frac{\partial^2 Q^+}{\partial \lambda^2_2}\bigg|_{r_0}\right) = \frac{1}{2} \sum_i \sum_j \frac{(n_{ij} - 1)^2}{\{1 + (n_{ij} - 1)\phi_i\}^2} \left[ \frac{1 + (n_{ij} - 1)\phi}{2} \right]
\]

\[
= \frac{1}{2} \sum_i \sum_j \frac{(n_{ij} - 1)^2}{\{1 + (n_{ij} - 1)\phi_i\}^2}.
\]
REFERENCES


Breslow, N.E. (1990(b)). Further studies in the variability of poick counts. Statistics in Medicine. 9, 615-626.


Weil, C.S. (1970). Selection of the valid number of sampling units and a consideration of their combination in toxicological studies involving reproduction, teratogenesis or carcinogenesis reproduction, teratogenesis. *Food and Cosmetics Toxicology*. 8, 177-182.


