2014

Behrens-Fisher Analogs for Discrete and Survival Data

S.M. Khurshid Alam
University of Windsor

Behrens-Fisher Analogs for Discrete and Survival Data

by

S. M. Khurshid Alam

A Dissertation
Submitted to the Faculty of Graduate Studies
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

Behrens-Fisher Analogs for Discrete and Survival Data

by

S. M. Khurshid Alam

APPROVED BY:

Dr. J. Koval, External Examiner
University of Western Ontario

Dr. Y. Aneja
Odette School of Business

Dr. M. Hlynka
Department of Mathematics and Statistics

Dr. A. Hussein
Department of Mathematics and Statistics

Dr. S. R. Paul, Advisor
Department of Mathematics and Statistics

13 June 2014

Author's Declaration of Originality

I hereby declare that this dissertation incorporates the outcome of joint research undertaken in collaboration with my supervisor Professor S. R. Paul. I also declare that no part of this dissertation has been published.

I certify that, to the best of my knowledge, my dissertation does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my dissertation, published or otherwise, are fully acknowledged in accordance with standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained written permission from the copyright owner(s) to include such material(s) in my dissertation and have included copies of such copyright clearances to the appendix.

I declare that this is a true copy of my dissertation, including any final revisions, as approved by my dissertation committee and the Graduate Studies Office, and that this dissertation has not been submitted for a higher degree to any other University or Institution.

Abstract

Discrete data often exhibit variation greater or smaller than predicted by a simple model. Negative binomial distribution and beta-binomial distribution are popular and widely used to accommodate the extra-Poisson and extra-binomial variations respectively in analyzing discrete data. Weibull distribution is one of the most popular distributions in survival data analysis. Often both discrete and survival data appear in groups and it may be of interest to compare certain characteristics of two groups of such data. The purpose of this dissertation is to deal with Behrens-Fisher analogs for data that follow negative binomial, beta-binomial and Weibull distributions.

We first develop six test procedures, namely, LR, $LR\,(bc)$, T^2, $T^2\,(bc)$, T_1 and T_N, for testing the equality of two negative binomial means assuming unequal dispersion parameters. A simulation study is conducted to compare the performance of the test procedures. Two sets of data are analyzed. For small to moderate sample sizes, the statistic T_1 shows best overall performance. For large sample sizes, all six statistics perform well and are found similar in terms of maintaining size and power.

We, then, develop eight test procedures, namely, LR, C_{ml}, C_{kmm}, C_{qb}, C_{qs}, C_{eq}, C_{rs} and C_{ars}, for testing the equality of proportions in two beta-binomial distributions where the dispersion parameters are assumed unknown and unequal. These test procedures are compared through simulation studies and data analysis. The LR test is observed to maintain the nominal level reasonably well accompanied with the best power performance. The next best is the performance of the statistic C_{eq} in terms of nominal level and power.

Last but not least, we develop four test procedures, namely, LR, C_{ml}, C_{cr} and C_{tg}, for testing the equality of scale parameters of two Weibull distributions where the shape parameters are unequal and compare these statistics through simulation studies and data analysis. For small sample sizes, the statistics LR and C_{ml} hold nominal level most effectively. The statistic C_{cr} shows highest power although its level is also higher (liberal). For moderate and large sample sizes the overall performance of the statistic LR is found to be superior to others.

Dedication

Dedicated to the Memory of My Grandmother.

Acknowledgements

All praises and glorifications are to Almighty Allah for all His blessings and divine gifts upon me and for giving me the opportunity and ability to carry out this research work.

I wish to express my sincere indebtedness and deepest sense of gratitude to my esteemed supervisor, Professor S. R. Paul for his supervision, guidance, encouragement and financial support from his NSERC grant throughout the preparation of this dissertation. The stimulation that I obtained from his spirit of academic research will guide me in my future career.

I am very grateful for having an exceptional doctoral committee and wish to thank Dr. M. Hlynka, Dr. A. Hussein and Dr. Y. Aneja for their critical review and constructive suggestions and comments which improved this dissertation. I would also like to express my gratitude to my external examiner Dr. J. Koval, Department of Epidemiology and Biostatistics, University of Western Ontario for his valuable suggestions and constructive criticisms. In addition, I extend my sincere gratitude to Dr. K. Y. Fung for her encouragement.

I would like to thank the Department of Mathematics and Statistics for providing me with the financial support in the form of Graduate Assistantships, Ontario Ministry of Training and Colleges for the Ontario Graduate Scholarship (OGS).

I must also thank my parents and younger brother and sisters for their support, patience, love and prayers that helped me reach where I am today. All words of gratitude are unable to acknowledge their real contribution. Finally, a very special thanks goes to my wife Shaheena Sultana for her all out support and encouragement.

S. M. Khurshid Alam

June 13, 2014

Windsor, Ontario, Canada.

Contents

List of Tables

Chapter 1

Introduction

Count data arise in many biological (Anscombe (1949); Bliss and Fisher (1953); Bliss and Owen (1958); McCaughran and Arnold (1976); Margolin et al. (1981); Ross and Preece (1985)) and epidemiological (Manton et al. (1981)) investigations. The Poisson distribution, which is characterized by the fact that the variance is equal to the mean of the distribution, is a very widely used and popular model in analysing count data. However, in practice, it often occurs that the variance of the data exceeds that which would be normally expected. This phenomenon of over-dispersion in count data is quite common in practice. For example, Bliss and Fisher (1953) present a set of count data consisting of the number of European red mites on apple leaves for which the mean and the variance are 1.2467 and 2.2737 respectively showing that the variance exceeds the mean. In order to properly accommodate this over-dispersion in the count data, a popular and convenient model, as an alternative to the Poisson distribution, is the negative binomial distribution (see Engel (1984); Breslow (1984); Margolin et al. (1989); Lawless (1987); Manton et al. (1981)). For more references on applications of the negative binomial distribution see Clark and Perry (1989). Different authors have used different parametrizations for the negative binomial distribution (see, for example, Paul and Plackett (1978); Barnwal and Paul (1988); Paul and Banerjee (1998); Piegorsch (1990)).

Scientists in various areas, for example, toxicology, teratology (Weil (1970); Kleinman (1973); Williams (1975); Paul (1982)) and other similar fields (Crowder (1978); Otake and Prentice (1984); Donovan et al. (1994); Gibson and Austin (1996)), frequently encounter data in the form of proportions. Binomial model is a basic model to deal with the data of such kind. It happens quite often that the proportion data exhibit lesser or greater variability than predicted by the simple binomial model and the reason for this variability depends on the form of study. Weil (1970) observes that if the experimental units of the data are litters of animals then 'litter effect', that is, the tendency of animals in the same litter to respond more similarly than animals from different litters contribute to greater variability than predicted by simple model. This effect of litter is known as 'heritability of a dichotomous trait' (see Elston (1977); Crowder (1982)) or intra-litter or intra-class correlation. Apart from this, in some situations the intra-class correlation provides an index of disease aggregation (see Ridout et al. (1999)). A number of parametric and semi-parametric models have been used to incorporate the extra-binomial variation in analyzing data in the form of proportions. Among the parametric models are beta-binomial model (Skellam (1948)), the correlated binomial model (Haseman and Kupper (1979)), and the additive and multiplicative binomial models (Altham (1978)). The correlated and the additive binomial models are identical and give a first order approximation of the beta-binomial model (Srivastava and Wu (1993)). The beta-binomial is a commonly used parametric model because it is easy to use, flexible and extends readily to more complex models with extraneous variance as a function of covariates (Chen and Kodell (1989)). Many authors (Paul (1982); Pack (1986b,a)) have shown the superiority of the beta-binomial model for analyzing proportions and model fitting. In model fitting the superiority of the beta-binomial model over the others is that the beta-binomial model is more sensitive for departure from the binomial model and the likelihood under this model is the easiest to maximize. The beta-binomial model is preferred to the others in analyzing data sets because it is easier to use. This model has been employed in analyzing consumer purchasing behaviour (Chatfield and Goodhart (1970)), in studies of dental caries in children (Weil (1970)) and in toxicological data (Williams (1975)). With all these advantages, the beta-binomial model, however, takes into account only the positive correlation or over-dispersion between littermates. But in practice, the littermates may compete with each other and as a result, negative correlation among the littermates may occur. To overcome this drawback of the beta-binomial model, Prentice (1986) proposed an extension of the beta-binomial model which allows both positive and negative correlations (or over and under dispersions) among the binary variates corresponding to the littermates, which is known as extended beta-binomial model. For analyzing data, the extended beta-binomial model with over and under dispersion has been used by many authors (Otake and Prentice (1984); Prentice (1986); Paul and Islam (1998)).

An alternative approach to accommodate the over/under dispersion is the use of semi-parametric models and the method of moments, which are known to be robust to misspecification of the variance structure. For binomial data with over/under dispersion several semi-parametric models have been proposed that require the assumption of only the first two moments. 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

Carroll (1987)), the double extended quasi-likelihood (Lee and Nelder (2001)) and others based on optimal quadratic estimating equations (Crowder (1987); Godambe and Thompson (1989)). For the estimation of the mean and the dispersion parameters, several estimators based on the method of moments have been proposed by Breslow (1990); Kleinman (1973); Moore (1986) and Srivastava and Wu (1993).

In many biomedical applications the primary endpoint of interest is the survival time or time to a certain event like time to death, time it takes for a patient to respond to a therapy, time from response until disease relapse (i.e., disease returns) etc. The importance of parametric models in the analysis of life time data is well known (Lawless (1982); Nelson (1982); Mann et al. (1974)). The use of Weibull distribution to describe the real phenomena has a long and rich history. The Weibull distribution has been considered to be a successful model for many product failure mechanisms. However, Lloyd (1967); Ku et al. (1972); Hammitt (2004) and McCool (1998), among others, have extended the use of the Weibull distribution to other branches of statistics, such as reliability, risks and quality control.

The traditional Behrens-Fisher problem is to test the equality of the means of two normal populations where the population variances are assumed unknown and possibly unequal. In various biological and epidemiological investigations over-dispersed data in the form of counts can appear in several independent samples where the variances are unequal and larger than the means. Likewise, in toxicological and other similar fields data may arise as the proportion of a certain characteristic in several groups where the intra-class correlations vary from group to group. Also, in life testing, reliability or survival analysis we may encounter data on different component types or cohorts, where assuming the Weibull model, with shape parameters unequal and unknown. Special cases, in all of the above three scenarios, are data in two groups where one group is treated as control, while, the other as the group of interest or treatment group. It may be of interest, in the case of two groups over-dispersed count data situation, to test the equality of two means assuming that the variances are unknown and unequal. Similarly, testing the equality of proportions between two groups considering unequal over-dispersion parameters may be the objective where the data are in the forms of proportions. In addition, for survival data, interest may focus on testing the equality of scale parameters of two Weibull distributions assuming unequal shape parameters. These, in the sense of testing the equality of two mean or scale parameters in the presence of unequal dispersion or shape parameters, can be considered as analogous to the traditional Behrens-Fisher problem.

The main purpose of this thesis is to construct procedures for testing (i) the equality of two mean parameters for over-dispersed count and proportion data with unequal dispersion parameters; (ii) equality of two scale parameters for Weibull distributed survival data with unequal shape parameters and (iii) to study the performance, thought simulation studies, of the test procedures. We compare the performances of the test procedures in terms of empirical size and power.

In Chapter 2, we discuss some preliminaries and review the theory of likelihood ratio and $C(\alpha)$ tests (Neyman (1959)). We also review the estimation procedures for the generalized linear models. In addition, we review the general bias correction results of Cox and Snell (1968) and present the Cordeiro and Klein (1994) formula for the biases of the maximum likelihood estimators in general parametric models.

In Chapter 3, we develop procedures for testing equality of two negative binomial means where the over-dispersion parameters are unequal. The test procedures, namely, the likelihood ratio test, the likelihood ratio test based on the bias corrected maximum likelihood estimates of the parameters, the score test, the score test based on the bias corrected maximum likelihood estimates of the parameters and two tests, similar to Welch's (1937) V statistic, constructed from the $C(\alpha)$ statistic based on the method of moments estimates of the nuisance parameters.

In Chapter 4, we derive parametric and semi-parametric procedures for testing the equality of two proportions in the presence of unequal dispersion parameters. The likelihood ratio test, the $C(\alpha)$ (score) test based on the maximum likelihood estimates of the nuisance parameters and the $C(\alpha)$ test based on Kleinman's (1973) method of moments estimates of the nuisance parameters are the parametric tests. The semi-parametric tests are, a $C(\alpha)$ test based on the quasi-likelihood and the method of moments estimates by Breslow (1990), a $C(\alpha)$ test based on the quasi-likelihood and the method of moments estimates by Srivastava and Wu (1993), a $C(\alpha)$ test based on the extended quasi-likelihood estimates, the Rao-Scott and the adjusted Rao-Scott tests by following Rao and Scott (1992). A simulation study is conducted to compare the relative performance, in terms of size and power, of the test procedures.

In Chapter 5, we deal with the Weibull distributed survival data and develop test procedures for testing the equality of scale parameters of two Weibull distributions where the shape parameters are assumed unequal. We develop a likelihood ratio test and three $C(\alpha)$ tests. Of the $C(\alpha)$ tests, one is based on the maximum likelihood estimates of the nuisance parameters, one is based on the method of moments estimates of the nuisance parameters by Cran (1988) and the other is based on the method of moments estimates of the nuisance parameters by Teimouri and Gupta (2013). A small scale simulation study is conducted to compare, on the basis of empirical size and power, the test procedures.

Finally, conclusions of the thesis with the summary of findings and a discussion of some topics for future study are presented in Chapter 6.

Chapter 2

Some Preliminaries and Review of Current Literature

2.1 Behrens-Fisher Problem

In hypothesis testing a frequently encountered problem is to test the equality of two population means and the test statistic varies according to the nature of the populations involved. A widely used assumption regarding the distribution of populations is normality. Different scenarios appear depending on the sample sizes and assumptions on the variances of the populations. The Behrens-Fisher (BF) problem arises when testing the equality of the means of two normal populations where both variances are unknown and possibly unequal. This problem has been well known since it was discovered by Behrens (1929). Fisher (1935) used the fiducial theory of statistical inference to justify Behrens' solution. A number of solutions, both parametric and non-parametric, have been proposed for the Behrens-Fisher problem (see for example, Fisher (1935, 1941); Scheffé (1943); Welch (1947); Aspin (1948); Cochran and Cox (1950); Qin (1991); Paul (1992); Dong (2004) and Tsui and Tang (2005)).

2.2 $\sqrt{n}$ Consistent Estimators

Definition: Consider $\left\{\hat{\theta}_n\right\}, n = 1, 2, \cdots$, a sequence of estimators. The sequence of estimates $\hat{\theta}_n$ is called $\sqrt{n}$ consistent if the quantity $|\hat{\theta}_n - \theta|\sqrt{n}$ remains bounded in probability as $n \to \infty$ (see Lehman (1999)).

Theorem: Let $\hat{\theta}_n$ be a sequence of estimates of θ, and $var(\hat{\theta}_n) = \mathrm{O}(\frac{1}{n})$. Then this sequence of estimates is $\sqrt{n}$ consistent.

Proof: According to Chebyshev's inequality, for a given $\epsilon > 0$,
$P\left(\left|\hat{\theta}_n - \theta\right|\sqrt{n} \leq \epsilon\right) \geq 1 - \dfrac{var(\hat{\theta}_n)n}{\epsilon^2}$.
If $\hat{\theta}_n$ is a sequence of maximum likelihood estimates, then according to the asymptotic properties of maximum likelihood estimates, the distribution of $\hat{\theta}_n$ is normal with mean θ and variance $\dfrac{1}{nI(\theta)}$, where $I(\theta)$ is the Fisher information matrix defined as $I(\theta) = E\left[\dfrac{\partial}{\partial\theta} f(y|\theta)\right]^2$, where $f(y|\theta)$ is the probability density function of Y. This means that $var(\hat{\theta}_n)$ tends to zero and n tends to infinity, that is, $var(\hat{\theta}_n)$ is $\mathrm{O}(n^{-1})$. This means that MLE is $\sqrt{n}$ consistent. The method of moments estimators are also $\sqrt{n}$ consistent estimates (Moore (1986)).

2.3 Likelihood Ratio Test

Let $Y_1, Y_2, \cdots, Y_n$ be a random sample of size n from a distribution that has probability density function $f(y|\lambda)$, where $\lambda = (\theta, \phi)' = (\theta_1, \theta_2, \cdots, \theta_k, \phi_1, \phi_2, \cdots, \phi_s)'$ is a $k+s$ component vector. The likelihood function of the parameter λ for the given data is given by $L(\lambda|Y_1, Y_2, \cdots, Y_n) = \prod_{i=1}^{n} f(y_i|\lambda)$. Our interest is to test the null hypothesis $H_0 : \theta = \theta_0 = (\theta_{10}, \theta_{20}, \cdots, \theta_{k0})'$ treating $\phi = (\phi_1, \phi_2, \cdots, \phi_s)'$ as a nuisance parameter. The likelihood ratio test is based on the ratio of the maximized likelihood function under the null hypothesis to that under the alternative hypothesis and is defined as

$$\Lambda = \frac{L\left(Y_1, Y_2, \cdots, Y_n, \theta_0, \hat{\phi}\right)}{L\left(Y_1, Y_2, \cdots, Y_n, \tilde{\theta}, \tilde{\phi}\right)}.$$

Under the null hypothesis, H_0, the quantity $LR = -2\ln\Lambda = 2(l_1 - l_0)$ is distributed as chi-squared with k degrees of freedom if the sample size n is large. Here l_0 is the maximized log-likelihood function under the null hypothesis and l_1 is the maximized log-likelihood function under the alternative hypothesis.

2.4 $C(\alpha)$ Test

The $C(\alpha)$ test is based on the partial derivatives of the log-likelihood function with respect to the nuisance parameters and the parameters of interest evaluated at the null hypothesis. Let l be the log-likelihood for the data. Define the partial derivatives of the log-likelihood which are evaluated at $\theta = \theta_0 = (\theta_{10}, \theta_{20}, \cdots, \theta_{k0})'$ as

$$\psi = \frac{\partial l}{\partial \theta}\Big|_{\theta=\theta_0} = \left[\frac{\partial l}{\partial \theta_1}, \frac{\partial l}{\partial \theta_2}, \cdots, \frac{\partial l}{\partial \theta_k}\right]'\Bigg|_{\theta=\theta_0}$$

and

$$\gamma = \frac{\partial l}{\partial \phi}\Big|_{\theta=\theta_0} = \left[\frac{\partial l}{\partial \phi_1}, \frac{\partial l}{\partial \phi_2}, \cdots, \frac{\partial l}{\partial \phi_s}\right]'\Bigg|_{\theta=\theta_0}.$$

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

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

is the information matrix with elements

$$I_{11} = E\left(-\frac{\partial^2 l}{\partial \theta \partial \theta'}\Big|_{\theta=\theta_0}\right),\ I_{12} = E\left(-\frac{\partial^2 l}{\partial \theta \partial \phi'}\Big|_{\theta=\theta_0}\right),\ \text{and}\ I_{22} = E\left(-\frac{\partial^2 l}{\partial \phi \partial \phi'}\Big|_{\theta=\theta_0}\right)$$

which are $(k \times k)$, $(k \times s)$ and $(s \times s)$ matrices respectively. The $C(\alpha)$ test is based on the adjusted score $S = \frac{\partial l}{\partial \theta} - B\frac{\partial l}{\partial \phi}$, where B is the matrix of partial regression coefficients that is obtained by regressing $\frac{\partial l}{\partial \theta}$ on $\frac{\partial l}{\partial \phi}$. Bartlett (1953a) showed that $B = I_{12}I_{22}^{-1}$ and the variance-covariance matrix of S is $I_{11.2} = I_{11} - I_{12}I_{22}^{-1}I_{21}$. Thus the distribution of the adjusted score, S is multivariate normal with mean vector 0 and variance covariance matrix $I_{11.2}$, that is, $S \sim MN(0,\ I_{11.2})$. Following Neyman (1959) the distribution of $S'I_{11.2}^{-1}S$ is chi-squared with k degrees of freedom. As this statistic involves nuisance parameters $\phi = (\phi_1, \phi_2, \cdots, \phi_s)'$, it is necessary to replace them by some suitable estimates for testing the null hypothesis. Moran (1970) suggested the use of some $\sqrt{n}$ consistent estimators of ϕ that are obtained from the data. Let $\tilde{\phi} = (\tilde{\phi}_1, \tilde{\phi}_2, \cdots, \tilde{\phi}_s)'$ be $\sqrt{n}$ consistent estimators of the parameters $\phi = (\phi_1, \phi_2, \cdots, \phi_s)'$. Hence following Neyman (1959) and replacing the nuisance parameters by some $\sqrt{n}$ consistent estimators, the test statistic becomes $\chi^2_{C(\alpha)} = \tilde{S}'\tilde{I}^{-1}_{11.2}\tilde{S}$, which is asymptotically distributed as chi-squared with k degrees of freedom. It is to be noted that if the nuisance parameter ϕ is replaced by its maximum likelihood estimate, $\hat{\phi}$, then the adjusted score function reduces to $S_i = \psi_i,\ i = 1,\ 2, \cdots, k-1$. The form of the $C(\alpha)$ statistic becomes $\hat{\psi}'\hat{I}^{-1}_{11.2}\hat{\psi}$, which is a score test (Rao (1948)). Again, the score test is asymptotically equivalent to the likelihood ratio test (Moran (1970); Cox and Hinkley (1974)).

2.5 Lindeberg Central Limit Theorem

The Central Limit Theorem: Let $X_1, X_2, \cdots, X_n$ denote the observations of a random sample from a distribution that has mean μ and positive variance σ^2. Then the random variable $Y_n = \dfrac{\left(\sum_{i=1}^{n} X_i - n\mu\right)}{\sqrt{n}\sigma} = \sqrt{n}\left(\bar{X}_n - \mu\right) / \sigma$ converges in distribution to a random variable which has a normal distribution with mean zero and variance 1 (Hogg et al. (2005)).

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

$$U_j = Z_j - \mu_j, \text{ if } \mid Z_j - \mu_j \mid \leq \epsilon s_m$$

$$U_j = 0, \text{ if } \mid Z_j - \mu_j \mid > \epsilon s_m$$

satisfy the condition $s_m \to \infty$ and $\frac{1}{s_m^2}\sum_{j=1}^{m} E\left(U_j^2\right) \to 1$, where $\mu_j = E\left(Z_j\right)$, $\sigma_j^2 = var\left(Z_j\right)$ and $s_m^2 = \sum_{j=1}^{m}\sigma_j^2$. This implies that every uniformly bounded sequence of mutually independent random variables $\{Z_j\}$ obeys the central limit theorem, provided that $s_m \to \infty$.

Theorem: Let $U_j\left(\beta,\, \phi\right),\ j = 1,\ 2, \cdots, k$, represent an unbiased estimating function for β_j and let $U_{k+1}\left(\beta,\, \phi\right)$ represent an unbiased estimating function for ϕ. Then, by the Lindeberg central limit theorem

$$\frac{1}{\sqrt{m}}\begin{pmatrix} U_j \\ U_{k+1}\end{pmatrix} \xrightarrow{d} N_{k+1}\left(0,\ \sum_0\right)$$

where $\sum_0$ denotes the limiting variance-covariance matrix of U_j and U_{k+1}. The details of the proof of this theorem are given in Moore (1985).

2.6 Generalized Linear Models

A generalized linear model (GLM) is the generalization of ordinary linear regression models to accommodate non-normal response variables. It consists of three components, namely, a random component, a systematic component, and a link function and can be described as follows

i) Suppose that the joint probability function of the response variable Y with mean μ can be written in the form

$$f(y;\theta) = exp\left[\phi^{-1}\left\{y\theta - b(\theta)\right\} + C(y,\phi)\right].$$

for some known functions $b(.)$, $C(.)$, canonical or natural parameter θ, and constant ϕ which may be known or a parameter to be estimated. This is said to be in canonical form.

ii) The systematic component relates a set of explanatory variables with a linear predictor in the form

$$\eta = \sum_{j=1}^{p} X_j \beta_j.$$

iii) The link function is a monotone differentiable function of the mean that connects the random and the systematic components. The model links the mean μ to the linear predictor η by $\eta = g(\mu)$, where the link function $g(.)$ is a monotone differentiable function.

The mean and the variance of Y are $E(Y) = b'(\theta)$ and $var(Y) = \phi b''(\theta)$ respectively (McCullagh and Nelder (1989, p. 29)).

2.7 Quasi-Likelihood

In numerous applications, the full distributional assumptions of the GLM are not always satisfied. To overcome the situation where the full distributional assumptions are not met, Wedderburn (1974) proposes the quasi-(log)likelihood, henceforth referred to as the quasi-likelihood (QL) model which takes into account only the first two moments of the random variable Y. The QL possesses properties similar to those of the log-likelihood and can be used for inferential purposes. Suppose $y_1, y_2, \cdots, y_n$ is a random sample of size n with mean and variance of the i^{th} observation $\mu = E(y_i)$ and $var(y_i) = \phi V(\mu)$ respectively, where V is some known function called the variance function and ϕ is a known constant or a parameter to be estimated. Then the quasi-likelihood of the data is defined as

$$Q(y;\mu) = -\frac{1}{2}\sum_{i=1}^{n}\frac{d_i(y;\mu)}{\phi}, \quad \text{where} \quad d_i(y;\mu) = 2\int_{\mu}^{y_i}\frac{(y_i - t)}{V(t)}dt$$

is the discrepancy between the observation and its mean and is called the deviance component of i^{th} observation. In GLM framework the parameter ϕ can be treated as dispersion parameter and its moments estimate can be obtained. As the quasi-likelihood is designed only for the estimation of the mean parameter, a maximum quasi-likelihood estimate of the dispersion parameter (ϕ) cannot be found. The value of ϕ is 1 for the binomial and the Poisson distributions. For the Poisson variable, the variance function is $V(m) = m$ where $E(y_i) = m$ and for the binomial variable, the variance function is $V(\pi) = n\pi(1-\pi)$ where $E(y_i) = n\pi$.

2.8 Extended Quasi-Likelihood

Though the quasi-likelihood facilitates the estimation of the mean (regression) parameter, it is not suitable for the estimation of the dispersion parameter. To implement the joint estimation of the mean parameter and the dispersion parameter from the same function, Nelder and Pregibon (1987) and Godambe and Thompson (1989) initiated the extended quasi-likelihood function (EQL) by adding a normalizing factor with the quasi-likelihood. This normalizing factor includes the dispersion parameter and the variance function. Taking into account the normalizing factor, the EQL is defined as

$$Q^{+}(y;\mu,\phi) = -\frac{1}{2}\sum_{i=1}^{n}\left[\frac{d_i(y;\mu)}{\phi} + \ln\{2\pi\phi V(y_i)\}\right],$$

where, the deviance function, $d_i(y;\mu)$, is defined previously, ϕ is the dispersion parameter and $V(y_i)$ is the variance function for the i^{th} observation. The product of the dispersion parameter ϕ and the variance function V gives the variance of datum. For example, for negative binomial distribution which is the over-dispersed Poisson distribution $var(y_i) = \mu(1 + c\mu) = \phi V(\mu)$, where $\phi = 1$ and $V(\mu) = \mu(1 + c\mu)$. For the beta binomial distribution which is the over-dispersed binomial distribution, $var(y_i) = n_i\pi(1-\pi)\{1 + (n_i - 1)\theta\} = \phi_i V(\pi)$, where $\phi_i = 1 + (n_i - 1)\theta$ is the dispersion parameter for i^{th} observation and $V(\pi) = n_i\pi(1-\pi)$. For over-dispersed binomial model the dispersion parameter θ and for over-dispersed Poisson model the dispersion parameter c can take positive as well as negative values. The extended quasi-likelihood, like the quasi-likelihood does not need a full distributional form, rather it only needs the form of the first two moments. The extended quasi-likelihood function possesses properties similar to log-likelihood function but the estimate of the over-dispersion parameter obtained by using Q^{+} can be robust to the maximum likelihood estimate.

2.9 Bias to Order n^{-1} of the Maximum Likelihood Estimator

For a single independent and identically distributed random sample with one unknown parameter, Bartlett (1953b) gave a simple expression for the bias to order $O(\frac{1}{n})$ of the maximum likelihood estimate (MLE). For the parameters of any distribution, Cox and Snell (1968) derived general results for the biases to the order n^{-1} of the maximum likelihood estimators. First they derived an expression for the bias of the MLE for distributions with a single unknown parameter and afterwards they extended the result for multi-parameter case. Cordeiro and Klein (1994) derived a modified expression of bias of order n^{-1} which is easier to implement. Below we explain the two cases, namely, (i) a single parameter n^{-1} bias of the maximum likelihood estimator and (ii) multi-parameter n^{-1} bias of the maximum likelihood estimator and present the Cordeiro and Klein (1994) version of formula for bias of multi-parameter distribution.

2.9.1 A Single Parameter n^{-1} Bias of the Maximum Likelihood Estimator

Let Y_i, $i = 1, 2, \cdots, n$ be random variables with probability density function $f(y_i, \theta)$, depending on unknown parameter θ. Then the log-likelihood function for data is

$$l(\theta) = \sum_{i=1}^{n} l_i(\theta),$$

where, $l_i(\theta) = \log\{f(y_i, \theta)\}$. The maximum likelihood estimating equation for θ is

$$l'(\hat{\theta}) = 0,$$

where, $\hat{\theta}$ is the maximum likelihood estimate of θ. The first order Taylor series expansion of $l'(\hat{\theta})$ at $\hat{\theta} = \theta$ is

$$l'(\theta) + (\hat{\theta} - \theta)l''(\theta) = 0.$$

Let us write, $U_i = \dfrac{\partial l_i}{\partial \theta}$, $V_i = \dfrac{\partial^2 l_i}{\partial \theta^2}$. Now, replacing $-l''(\theta)$ by its expectation gives the Fisher information I in the sample and writing $U = \sum_{i=1}^{n} U_i$, we obtain the standard first order expressions

$$\hat{\theta} - \theta \approx \frac{U}{I} \tag{2.1}$$

and

$$var\left(\hat{\theta}\right) = I^{-1}.$$

To attain an improved result using the second order Taylor series expansion of $l'(\hat{\theta})$ at θ we obtain (see Cox and Snell (1968)),

$$l'(\theta) + (\hat{\theta} - \theta)l''(\theta) + \frac{1}{2}(\hat{\theta} - \theta)^2 l'''(\theta) = 0.$$

Now, from the properties of the log-likelihood function and expected values of random variables we have

$$E(l'(\theta)) = 0,$$

$$E\left[(\hat{\theta}-\theta)l''(\theta)\right] = cov\left[(\hat{\theta}-\theta), l''(\theta)\right] + E(\hat{\theta}-\theta)E(l''(\theta))$$

and

$$E\left[(\hat{\theta}-\theta)^2 l'''(\theta)\right] = cov\left[(\hat{\theta}-\theta)^2, l'''(\theta)\right] + E\left[(\hat{\theta}-\theta)^2\right]E(l'''(\theta)).$$

By taking expectation on both sides of the second order Taylor series expansion and using the results obtained above we get

$$\begin{aligned} &E(\hat{\theta}-\theta)E(l''(\theta)) + cov\left[(\hat{\theta}-\theta),\ l''(\theta)\right] + \frac{1}{2}E\left[(\hat{\theta}-\theta)^2\right]E(l'''(\theta)) \\ &+ \frac{1}{2}cov\left[(\hat{\theta}-\theta)^2,\ l'''(\theta)\right] = 0. \end{aligned} \tag{2.2}$$

Now,

$$cov\left[(\hat{\theta}-\theta),\ l''(\theta)\right] \approx cov\left[\frac{U}{I},\ V\right] = \frac{1}{I}cov\left[\sum_{i=1}^{n} U_i,\ \sum_{i=1}^{n} V_i\right]$$

$$= \frac{1}{I}\sum_{i=1}^{n} cov(U_i,\ V_i), \quad \text{as} \quad cov(U_i,\ V_j) = 0 \quad \text{for} \quad i \neq j.$$

Thus, $cov\left[(\hat{\theta}-\theta),\ l''(\theta)\right] = \frac{1}{I}\sum_{i=1}^{n} E\left(U_i V_i\right) = \frac{J}{I},$

where $J = \sum_{i=1}^{n} E\left(U_i V_i\right),$ as $E\left(U_i\right) = E\left(\frac{\partial l_i}{\partial \theta}\right) = 0.$

Similarly,

$$cov\left[(\hat{\theta}-\theta)^2,\, l'''(\theta)\right] \approx cov\left[\frac{U^2}{I^2},\, W\right] = \frac{1}{I^2} cov\left[U^2,\, W\right]$$

$$= \frac{1}{I^2} cov\left[\left(\sum_{i=1}^{n} U_i\right)^2,\, \sum_{i=1}^{n} W_i\right]$$

$$= \frac{1}{I^2} \sum_{i=1}^{n} cov\left[U_i^2,\, W_i\right], \quad \text{as} \quad cov\left(U_i^2,\, W_j\right) = 0, \quad \text{for} \;\; i \neq j.$$

$$\text{Thus,} \;\; cov\left[(\hat{\theta}-\theta)^2,\, l'''(\theta)\right] = \frac{1}{I^2} \sum_{i=1}^{n} \left[E\left(U_i^2 W_i\right) - E\left(U_i^2\right) E\left(W_i\right)\right] = \frac{L}{I^2},$$

$$\text{where} \;\; L = \sum_{i=1}^{n} \left[E\left(U_i^2 W_i\right) - E\left(U_i^2\right) E\left(W_i\right)\right], \;\; W = \sum_{i=1}^{n} W_i = l'''(\theta) \;\; \text{and} \;\; W_i = \frac{\partial^3 l_i(\theta)}{\partial \theta^3}.$$

Since the terms I and L refer to a total over the sample, they are of order n which implies that $cov\left[(\hat{\theta}-\theta)^2,\, l'''(\theta)\right]$ is $O\left(n^{-1}\right)$. Finally, based on the results obtained above, the equation (2.2) can be written as

$-IE\left(\hat{\theta}-\theta\right) + \frac{J}{I} + \frac{1}{2} var\left(\hat{\theta}\right) K + O\left(n^{-1}\right) = 0$, where $K = E\left(l'''(\theta)\right) = E\left(\sum_{i=1}^{n} W_i\right)$. Thus, the bias to order n^{-1} of the maximum likelihood estimator of θ is

$$b\left(\hat{\theta}\right) = E\left(\hat{\theta}-\theta\right) = \frac{J}{I^2} + \frac{1}{2I} var\left(\hat{\theta}\right) K = \frac{J}{I^2} + \frac{1}{2I^2} K = \frac{1}{2I^2}\left(K + 2J\right),$$

as $var\left(\hat{\theta}\right) = I^{-1}$.

2.9.2 Multi-parameter n^{-1} Bias of the Maximum Likelihood Estimator

Let the p parameter random variable $Y_i,\ i = 1,\ 2, \cdots, n$ have probability density function $f(y_i; \theta_1, \theta_2, \cdots, \theta_p)$ and also let $\theta' = (\theta_1, \theta_2, \cdots, \theta_p)$. For $r, t, u = 1, 2, \cdots, p$ define

$$I_{rt} = E\left(-\sum_{i=1}^{n} V_{rt}^{(i)}\right), \quad J_{rtu} = E\left(-\sum_{i=1}^{n} W_{rtu}^{(i)}\right) \quad \text{and}$$

$$K_{r,tu} = E\left(-\sum_{i=1}^{n} U_r^{(i)} V_{tu}^{(i)}\right),$$

$$\text{where} \quad U_r^{(i)} = \frac{\partial l_i}{\partial \theta_r}, \quad V_{rt}^{(i)} = \frac{\partial^2 l_i}{\partial \theta_r \partial \theta_t} \quad \text{and} \quad W_{rtu}^{(i)} = \frac{\partial^3 l_i}{\partial \theta_r \partial \theta_t \partial \theta_u},$$

$$\text{with} \quad l_i = \log\left\{f(y_i; \theta_1, \theta_2, \cdots, \theta_p)\right\}.$$

As in equation (2.1), using the first order Taylor series expansion of the maximum likelihood estimating equations for $\theta_r, r = 1, 2, \cdots, p$, that is, $\left.\frac{\partial l}{\partial \theta_r}\right|_{\theta=\hat{\theta}}$, we obtain

$$\hat{\theta}_r - \theta_r = \sum_{i=1}^{n} M^{rs} U_s^{(i)}, \tag{2.3}$$

where M^{rs} is the $(r, s)^{th}$ element of the inverse of the informations matrix $I = (I_{rs})$ of order p and the summation convention is applied to multiple suffixes referring to parameter components. Again, the estimating equations for the second order Taylor series expansion of $\theta_s, s = 1, 2, \cdots, p$ are

$$\sum_{i=1}^{n}\left[U_r^{(i)} + \left(\hat{\theta}_s - \theta_s\right)V_{rs}^{(i)} + \frac{1}{2}\left(\hat{\theta}_t - \theta_t\right)\left(\hat{\theta}_u - \theta_u\right)W_{rtu}^{(i)}\right] = 0 \tag{2.4}$$

Now, on taking expectation of (2.4) and using (2.3) we obtain a set of simultaneous linear equations as

$$E\left(\hat{\theta}_s - \theta_s\right)I_{rs} = \frac{1}{2}M^{tu}\left(J_{rtu} + 2K_{t,ru}\right), \quad \text{for} \quad s = 1, 2, \cdots, p.$$

Finally, as a solution to the above set of equations, the biases to order n^{-1} of the maximum likelihood estimates of $\hat{\theta}_s,\ s = 1,\ 2,\ \cdots,\ p$ are

$$b_s\left(\hat{\theta}_1,\ \hat{\theta}_2, \cdots, \hat{\theta}_p\right) = E\left(\hat{\theta}_s - \theta_s\right) = \frac{1}{2}\sum_{r=1}^{p}\sum_{s=1}^{p}\sum_{t=1}^{p}M^{rs}M^{tu}\left(J_{rtu} + 2K_{t,ru}\right). \tag{2.5}$$

Cordeiro and Klein (1994) suggested an alternative form of (2.5) for biases of $\hat{\theta}_s,\ s = 1,\ 2,\ \cdots,\ p$ which is

$$b_s\left(\hat{\theta}_1,\ \hat{\theta}_2, \cdots, \hat{\theta}_p\right) = \sum_{r=1}^{p}M^{rs}\sum_{t=1}^{p}\sum_{u=1}^{p}\left(K_{rt}^{(u)} - \frac{1}{2}J_{rtu}\right)M^{tu}, \tag{2.6}$$

where $K_{rt}^{(u)} = \dfrac{\partial}{\partial\theta_u}E\left(V_{rt}\right);\ r,\ t,\ u\ = 1,\ 2,\ ,\cdots,\ p;\ s = 1,\ 2,\ ,\cdots,\ p.$

Chapter 3

Testing Equality of Two Negative Binomial Means in the Presence of Unequal Over-dispersion Parameters

3.1 Introduction

Count data arise in numerous biological (Anscombe (1949); Bliss and Fisher (1953); Bliss and Owen (1958); McCaughran and Arnold (1976); Margolin et al. (1981); Ross and Preece (1985)) and epidemiological (Manton et al. (1981)) investigations. The analysis of such count data is often based upon the assumption of some form of Poisson model. However, in practice, it often occurs that the variance of the data exceeds that which would be normally expected. This phenomenon of over-dispersion in count data is quite common in practice. For example, Bliss and Fisher (1953, p. 178) present a set of count data consisting of the number of European red mites on apple leaves for which the mean and the variance are 1.1467 and 2.2737, respectively, showing that the variance exceeds the mean. Also, two sample over-dispersed count data occur in practice where the variances are larger than the means and the two variances are unequal. For example, Saha (2011) presents a data set (originally given by Uusipaikka (2009)) on the number of cycles required to get pregnant for two groups of women, namely, smokers and non smokers. Table 3.18 shows the number of cycles required to get pregnant for smoker and non smoker women along with sample statistics. As can be seen, in these data, the sample means as well as the sample variances are unequal and the variance in each group is greater than the mean. A problem in this setting is to test the equality of the means when the two variances are possibly unequal. This problem is analogous to the traditional Behrens-Fisher problem. A lot of work has been done on the traditional Behrens-Fisher problem when data come from two normal populations with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2 are unknown and possibly unequal (see Best and Rayner (1987) ; and Paul (1992)). The purpose of this Chapter is to develop tests for the corresponding problem for over-dispersed count data. We develop six tests and compare these in terms of size and power.

For modelling count data with over-dispersion, a popular and convenient model is the negative binomial distribution (Manton et al. (1981); Saha (2011); Breslow (1984); Engel (1984); Margolin et al. (1989)). For more references on applications of the negative binomial distribution, see Clark and Perry (1989). Different authors have used different parametrizations for the negative binomial distribution (Paul and Plackett (1978); Barnwal and Paul (1988); Paul and Banerjee (1998); Piegorsch (1990)). Let Y be a negative binomial random variable with mean parameter μ and dispersion parameter c. We write $Y \sim NB(\mu, c)$, which has probability mass function

$$f(y|\mu, c) = Pr(Y = y|\mu, c) = \frac{\Gamma(y + c^{-1})}{y!\Gamma(c^{-1})} \left(\frac{c\mu}{1 + c\mu}\right)^y \left(\frac{1}{1 + c\mu}\right)^{c^{-1}}, \tag{3.1}$$

for $y = 0, 1, ..., \mu > 0$. Now, $Var(Y) = \mu(1 + \mu c)$ and $c > -1/\mu$. Since c can take a positive as well as a negative value, it is called a dispersion parameter rather than an over-dispersion parameter, and with this range of c, $f(y|\mu, c)$ is a valid probability function. Obviously, when $c = 0$, the variance of the $NB(\mu, c)$ distribution becomes that of the Poisson (μ) distribution. Moreover, it can be shown that the limiting distribution of the $NB(\mu, c)$ distribution, as $c \to 0$, is the Poisson (μ).

Let $Y_{11}, \ldots, Y_{1n_1}$ be a random sample from the negative binomial distribution $NB(\mu_1, c_1)$ and $Y_{21}, \ldots, Y_{2n_2}$ be a random sample from the negative binomial distribution $NB(\mu_2, c_2)$. Our problem is to test $H_0 : \mu_1 = \mu_2$, where c_1 and c_2 are unspecified. To test this hypothesis we develop six test procedures, namely, a likelihood ratio test LR, a likelihood ratio test based on the bias corrected maximum likelihood estimates of the nuisance parameters $LR(bc)$, a score test T^2, a score test based on the bias corrected estimates of the nuisance parameters $T^2(bc)$, a $C(\alpha)$ test based on the method of moments estimates of the nuisance parameters, T_1 with Welch's (Welch (1937)) degree of freedom correction, and a test T_N using the asymptotic normal distribution of T_1. These procedures are then compared in terms of size and power using simulations.

The test statistics are developed in Section 3.2. A simulation study is conducted in Section 3.3. A few examples illustrating and motivating the use of the procedures developed are given in Section 3.4 and a discussion follows in Section 3.5.

3.2 The Tests

As indicated in Section 3.1, we consider data Y_{ij}, $i = 1, 2$, $j = 1, 2, ..., n_i$, where $Y_{ij} \sim NB(\mu_i, c_i)$. Our interest is to test $H_0 : \mu_1 = \mu_2 = \mu$, where c_1 and c_2 are unspecified against the alternative $H_A : \mu_1 \neq \mu_2$, with c_1 and c_2 unspecified.

3.2.1 The Likelihood Ratio Test

The log-likelihood for the data of the two samples, based on the negative binomial model (3.1), apart from a constant, is

$$\begin{aligned} l = & \sum_{i=1}^{2}\sum_{j=1}^{n_i}\sum_{l=1}^{y_{ij}}\left[\log\left(1+c_i\left(l-1\right)\right)\right] + \sum_{i=1}^{2} y_{i+}\left[\log(\mu_i) - \log\{1+c_i\mu_i\}\right] \\ & -\sum_{i=1}^{2}\frac{n_i}{c_i}\log\{1+c_i\mu_i\}. \end{aligned} \tag{3.2}$$

Let $\hat{l}_0$ and $\hat{l}_1$ be the maximized log-likelihood under the null and the alternative hypothesis, respectively. Then, the likelihood ratio statistic $LR = 2\left(\hat{l}_1 - \hat{l}_0\right)$ has a $\chi^2(1)$ distribution, asymptotically as $n \to \infty$, where $n = n_1 + n_2$. Under H_0, the parameters μ, c_1 and c_2 are estimated by solving the maximum likelihood estimating equations

$$\begin{aligned} & \sum_{i=1}^{2}\frac{n_i(\bar{y}_i-\mu)}{\mu(1+\mu c_i)} = 0 \\ & \text{and} \\ & \frac{n_i}{c_i^2}\log(1+c_i\mu) + \frac{n_i(\bar{y}_i-\mu)}{c_i(1+c_i\mu)} - \sum_{j=1}^{n_i}\sum_{l=0}^{y_{ij}-1}\frac{1}{c_i(1+c_i l)} = 0, \;\; i = 1,\, 2 \end{aligned} \tag{3.3}$$

simultaneously or by directly maximizing (3.2) with respect to μ, c_1 and c_2. Denote these estimators by $\hat{\mu}_0$, $\hat{c}_{10}$ and $\hat{c}_{20}$. Under H_A, the maximum likelihood estimate of μ_i is $\bar{y}_i$, $i = 1,\ 2$. The maximum likelihood estimate of c_i, denoted by $\hat{c}_i$, is obtained by solving the maximum likelihood estimating equation

$$\frac{n_i}{c_i^2}\log(1+c_i\bar{y}_i) - \sum_{j=1}^{n_i}\sum_{l=0}^{y_{ij}-1}\frac{1}{c_i(1+c_il)} = 0,\ i = 1,\ 2.$$

3.2.2 The Likelihood Ratio Test Based on the Bias Corrected Maximum Likelihood Estimates

Let l_0 be the kernel of the log-likelihood under the null hypothesis and let $\theta = (\theta_1,\ \theta_2,\ \theta_3)' = (\mu,\ c_1,\ c_2)'$. Further, let $I_{rt} = E\left(-V_{rt}\right),\ J_{rtu} = E\left(W_{rtu}\right),$ and $K_{rt}^{(u)} = \frac{\partial}{\partial\theta_u}E\left(V_{rt}\right)$, where $V_{rt} = \frac{\partial^2 l_0}{\partial\theta_r\partial\theta_t}$ and $W_{rtu} = \frac{\partial^3 l_0}{\partial\theta_r\partial\theta_t\partial\theta_u}$; $r,\ t,\ u = 1,\ 2,\ 3$. Derivation of the above terms and the asymptotic biases of the maximum likelihood estimates of μ and c_i under the null hypothesis are given in the Appendix A.1. The results are summarized here. Now, from Appendix A.1 we have the following

$$I_{11} = \sum_{i=1}^{2}\left[\frac{n_i}{\mu(1+c_i\mu)}\right],\quad I_{22} = \frac{n_1}{c_1^4}\sum_{k=1}^{\infty}\frac{k!(b_1c_1)^{k+1}}{(k+1)d_{1k}},\quad I_{33} = \frac{n_2}{c_2^4}\sum_{k=1}^{\infty}\frac{k!(b_2c_2)^{k+1}}{(k+1)d_{2k}},$$

$$J_{111} = \sum_{i=1}^{2}\sum_{j=1}^{n_i}\frac{2(1+2c_i\mu)}{\mu^2(1+c_i\mu)^2},\quad J_{112} = \frac{n_1}{(1+c_1\mu)^2},\quad J_{113} = \frac{n_2}{(1+c_2\mu)^2},$$

$$J_{222} = \frac{2n_1}{c_1^2}\kappa_1 - \frac{n_1\mu(4+5c_1\mu)}{c_1^3(1+c_1\mu)^2} - \frac{2n_1}{c_1^3}\Delta_1,$$

$$J_{333} = \frac{2n_2}{c_2^2}\kappa_2 - \frac{n_2\mu(4+5c_2\mu)}{c_2^3(1+c_2\mu)^2} - \frac{2n_2}{c_2^3}\Delta_2,$$

$$K_{11}^{(1)} = \sum_{i=1}^{2} \left\{ \frac{n_i(1+2c_i\mu)}{\mu^2(1+c_i\mu)^2} \right\}, \quad K_{11}^{(2)} = \frac{n_1}{(1+c_1\mu)^2}, \quad K_{11}^{(3)} = \frac{n_2}{(1+c_2\mu)^2},$$

$$K_{22}^{(1)} = -\frac{n_1}{c_1^4} \sum_{k=1}^{\infty} \frac{k!\,(b_1c_1)^k\, b_1^2}{\mu^2 d_{1k}},$$

$$K_{22}^{(2)} = \frac{n_1}{c_1^5} \sum_{k=1}^{\infty} \frac{k!(c_1b_1)^{k+1}}{(k+1)d_{1k}} \left\{ 4 - \frac{(k+1)(2+c_1\mu)}{1+c_1\mu} + c_1 \left(\sum_{l=0}^{k} \frac{l}{1+c_1 l} \right) \right\},$$

$$K_{33}^{(1)} = -\frac{n_2}{c_2^4} \sum_{k=1}^{\infty} \frac{k!\,(b_2c_2)^k\, b_2^2}{\mu^2 d_{2k}}$$

and

$$K_{33}^{(3)} = \frac{n_2}{c_2^5} \sum_{k=1}^{\infty} \frac{k!(c_2b_2)^{k+1}}{(k+1)d_{2k}} \left\{ 4 - \frac{(k+1)(2+c_2\mu)}{1+c_2\mu} + c_2 \left(\sum_{l=0}^{k} \frac{l}{1+c_2 l} \right) \right\},$$

where $b_i = \dfrac{c_i\mu}{1+c_i\mu}$, $d_{ik} = \prod_{l=0}^{k}(1+c_i l)$, $\kappa_i = \dfrac{1}{c_i^2}\log(1+c_i\mu)$ and $\Delta_i = \sum_{y_{ij}=0}^{\infty} \sum_{l=0}^{y_{ij}-1} \dfrac{c_i^2 l^2 - c_i l - 1}{(1+c_i l)^3} Pr\,(y_{ij})$, for $i = 1,\ 2;\ j = 1,\ 2,\ \cdots,\ n_i$; with $Pr\,(y_{ij}) = f\,(y_{ij}|\mu,\ c_i)$ and $\Delta_i = 0$ if $y_{ij} = 0$. All other elements of I_{rt}, J_{rtu} and $K_{rt}^{(u)}$ are zeros. Now, following Cordeiro and Klein (1994), the biases of $\hat{\mu}_0$, $\hat{c}_{10}$ and $\hat{c}_{20}$ are

$$b_{\hat{\mu}_0}\,(\hat{\mu}_0,\ \hat{c}_{10},\ \hat{c}_{20}) = M^{11} \left(K_{11}^{(1)} - \frac{1}{2} J_{111} \right) M^{11} = 0,$$

$$b_{\hat{c}_{10}}\,(\hat{\mu}_0,\ \hat{c}_{10},\ \hat{c}_{20}) = M^{22} \left[-\frac{1}{2} J_{211} M^{11} + \left(K_{22}^{(2)} - \frac{1}{2} J_{222} \right) M^{22} \right]$$

and

$$b_{\hat{c}_{20}}\,(\hat{\mu}_0,\ \hat{c}_{10},\ \hat{c}_{20}) = M^{33} \left[-\frac{1}{2} J_{311} M^{11} + \left(K_{33}^{(3)} - \frac{1}{2} J_{333} \right) M^{33} \right]$$

respectively, where M^{rt} is the $(r,\ t)^{th}$ element of the inverse of the information matrix I. It is seen that the asymptotic bias of the maximum likelihood estimate of μ is zero. So, the bias corrected estimate of μ is $\hat{\mu}_0\,(bc) = \hat{\mu}_0$. The bias corrected estimate of c_i under the null hypothesis is $\hat{c}_{i0}\,(bc) = \hat{c}_{i0} - b_{\hat{c}_{i0}}\left(\hat{\mu}_0,\ \hat{c}_{10},\ \hat{c}_{20}\right),\ i = 1,\ 2.$

Under the alternative hypothesis, the estimate of μ_i is $\bar{y}_i$, which is unbiased. Yet, for ease and convenience of calculation let $\theta' = (\theta_1,\ \theta_2,\ \theta_3,\ \theta_4) = (\mu_1,\ \mu_2,\ c_1,\ c_2)$ and further, as in the case under the null hypothesis, let $I_{rt} = E\left(-V_{rt}\right),\ J_{rtu} = E\left(W_{rtu}\right)$, and $K_{rt}^{(u)} = \dfrac{\partial}{\partial\theta_u}E\left(V_{rt}\right)$, where $V_{rt} = \dfrac{\partial^2 l}{\partial\theta_r\partial\theta_t}$ and $W_{rtu} = \dfrac{\partial^3 l}{\partial\theta_r\partial\theta_t\partial\theta_u}$; $r,\,t,\,u = 1,\ 2,\ 3,\ 4$. Detailed derivation is given in the Appendix A.2 and from there we have the following

$$I_{11} = \frac{n_1}{\mu_1(1+c_1\mu_1)}, \quad I_{22} = \frac{n_2}{\mu_2(1+c_2\mu_2)},$$

$$I_{33} = \frac{n_1}{c_1^4}\sum_{k=1}^{\infty}\frac{k!(c_1b_1)^{k+1}}{(k+1)d_{1k}}, \quad I_{44} = \frac{n_2}{c_2^4}\sum_{k=1}^{\infty}\frac{k!(c_2b_2)^{k+1}}{(k+1)d_{2k}},$$

$$J_{111} = \frac{2n_1(1+2c_1\mu_1)}{\mu_1^2(1+c_1\mu_1)^2}, \quad J_{113} = \frac{n_1}{(1+c_1\mu_1)^2}, \quad J_{222} = \frac{2n_2(1+2c_2\mu_2)}{\mu_2^2(1+c_2\mu_2)^2},$$

$$J_{224} = \frac{n_2}{(1+c_2\mu_2)^2}, \quad J_{333} = \frac{2n_1}{c_1^2}\kappa_1 - \frac{n_1\mu_1(4+5c_1\mu_1)}{c_1^3(1+c_1\mu_1)^2} - \frac{2n_1}{c_1^3}\Delta_1,$$

$$J_{444} = \frac{2n_2}{c_2^2}\kappa_2 - \frac{n_2\mu_2(4+5c_2\mu_2)}{c_2^3(1+c_2\mu_2)^2} - \frac{2n_2}{c_2^3}\Delta_2,$$

$$K_{11}^{(1)} = \frac{n_1(1+2c_1\mu_1)}{\mu_1^2(1+c_1\mu_1)^2}, \quad K_{11}^{(3)} = \frac{n_1}{(1+c_1\mu_1)^2}, \quad K_{22}^{(2)} = \frac{n_2(1+2c_2\mu_2)}{\mu_2^2(1+c_2\mu_2)^2},$$

$$K_{22}^{(4)} = \frac{n_2}{(1+c_2\mu_2)^2}, \quad K_{33}^{(1)} = -\frac{n_1}{c_1^4}\sum_{k=1}^{\infty}\frac{k!(c_1b_1)^k b_1^2}{\mu_1^2 d_{1k}},$$

$$K_{33}^{(3)} = \frac{n_1}{c_1^5} \sum_{k=1}^{\infty} \frac{k!(c_1 b_1)^{k+1}}{(k+1) d_{1k}} \left\{ 4 - \frac{(k+1)(2+c_1\mu_1)}{1+c_1\mu_1} + c_1 \left(\sum_{l=0}^{k} \frac{l}{1+c_1 l} \right) \right\},$$

$$K_{44}^{(2)} = -\frac{n_2}{c_2^4} \sum_{k=1}^{\infty} \frac{k!(c_2 b_2)^k b_2^2}{\mu_2^2 d_{2k}}$$

and

$$K_{44}^{(4)} = \frac{n_2}{c_2^5} \sum_{k=1}^{\infty} \frac{k!(c_2 b_2)^{k+1}}{(k+1) d_{2k}} \left\{ 4 - \frac{(k+1)(2+c_2\mu_2)}{1+c_2\mu_2} + c_2 \left(\sum_{l=0}^{k} \frac{l}{1+c_2 l} \right) \right\}.$$

As in the previous section, following Cordeiro and Klein (1994), the biases of the estimates of c_1 and c_2 are

$$b_{\hat{c}_1}(\hat{\mu}_1,\, \hat{\mu}_2,\, \hat{c}_1,\, \hat{c}_2) = M^{33} \left\{ -\frac{1}{2} J_{311} M^{11} + \left(K_{33}^{(3)} - \frac{1}{2} J_{333} \right) M^{33} \right\} \quad \text{and}$$

$$b_{\hat{c}_2}(\hat{\mu}_1,\, \hat{\mu}_2,\, \hat{c}_1,\, \hat{c}_2) = M^{44} \left\{ -\frac{1}{2} J_{422} M^{22} + \left(K_{44}^{(4)} - \frac{1}{2} J_{444} \right) M^{44} \right\} \quad \text{respectively.}$$

Then, for $i = 1,\, 2$, the bias corrected estimate of c_i under the alternative hypothesis is $\hat{c}_i\,(bc) = \hat{c}_i - b_{\hat{c}_i}(\hat{\mu}_1,\, \hat{\mu}_2,\, \hat{c}_1,\, \hat{c}_2)$.

Let $\hat{l}_0\,(bc)$ and $\hat{l}_1\,(bc)$ be the maximized log-likelihood under the null and the alternative hypothesis, respectively, using the bias corrected estimates. Then, the likelihood ratio statistic using the bias corrected estimates is $LR\,(bc) = 2\left\{\hat{l}_1\,(bc) - \hat{l}_0\,(bc)\right\}$, which has a $\chi^2\,(1)$ distribution, asymptotically as $n \to \infty$, where $n = n_1 + n_2$.

3.2.3 The Score Test

Suppose that the alternative hypothesis is represented by $\mu_i = \mu + \phi_i$ with $\phi_2 = 0$. Then the null hypothesis $H_0 : \mu_1 = \mu_2$ reduces to $H_0 : \phi_i = 0$ for $i = 1,\, 2$, with μ,

c_1 and c_2 treated as nuisance parameters. The log-likelihood, apart from a constant, can then be written as

$$\begin{aligned} l = &\sum_{i=1}^{2}\sum_{j=1}^{n_i}\sum_{l=1}^{y_{ij}}\left[\log\left(1 + c_i\left(l-1\right)\right)\right] + \sum_{i=1}^{2} y_{i+}[\log(\mu+\phi_i) - \log\{1 + c_i(\mu+\phi_i)\}] \\ &-\sum_{i=1}^{2}\frac{n_i}{c_i}\log\{1 + c_i(\mu+\phi_i)\}. \end{aligned} \tag{3.4}$$

Now, define $\phi = \phi_1$ and $\theta = (\theta_1, \theta_2, \theta_3)' = (\mu, c_1, c_2)'$. Further, define $\psi = \psi(\theta) = \left.\frac{\partial l}{\partial \phi}\right|_{\phi=0}$ and $\gamma_j = \gamma_j(\theta) = \left.\frac{\partial l}{\partial \theta_j}\right|_{\phi=0}$, $j = 1,\ 2,\ 3$. Let $\hat{\theta}$ be some $\sqrt{n}$ consistent estimator of θ under the null hypothesis. Define $S(\hat{\theta}) = \psi(\hat{\theta}) - \sum_{j=1}^{3}\beta_j\gamma_j(\hat{\theta})$, where β_1, β_2 and β_3 are partial regression coefficients of ψ on γ_1, γ_2 and γ_3, respectively. Further, define $D = E\left[-\frac{\partial^2 l}{\partial \phi^2}\right]_{\phi=0}$, a 1×3 matrix A with j^{th} element $A_j = E\left[-\frac{\partial^2 l}{\partial\phi\partial\theta_j}\right]_{\phi=0}$ and 3×3 matrix B with $(i,j)^{th}$ element $B_{i,j} = E\left[-\frac{\partial^2 l}{\partial\theta_i\partial\theta_j}\right]_{\phi=0}$, $i,\ j = 1,\ 2,\ 3$. Then, the $C(\alpha)$ test statistic (Neyman (1959)) for testing $H_0 : \mu_1 = \mu_2$, where c_1 and c_2 are unspecified against the alternative $H_A : \mu_1 \neq \mu_2$, with c_1 and c_2 are unspecified is $T^2 = S^2 \,/\, \left(D - AB^{-1}A'\right)$, which has an approximate chi-square distribution with 1 degree of freedom, where in A, B and D the nuisance parameter θ is replaced by $\hat{\theta}$. Now, $\beta = (\beta_1,\ \beta_2,\ \beta_3)' = AB^{-1}$ and detailed calculation shows that $\beta_1 = A_1 \,/\, B_{1,1}$, $\beta_2 = 0$ and $\beta_3 = 0$. Then, $S\left(\hat{\theta}\right) = \psi(\hat{\theta}) - \beta_1\gamma_1(\hat{\theta}) = \psi\left(\hat{\theta}\right) - A_1\gamma_1\left(\hat{\theta}\right) / B_{1,1}$. If $\hat{\theta}$ is the maximum likelihood estimate of θ, which is $\sqrt{n}$ consistent, then the $C\left(\alpha\right)$ statistic reduces to the score test (Rao (1948)) statistic $T^2 = \psi^2 \,/\, \left(D - AB^{-1}A'\right)$ as $\gamma_1(\hat{\theta}) = 0$. Now, $\psi = \left.\frac{\partial l}{\partial \phi}\right|_{\phi=0} = \frac{n_1\left(\bar{y}_1 - \mu\right)}{\mu\left(1 + c_1\mu\right)}$. The derivation of the components of A, B and D are given in Appendix A.3 and the results are summarized below

$$A_1 = E\left[-\left.\frac{\partial^2 l}{\partial\phi_1\partial\mu}\right|_{\phi=0}\right] = \frac{n_1}{\mu\left(1 + c_1\mu\right)}, \quad A_2 = E\left[-\left.\frac{\partial^2 l}{\partial\phi_1\partial c_1}\right|_{\phi=0}\right] = 0,$$

$$A_3 = E\left[-\frac{\partial^2 l}{\partial\phi_1\partial c_2}\Big|_{\phi=0}\right] = 0, \quad D = E\left[-\frac{\partial^2 l}{\partial\phi_1^2}\Big|_{\phi=0}\right] = \frac{n_1}{\mu\left(1+c_1\mu\right)},$$

$$B_{1,1} = E\left[-\frac{\partial^2 l}{\partial\mu^2}\Big|_{\phi=0}\right] = \sum_{i=1}^{2}\left[\frac{n_i}{\mu\left(1+c_i\mu\right)}\right], \quad B_{1,2} = B_{2,1} = E\left[-\frac{\partial^2 l}{\partial\mu\partial c_1}\Big|_{\phi=0}\right] = 0,$$

$$B_{1,3} = B_{3,1} = E\left[-\frac{\partial^2 l}{\partial\mu\partial c_2}\Big|_{\phi=0}\right] = 0, \quad B_{2,2} = E\left[-\frac{\partial^2 l}{\partial c_1^2}\Big|_{\phi=0}\right] = \frac{n_1}{c_1^4}\sum_{k=1}^{\infty}\frac{k!(b_1c_1)^{k+1}}{(k+1)d_{1k}},$$

$$B_{2,3} = B_{3,2} = E\left[-\frac{\partial^2 l}{\partial c_1\partial c_2}\Big|_{\phi=0}\right] = 0 \quad \text{and} \quad B_{3,3} = E\left[-\frac{\partial^2 l}{\partial c_2^2}\Big|_{\phi=0}\right] = \frac{n_2}{c_2^4}\sum_{k=1}^{\infty}\frac{k!(b_2c_2)^{k+1}}{(k+1)d_{2k}}.$$

After detailed calculations and substitution of the nuisance parameters μ, c_1 and c_2 by their maximum likelihood estimates $\hat{\mu}_0$, $\hat{c}_{10}$ and $\hat{c}_{20}$ in ψ, A, B and D we obtain

$$T^2 = \sum_{i=1}^{2}\frac{n_i(\bar{y}_i - \hat{\mu}_0)^2}{\hat{\mu}_0(1+\hat{\mu}_0\hat{c}_{i0})},$$

which is asymptotically distributed as $\chi^2(1)$ as $n \to \infty$, where $n = n_1 + n_2$.

The score test statistic based on the bias corrected maximum likelihood estimates of the nuisance parameters is

$$T^2(bc) = \sum_{i=1}^{2}\frac{n_i(\bar{y}_i - \hat{\mu}_0)^2}{\hat{\mu}_0(1+\hat{\mu}_0\hat{c}_{i0}(bc))},$$

which also has, asymptotically, a $\chi^2\,(1)$ distribution, as $n \to \infty$, where $n = n_1 + n_2$.

3.2.4 $C(\alpha)$ Test Based on the Method of Moments Estimates of the Nuisance Parameters

Other $\sqrt{n}$ consistent estimates of the parameters μ, c_1 and c_2 can be used. Following Barnwal (1989), the following estimates are obtained.

The first equation of (3.3) can be written as

$$\mu = \sum_{i=1}^{2} w_i \bar{y}_i \,/ \sum_{i=1}^{2} w_i, \tag{3.5}$$

where $w_i = \dfrac{n_i}{\{\mu(1 + c_i\mu)\}}$. Let s_i^2 be the sample variance. Then, a method of moments estimate of $\mu(1 + c_i\mu)$ is s_i^2. Using this in (3.5), we obtain a method of moments estimate for μ as

$$\widetilde{\mu} = \sum_{i=1}^{2} w_i^* \bar{y}_i \,/ \sum_{i=1}^{2} w_i^*, \text{ where } w_i^* = \frac{n_i}{s_i^2}. \tag{3.6}$$

Using this estimate of μ, the moment estimate of c_i is $\widetilde{c_i} = \dfrac{s_i^2 - \widetilde{\mu}}{\widetilde{\mu}^2}$. Clearly, $\widetilde{\mu}$ satisfies $\left.\dfrac{\partial l}{\partial \mu}\right|_{\mu=\widetilde{\mu},\, c_1=\widetilde{c_1},\, c_2=\widetilde{c_2}} = 0$ and thus $\gamma_1(\widetilde{\theta}) = 0$, where $\widetilde{\theta} = (\widetilde{\mu}, \widetilde{c_1}, \widetilde{c_2})'$. With this and after detailed calculations and substitution of the nuisance parameters μ, c_1 and c_2 by their method of moments estimates $\widetilde{\mu}$, $\widetilde{c_1}$, $\widetilde{c_2}$ in ψ, A, B and D the $C(\alpha)$ statistic is obtained as

$$T_1^2 = \sum_{i=1}^{2} \frac{n_i(\bar{y}_i - \widetilde{\mu})^2}{\widetilde{\mu}(1 + \widetilde{c_i}\widetilde{\mu})}.$$

The distribution of T_1^2 is also asymptotic $\chi^2(1)$ as $n \to \infty$.

Note that by using the method of moments estimates of μ and c_i, the $C(\alpha)$ statistic T_1^2 can be written as $T_1^2 = \sum_{i=1}^{2} \frac{n_i(\bar{y}_i - \widetilde{\mu})^2}{s_i^2}$, where $\widetilde{\mu} = \sum_{i=1}^{2} w_i^* \bar{y}_i \,/\, \sum_{i=1}^{2} w_i^*$ and $s_i^2 = \tilde{\mu}(1 + \tilde{\mu} c_i)$. Now, putting $\widetilde{\mu}$ into T_1^2 and simplifying we obtain

$$T_1^2 = \frac{(\bar{y}_1 - \bar{y}_2)^2}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

By taking square-root of both sides of T_1^2 we obtain

$$T_1 = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}.$$

This is an interesting result as the form of T_1 is exactly the same as that of the V statistic of Welch (1937). Also, following the $C(\alpha)$ theory, the asymptotic distribution of T_1^2 is $\chi^2(1)$ and hence the asymptotic distribution of T_1 is standard normal. This suggests, following the theory of Welch (1937), that for small sample size ($n = n_1 + n_2$), the distribution of T_1 might be better approximated by a t distribution with degrees of freedom

$$f = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left(\frac{s_1^4}{n_1^2(n_1 - 1)} + \frac{s_2^4}{n_2^2(n_2 - 1)}\right)},$$

where $\bar{y}_i = \frac{\sum_{j=1}^{n_i} y_{ij}}{n_i}$ and $s_i^2 = \frac{\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2}{n_i - 1}$, although, the data are assumed to come from negative binomial distributions and not from normal distributions. Now, denote the asymptotic normal distribution of T_1 by T_N. It is of interest to see for what sample size n the behaviours of T_1 and T_N are similar.

3.3 Simulation Study

In this section we conduct a simulation study to compare, in terms of size and power, the six test statistics, LR, $LR(bc)$, T^2, $T^2\,(bc)$, T_1 and T_N.

To compare the statistics in terms of size, we considered $n_1 = n_2 = 5, 10, 20, 50$; $\mu_1 = \mu_2 = 1,\ 2,\ 5$, all combinations of $(c_1,\ c_2) = (0.05,\ 0.05)$, $(0.05,\ 0.10)$, $(0.05,\ 0.20)$, $(0.20,\ 0.20)$, $(0.20,\ 0.30)$, $(0.20,\ 0.40)$, $(0.40,\ 0.50)$, $(0.40,\ 0.80)$, and nominal levels $\alpha = 0.05,\quad 0.10$. Comparative results in terms of empirical size for $\mu_1 = \mu_2 = 1$, $\mu_1 = \mu_2 = 2$ and $\mu_1 = \mu_2 = 5$ are similar. So, we present results for only $\mu_1 = \mu_2 = 2$. For power comparison we considered same sample size and same combination of $(c_1,\ c_2)$ with $(\mu_1,\ \mu_2) = (2,\ 2.5)$, $(2,\ 3)$, $(2,\ 4)$, $(2,\ 6)$, $(2,\ 8)$, $(2,\ 10)$ and $(\mu_1,\ \mu_2) = (5,\ 7)$, $(5,\ 10)$, $(5,\ 15)$, $(5,\ 20)$, $(5,\ 25)$. However, comparative results for both α's are similar. Also comparative power for $(\mu_1,\ \mu_2) = (2,\ 2.5)$, $(2,\ 3)$, $(2,\ 4)$, $(2,\ 6)$, $(2,\ 8)$, $(2,\ 10)$ and those for $(5,\ 7)$, $(5,\ 10)$, $(5,\ 15)$, $(5,\ 20)$, $(5,\ 25)$ are similar. So, we present results only for $\alpha = 0.05$ and for $(\mu_1,\ \mu_2) = (2,\ 2.5)$, $(2,\ 3)$, $(2,\ 4)$, $(2,\ 6)$, $(2,\ 8)$, $(2,\ 10)$. These results for level and power are summarized in Tables 3.1 to 3.4.

For small sample sizes, for example, for $n_1 = n_2 = 5$, 10, the LR statistic holds level most effectively. The statistic T_N is liberal, which is expected. The other statistics are in general somewhat conservative. As the statistic T_N is liberal we omit this from comparison of power with the other statistics. For small to moderate over-dispersion parameters the statistic T_1 in general shows the largest power. For very large over-dispersion parameters this statistic is the least powerful and the statistic $T^2(bc)$ is the most powerful. For example, for $\mu_1 = 2$, $\mu_2 = 6$, $c_1 = .4$ and $c_2 = .8$ power of T_1 is 13% and that of $T^2(bc)$ is 22.1%.

For moderate sample size situations, ($n_1 = n_2 = 20$), the statistics T_1 and T_N perform best in both level and power except, again, when the over-dispersions are large in which case the other statistics do somewhat better in terms of power. For example, for $\mu_1 = 2$, $\mu_2 = 4$, $c_1 = .4$ and $c_2 = .8$, power of the statistics LR, $LR(bc)$, T^2, $T^2(bc)$ are .58, .57, .54 and .55 respectively and those of T_1 and T_N are .48 and .53 respectively.

For large sample sizes, for example, for $n_1 = n_2 = 50$, all six statistics do well in terms of level and their power performances are also similar. So, the asymptotic distribution of T_N works when $n_1 = n_2 = 50$. Power of all the statistics increase as the sample sizes increase.

The bias corrected statistics $LR(bc)$ and $T^2(bc)$ do not in general show improvement in terms of level and power over their uncorrected counterparts LR and T^2. The reason could be that the biases of the estimates of c_1 and c_2 are in general small (see Tables 3.5 - 3.8).

To show results for sparse data we present level and power results for $\mu_1 = \mu_2 = 1$ and $\alpha = .05$ in Tables 3.9 - 3.12. As mentioned above, the comparative level and power performances of all the statistics are similar to those of $\mu_1 = \mu_2 = 2$. Average bias of the estimates of c_1 and c_2 for 5000 simulation runs for $\mu_1 = \mu_2 = 1$ and different combinations of c_1 and c_2 are also given in Tables 3.13 - 3.16.

In all cases, empirical significance levels and power of the tests were based on 5000 iterations. Samples were generated from the negative binomial distribution of the form given in equation (3.1) using "*rnegbin*" in "*R*".

Samples for which maximum likelihood estimates did not converge and the estimates of c's found to be close to zero were discarded because small values of c's lead to large bias in the bias correction formula. Samples were discarded if c_i's were less than $0.00001 - 1/\max\{y_{ij}\}$, where $\max\{y_{ij}\}$ is the maximum value of y_{ij} for $i = 1,\ 2;\ \ j = 1,\ 2, \cdots, n_i$. Among the 5000 simulation runs, test statistics were computed for the samples after discarding the non convergent ones and those with small c_i's.

3.4 Examples

3.4.1 Example 1

Lawless (1987) presents a set of data, originally given by Gail et al. (1980), on the times to development of mammary tumours for a total of 48 female rats. Of these 48 rats, 23 were assigned to treatment group 1 (Retinoid) and the remaining 25 rats were assigned to treatment group 2 (Control). The data are presented in Table 3.17. The means for the two groups are 2.65 and 6.04 which are visibly different. The maximum likelihood estimates of the over-dispersion parameters c_1 and c_2 are 0.17 and 0.31 which are moderate but different. The values of the test statistics LR, $LR(bc)$, T^2, $T^2(bc)$, T_1 (degrees of freedom: df=35.66) and T_N with p-values in parenthesis are 13.39 (.0003), 13.33 (.0003), 9.62 (.0019), 9.57 (.002), -3.82 (.0005) and -3.82 (.0001), respectively.

All test statistics produce small p-values indicating strong difference between the means. The p-values of all the statistics are similar indicating rejection of equality at similar levels. This is in line with the simulation results given in Table 3.3 under the columns $(.20, .30)$ and $(.20, .40)$ and row $(2, 6)$ where $n_1 = n_2 = 20$.

Note, the sample sizes for these data are $n_1 = 23$ and $n_2 = 25$. In order to see the power performance of the statistics with very similar characteristics of these data, we did a simulation experiment using 5000 repeated samples with $n_1 = 23$, $n_2 = 25$, $\mu_1 = 2.65$, $\mu_2 = 6.04$, $c_1 = .17$ and $c_2 = .31$. The power of the test statistics LR, $LR(bc)$, T^2, $T^2(bc)$, T_1 and T_N were $97.2, 97.0, 96.8, 96.8, 96.8, 97.3$, respectively. The conclusion, again, is very similar. A further simulation experiment was conducted to see the performance of the test statistics when the means are closer to equality. Everything else remaining the same we used $\mu_1 = 2.65$ and $\mu_2 = 3$. The power of the test statistics LR, $LR(bc)$, T^2, $T^2\,(bc)$, T_1 and T_N were 9.4, 9.0, 8.4, 8.3, 8.8, 9.9, respectively. Here also, the performance of all the test statistics are similar, although, the statistics LR, $LR\,(bc)$, T_1 and T_N do somewhat better than the statistics T^2 and $T^2\,(bc)$. This finding is in line with the simulation results given in Table 3.3 under row (2, 2.5) and column (.20, .30).

3.4.2 Example 2

Saha (2011) presents data, originally given by Uusipaikka (2009), which are given in Table 3.18. The data refer to the number of cycles required for two groups of women, namely, the smoker and non-smoker group, to get pregnant. The means for the two groups are 4.20 and 2.97 which are visibly different. The maximum likelihood estimates of the over-dispersion parameters c_1 and c_2 are 0.47 and 0.38 which are moderate but different. The values of the test statistics LR, $LR(bc)$, T^2, $T^2(bc)$, $T_1(df = 122.66)$ and T_N with p-values in parenthesis are 13.92 (0.0002), 13.92 (0.0002), 15.30 (0.0001), 15.35 (0.0001), 3.14 (0.0021), 3.14 (0.0017), respectively.

As in example 1, all test statistics produce small p-values indicating strong difference between the means. Also, as in example 1, the p-values of all the statistics are similar indicating rejection of equality at similar levels. This is in line with the simulation results given in Table 3.4 under column $(.40, .50)$.

3.5 Discussion

In this Chapter we deal with a problem analogous to the Behrens-Fisher problem for normally distributed data. For testing equality of two population means for count data, we develop six test statistics, namely, a likelihood ratio test statistic LR, a likelihood ratio test statistic $LR(bc)$ based on the bias corrected maximum likelihood estimates of the dispersion parameters, a score test statistic T^2, a score test statistic $T^2(bc)$ based on the bias corrected maximum likelihood estimates of the dispersion parameters, a $C(\alpha)$ test statistic T_1 based on the method of moments estimates of the nuisance parameters and Welch's degree of freedom corrected t distribution and a statistic T_N using the asymptotic distribution of T_1.

Simulations and data analysis show no advantage of the bias corrected statistics $LR(bc)$ and $T^2(bc)$ over their uncorrected counterparts. For small to moderate sample sizes, in general, the statistic T_1 shows best overall performance in terms of size and power and it is easy of calculate. This is an interesting finding in that the $C(\alpha)$ statistic based on the method of moments estimates of the dispersion parameters is the well-known Welch's V statistic which performs well in terms of level and power. For large sample sizes, for example, for $n_1 = n_2 = 50$, all six statistics do well in terms of level and their power performances are also similar. So, the asymptotic distribution of T_N works when $n_1 = n_2 = 50$ or larger.

Table 3.1: Empirical level and power (%) of test statistics LR, $LR(bc)$, T^2, $T^2(bc)$, T_1 and T_N; based on 5000 iterations and $n_1 = 5, n_2 = 5, \alpha = 0.05$

		(c_1, c_2)							
(μ_1, μ_2)	Statistics	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(2, 2)	LR	5.1	4.5	5.2	4.6	5.0	5.7	5.0	5.4
	$LR(bc)$	4.0	3.2	3.8	3.8	4.4	5.4	4.7	5.3
	T^2	4.1	3.7	4.2	3.4	3.6	4.1	3.8	4.4
	$T^2(bc)$	4.2	3.8	4.4	3.7	3.8	4.7	4.6	5.4
	T_1	4.3	4.4	4.0	3.7	4.1	3.5	3.4	3.2
	T_N	8.7	9.1	8.5	8.4	8.9	8.4	7.8	8.0
(2, 2.5)	LR	5.7	6.4	6.2	5.4	5.7	6.3	5.8	6.1
	$LR(bc)$	4.5	5.3	5.1	4.3	5.0	6.2	6.2	6.0
	T^2	5.3	5.6	5.7	4.9	4.4	4.7	4.6	5.8
	$T^2(bc)$	5.6	6.1	6.6	5.5	4.7	5.4	5.3	6.2
	T_1	6.0	5.5	4.3	5.1	4.3	4.0	3.2	3.6
	T_N	11.7	10.9	10.2	10.1	9.7	9.7	8.3	8.2
(2, 3)	LR	7.5	7.0	9.4	7.1	5.8	6.3	6.0	6.6
	$LR(bc)$	6.4	6.3	8.6	6.3	5.1	6.5	5.8	7.0
	T^2	7.5	7.2	9.3	7.0	6.6	8.2	6.5	7.1
	$T^2(bc)$	7.8	7.4	10.0	7.2	7.2	8.7	7.5	8.8
	T_1	10.2	10.1	8.3	8.5	7.3	6.1	5.3	4.3
	T_N	19.5	17.7	16.0	16.0	14.4	13.5	12.6	10.1
(2, 4)	LR	15.4	15.6	15.9	12.0	12.9	12.3	8.4	7.4
	$LR(bc)$	13.9	14.4	14.0	10.9	12.7	11.1	8.2	7.6
	T^2	15.0	15.8	16.2	12.4	13.6	13.5	10.0	9.6
	$T^2(bc)$	15.1	16.5	17.5	13.1	14.5	14.8	11.6	11.9
	T_1	28.0	24.0	18.1	19.0	15.8	12.5	11.2	6.3
	T_N	43.2	38.7	32.8	33.0	28.9	25.1	22.8	16.4
(2, 6)	LR	35.1	33.5	35.7	24.0	22.8	23.6	16.0	15.3
	$LR(bc)$	31.6	30.3	32.8	21.0	19.9	22.0	14.7	15.5
	T^2	29.8	29.7	36.5	23.4	22.4	25.1	17.7	19.3
	$T^2(bc)$	29.9	30.5	37.6	24.2	23.4	27.1	19.6	22.1
	T_1	64.8	58.0	43.1	43.7	32.9	27.6	22.2	12.8
	T_N	81.1	77.6	66.8	65.6	56.6	50.9	42.8	30.7
(2, 8)	LR	46.6	50.5	51.6	37.0	36.3	38.2	25.6	24.7
	$LR(bc)$	37.7	45.2	47.7	29.6	30.9	32.7	21.0	21.8
	T^2	30.4	40.8	49.4	30.5	33.2	37.5	23.5	26.2
	$T^2(bc)$	30.9	41.4	50.3	31.2	34.8	39.6	25.0	30.4
	T_1	87.4	79.2	63.2	61.5	50.7	39.6	33.7	19.3
	T_N	96.6	93.6	85.1	84.4	75.2	67.1	60.1	43.4
(2, 10)	LR	64.0	60.6	61.0	53.9	48.4	45.9	33.3	32.4
	$LR(bc)$	45.3	51.0	53.5	41.8	37.8	40.4	26.1	28.8
	T^2	34.6	43.4	54.1	39.0	40.0	43.7	28.4	32.8
	$T^2(bc)$	34.6	42.9	55.9	39.0	41.0	45.7	30.6	35.8
	T_1	96.0	89.1	75.2	74.7	60.9	49.8	40.8	23.2
	T_N	99.5	98.3	93.4	92.7	85.9	78.8	70.2	52.4

Table 3.2: Empirical level and power (%) of test statistics LR, $LR(bc)$, T^2, $T^2(bc)$, T_1 and T_N; based on 5000 iterations and $n_1 = 10, n_2 = 10, \alpha = 0.05$

		(c_1, c_2)							
(μ_1, μ_2)	Statistics	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(2, 2)	LR	3.6	3.5	3.5	4.2	4.0	3.9	4.0	4.3
	$LR(bc)$	2.6	2.8	3.0	3.3	3.2	3.5	3.1	3.7
	T^2	2.8	2.2	2.1	2.6	2.3	2.4	2.4	2.8
	$T^2(bc)$	2.7	2.0	2.1	2.5	2.3	2.5	2.5	3.0
	T_1	4.7	4.3	4.2	4.0	4.7	3.9	4.0	4.5
	T_N	6.5	6.3	5.9	5.8	6.7	6.0	5.7	6.7
(2, 2.5)	LR	4.2	4.7	6.2	4.8	4.3	5.3	5.1	5.4
	$LR(bc)$	3.2	3.4	4.7	3.4	3.2	3.8	3.6	4.2
	T^2	3.5	4.2	5.4	4.0	3.2	4.1	3.8	4.2
	$T^2(bc)$	3.4	4.0	5.4	3.9	3.2	4.2	4.0	4.6
	T_1	9.6	9.2	7.2	8.3	7.0	6.6	6.3	5.0
	T_N	12.6	11.9	10.3	10.8	9.8	8.9	9.2	7.2
(2, 3)	LR	10.5	11.9	12.2	10.1	10.6	10.1	8.8	8.3
	$LR(bc)$	6.7	8.6	8.9	6.4	7.3	7.3	5.5	6.2
	T^2	8.0	10.0	9.9	6.8	7.4	8.0	5.8	6.8
	$T^2(bc)$	7.7	9.6	9.6	6.8	7.4	8.1	6.0	7.4
	T_1	23.9	19.5	18.5	18.0	14.9	11.9	10.9	8.5
	T_N	28.7	24.7	23.8	22.4	19.3	16.6	15.1	12.5
(2, 4)	LR	40.1	39.1	37.6	29.7	29.6	29.4	23.1	22.1
	$LR(bc)$	27.5	27.8	29.6	19.9	20.6	21.6	15.5	16.1
	T^2	23.1	24.4	26.6	15.2	17.0	19.0	12.1	15.4
	$T^2(bc)$	22.7	23.7	26.1	15.0	17.0	19.2	12.4	16.2
	T_1	61.3	57.4	47.4	46.5	38.7	35.3	28.9	19.7
	T_N	67.3	64.1	55.6	53.5	47.1	42.5	36.9	28.0
(2, 6)	LR	90.6	87.0	83.1	78.4	74.0	69.0	59.8	54.1
	$LR(bc)$	69.4	68.9	67.0	59.6	58.0	55.4	45.3	43.0
	T^2	31.8	38.0	50.3	27.2	34.6	38.9	27.4	34.0
	$T^2(bc)$	32.4	38.7	50.7	27.5	35.0	39.5	27.8	35.0
	T_1	97.0	95.5	89.1	87.9	81.3	73.2	64.2	46.2
	T_N	98.2	97.2	93.4	91.9	87.3	80.6	73.2	58.6
(2, 8)	LR	99.2	98.9	97.0	95.4	93.0	90.9	83.8	75.6
	$LR(bc)$	71.2	71.1	78.8	71.0	70.1	72.4	63.2	61.3
	T^2	23.7	30.1	46.7	29.3	38.2	48.8	35.2	45.3
	$T^2(bc)$	28.2	30.4	47.3	30.8	39.4	49.3	35.7	46.8
	T_1	100	99.8	98.3	97.9	95.0	90.6	84.0	64.7
	T_N	100	99.9	99.4	98.8	97.4	95.3	91.0	77.0
(2, 10)	LR	100	99.6	99.4	98.3	98.1	97.2	93.3	87.7
	$LR(bc)$	62.0	65.4	70.0	69.3	70.3	75.5	68.0	67.5
	T^2	13.9	24.3	43.3	21.7	35.3	44.8	38.2	49.2
	$T^2(bc)$	16.7	27.2	45.4	24.3	36.8	46.1	39.6	50.5
	T_1	100	100	99.8	99.7	98.4	96.5	92.5	77.5
	T_N	100	100	99.9	99.9	99.5	98.6	96.5	87.6

Table 3.3: Empirical level and power (%) of test statistics LR, $LR(bc)$, T^2, $T^2(bc)$, T_1 and T_N; based on 5000 iterations and $n_1 = 20, n_2 = 20, \alpha = 0.05$

		(c_1, c_2)							
(μ_1, μ_2)	Statistics	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(2, 2)	LR	3.8	3.9	4.6	4.7	4.8	4.9	5.1	5.1
	$LR(bc)$	3.2	3.2	4.0	4.1	4.4	4.7	4.8	4.9
	T^2	3.2	2.9	4.0	3.7	3.7	3.7	3.8	3.7
	$T^2(bc)$	3.2	2.9	3.9	3.6	3.6	3.6	3.8	3.8
	T_1	4.5	5.0	4.8	5.3	4.7	4.3	4.8	4.8
	T_N	5.3	5.9	5.3	6.0	5.9	5.3	5.4	5.6
(2, 2.5)	LR	14.0	12.5	12.8	11.9	11.8	12.2	10.5	10.7
	$LR(bc)$	12.2	11.2	11.2	11.1	11.1	11.3	10.1	10.4
	T^2	11.3	10.6	11.7	9.4	9.8	10.7	8.4	9.5
	$T^2(bc)$	10.9	10.4	11.4	9.4	9.7	10.7	8.5	9.6
	T_1	16.8	15.0	12.8	13.3	11.2	11.0	9.4	8.1
	T_N	18.8	16.6	14.7	14.9	12.6	12.6	11.2	9.4
(2, 3)	LR	37.7	36.8	35.1	32.0	30.3	29.3	25.1	22.8
	$LR(bc)$	33.8	33.8	32.8	30.3	29.0	27.7	24.4	22.5
	T^2	32.4	32.2	32.9	26.4	26.4	26.4	21.2	20.6
	$T^2(bc)$	32.0	32.0	32.6	26.2	26.4	26.4	21.4	20.9
	T_1	45.6	42.0	35.2	33.9	31.1	27.0	24.3	18.1
	T_N	48.7	45.1	38.1	36.7	34.4	30.0	27.0	20.9
(2, 4)	LR	87.2	84.3	81.0	76.8	73.9	71.2	63.5	58.4
	$LR(bc)$	84.2	81.6	78.8	74.7	72.8	70.2	62.4	57.7
	T^2	81.8	80.0	77.7	69.9	69.4	67.3	55.4	54.4
	$T^2(bc)$	81.6	79.7	77.5	69.8	69.4	67.4	55.8	55.0
	T_1	92.0	88.9	83.4	80.5	74.5	68.5	61.8	48.3
	T_N	93.1	90.4	85.5	82.9	77.3	71.8	64.8	53.4
(2, 6)	LR	100	99.9	99.9	99.7	99.2	98.7	96.2	92.6
	$LR(bc)$	98.7	99.0	99.0	99.0	98.4	97.8	95.6	92.3
	T^2	99.4	99.7	99.6	97.2	97.8	97.3	90.7	90.3
	$T^2(bc)$	99.7	99.7	99.7	97.7	97.9	97.5	91.6	90.7
	T_1	100	100	99.8	99.6	99.1	98.3	95.9	88.3
	T_N	100	100	99.8	99.7	99.3	98.6	96.9	90.5
(2, 8)	LR	100	100	100	100	100	99.9	99.9	99.3
	$LR(bc)$	98.0	98.3	99.1	99.0	98.7	99.4	99.1	98.6
	T^2	98.9	98.9	99.8	98.2	99.0	99.0	96.6	97.7
	$T^2(bc)$	99.2	99.6	99.8	99.1	99.4	99.4	97.4	97.9
	T_1	100	100	100	100	100	99.9	99.7	98.1
	T_N	100	100	100	100	100	100	99.8	98.7
(2, 10)	LR	100	100	100	100	100	100	100	100
	$LR(bc)$	98.3	98.0	98.7	98.9	99.1	99.0	98.8	99.0
	T^2	99.5	98.5	99.6	99.0	98.3	99.3	96.9	98.3
	$T^2(bc)$	99.8	99.4	99.9	99.5	99.2	99.6	98.0	98.8
	T_1	100	100	100	100	100	100	100	99.5
	T_N	100	100	100	100	100	100	100	99.7

Table 3.4: Empirical level and power (%) of test statistics LR, $LR(bc)$, T^2, $T^2(bc)$, T_1 and T_N; based on 5000 iterations and $n_1 = 50, n_2 = 50, \alpha = 0.05$

		(c_1, c_2)							
(μ_1, μ_2)	Statistics	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(2, 2)	LR	5.3	4.5	5.4	5.3	5.3	4.5	5.4	5.7
	$LR(bc)$	4.9	4.3	5.3	5.2	5.2	4.5	5.4	5.7
	T^2	5.0	4.3	5.1	5.0	5.0	4.2	4.9	5.2
	$T^2(bc)$	5.0	4.2	5.0	5.0	4.9	4.1	4.9	5.2
	T_1	5.1	4.8	4.5	4.9	4.9	5.1	5.2	5.0
	T_N	5.4	5.0	4.7	5.2	5.2	5.2	5.5	5.1
(2, 2.5)	LR	34.0	32.7	30.4	27.8	26.7	25.6	21.2	20.0
	$LR(bc)$	33.1	32.1	29.8	27.6	26.5	25.4	21.2	19.9
	T^2	33.3	32.1	30.2	26.9	26.0	25.2	20.3	20.1
	$T^2(bc)$	32.9	31.8	30.0	26.8	25.9	25.2	20.3	20.2
	T_1	35.1	34.0	29.0	26.5	25.8	23.1	20.1	15.5
	T_N	36.0	35.0	30.2	27.3	26.7	24.1	21.0	16.1
(2, 3)	LR	83.3	81.6	78.2	72.7	69.4	66.7	58.7	52.2
	$LR(bc)$	82.8	81.0	77.8	72.4	69.2	66.6	58.6	52.0
	T^2	83.0	81.0	78.0	71.3	68.7	66.3	57.2	52.2
	$T^2(bc)$	82.9	80.9	78.0	71.2	68.5	66.3	57.2	52.3
	T_1	83.3	82.2	76.1	72.9	67.4	64.1	56.9	46.7
	T_N	83.9	83.0	77.2	73.4	68.3	65.2	58.1	48.1
(2, 4)	LR	99.9	99.9	99.8	99.7	99.3	98.6	97.1	93.9
	$LR(bc)$	99.7	99.6	99.7	99.6	99.2	98.5	97.1	93.9
	T^2	99.9	99.9	99.8	99.6	99.2	98.5	97.0	94.0
	$T^2(bc)$	99.9	99.9	99.8	99.6	99.2	98.5	97.0	94.0
	T_1	100	100	99.8	99.4	99.2	98.5	96.3	90.9
	T_N	100	100	99.8	99.5	99.3	98.7	96.5	91.5
(2, 6)	LR	100	100	100	100	100	100	100	100
	$LR(bc)$	99.9	100	100	100	100	100	100	100
	T^2	100	100	100	100	100	100	100	100
	$T^2(bc)$	100	100	100	100	100	100	100	100
	T_1	100	100	100	100	100	100	100	100
	T_N	100	100	100	100	100	100	100	100
(2, 8)	LR	100	100	100	100	100	100	100	100
	$LR(bc)$	100	100	99.9	100	100	100	100	100
	T^2	100	100	100	100	100	100	100	100
	$T^2(bc)$	100	100	100	100	100	100	100	100
	T_1	100	100	100	100	100	100	100	100
	T_N	100	100	100	100	100	100	100	100
(2, 10)	LR	100	100	100	100	100	100	100	100
	$LR(bc)$	100	100	100	100	100	100	100	100
	T^2	100	100	100	100	100	100	100	100
	$T^2(bc)$	100	100	100	100	100	100	100	100
	T_1	100	100	100	100	100	100	100	100
	T_N	100	100	100	100	100	100	100	100

Table 3.5: Estimated bias of the estimates of dispersion parameters, c_1 and c_2, under null and alternative hypotheses; based on 5000 iterations and $n_1 = 5, n_2 = 5$

		(c_1, c_2)							
(μ_1, μ_2)	Par†	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(2, 2)	c_{10}	-0.0177	-0.0174	-0.0202	0.0234	-0.0352	0.0688	0.0716	-0.0080
	c_{20}	0.0529	0.0526	0.1023	-0.0119	-0.0086	-0.0241	0.0084	-0.0261
	c_{1a}	-0.0679	-0.0576	-0.0594	-0.0098	0.0071	0.0721	-0.0005	0.0039
	c_{2a}	-0.0097	-0.0218	-0.0083	0.0045	0.0628	0.0099	-0.0081	-0.0175
(2, 2.5)	c_{10}	-0.0179	-0.0154	-0.0225	0.0150	0.0075	0.0061	0.0099	-0.0011
	c_{20}	0.0115	0.0300	0.0629	0.0526	-0.0028	-0.0084	0.0070	0.0042
	c_{1a}	-0.0481	-0.0496	-0.0549	-0.0351	-0.0064	0.0167	0.0028	-0.0033
	c_{2a}	-0.0355	0.0218	-0.0415	-0.0377	-0.0188	0.0031	-0.0069	0.0055
(2, 3)	c_{10}	-0.0072	-0.0173	-0.0175	0.0127	0.0170	0.0083	0.0485	-0.0052
	c_{20}	0.0071	0.0216	0.0589	0.0492	0.0816	0.1154	0.1585	-0.0474
	c_{1a}	-0.0326	-0.0417	-0.0422	-0.0329	-0.0035	-0.0100	0.0243	-0.0318
	c_{2a}	-0.0082	-0.0147	-0.0366	-0.0299	0.0086	-0.0949	-0.0924	0.0077
(2, 4)	c_{10}	-0.0002	-0.0042	-0.0139	0.0138	0.0110	-0.0003	-0.0080	-0.0019
	c_{20}	0.0056	0.0256	0.0583	0.0482	0.1076	0.0614	-0.0013	-0.0127
	c_{1a}	0.0112	-0.0098	-0.0042	0.0333	0.0302	0.0135	0.0293	-0.0058
	c_{2a}	0.0014	-0.0213	-0.0317	-0.0503	-0.0077	-0.0011	0.0075	-0.0098
(2, 6)	c_{10}	0.0183	0.0033	-0.0095	0.0304	0.0099	0.0088	-0.0441	0.0032
	c_{20}	0.0281	0.0627	0.0024	0.0735	0.0319	0.0112	-0.0125	0.0033
	c_{1a}	0.0343	0.0708	0.0400	0.0712	0.0579	0.0807	0.0015	-0.0070
	c_{2a}	-0.0550	-0.0615	-0.0721	-0.0681	-0.0776	-0.0836	0.0873	0.0424
(2, 8)	c_{10}	0.0433	0.0223	-0.0089	0.0336	0.0219	0.0023	0.0091	-0.0073
	c_{20}	0.0483	0.0792	0.0227	0.0479	-0.0237	0.0076	-0.0022	-0.0091
	c_{1a}	0.0328	0.0477	0.0746	-0.0291	0.0422	0.0994	-0.0877	0.0434
	c_{2a}	-0.0470	-0.0565	-0.0727	-0.0642	-0.0771	-0.0845	0.0089	0.0870
(2, 10)	c_{10}	0.0711	0.0298	-0.0079	0.0551	0.0169	-0.0002	0.0422	0.0007
	c_{20}	0.0322	0.0702	-0.0355	0.0047	0.0058	0.0319	0.0025	-0.0579
	c_{1a}	0.0349	- 0.0501	0.0412	0.0092	0.0258	0.0341	0.0054	-0.0311
	c_{2a}	-0.0421	-0.0522	-0.0733	-0.0641	-0.0767	-0.0890	-0.0861	0.07151

Par† = Parameters

c_{10} is c_1 under null hypothesis

c_{20} is c_2 under null hypothesis

c_{1a} is c_1 under alternative hypothesis

c_{2a} is c_2 under alternative hypothesis

Table 3.6: Estimated bias of the estimates of dispersion parameters, c_1 and c_2, under null and alternative hypotheses; based on 5000 iterations and $n_1 = 10, n_2 = 10$

		(c_1, c_2)							
(μ_1, μ_2)	Par†	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(2, 2)	c_{10}	-0.0145	-0.0148	-0.0164	-0.0024	-0.0040	-0.0067	0.0112	0.0051
	c_{20}	-0.0113	-0.0069	0.0039	0.0002	0.0118	0.0303	0.0316	0.0093
	c_{1a}	-0.0315	-0.0303	-0.0319	-0.0229	-0.0228	-0.0232	-0.0132	-0.0157
	c_{2a}	-0.0546	-0.0555	-0.0563	-0.0564	-0.0575	-0.0570	-0.0579	-0.0521
(2, 2.5)	c_{10}	-0.0118	-0.0125	-0.0148	-0.0005	-0.0024	-0.0066	0.0136	0.0059
	c_{20}	-0.0138	-0.0102	0.0000	-0.0031	0.0049	0.0194	0.0279	0.0743
	c_{1a}	-0.0198	-0.0160	-0.0214	-0.0052	-0.0083	-0.0103	0.0102	0.0003
	c_{2a}	-0.0458	-0.0467	-0.0489	-0.0482	-0.0509	-0.0519	-0.0532	-0.0529
(2, 3)	c_{10}	-0.0065	-0.0086	-0.0110	0.0046	0.0003	-0.0029	0.0201	0.0082
	c_{20}	-0.0129	-0.0091	-0.0006	-0.0043	0.0048	0.0181	0.0230	0.0720
	c_{1a}	0.0067	0.0024	0.0006	0.0212	0.0124	0.0111	0.0446	0.0265
	c_{2a}	-0.0395	-0.0409	-0.0438	-0.0439	-0.0462	-0.0487	-0.0501	-0.0512
(2, 4)	c_{10}	0.0070	0.0046	-0.0035	0.0197	0.0126	0.0084	0.0271	0.0157
	c_{20}	-0.0100	-0.0056	0.0074	-0.0041	0.0053	0.0204	0.0209	0.0710
	c_{1a}	0.0758	0.0660	0.0464	0.097	0.0812	0.0740	0.1025	0.0835
	c_{2a}	-0.0318	-0.0342	-0.0382	-0.0375	-0.0412	-0.0446	-0.0469	-0.0510
(2, 6)	c_{10}	0.0473	0.0374	0.0193	0.0475	0.0406	0.0292	0.0500	0.0308
	c_{20}	-0.0083	-0.0011	0.0156	-0.0042	0.0077	0.0264	0.0188	0.0794
	c_{1a}	0.0317	0.0094	0.0068	0.0419	0.0084	0.0612	0.0255	0.0433
	c_{2a}	-0.0236	-0.0269	-0.0334	-0.0318	-0.0369	-0.0416	-0.0443	-0.0514
(2, 8)	c_{10}	0.0903	0.0743	0.0475	0.0862	0.0690	0.0524	0.0718	0.0456
	c_{20}	-0.0109	-0.0069	0.0178	-0.0085	0.0034	0.0259	0.0197	0.0821
	c_{1a}	0.0011	0.0057	-0.0318	0.0137	0.0472	0.0369	0.0758	0.0146
	c_{2a}	-0.0184	-0.0224	-0.0305	-0.0282	-0.0348	-0.0409	-0.0437	-0.0526
(2, 10)	c_{10}	0.0228	-0.0194	0.0680	0.1080	0.0886	0.0640	0.0824	0.0580
	c_{20}	-0.0128	-0.0078	0.0196	-0.0082	0.0039	0.0302	0.0124	0.0847
	c_{1a}	0.0128	0.0251	0.0413	0.0021	0.0655	0.02087	0.0419	0.0944
	c_{2a}	-0.0150	-0.0204	-0.0297	-0.0267	-0.0337	-0.0403	-0.0440	-0.0536

Par† = Parameters
c_{10} is c_1 under null hypothesis
c_{20} is c_2 under null hypothesis
c_{1a} is c_1 under alternative hypothesis
c_{2a} is c_2 under alternative hypothesis

Table 3.7: Estimated bias of the estimates of dispersion parameters, c_1 and c_2, under null and alternative hypotheses; based on 5000 iterations and $n_1 = 20, n_2 = 20$

		(c_1, c_2)							
(μ_1, μ_2)	Par†	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(2, 2)	c_{10}	-0.0093	-0.0097	-0.0107	-0.0051	-0.0064	-0.0078	0.0004	-0.0032
	c_{20}	-0.0094	-0.0078	-0.0032	-0.0055	0.0002	0.0049	0.0076	0.0297
	c_{1a}	-0.0190	-0.0185	-0.0201	-0.0187	-0.0186	-0.0198	-0.0180	-0.0183
	c_{2a}	-0.0267	-0.0269	-0.0280	-0.0279	-0.0286	-0.0292	-0.0296	-0.0289
(2, 2.5)	c_{10}	-0.0071	-0.0080	-0.0095	-0.0037	-0.0050	-0.0064	0.0013	-0.0029
	c_{20}	-0.0093	-0.0077	-0.0037	-0.0057	-0.0014	0.0036	0.0055	0.0253
	c_{1a}	-0.0049	-0.0083	-0.0093	-0.0059	-0.0072	-0.0091	-0.0054	-0.0092
	c_{2a}	-0.0218	-0.0226	-0.0239	-0.0240	-0.0251	-0.0263	-0.0271	-0.0280
(2, 3)	c_{10}	-0.0051	-0.0058	-0.0076	-0.0008	-0.0024	-0.0042	0.0033	-0.0013
	c_{20}	-0.0083	-0.0068	-0.0029	-0.0057	-0.0015	0.0032	0.0046	0.0228
	c_{1a}	0.0148	0.0124	0.0073	0.0148	0.0131	0.0082	0.0123	0.0066
	c_{2a}	-0.0189	-0.0198	-0.0216	-0.0214	-0.0230	-0.0245	-0.0256	-0.0276
(2, 4)	c_{10}	0.0040	0.0013	-0.0024	0.0079	0.0043	0.0014	0.0106	0.0035
	c_{20}	-0.0078	-0.0063	-0.0019	-0.0061	-0.0018	0.0029	0.0030	0.0229
	c_{1a}	0.0781	0.0654	0.0561	0.0763	0.0635	0.0554	0.0646	0.0479
	c_{2a}	-0.0150	-0.0163	-0.0187	-0.0183	-0.0206	-0.0225	-0.0241	-0.0271
(2, 6)	c_{10}	0.0336	0.0291	0.0197	0.0339	0.0265	0.0200	0.0313	0.0187
	c_{20}	-0.0083	-0.0080	-0.0044	-0.0089	-0.0065	-0.0010	-0.0033	0.0163
	c_{1a}	0.0042	0.0048	0.0284	0.0543	0.0218	0.0093	0.0243	0.0513
	c_{2a}	-0.0106	-0.0123	-0.0158	-0.0151	-0.0181	-0.0208	-0.0225	-0.0270
(2, 8)	c_{10}	0.0612	0.0564	0.0451	0.0598	0.0491	0.0413	0.0530	0.0350
	c_{20}	-0.0080	-0.0088	-0.0080	-0.0103	-0.0098	-0.0065	-0.0089	0.0081
	c_{1a}	0.0311	0.0247	0.0328	0.0988	0.0455	0.0077	0.0062	0.0429
	c_{2a}	-0.0085	-0.0100	-0.0140	-0.0136	-0.017	-0.0197	-0.0219	-0.0272
(2, 10)	c_{10}	0.0800	0.0783	0.0676	0.0772	0.0673	0.0616	0.0705	0.0518
	c_{20}	-0.0069	-0.0084	-0.0103	-0.0109	-0.0119	-0.0111	-0.0118	0.0008
	c_{1a}	0.0091	0.0118	0.0248	0.0448	0.0013	0.0054	0.0346	0.0288
	c_{2a}	-0.0072	-0.0090	-0.0128	-0.0126	-0.0160	-0.0192	-0.0217	-0.0276

Par† = Parameters
c_{10} is c_1 under null hypothesis
c_{20} is c_2 under null hypothesis
c_{1a} is c_1 under alternative hypothesis
c_{2a} is c_2 under alternative hypothesis

Table 3.8: Estimated bias of the estimates of dispersion parameters, c_1 and c_2, under null and alternative hypotheses; based on 5000 iterations and $n_1 = 50, n_2 = 50$

		(c_1, c_2)							
(μ_1, μ_2)	Par†	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(2, 2)	c_{10}	-0.0042	-0.0044	-0.0047	-0.0028	-0.0033	-0.0037	-0.0011	-0.0024
	c_{20}	-0.0042	-0.0036	-0.0020	-0.0027	-0.0011	0.0009	0.0016	0.0090
	c_{1a}	-0.0090	-0.0088	-0.0092	-0.0094	-0.0097	-0.0098	-0.0104	-0.0106
	c_{2a}	-0.0104	-0.0106	-0.0110	-0.0111	-0.0115	-0.0118	-0.0120	-0.0121
(2, 2.5)	c_{10}	-0.0036	-0.0038	-0.0043	-0.0024	-0.0028	-0.0034	-0.0009	-0.0025
	c_{20}	-0.0038	-0.0033	-0.0018	-0.0026	-0.0012	0.0007	0.0013	0.0080
	c_{1a}	-0.0031	-0.0038	-0.0044	-0.0052	-0.0061	-0.0063	-0.0077	-0.0084
	c_{2a}	-0.0086	-0.0089	-0.0095	-0.0095	-0.0101	-0.0107	-0.0111	-0.0117
(2, 3)	c_{10}	-0.0026	-0.0030	-0.0037	-0.0015	-0.0022	-0.0029	-0.0005	-0.0020
	c_{20}	-0.0036	-0.0030	-0.0016	-0.0026	-0.0011	0.0009	0.0009	0.0077
	c_{1a}	0.0141	0.0122	0.0076	0.0051	0.0027	0.0013	-0.0015	-0.0038
	c_{2a}	-0.0073	-0.0077	-0.0085	-0.0085	-0.0093	-0.0100	-0.0105	-0.0114
(2, 4)	c_{10}	0.0009	-0.0001	-0.0016	0.0016	0.0001	-0.0010	0.0019	-0.0008
	c_{20}	-0.0036	-0.0032	-0.0017	-0.0031	-0.0017	0.0002	0.0001	0.0074
	c_{1a}	0.0647	0.0605	0.0513	0.0351	0.0297	0.0248	0.0186	0.0107
	c_{2a}	-0.0058	-0.0063	-0.0073	-0.0073	-0.0083	-0.0092	-0.0098	-0.0112
(2, 6)	c_{10}	0.0130	0.0115	0.0080	0.0113	0.0087	0.0065	0.0099	0.0048
	c_{20}	-0.0037	-0.0040	-0.0039	-0.0044	-0.0040	-0.0029	-0.0029	0.0038
	c_{1a}	-0.0055	0.0122	0.0310	0.0382	-0.0086	0.0299	0.0511	0.0133
	c_{2a}	-0.0041	-0.0048	-0.0061	-0.0061	-0.0073	-0.0083	-0.0092	-0.0111
(2, 8)	c_{10}	0.0226	0.0206	0.0175	0.0208	0.0181	0.0153	0.0195	0.0123
	c_{20}	-0.0032	-0.0037	-0.0045	-0.0046	-0.0050	-0.0049	-0.0050	-0.0004
	c_{1a}	-0.0054	0.0291	0.0216	0.0428	0.0097	0.0278	0.0081	0.0317
	c_{2a}	-0.0033	-0.0040	-0.0055	-0.0055	-0.0068	-0.0079	-0.0090	-0.0112
(2, 10)	c_{10}	0.0301	0.0281	0.0248	0.0287	0.0257	0.0227	0.0271	0.0194
	c_{20}	-0.0027	-0.0034	-0.0045	-0.0046	-0.0053	-0.0057	-0.0060	-0.0033
	c_{1a}	0.0088	-0.0117	0.0052	0.0314	0.0054	0.0292	0.0214	0.0084
	c_{2a}	-0.0028	-0.0036	-0.0052	-0.0051	-0.0065	-0.0078	-0.0089	-0.0114

Par† = Parameters
c_{10} is c_1 under null hypothesis
c_{20} is c_2 under null hypothesis
c_{1a} is c_1 under alternative hypothesis
c_{2a} is c_2 under alternative hypothesis

Table 3.9: Empirical level and power (%) of test statistics LR, $LR(bc)$, T^2, $T^2(bc)$, T_1 and T_N; based on 5000 iterations and $n_1 = 5, n_2 = 5, \alpha = 0.05$

		(c_1, c_2)							
(μ_1, μ_2)	Statistics	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(1, 1)	LR	2.8	2.1	3.0	2.5	2.7	3.7	2.9	3.3
	$LR(bc)$	2.5	1.9	2.8	2.3	2.4	3.5	2.6	3.0
	T^2	2.7	2.4	3.5	3.3	3.3	3.8	3.0	3.4
	$T^2(bc)$	3.0	2.8	3.7	3.6	3.7	4.0	3.5	3.7
	T_1	3.9	3.5	3.3	3.3	3.4	3.2	2.9	2.6
	T_N	8.2	7.9	8.4	8.2	7.9	8.0	7.8	7.3
(1, 1.1)	LR	3.0	2.4	3.3	2.9	2.7	3.9	3.3	3.6
	$LR(bc)$	2.7	2.2	3.0	2.7	2.4	3.7	3.1	3.5
	T^2	2.8	2.6	3.6	3.5	3.4	4.1	3.2	3.7
	$T^2(bc)$	3.2	3.0	3.7	3.8	3.8	4.4	3.6	3.9
	T_1	4.1	3.6	3.7	3.6	3.9	3.3	3.1	2.7
	T_N	8.5	8.1	8.5	8.3	8.9	8.4	8.2	7.5
(1, 1.2)	LR	3.1	2.6	3.5	3.0	2.9	4.1	3.6	3.9
	$LR(bc)$	2.9	2.5	3.2	2.9	2.9	3.9	3.3	3.7
	T^2	2.9	2.8	3.9	4.0	3.6	4.2	3.5	4.0
	$T^2(bc)$	3.4	3.4	3.9	4.3	3.9	4.4	3.7	4.1
	T_1	4.4	4.5	4.0	4.2	4.0	3.7	3.6	2.8
	T_N	9.3	9.4	8.3	9.4	8.7	8.3	8.4	7.7
(1, 1.4)	LR	4.2	3.6	4.6	3.9	3.7	4.4	4.0	4.5
	$LR(bc)$	3.8	3.2	4.3	3.7	3.6	4.2	3.9	4.4
	T^2	4.3	3.8	4.7	4.5	4.4	4.6	3.9	4.5
	$T^2(bc)$	5.2	4.6	5.1	5.0	5.0	4.8	4.1	4.5
	T_1	6.7	5.4	5.4	5.6	4.5	4.3	4.1	3.0
	T_N	12.4	11.2	11.4	11.7	10.4	9.7	10.2	8.3
(1, 1.5)	LR	4.4	3.8	4.6	4.2	4.2	4.9	4.6	5.0
	$LR(bc)$	4.0	3.5	4.4	3.9	4.1	4.8	4.3	4.7
	T^2	4.4	4.0	4.9	4.6	4.6	4.9	4.5	5.0
	$T^2(bc)$	5.4	4.8	5.2	5.3	5.3	5.2	4.6	5.3
	T_1	7.3	7.0	5.6	6.3	5.6	4.9	4.4	3.8
	T_N	14.0	13.4	11.6	12.6	11.8	11.7	10.8	9.8
(1, 2)	LR	8.3	7.9	8.5	8.1	6.9	7.2	7.0	7.3
	$LR(bc)$	8.2	7.6	8.4	7.9	6.8	7.0	7.0	7.1
	T^2	9.4	8.9	10.0	9.7	7.4	7.7	7.6	8.0
	$T^2(bc)$	10.8	10.1	11.2	11.0	8.2	8.6	8.6	9.1
	T_1	17.0	14.6	11.7	12.2	10.0	9.0	7.8	5.4
	T_N	28.2	25.4	22.6	23.4	20.8	19.1	18.4	13.8

Table 3.10: Empirical level and power (%) of test statistics LR, $LR(bc)$, T^2, $T^2(bc)$, T_1 and T_N; based on 5000 iterations and $n_1 = 10, n_2 = 10, \alpha = 0.05$

		(c_1, c_2)							
(μ_1, μ_2)	Statistics	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(1, 1)	LR	4.1	4.2	4.2	4.2	4.4	4.6	3.9	4.2
	$LR(bc)$	3.4	3.6	3.5	3.3	3.6	4.2	3.5	4.0
	T^2	3.1	3.6	3.4	3.3	3.5	3.7	3.6	3.8
	$T^2(bc)$	3.1	3.4	3.3	3.2	3.5	4.2	3.9	4.1
	T_1	4.0	4.4	4.4	4.5	4.2	3.8	4.2	4.7
	T_N	6.5	6.9	6.2	6.4	6.4	6.4	6.4	6.6
(1, 1.1)	LR	4.7	4.9	4.7	4.3	4.6	4.6	4.4	4.4
	$LR(bc)$	4.2	4.2	3.9	3.5	4.0	4.3	4.0	4.3
	T^2	3.4	3.9	3.7	3.6	4.1	3.9	3.8	4.2
	$T^2(bc)$	3.5	4.0	3.9	3.4	4.4	4.4	4.2	4.6
	T_1	4.9	4.9	4.6	4.6	4.3	4.5	4.3	4.8
	T_N	7.0	7.0	6.6	7.0	6.7	6.6	6.5	6.9
(1, 1.2)	LR	4.8	5.1	4.8	4.6	4.7	4.8	4.7	4.7
	$LR(bc)$	4.3	4.3	3.9	4.2	4.2	4.4	4.4	4.3
	T^2	4.0	4.7	4.2	3.7	4.2	4.8	4.0	4.6
	$T^2(bc)$	4.1	4.8	4.2	3.9	4.5	5.1	4.3	5.0
	T_1	6.0	6.3	5.9	5.6	4.9	4.7	5.3	4.8
	T_N	8.5	8.7	8.6	7.9	7.6	6.9	7.8	6.9
(1, 1.4)	LR	5.2	5.7	6.6	5.4	5.0	5.9	5.6	5.2
	$LR(bc)$	4.5	5.0	5.6	4.8	4.5	5.3	5.1	4.8
	T^2	4.8	5.5	6.7	5.2	5.1	6.5	6.3	5.0
	$T^2(bc)$	4.8	5.6	6.8	5.3	5.2	6.7	6.6	5.4
	T_1	11.2	10.2	10.1	9.2	8.8	7.9	7.7	6.2
	T_N	14.5	14.0	13.3	12.5	12.3	11.1	10.8	9.1
(1, 1.5)	LR	6.4	6.5	7.7	6.4	6.2	6.9	5.6	5.3
	$LR(bc)$	5.1	5.6	6.8	5.7	5.5	6.1	5.1	5.0
	T^2	6.3	6.7	7.7	6.2	6.5	7.2	6.1	5.9
	$T^2(bc)$	6.6	6.8	8.0	6.5	6.7	7.5	6.4	6.3
	T_1	14.0	13.9	11.3	11.9	10.3	9.5	8.9	6.7
	T_N	18.4	18.0	14.8	16.2	14.5	13.2	13.2	10.2
(1, 2)	LR	15.5	16.6	17.0	11.5	13.6	16.9	11.0	10.5
	$LR(bc)$	13.9	14.8	15.1	10.4	12.2	15.0	10.1	9.2
	T^2	15.6	16.6	17.7	11.7	13.9	17.3	12.3	11.8
	$T^2(bc)$	15.9	17.0	18.0	12.0	14.5	17.8	13.0	12.7
	T_1	37.2	33.2	29.8	29.9	25.9	23.7	19.7	15.6
	T_N	43.8	39.6	37.4	36.8	33.0	31.1	26.8	22.2

Table 3.11: Empirical level and power (%) of test statistics LR, $LR(bc)$, T^2, $T^2(bc)$, T_1 and T_N; based on 5000 iterations and $n_1 = 20, n_2 = 20, \alpha = 0.05$

		(c_1, c_2)							
(μ_1, μ_2)	Statistics	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(1, 1)	LR	3.0	3.2	3.1	3.0	3.0	3.1	3.3	3.7
	$LR(bc)$	2.2	2.3	2.4	2.4	2.4	2.6	2.6	3.0
	T^2	2.2	2.1	2.5	2.3	2.1	2.4	2.3	2.8
	$T^2(bc)$	2.2	2.1	2.5	2.4	2.1	2.4	2.4	2.9
	T_1	5.2	5.0	5.0	4.3	4.8	4.9	4.5	4.9
	T_N	6.1	5.6	6.2	4.9	5.8	5.7	5.4	5.9
(1, 1.1)	LR	3.0	3.2	3.4	3.3	3.0	3.4	3.4	3.9
	$LR(bc)$	2.4	2.4	2.9	2.5	2.5	2.8	2.8	3.4
	T^2	2.4	2.3	2.6	2.7	2.3	2.6	2.5	2.9
	$T^2(bc)$	2.4	2.3	2.5	2.7	2.3	2.5	2.5	2.9
	T_1	5.4	5.3	5.2	5.0	6.1	5.7	5.1	5.3
	T_N	6.2	6.2	6.3	6.1	7.0	6.8	5.8	6.2
(1, 1.2)	LR	4.1	4.0	4.7	3.4	4.6	4.9	4.7	4.4
	$LR(bc)$	2.9	3.0	3.5	2.4	3.3	3.7	3.6	3.3
	T^2	3.7	3.6	4.2	2.7	3.8	4.3	3.7	3.5
	$T^2(bc)$	3.5	3.5	4.1	2.6	3.9	4.3	3.8	3.6
	T_1	8.7	8.1	7.0	8.2	7.1	7.7	7.0	6.9
	T_N	9.7	9.2	8.3	9.4	8.3	8.9	8.2	8.0
(1, 1.4)	LR	10.0	9.8	9.9	8.8	9.9	10.0	9.1	8.9
	$LR(bc)$	7.0	7.2	7.8	6.5	7.2	7.5	6.4	6.9
	T^2	7.9	8.2	8.1	6.3	7.0	7.9	6.3	6.7
	$T^2(bc)$	7.7	8.0	8.0	6.3	7.1	8.1	6.5	7.0
	T_1	18.1	17.4	16.4	17.0	15.4	14.5	13.7	10.1
	T_N	20.2	19.5	18.5	19.1	17.2	16.6	15.5	12.1
(1, 1.5)	LR	13.2	13.6	14.5	13.3	12.9	14.6	12.7	11.4
	$LR(bc)$	9.6	10.2	11.1	9.6	9.7	10.8	9.0	8.9
	T^2	9.8	10.7	11.2	9.1	9.6	11.5	8.4	8.7
	$T^2(bc)$	9.6	10.8	11.2	9.1	9.7	11.9	8.5	9.1
	T_1	26.2	24.3	23.0	21.7	20.8	20.4	17.3	12.9
	T_N	28.7	26.5	25.6	24.0	23.4	23.1	20.0	15.3
(1, 2)	LR	47.7	47.4	49.0	43.1	42.9	41.8	37.3	33.5
	$LR(bc)$	37.9	37.9	41.0	34.3	34.4	34.4	30.3	28.0
	T^2	32.5	34.3	37.8	27.2	29.8	31.8	23.2	25.3
	$T^2(bc)$	32.8	34.5	38.1	28.0	30.5	32.7	23.9	26.4
	T_1	68.3	65.4	61.2	59.8	56.5	51.2	45.8	36.3
	T_N	70.9	68.3	64.5	62.6	59.6	55.0	49.7	40.1

Table 3.12: Empirical level and power (%) of test statistics LR, $LR(bc)$, T^2, $T^2(bc)$, T_1 and T_N; based on 5000 iterations and $n_1 = 50, n_2 = 50, \alpha = 0.05$

		(c_1, c_2)							
(μ_1, μ_2)	Statistics	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(1, 1)	LR	3.7	3.3	4.1	4.3	4.5	4.1	5.3	4.7
	$LR(bc)$	3.4	3.0	3.7	4.1	4.3	4.0	5.1	4.6
	T^2	3.4	3.2	3.8	3.9	4.0	3.9	4.7	3.9
	$T^2(bc)$	3.4	3.2	3.7	3.9	4.0	3.9	4.7	3.9
	T_1	4.7	4.5	4.9	5.2	5.0	5.6	4.9	4.2
	T_N	5.1	4.9	5.2	5.6	5.3	5.9	5.1	4.8
(1, 1.1)	LR	5.2	6.2	6.0	6.2	6.4	6.9	6.7	6.3
	$LR(bc)$	4.7	5.7	5.5	6.0	6.1	6.4	6.5	6.2
	T^2	5.0	6.1	5.9	5.6	6.1	6.4	6.1	6.1
	$T^2(bc)$	4.9	6.0	5.9	5.6	6.1	6.4	6.1	6.2
	T_1	7.3	7.3	7.7	6.9	7.3	6.1	6.3	6.3
	T_N	7.6	7.7	8.1	7.4	7.9	6.5	6.7	6.5
(1, 1.2)	LR	12.0	11.4	12.0	12.6	12.5	11.5	11.8	11.6
	$LR(bc)$	10.8	10.1	11.1	11.9	11.9	11.0	11.3	11.2
	T^2	11.6	11.2	11.9	11.6	12.1	11.2	11.0	11.1
	$T^2(bc)$	11.5	11.0	11.8	11.5	12.1	11.2	11.1	11.3
	T_1	14.6	15.0	13.0	13.2	12.7	11.3	10.6	9.0
	T_N	15.2	15.6	13.6	14.0	13.2	11.8	11.1	9.7
(1, 1.4)	LR	34.8	36.6	36.5	34.1	33.5	32.5	30.6	27.3
	$LR(bc)$	32.7	34.4	34.6	32.9	32.2	31.4	30.0	26.8
	T^2	33.6	35.5	35.9	32.4	32.4	31.7	28.7	26.9
	$T^2(bc)$	33.6	35.3	35.8	32.3	32.4	31.7	28.7	27.1
	T_1	41.5	39.9	37.2	36.5	34.6	32.9	28.5	25.5
	T_N	42.4	40.9	38.1	37.6	35.5	34.0	29.8	26.4
(1, 1.5)	LR	51.5	50.0	48.6	48.2	45.2	45.2	42.9	39.0
	$LR(bc)$	48.9	47.5	46.6	46.6	43.8	43.9	42.2	38.6
	T^2	50.1	48.7	47.9	46.2	44.0	44.1	40.8	38.5
	$T^2(bc)$	49.9	48.5	47.7	46.1	43.9	44.1	40.9	38.8
	T_1	57.1	56.2	52.4	50.8	47.6	46.3	40.7	35.2
	T_N	57.9	57.2	53.6	51.7	48.7	47.6	41.4	36.1
(1, 2)	LR	96.3	95.5	95.3	94.1	92.2	90.9	88.3	84.4
	$LR(bc)$	94.6	94.4	93.9	93.2	91.4	90.2	87.6	84.2
	T^2	96.0	95.4	94.9	93.1	91.5	90.5	87.1	84.2
	$T^2(bc)$	95.8	95.2	94.9	93.1	91.5	90.6	87.2	84.3
	T_1	97.6	97.2	95.7	94.4	93.0	91.1	87.9	81.4
	T_N	97.7	97.4	96.0	94.7	93.5	91.5	88.5	82.4

Table 3.13: Estimated bias of the estimates of dispersion parameters, c_1 and c_2, under null and alternative hypotheses; based on 5000 iterations and $n_1 = 5, n_2 = 5$

		(c_1, c_2)							
(μ_1, μ_2)	Par†	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(1, 1)	c_{10}	-0.0172	-0.0193	-0.0214	0.0084	-0.0122	-0.0141	-0.0042	0.0093
	c_{20}	-0.0153	-0.0110	0.0003	-0.0054	0.0101	0.0263	0.0134	-0.0416
	c_{1a}	0.0345	0.0360	-0.0373	-0.0339	-0.0346	-0.0343	-0.0281	0.0514
	c_{2a}	-0.0510	-0.0555	-0.0628	0.0626	-0.0701	-0.0759	0.0822	-0.0633
(1, 1.1)	c_{10}	0.0122	-0.0154	-0.0191	0.0049	0.0075	-0.0145	0.0051	-0.0067
	c_{20}	-0.0150	-0.0099	0.0008	-0.0062	0.0035	0.0226	0.0307	0.1054
	c_{1a}	-0.0154	-0.0209	-0.0248	-0.0097	-0.0103	-0.0182	-0.0011	-0.0077
	c_{2a}	-0.0395	-0.0450	-0.0549	-0.0546	-0.0628	-0.0707	-0.0769	-0.0892
(1, 1.2)	c_{10}	-0.0011	-0.0073	-0.0138	0.0035	-0.0017	-0.0079	0.0116	-0.0041
	c_{20}	-0.0103	-0.0020	0.0127	-0.0026	0.0087	0.0292	0.0300	0.0949
	c_{1a}	0.0347	0.0199	0.0097	0.0410	0.0353	0.0203	0.0554	0.0307
	c_{2a}	-0.0326	-0.0385	-0.0510	-0.0491	-0.0592	-0.0684	-0.0767	-0.0927
(1, 1.4)	c_{10}	0.0274	0.0122	-0.0037	0.0190	0.0085	-0.0012	0.0155	-0.0057
	c_{20}	-0.0047	0.0063	0.0371	0.0058	0.0260	0.0587	0.0496	0.0855
	c_{1a}	-0.0072	0.0234	0.0192	0.0083	0.0319	0.0251	0.0111	-0.0328
	c_{2a}	-0.0264	-0.0350	-0.0503	-0.0463	-0.0585	-0.0714	-0.0785	-0.1019
(1, 1.5)	c_{10}	0.0475	0.0339	0.0045	0.0307	0.0137	0.0026	0.0180	-0.0021
	c_{20}	-0.0104	0.0032	0.0660	0.0088	0.0435	0.0956	0.0705	0.0248
	c_{1a}	0.0217	0.0055	0.0224	0.0313	0.0068	0.0049	0.0337	0.0295
	c_{2a}	-0.0208	-0.0325	-0.0529	-0.0447	-0.0621	-0.0749	-0.0834	-0.1100
(1, 2)	c_{10}	0.0471	0.0417	0.0080	0.0345	0.0146	0.0007	0.0152	-0.0069
	c_{20}	-0.0126	-0.0040	0.0357	0.0027	0.0539	0.0994	0.0699	0.0051
	c_{1a}	0.0058	0.0179	0.0018	0.0274	0.0318	0.0239	0.0873	0.0298
	c_{2a}	-0.0199	-0.0279	-0.0572	-0.0449	-0.0645	-0.0814	-0.0876	-0.0412

Par† = Parameters
c_{10} is c_1 under null hypothesis
c_{20} is c_2 under null hypothesis
c_{1a} is c_1 under alternative hypothesis
c_{2a} is c_2 under alternative hypothesis

Table 3.14: Estimated bias of the estimates of dispersion parameters, c_1 and c_2, under null and alternative hypotheses; based on 5000 iterations and $n_1 = 10, n_2 = 10$

		(c_1, c_2)							
(μ_1, μ_2)	Par†	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(1, 1)	c_{10}	-0.0009	-0.0035	0.0082	0.0234	-0.0614	0.0202	-0.0348	0.0624
	c_{20}	0.0597	-0.0627	-0.0742	0.1052	0.0398	-0.0095	0.1011	-0.0099
	c_{1a}	-0.0489	0.0276	-0.0050	-0.0246	-0.0737	0.0426	0.0568	0.0537
	c_{2a}	-0.0087	0.0211	0.0085	-0.0920	-0.0299	0.1000	0.0071	0.0084
(1, 1.1)	c_{10}	-0.0057	0.0034	-0.0063	0.0069	0.0206	0.0177	0.0623	0.0457
	c_{20}	0.0403	0.0527	0.0916	-0.0913	0.1052	0.0479	0.0085	0.0648
	c_{1a}	-0.0343	-0.0369	-0.0432	-0.0458	-0.0192	-0.0167	0.0147	0.0115
	c_{2a}	-0.0041	-0.0151	-0.0322	0.0428	-0.0819	-0.0777	-0.0748	-0.0471
(1, 1.2)	c_{10}	-0.0027	-0.0019	-0.0087	0.0263	-0.0077	0.0093	0.0596	0.0069
	c_{20}	0.0333	0.0399	0.0615	0.0554	-0.0721	0.0449	0.0568	-0.0515
	c_{1a}	-0.0485	-0.0311	-0.0460	-0.0122	0.0094	-0.0166	0.0270	-0.0314
	c_{2a}	-0.0393	-0.0281	0.0043	-0.0456	0.0329	-0.0492	-0.0059	0.0522
(1, 1.4)	c_{10}	0.0073	0.0026	-0.0041	0.0189	0.0183	0.0191	0.0106	0.0073
	c_{20}	0.0044	0.0256	0.0490	0.0522	0.0669	0.1103	0.0067	-0.0248
	c_{1a}	0.0036	-0.0097	-0.0144	0.0163	0.0101	0.0153	-0.0083	0.0418
	c_{2a}	-0.0218	-0.0134	-0.0082	-0.0456	-0.0384	-0.0524	0.0088	0.0122
(1, 1.5)	c_{10}	0.0023	-0.0019	-0.0006	0.0224	0.0219	0.0194	0.0515	-0.0233
	c_{20}	0.0110	0.0234	0.0436	0.0448	0.0557	0.0880	0.0333	0.0843
	c_{1a}	-0.0125	-0.0057	-0.0008	0.0278	0.0261	0.0293	0.0576	-0.0087
	c_{2a}	0.0091	0.0024	0.0257	-0.0427	-0.0318	-0.0422	-0.0083	0.0069
(1, 2)	c_{10}	0.0083	0.0048	0.0013	0.0288	0.0279	-0.0038	-0.0088	0.0471
	c_{20}	0.0050	0.0186	0.0456	0.0344	0.0558	-0.0061	-0.0010	0.1846
	c_{1a}	0.0492	0.0391	0.0352	0.0742	0.0627	0.0293	-0.0615	0.1071
	c_{2a}	-0.0587	-0.0590	-0.0604	-0.0599	-0.0596	0.0099	0.1008	-0.0520

Par† = Parameters
c_{10} is c_1 under null hypothesis
c_{20} is c_2 under null hypothesis
c_{1a} is c_1 under alternative hypothesis
c_{2a} is c_2 under alternative hypothesis

Table 3.15: Estimated bias of the estimates of dispersion parameters, c_1 and c_2, under null and alternative hypotheses; based on 5000 iterations and $n_1 = 20, n_2 = 20$

		(c_1, c_2)							
(μ_1, μ_2)	Par†	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(1, 1)	c_{10}	-0.0113	-0.0113	-0.0117	-0.0016	-0.0015	-0.0032	0.0139	0.0085
	c_{20}	-0.0082	-0.0058	0.0030	0.0016	0.0085	0.0201	0.0271	0.0660
	c_{1a}	-0.0279	-0.0247	-0.0282	-0.0172	-0.0165	-0.0170	-0.0056	-0.0087
	c_{2a}	-0.0492	-0.0492	-0.0476	-0.0476	-0.0462	-0.0448	-0.0424	-0.0363
(1, 1.1)	c_{10}	-0.0088	-0.0098	-0.0109	0.0012	-0.0007	-0.0015	0.0143	0.0103
	c_{20}	-0.0111	-0.0067	0.0010	-0.0010	0.0074	0.0161	0.0243	0.0610
	c_{1a}	-0.0148	-0.0181	-0.0224	-0.0103	-0.0107	-0.0094	0.0026	0.0014
	c_{2a}	-0.0461	-0.0451	-0.0441	-0.0443	-0.0433	-0.0420	-0.0404	-0.0339
(1, 1.2)	c_{10}	-0.0083	-0.0093	-0.0101	0.0011	-0.0002	-0.0025	0.0147	0.0105
	c_{20}	-0.0099	-0.0073	0.0005	-0.0028	0.0086	0.0141	0.0237	0.0547
	c_{1a}	-0.0159	-0.0152	-0.0139	-0.0041	-0.0031	-0.0058	0.0117	0.0137
	c_{2a}	-0.0425	-0.0424	-0.0414	-0.0418	-0.0340	-0.0399	-0.0382	-0.0336
(1, 1.4)	c_{10}	-0.0044	-0.0054	-0.0069	0.0046	0.0038	0.0016	0.0171	0.0139
	c_{20}	-0.0111	-0.0070	0.0000	-0.0034	0.0039	0.0134	0.0198	0.0503
	c_{1a}	0.0059	0.0069	0.0035	0.0216	0.0222	0.0214	0.0411	0.0396
	c_{2a}	-0.0374	-0.0374	-0.0371	-0.0372	-0.0367	-0.0362	-0.0353	-0.0322
(1, 1.5)	c_{10}	-0.0029	-0.0040	-0.0047	0.0076	0.0056	0.0045	0.0211	0.0132
	c_{20}	-0.0093	-0.0068	0.0010	-0.0030	0.0067	0.0133	0.0169	0.0490
	c_{1a}	0.0161	0.0164	0.0151	0.0394	0.0331	0.0336	0.0626	0.0421
	c_{2a}	-0.0354	-0.0353	-0.0353	-0.0351	-0.0350	-0.0345	-0.0341	-0.0311
(1, 2)	c_{10}	0.0116	0.0087	0.0059	0.0216	0.0197	0.0143	0.0344	0.0239
	c_{20}	-0.0089	-0.0058	0.0007	-0.0029	0.0049	0.0117	0.0138	0.0440
	c_{1a}	0.0019	0.0038	-0.0009	0.0053	0.0318	0.0077	0.0246	0.0344
	c_{2a}	-0.0278	-0.0283	-0.0289	-0.0289	-0.0294	-0.0299	-0.0300	-0.0292

Par† = Parameters

c_{10} is c_1 under null hypothesis

c_{20} is c_2 under null hypothesis

c_{1a} is c_1 under alternative hypothesis

c_{2a} is c_2 under alternative hypothesis

Table 3.16: Estimated bias of the estimates of dispersion parameters, c_1 and c_2, under null and alternative hypotheses; based on 5000 iterations and $n_1 = 50, n_2 = 50$

		(c_1, c_2)							
(μ_1, μ_2)	Par†	(.05, .05)	(.05, .10)	(.05, .20)	(.20, .20)	(.20, .30)	(.20, .40)	(.40, .50)	(.40, .80)
(1, 1)	c_{10}	-0.0064	-0.0068	-0.0070	-0.0030	-0.0036	-0.0040	0.0013	-0.0003
	c_{20}	-0.0065	-0.0054	-0.0023	-0.0032	-0.0001	0.0030	0.0055	0.0179
	c_{1a}	-0.0149	-0.0154	-0.0143	-0.0123	-0.0125	-0.0129	-0.0112	-0.0115
	c_{2a}	-0.0195	-0.0193	-0.0187	-0.0188	-0.0182	-0.0178	-0.0172	-0.0148
(1, 1.1)	c_{10}	-0.0061	-0.0062	-0.0069	-0.0028	-0.0034	-0.0038	0.00149	-0.0003
	c_{20}	-0.0063	-0.0052	-0.0025	-0.0031	-0.0003	0.0028	0.0050	0.0168
	c_{1a}	-0.0127	-0.0120	-0.0123	-0.0098	-0.0089	-0.0095	-0.0078	-0.0091
	c_{2a}	-0.0178	-0.0177	-0.0174	-0.0174	-0.0170	-0.0166	-0.0161	-0.0143
(1, 1.2)	c_{10}	-0.0054	-0.0057	-0.0064	-0.0024	-0.0031	-0.0036	0.0017	-0.0001
	c_{20}	-0.0062	-0.0051	-0.0025	-0.0032	-0.0005	0.0027	0.0051	0.0159
	c_{1a}	-0.0059	-0.0075	-0.0074	-0.0036	-0.0039	-0.0048	-0.0024	-0.0043
	c_{2a}	-0.0165	-0.0165	-0.0163	-0.0162	-0.0160	-0.0157	-0.0153	-0.0139
(1, 1.4)	c_{10}	-0.0039	-0.0045	-0.0050	-0.0010	-0.0017	-0.0025	0.0028	0.0003
	c_{20}	-0.0059	-0.0048	-0.0023	-0.0031	-0.0003	0.0026	0.0043	0.0153
	c_{1a}	0.0115	0.0115	0.0091	0.0141	0.0138	0.0088	0.0114	0.0080
	c_{2a}	-0.0145	-0.0144	-0.0144	-0.0144	-0.0144	-0.0143	-0.0141	-0.0132
(1, 1.5)	c_{10}	0.0030	-0.0037	-0.0045	-0.0004	-0.0010	0.0017	0.0037	0.0016
	c_{20}	-0.0058	-0.0047	-0.0020	-0.0030	-0.0003	0.0027	0.0041	0.0146
	c_{1a}	0.0232	0.0209	0.0202	0.0253	0.0225	0.0204	0.0240	0.0172
	c_{2a}	-0.0136	-0.0115	-0.0137	-0.0077	-0.0138	-0.0138	-0.0137	-0.0130
(1, 2)	c_{10}	0.0036	0.0028	0.0010	0.0075	0.0056	0.0037	0.0101	0.0065
	c_{20}	-0.0052	-0.0043	-0.0018	-0.0034	-0.0009	0.0020	0.0026	0.0132
	c_{1a}	0.0221	0.0052	0.0141	0.0091	0.0078	0.0438	0.0243	0.0066
	c_{2a}	-0.0107	-0.0109	-0.0113	-0.0112	-0.0116	-0.0119	-0.0120	-0.0121

Par† = Parameters

c_{10} is c_1 under null hypothesis

c_{20} is c_2 under null hypothesis

c_{1a} is c_1 under alternative hypothesis

c_{2a} is c_2 under alternative hypothesis

Table 3.17: Number of tumors for rats in treatment groups 1 and 2

No. of Tumors	Groups	
	Treatment Group	Control Group
0	2	0
1	7	4
2	4	2
3	2	3
4	2	2
5	4	1
6	2	2
7	0	2
8	0	0
9	0	3
10	0	1
11	0	3
12	0	1
13	0	1
$\bar{y}$	2.652	6.04
s^2	3.618	14.918

Table 3.18: Number of cycles required for smoker and non-smoker women to get pregnant

No. of Cycles	Smoker	Non-smoker
1	29	198
2	16	107
3	17	55
4	4	38
5	3	18
6	8	22
7	4	7
8	5	9
9	1	5
10	1	3
11	1	6
12	3	6
> 12	7	12
$\bar{y}$	4.2020	2.9650
s^2	13.8567	8.1245

Chapter 4

Testing Equality of Two Beta Binomial Proportions in the Presence of Unequal Dispersion Parameters

4.1 Introduction

Scientists in various areas, for example, toxicology (Weil (1970); Kleinman (1973); Williams (1975); Paul (1982)) and other similar fields (Crowder (1978); Otake and Prentice (1984); Donovan et al. (1994); Gibson and Austin (1996)), frequently encounter data in the form of proportions. Binomial model is a basic model to deal with the data of such kind. It happens quite often that the proportion data exhibit greater variability than predicted by the simple binomial model and the reason for this variability depends on the form of study. Weil (1970) observes that if the experimental units of the data are litters of animals then 'litter effect', that is, the tendency of animals in the same litter to respond more similarly than animals from different litters contribute to greater variability than predicted by the simple model. This effect of litter is known as intra-litter or intra-correlation coefficient. The assumption of some specific parametric model is frequently used to deal with correlated binomial data. Beta binomial model is a widely used model to accommodate the over-dispersion in proportion data (Williams (1975); Crowder (1978)). An extension to beta binomial model to include over-dispersion as well as under-dispersion was presented by Prentice (1986). In situations where the data are in the form of proportions with possible over-dispersion, we may be interested in testing the equality of proportions of a certain characteristic in two groups. Data of this form can be described as follows.

Suppose that there are 2 treatment groups, i^{th} group having m_i litters, $i = 1, 2$. Then, the data are of the form

Groups	Proportions
1	$y_{11}/n_{11}, y_{12}/n_{12}, \cdots, y_{1j}/n_{1j}, \cdots, y_{1m_1}/n_{1m_1}$
2	$y_{21}/n_{21}, y_{22}/n_{22}, \cdots, y_{2j}/n_{2j}, \cdots, y_{2m_2}/n_{2m_2}$

Here the size of the j^{th} litter in the i^{th} group is n_{ij} of which y_{ij} respond to the i^{th} treatment. Now, given a parameter p, $y_{ij}|p_i \sim binomial(n_{ij}, p_i)$ and p_i is a beta random variable having probability function

$$f(p_i) = \frac{1}{B(\alpha_i, \beta_i)} p_i^{\alpha_i - 1}(1 - p_i)^{\beta_i - 1}.$$

Then, unconditionally, y_{ij} has a beta-binomial distribution with probability function

$$Pr(y_{ij}) = \binom{n_{ij}}{y_{ij}} \frac{B(y_{ij} + \alpha_i, n_{ij} + \beta_i - y_{ij})}{B(\alpha_i, \beta_i)}. \tag{4.1}$$

The mean and variance of y_{ij} are $n_{ij}\left(\frac{\alpha_i}{\alpha_i+\beta_i}\right)$ and $\frac{n_{ij}\alpha_i\beta_i(\alpha_i+\beta_i+n_{ij})}{(\alpha_i+\beta_i)^2(\alpha_i+\beta_i+1)}$ respectively. Now, define $\pi_i = \frac{\alpha_i}{\alpha_i+\beta_i}$, $\omega_i = \frac{1}{\alpha_i+\beta_i}$ and $\theta_i = \frac{\omega_i}{1+\omega_i}$. Then the mean and variance of y_{ij} can be expressed as $n_{ij}\pi_i$ and $n_{ij}\pi_i(1-\pi_i)\left[1+(n_{ij}-1)\theta_i\right]$.

Taking the above reparameterization into account, the probability function can be written as

$$Pr\left(y_{ij}|\pi_i,\theta_i\right) = \binom{n_{ij}}{y_{ij}} \frac{\prod_{r=0}^{y_{ij}-1}\left[\pi_i(1-\theta_i)+r\theta_i\right]\prod_{r=0}^{n_{ij}-y_{ij}-1}\left[(1-\pi_i)(1-\theta_i)+r\theta_i\right]}{\prod_{r=0}^{n_{ij}-1}\left[(1-\theta_i)+r\theta_i\right]}. \quad (4.2)$$

We denote the probability function of the beta-binomial distribution in equation (4.2) as $BB(\pi_i,\theta_i)$. Our purpose is to test $H_0 : \pi_1 = \pi_2$ with θ_1 and θ_2 being unspecified.

Several parametric and semi-parametric procedures are available for testing homogeneity of proportions in presence of over-dispersion. Paul and Islam (1995) developed tests for testing the equality of several proportions when the over-dispersion parameters are equal. We extend this idea to testing equality of two beta-binomial proportions for possibly unequal over-dispersion parameters and develop parametric as well as semi-parametric tests. The parametric tests that we develop are a likelihood ratio test, a $C(\alpha)$ test based on the maximum likelihood estimates of nuisance parameters, a $C(\alpha)$ test based on the Kleinman's (1973) method of moments estimates of the nuisance parameters. Further, we develop a $C(\alpha)$ test based on the quasi-likelihood and the method of moments estimates of the nuisance parameters by Breslow (1990), a $C(\alpha)$ test based on the quasi-likelihood and the method of moments estimates of the nuisance parameters by Srivastava and Wu (1993), a $C(\alpha)$ test based on extended quasi-likelihood estimates of the nuisance parameters. We also develop two more statistics, namely, the Rao-Scott and the adjusted Rao-Scott statistics by following Rao and Scott (1992). All eight statistics are compared, in terms of empirical size and power, using a simulation study.

4.2 Parametric Tests

4.2.1 The Likelihood Ratio Test

Let $y_{i1}, y_{i2}, \cdots, y_{im_i}$ be a sample of size m_i from group i $(i = 1, 2)$. Then, $y_{ij} \sim BB(\pi_i, \theta_i)$ and the log-likelihood, apart from a constant, can be written as

$$\begin{aligned} l_i = \sum_{j=1}^{m_i} \Bigg[\sum_{r=0}^{y_{ij}-1} \log\{(1-\theta_i)\pi_i + r\theta_i\} + \sum_{r=0}^{n_{ij}-y_{ij}-1} \log\{(1-\theta_i)(1-\pi_i) + r\theta_i\} \\ - \sum_{r=0}^{n_{ij}-1} \log\{(1-\theta_i) + r\theta_i\} \Bigg]. \end{aligned} \tag{4.3}$$

Under H_A, the estimates of the parameters π_i and θ_i $(i = 1, 2)$ can be obtained by directly maximizing the above log-likelihood or by solving the pair of equations

$$\sum_{j=1}^{m_i} \left[\sum_{r=0}^{y_{ij}-1} \frac{(1-\theta_i)}{(1-\theta_i)\pi_i + r\theta_i} - \sum_{r=0}^{n_{ij}-y_{ij}-1} \frac{(1-\theta_i)}{(1-\theta_i)(1-\pi_i) + r\theta_i} \right] = 0$$

and

$$\sum_{j=1}^{m_i} \left[\sum_{r=0}^{y_{ij}-1} \frac{r-\pi_i}{(1-\theta_i)\pi_i + r\theta_i} + \sum_{r=0}^{n_{ij}-y_{ij}-1} \frac{r+\pi_i-1}{(1-\theta_i)(1-\pi_i) + r\theta_i} - \sum_{r=0}^{n_{ij}-1} \frac{r-1}{(1-\theta_i) + r\theta_i} \right] = 0$$

simultaneously.

Under the null hypothesis the log-likelihood function of the parameters π, θ_1 and θ_2, apart from a constant, can be written as

$$l_0 = \sum_{i=1}^{2}\sum_{j=1}^{m_i}\left[\sum_{r=0}^{y_{ij}-1}\log\{(1-\theta_i)\pi + r\theta_i\} + \sum_{r=0}^{n_{ij}-y_{ij}-1}\log\{(1-\theta_i)(1-\pi)+r\theta_i\} - \sum_{r=0}^{n_{ij}-1}\log\{(1-\theta_i)+r\theta_i\}\right].$$

Then the maximum likelihood estimates of the parameters π, θ_1, and θ_2 can be obtained directly by maximizing the above log-likelihood or by solving the following estimating equations

$$\sum_{i=1}^{2}\sum_{j=1}^{m_i}\left[\sum_{r=0}^{y_{ij}-1}\frac{(1-\theta_i)}{(1-\theta_i)\pi + r\theta_i} - \sum_{r=0}^{n_{ij}-y_{ij}-1}\frac{(1-\theta_i)}{(1-\theta_i)(1-\pi)+r\theta_i}\right] = 0,$$

$$\sum_{j=1}^{m_1}\left[\sum_{r=0}^{y_{1j}-1}\left\{\frac{r-\pi}{(1-\theta_1)\pi + r\theta_1}\right\} + \sum_{r=0}^{n_{1j}-y_{1j}-1}\left\{\frac{r+\pi-1}{(1-\theta_1)(1-\pi)+r\theta_1}\right\} - \sum_{r=0}^{n_{1j}-1}\left\{\frac{r-1}{(1-\theta_1)+r\theta_1}\right\} = 0\right.$$

and

$$\sum_{j=1}^{m_2}\left[\sum_{r=0}^{y_{2j}-1}\left\{\frac{r-\pi}{(1-\theta_2)\pi + r\theta_2}\right\} + \sum_{r=0}^{n_{2j}-y_{2j}-1}\left\{\frac{r+\pi-1}{(1-\theta_2)(1-\pi)+r\theta_2}\right\} - \sum_{r=0}^{n_{2j}-1}\left\{\frac{r-1}{(1-\theta_2)+r\theta_2}\right\} = 0\right.$$

simultaneously.

Let $\hat{l}_1$ be the maximized log-likelihood under the alternative hypothesis and $\hat{l}_0$ be the maximized log-likelihood under the null hypothesis. Then, the likelihood ratio statistic for testing $H_0 : \pi_1 = \pi_2$ against $H_A : \pi_1 \neq \pi_2$ with θ_1 and θ_2 unspecified, is $LR = 2(\hat{l}_1 - \hat{l}_0)$, which is distributed asymptotically (as $m \to \infty$, where $m = m_1 + m_2$) as a chi-squared with 1 degree of freedom.

4.2.2 $C(\alpha)$ (Score) Test

Suppose the alternative hypothesis is represented by $\pi_i = \pi + \phi_i$ with $\phi_2 = 0$. Then the null hypothesis $H_0 : \pi_1 = \pi_2$ reduces to $H_0 : \phi_1 = 0$ with π, θ_1 and θ_2 treated as nuisance parameters. In order to derive the $C(\alpha)$ statistic to test the equality of several odds ratios, Tarone (1985) used this technique. Barnwal and Paul (1988) used this technique to derive $C(\alpha)$ statistic for testing the equality of several Poisson means in the presence of negative binomial over-dispersion. Paul and Islam (1995) applied this procedure for testing the equality of several proportions in the presence of common over/under dispersion. With this reparameterizaton, the log-likelihood function under the alternative hypothesis, apart from a constant, can be written as

$$l = \sum_{i=1}^{2} \sum_{j=1}^{m_i} \left[\sum_{r=0}^{y_{ij}-1} \log \{(1-\theta_i)(\pi + \phi_i) + r\theta_i\} + \sum_{r=0}^{n_{ij}-y_{ij}-1} \log \{(1-\theta_i)(1-\pi-\phi_i) + r\theta_i\} \right. \\ \left. - \sum_{r=0}^{n_{ij}-1} \log \{(1-\theta_i) + r\theta_i\} \right]. \tag{4.4}$$

Now, define $\phi = \phi_1$, $\delta = (\delta_1,\ \delta_2,\ \delta_3)' = (\pi,\ \theta_1,\ \theta_2)'$, $\psi_1 = \left.\dfrac{\partial l}{\partial \phi_1}\right|_{\phi=0}$, $\gamma_1 = \left.\dfrac{\partial l}{\partial \delta_1}\right|_{\phi=0}$, $\gamma_2 = \left.\dfrac{\partial l}{\partial \delta_2}\right|_{\phi=0}$ and

$\gamma_3 = \frac{\partial l}{\partial \delta_3}\Big|_{\phi=0}$.

Let $\hat{\delta}$ be a $\sqrt{m}$ (m is the number of litters in a group) consistent estimator of δ. Then the $C(\alpha)$ statistic is based on the adjusted score

$$S = \psi_1 - \beta_1\gamma_1 - \beta_2\gamma_2 - \beta_3\gamma_3,$$

where β_1, β_2 and β_3 are partial regression coefficient of ψ_1 on γ_1, ψ_1 on γ_2, and ψ_1 on γ_3, respectively.

The structure of dispersion matrix of $(\phi,\ \pi,\ \theta_1,\ \theta_2)$ is

$$V = \begin{pmatrix} D & A \\ A' & B \end{pmatrix}$$

and the regression coefficients $\beta = (\beta_1, \beta_2, \beta_3) = AB^{-1}$ (Neyman (1959)), where D is 1×1, A is 1×3 and B is 3×3 with elements

$$D_{11} = E\left[-\frac{\partial^2 l}{\partial \phi_1^2}\Big|_{\phi=0}\right],$$

$$A_{11} = E\left[-\frac{\partial^2 l}{\partial \phi_1 \partial \pi}\Big|_{\phi=0}\right],\ A_{12} = E\left[-\frac{\partial^2 l}{\partial \phi_1 \partial \theta_1}\Big|_{\phi=0}\right],\ A_{13} = E\left[-\frac{\partial^2 l}{\partial \phi_1 \partial \theta_2}\Big|_{\phi=0}\right],$$

$$B_{11} = E\left[-\frac{\partial^2 l}{\partial \pi^2}\Big|_{\phi=0}\right],\ B_{12} = B_{21} = E\left[-\frac{\partial^2 l}{\partial \pi \partial \theta_1}\Big|_{\phi=0}\right],\ B_{13} = B_{31} = E\left[-\frac{\partial^2 l}{\partial \pi \partial \theta_2}\Big|_{\phi=0}\right],$$

$$B_{22} = E\left[-\frac{\partial^2 l}{\partial \theta_1^2}\Big|_{\phi=0}\right],\ B_{23} = B_{32} = E\left[-\frac{\partial^2 l}{\partial \theta_1 \partial \theta_2}\Big|_{\phi=0}\right] \text{ and}$$

$$B_{33} = E\left[-\frac{\partial^2 l}{\partial \theta_2^2}\Big|_{\phi=0}\right].$$

Substituting $\sqrt{m}$ consistent estimate of δ, that is, $\hat{\delta}$ in S, D, A and B, the $C(\alpha)$ statistic can be obtained as $S'(D - AB^{-1}A')^{-1}S$, which is approximately distributed as a chi-squared with 1 degree of freedom (Neyman (1959); Neyman and Scott (1966); Moran (1970)). If the maximum likelihood estimate of δ is used then $S = \psi_1$, and the $C(\alpha)$ statistic reduces to score statistic (Rao (1948)). Using the log-likelihood (4.4) we obtain

$$\psi_1 = \sum_{j=1}^{m_1} \left[\sum_{r=0}^{y_{1j}-1} \left\{ \frac{(1-\theta_1)}{(1-\theta_1)\,\pi + r\theta_1} \right\} - \sum_{r=0}^{n_{1j}-y_{1j}-1} \left\{ \frac{(1-\theta_1)}{(1-\theta_1)\,(1-\pi) + r\theta_1} \right\} \right],$$

$$\gamma_1 = \sum_{i=1}^{2} \sum_{j=1}^{m_i} \left[\sum_{r=0}^{y_{ij}-1} \left\{ \frac{(1-\theta_i)}{(1-\theta_i)\,\pi + r\theta_i} \right\} - \sum_{r=0}^{n_{ij}-y_{ij}-1} \left\{ \frac{(1-\theta_i)}{(1-\theta_i)\,(1-\pi) + r\theta_i} \right\} \right],$$

$$\gamma_2 = \sum_{j=1}^{m_1} \left[\sum_{r=0}^{y_{1j}-1} \left\{ \frac{(r-\pi)}{(1-\theta_1)\,\pi + r\theta_1} \right\} + \sum_{r=0}^{n_{1j}-y_{1j}-1} \left\{ \frac{(r+\pi-1)}{(1-\theta_1)\,(1-\pi) + r\theta_1} \right\} \right.$$
$$\left. - \sum_{r=0}^{n_{1j}-1} \left\{ \frac{(r-1)}{(1-\theta_1) + r\theta_1} \right\} \right]$$

and

$$\gamma_3 = \sum_{j=1}^{m_2} \left[\sum_{r=0}^{y_{2j}-1} \left\{ \frac{(r-\pi)}{(1-\theta_2)\,\pi + r\theta_2} \right\} + \sum_{r=0}^{n_{2j}-y_{2j}-1} \left\{ \frac{(r+\pi-1)}{(1-\theta_2)\,(1-\pi) + r\theta_2} \right\} \right.$$
$$\left. - \sum_{r=0}^{n_{2j}-1} \left\{ \frac{(r-1)}{(1-\theta_2) + r\theta_2} \right\} \right].$$

The derivation of the expected values of negative of the mixed partial derivatives is given in the Appendix **??** and the results are given below

$$D_{11} = (1-\theta_1)^2 \sum_{j=1}^{m_1} \left[\left\{ \sum_{r=1}^{n_{1j}} \frac{Pr\,(y_{1j} \geq r)}{\{(1-\theta_1)\,\pi + (r-1)\,\theta_1\}^2} \right\} + \left\{ \sum_{r=1}^{n_{1j}} \frac{Pr\,(y_{1j} \leq n_{1j} - r)}{\{(1-\theta_1)\,(1-\pi) + (r-1)\,\theta_1\}^2} \right\} \right],$$

$$A_{11} = (1-\theta_1)^2 \sum_{j=1}^{m_1} \left[\left\{ \sum_{r=1}^{n_{1j}} \frac{Pr\,(y_{1j} \geq r)}{\{(1-\theta_1)\,\pi + (r-1)\,\theta_1\}^2} \right\} + \left\{ \sum_{r=1}^{n_{1j}} \frac{Pr\,(y_{1j} \leq n_{1j} - r)}{\{(1-\theta_1)\,(1-\pi) + (r-1)\,\theta_1\}^2} \right\} \right],$$

$$A_{12} = (1-\theta_1) \sum_{j=1}^{m_1} \left[\sum_{r=1}^{n_{1j}} \frac{(r-\pi-1)\,Pr\,(y_{1j} \geq r)}{\{(1-\theta_1)\,\pi + (r-1)\,\theta_1\}^2} - \sum_{r=1}^{n_{1j}} \frac{(r+\pi-2)\,Pr\,(y_{1j} \leq n_{1j} - r)}{\{(1-\theta_1)\,(1-\pi) + (r-1)\,\theta_1\}^2} \right] + \sum_{j=1}^{m_1} \left[\sum_{r=1}^{n_{1j}} \frac{Pr\,(y_{1j} \geq r)}{\{(1-\theta_1)\,\pi + (r-1)\,\theta_1\}} - \sum_{r=1}^{n_{1j}} \frac{Pr\,(y_{1j} \leq n_{1j} - r)}{\{(1-\theta_1)\,(1-\pi) + (r-1)\,\theta_1\}} \right] = A_{21},$$

$$A_{13} = 0 = A_{31},$$

$$B_{11} = (1-\theta_1)^2 \sum_{j=1}^{m_1} \left[\sum_{r=1}^{n_{1j}} \frac{Pr\,(y_{1j} \geq r)}{\{(1-\theta_1)\,\pi + (r-1)\,\theta_1\}^2} + \sum_{r=1}^{n_{1j}} \frac{Pr\,(y_{1j} \leq n_{1j} - r)}{\{(1-\theta_1)\,(1-\pi) + (r-1)\,\theta_1\}^2} \right] + (1-\theta_2)^2 \sum_{j=1}^{m_2} \left[\sum_{r=1}^{n_{2j}} \frac{Pr\,(y_{2j} \geq r)}{\{(1-\theta_2)\,\pi + (r-1)\,\theta_2\}^2} + \sum_{r=1}^{n_{2j}} \frac{Pr\,(y_{2j} \leq n_{2j} - r)}{\{(1-\theta_2)\,(1-\pi) + (r-1)\,\theta_2\}^2} \right],$$

$$B_{12} = (1-\theta_1) \sum_{j=1}^{m_1} \left[\sum_{r=1}^{n_{1j}} \frac{(r-\pi-1)\,Pr\,(y_{1j} \geq r)}{\{(1-\theta_1)\,\pi + (r-1)\,\theta_1\}^2} - \sum_{r=1}^{n_{1j}} \frac{(r+\pi-2)\,Pr\,(y_{1j} \leq n_{1j} - r)}{\{(1-\theta_1)\,(1-\pi) + (r-1)\,\theta_1\}^2} \right] + \sum_{j=1}^{m_1} \left[\sum_{r=1}^{n_{1j}} \frac{Pr\,(y_{1j} \geq r)}{\{(1-\theta_1)\,\pi + (r-1)\,\theta_1\}} - \sum_{r=1}^{n_{1j}} \frac{Pr\,(y_{1j} \leq n_{1j} - r)}{\{(1-\theta_1)\,(1-\pi) + (r-1)\,\theta_1\}} \right] = B_{21},$$

$$B_{13} = (1-\theta_2)\sum_{j=1}^{m_2}\left[\sum_{r=1}^{n_{2j}}\frac{(r-\pi-1)\,Pr\,(y_{2j}\geq r)}{\{(1-\theta_2)\,\pi+(r-1)\,\theta_2\}^2}-\sum_{r=1}^{n_{2j}}\frac{(r+\pi-2)\,Pr\,(y_{2j}\leq n_{2j}-r)}{\{(1-\theta_2)\,(1-\pi)+(r-1)\,\theta_2\}^2}\right]$$
$$+\sum_{j=1}^{m_2}\left[\sum_{r=1}^{n_{2j}}\frac{Pr\,(y_{2j}\geq r)}{\{(1-\theta_2)\,\pi+(r-1)\,\theta_2\}}-\sum_{r=1}^{n_{2j}}\frac{Pr\,(y_{2j}\leq n_{2j}-r)}{\{(1-\theta_2)\,(1-\pi)+(r-1)\,\theta_2\}}\right]=B_{31},$$

$$B_{22} = \sum_{j=1}^{m_1}\left[\sum_{r=1}^{n_{1j}}\frac{(r-\pi-1)^2Pr(y_{1j}\geq r)}{\{(1-\theta_1)\pi+(r-1)\theta_1\}^2}+\sum_{r=1}^{n_{1j}}\frac{(r+\pi-2)^2Pr(y_{1j}\leq n_{1j}-r)}{\{(1-\theta_1)(1-\pi)+(r-1)\theta_1\}^2}\right.$$
$$\left.-\sum_{r=1}^{n_{1j}}\frac{(r-2)^2}{\{(1-\theta_1)+(r-1)\theta_1\}^2}\right],$$

$$B_{23} = 0 = B_{32}$$

and

$$B_{33} = \sum_{j=1}^{m_2}\left[\sum_{r=1}^{n_{2j}}\frac{(r-\pi-1)^2Pr(y_{2j}\geq r)}{\{(1-\theta_2)\pi+(r-1)\theta_2\}^2}+\sum_{r=1}^{n_{2j}}\frac{(r+\pi-2)^2Pr(y_{2j}\leq n_{2j}-r)}{\{(1-\theta_2)(1-\pi)+(r-1)\theta_2\}^2}\right.$$
$$\left.-\sum_{r=1}^{n_{2j}}\frac{(r-2)^2}{\{(1-\theta_2)+(r-1)\theta_2\}^2}\right].$$

Denoting the maximum likelihood estimate of δ as $\hat{\delta}_{ml}$ and using this estimate, the $C\,(\alpha)$ statistic which, in this case is the score statistic (Rao (1948)), is obtained as $C_{ml} = \psi_1^2/\left(D_{11} - AB^{-1}A'\right)$. Under the null hypothesis C_{ml} is distributed asymptotically (as $m \to \infty$, where $m = m_1 + m_2$) as a chi-squared with 1 degree of freedom.

4.2.3 $C(\alpha)$ Test Based on Kleinman's (1973) Method of Moments Estimates of Nuisance Parameters

The method of moments estimators of a two parameter model can be obtained by equating the sample mean and the sample variance to their expected values. In a beta binomial model Kleinman (1973) uses weighted average and weighted variance of sample proportions to equate to their respective expected values to find method of moments estimates of the parameters π and θ in case of a single sample.

Our objective here is to find method of moments estimates of the parameters π, θ_1 and θ_2 from two independent samples assuming equality of the proportions. Now, define $z_{ij} = \dfrac{y_{ij}}{n_{ij}}$, $w_{ij} = \dfrac{n_{ij}}{\pi(1-\pi)\{1+(n_{ij}-1)\theta_i\}}$; $i = 1, 2$; $j = 1, 2, \cdots, m_i$. Note, w_{ij} is the inverse of the variance of z_{ij}, under the null hypothesis. Then, the pooled weighted average of sample proportions is given by $\hat{\pi} = \dfrac{\sum_{i=1}^{2}\sum_{j=1}^{m_i} w_{ij} z_{ij}}{\sum_{i=1}^{2}\sum_{j=1}^{m_i} w_{ij}}$.

It can be seen that under H_0 $E(\hat{\pi}) = \pi$. Further, define $S_1 = \sum_{j=1}^{m_1} w_{1j}(z_{1j}-\pi)^2$ and $S_2 = \sum_{j=1}^{m_2} w_{2j}(z_{2j}-\pi)^2$. Again, under H_0, it can be seen that $E(S_1) = \sum_{j=1}^{m_1} \dfrac{w_{1j}\pi(1-\pi)}{n_{1j}} + \sum_{j=1}^{m_1} \dfrac{w_{1j}\pi(1-\pi)(n_{1j}-1)\theta_1}{n_{1j}}$ and $E(S_2) = \sum_{j=1}^{m_2} \dfrac{w_{2j}\pi(1-\pi)}{n_{2j}} + \sum_{j=1}^{m_2} \dfrac{w_{2j}\pi(1-\pi)(n_{2j}-1)\theta_2}{n_{2j}}$. Thus, following Kleinman (1973), the method of moments estimates of $\delta = (\pi, \theta_1, \theta_2)$ are obtained by solving the equations

$$\hat{\pi} = E(\hat{\pi}), \quad S_1 = E(S_1) \quad \text{and} \quad S_2 = E(S_2)$$

simultaneously for π, θ_1 and θ_2. These equations, can then be expressed as the following estimating equations

$$\sum_{i=1}^{2}\sum_{j=1}^{m_i} w_{ij}\left(z_{ij}-\pi\right)=0, \tag{4.5}$$

$$\sum_{j=1}^{m_1} w_{1j}\left(z_{1j}-\pi\right)^2-\sum_{j=1}^{m_1}\frac{w_{1j}\pi\left(1-\pi\right)}{n_{1j}}-\sum_{j=1}^{m_1}\frac{w_{1j}\pi\left(1-\pi\right)\left(n_{1j}-1\right)\theta_1}{n_{1j}}=0 \tag{4.6}$$

and

$$\sum_{j=1}^{m_2} w_{2j}\left(z_{2j}-\pi\right)^2-\sum_{j=1}^{m_2}\frac{w_{2j}\pi\left(1-\pi\right)}{n_{2j}}-\sum_{j=1}^{m_2}\frac{w_{2j}\pi\left(1-\pi\right)\left(n_{2j}-1\right)\theta_2}{n_{2j}}=0. \tag{4.7}$$

Denote the estimates of π, θ_1 and θ_2 obtained by solving the three equations (4.5), (4.6) and (4.7) by $\tilde{\delta}_{kmm}$. Substituting these estimates in S, A, B and D, the $C(\alpha)$ statistic based on Kleinman's method of moments estimates is obtained as $C_{kmm} = S^2 \,/\, \left(D_{11} - AB^{-1}A'\right)$, which follows, asymptotically, as $m \to \infty$, a chi-squared distribution with 1 degree of freedom, where $m = m_1 + m_2$.

4.3 Semi-parametric Tests

4.3.1 $C(\alpha)$ Test Based on the Quasi-Likelihood and the Method of Moments Estimates by Breslow (1990)

The quasi-likelihood (QL) of Wedderburn (1974) is based on the knowledge of the first two moments of the random variable. For an observation $z_{ij} = \dfrac{y_{ij}}{n_{ij}}$ with

$E\left(Z_{ij}\right)=\pi_i$ and $var\left(Z_{ij}\right)=\dfrac{\pi_i(1-\pi_i)}{n_{ij}}\left\{1+\left(n_{ij}-1\right)\theta_i\right\}$; $i=1,2;\ \ j=1,2,\cdots,m_i$, $0\le\pi_i\le 1$ and $\cap_j\left(\dfrac{-1}{n_{ij}-1}\right)<\theta_i<1$, the QL can be obtained as $Q_{ij}\left(z_{ij}\right)=\displaystyle\int_{z_{ij}}^{\pi_i}\frac{n_{ij}\left(z_{ij}-t\right)}{t\left(1-t\right)\left\{1+\left(n_{ij}-1\right)\theta_i\right\}}\,dt$. Here, $\cap_j\left(\dfrac{-1}{n_{ij}-1}\right)=\dfrac{-1}{max\left\{n_{ij}\right\}-1}$.

Now, as we have two independent samples, the quasi-likelihood, Q, for the data can be obtained as

$$\begin{aligned}Q\left(z_{ij};\ \pi_i,\ \theta_i\right)=&\sum_{i=1}^{2}\sum_{j=1}^{m_i}\left[\frac{1}{\left\{1+\left(n_{ij}-1\right)\theta_i\right\}}\left\{y_{ij}\log\left(\frac{\pi_i}{z_{ij}}\right)\right.\right.\\&\left.\left.+\left(n_{ij}-y_{ij}\right)\log\left(\frac{1-\pi_i}{1-z_{ij}}\right)\right\}\right].\end{aligned}\tag{4.8}$$

Considering the same reparameterization as in section 4.2.2, under $H_A:\pi_i=\pi+\phi_i$, with $\phi_2=0$ the quasi-likelihood function (4.8) takes the form

$$\begin{aligned}Q\left(z_{ij};\pi,\ \phi_i,\ \theta_i\right)=&\sum_{i=1}^{2}\sum_{j=1}^{m_i}\left[\frac{1}{\left\{1+\left(n_{ij}-1\right)\theta_i\right\}}\left\{y_{ij}\log\left(\frac{\pi+\phi_i}{z_{ij}}\right)\right.\right.\\&\left.\left.+\left(n_{ij}-y_{ij}\right)\log\left(\frac{1-\pi-\phi_i}{1-z_{ij}}\right)\right\}\right].\end{aligned}\tag{4.9}$$

Then, the quasi-likelihood score function for π, under H_0, is

$$g_1=\left.\frac{\partial Q}{\partial\pi}\right|_{\phi=0}=\sum_{i=1}^{2}\sum_{j=1}^{m_i}\left[\frac{1}{1+\left(n_{ij}-1\right)\theta_i}\left\{\frac{\left(z_{ij}-\pi\right)n_{ij}}{\pi\left(1-\pi\right)}\right\}\right],\tag{4.10}$$

which is an unbiased estimating function. As the quasi-likelihood is designed only for the estimation of the mean parameter, no such estimating functions for θ_1 and θ_2 can be obtained from Q. However, given π, the unbiased estimating functions to estimate θ_1 and θ_2 can be obtained by using the moment method (Breslow (1990); Moore and

Tsiatis (1991)). These estimating functions for θ_1 and θ_2 are obtained by equating the Pearson chi-squared statistics with their expected values and are given by

$$g_2 = \sum_{j=1}^{m_1} \left[\frac{1}{1+(n_{1j}-1)\,\theta_1} \left\{ \frac{(z_{1j}-\pi-\phi_1)^2\, n_{1j}}{(\pi-\phi_1)(1-\pi-\phi_1)} \right\} \right] - (n_1 - m_1) \tag{4.11}$$

and

$$g_3 = \sum_{j=1}^{m_2} \left[\frac{1}{1+(n_{2j}-1)\,\theta_2} \left\{ \frac{(z_{2j}-\pi-\phi_2)^2\, n_{2j}}{(\pi-\phi_2)(1-\pi-\phi_2)} \right\} \right] - (n_2 - m_2)\,, \tag{4.12}$$

where $n_i = \sum_{j=1}^{m_i} n_{ij};\ i = 1,\, 2$. Then, the method of moment estimates of π, θ_1 and θ_2, under the null hypothesis, based on the quasi-likelihood, are obtained by solving the following equations

$$\sum_{i=1}^{2} \sum_{j=1}^{m_i} \left[\frac{1}{1+(n_{ij}-1)\,\theta_i} \left\{ \frac{(z_{ij}-\pi)\, n_{ij}}{\pi(1-\pi)} \right\} \right] = 0, \tag{4.13}$$

$$\sum_{j=1}^{m_1} \left[\frac{1}{1+(n_{1j}-1)\,\theta_1} \left\{ \frac{(z_{1j}-\pi)^2\, n_{1j}}{\pi(1-\pi)} \right\} \right] - (n_1 - m_1) = 0 \tag{4.14}$$

and

$$\sum_{j=1}^{m_2} \left[\frac{1}{1+(n_{2j}-1)\,\theta_2} \left\{ \frac{(z_{2j}-\pi)^2\, n_{2j}}{\pi(1-\pi)} \right\} \right] - (n_2 - m_2) = 0 \tag{4.15}$$

simultaneously.

Now, define ϕ and δ as before and the following

$$\psi_{1qb} = \frac{\partial Q}{\partial \phi_1}\Big|_{\phi=0}, \ \ \gamma_{1qb} = \frac{\partial Q}{\partial \pi}\Big|_{\phi=0}, \ \gamma_{2qb} = g_2\Big|_{\phi=0} \text{ and } \gamma_{3qb} = g_3\Big|_{\phi=0}.$$

By the Lindeberg central-limit theorem (Moore (1985)), asymptotically, as $m_i \to \infty$, we have $\psi_{1qb}(\delta) \sim N\left(0,\ \Delta_{11qb} - \Delta_{12qb}\Delta_{22qb}^{-1}\Delta_{21qb}\right)$.

The dimensions of matrices Δ_{11qb}, Δ_{12qb}, Δ_{21qb}, and Δ_{22qb} are 1×1, 1×3, 3×1, and 3×3 respectively with elements

$$\Delta_{11qb} = E\left(-\frac{\partial^2 Q}{\partial \phi_1^2}\Big|_{\phi=0}\right),$$

$$\Delta_{12qb1} = E\left(-\frac{\partial^2 Q}{\partial \phi_1 \partial \pi}\Big|_{\phi=0}\right), \ \Delta_{12qb2} = E\left(-\frac{\partial^2 Q}{\partial \phi_1 \partial \theta_1}\Big|_{\phi=0}\right), \ \Delta_{12qb3} = E\left(-\frac{\partial^2 Q}{\partial \phi_1 \partial \theta_2}\Big|_{\phi=0}\right),$$

$$\Delta_{21qb1} = E\left(-\frac{\partial^2 Q}{\partial \pi \partial \phi_1}\Big|_{\phi=0}\right), \ \Delta_{21qb2} = E\left(-\frac{\partial g_2}{\partial \phi_1}\Big|_{\phi=0}\right), \ \Delta_{21qb3} = E\left(-\frac{\partial g_3}{\partial \phi_1}\Big|_{\phi=0}\right),$$

$$\Delta_{22qb11} = E\left(-\frac{\partial^2 Q}{\partial \pi^2}\Big|_{\phi=0}\right), \ \Delta_{22qb12} = E\left(-\frac{\partial^2 Q}{\partial \pi \partial \theta_1}\Big|_{\phi=0}\right), \ \Delta_{22qb13} = E\left(-\frac{\partial^2 Q}{\partial \pi \partial \theta_2}\Big|_{\phi=0}\right),$$

$$\Delta_{22qb21} = E\left(-\frac{\partial g_2}{\partial \pi}\Big|_{\phi=0}\right), \ \Delta_{22qb22} = E\left(-\frac{\partial g_2}{\partial \theta_1}\Big|_{\phi=0}\right), \ \Delta_{22qb23} = E\left(-\frac{\partial g_2}{\partial \theta_2}\Big|_{\phi=0}\right),$$

$$\Delta_{22qb31} = E\left(-\frac{\partial g_3}{\partial \pi}\Big|_{\phi=0}\right), \ \Delta_{22qb32} = E\left(-\frac{\partial g_3}{\partial \theta_1}\Big|_{\phi=0}\right) \quad \text{and}$$

$$\Delta_{22qb33} = E\left(-\frac{\partial g_3}{\partial \theta_2}\bigg|_{\phi=0}\right).$$

On taking the expectation of the negative of the second derivatives and after simplifications, the above quantities are obtained as

$$\Delta_{11qb} = \sum_{j=1}^{m_1}\left[\frac{n_{1j}}{\pi(1-\pi)\{1+(n_{1j}-1)\theta_1\}}\right] = d_{1qb},$$

$$\Delta_{12qb1} = \sum_{j=1}^{m_1}\left[\frac{n_{1j}}{\pi(1-\pi)\{1+(n_{1j}-1)\theta_1\}}\right] = d_{1qb},$$

$$\Delta_{12qb2} = 0,\;\; \Delta_{12qb3} = 0,$$

$$\Delta_{21qb1} = \sum_{j=1}^{m_1}\left[\frac{n_{1j}}{\pi(1-\pi)\{1+(n_{1j}-1)\theta_1\}}\right] = d_{1qb},$$

$$\Delta_{21qb2} = \sum_{j=1}^{m_1}\left[\frac{1-2\pi}{\pi(1-\pi)}\right] = c_{qb}m_1,\;\; \Delta_{21qb3} = 0,$$

$$\Delta_{22qb11} = \sum_{i=1}^{2}\sum_{j=1}^{m_i}\left[\frac{n_{ij}}{\pi(1-\pi)\{1+(n_{ij}-1)\theta_i\}}\right] = \sum_{i=1}^{2} d_{iqb} = d_{qb},$$

$$\Delta_{22qb12} = 0,\;\; \Delta_{22qb13} = 0,$$

$$\Delta_{22qb21} = \sum_{j=1}^{m_1}\left[\frac{1-2\pi}{\pi(1-\pi)}\right] = c_{qb}m_1,\;\; \Delta_{22qb22} = \sum_{j=1}^{m_1}\left[\frac{(n_{1j}-1)}{1+(n_{1j}-1)\theta_1}\right] = s_{1qb},$$

$$\Delta_{22qb23} = 0,\;\; \Delta_{22qb31} = \sum_{j=1}^{m_2}\left[\frac{1-2\pi}{\pi(1-\pi)}\right] = c_{qb}m_2,$$

$$\Delta_{22qb32} = 0 \text{ and } \Delta_{22qb33} = \sum_{j=1}^{m_2} \left[\frac{(n_{2j}-1)}{1+(n_{2j}-1)\,\theta_2} \right] = s_{2qb}.$$

Then, the matrices Δ_{11qb}, Δ_{12qb}, Δ_{21qb}, and Δ_{22qb} are

$$\Delta_{11qb} = d_{1qb}, \;\; \Delta_{12qb} = (d_{1qb} \;\, 0 \;\, 0)\,,$$

$$\Delta_{22qb} = \begin{pmatrix} d_{qb} & 0 & 0 \\ c_{qb}m_1 & s_{1qb} & 0 \\ c_{qb}m_2 & 0 & s_{2qb} \end{pmatrix} \text{ and } \Delta_{21qb} = \begin{pmatrix} d_{1qb} \\ c_{qb}m_1 \\ 0 \end{pmatrix}$$

respectively.

Using the estimates of π, θ_1 and θ_2 obtained from simultaneously solving the equations (4.13), (4.14) and (4.15), the $C(\alpha)$ statistic, based on the quasi-likelihood and Breslow's (1990(a)) method of moments estimates of the parameters π, θ_1 and θ_2, is $C_{qb} = \psi^2_{1qb}/\left(\Delta_{11qb} - \Delta_{12qb}\Delta^{-1}_{22qb}\Delta_{21qb}\right)$, which is distributed asymptotically, as $m \to \infty$, a chi-squared distribution with 1 degree of freedom, where $m = m_1 + m_2$.

4.3.2 $C(\alpha)$ Test Based on the Quasi-Likelihood and the Method of Moments Estimates by Srivastava and Wu (1993)

We follow the same procedure as in the preceding section except that for estimating the dispersion parameters θ_1 and θ_2, we use estimating functions proposed by Srivastava and Wu (1993). Here, like in the preceding section, the quasi-likelihood score, under the null hypothesis, which is the unbiased estimating function, to estimate the parameter π is

$$g_1 = \frac{\partial Q}{\partial \pi}\bigg|_{\phi=0} = \sum_{i=1}^{2}\sum_{j=1}^{m_i}\left[\frac{1}{1+(n_{ij}-1)\,\theta_i}\left\{\frac{(z_{ij}-\pi)\,n_{ij}}{\pi\,(1-\pi)}\right\}\right].$$

Given π, the estimating functions based on the method of moments estimates for θ_1 and θ_2 proposed by Srivastava and Wu (1993) are

$$g_2 = \sum_{j=1}^{m_1}\left[\frac{n_{1j}^2\,(z_{1j}-\pi-\phi_1)^2}{(\pi+\phi_1)\,(1-\pi-\phi_1)} - n_{1j}\left\{1+(n_{1j}-1)\,\theta_1\right\}\right] \tag{4.16}$$

and

$$g_3 = \sum_{j=1}^{m_2}\left[\frac{n_{2j}^2\,(z_{2j}-\pi-\phi_2)^2}{(\pi+\phi_2)\,(1-\pi-\phi_2)} - n_{2j}\left\{1+(n_{2j}-1)\,\theta_2\right\}\right]. \tag{4.17}$$

Under the null hypothesis, the $\sqrt{m}$ consistent estimates of π, θ_1 and θ_2 are obtained by solving the following equations

$$\sum_{i=1}^{2}\sum_{j=1}^{m_i}\left[\frac{1}{1+(n_{ij}-1)\,\theta_i}\left\{\frac{(z_{ij}-\pi)\,n_{ij}}{\pi\,(1-\pi)}\right\}\right] = 0, \tag{4.18}$$

$$\sum_{j=1}^{m_1}\left[\frac{n_{1j}^2\,(z_{1j}-\pi)^2}{\pi\,(1-\pi)} - n_{1j}\left\{1+(n_{1j}-1)\,\theta_1\right\}\right] = 0 \tag{4.19}$$

and

$$\sum_{j=1}^{m_2}\left[\frac{n_{2j}^2\,(z_{2j}-\pi)^2}{\pi\,(1-\pi)} - n_{2j}\left\{1+(n_{2j}-1)\,\theta_2\right\}\right] = 0 \tag{4.20}$$

simultaneously.

As in the previous section we denote the matrices involving the second derivatives (mixed) by Δ_{11qs}, Δ_{12qs}, Δ_{21qs} and Δ_{22qs}. Then, following similar steps as in the preceding section, the elements of the matrices are obtained as

$$\Delta_{11qs} = E\left(-\frac{\partial^2 Q}{\partial \phi_1^2}\bigg|_{\phi=0}\right),\ \Delta_{12qs1} = E\left(-\frac{\partial^2 Q}{\partial \phi_1 \partial \pi}\bigg|_{\phi=0}\right),\ \Delta_{12qs2} = E\left(-\frac{\partial^2 Q}{\partial \phi_1 \partial \theta_1}\bigg|_{\phi=0}\right),$$

$$\Delta_{12qs3} = E\left(-\frac{\partial^2 Q}{\partial \phi_1 \partial \theta_2}\bigg|_{\phi=0}\right),\ \Delta_{21qs1} = E\left(-\frac{\partial^2 Q}{\partial \pi \partial \phi_1}\bigg|_{\phi=0}\right),\ \Delta_{21qs2} = E\left(-\frac{\partial g_2}{\partial \phi_1}\bigg|_{\phi=0}\right),$$

$$\Delta_{21qs3} = E\left(-\frac{\partial g_3}{\partial \phi_1}\bigg|_{\phi=0}\right),\ \Delta_{22qs11} = E\left(-\frac{\partial^2 Q}{\partial \pi^2}\bigg|_{\phi=0}\right),\ \Delta_{22qs12} = E\left(-\frac{\partial^2 Q}{\partial \pi \partial \theta_1}\bigg|_{\phi=0}\right),$$

$$\Delta_{22qs13} = E\left(-\frac{\partial^2 Q}{\partial \pi \partial \theta_2}\bigg|_{\phi=0}\right),\ \Delta_{22qs21} = E\left(-\frac{\partial g_2}{\partial \pi}\bigg|_{\phi=0}\right),\ \Delta_{22qs22} = E\left(-\frac{\partial g_2}{\partial \theta_1}\bigg|_{\phi=0}\right),$$

$$\Delta_{22qs23} = E\left(-\frac{\partial g_2}{\partial \theta_2}\bigg|_{\phi=0}\right),\ \Delta_{22qs31} = E\left(-\frac{\partial g_3}{\partial \pi}\bigg|_{\phi=0}\right),\ \Delta_{22qs32} = E\left(-\frac{\partial g_3}{\partial \theta_1}\bigg|_{\phi=0}\right)$$

and $\Delta_{22qs33} = E\left(-\frac{\partial g_3}{\partial \theta_2}\bigg|_{\phi=0}\right)$.

After detained derivation we obtain

$$\Delta_{11qs} = \sum_{j=1}^{m_1}\left[\frac{n_{1j}}{\pi(1-\pi)\{1+(n_{1j}-1)\theta_1\}}\right] = d_{1qs},$$

$$\Delta_{12qs1} = \sum_{j=1}^{m_1} \left[\frac{n_{1j}}{\pi(1-\pi)\{1+(n_{1j}-1)\theta_1\}} \right] = d_{1qs},\ \ \Delta_{12qs2} = 0,\ \ \Delta_{12qs3} = 0,$$

$$\Delta_{21qs1} = \sum_{j=1}^{m_1} \left[\frac{n_{1j}}{\pi(1-\pi)\{1+(n_{1j}-1)\theta_1\}} \right] = d_{1qs},$$

$$\Delta_{21qs2} = \frac{(1-2\pi)}{\pi(1-\pi)} \sum_{j=1}^{m_1} n_{1j}\{1+(n_{1j}-1)\theta_1\} = c_{qs}s_{1qs},\ \ \Delta_{21qs3} = 0,$$

$$\Delta_{22qs11} = \sum_{i=1}^{2}\sum_{j=1}^{m_i} \left[\frac{n_{ij}}{\pi(1-\pi)\{1+(n_{ij}-1)\theta_i\}} \right] = \sum_{i=1}^{2} d_i = d_{qs},\ \ \Delta_{22qs12} = 0,$$

$$\Delta_{22qs13} = 0,\ \ \Delta_{22qs21} = \frac{(1-2\pi)}{\pi(1-\pi)} \sum_{j=1}^{m_1} n_{1j}\{1+(n_{1j}-1)\theta_1\} = c_{qs}s_{1qs},$$

$$\Delta_{22qs22} = \sum_{j=1}^{m_1} n_{1j}(n_{1j}-1) = e_{1qs},\ \ \Delta_{22qs23} = 0,$$

$$\Delta_{22qs31} = \frac{(1-2\pi)}{\pi(1-\pi)} \sum_{j=1}^{m_2} n_{2j}\{1+(n_{2j}-1)\theta_2\} = c_{qs}s_{2qs},\ \ \Delta_{22qs32} = 0$$

and

$$\Delta_{22qs33} = \sum_{j=1}^{m_2} n_{2j}(n_{2j}-1) = e_{2qs}.$$

Now, the matrices Δ_{11qs}, Δ_{12qs}, Δ_{21qs}, and Δ_{22qs} are

$$\Delta_{11qs} = d_{1qs},\ \ \Delta_{12qs} = (d_{1qs}\ \ 0\ \ 0),$$

$$\Delta_{22qs} = \begin{pmatrix} d_{qs} & 0 & 0 \\ c_{qs}s_{1qs} & e_{1qs} & 0 \\ c_{qs}s_{2qs} & 0 & e_{2qs} \end{pmatrix} \quad \text{and} \quad \Delta_{21qs} = \begin{pmatrix} d_{1qs} \\ c_{qs}s_{1qs} \\ 0 \end{pmatrix} \quad \text{respectively.}$$

Using the estimates of π, θ_1 and θ_2 obtained from simultaneously solving the equations (4.18), (4.19) and (4.20), the $C(\alpha)$ statistic, based on the quasi-likelihood and the method of moments estimates by Srivastava and Wu (1993), is $C_{qs} = \psi_{1qs}^2 / \left(\Delta_{11qs} - \Delta_{12qs}\Delta_{22qs}^{-1}\Delta_{21qs}\right)$, which is distributed, asymptotically, as χ^2 with 1 degree of freedom.

4.3.3 $C(\alpha)$ Test Based on the Extended Quasi-Likelihood Estimates

The quasi-likelihood function does not possess a property similar to the log-likelihood function with respect to the derivative of the dispersion parameter. Thus, the quasi-likelihood function facilitates the estimation of only the mean parameter and it is not suitable to estimate the dispersion parameter(s). In order to estimate the mean parameter as well as the dispersion parameter from the same function, Nelder and Pregibon (1987) and Godambe and Thompson (1989) propose the extended quasi-likelihood function (EQL). This function is implemented by adding a normalizing factor to the quasi-likelihood function.

The extended quasi-likelihood for an observation $z_{ij} = \dfrac{y_{ij}}{n_{ij}}$ with mean and variance specified in the previous section can be obtained from

$$Q^{+}\left(z_{ij};\pi_i,\ \theta_i\right)=-\frac{1}{2}\log\left(2k\right)-\frac{1}{2}\log\left[\frac{z_{ij}\left(1-z_{ij}\right)\left\{1+\left(n_{ij}-1\right)\theta_i\right\}}{n_{ij}}\right]$$
$$+\int_{z_{ij}}^{\pi_i}\frac{n_{ij}\left(z_{ij}-t\right)}{t\left(1-t\right)\left\{1+\left(n_{ij}-1\right)\theta_i\right\}}\,dt. \tag{4.21}$$

For two independent random samples $z_{ij},\ i=1,\ 2;\ j=1,\ 2,\cdots,\ m_i$, where $y_{ij}\sim BB\left(\pi_i,\ \theta_i\right)$, the extended quasi-likelihood, apart from a constant, can be written as

$$Q^{+}=C-\frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{m_i}\Bigg[\log\left\{1+(n_{ij}-1)\theta_i\right\}$$
$$-\frac{2}{1+(n_{ij}-1)\theta_i}\left\{y_{ij}\log\left(\frac{\pi_i}{z_{ij}}\right)+\left(n_{ij}-y_{ij}\right)\log\left(\frac{1-\pi_i}{1-z_{ij}}\right)\right\}\Bigg]. \tag{4.22}$$

Under the null hypothesis, $H_0:\pi_1=\pi_2=\pi$, the extended quasi-likelihood (4.22) takes the form

$$Q^{+}=C-\frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{m_i}\Bigg[\log\left\{1+(n_{ij}-1)\theta_i\right\}$$
$$-\frac{2}{1+(n_{ij}-1)\theta_i}\left\{y_{ij}\log\left(\frac{\pi}{z_{ij}}\right)+\left(n_{ij}-y_{ij}\right)\log\left(\frac{1-\pi}{1-z_{ij}}\right)\right\}\Bigg]. \tag{4.23}$$

The estimates of the parameters π, θ_1 and θ_2, under the null hypothesis, can be obtained by directly maximizing the extended quasi-likelihood function (4.23) or by solving the following estimating equations

$$\sum_{i=1}^{2}\sum_{j=1}^{m_i}\left[\frac{1}{1+(n_{ij}-1)\theta_i}\left\{\frac{y_{ij}}{\pi}-\frac{n_{ij}-y_{ij}}{1-\pi}\right\}\right]=0,$$

$$\sum_{j=1}^{m_1}\left[\frac{n_{1j}-1}{\{1+(n_{1j}-1)\theta_1\}^2}\left\{y_{1j}\log\left(\frac{z_{1j}}{\pi}\right)+(n_{1j}-y_{1j})\log\left(\frac{1-z_{1j}}{1-\pi}\right)\right.\right.$$
$$\left.-\frac{1+(n_{1j}-1)\theta_1}{2}\right]=0$$

and

$$\sum_{j=1}^{m_2}\left[\frac{n_{2j}-1}{\{1+(n_{2j}-1)\theta_2\}^2}\left\{y_{2j}\log\left(\frac{z_{2j}}{\pi}\right)+(n_{2j}-y_{2j})\log\left(\frac{1-z_{2j}}{1-\pi}\right)\right.\right.$$
$$\left.-\frac{1+(n_{2j}-1)\theta_2}{2}\right]=0$$

simultaneously.

Now, following the reparameterization as in section 4.2.2, the extended quasi-likelihood in terms of π, ϕ_1, ϕ_2, θ_1 and θ_2, under the alternative hypothesis, can be written as

$$\begin{aligned} Q^+ &= C-\frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{m_i}\left[\log\{1+(n_{ij}-1)\theta_i\}\right. \\ &\left.-\frac{2}{1+(n_{ij}-1)\theta_i}\left\{y_{ij}\log\left(\frac{\pi+\phi_i}{z_{ij}}\right)+(n_{ij}-y_{ij})\log\left(\frac{1-\pi-\phi_i}{1-z_{ij}}\right)\right\}\right]. \end{aligned} \tag{4.24}$$

Then, the $C(\alpha)$ statistic is based on the adjusted score $S_{eq}=\psi_{1eq}-\beta_{1eq}\gamma_{1eq}-\beta_{2eq}\gamma_{2eq}-\beta_{3eq}\gamma_{3eq}$, where $\psi_{1eq}=\left.\frac{\partial Q^+}{\partial\phi_1}\right|_{\phi=0}$ and β_{1eq}, β_{2eq} and β_{3eq} are the partial regression coefficient of ψ_{1eq} on γ_{1eq}, ψ_{1eq} on γ_{2eq}, and ψ_{1eq} on γ_{3eq} respectively. It can be seen that, for the extended quasi-likelihood, $\gamma_{1eq}=0$, $\gamma_{2eq}=0$, and $\gamma_{3eq}=0$. Thus $S_{eq}=\psi_{1eq}$.

The structure of the dispersion matrix ψ_{1eq} is

$$V_{eq} = \begin{pmatrix} D_{eq} & A_{eq} \\ A'_{eq} & B_{eq} \end{pmatrix},$$

where D_{eq} is 1×1, A_{eq} is 1×3 and B_{eq} is 3×3 with the following elements

$$D_{11eq} = E\left[-\frac{\partial^2 Q^+}{\partial \Phi_1^2}\Big|_{\Phi=0}\right],$$

$$A_{11eq} = E\left[-\frac{\partial^2 Q^+}{\partial \Phi_1 \partial \pi}\Big|_{\Phi=0}\right], \quad A_{12eq} = E\left[-\frac{\partial^2 Q^+}{\partial \Phi_1 \partial \theta_1}\Big|_{\Phi=0}\right], \quad A_{13eq} = E\left[-\frac{\partial^2 Q}{\partial \Phi_1 \partial \theta_2}\Big|_{\Phi=0}\right],$$

$$B_{11eq} = E\left[-\frac{\partial^2 Q^+}{\partial \pi^2}\Big|_{\Phi=0}\right], \quad B_{12eq} = E\left[-\frac{\partial^2 Q^+}{\partial \pi \partial \theta_1}\Big|_{\Phi=0}\right], \quad B_{13eq} = E\left[-\frac{\partial^2 Q^+}{\partial \pi \partial \theta_2}\Big|_{\Phi=0}\right],$$

$$B_{22eq} = E\left[-\frac{\partial^2 Q^+}{\partial \theta_1^2}\Big|_{\Phi=0}\right], \quad B_{23eq} = E\left[-\frac{\partial^2 Q^+}{\partial \theta_1 \partial \theta_2}\Big|_{\Phi=0}\right] \text{ and}$$

$$B_{33eq} = E\left[-\frac{\partial^2 Q^+}{\partial \theta_2^2}\Big|_{\Phi=0}\right].$$

After taking expectation of the negative of mixed partial derivatives and on simplification, the above quantities are obtained as follows

$$\psi_{1eq} = \sum_{j=1}^{m_1}\left[\frac{1}{1+(n_{1j}-1)\theta_1}\left\{\frac{y_{1j}}{\pi} - \frac{n_{1j}-y_{1j}}{1-\pi}\right\}\right],$$

$$D_{11eq} = \sum_{j=1}^{m_1}\left[\frac{1}{1+(n_{1j}-1)\theta_1}\left\{\frac{n_{1j}}{\pi(1-\pi)}\right\}\right],$$

$$A_{11eq} = \sum_{j=1}^{m_1} \left[\frac{1}{1 + (n_{1j} - 1)\theta_1} \left\{ \frac{n_{1j}}{\pi(1-\pi)} \right\} \right], \quad A_{12eq} = 0, \quad A_{13eq} = 0,$$

$$B_{11eq} = \sum_{i=1}^{2} \sum_{j=1}^{m_1} \left[\frac{1}{1 + (n_{ij} - 1)\theta_i} \left\{ \frac{n_{ij}}{\pi(1-\pi)} \right\} \right], \quad B_{12eq} = 0, \quad B_{13eq} = 0,$$

$$\begin{aligned} B_{22eq} = \frac{1}{2} \sum_{j=1}^{m_1} \Bigg[&- \frac{(n_{1j} - 1)^2}{\{1 + (n_{1j} - 1)\theta_1\}^2} \\ &- \frac{4\,(n_{1j} - 1)^2}{\{1 + (n_{1j} - 1)\theta_1\}^3} \left\{ E\left(y_{1j} \log\left(\frac{\pi}{z_{1j}}\right)\right) + n_{1j} E\left(\log\left(\frac{1-\pi}{1-z_{1j}}\right)\right) \right. \\ &\left. - E\left(y_{1j} \log\left(\frac{1-\pi}{1-z_{1j}}\right)\right) \right\} \Bigg], \end{aligned}$$

$$B_{23eq} = 0 \quad \text{and}$$

$$\begin{aligned} B_{33eq} = \frac{1}{2} \sum_{j=1}^{m_2} \Bigg[&- \frac{(n_{2j} - 1)^2}{\{1 + (n_{2j} - 1)\theta_2\}^2} \\ &- \frac{4\,(n_{2j} - 1)^2}{\{1 + (n_{2j} - 1)\theta_2\}^3} \left\{ E\left(y_{2j} \log\left(\frac{\pi}{z_{2j}}\right)\right) + n_{2j} E\left(\log\left(\frac{1-\pi}{1-z_{2j}}\right)\right) \right. \\ &\left. - E\left(y_{2j} \log\left(\frac{1-\pi}{1-z_{2j}}\right)\right) \right\} \Bigg]. \end{aligned}$$

If we denote the estimate of δ based on the extended quasi-likelihood function by $\tilde{\delta}_{eq}$ and substitute these estimates in ψ_{1eq}, D_{11eq}, A_{eq} and B_{eq}, then the $C(\alpha)$ test based on the extended quasi-likelihood is $C_{eq} = \psi_{1eq}^2 \,/\, \left(D_{11eq} - A_{eq} B_{eq}^{-1} A'_{eq}\right)$, which is distributed as chi-squared, asymptotically as $m \to \infty$, where $m = m_1 + m_2$, with 1 degree of freedom.

4.3.4 The Rao-Scott (RS) and the Adjusted Rao-Scott (ARS) statistics

Rao and Scott (1992) proposed a simple method which is based on the concepts of design effect and effective sample size for comparing independent groups of clustered binary data. They applied this method to a variety of biometrical problems including testing homogeneity of binomial proportions. Following Rao and Scott (1992) define

$$y_i = \sum_{j=1}^{m_i} y_{ij}, \quad n_i = \sum_{j=1}^{m_i} n_{ij}, \quad \hat{p}_i = \frac{y_i}{n_i}, \quad \hat{p} = \frac{\sum_{i=1}^{2} y_i}{\sum_{i=1}^{2} n_i},$$

$$v_i = m_i(m_i - 1)^{-1} n_i^{-2} \sum_{j=1}^{m_i} (y_{ij} - n_{ij}\hat{p}_i)^2, \quad d_i = \frac{n_i v_i}{\{\hat{p}_i(1 - \hat{p}_i)\}},$$

$$\tilde{y}_i = \frac{y_i}{d_i}, \quad \tilde{n}_i = \frac{n_i}{d_i}, \text{ and } \quad \tilde{p} = \frac{\sum_{i=1}^{2} \tilde{y}_i}{\sum_{i=1}^{2} \tilde{n}_i}, \quad i = 1, 2.$$

In sampling design literature the term d_i is called the design effect or the variance inflation factor where as the term $\tilde{n}_i$ is called the effective sample size (Kish (1965)). The RS statistic for testing the equality of two proportions is

$$C_{rs} = \sum_{i=1}^{2} \frac{(\tilde{y}_i - \tilde{n}_i\tilde{p})^2}{\{\tilde{n}_i\tilde{p}(1 - \tilde{p})\}}.$$

It is evident that rather than assuming any specific model for the intracluster correlations, the RS statistic uses the binomial model for the overall response y_i in the i^{th}

cluster. An adjustment for the variance inflation due to clustering or the design effect is also taken into consideration in this statistic. For this reason, the RS statistic is considered to be based on semi-parametric model. The RS statistic is asymptotically distributed as χ^2 with one degree of freedom under the null hypothesis. If the population variance inflation factors D_i's are equal, that is, $D_i = D$ for $i = 1, 2$, which is a special case, Rao and Scott (1992) recommend using $\tilde{y}_i = \frac{y_i}{d}$ and $\tilde{n}_i = \frac{n_i}{d}$, where d is a pooled estimate of D and is defined as

$$d = \left[\sum_{i=1}^{2} (1 - f_i) \frac{\hat{p}_i(1 - \hat{p}_i)}{\hat{p}(1 - \hat{p})} d_i \right],$$

with $f_i = \frac{n_i}{n}$ and $n = \sum_{i=1}^{2} n_i$. Taking this modification into account, the RS statistic takes the form $C_{ars} = \frac{\chi^2}{d}$, where

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

We call this modified RS statistic as adjusted Rao-Scott statistic (ARS) which, under the null hypothesis, is distributed asymptotically as a χ^2 with 1 degree of freedom.

4.4 Simulation Study

In this section we report on a simulation study to compare the test statistics LR, C_{ml}, C_{kmm}, C_{qb}, C_{qs}, C_{eq}, C_{rs} and C_{ars} in terms of empirical size and power. The litter sizes within each group were required to generate data for the simulation study. The litter sizes and the number of litters were chosen as those of the control group

($m_1 = 27$) and medium dose group ($m_2 = 21$) of Paul (1982). Thus, the litter sizes of group 1 were 12, 7, 6, 6, 7, 8, 10, 7, 8, 6, 11, 7, 8, 9, 2, 7, 9, 7, 11, 10, 4, 8, 10, 12, 8, 7, 8 and those of group 2 were 4, 4, 9, 8, 9, 7, 8, 9, 6, 4, 6, 7, 3, 13, 6, 8, 11, 7, 6, 10, 6. Data were generated using "*R*" through the function "*rbetabinom*" with arguments $size =$ the litter sizes (n_{ij}), $prob = \pi_i$, and $rho = \theta_i$, within group i, where $j = 1, 2, \cdots, n_i$, and $i = 1,\ 2$. For simultaneously solving the equations to obtain the estimates of parameters we used the "*R*" built-in function "*nleqslv*". For each simulation run, 3000 valid iterations were considered after discarding the non-convergent samples and those that produced out-of-range estimates of the parameters. The dispersion parameter, θ, of a beta-binomial distribution can take negative value as well. So, we considered the negative estimates of θ_1 and θ_2. Since $\cap_j \left(\dfrac{-1}{n_{ij}-1}\right) < \theta_i < 1$, an estimate of θ_i was considered out-of-range if $\hat{\theta}_i < \dfrac{-1}{\max\{n_{ij}\}-1} + 0.00001$, where $\max\{n_{ij}\}$ is the largest value of $\{n_{ij}\}$; $i = 1,\ 2; j = 1,\ 2, \cdots, m_i$. Convergence of estimating equations depend on the initial values provided for the parameters. If the same initial values were provided for all iterations then a large number of samples became non-convergent. To keep the number of convergent samples to its maximum we used method of moments estimates of parameters as initial values for each iteration. Also in order to avoid the undefined estimates we discarded samples if $y_{ij} = 0$ or $y_{ij} = n_{ij}$ for all $i = 1,\ 2; j = 1,\ 2, \cdots, m_i$. We observed that the number of discarded samples depend on the arguments, π and θ, of the function "*rbetabinom*" and this number was more for small values of π and close to the boundary values of θ (close to 0 or 1).

As we considered unequal dispersion parameters θ_1 and θ_2 in developing the test procedures, for computing the empirical levels and powers of the test statistics, we took four sets of combinations of (θ_1, θ_2) into account. The combinations were (θ_1, θ_2)

= (0.02, 0.02), (0.02, 0.05), (0.02, 0.10), (0.02, 0.20), (0.02, 0.30), (0.02, 0.40); (θ_1, θ_2) = (0.05, 0.05), (0.05, 0.10), (0.05, 0.20), (0.05, 0.30), (0.05, 0.40), (0.05, 0.50); (θ_1, θ_2) = (0.10, 0.20), (0.10, 0.15), (0.10, 0.20), (0.10, 0.30), (0.10, 0.40), (0.10, 0.50) and (θ_1, θ_2) = (0.20, 0.20),(0.20, 0.30), (0.20, 0.40), (0.20, 0.50), (0.20, 0.60), (0.20, 0.70).

In order to compute empirical levels we considered $\pi_1 = \pi_2 =$ 0.05, 0.10, 0.20, 0.30 for all of the above four combinations of (θ_1, θ_2) and the results corresponding to 5% nominal level are presented in Tables 4.1 to 4.4 . We observe the following features of the test statistics in maintaining the nominal levels.

a) For small values of $\pi_1 = \pi_2(0.05)$, the $C(\alpha)$ statistic based on the Kleinman's method of moment estimates of nuisance parameters, δ, show conservative behaviour in maintaining nominal level for all combinations of (θ_1, θ_2), from no difference between then to moderately large difference. For example, in Table 4.1 the level of C_{kmm} is 2.7% for (θ_1, θ_2) = (0.02, 0.02) and it raises to 4.1% for (θ_1, θ_2) = (0.02, 0.20). Tables 4.2, 4.3, and 4.4 also display the similar results for this test procedure.

b) The trends of the statistics C_{rs} and C_{ars} are also similar to that of C_{kmm} in maintaining nominal levels. For example, the empirical level of C_{rs} is 3.5% for (θ_1, θ_2) = (0.20, 0.20) and 4.1% for (θ_1, θ_2) = (0.20, 0.40) as seen in Table 4.4.

c) For moderate values of π, that is, for $\pi_1 = \pi_2 = 0.10$ and the first two combinations of (θ_1, θ_2), the statistic C_{kmm} behaves conservatively in maintaining nominal levels from no to moderate differences of the dispersion parameters. For example, in Table 4.2, we see that 3.6% is the empirical level of this test procedure for (θ_1, θ_2) = (0.05, 0.05) and 4.1% for (θ_1, θ_2) = (0.05, 0.20).

d) For the same values of the dispersion parameters θ_1 and θ_2, that is, when θ_1 and θ_2 are equal in all four combinations, all of the test procedures are conservative with respect to empirical levels if values of $\pi_1 = \pi_2$ are small ($\pi_1 = \pi_2 = 0.05$).

e) All test procedures maintain the nominal level reasonably well for all combinations of the dispersion parameters θ_1 and θ_2 apart from the above cases. The statistics C_{qb}, C_{qs} and C_{eq} maintain nominal level better than the other statistics. For example, in Table 4.3 for $\pi_1 = \pi_2 = 0.30$ and $(\theta_1,\ \theta 2) = (0.10,\ 0.10)$, the empirical levels of the statistics C_{qb}, C_{qs} and C_{eq} are 5.0%, 5.1% and 5.3% respectively which are close to the nominal level.

For all of the statistics we computed empirical powers corresponding to the 5% nominal significance level for combinations of $(\theta_1,\ \theta_2)$ mentioned earlier. In computing powers the values of $(\pi_1,\ \pi_2)$ taken into consideration for all combinations of $(\theta_1,\ \theta_2)$ were $(\pi_1,\ \pi_2)$ = (0.05, 0.10), (0.05, 0.15), (0.05, 0.20), (0.05, 0.25), (0.05, 0.30), (0.05, 0.35), (0.05, 0.40), (0.05, 0.45), (0.05, 0.50) and (0.05, 0.60) and the results are presented in Tables 4.5 to 4.8. The test statistics were of two categories, namely, parametric and semi-parametric. The parametric procedures were LR, C_{ml} and C_{kmm} and the semi-parametric procedures were C_{qb}, C_{qs}, C_{eq}, C_{rs} and C_{ars}. In what follows we present how the test procedures behave with respect to the empirical powers.

a) For specific values of $(\pi_1,\ \pi_2)$, the powers of all of the test procedures increase with the increase in the difference between the dispersion parameters, θ_1 and θ_2. For example, in Table 4.5, the power of the likelihood ratio statistic, LR, is 7.6% for $(\theta_1,\ \theta_2) = (0.02,\ 0.02)$ and it raises to 58.9% for $(\theta_1,\ \theta_2) = (0.02,\ 0.40)$ when $(\pi_1,\ \pi_2) = (0.05,\ 0.10)$.

b) The powers of all of the test procedures increase with the increase in the difference between π_1 and π_2. We see in Table 4.6 that for $(\theta_1,\ \theta_2) = (0.05,\ 0.20)$ the power of C_{ml} is 22.4% when $(\pi_1,\ \pi_2) = (0.05,\ 0.10)$ and the increase of π_2 to 0.30 leads to the empirical power increase to 81.3%.

c) Of the parametric test procedures, the likelihood ratio statistic maintains the highest power in all combinations of $(\theta_1,\ \theta_2)$ and for all values of $(\pi_1,\ \pi_2)$ which is followed by C_{kmm} and C_{ml} except in some situations where C_{ml} is as good as C_{kmm}. In Table 4.7, the powers of LR, C_{ml} and C_{kmm} are 7.8%, 7.4% and 7.4% respectively for $(\theta_1,\ \theta_2) = (0.10,\ 0.10)$ and $(\pi_1,\ \pi_2) = (0.05,\ 0.10)$ and the powers of these statistics are 75.1%, 73.4% and 74.3% for $(\pi_1,\ \pi_2) = (0.05,\ 0.35)$.

d) Among the test statistics based on the semi-parametric models, the Rao-Scott statistic exhibit the highest powers in all of the parameter combinations considered. Apart from the RS and ARS statistics, the test statistic based on the extended quasi-likelihood estimates of the nuisance parameters performs best in terms of power performance followed by C_{qs} and C_{qb} for all combinations of dispersion parameters, $(\theta_1,\ \theta_2)$ and all values of $(\pi_1,\ \pi_2)$. It can be seen in Table 4.8 under column of (0.20, 0.40) indicating values of $(\theta_1,\ \theta_2)$ corresponding to $(\pi_1,\ \pi_2) = (0.05,\ 0.10)$ that the powers of C_{eq}, C_{qs} and C_{qb} are 22.8%, 21.7% and 19.3% respectively and the powers of C_{rs} and C_{ars} are 26.3% and 25.6% respectively.

4.5 Examples and Discussion

We considered data from a toxicological experiment reported in Paul (1982). The data represent the number of live foetuses, n_{ij}, and the number of affected live foetuses by treatments, y_{ij}, for each of four doses of treatments, namely, control (C), low (L), medium (M) and high (H) dose groups. We present the data in Table 4.9 and perform an analysis for all possible pairs of treatment groups. The estimates of the parameters obtained by different methods that were required for calculating the test statistics are reported in Table 4.10 and the values of the test statistics along with the corresponding p-values are presented in Table 4.11. As in the simulation study we considered the litter sizes of those of control and medium dose groups, we focus on the values of test statistics for this particular treatment combination which is given in column 3 of Table 4.11. Though all of the test procedures reject the hypothesis of equality of proportions there are some remarkable features in the values of the test statistics. Among the three parametric test procedures the LR gives the largest value and consequently the smallest p-value which indicates that the likelihood ratio test procedure rejects the hypothesis most strongly. The LR statistic is followed by the $C(\alpha)$ statistic, C_{kmm} and the score statistic C_{ml}. Of the semi-parametric test procedures, the value of Rao-Scott statistic, C_{rs}, is the largest which indicates that this statistic is the most powerful. Other than the C_{rs} and C_{ars} within the category of semi-parametric tests, the C_{eq} statistic has the largest value which is followed by C_{qs} and C_{qb} which indicates the superiority of the $C(\alpha)$ based on the extended quasi-likelihood estimates of nuisance parameters over the $C(\alpha)$ statistics based on the quasi-likelihood and method of moments estimates.

While the C_{rs} and C_{ars} statistics are liberal in maintaining nominal levels for $\pi = 0.30$ and all combinations of the dispersion parameters θ_1 and θ_2, they show some conservative behaviour for small values of π. For moderate values of π these test procedures behave conservatively for small to moderate differences between the dispersion parameters but exhibit liberal behaviour for large differences. In addition, not only the performances with respect to powers are better but these statistics have additional features such as computational ease and simplicity. The likelihood ratio statistic is the only one that requires the estimates of parameters both under the null and the alternative hypotheses. This statistic shows some upward trend in maintaining nominal level for large values of π and large differences in the dispersion parameters. For small to moderate differences in the dispersion parameters and small to moderate values of π it shows some slightly conservative behaviour. But overall this statistic maintains nominal level very well. Among the parametric test procedures, the likelihood ratio statistic demonstrates highest power for all parameter combinations. The score statistic, C_{ml}, and the $C(\alpha)$ statistic C_{kmm}, require the estimates of the parameters only under the null hypothesis. The latter shows higher power than the former in all instances. Both are conservative for small differences in $(\theta_1,\, \theta_2)$ and small values of π. In other combinations these statistics are liberal. The $C(\alpha)$ statistic based on the extended quasi-likelihood, C_{eq}, performs best in terms of powers among C_{qb}, C_{qs} and C_{eq} and for this statistic the empirical level is closer to the nominal level. These three statistics are also conservative for some combinations of the dispersion parameters, particularly when the differences in θ_1 and θ_2 are small accompanied by the small values of π. In other instances they are somewhat liberal, although, overall, they maintain the nominal level reasonably well.

Table 4.1: Empirical level (%) of test statistics, LR, C_{ml}, C_{kmm}, C_{qb}, C_{qs}, C_{eq}, C_{rs}, C_{ars}; based on $3,000$ replications and $\alpha = 0.05$ (Control and Medium Dose Groups)

Test Statistic	$\pi_1 = \pi_2$	(θ_1, θ_2)					
	0.05	(.02, .02)	(.02, .05)	(.02, .10)	(.02, .20)	(.02, .30)	(.02, .40)
LR		3.5	3.7	4.2	4.8	5.3	5.8
C_{ml}		3.5	4.3	5.1	5.7	6.2	6.6
C_{kmm}		2.7	2.9	3.5	4.1	4.7	5.2
C_{qb}		3.7	4.1	5.2	5.6	5.9	6.4
C_{qs}		3.8	4.2	5.4	5.6	5.9	6.5
C_{eq}		4.0	4.3	4.8	5.2	5.5	5.9
C_{rs}		3.3	3.6	3.9	4.2	4.5	4.6
C_{ars}		3.4	3.6	4.1	4.4	4.7	4.8
	0.10						
LR		4.1	4.3	4.7	5.0	5.3	5.5
C_{ml}		4.4	4.6	5.0	5.4	5.6	5.7
C_{kmm}		3.8	3.9	4.3	4.6	5.0	5.2
C_{qb}		4.3	4.6	5.0	5.3	5.4	5.5
C_{qs}		3.9	4.3	5.1	5.3	5.5	5.6
C_{eq}		3.9	4.4	4.6	5.2	5.3	5.6
C_{rs}		3.8	4.5	4.7	4.9	5.0	5.4
C_{ars}		4.0	4.4	4.7	5.1	5.1	5.5
	0.20						
LR		4.4	4.5	4.8	4.8	5.1	5.3
C_{ml}		5.1	5.2	5.2	5.4	5.5	5.1
C_{kmm}		4.2	4.3	4.7	4.7	5.1	5.2
C_{qb}		4.5	4.6	5.0	5.1	5.2	5.6
C_{qs}		4.7	4.9	5.3	5.5	5.7	5.9
C_{eq}		4.8	5.0	5.2	5.3	5.5	5.6
C_{rs}		4.3	4.5	4.9	5.2	5.5	5.7
C_{ars}		4.4	4.6	5.1	5.3	5.7	6.3
	0.30						
LR		5.2	5.3	5.6	5.6	5.8	5.8
C_{ml}		4.9	5.3	5.4	5.6	5.9	6.2
C_{kmm}		5.2	5.4	5.7	5.7	6.2	6.6
C_{qb}		4.7	4.9	5.4	5.5	5.8	6.3
C_{qs}		5.0	5.1	5.4	5.6	6.1	6.5
C_{eq}		5.0	5.1	5.4	5.5	5.5	5.5
C_{rs}		5.5	5.6	5.9	6.0	6.3	6.7
C_{ars}		5.5	5.7	6.1	6.3	6.7	6.9

Table 4.2: Empirical level (%) of test statistics, LR, C_{ml}, C_{kmm}, C_{qb}, C_{qs}, C_{eq}, C_{rs}, C_{ars}; based on $3,000$ replications and $\alpha = 0.05$ (Control and Medium Dose Groups)

Test Statistic	$\pi_1 = \pi_2$	(θ_1, θ_2)					
	0.05	(.05, .05)	(.05, .10)	(.05, .20)	(.05, .30)	(.05, .40)	(.05, .50)
LR		3.3	3.5	4.0	4.6	5.2	5.6
C_{ml}		3.3	4.2	4.9	5.5	6.0	6.4
C_{kmm}		2.5	2.7	3.3	3.9	4.5	5.0
C_{qb}		3.5	3.9	5.0	5.4	5.7	6.2
C_{qs}		3.7	4.1	5.1	5.4	5.7	6.2
C_{eq}		3.8	4.1	4.6	5.0	5.3	5.7
C_{rs}		3.1	3.4	3.7	4.0	4.3	4.4
C_{ars}		3.2	3.4	3.9	4.2	4.5	4.6
	0.10						
LR		3.9	4.2	4.5	4.8	5.1	5.3
C_{ml}		4.2	4.4	4.8	5.2	5.4	5.5
C_{kmm}		3.6	3.7	4.1	4.4	4.8	5.0
C_{qb}		4.1	4.4	4.8	5.1	5.3	5.3
C_{qs}		3.7	4.2	4.8	5.1	5.2	5.4
C_{eq}		3.7	4.2	4.4	5.0	5.1	5.4
C_{rs}		3.6	4.3	4.5	4.7	4.8	5.2
C_{ars}		3.8	4.2	4.5	4.9	4.9	5.3
	0.20						
LR		4.2	4.3	4.6	4.6	4.9	5.2
C_{ml}		4.9	5.0	5.0	5.2	5.3	4.9
C_{kmm}		4.0	4.1	4.5	4.5	4.9	5.0
C_{qb}		4.3	4.4	4.8	4.9	5.0	5.4
C_{qs}		4.5	4.8	5.1	5.3	5.5	5.7
C_{eq}		4.6	4.8	5.0	5.1	5.3	5.4
C_{rs}		4.1	4.3	4.7	5.0	5.3	5.5
C_{ars}		4.2	4.4	4.9	5.1	5.5	6.1
	0.30						
LR		5.0	5.1	5.4	5.4	5.6	5.6
C_{ml}		4.7	5.1	5.2	5.4	5.7	6.0
C_{kmm}		5.0	5.2	5.5	5.5	6.0	6.4
C_{qb}		4.5	4.7	5.2	5.3	5.6	6.1
C_{qs}		4.8	4.8	5.4	5.4	6.0	6.3
C_{eq}		4.8	4.9	5.2	5.3	5.3	5.3
C_{rs}		5.3	5.4	5.7	5.8	6.1	6.5
C_{ars}		5.3	5.5	5.9	6.1	6.5	6.7

Table 4.3: Empirical level (%) of test statistics, LR, C_{ml}, C_{kmm}, C_{qb}, C_{qs}, C_{eq}, C_{rs}, C_{ars}; based on $3,000$ replications and $\alpha = 0.05$ (Control and Medium Dose Groups)

Test Statistic	$\pi_1 = \pi_2$	(θ_1, θ_2)					
	0.05	(.10, .10)	(.10, .15)	(.10, .20)	(.10, .30)	(.10, .40)	(.10, .50)
LR		3.8	4.0	4.5	5.1	5.6	6.1
C_{ml}		3.8	4.6	5.4	6.0	6.5	6.9
C_{kmm}		3.0	3.2	3.8	4.4	5.0	5.5
C_{qb}		4.0	4.4	5.5	5.9	6.2	6.7
C_{qs}		4.1	4.5	5.5	6.0	6.2	6.7
C_{eq}		4.3	4.6	5.1	5.5	5.8	6.1
C_{rs}		3.6	3.9	4.2	4.5	4.8	4.9
C_{ars}		3.7	3.9	4.4	4.5	5.0	5.1
	0.10						
LR		4.4	4.6	5.0	5.3	5.6	5.8
C_{ml}		4.7	4.9	5.3	5.7	5.9	6.0
C_{kmm}		4.1	4.2	4.7	4.9	5.3	5.5
C_{qb}		4.7	4.9	5.3	5.6	5.7	5.8
C_{qs}		4.2	4.6	5.3	5.6	5.7	5.9
C_{eq}		4.2	4.7	4.9	5.5	5.6	5.9
C_{rs}		4.1	4.8	5.0	5.2	5.3	5.7
C_{ars}		4.3	4.7	5.0	5.4	5.4	5.8
	0.20						
LR		4.7	4.8	5.1	5.1	5.4	5.6
C_{ml}		5.4	5.5	5.5	5.7	5.8	5.4
C_{kmm}		4.5	4.5	5.0	5.0	5.4	5.5
C_{qb}		4.8	4.9	5.3	5.4	5.5	5.9
C_{qs}		5.0	5.1	5.6	5.8	6.0	6.2
C_{eq}		5.1	5.3	5.5	5.6	5.8	5.9
C_{rs}		4.6	4.8	5.2	5.5	5.8	6.0
C_{ars}		4.7	4.9	5.4	5.6	6.0	6.6
	0.30						
LR		5.5	5.6	5.9	5.9	6.1	6.1
C_{ml}		5.2	5.6	5.7	5.9	6.2	6.5
C_{kmm}		5.5	5.7	6.0	6.0	6.5	6.9
C_{qb}		5.0	5.2	5.7	5.8	6.1	6.6
C_{qs}		5.1	5.4	5.7	5.9	6.3	6.7
C_{eq}		5.3	5.4	5.7	5.8	5.8	5.8
C_{rs}		5.8	5.9	6.2	6.3	6.6	7.0
C_{ars}		5.8	6.0	6.4	6.6	7.0	7.2

Table 4.4: Empirical level (%) of test statistics, LR, C_{ml}, C_{kmm}, C_{qb}, C_{qs}, C_{eq}, C_{rs}, C_{ars}; based on 3,000 replications and $\alpha = 0.05$ (Control and Medium Dose Groups)

Test Statistic	$\pi_1 = \pi_2$	(θ_1, θ_2)					
	0.05	(.20, .20)	(.20, .30)	(.20, .40)	(.20, .50)	(.20, .60)	(.20, .70)
LR		3.7	3.9	4.4	5.0	5.5	6.0
C_{ml}		3.7	4.5	5.3	5.9	6.4	6.8
C_{kmm}		2.9	3.1	3.7	4.3	4.9	5.4
C_{qb}		3.9	4.3	5.4	5.8	6.1	6.6
C_{qs}		4.1	4.3	5.6	5.9	6.0	6.7
C_{eq}		4.2	4.5	5.0	5.4	5.7	6.1
C_{rs}		3.5	3.8	4.1	4.4	4.7	4.8
C_{ars}		3.6	3.8	4.3	4.6	4.9	5.0
	0.10						
LR		4.3	4.5	4.9	5.2	5.5	5.7
C_{ml}		4.6	4.8	5.2	5.6	5.8	5.9
C_{kmm}		4.0	4.1	4.5	4.8	5.2	5.4
C_{qb}		4.5	4.8	5.2	5.5	5.6	5.7
C_{qs}		4.1	4.5	5.3	5.6	5.7	5.7
C_{eq}		4.1	4.6	4.8	5.4	5.5	5.8
C_{rs}		4.0	4.7	4.9	5.1	5.2	5.6
C_{ars}		4.2	4.6	4.9	5.3	5.3	5.7
	0.20						
LR		4.6	4.7	5.0	5.0	5.3	5.5
C_{ml}		5.3	5.4	5.4	5.6	5.7	5.3
C_{kmm}		4.4	4.5	4.9	4.9	5.3	5.4
C_{qb}		4.7	4.8	5.2	5.3	5.4	5.8
C_{qs}		4.9	5.0	5.5	5.8	6.1	6.2
C_{eq}		5.0	5.2	5.4	5.5	5.7	5.8
C_{rs}		4.5	4.7	5.1	5.4	5.7	5.9
C_{ars}		4.6	4.8	5.3	5.5	5.9	6.5
	0.30						
LR		5.4	5.5	5.8	5.8	6.0	6.0
C_{ml}		5.1	5.5	5.6	5.8	6.1	6.4
C_{kmm}		5.4	5.6	5.9	5.9	6.4	6.8
C_{qb}		4.9	5.1	5.6	5.7	6.0	6.5
C_{qs}		5.2	5.3	5.7	6.0	6.3	6.7
C_{eq}		5.2	5.3	5.6	5.7	5.7	5.7
C_{rs}		5.7	5.8	6.1	6.2	6.5	6.9
C_{ars}		5.7	5.9	6.3	6.5	6.9	7.1

Table 4.5: Empirical power (%) of test statistics, LR, C_{ml}, C_{kmm}, C_{qb}, C_{qs}, C_{eq}, C_{rs}, C_{ars}; based on 3,000 replications and $\alpha = 0.05$ (Control and Medium Dose Groups)

Test Statistic	(π_1, π_2)	(θ_1, θ_2)					
	(0.05, 0.10)	(.02, .02)	(.02, .05)	(.02, .10)	(.02, .20)	(.02, .30)	(.02, .40)
LR		7.6	12.5	22.9	41.1	52.1	58.9
C_{ml}		7.2	11.1	21.5	38.0	49.7	55.6
C_{kmm}		7.2	11.9	22.4	39.4	51.0	56.9
C_{qb}		6.2	9.8	17.6	30.1	42.4	49.7
C_{qs}		6.8	10.5	19.0	33.2	45.4	52.3
C_{eq}		6.9	10.7	21.0	35.3	46.8	52.9
C_{rs}		8.6	13.3	24.6	43.2	54.4	62.1
C_{ars}		8.3	12.9	23.9	42.4	53.6	61.3
	(0.05, 0.15)						
LR		18.1	27.0	39.1	52.1	69.0	85.4
C_{ml}		16.0	22.3	35.6	47.9	64.9	79.6
C_{kmm}		16.9	24.0	37.8	50.0	67.1	83.9
C_{qb}		11.9	17.2	28.7	42.6	55.7	71.2
C_{qs}		13.7	18.8	32.2	43.7	59.0	74.1
C_{eq}		14.3	20.1	33.1	45.9	61.2	75.5
C_{rs}		20.1	30.1	45.0	55.7	73.6	91.2
C_{ars}		19.8	28.7	43.2	53.2	72.4	87.9
	(0.05, 0.20)						
LR		33.4	43.2	53.9	73.1	82.7	96.7
C_{ml}		31.1	39.0	50.1	69.4	79.0	92.3
C_{kmm}		32.7	41.8	52.0	72.0	80.9	94.6
C_{qb}		24.4	31.9	43.5	58.9	67.4	81.8
C_{qs}		27.1	34.0	46.3	64.3	72.9	84.8
C_{eq}		28.3	36.5	48.1	66.2	74.9	86.3
C_{rs}		36.9	46.2	57.1	76.1	84.7	99.4
C_{ars}		35.1	44.8	55.4	74.2	83.4	98.0
	(0.05, 0.25)						
LR		51.2	58.9	72.2	86.7	95.2	100
C_{ml}		47.9	54.4	64.7	81.1	90.8	97.7
C_{kmm}		49.3	57.0	68.9	83.4	92.5	98.5
C_{qb}		38.3	45.7	56.1	72.0	81.4	90.4
C_{qs}		41.8	49.6	58.3	74.7	84.9	92.9
C_{eq}		44.7	51.2	60.2	76.7	87.0	95.7
C_{rs}		56.4	63.5	77.8	89.7	98.3	100
C_{ars}		53.6	61.2	75.3	88.0	97.1	100
	(0.05, 0.30)						
LR		65.6	72.4	84.5	96.7	100	100
C_{ml}		62.3	69.8	80.4	95.1	100	100
C_{kmm}		64.2	71.1	82.6	95.8	100	100
C_{qb}		54.7	62.7	74.1	89.9	96.4	100
C_{qs}		58.1	65.8	77.3	91.9	98.4	100
C_{eq}		60.1	66.5	78.0	94.2	99.7	100
C_{rs}		68.9	76.1	89.1	99.4	100	100
C_{ars}		67.1	74.3	87.3	97.8	100	100

Table 4.5: (continued)

Test Statistic	(π_1, π_2)	(θ_1, θ_2)					
	(0.05, 0.35)	(.02, .02)	(.02, .05)	(.02, .10)	(.02, .20)	(.02, .30)	(.02, .40)
LR		74.9	83.1	95.4	100	100	100
C_{ml}		73.2	80.2	92.8	100	100	100
C_{kmm}		74.1	81.9	93.7	100	100	100
C_{qb}		66.2	74.7	87.2	100	100	100
C_{qs}		69.7	77.2	89.7	100	100	100
C_{eq}		71.4	78.3	91.0	100	100	100
C_{rs}		77.8	87.1	97.6	100	100	100
C_{ars}		76.3	85.2	96.3	100	100	100
	(0.05, 0.40)						
LR		82.3	95.2	100	100	100	100
C_{ml}		78.3	91.2	100	100	100	100
C_{kmm}		80.1	93.3	100	100	100	100
C_{qb}		73.4	86.3	98.7	100	100	100
C_{qs}		75.5	88.7	100	100	100	100
C_{eq}		76.7	89.3	100	100	100	100
C_{rs}		85.2	98.3	100	100	100	100
C_{ars}		84.1	96.9	100	100	100	100
	(0.05, 0.45)						
LR		91.2	100	100	100	100	100
C_{ml}		88.0	100	100	100	100	100
C_{kmm}		89.3	100	100	100	100	100
C_{qb}		83.4	97.8	100	100	100	100
C_{qs}		85.9	99.7	100	100	100	100
C_{eq}		86.8	99.9	100	100	100	100
C_{rs}		94.1	100	100	100	100	100
C_{ars}		92.7	100	100	100	100	100
	(0.05, 0.50)						
LR		100	100	100	100	100	100
C_{ml}		100	100	100	100	100	100
C_{kmm}		100	100	100	100	100	100
C_{qb}		97.3	100	100	100	100	100
C_{qs}		99.9	100	100	100	100	100
C_{eq}		100	100	100	100	100	100
C_{rs}		100	100	100	100	100	100
C_{ars}		100	100	100	100	100	100
	(0.05, 0.60)						
LR		100	100	100	100	100	100
C_{ml}		100	100	100	100	100	100
C_{kmm}		100	100	100	100	100	100
C_{qb}		100	100	100	100	100	100
C_{qs}		100	100	100	100	100	100
C_{eq}		100	100	100	100	100	100
C_{rs}		100	100	100	100	100	100
C_{ars}		100	100	100	100	100	100

Table 4.6: Empirical power (%) of test statistics, LR, C_{ml}, C_{kmm}, C_{qb}, C_{qs}, C_{eq}, C_{rs}, C_{ars}; based on $3,000$ replications and $\alpha = 0.05$ (Control and Medium Dose Groups)

Test Statistic	(π_1, π_2)	(θ_1, θ_2)					
	(0.05, 0.10)	(.05, .05)	(.05, .10)	(.05, .20)	(.05, .30)	(.05, .40)	(.05, .50)
LR		8.0	13.3	23.9	42.6	53.4	60.3
C_{ml}		7.6	11.9	22.4	39.5	50.9	57.0
C_{kmm}		7.6	12.7	23.6	40.9	52.2	58.3
C_{qb}		6.6	10.6	18.7	31.7	43.6	51.1
C_{qs}		7.2	11.3	20.9	36.3	47.1	54.0
C_{eq}		7.3	11.5	22.2	36.8	48.0	54.3
C_{rs}		9.0	14.1	25.5	44.7	55.6	63.5
C_{ars}		8.7	13.7	25.0	43.9	54.8	62.8
	(0.05, 0.15)						
LR		18.5	27.8	40.2	53.5	70.2	86.8
C_{ml}		16.4	23.1	36.6	49.4	66.3	81.0
C_{kmm}		17.3	24.8	38.9	51.5	68.1	85.3
C_{qb}		12.3	18.0	29.8	43.9	56.9	72.6
C_{qs}		13.9	20.1	33.7	46.4	61.1	75.8
C_{eq}		14.7	20.9	34.3	47.4	62.6	76.9
C_{rs}		20.5	31.0	46.2	57.2	74.8	92.6
C_{ars}		20.2	29.5	44.5	54.7	73.6	89.5
	(0.05, 0.20)						
LR		33.8	44.0	55.0	74.6	83.9	98.3
C_{ml}		31.5	39.8	51.2	70.9	80.3	93.7
C_{kmm}		33.1	42.6	53.0	73.5	82.1	96.0
C_{qb}		24.8	32.7	44.7	60.3	68.7	83.2
C_{qs}		27.6	35.7	48.0	65.9	75.1	85.9
C_{eq}		28.7	37.3	49.1	67.7	76.1	87.7
C_{rs}		37.3	47.1	58.0	77.6	85.9	100
C_{ars}		35.5	45.6	56.5	75.7	84.7	99.6
	(0.05, 0.25)						
LR		51.6	59.7	73.1	88.2	96.4	100
C_{ml}		48.3	55.2	65.8	82.6	92.1	99.3
C_{kmm}		49.7	57.8	70.0	84.9	93.7	99.9
C_{qb}		38.7	46.5	57.3	73.6	82.6	91.8
C_{qs}		42.3	50.2	60.1	76.5	85.9	95.7
C_{eq}		45.1	52.0	61.3	78.2	88.2	97.3
C_{rs}		56.8	64.3	79.0	91.2	99.7	100
C_{ars}		54.0	62.1	76.5	89.5	98.3	100
	(0.05, 0.30)						
LR		66.0	73.2	85.6	98.2	100	100
C_{ml}		62.7	70.6	81.3	96.7	100	100
C_{kmm}		64.6	72.0	83.5	97.3	100	100
C_{qb}		55.1	63.5	75.2	91.4	97.8	100
C_{qs}		58.6	66.2	78.1	93.7	100	100
C_{eq}		60.5	67.3	79.2	95.7	100	100
C_{rs}		69.3	76.9	91.0	100	100	100
C_{ars}		67.5	75.1	88.4	99.3	100	100

Table 4.6: (continued)

Test Statistic	(π_1, π_2)	(θ_1, θ_2)					
	(0.05, 0.35)	(.05, .05)	(.05, .10)	(.05, .20)	(.05, .30)	(.05, .40)	(.05, .50)
LR		75.3	83.9	96.5	100	100	100
C_{ml}		73.6	81.0	93.7	100	100	100
C_{kmm}		74.5	82.7	94.9	100	100	100
C_{qb}		66.6	75.5	88.1	100	100	100
C_{qs}		70.3	77.9	90.8	100	100	100
C_{eq}		71.8	79.1	92.3	100	100	100
C_{rs}		78.2	87.9	98.9	100	100	100
C_{ars}		76.7	86.1	97.5	100	100	100
	(0.05, 0.40)						
LR		82.7	96.0	100	100	100	100
C_{ml}		78.7	92.2	100	100	100	100
C_{kmm}		80.5	94.1	100	100	100	100
C_{qb}		73.8	87.1	100	100	100	100
C_{qs}		76.8	89.4	100	100	100	100
C_{eq}		77.1	90.1	100	100	100	100
C_{rs}		85.6	99.2	100	100	100	100
C_{ars}		84.5	97.9	100	100	100	100
	(0.05, 0.45)						
LR		91.6	100	100	100	100	100
C_{ml}		88.4	100	100	100	100	100
C_{kmm}		89.7	100	100	100	100	100
C_{qb}		83.8	98.9	100	100	100	100
C_{qs}		86.1	100	100	100	100	100
C_{eq}		87.2	100	100	100	100	100
C_{rs}		94.5	100	100	100	100	100
C_{ars}		93.1	100	100	100	100	100
	(0.05, 0.50)						
LR		100	100	100	100	100	100
C_{ml}		100	100	100	100	100	100
C_{kmm}		100	100	100	100	100	100
C_{qb}		98.1	100	100	100	100	100
C_{qs}		100	100	100	100	100	100
C_{eq}		100	100	100	100	100	100
C_{rs}		100	100	100	100	100	100
C_{ars}		100	100	100	100	100	100
	(0.05, 0.60)						
LR		100	100	100	100	100	100
C_{ml}		100	100	100	100	100	100
C_{kmm}		100	100	100	100	100	100
C_{qb}		100	100	100	100	100	100
C_{qs}		100	100	100	100	100	100
C_{eq}		100	100	100	100	100	100
C_{rs}		100	100	100	100	100	100
C_{ars}		100	100	100	100	100	100

Table 4.7: Empirical power (%) of test statistics, LR, C_{ml}, C_{kmm}, C_{qb}, C_{qs}, C_{eq}, C_{rs}, C_{ars}; based on 3,000 replications and $\alpha = 0.05$ (Control and Medium Dose Groups)

Test Statistic	(π_1, π_2)	(θ_1, θ_2)					
	(0.05, 0.10)	(.10, .10)	(.10, .15)	(.10, .20)	(.10, .30)	(.10, .40)	(.10, .50)
LR		7.8	13.0	23.5	42.3	53.2	60.1
C_{ml}		7.4	11.6	22.1	39.2	50.9	56.8
C_{kmm}		7.4	12.4	23.0	40.6	52.1	58.1
C_{qb}		6.4	10.3	18.2	31.3	43.5	51.0
C_{qs}		7.0	11.1	20.8	35.7	47.5	53.9
C_{eq}		7.1	11.2	21.6	36.5	48.0	54.2
C_{rs}		8.8	13.8	25.2	44.4	55.5	63.5
C_{ars}		8.5	13.4	24.5	43.6	54.7	62.8
	(0.05, 0.15)						
LR		18.3	27.5	39.7	53.3	70.2	86.6
C_{ml}		16.2	22.8	36.2	49.1	66.1	80.8
C_{kmm}		17.1	24.5	38.4	51.2	68.2	85.4
C_{qb}		12.1	17.7	29.3	43.8	56.9	72.5
C_{qs}		13.9	20.0	32.5	46.0	60.1	75.7
C_{eq}		14.5	20.6	33.7	47.1	62.4	76.8
C_{rs}		20.3	30.6	45.6	56.9	74.7	92.4
C_{ars}		20.0	29.2	43.2	54.4	73.5	89.3
	(0.05, 0.20)						
LR		33.6	43.7	54.5	74.3	83.9	98.0
C_{ml}		31.3	39.5	50.7	70.6	80.2	93.6
C_{kmm}		32.9	42.3	52.6	73.2	82.1	95.9
C_{qb}		24.6	32.4	44.1	60.1	68.5	83.1
C_{qs}		27.4	35.2	47.2	65.9	75.3	85.8
C_{eq}		28.5	37.0	48.1	67.4	76.1	87.7
C_{rs}		37.1	46.7	57.7	77.3	85.8	100
C_{ars}		35.3	45.3	56.0	75.4	84.6	99.5
	(0.05, 0.25)						
LR		51.4	59.4	72.8	87.9	96.4	100
C_{ml}		48.1	54.9	65.3	82.3	92.0	99.0
C_{kmm}		49.5	57.5	69.5	84.6	93.6	99.8
C_{qb}		38.5	46.2	56.7	73.2	82.6	91.8
C_{qs}		42.8	50.4	59.1	76.2	86.6	96.1
C_{eq}		44.9	51.7	60.8	77.9	88.1	97.1
C_{rs}		56.6	64.0	78.4	90.9	99.5	100
C_{ars}		53.8	61.7	75.9	89.2	98.3	100
	(0.05, 0.30)						
LR		65.8	72.9	85.1	97.9	100	100
C_{ml}		62.5	70.3	81.0	96.3	100	100
C_{kmm}		64.4	71.6	83.2	97.0	100	100
C_{qb}		54.9	63.2	74.7	91.1	97.8	100
C_{qs}		58.8	66.5	77.9	93.8	100	100
C_{eq}		60.3	67.0	78.6	95.4	100	100
C_{rs}		69.1	76.6	89.7	100	100	100
C_{ars}		67.3	74.8	87.9	99.3	100	100

Table 4.7: (continued)

Test Statistic	(π_1, π_2)	(θ_1, θ_2)					
	(0.05, 0.35)	(.10, .10)	(.10, .15)	(.10, .20)	(.10, .30)	(.10, .40)	(.10, .50)
LR		75.1	83.6	95.0	100	100	100
C_{ml}		73.4	80.7	93.4	100	100	100
C_{kmm}		74.3	82.4	94.4	100	100	100
C_{qb}		66.4	75.2	87.8	100	100	100
C_{qs}		70.4	77.7	90.5	100	100	100
C_{eq}		71.6	78.8	91.7	100	100	100
C_{rs}		78.0	87.6	98.5	100	100	100
C_{ars}		76.5	85.7	97.0	100	100	100
	(0.05, 0.40)						
LR		82.5	95.7	100	100	100	100
C_{ml}		78.5	91.7	100	100	100	100
C_{kmm}		80.3	93.8	100	100	100	100
C_{qb}		73.6	86.8	100	100	100	100
C_{qs}		75.8	89.1	100	100	100	100
C_{eq}		76.9	89.8	100	100	100	100
C_{rs}		85.4	98.8	100	100	100	100
C_{ars}		84.3	97.4	100	100	100	100
	(0.05, 0.45)						
LR		91.3	100	100	100	100	100
C_{ml}		88.2	100	100	100	100	100
C_{kmm}		89.5	100	100	100	100	100
C_{qb}		83.6	98.5	100	100	100	100
C_{qs}		86.8	100	100	100	100	100
C_{eq}		87.0	100	100	100	100	100
C_{rs}		94.3	100	100	100	100	100
C_{ars}		92.9	100	100	100	100	100
	(0.05, 0.50)						
LR		100	100	100	100	100	100
C_{ml}		100	100	100	100	100	100
C_{kmm}		100	100	100	100	100	100
C_{qb}		100	100	100	100	100	100
C_{qs}		100	100	100	100	100	100
C_{eq}		100	100	100	100	100	100
C_{rs}		100	100	100	100	100	100
C_{ars}		100	100	100	100	100	100
	(0.05, 0.60)						
LR		100	100	100	100	100	100
C_{ml}		100	100	100	100	100	100
C_{kmm}		100	100	100	100	100	100
C_{qb}		100	100	100	100	100	100
C_{qs}		100	100	100	100	100	100
C_{eq}		100	100	100	100	100	100
C_{rs}		100	100	100	100	100	100
C_{ars}		100	100	100	100	100	100

Table 4.8: Empirical power (%) of test statistics, LR, C_{ml}, C_{kmm}, C_{qb}, C_{qs}, C_{eq}, C_{rs}, C_{ars}; based on $3,000$ replications and $\alpha = 0.05$ (Control and Medium Dose Groups)

Test Statistic	(π_1, π_2)	(θ_1, θ_2)					
	(0.05, 0.10)	(.20, .20)	(.20, .30)	(.20, .40)	(.20, .50)	(.20, .60)	(.20, .70)
LR		8.3	13.7	24.6	43.2	54.0	61.0
C_{ml}		7.9	12.3	23.2	40.1	51.7	57.7
C_{kmm}		7.8	13.1	24.1	41.5	53.0	58.9
C_{qb}		7.0	11.0	19.3	32.2	44.3	51.6
C_{qs}		7.4	11.8	21.7	36.6	47.8	54.9
C_{eq}		7.6	12.0	22.8	37.5	48.7	55.1
C_{rs}		9.3	14.5	26.3	45.3	56.3	64.1
C_{ars}		9.0	14.1	25.6	44.6	55.5	63.4
	(0.05, 0.15)						
LR		18.8	28.2	40.8	54.2	70.9	87.6
C_{ml}		17.7	23.5	37.3	50.0	66.8	81.5
C_{kmm}		17.7	25.2	39.5	52.3	69.1	86.0
C_{qb}		12.6	18.4	30.5	44.7	57.6	73.4
C_{qs}		14.3	20.5	33.6	47.1	61.8	76.7
C_{eq}		15.0	21.3	34.8	48.0	63.1	77.4
C_{rs}		20.8	31.4	46.6	57.8	75.4	93.4
C_{ars}		20.4	29.9	44.9	55.3	74.4	90.0
	(0.05, 0.20)						
LR		34.1	44.4	55.6	75.2	84.6	98.9
C_{ml}		31.8	40.2	51.8	71.5	80.9	94.2
C_{kmm}		33.4	43.0	53.7	74.1	82.8	96.8
C_{qb}		25.1	33.1	45.2	61.0	69.3	83.9
C_{qs}		27.8	36.1	48.2	66.4	75.7	87.5
C_{eq}		29.0	37.7	49.8	68.3	76.8	88.4
C_{rs}		37.6	47.4	58.3	78.3	86.6	100
C_{ars}		35.8	46.0	57.1	76.1	85.3	99.9
	(0.05, 0.25)						
LR		51.9	60.1	73.8	88.8	97.1	100
C_{ml}		48.6	55.6	65.9	83.2	92.7	99.9
C_{kmm}		50.0	58.2	70.6	85.5	94.4	100
C_{qb}		39.1	46.9	57.8	74.1	83.3	92.6
C_{qs}		42.8	50.6	60.2	77.4	86.6	96.8
C_{eq}		45.5	52.4	61.9	78.8	88.8	97.9
C_{rs}		57.1	64.7	79.5	91.8	100	100
C_{ars}		54.3	62.4	77.1	90.1	99.3	100
	(0.05, 0.30)						
LR		66.3	73.6	86.2	98.9	100	100
C_{ml}		63.0	71.0	82.1	97.2	100	100
C_{kmm}		64.9	72.3	84.3	97.9	100	100
C_{qb}		55.5	63.9	75.9	92.0	98.5	100
C_{qs}		58.7	66.9	78.6	94.8	100	100
C_{eq}		60.8	67.7	79.7	96.3	100	100
C_{rs}		69.6	77.3	90.8	100	100	100
C_{ars}		67.8	75.5	89.0	99.9	100	100

Table 4.8: (continued)

Test Statistic	(π_1, π_2)	(θ_1, θ_2)					
	(0.05, 0.35)	(.20, .20)	(.20, .30)	(.20, .40)	(.20, .50)	(.20, .60)	(.20, .70)
LR		75.6	84.3	97.1	100	100	100
C_{ml}		73.9	81.4	94.5	100	100	100
C_{kmm}		74.9	83.1	95.4	100	100	100
C_{qb}		66.9	75.9	88.9	100	100	100
C_{qs}		70.6	78.7	91.4	100	100	100
C_{eq}		72.1	79.5	92.8	100	100	100
C_{rs}		78.5	88.3	99.3	100	100	100
C_{ars}		77.0	86.4	98.0	100	100	100
	(0.05, 0.40)						
LR		83.1	96.4	100	100	100	100
C_{ml}		79.0	92.4	100	100	100	100
C_{kmm}		80.8	94.5	100	100	100	100
C_{qb}		74.1	87.6	100	100	100	100
C_{qs}		76.7	89.8	100	100	100	100
C_{eq}		77.4	90.5	100	100	100	100
C_{rs}		85.9	99.7	100	100	100	100
C_{ars}		84.8	98.2	100	100	100	100
	(0.05, 0.45)						
LR		91.8	100	100	100	100	100
C_{ml}		88.7	100	100	100	100	100
C_{kmm}		90.0	100	100	100	100	100
C_{qb}		84.1	99.1	100	100	100	100
C_{qs}		86.7	100	100	100	100	100
C_{eq}		87.5	100	100	100	100	100
C_{rs}		94.8	100	100	100	100	100
C_{ars}		93.4	100	100	100	100	100
	(0.05, 0.50)						
LR		100	100	100	100	100	100
C_{ml}		100	100	100	100	100	100
C_{kmm}		100	100	100	100	100	100
C_{qb}		98.1	100	100	100	100	100
C_{qs}		100	100	100	100	100	100
C_{eq}		100	100	100	100	100	100
C_{rs}		100	100	100	100	100	100
C_{ars}		100	100	100	100	100	100
	(0.05, 0.60)						
LR		100	100	100	100	100	100
C_{ml}		100	100	100	100	100	100
C_{kmm}		100	100	100	100	100	100
C_{qb}		100	100	100	100	100	100
C_{qs}		100	100	100	100	100	100
C_{eq}		100	100	100	100	100	100
C_{rs}		100	100	100	100	100	100
C_{ars}		100	100	100	100	100	100

Table 4.9: Data from Toxicological experiment (Paul (1982)). (i) Number of live foetuses affected by treatment. (ii) Total number of live foetuses.

Dose Group																												
Control (C)	(i)	1	1	4	0	0	0	0	0	1	0	2	0	5	2	1	2	0	0	1	0	0	0	0	3	2	4	0
	(ii)	12	7	6	6	7	8	10	7	8	6	11	7	8	9	2	7	9	7	11	10	4	8	10	12	8	7	1
Low (L)	(i)	0	1	1	0	2	0	1	0	1	0	0	3	0	0	1	5	0	0	3								
	(ii)	5	11	7	9	12	8	6	7	6	4	6	9	6	7	5	9	1	6	9								
Medium (M)	(i)	2	3	2	1	2	3	0	4	0	0	4	0	0	6	6	5	4	1	0	3	6						
	(ii)	4	4	9	8	9	7	8	9	6	4	6	7	3	13	6	8	11	7	6	10	6						
High (H)	(i)	1	0	1	0	1	0	1	1	2	0	4	1	1	4	2	3	1										
	(ii)	9	10	7	5	4	6	3	8	5	4	4	5	3	8	6	8	6										

Table 4.10: Estimates of parameters obtained by different methods for treatment combinations of toxicological data in Table 4.9.

	Treatment Combinations					
	CL	CM	CH	LM	LH	MH
$\hat{\pi}_0$	0.1418	0.2354	0.1956	0.2377	0.1820	0.3165
$\hat{\theta}_{10}$	0.2048	0.3164	0.2637	0.2344	0.1488	0.3079
$\hat{\theta}_{20}$	0.1122	0.3081	0.0993	0.3077	0.098	0.1723
$\hat{\pi}_{1a}$	0.1442	0.1442	0.1442	0.1272	0.1272	0.3505
$\hat{\pi}_{2a}$	0.1272	0.3505	0.2387	0.3505	0.2387	0.2387
$\hat{\theta}_{1a}$	0.2069	0.2069	0.2069	0.1054	0.1054	0.3155
$\hat{\theta}_{2a}$	0.1054	0.3155	0.1132	0.3155	0.1132	0.1132
$\hat{\pi}_{kmm}$	0.1367	0.1858	0.1741	0.1704	0.1611	0.2824
$\hat{\theta}_{1kmm}$	0.2497	0.1684	0.1792	0.0602	0.0644	0.4369
$\hat{\theta}_{2kmm}$	0.0862	0.8390	0.3247	0.9587	0.3842	0.1369
$\hat{\pi}_{qbmm}$	0.1364	0.1856	0.1739	0.1698	0.1602	0.2835
$\hat{\theta}_{1qbmm}$	0.2679	0.1822	0.1937	0.0728	0.0775	0.4683
$\hat{\theta}_{2qbmm}$	0.1001	0.8957	0.3656	1.0264	0.4330	0.1614
$\hat{\pi}_{qsmm}$	0.1369	0.1931	0.1751	0.1896	0.1701	0.2889
$\hat{\theta}_{1qsmm}$	0.1637	0.1177	0.1231	0.0651	0.0746	0.2360
$\hat{\theta}_{2qsmm}$	0.1162	0.5016	0.1726	0.5187	0.1836	0.1009
$\hat{\pi}_{eql}$	0.1362	0.2051	0.1814	0.1959	0.1728	0.2873
$\hat{\theta}_{1eql}$	0.1987	0.2334	0.2139	0.1413	0.1247	0.4437
$\hat{\theta}_{2eql}$	0.1153	0.5602	0.1866	0.5813	0.1947	0.1948

Table 4.11: Test statistics and p-values for treatment combinations of toxicological data in Table 4.9.

	Treatment Combinations					
	CL	CM	CH	LM	LH	MH
LR	0.1461 (0.7023)	7.0252 (0.0080)	2.0858 (0.1487)	8.2770 (0.0040)	2.9201 (0.0875)	1.8422 (0.1747)
C_{ml}	0.0079 (0.9290)	6.4643 (0.0110)	2.4972 (0.1141)	7.6209 (0.0058)	2.9242 (0.0873)	0.5198 (0.4709)
C_{kmm}	0.0902 (0.7639)	6.7148 (0.0096)	2.3268 (0.1272)	8.7917 (0.0030)	3.2022 (0.0735)	1.4289 (0.2319)
C_{qb}	0.1022 (0.7492)	4.8936 (0.0270)	1.7639 (0.1841)	5.5332 (0.0187)	2.3777 (0.1231)	1.2462 (0.2643)
C_{qs}	0.1064 (0.7443)	5.7404 (0.0166)	2.3006 (0.1293)	7.7773 (0.0053)	2.8571 (0.0910)	1.8122 (0.1782)
C_{eq}	0.1074 (0.7432)	5.7884 (0.0161)	1.9502 (0.1626)	6.7972 (0.0091)	2.7264 (0.0987)	1.1893 (0.2755)
C_{rs}	0.0055 (0.9409)	8.9301 (0.0028)	1.9382 (0.1639)	7.9644 (0.0048)	1.8599 (0.1726)	2.0157 (0.1557)
C_{ars}	0.0055 (0.9408)	8.4698 (0.0036)	1.8153 (0.1779)	8.2026 (0.0042)	1.8305 (0.1761)	2.0620 (0.1510)

Chapter 5

Testing Equality of Scale Parameters of Two Weibull Distributions in the Presence of Unequal Shape Parameters

5.1 Introduction

Weibull distribution has long history in describing real phenomena since its initiation by the Swedish physicist Waloddi Weibull and is one of the most popular parametric distributions in survival analysis. This distribution has been considered as an appropriate model in reliability studies and life-testing experiments and thus has versatile use in the fields such as engineering, manufacturing, aeronautics and bio-medical sciences among others. Of all the parametric models this distribution has the unique feature that in addition to being proportional it is simultaneously an accelerated failure-time model (AFT) (Carroll (2003). Many authors illustrate the use and applications of Weibull models in numerous areas of Statistics. For example, Lloyd (1967); Ku et al. (1972); McCool (1998) and Hammitt (2004) focus on the applications of this distribution in the fields such as reliability, risks and quality control. Cohen (1965); Sirvanci and Yang (1984) discuss the maximum likelihood estimation procedure of the parameters under complete and various censoring samples, Harter and Moore (1965) deal with the joint maximum-likelihood estimation from complete and censored samples, Cohen et al. (1984), and Cran (1988) consider the moments estimation of Weibull parameters. In a recent article, Teimouri and Gupta (2013) propose a consistent and closed form estimator for shape parameter which is independent of the scale parameter.

Very often lifetime or survival time data that are collected in the form of two independent samples are assumed to have come from two Weibull distributions. In such situation it may be of interest to test the equality of scale parameters of two Weibull distributions which eventually is equivalent to testing the equality of reliability at a certain time. Lawless (1982) shows that the ratio of the p^{th} quantile of two Weibull distributions is the same as the ratio of their scale parameters when the shape parameters are equal and proposes a test statistic based on this property for testing equality of two Weibull scale parameters when the shape parameters are equal. Thoman and Bain (1969); Schafer and Sheffield (1976) present statistics for testing the equality of scale parameters of two Weibull distributions when the shape parameters are equal. These statistics are based on the maximum likelihood estimates of the parameters. Thoman and Bain (1969) mention that testing equality of two scale parameters where the shape parameters are assumed to be equal is equivalent to testing the equality of two Weibull means. But this is not the case when the shape parameters are not assumed to be equal. McCool (1979, 1982) proposes a test procedure which is based on the ratio of the maximum likelihood estimators of the shape parameters. In an unpublished thesis, Thiagarajah (1992) (Department of Mathematics and Statistics, University of Windsor), derives a $C(\alpha)$ statistic and compares it with the statistics proposed by Lawless (1982) and McCool (1979, 1980).

The assumption that the shape parameters are equal is not always satisfied in practice while testing the equality of two Weibull scale parameters. Also we observe that the available test procedures are based on the maximum likelihood estimates of the parameters. Apart from the maximum likelihood estimates of the parameters several methods of moments estimators are proposed by different authors. Our objective, in this Chapter, is to develop test procedures to test the equality of scale parameters of two Weibull distributions where the shape parameters are assumed unequal and unknown and compare the performance of these test procedures. We compare the performance, through simulation studies, in terms of empirical size and power of the test procedures. The test procedures are developed in section 5.2, simulation studies are presented in section 5.3 and illustrative examples and discussion have been given in section 5.4.

Let the random variable Y follow a two parameter Weibull distribution with shape parameter β and scale parameter α. Then the probability density function of Y can be written as

$$f(y) = \frac{\beta}{\alpha}\left(\frac{y}{\alpha}\right)^{(\beta-1)} exp\left[-\left(\frac{y}{\alpha}\right)^{\beta}\right];\;\; y \geq 0;\; \beta,\; \alpha > 0. \tag{5.1}$$

We denote this distribution as $Y \sim Weibull(\beta,\ \alpha)$. The shape parameter β possesses the role to determine how the curve of the Weibull density looks.

5.2 The Tests

Let $y_{11}, y_{12}, \cdots, y_{1n_1}$ be a random sample to size n_1 from a two parameter Weibull distribution with shape parameter β_1 and scale parameter α_1 and $y_{21}, y_{22}, \cdots, y_{2n_2}$ be another independent random sample of size n_2 from the same distribution having shape parameter β_2 and scale parameter α_2. Our objective is to test the null hypothesis H_0: $\alpha_1 = \alpha_2$, where β_1 and β_2 are unknown and unequal. We develop the following test procedures to test this hypothesis (i) a likelihood ratio test, (ii) a $C(\alpha)$ test based on the maximum likelihood estimates of the nuisance parameters, (iii) a $C(\alpha)$ test based on the method of moments estimates of the nuisance parameters by Cran (1988) and (iv) a $C(\alpha)$ test based on the method of moments estimates of the nuisance parameters by Teimouri and Gupta (2013).

5.2.1 The Likelihood Ratio Test

The log-likelihood function for the data from two samples can be written as

$$l_1 = \sum_{i=1}^{2} \left[n_i \log\left(\frac{\beta_i}{\alpha_i}\right) + (\beta_i - 1)\left\{\sum_{j=1}^{n_i} \log(y_{ij}) - n_i \log(\alpha_i)\right\} - \frac{\sum_{j=1}^{n_i} y_{ij}^{\beta_i}}{\alpha_i^{\beta_i}} \right].$$

The unrestricted maximum likelihood estimates of the parameters α_i and β_i; $i = 1, 2$ can be obtained by solving the following equations

$$-\frac{n_i\beta_i}{\alpha_i}+\frac{\beta_i}{\alpha_i^{\beta_i+1}}\sum_{j=1}^{n_i}y_{ij}^{\beta_i}=0$$

and

$$\frac{n_i}{\beta_i}+\sum_{j=1}^{n_i}\log\left(y_{ij}\right)-n_i\log\left(\alpha_i\right)+\frac{\log(\alpha_i)}{\alpha_i^{\beta i}}\sum_{j=1}^{n_i}y_{ij}^{\beta_i}-\frac{1}{\alpha_i^{\beta_i}}\sum_{j=1}^{n_i}y_{ij}^{\beta_i}\log\left(y_{ij}\right);\ i=1,2,$$

simultaneously.

Under the null hypothesis $\alpha_1=\alpha_2=\alpha$ the log-likelihood function is

$$l_0=\sum_{i=1}^{2}\left[n_i\log\left(\frac{\beta_i}{\alpha}\right)+(\beta_i-1)\left\{\sum_{j=1}^{n_i}\log\left(y_{ij}\right)-n_i\log\left(\alpha\right)\right\}-\frac{\sum_{j=1}^{n_i}y_{ij}^{\beta_i}}{\alpha^{\beta_i}}\right].$$

The maximum likelihood estimates of the parameters α,β_1 and β_2 are obtained from the solution of the following equations

$$\sum_{i=1}^{2}\left[-\frac{n_i\beta_i}{\alpha}+\frac{\beta_i}{\alpha^{\beta_i+1}}\sum_{j=1}^{n_i}y_{ij}^{\beta_i}\right]=0,$$

$$\frac{n_1}{\beta_1}+\sum_{j=1}^{n_1}\log\left(y_{1j}\right)-n_1\log\left(\alpha\right)+\frac{\log\left(\alpha\right)}{\alpha^{\beta_1}}\sum_{j=1}^{n_1}y_{1j}^{\beta_1}-\frac{1}{\alpha^{\beta_1}}\sum_{j=1}^{n_1}y_{1j}^{\beta_1}\log\left(y_{1j}\right)=0$$

and

$$\frac{n_2}{\beta_2}+\sum_{j=1}^{n_2}\log\left(y_{2j}\right)-n_2\log\left(\alpha\right)+\frac{\log\left(\alpha\right)}{\alpha^{\beta_2}}\sum_{j=1}^{n_2}y_{2j}^{\beta_2}-\frac{1}{\alpha^{\beta_2}}\sum_{j=1}^{n_2}y_{2j}^{\beta_2}\log\left(y_{2j}\right)=0$$

simultaneously.

Now, let $\hat{l}_1$ be the maximized log-likelihood under the alternative hypothesis and $\hat{l}_0$

be the maximized log-likelihood under the null hypothesis. Then the likelihood ratio test statistic is $LR = 2\left(\hat{l}_1 - \hat{l}_0\right)$; which, asymptotically as $n \to \infty$, where$n = n_1 + n_2$, follows a χ^2 distribution with 1 degree of freedom.

5.2.2 The $C(\alpha)$ (Score) Test Based on Maximum Likelihood Estimates

Suppose the alternative hypothesis is represented by $\alpha_i = \alpha + \phi_i$, $i = 1, 2$, with $\phi_2 = 0$. Then the null hypothesis, $H_0 : \alpha_1 = \alpha_2$, can be written as $H_0 : \phi_i = 0$, $i = 1, 2$, where $\alpha, \beta_1,$ and β_2 are treated as nuisance parameters. The log-likelihood can then be written as

$$l = \sum_{i=1}^{2} \left[n_i \log\left(\frac{\beta_i}{\alpha + \phi_i}\right) + (\beta_i - 1)\left\{\sum_{j=1}^{n_i} \log(y_{ij}) - n_i \log(\alpha + \phi_i)\right\} - \frac{\sum_{j=1}^{n_i} y_{ij}^{\beta_i}}{(\alpha + \phi_i)^{\beta_i}} \right]. \tag{5.2}$$

Now, define $\phi = \phi_1,\ \delta = (\alpha, \beta_1, \beta_2)',\ \psi_1 = \frac{\partial l}{\partial \phi_1}\Big|_{\phi=0},\ \gamma_1 = \frac{\partial l}{\partial \delta_1}\Big|_{\phi=0},\ \gamma_2 = \frac{\partial l}{\partial \delta_2}\Big|_{\phi=0}$ and $\gamma_3 = \frac{\partial l}{\partial \delta_3}\Big|_{\phi=0}$. Let that $\hat{\delta}$ be a $\sqrt{n}$ (n is the size of the sample in a group) consistent estimator of δ. Then, the $C(\alpha)$ statistic is based on the adjusted score $S_1 = \psi_1 - \beta_1\gamma_1 - \beta_2\gamma_2 - \beta_3\gamma_3$, where β_1, β_2 and β_3 are partial regression coefficient of ψ_1 on γ_1, ψ_1 on γ_2, and ψ_1 on γ_3, respectively.

The structure of dispersion matrix of $(\phi, \alpha, \beta_1, \beta_2)$ is

$$V = \begin{pmatrix} D & A \\ A' & B \end{pmatrix}$$

and the regression coefficients $\beta = (\beta_1, \beta_2, \beta_3) = AB^{-1}$ (Neyman (1959)), where D is 1×1, A is 1×3 and B is 3×3 with elements

$$D_{11} = E\left[-\frac{\partial^2 l}{\partial \phi_1^2}\Big|_{\phi=0}\right],$$

$$A_{11} = E\left[-\frac{\partial^2 l}{\partial \phi_1 \partial \alpha}\Big|_{\phi=0}\right], \; A_{12} = E\left[-\frac{\partial^2 l}{\partial \Phi_1 \partial \beta_1}\Big|_{\phi=0}\right], \; A_{13} = E\left[-\frac{\partial^2 l}{\partial \phi_1 \partial \beta_2}\Big|_{\phi=0}\right],$$

$$B_{11} = E\left[-\frac{\partial^2 l}{\partial \alpha^2}\Big|_{\phi=0}\right], \; B_{12} = E\left[-\frac{\partial^2 l}{\partial \alpha \partial \beta_1}\Big|_{\phi=0}\right], \; B_{13} = E\left[-\frac{\partial^2 l}{\partial \alpha \partial \beta_2}\Big|_{\phi=0}\right],$$

$$B_{22} = E\left[-\frac{\partial^2 l}{\partial \beta_1^2}\Big|_{\Phi=0}\right], \; B_{23} = E\left[-\frac{\partial^2 l}{\partial \beta_1 \partial \beta_2}\Big|_{\phi=0}\right] \text{ and } B_{33} = E\left[-\frac{\partial^2 l}{\partial \beta_2^2}\Big|_{\phi=0}\right].$$

Substituting $\sqrt{n}$ consistent estimate of δ, that is, $\hat{\delta}$ in S_1, D, A, and B, the $C(\alpha)$ statistic can be obtained as $S_1'(D - AB^{-1}A')^{-1}S_1$, which is approximately distributed as the chi-squared with 1 degree of freedom (Neyman (1959); Neyman and Scott (1966); Moran (1970)). Using the above log-likelihood, that is, equation (5.2) and after simplification we obtain the above elements as follows

$$\psi_1 = -\frac{n_1\beta_1}{\alpha} + \frac{\beta_1}{\alpha^{\beta_1+1}} \sum_{j=1}^{n_1} y_{1j}^{\beta_1},$$

$$\gamma_1 = \sum_{i=1}^{2} \left[-\frac{n_i\beta_i}{\alpha} + \frac{\beta_i}{\alpha^{\beta_i+1}} \sum_{j=1}^{n_i} y_{ij}^{\beta_i}\right],$$

$$\gamma_2 = \frac{n_1}{\beta_1} + \sum_{j=1}^{n_{1j}} \log\left(y_{1j}\right) - n_1 \log\left(\alpha\right) + \frac{\log\left(\alpha\right)}{\alpha^{\beta_1}} \sum_{j=1}^{n_1} y_{1j}^{\beta_1} - \frac{1}{\alpha^{\beta_1}} \sum_{j=1}^{n_1} y_{1j}^{\beta_1} \log\left(y_{1j}\right),$$

$$\gamma_3 = \frac{n_2}{\beta_2} + \sum_{j=1}^{n_{2j}} \log\left(y_{2j}\right) - n_2 \log\left(\alpha\right) + \frac{\log\left(\alpha\right)}{\alpha^{\beta_2}} \sum_{j=1}^{n_2} y_{2j}^{\beta_2} - \frac{1}{\alpha^{\beta_2}} \sum_{j=1}^{n_2} y_{2j}^{\beta_2} \log\left(y_{2j}\right),$$

$$D_{11} = -\frac{n_1\beta_1}{\alpha^2} + \frac{n_1\beta_1\left(\beta_1+1\right)E\left(y_{1j}^{\beta_1}\right)}{\alpha^{\beta_1+2}},$$

$$A_{11} = -\frac{n_1\beta_1}{\alpha^2} + \frac{n_1\beta_1\left(\beta_1+1\right)E\left(y_{1j}^{\beta_1}\right)}{\alpha^{\beta_1+2}},$$

$$A_{12} = \frac{n_1}{\alpha} - \frac{n_1\left\{1-\beta_1\log\left(\alpha\right)\right\}E\left(y_{1j}^{\beta_1}\right)}{\alpha^{\beta_1+1}} - \frac{n_1\beta_1 E\left\{y_{1j}^{\beta_1}\log\left(y_{1j}\right)\right\}}{\alpha^{\beta_1+1}}, \quad A_{13} = 0,$$

$$B_{11} = \sum_{j=1}^{n_i}\left[-\frac{n_i\beta_i}{\alpha^2} + \frac{n_i\beta_i\left(\beta_i+1\right)E\left(y_{ij}^{\beta_i}\right)}{\alpha^{\beta_i+2}}\right],$$

$$B_{12} = \frac{n_1}{\alpha} - \frac{n_1\left\{1-\beta_1\log\left(\alpha\right)\right\}E\left(y_{1j}^{\beta_1}\right)}{\alpha^{\beta_1+1}} - \frac{n_1\beta_1 E\left\{y_{1j}^{\beta_1}\log\left(y_{1j}\right)\right\}}{\alpha^{\beta_1+1}},$$

$$B_{13} = \frac{n_2}{\alpha} - \frac{n_2\left\{1-\beta_2\log\left(\alpha\right)\right\}E\left(y_{2j}^{\beta_2}\right)}{\alpha^{\beta_2+1}} - \frac{n_2\beta_2 E\left\{y_{2j}^{\beta_2}\log\left(y_{2j}\right)\right\}}{\alpha^{\beta_2+1}},$$

$$B_{22} = \frac{n_1}{\beta_1^2} + \frac{n_1\left\{\log\left(\alpha\right)\right\}^2 E\left(y_{1j}^{\beta_1}\right) + n_1 E\left\{y_{1j}^{\beta_1}\left(\log\left(y_{1j}\right)\right)^2\right\}}{\alpha^{\beta_1}}, \quad B_{23} = 0 \text{ and}$$

$$B_{33} = \frac{n_2}{\beta_2^2} + \frac{n_2\left\{\log\left(\alpha\right)\right\}^2 E\left(y_{2j}^{\beta_2}\right) + n_2 E\left\{y_{2j}^{\beta_2}\left(\log\left(y_{2j}\right)\right)^2\right\}}{\alpha^{\beta_2}}.$$

If the maximum likelihood estimate of δ is used then $S_1 = \psi_1$, and the $C(\alpha)$ statistic reduces to score statistic (Rao (1948)). Denoting the maximum likelihood estimate of δ as $\hat{\delta}_{ml}$ and using this estimate, the $C(\alpha)$ statistic which, in this case is the score statistic (Rao (1948)), is obtained as $C_{ml} = \psi_1^2 \,/\, \left(D_{11} - AB^{-1}A'\right)$. Under the null hypothesis C_{ml} is distributed, asymptotically as $n \to \infty$, as chi-squared with 1 degree of freedom.

5.2.3 The $C(\alpha)$ Test Based on Method of Moments Estimates by Cran (1988)

Cran (1988) proposes moments estimates of the parameters for three-parameter Weibull distribution and apply this procedure for two-parameter model considering the location parameter as zero. The estimate of the shape parameter β is independent of the scale parameter α and the estimates of the parameters for a single sample are $\tilde{\beta}_c = \dfrac{ln\,(2)}{ln\,(\bar{m}_1) - ln\,(\bar{m}_2)}$ and $\tilde{\alpha}_c = \dfrac{\bar{m}_1}{\Gamma\left(1 + \frac{1}{\tilde{\beta}_c}\right)}$, where $\bar{m}_k = \sum_{r=0}^{n-1} \left(1 - \dfrac{r}{n}\right)^k \left\{y_{(r+1)} - y_{(r)}\right\}$, with $y_{(0)} = 0$ and $y_{(r)}$ is the r^{th} ordered observation. As we have two independent samples, we estimate the parameters β_i and α_i, $i = 1, 2$ using this method and then obtain the estimates of variances $Var\,(y_i) = \alpha_i^2 \left[\Gamma\left(1 + \dfrac{2}{\beta_i}\right) - \left\{\Gamma\left(1 + \dfrac{1}{\beta_i}\right)\right\}^2\right]$ by substituting the estimates. An estimate of $\tilde{\alpha}_c$, under null hypothesis of equality of the scale parameters, is then obtained by $\tilde{\alpha}_c = \dfrac{\sum_{i=1}^{2} w_i \tilde{\alpha}_i}{\sum_{i=1}^{2} w_i}$, where $w_i = \dfrac{n_i}{Var\,(y_i)}$, $i = 1, 2$. Then the $C\,(\alpha)$ statistic based on the method of moments estimates of nuisance parameters by Cran (1988) is obtained as $C_{cr} = S_1^2 \,/\, \left(D_{11} - AB^{-1}A'\right)$, where in S_1, A, B and D, the α, β_1 and β_2

are replaced by $\tilde{\alpha}_c$, $\tilde{\beta}_{1c}$ and $\tilde{\beta}_{2c}$. Under the null hypothesis C_{cr} follows, asymptotically as $n \to \infty$, where, $n = n_1 + n_2$, a chi-squared distribution with 1 degree of freedom.

5.2.4 The $C(\alpha)$ Test Based on Method of Moments Estimates by Teimouri and Gupta (2013)

In a recent article, Teimouri and Gupta (2013) propose a method of moments estimate of the shape parameter of a three-parameter Weibull distribution and without loss of generality apply this method to a two-parameter Weibull distribution for estimating the shape parameter. The special feature of this estimate is that it is independent of the scale parameter α and depends only on the coefficient of variation statistic. For a random sample $y_1, y_2, \cdots, y_n$, of size n from a two-parameter Weibull distribution with scale parameter α and shape parameter β, the Teimouri and Gupta (2013) method of moments estimate of β is $\tilde{\beta} = \dfrac{-ln2}{ln\left[1 - \frac{r}{\sqrt{3}} CV \sqrt{\frac{n+1}{n-1}}\right]}$, where r denotes the sample correlation between y_i and their ranks and CV is the sample coefficient of variation. Our objective here is to estimate α, β_1 and β_2 from two independent samples of Weibull distributions. Following Teimouri and Gupta (2013) we obtain the estimates of β_1 and β_2 which are independent of α and then obtain the estimate of α_i by equating $E(Y_i)$ to $\bar{y}_i$, $i = 1, 2$ and then take a weighted average of α_is, as in the previous section, in order to obtain the common estimate of α under the null hypothesis. We denote the estimate of δ so obtained by $\tilde{\delta}_{tg}$. Then the $C(\alpha)$ statistic based on the method of moments estimates of nuisance parameters by Teimouri and Gupta (2013) is obtained as $C_{tg} = S_1^2 / \left(D_{11} - AB^{-1}A'\right)$. The distribution of C_{tg} is also asymptotically $\chi^2(1)$ when $n \to \infty$, where $n = n_1 + n_2$.

5.3 Simulation Studies

We conduct a simulation study to compare the performance of the test procedures, namely, LR, C_{ml}, C_{cr} and C_{tg} that were developed in section 5.2. The performance of the test procedures are compared on the basis of empirical level and power. To compare the statistics in terms of empirical level we consider sample sizes $n_1 = n_2 = 5,\ 10,\ 20,\ 50$, values of scale parameters $\alpha_1 = \alpha_2 = 5,\ 8,\ 10,\ 12$ and the combinations of the values of shape parameters $(\beta_1,\ \beta_2) = (3,\ 3),\ (3,\ 3.5),\ (3,\ 4),\ (3,\ 4.5),\ (3,\ 5),\ (3,\ 5.5),\ (3,\ 6)$ and nominal levels $\alpha = 0.05,\ 0.10$. In the cases of $\alpha_1 = \alpha_2 = 5$, $\alpha_1 = \alpha_2 = 8$, $\alpha_1 = \alpha_2 = 10$ and $\alpha_1 = \alpha_2 = 12$ the comparative results in terms of empirical size are similar. So, we present results of empirical sizes for only $\alpha_1 = \alpha_2 = 10$. In order to compare powers we consider same sample size and same combination of $(\beta_1,\ \beta_2)$ with the combinations of the scale parameters $(\alpha_1,\ \alpha_2) = (5,\ 6),\ (5,\ 7),\ (5,\ 8),\ (5,\ 9),\ (5,\ 10),\ (5,\ 12)$, $(\alpha_1,\ \alpha_2) = (8,\ 9),\ (8,\ 10),\ (8,\ 11),\ (8,\ 12),\ (8,\ 13),\ (8,\ 15)$, $(\alpha_1,\ \alpha_2) = (10,\ 11),\ (10,\ 12),\ (10,\ 13),\ (10,\ 14),\ (10,\ 15),\ (10,\ 17)$ and $(\alpha_1,\ \alpha_2) = (12,\ 13),\ (12,\ 14),\ (12,\ 15),\ (12,\ 16),\ (12,\ 17),\ (12,\ 18)$. The comparative results in power comparisons are also similar for all combinations of the scale parameters. So, as representative results, we present results for nominal level $\alpha = 0.05$ and for $(\alpha_1,\ \alpha_2) = (10,\ 11),\ (10,\ 12),\ (10,\ 13),\ (10,\ 14),\ (10,\ 15),\ (10,\ 17)$. Results for level and power are summarized in Tables 5.1 to 5.4 and we observe the following features of the test statistics in maintaining empirical level and power

a) In case of small sample sizes, that is, for $n_1 = n_2 = 5,\ 10$ the statistics C_{cr} and C_{tg} show liberal behaviour in maintaining nominal level for all combinations of β_1 and β_2. The statistics C_{ml} and LR hold level effectively for no difference to small difference of the shape parameters but for moderate to large differences of β_1 and β_2 these statistics show somewhat liberal behaviour. For example, in Table 5.1, the empirical level of C_{cr} is 6.2% for $(\beta_1,\ \beta_2) = (3,\ 3)$ and that of C_{tg} is 6.8% for $(\beta_1,\ \beta_2) = (3,\ 6)$ and in Table 5.2, the empirical level of C_{ml} is 4.7% for $(\beta_1,\ \beta_2) = (3,\ 4)$. The statistic C_{cr} appears with the largest power for small sample sizes and for all combinations of $(\beta_1,\ \beta_2)$, followed by C_{tg}, LR and C_{ml}. The power exhibits an increasing trend with the increase in the departure between $(\beta_1,\ \beta_2)$ within the same combination of $(\alpha_1,\ \alpha_2)$. For instance, for $(\alpha_1,\ \alpha_2) = (10,\ 13)$, $(\beta_1,\ \beta_2) = (3,\ 4)$, the power of C_{cr} is 30.5% and that of C_{ml} is 25.5%.

b) The empirical levels of all of the four statistics are close to the nominal level for the small departures between the shape parameters, though, these statistics demonstrate somewhat liberal behaviour for large departures between β_1 and β_2 in the moderate sample sizes situations, that is, when $n_1 = n_2 = 20$. For example, it can be seen from column 4 of Table 5.3 that the empirical levels of LR, C_{ml}, C_{cr} and C_{tg} are 4.6%, 4.8%, 4.8% and 4.9% respectively. The test procedures based on the maximum likelihood estimates of the parameters, that is, LR and C_{ml} demonstrate larger power, with power of LR being the largest, than the statistics based on method of moments estimates, that is, C_{cr} and C_{tg}, with power of C_{tg} being the smallest, for all values of $(\alpha_1,\ \alpha_2)$ and all combinations of $(\beta_1,\ \beta_2)$. For example, for $(\alpha_1,\ \alpha_2) = (10,\ 14)$ and $(\beta_1,\ \beta_2) = (3,\ 5)$, the powers of LR, C_{ml}, C_{cr} and C_{tg} are 80.9%, 77.3%, 74.1% and 70.6%.

c) For large sample sizes, that is, for $n_1 = n_2 = 50$, the empirical levels of all statistics exhibit more closeness to nominal level than other sample size situations, though, for large differences in the shape parameters these statistics are a bit liberal. Again, like the moderate sample size situations, the statistics based on the maximum likelihood estimates of parameters appear with higher power than test procedures based on the method of moments estimates.

For each combination of sample sizes (n_1, n_2) and the specified values of scale (α_1, α_2) and shape (β_1, β_2) parameters, samples were generated from the Weibull distribution as per equation (5.1) using the built in function "*rweibull*" in R. Also the simultaneous solution of the equations were obtained through the function "*nleqslv*" in R. The empirical significance levels and powers of the test procedures were obtained from 3000 valid simulation runs after discarding the non-convergent samples and samples that produced out-of-range estimates of the parameters. Cohen and Whitten (1982) mention that the usual asymptotic properties of maximum likelihood do not hold unless $\beta > 2$. So, we discarded samples for which the value of maximum likelihood estimate of either β_1 or β_2 were obtained to be less than 2.5. We observed that though convergence in finding estimates of parameters was obtained, the values of the test statistics were unusual when the estimates of parameters were very high. In order to avoid this problem, we discarded samples for which the estimates of parameters were found to be more than three times the specified values of the parameters.

5.4 Examples

Lawless (1982) presents a set of data (originally given by McCool (1979)) that represents the times to fatigue failure in units of millions of cycles of 10 high-speed turbine engine bearings made out of five different compounds. The data are given in Table 5.5. We conduct a pairwise comparison of the five different compound types. The maximum likelihood estimates of parameters, under both alternative and null hypotheses, and the methods of moments estimates are presented in Table 5.6 and the values of the test statistics along with the corresponding p-values are given in Table

5.7. In all combinations of pairs of compound types, the value of C_{cr} is the largest with the smallest p-value which is followed by C_{tg}, LR and C_{ml} in decreasing order. In particular, if we focus on the results of the comparison of compound types I and V then we see that the values of the test statistics LR, C_{ml}, C_{cr} and C_{tg} with the p-values within the parentheses are $3.4073\,(0.0649)$, $2.6298\,(0.1049)$, $5.1513\,(0.0232)$ and $4.5956\,(0.0321)$ respectively. This is in agreement with the simulation results presented in Table 5.2 where we see that for $n_1 = n_2 = 10$ the statistic C_{cr} demonstrates the largest power which is followed by C_{tg}, LR and C_{ml}.

5.5 Discussion

In this chapter we dealt with survival data that follow Weibull distribution and we developed four test procedures to test the equality of scale parameters of two Weibull distributions where the shape parameters are assumed unknown and unequal. The test procedures we developed are, a likelihood ratio statistic LR, a $C(\alpha)$ (score) statistic based on maximum likelihood estimates of the nuisance parameters C_{ml}, a $C(\alpha)$ statistic based on method of moments estimates of the nuisance parameters by Cran (1988) C_{cr} and a $C(\alpha)$ statistic based on method of moments estimates of the nuisance parameters by Teimouri and Gupta (2013) C_{tg}. A comparative study, through Monte-Carlo simulation, was conducted to observe the performance of the test procedures. Empirical significance level and power were considered as the tools for measuring performance. In general, for small sample sizes the statistics C_{cr} and C_{tg} were found liberal where as the statistics C_{ml} and LR were found to hold nominal level effectively. The statistic C_{cr} exhibits largest power followed by C_{tg}, LR and C_{ml}. All four statistics hold reasonably well empirical level in case of moderate and large sample sizes situations. In large sample situations the likelihood ratio statistic appears with the largest power in all parameter combinations followed by C_{ml}, C_{cr} and C_{tg}.

Table 5.1: Empirical level and power (%) of test statistics LR, C_{ml}, C_{cr} and C_{tg}; based on 3000 iterations and $n_1 = 5, n_2 = 5, \alpha = 0.05$

		(β_1, β_2)						
(α_1, α_2)	Statistics	$(3, 3)$	$(3, 3.5)$	$(3, 4)$	$(3, 4.5)$	$(3, 5)$	$(3, 5.5)$	$(3, 6)$
(10, 10)	LR	4.8	5.3	5.6	6.0	6.2	6.0	6.1
	C_{ml}	4.4	4.8	5.1	5.6	5.8	5.7	5.8
	C_{cr}	6.2	6.5	6.8	7.2	7.3	7.1	7.5
	C_{tg}	5.5	5.9	6.3	6.7	6.7	6.4	6.8
(10, 11)	LR	6.4	7.0	7.4	7.9	8.6	9.1	9.8
	C_{ml}	6.0	6.5	6.9	7.4	8.0	8.5	9.2
	C_{cr}	7.6	8.4	8.8	9.5	10.1	10.6	11.0
	C_{tg}	6.9	7.6	8.0	8.6	9.3	9.8	10.3
(10, 12)	LR	10.8	11.8	12.5	13.4	14.5	15.6	16.3
	C_{ml}	10.2	11.0	11.7	12.6	13.5	14.5	15.7
	C_{cr}	12.4	13.8	14.5	15.6	16.6	17.6	18.4
	C_{tg}	11.5	12.7	13.4	14.4	15.4	16.4	17.4
(10, 13)	LR	16.4	17.7	18.9	20.1	21.6	23.1	24.1
	C_{ml}	15.5	16.7	17.7	18.9	20.2	21.5	23.0
	C_{cr}	18.6	20.4	22.5	23.0	24.4	25.7	26.9
	C_{tg}	17.4	18.9	20.1	21.4	22.8	24.3	25.7
(10, 14)	LR	23.9	25.6	27.3	28.9	30.8	32.7	34.1
	C_{ml}	22.7	24.2	25.6	27.1	28.9	30.5	32.6
	C_{cr}	26.9	29.1	31.8	32.6	34.4	36.2	37.6
	C_{tg}	25.3	27.1	29.0	30.6	32.4	34.3	36.0
(10, 15)	LR	34.6	36.9	39.2	41.5	44.0	46.8	48.6
	C_{ml}	33.0	34.9	36.7	38.9	41.4	43.7	46.5
	C_{cr}	38.7	41.5	44.7	46.5	49.2	51.7	53.8
	C_{tg}	36.6	38.9	41.4	43.8	46.4	49.1	51.3
(10, 17)	LR	53.6	57.0	60.4	63.1	66.0	69.2	71.3
	C_{ml}	51.4	53.9	56.3	59.0	61.9	64.6	67.7
	C_{cr}	59.1	63.5	67.3	69.5	72.5	75.4	77.8
	C_{tg}	56.3	60.1	63.7	66.6	69.5	72.6	75.1

Table 5.2: Empirical level and power (%) of test statistics LR, C_{ml}, C_{cr} and C_{tg}; based on 3000 iterations and $n_1 = 10, n_2 = 10, \alpha = 0.05$

		(β_1, β_2)						
(α_1, α_2)	Statistics	$(3, 3)$	$(3, 3.5)$	$(3, 4)$	$(3, 4.5)$	$(3, 5)$	$(3, 5.5)$	$(3, 6)$
(10, 10)	LR	4.6	5.0	5.4	5.7	5.8	6.1	6.4
	C_{ml}	4.1	4.4	4.7	5.1	5.5	5.5	6.0
	C_{cr}	5.9	6.2	6.4	6.5	6.6	7.0	7.4
	C_{tg}	5.4	5.7	6.0	6.2	6.2	6.5	6.9
(10, 11)	LR	11.9	13.5	14.8	15.7	16.5	17.3	18.4
	C_{ml}	11.2	12.6	13.8	14.7	15.6	16.4	17.4
	C_{cr}	13.9	15.9	16.2	17.8	18.8	19.6	20.4
	C_{tg}	12.8	14.6	15.8	16.8	17.1	18.5	19.3
(10, 12)	LR	17.6	19.9	21.8	23.9	26.2	28.3	30.3
	C_{ml}	16.6	18.6	20.4	22.4	24.6	26.6	29.0
	C_{cr}	19.8	22.7	24.6	27.0	29.2	31.2	33.3
	C_{tg}	18.6	21.2	23.0	25.2	27.5	29.5	31.7
(10, 13)	LR	24.2	27.0	29.4	31.9	34.5	37.0	39.3
	C_{ml}	22.9	25.5	27.7	29.9	32.6	34.9	37.6
	C_{cr}	27.0	30.5	33.8	35.7	38.3	40.6	42.9
	C_{tg}	25.5	28.5	31.0	33.4	36.3	38.7	41.5
(10, 14	LR	33.0	36.1	39.1	41.9	44.9	47.9	50.8
	C_{ml}	31.4	34.2	37.0	39.4	42.6	45.1	48.6
	C_{cr}	36.6	40.4	44.4	46.5	49.6	52.4	55.0
	C_{tg}	34.7	37.9	41.1	43.8	47.2	49.9	53.2
(10, 15)	LR	46.1	49.7	53.4	57.0	61.0	64.6	67.9
	C_{ml}	44.1	47.3	50.6	53.8	57.9	60.8	65.3
	C_{cr}	50.7	55.4	59.9	63.0	67.5	70.5	74.0
	C_{tg}	48.4	52.2	56.1	59.7	64.2	67.4	71.2
(10, 17)	LR	84.3	89.4	98.0	99.9	100	100	100
	C_{ml}	82.6	86.4	94.1	99.8	100	100	100
	C_{cr}	91.2	97.1	100	100	100	100	100
	C_{tg}	87.9	91.9	99.9	100	100	100	100

Table 5.3: Empirical level and power (%) of test statistics LR, C_{ml}, C_{cr} and C_{tg}; based on 3000 iterations and $n_1 = 20, n_2 = 20, \alpha = 0.05$

		(β_1, β_2)						
(α_1, α_2)	Statistics	$(3, 3)$	$(3, 3.5)$	$(3, 4)$	$(3, 4.5)$	$(3, 5)$	$(3, 5.5)$	$(3, 6)$
(10, 10)	LR	4.5	4.6	4.6	4.8	5.1	5.2	5.4
	C_{ml}	4.6	4.8	4.8	5.2	5.4	5.4	5.7
	C_{cr}	4.8	4.8	5.1	5.4	5.7	5.8	5.9
	C_{tg}	4.8	4.9	5.1	5.5	5.7	5.9	6.1
(10, 11)	LR	21.7	25.6	27.3	30.3	32.8	33.1	33.8
	C_{ml}	20.1	23.6	26.4	29.0	30.4	31.0	32.1
	C_{cr}	18.7	21.6	24.7	27.2	28.6	29.1	30.6
	C_{tg}	17.6	20.1	22.8	25.1	27.3	27.5	29.1
(10, 12)	LR	32.4	37.7	41.7	45.1	49.1	52.0	55.6
	C_{ml}	30.7	35.6	39.4	42.5	46.7	49.2	53.2
	C_{cr}	28.9	33.6	37.5	40.7	44.5	47.2	50.7
	C_{tg}	27.1	31.5	35.3	38.5	42.2	44.8	48.5
(10, 13)	LR	46.2	50.9	56.6	60.5	63.5	66.8	71.0
	C_{ml}	43.6	47.9	52.5	56.9	60.9	63.7	68.2
	C_{cr}	41.4	45.6	50.2	54.7	57.8	60.9	65.2
	C_{tg}	39.0	43.4	47.2	51.6	54.7	57.7	62.3
(10, 14)	LR	62.7	70.4	76.8	77.9	80.9	84.4	88.6
	C_{ml}	59.7	65.8	72.3	74.1	77.3	81.2	86.2
	C_{cr}	56.3	62.5	69.2	71.1	74.1	78.0	82.9
	C_{tg}	53.4	60.0	65.9	67.5	70.6	74.2	79.6
(10, 15)	LR	75.8	81.6	87.4	92.1	97.9	100	100
	C_{ml}	72.7	77.3	83.0	87.7	93.4	97.7	100
	C_{cr}	69.3	74.1	79.5	84.1	89.1	94.1	98.5
	C_{tg}	66.1	70.8	76.0	80.2	85.0	89.5	95.2
(10, 17)	LR	100	100	100	100	100	100	100
	C_{ml}	100	100	100	100	100	100	100
	C_{cr}	100	100	100	100	100	100	100
	C_{tg}	100	100	100	100	100	100	100

Table 5.4: Empirical level and power (%) of test statistics LR, C_{ml}, C_{cr} and C_{tg}; based on 3000 iterations and $n_1 = 50, n_2 = 50, \alpha = 0.05$

		(β_1, β_2)						
(α_1, α_2)	Statistics	$(3, 3)$	$(3, 3.5)$	$(3, 4)$	$(3, 4.5)$	$(3, 5)$	$(3, 5.5)$	$(3, 6)$
(10, 10)	LR	4.8	4.9	5.3	5.5	5.5	5.7	6.1
	C_{ml}	5.0	5.2	5.4	5.5	5.7	5.8	6.1
	C_{cr}	5.1	5.2	5.4	5.7	5.7	5.8	6.3
	C_{tg}	5.1	5.3	5.6	5.8	6.0	6.2	6.4
(10, 11)	LR	38.0	44.5	48.5	52.4	57.3	59.9	62.9
	C_{ml}	34.2	41.4	46.4	50.4	54.0	57.1	60.0
	C_{cr}	32.0	38.1	43.6	47.5	51.1	54.0	57.7
	C_{tg}	29.9	35.7	40.6	44.3	48.4	51.3	55.0
(10, 12)	LR	59.4	66.6	72.1	76.9	81.1	84.3	89.4
	C_{ml}	56.9	63.8	69.0	72.9	77.5	80.4	85.9
	C_{cr}	53.8	60.7	66.2	69.9	74.1	77.5	82.6
	C_{tg}	50.6	57.3	63.2	66.9	70.6	73.9	79.5
(10, 13)	LR	81.4	87.9	93.9	97.6	99.4	99.8	100
	C_{ml}	77.7	83.1	88.5	93.1	96.3	97.8	99.9
	C_{cr}	74.2	79.7	85.4	90.0	92.7	94.7	98.0
	C_{tg}	70.5	76.6	81.3	85.7	88.8	90.8	94.4
(10, 14)	LR	96.7	100	100	100	100	100	100
	C_{ml}	93.3	99.8	100	100	100	100	100
	C_{cr}	88.6	95.6	99.8	100	100	100	100
	C_{tg}	84.6	92.3	97.1	98.9	100	100	100
(10, 15)	LR	100	100	100	100	100	100	100
	C_{ml}	100	100	100	100	100	100	100
	C_{cr}	99.9	100	100	100	100	100	100
	C_{tg}	98.2	100	100	100	100	100	100
(10, 17)	LR	100	100	100	100	100	100	100
	C_{ml}	100	100	100	100	100	100	100
	C_{cr}	100	100	100	100	100	100	100
	C_{tg}	100	100	100	100	100	100	100

Table 5.5: Failure Times of Different Bearing Specimens

Type of Compound				
I	*II*	*III*	*IV*	*V*
3.03	3.19	3.46	5.88	6.43
5.53	4.26	5.22	6.74	9.97
5.60	4.47	5.69	6.90	10.39
9.30	4.53	6.54	6.98	13.55
9.92	4.67	9.16	7.21	14.45
12.51	4.69	9.40	8.14	14.72
12.95	5.78	10.19	8.59	16.81
15.21	6.79	10.71	9.80	18.39
16.04	9.37	12.58	12.28	20.84
16.84	12.75	13.41	25.46	21.51

Table 5.6: Estimates of parameters obtained by different methods for compound combinations of bearing specimens data in Table 5.5

	Compound Type Combinations									
	(I,II)	(I,III)	(I,IV)	(I,V)	(II,III)	(II,IV)	(II,V)	(III,IV)	(III,V)	(IV,V)
$\hat{\alpha}_0$	9.0056	10.4848	11.7213	14.7887	8.5093	7.1645	9.5075	8.6567	13.8549	15.0676
$\hat{\beta}_{10}$	1.8385	2.2491	2.5351	2.4628	2.3276	2.3718	2.1549	2.3160	2.2909	1.9228
$\hat{\beta}_{20}$	2.2376	3.2077	1.9758	3.1844	2.6780	1.5713	1.4804	1.3236	2.8340	3.2844
$\hat{\alpha}_{1a}$	12.0607	12.0607	12.0607	12.0607	6.8596	6.8596	6.8596	9.6847	9.6847	7.5107
$\hat{\alpha}_{2a}$	6.8596	9.6847	7.5107	16.3507	9.6847	7.5107	16.3507	7.5107	16.3507	16.3507
$\hat{\beta}_{1a}$	2.5881	2.5881	2.5881	2.5881	2.3202	2.3202	2.3202	3.1324	3.1324	4.0912
$\hat{\beta}_{2a}$	2.3202	3.1324	4.0912	3.6518	3.1324	4.0912	3.6518	4.0912	3.6518	3.6518
$\hat{\alpha}_{cr}$	8.2096	10.4267	11.6395	14.1747	8.0635	7.6565	9.3231	10.0022	11.7626	14.1524
$\hat{\beta}_{1cr}$	2.4941	2.4941	2.4941	2.4941	2.6956	2.6956	2.6956	3.0152	3.0152	2.5244
$\hat{\beta}_{2cr}$	2.6956	3.0152	2.5244	3.5348	3.0152	2.5244	3.5348	2.5244	3.5348	3.5348
$\hat{\alpha}_{tg}$	7.9949	9.3692	11.7085	12.7716	7.7511	8.2172	9.0596	9.7884	10.6421	14.1160
$\hat{\beta}_{1tg}$	2.0733	2.0733	2.0733	2.0733	2.2457	2.2457	2.2457	2.5192	2.5192	2.0992
$\hat{\beta}_{2tg}$	2.2457	2.5192	2.0992	2.9636	2.5192	2.0992	2.9636	2.0992	2.9636	2.9636

Table 5.7: Test statistics along with p-values for compound combinations of bearing specimens data in Table 5.5

	Compound Type Combinations									
	(I, II)	(I, III)	(I, IV)	(I, V)	(II, III)	(II, IV)	(II, V)	(III, IV)	(III, V)	(IV, V)
LR	7.0443 (0.0080)	1.6233 (0.2026)	0.5067 (0.4766)	3.4073 (0.0649)	3.4310 (0.0640)	7.4859 (0.0062)	18.8333 (0.0000)	1.0934 (0.2957)	10.1554 (0.0014)	2.9559 (0.0856)
C_{ml}	5.4028 (0.0201)	1.4351 (0.2309)	0.1191 (0.7300)	2.6298 (0.1049)	2.7254 (0.0988)	3.6041 (0.0576)	8.5357 (0.0035)	0.8114 (0.3677)	5.8944 (0.0152)	2.6429 (0.1040)
C_{cr}	18.7256 (0.0000)	2.0846 (0.1488)	1.0047 (0.3162)	5.1513 (0.0232)	4.7590 (0.0291)	29.3146 (0.0000)	171.7005 (0.0000)	2.9979 (0.0834)	30.3739 (0.0000)	4.0519 (0.0441)
C_{tg}	11.6447 (0.0006)	1.9165 (0.1662)	0.9138 (0.3391)	4.5956 (0.0321)	3.7401 (0.0531)	9.0158 (0.0027)	90.3166 (0.0000)	1.1887 (0.2756)	29.2231 (0.0000)	3.0675 (0.0799)

Chapter 6

Summary and Recommendations for Future Research Topics

6.1 Summary

Often count data with extra Poisson variation, binary data having extra-binomial variation and Weibull distributed survival data may appear as two or more groups. In the scenario of two groups data, one group may be the control group and the other one the treatment group. It may be of interest to compare some characteristic of the two groups of data. For over-dispersed count data, considering negative binomial model, the objective may be to test the equality of means of two groups assuming that the over-dispersion parameters are unequal. When data in the form of proportion fit the beta-binomial model it may be of interest to test the equality of proportions in two groups with unequal dispersion parameters. In case of Weibull distributed survival data, testing the equality of scale parameters of two groups where the shape parameters are unequal may be the objective.

In Chapter 3, we developed six test procedures, namely, $LR, LR\,(bc)\,, T^2, T^2\,(bc)\,, T_1$ and T_N, for testing the equality of two negative binomial means in the presence of unequal dispersion parameters. We also studied, through simulation studies, the performance of the test procedures for small, moderate and large sample size situations. Performance of the test procedures were compared on the basis of empirical size and power. We did not find reason to recommend the bias corrected statistics, $LR\,(bc)$ and $T^2\,(bc)$, over their uncorrected counterparts because these had not shown improvement in power and closeness to nominal level. In addition, bias corrected statistics are more computational intensive. For small to moderate sample sizes the statistic T_1 is recommended as this statistic shows best performance in maintaining size and power for most combinations of $(\mu_1,\ \mu_2)$ and $(c_1,\ c_2)$. Apart from this, T_1 is easy to understand and compute. In large sample size situations all six statistics maintain nominal level reasonably well and the power performance are found similar. Therefore, for large samples, no substantial advantage has been found to use one over others except that T_N is the easiest to implement.

In Chapter 4, we derived parametric as well as semi-parametric statistics for testing the equality of two proportions in the presence of unequal dispersion parameters. We then, through simulation studies, compared these statistics in terms of size and power. The parametric statistics are LR, C_{ml}, and C_{kmm}. Of these three C_{ml} and C_{kmm} are $C\,(\alpha)$ statistics. The semi-parametric statistics are C_{qb}, C_{qs}, C_{eq}, C_{ar} and C_{ars} among which the first three are $C\,(\alpha)$ statistics. Among the parametric tests LR, in general, maintains nominal level very well and its power performance is the best. This statistic needs estimates of the parameters both under the null and the alternative hypotheses. Both of the statistics, C_{ml} and C_{kmm}, need the estimates of the parameters only under the null hypothesis and are similar in maintaining nominal level. These are conservative for small differences in dispersion parameters, (θ_1, θ_2), and small values of proportions, (π_1, π_2), but liberal otherwise. In regards to power, the statistic C_{kmm} performs better than C_{ml} in all parameter combinations. However, C_{kmm} is based on the method of moments estimate and is computationally less intensive than C_{ml} which is based on the maximum likelihood estimate. We, thus, recommend C_{kmm} among the parametric statistics. With small differences in the dispersion parameters, θ_1 and θ_2, and for small values of $\pi_1 = \pi_2$, the semi-parametric $C(\alpha)$ statistics C_{qb}, C_{qs} and C_{eq} behave conservatively. Overall, the statistic based on the extended quasi-likelihood estimates of the nuisance parameters, C_{eq}, maintains nominal level well and the power performance is the best for this statistic. In addition, this statistic is easy to implement because the estimates of the mean and dispersion parameters are obtained from a single function. On the other hand, in addition to the quasi-likelihood function the use of method of moments is required for finding mean and dispersion parameters while computing the statistics C_{qb} and C_{qs}. Therefore, the statistic C_{eq} can be recommended for use among the semi-parametric procedures.

In Chapter 5, we constructed a likelihood ratio test LR, and three $C(\alpha)$ tests C_{ml}, C_{cr} and C_{tg}, for testing the equality of scale parameters of two Weibull distributions where the shape parameters are unequal. For small sample sizes the statistics LR and C_{ml} hold effective nominal level while C_{cr} and C_{tg} are liberal. The statistic C_{cr} performs best in maintaining power. For moderate and large sample sizes, LR exhibits the highest powers for all instances though all four statistics hold nominal level effectively. We, thus, taking into account the ease of computation and implementation, recommend the use of C_{cr} for small sample sizes. For moderate and large sample sizes the statistic LR is recommended.

6.2 Recommendations for Future Research

6.2.1 Behrens-Fisher Analogs for Zero-Inflated Discrete Data

Often unbounded count data exhibit an excess of zeroes compared to what is expected from negative binomial model and the bounded count data appear with number of zeroes more than expected from beta-binomial model. To accommodate this inflated number of zeroes, it is assumed that the distribution is a mixture of the original distribution and a degenerate distribution at zero. If the original distribution is negative binomial then the model that accounts the extra zeroes is called the zero-inflated negative binomial model. Similarly, the model that incorporates excess zeroes in bounded counts is called zero-inflated beta-binomial model.

The probability function of a zero-inflated negative random variable Y can be written as

$$f\left(y|\mu,\ c,\ \eta\right) = Pr\left(Y = y|\mu,\ c,\ \eta\right)$$
$$= \begin{cases} \eta + (1-\eta)\left(\dfrac{1}{1+c\mu}\right)^{c^{-1}} & \text{if } y = 0; \\ (1-\eta)\dfrac{\Gamma\left(y+c^{-1}\right)}{y!\Gamma\left(c^{-1}\right)}\left(\dfrac{c\mu}{1+c\mu}\right)^{y}\left(\dfrac{1}{1+c\mu}\right)^{c^{-1}} & \text{if } y = 1, 2, \cdots, \end{cases}$$

where $0 \leq \eta \leq 1$.

The mean and variance of Y are $E\left(Y\right) = (1-\eta)\,\mu$ and $\text{var}\left(Y\right) = (1-\eta)\,\mu\left[1+\mu\left(\eta+c\right)\right]$ respectively. For two groups of negative binomial data with excess number of zeroes, one possible area of further research is to derive test procedures for testing equality of means where the dispersion parameters are assumed unknown and unequal considering that η's from both groups are equal. Another possible area of further research is to develop tests for testing the simultaneous equality of μ_1, μ_2 and η_1, η_2 where c_1 and c_2 are unequal.

The probability function of a zero-inflated beta-binomial random variable Y with index n can be expressed as

$$f(y|n,\ \pi,\ \theta,\ \eta) = Pr(Y = y|n,\ \pi,\ \theta,\ \eta)$$
$$= \begin{cases} \eta + (1-\eta)BB(\pi,\ \theta) & \text{if } y = 0; \\ (1-\eta)BB(\pi,\ \theta) & \text{if } y = 1, 2, \cdots, n, \end{cases}$$

where $0 \leq \eta \leq 1$ and $BB(\pi,\ \theta)$ is the probability function of a beta-binomial variate given in equation 4.2. The mean and variance of beta-binomial variate, Y are $E(Y) = (1-\eta)\,n\pi$ and $\text{var}(Y) = (1-\eta)\,n\pi\,[\eta n\pi + (1-\pi)\{1 + (n-1)\,\theta\}]$ respectively. A consideration can be taken into account as a further research topic that involves in developing test procedures to test the equality of proportions $(\pi_1,\ \pi_2)$ for two groups of beta-binomial data having excess number of zeroes with common η and unequal θ_1, θ_2. Likewise, developing test methods for simultaneously testing the equality of proportions $(\pi_1,\ \pi_2)$ and equality of zero inflation parameters $(\eta_1,\ \eta_2)$ in the presence of unequal dispersion parameters $(\theta_1,\ \theta_2)$ can be considered as a further research topic.

6.2.2 Behrens-Fisher Analog for Censored Survival Data

In Chapter 5, we developed test procedures for testing the equality of scale parameters for two Weibull distributions where the shape parameters are unequal in the complete sample scenario. In fact, in most instances, censoring is evident in survival data. We, therefore, suggest to develop test statistics, as a further research topic, for testing the equality of scale parameters of two Weibull distributions in the presence of unequal shape parameters in different censoring scenarios.

Appendix A

A.1 Derivation of the Biases of Maximum Likelihood Estimates of the Parameters Under the Null Hypothesis

For the negative binomial model (3.1), under the null hypothesis H_0: $\mu_1 = \mu_2 = \mu$, the kernel of the log-likelihood is

$$l_0 = \sum_{i=1}^{2} \sum_{j=1}^{n_i} \left[\sum_{l=0}^{y_{ij}-1} \log\left(\frac{1 + c_i l}{c_i}\right) + y_{ij} \log c_i + y_{ij} \log \mu - y_{ij} \log(1 + c_i \mu) \right.$$

$$\left. - c_i^{-1} \log(1 + c_i \mu) \right] \quad \text{(A.1)}$$

A.1.1 Derivation of the Second Order Quantities (V_{rt}) and the Elements of the Fisher Information Matrix (I_{rt})

Using the above log-likelihood, we obtain

$$V_{11} = \frac{\partial^2 l_0}{\partial \mu^2} = \sum_{i=1}^{2} \sum_{j=1}^{n_i} \left[\frac{-y_{ij} - 2y_{ij}c_i\mu + c_i\mu^2}{\mu^2(1+c_i\mu)^2} \right],$$

$$V_{12} = V_{21} = \frac{\partial^2 l_0}{\partial \mu \partial c_1} = \sum_{j=1}^{n_1} \left[\frac{-y_{1j} + \mu}{(1+c_1\mu)^2} \right], \quad V_{13} = V_{31} = \frac{\partial^2 l_0}{\partial \mu \partial c_2} = \sum_{j=1}^{n_2} \left[\frac{-y_{2j} + \mu}{(1+c_2\mu)^2} \right],$$

$$V_{22} = \frac{\partial^2 l_0}{\partial c_1^2} = -\frac{2n_1}{c_1^3} \log(1+c_1\mu) + \frac{n_1\mu}{c_1^2(1+c_1\mu)} - \frac{n_1(\bar{y}_1 - \mu)(1+2c_1\mu)}{c_1^2(1+c_1\mu)^2}$$

$$+ \sum_{j=1}^{n_1} \sum_{l=0}^{y_{1j}-1} \left\{ \frac{1+2c_1 l}{c_1^2(1+c_1 l)^2} \right\}, \quad V_{23} = V_{32} = \frac{\partial^2 l_0}{\partial c_1 \partial c_2} = 0$$

and

$$V_{33} = \frac{\partial^2 l_0}{\partial c_2^2} = -\frac{2n_2}{c_2^3} \log(1+c_2\mu) + \frac{n_2\mu}{c_2^2(1+c_2\mu)} - \frac{n_2(\bar{y}_2 - \mu)(1+2c_2\mu)}{c_2^2(1+c_2\mu)^2}$$

$$+ \sum_{j=1}^{n_2} \sum_{l=0}^{y_{2j}-1} \left\{ \frac{1+2c_2 l}{c_2^2(1+c_2 l)^2} \right\}.$$

Now, under the null hypothesis, using $E(y_{ij}) = \mu$, for $i = 1,\ 2;\ j = 1,\ 2,\ \cdots,\ n_i$, we have

$$I_{11} = E(-V_{11}) = \sum_{i=1}^{2}\left[\frac{n_i}{\mu(1+c_i\mu)}\right],$$

$$I_{12} = I_{21} = E(-V_{12}) = 0, \qquad I_{13} = I_{31} = E(-V_{13}) = 0,$$

$$I_{22} = E(-V_{22}) = \frac{2n_1}{c_1^3}\log(1+c_1\mu) - \frac{n_1\mu}{c_1^2(1+c_1\mu)}$$

$$-\frac{2}{c_1^2}E\left\{\sum_{j=1}^{n_1}\sum_{l=0}^{y_{1j}-1}\frac{1}{(1+c_1 l)}\right\} + \frac{1}{c_1^4}E\left\{\sum_{j=1}^{n_1}\sum_{l=0}^{y_{1j}-1}\frac{c_1^2}{(1+c_1 l)^2}\right\},$$

$$I_{23} = I_{32} = E(-V_{23}) = 0,$$

and

$$I_{33} = E(-V_{33}) = \frac{2n_2}{c_2^3}\log(1+c_2\mu) - \frac{n_2\mu}{c_2^2(1+c_2\mu)}$$

$$-\frac{2}{c_2^2}E\left\{\sum_{j=1}^{n_2}\sum_{l=0}^{y_{2j}-1}\frac{1}{(1+c_2 l)}\right\} + \frac{1}{c_2^4}E\left\{\sum_{j=1}^{n_2}\sum_{l=0}^{y_{2j}-1}\frac{c_2^2}{(1+c_2 l)^2}\right\}.$$

Since $E\left(\frac{\partial l}{\partial c_i}\right) = 0;$ we have $E\left\{\sum_{j=1}^{n_i}\sum_{l=0}^{y_{ij}-1}\frac{1}{(1+c_i l)}\right\} = \frac{n_i}{c_i}\log(1+c_i\mu).$

Further, following Fisher (1941), we have $E\left\{\sum_{j=1}^{n_i}\sum_{l=0}^{y_{ij}-1}\frac{c_i^2}{(1+c_i l)^2}\right\} = n_i\sum_{k=0}^{\infty}\frac{k!(c_i b_i)^{k+1}}{(k+1)d_{ik}}$

where $b_i = \frac{c_i\mu}{1+c_i\mu}$ and $d_{ik} = \prod_{l=0}^{k}(1+c_i l),\ i = 1,\ 2.$

We, thus obtain

$$I_{22} = \frac{n_1}{c_1^4} \sum_{k=1}^{\infty} \frac{k!(b_1 c_1)^{k+1}}{(k+1)d_{1k}} \quad \text{and} \quad I_{33} = \frac{n_2}{c_2^4} \sum_{k=1}^{\infty} \frac{k!(b_2 c_2)^{k+1}}{(k+1)d_{2k}}.$$

A.1.2 Derivation of the Third Order Quantities (W_{rtu}) and the Expected Values (J_{rtu})

The third order derivatives are $W_{rtu} = \dfrac{\partial^3 l_0}{\partial\theta_r \partial\theta_t \partial\theta_u}$; $r,\ t,\ u = 1,\ 2,\ 3$ which we obtain as follows

$$W_{111} = \frac{\partial^3 l_0}{\partial \mu^3} = \sum_{i=1}^{2} \sum_{j=1}^{n_i} \left[\frac{2y_{ij}}{\mu^3} - \frac{2y_{ij}c_i^3}{(1+c_i\mu)^3} - \frac{2c_i^2}{(1+c_i\mu)^3} \right],$$

$$W_{112} = W_{121} = W_{211} = \frac{\partial^3 l_0}{\partial \mu^2 \partial c_1} = \sum_{j=1}^{n_1} \left[\frac{1 - c_1\mu + 2y_{1j}c_1}{(1+c_1\mu)^3} \right],$$

$$W_{113} = W_{131} = W_{311} = \frac{\partial^3 l_0}{\partial \mu^2 \partial c_2} = \sum_{j=1}^{n_2} \left[\frac{1 - c_2\mu + 2y_{2j}c_2}{(1+c_2\mu)^3} \right],$$

$$W_{122} = W_{212} = W_{221} = \frac{\partial^3 l_0}{\partial \mu \partial c_1^2} = \sum_{j=1}^{n_1} \left[\frac{2\mu(y_{1j} - \mu)}{(1+c_1\mu)^3} \right],$$

$$W_{133} = W_{313} = W_{331} = \frac{\partial^3 l_0}{\partial \mu \partial c_2^2} = \sum_{j=1}^{n_2} \left[\frac{2\mu(y_{2j} - \mu)}{(1+c_2\mu)^3} \right],$$

$$W_{222} = \frac{\partial^3 l_0}{\partial c_1^3} = \frac{6n_1}{c_1^4}\log(1+c_1\mu) - \frac{n_1\mu(4+5c_1\mu)}{c_1^3(1+c_1\mu)^2} + \frac{2n_1(\bar{y}_1-\mu)(1+3c_1\mu+3c_1^2\mu^2)}{c_1^3(1+c_1\mu)^3}$$

$$-\frac{4}{c_1^3}\sum_{j=1}^{n_1}\sum_{l=0}^{y_{1j}-1}\frac{1}{(1+c_1 l)} - \frac{2}{c_1^3}\sum_{j=1}^{n_1}\sum_{l=0}^{y_{1j}-1}\frac{c_1^2 l^2 - c_1 l - 1}{(1+c_1 l)^3},$$

$$W_{233} = W_{323} = W_{332} = \frac{\partial^3 l_0}{\partial c_1 \partial c_2^2} = 0, \quad W_{322} = W_{232} = W_{223} = \frac{\partial^3 l_0}{\partial c_2 \partial c_1^2} = 0,$$

$$W_{333} = \frac{\partial^3 l_o}{\partial c_2^3} = \frac{6n_2}{c_2^4}\log(1+c_2\mu) - \frac{n_2\mu(4+5c_2\mu)}{c_2^3(1+c_2\mu)^2} + \frac{2n_2(\bar{y}_2-\mu)(1+3c_2\mu+3c_2^2\mu^2)}{c_2^3(1+c_2\mu)^3}$$

$$-\frac{4}{c_2^3}\sum_{j=1}^{n_2}\sum_{l=0}^{y_{2j}-1}\frac{1}{(1+c_2 l)} - \frac{2}{c_2^3}\sum_{j=1}^{n_2}\sum_{l=0}^{y_{2j}-1}\frac{c_2^2 l^2 - c_2 l - 1}{(1+c_2 l)^3} \text{ and}$$

$$W_{123} = W_{132} = W_{213} = W_{231} = W_{312} = W_{321} = \frac{\partial^3 l}{\partial \mu \partial c_1 \partial c_2} = 0$$

Now, the expected values of the above third order quantities are $J_{rtu} = E(W_{rtu})\,;\; r,\ t,\ u = 1,\ 2,\ 3$, which we obtain as follows

$$J_{111} = \sum_{i=1}^{2}\sum_{j=1}^{n_i}\frac{2(1+2c_i\mu)}{\mu^2(1+c_i\mu)^2}, \quad J_{112} = \frac{n_1}{(1+c_1\mu)^2}, \quad J_{113} = \frac{n_2}{(1+c_2\mu)^2}, \quad J_{122} = 0,$$

$$J_{123} = 0, \quad J_{133} = 0,$$

$$J_{222} = \frac{6n_1}{c_1^4}\log(1+c_1\mu) - \frac{n_1\mu(4+5c_1\mu)}{c_1^3(1+c_1\mu)^2} - \frac{4}{c_1^3}E\left\{\sum_{j=1}^{n_1}\sum_{l=0}^{y_{1j}-1}\frac{1}{1+c_1 l}\right\}$$

$$-\frac{2}{c_1^3}E\left\{\sum_{j=1}^{n_1}\sum_{l=0}^{y_{1j}-1}\frac{c_1^2l^2-c_1l-1}{(1+c_1l)^3}\right\},$$

$$J_{222} = \frac{2n_1}{c_1^2}\kappa_1 - \frac{n_1\mu(4+5c_1\mu)}{c_1^3(1+c_1\mu)^2} - \frac{2n_1}{c_1^3}\Delta_1, \quad J_{223}=0, \quad J_{233}=0$$

and

$$J_{333} = \frac{2n_2}{c_2^2}\kappa_2 - \frac{n_2\mu(4+5c_2\mu)}{c_2^3(1+c_2\mu)^2} - \frac{2n_2}{c_2^3}\Delta_2,$$

where $\kappa_i = \frac{1}{c_i^2}\log(1+c_i\mu)$, and $\Delta_i = \sum_{y_{ij}=0}^{\infty}\sum_{l=0}^{y_{ij}-1}\frac{c_i^2l^2-c_il-1}{(1+c_il)^3}Pr(y_{ij});\ i=1,\ 2;\ j=1,\ 2,\ \cdots,\ n_i$.

A.1.3 Derivation of the Partial Derivatives of the Expected Values of Second Order Quantities $\left(K_{rt}^{(u)}\right)$

In addition to the expectations of second order and third order quantities, Cordeiro and Klein (1994) formula of biases of the maximum likelihood estimates requires partial derivatives of the expected values of second order quantities, that is, $K_{rt}^{(u)} = \frac{\partial}{\partial\theta_u}E(V_{rt}),\ r,\ t,\ u=1,\ 2,\ 3.$ These partial derivatives are obtained as follows

$$K_{11}^{(1)} = \frac{\partial}{\partial \mu} E\left(V_{11}\right) = \sum_{i=1}^{2} \left\{ \frac{n_i(1+2c_i\mu)}{\mu^2(1+c_i\mu)^2} \right\},$$

$$K_{11}^{(2)} = \frac{\partial}{\partial c_1} E\left(V_{11}\right) = \frac{n_1}{(1+c_1\mu)^2}, \quad K_{11}^{(3)} = \frac{\partial}{\partial c_2} E\left(V_{11}\right) = \frac{n_2}{(1+c_2\mu)^2}$$

$$K_{12}^{(1)} = K_{21}^{(1)} = \frac{\partial}{\partial \mu} E\left(V_{12}\right) = 0, \quad K_{12}^{(2)} = K_{21}^{(2)} = \frac{\partial}{\partial c_1} E\left(V_{12}\right) = 0,$$

$$K_{12}^{(3)} = K_{21}^{(3)} = \frac{\partial}{\partial c_2} E\left(V_{12}\right) = 0, \quad K_{13}^{(1)} = K_{31}^{(1)} = \frac{\partial}{\partial \mu} E\left(V_{13}\right) = 0,$$

$$K_{13}^{(2)} = K_{31}^{(2)} = \frac{\partial}{\partial c_1} E\left(V_{13}\right) = 0, \quad K_{13}^{(3)} = K_{31}^{(3)} = \frac{\partial}{\partial c_2} E\left(V_{13}\right) = 0,$$

$$K_{22}^{(1)} = \frac{\partial}{\partial \mu} E\left(V_{22}\right) = -\frac{n_1}{c_1^4} \sum_{k=1}^{\infty} \frac{k!\left(b_1c_1\right)^k b_1^2}{\mu^2 d_{1k}},$$

$$K_{22}^{(2)} = \frac{\partial}{\partial c_1} E\left(V_{22}\right) = \frac{n_1}{c_1^5} \sum_{k=1}^{\infty} \frac{k!(c_1b_1)^{k+1}}{(k+1)d_{1k}} \left\{ 4 - \frac{(k+1)(2+c_1\mu)}{1+c_1\mu} + c_1 \left(\sum_{l=0}^{k} \frac{l}{1+c_1 l} \right) \right\},$$

$$K_{22}^{(3)} = \frac{\partial}{\partial c_2} E\left(V_{22}\right) = 0, \quad K_{23}^{(1)} = K_{32}^{(1)} = \frac{\partial}{\partial \mu} E\left(V_{23}\right) = 0,$$

$$K_{23}^{(2)} = K_{32}^{(2)} = \frac{\partial}{\partial c_1} E\left(V_{23}\right) = 0, \quad K_{23}^{(3)} = K_{32}^{(3)} = \frac{\partial}{\partial c_2} E\left(V_{23}\right) = 0,$$

$$K_{33}^{(1)} = \frac{\partial}{\partial \mu} E\left(V_{33}\right) = -\frac{n_2}{c_2^4} \sum_{k=1}^{\infty} \frac{k!\left(b_2c_2\right)^k b_2^2}{\mu^2 d_{2k}}, \quad K_{33}^{(2)} = \frac{\partial}{\partial c_1} E\left(V_{33}\right) = 0$$

and

$$K_{33}^{(3)} = \frac{\partial}{\partial c_2} E\left(V_{33}\right) = \frac{n_2}{c_2^5} \sum_{k=1}^{\infty} \frac{k!(c_2 b_2)^{k+1}}{(k+1)d_{2k}} \left\{ 4 - \frac{(k+1)(2+c_2\mu)}{1+c_2\mu} + c_2 \left(\sum_{l=0}^{k} \frac{l}{1+c_2 l} \right) \right\}.$$

A.1.4 The Biases of the Estimates of the Parameters

Now, following Cordeiro and Klein (1994), the biases of the maximum likelihood estimates of μ, c_1, and c_2 are

$$b_{\hat{\mu}}\left(\hat{\mu},\ \hat{c}_1,\ \hat{c}_2\right) = \sum_{r=1}^{3} M^{1r} \sum_{t=1}^{3} \sum_{u=1}^{3} \left(K_{rt}^{(u)} - \frac{1}{2} J_{rtu} \right) M^{tu},$$

$$b_{\hat{c}_1}\left(\hat{\mu},\ \hat{c}_1,\ \hat{c}_2\right) = \sum_{r=1}^{3} M^{2r} \sum_{t=1}^{3} \sum_{u=1}^{3} \left(K_{rt}^{(u)} - \frac{1}{2} J_{rtu} \right) M^{tu}$$

and

$$b_{\hat{c}_2}\left(\hat{\mu},\ \hat{c}_1,\ \hat{c}_2\right) = \sum_{r=1}^{3} M^{3r} \sum_{t=1}^{3} \sum_{u=1}^{3} \left(K_{rt}^{(u)} - \frac{1}{2} J_{rtu} \right) M^{tu}, \text{respectively.}$$

The above biases, after simplifications, take the forms

$$b_{\hat{\mu}}\left(\hat{\mu},\ \hat{c}_1,\ \hat{c}_2\right) = M^{11} \left(K_{11}^{(1)} - \frac{1}{2} J_{111} \right) M^{11} = 0,$$

$$b_{\hat{c}_1}\left(\hat{\mu},\ \hat{c}_1,\ \hat{c}_2\right) = M^{22} \left[-\frac{1}{2} J_{211} M^{11} + \left(K_{22}^{(2)} - \frac{1}{2} J_{222} \right) M^{22} \right]$$

and

$$b_{\hat{c}_2}\left(\hat{\mu},\ \hat{c}_1,\ \hat{c}_2\right) = M^{33} \left[-\frac{1}{2} J_{311} M^{11} + \left(K_{33}^{(3)} - \frac{1}{2} J_{333} \right) M^{33} \right]$$

respectively, where M^{rt} is the $(r,\ t)^{th}$ element of the inverse of the information matrix I.

A.2 Derivation of the Biases of Maximum Likelihood Estimates of the Parameters Under the Alternative Hypothesis

Under the alternative hypothesis we have four parameters to estimate, namely, $\mu_1,\ \mu_2,\ c_1$ and c_2, that is, under the alternative hypothesis $\theta' = (\theta_1,\ \theta_2,\ \theta_3,\ \theta_4) = (\mu_1,\ \mu_2,\ c_1,\ c_2)$. The estimates of μ_1 and μ_2 are $\bar{y}_1$ and $\bar{y}_2$ respectively, that is, under the alternative hypothesis the MLEs of μ_1 and μ_2 are unbiased. The second order and third order derivatives are $V_{rt} = \dfrac{\partial^2 l}{\partial\theta_r\partial\theta_t}$ and $W_{rtu} = \dfrac{\partial^3 l}{\partial\theta_r\partial\theta_t\partial\theta_u}$, $r,\ t,\ u = 1,\ 2,\ 3,\ 4$, and the partial derivatives of expected values of the second order terms are $K_{rt}^{(u)} = \dfrac{\partial}{\partial\theta_u} E\left(V_{rt}\right)$. The $(r,\ t)^{th}$ element of the Fisher information matrix is $I_{rt} = E\left(-V_{rt}\right)$ and the expected value of the $(r,\ t,\ u)^{th}$ term of the third order derivatives is $J_{rtu} = E\left(W_{rtu}\right)$. Following similar steps as in the previous section we obtain

$$V_{11} = \sum_{j=1}^{n_1} \left\{ \frac{-y_{1j} - 2y_{1j}c_1\mu_1 + c_1\mu_1^2}{\mu_1^2(1+c_1\mu_1)^2} \right\}, \quad V_{13} = V_{31} = \sum_{j=1}^{n_1} \left[\frac{-(y_{1j}-\mu_1)}{(1+c_1\mu_1)^2} \right],$$

$$V_{22} = \sum_{j=1}^{n_2} \left\{ \frac{-y_{2j} - 2y_{2j}c_2\mu_2 + c_2\mu_2^2}{\mu_2^2(1+c_2\mu_2)^2} \right\}, \quad V_{24} = V_{42} = \sum_{j=1}^{n_2} \left[\frac{-(y_{2j}-\mu_2)}{(1+c_2\mu_2)^2} \right],$$

$$V_{33} = -\frac{2n_1}{c_1^3}\log(1+c_1\mu_1) + \frac{n_1\mu_1}{c_1^2(1+c_1\mu_1)} - \frac{n_1\left(\bar{y}_1-\mu_1\right)\left(1+2c_1\mu_1\right)}{c_1^2(1+c_1\mu_1)^2}$$

$$+\sum_{j=1}^{n_1}\sum_{l=0}^{y_{1j}-1} \left\{ \frac{1+2c_1 l}{c_1^2(1+c_1 l)^2} \right\},$$

$$V_{44} = -\frac{2n_2}{c_2^3}\log(1+c_2\mu_2) + \frac{n_2\mu_2}{c_2^2(1+c_2\mu_2)} - \frac{n_2\left((\bar{y}_2-\mu_2)\left(1+2c_2\mu_2\right)\right)}{c_2^2(1+c_2\mu_2)^2}$$

$$+\sum_{j=1}^{n_2}\sum_{l=0}^{y_{2j}-1}\left\{\frac{1+2c_2 l}{c_2^2(1+c_2 l)^2}\right\},$$

$$I_{11} = \frac{n_1}{\mu_1(1+c_1\mu_1)}, \quad I_{22} = \frac{n_2}{\mu_2(1+c_2\mu_2)},$$

$$I_{33} = \frac{n_1}{c_1^4}\sum_{k=1}^{\infty}\frac{k!(c_1b_1)^{k+1}}{(k+1)d_{1k}}, \quad I_{44} = \frac{n_2}{c_2^4}\sum_{k=1}^{\infty}\frac{k!(c_2b_2)^{k+1}}{(k+1)d_{2k}},$$

$$J_{111} = \frac{2n_1(1+2c_1\mu_1)}{\mu_1^2(1+c_1\mu_1)^2}, \quad J_{113} = \frac{n_1}{(1+c_1\mu_1)^2},$$

$$J_{222} = \frac{2n_2(1+2c_2\mu_2)}{\mu_2^2(1+c_2\mu_2)^2}, \quad J_{224} = \frac{n_2}{(1+c_2\mu_2)^2},$$

$$J_{333} = \frac{2n_1}{c_1^2}\kappa_1 - \frac{n_1\mu_1(4+5c_1\mu_1)}{c_1^3(1+c_1\mu_1)^2} - \frac{2n_1}{c_1^3}\Delta_1,$$

$$J_{444} = \frac{2n_2}{c_2^2}\kappa_2 - \frac{n_2\mu_2(4+5c_2\mu_2)}{c_2^3(1+c_2\mu_2)^2} - \frac{2n_2}{c_2^3}\Delta_2,$$

$$K_{11}^{(1)} = \frac{n_1(1+2c_1\mu_1)}{\mu_1^2(1+c_1\mu_1)^2}, \quad K_{11}^{(3)} = \frac{n_1}{(1+c_1\mu_1)^2}, \quad K_{22}^{(2)} = \frac{n_2(1+2c_2\mu_2)}{\mu_2^2(1+c_2\mu_2)^2},$$

$$K_{22}^{(4)} = \frac{n_2}{(1+c_2\mu_2)^2}, \quad K_{33}^{(1)} = -\frac{n_1}{c_1^4}\sum_{k=1}^{\infty}\frac{k!(c_1b_1)^k b_1^2}{\mu_1^2 d_{1k}},$$

$$K_{33}^{(3)} = \frac{n_1}{c_1^5}\sum_{k=1}^{\infty}\frac{k!(c_1b_1)^{k+1}}{(k+1)d_{1k}}\left\{4 - \frac{(k+1)(2+c_1\mu_1)}{1+c_1\mu_1} + c_1\left(\sum_{l=0}^{k}\frac{l}{1+c_1 l}\right)\right\},$$

$$K_{44}^{(2)} = -\frac{n_2}{c_2^4}\sum_{k=1}^{\infty}\frac{k!(c_2b_2)^k b_2^2}{\mu_2^2 d_{2k}}$$

and

$$K_{44}^{(4)} = \frac{n_2}{c_2^5}\sum_{k=1}^{\infty}\frac{k!(c_2b_2)^{k+1}}{(k+1)d_{2k}}\left\{4-\frac{(k+1)(2+c_2\mu_2)}{1+c_2\mu_2}+c_2\left(\sum_{l=0}^{k}\frac{l}{1+c_2 l}\right)\right\}.$$

As in the previous section, following Cordeiro and Klein (1994), the biases of the estimates of c_1 and c_2 are

$$b_{\hat{c}_1}\left(\hat{\mu}_1,\ \hat{\mu}_2,\ \hat{c}_1,\ \hat{c}_2\right) = \sum_{r=1}^{4} M^{3r}\sum_{t=1}^{4}\sum_{u=1}^{4}\left(K_{rt}^{(u)}-\frac{1}{2}J_{rtu}\right)M^{tu}$$
$$= M^{33}\sum_{t=1}^{4}\sum_{u=1}^{4}\left(K_{3t}^{(u)}-\frac{1}{2}J_{3tu}\right)M^{tu}$$

and

$$b_{\hat{c}_2}\left(\hat{\mu}_1,\ \hat{\mu}_2,\ \hat{c}_1,\ \hat{c}_2\right) = \sum_{r=1}^{4} M^{4r}\sum_{t=1}^{4}\sum_{u=1}^{4}\left(K_{rt}^{(u)}-\frac{1}{2}J_{rtu}\right)M^{tu}$$
$$= M^{44}\sum_{t=1}^{4}\sum_{u=1}^{4}\left(K_{4t}^{(u)}-\frac{1}{2}J_{4tu}\right)M^{tu}$$

respectively, which, after simplification, take the forms

$$b_{\hat{c}_1}\left(\hat{\mu}_1,\ \hat{\mu}_2,\ \hat{c}_1,\ \hat{c}_2\right) = M^{33}\left\{-\frac{1}{2}J_{311}M^{11}+\left(K_{33}^{(3)}-\frac{1}{2}J_{333}\right)M^{33}\right\} \quad \text{and}$$

$$b_{\hat{c}_2}\left(\hat{\mu}_1,\ \hat{\mu}_2,\ \hat{c}_1,\ \hat{c}_2\right) = M^{44}\left\{-\frac{1}{2}J_{422}M^{22}+\left(K_{44}^{(4)}-\frac{1}{2}J_{444}\right)M^{44}\right\}$$

respectively.

A.3 Derivation of the Terms Needed for the Score Test

After the reparameterization $\mu_i = \mu + \phi_i,\ i = 1,\ 2$ the log-likelihood, apart from a constant, can be written as

$$l = \sum_{i=1}^{2} \sum_{j=1}^{n_i} \left[\sum_{l=0}^{y_{ij}-1} \log\left(\frac{1+c_i l}{c_i}\right) + y_{ij}\log(c_i) + y_{ij}\log(\mu+\phi_i) \right.$$

$$\left. -y_{ij}\log\{1+c_i(\mu+\phi_i)\} - \frac{1}{c_i}\log\{1+c_i(\mu+\phi_i)\} \right].$$

We then obtain

$$\left.\frac{\partial l}{\partial \phi_1}\right|_{\phi=0} = \frac{n_1(\bar{y}_1-\mu)}{\mu(1+c_1\mu)}, \quad \left.\frac{\partial l}{\partial \mu}\right|_{\phi=0} = \sum_{i=1}^{2}\frac{n_i(\bar{y}_i-\mu)}{\mu(1+c_i\mu)},$$

$$\left.\frac{\partial l}{\partial c_i}\right|_{\phi=0} = \frac{n_i(\bar{y}_i-\mu)}{c_i(1+c_i\mu)} + \frac{n_i}{c_i^2}\log(1+c_i\mu) - \sum_{j=1}^{n_i}\sum_{l=0}^{y_{ij}-1}\frac{1}{c_i(1+c_i l)}, \quad i = 1,\ 2,$$

$$\left.\frac{\partial^2 l}{\partial \phi_1^2}\right|_{\phi=0} = \sum_{j=1}^{n_1}\left[-\frac{y_{1j}}{\mu^2} + \frac{y_{1j}c_1^2}{(1+c_1\mu)^2} + \frac{c_1}{(1+c_1\mu)^2}\right],$$

$$\left.\frac{\partial^2 l}{\partial \phi_1 \partial \mu}\right|_{\phi=0} = \sum_{j=1}^{n_1}\left[-\frac{y_{1j}}{\mu^2} + \frac{y_{1j}c_1^2}{(1+c_1\mu)^2} + \frac{c_1}{(1+c_1\mu)^2}\right],$$

$$\left.\frac{\partial^2 l}{\partial \phi_1 \partial c_1}\right|_{\phi=0} = \sum_{j=1}^{n_1}\left[\frac{-y_{1j}+\mu}{(1+c_1\mu)^2}\right], \quad \left.\frac{\partial^2 l}{\partial \phi_1 \partial c_2}\right|_{\phi=0} = 0,$$

$$\frac{\partial^2 l}{\partial \mu^2}\Big|_{\phi=0} = \sum_{i=1}^{2}\sum_{j=1}^{n_i}\left[-\frac{y_{ij}}{\mu^2} + \frac{y_{ij}c_i^2}{(1+c_i\mu)^2} + \frac{c_i}{(1+c_i\mu)^2}\right],$$

$$\frac{\partial^2 l}{\partial \mu \partial c_1}\Big|_{\phi=0} = \sum_{j=1}^{n_1}\left[\frac{-y_{1j}+\mu}{(1+c_1\mu)^2}\right], \quad \frac{\partial^2 l}{\partial \mu \partial c_2}\Big|_{\phi=0} = \sum_{j=1}^{n_2}\left[\frac{-y_{2j}+\mu}{(1+c_2\mu)^2}\right],$$

$$\frac{\partial^2 l}{\partial c_1^2}\Big|_{\phi=0} = -\frac{2n_1}{c_1^3}\log(1+c_1\mu) + \frac{n_1\mu}{c_1^2(1+c_1\mu)} - \frac{n_1(\bar{y}_1-\mu)(1+2c_1\mu)}{c_1^2(1+c_1\mu)^2}$$

$$+\sum_{j=1}^{n_1}\sum_{l=0}^{y_{1j}-1}\left\{\frac{1+2c_1 l}{c_1^2(1+c_1 l)^2}\right\},$$

$$\frac{\partial^2 l}{\partial c_1 \partial c_2}\Big|_{\phi=0} = 0$$

$$\frac{\partial^2 l}{\partial c_2^2}\Big|_{\phi=0} = -\frac{2n_2}{c_2^3}\log(1+c_2\mu) + \frac{n_2\mu}{c_2^2(1+c_2\mu)} - \frac{n_2(\bar{y}_2-\mu)(1+2c_2\mu)}{c_2^2(1+c_2\mu)^2}$$

$$+\sum_{j=1}^{n_2}\sum_{l=0}^{y_{2j}-1}\left\{\frac{1+2c_2 l}{c_2^2(1+c_2 l)^2}\right\}.$$

The expected values of the negative of mixed partial derivatives are obtained as follows

$$A_1 = E\left[-\frac{\partial^2 l}{\partial \phi_1 \partial \mu}\Big|_{\phi=0}\right] = \frac{n_1}{\mu(1+c_1\mu)}, \quad A_2 = E\left[-\frac{\partial^2 l}{\partial \phi_1 \partial c_1}\Big|_{\phi=0}\right] = 0,$$

$$A_3 = E\left[-\frac{\partial^2 l}{\partial \phi_1 \partial c_2}\Big|_{\phi=0}\right] = 0, \quad D = E\left[-\frac{\partial^2 l}{\partial \phi_1^2}\Big|_{\phi=0}\right] = \frac{n_1}{\mu(1+c_1\mu)},$$

$$B_{1,1} = E\left[-\frac{\partial^2 l}{\partial \mu^2}\bigg|_{\phi=0} \right] = \sum_{i=1}^{2}\left[\frac{n_i}{\mu\,(1+c_i\mu)} \right], \quad B_{1,2} = B_{2,1} = E\left[-\frac{\partial^2 l}{\partial \mu \partial c_1}\bigg|_{\phi=0} \right] = 0,$$

$$B_{1,3} = B_{3,1} = E\left[-\frac{\partial^2 l}{\partial \mu \partial c_2}\bigg|_{\phi=0} \right] = 0, \quad B_{2,2} = E\left[-\frac{\partial^2 l}{\partial c_1^2}\bigg|_{\phi=0} \right] = \frac{n_1}{c_1^4}\sum_{k=1}^{\infty}\frac{k!(b_1c_1)^{k+1}}{(k+1)d_{1k}},$$

$$B_{2,3} = B_{3,2} = E\left[-\frac{\partial^2 l}{\partial c_1 \partial c_2}\bigg|_{\phi=0} \right] = 0 \quad \text{and} \quad B_{3,3} = E\left[-\frac{\partial^2 l}{\partial c_2^2}\bigg|_{\phi=0} \right] = \frac{n_2}{c_2^4}\sum_{k=1}^{\infty}\frac{k!(b_2c_2)^{k+1}}{(k+1)d_{2k}}.$$

Appendix B

B.1 Expected Values of Negative of the Mixed Partial Derivatives in Beta Binomial Model

The kernel of the log-likelihood, after the reparameterization $\pi_i = \pi + \phi_i$, $i = 1, 2$, of a beta binomial model, is

$$l = \sum_{i=1}^{2} \sum_{j=1}^{m_i} \left[\sum_{r=0}^{y_{ij}-1} \log\{(1-\theta_i)(\pi+\phi_i) + r\theta_i\} + \sum_{r=0}^{n_{ij}-y_{ij}-1} \log\{(1-\theta_i)(1-\pi-\phi_i) + r\theta_i\} - \sum_{r=0}^{n_{ij}-1} \log\{(1-\theta_i) + r\theta_i\} \right].$$

We then obtain

$$\frac{\partial l}{\partial \phi_1}\Big|_{\phi=0} = \sum_{j=1}^{m_1}\left\{\left[\sum_{r=0}^{y_{1j}-1}\frac{(1-\theta_1)}{(1-\theta_1)\,\pi + r\theta_1}\right\} - \sum_{r=0}^{n_{1j}-y_{1j}-1}\left\{\frac{(1-\theta_1)}{(1-\theta_1)\,(1-\pi) + r\theta_1}\right\}\right],$$

$$\frac{\partial l}{\partial \pi}\Big|_{\phi=0} = \sum_{i=1}^{2}\sum_{j=1}^{m_i}\left[\sum_{r=0}^{y_{ij}-1}\left\{\frac{(1-\theta_i)}{(1-\theta_i)\,\pi + r\theta_i}\right\} - \sum_{r=0}^{n_{ij}-y_{ij}-1}\left\{\frac{(1-\theta_i)}{(1-\theta_i)\,(1-\pi) + r\theta_i}\right\}\right],$$

$$\frac{\partial l}{\partial \theta_i}\Big|_{\phi=0} = \sum_{j=1}^{m_i}\left[\sum_{r=0}^{y_{ij}-1}\left\{\frac{(r-\pi)}{(1-\theta_i)\,\pi + r\theta_i}\right\} + \sum_{r=0}^{n_{ij}-y_{ij}-1}\left\{\frac{(r+\pi-1)}{(1-\theta_i)\,(1-\pi) + r\theta_i}\right\}\right.$$
$$\left. - \sum_{r=0}^{n_{ij}-1}\left\{\frac{(r-1)}{(1-\theta_i) + r\theta_i}\right\}\right], \;\; i = 1,\; 2,$$

$$\frac{\partial^2 l}{\partial \phi_1^2}\Big|_{\phi=0} = \sum_{j=1}^{m_1}\left[\sum_{r=0}^{y_{1j}-1}\left\{-\frac{(1-\theta_1)^2}{[(1-\theta_1)\,\pi + r\theta_1]^2}\right\} - \sum_{r=0}^{n_{1j}-y_{1j}-1}\left\{\frac{(1-\theta_1)^2}{[(1-\theta_1)\,(1-\pi) + r\theta_1]^2}\right\}\right],$$

$$\frac{\partial^2 l}{\partial \phi_1 \partial \pi}\Big|_{\phi=0} = \sum_{j=1}^{m_1}\left[\sum_{r=0}^{y_{1j}-1}\left\{-\frac{(1-\theta_1)^2}{[(1-\theta_1)\,\pi + r\theta_1]^2}\right\} - \sum_{r=0}^{n_{1j}-y_{1j}-1}\frac{(1-\theta_1)^2}{[(1-\theta_1)\,(1-\pi) + r\theta_1]^2}\right],$$

$$\frac{\partial^2 l}{\partial \phi_1 \partial \theta_1}\Big|_{\phi=0} = \sum_{j=1}^{m_1}\left[\sum_{r=0}^{y_{1j}-1}\left\{-\frac{r}{[(1-\theta_1)\,\pi + r\theta_1]^2}\right\} + \sum_{r=0}^{n_{1j}-y_{1j}-1}\left\{\frac{r}{[(1-\theta_1)\,(1-\pi) + r\theta_1]^2}\right\}\right],$$

$$\frac{\partial^2 l}{\partial \phi_1 \partial \theta_2}\Big|_{\phi=0} = 0,$$

$$\frac{\partial^2 l}{\partial \pi^2}\Big|_{\phi=0} = \sum_{i=1}^{2}\sum_{j=1}^{m_i}\left[\sum_{r=0}^{y_{ij}-1}\left\{-\frac{(1-\theta_i)^2}{[(1-\theta_i)\pi + r\theta_i]^2}\right\} - \sum_{r=0}^{n_{ij}-y_{ij}-1}\left\{\frac{(1-\theta_i)^2}{[(1-\theta_i)(1-\pi) + r\theta_i]^2}\right\}\right],$$

$$\frac{\partial^2 l}{\partial \pi \partial \theta_1}\Big|_{\phi=0} = \sum_{j=1}^{m_1}\left[\sum_{r=0}^{y_{1j}-1}\left\{-\frac{r}{[(1-\theta_1)\pi + r\theta_1]^2}\right\} + \sum_{r=0}^{n_{1j}-y_{1j}-1}\left\{\frac{r}{[(1-\theta_1)(1-\pi) + r\theta_1]^2}\right\}\right],$$

$$\frac{\partial^2 l}{\partial \pi \partial \theta_2}\Big|_{\phi=0} = 0,$$

$$\frac{\partial^2 l}{\partial \theta_1^2}\Big|_{\phi=0} = \sum_{j=1}^{m_1}\left[\sum_{r=0}^{y_{1j}-1}\left\{-\frac{(r-\pi)^2}{[(1-\theta_1)\pi + r\theta_1]^2}\right\} - \sum_{r=0}^{n_{1j}-y_{1j}-1}\left\{\frac{(r+\pi-1)^2}{[(1-\theta_1)(1-\pi) + r\theta_1]^2}\right\} + \sum_{r=0}^{n_{1j}-1}\left\{\frac{(r-1)^2}{[(1-\theta_1) + r\theta_1]^2}\right\}\right],$$

$$\frac{\partial^2 l}{\partial \theta_1 \partial \theta_2}\Big|_{\phi=0} = 0 \quad \text{and}$$

$$\frac{\partial^2 l}{\partial \theta_2^2}\Big|_{\phi=0} = \sum_{j=1}^{m_2}\left[\sum_{r=0}^{y_{2j}-1}\left\{-\frac{(r-\pi)^2}{[(1-\theta_2)\pi + r\theta_2]^2}\right\} - \sum_{r=0}^{n_{2j}-y_{2j}-1}\left\{\frac{(r+\pi-1)^2}{[(1-\theta_2)(1-\pi) + r\theta_2]^2}\right\} + \sum_{r=0}^{n_{2j}-1}\left\{\frac{(r-1)^2}{[(1-\theta_2) + r\theta_2]^2}\right\}\right].$$

In order to obtain the expected values of the negative of the mixed partial derivative, we need to evaluate the terms

$$E\left[\sum_{r=0}^{y_{ij}-1}\left\{\frac{(1-\theta_i)^2}{[(1-\theta_i)\pi + r\theta_i]^2}\right\}\right], \quad E\left[\sum_{r=0}^{n_{ij}-y_{ij}-1}\left\{\frac{(1-\theta_i)^2}{[(1-\theta_i)(1-\pi) + r\theta_i]^2}\right\}\right],$$

$$E\left[\sum_{r=0}^{y_{ij}-1}\left\{\frac{r}{\left[\left(1-\theta_i\right)\pi+r\theta_i\right]^2}\right\}\right],\quad E\left[\sum_{r=0}^{n_{ij}-y_{ij}-1}\left\{\frac{r}{\left[\left(1-\theta_i\right)\left(1-\pi\right)+r\theta_i\right]^2}\right\}\right],$$

$$E\left[\sum_{r=0}^{y_{ij}-1}\left\{\frac{\left(r-\pi\right)^2}{\left[\left(1-\theta_i\right)\pi+r\theta_i\right]^2}\right\}\right],\quad E\left[\sum_{r=0}^{n_{ij}-y_{ij}-1}\left\{\frac{\left(r+\pi-1\right)^2}{\left[\left(1-\theta_i\right)\left(1-\pi\right)+r\theta_i\right]^2}\right\}\right] \quad \text{and}$$

$$E\left[\sum_{r=0}^{n_{ij}-1}\left\{\frac{\left(r-1\right)^2}{\left[\left(1-\theta_i\right)+r\theta_i\right]^2}\right\}\right].$$

Again,

$$E\left[\sum_{r=0}^{y_{ij}-1}\left\{\frac{r}{\left[\left(1-\theta_i\right)\pi+r\theta_i\right]^2}\right\}\right]=E\left[\sum_{r=0}^{y_{ij}-1}\left\{\frac{\left(1-\theta_i\right)\left(r-\pi\right)}{\left[\left(1-\theta_i\right)\pi+r\theta_i\right]^2}+\frac{1}{\left[\left(1-\theta_i\right)\pi+r\theta_i\right]}\right\}\right]$$

and

$$E\left[\sum_{r=0}^{n_{ij}-y_{ij}-1}\left\{\frac{r}{\left[\left(1-\theta_i\right)\left(1-\pi\right)+r\theta_i\right]^2}\right\}\right]$$

$$=E\left[\sum_{r=0}^{n_{ij}-y_{ij}-1}\left\{\frac{\left(1-\theta_i\right)\left(r+\pi-1\right)}{\left[\left(1-\theta_i\right)\left(1-\pi\right)+r\theta_i\right]^2}+\frac{1}{\left[\left(1-\theta_i\right)\left(1-\pi\right)+r\theta_i\right]}\right\}\right].$$

Now,

$$E\left[\sum_{r=0}^{y_{ij}-1}\left\{\frac{1}{\left[(1-\theta_i)\,\pi+r\theta_i\right]^2}\right\}\right]=\sum_{y_{ij}=0}^{n_{ij}}\left[\sum_{r=0}^{y_{ij}-1}\left\{\frac{1}{\left[(1-\theta_i)\,\pi+r\theta_i\right]^2}\right\}\right]Pr\left(y_{ij}\right)$$

$$\begin{aligned}
&= 0+\\
&\frac{Pr\left(y_{ij}=1\right)}{\left[(1-\theta_i)\,\pi+0\theta_i\right]^2}+\\
&\frac{Pr\left(y_{ij}=2\right)}{\left[(1-\theta_i)\,\pi+0\theta_i\right]^2}+\frac{Pr\left(y_{ij}=2\right)}{\left[(1-\theta_i)\,\pi+1\theta_i\right]^2}+\\
&\frac{Pr\left(y_{ij}=3\right)}{\left[(1-\theta_i)\,\pi+0\theta_i\right]^2}+\frac{Pr\left(y_{ij}=3\right)}{\left[(1-\theta_i)\,\pi+1\theta_i\right]^2}+\frac{Pr\left(y_{ij}=3\right)}{\left[(1-\theta_i)\,\pi+2\theta_i\right]^2}\\
&\frac{Pr\left(y_{ij}=4\right)}{\left[(1-\theta_i)\,\pi+0\theta_i\right]^2}+\frac{Pr\left(y_{ij}=4\right)}{\left[(1-\theta_i)\,\pi+1\theta_i\right]^2}+\frac{Pr\left(y_{ij}=4\right)}{\left[(1-\theta_i)\,\pi+2\theta_i\right]^2}+\frac{Pr\left(y_{ij}=4\right)}{\left[(1-\theta_i)\,\pi+3\theta_i\right]^2}+\\
&\vdots\\
&+\frac{Pr\left(y_{ij}=n_{ij}\right)}{\left[(1-\theta_i)\,\pi+0\theta_i\right]^2}+\frac{Pr\left(y_{ij}=n_{ij}\right)}{\left[(1-\theta_i)\,\pi+1\theta_i\right]^2}+\cdots+\frac{Pr\left(y_{ij}=n_{ij}\right)}{\left[(1-\theta_i)\,\pi+\left(n_{ij}-1\right)\theta_i\right]^2}\\
&=\frac{Pr\left(y_{ij}\geq 1\right)}{\left[(1-\theta_i)\,\pi+(1-1)\,\theta_i\right]^2}+\frac{Pr\left(y_{ij}\geq 2\right)}{\left[(1-\theta_i)\,\pi+(2-1)\,\theta_i\right]^2}+\cdots+\frac{Pr\left(y_{ij}\geq n_{ij}\right)}{\left[(1-\theta_i)\,\pi+\left(n_{ij}-1\right)\theta_i\right]^2}.
\end{aligned}$$

$$E\left[\sum_{r=0}^{y_{ij}-1}\left\{\frac{1}{\left[(1-\theta_i)\,\pi+r\theta_i\right]^2}\right\}\right]=\sum_{r=1}^{n_{ij}}\frac{Pr\left(y_{ij}\geq r\right)}{\left[(1-\theta_i)\,\pi+(r-1)\,\theta_i\right]^2}.$$

$$E\left[\sum_{r=0}^{n_{ij}-y_{ij}-1}\left\{\frac{1}{\left[(1-\theta_i)(1-\pi)+r\theta_i\right]^2}\right\}\right]=\sum_{y_{ij}=0}^{n_{ij}}\left[\sum_{r=0}^{n_{ij}-y_{ij}-1}\left\{\frac{1}{\left[(1-\theta_i)(1-\pi)+r\theta_i\right]^2}\right\}\right]Pr\left(y_{ij}\right)$$

$$\begin{aligned}
&=\frac{Pr\left(y_{ij}=0\right)}{\left[(1-\theta_i)(1-\pi)+0\theta_i\right]^2}+\frac{Pr\left(y_{ij}=0\right)}{\left[(1-\theta_i)(1-\pi)+1\theta_i\right]^2}+\cdots+\frac{Pr\left(y_{ij}=0\right)}{\left[(1-\theta_i)(1-\pi)+(n_{ij}-1)\theta_i\right]^2}+\\
&\frac{Pr\left(y_{ij}=1\right)}{\left[(1-\theta_i)(1-\pi)+0\theta_i\right]^2}+\frac{Pr\left(y_{ij}=1\right)}{\left[(1-\theta_i)(1-\pi)+1\theta_i\right]^2}+\cdots+\frac{Pr\left(y_{ij}=1\right)}{\left[(1-\theta_i)(1-\pi)+(n_{ij}-2)\theta_i\right]^2}+\\
&\frac{Pr\left(y_{ij}=2\right)}{\left[(1-\theta_i)(1-\pi)+0\theta_i\right]^2}+\frac{Pr\left(y_{ij}=2\right)}{\left[(1-\theta_i)(1-\pi)+1\theta_i\right]^2}+\cdots+\frac{Pr\left(y_{ij}=2\right)}{\left[(1-\theta_i)(1-\pi)+(n_{ij}-3)\theta_i\right]^2}+\\
&\vdots\\
&+\frac{Pr\left(y_{ij}=n_{ij}-2\right)}{\left[(1-\theta_i)(1-\pi)+0\theta_i\right]^2}+\frac{Pr\left(y_{ij}=n_{ij}-2\right)}{\left[(1-\theta_i)(1-\pi)+1\theta_i\right]^2}+\\
&\frac{Pr\left(y_{ij}=n_{ij}-1\right)}{\left[(1-\theta_i)(1-\pi)+0\theta_i\right]^2}+0
\end{aligned}$$

$$=\frac{Pr\left(y_{ij}\le n_{ij}-1\right)}{\left[(1-\theta_i)\pi+(1-1)\theta_i\right]^2}+\frac{Pr\left(y_{ij}\le n_{ij}-2\right)}{\left[(1-\theta_i)\pi+(2-1)\theta_i\right]^2}+\cdots+\frac{Pr\left(y_{ij}\le 0\right)}{\left[(1-\theta_i)\pi+(n_{ij}-1)\theta_i\right]^2}.$$

$$E\left[\sum_{r=0}^{n_{ij}-y_{ij}-1}\left\{\frac{1}{\left[(1-\theta_i)(1-\pi)+r\theta_i\right]^2}\right\}\right]=\sum_{r=1}^{n_{ij}}\frac{Pr\left(y_{ij}\le n_{ij}-r\right)}{\left[(1-\theta_i)(1-\pi)+(r-1)\theta_i\right]^2}.$$

$$E\left[\sum_{r=0}^{y_{ij}-1}\left\{\frac{(r-\pi)}{\left[(1-\theta_i)\,\pi+r\theta_i\right]^2}\right\}\right]=\sum_{y_{ij}=0}^{n_{ij}}\left[\sum_{r=0}^{y_{ij}-1}\left\{\frac{(r-\pi)}{\left[(1-\theta_i)\,\pi+r\theta_i\right]^2}\right\}\right]Pr\left(y_{ij}\right)$$

$$\begin{aligned}
&= 0+\\
&\frac{(0-\pi)\,Pr\left(y_{ij}=1\right)}{\left[(1-\theta_i)\,\pi+0\theta_i\right]^2}+\\
&\frac{(0-\pi)\,Pr\left(y_{ij}=2\right)}{\left[(1-\theta_i)\,\pi+0\theta_i\right]^2}+\frac{(1-\pi)\,Pr\left(y_{ij}=2\right)}{\left[(1-\theta_i)\,\pi+1\theta_i\right]^2}+\\
&\frac{(0-\pi)\,Pr\left(y_{ij}=3\right)}{\left[(1-\theta_i)\,\pi+0\theta_i\right]^2}+\frac{(1-\pi)\,Pr\left(y_{ij}=3\right)}{\left[(1-\theta_i)\,\pi+1\theta_i\right]^2}+\frac{(2-\pi)\,Pr\left(y_{ij}=3\right)}{\left[(1-\theta_i)\,\pi+2\theta_i\right]^2}\\
&\frac{(0-\pi)\,Pr\left(y_{ij}=4\right)}{\left[(1-\theta_i)\,\pi+0\theta_i\right]^2}+\frac{(1-\pi)\,Pr\left(y_{ij}=4\right)}{\left[(1-\theta_i)\,\pi+1\theta_i\right]^2}+\frac{(2-\pi)\,Pr\left(y_{ij}=4\right)}{\left[(1-\theta_i)\,\pi+2\theta_i\right]^2}+\frac{(3-\pi)\,Pr\left(y_{ij}=4\right)}{\left[(1-\theta_i)\,\pi+3\theta_i\right]^2}+\\
&\vdots\\
&+\frac{(0-\pi)\,Pr\left(y_{ij}=n_{ij}\right)}{\left[(1-\theta_i)\,\pi+0\theta_i\right]^2}+\frac{(1-\pi)\,Pr\left(y_{ij}=n_{ij}\right)}{\left[(1-\theta_i)\,\pi+1\theta_i\right]^2}+\cdots+\frac{(n_{ij}-1-\pi)\,Pr\left(y_{ij}=n_{ij}\right)}{\left[(1-\theta_i)\,\pi+(n_{ij}-1)\,\theta_i\right]^2}\\
&=\frac{(0-\pi)\,Pr\left(y_{ij}\geq 1\right)}{\left[(1-\theta_i)\,\pi+(1-1)\,\theta_i\right]^2}+\frac{(1-\pi)\,Pr\left(y_{ij}\geq 2\right)}{\left[(1-\theta_i)\,\pi+(2-1)\,\theta_i\right]^2}+\cdots+\frac{(n_{ij}-1-\pi)\,Pr\left(y_{ij}\geq n_{ij}\right)}{\left[(1-\theta_i)\,\pi+(n_{ij}-1)\,\theta_i\right]^2}.
\end{aligned}$$

$$E\left[\sum_{r=0}^{y_{ij}-1}\left\{\frac{(r-\pi)}{\left[(1-\theta_i)\,\pi+r\theta_i\right]^2}\right\}\right]=\sum_{r=1}^{n_{ij}}\frac{(r-\pi-1)\,Pr\left(y_{ij}\geq r\right)}{\left[(1-\theta_i)\,\pi+(r-1)\,\theta_i\right]^2}.$$

Similarly

$$E\left[\sum_{r=0}^{n_{ij}-y_{ij}-1}\left\{\frac{(r+\pi-1)}{\left[(1-\theta_i)(1-\pi)+r\theta_i\right]^2}\right\}\right]=\sum_{r=1}^{n_{ij}}\frac{(r+\pi-2)\,Pr\left(y_{ij}\leq n_{ij}-r\right)}{\left[(1-\theta_i)(1-\pi)+(r-1)\,\theta_i\right]^2},$$

$$E\left[\sum_{r=0}^{y_{ij}-1}\left\{\frac{(r-\pi)^2}{\left[(1-\theta_i)\,\pi+r\theta_i\right]^2}\right\}\right]=\sum_{r=1}^{n_{ij}}\frac{(r-\pi-1)^2\,Pr\left(y_{ij}\geq r\right)}{\left[(1-\theta_i)\,\pi+(r-1)\,\theta_i\right]^2},$$

$$E\left[\sum_{r=0}^{n_{ij}-y_{ij}-1}\left\{\frac{(r+\pi-1)^2}{\left[(1-\theta_i)(1-\pi)+r\theta_i\right]^2}\right\}\right]=\sum_{r=1}^{n_{ij}}\frac{(r+\pi-2)^2\,Pr\left(y_{ij}\leq n_{ij}-r\right)}{\left[(1-\theta_i)(1-\pi)+(r-1)\,\theta_i\right]^2}$$

and

$$E\left[\sum_{r=0}^{n_{ij}-1}\left\{\frac{(r-1)^2}{\left[(1-\theta_i)+r\theta_i\right]^2}\right\}\right]=\sum_{r=1}^{n_{ij}}\frac{(r-2)^2}{\left[(1-\theta_i)+(r-1)\,\theta_i\right]^2}.$$

Bibliography

Altham, P. M. E. (1978). Two generalizations of the binomial distribution. *Applied Statistics*, 27:162–167.

Anscombe, F. J. (1949). The statistical analysis of insect counts based on the negative binomial distribution. *Biometrics*, 5:165–173.

Aspin, A. A. (1948). An examination and further development of a formula arising in the problem of comparing two mean values. *Biometrika*, 35:88–96.

Barnwal, R. K. (1989). Analysis of one-way layout of count data. Unpublished Ph.D. thesis.

Barnwal, R. K. and Paul, S. R. (1988). Analysis of one-way layout of count data with negative binomial variation. *Biometrika*, 75:215–222.

Bartlett, M. S. (1953a). Approximate confidence intervals. *Biometrika*, 40:12–19.

Bartlett, M. S. (1953b). Approximate confidence intervals.ii. more than one unknown parameter. *Biometrika*, 40:306 – 317.

Behrens, W. V. (1929). Ein beitrag zur fehlerberechung beiwenigen beobachtungen. *Landwirtsch, Jahrbcher*, 68:807–837.

Best, D. J. and Rayner, J. C. W. (1987). Welch's approximate solution for the Behrens-Fisher problem. *Technometrics*, 29:205–210.

Bliss, C. I. and Fisher, R. A. (1953). Fitting the negative binomial distribution to biological data. *Biometrics*, 9:176–200.

Bliss, C. I. and Owen, A. R. G. (1958). Negative binomial distribution with a common k. *Biometrika*, 45:37–58.

Breslow, N. E. (1984). Extra-Poisson variation in log-linear models. *Applied Statistics*, 33:38–44.

Breslow, N. E. (1990). Tests of hypotheses in over-dispersed Poisson regression and other quasi-likelihood models. *Journal of the American Statistical Association*, 85:565–571.

Carroll, K. J. (2003). On the use and utility of the Weibull model in the analysis of survival data. *Controlled Clinical Trials*, 24:682–701.

Chatfield, C. and Goodhart, G. J. (1970). The beta-binomial model for consumer purchasing behaviour. *Applied Statistics*, 19:240–250.

Chen, J. J. and Kodell, R. L. (1989). Quantitative risk assessment for teratological effects. *Journal of the American Statistical Association*, 85:565–571.

Clark, S. and Perry, J. (1989). Estimation of the negative binomial parameter κ by maximum quasi-likelihood. *Biometrics*, 45:309–316.

Cochran, W. G. and Cox, G. M. (1950). *Experiment designs.* John Wiley & Sons, Inc., New York.

Cohen, A. C. (1965). Maximum likelihood estimation in the Weibull distribution based on complete and censored samples. *Technometrics*, 5:579–588.

Cohen, A. C. and Whitten, B. (1982). Modified maximum likelihood and modified moment estimators for the three-parameter Weibull distribution. *Communications in Statistics-Theory and Methods*, 11:2631–2656.

Cohen, A. C., Whitten, B. J., and Ding, Y. (1984). Modified moment estimation for the three-parameter Weibull distribution. *Journal of Quality Technology*, 16:159–167.

Cordeiro, G. M. and Klein, R. (1994). Bias correction in ARMA models. *Statistics and Probability Letters*, 19:169–176.

Cox, D. R. and Hinkley, D. V. (1974). *Theoretical Statistics.* Chapman and Hall, London.

Cox, D. R. and Snell, E. J. (1968). A general definition of residuals (with discussion). *Journal of the Royal Statistical Society. Series B (Methodological)*, 30:248–275.

Cramer, H. (1946). *Mathematical Methods of Statistics.* Princeton University Press, USA.

Cran, G. W. (1988). Moment estimators for the 3-parameter Weibull distribution. *IEEE Transactions on Reliability*, 37:360–363.

Crowder, M. J. (1978). Beta-binomial ANOVA for proportions. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 27:34–37.

Crowder, M. J. (1982). On weighted least squares and some variants. Technical Report in Statistics 13, Surrey University.

Crowder, M. J. (1987). On linear and quadratic estimating functions. *Biometrika*, 74:591–597.

Davidian, M. and Carroll, R. J. (1987). Variance function estimation. *Journal of the American Statistical Association*, 82:1079–1091.

Dong, L. B. (2004). The Behrens-Fisher problem: an empirical likelihood approach. Econometrics Working Paper EWP 0404, University of Victoria.

Donovan, A. M., Ridout, M. S., and James, D. J. (1994). Assessment of somaclonal variation in apple *ii*. rooting ability and shoot proliferation. *Journal of Horticultural Science*, 69:115–122.

Elston, R. C. (1977). Response to query, consultants corner. *Biometrics*, 33:232–233.

Engel, J. (1984). Models for response data showing extra-Poisson variation. *Statistica Neerlandica*, 38:159–167.

Fisher, R. A. (1935). The fiducial argument in statistical inference. *Annals of Eugenics*, 6:391–398.

Fisher, R. A. (1941). The negative binomial distribution. *Annals of Eugenics*, 11:182–187.

Gail, M. H., Santner, T. J., and Brown, C. C. (1980). An analysis of comparative carcinogenesis experiments based on multiple times to tumor. *Biometrics*, 36:255–266.

Gibson, G. J. and Austin, E. J. (1996). Fitting and testing spatio-temporal stochastic models with applications in plant pathology. *Plant Pathology*, 45:172–184.

Godambe, V. P. and Thompson, M. E. (1989). An extension of quasi-likelihood estimation. *Journal of Statistical Planning and Inference*, 22:137–152.

Hammitt, J. K. (2004). Effects of disease type and latency on the value of mortality risk. *The Journal of Risk and Uncertainty*, 28:73–95.

Harter, H. L. and Moore, A. H. (1965). Maximum likelihood estimation of the parameters of the gamma and Weibull populations from censored samples. *Technometrics*, 7:639–643.

Haseman, J. K. and Kupper, L. L. (1979). Analysis of dichotomous response data from certain toxicological experiments. *Biometrics*, 35:281–293.

Hogg, R. V., McKean, J. W., and Craig, A. T. (2005). *Introduction to Mathematical Statistics*. Prentice Hall, New Jersey, Sixth edition.

Kish, L. (1965). *Survey Sampling*. John Wiley & Sons Inc, New York.

Kleinman, J. C. (1973). Proportions with extraneous variance: single and independent samples. *Journal of the American Statistical Association*, 8:46–54.

Ku, P. M., Anderson, E. L., and Carper, H. J. (1972). Some considerations in rolling fatigue evaluation. *American Society of Lubrication Engineers Transactions*, 15:113–129.

Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. John Wiley & Sons Inc, New York.

Lawless, J. F. (1987). Negative binomial and mixed Poisson regression. *The Canadian Journal of Statistics*, 15:209–225.

Lee, Y. and Nelder, J. A. (2001). Hierarchical generalized linear models: A synthesis of generalized linear models, random effect models and structured dispersions. *Biometrika*, 88:987–1006.

Lehman, E. L. (1999). *Elements of Large Sample Theory*. Springer, USA.

Lloyd, N. S. (1967). Weibull probability paper. *Industrial Quality Control*, 23:452–453.

Mann, N. R., Schafer, R. E., and Singpurwalla, N. D. (1974). *Methods for Statistical Analysis of Reliability and Lifetime Data*. John Wiley & Sons Inc, New York.

Manton, K. G., Woodbury, M. A., and Stallard, E. (1981). A variance components approach to categorical data models with heterogeneous cell populations: analysis of spatial gradients in lung cancer mortality rates in North Carolina counties. *Biometrics*, 37:259–269.

Margolin, B., Kaplan, N., and Zeiger, E. (1981). Statistical analysis of the ames salmonella/microsome test. *Proceedings of the National Academy of Sciences, U.S.A.*, 76:3779–3783.

Margolin, B. H., Kim, B. S., and Risko, K. J. (1989). The ames salmonella/microsome mutagenicityassay: Issues of inference and validation. *Journal of the American Statistical Association*, 84:651–661.

McCaughran, D. A. and Arnold, D. W. (1976). Statistical models for members of implantation sites and embryonic deaths in mice. *Toxicology and Applied Pharmacology*, 38:325–333.

McCool, J. I. (1979). Analysis of single classification experiments based on censored samples from the two-parameter Weibull distribution. *Journal of Statistical Planning and Inference*, 3:39–68.

McCool, J. I. (1980). Confidence limits for Weibull regression with censored data. *IEEE Transactions on Reliability*, R29:145–150.

McCool, J. I. (1982). Analysis of sets of two-parameter weibull data arising in rolling contact endurance testing. volume ASTM, STP 771, pages 293–319, Baltimore, MD. American Society for Testing and Materials.

McCool, J. I. (1998). Inference on the Weibull location parameter. *Journal of Quality Technology*, 30:119–126.

McCullagh, P. (1983). Quasi-likelihood functions. *Annals of Statistics*, 11:59–69.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*. Chapman and Hall, London.

Moore, D. F. (1985). *Methods of moments estimators for over-dispersed counts and proportions*. Ph.D. thesis, University of Washington, Seattle.

Moore, D. F. (1986). Asymptotic properties of moments estimators for over-dispersed counts and proportions. *Biometrika*, 73:583–588.

Moore, D. F. and Tsiatis, A. (1991). Robust estimation of the variance in moment methods for extra-binomial and extra-Poisson variation. *Biometrics*, 47:383–401.

Moran, P. A. P. (1970). On asymptotically optimal tests of composite hypotheses. *Biometrika*, 57:45–75.

Nelder, J. A. and Pregibon, D. (1987). An extended quasi-likelihood function. *Biometrika*, 74:221–232.

Nelson, W. (1982). *Applied Life Data Analysis*. John Wiley & Sons Inc, New York.

Neyman, J. (1959). *Optimal asymptotic tests for composite hypotheses*. In Probability and Statistics. John Wiley & Sons Inc, New York.

Neyman, J. and Scott, E. L. (1966). On the use of $C(\alpha)$ optimal tests of composite hypotheses. *Bulletin of the International Statistical Institute*, 41:477–497.

Otake, M. and Prentice, R. L. (1984). The analysis of chromosomally aberrant cells based on beta-binomial distribution. *Radiation Research*, 98:456–470.

Pack, S. E. (1986a). *The analysis of proportion from toxicological experiments*. Ph.D. thesis, University of Kent, Canterbury.

Pack, S. E. (1986b). Hypothesis testing for proportions with over-dispersion. *Biometrics*, 42:967–972.

Paul, S. R. (1982). Analysis of proportions of affected foetuses in teratological experiments. *Biometrics*, 38:361–370.

Paul, S. R. (1992). Comment on Best and Rayner(1987). *Technometrics, Letters to the editor*, 34:249–250.

Paul, S. R. and Banerjee, T. (1998). Analysis of two-way layout of count data involving multiple counts in each cell. *Journal of the American Statistical Association*, 93:1419–1429.

Paul, S. R. and Islam, A. S. (1995). Analysis of proportions in the presence of over-/under-dispersion. *Biometrics*, 51:1400–1410.

Paul, S. R. and Islam, A. S. (1998). Joint estimation of the mean and dispersion parameters in the analysis of proportions: a comparison of efficiency and bias. *The Canadian Journal of Statistics*, 26:83–94.

Paul, S. R. and Plackett, R. L. (1978). Inference sensitivity for Poisson mixtures. *Biometrika*, 65:591–602.

Piegorsch, W. W. (1990). Maximum likelihood estimation for the negative binomial dispersion parameter. *Biometrics*, 46:863–867.

Prentice, R. L. (1986). Binary regression using an extended beta-binomial distribution, with discussion of correlation induced by covariate measurement errors. *Journal of the American Statistical Association*, 81:321–327.

Qin, J. (1991). Likelihood and empirical likelihood ratio confidence intervals in two sample semi-parametric models. Technical Report Series Stat-91-6, University of Waterloo.

Rao, C. R. (1948). Large sample tests of statistical hypotheses concerning several parameters with application to problems of estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 44:50–57.

Rao, J. N. K. and Scott, A. J. (1992). A simple method for the analysis of clustered binary data. *Biometrics*, 48:577–585.

Ridout, M. S., Demetrio, C. G. B., and Firth, D. (1999). Estimating intraclass correlation for binary data. *Biometrics*, 55:137–148.

Ross, G. J. S. and Preece, D. A. (1985). The negative binomial distribution. *The Statistician*, 34:323–336.

Saha, K. K. (2011). Interval estimation of the over-dispersion parameter in the analysis of one-way layout of count data. *Statistics in Medicine*, 30:39–51.

Schafer, R. E. and Sheffield, T. S. (1976). On procedures for comparing two Weibull populations. *Technometrics*, 18:231–235.

Scheffé, H. (1943). On solutions of the Behrens-Fisher problem based on the t-distribution. *Annals of Mathematical Statistics*, 14:35–44.

Sirvanci, M. and Yang, G. (1984). Estimation of the Weibull parameters under type *I* censoring. *Journal of the American Statistical Association*, 79:183–187.

Skellam, J. G. (1948). A probability distribution derived from the binomial distribution by regarding the probability of success as a variable between the sets of trials. *Journal of the Royal Statistical Society, Series B (Methodological)*, 10:257–261.

Srivastava, M. S. and Wu, Y. (1993). Local efficiency of moment estimators in beta-binomial model. *Communications in Statistics: Theory and Methods*, 22:2471–2490.

Tarone, R. E. (1985). On heterogeneity tests based on efficient scores. *Biometrika*, 72:91–95.

Teimouri, M. and Gupta, A. K. (2013). On the three-parameter Weibull distribution shape parameter estimation. *Journal of Data Science*, 11:403–414.

Thiagarajah, K. (1992). *Inference procedures in some lifetime models*. Ph.D. thesis, University of Windsor, Canada.

Thoman, D. R. and Bain, L. J. (1969). Two sample tests in the Weibull distribution. *Technometrics*, 11:805–815.

Tsui, K.-W. and Tang, S. (2005). Distributional property of the generalized p-value for the Behrens-Fisher problem with applications to multiple testing. Technical Report Report No. 1111, University of Wisconsin, Madison.

Uusipaikka, E. (2009). *Confidence intervals in generalized regression models*. Chapman & Hall/CRC:, USA.

Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models and the Gauss-Newton method. *Biometrika*, 61:439–447.

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. *Food and Cosmetics Toxicology*, 8:177–182.

Welch, B. L. (1937). The significance of the difference between two means when the population variances are unequal. *Biometrika*, 29:350–362.

Welch, B. L. (1947). The generalization of "Student's" problem when several different population variances are involved. *Biometrika*, 34:28–35.

Williams, D. A. (1975). The analysis of binary responses from toxicological experiments involving reproduction and teratogenicity. *Biometrics*, 31:949–952.

Williams, D. A. (1982). Extra-binomial variation in logistic linear models. *Applied Statistics*, 31:144–148.

Vita Auctoris

The author was born in Faridpur, Bangladesh in 1969. He obtained his B.Sc. (Honours) and M.Sc. in Statistics from Jahangirnagar University, Bangladesh in 1991 and 1992 respectively. In 1996 he joined as a Lecturer in the Department of Statistics, Shahjalal University of Science and Technology, Sylhet, Bangladesh and worked there before coming to the University of Windsor in 2005 for pursuing his Master's Degree. The author obtained his M.Sc. in Statistics from University of Windsor, Canada in 2007. He is currently a candidate for the Doctor of Philosophy in Statistics at the University of Windsor and will graduate in 2014.