5-2016

Adjusted Empirical Likelihood Method for Comparison of Treatment Effects in Linear Model Setting

Xi Kang
Montclair State University

MONTCLAIR STATE UNIVERSITY

Adjusted Empirical Likelihood Method for Comparison of Treatment Effects in a Linear Model Setting

by

Xi Kang

A Master's Thesis Submitted to the Faculty of

Montclair State University

In Partial Fulfillment of the Requirements

For the Degree of

Master of Statistics

May 2016

College/School The College of Science and Mathematics

Department Mathematical Sciences

Thesis Committee:

[redacted]

Dr. Haiyan Su

[redacted]

Dr. Andrew McDougall

[redacted]

Dr. Andrada Ivanescu
Committee Member

Abstract

Empirical likelihood is a nonparametric method of statistical inference which was introduced by Owen. It allows the data analyst to use it without making distribution assumptions. Empirical likelihood method has been widely used not only for nonparametric models but also for semi-parametric models, with the effectiveness of the likelihood approach and good power properties. However, when the sample size is small or the dimension is high, the method is poorly calibrated, producing tests that generally have a higher type I error. In addition, it suffers from a limiting convex hull constraint. Many statisticians have proposed methods to address the performance. We explore the method proposed by Chen which makes an adjustment on empirical likelihood method. This thesis derives an adjusted empirical likelihood-based method for comparing two treatment effects in a linear model setting. We use the adjusted empirical likelihood-based method to make inference for the difference by comparing the parameters in two linear models. Our method is free of the assumptions of normally distributed and homogeneous errors, and equal sample size. In addition, the adjusted empirical likelihood method is Bartlett correctable. We apply the Bartlett correction procedure to further improve the coverage of our proposed method. Simulation experimental are used to illustrate that our method outperforms the published ones and also empirical likelihood-based method. This method can be extended into multiple treatment effects comparison.

ADJUSTED EMPIRICAL LIKELIHOOD METHOD FOR COMPARISON OF TREATMENT EFFECTS IN A LINEAR MODEL SETTING

A THESIS

Submitted in partial fulfillment of the requirements

For the degree of Master of Statistics

by

XI KANG

Montclair State University

Montclair, New Jersey

May, 2016

Acknowledgments

I would like to express my deepest appreciation to my thesis sponsor, Dr. Haiyan Su, for her support throughout this study. Without her guidance and persistent help, this thesis would not have been possible.

I would like to thank my committee members, Dr. Andrew McDougall and Dr. Andrada Ivanescu, for many helpful feedback and encouragement they have given to me during the thesis study.

Finally, I would like to thank my friends and family for their continuous support and love. None of this would have been possible without them.

Contents

1 Introduction

1.1 What is empirical likelihood?

Likelihood methods are very effective and flexible. They can be used to find efficient estimators, to construct tests with good power properties and to offset or even correct for data problems such as incompletely observed, distorted, or sampled with a bias. Likelihood can be used to pool information from different data sources. In parametric likelihood methods, we assume that the joint distribution of all available data has a known form, apart from one or more unknown quantities. But a problem with parametric likelihood inferences is that we might not know which parametric family to use. And there is no reason to suppose that a newly encountered data set belongs to any of the well studied parametric families. This misspecification may cause likelihood-based estimates to be inefficient, even the corresponding confidence intervals and test may have high type-I error and low power.

Many statisticians have chosen nonparametric inferences to avoid specifying a parametric family for the data. These methods include the jackknife, the infinitesimal jackknife, several versions of the bootstrap, and especially, empirical likelihood. These nonparametric methods give confidence intervals and tests with validity not depending on strong distributional assumptions. The use of nonparametric methods is in line with John Tukey's quote "It is better to be approximately right, than exactly wrong". But when we contemplate replacing a parametric method by a nonparametric one, we need to consider that sometimes the improved generality comes at a cost of reduced power. Among these methods mentioned above, empirical likelihood arises because it combines the reliability of the nonparametric methods with the flexibility and effectiveness of the likelihood approach. It can be combined effectively with bootstrap and as well as parametric on some problems.

Empirical likelihood(EL) is a nonparametric method of statistical inference which was introduced by Owen (1988,1990). He extended earlier work of Thomas and Grunkemeier (1975) who employed a nonparametric likelihood ratio idea to construct confidence intervals for the survival function. That work in turn builds on nonparametric maximum likelihood estimation which has a long history in survival analysis. Owen proposed this method for the univariate mean and some other statistics such as for multivariate mean, for quantiles, for Kernel smooths, for right censoring as well as others. For example, given a random variable $X_i \sim F(x, \theta)$ with a parameter $\theta \in \Omega$, let ω_i be the weight that F places on observation X_i, then an empirical likelihood ratio for testing the null hypothesis $H_0 : \mu_0(F) = \mu$, where $\mu_0(F)$ is the expectation with respect to the distribution F, can be written as follows:

$$R(\mu) = max\left\{\prod_{i=1}^{n} n\omega_i \,\middle|\, \sum_{i=1}^{n} \omega_i X_i = \mu, \omega_i \geq 0, \sum_{i=1}^{n} \omega_i = 1\right\}. \tag{1.1}$$

According to his univariate empirical likelihood theorem (Univariate ELT), if $0 < Var(X_i) < \infty$, then $-2\log(R(\mu_0))$ converges in distribution to χ_1^2 as $n \to \infty$, where $\mu_0 = E(X_i)$.

EL is a data-driven technique with the advantage of automatically determining shape of confidence region. We use the following example with data from Larsen and Marx (1986) by Owen (1990) as an illustration. Eleven male ducks, each a second generation cross between mallard and pintail, were examined. Their plumage was rated on a scale from 0 (completely mallardlike) to 20 (completely pintaillike) and their behavior was similarly rated on a scale from 0 (mallard) to 15 (pintail). Figure 1 shows the data, together with nested empirical likelihood confidence contours for the mean. The confidence contours are presented for nominal confidence levels: 0.50, 0.90, 0.95, 0.99, taken from 20/9 times the $F_{2,9}$ distribution. An asterisk marks the sample mean. Figure 2 shows the same data with the contours taken from a scaled $F_{2,9}$ distribution for Hotelling's T^2 statistic. These are parametric likelihood ratio contour assuming a bivariate normal distribution with unknown mean and variance.

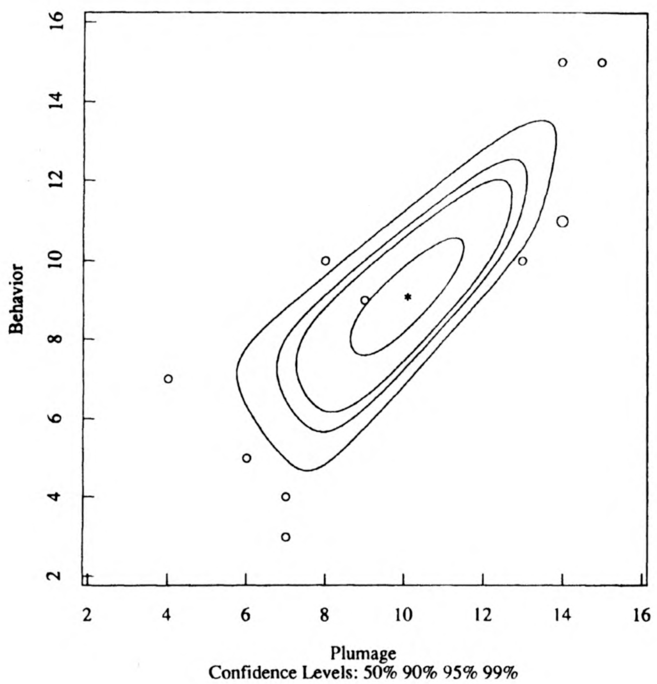

Figure 1: Empirical likelihood contours

Suppose we have $x_1, x_2, ..., x_n$ as a random sample from a nonparametric population $F(x)$ such that $x \in R^m$ with dimension m. The problem of interest is inference on the p-dimensional parameter $\theta = \theta(F)$. Assume that the general estimating equations is defined by

$$E(g(X; \theta)) = 0 \tag{1.2}$$

for a q-dimensional estimating function g and a p-dimensional parameter θ. Let p_i be the probability that distribution function F assigned to each point x_i, then the

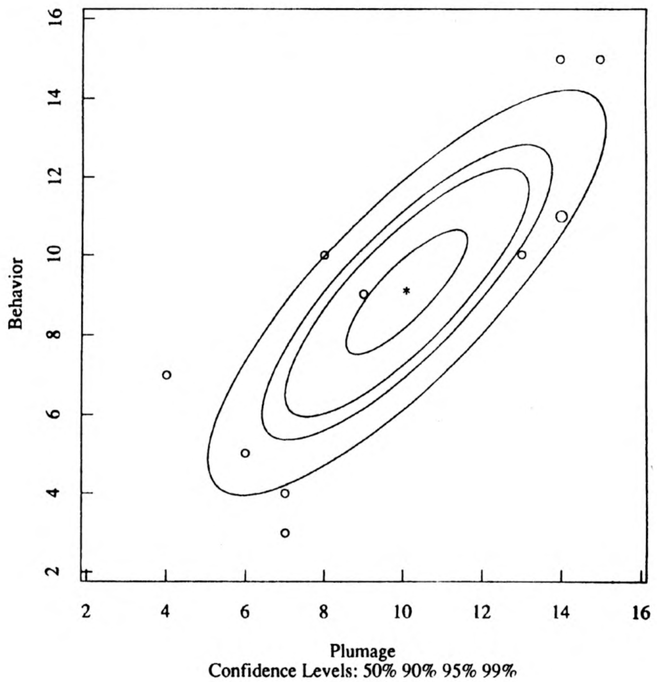

Figure 2: Normal likelihood contours

empirical likelihood function of θ is defined as

$$L_n(\theta) = \sup\left\{ \prod_{i=1}^{n} p_i : p_i \geq 0, \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i g(x_i;\theta) = 0 \right\}. \tag{1.3}$$

And the empirical log-likelihood ratio function is defined to be

$$R_n(\theta) = -2\log(n^n \times L_n(\theta)). \tag{1.4}$$

It converges to a chi-square distribution with q degrees of freedom (Owen 2001, and Qin and Lawless 1994).

Empirical likelihood method has many nice statistical properties. It allows the data analyst to use likelihood methods, without having to assume that the data come from a known family of distributions, as long as it is independent and identically distributed. In particular, EL method does not involve asymptotic variance estimation, which may be complicated in nonparametric or semi-parametric models. The confidence region has a data-determined shape, thereby better reflecting the true shape of the underlying distribution. It performs well even when the distribution is asymmetric or censored. EL methods are also useful since they can easily incorporate constraints and prior information. This method provides a versatile approach that may be applied to perform inference for a wide variety of functionals of interest. It has been employed in a number of different areas of statistics. A brief examination of the literature on empirical likelihood produces applications including quantiles of weakly dependent processes by Chen and Wong (1993), inference in missing data problems by Qin and Zhang (2007), estimation of variogram model parameters by Nordman and Caragea (2008), empirical likelihood for regression by Chen and Keilegom (2009), empirical likelihood for comparison of treatment effects by Su and Liang (2009). Many recent applications of EL in a variety of situations such as construction of simultaneous confidence band for right censored data, regression analysis, weighted EL, can be found in Owen (2001), Wang and Rao (2001, 2002), McKeague and Zhao (2002, 2006), Li and Wang (2003), Zhao (2005, 2010), Glenn and Zhao (2007), Zhao and Chen (2008). The flexibility and effectiveness of the empirical likelihood approach, as well as its relationship to many standard parametric procedures, make it a useful and interesting tool for many problems.

1.2 What is adjusted empirical likelihood?

The advantages of the EL methods over normal approximation (NA) based method have been demonstrated in the Chapter 1.1. However, when the sample size is small, or the dimension of the accompanying estimating function is high, the coverage probabilities of the EL confidence regions are often lower than the nominal value (under-coverage problem; DiCiccio, Hall, and Romano (1991); Owen (2001); Chen and Liu (2010)). Computing a profile empirical likelihood function involves constrained maximization and it is a key step in applications of empirical likelihood. Yet in some situations, there is no solution to the required numerical problem. In this case, the convention is to assign a zero value to the profile empirical likelihood. This strategy has at least two limitations. First, it is numerically difficult to determine that there is no solution; secondly, no information is provided on the relative plausibility of the parameter values where the likelihood is set to zero.

The adjusted empirical likelihood (AEL) proposed by Chen and Variyath and Abrasham (2008) tackles the low precision of the chi-square approximation with small sample size and also the empty set problems simultaneously. They proposed a novel adjustment to the empirical likelihood that retains the optimality properties, and guarantees a sensible value of the likelihood at any parameter value. The adjusted empirical likelihood is obtained by adding a pseudo-observation into the data set. This approach offers several key benefits in both ease of computation and accuracy. Its principal utility is to overcome the difficulty arising when the estimating equations have no solution; a solution is required in the EL approach. By using a conventional level of adjustment, Chen, Variyath and Abrasham found the AEL improves the approximation precision of the chi-square limiting distribution.

Figure 3 is a simple example to illustrate the convex hull problem and the adjustment from Chen et al. (2008). There are 50 observations generated from an independent bivariate standard normal distribution. They compute the profile likelihood at $(\mu_1, \mu_2) = (2, 2)$. The left side of Figure 3 gives the plot of g values and it is seen that the convex hull does not contain 0. The right side of Figure 3 gives the plot of g values with an artificial observation $g_{n+1} = -a_n \bar{g}_n$, where $a_n = \log(n)/2$. The convex hull is expanded and 0 is an interior points.

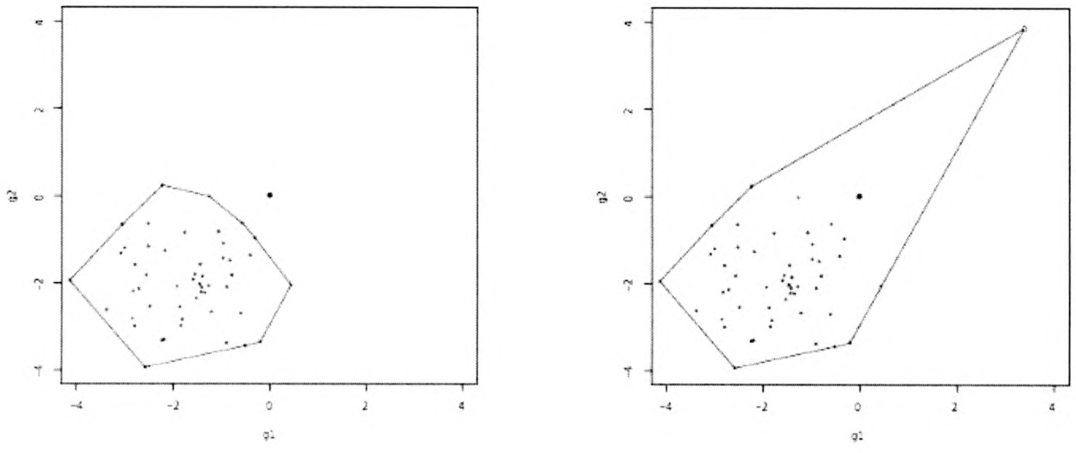

Figure 3: Convex hull (left) and adjusted convex hull with $a_n = log(n)/2$ (right). The bold dot is (0,0).

Many statisticians looked into the level of adjustment to empirical likelihood. Emerson and Owen (2009) discussed the level of adjustment for inference on multivariate population mean. Their simulation studies show that the AEL has better precision, and especially under linear and asset-pricing models. Chen and Liu (2010) showed that with a specific level of adjustment, the adjusted empirical likelihood achieves the high-order precision of the Bartlett correction. In addition, their simulation results indicated that the confidence regions by the adjusted empirical likelihood have the comparable coverage probabilities or substantially more accurate than the original empirical likelihood enhanced by the Bartlett correction. Wang, Chen and Pu (2015) showed that the general AEL is Bartlett-correctable and proposed a two-stage procedure for constructing accurate confidence regions.

1.3 Motivation

In clinical trials, related medical studies and biomedical studies, physicians and medical researchers are often interested in evaluating the difference between two treatments in order to justify the effect of a new medicine or a new cure. Statistical analysis usually provides important reference to the quantitative evaluation of medical advantages of one treatment over another. Many methods have been proposed to evaluate the difference in special cases. For example, Behrens-Fisher problems (1974) is a powerful and popular tool to study the difference between the means of two independent and normally distributed populations. However, in some situations, comparison of the response means from two populations ignores the fact that two populations may not be identical and may not normally distributed. Measuring a treatment effect may need to take into account the effect of other covariates. Bhuyan and Majumder (1996) gave another simple example. Such concern gives rise to a comparison of coefficients in linear regression models. Comparisons of treatment effects in linear regressions are quite popular since the comparison controls other covariates through the regression model. In Su and Liang's (2009) paper, they proposed an empirical likelihood-based method for comparing treatment effects by testing equality of coefficients in linear models. This method shows advantages in terms of power over other methods such as the normal approximation-based method, Weerahandi test, Dupont and Plummer test. The advantages of AEL method in chapter 1.2 triggered our research interest to focus on the development of improving the coverage probability to Su and Liang's (2009) research by using adjusted empirical likelihood-based method to compare treatment effects. We also planned to use Bartlett correction to improve AEL method.

This thesis is organized as follows. In Chapter 2, we review the treatments comparison using empirical likelihood-based test and derived the adjusted empirical likelihood method for the treatments comparison. Chapter 3 reports the results of simulation experiments. Chapter 4 presents the results of the proposed method on a drug study. Chapter 5 discusses the improvement and conclusion. The proof of the theoretical result is given in the Appendix.

2 Treatment Effects Comparison Test

2.1 Background

The difference between two treatments can be described as the difference of the parameters in two linear models as shown in Su and Liang (2009). Many methods have been proposed to evaluate the equality of linear regression models. A pioneering work on testing linear regression equality was done by Chow (1960), in which he proposed a statistical ratio to conduct the hypothesis test where the difference between the sum of residual squares assuming equality as numerator and the sum of residual squares without assuming equality as denominator. Under the null hypothesis, the resulting ratio was shown to follow an F distribution. However, the significance level of the test is considerably affected by even moderate heteroscedasticity when both sample sizes are small. Since then, a number of authors have been proposed the various versions of the Chow test. Schmidt and Sickles (1977) provided a formula to calculate the exact tail probability of the Chow test under the known ratio of two variances. Ali and Silver (1985) proposed two relatively robust tests on the basis of the Chow test and likelihood ratio statistics. However, the distribution of their tests need to be approximated using the moments of statistics under the null hypothesis since unknown distributions. Cornerly and Mansfield (1998) presented an approximate test. Their method provides an alternative for comparing heteroscedastic regression models. They replaced the pooled residual variance in the denominators with a weighted average of the residual variances from each group in Chow statistic. Dupont and Plummer (1998) developed an intuitive test by comparing slopes of two linear regressions to calculate sample size and power function, which has been applied to clinical trials, and qualified its applicability when the two error terms have the same variance. Yang and Zhao (2007) proposed a test of treatment effect using weighted log rank tests with empirical likelihood method.

All the tests mentioned above were derived from the likelihood principle and therefore need assumptions on distributions. The results of these tests were obtained subject to additional information about the ratio of the variances, that is, either the ratio is known, or the magnitude of the variance is of the same order. These assumptions are not always satisfied or at least need to be diagnosed. Su and Liang (2009) proposed an empirical likelihood-based method to make inference for the difference. Their test is free of these assumptions on the basis of the empirical likelihood principle and is shown to perform better than other normal-based tests. In this chapter, we will review the empirical likelihood-based test and derive the adjusted empirical likelihood-based test. Since Su and Liang (2009) already proved that EL is better performed than other methods such as empirical t-test, Weerahandi test and Dupont and Plummer test, we will just compare the AEL method with EL method.

2.2 Empirical likelihood-based test

Suppose that we observed samples of independent observations from the models

$$\begin{cases} y_1 = x_1^T \beta_1 + \varepsilon_1, \\ y_2 = x_2^T \beta_2 + \varepsilon_2, \end{cases} \tag{2.1}$$

where ε_1 and ε_2 are two independent random errors with a mean of zero and variances σ_1^2 and σ_2^2 respectively, and β_1 and β_2 are two unknown parameter vectors of length p, which indicate the treatment effects. We were interested in testing the hypothesis $H_0 : \beta_1 = \beta_2$ against $H_1 : \beta_1 \neq \beta_2$.

Let $(x_{11}, y_{11}), ..., (x_{1n_1}, y_{1n_1}), (x_{21}, y_{21}), ..., (x_{2n_2}, y_{2n_2})$ be the independent samples from model (2.1), where each y_i is regarded as the response of variable x_i. Let X_1

$= (x_{11}, ..., x_{1n_1})^T$, $X_2 = (x_{21}, ..., x_{2n_2})^T$, and similarly for Y_1 and Y_2, ε_1 and ε_2, then model (2.1) can be written in the form of

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}. \tag{2.2}$$

Let $Y = (Y_1^T, Y_2^T)^T$, X be the block diagonal matrix diag(X_1, X_2), $\beta = (\beta_1^T, \beta_2^T)^T$, and $\varepsilon = (\varepsilon_1^T, \varepsilon_2^T)^T$. Model (2.2) can futher be expressed as

$$Y = X_1^* \beta_1^* + X_2^* \beta_2 + \varepsilon, \tag{2.3}$$

where $X_1^* = (X_1^T, 0)^T$, $X_2^* = (X_1^T, X_2^T)^T$, and $\beta_1^* = \beta_1$ - β_2.

Therefore, null hypothesis H_0 is equivalent to $\beta_1^* = 0$. We now treat β_2 as a nuisance parameter. Our estimation method, motivated by partial regression plots, is to form partial residual vectors adjusting the influence of X_2^*, that is, Y - $E(Y|X_2^*)$ and X_1^* - $E(X_1^*|X_2^*)$. It follows from (2.3) that

$$\{Y - E(Y|X_2^*)\} = \{X_1^* - E(X_1^*|X_2^*)\}\beta_1^* + \varepsilon. \tag{2.4}$$

If $E(Y|X_2^*)$ and $E(X_1^*|X_2^*)$ are known, we can define estimates of β_1^* in a standard way. We denote $\tilde{S} = S$ - $E(S|X_2^*)$ and $\hat{S} = S$ - $\hat{E}(S|X_2^*)$ for any vector S. So, $\tilde{Y} =$ Y - $E(Y|X_2^*)$ and $\hat{Y} = Y$ - $\hat{E}(Y|X_2^*)$.

Let $\hat{E}(Y|X_2^*)$ and $\hat{E}(X_1^*|X_2^*)$ be the least square estimates of $E(Y|X_2^*)$ and $E(X_1^*|X_2^*)$ respectively. We may estimate β_1^* by solving the estimating equation

$$\frac{1}{n}\sum_{i=1}^{n} \hat{X}_{1i}^{*T}(\hat{Y}_i - \hat{X}_{1i}^*\beta_1^*) = 0$$

as $E[\tilde{X}_1^{*T}(\tilde{Y} - \tilde{X}_1^*\beta_1^*)] = 0$. This statement leads to the definition of Su and Liang's (2009) empirical likelihood ratio as follows.

Let F be the distribution function which assigns probability p_i at points $(Y_i, X_{1i}^*, X_{2i}^*)$. The empirical likelihood ratio function for β_1^* is therefore defined as

$$R_n(\beta_1^*) = \sup\left\{\prod_{i=1}^n np_i \middle| \sum_{i=1}^n p_i \hat{X}_{1i}^{*T}(\hat{Y}_i - \hat{X}_{1i}^*\beta_1^*) = 0, p_i \geq 0, \sum_{i=1}^n p_i = 1\right\}. \quad (2.5)$$

2.3 Our proposed adjusted empirical likelihood-based test

For any given θ , the likelihood ratio function $R_n(\theta)$ is well defined only if the convex hull of

$$\{g(x_i;\theta) : i = 1, 2, ..., n\} \quad (2.6)$$

contains the q-dimentional vector 0. When n is not large, or when a good candidate value of θ is not available, this convex hull often fails to contain 0. Blindly setting $L_n(\theta) = 0$ as suggested in the literature fails to provide information on whether θ is grossly unfit to the data or is in fact only slightly off an appropriate value.

Chen, Variyath and Abraham (2008) proposed adjusted empirical likelihood to the above issue by adding a pseudo observation. Let $g_i = g(x_i;\theta), i = 1, ..., n$, and

$$g_{n+1} = -a_n\bar{g}_n = -a_n n^{-1}\sum_{i=1}^n g_i \quad (2.7)$$

for some $a_n > 0$. They recommend to take $a_n = \log(n)/2$. The adjusted empirical likelihood is then defined as

$$L_n(\theta; a_n) = \sup\left\{\prod_{i=1}^{n+1} p_i : p_i \geq 0, \sum_{i=1}^{n+1} p_i = 1, \sum_{i=1}^{n+1} p_i g_i = 0\right\}. \quad (2.8)$$

and the adjusted empirical likelihood ratio function as

$$R_n(\theta; a_n) = -2\log{(n+1)^{n+1}} L_n(\theta; a_n).$$

Applying the adjusted empirical likelihood method in the treatment comparison case described in Chapter 2.2, the empirical likelihood ratio function for β_1^* is therefore defined as

$$R_n(\beta_1^*) = \sup\left\{\prod_{i=1}^{n+1}(n+1)p_i \middle| \sum_{i=1}^{n+1} p_i g_i = 0, \sum_{i=1}^{n+1} p_i = 1, p_i \geq 0\right\}, \tag{2.9}$$

where $g_i(\beta) = \hat{X}_{1i}^{*T}(\hat{Y}_i - \hat{X}_{1i}^{*}\beta_1^*)$ for $i = 1, 2, ..., n$ and $g_{n+1}(\beta) = -a_n n^{-1}\sum_{i=1}^{n} g_i = -a_n n^{-1}\hat{X}_{1i}^{*T}(\hat{Y}_i - \hat{X}_{1i}^{*}\beta_1^*) = \frac{-\log(n)}{2n}\sum_{i=1}^{n} g_i = -a_n n^{-1}\hat{X}_{1i}^{*T}(\hat{Y}_i - \hat{X}_{1i}^{*}\beta_1^*)$ for $i = n+1$. Then the model (2.9) can be rewritten as

$$R_n(\beta_1^*) = \sup\left\{\prod_{i=1}^{n+1}(n+1)p_i \middle| \right.$$

$$\left.\sum_{i=1}^{n} p_i \hat{X}_{1i}^{*T}(\hat{Y}_i - \hat{X}_{1i}^{*}\beta_1^*) - p_{n+1}\frac{\log(n)}{2n}\sum_{i=1}^{n}\hat{X}_{1i}^{*T}(\hat{Y}_i - \hat{X}_{1i}^{*}\beta_1^*) = 0, \sum_{i=1}^{n+1} p_i = 1, p_i \geq 0\right\}. \tag{2.10}$$

Theorem1. Assuming that $E(\|X_k\|^4) < \infty$, $E(\|X_k\|^2 Y_k^2) < \infty$ and $E(X_k^T X_k)$ are nonsingular for $k = 1, 2$, then $-2\log R_n(\beta_1^*)$ converges in distribution to a chi-squared distribution with p degrees of freedom.

Based on Theorem 1, we can obtain an estimate of β_1^* and the associated $100(1-\alpha)\%$ confidence region:

$$\{\beta_1^* : -2\log\{R_n(\beta_1^*)\} \leqslant c_\alpha\}, \tag{2.11}$$

where c_α is the $1-\alpha$ quantile of the χ_p^2 distribution satisfying $P(\chi_p^2 < c_\alpha) = 1-\alpha$. The proof of theorem 1 is given in the appendix.

2.4 Bartlett correction on EL and AEL

Bartlett (1937) pioneered the correction to the likelihood ratio statistic in the context of comparing the variances of several populations. For regular problems, Lawley (1956), through a heroic series of calculations, obtained a general formula for the null expected value of likelihood ratio and demonstrated that all cumulants of the Bartlett-corrected statistic for testing a composite hypothesis agree with those of the reference χ^2 distribution with error of order $n^{-3/2}$. Alternative expressions for the Bartlett corrections were developed by DiCiccio and Stern (1993), McCullagh and Cox (1986), and Skovgaard (2001).

The empirical likelihood confidence regions have data-driven shape and are Bartlett-correctable (DiCicco, Hall and Romano (1991)). To improve the precision of the coverage probability when sample size is not large, we may replace chi-square distribution by bootstrap calibration or by high-order approximation via the Bartlett correction (Chen and Cui (2006,2007)).

Chen and Liu (2010) proved that AEL is also Bartlett-correctable. With the Bartlett correction factor B_c, we can obtain an approximate $100(1-\alpha)$ % EL confidence region of θ

$$\mathcal{I}_{B_cEL}(\theta) = \{\theta : W_n(\theta) \leqslant c_\alpha\}, \tag{2.12}$$

where $W_n(\theta) = R_n(\theta) - inf_\theta R_n(\theta)$, $R_n(\theta)$ is defined in equation (1.4), and c_α is the $(1-\alpha)$th quantile of χ^2_p distribution.

In other words, by applying the Bartlett correction into formula (2.12), we have:

$$Pr(-2\log\{R_n(\beta_1^*)\} < c_\alpha(1+n^{-1}B_c)) = 1-\alpha+O(n^{-b}), b \in (1,2].$$

They specified the level of adjustment and proposed estimation of the Bartlett correction factor B_c as

$$B_c = \frac{\alpha_4}{2\alpha_2^2} - \frac{\alpha_3^2}{3\alpha_2^3}. \tag{2.13}$$

where $\alpha_2 = Eg(X;\theta)^2$, $\alpha_3 = Eg(X;\theta)^3$ and $\alpha_4 = Eg(X;\theta)^4$. The estimators given in the following table were used to construct an estimator of B_c:

Parameter	Estimator	Expression
α_2	$\tilde{\alpha}_2$	$n\hat{\alpha}_2/(n-1)$
α_4	$\tilde{\alpha}_4$	$(n\hat{\alpha}_4 - 6\tilde{\alpha}_2^2)/(n-4)$
α_2^2	$\tilde{\alpha}_{22}$	$\tilde{\alpha}_2^2 - \tilde{\alpha}_4/n$
α_3	$\tilde{\alpha}_3$	$n\hat{\alpha}_3/(n-3)$
α_3^2	$\tilde{\alpha}_{33}$	$\tilde{\alpha}_3^2 - (\hat{\alpha}_6 - \tilde{\alpha}_3^2)/n$
α_2^3	$\tilde{\alpha}_{222}$	$\tilde{\alpha}_2^3$

In this thesis, we use equation (2.13) and the estimation table above to estimate Bartlett correction factor B_c.

3 Simulation Study

In this Chapter, we report results from extensive simulation experiments to evaluate the performance of the proposed methods with finite sample sizes. For comparison purposes, we carry out simulation based on model (2.1) for empirical likelihood-based method and adjusted empirical likelihood-based method since Su and Liang (2009) already proved that empirical likelihood-based method is better than other methods. To improve our proposed method, we also carry out Bartlett corrected AEL-based method in our simulation. The coverage probability of the true β_1^* was reported by using EL method, AEL method, and the Bartlett corrected AEL method.

We set the confidence interval to be 95% with type-I error $\alpha = 0.05$ and ran each simulation 1000 times, with the results being the percentage of confidence intervals derived from EL, AEL and bartlet corrected AEL covering true β_1^* of the 1000 simulations. We then compare the performance of the proposed test with empirical likelihood-based test.

Example 1. X_1 and X_2 are generated as follows. Let $u_1 \sim U(1, 10)$ and $u_2 \sim U(1, 10)$, then $X_1 = \text{Exp}(u_1/4) + \epsilon_1$ and $X_2 = \text{Exp}(u_2/4) + \epsilon_2$, where $\epsilon_j \sim N(0, 4)$. Let $Y_j = 1 + X_j\beta_j + \varepsilon_j$, where ε_j are the error terms with a mean of 0 and a variance of σ_j^2. Sample sizes are $n_1 = 15$ and $n_2 = 15$. We conduct simulations, for seven different δ , the true difference between β_1 and β_2 : 0, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30. We consider the following ten cases for the error terms ε_j. Cases 1 - 5 are having equal variances $\sigma_1^2 = \sigma_2^2 = 1$ for two independent errors and cases 6 - 10 are having different variances, $\sigma_1^2 = 1$ and $\sigma_2^2 = 5$.

Case 1: ε_1 and ε_2 independently follow $N(0, \sigma_j^2), j = 1, 2. \sigma_1^2 = \sigma_2^2 = 1$. This is to check the proposed method under normality assumption.

Case 2: ε_1 and ε_2 independently follow $U(-1, 1)\sqrt{3}\sigma_j, j = 1, 2. \sigma_1^2 = \sigma_2^2 = 1$. This is to check the proposed test when the normal distribution assumption is violated.

Case 3: ε_1 and ε_2 independently follow $(\chi^2_{(1)} - 1) * \sigma_j, j = 1, 2. \sigma_1^2 = \sigma_2^2 = 1$. This is to check the proposed test when errors follow χ^2 distribution.

Case 4: ε_1 and ε_2 independently follow $t_{(5)} * \sigma_j, j = 1, 2. \sigma_1^2 = \sigma_2^2 = 1$. This is to check the proposed test when errors follow t distribution.

Case 5: ε_1 and ε_2 independently follow $(\text{Exp}(1) - 1) * \sigma_j, j = 1, 2. \sigma_1^2 = \sigma_2^2 = 1$. This is to check the proposed test when errors follow Exp distribution.

Case 6: ε_1 and ε_2 independently follow $N(0, \sigma_j^2)$ for $j = 1, 2$. σ_1 and σ_2 are different with $\sigma_1^2 = 1$ and $\sigma_2^2 = 5$. This design is to check the performance of the proposed test when the variances of two errors are unequal.

Case 7: ε_1 and ε_2 independently follow $U(-1, 1)\sqrt{3}\sigma_j, j = 1, 2$. $\sigma_1^2 = 1$ and $\sigma_2^2 = 5$.

Case 8: ε_1 and ε_2 independently follow $(\chi^2_{(1)} - 1) * \sigma_j, j = 1, 2$. $\sigma_1^2 = 1$ and $\sigma_2^2 = 5$.

Case 9: ε_1 and ε_2 independently follow $t_{(5)} * \sigma_j, j = 1, 2$. $\sigma_1^2 = 1$ and $\sigma_2^2 = 5$.

Case 10: ε_1 and ε_2 independently follow $(\text{Exp}(1) - 1) * \sigma_j, j = 1, 2$. $\sigma_1^2 = 1$ and $\sigma_2^2 = 5$.

Table 1 lists the coverage probability for empirical likelihood-based test, adjusted empirical likelihood-based test and Bartlett corrected AEL-based test under different error cases 1 - 5 with equal variance $\sigma_1^2 = \sigma_2^2 = 1$. Under the equal variances condition and small sample sizes, we observe that AEL shows higher coverage probability than EL method, but both methods are not close to the nominal level. The Bartlett corrected AEL method further improves the AEL coverage probability and is closer to the nominal 95% level.

Table 2 lists the simulation results under different variances of the model error for cases 6 - 10. Here $\sigma_1^2 = 1$ and $\sigma_2^2 = 5$. From Table 2, it is observed that the adjusted empirical likelihood-based test has higher coverage probabilities. Bartlett corrected AEL coverage probability further improved the AEL coverage probability and are the

Table 1: Simulation results of coverage probability for $n_1 = n_2 = 15$, $\sigma_1^2 = \sigma_2^2 = 1$. EL-empirical likelihood based test; AEL- adjusted empirical likelihood based test; BcAEL-Bartlett corrected adjusted empirical likelihood.

Case	Method	β_1^*						
		0	0.05	0.1	0.15	0.2	0.25	0.3
1-Normal	EL	0.880	0.896	0.877	0.891	0.883	0.884	0.868
	AEL	0.906	0.922	0.903	0.914	0.906	0.902	0.892
	BcAEL	0.929	0.942	0.947	0.935	0.933	0.933	0.915
2-Uniform	EL	0.902	0.901	0.888	0.900	0.899	0.915	0.908
	AEL	0.921	0.919	0.915	0.923	0.916	0.933	0.930
	BcAEL	0.941	0.938	0.950	0.946	0.933	0.949	0.942
3-Chisquare	EL	0.870	0.867	0.856	0.866	0.855	0.861	0.837
	AEL	0.898	0.901	0.892	0.904	0.876	0.891	0.877
	BcAEL	0.953	0.951	0.940	0.940	0.928	0.958	0.932
4-t	EL	0.873	0.889	0.859	0.883	0.885	0.864	0.900
	AEL	0.899	0.910	0.885	0.912	0.912	0.891	0.923
	BcAEL	0.935	0.938	0.939	0.946	0.952	0.931	0.951
5-Exponential	EL	0.874	0.865	0.867	0.857	0.859	0.857	0.848
	AEL	0.894	0.898	0.889	0.883	0.894	0.886	0.876
	BcAEL	0.943	0.934	0.948	0.939	0.930	0.925	0.925

closest to nominal level among the three methods for all error cases and true β_1^*.

Example 2. We generated X_1 and X_2 same as Example 1 but with different sample size $n_1 = 25$ and $n_2 = 15$. We considered all the ten error cases in Example 1.

Table 3 lists the coverage probability for empirical likelihood-based test, adjusted empirical likelihood-based test and Bartlett corrected AEL-based test with different sample size, under different error cases 1 - 5 with equal variance $\sigma_1^2 = \sigma_2^2 = 1$. Similar results were observed, Bartlett corrected AEL method gave the much better coverage probability comparing to AEL and EL methods.

Table 4 lists the similar simulation results under respective different error cases 6 - 10 but with $\sigma_1^2 = 1$ and $\sigma_2^2 = 5$, when sample sizes are different $n_1 = 25$ and $n_2 = 15$. Again, simulation results shows that AEL performs better than EL but not is as good as BcAEL. Also we notice that the performance of the Bartlett corrected

Table 2: Simulation results of coverage probability for $n_1 = n_2 = 15$, $\sigma_1^2 = 1, \sigma_2^2 = 5$. EL-empirical likelihood based test; AEL- adjusted empirical likelihood based test; BcAEL-Bartlett corrected adjusted empirical likelihood.

Case	Method	β_1^*						
		0	0.05	0.1	0.15	0.2	0.25	0.3
6-Normal	EL	0.885	0.884	0.858	0.890	0.934	0.927	0.926
	AEL	0.907	0.905	0.881	0.907	0.946	0.944	0.942
	BcAEL	0.926	0.934	0.923	0.933	0.967	0.956	0.959
7-Uniform	EL	0.917	0.900	0.870	0.893	0.938	0.931	0.915
	AEL	0.929	0.916	0.887	0.913	0.958	0.947	0.936
	BcAEL	0.939	0.941	0.916	0.937	0.968	0.962	0.951
8-Chisquare	EL	0.846	0.844	0.817	0.840	0.892	0.874	0.877
	AEL	0.877	0.870	0.843	0.869	0.913	0.903	0.904
	BcAEL	0.918	0.902	0.880	0.891	0.955	0.944	0.948
9-t	EL	0.904	0.893	0.877	0.874	0.900	0.910	0.905
	AEL	0.922	0.913	0.899	0.904	0.921	0.934	0.930
	BcAEL	0.944	0.949	0.942	0.944	0.938	0.962	0.955
10-Exponential	EL	0.871	0.870	0.842	0.861	0.899	0.904	0.879
	AEL	0.896	0.891	0.875	0.884	0.919	0.920	0.903
	BcAEL	0.921	0.922	0.913	0.909	0.954	0.952	0.945

method seems to work better when the error distribution is symmetric than skewed distribution. This may be due to the different sample sizes (both are small) from the two models in this simulation setting.

Example 3. We generated X_1 and X_2 same as Example 1 but with larger sample size $n_1 = 50$ and $n_2 = 30$. We considered all the ten error cases in Example 1.

Table 5 lists the coverage probability for empirical likelihood-based test, adjusted empirical likelihood-based test and Bartlett corrected AEL-based test with different but larger sample size, under different error cases 1 - 5 with equal variance $\sigma_1^2 = \sigma_2^2 = 1$.

Table 6 lists the similar simulation results under respective different error cases 6 - 10 but with $\sigma_1^2 = 1$ and $\sigma_2^2 = 5$.

Simulation results in Table 5 and Table 6 show that AEL performs better than

Table 3: Simulation results of coverage probability for $n_1 = 25, n_2 = 15$, $\sigma_1^2 = \sigma_2^2 = 1$. EL-empirical likelihood based test; AEL- adjusted empirical likelihood based test; BcAEL-Bartlett corrected adjusted empirical likelihood.

Case	Method	β_1^*						
		0	0.05	0.1	0.15	0.2	0.25	0.3
1-Normal	EL	0.896	0.866	0.887	0.896	0.901	0.858	0.905
	AEL	0.918	0.893	0.914	0.924	0.916	0.883	0.926
	BcAEL	0.953	0.908	0.939	0.948	0.937	0.929	0.948
2-Uniform	EL	0.895	0.906	0.904	0.885	0.921	0.883	0.899
	AEL	0.908	0.924	0.920	0.915	0.933	0.901	0.911
	BcAEL	0.929	0.937	0.937	0.934	0.932	0.931	0.934
3-Chisquare	EL	0.844	0.869	0.867	0.873	0.869	0.860	0.861
	AEL	0.872	0.892	0.884	0.895	0.891	0.894	0.887
	BcAEL	0.910	0.937	0.932	0.938	0.926	0.945	0.926
4-t	EL	0.895	0.891	0.886	0.886	0.901	0.895	0.894
	AEL	0.912	0.916	0.904	0.896	0.917	0.916	0.912
	BcAEL	0.947	0.943	0.940	0.924	0.937	0.951	0.948
5-Exponential	EL	0.878	0.865	0.857	0.863	0.868	0.850	0.857
	AEL	0.901	0.885	0.885	0.885	0.891	0.875	0.883
	BcAEL	0.931	0.939	0.920	0.930	0.917	0.925	0.928

EL but not as good as BcAEL. In addition, when the sample size difference is larger, the Bartlett corrected AEL coverage probability is a little off from the nominal level.

Based on the simulation results with small samples, it can be seen that AEL-based method improves the coverage probability over EL method and Bartlett correction further improves the coverage probability of AEL method.

We also noticed that both the AEL method and BcAEL method have the coverage probability closer to nominal level when the error has a symmetric distribution such as Normal distribution, Uniform distribution and t distribution than the cases when the error has a skewed distribution like chi-square and exponential distributions. Also the proposed method perform better when the two sample sizes are similar. This could be due to the small sample sizes in our simulation experiments.

Overall, under all the error distribution in our simulation, combined with different

Table 4: Simulation results of coverage probability for $n_1 = 25, n_2 = 15$, $\sigma_1^2 = 1, \sigma_2^2 = 5$. EL-empirical likelihood based test; AEL- adjusted empirical likelihood based test; BcAEL-Bartlett corrected adjusted empirical likelihood.

Case	Method	β_1^*						
		0	0.05	0.1	0.15	0.2	0.25	0.3
6-Normal	EL	0.883	0.890	0.856	0.886	0.896	0.899	0.882
	AEL	0.901	0.908	0.874	0.899	0.915	0.910	0.903
	BcAEL	0.934	0.937	0.925	0.928	0.942	0.939	0.928
7-Uniform	EL	0.893	0.911	0.845	0.886	0.907	0.914	0.922
	AEL	0.911	0.925	0.861	0.902	0.923	0.929	0.930
	BcAEL	0.928	0.943	0.901	0.926	0.936	0.950	0.951
8-Chisquare	EL	0.856	0.820	0.792	0.833	0.856	0.858	0.851
	AEL	0.870	0.844	0.813	0.856	0.876	0.871	0.868
	BcAEL	0.899	0.874	0.855	0.888	0.912	0.902	0.903
9-t	EL	0.889	0.892	0.836	0.882	0.872	0.897	0.889
	AEL	0.902	0.907	0.861	0.907	0.896	0.912	0.902
	BcAEL	0.942	0.947	0.932	0.942	0.930	0.949	0.938
10-Exponential	EL	0.857	0.871	0.802	0.856	0.880	0.877	0.872
	AEL	0.878	0.880	0.822	0.876	0.902	0.893	0.890
	BcAEL	0.898	0.910	0.877	0.903	0.924	0.920	0.915

sample sizes or equal sample sizes, equal variances or unequal variances, the BcAEL gives the best coverage probability. Thus we recommend the Bartlett corrected AEL method in these situations.

Table 5: Simulation results of coverage probability for $n_1 = 50, n_2 = 30$, $\sigma_1^2 = \sigma_2^2 = 1$. EL-empirical likelihood based test; AEL- adjusted empirical likelihood based test; BcAEL-Bartlett corrected adjusted empirical likelihood.

Case	Method	β_1^*						
		0	0.05	0.1	0.15	0.2	0.25	0.3
1-Normal	EL	0.904	0.917	0.888	0.912	0.905	0.907	0.912
	AEL	0.913	0.930	0.901	0.919	0.911	0.916	0.928
	BcAEL	0.925	0.941	0.923	0.936	0.923	0.932	0.934
2-Uniform	EL	0.904	0.923	0.919	0.901	0.903	0.910	0.898
	AEL	0.916	0.932	0.930	0.911	0.930	0.920	0.908
	BcAEL	0.928	0.942	0.936	0.930	0.943	0.933	0.915
3-Chisquare	EL	0.876	0.887	0.864	0.875	0.890	0.889	0.878
	AEL	0.888	0.895	0.883	0.891	0.898	0.901	0.896
	BcAEL	0.918	0.935	0.910	0.920	0.924	0.927	0.926
4-t	EL	0.907	0.899	0.912	0.915	0.907	0.891	0.892
	AEL	0.921	0.908	0.919	0.920	0.914	0.905	0.904
	BcAEL	0.942	0.928	0.937	0.938	0.932	0.930	0.923
5-Exponential	EL	0.876	0.873	0.902	0.871	0.885	0.875	0.877
	AEL	0.883	0.885	0.910	0.882	0.890	0.885	0.890
	BcAEL	0.907	0.913	0.932	0.904	0.918	0.916	0.909

Table 6: Simulation results of coverage probability for $n_1 = 50, n_2 = 30$, $\sigma_1^2 = 1, \sigma_2^2 = 5$. EL-empirical likelihood based test; AEL- adjusted empirical likelihood based test; BcAEL-Bartlett corrected adjusted empirical likelihood.

Case	Method	β_1^*						
		0	0.05	0.1	0.15	0.2	0.25	0.3
6-Normal	EL	0.896	0.897	0.912	0.925	0.904	0.914	0.914
	AEL	0.910	0.906	0.918	0.931	0.912	0.919	0.923
	BcAEL	0.939	0.937	0.939	0.951	0.929	0.939	0.939
7-Uniform	EL	0.915	0.914	0.925	0.924	0.917	0.904	0.921
	AEL	0.923	0.926	0.934	0.930	0.926	0.910	0.930
	BcAEL	0.940	0.948	0.947	0.946	0.935	0.932	0.944
8-Chisquare	EL	0.874	0.855	0.870	0.855	0.885	0.881	0.882
	AEL	0.883	0.867	0.885	0.867	0.896	0.888	0.888
	BcAEL	0.901	0.896	0.914	0.899	0.919	0.908	0.909
9-t	EL	0.897	0.907	0.919	0.899	0.904	0.908	0.897
	AEL	0.903	0.923	0.929	0.908	0.912	0.917	0.908
	BcAEL	0.937	0.952	0.950	0.932	0.936	0.937	0.940
10-Exponential	EL	0.865	0.860	0.894	0.897	0.879	0.900	0.903
	AEL	0.870	0.870	0.905	0.908	0.886	0.909	0.915
	BcAEL	0.906	0.906	0.931	0.926	0.902	0.925	0.931

4 Drug Study

In this chapter, we present an illustrative analysis with a real-data example from Hocking (2003) which is a medical experiment involving four drugs to measure their effects on the response y to a particular stimulus. Since the individuals in this study may not be identical in their responses to the drugs, their response x to the stimulus prior to taking the drug was measured. For each drug, there are 9 observations (x, y). Figure 4 shows these records, together with the linear model fitting of each drug. It can be seen that squares, octagons, triangles, and diamonds indicate the observed values with regard to drugs A, B, C, D respectively and the solid, slashed, dotted, and broken-dotted lines correspond to the linear model fitting for drugs A, B, C, and D respectively.

Be observing y values in the four treatments, one may see that the treatment effects in Groups A and C are stronger than those in Groups B and D. We are interested in whether these effects are significant and how we can provide a statistical justification for them.

We therefore fit the model $\hat{y}_{ij} = \alpha_j + x_{ij}\beta_j$ and apply the three tests to this data set for a pair comparison. We obtain the linear regression, empirical likelihood-based, adjusted empirical likelihood-based and Bartlett corrected adjusted empirical likelihood-based confidence intervals for the difference of the treatment effects for each pair in Table 7.

These result shown in Table 7 indicate that the differences between drug A and B, A and D, C and D are statistically significant since 0 is not included in the 95% confidence interval. This result is consistent with Figure 4 and Su and Liang's (2009) paper result. Although the results from the three tests are all the same, the results from the Bartlett corrected AEL should be considered more reliable than the others

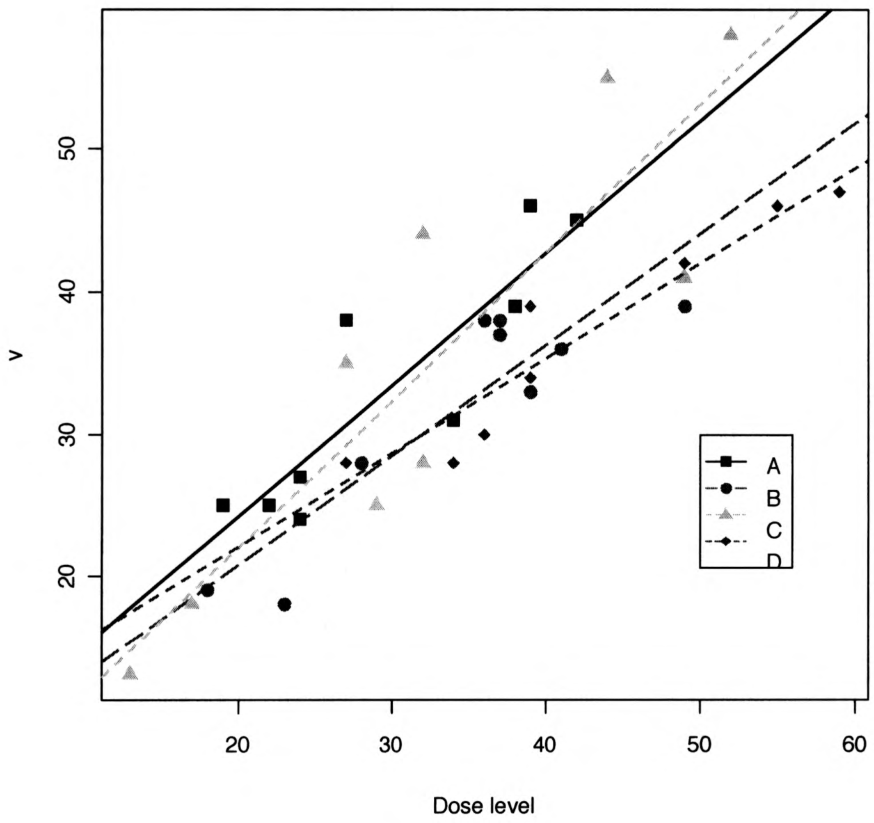

Figure 4: Data from a drug study. Responses of treatment effect against dose level for four treatment agents: A(square), B(circle), C(triangle), D(diamond), and the associated four linear model fittings.

Table 7: Confidence interval of comparing treatment effects against dose level for four treatment using empirical likelihood test (EL), adjusted empirical likelihood test (AEL) and Bartlett corrected AEL method.

Confidence	Test		
Interval	EL	AEL	BcAEL
A vs B	(0.061 0.266)	(0.050 0.277)	(0.039 0.289)
A vs C	(-0.130 0.183)	(-0.155 0.207)	(-0.163 0.216)
A vs D	(0.078 0.268)	(0.062 0.279)	(0.062 0.294)
B vs C	(-0.295 0.008)	(-0.309 0.008)	(-0.331 0.008)
B vs D	(-0.040 0.098)	(-0.048 0.109)	(-0.052 0.112)
C vs D	(0.041 0.323)	(0.041 0.343)	(0.041 0.363)

based on our simulation results.

5 Discussion and Conclusion

In this thesis, we derived adjusted empirical likelihood-based method and Bartlett correction of AEL-based method for comparing treatment effects by testing equality of coefficients in linear regression models. Our simulation results indicate that the confidence intervals constructed by the adjusted empirical likelihood have coverage probabilities comparable to or more accurate than the original empirical likelihood. And our proposed Bartlett corrected adjusted empirical likelihood-based method shows the best performance.

We have shown that the BcAEL method for comparison of treatment effects in a linear model setting is reliable in detecting the difference of the parameters of interest. The computation of the proposed test is simple, and the theoretical results do not need any distribution assumption nor the homoscedasticity assumption. This makes it applicable to real studies in detecting and comparing two treatment effects.

Bartlett correction factor can also be estimate from bootstrap. In the future, we can use the bootstrap to get the factor B_c and compare with the estimate recommended by Chen and Liu (2010) to find out a better performance.

In this thesis, we only studied the case of linear models. In the future, we may consider two partially linear models, both of which take the form

$$Y = X^T\beta + g(Z) + \varepsilon,$$

where $g(.)$ is an unknown smoothing function and Z is a covariate, which nonlinearly contributes to Y. We may use the same idea to study the difference between two linear parameters(βs).

References

[1] Ali, M. M., Silver, J. L., "Tests for equality between sets of coefficients in two linear regressions under heteroscedasticity," *Journal of American Statistical Association*, 80, 730-735, 1985.

[2] Bartlett, M. S., "Properties of sufficiency and statistical tests," *Proceedings of the Royal Society of London*, A, 160, 268-282, 1937.

[3] Bhuyan, K. C., Majumder, A. K., "Testing equality of regressions when error variances are heterogeneous to study the production differentials of rice in Bangladesh," *Biometrica*, J, 38, 819-828, 1996.

[4] Chen, S. X., "On the accuracy of empirical likelihood confidence regions for linear regression model," *Annals of the Institute of Statistical Mathematics*, 45, 621-637, 1993.

[5] Chen, S. X., Cui, H. J., "On Bartlett correction of empirical likelihood in the presence of nuisance parameters," *Biometrika*, 93, 215-220, 2006.

[6] Chen, S. X., Cui, H. J., "On the second-order properties of empirical likelihood with moment restrictions," *Journal of Econometrics*, 141, 492-516, 2007.

[7] Chen, J., Variyath, A. M., Abraham, B., "Adjusted empirical likelihood and its properties," *Journal of Computational and Graphical Statistics*, 17, 426-443, 2008.

[8] Chen, J., Liu, Y., "Adjusted empirical likelihood with high-order precision," *The Annals of Statistics*, Vol. 38, 1341-1362, 2010.

[9] Chen, J., Wang, L., Pu, X., "Resampling calibrated adjusted empirical likelihood," *The Canadian Journal of Statistics*, Vol. 43, 42-59, 2015.

[10] Chen, S. X., Keilegom, I. V., "A Review on Empirical Likelihood Methods for Regression," *Test*, 18, Vol. 3, 448-451, 2009.

[11] Chow, G. C., "Tests of equality between sets of coefficients in two linear regressions," *Econometrica*, 28, 591-605, 1960.

[12] Conerly, M. D., Mansfield, E. R., "An approximate test for comparing heteroscedastic regression models," *Journal of American Statistical Association*, 83, 811-817, 1988.

[13] DiCiccio, T., Hall, P., Romano, J., "Empirical likelihood is Bartlett correctable," *The Annals of Statistics*, 19, 1053-1061, 1991.

[14] Dupont, W. D., Plummer, W. D., "Power and sample size calculations for studies involving linear regression," *Controlled Clinical Trials*, 19, 589-601, 1998.

[15] Efron, B., "Nonparametric standard errors and confidence intervals (with discussion)," *The Canadian Journal of Statistics*, 9, 139-172, 1981.

[16] Emerson, S. A., Owen, A. B., "Calibration of the empirical likelihood method for a vector mean," *Electronic Journal of Statistics*, 3, 1161-1192, 2009.

[17] Glenn, N. L., Zhao, Y., "Weighted empirical likelihood estimates and their robustness properties," *Computational Statistics and Data Analysis*, 5, 5130-5141, 2007.

[18] Hocking, R. R., "*Methods and applications of linear models, regression and analysis of variance*," 2, John Wiley & Sons, N.J., 2003.

[19] Larsen, R. J., Marx, M. L., "*An introduction to mathematical statistics and its applications*," Prentice-Hall, Englewood Cliffs, N.J., 1986.

[20] Lawley, D. N., "A general method for approximating the distribution of likelihood ratio criteria," *Biometrika*, 43, 295-303, 1956.

[21] Li, G., Wang, Q. H., "Empirical likelihood regression analysis for right censored data," *Statistica Sinica*, 13, 51-68, 2003.

[22] Liang, H., Wang, S., Carroll, R.J., "Partially linear models with missing response variables and error-prone covariates," *Biometrika*, 94, 185-198, 2007.

[23] McCullagh, P., "*Tensor methods in statistics*," Chapman and Hall, London, 1987.

[24] McKeague, I. W., Zhao, Y., "Simultaneous confidence bands for ratios of survival functions via empirical likelihood," *Statistics & Probability Letters*, 60, 405-415, 2002.

[25] McKeague, I. W., Zhao, Y., "Width-scaled confidence bands for survival functions," *Statistics & Probability Letters*, 76, 327-339, 2006.

[26] Owen, A. B., "Empirical likelihood ratio confidence intervals for a single functional," *Biometrika*, 75, 237-249, 1988.

[27] Owen, A. B., "Empirical likelihood confidence regions," *The Annals of Statistics*, 18, 90-120, 1990.

[28] Owen, A. B., "Empirical likelihood for linear models," *The Annals of Statistics*, 19, 1725-1747, 1991.

[29] Owen, A. B., "*Empirical Likelihood*," Chapman & Hall/CRC, New York, 2001.

[30] Qin, J., Lawless, J., "Empirical likelihood and general equations," *The Annals of Statistics*, 22, 300-325, 1994.

[31] Qin, J., Zhang, B., "Empirical-likelihood-based inference in missing response problems and its application in observational studies," *Journal of the Royal Statistical Society*, B, 69, 101-122, 2007.

[32] Schmidt, P., Sickles, R., "Some further evidence on the use of the Chow test under heteroskedasticity," *Econometrica*, 45, 1293-1298, 1977.

[33] Su, H., Liang, H., "An empirical likelihood-based method for comparison of treatment effects-Test of equality of coefficients in linear models," *Computational Statistics and Data Analysis*, 54, 1079-1088, 2009.

[34] Thomas, D. R., Grunkemeier, G. L., "A Guide to Elementary Number TheoryConfidence interval estimation of survival probabilities for censored data," *Journal of the American Statistical Association*, 70, 865-871, 1975.

[35] Wang, Q. H., Rao, J. N. K., "Empirical likelihood for linear regression models under imputation for missing responses," *Canadian Journal of Statistics*, 29, 597-608, 2001.

[36] Wang, Q. H., Rao, J. N. K., "Empirical likelihood-based inference in linear models with missing data," *Scandinavian Journal of Statistics*, 29, 563-576, 2002.

[37] Zhao, Y., Chen, F., "Empirical likelihood inference for censored median regression model via nonparametric kernel estimation," *Journal of Multivariate Analysis*, 99, 215-231, 2005.

[38] Zhao, Y., "Regression analysis for long-term survival rate via empirical likelihood," *Journal of Nonparametric Statistics*, 17, 995-1007, 2005.

[39] Zhao, Y., "Semiparametric inference for transformation models via empirical likelihood," *Journal of Multivariate Analysis*, 101, 1846-1858, 2010.

6 Appendix

Proof of Theorem 1

Suppose $X_1^* = \gamma_0 + X_2^*\gamma_1 + \epsilon_1$, which along with (2.4) indicates that $Y = \nu_0 + X_2^*\nu_1 + \epsilon_1\beta_1^* + \varepsilon$. Let $\widehat{\gamma_k}$ and $\widehat{\nu_k}$ be the least squares estimators of γ_k and ν_k for $k = 0, 1$, respectively. Then $E(X_1^*|X_2^*) = \gamma_0 + X_2^*\gamma_1$ and $E(Y|X_2^*) = \nu_0 + X_2^*\nu_2$.

Let $\Omega_i = \{X_{1i}^* - \hat{E}(X_{1i}^*|X_{2i}^*)\}^T[Y_i - \hat{E}(Y_i|X_{2i}^*) - \{X_{1i}^* - \hat{E}(X_{1i}^*|X_{2i}^*)\}\beta_1^*]$ and $\tilde{\Omega}_i = \{X_{1i}^* - E(X_{1i}^*|X_{2i}*)\}^T[Y_i - E(Y_i|X_{2i}^*) - \{X_{1i}^* - E(X_{1i}^*|X_{2i}^*)\}\beta_1^*]$ for $i = 1, ..., n$. A standard simplification as in Owen (2001) yields that

$$p_i = \frac{1}{n(1 + a^T\Omega_i)}, \tag{A.1}$$

for $i = 1, ..., n$, where a is the solution of the equation

$$\sum_{i=1}^{n} \frac{\Omega_i}{1 + a^T\Omega_i} = 0, \tag{A.2}$$

A direct calculation yields that

$$\begin{aligned}\Omega_i &= (\hat{X}_{1i} - \tilde{X}_{1i})^T[(\hat{Y}_i - \tilde{Y}_i) - (\hat{X}_{1i} - \tilde{X}_{1i})\beta_1^* + (\tilde{Y}_i - \tilde{X}_{1i})\beta_1^*] \\ &+ \tilde{X}_{1i}^T[(\hat{Y}_i - \tilde{Y}_i) - (\hat{X}_{1i} - \tilde{X}_{1i})\beta_i^* + (\tilde{Y}_i - \tilde{X}_{1i})\beta_1^*] \\ &= \tilde{\Omega}_i + o_p(1). \end{aligned} \tag{A.3}$$

Mimicking the proof Theorem 3.2 of Owen (2001), we have $||a|| = O_p(n^{-1/2})$ and $max_{1\leqslant i\leqslant n}||\Omega_i|| = o_p(n^{1/2})$. By (A.3), we have $max_{1\leqslant i\leqslant n}||\Omega_i|| \leqslant ||\tilde{\Omega}_i|| + o_p(1) = o_p(n^{1/2})$.

Using the same argument as the proof of Theorem 4 in Liang et al.(2007), we have

$$-2\log R_n(\beta_1^*) = \sum_{i=1}^{n} a^T\Omega_i\Omega_i^T a + o_p(1) = \sum_{i=1}^{n}(a^T\Omega_i)^2 + o_p(1)$$

$$= (n^{-1/2}\sum_{i=1}^{n}\Omega_i)^T(n^{-1}\sum_{i=1}^{n}\Omega_i\Omega_i^T)^{-1}(n^{-1/2}\sum_{i=1}^{n}\Omega_i) + o_p(1).$$

To show Theorem 1. We first show that $n_1^{-1/2}\sum_{i=1}^{n_1}\Omega_i$ and $n_1^{-1/2}\sum_{i=1}^{n_1}\tilde{\Omega}_i$ have the same limiting normal distribution, $n_2^{-1/2}\sum_{i=n_1+1}^{n}\Omega_i$ and $n_2^{-1/2}\sum_{i=n_1+1}^{n}\tilde{\Omega}_i$ have the same limiting normal distribution, $n_1^{-1}\sum_{i=1}^{n_1}\Omega_i\Omega_i^T$ and $n_1^{-1}\sum_{i=1}^{n_1}\tilde{\Omega}_i\tilde{\Omega}_i^T$ have the same limiting value, and $n_2^{-1}\sum_{i=n_1+1}^{n+1}\Omega_i\Omega_i^T$ and $n_2^{-1}\sum_{i=n_1+1}^{n+1}\tilde{\Omega}_i\tilde{\Omega}_i^T$ have the same limiting value. Without loss of generality, we will prove it for the first group, that is, for $i = 1, ..., n_1$. The second group, $i = n_1 + 1, ..., n + 1$, can be proved similarly.

Note that $\tilde{\Omega}_i - \Omega_i$ can be decomposed as:

$$\tilde{X}_{1i}^{*T}[\hat{E}(Y_i|X_{2i}^*) - E(Y_i|X_{2i}^*) - \{\hat{E}(X_{1i}^*|X_{2i}^*) - E(X_{1i}^*|X_{2i}^*)\}\beta_1^*]$$

$$-\{\hat{E}(X_{1i}^*|X_{2i}^*) - E(X_{1i}^*|X_{2i}^*)\}^T[\hat{E}(Y_i|X_{2i}^*) - E(Y_i|X_{2i}^*) - \{\hat{E}(X_{1i}^*|X_{2i}^*) - E(X_{1i}^*|X_{2i}^*)\}\beta_1^*]$$

$$+\{\hat{E}(X_{1i}^*|X_{2i}^*) - E(X_{1i}^*|X_{2i}^*)\}^T(\tilde{Y}_i - \tilde{X}_{1i}^*\beta_1^*). \quad (A.4)$$

Note that

$$\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\tilde{X}_{1i}^{*T}\{E(Y_i|X_{2i}^*) - E(Y_i|X_{2i}^*)\}$$

$$= \sqrt{n_1}(\hat{\nu}_1 - \nu_1)\frac{1}{n_1}\sum_{i=1}^{n_1}\epsilon_{1i}X_{2i}^* + \sqrt{n_1}(\hat{\nu}_0 - \nu_0)\frac{1}{n_1}\sum_{i=1}^{n_1}\epsilon_{1i}.$$

The right hand side is of order $o_p(1)$ since $\sqrt{n_1}(\hat{\nu}_1 - \nu_1) = Op(1)$ for $k = 0, 1$, $1/n_1\sum_{i=1}^{n_1}\epsilon_{1i}X_{2i}^* = o_p(1)$, and $1/n_1\sum_{i=1}^{n_1}\epsilon_{1i} = o_p(1)$.

It follows that $n_1^{-1/2}\sum_{i=1}^{n_1}\tilde{X}_{1i}^{*T}\hat{E}(Y_i|X_{2i}^*) - E(Y_i|X_{2i}^*) = o_p(1)$.

In the same way, we obtain the following statements.

$$\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\tilde{X}_{1i}^{*T}\{\hat{E}(X_{1i}^*|X_{2i}^*)-E(X_{1i}^*|X_{2i}^*)\}\beta_1^*=o_p(1).$$

$$\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\{\hat{E}(X_{1i}^*|X_{2i}^*)-E(X_{1i}^*|X_{2i}^*)\}^T\{\hat{E}(Y_i|X_{2i}^*)-E(Y_i|X_{2i}^*)\}=o_p(1).$$

$$\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\{\hat{E}(X_{1i}^*|X_{2i}^*)-E(X_{1i}^*|X_{2i}^*)\}^T\{\hat{E}(X_{1i}^*|X_{2i}^*)-E(X_{1i}^*|X_{2i}^*)\}\beta_1^*=o_p(1).$$

$$\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\{\hat{E}(X_{1i}^*|X_{2i}^*)-E(X_{1i}^*|X_{2i}^*)\}^T\tilde{Y}_i=o_p(1).$$

$$\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\{\hat{E}(X_{1i}^*|X_{2i}^*)-E(X_{1i}^*|X_{2i}^*)\}^T\tilde{X}_{1i}^*\beta_1^*=o_p(1).$$

These results imply that $n_1^{-1/2}\sum_{i=1}^{n_1}\Omega_i$ and $n_1^{-1/2}\sum_{i=1}^{n_1}\tilde{\Omega}_i$ asymptotically have the same normal distribution, and $n_1^{-1}\sum_{i=1}^{n_1}\Omega_i\Omega_i^T$ and $n_1-1\sum_{i=1}^{n_1}\tilde{\Omega}_i\tilde{\Omega}_i^T$ have the same limiting value.

In addition, we know that $(n_1^{-1/2}\sum_{i=1}^{n_1}\tilde{\Omega}_i)^T(n_1^{-1}\sum_{i=1}^{n_1}\tilde{\Omega}_i\tilde{\Omega}_i^T)(n_1^{-1/2}\sum_{i=1}^{n_1}\tilde{\Omega}_i)$ converges to χ_p^2 in distribution. It follows that

$$(n_1^{-1/2}\sum_{i=1}^{n_1}\Omega_i)^T(n_1^{-1}\sum_{i=1}^{n_1}\Omega_i\Omega_i^T)(n_1^{-1/2}\sum_{i=1}^{n_1}\Omega_i)$$

converges to χ_p^2 in distribution.

Furthermore, we have

$$\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n}\tilde{\Omega}_i=\frac{\sqrt{n_1}}{\sqrt{n}}\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\tilde{\Omega}_i+\frac{\sqrt{n_2}}{\sqrt{n}}\frac{1}{\sqrt{n_2}}\sum_{i=n_1+1}^{n}\tilde{\Omega}_i,$$

which converges to a summand of two independent normal random variables with two weights, that is, $\sqrt{\xi}Z_1+\sqrt{1-\xi}Z_2$, where $\xi=\lim_{n\to\infty}n_1/n$. So we know that

$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\Omega}_i$ converges to a normal distribution, $N(0,\Sigma)$, where $\Sigma = cov(\tilde{\Omega}_1)$.

On the other hand,

$$\frac{1}{n}\sum_{i=1}^{n}\tilde{\Omega}_i\tilde{\Omega}_i^T = \frac{n_1}{n}\frac{1}{n_1}\sum_{i=1}^{n_1}\tilde{\Omega}_i\tilde{\Omega}_i^T + \frac{n_2}{n}\frac{1}{n_2}\sum_{i=n_1+1}^{n}\tilde{\Omega}_i\tilde{\Omega}_i^T \to \xi\Sigma + (1-\xi)\Sigma = \Sigma.$$

As a consequence, we conclude that

$$(n^{-1/2}\sum_{i=1}^{n}\Omega_i)^T(n^{-1}\sum_{i=1}^{n}\Omega_i\Omega_i^T)(n^{-1/2}\sum_{i=1}^{n}\tilde{\Omega}_i) \sim \chi_p^2.$$

Since $n^{-1/2}\sum_{i=1}^{n}\Omega_i$ and $n^{-1/2}\sum_{i=1}^{n}\tilde{\Omega}_i$ have the same limiting normal distribution, and $n^{-1}\sum_{i=1}^{n}\Omega_i\Omega_i^T$ and $n^{-1}\sum_{i=1}^{n}\tilde{\Omega}_i\tilde{\Omega}_i^T$ have the same limiting value.

With the definition of $g_i(\beta) = \hat{X}_{1i}^{*T}(\hat{Y}_i - \hat{X}_{1i}^{*}\beta_1^{*})$ for $i = 1, 2, ..., n$ and $g_{n+1}(\beta) = \frac{-\log(n)}{2n}\sum_{i=1}^{n} g_i$, we can follow the same procedure of proof in Chen, Variyath and Bovas (2008) to show that the $-2\log\mathcal{R}_n(\beta_1^*)$ converges to χ_p^2 as $n \to \infty$. The proof is thus complete.