USING THE BOOTSTRAP TO ANALYZE VARIABLE STARS DATA

A Dissertation

by

MICKEY P. DUNLAP

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2004

Major Subject: Statistics

USING THE BOOTSTRAP TO ANALYZE

VARIABLE STARS DATA

A Dissertation

by

MICKEY P. DUNLAP

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

Jeffrey D. Hart
(Chair of Committee)

Francis J. Narcowich
(Member)

Willa Chen
(Member)

Thomas E. Wehrly
(Member)

Michael T. Longnecker
(Head of Department)

December 2004

Major Subject: Statistics

ABSTRACT

Using the Bootstrap to Analyze

Variable Stars Data. (December 2004)

Mickey P. Dunlap B.S., University of Tennessee at Martin;

M.S., Mississippi State University

Chair of Advisory Committee: Dr. Jeffrey D. Hart

Often in statistics it is of interest to investigate whether or not a trend is significant. Methods for testing such a trend depend on the assumptions of the error terms such as whether the distribution is known and also if the error terms are independent. Likelihood ratio tests may be used if the distribution is known but in some instances one may not want to make such assumptions. In a time series, these errors will not always be independent. In this case, the error terms are often modelled by an autoregressive or moving average process. There are resampling techniques for testing the hypothesis of interest when the error terms are dependent, such as, model-based bootstrapping and the wild bootstrap, but the error terms need to be whitened. In this dissertation, a bootstrap procedure is used to test the hypothesis of no trend for variable stars when the error structure assumes a particular form. In some cases, the bootstrap to be implemented is preferred over large sample tests in terms of the level of the test. The bootstrap procedure is able to correctly identify the underlying distribution which may not be χ^2.

To Mom and Dad

ACKNOWLEDGEMENTS

I would especially like to thank Dr. Jeffrey D. Hart for his help in completing my Ph.D. over the semesters. His guidance and expertise is greatly appreciated. I would also like to thank my committee members, Dr. Fran Narcowich, Dr. Willa Chen, Dr. Thomas E. Wehrly and Dr. Marina Vannucci for serving on my committee.

There are many others in the department at Texas A&M that have given advice in one way or another . Special thanks to Dr. Jane Harvill for the motivation and support to complete this degree.

Finally, thanks most of all to my parents and family. Without the much needed encouragement, I wouldn't have made it this far.

TABLE OF CONTENTS

		Page
ABSTRACT		iii
DEDICATION		iv
ACKNOWLEDGEMENTS		v
TABLE OF CONTENTS		vi
LIST OF TABLES		viii
LIST OF FIGURES		ix
CHAPTER		
I	INTRODUCTION	1
II	RELATED LITERATURE	4
	2.1 Model Fitting	4
	2.2 Bootstrap	6
	2.3 Edgeworth Expansion	7
III	BOOTSTRAPPING THE TEST STATISTIC	10
	3.1 Test Statistic	10
	3.2 Edgeworth Expansion of a Test Statistic	12
	3.3 Edgeworth Expansion of S_n	16
	3.4 Bootstrap Procedure	20
	3.5 Bootstrap Distributions	21
	3.6 Bootstrap Simulation	24
IV	ESTIMATION	33
	4.1 Method of Moments	33
	4.2 Consistency of Estimates	34
V	APPLICATION OF THE BOOTSTRAP PROCEDURE	42
	5.1 R. Aquilae	43
	5.2 R. Bootis	48
	5.3 R. Canum Venaticorum	52

CHAPTER Page

5.4 W. Draconis 56
5.5 Y. Aquarii 60

VI CONCLUSION 64

REFERENCES 65

APPENDIX A 66

APPENDIX B 70

APPENDIX C 74

VITA 81

LIST OF TABLES

TABLE		Page
1	Simulation results for comparing the bootstrap and large sample test Type I error rates for r = 0, 0.5 and 1.	30
2	Simulation results for comparing the bootstrap and large sample test Type I error rates for n = 100, 150 and 200 when r = 0.	31
3	Simulation results for comparing the bootstrap and large sample test Type I error rates for r = 2, 5, and 10.	32
4	Maximum likelihood estimates and standard errors of the regression coefficients for R. Aquilae.	44
5	Maximum likelihood estimates and standard errors of the regression coefficients for R. Bootis.	48
6	Maximum likelihood estimates and standard errors of the regression coefficients for R. Canum Venaticorum.	52
7	Maximum likelihood estimates and standard errors of the regression coefficients for W. Draconis.	56
8	Maximum likelihood estimates and standard errors of the regression coefficients for Y. Aquarii.	60

LIST OF FIGURES

FIGURE Page

1 Quantile plots for test statistics. $I \sim \sqrt{r}N(0,1)$ and $\epsilon \sim N(0,1)$. Clockwise from upper left, r = 0, 0.5 and 1. 27

2 Quantile plots for test statistics. $I \sim \sqrt{r}\exp(-1,1)$ and $\epsilon \sim \exp(-1,1)$. Clockwise from upper left, r = 0, 0.5 and 1. 28

3 Quantile plots for test statistics. $I \sim \sqrt{r}\exp(-1,1)$ and $\epsilon \sim N(0,1)$. Clockwise from upper left, r = 0, 0.5 and 1. 29

4 Plot of pseudo-periods for R. Aquilae. 45

5 Density estimate of the bootstrapped test statistic for R. Aquilae. . . 46

6 Quantile plot of the bootstrapped test statistic for R. Aquilae. Vertical and horizontal lines mark the 90th, 95th, 98th and 99th percentiles of the bootstrapped test statistic. 47

7 Plot of pseudo-periods for R. Bootis. 49

8 Density estimate of the bootstrapped test statistic for R. Bootis. . . . 50

9 Quantile plot of the bootstrapped test statistic for R. Bootis. Vertical and horizontal lines mark the 90th, 95th, 98th and 99th percentiles of the bootstrapped test statistic. 51

10 Plot of pseudo-periods for R. Canum Venaticorum. 53

11 Density estimate of the bootstrapped test statistic for R. Canum Venaticorum. 54

12 Quantile plot of the bootstrapped test statistic for R. Canum Venaticorum. Vertical and horizontal lines mark the 90th, 95th, 98th and 99th percentiles of the bootstrapped test statistic. 55

13 Plot of pseudo-periods for W. Draconis. 57

14 Density estimate of the bootstrapped test statistic for W. Draconis. . 58

FIGURE Page

15 Quantile plot of the bootstrapped test statistic for W. Draconis. Vertical and horizontal lines mark the 90th, 95th, 98th and 99th percentiles of the bootstrapped test statistic. 59

16 Plot of pseudo-periods for Y. Aquarii. 61

17 Density estimate of the bootstrapped test statistic for Y. Aquarii. . . . 62

18 Quantile plot of the bootstrapped test statistic for Y. Aquarii. Vertical and horizontal lines mark the 90th, 95th, 98th and 99th percentiles of the bootstrapped test statistic. 63

CHAPTER I

INTRODUCTION

Often in statistics it is of interest to investigate whether or not a trend is significant. Consider a model of the form

$$Y_i = m(x_i) + Z_i, \quad i = 1, \ldots, n, \tag{1.1}$$

for which we are interested in testing $H_0 : m(x) = \text{constant}, \forall\ x$. The manner in which the tests are carried out depends on the nature of the error terms, Z_i. The most common setting is to assume that $Z_1, \ldots, Z_n$ are independent and identically distributed (i.i.d.) N(0, σ^2) random variables.

If the forms of the error distribution and $m(\cdot)$ are known, likelihood ratio tests can be used to test hypotheses of interest. However, in a time series, for example, errors are not always independent, and may be modelled by an autoregressive or moving average process for instance. Again, it may be possible to assume that the distribution is normal. In some instances, one may not want to make such assumptions about the distribution of the error terms and therefore some form of resampling such as bootstrapping or the jackknife could be used to test hypotheses about the model.

An example of a situation where interest centers on testing constancy of a function arises in the analysis of variable star data. For a given variable star, astronomers observe the times at which the star's brightness achieves a minimum or maximum. Koen & Lombard (2001) propose a model which utilizes both the minima and maxima. Denoting by T_i the ith time at which the brightness of a star achieves a maximum,

---

The format and style follow that of *Journal of the American Statistical Association.*

then the ith pseudo-period is $Y_i = T_i - T_{i-1}$. The term pseudo-period is used to distinguish between a single time interval between maxima and the actual period, which is the long-run average of the pseudo-periods. The model that we want to consider is just as in (1.1) with $m(\cdot)$ being a polynomial of specified degree with x_i being the time of maximum brightness and the error structure as follows:

$$Z_i = I_i + \epsilon_i - \epsilon_{i-1}, \quad i = 1, \ldots, n,$$

where I_i represents random intrinsic variation peculiar to the star and ϵ_i is the measurement error at time T_i. It is assumed that $I_1, \ldots, I_n \sim (0, \sigma_I^2)$ i.i.d., $\epsilon_0, \ldots, \epsilon_n \sim (0, \sigma_\epsilon^2)$ i.i.d., and the process $\{I_i\}$ is independent of $\{\epsilon_i\}$. Note that there is no mention of the distributional form for I or ϵ. The covariance matrix of Z has the form

$$\begin{aligned}
\Sigma &= \begin{pmatrix} \sigma_I^2 + 2\sigma_\epsilon^2 & -\sigma_\epsilon^2 & 0 & \cdots & 0 \\ -\sigma_\epsilon^2 & \sigma_I^2 + 2\sigma_\epsilon^2 & -\sigma_\epsilon^2 & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & 0 & \ddots & \ddots & -\sigma_\epsilon^2 \\ 0 & \cdots & 0 & -\sigma_\epsilon^2 & \sigma_I^2 + 2\sigma_\epsilon^2 \end{pmatrix} \\
&= \sigma_\epsilon^2 \begin{pmatrix} r+2 & -1 & 0 & \cdots & 0 \\ -1 & r+2 & -1 & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & 0 & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & -1 & r+2 \end{pmatrix} \\
&= \sigma_\epsilon^2 A \qquad (1.2)
\end{aligned}$$

where $r = \sigma_I^2 / \sigma_\epsilon^2$.

The goals of this dissertation are as follows:

1. Develop a test statistic for testing H_0.
2. Obtain the large-sample distribution of the test statistic under H_0.
3. Obtain estimates of the covariance parameters which the distribution of the test statistic depends on.
4. Determine what moments of I and ϵ the test statistic depends on.
5. Use knowledge of 4. to develop a bootstrap procedure for approximating the small-sample distribution of the test statistic.
6. Apply the test procedure to variable star data.

CHAPTER II

RELATED LITERATURE

To begin testing H_0, a suitable test statistic and the distribution of this test statistic under H_0 will be needed. The distribution will depend on the parameters in the covariance matrix (1.2) and also on the distributions of I and ϵ. In this case, a distribution will not be imposed upon I and ϵ.

2.1 Model Fitting

It will first need to be determined how to fit the model, $m(x)$. The method chosen to fit the model determines the residuals, which in turn determine the estimates of the parameters in the covariance matrix. One possibility for fitting the model that comes to mind is to use a nonparametric approach. This would involve estimating a smoothing parameter. Opsomer, Wang & Yang (2001) point out that ignoring correlation between the errors causes the commonly used automatic tuning parameter selection methods, such as cross-validation or plug-in, to break down. Furthermore, they state that a wrong choice of the smoothing parameter can lead to an estimated correlation that does not reflect the true correlation in the random error. Some authors make modifications in bandwidth selection techniques by modelling the correlation structure parametrically and use this estimate to adjust for the bandwidth selection.

Another possible approach is to fit a polynomial of specified degree, with the degree being analagous to the smoothing parameter mentioned in the nonparametric approach. To use ordinary least squares (OLS), we first rewrite the original model in matrix form:

$$\mathbf{Y} = X\boldsymbol{\beta} + \mathbf{Z}.$$

This results in estimates of $\boldsymbol{\beta}$ as

$$\hat{\boldsymbol{\beta}}_{\text{OLS}} = (X'X)^{-1}X'\mathbf{Y}.$$

Recalling that the covariance matrix of $\mathbf{Z}$ has the structure as in (1.2), generalized least squares (GLS) performs a transformation so that the errors are at least approximately uncorrelated. In this case, $\mathbf{Z}^* = A^{-1/2}\mathbf{Z}$ will have covariance $\sigma_\epsilon^2 I$. Carrying this transformation throughout, the transformed model becomes

$$\mathbf{Y}^* = X^*\boldsymbol{\beta} + \mathbf{Z}^*. \tag{2.1}$$

Applying ordinary least squares (OLS) to this model yields

$$\begin{aligned} \hat{\boldsymbol{\beta}}_{\text{GLS}} &= (X^{*\prime}X^*)^{-1}X^{*\prime}\mathbf{Y}^* \\ &= (X'A^{-1}X)^{-1}X'A^{-1}\mathbf{Y}. \end{aligned} \tag{2.2}$$

One should note, however, that if r is unknown, $\hat{\boldsymbol{\beta}}_{\text{GLS}}$ is not usuable in practice. According to Fang & Koreisha (2001), both $\hat{\boldsymbol{\beta}}_{\text{OLS}}$ and $\hat{\boldsymbol{\beta}}_{\text{GLS}}$ are consistent estimates of $\boldsymbol{\beta}$ when the errors are correlated. If X is non-stochastic, the columns of X are linearly independent, and $\lim_{n\to\infty}(\frac{1}{n}X'X)$ is finite and non-singular, then $\hat{\boldsymbol{\beta}}_{\text{OLS}}$ is a consistent estimate of $\boldsymbol{\beta}$. The off-diagonal elements of A do not bias the ordinary least squares estimates but even fairly small off-diagonal elements can cause the variance of the estimates to increase substantially (Sen & Srivastava 1990). Another consideration when comparing ordinary and generalized least squares estimates is that A is not generally known exactly so it must be estimated. This obviously complicates matters a great deal. Fortunately, for the covariance structure of interest, *i.e*, (1.2), there are only two parameters to estimate, r and σ_ϵ^2, and only r must be estimated in A. If we imposed no structure whatsoever on the covariance matrix, there would be $n(n+1)/2$ elements, which could result in very unreliable estimates.

2.2 Bootstrap

The bootstrap involves drawing samples from the observed data to approximate a statistic's sampling distribution. The basic idea is to treat the sample as if it were the population so that we may make inferences without assuming a distribution, such as the normal.

Let $X_1, \ldots, X_n$ be a random sample of size n. The bootstrap sample, $X_1^*, \ldots, X_n^*$, is a random sample with replacement from $\{X_1, \ldots, X_n\}$. For each resample, statistics of interest such as the sample mean can be obtained in order to approximate their sampling distributions. An advantage of the bootstrap over other resampling methods, such as the jackknife, is that the number of resamples is not limited to the sample size. The bootstrap can be used to obtain confidence intervals and perform hypothesis tests.

Hall (1992) points out that the bootstrap performs best when using a pivotal quantity, such as $\sqrt{n}(\hat{\theta}-\theta_0)/\hat{\sigma}$, rather than a nonpivotal quantity such as, $\sqrt{n}(\hat{\theta}-\theta_0)$. The error in the bootstrap is generally only $O_p(n^{-1})$ when using a pivotal quantity but $O_p(n^{-1/2})$ when using a nonpivotal quantity.

Since the basic bootstrap assumes the sample is random, modifications are needed when the data are dependent. There are various ways to bootstrap when the variables are dependent. Davison & Hinkley (1997) mention methods such as model-based resampling, block resampling and wild bootstrap. With model-based resampling, the idea is to create a whitened sequence, $\{W_t\}$, from the original data. How the data are whitened depends on the model used, such as AR(p), MA(q), etc. A bootstrap sample is generated by reconstructing the time series based on the whitened data. This may require a burn-in period to insure the bootstrap sample is stationary. For block resampling, b blocks of length l, $(B_1, \ldots, B_b)$, are created based on the original data. The blocks are then resampled and placed end-to-end to create a new series. The series generated here may be less dependent than the original data. For the wild bootstrap, an i.i.d. sequence $\{\eta_t\}$ with first three moments 0, 1 and 1 is needed. After whitening the original data just as in model-based resampling, the whitened data is multiplied by the i.i.d. sequence to yield $\{\eta_t W_t\}$. The bootstrap sample is generated just as in model-based resampling except that it is not based on the whitened data but rather a modified version of the whitened data.

2.3 Edgeworth Expansion

Often times, the exact distribution of a test statistic is unknown and inferences depend upon the approximate or limiting distribution, which may, for example, be normal or chi-square. Let $X_1, \ldots, X_n$ be i.i.d. random variables with mean μ and finite variance σ^2. It is well known, according to the central limit theorem, that

$$S_n = \sqrt{n}(\bar{X} - \mu)/\sigma$$

is approximately normally distributed with mean 0 and variance 1. The Edgeworth expansion is a mathematical procedure which attempts to improve upon the approximation given by the central limit theorem. Hall (1992) has shown that after expanding the cumulative distribution function (cdf) of S_n in a power series in $n^{-1/2}$, the Edgeworth expansion of $\mathrm{P}(S_n < x)$ takes the form

$$\Phi(x) + n^{-1/2} p_1(x)\phi(x) + \cdots + n^{-j/2} p_j(x)\phi(x) + \cdots$$

where $\Phi(x)$ and $\phi(x)$ are the respective cdf and probability density function for the standard normal distribution. The quantities, $p_j(x)$, are functions of the moments of $\bar{X} - \mu$ and Hermite polynomials. For example,

$$p_1(x) = -\{A_1/\sigma + \frac{1}{6}A_2\sigma^{-3}(x^2 - 1)\}$$

where

$$\begin{aligned} & \mathrm{E}(S_n) = n^{-1/2}A_1 + O(n^{-1}) \\ & \mathrm{E}(S_n^2) - (\mathrm{E}(S_n))^2 = \sigma^2 + O(n^{-1}) \\ \text{and} \qquad & \mathrm{E}(S_n^3) - 3\mathrm{E}(S_n^2)\mathrm{E}(S_n) + 2(\mathrm{E}(S_n))^3 = n^{-1/2}A_2 + O(n^{-1}). \end{aligned} \tag{2.3}$$

The analogous Edgeworth expansion for the cdf of the bootstrap test statistic, S_n^*, has the form

$$\Phi(x) + n^{-1/2}\hat{p}_1(x)\phi(x) + \cdots + n^{-j/2}\hat{p}_j(x)\phi(x) + \cdots$$

where $\hat{p}_j(x)$ is the result of using the bootstrap to estimate the moments of S_n. Therefore,

$$\begin{aligned} P(S_n < x) - P(S_n^* < x) &= n^{-1/2}p_1(x)\phi(x) - n^{-1/2}\hat{p}_1(x)\phi(x) + O(n^{-1}) \\ &= n^{-1/2}(p_1(x) - \hat{p}_1(x))\phi(x) + O_p(n^{-1}). \end{aligned}$$

The bootstrap procedure will then be more accurate than the large sample test if the moments of S_n can be estimated efficiently. Typically, when a pivotal quantity is used, $(p_1(x) - \hat{p}_1(x))$ is $O_p(n^{-1/2})$ yielding an error in the bootstrap of only $O_p(n^{-1})$.

If the first and third moments of S_n were 0, the bootstrap procedure would have no major advantage over using a large sample test. However, when, in particular, there is skewness, the bootstrap procedure can be quite useful.

CHAPTER III

BOOTSTRAPPING THE TEST STATISTIC

In order to test the hypothesis of interest, a test statistic needs to be chosen. There are various test statistics which are suitable for testing no trend. The choice of the test statistic is important since it can greatly influence results such as the performance of the bootstrap. In this chapter, we want to consider possible test statistics as well as properties of the bootstrap scheme to be used.

3.1 Test Statistic

In multiple regression, the F statistic used is

$$\frac{\sum_{i=1}^{n}(\widehat{m(x_i)} - \bar{y})^2}{\sum_{i=1}^{n} \hat{Z}_i^2}.$$

For the model,

$$\mathbf{Y} = X\boldsymbol{\beta} + \mathbf{Z},$$

the usual F statistic is equivalent to

$$F = \frac{(J\mathbf{Y} - X\hat{\boldsymbol{\beta}}_{\text{OLS}})'(J\mathbf{Y} - X\hat{\boldsymbol{\beta}}_{\text{OLS}})/df_1}{(\mathbf{Y} - X\hat{\boldsymbol{\beta}}_{\text{OLS}})'(\mathbf{Y} - X\hat{\boldsymbol{\beta}}_{\text{OLS}})/df_2} \tag{3.1}$$

where $J = \frac{1}{n}\mathbf{1}\mathbf{1}'$. The numerator and denominator degrees of freedom in this case, df_1 and df_2, correspond to the respective ranks of the following matrices:

$$(J - X(X'X)^{-1}X^{-1})^2 \text{ and } (\mathbf{I}_n - X(X'X)^{-1}X^{-1})^2,$$

where $\mathbf{I}_n$ is the identity matrix of size n.

The above test statistic (3.1) is for error terms that are independent. Another possibility hinges upon (2.1), which accounts for the covariance of the error terms. By applying the necessary transformation to $\mathbf{Y}$ and the design matrix X the test statistic now has the form

$$F^* = \frac{(J\mathbf{Y}^* - X^*\hat{\boldsymbol{\beta}}_{\text{GLS}})'(J\mathbf{Y}^* - X^*\hat{\boldsymbol{\beta}}_{\text{GLS}})/df_1^*}{(\mathbf{Y}^* - X^*\hat{\boldsymbol{\beta}}_{\text{GLS}})'(\mathbf{Y}^* - X^*\hat{\boldsymbol{\beta}}_{\text{GLS}})/df_2^*}$$

where J is as above, $\mathbf{Y}^* = A^{-1/2}\mathbf{Y}$ and $X^* = A^{-1/2}X$. In this case, the degrees of freedom are the ranks of

$$(JA^{-1/2} - X^*(X^{*\prime}X^*)^{-1}X^{*\prime}A^{-1/2})^2$$

$$\text{and} \quad (A^{-1/2} - X^*(X^{*\prime}X^*)^{-1}X^{*\prime}A^{-1/2})^2$$

respectively. If A contains unknown parameters, then A may be replaced by an estimator, $\hat{A}$.

Another choice for the test statistic is to take a likelihood ratio approach. Begin by assuming that the error terms are normally distributed with mean $\mathbf{0}$ and covariance matrix $\Sigma = \sigma_\epsilon^2 A$. The log-likelihood has the form

$$l(\boldsymbol{\beta}, \sigma_\epsilon^2, r) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma_\epsilon^2) - \frac{1}{2}\log(|\,\hat{A}\,|) - \frac{1}{2\sigma_\epsilon^2}(\mathbf{Y} - X\boldsymbol{\beta})'A^{-1}(\mathbf{Y} - X\boldsymbol{\beta}).$$

The log-likelihood ratio statistic is given by

$$\log(\Lambda) = -l(\hat{\boldsymbol{\beta}}, \hat{\sigma}_\epsilon^2, \hat{r}) + l(\hat{\boldsymbol{\beta}}_0, \hat{\sigma}_{\epsilon 0}^2, \hat{r}_0)$$

where $\hat{\boldsymbol{\beta}} = (\hat{\beta}_{00}, 0, \ldots, 0)'$, $\hat{\sigma}_{\epsilon 0}^2$ and $\hat{r}_0$ are estimates under H_0 and $\hat{\boldsymbol{\beta}}$,$\hat{\sigma}_\epsilon^2$ and $\hat{r}$ are unrestricted estimates. Alternatively the statistic can be transformed by

$$\begin{aligned}
-2\log(\Lambda) &= 2l(\hat{\boldsymbol{\beta}}, \hat{\sigma}_\epsilon, \hat{r})) - 2l(\hat{\boldsymbol{\beta}}, \hat{\sigma}_{\epsilon 0}, \hat{r}_0)) \\
&= -n\log(\hat{\sigma}^2_\epsilon) - \log(|\,\hat{A}\,|) - \frac{n}{2} \\
&\qquad +n\log(\hat{\sigma}^2_{\epsilon 0}) + \log(|\,\hat{A}_0\,|) + \frac{n}{2} \\
&= n\log\left(\frac{\hat{\sigma}^2_{\epsilon 0}}{\hat{\sigma}^2_\epsilon}\right) + \log\left(\frac{|\,\hat{A}_0\,|}{|\,\hat{A}\,|}\right). \qquad (3.2)
\end{aligned}$$

As mentioned by Davison & Hinkley (1997), in most cases, $-2\log(\Lambda)$ is distributed approximately χ^2_d under H_0, where d is the difference in the dimension of the parameter space under H_a and H_0. In our case, d is the degree of the polynomial fitted under H_a. The distribution does not depend upon any unknown quantities and hence $-2\log(\Lambda)$ is an approximate pivotal quantity, which makes it ideal for bootstrap purposes.

3.2 Edgeworth Expansion of a Test Statistic

An Edgeworth expansion will be used to determine which moments need to be estimated in order to carry out the bootstrap procedure. Let $x_1, \ldots, x_n$ be evenly spaced design points on (0,1) and let ϕ be a function defined on [0,1] that has the following properties:

1. ϕ is continuously differentiable.

2. $\sum_{i=1}^n \phi(x_i) = 0$.

3. $\int_0^1 \phi^2(x)\,dx = 1$.

Consider a test statistic of the form

$$T_n = \frac{\sum_{i=1}^n \phi(x_i) Y_i}{\sqrt{\hat{\sigma}^2_I A_n + \hat{\sigma}^2_\epsilon B_n}}. \qquad (3.3)$$

Under H_0, the numerator of T_n will have the form

$$\mu \sum_{i=1}^{n} \phi(x_i) + \sum_{i=1}^{n} \phi(x_i) Z_i = \sum_{i=1}^{n} \phi(x_i) Z_i$$

where $Z_j = I_j + \epsilon_j - \epsilon_{j-1}, j = 1, \ldots, n$. Standardized estimates of regression coefficients, for example, take on the form (3.3). The Edgeworth expansion will be developed using ideas from Hall on the assumption that the data are, if not independent, only weakly dependent. In practice, the test statistic will take the form (3.2) but for theoretical facility the Edgeworth expansion will be based on (3.3). Ultimately, both expressions should depend on the same moments in the expansion. The variance estimators are defined as follows:

$$\hat{\sigma}_I^2 = \frac{1}{n} \sum_{i=1}^{n} \hat{Z}_i^2 + \frac{2}{n} \sum_{i=2}^{n} \hat{Z}_i \hat{Z}_{i-1} \text{ and } \hat{\sigma}_\epsilon^2 = \frac{-1}{n} \sum_{i=2}^{n} \hat{Z}_i \hat{Z}_{i-1},$$

where $\hat{Z}_1, \ldots, \hat{Z}_n$ are residuals from a fitted model. The quantities A_n and B_n are

$$A_n = \sum_{i=1}^{n} \phi^2(x_i) \text{ and } B_n = 2 \left[A_n - \sum_{i=1}^{n-1} \phi(x_i) \phi(x_{i+1}) \right].$$

We will assume that H_0 is true (meaning the regression function is constant) and that $\sigma_I^2 > 0$. In this case we have

$$T_n = \frac{\sum_{i=1}^{n} \phi(x_i)(I_i + \epsilon_i - \epsilon_{i-1})/\sqrt{n}}{\sqrt{\hat{\sigma}_I^2 A_n/n + \hat{\sigma}_\epsilon^2 B_n/n}}.$$

Now, $A_n/n = 1 + O(n^{-1})$ since ϕ is continuously differentiable and $\int_0^1 \phi^2(x)dx = 1$. Also,

$$\begin{aligned} \frac{A_n}{n} - \frac{1}{n} \sum_{i=1}^{n-1} \phi(x_i)\phi(x_{i+1}) &= \frac{A_n}{n} - \frac{1}{n} \sum_{i=1}^{n-1} \phi(x_i) \left[\phi(x_i) + n^{-1} \phi'(\tilde{x}_i) \right] \\ &= n^{-1} \phi^2(x_n) - \frac{1}{n^2} \sum_{i=1}^{n-1} \phi(x_i) \phi'(\tilde{x}_i) \end{aligned}$$

$$
\begin{aligned}
&= n^{-1}\phi^2(x_n) - \frac{1}{n}\int_0^1 \phi(x)\phi'(x)\,dx + O(n^{-2}) \\
&= \frac{\phi^2(0)+\phi^2(1)}{2n} + O(n^{-2})
\end{aligned}
$$

where $\tilde{x}_i$ is in a neighborhood of x_i. Now consider

$$
\begin{aligned}
\sum_{i=1}^{n}\phi(x_i)(\epsilon_i - \epsilon_{i-1}) &= \sum_{i=1}^{n}\phi(x_i)\epsilon_i - \sum_{i=0}^{n-1}\phi(x_{i+1})\epsilon_i \\
&= \phi(x_n)\epsilon_n - \phi(x_1)\epsilon_0 + \sum_{i=1}^{n-1}[\phi(x_i) - \phi(x_{i+1})]\epsilon_i \\
&= \phi(x_n)\epsilon_n - \phi(x_1)\epsilon_0 - \frac{1}{n}\sum_{i=1}^{n-1}\phi'(\tilde{x}_i)\epsilon_i.
\end{aligned}
$$

Combining the above results implies that

$$
T_n = S_n + \frac{[\phi(x_n)\epsilon_n - \phi(x_1)\epsilon_0]}{\sqrt{n}\hat{\sigma}_I} + O_p(n^{-1}), \tag{3.4}
$$

where

$$
S_n = \frac{\sum_{i=1}^{n}\phi(x_i)I_i/\sqrt{n}}{\hat{\sigma}_I}.
$$

Therefore, the effect of the ϵ_i's is negligible, and we see that

$$
T_n - S_n = O_p(n^{-1/2}).
$$

A standard central limit theorem implies that $T_n \xrightarrow{\mathcal{D}} N(0,1)$.

More interesting is what (3.4) says about our bootstrap scheme. To be able to use existing results, the second summand on the right hand side of equation (3.4) is potentially troublesome. Why? Well, that term is $O_p(n^{-1/2})$, which is the same size as the first order error term in an Edgeworth expansion. But using a function ϕ with the property that $\phi(0) = \phi(1) = 0$ effectively eliminates that term, since then

$$\frac{[\phi(x_n)\epsilon_n - \phi(x_1)\epsilon_0]}{\sqrt{n}\hat{\sigma}_I} = O_p(n^{-3/2}).$$

Examples of functions that satisfy 1.-3. and $\phi(0) = \phi(1) = 0$ are $\sin(2\pi jx)$, $j = 1, 2, \ldots$. For such functions we have

$$T_n = S_n + \delta_n,$$

where $\delta_n = O_p(n^{-1})$. For a positive sequence $\{\gamma_n\}$ we have

$$\begin{aligned} P(S_n \le x - \gamma_n \cap \delta_n \le \gamma_n) \le \quad & P(T_n \le x) \le P(S_n \le x + \gamma_n \cup \delta_n \le -\gamma_n) \Longrightarrow \\ P(S_n \le x - \gamma_n) - P(\delta_n > \gamma_n) \le \quad & P(T_n \le x) \le P(S_n \le x + \gamma_n) + P(\delta_n \le -\gamma_n). \end{aligned}$$

Suppose we assume that S_n has a density g_n such that $g_n(x) \le C$ for all n and x. Then $P(S_n \le x \pm \gamma_n) = P(S_n \le x) \pm \gamma_n g_n(\tilde{x}_n) = P(S_n \le x) + O(\gamma_n)$, uniformly in x. In order for this bound to be useful, we need to take γ_n to be smaller than $n^{-1/2}$. However, γ_n needs to be bigger than n^{-1} in order to guarantee that $P(|\delta_n| > \gamma_n)$ tends to 0.

Now, examining δ_n and using Markov's inequality, we can show that if the ϵ_i's are i.i.d. with $\mathrm{E}(\epsilon_i^{2k}) < \infty$ and the I_i's are i.i.d. with $\mathrm{E}(I_i^{2k}) < \infty$ for some positive integer k, then

$$P(|\delta_n| > \gamma_n) = O\left(\frac{1}{(n\gamma_n)^{2k}}\right).$$

Suppose $\gamma_n = Cn^{-\alpha}$. Then for a given k, the two errors are balanced when $2k(1-\alpha) = \alpha$, or when $\alpha = 2k/(2k+1)$. So, we have shown that

$$P(T_n \le x) = P(S_n \le x) + O(n^{-2k/(2k+1)}) \tag{3.5}$$

when I_i and ϵ_i each has $2k$ moments.

Given an Edgeworth expansion for $P(S_n \leq x)$, the error term in (3.5) will dominate all but the leading error term of order $n^{-1/2}$. However, this is not all bad since the error in the bootstrap is typically $O_p(n^{-1})$. With enough moments, the error term in (3.5) will be arbitrarily close to n^{-1}, and more importantly, even if k is only 1, we will have a rigorous proof that the bootstrap provides smaller error than simply comparing T_n with percentiles of the standard normal.

Now let T_n^* and S_n^* be our bootstrap versions of T_n and S_n. Then so long as we use an appropriate absolutely continuous distribution for the I_i^*'s, the distribution of S_n^* will satisfy all the conditions above, and we will have

$$P(T_n^* \leq x) = P(S_n^* \leq x) + O_p(n^{-2k/(2k+1)}),$$

which is enough to ensure that

$$P(T_n^* \leq x) = P(T_n \leq x) + O_p(n^{-2k/(2k+1)}).$$

3.3 Edgeworth Expansion of S_n

In this section, it is of interest to find an Edgeworth expansion for a statistic of the form

$$S_n = \sqrt{n}\frac{\frac{1}{n}\sum_{i=1}^{n}\phi(x_i)I_i}{\check{\sigma}_I^2}$$

where $I_1, \ldots, I_n \sim (0, \sigma_I^2)$ and $\phi(x_i)$ has the properties as defined in Section 3.2. Also, $\check{\sigma}_I^2$ is

$$\check{\sigma}_I^2 = \frac{1}{n-1}\sum_{i=2}^{n} Z_i^2 + \frac{2}{n-1}\sum_{i=2}^{n} Z_i Z_{i-1}.$$

Note that the estimator $\hat{\sigma}_I^2$ used in practice has residuals $\hat{Z}_i$ in place of Z_i. However, it is easily shown that $\hat{\sigma}_I = \check{\sigma}_I + O_p(n^{-1})$. The $O_p(n^{-1})$ term may be absorbed into the $O(n^{-2k/(2k+1)})$ term in the previous section.

The Edgeworth expansion in Chapter 3 of Hall's book depends upon the first three moments of S_n. To obtain moments of S_n, first write S_n as

$$\sqrt{n}\frac{X}{\sqrt{Y}}$$

where $X = \frac{1}{n}\sum_{i=1}^{n}\phi(x_i)I_i$ and $Y = \check{\sigma}_I^2$. If $Y^{-1/2}$ is expanded in a Taylor series about $y_0 = \mathrm{E}(Y^2) = \sigma_I^2$,

$$\sqrt{n}\frac{X}{\sqrt{Y}} \approx \sqrt{n}\frac{X}{\sqrt{y_0}}\Big(1 - \frac{Y - y_0}{2y_0}\Big).$$

Let $\kappa_I = \mathrm{E}(I^3)$, $\kappa_\epsilon = \mathrm{E}(\epsilon^3)$, $\gamma_I = \mathrm{E}(I^4)$ and $\gamma_\epsilon = \mathrm{E}(\epsilon^4)$. Since $\mathrm{E}(X) = 0$,

$$\begin{aligned}
\mathrm{E}(S_n) &\approx \frac{-\sqrt{n}}{2y_0^{3/2}}\mathrm{E}(XY) \\
&= \frac{-\sqrt{n}}{2y_0^{3/2}}\Big(\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=2}^{n}\phi(x_i)\mathrm{E}(I_iZ_j^2) \\
&\qquad + \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=2}^{n}\phi(x_i)\mathrm{E}(I_iZ_jZ_{j-1})\Big) \\
&= \frac{-\sqrt{n}}{2y_0^{3/2}}\frac{1}{n(n-1)}\sum_{i=2}^{n}\phi(x_i)\mathrm{E}(I_i^3) \\
&= \frac{-\kappa_I}{2\sqrt{n}y_0^{3/2}}\frac{1}{n-1}\sum_{i=2}^{n}\phi(x_i) \\
&= \frac{-\kappa_I}{2\sqrt{n}\sigma_I^3}O(n^{-1}).
\end{aligned}$$

Futhermore,

$$\begin{aligned}
S_n^2 &\approx \frac{nX^2}{\sigma_I^2}\Big(1 - \frac{Y - \sigma_I^2}{2\sigma_I^2}\Big)^2 \\
&= \frac{nX^2}{\sigma_I^2}\Big(1 - \frac{Y - \sigma_I^2}{\sigma_I^2} + \frac{(Y - \sigma_I^2)^2}{4\sigma_I^4}\Big)
\end{aligned}$$

which yields

$$\begin{aligned}
\mathrm{E}(S_n^2) &\approx \frac{n}{\sigma_I^2}\mathrm{E}(X^2) - \frac{n}{\sigma_I^4}\mathrm{E}(X^2(Y-\sigma_I^2)) + O(n^{-1}) \\
&= \frac{1}{n}\sum_{i=1}^{n}\phi^2(x_i) - \frac{n}{\sigma_I^4}\mathrm{E}(X^2(Y-\sigma_I^2)) + O(n^{-1}) \\
&= 1 - \frac{n}{\sigma_I^4}\mathrm{E}(X^2(Y-\sigma_I^2)) + O(n^{-1}).
\end{aligned}$$

Define $\delta_k = \epsilon_k - \epsilon_{k-1}$. Then

$$\begin{aligned}
Y - \sigma_I^2 &= \tilde{\sigma}_I^2 - \sigma_I^2 + A_n + B_n + C_n + D_n \\
&\quad + \frac{1}{n-1}\sum_{k=2}^{n}(\delta_k^2 + 2\delta_k\delta_{k-1})
\end{aligned}$$

where

$$\begin{aligned}
\tilde{\sigma}_I^2 &= \frac{1}{n-1}\sum_{k=2}^{n} I_k^2, \\
A_n &= \frac{2}{n-1}\sum_{k=2}^{n} I_k\delta_k, \\
B_n &= \frac{2}{n-1}\sum_{k=2}^{n} I_k I_{k-1}, \\
C_n &= \frac{2}{n-1}\sum_{k=2}^{n} I_k\delta_{k-1} \\
\text{and } D_n &= \frac{2}{n-1}\sum_{k=2}^{n} I_{k-1}\delta_k.
\end{aligned}$$

Since $\mathrm{E}(X^2A_n) = \mathrm{E}(X^2B_n) = \mathrm{E}(X^2C_n) = \mathrm{E}(X^2D_n) = 0$,

$$\mathrm{E}(X^2(Y-\sigma_I^2)) \approx \mathrm{E}(X^2(\tilde{\sigma}_I^2 - \sigma_I^2)) + \mathrm{E}(X^2(A_n + B_n + C_n + D_n))$$

$$
\begin{aligned}
&= \frac{1}{n^2(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=2}^{n} \phi(x_i)\phi(x_j)\mathrm{E}(I_i I_j (I_k^2 - \sigma_I^2)) \\
&= \frac{1}{n^2(n-1)} \sum_{i=1}^{n} \phi^2(x_i)(\mathrm{E}(I_i^4) - \sigma_I^4) \\
&= O(n^{-2}).
\end{aligned}
$$

So, $\mathrm{E}(S_n^2) = 1 + O(n^{-1})$. The constant $A_1 = 0$ when $\int_0^1 \phi(x)dx = 0$, which it must be or other aspects of our argument break down (see Section 3.2). Therefore, from Hall's notation

$$
\sigma^2 = \mathrm{E}(S_n^2) - (\mathrm{E}(S_n))^2 = 1 + O(n^{-1}).
$$

Similarly,

$$
\begin{aligned}
S_n^3 &\approx \frac{n^{3/2}X^3}{y_0^{3/2}}\Big(1 - \frac{Y - y_0}{2y_0}\Big)^3 \\
&= \frac{n^{3/2}X^3}{y_0^{3/2}}\Big(1 - 3\frac{Y - y_0}{2y_0} + O_p(n^{-1})\Big).
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}(n^{3/2}X^3) &= \frac{n^{3/2}}{n^3} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \phi(x_i)\phi(x_k)\phi(x_k) I_i I_j I_k \\
&= \frac{\kappa_I}{n\sqrt{n}} \sum_{i=1}^{n} \phi^3(x_i) \\
&= \frac{\kappa_I}{\sqrt{n}} \int_0^1 \phi^3(x)dx + O(n^{-3/2}).
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}(n^{3/2}X^3(Y - y_0)) &= \mathrm{E}(n^{3/2}X^3(\tilde{\sigma}_I^2 - y_0)) + \mathrm{E}(n^{3/2}X^3(A_n + B_n + C_n + D_n)) \\
&= n^{3/2}\mathrm{E}(X^3(\tilde{\sigma}_I^2 - y_0)) + n^{3/2}\mathrm{E}(X^3 B_n) \\
&= \frac{n^{3/2}}{n^3(n-1)}\Big(\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{l=2}^{n} \phi(x_i)\phi(x_j)\phi(x_k) I_i I_j I_k (I_l^2 - y_0)
\end{aligned}
$$

$$
\begin{aligned}
&\quad + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=2}^{n}\phi(x_i)\phi(x_j)\phi(x_k)I_iI_jI_kI_lI_{l-1}\Big)\\
&= O(n^{-3/2}) + \frac{n^{3/2}}{n^3(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\phi^3(x_i)\mathrm{E}(I_i^3I_j^2)\\
&= \frac{n^{3/2}(n-1)}{n^2(n-1)}\kappa_I\sigma_I^2\frac{1}{n}\sum_{i=1}^{n}\phi^3(x_i) + O(n^{-3/2})\\
&= \frac{n^{3/2}(n-1)}{n^2(n-1)}\kappa_I\sigma_I^2\Big[\int_0^1\phi^3(x)dx + O(n^{-1})\Big] + O(n^{-3/2})\\
&= \frac{1}{\sqrt{n}}\kappa_I\sigma_I^2\int_0^1\phi^3(x)dx + O(n^{-3/2}).
\end{aligned}
$$

This means that

$$\mathrm{E}(S_n^3) = -\frac{1}{2\sqrt{n}}\frac{\kappa_I}{\sigma_I^3}\int_0^1\phi^3(x)dx + O(n^{-1}).$$

Therefore, from (2.3), we can write

$$P(S_n \leq x) = \Phi(x) + \phi(x)n^{-1/2}(1/12)(\kappa_I/\sigma_I^3)\int_0^1\phi^3(x)dx + O(n^{-1}).$$

3.4 Bootstrap Procedure

For linear time series models, the usual bootstrap approach is to (approximately) whiten the series, and to then resample from residuals. Ordinarily when applying a bootstrap procedure, the resamples are from the original data. However, for the model of interest (1.1), the Z_i's cannot be whitened as in the case for ARMA processes.

Note that the covariance structure of $Z_1, \ldots, Z_n$ is identical to that of an MA(1) process. However, application of the linear filter that would whiten an MA(1) process with the same covariance function as $\{Z_i\}$ does not whiten $\{Z_i\}$. We must, therefore, use a different approach to bootstrap. A single bootstrap sample will be generated from distributions which have the same first 3 estimated moments as I and ϵ.

1. Generate $I_1^*, I_2^*, \ldots, I_n^* \sim F_I$ such that $\mathrm{E}(I_i^*) = 0$, $\mathrm{E}(I_i^{*2}) = \hat{r}\hat{\sigma}_\epsilon^2$ and $\mathrm{E}(I_i^{*3}) = \hat{\kappa}_I$.

2. Generate $\epsilon_0^*, \epsilon_1^*, \ldots, \epsilon_n^* \sim F_\epsilon$ such that $\mathrm{E}(\epsilon_i^*) = 0$,$\mathrm{E}(\epsilon_i^{*2}) = \hat{\sigma}_\epsilon^2$ and $\mathrm{E}(\epsilon_I^{*3}) = \hat{\kappa}_\epsilon$

   where $\hat{r}, \hat{\sigma}_\epsilon, \hat{\kappa}_I$and $\hat{\kappa}_\epsilon$ are estimates of parameters of interest. Estimation of these parameters will be investigated in Chapter IV.

   A bootstrap sample, $Z_1^*, Z_2^*, \ldots, Z_n^*$, will be constructed by forming

$$Z_i^* = I_i^* + \epsilon_i^* - \epsilon_{i-1}^*, \quad i = 1, 2, \ldots, n.$$

For each bootstrap sample, the test statistic will be calculated. The process is repeated as many times as desired. We can then obtain approximate percentiles of the test statistic under the null hypothesis.

The motivation for this method is that resampling from any distributions whose first three moments are $\sqrt{n}$ consistent for the corresponding population moments will produce an error of order n^{-1} in the bootstrap sampling distribution. Even were one able to sample from $\sqrt{n}$ consistent empirical distributions, the bootstrap error would not generally be less than $O(n^{-1})$. This is because estimation error of order $n^{-1/2}$ dominates higher order terms in an Edgeworth expansion. So, in an asymptotic sense, the proposed bootstrap scheme should work just as well as sampling from empirical distributions.

3.5 Bootstrap Distributions

There are various distributions which can be used in the bootstrap process so that the first three moments are 0, c_1 and c_2. For instance, a discrete distribution of the form

X	x_1	x_2
$p(X)$	p_1	$1-p_1$

can be used. Values of x_1, x_2 and p_1 need to be found such that E(X)=0, E(X^2)=c_1 and E(X^3)=c_2. In general,

$$\begin{aligned} p_1 &= -\frac{x_2}{x_1 - x_2}, \\ x_1 &= -\frac{c_1}{x_2}, \\ \text{and} \quad x_2 &= \frac{c_2 \pm \sqrt{c_2^2 + 4c_1^2}}{2c_1} \end{aligned}$$

In the case of the wild bootstrap, it is common to use the above distribution for which c_1 and c_2 are both equal to 1.

An alternative to using the discrete distribution above is to use a mixture of normals. We have the following pdf for a mixture of 2 normal random variables with the same standard deviation, σ:

$$f(x) = p_1 \frac{1}{\sigma} \phi\left(\frac{x-\mu_1}{\sigma}\right) + (1-p_1)\frac{1}{\sigma}\phi\left(\frac{x-\mu_2}{\sigma}\right).$$

We want to find p_1, μ_1, μ_2 and σ such that

$$\begin{aligned} 0 &= p_1\mu_1 + (1-p_1)\mu_2 \\ c_1 &= p_1(\sigma^2 + \mu_1^2) + (1-p_1)(\sigma^2 + \mu_2^2) \\ c_2 &= p_1(3\mu_1\sigma^2 + \mu_1^3) + (1-p_1)(3\mu_2\sigma^2 + \mu_2^3) \end{aligned}$$

This yields the following values of p_1, μ_1, μ_2 and σ^2:

$$
\begin{aligned}
p_1 &= -\frac{\mu_2}{\mu_1 - \mu_2} \\
\mu_1 &= \frac{-\mu_2^2 \pm \sqrt{\mu_2^4 - 4\mu_2 c_2}}{2\mu_2} \\
\text{and } \sigma^2 &= c_1 - \mid \mu_1\mu_2 \mid .
\end{aligned}
$$

We need only choose μ_2 to obtain the rest of the values. There are an infinite number of possibilities that will satisfy the conditions. Note that μ_2 will need to be chosen such that $c_1 >\mid \mu_1\mu_2 \mid$.

Yet another possibility is to generate data from a shifted gamma. If we let $X \sim \text{Gamma}(\alpha, \beta)$ and $Y = X - \alpha\beta$, then

$$
\begin{aligned}
\mathrm{E}(Y) &= 0 \\
\mathrm{E}(Y^2) &= \alpha\beta^2 \\
\mathrm{E}(Y^3) &= 2\alpha\beta^3.
\end{aligned}
$$

The parameters α and β then need to found in order to satisfy

$$\alpha\beta^2 = c_1 \text{ and } 2\alpha\beta^3 = c_2.$$

This yields

$$\beta = \frac{c_2}{2c_1^2} \text{ and } \alpha = \frac{2c_1^3}{c_2^2}.$$

Of course, α and β can only be positive but c_2 can be negative. To take this into account, α and β can be found based on $\mid c_2 \mid$ and then the distribution is reflected about 0 if c_2 is negative.

3.6 Bootstrap Simulation

In this section, we want to compare the Type I error probabilities of the bootstrap procedure to the large sample test. Consider the test statistic mentioned in Section 3.2:

$$T_n = \frac{\sum_{i=1}^{n} \phi(x_i) Y_i}{\sqrt{\hat{\sigma}_I^2 A_n + \hat{\sigma}_\epsilon^2 B_n}}.$$

$Y_1, \ldots, Y_n$ were generated under the assumption that β_0 and β_1 are both 0 for the model

$$Y_i = \beta_0 + \beta_1 x_i + Z_i,$$

where the x_i's were taken to be equally spaced points on (0,1). The distribution of I and ϵ were generated in three ways:

1. $I \sim \sqrt{r} N(0,1)$ and $\epsilon \sim N(0,1)$
2. $I \sim \sqrt{r} \exp(-1,1)$ and $\epsilon \sim \exp(-1,1)$
3. $I \sim \sqrt{r} \exp(-1,1)$ and $\epsilon \sim N(0,1)$

where $\exp(a,b)$ represents a shifted exponential having probability density function

$$b^{-1} e^{(x-a)/b} I_{(a,\infty)}(x), \;\; b > 0.$$

In the cases where I and ϵ are generated from a shifted exponential, the reason for the shift is so that their population means are 0. Other considerations in the simulation were:

- Choices for $\phi(x) : \sin(2\pi x)$ and $x - \bar{x}$.
- Choices for α, the nominal Type I error probability: 0.05 and 0.1.
- When comparing values of r, the sample size was taken to be n=75 since the variable stars to be analyzed later have sample sizes comparable to this.

The simulation procedure was carried out in the following manner:

1. 1000 samples from each of the combinations for the distribution of I and ϵ as mentioned aboved were generated. For each sample:

   (a) The test statistic for the original data, T_n, was computed as well as moment estimates of σ_I^2, σ_ϵ^2, κ_I and κ_ϵ. Denote these estimates as $\tilde{\sigma}_I^2$, $\tilde{\sigma}_\epsilon^2$, $\tilde{\kappa}_I$ and $\tilde{\kappa}_\epsilon$ respectively. See Section 4.1 for a discussion of the exact forms of the moment estimates and their properties.

   (b) For the large sample test, H_0 is rejected if $\mid T_n \mid$ is larger than $\Phi^{-1}(1-\alpha/2)$ which is the $1 - \alpha/2$ percentile of a standard normal distribution.

   (c) For the bootstrap test, H_0 is rejected if $\mid T_n \mid$ is larger than $\mid T_n^* \mid_{1-\alpha}$. Here, $\mid T_n^* \mid_{1-\alpha}$ represents the $(1-\alpha)$ percentile of the bootstrap test statistics generated after taking the absolute value. In this case, 500 bootstrap statistics were generated and therefore, with $\alpha = 0.05$, $\mid T_n \mid$ would be compared to the 475th. Each bootstrap test statistic was computed in the following manner:

      i. Generate $I_1^*, \ldots, I_n^*$ from a shifted gamma distribution with first 3 moments equal to 0, $\tilde{\sigma}_I^2$ and $\tilde{\kappa}_I$.

      ii. Generate $\epsilon_0^*, \ldots, \epsilon_n^*$ from a shifted gamma distribution with first 3 moments equal to 0, $\tilde{\sigma}_\epsilon^2$ and $\tilde{\kappa}_\epsilon$.

iii. Generate a bootstrap sample, $Z_1^*, \ldots, Z_n^*$ as

$$Z_i^* = I_i^* + \epsilon_i^* - \epsilon_{i-1}^*, \quad i = 1, \ldots, n.$$

iv. Compute a single bootstrap test statistic, T_n^*, from (3.3) using the bootstrap sample.

2. Compute the percentage of rejections for the large sample test as well as the bootstrap test. This yields the approximate Type I error rate.

The following tables summarize results of the simulations. For Table 1, $r = 0$, 0.5,1 and $n = 75$. In this case, $\phi(x) = \sin(2\pi x)$ and $\phi(x) = x - \bar{x}$ were both used. The latter corresponds to the Type I error rates in parentheses. Above the columns are the nominal Type I error rates, $\alpha = 0.05$ and $\alpha = 0.1$. In some instances, the Type I error rate for both tests is approximately 2α but in most cases, it is fairly close to α. For the case $I \sim \sqrt{r}$ exp(-1,1) and $\epsilon \sim$ N(0,1) with $r = 0.5$, the Type I error using the standard normal stays at 2α while the bootstrap method is much closer to α. Except for when $r = 0$, the bootstrap probabilities are less than those for the standard normal. Figures 1-3 correspond to quantile plots of the 1000 test statistics generated using $\phi(x) = \sin(2\pi x)$. Figure 1 gives quantile plots of 1000 test statistics in each of the cases: $r = 0, 0.5$ and 1. The three plots support the idea that when r=0, the test statistics do not follow a normal distribution and therefore, the large sample test should not be used. However, when $r = 1$, the percentage of test statistics in the tails that deviate from the standard normal lessens. The quantile plots in Figure 2 behave similarly to those in Figure 1. However, in particular, for the case $r = 1$, there is a larger percentage of test statistics in the tails, likely due to the skewness of the distribution of I. Interestingly, Figure 3 does not show the test statistics becoming closer to the standard normal as was the case in Figures 1 and 2..

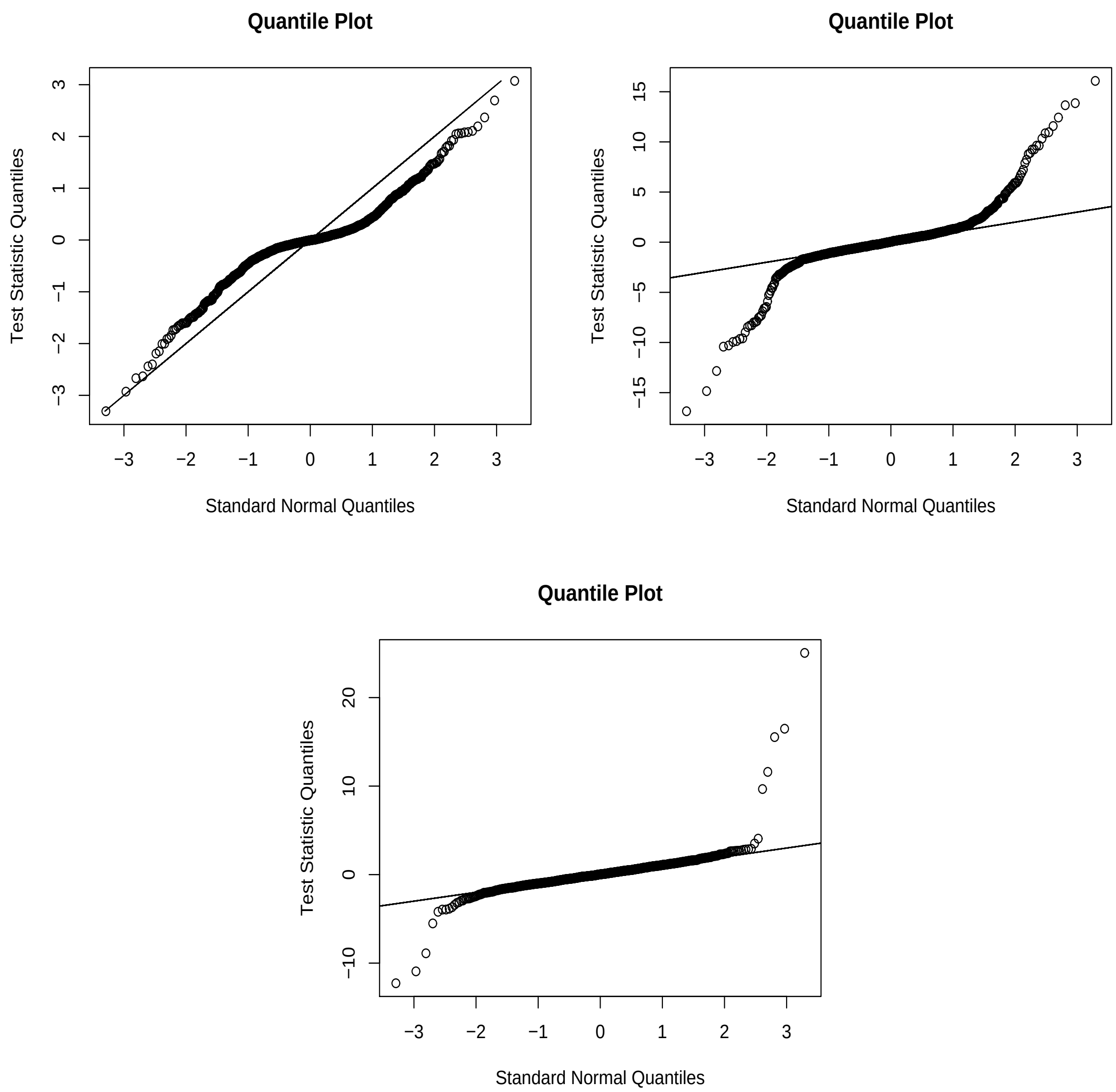

Figure 1. Quantile plots for test statistics. $I \sim \sqrt{r}N(0,1)$ *and* $\epsilon \sim N(0,1)$. *Clockwise from upper left,* r = *0, 0.5 and 1.*

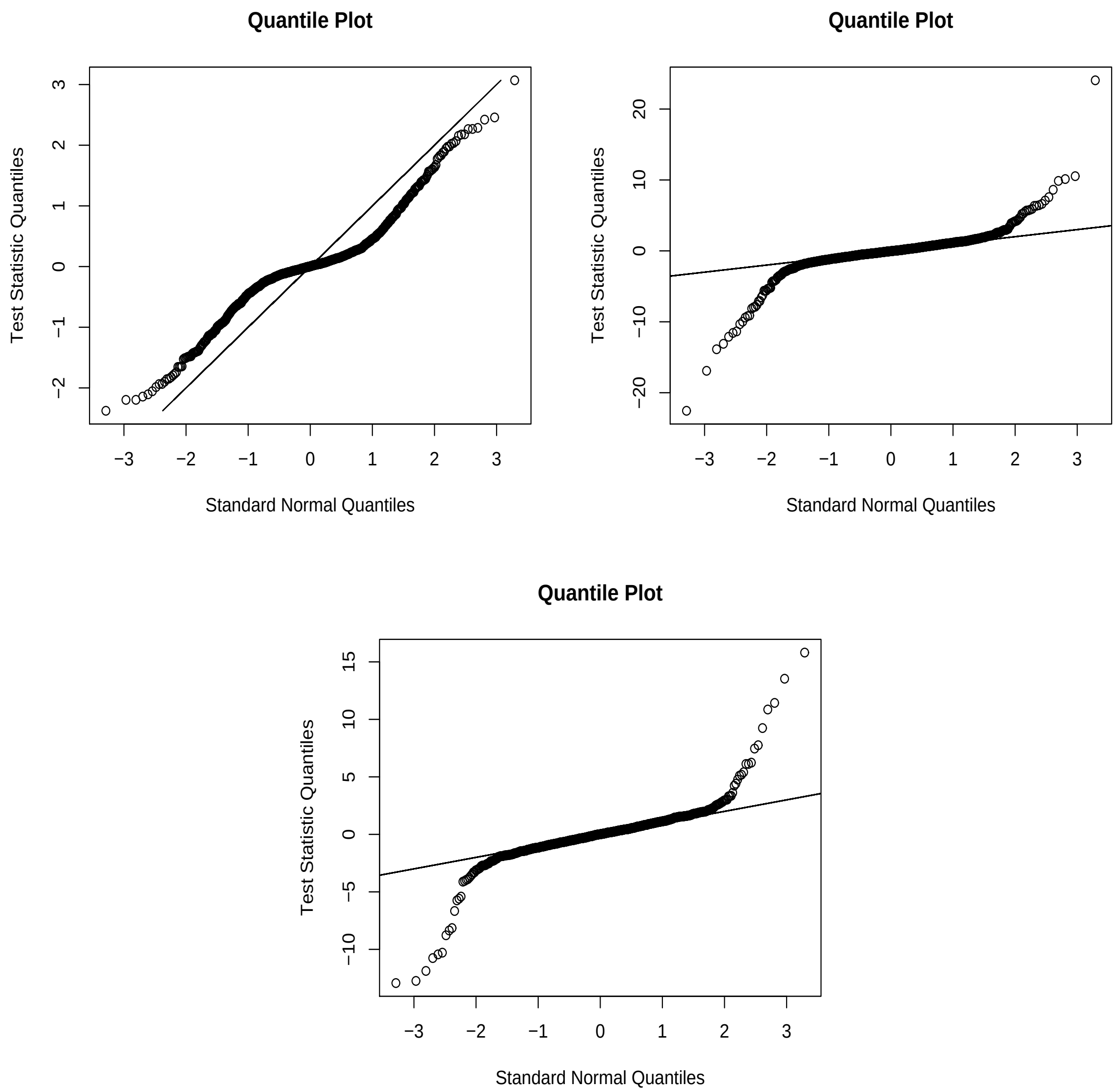

Figure 2. Quantile plots for test statistics. $I \sim \sqrt{r}\exp(-1,1)$ *and* $\epsilon \sim \exp(-1,1)$. *Clockwise from upper left, r = 0, 0.5 and 1.*

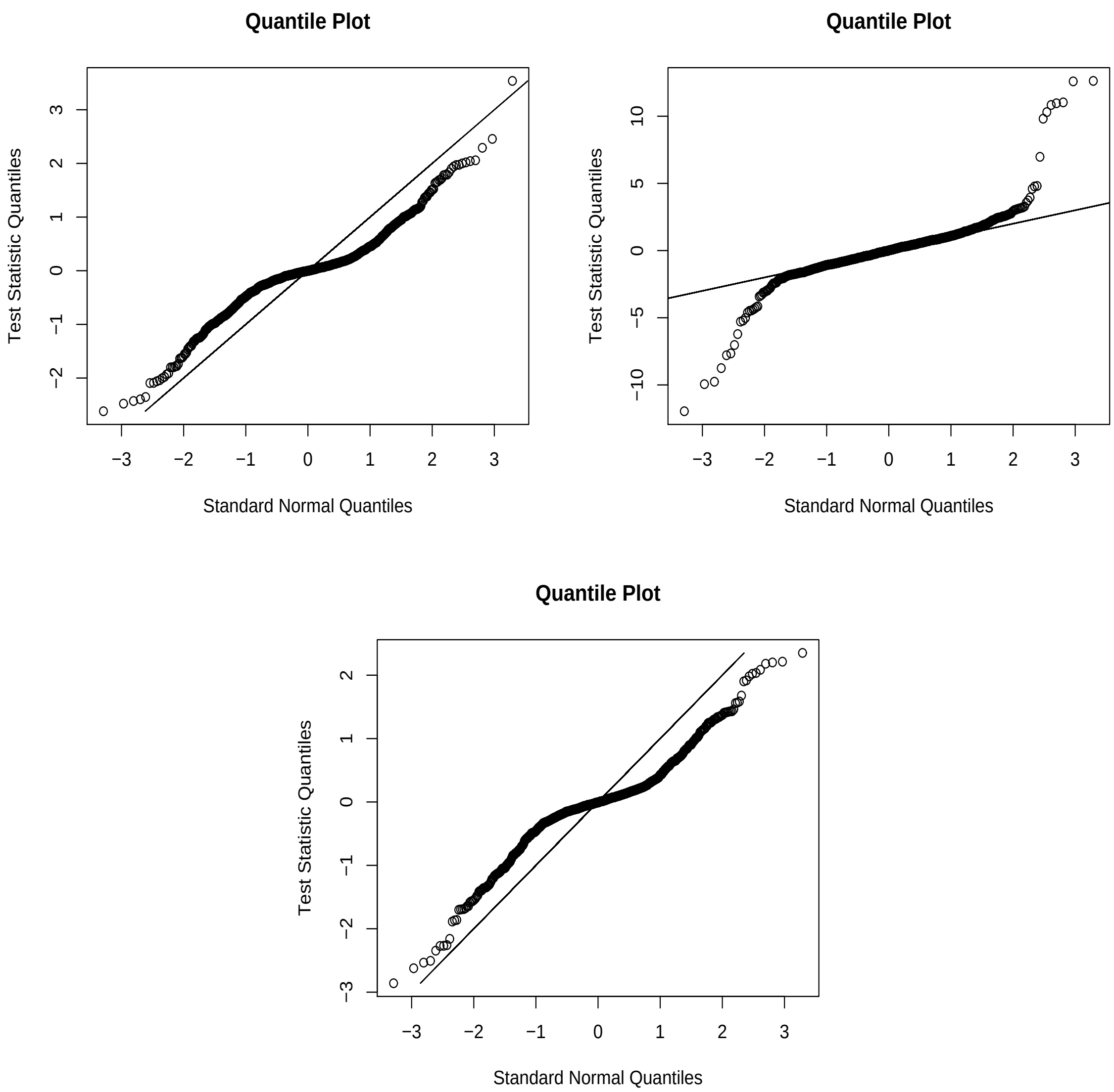

Figure 3. Quantile plots for test statistics. $I \sim \sqrt{r}\exp(-1,1)$ *and* $\epsilon \sim N(0,1)$. *Clockwise from upper left,* $r = 0$, *0.5 and 1.*

Table 1. Simulation results for comparing the bootstrap and large sample test Type I error rates for r = 0, 0.5 and 1.

Dist'n		Bootstrap		N(0,1)	
of I & ϵ	r	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$
$I \sim \sqrt{r}$ N(0,1)	0	0.056	0.098	0.02	0.036
		(0.036)	(0.084)	(0.011)	(0.027)
$\epsilon \sim$ N(0,1)	0.5	0.126	0.154	0.169	0.202
		(0.088)	(0.143)	(0.089)	(0.137)
	1	0.03	0.067	0.078	0.117
		(0.057)	(0.099)	(0.071)	(0.115)
$I \sim \sqrt{r}\,\exp(-1, 1)$	0	0.058	0.107	0.021	0.041
		(0.041)	(0.078)	(0.014)	(0.034)
$\epsilon \sim \exp(-1, 1)$	0.5	0.106	0.13	0.14	0.186
		(0.102)	(0.152)	(0.095)	(0.15)
	1	0.059	0.099	0.098	0.161
		(0.066)	(0.111)	(0.069)	(0.118)
$I \sim \sqrt{r}\,\exp(-1, 1)$	0	0.051	0.089	0.02	0.039
		(0.051)	(0.09)	(0.02)	(0.039)
$\epsilon \sim$ N(0,1)	0.5	0.053	0.097	0.106	0.162
		(0.084)	(0.127)	(0.082)	(0.124)
	1	0.04	0.098	0.017	0.029
		(0.063)	(0.123)	(0.078)	(0.143)

In Table 2, $\phi(x)$ was taken to be $\sin(2\pi x), r = 0$ and $n = 100, 150, 200$. It is of interest to see the effect of sample size on the probabilities. As can be seen in the table below, when n increases, the bootstrap probabilities remain close to α while the large sample results are approximately $\alpha/2$. This illustrates that the bootstrap is the preferred procedure when $r = 0$ as opposed to the large sample test.

Table 2. Simulation results for comparing the bootstrap and large sample test Type I error rates for n = 100, 150 and 200 when r = 0.

Dist'n of I & ϵ	n	Bootstrap $\alpha = 0.05$	Bootstrap $\alpha = 0.1$	N(0,1) $\alpha = 0.05$	N(0,1) $\alpha = 0.1$
$I \sim \sqrt{r}N(0,1)$	100	0.046	0.107	0.013	0.033
$\epsilon \sim N(0,1)$	150	0.058	0.117	0.017	0.042
	200	0.048	0.099	0.018	0.037
$I \sim \sqrt{r}\exp(-1,1)$	100	0.05	0.093	0.013	0.036
$\epsilon \sim \exp(-1,1)$	150	0.055	0.102	0.017	0.045
	200	0.049	0.091	0.019	0.038
$I \sim \sqrt{r}\exp(-1,1)$	100	0.057	0.101	0.024	0.045
$\epsilon \sim N(0,1)$	150	0.05	0.102	0.018	0.041
	200	0.057	0.103	0.022	0.045

Table 3 is from a simulation to determine the effect of increasing r away from 0. $\phi(x)$ is taken as $\sin(2\pi x), n = 75$ and $r = 2, 5, 10$. As the table shows, all probabilities are approximately α, but also the bootstrap probabilities are all smaller than those from the standard normal, and in some instances, the large sample test is closer to α. This is likely due to the fact that the lag 1 correlation decreases to 0 as r becomes larger.

Table 3. Simulation results for comparing the bootstrap and large sample test Type I error rates for r = 2, 5, and 10.

Dist'n		Bootstrap		N(0,1)	
of I & ϵ	r	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$
I $\sim \sqrt{r}N(0,1)$	2	0.029	0.084	0.077	0.123
$\epsilon \sim N(0,1)$	5	0.033	0.079	0.042	0.093
	10	0.043	0.091	0.054	0.104
$I \sim \sqrt{r}\exp(-1,1)$	2	0.026	0.075	0.06	0.12
$\epsilon \sim \exp(-1,1)$	5	0.045	0.097	0.059	0.107
	10	0.039	0.088	0.046	0.108
$I \sim \sqrt{r}\exp(-1,1)$	2	0.029	0.07	0.059	0.111
$\epsilon \sim \exp(-1,1)$	5	0.038	0.077	0.05	0.091
	10	0.037	0.089	0.05	0.105

CHAPTER IV

ESTIMATION

Based on the Edgeworth expansion in Section 3.3, the only moments of I and ϵ that are of particular interest are the second and third moments: σ_I^2, σ_ϵ^2, κ_I and κ_ϵ. These parameters need to be estimated as efficiently as possible to effectively bootstrap the test statistic.

4.1 Method of Moments

For method of moments, population moments are set equal to sample moments, resulting in a set of equations which can be used to solve for the parameters of interest. In this case, we note first that

$$\begin{aligned}
\mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n} Z_i^2\right) &= \sigma_I^2 + 2\sigma_\epsilon^2 \\
\mathrm{E}\left(\frac{1}{n}\sum_{i=2}^{n} Z_i Z_{i-1}\right) &= -\sigma_\epsilon^2 \\
\mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n} Z_i^3\right) &= \kappa_I \\
\text{and}\quad \mathrm{E}\left(\frac{1}{n}\sum_{i=2}^{n} Z_i^2 Z_{i-1}\right) &= \kappa_\epsilon .
\end{aligned}$$

Note that Z_i must be replaced by residuals $\hat{Z}_i$ to obtain the following moment estimates:

$$\tilde{\sigma}_\epsilon^2 = -\frac{1}{n}\sum_{i=2}^{n} \hat{Z}_i \hat{Z}_{i-1},$$

$$\begin{aligned}
\tilde{\sigma}_I^2 &= \frac{1}{n}\sum_{i=1}^{n}\hat{Z}_i^2 + \frac{2}{n}\sum_{i=2}^{n}\hat{Z}_i\hat{Z}_{i-1}, \\
\tilde{\kappa}_I &= \frac{1}{n}\sum_{i=1}^{n}\hat{Z}_i^3, \\
\tilde{\kappa}_\epsilon &= \frac{1}{n}\sum_{i=2}^{n}\hat{Z}_i^2\hat{Z}_{i-1}.
\end{aligned}$$

and thus $\tilde{r} = \tilde{\sigma}_I^2/\tilde{\sigma}_\epsilon^2$.

The advantage of using moment estimates is that they are $\sqrt{n}$ consistent and easy to compute. Two disadvantages are that we are not guaranteed to obtain estimates in the parameter space and moment estimators are often inefficient. The third population moments, κ_I and κ_ϵ, may be negative but σ_ϵ^2 and σ_I^2 cannot be. Therefore, we need to modify the moment estimates for σ_ϵ^2 and σ_I^2 in the following manner:

$$\begin{aligned}
\tilde{\sigma}_\epsilon^2 &= \max(0, -\frac{1}{n}\sum_{i=2}^{n}\hat{Z}_i\hat{Z}_{i-1}) \\
\tilde{\sigma}_I^2 &= \max(0, \frac{1}{n}\sum_{i=1}^{n}\hat{Z}_i^2 - 2\tilde{\sigma}_\epsilon^2).
\end{aligned}$$

4.2 Consistency of Estimates

Method of moments estimates are guaranteed to be consistent when the corresponding population moments exist. Furthermore, they are functions of quantities that are sample means. However, it is not clear whether $\hat{r}$, $\hat{\sigma}_\epsilon^2$ and $\hat{\boldsymbol{\beta}}$ are consistent or not. Maximum likelihood estimates are special cases of M-estimators. Furrer (2002) discusses four hypotheses necessary for M-estimators to be consistent for dependent random variables. To establish consistency for $\hat{r}$ and other estimates, the hypotheses stated by Furrer will be verified. In addition to the model assumptions already stated in Chapter I, the only conditions needed to establish consistency are the following:

1. $\mathrm{E}(I^{2\alpha}) < \infty$ for some $\alpha > 1$.

2. $\mathrm{E}(\epsilon^{2\beta}) < \infty$ for some $\beta > 1$.

3. $\sigma_I^2 > 0$.

4. The ith column of X is $U_i(x_1), \ldots, U_i(x_n)$, where $U_1, \ldots, U_m$ are orthogonal polynomials.

Define $G_n(r, \sigma_\epsilon^2, \boldsymbol{\beta}) = l(r, \sigma_\epsilon^2, \boldsymbol{\beta})/n$, where

$$l(r, \sigma_\epsilon^2, \boldsymbol{\beta}) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma_\epsilon^2) - \frac{1}{2}\log(|\, A\,|) - \frac{1}{2\sigma_\epsilon^2}(\mathbf{Y} - X\boldsymbol{\beta})' A^{-1}(\mathbf{Y} - X\boldsymbol{\beta}),$$

and let $\mathbf{Z} = \mathbf{Y} - X\boldsymbol{\beta}$. In Furrer's paper, he mentions that that the parameters must be in an open interval. The implications of this are that his results cannot be applied to the case $r = 0$. Aside from this detail, there are four hypotheses that need to be satisfied for the mles to be consistent:

1. G_n is almost surely twice differentiable with respect to $\boldsymbol{\theta} = r, \sigma_\epsilon^2, \boldsymbol{\beta}$.

2. $\partial G_n / \partial \boldsymbol{\theta}_i \longrightarrow 0$ almost surely.

3. For $i, j = 1, \ldots, n$, $H_n^{ij}(\boldsymbol{\theta}) = \partial G_n / \partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_j$ is almost surely L_n^{ij}-Lipschitz continuous in $\boldsymbol{\theta}$.

4. The matrix $H(\boldsymbol{\theta})$ converges almost surely to a negative definite matrix.

The first hypothesis is verified since the first and second partial derivatives of G_n with respect to $\boldsymbol{\theta}$ exist. They are

$$\begin{aligned}\frac{\partial G_n}{\partial \sigma_\epsilon^2} &= -\frac{1}{2\sigma_\epsilon^2} + \frac{1}{2n\sigma_\epsilon^4}(\mathbf{Y} - X\boldsymbol{\beta})' A^{-1}(\mathbf{Y} - X\boldsymbol{\beta}) \\ &= -\frac{1}{2\sigma_\epsilon^2} + \frac{1}{2n\sigma_\epsilon^4}\mathbf{Z}' A^{-1}\mathbf{Z},\end{aligned}$$

$$\frac{\partial G_n}{\partial \boldsymbol{\beta}} = \frac{1}{n\sigma_\epsilon^2} X' A^{-1}\mathbf{Z},$$

$$\frac{\partial G_n}{\partial r} = -\frac{1}{2n}\mathrm{tr}(A^{-1}) + \frac{1}{2n\sigma_\epsilon^2}\mathbf{Z}' A^{-2}\mathbf{Z},$$

$$\frac{\partial^2 G_n}{\partial \sigma_\epsilon^4} = \frac{1}{2\sigma_\epsilon^4} - \frac{1}{n\sigma_\epsilon^6}\mathbf{Z}' A^{-1}\mathbf{Z},$$

$$\frac{\partial^2 G_n}{\partial \boldsymbol{\beta}^2} = -\frac{1}{n\sigma_\epsilon^2} X' A^{-1} X,$$

$$\frac{\partial^2 G_n}{\partial r^2} = \frac{1}{2n}\mathrm{tr}(A^{-2}) - \frac{1}{n\sigma_\epsilon^2}\mathbf{Z}' A^{-3}\mathbf{Z},$$

$$\frac{\partial^2 G_n}{\partial \sigma_\epsilon^2 \partial \boldsymbol{\beta}} = -\frac{1}{n\sigma_\epsilon^4} X' A^{-1}\mathbf{Z},$$

$$\frac{\partial^2 G_n}{\partial \sigma_\epsilon^2 \partial r} = -\frac{1}{2n\sigma_\epsilon^4}\mathbf{Z}' A^{-2}\mathbf{Z}$$

$$\text{and} \quad \frac{\partial^2 G_n}{\partial \boldsymbol{\beta} \partial r} = -\frac{1}{n\sigma_\epsilon^2} X' A^{-2}\mathbf{Z}.$$

His second hypothesis is that the first partial derivatives evaluated at the true parameter values converge almost surely to 0 as $n \to \infty$. Appendices A and B give proofs that the linear and quadratic forms of interest converge almost surely to their means which are

$$
\begin{aligned}
\mathrm{E}\left(\frac{\partial G_n}{\partial \sigma_\epsilon^2}\right) &= \mathrm{E}\Big(-\frac{1}{2\sigma_\epsilon^2} + \frac{1}{2n\sigma_\epsilon^4}\mathbf{Z}'A^{-1}\mathbf{Z}\Big) \\
&= -\frac{1}{2\sigma_\epsilon^2} + \frac{1}{2n\sigma_\epsilon^4}\sigma_\epsilon^2 \mathrm{tr}(A^{-1}A) \\
&= -\frac{1}{2\sigma_\epsilon^2} + \frac{1}{2\sigma_\epsilon^2} \\
&= 0,
\end{aligned}
$$

$$
\mathrm{E}\left(\frac{\partial G_n}{\partial \boldsymbol{\beta}}\right) = \mathrm{E}\Big(\frac{1}{n\sigma_\epsilon^2}X'A^{-1}\mathbf{Z}\Big) = 0,
$$

and finally

$$
\begin{aligned}
\mathrm{E}\left(\frac{\partial G_n}{\partial r}\right) &= \mathrm{E}\Big(-\frac{1}{2n}\mathrm{tr}(A^{-1}) + \frac{1}{2n\sigma_\epsilon^2}\mathbf{Z}'A^{-2}\mathbf{Z}\Big) \\
&= -\frac{1}{2n}\mathrm{tr}(A^{-1}) + \frac{1}{2n\sigma_\epsilon^2}\sigma_\epsilon^2\mathrm{tr}(A^{-2}A) \\
&= -\frac{1}{2n}\mathrm{tr}(A^{-1}) + \frac{1}{2n}\mathrm{tr}(A^{-1}) \\
&= 0.
\end{aligned}
$$

$$
\begin{aligned}
\text{Denote } H(\boldsymbol{\theta}) &= \begin{pmatrix} \partial^2 G_n/\partial\boldsymbol{\beta}^2 & \partial^2 G_n/\partial\sigma_\epsilon^2\partial\boldsymbol{\beta} & \partial^2 G_n/\partial r\partial\boldsymbol{\beta} \\ \partial^2 G_n/\partial\sigma_\epsilon^2\partial\boldsymbol{\beta} & \partial^2 G_n/\partial\sigma_\epsilon^4 & \partial^2 G_n/\partial\sigma_\epsilon^2\partial r \\ \partial^2 G_n/\partial\sigma_\epsilon^2\partial r & \partial^2 G_n/\partial r\partial\boldsymbol{\beta} & \partial^2 G_n/\partial r^2 \end{pmatrix} \\
&= \begin{pmatrix} -\frac{1}{n\sigma_\epsilon^2}(X'A^{-1}X) & -\frac{1}{n\sigma_\epsilon^2}X'A^{-1}\mathbf{Z} & -\frac{1}{n\sigma_\epsilon^2}X'A^{-2}\mathbf{Z} \\ -\frac{1}{n\sigma_\epsilon^2}X'A^{-1}\mathbf{Z} & \frac{1}{2\sigma_\epsilon^4} - \frac{1}{n\sigma_\epsilon^6}\mathbf{Z}'A^{-1}\mathbf{Z} & -\frac{1}{2n\sigma_\epsilon^4}\mathbf{Z}'A^{-2}\mathbf{Z} \\ -\frac{1}{n\sigma_\epsilon^2}X'A^{-2}\mathbf{Z} & -\frac{1}{2n\sigma_\epsilon^4}\mathbf{Z}'A^{-2}\mathbf{Z} & \frac{1}{2n}\mathrm{tr}(A^{-2}) - \frac{1}{n\sigma_\epsilon^2}\mathbf{Z}'A^{-3}\mathbf{Z} \end{pmatrix}
\end{aligned}
$$

To show that the third hypothesis is satisfied, it is sufficient to show that $\partial(H(\boldsymbol{\theta}))/\partial\boldsymbol{\theta}$ exists and is finite almost surely.

$$\partial H_n(\boldsymbol{\theta})/\partial\boldsymbol{\beta} = \begin{pmatrix} 0 & \frac{1}{n\sigma_\epsilon^2}X'A^{-1}X & \frac{1}{n\sigma_\epsilon^2}X'A^{-2}X \\ \frac{1}{n\sigma_\epsilon^2}X'A^{-1}X & \frac{2}{n\sigma_\epsilon^6}X'A^{-1}\mathbf{Z} & \frac{1}{2n\sigma_\epsilon^4}X'A^{-2}\mathbf{Z} \\ \frac{1}{n\sigma_\epsilon^2}X'A^{-2}X & \frac{1}{2n\sigma_\epsilon^4}X'A^{-2}\mathbf{Z} & \frac{1}{n\sigma_\epsilon^2}X'A^{-3}\mathbf{Z} \end{pmatrix},$$

$$\partial H_n(\boldsymbol{\theta})/\partial\sigma_\epsilon^2 = \begin{pmatrix} \frac{1}{n\sigma_\epsilon^4}(X'A^{-1}X) & \frac{1}{n\sigma_\epsilon^4}X'A^{-1}\mathbf{Z} & \frac{1}{n\sigma_\epsilon^4}X'A^{-2}\mathbf{Z} \\ \frac{1}{n\sigma_\epsilon^4}X'A^{-1}\mathbf{Z} & -\frac{1}{\sigma_\epsilon^6} + \frac{3}{n\sigma_\epsilon^6}\mathbf{Z}'A^{-1}\mathbf{Z} & \frac{1}{n\sigma_\epsilon^6}\mathbf{Z}'A^{-2}\mathbf{Z} \\ \frac{1}{n\sigma_\epsilon^4}X'A^{-2}\mathbf{Z} & \frac{1}{n\sigma_\epsilon^6}\mathbf{Z}'A^{-2}\mathbf{Z} & \frac{1}{n\sigma_\epsilon^4}\mathbf{Z}'A^{-3}\mathbf{Z} \end{pmatrix},$$

$$\text{and } \partial H_n(\boldsymbol{\theta})/\partial r = \begin{pmatrix} \frac{1}{n\sigma_\epsilon^2}(X'A^{-2}X) & \frac{1}{n\sigma_\epsilon^2}X'A^{-2}\mathbf{Z} & \frac{2}{n\sigma_\epsilon^2}X'A^{-3}\mathbf{Z} \\ \frac{1}{n\sigma_\epsilon^2}X'A^{-2}\mathbf{Z} & \frac{1}{\sigma_\epsilon^6} - \frac{1}{n\sigma_\epsilon^6}\mathbf{Z}'A^{-2}\mathbf{Z} & \frac{1}{2n\sigma_\epsilon^4}\mathbf{Z}'A^{-3}\mathbf{Z} \\ \frac{2}{n\sigma_\epsilon^2}X'A^{-3}\mathbf{Z} & \frac{1}{n\sigma_\epsilon^4}\mathbf{Z}'A^{-3}\mathbf{Z} & \frac{-3}{n\sigma_\epsilon^2}\mathbf{Z}'A^{-4}\mathbf{Z} \end{pmatrix}.$$

The elements in the above matrices have 3 general forms: $\frac{1}{n}X'A^{-k}\mathbf{Z}$, $\frac{1}{n}\mathbf{Z}'A^{-k}\mathbf{Z}$ and $\frac{1}{n}X'A^{-k}X$. Appendices B and C show that $\frac{1}{n}X'A^{-k}\mathbf{Z}$ and $\frac{1}{n}\mathbf{Z}'A^{-k}\mathbf{Z}$ both converge almost surely to their respective means. To show that the ijth element of $\frac{1}{n}X'A^{-k}X$ is less than ∞ follows the same argument as in Appendix B for showing $\frac{1}{n}X'A^{-k}\mathbf{Z}$ converges almost surely. That is, the ijth element of $\frac{1}{n}X'A^{-k}X$ can be expressed as

$$\frac{1}{n}\sum_{r=1}^{n}\sum_{s=1}^{n}U_i(x_r)[A^{-k}]_{rs}U_j(x_s).$$

By assumption, U_i and U_j are orthogonal polynomials, and hence $\mid U_i(x) \mid < c < \infty$ and $\mid U_j(x) \mid < c$ for all $x \in (0,1)$. It then follows that the ijth element is bounded in absolute value by

$$\frac{c^2}{n}\sum_{r=1}^{n}\sum_{s=1}^{n}[A^{-k}]_{rs}$$

since the elements of A^{-1} are all positive. The limit as $n \to \infty$ of the last quantity is less than ∞ when $k = 1$ from results in Appendix A. A similar proof can be obtained for when $k = 2$.

To show that the last hypothesis holds, first take the expectation of the elements of $H(\boldsymbol{\theta}))$ which yields.

$$\mathrm{E}(H(\boldsymbol{\theta})) = \begin{pmatrix} -\frac{1}{n\sigma_\epsilon^2}(X'A^{-1}X) & 0 & 0 \\ 0 & \frac{-1}{2\sigma_\epsilon^4} & \frac{-1}{2n\sigma_\epsilon^2}\mathrm{tr}(A^{-1}) \\ 0 & \frac{-1}{2n\sigma_\epsilon^2}\mathrm{tr}(A^{-1}) & \frac{-1}{2n}\mathrm{tr}(A^{-2}) \end{pmatrix}.$$

To show that $\mathrm{E}(H(\boldsymbol{\theta}))$ is negative definite is equivalent to showing that $-\mathrm{E}(H(\boldsymbol{\theta}))$ is positive definite. The determinant of $-\mathrm{E}(H(\boldsymbol{\theta}))$ is

$$\begin{aligned} & \left|\frac{1}{n\sigma_\epsilon^2}(X'A^{-1}X)\right|\left(\frac{1}{2\sigma_\epsilon^4}\Big(\frac{1}{2n}\mathrm{tr}(A^{-2})\Big) - \Big(\frac{1}{2n\sigma_\epsilon^2}\mathrm{tr}(A^{-1})\Big)^2\right) \\ = & \left|\frac{1}{n\sigma_\epsilon^2}(X'A^{-1}X)\right|\left(\frac{1}{4n\sigma_\epsilon^4}\mathrm{tr}(A^{-2}) - \frac{1}{4n^2\sigma_\epsilon^4}\Big(\mathrm{tr}(A^{-1})\Big)^2\right) \end{aligned}$$

and so it only needs to be shown that

1. $\mid -\mathrm{E}(H(\boldsymbol{\theta})) \mid \longrightarrow$ positive constant

2. $\frac{1}{4n\sigma_\epsilon^4}\mathrm{tr}(A^{-2}) - \frac{1}{4n\sigma_\epsilon^4}\mathrm{tr}(A^{-1})^2 \longrightarrow$ positive constant

3. $\frac{1}{2n}\mathrm{tr}(A^{-2}) \longrightarrow$ positive constant

It is straightforward to show that

$$\frac{1}{n}X'A^{-1}X \rightarrow \frac{1+e^{-\lambda}}{(1-e^{-\lambda})(e^{\lambda}-e^{-\lambda})}\mathbf{I}_{p+1}$$

assuming the columns of X are orthogonal polynomials and design points are evenly spaced. $\mathbf{I}_{p+1}$ is a $p+1$ by $p+1$ Identity matrix and $\lambda = \cosh^{-1}((r+2)/2)$ as defined in Appendix A. This along with 2. from above implies 1. For a more detailed proof, see Appendix A.

Below are proofs of 2. and 3. The expression for the ijth element of A^{-1} can be found in Appendix A.

$$\begin{aligned}
\frac{1}{n}\mathrm{tr}(A^{-2}) &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}[A^{-1}]_{ij}^{2} \\
&= \frac{1}{n}\Big(\frac{\sinh(i\lambda)\sinh((n-i+1)\lambda)}{\sinh(\lambda)\sinh((n+1)\lambda)}\Big)^{2} \\
&\qquad + \frac{2}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{\sinh(i\lambda)\sinh((n-j+1)\lambda)}{\sinh(\lambda)\sinh((n+1)\lambda)} \\
&\longrightarrow (e^{\lambda}-e^{-\lambda})^{2} + 2(e^{\lambda}-e^{-\lambda})(e^{2\lambda}-1) \\
&= (e^{\lambda}-e^{-\lambda})^{-2}\Big(\frac{1+e^{2\lambda}}{e^{2\lambda}-1}\Big).
\end{aligned}$$

Therefore,

$$\begin{aligned}
& \frac{1}{n}\mathrm{tr}(A^{-2}) - \frac{1}{n^2}(\mathrm{tr}(A^{-1})^2 \\
\longrightarrow \quad & (e^{\lambda} - e^{-\lambda})^{-2}\Big(\frac{1+e^{2\lambda}}{e^{2\lambda}-1}\Big) - (e^{\lambda} - e^{-\lambda})^{-2} \\
= \quad & (e^{\lambda} - e^{-\lambda})^{-2}\Big(\frac{1+e^{2\lambda}-(e^{2\lambda}-1)}{e^{2\lambda}-1}\Big) \\
= \quad & \frac{2}{(e^{\lambda} - e^{-\lambda})^{2}(e^{2\lambda}-1)} \\
> \quad & 0.
\end{aligned}$$

CHAPTER V

APPLICATION OF THE BOOTSTRAP PROCEDURE

In this chapter, the bootstrap procedure is applied to data from five stars. These stars are the same ones analyzed by Koen & Lombard (2001). They are R. Aquilae, R. Bootis, R. Canum Venaticorum, W. Draconis and Y. Aquarii. The following is an outline of the bootstrap procedure used to analyze the data for the 5 stars:

1. Fit a sixth degree polynomial to the original data. Note: In testing for trend, our model for the mean does not have to be correct in order for the test to be valid or powerful. So long as the true degree is less than six, the coefficients will be unbiased. A degree of six usually suffices for purposes of power.

2. The restricted and unrestricted maximum likelihood estimates are used to obtain the test statistic for the original data. The form of the test statistic is the same as (3.2).

3. Perform the following bootstrap as many times as desired. In this case, the number of bootstrap replications was 1000.

   (a) Generate data for I and ϵ whose first 3 population moments correspond to the first 3 estimates of the moments from the original data. The bootstrap distribution used here is a mixture of normals.

   (b) Based on the data that were generated, obtain the test statistic as was done for the original data.

4. Approximate the p-value by the proportion of all bootstrap statistics that exceed the actual test statistic.

5.1 R. Aquilae

The variable star R. Aquilae has 86 observations. Figure 4 is a plot of the period versus the time of maximum brightness. The solid line is the result of fitting a sixth degree polynomial. The likelihood ratio test statistic is 118.35. The p-value is less than 0.0001. In comparison, the p-value based on a χ^2_d distribution with d=6 is also less than 0.0001. Figure 5 is a density estimate based on the 1000 bootstrapped test statistics. A χ^2_6 density is plotted for comparison. It is known that the mode, or the point at which the χ^2 achieves a maximum is $M-2$ where M is the degrees of freedom. In the plot, χ^2_M is also plotted for comparison. M was chosen by determining the approximate value at which the maximum is attained from the density of the bootstrapped test statistics. In this case, $M \approx 6.38$. From the figure, it is apparent that although the mode is approximately the same for the density of the bootstrapped test statistics and a χ^2_6 distribution, there is, however, quite a discrepency between the distributions. Figure 6 provides a quantile plot of the 1000 bootstrapped test statistics generated to further compare the test statistic from the original data to the bootstrapped distribution. In this case, the original test statistic is much larger than even the 99th percentile in this case. The maximum likelihood estimates in Table 4 help to illustrate which coefficients are possibly significant. Obviously the standardized estimate of β_1 stands out the most as it should judging from Figure 4. The standard errors are the square root of the diagonal elements of the matrix

$$\hat{\sigma}^2_\epsilon(X'\hat{A}^{-1}X)^{-1}.$$

Table 4. Maximum likelihood estimates and standard errors of the regression coefficients for R. Aquilae.

coefficient	estimate	standard error	standardized estimate
β_1	-96.1764	3.5085	-27.4122
β_2	5.6147	3.6058	1.5571
β_3	-18.7091	3.6755	-5.0902
β_4	-1.5182	3.7644	-0.4033
β_5	-0.9508	3.7893	-0.2509
β_6	1.0282	3.8641	0.2661

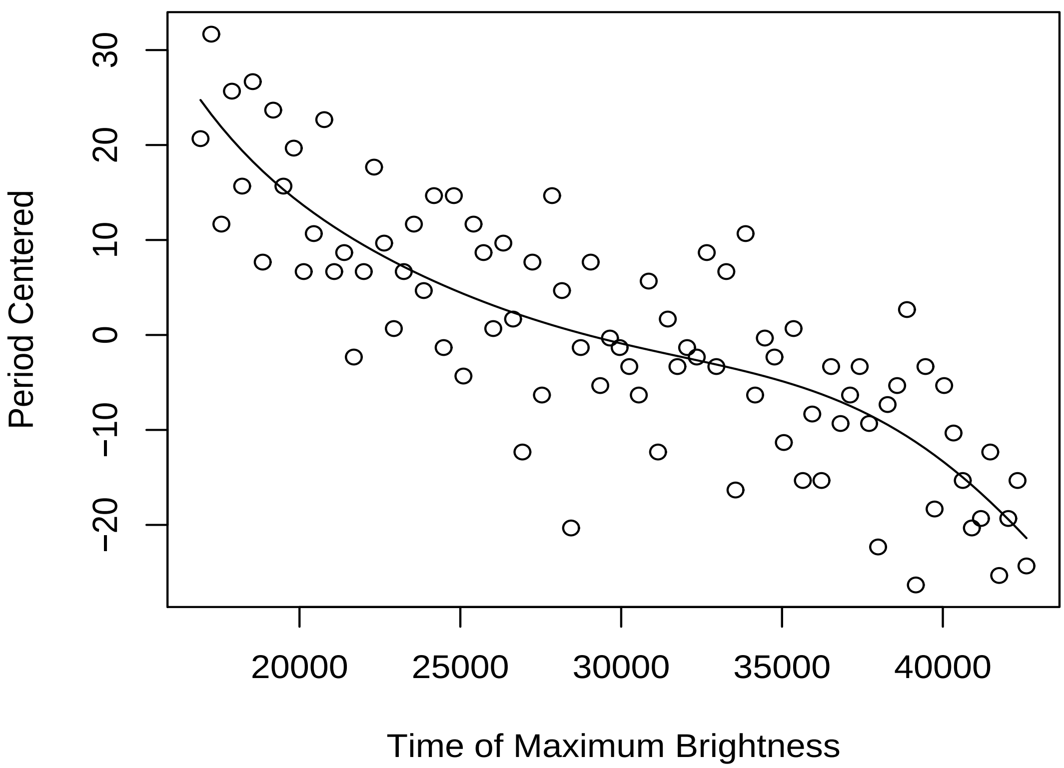

Figure 4. Plot of pseudo-periods for R. Aquilae.

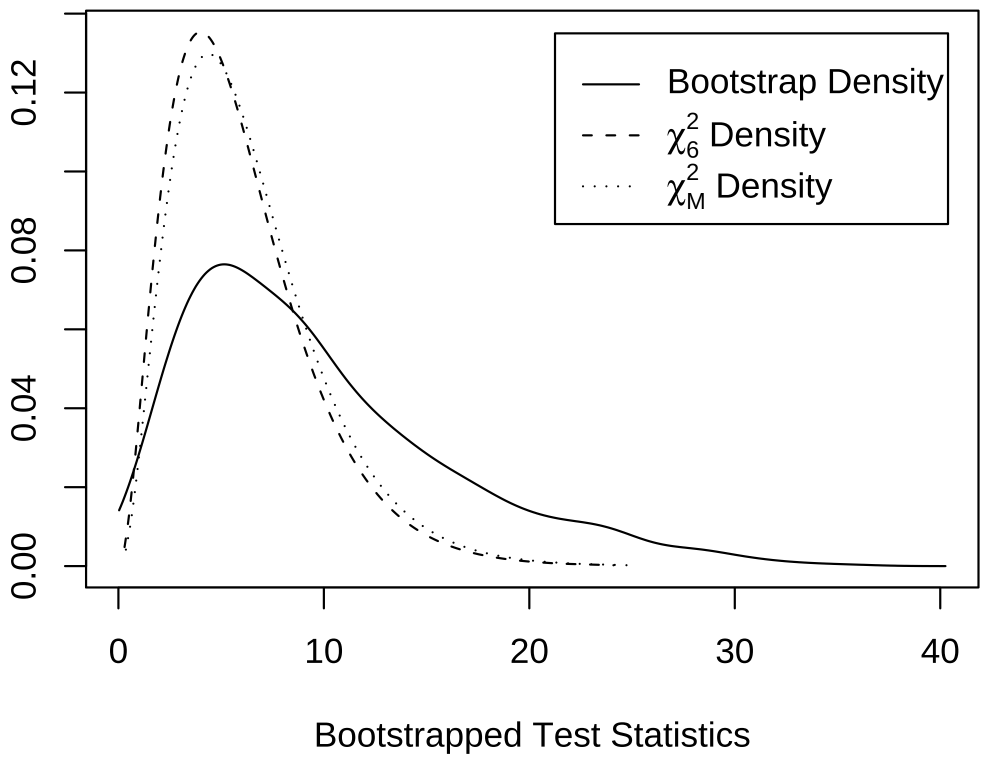

Figure 5. Density estimate of the bootstrapped test statistic for R. Aquilae.

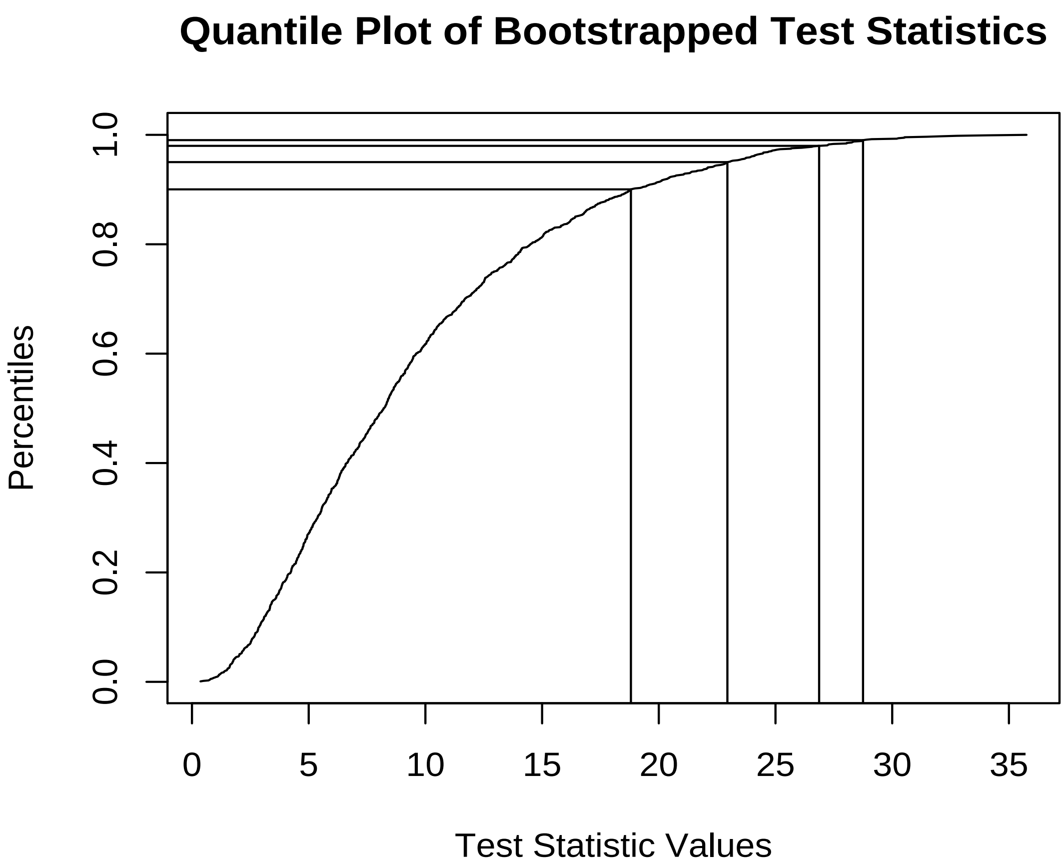

Figure 6. *Quantile plot of the bootstrapped test statistic for R. Aquilae. Vertical and horizontal lines mark the 90th, 95th, 98th and 99th percentiles of the bootstrapped test statistic.*

5.2 R. Bootis

The variable star R. Bootis has 115 observations. Figure 7 is a plot of the pseudo-period versus the time of maximum brightness. The solid line is the result of fitting a sixth degree polynomial. The observed test statistic obtained is 15.3238. The p-value in this case was 0.1997 whereas the p-value from χ^2_6 is 0.0179. Figure 8 illustrates the possible reason for the difference in p-values. The value of M in this case is approximately 6.48. The quantile plot in Figure 9 further illustrates that the observed test statistic is not extremely large, in fact less than the 90th percentile. The maximum likelihood estimates in Table 5 illustrate which coefficients might be significant.

Table 5. Maximum likelihood estimates and standard errors of the regression coefficients for R. Bootis.

coefficient	estimate	standard error	standardized estimate
β_1	1.183	1.9646	0.6022
β_2	-1.8736	2.0508	-0.9136
β_3	0.477	2.1058	0.2265
β_4	2.6005	2.1817	1.192
β_5	-8.8203	2.1851	-4.0366
β_6	5.9707	2.2441	2.6607

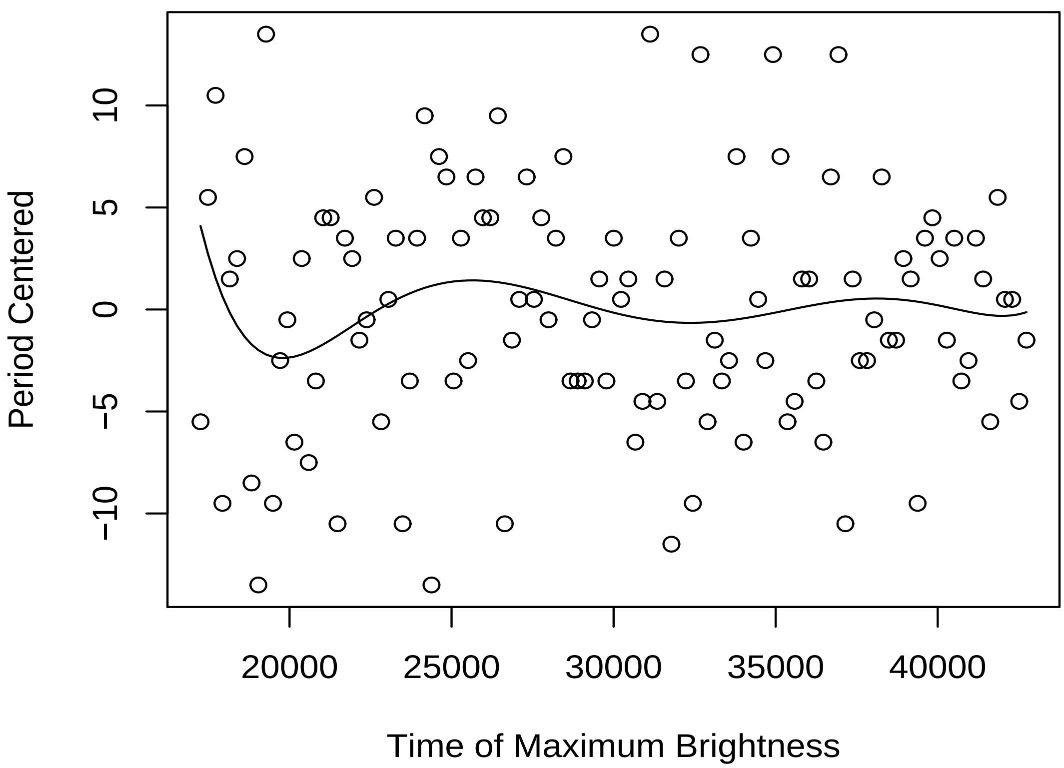

Figure 7. Plot of pseudo-periods for R. Bootis.

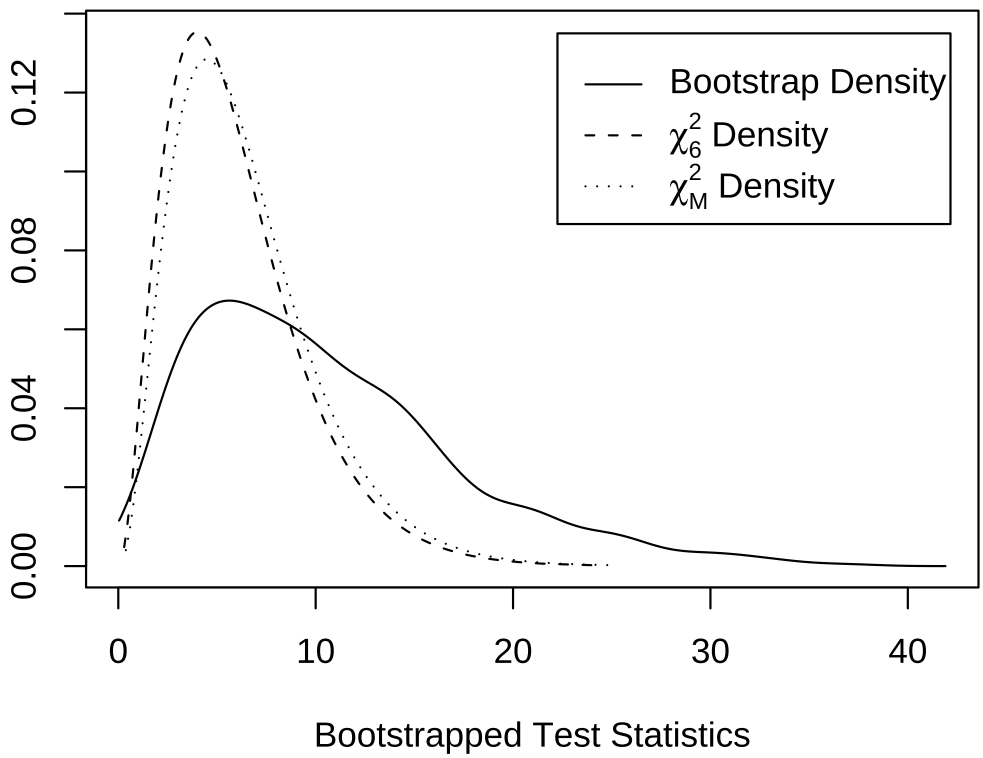

Figure 8. Density estimate of the bootstrapped test statistic for R. Bootis.

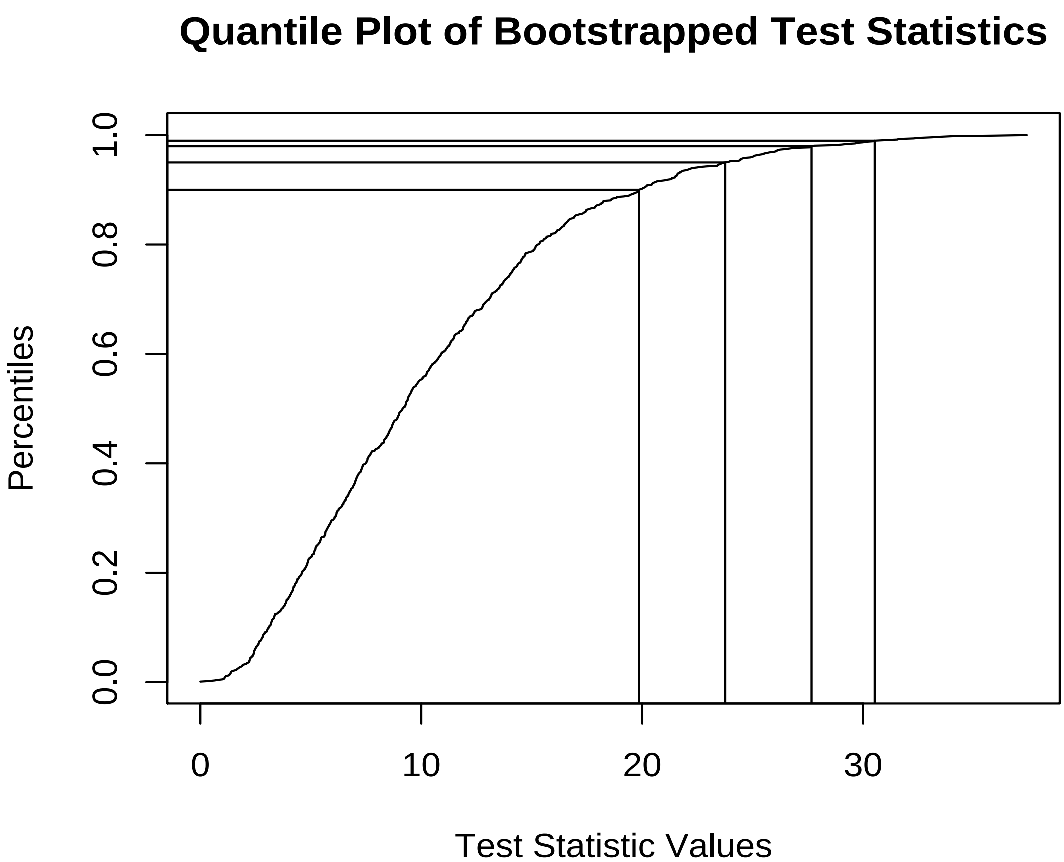

Figure 9. *Quantile plot of the bootstrapped test statistic for R. Bootis. Vertical and horizontal lines mark the 90th, 95th, 98th and 99th percentiles of the bootstrapped test statistic.*

5.3 R. Canum Venaticorum

The variable star R. Canum Venaticorum has 78 observations. Figure 10 is a plot of the period versus the time of maximum brightness. The solid line is the result of fitting a sixth degree polynomial. The test statistic obtained for the original data was 13.7379 and the p-value was found to be 0.1381. The corresponding p-value from χ^2_6 is 0.0327. Figure 11 shows larger values in the tail of the bootstrapped density than would is the case for a χ^2_6 distribution. The value of M in this case is approximately 6.41, again close to 6. In Figure 12, we see that the test statistic is slightly less than the 90th percentile. The maximum likelihood estimates are given in Table 6 below:

Table 6. Maximum likelihood estimates and standard errors of the regression coefficients for R. Canum Venaticorum.

coefficient	estimate	standard error	standardized estimate
β_1	10.597	5.9608	1.7778
β_2	-12.5934	6.0146	-2.0938
β_3	0.6104	6.0609	0.1007
β_4	13.762	6.1126	2.2514
β_5	0.7478	6.1467	0.1217
β_6	-16.5324	6.1936	-2.6693

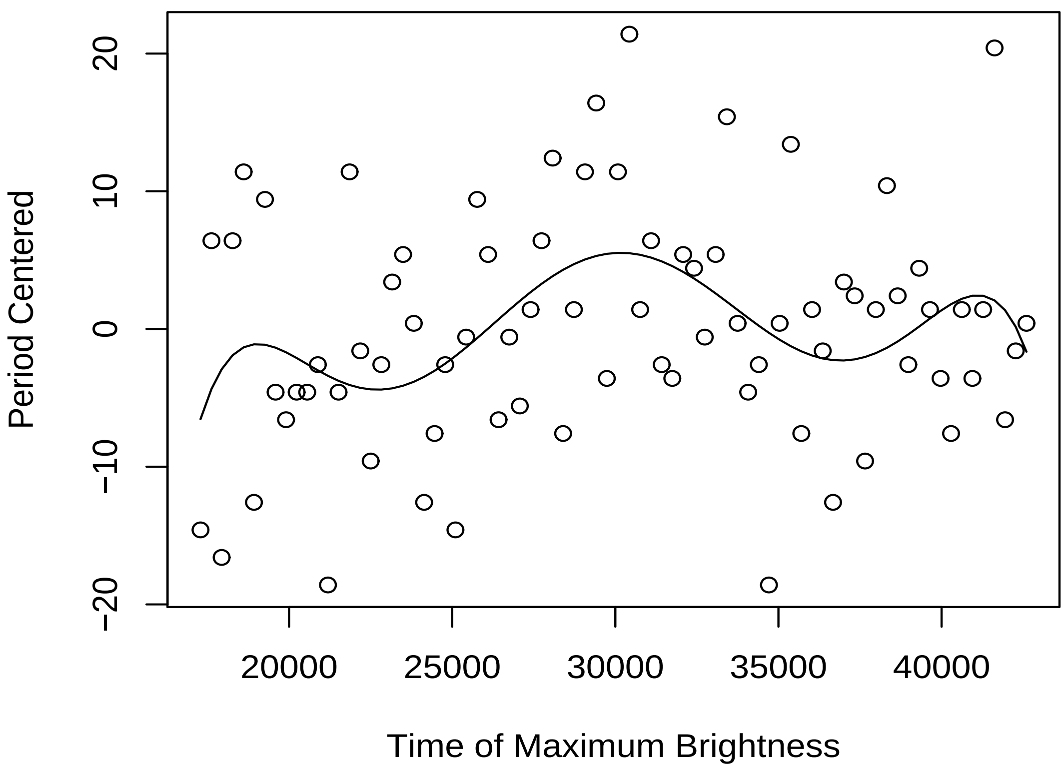

Figure 10. Plot of pseudo-periods for R. Canum Venaticorum.

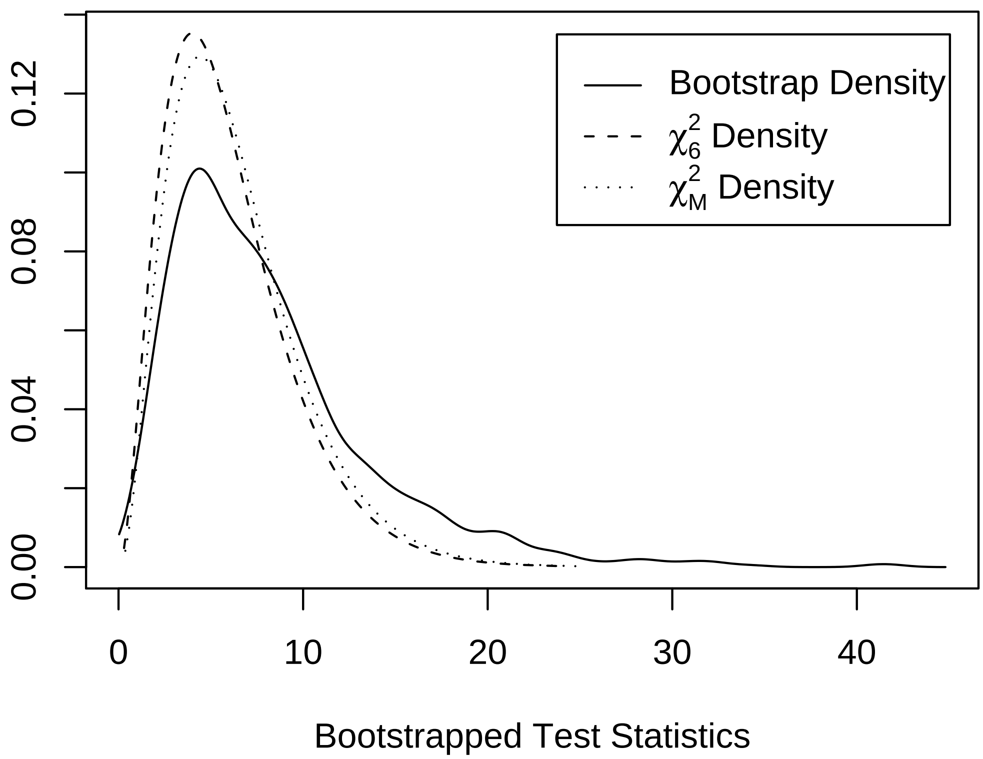

Figure 11. *Density estimate of the bootstrapped test statistic for R. Canum Venaticorum.*

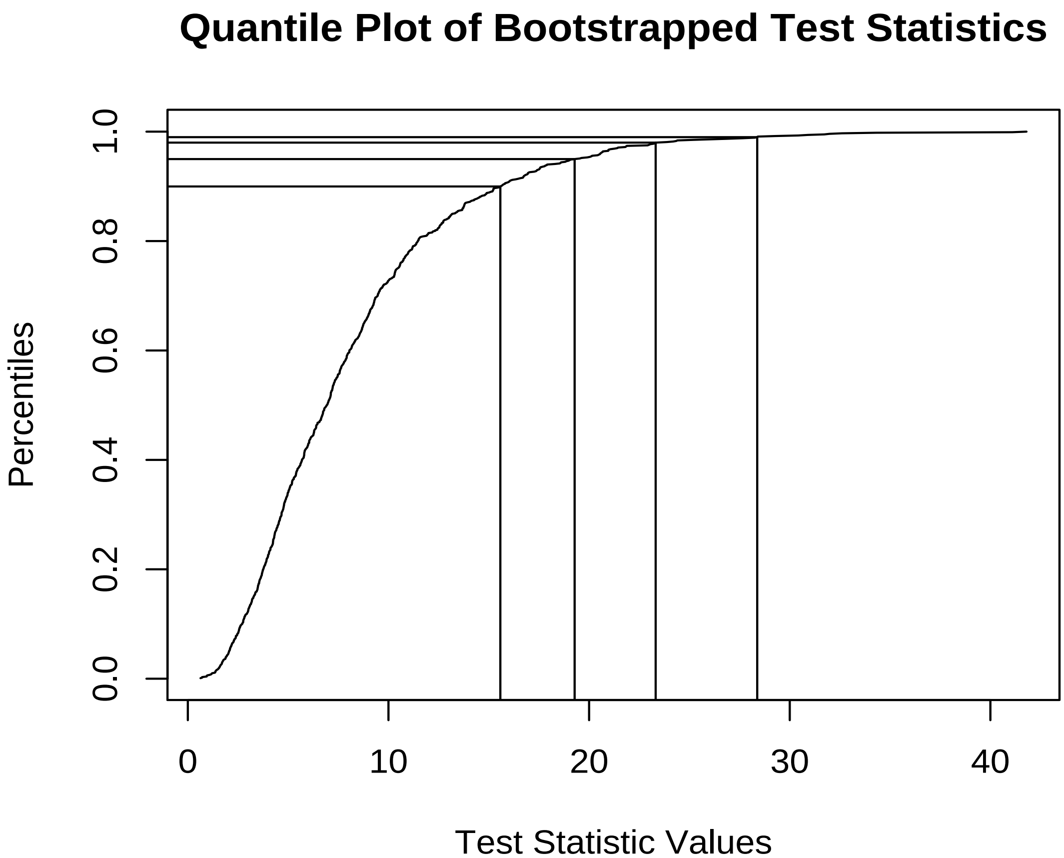

Figure 12. Quantile plot of the bootstrapped test statistic for R. Canum Venaticorum. Vertical and horizontal lines mark the 90th, 95th, 98th and 99th percentiles of the bootstrapped test statistic.

5.4 W. Draconis

The variable star W. Draconis has 98 observations. There is one missing observation however. To account for this, the correlation matrix is modified so that after deleting the ith observation the correlation between observation $i-1$ and $i+1$ is 0. Figure 13 is a plot of the pseudo-period versus the time of maximum brightness. The solid line is the result of fitting a sixth degree polynomial. The test statistic obtained from the original data was 75.8866 and the p-value is less than 0.0001. The p-value using χ^2_6 is also less than 0.0001. Figure 14 shows that the density of the bootstrapped test statistic is quite close to that of a χ^2_6. The value M in this case is 6.08. Figure 15 illustrates that the observed test statistic is much larger than the 99th percentile. The maximum likelihood estimates in Table 7 help to illustrate which coefficients could be significant.

Table 7. Maximum likelihood estimates and standard errors of the regression coefficients for W. Draconis.

coefficient	estimate	standard error	standardized estimate
β_1	55.4072	1.5378	36.0306
β_2	8.6105	0.7072	12.1762
β_3	12.6113	1.8423	6.8453
β_4	-0.5929	1.7058	-0.3476
β_5	8.4405	1.7141	4.9241
β_6	7.4912	2.0628	3.6316

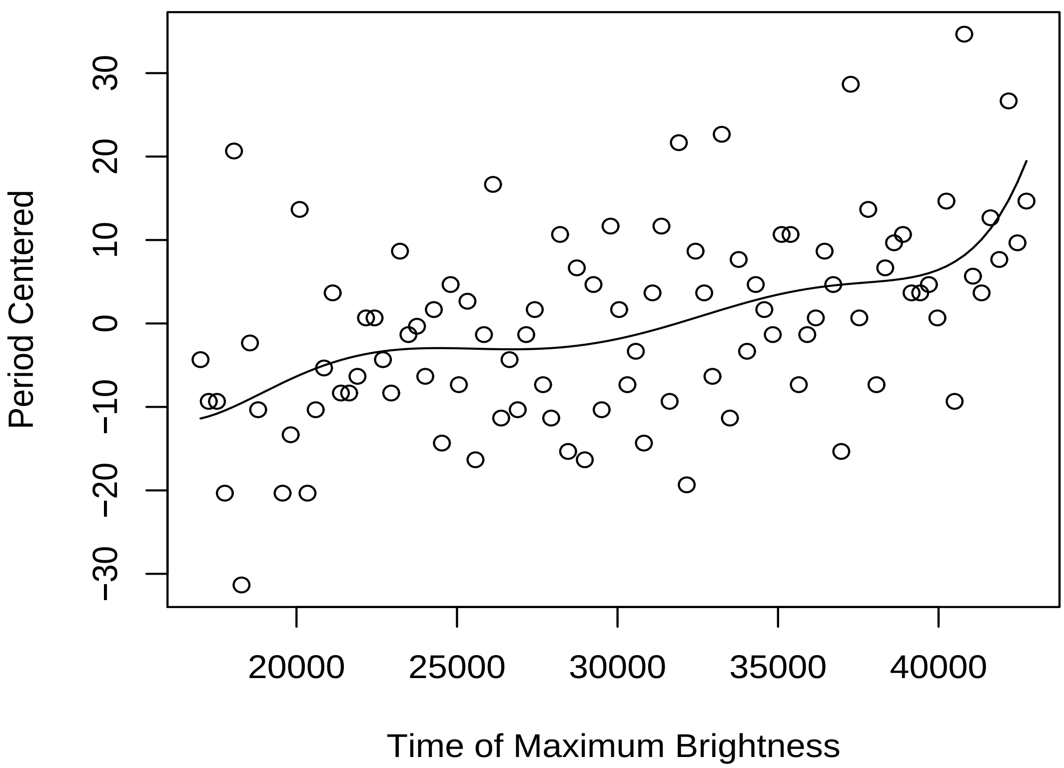

Figure 13. Plot of pseudo-periods for W. Draconis.

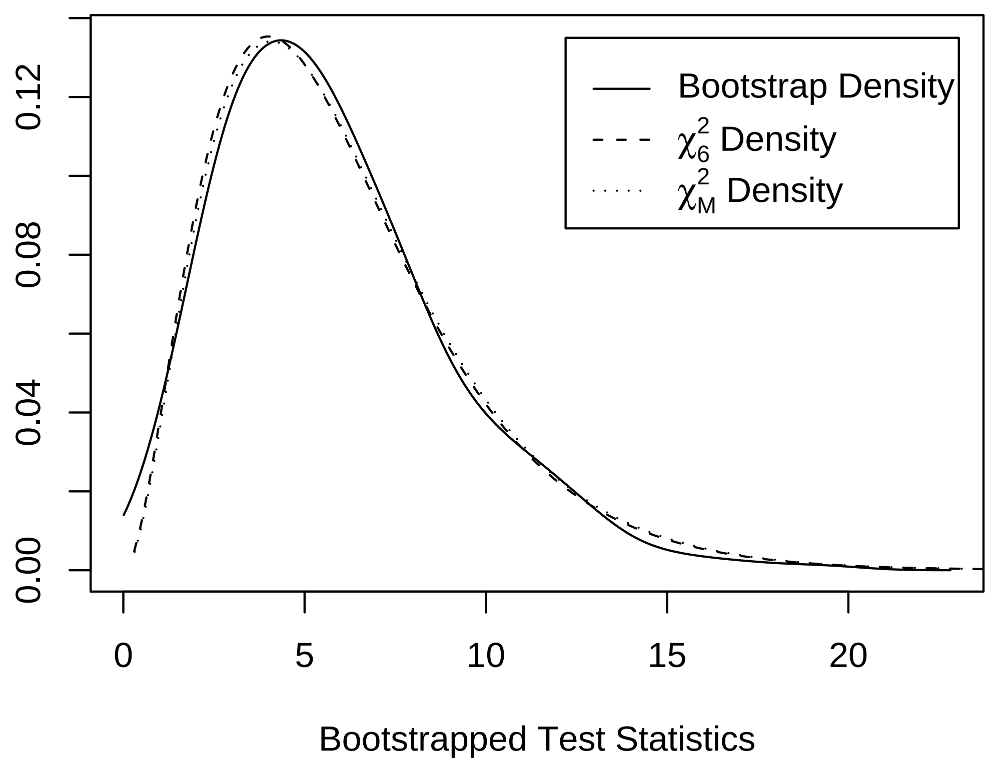

Figure 14. Density estimate of the bootstrapped test statistic for W. Draconis.

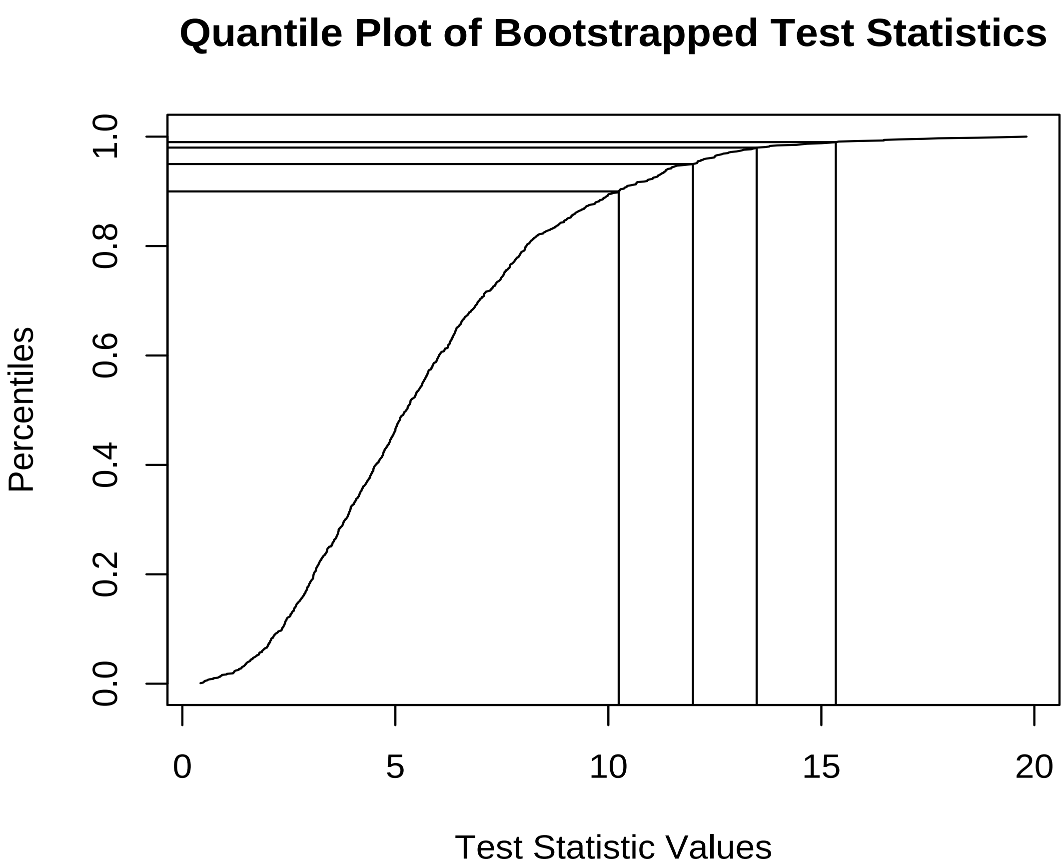

Figure 15. Quantile plot of the bootstrapped test statistic for W. Draconis. Vertical and horizontal lines mark the 90th, 95th, 98th and 99th percentiles of the bootstrapped test statistic.

5.5 Y. Aquarii

The variable star Y. Aquarii has 68 observations. There are missing observations in this dataset as well. After the fifth observation the next five observations are missing. The correlation matrix was modified as before to account for the missing observations. Figure 16 is a plot of the pseudo-period versus the time of maximum brightness. The solid line is the result of fitting a sixth degree polynomial. The test statistic obtained from the original data is 22.3489 while the p-value is 0.0250. The p-value from χ^2_6 is 0.0011. Figure 17 shows that there is a large discrepency between the density of the bootstrapped test statistics and the χ^2_6 density. In this case, M is 15.44. Figure 18 provides a quantile plot of the bootstrapped test statistics which also shows that the observed test statistic lies near the 98th percentile. Table 8 below provides maximum likelihood estimates as well as standard errors to illustrate which coefficients may be significant.

Table 8. Maximum likelihood estimates and standard errors of the regression coefficients for Y. Aquarii.

coefficient	estimate	standard error	standardized estimate
β_1	9.72	2.1094	4.6079
β_2	14.6017	1.7764	8.2199
β_3	-4.574	2.7582	-1.6583
β_4	2.4337	3.3236	0.7323
β_5	-7.6298	2.5484	-2.9939
β_6	-10.2411	3.0991	-3.3046

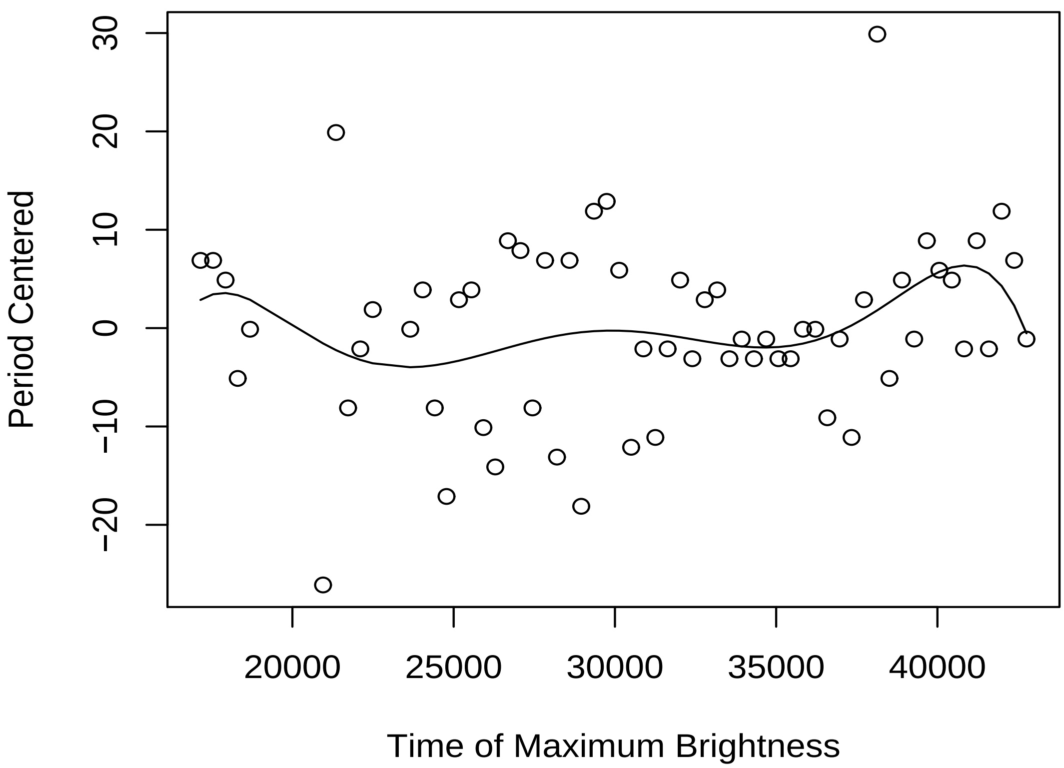

Figure 16. Plot of pseudo-periods for Y. Aquarii.

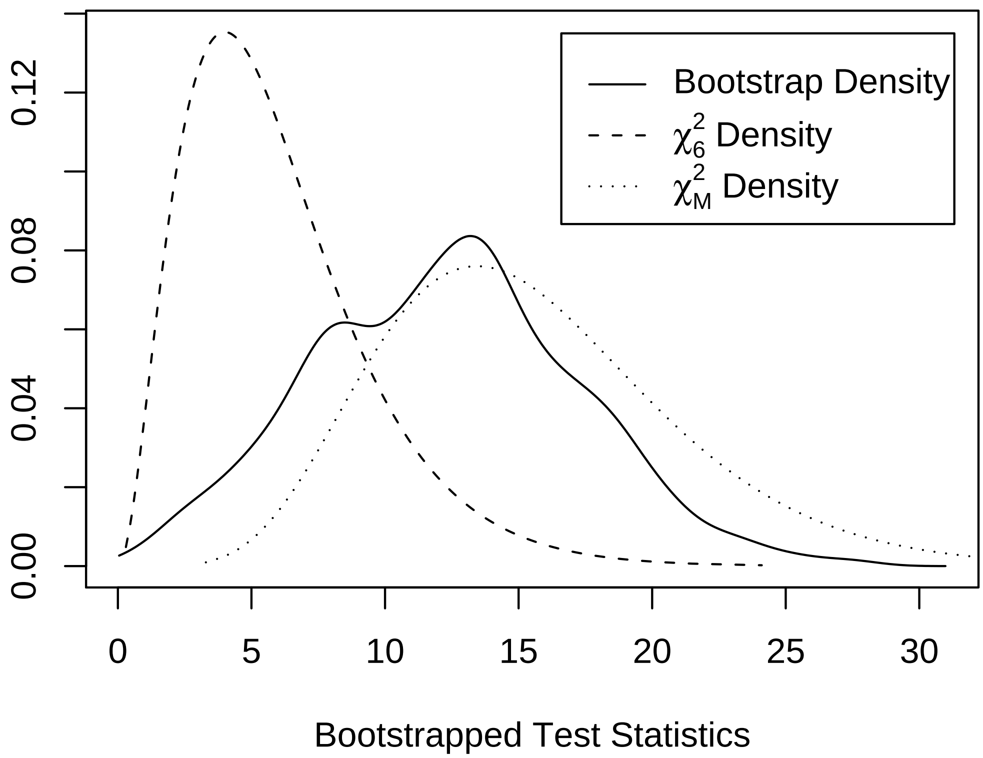

Figure 17. Density estimate of the bootstrapped test statistic for Y. Aquarii.

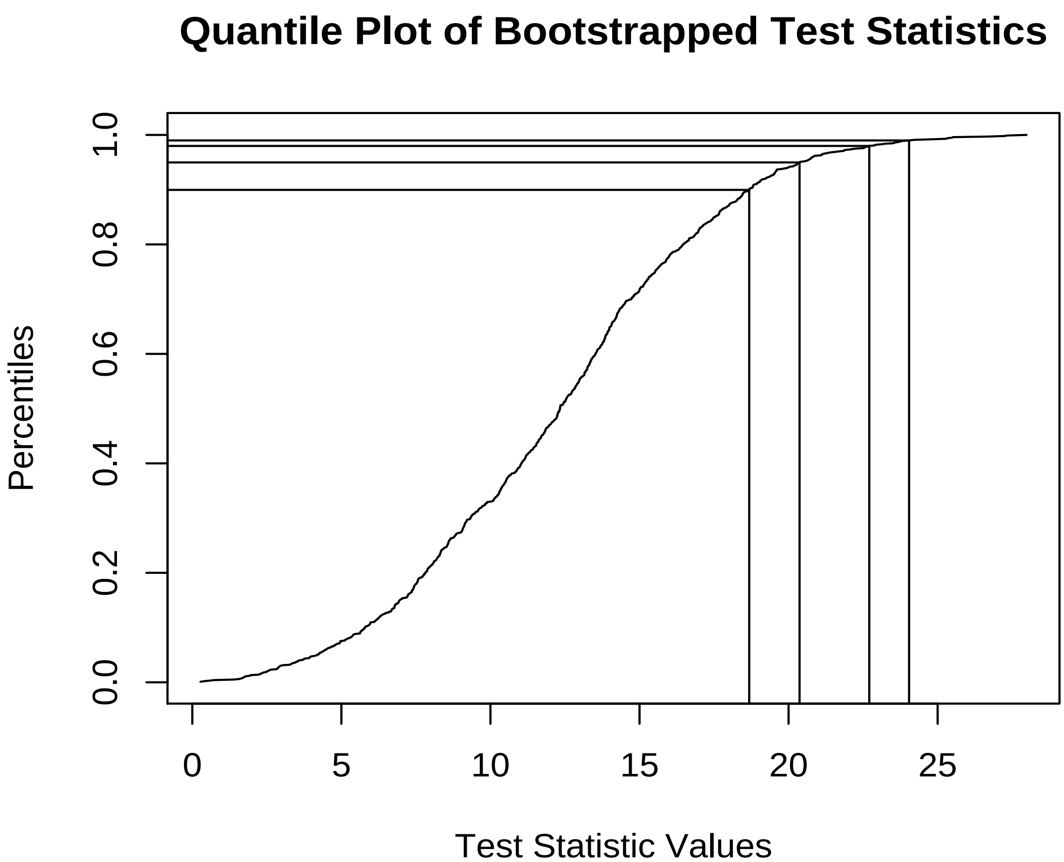

Figure 18. *Quantile plot of the bootstrapped test statistic for Y. Aquarii. Vertical and horizontal lines mark the 90th, 95th, 98th and 99th percentiles of the bootstrapped test statistic.*

CHAPTER VI

CONCLUSION

This dissertation discussed using a bootstrap procedure for a test of no trend when the error process has a particular dependence structure. It has been shown that only the first three moments of I and ϵ need to be estimated as efficiently as possible for the bootstrap procedure to be effective. Simulations show that for the case $r = 0$, *i.e.*, when a star's intrinsic variance vanishes, the large sample test is not appropriate. From the analysis of the five variable stars, the bootstrap distribution of the test statistic may be approximately χ^2 but this is not always the case. There can be quite a discrepency in results between the bootstrap and large sample test.

REFERENCES

Chung, K.L. (1974), *A Course in Probability Theory*, 2nd ed. New York: Academic Press.

Davison, A.C. and Hinkley, D.V. (1997), *Bootstrap Methods and Their Applications*, England: Cambridge University Press.

Fang, Y. and Koreisha, S. (2001), "Generalized Least Squares with Misspecified Serial Correlation Structures," *Journal of the Royal Statistical Society: Series B*, 63, 515-531.

Furrer, R. (2002), "M-Estimation for Dependent Random Variables," *Statistics & Probability Letters*, 57, 337-341.

Hall, P. (1992), *The Bootstrap and Edgeworth Expansion*, New York: Springer-Verlag.

Hu, G.Y. and O'Connell, R.R. (1996), "Analytical Inversion of Symmetric Tridiagonal Matrices," *Journal of Physics A: Mathematics and General*, 29, 1511-1513.

Koen, C. and Lombard, F. (2001), "The Analysis of Indexed Astronomical Time Series - VII. Simulations Use of Times of Maxima and Minima to Test for Period Changes in Long-Period Variables," *Monthly Notices of the Royal Astronomical Society*, 325, 1124 - 1132.

Opsomer,J., Wang, Y. and Yang, Y. (2001), "Nonparametric Regression with Correlated Errors," *Statistical Science*, 16, 134-153.

Sen, A. and Srivastava, M. (1990), *Regression Analysis: Theory, Methods and Applications*, New York: Springer-Verlag.

Serfling, R. (1980), *Approximation Theorems of Mathematical Statistics*, New York: John Wiley & Sons, Inc.

Stout, W.F. (1974), *Almost Sure Convergence*, New York: Academic Press.

APPENDIX A

MISCELLANEOUS RESULTS

In this section, we want to establish a few things:

1. The form of the ijth element of A^{-1}.
2. The elements of $\frac{1}{n}X'A^{-1}X \to$ positive constants.
3. $\lim_{n\to\infty} \frac{c^2}{n} \sum_{r=1}^{n} \sum_{s=1}^{n} [A^{-k}]_{rs} < \infty$.

Hu & O'Connell (1996) give an expression for the inverse of a symmetric tridiagonal matrix similar to that of A which we are interested in. After simplifying their results, the i,jth element of A^{-1} can be shown to be

$$\frac{\sinh(\min(i,j)\lambda)\sinh((n-\max(i,j)+1)\lambda)}{\sinh(\lambda)\sinh((n+1)\lambda)} \tag{A.1}$$

where $\lambda = \cosh^{-1}((r+2)/2)$. Note that $\lambda = 0$ only if $r = 0$.

Proof of 2.: Let the elements of A^{-1} be b_{ij}^n. Then

$$\begin{aligned}
(\frac{1}{n}X'A^{-1}X)_{rs} &= \frac{1}{n}\sum_{i=1}^{n} b_{ii}^n \phi_r(x_i)\phi_s(x_i) + \frac{2}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} b_{ij}^n \phi_r(x_i)\phi_s(x_j) \\
&= (e^{\lambda} - e^{-\lambda})^{-1} \int_0^1 \phi_r(x)\phi_s(x)dx + O\Big(\frac{1}{n}\Big) \\
&\quad + \frac{2}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} b_{ij}^n \phi_r(x_i)\phi_s(x_j).
\end{aligned}$$

Now,

$$\begin{aligned}
& \frac{2}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} b_{ij}^{n}\phi_r(x_i)\phi_s(x_j) \\
= \; & (e^{\lambda}-e^{-\lambda})^{-1}(1+c_n e^{-2(n+1)\lambda}) \\
& \quad \times\frac{2}{n}\sum_{i=1}^{n-1}\phi_r(x_i)\sum_{j=i+1}^{n} e^{-(j-i)\lambda}\phi_s(x_j)+O\Big(\frac{1}{n}\Big) \\
= \; & (e^{\lambda}-e^{-\lambda})^{-1}\frac{2}{n}\sum_{i=1}^{n-1}\phi_r(x_i)\sum_{k=1}^{n-i} e^{-k\lambda}\phi_s(x_{k+i})+O\Big(\frac{1}{n}\Big)
\end{aligned}$$

where $c_n=(1-a_n)^{-2}$ with $a_n\in[0,e^{-2(n+1)\lambda}]$. Let $m_n=o(n)$, but $m_n\to\infty$. Then we have

$$\begin{aligned}
\frac{2}{n}\sum_{i=1}^{n-1}\phi_r(x_i)\sum_{k=1}^{n-i} e^{-k\lambda}\phi_s(x_{k+i}) &= \frac{2}{n}\sum_{i=1}^{n-1}\phi_r(x_i)\sum_{k=1}^{m_n} e^{-k\lambda}\phi_s(x_{k+i}) \\
&+\frac{2}{n}\sum_{i=1}^{n-1}\phi_r(x_i)\sum_{k=m_n+1}^{n-i} e^{-k\lambda}\phi_s(x_{k+i})
\end{aligned}$$

The absolute value of the second term is bounded by

$$2C^2\sum_{k=m_n+1}^{\infty} e^{-k\lambda}=\frac{2C^2e^{-(m_n+1)\lambda}}{1-e^{-\lambda}}$$

where $\mid \phi_l \mid\leq C$. The last quantity tends to 0 since $m_n\to\infty$.

$$\begin{aligned}
\frac{2}{n}\sum_{i=1}^{n-1}\phi_r(x_i)\sum_{k=1}^{m_n} e^{-k\lambda}\phi_s(x_{k+i}) &= \frac{2}{n}\sum_{i=1}^{n-1}\phi_r(x_i)\sum_{k=1}^{m_n} e^{-k\lambda}(\phi_s(x_i)+\frac{k}{n}\phi'(x_{i,k})) \\
&= \frac{2}{n}\sum_{i=1}^{n-1}\phi_r(x_i)\phi_s(x_i)\sum_{r=1}^{m_n} e^{-r\lambda}+O\Big(\frac{m_n}{n}\Big) \\
&= 2\frac{e^{-\lambda}}{1-e^{-\lambda}}\int_0^1\phi_r(x)\phi_s(x)dx+O\Big(\frac{m_n}{n}\Big)+O\Big(e^{-m_n}\Big).
\end{aligned}$$

It follows that the elements of $\frac{1}{n}X'A^{-1}X$ converge to

$$
\begin{aligned}
& \int_0^1 \phi_r(x)\phi_s(x)dx(e^{\lambda}-e^{-\lambda})^{-1}\Big(1+\frac{2e^{-\lambda}}{1-e^{-\lambda}}\Big) \\
= & \int_0^1 \phi_r(x)\phi_s(x)dx\frac{1+e^{-\lambda}}{(1-e^{-\lambda})(e^{\lambda}-e^{-\lambda})}
\end{aligned}
$$

with error of order $O(\frac{\log n}{n})$ if $m_n = \log n$.

Proof of 3: It has been noted in Appendix B that the ijth element of A^{-1} is less than

$$C_\lambda e^{\min(i,j)\lambda}e^{-\max(i,j)\lambda}$$

where $C_\lambda = (e^{\lambda}-e^{-\lambda})^{-1}(1-e^{-2\lambda})^{-1}$, which implies that

$$
\begin{aligned}
& \frac{c^2}{n}\sum_{r=1}^{n}\sum_{s=1}^{n}[A^{-1}]_{rs} \\
< \; & C_\lambda \frac{c^2}{n}\sum_{r=1}^{n}\sum_{s=1}^{n} e^{\min(r,s)\lambda}e^{-\max(r,s)\lambda} \\
= \; & C_\lambda \frac{c^2}{n}\Big(\sum_{r=1}^{n}\sum_{s=1}^{r} e^{\min(r,s)\lambda}e^{-\max(r,s)\lambda} \\
& \qquad +\sum_{r=1}^{n-1}\sum_{s=r+1}^{n} e^{\min(r,s)\lambda}e^{-\max(r,s)\lambda}\Big) \\
= \; & C_\lambda \frac{c^2}{n}\Big(\sum_{r=1}^{n}\sum_{s=1}^{r} e^{s\lambda}e^{-r\lambda} \\
& \qquad +\sum_{r=1}^{n-1}\sum_{s=r+1}^{n} e^{r\lambda}e^{-s\lambda}\Big) \\
= \; & C_\lambda \frac{c^2}{n}\frac{ne^{2\lambda}-n+2e^{\lambda}e^{-n\lambda}-2e^{\lambda}}{(e^{\lambda}-1)^2} \\
= \; & C_\lambda c^2 \frac{e^{2\lambda}-1+(2/n)e^{\lambda}e^{-n\lambda}-(2/n)e^{\lambda}}{(e^{\lambda}-1)^2}.
\end{aligned}
$$

Taking the limit as n tends to ∞ yields

$$\frac{C_\lambda c^2(e^{2\lambda}-1)}{(e^\lambda-1)^2}$$

which is less than ∞.

APPENDIX B

ALMOST SURE CONVERGENCE OF LINEAR FORMS

We want to show that each element of $\frac{1}{n}X'A^{-1}\mathbf{Z}$ and $\frac{1}{n}X'A^{-2}\mathbf{Z}$ converges almost surely, where

$$Z_i = I_i + \epsilon_i - \epsilon_{i-1}.$$

X is the design matrix with dimension n x p. If it can be shown that $\frac{1}{n}X'A^{-1}\mathbf{I}$, $\frac{1}{n}X'A^{-1}\epsilon_{\mathbf{1}}$ and $\frac{1}{n}X'A^{-1}\epsilon_{\mathbf{0}}$ all converge almost surely to 0, then it also holds for $\frac{1}{n}X'A^{-1}\mathbf{Z}$. In fact, if it holds for any one of the expressions, it holds for the others. The ith element of $\frac{1}{n}X'A^{-1}\mathbf{I}$ can be written as

$$\mathbf{C}_i\mathbf{I} = \sum_{j=1}^{n} C_{ij}I_j$$

where $I_1, \ldots, I_n$ are i.i.d. random variables with mean 0 and finite variance. $\mathbf{C}_k$ is the kth row of the matrix $C = X'A^{-1}/n$, where A is as defined in (1.2). Let $x_i = \frac{i-1/2}{n}, i = 1, \ldots, n$. Denote the ikth element of X as $U_k(x_i), j = 1, \ldots, 6$, where we assume that $U_1, \ldots, U_6$ are orthogonal polynomials. We know that $\mid U_k(x_i) \mid < c_1$ for all x_i and $k = 1, \ldots, 6$, where c_1 is some constant less than ∞. According to Stout (1974, p. 231), in order that the sum above converges almost surely to 0, it is sufficient to show that

$$\sum_{j=1}^{\infty} C_{ij}^2 < Kn^{-\alpha} \quad \text{and} \quad C_{ij}^2 < Kj^{-1}$$

where $\alpha > 0$, $K < \infty$ and $k \geq 1$.

$$\begin{aligned}
\sum_{j=1}^{\infty} C_{ij}^2 &= \sum_{j=1}^{n}\Big(\frac{1}{n}\sum_{k=1}^{n} X_{ki}[A^{-1}]_{kj}\Big)^2 \\
&\leq \frac{c_1^2}{n^2}\sum_{j=1}^{n}\Big(\sum_{k=1}^{n}[A^{-1}]_{kj}\Big)^2 \\
&= \frac{c_1^2}{n^2}\Big(\sum_{j=1}^{n}\sum_{k=1}^{n}[A^{-1}]_{kj}^2 \\
&\qquad +2\sum_{j=1}^{n}\sum_{k=1}^{n-1}\sum_{l=k+1}^{n}[A^{-1}]_{kj}[A^{-1}]_{lj}\Big)
\end{aligned}$$

Each of these summations will be simplified separately. Appendix A gives the expression for the ijth element of A^{-1}. Denote $D_n(\lambda)$ as $\sinh(\lambda)\sinh((n+1)\lambda)$. Since $\sinh(x) < e^x$, the ijth element of A^{-1} is less than

$$\begin{aligned}
& \frac{e^{\min(i,j)\lambda}e^{-\max(i,j)\lambda}e^{(n+1)\lambda}}{D_n(\lambda)} \\
\leq\ & C_\lambda e^{\min(i,j)\lambda}e^{-\max(i,j)\lambda}
\end{aligned}$$

where C_λ is defined in Appendix A. Therefore,

$$\begin{aligned}
& \sum_{j=1}^{n}\sum_{k=1}^{n}[A^{-1}]_{kj}^2 \\
\leq\ & C_\lambda^2\sum_{j=1}^{n}\sum_{k=1}^{n} e^{2\min(i,j)\lambda}e^{-2\max(i,j)\lambda} \\
=\ & C_\lambda^2\Big(\sum_{j=1}^{n}\sum_{k=1}^{j} e^{2\min(j,k)\lambda}e^{-2\max(j,k)\lambda} \\
& \qquad +\sum_{j=1}^{n-1}\sum_{k=j+1}^{n} e^{2\min(j,k)\lambda}e^{-2\max(j,k)\lambda}\Big)
\end{aligned}$$

$$
\begin{aligned}
&= C_\lambda^2 e^{-2\lambda}\Big(\sum_{j=1}^{n}\sum_{k=1}^{j} e^{2k\lambda}e^{-2j\lambda} \\
&\qquad + \sum_{j=1}^{n-1}\sum_{k=j+1}^{n} e^{2j\lambda}e^{-2k\lambda}\Big) \\
&= C_\lambda^2 \frac{n(e^{4\lambda}-1) - 2e^{2\lambda}(1-e^{-2n\lambda})}{(e^{2\lambda}-1)^2} \\
&\leq C_2 n
\end{aligned}
$$

for a positive constant C_2.

The second summation is

$$
\begin{aligned}
& 2\sum_{j=1}^{n}\sum_{k=1}^{n-1}\sum_{l=k+1}^{n}[A^{-1}]_{kj}[A^{-1}]_{lj} \\
< \; & 2C_\lambda^2\Big(\sum_{j=1}^{n}\sum_{k=1}^{n-1}\sum_{l=k+1}^{n} e^{\min(j,k)\lambda}e^{-\max(j,k)\lambda}e^{\min(j,l)\lambda}e^{-\max(j,l)\lambda}\Big) \\
= \; & 2C_\lambda^2\Big(\sum_{j=1}^{n-1}\sum_{k=j}^{n-1}\sum_{l=k+1}^{n} e^{\min(j,k)\lambda}e^{-\max(j,k)\lambda}e^{\min(j,l)\lambda}e^{-\max(j,l)\lambda} \\
& \quad + \sum_{k=1}^{n-1}\sum_{j=k+1}^{n}\sum_{l=j}^{n} e^{\min(j,k)\lambda}e^{-\max(j,k)\lambda}e^{\min(j,l)\lambda}e^{-\max(j,l)\lambda} \\
& \quad + \sum_{k=1}^{n-2}\sum_{l=k+1}^{n-1}\sum_{j=l+1}^{n} e^{\min(j,k)\lambda}e^{-\max(j,k)\lambda}e^{\min(j,l)\lambda}e^{-\max(j,l)\lambda}\Big) \\
= \; & 2C_\lambda^2\Big(\sum_{j=1}^{n-1}\sum_{k=j}^{n-1}\sum_{l=k+1}^{n} e^{2j\lambda}e^{-k\lambda}e^{-l\lambda} \\
& \quad + \sum_{k=1}^{n-1}\sum_{j=k+1}^{n}\sum_{l=j}^{n} e^{k\lambda}e^{-l\lambda} \\
& \quad + \sum_{k=1}^{n-2}\sum_{l=k+1}^{n-1}\sum_{j=l+1}^{n} e^{k\lambda}e^{-2j\lambda}e^{l\lambda}\Big)
\end{aligned}
$$

It is easily verified that the last expression is bounded by $C_3 n$, where C_3 is a positive constant. Combining both summations,

$$\sum_{j=1}^{\infty} C_{ij}^2 \leq \frac{c_1^2}{n^2}(C_2 n + C_3 n) \leq \frac{C_4}{n},$$

which verifies one of Stout's two conditions.

Finally, it needs to be shown that $C_{ij}^2 < Kj^{-1}$.

$$\begin{aligned}
\mid C_{ij} \mid &\leq \frac{1}{n}\sum_{r=1}^{n} \mid U_i(x_r) \mid [A^{-1}]_{rj} \\
&\leq \frac{c_1}{n}\sum_{r=1}^{n}[A^{-1}]_{rj} \\
&= \frac{c_1}{n}\Big(\sum_{r=1}^{j}[A^{-1}]_{rj} + \sum_{r=1}^{n-j}[A^{-1}]_{(r+j),j}\Big).
\end{aligned}$$

Arguing as before, each of these two sums is bounded by a constant, call it C_5, and hence $\mid C_{ij} \mid \leq \frac{2C_1C_5}{n}$. Therefore

$$C_{ij}^2 \leq C_6/n^2 \leq C_6/(nj) \leq C_6/j.$$

A similar proof can be obtained when A^{-2} is used instead of A^{-1}.

APPENDIX C

ALMOST SURE CONVERGENCE OF QUADRATIC FORMS

Consider

$$\frac{-1}{2\sigma_\epsilon^2} + \frac{1}{2n\sigma_\epsilon^4}\mathbf{Z}'A^{-1}\mathbf{Z}$$

where

$$\begin{aligned}
\mathbf{Z}'A^{-1}\mathbf{Z} &= \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}^n (I_i + \epsilon_i - \epsilon_{i-1})(I_j + \epsilon_j - \epsilon_{j-1}) \\
&= Q_I + 2\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}^n I_i \epsilon_j - 2\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}^n I_i \epsilon_{j-1} + Q_\epsilon \\
&\quad -2\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}^n \epsilon_i \epsilon_{j-1} + \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}^n \epsilon_{i-1}\epsilon_{j-1}
\end{aligned}$$

where Q_I and Q_ϵ are quadratic forms equal to $\mathbf{I}'A^{-1}\mathbf{I}$ and $\epsilon' A^{-1}\epsilon$. We have

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}^n \epsilon_i \epsilon_{j-1} &= \sum_{i=1}^{n}\sum_{k=0}^{n-1} b_{i(k+1)}^n \epsilon_i \epsilon_k \\
&= \sum_{i=1}^{n}\sum_{k=0}^{n} b_{i(k+1)}^n \epsilon_i \epsilon_k
\end{aligned}$$

where $b_{i(n+1)} = 0 \;\forall\; i$. Now we have that the last term is

$$\sum_{i=1}^{n}\sum_{k=1}^{n} \tilde{b}_{ik}\epsilon_i\epsilon_k + \sum_{i=1}^{n} b_{i1}^n \epsilon_i \epsilon_0,$$

where

$$\tilde{b}_{ik} = \begin{cases} b_{i(k+1)}^n, & k = 1, \ldots, n-1, \\ 0, & k = n. \end{cases}$$

We know that

$$\mathrm{E}(\mathbf{Z}'A^{-1}\mathbf{Z}) = \mathrm{tr}(A^{-1}\sigma_\epsilon^2 A) = n\sigma_\epsilon^2,$$

$$\mathrm{E}\Big(\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}^n I_i \epsilon_j\Big) = 0 \quad \text{and} \quad \mathrm{E}\Big(\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}^n I_i \epsilon_{j-1}\Big) = 0,$$

which together imply that

$$\mathrm{E}\Big(\sum_{i=1}^{n} b_{ii}^n (I_i^2 + \epsilon_i^2 + \epsilon_{i-1}^2) - 2\sum_{i=1}^{n-1} b_{i(i+1)}^n \epsilon_i^2\Big) = n\sigma_\epsilon^2.$$

Below is a proof that b_{ij}^n is essentially free of n. For $i \leq j$, by (A.1),

$$\begin{aligned}
b_{ij}^n &= (e^\lambda - e^{-\lambda})^{-1}(1 - e^{-2(n+1)\lambda})^{-1}(e^{i\lambda} - e^{-i\lambda})(e^{-j\lambda} - e^{j\lambda}e^{-2(n+1)\lambda}) \\
&= (e^\lambda - e^{-\lambda})^{-1}(1 + e^{-2(n+1)\lambda}(1 - a_n)^{-2})(e^{-(j-i)\lambda} - e^{-(i+j)\lambda} \\
&\qquad - e^{-2(n+1)\lambda}(e^{(i+j)\lambda} - e^{(j-i)\lambda})) \\
&= (e^\lambda - e^{-\lambda})^{-1}(e^{-(j-i)\lambda} - e^{-(i+j)\lambda}) + (e^\lambda - e^{-\lambda})^{-1}e^{-2(n+1)\lambda} \\
&\qquad \times c_n(e^{-(j-i)\lambda} - e^{-(i+j)\lambda}) + d_n e^{-2(n+1)\lambda}(e^{(i+j)\lambda} - e^{-(j-i)\lambda}),
\end{aligned}$$

where $a_n \in [0, e^{-2(n+1)\lambda}]$, $c_n = (1 - a_n)^{-2}$ and $d_n = -(e^\lambda - e^{-\lambda})^{-1}(1 + c_n e^{-2(n+1)\lambda})$.

It will be shown that

1. $\frac{1}{n}\sum_{i=1}^n b_{ii}^n I_i^2 \xrightarrow{\text{a.s}} \sigma_I^2 \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n b_{ii}^n$
2. $\frac{1}{n}\sum_{i=1}^n b_{ii}^n \epsilon_i^2 \xrightarrow{\text{a.s}} \sigma_\epsilon^2 \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n b_{ii}^n$
3. $\frac{1}{n}\sum_{i=1}^n b_{i(i+1)}^n \epsilon_i^2 \xrightarrow{\text{a.s}} \sigma_\epsilon^2 \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n-1} b_{i(i+1)}^n$
4. $\frac{1}{n}\sum_{i\neq j} b_{ij}^n I_i I_j \xrightarrow{\text{a.s}} 0$
5. $\frac{1}{n}\sum_{i=1}^n \sum_{j=1}^n b_{ij}^n I_i \epsilon_j \xrightarrow{\text{a.s}} 0$

6. $\frac{1}{n}\sum_{i\neq j}^{n} b_{ij}^{n}\epsilon_i\epsilon_j \xrightarrow{\text{a.s}} 0$

7. $\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{b}_{ij}^{n}\epsilon_i\epsilon_j \xrightarrow{\text{a.s}} 0$

8. $\frac{1}{n}\sum_{i=1}^{n} b_{ii}^{n}\epsilon_i\epsilon_0 \xrightarrow{\text{a.s}} 0.$

If the above hold, this will show that

$$\frac{1}{2n\sigma_\epsilon^4}\mathbf{Z}'A^{-1}\mathbf{Z} \xrightarrow{\text{a.s}} \frac{1}{2\sigma_\epsilon^2},$$

since

$$\sum_{i=1}^{n} b_{ii}^{n}(\sigma_I^2 + 2\sigma_\epsilon^2) - 2\sigma_\epsilon^2\sum_{i=1}^{n} b_{i(i+1)}^{n} = n\sigma_\epsilon^2.$$

Proof of 1.:

$$b_{ii}^{n} = \frac{1-e^{-2i\lambda}}{e^{\lambda}-e^{-\lambda}}(1 + c_n e^{-2(n+1)\lambda}) + d_n e^{-2(n+1)\lambda}(e^{2i\lambda}-1)$$

and therefore

$$\frac{1}{n}\sum_{i=1}^{n} b_{ii}^{n} I_i^2 = \frac{1 + c_n e^{-2(n+1)\lambda}}{e^{\lambda}-e^{-\lambda}}\frac{1}{n}\sum_{i=1}^{n}(1-e^{-2i\lambda})I_i^2 + d_n e^{-2(n+1)\lambda}\frac{1}{n}\sum_{i=1}^{n}(e^{2i\lambda}-1)I_i^2.$$

Rewrite

$$\frac{1}{n}\sum_{i=1}^{n}(1-e^{-2i\lambda})I_i^2 = \frac{1}{n}\sum_{i=1}^{n}(1-e^{-2i\lambda})Z_i + \sigma_I^2\frac{1}{n}\sum_{i=1}^{n}(1-e^{-2i\lambda}),$$

where $Z_i = I_i^2 - \sigma_I^2$. Corollary to Theorem 5.4.1 in Chung (1974) now implies that

$$\frac{1}{n}\sum_{i=1}^{n}(1-e^{-2i\lambda})Z_i \xrightarrow{\text{a.s}} 0$$

so long as E $\mid Z_i \mid < \infty$, and hence

$$\frac{1}{n}\sum_{i=1}^{n}(1-e^{-2i\lambda})I_i^2 \overset{\text{a.s.}}{\rightarrow} \sigma_I^2.$$

Now,

$$d_n e^{-2(n+1)\lambda}\frac{1}{n}\sum_{i=1}^{n}(e^{2i\lambda}-1)I_i^2 = d_n e^{-2(n+1)\lambda}\frac{1}{n}\sum_{i=1}^{n}(e^{2i\lambda}-1)Z_i + O\Big(\frac{1}{n}\Big).$$

Define $V_i = (e^{2i\lambda}-1)Z_i$ and apply corollary to Theorem 5.4.1 to $e^{-2n\lambda}\frac{1}{n}\sum_{i=1}^{n} V_i$. In Chung's notation, $a_i = ie^{2i\lambda}$ and let $\phi(v) = \mid v \mid^{\alpha}$ for $1 < \alpha \leq 2$.

$$\begin{aligned}
\sum_{i=1}^{\infty}\frac{\mathrm{E}\mid V_i\mid^{\alpha}}{(ie^{2i\lambda})^{\alpha}} &= \mathrm{E}\mid Z_1\mid^{\alpha}\sum_{i=1}^{\infty}\frac{(e^{2i\lambda}-1)^{\alpha}}{i^{\alpha}e^{2i\lambda\alpha}} \\
&= \mathrm{E}\mid Z_1\mid^{\alpha}\sum_{i=1}^{\infty}\frac{(1-e^{-2i\lambda})^{\alpha}}{i^{\alpha}} \\
&< \infty.
\end{aligned}$$

Therefore, by corollary to Theorem 5.4.1,

$$d_n e^{-2(n+1)\lambda}\frac{1}{n}\sum_{i=1}^{n}(e^{2i\lambda}-1)Z_i \xrightarrow{\text{a.s.}} 0.$$

Combining the above results shows that

$$\frac{1}{n}\sum_{i=1}^{n} b_{ii}^n I_i^2 \xrightarrow{\text{a.s.}} (e^{\lambda}-e^{-\lambda})^{-1}\sigma_I^2$$

under the single condition that $\mathrm{E}\mid Z_1\mid^{\alpha}<\infty$ for some $1<\alpha\leq 2$. Propositions 2. and 3. are proven in exactly the same manner under the condition that $\mathrm{E}\mid \epsilon_1^2-\sigma_\epsilon^2\mid^{\alpha}<\infty$ for some $1<\alpha\leq 2$.

Proof of 4.: Consider the quadratic form

$$Q_n = \frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} e^{-(j-i)\lambda}I_iI_j,$$

where $\lambda > 0$ and $I_1, \ldots, I_n$ are i.i.d. with $E(I_i) = 0$. Write

$$Q_n = \frac{1}{n} \sum_{i=1}^{n-1} I_i \sum_{j=1}^{n-i} e^{-j\lambda} I_{i+j}.$$

So long as $E(|I_i|) < \infty$, Theorem 5.4.1, p. 124 of Chung implies that each of the series $\sum_{j=1}^{m} e^{-j\lambda} I_{i+j}$, $i = 1, \ldots, n-1$, converges almost surely, and hence we may write

$$Q_n = \frac{1}{n} \sum_{i=1}^{n-1} I_i V_i - \frac{1}{n} \sum_{i=1}^{n-1} I_i \sum_{j=n-i+1}^{\infty} e^{-j\lambda} I_{i+j}, \tag{C.1}$$

where $V_i = \sum_{j=1}^{\infty} e^{-j\lambda} I_{i+j}$, $i = 1, 2, \ldots$.

Now, (C.1) may be written

$$\begin{aligned} Q_n &= \frac{1}{n} \sum_{i=1}^{n-1} I_i V_i - \frac{1}{n} \sum_{i=1}^{n-1} I_i \sum_{r=1}^{\infty} e^{-(r+n-i)\lambda} I_{n+r} \\ &= \frac{1}{n} \sum_{i=1}^{n-1} I_i V_i - \frac{1}{n} \sum_{s=1}^{n-1} e^{-s\lambda} I_{n-s} \sum_{r=1}^{\infty} e^{-r\lambda} I_{n+r}. \end{aligned}$$

We will show that each of the two quantities on the right hand side of the last expression tend to 0 almost surely. We use the Borel-Cantelli lemma to prove that the second tends to 0. We have

$$\sum_{n=1}^{\infty} P\left(\frac{1}{n} \left| \sum_{s=1}^{n-1} e^{-s\lambda} I_{n-s} \sum_{r=1}^{\infty} e^{-r\lambda} I_{n+r} \right| > \epsilon \right) \le \sum_{n=1}^{\infty} \frac{\mathrm{E}(S_n^2)\mathrm{E}(T_n^2)}{n^2 \epsilon^2}, \tag{C.2}$$

where $S_n = \sum_{s=1}^{n-1} e^{-s\lambda} I_{n-s}$, $T_n = \sum_{r=1}^{\infty} e^{-r\lambda} I_{n+r}$ and we have used the fact that S_n and T_n are independent. Now,

$$\mathrm{E}(T_n^2) = \sigma_I^2 \sum_{r=1}^{\infty} e^{-2r\lambda} = \frac{\sigma_I^2 e^{-2\lambda}}{1 - e^{-2\lambda}}.$$

Similarly, $\mathrm{E}(S_n^2)$ is bounded by the same constant, and so the right hand side of (C.2) converges for each $\epsilon > 0$.

We turn now to $\sum_{i=1}^{n-1} I_i V_i/n$, to which we apply Theorem E, p. 27 of Serfling (1980). First note that $\mathrm{E}(I_i V_i) = 0$ since I_i and V_i are independent. Also, $I_i V_i$, $i = 1, 2, \ldots$ are uncorrelated, since (for $i < j$)

$$\mathrm{Cov}(I_i V_i, I_j V_j) = \mathrm{E}(I_i V_i I_j V_j) = \mathrm{E}(I_i)\mathrm{E}(I_j V_i V_j) = 0.$$

We have also

$$\mathrm{Var}(I_i V_i) = \mathrm{E}(I_i^2)\mathrm{E}(V_i^2) = \sigma_I^4 \sum_{r=1}^{\infty} e^{-2r\lambda},$$

and hence $\sum_{i=1}^{\infty} \mathrm{Var}(I_i V_i)(\log i)^2/i^2$ converges. It follows that $\sum_{i=1}^{n-1} I_i V_i/n$ converges almost surely to 0. Looking back at the proof, the only condition needed is that $\mathrm{Var}(I_i) < \infty$.

6. and 7. are proven in the same manner with the only condition being that $\sigma_\epsilon^2 < \infty$.

Proof of 8.:

$$\begin{aligned} & \frac{1}{n}\sum_{i=1}^{n} b_{ii}^n \epsilon_i \epsilon_0 \\ = \quad & \epsilon_0 \frac{1}{n}(e^{\lambda} - e^{-\lambda})^{-1}(1 + c_n e^{-2(n+1)\lambda}) \sum_{i=1}^{n}(e^{-(i-1)\lambda} - e^{-(i+1)\lambda})\epsilon_i \\ & \qquad + \epsilon_0 \frac{1}{n} d_n \epsilon^{-2(n+1)\lambda} \sum_{i=1}^{n}(e^{(i+1)\lambda} - e^{(i-1)\lambda})\epsilon_i. \end{aligned}$$

It is enough to show that

$$\frac{1}{n}\sum_{i=1}^{n} e^{-i\lambda}\epsilon_i \xrightarrow{\text{a.s.}} 0,$$

since

$$\begin{aligned}
\frac{1}{n}e^{-(n+1)\lambda}\sum_{i=1}^{n}e^{i\lambda}\epsilon_i &= e^{-\lambda}\frac{1}{n}\sum_{i=1}^{n}e^{(n-i)\lambda}\epsilon_i \\
&= e^{-\lambda}\frac{1}{n}\sum_{r=0}^{n-1}e^{-r\lambda}\epsilon_{n-r}.
\end{aligned}$$

The Borel-Cantelli lemma can be used to show that either of these sums converges almost surely to 0, so long as $\sigma_\epsilon^2 < \infty$.

Proof of 5.: Consider

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}e^{-|i-j|\lambda}I_i\epsilon_j.$$

Using similar arguments to previous, if it can be shown that this term tends to 0 almost surely, then clearly

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}b_{ij}^{n}I_i\epsilon_j \xrightarrow{\text{a.s.}} 0.$$

The sum can be written as

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}I_i\epsilon_i + \frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}e^{-(j-i)\lambda}I_i\epsilon_j + \frac{1}{n}\sum_{i=2}^{n}\sum_{j=1}^{i-1}e^{-(i-j)\lambda}I_i\epsilon_j \\
= \ &\frac{1}{n}\sum_{i=1}^{n}I_i\epsilon_i + \frac{1}{n}\sum_{i=1}^{n-1}I_i\sum_{j=i+1}^{n}e^{-(j-i)\lambda}\epsilon_j + \frac{1}{n}\sum_{j=1}^{n-1}\epsilon_j\sum_{i=j+1}^{n}e^{-(i-j)\lambda}I_i.
\end{aligned}$$

$\frac{1}{n}\sum_{i=1}^{n}I_i\epsilon_i \xrightarrow{\text{a.s.}} 0$ since $I_i\epsilon_i, i = 1,\ldots,n$, are i.i.d. and $\sigma_\epsilon^2 < \infty$, $\sigma_I^2 < \infty$. The other 2 terms can be handled in the same manner as those in the proof of 4.

VITA

Mickey P. Dunlap was born in Union City, Tennessee on August 31, 1972, the son of Mickey C. and Patricia A. Dunlap. He entered the Doctorate program in statistics at Texas A&M University in College Station, Texas which he completed in December 2004. At Mississippi State University, he completed a Master of Science degree in statistics in 1999. A Bachelor of Science degree in mathematics at the University of Tennessee at Martin was completed in 1997. His permenent address is

530 Hazelwood Road

Martin, TN 38237