

Alternative Methods to
Parametric Significance
Testing in Linear Regression and
ANOVA

By

Nhlanhla Makhanya

10308343

Submitted in partial fulfilment of the requirements for the degree

Magister Scientiae (Mathematical Statistics)

in the

Faculty of Natural and Agricultural Sciences

at the

University of Pretoria

Date of Submission

August 2015

Promoter: Prof F.E Steffens

Dr L Fletcher

Declaration

I Nhlanhla Makhanya declare that the dissertation, which I hereby submit for the degree Magister Scientiae (Mathematical Statistics) at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution

Signature

Date

Abstract

Alternative Methods to Parametric Significance Testing in Linear Regression and ANOVA

N.W. Makhanya

Supervisor: Prof F.E. Steffens & Dr L. Fletcher

Department of Statistics

University of Pretoria,

August 2015

The aim of the study was to survey permutation tests, bootstrapping and jackknife methods and their application to significance testing of regression coefficients in linear regression analysis. A Monte Carlo simulation study was performed in order to compare the different methods in terms of empirical probability of type 1 error, power of a test and confidence interval where coverage and average length of confidence interval were used as measures of comparison. The empirical probability of type 1 error and power of a test were used to compare permutation tests, bootstrapping and parametric methods, while the confidence intervals were used to compare jackknife, bootstrap as well as the parametric method. These comparisons were performed in order to investigate the effect of (1) sample size (2) when errors are normally, uniformly and lognormally distributed (3) when the number of explanatory variables is 1, 2 and 5. (4) When the correlation coefficient between the explanatory variables is 0, 0.5 and 0.9. The results obtained from the Monte Carlo simulation study showed that permutation and bootstrap methods produced similar probability of type 1 error results while the parametric methods understated probability of type 1 error when errors are lognormally distributed. In the absence of multicollinearity all the methods were almost equally powerful and in presence of multicollinearity they all suffered equally in terms of power. The jackknife produced poor result in terms of $100(1 - \alpha)\%$ confidence interval while the bootstrap produced reasonable results especially for larger sample sizes. The improvement was observed under the jackknife method when percentile based intervals were applied. It was concluded that permutation tests as well as bootstrap methods are good alternative methods to use in significance testing in regression and ANOVA.

Acknowledgements

I would like to first thank Prof Steffens for all the support given to me since my arrival at the University of Pretoria. He offered me all the support I needed and did not hesitate to write me a motivation for the EUROSA scholarship which was an opportunity to do some of my Masters coursework at Leuven University (Belgium). While am still on EUROSA I would like to thank all parties involved in making the partnership between the European Commission and South African Department of Education a success. This was a once in a lifetime opportunity and an eye opener where students all over the world converge in order to learn together and from one another. I would also like to thank Prof Steffens and Dr L Fletcher who helped me to come up with such an interesting project. If it wasn't for the pair I'm certain that I wouldn't have made it this far. I had a number of instances where I wanted give up because of the challenges that life threw at me. In most cases I went to meetings feeling down and every time I walked out of those meeting feeling more positive. I would also like to thank my wife Sibusisiwe Makhanya for all the support she gave me and for being strong enough when our boy (Olwethu Makhanya) was diagnosed with cerebral palsy. I would also like to thank my managers Stefan Steyn and William Doku at FirstRand for their support and words of encouragement. Last but not least I'd like to thank my family members, friends and colleagues who were with me through it all.

.....
Promoter 1: Prof F.E Steffens

.....
Promoter 2: Dr L Fletcher

Table of Contents

Declaration.....	ii
Abstract.....	iii
Acknowledgements.....	iv
Table of Figures.....	viii
List of Tables	x
Introduction	1
Literature Survey.....	3
2.1 Introduction	3
2.2 Permutation Tests.....	3
2.3 Jackknife.....	6
2.4 Bootstrap	8
2.5 Conclusion.....	10
Significance Testing Using Permutation, Jackknife and Bootstrap in Regression.....	11
3.1 Introduction	11
3.2 Simple Linear Regression	11
3.2.1 Significance testing using permutation tests in simple linear regression.....	11
3.2.2 Significance testing using jackknife method in simple linear regression.....	12
3.2.3 Significance testing using the bootstrapping method in simple linear regression	14
3.2.3.1 Bootstrapping pairs.....	15
3.2.3.2 Bootstrapping residuals	17
3.3 Multiple Linear Regression	19
3.3.1 Significance testing using permutation tests in multiple linear regression.....	20
3.3.1.1 Exact permutation tests	21
3.3.1.2 Permuting raw data	23
3.3.1.3 Permutation of residuals.....	25
3.3.2 Significance testing using bootstrap in multiple linear regression	29
3.4 Analysis of Variance (ANOVA).....	30
3.5 Conclusion.....	31
Confidence Intervals and Percentile Based Intervals.....	32
4.1 Introduction	32
4.2 Confidence Interval by Normal Approach.....	32
4.2.1 Bootstrap confidence interval by Normal Approach	32

4.2.2 Jackknife confidence interval by Normal Approach	36
4.3 Percentile Based Interval	37
4.3.1 Bootstrap Percentile Based interval	38
4.3.2 Jackknife Percentile Based interval.....	38
4.4 Conclusion.....	39
Research Methodology and Design	40
5.1 Introduction	40
5.2 Problem Statement.....	40
5.3 Research Procedure	40
5.4 Conclusion.....	45
Simulation Results.....	46
6.1 Introduction	46
6.2 Probability of type 1 error.....	46
6.2.1 Probability of type 1 error for independent explanatory variables.....	48
6.2.2 Probability of type 1 error in the presence of multicollinearity ($\rho = 0.5$)	59
6.2.3 Probability of type 1 error in the presence of increased multicollinearity ($\rho = 0.9$).....	66
6.3 Power of a test.....	73
6.3.1 Power under independent explanatory variables	73
6.3.2 Power in the presence of multicollinearity.....	83
6.3.3 Power in the presence of increased multicollinearity	90
6.4 Confidence Intervals	97
6.4.1 Normally distributed errors	104
6.4.2 Uniformly distributed errors.....	105
6.4.3 Lognormally distributed errors	105
6.5 Percentile Based Interval	106
6.5.1 Normally distributed errors	114
6.5.2 Uniformly distributed errors.....	115
6.5.3 Lognormally distributed errors	116
6.6 Conclusion.....	117
Conclusion.....	118
References	120
Appendix A – Probability of type 1 error	124
Appendix B – Power of a test.....	127

Appendix C – SAS Programs	130
Probability of type 1 and Power Program.....	130
Confidence Intervals	144
Graph and Confidence Interval tables Preparation	163
Probability of type 1 Error Graphs	164
Power Graphs.....	166

Table of Figures

Figure 1: Type 1 error for $p = 1$ and $\varepsilon \sim \text{Normal}(0, 1)$ on wider vertical axis	48
Figure 2: Type 1 error for $p = 1$ and $\varepsilon \sim \text{Normal}(0, 1)$ on narrow vertical axis	48
Figure 3: Type 1 error for $p = 1$ and $\varepsilon \sim \text{Uniform}(0, 1)$ on wider vertical axis	49
Figure 4: Type 1 error for $p = 1$ and $\varepsilon \sim \text{Uniform}(0, 1)$ on narrow vertical axis	49
Figure 5: Type 1 error for $p = 1$ and $\varepsilon \sim \text{LOGN}(0, 1)$ on wider vertical axis	50
Figure 6: Type 1 error for $p = 1$ and $\varepsilon \sim \text{LOGN}(0, 1)$ on narrow vertical axis	50
Figure 7: Type 1 error for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0$ on wider vertical axis	51
Figure 8: Type 1 error for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0$ on narrow vertical axis	51
Figure 9: Type 1 error for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0$ on wider vertical axis	52
Figure 10: Type 1 error for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0$ on narrow vertical axis	52
Figure 11: Type 1 error for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on wider vertical axis	53
Figure 12: Type 1 error for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on narrow vertical axis	53
Figure 13: Type 1 error for $p = 5$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0$ on wider vertical axis	54
Figure 14: Type 1 error for $p = 5$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0$ on narrow vertical axis	54
Figure 15: Type 1 error for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0$ on wider vertical axis	55
Figure 16: Type 1 error for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0$ on narrow vertical axis	55
Figure 17: Type 1 error for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on wider vertical axis	56
Figure 18: Type 1 error for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on narrow vertical axis	56
Figure 19: Type 1 error for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.5$ on wider vertical axis	59
Figure 20: Type 1 error for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis	59
Figure 21: Type 1 error for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on wider vertical axis	60
Figure 22: Type 1 error for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis	60
Figure 23: Type 1 error for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on wider vertical axis	61
Figure 24: Type 1 error for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis	61
Figure 25: Type 1 error for $p = 5$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.5$ on wider vertical axis	62
Figure 26: Type 1 error for $p = 5$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis	62
Figure 27: Type 1 error for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on wider vertical axis	63
Figure 28: Type 1 error for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis	63
Figure 29: Type 1 error for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on wider vertical axis	64
Figure 30: Type 1 error for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis	64
Figure 31: Type 1 error for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.9$ on wider vertical axis	66
Figure 32: Type 1 error for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis	66
Figure 33: Type 1 error for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on wider vertical axis	67
Figure 34: Type 1 error for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis	67
Figure 35: Type 1 error for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on wider vertical axis	68
Figure 36: Type 1 error for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis	68
Figure 37: Type 1 error for $p = 5$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.9$ on wider vertical axis	69
Figure 38: Type 1 error for $p = 5$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis	69
Figure 39: Type 1 error for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on wider vertical axis	70
Figure 40: Type 1 error for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis	70
Figure 41: Type 1 error for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on wider vertical axis	71
Figure 42: Type 1 error for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis	71

Figure 43: Power for $p = 1$ and $\varepsilon \sim \text{Normal}(0, 1)$ on wider vertical axis.....	73
Figure 44: Power for $p = 1$ and $\varepsilon \sim \text{Normal}(0, 1)$ on narrow vertical axis.....	73
Figure 45: Power for $p = 1$ and $\varepsilon \sim \text{Uniform}(0, 1)$ on wider vertical axis	74
Figure 46: Power for $p = 1$ and $\varepsilon \sim \text{Uniform}(0, 1)$ on narrow vertical axis.....	74
Figure 47: Power for $p = 1$ and $\varepsilon \sim \text{LOGN}(0, 1)$ on wider vertical axis.....	75
Figure 48: Power for $p = 1$ and $\varepsilon \sim \text{LOGN}(0, 1)$ on narrow vertical axis	75
Figure 49: Power for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0$ on wider vertical axis	76
Figure 50: Power for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0$ on narrow vertical axis.....	76
Figure 51: Power for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0$ on wider vertical axis.....	77
Figure 52: Power for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0$ on narrow vertical axis	77
Figure 53: Power for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on wider vertical axis	78
Figure 54: Power for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on narrow vertical axis.....	78
Figure 55: Power for $p = 5$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0$ on wider vertical axis	79
Figure 56: Power for $p = 5$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0$ on narrow vertical axis.....	79
Figure 57: Power for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0$ on wider vertical axis.....	80
Figure 58: Power for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0$ on narrow vertical axis	80
Figure 59: Power for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on wider vertical axis.....	81
Figure 60: Power for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on narrow vertical axis.....	81
Figure 61: Power for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.5$ on wider vertical axis	83
Figure 62: Power for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis.....	83
Figure 63: Power for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on wider vertical axis	84
Figure 64: Power for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis	84
Figure 65: Power for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on wider vertical axis	85
Figure 66: Power for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis.....	85
Figure 67: Power for $p = 5$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.5$ on wider vertical axis	86
Figure 68: Power for $p = 5$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis.....	86
Figure 69: Power for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on wider vertical axis.....	87
Figure 70: Power for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis	87
Figure 71: Power for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on wider vertical axis	88
Figure 72: Power for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis.....	88
Figure 73: Power for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.9$ on wider vertical axis	90
Figure 74: Power for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis.....	90
Figure 75: Power for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on wider vertical axis.....	91
Figure 76: Power for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis	91
Figure 77: Power for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on wider vertical axis	92
Figure 78: Power for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis.....	92
Figure 79: Power for $p = 5$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.9$ on wider vertical axis	93
Figure 80: Power for $p = 5$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis.....	93
Figure 81 : Power for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on wider vertical axis.....	94
Figure 82: Power for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis	94
Figure 83: Power for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on wider vertical axis	95
Figure 84: Power for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis.....	95

List of Tables

Table 1: Confidence Intervals when errors are normally distributed.....	99
Table 2: Confidence Intervals when errors are uniformly distributed	101
Table 3: Confidence Intervals when errors are lognormally distributed	103
Table 4: Percentile based interval when errors are normally distributed	109
Table 5: Percentile based interval when errors are uniformly distributed	111
Table 6: Percentile based interval when errors are lognormally distributed	113
Table 7: Type 1 error when errors are normally distributed	124
Table 8 :Type 1 error when errors are uniformly distributed.....	125
Table 9: Type 1 error when errors are lognormally distributed	126
Table 10: Power of a test when errors are normally distributed	127
Table 11 : Power of a test when errors are uniformly distributed	128
Table 12: Power of a test when errors are lognormally distributed	129

Chapter 1

Introduction

In this project non-parametric/resampling methods will be explored, in particular permutation tests, jackknife and bootstrapping methods and their application to significance testing in regression analysis. The goal of this project is to explore which methods work best in terms of probability of type 1 error, power and confidence intervals and to make recommendations on the appropriate method to use under different conditions namely:

- When errors are normally, uniformly and lognormally distributed;
- Varying number of covariables present in the explanatory variables; and
- Effect of multicollinearity among those exploratory variables

For a very long time parametric methods which relied heavily on normal theory were applied in many situations with inexperienced statistical practitioners applying them without testing whether embedded assumptions are satisfied. Aldrich (2005) gives a good account of the development of regression theory focussing on Fisher's contribution as one of the pioneers. Although R.A Fisher was firstly a pioneer of regression analysis under normal theory, he is also well known for his exploration of permutation/randomisations tests in the "*Design of experiments*" book in 1935 (Kempthorne, 1955). Jackknife and bootstrapping methods emerged much later with the jackknife method first introduced in 1949 (Quenouille, 1949) and bootstrap introduced in 1979 (Efron, 1979). This study will consist of the following:

Chapter 2 will cover the literature review with focus on the history of permutation tests, jackknife and bootstrapping methods.

Chapter 3 will cover significance testing using permutation, jackknife and bootstrap methods in regression analysis.

Chapter 4 will cover the application of permutation tests, jackknife and bootstrap in confidence intervals and percentile based intervals.

Chapter 5 will cover the research methodologies which will detail how these methods were applied in this project.

Chapter 6 will cover the result obtained through simulations and their discussions.

Chapter 7 will be the conclusion detailing the alignment of the findings from the results and similar studies done previously as well as recommendations on the appropriate method for different conditions.

Chapter 2

Literature Survey

2.1 Introduction

This chapter will address the early development of the chosen non-parametric methods, namely permutation test, jackknife and bootstrap methods. The work covered in this chapter is limited to the original use of the methods excluding their use in regression analysis. This aspect will be explored in the next chapter. In the permutation test, earlier work by Neyman in 1923 (Neyman, Dabrowska, Speed, 1990), Pitman (1937a, 1937b, 1938), Fisher in 1935 (Fisher, 1937), Kempthorne (1955) and many others who contributed meaningfully in the early work on the topic will be explored. In jackknife the main authors would include Quenouille (1949), Tukey (1958), Miller (1974) and others. Bootstrap was mainly developed by Efron (1979).

The structure of this chapter will be to follow the history of these methods in their sequence, starting with the oldest of the three, namely permutation tests, followed by the jackknife technique and bootstrap methods. The latter two took the world by storm after computers became available, unlike the previous methods which also relied heavily on the computer although they were very difficult to access at the time. The unavailability of computers was the primary factor in the difficulty of using permutation tests.

2.2 Permutation Tests

Many authors, including Welch (1990), Freedman et al. (1983) and Kennedy (1995) pointed out that the history of permutation tests dates back to 1935 as indicated in R.A.Fisher's book titled *Design of experiments*. Kempthorne (1955) gave a more

detailed description of Fisher's discovery by stating the following: "it seems to have escaped recognition that the physical act of randomisation..., affords the means, in respect of any body of data, of examining the wider hypothesis in which no normality distribution is implied".

Anderson (2001) explained in his paper that Fisher described an experiment to compare growth rate of self-fertilised and cross fertilised *Zea* maize plants. There were in total 15 pairs of *Zea* maize plants where each i^{th} pair can be denoted as (x_{Ai}, x_{Bi}) where A and B represent self-fertilised and cross-fertilised plants respectively. The pairs are randomly allocated to the plot. The test statistic used was the difference in growth rate say $d_i = x_{Ai} - x_{Bi}$. Let D denote the distribution of sum of differences from all possible outcomes under the null hypothesis and the sum of observed differences by D_{obs} . The p value was then computed as $p\ value = \frac{\text{no.of values where } |D| \geq D_{obs}}{\text{possible no.of values of } |D|}$. Kempthorne (1955) further mentions Fisher's findings that he observed only 1726 of 2^{15} combinations of 15 pairs from two samples which had differences greater than or equal to the observed difference. This gave a significance level of $\frac{1726}{2^{15}} = 0.05267$ while the normal theory produced a significance level of 0.05158 after applying the Yates correction for continuity.

A well-structured and comprehensive work came two years after Fisher's work in 1937 by Pitman (1937a, 1937b, 1938) who produced a series of papers titled *Significance tests which may be applied to samples from any population* I, II, III. Papers in this order dealt with tests of significance of difference of means which is an equivalence of *Student t* under normal theory, significance tests for linear dependencies which is an equivalence of the correlation coefficient under normal theory and lastly the analysis of variance test of significance which is an equivalence of *F* tests under normal theory. In his papers he acknowledged Fisher as the first author to come up with the idea of permutation tests by saying that "the idea is not new, it seems to be implicit in all Fisher's writings".

Other authors such as te Braak (1992) and Anderson (2001) contend that permutation tests date further back to 1923 by Neyman. However, many authors failed to consider the Neyman paper because it was written in Polish with a summary

only in German. It was not until 1990 when the majority of the statistical community was exposed to the earlier work by Neyman (1923) when Dabrowska and Speed translated Neyman's paper from Polish to English. Rubin (1990) discussed the work that is contained in Neyman's paper that was published in 1923 and also clarified contributions by Neyman and R.A. Fisher on the topic. Rubin mentioned the following: "To my knowledge, this paper represents the first attempt to evaluate, formally and informally, the repeated-sampling properties of statistics over their nonnull randomisation distributions, and so I believe this contribution is uniquely and distinctly Neyman's". He also mentioned that "Fisher's mode of randomisation-based inference in experiments was distinctly different from Neyman's". He concluded section 6 of the same paper by mentioning that "Neyman's prescription offers a general plan for evaluating proposed procedures, whereas Fisher's prescription directly provides distribution free p-values for sharp null hypothesis. I find the approaches complimentary".

Kempthorne (1955) discussed the randomisation tests in the linear models for experimental designs in cases of completely randomised design, randomised block designs as well as Latin squares with much emphasis on completely randomised design. His paper generally addressed the comparative experiment where for example, the difference between two yields is established after different fixed treatments have been applied to the plot. In his paper he showed how randomisation can be used to make the statistical inference in the comparative experiment. He does not deal with the type of comparative experiment with a population for which inference needs to be made by using a sample for that population. As Pitman (1937a, 1937b, and 1938) indicated, the randomisation theory can be applied to samples without any knowledge of population.

All the work done by the authors mentioned here and other authors of that time was mainly centred on the design of experiments and observational studies. The use of permutation tests is permissible by randomisation in design of experiments and the exchangeability in observational studies. Their use however was not practical because they are computer intensive and the lack of computer availability made it practically impossible to perform them in real problems, even those of small

samples. Pitman (1938) showed that the F test is an approximation to permutation tests which is an exact test; hence their use was limited to justifying the validity of the ANOVA F test.

2.3 Jackknife

In 1949 Quenouille(1949) introduced the jackknife technique to eliminate the bias on the serially correlated estimators in time series by splitting the sample into two groups. He explored the splitting of the sample size $n=kg$ into g groups of size k . In his review of the jackknife technique Miller(1974) pointed out that much of the theory developed for jackknife technique uses $n=g$ and $k=1$. Quenouille introduced the use of jackknife as a technique for bias reduction while Tukey (1958) introduced it as a technique to obtain robust interval estimation. Adopting the Huber (2002) notation, the idea of the jackknife technique is as follows:

DEFINITION. Let $x_1, x_2 \dots \dots x_n$ be a sample of independent and identically distributed random variables and Let $T_n = T_n(x_1, x_2, \dots x_n)$ be an arbitrary statistic based on sample size n and $T_{i,n-1} = T_{n-1}(x_1, x_2, \dots x_{i-1}, x_{i+1}, \dots, x_n)$ be a statistic based on sample size $n - 1$ with an observation x_i omitted. A bias reduced estimate T_n^* estimates the same quantity as T_n based on the sample of size n . Define

$$T_{i,n}^* = nT_n - (n - 1)T_{i,n-1} \quad (2.1)$$

The estimator $T_{i,n}^*$ above measures the contribution of x_i to T_n .

The Quenouille's estimator for the mean

$$T_n^* = n^{-1} \sum_{i=1}^n T_{i,n}^* \quad (2.2)$$

has a smaller bias than T_n .

Tukey pioneered the use of jackknife for interval estimation. He wanted to assess the influence of the individual point or observation on the statistic of interest and tools for estimating its variability (Huber; 2002). The tool for bias reduction developed by Quenouille (1949,1956) was best suited for his purpose. He referred to the technique as “jackknife”. Tukey established that

$$\frac{1}{n(n-1)} \sum_{i=1}^n (T_{i,n}^* - T_n^*)^2 \quad (2.3)$$

is a variance estimate for both T_n and T_n^* . He further proved that the statistic

$$\frac{\sqrt{n}(T_{i,n}^* - T_n^*)}{\left(\frac{1}{n(n-1)} \sum_{i=1}^n (T_{i,n}^* - T_n^*)^2\right)^{1/2}} \quad (2.4)$$

has an approximate t distribution with $n - 1$ degrees of freedom and for large n the statistic is approximately normally distributed, which is useful for robust interval estimation. Although this result has been shown for a special case where $g = n$ and $k = 1$ it also holds for a general case where there are g groups of k observations. In such a case n is replaced by g in the above formula. This was a very important contribution by Tukey that lead to the technique being called Tukey-Quenouille jackknife in recognition of the authors’ contribution. Brillinger (1964) used maximum likelihood methods to show that the statistic (2.4) has a limiting t -distribution with $g - 1$ degrees of freedom if g is fixed and k tends to infinity. He also outlined the procedure for constructing approximate confidence intervals for the parameter being estimated by an estimator (2.2). This is done by calculating the maximum likelihood estimates which are independent and asymptotically normal for each of g groups of k observations. These estimates can be used for constructing the t statistic and their associated confidence intervals. The smaller mean square error is achieved by increasing the number of groups rather than their sizes (Miller, 1974).

Some of the authors who made further contributions to the technique for bias reduction are Schucany, Gray and Owen (1971) in the second-order jackknife (Miller,

1974). In the same paper Schucany, Gray and Owen further addressed the cases where bias is of order greater than 2 to provide the jackknife generalisation (Miller, 1974). Durbin (1959) applied the jackknife to ratio estimation by investigating what happens to the estimator

$$T = 2T_1 - \frac{1}{2}(T_2 + T_3) \quad (2.5)$$

where T_1 is an estimate based on n observations and T_2 and T_3 are estimates based on two groups of size $\frac{1}{2}n$. The estimator T has bias of $O(n^{-2})$. The ratio estimator y/x is applied to the simple linear regression $y = \alpha + \beta x + \varepsilon$, where x is normally distributed. By ignoring the $O(n^{-4})$ Durbin (1959) established that the estimator T has a smaller bias and variance than the normal ratio estimator $R = \bar{Y}/\bar{X}$. When x has a Gamma distribution and coefficient of variation less than $\frac{1}{4}$ he established that the estimator T above has smaller bias and mean square error compared to the estimator R , but as pointed out by Quenouille (1949 and 1956) the reduction in bias often leads to the increase in variance. That increase is however, of smaller order compared to the variance itself. The variance of t is slightly higher than that of R . This is not of major concern since the estimator with bias and smaller mean square error is preferable. The optimum number of groups for the jackknife is n for the normal distribution i.e. each observation in its own group (Rao, 1965). This also holds for the gamma distribution (Rao and Webster, 1966).

2.4 Bootstrap

Since its inception by Quenouille (1949) the jackknife gained momentum for bias reduction and variance estimation. It was not until 1979 that Efron defined another robust non-parametric method similar to jackknife and referred to it as bootstrap. The two techniques are very similar: jackknife can be viewed as drawing samples of size $n-1$ without replacement while bootstrapping can be viewed as drawing the samples of size n with replacement. Since its inception, bootstrap gained a lot of ground and

became a widely used technique (Efron, 1979) as it proved to be more reliable than the traditional jackknife. Efron further showed that the jackknife method can be viewed as a linear expansion to the approximation of bootstrap. The problem statement in bootstrapping is:

Given the random variables $\mathbf{X} = (X_1, X_2, \dots, X_n)$ its realization $x = (x_1, x_2, \dots, x_n)$ from the unknown distribution function F . The bootstrap estimates in case of one sample can be estimated in three steps as outlined in Efron (1979) namely:

1. Construct the empirical distribution function \hat{F} , putting mass $1/n$ at each point x_1, x_2, \dots, x_n .
2. With \hat{F} fixed, draw an independent random sample of size n from \hat{F}
3. Estimate the sampling distribution of the statistic of interest using the bootstrap distribution of the statistic of interest.

If the empirical distribution $\hat{F} = F$ then the results derived from the bootstrap distribution should be close to the one calculated from the original sampling distribution (Efron, 1979). The effectiveness of the bootstrap estimator as estimator depends on the choice of the statistic being measured. Three methods of calculating the bootstrap distribution as outlined in Efron (1979) are:

1. Direct theoretical calculations.
2. Monte Carlo approximation to the bootstrap distribution.
3. Taylor series expansion methods can be used to obtain the approximate mean and variance of the bootstrap distribution.

The third method is the same as using a certain form of jackknife (Efron, 1979). However, the bootstrap method gained a lot of popularity because it can be easily implemented on the computer. The other work that Efron covered in his 1979 paper was the use of bootstrapping method in sample median estimation where the jackknife is known to fail. This was another reason why bootstrapping was seen as superior to the standard jackknife method. In this problem standard jackknife does not produce the asymptotically consistent variance for the sample median while the bootstrap produces the correct asymptotic variance. Bootstrapping also outperforms

the crossvalidation and other nonparametric methods in estimating the error rates in discriminant analysis (Efron, 1979).

2.5 Conclusion

In this chapter permutation tests were discussed from their inception. Their use at the time was limited by the unavailability of computers and they served as a motivation for the validity of ANOVA F tests. Most statisticians at that time were happy using the normal theory in dealing with significance testing. In design of experiment problems, permutation tests had lot of grounds and were perfectly justifiable under randomisation. In 1945 a bias reduction method by Quenouille emerged which Tukey later named “jackknife”. The development of jackknife applied to time series problems but it never gained momentum in this field. It however, quickly found its use in a wider spectrum of problems which Miller reviewed (1974). After carefully studying the jackknife Efron (1979) found some limitations of jackknife in its form as developed by Quenouille. This gave birth to a new method closely related to the jackknife which Efron (1979) created in his paper titled: *Bootstrap another look at jackknife*. This method gained a lot of ground because of its flexibility and it could be applied to a much wider spectrum of problems.

Chapter 3

Significance Testing Using Permutation, Jackknife and Bootstrap in Regression.

3.1 Introduction

This chapter explores significance testing using permutation tests, the jackknife and bootstrap methods in regression. The simple linear regression case is explored in some detail for all the methods. In cases where extension from simple linear to multiple linear regression is straight forward the procedure/algorithm is briefly discussed as would be seen to be the case for bootstrap and jackknife. The multiple regression case in permutation tests does not follow directly from the simple linear regression (Kennedy, 1995) unlike jackknife and bootstrap.

3.2 Simple Linear Regression

3.2.1 Significance testing using permutation tests in simple linear regression

There is a general consensus on the application of permutation tests in simple linear regression where Y is regressed against X (Anderson, 2001) and (Kennedy et al., 1996). In the equation

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (3.1)$$

where β_0 is the intercept, β_1 is the slope of the straight line, X_i is considered fixed and ε_i is the random error. The null hypothesis of interest is $H_0: \beta_1 = 0$ and the alternative is $H_1: \beta_1 \neq 0$. Under normal circumstances the significance test that would be applied to test the null hypothesis is the *Student t* test i.e. let n be a sample size,

$se(\hat{\beta}_1)$ be a standard error of $\hat{\beta}_1$, α be level of significance then the two sided t statistic is

$$t_0 = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)} \sim t_{n-2}. \quad (3.2)$$

The null hypothesis is rejected if $|t_0| \geq t_{\alpha/2; n-2}$ or if the probability of obtaining the test statistic that is at least as extreme as the observed t value i.e. the p -value is less than the level of significance α .

In an event where errors in (3.1) are not identically and independently normally distributed with constant variance, the normality assumption in the errors can be substituted with an exchangeability assumption and a permutation test applied as follows:

1. Compute the t value, say t_0 on the original sample.
2. Generate permuted samples by permuting either X or Y .
3. Compute the t values for each permuted sample (say t_i).
4. Construct the distribution of the resulting t values.
5. Plot the value of t_0 on the distribution constructed in step 4.

The null hypothesis is rejected if the proportion of t values greater than or equal to t_0 in absolute value, are less than the level of significance α . In practice the sample sizes are such that it is not computationally feasible to use all possible permutations. Step 2) will then be replaced by taking a random sample of permutations.

3.2.2 Significance testing using jackknife method in simple linear regression

The jackknife method developed by Quenouille (1949) for bias reduction and Tukey (1958) for variance estimation provides another resampling method used for statistical significance test under non-normality assumptions. Jackknife is closely related to the bootstrapping method (Efron, 1979). Here the significance testing will

be explored by using the jackknife method. The two choices that will be explored in this section are traditional Quenouille (1949) methods by delete 1 jackknife from the original sample and delete d jackknife (Efron and Gong, 1983). As the names suggest, deleting 1 jackknife deletes one observation at a time and computes the estimates of interest i.e. each statistic is computed from the n samples of size $n-1$ each, while the delete d jackknife has samples of size $n - d$. Wu (1986) also proposes using the weighted jackknife estimates which is robust for the variance of $\hat{\beta}_1$ against the effects of heteroskedastic errors.

The simple linear regression model (3.1) will now be considered. The steps for delete 1 jackknife significance testing for null hypothesis ($H_0: \beta_1 = 0$ vs $H_1: \beta_1 \neq 0$) is as follows:

1. Regress Y on X using equation (3.1) to obtain the least squares estimates.

$\hat{\beta}_0, \hat{\beta}_1$ and the standard error $se(\hat{\beta}_1)$, then compute the t statistic $t_0 = \hat{\beta}_1 / se(\hat{\beta}_1)$

2. Order the pairs (X_r, Y_r) $r = 1, \dots, n$ of observations in a sequence.
3. Omit the r^{th} pair (X_r, Y_r) at a time and draw a sample of the remaining $n - 1$ pairs.
4. Compute the least squares estimator $\hat{\beta}_1^{Jr}$ $r = 1, \dots, n$ and its standard error from the sample with the r^{th} pair omitted.
5. Compute the respective the t statistic as $t^{Jr} = \frac{\hat{\beta}_1^{Jr} - \hat{\beta}_1}{se(\hat{\beta}_1^{Jr})}$, $r = 1, \dots, n$ for the sample.
6. Repeat step 3 to 5 and generate the empirical distribution of t^{Jr} .
7. The t statistic t_0 from the original sample is plotted on the distribution of t^{Jr} .

The null hypothesis is rejected against a two sided alternative if the absolute value of t_0 is greater than or equal to the $100 \left(1 - \frac{\alpha}{2}\right)$ percentile of jackknife distribution. The delete d jackknife follows a similar algorithm. The only difference is at step 3 where d -observations are deleted each time resulting in smaller sample sizes ($d > 1$) for

each jackknife sample. The sample sizes are $n - d$ and there would be $S = \binom{n}{d}$ samples.

3.2.3 Significance testing using the bootstrapping method in simple linear regression

As discussed in the previous chapter, bootstrapping methods are well known for their credibility in estimating confidence intervals. Their use in significance testing was not studied in great detail until almost a decade later. Among the first authors to explore the use of bootstrap method in significance testing were Hinkley (1988), Hall and Wilson (1991). The close relationship between confidence intervals and hypothesis testing made studying hypothesis testing using bootstrap tests easier since the use of bootstrapping in confidence interval was studied in great details.

Hall and Wilson (1991) gave two guidelines for the use of bootstrapping in hypothesis testing. “The first guideline says that care should be taken to ensure that even if the data might be drawn from the population that fails to satisfy H_0 , resampling is done in the way that reflects H_0 ” (Hall and Wilson, 1991). “The second guideline argues that bootstrap hypothesis testing should use methods that are already recognised as having good features in the closely related problem of confidence interval construction” (Hall and Wilson, 1991). The second guideline immediately connects bootstrap hypothesis testing with bootstrap confidence intervals where great strides had been made in terms of theory. Hall and Wilson (1991) argue that the first guideline has a direct impact on the power of a test. This implies that not adhering to the first guideline, compromises the power of the test against the alternative hypothesis. The second guideline does not have a direct impact on power. However its non-adherence has a huge impact on the coverage accuracy. This reason shaped the direction for hypothesis testing using bootstrapping methods. Hall and Wilson (1991) mentioned that these guidelines had been implicit in the work done prior to 1991 but due to heavy theory in the papers presented before, it could have easily escaped the attention of biometricians with a resulting non adherence to the guideline. In both bootstrapping methods to be

considered in the next section both possibilities of resampling the observations and residuals will be explored

Consider the regression equation (3.1).

There are two choices on what to bootstrap, namely;

- 1) Bootstrap pairs (X_i^*, Y_i^*) $i = 1, 2, \dots, n$ of size n with replacement from $(X_1, Y_1), (X_2, Y_2); \dots (X_n, Y_n)$.
- 2) Bootstrap the residuals ε_i^* $i = 1, 2, \dots, n$ of size n with replacement from $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_n$

The choice of an appropriate method is determined by the practitioner and depends on the properties of the dataset. There is no single method that works well in all different situations as both methods have their advantages and disadvantages. This implies that certain assumptions are required before the choice of method to use is made. The two methods will be dealt with in some detail in the next subsections.

3.2.3.1 Bootstrapping pairs

Method 1 (bootstrapping pairs) which was proposed by Freedman (1981) implicitly means that there is no conditioning on X_t because each bootstrap sample has a different X_t . It is on this note that Hinkley (1988) criticized this method after the observation that method 1 might not be suitable for the regression problem since the inference in regression model is made conditional to say $G = (x_1, x_2, \dots, x_n)$ and by bootstrapping the X 's one runs a risk of simulating a dataset whose distribution $G^* = (x_1^*, x_2^*, \dots, x_n^*)$ is very different from G . Despite Hinkley's view, bootstrapping pairs is a very useful tool. MacKinnon (2002) outlines the advantages of this method as follows;

- Bootstrapping pairs works well even when there is a presence of heteroskedasticity in the residuals.

- It works well for dynamic models i.e. regressing the dependent variable on the lagged dependent variables.
- It can be applied in a wide range of models.
- In multivariate models the pairs and residual bootstrap can be combined by organising residuals as a matrix and applying pairs on its rows, thus preserving cross-equation correlations.

The general procedure of significance testing in a regression model in the presence of heteroskedasticity is outlined below:

Consider simple linear regression

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \quad u_i = \sigma_i \varepsilon_i, \quad (3.3)$$

where σ_i^2 is the variance of the error terms, here the dependent and independent variable pair is sampled and each bootstrap regression equation is

$$Y_i^* = \beta_0 + \beta_1 X_i^* + u_i^* \quad (3.4)$$

where each pair (X_i^*, Y_i^*) is drawn from the pairs (X_i, Y_i) where $i = 1, 2, \dots, n$. Suppose that the hypothesis of interest is $H_0 : \beta_1 = 0$ and $H_1 : \beta_1 \neq 0$ as is generally the case. The least squares estimate $\hat{\beta}_1$ is calculated for the regression equation. A usual t test is computed

$$t_0 = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)} \quad (3.5)$$

where $se(\hat{\beta}_1)$ is the standard error of $\hat{\beta}_1$. Suppose α is the level of significance. The same procedure is applied B times on B bootstrap samples calculating the least squares estimators (say $\hat{\beta}_{1,j}^*$) and their corresponding standard errors ($se(\hat{\beta}_{1,j}^*)$) then the t statistic is calculated for each j^{th} bootstrap sample as follows

$$t_j^* = \frac{\hat{\beta}_{1,j}^* - \hat{\beta}_1}{se(\hat{\beta}_{1,j}^*)}$$

according to Hall and Wilson (1991) second guideline which states that the test statistics should be based on $\frac{\hat{\theta}^* - \hat{\theta}}{\hat{\sigma}^*}$ where $\hat{\theta}^*$ is a bootstrap estimate, $\hat{\theta}$ is the estimate based on the original sample and $\hat{\sigma}^*$ is the standard error of a bootstrap estimate, in this study the above quantities are given by $\hat{\beta}_{1,j}^*$, $\hat{\beta}_1$ and $se(\hat{\beta}_{1,j}^*)$.

The empirical distribution of bootstrapped *t statistic* t_j^* , $j = 1, \dots, B$ is constructed and the value of t_0 is plotted on the bootstrapped distribution. The null hypothesis of β_1 is rejected in favour of alternative if the absolute value of t_0 is greater than or equal to the $100 \left(1 - \frac{\alpha}{2}\right)$ percentile of the bootstrap distribution.

3.2.3.2 Bootstrapping residuals

The bootstrapping residuals have been generally accepted by many authors and the method allows for direct comparison with other re-sampling methods like permutation test. A practitioner needs to answer two questions before applying the bootstrapping residual method, namely:

- Are the errors identically distributed?
- Are the errors independent?

If the answer is yes to both questions, then the bootstrapping residuals generally outperform the bootstrapping pairs method. From this point of view the bootstrapping residuals methods will be dealt with in detail. Two methods of bootstrapping residuals will be considered, namely restricted residuals and unrestricted residuals.

Consider equation (3.1)

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad i = 1, 2, \dots, n \quad (3.6)$$

with ε_i identically distributed with mean zero and a constant variance. The test of significance for the least squares estimator $\hat{\beta}_1$ can be conducted firstly in a way that

reflects the null hypothesis even if the bootstrap samples are drawn for the population that fails to satisfy the null hypothesis (Hall and Wilson, 1991). Paparoditis and Politis (2005) outline the steps as follows:

1. Regress Y on X as in equation (3.8) to obtain the least squares estimates $\hat{\beta}_0, \hat{\beta}_1$ and the standard error of $\hat{\beta}$ then compute the t statistic t_0 .
2. Assume that the null hypothesis is true, then fit the constant model to obtain the estimate $\hat{\beta}_0$.
3. Estimate the residuals under null hypothesis ($H_0 : \beta_1 = 0$) as $\hat{\varepsilon}_i = y_i - \hat{\beta}_0$.
4. Sample the residuals ε_i^* $i = 1, 2, \dots, n$ with replacement from the empirical distribution $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_n$.
5. Generate the corresponding Y observations from $Y_i^* = \hat{\beta}_0 + \varepsilon_i^*$.
6. Regress the bootstrapped Y i.e. $E(Y^*|X) = \beta_0 + \beta_1 X$ to obtain the least squares estimator $\hat{\beta}_1^*$ and its standard error from the bootstrap sample.
7. Repeat step 4 to 6 B times and compute t_j^* statistic for each bootstrap sample and generate the empirical distribution of the t_j^* .
8. Plot t_0 on the bootstrapped distribution.

The null hypothesis of β_1 is rejected by the bootstrap method if the absolute value of t_0 is greater than or equal to the $100 \left(1 - \frac{\alpha}{2}\right)$ bootstrap percentile $t_{1-\frac{\alpha}{2}}^*$. Paparoditis and Politis (2005) also provide an alternative way of applying the bootstrapping residual. In that case the null hypothesis is not imposed in the way that residuals are bootstrapped. This is referred to as an unrestricted residual bootstrapping (te Braak, 1992). The bootstrap residuals ε_i^* $i = 1, 2, \dots, n$ are sampled from the empirical sample of residual $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_n$ generated as

$$\hat{\varepsilon}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i$$

The choice of method to use for bootstrapping residuals (restricted vs. unrestricted) depends on the choice of test statistic. In the above procedure (step 7) the t -statistic is used as a pivotal quantity. In this case the results from both methods are

approximately equivalent, regardless of whether the null hypothesis is true or not. According to Hall and Wilson (1991) the first guideline for bootstrap hypothesis testing is that the first method is appropriate. Many authors have advocated the use of pivotal statistic in the bootstrap hypothesis testing because they are robust enough to be insensitive to the choice of the set of residuals used in re-sampling step (Paparoditis and Politis, 2005).

If a non-pivotal statistic is used in step 7 of the above procedure, the choice of method to use is extremely critical because if the null hypothesis (H_0) is true then the two cases are approximately equivalent and the restricted residuals are generally preferable. However if the alternative hypothesis (H_1) is true then the first case (restricted) can be very erratic and the unrestricted residuals are then preferable (Paparoditis and Politis, 2005). The reason is that if the alternative hypothesis is true, then the empirical variance of restricted residuals is greater than that of the unrestricted residuals based on the least squares estimator. Although both methods achieve a desired level of test, there are power concerns when the restricted residuals method is applied because it “may fail to identify the optimum critical region for the test under alternative hypothesis” (Paparoditis and Politis, 2005). It is for this reason that the pivotal statistic is recommended for this type of bootstrap hypothesis testing.

3.3 Multiple Linear Regression

The next sections will be outlining the use of permutation tests, jackknife and bootstrapping methods in the significance testing in multiple regression. In matrix notation the model is described as follows,

$$Y = X\beta + Z\gamma + \varepsilon \quad (3.7)$$

where Y is an $(n \times 1)$ response vector, X and Z are $(n \times p)$ and $(n \times q)$ matrices of explanatory variables respectively containing β and γ of size p and q unknown regression coefficients and ε the error distributed with mean zero and variance covariance matrix $I\sigma^2$. The null hypothesis test is $H_0: \beta = \mathbf{0}$ vs $H_1: \beta \neq \mathbf{0}$ ie. the effects of variable X in predicting Y in the presence of the variable Z is zero.

The following section will begin by showing why the extension from simple linear regression is not straight forward.

3.3.1 Significance testing using permutation tests in multiple linear regression

In section 3.2.1 significance testing in simple linear regression using permutation test was considered. In simple linear regression the *t-statistic* was used to derive the *p-value*, however other equivalent choices of the statistic can be considered in deriving the *p-values*. For a simple linear regression model permuting the dependent variable or permuting the explanatory variable was equally correct.

$$S_{xy} = \sum_{i=1}^n (x_i y_i) \quad (3.8)$$

is monotonically related to the least squares estimate

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} \quad (3.9)$$

the *t statistic*

$$t = \frac{S_{xy}}{S_{xx} se(\hat{\beta}_1)} \quad (3.10)$$

and the square of the correlation coefficient

$$r^2 = \frac{S_{xy}^2}{S_{xx} S_{yy}} \quad (3.11)$$

When permuting X variable only S_{xy} the numerator is affected hence any of the test statistics outlined above are “equivalent” statistics under permutation (Kennedy and Cade, 1996). This relationship only works for simple linear regression. It does not hold under multiple linear regression because of the presence of the covariables. The theory takes a different direction when multiple linear regression is considered.

In multiple regression there are many choices of what to permute, unlike in simple linear regression where the response variable Y is permuted (Anderson & Robinson, 2001). Several permutation methods in multiple linear regression have been explored in the literature. Manly (1991), te Braak (1992), Kennedy (1995), Kennedy and Cade (1996), Anderson and Legendre (1999), Anderson and Robinson (2001) are among those who gave an account of these different methods. Anderson and Legendre (1999) and Kennedy (1995) applied Monte Carlo methods in different permutation tests and found that the tests were not exact and none of them gave identical results. Anderson and Robinson (2001) gave a theoretically exact permutation test for the null hypothesis.

3.3.1.1 Exact permutation tests

This subsection will follow the work of Anderson and Robinson (2001). Permutation tests considered in section 3.2.1 are exact tests and Anderson and Robinson (2001) state that even when n is large and it is not possible to consider all $n!$ permutations, taking a subset $m \leq n!$ the test is still exact. The multiple regression exact tests are often not possible in practice because they require knowledge of the relationship of the response variable and the explanatory covariable under the null hypothesis (Anderson and Robinson, 2001). Exact permutation tests in multiple regression proceed as follows:

Consider the following univariate case of equation (3.7). Without loss of generality if Y, X and Z are standardised variables with mean zero i.e.

$$Y_i = \beta_1 X_i + \beta_2 Z_i + \varepsilon_i \quad (3.12)$$

According to equation (3.12) when the null hypothesis is true, Y is directly related to Z and only related to X through Z giving the two equations

$$Y_i = \delta Z_i + \mu_i \quad (3.13)$$

and

$$X_i = \gamma Z_i + u_i. \quad (3.14)$$

Instead of the *t-statistic* Anderson and Robinson (2001) used the squared correlation coefficient as a test statistic when examining the different permutation methods, which uses the relationship of X and Z in equation (3.14). Anderson and Legendre (2001) showed that the squared correlation coefficient is a pivotal test statistic and also showed the relationship between the squared correlation coefficient and the *t-statistic*.

Assuming that the null hypothesis is true as it is the case in Hall and Wilson (1991) guideline. Under the null hypothesis $H_0: \beta_1 = 0$, $Y_i = \delta Z_i + \mu_i$ and adopting the notation used by Anderson and Robinson (2001) the hypothesis testing is based on

$$r^2 = \frac{\sum\{(Y_i - \hat{\delta}Z_i)(X_i - \gamma Z_i)\}^2}{\sum(Y_i - \hat{\delta}Z_i)^2 \sum(X_i - \gamma Z_i)^2} \quad (3.15)$$

with $\hat{\delta} = S_{zy}/S_{zz}$ and for fixed X and Z, γ is known. (3.15) provides the means to do

the partial test for the relationship between Y and Z in the presence of other relationship of X and Z. If δ is known then the exact test can be obtained by permuting Y conditioned on X. The residuals can be calculated from the known relationship between Y and X and from the distribution of true errors μ_i , $i = 1, \dots, n$ permute the residuals μ_i^π , $i = 1, \dots, n$. Using permuted residuals under the null hypothesis the true observations Y conditional on X can be generated as

$$Y_{i(\pi)} = \delta Z_i + \mu_i^\pi \quad (3.16)$$

where the π superscripted variable is the permuted variable and the π subscripted variables are derived from the permuted and non-permuted variables. The test statistic under the null hypothesis for an exact test uses equation (3.15) with $\hat{\delta}_\pi = \frac{\sum Y_{i(\pi)} X_i}{\sum X_i^2}$ and $Y_{i(\pi)}$ i.e.

$$r_\pi^2 = \frac{\sum \{(Y_{i(\pi)} - \delta_\pi Z_i)(X_i - \gamma Z_i)\}^2}{\sum (Y_{i(\pi)} - \delta_\pi Z_i)^2 \sum (X_i - \gamma Z_i)^2} \quad (3.17)$$

where $Y_{i(\pi)} - \delta_\pi X_i$ is the residual of $Y_{i(\pi)}$ without the effect of X . The p-value of an exact test of hypothesis is the proportion of r^2 calculated from the $n!$ permutations that are greater than or equal to the r^2 calculated from the original sample. This test depends on the order statistic of the original Y observations (Anderson and Robinson, 2001). However as mentioned earlier this exact test assumes that δ is known which is not the case in reality, making it impossible to conduct this test in a practical situation. This resulted in the reliance on approximate permutation tests to perform significance testing in multiple regression. Approximate permutation tests fall within two categories namely permuting raw data and permuting residuals. There are choices on what to permute, namely raw data permutation and the model under which residuals are permuted. These will be dealt with in the next subsection.

3.3.1.2 Permuting raw data

Two methods of permutation tests that have been studied in literature are permutation of the dependent variable Y which authors such as Kennedy and Cade (1996) and Anderson and Robinson (2001) attribute to Manly (1991), and permutation of the X variable which te Braak (1992) and Kennedy and Cade (1996) have mainly attributed to Collins (1987) and Oja (1987). From equation (3.14) X , Y and Z are multivariate variables but for notational ease X , Y and Z are restricted to a univariate case. The univariate version of equation (3.14) is equation (3.15). Without

loss of generality Y , X and Z are standardised variables with mean zero. The hypothesis test of interest is $H_0: \beta_1 = 0$ against the two sided alternative.

Permute the X variable

The following algorithm can be used for test of hypothesis when the X covariable is permuted. From the multiple regression equation (3.12):

1. Fit the regression model to the data and obtain the least squares estimate, its standard error and compute the t statistic.
2. Permute the X variable to generate the permuted X^π observations.
3. Regress the original Y on the permuted X and original Z .
4. Repeat step 2 and 3 a number of times to obtain the t -statistic distribution.
5. Plot the t -statistic computed from the original data on the distribution of the permuted t -statistic. The p -value is the proportion of $t_{i(\pi)} \geq t_0$.

This method has been criticised by te Braak (1992) and Welch (1990) because it does not preserve the collinearity between the predictor variables X and Z as it violates the ancillary principle. Kennedy and Cade (1996) however point out that if the pivotal statistic is employed, the collinearity in the original sample should not affect the test. Under a simple linear regression, the least squares estimate is equivalent in randomisation to a t -test (refer to equations 3.8 to 3.11). According to Kennedy and Cade (1996), Oja and Collins extended this equivalence to the multiple regression where it is no longer applicable. The least squares estimate of the parameters in equation (3.14) is given by

$$\hat{\beta} = \frac{S_{xy}S_{yy} - S_{zy}S_{xz}}{S_{xx}S_{yy} - (S_{xz})^2}. \quad (3.18)$$

Clearly the denominator of the least squares estimate is affected by the permutation of the X variable in this case, hence the equivalence relation does not hold anymore.

This method is only useful in repeated samples if it uses a pivotal statistic (Kennedy and Cade, 1996).

Permute the Y variable

Permuting X method is not commonly used primarily because of its violation of the “ancillary principle”¹ because it ignores the collinearity between X and Z (Welch, 1990). The alternative method of permuting a raw variable which does not violate the ancillary principle is permuting the Y variable method. The same procedure used in the permute X method is applicable to permute Y method with the exception of step 3. In step 3 the permuted Y variable is regressed against X and Z (unpermuted) to obtain the least squares estimate of β . The preference of the permute Y method over permute X methods arises from this point since it preserves collinearity between X and Z (Anderson and Legendre, 1999). Kennedy and Cade (1996) were not in favour of this method claiming that this method is not justified when the nuisance parameter β_2 in equation (3.12) is not equal to zero. Anderson and Legendre (1999) are of a different view when they mentioned that “The rationale for this method is that the permutable units for the test are the original Y values, independent of any model which might be imposed: that is any value of Y could have been observed associated with any combination of paired values (X,Z) ”. The common criticism from Anderson and Legendre (1999) and Kennedy and Cade (1996) was Manly’s use of non-pivotal statistic (least squares estimate $\hat{\beta}_1$).

3.3.1.3 Permutation of residuals

Two methods of permuting residuals will be discussed in this subsection namely: permutation of residuals under the reduced model and permutation of residuals under the full model.

¹ The principle underlying hypothesis testing is to compare what was actually observed with what could have been observed in hypothetical repetitions of the original experiment, under the null hypothesis (Sandved, 1966)

Permuting residuals under the reduced model

The methods by Freedman and Lane (1983) and Kennedy (1995) are two closely related methods which both permute the residuals under the reduced model. Anderson and Legendre (1999) outline the procedure as follows:

1. Regress Y on both X and Z as in equation (3.12) to obtain the least squares estimates of $\hat{\beta}_1$ of β_1 and t -statistic t_0 .
2. Regress Y on Z as in equation (3.13) and obtain the least squares estimate $\hat{\delta}$ of δ and residuals $\hat{\mu}_i$.
3. Permute the residuals μ_i^π from the distribution of $\hat{\mu}_i$ $i = 1, \dots, n$.
4. Generate the new pseudo values of Y by using the residuals found in step 3 using (3.13) i.e. $Y_{i(\pi)} = \hat{\delta} Z_i + \mu_i^\pi$.
5. The new Y observations found in step 4 are then regressed against X and Z using equation (3.14) i.e. $E(Y_{i(\pi)}) = \beta_{1(\pi)}X_i + \beta_{2(\pi)}Z_i$ to obtain the estimate $\hat{\beta}_{1(\pi)}$ and $t_{j(\pi)} = \hat{\beta}_{1(\pi)} / se(\hat{\beta}_{1(\pi)})$ the t -statistic.
6. Step 2 to 5 is repeated many times to obtain the distribution of the permutation t -statistic.
7. The p -value of the two sided t -test is calculated as the proportion of the permutation t statistic greater than or equal in absolute value to the reference t -statistic from the full model with the original variables.

Instead of the t -statistic Anderson and Robinson (2001) used the squared correlation coefficient as a test statistic when examining the different permutation methods, which uses the relationship of X and Z in equation (3.16) i.e.

$$r_\pi^2 = \frac{\sum\{(Y_{i(\pi)} - \delta_\pi Z_i)(X_i - \gamma Z_i)\}^2}{\sum(Y_{i(\pi)} - \delta_\pi Z_i)^2 \sum(X_i - \gamma Z_i)^2} \quad (3.19)$$

Freedman and Lane (1983) specified that the sample size has to be large, there should be no outliers in the data and the degree of collinearity between X and Z should be low. Kennedy (1995) explores a method similar to Freedman and Lane's method but his paper does not provide details on how to apply this method. Anderson and Legendre (1999) outline the detailed procedure that can be followed when this method is applied as follows:

1. Regress Y on both X and Z as in equation (3.12) to obtain the least squares estimates $\hat{\beta}_1$ of β_1 and t -statistic t_0 .
2. Regress Y on Z as in equation (3.13) and obtain the least squares estimate $\hat{\delta}$ of δ and residuals $\hat{\mu}_i$.
3. Permute the residuals μ_i^π from the distribution of $\hat{\mu}_i, i = 1, \dots, n$.

These three steps are exactly the same as the Freedman and Lane method. The two methods differ from step 4.

4. Original X variable is regressed against Z using equation (3.14) to obtain the residuals $u_i, i = 1, \dots, n$. These residuals remain fixed across all different permutations.
5. Regress the permuted residuals in step three against residuals found in step 4 according to the model $E(\mu_i^\pi) = \beta_{1(\pi)}u_i$ to obtain the least squares estimate $\hat{\beta}_{1(\pi)}$ and t test statistic $t_{j(\pi)}$.
6. Repeat step 3 and step 5 many times and compute the t -statistic under permutation each time.
7. The p -value of the two sided t -test is calculated as the proportion of the permutation t -statistic greater than or equal in absolute value to the reference t -statistic from the full model with the original variables

Kennedy's method gives the same slope coefficient as Freedman and Lane's method but different values of the t -statistic under permutation (Anderson and Legendre, 1999). However the results obtained from both models are similar.

Permuting residuals under the full model

The second type of permutation of residuals does not assume that the null hypothesis is true, unlike the Kennedy and Freedman and Lane methods. In this method permutation is performed under the alternative hypothesis. It was first presented by te Braak (1992) as permutation analogue of a bootstrap significance test. Under the alternative hypothesis the exchangeability of residuals requirement is not required (te Braak, 1992). It differs from the bootstrapping residuals method in that in bootstrapping the residuals are sampled with replacement while in permutation methods they are sampled without replacement. The application of this method was detailed by Anderson and Legendre (1999) as follows:

1. Y is regressed on both X and Z as in equation (3.12) to obtain the estimates $\hat{\beta}_1, \hat{\beta}_2$ of β_1, β_2 respectively, t statistic for $\hat{\beta}_1$ and the residuals $\hat{\varepsilon}_{Y|XZ}$.
2. From the distribution of $\hat{\varepsilon}_{Y|XZ}$ draw the sample of residuals $\varepsilon_{Y|XZ}^\pi$ without replacement.
3. The new values of Y are calculated from equation (3.12) by replacing the unknown errors ε_i by the permuted residuals from step 2.
4. The new values of Y say Y_π are regressed on X and Z to find $\hat{\beta}_{1(\pi)}$, and its standard error used to compute the t statistic under permutation as follows

$$t_{j(\pi)} = \frac{\hat{\beta}_{1(\pi)} - \hat{\beta}_1}{se(\hat{\beta}_{1(\pi)})}$$

Alternatively in step 4 the permuted residuals can be regressed against X and Y to obtain the $\hat{\beta}_{1(\tau)}$ different from the one in step 4 and the t statistic computed as

$$t_{j(\tau)} = \frac{\hat{\beta}_{1(\tau)}}{se(\hat{\beta}_{1(\tau)})}$$

(Anderson and Legendre, 1999).

5. Steps 2 to 4 are repeated many times to generate the distribution of the t -statistic under permutation.

6. The *t*-statistic calculated for the original sample is placed on the distribution of the *t*-statistic from the permuted data. The *p* value is computed as the proportion of the *t*-statistic from the permuted samples greater than or equal in absolute value to the *t*-statistic from the original sample i.e. $p \text{ value} = P(t_{\tau} \geq t_0)$.

In all the permutation test methods the null hypothesis is rejected if the *p* value is less than the level of significance α .

3.3.2 Significance testing using bootstrap in multiple linear regression

Bootstrap hypothesis testing in multiple regression extends directly from the simple linear regression similar to the jackknife method. The two methods or procedures of bootstrapping namely bootstrapping pairs and bootstrapping residuals are applicable in a multiple regression setup. Equation (3.12) written in matrix form can be considered as in equation (3.7) and the interest is in testing the effect of *X* in explaining *Y* in the presence of the covariable *Z*. In the presence of heteroscedasticity in the errors, MacKinnon (2002) suggests using the bootstrapping pairs method of Freedman (1981) to handle this problem. In this method each *Y* observation sampled is sampled together with the *X* row to form

$$Y_i^* = X_i^{*\prime} \beta + \varepsilon_i^* \quad (3.20)$$

The row vector (X_i^*, Y_i^*) is of the exact form as (X_i, Y_i) the original row vector, for $i = 1, 2, \dots, n$. The method follows in a straightforward way from the simple linear regression case. A bootstrapping residuals method has more flexibility and allows for direct comparison with permutation tests (te Braak, 1992). The two choices of bootstrapping residuals (restricted and unrestricted bootstrapping) as outlined by Paparodotis and Politis, (2005) still hold in multiple regression as they do in the simple linear regression case.

3.4 Analysis of Variance (ANOVA)

In the previous sections permutation tests, jackknife and bootstrapping were considered for significance testing of a partial regression coefficient in the presence of other covariables. The same procedure can be used in the analysis of variance. Consider the one factor ANOVA model

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad i = 1, 2, \dots, k; j = 1, \dots, n \quad (3.21)$$

where y_{ij} is the (ij) observation, μ is the common mean across all factors, α_i is the i^{th} factor and ε_{ij} is the random error component. For the case of equal weightings the constraints on α_i is;

$$\sum_{i=1}^k \alpha_i = 0$$

Hence one of the k parameters α_i can be derived from other $k - 1$ parameters α_i in the following manner:

$$\alpha_k = -\alpha_1 - \alpha_2 - \dots - \alpha_{k-1}$$

(Kutner et al, 2005). The ANOVA is concerned with testing the equality of the k factors namely $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ vs H_1 : at least one of $\alpha_i \neq 0$. Equation (3.21) can be written in a regression equation by defining the $k - 1$ dummy variables resulting in a design matrix of the form;

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ \vdots & 1 & 0 \\ \vdots & 0 & 1 \\ \vdots & \vdots & \vdots \\ \vdots & 0 & 1 \\ \vdots & -1 & -1 \\ \vdots & \vdots & \vdots \\ 1 & -1 & -1 \end{bmatrix} \quad (3.22)$$

The column vector $\boldsymbol{\beta} = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \end{bmatrix}$

for an example of $k = 3$. By writing the design matrix in the form of equation (3.22) it allows for the ANOVA model (3.21) to be written in a regression form as follows:

$$Y = X\boldsymbol{\beta} + \varepsilon \quad (3.23)$$

Once the ANOVA model is represented as a regression equation (3.23) the null hypothesis $H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$ vs H_1 : at least one of $\beta_i \neq 0, i = 1, 2, \dots, p$.

The appropriate test statistic is the Fisher's F statistics given by $F = \frac{SSR/p}{SSE/(n-p-1)}$ for the linear model which includes the intercept. The same idea was described in the previous section of resampling appropriately for a relevant method and computing the resampled F statistic F^* each time. The p -value is thus computed as $p = \frac{(\text{no of } F^* \geq F)}{(\text{total no. of } F^*)}$ (Anderson, 2001).

3.5 Conclusion

In this chapter different resampling techniques were explored in some detail. In permutation methods there is a wide variety of choices of what to permute in multiple linear regression, namely the response variable Y , the explanatory variable X or residuals under the full model (te Braak, 1992) or under the reduced model (Freedman and Lane 1983). A general consensus in simple linear regression is permuting the response variable Y . In the jackknife method the two methods considered were *delete 1* and *delete d* methods. In bootstrapping there are two choices on whether to bootstrap the residuals or bootstrap the observations.

Chapter 4

Confidence Intervals and Percentile Based Intervals

4.1 Introduction

This chapter will explore the application of bootstrap and jackknife $100(1 - \alpha)$ confidence intervals as well as the $\alpha/2$ and $(1 - \alpha/2)$ percentile intervals. Parametric confidence intervals can be used whenever the assumption of normality is satisfied while the percentile interval does not require the normality assumption. It involves construction of an empirical distribution of least squares estimates of the regression coefficient computed from bootstrap/jackknife samples to obtain the upper $1 - \alpha/2$ and lower $\alpha/2$ limits in the case of two sided percentile interval.

4.2 Confidence Interval by Normal Approach

Another form of significance testing for regression coefficients is through the computation of confidence intervals and the null hypothesis is rejected whenever the value under the null hypothesis does not fall within the interval limits.

4.2.1 Bootstrap confidence interval by Normal Approach

It is in the area of confidence intervals where the bootstrapping technique proved to be a very useful tool in statistics. Bootstrapping confidence intervals has been studied in great details hence there are many ways to construct bootstrap confidence interval than there are to perform bootstrap tests (MacKinnon, 2002). Among many bootstrapping confidence interval methods the two which are constructed in a very similar way as the standard normal confidence interval are bootstrap t interval and interval based on bootstrap standard errors. The latter is referred to as bootstrap

confidence interval by normal approach by Sahinler and Topuz (2007). Bootstrap t interval uses the bootstrap test statistic

$$t_j^* = \frac{\hat{\beta}_j^* - \hat{\beta}}{se(\hat{\beta}_j^*)}, \quad j = 1, 2, \dots, B \quad (4.1)$$

Where $\hat{\beta}_j^*$ is estimated from each bootstrap sample together with its standard error $se(\hat{\beta}_j^*)$, $\hat{\beta}$ the estimate of β and $se(\hat{\beta})$ the standard error of $\hat{\beta}$ are estimated from the original sample. If $se(\hat{\beta}_j^*)$ is replaced by $se(\hat{\beta})$ in computing test statistic t_j^* , the resulting test statistic t_j^* is not asymptotically pivotal (MacKinnon, 2002). The bootstrap t interval is then given by

$$\hat{\beta} - t_{1-\alpha/2}^* se(\hat{\beta}); \hat{\beta} - t_{\alpha/2}^* se(\hat{\beta}) \quad (4.2)$$

where $t_{\alpha/2}^*$ and $t_{1-\alpha/2}^*$ are the $(\alpha/2)\%$ and $(1 - \alpha/2)\%$ quantiles of the ordered t_j^* . This method works well if the test statistic on which it is based is approximately pivotal (MacKinnon, 2002). Bootstrapping t can lead to intervals that are skewed since it may be asymmetric about zero. Efron and Tibshirani (1993) stated that this method cannot be trusted for setting up confidence interval for a correlation coefficient.

The other widely used method is constructing interval based on bootstrapped standard errors (MacKinnon, 2002) and this is the method which will be detailed further following the work that was published by Sahinler and Topuz (2007) which they referred to as bootstrapping confidence interval by normal approach. This approach will be outlined below.

Sahinler and Topuz (2007) described an algorithm for bootstrapping residuals, estimating the regression coefficients, computing the standard error of the regression coefficients and finally computing the confidence interval by normal approach.

1. Regress \mathbf{Y} on \mathbf{X} as in equation (3.23) to obtain the least squares estimates $\hat{\boldsymbol{\beta}}$.
2. Estimate the vector of residuals as $\hat{\boldsymbol{\varepsilon}} = \mathbf{Y} - \hat{\mathbf{Y}}$.
3. Randomly select the random sample of residuals ε_i^* $i = 1, 2, \dots, n$ with replacement from the vector of residuals $\hat{\boldsymbol{\varepsilon}}$.
4. Generate the bootstrap \mathbf{Y}^* observations using the bootstrapped sample of residuals from residual vector $\hat{\boldsymbol{\varepsilon}}$ from step 3 above and ordinary least squares estimate obtained from step 1 i.e $\mathbf{Y}^* = \mathbf{X}\hat{\boldsymbol{\beta}} + \boldsymbol{\varepsilon}^*$.
5. Regress the bootstrapped \mathbf{Y} on fixed \mathbf{X} i.e $E(\mathbf{Y}^*|\mathbf{X}) = \mathbf{X}\boldsymbol{\beta}$ to obtain the least squares estimator $\hat{\boldsymbol{\beta}}^{br}$ and its standard error $se(\hat{\boldsymbol{\beta}}^{br})$ from the r^{th} bootstrap sample.
6. Repeat step 3 to 6 for $r = 1, 2, \dots, B$.
7. Generate the empirical distribution of $\hat{\boldsymbol{\beta}}^{br}$ say $F(\hat{\boldsymbol{\beta}}^{br})$.

A similar approach can be followed for the bootstrap based on resampling observations rather than residuals which is generally applicable whenever the regressors are as random as the dependent variable. The algorithm is as follows:

1. Draw a sample of size n from the elements of the \mathbf{Y} vector and corresponding rows of the \mathbf{X} matrix.
2. Regress the resulting \mathbf{Y} vector on the \mathbf{X} matrix to obtain ordinary least squares estimates of regression coefficients.
3. Repeat step 1 and step 2 for $r = 1, 2, \dots, B$.
4. Generate the empirical distribution of $\hat{\boldsymbol{\beta}}^{br}$ say $F(\hat{\boldsymbol{\beta}}^{br})$.

For both techniques (bootstrapping residuals and bootstrapping observations) $\hat{\beta}^b$ the bootstrap estimator of β is then computed as the mean of the individual ordinary least squares regression coefficient estimates computed in each of the B bootstrap samples considered i.e.

$$\hat{\beta}^b = \frac{1}{B} \sum_{r=1}^B \hat{\beta}^{br} \quad (4.3)$$

and its standard error is given by

$$se(\hat{\beta}^b) = \left(\frac{1}{B-1} \sum_{r=1}^B (\hat{\beta}^{br} - \hat{\beta}^b)(\hat{\beta}^{br} - \hat{\beta}^b)' \right)^{1/2} \quad (4.4)$$

The two sided $100(1 - \alpha)\%$ bootstrap confidence by normal approach limits computed as

$$\hat{\beta}^b - t_{(n-p, 1-\frac{\alpha}{2})} se(\hat{\beta}^b) < \beta < \hat{\beta}^b + t_{(n-p, 1-\frac{\alpha}{2})} se(\hat{\beta}^b) \quad (4.5)$$

Due to the *t distribution* being assumed, the $\alpha/2$ and $(1 - \alpha/2)$ quantiles of the distribution are symmetrical around the origin. The upper limit depends on the lower critical value and vice versa i.e. $t_{\alpha/2} = -t_{1-\alpha/2}$. The equation 4.3 can therefore be rewritten as

$$\hat{\beta}^b - t_{(n-p, 1-\frac{\alpha}{2})} se(\hat{\beta}^b) < \beta < \hat{\beta}^b - t_{(n-p, \frac{\alpha}{2})} se(\hat{\beta}^b) \quad (4.6)$$

A null hypothesis of the regression coefficient is rejected whenever the value of the coefficient under the null hypothesis falls outside of the confidence limits.

4.2.2 Jackknife confidence interval by Normal Approach

When a jackknife technique is applied in regression, a general way of testing a hypothesis is through a confidence interval. It can be done by following the algorithm given by Sahinler and Topuz (2007), which is as follows: Given the regression model in equation (3.23).

1. Draw the sample (X_i, Y_i) of size n and number from $i = 1, 2, \dots, n$.
2. Omit one row of the pair (X_i, Y_i) starting with the first pair.
3. Compute the least squares estimate from the $(n - 1)$ elements and label the estimate say $\hat{\beta}^{Ji}$.
4. Repeat step 2 and 3 n times by deleting 1 observation at the time.
5. Construct the distribution of jackknife estimates $\hat{\beta}^{J1}, \hat{\beta}^{J2}, \dots, \hat{\beta}^{Jn}$.
6. Compute the jackknife regression coefficient estimates as the mean of the $\hat{\beta}^{Ji}$ s i.e.

$$\hat{\beta}^J = \frac{1}{n} \sum_{i=1}^n \hat{\beta}^{Ji} \quad (4.7)$$

with its standard error given by

$$se(\hat{\beta}^J) = \left(\frac{n-1}{n} \sum_{i=1}^n (\hat{\beta}^{Ji} - \hat{\beta}^J)(\hat{\beta}^{Ji} - \hat{\beta}^J)' \right)^{1/2} \quad (4.8)$$

(Friedl and Stampfer, 2002).

The jackknife bias is measured by

$$bias(\hat{\beta}^J) = (n - 1)(\hat{\beta}^J - \hat{\beta}) \quad (4.9)$$

The jackknife confidence interval is given by

$$\hat{\beta}^J - t_{(n-p, 1-\frac{\alpha}{2})} se(\hat{\beta}^J) < \beta < \hat{\beta}^J + t_{(n-p, 1-\frac{\alpha}{2})} se(\hat{\beta}^J) \quad (4.10)$$

Friedl and Stampfer (2002) outlined two deficiencies of the jackknife technique, namely:

1. The jackknife variance estimator is not consistent for the sample median because the sample median cannot be approximated linearly.
2. Jackknife does not directly provide distribution estimators.

However, Friedl and Stampfer (2002) claim that these deficiencies can be overcome by the *delete d* method. The delete *d* method can be applied using a similar algorithm except that *d* rows/observations are deleted at a time instead of deleting 1. The resulting number of samples is $S = \binom{n}{d}$ of size $(n - d)$ each. The variance of the *delete d* method is given by

$$se(\hat{\beta}^{J-d}) = \left(\frac{n-d}{Sd} \sum_{i=1}^S (\hat{\beta}^{J_i} - \hat{\beta}^J)(\hat{\beta}^{J_i} - \hat{\beta}^J)' \right)^{1/2} \quad (4.11)$$

The fraction *d* and *n* has to be large enough for (4.11) to be consistent (Friedl and Stampfer, 2002). It can be noted that for $d = 1$, equation (4.11) is equivalent to (4.8) i.e. $d = 1$ is a special case of delete *d* method.

4.3 Percentile Based Interval

In real life problems it is often not appropriate to assume a symmetric distribution of the data, especially when there is a scarcity of data. In this instance a $100(1 - \alpha)\%$ confidence interval can be very misleading by overstating or understating the length of the confidence limits due to the use of the *t* distribution when the data are not

symmetric around any point. Here $\alpha/2$ and $(1 - \alpha/2)$ percentiles intervals give more reliable estimates because the percentile intervals are derived from the distribution itself regardless of the shape. If the distribution is skewed to one side the percentile interval will produce the smaller length of the confidence interval that is concentrated to the side where there are more data points (Efron and Tibshirani ,1993).

4.3.1 Bootstrap Percentile Based interval

The percentile interval makes use of the distribution of bootstrap regression coefficients generated from each of B bootstrap samples. For the regression model considered in Section 4.2 above the observed values of the bootstrap regression coefficient $\hat{\beta}^{br}$ are ordered in an ascending order. The $\hat{\beta}^{br}(lower)$ is the lower $\left(\frac{\alpha}{2}\right) B$ bootstrap estimate of the regression coefficient and $\hat{\beta}^{br}(upper)$ is the $\left(1 - \frac{\alpha}{2}\right) B$ bootstrap estimate of the coefficient. The percentile based interval therefore becomes

$$\hat{\beta}^{br}(lower) < \beta < \hat{\beta}^{br}(upper). \quad (4.12)$$

The null hypothesis is rejected if the value of the regression coefficient does not fall within the limits of the percentile interval. The percentile based bootstrap can be improved by what Efron and Tibshirani (1993) refer to as bias-corrected and accelerated (BC_a) and approximate bootstrap confidence intervals (ABC). These methods will not be described further here because a similar extension of jackknife is not possible (Efron, 2003).

4.3.2 Jackknife Percentile Based interval

The application of confidence intervals in jackknife does not produce reliable confidence limits (Efron and Tibshirani, 1993). It is for this reason that it is more appropriate to compute percentile intervals of jackknife. The formulation of percentile intervals is similar to the bootstrap interval above. For both the *delete 1* and *delete d*

methods, the computed estimates of the regression coefficient are ranked from the smallest to the largest. For the *delete 1* method $\hat{\beta}^J(lower)$ is the $\left(\frac{\alpha}{2}\right)^{th}$ jackknife estimate of regression coefficient and $\hat{\beta}^J(upper)$ is the $\left(1 - \frac{\alpha}{2}\right)^{th}$ of jackknife estimate of regression coefficient. For *delete d* method the lower and upper limits indices are

$$lower = \left(\frac{\alpha}{2}\right)S \text{ and } upper = \left(1 - \frac{\alpha}{2}\right)S \quad (4.13)$$

Therefore the percentile interval is expressed as

$$\hat{\beta}^J(lower) < \beta < \hat{\beta}^J(upper) \quad (4.14)$$

which can also be written as

$$CDFJ\left(\frac{\alpha}{2}\right); CDFJ\left(1 - \frac{\alpha}{2}\right)$$

where CDFJ () is the empirical distribution function based on $\hat{\beta}^J$ (Sahinler and Topuz, 2007 and Wu, 1986).

4.4 Conclusion

Sections 4.2.1 and 4.2.2 showed that both techniques (bootstrap and jackknife) can be used to first compute the estimate of the regression coefficient and the standard errors which are used in the derivation of the confidence intervals. Completely non-parametric percentile intervals are a viable method for both jackknife and bootstrap.

Chapter 5

Research Methodology and Design

5.1 Introduction

This chapter describes the methodology followed in a simulation study of the bootstrap, jackknife and permutation test for significance testing in regression analysis which was discussed in Chapter 3 and computation of confidence intervals and percentile intervals as discussed in Chapter 4. This chapter will be organised as follows: the introduction, problem statement, research procedure and conclusion.

5.2 Problem Statement

Bootstrap, jackknife and permutation tests are explored as alternative methods to parametric methods, especially when the errors are not normally distributed. Parametric, bootstrap and permutation significance testing for regression coefficient are compared in terms of power and probability of type 1 error. In the second part of the study parametric, bootstrap and jackknife confidence intervals are compared in terms of the probability of including the correct value of β_1 when the correct value of $\beta_1 = 0$ and $\beta_1 = 2$ (coverage) and the average length of the interval in a Monte Carlo simulation study.

5.3 Research Procedure

In this section the performance of parametric, bootstrap and permutation tests are compared in testing the significance of partial regression coefficient i.e. $H_0: \beta_1 = 0$ under the model

$$Y = X\beta + \varepsilon \quad (5.1)$$

where X is $n \times (p + 1)$ with the first column containing 1's.

The Monte Carlo simulation study for significance testing of a single partial regression coefficient was investigated under different conditions to ascertain the robustness of the techniques under investigation. (1) The effect of a sample size was considered ranging from a very small sample size to a relatively large sample size, (2) number of covariates in the model, (3) the distribution of the added random error (4) and the effect of collinearity among the independent variables in the model when $p > 1$. The exact choices of the factors outlined above are given below namely:

1. The number of independent variables was $p = (1, 2, 5)$ resulting in the following regression equations:

$$Y_1 = \beta_0 + \beta_1 X_1 + \varepsilon$$

$$Y_2 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

$$Y_5 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon$$

The interest is to establish the impact that that increase in the number of covariables has on the probability of type 1 error and power of a test.

2. The sample sizes $n = (5p, 10p, 25p, 50p)$.

The aim is to ensure that we can study the behaviour of probability of type 1 error and power of a test for a wider range of sample size. This will help to ascertain how different methods behave from the smallest sample size of 5 to the largest sample size of 250.

3. Error distributions considered are $\varepsilon \sim$ (standard normal, uniform (0, 1), lognormal).

The aim is to establish the influence of error distribution has to the probability of type 1 error and power of a test for different methods to see what methods are robust enough when errors are deviating from normality.

4. The collinearity structure of the independent variables was correlated with $\rho = (0, 0.5, 0.9)$.

The aim is to establish what happens to the probability of type 1 error and power of a test when there are inherent shortcomings in the data.

Below are the specific techniques assessed under the factors outlined above.

1. Bootstrapping the observations.
2. Bootstrapping the residuals.
3. Permuting the observations.
4. Freedman and Lane permutation method.

The jackknife method was considered but it did not yield positive results and there was a shortage of literature where the jackknife method is used for significance testing. It was therefore assessed through confidence intervals computation.

SAS statistical software was used in the analysis. However there were no readily available procedures in SAS that could perform these techniques. SAS macros have been developed which are useful in testing the significance of differences in the statistic of choice such as mean between two populations. These macros are not for the application of bootstrap and permutation tests in the regression model. The SAS Iterated Matrix Language procedure (Proc IML) was used for programming. The \mathbf{W} matrix used to construct the \mathbf{X} matrix was generated independently from Uniform (0,1) distribution using the randgen call routine with randseed initialised to ensure that the sampling process can be replicated to obtain the same sample. The randgen call routine “uses the Marsenne-Twister random number generator developed by Matsumoto and Nishimura (1998)”. Matrix \mathbf{X} was constructed as follows

$$\mathbf{X} = \mathbf{1W} = \begin{pmatrix} 1 & w_{11} & \cdots & w_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & w_{n1} & \cdots & w_{np} \end{pmatrix} \quad (5.2)$$

The correlation structure among the independent variables was introduced through using the square root of the correlation matrix \mathbf{R} as follows

$$\mathbf{R}^{1/2} = \begin{bmatrix} 1 & \cdots & \rho \\ \vdots & \ddots & \vdots \\ \rho & \cdots & 1 \end{bmatrix}^{1/2} \quad (5.3)$$

where matrix \mathbf{R} is a $(p \times p)$ matrix. Matrix \mathbf{X}_R ; the matrix with correlated covariables was constructed as follows

$$\mathbf{X}_R = \mathbf{1}\mathbf{W}_R = \mathbf{1}\mathbf{W}\mathbf{R}^{1/2}. \quad (5.4)$$

Under multicollinearity the regression model is

$$\mathbf{Y} = \mathbf{X}_R\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \quad (5.5)$$

The added random error distributions were generated using SAS randgen call routine with a fixed random seed. Initially the number of Monte Carlo simulations were chosen to be 1000 and this resulted in empirical estimates of probability of type 1 error as large as 20% for very small sample sizes. This meant that the number of simulations had to be increased. 5000 Monte Carlo simulations were generated with the number of bootstrap samples equal to 999 as it is desirable to have $\alpha(B + 1)$ that is an integer (MacKinnon, 2002). MacKinnon (2002) showed that the loss of power is generally small when $B \geq 999$. The number of permutations was chosen to be equal to the number of bootstrap samples for comparison purposes. The empirical probability of type 1 errors were calculated from the 5000 Monte Carlo simulated data where $\beta_1 = 0$. The level of significance was chosen to be 5%. Under the null hypothesis the data was generated for $p = 1, 2, 5$ from the equations

$$Y_1 = \beta_0 + \varepsilon$$

$$Y_2 = \beta_0 + \beta_2 X_2 + \varepsilon$$

$$Y_5 = \beta_0 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon$$

For each simulation the reference *t*-statistic was computed from the originally simulated data using equation (3.2) using the regression module on proc iml. This reference *t*-statistic was passed to all the modules for bootstrap and permutation methods. Within the bootstrap and permutation modules the *t*-statistics were computed 999 times to obtain the empirical distribution of the *t*-statistic. The *p*-value of the two sided *t*-test was then calculated as the proportion of the permutation/bootstrap *t*-statistic greater than or equal in absolute value to the reference *t*-statistic. This process was repeated 5000 times resulting in 5000 *p*-values for each method. The probability of type 1 error was then computed as the proportion of times that the *p*-value is less than or equal to the level of significance (0.05). It was expected that the number of *p*-values which are significant would be around 250.

The power of the test was examined under the same conditions with $\beta_1 = 2$ (null hypothesis false) was chosen as an appropriate value to have the null hypothesis rejected quite frequently but not all the time. Under the alternative the data was generated for $p = 1,2,5$ from the equations

$$Y_1 = \beta_0 + 2X_1 + \varepsilon$$

$$Y_2 = \beta_0 + 2X_1 + \beta_2 X_2 + \varepsilon$$

$$Y_5 = \beta_0 + 2X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon$$

The same process used in computing the probability of type 1 error was followed for the computation of power.

The two-sided confidence and percentile intervals were computed under the null hypothesis i.e. $\beta_1 = 0$ and also when $\beta_1 = 2$. Due to the program running much longer for confidence and percentile intervals the number of Monte Carlo simulations was kept to 1000. The confidence interval data generation involved the two sets of equations one with $\beta_1 = 0$ and the other with $\beta_1 = 2$ which are outlined above. Note that it was not necessary to work with two sets of β_1 , one value of β_1 would suffice. Also note that it was not necessary to work with these specific values of β_1 . The difference between the probability type 1 error/power and confidence interval is that for the probability of type 1 error and power the null hypothesis is assumed to be

true within the module while on confidence interval the estimation of the regression coefficients are taking place without assuming the null hypothesis. The coverage of the confidence interval was performed by computing the number of times that the correct value of β_1 was within the confidence limits. The average length of a confidence interval was computed by first computing the length of interval (upper confidence limit-lower confidence limit) within each module. This process was repeated 1000 times producing 1000 values of length of an interval. An average was then computed resulting in the average length of an interval.

5.4 Conclusion

This chapter explained the methodology employed in the comparative study of jackknife, bootstrap and permutation tests as alternatives to the parametric methods. SAS proc IML using predefined modules for the implementation of these techniques was employed as a statistical tool in the absence of readily available SAS procedures. Prog Gplot was utilised for probability of type 1 error graph plots and power of a test plots.

Chapter 6

Simulation Results

6.1 Introduction

In this chapter results obtained during analysis using the methodology outlined in Chapter 5 of this document will be presented and their significance discussed. The tables of probability of type 1 error and power used to generate the graphs can be found on the appendix. The chapter will be organised in the following way: (1) probability of type 1 error investigation, (2) power of the tests and (3) confidence interval comparisons and (4) a conclusion.

6.2 Probability of type 1 error

This section will be split into three parts: Probability of type 1 error for the null hypothesis $H_0: \beta_1 = 0$ vs $H_1: \beta_1 \neq 0$ will be presented for cases where:

1. Explanatory variables are independent from each other.
2. Explanatory variables are correlated to each other.
3. Effect of increasing correlation within explanatory variables.

The following models are considered

$$Y_1 = \beta_0 + \beta_1 X_1 + \varepsilon$$

$$Y_2 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

$$Y_5 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon$$

The results for each of the probability of type 1 error graphical representation are presented twice to portray two findings. The probability of type 1 error is stable around the chosen significance level 0.05 although the type one error is not exactly the same for all the methods. The former is made more visible through the expansion of the y axis to cover from zero to 1 and the latter is achieved by allowing SAS to automatically select the y axis scale. The tables showing the probability of

type 1 error and power of a test results are contained in Appendix A and B respectively. The bootstrapping observation method produced singular matrices for $\hat{\beta}_1$ computations when the sample size is $5p$ for cases where $p=1$ and $p=2$. This is because in sampling 999 times with replacement in a “population” of say 5, each pair has 20% chance of being selected, this result in a same pair being repeated in at least one of the samples. This results in problem when the matrices are inverted hence no results are displayed for this method.

6.2.1 Probability of type 1 error for independent explanatory variables

Probability of type 1 error results for $p=1$ Normal and $\rho=0.0$

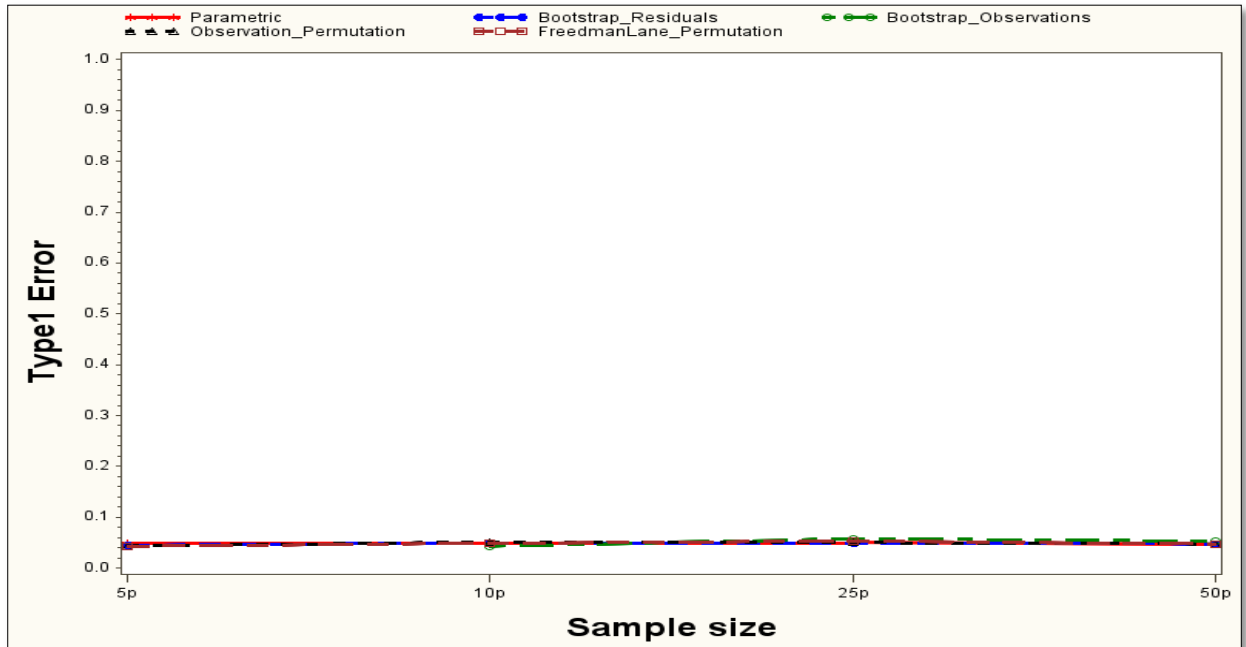


Figure 1: Type 1 error for $p = 1$ and $\epsilon \sim \text{Normal}(0, 1)$ on wider vertical axis

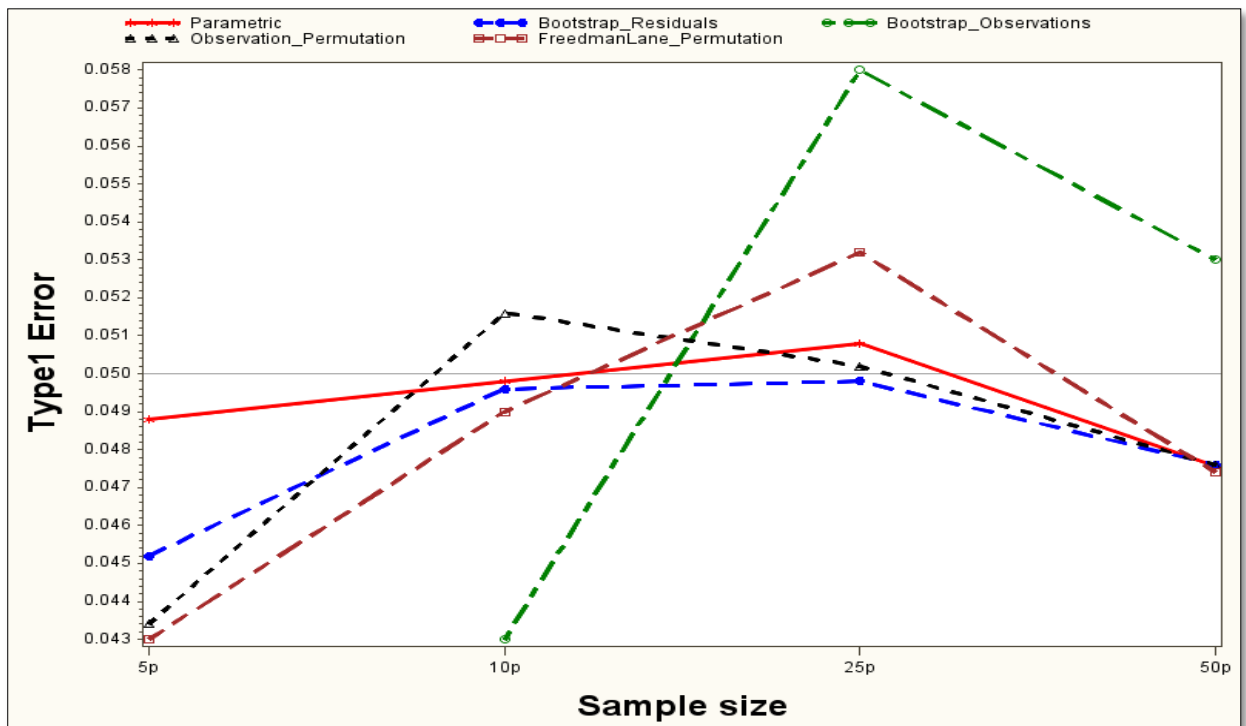


Figure 2: Type 1 error for $p = 1$ and $\epsilon \sim \text{Normal}(0, 1)$ on narrow vertical axis

Probability of type 1 error results for $p=1$ Uniform and $\rho=0.0$

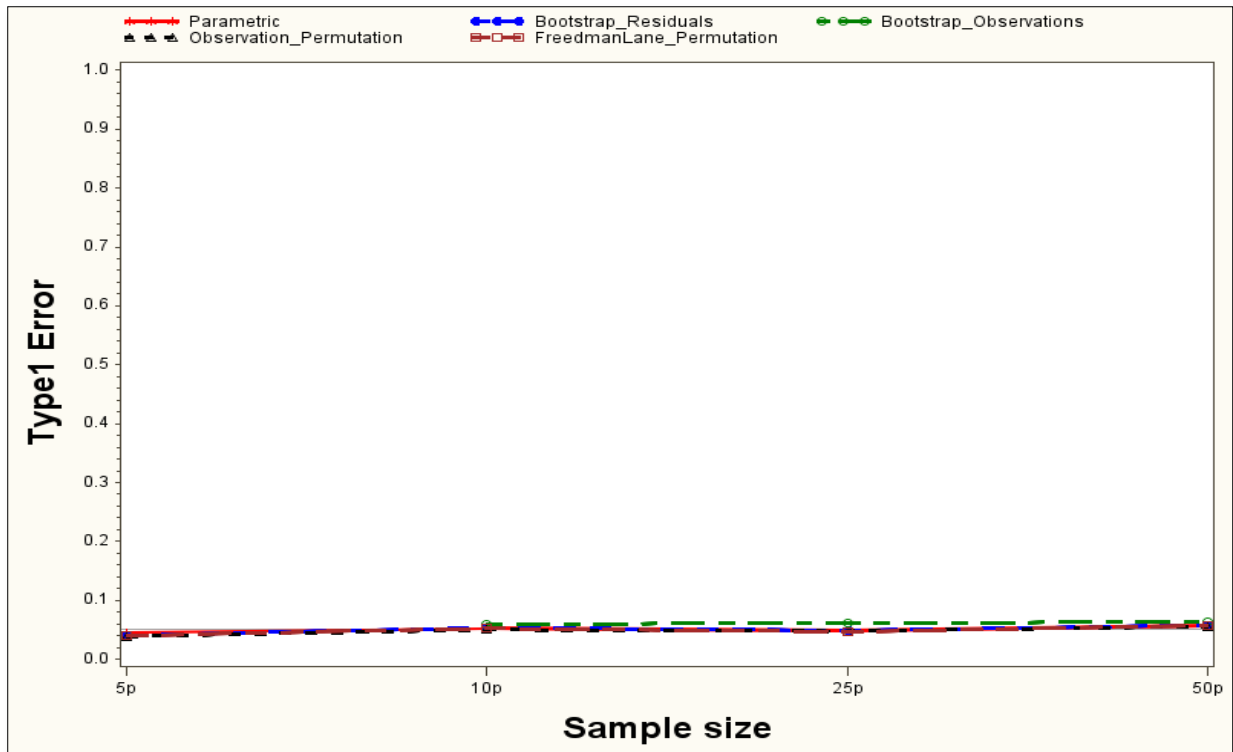


Figure 3: Type 1 error for $p = 1$ and $\varepsilon \sim \text{Uniform}(0, 1)$ on wider vertical axis

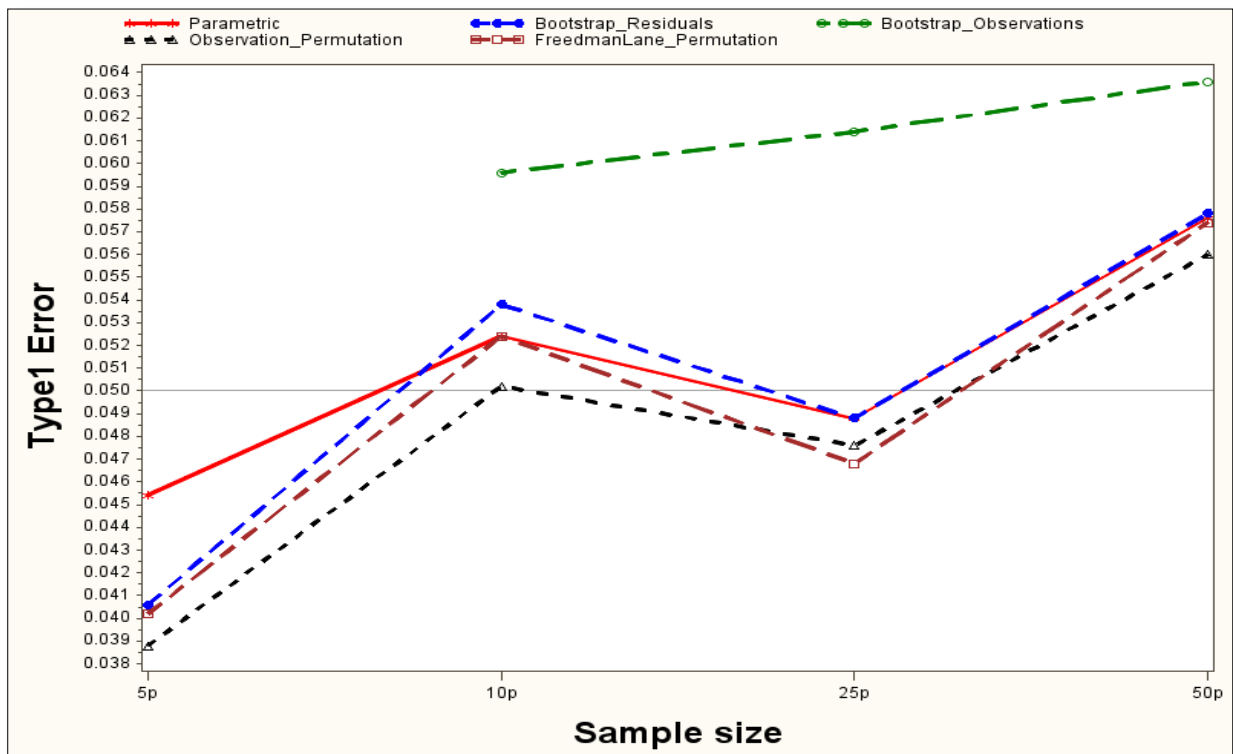


Figure 4: Type 1 error for $p = 1$ and $\varepsilon \sim \text{Uniform}(0, 1)$ on narrow vertical axis

Probability of type 1 error results for $p=1$ LOGN and $\rho=0.0$

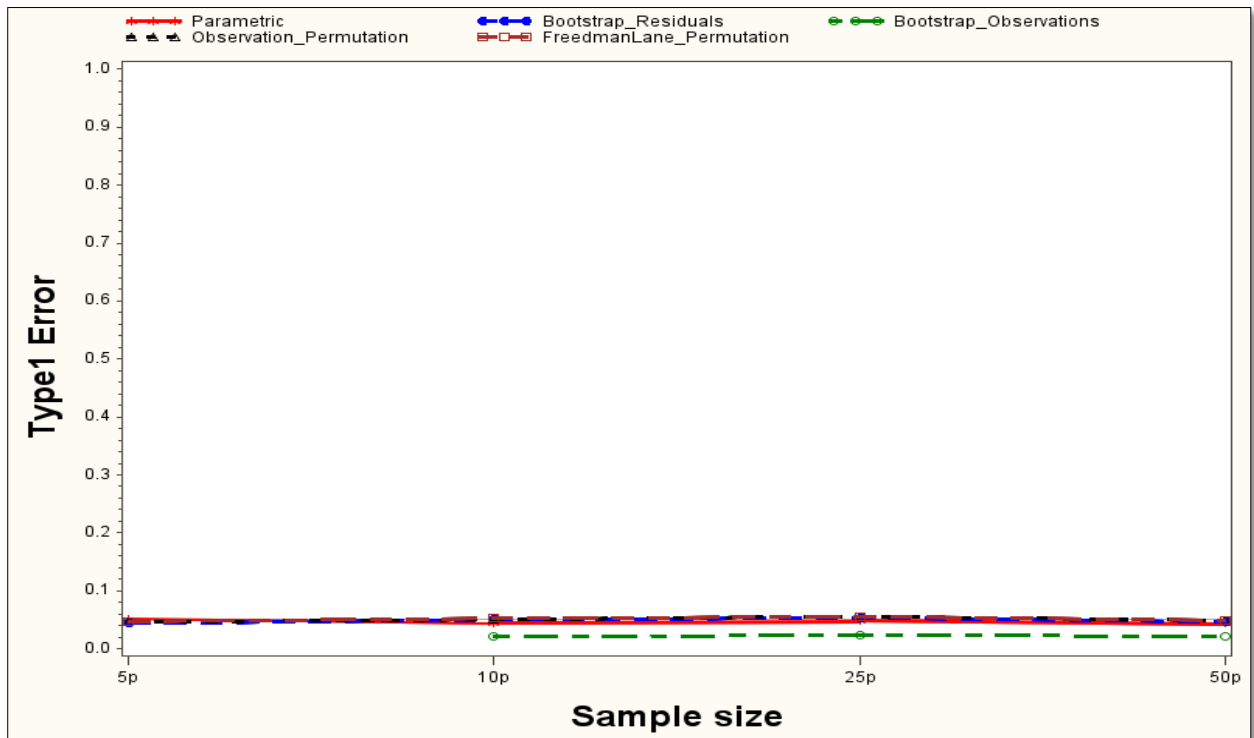


Figure 5: Type 1 error for $p = 1$ and $\epsilon \sim \text{LOGN}(0, 1)$ on wider vertical axis

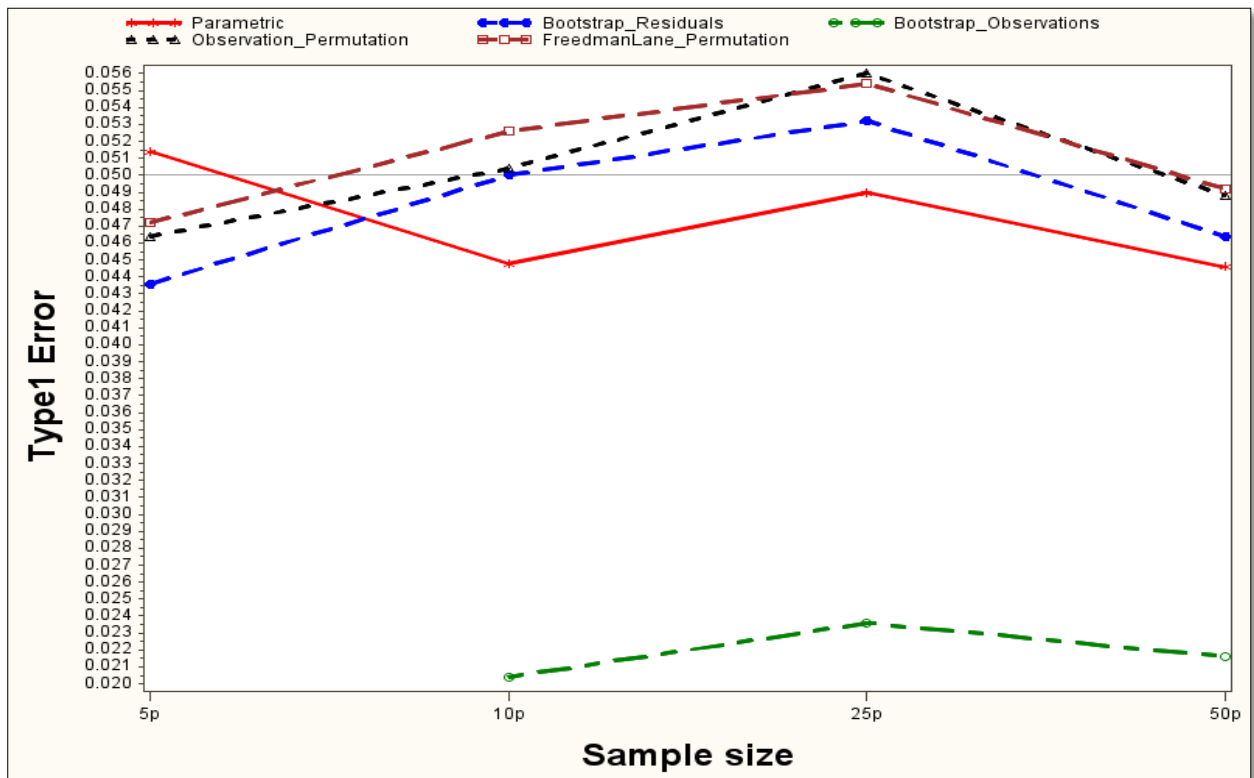


Figure 6: Type 1 error for $p = 1$ and $\epsilon \sim \text{LOGN}(0, 1)$ on narrow vertical axis

Probability of type 1 error results for $p=2$ Normal and $\rho=0.0$

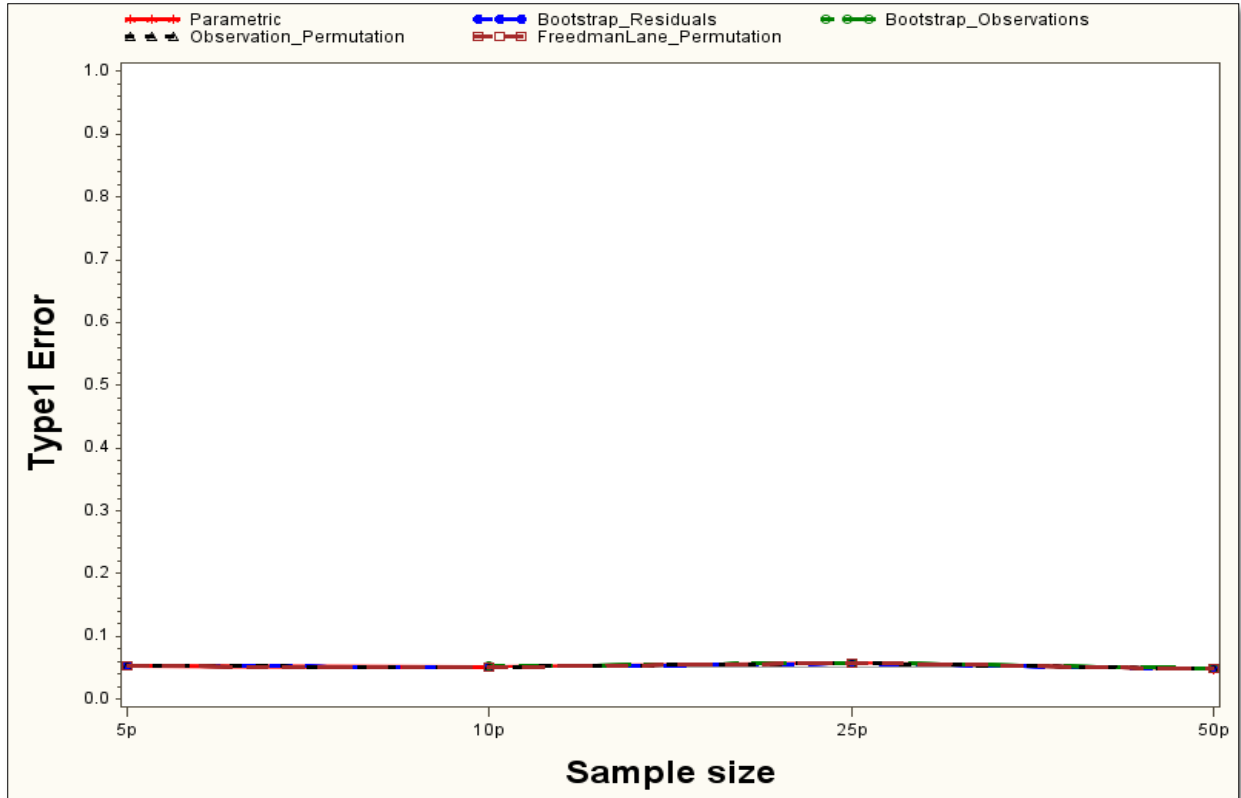


Figure 7: Type 1 error for $p = 2, \varepsilon \sim Normal(0, 1)$ and $\rho = 0$ on wider vertical axis

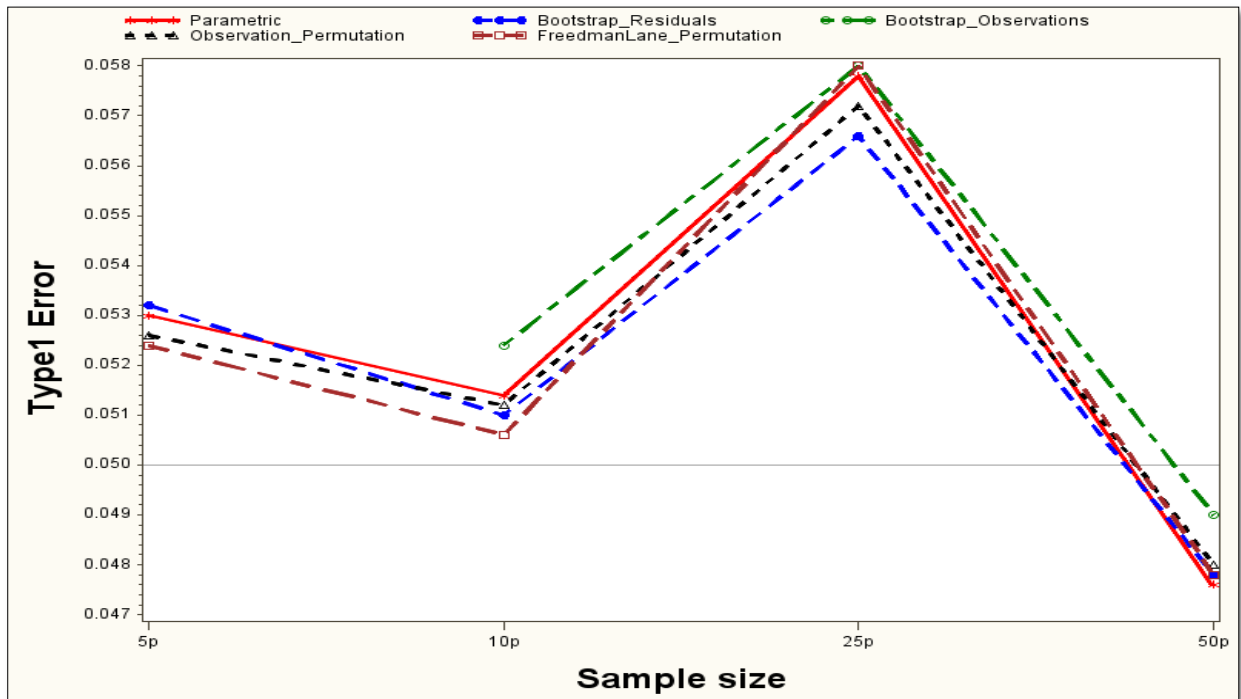


Figure 8: Type 1 error for $p = 2, \varepsilon \sim Normal(0, 1)$ and $\rho = 0$ on narrow vertical axis

Probability of type 1 error results for $p=2$ Uniform and $\rho=0.0$

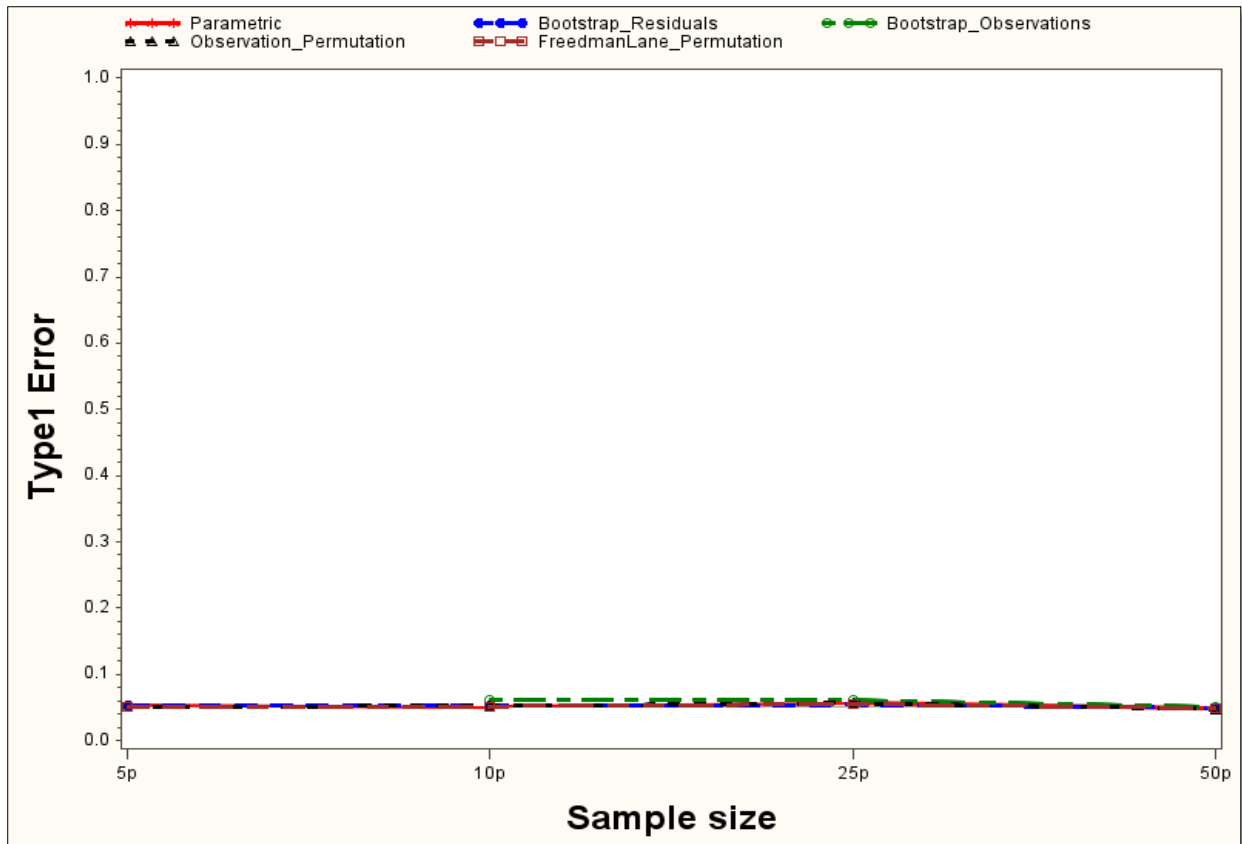


Figure 9: Type 1 error for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0$ on wider vertical axis

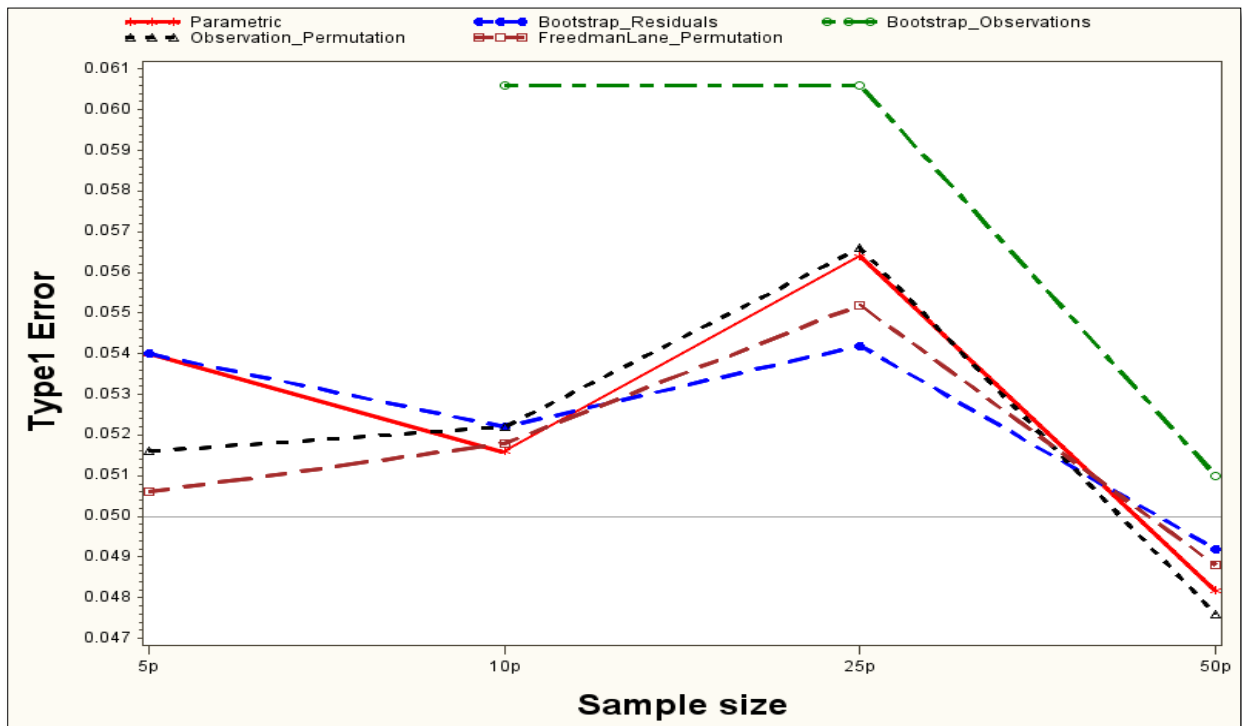


Figure 10: Type 1 error for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0$ on narrow vertical axis

Probability of type 1 error results for $p=2$ LOGN and $\rho=0.0$

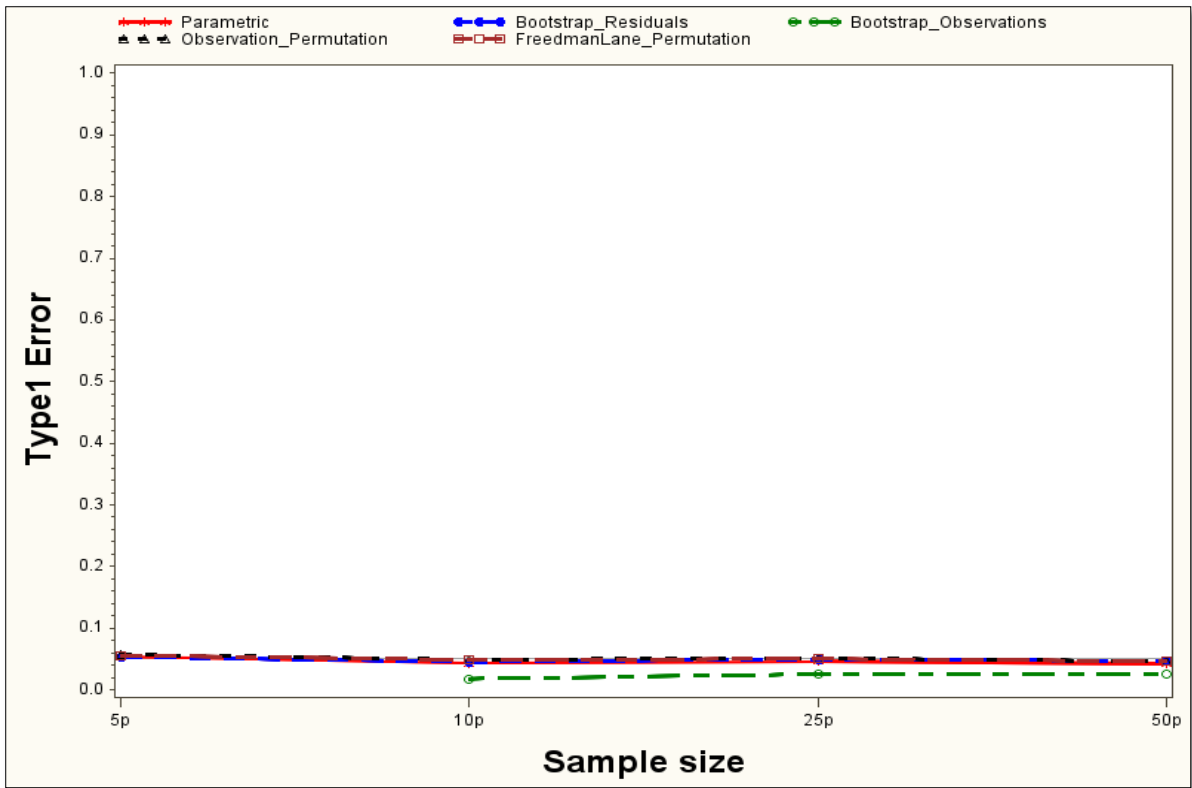


Figure 11: Type 1 error for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on wider vertical axis

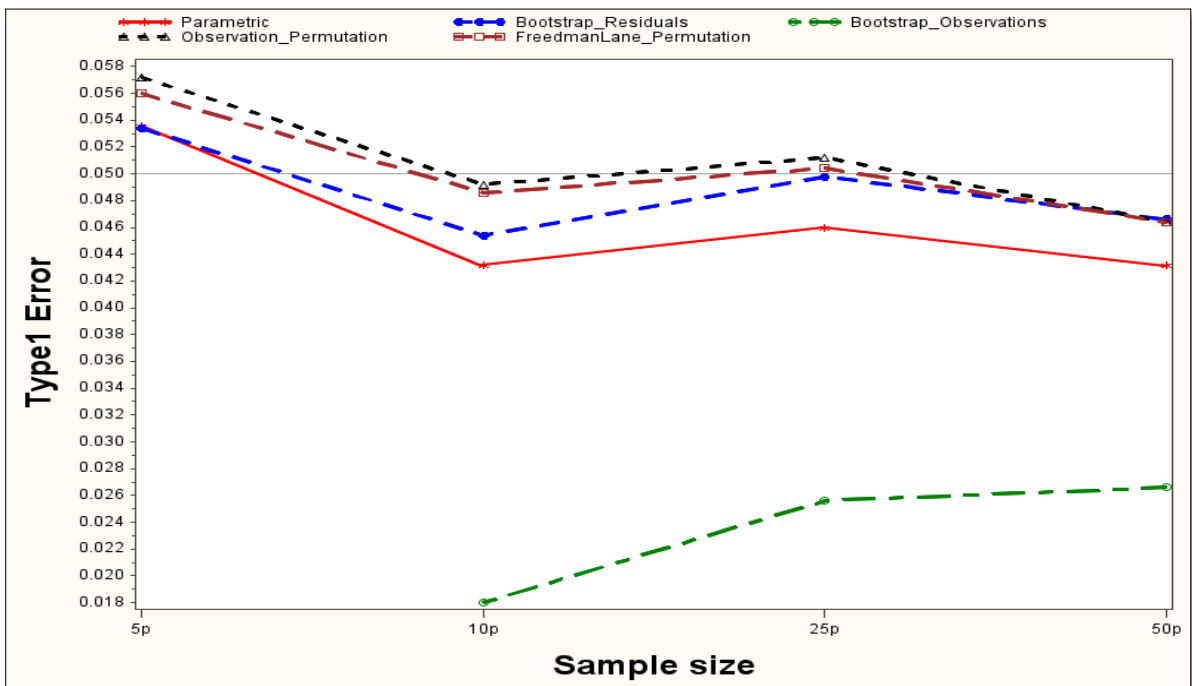


Figure 12: Type 1 error for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on narrow vertical axis

Probability of type 1 error results for $p=5$ Normal and $\rho=0.0$

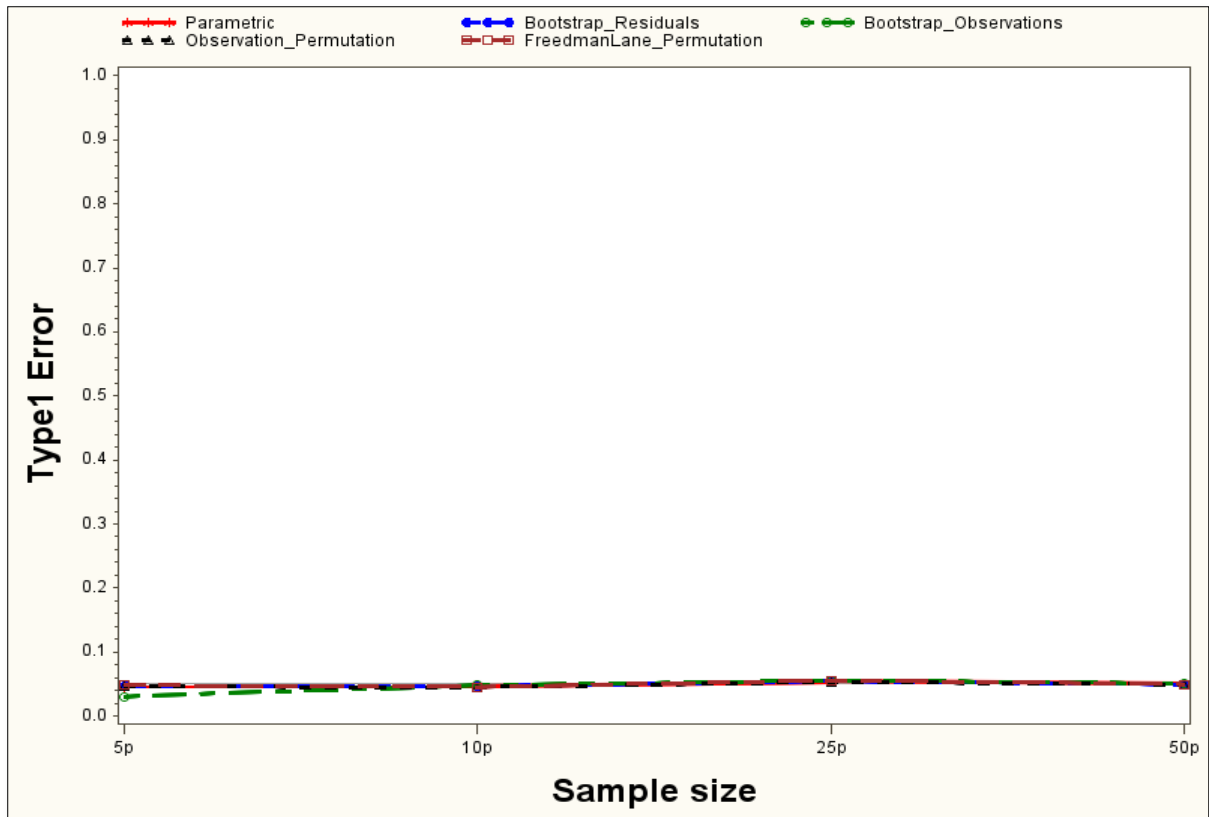


Figure 13: Type 1 error for $p = 5$, $\varepsilon \sim Normal(0, 1)$ and $\rho = 0$ on wider vertical axis

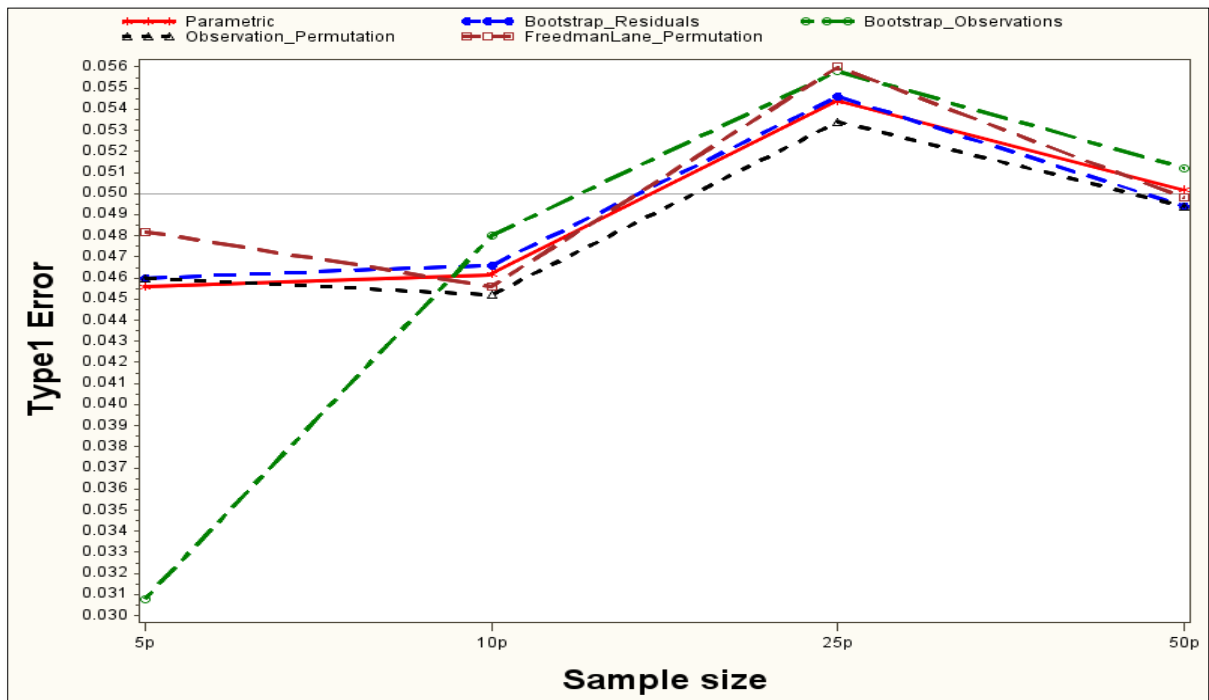


Figure 14: Type 1 error for $p = 5$, $\varepsilon \sim Normal(0, 1)$ and $\rho = 0$ on narrow vertical axis

Probability of type 1 error results for $p=5$ Uniform and $\rho=0.0$

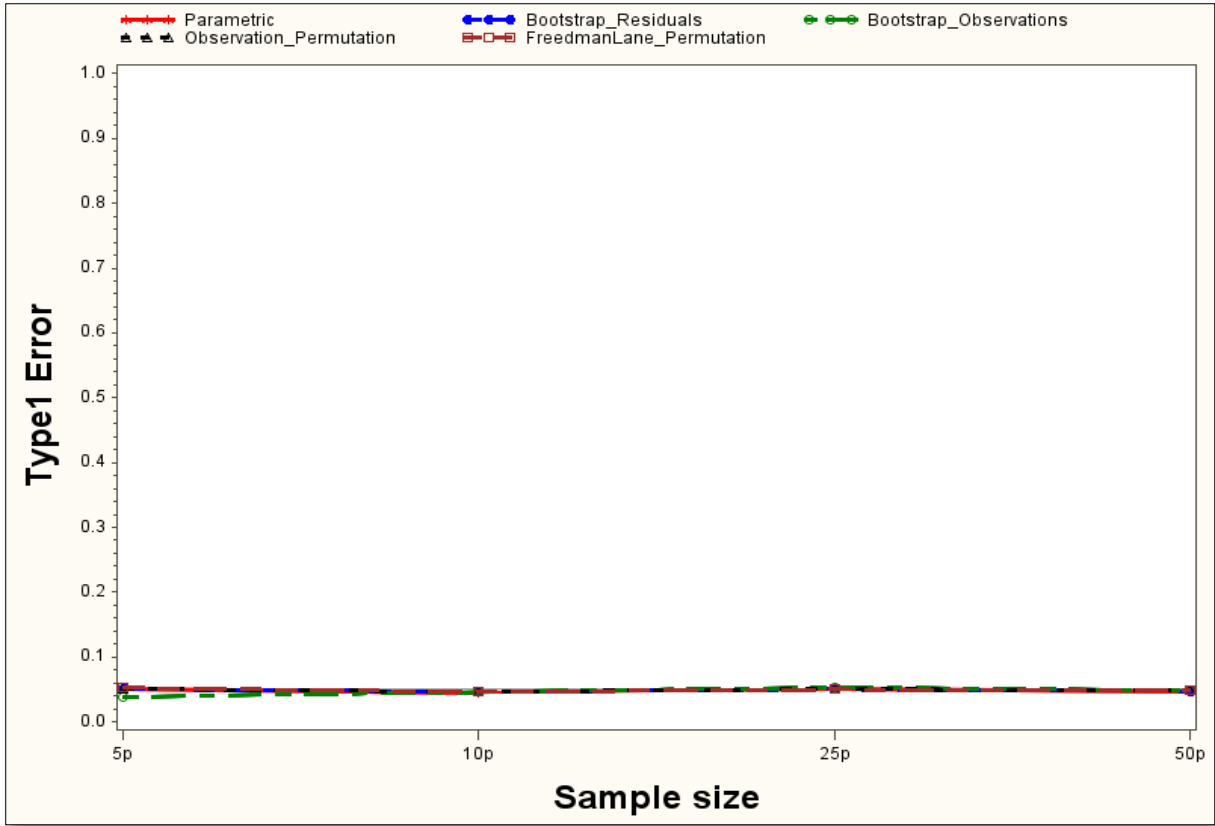


Figure 15: Type 1 error for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0$ on wider vertical axis

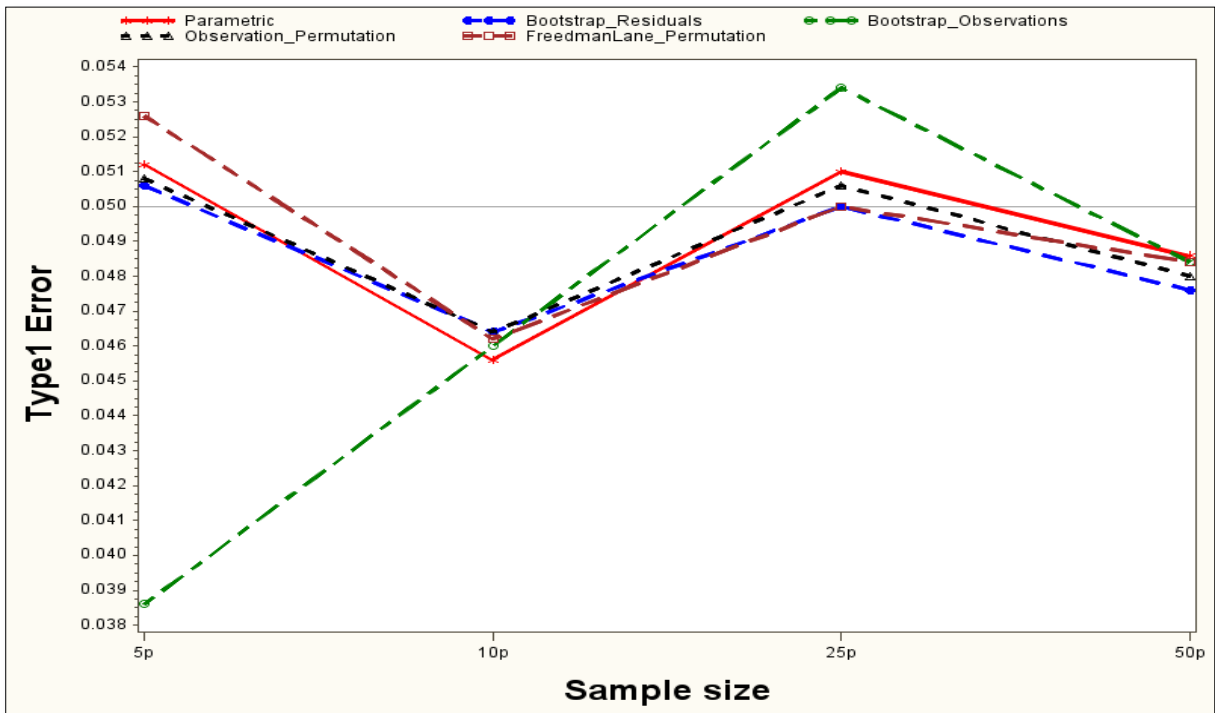


Figure 16: Type 1 error for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0$ on narrow vertical axis

Probability of type 1 error results for $p=5$ LOGN and $\rho=0.0$

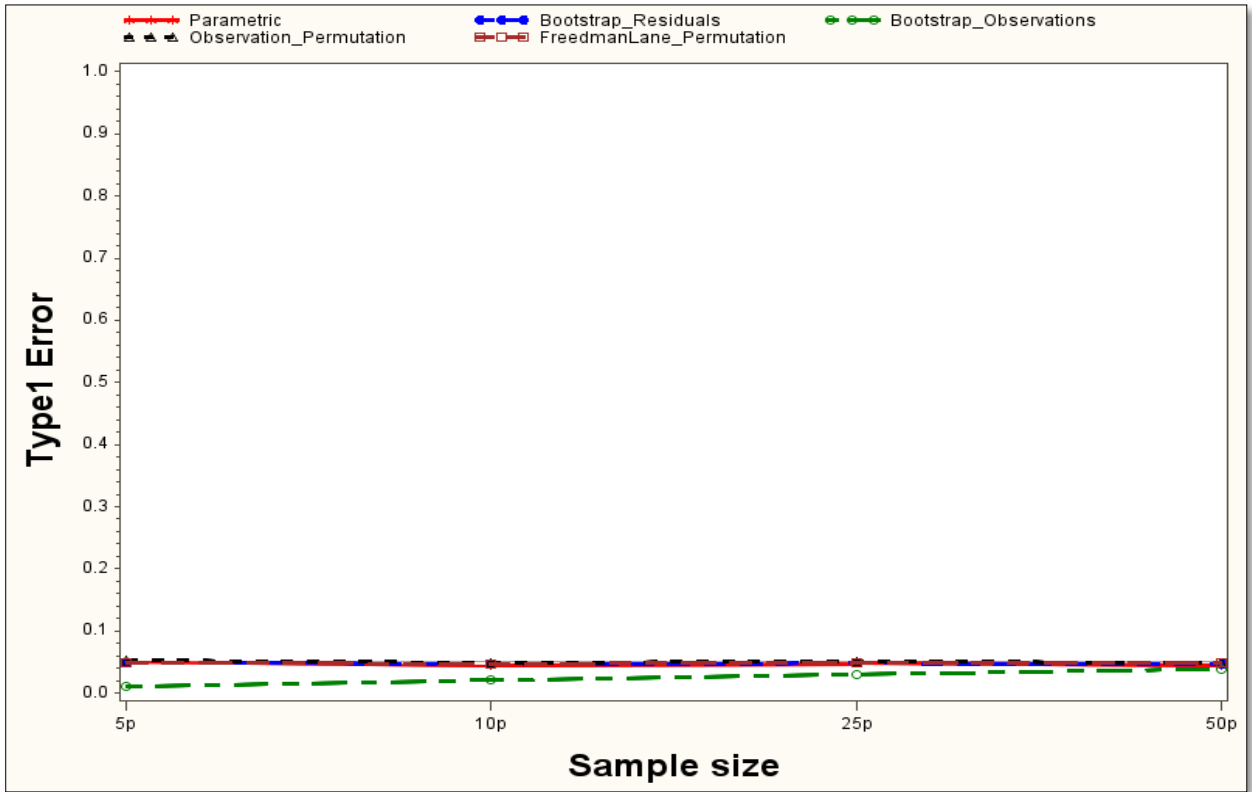


Figure 17: Type 1 error for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on wider vertical axis

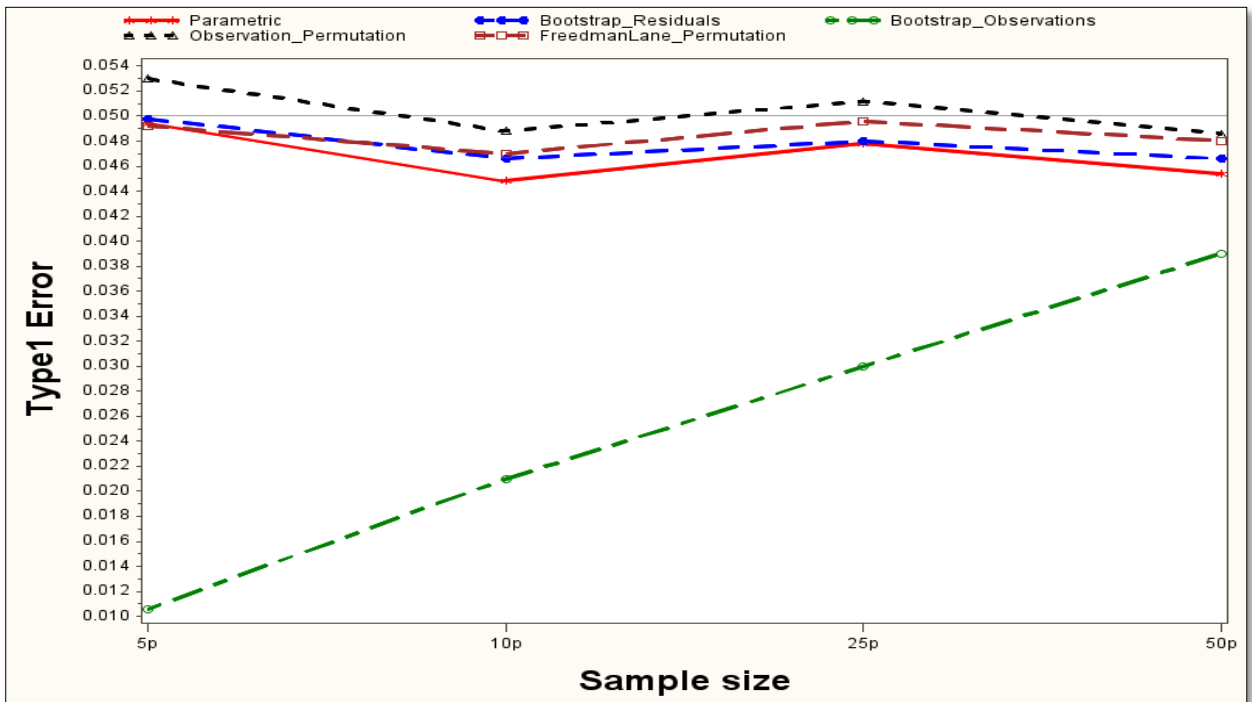


Figure 18: Type 1 error for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on narrow vertical axis

When errors are normally distributed, no method consistently outperforms others. For $p = 1$ all methods produced a probability of type 1 error bounded between 0.043 and 0.058 both inclusive. The minimum and maximum probability of type 1 error of 0.043 and 0.058 are both obtained under bootstrapping observations when the sample size is $10p$ and $25p$ respectively. When $p = 2$ all methods slightly overstate the probability of type 1 error for smaller sample sizes. The biggest overstatement of the probability of type 1 error is 0.058 which is observed when sample size is $25p$ for both bootstrapping observations and Freedman and Lane permutation method. It is difficult to point out which method performed best when $p = 5$. It can be observed that in all cases where errors are normally distributed the maximum probabilities of type 1 error were attained at sample size $25p$ with the Freedman and Lane permutation method being the biggest at 0.056. Generally all methods moved in the same direction (increasing/decreasing) with the increase in sample size. The minimum probability of type 1 error of 0.0308 is observed for bootstrapping observation when the sample size is $5p$.

When errors are distributed uniformly the case where $p = 1$ shows a consistent understatement of the probability of type 1 error across all resampling methods at sample size $5p$. From size $10p$ there is a close alignment of all the methods in stating the probabilities of type 1 error with an exception of bootstrapping observations method which consistently overstates the probability of type 1 error for sample sizes $10p$ to $50p$. For $p = 2$ the bootstrapping observations method produced probability of type 1 error of 0.0606 for samples sizes $10p$ and $25p$ and dropped to 0.051 at sample size $50p$ which is bringing it in line with the rest of the methods. The rest of the methods produced similar probability of type 1 error throughout the sample sizes. Similar values of probability of type 1 error errors were observed when $p = 2$ at size $10p$ and $50p$ are comparable to those observed when $p = 1$ at the same sizes. This is an indication that the introduction of an extra explanatory variable independent of the first variable does not impact the probability of type 1 error for $\beta_1 = 0$.

The close alignment between all methods continue to be prevalent across methods including the bootstrapping observations method which was slightly different from the rest for $p = 1$ and $p = 2$ when errors are uniformly distributed. For the bootstrapping observations the close alignment to the rest of the methods starts from the sample size of $10p$, for the sample size of $5p$ the probability of type 1 error is the lowest at 0.0386 for bootstrapping observations method.

When errors are lognormally distributed all methods except bootstrapping observations are closely aligned and with minimal fluctuations around the true probability of type 1 error of 0.05 across sample sizes and in all cases (i.e. explanatory variables cases $p = 1, 2$ and 5). The bootstrapping observations method consistently understates the probability of type 1 error across all cases.

6.2.2 Probability of type 1 error in the presence of multicollinearity ($\rho = 0.5$)

Probability of type 1 error results for $p=2$ Normal and $\rho=0.50$

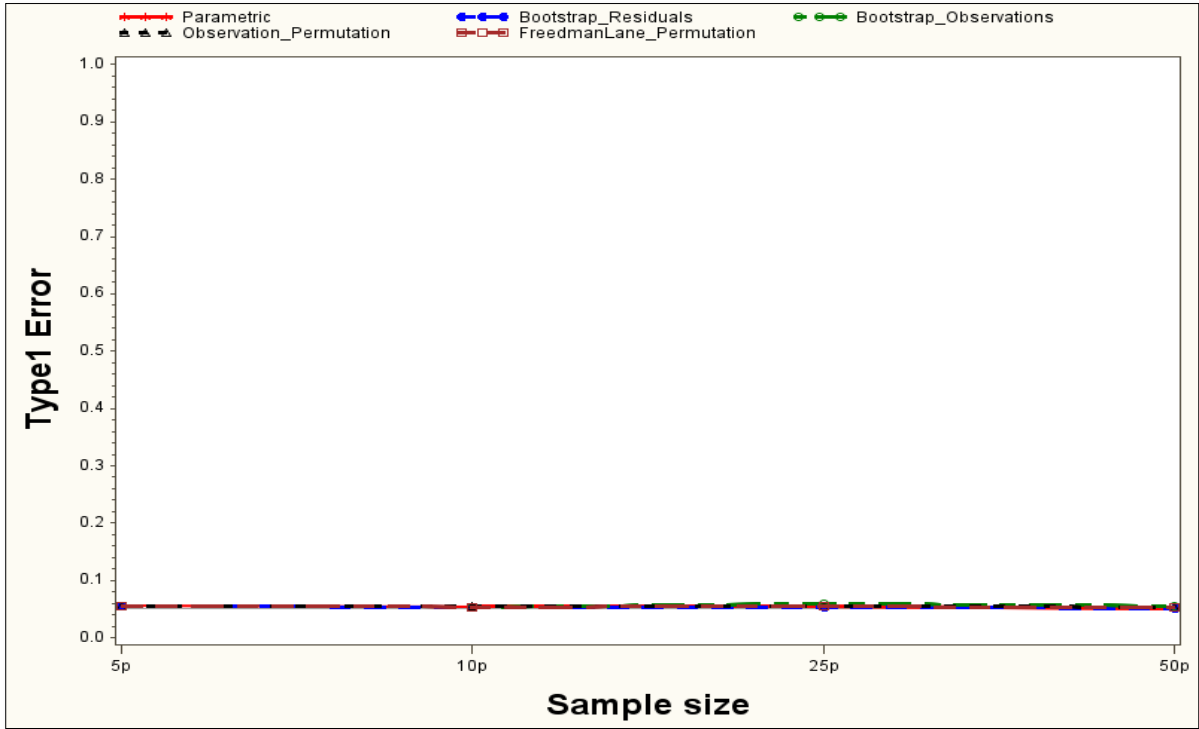


Figure 19: Type 1 error for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.5$ on wider vertical axis

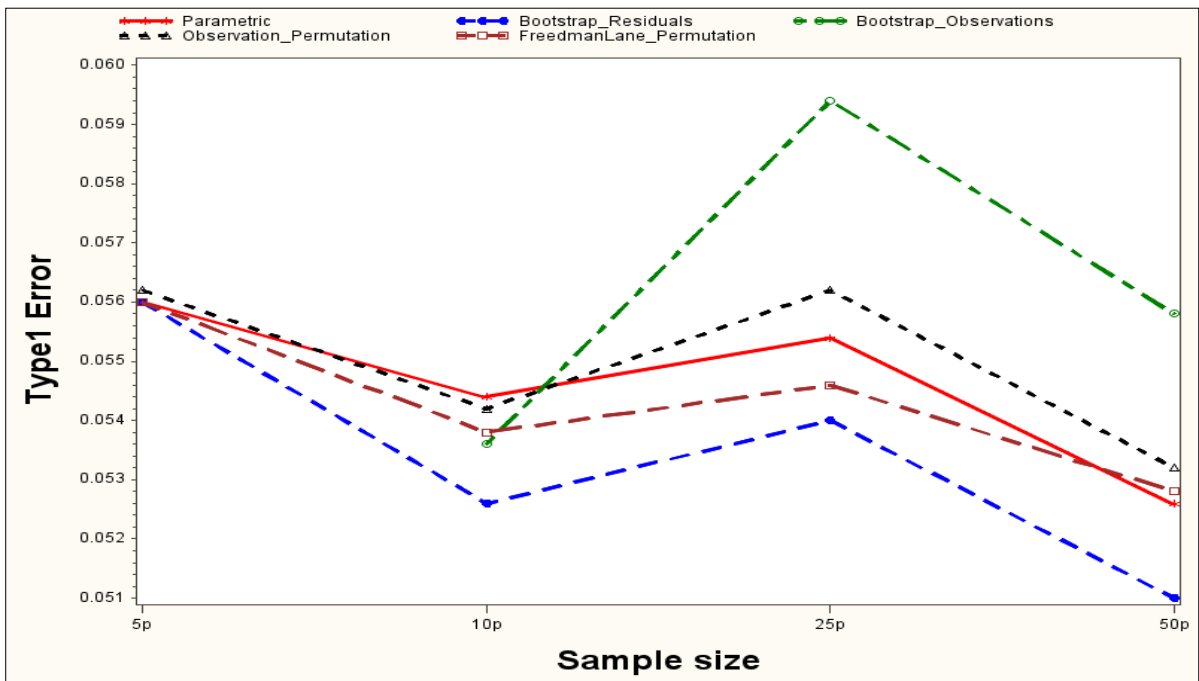


Figure 20: Type 1 error for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis

Probability of type 1 error results for $p=2$ Uniform and $\rho=0.50$

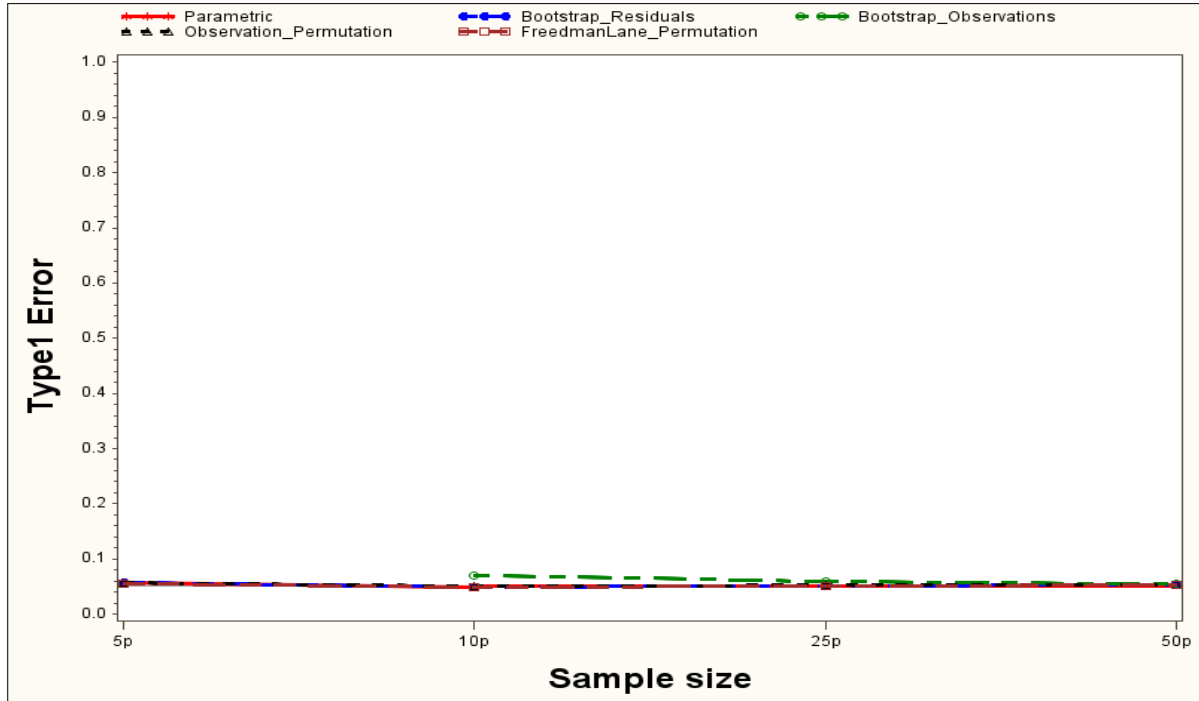


Figure 21: Type 1 error for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on wider vertical axis

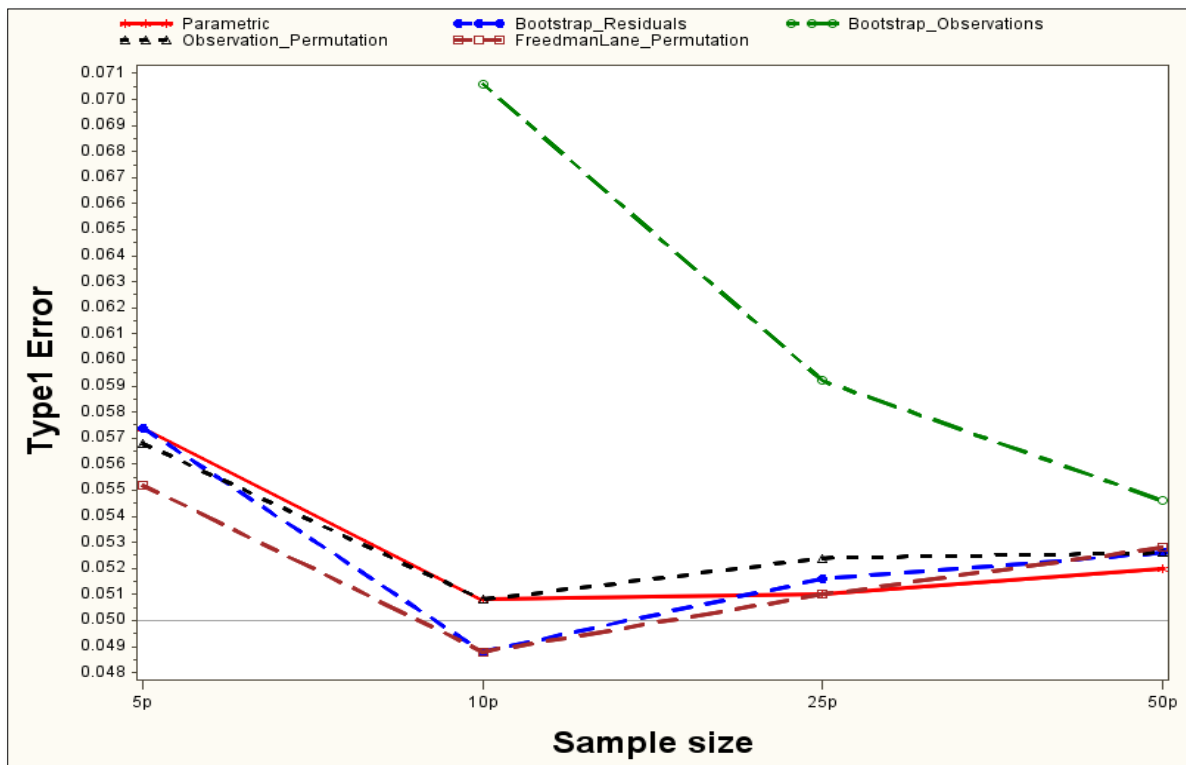


Figure 22: Type 1 error for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis

Probability of type 1 error results for $p=2$ LOGN and $\rho=0.50$

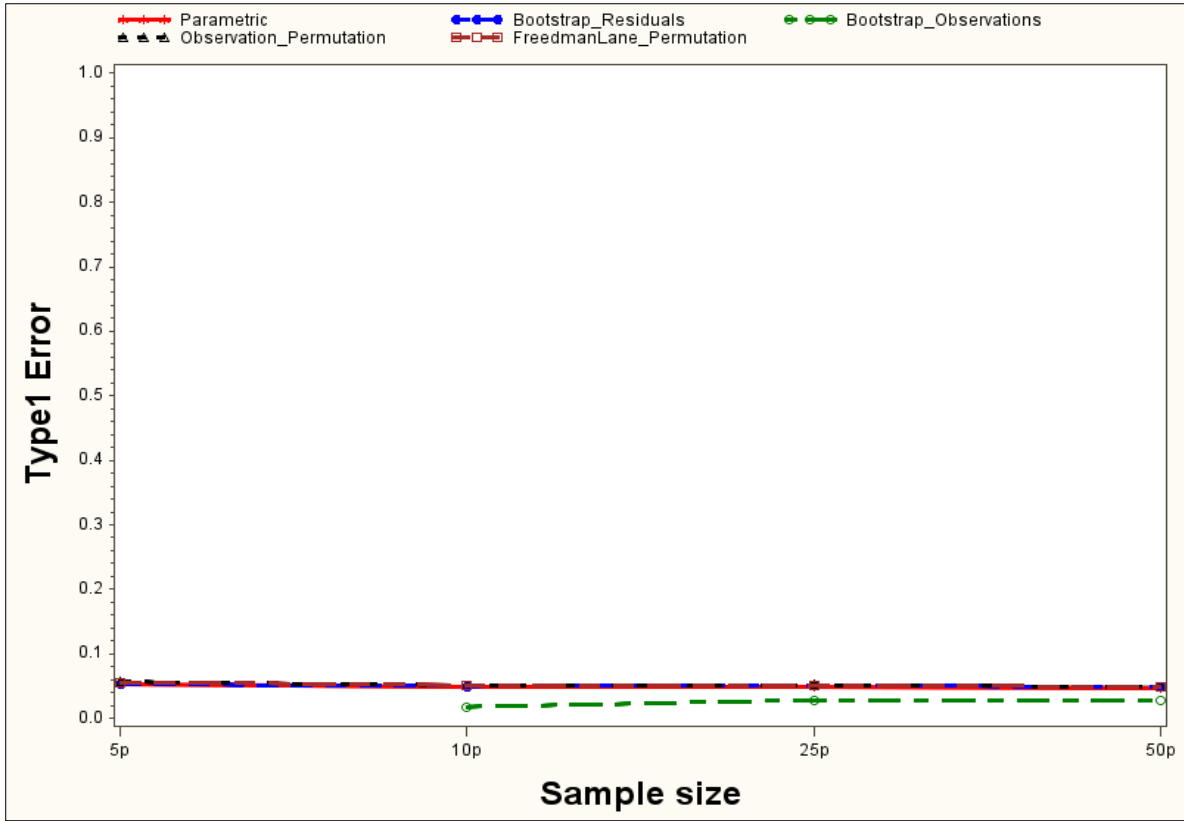


Figure 23: Type 1 error for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on wider vertical axis

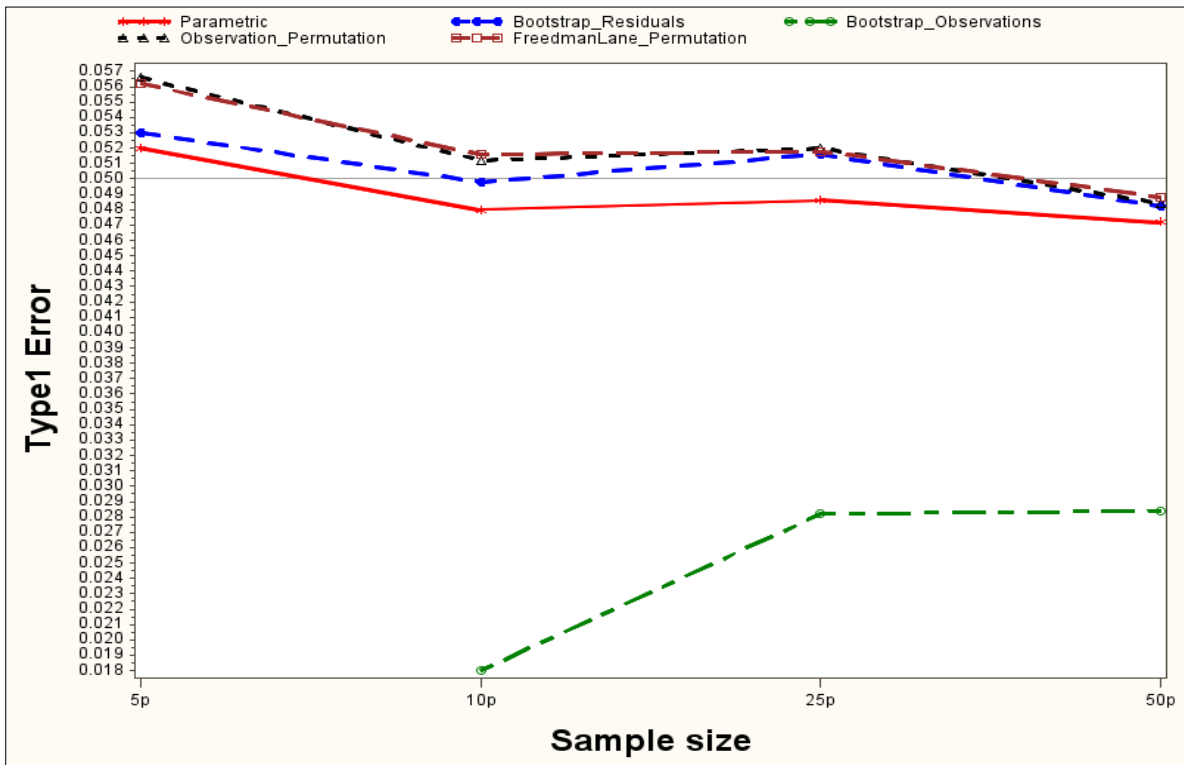


Figure 24: Type 1 error for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis

Probability of type 1 error results for $p=5$ Normal and $\rho=0.50$

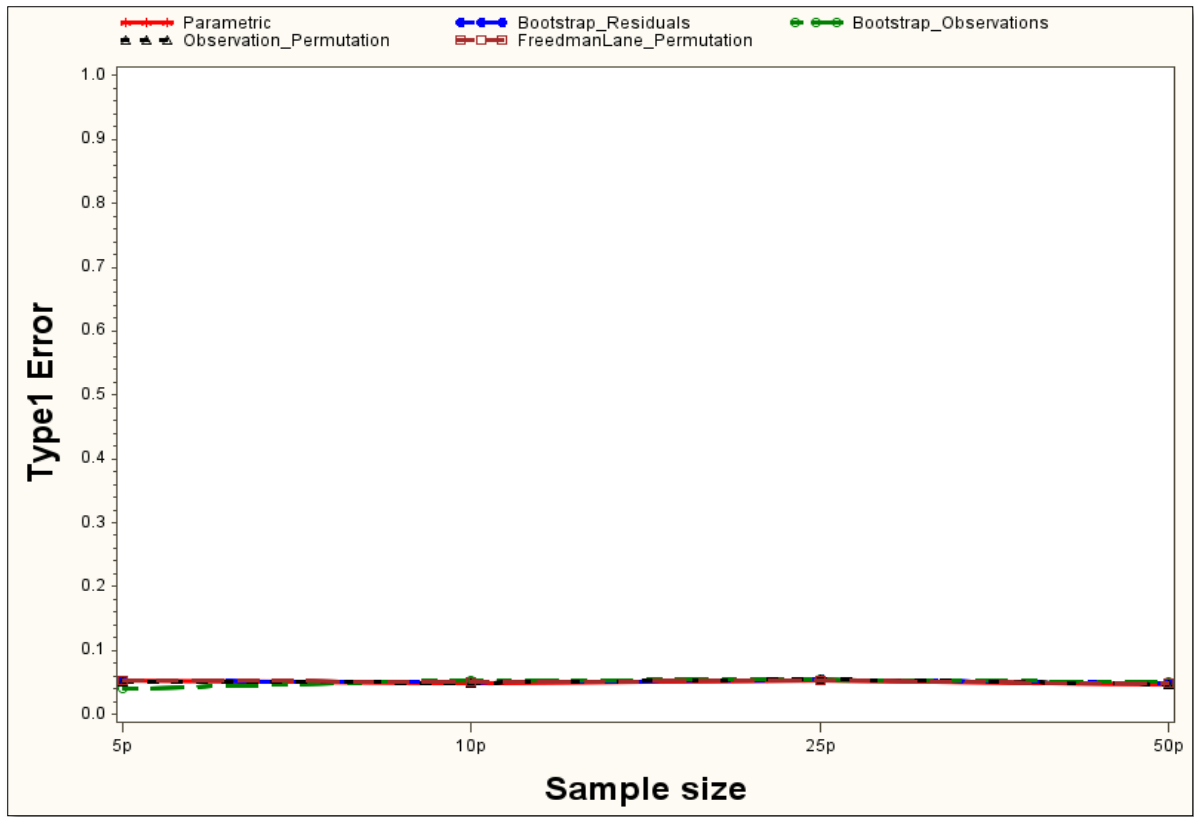


Figure 25: Type 1 error for $p = 5$, $\varepsilon \sim Normal(0, 1)$ and $\rho = 0.5$ on wider vertical axis

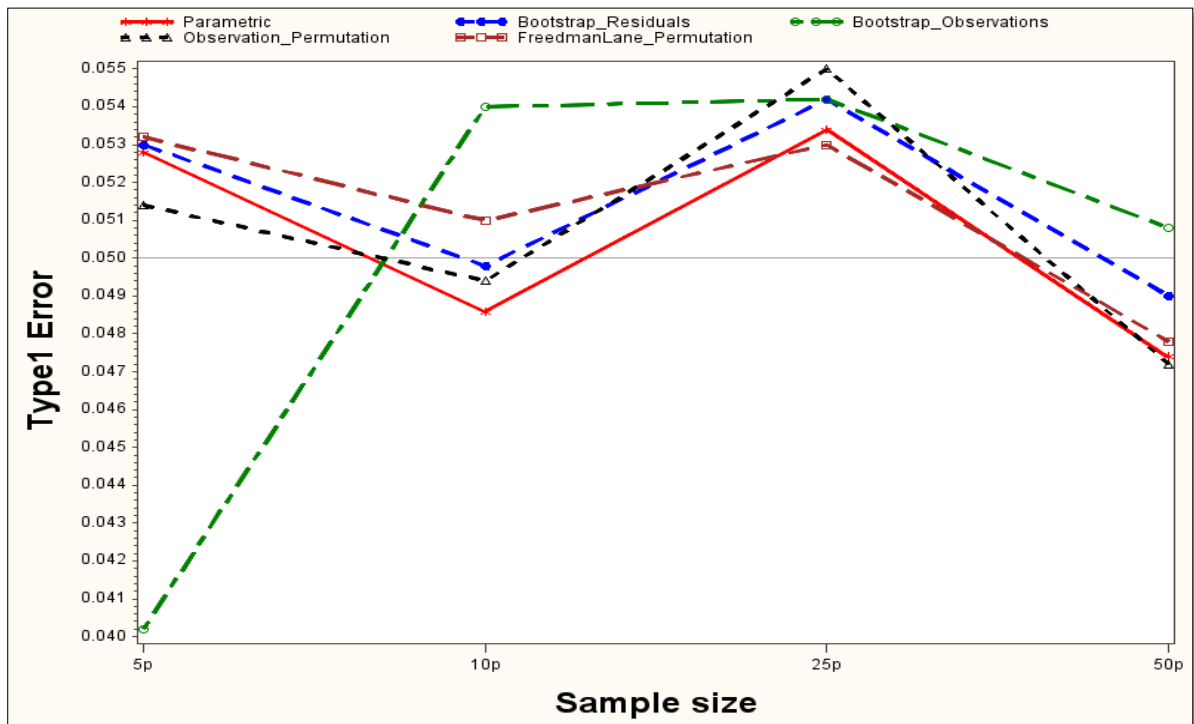


Figure 26: Type 1 error for $p = 5$, $\varepsilon \sim Normal(0, 1)$ and $\rho = 0.5$ on narrow vertical axis

Probability of type 1 error results for $p=5$ Uniform and $\rho=0.50$

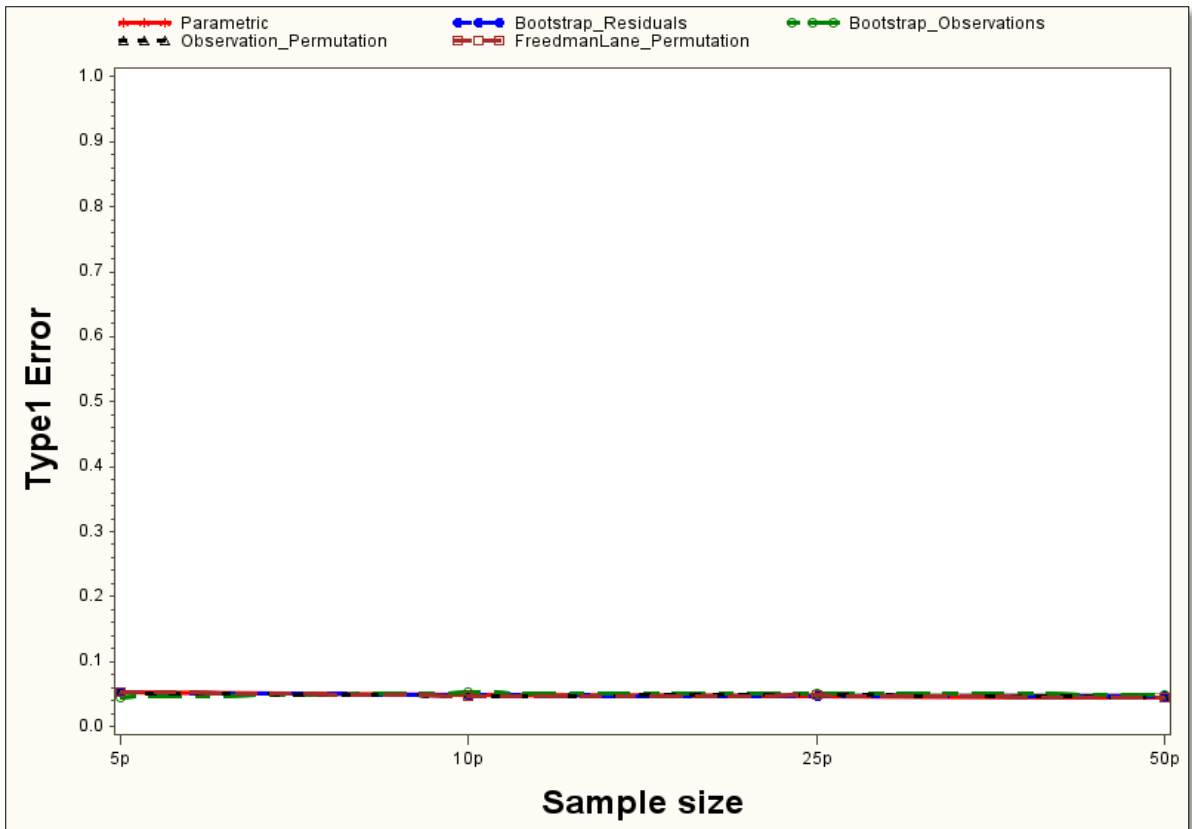


Figure 27: Type 1 error for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on wider vertical axis

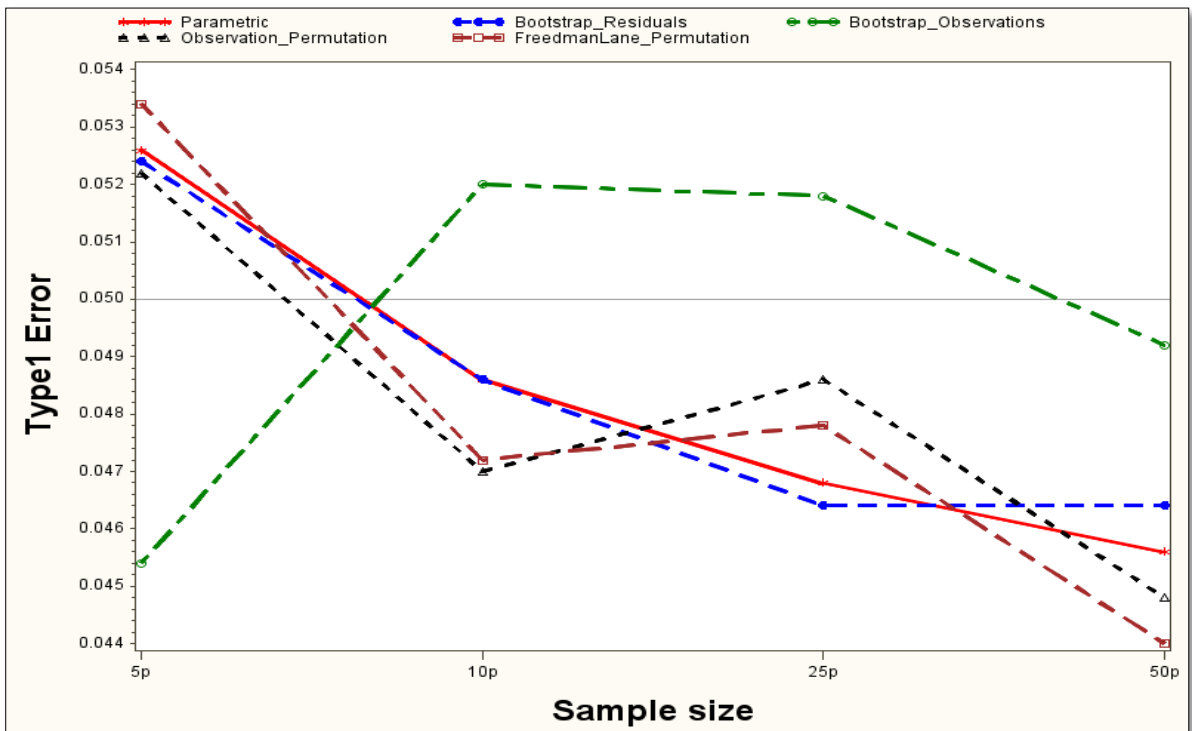


Figure 28: Type 1 error for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis

Probability of type 1 error results for $p=5$ LOGN and $\rho=0.50$

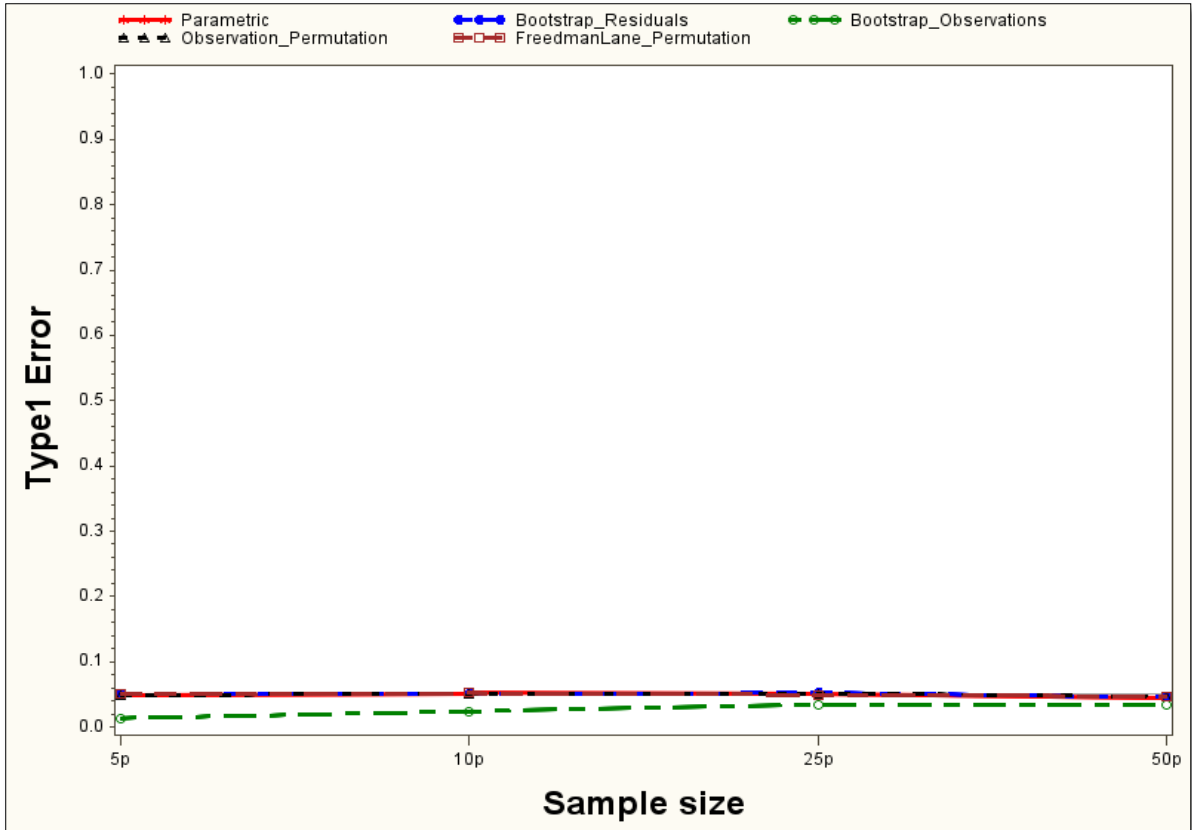


Figure 29: Type 1 error for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on wider vertical axis

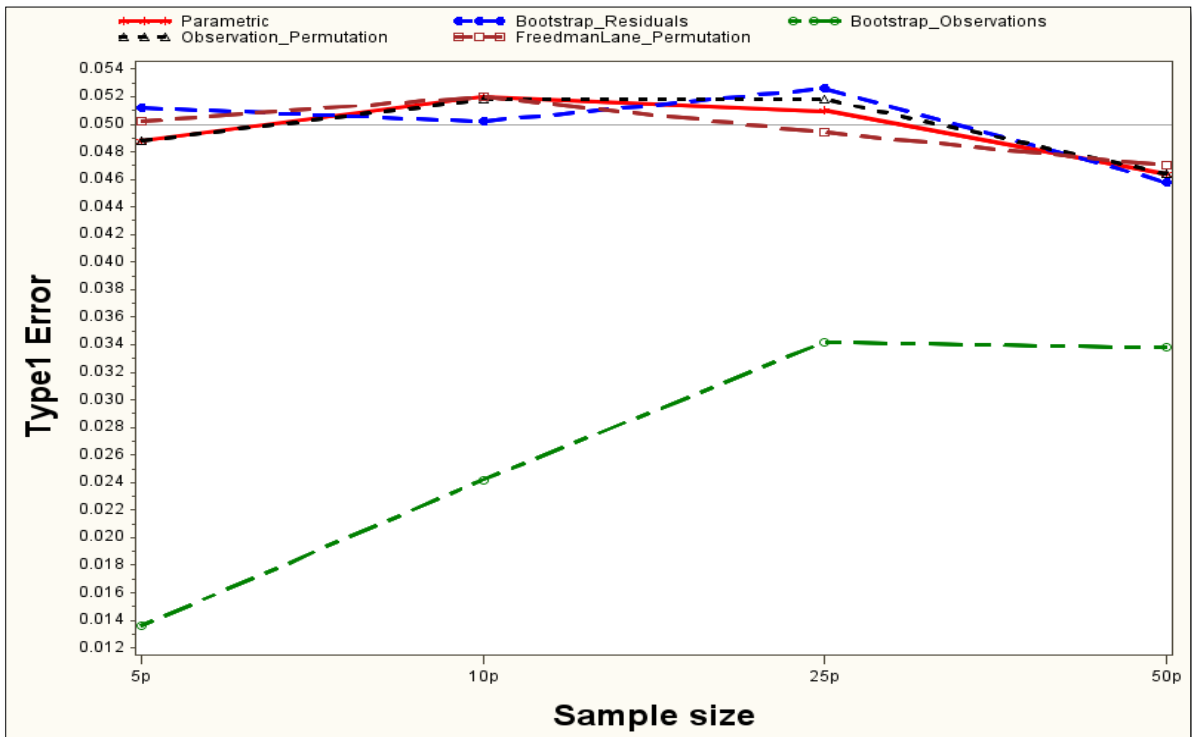


Figure 30: Type 1 error for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis

The introduction of a correlation structure in the independent variables affected the profile of the probability of type 1 error produced when errors are normally distributed for the case where $p = 2$. In the presence of multicollinearity, all methods consistently overstated the probability of type 1 error for all considered sample sizes. The maximum probability of type 1 error of 0.0594 was observed for bootstrapping observations at a sample size $25p$. The minimum observed probability of type 1 was 0.0526 which was observed under bootstrapping residuals method at a sample size $10p$.

For the case where $p = 5$, the probability of type 1 error produced by all methods is closely aligned for all sample sizes considered with an exception of bootstrapping observations method. This is an indication that the increase in the number of explanatory variables did not affect the probability of type 1 error for β_1 when errors are normally distributed even in the presence of multicollinearity.

The introduction of a correlation structure in the explanatory variables when errors are uniformly distributed in the case where $p = 2$ did not significantly affect the result probability of type 1 error. A slight shift upward in the probability of type 1 error is observed in all methods. The minimum and maximum probability of type 1 error obtained increased from 0.0476 to 0.0488 and 0.0606 to 0.0706 respectively. The maximum in both cases ($\rho = 0$ and $\rho = 0.5$) was attained for the bootstrapping observations method which is observed to consistently overstate the probability of type 1 error.

For the case where $p = 5$ there is very minimal difference in the type1 errors produced under multicollinearity and independence of explanatory variables. This indicates that multicollinearity of 0.5 among the explanatory variables does not distort the data enough to have an effect on probability of type 1 error.

An improvement in the type1 errors across all the methods was observed after the introduction of multicollinearity for the case $p = 2$ when errors are lognormally distributed. All the methods considered showed a much closer alignment as the samples sizes increased. A similar trend was also observed for $p = 5$.

6.2.3 Probability of type 1 error in the presence of increased multicollinearity ($\rho = 0.9$)

Probability of type 1 results for $p=2$ Normal and $\rho=0.90$

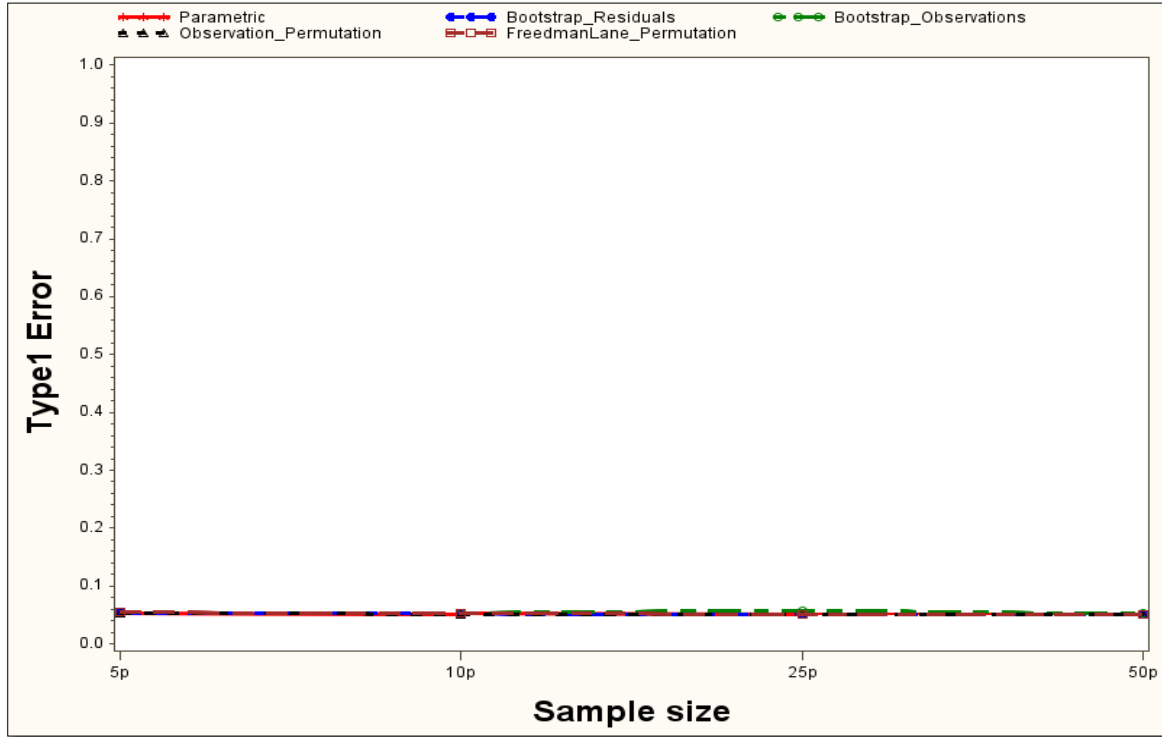


Figure 31: Type 1 error for $p = 2$, $\varepsilon \sim Normal(0, 1)$ and $\rho = 0.9$ on wider vertical axis

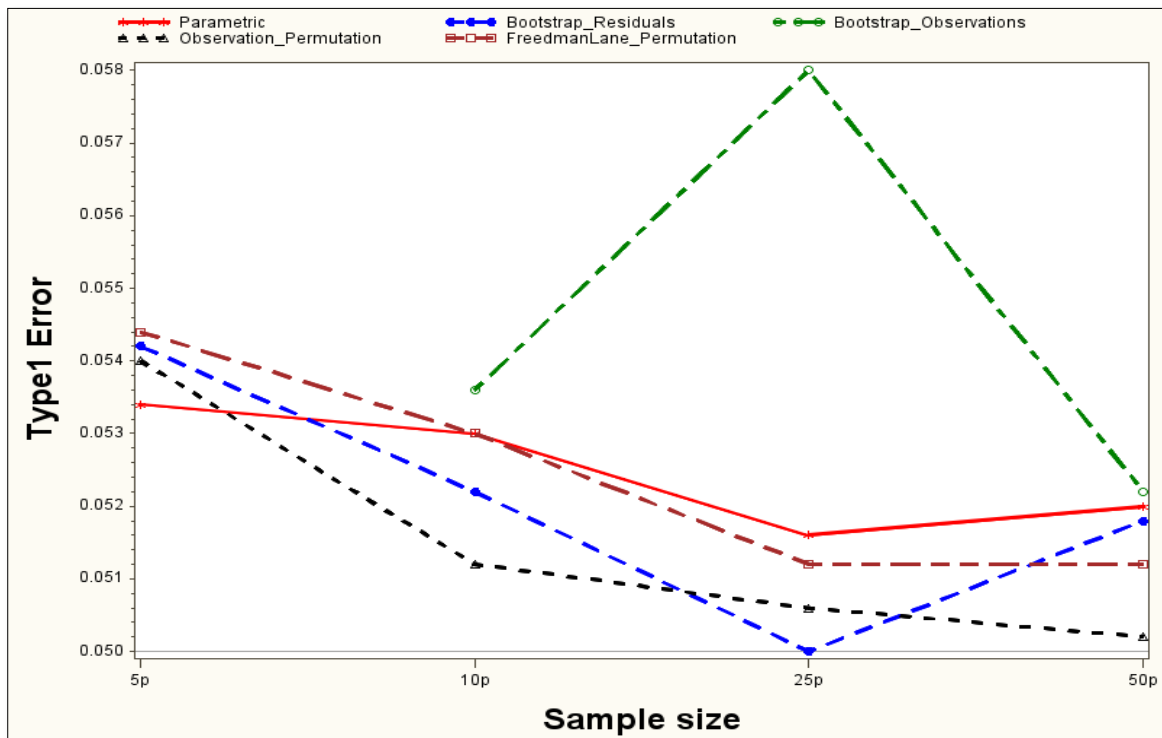


Figure 32: Type 1 error for $p = 2$, $\varepsilon \sim Normal(0, 1)$ and $\rho = 0.9$ on narrow vertical axis

Probability of type 1 results for $p=2$ Uniform and $\rho=0.90$

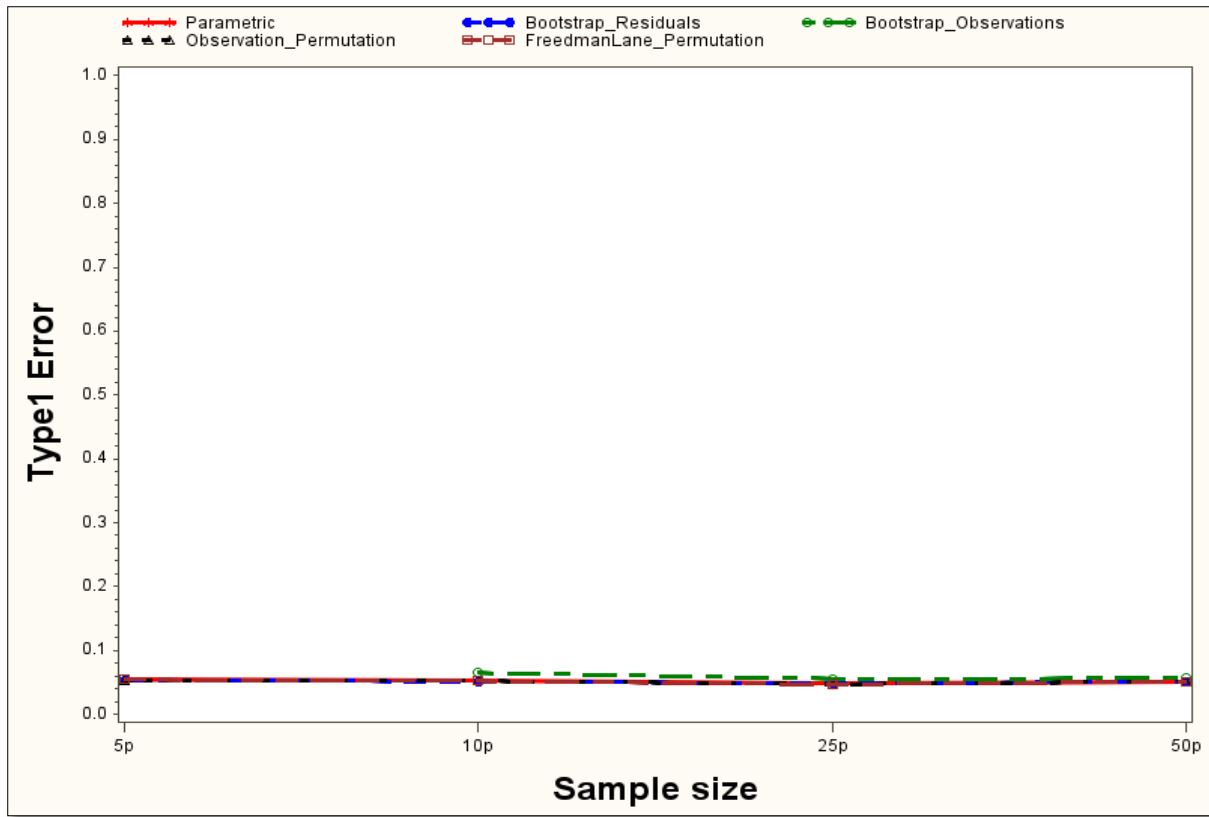


Figure 33: Type 1 error for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on wider vertical axis

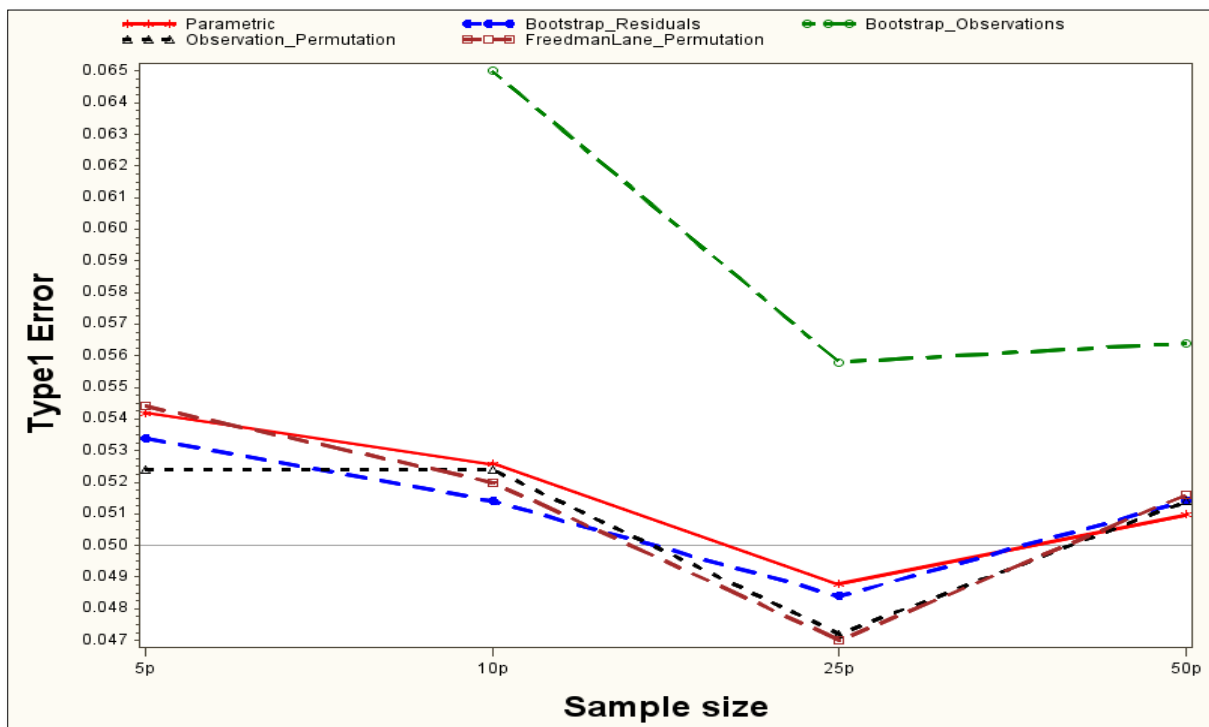


Figure 34: Type 1 error for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis

Probability of type 1 results for $p=2$ LOGN and $\rho=0.90$

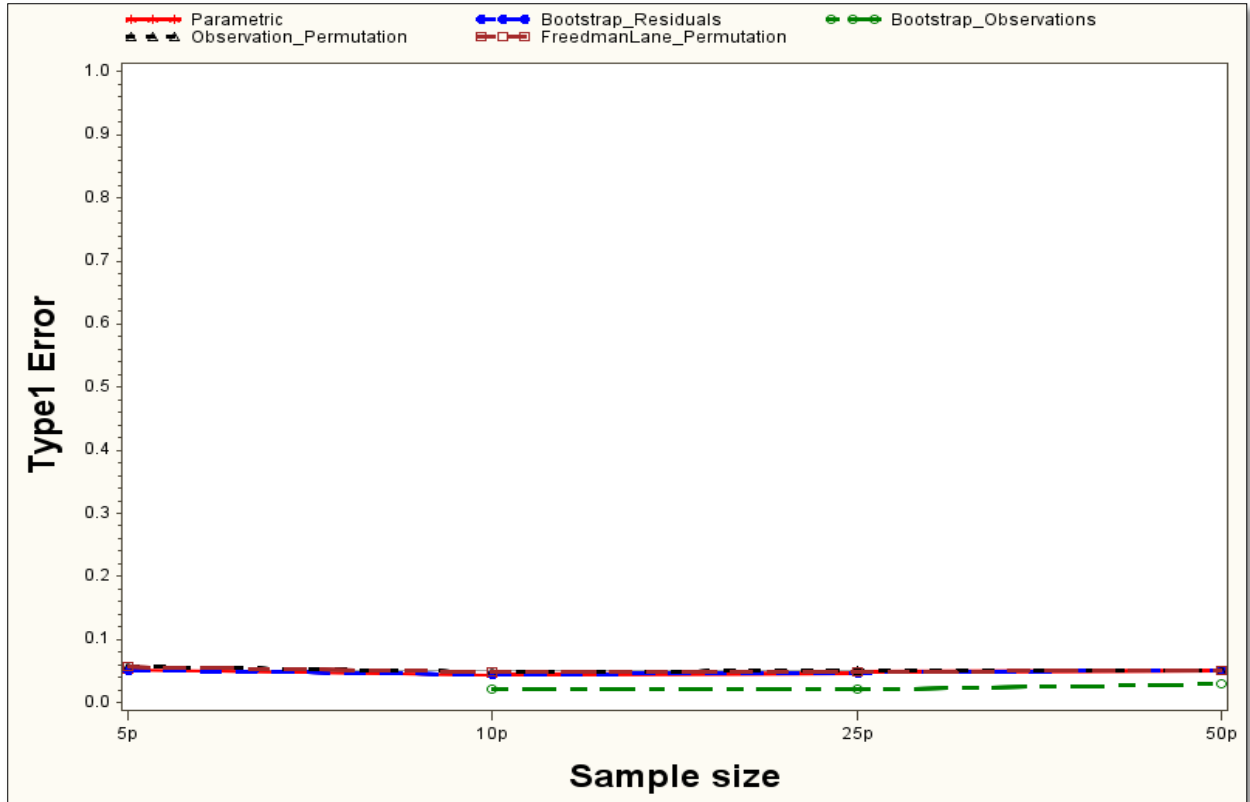


Figure 35: Type 1 error for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on wider vertical axis

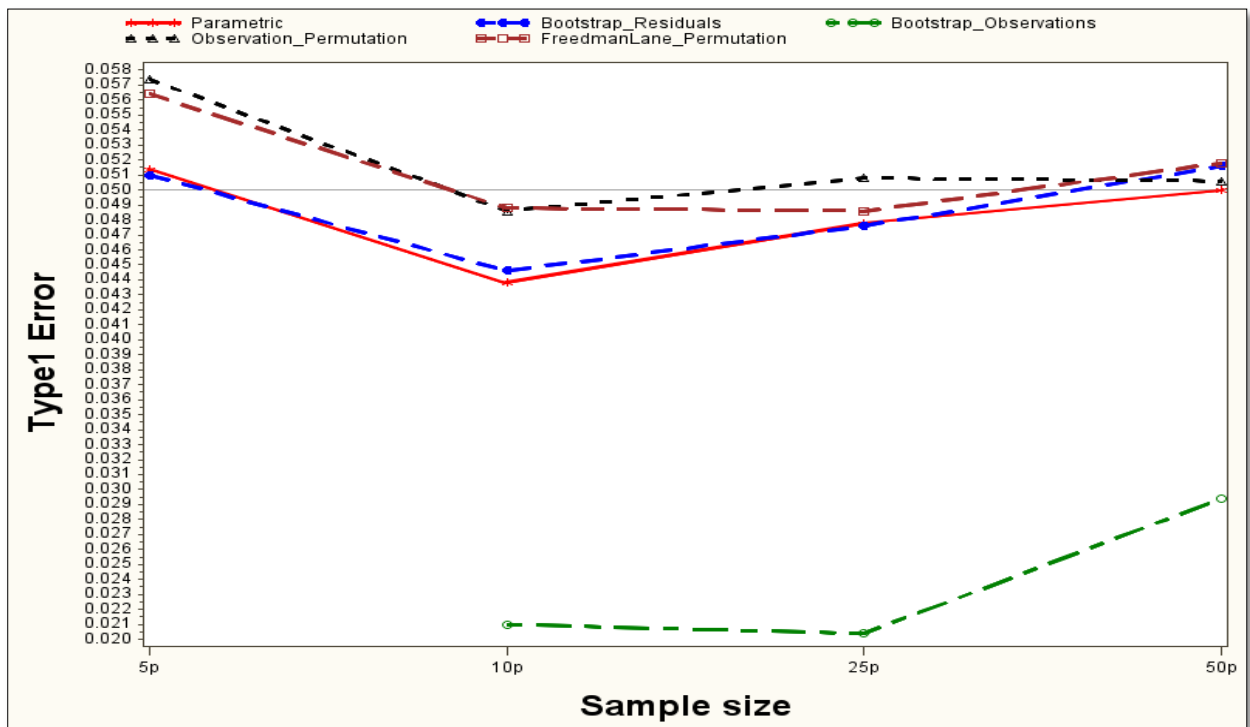


Figure 36: Type 1 error for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis

Probability of type 1 results for $p=5$ Normal and $\rho=0.90$

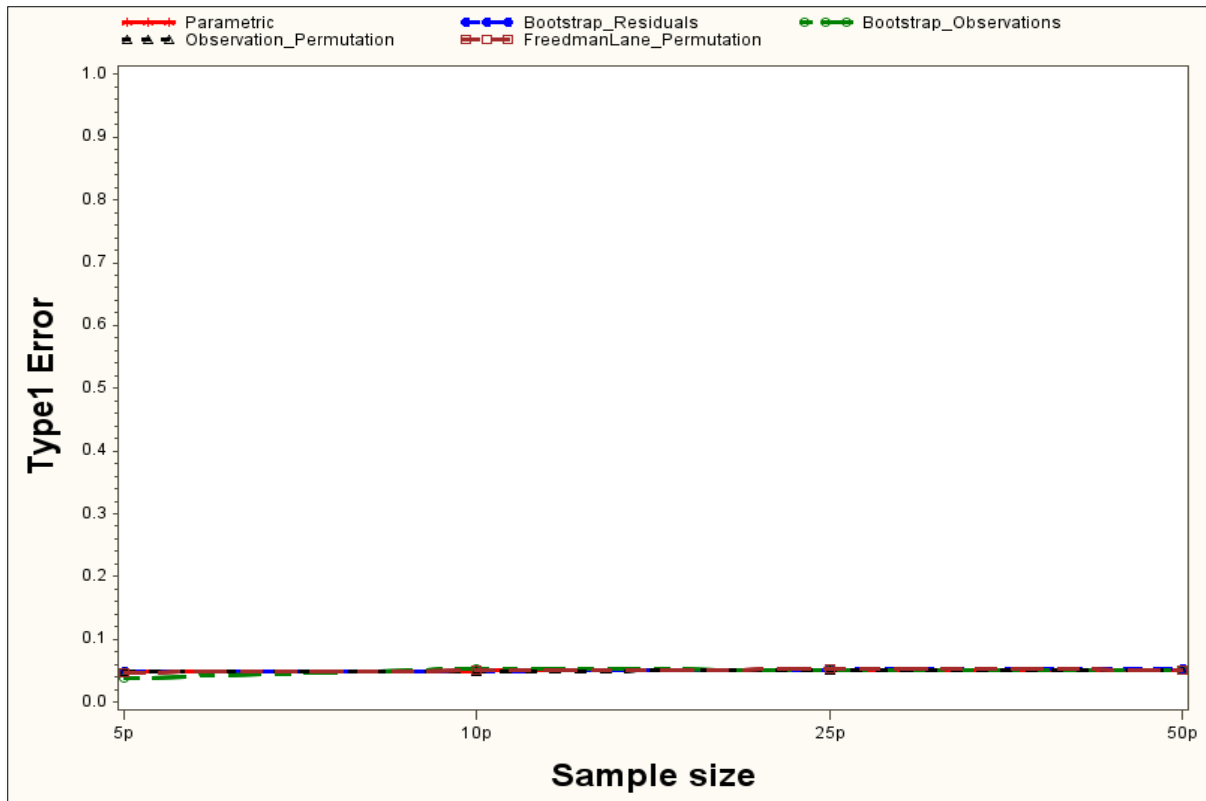


Figure 37: Type 1 error for $p = 5$, $\varepsilon \sim Normal(0, 1)$ and $\rho = 0.9$ on wider vertical axis

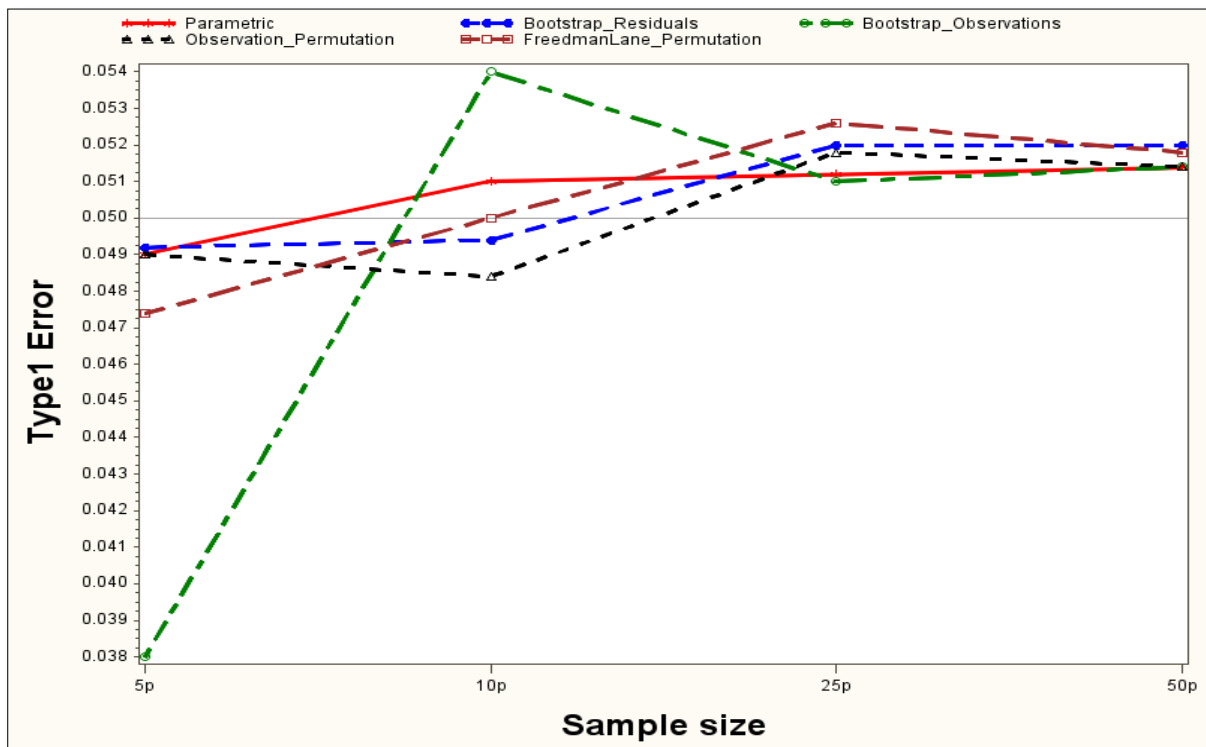


Figure 38: Type 1 error for $p = 5$, $\varepsilon \sim Normal(0, 1)$ and $\rho = 0.9$ on narrow vertical axis

Probability of type 1 results for $p=5$ Uniform and $\rho=0.90$

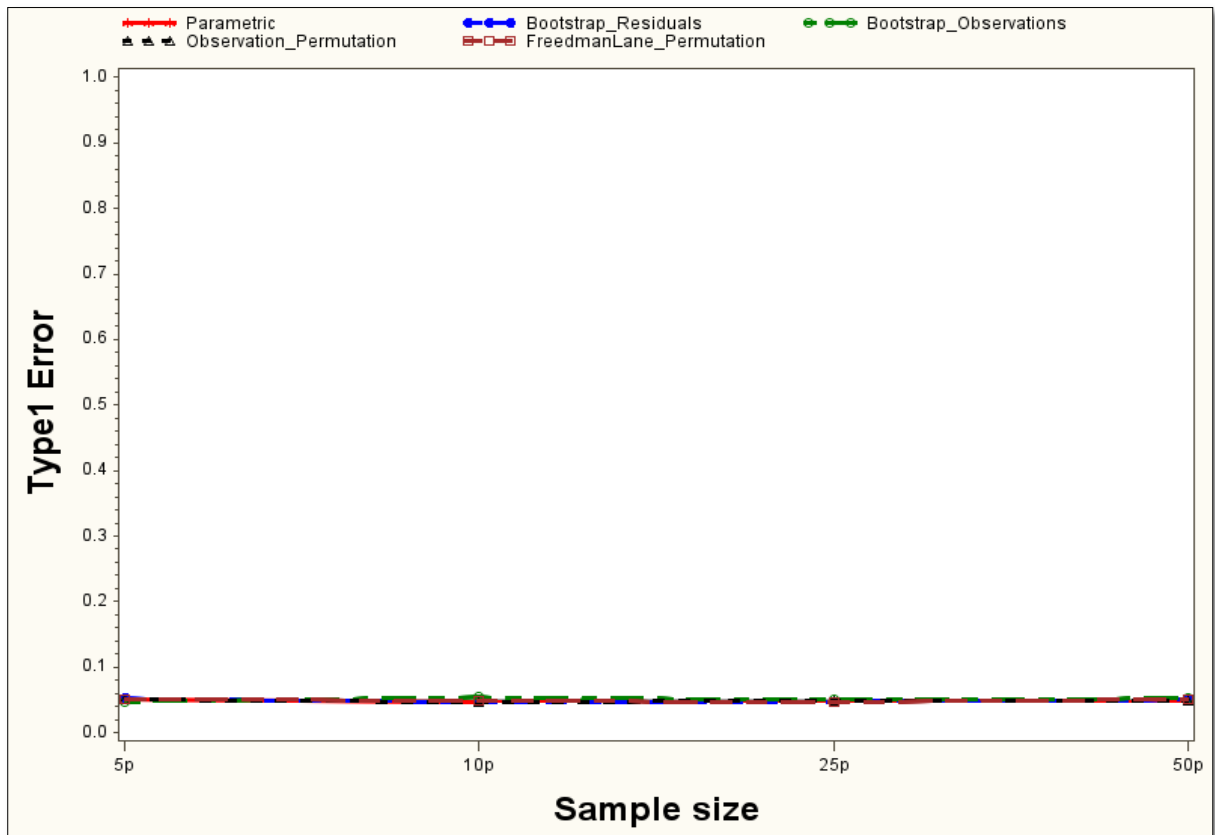


Figure 39: Type 1 error for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on wider vertical axis

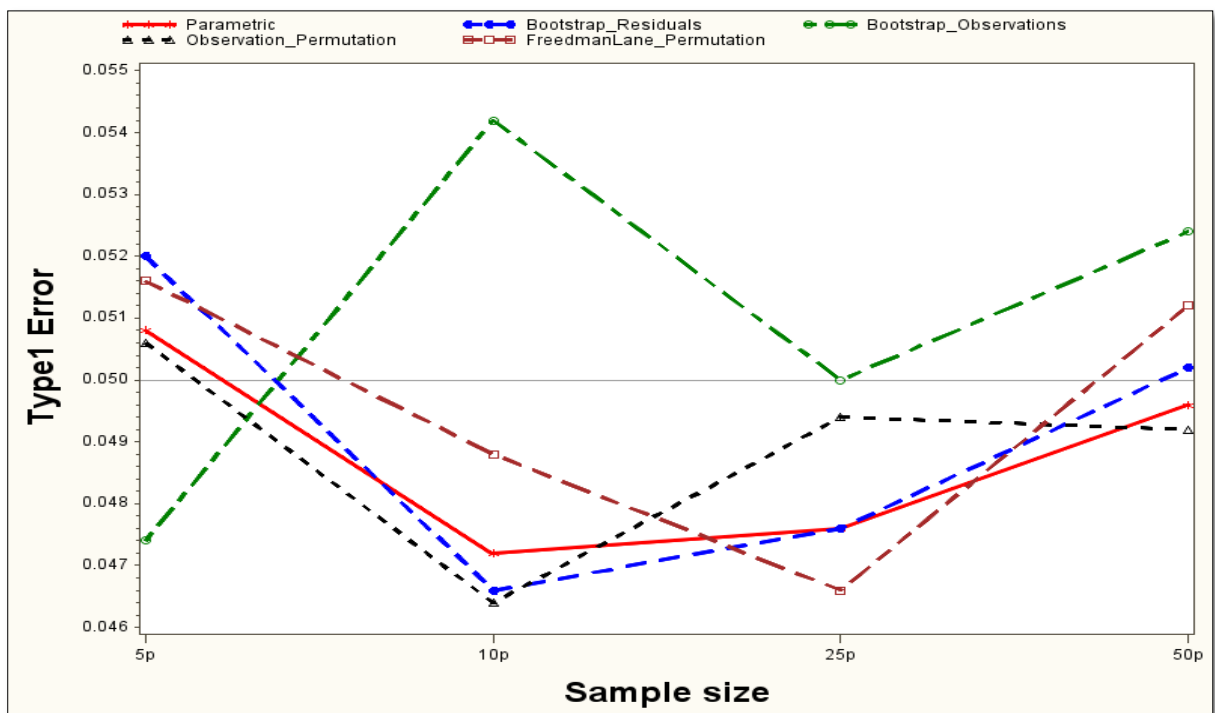


Figure 40: Type 1 error for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis

Probability of type 1 results for $p=5$ LOGN and $\rho=0.90$

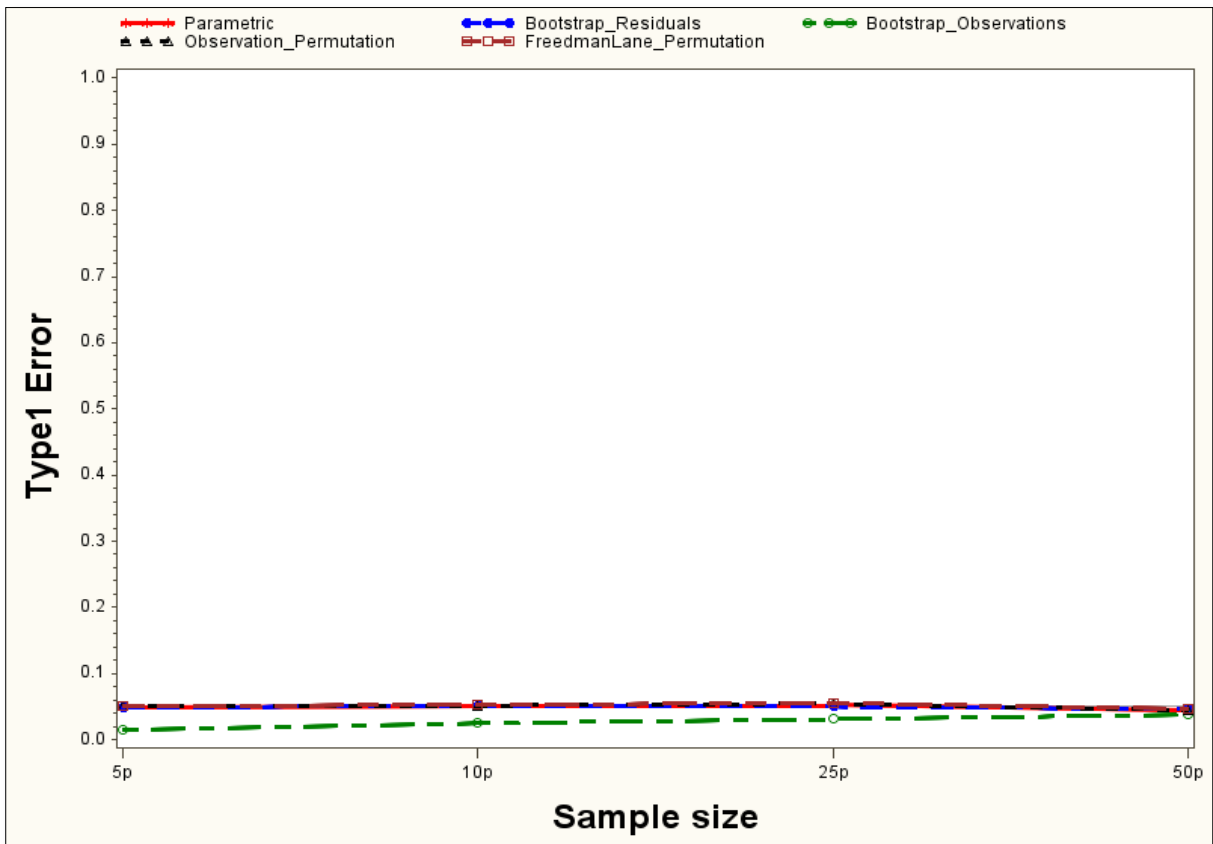


Figure 41: Type 1 error for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on wider vertical axis

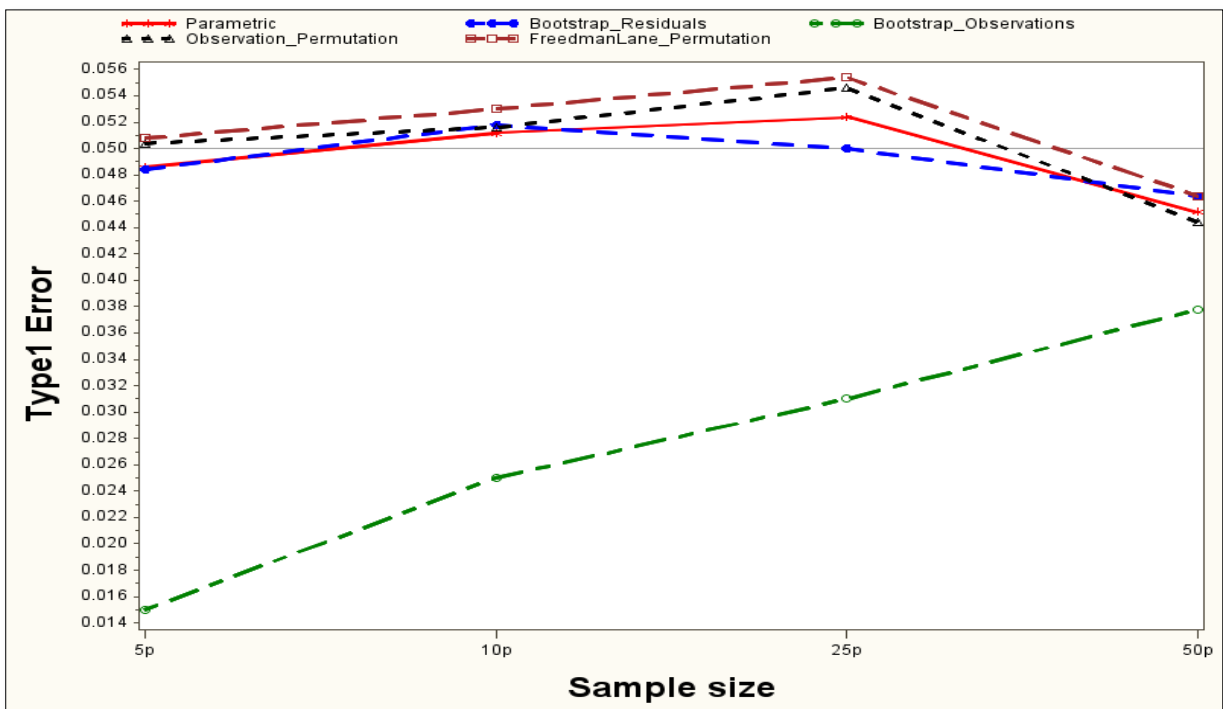


Figure 42: Type 1 error for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis

The increasing multicollinearity in the explanatory variables for the regression model with normally distributed errors for $p = 2$ did not affect the probabilities of type 1 error for all sample sizes and across all methods. There are cases where slight random fluctuations are observed when the multicollinearity is increased from 0.5 to 0.9 however none of these are significant. A similar trend is observed for $p = 5$.

For the case of $p = 2$ when errors are uniformly distributed the trend is similar to that of 0.5 multicollinearity where all other methods are closely aligned to each other with an exception of bootstrapping observations method. This method is overstating the probability of type 1 error more than the rest of the other methods, more especially for smaller sample sizes. Convergence with other methods is observed as the sample size increases. For the case of $p = 5$ there is a much closer alignment of bootstrapping observations with other methods more especially from sample size of $10p$. This is an indication that all the methods converge for larger sample sizes.

When errors are lognormally distributed the trends observed under the multicollinearity of 0.5 are still evident even when there has been an increase of multicollinearity to 0.9. Bootstrapping observation continues to severely understate the probability of type 1 error while the other methods are closely aligned around the true 0.05 probability of type 1 error. This is the case for both $p = 2$ and $p = 5$. The minimum and maximum probability of type 1 error achieved by bootstrapping observations are 0.021 and 0.0294 respectively for $p = 2$ and 0.015 and 0.0378 respectively for $p = 5$. It can be observed that though the bootstrapping observations understates the probability of type 1 error, this understatement of the probability of type 1 error improves as the sample size increase.

6.3 Power of a test

In all the cases power is examined when H_1 is $\beta_1 = 2$, Table of result can be seen at Appendix B.

6.3.1 Power under independent explanatory variables

Power results for $p=1$ Normal $\rho=0.0$

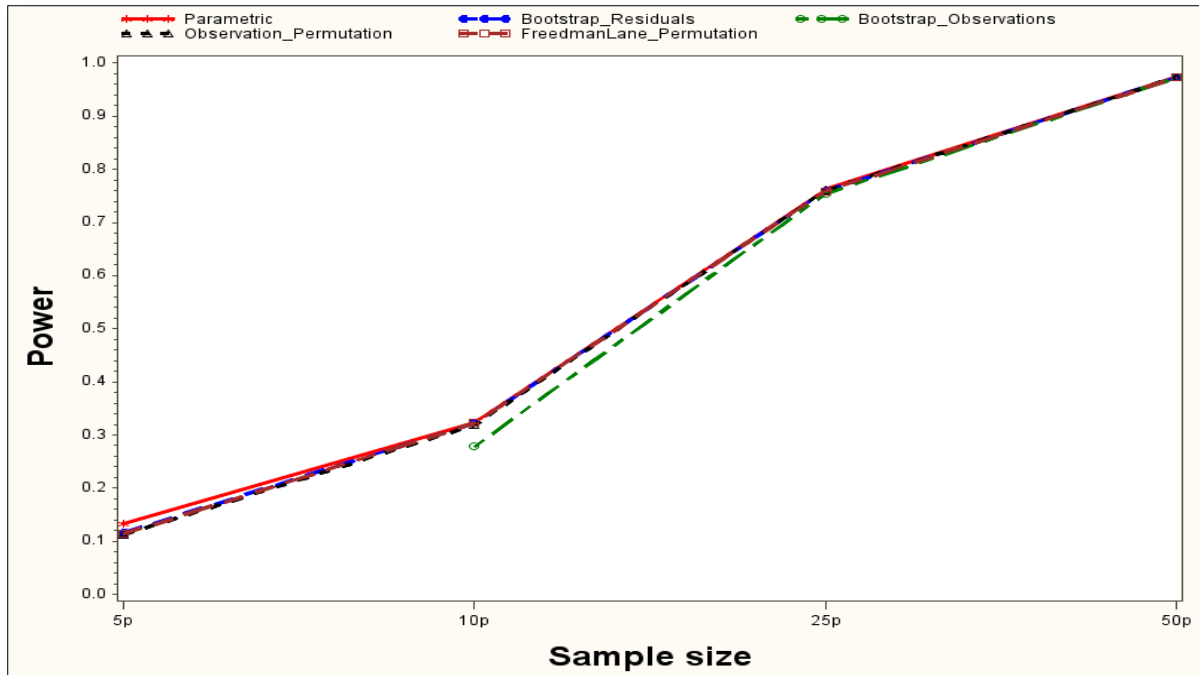


Figure 43: Power for $p = 1$ and $\varepsilon \sim Normal(0, 1)$ on wider vertical axis

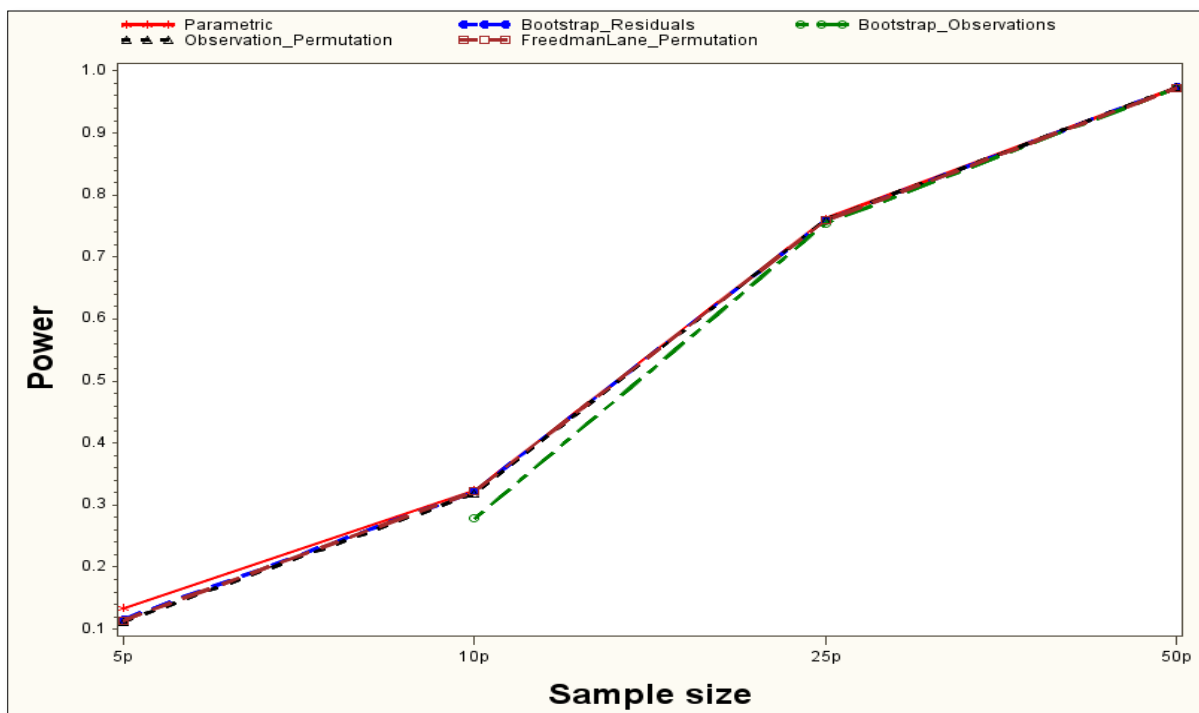


Figure 44: Power for $p = 1$ and $\varepsilon \sim Normal(0, 1)$ on narrow vertical axis

Power results for $p=1$ Uniform $\rho=0.0$

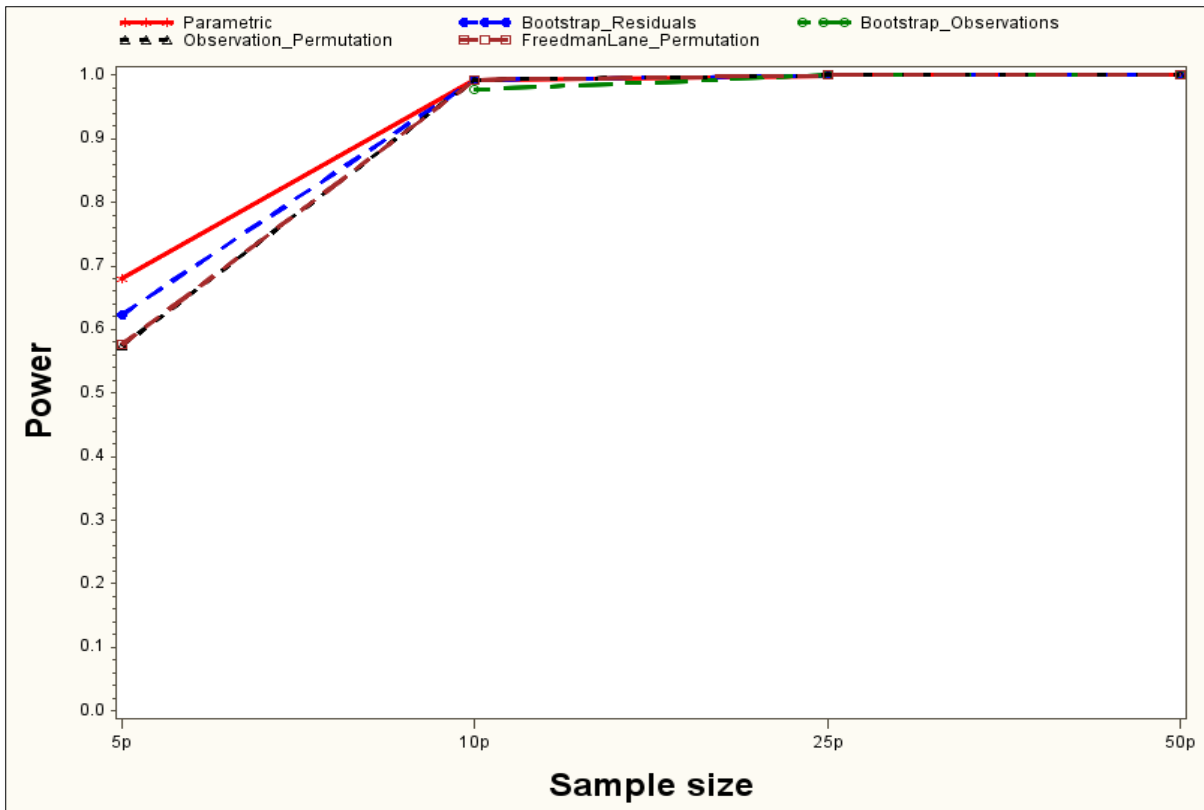


Figure 45 Power for $p = 1$ and $\varepsilon \sim \text{Uniform}(0, 1)$ on wider vertical axis

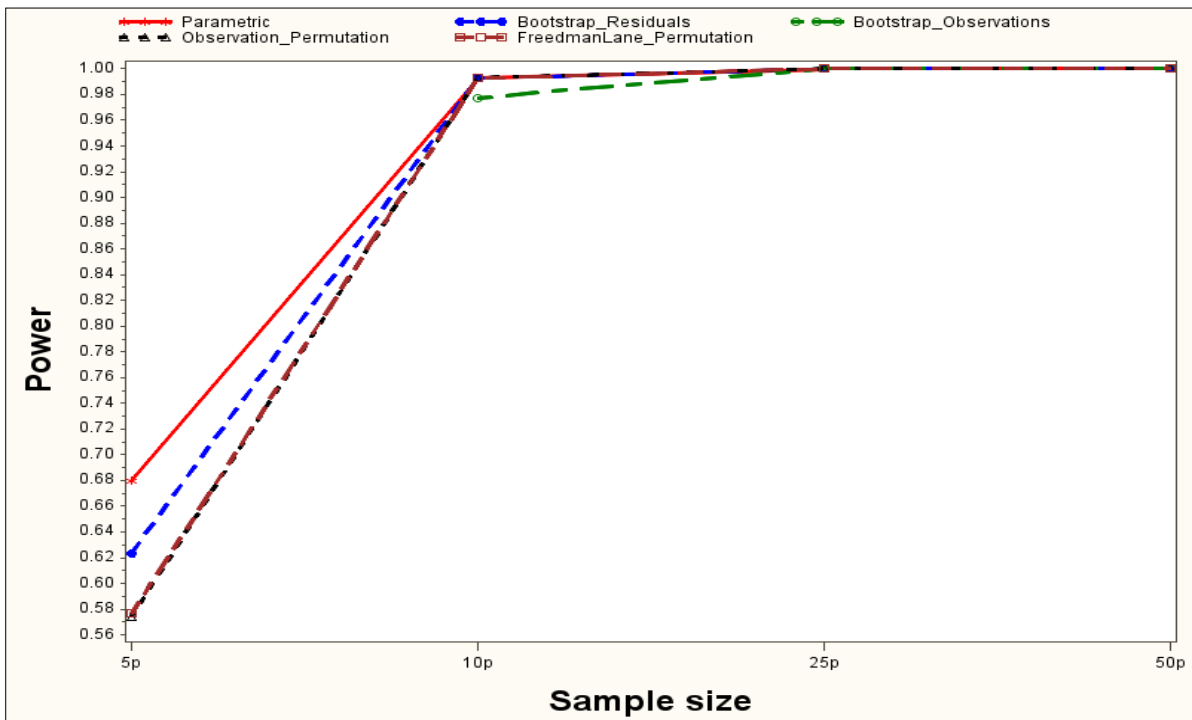


Figure 46: Power for $p = 1$ and $\varepsilon \sim \text{Uniform}(0, 1)$ on narrow vertical axis

Power results for $p=1$ LOGN $\rho=0.0$

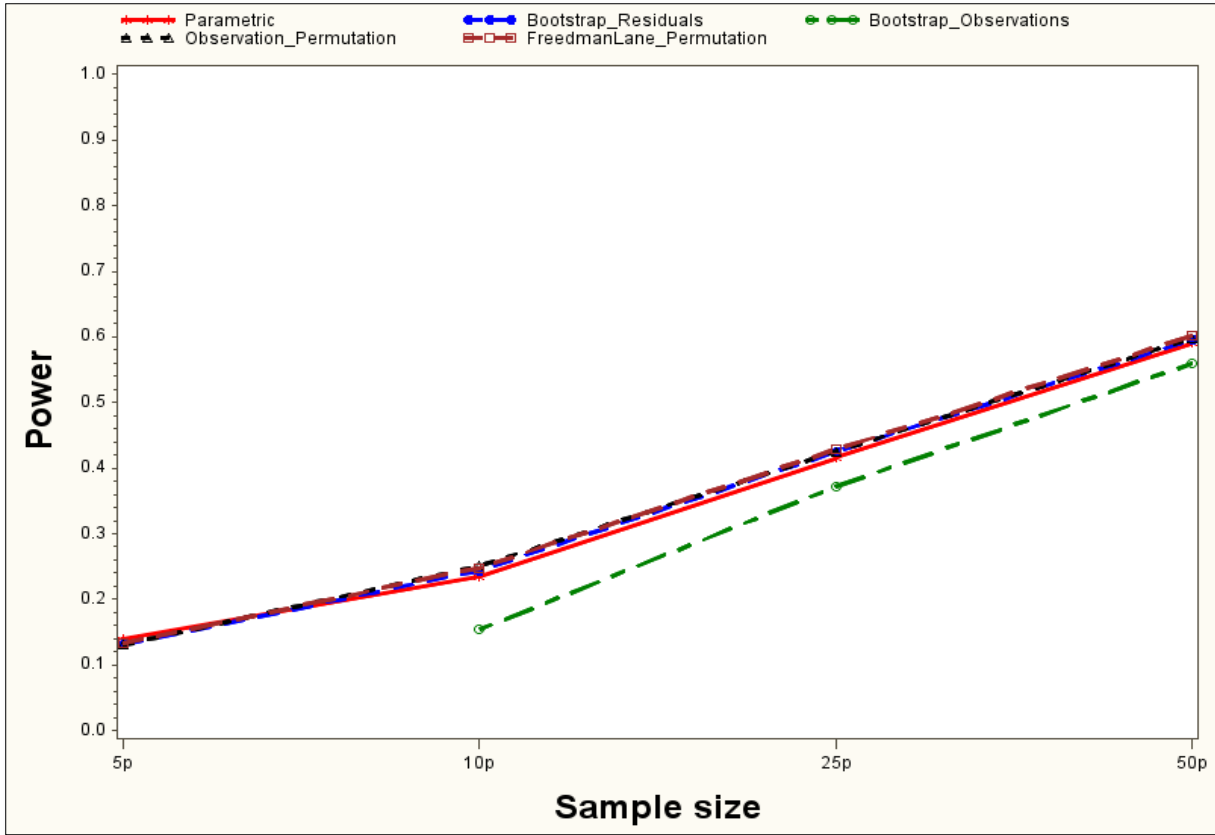


Figure 47: Power for $p = 1$ and $\varepsilon \sim \text{LOGN}(0, 1)$ on wider vertical axis

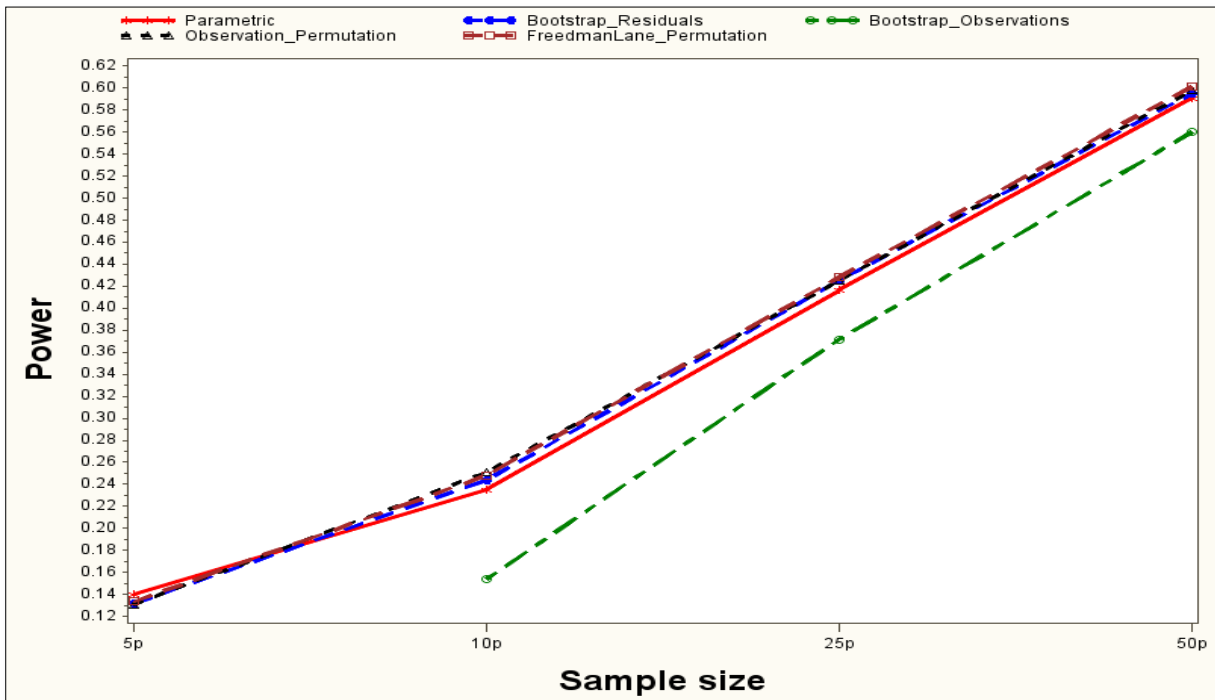


Figure 48: Power for $p = 1$ and $\varepsilon \sim \text{LOGN}(0, 1)$ on narrow vertical axis

Power results for $p=2$ Normal $\rho=0.0$

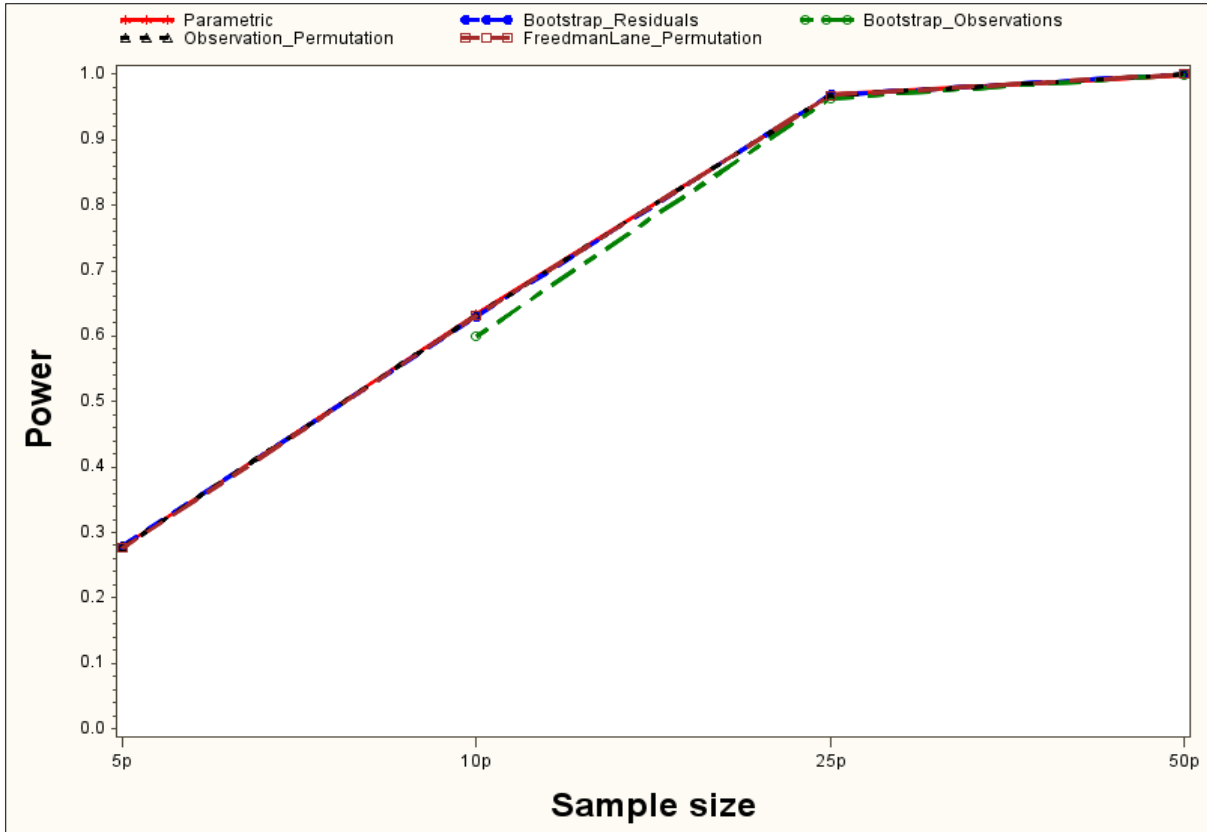


Figure 49: Power for $p = 2$, $\varepsilon \sim Normal(0, 1)$ and $\rho = 0$ on wider vertical axis

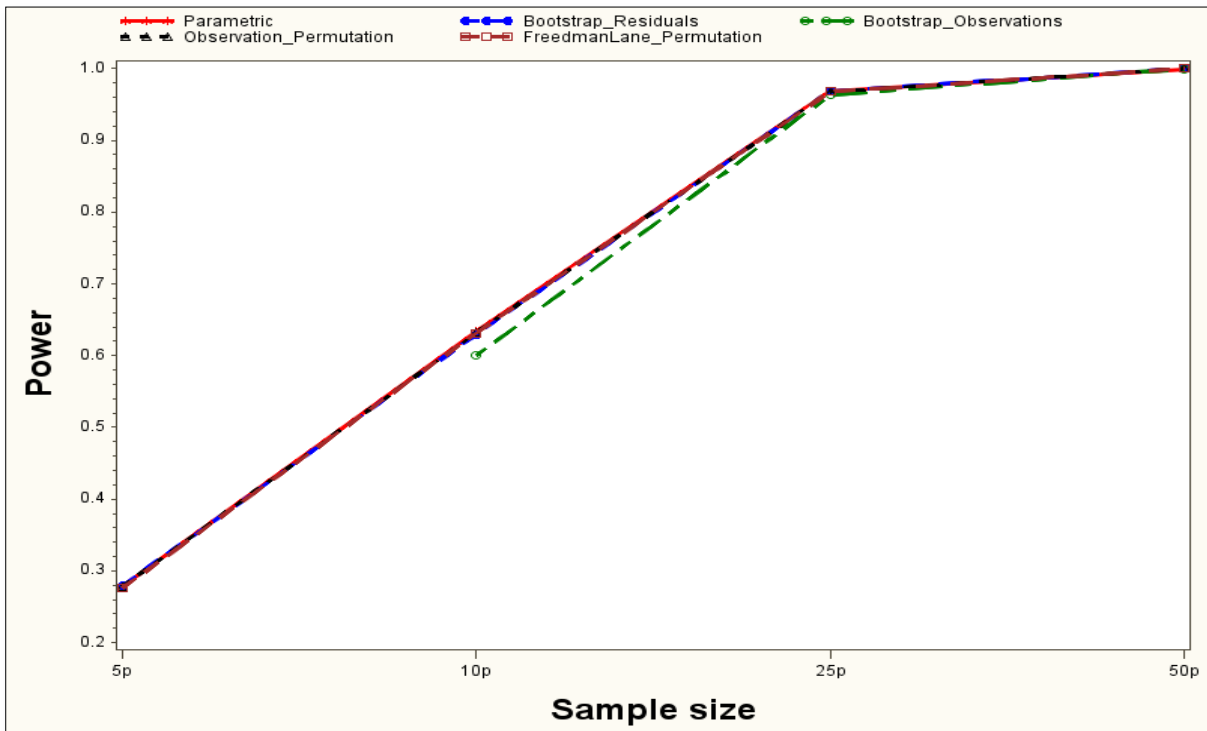


Figure 50: Power for $p = 2$, $\varepsilon \sim Normal(0, 1)$ and $\rho = 0$ on narrow vertical axis

Power results for $p=2$ Uniform $\rho=0.0$

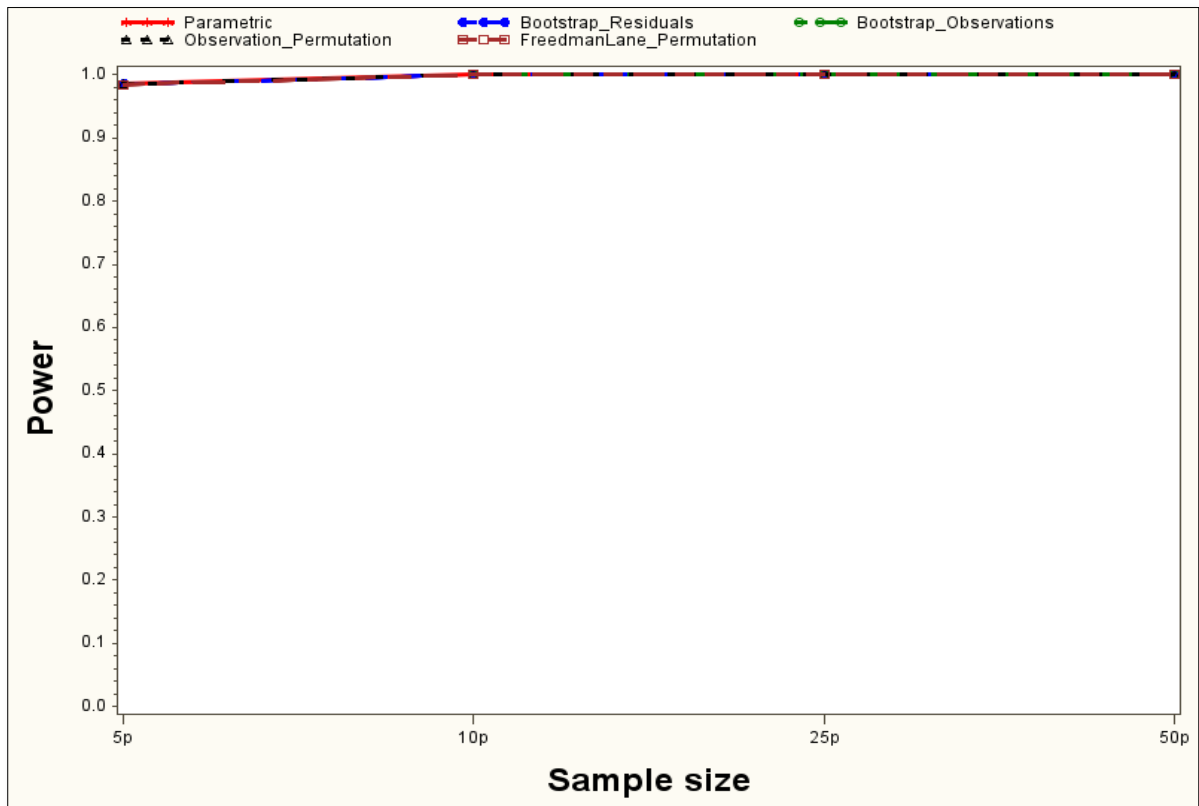


Figure 51: Power for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0$ on wider vertical axis

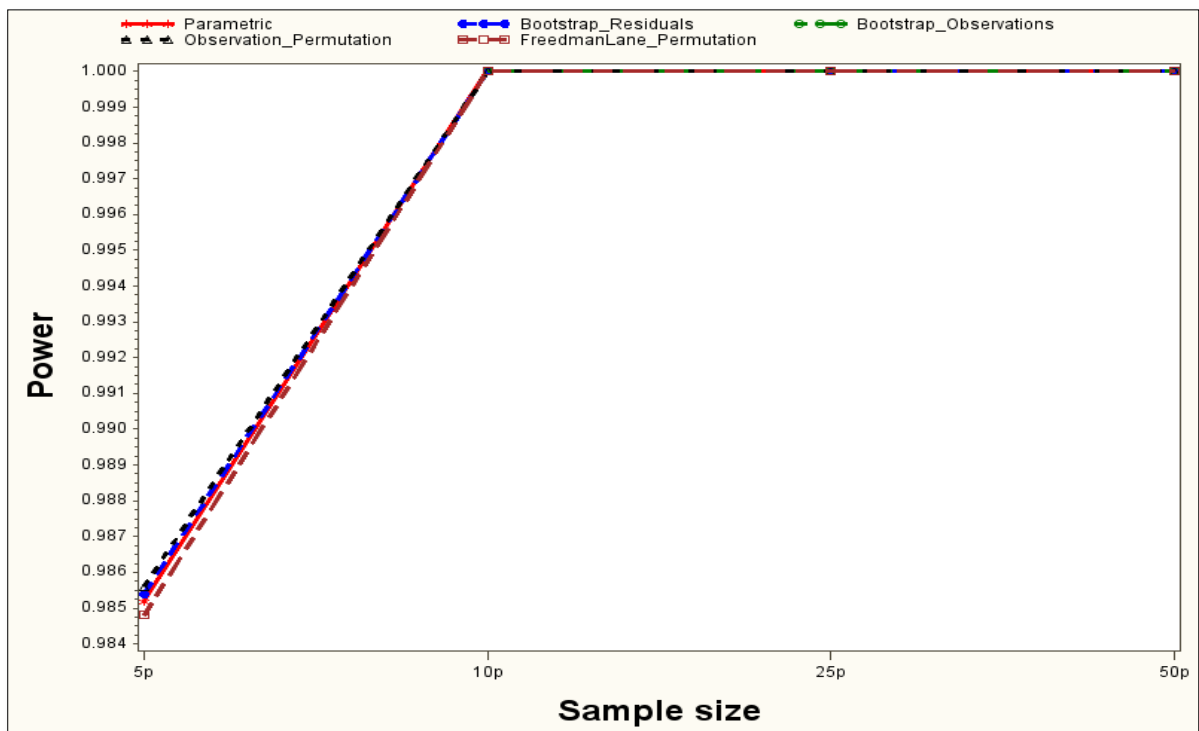


Figure 52: Power for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0$ on narrow vertical axis

Power results for $p=2$ LOGN $\rho=0.0$

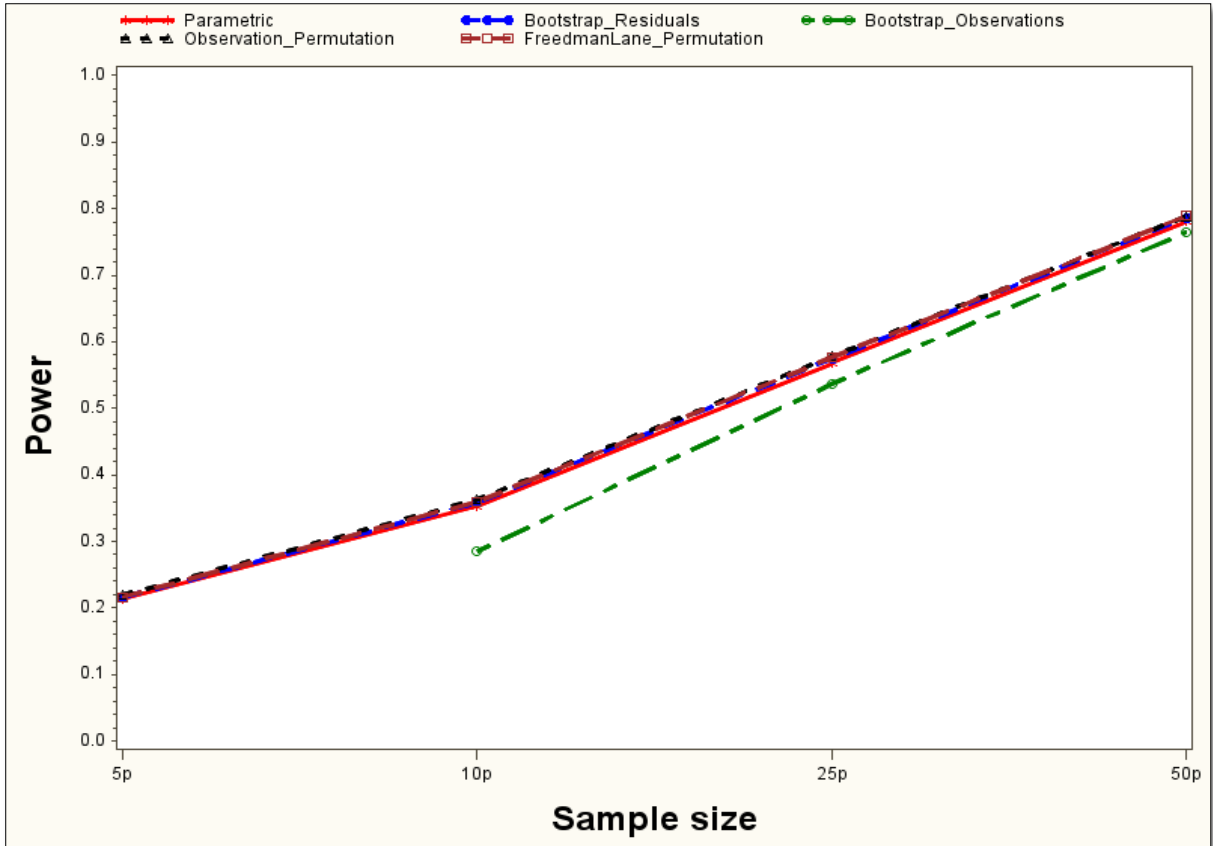


Figure 53: Power for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on wider vertical axis

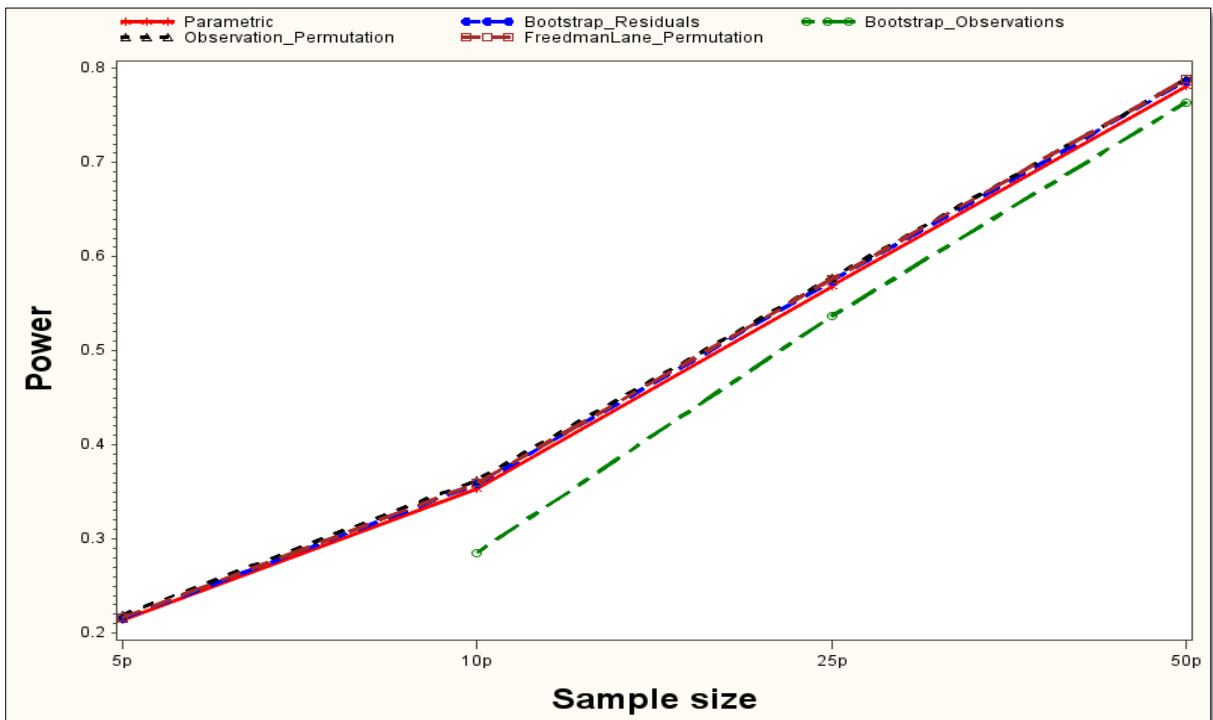


Figure 54: Power for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on narrow vertical axis

Power results for $p=5$ Normal $\rho=0.0$

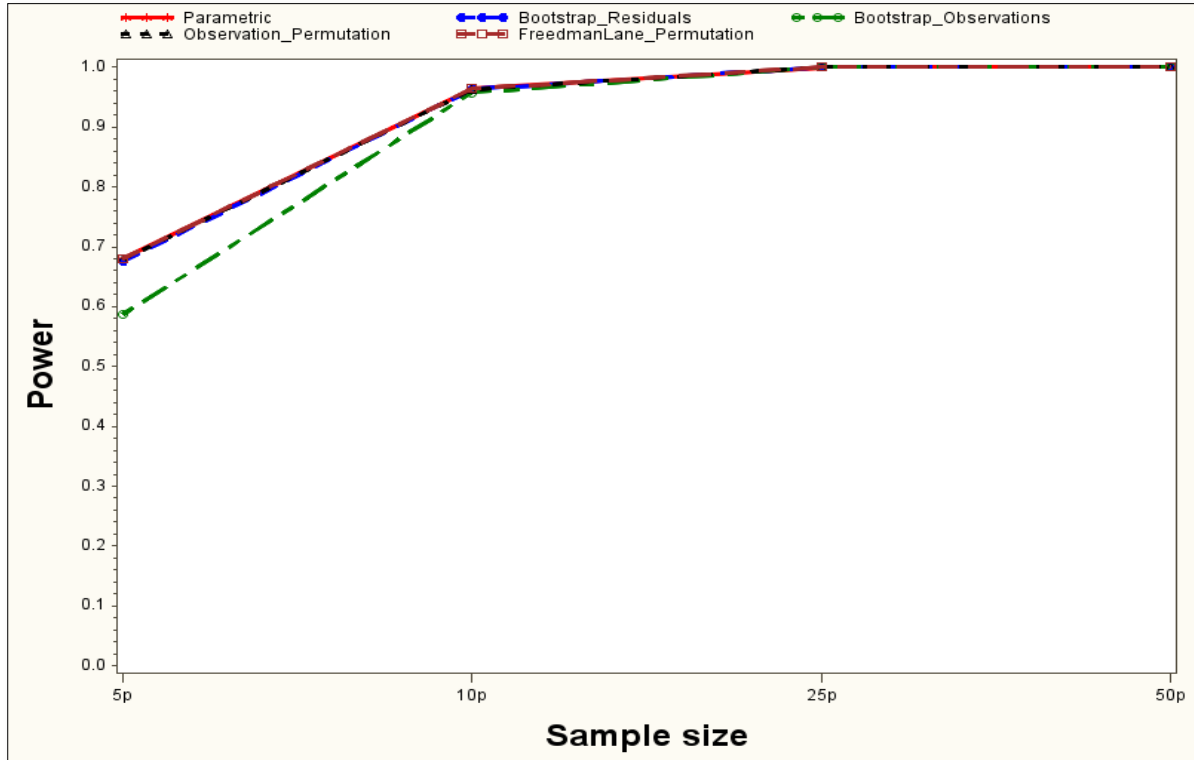


Figure 55: Power for $p = 5$, $\epsilon \sim Normal(0, 1)$ and $\rho = 0$ on wider vertical axis

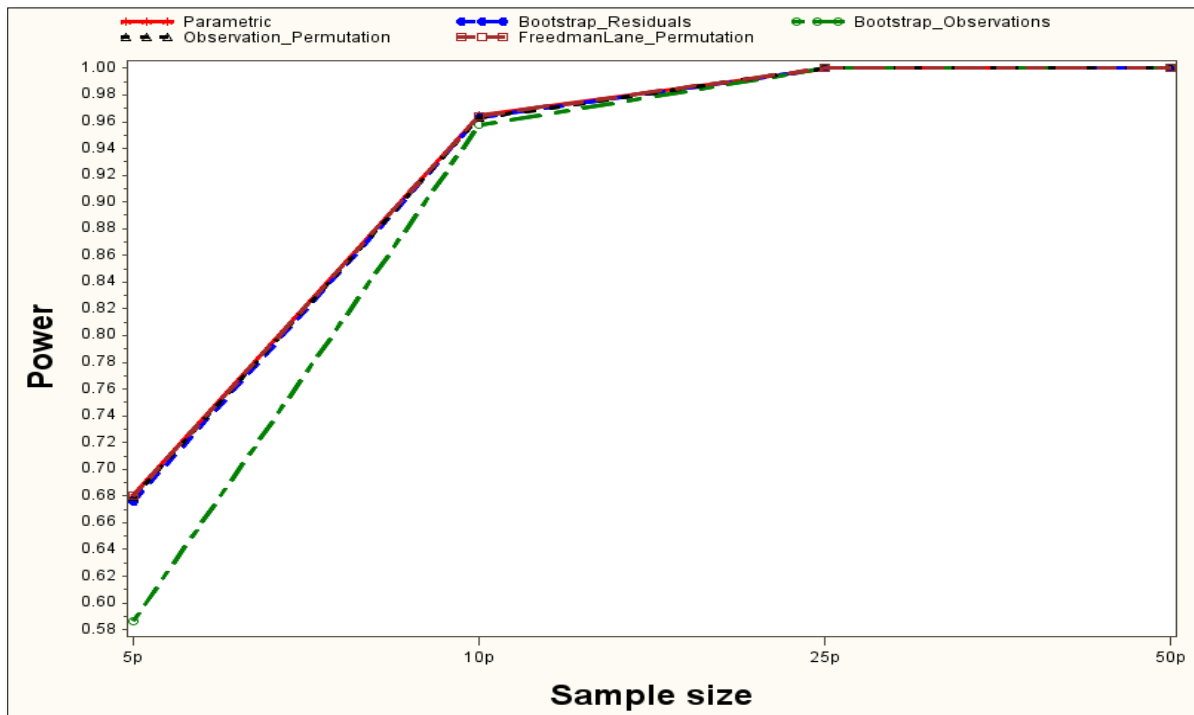


Figure 56: Power for $p = 5$, $\epsilon \sim Normal(0, 1)$ and $\rho = 0$ on narrow vertical axis

Power results for $p=5$ Uniform $\rho=0.0$

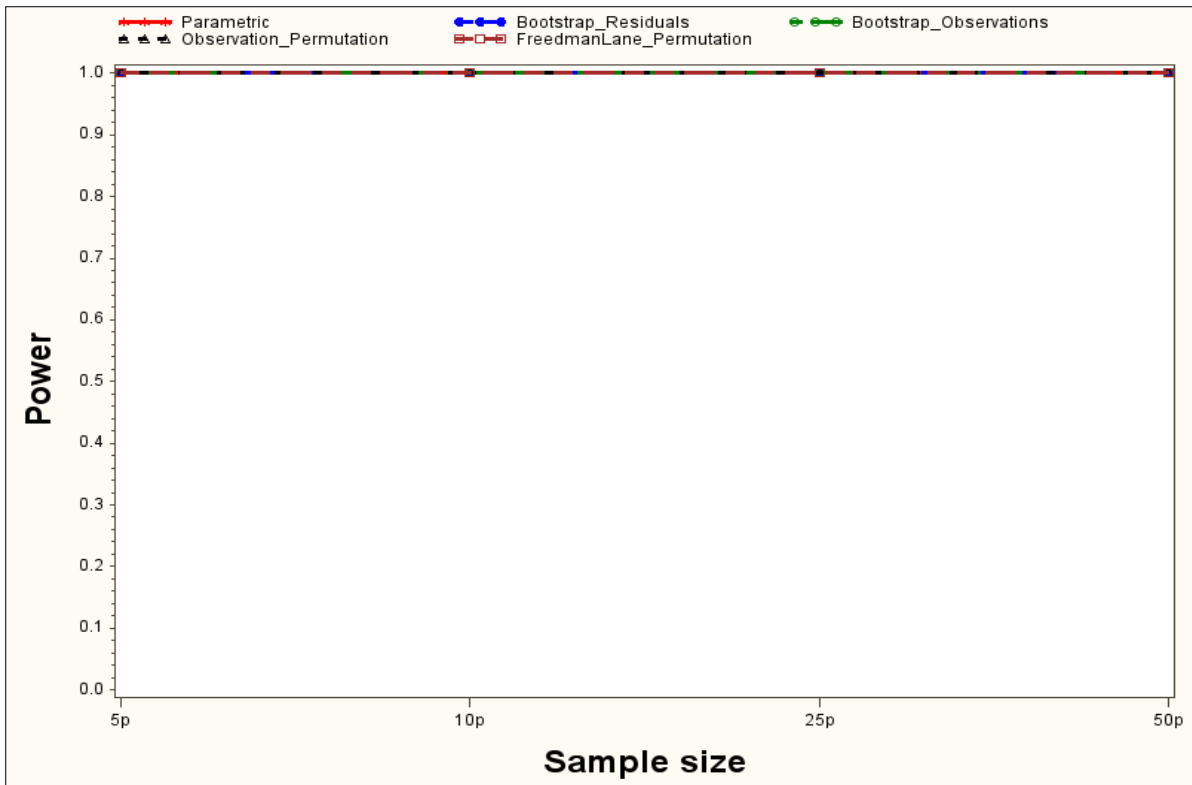


Figure 57: Power for $p = 5$, $\varepsilon \sim Uniform(0, 1)$ and $\rho = 0$ on wider vertical axis

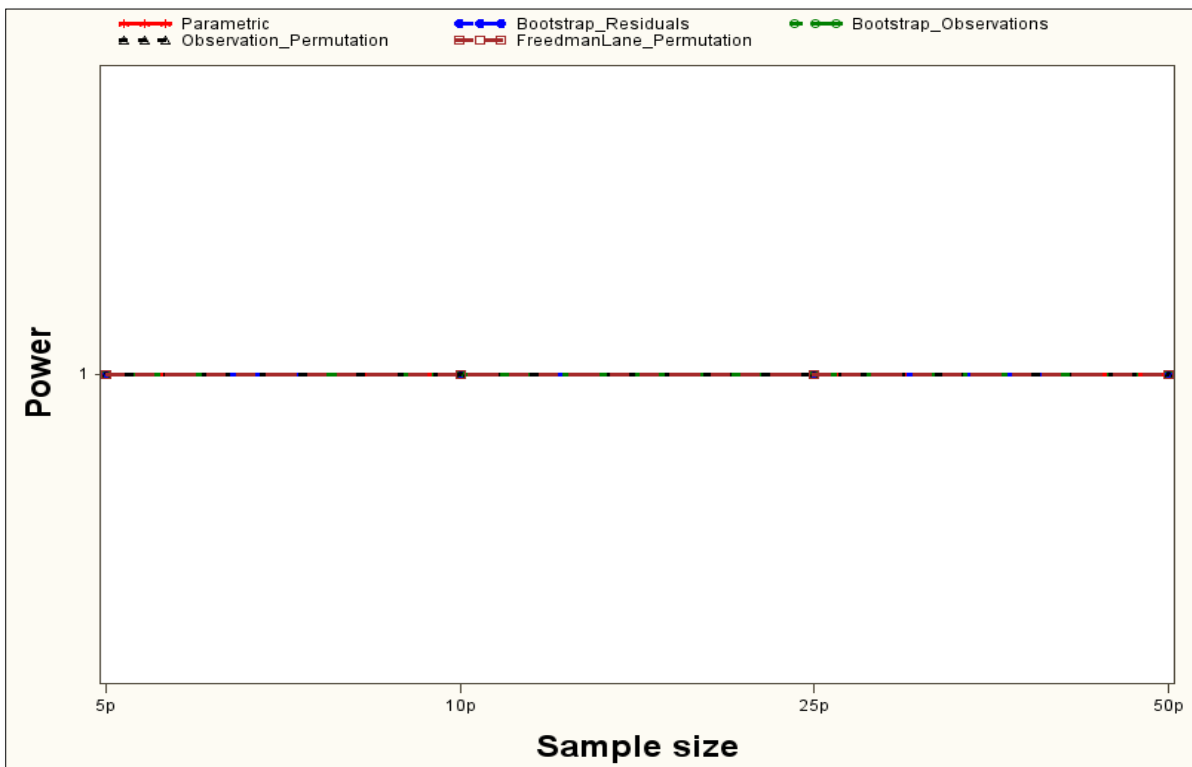


Figure 58: Power for $p = 5$, $\varepsilon \sim Uniform(0, 1)$ and $\rho = 0$ on narrow vertical axis

Power results for $p=5$ LOGN $\rho=0.0$

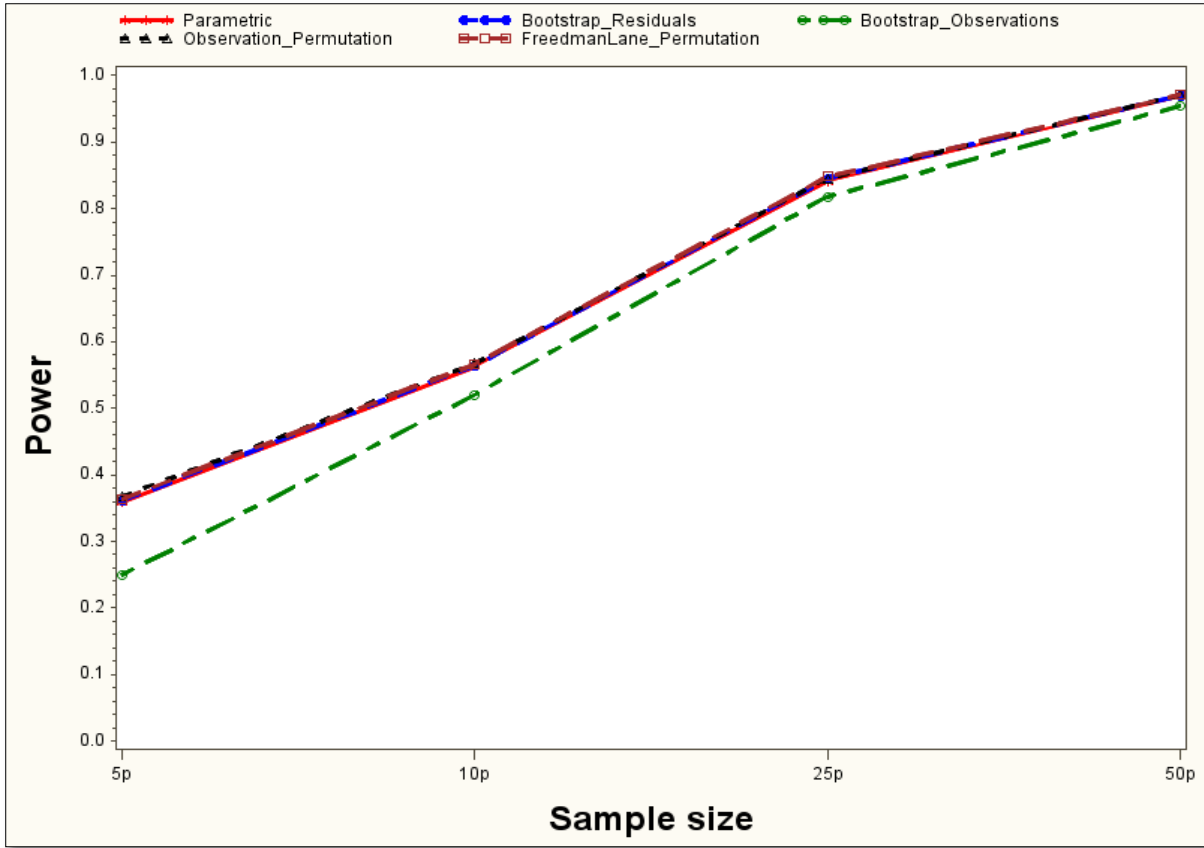


Figure 59: Power for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on wider vertical axis

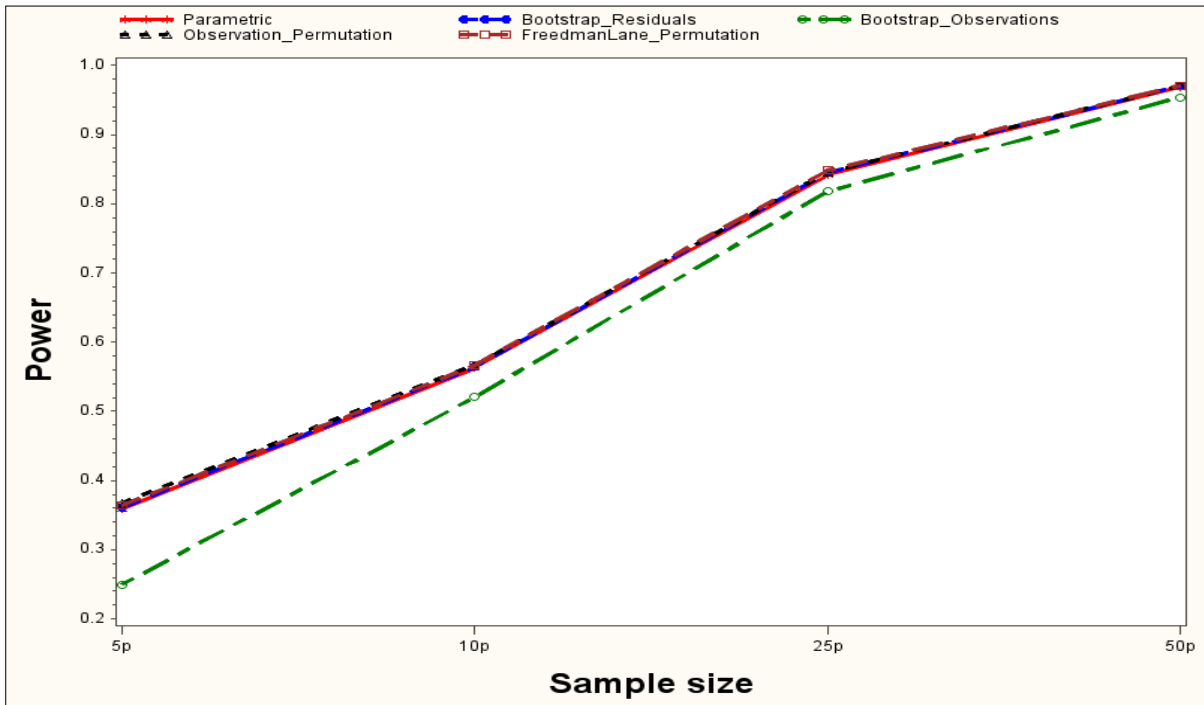


Figure 60: Power for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0$ on narrow vertical axis

There were no substantial power differences when errors are normally distributed. As could be generally expected, power increased with increase in sample size. The bootstrapping observations method is the least powerful method when the sample size is $10p$ however from the sample size $25p$ there are no visible differences for $p = 1$ and $p = 2$. For $p = 5$ the bootstrapping observations method is least powerful at sample size $5p$.

For uniformly distributed errors there are differences only for smaller sample sizes and from the sample size of $25p$ and $10p$ for $p = 1$ and $p = 2$ respectively the maximum power of approximately 1 was attained. For $p = 1$ at sample size $5p$ the parametric method is the best while permuting observations is the worst. Bootstrapping observations is slightly less powerful at sample size $10p$. For $p = 2$ and $p = 5$ all methods are equally powerful.

All the methods are almost equally powerful when errors have a lognormal distribution with an exception of bootstrapping observations which is consistently less powerful for all cases $p = 1, 2$ and 5 . It can be seen that the parametric method is slightly less powerful than the other methods excluding the bootstrapping observations method.

It can be noted that for lognormal errors all methods are less powerful compared to normal and uniform errors. The largest attainable power when $p = 1$ in case of lognormally distributed errors is approximately 0.6. A power of 0.97 is only attainable when $p = 5$ under lognormally distributed errors and this power is attained for $p = 1$ in case of normally distributed errors.

6.3.2 Power in the presence of multicollinearity

Power results for $p=2$ Normal $\rho=0.50$

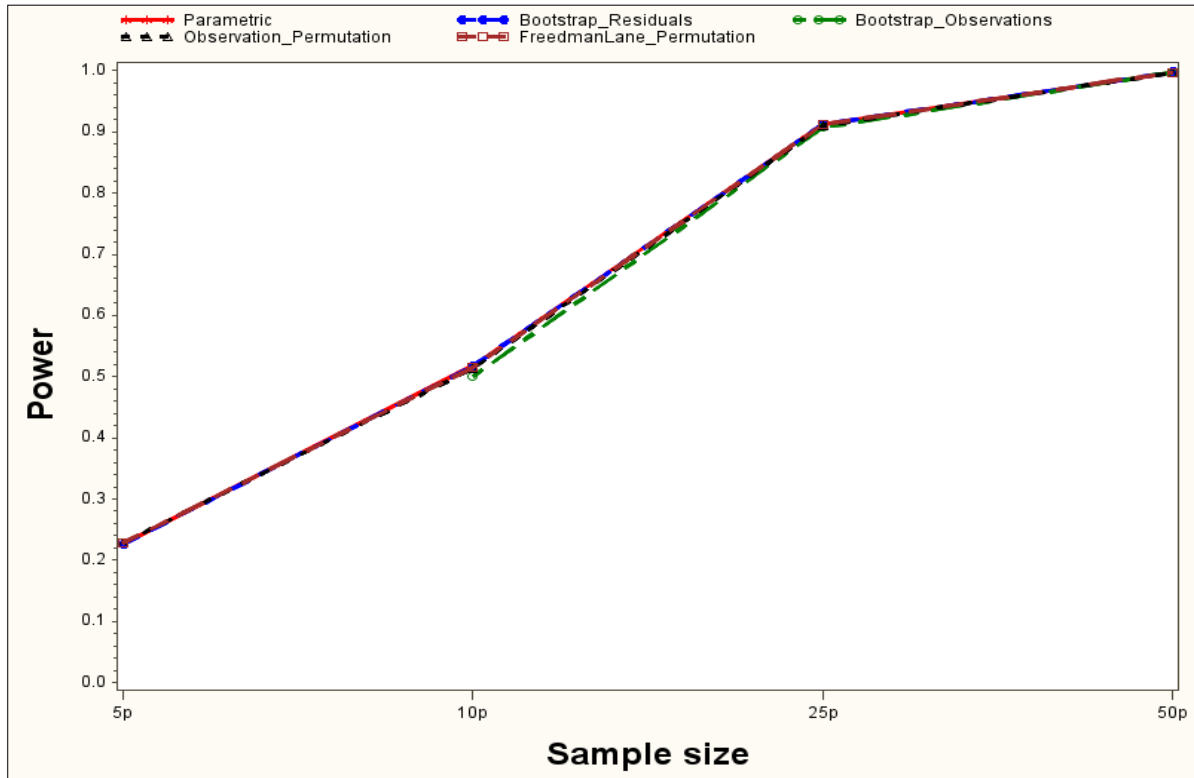


Figure 61: Power for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.5$ on wider vertical axis

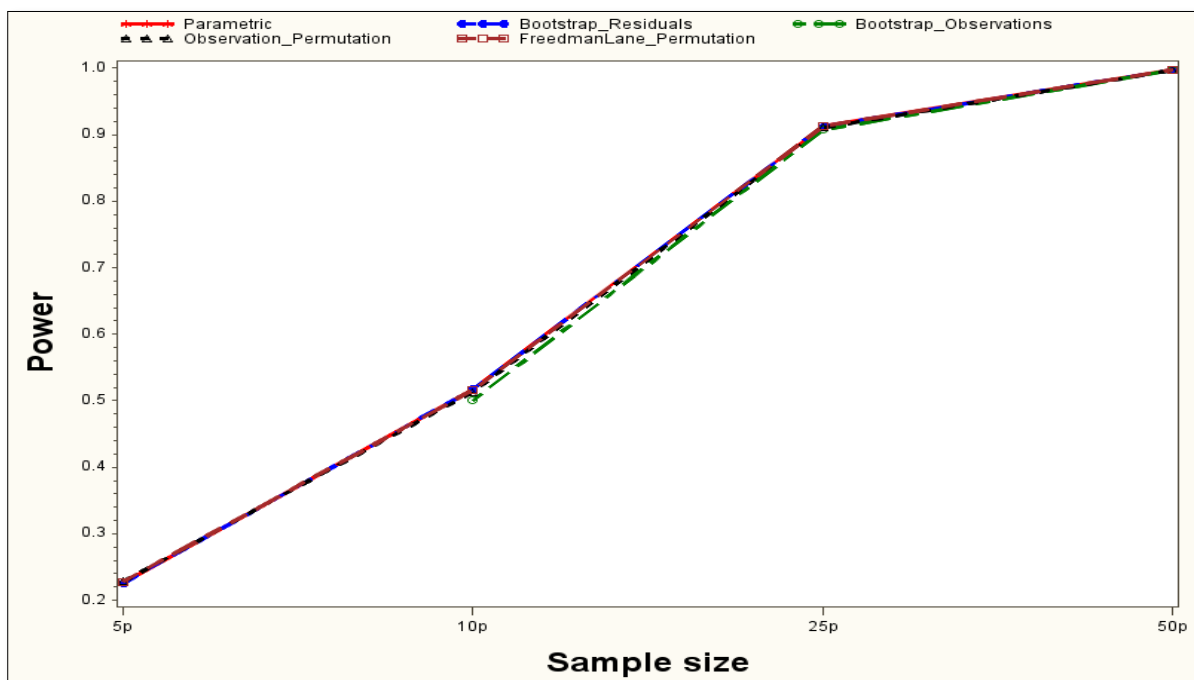


Figure 62: Power for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis

Power results for $p=2$ Uniform $\rho=0.50$

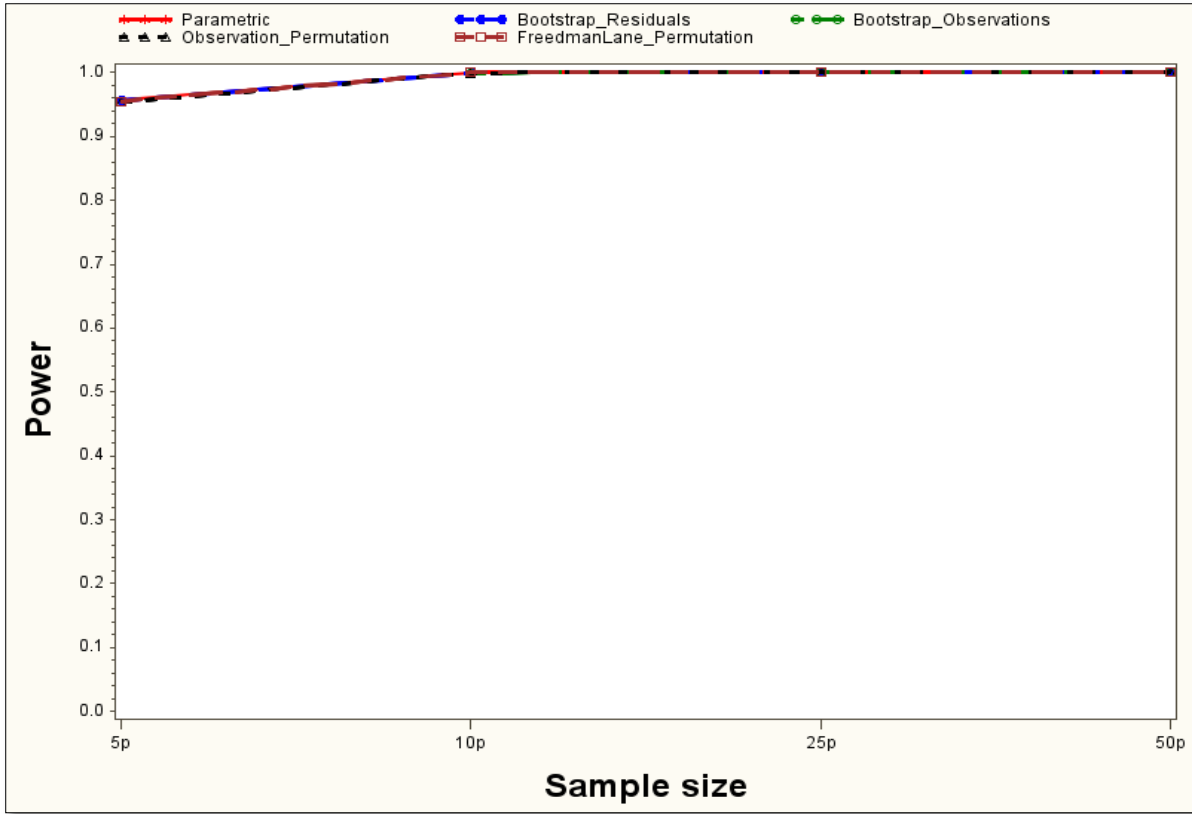


Figure 63: Power for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on wider vertical axis

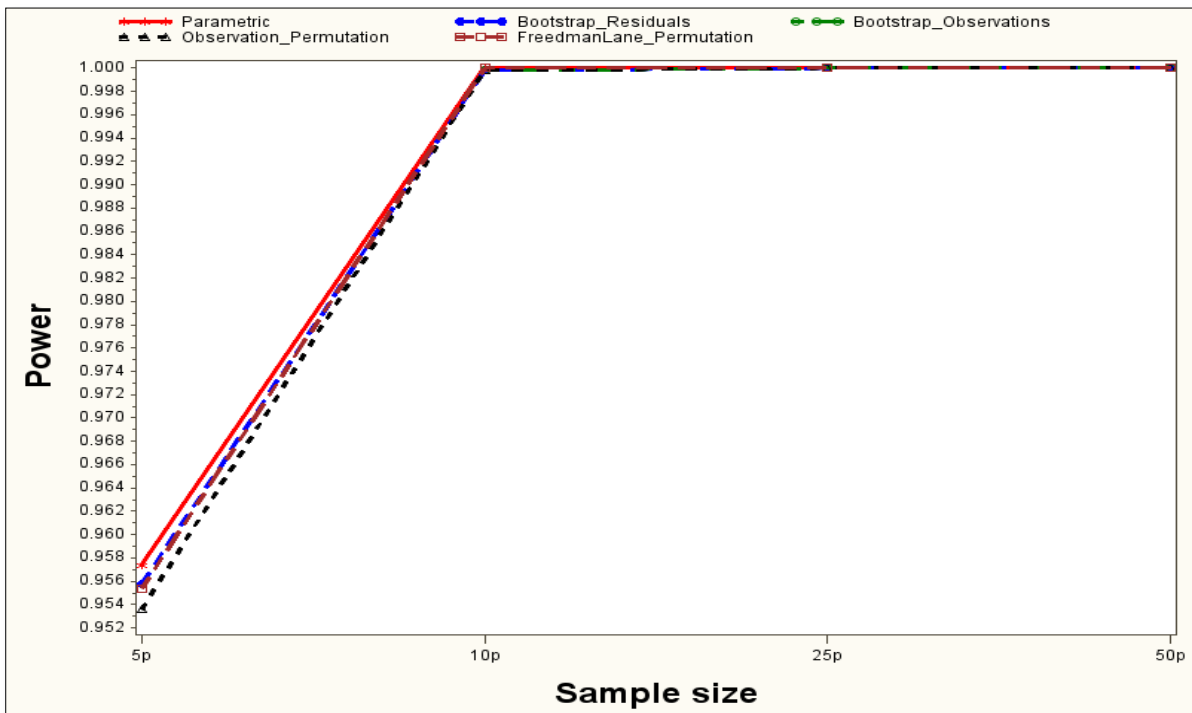


Figure 64: Power for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis

Power results for $p=2$ LOGN $\rho=0.50$

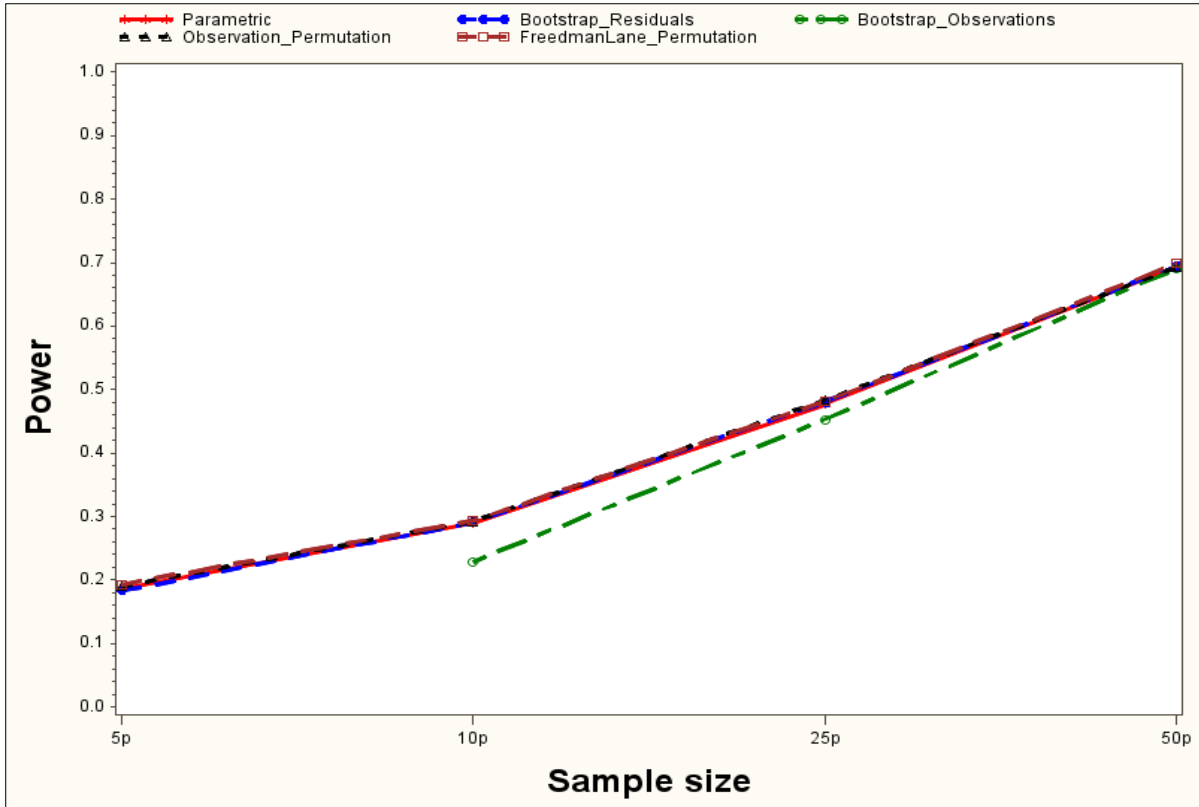


Figure 65: Power for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on wider vertical axis

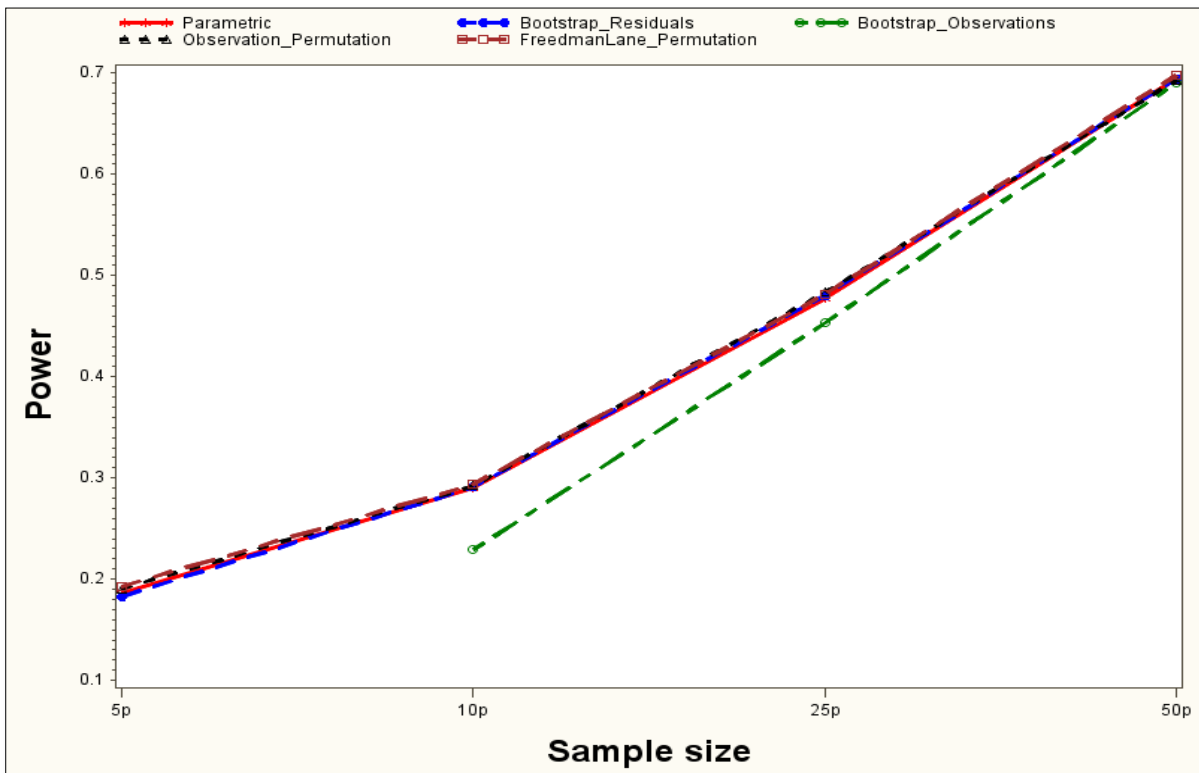


Figure 66: Power for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis

Power results for $p=5$ Normal $\rho=0.50$

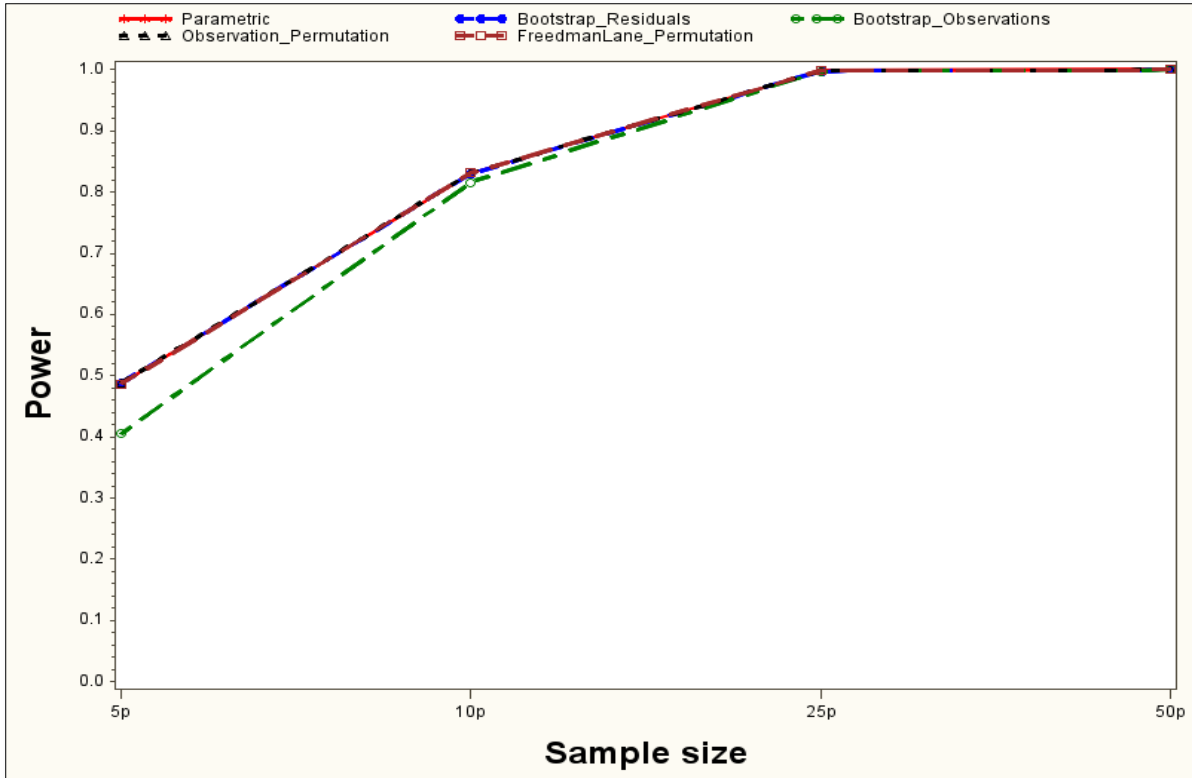


Figure 67: Power for $p = 5$, $\varepsilon \sim Normal(0, 1)$ and $\rho = 0.5$ on wider vertical axis

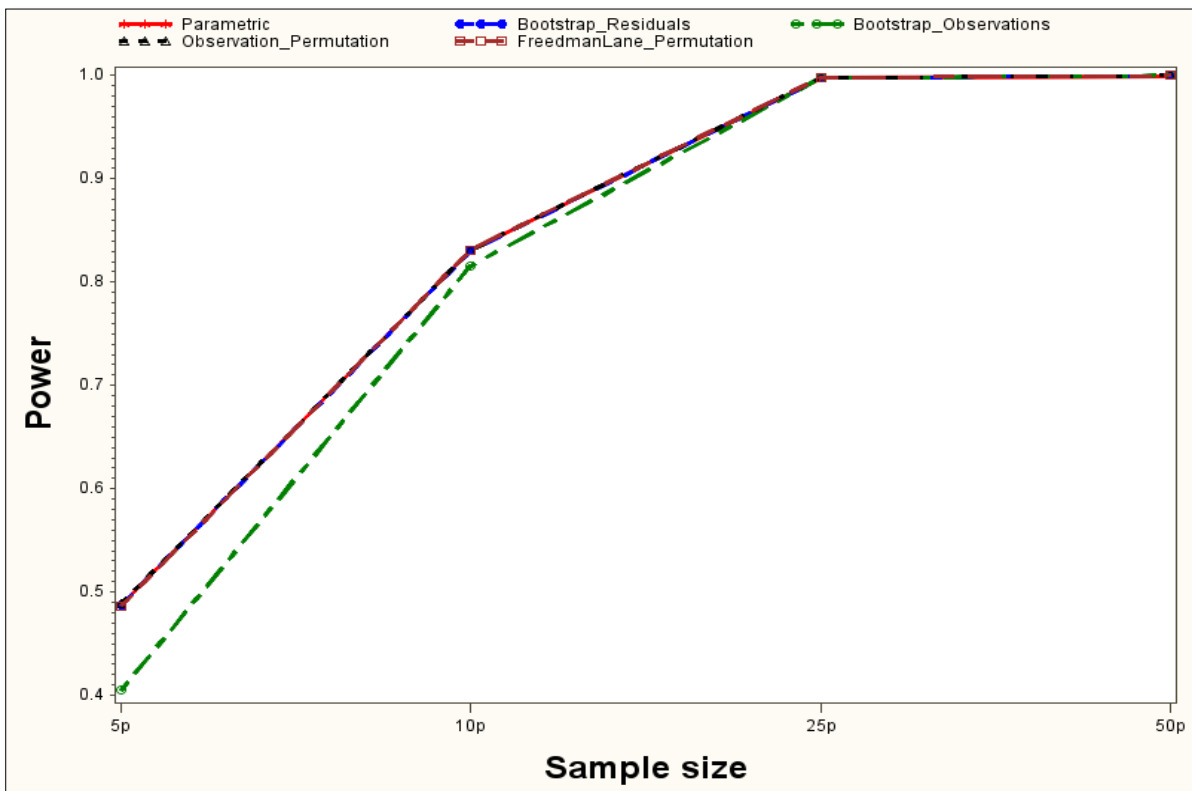


Figure 68: Power for $p = 5$, $\varepsilon \sim Normal(0, 1)$ and $\rho = 0.5$ on narrow vertical axis

Power results for $p=5$ Uniform $\rho=0.50$

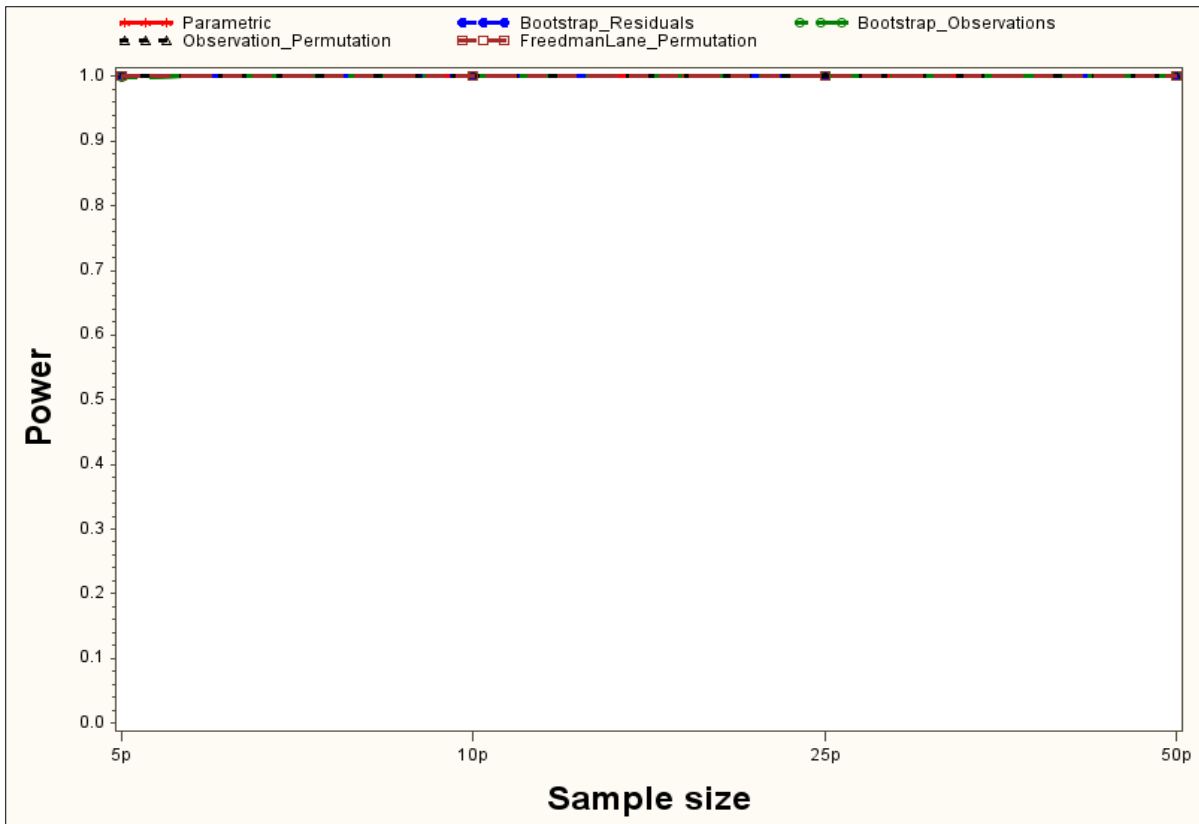


Figure 69: Power for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on wider vertical axis

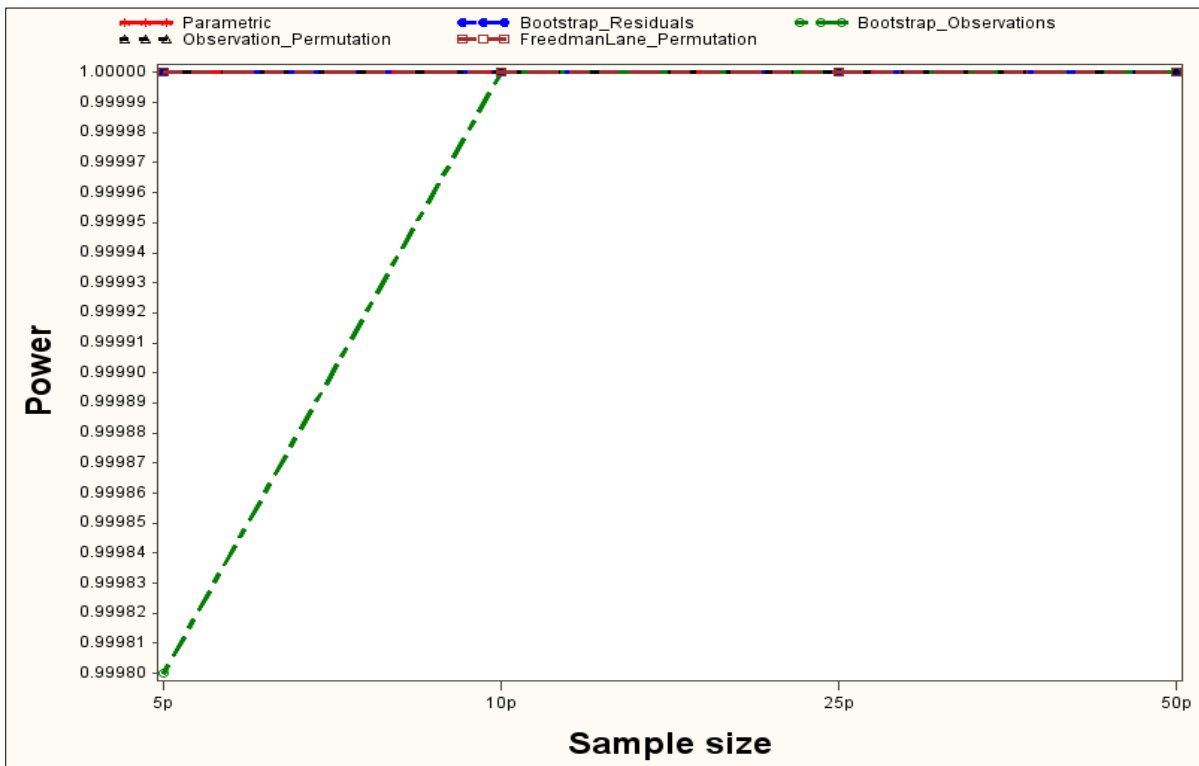


Figure 70: Power for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis

Power results for $p=5$ LOGN $\rho=0.50$

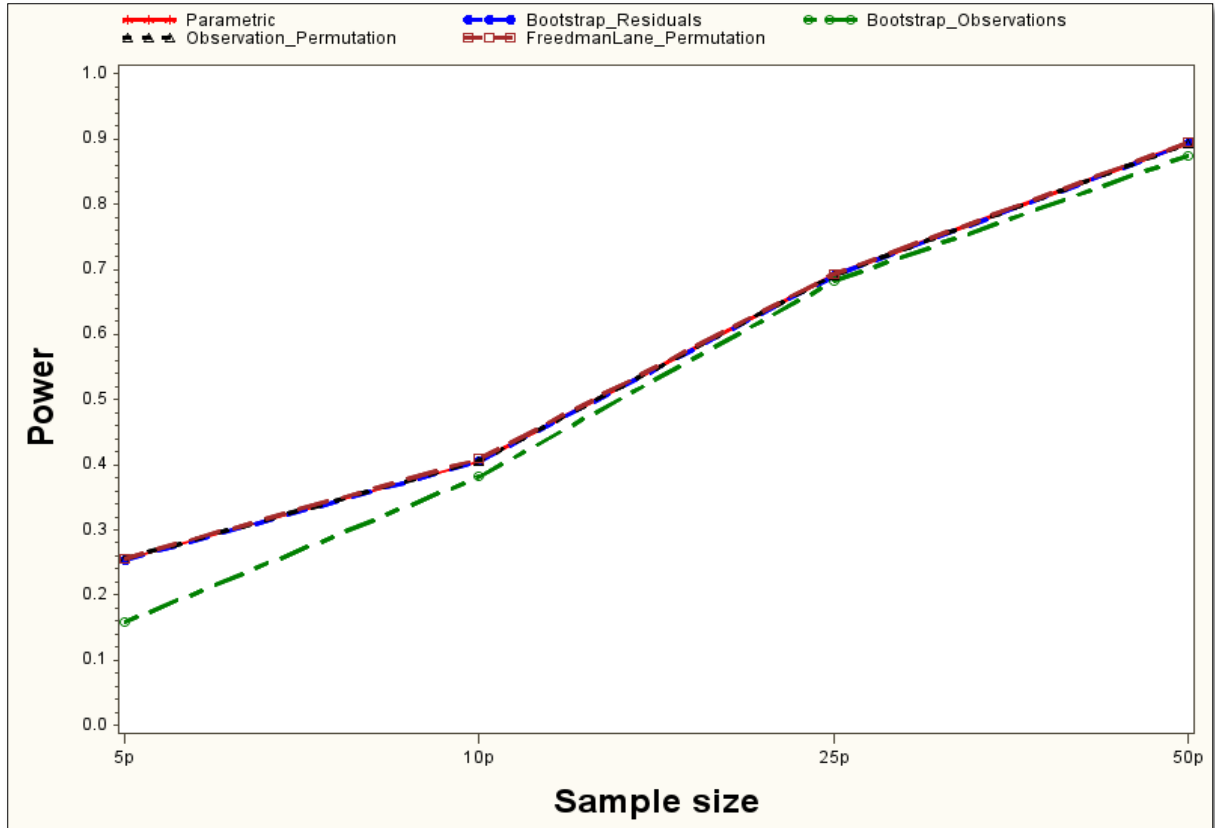


Figure 71: Power for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on wider vertical axis

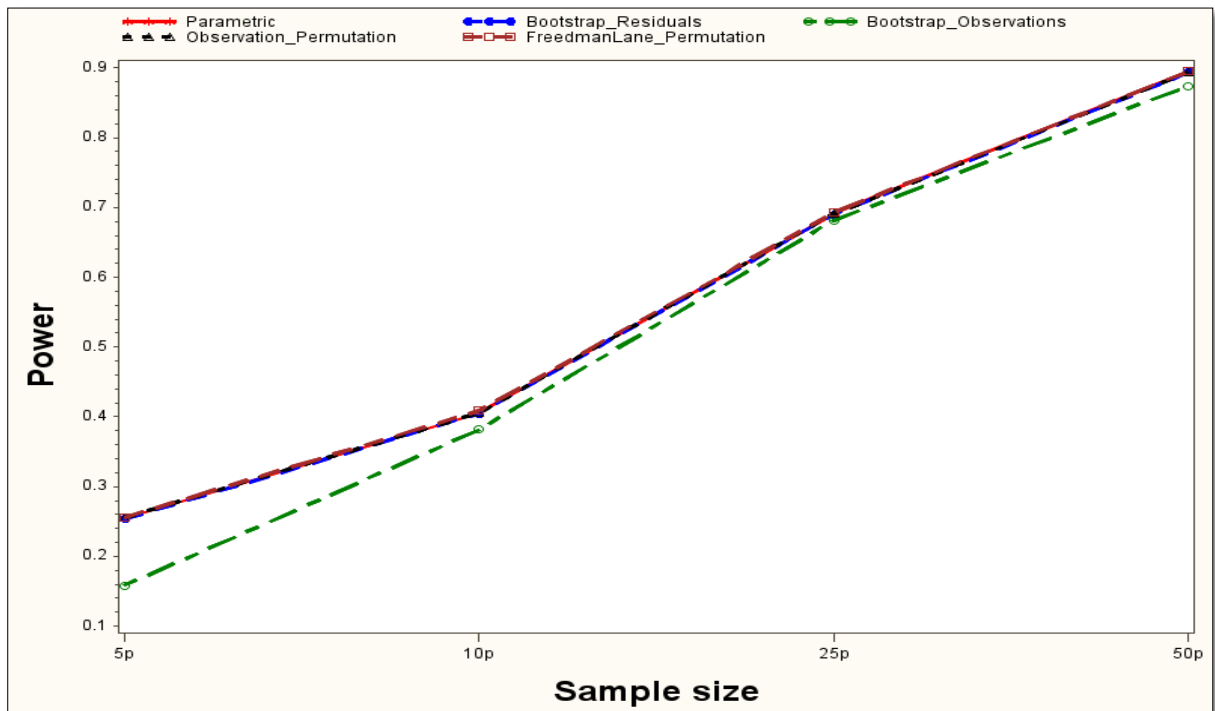


Figure 72: Power for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.5$ on narrow vertical axis

The impact of an increase in collinearity among explanatory variables was the slight decrease in power of all methods under consideration. For normally distributed errors there are no significant differences in power across all methods. For $p = 2$ the impact of the introduction of multicollinearity on the minimum and maximum power is a decrease in minimum power from 0.2758 to 0.226 and maximum power from 1 to 0.9978 across all methods. For $p = 5$ a similar observation was made only for the minimum power where it decreased from 0.5866 to 0.4048 while the maximum was not affected. However while the maximum of 1 was reached at sample size $25p$ under independent explanatory variables, the same power was only attained at sample size $50p$.

For uniformly distributed errors the small decrease brought by the introduction of multicollinearity can be observed under the case where $p = 2$ on small sample sizes only. In the case of $p = 2$ the worse power attained for sample size $5p$ is 0.9536 compared to 0.9848 attained in the absence of multicollinearity. For $p = 5$ the power remains unaffected by an introduction of multicollinearity. There is no one method that outright outperforms others.

The power possessed by different methods tends to suffer when errors are lognormally distributed but the decrease in power ranges between 11% and 13% for same sample sizes. The maximum power attained for cases $p = 2$ and $p = 5$ are 0.69 and 0.87 respectively compared to the maximum power of 0.76 and 0.95. Bootstrapping observations methods is generally the least powerful of all methods.

6.3.3 Power in the presence of increased multicollinearity

Power results for $p=2$ Normal $\rho=0.90$

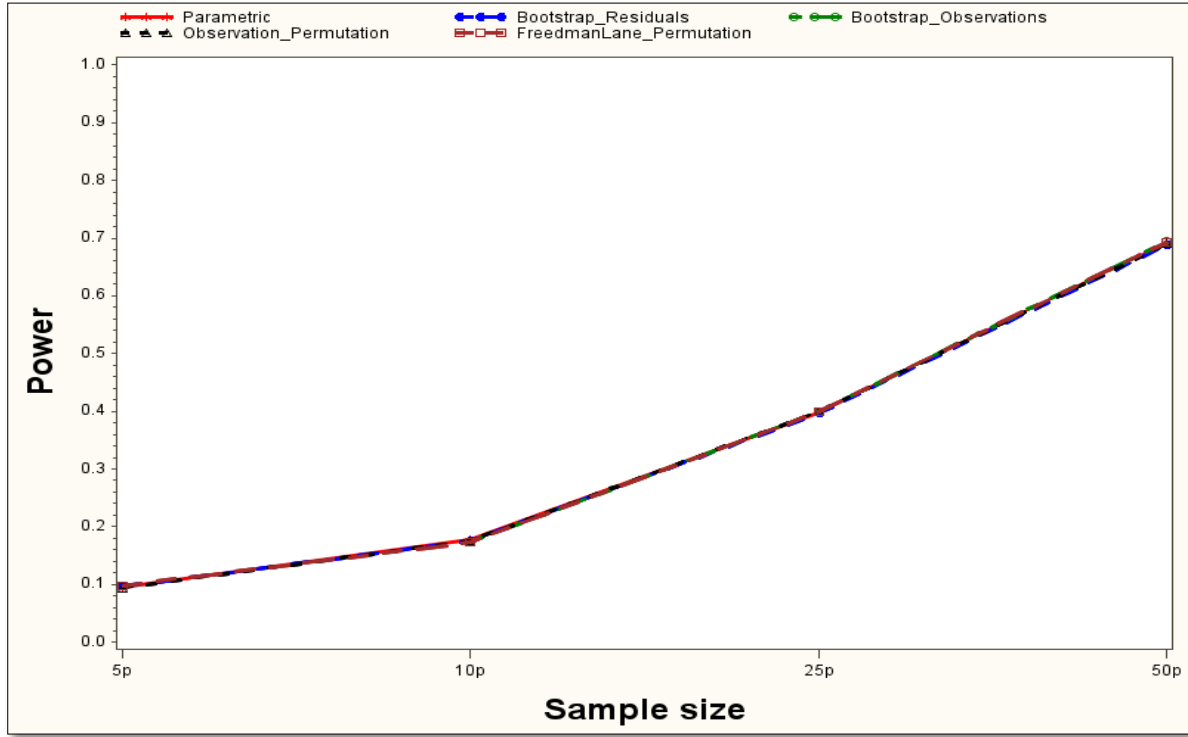


Figure 73: Power for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.9$ on wider vertical axis

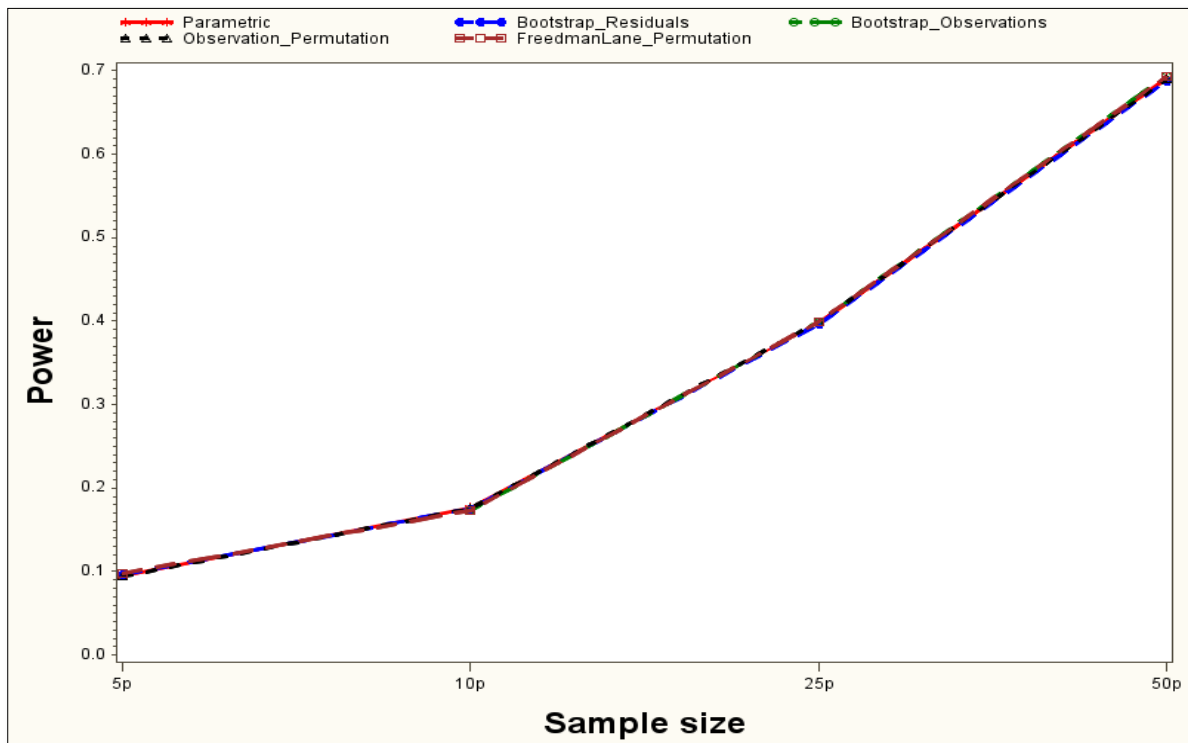


Figure 74: Power for $p = 2$, $\varepsilon \sim \text{Normal}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis

Power results for $p=2$ Uniform $\rho=0.90$

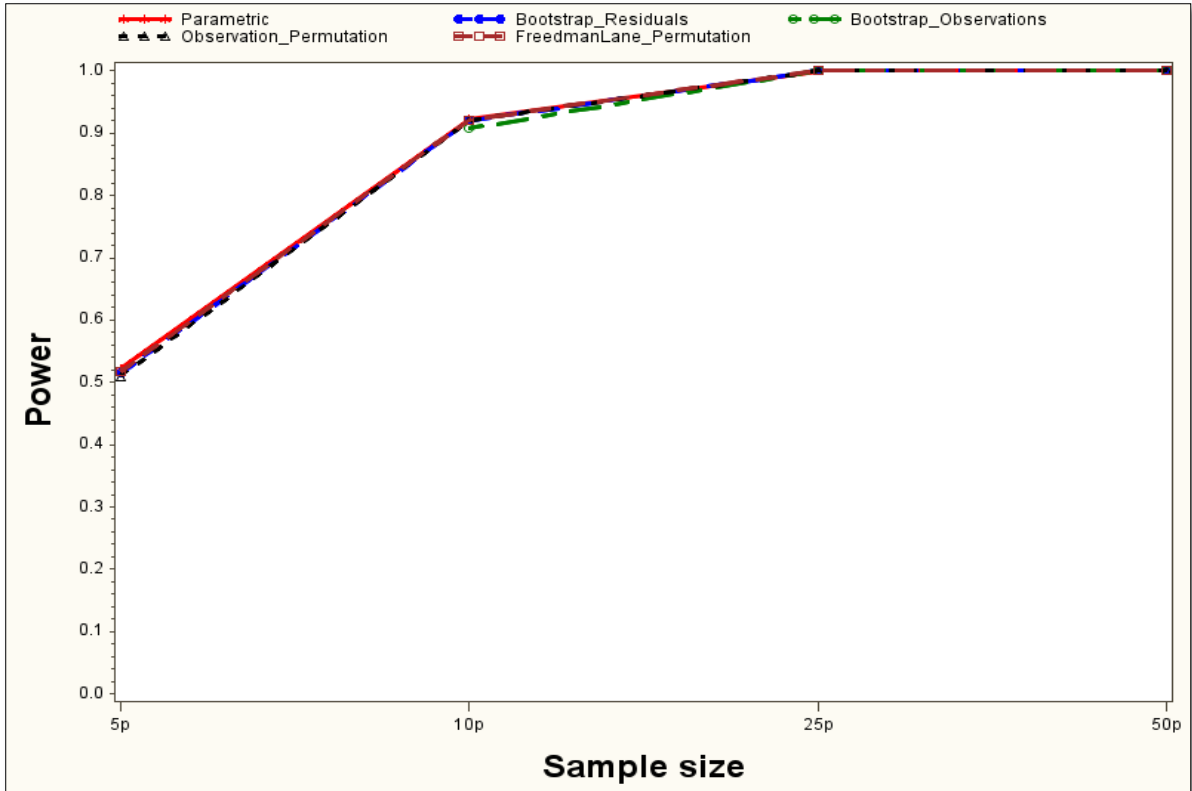


Figure 75: Power for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on wider vertical axis

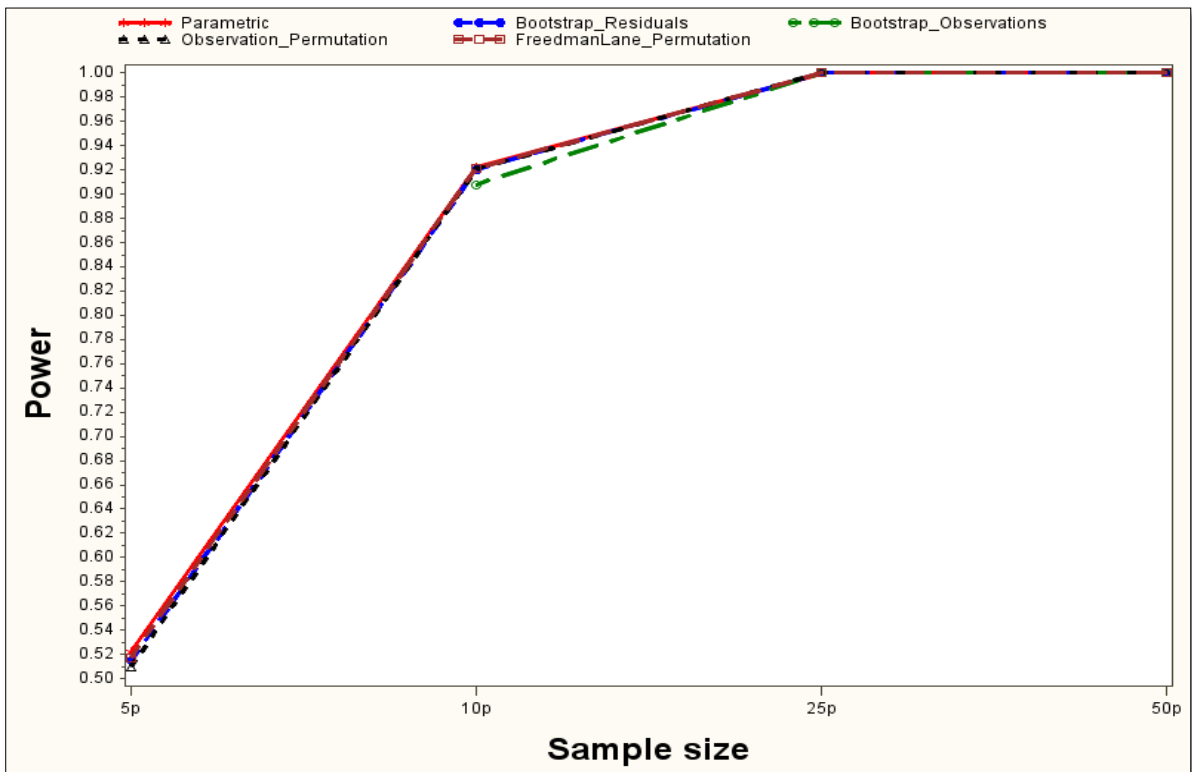


Figure 76: Power for $p = 2$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis

Power results for $p=2$ LOGN $\rho=0.90$

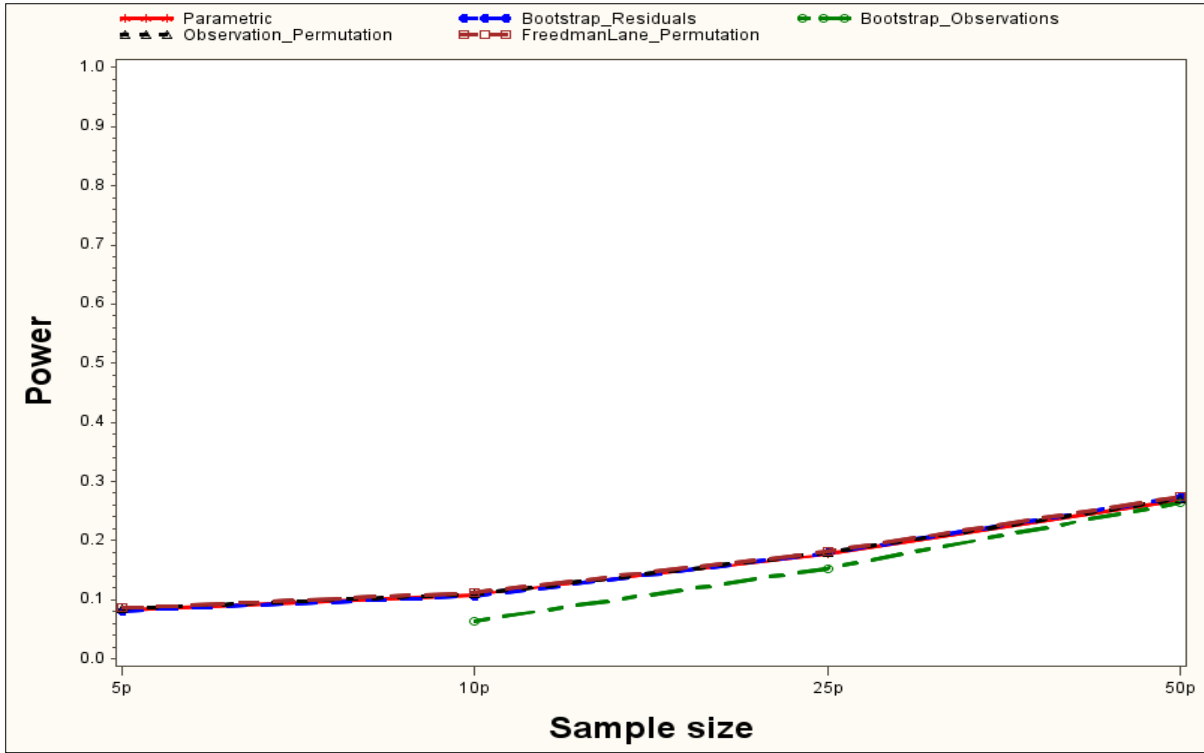


Figure 77: Power for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on wider vertical axis

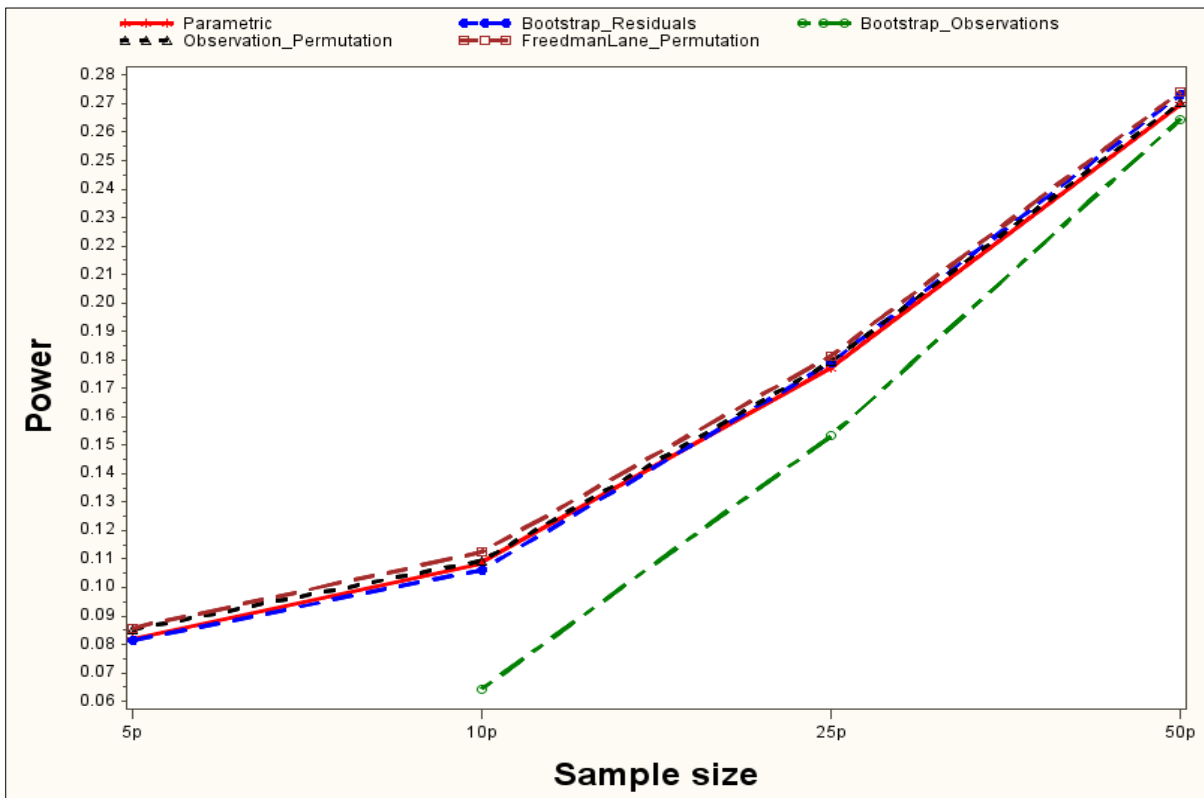


Figure 78: Power for $p = 2$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis

Power results for $p=5$ Normal $\rho=0.90$

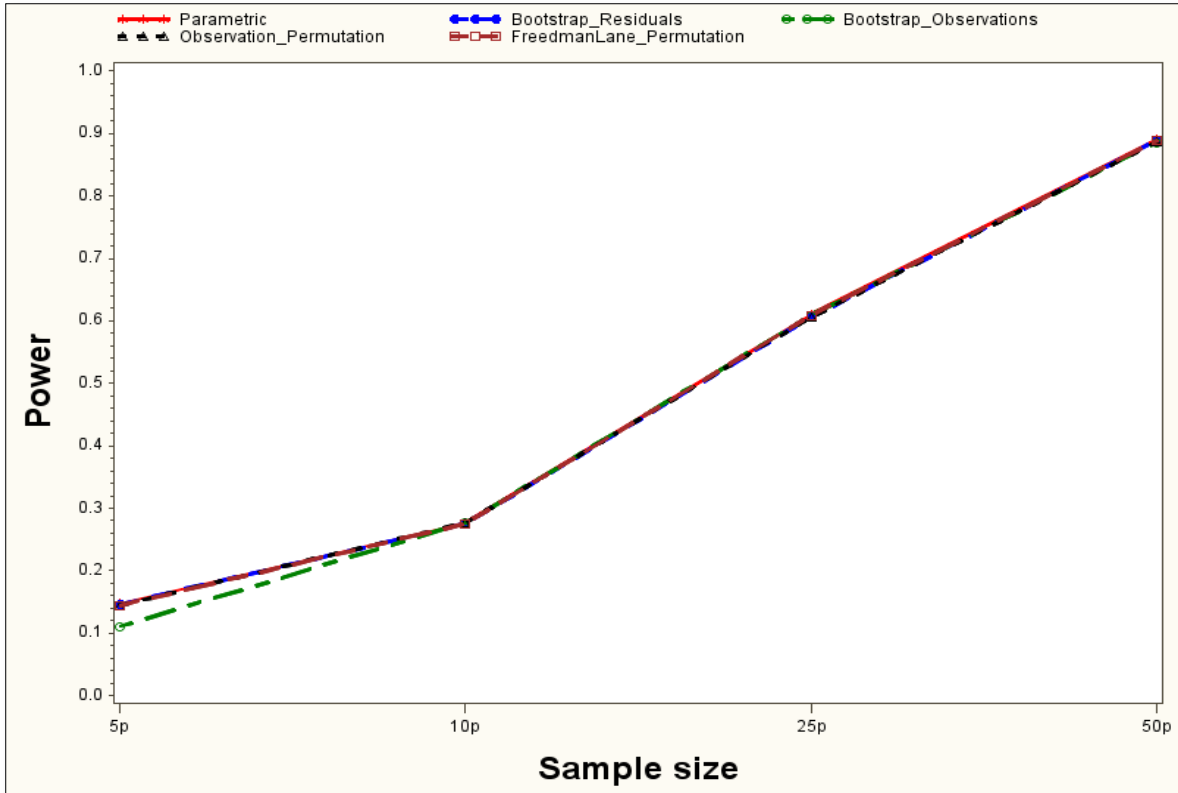


Figure 79: Power for $p = 5$, $\varepsilon \sim Normal(0, 1)$ and $\rho = 0.9$ on wider vertical axis

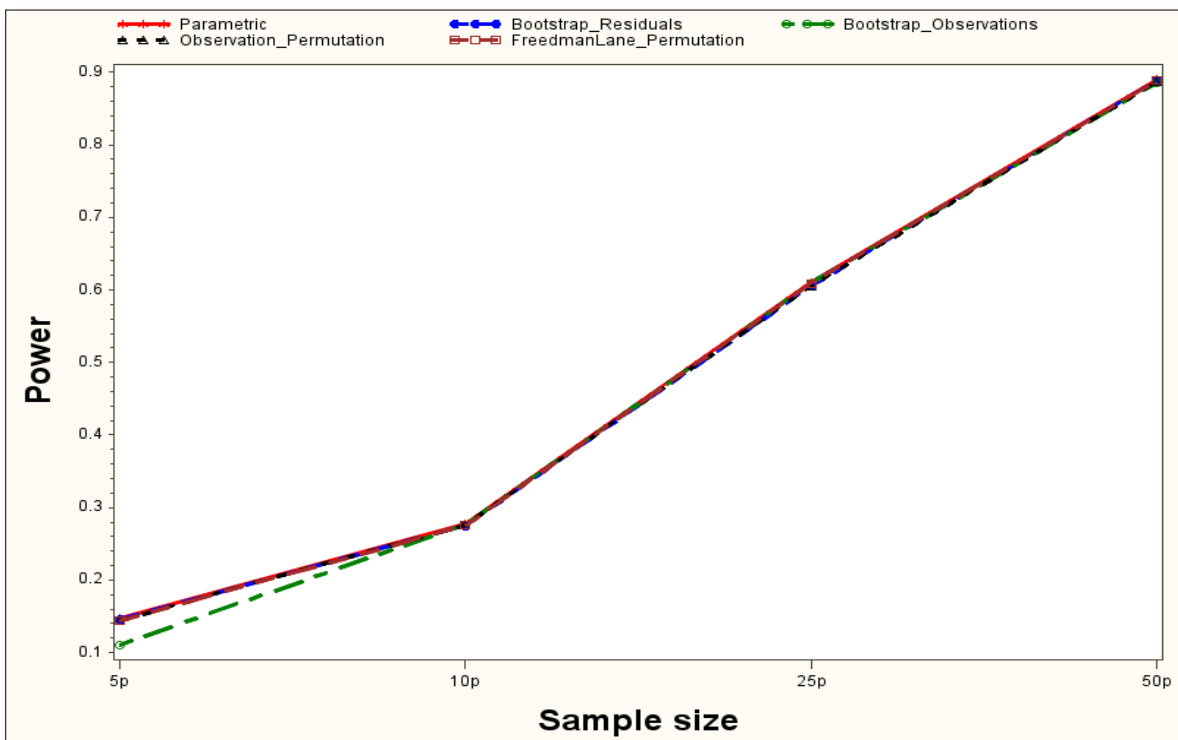


Figure 80: Power for $p = 5$, $\varepsilon \sim Normal(0, 1)$ and $\rho = 0.9$ on narrow vertical axis

Power results for $p=5$ Uniform $\rho=0.90$

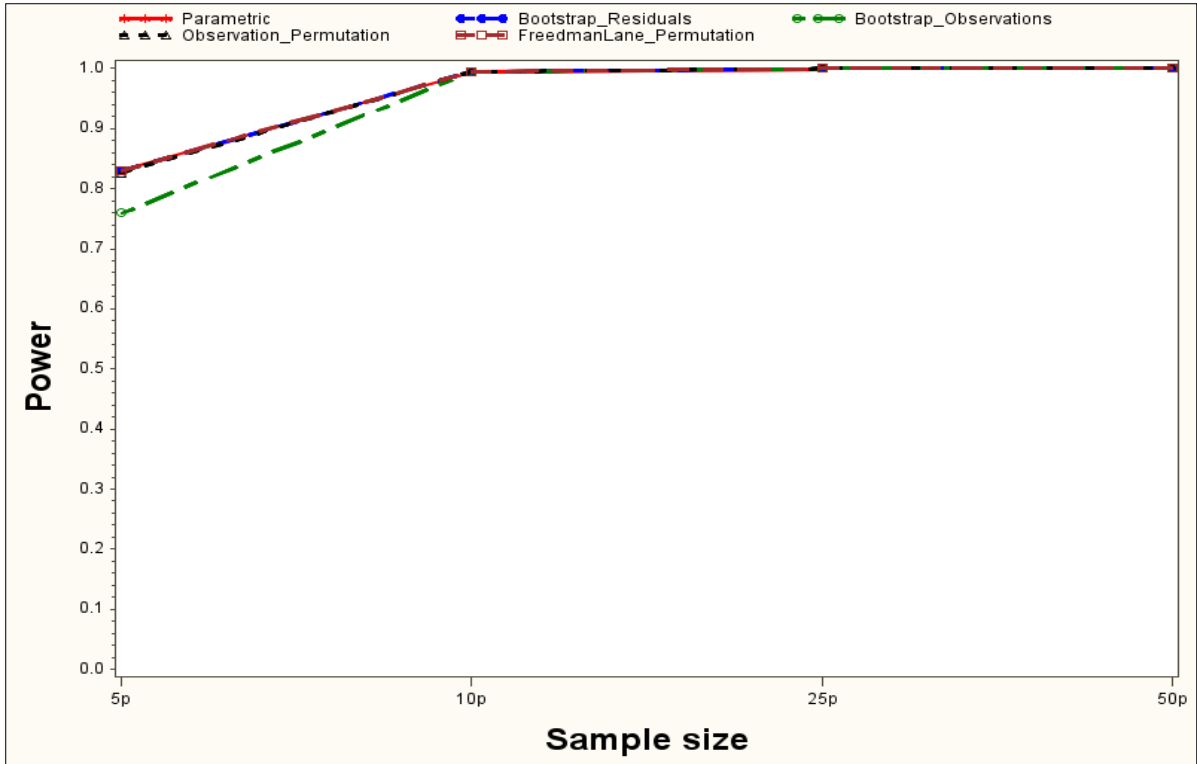


Figure 81 : Power for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on wider vertical axis

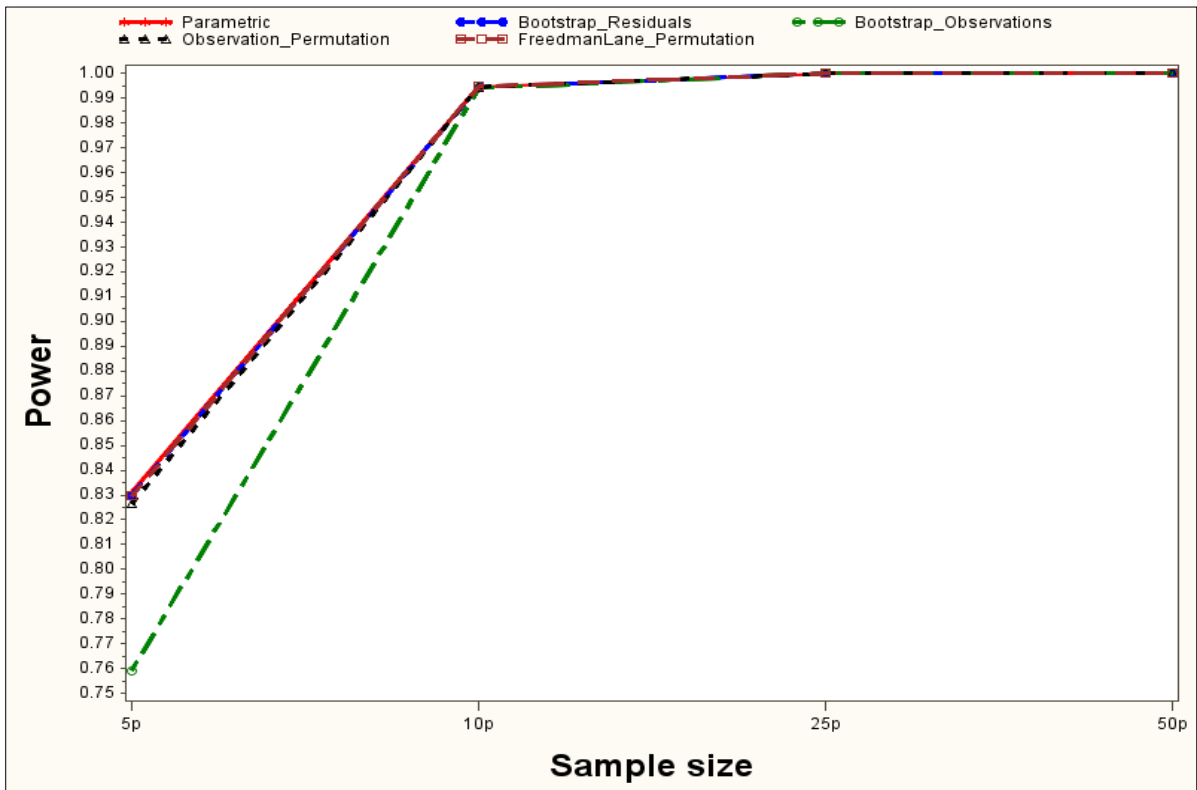


Figure 82: Power for $p = 5$, $\varepsilon \sim \text{Uniform}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis

Power results for $p=5$ LOGN $\rho=0.90$

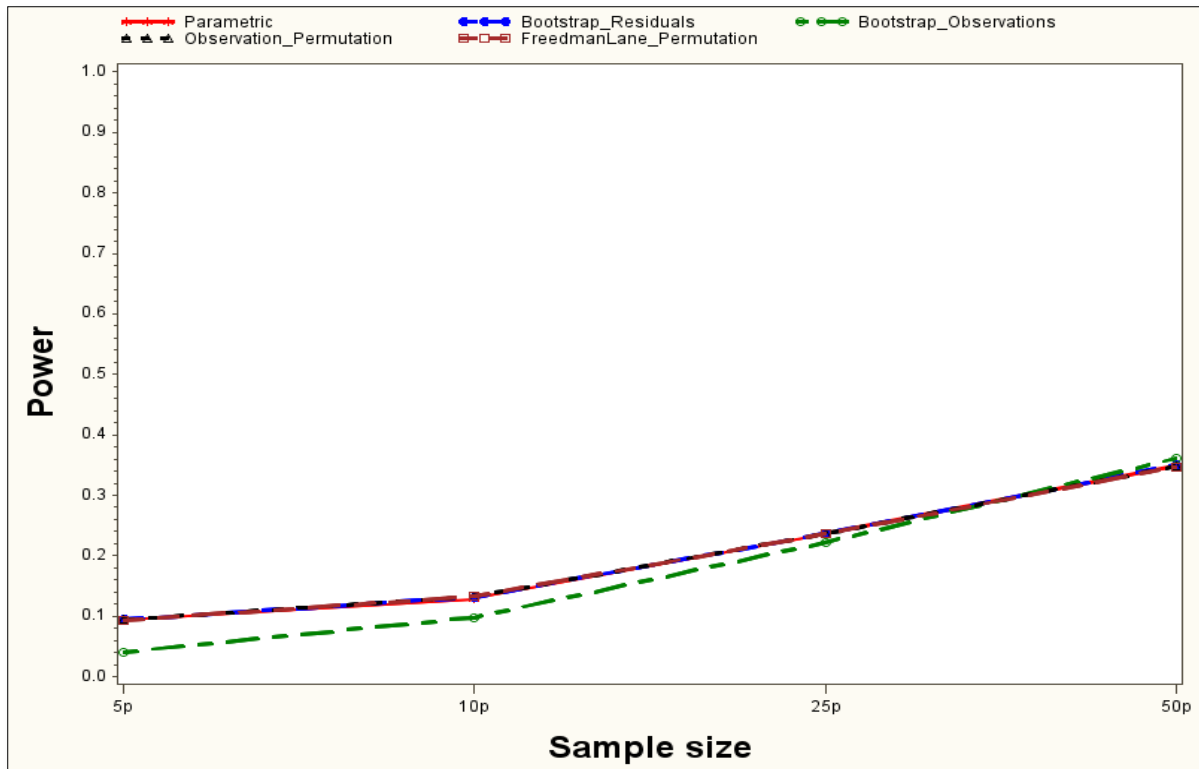


Figure 83: Power for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on wider vertical axis

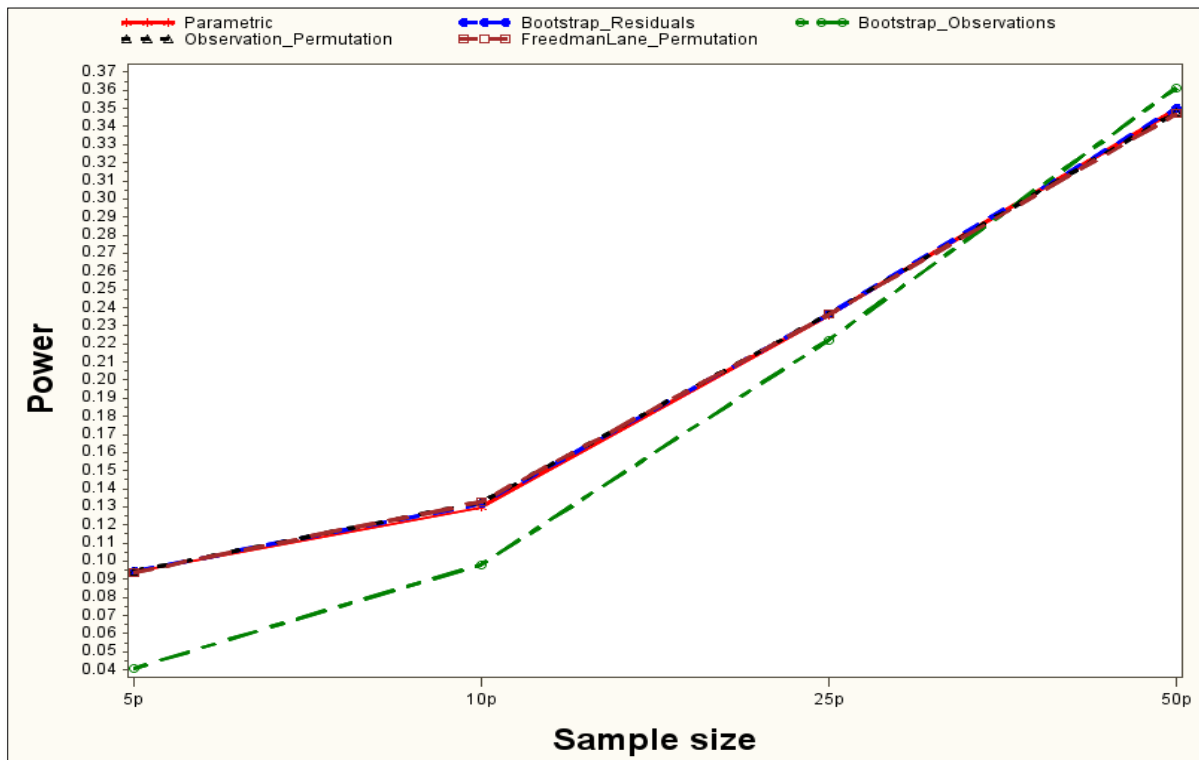


Figure 84: Power for $p = 5$, $\varepsilon \sim \text{LOGN}(0, 1)$ and $\rho = 0.9$ on narrow vertical axis

The effect of increasing multicollinearity from 0.5 to 0.9 resulted in a further reduction of power across all methods. For the case where errors are normally distributed the resulting maximum power decreased from 0.997 to 0.6896 when $p = 2$ and when $p = 5$ the maximum power decreased from 1 to 0.89.

For uniformly distributed errors the difference between the methods is still not visually possible regardless of scale. The reduction in power across all methods is only observed for the smaller sample sizes $5p$ and $10p$ for both cases of $p = 2$ and $p = 5$. When errors are uniformly distributed the multicollinearity does not affect the power of a test for bigger sample sizes.

The loss of power due to increase in multicollinearity when errors are lognormally distributed is substantial since the maximum attained power is 0.27 for $p = 2$ and 0.34 for $p = 5$. The least powerful is the bootstrapping observations for all sample sizes. There is no substantial difference in power attained across all other methods.

6.4 Confidence Intervals

The 95% confidence interval will be presented comparing the results obtained under parametric, jackknife and bootstrap methods under the following conditions

1. Residuals are normally, uniformly and lognormally distributed.
2. Correlation structure within the explanatory variables is 0, 0.5 and 0.9.
3. Number of covariables is 1, 2, and 5.

The number of Monte Carlo simulation used was 1000 and number of bootstrap replications used was $B = 999$. The jackknife and bootstrap confidence interval was computed using the methodology outlined in Chapter 4 while the parametric method is referring to the standard normal confidence intervals. The comparison will be performed in terms of the 95% confidence interval coverage as well as the average length of the confidence intervals. The 95% confidence interval coverage computed for the cases where $\beta_1 = 0$ and $\beta_1 = 2$. Where the coverage is the assessment of how often does the true value of β_1 is within the confidence limits. The average length of confidence interval was computed by averaging the length of confidence interval calculated for each of the 1000 Monte Carlo simulations.

The bootstrapping observation method produced singular matrices for $\hat{\beta}_1$ computations when the sample size is $5p$ for cases where $p = 1$ and $p = 2$. This is because in sampling 999 times with replacement in a “population” of say 5, each pair has 20% chance of being selected, this result in a same pair being repeated in at least one of the samples. This results in problem when the matrices are inverted hence no results are displayed for this method. The SAS program provided an error while running the jackknife delete 2 method when the number of explanatory variables was 5 and the sample size $50p$. The error had to do with the size of the memory being insufficient “(ERROR: (execution) Unable to allocate sufficient memory. At least 2147483647 more bytes required.)”. The results are therefore not presented for that case.

				95% Confidence Interval Coverage							Average Length of 95% Confidence Limits				
no_of_par	rho	error	Beta1	N	Ssize	parametric	Jackknife delete 1	Jackknife delete 2	Bootstrap residual	Bootstrap observations	Parametric	Jackknife delete 1	Jackknife delete 2	Bootstrap residual	Bootstrap observations
1	0	Normal	0	5p	5	93	95	98	88	0	11.700	17.665	28.908	9.074	
				10p	10	88	89	90	85	90	5.457	6.126	6.338	4.879	6.103
				25p	25	74	76	76	71	73	2.932	3.029	3.036	2.810	2.894
				50p	50	48	50	50	47	48	1.995	2.032	2.033	1.954	1.976
			2	5p	5	99	99	100	98	0	11.700	17.665	28.908	9.074	
				10p	10	100	100	100	100	100	5.457	6.126	6.338	4.879	6.103
				25p	25	100	100	100	100	100	2.932	3.029	3.036	2.810	2.894
				50p	50	100	100	100	100	100	1.995	2.032	2.033	1.954	1.976
2	0	Normal	0	5p	10	91	93	94	86	0	5.866	7.112	7.500	4.902	
				10p	20	79	80	81	74	78	3.451	3.696	3.720	3.180	3.506
				25p	50	51	53	53	48	50	2.021	2.075	2.076	1.956	2.006
				50p	100	19	20	20	17	19	1.389	1.408	1.408	1.366	1.382
			2	5p	10	100	100	100	99	0	5.866	7.112	7.500	4.902	
				10p	20	100	100	100	100	100	3.451	3.696	3.720	3.180	3.506
				25p	50	100	100	100	100	100	2.021	2.075	2.076	1.956	2.006
				50p	100	100	100	100	100	100	1.389	1.408	1.408	1.366	1.382
5	0	Normal	0	5p	25	77	84	84	69	82	3.290	3.770	3.803	2.866	3.626
				10p	50	54	59	59	50	54	2.074	2.208	2.211	1.941	2.101
				25p	125	12	13	13	11	12	1.252	1.283	1.283	1.221	1.251
				50p	250	2	2		2	3	0.870	0.880		0.859	0.868
			2	5p	25	100	100	100	100	100	3.290	3.770	3.803	2.866	3.626
				10p	50	100	100	100	100	100	2.074	2.208	2.211	1.941	2.101
				25p	125	100	100	100	100	100	1.252	1.283	1.283	1.221	1.251
				50p	250	100	100		100	100	0.870	0.880		0.859	0.868

2	0.5	Normal	0	5p	10	93	93	95	87	0	6.716	8.076	8.525	5.610	
				10p	20	82	83	83	78	81	3.974	4.248	4.276	3.664	4.031
				25p	50	60	62	62	58	59	2.329	2.380	2.381	2.255	2.304
				50p	100	31	31	31	30	30	1.604	1.622	1.622	1.578	1.593
			2	5p	10	100	100	100	100	0	6.716	8.076	8.525	5.610	
				10p	20	100	100	100	100	100	3.974	4.248	4.276	3.664	4.031
				25p	50	100	100	100	100	100	2.329	2.380	2.381	2.255	2.304
				50p	100	100	100	100	100	100	1.604	1.622	1.622	1.578	1.593
5	0.5	Normal	0	5p	25	86	88	89	79	88	4.247	4.809	4.851	3.699	4.635
				10p	50	68	72	73	64	69	2.671	2.837	2.841	2.501	2.700
				25p	125	29	31	31	27	29	1.614	1.659	1.659	1.574	1.616
				50p	250	6	7		5	6	1.123	1.135		1.108	1.118
			2	5p	25	100	100	100	100	100	4.247	4.809	4.851	3.699	4.635
				10p	50	100	100	100	100	100	2.671	2.837	2.841	2.501	2.700
				25p	125	100	100	100	100	100	1.614	1.659	1.659	1.574	1.616
				50p	250	100	100		100	100	1.123	1.135		1.108	1.118
2	0.9	Normal	0	5p	10	95	96	97	92	0	13.376	16.119	17.045	11.171	
				10p	20	92	92	92	90	91	7.895	8.461	8.516	7.275	8.014
				25p	50	87	88	88	86	86	4.619	4.717	4.721	4.475	4.567
				50p	100	76	77	77	75	76	3.183	3.225	3.225	3.134	3.169
			2	5p	10	99	99	99	99	0	13.376	16.119	17.045	11.171	
				10p	20	99	100	100	99	99	7.895	8.461	8.516	7.275	8.014
				25p	50	100	100	100	100	100	4.619	4.717	4.721	4.475	4.567
				50p	100	100	100	100	100	100	3.183	3.225	3.225	3.134	3.169
5	0.9	Normal	0	5p	25	95	95	96	91	95	9.270	10.543	10.633	8.075	10.141
				10p	50	90	91	91	87	90	5.858	6.203	6.213	5.488	5.909
				25p	125	80	81	81	78	80	3.548	3.650	3.650	3.460	3.555
				50p	250	66	67		65	67	2.467	2.495		2.435	2.458
			2	5p	25	99	99	99	99	100	9.270	10.543	10.633	8.075	10.141
				10p	50	100	100	100	100	100	5.858	6.203	6.213	5.488	5.909
				25p	125	100	100	100	100	100	3.548	3.650	3.650	3.460	3.555
				50p	250	100	100		100	100	2.467	2.495		2.435	2.458

Table 1: Confidence Intervals when errors are normally distributed

						95% Confidence Interval Coverage					Average Length of 95% Confidence Limits				
no_of_p r	rho	Error	Beta 1	n	Ssize	parametri c	Jackknif e delete 1	Jackknif e delete 2	Bootstra p residual	Bootstrap observatio n	Parametri c	Jackknif e delete 1	Jackknif e delete 2	Bootstra p residual	Bootstrap observatio ns
1	0	Unifor m	0	5p	5	76	82	93	62	0	3.472	5.283	8.647	2.684	
				10p	10	30	40	42	21	38	1.556	1.733	1.795	1.391	1.733
				25p	25	3	3	3	3	3	0.848	0.883	0.885	0.812	0.841
				50p	50	2	3	3	3	3	0.576	0.584	0.585	0.564	0.569
			2	5p	5	100	100	100	100	0	3.472	5.283	8.647	2.684	
				10p	10	100	100	100	100	100	1.556	1.733	1.795	1.391	1.733
				25p	25	100	100	100	100	100	0.848	0.883	0.885	0.812	0.841
				50p	50	100	100	100	100	100	0.576	0.584	0.585	0.564	0.569
2	0	Unifor m	0	5p	10	38	52	57	27	0	1.751	2.136	2.263	1.466	
				10p	20	5	7	7	5	6	1.007	1.087	1.094	0.928	1.026
				25p	50	3	3	3	3	3	0.586	0.601	0.601	0.568	0.581
				50p	100	3	3	3	3	3	0.402	0.408	0.408	0.395	0.400
			2	5p	10	100	100	100	100	0	1.751	2.136	2.263	1.466	
				10p	20	100	100	100	100	100	1.007	1.087	1.094	0.928	1.026
				25p	50	100	100	100	100	100	0.586	0.601	0.601	0.568	0.581
				50p	100	100	100	100	100	100	0.402	0.408	0.408	0.395	0.400
5	0	Unifor m	0	5p	25	3	6	6	4	5	0.947	1.091	1.101	0.826	1.046
				10p	50	2	2	2	3	2	0.601	0.637	0.638	0.564	0.606
				25p	125	2	2	2	3	3	0.363	0.371	0.371	0.354	0.362
				50p	250	3	3	3	3	3	0.252	0.255	0.249	0.252	
			2	5p	25	100	100	100	100	100	0.947	1.091	1.101	0.826	1.046
				10p	50	100	100	100	100	100	0.601	0.637	0.638	0.564	0.606
				25p	125	100	100	100	100	100	0.363	0.371	0.371	0.354	0.362
				50p	250	100	100	100	100	100	0.252	0.255	0.249	0.252	

2	0.5	Unifor m	0	5p	10	50	61	65	37	0	1.994	2.401	2.552	1.668	
				10p	20	9	13	13	8	11	1.162	1.251	1.259	1.071	1.183
				25p	50	3	3	3	4	3	0.678	0.696	0.697	0.658	0.673
				50p	100	2	2	2	3	3	0.464	0.470	0.470	0.456	0.461
			2	5p	10	100	100	100	100	0	1.994	2.401	2.552	1.668	
				10p	20	100	100	100	100	100	1.162	1.251	1.259	1.071	1.183
				25p	50	100	100	100	100	100	0.678	0.696	0.697	0.658	0.673
				50p	100	100	100	100	100	100	0.464	0.470	0.470	0.456	0.461
5	0.5	Unifor m	0	5p	25	10	17	17	7	14	1.216	1.393	1.405	1.060	1.335
				10p	50	2	2	2	3	2	0.775	0.822	0.823	0.726	0.782
				25p	125	2	2	2	2	2	0.468	0.477	0.477	0.456	0.466
				50p	250	3	3		3	3	0.325	0.329		0.321	0.324
			2	5p	25	100	100	100	100	100	1.216	1.393	1.405	1.060	1.335
				10p	50	100	100	100	100	100	0.775	0.822	0.823	0.726	0.782
				25p	125	100	100	100	100	100	0.468	0.477	0.477	0.456	0.466
				50p	250	100	100		100	100	0.325	0.329		0.321	0.324
2	0.9	Unifor m	0	5p	10	84	86	89	75	0	3.947	4.752	5.056	3.299	
				10p	20	60	66	66	53	62	2.307	2.492	2.508	2.126	2.356
				25p	50	18	19	19	16	18	1.346	1.386	1.387	1.303	1.339
				50p	100	2	2	2	2	2	0.922	0.935	0.935	0.907	0.916
			2	5p	10	100	100	100	100	0	3.947	4.752	5.056	3.299	
				10p	20	100	100	100	100	100	2.307	2.492	2.508	2.126	2.356
				25p	50	100	100	100	100	100	1.346	1.386	1.387	1.303	1.339
				50p	100	100	100	100	100	100	0.922	0.935	0.935	0.907	0.916
5	0.9	Unifor m	0	5p	25	68	74	74	58	73	2.668	3.033	3.061	2.326	2.919
				10p	50	36	40	40	31	37	1.701	1.804	1.807	1.594	1.716
				25p	125	3	3	3	3	3	1.027	1.048	1.048	1.001	1.023
				50p	250	3	3		3	3	0.713	0.721		0.705	0.711
			2	5p	25	100	100	100	100	100	2.668	3.033	3.061	2.326	2.919
				10p	50	100	100	100	100	100	1.701	1.804	1.807	1.594	1.716
				25p	125	100	100	100	100	100	1.027	1.048	1.048	1.001	1.023
				50p	250	100	100		100	100	0.713	0.721		0.705	0.711

Table 2: Confidence Intervals when errors are uniformly distributed

no_of_p ar	rho	Error	Beta	n	Ssize	95% Confidence Interval Coverage					Average Length of 95% Confidence Limits				
						parametri c	Jackknif e delete 1	Jackknif e delete 2	Bootstra p residual	Bootstrap observation s	0	Parametri c	Jackknif e delete 1	Jackknif e delete 2	Bootstra p residual
1	0	Lognorma 1	0	5p	5	93	95	98	89	0	19.684	30.578	48.199	15.274	
				10p	10	90	92	93	87	92	9.782	10.678	11.082	8.749	10.839
				25p	25	87	86	86	85	85	5.619	5.682	5.697	5.392	5.484
				50p	50	78	76	76	76	75	3.938	3.915	3.917	3.851	3.822
			2	5p	5	98	99	100	97	0	19.684	30.578	48.199	15.274	
				10p	10	100	100	100	99	100	9.782	10.678	11.082	8.749	10.839
				25p	25	100	100	100	100	100	5.619	5.682	5.697	5.392	5.484
				50p	50	100	100	100	100	100	3.938	3.915	3.917	3.851	3.822
2	0	Lognorma 1	0	5p	10	93	95	96	88	0	10.613	12.426	13.202	8.885	
				10p	20	88	89	89	85	88	6.446	6.711	6.758	5.937	6.427
				25p	50	80	80	80	79	77	4.029	4.031	4.034	3.902	3.917
				50p	100	66	65	65	65	64	2.863	2.828	2.829	2.817	2.783
			2	5p	10	100	100	100	100	0	10.613	12.426	13.202	8.885	
				10p	20	99	100	100	99	100	6.446	6.711	6.758	5.937	6.427
				25p	50	100	100	100	100	100	4.029	4.031	4.034	3.902	3.917
				50p	100	100	100	100	100	100	2.863	2.828	2.829	2.817	2.783
5	0	Lognorma 1	0	5p	25	88	92	93	83	92	6.330	7.167	7.235	5.509	6.947
				10p	50	81	83	83	78	81	4.088	4.274	4.281	3.828	4.086
				25p	125	61	60	60	60	59	2.610	2.628	2.628	2.543	2.568
				50p	250	41	40	40	41	39	1.829	1.824	1.824	1.805	1.802
			2	5p	25	100	100	100	100	100	6.330	7.167	7.235	5.509	6.947
				10p	50	100	100	100	100	100	4.088	4.274	4.281	3.828	4.086
				25p	125	100	100	100	100	100	2.610	2.628	2.628	2.543	2.568
				50p	250	100	100	100	100	100	1.829	1.824	1.824	1.805	1.802

2	0.5	Lognorma 1	0	5p	10	95	96	97	90	0	12.047	14.117	15.024	10.078	
				10p	20	89	89	89	87	89	7.434	7.743	7.799	6.846	7.428
				25p	50	83	84	84	82	81	4.642	4.611	4.614	4.499	4.486
				50p	100	73	72	72	73	70	3.310	3.232	3.233	3.257	3.183
			2	5p	10	100	100	100	99	0	12.047	14.117	15.024	10.078	
				10p	20	100	100	100	100	100	7.434	7.743	7.799	6.846	7.428
				25p	50	100	100	100	100	100	4.642	4.611	4.614	4.499	4.486
				50p	100	100	100	100	100	100	3.310	3.232	3.233	3.257	3.183
5	0.5	Lognorma 1	0	5p	25	91	94	95	86	94	8.186	8.964	9.055	7.121	8.770
				10p	50	86	89	90	84	88	5.260	5.458	5.468	4.927	5.232
				25p	125	73	73	73	72	72	3.368	3.405	3.406	3.283	3.328
				50p	250	59	56		58	55	2.362	2.344		2.328	2.316
			2	5p	25	100	100	100	99	100	8.186	8.964	9.055	7.121	8.770
				10p	50	100	100	100	100	100	5.260	5.458	5.468	4.927	5.232
				25p	125	100	100	100	100	100	3.368	3.405	3.406	3.283	3.328
				50p	250	100	100		100	100	2.362	2.344		2.328	2.316
2	0.9	Lognorma 1	0	5p	10	97	97	98	94	0	24.005	28.246	30.070	20.070	
				10p	20	95	95	95	92	95	14.777	15.540	15.651	13.608	14.880
				25p	50	93	94	94	92	92	9.209	9.139	9.146	8.922	8.895
				50p	100	91	90	90	90	89	6.567	6.506	6.507	6.468	6.408
			2	5p	10	99	100	100	99	0	24.005	28.246	30.070	20.070	
				10p	20	99	100	100	98	100	14.777	15.540	15.651	13.608	14.880
				25p	50	99	100	100	99	100	9.209	9.139	9.146	8.922	8.895
				50p	100	100	100	100	100	100	6.567	6.506	6.507	6.468	6.408
5	0.9	Lognorma 1	0	5p	25	96	98	98	94	98	17.832	19.383	19.578	15.524	19.005
				10p	50	94	95	95	93	94	11.531	11.987	12.008	10.805	11.503
				25p	125	91	92	92	90	91	7.404	7.505	7.507	7.219	7.336
				50p	250	89	88		88	87	5.187	5.134		5.114	5.071
			2	5p	25	99	100	100	99	100	17.832	19.383	19.578	15.524	19.005
				10p	50	99	100	100	99	100	11.531	11.987	12.008	10.805	11.503
				25p	125	99	100	100	99	100	7.404	7.505	7.507	7.219	7.336
				50p	250	99	100		99	99	5.187	5.134		5.114	5.071

Table 3: Confidence Intervals when errors are lognormally distributed

6.4.1 Normally distributed errors

The confidence interval was compared for $\beta_1 = 0$ and for $\beta_1 = 2$. Under normally distributed errors and independent explanatory variables, the confidence intervals consistently over-cover for $\beta_1 = 2$ with the least coverage of 98% being of observed under bootstrapping residuals. For $\beta_1 = 0$ a reasonable coverage is observed for the sample size equal $5p$ and $10p$ when $p = 1$ and $p = 2$. The coverage becomes rapidly poorer with the increase in sample size for $\beta_1 = 0$ while the average length of the confidence interval becomes narrower with an increase in sample size. The jackknife methods resulted in higher coverage compared to bootstrap methods; however bootstrap methods have a narrower average length of confidence interval. In all methods, delete 2 jackknife methods have the largest average length of confidence interval of 28.9 and the highest coverage of 98% for $\beta_1 = 0$ which is observed when $p = 1$. The true coverage of 95% for $\beta_1 = 0$ is only attained by delete 1 jackknife methods when $p = 1$ and sample size $5p$. For the case where there are 5 regressors and sample size is $50p$ the coverage of the confidence interval is extremely poor for $\beta_1 = 0$. Coverage is 2% across all methods with an exception of bootstrapping observations which has coverage of 3% while the average length of the confidence interval is very narrow, ranging from 0.859 to 0.88 for $\beta_1 = 0$.

The increase in correlation among the regressors generally improves the confidence interval coverage while increasing the average length of the confidence interval. When $\rho = 0.5$ and $p = 2$ the true coverage of 95% is obtained for the jackknife delete 2 method when sample size is $5p$. The jackknife delete 1 method overstates the coverage by 1% when the sample is $5p$ and provides the true coverage of 95% when sample size is $10p$ together with a delete 2 jackknife. The true coverage however is achieved at a cost of a wider average length of the confidence interval. The bootstrap residual and observation methods provide better average lengths of the confidence interval compared to both delete 1 and delete 2 jackknife methods under the same conditions.

The increase of the correlation coefficient between regressors from 0.5 to 0.9 significantly widens the average length of the confidence interval from 6.7 to 13.38

and 8.5 to 17.04 under parametric and delete 2 jackknife respectively for a combination of $p = 2$ and sample size $5p$. In general the relationship between confidence interval coverage and the average length of the confidence interval is that they are directly proportional to one another. There is a notable improvement in coverage in the presence of multicollinearity which is achieved at an expense of wider confidence interval.

6.4.2 Uniformly distributed errors

When errors are uniformly distributed and $\rho = 0$ the confidence interval coverage is consistently understated across all methods for the case where $\beta_1 = 0$ except for the jackknife delete 2 method when $p = 1$ and sample size $5p$ where the coverage is understated by 2%. For the case $\beta_1 = 2$ the coverage is always 100%. The level at which the coverage is understated is more severe for the uniformly distributed errors compared to the normally distributed errors because the average length of the confidence interval is much narrower. For $p = 5$ a 2% coverage is observed for parametric and bootstrapping observation methods at a sample size $25p$ which was never observed under normality. Instances of poor coverage are observed when $\rho = 0$ across all methods. The delete 1 and delete 2 jackknife methods give better coverage at small sample sizes and coverage deteriorates at a rapid rate as the sample size increases.

As under normality, the coverage improves with the increase in correlation among explanatory variables at an expense of wider average length of confidence interval. The best coverage of 89% was observed under jackknife delete 2 method when $p = 2$ and $\rho = 0.9$. This coverage was attained at an expense of the widest average length of confidence interval of 5.06. The bootstrap residual method produced the shortest average length of confidence interval in all cases as compared to the rest other the methods considered.

6.4.3 Lognormally distributed errors

For the lognormally distributed errors a trade-off between the coverage and average length of confidence interval is observed. There is a significant improvement in

coverage of all methods for $\beta_1 = 0$ compared to normally and uniformly distributed errors. When $\rho = 0$ the true coverage of 95% is observed for jackknife delete 1 method when $p = 1$, $p = 2$ and sample size $5p$. In both instances this method has the second widest average length of a confident interval of 30.58 for $p = 1$ and 12.43 for $p = 2$. The largest overstatement of coverage (98%) is observed under delete 2 jackknife together with an average distance of confidence interval of 48.2 which is observed when $p = 1$, $\rho = 0$ and a sample size of $5p$. When $\rho = 0$ the biggest understatement of coverage observed was 39% coverage which was seen under bootstrapping observations when $p = 5$ and sample size $50p$ and the observed average length of interval was 1.802 which is the smallest average length.

The introduction of a correlation factor of 0.5 among the explanatory variables produced an improved coverage particularly for $\beta_1 = 0$. The worst coverage of 55% is observed under bootstrapping observations when $p = 5$ and the sample size $50p$ with the average lengths of the confidence interval being 2.3. An overstatement of coverage (98%) for $\beta_1 = 0$ is observed under delete1, delete2 jackknife and bootstrapping observation when $p = 5$ and $\rho = 0.9$ with the corresponding average length of the confidence interval of 19.38, 19.56 and 19.01 respectively. The true coverage (95%) for $\beta_1 = 0$ is attained under parametric, jackknife delete 1, jackknife delete 2 and bootstrapping observations when $p = 2$ and $\rho = 0.9$. At this instance the parametric method has the smallest average length of confidence interval of 14.78 followed by bootstrapping observation method with an average length of confidence interval of 14.89. The true coverage (95%) is also attained under jackknife delete 1 and jackknife delete 2 methods when $p = 5$ and $\rho = 0.9$ with the parametric and bootstrapping observations methods producing coverage of 94%. For $\rho = 0.5$ and $\rho = 0.9$ bootstrapping residuals method is always providing a slightly less coverage for all sample sizes except sample size $50p$.

6.5 Percentile Based Interval

Percentile based intervals are presented for jackknife and bootstrapping methods. For consistency the comparison is performed on the general percentile method

described in Chapter 4 and not on methods that improve the percentile based confidence interval for bootstrapping because the same improvement can't be performed for the jackknife method (Efron, 2003).

				Coverage of Percentile Based Intervals						Average Length of Percentile Limits			
no_of_par	rho	error	Beta1	N	Ssize	Jackknife delete 1	Jackknife delete 2	Bootstrap residual	Bootstrap observations	Jackknife delete 1	Jackknife delete 2	Bootstrap residual	Bootstrap observations
1	0	Normal	0	5p	5	57	85	76	0	3.948	13.177	5.500	
				10p	10	45	54	78	82	1.640	2.416	4.158	5.092
				25p	25	33	28	67	70	0.722	0.879	2.671	2.762
				50p	50	37	32	45	47	0.304	0.434	1.913	1.937
			2	5p	5	88	97	94	0	3.948	13.177	5.500	
				10p	10	91	94	100	100	1.640	2.416	4.158	5.092
				25p	25	97	98	100	100	0.722	0.879	2.671	2.762
				50p	50	99	99	100	100	0.304	0.434	1.913	1.937
2	0	Normal	0	5p	10	50	62	78	0	1.861	2.808	4.073	
				10p	20	32	33	71	74	0.914	1.190	2.966	3.289
				25p	50	39	34	47	48	0.307	0.445	1.914	1.966
				50p	100	41	38	17	18	0.157	0.212	1.356	1.373
			2	5p	10	93	96	99	0	1.861	2.808	4.073	
				10p	20	96	97	100	100	0.914	1.190	2.966	3.289
				25p	50	99	99	100	100	0.307	0.445	1.914	1.966
				50p	100	100	100	100	100	0.157	0.212	1.356	1.373
5	0	Normal	0	5p	25	33	32	65	79	0.902	1.091	2.698	3.445
				10p	50	40	34	48	52	0.326	0.476	1.896	2.056
				25p	125	41	39	11	12	0.123	0.173	1.213	1.246
				50p	250	46		2	3	0.061		0.858	0.868
			2	5p	25	96	97	100	100	0.902	1.091	2.698	3.445
				10p	50	98	98	100	100	0.326	0.476	1.896	2.056
				25p	125	100	100	100	100	0.123	0.173	1.213	1.246
				50p	250	100		100	100	0.061		0.858	0.868

2	0.5	Normal	0	5p	10	50	65	81	0	2.139	3.191	4.661	
				10p	20	37	39	74	78	1.075	1.376	3.415	3.790
				25p	50	39	33	56	58	0.353	0.523	2.206	2.258
				50p	100	42	39	28	29	0.183	0.249	1.564	1.584
			2	5p	10	92	96	99	0	2.139	3.191	4.661	
				10p	20	94	95	100	100	1.075	1.376	3.415	3.790
				25p	50	98	99	100	100	0.353	0.523	2.206	2.258
				50p	100	100	100	100	100	0.183	0.249	1.564	1.584
5	0.5	Normal	0	5p	25	38	39	75	85	1.176	1.401	3.476	4.405
				10p	50	38	32	62	68	0.421	0.632	2.442	2.647
				25p	125	40	36	28	29	0.159	0.229	1.564	1.608
				50p	250	45		5	6	0.079		1.108	1.119
			2	5p	25	94	96	100	100	1.176	1.401	3.476	4.405
				10p	50	97	98	100	100	0.421	0.632	2.442	2.647
				25p	125	100	100	100	100	0.159	0.229	1.564	1.608
				50p	250	100		100	100	0.079		1.108	1.119
2	0.9	Normal	0	5p	10	61	73	88	0	4.263	6.340	9.286	
				10p	20	51	57	87	88	2.137	2.746	6.771	7.526
				25p	50	38	36	85	86	0.700	1.032	4.379	4.475
				50p	100	40	37	74	74	0.362	0.493	3.105	3.145
			2	5p	10	84	91	97	0	4.263	6.340	9.286	
				10p	20	85	89	99	99	2.137	2.746	6.771	7.526
				25p	50	89	92	100	100	0.700	1.032	4.379	4.475
				50p	100	93	94	100	100	0.362	0.493	3.105	3.145
5	0.9	Normal	0	5p	25	55	60	89	94	2.596	3.077	7.589	9.645
				10p	50	38	40	86	89	0.921	1.390	5.360	5.797
				25p	125	40	37	77	80	0.350	0.502	3.441	3.532
				50p	250	46		65	66	0.174		2.436	2.459
			2	5p	25	84	86	99	100	2.596	3.077	7.589	9.645
				10p	50	84	88	100	99	0.921	1.390	5.360	5.797
				25p	125	90	91	100	100	0.350	0.502	3.441	3.532
				50p	250	95		100	100	0.174		2.436	2.459

Table 4: Percentile based interval when errors are normally distributed

						Coverage of Percentile Based Intervals				Average Length of Percentile Limits			
no_of_par	rho	Error	Beta1	N	Ssize	Jackknife delete 1	Jackknife delete 2	Bootstrap residual	Bootstrap observations	Jackknife delete 1	Jackknife delete 2	Bootstrap residual	Bootstrap observations
1	0	Uniform	0	5p	5	33	61	36	0	1.173	3.968	1.626	
				10p	10	29	18	15	23	0.457	0.689	1.178	1.447
				25p	25	31	27	4	4	0.194	0.251	0.770	0.802
				50p	50	36	31	3	3	0.083	0.119	0.552	0.557
			2	5p	5	99	100	100	0	1.173	3.968	1.626	
				10p	10	100	100	100	100	0.457	0.689	1.178	1.447
				25p	25	100	100	100	100	0.194	0.251	0.770	0.802
				50p	50	100	100	100	100	0.083	0.119	0.552	0.557
2	0	Uniform	0	5p	10	26	20	20	0	0.554	0.845	1.211	
				10p	20	30	24	5	6	0.258	0.349	0.863	0.962
				25p	50	38	34	3	3	0.086	0.123	0.555	0.569
				50p	100	43	39	4	3	0.043	0.059	0.392	0.397
			2	5p	10	100	100	100	0	0.554	0.845	1.211	
				10p	20	100	100	100	100	0.258	0.349	0.863	0.962
				25p	50	100	100	100	100	0.086	0.123	0.555	0.569
				50p	100	100	100	100	100	0.043	0.059	0.392	0.397
5	0	Uniform	0	5p	25	29	24	5	5	0.251	0.312	0.774	0.993
				10p	50	39	33	4	2	0.092	0.132	0.550	0.594
				25p	125	42	40	3	3	0.034	0.048	0.351	0.360
				50p	250	45		3	3	0.017		0.249	0.252
			2	5p	25	100	100	100	100	0.251	0.312	0.774	0.993
				10p	50	100	100	100	100	0.092	0.132	0.550	0.594
				25p	125	100	100	100	100	0.034	0.048	0.351	0.360
				50p	250	100		100	100	0.017		0.249	0.252

2	0.5	Uniform	0	5p	10	28	23	26	0	0.627	0.945	1.379	
				10p	20	29	24	8	10	0.306	0.403	0.995	1.111
				25p	50	37	33	4	4	0.102	0.145	0.642	0.660
				50p	100	43	40	3	2	0.052	0.069	0.452	0.458
			2	5p	10	100	100	100	0	0.627	0.945	1.379	
				10p	20	100	100	100	100	0.306	0.403	0.995	1.111
				25p	50	100	100	100	100	0.102	0.145	0.642	0.660
				50p	100	100	100	100	100	0.052	0.069	0.452	0.458
5	0.5	Uniform	0	5p	25	27	23	6	13	0.332	0.403	0.994	1.271
				10p	50	38	33	4	3	0.121	0.175	0.708	0.767
				25p	125	42	40	3	3	0.045	0.063	0.453	0.463
				50p	250	45		3	3	0.023		0.321	0.324
			2	5p	25	100	100	100	100	0.332	0.403	0.994	1.271
				10p	50	100	100	100	100	0.121	0.175	0.708	0.767
				25p	125	100	100	100	100	0.045	0.063	0.453	0.463
				50p	250	100		100	100	0.023		0.321	0.324
2	0.9	Uniform	0	5p	10	37	48	64	0	1.238	1.891	2.723	
				10p	20	30	25	48	58	0.606	0.804	1.975	2.213
				25p	50	39	35	15	17	0.203	0.288	1.274	1.314
				50p	100	41	37	2	2	0.102	0.137	0.898	0.909
			2	5p	10	95	98	99	0	1.238	1.891	2.723	
				10p	20	98	99	100	100	0.606	0.804	1.975	2.213
				25p	50	100	100	100	100	0.203	0.288	1.274	1.314
				50p	100	100	100	100	100	0.102	0.137	0.898	0.909
5	0.9	Uniform	0	5p	25	30	28	54	70	0.722	0.879	2.181	2.782
				10p	50	37	31	29	36	0.267	0.386	1.555	1.685
				25p	125	43	40	3	3	0.100	0.140	0.994	1.017
				50p	250	45		3	3	0.050		0.704	0.711
			2	5p	25	98	98	100	100	0.722	0.879	2.181	2.782
				10p	50	100	100	100	100	0.267	0.386	1.555	1.685
				25p	125	100	100	100	100	0.100	0.140	0.994	1.017
				50p	250	100		100	100	0.050		0.704	0.711

Table 5: Percentile based interval when errors are uniformly distributed

						Coverage of Percentile Based Intervals				Average Length of Percentile Limits			
no_of_par	rho	Error	Beta1	N	Ssize	Jackknife delete 1	Jackknife delete 2	Bootstrap residual	Bootstrap observations	Jackknife delete 1	Jackknife delete 2	Bootstrap residual	Bootstrap observations
1	0	Lognormal	0	5p	5	59	84	76	0	6.847	21.362	9.268	
				10p	10	50	59	82	83	2.949	4.128	7.597	8.802
				25p	25	41	43	83	78	1.503	1.725	5.244	5.141
				50p	50	39	33	76	70	0.475	0.998	3.836	3.724
			2	5p	5	88	97	93	0	6.847	21.362	9.268	
				10p	10	89	93	99	99	2.949	4.128	7.597	8.802
				25p	25	95	95	99	100	1.503	1.725	5.244	5.141
				50p	50	91	96	100	100	0.475	0.998	3.836	3.724
2	0	Lognormal	0	5p	10	56	66	83	0	3.337	4.831	7.509	
				10p	20	42	45	82	81	1.793	2.186	5.669	5.916
				25p	50	41	36	79	72	0.500	1.024	3.895	3.807
				50p	100	43	39	66	60	0.250	0.404	2.834	2.754
			2	5p	10	92	95	99	0	3.337	4.831	7.509	
				10p	20	94	95	99	99	1.793	2.186	5.669	5.916
				25p	50	92	96	100	100	0.500	1.024	3.895	3.807
				50p	100	93	95	100	100	0.250	0.404	2.834	2.754
5	0	Lognormal	0	5p	25	43	45	80	87	1.859	2.140	5.276	6.485
				10p	50	43	38	77	76	0.556	1.064	3.806	3.970
				25p	125	45	40	60	57	0.189	0.339	2.560	2.547
				50p	250	47	41	41	40	0.091		1.821	1.800
			2	5p	25	93	94	100	100	1.859	2.140	5.276	6.485
				10p	50	91	96	100	100	0.556	1.064	3.806	3.970
				25p	125	95	97	100	100	0.189	0.339	2.560	2.547
				50p	250	99		100	100	0.091		1.821	1.800

2	0.5	Lognormal	0	5p	10	57	68	84	0	3.837	5.511	8.541	
				10p	20	46	50	85	81	2.092	2.526	6.554	6.830
				25p	50	40	37	82	77	0.566	1.182	4.509	4.355
				50p	100	42	37	73	67	0.286	0.461	3.296	3.153
			2	5p	10	90	95	98	0	3.837	5.511	8.541	
				10p	20	91	92	99	99	2.092	2.526	6.554	6.830
25p	50	90		96	100	100	0.566	1.182	4.509	4.355			
50p	100	91		93	100	100	0.286	0.461	3.296	3.153			
5	0.5	Lognormal	0	5p	25	48	51	85	90	2.344	2.687	6.852	8.234
				10p	50	42	40	83	84	0.702	1.382	4.922	5.081
				25p	125	45	40	73	68	0.241	0.435	3.325	3.292
				50p	250	45		59	54	0.116		2.356	2.315
			2	5p	25	91	93	99	100	2.344	2.687	6.852	8.234
				10p	50	86	95	100	100	0.702	1.382	4.922	5.081
25p	125	89		92	100	100	0.241	0.435	3.325	3.292			
50p	250	97			100	100	0.116		2.356	2.315			
2	0.9	Lognormal	0	5p	10	65	75	90	0	7.649	10.954	16.984	
				10p	20	60	64	90	90	4.197	5.075	13.019	13.677
				25p	50	43	50	92	89	1.118	2.331	8.938	8.635
				50p	100	41	39	90	87	0.570	0.912	6.528	6.342
			2	5p	10	85	90	97	0	7.649	10.954	16.984	
				10p	20	82	86	98	99	4.197	5.075	13.019	13.677
25p	50	78		87	99	99	1.118	2.331	8.938	8.635			
50p	100	78		82	100	100	0.570	0.912	6.528	6.342			
5	0.9	Lognormal	0	5p	25	62	65	93	95	5.072	5.802	14.971	17.837
				10p	50	45	54	92	91	1.510	3.059	10.800	11.151
				25p	125	44	40	90	88	0.532	0.951	7.315	7.254
				50p	250	47		89	86	0.256		5.181	5.068
			2	5p	25	82	84	99	99	5.072	5.802	14.971	17.837
				10p	50	75	85	99	99	1.510	3.059	10.800	11.151
25p	125	74		78	99	100	0.532	0.951	7.315	7.254			
50p	250	82			99	99	0.256		5.181	5.068			

Table 6: Percentile based interval when errors are lognormally distributed

6.5.1 Normally distributed errors

For $\beta_1 = 0$ and $p = 1$ the best coverage of 85% is observed under the jackknife delete 2 method when the samples size is $5p$. The same method produced the largest average length of the confidence interval of 13.18. An increase in sample size resulted in all the methods showing a decrease in both the coverage and average length of an interval.

For $\beta_1 = 2$ and $p = 1$, jackknife methods provided coverage that is closest to the true coverage of 95% and a shorter average length of an interval than bootstrap methods while the bootstrap methods are consistently overstating the coverage. The coverage increased with the increase in sample size for this case.

Bootstrapping methods are observed to be performing better than jackknife methods in terms of coverage for small sample sizes and thereafter deteriorate rapidly as the sample size increase for the case of $\beta_1 = 0$ when $p = 2$ and $p = 5$. For $p = 2$ and $p = 5$ bootstrapping residuals method showed a 78% and 65% coverage at a sample size of $5p$ and 17% and 2% at a sample size of $50p$ respectively. Meanwhile jackknife delete 1 showed 50% and 33% coverage at sample size $5p$ and 41% and 46% at sample size $50p$ respectively. For $\beta_1 = 2$ all methods are overstating the coverage when $p = 2$ and $p = 5$ with an exception of jackknife delete 1 method which provided a coverage of 93% for the sample size $5p$.

The introduction of $\rho = 0.5$ between explanatory variables showed an increase in both the coverage and average length of an interval for the case where is $\beta_1 = 0$. An improvement in the coverage of all methods was observed at an expense of an increased length of an interval, however increases are not very significant compared to the same case when the explanatory variables were independent. Bootstrap methods provided very consistent coverage across all sample sizes. For $\beta_1 = 0$ bootstrapping residuals showed the best coverage of 81% when $p = 2$ while bootstrapping observations showed the best coverage of 85% when $p = 5$.

For $\beta_1 = 2$ the both bootstrapping methods consistently overstate the coverage while the jackknife methods fluctuate around the 95% coverage while providing a shorter length of interval. The best coverage observed under the jackknife delete 2

method is 95% which was observed at sample size $10p$ with 1.4 as the corresponding lengths when $p = 2$.

An increase in multicollinearity among the explanatory variables from $\rho = 0.5$ to $\rho = 0.9$ resulted in the improved coverage for bootstrapping methods for both cases ($\beta_1 = 0$ and $\beta_1 = 2$) when $p = 2$ and $p = 5$. This improvement however came at the expense of the significantly increased length of the confidence interval. For $p = 2$ the widest length of an interval is 9.3 when $\rho = 0.9$, from 4.7, when $\rho = 0.5$ while the shortest length of an interval is 0.36 from 0.18 these are both observed for sample sizes $5p$ and $50p$. For $\beta_1 = 2$ coverage under jackknife methods improves substantially as the sample size increase.

6.5.2 Uniformly distributed errors

In the case where all explanatory variables are independent, the coverage of the percentile based confidence interval is understated across all methods considered for $\beta_1 = 0$. The two jackknife methods are performing better in both coverage and average length of an interval than their bootstrap counterpart. The best coverage observed is 61% was observed under the jackknife delete 2 method. This coverage is observed when $p = 1$. The worst coverage of 2% is observed under bootstrapping observations when $p = 5$ and sample size $10p$. The worst coverage observed under the jackknife method is 20% which was observed under the jackknife delete 2 method when $p = 2$ and sample size $5p$. For corresponding sample sizes the jackknife methods consistently showed shorter lengths of interval except for one case where jackknife delete 2 method is 4.0 which is observed when $p = 1$ and sample size $5p$. The shortest length of interval (0.02) was observed under the jackknife delete 1 method when $p = 5$ and sample size $50p$, compared to 0.25 and 0.252 observed under bootstrapping residuals and bootstrapping observation respectively.

An introduction of multicollinearity had very minimal impact on the coverage as well as in the average length of an interval across all methods. The only notable result is

the achievement of 95% coverage for $\beta_1 = 2$ which is observed under jackknife delete 1 method when $p = 5$ and $\rho = 0.9$.

6.5.3 Lognormally distributed errors

The results are showing that when independent variables are uncorrelated there is a general understatement of coverage across all methods for $\beta_1 = 0$. Bootstrapping methods are outperforming the jackknife methods in coverage while the jackknife methods are showing shorter average length of interval. The best coverage of 87% is observed under bootstrapping observations when $p = 5$ and the sample size $5p$.

For $\beta_1 = 2$, jackknife methods are showing coverage closer to the true coverage of 95% while bootstrapping methods are consistently overstating the coverage. True coverage of 95% percent were observed for both jackknife methods when $p = 1$ and sample size $25p$. Jackknife delete 2 method also achieved the true coverage of 95% for sample sizes $5p$, $10p$ and $50p$ when $p = 2$. The only case where a true coverage was not achieved in this instance was when the sample size was $25p$, even then the coverage attained is 96%.

The introduction of a correlation structure with $\rho = 0.5$ sees the improvement in the bootstrapping methods in terms of coverage and worsening of the average length of an interval. The best coverage of 90% is attained for bootstrapping observations when $p = 5$ and sample size $5p$ for $\beta_1 = 0$ while the average length of an interval increased to 8.2. For the case where for $\beta_1 = 2$ the introduction of the correlation factor among the explanatory variables did not affect the coverage achieved under the jackknife methods.

Increase of ρ from 0.5 to 0.9 resulted in a huge increase in the average length of the interval. The largest increase from 6.9 to 17 was observed for bootstrapping residuals when $p = 2$ and sample size $5p$.

6.6 Conclusion

The 95% confidence interval coverage was easily attainable due to longer average length of the confidence interval compared to the percentile based coverage where relatively shorter lengths of intervals were observed. The unsuitability of the application of confidence interval in the jackknife methods was because it produces poor coverage and wider lengths of interval. Percentile based confidence intervals are more appropriate for jackknife methods. The length of an interval decreases with an increase in sample size regardless of the presence of multicollinearity in the data. However there is no definite conclusion on the behaviour of coverage when the sample size increases. Multicollinearity in the data results in inflated length of confidence intervals. It should be noted that though in some instances jackknife methods performed better than bootstrapping methods in terms of the coverage and average length on an interval, this is not a conclusive result. The bootstrap percentile can be improved while jackknife cannot (Efron, 2003).

Chapter 7

Conclusion

In this project the Monte Carlo study compared resampling methods in terms of probability of type 1 error, power of a test as well as confidence intervals. The objective was to explore permutation tests, bootstrapping and jackknife methods as possible alternative for the parametric methods in significance testing of regression coefficients. These comparisons were performed in order to investigate the effect of (1) sample size (2) when errors are normally, uniformly and lognormally distributed (3) when the number of explanatory variables is 1, 2 and 5. (4) When the correlation coefficient between the explanatory variables is 0, 0.5 and 0.9. By examining probability of type 1 error results it was clear that parametric methods, bootstrapping residuals, bootstrapping observations, permuting observations and Freedman and Lane permutations provided reasonable probability of type 1 errors which was within $\hat{p} \pm 1.96\sqrt{\hat{p}(1-\hat{p})/n}$. The probability of type 1 error was not affected by all the factors outlined above.

The power results indicated that these methods are almost equally powerful except for bootstrapping observations method when errors are lognormally distributed where it was observed to be less powerful. One of the reasons for the failure of bootstrapping observations in certain instances was partly due to the fact that this method is appropriate for dynamic models where there is heteroscedasticity in the errors (MacKinnon, 2002), while in this study design the errors were independently identically distributed.

Based on the probability of type 1 error and power results all methods considered in the study are “equally” appropriate except for the bootstrapping observations when errors are lognormally distributed. There is no outright method that is superior.

The presence of multicollinearity did not affect the probability of type 1 error while it compromised the power of the tests. This is more prevalent when errors are normally

or lognormally distributed. When errors are uniformly distributed the impact is not as severe. In all probability of type 1 error analyses it is clear that it is safe to use any of the methods considered in this project. In this project it became clear that the inherent flaws in the data cannot be remedied by better/more robust methods i.e if there is a presence of multicollinearity in the data the use of permutation or bootstrapping does not necessarily give better results than those given by parametric methods.

In this project power of a test was investigated for an extreme case of $\beta_1 = 2$ for different sample sizes. For future research the analysis of power of a test can be performed for incremental sizes of β_1 say 0.5, 1, 1.5 and 2. In future studies it would be worth considering more extreme cases where errors are sampled from a distribution that has much higher kurtosis such a double exponential.

References

- Aldrich, J (2005). Fisher and Regression. *Statistical Science*. **20**:401-417
- Anderson, M. J. (2001). Permutation tests for univariate or multivariate analysis of variance and regression. *Canadian Journal of fisheries and aquatic sciences*.**58**: 626-639.
- Anderson, M. J and Legendre, P. (1999). An empirical comparison of permutation methods for tests of partial regression coefficients in linear model. *Journal of Computation and simulation*,**62**:271-303.
- Anderson, M. J. and Robinson, J. (2001). Permutation tests for linear models. *Aust. N. Z. J. Stat*, 43: 5-88
- Brillinger D.R. (1964). The asymptotic behaviour of Tukey's general method of setting approximate confidence interval limits (The jackknife) when applied to maximum likelihood estimates. *Review of the International Statistical Institute*, **32**:202-206
- Collins, M. F. (1987). A permutation test for planar regression. *Australian journal of statistics*, **29**: 303-308.
- Durbin J. (1959), A Note on the Application of Quenouille's Method of Bias Reduction to the Estimation of Ratios, *Biometrika*, **46**:477-480.
- Efron B. (1979), Bootstrap Methods: Another look at the Jackknife, *The Annals of Statistics*, **5**:1-26
- Efron, B. and Gong, G. (1983). A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation, *The American Statistician*, **37**:36-48
- Efron, B. and Tibshirani, R. (1993). An introduction to the Bootstrap. Chapman and Hall London
- Efron, B (2003). Second thought on the Bootstrap, *Statistical Science*, **18**:135-140.
- Fisher, R. A (1937). The design of experiments (2nd Edition), Oliver and Boyd, Edinburgh.
- Freedman, D (1981). Bootstrapping Regression Models, *Annals of Statistics*, **9**:1218-1228.
- Freedman, D. and Lane, D. (1983). A nonstochastic interpretation of reported significance levels, *Journal of business and economic statistics*, **1**: 292-298.

- Friedl, H. and Stampfer, E. Jackknife Resampling (2002), *Encyclopedia of Environmetrics*, **2**:1089-1098
- Hall P., Wilson S.R. (1991). Two guidelines for bootstrapping hypothesis testing, *Biometrics*, **47**:757-762.
- Hinkley D.V. (1988), Bootstrap Methods, *Journal of Royal Statistical Society. Series B (methodological)*, **50**:321-337.
- Huber P. J. (2002), John W. Tukey's contribution to robust statistics, *The Annals of Statistics*, **30**:1640-1648.
- Kempthorne, O. (1955). The randomisation theory of experimental inference. *Journal of the American Statistical Association*, **50**:946-967.
- Kutner M.H., Nachtsheim C.J., Neter J. and Li W. (2005), *Applied Linear Statistical Models, 5th Edition.*, McGraw-Hill.
- Kennedy, P. E. and Cade, B. S. (1996). Randomisation tests in multiple regression. *Communications in Statistics-Computation and Simulation*, **25**: 923-936.
- Kennedy, P. E. (1995). Randomisation tests in econometrics. *Journal of Business and Economic Statistics*, **13**:85-94.
- MacKinnon J.G. (2002). Bootstrap inference in econometrics, *Canadian Journal of Economics*, **35**: 615-645.
- Manly, B. F. (1991). Randomization and Monte Carlo methods in biology, 1991. *Chapman Hall, New York*.
- Matsumoto, M., & Nishimura, T. (1998). Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, **8**: 3-30.
- Miller R.G. (1974), The Jackknife- A Review, *Biometrika*, **61**:1-15.
- Neyman J.S., Dabrowska D.M., Speed T.P. (1990). On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Statist. Sci.* **5**:465-472.
- Oja, H. (1987). On permutation tests in multiple regression and analysis of covariance problems. *Australian Journal of Statistics*, **29**: 91-100.

Paparoditis E, Politis N. (2005). Bootstrap hypothesis testing in regression models, *Statistics and Probability Letters*, **74**: 356-365.

Pitman E. J. G. (1937a), Significance tests which can be applied to samples from any populations. *Supplement to the Journal of the Royal Statistical Society*,**4**:119-130

Pitman E. J. G. (1937b), Significance tests which can be applied to samples from any populations II. The correlation coefficient test, *Supplement to the Journal of the Royal Statistical Society*, **4**:225-232

Pitman E. J. G. (1938), Significance tests which can be applied to samples from any populations III. The analysis of variance test, *Biometrika*, **29**:322-335

Quenouille M.H. (1949), Approximate tests of correlation in time-series. *Journal of Royal Statistical Society. Series B (Methodological)*, **11**:68-84

Quenouille M.H. (1956), Notes on Bias in Estimation, *Biometrika*, **43**:353-360

Rao J. N. K (1965), A Note on Estimation of Ratios by Quenouille's Method, *Biometrika*, **52**:647-649

Rao J.N.K, Webster J.T (1966). On two methods of bias reduction in the estimation of ratios, *Biometrika*, **53**:571-577

Rubin, D.B. (1990). On the application of probability theory to agricultural experiments. Essay on principles. Section 9. Comment: Neyman(1923) and Causal inference in experiments and observational studies. *Statist. Sci.* **5**:472-480.

Sandved, E. (1966). A principle for conditioning on an ancillary statistic. *Scandinavian Actuarial Journal*, **50**: 39-47.

Schucany W.R, Gray H.L, Owen D.B. (1971) On bias reduction in estimation. *Journal of the American Statistical Association*, **66**:524-533

Sahinler,S,Topuz,D. (2007). Bootstrap and Jackknife resampling algorithms for estimation of regression parameters. *Journal of Applied Quantitative Methods*, **2**:188-199

te Braak C.J.F (1992). Permutation versus bootstraps significance tests in multiple regression and ANOVA. In: *Bootstrapping and Related techniques* (Eds., Jockel, H.-H, Rothe, G and Sendler,W.) pp 79-86. Berlin: Springer-Verlag.

Tukey J.W (1958). Bias and confidence in not-quite large samples. *Annals of Mathematical Statistics*, **29**: 614.

Welch, J. W. (1990). Construction of permutation tests. *Journal of the American Statistical Association*, **85**: 693-698.

Wu C. F. J. (1986), Jackknife, bootstrap and other resampling methods in regression analysis. *The annals of Statistics*, **14**: 1261-1295

Appendix A - Probability of type 1 error

Type1 error when errors are normally distributed

p	Rho	n	Sample size	Parametric	Bootstrapping residuals	Bootstrapping observations	Permuting observations	Freedman & Lane Permutation
1	0	5p	5	0.0488	0.0452		0.0434	0.043
		10p	10	0.0498	0.0496	0.043	0.0516	0.049
		25p	25	0.0508	0.0498	0.058	0.0502	0.0532
		50p	50	0.0476	0.0476	0.053	0.0476	0.0474
2	0	5p	10	0.053	0.0532		0.0526	0.0524
		10p	20	0.0514	0.051	0.0524	0.0512	0.0506
		25p	50	0.0578	0.0566	0.058	0.0572	0.058
		50p	100	0.0476	0.0478	0.049	0.048	0.0478
	0.5	5p	10	0.056	0.056		0.0562	0.056
		10p	20	0.0544	0.0526	0.0536	0.0542	0.0538
		25p	50	0.0554	0.054	0.0594	0.0562	0.0546
		50p	100	0.0526	0.051	0.0558	0.0532	0.0528
	0.9	5p	10	0.0534	0.0542		0.054	0.0544
		10p	20	0.053	0.0522	0.0536	0.0512	0.053
		25p	50	0.0516	0.05	0.058	0.0506	0.0512
		50p	100	0.052	0.0518	0.0522	0.0502	0.0512
5	0	5p	25	0.0456	0.046	0.0308	0.046	0.0482
		10p	50	0.0462	0.0466	0.048	0.0452	0.0456
		25p	125	0.0544	0.0546	0.0558	0.0534	0.056
		50p	250	0.0502	0.0494	0.0512	0.0494	0.0498
	0.5	5p	25	0.0528	0.053	0.0402	0.0514	0.0532
		10p	50	0.0486	0.0498	0.054	0.0494	0.051
		25p	125	0.0534	0.0542	0.0542	0.055	0.053
		50p	250	0.0474	0.049	0.0508	0.0472	0.0478
	0.9	5p	25	0.049	0.0492	0.038	0.049	0.0474
		10p	50	0.051	0.0494	0.054	0.0484	0.05
		25p	125	0.0512	0.052	0.051	0.0518	0.0526
		50p	250	0.0514	0.052	0.0514	0.0514	0.0518

Table 7: Type 1 error when errors are normally distributed

Type 1 error when errors are uniformly distributed

p	Rho	n	Sample size	Parametric	Bootstrapping residuals	Bootstrapping observations	Permuting observations	Freedman & Lane Permutation
1	0	5p	5	0.0454	0.0406		0.0388	0.0402
		10p	10	0.0524	0.0538	0.0596	0.0502	0.0524
		25p	25	0.0488	0.0488	0.0614	0.0476	0.0468
		50p	50	0.0576	0.0578	0.0636	0.056	0.0574
2	0	5p	10	0.054	0.054		0.0516	0.0506
		10p	20	0.0516	0.0522	0.0606	0.0522	0.0518
		25p	50	0.0564	0.0542	0.0606	0.0566	0.0552
		50p	100	0.0482	0.0492	0.051	0.0476	0.0488
	0.5	5p	10	0.0574	0.0574		0.0568	0.0552
		10p	20	0.0508	0.0488	0.0706	0.0508	0.0488
		25p	50	0.051	0.0516	0.0592	0.0524	0.051
		50p	100	0.052	0.0526	0.0546	0.0526	0.0528
	0.9	5p	10	0.0542	0.0534		0.0524	0.0544
		10p	20	0.0526	0.0514	0.065	0.0524	0.052
		25p	50	0.0488	0.0484	0.0558	0.0472	0.047
		50p	100	0.051	0.0514	0.0564	0.0514	0.0516
5	0	5p	25	0.0512	0.0506	0.0386	0.0508	0.0526
		10p	50	0.0456	0.0464	0.046	0.0464	0.0462
		25p	125	0.051	0.05	0.0534	0.0506	0.05
		50p	250	0.0486	0.0476	0.0484	0.048	0.0484
	0.5	5p	25	0.0526	0.0524	0.0454	0.0522	0.0534
		10p	50	0.0486	0.0486	0.052	0.047	0.0472
		25p	125	0.0468	0.0464	0.0518	0.0486	0.0478
		50p	250	0.0456	0.0464	0.0492	0.0448	0.044
	0.9	5p	25	0.0508	0.052	0.0474	0.0506	0.0516
		10p	50	0.0472	0.0466	0.0542	0.0464	0.0488
		25p	125	0.0476	0.0476	0.05	0.0494	0.0466
		50p	250	0.0496	0.0502	0.0524	0.0492	0.0512

Table 8 :Type 1 error when errors are uniformly distributed

Type1 error when errors are lognormally distributed

p	Rho	n	Sample size	Parametric	Bootstrapping residuals	Bootstrapping observations	Permuting observations	Freedman & Lane Permutation
1	0	5p	5	0.0514	0.0436		0.0464	0.0472
		10p	10	0.0448	0.05	0.0204	0.0504	0.0526
		25p	25	0.049	0.0532	0.0236	0.056	0.0554
		50p	50	0.0446	0.0464	0.0216	0.0488	0.0492
2	0	5p	10	0.0536	0.0534		0.0572	0.056
		10p	20	0.0432	0.0454	0.018	0.0492	0.0486
		25p	50	0.046	0.0498	0.0256	0.0512	0.0504
		50p	100	0.0432	0.0466	0.0266	0.0464	0.0464
	0.5	5p	10	0.052	0.053		0.0566	0.0562
		10p	20	0.048	0.0498	0.018	0.0512	0.0516
		25p	50	0.0486	0.0516	0.0282	0.052	0.0518
		50p	100	0.0472	0.0482	0.0284	0.0484	0.0488
	0.9	5p	10	0.0514	0.051		0.0574	0.0564
		10p	20	0.0438	0.0446	0.021	0.0486	0.0488
		25p	50	0.0478	0.0476	0.0204	0.0508	0.0486
		50p	100	0.05	0.0516	0.0294	0.0506	0.0518
5	0	5p	25	0.0494	0.0498	0.0106	0.053	0.0492
		10p	50	0.0448	0.0466	0.021	0.0488	0.047
		25p	125	0.0478	0.048	0.03	0.0512	0.0496
		50p	250	0.0454	0.0466	0.039	0.0486	0.048
	0.5	5p	25	0.0488	0.0512	0.0136	0.0488	0.0502
		10p	50	0.052	0.0502	0.0242	0.0518	0.052
		25p	125	0.051	0.0526	0.0342	0.0518	0.0494
		50p	250	0.0464	0.0458	0.0338	0.0464	0.047
	0.9	5p	25	0.0486	0.0484	0.015	0.0504	0.0508
		10p	50	0.0512	0.0518	0.025	0.0516	0.053
		25p	125	0.0524	0.05	0.031	0.0546	0.0554
		50p	250	0.0452	0.0464	0.0378	0.0444	0.0464

Table 9: Type 1 error when errors are lognormally distributed

Appendix B – Power of a test

Power when errors are normally distributed

p	Rho	n	Sample size	Parametric	Bootstrapping residuals	Bootstrapping observations	Permuting observations	Freedman & Lane Permutation
1	0	5p	5	0.1336	0.1162		0.112	0.114
		10p	10	0.324	0.3222	0.2782	0.3188	0.3224
		25p	25	0.7618	0.7592	0.7544	0.7602	0.759
		50p	50	0.9734	0.9734	0.9726	0.9734	0.9724
2	0	5p	10	0.2782	0.279		0.2782	0.2758
		10p	20	0.6338	0.63	0.6002	0.632	0.6316
		25p	50	0.969	0.9686	0.9636	0.9678	0.968
		50p	100	1	1	0.9998	1	1
	0.5	5p	10	0.226	0.2258		0.2282	0.228
		10p	20	0.5178	0.5172	0.5002	0.5132	0.5158
		25p	50	0.9126	0.912	0.907	0.9104	0.9122
		50p	100	0.9976	0.9978	0.9974	0.997	0.9976
	0.9	5p	10	0.0946	0.0964		0.0942	0.0982
		10p	20	0.177	0.1752	0.1732	0.1756	0.173
		25p	50	0.3984	0.3962	0.399	0.3992	0.399
		50p	100	0.6924	0.688	0.6938	0.6896	0.6934
5	0	5p	25	0.6806	0.6758	0.5866	0.679	0.68
		10p	50	0.9646	0.9632	0.9576	0.9628	0.9644
		25p	125	1	1	1	1	1
		50p	250	1	1	1	1	1
	0.5	5p	25	0.4872	0.4878	0.4048	0.4884	0.4866
		10p	50	0.831	0.8302	0.816	0.8312	0.8314
		25p	125	0.9978	0.9976	0.9976	0.998	0.998
		50p	250	1	1	1	1	1
	0.9	5p	25	0.147	0.1462	0.111	0.1444	0.1444
		10p	50	0.2774	0.275	0.2766	0.2764	0.2752
		25p	125	0.6102	0.6058	0.6092	0.606	0.6092
		50p	250	0.8902	0.8888	0.8858	0.887	0.8886

Table 10: Power of a test when errors are normally distributed

Power when errors are uniformly distributed

p	Rho	n	Sample size	Parametric	Bootstrapping residuals	Bootstrapping observations	Permuting observations	Freedman & Lane Permutation
1	0	5p	5	0.6798	0.6236		0.5742	0.5764
		10p	10	0.9928	0.9926	0.9776	0.993	0.9928
		25p	25	1	1	1	1	1
		50p	50	1	1	1	1	1
2	0	5p	10	0.9852	0.9854		0.9856	0.9848
		10p	20	1	1	1	1	1
		25p	50	1	1	1	1	1
		50p	100	1	1	1	1	1
	0.5	5p	10	0.9574	0.9558		0.9536	0.9554
		10p	20	1	0.9998	0.9998	0.9998	1
		25p	50	1	1	1	1	1
		50p	100	1	1	1	1	1
	0.9	5p	10	0.5218	0.5144		0.51	0.5174
		10p	20	0.9222	0.9202	0.908	0.9208	0.9216
		25p	50	1	1	1	1	1
		50p	100	1	1	1	1	1
5	0	5p	25	1	1	1	1	1
		10p	50	1	1	1	1	1
		25p	125	1	1	1	1	1
		50p	250	1	1	1	1	1
	0.5	5p	25	1	1	0.9998	1	1
		10p	50	1	1	1	1	1
		25p	125	1	1	1	1	1
		50p	250	1	1	1	1	1
	0.9	5p	25	0.8308	0.8298	0.7592	0.8266	0.8296
		10p	50	0.9946	0.9944	0.9942	0.9946	0.9948
		25p	125	1	1	1	1	1
		50p	250	1	1	1	1	1

Table 11 : Power of a test when errors are uniformly distributed

Power when errors are lognormally distributed

p	Rho	n	Sample size	Parametric	Bootstrapping residuals	Bootstrapping observations	Permuting observations	Freedman & Lane Permutation
1	0	5p	5	0.1398	0.1318		0.1308	0.1334
		10p	10	0.2354	0.2438	0.154	0.251	0.2482
		25p	25	0.4166	0.4254	0.3718	0.4254	0.4282
		50p	50	0.591	0.5954	0.56	0.598	0.6016
2	0	5p	10	0.2138	0.2154		0.2192	0.2166
		10p	20	0.3534	0.3584	0.2856	0.3628	0.3594
		25p	50	0.5686	0.5744	0.5368	0.5778	0.5762
		50p	100	0.7812	0.7866	0.7644	0.7886	0.789
	0.5	5p	10	0.186	0.183		0.1904	0.193
		10p	20	0.2902	0.291	0.229	0.2924	0.2938
		25p	50	0.4776	0.4804	0.453	0.4834	0.4814
		50p	100	0.6942	0.6946	0.6904	0.693	0.6982
	0.9	5p	10	0.082	0.0814		0.0852	0.0858
		10p	20	0.1088	0.1062	0.0644	0.1094	0.1124
		25p	50	0.1772	0.179	0.1532	0.1794	0.1814
		50p	100	0.27	0.2732	0.2646	0.2704	0.2742
5	0	5p	25	0.3594	0.3612	0.25	0.3676	0.3638
		10p	50	0.5632	0.5642	0.5204	0.5672	0.5668
		25p	125	0.8414	0.845	0.8182	0.8456	0.848
		50p	250	0.9692	0.97	0.9536	0.9706	0.9708
	0.5	5p	25	0.2546	0.2534	0.1584	0.2556	0.2562
		10p	50	0.4056	0.4048	0.3816	0.4056	0.4088
		25p	125	0.6908	0.6898	0.6816	0.691	0.6924
		50p	250	0.8944	0.8932	0.8742	0.8932	0.8952
	0.9	5p	25	0.094	0.0946	0.041	0.0944	0.0938
		10p	50	0.1298	0.1318	0.0982	0.1326	0.1328
		25p	125	0.236	0.2364	0.2222	0.2364	0.2362
		50p	250	0.3496	0.3504	0.3614	0.348	0.347

Table 12: Power of a test when errors are lognormally distributed

Appendix C – SAS Programs

Probability of type 1 and Power Program

```

* Masters dissertation program;

libname nhla "C:\Users\Sibu\Documents\Nhlanhla";

/*Appendix C – SAS Programs*/
/**/
/*Type 1 and Power Program*/
/** Masters dissertation program;*/

proc iml;

*starting the standard regression analysis;

  start regress1 (x,y)
Global (t1,pvalue_b1,F,crit,LCLOri,UCLOri,dist_Ori,betaOri); /*
begin module */
  alpha=0.05;

  xpxi=inv(t(x)*x);          /* inverse of X'X          */
  beta=xpxi*(t(x)*y);       /* parameter estimate     */
  yhat=x*beta;              /* predicted values       */
  resid=y-yhat;             /* residuals              */
  sse=ssq(resid);           /* SSE                    */
  n_0=nrow(x);              /* sample size            */
  dfe=nrow(x)-ncol(x);     /* error DF               */
  mse=sse/dfe;              /* MSE                    */
  cssy=ssq(y-sum(y)/n_0);   /* corrected total SS    */
  rsquare=(cssy-sse)/cssy;  /* RSQUARE                */
  ssr=t(beta)*xpxi*beta;
  msr=ssr/(ncol(x)-1);
  F=msr/mse;
  stdb=sqrt(vecdiag(xpxi)*mse); /* std of estimates     */
  t=beta/stdb; /* parameter t tests */

  t1=t[2];
  prob=1-probf(t#t,1,dfe); /* p-values             */
  pvalue_b1=prob[2];

finish regress1; /* end module */
store module=(regress1);

*-----;

```

```

/* The beginning of Hypothesis testing using bootstrapping and
Permutation test*/

*bootstrapping residuals;

start regress_br(x,y,n) Global(t1,pvalue_br,F); /*
begin module */
n=nrow(x);
rep=999;
t1_br=j(rep,1,1);
beta1_br=J(rep,1,1);
stdb1_br=J(rep,1,1);
f=j(n,1,1);

p_br=0;
pf=0;

k=j(n,1);

xpxi=inv(t(x)*x); /* inverse of X'X */
beta=xpxi*(t(x)*y); /* parameter estimate */
beta[2,1]=0; /*under null hypothesis beta1=0*/
yhat=x*beta; /* predicted values */
resid=y-yhat;

do j=1 to rep;

call randgen(k,"uniform");
f=ceil(k#n);

v=resid[f];

ybr=yhat+v;

xbrpxbri=inv(t(x)*x); /* inverse of X'X */
beta_br=xbrpxbri*(t(x)*ybr); /* parameter estimate */
ybrhat=x*beta_br; /* predicted values */
resid_br=ybr-ybrhat; /* residuals */
*/
sse_br=ssq(resid_br); /* SSE */
*/
dfe_br=nrow(x)-ncol(x); /* error DF */
*/
mse_br=sse_br/dfe_br; /* MSE */
*/
ssr=t(beta_br)*xbrpxbri*beta_br;
msr=ssr/(ncol(x)-1);
Fbr=msr/mse_br;
stdb_br=sqrt(vecdiag(xbrpxbri)*mse_br); /* std of estimates */
*/
t_br=beta_br/stdb_br; /* parameter t tests */
*/

```

```

t1_br[j]=t_br[2];
beta1_br[j]=beta_br[2] ;
stdb1_br[j]=stdb_br[2] ;

if abs(t1_br[j]) >= abs(t1) then p_br=p_br+1;
if Fbr>=F then pf=pf+1;

end;

pvalue_br=p_br/rep;
pvalueFbr=pf/rep;

finish regress_br;          /* end module          */
store module=(regress_br);

*-----
*-----

*bootstrapping observations;

start regress_b(x,y,n)Global(t1,pvalue_b,F); /* begin module */
n=nrow(x);
h=j(n,1,1);
rep=999;
t1_b=j(rep,1,1);
*Fb=j(rep,1,1);
beta1_b=J(rep,1,1);
stdb1_b=J(rep,1,1);

p_b=0;
pf=0;

v=j(n,1);

xpxi=inv(t(x)*x);          /* inverse of X'X          */
beta=xpxi*(t(x)*y);

do j=1 to rep;
call randgen(v,"uniform");
h=ceil(v#n);

xb=x[h,];
yb=y[h,];

xpxbi=inv(t(xb)*xb);          /* inverse of X'X          */
beta_b=xpxbi*(t(xb)*yb);      /* parameter estimate      */
ybhat=xb*beta_b;              /* predicted values        */
resid_b=yb-ybhat;             /* residuals                */
sse_b=ssq(resid_b);           /* SSE                      */
n_b=nrow(x);                  /* sample size              */

```

```

    dfe_b=nrow(x)-ncol(x);          /* error DF          */
    mse_b=sse_b/dfe_b;             /* MSE              */
*/
    ssr=t(beta_b)*xbpxbi*beta_b;
    msr=ssr/(ncol(x)-1);
    Fb=msr/mse_b;

    stdb_b=sqrt(vecdiag(xbpxbi)*mse_b); /* std of estimates
*/
    t_b=(beta_b-beta)/stdb_b;        /* parameter t
tests */

    t1_b[j]=t_b[2];
    betal_b[j]=beta_b[2];
    stdbl_b[j]=stdb_b[2];

    if abs(t1_b[j]) >= abs(t1) then p_b=p_b+1;
    if Fb>=F then pf=pf+1;

end;
    pvalue_b=p_b/rep;
    pvalueFb=pf/rep;

finish regress_b; /* end module */

*-----
-----;
*permutation tests;

start regress_p(x,y,n) Global (t1,pvalue_p,F); /* begin module */
perm=999;

u = j(n,1);          /* allocate vector */

t1_p=j(perm,1,1);
*Fp=j(perm,1,1);
betal_p=J(perm,1,1);
stdbl_p=J(perm,1,1);

p_p=0;
pf=0;

do i = 1 to perm;

    call randgen(u, "uniform"); /* fill u with random uniform */

    yp = y[ rank(u) ];          /* permute y use RANK(u) to permute
data */

```

```

xpxi=inv(t(x)*x);          /* inverse of X'X          */
beta_p=xpxi*(t(x)*yp);    /* parameter estimate     */
yphat=x*beta_p;          /* predicted values       */
  resid=yp-yphat;        /* residuals              */
  sse=ssq(resid);        /* SSE                     */
  n_p=nrow(x);          /* sample size            */
  dfe=nrow(x)-ncol(x);  /* error DF               */
  mse=sse/dfe;          /* MSE                     */
  cssyp=ssq(yp-sum(yp)/n_p); /* corrected total SS    */
  rsquare=(cssyp-sse)/cssyp; /* RSQUARE                */
  ssr=t(beta_p)*xpxi*beta_p;
  msr=ssr/(ncol(x)-1);
  Fp=msr/mse;
  stdb_p=sqrt(vecdiag(xpxi)*mse); /* std of estimates      */
  t_p=beta_p/stdb_p;      /* parameter t tests     */
*/

  t1_p[i]=t_p[2];
  betal_p[i]=beta_p[2];
  stdbl_p[i]=stdb_p[2];

  if abs(t1_p[i]) >= abs(t1) then p_p=p_p+1;
  if Fp>=F then pf=pf+1;

  end;
  pvalue_p=p_p/perm;
  pvalueFp=pf/perm;

finish regress_p;          /* end module            */

*-----
-----;

start delcol(x,i);        /*starting the function to delete column*/
return(x[,setdif(1:ncol(x),i)]);
finish;

start regress_FL(x,y,n) Global (t1,pvalue_FL,F); /* begin module */
perm=999;

u = j(n,1);              /* allocate vector */

t1_FL=j(perm,1,1);
betal_FL=J(perm,1,1);
stdbl_FL=J(perm,1,1);

X1=x[,1:2];
x2=delcol(x,2);

p_p=0;
pf=0;

```



```

do i = 1 to perm;

    xpxi=inv(t(x2)*x2);          /* inverse of X'X          */
    beta=xpxi*(t(x2)*y);        /* parameter estimate      */
    yhat=x2*beta;               /* predicted values        */
    resid=y-yhat;               /* residuals                */

    call randgen(u, "uniform"); /* fill u with random uniform */

    resid_p = resid[ rank(u) ]; /* permute residuals use
RANK(u) to permute data */

    yp=x2*beta+resid_p;

    xpxi=inv(t(x)*x);          /* inverse of X'X          */
    beta_p=xpxi*(t(x)*yp);     /* parameter estimate      */
    yphat=x*beta_p;           /* predicted values        */
    resid=yp-yphat;           /* residuals                */
    sse=ssq(resid);           /* SSE                      */
    n_p=nrow(x);              /* sample size              */
    dfe=nrow(x)-ncol(x);     /* error DF                 */
    mse=sse/dfe;              /* MSE                      */
    cssyp=ssq(yp-sum(yp)/n_p); /* corrected total SS      */
    rsquare=(cssyp-sse)/cssyp; /* RSQUARE                  */
    ssr=t(beta_p)*xpxi*beta_p;
    msr=ssr/(ncol(x)-1);
    FpFL=msr/mse;
    stdb_p=sqrt(vecdiag(xpxi)*mse); /* std of estimates      */
    t_p=beta_p/stdb_p;        /* parameter t tests      */
*/

    t1_FL[i]=t_p[2];
    beta1_FL[i]=beta_p[2];
    stdb1_FL[i]=stdb_p[2];

    if abs(t1_FL[i]) >= abs(t1) then p_p=p_p+1;
    if FpFL >=F then pf=pf+1;

end;
    pvalue_FL=p_p/perm;
    pvalueFpFL=pf/perm;

    finish regress_FL;          /* end module
*/

store module=regress1;
store module=(regress_br regress_b regress_p regress_FL delcol);
store module=(regress_br);
quit;

```

```

%macro ssize(size,er,h0,p,rho,pr);

proc iml;
load module=regress1;

no=5000;

/*pvalues declarations*/
pval_N&size=j(no,1,.);
pval_br&size=j(no,1,.);
pval_b&size=j(no,1,.);
pval_p&size=j(no,1,.);
pval_FL&size=j(no,1,.);

meth="&er";
alpha=0.05;
h0="&h0"; /*Null hypothesis*/

call randseed(12);

do k=1 to no;

    n=&p*&size;      /* sample size */
/*    A={10,0,20,30,40,50};*/
        b=j(&p+1,1,1);
/*    b=A;*/

    *b[1]=0;
    if h0="true" then b[2]=0;
    else b[2]=2;

    R=j(&p,&p,1);    *correlation matrix;

    if &p > 1 then do;
        do i=1 to &p;
            do j=1 to &p;
                if i ^= j then R[i,j]=&rho;
            end;
        end;
    end;

    e&size.&er=j(n,1,.);
    w=J(n,&p,1);

    call randgen(e&size.&er,"&er");
    call randgen(w,"Uniform");
    x=j(n,1,1)||w*root(R);

```

```

    y = x*b+e&size.&er;

    *reset noprint;
    run regress1(x,y);
    run regress_br(x,y,n);
    if (&size>5 & (&p=1 | &p=2))|&p=5 then run regress_b(x,y,n);
    run regress_p(x,y,n);
    run regress_FL(x,y,n);

    pval_N&size[k]=pvalue_b1;
    pval_br&size[k]=pvalue_br;
    if (&size>5 & (&p=1 | &p=2))|&p=5 then
pval_b&size[k]=pvalue_b;
    pval_p&size[k]=pvalue_p;
    pval_FL&size[k]=pvalue_FL;

    end;

    if h0="true" then do;

        type1_N&size=sum(pval_N&size<alpha)/no;
        if pval_br&size=. then type1_br&size=.; else
    type1_br&size=sum(pval_br&size<alpha)/no;
        if (&size>5 & (&p=1 | &p=2))|&p=5 then
    type1_b&size=sum(pval_b&size<alpha)/no;else type1_b&size=.;
        type1_p&size=sum(pval_p&size<alpha)/no;
        type1_FL&size=sum(pval_FL&size<alpha)/no;

    end;

    else do;
        power_N&size=sum(pval_N&size<alpha)/no;
        if pval_br&size=. then power_br&size=.; else
    power_br&size=sum(pval_br&size<alpha)/no;
        if (&size>5 & (&p=1 | &p=2))|&p=5 then
    power_b&size=sum(pval_b&size<alpha)/no;else power_b&size=.;
        power_p&size=sum(pval_p&size<alpha)/no;
        power_FL&size=sum(pval_FL&size<alpha)/no;

    end;

    if meth="Normal" then error=0;
    else if meth="Uniform" then error=1;
    else error=2;

    if h0="false" then do;

        pw&size=error||&p||&size||n||&rho||power_N&size||power_br&size
|power_b&size||power_p&size||power_FL&size;

```

```

    pn={"Error", "no of
par", "size", "Ssize", "Rho", "PowerNormal", "PowerBootR", "PowerBootO", "P
owerPerm", "PowerFL"};
    create nhla.power&size.&er.&pr from    pw&size[colname=pn];
    append from    pw&size;

end;

else do;

    tp&size=error||&p||&size||n||&rho||type1_N&size||type1_br&size
||type1_b&size||type1_p&size||type1_FL&size;

    tn={"Error", "no of
par", "size", "Ssize", "Rho", "Type1Normal", "Type1BootR", "Type1BootO", "T
ypelPerm", "Type1FL"};
    create nhla.type1&size.&er.&pr from    tp&size[colname=tn];
    append from    tp&size;

end;

quit;
%mend ssize;
%ssize(5,Normal,true,1,0,0);
%ssize(10,Normal,true,1,0,0);
%ssize(25,Normal,true,1,0,0);
%ssize(50,Normal,true,1,0,0);

%ssize(5,Uniform,true,1,0,0);
%ssize(10,Uniform,true,1,0,0);
%ssize(25,Uniform,true,1,0,0);
%ssize(50,Uniform,true,1,0,0);

%ssize(5,LOGN,true,1,0,0);
%ssize(10,LOGN,true,1,0,0);
%ssize(25,LOGN,true,1,0,0);
%ssize(50,LOGN,true,1,0,0);

%ssize(5,Normal,false,1,0,0);
%ssize(10,Normal,false,1,0,0);
%ssize(25,Normal,false,1,0,0);
%ssize(50,Normal,false,1,0,0);

%ssize(5,Uniform,false,1,0,0);
%ssize(10,Uniform,false,1,0,0);
%ssize(25,Uniform,false,1,0,0);
%ssize(50,Uniform,false,1,0,0);

%ssize(5,LOGN,false,1,0,0);
%ssize(10,LOGN,false,1,0,0);
%ssize(25,LOGN,false,1,0,0);
%ssize(50,LOGN,false,1,0,0);

```

```

/*ods html
body="C:\Users\nhlanhla\Documents\important\dessertation\results"*/
/*
gpath="C:\Users\nhlanhla\Documents\important\dessertation\results\"
style=journal;*/
%macro comp(er,rho,p);

data nhla.power_&er.&rho;
length n $3;
set nhla.power5&er.&rho; nhla.power10&er.&rho; nhla.power25&er.&rho;
nhla.power50&er.&rho;
if size=5 then n="5p";
else if size=10 then n="10p";
else if size=25 then n="25p";
else if size=50 then n="50p";

run;

data nhla.type1_&er.&rho;
length n $3;
set nhla.type15&er.&rho; nhla.type110&er.&rho; nhla.type125&er.&rho;
nhla.type150&er.&rho;
if size=5 then n="5p";
else if size=10 then n="10p";
else if size=25 then n="25p";
else if size=50 then n="50p";
run;

proc datasets;
delete type15&er.&rho; type110&er.&rho; type125&er.&rho;
type150&er.&rho; power5&er.&rho; power10&er.&rho; power25&er.&rho;
power50&er.&rho;

%mend comp;
%comp (Normal,0,1);
%comp (Uniform,0,1);
%comp (LOGN,0,1);

data nhla.type1_p1;
set nhla.type1_Normal0 nhla.type1_Uniform0 nhla.type1_LOGN0;
run;
data nhla.power_p1;
set nhla.power_Normal0 nhla.power_Uniform0 nhla.power_LOGN0;
run;

ods html close;
*p=2;

%ssize(5,Normal,true,2,0,0);
%ssize(10,Normal,true,2,0,0);
%ssize(25,Normal,true,2,0,0);
%ssize(50,Normal,true,2,0,0);

```

```
%ssize(5,Uniform,true,2,0,0);
%ssize(10,Uniform,true,2,0,0);
%ssize(25,Uniform,true,2,0,0);
%ssize(50,Uniform,true,2,0,0);
```

```
%ssize(5,LOGN,true,2,0,0);
%ssize(10,LOGN,true,2,0,0);
%ssize(25,LOGN,true,2,0,0);
%ssize(50,LOGN,true,2,0,0);
```

```
%ssize(5,Normal,false,2,0,0);
%ssize(10,Normal,false,2,0,0);
%ssize(25,Normal,false,2,0,0);
%ssize(50,Normal,false,2,0,0);
```

```
%ssize(5,Uniform,false,2,0,0);
%ssize(10,Uniform,false,2,0,0);
%ssize(25,Uniform,false,2,0,0);
%ssize(50,Uniform,false,2,0,0);
```

```
%ssize(5,LOGN,false,2,0,0);
%ssize(10,LOGN,false,2,0,0);
%ssize(25,LOGN,false,2,0,0);
%ssize(50,LOGN,false,2,0,0);
```

```
%comp(Normal,0,2);
%comp (Uniform,0,2);
%comp (LOGN,0,2);
```

```
*-----;
-----;
```

```
%ssize(5,Normal,true,2,0.5,50);
%ssize(10,Normal,true,2,0.5,50);
%ssize(25,Normal,true,2,0.5,50);
%ssize(50,Normal,true,2,0.5,50);
```

```
%ssize(5,Uniform,true,2,0.5,50);
%ssize(10,Uniform,true,2,0.5,50);
%ssize(25,Uniform,true,2,0.5,50);
%ssize(50,Uniform,true,2,0.5,50);
```

```
%ssize(5,LOGN,true,2,0.5,50);
%ssize(10,LOGN,true,2,0.5,50);
%ssize(25,LOGN,true,2,0.5,50);
%ssize(50,LOGN,true,2,0.5,50);
```

```
%ssize(5,Normal,false,2,0.5,50);
%ssize(10,Normal,false,2,0.5,50);
%ssize(25,Normal,false,2,0.5,50);
%ssize(50,Normal,false,2,0.5,50);
```

```
%ssize(5,Uniform,false,2,0.5,50);
%ssize(10,Uniform,false,2,0.5,50);
%ssize(25,Uniform,false,2,0.5,50);
%ssize(50,Uniform,false,2,0.5,50);
```

```
%ssize(5,LOGN,false,2,0.5,50);
%ssize(10,LOGN,false,2,0.5,50);
%ssize(25,LOGN,false,2,0.5,50);
%ssize(50,LOGN,false,2,0.5,50);
```

```
*-----;
```

```
%comp(Normal,50,2);
%comp (Uniform,50,2);
%comp (LOGN,50,2);
```

```
*-----
-----;
```

```
%ssize(5,Normal,true,2,0.9,90);
%ssize(10,Normal,true,2,0.9,90);
%ssize(25,Normal,true,2,0.9,90);
%ssize(50,Normal,true,2,0.9,90);
```

```
%ssize(5,Uniform,true,2,0.9,90);
%ssize(10,Uniform,true,2,0.9,90);
%ssize(25,Uniform,true,2,0.9,90);
%ssize(50,Uniform,true,2,0.9,90);
```

```
%ssize(5,LOGN,true,2,0.9,90);
%ssize(10,LOGN,true,2,0.9,90);
%ssize(25,LOGN,true,2,0.9,90);
%ssize(50,LOGN,true,2,0.9,90);
```

```
%ssize(5,Normal,false,2,0.9,90);
%ssize(10,Normal,false,2,0.9,90);
%ssize(25,Normal,false,2,0.9,90);
%ssize(50,Normal,false,2,0.9,90);
```

```
%ssize(5,Uniform,false,2,0.9,90);
%ssize(10,Uniform,false,2,0.9,90);
%ssize(25,Uniform,false,2,0.9,90);
%ssize(50,Uniform,false,2,0.9,90);
```

```
%ssize(5,LOGN,false,2,0.9,90);
%ssize(10,LOGN,false,2,0.9,90);
%ssize(25,LOGN,false,2,0.9,90);
%ssize(50,LOGN,false,2,0.9,90);
```

```

%comp(Normal,90,2);
%comp (Uniform,90,2);
%comp (LOGN,90,2);

data nhla.type1_p2;
set nhla.type1_Normal0 nhla.type1_Uniform0 nhla.type1_LOGN0
nhla.type1_Normal50 nhla.type1_Uniform50 nhla.type1_LOGN50
nhla.type1_Normal90 nhla.type1_Uniform90 nhla.type1_LOGN90;
run;
data nhla.power_p2;
set nhla.power_Normal0 nhla.power_Uniform0 nhla.power_LOGN0
nhla.power_Normal50 nhla.power_Uniform50 nhla.power_LOGN50
nhla.power_Normal90 nhla.power_Uniform90 nhla.power_LOGN90;
run;
*p=5;

%ssize(5,Normal,true,5,0,0);
%ssize(10,Normal,true,5,0,0);
%ssize(25,Normal,true,5,0,0);
%ssize(50,Normal,true,5,0,0);

%ssize(5,Uniform,true,5,0,0);
%ssize(10,Uniform,true,5,0,0);
%ssize(25,Uniform,true,5,0,0);
%ssize(50,Uniform,true,5,0,0);

%ssize(5,LOGN,true,5,0,0);
%ssize(10,LOGN,true,5,0,0);
%ssize(25,LOGN,true,5,0,0);
%ssize(50,LOGN,true,5,0,0);

%ssize(5,Normal,false,5,0,0);
%ssize(10,Normal,false,5,0,0);
%ssize(25,Normal,false,5,0,0);
%ssize(50,Normal,false,5,0,0);

%ssize(5,Uniform,false,5,0,0);
%ssize(10,Uniform,false,5,0,0);
%ssize(25,Uniform,false,5,0,0);
%ssize(50,Uniform,false,5,0,0);

%ssize(5,LOGN,false,5,0,0);
%ssize(10,LOGN,false,5,0,0);
%ssize(25,LOGN,false,5,0,0);
%ssize(50,LOGN,false,5,0,0);

%comp(Normal,0,5);
%comp (Uniform,0,5);
%comp (LOGN,0,5);

```



```

*-----
-----;

%ssize(5,Normal,true,5,0.5,50);
%ssize(10,Normal,true,5,0.5,50);
%ssize(25,Normal,true,5,0.5,50);
%ssize(50,Normal,true,5,0.5,50);

%ssize(5,Uniform,true,5,0.5,50);
%ssize(10,Uniform,true,5,0.5,50);
%ssize(25,Uniform,true,5,0.5,50);
%ssize(50,Uniform,true,5,0.5,50);

%ssize(5,LOGN,true,5,0.5,50);
%ssize(10,LOGN,true,5,0.5,50);
%ssize(25,LOGN,true,5,0.5,50);
%ssize(50,LOGN,true,5,0.5,50);

%ssize(5,Normal,false,5,0.5,50);
%ssize(10,Normal,false,5,0.5,50);
%ssize(25,Normal,false,5,0.5,50);
%ssize(50,Normal,false,5,0.5,50);

%ssize(5,Uniform,false,5,0.5,50);
%ssize(10,Uniform,false,5,0.5,50);
%ssize(25,Uniform,false,5,0.5,50);
%ssize(50,Uniform,false,5,0.5,50);

%ssize(5,LOGN,false,5,0.5,50);
%ssize(10,LOGN,false,5,0.5,50);
%ssize(25,LOGN,false,5,0.5,50);
%ssize(50,LOGN,false,5,0.5,50);

%comp(Normal,50,5);
%comp(Uniform,50,5);
%comp(LOGN,50,5);

*-----
-----;

%ssize(5,Normal,true,5,0.9,90);
%ssize(10,Normal,true,5,0.9,90);
%ssize(25,Normal,true,5,0.9,90);
%ssize(50,Normal,true,5,0.9,90);

%ssize(5,Uniform,true,5,0.9,90);
%ssize(10,Uniform,true,5,0.9,90);
%ssize(25,Uniform,true,5,0.9,90);
%ssize(50,Uniform,true,5,0.9,90);

%ssize(5,LOGN,true,5,0.9,90);

```

```

%ssize(10,LOGN,true,5,0.9,90);
%ssize(25,LOGN,true,5,0.9,90);
%ssize(50,LOGN,true,5,0.9,90);

%ssize(5,Normal,false,5,0.9,90);
%ssize(10,Normal,false,5,0.9,90);
%ssize(25,Normal,false,5,0.9,90);
%ssize(50,Normal,false,5,0.9,90);

%ssize(5,Uniform,false,5,0.9,90);
%ssize(10,Uniform,false,5,0.9,90);
%ssize(25,Uniform,false,5,0.9,90);
%ssize(50,Uniform,false,5,0.9,90);

%ssize(5,LOGN,false,5,0.9,90);
%ssize(10,LOGN,false,5,0.9,90);
%ssize(25,LOGN,false,5,0.9,90);
%ssize(50,LOGN,false,5,0.9,90);

%comp(Normal,90,5);
%comp(Uniform,90,5);
%comp(LOGN,90,5);

data nhla.type1_p5;
set nhla.type1_Normal0 nhla.type1_Uniform0 nhla.type1_LOGN0
nhla.type1_Normal50 nhla.type1_Uniform50 nhla.type1_LOGN50
nhla.type1_Normal90 nhla.type1_Uniform90 nhla.type1_LOGN90;
run;

data nhla.power_p5;
set nhla.power_Normal0 nhla.power_Uniform0 nhla.power_LOGN0
nhla.power_Normal50 nhla.power_Uniform50 nhla.power_LOGN50
nhla.power_Normal90 nhla.power_Uniform90 nhla.power_LOGN90;
run;

```

Confidence Intervals

```

* Masters dissertation program;
libname nhla "C:\Users\Sibu\Documents\Nhlanhla";
proc iml;

*starting the standard regression analysis;

    start regress1 (x,y)
Global (t1,pvalue_b1,F,crit,LCLori,UCLori,dist_Ori,betaOri,beta1);
/* begin module */
    alpha=0.05;

```

```

xpxi=inv(t(x)*x);          /* inverse of X'X          */
beta=xpxi*(t(x)*y);       /* parameter estimate     */
yhat=x*beta;              /* predicted values       */
  resid=y-yhat;           /* residuals              */
  sse=ssq(resid);         /* SSE                    */
  n_0=nrow(x);            /* sample size            */
  dfe=nrow(x)-ncol(x);   /* error DF              */
  mse=sse/dfe;           /* MSE                    */
  cssy=ssq(y-sum(y)/n_0); /* corrected total SS    */
  rsquare=(cssy-sse)/cssy; /* RSQUARE               */
  ssr=t(beta)*xpxi*beta;
  msr=ssr/(ncol(x)-1);
  F=msr/mse;
  stdb=sqrt(vecdiag(xpxi)*mse); /* std of estimates     */
  t=beta/stdb;          /* parameter t tests    */

  prob=1-probf(t#t,1,dfe); /* p-values             */
  pvalue_b1=prob[2];

  crit=tinv(1-(alpha/2),dfe); /* t critical value for
95% confidence interval*/
  LCL=beta - crit#stdb;
  UCL=beta + crit#stdb;
  beta1=beta[2];
  t1=t[2];
  LCLOri=lcl[2];
  UCLOri=UCL[2];
  dist_ori=ucl[2]-lcl[2];

  finish regress1;          /* end module           */

*-----
-----;
  *starting a module to determine the quantile/percentiles;
/** Qntl: compute quantiles (Defn. 5 from the UNIVARIATE doc) **/
/** Arguments:
  q  upon return, q contains the specified sample quantiles of
      the data.
  x  is a matrix. The module computes quantiles for each column.
  p  specifies the quantiles. For example, 0.5 specifies the
      median, whereas {0.25 0.75} specifies the first and
      third quartiles.
This module does not handle missing values in the data. **/
start Qntl(q, x, p);       /** definition 5 from UNIVARIATE doc **/
  n = nrow(x);             /** assume nonmissing data **/
  q = j(ncol(p), ncol(x)); /** allocate space for return values **/
  do j = 1 to ncol(x);     /** for each column of x... **/
    h = x[,j];
    call sort(h,1);        /** sort the values **/
  do i = 1 to ncol(p);    /** for each quantile **/

```

```

    k = n*p[i];          /** find position in ordered data **/
    k1 = int(k);        /** find indices into ordered data **/
    k2 = k1 + 1;
    g = k - k1;
    if g>0 then
        q[i,j] = h[k2]; /** return a data value **/
    else                /** average adjacent data **/
        q[i,j] = (h[k1]+h[k2])/2;
    end;
end;
finish;
store module = qntl;
*-----
-----;

start delrow(x,i);     /*starting the function to delete rows*/
return(x[setdif(1:nrow(x),i),]);
finish delrow;

start regress_j(x,y)
Global(t1,pvalue_j,F,crit,q,LCLJ,UCLJ,BetalJ,PercLJ,PercUJ,betaJ1,di
st_J,distP_J)        ;    /* starting the regression function*/
n=nrow(x);
p_j=0;
pf=0;
t1_j=J(nrow(x),1,1);
betal_j=J(nrow(x),1,1);
stdbl_j=J(nrow(x),1,1);

do i=1 to nrow(x);

xj=delrow(x,i);
yj=delrow(y,i);

    xjpxji=inv(t(xj)*xj);          /* inverse of X'X          */
    beta_j=xjpxji*(t(xj)*yj);     /* parameter estimate     */
    yjhat=xj*beta_j;             /* predicted values       */
    resid_j=yj-yjhat;            /* residuals               */
    sse_j=ssq(resid_j);          /* SSE                     */
    n_j=nrow(xj);                /* sample size             */
*/
    dfe_j=nrow(xj)-ncol(xj);      /* error DF                */
    mse_j=sse_j/dfe_j;           /* MSE                     */
*/
    cssb_j=ssq(yj-sum(yj)/n_j);   /* corrected total SS     */
*/
    rsquare_j=(cssb_j-sse_j)/cssb_j; /* RSQUARE                */
*/
    ssr=t(beta_j)*xjpxji*beta_j;
    msr=ssr/(ncol(xj)-1);
    F_j=msr/mse_j;

```

```

        stdb_j=sqrt(vecdiag(xjpxji)*mse_j); /* std of estimates
*/
        t_j=beta_j/stdb_j;                /* parameter t tests
*/

        /* creating the vectors of betal separately*/
        t1_j[i]=t_j[2];
        betal_j[i]=beta_j[2] ;
        stdbl_j[i]=stdb_j[2] ;
end;

*Estimate of confidence interval;

betalJ=beta1_j[:]; /*Jackknife estimate of beta1*/
varbj=(n-1)/n *sum(vecdiag((beta1_j-betalJ) * t(beta1_j-betalJ)));
/* Jackknife variance estimate*/
stdJ=sqrt(varbj); /* standard error of jackknife estimate*/
LCLJ=beta1J - crit#stdJ; /* Lower Normal 95% confidence interval of
beta 1*/
UCLJ=beta1J + crit#stdJ; /* Upper Normal 95% confidence interval of
beta 1*/

p={0.025 0.975}; /* alpha/2 and 1-alpha/2 percentiles*/
run qntl(q, beta1_j, p); /* call module to compute sample
quantiles **/
call sort (beta1_j,1);

PercLJ=q[1]; /* Lower Percentile confidence limit*/
PercUJ=q[2]; /* Upper Percentile confidence limit*/

dist_J=uclj-lclj;
distP_J=percUJ-percLJ;

finish regress_j ;

/*-----*/
-----*/

start delrow2(x,g);
return(x[setdif(1:nrow(x),g),]);
finish;

start regress_j2 (x,y,n)
Global (t1,pvalue_j2,F,crit,LCLJ2,UCLJ2,BetalJ2,PercLJ2,PercUJ2,betaJ
2,dist_J2,distP_J2);
n=nrow(x);
j=2;
p_j2=0;
pf=0;
h=comb(n,j) ;

```

```

t1_j2=j(h,1,1);
*Fj2=j(h,1,1);
beta1_j2=j(h,1,1);
stdb1_j2=j(h,1,1);

u=1:nrow(x) ; /*print u; */
l=0;

do i=1 to nrow(x)-(j-1);
  do k=1 to nrow(x)-i;
    r=k+i;
    g=u[i||r];
    l=l+1;

c=delrow2(x,g);
d=delrow2(y,g); *c=x in this case and d equals y;

cpci=inv(t(c)*c);          /* inverse of X'X          */
beta_j2=cpci*(t(c)*d);    /* parameter estimate    */
bhat=c*beta_j2;          /* predicted values      */
resid_j2=d-bhat;         /* residuals              */
sse_j2=ssq(resid_j2);    /* SSE                    */
*/
n_j2=nrow(c);            /* sample size           */
*/
dfe_j2=nrow(c)-ncol(c);  /* error DF              */
mse_j2=sse_j2/dfe_j2;    /* MSE                   */
*/
cssb_j2=ssq(d-sum(d)/n_j2); /* corrected total SS   */
*/
rsquare_j2=(cssb_j2-sse_j2)/cssb_j2; /* RSQUARE              */
*/
ssr=t(beta_j2)*cpci*beta_j2;
msr=ssr/(ncol(c)-1);
stdb_j2=sqrt(vecdiag(cpci)*mse_j2); /* std of estimates     */
*/
t_j2=beta_j2/stdb_j2;    /* parameter t tests    */

t1_j2[1]=t_j2[2];
beta1_j2[1]=beta_j2[2] ;
stdb1_j2[1]=stdb_j2[2] ;

if abs(t1_j2[1]) >= abs(t1) then p_j2=p_j2+1;

end;

end;
pvalue_j2=p_j2/h;

*Estimate of confidence interval;

```

```

beta1J2=beta1_j2[:]; /*Jackknife estimate of beta1*/
varbj2=((n-2)/(2*1))*sum(vecdiag((beta1_j2-beta1J2)* t(beta1_j2-
beta1J2))); /* Jackknife variance estimate*/
stdJ2=sqrt(varbj2); /* standard error of jackknife estimate*/
LCLJ2=beta1J2 - crit#stdJ2; /* Lower Normal 95% confidence interval
of beta 1*/
UCLJ2=beta1J2 + crit#stdJ2; /* Upper Normal 95% confidence interval
of beta 1*/

p={0.025 0.975}; /* alpha/2 and 1-alpha/2 percentiles*/
call qntl(q, beta1_j2, p); /* call module to compute sample
quantiles **/
call sort (beta1_j2,1);

PercLJ2=q[1]; /* Lower Percentile confidence limit*/
PercUJ2=q[2]; /* Upper Percentile confidence limit*/

dist_J2=uclj2-lclj2;
distP_J2=percUJ2-percLJ2;

finish regress_j2 ;

/*-----*/
*bootstrapping residuals;

start param_br(x,y,n)
Global (crit,LCLBr,UCLBr,Beta1Br,PercLBr,PercUBr,betaBr,dist_Br,distP
_Br,beta1); /* begin module */
n=nrow(x);
rep=999;
beta1_bpr=J(rep,1,1);
stdb1_bpr=J(rep,1,1);
f=j(n,1,1);

k=j(n,1);

xtxi=inv(t(x)*x); /* inverse of X'X */
betabpr=xtxi*(t(x)*y); /* parameter estimate */

yhat=x*betabpr; /* predicted values */
resid=y-yhat;

do j=1 to rep;

call randgen(k,"uniform");
f=ceil(k#n);

```

```

v=resid[f];

ybpr=yhat+v;

    xbrtxbri=inv(t(x)*x);          /* inverse of X'X          */
    beta_bpr=xbrtxbri*(t(x)*ybpr); /* parameter estimate
*/
    ybprhat=x*beta_bpr;           /* predicted values      */
    resid_bpr=ybpr-ybprhat;      /* residuals
*/
    sse_bpr=ssq(resid_bpr);      /* SSE
*/
    n_br=nrow(x);                /* sample size
*/
    dfe_bpr=nrow(x)-ncol(x);     /* error DF
*/
    mse_bpr=sse_bpr/dfe_bpr;     /* MSE
*/

    betal_bpr[j]=beta_bpr[2] ;

end;

    *Bootstrap betal estimate and its confidence interval;
    betalBr=betal_bpr[:];
    varBr= (1/rep)*sum(vecdiag((betal_bpr-betalBr)* t(betal_bpr-
    betalBr)));
    stdBr=sqrt(varBr);
    LCLBr=betalBr - crit#stdBr; /* Lower Normal 95% confidence interval
of beta 1*/
    UCLBr=betalBr + crit#stdBr; /* Upper Normal 95% confidence interval
of beta 1*/

p={0.025 0.975}; /* alpha/2 and 1-alpha/2 percentiles*/
call qntl(q, betal_bpr, p); /* call module to compute sample
quantiles */
call sort (betal_bpr,1);

PercLBr=q[1]; /* Lower Percentile confidence limit*/
PercUBr=q[2]; /* Upper Percentile confidence limit*/
dist_Br=uclbr-lclBr;
distP_Br=percUBr-percLBr;

finish param_br; /* end module */

*-----
-----

*bootstrapping observations;

start
param_b(x,y,n,p) Global (t1,F,crit,LCLb,UCLb,Betalb,PercLB,PercUB,beta
B,dist_B,distP_B); /* begin module */
n=nrow(x);

```



```

h=j(n,1,.);

if (n=5 & p=1) | (n=10 & p=2) then do;
dist_B=.;
distP_B=.;
LCLb=.;
UCLb=.;
Beta1b=.;
PercLB=.;
PercUB=.;
betaB=.;
end;

else do;
rep=999;

beta1_bp=J(rep,1,1);

v=j(n,1);

do j=1 to rep;
call randgen(v,"uniform");
h=ceil(v#n);

xb=x[h,];
yb=y[h,];

    xbtxbi=inv(t(xb)*xb);          /* inverse of X'X          */
    beta_bp=xbtxbi*(t(xb)*yb);     /* parameter estimate     */
    ybphat=xb*beta_bp;            /* predicted values      */

beta1_bp[j]=beta_bp[2];

    end;

*Bootstrap beta1 estimate and it confidence interval;
beta1B=beta1_bp[:];
varB= (1/rep)*sum(vecdiag((beta1_bp-beta1B)* t(beta1_bp-beta1B)));
cc= (beta1_bp-beta1B)* t(beta1_bp-beta1B);
stdB=sqrt(varB);
LCLB=beta1B - crit#stdB; /* Lower Normal 95% confidence interval of
beta 1*/
UCLB=beta1B + crit#stdB; /* Upper Normal 95% confidence interval of
beta 1*/

p={0.025 0.975}; /* alpha/2 and 1-alpha/2 percentiles*/
call qntl(q, beta1_bp, p); /* call module to compute sample
quantiles **/
call sort (beta1_bp,1);

PercLB=q[1]; /* Lower Percentile confidence limit*/
PercUB=q[2]; /* Upper Percentile confidence limit*/

```

```

        method=j(3,1,4);
beta11=j(3,1,betalb);
betaB=method||beta11||(./LCLB//PercLB)||(.//UCLB//PercUB);
*print betaB;

dist_B=uclB-lclB;
distP_B=percUB-percLB;

end;

finish param_b; /* end module */

/*End of bootstrap and jackknife resampling algorithms for
regression parameters estimation*/

*-----;

store module=regress1;
store module=regress_j;
store module=delrow;
store module=delrow2;
store module=(regress_j2 param_br param_b);

quit;

%macro ssize(size,er,h0,p,rho,pr);

proc iml;
load module=regress1;
no=1000;

/*confidence interval declaration*/
LCLOri&size=j(no,1,.);
UCLOri&size=j(no,1,.);
LCLJ&size=j(no,1,.);
UCLJ&size=j(no,1,.);
LCLJ2&size=j(no,1,.);
UCLj2&size=j(no,1,.);
LCLBr&size=j(no,1,.);
UCLBr&size=j(no,1,.);
LCLB&size=j(no,1,.);
UCLB&size=j(no,1,.);

```

```

dist_Ori&size=j(no,1,.);
dist_J&size=j(no,1,.);
dist_J2&size=j(no,1,.);
dist_B&size=j(no,1,.);
dist_Br&size=j(no,1,.);

PercLJ&size=j(no,1,.);
PercUJ&size=j(no,1,.);
PercLJ2&size=j(no,1,.);
PercUJ2&size=j(no,1,.);
PercLBr&size=j(no,1,.);
PercUBr&size=j(no,1,.);
PercLB&size=j(no,1,.);
PercUB&size=j(no,1,.);

distP_J&size=j(no,1,.);
distP_J2&size=j(no,1,.);
distP_B&size=j(no,1,.);
distP_Br&size=j(no,1,.);

meth="&er";
alpha=0.05;
h0="&h0"; /*Null hypothesis*/

call randseed(89721);

do k=1 to no;

n=&p*&size;      /* sample size */
b=j(&p+1,1,1);

if h0="true" then  b[2]=0;
else b[2]=2;

R=j(&p,&p,1);    *correlation matrix;

if &p > 1 then do;
  do i=1 to &p;
    do j=1 to &p;
      if i ^= j then  R[i,j]=&rho;
    end;
  end;
end;

e&size.&er=j(n,1,.);
w=J(n,&p,1);

p=&p;

call randgen(e&size.&er,"&er");

```

```

call randgen (w, "Uniform");
x=j(n,1,1)||w*root(R);
y = x*b+e&size.&er;

run regress1(x,y);
run regress_j(x,y);
if &size ^ = 50 & &p ^ = 5 then do;
run regress_j2(x,y,n);end;
run param_br(x,y,n);
run param_b(x,y,n,p);

/*confidence interval declaration*/
LCLOri&size[k]=LCLOri;
UCLOri&size[k]=UCLOri;
LCLJ&size[k]=LCLJ;
UCLJ&size[k]=UCLJ;
LCLJ2&size[k]=LCLJ2;
UCLj2&size[k]=UCLj2;
LCLBr&size[k]=LCLBr;
UCLBr&size[k]=UCLBr;
LCLB&size[k]=LCLB;
UCLB&size[k]=UCLB;

dist_Ori&size[k]=dist_Ori;
dist_J&size[k]=dist_J;
dist_J2&size[k]=dist_J2;
dist_B&size[k]=dist_B;
dist_Br&size[k]=dist_Br;

PercLJ&size[k]=PercLJ;
PercUJ&size[k]=PercUJ;
PercLJ2&size[k]=PercLJ2;
PercUJ2&size[k]=PercUJ2;
PercLBr&size[k]=PercLBr;
PercUBr&size[k]=PercUBr;
PercLB&size[k]=PercLB;
PercUB&size[k]=PercUB;

distP_J&size[k]=distP_J;
distP_J2&size[k]=distP_J2;
distP_B&size[k]=distP_B;
distP_Br&size[k]=distP_Br;

end;

Ave_LCLOri&size = LCLOri&size[:];
Ave_UCLOri&size= UCLOri&size[:];
Ave_LCLJ&size=LCLJ&size[:];
Ave_UCLJ&size=UCLJ&size[:];
Ave_LCLJ2&size=LCLJ2&size[:];

```

```

Ave_UCLj2&size=UCLj2&size[:];
Ave_LCLBr&size=LCLBr&size[:];
Ave_UCLBr&size=UCLBr&size[:];
Ave_LCLB&size=LCLB&size[:];
Ave_UCLB&size=UCLB&size[:];

Ave_dist_Ori&size=dist_Ori&size[:];
Ave_dist_J&size=dist_J&size[:];
Ave_dist_J2&size=dist_J2&size[:];
Ave_dist_B&size=dist_B&size[:];
Ave_dist_Br&size=dist_Br&size[:];

Ave_PercLJ&size=PercLJ&size[:];
Ave_PercUJ&size=PercUJ&size[:];
Ave_PercLJ2&size=PercLJ2&size[:];
Ave_PercUJ2&size=PercUJ2&size[:];
Ave_PercLBr&size=PercLBr&size[:];
Ave_PercUBr&size=PercUBr&size[:];
Ave_PercLB&size=PercLB&size[:];
Ave_PercUB&size=PercUB&size[:];

Ave_distP_J&size=distP_J&size[:];
Ave_distP_J2&size=distP_J2&size[:];
Ave_distP_B&size=distP_B&size[:];
Ave_distP_Br&size=distP_Br&size[:];

if h0="true" then do;

CIOri&size=100*(sum(LCLOri&size <= 0 <= UCLOri&size)/no);
CIJ&size= 100*(sum(LCLJ&size <= 0 <= UCLJ&size)/no);
CIJ2&size= 100*(sum(LCLJ2&size <= 0 <= UCLJ2&size)/no);
CIBr&size= 100*(sum(LCLBr&size <= 0 <= UCLBr&size)/no);
CIB&size= 100*(sum(LCLB&size <= 0 <= UCLB&size)/no);

PIJ&size= 100*(sum(PercLJ&size <= 0 <= PercUJ&size)/no);
PIJ2&size= 100*(sum(PercLJ2&size <= 0 <= PercUJ2&size)/no);
PIBr&size= 100*(sum(PercLBr&size <= 0 <= PercUBr&size)/no);
PIB&size= 100*(sum(PercLB&size <= 0 <= PercUB&size)/no);

end;

else do;

CIOri&size=100*(sum(LCLOri&size <= 2 <= UCLOri&size)/no);
CIJ&size= 100*(sum(LCLJ&size <= 2 <= UCLJ&size)/no);
CIJ2&size= 100*(sum(LCLJ2&size <= 2 <= UCLJ2&size)/no);
CIBr&size= 100*(sum(LCLBr&size <= 2 <= UCLBr&size)/no);
CIB&size= 100*(sum(LCLB&size <= 2 <= UCLB&size)/no);

```

```

PIJ&size= 100*(sum(PercLJ&size <=      2 <= PercUJ&size)/no);
PIJ2&size= 100*(sum(PercLJ2&size <=    2 <= PercUJ2&size)/no);
PIBr&size= 100*(sum(PercLBr&size <=    2 <= PercUBr&size)/no);
PIB&size= 100*(sum(PercLB&size <=     2 <= PercUB&size)/no);
  end;

if &size = 50 & &p = 5 then do;
CIJ2&size=.;Ave_dist_J2&size=.;
PIJ2&size=.;Ave_distP_J2&size=.;
end;

if meth="Normal" then error=0;
else if meth="Uniform" then error=1;
else error=2;

rn={"Beta1","no of
par","Size","Ssize","rho","method","CIOri","AveDistOri","CIJ","AveDi
stJ","CIJ2","AveDistJ2","CIBr","AveDistBr","CIB","AveDistB","PIJ","A
veDistPJ","PIJ2","AveDistPJ2","PIBr","AveDistPBr","PIB","AveDistPB"}
;

if h0="false" then do;

beta1=2;
ciF&size=beta1||&p||&size||n||&rho||error||CIOri&size
||Ave_dist_Ori&size||CIJ&size||Ave_dist_J&size||
CIJ2&size||Ave_dist_J2&size|| CIBr&size||Ave_dist_Br&size|| CIB&size
||Ave_dist_B&size||PIJ&size
||Ave_distP_J&size||PIJ2&size||Ave_distP_J2&size|| PIBr&size
||Ave_distP_Br&size||PIB&size||Ave_distP_B&size ;

create confIF&size.&er.&pr from ciF&size[colname=rn];
append from ciF&size;

end;

else do;

beta1=0;
ciT&size=beta1||&p||&size||n||&rho||error||CIOri&size
||Ave_dist_Ori&size||CIJ&size||Ave_dist_J&size||
CIJ2&size||Ave_dist_J2&size|| CIBr&size||Ave_dist_Br&size|| CIB&size
||Ave_dist_B&size||PIJ&size
||Ave_distP_J&size||PIJ2&size||Ave_distP_J2&size|| PIBr&size
||Ave_distP_Br&size||PIB&size||Ave_distP_B&size ;

create confIT&size.&er.&pr from ciT&size[colname=rn];
append from ciT&size;
end;

```

```

quit;
%mend ssize;
%ssize(5,Normal,true,1,0,0);
%ssize(10,Normal,true,1,0,0);
%ssize(25,Normal,true,1,0,0);
%ssize(50,Normal,true,1,0,0);

%ssize(5,Uniform,true,1,0,0);
%ssize(10,Uniform,true,1,0,0);
%ssize(25,Uniform,true,1,0,0);
%ssize(50,Uniform,true,1,0,0);

%ssize(5,LOGN,true,1,0,0);
%ssize(10,LOGN,true,1,0,0);
%ssize(25,LOGN,true,1,0,0);
%ssize(50,LOGN,true,1,0,0);

%ssize(5,Normal,false,1,0,0);
%ssize(10,Normal,false,1,0,0);
%ssize(25,Normal,false,1,0,0);
%ssize(50,Normal,false,1,0,0);

%ssize(5,Uniform,false,1,0,0);
%ssize(10,Uniform,false,1,0,0);
%ssize(25,Uniform,false,1,0,0);
%ssize(50,Uniform,false,1,0,0);

%ssize(5,LOGN,false,1,0,0);
%ssize(10,LOGN,false,1,0,0);
%ssize(25,LOGN,false,1,0,0);
%ssize(50,LOGN,false,1,0,0);

ods html
body="C:\Users\nhlanhla\Documents\important\dessertation\results"

gpath="C:\Users\nhlanhla\Documents\important\dessertation\results\"
style=journal;
%macro comp(er,rho,p);

data CI_&er.&rho;
length n $3;
set confit5&er.&rho confit10&er.&rho confit25&er.&rho
confit50&er.&rho confif5&er.&rho confif10&er.&rho confif25&er.&rho
confif50&er.&rho;
if size=5 then n="5p";
else if size=10 then n="10p";
else if size=25 then n="25p";
else if size=50 then n="50p";
drop size;

```

```
run;
```

```
/*proc datasets;*/
/*confit5&er.&rho confit10&er.&rho confit25&er.&rho
confit50&er.&rho confif5&er.&rho confif10&er.&rho confif25&er.&rho
confif50&er.&rho;*/
```

```
%mend comp;
```

```
%comp (Normal, 0, 1);
```

```
%comp (Uniform, 0, 1);
```

```
%comp (LOGN, 0, 1);
```

```
data nhla.conf_int_p1;
```

```
set CI_Normal0 CI_Uniform0 CI_LOGN0;
```

```
run;
```

```
*p=2;
```

```
%ssize(5, Normal, true, 2, 0, 0);
```

```
%ssize(10, Normal, true, 2, 0, 0);
```

```
%ssize(25, Normal, true, 2, 0, 0);
```

```
%ssize(50, Normal, true, 2, 0, 0);
```

```
%ssize(5, Uniform, true, 2, 0, 0);
```

```
%ssize(10, Uniform, true, 2, 0, 0);
```

```
%ssize(25, Uniform, true, 2, 0, 0);
```

```
%ssize(50, Uniform, true, 2, 0, 0);
```

```
%ssize(5, LOGN, true, 2, 0, 0);
```

```
%ssize(10, LOGN, true, 2, 0, 0);
```

```
%ssize(25, LOGN, true, 2, 0, 0);
```

```
%ssize(50, LOGN, true, 2, 0, 0);
```

```
%ssize(5, Normal, false, 2, 0, 0);
```

```
%ssize(10, Normal, false, 2, 0, 0);
```

```
%ssize(25, Normal, false, 2, 0, 0);
```

```
%ssize(50, Normal, false, 2, 0, 0);
```

```
%ssize(5, Uniform, false, 2, 0, 0);
```

```
%ssize(10, Uniform, false, 2, 0, 0);
```

```
%ssize(25, Uniform, false, 2, 0, 0);
```

```
%ssize(50, Uniform, false, 2, 0, 0);
```

```
%ssize(5, LOGN, false, 2, 0, 0);
```

```
%ssize(10, LOGN, false, 2, 0, 0);
```

```
%ssize(25, LOGN, false, 2, 0, 0);
```

```
%ssize(50, LOGN, false, 2, 0, 0);
```

```
%comp (Normal, 0, 2);
```

```
%comp (Uniform, 0, 2);
```

```
%comp (LOGN, 0, 2);
```



```

*-----;
-----;

%ssize(5,Normal,true,2,0.5,50);
%ssize(10,Normal,true,2,0.5,50);
%ssize(25,Normal,true,2,0.5,50);
%ssize(50,Normal,true,2,0.5,50);

%ssize(5,Uniform,true,2,0.5,50);
%ssize(10,Uniform,true,2,0.5,50);
%ssize(25,Uniform,true,2,0.5,50);
%ssize(50,Uniform,true,2,0.5,50);

%ssize(5,LOGN,true,2,0.5,50);
%ssize(10,LOGN,true,2,0.5,50);
%ssize(25,LOGN,true,2,0.5,50);
%ssize(50,LOGN,true,2,0.5,50);

%ssize(5,Normal,false,2,0.5,50);
%ssize(10,Normal,false,2,0.5,50);
%ssize(25,Normal,false,2,0.5,50);
%ssize(50,Normal,false,2,0.5,50);

%ssize(5,Uniform,false,2,0.5,50);
%ssize(10,Uniform,false,2,0.5,50);
%ssize(25,Uniform,false,2,0.5,50);
%ssize(50,Uniform,false,2,0.5,50);

%ssize(5,LOGN,false,2,0.5,50);
%ssize(10,LOGN,false,2,0.5,50);
%ssize(25,LOGN,false,2,0.5,50);
%ssize(50,LOGN,false,2,0.5,50);

*-----;

%comp(Normal,50,2);
%comp(Uniform,50,2);
%comp(LOGN,50,2);
*-----;
-----;

%ssize(5,Normal,true,2,0.9,90);
%ssize(10,Normal,true,2,0.9,90);
%ssize(25,Normal,true,2,0.9,90);
%ssize(50,Normal,true,2,0.9,90);

%ssize(5,Uniform,true,2,0.9,90);
%ssize(10,Uniform,true,2,0.9,90);
%ssize(25,Uniform,true,2,0.9,90);
%ssize(50,Uniform,true,2,0.9,90);

```

```

%ssize(5,LOGN,true,2,0.9,90);
%ssize(10,LOGN,true,2,0.9,90);
%ssize(25,LOGN,true,2,0.9,90);
%ssize(50,LOGN,true,2,0.9,90);

%ssize(5,Normal,false,2,0.9,90);
%ssize(10,Normal,false,2,0.9,90);
%ssize(25,Normal,false,2,0.9,90);
%ssize(50,Normal,false,2,0.9,90);

%ssize(5,Uniform,false,2,0.9,90);
%ssize(10,Uniform,false,2,0.9,90);
%ssize(25,Uniform,false,2,0.9,90);
%ssize(50,Uniform,false,2,0.9,90);

%ssize(5,LOGN,false,2,0.9,90);
%ssize(10,LOGN,false,2,0.9,90);
%ssize(25,LOGN,false,2,0.9,90);
%ssize(50,LOGN,false,2,0.9,90);

%comp(Normal,90,2);
%comp(Uniform,90,2);
%comp(LOGN,90,2);
data nhla.conf_int_p2;
set CI_Normal0 CI_Uniform0 CI_LOGN0 CI_Normal50 CI_Uniform50
CI_LOGN50 CI_Normal90 CI_Uniform90 CI_LOGN90;
run;

*p=5;

%ssize(5,Normal,true,5,0,0);
%ssize(10,Normal,true,5,0,0);
%ssize(25,Normal,true,5,0,0);
%ssize(50,Normal,true,5,0,0);

%ssize(5,Uniform,true,5,0,0);
%ssize(10,Uniform,true,5,0,0);
%ssize(25,Uniform,true,5,0,0);
%ssize(50,Uniform,true,5,0,0);

%ssize(5,LOGN,true,5,0,0);
%ssize(10,LOGN,true,5,0,0);
%ssize(25,LOGN,true,5,0,0);
%ssize(50,LOGN,true,5,0,0);

%ssize(5,Normal,false,5,0,0);
%ssize(10,Normal,false,5,0,0);
%ssize(25,Normal,false,5,0,0);
%ssize(50,Normal,false,5,0,0);

```

```
%ssize(5,Uniform,false,5,0,0);
%ssize(10,Uniform,false,5,0,0);
%ssize(25,Uniform,false,5,0,0);
%ssize(50,Uniform,false,5,0,0);
```

```
%ssize(5,LOGN,false,5,0,0);
%ssize(10,LOGN,false,5,0,0);
%ssize(25,LOGN,false,5,0,0);
%ssize(50,LOGN,false,5,0,0);
```

```
%ssize(50,Uniform,true,5,0,0);
%ssize(50,LOGN,true,5,0,0);
%ssize(50,Normal,false,5,0,0);
%ssize(50,Uniform,false,5,0,0);
%ssize(50,LOGN,false,5,0,0);
```

```
%comp(Normal,0,5);
%comp(Uniform,0,5);
%comp(LOGN,0,5);
```

```
*-----;
-----;
```

```
%ssize(5,Normal,true,5,0.5,50);
%ssize(10,Normal,true,5,0.5,50);
%ssize(25,Normal,true,5,0.5,50);
%ssize(50,Normal,true,5,0.5,50);
```

```
%ssize(5,Uniform,true,5,0.5,50);
%ssize(10,Uniform,true,5,0.5,50);
%ssize(25,Uniform,true,5,0.5,50);
%ssize(50,Uniform,true,5,0.5,50);
```

```
%ssize(5,LOGN,true,5,0.5,50);
%ssize(10,LOGN,true,5,0.5,50);
%ssize(25,LOGN,true,5,0.5,50);
%ssize(50,LOGN,true,5,0.5,50);
```

```
%ssize(5,Normal,false,5,0.5,50);
%ssize(10,Normal,false,5,0.5,50);
%ssize(25,Normal,false,5,0.5,50);
%ssize(50,Normal,false,5,0.5,50);
```

```
%ssize(5,Uniform,false,5,0.5,50);
%ssize(10,Uniform,false,5,0.5,50);
%ssize(25,Uniform,false,5,0.5,50);
%ssize(50,Uniform,false,5,0.5,50);
```

```
%ssize(5,LOGN,false,5,0.5,50);
%ssize(10,LOGN,false,5,0.5,50);
```

```

%ssize(25,LOGN,false,5,0.5,50);
%ssize(50,LOGN,false,5,0.5,50);

%ssize(50,Normal,true,5,0.5,50);
%ssize(50,Uniform,true,5,0.5,50);
%ssize(50,LOGN,true,5,0.5,50);
%ssize(50,Normal,false,5,0.5,50);
%ssize(50,Uniform,false,5,0.5,50);
%ssize(50,LOGN,false,5,0.5,50);

%comp(Normal,50,5);
%comp(Uniform,50,5);
%comp(LOGN,50,5);

*-----
-----;

%ssize(5,Normal,true,5,0.9,90);
%ssize(10,Normal,true,5,0.9,90);
%ssize(25,Normal,true,5,0.9,90);
%ssize(50,Normal,true,5,0.9,90);

%ssize(5,Uniform,true,5,0.9,90);
%ssize(10,Uniform,true,5,0.9,90);
%ssize(25,Uniform,true,5,0.9,90);
%ssize(50,Uniform,true,5,0.9,90);

%ssize(5,LOGN,true,5,0.9,90);
%ssize(10,LOGN,true,5,0.9,90);
%ssize(25,LOGN,true,5,0.9,90);
%ssize(50,LOGN,true,5,0.9,90);

%ssize(5,Normal,false,5,0.9,90);
%ssize(10,Normal,false,5,0.9,90);
%ssize(25,Normal,false,5,0.9,90);
%ssize(50,Normal,false,5,0.9,90);

%ssize(5,Uniform,false,5,0.9,90);
%ssize(10,Uniform,false,5,0.9,90);
%ssize(25,Uniform,false,5,0.9,90);
%ssize(50,Uniform,false,5,0.9,90);

%ssize(5,LOGN,false,5,0.9,90);
%ssize(10,LOGN,false,5,0.9,90);
%ssize(25,LOGN,false,5,0.9,90);
%ssize(50,LOGN,false,5,0.9,90);

%ssize(50,Normal,true,5,0.9,90);
%ssize(50,Uniform,true,5,0.9,90);
%ssize(50,LOGN,true,5,0.9,90);
%ssize(50,Normal,false,5,0.9,90);
%ssize(50,Uniform,false,5,0.9,90);
%ssize(50,LOGN,false,5,0.9,90);

```

```

%comp(Normal, 90, 5);
%comp (Uniform, 90, 5);
%comp (LOGN, 90, 5);

data nhla.conf_int_p5;
set CI_Normal0 CI_Uniform0 CI_LOGN0 CI_Normal50 CI_Uniform50
CI_LOGN50 CI_Normal90 CI_Uniform90 CI_LOGN90;
run;
ods html close;

```

Graph and Confidence Interval tables Preparation

```

data nhla.confidence_int;
set nhla.conf_int_p1 nhla.conf_int_p2 nhla.conf_int_p5;
run;

data nhla.conf;
retain "no of par"n rho method beta1 n ssize ciori cij cij2 cibr cib
AveDistOri AveDistJ AveDistj2 AveDistBr AveDistB;
set nhla.confidence_int;

keep "no of par"n rho method beta1 n ssize ciori cij cij2 cibr cib
AveDistOri AveDistJ AveDistj2 AveDistBr AveDistB;
run;
proc sort data=nhla.conf; by method rho ;run;
data nhla.perc;
retain "no of par"n rho method beta1 n ssize piori pij pij2 pibr pib
AveDistJ AveDistj2 AveDistBr AveDistB;
set nhla.confidence_int;

keep "no of par"n rho method beta1 n ssize pij pij2 pibr pib
AveDistpJ AveDistpj2 AveDistpBr AveDistpB;
run;
proc sort data=nhla.perc; by method rho ;run;

data nhla.power;
set nhla.power_p1 nhla.power_p2 nhla.power_p5;
run;
data nhla.type1;
set nhla.type1_p1 nhla.type1_p2 nhla.type1_p5;
run;

data nhla.powerN nhla.powerU nhla.powerL;
set nhla.power;
retain error no of par rho n ssize powerNormal -- powerFL;
if error=0 then output nhla.powerN; else
if error=1 then output nhla.powerU; else
output nhla.powerL;

```

```

run;

data nhla.type1N nhla.type1U nhla.type1L;
set nhla.type1;
retain error no_of_par rho n ssize type1Normal -- type1FL;
drop size ;
if error=0 then output nhla.type1N;else
if error=1 then output nhla.type1U; else
output nhla.type1L;
run;

```

Probability of type 1 Error Graphs

```

goptions reset=all border;

symbol1 interpol=join
value=star
      cv=red
      ci=red
      width=2;
symbol2 interpol=join
value=dot
      cv=blue
      ci=blue
      line=3
      width=2;
symbol3 interpol=join
value=circle
      cv=green
      ci=green
      line=8
      width=2;
symbol4 interpol=join
value=triangle
      cv=black
      ci=black
      line=33
      width=2;
symbol5 interpol=join
value=square
      cv=brown
      ci=brown
      line=29
      width=2;

legend1 label=none
      position=(top center outside)
      offset=(0,0)
      mode=reserve;

axis1 label=(angle=90 h=2 color=black "Power") ;
axis2 label=(c=black h=2 "Sample size")
      order=(5 10 25 50)
      value=("5p" "10p" "25p" "50p") ;
axis3 label=(angle=90 h=2 color=black "Type1 Error") ;

```

```
axis4 order =(0 to 1 by 0.1) label=(angle=90 h=2 color=black "Type1 Error")
;
```

```
data T1N0 T1U0 T1L0 T2N0 T2U0 T2L0 T2N50 T2U50 T2L50 T2N90 T2U90 T2L90 T5N0
T5U0 T5L0 T5N50 T5U50 T5L50 T5N90 T5U90 T5L90;
set nhla.type1(rename=(Type1Normal=Parametric
Type1BootR=Bootstrap_Residuals Type1BootO=Bootstrap_Observations
Type1Perm=Observation_Permutation Type1FL=FreedmanLane_Permutation));
if no_of_par=1 then do;
  if (error=0 & rho=0) then output T1N0;
  else if (error=1 & rho=0) then output T1U0;
  else if (error=2 & rho=0) then output T1L0;
end;

if no_of_par=2 then do;
  if rho=0 then do;
    if error =0 then output T2N0;
    else if error =1 then output T2U0;
    else output T2L0;
  end;
  else if rho=0.5 then do;
    if error =0 then output T2N50;
    else if error =1 then output T2U50;
    else output T2L50;
  end;

  else if rho=0.9 then do;
    if error =0 then output T2N90;
    else if error =1 then output T2U90;
    else output T2L90;
  end;
end;

if no_of_par=5 then do;
  if rho=0 then do;
    if error =0 then output T5N0;
    else if error =1 then output T5U0;
    else output T5L0;
  end;
  else if rho=0.5 then do;
    if error =0 then output T5N50;
    else if error =1 then output T5U50;
    else output T5L50;
  end;

  else if rho=0.9 then do;
    if error =0 then output T5N90;
    else if error =1 then output T5U90;
    else output T5L90;
  end;
end;

run;
```

```

%macro comp(data,er,rho,p);

title "Probability of type 1 results for p=&p &er and rho=0.&rho";
proc gplot data=&DATA;
plot (Parametric--FreedmanLane_Permutation)*size/overlay haxis=axis2
vaxis=axis4 vref=0.05 legend=legend1 ;
plot (Parametric--FreedmanLane_Permutation)*size/overlay haxis=axis2
vaxis=axis3 vref=0.05 legend=legend1 ;

run;

%mend;

%comp (T1N0,Normal,0,1);
%comp (T1U0,Uniform,0,1);
%comp (T1L0,LOGN,0,1);
%comp (T2N0,Normal,0,2);
%comp (T2U0,Uniform,0,2);
%comp (T2L0,LOGN,0,2);
%comp (T5N0,Normal,0,5);
%comp (T5U0,Uniform,0,5);
%comp (T5L0,LOGN,0,5);

%comp (T2N50,Normal,50,2);
%comp (T2U50,Uniform,50,2);
%comp (T2L50,LOGN,50,2);
%comp (T5N50,Normal,50,5);
%comp (T5U50,Uniform,50,5);
%comp (T5L50,LOGN,50,5);

%comp (T2N90,Normal,90,2);
%comp (T2U90,Uniform,90,2);
%comp (T2L90,LOGN,90,2);
%comp (T5N90,Normal,90,5);
%comp (T5U90,Uniform,90,5);
%comp (T5L90,LOGN,90,5);

```

Power Graphs

```

goptions reset=all border;

symbol1 interpol=join
value=star
cv=red
ci=red
width=2;
symbol2 interpol=join
value=dot
cv=blue
ci=blue
line=3
width=2;
symbol3 interpol=join
value=circle
cv=green
ci=green
line=8
width=2;

```



```

symbol4 interpol=join
value=triangle
    cv=black
    ci=black
    line=33
    width=2;
symbol5 interpol=join
value=square
    cv=brown
    ci=brown
    line=29
    width=2;

legend1 label=none
    position=(top center outside)
    offset=(0,0)
    mode=reserve;

axis1 label=(angle=90 h=2 color=black "Power") ;
axis2 label=(c=black h=2 "Sample size")
    order=(5 10 25 50)
    value=("5p" "10p" "25p" "50p") ;
axis3 label=(angle=90 h=2 color=black "Type1 Error") ;
axis4 order =(0 to 1 by 0.1) label=(angle=90 h=2 color=black "Power") ;

data P1N0 P1U0 P1L0 P2N0 P2U0 P2L0 P2N50 P2U50 P2L50 P2N90 P2U90 P2L90 P5N0
P5U0 P5L0 P5N50 P5U50 P5L50 P5N90 P5U90 P5L90;
set nhla.POWER(rename=(PowerNormal=Parametric
PowerBootR=Bootstrap_Residuals PowerBootO=Bootstrap_Observations
PowerPerm=Observation_Permutation PowerFL=FreedmanLane_Permutation));
if no_of_par=1 then do;
    if (error=0 & rho=0) then output P1N0;
    else if (error=1 & rho=0) then output P1U0;
    else if (error=2 & rho=0) then output P1L0;
end;

if no_of_par=2 then do;
    if rho=0 then do;
        if error =0 then output P2N0;
        else if error =1 then output P2U0;
        else output P2L0;
    end;
    else if rho=0.5 then do;
        if error =0 then output P2N50;
        else if error =1 then output P2U50;
        else output P2L50;
    end;

    else if rho=0.9 then do;
        if error =0 then output P2N90;
        else if error =1 then output P2U90;
        else output P2L90;
    end;

```

```

end;

if no_of_par=5 then do;
  if rho=0 then do;
    if error =0 then output P5N0;
    else if error =1 then output P5U0;
    else output P5L0;
  end;
  else if rho=0.5 then do;
    if error =0 then output P5N50;
    else if error =1 then output P5U50;
    else output P5L50;
  end;

  else if rho=0.9 then do;
    if error =0 then output P5N90;
    else if error =1 then output P5U90;
    else output P5L90;
  end;

end;

end;

run;

%macro comp(data,er,rho,p);
title "Power results for p=&p &er rho=0.&rho";

proc gplot data=&DATA;
plot (Parametric--FreedmanLane_Permutation)*size/overlay haxis=axis2
vaxis=axis4 legend=legend1;
plot (Parametric--FreedmanLane_Permutation)*size/overlay haxis=axis2
vaxis=axis1 legend=legend1;

run;

%mend;

%comp (P1N0,Normal,0,1);
%comp (P1U0,Uniform,0,1);
%comp (P1L0,LOGN,0,1);
%comp (P2N0,Normal,0,2);
%comp (P2U0,Uniform,0,2);
%comp (P2L0,LOGN,0,2);
%comp (P5N0,Normal,0,5);
%comp (P5U0,Uniform,0,5);
%comp (P5L0,LOGN,0,5);

%comp (P2N50,Normal,50,2);
%comp (P2U50,Uniform,50,2);
%comp (P2L50,LOGN,50,2);
%comp (P5N50,Normal,50,5);
%comp (P5U50,Uniform,50,5);
%comp (P5L50,LOGN,50,5);

%comp (P2N90,Normal,90,2);
%comp (P2U90,Uniform,90,2);
%comp (P2L90,LOGN,90,2);
%comp (P5N90,Normal,90,5);

```

```
%comp (P5U90,Uniform,90,5);  
%comp (P5L90,LOGN,90,5);
```