Non-Parametric Bootstrap Methods in Analyzing CD4 Data

Nana Boateng, Qirui Huang, Akwasi Kusi-Appiah

The University of Memphis

December 12, 2012

Introduction

Regression Analysis

Bootstrap Graphs

Conclusion

Background

- The $CD4$ data frame has 20 rows and 2 columns. $CD4$ cells are carried in the blood as part of the human immune system. One of the effects of the HIV virus is that these cells die.

- The count of $CD4$ cells is used in determining the onset of full-blown AIDS in a patient. In this study of the effectiveness of a new anti-viral drug on HIV, 20 HIV-positive patients had their $CD4$ counts recorded and then were put on a course of treatment with this drug.

- After using the drug for one year, their $CD4$ counts were again recorded. The aim of the experiment was to show that patients taking the drug had increased $CD4$ counts which is not generally seen in HIV-positive patients.
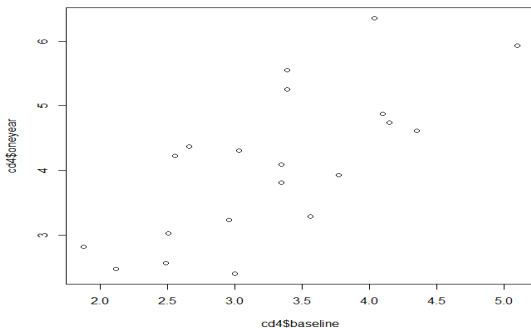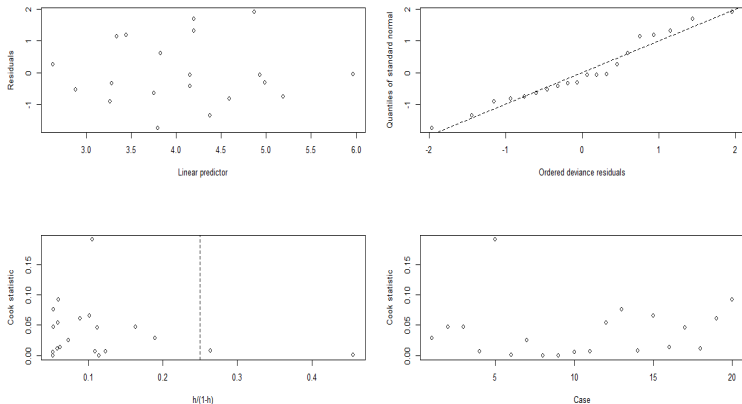
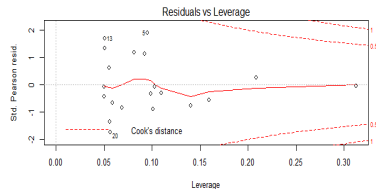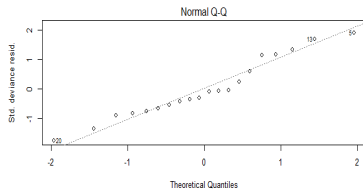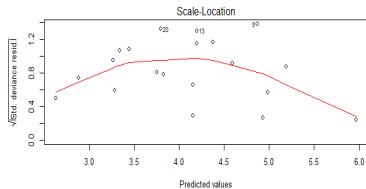# Regression of Baseline year on after oneyear

Figure : Regression of Baseline year on Data after oneyear
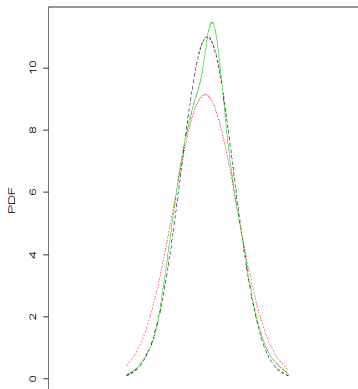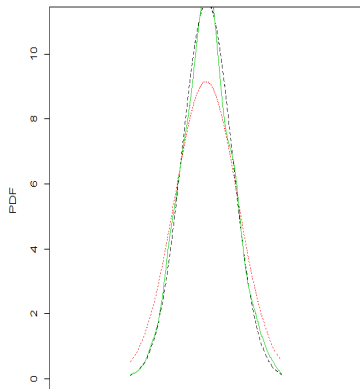
Diagnostic Plots of Residuals

Figure : Regression

Diagnostic Plots of Residuals

Figure : Regression
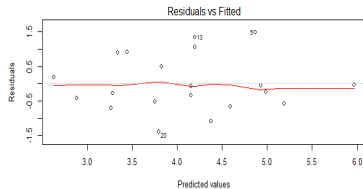
The plot of the the data after oneyear versus baseline year shows
there is some association between the two.We further derive a

regression model which shows

Kernel Density, $N(b, v)$ and $N(0, V_l)$

▶ The approximation of the probability distribution function of t^* can be obtained using simulation results

▶ The density estimates plot shows the effect of small sample size on the accuracy of the normal distributions

▶ The density estimates for $t^* - t$ based on 999 non-parametric simulations for the entire CD4 data set.
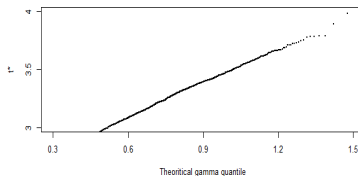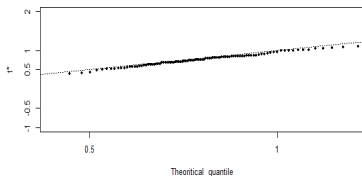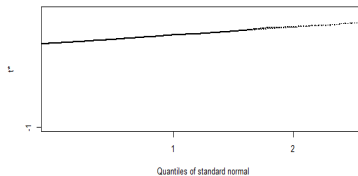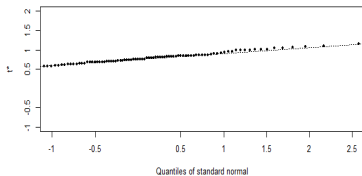
Bootsrap Average of Simulations
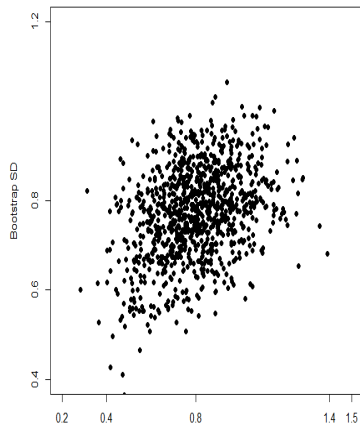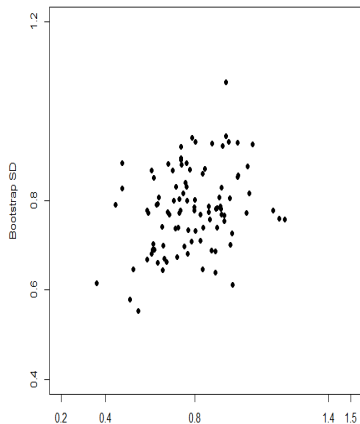
The qualitative feature to be read from these plots is that data standard deviation is proportional to data average
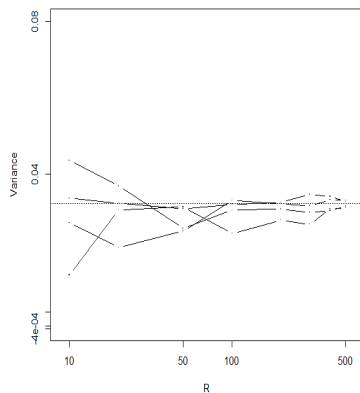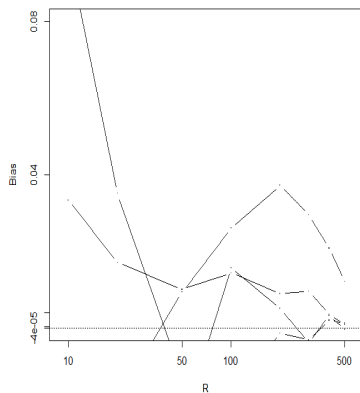
Histogram plots of Difference in Means with Corresponding Normal Density
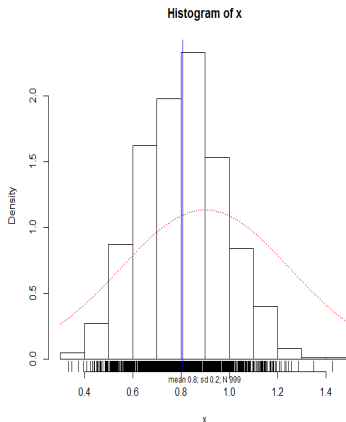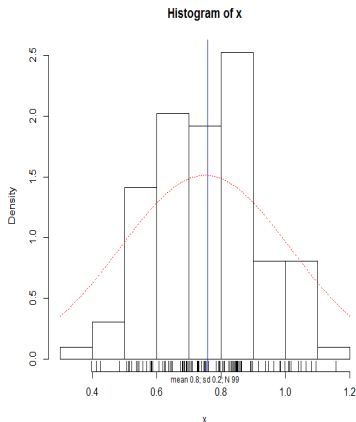
Quantile Plots of Standard Normal and Qgamma

Empirical Approximations

▶ The empirical approximations are justified by the law of large numbers

▶ The bias of the bootstrap sample converges to the exact value under the fitted model as R,increases

Conclusion

A paired t test of the CD4 data proved the alternative hypothesis to be true that there is significant difference between the mean of the CD4 counts in the baseline year and the CD4 counts after oneyear.This shows the anti-viral drug must have contributed significantly to theincrease in CD4 counts