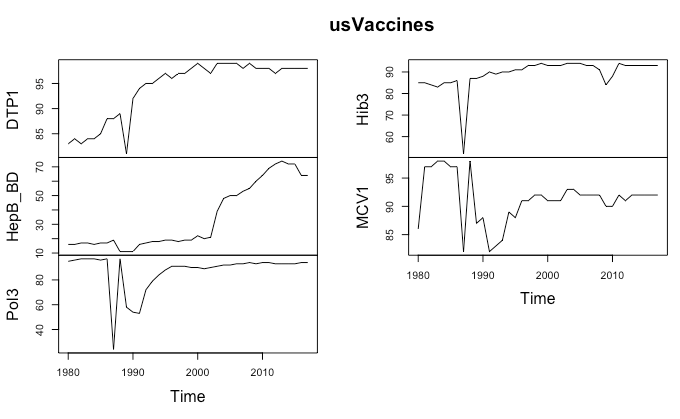
Sana Khan

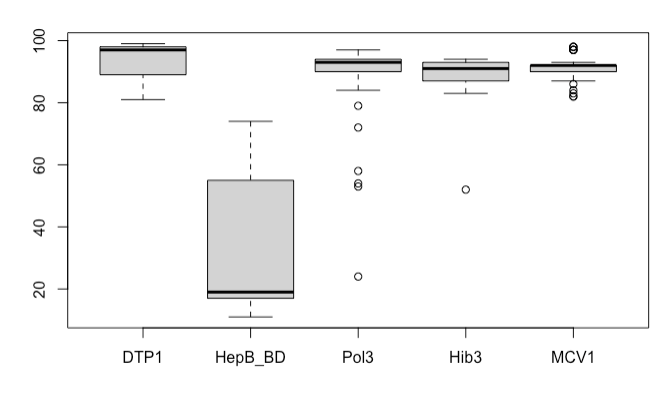
Final Exam

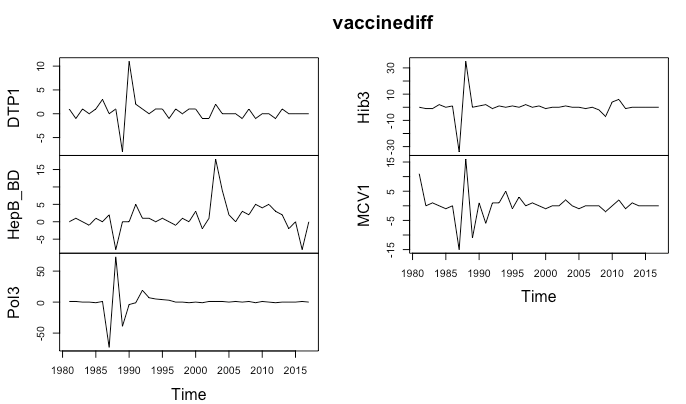
IST-772

1. How have U.S. vaccination rates varied over time? Are vaccination rates increasing or decreasing? Which vaccination has the highest rate at the conclusion of the time series? Which vaccination has the lowest rate at the conclusion of the time series? Which vaccine has the greatest volatility?

The data provided in the us.Vaccines data set shows vaccine rates reported by the WHO from 1980 – 2017 for Diphtheria/Pertussis/Tetanus vaccine; Hepatitis B, (Birth Dose), Pol3 (Polio third dose), Hib3 (Influenza third dose) and MCV1 (Measles first dose).

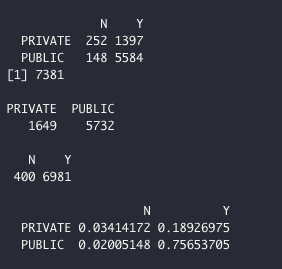






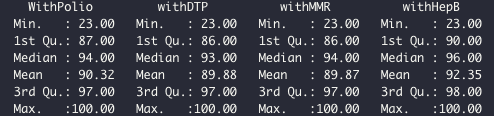
For US vaccination rates, all vaccines showed a lot of variability between 1985 and 1990 but after the 1990’s the rates had stabilized and had a general upward trend. Hep B has another change in the early 2000s, with a spike which then falls. By the end of the study in 2017, all vaccines except Hep B had 90% vaccination rates, with DTP being the highest at 99% and the Flu vaccine the second highest at 98%. Hep B had the lowest vaccination rates, about 70% at the end of the study. Polio also had the most volatility, which is evident in the boxplot and the changepoint which had 7 changes in the mean and 16 in the variance.

1. What proportion of public schools reported vaccination data? What proportion of private schools reported vaccination data? Was there any credible difference in overall reporting proportions between public and private schools?



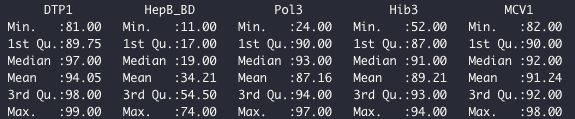
76% of public schools and 19% of private schools reported vaccination data in 2013. Yes, there is a credible difference in reporting proportions, with public schools having almost 4 times the reporting rate as private schools.

1. What are 2013 vaccination rates for individual vaccines (i.e., DOT, Polio, MMR, and HepB) in California public schools? How do these rates for individual vaccines in California districts compare with overall US vaccination rates (make an informal comparison to the final observations in the time series)?



California School Vaccination Rates

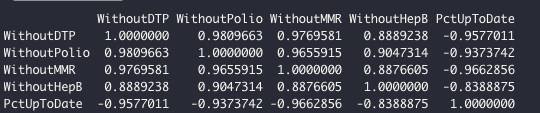
The mean vaccination rates for California public schools in 2013 was 90% for Polio, DTP and MMR. The mean vaccination rate for Hep3 was 92%.

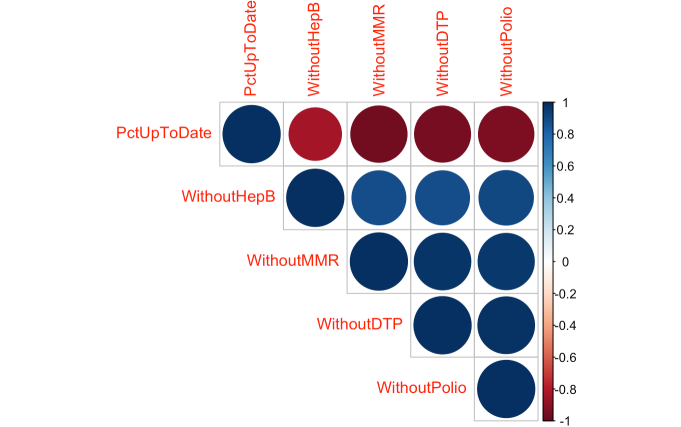


US Vaccination Rates

The mean vaccination rates for DTP was 94%, HepB was 34% and Polio was 97%. The two populations have very different samples, with one being a school district students K-12 and the other being the US population. Compared to the US, DTP was close but a bit lower for the school data set. Polio vaccination rates for the school population was higher than US data. The school data set did outperform the US HepB vaccination rates with a 92% compared to 34%.

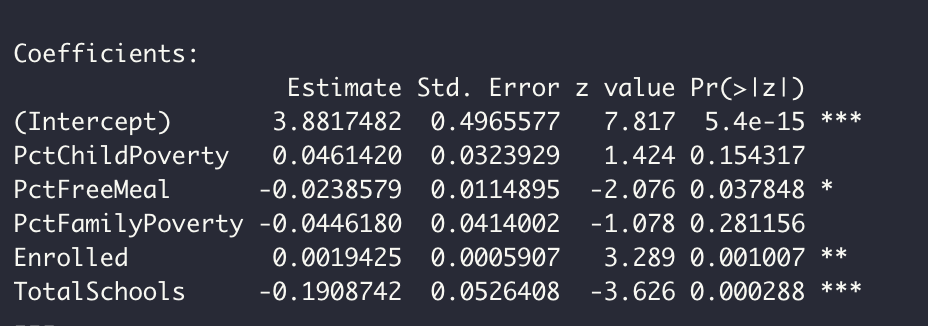
1. Among districts, how are the vaccination rates for individual vaccines related? In other words, if students are missing one vaccine are they missing all of the others?



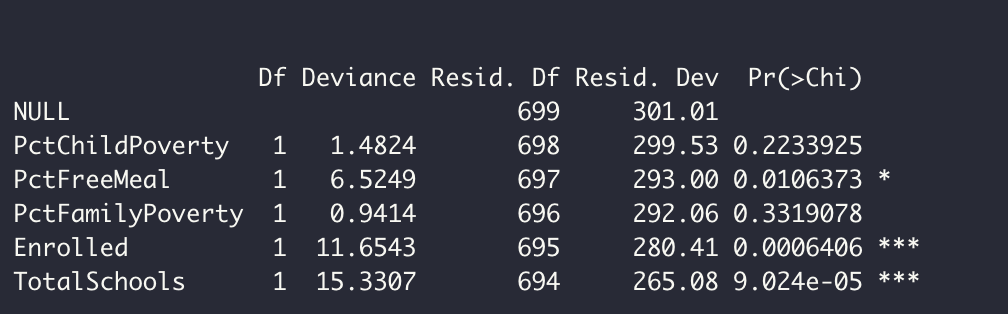


Vaccination rates are highly correlated, with the highest being among DTP, Polio and Flu. HepB is not as highly correlated as the others, but still has a strong correlation with the other vaccines as well. This correlation matrix shows that if one vaccine is missing than the chances of other vaccines being missed are high.

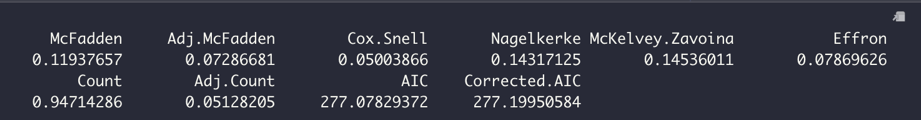
1. What variables predict whether or not a district’s reporting was complete?



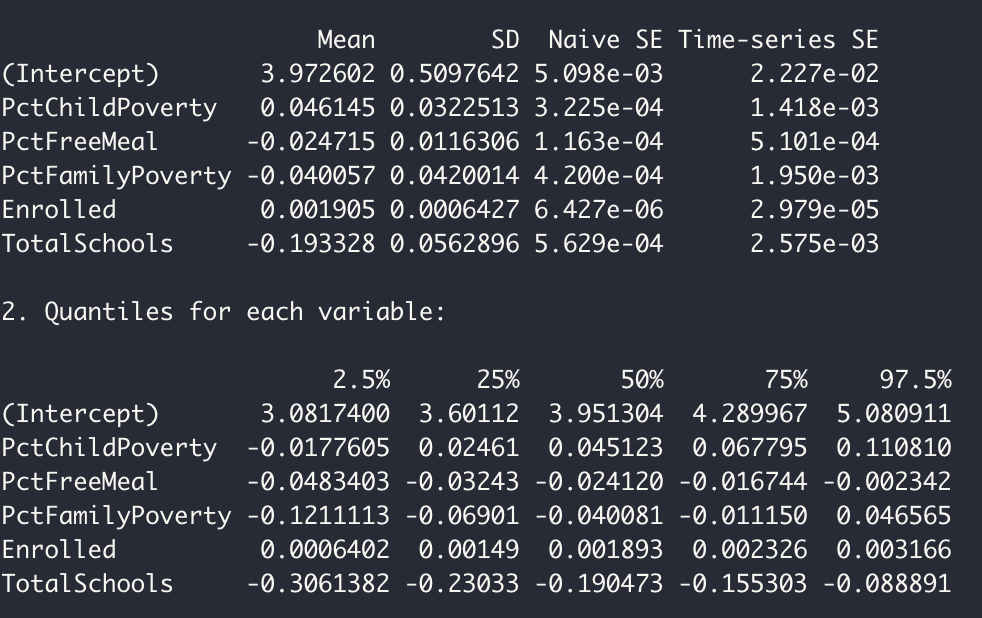
Frequentist Model



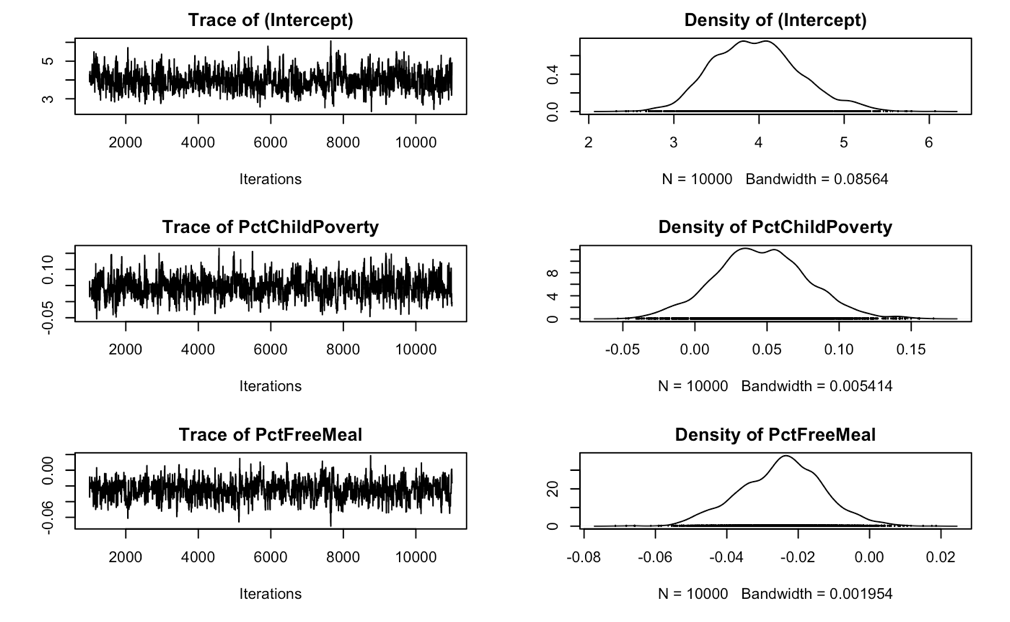
Chi-Square

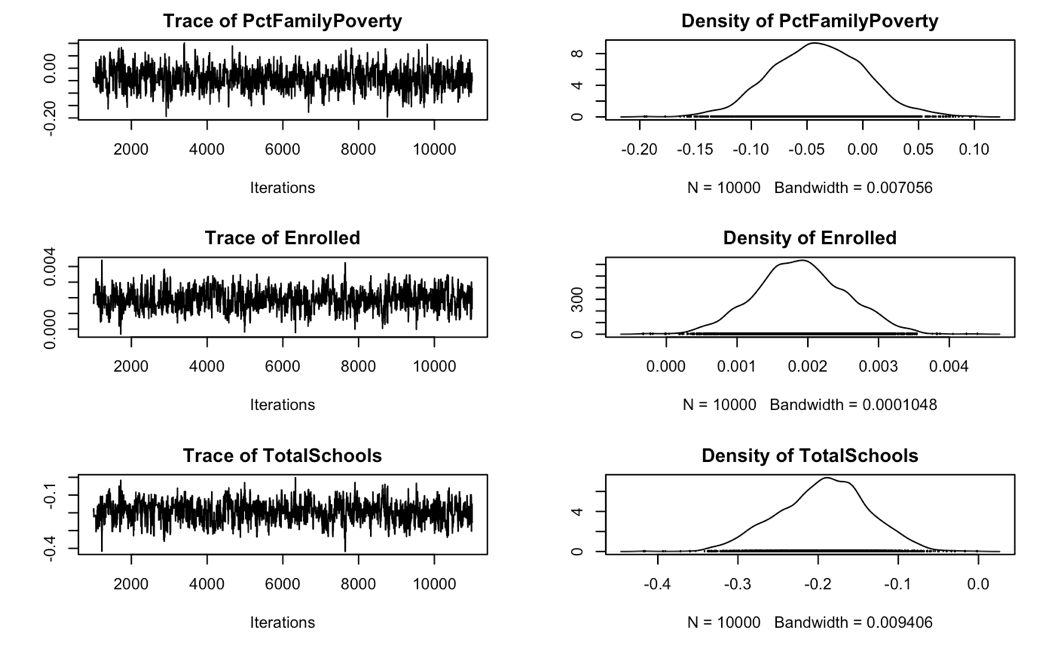


PseudoR2



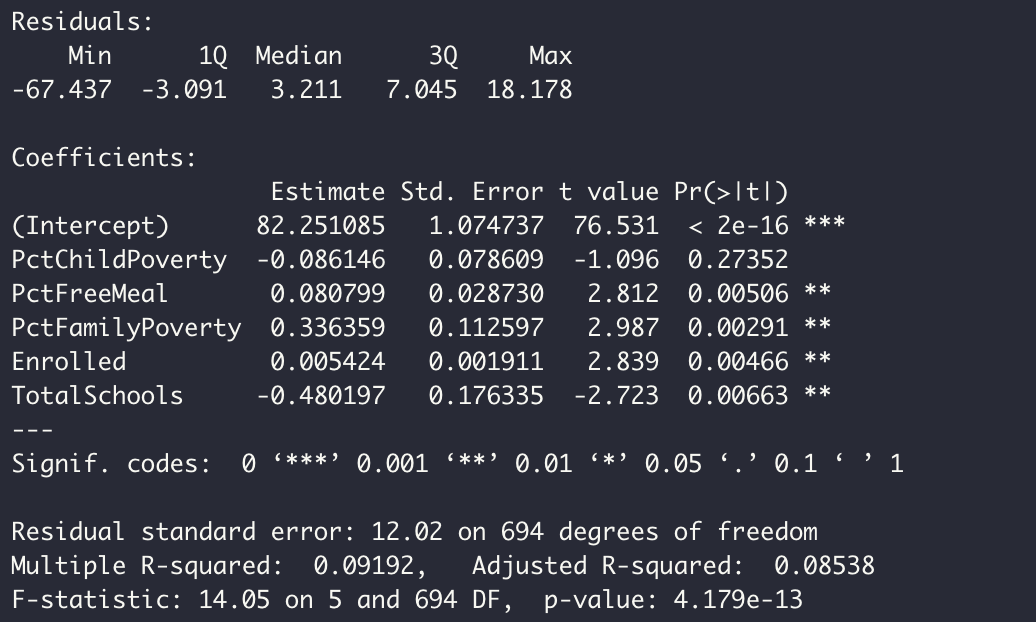
Bayesian

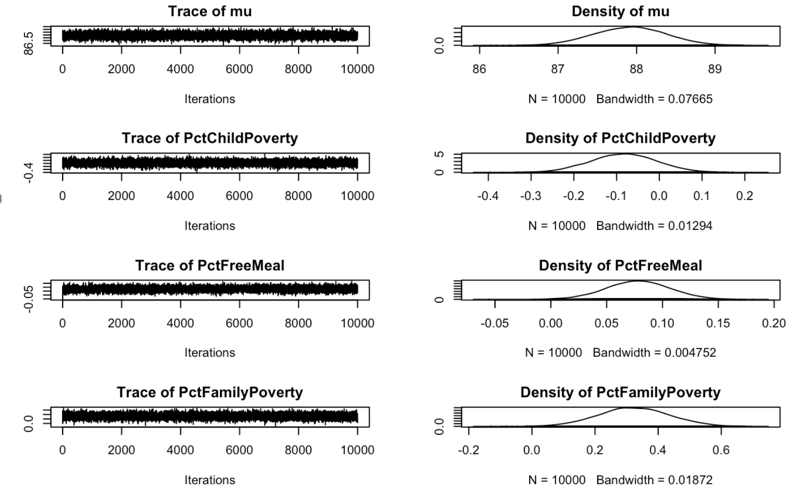


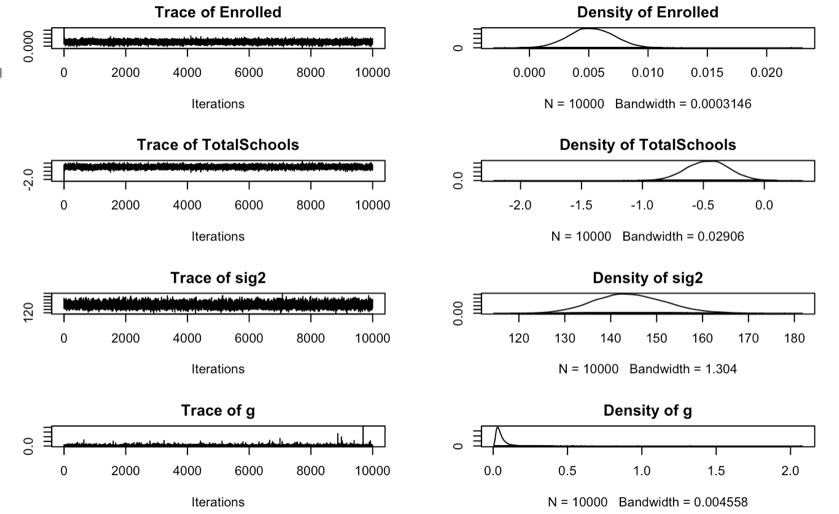


The variables being used as predictors for a districts vaccine reporting are PctChildPoverty, PctFreeMeal, PctFamilyPoverty, Enrolled, and TotalSchools. The intercept, PctFreeMeal, Enrolled and TotalSchools are significant because the p-value is less than our alpha. Therefore, we can reject the null hypothesis that there is no correlation between these variables and the districts reporting numbers. This is supported by the Bayesian model, where the density for those variable do not overlap with 0. PctChildPoverty, PctFreeMeal, PctFamilyPoverty have a p-value greater than alpha so we fail to reject the null hypothesis. This is also further supported by the density plots for each which overlap with 0.

1. What variables predict the percentage of all enrolled students with completely up-to-date vaccines?

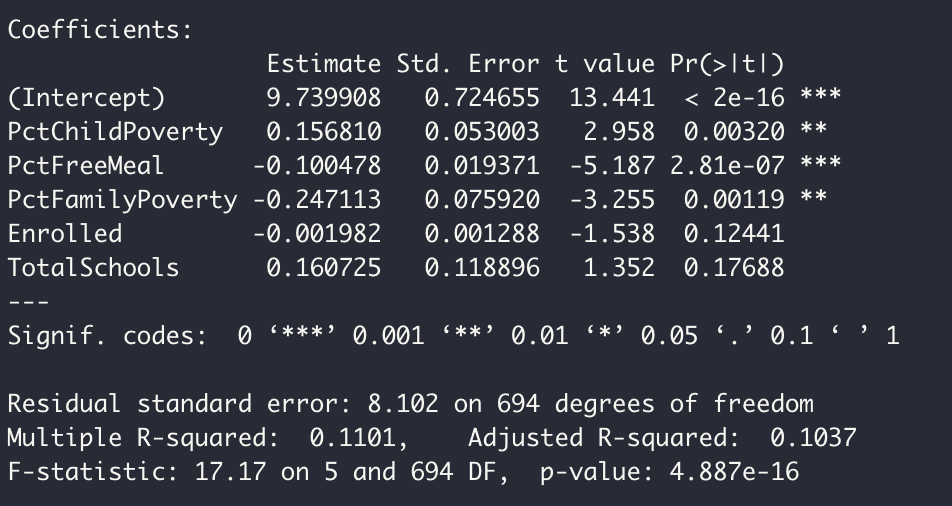


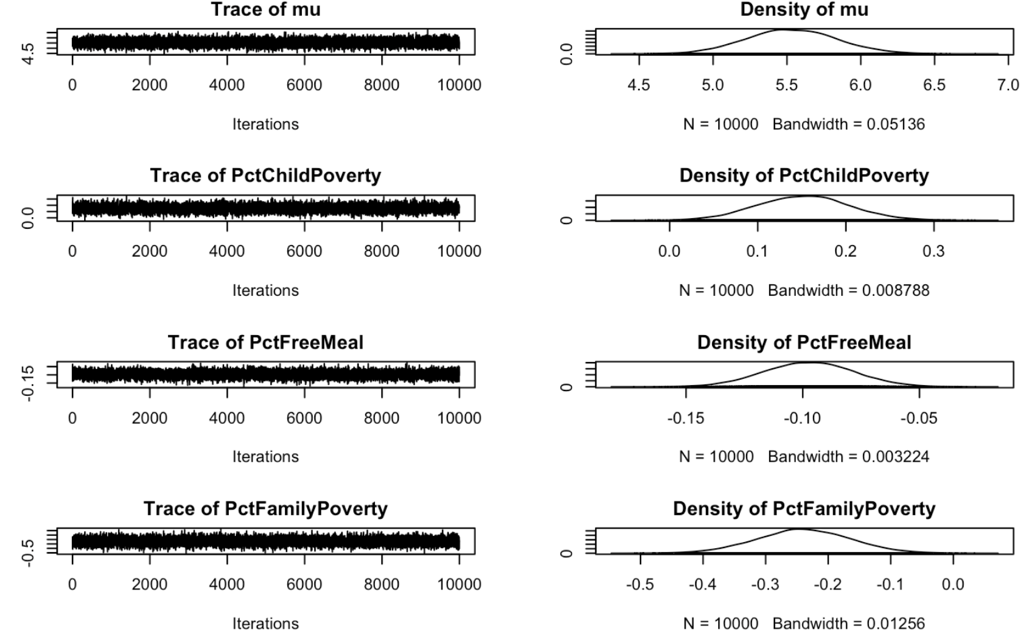


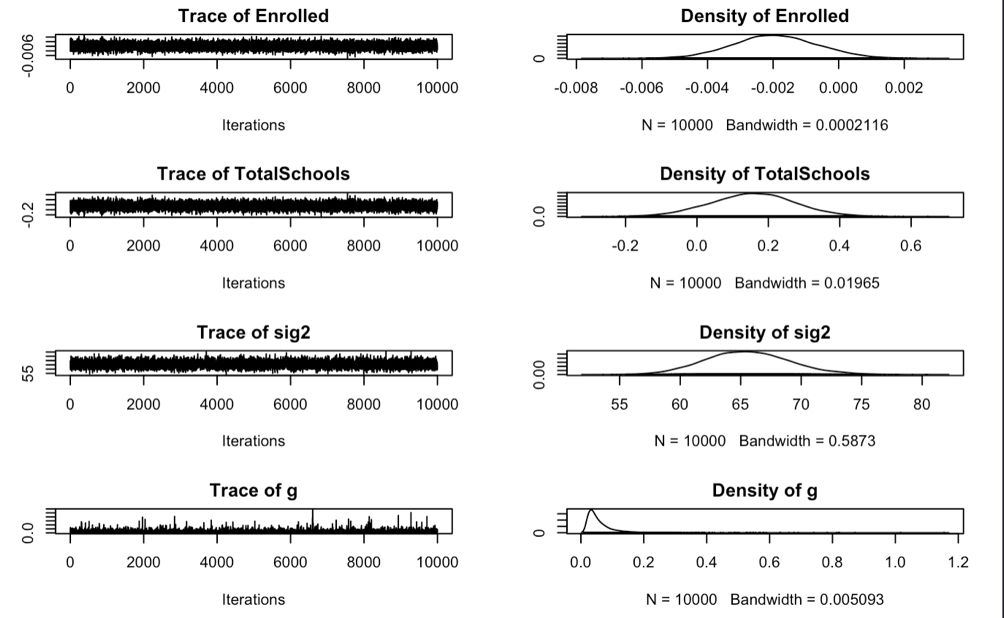


The variables, PctFreeMeal, PctFamilyPoverty, Enrolled and TotalSchools were significant in predicting the number of enrolled students with up to date vaccines. The p-value for those variables is less than our alpha so we can therefore reject the null hypothesis that these variables do not have a correlation on reporting the percentage of all enrolled students with completely up-to-date vaccines. The Bayesian plots also support what the frequentist model shows, as there is no overlap with 0 in the plots. The variables PctChildPoverty and PctFreeMeal have a p-value greater than alpha and overlap with 0 so we fail to reject the null hypothesis.

1. What variables predict the percentage of all enrolled students with belief exceptions?







The variables that were statistically significant in predicting the percentage of all enrolled students with belief exceptions are PctChildPoverty, PctFreeMeal, PctFamilyPoverty. These variables had a p-value that was less than alpha. The Bayesian plot also shows those variables do not cross 0, so we can therefore reject the null hypothesis. For the variables Enrolled and TotalSchools the p-value is greater than alpha and the HDI does cross 0. We fail to reject the null hypothesis that those variables do have an influence on number of enrolled students with belief expectations.

1. What’s the big picture, based on all of the foregoing analyses? The staff member in the state legislator’s office is **interested to know how to allocate financial assistance to school districts to improve both their vaccination rates and their reporting compliance**. What have you learned from the data and analyses that might inform this question?

Based on the data provided, California school districts had at least 90% of students vaccinated for DTP1, HepB, Pol3, Hib3 and MCV1. Public schools also have a much higher rate for reporting vaccines, almost 4x higher than private schools. Private schools have a lower rate of reporting and potentially lower rates overall. This could be due to the fact that these parents are wealthier and may have objections to having their children vaccinated. These parents may also have the resources to tend to a child who could get sick more easily than a non affluent family. The models also showed that variables like PctFreeMeal, Enrolled and TotalSchools influenced a districts reporting. Variables like PctFreeMeal, PctFamilyPoverty, Enrolled and TotalSchools are also correlated to the number of students who have up to date vaccines. The data also shows that vaccination rates are correlated, where if a student has missed one vaccine, they have most likely missed all. It also seems that students with belief exceptions have an impact on vaccination rates and reporting and that the variables PctFreeMeal, PctFamilyPoverty and PctChildPoverty are correlated to having a belief exception.

The areas to focus on are getting reporting and vaccination rates up in private schools, which could be difficult since these schools tend to have more students with belief exceptions. Vaccinations should also start at a younger age because as noted above if one vaccine is missed the chances of others being missed is highly correlated. To get vaccination rates up, programs that focus on providing free meals and helping families in poverty should also be considered.