Abusers of Cocaine in Virginia Less Likely to Rehabilitate?

Summary

In the analysis of annual data of people discharged from substance abuse treatment programs, we found that young people addicted to alchohol and marijuana are more likely to complete their treatments, compared to those addicted to “hard” drugs, including Heroin, Opiates, Methamphetamine, and Cocaine. Apart from that, we also came up with evidence that some American states indeed have better treatment programs, and the completion rates in these states were much higher.

Introduction

As it is known that the abuse of substances(drugs, achohol, cocaine, etc.) causes serious health problems and affects the quality of lives, some abusers are actively involved in abuse treatment programs. We used the data provided by Treatment Episode Data Set – Discharges (TEDS-D), trying to explore whether young people addicted to “hard” drugs(Heroin, Opiates, Methamphetamine, Cocaine) are less likely to complete treatment programs. We also want to know whether there are significant differences among American states in the efficiency of treatment programs. These are administrative records data, collected using longitudinal time method.

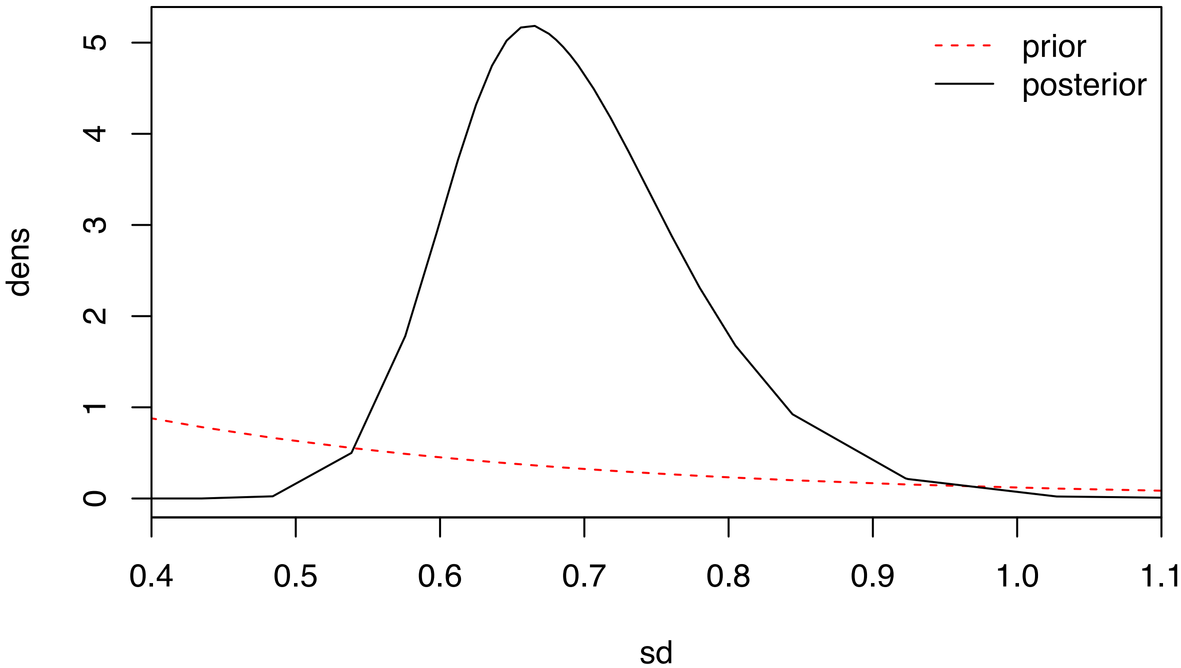
Methods

We used the dataset released by ICPSR. In the study during 2011, administrative records method was used to collect data from substance abuse treatment programs as reported to state substance abuse agencies. As an exploratory tool, based on historical data, we assumed penalized complexity prior distribution for our nested random effects, American states and towns. Specifically, for states, we assumed the probability of point estimates greater than 0.9 being 0.05, and for towns , we assumed $P(\mu > 0.8) = 0.05$. After obtaining the data, we used INLA to perform the approximate Bayesian inference for the posterior distribution, which is proportional to prior distribution. We used glmm method to fit a logistic model, since we are willing to know the relationship between the binary variable(completed the program or not) and a combination of fixed effects(substances, age, gender, ethnicity, homeless or not) and random effects(states, town). All main effects were retained, eliminating effects not significantly interacted with them. Our model is as follows:

$$\log(odds) = X\beta + S_i + T_{ij} \tag{2}$$

Results

The odds of treatment completion table shows that holding the other effects constant, we can fit a logistic model with respect to the fixed effect, substances(alchohol, heroin, etc.). In this model, we treated marijuana as the reference group, the odds ratio between marijuana and alchohol is 1:1.642, which means people in the alchohol group were more likely to complete the treatment. Also, based on the data, there is a 95% probability that the true effect mean is in the interval [1.608, 1.677]. As for the other four groups(Heroin, Opiates, Methamphetamine, Cocaine), the odds are relatively lower than the odds of marijuana group. Hence, we have found some evidence that people addicted to “hard” drugs are less likely to complete their treatments, compared with marijuana, and alchohol as well. Then we want to analyze the differences in the efficiency of treatment programs among states in the U.S.. Indeed, we found that treatment programs in some states/towns performed much better than others. As shown in the table below, we found that the point estimate(posteior mean) of Virginia, -2.9, is the lowest of all states. And there is a 95% probability that the estimate falls within the credible interval [-3.2, -2.5]. Following are New Mexico(-1.1), North California(-0.8), etc. This indicates that programs in these states showed very low efficiency in treating people, and they negatively affected the average rate of program completion over the country. In contrast, it looks like some other states, such as Colorado, Florida, Massachusetts, etc., have much more efficient treatment programs and contributed positive effects to the overall average completion rate. Especially for Florida, its point estimate, 1.0, is significantly 3.9 higher than that of Virginia, and the credible interval [0.7, 1.3] confirms that with a probablity of 95%, the estimate would be at least 0.7.



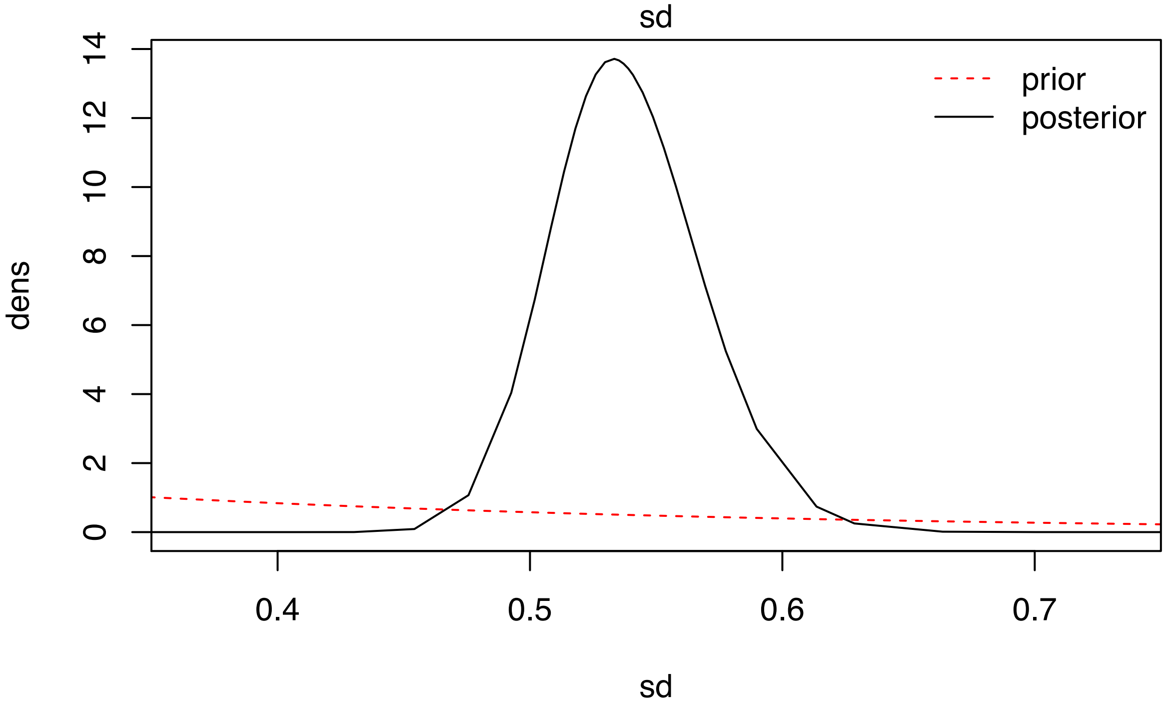


Table 2: Posterior Means and Credible Intervals for Model Effects

	0.5quant	0.025quant	0.975quant
(Intercept)			
(Intercept)	0.681	0.549	0.844
SUB1			
ALCOHOL	1.642	1.608	1.677
HEROIN	0.898	0.875	0.921
OTHER OPIATES AND SYNTHET	0.924	0.897	0.952
METHAMPHETAMINE	0.982	0.944	1.022
COCAINE/CRACK	0.876	0.834	0.920
GENDER			
FEMALE	0.895	0.880	0.910
raceEthnicity			
Hispanic	0.829	0.810	0.849
BLACK OR AFRICAN AMERICAN	0.685	0.669	0.702
AMERICAN INDIAN (OTHER TH	0.729	0.680	0.782
OTHER SINGLE RACE	0.864	0.810	0.921
TWO OR MORE RACES	0.851	0.790	0.917
ASIAN	1.133	1.038	1.236
NATIVE HAWAIIAN OR OTHER	0.847	0.750	0.955
ASIAN OR PACIFIC ISLANDER	1.451	1.225	1.720
ALASKA NATIVE (ALEUT, ESK	0.844	0.623	1.144
homeless			
TRUE	1.015	0.983	1.048
SD			
STFIPS	0.685	0.558	0.880
TOWN	0.537	0.485	0.601

Table 3: Posterior Means and Credible Intervals for States

ID	mean	0.025q	0.975q	ID	mean	0.025q	0.975q
ALABAMA	0.2	-0.3	0.8	MONTANA	-0.2	-1.0	0.6
ALASKA	0.0	-0.8	0.8	NEBRASKA	0.8	0.4	1.2
ARIZONA	0.0	-1.3	1.3	NEVADA	-0.1	-0.8	0.5
ARKANSAS	-0.1	-0.7	0.5	NEW HAMPSHIRE	0.2	-0.3	0.7
CALIFORNIA	-0.3	-0.6	0.0	NEW JERSEY	0.5	0.2	0.8
COLORADO	0.5	0.1	1.0	NEW MEXICO	-1.2	-1.9	-0.5
CONNECTICUT	0.1	-0.4	0.7	NEW YORK	-0.3	-0.6	0.0
DELAWARE	1.0	0.7	1.3	NORTH CAROLINA	-0.8	-1.1	-0.5
WASHINGTON DC	-0.3	-0.6	0.1	NORTH DAKOTA	-0.3	-1.0	0.4
FLORIDA	1.0	0.7	1.4	OHIO	-0.2	-0.6	0.1
GEORGIA	-0.2	-0.9	0.4	OKLAHOMA	0.6	0.0	1.1
HAWAII	0.2	-0.6	1.0	OREGON	0.1	-0.3	0.5
IDAHO	-0.2	-1.0	0.6	PENNSYLVANIA	0.0	-1.3	1.3
ILLINOIS	-0.5	-0.8	-0.2	RHODE ISLAND	-0.2	-0.6	0.3
INDIANA	-0.1	-0.9	0.8	SOUTH CAROLINA	0.4	0.0	0.7
IOWA	0.4	0.1	0.7	SOUTH DAKOTA	0.5	-0.3	1.3
KANSAS	-0.2	-0.6	0.1	TENNESSEE	0.3	-0.2	0.7
KENTUCKY	-0.1	-0.5	0.2	TEXAS	0.6	0.3	0.9
LOUISIANA	-0.5	-1.0	-0.1	UTAH	0.1	-0.5	0.7
MAINE	0.1	-0.7	1.0	VERMONT	-0.2	-1.1	0.6
MARYLAND	0.5	0.2	0.8	VIRGINIA	-2.9	-3.3	-2.5

ID	mean	0.025q	0.975q	ID	mean	0.025q	0.975q
MASSACHUSETTS	0.8	0.4	1.2	WASHINGTON	-0.1	-0.5	0.3
MICHIGAN	-0.4	-0.7	0.0	WEST VIRGINIA	0.0	-1.3	1.3
MINNESOTA	0.4	0.0	0.9	WISCONSIN	0.0	-1.3	1.3
MISSISSIPPI	0.0	-1.3	1.3	WYOMING	0.0	-1.3	1.3
MISSOURI	-0.4	-0.7	-0.1	PUERTO RICO	0.6	-0.1	1.3

Appendix

```r
library("nlme")
library("INLA")
data("MathAchieve", package = "MEMSS")
mod <- lme(MathAch ~ Minority + Sex + SES, random = ~1 | School,
           data=MathAchieve)
tab1 <- Pmisc::lmeTable(mod)[,-3]
knitr::kable(tab1, digits = 2, escape=FALSE,
             caption = "Linear Mixed Effect Table")

download.file("http://pbrown.ca/teaching/appliedstats/data/drugs.rds",
              "drugs.rds")
xSub = readRDS("drugs.rds")

forInla = na.omit(xSub)
forInla$y = as.numeric(forInla$completed)
library("raster")
library("INLA")
library("Pmisc")
library("Hmisc")
library("data.table")
ires = inla(y ~ SUB1 + GENDER + raceEthnicity + homeless +
              f(STFIPS,hyper=list(prec=list(
                prior='pc.prec', param=c(0.9, 0.05)))) +
              f(TOWN, hyper=list(prec=list(
                prior='pc.prec', param=c(0.8, 0.05)))),
            data=forInla,
            family='binomial',
            inla.link.logit,
            control.inla = list(strategy='gaussian',
                                int.strategy='eb'))

sdState = Pmisc::priorPostSd(ires)

do.call(matplot, sdState$STFIPS$matplot)
do.call(legend, sdState$legend)

do.call(matplot, sdState$TOWN$matplot)
do.call(legend, sdState$legend)

toPrint = as.data.frame(rbind(exp(ires$summary.fixed[, c(4, 3, 5)]),
                              sdState$summary[, c(4, 3, 5)]))
sss = "^(raceEthnicity|SUB1|GENDER|homeless|SD)
(.[[:digit:]]+.[[:space:]]+| for )?"
toPrint = cbind(variable = gsub(paste0(sss, ".*"),
                                "\\1",
                                rownames(toPrint)),
                category = substr(gsub(sss, "",
                                       rownames(toPrint)), 1, 25),
                toPrint)

Pmisc::mdTable(toPrint, digits = 3, mdToTex = TRUE,
```

```r
               guessGroup = TRUE,
               caption = "Posterior Means and Credible Intervals
               for Model Effects")

ires$summary.random$STFIPS$ID = gsub("[[:punct:]]|[[:digit:]]",
                                     "", ires$summary.random$STFIPS$ID)
ires$summary.random$STFIPS$ID = gsub("DISTRICT OF COLUMBIA",
                                     "WASHINGTON DC", ires$summary.random$STFIPS$ID)
toprint = cbind(ires$summary.random$STFIPS[1:26, c(1, 2, 4, 6)],
                ires$summary.random$STFIPS[-(1:26),
                                           c(1, 2, 4, 6)])
colnames(toprint) = gsub("uant", "", colnames(toprint))
knitr::kable(toprint, digits = 1,
             caption = "Posterior Means and Credible
             Intervals for States")
```