

Assignment 3

Litong Zhang

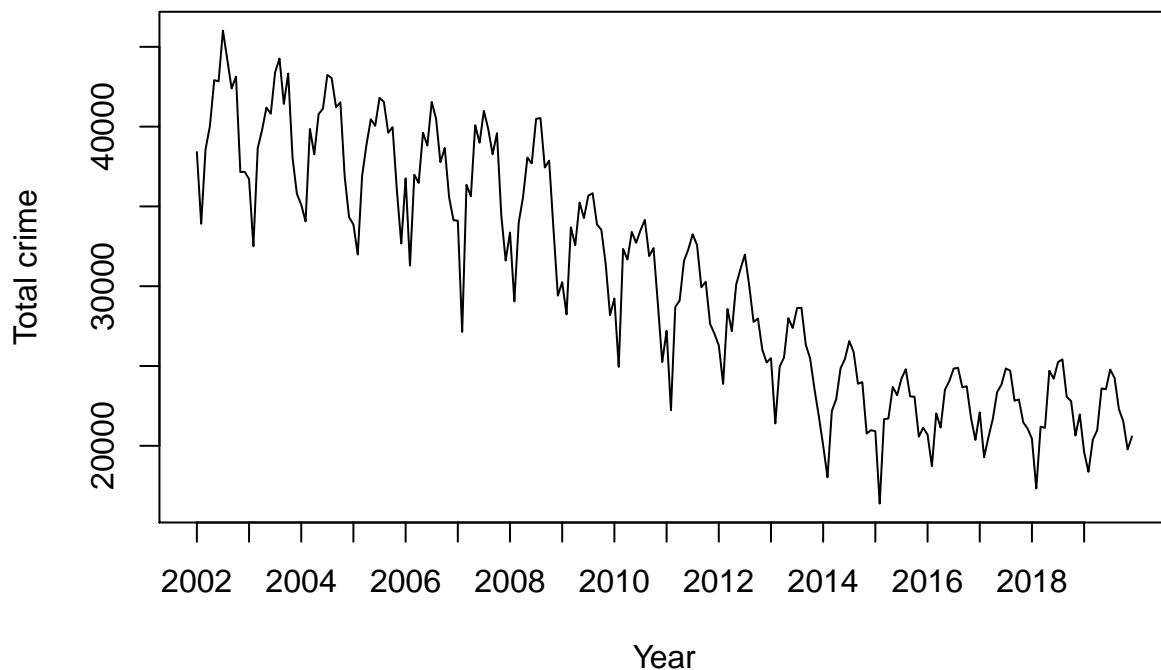
4/25/2021

Exercise 1

```
setwd("~/Desktop/613")
pop <- read.csv('population.csv')
crime_long <- read.csv('crime_long.csv')
officer <- read.csv('officers.csv')
```

Exercise 2

```
crime_ts <- aggregate(crimes ~ crime_month, crime_long, sum)
crime_ts <- ts(crime_ts$crimes, frequency = 12, start = c(2002, 1))
plot.ts(crime_ts, xlab = 'Year', ylab = 'Total crime', xaxt = 'n')
axis(side = 1, at = 2002:2019)
```



```
crime_long <- aggregate(crimes ~ crime_month + district + crime_type, crime_long, sum)
crime_drug <- crime_long[crime_long$crime_type == 'drug', c('crime_month', 'district', 'crimes')]
```

```

crime_property <- crime_long[crime_long$crime_type == 'property', c('crime_month', 'district', 'crimes')]
crime_violent <- crime_long[crime_long$crime_type == 'violent', c('crime_month', 'district', 'crimes')]
crime_other <- crime_long[crime_long$crime_type == 'other', c('crime_month', 'district', 'crimes')]
crime <- unique(crime_long[,c('crime_month', 'district')])
crime <- cbind.data.frame(
  crime,
  drug = crime_drug$crimes,
  property = crime_property$crimes,
  violent = crime_violent$crimes,
  other = crime_other$crimes
)
pop_crime <- merge(pop, crime, by.x = c('month', 'district'), by.y = c('crime_month', 'district'))

pop_crime$crime <- rowSums(pop_crime[,c('drug', 'property', 'violent', 'other')])
pop_crime$crime_per <- pop_crime$crime / pop_crime$tot_pop
pop_crime$violent_per <- pop_crime$violent / pop_crime$tot_pop
pop_crime$property_per <- pop_crime$property / pop_crime$tot_pop
pop_crime$black <- pop_crime$tot_black / pop_crime$tot_pop
pop_crime$white <- pop_crime$tot_white / pop_crime$tot_pop
pop_crime$hisp <- pop_crime$tot_hisp / pop_crime$tot_pop

```

Exercise 3

```

officer_pop_crime <- merge(officer, pop_crime, by.x = c('month', 'unit'), by.y = c('month', 'district'))
# arrest ~ tenure + crime + p50_inc + black + white + hisp
lm3 <- lm('arrest ~ tenure + crime + p50_inc + black + white + hisp', officer_pop_crime)

```

Exercise 4

```

# arrest ~ tenure + crime + p50_inc + black + white + hisp + factor(unit) + factor(month)
lm4 <- lm('arrest ~ tenure + crime + p50_inc + black + white + hisp + factor(unit) + factor(month)', of

```

Exercise 5

```

# There are too many NUID. The computation power cannot deal with such a large amount of fixed effects.
# Therefore, I take only the 10% sample for the following analysis.
officer_pop_crime <- officer_pop_crime[officer_pop_crime$NUID <= 3200,]
# arrest ~ NUID + tenure + crime + p50_inc + black + white + hisp + factor(unit) + factor(month)
# within (based on monthly average for each NUID):
officer_pop_crime <- cbind(
  officer_pop_crime,
  with(officer_pop_crime,
    data.frame(
      arrest_d = arrest - ave(arrest, NUID),
      tenure_d = tenure - ave(tenure, NUID),
      crime_d = crime - ave(crime, unit),
      p50_inc_d = p50_inc - ave(p50_inc, unit),

```

```

    black_d = black - ave(black, unit),
    white_d = white - ave(white, unit),
    hisp_d = hisp - ave(hisp, unit)
  )
)
)
lm5_within <- lm('arrest_d ~ tenure_d + crime_d + p50_inc_d + black_d + white_d + hisp_d + factor(unit)')
# between (based on monthly average for each NUID):
pop_crime_mean <- officer_pop_crime[,c('unit', 'crime', 'p50_inc', 'black', 'white', 'hisp')]
pop_crime_mean <- aggregate(. ~ unit, pop_crime_mean, mean)
colnames(pop_crime_mean)[-1] <- paste(colnames(pop_crime_mean)[-1], '_mean', sep = '')
officer_pop_crime_mean <- officer_pop_crime[,c('unit', 'NUID', 'tenure', 'arrest')]
officer_pop_crime_mean <- aggregate(. ~ NUID + unit, officer_pop_crime_mean, mean)
officer_pop_crime_mean <- merge(officer_pop_crime_mean, pop_crime_mean, by = 'unit')
lm5_between <- lm('arrest ~ tenure + crime_mean + p50_inc_mean + black_mean + white_mean + hisp_mean + factor(unit)')
# first difference:
officer_pop_crime_d <- officer_pop_crime[,c('unit', 'NUID', 'month', 'tenure', 'arrest')]
officer_pop_crime_d$month <- as.Date(officer_pop_crime_d$month)
officer_pop_crime_d <- officer_pop_crime_d[order(officer_pop_crime_d$NUID, officer_pop_crime_d$month),]
officer_pop_crime_d <- tail(officer_pop_crime_d, -1) - head(officer_pop_crime_d, -1)
colnames(officer_pop_crime_d)[-1:3] <- c('td_tenure', 'td_arrest')
officer_pop_crime_d <- rbind(c(0, 1, 0, 0, 0), officer_pop_crime_d)
officer_pop_crime_d[officer_pop_crime_d$NUID != 0, c('td_tenure', 'td_arrest')] <- NA
officer_pop_crime_d <- officer_pop_crime_d[, c('td_tenure', 'td_arrest')]
officer_pop_crime <- cbind(officer_pop_crime, officer_pop_crime_d)
lm5_fd <- lm('td_arrest ~ td_tenure + crime + p50_inc + black + white + hisp + factor(unit)', officer_pop_crime)
# GMM:
# do not use the function since there are too many dummies:
gmm <- function(y, X, Z = X){
  inv_tX_X <- solve(t(X) %*% X)
  tX_y <- t(X) %*% y
  if (nargs() == 2){
    # solve(t(X) %*% X) %*% (t(X) %*% y)
    inv_tX_X %*% tX_y
  } else{
    # solve((t(Z) %*% X) %*% solve(t(X) %*% X) %*% (t(X) %*% Z)) %*% ((t(Z) %*% X) %*% solve(t(X) %*% X) %*% tX_y)
    tZ_X <- t(Z) %*% X
    tX_Z <- t(X) %*% Z
    tZ_X_inv_tX_X <- tZ_X %*% inv_tX_X
    solve(tZ_X_inv_tX_X %*% tX_Z) %*% (tZ_X_inv_tX_X %*% tX_y)
  }
}
y <- matrix(officer_pop_crime$arrest)
X <- officer_pop_crime[,c('tenure', 'crime', 'p50_inc', 'black', 'white', 'hisp', 'unit', 'month', 'NUID')]
unit <- unique(officer_pop_crime$unit)[order(unique(officer_pop_crime$unit))]
X$intercept <- 1
for (i in 2:length(unit)){
  X[,paste('unit', i, sep = '.')] <- 0
  X[X$unit == unit[i], paste('unit', i, sep = '.')] <- 1
}
month <- unique(officer_pop_crime$month)[order(unique(officer_pop_crime$month))]
for (i in 2:length(month)){
  X[,paste('month', i, sep = '.')] <- 0

```

```

X[X$month == month[i], paste('month', i, sep = '.')] <- 1
}
NUID <- unique(officer_pop_crime$NUID)[order(unique(officer_pop_crime$NUID))]
for (i in 2:length(NUID)){
  X[,paste('NUID', i, sep = '.')] <- 0
  X[X$NUID == NUID[i], paste('NUID', i, sep = '.')] <- 1
}
X <- X[, !(colnames(X) %in% c('unit', 'month', 'NUID'))]
X <- as.matrix(X)
tX_y <- t(X) %*% y
tX_X <- t(X) %*% X
inv_tX_X <- solve(tX_X, tol = 1e-19)
delta <- inv_tX_X %*% tX_y
head(delta, 7)

```

```

##           [,1]
## tenure    -8.639279e-03
## crime     -5.511617e-06
## p50_inc    1.793288e-06
## black     -3.231274e-01
## white     -7.650944e-01
## hisp      -6.686149e-01
## intercept  9.936447e-01

```