Covid 19-India

LiJin Joo lijin.joo at gmail.com

2020-08-23

In this section, I will reproduce the estimates for indian COVID 19 cases between March 19 and May 17 in (Basu et al. 2020). You may find all necessary informatioin at ("COVID19-India Study Group" 2020).

Down laoding the data files raw_data 1-5 from (Group 2020), we can prepare a series of incidence as follows.

```
data1=read.csv("raw_data1.csv", header=T)
data2=read.csv("raw_data2.csv", header=T)
data3=read.csv("raw_data3.csv", header=T)
data4=read.csv("raw_data4.csv", header=T)
data5=read.csv("raw_data5.csv", header=T)

series1<-as.Date(data1[,3], format = "%d/%m/%y")
series2<-as.Date(data2[,3], format = "%d/%m/%y")
series3<-as.Date(data3[,3], format = "%d/%m/%y")
series4<-as.Date(data4[,3], format = "%d/%m/%y")
series5<-as.Date(data5[,3], format = "%d/%m/%y")

all_cases <- c(series1, series2, series3, series4, series5)
cases_covid19_india<-as.numeric(table(all_cases))
dates_covid19_india <- names(table(all_cases))

head(cases_covid19_india, 10)
```

```
##  [1]  1  1  1  2  1 22  2  1  3  5
```

```
head(dates_covid19_india, 10)
```

```
##  [1] "2020-01-30" "2020-02-02" "2020-02-03" "2020-03-02" "2020-03-03"
##  [6] "2020-03-04" "2020-03-05" "2020-03-06" "2020-03-07" "2020-03-08"
```

From their supplment, they explained how they obtained the distribution of serial intervals as follows:

> We use the "parametric_SI" estimation method and a 5-day window ("estimate_R" function, which was used to describe the progression of the outbreak in Wuhan). We also use a gamma distribution prior with a mean of 7 days and a standard deviation of 4.5 days, based on research by Wu and colleagues, for the generation time (a distribution of the onset of disease used to estimate R).

We estimate Rt for Covid19 of India as follows.

```
    t_start <- seq(2, length(cases_covid19_india) - 4)
    t_end   <- t_start + 4

    Rt_covid19 <- EpiEstim::estimate_R(incid = cases_covid19_india, method = "parametric_si",
                      config = make_config(list(mean_si = 7, std_si = 4.5, si_parametric_distr = "G",
                             t_start = t_start, t_end = t_end, seed = 123)))
```

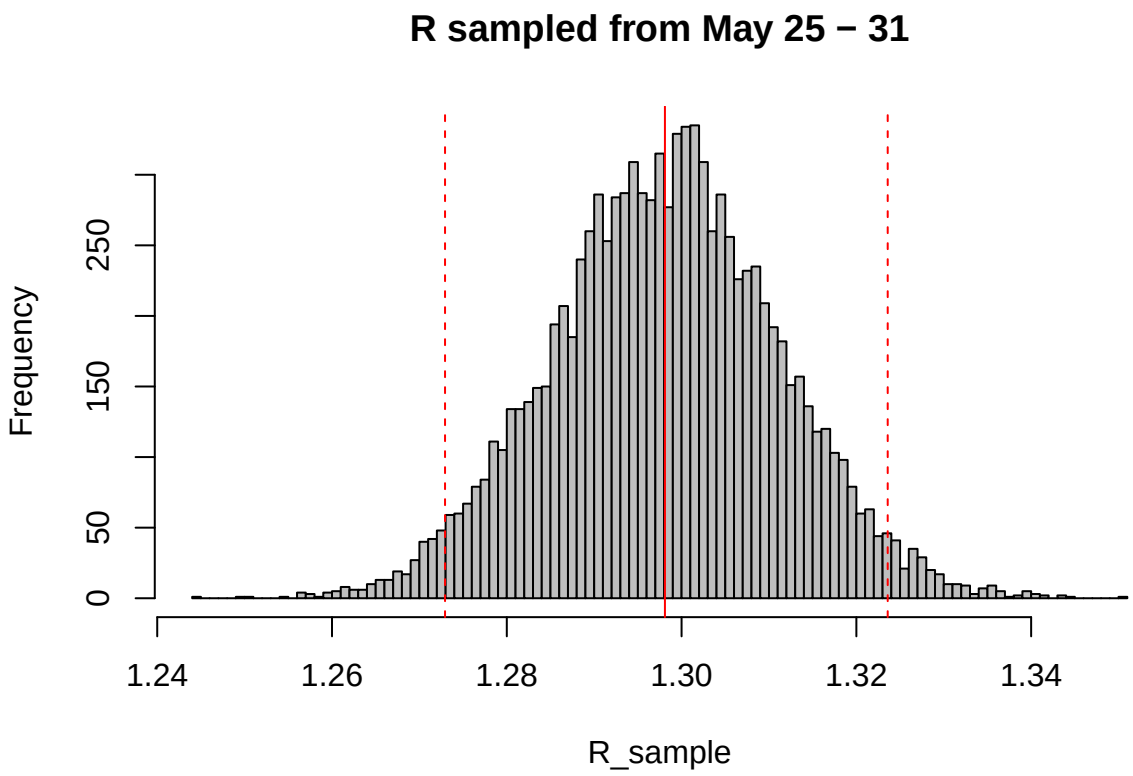
We see that the estimates are very close to the numbers from (Basu et al. 2020) below.

> "The estimated effective reproduction number R for India was 3.36 (95% confidence interval (CI): [3.03, 3.71]) on March 24, whereas the average of estimates from May 25 - May 31 stands at 1.27 (95% CI: [1.26, 1.28])."

```
tibble::tibble(
       date_num = Rt_covid19$dates
   ) %>% dplyr::left_join(
       Rt_covid19$R, by = c("date_num" = "t_end")
   ) %>%
       dplyr::select(
           date_num, t_start, 'Mean(R)', 'Quantile.0.025(R)', 'Quantile.0.975(R)'
       ) %>%
       tibble::add_column(date = dates_covid19_india) %>%
       dplyr::select(-date_num) %>%
       dplyr::select(date, tidyselect::everything()) %>%
       dplyr::slice(c(24:27, 87:94))
```

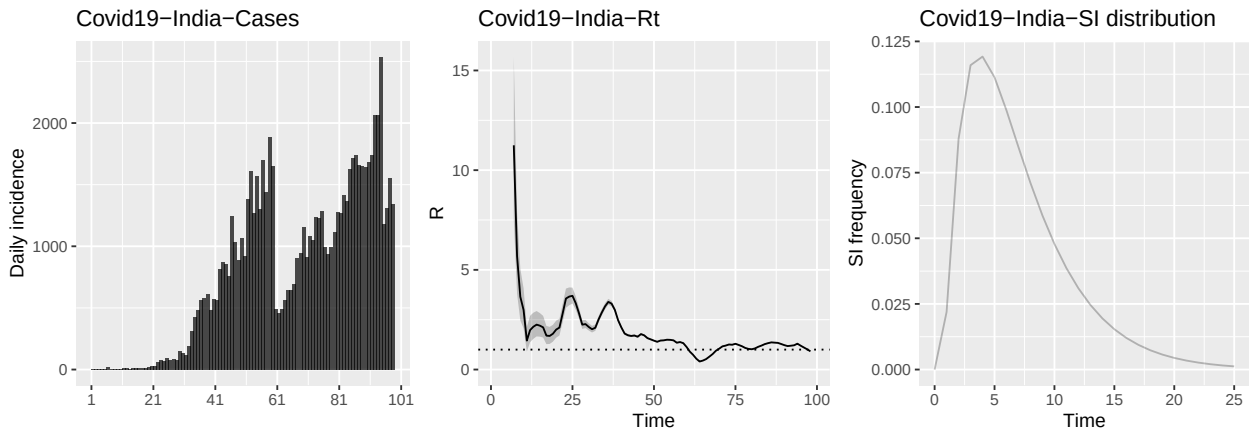
```
## # A tibble: 12 x 5
##    date       t_start `Mean(R)` `Quantile.0.025(R)` `Quantile.0.975(R)`
##    <chr>        <dbl>     <dbl>               <dbl>               <dbl>
##  1 2020-03-22      20      3.66                3.23                4.12
##  2 2020-03-23      21      3.70                3.31                4.11
##  3 2020-03-24      22      3.33                3.00                3.67
##  4 2020-03-25      23      2.82                2.56                3.11
##  5 2020-05-24      83      1.35                1.32                1.38
##  6 2020-05-25      84      1.34                1.31                1.37
##  7 2020-05-26      85      1.28                1.26                1.31
##  8 2020-05-27      86      1.22                1.20                1.25
##  9 2020-05-28      87      1.18                1.15                1.20
## 10 2020-05-29      88      1.20                1.17                1.22
## 11 2020-05-30      89      1.22                1.19                1.24
## 12 2020-05-31      90      1.30                1.27                1.32
```

```
win <- 89
R_median <- Rt_covid19$R$`Median(R)`[win]
R_CrI <- c(Rt_covid19$R$`Quantile.0.025(R)`[win], Rt_covid19$R$`Quantile.0.975(R)`[win])
set.seed(2019)
R_sample <- sample_posterior_R(Rt_covid19, n = 10000, window = win)

hist(R_sample, col = "grey", breaks=100, main = "R sampled from May 25 - 31")
abline(v = R_median, col = "red")
abline(v = R_CrI, col = "red", lty = 2)
```



The three graphs show the trends of COVID19 cases and Rt as well as SI density.

```
library(incidence)
case_covid19 <- plot(as.incidence(cases_covid19_india)) + ggtitle("Covid19-India-Cases")
rt_covid19 <- plot(Rt_covid19, "R") + theme(legend.position = "none") + ggtitle("Covid19-India-Rt")
si_covid19 <- plot(Rt_covid19, "SI") + theme(legend.position = "none") + ylab("SI frequency") + ggtitle

grid.arrange(case_covid19, rt_covid19, si_covid19, nrow = 1)
```



Reference

Basu, Deepankar, Maxwell Salvatore, Debashree Ray, Michael Kleinsasser, Soumik Purkayastha, Rupam Bhattacharyya, and Bhramar Mukherjee. 2020. "A Comprehensive Public Health Evaluation of Lockdown as a Non-Pharmaceutical Intervention on Covid-19 Spread in India: National Trends Masking State Level Variations." *medRxiv*. Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/2020.05.25.20113043.

"COVID19-India Study Group." 2020. https://umich-biostatistics.shinyapps.io/covid19/.

Group, COVID-19 India Org Data Operations. 2020. “COVID19-India Api.” https://api.covid19india.org/.