Statistical Estimation of the Reproductive Number From Case Notification Data

Laura F. White, Carlee B. Moser, Robin N. Thompson, and Marcello Pagano

Contents

WEB APPENDIX

R Code for Conducting Example Analyses

```r
# Analysis of sample epidemic data #
#####################################
install.packages("EpiEstim")
install.packages("R0")
library("EpiEstim")
library("incidence")
library("R0")
library("ggplot2")
library("reshape2")

# data from SARS that is published and Influenza in LaGloria
setwd("<location of data>")

lg <- read.csv("LaGloria.csv")
names(lg) <- c("dayNum","N")
sars.hk <- read.table("SARS\\hongkong.txt")
names(sars.hk) <- c("dayNum","N")
sars.hk$dayNum <- sars.hk$dayNum-min(sars.hk$dayNum)+1

# make sure there are 0's on days with no data
sars.hk2 <- cbind(c(rep(0,max(sars.hk$dayNum))),c(rep(0,max(sars.hk$dayNum))))
sars.hk2[,1] <- c(1:max(sars.hk$dayNum))
sars.hk2[sars.hk$dayNum,2] <- sars.hk$N
sars.hk2 <- as.data.frame(sars.hk2)
names(sars.hk2) <- c("dayNum","N")

# SI's for SARS and Influenza
# from Lipsitch et al (2003)
sars.si <- generation.time("weibull",c(8.4,3.5))
# from Cowling et al (2009)
flu.si <- generation.time("gamma",c(3.6,1.6))

# calculation of basic reproductive number, R0 #
################################################
# calculate R0 using White and Pagano #
########################################
# SARS outbreak #
##################
R0.ML.sars <- est.R0.ML(sars.hk$N,t=sars.hk$dayNum,begin=7,end=39,GT=sars.si)
```

```
# get estimates of R0 using varying amounts of data #
R0.ML.sars.ests <- NULL
for(j in 12:39){
  R0.ML.sars.ests[j] <-
est.R0.ML(sars.hk$N,t=sars.hk$dayNum,begin=7,end=as.numeric(j),GT=sars.si)$R
}

# Influenza 2009 data #
#########################
R0.ML.flu <- est.R0.ML(lg$N,t=lg$dayNum,begin=1,end=15,GT=flu.si)

# using varying amounts of data #
R0.ML.flu.ests <- NULL
for(i in 7:15){
  R0.ML.flu.ests[i] <- est.R0.ML(lg$N,t=lg$dayNum,begin=1,end=as.numeric(i),GT=flu.si)$R
}

# calculate R0 using sequential bayesian approach:
# start on day 7, once outbreak is established
# SARS #
R0.SB.sars <- est.R0.SB(sars.hk2$N[7:60],GT=sars.si)

# influenza #
R0.SB.flu <- est.R0.SB(lg$N,GT=flu.si)

# Time-varying reproductive numbers #
####################################
# case reprodcutive number #
# Wallinga and Teunis estimator (2004) #
# SARS #
WT.Rt.sars <- est.R0.TD(sars.hk$N,t=sars.hk$dayNum,GT=sars.si,begin=1,end=96,nsim=1000)

# influenza #
WT.Rt.flu <- est.R0.TD(lg$N,t=lg$dayNum,GT=flu.si,nsim=1000,begin=1,end=34)

# Instantaneous reproductive number #
# no smoothing #
# SARS #
Inst.Rt.sars <- estimate_R(sars.hk2$N,method="parametric_si",
                   config = make_config(list(mean_si = sars.si$mean,
                     std_si = sars.si$sd,
                     t_start=c(3:length(sars.hk2$N)-1),
                     t_end=c(3:length(sars.hk2$N)-1)+1))
)

# influenza #
```

```
Inst.Rt.flu <- estimate_R(lg$N,method="parametric_si",
                config = make_config(list(
                  mean_si = flu.si$mean,
                  std_si = flu.si$sd,
                  t_start=c(3:length(lg$N)-1),
                  t_end=c(3:length(lg$N)-1)+1))
)


#############################
## Plot all results     ##
#############################
# Plot R0 estimates over the epidemic growth period #
###########################################################
# create dataframe with dayNum, N, R0est, Method for each day of epidemic period
flu.R0.results <- as.data.frame(cbind(lg[1:15,],c(NA,R0.SB.flu$R),R0.ML.flu.ests))
names(flu.R0.results) <- c("dayNum","N","R0.SB","R0.ML")
tmp <- melt(flu.R0.results,id="dayNum")
flu.R0.long <- dcast(tmp,dayNum~variable)
flu.R0.long <-
as.data.frame(rbind(as.matrix(flu.R0.long[,1:3]),as.matrix(flu.R0.long[,c(1:2,4)])))
names(flu.R0.long) <- c("dayNum","N","R0")
flu.R0.long$Method <- c(rep("SB",15),rep("ML",15))

# separately for SARS results
sars.R0.results <-
as.data.frame(cbind(sars.hk2[1:39,],c(rep(NA,7),R0.SB.sars$R),R0.ML.sars.ests))
names(sars.R0.results) <- c("dayNum","N","R0.SB","R0.ML")
tmp <- melt(sars.R0.results,id="dayNum")
sars.R0.long <- dcast(tmp,dayNum~variable)
sars.R0.long <-
as.data.frame(rbind(as.matrix(sars.R0.long[,1:3]),as.matrix(sars.R0.long[,c(1:2,4)])))
names(sars.R0.long) <- c("dayNum","N","R0")
sars.R0.long$Method <- c(rep("SB",39),rep("ML",39))


# plot R0 results with incidence data through epidemic period #
#################################################################
### Plot the results ###
flu.plot <- ggplot(flu.R0.long,aes(x=dayNum,y=N/2))+
  geom_bar(stat="identity")+labs(y="Number of cases",x="Day of outbreak",title="(a)")+
  geom_line(aes(y=R0*8,linetype=Method),size=1)+
  scale_y_continuous(sec.axis=sec_axis(trans=~./8,name=expression(hat(R)[0])))+
  geom_hline(yintercept=8,color="gray")
flu.plot
```

```
# SARS
sars.plot <- ggplot(sars.R0.long,aes(x=dayNum,y=N/2))+
  geom_bar(stat="identity")+labs(y="Number of cases",x="Day of outbreak",title="(b)")+
  geom_line(aes(y=R0*8,linetype=Method),size=1)+
  scale_y_continuous(sec.axis=sec_axis(trans=~./8,name=expression(hat(R)[0])))+
  geom_hline(yintercept=8,color="gray")
sars.plot


# Plot of incidence curves  with time-varying Rt estimates #
################################################################
# set up data for ggplot
flu.results <- as.data.frame(cbind(lg,WT.Rt.flu$R,
                    c(NA,Inst.Rt.flu$R[,3],NA)))
names(flu.results) <- c("dayNum","N","WT.Rt","Inst.Rt")
tmp <- melt(flu.results,id="dayNum")
flu.long <- dcast(tmp,dayNum~variable)
flu.long <- as.data.frame(rbind(as.matrix(flu.long[,1:3]),as.matrix(flu.long[,c(1:2,4)])))
names(flu.long) <- c("dayNum","N","Rt")
flu.long$Method <- c(rep("WT",34),rep("Inst",34))

sars.results <- as.data.frame(cbind(sars.hk2,WT.Rt.sars$R,
                    c(NA,Inst.Rt.sars$R[,3],NA)))
names(sars.results) <- c("dayNum","N","WT.Rt","Inst.Rt")
tmp <- melt(sars.results,id="dayNum")
sars.long <- dcast(tmp,dayNum~variable)
sars.long <- as.data.frame(rbind(as.matrix(sars.long[,1:3]),as.matrix(sars.long[,c(1:2,4)])))
names(sars.long) <- c("dayNum","N","Rt")
sars.long$Method <- c(rep("WT",96),rep("Inst",96))

### Plot the results ###
flu.plot <- ggplot(flu.long,aes(x=dayNum,y=N/2))+
  geom_bar(stat="identity")+labs(y="Number of cases",x="Day of outbreak",title="(a)")+
  geom_line(aes(y=Rt*8,linetype=Method),size=1)+
  scale_y_continuous(sec.axis=sec_axis(trans=~./8,name=expression(hat(R)[t])))+
  geom_hline(yintercept=8,color="gray")
flu.plot

sars.plot <- ggplot(sars.long,aes(x=dayNum,y=N/2))+
  geom_bar(stat="identity")+labs(y="Number of cases",x="Day of outbreak",title="(b)")+
  geom_line(aes(y=Rt*8,linetype=Method),size=1)+
  scale_y_continuous(sec.axis=sec_axis(trans=~./8,name=expression(hat(R)[t])))+
  geom_hline(yintercept=8,color="gray")
sars.plot
```