Will the COVID-19 cause the excess death for human diseases?

Evidence based on four diseases data in Ontario: Malignant neoplasms, disease of the heart, accidents, and chronic lower respiratory diseases.

Introduction

2020 Covid-19 pandemic has changed human's life in so many ways. In this project, we want to investigate that was there any excess mortality caused by the Covid-19 from either of the four types of disease, which is malignant neoplasms, disease of the heart, accidents, and chronic lower respiratory diseases respectively. To find out the excess mortality caused by Covid-19, the method we use is comparing each disease total mortality data under the forecasting NO Covid circumstance and the reality circumstance (under the Covid). The forecasting NO Covid circumstance is the model with 100 observations that formed by fitting the pre-Covid dataset. The dataset is the mortality data for each of the four types of disease in Ontario Canada from March to November 2020, which can be accessed from Statistics Canada *https://open.canada.ca/data/en/dataset/aed00edc-26ad-414c-8aa3-82212059ef8a.*

Method & Model

To investigate the impact of COVID-19 on mortality rate of four diseases based on the dataset of weekly mortality data of each disease in Ontario, Canada. We assume that Y, the expected number of death people caused by each of the four diseases every week, follows the poisson distribution, since the variable Y is positive and discrete for both. Thus, the mean and variance of Y are both λ_i since Y follows poisson distribution.

$$Yi \sim Poisson(\lambda_i)$$

Based on the non-linear weekly pattern of mortality caused by each of four diseases (Figure2.1 to Figure2.4), we use the semi-parametric model to fit our data:

$$log\,(\lambda_i) = X_i\beta + U_{t_i} + V_i$$

λ_i refers to expected number of people died caused by either of the four types of disease; $\log\lambda_i$ is the corresponding log mortality. $X_i\beta$ is the fixed effect. X_i denotes the covariates defined by the four trigonometric functions for weekly pattern; β are their corresponding parameters; U_{t_i} is the random walk of order 2 such as the random effect smoothing with time for smoothing purpose with variance σ^2; V_i is the random effect of time.

For U_{t_i}, we apply the pc.prec for precision

$$U_{t_i} \sim N(0, \sigma_U^2)$$

$$Prob\ (\sigma_u > 0.001) = 0.5$$

In general, we have the following distribution

$$[U_1 \ldots U_T]^T \sim RW2(0, \sigma_U^2)$$

RW(2), Random slope:

$$U_{t+1} | U_k < t \sim N(-2U_t + U_{t-1}, \sigma_U^2)$$

$$(U_{t+1} - U_t) - (U_t - U_{t-1}) \sim N(0, \sigma_U^2)$$

$$U_{t+1} - 2U_t + U_{t-1} \sim N(0, \sigma_U^2)$$

For V_i the random effect we apply the pc.prec for precision set

$$V_i \sim N(0, \sigma_V^2)$$

$$Prob\ (\sigma_V > 0.0001) = 0.5$$

X_i denotes the covariates defined by the four trigonometric functions, cos12, sin12, cos6 and sin6, for weekly pattern with their own formulas:

*cos12 = cos(2*π*dataInt/365.25)*

*sin12 = sin(2*π* dataInt/365.25)*

*cos6 = cos(2*2*π* dataInt/365.25)*

*sin6 = sin(2*2*π* dataInt/365.25)*

The dataInt is the numerical representation of time point by normalizing the time variable with one standard deviation equals to 365.25 days.

Result

Figure2.1 Mortality of heart disease

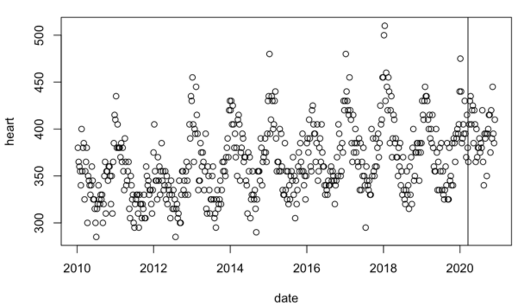


Figure2.2 Mortality of neoplasms

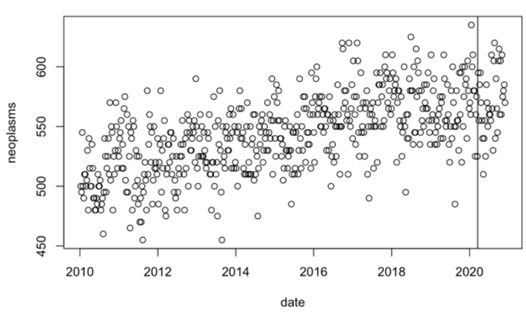


Figure2.3 Mortality of respiratory disease

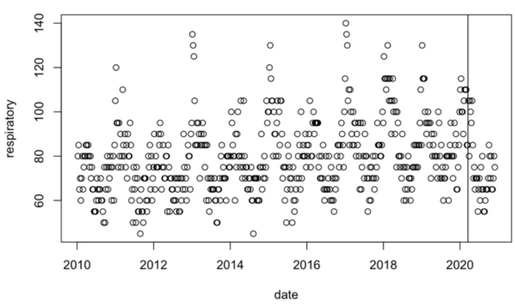
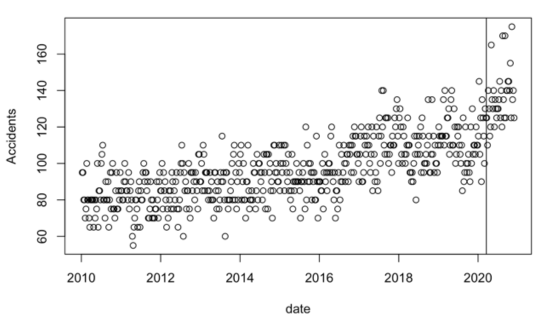


Figure2.4 Mortality of accidents



These four figures show the distribution of weekly number of death people caused by heart disease (Figure2.1), malignant neoplasms (Figure2.2), respiratory disease (Figure2.3), and accidents (Figure2.4) respectively from 2010/1/9 to 2020/11/28 in Ontario. We marked the special time point 2020/03/17 for four diseases since it is the start time point for the COVID-19 pandemic. All the data after this time point may get the effects of COVID-19.

To investigate how will the COVID -19 affects the mortality of those four diseases, we can focus on the pattern after the time point 2020/03/17. Based on those plots, we can only assume that after the explosion of COVID-19, there is a distinct increase mortality of accident (Figure2.1) and a slightly decrease mortality of respiratory disease (Figure2.3). To the heart disease and neoplasms, there is not enough evidence to tell how the mortality different after the COVID shown in the plot.

The dataset we use has some special case that the data after February 1st, 2020, is the explosion of COVID-19, and each year the data from 21 December to 12 January is not follow the normal pattern since the health system will be disrupted by Christmas holiday period. Therefore, we mark those special cases data as red in the plot and blue point are the normal case. Figure 2.5 to Figure2.8 is the distribution of four diseases mortality after fitting the GAM models

Figure2.5 Model of heart disease mortality

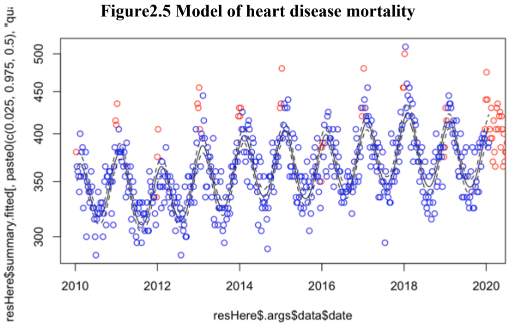


Figure2.6 Model of neoplasms mortality

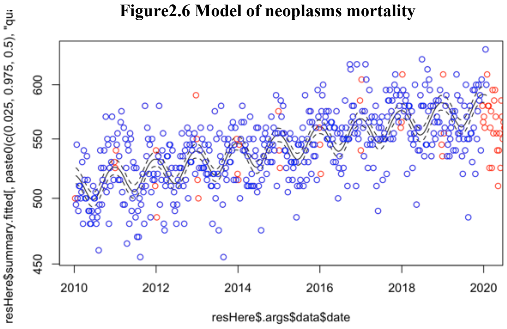


Figure2.7 Model of respiratory disease mortality

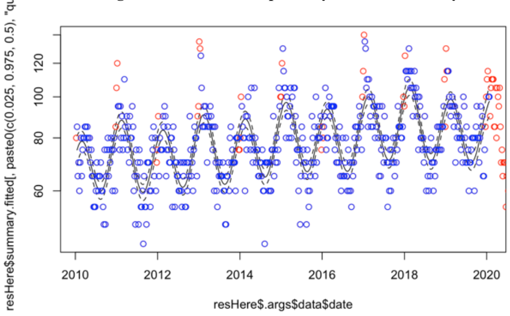


Figure2.8 Model of accident mortality

To find out the excess mortality caused by Covid-19, the method we use is comparing each disease total mortality data under the forecasting NO-Covid circumstance and the reality circumstance (under the Covid). The forecasting NO Covid circumstance is the model with 100 observations that formed by fitting the pre-Covid dataset and marked as the red dot. The black pattern is the distribution of reality mortality. We will find out whether the excess mortality by the difference between the NO-Covid circumstance (red dot) and Covid circumstance (black pattern) after the timepoint February 2020.

Figure2.9 NO-Covid VS Covid for heart disease

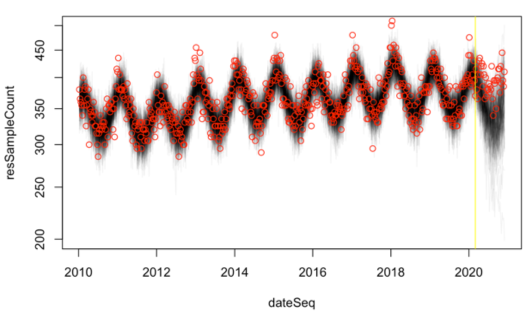


Figure2.10 NO-Covid VS Covid for neoplasms

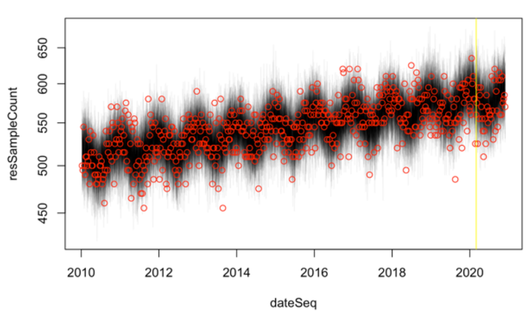


Figure2.11 NO-Covid VS Covid for respiratory disease

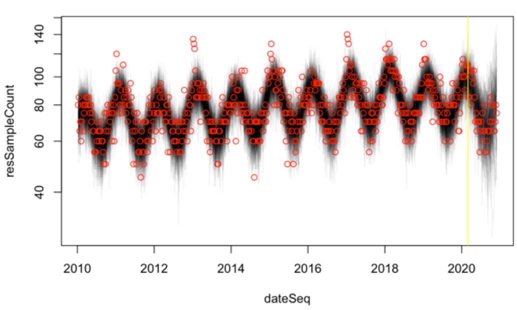
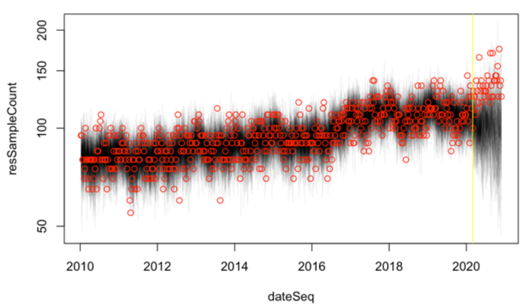


Figure2.12 NO-Covid VS Covid for accident



The Figure2.13 to Figure2.16 will zoom in the time period after February 2020 and show the excess mortality caused by COVID-19 with the y-axis of excess mortality. The blue pattern shows the differences of four diseases mortality between NO-COVID and COVID circumstance over time in 2020. Each figure has the baseline of zero which means no excess death caused by pandemic, the blue pattern above the baseline zero represents the excess mortality of the disease caused by the COVID-19, the pandemic make the disease worse; the blue pattern under the baseline zero represents less mortality of the disease caused by the COVID-19, the pandemic make the disease better than before.

Figure2.13 shows the mortality difference of heart disease over 2020 from Feburary to December. It is obvious to see that the COVID-19 caused large excess mortality of heart disease compared with No-COVID circumstance from April to December 2020. More precisely, the Table 2.1 shows the quantile of heart disease excess mortality.

We firstly focus on the value of 25 quantile and 75 quantile, it means that for all dataset of heart disease excess mortality, 25 percent of the values fall below than 2028.25 and 25 percent of the values larger than 5804.00, and the median of the excess mortality is 4098.00. Therefore, since almost 75% scale of data is positive excess mortality, we can say that **the COVID-19 pandemic did make people with heart disease get worse and more serious, because much more people died of heart disease under the pandemic compared with NO-pandemic, especially from April to November 2020.**

Figure2.13 NO-Covid VS Covid for heart disease mortality in 2020

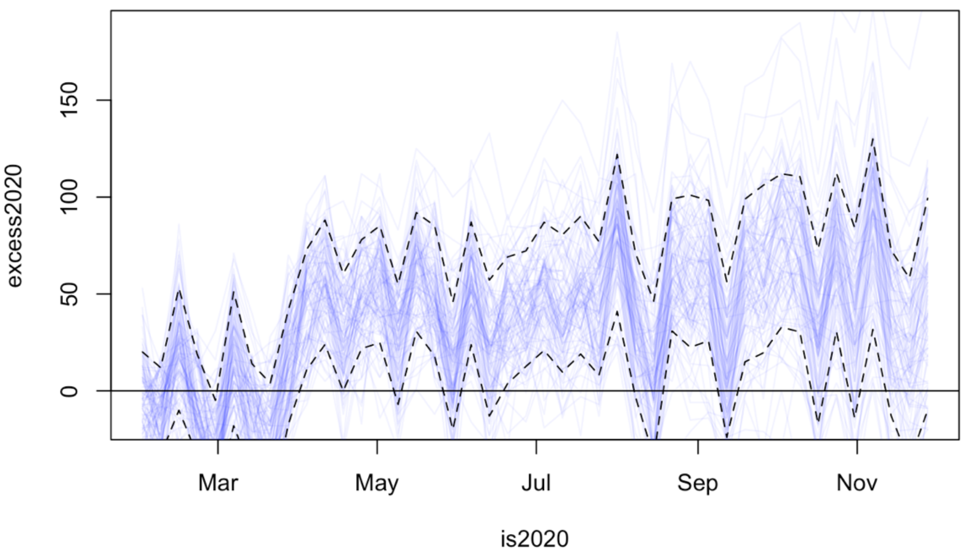


Table 2.1 Quantile of heart disease excess mortality

Quantile	0%	25%	50%	75%	100%
excess mortality 2020	−3139.00	2028.25	4098.00	5804.00	8703.00

Figure2.14 shows the mortality difference of neoplasms in 2020. It is diffcult to tell that the COVID-19 make any effects to neoplasms mortality from Feburary to December 2020 since all the excess mortality is around the baseline zero. More precisely, the Table 2.2 shows the quantile of neoplasms excess mortality.

We firstly focus on the value of 25 quantile and 75 quantile, it means that for all dataset of neoplasms excess mortality, 25 percent of the values fall below than -2551.50 and 25 percent of the values larger than 1021.25, and the median of the excess mortality is -1227. Therefore, since the similar scale of the positive and negative excess mortality, we can say that **the COVID-19 pandemic will not affect people with neoplasms a lot**, since the **growth** in the number of deaths of neoplasms in a panemic circumstance is almost equal with the **reduction** in the number of deaths of neoplasms in a NO-pandemic circumstance from Feburary to December 2020. However, May 2020 is a special case of neoplasma, there are less people will die by neoplasms under the COVID-19 circumstance compared with NO-Covid in May 2020.

Figure2.14 NO-Covid VS Covid for neoplasms mortality in 2020

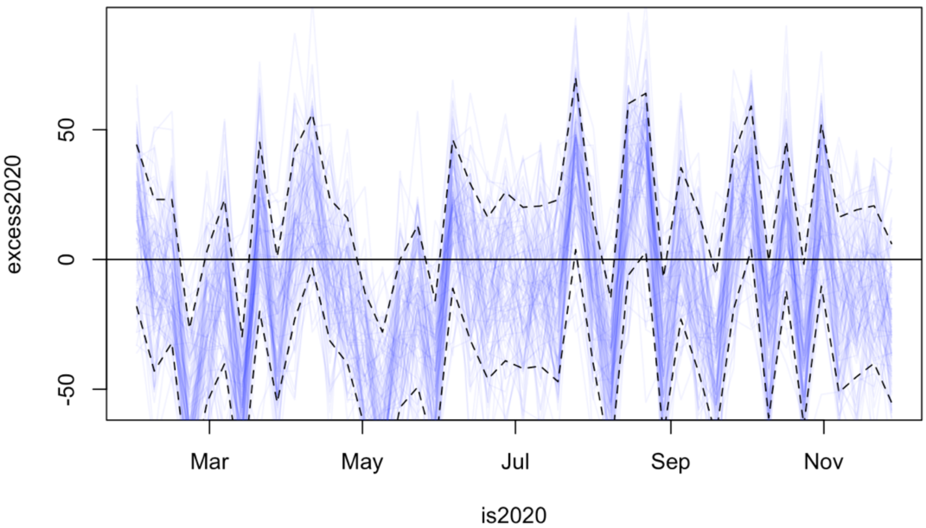


Table 2.2 Quantile of neoplasms excess mortality

Quantile	0%	25%	50%	75%	100%
excess mortality 2020	−6031.00	−2551.50	−1227.00	1021.25	4299.00

Figure2.15 shows the mortality difference of respiratory disease under the pandemic and NO-pandemic circumstance in 2020. It is diffcult to tell that the COVID-19 make any effects to respiratory disease mortality from Feburary to December 2020 since all the excess mortality is around the baseline zero with the small range from 33 to -30. More precisely, the Table 2.3 shows the quantile of neoplasms excess mortality.

We firstly focus on the value of 25 quantile and 75 quantile, it means that for all dataset of neoplasms excess mortality, 25 percent of the values fall below than -949.50 and 25 percent of the values larger than 564.75, and the median of the excess mortality is 450.5. Therefore, we can say that **the COVID-19 pandemic will not affect people with respiratory disease a lot**, since the small **growth** in the number of deaths of respiratory disease in a pandemic circumstance can offset with the small reduction in the number of deaths of respiratory disease in a NO-pandemic circumstance from Feburary to December in 2020.

Figure2.15 NO-Covid VS Covid for respiratory disease mortality in 2020

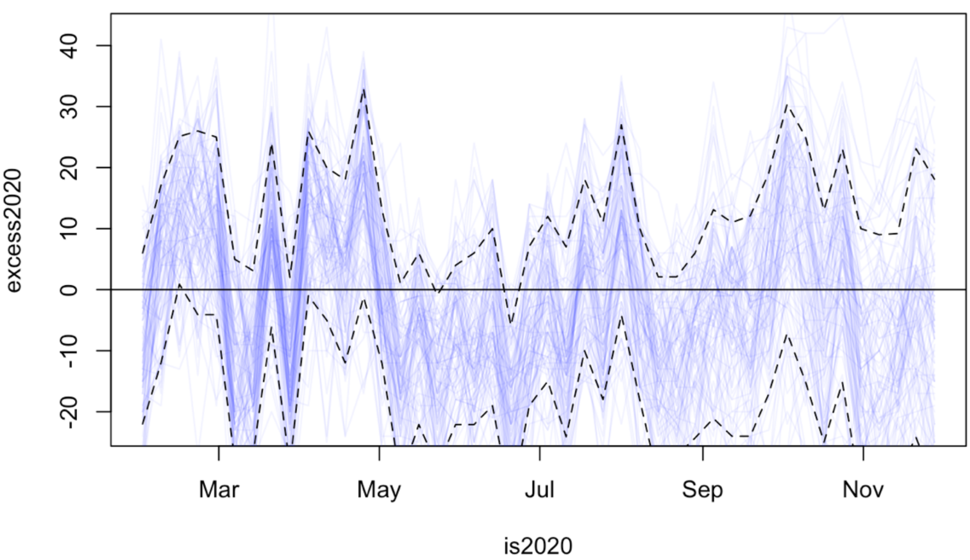


Table 2.3 Quantile of respiratory disease excess mortality

Quantile	0%	25%	50%	75%	100%
excess mortality 2020	−1890.00	−949.50	450.50	564.75	4299.00

Figure2.16 shows the mortality difference of accident over 2020 from Feburary to December. It is obvious to see that the COVID-19 caused many excess mortality of heart disease from April to December 2020, expecially October and November. More precisely, the Table 2.4 shows the quantile of accident excess mortality.
We firstly focus on the value of 25 quantile and 75 quantile, it means that for all dataset of heart disease excess mortality, 25 percent of the values fall below than 1910.75 and 25 percent of the values larger than 3472, and the median of the excess mortality is 2598. Therefore, since almost 75% scale of data is positive excess mortality, we can say that **the COVID-19 pandemic did make the accidents get more and more serious, because much more people died of accident under the pandemic compared with NO-pandemic from Feburary to December,** especially May and August in 2020**.**

Figure2.16 NO-Covid VS Covid for accident mortality in 2020

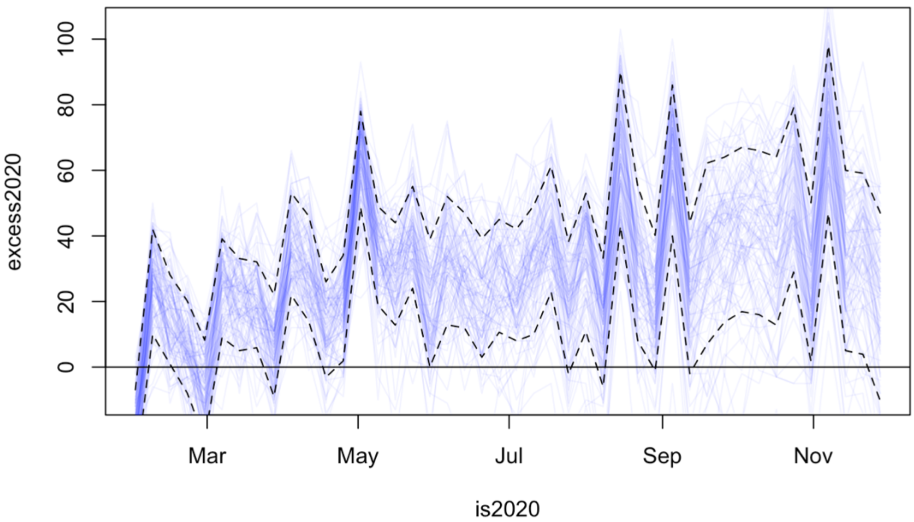


Table 2.4 Quantile of accident excess mortality

Quantile	0%	25%	50%	75%	100%
excess mortality 2020	−2053.00	1910.75	2598.00	3472.00	6950.00

Summary

To investigate that will the Covid-19 cause any excess mortality from either of the four types of diseases, which is malignant neoplasms, disease of the heart, accidents, and chronic lower respiratory diseases respectively, the method we use is comparing each disease mortality data under the forecasting NO-Covid circumstance and the reality circumstance (under the Covid). The first result we get is that the COVID-19 pandemic will cause much more people die by heart disease compared with No-COVID circumstance, especially from April to November 2020. Secondly, a little bit more people will die in accident in pandemic compared with NO-COVID circumstance from February to December in 2020. Thirdly, for the neoplasms, the COVID-19 pandemic will not affect people with neoplasms a lot expect May, there are less people will die by neoplasms under the COVID-19 circumstance compared with NO-Covid in May 2020. Lastly, for the respiratory disease, the COVID-19 pandemic will not affect people with respiratory disease a lot, since the small **growth** in the number of deaths of respiratory disease in a pandemic circumstance can offset with the small reduction in the number of deaths of respiratory disease in a NO-pandemic circumstance from Feburary to December in 2020.

Appendix

```r
#install.packages("R.utils")
library(R.utils, quietly = TRUE)

#install.packages("Pmisc", repos="http://R-Forge.R-project.org")
library("Pmisc", quietly = TRUE)

deadFile = Pmisc::downloadIfOld("https://www150.statcan.gc.ca/n1/tbl
/csv/13100810-eng.zip")
(deadFileCsv = deadFile[which.max(file.info(deadFile)$size)])

x = read.csv(deadFileCsv)
x[1:2, ]

x$date = as.Date(as.character(x[[grep("DATE", names(x))]]))
x$province = gsub("[,].*", "", x$GEO)
# remove 2021 data, which appears incomplete
x = x[x$date < as.Date("2020/12/01") & x$province =="Ontario", ]
for (D in c("heart", "neoplasms", "respiratory", "Accidents")) {
  plot(x[grep(D, x$Cause), c("date", "VALUE")], ylab = D)
  abline(v = as.Date("2020/03/17")) }
```

#heart

```r
dateSeq = sort(unique(x$date))
table(diff(dateSeq))

dateSeqInt = as.integer(dateSeq)
x$dateInt = x$dateIid = as.integer(x$date)
x$cos12 = cos(2 * pi * x$dateInt/365.25)
```

```
x$sin12 = sin(2 * pi * x$dateInt/365.25)
x$sin6 = sin(2 * 2 * pi * x$dateInt/365.25)
x$cos6 = cos(2 * 2 * pi * x$dateInt/365.25)
x$dayOfYear = as.Date(gsub("^[[:digit:]]+", "0000",x$date))
x$christmasBreak = (x$dayOfYear >= as.Date("0000/12/21")) |(x$dayOfYe
ar <= as.Date("0000/01/12"))
xSub = x[grepl("heart", x$Cause, ignore.case = TRUE) & x$province ==
"Ontario", ]
xPreCovid = xSub[xSub$date < as.Date("2020/02/01") & (!xSub$christmas
Break), ]

library("INLA")

resHere = inla(VALUE ~ cos12 + cos6 + sin12 + sin6 +
                f(dateInt, model = "rw2", values = dateSeqInt,prior =
 "pc.prec", param = c(0.001, 0.5)) +
                f(dateIid, values = dateSeqInt, prior = "pc.prec", pa
ram = c(0.0001, 0.5)),
              data = xPreCovid, family = "poisson",
              control.compute = list(config = TRUE), control.predicto
r = list(compute = TRUE))

matplot(resHere$.args$data$date, resHere$summary.fitted[, paste0(c
(0.025, 0.975, 0.5), "quant")], type = "l",lty = c(2, 2, 1), col = "b
lack", log = "y", ylim = range(xSub$VALUE))
points(xSub$date, xSub$VALUE, col = "red")
points(xPreCovid$date, xPreCovid$VALUE, col = "blue")
```

```
matplot(dateSeq, resHere$summary.random$dateInt[, paste0(c(0.025,0.9
75, 0.5), "quant")], type = "l", lty = c(2,2, 1), col = "black")
```

```
toPredict = cbind(`(Intercept):1` = 1, `cos12:1` = cos(2 *pi * dateSe
qInt/365.25), `sin12:1` = sin(2 * pi * dateSeqInt/365.25), `cos6:1` =
 cos(2 * pi * dateSeqInt *2/365.25),`sin6:1` = sin(2 * pi * dateSeqIn
t *2/365.25))
dateIntSeq = paste0("dateInt:", 1:length(dateSeqInt))
dateIidSeq = paste0("dateIid:", 1:length(dateSeqInt))
set.seed(123456)
resSample = inla.posterior.sample(n = 100, resHere)
resSampleFitted = lapply(resSample, function(xx) {toPredict %*% xx$la
```

```
tent[colnames(toPredict), ] + xx$latent[dateIntSeq, ] + xx$latent[dat
eIidSeq,]})
resSampleFitted = do.call(cbind, resSampleFitted)
resSampleLambda = exp(resSampleFitted)
resSampleCount = matrix(rpois(length(resSampleLambda),
                            resSampleLambda), nrow(resSampleLambda),
ncol(resSampleLambda))
matplot(dateSeq, resSampleCount, col ="#00000010",type = "l", lty =
1, log = "y")

points(xSub[, c("date", "VALUE")], col = "red")
abline(v = as.Date("2020/03/01"), col = "yellow")
```

```
is2020 = dateSeq[dateSeq >= as.Date("2020/2/1")]
sample2020 = resSampleCount[match(is2020, dateSeq),]
count2020 = xSub[match(is2020, xSub$date), "VALUE"]
excess2020 = count2020 - sample2020
matplot(is2020, excess2020, type = "l", lty = 1, col = "#0000FF10",
       ylim = range(-10, quantile(excess2020, c(0.1, 0.999))))
matlines(is2020, t(apply(excess2020, 1, quantile, prob = c(0.1,0.
9))), col = "black", lty = 2)
abline(h = 0)
```

```
quantile(apply(excess2020, 1, sum))

##       0%      25%      50%      75%     100%
## -3139.00  2028.25  4098.00  5804.00  8703.00
```

#neoplasms

```
dateSeq = sort(unique(x$date))
table(diff(dateSeq))

dateSeqInt = as.integer(dateSeq)
x$dateInt = x$dateIid = as.integer(x$date)
x$cos12 = cos(2 * pi * x$dateInt/365.25)
x$sin12 = sin(2 * pi * x$dateInt/365.25)
x$sin6 = sin(2 * 2 * pi * x$dateInt/365.25)
x$cos6 = cos(2 * 2 * pi * x$dateInt/365.25)
x$dayOfYear = as.Date(gsub("^[[:digit:]]+", "0000",x$date))
```

```r
x$christmasBreak = (x$dayOfYear >= as.Date("0000/12/21")) |(x$dayOfYear <= as.Date("0000/01/12"))
xSub = x[grepl("neoplasms", x$Cause, ignore.case = TRUE) & x$province == "Ontario", ]
xPreCovid = xSub[xSub$date < as.Date("2020/02/01") & (!xSub$christmasBreak), ]

#library("INLA")
resHere = inla(VALUE ~ cos12 + cos6 + sin12 + sin6 +
                f(dateInt, model = "rw2", values = dateSeqInt,prior = "pc.prec", param = c(0.001, 0.5)) +
                f(dateIid, values = dateSeqInt, prior = "pc.prec", param = c(0.0001, 0.5)),
              data = xPreCovid, family = "poisson",
              control.compute = list(config = TRUE), control.predictor = list(compute = TRUE))

matplot(resHere$.args$data$date, resHere$summary.fitted[, paste0(c(0.025, 0.975, 0.5), "quant")], type = "l",lty = c(2, 2, 1), col = "black", log = "y", ylim = range(xSub$VALUE))
points(xSub$date, xSub$VALUE, col = "red")
points(xPreCovid$date, xPreCovid$VALUE, col = "blue")
```

```r
matplot(dateSeq, resHere$summary.random$dateInt[, paste0(c(0.025,0.975, 0.5), "quant")], type = "l", lty = c(2,2, 1), col = "black")
```

```r
toPredict = cbind(`(Intercept):1` = 1, `cos12:1` = cos(2 *pi * dateSeqInt/365.25), `sin12:1` = sin(2 * pi * dateSeqInt/365.25), `cos6:1` = cos(2 * pi * dateSeqInt *2/365.25),`sin6:1` = sin(2 * pi * dateSeqInt *2/365.25))
dateIntSeq = paste0("dateInt:", 1:length(dateSeqInt))
dateIidSeq = paste0("dateIid:", 1:length(dateSeqInt))
set.seed(123456)
resSample = inla.posterior.sample(n = 100, resHere)
resSampleFitted = lapply(resSample, function(xx) {toPredict %*% xx$latent[colnames(toPredict), ] + xx$latent[dateIntSeq, ] + xx$latent[dateIidSeq,]})
resSampleFitted = do.call(cbind, resSampleFitted)
resSampleLambda = exp(resSampleFitted)
resSampleCount = matrix(rpois(length(resSampleLambda),
```

```
                                    resSampleLambda), nrow(resSampleLambda),
ncol(resSampleLambda))
matplot(dateSeq, resSampleCount, col ="#00000010",type = "l", lty =
1, log = "y")

points(xSub[, c("date", "VALUE")], col = "red")
abline(v = as.Date("2020/03/01"), col = "yellow")
```

```
is2020 = dateSeq[dateSeq >= as.Date("2020/2/1")]
sample2020 = resSampleCount[match(is2020, dateSeq),]
count2020 = xSub[match(is2020, xSub$date), "VALUE"]
excess2020 = count2020 - sample2020
matplot(is2020, excess2020, type = "l", lty = 1, col = "#0000FF10",
        ylim = range(-10, quantile(excess2020, c(0.1, 0.999))))
matlines(is2020, t(apply(excess2020, 1, quantile, prob = c(0.1,0.
9))), col = "black", lty = 2)
abline(h = 0)
```

```
quantile(apply(excess2020, 1, sum))

##       0%      25%      50%      75%     100%
## -6031.00 -2551.50 -1227.00  1021.25  4399.00
```

#respiratory

```
dateSeq = sort(unique(x$date))
table(diff(dateSeq))

## 568

dateSeqInt = as.integer(dateSeq)
x$dateInt = x$dateIid = as.integer(x$date)
x$cos12 = cos(2 * pi * x$dateInt/365.25)
x$sin12 = sin(2 * pi * x$dateInt/365.25)
x$sin6 = sin(2 * 2 * pi * x$dateInt/365.25)
x$cos6 = cos(2 * 2 * pi * x$dateInt/365.25)
x$dayOfYear = as.Date(gsub("^[[:digit:]]+", "0000",x$date))
x$christmasBreak = (x$dayOfYear >= as.Date("0000/12/21")) |(x$dayOfYe
ar <= as.Date("0000/01/12"))
xSub = x[grepl("respiratory", x$Cause, ignore.case = TRUE) & x$provin
ce == "Ontario", ]
```

```r
xPreCovid = xSub[xSub$date < as.Date("2020/02/01") & (!xSub$christmas
Break), ]

#library("INLA")
resHere = inla(VALUE ~ cos12 + cos6 + sin12 + sin6 +
                f(dateInt, model = "rw2", values = dateSeqInt,prior =
 "pc.prec", param = c(0.001, 0.5)) +
                f(dateIid, values = dateSeqInt, prior = "pc.prec", pa
ram = c(0.0001, 0.5)),
              data = xPreCovid, family = "poisson",
              control.compute = list(config = TRUE), control.predicto
r = list(compute = TRUE))
```

```r
matplot(resHere$.args$data$date, resHere$summary.fitted[, paste0(c
(0.025, 0.975, 0.5), "quant")], type = "l",lty = c(2, 2, 1), col = "b
lack", log = "y", ylim = range(xSub$VALUE))
points(xSub$date, xSub$VALUE, col = "red")
points(xPreCovid$date, xPreCovid$VALUE, col = "blue")
```

```r
matplot(dateSeq, resHere$summary.random$dateInt[, paste0(c(0.025,0.9
75, 0.5), "quant")], type = "l", lty = c(2,2, 1), col = "black")
```

```r
toPredict = cbind(`(Intercept):1` = 1, `cos12:1` = cos(2 *pi * dateSe
qInt/365.25), `sin12:1` = sin(2 * pi * dateSeqInt/365.25), `cos6:1` =
 cos(2 * pi * dateSeqInt *2/365.25),`sin6:1` = sin(2 * pi * dateSeqIn
t *2/365.25))
dateIntSeq = paste0("dateInt:", 1:length(dateSeqInt))
dateIidSeq = paste0("dateIid:", 1:length(dateSeqInt))
set.seed(123456)
resSample = inla.posterior.sample(n = 100, resHere)
resSampleFitted = lapply(resSample, function(xx) {toPredict %*% xx$la
tent[colnames(toPredict), ] + xx$latent[dateIntSeq, ] + xx$latent[dat
eIidSeq,]})
resSampleFitted = do.call(cbind, resSampleFitted)
resSampleLambda = exp(resSampleFitted)
resSampleCount = matrix(rpois(length(resSampleLambda),
                              resSampleLambda), nrow(resSampleLambda),
ncol(resSampleLambda))
matplot(dateSeq, resSampleCount, col ="#00000010",type = "l", lty =
1, log = "y")
```

```
points(xSub[, c("date", "VALUE")], col = "red")
abline(v = as.Date("2020/03/01"), col = "yellow")
```

```
is2020 = dateSeq[dateSeq >= as.Date("2020/2/1")]
sample2020 = resSampleCount[match(is2020, dateSeq),]
count2020 = xSub[match(is2020, xSub$date), "VALUE"]
excess2020 = count2020 - sample2020
matplot(is2020, excess2020, type = "l", lty = 1, col = "#0000FF10",
       ylim = range(-10, quantile(excess2020, c(0.1, 0.999))))
matlines(is2020, t(apply(excess2020, 1, quantile, prob = c(0.1,0.
9))), col = "black", lty = 2)
abline(h = 0)
```

```
quantile(apply(excess2020, 1, sum))

##       0%      25%      50%      75%     100%
## -1890.00  -949.50  -450.50   564.75  1737.00
```

#Accidents

```
dateSeq = sort(unique(x$date))
table(diff(dateSeq))

dateSeqInt = as.integer(dateSeq)
x$dateInt = x$dateIid = as.integer(x$date)
x$cos12 = cos(2 * pi * x$dateInt/365.25)
x$sin12 = sin(2 * pi * x$dateInt/365.25)
x$sin6 = sin(2 * 2 * pi * x$dateInt/365.25)
x$cos6 = cos(2 * 2 * pi * x$dateInt/365.25)
x$dayOfYear = as.Date(gsub("^[[:digit:]]+", "0000",x$date))
x$christmasBreak = (x$dayOfYear >= as.Date("0000/12/21")) |(x$dayOfYe
ar <= as.Date("0000/01/12"))
xSub = x[grepl("accidents", x$Cause, ignore.case = TRUE) & x$province
 == "Ontario", ]
xPreCovid = xSub[xSub$date < as.Date("2020/02/01") & (!xSub$christmas
Break), ]

library("INLA")
```

```
resHere = inla(VALUE ~ cos12 + cos6 + sin12 + sin6 +
                f(dateInt, model = "rw2", values = dateSeqInt,prior =
 "pc.prec", param = c(0.001, 0.5)) +
                f(dateIid, values = dateSeqInt, prior = "pc.prec", pa
ram = c(0.0001, 0.5)),
              data = xPreCovid, family = "poisson",
              control.compute = list(config = TRUE), control.predicto
r = list(compute = TRUE))

matplot(resHere$.args$data$date, resHere$summary.fitted[, paste0(c
(0.025, 0.975, 0.5), "quant")], type = "l",lty = c(2, 2, 1), col = "b
lack", log = "y", ylim = range(xSub$VALUE))
points(xSub$date, xSub$VALUE, col = "red")
points(xPreCovid$date, xPreCovid$VALUE, col = "blue")
```

```
matplot(dateSeq, resHere$summary.random$dateInt[, paste0(c(0.025,0.9
75, 0.5), "quant")], type = "l", lty = c(2,2, 1), col = "black")
```

```
toPredict = cbind(`(Intercept):1` = 1, `cos12:1` = cos(2 *pi * dateSe
qInt/365.25), `sin12:1` = sin(2 * pi * dateSeqInt/365.25), `cos6:1` =
 cos(2 * pi * dateSeqInt *2/365.25),`sin6:1` = sin(2 * pi * dateSeqIn
t *2/365.25))
dateIntSeq = paste0("dateInt:", 1:length(dateSeqInt))
dateIidSeq = paste0("dateIid:", 1:length(dateSeqInt))
set.seed(123456)
resSample = inla.posterior.sample(n = 100, resHere)
resSampleFitted = lapply(resSample, function(xx) {toPredict %*% xx$la
tent[colnames(toPredict), ] + xx$latent[dateIntSeq, ] + xx$latent[dat
eIidSeq,]})
resSampleFitted = do.call(cbind, resSampleFitted)
resSampleLambda = exp(resSampleFitted)
resSampleCount = matrix(rpois(length(resSampleLambda),
                            resSampleLambda), nrow(resSampleLambda),
ncol(resSampleLambda))
matplot(dateSeq, resSampleCount, col ="#00000010",type = "l", lty =
1, log = "y")

points(xSub[, c("date", "VALUE")], col = "red")
abline(v = as.Date("2020/03/01"), col = "yellow")
```

```
is2020 = dateSeq[dateSeq >= as.Date("2020/2/1")]
sample2020 = resSampleCount[match(is2020, dateSeq),]
count2020 = xSub[match(is2020, xSub$date), "VALUE"]
excess2020 = count2020 - sample2020

matplot(is2020, excess2020, type = "l", lty = 1, col = "#0000FF10",
        ylim = range(-10, quantile(excess2020, c(0.1, 0.999))))
matlines(is2020, t(apply(excess2020, 1, quantile, prob = c(0.1,0.
9))), col = "black", lty = 2)
abline(h = 0)
```

```
quantile(apply(excess2020, 1, sum))

##       0%      25%      50%      75%     100%
## -2053.00  1910.75  2598.00  3472.00  6950.00
```