ITS Health Intervention

Lingfeng Shi

2024-02-26

Install packages required for the analysis (uncomment if needed)

install.packages("lmtest") ; install.packages("Epi") install.packages("tsModel"); install.packages("vcd")

load the packages

library(foreign) library(tsModel) library("lmtest") library("Epi") library("splines") library("vcd")

read data from csv file

data <- read.csv("sicily.csv") head(data) data

1. Descriptive Statistics

compute the standardized rates

data$rate < -with(data, aces/stdpop*10^5)#starttheplot, excludingthepointsandthex-axisplot(data$rate,type="n",ylim=c(00 ylab="Std rate x 10,000", bty="l",xaxt="n") # shade the post intervention period grey rect(36,0,60,300,col=grey(0.9),border=F # plot the observed rate for pre-intervention period points(data$rate[data$smokban==0],cex=0.7) #specify the x-axis (i.e. time units) axis(1,at=0:5*12,labels=F) axis(1,at=0:4*12+6,tick=F,labels=2002:2006) # add a title title("Sicily, 2002-2006")

summary(data)

#tabulate aces before and after the smoking ban summary(data$aces[data$smokban==0]) summary(data$aces[data$smokban==1]) summary(data$rate[data$smokban==0]) summary(data$rate[data$smokban==1])

2. Regression model

#Poisson with the standardised population as an offset model1 <- glm(aces ~ offset(log(stdpop)) + smokban + time, family=poisson, data) summary(model1) summary(model1)$dispersion round(ci.lin(model1, Exp = T), 3) datanew < -data.frame(stdpop = mean(data$stdpop),smokban=rep(c(0,1),c(360,240)), time= 1:600/10,month=rep(1:120/10,5)) # We generate predicted values based on the model in order to create a plot pred1 <- predict(model1,type="response",datanew)/mean(data$stdpop)*10^5#Thiscanthenbeplottedalongwithascattergraph(s rate x 10,000", bty="l",xaxt="n") rect(36,0,60,300,col=grey(0.9),border=F) points(data$rate, cex = 0.7)axis(1, at = 0 : 5 * 12, labels = F)axis(1, at = 0 : 4 * 12 + 6, tick = F, labels = 2002 : 2006)lines((1 : 600/10), pred1, col = 2)title("Sicily, 2002 − 2006")datanew < −data.frame(stdpop = mean(data$stdpop),smokban=0,time=1:600/10, month=rep(1:120/10,5)) # generate predictions under the counterfactual scenario and add it to the plot pred1b <- predict(model1,datanew,type="response")/mean(data$stdpop)*

10^5*lines*(*datanew*time,pred1b,col=2,lty=2) # return the data frame to the scenario including the intervention datanew <- data.frame(stdpop=mean(data$stdpop),smokban=rep(c(0,1),c(360,240)), time= 1:600/10,month=rep(1:120/10,5))

#3. methodological issues

#Model checking and autocorrelation # Check the residuals by plotting against time res2 <- residuals(model2,type="deviance") plot(data$time,res2,ylim=c(-5,10),pch=19,cex=0.7,col=grey(0.6), main="Residuals over time",ylab="Deviance residuals",xlab="Date") abline(h=0,lty=2,lwd=2) # Further check for autocorrelation by examining the autocorrelation and # partial autocorrelation functions acf(res2) pacf(res2)

#4. adjusting for seasonality

There are various ways of adjusting for seasonality - here we use harmonic

terms specifying the number of sin and cosine pairs to include (in this

case 2) and the length of the period (12 months)

model3 <- glm(aces ~ offset(log(stdpop)) + smokban + time + harmonic(month,2,12), family=quasipoisson, data) summary(model3) summary(model3)$dispersion round(ci.lin(model3, Exp = T), 3) #EFFECTS ci.lin(model3, Exp = T)["smokban", 5 : 7] #TREND exp(coef(model3)["time"] * 12) #We again check the model and autocorrelation functions res3 < -residuals(model3, type = "deviance") plot(res3, ylim = c(-5, 10), pch = 19, cex = 0.7, col = grey(0.6), main = "Residuals over time", ylab = "Deviance residuals", xlab = "Date") abline(h = 0, lty = 2, lwd = 2) acf(res3) pacf(res3) #predict and plot of the seasonally adjusted model pred3 < -predict(model3, type = "response", datanew) / mean(data$stdpop) *10^5 plot(data$rate, type = "n", ylim = c(120, 300), xlab = "Year", ylab = "Std rate x 10,000", bty = "l", xaxt = "n") rect(36, 120, 60, 300, col = grey(0.9), border = F) points(data$rate,cex=0.7) axis(1,at=0:5*12,labels=F) axis(1,at=0:4*12+6,tick=F,labels=2002:2006) lines(1:600/10,pred3,col=2) title("Sicily, 2002-2006") # it is sometimes difficult to clearly see the change graphically in the # seasonally adjusted model, therefore it can be useful to plot a straight # line representing a 'deseasonalised' trend # this can be done by predicting all the observations for the same month, in # this case we use June pred3b <- predict(model3,type="response",transform(datanew,month=6))/ mean(data$stdpop)*10^5 #this can then be added to the plot as a dashed line lines(1:600/10,pred3b,col=2,lty=2)

5. Additional material

add a change-in-slope

we parameterize it as an interaction between time and the ban indicator

model4 <- glm(aces ~ offset(log(stdpop)) + smokban*time + harmonic(month,2,12), family=quasipoisson, data) summary(model4) round(ci.lin(model4,Exp=T),3) # predict and plot the 'deseasonalised' trend # compare it with the step-change only model pred4b <- predict(model4,type="response",transform(datanew,month=6))/ mean(data$stdpop) * 10^5 plot(data$rate,type="n",ylim=c(120,300),xlab="Year",ylab="Std rate x 10,000", bty="l",xaxt="n") rect(36,120,60,300,col=grey(0.9),border=F) points(data$rate,cex=0.7) axis(1,at=0:5*12,labels=F) axis(1,at=0:4*12+6,tick=F,labels=2002:2006) lines(1:600/10,pred3b,col=2) lines(1:600/10,pred4b,col=4) title("Sicily, 2002-2006") legend("topleft",c("Step-change only","Step-change + change-in-slope"),lty=1, col=c(2,4),inset=0.05,bty="n",cex=0.7) # test if the change-in-slope improve the fit # the selected test here is an F-test, which accounts for the overdispersion, # while in other cases a likelihood ratio or wald test can be applied anova(model3,model4,test="F") # not surprisingly, the p-value is similar to that of the interaction term