

# Are Time Series Models Adaptive to Sudden Swings in Annual Premiums for Smoker and Non-smoker Class Policyholders? (1999-2009)

## A Time Series Modeling Report by Yin Kwong (John) Lee

### Introduction

The affordability of whole life insurance has been a raising concern to families and children in different economic classes. We would like to take a time series approach to model the premiums received by the insurance companies in general or the amount of premiums paid by the policyholders. On one hand, this is important for life insurance companies to set rates and maintain adequate reserves to pay off the claims. On the other hand, premium rate has a pivotal indication on the economic well-being of the society such that during prospective economic times, people can afford rigid insurance rates, and during economic downturns they cannot. The “uslapseagent” dataset within the “CASdatasets” package is explored in this project. We do some exploratory data analysis and found that the annual premiums for smokers and non-smokers seem to decline from 2006 onward. Why? From an insurance perspective, the underwriting scheme after 2006 may be different.

Underwriting is a process to determine the risk classes of policyholders. For instance, a person classified as “high-risk” before 2006 may be classified as “standard-risk” after 2006 and therefore get charged a “standard-risk” premium rate rather than a “high-risk” premium rate. From an economic perspective, people may be more aware of risky habits. For example, more people follow family habits of smoking before 2006, but after 2006, the society dynamics may have changed and people may be more cautious of such activity, thus lowering their risk status and their respective premium rates. Since the internal system (e.g. a company’s underwriting system) and external structure (e.g. society’s value and habits) are constantly changing, in a long run, it is hard to fixate our premium analyses solely on these parameters. Therefore, we take on a more consistent approach to estimate the annual premiums with respect to time. We would like to know are sudden swings in premium rates (e.g. the sudden dropdown after 2006) estimatable by time series models, and whether risk statuses (e.g. smoker and non-smoker) matter.

### Loading packages

```
library(ggplot2)
library(dplyr)
library(gridExtra)
library(xts)
library(sp)
library(zoo)
library(CASdatasets)
library(tseries)
library(forecast)
```

### Data Pre-processing

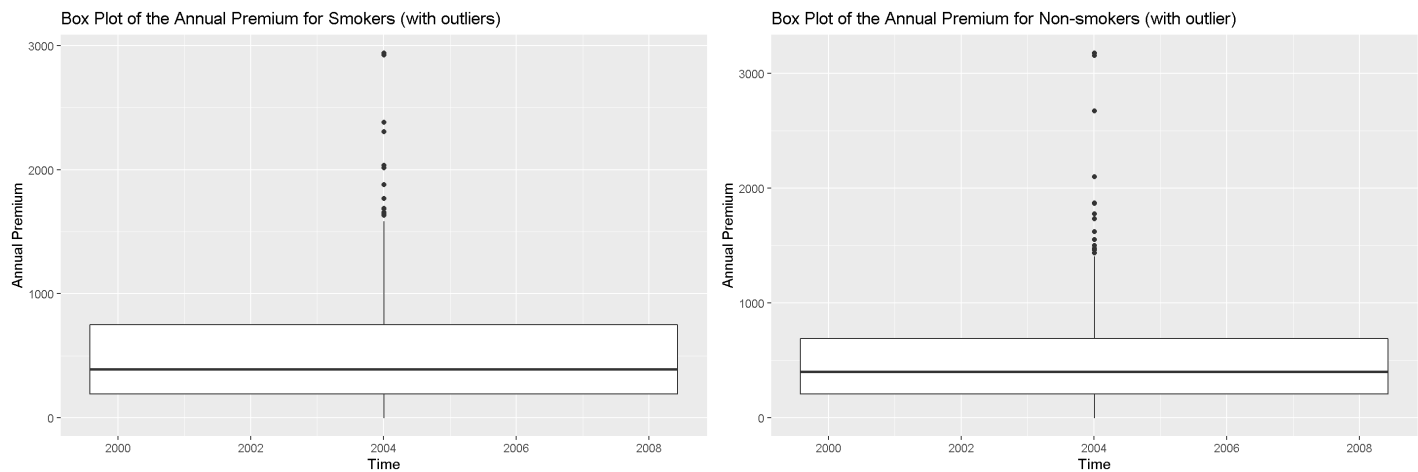
To begin, we only consider the variables “issue.date”, “risk.state”, and “annual.premium” for whole life policies that require ANNUAL PAYMENTS and those that DO NOT COVER ACCIDENTAL DEATHS. “issue.date” indicates the date that the whole life policy is issued; “risk.state” classifies the policyholder as a smoker or a non-smoker; “annual.premium” specifies the standardized annual premium that the policyholders pay. The annual premium

amount is standardized with mean=560.88 and standard deviation=526.58. We may transform the annual premium amount back to its non-standardized form for the sake of clarity and crudity. We did exactly this. We start by filtering the "premium.frequency" status to be "Annual" to contain policies with only annual premium payments, and distill the "acc.death.rider" status to be "NoRider" to limit the policies to contain no accidental death coverages. Only policies issued from 1999 to 2009 are used. Later, we separate the data into two categories: One containing whole life insurance policies issued to smokers and the other one containing whole life insurance policies issued to non-smokers. We then compare the annual premiums of the policies written on the same days for smokers and non-smokers to avoid time bias. Each day, annual premiums of policies are averaged to represent the annual premiums written on that day. We have now reduced the number of data points for the "Smoker" class from 10843 to 333 and for the "NonSmoker" class from 18474 to 333. The data is cleaned and we move on to exploring them. To remove the extreme outliers, we replace the annual premium data points that lie outside of the 1.5\*IQR limit with the 5th percentile and 95th percentile values.

```
data(uslapseagent)
#Filtering the policies issued from 1999 to 2009 with ANNUAL PAYMENTS and NO ACCIDENTAL DEATH COVE
RAGES
uslapseagent=uslapseagent%>%filter(issue.date>="1999-02-02",premium.frequency=="Annual"&acc.death.
rider=="NoRider")
# Select the variables to be considered
uslapseagent=uslapseagent[,c("issue.date","risk.state","annual.premium")]
# De-standardizing the premium amount for the sake of clarity and crudity
uslapseagent$annual.premium=uslapseagent$annual.premium*526.58+560.88
# Filtering the data so that one data frame contains policies written to smokers and the other one
contains policies written to non-smokers
smoker.prem=uslapseagent%>%filter(risk.state=="Smoker")
nonsmoker.prem=uslapseagent%>%filter(risk.state=="NonSmoker")
# We unify the dates for which the policies are issued to both smoker and non-smokers to avoid tim
e bias. There may be multiple policies issued to both smokers and non-smokers in one day.
common.dates=as.Date(intersect(factor(smoker.prem$issue.date),factor(nonsmoker.prem$issue.date)))
smoker.prem=smoker.prem%>%filter(issue.date%in%common.dates)
nonsmoker.prem=nonsmoker.prem%>%filter(issue.date%in%common.dates)

# Averaging multiple policies issued to both smokers and non-smokers in one day
avgmultipol=function(x){
  dupInd=which(duplicated(x[,1]))
  for (i in 1:length(dupInd)){
    dupAvg=mean(x[,3][which(x[,1]==x[,1][dupInd[i]])])
    x[,3][which(x[,1]==x[,1][dupInd[i]])]=dupAvg
  }
  x[!duplicated(x[,1]),]
}
smoker.prem=avgmultipol(smoker.prem)
nonsmoker.prem=avgmultipol(nonsmoker.prem)

# Plotting boxplots to examine the outliers of the annual premium amount for smokers and non-smoke
rs
plt1=ggplot(data=smoker.prem,aes(x=issue.date,y=annual.premium,group=1))+geom_boxplot()+labs(title
="Box Plot of the Annual Premium for Smokers (with outliers)",x="Time",y="Annual Premium")
plt2=ggplot(data=nonsmoker.prem,aes(x=issue.date,y=annual.premium,group=1))+geom_boxplot()+labs(ti
tle="Box Plot of the Annual Premium for Non-smokers (with outlier)",x="Time",y="Annual Premium")
grid.arrange(plt1,plt2,ncol=2)
```



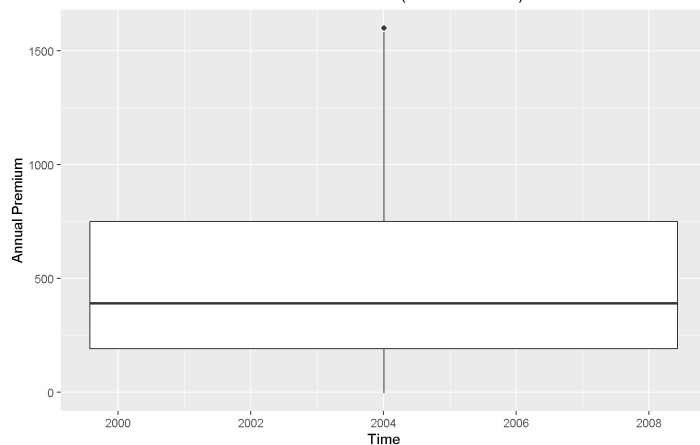
```
# Removing the outliers by capping values at the 5% and the 95% quantiles
# set inter-quantile values
s.qnt=quantile(smoker.prem$annual.premium,probs=c(0.25,0.75))
ns.qnt=quantile(nonsmoker.prem$annual.premium,probs=c(0.25,0.75))
# set replacement values
s.caps=quantile(smoker.prem$annual.premium,probs=c(0.05,0.95))
ns.caps=quantile(nonsmoker.prem$annual.premium,probs=c(0.05,0.95))
# set cutoff values
s.lim=1.5*IQR(smoker.prem$annual.premium)
ns.lim=1.5*IQR(nonsmoker.prem$annual.premium)
# replacing data with capping
smoker.prem$annual.premium[smoker.prem$annual.premium>s.qnt[[2]]+s.lim]=s.caps[[2]]
nonsmoker.prem$annual.premium[nonsmoker.prem$annual.premium>ns.qnt[[2]]+ns.lim]=ns.caps[[2]]
smoker.prem$annual.premium[smoker.prem$annual.premium<s.qnt[[1]]-s.lim]=s.caps[[1]]
nonsmoker.prem$annual.premium[nonsmoker.prem$annual.premium<ns.qnt[[1]]-ns.lim]=ns.caps[[1]]
```

## Exploratory Data Analysis

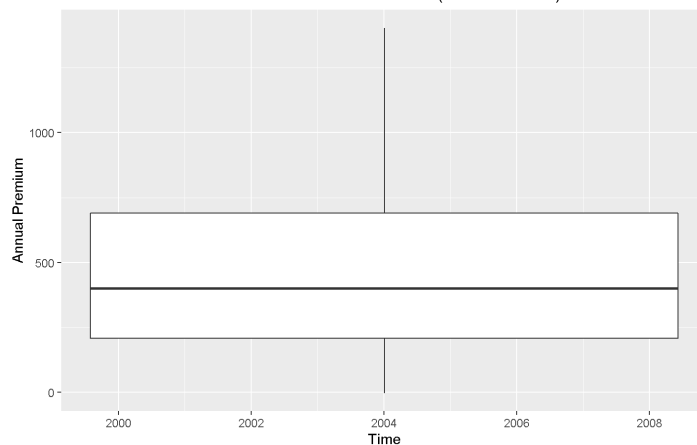
We explore the processed data with boxplots and notice that the annual premium outliers for both smoker and non-smoker groups are removed. We use this processed data to plot the series of annual premiums against time, and observe that the annual premiums for smokers seem to show a declining trend as that for non-smokers seem to also show a downward trend, but a more minor one. Annual premiums for smokers fluctuated at a [0,1650] band level with a seemingly constant mean and variance until 2006, then fluctuated at a [0,750] level afterwards up to 2009. Annual premiums for non-smokers behaves similarly as the annual premiums for smokers until 2006 with also a seemingly constant mean and variance, then fluctuated less rapidly at a slightly downward trend towards 2009. There do not seem to be a consistent trend or a seasonal pattern in the series. We then plot the correlograms of the annual premiums and find that for both smokers and non-smokers, there is no significant correlations between annual premiums at different time lags.

```
#Visualizing with boxplots
plt3=ggplot(data=smoker.prem,aes(x=issue.date,y=annual.premium,group=1))+geom_boxplot()+geom_boxplot()+labs(title="Box Plot of the Annual Premium for Smokers (without outliers)",x="Time",y="Annual Premium")
plt4=ggplot(data=nonsmoker.prem,aes(x=issue.date,y=annual.premium,group=1))+geom_boxplot()+geom_boxplot()+labs(title="Box Plot of the Annual Premium for Non-Smokers (without outliers)",x="Time",y="Annual Premium")
grid.arrange(plt3,plt4,ncol=2)
```

Box Plot of the Annual Premium for Smokers (without outliers)

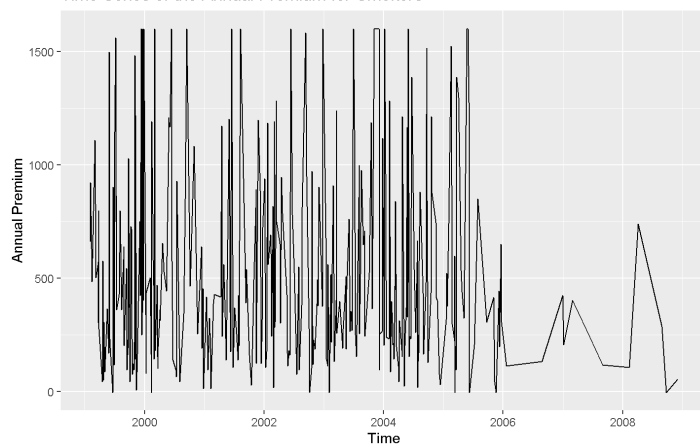


Box Plot of the Annual Premium for Non-Smokers (without outliers)

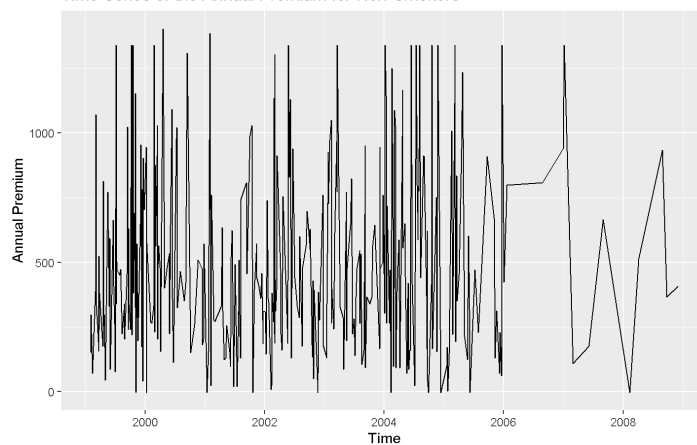
*# Visualizing with time plots*

```
plt5=ggplot(data=smoker.prem,aes(x=issue.date,y=annual.premium,group=1))+geom_line()+labs(title="T
ime Series of the Annual Premium for Smokers",x="Time",y="Annual Premium")
plt6=ggplot(data=nonsmoker.prem,aes(x=issue.date,y=annual.premium,group=1))+geom_line()+labs(title
="Time Series of the Annual Premium for Non-Smokers",x="Time",y="Annual Premium")
grid.arrange(plt5,plt6,ncol=2)
```

Time Series of the Annual Premium for Smokers

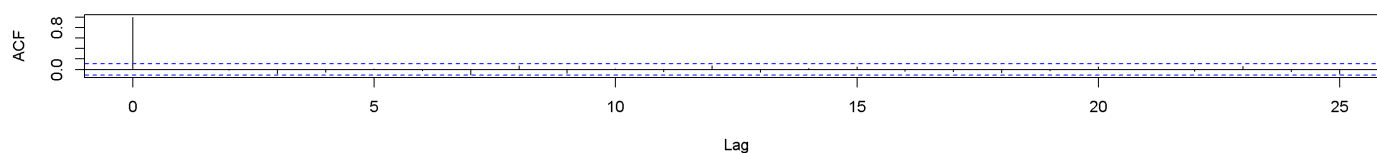


Time Series of the Annual Premium for Non-Smokers

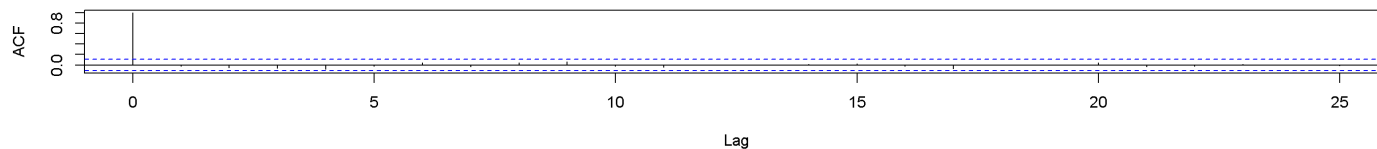
*# Visualizing correlogram*

```
par(mfrow=c(2,1))
acf(smoker.prem$annual.premium, type=c("correlation"),main="Correlogram of Annual Premiums for Sm
okers")
acf(nonsmoker.prem$annual.premium, type=c("correlation"),main="Correlogram of Annual Premiums for
Non-smokers")
```

Correlogram of Annual Premiums for Smokers



Correlogram of Annual Premiums for Non-smokers



### Testing for Stationarity

To fit stationary models for further analyses, we require the data to be stationary. From the Augmented Dickey-Fuller (ADF) Test, we reject the null hypothesis of non-stationarity for smokers and non-smokers annual premium series at p-value=0.01, and conclude that both series are stationary. The sharp decay in the autocorrelation plots above is also a phenomenon of a stationary time series.

```
# ADF test for smokers
adf.test(smoker.prem$annual.premium,alternative="stationary")
```

```
##
## Augmented Dickey-Fuller Test
##
## data:  smoker.prem$annual.premium
## Dickey-Fuller = -8.2987, Lag order = 6, p-value = 0.01
## alternative hypothesis: stationary
```

```
# ADF test for non-smokers
adf.test(nonsmoker.prem$annual.premium,alternative="stationary")
```

```
##
## Augmented Dickey-Fuller Test
##
## data:  nonsmoker.prem$annual.premium
## Dickey-Fuller = -7.6796, Lag order = 6, p-value = 0.01
## alternative hypothesis: stationary
```

### Fitting Stationary ARMA(p,q) Model

The Akaike Information Criterion (AIC) is a quantity that measures the predictability of a model while penalizing large number of parameters to discourage overfitting. It has a maximum likelihood component denoted by  $\hat{L}$  and a number of parameters component denoted by  $k$ . The quantity is written as  $AIC = 2k - 2\ln(\hat{L})$ . Here, we use AIC to select the best ARMA model. To do this, we create AIC tables to compare the AIC values between different ARMA models, for fitting annual premiums for smokers and non-smokers.

```

# Using AIC to select p and q
# Creating a function to construct AIC table for smoker annual premiums
smokerAIC=function(data,P,Q){
  table=matrix(NA,(P+1),(Q+1))
  for(p in 0:P){
    for(q in 0:Q){
      table[p+1,q+1]=arima(smoker.prem$annual.premium,order=c(p,0,q))$aic
    }
  }
  dimnames(table)=list(paste("AR",0:P,sep=""),paste("MA",0:Q,sep=""))
  table
}
# Creating a function to construct AIC table for non-smoker annual premiums
nonsmokerAIC=function(data,P,Q){
  table=matrix(NA,(P+1),(Q+1))
  for(p in 0:P){
    for(q in 0:Q){
      table[p+1,q+1]=arima(nonsmoker.prem$annual.premium,order=c(p,0,q))$aic
    }
  }
  dimnames(table)=list(paste("AR",0:P,sep=""),paste("MA",0:Q,sep=""))
  table
}

```

We inspect the AIC table for smoker annual premiums and observe the ARMA(0,0) model has the lowest AIC value at 5002.112.

```
as.data.frame(smokerAIC(smoker.prem$annual.premium,5,5))
```

##	MA0	MA1	MA2	MA3	MA4	MA5
## AR0	5002.112	5004.108	5005.992	5005.617	5006.168	5008.008
## AR1	5004.108	5006.106	5003.374	5005.076	5008.020	5009.972
## AR2	5006.000	5003.346	5005.294	5006.920	5006.864	5008.542
## AR3	5005.777	5005.170	5006.987	5004.430	5006.424	5007.669
## AR4	5006.925	5008.838	5007.446	5006.094	5007.588	5009.681
## AR5	5008.839	5010.813	5010.190	5007.669	5009.752	5011.068

We inspect the AIC table for non-smoker annual premiums and observe the ARMA(4,3) model has the lowest AIC value at 4864.730.

```
as.data.frame(nonsmokerAIC(nonsmoker.prem$annual.premium,5,5))
```

##	MA0	MA1	MA2	MA3	MA4	MA5
## AR0	4864.735	4866.499	4867.418	4867.967	4867.935	4869.885
## AR1	4866.523	4865.855	4867.195	4868.742	4869.916	4871.278
## AR2	4867.733	4867.110	4867.845	4871.083	4868.498	4867.559
## AR3	4868.441	4868.487	4869.697	4869.914	4864.754	4873.854
## AR4	4867.941	4869.631	4871.583	4864.730	4867.455	4873.653
## AR5	4869.481	4870.893	4872.278	4873.637	4869.882	4873.931

Here, we fit ARMA(0,0) model for smokers annual premium based on the smallest AIC measure. Note that this is a white noise model.

```
smoker.prem.arma00=arima(smoker.prem$annual.premium,order=c(0,0,0))
smoker.prem.arma00
```

```
##
## Call:
## arima(x = smoker.prem$annual.premium, order = c(0, 0, 0))
##
## Coefficients:
##      intercept
##      524.2176
## s.e.      24.0868
##
## sigma^2 estimated as 193198:  log likelihood = -2499.06,  aic = 5002.11
```

Here, we fit ARMA(4,3) model for non-smokers annual premium based on the smallest AIC measure.

```
nonsmoker.prem.arma43=arima(nonsmoker.prem$annual.premium,order=c(4,0,3))
nonsmoker.prem.arma43
```

```
##
## Call:
## arima(x = nonsmoker.prem$annual.premium, order = c(4, 0, 3))
##
## Coefficients:
##      ar1      ar2      ar3      ar4      ma1      ma2      ma3 intercept
##      0.2344 -0.7734  0.5234 -0.0734 -0.2619  0.7613 -0.6368  487.0793
## s.e.  0.2346  0.1191  0.2260  0.0626  0.2313  0.0875  0.2306  15.1433
##
## sigma^2 estimated as 121057:  log likelihood = -2423.36,  aic = 4864.73
```

We examine the roots of the AR and MA components to check for causality and invertibility, respectively. We examine causality to check if the time series is dependent on past values, and look out for invertibility to distinguish unique or non-unique models. Although the ARMA (0,0) model is suggested for estimating the annual premium for smokers, it does not serve for causality or invertibility because of its white noise properties. To state simply, its roots do not exist.

*Fitting ARMA(2,1) model for smokers annual premium and examining its roots*

We examine the ARMA(2,1) model instead, which has the second smallest AIC value at 5003.346. Here, the AR roots outside of the unit circle implies a causal process, as the MA roots outside of the unit circle implies an invertible process.

```
# Fitting ARMA(2,1) model for smokers annual premium
smoker.prem.arma21=arima(smoker.prem$annual.premium,order=c(2,0,1))
# Exploring the AR roots to determine causality
polyroot(c(1,-coef(smoker.prem.arma21)[c("ar1","ar2")]))
```

```
## [1] 1.097109+0i 22.365811+0i
```

```
# Exploring the MA roots to determine invertibility
polyroot(c(1, -coef(smoker.prem.arma21)[c("ma1")]))
```

```
## [1] -1.033007+0i
```

### *Roots of the ARMA(4,3) model for estimating non-smokers annual premium*

We examine the AR roots of the suggest ARMA(4,3) model, and realize that not all the roots are outside of the unit circle, implying that this model is not causal. Similarly, not all MA roots lie outside the unit circle, indicating that this model is not invertible. Without causality and invertibility, the model may be at strong disadvantage for its predictability and reliability. Thus, we select another model to fit the non-smoker annual premium series that remedies these concerns.

```
# Exploring the AR roots to determine causality
polyroot(c(1, -coef(nonsmoker.prem.arma43)[c("ar1", "ar2", "ar3", "ar4")]))
```

```
## [1] -0.201471+1.025135i -0.201471-1.025135i  2.459412+0.000000i
## [4]  5.077820-0.000000i
```

```
# Exploring the MA roots to determine invertibility
polyroot(c(1, -coef(nonsmoker.prem.arma43)[c("ma1", "ma2", "ma3")]))
```

```
## [1]  0.9935037+0.9984848i -0.7915301+0.0000000i  0.9935037-0.9984848i
```

### *Fitting ARMA(1,1) model for non-smokers annual premium and examining its roots*

We examine the ARMA(1,1) model instead, which has the third smallest AIC value at 4865.855. Here, the AR roots outside of the unit circle implies a causal process, as the MA roots outside of the unit circle implies an invertible process.

```
# Fitting ARMA(1,1) model for non-smokers annual premium
nonsmoker.prem.arma11=arima(nonsmoker.prem$annual.premium, order=c(1,0,1))
# Exploring the AR roots to determine causality
polyroot(c(1, -coef(nonsmoker.prem.arma11)[c("ar1")]))
```

```
## [1] 1.361402+0i
```

```
# Exploring the MA roots to determine invertibility
polyroot(c(1, -coef(nonsmoker.prem.arma11)[c("ma1")]))
```

```
## [1] -1.252775+0i
```

Ultimately, we choose to fit the smokers annual premium with the ARMA(2,1) model and to fit the non-smokers annual premium with the ARMA(1,1) model. We learned that although AIC is a fair measure for model complexity vs predictability, it is also beneficial to select models that are causal and invertible for the sake of predictability and reliability.

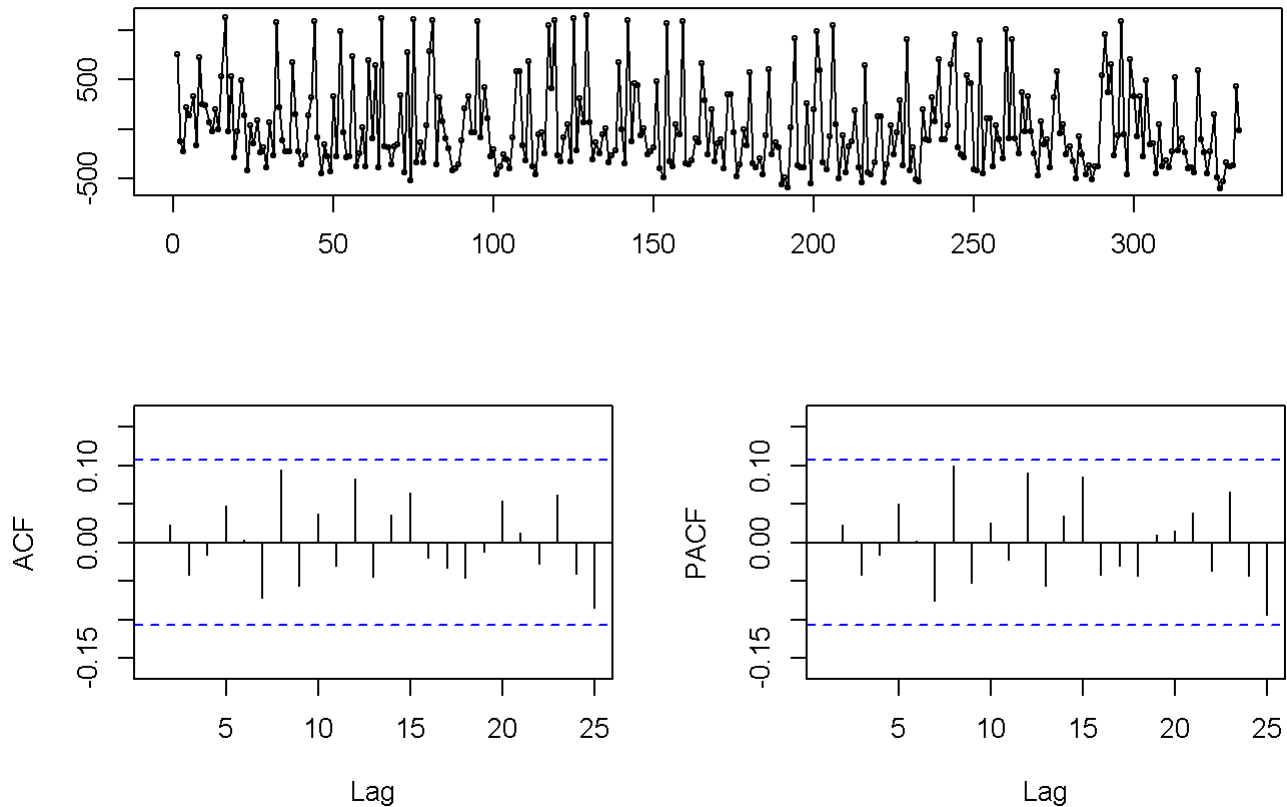


### Residual Analysis

The residuals for both the ARMA(2,1) and ARMA(1,1) models fall within the bounds in the autocorrelation (ACF) graph and do not seem to have great significance. This shows signs of accuracy for fitting the ARMA(2,1) model on smokers annual premium and the ARMA(1,1) model on non-smokers annual premium.

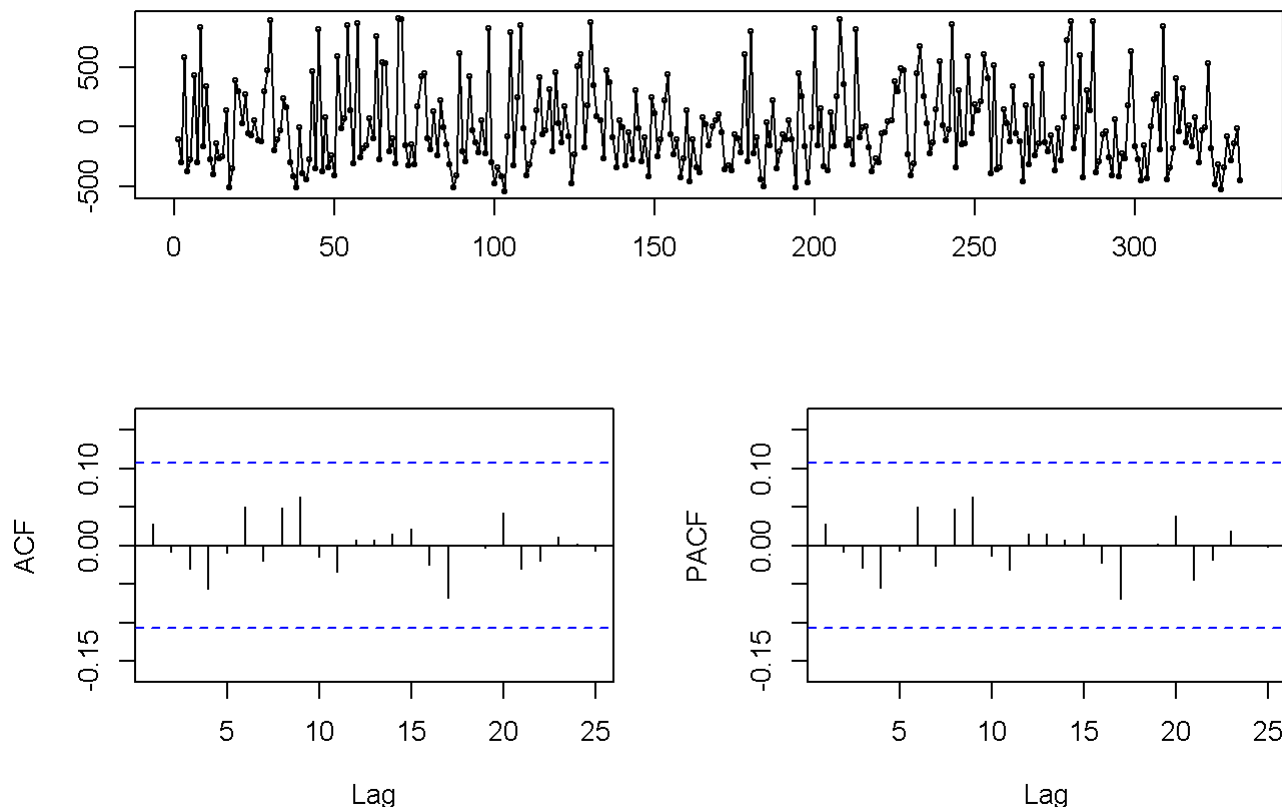
```
# Residuals on fitting ARMA(2,1) model on smokers annual premium
tsdisplay(residuals(smoker.prem.arma21),main="ARMA(2,1) Residuals for Fitting Smokers Annual Premium")
```

#### ARMA(2,1) Residuals for Fitting Smokers Annual Premium



```
# Residuals on fitting ARMA(1,1) model on non-smokers annual premium
tsdisplay(residuals(nonsmoker.prem.arma11),main="ARMA(1,1) Residuals for Fitting Non-smokers Annual Premium")
```

### ARMA(1,1) Residuals for Fitting Non-smokers Annual Premium

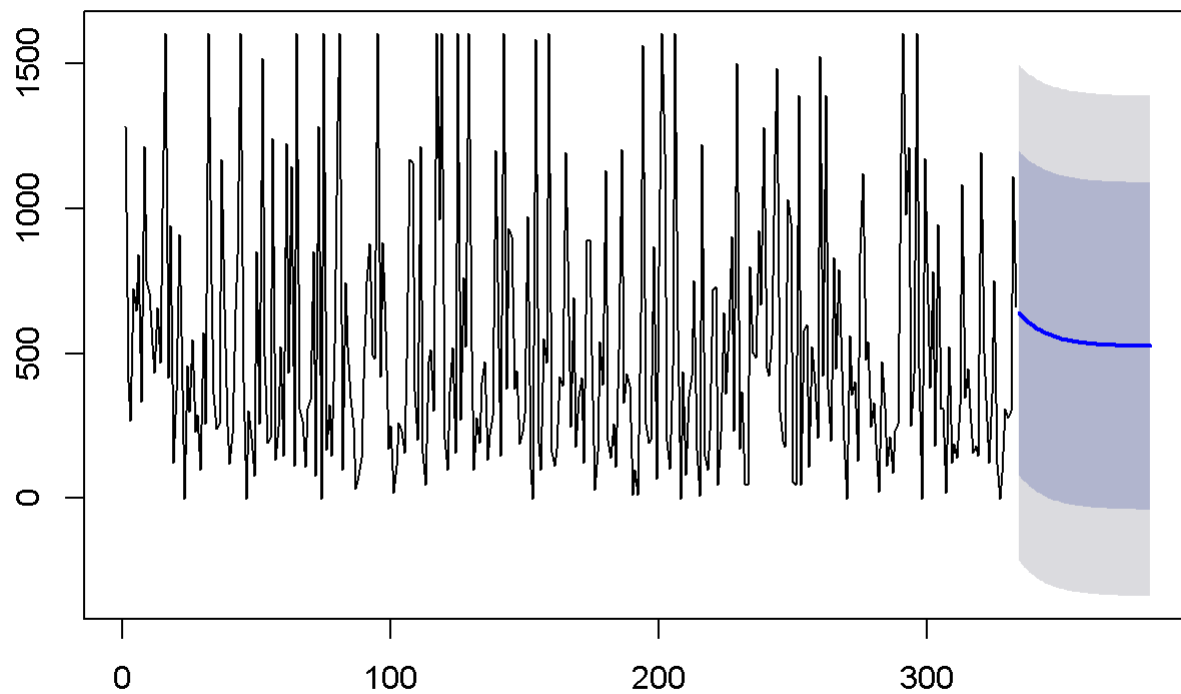


#### Model Forecast and Outlook

According to the earlier time series plot, it is reasonable for our outlook for the annual premiums to decrease in the future. This is because of the drastic drop in smokers and non-smokers annual premiums beyond 2006. The ARIMA(2,1) forecast for the smokers annual premium in the next 50 days seems to decrease. A same phenomenon is observed for the ARIMA(1,1) forecast for the non-smokers annual premium. These forecasts align with our expectation for the directions of the smokers and non-smokers annual premium in the future.

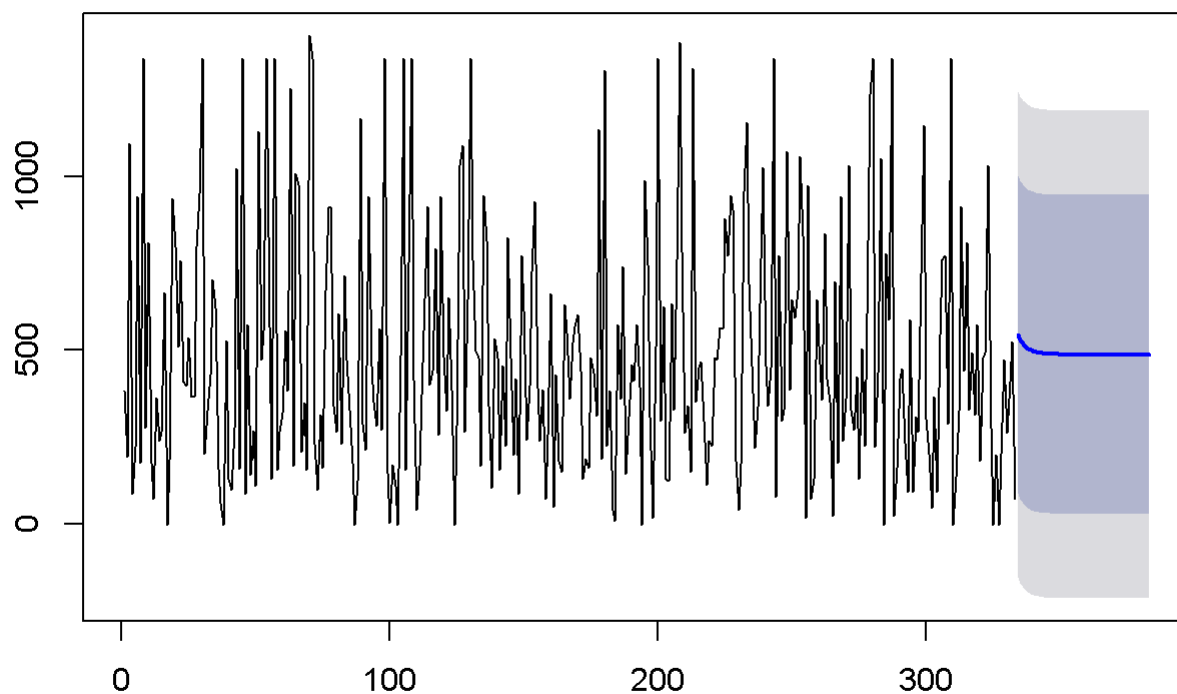
```
# The average annual premiums for whole policies written for smokers in the next 50 days
plot(forecast(smoker.prem.arma21,h=50))
```

## Forecasts from ARIMA(2,0,1) with non-zero mean



```
# The average annual premiums for whole policies written for non-smokers in the next 50 days  
plot(forecast(nonsmoker.prem.arma11,h=50))
```

## Forecasts from ARIMA(1,0,1) with non-zero mean

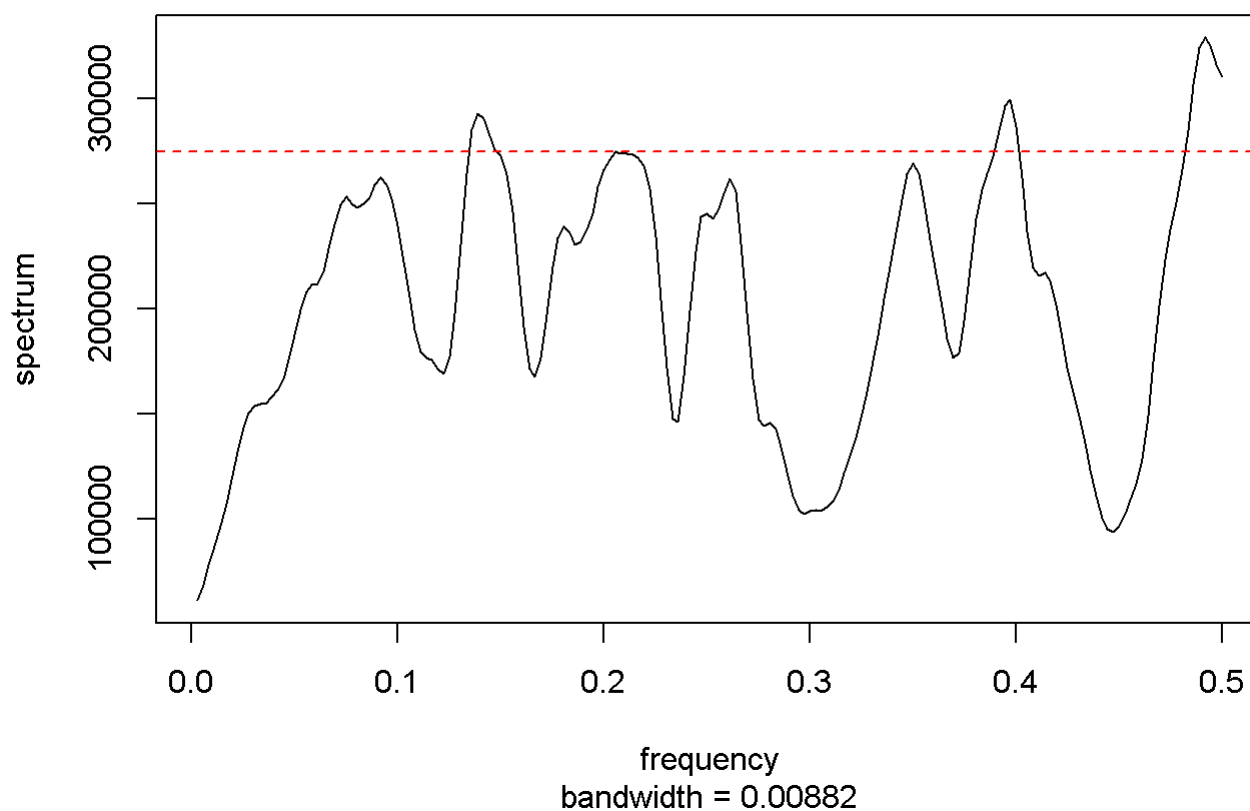


### Spectral Analysis

We smoothed the original periodograms 3 times with different moving average smoothers in order to remove the background noise of the time series. The frequency of each periodogram is 10 cycles per unit time as the period for each periodogram is 1/10. The spectral density for the smokers annual premium consists of 7 distinct peaks with different spaces.; this means that the annual periodic component is sinusoidal. The peaks are centered around the 275000 level of the power spectrum (denoted by the red dashed line), and seem to occur during irregular cycles. The spectral density for the non-smokers annual premium consists of 3 major peaks at around 0.14, 0.21, and 0.33 frequency levels with the maximum hitting around 200000 power spectrum. Each of the hiking patterns seems to take up one cycle. We compare both periodograms and find that the peak powers are more extreme for the smokers annual premium. This means that the autocovariances and thus, the autocorrelations for smokers annual premium are higher.

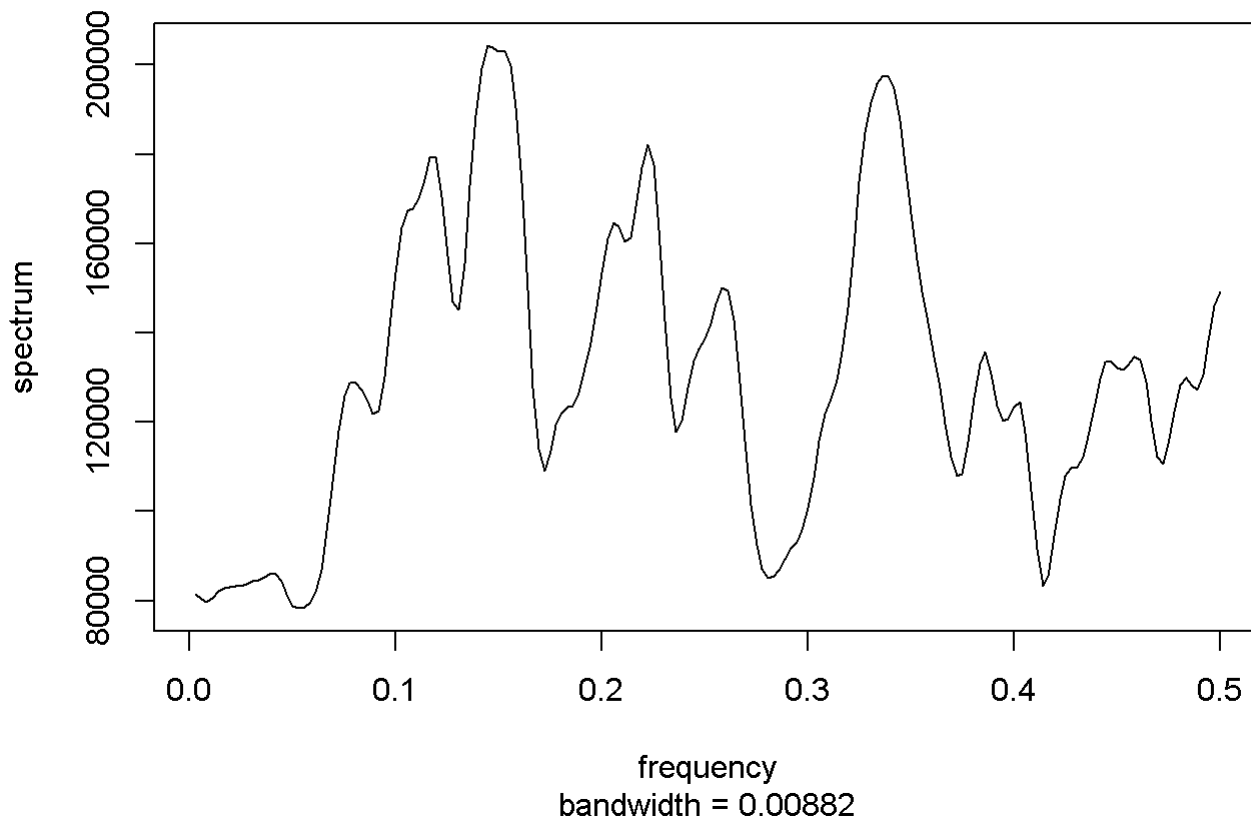
```
# Smoothed periodogram of smokers annual premium
spectrum(smoker.prem$annual.premium, log="no", spans=c(5,10), main="Smoothed Periodogram of Annual
Premiums for Smokers")
abline(h=275000, lty="dashed", col="red")
```

## Smoothed Periodogram of Annual Premiums for Smokers



```
# Smoothed periodogram of non-smokers annual premium  
spectrum(nonsmoker.prem$annual.premium, log="no", spans=c(5,10), main="Smoothed Periodogram of Annual Premiums for Non-smokers")
```

## Smoothed Periodogram of Annual Premiums for Non-smokers



### Estimating Trend by Loess Smoothing

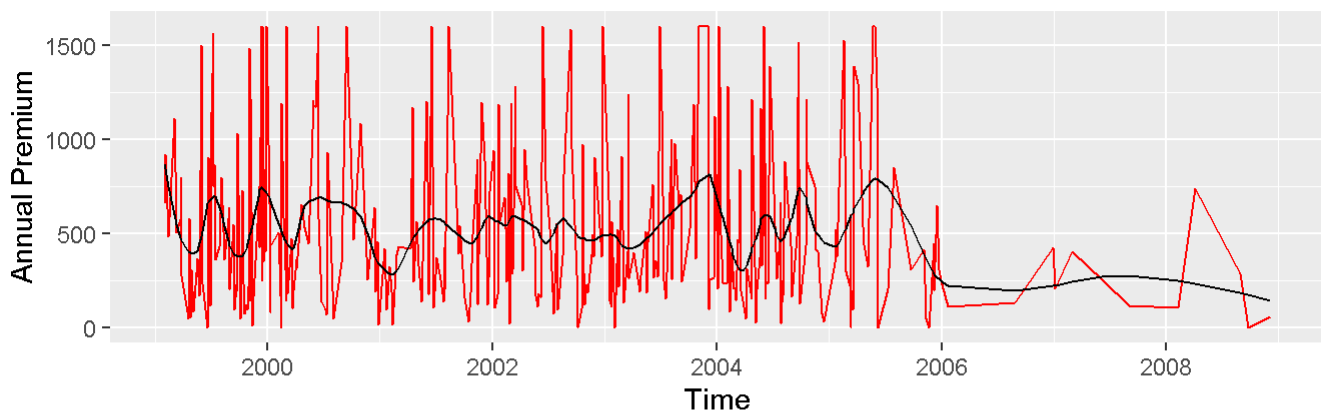
We set span=0.1 here and observe that the loess smoothing method estimates the trend of the original series very well. For smokers annual premium, loess follows the fluctuating trend from 1999 to 2006 then flags at a lower level thereafter. For non-smokers annual premium, loess fluctuates similarly until 2006, then rises to catch the hike before dropping down to a lower level.

```
# Loess smoothing for smoker annual premium
smoker.loess=loess(annual.premium~as.numeric(issue.date),data=smoker.prem,span=0.1)
plt7=ggplot(data=smoker.prem,aes(x=issue.date,y=annual.premium,group=1))+geom_line(color="red")+geom_line(aes(x=smoker.prem$issue.date,y=smoker.loess$fitted),color="black",linetype="solid")+labs(title="Time Series of the Annual Premium for Smokers",x="Time",y="Annual Premium")
# Loess smoothing for non-smoker annual premium
nonsmoker.loess=loess(annual.premium~as.numeric(issue.date),data=nonsmoker.prem,span=0.1)
plt8=ggplot(data=nonsmoker.prem,aes(x=issue.date,y=annual.premium,group=1))+geom_line(color="red")+geom_line(aes(x=nonsmoker.prem$issue.date,y=nonsmoker.loess$fitted),color="black",linetype="solid")+labs(title="Time Series of the Annual Premium for Non-smokers",x="Time",y="Annual Premium")
grid.arrange(plt7,plt8,nrow=2)
```

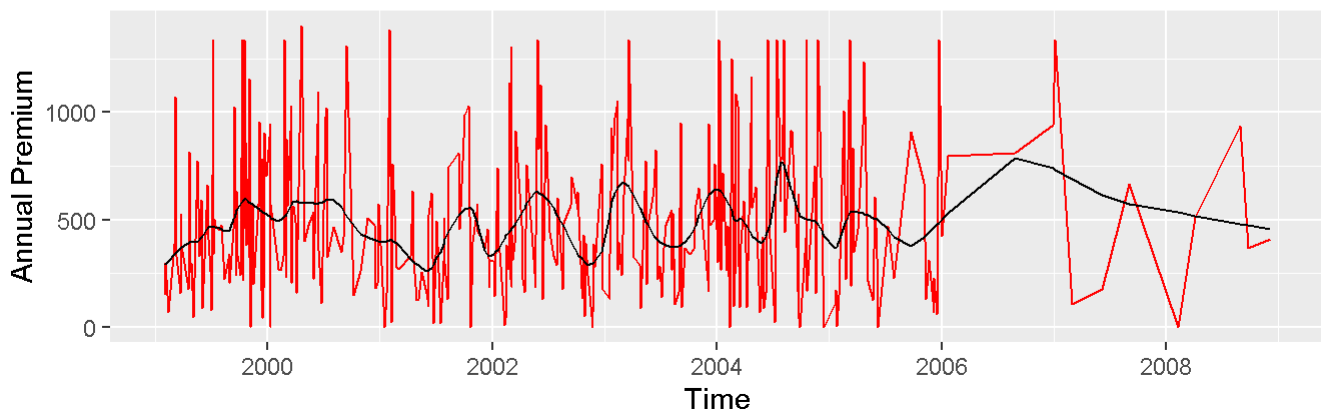
```
## Warning: Use of `smoker.prem$issue.date` is discouraged. Use `issue.date`
## instead.
```

```
## Warning: Use of `nonsmoker.prem$issue.date` is discouraged. Use `issue.date`
## instead.
```

### Time Series of the Annual Premium for Smokers



### Time Series of the Annual Premium for Non-smokers



### Conclusion

After our analysis, we found ARMA(2,1) as the proper model for estimating the smokers annual premiums, and ARMA(1,1) as the proper model for estimating the non-smokers annual premiums. Both models are fitted to forecast a resulting downward trend for annual premiums for different risk status, which is what we expect. This means that whether or not the insurance companies change their underwriting schemes, we are still able to predict the trend of the annual premiums based on the passage of time. In other words, as time goes on, the behaviors of the society may change, but we are still able to get adaptable results in time series that reflects the changes. In the above case, the downward forecasts reflect changes in society's behaviors, and because smokers annual premium (not just non-smokers annual premium ) also displays such forecast property, these behaviors are not only caused by a reduction in smoking activities. To conclude, time series models, at least the ones listed, are adaptive to sudden swings in annual premiums for smoker and non-smoker class policyholders.

### References

- (1) Dataset used in Milhaud and Dutang (2018), Lapse tables for lapse risk management in insurance: a competing risk approach, European Actuarial Journal, 2018, Volume 8, Issue 1.
- (2) Xavier Milhaud, Christophe Dutang. Lapse tables for lapse risk management in insurance: a competing risk approach. European Actuarial Journal, Springer, 2018, 8 (1), pp.97-126.
- (3) <http://dutangc.free.fr/pub/RRepos/web/CASdatasets-manual.pdf>  
(<http://dutangc.free.fr/pub/RRepos/web/CASdatasets-manual.pdf>)
- (4) <https://www.iii.org/press-release/life-insurance-premiums-expected-to-decline-by-4-percent-in-2007-says-the-insurance-information-institute-091906> (<https://www.iii.org/press-release/life-insurance-premiums-expected-to-decline-by-4-percent-in-2007-says-the-insurance-information-institute-091906>)