Assignment5 Yunzhi Wang

Part1: forecast beer sales for all months of 1990 1A

```
knitr::opts_chunk$set(
    echo = TRUE,
    message = FALSE,
    warning = FALSE
)
library(TSA)

## Loading required package: leaps

## Warning: package 'leaps' was built under R version 3.3.2

## Loading required package: locfit

## locfit 1.5-9.1    2013-03-22

## Loading required package: mgcv

## Warning: package 'mgcv' was built under R version 3.3.2

## Loading required package: nlme

## Warning: package 'nlme' was built under R version 3.3.2

## This is mgcv 1.8-22. For overview type 'help("mgcv-package")'.

## Loading required package: tseries

## Warning: package 'tseries' was built under R version 3.3.2

##
## Attaching package: 'TSA'

## The following objects are masked from 'package:stats':
##
##     acf, arima

## The following object is masked from 'package:utils':
##
##     tar

library(tseries)
library(forecast)

## Warning: package 'forecast' was built under R version 3.3.2

## Warning in as.POSIXlt.POSIXct(Sys.time()): unknown timezone 'default/
## America/Chicago'
```

```
##
## Attaching package: 'forecast'

## The following object is masked from 'package:nlme':
##
##     getResponse

library(timeSeries)

## Loading required package: timeDate

##
## Attaching package: 'timeDate'

## The following objects are masked from 'package:TSA':
##
##     kurtosis, skewness

library(hydroGOF)

## Warning: package 'hydroGOF' was built under R version 3.3.2

## Loading required package: zoo

## Warning: package 'zoo' was built under R version 3.3.2

##
## Attaching package: 'zoo'

## The following object is masked from 'package:timeSeries':
##
##     time<-

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

##
## Attaching package: 'hydroGOF'

## The following object is masked from 'package:locfit':
##
##     cp

data(beersales)
#days in 01/1975- 12/1989
d <- beersales[1:180]
y <- beersales[181:192]
beersales

##          Jan     Feb     Mar     Apr     May     Jun     Jul     Aug
## 1975 11.1179  9.8413 11.5732 13.0097 13.4182 14.4418 14.7534 13.8816
## 1976 10.8633 11.0000 10.9934 12.9140 13.5853 14.1553 15.0056 14.8590
```

```
## 1977 10.0067 10.4321 14.5477 14.2748 14.9986 15.7100 14.7980 14.6431
## 1978 10.6897 11.0093 14.7983 13.5984 14.9606 15.8187 15.2871 16.2773
## 1979 12.3244 12.0133 15.0094 14.9562 15.9268 15.5702 15.1282 15.5625
## 1980 12.5357 12.6446 14.0848 14.3271 16.1862 16.6604 17.0810 16.2811
## 1981 12.0798 12.4126 15.0092 15.4733 16.9966 17.2933 17.3701 16.2422
## 1982 11.9036 12.9126 15.6815 15.8119 16.5611 17.2255 16.1033 16.2590
## 1983 12.5696 12.6644 15.0723 15.5742 16.8397 17.0121 16.8476 17.3471
## 1984 12.4214 12.5443 15.3242 15.0629 16.8656 17.2300 17.3288 16.9654
## 1985 13.5114 12.7501 14.4642 15.8558 17.6043 16.1731 16.6319 16.0352
## 1986 13.9861 13.0120 14.6625 16.0165 17.1046 16.5952 17.0626 16.3092
## 1987 13.6094 13.7362 15.3119 15.9071 16.1350 16.6147 17.0362 15.8162
## 1988 13.8006 13.9416 15.2575 15.2452 16.4849 17.0435 16.4097 16.2246
## 1989 14.0913 13.1950 15.4059 14.8754 16.7768 16.9378 16.2259 17.4078
## 1990 14.2600 13.3800 15.8900 15.2300 16.9100 16.8854 17.0000 17.4000
##          Sep     Oct     Nov     Dec
## 1975 12.5123 11.8983 10.6088 11.5874
## 1976 13.4387 12.2184 10.5208 10.8335
## 1977 12.8878 11.6235 11.4853 11.5065
## 1978 13.9370 13.3270 12.0353 11.5670
## 1979 13.7112 13.6425 12.5158 11.7629
## 1980 14.5118 14.1594 12.5120 12.3830
## 1981 14.6808 13.8444 12.3871 12.9072
## 1982 14.8834 13.8291 13.1376 12.2662
## 1983 14.8442 13.8536 12.7904 11.9797
## 1984 13.6582 14.2932 12.4037 11.3818
## 1985 13.5914 14.0102 12.3939 12.1101
## 1986 14.0156 14.6417 12.4761 12.8391
## 1987 14.3066 14.4671 12.5856 12.3225
## 1988 14.4386 13.9469 13.2062 12.2347
## 1989 14.7684 14.3167 13.4048 12.0999
## 1990 14.7500 15.7700 14.5400 13.2200
```

```r
dts <- ts(d)
m <- auto.arima(dts, seasonal = FALSE)
mm <- forecast(m, h = 12)
mse(mm$mean, y)
```

```
## [1] 15.96367
```

1B

```r
#forecast Jan 1990
m1 <- forecast(m, h=1)
m1
```

```
##     Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## 181       11.97945 10.62865 13.33025 9.913577 14.04532
```

```r
as.numeric(m1$mean)
```

```
## [1] 11.97945
```

```r
n1 <- c(d,as.numeric(m1$mean))
#forecast Feb 1990
n1ts <- ts(n1)
n1arima <- auto.arima(n1ts)
m2 <- forecast(n1arima, h=1)
n2 <- c(n1,as.numeric(m2$mean))
#Mar 1990
n2ts <- ts(n2)
n2arima <- auto.arima(n2ts)
m3 <- forecast(n2arima, h=1)
n3 <- c(n2,as.numeric(m3$mean))
#Apr 1990
n3ts <- ts(n3)
n3arima <- auto.arima(n3ts)
m4 <- forecast(n3arima, h=1)
n4 <- c(n3,as.numeric(m4$mean))
#may 1990
n4ts <- ts(n4)
n4arima <- auto.arima(n4ts)
m5 <- forecast(n4arima, h=1)
n5 <- c(n4,as.numeric(m5$mean))
#Jun 1990
n5ts <- ts(n5)
n5arima <- auto.arima(n5ts)
m6 <- forecast(n5arima, h=1)
n6 <- c(n5,as.numeric(m6$mean))
#JUL 1990
n6ts <- ts(n6)
n6arima <- auto.arima(n6ts)
m7 <- forecast(n6arima, h=1)
n7 <- c(n6,as.numeric(m7$mean))
#Aug 1990
n7ts <- ts(n7)
n7arima <- auto.arima(n7ts)
m8 <- forecast(n7arima, h=1)
n8 <- c(n7,as.numeric(m8$mean))
#Sep 1990
n8ts <- ts(n8)
n8arima <- auto.arima(n8ts)
m9 <- forecast(n8arima, h=1)
n9 <- c(n8,as.numeric(m9$mean))
#Oct 1990
n9ts <- ts(n9)
n9arima <- auto.arima(n9ts)
m10 <- forecast(n9arima, h=1)
n10 <- c(n9,as.numeric(m10$mean))
#Nov 1990
n10ts <- ts(n10)
n10arima <- auto.arima(n10ts)
m11 <- forecast(n10arima, h=1)
```

```r
n11 <- c(n10,as.numeric(m11$mean))
#Dec 1990
n11ts <- ts(n11)
n11arima <- auto.arima(n11ts)
m12 <- forecast(n11arima, h=1)
n12 <- c(n11,as.numeric(m12$mean))
```

1C: mse in 1A is 15.96367, while the mse in 1B is 15.96326, which shows that the second approach(h=1 in all forecast) is better to forecst in forecasting.

```r
#1A
mse(mm$mean, y)

## [1] 15.96367

#1B
nn <- n12[181:192]
mse(nn,y)

## [1] 15.96326
```

Part2

```r
ddts <- ts(d, frequency = 12)
mk <- auto.arima(ddts, seasonal = TRUE)
mk

## Series: ddts
## ARIMA(4,1,2)(2,1,2)[12]
##
## Coefficients:
##          ar1      ar2     ar3      ar4      ma1     ma2    sar1    sar2
##       0.5103  -0.1662  0.1032  -0.3966  -1.1757  0.3125  0.6838  -0.592
## s.e.  0.1453   0.0986  0.0863   0.0789   0.1493  0.1421  0.1451   0.165
##          sma1    sma2
##       -1.1967  0.5849
## s.e.   0.1394  0.2087
##
## sigma^2 estimated as 0.2837:  log likelihood=-134.55
## AIC=291.1   AICc=292.81   BIC=325.4

mkm <- forecast(mk, h = 12)
mkm

##        Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## Jan 16       13.81601 13.13331 14.49871 12.77191 14.86011
## Feb 16       13.07707 12.35715 13.79698 11.97605 14.17808
## Mar 16       14.96181 14.23546 15.68817 13.85095 16.07268
## Apr 16       15.58503 14.83785 16.33220 14.44232 16.72774
## May 16       17.24847 16.49698 17.99996 16.09917 18.39777
## Jun 16       16.86360 16.10993 17.61727 15.71096 18.01624
## Jul 16       16.95571 16.19987 17.71156 15.79974 18.11168
```

```
## Aug 16       17.02231 16.26451 17.78012 15.86336 18.18127
## Sep 16       14.28619 13.51600 15.05638 13.10828 15.46410
## Oct 16       14.55136 13.75967 15.34305 13.34057 15.76214
## Nov 16       12.89695 12.09174 13.70216 11.66548 14.12841
## Dec 16       12.30127 11.48554 13.11699 11.05372 13.54881
```

```r
mse(as.numeric(mkm$mean),y)
```

```
## [1] 0.565002
```

Part3 The mse in the part 2 is 0.565002, which is greatly smaller than the two approaches in Part1, which suggest that the the approach in Part 2 is better to forecast beer sales for each month of 1990.