Time Series Final Exam

104304033 統計四 劉書宏

Q1 U.S. tobacco production from 1871 to 1985 is shown in worksheet

(a) ARIMA model

Series: data

ARIMA(0,1,1) with drift

Coefficients:

ma1 drift

-0.594 11.6424

s.e. 0.097 7.6008

sigma^2 estimated as 39571: log likelihood=-764.36

AIC=1534.72 AICc=1534.94 BIC=1542.93

Training set error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set 0.3101048 196.3133 140.4715 -1.298765 11.51846 0.8818144 0.06338094

By using auto.arima and seasonal = false, the data was fitted with ARIMA(0,1,1)

(b) Forecast next 10 years with 95% forecast limits

|  |  |  |  |
| --- | --- | --- | --- |
| Year | predict\_10.mean | 95%lower | 95%upper |
| 1986 | 1688.827 | 1298.941 | 2078.713 |
| 1987 | 1700.469 | 1279.67 | 2121.268 |
| 1988 | 1712.112 | 1262.521 | 2161.703 |
| 1989 | 1723.754 | 1247.107 | 2200.402 |
| 1990 | 1735.397 | 1233.148 | 2237.645 |
| 1991 | 1747.039 | 1220.433 | 2273.645 |
| 1992 | 1758.681 | 1208.796 | 2308.567 |
| 1993 | 1770.324 | 1198.104 | 2342.543 |
| 1994 | 1781.966 | 1188.253 | 2375.68 |
| 1995 | 1793.609 | 1179.153 | 2408.065 |

the forecast values are increasing, but the interval are becoming big.

Q2 U.S. monthly employment figures for young men between 16 and 19 of age from 1971 to 1981

(a)

Series: data

ARIMA(2,0,3) with non-zero mean

Coefficients:

ar1 ar2 ma1 ma2 ma3 mean

1.6125 -0.6344 -1.0692 -0.3389 0.5381 821.0805

s.e. 0.1780 0.1722 0.1609 0.1486 0.0788 54.7779

sigma^2 estimated as 14318: log likelihood=-816.62

AIC=1647.24 AICc=1648.14 BIC=1667.42

Training set error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set 4.319159 116.9058 79.73285 -1.333272 9.500921 0.7908687 0.009577191

Using auto.arima and add seasonal = true ,then get the ARIMA (2,0,3)..

(b)

|  |  |  |  |
| --- | --- | --- | --- |
| Index | mean | 99% lower | 99% upper |
| 1 | 927.4282 | 619.2126 | 1235.644 |
| 2 | 900.9468 | 550.1774 | 1251.716 |
| 3 | 901.2175 | 549.1709 | 1253.264 |
| 4 | 899.6357 | 547.4077 | 1251.864 |
| 5 | 896.9133 | 542.7228 | 1251.104 |
| 6 | 893.5268 | 535.4121 | 1251.642 |
| 7 | 889.7931 | 526.4135 | 1253.173 |
| 8 | 885.9208 | 516.624 | 1255.218 |
| 9 | 882.0453 | 506.7067 | 1257.384 |
| 10 | 878.2526 | 497.0882 | 1259.417 |
| 11 | 874.5955 | 488.0125 | 1261.178 |
| 12 | 871.1043 | 479.5986 | 1262.61 |

the forecast values are strictly decreasing.

Q3 The Lydia Pinkham annual advertising (xt) and sales (yt) data

(a) Transfer function Noise model

|  |
| --- |
|  |
| |  | | --- | |  | |

Call:

arimax(x = data2$yt, order = c(1, 0, 0), include.mean = FALSE, xtransf = data.frame(data2$lag3x),

transfer = list(c(2, 0)))

Coefficients:

ar1 data2.lag3x-AR1 data2.lag3x-AR2 data2.lag3x-MA0

0.9891 -0.3611 -0.9603 -0.0921

s.e. 0.0111 0.0820 0.0928 0.0834

sigma^2 estimated as 50397: log likelihood = -350.39, aic = 708.77

I take lag3 in transfer function.

(b)

|  |  |  |  |
| --- | --- | --- | --- |
| Month | 95% lower | mean | 95% upper |
| 1 | 1012.355 | 1354.004 | 1695.652 |
| 2 | 743.0841 | 1324.35 | 1905.615 |
| 3 | 483.138 | 1300.627 | 2118.115 |
| 4 | 230.1466 | 1281.648 | 2333.149 |
| 5 | -16.307 | 1266.465 | 2549.237 |
| 6 | -256.915 | 1254.319 | 2765.552 |
| 7 | -492.724 | 1244.602 | 2981.927 |
| 8 | -724.998 | 1236.828 | 3198.654 |
| 9 | -955.092 | 1230.609 | 3416.31 |
| 10 | -1184.36 | 1225.634 | 3635.632 |
| 11 | -1414.13 | 1221.654 | 3857.439 |
| 12 | -1645.64 | 1218.47 | 4082.58 |

the forecast values are decreasing.

Q4 The daily closing stock of Duke Energy Corporation between January 3, 2002, and August 31, 2002

(a)

Series: pre\_data

ARIMA(0,1,0)

sigma^2 estimated as 0.51: log likelihood=-138.52

AIC=279.05 AICc=279.08 BIC=281.9

Training set error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set -0.08294729 0.7113454 0.5348822 -0.2643994 1.553825 0.9928207 0.02953077

Before intervention occurred, I used the first 129 rows to make an arima model.

Because the arima model (0,1,0), so next we will take difference to do the intervention model.

Before intervention occurred, I give the data a value 0, and give the data after the intervention occurred 1.

Call:

arimax(x = diff(ts(Q4)), order = c(0, 0, 1), include.mean = FALSE, xtransf = St[2:T],

transfer = list(c(0, 0)))

Coefficients:

ma1 T1-MA0

-0.0044 -0.0641

s.e. 0.0732 0.1409

sigma^2 estimated as 0.74: log likelihood = -209.29, aic = 422.57

(b)

The effect of the intervention causes the stock price dropping sharply.

(c)

|  |  |  |  |
| --- | --- | --- | --- |
| Day | 95% lower | mean | 95% upper |
| 1 | -0.07695 | -0.06402 | -0.05108 |
| 2 | -0.0796 | -0.06402 | -0.04843 |
| 3 | -0.08186 | -0.06402 | -0.04617 |
| 4 | -0.08387 | -0.06402 | -0.04416 |
| 5 | -0.08569 | -0.06402 | -0.04234 |
| 6 | -0.08737 | -0.06402 | -0.04066 |
| 7 | -0.08894 | -0.06402 | -0.03909 |
| 8 | -0.09041 | -0.06402 | -0.03762 |
| 9 | -0.0918 | -0.06402 | -0.03623 |
| 10 | -0.09313 | -0.06402 | -0.0349 |
| 11 | -0.0944 | -0.06402 | -0.03363 |
| 12 | -0.09562 | -0.06402 | -0.03241 |

the forecast values are all the same.

R code

library(openxlsx)

library(tseries)

library(forecast)

### Q1

Q1 = read.xlsx("2019Spring\_Time\_Series\_Analysis\_final\_exam\_0620.xlsx", sheet = 1)

data = ts(Q1$Tobacco.Production)

model = auto.arima(data, seasonal = FALSE, test = "adf", ic = "aic")

summary(model)

predict\_10 = forecast(model, h = 10)

plot(predict\_10)

forecast\_result = data.frame(Year = c(1986:1995), predict\_10$mean, predict\_10$lower, predict\_10$upper)

forecast\_result = forecast\_result[,c(1,2,4,6)]

write.csv(forecast\_result, file = "Q1.csv")

### Q2

Q2 = read.xlsx("2019Spring\_Time\_Series\_Analysis\_final\_exam\_0620.xlsx", sheet = 2)

data = ts(Q2)

model = auto.arima(data, seasonal = TRUE, test = "adf", ic = "aic")

summary(model)

forecast\_result = forecast(model, h = 12, level = 99)

plot(forecast\_result)

forecast = data.frame(Index = c(1:12), forecast\_result$mean, forecast\_result$lower, forecast\_result$upper)

write.csv(forecast, file = "Q2.csv")

### Q3

Q3 = read.xlsx("2019Spring\_Time\_Series\_Analysis\_final\_exam\_0620.xlsx", sheet = 3)

library(TSA)

ts.xt<-ts(Q3$Ad)

lag3.x<-lag(ts.xt,-3)

ts.yt<-ts(Q3$sales)

dat3<-cbind(ts.xt,lag3.x,ts.yt)

dimnames(dat3)[[2]]<-c("xt","lag3x","yt")

data2<-na.omit(as.data.frame(dat3))

visc.tf <- arimax(data2$yt, order=c(1,0,0), xtransf=data.frame(data2$lag3x),

transfer=list(c(2,0)), include.mean = FALSE)

visc.tf

model = as.vector(fitted(visc.tf))

forecast\_result = forecast(model, h = 12, level = 95)

plot(forecast\_result)

forecast = data.frame(Index = c(1:12), forecast\_result$lower, forecast\_result$mean, forecast\_result$upper)

write.csv(forecast, file = "Q3.csv")

### Q4

Q4 = read.xlsx("2019Spring\_Time\_Series\_Analysis\_final\_exam\_0620.xlsx", sheet = 4)

prevention = Q4[1:129,]

pre\_data = ts(prevention)

model1 = auto.arima(pre\_data, seasonal = FALSE, test = "adf", ic = "aic")

summary(model1)

T <- nrow(Q4)

St <- c(rep(0,129),rep(1,(T-129)))

sales.tf<-arimax(diff(ts(Q4)), order=c(0,0,1), xtransf= St[2:T],

transfer=list(c(0,0)), include.mean = FALSE)

sales.tf

model = as.vector(fitted(sales.tf))

forecast\_result = forecast(model, h = 12, level = 95)

plot(forecast\_result)

forecast = data.frame(Index = c(1:12), forecast\_result$lower, forecast\_result$mean, forecast\_result$upper)

write.csv(forecast, file = "Q4.csv")

