Midterm Exam

FA-542

Due: October 27, 2023 at 10:00am

Problem 1 (20pt)
Let p_t be the log price of an asset at time t. Assume that the log price follows the model

$$p_t = 0.001 + p_{t-1} + a_t, \quad a_t \sim N(0, 0.16)$$

where $N(\mu, \sigma^2)$ denotes normal distribution with mean μ and variance σ^2. Assume further that $p_{200} = 4.551$.

a) Compute the 95% interval forecast for p_{201} at the forecast origin $t = 200$.

b) Compute the 2-step ahead point forecast and its standard error for p_{202} at the forecast origin $t = 200$.

c) What is the 100-step ahead forecast for p_{300} at the forecast origin $t = 200$?

Problem 2 (20pt)
Suppose that the quarterly growth rates rt of an economy follows the model

$$r_t = 0.006 + 0.168 r_{t-1} + 0.338 r_{t-2} - 0.189 r_{t-3} + a_t, \quad a_t \sim N(0, 0.0016)$$

a) What is the expected growth rate of r_t?

b) Does the model imply existence of business cycles? Why?

c) What is the average length of business cycles of the economy, if any?

Problem 3 (20pt)
The quarterly gross domestic product implicit price deflator is often used as a measure of inflation. The file **q-gdpdef.txt** contains the data for the United States from the first quarter of 1947 to the last quarter of 2008. Data format is year, month, day, and deflator. The data are seasonally adjusted and equal to 100 for year 2000.

a) Build an ARIMA model for the series and check the validity of the fitted model.

b) Use the fitted model to predict the inflation for each quarter of 2009.

Problem 4 (20pt)
You can use the quantmod package in R to obtain financial data.

a) Download daily price data for January 1, 2018 through October 23, 2023 of McDonald's Corp (MCD) stock from Yahoo Finance.

b) Build a time series model for this data.

c) Evaluate its performance. Justify your choices.

Problem 5 (20pt)
Consider the monthly U.S. unemployment rates from January 1947 to March 2016. Due to strong serial dependence, we analyze the differenced series $x_t = r_t - r_{t-1}$, where r_t is the seasonally adjusted unemployment rate. Answer the following questions, using the R output listed below the questions. Note: A fitted ARIMA model should include residual variance.

a) The auto.arima command in R specifies an ARIMA(2,0,2) model for x_t. The fitted model is referred to as **m1** in the output. Write down the fitted model.

b) Model checking shows two large outliers. An ARIMA(2,0,2) model with two outliers are then specified, **m3**. Write down the fitted model.

c) Model checking shows some serial correlations at lags 12 and 24. A seasonal model is then employed and called **m4**. Write down the fitted model.

d) The outliers remain in the seasonal model. Therefore, a refined model is used and called **m5**. Write down the fitted model.

e) Based on the model checking statistics provided, are there serial correlations in the residuals of model **m5**? Why?

f) Among models **m1**, **m3**, **m4**, and **m5**, which model is preferred under the in-sample fit? Why?

g) If root mean squares of forecast errors are used in out-of-sample prediction, which model is preferred? Why?

h) If mean absolute forecast errors are used in out-of-sample comparison, which model is selected?

i) Consider models **m1** and **m3**. State the impact of outliers on in-sample fitting.

j) Again, consider models **m1** and **m3**. State the impact of outliers on out-of-sample predictions.

R output: Problem 5

```
> UNRATE <- read.table("UNRATE-1.txt",header=T)
> rate <- as.numeric(UNRATE[,4])
> xt <- diff(rate)
> require(forecast)
> auto.arima(xt)
Series: xt
ARIMA(2,0,2) with zero mean
```

```
Coefficients:
         ar1      ar2      ma1     ma2
      1.6546  -0.7753  -1.6288  0.8440
s.e.  0.0427   0.0468   0.0420  0.0477

sigma^2 = 0.03857:  log likelihood = 172.36
AIC=-334.71   AICc=-334.64   BIC=-311.18
>
> m1 <- arima(xt,order=c(2,0,2),include.mean=F)
> m1

Call:
arima(x = xt, order = c(2, 0, 2), include.mean = F)

Coefficients:
         ar1      ar2      ma1     ma2
      1.6546  -0.7753  -1.6288  0.8440
s.e.  0.0427   0.0468   0.0420  0.0477

sigma^2 estimated as 0.03838:  log likelihood = 172.36,  aic = -336.71
>
> which.min(m1$residuals)
[1] 22
>
> i22 <- rep(0,818)
> i22[22]=1
>
> m2 <- arima(xt,order=c(2,0,2),xreg=i22,include.mean=F)
> m2

Call:
arima(x = xt, order = c(2, 0, 2), xreg = i22, include.mean = F)

Coefficients:
         ar1      ar2      ma1     ma2     xreg
      1.6953  -0.7965  -1.6286  0.8164  -1.5038
s.e.  0.0454   0.0477   0.0484  0.0509   0.1837

sigma^2 estimated as 0.03545:  log likelihood = 204.92,  aic = -399.84
>
> which.max(m2$residuals)
[1] 21
>
> i21 <- rep(0,818)
> i21[21]=1
> out <- cbind(i22,i21)
> m3 <- arima(xt,order=c(2,0,2),xreg=out,include.mean=F)
> m3

Call:
arima(x = xt, order = c(2, 0, 2), xreg = out, include.mean = F)

Coefficients:
         ar1      ar2      ma1     ma2      i22     i21
      1.6901  -0.7909  -1.6128  0.8014  -1.5302  1.1472
s.e.  0.0466   0.0504   0.0534  0.0592   0.1755  0.1757

sigma^2 estimated as 0.03368:  log likelihood = 225.86,  aic = -439.72
```

```
>
> Box.test(m3$residuals,lag=12,type='Ljung')

        Box-Ljung test

data:  m3$residuals
X-squared = 31.83, df = 12, p-value = 0.00147

>
> m4 <- arima(xt,order=c(2,0,2),seasonal=list(order=c(1,0,1),period=12),include.mean=F)
> m4

Call:
arima(x = xt, order = c(2, 0, 2), seasonal = list(order = c(1, 0, 1), period = 12),
    include.mean = F)

Coefficients:
         ar1      ar2      ma1     ma2    sar1     sma1
      1.2357  -0.3608  -1.2354  0.5151  0.5542  -0.8220
s.e.  0.2413   0.2221   0.2241  0.1702  0.0662   0.0473

sigma^2 estimated as 0.03538:  log likelihood = 204.21,  aic = -396.43
>
> m5 <- arima(xt,order=c(2,0,2),seasonal=list(order=c(1,0,1),period=12),include.mean=F,xreg=ou
> m5

Call:
arima(x = xt, order = c(2, 0, 2), seasonal = list(order = c(1, 0, 1), period = 12),
    xreg = out, include.mean = F)

Coefficients:
         ar1      ar2      ma1     ma2    sar1     sma1      i22     i21
      1.5743  -0.6591  -1.4869  0.6720  0.5488  -0.8208  -1.4762  1.1441
s.e.  0.1159   0.1110   0.1111  0.0913  0.0659   0.0448   0.1620  0.1616

sigma^2 estimated as 0.03062:  log likelihood = 263.2,  aic = -510.4
>
> Box.test(m5$residuals,lag=24,type='Ljung')

        Box-Ljung test

data:  m5$residuals
X-squared = 27.826, df = 24, p-value = 0.2674

>
> source("backtest.R")
> backtest(m1,xt,750,include.mean=F)
[1] "RMSE of out-of-sample forecasts"
[1] 0.1621524
[1] "Mean absolute error of out-of-sample forecasts"
[1] 0.1242145
> backtest(m3,xt,750,include.mean=F,xre=out)
[1] "RMSE of out-of-sample forecasts"
[1] 0.163189
[1] "Mean absolute error of out-of-sample forecasts"
[1] 0.1237345
> backtest(m4,xt,750,include.mean=F)
[1] "RMSE of out-of-sample forecasts"
[1] 0.1499355
```

```
[1] "Mean absolute error of out-of-sample forecasts"
[1] 0.1164277
> backtest(m5,xt,750,include.mean=F,xre=out)
[1] "RMSE of out-of-sample forecasts"
[1] 0.1493882
[1] "Mean absolute error of out-of-sample forecasts"
[1] 0.1164356
```