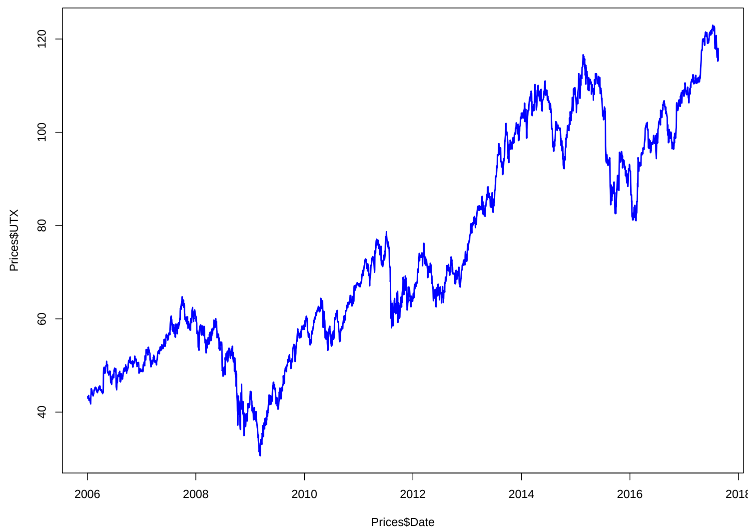
ORIE 4630: Spring Term 2019
Homework #4 Solutions

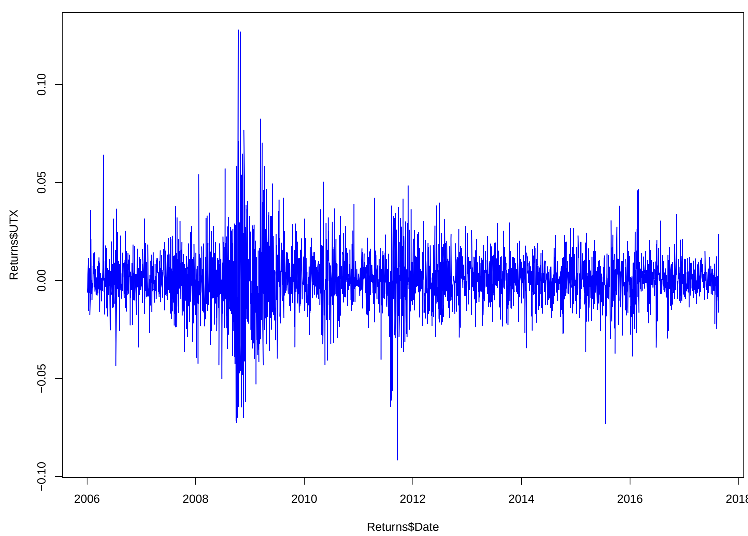
Question 1. [10 points]

Output from line 5:



This plot of the log prices shows non-stationarity. The mean is not constant, nor is the variance constant.

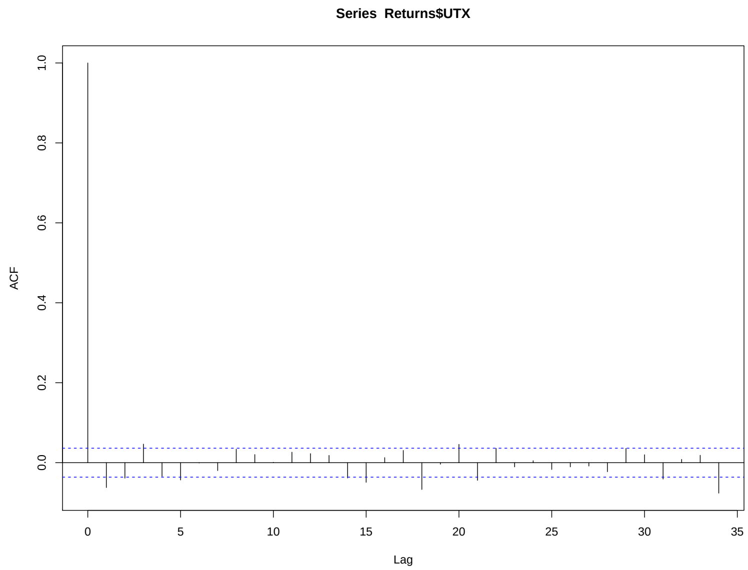
Output from line 6:



The plot of the log gross returns shows non-stationarity. Although the mean appears to be constant, there is volatility clustering, particularly in the years 2008-2009, so the variance is not constant.

Question 2. [5 points]

Output from line 7:



The plot of the sample ACF shows evidence serial correlation; for example, it appears from the plot that $\rho(1)$ is non-zero, since $\hat{\rho}(1)$ is beyond the error bars.

Question 3. [5 points]

Output from line 8:

```
> Box.test(Returns$UTX, lag=12, type="Ljung-Box")

Box-Ljung test

data:  Returns$UTX
X-squared = 40.228, df = 12, p-value = 6.589e-05
```

The test does show evidence of serial correlation. Since the p-value is very small , there is extremely strong evidence against the null hypothesis $H_0 : \rho(1) = \cdots = \rho(12) = 0$ in favor of the alternative hypothesis H_A : at least one of $\rho(1), \ldots, \rho(12)$ is non-zero.

Question 4. [15 points]

Output from line 14:

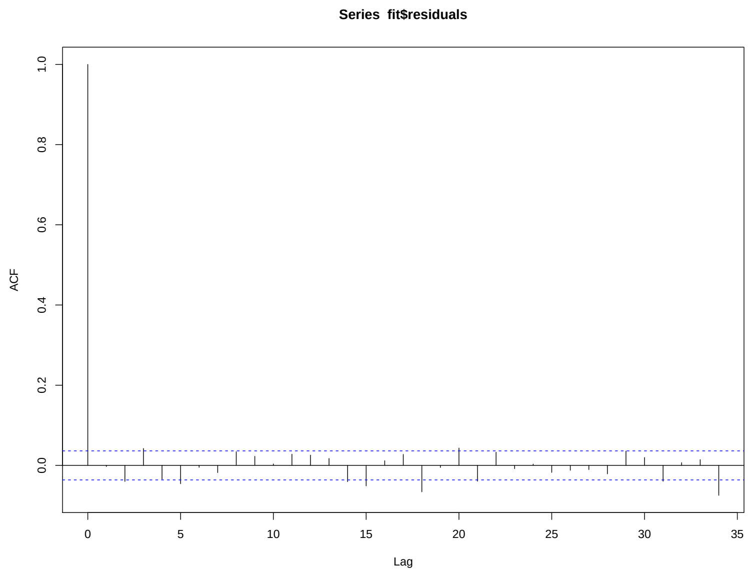
```
> fit

Call:
arima(x = Returns$UTX, order = c(1, 0, 0))
```

```
Coefficients:
          ar1  intercept
      -0.0622      3e-04
s.e.   0.0184      3e-04

sigma^2 estimated as 0.000215:  log likelihood = 8206.15,  aic = -16406.3
```

Output from line 15:



The plot of the sample ACF based on the residuals from the AR(1) model shows evidence of serial correlation; for example, it appears from the plot that both $\rho(2)$ and $\rho(3)$ are non-zero, since both $\hat{\rho}(2)$ and $\hat{\rho}(3)$ are beyond the standard error bars.

Output from line 16:

```
> Box.test(fit$residuals, lag=7, type="Ljung-Box", fitdf=1)

Box-Ljung test

data:  fit$residuals
X-squared = 20.529, df = 6, p-value = 0.002228
```

The test does show evidence of serial correlation. Since the p-value 0.002228 is small, there is strong evidence against the null hypothesis $H_0 : \rho(1) = \cdots = \rho(7) = 0$ in favor of the alternative hypothesis H_A : at least one of $\rho(1), \ldots, \rho(7)$ is non-zero.

If the AR(1) model was appropriate, then the disturbances in the model, i.e., the ϵ_ts, should be a white noise process. The residuals from the model, i.e., the $\hat{\epsilon}_t$s, which estimate the ϵ_ts, show evidence of autocorrelation. Hence, there is evidence from the plot and the test that the ϵ_ts are not a white noise process, so the returns appear not to follow an AR(1) model.

Question 5. [15 points]

Output from line 32:

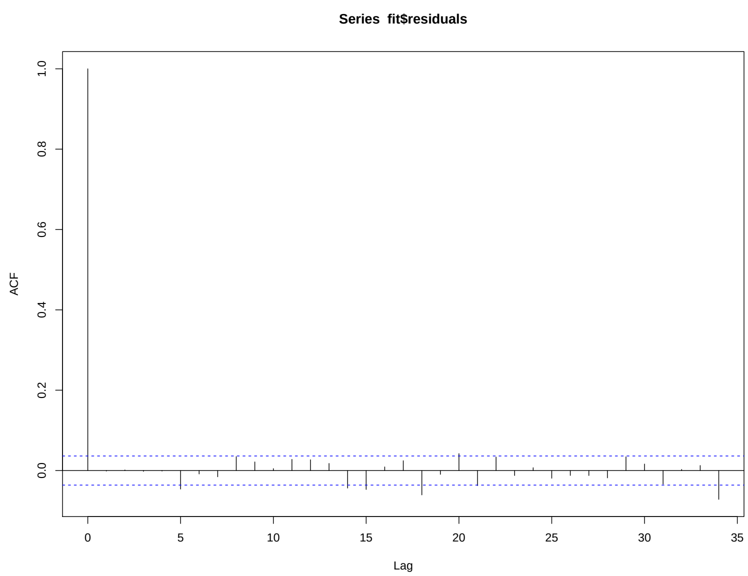
```
> fit

Call:
arima(x = Returns$UTX, order = c(4, 0, 0))

Coefficients:
          ar1      ar2     ar3      ar4  intercept
      -0.0619  -0.0413  0.0397  -0.0303      3e-04
s.e.   0.0185   0.0185  0.0185   0.0185      2e-04

sigma^2 estimated as 0.000214:  log likelihood = 8212.7,  aic = -16413.41
```

Output from line 33:



The plot of the sample ACF based on the residuals from the AR(4) model shows evidence of serial correlation; it appears from the plot that $\rho(5)$ is non-zero, since $\hat{\rho}(5)$ is beyond the error bars..

Output from line 34:

```
> Box.test(fit$residuals, lag=15, type="Ljung-Box", fitdf=4)

Box-Ljung test

data:  fit$residuals
X-squared = 29.628, df = 11, p-value = 0.001812
```

The test shows evidence of serial correlation. Since the p-value 0.001812 is small, there is evidence against the null hypothesis $H_0 : \rho(1) = \cdots = \rho(15) = 0$ in favor of the alternative hypothesis H_A : at least one of $\rho(1), \ldots, \rho(15)$ is non-zero.

If the AR(4) model was appropriate, then the disturbances in the model, i.e., the ϵ_ts, should be a white noise process. The residuals from the model, i.e., the $\hat{\epsilon}_t$s, which estimate the ϵ_ts, show evidence of autocorrelation. Hence, there is evidence from the plot and the test that the ϵ_ts are not a white noise process; the AR(4) model for the returns is not supported.

Question 6. [15 points]

Output from line 53:

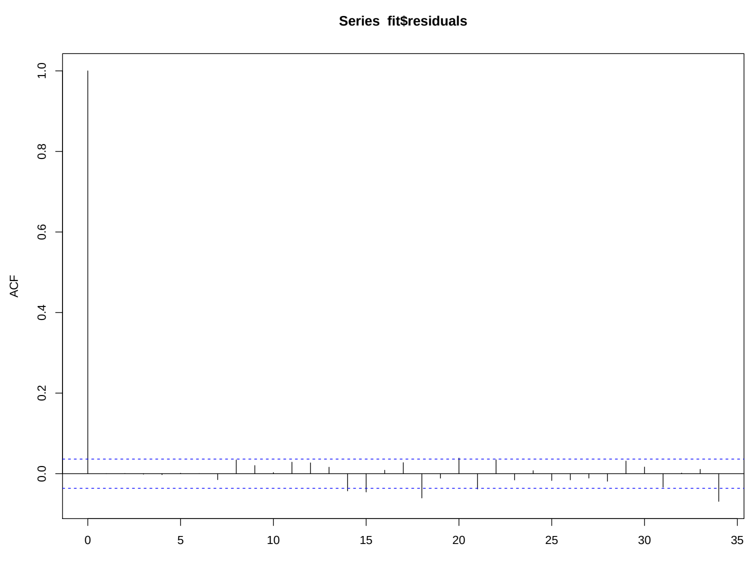
```
> fit

Call:
arima(x = Returns$UTX, order = c(0, 0, 5))

Coefficients:
          ma1      ma2     ma3      ma4      ma5  intercept
      -0.0636  -0.0367  0.0456  -0.0322  -0.0467      3e-04
s.e.   0.0185   0.0185  0.0185   0.0179   0.0186      2e-04

sigma^2 estimated as 0.0002135:  log likelihood = 8215.85,  aic = -16417.7
```

Output from line 54:



The plot of the sample ACF based on the residuals from the AR(5) model shows no evidence of serial correlation; none of $\hat{\rho}(1), \ldots, \hat{\rho}(12)$ are beyond the error bars.

Output from line 55:

```
> Box.test(fit$residuals, lag=13, type="Ljung-Box", fitdf=5)

Box-Ljung test

data:  fit$residuals
X-squared = 10.667, df = 8, p-value = 0.2213
```

The test shows no evidence of serial correlation. Since the p-value 0.2213 is large, there is no evidence against the null hypothesis $H_0 : \rho(1) = \cdots = \rho(13) = 0$ in favor of the alternative hypothesis H_A : at least one of $\rho(1), \ldots, \rho(13)$ is non-zero.

If the MA(5) model was appropriate, then the disturbances in the model, i.e., the ϵ_ts, should be a white noise process. The residuals from the model, i.e., the $\hat{\epsilon}_t$s, which estimate the ϵ_ts, show no evidence of autocorrelation. Hence, there is no evidence from the plot and the test that the ϵ_ts are not a white noise process, so the MA(5) model for the returns is supported.

Question 7. [**5 points**]
The AR(4) model has AIC$=-16413.41$, while the MA(5) model has AIC$=-16417.7$. A smaller value of AIC indicates better fit, so the MA(5) model would be preferred.

Question 8. [**10 points**]
Output from line 64:

```
> auto.arima(Returns$UTX, max.p=20, max.q=20, d=0, ic="aic")
Series: Returns$UTX
ARIMA(2,0,4) with zero mean

Coefficients:
         ar1      ar2      ma1     ma2     ma3      ma4
      0.3626  -0.8254  -0.4254  0.8118  0.0010  -0.0572
s.e.  0.0746   0.0809   0.0764  0.0839  0.0242   0.0218

sigma^2 estimated as 0.000214:  log likelihood=8215.72
AIC=-16417.44   AICc=-16417.41   BIC=-16375.57
```

The model identified by the `auto.arima()` function is the ARMA(2,2) model having zero mean. The ARMA(2,4) model has AIC$=-16417.44$, while the MA(5) model has AIC$=-16417.7$. Thus, based on the AIC information criterion, preference would be for the MA(5) model.

9. [10 points]

Output from line 68:

```
> fc
     Point Forecast    Lo 75    Hi 75    Lo 90    Hi 90
2929       115.6406 114.5621 116.7191 114.0984 117.1827
2930       115.5484 114.0454 117.0514 113.3992 117.6975
2931       115.5451 113.7246 117.3656 112.9420 118.1481
2932       115.5949 113.4883 117.7016 112.5827 118.6072
2933       115.5541 113.2038 117.9044 112.1934 118.9147
2934       115.5635 112.9931 118.1339 111.8881 119.2388
2935       115.5763 112.7993 118.3533 111.6056 119.5471
```

The point prediction for period $t = 2932$ is \$115.5949; the 75% prediction interval is $(112.9931, 118.1339)$.

10. [10 points]

Output from line 71: