Module 2: Time series concepts HW

(The attached PDF file has better formatting.)

*Homework assignment: equally weighted moving average*

This homework assignment uses the material on pages 14-15 (“A moving average”).

Let $Y_t = 1/5 \times (\epsilon_t + \epsilon_{t-1} + \epsilon_{t-2} + \epsilon_{t-3} + \epsilon_{t-4})$ and $\sigma^2_e = 100$.

A. What is $\gamma_{t,t}$, the variance of $Y_t$?
B. What is $\gamma_{t,t-3}$, the covariance of $Y_t$ and $Y_{t-3}$?

Write out the derivations in the format shown on page 15.

Module 3: Trends HW

(The attached PDF file has better formatting.)

*Homework assignment: MA(1) Process: Variance of mean*

Five MA(1) processes with 50 observations are listed below. The variance of $\epsilon_t$ is 1.

A. For each process, what is the variance of $\overline{y}$, the average of the Y observations?
B. How does the pattern of the first time series differ from that of the last time series?
C. Explain intuitively why oscillating patterns have lower variances of their means.

1. $Y_t = \mu + e_t + e_{t-1}$
2. $Y_t = \mu + e_t + ½\, e_{t-1}$
3. $Y_t = \mu + e_t$
4. $Y_t = \mu + e_t – ½\, e_{t-1}$
5. $Y_t = \mu + e_t – e_{t-1}$

(See page 50 of the Cryer and Chan text, Exercise 3.2)

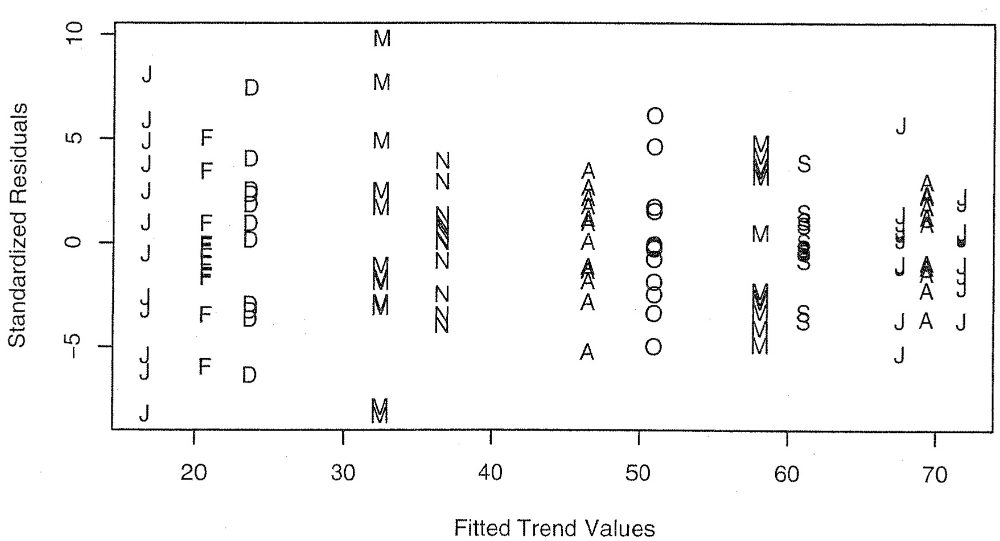
TS Module 4: Regression methods HW

(The attached PDF file has better formatting.)

*Homework assignment: Residuals*

Cryer and Chan show the following Exhibit 3.10.

**Exhibit 3.10 Standardized Residuals versus Fitted Values for the Temperature Seasonal Means Model**




A. Temperatures vary by month. Why aren't the residuals high in July and August and low in January and February?
B. An actuary looking at this exhibit concludes that temperatures vary more in December than in November. Why is this not a reasonable conclusion?
C. A toy store has high sales in December and low sales in February. Would you expect the residuals to have a higher variance in December or February?
D. Why does the reasoning for toy sales not apply to daily temperature?

For Part C, suppose expected sales are $500,000 in December and $50,000 in February. Think of December as 10 February's packed into one month. What is the ratio of the variances if the 10 February's are independent? What is the ratio if the 10 February's are perfectly correlated?

For Part D, the daily temperature is in arbitrary units. A day in August is not like two days or ten days of February packed together.

TS Module 5: Stationary processes HW

(The attached PDF file has better formatting.)

*Homework assignment: general linear process*

A time series has the form $Y_t = \epsilon_t + \phi \times \epsilon_{t-1} - \phi^2 \times \epsilon_{t-2} + \phi^3 \times \epsilon_{t-3} - \ldots$

*The plus and minus signs alternate.* $\phi = 0.2$ and $\sigma^2_e = 9$.

A. What is $\gamma_0$, the variance of $Y_t$? Show the derivation.
B. What is $\gamma_1$, the covariance of $Y_t$ and $Y_{t-1}$? Show the derivation.
C. What is $\rho_2$, the correlation of $Y_t$ and $Y_{t-2}$? Show the derivation.

(Show the algebra for the derivations. One or two lines is sufficient for each part.)

TS Module 6: Stationary autoregressive processes HW

(The attached PDF file has better formatting.)

*Homework assignment: AR(2) process*

An AR(2) process has $\phi_1$ = 0.2 or –0.2 and $\phi_2$ = ranging

- from 0.2 to 0.7 in steps of 0.1.
- from –0.2 to –0.9 in steps of –0.1.

Complete the table below, showing $\rho_1$ and $\rho_2$, the autocorrelations of lags 1 and 2. Use an Excel spreadsheet (or other software) and form the table by coding the cell formulas. Print the Excel spreadsheet and send it in as your homework assignment.

| $\phi_1$ | $\phi_2$ | $\rho_1$ | $\rho_2$ | $\phi_1$ | $\phi_2$ | $\rho_1$ | $\rho_2$ |
|---|---|---|---|---|---|---|---|
| 0.2 | 0.2 | | | -0.2 | 0.2 | | |
| 0.2 | 0.3 | | | -0.2 | 0.3 | | |
| 0.2 | 0.4 | | | -0.2 | 0.4 | | |
| 0.2 | 0.5 | | | -0.2 | 0.5 | | |
| 0.2 | 0.6 | | | -0.2 | 0.6 | | |
| 0.2 | 0.7 | | | -0.2 | 0.7 | | |
| 0.2 | -0.2 | | | -0.2 | -0.2 | | |
| 0.2 | -0.3 | | | -0.2 | -0.3 | | |
| 0.2 | -0.4 | | | -0.2 | -0.4 | | |
| 0.2 | -0.5 | | | -0.2 | -0.5 | | |
| 0.2 | -0.6 | | | -0.2 | -0.6 | | |
| 0.2 | -0.7 | | | -0.2 | -0.7 | | |
| 0.2 | -0.8 | | | -0.2 | -0.8 | | |
| 0.2 | -0.9 | | | -0.2 | -0.9 | | |

TS Module 7: stationary mixed processes HW

(The attached PDF file has better formatting.)

*Homework assignment: mixed autoregressive moving average process*

An ARMA(1,1) process has $\sigma^2 = 1$, $\theta_1 = 0.4$, and $\phi_1 = 0.6$.

A. What is the value of $\gamma_0$?
B. What is the value of $\gamma_1$?
C. What is the value of $\rho_1$?
D. What is the value of $\rho_2$?

TS Module 8: Non-stationary time series basics HW

(The attached PDF file has better formatting.)

*Homework assignment: Stationarity through differencing and logarithms*

Automobile liability claim severities have a geometric trend of +8% per annum.
The average claim severity in year t is the average claim severity in year t-1, plus or minus a random error term.

A. Is the time series of average claim severities stationary?
B. Is the first difference of this time series stationary?
C. Is the second difference of this time series stationary?
D. Is the logarithm of this time series stationary?
E. What transformation makes the time series stationary?

TS Module 9: Non-stationary ARIMA time series HW

(The attached PDF file has better formatting.)

*Homework assignment: Non-stationary autoregressive process*

A time series $Y_t = \beta \times Y_{t-1} + \epsilon_t$ has $\sigma_\varepsilon^2$ = 3, where *k* is a constant. (The textbook has $\beta = 3$.)

A. What is the variance of $Y_t$ as a function of β and *t*?
B. What is $\rho(y_t, y_{t-k})$ as a function of β, k, and *t*?

See equations 5.1.4 and 5.1.5 on page 89. Show the derivations for an arbitrary β.

TS Module 10: autocorrelation functions HW

(The attached PDF file has better formatting.)

*Homework assignment: Sample autocorrelations*

A time series has ten elements: {10, 8, 9, 11, 13, 12, 10, 8, 7, 12}.

A. What is the sample autocorrelation of lag 1?
B. What is the sample autocorrelation of lag 2?
C. What is the sample autocorrelation of lag 3?

Show the derivations with a table like the one below. Remember to use the proper number of terms in the numerator, depending on the lag.

| *Entry* | *Entry* | *Deviation* | *Deviation Squared* | *Cross Product Lag1* | *Cross Product Lag2* | *Cross Product Lag3* |
|---|---|---|---|---|---|---|
| 1 | 10 | | | | | |
| 2 | 8 | | | | | |
| 3 | 9 | | | | | |
| 4 | 11 | | | | | |
| 5 | 13 | | | | | |
| 6 | 12 | | | | | |
| 7 | 10 | | | | | |
| 8 | 8 | | | | | |
| 9 | 7 | | | | | |
| 10 | 12 | | | | | |
| *Avg/tot* | | | | | | |
| *Autocorr* | | | | | | |

TS Module 11: simulated and actual time series HW

(The attached PDF file has better formatting)

*Homework assignment: Partial autocorrelations*

[Partial autocorrelations are covered in Module 10, along with sample autocorrelations.]

- A stationary ARMA process has $\rho_2 = 0.20$.
- $\rho_1$ ranges from 0.2 to 0.7 in units of 0.1.

A. Graph the partial autocorrelation of lag 2 ($\phi_{22}$) as a function of $\rho_1$.
B. Explain why the partial autocorrelation is positive for low $\rho_1$ and negative for high $\rho_1$.

TS Module 12: Parameter estimation method of moments HW

(The attached PDF file has better formatting.)

*Homework assignment: Method of moments*

An ARMA(1,1) process has $r_1 = -0.25$ and $r_2 = -0.125$.

A. What is $\phi$, the autoregressive parameter?
B. What is $\theta$, the moving average parameter?

TS Module 13: Parameter estimation least squares HW

(The attached PDF file has better formatting.)

*Homework assignment: Estimating parameters by regression*

An AR(1) process has the following values:

0.44 1.05 0.62 0.72 1.08 1.24 1.42 1.35 1.50

A. Estimate the parameter $\phi$ by regression analysis.
B. What are 95% confidence intervals for the value of $\phi$?
C. You initially believed that $\phi$ is 50%. Should you reject this assumption?

The time series course does not teach regression analysis. You are assumed to know how to run a regression analysis, and you must run regressions for the student project.

Use the Excel *REGRESSION* add-in. The 95% confidence interval is the estimated $\beta$ ± the *t*-value × the standard error of $\beta$. The *t*-value depends on the number of observations. Excel has a built-in function giving the *t*-value for a sample of N observations.

TS Module 14: Model diagnostics HW

(The **two** attached PDF files have better formatting.)

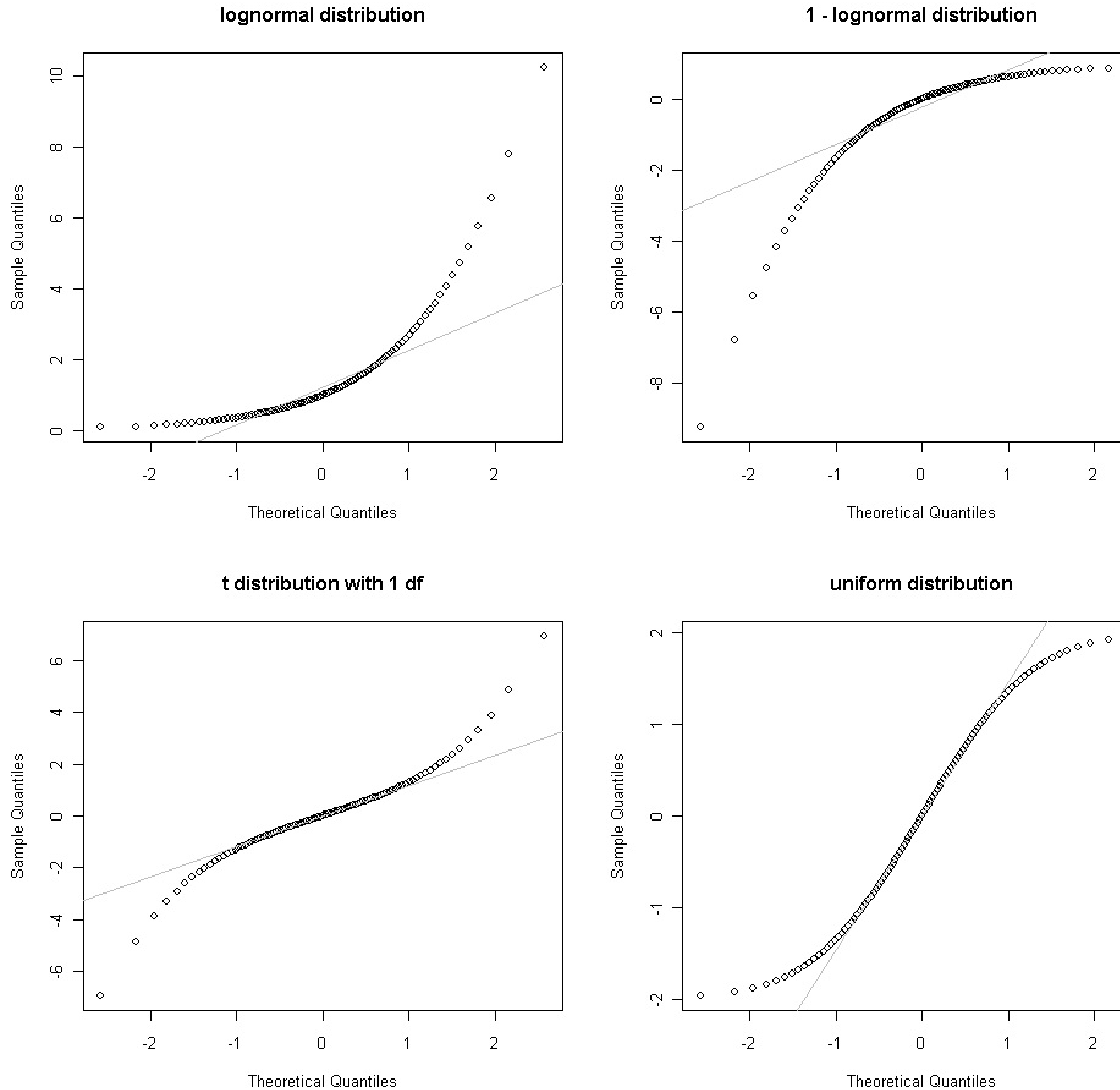
*Homework assignment: quantile comparison (q-q) plots*

Quantile comparison plots are explained in the regression analysis on-line course, and they are used also in the time series course. If this homework assignment is difficult, review the module of quantile comparison plots in the regression analysis course. (Module 3, "Quantile comparison plots," on pages 34-37, especially Figures 3.8 and 3.9; you can search for *quantile comparison plots* or *q-q plots* on the internet to see several examples.)

The four figures below show quantile comparison plots for four distributions. For each one

A. Is the distribution symmetric, right skewed, or left skewed? Explain how the quantile comparison plot shows this.
B. If the distribution is symmetric, is it heavy tailed or thin tailed? Explain how the quantile comparison plot shows this.

Quantile comparison plots are a useful tool for actuarial work, so it is worth knowing how to use them. For your student project, you may test if the residuals of an ARIMA process are normally distributed by forming a quantile comparison plot.


lognormal distribution
1 - lognormal distribution
t distribution with 1 df
uniform distribution
Sample Quantiles
Theoretical Quantiles

TS Module 14: Model diagnostics HW

(The attached PDF file has better formatting.)

*Homework assignment: quantile comparison (q-q) plots*

Quantile comparison plots are explained in the regression analysis on-line course, and they are used also in the time series course. If this homework assignment is difficult, review the module of quantile comparison plots in the regression analysis course. (Module 3, "Quantile comparison plots," on pages 34-37, especially Figures 3.8 and 3.9; you can search for *quantile comparison plots* or *q-q plots* on the internet to see several examples.)

The four figures below show quantile comparison plots for four distributions. For each one

A. Is the distribution symmetric, right skewed, or left skewed? Explain how the quantile comparison plot shows this.
B. If the distribution is symmetric, is it heavy tailed or thin tailed? Explain how the quantile comparison plot shows this.

Quantile comparison plots are a useful tool for actuarial work, so it is worth knowing how to use them. For your student project, you may test if the residuals of an ARIMA process are normally distributed by forming a quantile comparison plot.

TS Module 15: Forecasting basics HW

(The attached PDF file has better formatting.)

*Homework assignment: ARIMA(1,1,0) forecasts*

An ARIMA(1,1,0) process has 40 observations $y_t$, t = 1, 2, …, 40, with $y_{40}$ = 60 and $y_{39}$ = 50.

This time series is *not* stationary, but its *first differences* are a stationary AR(1) process.

The parameter $\theta_0$ of the stationary AR(1) time series of first differences is 5.

The 1 period ahead forecast $\hat{y}_{40}(1)$ is 60.

We determine the 2 period ahead forecast $\hat{y}_{40}(2)$.

A. What is the most recent value of the autoregressive model of first differences? Derive this value from the most recent two values of the ARIMA(1,1,0) process.
B. What is the one period ahead forecast of the first differences? Derive this value from the the one period ahead forecast of the ARIMA(1,1,0) process.
C. What is the parameter $\phi_1$ of the AR(1) process of first differences? Derive this parameter from the 1 period ahead forecast.
D. What is the two periods ahead forecast of the AR(1) process of first differences? Use the parameter of the AR(1) process.
E. What is the two periods ahead forecast of the ARIMA(1,1,0) process? Derive this from the two periods ahead forecast of the AR(1) process.

TS Module 16: ARIMA Forecasting HW

(The attached PDF file has better formatting.)

*Homework assignment: ARIMA(0,1,1) process*

The estimated (forecast) and actual values for Periods 48, 49, and 50 of an ARIMA(0,1,1) process are shown below.

Forecasts are *one period ahead* forecasts: $\hat{y}_{48}(1)$ for Period 49 and $\hat{y}_{49}(1)$ for Period 50.

| *Period* | *Forecast* | *Actual* |
|---|---|---|
| 48 | 70.5 | 71.5 |
| 49 | 72.0 | 74.0 |
| 50 | 73.0 | 74.8 |

- The estimated and actual values are for the ARIMA(0,1,1) time series.
- The values of $\mu$ and $\theta_1$ are for the ARMA model of first differences. (Cryer and Chan use $\theta$ for an ARMA process, not $\theta_1$.)

To solve the homework assignment, use the following steps:

- Determine the residual for each period for the ARIMA(0,1,1) model.
- These are also the residuals for the ARMA process of first differences.
- Determine the forecasts and actual values for the MA(1) process of first differences for the last two periods.
- Write equations for these forecasts in terms of the mean, the residual in the previous period, and $\theta_1$. Remember that $\theta_1$ is the *negative* of the moving average parameter.
- You have a pair of linear equations with two unknowns, $\mu$ and $\theta_1$.
- Solve for $\mu$ and $\theta_1$ and verify that these values give the forecasts in the table.
- Use the residual for Period 50 and the values of $\mu$ and $\theta_1$ to forecast Period 51.
- Derive the forecast for the original ARIMA time series for Period 51.

A. What is the mean $\mu$ of the ARMA model of first differences?
B. What is $\theta_1$ of the ARMA model of first differences?
C. What is the forecasted value of the ARIMA(0,1,1) process for Period 51?

Show the derivations of the parameters and the forecast.

TS Module 17: Forecasting bounds HW

(The attached PDF file has better formatting.)

*Homework assignment: Random walk with drift*

An insurer's capital follows a random walk with a drift of $10 million a month and a volatility of $40 million a month. The initial capital is $200 million.

A random walk is an ARIMA(0,1,0) process. The capital *changes* are a white noise process, with a mean $\mu$ of $10 million a month and a $\sigma$ of $40 million a month.

A. What is the distribution of capital after one month? (What is the type of distribution, such as normal, lognormal, uniform, or something else? Use the characteristics of a white noise process. What is the mean of the distribution after one month? Use the starting capital and the drift. What is the standard deviation after one month? The volatility is the standard deviation per unit of time, not the variance per unit of time. It is $\sigma$, not $\sigma^2$.)
B. What is the distribution of capital after six months? (The serial correlation is zero, so the capital changes are independent and the variances are additive. Derive $\sigma^2$ for one month from $\sigma$, add the $\sigma^2$'s for six months, and derive the $\sigma$ after six months.)
C. What is the distribution of capital after one year?
D. What are the probabilities of insolvency at the end of six months and one year? (You have a distribution with a mean $\mu$ and a standard deviation $\sigma$. Find the probability that a random draw from this distribution is less than zero. Use the cumulative distribution function of a standard normal distribution. Excel has a built-in function for this value.)
E. At what time in the future is the probability of insolvency greatest? (Write an equation for the probability as a function of (i) the mean of the distribution at time *t* and (ii) the standard deviation of the distribution at time *t*. To maximize this probability, set its first derivative to zero, and solve for *t*.)

TS Module 18: Forecast updates and weights HW

(The attached PDF file has better formatting.)

*Homework assignment: ARIMA(0,1,1) forecasts*

An ARIMA(0,1,1) model for a time series of 100 observations, $y_t$, t = 1, 2, …, 100, has $\theta_1 = 0.4$.

- The forecast of the next observation, $y_{101}$, is 25.
- The actual value of $y_{101}$ is 26.
- The forecast of the next observation, $y_{102}$, is 26.
- The actual value of $y_{102}$ is 26.

We continue to use the same ARIMA model. That is, we don't re-estimate the parameters with the additional data. We forecast $y_{103}$, the ARIMA value in the next period.

A. From the actual and forecasted values of $y_{101}$, derive the residual for the ARMA model of the first differences.
B. From the actual value of $y_{101}$ and the forecasted value of $y_{102}$, derive the forecasted value for Period 102 for the ARMA model of the first differences.
C. This forecasted value for Period 102 is a function of $\mu$, $\theta_1$, and the residual for Period 101. Derive the $\mu$ (mean) of the ARMA model of first differences.
D. From the actual and forecasted values of $y_{102}$, derive the residual for the ARMA model of the first differences for Period 102.
E. Using this residual, determine the forecasted first difference for the next period.
F. From the forecasted first difference, derive the forecasted value of the original time series.

The values of $\mu$ and $\theta_1$ are the coefficients of the ARMA process for the first differences. (Cryer and Chan use $\theta$ for an MA(1) process, not $\theta_1$.)

TS Module 19: Seasonal models basics HW

(The attached PDF file has better formatting.)

*Homework assignment: auto insurance seasonality*

This homework assignment applies ARIMA modeling to auto insurance policy counts. The same logic applies to other lines of business as well. The homework assignment has three parts, each adding a piece to the insurance scenario.

*Part A:* Seasonality

The population in State W, the number of drivers, and the number of cars are stable.

- An auto insurer sells twelve month policies in State W.
  - Its renewal rate is 90% on average.
  - The actual renewal rate varies with its prices and those of its competitors.

For Part A of the homework assignment, all insurers charge the same state-made rates.

- Random fluctuation determines how much new business each insurer writes.
- You model policies written each month (both new and renewal) with an ARIMA process.

A. How would you model renewal policies? Do you use a $\phi$ (autoregressive) parameter or a $\theta$ (moving average) parameter?
B. Is the process stationary?
C. If the renewal rate were 100%, would the process be stationary?

*Intuition:* If the insurer writes more policies in January 20X1, it writes more policies in January 20X2, January 20X3, and so forth. A random fluctuation dies our slowly. What is the autocorrelation function? Is this a stationary time series?

*Part B:* We add a free market to this exercise.

- The insurer competes in a free market.
- The insurer revises its rates each year, and its competitors revise their rates at other dates during the year.

New policies sold (new business production) depends on the insurer's relative rate level compared with its competitors. If its rates are lower (higher) than its competitors', it writes more (fewer) new policies.

- The insurer does not expect higher or lower rates than its competitors charge.
- At any time, its rates may be higher or lower than its competitors charge, so its market share may grow or shrink.

D. How would you change the model? Do you add an autoregressive or a moving average term? Note that rate changes occur once a year, so if the insurer has high (low) rates now, it will probably have high (low) rates next month.
E. Does the free market increase or decrease the variance of the process? (When firms charge different prices and revise their prices periodically, is market share more or less variable?)

*Part C:* The insurer revises rates if its policy count is higher or lower than expected.

- If the insurer writes more policies than expected, it is afraid that its rates are too low, and it files for a rate increase.
- If the insurer writes fewer policies than expected, it is afraid that its rates are too high, and it files for a rate decrease.

F. If this rate change occurs immediately, how should you change the model? Do you add an autoregressive or a moving average term? (In practice, rate filings take several months to be approved. Assume this rate change takes effect immediately.)
G. If this rate change has a one month lag, how should you change the model? (For this part of the homework assignment, assume the insurer compares its actual vs expected policy count at the end of each month and changes the rate beginning either the next day or one month later. Use whichever assumption you prefer.)

TS Module 20 Seasonal models advanced HW

(The attached PDF file has better formatting.)

*Homework assignment: auto sales*

New auto sales vary by month. New models are released in October, in time for holiday sales of November and December. Sales in January and February are low, since cold weather and the high sales in the previous months leave low demand for new cars.

*Part A:* Compare the seasonality of auto insurance in the homework assignment for the previous module with the seasonality of auto sales in this homework assignment. We can model seasonality by a seasonal autoregressive parameter or by seasonally adjusted data.

- Monthly temperature is seasonal. But if 20X8 has a particularly cold January, we don't expect January 20X9 to be colder than usual.
- Insurance renewals depend on the number of policies written one year ago (or six months ago). If January 20X8 has a higher than usual policy count, we expect a higher than usual policy count in January 20X9 as well.

1. How would you model the seasonality of auto insurance?
2. How would you model the seasonality of auto sales?

*Part B:* Cars are durable and expensive goods.

Cars last several years. If a consumer buys a new car in year X, his or her demand for new cars is low in year X+1 and rises steadily lover subsequent years.

Cars are expensive. In tough economic times (recessions, high unemployment), demand for new cars is weak. In prosperous years, built-up demand for new cars causes high sales.

GDP shows weak cycles. Durable goods (cars, housing) show strong cycles.

1. How would you model the long-term cycles of auto sales? Use an annual model, not a monthly model, so you can avoid the monthly seasonality.
2. Assume GDP follows a weak cycle:
   a. Consumers who have not bought a new car recently are more likely to buy one if the economy is strong.
   b. Consumers who have bought a new are recently are unlikely to buy a second one for several years.
   c. Explain how you might model this process.

This homework assignment does not have a single correct answer. Economists model the sales of durable goods many ways. This homework assignment asks you to demonstrate that you understand how autoregressive and moving average processes work.