Solutions to Exam 2

1(a). From the definition of the data generation process for $y_t$, it follows that

$$w_t = y_t - \theta y_{t-1} - \phi w_{t-1}$$  \hspace{1cm} (1)

Recursive back substitution in (1) yields

$$w_t = y_t - \theta y_{t-1} - \phi(y_{t-1} - \theta y_{t-2} - \phi w_{t-2})$$

$$= y_t - (\theta + \phi)y_{t-1} + \phi\theta y_{t-2} + \phi^2 w_{t-2}$$

$$= y_t - (\theta + \phi)y_{t-1} + \phi\theta y_{t-2} + \phi^2(y_{t-2} - \theta y_{t-3} - \phi w_{t-3})$$

$$= \sum_{i=0}^{3} \eta_{3,i} y_{t-i} - \phi^3 w_{t-3}$$  \hspace{1cm} (2)

where $\eta_{3,0} = 1$, $\eta_{3,i} = (\theta + \phi)(-1)^i \phi^{i-1}$ for $i = 1, 2$, and $\eta_{3,3} = (-1)^3 \phi^2 \theta$. This process can be continued to give

$$w_t = \sum_{i=0}^{t} \eta_{t,i} y_{t-i} + (-\phi)^t w_0$$  \hspace{1cm} (3)

where $\eta_{t,0} = 1$, $\eta_{t,i} = (-1)^i (\theta + \phi)(\phi)^{i-1}$ for $i = 1, 2, \ldots, t-1$ and $\eta_{t,t} = (-1)^t \phi^{t-1} \theta$.

1(b). To calculate $E[y_{t+s}|Y_t, w_0 = \bar{w}]$, we derive an equation that relates $y_{t+s}$ to $Y_t$. If we set

$$v_t = w_t + \phi w_{t-1}$$  \hspace{1cm} (4)

then the data generation process can be written as

$$y_t = \theta y_{t-1} + v_t$$  \hspace{1cm} (5)

and so

$$y_{t+s} = \theta y_{t+s-1} + v_{t+s}$$  \hspace{1cm} (6)

Repeated back substitution in (6) for $y$ yields

$$y_{t+s} = \theta(\theta y_{t+s-2} + v_{t+s-1}) + v_{t+s}$$

$$= \sum_{i=0}^{s-1} \theta^i v_{t+s-i} + \theta^s y_t$$  \hspace{1cm} (7)

Therefore, we have

$$E[y_{t+s}|Y_t, w_0 = \bar{w}] = \sum_{i=0}^{s-1} \theta^i E[v_{t+s-i}|Y_t, w_0 = \bar{w}] + \theta^s E[y_t|Y_t, w_0 = \bar{w}]$$  \hspace{1cm} (8)
Now \( y_t \in Y_t \) and so

\[
E[y_t | Y_t, u_0 = \bar{w}] = y_t
\]  
(9)

From (4), it follows that

\[
E[y_{t+s-i} | Y_t, u_0 = \bar{w}] = E[w_{t+s-i} | Y_t, u_0 = \bar{w}] + \phi E[w_{t+s-i-1} | Y_t, u_0 = \bar{w}]
\]  
(10)

From part (a), it follows that \( w_t \) can be calculated from \( Y_t \) and \( u_0 \), and so

\[
E[w_t | Y_t, u_0 = \bar{w}] = w_t(\bar{w})
\]  
(11)

However, since \( w_t \sim IN(0, \sigma^2_w) \), knowledge of \( w_t \) does not help us forecast \( w_{t+j} \) for \( j > 0 \) and so

\[
E[w_{t+s-i} | Y_t, u_0 = \bar{w}] = \phi w(t)(\bar{w}), \quad \text{for } i = s - 1
\]

\[
= 0, \quad \text{for } i = 0, 1, \ldots s - 2
\]  
(12)

(13)

Combining (8)-(13), we obtain:

\[
E[y_{t+s} | Y_t, u_0 = \bar{w}] = \theta^s y_t + \theta^{s-1} \phi w(t)(\bar{w})
\]  
(14)

1. (c) Since \( |\theta| < 1 \), it follows that \( \lim_{m \to \infty} \theta^m = 0 \) and so it follows from (7) that

\[
y_t = \lim_{m \to \infty} \sum_{i=0}^{m} \theta^i v_{t-i}
\]  
(15)

Therefore, we have

\[
E[y_t] = \lim_{m \to \infty} \sum_{i=0}^{m} \theta^i E[v_{t-i}]
\]

Since \( E[w_t] = 0 \), it follows from (4) that \( E[v_t] = 0 \) and so \( E[y_t] = 0 \), using (15). A comparison of the conditional and unconditional expectations indicates that they are not equal for finite \( s \) in general. However as \( s \to \infty \), \( \theta^s \to 0 \) because \( |\theta| < 1 \) and so \( E[y_{t+s} | Y_t, u_0 = \bar{w}] \to E[y_t] \).

2(a). The condition for stationarity of an AR(p) process is that the roots of the autoregressive polynomial lie outside the unit circle. In this example, the condition for the AR polynomial for \( x_t \) reduces to \( (1 - 0.5m) = 0 \) which implies a root of \( m = 2 \). Since \( w_t \sim IN(0, \sigma^2_w) \), \( x_t \) is a stationary process. However, \( v_t = \alpha_0 + \delta_0 t + x_t \) and so \( E[v_t] = \alpha_0 + \delta_0 t + E[x_t] = \alpha_0 + \delta_0 t \) because \( E[x_t] = 0 \). Clearly, the mean of \( v_t \) depends on \( t \) - assuming \( \delta_0 \neq 0 \) - and so \( v_t \) is not stationary.

2(b). Note that the moments of the error process do not depend on \( t \). Therefore, the condition for stationarity of the VAR is that the roots of \( |\Theta(m)| = 0 \) lie outside the unit circle where \( \Theta(m) \) is the autoregressive polynomial. In this example, the condition reduces to \( d(m) = 0 \) where

\[
d(m) = |I_2 - \Theta m|
\]

For the given value of \( \Theta \), \( d(m) = (1 - 0.5m)(1 - 0.5m) \) and so \( d(m) = 0 \) has the repeated root of two. Therefore, the condition for stationarity is satisfied.
3(a). Let \( \hat{E}[v_{1,t+s}|V_{1,t},V_{2,t}] \) denote the linear forecast of \( v_{1,t+s} \) given \( V_{1,t} = (v_{1,t}, v_{1,t-1}, \ldots, v_{1,1}) \) and \( V_{2,t} = (v_{2,t}, v_{2,t-1}, \ldots, v_{2,1}) \), and \( \hat{E}[v_{1,t+s}|V_{1,t}] \) denote the linear forecast of \( v_{1,t+s} \) given \( V_{1,t} \). Further define the mean square errors of these forecasts by \( MSE\{\hat{E}[v_{1,t+s}|V_{1,t},V_{2,t}]\} = E[(v_{1,t+s} - \hat{E}[v_{1,t+s}|V_{1,t},V_{2,t}])^2] \) and \( MSE\{\hat{E}[v_{1,t+s}|V_{1,t}]\} = E[(v_{1,t+s} - \hat{E}[v_{1,t+s}|V_{1,t}])^2] \). \( v_2 \) is said to not Granger cause \( v_1 \) if \( MSE\{\hat{E}[v_{1,t+s}|V_{1,t},V_{2,t}]\} = MSE\{\hat{E}[v_{1,t+s}|V_{1,t}]\} \) for all \( s > 0 \).

3(b) Let the VAR(2) representation be:

\[
v_t = \Theta_1 v_{t-1} + \Theta_2 v_{t-2} + \nu_t
\]

where \( v_t = (v_{1,t}, v_{2,t}) \). If \( v_2 \) does not Granger cause \( v_1 \) then the evolution of \( v_1 \) must not depend on \( v_2 \). Within the VAR(2) framework, this is equivalent to the restriction that \( \{\Theta_1\}_{1,2} = \{\Theta_2\}_{1,2} = 0 \) where \( \{\Theta_i\}_{i,j} \) is the \( i - j \)th element of \( \Theta_i \). Therefore, the null of Granger non-causality can tested against the alternative of Granger causality by testing:

\[
H_0 : \{\Theta_1\}_{1,2} = \{\Theta_2\}_{1,2} = 0 \iff v_2 \text{ does not Granger cause } v_1
\]

\[
H_1 : \{\Theta_j\}_{1,2} \neq 0, \text{ for at least one } j \iff v_2 \text{ does Granger cause } v_1
\]

This test can be performed using the Wald, LR or LM statistic.

3(c). Granger non-causality is related to the concept of strong exogeneity but not weak exogeneity. To be more specific, suppose that the data is \( \{v_t; t = 1,2 \ldots T\} \) and \( v_t = (y_t', x_t') \). Also let \( Y_t = (y_t, y_{t-1}, \ldots y_1) \) and \( X_t = (x_t, x_{t-1}, \ldots x_1) \). Let the joint pdf of \( v_t \) given \( Y_t \) and \( X_t \) be \( p(v_t|X_{t-1}, Y_{t-1}; \lambda, \mu) \) where \( \lambda \) and \( \mu \) are parameter vectors. \( x_t \) is said to strongly exogenous for estimation of \( \lambda \) if:

\[
p(v_t|X_{t-1}, Y_{t-1}; \lambda, \mu) = p(y_t|x_t, X_{t-1}, Y_{t-1}; \lambda)p(x_t|X_{t-1}; \mu)
\]

where \( p(y_t|x_t, X_t, Y_t; \lambda) \) is the conditional pdf for \( y_t \) given \( x_t, X_t \) and \( Y_t \), and \( p(x_t|X_{t-1}; \mu) \) is the conditional pdf of \( x_t \) given \( X_{t-1} \). Note that this factorization includes the restriction that the conditional distribution of \( x_t \) given \( X_{t-1} \) and \( Y_{t-1} \) does not depend on \( Y_{t-1} \) which, using the definitions in Engle, Hendry and Richard (1983), means that \( y \) does not Granger cause \( x \).