Time Series: M.Phil. Econometrics, Hilary Term 2014
Solutions for week 3 problem set

Oleg I. Kitov

February 5, 2014

Useful Concepts

Definition 1. (Stationarity) The concept of stationarity is related to the full joint distribution of a process. A process $(X_t)_{t\in\mathbf{N}}$ is stationary if for all $t, s \in \mathbf{N}$ it holds that $(X_1, ...X_s) \stackrel{\text{d}}{=} (X_{t+1}, ...X_{t+s})$.

Definition 2. (Covariance (Weak) Stationarity) This concept is related to the first two moments of a process. A process $(X_t)_{t\in\mathbf{N}}$ is covariance stationary if for all $t, s \in \mathbf{N}$ it holds that:

(i) $\text{E}[X_t] = \mu$

(ii) $\text{Cov}(X_t, X_{t+s}) = \text{E}[(X_t - \mu)(X_{s+t} - \mu)] = \phi_s$ (dependent on s but not t)

Neither stationarity nor covariance-stationarity are stronger than the other (see Question 1).

Definition 3. (Gaussian process) A process $(X_t)_{t\in\mathbf{N}}$ is Gaussian (jointly normal) if for all $t, s \in \mathbf{N}$ the sequence given by $(X_{t+1}, ..., X_{t+s})$ has a Gaussian distribution. The Gaussian distribution is characterised by the first two moments, this means that covariance stationarity + jointly normal distribution implies strict stationarity.

Definition 4. (Martingales) A process $(X_t)_{t\in\mathbf{N}}$ is a martingale if for all $t \in \mathbf{N}$ it holds that:

(i) $\text{E}[|X_t|] < \infty$

(ii) $\text{E}[X_t|X_{t-1}, ...X_0] = X_{t-1}$

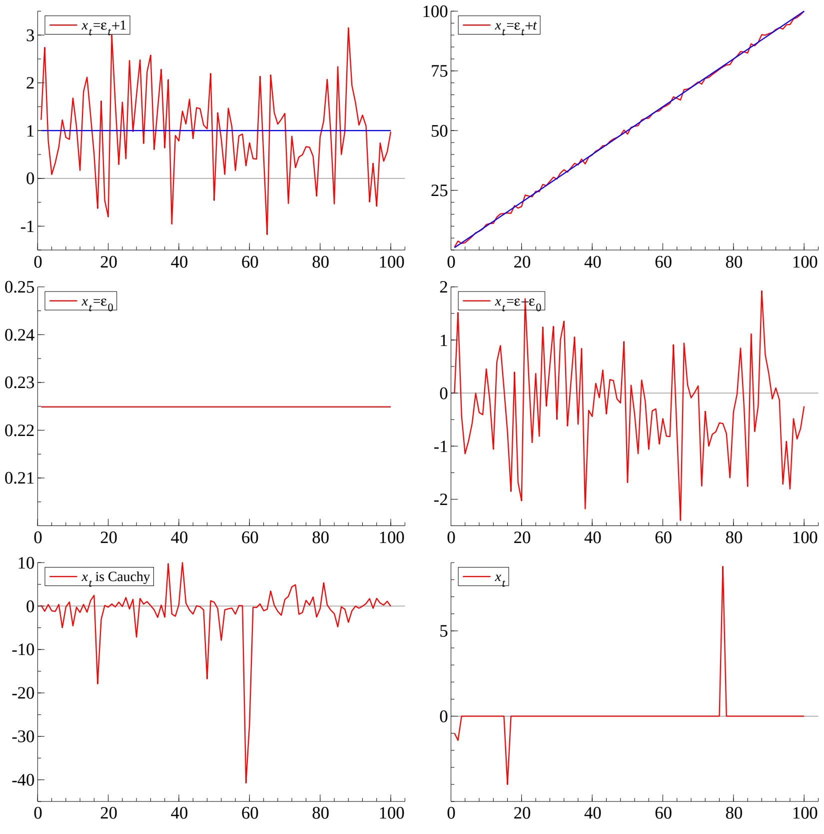
Definition 5. (Martingale Difference Sequences (MDS)) A process $(\epsilon_t)_{t\in\mathbf{N}}$ is a martingale difference sequence if $\text{E}[|\epsilon_t|] < \infty$ and for all $t \in \mathbf{N}$ $\text{E}[\epsilon_t|\epsilon_{t-1}, ...\epsilon_1] = 0$ for all $t > 1$. If X_t is a martingale, then $X_t - X_{t-1}$ is a martingale difference sequence.

Question 1

Let ϵ_t, $t = 0, 1, 2, \ldots$ be independent $\mathsf{N}(0, 1)$. For each of the following series compute first and second moment, autocovariance. Discuss whether their properties with respect to stationarity, covariance stationarity and martingale difference sequence.

Common mistake 1. You have to consider a general case for the covariance, i.e. $Cov(x_t, x_s)$ for $s \geq 1$ and all possible values of t.

First, let's have a look at the plots of the time series under consideration and try to get an idea about their stationarity. Below are the plots for the series, constructed using a single random draw from the Normal distribution (a-d), Cauchy (e) and the discrete case (f). What can we about the distribution of the processes just by looking at the plot? Does the series exhibit mean-reverting properties? Does the variance seem to change with time? (Note that process (b) is an example of a trend-stationary series, i.e. a process that is stationary around a trend - a de-trended series would look like (a)). We probably cannot say all that much about higher moments, so we have to compute (at least) the first two moments, explicitly.



(a) $x_t = \epsilon_t + 1$.
First and second moment:

$$E(x_t) = 1,$$

$$V(x_t) = V(\epsilon_t + 1) = V(\epsilon_t) = 1$$

Autocovariance:

$$\begin{aligned} K_{xx}(t,s) &= E[(x_t-\mu_t)(x_s-\mu_s)] \\ &= E[(x_t-1)(x_s-1)] \\ &= E(x_t x_s) - E(x_t) - E(x_s) + 1 \\ &= E[(\epsilon_t+1)(\epsilon_s+1)] - 1 - 1 + 1 \\ &= E(\epsilon_t\epsilon_s) + E(\epsilon_t) + E(\epsilon_s) + 1 - 1 = 0 \end{aligned}$$

We can use these results and write

$$\begin{pmatrix} x_t \\ x_{t+s} \end{pmatrix} \overset{\mathrm{d}}{=} \mathsf{N}\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

Stationarity: Since the joint distribution of x_t and x_{t+s} does not depend on t for any s, the process is stationary. This results follows from the joint-normality of the distribution above. Based on the first two moments, the process is also covariance stationary.

(b) $x_t = \epsilon_t + t$
First and second moments:

$$\begin{aligned} E(x_t) &= E(\epsilon_t + t) = t \\ V(x_t) &= V(\epsilon_t + t) \\ &= E[(\epsilon_t+t)(\epsilon_t+t)] - E[\epsilon_t+t]^2 \\ &= E(\epsilon_t^2) + 2tE(\epsilon_t) + E(t^2) - t^2 = 1 \end{aligned}$$

Autocovariance:

$$\begin{aligned} K_{xx}(t,s) &= E[(x_t-\mu_t)(x_s-\mu_s)] \\ &= E[(\epsilon_t+t-t)(\epsilon_s+s-s)] \\ &= E(\epsilon_t\epsilon_s) = 0 \end{aligned}$$

We can write this as:

$$\begin{pmatrix} x_t \\ x_{t+s} \end{pmatrix} \overset{\mathrm{d}}{=} \mathsf{N}\left\{ \begin{pmatrix} t \\ t+s \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

Note that we can write the joint process in this notation only because is it a normal distribution, which happens to be fully described by its first two moments. This is not the case for all distributions (e.g. part (f)), so you have to be careful when stating this strong result.

Stationarity: The process is not stationary, as the means depend on the value of t. Is is also not covariance stationary, as $E(x_t)$ is not constant

(c) $x_t = \epsilon_0$
First and second moment:

$$E(x_t) = E(\epsilon_0) = 0$$
$$V(x_t) = V(\epsilon_0) = 1$$

Autocovariance:

$$\begin{aligned} K_{xx}(t,s) &= E[(x_t-\mu_t)(x_s-\mu_s)] \\ &= E(\epsilon_0^2) = 1 \end{aligned}$$

Stationarity: The process is stationary as well as covariance stationary.

(d) $x_t = \epsilon_t - \epsilon_0$

This one is tricky, since you here you need to consider two case: $t = 0$ and $t \neq 0$. The process is only stationary if it's distribution does not change for all t. Even if a single time period exhibits different properties, the process is non-stationary. Let's show this by finding the first two moments:

$$E(x_t) = E(\epsilon_t - \epsilon_0) = 0$$

$$\begin{aligned} V(x_t) &= V(\epsilon_t - \epsilon_0) \\ &= E[(\epsilon_t - \epsilon_0)(\epsilon_t - \epsilon_0)] - E(\epsilon_t - \epsilon_0)^2 \\ &= E(\epsilon_t^2) - 2E(\epsilon_0\epsilon_t) + E(\epsilon_0^2) - 0 \\ &= \begin{cases} 2 & \text{for } t \geq 1 \\ 0 & \text{for } t = 0 \end{cases} \end{aligned}$$

Autocovariance:

$$\begin{aligned} K_{xx}(t,s) &= E[(\epsilon_t - \epsilon_0 - 0)(\epsilon_s - \epsilon_0 - 0)] \\ K_{xx}(t,s) &= E(\epsilon_t\epsilon_s) - E(\epsilon_0\epsilon_t) - E(\epsilon_0\epsilon_s) + E(\epsilon_0^2) \\ K_{xx}(t,s) &= \begin{cases} 1 & \text{for } t,s \geq 1 \\ 0 & \text{for } t = 0 \end{cases} \end{aligned}$$

Stationarity: Clearly, $\{X_t\}_{t\geq 1}$ is stationary and covariance stationary, but $\{X_t\}_{t\geq 0}$ is neither stationary, nor covariance stationary.

(e) x_t independent Cauchy distributed (same at t(1)).
Cauchy distribution is given by:

$$f(x) = \frac{1}{\pi(1 + x^2)}$$

For a Cauchy random variable, mean and variance are not defined, neither are higher moments (Take the integral to compute mean and variance from first principles to see this). Thus, while the process is stationary (distribution does not depend on t), it is *not* covariance stationary (as variance and covariance have to be finite).

(f) x_t takes the values $-\sqrt{t}$, 0, and $\sqrt{t}$ with probabilities $\frac{1}{2t}$, $1 - \frac{1}{t}$ and $\frac{1}{2t}$ respectively.
This one is also tricky, since you need to compute higher moments to show non-stationarity. But first, let concentrate on the first and second moments:

$$E(x_t) = \frac{1}{2t}(-\sqrt{t}) + \left(1 - \frac{1}{t}\right)0 + \frac{1}{2t}(\sqrt{t}) = 0$$

$$\begin{aligned} V(x_t) &= E(x_t^2) - E(x_t)^2 = \frac{1}{2t}(-\sqrt{t})^2 + \left(1 - \frac{1}{t}\right)0^2 + \frac{1}{2t}(\sqrt{t})^2 - 0 \\ &= 2\frac{1}{2t}t = 1 \end{aligned}$$

Autocovariance:

$$\begin{aligned} K_{xx}(t,s) &= E[(x_t-\mu_t)(x_s-\mu_s)] \\ &= E(x_t x_s) \\ &= \frac{1}{2t}(-\sqrt{t})\left[\frac{1}{2s}(-\sqrt{s})+\left(1-\frac{1}{s}\right)0+\frac{1}{2s}(\sqrt{s})\right] \\ &\quad +\left(1-\frac{1}{t}\right)0\left[\frac{1}{2s}(-\sqrt{s})+\left(1-\frac{1}{s}\right)0+\frac{1}{2s}(\sqrt{s})\right] \\ &\quad +\frac{1}{2t}(\sqrt{t})\left[\frac{1}{2s}(-\sqrt{s})+\left(1-\frac{1}{s}\right)0+\frac{1}{2s}(\sqrt{s})\right]=0 \end{aligned}$$

Stationarity: Note that we cannot write the same results for the joint distribution as in cases (a) and (b), since this is not a standard distribution and we don't know all its moments. It is, in fact, not described by only the first two moments, so we need to dig deeper. However, first note that the process is clearly covariance stationary. For (strict) stationarity we require the full joint distribution to be independent of t. To show this process is not stationary need to consider higher moments:

$$E(x_t^4) = \frac{1}{2t}(-\sqrt{t})^4+\frac{1}{2t}(\sqrt{t})^4 = t$$

By definition, the process is not stationary.

Question 2

Consider the regression equation

$$Y_i = \alpha X_i + \beta Z_i + u_i$$

Derive the least squares estimator for α.

Deriving the least squares estimator for a model with two regressors can be done in various ways. The direct (crude) approach would involve differentiating SSR and solving a bunch of equations with two unknowns. This is fine, but the computations are long, the final results are cumbersome and not so easy to interpret. Instead, we can use the orthogonalization approach, which will make our notation neat and simplify most calculations (you would have probably seen this in the first term with slightly different notation).

Let's introduce a new variable, a residual from the regression of X_i on Z_i:

$$(X_i|Z_i) = X_i - \frac{\sum_{j=1}^N X_j Z_j}{\sum_{j=1}^N Z_j^2} Z_i$$

The residuals from the regression on Y_i on X_i and Z_i respectively are denoted $(Y_i|X_i)$ and $(Y_i|Z_i)$. Note that the coefficients are the usual OLS estimators. Then it is easy to check that $(X_i|Z_i)$ are orthogonal to Z_i for all i (check this). So we can write the original regression in terms of these two orthogonal regressors:

$$Y_i = \alpha\,(X_i|Z_i) + \gamma Z_i + u_i$$

where $\gamma = \beta + \frac{\sum_{j=1}^N X_j Z_j}{\sum_{j=1}^N Z_j^2}\alpha$. Notice that the coefficient of Z_i is not longer β, but a linear combination of the coefficients from the original regression. We can check that the new regression is in fact equivalent to the original one:

$$\begin{aligned} Y_i &= \alpha\left(X_i|Z_i\right) + \gamma Z_i + u_i \\ &= \alpha\left(X_i - \frac{\sum_{j=1}^N X_j Z_j}{\sum_{j=1}^N Z_j^2} Z_i\right) + \left(\beta + \frac{\sum_{j=1}^N X_j Z_j}{\sum_{j=1}^N Z_j^2}\alpha\right) Z_i + u_i \\ &= \alpha X_i + \beta Z_i + u_i \end{aligned}$$

Since the two representations are equivalent, then the sum of squared residuals should be identical. Hence, we can write:

$$S(\alpha, \beta) = \sum_{i=1}^N \left(Y_i - \alpha\left(X_i|Z_i\right) - \gamma Z_i\right)^2$$

And the corresponding first order conditions with respect to α and β are:

$$\begin{aligned} 0 &= \frac{\partial}{\partial \alpha} S(\alpha, \beta) = -2\sum_{i=1}^N \left(Y_i - \alpha\left(X_i|Z_i\right) - \gamma Z_i\right)\left(X_i|Z_i\right) = -2\sum_{i=1}^N \left(Y_i - \alpha\left(X_i|Z_i\right)\right)\left(X_i|Z_i\right) \\ 0 &= \frac{\partial}{\partial \beta} S(\alpha, \beta) = -2\sum_{i=1}^N \left(Y_i - \alpha\left(X_i|Z_i\right) - \gamma Z_i\right) Z_i = -2\sum_{i=1}^N \left(Y_i - \gamma Z_i\right) Z_i \end{aligned}$$

Notice that the last two equalities in each of the equations are obtained using the orthogonality result, that is $\sum_{i=1}^N (Z_i)(X_i|Z_i) = 0$. These first order conditions are independent, since each contain only one unknown. This makes the computation much easier. From the first equation we find α in one step by re-arranging the terms:

$$\hat{\alpha} = \frac{\sum_{i=1}^N Y_i\left(X_i|Z_i\right)}{\sum_{i=1}^N \left(X_i|Z_i\right)^2}$$

Compare this with the result from the direct calculation:

$$\hat{\alpha} = \frac{\left(\sum_{i=1}^N Z_i^2\right)\left(\sum_{i=1}^N X_i Y_i\right) - \left(\sum_{i=1}^N X_i Z_i\right)\left(\sum_{i=1}^N Z_i Y_i\right)}{\left(\sum_{i=1}^N X_i^2\right)\left(\sum_{i=1}^N Z_i^2\right) - \left(\sum_{i=1}^N X_i Z_i\right)^2}$$

I think the former looks much neater and is has a straightforward interpretation - α is the regression coefficient of regressing Y_i on the residuals from the regression of X_i on Z_i. I wouldn't even know how to interpret the second equation by just looking at the formula (even thought the two representations are of course equivalent). Finally, the coefficient γ is given by

$$\hat{\gamma} = \frac{\sum_{i=1}^N Y_i Z_i}{\sum_{i=1}^N Z_i^2}$$

which is also the regression coefficient of regressing Y_i on Z_i. Even though this was not asked in the question, we will need $\hat{\gamma}$ at a later stage.

(b) What is the least squares estimator of β?

Using the same line of argument as in (a) and by orthogonalising Z_i with respect to X_i we can conclude that:

$$\hat{\beta} = \frac{\sum_{i=1}^N Y_i\left(Z_i|X_i\right)}{\sum_{i=1}^N \left(Z_i|X_i\right)^2}$$

(c) What is the t-statistic for testing $\alpha = 0$?

Orthogonalisation makes the calculation of standard errors of the estimators much easier as well. Recall that the covariance matrix is of the form:

$$s^2\left(\tilde{X}'\tilde{X}\right) = \frac{s^2}{\left(\sum_{i=1}^N X_i^2\right)\left(\sum_{i=1}^N Z_i^2\right) - \left(\sum_{i=1}^N X_i Z_i\right)^2} \begin{pmatrix} \sum_{i=1}^N Z_i^2 & -\sum_{i=1}^N X_i Z_i \\ -\sum_{i=1}^N X_i Z_i & \sum_{i=1}^N X_i^2 \end{pmatrix}$$

where $\tilde{X}$ is a matrix with X_i and Z_i, and s^2 is given by:

$$s^2 = \frac{1}{N-2}\sum_{i=1}^N \left(Y_i - \hat{\alpha}X_i - \hat{\beta}Z_i\right)^2 = \frac{1}{N-2}\sum_{i=1}^N (Y_i|X_i, Z_i)$$

Now, we make use of orthogonalisation once again, so that X_i becomes $(X_i|Z_i)$. The above equation becomes:

$$\begin{aligned} s^2\left(\tilde{X}'\tilde{X}\right) &= \frac{s^2}{\left(\sum_{i=1}^N X_i^2\right)\left(\sum_{i=1}^N Z_i^2\right) - \left(\sum_{i=1}^N X_i Z_i\right)^2} \begin{pmatrix} \sum_{i=1}^N Z_i^2 & -\sum_{i=1}^N X_i Z_i \\ -\sum_{i=1}^N X_i Z_i & \sum_{i=1}^N X_i^2 \end{pmatrix} \\ &= \frac{s^2}{\left(\sum_{i=1}^N (X_i|Z_i)^2\right)\left(\sum_{i=1}^N Z_i^2\right) - \left(\sum_{i=1}^N (X_i|Z_i)\, Z_i\right)^2} \\ &\times \begin{pmatrix} \sum_{i=1}^N Z_i^2 & -\sum_{i=1}^N (X_i|Z_i)\, Z_i \\ -\sum_{i=1}^N (X_i|Z_i)\, Z_i & \sum_{i=1}^N (X_i|Z_i)^2 \end{pmatrix} \\ &= s^2 \begin{pmatrix} \left(\sum_{i=1}^N (X_i|Z_i)^2\right)^{-1} & 0 \\ 0 & \left(\sum_{i=1}^N Z_i^2\right)^{-1} \end{pmatrix} \end{aligned}$$

as we used the property that $(X_i|Z_i)$ are orthogonal to Z_i, so the cross-product of the two disappears. Now, we recall that the t-statistic for testing $\alpha = 0$ is given by:

$$\begin{aligned} Z_{\alpha=0} &= \frac{\hat{\alpha} - 0}{se(\hat{\alpha})} \\ &= \left(\frac{\sum_{i=1}^N Y_i\,(X_i|Z_i)}{\sum_{i=1}^N (X_i|Z_i)^2}\right)\left(s^2\left(\sum_{i=1}^N (X_i|Z_i)^2\right)^{-1}\right)^{-1/2} \\ &= \frac{\sum_{i=1}^N Y_i\,(X_i|Z_i)}{\sqrt{s^2\sum_{i=1}^N (X_i|Z_i)^2}} \end{aligned}$$

where $se(\hat{\alpha})$ is given by the square root of first entry in the matrix $s^2\left(\tilde{X}'\tilde{X}\right)$. Again this equation has a much nicer interpretation in terms of residuals from partial and full regressions.

(d) Show that $(n-2)\,s^2 = (n-1)\,s^2_{\alpha=0}\left(1 - \hat{\lambda}^2\right)$, where $s^2_{\alpha=0}$ is the variance estimator in the restricted model where $\alpha = 0$. What is $\hat{\lambda}$?
This is just one method of doing this - perhaps a long one, but you get to see the computation using orthogonalization in its full glory. I will evaluate the expression on the left

hand side by re-arranging the equation for s^2 using the notation I introduced earlier. But first we have to make an important observation: $\sum_{i=1}^{N} Y_i \left(X_i|Z_i\right) = \sum_{i=1}^{N} \left(Y_i|Z_i\right)\left(X_i|Z_i\right)$

$$\begin{aligned}
\sum_{i=1}^{N} \left(Y_i|Z_i\right)\left(X_i|Z_i\right) &= \sum_{i=1}^{N} \left(Y_i - \frac{\sum_{j=1}^{N} Y_j Z_j}{\sum_{j=1}^{N} Z_j^2} Z_i\right)\left(X_i|Z_i\right) \\
&= \sum_{i=1}^{N} Y_i \left(X_i|Z_i\right) - \sum_{i=1}^{N} \left(\frac{\sum_{j=1}^{N} Y_j Z_j}{\sum_{j=1}^{N} Z_j^2} Z_i \left(X_i|Z_i\right)\right) \\
&= \sum_{i=1}^{N} Y_i \left(X_i|Z_i\right)
\end{aligned}$$

again because of orthogonality. We will use that fact in the following expansion of $(n-2)\, s^2$:

$$\begin{aligned}
(N-2)\, s^2 &= \sum_{i=1}^{N} \left(Y_i - \hat{\alpha} X_i - \hat{\beta} Z_i\right)^2 \\
&= \sum_{i=1}^{N} \left(Y_i - \hat{\alpha}\left(X_i|Z_i\right) - \hat{\gamma} Z_i\right)^2 \\
&= \sum_{i=1}^{N} \left(\left(Y_i|Z_i\right) - \hat{\alpha}\left(X_i|Z_i\right)\right)^2 \\
&= \sum_{i=1}^{N} \left(Y_i|Z_i\right)^2 - 2\hat{\alpha} \sum_{i=1}^{N} \left(Y_i|Z_i\right)\left(X_i|Z_i\right) + \hat{\alpha}^2 \sum_{i=1}^{N} \left(X_i|Z_i\right)^2 \\
&= \sum_{i=1}^{N} \left(Y_i|Z_i\right)^2 - 2\frac{\sum_{i=1}^{N} \left(Y_i|Z_i\right)\left(X_i|Z_i\right)}{\sum_{i=1}^{N} \left(X_i|Z_i\right)^2} \sum_{i=1}^{N} \left(Y_i|Z_i\right)\left(X_i|Z_i\right) \\
&\quad + \left(\frac{\sum_{i=1}^{N} \left(Y_i|Z_i\right)\left(X_i|Z_i\right)}{\sum_{i=1}^{N} \left(X_i|Z_i\right)^2}\right)^2 \sum_{i=1}^{N} \left(X_i|Z_i\right)^2 \\
&= \sum_{i=1}^{N} \left(Y_i|Z_i\right)^2 - \frac{\left(\sum_{i=1}^{N} \left(Y_i|Z_i\right)\left(X_i|Z_i\right)\right)^2}{\sum_{i=1}^{N} \left(X_i|Z_i\right)^2} \\
&= \sum_{i=1}^{N} \left(Y_i|Z_i\right)^2 \left(1 - \frac{\left(\sum_{i=1}^{N} \left(Y_i|Z_i\right)\left(X_i|Z_i\right)\right)^2}{\sum_{i=1}^{N} \left(Y_i|Z_i\right)^2 \sum_{i=1}^{N} \left(X_i|Z_i\right)^2}\right) \\
&= \sum_{i=1}^{N} \left(Y_i|Z_i\right)^2 \left(1 - \hat{\lambda}^2\right) \\
&= (N-1)\, s_{\alpha=0}^2 \left(1 - \hat{\lambda}^2\right)
\end{aligned}$$

where

$$s_{\alpha=0}^2 = (N-1)^{-1} \sum_{i=1}^{N} \left(Y_i - \gamma Z_i\right)^2 = (N-1)^{-1} \sum_{i=1}^{N} \left(Y_i|Z_i\right)^2$$

is the standard error from the regression of Y_i on Z_i, i.e. where the X_i variable has been omitted, and

$$\hat{\lambda}^2 = C\hat{orr}\left(Y_i, X_i|Z_i\right)^2 = \frac{\left(\sum_{i=1}^N \left(Y_i|Z_i\right)\left(X_i|Z_i\right)\right)^2}{\sum_{i=1}^N \left(Y_i|Z_i\right)^2 \sum_{i=1}^N \left(X_i|Z_i\right)^2}$$

is by definition the correlation between Y_i and X_i given Z_i.

(e) Write the t-statistic in the form $(n-2)^{1/2}\,\hat{\lambda}/\left(1-\hat{\lambda}^2\right)^{1/2}$

Combine the results from (c) and (d) and use the result that $\sum_{i=1}^N Y_i\left(X_i|Z_i\right) = \sum_{i=1}^N \left(Y_i|Z_i\right)\left(X_i|Z_i\right)$ to get:

$$\begin{aligned}
Z_{\alpha=0} &= \frac{\sum_{i=1}^N \left(Y_i|Z_i\right)\left(X_i|Z_i\right)}{\sqrt{s^2 \sum_{i=1}^N \left(X_i|Z_i\right)^2}} \\
&= \frac{\sum_{i=1}^N \left(Y_i|Z_i\right)\left(X_i|Z_i\right)}{\sqrt{(N-2)^{-1}\, s_{\alpha=0}^2 \left(1-\hat{\lambda}^2\right) \sum_{i=1}^N \left(X_i|Z_i\right)^2}} \\
&= \frac{\sum_{i=1}^N \left(Y_i|Z_i\right)\left(X_i|Z_i\right)}{\sqrt{(N-2)^{-1} \sum_{i=1}^N \left(Y_i|Z_i\right)^2 \left(1-\hat{\lambda}^2\right) \sum_{i=1}^N \left(X_i|Z_i\right)^2}} \\
&= \frac{\sum_{i=1}^N \left(Y_i|Z_i\right)\left(X_i|Z_i\right)}{\sqrt{(N-2)^{-1} \left(1-\hat{\lambda}^2\right)}\sqrt{\sum_{i=1}^N \left(Y_i|Z_i\right)^2 \sum_{i=1}^N \left(X_i|Z_i\right)^2}} \\
&= (N-2)^{1/2}\, \frac{\hat{\lambda}}{\left(1-\hat{\lambda}^2\right)}
\end{aligned}$$

(f) What is the LR-statistic for testing $\alpha = 0$?

By definition, the LR-statistic is given by

$$\text{LR} = -Nlog\left(\frac{\sigma_U^2}{\sigma_R^2}\right)$$

where

$$\begin{aligned}
\sigma_U^2 &= \frac{1}{N}\sum_{i=1}^N \left(Y_i|X_i, Z_i\right)^2 \\
\sigma_R^2 &= \frac{1}{N}\sum_{i=1}^N \left(Y_i|Z_i\right)^2
\end{aligned}$$

(g) Write the LR-statistic on the form $-Nlog\left(1-\hat{\lambda}^2\right)$

Recall, from part (d) that

$$\sigma_U^2 = \frac{1}{N}\sum_{i=1}^N \left(Y_i|Z_{,i}\right)^2 \left(1-\hat{\lambda}^2\right) = \sigma_R^2\left(1-\hat{\lambda}^2\right)$$

and the result follows immediately.

Use the above results to analyse the following time series models. First

$$Y_t = \alpha Y_{t-1} + \mu + \varepsilon_t$$

From the estimation point of view, there is not difference we are considering a cross-section or a time-series, so we can use all of the above results and make the necessary substitution of notation and get the required results almost immediately. You don't really need to derive all the results again.

(h) What is the least squares estimator for μ?

We can use the result from part (a) and make a substitution of the variables: $(i, n, Y_i, X_{i,} Z_i) \rightarrow (t, T, Y_t, 1, Y_{t-1})$ to get the result:

$$\hat{\mu} = \frac{\sum_{t=1}^{T} Y_t \left(1_t | Y_{t-1}\right)}{\sum_{t=1}^{T} \left(1_t | Y_{t-1}\right)^2}$$

(i) What is the t-statistic for testing $\mu = 0$

Make the same substitution as in (h) and use the result from (c) to get:

$$Z_{\mu=0} = \frac{\sum_{t=1}^{T} Y_t \left(1_t | Y_{t-1}\right)}{\sqrt{s^2 \sum_{t=1}^{T} \left(1_t | Y_{t-1}\right)^2}}$$

and

$$s^2 = \frac{1}{T-2} \sum_{t=1}^{T} \left(Y_t - \hat{\mu} - \hat{\alpha} Y_{t-1}\right)^2 = \frac{1}{T-2} \sum_{t=1}^{T} \left(Y_t | 1, Y_{t-1}\right)$$

Then consider

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t$$

(j) What is the least squares estimate for α_2?

Make the substitution $(i, n, Y_i, X_{i,} Z_i) \rightarrow (t, T, X_t, X_{t-2}, X_{t-1})$ applied to part (a) of the question:

$$\hat{\alpha}_2 = \frac{\sum_{t=1}^{T} X_t \left(X_{t-2} | X_{t-1}\right)}{\sum_{t=1}^{T} \left(X_{t-2} | X_{t-1}\right)^2}$$

(k) What is the LR-statistic for testing $\alpha_2 = 0$?

Make the substitution $(i, n, Y_i, X_{i,} Z_i) \rightarrow (t, T, X_t, X_{t-2}, X_{t-1})$ applied to part (f) of the question:

$$\text{LR} = -N log \left(\frac{\sigma_U^2}{\sigma_R^2}\right)$$

where

$$\sigma_U^2 = \frac{1}{T}\sum_{t=1}^{T}(X_t|X_{t-1}, X_{t-2})^2$$
$$\sigma_R^2 = \frac{1}{T}\sum_{t=1}^{T}(X_T|X_{t-1})^2$$

Finally, consider

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \mu + \epsilon_t$$

(l) What is the link between this regression and the partial correlogram?

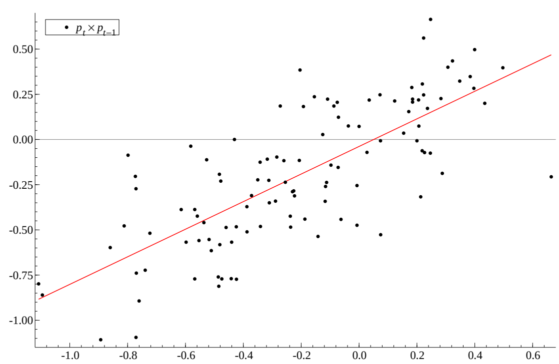
The second order sample partial autocorrelation in the partial correlogram equals the partial correlation $Corr\,(Y_t, Y_{t-2}|Y_{t-1}, 1)$, which is the partial correlation associated with the coefficient α_2. And recall the results from part (d), where we linked the estimators to partial correlations. In addition α_1 is the effect of Y_{t-1} on Y_t after partialling out the effects of Y_{t-2} and a constant on Y_t. Partial correlogram, on the other hand, plots correlation between Y_t and Y_{t-1}, while holding Y_{t-2} constant. Try to work with residuals of partial regressions and relate them to the estimates of the coefficients in the full regression, similarly to what we did in part (a).

Question 3

There is not much to say about this question. You may want to have a look at the code that generates the plot and the regression coefficients if you don't want to use the menus too much. The regression is given by:

$$p_t = 0.76 p_{t-1} - 0.039 + u_t$$

where u_t is the noise component. The intercept is the value that p_t takes when $p_{t-1} = 0$, note that is very close to zero and so the regression line (almost) passes through the origin. The slope is 0.76, which means that out best prediction for the price next period is 0.76 times the price in the current period.



Question 4

(a) Compute long-run mean of the autoregression

As seen above, an autoregressive equation with an intercept

$$Y_t = \mu + \alpha Y_{t-1} + \epsilon_t$$

can be rewritten for $\alpha \neq 0$ as

$$Y_t - \frac{\mu}{1-\alpha} = \alpha \left(Y_{t-1} - \frac{\mu}{1-\alpha} \right) + \epsilon_t$$

We use the term long-run mean to refer to the fraction $\frac{\mu}{1-\alpha}$. Consequently, the long-run mean in our case is given by

$$\begin{aligned} \mu_Y &= \frac{\mu}{1-\alpha} \\ &= \frac{6.78}{1-0.20} \\ &= 8.475 \end{aligned}$$

(b) Construct a likelihood ratio test for no serial correlation

The Likelihood Ratio test statistic can be calculated as

$$\begin{aligned} \mathsf{LR} &= -2\left(\hat{\ell}_R - \hat{\ell}\right) \\ &= -2(-123.01 + 120.67) \\ &= 4.68 \end{aligned}$$

This value needs to be compared against the critical value of a χ^2_1-distribution, which is 3.84. Since $4.68 > 3.84$, the hypothesis that $\alpha = 0$ could be rejected.

(c) Compare with t-test.

The t-test statistic for the same hypothesis is

$$\begin{aligned} \mathsf{t} &= \frac{0.20}{0.09} \\ &= 2.22 \end{aligned}$$

No sample size is given in the exercise. If we assume a large sample, we can again reject the null hypothesis at the 5% critical value since $2.22 > 1.96$.

(d) Interpret Figure 1.1.
Figures a) and c) show a small but observable degree of persistence in the log quantities of whiting sold. The correlogram and the partial correlogram in Figure b) suggest that indeed one lag of the dependent variable should be included in the model, even though the first bars are only marginally outside the confidence bands. This fits well our earlier estimation results, which suggested a small but significant coefficient on the lagged dependent variable. Figure d) finally shows that the errors are roughly normally distributed, though slightly skewed to the right. This confirms the visual impression from the residuals in Figure c) that negative residuals appear to be fewer but larger in absolute value.