STOCKHOLMS UNIVERSITET,
MATEMATISKA INSTITUTIONEN,
Avd. Matematisk statistik

Suggested solutions:
Econometric methods (MT5014)
2021-01-07

Problem 1

We use restricted/unrestricted regression (ch. 3.4 in Tyrcha et al.). The unrestricted model is

$$Y_j = \beta_0 + \beta_1 X_{1j} + \beta_2 X_{2j} + \beta_3 X_{3j} + \beta_4 X_{4j} + \beta_5 X_{5j} + \varepsilon_j.$$

Call the RSS of this model RSS_U. The hypothesis can be written as

$$\beta_3 = 3 + 3\beta_2$$

and under this hypothesis we can write the model as

$$Y_j = \beta_0 + \beta_1 X_{1j} + \beta_2 X_{2j} + (3 + 3\beta_2) X_{3j} + \beta_4 X_{4j} + \beta_5 X_{5j} + \varepsilon_j$$

which is equivalent to

$$Y_j - 3X_{3j} = \beta_0 + \beta_1 X_{1j} + \beta_2 (X_{2j} + 3X_{3j}) + \beta_4 X_{4j} + \beta_5 X_{5j} + \varepsilon_j,$$

which is thus the restricted model corresponding to the hypothesis. Call the RSS of this model RSS_R. Under the hypothesis it holds that

$$F = \frac{(RSS_R - RSS_U)/1}{RSS_U/(100-6)} = \frac{RSS_R - RSS_U}{RSS_U/94}$$

is $F(1, 94)$-distributed (cf. p 54); and given the values of RSS_R and RSS_U we can therefore use the usual procedure to see if the data supports rejecting the hypothesis or not.

Problem 2

(A)

$$\begin{aligned}
E[x_t] &= a + bE[x_{t-1}] + cE[\epsilon_t] + dE[\epsilon_{t-1}] \\
E[x_t] &= a + bE[x_t] \\
E[x_t] &= \frac{a}{1-b}
\end{aligned}$$

where the second equality follows from $\{\epsilon_t\}$ being a white noise series and $E[x_t] = E[x_{t-1}]$ since $\{x_t\}$ is assumed weakly stationary. Similarly,

$$\begin{aligned} V[x_t] &= V[a + c\epsilon_t] + V[bx_{t-1} + d\epsilon_{t-1}] \\ &= c^2 + b^2 V[x_{t-1}] + d^2 V[\epsilon_{t-1}] + 2bdCov[x_{t-1}, \epsilon_{t-1}] \end{aligned}$$

where the first equality follows since ϵ_t is independent of ϵ_{t-1} and x_{t-1}, and the second equality follows since $V[\epsilon_t] = 1$. Now

$$\begin{aligned} Cov[x_{t-1}, \epsilon_{t-1}] &= E[x_{t-1}\epsilon_{t-1}] - E[x_{t-1}]E[\epsilon_{t-1}] \\ &= E[x_{t-1}\epsilon_{t-1}] \\ &= E[(a + bx_{t-2} + c\epsilon_{t-1} + d\epsilon_{t-2})\epsilon_{t-1}] \\ &= cE[\epsilon_{t-1}^2] \\ &= cV[\epsilon_{t-1}] \\ &= c, \end{aligned}$$

since $E[\epsilon_{t-1}] = 0$ and $V[\epsilon_{t-1}] = 1$. Finally, combining the above expressions we have that

$$\begin{aligned} V[x_t] &= c^2 + b^2 V[x_{t-1}] + d^2 V[\epsilon_{t-1}] + 2bcd \\ &= c^2 + b^2 V[x_t] + d^2 + 2bcd \\ &= \frac{c^2 + d^2 + 2bcd}{1 - b^2}, \end{aligned}$$

where we also used the weak stationarity of $\{x_t\}$.

(B) Yes. Since $\{\epsilon_t\}$ is a white noise series, it is strictly stationary. Hence, $\{x_t\}$ consists of a strictly stationary series shifted by a constant, again resulting in a strictly stationary series.

(C) We have

$$\begin{aligned} x_{t+2} &= a + bx_{t+1} + c\epsilon_{t+2} + d\epsilon_{t+1} \\ &= a + b(a + bx_t + c\epsilon_{t+1} + d\epsilon_t) + c\epsilon_{t+2} + d\epsilon_{t+1} \\ &= a(1 + b) + b^2 x_t + bd\epsilon_t + (bc + d)\epsilon_{t+1} + c\epsilon_{t+2}. \end{aligned}$$

Conditioned on x_t and ϵ_t, the sum of the first three terms on the right hand side of the last equality is a constant. Further, the sum of the two last terms is Gaussian, since the sum of two Gaussian random variables again is Gaussian. Hence, we have that

$$\begin{aligned} x_{t+2}|x_t, \epsilon_t &\sim N(\mu, \sigma^2), \quad \text{with} \\ \mu &= a(1 + b) + b^2 x_t + bd\epsilon_t \\ \sigma^2 &= (bc + d)^2 + c^2 \end{aligned}$$

since ϵ_{t+1} and ϵ_{t+2} are independent, $E[\epsilon_{t+1}] = E[\epsilon_{t+2}] = 0$ and $V[\epsilon_{t+1}] = V[\epsilon_{t+2}] = 1$.

Problem 3

Using $\tilde{X}_j = aX_j$, we write the model fitted with measurement error as $Y_j = \hat{\alpha} + \hat{\beta}\tilde{X}_j + e_j$. It is now clear that the OLS-estimate for β with measurement error is given by

$$\begin{aligned}
\hat{\beta} &= \frac{\sum_j (\tilde{X}_j - \overline{\tilde{X}})(Y_j - \overline{Y})}{\sum_j (\tilde{X}_j - \overline{\tilde{X}})^2} \\
&= \frac{\sum_j (aX_j - a\overline{X})(Y_j - \overline{Y})}{\sum_j (aX_j - a\overline{X})^2} \\
&= \frac{1}{a} \frac{\sum_j (X_j - \overline{X})(Y_j - \overline{Y})}{\sum_j (X_j - \overline{X})^2} \\
&= \frac{1}{a} \hat{\beta}_{\text{no error}}.
\end{aligned}$$

where $\hat{\beta}_{\text{no error}}$ is the OLS estimate without measurement error. Similarly, the OLS estimate for α with measurement error is given by

$$\begin{aligned}
\hat{\alpha} &= \overline{Y} - \hat{\beta}\overline{\tilde{X}} \\
&= \overline{Y} - \hat{\beta}\overline{aX} \\
&= \overline{Y} - \hat{\beta}a\overline{X} \\
&= \overline{Y} - \frac{1}{a}\hat{\beta}_{\text{no error}} a\overline{X} \\
&= \overline{Y} - \hat{\beta}_{\text{no error}}\overline{X} \\
&= \hat{\alpha}_{\text{no error}}.
\end{aligned}$$

(We conclude that the measurement error affects only the estimate of β.)

Problem 4

(A) The 1 step ahead forecast is

$$\begin{aligned}
\hat{r}_8(1) &= E[r_9|F_8] \\
&= E[0.2 + 0.4r_8 + 1a_8 + a_9|F_8] \\
&= 0.2 + 0.4r_8 + a_8 + E[a_9|F_8] \\
&= 0.2 + 0.4r_8 + a_8 \\
&= 0.2 + 0.4 * 0.4 + 0.1 = 0.4600.
\end{aligned}$$

.

(B) The 2 step ahead forecast is

$$\begin{aligned}\hat{r}_8(2) &= E[r_{10}|F_8] \\ &= E[\phi_0 + 0.4r_9 + \theta a_9 + a_{10}|F_8] \\ &= 0.2 + E[0.4r_9|F_8] \\ &= 0.2 + 0.4E[r_9|F_8] \\ &= 0.2 + 0.4\hat{r}_8(1) \\ &= 0.2 + 0.4 * 0.4600 = 0.3840.\end{aligned}$$

The forecast error is

$$\begin{aligned}e_8(2) &= r_{10} - \hat{r}_8(2) \\ &= (0.2 + 0.4r_9 + \theta a_9 + a_{10}) - (0.2 + 0.4 * 0.4600) \\ &= (0.4r_9 + a_9 + a_{10}) - 0.4 * 0.4600 \\ &= 0.4r_9 + a_9 + a_{10} - 0.4 * 0.4600 \\ &= 0.4(0.2 + 0.4r_8 + a_8 + a_9) + a_9 + a_{10} - 0.4 * 0.4600 \\ &= 0.4(0.2 + 0.4 * 0.4 + 0.1 + a_9) + a_9 + a_{10} - 0.4 * 0.4600 \\ &= C + 1.4a_9 + a_{10}\end{aligned}$$

where C is an easily found constant that we do not need ($C = 0$). Hence, $V(e_8(2)) = V(1.4a_9 + a_{10}) = (1.4^2 + 1)\sigma^2 = 2.96\sigma^2$.

(C) Using the parameter values in the problem we find $r_9 = 0.4r_8 + a_9$. Similarly, $r_{10} = 0.4r_9 + a_{10} = 0.4(0.4r_8 + a_9) + a_{10} = 0.4^2 r_8 + 0.4a_9 + a_{10}$ and then $r_{11} = 0.4(0.4^2 r_8 + 0.4a_9 + a_{10}) + a_{11} = 0.4^3 r_8 + 0.4^2 a_9 + 0.4a_{10} + a_{11}$. Continuing this way we find, for arbitrary large l,

$$r_{8+l} = 0.4^l r_8 + 0.4^{l-1} a_9 + 0.4^{l-2} a_{10} + 0.4^{l-3} a_{11} + ... + 0.4^0 a_{8+l}.$$

The forecast $\hat{r}_8(l)$ can be computed recursively (cf. p. 56 in Tsay) as a constant depending on $r_8 = 0.4$ (we do not need to know the value of this constant however). Hence, the variance of the forecast error with lag l is given by

$$\begin{aligned}V(e_8(l)) &= V(r_{8+l} - \hat{r}_8(l)) \\ &= V(0.4^l r_8 + 0.4^{l-1} a_9 + 0.4^{l-2} a_{10} + 0.4^{l-3} a_{11} + ... + 0.4^0 a_{8+l}) \\ &= V(0.4^{l-1} a_9 + 0.4^{l-2} a_{10} + 0.4^{l-3} a_{11} + ... + 0.4^0 a_{8+l}) \\ &= 0.16^{l-1} V(a_9) + 0.16^{l-2} V(a_{10}) + 0.16^{l-3} V(a_{11}) + ... + 0.16^0 V(a_{8+l}) \\ &= \sigma^2(0.16^{l-1} + 0.16^{l-2} + 0.16^{l-3} + ... + 0.16^0).\end{aligned}$$

Since the sum in the parenthesis above converges to $1/(1 - 0.16) = 1.1905$ as $l \to \infty$, it holds that $\lim_{l\to\infty} V(e_8(l)) = 1.1905\sigma^2$.

Problem 5

The GLS formula (p. 89 in the compendium) is in this case

$$[\hat{\alpha}_{GLS}, \hat{\beta}_{GLS}]^T = (\boldsymbol{X}^T \boldsymbol{\Omega}^{-1} \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{\Omega}^{-1} \boldsymbol{Y}$$

where $\boldsymbol{\Omega} = \text{diag}(1^2, 2^2, ..., n^2)$, and

$$\boldsymbol{X} = \begin{bmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix}$$

$$\boldsymbol{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}.$$

Let $\boldsymbol{\beta} = [\alpha, \beta]^T$. Using the introduced notation we may write the model as $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. We find

$$\begin{aligned} [\hat{\alpha}_{GLS}, \hat{\beta}_{GLS}]^T &= (\boldsymbol{X}^T\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{\Omega}^{-1}\boldsymbol{Y} \\ &= (\boldsymbol{X}^T\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{\Omega}^{-1}(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) \\ &= \boldsymbol{\beta} + (\boldsymbol{X}^T\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{\Omega}^{-1}\boldsymbol{\varepsilon} \\ &= [\alpha, \beta]^T + (\boldsymbol{X}^T\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{\Omega}^{-1}\boldsymbol{\varepsilon} \end{aligned}$$

Using classical assumption $E[\boldsymbol{\varepsilon}|\mathbf{X}] = \mathbf{0}$ we find that $E[[\hat{\alpha}_{GLS}, \hat{\beta}_{GLS}]|\mathbf{X}] = [\alpha, \beta]$, and it follows that the estimators are unbiased. We use the formula on p. 90 (and that in our case $\sigma^2 = 1$) to find

$$V([\hat{\alpha}_{GLS}, \hat{\beta}_{GLS}]^T|\mathbf{X}) = (\mathbf{X}^T\mathbf{\Omega}^{-1}\mathbf{X})^{-1}$$

Using the expressions for $\mathbf{\Omega}$ and $\mathbf{X}$ above with $n = 3$ and the values for X_i given in the question, and basic matrix calculations we find: $\boldsymbol{\Omega}^{-1} = \text{diag}(1/1^2, 1/2^2, 1/3^2) = \text{diag}(1/2, 1/4, 1/9)$ and

$$\boldsymbol{X} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \end{bmatrix}.$$

More calculations give

$$(\mathbf{X}^T\mathbf{\Omega}^{-1}\mathbf{X}) = \begin{bmatrix} \frac{49}{36} & \frac{70}{36} \\ \frac{70}{36} & \frac{136}{36} \end{bmatrix}$$

and we conclude that

$$V([\hat{\alpha}_{GLS}, \hat{\beta}_{GLS}]^T|\mathbf{X}) = (\mathbf{X}^T\mathbf{\Omega}^{-1}\mathbf{X})^{-1} = \begin{bmatrix} \frac{136}{49} & \frac{-10}{7} \\ \frac{-10}{7} & 1 \end{bmatrix}$$

Hence, $V(\hat{\beta}_{GLS}|\boldsymbol{X}) = 1$.

Problem 6

Using basic probability (see also p. 111 in Tyrcha et al.) theory we find the moment conditions corresponding to the first and second moments to be

$$\mathbb{E}\left(V_j - \frac{\theta_0 + 15}{2}\right) = 0, \qquad \mathbb{E}\left(V_j^2 - \frac{\theta_0^2 + 15\theta_0 + 15^2}{3}\right) = 0,$$

where we know that $\theta_0 < 15$ (since the RV is distributed uniformly on the interval $[\theta_0, 15]$ this must be so).

Using the first moment as moment condition means that we estimate θ_0 by setting $\sum_{j=1}^{10} \frac{1}{10}\left(v_j - \frac{\hat{\theta}_0+15}{2}\right) = 0$, where the summation is taken over the sample, i.e. $v_1 = 8.87, v_2 = 5.99$ and so on (note that 10 is the sample size). Solving this equation gives $\hat{\theta}_0 = 2\sum_{j=1}^{10} \frac{v_j}{10} - 15 = 2\bar{v} - 15 = 2.514$.

Using notation from p. 115 (Tyrcha et al.) we find that the moments above correspond to $f_1(v_j, \theta) = v_j - \frac{\theta+15}{2}$ and $f_2(v_j, \theta) = v_j^2 - \frac{\theta^2+15\theta+15^2}{3}$. When W is the identity matrix we find (similar to Example 7.8 in Tyrcha et al.)

$$\begin{aligned} Q_{10}(\theta) &= \left(\sum_{j=1}^{10} v_j - \frac{\theta+15}{2}\right)^2 + \left(\sum_{j=1}^{10} v_j^2 - \frac{\theta^2+15\theta+15^2}{3}\right)^2. \\ &= \left(8.7570 - \frac{\theta+15}{2}\right)^2 + \left(88.3580 - \frac{\theta^2+15\theta+15^2}{3}\right)^2. \end{aligned}$$

The estimate that we are looking for is now given by $\hat{\theta}_0 = \text{argmin}_{\theta<15} Q_n(\theta)$ (easily solved with a computer); where the restriction $\theta < 15$ comes from the fact that 15 is the upper limit of the interval on which V_j is uniformly distributed (we could also have used the restriction $\theta \leq 4.27$ which is the smallest value in the sample).