Watson

(30) 1. Suppose that $u_t = v_t + e_t$ where $v_t = \phi_1 v_{t-1} + \phi_2 v_{t-2} + \varepsilon_t$, $e_t = \eta_t - \theta\eta_{t-1}$, and where $\varepsilon_t \sim \text{iid}(0, \sigma_\varepsilon^2)$, $\eta_t \sim \text{iid}(0, \sigma_\eta^2)$ and ε_t and η_j are independent for all t and j. You may assume v_t and e_t are stationary.

(10) (a) Show that u_t follows an ARMA(2,3) process.

(10) (b) Derive an expression showing the autocovariance generating function of u as a function of ϕ_1, ϕ_2, θ, σ_ε^2 and σ_η^2.

(10) (c) Suppose that $y_t = \mu + u_t$, and let $\bar{Y} = \frac{1}{T}\sum_{t=1}^{T} y_t$. Suppose that $\phi_1 = 1.3$, $\phi_2 = -0.6$, θ=0.8, $\sigma_\varepsilon^2 = 4$, and $\sigma_\eta^2 = 1$. Suppose that $T = 100$ and $\bar{Y} = 10$. Construct a 95% confidence interval for μ.

(30) 2. Suppose that $y_t = \phi y_{t-1} + \varepsilon_t$ where $\varepsilon_t \sim \text{iid}(0,1)$, and where $\phi = 1$ and $y_0 = 0$. Let $\hat{\phi}$ denote the OLS estimator of ϕ, and let $\hat{\varepsilon}_t = y_t - \hat{\phi} y_{t-1}$ denote the OLS residuals. I want you to show that $T^{-1}\sum_{t=1}^{T} \hat{\varepsilon}_t^2 \xrightarrow{p} 1$, and this will be done in parts(a)-(d).

(5) (a) Show that $T^{-1}\sum_{t=1}^{T} \hat{\varepsilon}_t^2 = T^{-1}\sum_{t=1}^{T} \varepsilon_t^2 + (\hat{\phi}-1)^2 T^{-1}\sum_{t=1}^{T} y_{t-1}^2 - 2(\hat{\phi}-1)T^{-1}\sum_{t=1}^{T} \varepsilon_t y_{t-1}$

(5) (b) Show that $T^{-1}\sum_{t=1}^{T} \varepsilon_t^2 \xrightarrow{p} 1$.

(10) (c) Show that $(\hat{\phi}-1)^2 T^{-1}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{p} 0$.

(10) (d) Show that $2(\hat{\phi}-1)T^{-1}\sum_{t=1}^{T} \varepsilon_t y_{t-1} \xrightarrow{p} 0$.

(30) 3. Suppose that $y_t = x_t + \varepsilon_t$, where $x_t = 0.8x_{t\text{-}1} + e_t$, and were

$$\begin{bmatrix} \varepsilon_t \\ e_t \end{bmatrix} \sim iidN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \right).$$ Suppose you know that $x_0 = 2$ and $y_1 = 6$.

(7) (a) Derive the minimum mean square error estimate of x_1.

(5) (b) What is the mean squared error of the estimate in (a)?

Now suppose that $\{\varepsilon_t\}$ and $\{e_t\}$ are mutually independent iid processes with (i) $\varepsilon_t = -2$ with probability 0.5 and $\varepsilon_t = 2$ with probability 0.5, and (ii) $e_t = -1$ with probability 0.5 and $e_t = 1$ with probability 0.5. Suppose you know that $x_0 = 2$ and $y_1 = 6$

(8) (c) Derive the <u>linear</u> minimum mean square error estimate of x_1.

(5) (d) What is the mean squared error of this estimate?

(5) (e) Is the estimate in (d) the minimum mean squared estimate? Explain.

Honore

Problem 1 (10 points)

You have a sample of n independent observations. For each observation, you have a discrete variable, y, which takes values 0 and 1. You also have a regressor, x. Suppose you estimate a logit model by maximum likelihood in order to characterize the relationship between y and x (you use x and a constant as explanatory variables). Let α be the constant and let β be the coefficient on x. Suppose you estimate α to be 0 and β to be 1.

What would be your estimate of $\frac{dP(y=1|x)}{dx}$ at $x = 1$.

Problem 2 (35 points)

Let y be the variable of interest and let x be a vector of regressors (including a constant). Assume that the conditional distribution of y given x is the exponential distribution with mean given by $\exp(x'\beta)$, where β is an unknown vector of parameters.

This implies that $V[y|x] = \exp(2x'\beta)$ and that the probability distribution of y given x is

$$f(y|x) = \left\{ \begin{array}{cc} \exp\left(-y\exp\left(-x'\beta\right)\right) & for \quad y \geq 0 \\ 0 & for \quad y < 0 \end{array} \right|$$

Assume that you have a random sample of observations of (y_i, x_i).

In answering the following questions, you do *not* have to worry about verifying the regularity conditions.

1. (9 points) Derive the asymptotic distribution of the non-linear least squares estimator of β:
$$\widehat{\beta}_{NLLS} = \arg\min_b \sum_{i=1}^{n} \left(y_i - \exp(x_i'b)\right)^2$$

2. (9 points) Derive the asymptotic distribution of the maximum likelihood estimator of β.

3. (17 points) Consider the following non–random sampling scheme. You pick n "clusters" (think of them as villages). For each cluster, you then get 2 observations. That gives you a total of $2 \cdot n$ observations. You are willing to assume that the observations are independent across clusters, but not within clusters. Let $\widetilde{\beta}_{MLE}$ be the maximum likelihood estimator that incorrectly assumes that the observations are independent. Find the asymptotic distribution of $\widetilde{\beta}_{MLE}$ calculated on the basis of all $2{\cdot}n$ observations. The asymptotic distribution will depend on the correlation structure within the cluster. Don't worry about that. We only want your answer to be sufficiently precise that you can use it to write a program to estimate the variance of the asymptotic distribution.

Problem 3 (30 points)

Consider observations of (y_{it}, x_{it}) from the linear panel data model

$$y_{i,t} = x_{it}'\beta + \gamma\left(y_{i,t-1} - y_{i,t-2}\right) + \alpha_i + \varepsilon_{it}, \quad t = 1, ..., T, \quad i = 1, ..., N$$

where α_i is an unobserved individual–specific effect. No assumption is made on the relationship between α_i and x_{it}. (Note that you do not observe $y_{i,0}$).

1. (15 points) Suppose that the dimensionality of x_{it} is 2 and

$$E\left[\varepsilon_{is} x_{it}\right] = 0 \qquad \text{for all } t \leq s \tag{1}$$

   What is the minimum T such that β and γ are identified? Is the model over–identified for that T? Explain and explicitly state any additional "regularity conditions" that you assume.

2. (15 points) Suppose that the dimensionality of x_{it} is 2 and

$$E\left[\varepsilon_{is} x_{it}\right] = 0 \qquad \text{for all } t \leq s + 1 \tag{2}$$

   (loosely speaking, we allow feedback from todays ε to future values of x, but we are willing to assume that it takes two periods for this to happen). What is the minimum T such that β and γ are identified? Is the model over–identified for that T? Explain and explicitly state any additional "regularity conditions" that you assume.

Problem 4 (15 points)

This problem is concerned with bounding treatment effects. Suppose that D is a random variable that indicates whether an individual has been "treated". If $D = 1$, the individual has received treatment and we observe a random variable Y_1. If $D = 0$, the individual has not received treatment and we observe a random variable Y_0. We do not observe Y_0 if $D = 1$, and we do not observe Y_1 if $D = 0$. This is the standard notation in this literature.

Suppose that it is known that $P\left(-1 \leq Y_0 \leq \frac{1}{2}Y_1 \leq 1\right) = 1$. Use this to construct bounds on the average treatment effect, $E\left[Y_1 - Y_0\right]$.