This exam consists of 4 pages. Please make sure that you have all pages before beginning.
Q1. (20) (Econ 307) Suppose \( Y_1, \ldots, Y_n \) are IID discrete random variables with
\[
\Pr(Y_i = 0) = \theta_0, \quad \Pr(Y_i = 1) = \theta_1, \quad \Pr(Y_i = 2) = \theta_2
\]
where \( \theta_j \geq 0 \) and \( \theta_0 + \theta_1 + \theta_2 = 1 \).

(a) Calculate \( E[Y_i] \) and \( E[Y_i^2] \) and use the results to derive a method of moments estimator for the parameters \((\theta_1, \theta_2)\).

(b) Show that the maximum likelihood estimator for \( \theta = (\theta_0, \theta_1, \theta_2) \) is
\[
\hat{\theta}_0 = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(Y_i = 0) \\
\hat{\theta}_1 = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(Y_i = 1) \\
\hat{\theta}_2 = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(Y_i = 2)
\]
(Hint: be sure to impose the constraint that \( \theta_0 + \theta_1 + \theta_2 = 1 \).)

(c) Prove directly that \( \hat{\theta}_0 \) is consistent for \( \theta_0 \), i.e., \( \hat{\theta}_0 \xrightarrow{p} \theta_0 \), without using the general result about consistency of MLE.

(d) Suppose that we take a Bayesian approach and use the following prior for \( \theta \):
\[
p(\theta) = c \theta_0^{\alpha_0-1} \cdot \theta_1^{\alpha_1-1} \cdot \theta_2^{\alpha_2-1},
\]
for \( \theta_j \geq 0 \) and \( \theta_0 + \theta_1 + \theta_2 = 1 \). Here, \( \alpha_j \) are some constants, and \( c \) is a normalizing constant. Derive the posterior (up to a normalizing constant) and show that this choice of prior is a conjugate prior.

Q2. (20) (Econ 309A) Consider a random variable \( y_i \) whose conditional distribution given \( x_i \) is characterized by the probability density function
\[
f(y_i|x_i, \beta) = \frac{1}{x_i^\beta} \exp \left( -\frac{y_i}{x_i^\beta} \right) \mathbb{1}(y_i > 0)
\]
You observe a random sample of \( \{(y_i, x_i)\}_{i=1}^{n} \) where \( x_i \in \mathbb{R}_+ \) and \( \beta \in \mathbb{R}_+ \).
(a) Let $e_i(\beta) := y_i - E[y_i | x_i]$. Then, we have the conditional moment restriction,

$$ E[e_i(\beta) | x_i] = 0. $$

Find an optimal instrumental variable (IV) and construct the unconditional moment restriction that exploits the optimal IV.

(b) Show that the GMM estimator that uses the unconditional moment restriction with the optimal IV is identical to the maximum likelihood estimator (MLE). Recall that the MLE maximizes the log-likelihood

$$ \sum_{i=1}^{n} \log f(y_i | x_i, \beta) $$

(Hint: you need to obtain neither the MLE nor the GMM estimator.)

Q3. (20) (Econ 8320) Consider an M-estimation where you want to minimize the objective function $\sum_{i=1}^{n} q(w_i, \theta)$ with respect to $\theta$, and the usual regularity conditions we discuss in the class hold. Suppose that you start with $\hat{\theta}^{(0)}$, a consistent estimator of $\theta_0$, and obtain in the next iteration

$$ \hat{\theta}^{(1)} = \hat{\theta}^{(0)} - \left[ \sum_{i=1}^{n} H(w_i, \hat{\theta}^{(0)}) \right]^{-1} \sum_{i=1}^{n} s(w_i, \hat{\theta}^{(0)}), $$

where $s(w, \theta) = \nabla_{\theta} q(w, \theta)$, and $H(w, \theta) = \partial^2 q(w, \theta) / \partial \theta \partial \theta'$.

(i) When is $\hat{\theta}^{(1)}$ a consistent estimator of $\theta_0$? Give out key conditions for consistency of $\hat{\theta}^{(1)}$.

(ii) Assume that $\sqrt{n}(\hat{\theta}^{(0)} - \theta_0)$ has a proper limit distribution. Let $\hat{\theta}$ be the estimator that minimizes $\sum_{i=1}^{n} q(w_i, \theta)$, i.e., $\sum_{i=1}^{n} s(w_i, \hat{\theta}) = 0$. Prove that $\sqrt{n}(\hat{\theta}^{(1)} - \theta_0)$ and $\sqrt{n}(\hat{\theta} - \theta_0)$ have the same asymptotic distribution.
Part II. (60) This part contains two sets of questions from 9310 and 9330 respectively. Each set has 2 questions. Students must answer both questions in one set of their choice, and choose one question from the other set, so they must answer three questions in total from part II. Each question in Part II accounts for 20.

Questions from Econ 9310

Q4. (20) Consider the following bivariate VARMA(1,1) model

\[
\begin{bmatrix}
y_{1t} \\
y_{2t}
\end{bmatrix} = \begin{bmatrix}
0.5 & 0 \\
0.5 & 0.5
\end{bmatrix} \begin{bmatrix}
y_{1t} \\
y_{2t}
\end{bmatrix} + \begin{bmatrix}
0.5 & 0 \\
0 & 0.5
\end{bmatrix} \begin{bmatrix}
u_{1t} \\
u_{2t}
\end{bmatrix} + \begin{bmatrix}
u_{1,t-1} \\
u_{2,t-1}
\end{bmatrix},
\]

where \( u_t = [u_{1t} \ u_{2t}]' \) is iid with zero mean and positive definite covariance matrix \( \Sigma \).

(a) Find reduced-form moving average coefficients \( \Theta_0 \) and \( \Theta_1 \) in

\[ y_t = \Theta_0 u_t + \Theta_1 u_{t-1} + \Theta_2 u_{t-2} \ldots \]

(b) Suppose that the reduced-form shock \( (u_t) \) and structural shock \( (\varepsilon_t) \) are related via \( u_t = H \varepsilon_t \), where \( H = [H_1 \ H_2] \) is a 2 \times 2 matrix whose (1,1) element is normalized to be one, i.e., \( H_{11} = 1 \), and \( H_1 \) and \( H_2 \) denote the first and second columns of \( H \), respectively.

To identify structural impulse responses to one of the structural shocks, say \( \varepsilon_{1t} \), suppose that there is an instrument, \( z_t \), that satisfies (i) \( E(z_t \varepsilon_{1t}) = \alpha \neq 0 \) and (ii) \( E(z_t \varepsilon_{2t}) = 0 \). Write \( E(u_t z_t) \) in terms of \( \alpha \) and \( H = [H_1 \ H_2] \).

(c) Based on (b), write down moment conditions for identifying structural impulse responses at horizons 0 and 1.

Q5. (20) Consider GMM estimation of a forecasting model of two-quarter-ahead inflation:

\[ \pi_{t+2} = \beta \pi_t + u_{t+2}, \]

where \( \pi_t \) denote inflation at time \( t \), \( E(u_{t+2}|\Omega_t) = 0 \) and \( \Omega_t \) denotes the set of information available at time \( t \).

(a) Let \( z_t \) denote a vector of instruments. Write down unconditional moment conditions for estimating \( \beta \). Provide examples of the instruments that can be used.

(b) Write down the rank condition for identification for the moment conditions in part (a). Is it likely to be satisfied? Explain why or why not.

(c) What is the optimal weighting matrix for the moment conditions in part (a)? Explain.

(d) What is the asymptotic covariance matrix of the GMM estimator based on the moment conditions in part (a)?
Questions from Econ 9330

Q6. (20) Suppose that we date our observations starting at \( t = 0 \), so that \( y_{i0} \) is the first observation on \( y \). For \( t = 1, \ldots, T \), we are interested in the dynamic unobserved effects model

\[
P(y_{it} = 1|y_{i,t-1}, \ldots, y_{i0}, z_i, c_i) = G(z_i\delta + \rho y_{i,t-1} + c_i),
\]

where \( z_{it} \) is a vector of contemporaneous explanatory variables, \( z_i = (z_{i1}, \ldots, z_{ iT}) \), and \( G \) can be the probit or logit function. The \( z_{it} \)'s are assumed to satisfy a strict exogeneity assumption (conditional on \( c_i \)).

(i) Which parameter in the model captures state dependence?

(ii) Could one claim that in general \( P(y_{it} = 1|y_{i,t-1}, z_i) = P(y_{it} = 1|z_i) \)? Why or Why not?

(iii) One can always write

\[
f(y_1, y_2, \cdots, y_T|y_0, z, c; \beta) = \prod_{t=1}^{T} G(z_i\delta + \rho y_{t-1} + c)^{y_i} [1 - G(z_i\delta + \rho y_{t-1} + c)]^{1-y_i}.
\]

With fixed-T asymptotics, can we use (6.2) to construct a log-likelihood function to estimate \( \beta \) consistently while treating \( c_i \) as parameters? Justify your claim.

(iv) What would you do to use the density in (6.2) to consistently estimate \( \beta \)? Discuss any issue arising from your suggested approach.

Q7. (20) Consider a binary treatment with potential outcomes \( Y_1 \) and \( Y_0 \). The econometrician observes the treatment indicator \( D \) and the observed outcome \( Y = DY_1 + (1-D)Y_0 \) for \( n \) individuals. In addition, he observes a covariate vector \( X \) for each individual.

(i) Describe the conditional independence (CI) assumption and the overlap assumption used in literature to identify the average treatment effect (ATE) and the average treatment effect on the treated (ATT).

(ii) Provide two alternative estimators for ATE and two for ATT and give as much detail as you can.