**ECMT Qualifying Exam (June, 2009)**

**Part I**

1. Let $X_1, ..., X_n$ be i.i.d. random variables with $E(X_i^6) < \infty$, $\mu = E(X_i)$, $\sigma^2 = E[(X_i - \mu)^2]$ and $\kappa_3 = E[(X_i - \mu)^3]$.

   Define $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, $S_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2$ and $K_n^3 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^3$.

   (a) Show that $\bar{X}_n$, $S_n^2$ and $K_n^3$ are consistent estimators for $\mu$, $\sigma^2$ and $\kappa_3$, respectively.

   (b) Derive the asymptotic distribution of $n^{1/2}[S_n^2 - \sigma^2]$.

2. Consider a linear regression model

$$Y = X\beta + u,$$

   where $Y$ and $u$ are $n \times 1$, $X$ is $n \times K$ and $\beta$ is $K \times 1$. Assume $u \sim N(0, \sigma^2 I_n)$, where $I_n$ is an identity matrix of dimension $n \times n$. Also, assume that $X$ is non-stochastic and that $X$ has full column rank (so that $X'X)^{-1}$ exist).

   Consider the problem of testing the null hypothesis of $R\beta - q = 0$, where $R$ is $J \times K$, and $q$ is $J \times 1$ $(J < K)$. Naturally, we can use $d = R\hat{\beta} - q$ as the basis of our test statistic, where $\hat{\beta}$ is the OLS estimator of $\beta$.

   (a) Compute $E(\hat{\beta})$ and $Var(\hat{\beta})$ under $H_0$.

   (b) What is the distribution of $d$ under $H_0$ ? Justify your answer.

   (c) Derive a $F$ test statistic based on a quadratic form of $d$.

   (d) Consider the case of $K = 3$, $J = 2$ and test $H_0$: $\beta_1 = 0$ and $\beta_2 + \beta_3 = 1$. What is $R$ and $q$ in $d = R\hat{\beta} - q$?

   (e) Now assume that $X$ is stochastic and the distribution of $u$ is unknown (not necessarily normally distributed). You can assume i.i.d. data. and assume that $n$ is large. Discuss whether the results you obtained in (a), (b) and (c) are still valid. Justify your answer.

3. Consider the following regression model:

$$Y = X_1\beta_1 + X_2\beta_2 + u,$$

where $Y$ and $u$ are of dimension $n \times 1$, $X_1$ and $X_2$ are dimension $n \times k_1$ and $n \times k_2$, respectively.

(a) Write the OLS estimator of $\beta_1$ in terms of $X_1$, $X_2$ and $Y$.

(b) Assuming that $X_1$ and $X_2$ are non-stochastic, compute $E(\hat{\beta}_1)$ and $Var(\hat{\beta}_1)$ using the expression of $\hat{\beta}_1$ obtained from (a).

(c) Now let us assume that $X_1$ and $X_2$ are stochastic, $u_i$ is i.i.d. $(0, \sigma^2)$ and independent of $X_1$ and $X_2$.

Show that $\hat{\beta}_1 \rightarrow \beta_1$ in probability.

You can apply law of large numbers arguments provided you apply it correctly and state the (moment) conditions you need on $X_{1i}$ and $X_{2i}$ when proving your result.

(d) Derive the asymptotic distribution of $\sqrt{n}(\hat{\beta}_1 - \beta_1)$. Again, state conditions you may need in order to derive your result.

2009 Econometrics Qualifying Exam

Part II

[1] Consider a structural equation $y = Y_1\gamma_1 + X_1\beta_1 + u_1 = Z_1\delta_1 + u_1$, and its reduced form equations

$$y = X_1\pi_{11} + X_2\pi_{21} + v_1, \qquad Y_1 = X_1\Pi_{12} + X_2\Pi_{22} + V_1$$

where $Y_1$: $T\times G_1$ is the data matrix of endogenous regressors, $X_1$: $T\times K_1$ and $X_2$: $T\times K_2$ are the included and excluded exogenous variables, respectively. Other standard assumptions apply.

(a) What is the necessary (order) condition for the identifiability of the structural equation?

(b) Describe the first and second stages of two-stage least squares estimation method.

(c) Present an alternative interpretation of the 2SLS estimator (excluding IV and GMM interpretations).

(d) We wish to test $H_0$: $cov(Y_{t1}, u_{t1}) = 0$ against $H_1$: $cov(Y_{t1}, u_{t1}) \neq 0$

(i) Show that this null hypothesis is equivalent to the null hypothesis $H_0$: $\theta = \gamma_1$, where $\theta \equiv \Omega_{22}^{-1}\Omega_{21}$.

(ii) Construct a test of the null hypothesis as specified in (i). Explain the idea and derivation of the test statistic. State clearly the null distribution of your test statistic.

[2] Consider a *bivariate* SVAR(2) and its reduced form VAR

(1) $\Gamma y_t = c_0 + C_1 y_{t-1} + C_2 y_{t-2} + u_t, \quad u_t \sim i.i.d.(0, \Sigma)$

(2) $y_t = a_0 + A_1 y_{t-1} + A_2 y_{t-2} + v_t, \quad v_t \sim i.i.d.(0, \Omega)$

where $y_t$, $y_{t-1}$, $y_{t-2}$, $u_t$ and $v_t$ are $2\times 1$ vectors, and $\Gamma$ and $C_i$ are $2\times 2$ matrices of coefficients.

(a) Show that the GLS estimator of the coefficients in (2) is identical to the OLS estimator of the coefficients in each equation.

(b) Derive the expression of the impulse response function of $y_{1,t+s}$ to the innovation $v_{2t}$ in terms of the elements $b_{s,ij}$ of matrix $B_s$, which is the coefficient matrix of the moving average representation of (2).

(c) Derive the expression of the *orthogonalized* impulse response function of $y_{1,t+s}$ to the innovation $v_{1t}$. Describe the steps to follow to compute this orthogonalized IRF.

(d) Let $\hat{y}_{t+2|t}$ be the minimum mean squared error forecast of $y_{t+2}$ conditional on information available at time $t$. Derive the variance-covariance matrix of the forecast error $y_{t+2} - \hat{y}_{t+2|t}$ vector and show the fractions of the variation of forecast error of $\hat{y}_{1,t+2|t}$ due to the variation in shocks $v_{1t}$ and $v_{2t}$. [Note that this question is asking specifically about two period ahead forecast.]

(e) Suppose that SVAR in (1) is a recursive model with $\gamma_{12}=0$, $\gamma_{11}=\gamma_{22}=1$ and a diagonal $\Sigma$. Show that this recursive model is just identified by showing that the only *admissible* transformation matrix is an identity matrix. That is, a transformation matrix $Q$ such that $\tilde{\Gamma}(L) = Q\Gamma(L)$ and $\tilde{\Sigma} = Q\Sigma Q'$ belong to the same recursive model must be $Q=I$.

[3] Suppose you are asked to analyze the academic performance of the first year graduate students. You want to set up a regression model to answer the following questions:

(Q1) Is the GRE score a significant predictor of the academic performance?

(Q2) Is there any difference in academic performance between male and female students?

(Q3) Is there any difference in the predictability of GRE between male and female students?

(a) Suppose you measure the academic performance by the average scores of all exams in the first year courses.

(i) Specify your econometric model, including the assumptions about statistical properties.

(ii) Specify the null and alternative hypotheses in terms of the model parameters you specified to test (Q2) and to test (Q3). In each case, specify clearly the test statistic, its null distribution, and how to draw the conclusion.

(b) Suppose that a significant proportion of incoming students dropped out and you don't have their exam scores though you have the data on their GRE scores and gender. [Use hint #3 for this question]

(i) Specify your econometric model, including the assumptions about statistical properties.

(ii) Describe how you would estimate your model. If you are going to use the MLE, specify the log-likelihood function.

(iii) How do you compute the marginal effect of the GRE score on the average of observed dependent variable? How about the marginal effect of gender variable?

(c) Now you want to measure the academic performance by passing or not passing the qualifying exams. You decided to use a binary variable ($y_i = 0$ if not passed all qualifiers, and $y_i = 1$ if passed all qualifiers). Assume that there were no drop-outs for this question. [Use hint #3 for this question]

(i) Specify your econometric model, including the assumptions about statistical properties.

(ii) Describe how you would estimate your model. If you are going to use the MLE, specify the log-likelihood function.

(iii) Compute the marginal effect of the GRE score and gender on the probability of passing all qualifying exams.

---

Hints:

(1) $(A \otimes B)' = (A' \otimes B')$, $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$, and $(A \otimes B)(C \otimes D) = (AC \otimes BD)$

(2) Let $\Omega$ be a symmetric positive definite matrix. Then, there exists a *unique* lower triangular nonsingular matrix $C$ such that $\Omega = CC'$. The diagonal elements of $C$ are all strictly positive. It can also be written as $\Omega = WDW'$, where $D$ is a diagonal matrix with diagonal elements $d_{ii} = c_{ii}^2$ and $W$ is a lower triangular matrix with the principal diagonal elements being equal to 1, and it is obtained by dividing each column of $C$ by its principal diagonal element ($w_{ij} = c_{ij}/c_{jj}$).

(3) The cdf and pdf of a logistic random variable $Z$ are $F(z) = 1/(1 + e^{-z})$ and $f(z) = F(z)[1 - F(z)]$. The means of $Z$ truncated at a constant c are

$$E(Z \mid Z \leq c) = \frac{cF(c) + ln(1 - F(c))}{F(c)}, \qquad E(Z \mid Z > c) = -\frac{cF(c) + ln(1 - F(c))}{1 - F(c)}$$