Questions from Previous Final Exams

1. (1996) In the non-linear regression model;
   \[ y_i = g_i(\theta) + \varepsilon_i \quad i = 1, 2, \ldots, N \]
   with \( \varepsilon_i \) independent with mean zero and variance \( \sigma^2 \). The condition;
   \[ \lim_{N \to \infty} \sum_{i=1}^{N} (g_i(\theta) - g_i(\theta'))^2 = \infty \]
   for \( \theta' \neq \theta \) is necessary and sufficient for the existence of a consistent estimator of \( \theta \). Interpret this condition in the context of the linear regression model.

2. (1996) The ML estimates of \( \theta \), a \((2 \times 1)\) vector of parameters, are \((6, 2)\)' and their asymptotic normal distribution has variance matrix;
   \[
   \begin{pmatrix}
   2 & 1 \\
   1 & 2
   \end{pmatrix}
   \]
   It turns out that the really interesting parameter is;
   \[ \gamma = \theta_1 + \frac{1}{2} \theta_2^2 \]
   What is the ML estimate of \( \gamma \), and what is asymptotic variance? How would you test the hypothesis that \( \gamma = 6 \)?

3. (1996) Consider the system;
   \[ y_1 = \alpha_1 + \beta_1 y_2 + \varepsilon_1 \]
   \[ y_2 = \alpha_2 + \beta_2 x + \varepsilon_2 \]
   with \( E \varepsilon_k^2 = \sigma_k^2 \) for \( k = 1, 2 \) and \( E \varepsilon_1 \varepsilon_2 = \sigma_{12} \). Identify briefly whether these equations are identified and how the parameters might be estimated. Now suppose that you have reason to believe \( \sigma_{12} = 0 \). How might you estimate the parameters under this assumption? In particular, consider the properties of the OLS estimator. How might you test the restriction that \( \sigma_{12} = 0 \)?

4. (1997) Consider the trinomial distribution;
   \[ P[y = 0] = p_0, P[y = 1] = p_1, P[y = 2] = p_2 = 1 - p_0 - p_1 \]
   a. What are the mean and variance of \( y \)?
   b. What are the MLE for \( p_0, p_1 \) and \( p_2 \) given a sample of \( N \) draws from this distribution?
   c. What is the asymptotic distribution of the MLE?
   d. Entropy, a measure of the expected information value of a draw from a distribution, is defined as
      \[ R = -p_0 \ln p_0 - p_1 \ln p_1 - p_2 \ln p_2 \]
      Find a good estimator for \( R \) and give its asymptotic distribution.

5. Suppose
   \[ y_t = \alpha + \beta t + \varepsilon_t \]
   \[ x_t = \gamma t + u_t \]
   where \( t = 1, 2, \ldots, T \) and \( E(\varepsilon_t) = E(u_t) = E(\varepsilon_t u_t) = 0 \). You regress \( y \) on \( x \) and a constant, getting a slope coefficient \( b \). Find \( \text{plim} b \). In response to a criticism, you expand your model to;
   \[ y_t = \lambda + \delta t + \eta x_t + \xi_t \]
   and fit again. Find \( \text{plim} \hat{\lambda} \) and \( \text{plim} \hat{\eta} \).

6. Consider the model;
\[ y_i = X_i \beta + \varepsilon_i \quad i = 1, 2, \ldots, N \]

Here \( X_i \) is \((1 \times K)\). We assume that
\[
E(\varepsilon_i) = 0
\]
\[
Var(\varepsilon_i) = \begin{cases} 1 & \text{with probability } \frac{1}{2} \\ 99 & \text{with probability } \frac{1}{2} \end{cases}
\]

That is, about \( \frac{1}{2} \) of the observations have variance 1 and \( \frac{1}{2} \) have 99 and you do not know which are which.

a. Is \( \hat{\beta}_{OLS} \) unbiased? BLUE? Consistent?
b. What is the asymptotic distribution of \( \hat{\beta}_{OLS} \)?
c. Now, suppose you acquire information on which observations have variance 1 and which 99. For simplicity, assume the rows of \( X \) are the same for the 2 variance groups, so we can write
\[
\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \beta + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}
\]

with \( Var(\varepsilon_1) = \frac{1}{2} I_N \) and \( Var(\varepsilon_2) = 99I_N \). Can you use this information to make a better estimator? What is it? Compare the asymptotic variance of your new estimator with that of \( \hat{\beta}_{OLS} \).