1. Take the model \( y_i = x_{1i} \beta_1 + x_{2i} \beta_2 + e_i \) with \( E x_i e_i = 0 \). Suppose that \( \beta_1 \) is estimated by regressing \( y_i \) on \( x_{1i} \) only. Find the probability limit of this estimator. In general, is it consistent for \( \beta_1 \)? If not, under what conditions is this estimator consistent for \( \beta_1 \)?

2. Of the variables \( (y_i^*, y_i, x_i) \) only the pair \( (y_i, x_i) \) are observed. In this case, we say that \( y_i^* \) is a latent variable. Suppose

\[
\begin{align*}
y_i^* &= x_i \beta + e_i \\
E(x_i e_i) &= 0 \\
y_i &= y_i^* + u_i
\end{align*}
\]

where \( u_i \) is a measurement error satisfying

\[
\begin{align*}
E(x_i u_i) &= 0 \\
E(y_i^* u_i) &= 0
\end{align*}
\]

Let \( \tilde{\beta} \) denote the OLS coefficient from the regression of \( y_i \) on \( x_i \).

(a) Is \( \beta \) the coefficient from the linear projection of \( y_i \) on \( x_i \)?

(b) Is \( \tilde{\beta} \) consistent for \( \beta \) as \( n \to \infty \)?

(c) Find the asymptotic distribution of \( \sqrt{n}(\tilde{\beta} - \beta) \) as \( n \to \infty \).

3. Take the model \( Y = X \beta + e \) where \( \beta \) is subject to the restriction \( R' \beta = r \) where \( R \) is a known \( k \times s \) matrix, \( r \) is a known \( s \times 1 \) vector, \( 0 < s < k \), and \( \text{rank}(R) = s \). The restricted least-squares estimator of \( \beta \) is

\[
\tilde{\beta} = \arg \min_{\beta: R' \beta = r} S_n(\beta)
\]

\[
S_n(\beta) = (Y - X \beta)' (Y - X \beta).
\]

(a) Explain why \( \tilde{\beta} \) solves the minimization of the Lagrangian

\[
L(\beta, \lambda) = \frac{1}{2} S_n(\beta) + \lambda' (R' \beta - r)
\]

where \( \lambda \) is \( s \times 1 \).
(b) Show that the solution is
\[ \begin{align*}
\hat{\beta} &= \tilde{\beta} - (X'X)^{-1} R [R' (X'X)^{-1} R]^{-1} (R'\hat{\beta} - r) \\
\hat{\lambda} &= [R' (X'X)^{-1} R]^{-1} (R'\hat{\beta} - r)
\end{align*} \]
where
\[ \hat{\beta} = (X'X)^{-1} X' Y \]
is the unconstrained OLS estimator.

(c) Verify that \( R'\tilde{\beta} = r \).

(d) Show that if \( R'\beta = r \) is true, then
\[ \hat{\beta} - \beta = \left( I_k - (X'X)^{-1} R [R' (X'X)^{-1} R]^{-1} R' \right) (X'X)^{-1} X'e. \]

(e) Under the standard assumptions plus \( R'\beta = r \), find the asymptotic distribution of \( \sqrt{n} (\hat{\beta} - \beta) \) as \( n \to \infty \).

(f) Find an appropriate formula to calculate standard errors for the elements of \( \tilde{\beta} \).

4. The ASCII datafile “invest.dat” is a 565 \times 4 matrix containing data on 565 U.S. firms extracted from Compustat for the year 1987. The variables, in order, are

- \( I_i \) Investment to Capital Ratio (multiplied by 100).
- \( Q_i \) Total Market Value to Asset Ratio (Tobin’s Q).
- \( C_i \) Cash Flow to Asset Ratio.
- \( D_i \) Long Term Debt to Asset Ratio.

The flow variables are annual sums for 1987. The stock variables are beginning of year.

(a) Estimate a linear regression of \( I_i \) on the other variables. Calculate appropriate standard errors.

(b) Calculate asymptotic confidence intervals for the coefficients.

(c) This regression is related to Tobin’s \( q \) theory of investment, which suggests that investment should be predicted solely by \( Q_i \). Thus the coefficient on \( Q_i \) should be positive and the others should be zero. Test the joint hypothesis that the coefficients on \( C_i \) and \( D_i \) are zero. Test the hypothesis that the coefficient on \( Q_i \) is zero. Are the results consistent with the predictions of the theory?

(d) Now try a non-linear (quadratic) specification. Regress \( I_i \) on \( Q_i, C_i, D_i, Q_i^2, C_i^2, D_i^2, Q_iC_i, Q_iD_i, C_iD_i \). Test the joint hypothesis that the six interaction and quadratic coefficients are zero.