ECON 510

Introduction to Probability and Statistics II

Homework I

1) Let X_i be $1\times k$ row vector of regressor matrix X for each i. Consider following estimator for the equation $y = X\beta + \varepsilon$:

$$\hat{\beta} = \arg\min_{\beta}\left(\sum_{i=1}^{N}\left(y_i - X_i\beta\right)^2 + \lambda\sum_{j=1}^{k}\beta_j^2\right)$$

a) Derive a close form solution for $\hat{\beta}$ given any value of λ. (Hint: using matrix notation may simplify the derivation)

b) Under what condition this new estimator is equivalent to classic OLS estimator.

c) Now assume X is orthonormal (only for part c) i.e. $X'X = I$ then find the relation between OLS and this new estimator. Hint: result will be a function of OLS estimate and λ.

d) Does this estimator unbiased? If not compute the bias in terms of $X,\ \lambda\ and\ \beta$.

e) Compute the variance of $\hat{\beta}$

f) Show that even if $X'X$ is not invertible, for each positive value of λ, $\hat{\beta}$ is uniquely determined. (bonus).

2) Consider the estimator $\hat{\beta} = (Z'AX)^{-1}Z'Ay$, where A is a symmetric $N\times N$ matrix of constants, in the multiple regression model $y = X\beta + \varepsilon$ where y is an $N\times 1$ vector, X is an $N\times k$ matrix, Z is an $N\times k$ matrix, β is a $k\times 1$ vector and u is an $N\times 1$ vector.

We assume that $u\,|\,Z \sim N(0,\Sigma)$ where $\Sigma \neq \sigma^2 I$. For asymptotic theory assume that necessary laws of large numbers and central limit theorems can be applied.

a) Obtain $E\left[\hat{\beta}\right]$ and $\operatorname{var}(\hat{\beta})$. For this part only assume that Z and X are nonstochastic.

b) Show that $\hat{\beta}$ is consistent, stating any necessary assumptions.

c) Obtain the limit distribution of $\left(\sqrt{N}(\hat{\beta} - \beta)\right)$

3) Let $R\beta = r$ be linear restriction on coefficient vector β. Let $\hat{\beta}$ be unrestricted estimator of β and $\hat{\beta}_R$ be restricted estimator. Show that $\text{cov}[\hat{\beta}, \hat{\beta}_R \mid X] = Var[\hat{\beta}_R \mid X]$

4) Consider regression equation $y = X\beta + \varepsilon$ where all classical linear regression assumptions hold and $\text{var}(\varepsilon \mid X) = \sigma^2$. Define residual sum of square as $RSS = (y - X\hat{\beta})'(y - X\hat{\beta})$. Find the distribution of following statistic RSS / σ^2 .

5) Consider a simple bivariate regression $y = \alpha + \beta x + \varepsilon$ with standard assumptions. Further assume y *and* x has same marginal distribution, i.e. $x \sim D(\mu, \sigma^2)$ and $y \sim D(\mu, \sigma^2)$ where D is a distribution, but $y \neq x$ in vector sense.

a) Show that $\hat{\beta} < 1$

b) Now, suppose we run $x = \alpha^* + \beta^* y + \varepsilon^*$. Compare the regression coefficients with ones in previous regression.

6) Suppose, we have a regression model $y = X\beta + \varepsilon$ with standard assumptions. We know that from $\sigma^2 = E\left[\varepsilon^2\right]$ the variance of ε is a moment and a natural estimator for the variance σ^2 is sample variance $\tilde{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N}\varepsilon_i^2$. However, in practice we cannot observe true residual process, thus, it is not feasible. Whereas we could use following feasible estimator for the variance term: $\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N}\hat{\varepsilon}_i^2$ where $\hat{\varepsilon}$ is residuals from OLS regression. Now prove following statement: the feasible estimator $\hat{\sigma}^2$ is always smaller than infeasible estimator $\tilde{\sigma}^2$

7) Short questions

a) Let $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$ and assume $X_1'X_2 = 0$. Show $P = P_1 + P_2$

b) Show that if X contains a constant term, then $\frac{1}{N}\sum_{i=1}^{N}\hat{y}_i = \bar{y}$ thus, sample mean of predicted y is equal to actual sample mean of y.

c) Prove that R^2 is square of sample correlation between y and $\hat{y}$.

8) Consider following simple linear regression model:

$$y_i = x_i\beta + \varepsilon_i$$
$$E\left[\varepsilon_i \mid x_i\right] = 0$$

where $x_i \in R$. define following estimators:

$$\hat{\beta} = \frac{\sum_{i=1}^{N} x_i y_i}{\sum_{i=1}^{N} x_i^2} \text{ and } \tilde{\beta} = \sum_{i=1}^{N} \frac{y_i}{x_i}$$

a) Under stated assumptions**,** are these estimators consistent for β ?

b) Are there conditions in which either estimator is efficient?