Econometrics A
Practice Problems #1

1. Suppose that $(X,Y)$ is a random vector, where $X \geq 1$, $Y \leq M$ (each with probability one) and $E[Y|X] = \theta X$. Let $Z = Y/X$.

(a) Is $Z$ a random variable? Why?
(b) Show that $E[Z] = \theta$.
(c) Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be an i.i.d. sample from $(X,Y)$. Let $\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n \frac{Y_i}{X_i}$.
   Show that:
   i. $\hat{\theta}_n$ is an unbiased estimator of $\theta$.
   ii. $\hat{\theta}_n$ is a consistent estimator of $\theta$.

2. Let $X_1, X_2, \ldots, X_n$ be an i.i.d. sample from $X$, where $\text{Var}[X] < \infty$. Suppose you are interested in estimating $E[X]$. Consider two different estimators,

$$\hat{\theta}_1, n = \bar{X}_n + \frac{1}{n}$$
$$\hat{\theta}_2, n = \frac{1}{4} X_1 + \frac{3}{4} X_n .$$

(a) Show that $\hat{\theta}_1,n$ is a biased estimator of $E[X]$.
(b) Show that $\hat{\theta}_2,n$ is an unbiased estimator of $E[X]$.
(c) What is $\text{Var}(\hat{\theta}_1,n)$?
(d) What is $\text{Var}(\hat{\theta}_2,n)$? How does it compare with the answer to part (c)?
(e) Show that $\hat{\theta}_1,n$ is a consistent estimator of $E[X]$.

3. Let $X_1, X_2, \ldots, X_n$ be an i.i.d. sample from $X$, where $X$ is a binary random variable. Assume that $p = P\{X = 1\}$ is such that $0 < p < 1$. Let

$$\gamma = \frac{p}{1-p}$$

and define

$$\hat{\gamma}_n = \frac{\bar{X}_n}{1 - \bar{X}_n} .$$

(a) Can you show that $\hat{\gamma}_n$ is an unbiased estimator of $\gamma$?
(b) Show that $\hat{\gamma}_n$ is a consistent estimator of $\gamma$. 
4. Suppose

\[ Y = \beta_0 + \beta_1 X + U, \]

where \( Y \) is a binary random variable. Suppose further that \( E[U|X] = 0 \) and \( 0 < \text{Var}[X] < \infty \).

(a) What is \( E[Y|X] \)? What is \( P\{Y = 1|X\} \)?

(b) What is \( \text{Var}[Y|X] \)?

(c) What is \( \text{Var}[U|X] \)? Is the model homoskedastic or heteroskedastic?

(d) Let \((Y_1, X_1), \ldots, (Y_n, X_n)\) be a i.i.d. sample from \((Y, X)\). In addition to the assumptions above, suppose that \( E[X^4] < \infty \). Assume that the sample size, \( n \), is large.

i. How would you test the null hypothesis that \( \beta_1 = 0 \) versus the alternative that \( \beta_1 \neq 0 \) at the 5% significance level?

ii. How would you compute the \( p \)-value for the test in part (i)?

iii. Construct a (two-sided) confidence interval for \( \beta_1 \) at the 5% significance level?

iv. What coverage property does the interval in part (iii) satisfy?