

Problem Set 01
Linear Models -- Fall 2015
Due date: 2015-09-19

Problems sets are due at the start of class on the due date. You may either hand write or type up and print the solutions, but we will not accept e-mail solution sets except in exceptional circumstances. You may discuss problem sets with other students, but must write up your own solutions. This means that you should have no need to look at other student's final written solutions. Many of these problems come from a variety of textbooks, which are referenced in the problems. These are for citation purposes and not because you will need to consult the text itself (though you may feel free to do so).

1. [Sheather 2009, pg 39] A story by James R. Hagerty entitled *With Buyers Sidelined, Home Prices Slide* published in the Thursday October 25, 2007 edition of the *Wall Street Journal* contained data on so-called fundamental housing indicators in major real estate markets across the US. The author argues that... *prices are generally falling and overdue loan payments are piling up*. Thus, we shall consider data presented in the article on:

Y = Percentage change in average price from July 2006 to July 2007 (based on the S&P/Case-Shiller national housing index); and
x = Percentage of mortgage loans 30 days or more overdue in latest quarter (based on data from Equifax and Moody's).

The data are available at euler.stat.yale.edu/~tba3/psets/pset01/data/indicators.txt. Fit a simple linear regression model with an intercept to the data.

- (a) Find a 95% confidence interval for the slope of the regression model, β . On the basis of this confidence interval decide whether there is evidence of a significant negative linear association.
- (b) Use the fitted regression model to estimate $\mathbb{E}(Y|X = 4)$. Find a 95% confidence interval for $\mathbb{E}(Y|X = 4)$. Is 0% a feasible value for $\mathbb{E}(Y|X = 4)$? Give a reason to support your answer.

2. The Cramér–Rao bound gives a lower bound on the variance of any unbiased estimator for a given deterministic parameter. Specifically, if $\hat{\theta}$ is an unbiased estimator of θ , then the following inequality holds:

$$\mathbb{V}\hat{\theta} \geq \frac{1}{I(\theta)} \quad (1)$$

Where $I(\theta)$ is the Fisher information. The Fisher information is defined in terms of the log-likelihood ℓ :

$$I(\theta) = \mathbb{E} \left[\left(\frac{\partial \ell}{\partial \theta} \right)^2 \right] \quad (2)$$

Calculate the Fisher information for β in the simple linear regression model with no intercept and normal, i.i.d. errors. Use this to establish that the MLE estimator for this model achieves the Cramér–Rao bound (i.e., it has the lowest allowable variance amongst the class of unbiased estimators).

2.