1. Consider a random variable $Y$ with finite variance. Show that for any other random variable $X$, the function of $X$ with finite variance, say $g$, that minimizes the mean square error $E(Y - g(X))^2$ is the conditional expectation $E(Y|X)$.

2. Compute the covariance between the ordinary least squares estimators of slope and intercept, $b_1$ and $b_0$ in a simple linear regression.

3. The data set from example 2.1 in the text can be downloaded from the course homepage. Use PROC REG in SAS to estimate the least squares regression line, the error variance and compute $R^2$. Test whether there is any significant linear relationship between Market Rate and Accounting Rate. Also, test the hypothesis that the slope of the regression line is different from zero. Plot the residuals versus the explanatory variable and check for any pattern that may indicate that the residuals do not form a random scatter.

4. We have derived finite sample properties of the least squares estimators (under the model assumptions) such as unbiasedness, variance going to zero if $S_{xx}$ goes to $\infty$, etc. To have a better idea about the properties, in particular what happens if model assumptions are violated we want to run a simulation experiment. I will help you with PROC IML codes (the handout given in the class) but you are welcome to use any software of your choice.

Let the sample size, $n$, be 30 and the $x$ values can be found from the course homepage, datasets section. Suppose we generate 1000 different samples of size $n$ from the model

$$y_i = 1 + 0.5x_i + e_i, \ i = 1, 2, \ldots, 30,$$

where $e_i$ are independently and identically distributed as $N(0, .25)$. Based on each sample compute the least squares estimates of $\beta_0(=1), \beta_1(=0.5)$ and $\sigma^2(= .25)$. Let the estimates from the $j$th sample be $(b_{0,j}, b_{1,j}, s_j^2), \ j = 1, 2, \ldots, 1000$. Presumably these are random samples from the finite sample distribution of the least squares estimators for this particular values of the population parameter and the given values of $x$. Thus, for example, according to our theory $b_{1,1}, b_{1,2}, \ldots, b_{1,1000}$ constitutes an i.i.d sample from the distribution $b_1 \sim N(0.5, 0.25/S_{xx})$. Also, the covariance of $b_{0,j}$ and $b_{1,j}$ is equal to your computed value in question 2. Compute sample mean, variance and covariances of the samples of $(b_{0,j}, b_{1,j}, s_j^2)$ to compare with your theory. Also, look at the histograms of the individuals samples of slope, intercept and error variance estimators and comment on the shape of the distribution.

Now we want to check how does violation of model assumptions affect the properties of the estimators. One critical assumption is that the errors are uncorrelated. Suppose for each sample we generate the errors using the following scheme

$$e_1 \sim N(0, 0.25/(1 - \rho^2)), \ e_i = \rho e_{i-1} + u_i, \ i = 2, 3, \ldots, 30,$$

where the $u_i$ are i.i.d. $N(0, 0.25)$ and $-1 < \rho < 1$ is a real constant. Take $\rho = -0.5, 0.5, 0.99$. Note that $\rho = 0$ gives you back your first model. Repeat the experiment with the given values of $\rho$ and comment on how the finite sample properties of the estimators depend on $\rho$. Another critical assumption is the assumption of normality. Suppose in the first model (uncorrelated errors) the errors are independently and identically distribute as $0.25\chi^2_2 - 0.5$. Repeat your experiment and compare your results with those from the normal model.