UST 803
MONTE-CARLO SIMULATIONS
COMPUTER LAB #2

Monte Carlo Simulation of Linear Regression Model: N is number of data points, X is the independent (explanatory) variable, Y is the dependent variable, u is the stochastic disturbance, β_1 is the Y intercept and β_2 is the slope. The sample estimates of the parameters are: βhat1 and βhat2.
We create a vector of random numbers from a normal distribution with a specified mean μ and standard deviation σ.

Enter the mean m: $\mu := 0$

Enter the standard deviation s: $\sigma := 4$

The random deviates: $u := \mathrm{rnorm}(N, \mu, \sigma)$

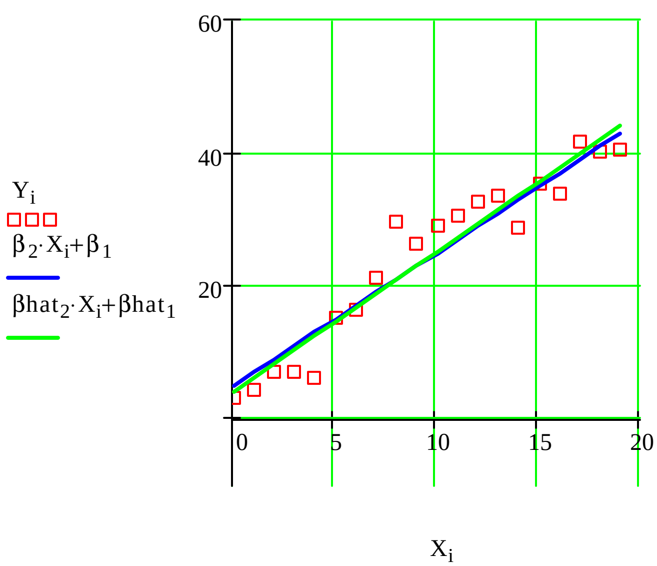
$\beta_2 := 2$ $\beta_1 := 5$

$i := 0 .. N - 1$

$X_i := i$

$Y_i := \beta_2 \cdot X_i + \beta_1 + u_i$

$\beta hat_1 := \mathrm{intercept}(X, Y)$ $\beta hat_2 := \mathrm{slope}(X, Y)$



Enter the number of random deviates N. Then click on N=20 below and click F9 (compute). Type the estimates bhat in two arrays. Repeat the experiment 10 times. Each time record the estimates.

$N \equiv 20$

$\beta hat_1 = 4.088$

$\beta hat_2 = 2.113$

Below are my estimates.

$$\beta hat_1 := \begin{pmatrix} 4.088 \\ 4.563 \\ 6.268 \\ 5.906 \\ 3.086 \\ 6.091 \\ 7.565 \\ 4.481 \\ 6.954 \\ 4.896 \end{pmatrix} \qquad \beta hat_2 := \begin{pmatrix} 2.113 \\ 1.951 \\ 1.878 \\ 1.806 \\ 2.050 \\ 1.876 \\ 1.701 \\ 2.108 \\ 1.963 \\ 2.071 \end{pmatrix}$$

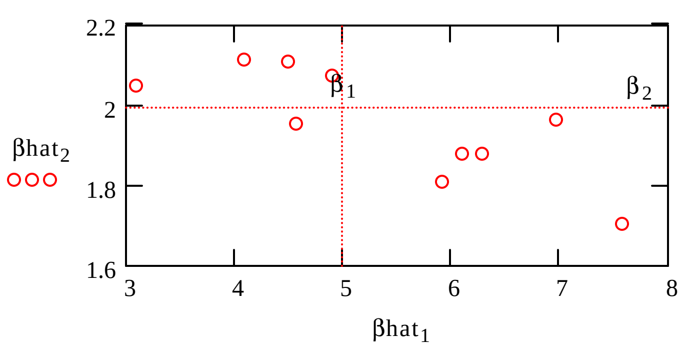
Next we calculate the means of the sampling estimates and compare them to the population values: β_1 = 5 and β_2 =2. This is an experimental check on the poperty of OLS estimates to be unbiased. We then compute the standard deviations for thos estimates and the covariance. Note that it is negative because the mean of X is positive.

$\mathrm{mean}(\beta hat_1) = 5.390$ $\qquad$ $\mathrm{mean}(\beta hat_2) = 1.952$

$\mathrm{Stdev}(\beta hat_1) = 1.393$ $\qquad$ $\mathrm{Stdev}(\beta hat_2) = 0.137$

$$\mathrm{cvar}(\beta hat_1, \beta hat_2) = -0.136$$

The significance of negative covariance is that there is a tendency that: (1) the slope estimate will be larger than the population slope and the intercept estimate will be lower than the population intercept or (2) the slope estimate will be smaller than the population slope and the intercept estimate will be larger than the population intercept. You can see this in the graph below.



$r := \mathrm{corr}(X, Y)$ $\quad r = 0.961$

$$\sigma hat := \sqrt{\frac{\mathrm{length}(X)}{\mathrm{length}(X) - 2} \cdot \mathrm{var}(Y) \cdot \left(1 - r^2\right)}$$

$\sigma hat = 3.705$

$$se\beta_1(X, Y) \equiv \sqrt{\frac{\mathrm{var}(Y)}{\mathrm{var}(X)} \cdot \frac{\mathrm{mean}\left(\overrightarrow{X^2}\right)}{\mathrm{length}(X) - 2} \cdot \left(1 - \mathrm{corr}(X, Y)^2\right)}$$

$$se\beta_2(X, Y) \equiv \sqrt{\frac{\mathrm{var}(Y)}{\mathrm{var}(X)} \cdot \frac{1}{\mathrm{length}(X) - 2} \cdot \left(1 - \mathrm{corr}(X, Y)^2\right)}$$

$$cov\beta1\beta2(X, Y) \equiv \frac{-\mathrm{var}(Y)}{\mathrm{var}(X)} \cdot \mathrm{mean}(X) \cdot \left(1 - \mathrm{corr}(X, Y)^2\right) \cdot \frac{1}{\mathrm{length}(X) - 2}$$

$se\beta_1(X, Y) = 1.597$ $\quad se\beta_2(X, Y) = 0.144$ $\quad cov\beta1\beta2(X, Y) = -0.196$