

AGEC 643
Homework 1
Due Sept. 25, 2015

Be sure to document all cells with formulas that you type using the =VFORMULA() function.

1. Use the data in 2015HWK1.XLSX, worksheet Data 1, and answer the steps below. Print the answers like an economist, i.e., minimize pages of paper.
 - a. Calculate the summary statistics for all variables
 - b. Estimate trend regressions
 - c. Forecast variables for 5 years using trend regression
 - d. Calculate the correlation matrix for the variables
 - e. Calculate the covariance matrix for the variables
 - f. Calculate the square root of the correlation matrix; call it R
 - g. Transpose the R matrix
 - h. Calculate the product of RR' , check if it equals the original correlation matrix and print both matrices
 - i. Sort the original data on column 1 from low to high, carrying along all the other columns
 - j. Convert the matrix of data to a column vector
 - k. Next reverse the steps in j and create a matrix from the column vector
 - l. Create a diagonal matrix (10 x 10) of 1's

2. Use the data in 2015HWK1.XLSX, workbook Data 3.
Use Theta for the univariate parameter estimator to calculate the parameters for the Y variable; use the MLE option. Add the EMP distribution and be sure to add in the common random number in the EMP function. Simulate the distributions for 500 iterations and then use CDFDEV to determine the best distribution to simulate the random variable Y. Report the CDFDEV stats and the UPES parameters in a table on one page.

Calculate the parameters for normal, uniform, gamma, and exponential by hand. Did you get the same answers?

3. Use the data in 2015HWK1.XLSX, workbook Data 2 and run the regression below.

Estimate the multiple regression model using the Simetar function. Forecast the dependent variable for three years given these exogenous values below and report the Y values:

	<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>
X ₁	4	3	3.5
X ₂	74	75	76
X ₃	40	41	42
X ₄	9.7	9.7	9.7

X₅ 3.9 2.9 2.7

- a. Simulate the dependent variables for 3 years using the Y-hats in step b using the std. dev. of the residuals, assuming the residuals are normally distributed. Program your own formulas using $Y\text{-hat} + \text{Std Dev} * \text{SND}$, i.e., do not use the random value cells (NORM) in the regression, I want you to do it on your own.
 - b. Report the PDF and CDF of the second year's values in steps c.
 - c. Next calculate the beta-hat vector for the multiple regression using matrix algebra. Be sure to include an intercept.
4. Develop a simple simulation model to estimate the PDF and CDF of total net returns for a business. This is a fun example of what can be done easily with simulation. Assume it is a farm with three crops. Use the data in 2015HWK1.XLSX Data 4 to estimate the yield and price distributions. Simulate these random variables as Empirical (independent). You have to estimate the EMP parameters by hand. Estimate them as Fractions from the Mean.

	<u>X₁</u>	<u>X₂</u>	<u>X₃</u>
Acres Planted to each Crop	25	50	40
Average Yield	60	20	42
Average Price	7.00	6.00	4.50
Variable Cost/Bushel	3.00	6.00	4.00
Fixed Cost (\$/acre)	100.	90.	30.

KOVs:

- Calculate total revenue
- Calculate total costs
- Calculate total net returns

Simulate the model and report summary statistics for all stochastic variables and the KOVs. Print the Parameter estimation for the stochastic variables, CDF and PDF for the net returns distribution.

5. I want you to test the power of Latin hypercube simulation for reducing the number of iterations necessary to reduce the number of iterations required to get an accurate sample. Do the following 5 simulations using UNIFORM(0,1) and RAND():
 - 100 iterations
 - 200 iterations
 - 500 iterations
 - 1000 iterations
 - 5000 iterations

For each sample print the summary statistics and a CDF of the results on the same sheet of paper. Actually if you can put two or more samples on one sheet I would prefer it.

Next I want you to test the 5,000 iteration test of Rand and Uniform for autocorrelation. Here is how to do this: For each sampling procedure break the sample into 10 columns of data, i.e., 1-500 in column A, 501-1000 in column B, and so on till 4500-5000 in the last column. Calculate a correlation matrix on the 10 columns. Next run a regression using column 10 as the endogenous variable and columns 1-9 as the exogenous variables.

Print the correlation matrices and the regression results and explain the results.