Economics 487

Homework #2 Solution Key
Random Variables and the Constant Expected Return Model

A. Excel Exercises: *(10 points)*

1. Download the Excel file "hw2a.xls". The file contains the probability distributions for four discrete random variables. Think of the random variables as representing the future annual return on A.G. Edwards, Anaheuser Busch, Boeing, and Monsanto stock. The distributions of these returns are based on (i.e., conditional on) the future state of the economy (e.g., depression, recession, normal, mild boom, or major boom). To help you learn more about Excel, the answers to the below questions have been answered for the A.G. Edwards case (R_A), although you should still comment on the A.G. Edwards results.
   a. Plot the distributions for each random variable (make a bar chart). Paste to a Word file. Comment on any differences or similarities between the distributions.
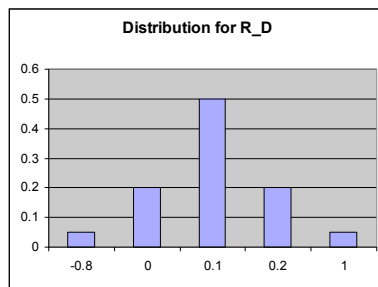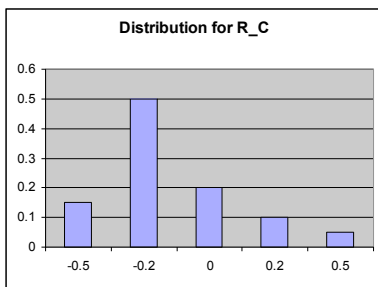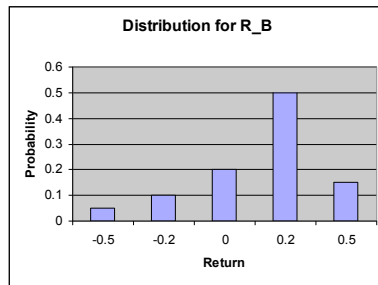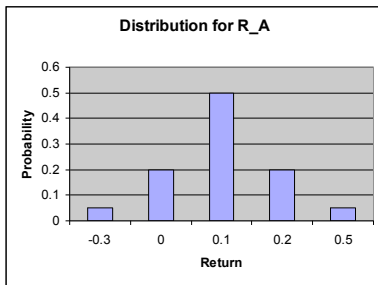
A and D are fairly similar in that they are symmetric around the mean, although D is more spread out. B and C are mirror images of each other, with B being negatively skewed (lopsided) and C being positively skewed.

   b. For each random variable, compute the expected value, variance, standard deviation, skewness, and kurtosis. Paste your table of these summary statistics to a Word file and comment.

Distribution Shape	R_A	R_B	R_C	R_D
Mean	0.100	0.130	-0.130	0.100
Variance	0.020	0.057	0.057	0.085
Std. Deivation	0.141	0.239	0.239	0.292
Skewness	0.000	-0.642	0.642	0.000
Kurtosis	6.500	3.663	3.663	9.087

The numbers confirm the visual impressions.

2. Suppose X is a Normally distributed random variables with mean 10 and variance 24. I.e., $X \sim N(10,24)$.
   a. Using Excel, calculate $Pr[X>14]$. Write out the Excel function you used to make this calculation.
   b. Calculate $Pr[8<X<20]$. Again, write out the Excel function.
   c. Find the probability that X takes on a value that is at least plus or minus 6 away from its mean. Again, write out how you arrive at your answer.
   d. Suppose y is a constant defined such that $Pr[X>y]=0.10$. What is y? Again, explain how you arrived at this answer.
   Hint: Check out the description of functions NORMDIST and NORMINV using the Microsoft Excel Help feature.

   A.   0.207108
   B.   0.637841
   C.   0.220671
   D.   16.27829

   A. =1-NORMDIST(14,10,SQRT(24),TRUE)
   B. =NORMDIST(20,10,SQRT(24),TRUE)-NORMDIST(8,10,SQRT(24),TRUE)
   C. =1-(NORMDIST(10+6,10,SQRT(24),TRUE)-NORMDIST(10-6,10,SQRT(24),TRUE))
   D. =NORMINV(0.9,10,SQRT(24))

3. Consider the possible rates of return that you might obtain over the next year. You can invest in stock A or stock B.

| State of economy	Probability	Return on stock A	Return on stock B
Recession	0.3	-10%	0%
Normal	0.4	20%	10%
Boom	0.3	50%	10%

   a. Determine the expected return, variance, and standard deviation for stocks A and B.
   b. Determine the covariance and correlation between the returns on stocks A and B.
   c. Determine the expected return and standard deviation of an equally weighted portfolio of stocks A and B. *Hint: an equally weighted portfolio return is calculated by adding 50% of the return for stock A to 50% of the return to stock*

B. Think of it as having $50 in asset A and $50 is asset B instead of $100 in one or the other.

Prob	A	B	EW Portfolio
0.3	-0.1	0	-0.05
0.4	0.2	0.1	0.15
0.3	0.5	0.1	0.3

mean	0.2	0.07	0.135
variance	0.054	0.0021	0.018525
std. Dev.	0.232379001	0.045825757	0.136106576
covariance	0.009		
correlation	0.845154255		

4. Download the Excel file "hw2b.xls". The file contains price data for Microsoft and IBM stock. Compute the continuously compounded monthly returns $r_t = \ln(P_t / P_{t-1})$.

   a. Make a time plot of monthly returns (use basic line graph with no markers), putting the returns for Microsoft and IBM on the same plot. Paste to Word and comment on the salient features of the plot.
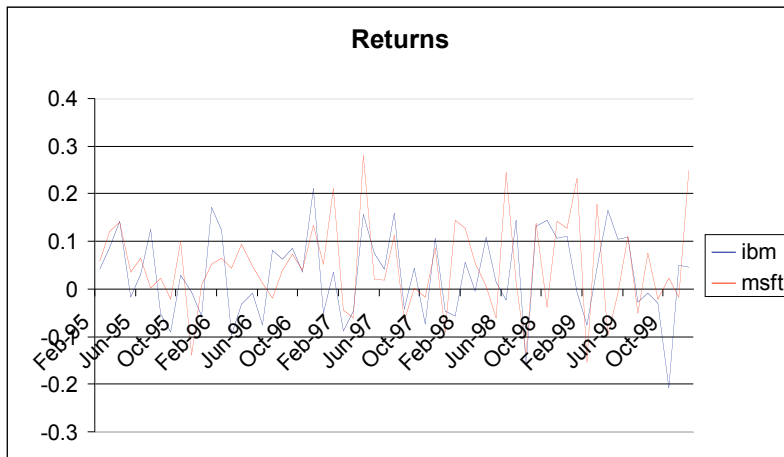

Both series appear individually quite random, but they do seem to fluctuate together somewhat. MSFT is noticeably more volatile than IBM.

   b. Make a scatter plot of the monthly returns on Microsoft against the returns on IBM. Paste. Briefly comment on the direction and strength of a linear association between the returns on the two stocks.
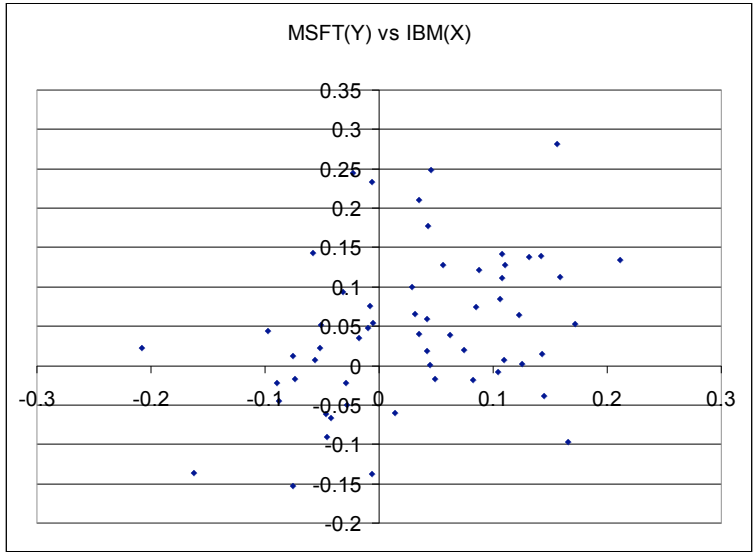
Most of the returns are in the upper right quadrant (+,+). There appears to be a weak positive linear relationship between the two return series.

5. Use **Tools/Data Analysis/Histogram** to compute histograms of the returns on Microsoft and IBM. Paste. Comment on the shape characteristics of the histograms (i.e., center, spread, symmetry, and tail thickness). *Note: you can use the default bins given by Excel.*
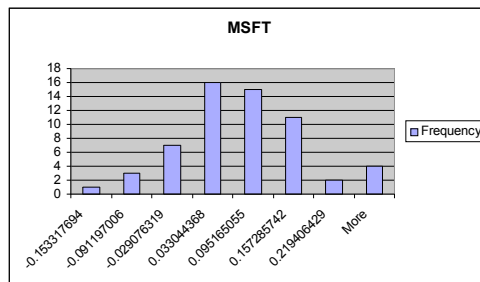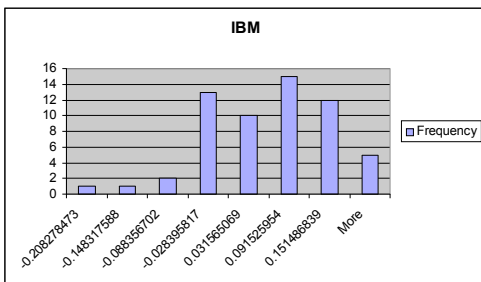
*MSFT appears more symmetric. It also has more large positive observations. IBM appears more uniform near the mean.*

6. Compute sample descriptive statistics for the historical returns on Microsoft and IBM. Use Excel functions AVERAGE, VAR, STDEV, SKEW, KURT, COVAR, and CORREL to compute these statistics. Paste a table of results to Word. Briefly compare and contrast the descriptive statistics for the two stocks.

|	IBM	MSFT
Average	0.030426	0.046705
variance	0.007927	0.009283
std. Dev	0.089033	0.09635
skew	-0.22403	0.28107
kurt	2.684012	3.041439
covar	0.002916	
correl	0.345834	

The mean and standard deviation for MSFT are higher than for IBM. In both cases, the sample kurtosis is close to the true value for a normal distribution (i.e., 3). The positive correlation coefficient confirms a reasonably large positive linear relationship between the two returns evident in the scatterplot above.

7. Using **Tools/Data Analysis/Random Number Generation** simulate 60 observations (number of variables is 1 and number of random numbers is 60) from the constant expected return (CER) model for Microsoft only. Assume that the mean of the normal distribution is equal to the sample mean and the standard deviation is equal to the sample standard deviation.
   a. Make a time plot of the simulated monthly returns. Paste. Briefly comment on the salient features of the plot. Does the simulated data look like the actual data?
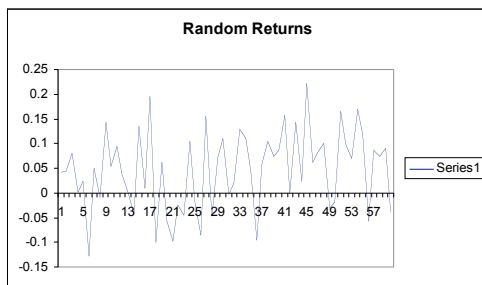
Everyone's simulated series will look somewhat different, but the general properties will be the same. The returns will typically be positive and appear random, with more realizations closer to the mean than far away. This looks similar to the actual data.

   b. Compute a histogram of the simulated data. Paste. Does it look like the histogram of the actual data?
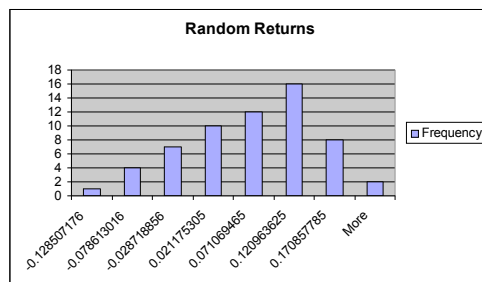
Again, everyone's histogram will look somewhat different. But, it should look like it could come from a normal distribution, without looking exactly like a "bell curve."

c. Compute descriptive statistics for the simulated data. Paste. Are these similar to the descriptive statistics for the actual data?

	Random Returns	MSFT
Average	0.046812	0.046705
variance	0.006382	0.009283
std. Dev	0.079885	0.09635
skew	-0.14313	0.28107
kurt	2.463388	3.041439

These sample statistics are fairly similar. Interestingly, even though we know the random returns really are normal, the sample kurtosis for the MSFT returns is closer to the true kurtosis for the normal case than the random returns.

8. Using the sample statistics computed in exercise 6, give estimates of the parameters of the CER model:

$$R_{it} = \mu_i + \varepsilon_{it}, \quad i = \text{Microsoft, IBM}$$
$$\varepsilon_{it} \sim i.i.d.N(0, \sigma_i^2)$$
$$corr(\varepsilon_{it}, \varepsilon_{jt}) = \rho_{ij}$$

a. Compute estimated standard errors and 95% confidence intervals for the μ_i's, σ_i^2's, and ρ_{ij} using the formulas given in class. Comment on the precision of your estimates. *Hint: You can calculate the least squared estimates for mean and variance using the Excel functions AVERAGE, VAR, and COVAR. For example, AVERAGE gives you the sample mean, which is the least squares estimate of the true population mean. You can say how you arrived at the point estimates using the Excel functions, but you do not have to derive the equivalence of the sample mean to the least squares estimate. That is, you can just state that the least squares estimate is the sample mean. In terms of calculating the 95% confidence bands for, say, mu, "mu+/-2*Standard Error(mu)" is the correct formula. Remember the 2 refers to the critical value for a t-distribution with 58 degrees of freedom. That is, only 5% of observations from a t-distributed random variable are greater than 2 in absolute value. This is pretty similar to a standard normal distribution, where only 5% of observations are greater than 1.96 in absolute value. The slightly higher critical value reflects the greater variance in a statistic that makes use of an estimated standard error rather than the true standard error (which has no variance, but is just some number).*

	IBM	MSFT
Average	**0.030426**	**0.046705**
SE	0.011591	0.012544
95% lower	*0.007243*	*0.021618*
95% upper	*0.053608*	*0.071793*
variance	**0.007927**	**0.009283**
SE	0.001472	0.001724
95% lower	*0.004983*	*0.005836*
95% upper	*0.010871*	*0.012731*
correl	**0.345834**	
SE	0.115602	
95% lower	*0.11463*	
95% upper	*0.577038*	

The means and standard deviations are not particularly precise, but the positive correlation is reasonably precise.