ECONOMICS 422

REVIEW OF PROBABILITY AND RANDOM VARIABLES

R. W. Parks

I assume that you have had exposure to basic probability theory and statistics in prerequisite courses. Chapter 4 from Haley and Schall (in the readings packet) provides a review of this material emphasizing those ideas that are especially useful in this course. The textbook reviews some of these concepts as well. Give particular attention to the following topics:

1. Probability Distributions and Random Variables.

What two sets of information are required to specify the distribution of the random variable X? What is the distinction between discrete and continuous random variables? How do their probability distributions differ?

2. Expectations and Moments

You need to know how to compute means and variances from the information provided by the probability distribution. Note that this is not the same thing as computing <u>an estimate</u> of the mean or variance based on a sample. Be sure to learn the rules for computing the means and variances of simple linear functions of a random variable.

3. Combinations of Random Variables

This material is probably the least familiar, but it is the material that will be most used in 422. Study it with particular care. Work through the examples in Haley and Schall. Ask questions if there is something that you don't understand.

4. Expected utility will be covered in class and there is a chapter from Varian in the reading packet. This material is not covered in the textbook.

Test your understanding by working the set of review problems.

Review Problems on Random Variables

1. A random variable X has the following distribution:

Row 1	Possible values of X:	0	1	2	3	4
Row 2	Probabilities	.2	.1	.2	.3	.2
Row 3	Values of Y=3X-1:					
Row 4	Values of $Q=X^2$					

a. What two sets of information are required to specify the distribution of the random variable X?
b. Compute E (X).
c. Compute V(X).
d. What is the probability distribution of the new random variable Y=3X-1. [Hint: see part a.]
e. Compute E(Y) and V(Y).
f. Find the probability distribution of the random variable X^2 and compute $E(X^2)$.
g. What is the value of $E(9X^2-6X+1)$. [Hint: the easy way is to use the results above and the simple theorems.]
h. What is the value of $E[(X-E(X))^2]$? What is another name for this expression? What does it measure?

2. A producer has the following cost function:

Output X	500	700	1000	1500
Average Cost AC	2	3.5	6	9

Price is a random variable with the probability distribution:

Price P	10	12	15	20
Probabilities	.5	.3	.1	.1

The producer wants to maximize expected profits. What output should he/she choose?

3. An investor can choose between assets A1 and A2. The probability distributions of the rates of return for these assets are given below.

R1 (%)	10	15	17	20
Probabilities	.4	.3	.2	.1

and

R2 (%)	10	15	17	20
Probabilities	.3	.3	.2	.2

a. Compute the expected return for both assets.
b. Compute the variance for both assets.
c. Based on the expected return vs. variance criterion, which asset would you prefer and why? Do you see any other basis for preferring one asset to the other?
d. Compute the expected return and variance of a portfolio that has portfolio weights 0.4 and 0.6.

4. Random variables X and Y have the following joint distribution:

X	Y	Prob(X,Y)
0	1	.4
0	-1	.3
1	0	.2
1	1	.1

a. First convert this information into a 2-way table showing the joint X,Y probability distribution.
b. Compute E(X), E(Y), V(X), V(Y),E(XY), Cov(X,Y), and the correlation coefficient.

5. For two jointly distributed random variables X and Y (not necessarily those of problem 4),
a. If V(X)=0 <u>or</u> V(Y)=0, what can you say about Cov(X,Y)? Explain.
b. Show, using the definition, that Cov(aX,bY)=abCov(X,Y).
c. Show, using the definition, that Cov(X+Y,Z)= Cov(X,Z)+Cov(Y,Z).

6. To test your understanding of expected utility concepts, try the following problem.

A consumer has the utility function u=U(W)=ln(W) where u represents the level of utility, U(W) represents the general form of a utility function, W is the consumer's wealth, and ln(W) represents a specific utility function that specifies that the value of the utility index associated with a given level of wealth, W, is given by the natural log (base e) of W.

The consumer's wealth is not fixed, but is a random variable given by the probability distribution:

Possible value of W:	100,000	150,000	200,000	250,000
Probability of W:	0.2	.5	0.2	0.1

a. Compute the consumer's expected wealth i.e. E(W).

b. Extend the table to show the possible values of U(W).
c. Compute the expected utility of wealth, i.e. .E[U(W)]
d. Prove that E[U(W)]≠U(E(W)).
e. Find the certainty equivalent wealth, that is the wealth, which if held for certain, would be regarded as equivalent in the utility sense to the random wealth distribution given above. i.e. find W_c such that $U(W_c)$=E[U(W)].

Economics 422: Random Variable Review

Combinations of Random Variables

Consider the following joint probability distribution for random variables X and Y. To be concrete you may allow X to represent the price next period of firm 1's stock and Y to represent the price next period of firm 2's stock. Suppose that there are 16 possible states of the world, and for each state of the world there is a corresponding X value and a Y value. The joint (X,Y) distribution is given by the following table.

		Y	Price of	firm 2		
		30	40	50	60	Marginal X
X	30	.01	.05	.04	.05	.15
Price of	40	.05	.10	.10	.05	.30
firm 1	50	.10	.15	.10	.05	.40
	60	.05	.06	.03	.01	.15
	Marginal Y	.21	.36	.27	.16	1.0

If you are interested in X alone (or in Y alone), the univariate distribution of X is given by its list of values together with the marginal X probabilities. From this X distribution you can compute E(X)=45.5, V(X)=84.75. Similarly for Y we have E(Y)=43.80 and V(Y)=97.56.
Make sure that you can reproduce these calculations.

Now suppose that you have a portfolio involving one share each of firm 1 and firm 2. The value of your portfolio is the given by the random variable Z defined as Z=X+Y. For each of the 16 states of the world there is an X value and a Y value, hence there is a Z value. We can construct a list of the Z values for each of the 16 states (row-wise):

Z=	X+Y
60	.01
70	.05
80	.04
90	.05
70	.05
80	.10
90	.10
100	.05
80	.10
90	.15
100	.10
110	.05
90	.05
100	.06
110	.03
120	.01
	1.00

Although this table contains all the relevant information about Z, it is not strictly speaking Z's probability distribution because the list of possible values contains duplicates. By combining the duplicates we obtain Z's probability distribution:

Z values	Prob's
60	.01
70	.10
80	.24
90	.35
100	.21
110	.08
120	.01
	1.0

From this distribution we can compute E(Z)=89.30 and V(Z)= 134.51 in the usual way. ***Be sure to check this.***

We are interested in the relationship between the moments of the original X and Y variables and those of Z. The theorems regarding expectations and variances tell us that:

E(Z)=E(X)+E(Y) ***Verify this***

V(Z)=V(X)+V(Y)+2Cov(X,Y).

Notice that V(Z)<V(X)+V(Y) as computed above. This is explained by the fact that for the joint distribution of X and Y given above, the covariance between these two variables is negative, i.e. there is a tendency for Y to be below its mean when X is above its mean and vice versa. As a consequence there is a tendency for Y's variability to offset X's and vice versa. Thus the variance of their sum is less than the sum of their variances. This example should give you a sense of why covariance is important in this kind of problem.

To compute the covariance, you can use the definition or its equivalent expression:

Cov(X,Y)=E[(X-E(X))(Y-E(Y))]=E(XY)-E(X)E(Y). The latter is usually simpler.

To compute E(XY) we need the probability weighted sum of the XY products for each of the 16 cells in the X,Y distribution:

$$E(XY) = \sum_i \sum_j x_i y_j \cdot \Pr(X = x_i, Y = y_j)$$

=900(.01)+1200(.05)+1500(.04) etc. for all 16 cells=1969.

Cov(X,Y)=E(XY)-E(X)E(Y)=1969-(45.5)(43.8)=-23.9

Now verify the above expression for the variance of Z.

There are times when it is convenient to use a measure of linear association between two variables that does not depend on units as the covariance does. The correlation coefficient is defined as

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

where the denominator is the product of the standard deviations for the two variables. The correlation coefficient takes values in the interval [-1,+1], and its sign, as with the covariance, indicates whether the variables have a positive or negative (inverse) association.