**The Binomial Distribution**

Introduction: Go to *D2L > Drop Box > Classwork* for today. Download the Excel spreadsheet workbook. Save all of your work in this Excel Workbook. Save the workbook as .xls file extension and title it "*CW7.2\_yourname.*xls", dropping the quotation marks and replacing *yourname* with your own name. When you are finished with the assignment, go to *D2L > Drop Box > Classwork* and upload your workbook.

Preamble: Recall a binomial random variable, **X**, is defined as the number of successes in *n* independent trials where each trial has a probability *p* of being a success.

For example, suppose the probability of a student passing an exam is 0.7 (p, *probability of success*)and we sample 5 students (n, *number of trials*). Further suppose we are interested in the probability exactly 2 of the students pass the exam. In other words, if we let *s* stand for success and *f* stand for failure, we are interested in the probability of the sequence

*s-s-f-f-f*

Since the probability of success is the complement of the probability of failure (i.e., probability is failure is 1 – 0.7) and since each trial is independent from the others, the probability of this sequence can be found using the multiplication law of probability,

(0.7)\*(0.7)\*(1-0.7)\*(1-0.7)\*(1-0.7) =

(0.7)2 \* (1-0.7)5 – 2

However, this does not take into account the different ways of obtaining 2 success from 5 trials. (We could for instance get the sequence *s-f-s-f-f*, *s-f-f-s-f*, etc. ) From the combination formula, the number of ways to select 2 objects from a total of 5 objects is given by,

If we put these two pieces of information together, we can calculate the probability of exactly two students passing the exam as,

P(**X** = 2 successes) = \* (0.7)2 \* (1-0.7)5 – 2

= 0.1323

= 13.23 %

In general, the probability a binomial random variable **X** with parameters *n* and *p* equals a certain value *x* is given by,

P(**X** = *x*) (*Formula 1)*

For this classwork, you will also need to recall the probability definition of the expected value of a random variable and the shortcut formula for the expected value of a binomial random variable. We will use both formulas to calculate the expected value of a binomial random variable in this classwork.

The probability definition for the expected value of a random variable is,

(*Formula 2*)

The shortcut formula for the expected value of a binomial random variable is,

(*Formula 3*)

We will also be calculating the variance and standard deviation of a binomial distribution, so we will need to remember the probability definition for the variance of a random variable and the shortcut formula for the variance of a binomial random variable.

The probability definition for the variance of a random variable is,

(*Formula 4*)

The shortcut formula for the variance of a binomial random variable is,

(*Formula 5*)

Instructions: On D2Lyou will find an Excel template for this assignment. You will be using Excel to calculate the probability distribution of a binomial random variable for various values of the parameter *n*. You will then plot the resulting distribution’s histogram.

The first column of the spreadsheet gives you number of successes whose probability you are trying to find.

In the next column *n C x,* you will use the COMBIN(a, b) function to find the appropriate combination for this value of x. This is the first term in Formula 1. Note, the COMBIN( ) function has two arguments:*a*, the number of distinct objects are you are selecting from and **b**, the number of object you are selecting. In this case, *a* will be the number of trials *n* and *b* will be the number of successes desired *x*.

In the next column *Success,* you will calculate the probability of success for this value of x. This is the second termin Formula 1. Note in the case of 0, this will work out to be 1 in order for the formula to check out.

In the next column *Failure*, you will calculate the probability of failure for this value of x. This is the third term in Formula 1. Note in the case of n, this will work out to be 1 in order for the formula to check out.

In the final column *P(X = x),* you will put all the pieces of Formula 1 together. After you have completed the final column, plot these values against the number of successes *x* and make a histogram.

**Note the workbook has multiple tabs at the bottom for different values of *n.* After you are finished each spreadsheet in the workbook, upload your spreadsheet in the appropriate spot on Blackboard.**

*Note:* We will do the first spreadsheet in class!

**Problems**

1. For each spreadsheet in the workbook, construct the binomial distribution for the given value of number of trials *n* and probability of success *p* by following the steps given below:

a. In Column B, use the COMBIN() function to calculate the number of ways to get x success in n trials.

b. In Column C, use cell references and formulas to calculate the probability of x successes.

c. In Column D, use cell references and formulas to calculate the probability n – x failures.

d. In Column E, multiply the results of Column B, C & D.

2. After you have constructed the binomial distribution in each spreadsheet of the workbook, insert a

probability histogram on each spreadsheet.

3. Calculate the expected value of the binomial distribution in two ways:

a. Use the probability definition of an expected value (*Formula 2*)

b. Use the shortcut formula for a binomial random variable. (*Formula 3*)

4. Calculate the variance of the binomial distribution in two ways:

a. Use the probability definition of variance. (*Formula 4*)

b. Use the shortcut formula for a binomial random variable. (*Formula 5*)