Chapter 4
Discrete Random Variables

4.1 Introduction

Chapter 4 introduces discrete random variables to the reader. Three discrete random variables, the binomial, poisson, and hypergeometric distribution, are introduced. Both individual and cumulative probabilities can be found for the discrete random variables using several different **Excel** functions. Individual probabilities are the exact probability of an outcome and will answer the "equal to" probability questions found in the text. Cumulative probabilities are the "at most" probabilities that are found in the cumulative tables listed in the text. Both are useful probabilities and will be explained below. Please note that only the binomial random variable is covered in *A First Course in Statistics*.

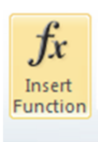
The following examples from *Statistics* are solved with **Excel** in this chapter:

Excel Companion		
Exercise	**Page**	**Statistics Text**
4.1	67	Example 4.13
4.2	70	Example 4.14
4.3	74	Example 4.15

4.2 Calculating Binomial Probabilities

To use the binomial probability function within **Excel**, we begin by opening up **Excel** and placing the cursor on any cell in the blank worksheet. We click on the **Formulas** tab and then select the **Insert Function** icon shown in Figure 4.1.

Figure 4.1



This opens the Insert Function menu shown in Figure 4.2. Click on the arrow to select that **All** categories are being used and scroll down until you reach the **BINOM.DIST** function. **Highlight** the **BINOM.DIST** function and click **OK**.

Figure 4.2

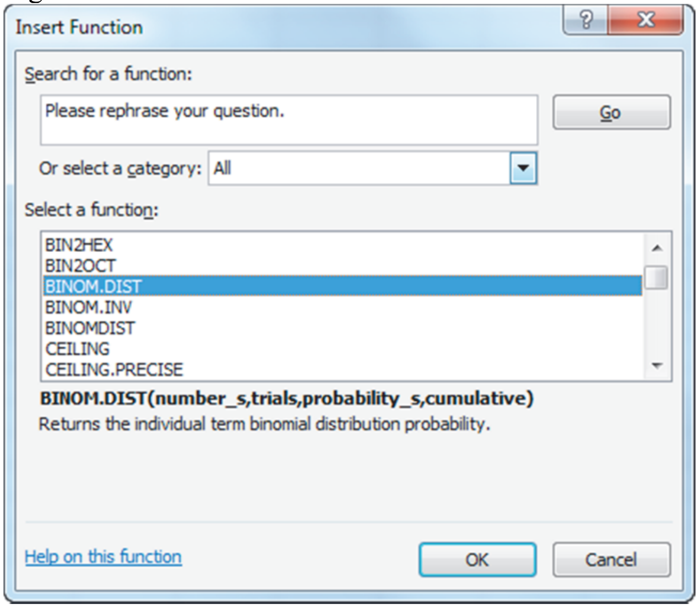
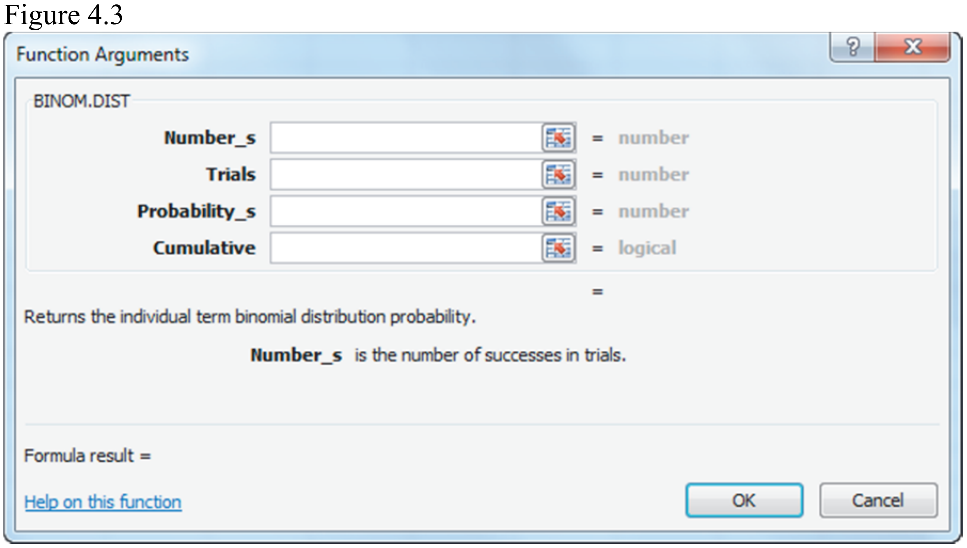


Figure 4.3



The **BINOM.DIST** function requires the user to enter the number of successes (**Number_s)**, the sample size (**Trials),** the probability of a success (**Probability_s)**, and the type of probability desired (**Cumulative – either True or False**). For most applications, the cumulative probability option should be selected (**Cumulative = True**) in order to maximize the information that **Excel** will offer. **Click OK** to finish. We illustrate with the next example.

Exercise 4.1: As an example, we turn to Example 4.13 from the *Statistics* text.

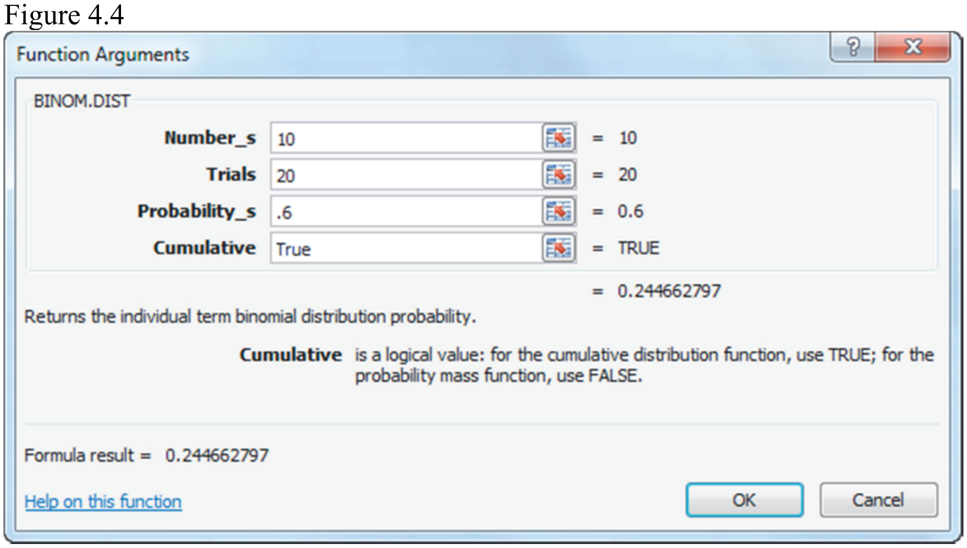
Suppose a poll of 20 employees is taken in a large company. The purpose is to determine x, the number who favor unionization. Suppose that 60% of all the company's employees favor unionization.

a. Find the mean and standard deviation of x.
b. Find the probability that $x \leq 10$.
c. Find the probability that $x > 12$.
d. Find the probability that $x = 11$.

Solution:

We utilize the **BINOM.DIST** function to solve parts b-d of this problem. We identify in the problem that the sample size is **n=20** and the probability of a success is **p=.60**. We enter both of these values in the appropriate locations in the **BINOM.DIST** menu shown if Figure 4.4. In order to solve part b, we need to find the cumulative probability for the value of x=10 successes. We specify the number of successes to be **x=10** and we enter **True** option in the Cumulative box to signify that we want a cumulative probability.

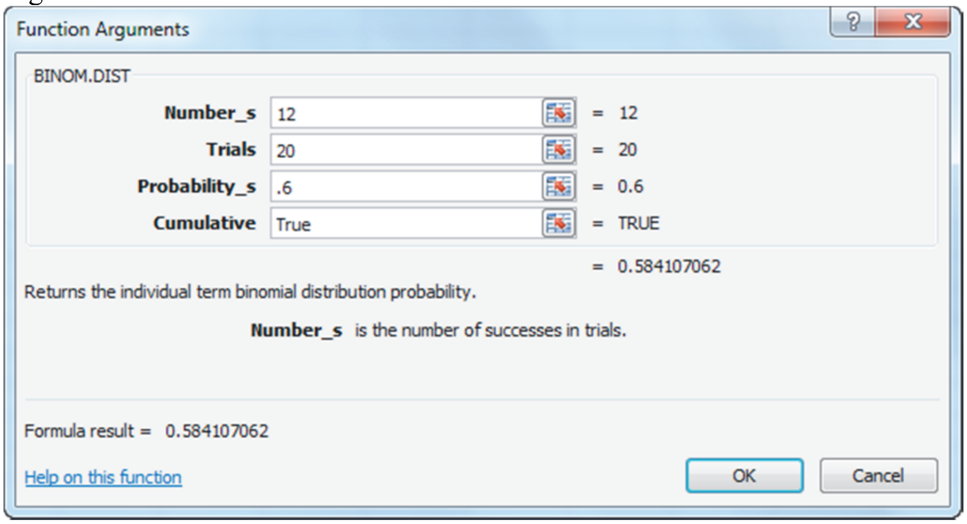
Figure 4.4



We note that the probability is shown in Figure 4.4 and then placed in our worksheet cell once we click OK. That probability is given as 0.244663.

Part c asks us to find the $P(X > 12)$. As we saw in the text, there are two methods for finding this probability. We could use the **BINOM.DIST** function and find the exact probabilities (Cumulative=False) for the values of x=13, 14, 15,…., 20 and then add the results together. A smarter and more efficient method is to recognize that the probability that we want, $P(X > 12)$, can be found by finding $1 - P(X \leq 12)$. By using the cumulative feature of the **BINOM.DIST**, we substitute the value of $X = 12$ into the **Number_s** box and find the result shown on the next page.

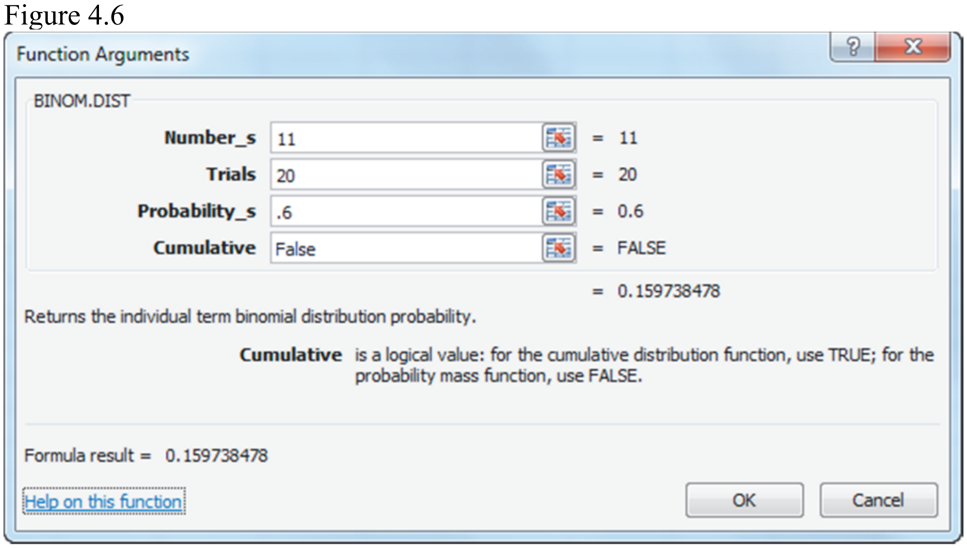
Figure 4.5



We find that $P(X > 12) = 1 - P(X \leq 12) = 1 - 0.58411 = 0.41589$.

Lastly, we find the exact probability asked for in part d by using the individual probability option (**Cumulative=False**). To find $P(X = 11)$, we use **11** as the number of successes and select **False** for the Cumulative option. The results are shown in Figure 4.6.

Figure 4.6

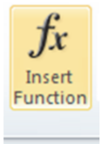


We find $P(X = 11) = 0.15974$. Compare these values found with the solutions shown in the text.

4.3 Calculating Poisson Probabilities

To use the poisson probability function within **Excel**, we begin by opening up **Excel** and placing the cursor on any cell in the blank worksheet. We click on the **Formulas** tab and then select the **Insert Function** icon shown in Figure 4.7.

Figure 4.7



This opens the Insert Function menu shown in Figure 4.8. Click on the arrow to select that **All** categories are being used and scroll down until you reach the **POISSON.DIST** function. **Highlight** the **POISSON.DIST** function and click **OK**.

Figure 4.8

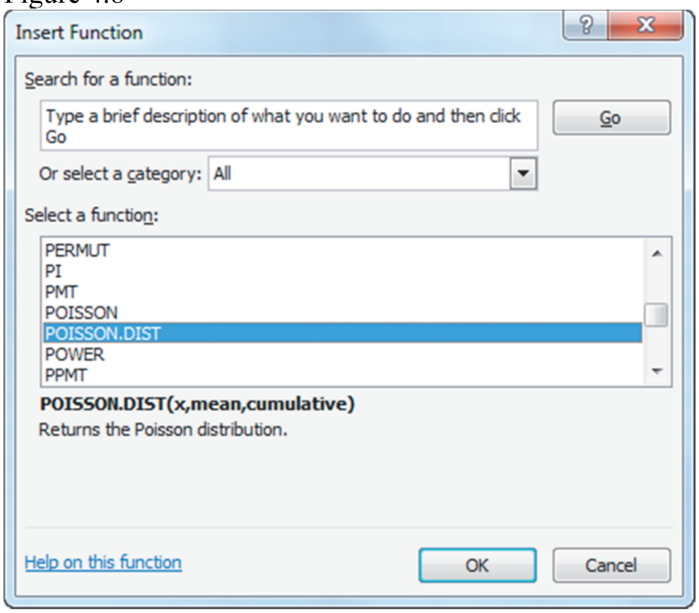
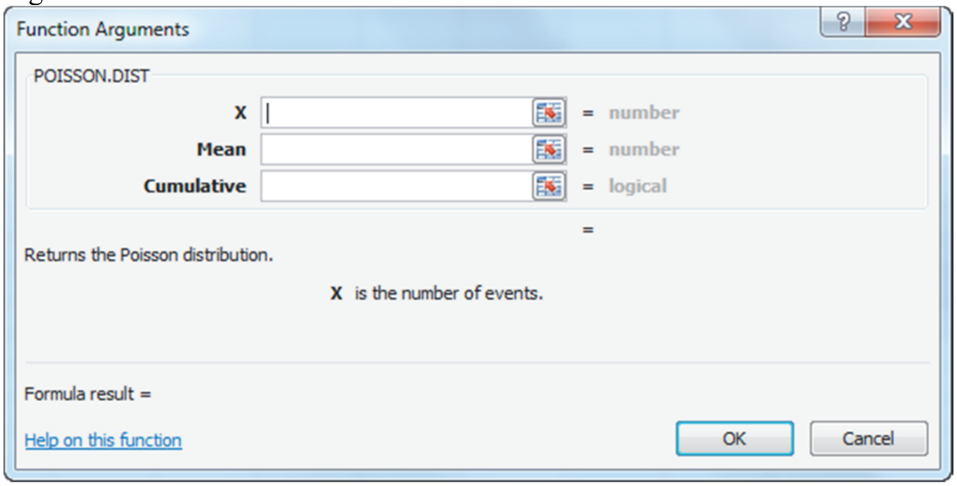


Figure 4.9



The **POISSON.DIST** function requires the user to enter a value of X to find a probability for, (**X**), the mean of the distribution (**Mean),** and the type of probability desired (**Cumulative – either True or False**). For most applications, the cumulative probability option should be selected (**Cumulative = True**) in order to maximize the information that **Excel** will offer. **Click OK** to finish. We illustrate with the next example.

Exercise 4.2: As an example, we turn to Example 4.14 from the *Statistics* text.

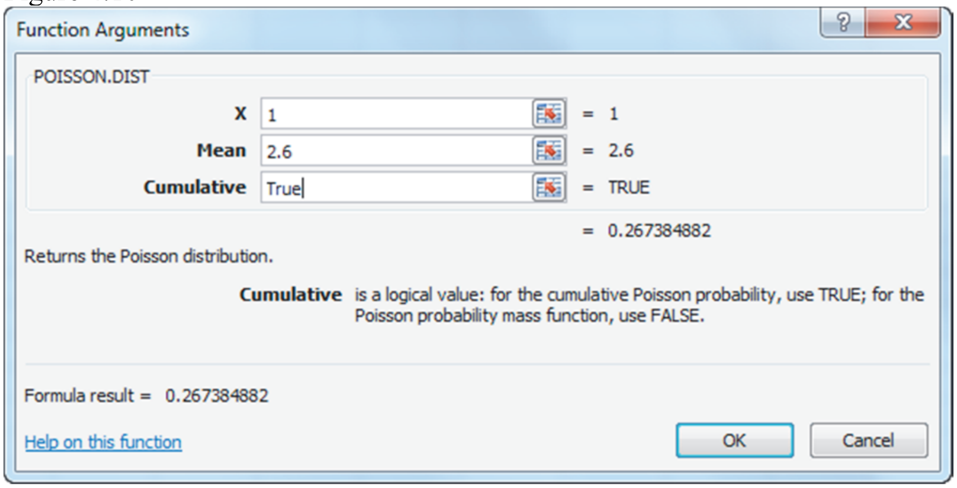
Ecologists often use the number of reported sightings of a rare species of animal to estimate the remaining population size. For example, suppose the number x of reported sightings per week of blue whales is recorded. Assume that x has (approximately) a **POISSON.DIST** probability distribution. Furthermore, assume that the average number of weekly sightings is 2.6.

a. Find the mean and standard deviation of x, the number of blue-whale sightings per week.
b. Find the probability that fewer than two sightings are made during a given week.
c. Find the probability that more than five sightings are made during a given week.
d. Find the probability that exactly five sightings are made during a given week.

Solution:

We utilize the **POISSON.DIST** function to solve parts b-d of this problem. We identify in the problem that the mean of the distribution is Mean=2.6. In order to solve part b, we need to find the cumulative probability for the value of x=1 (since $P(X < 2) = P(X \leq 1)$ and we can utilize the cumulative function). We specify the **x=1** and we enter **True** option in the Cumulative box to signify that we want a cumulative probability.

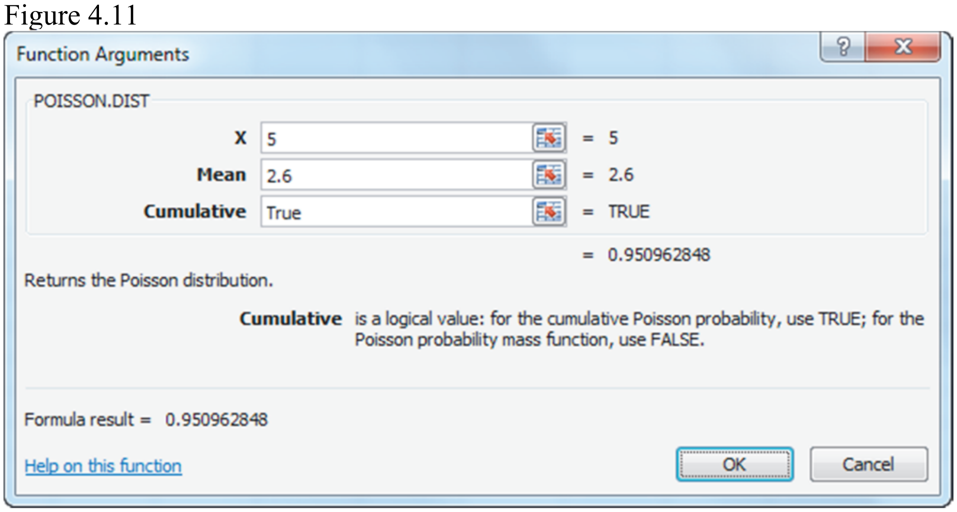
Figure 4.10



We note that the probability is shown in Figure 4.10 and then placed in our worksheet cell once we click OK. That probability is given as 0.267395.

Part c asks us to find the P(X > 5). As we saw in the text, there are two methods for finding this probability. We could use the **POISSON.DIST** function and find the exact probabilities (Cumulative=False) for the values of x=5, 6, 7, 8, … and then add the results together. A smarter and more efficient method is to recognize that the probability that we want, P(X > 5), can be found by finding 1 – P(X ≤ 5). By using the cumulative feature of the **POISSON.DIST** function, we substitute the value of **X=5** into the proper location and find the result shown in Figure 4.11.

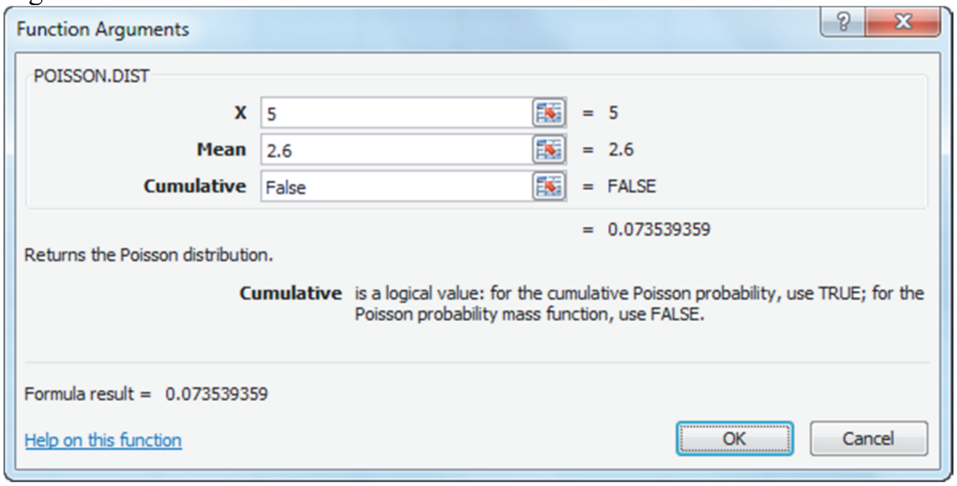
Figure 4.11



We find that P(X > 5) = 1 - P(X ≤ 5) = 1 - 0.950963 = 0.049037.

Lastly, we find the exact probability asked for in part d by using the individual probability option (**Cumulative=False**). To find P(X = 5), we use **5** as the number of successes and select **False** for the Cumulative option. The results are shown in Figure 4.12.

Figure 4.12

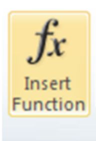


We find P(X = 5) = 0.073539. Compare this value found with the solution shown in the text.

4.4 Calculating Hypergeometric Probabilities

To use the hypergeometric probability function within **Excel**, we begin by opening up **Excel** and placing the cursor on any cell in the blank worksheet. We click on the **Formulas** tab and then select the **Insert Function** icon shown in Figure 4.13.

Figure 4.13



This opens the Insert Function menu shown in Figure 4.14. Click on the arrow to select that **All** categories are being used and scroll down until you reach the **HYPGEOM.DIST** function. **Highlight** the **HYPGEOM.DIST** function and click **OK**.

Figure 4.14

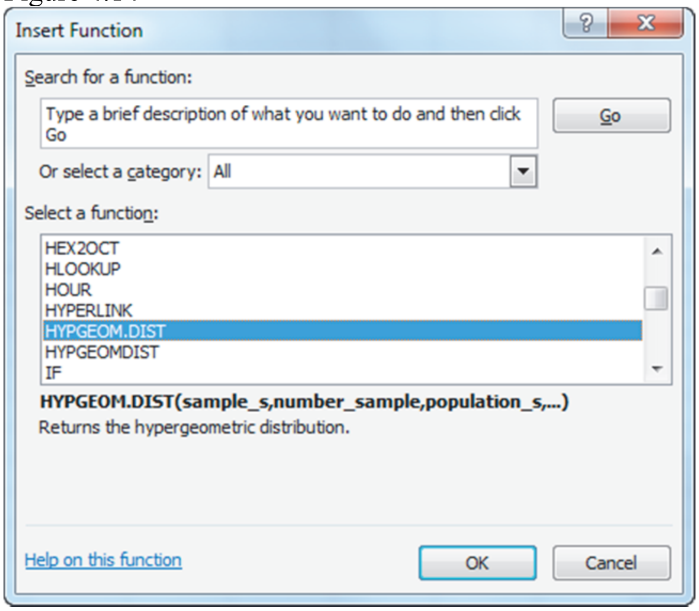
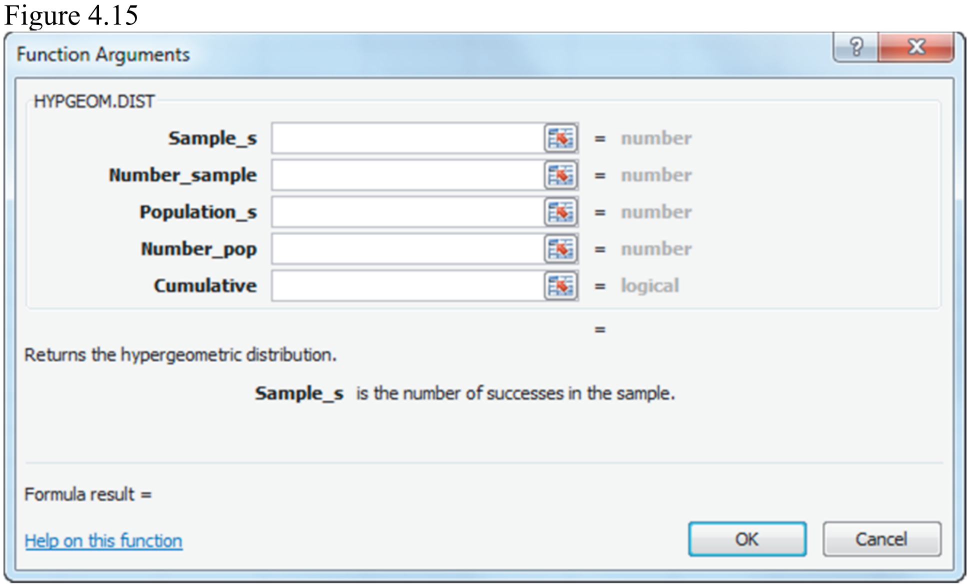


Figure 4.15



The **HYPGEOM.DIST** function requires the user to enter the size of the population (**Number_pop**), the number of successes in the population (**Population_s**), the sample size (**Number_sample),** the number of successes in the sample (**Sample_s**), and the type of probability desired (**Cumulative – either True or False**). For most applications, the cumulative probability option should be selected (**Cumulative = True**) in order to maximize the information that **Excel** will offer. **Click OK** to finish. We illustrate with the next example.

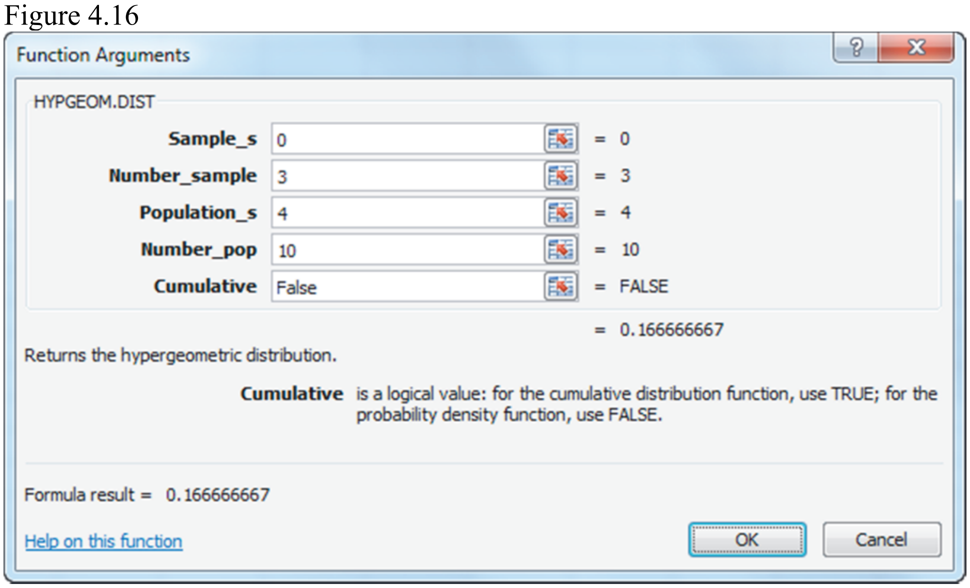
Exercise 4.3: As an example, we turn to Example 4.15 from the *Statistics* text.

Suppose a professor randomly selects three new teaching assistants from a total of 10 applicants – six male and four female students. Let x be the number of females who are hired.

a. Find the mean and standard deviation of x.
b. Find the probability that no females are hired.

To find the probability in part b, we use the hypergeometric distribution with $N = 10$ (**Number_pop=10**), $n = 3$ (**Number_sample=3**), $r = 4$ (**Population_s=4**), and **Cumulative = False**. To find the probability that no females are hired, we use $x = 0$ (**Sample_s=0**) in the **HYPGEOM.DIST** menu shown in Figure 4.16.

Figure 4.16



We note that the probability is shown in Figure 4.16 and then placed in our worksheet cell once we click OK. That probability is given as 0.1666667. Compare this probability to the one found in the text.

4.5 Technology Lab

The Technology Lab consists of problems for the student to practice the techniques presented in each lesson. Each problem is taken from the homework exercises within the *Statistics* text and includes an **Excel** data set that should be used to create the desired output. The completed output has been included with each problem so that the student can verify that he/she is generating the correct output.

1. **Where will you get your next pet?** According to an Associated Press/PEtside.com poll, half of all pet owners would get their next dog or cat from a shelter. (*USAToday*, May 12, 2010.) Consider a random sample of 10 pet owners and define x as the number of pet owners who would acquire their next dog or cat from a shelter. Assume x is a binomial random variable.

   a. Find $P(x = 7)$
   b. Find $P(x \leq 3)$
   c. Find $P(x > 8)$

   Excel Output

	A	B
1	1 a	0.117188
2	1 b	0.171875
3	1 c	0.010742

2. **Airline Fatalities.** U.S. airlines average about 1.6 fatalities per month (*Statistical Abstract of the United States: 2010*). Assume the probability distribution for x, the number of fatalities per month, can be approximated by a poisson probability distribution.

   a. What is the probability that no fatalities will occur during any given month?
   b. What is the probability that one fatality will occur during a month?
   c. Find $E(x)$ and the standard deviation of x.

   Excel Output

	A	B
1	2 a	0.20189652
2	2 b	0.32303443

3. **Testing for spoiled wine**. Suppose that you are purchasing cases of wine (12 bottles per case) and that, periodically, you select a test case to determine the adequacy of the bottles' seal. To do this, you randomly select and test 3 bottles in the case. If a case contains 1 spoiled bottle of wine, what is the probability that this bottle will turn up in your sample?

   Excel Output

	A	B
1	3	0.25