36-220 Lab#3
Probability

Week of September 19, 2005

Please write your name below, tear off this front page and give it to a teaching assistant as you leave the lab. It will be a record of your participation in the lab. Please remember to include your section (A, B or C). Keep the rest of your lab write-up as a reference for doing homework and studying for exams.

Name:

Section:

- The symbol ♣ at the beginning of a question means that, after you answer that question, you should raise your hand and have either the TA or lab assistant review your answer. Once they have reviewed your work they will place a check in the appropriate space in the table below. The purpose of this check is to be sure you have answered the question correctly.
- Try to complete as much of the lab exercise as possible. We understand that students work at different paces and have tried to structure the exercise so that it can be completed in the allotted time. If you work systematically through the handout and still don't complete every question don't worry. The important thing is that you understand what you are doing. Nonetheless, you are encouraged to complete the lab on your own.

Check-Problem ♣	Instructor's Initials
Question I(d)	
Question II(d)	
Question II(h)	
Question III(c)	

36-220 Lab #3
Discrete Probability Distributions

Part I A quality control specialist examines a shipment of 16 semiconductor wafers and either rejects or accepts each of the wafers. Suppose that in the population of wafers produced in the factory, the probability of a wafer being unacceptable is 0.38. The MINITAB command **Calc** $\succ$ **Probability Distributions** $\succ$ **Binomial** will be useful.

(a) Explain the circumstances needed for the Binomial distribution to be a reasonable model in this case.

(b) Assuming that the binomial distribution is reasonable, what is the probability that exactly 7 of the 16 wafers are unacceptable?

(c) What is the probability that 8 or fewer are unacceptable?

♣ (d) What is the probability that 8 or more are unacceptable?

(e) What is the probability that from 5 to 7 are unacceptable?

Part II

We are going to generate random data from the binomial distribution. To do this select **Calc** $\succ$ **Random Data** $\succ$ **Binomial**. In the box that appears, indicate that you want to generate 200 rows of data and store the results in column C1. Indicate that you want to have 10 trials with a .2 probability of success. Click **OK**. You now have 200 observations in column 1. Each observation is an independent random draw from a binomial distribution with $n = 10$ and $p = .2$.

1. Label the first column by typing *BIN1* in the cell at the top of the column.

2. You need to generate random data from a binomial distribution two more times. First, change the probability of success to .5 and store the results in column C2. Then, change the probability of success to .8 and store the results in column C3.

3. Label the second and third columns by typing *BIN2* and *BIN3* in the cells at the top of the columns.

4. We want to compare the distribution of the three sets of numbers. To do this, choose **Graph** $\succ$ **Histogram**. Select a histogram of type **Simple** on the next screen. Then, select all three variables *BIN1, BIN2, BIN3* in order to make a graph for each. Before exiting, click on the button **Multiple Graphs**, and select the buttons **In separate panels on the same graph**, **Same Y**, and **Same X, including same bins**. Click on **OK**, and you should see three bar charts giving the distributions of each. Note that these truly are bar charts, and not histograms as we have defined them in this course.

   (a) In the space below, sketch the distribution of each of the 3 variables.

(b) Describe the distribution of *BIN1*. Compare it to the distribution of *BIN2*.

(c) Compare the distribution of *BIN1* to the distribution of *BIN3*.

♣ (d) If the sample size for *BIN1* were very large, what would the the graph look like? If the sample sizes for both *BIN1* and *BIN3* were very large, how do you think the graphs of *BIN1* and *BIN3* would compare? (Here, "sample size" refers to the number of rows in your table, or 200.)

Select **Window** $\succ$ **Close All Graphs**. Click **OK**.

(e) Using the information about the Binomial distributions given above, calculate $\mathbf{E}[Y]$ for each of the three random variables: *BIN1*, *BIN2* and *BIN3*. This tells you the population mean for each variable. Then calculate $\mathrm{Var}(Y)$ and take the square root to find the population standard deviation for each of the three variables.

(f) In Question #5, you calculated the population mean and standard deviation based on your knowledge of n and p. Now let's compare the population values to the means and standard deviations of your three samples. Select **Stat** $\succ$ **Basic Statistics** $\succ$ **Display Descriptive Statistics**. Select *BIN1*, *BIN2* and *BIN3*. Click **OK**.

(g) How do the sample means and sample standard deviations compare to their population counterparts?

♣ (h) How do you think the sample means would compare to the population means if the sample size were 1000 instead of 200?

Part III A random variable X has the Poisson(λ) distribution if the possible values for X are $0, 1, 2, \ldots$ and

$$\Pr(X = k) = \frac{\exp(-\lambda)\,\lambda^k}{k!} \qquad k = 0, 1, 2, \ldots$$

In this case, $\mathbf{E}[X] = \lambda$ and $\mathrm{Var}(X) = \lambda$.

Suppose that the rate of emission of a sub-atomic particle from a particular specimen is, on average, 15 particles per second. Assume the emission of particles follows a Poisson distribution. The Minitab command **Calc** $\succ$ **Probability Distributions** $\succ$ **Poisson** will be useful.

(a) What value should be chosen for λ?

(b) What is the probability that 15 or fewer particles will be emitted in one second?

♣ (c) What is the probability that 15 or more particles will be emitted in one second?

(d) Find a number c such that approximately 95% of the time fewer than c particles will be emitted in one second. Hint: Begin by expressing the problem in terms of a random variable X, then use the **Inverse Cumulative Probability** option from **Calc** $\succ$ **Probability Distributions** $\succ$ **Poisson**.