MAST90105: Lab and Workshop 4

1 Lab

In this lab., Mathematica will be used because it provides very good facilities for examining the shape of distributions as the numbers which define them change.

A Mathematica notebook called Lab4.nb will be the base for you to explore the examples from Lectures this week further. Please start by opening this notebook and from the `Evaluation` menu select the item `Evaluate Notebook`. This will cause all of the commands to be evaluated - this is needed to activate the commands that you need for your work in this lab.

1. In the section of the notebook labelled Skewness and Kurtosis

   a. use the sliders in the graph to find the skewness and kurtosis recording them in the following table for the **Binomial Distribution**:

	p=0.1	p=0.3	p=0.5	p=0.7	p=0.9
n=1					
n=10					
n=50					
n=100					
n=250					
n=500					

   b. Comment on the trends you find in your table, in particular across the rows and down the columns

   c. Copy the Mathematica `Manipulate` command into the cell below the Skewness and Kurtosis graph. Alter the `Manipulate` command for the Negative Binomial distribution. Record the results in the **Negative Binomial** table:

	p=0.1	p=0.3	p=0.5	p=0.7	p=0.9
n=1					
n=10					
n=50					
n=100					
n=250					
n=500					

d. Comment on the trends you find in the table, in particular across the rows and down the columns

2. In the section of the Notebook labelled Distance between distributions for Sampling with and Without Replacement, `n` stands for the sample size, `p` for the probability of success (whatever that might be in the sampling context) and `t` for the sample size.

a. Use the sliders, to take the sample size `n` to be 500, 1000 and 3000. Record the three distances between the distributions for each of the combinations in the table:

	p=0.1	p=0.3	p=0.5	p=0.7	p=0.9
t=n+100					
n=n+500					
n=1,500					
n=n+2,000					
n=n+10,000					
n=16,000,000					

b. Comment on the trends in the table, in particular across the rows and columns

c. Looking at the Mathematica code, what is the purpose of defining the variables `lower` and `upper`? In particular, what is the significance of `n*p` and `(n*p*(1-p))hat5`? (Hint: look at the definitions of `hyperg` and `binom` and then work out which entries from these vectors are plotted and which are used to compute the distance)

d. Why is the variable `total` introduced?

e. (*challenging*) Can you relate the formula for `mpd` to what was claimed in lectures, namely the maximum difference between probabilities for the Binomial and the Hypergeometric? (hint: use the fact that probabilities add to one, and think about the set of numbers for which the Hypergeomtric probability is greater than the Binomial probability and vice-versa.)

3. **The objective of this question is to use Mathematica to look at the actual distance between Binomial and Poisson distributions.**

   a. The bound in Lectures said that the maximum difference between any Binomial and the corresponding Poisson probability, denoted D in the Mathematica plots, was bounded by p, the probability of success in the Binomial.

   b. Choosing a variety of values for p between 0 and 1, record the values of the maximum probability difference, D, for different values of n. Comment, particularly on whether the relationship with n is monotone.

   c. Find the maximum difference that you can between D and p and write it down. What do the plots show about the guidance that is given in many textbooks that "the Poisson approximation to Binomial can be used when n is large, p is small and np is moderate"?

   d. Try to find n and p that minimise the ratio between p and p. Report on your findings.

2 Workshop

1. A warranty is written on a product worth \$10,000 so that the buyer is given \$8000 if it fails in the first year, \$6000 if it fails in the second, \$4000 if it fails in the third, \$2000 if it fails in the fourth, and zero after that. Its probability of failing in a year is 0.1; failures are independent of those of other years. What is the expected value of the warranty?

2. Define the pmf and give the values of μ and σ^2 when the moment-generating function (mgf) of X is defined by

   a. $M(t) = \left(\frac{0.6e^t}{1-0.4e^t}\right)^2, \quad t < -\ln(0.4).$

3. Let X equal the number of people selected at random that you must ask in order to find someone with the same birthday as yours. Assuming each day of the year is equally likely (and ignoring February 29).

a. What probability distribution does X have? Namely, what is the pmf of X?

b. Give the mean and variance of X.

c. Find $P(X > 400)$ and $P(X < 300)$.

4. Let X equal the number of flips of a fair coin that are required to observe the same face on consecutive flips. Find the pmf of X.

5. Suppose that a basketball player can make a free throw 60% of the time. Let X equal the minimum number of free throws that this player must attempt to make a total of 10 shots,

   a. What probability distribution does X have? Namely, what is the pmf of X?

   b. Give the mean and variance of X.

   c. Find $P(X = 16)$.

6. Let X have a Poisson distribution with a variance of 3. Find $P(X = 2)$.

7. Flaws in a certain type of drapery material appear on the average of one in 150 square feet. If we assume the Poisson distribution, find the probability of at most one flaw in 225 square feet.

8. Suppose that the probability of suffering a side effect from a certain flu vaccine is 0.005. Also suppose 1000 persons are inoculated.

   a. Find the exact probability that at most 1 person suffers using a binomial distribution.

   b. Find approximately the probability that at most 1 person suffers using a Poisson distribution.

9. A hospital obtains 40% of its flu vaccine from Company A, 50% from Company B, and 10% from Company C. From past experience it is known that 3% of the vials from A are ineffective, 2% from B are ineffective, and 5% from C are ineffective. The hospital test 5 vials from each shipment. If at least one of the five is ineffective, find the conditional probability of that shipment coming from C.

10. If X has a Poisson distribution so that $3P(X = 1) = P(X = 2)$, find $P(X = 4)$.

11. One of four different prizes was randomly put into each box of a cereal. If a family decided to buy this cereal until it obtained at least one of each of the four different prizes, what is the expected number of boxes of cereal that must be purchased?