Lab Project 5: The Normal approximation to Binomial distribution

Course : Introduction to Probability and Statistics, Math 113 Section 3234 Instructor: Abhijit Champanerkar Date: Nov 21st 2012



In this project we will compare the binomial distribution, its approximation using the normal distribution and the approximation using the continuity correction. We will also learn commands to find probability for the binomial distribution.

Binomial Distributions using R

For the Binomial Distribution let p be the probability of a success and q = 1 - p be the probability of failure. The probability of exactly k successes in n trials is given by $P(k) = p^k q^{n-k} {}_n C_k$. The mean $\mu = n p$ and the standard deviation $\sigma = n p (1-p)$.

The **R** command |dbinom(k,size=n,prob=p)| gives the probability P(k). For example probability of getting 4 heads when 7 coins are tossed is:

> dbinom(4,size=7,prob=0.5)
[1] 0.2734375
> 0.5^4*0.5^3*choose(7,4) #check answer using formula
[1] 0.2734375

The **R** command [pbinom(k,size=n,prob=p)] gives the probability for the binomial distribution for **at most** k successes. This can also be done by summing P(k) for k from 0 to n. For example the probability of getting at most 4 heads when 7 coins are tossed is:

```
> pbinom(4,size=7,prob=0.5)
[1] 0.7734375
> sum(dbinom(0:4,size=7,prob=0.5)) #check answer by adding
[1] 0.7734375
```

Example 1 The probability of getting between 3 and 6 heads when 7 coins are tossed is given by:

```
> pbinom(6,size=7,prob=0.5)-pbinom(2,size=7,prob=0.5) #Note the 2 instead of 3
[1] 0.765625
> sum(dbinom(3:6,size=7,prob=0.5)) #check answer by adding
[1] 0.765625
```

Example 2 Use the pbinom command to find the probability of getting between 5 and 15 heads when 25 coins are tossed. (answer = 0.8847833)

Normal Distributions using R

The command pnorm(x,mean=0,sd=1) gives the probability for that the z-value is less than x i.e. the cumulative area on the left of a x for a standard normal distribution. The area which pnorm computes is shown here. For example, probability of getting a number less than 1 in the standard normal distribution is:



> pnorm(1,mean=0,sd=1)
[1] 0.8413447

The command pnorm(x,mean=m,sd=s) gives the probability for selecting a number less than x from a normal distribution with mean m and standard deviation s.

Example 3 The probability of getting a number between 1 and 4 in the a normal distribution with mean 2 and standard distribution 0.7 is given by:

```
> pnorm(4,mean=2,sd=.7)-pnorm(1,mean=2,sd=0.7)
[1] 0.9978535
```

Example 4 Use the pnorm command to find the probability of getting a number between 5 and 15 heads for a normal distribution with mean 8 and standard deviation 4. (answer = 0.7333135).

Approximating the Binomial distribution

Now we are ready to approximate the binomial distribution using the normal curve and using the continuity correction.

Example 5 Suppose 35% of all households in Carville have three cars, what is the probability that a random sample of 80 households in Carville will contain **at least 30 households** that have three cars.

Solution : For this problem n = 80 and p = 35% = 0.35, q = 0.65. The mean $\mu = n \times p = 80 \times 0.35 = 28$ and the standard deviation $\sigma = \sqrt{n \times p \times q} = \sqrt{80 \times 0.35 \times 0.65} = 4.26$.

Using binomial distribution:

> pbinom(80,size=80,prob=0.35)- pbinom(29,size=80,prob=0.35)
[1] 0.3588295

Using the normal distribution:

> 1-pnorm(30,mean=28,sd=4.26)
[1] 0.319362

Since we need to find probability that at least 30 households that have three cars, we need to use 30 - 0.5 in the continuity correction.

Using continuity correction:

> 1-pnorm(29.5,mean=28,sd=4.26)
[1] 0.3623769

You can see that the answer using continuity correction is much closer to the actual value !

Questions

About two out of every three gas purchases at Cheap Gas station are paid for by credit cards. 480 customers buying gas at this station are randomly selected. Find the following probabilities using the binomial distribution, normal approximation and using the continuity correction.

- 1. Find n, p, q, the mean and the standard deviation.
- 2. Find the probability that **greater than 300** will pay for their purchases using credit card.
- 3. Find the probability that **between 220 to 320** will pay for their purchases using credit card.
- 4. Generate a random number using the command floor(rnorm(1, mean=200, sd=50)). Write this number down. Lets call it N. (This number will be different for each student.)
- 5. Find the probability that at most N customers will pay for their purchases using credit card.

Write the answers you found for the above questions in the tear off sheet on the next page and hand in this sheet.

Lab Project 5

Please write your name, fill in the values, tear off and hand to instructor.

Name: _____

Write the answers you found for the above questions in the given tables and hand in this sheet.

Question 1

n	p	q	Mean μ	SD σ

Questions 2,3,4,5

Probability	Binomial	Normal	Correction
p(x > 300)			
$p(220 \le x \le 320)$			
Random Number N		Do not write	Do not write
$p(x \le N)$			