R-assignment 4

Math 361

(Read page 41-45 in "Using R for Introductory Statistics" by J. Verzani and look at the lecture notes).

Turn in the R-code together with the results.

1. Flip a fair coin 5 times. Let the random variable X be the number of heads observed.

(A) Find the probability distribution of X by filling out the following table:

x	0	1	2	3	4	5
$P(X = x)$						

(B) Find the mean and standard deviation of X by using R

(C) Now use R to simulate flipping a fair coin 5 times and count the number of heads, X. Perform this experiment $n = 100$ times. Use the **sample** command to do this.

Fill out the following table:

x	0	1	2	3	4	5
$n(x)$: frequency						
$h(x)$: relative frequency						

(D) Compute the sample mean and sample variance in part (C). What do you observe?

Problem 2 on next page.

Written by
Dr. John Palmer

There is doubtless some truth to the old saw "a camel is a horse designed by a committee". However, in the first part of this lab we will explore the possibility that for certain decisions a committee may demonstrate better judgement than the individuals who make up the committee. Suppose that you are writing a new constitution and must make a decision about the size of a jury (N) and the rule by which it comes to a verdict (conviction takes place when k or more jurors vote to convict). In your country a typical citizen will vote to convict an innocent defendent 25% of the time. The same typical citizen will vote to convict a guilty defendent 80% of the time. Let $P(C|I)$ denote the probability that a *jury* convicts an innocent defendent. Let $P(C|G)$ be the probability that a jury votes to convict a guilty defendent.

Problem 1. What is the smallest jury size, N, so that there is a k for which,

$$P(C|I) \leq .001$$

and

$$P(C|G) \geq .99.$$

You want to make the mistake of convicting an innocent at most 1 time in 1000 and convict the guilty at a rate of 99%. Notice that making k larger makes the probability of conviction in either case smaller; this will help the first inequality to be true and make it harder for the second inequality to be true.

Now suppose the number of jurors is N and it takes k or more votes to convict. Let X_I be the number of jurors that vote to convict in the trial of an innocent. Then it is natural to model X_I as a binomial random variable of size N and probability $p = .25$. We have,

$$P(C|I) = \sum_{j=k}^{N} P(X_I = j).$$

The program 'R' makes it easy to calculate this binomial tail. It you enter,

$$pbinom(k, N, .25)$$

for some choice of $k \leq N$, R computes,

$$\sum_{j=0}^{k} \binom{N}{j} (.25)^j (.75)^{N-j},$$

which is $P(X_I \leq k)$. The probability we are interested in is a complementary probability,

$$P(X_I \geq k) = 1 - pbinom(k-1, N, .25).$$

Here is one way to start trying to answer the first question. Choose $N = 12$. Enter,

$$i = 1:12$$

$$x = 1:12$$

Then,

$$for \text{ (n in i) } x[n] = 1 - pbinom(n-1, 12, .25).$$

This produces a vector whose 7^{th} entry (for example) is the probability of a jury of 12 convicting an innocent provided 7 or more jurors must vote to convict. You can test if any entry is less than .001 by typing,

$$x <= .001$$

If you do this you will see a string of 12 answers TRUE or FALSE. The first TRUE answer occurs in the 9^{th} slot. Is $k = 9$ good enough to satisfy the other condition $P(C|G) \geq .99$? If X_G is the number of votes to convict a guilty defendent then X_G is naturally modeled as a binomial random variable with parameters N and $p = .8$. The probability that a jury with 12 jurors that requires 9 or more to convict will vote to convict a guilty defendent is,

$$1 - pbinom(8, 12, .8)$$

This is about .79 and this shows that 12 jurors will not do the trick. Hint: you do not need to consider $N > 25$.