STAT 325 Introduction to Probability Models Spring 2010

HW3 due Tues Apr 6

Topics: Computer simulations

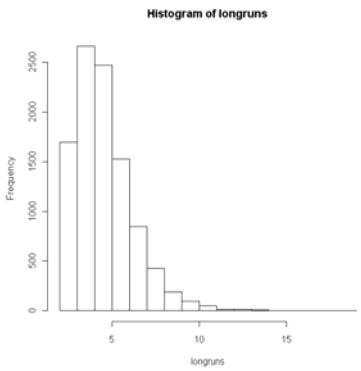
1. Reconsider the "longest run of heads" example. We will continue to approximate the probability that a sequence of n tosses of a fair coin will produce a run of more than 5 heads in a row.

a) What do you expect to happen to this probability as the number of tosses (n) increases? Explain.

The probability should increase, because with a larger number of tosses, there are more opportunities for a long run of heads.

b) Modify the R code that we produced in class (available under the "R code" link on our course webpage) to simulate $N = 10{,}000$ repetitions of $n = 50$ tosses each. Submit a histogram of the 10,000 longest runs of heads, and report the approximate probability that you obtain.

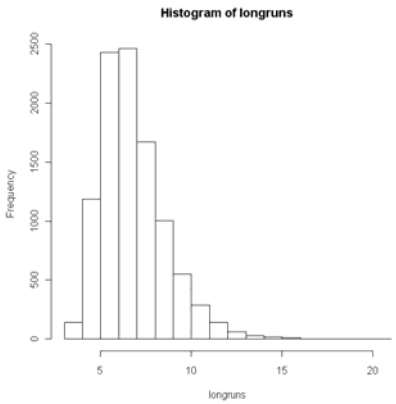
The histogram from my simulation is shown here (your results should be similar but not identical):



My simulation produced an approximate probability of 3170/10000 = .3170.

c) Repeat b) for a sequence of $n = 250$ tosses.

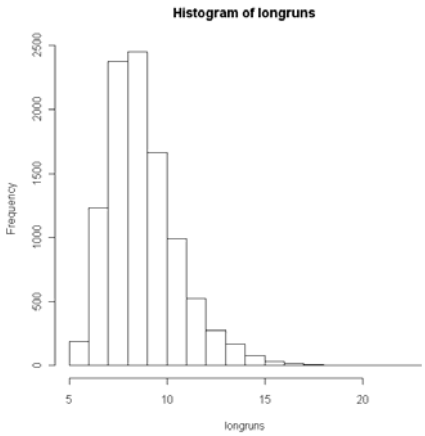
The histogram from my simulation is shown here:



My simulation produced an approximate probability of 8670/10000 = .8670.

d) Repeat b) for a sequence of n = 1000 tosses.

The histogram from my simulation is shown here:



My simulation produced an approximate probability of 9995/10000 = .9995.

e) Produce and interpret a 95% confidence interval for the actual probability in part d).

We can be 95% confident that the actual probability (of obtaining a run of more than 5 heads in a sequence of 1000 tosses) is within $.9995 \pm 1.96\sqrt{\frac{.9995 \times .0005}{10{,}000}}$, which is .9995 ± .0004, which is the interval (.9991, .9999).

f) Does your analysis support your prediction in a) about the effect of sample size on the probability of seeing a run of more than 5 consecutive heads? Explain.

Yes. The larger the number of tosses, the larger the probability of obtaining a run of more than 5 heads.

2. Consider again the example about determining the winner of a sports series. Continue to assume that team A has probability .6 of beating team B in any one game.

a) Modify the R program that we developed in class (available under the "R code" link on our course webpage) so a team needs to win k (a positive integer) more games than the other team in order to win the series. Describe the modification that you make.

Two modifications are needed. One if for the while command to check whether the difference in number of games won is less than k, rather than checking whether the difference is less than 2. The other modification is that in declaring the winner, we must check whether the number of A wins is equal to the number of B wins plus k, rather than plus 2, and the same for checking if B is the winner.

b) If you are a supporter of team A, would you prefer to have a large or small value of k, or do you not have a preference? In other words, do you expect the probability that team A wins the series to increase, decrease, or stay the same as k increases? Explain.

Team A's probability of winning the series should increase as k increases. Because A is the better team (with probability .6 of winning any one game), a longer series makes it more likely that A's superiority will come through.

c) Run the program with $N = 10{,}000$ repetitions using $k = 3$. Report the approximate probability that team A wins the series.

Answers will vary slightly, of course, but I obtained an approximate probability of .7771.

d) Produce a 95% confidence interval for this probability in part c).

The margin-of-error for my approximate probability is $1.96\sqrt{\frac{.7771 \times .2229}{10{,}000}} \approx .0082$, so a 95% confidence interval for the exact probability is .7771 ± .0082, which is (.7689, .7853).

e) Repeat c) using $k = 5$.

I obtained an approximate probability of .8845.

f) Repeat c) using $k = 7$.

I obtained an approximate probability of .9482.

g) Does your analysis support your prediction in b) about the effect of the value of k on the probability that team A wins the series? Explain.

Yes. The larger the value of k (i.e., the larger margin of victory needed for the series to end), the greater the probability that the better team (A) wins the series.