STAT 550, Fall 2009
Homework
due Friday October 30

Show all work.

The numbers refers to *Higgins and Keller-McNulty*, if not specified otherwise. The 2 problems use R.

1. p. 89, 2.7-10 REVISED
Let the random variable X denote the sum of the face values on a pair of dice. In R, simulate a random sample of this random variable for samples of size
$n = 10, 20, 30, 50, 75, 100, 200, 300, 400, 800$, and 1000. In generating the samples, use a new set of random numbers for each sample size by using a different `seed` in the `set.seed` function. Obtain the sample means $\overline{X}_n$ for each value of n. Construct the intervals $\mu \pm 2\sigma/\sqrt{n}$ where μ is $E(X)$ and σ is $STD(X)$. You do not have to plot these in R, but you can observe how $\overline{X}$ gets close to the mean. You should have 11 sample means and 11 intervals.

To simulate the roll of a pair of dice, you can use the R `sample` function,

```
> sample(c(1,2,3,4,5,6), size=2, replace=T)
```

You can then sum these two outcomes using,

```
> sum(sample(c(1,2,3,4,5,6), size=2, replace=T))
```

For each n, you can use a loop to generate n sums of the face values of a pair of dice and then use the R `mean` function to obtain the sample mean $\overline{X}_n$. You will have to use a loop for each n with a different `seed`. This means you will have 11 loops. Or, you might be able to figure out how to do this with more than one loop?

Be sure to use the R function `set.seed` so that you can reproduce your results. You should include your R code (call to functions) and output.

2. p. 110: 3.1-12. REVISED
An experiment consists of tossing a die 200 times. Let the value of 3 be a success. Let the random variable Y be the number of successes rolled in 200 die tosses. The proportion of successes is $\hat{P} = \frac{Y}{n}$ where $n = 200$. To compute the empirical probability of a 3, we could use the sample function,

```
> sum(sample(c(1,0), size=200, replace=T, prob=c(1/6,5/6)))/200
```

or the `rbinom` function as in Lab2.

(a) Compute the two-standard deviation interval for $\hat{P}$. Does the empirical probability fall in the two-standard deviation interval?

(b) Repeat the 200 die tosses 20 times (using a different `seed`). How many of the 20 empirical probabilities fall in the two-standard deviation interval?

Be sure to use the R function `set.seed` so that you can reproduce your results. You should include your R code (call to functions) and output.