Assignment #7 contains problems about sampling distributions. Here are the assigned problems.

From Chapter 5, problems 5.4, 5.20, 5.22, and 5.34.

**Problem 5:** Load in the `prob.R` source file as on a previous assignment.

Use R to plot the sampling distribution of the sample mean for a bimodal distribution with mean $\mu = 200$, standard deviation $\sigma = 72$, and additional parameter $d = 65$, for sample sizes $n = 1, 2, 4, 16, 25, 49,$ and $100$. For each sample size, draw a vertical line that is two standard errors to the right of the mean and compare the area under the true sampling distribution with that under the normal approximation.

The R code for this for $n = 1$ is as follows.

```r
n = 1
gbimod(n,200,72,65)
b = 200 + 1.96*72/sqrt(n)
abline(v=b)
```

Answer the following questions based on examining the graphs for each $n$ visually.

(a) For which values of $n$ would you say that the area to the right of $b$ under the actual sampling distribution is substantially different from the area under the normal curve?

(b) When the area is substantially different, is the area under the normal curve an overestimate or an underestimate of the true area?

**Problem 6:** Repeat the previous problem but use the skewed distribution `gskew` instead of the bimodal distribution `gbimod` with the same mean $\mu = 200$ and standard deviation $\sigma = 72$, but using skewness coefficient 2.99.

The R code for this for $n = 1$ is as follows.

```r
n = 1
gskew(n,200,72,2.99)
b = 200 + 1.96*72/sqrt(n)
abline(v=b)
```

**Problem 7:** The central limit theorem says that the distribution of the sampling distribution of the sample mean is approximately normal when $n$ is large enough. Which of the following two distributions would require a larger $n$ for the approximation to be accurate: (1) a strongly skewed distribution, or (2) a symmetric distribution with a non-normal shape. Briefly explain using the previous two problems for guidance.

There are 7 problems, so this assignment is worth 35 HW points. Please do additional problems on your own for extra practice if you feel you need to. The homework is due either in lecture or to your TA by 5pm on Thursday. You do not need to include R plots with your assignment.