Homework 5

1. Consider a set of N people, each of whom has an opinion about whether Brian should be the chair of a new biostatistics department committee. The opinion of individual i is represented by an indicator: $b_i = 1$ if individual i is for Brian and $b_i = 0$ if individual i is against Brian. $B = \sum_{i=1}^{N} b_i/N$ is the proportion of the population in favor of Brian for chair. To estimate the population proportion, a random sample of n is chosen and polled. The sample is random in the sense that all $\begin{pmatrix} N \\ n \end{pmatrix}$ samples are equally likely. The proportion of those polled who support Brian, $\bar{b} = \sum_{i=1}^{n} b_i/n$, is used to estimate the true population proportion. What is the mean and variance of $\bar{b}$? (Note: this is sampling without replacement from a finite population, different from the usual sampling scheme where we select n observations from an infinite pool. Hint: Define a random variable $I_i = 1$ if individual i is selected into the sample, $I_i = 0$ otherwise.)

2. Suppose $X_1, \cdots, X_n$ are i.i.d. with mean ξ, and suppose that $E|X_i|^k < \infty$, so that the k^{th} central moment $\mu_k = E(X_i - \xi)^k$ exists. Show that the k^{th} sample moment $M_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k$ converges in probability to μ_k.

3. Let $X_1, ..., X_n$ be a random sample from a population with pdf $f_X(x) = \frac{1}{\theta} I\{0 < x < \theta\}$. Show that $\frac{X_{(1)}}{X_{(n)}} \perp X_{(n)}$.

4. A researcher measures the pulse of n subjects while the subjects are resting and again while the subjects are exercising. For subject i, R_i is the resting pulse and E_i is the pulse measured while exercising. Assume that the pairs (R_i, E_i), $i = 1, ..., n$ are iid draws from the same underlying joint distribution. [Assume that the CLT holds for $\left(\bar{R}, \bar{E}\right)$.]

   (a) The researcher speculates that E is a location-scale transformation of R. If this is true, what is the limiting distribution of the ratio $\bar{R}/\bar{E}$?

   (b) The researcher collects data on the pulse while sleeping, S, for the same n patients. This time the researcher thinks that, for $Y_1, ..., Y_n$ iid $N(0, 1)$, $E_i = R_i + Y_i$ and $S_i = R_i - Y_i$. What is the limiting distribution of $\bar{S}/\bar{E}$?

5. Assuming $\bar{X}$ has a limiting normal distribution, what is the limiting distribution of $\bar{X}^3 - \bar{X}^2$?