1. Moment generating functions:

   (a) Obtain the MGFs for the χ^2, exponential, and Gamma distributions.

   (b) Find the distributions of $\sum_{i=1}^n Y_i$, where $Y_1, \ldots, Y_n$ are i.i.d. χ^2 and i.i.d. exponential.

   (c) Based on the MGFs, how are the χ^2 and Gamma distributions related?

2. Let $C_k(Y_1, \ldots, Y_n) = (\bar{Y} - a_k\sigma/\sqrt{n}, \bar{Y} + a_k\sigma/\sqrt{n})$ for $k \in \{1, 2\}$, where $a_1 = z_{1-\alpha/2}$ (the approximate normal interval) and $a_2 = 1/\sqrt{\alpha}$ (the Chebyshev interval). Make a plot of the relative interval width of C_1 to C_2 as a function of α. Then, via simulation, find the coverage rates of C_1 and C_2 for $\alpha = 0.80$ and $n \in \{1, 10\}$ when

   (a) $Y_1, \ldots, Y_n \sim$ i.i.d. $N(\mu, \sigma^2)$;

   (b) $Y_1, \ldots, Y_n \sim$ i.i.d. double exponential with variance 2.

   (c) $Y_1, \ldots, Y_n \sim$ i.i.d. beta(.1, .5).

   Include your code as an appendix to your homework. Discuss your results, and your thoughts on the robustness of the z-interval when the data are not normal (importantly, in this exercise, we are using the true variance of the population instead of an estimate).

3. CLT:

   (a) Let $Y_1, \ldots, Y_n \sim$ i.i.d. $N(0, \sigma^2)$. Compute the mean and variance of $\overline{Y^2} = \sum Y_i^2/n$.

   (b) Prove that $\overline{Y^2}$ converges in probability to something (and say what that something is).

   (c) If you were to simulate many datasets of size n, compute $\overline{Y^2}$ for each of them, and then make a histogram of the $\overline{Y^2}$ values, what distribution would the histogram resemble and why? Be specific, and identify the mean and variance of this distribution.

4. Weighted estimates: Sometimes our measurements of a quantity of interest have differing levels of precision. Let $\{Y_i : i \in \mathbb{N}\}$ be a vector of independent real-valued random variables with $\mathrm{E}[Y_i] = \mu$ and $\mathrm{Var}[Y_i] = a_i\sigma^2$ where the a_i's are known values.

   (a) Find the mean and variance of $\bar{Y}_w = \sum_{i=1}^n w_i Y_i$, where the w_i's are constants that sum to one.

   (b) Find the values of the w_i's that minimize the variance of $\bar{Y}_w$.

   (c) Obtain a WLLN for $\bar{Y}_w$, where the w_i's are the optimal values.

5. CDF: Let $Y_1, \ldots, Y_n$ be an i.i.d. sample from a distribution P with CDF F.

   (a) For a point $y \in \mathbb{R}$, find an unbiased and consistent estimate $\hat{F}(y)$ of $F(y)$. Show that it is unbiased, explain why it is consistent, and calculate its variance.

   (b) Find an approximation to the large-sample distribution of $\hat{F}(y)$, and use this to obtain an approximate 95% CI for $F(y)$.