Homework 5

Due Thursday, October 9

1. Let $Y_1, \dots, Y_n$ be iid Poisson(λ). Find $E(\overline{Y})$ and se($\overline{Y}$), and use these information to create a Z-score based on $\overline{Y}$ (see Lecture 8, page 11).

2. Let $Y_1, \dots, Y_n$ be iid Gamma($\alpha, 1$). Find $E(\overline{Y})$ and se($\overline{Y}$), and use these information to create a Z-score based on $\overline{Y}$.

3. Let $Y \sim$ Binomial(n, p), and consider a statistic Y/n. Find $E(Y/n)$ and se(Y/n), and use these information to create a Z-score based on Y/n.

4. Let Y_{1i} with $i = 1, \dots, n_1$ be iid $N(\mu_1, \sigma_1^2)$, and let Y_{2j} with $j = 1, \dots, n_j$ be iid $N(\mu_2, \sigma_2^2)$, with independence between Y_{1i}'s and Y_{2j}'s. Let $\overline{Y}_1 = \sum_{i=1}^{n_1} Y_{1i}/n_1$ and $\overline{Y}_2 = \sum_{j=1}^{n_2} Y_{2j}/n_2$. Find (a) se($\overline{Y}_1 + \overline{Y}_2$) and (b) se($\overline{Y}_1 - \overline{Y}_2$).

5. Let $Y_1, \dots, Y_n$ be iid, where each Y has a pdf $f_Y(y) = 3y^2$, $0 \le y \le 1$. Find $E(\overline{Y})$ and se($\overline{Y}$).

6. Let $Y_1, \dots, Y_n$ be iid, where each Y has a pdf $f_Y(y) = 4y^2 e^{-2y}$, $y > 0$. Find $E(\overline{Y})$ and se($\overline{Y}$).

7. A (sample) coefficient of variation (CV) is a statistic defined as $CV = S/\overline{Y}$. Although it is extremely difficult to derive a sampling distribution directly, we can find a related distribution under special circumstances:
   Let $Y_1, \dots, Y_n$ be iid $N(0, \sigma^2)$. Show that $n(\overline{Y})^2/S^2$ follows an F-distribution and determine its d.f.s.
   (Consequently, $S^2/[n(\overline{Y})^2]$ also follows an F-distribution and is directly related to CV).

8. Let $Y_1, \dots, Y_n$ be iid $N(0, \sigma^2)$. If $n = 4$, find $P(2\overline{Y} > S)$ (Hint: Recall the composition of T - Lecture 8, page 17 - and rearrange the terms inside the probability. Then you may use R to compute the probability).

9. Suppose that we have the same set up as in Problem 4. In addition, let

$$S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (Y_{1i} - \overline{Y}_1)^2 \quad \text{and} \quad S_2^2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (Y_{2j} - \overline{Y}_2)^2$$

   If $\sigma_1^2 = \sigma_2^2$, and if $n_1 = 10$ and $n_2 = 9$, find $P(S_1 > S_2)$ (Hint: Recall the composition of F - Lecture 8, pages 19-20 - and rearrange the terms inside the probability. Then you may use R to compute the probability).

10. Please find the dataset `hw5.10.txt`. Using R, find (a) $\overline{Y}$, (b) S, (c) the sample median M (please verify with the formula for M), (d) IQR and the sample range R, (e) the 6-th order statistic, (f) the 84-th sample quantile.

11. Repeat Problem 10 with the dataset `hw5.11.txt`.

12. Repeat Problem 10 with the dataset `hw5.12.csv`.