Solutions to Assignment #6

1. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a normal($\mu, \sigma^2$) distribution, where $\mu$ and $\sigma^2$ are unknown. Suppose that $n = 17$ and that the values of $X_1, X_2, \ldots, X_n$ add up to 79.90 and that the sum of the square difference from the values to the sample mean is 92.16. Give a 90% confidence interval for $\mu$.

**Solution:** The sample mean for the data is $\bar{X}_n = 4.7$ and the sample variance is $S_n^2 = 5.76$, so that $S_n = 2.4$. Thus, a 90% confidence interval for $\mu$ is given by

$$
\left(4.7 - t_{\alpha/2} \frac{2.4}{\sqrt{17}}, 4.7 + t_{\alpha/2} \frac{2.4}{\sqrt{17}}\right),
$$

where $\alpha = 0.1$ in this case. Thus, using MS Excel, R or looking up values in a table we obtain that $t_{\alpha/2} \approx 1.75$, where we have used the fact that the number of degrees of freedom is 16 in this case. Thus, a 90% confidence interval for the mean is $(4.7 - 1.02, 4.7 + 1.02)$, or $(3.68, 5.72)$. \(\square\)

2. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a normal($\mu, \sigma^2$) distribution, and let $S_n^2$ denote the sample variance. Since $S_n^2$ is an unbiased estimator for $\sigma^2$, $E(S_n^2) = \sigma^2$. Compute $\text{var}(S_n^2)$; that is, compute the variance of the sampling distribution of $S_n^2$.

**Suggestion:** Use the knowledge that you have about the distribution of $\frac{(n-1)S_n^2}{\sigma^2}$.

**Solution:** Using the fact that the variance of a $\chi^2(n-1)$ distribution is $2(n-1)$ we obtain that

$$\text{var} \left( \frac{(n-1)S_n^2}{\sigma^2} \right) = 2(n-1),$$

since

$$\frac{(n-1)}{\sigma^2} S_n^2 \sim \chi^2(n-1).$$

It then follows that

$$\frac{(n-1)^2}{\sigma^4} \text{var}(S_n^2) = 2(n-1),$$
from which we get that
\[ \text{var}(S_n^2) = \frac{2\sigma^4}{n-1}. \]

\[ \square \]

3. Let \( X_1, X_2, \ldots, X_n \) be a random sample from a distribution with finite variance, \( \sigma^2 \). Show that
\[ E(S_n) \leq \sigma, \]
where \( S_n \) denotes the positive square root of the sample variance. Furthermore, prove that \( E(S_n) < \sigma \) if \( \text{var}(S_n) \neq 0 \).

**Solution:** Observe that
\[
0 \leq \text{var}(S_n) = E(S_n^2) - [E(S_n)]^2 = \sigma^2 - [E(S_n)]^2.
\]
It then follows that
\[
[E(S_n)]^2 \leq \sigma^2,
\]
from which the result follows. Note that if \( \text{var}(S_n) \neq 0 \), we get strict inequality.

\[ \square \]

4. Suppose we are sampling from a Bernoulli(\( p \)) distribution. Approximately, what should the sample size, \( n \), be so that a 90% confidence interval for the parameter \( p \) has length at most 0.02.

**Solution:** An approximate 90% confidence interval for \( p \), based on the Central Limit Theorem, is based on the approximation
\[
P\left( |\hat{p}_n - p| < z_{\alpha/2} \frac{\sqrt{p(1-p)}}{\sqrt{n}} \right) \approx P(|Z| < z_{\alpha/2}) = 0.90
\]
for large values of \( n \). We then have that \( z_{\alpha/2} \approx 1.65 \). Observe that \( p(1-p) \) is at most 1/4. Consequently, the length of the interval is at most
\[
2 \frac{z_{\alpha/2}}{2\sqrt{n}} = z_{\alpha/2} \frac{1}{\sqrt{n}}.
\]
We want this length to be at most 0.02. Therefore

\[ z_{\alpha/2} \frac{1}{\sqrt{n}} < 0.02, \]

from which we get that

\[ \frac{\sqrt{n}}{z_{\alpha/2}} > 50, \]

or

\[ n > (50z_{\alpha/2})^2. \]

Hence, \( n \) should be at least 6,807.

5. Suppose we are sampling from a Poisson(\( \lambda \)) distribution. A sample of 200 observations from this distribution has mean equal to 3.4. Construct and approximate 90% confidence interval for \( \lambda \).

**Solution:** We may use the result obtained in Problem 5 of Assignment #5:

\[
\left( \left( \sqrt{\frac{Y}{n}} - \frac{z_{\alpha/2}}{2\sqrt{n}} \right)^2, \left( \sqrt{\frac{Y}{n}} + \frac{z_{\alpha/2}}{2\sqrt{n}} \right)^2 \right)
\]

where \( Y/n \) is the sample mean, which in this case is 3.4 and \( z_{\alpha/2} \) is 1.65. This yields an approximate 90% confidence interval for \( \lambda \) to be (3.2, 3.6).