Problem Set #6 Answers

Nobody caught the error in the problem. As it is written now, I ask you to describe the distribution of $\sigma_{\overline{X}}$. Finding the expected value and variance of this statistic is trivial. In class we found that $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$. Neither σ or n is a random variable. Thus, $E\left(\frac{\sigma}{\sqrt{n}}\right) = \frac{\sigma}{\sqrt{n}}$ and $var\left(\frac{\sigma}{\sqrt{n}}\right) = 0$.

What I meant to ask was to find the expected value and variance of S^2, the sample variance.

Recall that the sample variance is defined as follows:

$$S^2 = \frac{1}{n}\sum_i \left(X_i - \overline{X}\right)^2$$

The sample variance S^2 is a statistic, so like the sample mean, it will have a distribution. Here, we will discuss only the first raw moment of this distribution. To find the expected value of S^2, start with the squared difference between the first observation ($i = 1$) and the sample mean.

$$\begin{aligned}
V_1 &= \left(X_1 - \overline{X}\right)^2 \\
&= \left(X_1 - \mu + \mu - \overline{X}\right)^2 \\
&= \left[(X_1 - \mu) - \left(\overline{X} - \mu\right)\right]^2 \\
&= \left[(X_1 - \mu) - \left(\frac{\sum_i X_i}{n} - \mu\right)\right]^2 \\
&= (X_1 - \mu)^2 - 2\left(X_1 - \mu\right)\left(\frac{\sum_i X_i}{n} - \mu\right) + \left(\frac{\sum_i X_i}{n} - \mu\right)^2
\end{aligned}$$

To find the expectation of V_1, recall that the expectation of a sum equals the sum of the expectations of the individual terms, so we can consider each of the three terms separately. We have seen results for first and third terms already:

$$E\left(X_1 - \mu\right)^2 = \sigma^2$$

$$E\left(\frac{\sum_i X_i}{n} - \mu\right)^2 = \frac{\sigma^2}{n}$$

Regarding the second term, because of the random sample, we know there is no tendency for all subsequent draws to be above or below the mean depending on whether the first draw is above the mean. Therefore, we need only worry about the contribution of first value ($i = 1$) for X to the sample mean.

$$\begin{aligned}
E\left[2\left(X_1 - \mu\right)\left(\frac{\sum_i X_i}{n} - \mu\right)\right] &= \frac{2}{n}E\left(X_1 - \mu\right)^2 \\
&= \frac{2\sigma^2}{n}
\end{aligned}$$

Putting all three pieces together:

$$
\begin{aligned}
E\left(V_1\right) &= E\left(X_1-\mu\right)^2 - 2E\left[\left(X_1-\mu\right)\left(\frac{\sum_i X_i}{n}-\mu\right)\right] + E\left(\frac{\sum_i X_i}{n}-\mu\right)^2 \\
&= \sigma^2 - \frac{2\sigma^2}{n} + \frac{\sigma^2}{n} \\
&= \sigma^2\left(1-\frac{1}{n}\right)
\end{aligned}
$$

Now let's return to the expectation of S^2. Note that $S^2 = \frac{1}{n}\sum_i V_i$. Because all the V_i's are drawn from the same distribution, the expectation of each equals the expectation of V_1. Thus,

$$
\begin{aligned}
E\left(S^2\right) &= \frac{1}{n}\sum_i E\left(V_1\right) \\
&= \sigma^2\left(1-\frac{1}{n}\right)
\end{aligned}
$$

We turn now to the $var(S^2)$.

$$
\begin{aligned}
Var\left(S^2\right) &= E\left(S^2\right)^2 - \left[E\left(S^2\right)\right]^2 \\
&= E\left(S^2\right)^2 - \sigma^4\left(1-\frac{1}{n}\right)^2 \\
&= E\left[EX^2-(EX)^2\right]^2 - \sigma^4\left(1-\frac{1}{n}\right)^2 \\
&= E\left[\frac{\sum_i X_i^2}{n} - \left(\frac{\sum_i X_i}{n}\right)^2\right]^2 - \sigma^4\left(1-\frac{1}{n}\right)^2 \\
&= E\left[\frac{1}{n^2}\left(\sum_i X_i^2\right)^2 - \frac{2}{n^3}\left(\sum_i X_i^2\right)\left(\sum_i X_i\right)^2 + \frac{1}{n^4}\left(\sum_i X_i\right)^4\right] - \sigma^4\left(1-\frac{1}{n}\right)^2
\end{aligned}
$$

This is a tedious math problem. Setting up the problem up to this point is sufficient. The answer is as follows:

$$
Var\left(S^2\right) = \frac{n-1}{n^3}\left[(n-1)\,3-(n-3)\right]\sigma^4
$$