Problem Set 4 Solutions

Problem 1 in Lecture 4

(a)

$$\begin{aligned}
\bar{X}_{n+1} &= \frac{1}{n+1}\sum_{i=1}^{n+1} X_i \\
&= \frac{1}{n+1}\left(n\frac{1}{n}\sum_{i=1}^{n} X_i + X_{n+1}\right) \\
&= \frac{1}{n+1}\left(n\bar{X}_n + X_{n+1}\right)
\end{aligned}$$

(b)

$$\begin{aligned}
s_{n+1}^2 &= \frac{1}{n}\sum_{i=1}^{n+1}\left(X_i - \bar{X}_{n+1}\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_{n+1}\right)^2 + \frac{1}{n}\left(X_{n+1} - \bar{X}_{n+1}\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n + \bar{X}_n - \bar{X}_{n+1}\right)^2 \\
&+ \frac{1}{n}\left(\frac{n}{n+1}X_{n+1} - \frac{1}{n+1}\sum_{i=1}^{n} X_i\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n - \frac{1}{n+1}\left(X_{n+1} - \bar{X}_n\right)\right)^2 \\
&+ \frac{1}{n}\left(\frac{n}{n+1}\left(X_{n+1} - \bar{X}_n\right)\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 - 2\frac{1}{n+1}\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)\left(X_{n+1} - \bar{X}_n\right) \\
&+ \frac{1}{n+1}\left(X_{n+1} - \bar{X}_n\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 + \frac{1}{n+1}\left(X_{n+1} - \bar{X}_n\right)^2 \\
&= \frac{1}{n}\left((n-1)s_n^2 + (n/(n+1))(X_{n+1} - \bar{X}_n)^2\right)/n
\end{aligned}$$

Problem 2 in Lecture 4

The natural estimator is $\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^n X_i^k$. This estimator is unbiased under $i.i.d$ condition since $E(\hat{\mu}_k) = E(\frac{1}{n}\sum_{i=1}^n X_i^k) = \frac{1}{n}\sum_{i=1}^n E(X_i^k) = E(X^k)$.

Problem 3 in Lecture 4

The natural estimator is $\hat{m}_k = \frac{1}{n}\sum_{i=1}^n (X_i - \hat{\mu})$. In general, it is biased. For instance, when $k = 2$, we know it is biased.
Note that there is no clear relationship between the natural estimators for centered moments and their expectation. For instance, $E\hat{m}_2 = \frac{n-1}{n}\mu_2$, $E\hat{m}_3 = \frac{(n-1)(n-2)}{n^2}\mu_3$, and $E\hat{m}_4 = \frac{(n-1)(3(2n-3)\mu_2^2+(n^2-3n+3)\mu_4)}{n^3}$.

Problem 4 in Lecture 4

Note that here, as an estimator is not being centered to any sample average, we do not have to use the usual trick(like subtracting and adding the same moment). Therefore the calculation is straightforward as following:

$$\begin{aligned}
Var(\hat{\mu}_k) &= E[\left(\frac{1}{n}\sum_{i=1}^n X_i^k - \mu_k\right)^2] \\
&= \frac{1}{n^2}E[\sum_{i=1}^n \left(X_i^k - \mu_k\right)^2] + \frac{1}{n^2}E[\sum_{i\neq j}\left(X_i^k - \mu_k\right)\left(X_j^k - \mu_k\right)] \\
&= \frac{1}{n^2}\sum_{i=1}^n E[\left(X_i^k - \mu_k\right)^2] + \frac{1}{n^2}\sum_{i\neq j} E\left(X_i^k - \mu_k\right)E\left(X_j^k - \mu_k\right) \\
&= \frac{1}{n}E[\left(X_i^k - \mu_k\right)^2] \\
&= \frac{1}{n}(\mu_{2k} - \mu_k^2) \\
&= \frac{1}{n}Var(X^k)
\end{aligned}$$

This is finite when μ_{2k} is finite.

Problem 6 in Lecture 4

Note that $\sqrt{\cdot}$ is a concave function. Thus, Jensen's inequality implies

$$E(s_n) = E(\sqrt{s_n^2}) \leq \sqrt{E(s_n^2)} = \sigma.$$

Problem 8 in Lecture 4

$$
\begin{aligned}
\hat{\sigma^2} &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}((X_i - \mu) + (\mu - \bar{X}))^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 - 2(X_i - \mu)(\bar{X} - \mu)) + (\bar{X} - \mu)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 - (\bar{X}_n - \mu)^2.
\end{aligned}
$$

Problem 9 in Lecture 4

$$
\begin{aligned}
cov(\bar{X}_n, \hat{\sigma}^2) &= E((\bar{X}_n - E(\bar{X}_n))(\hat{\sigma}^2 - E(\hat{\sigma}^2))) \\
&= E((\bar{X}_n - E(\bar{X}_n))\hat{\sigma}^2)) \\
&= E\left((\bar{X}_n - \mu)\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 - (\bar{X}_n - \mu)^2\right)\right) \\
&= E\left((\bar{X}_n - \mu)\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2\right) - E\left((\bar{X}_n - \mu)^3\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}E\left((X_i - \mu)^3\right) + \frac{1}{n^2}\sum_{i\neq j}E\left((X_i - \mu)^2\right)E\left((X_j - \mu)\right) \\
&- \frac{1}{n^2}E(X - \mu)^3 \\
&= \frac{n-1}{n^2}E\left((X_i - \mu)^3\right)
\end{aligned}
$$

The last term is zero when the third centered moment is zero.

Problem 10 in Lecture 4

(a)

$$
E(\bar{X}_n) = \frac{1}{n}\sum_{i=1}^{n}\mu_i
$$

(b)

$$
\begin{aligned}
Var(\bar{X}_n) &= E\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu_i\right)^2\right] \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{n} E\left[(X_i-\mu_i)^2\right] + \sum_{i\neq j} E\left[(X_i-\mu_i)(X_{ij}-\mu_j)\right]\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i)
\end{aligned}
$$

Problem 12 in Lecture 2

Let's express $Q = \sum_{i=1}^{r} X_i^2$ where $X_i \sim N(0,1)$. Observe that $E(X_i^2) = 1$ and $E(X_i^4) = 3$.

$$
E(Q) = E(\sum_{i=1}^{r} X_i^2) = \sum_{i=1}^{r} E(X_i^2) = r
$$

$$
Var(Q) = Var(\sum_{i=1}^{r} X_i^2) = \sum_{i=1}^{r} Var(X_i^2) = \sum_{i=1}^{r}[E(X_i^4) - E(X_i^2)^2] = \sum_{i=1}^{r}(3-1) = 2r
$$

The second equality of variance calculation holds since X_i^2's are independent.

Problem 14 in Lecture 2

(a)

$$
\begin{aligned}
E(\bar{X}_n - \bar{Y}_n) &= E(\bar{X}_n) - E(\bar{Y}_n) \\
&= E(n_1^{-1}\sum_{i=1}^{n_1} X_i) - E(n_2^{-1}\sum_{i=1}^{n_2} Y_i) \\
&= n_1^{-1}\sum_{i=1}^{n_1} E(X_i) - n_2^{-1}\sum_{i=1}^{n_2} E(Y_i) \\
&= \mu_X - \mu_Y
\end{aligned}
$$

(b)

$$
\begin{aligned}
Var(\bar{X}_n - \bar{Y}_n) &= Var(\bar{X}_n) + V(\bar{Y}_n) \\
&= Var(n_1^{-1}\sum_{i=1}^{n_1} X_i) + Var(n_2^{-1}\sum_{i=1}^{n_2} Y_i) \\
&= n_1^{-2}\sum_{i=1}^{n_1} Var(X_i) + n_2^{-2}\sum_{i=1}^{n_2} Var(Y_i) \\
&= \frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}
\end{aligned}
$$

(c) Since X_i's and Y_i's are mutually independent and follow normal distribution, $\bar{X}_n - \bar{Y}_n$, which is a linear combination of X_i's and Y_i's is also normally distributed. We have calculated the mean and variance in part (a) and (b). Hence, $\bar{X}_n - \bar{Y}_n \sim N(\mu_X - \mu_Y, \frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2})$.