Solutions to Assignment #14

1. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a Bernoulli($p$) distribution with $0 < p < 1$. We have seen that $\hat{p} = \frac{1}{n} \sum_{i=1}^{n} X_i$ is the MLE for $p$. Compute the mean squared error, $\text{MSE}(\hat{p})$, of $\hat{p}$.

**Solution:** Observe that $\hat{p}$ is an unbiased estimator for $p$. It then follows that

$$\text{MSE}(\hat{p}) = \text{var}(\hat{p}) = \frac{p(1 - p)}{n}.$$ 

□

2. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a distribution with mean $\mu$ and variance $\sigma^2$.

(a) For non–negative constants $a_1, a_2, \ldots, a_n$, define

$$W = \sum_{i=1}^{n} a_i X_i.$$ (1)

Prove that $W$ is an unbiased estimator for $\mu$ if and only if $\sum_{i=1}^{n} a_i = 1$.

**Solution:** The estimator $W = \sum_{i=1}^{n} a_i X_i$ is an unbiased estimator for $\mu$ if and only if $E(W) = \mu$, if and only if,

$$\sum_{i=1}^{n} a_i E(X_i) = \mu,$$

if and only if

$$\sum_{i=1}^{n} a_i \mu = \mu,$$

if and only if

$$\sum_{i=1}^{n} a_i = 1,$$

which was to be shown. □
(b) Out of all the unbiased estimators of $\mu$ of the form in (1), find the one which has the smallest possible variance. Calculate the variance of that estimator.

**Solution:** For any estimator, $W$, of $\mu$, which is of the form

$$W = \sum_{i=1}^{n} a_i X_i,$$

the variance is given by

$$\text{var}(W) = \sum_{i=1}^{n} a_i^2 \text{var}(X_i) = \sigma^2 \sum_{i=1}^{n} a_i^2,$$

since the $X_i$s are independent. Thus, we need to minimize the function

$$f(a_1, a_2, \ldots, a_n) = \sum_{i=1}^{n} a_i^2$$

subject to the constraint

$$g(a_1, a_2, \ldots, a_n) = \sum_{i=1}^{n} a_i = 1.$$

We therefore use the method of Lagrange multipliers; that is, we find $\lambda$ and $a_1, a_2, \ldots, a_n$ such that

$$\nabla f(a_1, a_2, \ldots, a_n) = \lambda \nabla g(a_1, a_2, \ldots, a_n),$$

which leads to the equations

$$2a_i = \lambda \quad \text{for} \quad i = 1, 2, \ldots, n,$$

or

$$a_i = \frac{\lambda}{2} \quad \text{for} \quad i = 1, 2, \ldots, n.$$

Substituting these into the constraint equation,

$$g((a_1, a_2, \ldots, a_n) = 1,$$

yields

$$n \frac{\lambda}{2} = 1.$$
from which we get that
\[ a_i = \frac{1}{n} \quad \text{for } i = 1, 2, \ldots, n. \]

Thus, the estimator
\[ W = \sum_{i=1}^{n} \frac{1}{n} X_i = \bar{X}_n \]
provides a critical point for the variance over all estimators of the form
\[ W = \sum_{i=1}^{n} a_i X_i, \]
with \( \sum_{i=1}^{n} a_i = 1. \) To see that \( \bar{X}_n \) yields the smallest variance out of all estimators in the simplex defined by \( \sum_{i=1}^{n} a_i = 1, \) we compare var(\( \bar{X}_n \)) with the variance at the corners of the simplex; namely,
\[ \text{var}(X_i) = \sigma^2, \quad \text{for } i = 1, 2, \ldots, n. \]

Comparing these with
\[ \text{var}(\bar{X}_n) = \frac{\sigma^2}{n}, \]
we see that \( \bar{X}_n \) has the smallest possible variance. \( \square \)

3. Let \( X_1, X_2, \ldots, X_n \) denote a random sample from a normal distribution with mean \( \mu \) and variance \( \sigma^2. \)

Compute the efficiency, \( \text{eff}(\hat{\sigma}^2, S_n^2) \), of \( \hat{\sigma}^2 \), the MLE for \( \sigma^2 \), relative to the sample variance, \( S_n^2 \). What do you conclude?

**Solution**: Since \( \hat{\sigma}^2 \) is not an unbiased estimator of \( \sigma^2 \), in this problem, it makes more sense to look at the ratio of the MSEs; that is, we consider the relative efficiency defined by
\[ \text{eff}(\hat{\sigma}^2, S_n^2) = \frac{\text{MSE}(\hat{\sigma}^2)}{\text{MSE}(S_n^2)}. \]
where

\[ \text{MSE}(S_n^2) = \text{var}(S_n^2) \]

since \( S_n^2 \) is an unbiased estimator of \( \sigma^2 \).

Using the fact that \( \frac{n-1}{\sigma^2} S_n^2 \) has a \( \chi^2 \) with \( n - 1 \) degrees of freedom we obtain that

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

from which we get that

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

or

\[ \frac{(n-1)^2}{\sigma^4} \text{var}(S_n^2) = 2(n-1), \]

from which we get that

\[ \text{var}(S_n^2) = \frac{2\sigma^4}{n-1}. \]

We therefore have that

\[ \text{MSE} \left( S_n^2 \right) = \frac{2\sigma^4}{n-1}. \]

Next, we compute the MSE of \( \hat{\sigma}^2 \).

Observe first that \( \hat{\sigma}^2 = \frac{n-1}{n} S_n^2 \), so that

\[ E(\hat{\sigma}^2) = \frac{n-1}{n} E(S_n^2) = \frac{n-1}{n} \sigma^2. \]

Thus \( \hat{\sigma}^2 \) is biased with

\[ \text{bias}(\hat{\sigma}^2) = E(\hat{\sigma}^2) - \sigma^2 = -\frac{\sigma^2}{n}. \]
Next, we compute the variance of $\hat{\sigma}^2$:

\[
\text{var}(\hat{\sigma}^2) = \text{var}\left(\frac{n - 1}{n} S_n^2\right) \\
= \frac{(n - 1)^2}{n^2} \text{var}(S_n^2) \\
= \frac{(n - 1)^2}{n^2} \cdot \frac{2\sigma^4}{n - 1} \\
= \frac{2(n - 1)}{n^2} \cdot \sigma^4.
\]

Thus,

\[
\text{MSE}(\hat{\sigma}^2) = \text{var}(\hat{\sigma}^2) + (\text{bias}(\hat{\sigma}^2))^2 \\
= \frac{2(n - 1)}{n^2} \sigma^4 + \frac{1}{n^2} \sigma^2 \\
= \frac{2n - 1}{n^2} \sigma^4.
\]

Thus,

\[
\text{eff}(\hat{\sigma}^2, S_n^2) = \frac{\text{MSE}(\hat{\sigma}^2)}{\text{MSE}(S_n^2)} \\
= \frac{2n - 1}{n^2} \sigma^4 \\
= \frac{2}{n - 1} \sigma^4 \\
= \frac{(2n - 1)(n - 1)}{2n^2} \\
= 1 - \frac{1}{2n} \left(3 - \frac{1}{n}\right).
\]

Hence,

\[
\frac{\text{MSE}(\hat{\sigma}^2)}{\text{MSE}(S_n^2)} < 1, \quad \text{for all } n,
\]

which shows that the MLE $\hat{\sigma}^2$ has a smaller mean square error that the unbiased estimator $S_n^2$. \qed
4. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a Poisson distribution with parameter $\lambda$.

(a) Show that the sample mean, $\bar{X}_n$, and the sample variance, $S_n^2$, are unbiased estimators of $\lambda$.

**Solution:** The mean and variance of a Poisson distribution with parameter $\lambda$ are both equal to $\lambda$. It then follows that the sample mean, $\bar{X}_n$, and the sample variance, $S_n^2$, are both unbiased estimators of $\lambda$. \hfill \square

(b) Compute the efficiency, $\text{eff}(\bar{X}_n, S_n^2)$, of $\bar{X}_n$ relative to $S_n^2$. What do you conclude?

**Solution:** We need to compute the variance of $S_n^2$. In order to do this we apply the formula

$$\text{var}(S_n^2) = \frac{1}{n} \left( \mu_4 - \frac{n-3}{n-1} \mu_2 \right), \quad (2)$$

where $\mu_2$ denotes the second central moment, or variance, of the distribution and $\mu_4$ is the fourth central moment. We therefore have from (2) that

$$\text{var}(S_n^2) = \frac{1}{n} \left( E[(X - \lambda)^4] - \frac{n-3}{n-1} \lambda^2 \right), \quad (3)$$

where $X \sim \text{Poisson}(\lambda)$. Next, compute

$$E[(X - \lambda)^4] = E[X^4 - 4\lambda X^3 + 6\lambda^2 X^2 - 4\lambda^3 X + \lambda^4]$$

$$= E(X^4) - 4\lambda E(X^3) + 6\lambda^2 E(X^2) - 4\lambda^3 E(X) + \lambda^4,$$

where we have used the linearity of the expectation operator. Thus,

$$E[(X - \lambda)^4] = E(X^4) - 4\lambda E(X^3) + 6\lambda^2 E(X^2) - 3\lambda^4. \quad (4)$$

Next, compute the moments $E(X^2)$, $E(X^3)$ and $E(X^4)$ by using the moment generating function

$$M_X(t) = e^{\lambda(e^t-1)}, \quad \text{for all } t \in \mathbb{R}.$$ 

Taking the first four derivatives we obtain

$$M'_X(t) = \lambda e^t e^{\lambda(e^t-1)},$$

$$M''_X(t) = (\lambda^2 e^{2t} + \lambda e^t) e^{\lambda(e^t-1)},$$

$$M'''_X(t) = (\lambda^3 e^{3t} + 2\lambda^2 e^{2t} + \lambda e^t) e^{\lambda(e^t-1)},$$

$$M''''_X(t) = (\lambda^4 e^{4t} + 6\lambda^3 e^{3t} + 6\lambda^2 e^{2t} + 3\lambda^2 e^t) e^{\lambda(e^t-1)}.$$
\[ M''_X(t) = (\lambda^3 e^{3t} + 3\lambda^2 e^{2t} + \lambda e^t) e^{\lambda(e^t - 1)}, \]

and
\[ M^{(4)}_X(t) = (\lambda^4 e^{4t} + 6\lambda^3 e^{3t} + 7\lambda^2 e^{2t} + \lambda e^t) e^{\lambda(e^t - 1)}. \]

It then follows that the moments of \( X \) are
\[ E(X^2) = M''_X(0) = \lambda^2 + \lambda, \]
\[ E(X^3) = M'''_X(0) = \lambda^3 + 3\lambda^2 + \lambda, \]
\[ E(X^4) = M^{(4)}_X(0) = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda. \]

Thus, using equation (4), we obtain that
\[ E[(X - \lambda)^4] = 3\lambda^2 + \lambda. \]

We therefore get from (3) that
\[ \text{var}(S^2_n) = \frac{1}{n} \left( 3\lambda^2 + \lambda - \frac{n - 3\lambda^2}{n - 1} \right) \]
\[ = \frac{1}{n} \left( \frac{2n}{n - 1} \lambda^2 + \lambda \right) \]

The variance of \( X_n \) is
\[ \text{var}(X_n) = \frac{\lambda}{n}. \]

We then have that
\[ \text{eff}(\overline{X}_n, S^2_n) = \frac{\text{var}(\overline{X}_n)}{\text{var}(S^2_n)} = \frac{n - 1}{n - 1 + 2n\lambda} < 1 \]

for all \( n \). Consequently, \( \overline{X}_n \) is a more precise estimator of \( \lambda \) than the sample variance. \( \square \)

5. Let \( X_1, X_2, \ldots, X_n \) denote a random sample from a uniform distribution over the interval \([0, \theta]\) for some parameter \( \theta > 0 \).

We saw in Problem 4 of Assignment #12 that \( W = \max\{X_1, X_2, \ldots, X_n\} \) is the MLE for \( \theta \). Determine whether \( W \) is unbiased or not.
Solution: We need to compute the expected value of $W$:

$$E_n(W) = \int_{-\infty}^{\infty} w f_w(w \mid \theta) \, dw,$$

where $f_w(w \mid \theta)$ is the pdf of $W$. To determine the pdf of $W$, we first compute the cdf

$$F_w(w \mid \theta) = P(W \leq w)$$

$$= P(\max\{X_1, X_2, \ldots, X_n\} \leq w)$$

$$= P(X_1 \leq w, X_2 \leq w, \ldots, X_n \leq w)$$

$$= P(X_1 \leq w) \cdot P(X_2 \leq w) \cdots P(X_n \leq w),$$

where we have used the independence of the $X_i$s. Consequently, since the $X_i$s are also identically distributed,

$$F_w(w \mid \theta) = [F_X(w)]^n.$$

it then follows that

$$f_w(w \mid \theta) = n[F_X(w)]^{n-1} f_X(w \mid \theta),$$

where $f_X(w \mid \theta)$ is the pdf of $X \sim \text{uniform}[0, \theta]$; namely,

$$f_x(w \mid \theta) = \begin{cases} \frac{1}{\theta} & \text{if } 0 \leq w \leq \theta; \\ 0 & \text{otherwise}, \end{cases}$$

so that

$$F_x(w \mid \theta) = \begin{cases} 0 & \text{if } w \leq 0; \\ \frac{w}{\theta} & \text{if } 0 < w \leq \theta; \\ 1 & \text{if } w > \theta. \end{cases}$$

Consequently, by equation (6),

$$f_w(w \mid \theta) = \begin{cases} \frac{n w^{n-1}}{\theta^n} & \text{if } 0 \leq w \leq \theta; \\ 0 & \text{otherwise}, \end{cases}$$
It then follows from 5) that

\[ E_\theta(W) = \int_0^\theta \frac{nw^n}{\theta^n} \, d\theta = \frac{n}{n + 1} \theta. \]

We then see that \( E_\theta(W) \neq \theta \), which shows that \( W \) is not an unbiased estimator of \( \theta \). \( \square \)