Homework 3 Solutions

PSE 6.3

$$\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k$$

You would expect for $\hat{\mu}_k$ to be biased. In fact, we have already shown this for $k = 2$. You can follow the same route as in that case as well by noticing that

$$\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n}\Big((X_i - \mu) - (\bar{X} - \mu)\Big)^k$$

As far as I can tell, this seems like quite a tedious calculation, even for small values of k such as $k = 3$, but you can notice that essentially the same issue will come up as in the case with $k = 2$. If you raise the term inside parentheses to the kth power, the first term will involve $n^{-1}\sum_{i=1}^{n}(X_i - \mu)^k$, and the expectation of this term will be equal to μ_k. However, there will be left over terms that have non-zero expectations implying that $\hat{\mu}_k$ will be biased.

PSE 6.6

Let $X = s^2$, then we have from Jensen's inequality that

$$\sqrt{\mathrm{E}[X]} \geq \mathrm{E}\left[\sqrt{X}\right]$$

which holds because square root is a concave function. Thus,

$$\sqrt{\mathrm{E}[s^2]} = \sqrt{\sigma^2} = \sigma \geq \mathrm{E}\left[\sqrt{s^2}\right] = \mathrm{E}[s]$$

which is the result that we were trying to show and uses that $\mathrm{E}[s^2] = \sigma^2$ (i.e., that s^2 is unbiased for σ^2).

PSE 6.9

a) $\hat{\theta} = \exp(\bar{X})$

b) $\hat{\theta} = \log\left(\frac{1}{n}\sum_{i=1}^{n}\exp(X_i)\right)$

c) $\hat{\theta} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}X_i^4}$

d) It's helpful to re-write $\mathrm{var}(X^2) = \mathrm{E}\left[(X^2)^2\right] - \mathrm{E}[X^2]^2$ which suggests the following estimator

$$\widehat{\mathrm{var}}(X^2) = \frac{1}{n}\sum_{i=1}^{n}X_i^4 - \left(\frac{1}{n}\sum_{i=1}^{n}X_i^2\right)^2$$

PSE 6.12

a)

$$\mathrm{sd}(\bar{X}) = \sqrt{\mathrm{var}(\bar{X})} = \frac{\sigma}{\sqrt{n}}$$

b) We want $\mathrm{sd}(\bar{X}) < \tau$. In order for this to condition to be met, we need that

$$n > \frac{\sigma^2}{\tau^2}$$

which holds just from plugging in the expression from part (a) for $\mathrm{sd}(\bar{X})$ and re-arranging (and squaring). This implies that we need n to be larger if σ^2 is higher or if τ is smaller (i.e., a tighter bound).

PSE 6.15

a)

$$\begin{aligned} \mathrm{E}\left[\bar{X} - \bar{Y}\right] &= \mathrm{E}\left[\bar{X}\right] - \mathrm{E}\left[\bar{Y}\right] \\ &= \mathrm{E}[X] - \mathrm{E}[Y] \\ &= \mu_X - \mu_Y \end{aligned}$$

b)

$$\begin{aligned} \mathrm{var}(\bar{X} - \bar{Y}) &= \mathrm{var}(\bar{X}) + \mathrm{var}(\bar{Y}) - 2 \underbrace{\mathrm{cov}(\bar{X}, \bar{Y})}_{=0 \text{ by independence}} \\ &= \frac{\mathrm{var}(X)}{n_1} + \frac{\mathrm{var}(Y)}{n_2} \\ &= \frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2} \end{aligned}$$

c) Because X_i and Y_i are all independent and follow a normal distribution (hence $n_1^{-1}X_i$ and $n_2^{-1}Y_i$ are also all independent and normally distributed), this means that $\bar{X} - \bar{Y}$ is equal to the sum of independent and normally distributed random variables. Since this follows a normal distribution, all that remains is calculate its mean and variance, which is what we did in parts (a) and (b); thus,

$$\bar{X} - \bar{Y} \sim N\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}\right)$$

d) We can estimate σ_X^2 and σ_Y^2 by

$$\hat{\sigma}_X^2 = \frac{1}{n_1}\sum_{i=1}^{n_1}(X_i - \bar{X})^2 \quad \text{and} \quad \hat{\sigma}_Y^2 = \frac{1}{n_2}\sum_{i=1}^{n_2}(Y_i - \bar{Y})^2$$

and, therefore, this suggests (based on the answer to part (b)) estimating $\mathrm{var}(\bar{X} - \bar{Y})$ by

$$\widehat{\mathrm{var}}(\bar{X} - \bar{Y}) = \frac{\hat{\sigma}_X^2}{n_1} - \frac{\hat{\sigma}_Y^2}{n_2}$$

Additional Question

```r
# roll function
roll <- function() {
  sample(seq(1,6), size=1)
}

# function to generate sample
generate_sample <- function(n) {
  this_sample <- c()
  for (i in 1:n) {
    this_sample[i] <- roll()
  }
  this_sample
}

# function to run monte carlo simulations
# this a list with estimated bias,
# sampling variance, and mean squared error
mc_sim <- function(n, nsims=1000) {
  est_vec <- c()
  for (i in 1:nsims) {
    est_vec[i] <- mean(generate_sample(n))
  }

  # compute bias, sampling variance, and mse
  bias <- mean(est_vec) - 3.5
  v <- var(est_vec)
  mse <- bias^2 + v

  # return bias, sampling variance, and mse
  data.frame(bias=bias, variance=v, mse=mse)
}
```

n=2

```r
set.seed(123) # for replicability
mc_sim(2)
```

```
##      bias variance      mse
## 1 -0.0215 1.516805 1.517267
```

n=10

```r
mc_sim(10)
```

```
##     bias variance      mse
## 1 -0.025 0.313989 0.314614
```

n=50

```
mc_sim(50)
```

```
##      bias   variance        mse
## 1 0.00424 0.05472234 0.05474032
```

n=1000

```
mc_sim(1000)
```

```
##        bias    variance         mse
## 1 -0.000447 0.003127549 0.003127749
```

These results seem in line with our theory from class. The estimated bias when $n = 2$ does seem somewhat higher than for other values of n, but, if you were to increase the number of Monte Carlo simulations (say to 10,000), this should be closer to 0. The results for sampling variance and mean squared error are more clear — both are decreasing with the sample size.

Part 3

For this part, we will write a new version of the `mc_sim` function to calculate $\hat{\beta} = \log(\bar{X})$.

```
mc_sim2 <- function(n, nsims=1000) {
  est_vec <- c()
  for (i in 1:nsims) {
    # this is the only line changed here
    est_vec[i] <- log(mean(generate_sample(n)))
  }

  # compute bias, sampling variance, and mse
  bias <- mean(est_vec) - log(3.5)
  v <- var(est_vec)
  mse <- bias^2 + v

  # return bias, sampling variance, and mse
  data.frame(bias=bias, variance=v, mse=mse)
}
```

n=10

```
mc_sim2(10)
```

```
##          bias   variance        mse
## 1 -0.01373927 0.02529845 0.02548722
```

n=1000

```r
mc_sim2(1000)
```

```
##              bias     variance          mse
## 1 -5.828971e-05 0.0002370017 0.0002370051
```

We know from Jensen's inequality that $\hat{\beta}$ should be biased (because it is a concave function of $\bar{X}$). Indeed, we see that the bias is notably lower when $n = 1000$ than when $n = 10$. As in the previous part, sampling variance and mean squared error notably decrease with the larger sample size.