MATH 392 Problem Set 5

Sam D. Olson

Time Spent

$11\frac{1}{2}$ hours

Exercise from the book

8.5 #4

(Q): Suppose that $X_1, ..., X_n$ form a random sample from the normal distribution with unkown mean μ and unknown variance σ^2. How large a random sample must be taken in order that there will be a confidence interval for μ with confidence coefficient 0.95 and length less than 0.01σ?

(A):

Note: $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$ has a standard normal distribution.

It then follows:

$$Pr\left(-1.96 < \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} < 1.96\right) = 0.95$$

We isolate the μ in the above relation, giving us:

$$Pr\left(\bar{X}_n - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X}_n + 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

THus, we now have the confidence interval: $\left(\bar{X}_n - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X}_n + 1.96\frac{\sigma}{\sqrt{n}}\right)$

As the above interval included $\bar{X}_n$ on both ends of the interval (taking it as a constant for our following observation), we may then note the length of this interval is $1.96\frac{\sigma}{\sqrt{n}} \cdot 2 = 3.92\frac{\sigma}{\sqrt{n}}$

If we want a length less than 0.01σ, then we have:

$3.92\frac{\sigma}{\sqrt{n}} < 0.01\sigma \rightarrow \frac{3.92}{\sqrt{n}} < 0.01$

Thus we have, for n positive (given the context of the problem):

$3.92 < \sqrt{n} \cdot 0.01 \rightarrow 392 < \sqrt{n} \rightarrow 392^2 < n$

Taking note of the strict inequality (and the fact $392^2 = 153,664$), we need an n sample of size *at least* 153,665 to satisfy the stated condition.

Case Study: German Tank Problem

Let's pick up the example that we began in class but make the simplifying assumption that we're studying a process where our sample is drawn from the continuous distribution, $X_1, X_2, \ldots, X_n \sim \text{Unif}(0, \theta)$, but we're still interested in estimating θ. The MLE and Method of Moments estimators are the same:

$$\hat{\theta}_{MLE} = \max(X_1, X_2, \ldots, X_n) = X_{max}$$
$$\hat{\theta}_{MOM} = 2\bar{X}$$

1. Calculate the bias of each estimator. If either one is biased, propose an additional estimator that corrects that bias (in the spirit of how s^2 is the bias-corrected version of $\hat{\sigma}^2$) . What happens to the bias of these estimators as sample size grows? Plot the relationship between sample size and bias for each estimator (two lines on one plot).

(A):

For θ_{MLE} we have:

Note, the sampling distribution to θ_{MLE} is given by:

$$f_{\theta_{MLE}}(x) = \begin{cases} n\frac{x^{n-1}}{\theta^n} \ for \ 0 \leq x \leq \theta \\ 0 \ otherwise \end{cases}$$

Thus, we may evaluate $E_{\theta_{MLE}}(X) = \int\limits_0^\theta x f_{\theta_{MLE}} dx = \int\limits_0^\theta x n\frac{x^{n-1}}{\theta^n} dx = \frac{nx^{n+1}}{(n+1)\theta^n}\Big|_0^\theta = \frac{n}{n+1}\theta$

Thus, we have:

$Bias_{\theta_{MLE}} = E_{\theta_{MLE}}\left(\hat{\theta}\right) - \theta = \frac{n}{n+1}\theta - \theta = -\frac{1}{n+1}\theta$

(B):

For θ_{MOM} we have:

$E_{\theta_{MOM}}(X) = \int\limits_0^\theta x\frac{1}{\theta} dx = \frac{x^2}{2\theta}\Big|_0^\theta = \frac{\theta}{2}$

$Bias_{\theta_{MOM}} = E_{\theta_{MOM}}\left(\hat{\theta}\right) - \theta = \frac{2\theta}{2} - \theta = \theta - \theta = 0$

(C): Unbiased MLE estimator

Let θ_{adj} be the adjusted MLE estimator. We aim to show θ_{MLE} is bias adjusted.

Define θ_{adj} as follows:

$\theta_{adj} = \frac{n+1}{n}max(X_1, \ldots, X_n) = \frac{n+1}{n}X_{max}$

Note, and referring back to the Bias formulation in (A):

$Bias_{\theta_{adj}} = E_\theta\left(\theta_{adj}\right) - \theta = \frac{n+1}{n}\frac{n}{n+1}\theta - \theta = \theta - \theta = 0$

Thus, we note θ_{adj} is the bias adjusted MLE estimator.

Plot

```
theta <- 250
n <- (1:5000)

mle_bias <- rep(NA, 5000)
mom_bias <- rep(NA, 5000)
for (i in 1:5000){
  mle_bias[i] <- (1 / (n[i] + 1)) * theta
  mom_bias <- 0
}
df <- data.frame(n,mle_bias,mom_bias)

ggplot(df, aes(n)) +
  geom_line(aes(y=mle_bias), colour="red") +
  geom_line(aes(y=mom_bias), colour="green") +
  geom_text(x=200, y=20, label="MLE") +
  geom_text(x=2000, y=5, label="MOM, MLE adjusted") +
  ggtitle("Bias of MLE, MOM, and adjusted MLE Estimators") +
  labs(y="Bias", x = "n (observations)")
```
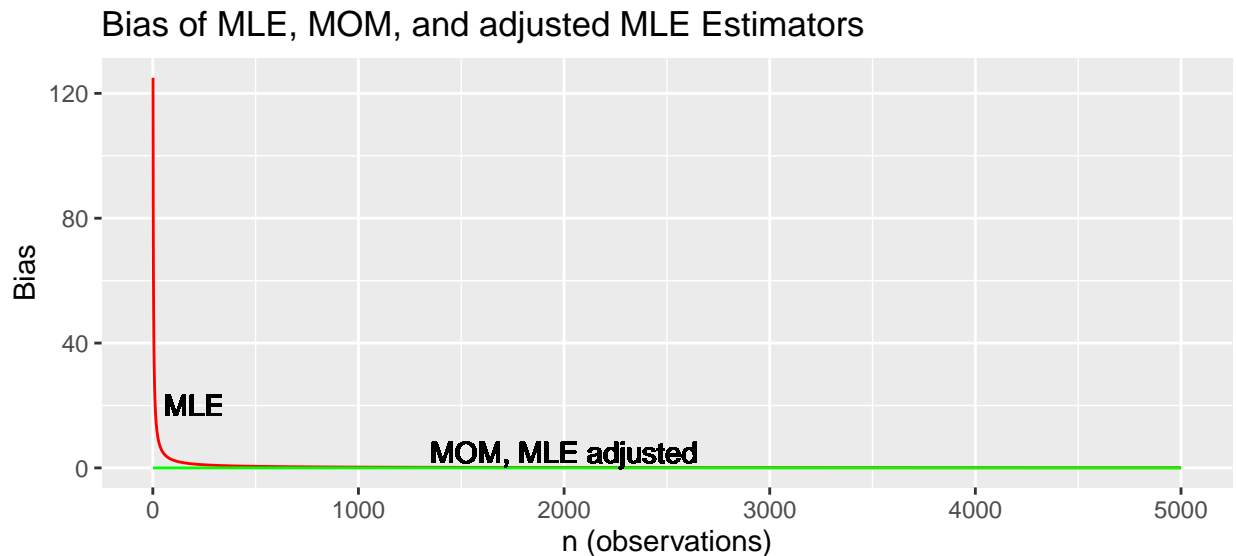

The above results indicate the MLE estimate is biased (but unbiased asymptotically as n increases) and the MOM estimate is unbiased.

2. Calculate the variance of each estimator (including any new bias-corrected ones). What happens as sample size grows? Create an analogous plot to the one above.

(A):

For θ_{MLE} we have:

$$E_{\theta_{MLE}}(X^2) = \int_0^\theta x^2 f_{\theta_{MLE}} dx = \int_0^\theta x^2 n\frac{x^{n-1}}{\theta^n} dx = \left.\frac{nx^{n+2}}{(n+2)\theta^n}\right|_0^\theta = \frac{n}{n+2}\theta^2$$

$$Var_{\theta_{MLE}} = E_{\theta_{MLE}}\left(\hat{\theta}^2\right) - \left(E_{\theta_{MLE}}\left(\hat{\theta}\right)\right)^2$$

$$= \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\theta\right)^2 = \frac{n}{(n+1)^2(n+2)}\theta^2$$

3

(B):

For θ_{MOM} we have:

$$E_{\theta_{MOM}}(X^2) = \int_0^\theta x^2 \tfrac{1}{\theta} dx = \tfrac{x^3}{3\theta}\Big|_0^\theta = \tfrac{\theta}{3}$$

Thus, we have:

$$Var_{\theta_{MOM}} = E_{\theta_{MOM}}\left(\hat\theta^2\right) - \left(E_{\theta_{MOM}}\left(\hat\theta\right)\right)^2$$

$$= \tfrac{\theta^2}{3} - \left(\tfrac{\theta}{2}\right)^2 = \tfrac{\theta^2}{3} - \tfrac{\theta^2}{4} = \tfrac{\theta^2}{12}$$

Thus:

$$V_\theta\left(\hat\theta_n\right) = \tfrac{4}{n^2} n V_\theta(X) = \tfrac{4}{n^2} n \tfrac{\theta^2}{12} = \tfrac{\theta^2}{3n}$$

(C):

For θ_{adj} we have:

$$Var_\theta(\theta_{adj}) = Var_\theta\left(\tfrac{n+1}{n}\theta_{MLE}\right) = \left(\tfrac{n+1}{n}\right)^2 Var_\theta(\theta_{MLE})$$

As we again solved the variance of the initial MLE estimator above, we have:

$$Var_\theta(\theta_{adj}) = \left(\tfrac{n+1}{n}\right)^2 \cdot \tfrac{n}{(n+1)^2(n+2)}\theta^2 = \tfrac{1}{n(n+2)}\theta^2$$

Plot & Commentary

```
theta <- 250
n <- (1:50)

mle_var <- rep(NA, 50)
mom_var <- rep(NA, 50)
mle_adj_var <- rep(NA, 50)
for (i in 1:50){
  mle_var[i] <- ((theta ^ 2) * (n[i]) )/ (((n[i] + 1) ^ 2) * (n[i] + 2))
  mom_var[i] <- (theta ^ 2) / (3 * n[i])
  mle_adj_var[i] <- (theta ^ 2) / ((n[i] * (n[i] +2)))
}

df <- data.frame(n,mle_var,mom_var, mle_adj_var)

ggplot(df, aes(n)) +
  geom_line(aes(y=mle_var), colour="red") +
  geom_line(aes(y=mom_var), colour="green") +
  geom_line(aes(y=mle_adj_var), colour="blue") +
  geom_text(x=2.5, y=1000, label="MLE") +
  geom_text(x=5, y=20000, label="MOM") +
  geom_text(x=10, y=5000, label="MLE adj.") +
  ggtitle("Variance of MLE, MOM, and adjusted MLE Estimators") +
  labs(y="variance", x = "n (observations)")
```

Variance of MLE, MOM, and adjusted MLE Estimators
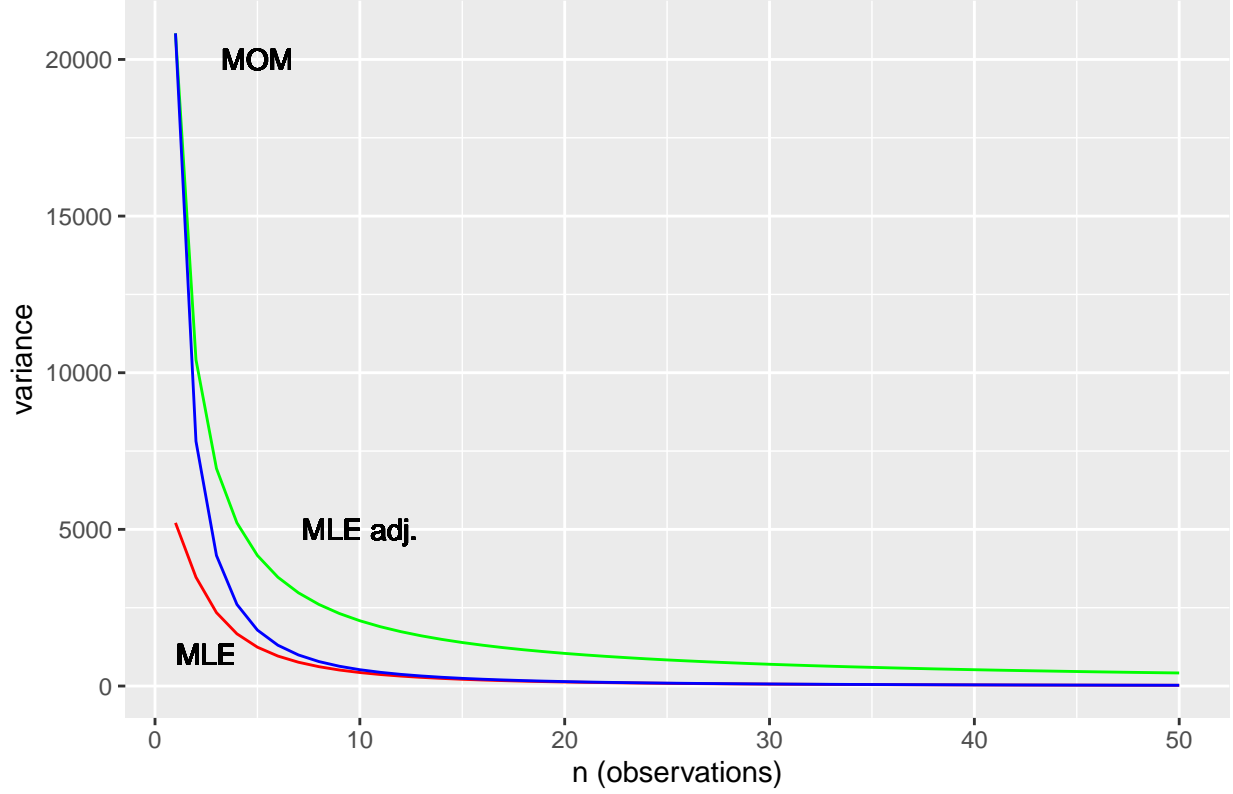
As the sample size grows, the variance of all three estimators converge to 0, though the mle estimator converges fastest (relatively), followed by the adjusted mle estimator.

3. Combine the notions of bias and variance into a third plot that shows how the Mean Squared Error changes as a function of sample size. Based on this plot, which estimator would you use and why?

Note, the formula for the M.S.E. is given as follows:

$$M.S.E. \left(\hat{\theta} \right) = Var(\hat{\theta}) + Bias(\hat{\theta})^2$$

Thus, using the above relations we have:

(A):

$$M.S.E.\,(\theta_{MLE}) = Var(\theta_{MLE}) + Bias(\theta_{MLE})^2 = \frac{n}{(n+1)(n+2)}\theta^2 + \left(-\frac{1}{n+1}\theta \right)^2 = \frac{n\theta^2}{(n+1)(n+2)} + \frac{\theta^2}{(n+1)^2} = \frac{2\theta^2}{(n+1)(n+2)}$$

(B):

Similarly, for the MOM we have:

$$M.S.E.\,(\theta_{MOM}) = Var(\theta_{MOM}) + Bias(\theta_{MOM})^2 = \frac{\theta^2}{3n} + 0^2 = \frac{\theta^2}{3n}$$

(C):

For the adjusted MLE estimator, we first note it has no bias, like the MOM estimator in (B). Thus, it follows:

$$M.S.E.\,(\theta_{adj}) = Var(\theta_{adj}) + Bias(\theta_{adj})^2 = \frac{1}{n(n+2)}\theta^2 + 0^2 = \frac{1}{n(n+2)}\theta^2$$

Plot & Commentary

```r
theta <- 250
n <- (1:50)

mle_mse <- rep(NA, 50)
mom_mse <- rep(NA, 50)
mle_adj_mse <- rep(NA, 50)
for (i in 1:50){
  mle_mse[i] <- ((theta ^ 2) * 2 )/ ((n[i] + 1) * (n[i] + 2))
  mom_mse[i] <- (theta ^ 2) / (3 * n[i])
  mle_adj_mse[i] <- (theta ^ 2) / ((n[i] * (n[i] +2)))
}

df <- data.frame(n,mle_mse,mom_mse, mle_adj_mse)

ggplot(df, aes(n)) +
  geom_line(aes(y=mle_mse), colour="red") +
  geom_line(aes(y=mom_mse), colour="green") +
  geom_line(aes(y=mle_adj_mse), colour="blue") +
  geom_text(x=3.5, y=500, label="MLE adj.") +
  geom_text(x=20, y=2500, label="MOM") +
  geom_text(x=8, y=3000, label="MLE") +
  ggtitle("MSE of MLE, MOM, and adjusted MLE Estimators") +
  labs(y="MSE", x = "n (observations)")
```


MSE of MLE, MOM, and adjusted MLE Estimators
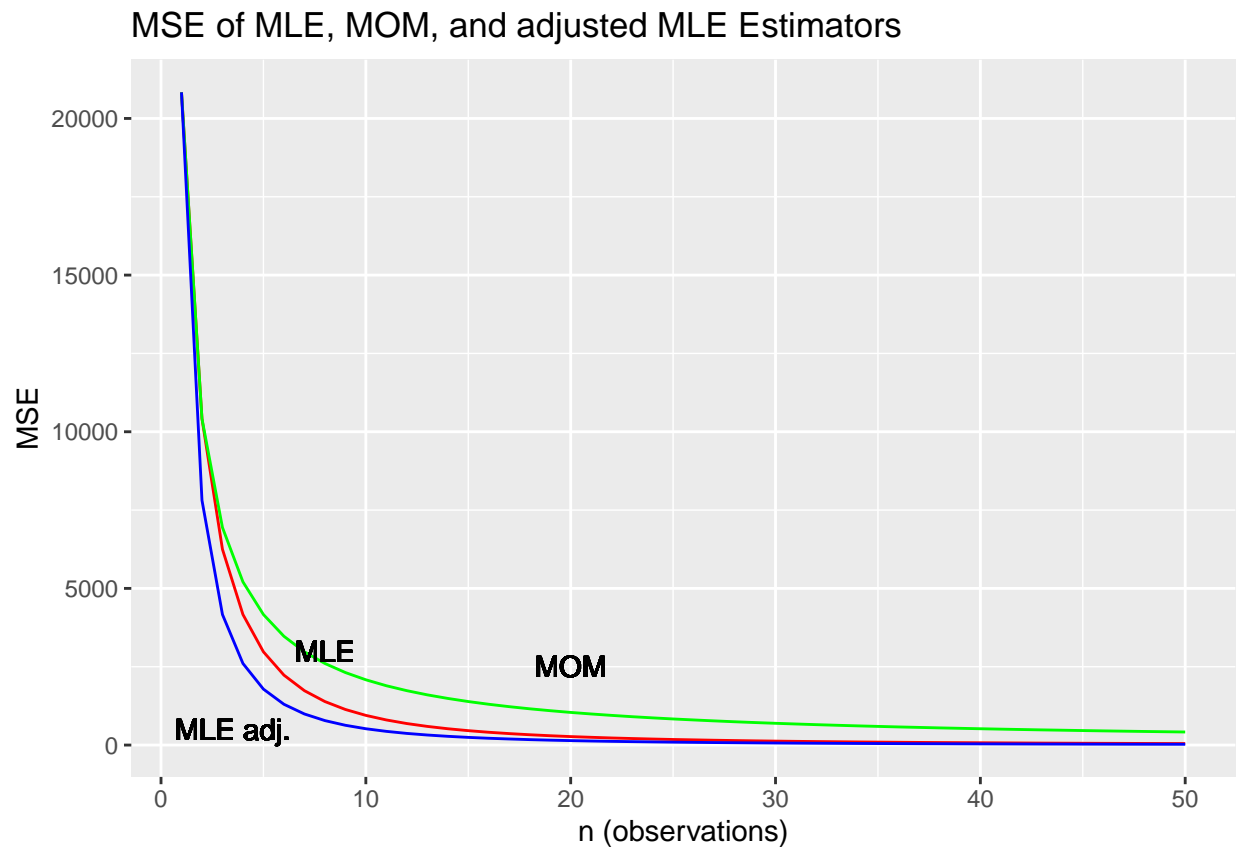
Note, as shown in the above plot, the adjusted MLE estimate converges 'fastest' to 0. Thus, we would anticipate that for a smaller number of observations, the adjusted MLE will produce estimates with lower MSE than the unadjusted MLE and the MOM estimators.

Though, if I'm being cheeky, for $n = 1$, both the adjusted MLE and the MOM have the same MSE. (In that instance... flip a coin? Or in all seriousness, get one more observation and use the adjusted MLE.)

4. Using the method that we saw in class based on Markov's Inequality, assess whether each of these estimators is consistent.

We can demonstrate consistency for each estimator by showing the following two relations:

(1): $\lim_{n \to \infty} Var\left(\hat{\theta}_n\right) = 0$

(2): $\lim_{n \to \infty} Bias\left(\hat{\theta}_n\right) = 0$

Thus, if both (1) and (2) hold for an estimator $\hat{\theta}_n$, we may take advantage of Markov's Inequality, noting that for any constant c $(c > 0)$:

$\lim_{n \to \infty} Pr\left(\mid \hat{\theta}_n - \theta \mid > c\right) = 0$ and we say $\hat{\theta}_n$ is a consistent estimator.

(A):

$Bias_{\theta_{MLE}} = E_{\theta_{MLE}}\left(\hat{\theta}\right) - \theta = \frac{n}{n+1}\theta - \theta = -\frac{1}{n+1}\theta$

Note, as $n \to \infty$, $Bias_{\theta_{MLE}} \to 0$. Thus, we say θ_{MLE} is assymptotically unbiased. We have thus satisfied condition (1).

For condition (2), we note the variance of θ_{MLE}:

$V_\theta\left(\hat{\theta}_{MLE}\right) = \frac{n}{(n+1)^2(n+2)}\theta^2$

Taking the limit as $n \to \infty$ shows $Var_{\theta_{MLE}} \to 0$ as well. Thus we have satisfied condition (2).

As conditions (1) and (2) are satisfied, we thus note, for any constant c $(c > 0)$:

$\lim_{n \to \infty} Pr\left(\mid \hat{\theta}_{MLE} - \theta \mid > c\right) = 0$ and we say $\hat{\theta}_{MLE}$ is a consistent estimator.

(B):

Note, for the estimator θ_{MOM}, the estimator is unbiased, thus we satisfy condition (1).

For condition (2) we note the variance of θ_{MOM}:

$V_\theta\left(\hat{\theta}_{MOM}\right) = \frac{\theta^2}{3n}$

For $n \to \infty$, $Var\left(\hat{\theta}_{MOM}\right) \to 0$.

As conditions (1) and (2) are satisfied, we may then note, for any constant c $(c > 0)$:

$\lim_{n \to \infty} Pr\left(\mid \hat{\theta}_{MOM} - \theta \mid > c\right) = 0$ and we say $\hat{\theta}_{MOM}$ is a consistent estimator.

(C):

Note, for the estimator θ_{adj}, the estimator is unbiased, thus we satisfy condition (1).

Thus, to satisfy condition (2) we note the variance of θ_{adj}:

$Var_\theta\left(\theta_{adj}\right) = \frac{1}{n(n+2)}\theta^2$

We then note, as $n \to \infty$, $Var_\theta\left(\theta_{adj}\right) \to 0$. Thus we have satisfied condition (2).

As conditions (1) and (2) have been satisfied, then we can say, for any constant c $(c > 0)$:

$\lim_{n \to \infty} Pr\left(\mid \hat{\theta}_{adj} - \theta \mid > c\right) = 0$ and we say $\hat{\theta}_n$ is a consistent estimator.

5. What is the sampling distribution of each statistic? For the MOM, consider both the Irwin-Hall distribution and a sensible approximation based on the Central Limit Theorem.

(A):

Note, the cdf of $\hat{\theta}_{MLE}$ is given as follows:

$$f(u) = \left\{ \begin{array}{c} 0 \ for \ u \leq \theta \\ \left(\frac{u}{\theta}\right)^n \ for \ 0 < u < \theta \\ 1 \ for \ u \geq \theta \end{array} \right\}$$

Thus, to determine the pdf of $\hat{\theta}_{MLE}$, we derive the above equation with respect to u. Thus gives us:

$\frac{\partial\left(\left(\frac{u}{\theta}\right)^n\right)}{\partial u} = \frac{nu^{n-1}}{\theta^n}$

Thus, for $0 < u < \theta$ the pdf of $\hat{\theta}_{MLE}$ is $\frac{nu^{n-1}}{\theta^n}$

(B):

To determine the sampling distribution of $\hat{\theta}_{MOM}$, first note the following:

For $i = 1, 2, ...$

$X_i \sim Unif(0, \theta)$

$Y_i \sim Unif(0, 1)$

$U = \sum\limits_{i=1}^{n} Y_i \sim IrwinHall(n)$

The follows holds between X and Y:

$\frac{1}{\theta}X \equiv Y \sim Unif(0, 1)$

It then follows that:

$\sum\limits_{i=1}^{n} Y_i = \frac{1}{\theta}\sum\limits_{i=1}^{n} X_i \sim IrwinHall(n)$

Note:

$\hat{\theta}_{MOM} = 2\bar{X} = 2\frac{1}{n}\sum\limits_{i=1}^{n} X_i$

Thus:

$\frac{\hat{\theta}_{MOM}}{\theta} = \frac{2}{\theta n}\sum\limits_{i=1}^{n} X_i = \frac{2}{\theta n}U \sim \frac{2}{n}IrwinHall(n)$

It then follows, removing the denominator from the above relation, that:

$\hat{\theta}_{MOM} \sim \frac{2\theta}{n}IrwinHall(n)$

Or, put differently:

$\frac{n\hat{\theta}_{MOM}}{2\theta} \sim IrwinHall(n)$

Approximation

Via application of the CLT, we may note that the IrwinHall(n) distribution converges to the standard normal as $n \to \infty$. Thus, bearing in mind the above relations, we may say that as $n \to \infty$, $\frac{n\hat{\theta}_{MOM}}{2\theta} \sim N(0, 1)$, or phrase the relation alternatively in terms of $\hat{\theta}_{MOM}$

To that end, we note that the distribution of $\hat{\theta}_{MOM}$ converges to the normal distribution with mean $E(\hat{\theta}_{MOM}) = \theta$ and variance $Var(\hat{\theta}_{MOM}) = \frac{\theta^2}{3n}$ (noting our prior work in parts 1 and 2 (B), resp.).

Thus, we may say, as $n \to \infty$, $\hat{\theta}_{MOM} \sim N\left(\theta, \frac{\theta^2}{3n}\right)$, which we will use for the plot below.

(C):

Similar to part (A), note the cdf of $\hat{\theta}_{adj}$ is given as follows:

$$f(u) = \left\{ \begin{array}{c} 0 \ for \ u \leq \theta \\ \frac{n+1}{n} \left(\frac{u}{\theta}\right)^n \ for \ 0 < u < \theta \\ 1 \ for \ u \geq \theta \end{array} \right\}$$

Thus, to determine the pdf of $\hat{\theta}_{adj}$, we derive the above equation with respect to u. Thus gives us:

$\frac{\partial\left(\frac{n+1}{n}\left(\frac{u}{\theta}\right)^n\right)}{\partial u} = \frac{(n+1)u^{n-1}}{\theta^n}$

Thus, for $0 < u < \theta$ the pdf of $\hat{\theta}_{adj}$ is $\frac{(n+1)u^{n-1}}{\theta^n}$

6. Create a plot of the sampling distribution of each estimator using $n = 10$. Construct the empirical distribution via simulation and overlay the appropriate exact or approximate analytical form (each plot should be a curve overlayed on a histogram. See slides.)
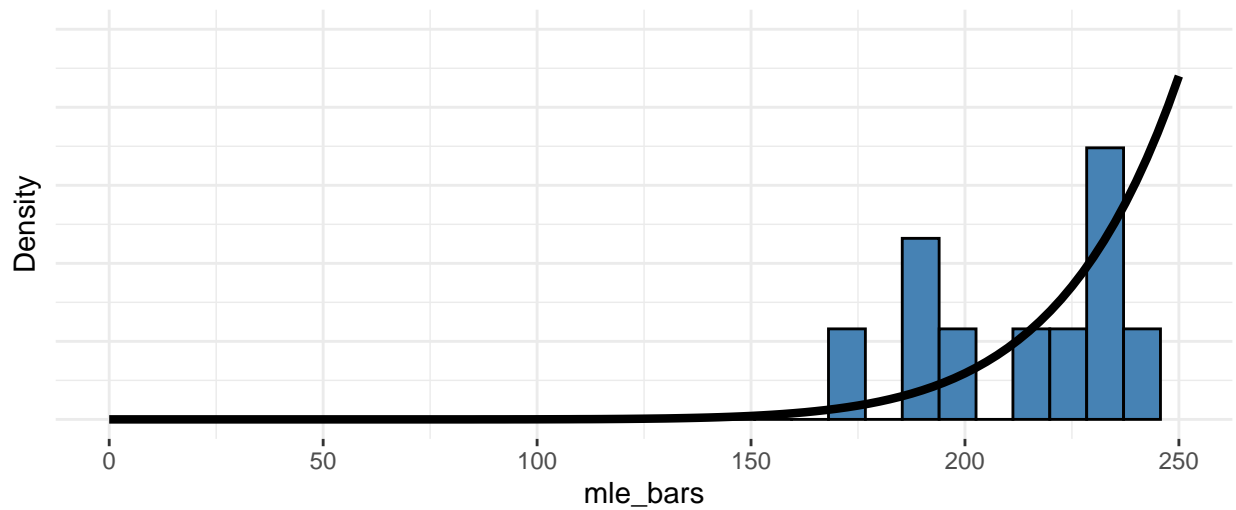
(A):

```
set.seed(84)
theta <- 250
n <- 10
x <- (0:250)

mle_bars <- rep(NA, 10)
for (i in 1:10) {
  x <- rdunif(n, theta, a = 0)
  mle_bars[i] <- max(x)
}

eq = function(x){((n +1)* (x^(n-1))) / (theta ^ n)}

ggplot(data.frame(mle_bars), x=c(0, 250), aes(mle_bars)) +
  geom_histogram(aes(y=..density..),
                 fill = "steelblue", color = "black") +
  theme_bw()  +
  theme(panel.border = element_blank(),
        axis.line.y = element_blank(),
        axis.ticks.y = element_blank(),
        axis.text.y = element_blank()) +
  ylab("Density") +
  stat_function(fun = eq, lwd = 1.5) +
  ylim(0, .05) +
  xlim(0, 250) +
  ggtitle("Distribution of MLE Estimator")
```
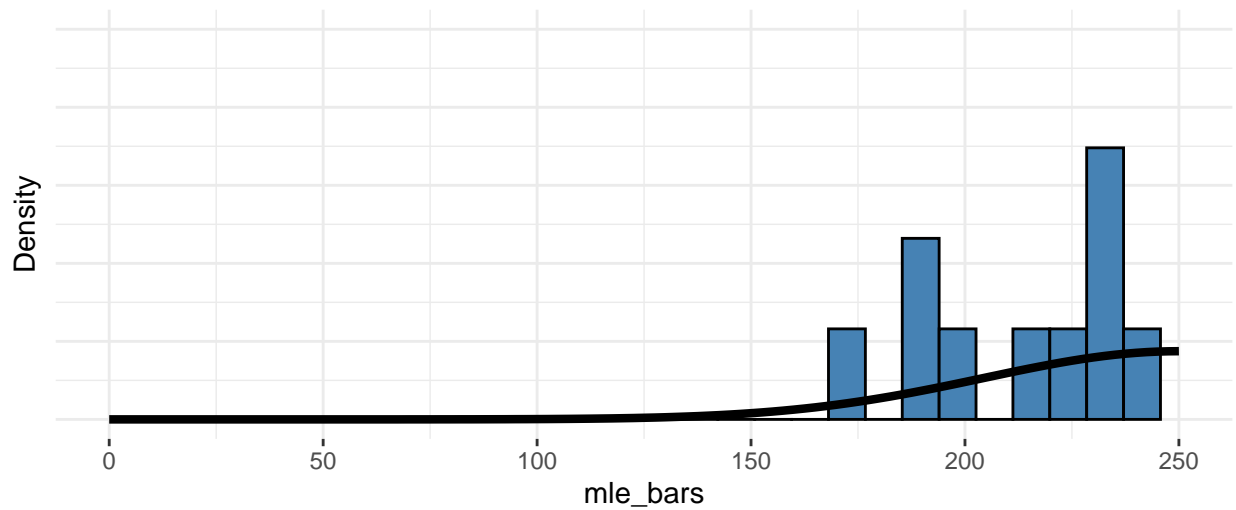
Distribution of MLE Estimator

(B):

```
set.seed(84)
theta <- 250
n <- 10
x <- (0:250)

mom_bars <- rep(NA, 10)
for (i in 1:10) {
  x <- rdunif(n, theta, a = 0)
  mom_bars[i] <- 2 * mean (x)
}

### Use prior problems to substitute mean and variance with mean and variance of the MOM estimator

ggplot(data.frame(mom_bars), x=c(0, 250), aes(mle_bars)) +
  geom_histogram(aes(y=..density..),
                 fill = "steelblue", color = "black") +
  theme_bw()  +
  theme(panel.border = element_blank(),
        axis.line.y = element_blank(),
        axis.ticks.y = element_blank(),
        axis.text.y = element_blank()) +
  ylab("Density") +
  stat_function(fun = dnorm,
                args = list(mean = theta, sd = sqrt(theta ^2 / (3 * n))),
                lwd = 1.5)+
  ylim(0, .05) +
  xlim(0, 250) +
  ggtitle("Distribution of MOM Estimator w/Normal")
```

Distribution of MOM Estimator w/Normal

Commentary

In particular, comparing the curvature in the above graph to the other estimators provides more information supporting the use of an MLE estimator, as even asymptotically, the MOM estimator will tend to understimate the value of θ more so than the MLE estimators.
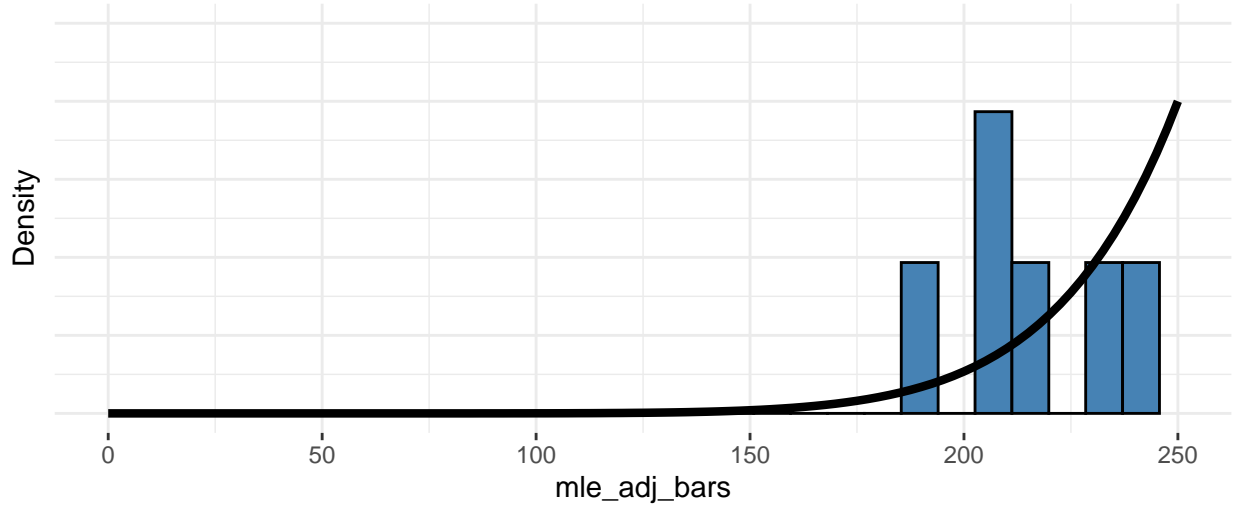
(C):

```
set.seed(84)
theta <- 250
n <- 10
x <- (0:250)

mle_adj_bars <- rep(NA, 10)
for (i in 1:10) {
  x <- rdunif(n, theta, a = 0)
  mle_adj_bars[i] <- ((n+1) / n ) * max(x)
}

eq = function(x){(n * (x^(n-1))) / (theta ^ n)}


ggplot(data.frame(mle_adj_bars), x=c(0, 250), aes(mle_adj_bars)) +
  geom_histogram(aes(y=..density..),
                 fill = "steelblue", color = "black") +
  theme_bw()  +
  theme(panel.border = element_blank(),
        axis.line.y = element_blank(),
        axis.ticks.y = element_blank(),
        axis.text.y = element_blank()) +
  ylab("Density") +
  stat_function(fun = eq, lwd = 1.5) +
  ylim(0, .05) +
  xlim(0, 250) +
  ggtitle("Distribution of adjusted MLE Estimator")
```

11

Distribution of adjusted MLE Estimator

7. Form two different 95% confidence intervals for θ by using pivotal statistics inspired by each estimator.

(1):

Let our pivotal statistic for $\hat{\theta}_{MLE}$ be:

$\eta = \frac{max(X_i)}{\theta} = \frac{\hat{\theta}_{MLE}}{\theta}$

We need to verify this is a pivotal statistic. To do so, we first note:

$\hat{\theta}_{MLE} \in (0, 1)$

Then, taking advantage of the prior cdf of $\hat{\theta}_{MLE}$, we may say:

$Pr(\frac{max(X_i)}{\theta} \leq c) = Pr(max(X_i) \leq \theta c)$

Taking advantage of the independent of $X_1, ..., X_n$, we may then note:

$Pr(\frac{max(X_i)}{\theta} \leq c) = \prod_{i=1}^{n} Pr(X_i \leq \theta c) = \prod_{i=1}^{n} \frac{\theta c}{\theta} = \prod_{i=1}^{n} c$

It then follows that:

$Pr(\frac{max(X_i)}{\theta} \leq c) = c^n$ for $c \in (0, 1)$

Note, the pcd doesn't depend on the parameter θ, and thus we've verified our statistic $\eta = \frac{max(X_i)}{\theta} = \frac{\hat{\theta}_{MLE}}{\theta}$ is a pivotal statistic.

Then, we construct the 95% confidence interval.

To do so, note $\alpha = 0.05$, giving us:

$Pr(q_1 < \frac{max(X_i)}{\theta} < q_2) = 0.95$

Work in progress

We derive the cdf above with respect to n to get our pdf of the pivotal statistic:

$\frac{\partial x^n}{\partial n} = nx^{n-1}$

Thus we may start to evaluate the confidence interval:

$Pr(q_1 < \frac{max(X_i)}{\theta} < q_2) = 0.95$

To evaluate this, we note, for $q_1 = 0.025 = \frac{\alpha}{2}$ and $q_2 = 0.975 = 1 - \frac{\alpha}{2}$:

$Pr(q_1 < \frac{max(X_i)}{\theta}) = 0.025$

$\int_0^{q_1} nx^{n-1} = 0.025$

Solving for q_1 gives us:

$q_1 = \sqrt{0.025}$ as $(q_1^n = 0.025)$

and

$Pr(\frac{max(X_i)}{\theta} < q_2) = 0.975$

$\int_0^{q_2} nx^{n-1} = 0.975$

Solving for q_2 gives us:

$q_2 = \sqrt{0.975}$ as $(q_2^n = 0.025)$

Which then gives us:

$Pr(\sqrt{0.025} < \frac{max(X_i)}{\theta} < \sqrt{0.975}) = 0.95$

Isolating $max(X_i)$ gives us:

$Pr(\theta\sqrt{0.025} < max(X_i) < \theta\sqrt{0.975}) = 0.95$

Substituting our estimator then gives us:

$Pr(\theta\sqrt{0.025} < \hat{\theta}_{MLE} < \theta\sqrt{0.975}) = 0.95$

Giving us the 95% confidence interval of

$(\theta\sqrt{0.025}, \theta\sqrt{0.975})$

Note, as the pivotal statistic does not include the parameter n, we may construct the same pivotal statistic to use for the construction of a 95% confidence interval for $\hat{\theta}_{adj}$. We may also extend this observation to say: Our pivotal statistics and confidence intervals for the estimators $\hat{\theta}_{MLE}$ and $\hat{\theta}_{adj}$ are the same.

(2):

For $\hat{\theta}_{MOM}$, we need to construct a different pivotal statistic.

To that end, note:

$\frac{n\hat{\theta}_{MOM}}{2\theta} \sim IrwinHall(n)$

However, we want a pivotal statistic that does not include the parameter n.

Thus, note:

$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

Inserting this into the above relation gives us:

$\frac{n\hat{\theta}_{MOM}\frac{1}{n}\sum_{i=1}^{n} X_i}{2\theta} = \frac{\hat{\theta}_{MOM}\sum_{i=1}^{n} X_i}{2\theta}$

Note:

$\frac{1}{\theta}\sum_{i=1}^{n} X_i \sim IrwinHall(n)$

The Jist

The statistic $\frac{\hat{\theta}_{MOM}}{2\theta^2}$ is composed of the product of two statistics whose distribution is given by the Irwin Hall distribution (Irwin Hall(n)). Taking advantage of the Central Limit Theorem, each of the Irwin Hall distributions converge to the Standard normal. Thus, the product of these two distributions is the χ^2 distribution with n degrees of freedom.

Thus, by also including the factor $\frac{1}{\theta}$ into the relation noted prior, it holds that:

$\frac{\hat{\theta}_{MOM}}{2\theta^2} \sim \chi_n^2$

We then note, the statistic $\frac{\hat{\theta}_{MOM}}{2\theta^2}$ does not depend on the parameters of its distribution, namely n.

Constructing the 95% confidence interval using the above information gives us:

$Pr(q_1 < \frac{\hat{\theta}_{MOM}}{2\theta^2} < q_2) = 1 - \alpha = 0.95$

Isolating $\hat{\theta}_{MOM}$ gives us the relation:

$Pr(2\theta^2 \cdot q_1 < \hat{\theta}_{MOM} < 2\theta^2 \cdot q_2) = 0.95$

Thus, we have the confidence interval as:

$\left(2\theta^2 q_1, 2\theta^2 q_2\right)$, where q_1 and q_2 are the quantiles (0.025 and 0.975 resp.) of the χ^2 distribution with n degrees of freedom.