A2Q1

1. $E(\tilde{\theta}_1) = E(2\overline{X}) = 2E(\overline{X}) = 2E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$

$= \frac{2}{n} E\left(\sum_{i=1}^{n} X_i\right) = \frac{2}{n}\sum_{i=1}^{n}(E(X_i)) = \frac{2}{n} n \left(\frac{0+\theta}{2}\right)$

$= 2\left(\frac{\theta}{2}\right) = \boxed{\theta}$

$\tilde{\theta}_1$ is an unbiased estimator of θ

---

2. $\tilde{\theta}_2 = \max(X_1, \cdots, X_n)$

CDF of $\tilde{\theta}_2$:

$F_{\tilde{\theta}_2}(y) = P(\max(X_1, \cdots, X_n) \leq y)$

$= P(X_1 \leq y \cup X_2 \leq y \cdots \cup X_n \leq y)$

$= \prod_{i=1}^{n} P(X_i \leq y)$ because X_i are independent

$= (y/\theta)^n = y^n/\theta^n$

$f_{\tilde{\theta}_2}(y) = \frac{d}{dy} F_{\tilde{\theta}_2}(y) = n(y^{n-1}/\theta^n)$, $0 < y < \theta$

$E(\tilde{\theta}_2) = \int_0^\theta y f_{\tilde{\theta}_2}(y)\,dy = \int_0^\theta n\left(\frac{y^{n-1}}{\theta^n}\right) y\,dy$

$= \int_0^\theta \frac{n}{\theta^n} y^n\,dy = \frac{n}{\theta^n}\left[\frac{y^{n+1}}{n+1}\right]_{y=0}^{y=\theta}$

$= \frac{n}{\theta^n}\frac{\theta^{n+1}}{n+1} = \boxed{\frac{n\theta}{n+1}} \neq \theta$

$\tilde{\theta}_2$ is a biased estimator of θ

---

3. $Var(\tilde{\theta}_1) = Var(2\overline{X})$

$= 2^2 Var(\overline{X})$

$= 4 Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$

$= \frac{4}{n^2} Var\left(\sum_{i=1}^{n} X_i\right)$

$= \frac{4}{n^2}\sum_{i=1}^{n}(Var(X_i))$ because X_i are independent

$$= \frac{4}{n^2} n\left(\frac{(\theta-0)^2}{12}\right)$$

$$= \frac{4}{n}\frac{\theta^2}{12} = \boxed{\frac{\theta^2}{3n}}$$

4.

```r
set.seed(3052020)
sink("simulation.txt")

n = 100
theta = 5
theta1Samples = c()
theta2Samples = c()
numSamples = 10000

for (i in 1:numSamples) {
    sample = runif(n, 0, theta)
    theta1Samples[i] = 2 * sum(sample) / n
    theta2Samples[i] = max(sample)
}

cat("mean of theta1Samples: ", mean(theta1Samples), "\n")
cat("variance of theta1Samples: ", var(theta1Samples), "\n")
cat("theoretical mean obtained from E(theta1): ", theta, "\n")
cat("theoretical variance obtained from Var(theta1): ", theta ^ 2 / (3 * n), "\n")
cat("======", "\n")

cat("mean of theta2Samples: ", mean(theta2Samples), "\n")
cat("variance of theta2Samples: ", var(theta2Samples), "\n")
cat("======", "\n")

realTheta = rep(theta, numSamples)
mse1 = sum((theta1Samples - realTheta) ^ 2) / numSamples
mse2 = sum((theta2Samples - realTheta) ^ 2) / numSamples

cat("approximation of mse(theta1): ", mse1, "\n")
cat("approximation of mse(theta2): ", mse2, "\n")

sink()
```

```
mean of theta1Samples:  5.000162
variance of theta1Samples:  0.08273941
theoretical mean obtained from E(theta1):  5
theoretical variance obtained from Var(theta1):  0.08333333
======
mean of theta2Samples:  4.951235
variance of theta2Samples:  0.002387045
======
approximation of mse(theta1):  0.08273116
approximation of mse(theta2):  0.004764837
```

$\widehat{MSE}(\theta_1) = 0.08273116$

Comparison: The sample mean & sample variance of the 10000 samples of $\tilde{\theta}_1$ are closer to the theoretical results I've obtained for $E(\tilde{\theta}_1)$ and $Var(\tilde{\theta}_1)$

5. $\widehat{MSE}(\theta_2) = 0.004764837$

6. $\tilde{\theta}_1$ over $\tilde{\theta}_2$: The sample mean & sample variance of the 10000 samples of $\tilde{\theta}_1$ are closer to the theoretical results I've obtained for $E(\tilde{\theta}_1)$ and $Var(\tilde{\theta}_1)$. Also, $\tilde{\theta}_1$ is an unbiased estimator of θ whereas $\tilde{\theta}_2$ is biased.

$\tilde{\theta}_2$ over $\tilde{\theta}_1$: From simulation, $\widehat{MSE}(\tilde{\theta}_2) < \widehat{MSE}(\tilde{\theta}_1)$