STAT 512 su 2021 hw 9

1. Suppose $Y_1, \dots, Y_n \overset{\text{ind}}{\sim} \text{Gamma}(2, \beta)$.

   (a) Write down the likelihood function for β based on $Y_1, \dots, Y_n$.

   We have

   $$\begin{aligned}\mathcal{L}(\beta; X_1, \dots, X_n) &= \prod_{i=1}^{n} \frac{1}{\Gamma(2)\beta^2} Y_i \exp\left[-\frac{Y_i}{\beta}\right] \\ &= \left[\frac{1}{\Gamma(2)\beta^2}\right]^n \prod_{i=1}^{n} Y_i \cdot \exp\left[-\frac{\sum_{i=1}^n Y_i}{\beta}\right].\end{aligned}$$

   (b) Write down the log-likelihood function for β based on $Y_1, \dots, Y_n$.

   We have

   $$\ell(\beta, X_1, \dots, X_n) = -2n \log \beta + \sum_{i=1}^{n} \log Y_i - \frac{\sum_{i=1}^n Y_i}{\beta}.$$

   (c) Find an expression for the MLE of β.

   Taking the derivative with respect to β of the log-likelihood and setting this equal to zero gives

   $$\frac{d}{d\beta}\ell(\beta, X_1, \dots, X_n) = -\frac{2n}{\beta} + \frac{\sum_{i=1}^n Y_i}{\beta^2} \overset{\text{set}}{=} 0 \quad \iff \quad \beta = \frac{\bar{Y}_n}{2},$$

   so the MLE for β is $\hat{\beta} = \bar{Y}_n/2$.

   (d) Give the MoMs estimator of β.

   The first population moment of the $\text{Gamma}(2, \beta)$ distribution is 2β, so we set $m_1' = 2\beta$, which gives $\beta = m_1'/2$, so the MoMs estimator of β is

   $$\bar{\beta} = \frac{\bar{Y}_n}{2},$$

   which is the same as the MLE.

2. Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Uniform}(\mu - \theta, \mu + \theta)$.

   (a) Give expressions for the MoMs estimators, call them $\bar{\mu}$ and $\bar{\theta}$, of μ and θ.
   Hint: Based on a sample of size $n = 5$ with the values

   ```
   -0.44 0.27 3.08 0.90 0.51
   ```

the MoMs estimator of θ should have the value `2.062598` *(to double-check your answer).*

To find the MoMs estimators of μ and θ, we set the first two sample moments equal to the first two population moments. We have

$$m'_1 = \mu'_1 = \mu$$
$$m'_2 = \mu'_2 = \frac{[(\mu + \theta) - (\mu - \theta)]^2}{12} + \mu^2 = \frac{\theta^2}{3} + \mu^2.$$

Solving the system of equations for μ and θ gives

$$\bar{\mu} = m'_1$$
$$\bar{\theta} = \sqrt{3(m'_2 - (m'_1)^2)}.$$

(b) Now consider the estimators of θ and μ given by

$$\hat{\theta}_{\text{unbiased}} = \left(\frac{n+1}{n-1}\right)\frac{X_{(n)} - X_{(1)}}{2} \quad \text{and} \quad \hat{\mu} = \frac{X_{(1)} + X_{(n)}}{2},$$

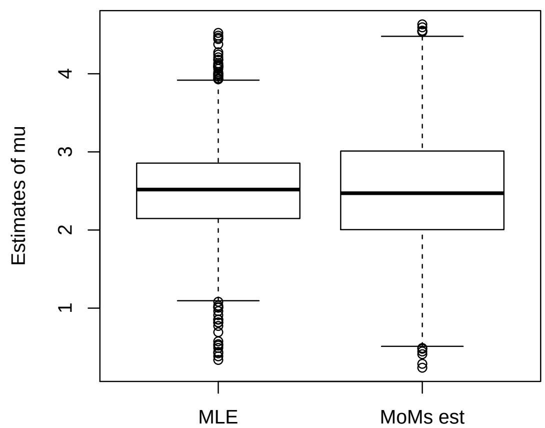
where $\hat{\theta}_{\text{unbiased}}$ is a bias-corrected, or *de-biased* version of the MLE for θ. Run a simulation in which $1{,}000$ random samples of size $n = 5$ are drawn from the $\text{Uniform}(\mu-\theta, \mu+\theta)$ distribution with $\mu = 2.5$ and $\theta = 3$, and compute the estimators $\bar{\mu}$, $\hat{\mu}$, $\bar{\theta}$, and $\hat{\theta}_{\text{unbiased}}$ on each random sample. Then:

i. Make side-by-side boxplots of the $1{,}000$ values of $\bar{\mu}$ and $\hat{\mu}$. You can take a photo of your screen to merge into your pdf to upload. Include R code.

The following R code runs the simulation:

```
mu <- 2.5
theta <- 3
n <- 5
S <- 1000

theta.mle.unbiased <- numeric(S)
mu.mle <- numeric(S)
theta.mom <- numeric(S)
mu.mom <- numeric(S)

for( s in 1:S )
{

  X <- runif(n,mu - theta, mu + theta)

  X1 <- min(X)
```

```
  Xn <- max(X)

  theta.mle.unbiased[s] <- (n+1)/(n-1)*(Xn - X1)/2
  mu.mle[s] <- (X1 + Xn)/2

  theta.mom[s] <- sqrt( 3*(mean(X^2) - mean(X)^2))
  mu.mom[s] <- mean(X)

}
boxplot(mu.mle,mu.mom)
```



ii. Compare the mean squared error of $\bar{\mu}$ and $\hat{\mu}$ based on the 1,000 simulated data sets and comment on which estimator you think is a better estimator of μ.

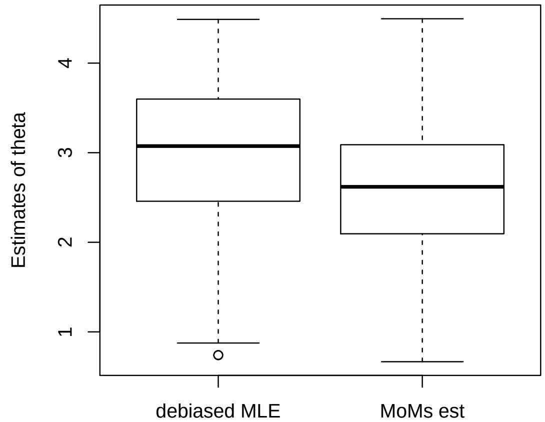
Running

```
mean((mu.mom - mu)^2)
mean((mu.mle - mu)^2)
```

shows that the MSE $\hat{\mu}$ is smaller than MSE $\bar{\mu}$, so $\hat{\mu}$ appears to be a better estimator.

iii. Make side-by-side boxplots of the 1,000 values of $\bar{\theta}$ and $\hat{\theta}_{\text{unbiased}}$. You can take a photo of your screen to merge into your pdf to upload. Include R code.

The boxplots of the $\bar{\theta}$ and $\hat{\theta}_{\text{unbiased}}$ should look like this:



iv. Compare the mean squared error of $\bar{\theta}$ and $\hat{\theta}_{\text{unbiased}}$ based on the 1,000 simulated data sets and comment on which estimator you think is a better estimator of θ.

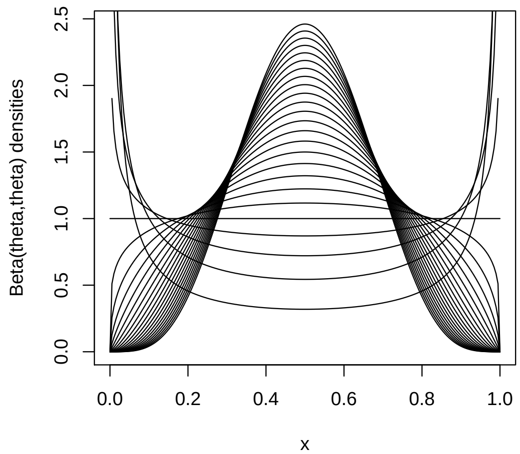
Running

```
mean((theta.mom - theta)^2)
mean((theta.mle.unbiased - theta)^2)
```

shows that $\text{MSE}\,\hat{\theta}_{\text{unbiased}}$ is smaller than $\text{MSE}\,\bar{\theta}$ so $\hat{\theta}_{\text{unbiased}}$ appears to be a better estimator.

Hint: Look at the code from hw 7 for setting up the simulation. To make side-by-side boxplots, just use `boxplot(mu.mle,mu.mom)`*, for example.*

3. Let $X_1, \ldots, X_n \overset{\text{ind}}{\sim} \text{Beta}(\theta, \theta)$, $\theta > 0$. The $\text{Beta}(\theta, \theta)$ pdf is plotted below for all $\theta \in \{.2, .4, .6, \ldots, 5\}$:



(a) Find the MoMs estimator of θ by setting $m_2' = \mu_2'$ (note that $\mu_1' = 1/2$ for all θ).

Writing

$$m_2' = \frac{\theta^2}{(2\theta)^2(2\theta+1)} + \left(\frac{\theta}{\theta+\theta}\right)^2 = \frac{1+\theta}{2+4\theta}$$

and solving for θ gives

$$\bar{\theta} = \frac{2m_2' - 1}{1 - 4m_2'}.$$

(b) Write down the likelihood function $\mathcal{L}(\theta; X_1, \ldots, X_n)$.

The likelihood function is given by

$$\begin{aligned} \mathcal{L}(\theta, X_1, \ldots, X_n) &= \prod_{i=1}^{n} \frac{\Gamma(2\theta)}{\Gamma(\theta)\Gamma(\theta)} X_i^{\theta-1}(1-X_i)^{\theta-1} \\ &= \left[\frac{\Gamma(2\theta)}{\Gamma(\theta)\Gamma(\theta)}\right]^n \left[\prod_{i=1}^{n} X_i(1-X_i)\right]^{\theta-1} \end{aligned}$$

for all $0 < X_1, \ldots, X_n < 1$.

(c) Argue that $\prod_{i=1}^n X_i(1-X_i)$ is a sufficient statistic for θ.

Since we can write down the joint pdf of $X_1, \ldots, X_n$, whic is the same as the likelihood function, as the product of a function of $\prod_{i=1}^n X_i(1-X_i)$ and θ and a function of only the data, we can claim, by the factorization theorem, that $\prod_{i=1}^n X_i(1-X_i)$ is a suficient statistic for θ.

(d) Write down the log-likelihood function $\ell(\theta; X_1, \ldots, X_n)$.

The log-likelihood is given by

$$\ell(\theta; X_1, \ldots, X_n) = n \log \Gamma(2\theta) - 2n \log \Gamma(\theta) + (\theta - 1) \sum_{i=1}^{n} [\log X_i + \log(1 - X_i)]$$

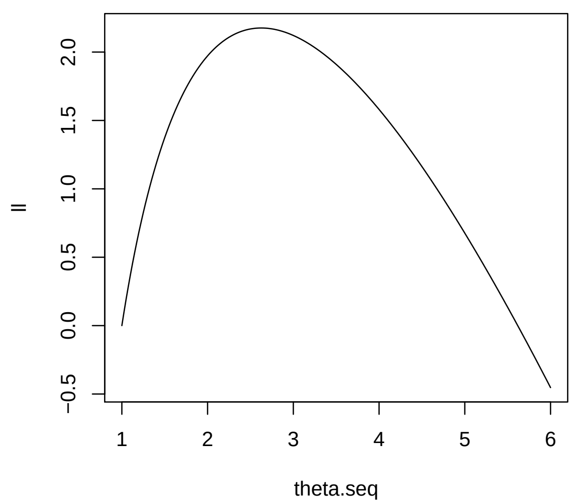
for $0 < X_1, \ldots, X_n < 1$.

(e) Use the following code to plot the log-likelihood function based an observed random sample:

```
X <- c(0.39, 0.48, 0.78, 0.49, 0.52, 0.66, 0.44, 0.43, 0.08, 0.25)
n <- length(X)
theta.seq <- seq(1,6,length=201)
ll <- n*log(gamma(2*theta.seq))-2*n*log(gamma(theta.seq))+(theta.seq-1)*sum(log(X)+log(1-X))
plot(ll ~ theta.seq,type="l")
```

Give your best guess of the value of the MLE $\hat{\theta}$ of θ by looking at the plot.

The plot looks like



The value in `theta.seq` at which the log-likelihood is the largest is the value 2.625. An acceptable answer is "around 2.6".

(f) Compute the value of the MoMs estimator on the same data.

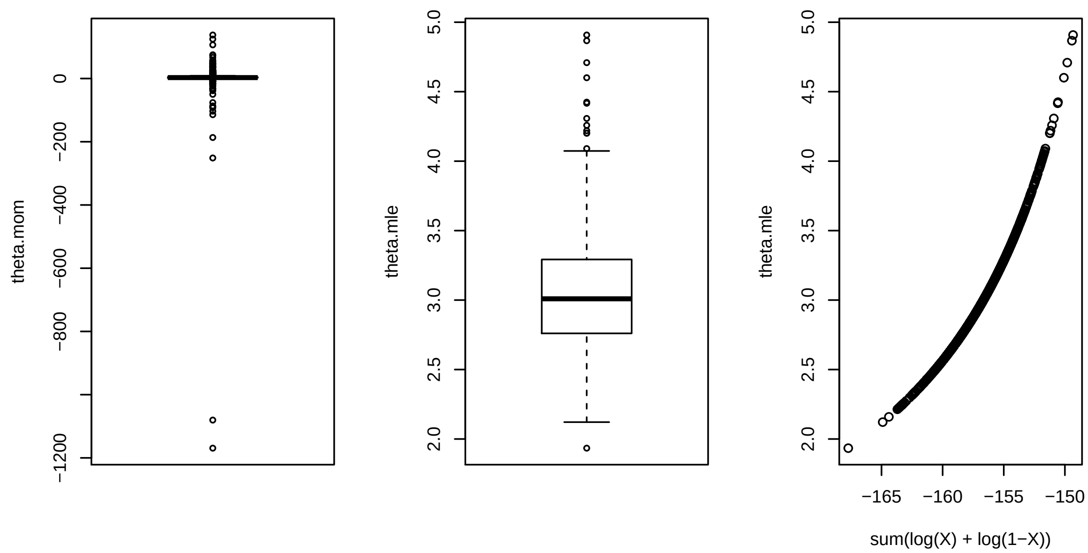
The MoMs estimator has the value -11.31315, given by

```
m2 <- mean(X^2)
theta.mom <- (2 * m2 - 1) / ( 1 - 4 * m2)
```

(g) What is wrong with the value of the MoMs estimator?

The value of θ should be positive. That is, the parameter space is $\Theta = (0, \infty)$, and this value fall outside of the parameter space!

(h) A simulation was run in which 1,000 random samples of size 100 were drawn from the Beta(3, 3) distribution, and the MLE and MoMs estimator were computed on each data set. The figure below shows boxplots of the 1,000 values of the MLE and the MoMs estimator as well as a plot of the values of the MLE versus the value of $\log\left(\prod_{i=1}^n X_i(1 - X_i)\right)$ over the 1,000 simulations.



i. Which estimator, the MLE or the MoMs estimator, appears more reliable?

It appears that the MoMs estimator often takes negative values and sometimes values which are very extreme. It looks like a very unstable estimator. In contrast, the MLE takes values which are centered around the true value $\theta = 3$ and which are always contained inside the parameter space $(0, \infty)$.

ii. What is the significance of the plot on the far right? Why would I show it?

We learned that the MLE is always a function of a sufficient statistic, which is a point in their favor. Though we could not write down the MLE in closed form, we can see from this plot that the MLE for θ really *is* a function of the sufficient statistic

$$\prod_{i=1}^n X_i(1 - X_i) = \exp\left[\sum_{i=1}^n (\log X_i + \log(1 - X_i))\right],$$

because the points fall on a curve. We cannot write down the formula for this curve in a nice way, but it nevertheless represents the function relating $\hat{\theta}$ to $\prod_{i=1}^n X_i(1 - X_i)$.

4. Suppose you observe the following independent realizations of a random variable X:

```
0.46 0.64 0.59 0.93 1.63 0.61 1.75 0.61 0.78 0.81
1.00 0.59 0.77 1.36 1.79 1.15 0.44 0.81 0.71 0.34
```

(a) Assuming that X follows a Gamma distribution, find the MoMs and the MLEs of (α, β). Use R as in the *Lec 11 R code* to find the MLEs.

The following code computes the maximum likelihood estimators:

```
X <- c( 0.46, 0.64, 0.59, 0.93, 1.63, 0.61, 1.75,
        0.61, 0.78, 0.81, 1.00, 0.59, 0.77, 1.36,
        1.79, 1.15, 0.44, 0.81, 0.71, 0.34)
n <- length(X)

negll <- function(x,data)
{
  alpha <- x[1]
  beta <- x[2]
  negll <- - sum( dgamma(data, shape = alpha, scale = beta, log = TRUE ))
  return(negll)
}

# Use the optim() function to find the MLEs of alpha and beta
mle <- optim(par=c(1.4,2),fn = negll, data = X)$par
alpha.hat <- mle[1]
beta.hat <- mle[2]
```

The values of the MLEs are $\hat{\alpha} = 4.989399$ and $\hat{\beta} = 0.1780829$. The MoMs estimators are given by

$$\bar{\alpha} = (m_1')^2/(m_2' - (m_1')^2)$$
$$\bar{\beta} = m_2' - (m_1')^2/m_1'.$$

The R code

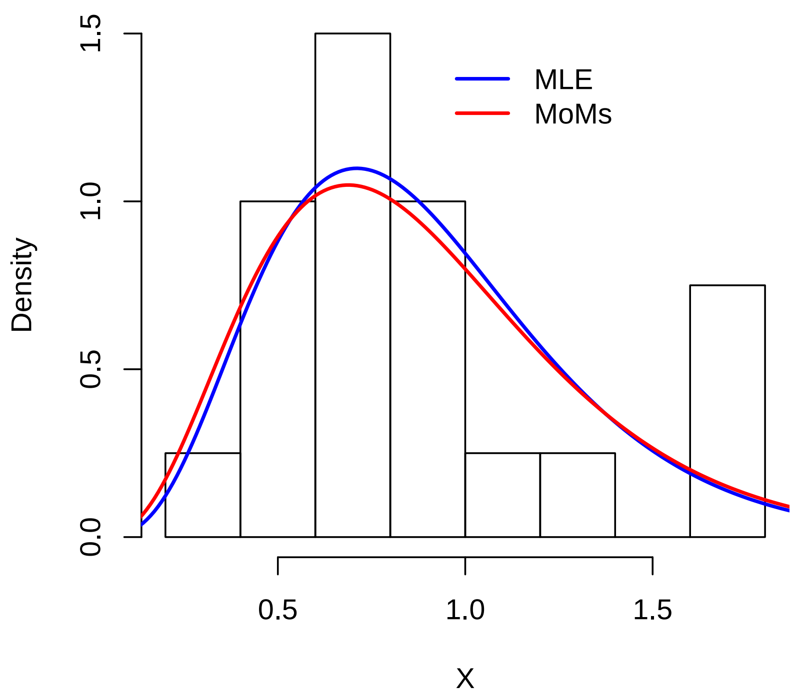
```
m1 <- mean(X)
m2 <- mean(X^2)
alpha.bar <- m1^2 / (m2 - m1^2)
beta.bar <- (m2 - m1^2) / m1
```

gives that the values of the MoMs estimators are $\bar{\alpha} = 4.434695$ and $\bar{\beta} = 0.200352$.

(b) Make a histogram of the realizations of X and overlay the pdfs of the Gamma distributions under the MoMs estimators and the MLEs of (α, β). Refer to the *Lec 11 R code*.

The following code produces the plot:

```
x.seq <- seq(0,3,length=300)
hist(X,freq=FALSE,ylim=c(0,1.5),main="")
lines(dgamma(x.seq,shape=alpha.hat,scale=beta.hat)~x.seq,col="blue",lwd=2)
lines(dgamma(x.seq,shape=alpha.bar,scale=beta.bar)~x.seq,col="red",lwd=2)
x.pos <- grconvertX(.5,from="nfc",to="user")
y.pos <- grconvertY(.8,from="nfc",to="user")
legend(x = x.pos, y = y.pos, legend = c("MLE","MoMs") ,
       col = c("blue","red"), lty=c(1,1), lwd=c(2,2), bty="n")
```



(c) Comment on the difference between the MoMs estimator and the MLE.

The MoMs estimators and the MLEs were quite similar for these data.

(d) Find the MLE for $P(X < 1)$.

To find the MLE for $P(X < 1)$, we compute the cdf of the Gamma$(\hat{\alpha}, \hat{\beta})$ distribuiton at 1, where $\hat{\alpha}$ and $\hat{\beta}$ are the MLEs of α and β. We get

`pgamma(1,shape=alpha.hat,scale=beta.hat) = 0.6619232`.

5. Suppose you observe the following independent realizations of a random variable X:

```
8.33  8.20 11.35  3.05 12.62  9.20 15.68 12.52  4.98 10.44
9.21  8.49  5.14  3.55  5.78  3.97 15.75  7.92  5.02 6.75
```

Assume that X follows a distribution with the density

$$f_X(x; a, b) = \frac{a}{b}\left(\frac{x}{b}\right)^{a-1} \exp\left[-\left(\frac{x}{b}\right)^{a}\right] \mathbb{1}(x > 0).$$

(a) Write down the likelihood function for a and b based on $X_1, \dots, X_n$.

We have

$$\begin{aligned}\mathcal{L}(a, b; X_1, \dots, X_n) &= \prod_{i=1}^{n} \frac{a}{b}\left(\frac{X_i}{b}\right)^{a-1} \exp\left[-\left(\frac{X_i}{b}\right)^{a}\right] \\ &= \frac{a^n}{b^{na}}\left(\prod_{i=1}^{n} X_i\right)^{a-1} \exp\left[-\frac{1}{b^a}\sum_{i=1}^{n} X_i^a\right].\end{aligned}$$

(b) Write down the log-likelihood function for a and b based on $X_1, \dots, X_n$.

We have

$$\ell(a, b; X_1, \dots, X_n) = n\log a - an\log b + (a-1)\sum_{i=1}^{n}\log X_i - \frac{1}{b^a}\sum_{i=1}^{n} X_i^a.$$

(c) Find the MLEs of (a, b). Use R as in the *Lec 11 R code* to find the MLEs.

The following code computes the MLEs based on the data:

```
X <- c(8.33,  8.20, 11.35,  3.05, 12.62,  9.20,
       15.68, 12.52,  4.98, 10.44, 9.21,  8.49,
       5.14,  3.55,  5.78,  3.97, 15.75, 7.92,
       5.02, 6.75)
n <- length(X)

negll <- function(x,data)
{
  a <- x[1]
  b <- x[2]
```

```
  ll <- n * log(a) - n*a*log(b) + (a - 1) * sum(log(X)) - (1/b^a) * sum(X^a)
  negll <- -ll
  return(negll)
}

# Use the optim() function to find the MLEs of alpha and beta
mle <- optim(par=c(1,2),fn = negll, data = X)$par
a.hat <- mle[1]
b.hat <- mle[2]
```

We get the values $\hat{a} = 2.465044$ and $\hat{b} = 9.50044$.

(d) Find the MLE of b^2.

The MLE of b^2 is simply $\hat{b}^2 = (9.50044)^2 = 90.25836$.

Optional (do not turn in) problems for additional study from Wackerly, Mendenhall, Scheaffer, 7th Ed.:

- 9.74, 9.75, 9.78
- 9.80, 9.81, 9.82, 9.92