1)
Show that the MLE of u satisfies

$$\cos(\hat{u}) \sum_{i=1}^{n} (D_i) - \sin(\hat{u}) \sum_{i=1}^{n} \cos(D_i) = 0$$

Von Mises distribution on density $[0, 2\pi)$ is:

$$f(\theta; k, u) = \frac{1}{2\pi I_0(k)} \exp(k\cos(\theta - u))$$

$0 \leq u < 2\pi$, $k \geq 0$, $I_0(k)$ is a 0^{th} order modified Bessel function of the first kind

$$I_0(k) = \frac{1}{2\pi} \int_0^{2\pi} \exp[k\cos(\theta)]\, d\theta$$

MLE

likelihood

$$\prod_{i=1}^{n} f(\theta; k, u)$$

note: $D_i = \theta_i$

$$= \prod_{i=1}^{n} \frac{1}{2\pi I_0(k)} e^{(k\cos(D_i - u))}$$

[expanding the $\prod_{i=1}^{n}$]

$$= \left(\frac{1}{2\pi I_0(k)}\right)^n e^{(\sum_{i=1}^{n} k\cos(D_i - u))}$$

A3 Q1 a)

$$= \frac{1}{(2\pi I_0(k))^n} e^{\left(\sum_{i=1}^{n} k\cos(D_i - \mu)\right)}$$

$$l(\mu) = \ln L(\mu)$$

$$l(\mu) = -n\ln(2\pi I_0(k)) + \sum_{i=1}^{n} k\cos(D_i - \mu)$$

taking the first derivitive we get

$$l'(\mu) = k\sum_{i=1}^{n} \sin(D_i - \mu)$$

using hint we know that $\sin(x-y)$

$= \sin(x)\cos(y) - \cos(x)\sin(y)$

thus

$$l'(\mu) = k\sum_{i=1}^{n} (\sin(D_i)\cos(\mu) - \cos(D_i)\sin(\mu))$$

$$= k\left(\cos(\mu)\sum_{i=1}^{n}\sin(D_i) - \sin(\mu)\sum_{i=1}^{n}\cos(D_i)\right)$$

Find MLE

$l'(\mu) = 0$

$$0 = k\left(\cos(\mu)\sum_{i=1}^{n}\sin(D_i) - \sin(\mu)\sum_{i=1}^{n}\cos(D_i)\right)$$

$$0 = \cos(\mu)\sum_{i=1}^{n}\sin(D_i) - \sin(\mu)\sum_{i=1}^{n}\cos(D_i)$$

$$\frac{\sin(\mu)}{\cos(\mu)} = \frac{\sum_{i=1}^{n} \sin(D_i)}{\sum_{i=1}^{n} \cos(D_i)}$$

$$\tan(\mu) = \frac{\sum_{i=1}^{n} \sin(D_i)}{\sum_{i=1}^{n} \cos(D_i)}$$

$$\hat{\mu} = \tan^{-1}\left(\frac{\sum_{i=1}^{n} \sin(D_i)}{\sum_{i=1}^{n} \cos(D_i)}\right)$$

Verify maximization

$$l''(\mu) = -k \sum_{i=1}^{n} \cos(D_i - \mu) < 0$$

Using identity $\cos(x-y) = \cos(x)\cos(y) + (\sin(x)\sin(y))$

$$-k \sum_{i=1}^{n} (\cos(D_i)\cos(\mu) + \sin(D_i)\sin(\mu)) < 0$$

$$\cos(\mu) \sum_{i=1}^{n} \cos(D_i) + \sin(\mu) \sum_{i=1}^{n} \sin(D_i) > 0$$

Both $\cos(\mu) \sum_{i=1}^{n} \cos(D_i)$ and $\sin(\mu) \sum_{i=1}^{n} \sin(D_i)$ will always be > 0 due to bounds of μ and D_i so $\hat{\mu}$ maximizes.

A3 Q1 B pg 1/2

Prior density on (k, μ):

$$\pi(k,\mu) = \frac{\lambda}{2\pi} \exp(-\lambda k) \quad k > 0 \text{ and}$$

$$0 \leq \mu < 2\pi)$$

Show that posterior density of k is

$$\pi(k \mid d_1, \dots, d_n) = c(d_1, \dots, d_n) \frac{\exp(-\lambda k) I_0(rk)}{(I_0(k))^n}$$

where, $r = \left\{ \left(\sum_{i=1}^{n} \cos(d_i) \right)^2 + \left(\sum_{i=1}^{n} \sin(d_i) \right)^2 \right\}^{1/2}$

$$\pi(k, \mu \mid d_1, \dots, d_n) = c(d_1, \dots, d_n)\, \pi(k,\mu)\, L(k,\mu).$$

$$= c(d_1, \dots, d_n) \frac{\lambda}{2\pi} e^{(-\lambda k)} \left(\frac{1}{(2\pi I_0(k))^n} \right) e^{\left(k \sum_{i=1}^{n} \cos(d_i - \mu)\right)}$$

find $\pi(k \mid d_1, \dots, d_n)$

$$= \int_0^{2\pi} c(d_1, \dots, d_n) \frac{\lambda}{2\pi} e^{(-\lambda k)} \frac{1}{(2\pi)^n} \frac{1}{(I_0 k)^n} e^{\left(\mu \sum_{i=1}^{n} \cos(d_i - \mu)\right)} \, d\mu$$

$$= \int_0^{2\pi} c(d_1, \dots, d_n) \frac{\lambda}{(2\pi)^{n+1}} \frac{1}{(I_0(k)^n)} e^{-\lambda k + \mu \sum_{i=1}^{n} \cos(d_i - \mu))} \, d\nu$$

A3 Q1 b pg 2/2

$$= c(d_1,\dots,d_n)\frac{1}{(I_0(k))^n}\frac{\lambda}{(2\pi)^{n+1}}\int_0^{2\pi} e^{-\lambda k + kr\cos(\theta-u))}\,du$$

$$= c(d_1,\dots,d_n)\frac{1}{(I_0(k))^n}\frac{\lambda}{(2\pi)^{n+1}} e^{-\lambda k}\int_0^{2\pi} e^{kr\cos(\theta-u))}\,du$$

use change of variable

$u' = \theta - u$ / since $\frac{du'}{du} = -1$ we know

$du = -du'$ new bounds $[\theta, \theta - 2\pi]$

$$= c(d_1,\dots,d_n)\frac{1}{(I_0(k))^n}\frac{\lambda}{(2\pi)^{n+1}} e^{-\lambda k}\int_{\theta-2\pi}^{\theta} e^{(kr\cos(u'))}\,du'$$

which is proportional to

$$\propto c(d_1,\dots,d_n)(I_0(k))^{-n} e^{-\lambda k} I_0(rk)$$

finally since $\cos(u')$ and as a result

$e^{\cos(w)}$ is a periodic function (period $= 2\pi$)

we can conclude

$$\pi(k \mid d_1,\dots,d_n) = c(d_1,\dots,d_n)\frac{e^{(-\lambda k)} I_0(rk)}{(I_0(k))^n}$$

q1

q1 c

```r
windData <- scan("wind.txt") # load data
k_val <- 1:1000 / 1000 * (2 * pi)
windRadData <- windData * (pi/180) # Transform into radians

#find k_hat
k_hat <- sqrt(sum(cos(windRadData))^2 + sum(sin(windRadData))^2) / 149 *
  (2 - sqrt(sum(cos(windRadData))^2 + sum(sin(windRadData))^2)^2 / 149^2) /
  (1 - sqrt(sum(cos(windRadData))^2 + sum(sin(windRadData))^2)^2/149^2)

# find MLE
MU <- atan2(sum(sin(windRadData)), sum(cos(windRadData)))

#log prior
find_log_prior <- function(k_val) {
  return (log(1 / (2*pi) * exp(-1 * k_val))) # lambda = 1
}

#log log_likelihood
find_log_likelihood <- function(k_val, x_val) {
  return(log(besselI(sqrt(sum(cos(windRadData))^2 + sum(sin(windRadData))^2)*k_val,0)
             / (2*pi)^(n-1) / (besselI(k_val,0))^149))
}
```

q1 d

```r
x <- runif(200,0,2*pi)
k_val <- 1:1000 / 1000 * (2 * pi)
Probabilities <- NULL



find_f <- function(k_val) {
  return (exp(-k_val) * besselI(sqrt(sum(cos(x))^2 + sum(sin(x))^2)*k_val,0)
          / (2*pi*besselI(k_val,0))^149)
}

find_g <- exp(-max(log(find_f(k_val))))

find_g_post <- function(k_val) {
  find_g*find_f(k_val)
}


for (index in 1:9 / 10) {
  probability <- index*(2*pi)^(-149) / (index*(2*pi)^(-149)+(1-index)*
                                          (sum((c(1/2,rep(1,9998),1/2)) *
                                                 (find_g_post(k_val))) / 1000)/find_g)
```

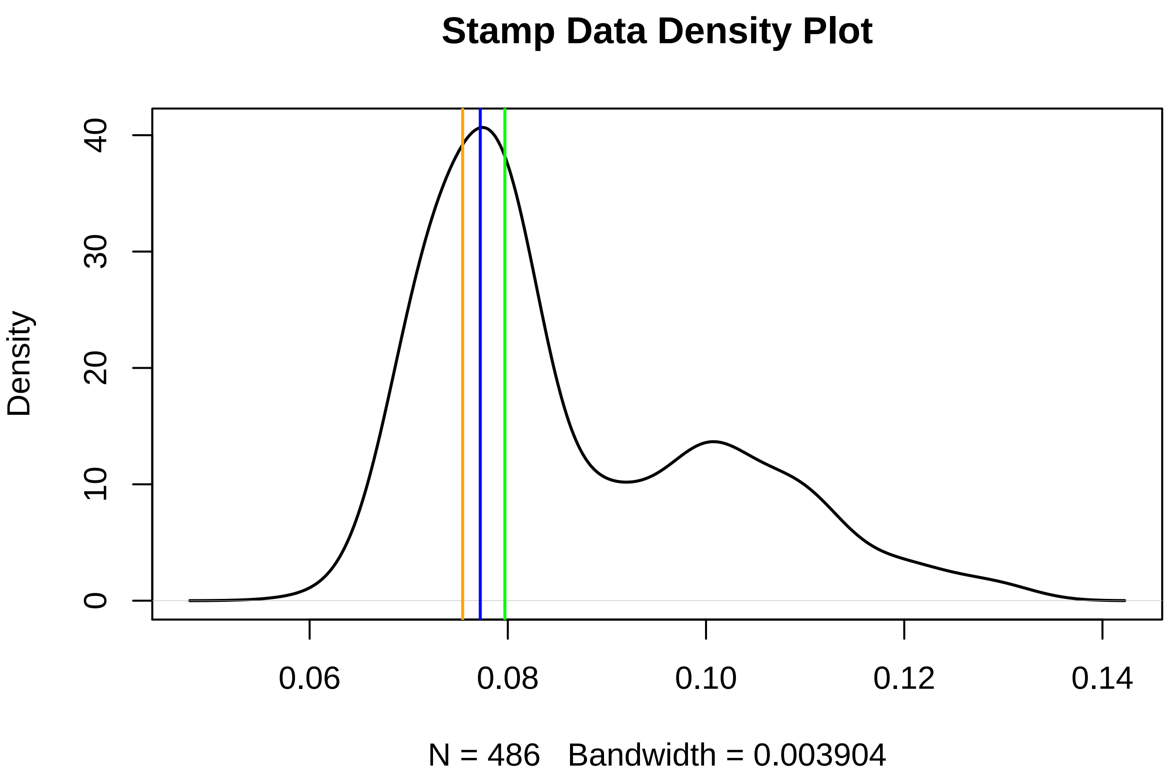
```
  Probabilities <- c(Probabilities, probability)
}
Probabilities
```

```
## [1] 0.3066565 0.4987832 0.6304457 0.7263063 0.7992201 0.8565458 0.9027995
## [8] 0.9409064 0.9728447
```

question 2

2a

```
venter <- function(x, tau=1/2) { # venter function provided
   x <- sort(x)
   n <- length(x)
   m <- ceiling(tau*n)
   x1 <- x[1:(n-m+1)]
   x2 <- x[m:n]
   j <- c(1:(n-m+1))
   len <- x2-x1
   k <- min(j[len==min(len)])
   (x[k]+x[k+m-1])/2
}

stampdata <- scan("stamp.txt") # loading in stamp data
plot(density(stampdata), lwd = 1.5,  main = "Stamp Data Density Plot")+
abline(v = venter(stampdata, 0.625), lwd = 1.5, col= "blue") +
abline(v = venter(stampdata, 0.675), lwd = 1.5, col= "green")+
abline(v = venter(stampdata), lwd = 1.5, col= "orange")
```




```
## integer(0)
```

As we can see the density plot has one global maximum which reaches its peak around the blue verticle line.Thus we can say that when tau is around 0.625 the estimate makes sense.

2b

```r
calc_mse <- function(sample_size, tau, alpha){

  simulation <- NULL

  for (index in 1:10000){ #10000 specified in question

    x <- rgamma(sample_size, alpha)
    simulation[index] <- venter(x, tau)

  }

  mean_squre_error <- mean((simulation-(alpha-1))^2)
  return (mean_squre_error)

}
```

simulations

```r
# sample size: 100 tau:0.1 alpha:2
calc_mse(100, 0.1, 2)
```

```
## [1] 0.2687202
```

```r
# sample size: 1000 tau:0.1 alpha:2
calc_mse(1000,0.1, 2)
```

```
## [1] 0.07069708
```

```r
# sample size: 100 tau:0.5 alpha:2
calc_mse(100,0.5, 2)
```

```
## [1] 0.1220437
```

```r
# sample size: 1000 tau:0.5 alpha:2
calc_mse(1000,0.5, 2)
```

```
## [1] 0.04919306
```

```r
# sample size: 100 tau:0.1 alpha:10
calc_mse(100,0.1, 10)
```

```
## [1] 1.774241
```

```r
# sample size: 1000 tau:0.1 alpha:10
calc_mse(1000,0.1, 10)
```

```
## [1] 0.5495351
```

```r
# sample size: 100 tau:0.5 alpha:10
calc_mse(100, 0.5, 10)
```

```
## [1] 0.7003582
```

```r
# sample size: 1000 tau:0.5 alpha:10
calc_mse(1000,0.5, 10)
```

```
## [1] 0.1634849
```

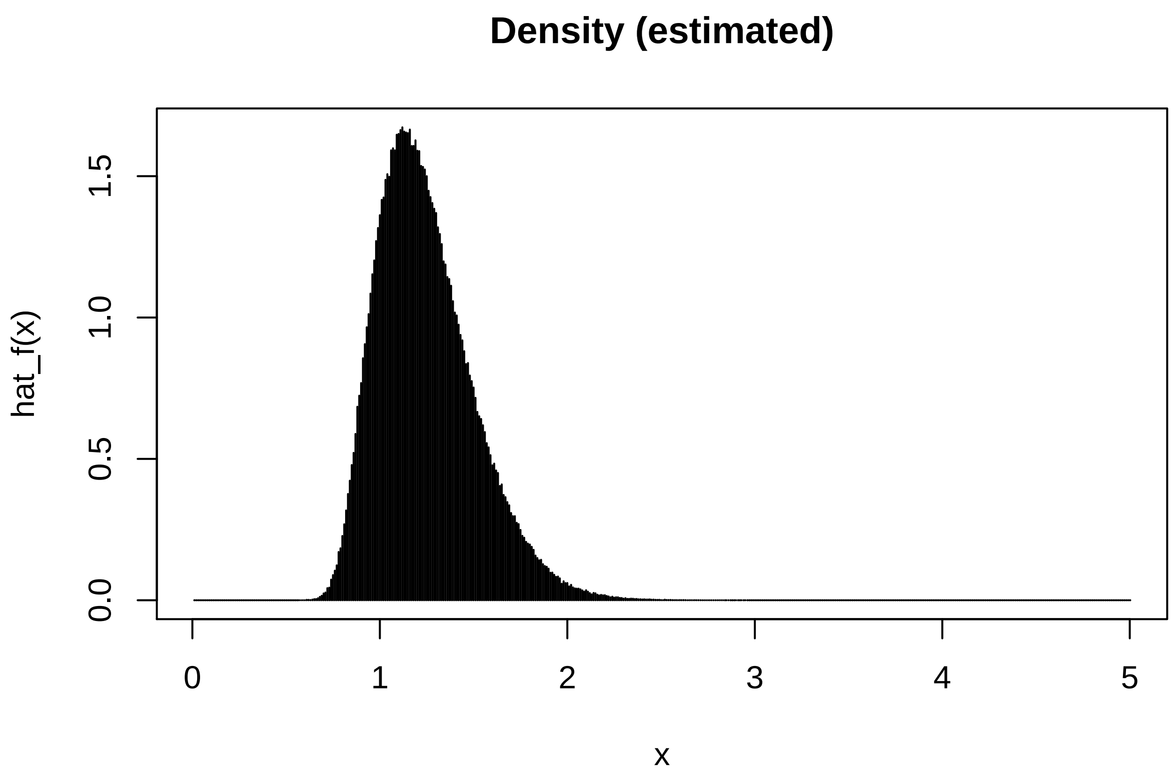
For both alpha values 2 and 8 we estimated the mse of the venter estimator for different tau values (.5 and

.1) and different sample size values(1000 and 100). From these 8 simulation, we can see that for both alpha values 2 and 8 that the mse values were lowest when the sample size was 1000 and the value of tau is 0.5. So we can say that the venter estimator where tau is 0.5 and sample size is 1000 is best when comparing mse.

2c

Before continuing it is important to know that each part of of the summation is a Gamma density function which we can use the dgamma function for. This includes the shape parameter, scale parameter and rate parameter.

```
fcn_fx <- function(alpha, tau, sample_size, x){
  summ <- NULL

  for (i in 1:10000){

    rgamma_val <- rgamma(sample_size, alpha)
    mu_val <- venter(rgamma_val, tau)
    T_val <- sum(rgamma_val)
    summ <- c(summ, dgamma(x, sample_size*alpha,  T_val/mu_val))
  }
  return (mean(summ))

}




x <- c(1:500)/100
values <- NULL
for (i in x){
values <- c(values, fcn_fx(2,0.5,100,i))
}
plot(x, values, type = "h",
main = "Density (estimated)", ylab = "hat_f(x)")
```


Density (estimated)
hat_f(x)
0.0
0.5
1.0
1.5
0
1
2
3
4
5
x