Ke Deng
1004957488
3/26/2021

Q1 (a)

(a) Show that the MLE of μ satisfies

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

and give an explicit formula for $\hat{\mu}$ in terms of $\sum_{i=1}^{n}\sin(D_i)$ and $\sum_{i=1}^{n}\cos(D_i)$. (Hint: Use the formula $\sin(x-y) = \sin(x)\cos(y) - \cos(x)\sin(y)$. Also verify that your estimator actually maximizes the likelihood function.)

$$f(\theta;\kappa,\mu) = \frac{1}{2\pi I_0(\kappa)}\exp(\kappa\cos(\theta-\mu)) \qquad I_0(\kappa) = \frac{1}{2\pi}\int_0^{2\pi}\exp[\kappa\cos(\theta)]\,d\theta.$$

$f(\theta;k,\mu) = \frac{1}{2\pi I_0(k)} e^{k\cos(\theta-\mu)}$ $\quad 0 \le \mu < 2\pi, \; k \ge 0$

$L(\mu) = \left(\frac{1}{2\pi I_0(k)}\right)^n e^{k\sum_{i=1}^{n}\cos(D_i-\mu)}$ $\qquad \cos(\hat{\mu})\sum_{i=1}^{n}\sin(D_i) - \sin(\hat{\mu})\sum_{i=1}^{n}\cos(D_i) = 0$

$\ln L(\mu) = l(\mu) = -n\ln(2\pi I_0(k)) + k\sum_{i=1}^{n}\cos(D_i-\mu) = 0$

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

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

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

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

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

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

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

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

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

$l''(\hat{\mu}) = -k\sum_{i=1}^{n}\cos(D_i-\hat{\mu}) < 0$ $\quad \Rightarrow$ therefore $\hat{\mu} = \arctan\left(\frac{\sum_{i=1}^{n}\sin(D_i)}{\sum_{i=1}^{n}\cos(D_i)}\right)$ maximizes

Q1(b) (b) Suppose that we put the following prior density on (κ, μ):

$$\pi(\kappa, \mu) = \frac{\lambda}{2\pi}\exp(-\lambda\kappa) \quad \text{for } \kappa > 0 \text{ and } 0 \le \mu < 2\pi.$$

Show that the posterior (marginal) density of κ is

$$\pi(\kappa|d_1, \cdots, d_n) = c(d_1, \cdots, d_n)\frac{\exp(-\lambda\kappa)I_0(r\kappa)}{[I_0(\kappa)]^n}$$

where

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

(Hint: To get the marginal posterior density of κ, you need to integrate the joint posterior over μ (from 0 to 2π). The trick is to write

$$\sum_{i=1}^n \cos(d_i - \mu) = r\cos(\theta - \mu)$$

for some θ.)

$\pi(\kappa,\mu) = \frac{\lambda}{2\pi} e^{-\lambda\kappa}$ for $\kappa > 0$, $0 \le \mu \le 2\pi$ WTS $= \pi(\kappa|d_1 \cdots d_n) = c(d_1 \cdots d_n)\frac{e^{-\lambda\kappa} I_0(r\kappa)}{(I_0(\kappa))^n}$

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

WTS $\sum_{i=1}^n \cos(d_i - \mu) = r\cos(\theta - \mu)$

$= r(\cos\theta\cos\mu + \sin\theta\sin\mu)$

$= \cos(\mu)(r\cos(\theta)) + \sin(\mu)\, r\sin(\theta)$

$\Updownarrow$

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

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

$r\cos\theta = \sum_{i=1}^n \cos(d_i)$

$r\sin\theta = \sum_{i=1}^n \sin(d_i)$

$(r\cos\theta)^2 + (r\sin\theta)^2 = r^2(\underbrace{\cos^2\theta + \sin^2\theta}_{1})$

$= r^2$

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

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

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

given prior

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

$= \frac{\lambda}{2\pi}\left(\frac{1}{2\pi I_0(\kappa)}\right)^n e^{-\lambda\kappa + \kappa\sum_{i=1}^n \cos(D_i - \mu)}$ $\rightarrow r\cos(\theta - \mu)$

$= \frac{\lambda}{2\pi}\left(\frac{1}{2\pi I_0(\kappa)^n}\right) e^{-\kappa(\lambda - r\cos(\theta - \mu))}$ θ from Q1(a) $= \arctan\left(\frac{\sum_{i=1}^n \sin(d_i)}{\sum_{i=1}^n \cos(d_i)}\right)$

$\pi(\kappa|d_1 \cdots d_n) \propto \int_0^{2\pi} \frac{\lambda}{2\pi}\left(\frac{1}{2\pi I_0(\kappa)^n}\right) e^{-\kappa(\lambda - r\cos(\theta - \mu))}\, d\mu$

$= \frac{\lambda}{2\pi}\left(\frac{1}{2\pi I_0(\kappa)^n}\right) e^{-\lambda\kappa}\int_0^{2\pi} e^{r\cos(\theta - \mu)}\, d\mu$ let $u = \theta - \mu$ $du = -1$

$$\pi(\kappa|d_1, \cdots, d_n) = c(d_1, \cdots, d_n)\frac{\exp(-\lambda\kappa)I_0(r\kappa)}{[I_0(\kappa)]^n}$$

$= -\lambda e^{-\kappa\lambda}\left(\frac{1}{2\pi I_0(\kappa)^n}\right) I_0(r\kappa)$

$= \frac{-\lambda}{(2\pi)^n}\, \frac{e^{-\lambda\kappa} I_0(r\kappa)}{I_0(\kappa)^n}$

$= C(d_1 \cdots d_n)\frac{e^{-\lambda\kappa} I_0(r\kappa)}{I_0(\kappa)^n}$

Q1(c)

Q1 c) #define functions

```
bee <- scan("bees.txt")

logfunction <- function(x, kappa){
  n = length(x)
  s = sqrt(sum(cos(x))^2 + (sum(sin(x)))^2)

  log <- -n*log(2*pi) - n * log(besselI(kappa, 0)) + kappa*sum(cos(x-mean(x)))
  log
}

prenorm <- function(x, kappa, lambda){
  r <- logfunction(x, kappa)
  r <- r - log(lambda) - log(2*pi) - lambda * kappa
  r <- r - max(r)

  result <- exp(r)
  result
}
```

#khat value

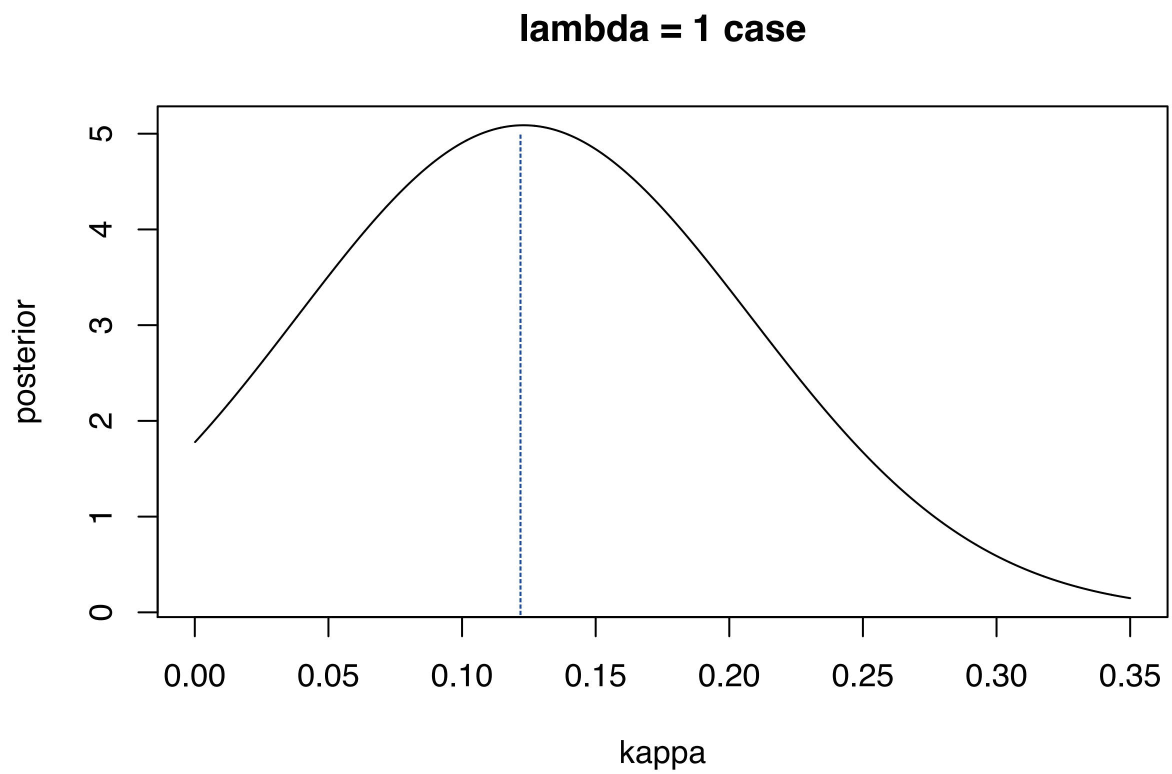
```
x <- bee/180*pi

r <- sqrt(sum(cos(x))^2 + (sum(sin(x)))^2)
n <- length(x)
khat <- (r/n * (2 - r^2/n^2)) / (1 - r^2/n^2)
print(khat)
```

```
## [1] 0.1525488
```

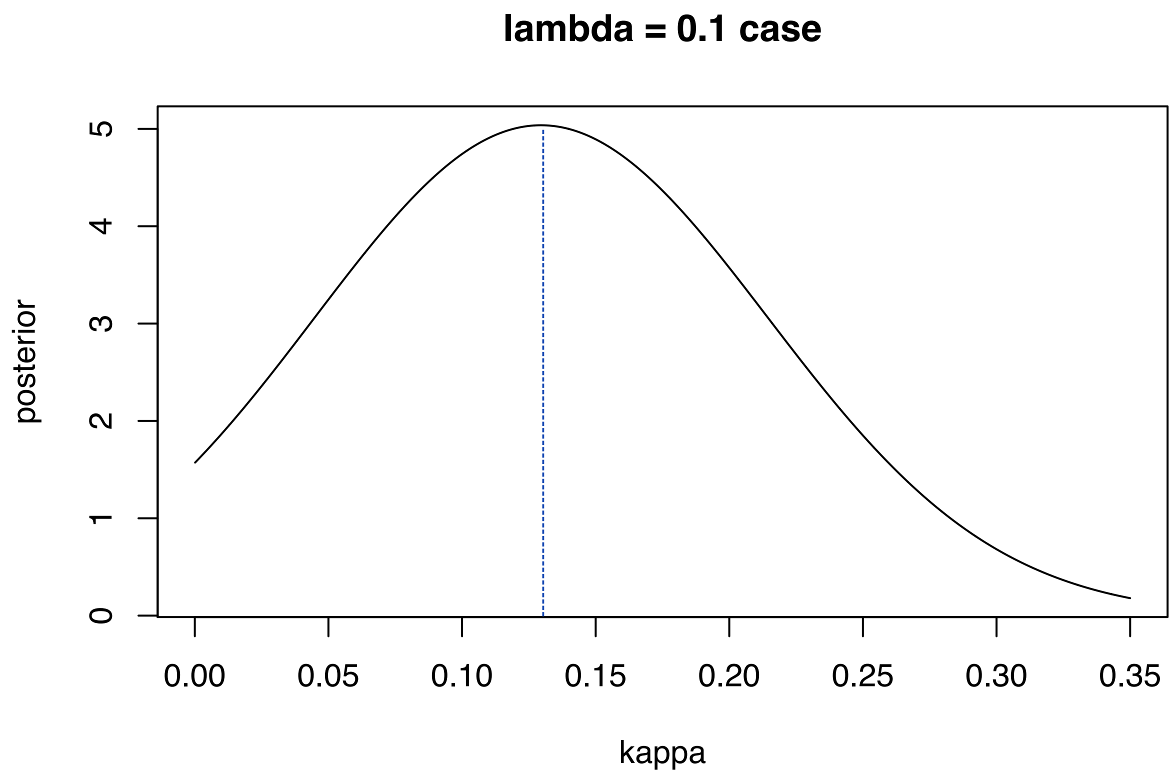
#lambda = 1 case

```
kappa <- c(1:3500)/10000
lambda <- 1
posterior <- prenorm(x,kappa, lambda)
mult <- c(1/2,rep(1, 3499),1/2)
norm <- sum(mult*posterior)/10000
posterior <- posterior/norm
plot(kappa, posterior,type="l",ylab="posterior", main="lambda = 1 case")
```



#lambda = 0.1 case

```
kappa <- c(1:3500)/10000
lambda <- 0.1
posterior <- prenorm(x,kappa, lambda)
mult <- c(1/2,rep(1, 3499),1/2)
norm <- sum(mult*posterior)/10000
posterior <- posterior/norm
plot(kappa, posterior,type="l",ylab="posterior", main="lambda = 0.1 case")
```



From the graph of lambda = 1 and lambda =0.1 above, we can see that both have a max posterior density at around 0.13, which helps eliminate the possibility that k = 0. Also by using the formula, we know that k_hat is approximately 0.15, which also matches the graph we have here.

$$\int_0^\infty \int_0^{2\pi} L(k,\mu)\,\pi(k,\mu)\,d\mu\,dk$$

$$= \int_0^\infty \frac{e^{-k I_0(rk)}}{(I_0(k))^n}\,dk = \frac{\text{norm}}{e^{-\max\left(\ln\left(r\frac{e^{-k I_0(rk)}}{I_0(k)^n}\right)\right)}} \longrightarrow \text{from Q1 c)}$$

Q1 d)

$$= \frac{\text{norm}}{e^{513.55}}$$

```r
probability <- function(theta){
  (theta * (2*pi)^(-n))/(theta*(2*pi)^(-n)+(1-theta)*(norm/exp(513.55)))
}

theta <- 0.1
print(probability(theta))
```

```
## [1] 0.5503128
```

```r
theta <- 0.2
print(probability(theta))
```

```
## [1] 0.7335804
```

```r
theta <- 0.3
print(probability(theta))
```

```
## [1] 0.8251824
```

```r
theta <- 0.4
print(probability(theta))
```

```
## [1] 0.8801334
```

```r
theta <- 0.5
print(probability(theta))
```

```
## [1] 0.9167632
```

```r
theta <- 0.6
print(probability(theta))
```

```
## [1] 0.9429252
```

```r
theta <- 0.7
print(probability(theta))
```

```
## [1] 0.9625456
```

```r
theta <- 0.8
print(probability(theta))
```

```
## [1] 0.9778052
```

```r
theta <- 0.9
print(probability(theta))
```

```
## [1] 0.9900125
```

Q2 a) (a) Suppose we have data $(x_1, y_1), \cdots, (x_n, y_n)$ and we define

$$z_i = a + bx_i + y_i \quad (i = 1, \cdots, n)$$

for some constants a and b. Suppose that $\widehat{\beta}_0$ and $\widehat{\beta}_1$ minimize

$$\text{median}\{(y_i - \beta_0 - \beta_1 x_i)^2 : i = 1, \cdots, n\}$$

and $\widetilde{\beta}_0$ and $\widetilde{\beta}_1$ minimize

$$\text{median}\{(z_i - \beta_0 - \beta_1 x_i)^2 : i = 1, \cdots, n\}.$$

What is the relationship between $(\widehat{\beta}_0, \widehat{\beta}_1)$ and $(\widetilde{\beta}_0, \widetilde{\beta}_1)$?

$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$

$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$

$z_i = a + bx_i + y_i$

$= a + bx_i + (\beta_0 + \beta_1 x_i)$

$= (a + \beta_0) + (b + \beta_1) x_i$

$\tilde{\beta}_0 = a + \hat{\beta}_0$

$\tilde{\beta}_1 = b + \hat{\beta}_1$

$\underline{(\tilde{\beta}_0, \tilde{\beta}_1) = (a + \hat{\beta}_0,\ b + \hat{\beta}_1)}$ $\square$

Q2(b) (b) Show that if $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ for $i = 1, \cdots, n$ then the bias and variance of the LMS estimators does not depend on β_0 and β_1; in other words, they depend only on $\{x_i\}$ and $\{\varepsilon_i\}$. (Hint: Use the result of part (a).)

from question 2a) $\tilde{\beta}_0 = a + \hat{\beta}_0$

$\tilde{\beta}_1 = b + \hat{\beta}_1$ $\qquad z_i = a + bx_i + y_i$

$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$

$= z_i - a - bx_i$ $\Big\}$ $a = -\beta_0 \quad b = -\beta_1$

$a = \beta_0 = \tilde{\beta}_0 - \hat{\beta}_0 \qquad \beta_0 = -a = \hat{\beta}_0 - \tilde{\beta}_0 = y_i - \varepsilon_i + (\tilde{\beta}_1 - \hat{\beta}_1) x_i$

$\beta_1 = -b = \hat{\beta}_1 - \tilde{\beta}_1 = \dfrac{z_i - \varepsilon_i + (\tilde{\beta}_0 - \hat{\beta}_0)}{x_i}$

therefore a, b do not depend on the true parameter, but on $\{x_i\}$ and $\{\varepsilon_i\}$

Q2 c)

```{r}
#install.package("MASS")
n <- c(50, 100, 500, 1000, 5000)

library(MASS)
a <- function(n){
  rep <- 500
  x <- c(1:n)/n
  beta <- NULL

for (i in 1:rep){
  y = rnorm(n)
  r = lmsreg(y~x)
  beta = c(beta, r$coef[2])
}
  return (var50<- var(beta))
}


var <- NULL
for(i in 1:length(n)){
  var = c(var, a(n[i]))
}

norm_var <- log(var)
n <- log(n)
coef(lm(norm_var~ n))

```

```
 (Intercept)           n
   3.2352166  -0.7503765
```

Q2 d) Cauchy errors

```{r}
n <- c(50, 100, 500, 1000, 5000)
var_cauchy <- NULL
for(i in 1:length(n)){
var_cauchy = c(var_cauchy, a(n[i]))
}
var <- log(var_cauchy)
n <- log(n)
coef(lm(var~n))

```

```
 (Intercept)           n
   3.0362330  -0.7118857
```

the a values here are pretty similar, suggesting both ways of calculating are reliable