stats100b_hw4

Takao

5/2/2022

1

```
knitr::include_graphics("/Users/takaooba/Desktop/Screen Shot 2022-11-28 at 4.40.52 PM.png")
```

1. (edited from Problem 55, page 324 of Rice). The following counts for the progeny of self-fertilized heterozygotes were reported and a genetic model was proposed.

a. Find a formula for the MLE of θ using the notations t_1, t_2, t_3, t_4.
b. Find the Fisher information of this genetic model.
c. What is the sample size reported in the table? Give the asymptotic variance of MLE.
d. Give the estimated standard error of MLE using the notations t_1, t_2, t_3, t_4.
e. Plug in the data value of t_1, t_2, t_3, t_3 and report the approximate 95% confidence interval of θ centered at MLE.

Type	Starchy green	Starchy white	Sugary green	Sugary white
Count	t_1=1997	t_2=906	$t_3 = 904$	$t_4 = 32$
Genetic Model	0.25(2+θ)	0.25(1-θ)	0.25(1-θ)	0.25θ

a

By definition

$$n = t_1 + t_2 + t_3 + t_4$$

To find the joint likelihood function, we multiply by its occurence.

$$L(\theta) = 0.25^n \times (2+\theta)^{t_1} \times (1-\theta)^{t_2} \times (1-\theta)^{t_3} \times \theta^{t_4}$$

Taking the log of the joint likelihood function, we have that

$$l(\theta) = n\log(0.25) + t_1\log(2+\theta) + t_2\log(1-\theta) + t_3\log(1-\theta) + t_4\log(\theta)$$

Taking the derivative of all and setting it equal to zero, we have that

$$\frac{d}{d\theta}l(\theta) = \frac{t_1}{2+\theta} - \frac{t_2+t_3}{1-\theta} + \frac{t_4}{\theta} = 0$$

Simplifying, we have that

$$t_1\theta - t_1\theta^2 - 2(t_2+t_3)\theta - (t_2+t_3)\theta^2 + 2t_4 - t_4\theta - t_4\theta^2 = 0$$

Further,

$$n\theta^2 - (t_1 - 2t_2 - 2t_3 - t_4)\theta - 2t_4 = 0$$

Using quadratic formula, we have that

$$\hat{\theta}_{MLE} = \frac{t_1 - 2t_2 - 2t_3 - t_4 \pm \sqrt{(t_1 - 2t_2 - 2t_3 - t_4)^2 + 8(t_1+t_2+t_3+t_4)t_4}}{2(t_1+t_2+t_3+t_4)}$$

b

Taking log of all the terms, we have

$$\log(0.25) + \log(2+\theta),\ \log(0.25) + \log(1-\theta),\ \log(0.25) + \log(1-\theta),\ \log(0.25) + \log(\theta)$$

Taking the derivative, we have that

$$\frac{1}{2+\theta},\ -\frac{1}{1-\theta},\ -\frac{1}{1-\theta},\ \frac{1}{\theta}$$

By definition, we take the square of each component

$$\frac{1}{(2+\theta)^2}, \frac{1}{(1-\theta)^2}, \frac{1}{(1-\theta)^2}, \frac{1}{\theta^2}$$

Taking the expectation of the squares, we have

$$I(\theta) = E[S(\theta)^2] = \frac{1}{(2+\theta)^2} \times 0.25(2+\theta) + \frac{1}{(1-\theta)^2} \times 0.25(1-\theta) + \frac{1}{(1-\theta)^2} \times 0.25(1-\theta) + \frac{1}{\theta^2} \times 0.25\theta$$

Through simplification, we have that

$$I(\theta) = 0.25[\frac{1}{2+\theta} + \frac{2}{1-\theta} + \frac{1}{\theta}]$$

c

```
t_1 <- 1997
t_2 <- 906
t_3 <- 904
t_4 <- 32
n <- t_1 + t_2 + t_3 + t_4
n
```

```
## [1] 3839
```

Through plugging in values in function, we can find the variance

$$Var(\theta_{MLE}) = \frac{1}{nI(\theta_{MLE})} = \frac{1}{0.25n} \times [\frac{1}{2+\theta_{MLE}} + \frac{2}{1-\theta_{MLE}} + \frac{1}{\theta_{MLE}}]^{-1}$$

d

The estimated standard error is taken simply through taking the square root, thus

$$\sqrt{\frac{1}{0.25n} \times [\frac{1}{2+\theta_{MLE}} + \frac{2}{1-\theta_{MLE}} + \frac{1}{\theta_{MLE}}]^{-1}}$$

where the theta value is given through part A,

$$\hat{\theta}_{MLE} = \frac{t_1 - 2t_2 - 2t_3 - t_4 \pm \sqrt{(t_1 - 2t_2 - 2t_3 - t_4)^2 + 8(t_1 + t_2 + t_3 + t_4)t_4}}{2(t_1 + t_2 + t_3 + t_4)}$$

e

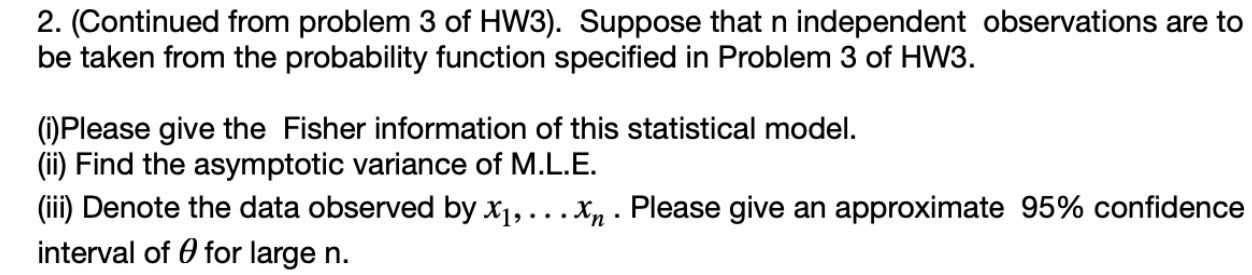
To find the confidence interval, we have

```
part <- t_1-2*t_2-2*t_3-t_4
theta_MLE <- (part+sqrt(part^2 + 8*n*t_4)) /(2*n)
I_theta <- 0.25*(1/(2+theta_MLE) + 2/(1-theta_MLE) + 1/theta_MLE)
asy_var <- 1/(n*I_theta)
cat("The 95% Confidence Interval is " , theta_MLE - qnorm(0.975)*sqrt(asy_var), theta_MLE+qnorm(0.975)*s
```

```
## The 95% Confidence Interval is  0.0242692 0.0471554
```

2

```
knitr::include_graphics("/Users/takaooba/Desktop/Screen Shot 2022-11-28 at 4.40.56 PM.png")
```



This is a discrete case ### i Taking the log of each component, we have that

$$\log(2) - \log(3) + \log(\theta), \log(1) - \log(3) + \log(\theta), log(1-\theta)$$

We can take the derivative of this and get

$$\frac{1}{\theta}, \frac{1}{\theta}, -\frac{1}{1-\theta}$$

Taking the square of each component, we have that

$$\frac{1}{\theta^2}, \frac{1}{\theta^2}, \frac{1}{(1-\theta)^2}$$

To find the fisher information, we will take the expected value of this squared values

$$I(\theta) = E[S(\theta)^2] = \frac{1}{\theta^2} \times \frac{2}{3}\theta + \frac{1}{\theta^2} \times \frac{1}{3}\theta + \frac{1}{(1-\theta)^2} \times (1-\theta)$$

Through simplification, we have that

$$I(\theta) = \frac{1}{\theta} + \frac{1}{1-\theta}$$

ii

The variance can be found by

$$Var(\theta_{MLE}) = (\frac{n}{\theta_{MLE}} + \frac{n}{1-\theta_{MLE}})^{-1}$$

iii

Constructing the confidence interval will be given through

$$[\theta_{MLE} - Z_{0.975} \times \sqrt{(\frac{n}{\theta_{MLE}} + \frac{n}{1-\theta_{MLE}})^{-1}}, \theta_{MLE} + Z_{0.975} \times \sqrt{(\frac{n}{\theta_{MLE}} + \frac{n}{1-\theta_{MLE}})^{-1}}]$$

3

```
knitr::include_graphics("/Users/takaooba/Desktop/Screen Shot 2022-11-28 at 4.41.02 PM.png")
```

3. Find the Kullback–Leibler "divergence $D_{KL}(P||Q)$ of P from Q when P and Q are defined according to the following table of probability mass function .

$x =$	$P(X = x \mid \theta = 1)$	$P(X = x \mid \theta = 2)$	$P(X = x \mid \theta = 3)$
0	0.5	0.3	0.2
1	0.2	0.3	0.2
3	0.1	0.1	0.3
5	0.2	0.3	0.3

(i) P= the distribution defined by $\theta = 1$, Q= the distribution defined by $\theta = 2$

(ii) P= the distribution defined by $\theta = 2$, Q= the distribution defined by $\theta = 1$

(iii) P= the distribution defined by $\theta = 1$, Q= the distribution defined by $\theta = 3$

(iv) From (i) and (iii), verify that if the true distribution of a random variable X is given by $\theta = 1$, then $max_{\theta=1,2,3}\ E(logp(X \mid \theta))$ is achieved at $\theta = 1$

i

Plugging in the numbers we have that

$$D_{KL}(P||Q) = \sum_x P(x) \log(\frac{P(x)}{Q(x)}) = 0.5 \log(\frac{0.5}{0.3}) + 0.2 \log(\frac{0.2}{0.3}) + 0.1 \log(\frac{0.1}{0.1}) + 0.2 \log(\frac{0.2}{0/3})$$

```
0.5*log(0.5/0.3) + 0.2*log(0.2 / 0.3) + 0.1*log(0.1/0.1) + 0.2*log(0.2/0.3)
```

```
## [1] 0.09322677
```

ii

This is basically the same, just different numbers

$$D_{KL}(P||Q) = \sum_x P(x) \log(\frac{P(x)}{Q(x)}) = 0.3 \log(\frac{0.3}{0.5}) + 0.3 \log(\frac{0.3}{0.2}) + 0.1 \log(\frac{0.1}{0.1}) + 0.3 \log(\frac{0.3}{0.2})$$

```
0.3*log(0.3/0.5) + 0.3*log(0.3/0.2) + 0.1*log(0.1/0.1) + 0.3*log(0.3/0.2)
```

```
## [1] 0.09003138
```

iii

Plugging in new numbers, we have

$$D_{KL}(P||Q) = \sum_x P(x) \log(\frac{P(x)}{Q(x)}) = 0.5 \log(\frac{0.5}{0.2}) + 0.2 \log(\frac{0.2}{0.2}) + 0.1 \log(\frac{0.1}{0.3}) + 0.2 \log(\frac{0.2}{0.3})$$

```
0.5*log(0.5/0.2) + 0.2*log(0.2/0.2) + 0.1*log(0.1/0.3) + 0.2*log(0.2/0.3)
```

```
## [1] 0.2671911
```

iv

The expectation of the log of the probability is given through

$$E(\log[P(X|\theta)]) = \sum_x P(x|\theta = 1) \log[P(x|\hat{\theta})]$$

Plug in the numbers for the thetas

For theta = 1

```
0.5*log(0.5) + 0.2*log(0.2) + 0.1*log(0.1) + 0.2*log(0.2)
```

```
## [1] -1.220607
```

For theta = 2

```
0.5*log(0.3) + 0.2*log(0.3) + 0.1*log(0.1) + 0.2*log(0.3)
```

```
## [1] -1.313834
```

For theta = 3

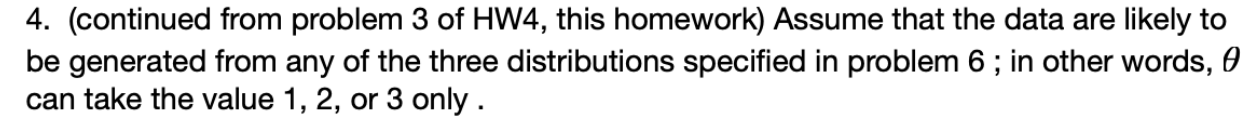
```
0.5*log(0.2) + 0.2*log(0.2) + 0.1*log(0.3) + 0.2*log(0.3)
```

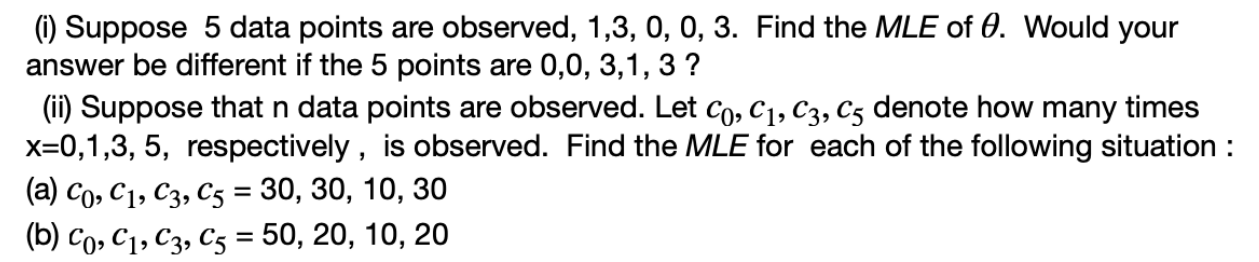
```
## [1] -1.487798
```

Compare the results and we notice that the maximum is achieved at theta = 1.

4

```
knitr::include_graphics("/Users/takaooba/Desktop/Screen Shot 2022-11-28 at 4.41.07 PM.png")
```





i

Plugging in the given data points, we have that

```
data <- c(2,1,2,0)
prob_table <- matrix(c(0.5, 0.3, 0.2,
                       0.2, 0.3, 0.2,
                       0.1, 0.1, 0.3,
                       0.2, 0.3, 0.3),nrow=4,ncol=3,byrow = T)
```

```r
ll <- function(data){
  log_likelihood <- rep(0,3)
  for(i in 1:3){
    holder <- 0
    for(j in 1:4){
      holder <- holder + data[j]*log(prob_table[j,i])
    }
    log_likelihood[i] <- holder
  }
  return(log_likelihood)
}

cat("Theta_MLE = ", which.max(ll(data)))
```

```
## Theta_MLE =  3
```

The answer will not be different if the 5 points are the given one.

ii

a From the given data, we have that

```r
data <- c(30, 30, 10, 30)
cat("Theta_MLE = ", which.max(ll(data)))
```

```
## Theta_MLE =  2
```

```r
data <- c(50, 20, 10, 20)
cat("Theta_MLE =", which.max(ll(data)) )
```

b

```
## Theta_MLE = 1
```

5

```r
knitr::include_graphics("/Users/takaooba/Desktop/Screen Shot 2022-11-28 at 4.41.12 PM.png")
```

5. (continued from the geometric distribution problem 1 and problem 7 of HW3.) A sample of size n=100 was obtained. The frequencies of x=1, 2, etc in the sample are shown in the table below. Find a confidence interval of θ using the following instruction.

(i) By the original bootstrap method. Use 1000 bootstrap runs. Please also show a histogram of the 5000 values of $\hat{\theta}^*_{mle}$ you generate during the simulation.

(ii) By the bootstrap-t method. Use 1000 bootstrap runs. Please also show a histogram of the 1000 values of Δ^* you generate during the simulation

x=	1	2	3	4	5	6	7	8	9	10
Counts	32	19	16	16	6	1	5	2	2	1

%%The dataset was generated by Allen Kei :

2 2 8 1 4 1 2 5 5 1 2 10 1 7 2 9 2 1 1 3 1 6 1 7 3 4 2 1 1
3 2 1 4 7 3 3 3 4 1 7 1 4 2 3 1 1 3 3 4 1 3 3 1 2 1 1 3 1
1 3 5 4 4 2 4 4 4 3 2 2 1 7 4 1 2 1 2 2 3 1 3 4 9 2 1 4 1
1 8 1 1 2 1 5 4 5 4 2 1 5

%%

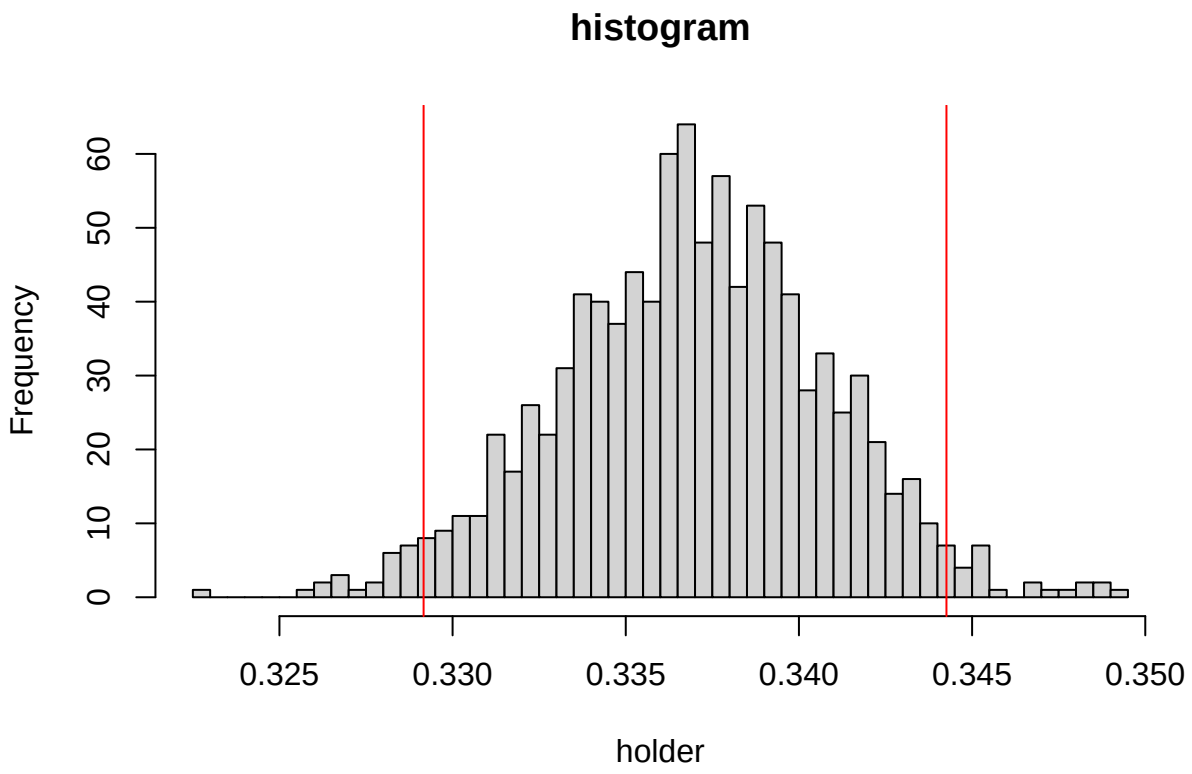
i

```
set.seed(1)
data <- rgeom(100,0.3) + 1
Bootstrap_length <- 1000

MLE <- function(data){
  return(length(data)/sum(data))
}

asymptotic_variance <- function(data){
  n <- length(data)
  p_MLE <- MLE(data)
  asy_var <- 1/(n/p_MLE^2 + n/(p_MLE*(1-p_MLE)))
  return(asy_var)
}

Bootstrap <- function(data, Bootstrap_length){
  p_MLE <- MLE(data)
  holder <- numeric(Bootstrap_length)
  e.s.e <- sqrt(asymptotic_variance(data))

  for(iter in 1:Bootstrap_length){
    sampling_data <- rgeom(5000, p_MLE) + 1
    holder[iter] <- MLE(sampling_data)
  }
  hist(holder, main = "histogram", breaks = 40)
  abline(v = quantile(holder, 0.025), col = "red");abline(v = quantile(holder, 0.975), col = "red");
  delta_under_bar <- quantile(holder, 0.025) - p_MLE
```

```
  delta_over_bar <- quantile(holder, 0.975) - p_MLE

  CI_lower <- p_MLE - delta_over_bar
  CI_upper <- p_MLE - delta_under_bar

  cat("95% CI (Bootstrap) is (", CI_lower, ", ", CI_upper, ")\n")
}

studentized_Bootstrap <- function(data, Bootstrap_length){
  p_MLE <- MLE(data)
  e.s.e <- sqrt(asymptotic_variance(data))
  holder <- numeric(Bootstrap_length)

  for(iter in 1:Bootstrap_length){
    sampling_data <- rgeom(1000, p_MLE) + 1
    p_star <- MLE(sampling_data)
    e.s.e_star <- sqrt(asymptotic_variance(sampling_data))
    holder[iter] <- (p_star - p_MLE) / e.s.e_star
  }
  hist(holder, main = "histogram", breaks = 40);
  abline(v = quantile(holder, 0.025), col = "red"); abline(v = quantile(holder, 0.975), col = "red");

  CI_lower <- p_MLE - quantile(holder, 0.975) * e.s.e
  CI_upper <- p_MLE - quantile(holder, 0.025) * e.s.e
  cat("95% CI (studentized Bootstrap) is (", CI_lower, ", ", CI_upper, ")\n" )
}
```

The red line corresponds to the quantiles

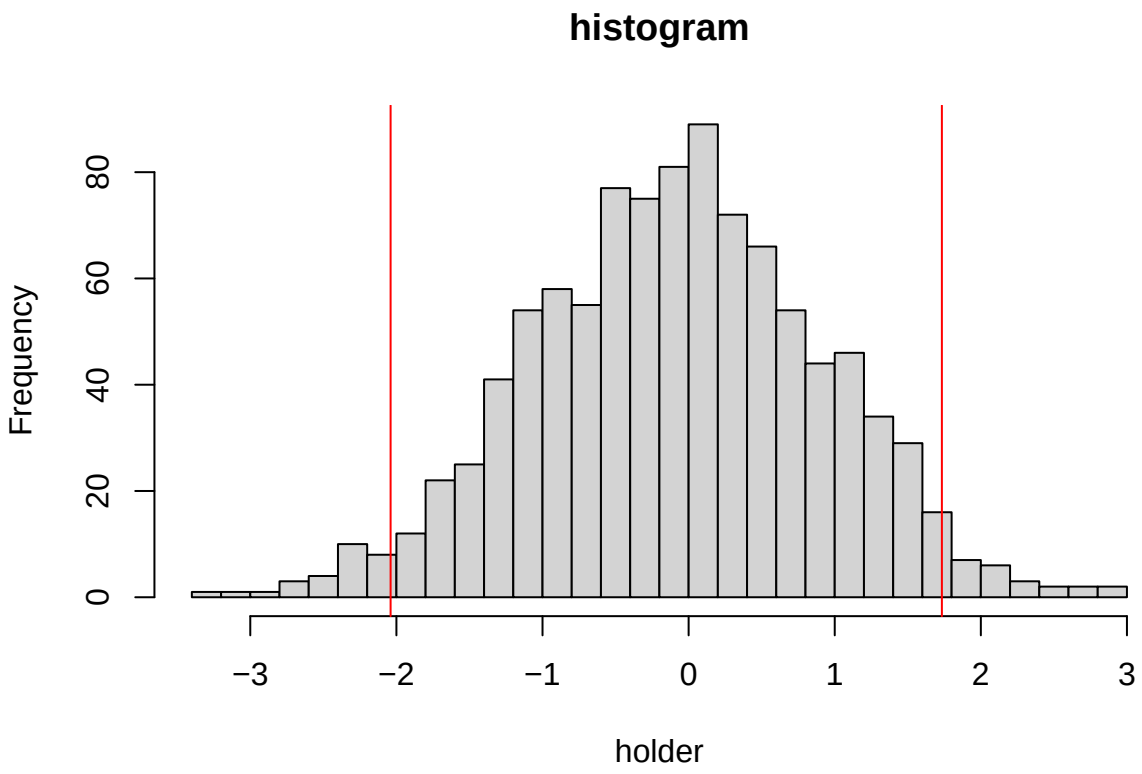
```
Bootstrap(data,Bootstrap_length)
```




```
## 95% CI (Bootstrap) is ( 0.3291411 ,  0.3442378 )
```

ii

```
studentized_Bootstrap(data,Bootstrap_length)
```




```
## 95% CI (studentized Bootstrap) is ( 0.289168 ,  0.3926306 )
```

6

```
knitr::include_graphics("/Users/takaooba/Desktop/Screen Shot 2022-11-28 at 4.41.16 PM.png")
```

6.(continued from problem 1 of HW4, this HW)
Find a 95% confidence interval of θ.

(i) By the original bootstrap method. Use 1000 bootstrap runs. Please also show a histogram of the 5000 values of $\hat{\theta}^*_{mle}$ you generate during the simulation.

(ii) By the bootstrap-t method. Use 1000 bootstrap runs. Please also show a histogram of the 1000 values of Δ^* you generate during the simulation

[hint : how to generate t_1, t_2, t_3, t_4 in each bootstrap run ? Multinomial]

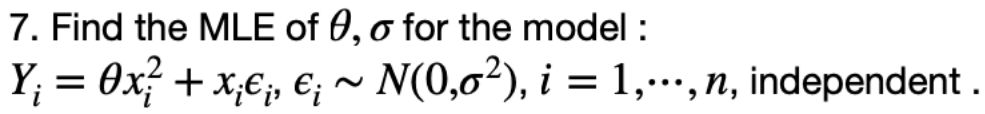
i

```
Bootstrap_length <- 1000
data <- c(1997, 906, 904, 32)

MLE <- function(data){
  n <- sum(data)
  part <- data[1] - 2*data[2] - 2*data[3] - data[4]
  theta_MLE <- (part + sqrt(part^2 + 8*n*data[4])/(2*n))
  return(theta_MLE)
}

asymptotic_variance <- function(data){
  n <- sum(data)
  theta_MLE <- MLE(data)
  I_theta <- 0.25*(1/(2 + theta_MLE) + 2/(1-theta_MLE) + 1/theta_MLE)
  asy_var <- 1/(n*I_theta)
  return(asy_var)
}

Bootstrap <- function(data, Bootstrap_length){
  theta_MLE <- MLE(data)
  holder <- numeric(Bootstrap_length)
  e.s.e <- sqrt(asymptotic_variance(data))

  for(iter in 1:Bootstrap_length){
    sampling_data <- c(rmultinom(1,4000, c(0.25*(2+theta_MLE), 0.25*(1-theta_MLE), 0.25*(1-theta_MLE), 0
    holder[iter] <- MLE(sampling_data)
  }

  hist(holder, main = "histogram", breaks = 40);
  abline(v = quantile(holder, 0.025), col = "red"); abline(v = quantile(holder, 0.975), col = "red");

  Delta_star <- holder - theta_MLE
  CI_lower <- theta_MLE - quantile(Delta_star, 0.975)
  CI_upper <- theta_MLE - quantile(Delta_star, 0.025)

  cat("e.s.e = ", e.s.e, "\n")
  cat("theta* 2.5 percentiles = ", quantile(holder, 0.025), ",", "theta*97.5 percentiles = ", quantile(h
  cat("95% CI (Bootstrap) is (", CI_lower, ",", CI_upper, ")\n")
  cat("95% CI (asy Var) is (", theta_MLE - qnorm(0.975)*e.s.e, ", ", theta_MLE + qnorm(0.975)*e.s.e, ")
}

studentized_Bootstrap <- function(data, Bootstrap_length){
  theta_MLE <- MLE(data)
  holder <- numeric(Bootstrap_length)
  e.s.e <- sqrt(asymptotic_variance(data))


  for(iter in 1:Bootstrap_length){
    sampling_data <- c(rmultinom(1,4000, c(0.25*(2+theta_MLE), 0.25*(1-theta_MLE), 0.25*(1-theta_MLE), 0
    theta_star <- MLE(sampling_data)
    e.s.e_star <- sqrt(asymptotic_variance((sampling_data)))
```

```
    holder[iter] <- (theta_star - theta_MLE) / e.s.e_star
  }
  hist(holder, main = "histogram of Delta*", breaks = 40);
  abline(v = quantile(holder, 0.025), col = "red"); abline(v = quantile(holder, 0.975),col = "red");

  CI_lower <- theta_MLE - quantile(holder, 0.975)*e.s.e
  CI_upper <- theta_MLE - quantile(holder, 0.025)*e.s.e

  cat("e.s.e = ", e.s.e, "\n")
  cat("Delta* 2.5 percentiles = ", quantile(holder, 0.025), ",", "Delta* 97.5 percentiles = ", quantile
  cat("95% CI (studentized Bootstrap) is (", CI_lower, ",", CI_upper, ")\n")
  cat("95% CI (asy Var) is (", theta_MLE-qnorm(0.975)*e.s.e, ",", theta_MLE + qnorm(0.975)*e.s.e, ")")
}
```

7

```
knitr::include_graphics("/Users/takaooba/Desktop/Screen Shot 2022-11-28 at 4.41.20 PM.png")
```



By assumption, we have that

$$\epsilon_i \sim N(0, \sigma^2)$$

Transfer Normality from epsilon to Y:

$$Y_i = \theta X_i^2 + X_i \epsilon_i \sim N(X_i \times 0 + \theta X_i^2, X_i^2 \sigma^2) = N(\theta X_i^2, X_i^2 \sigma^2)$$

To make the joint likelihood function, we have

$$L(\theta, \sigma; y) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \times \frac{1}{x_i \sigma} \times \exp[-\frac{1}{2}(\frac{y_i - \theta x_i^2}{x_i \sigma})^2]$$

The joint log-likelihood function is given through:

$$l(\theta, \sigma; y) = \sum_{i=1}^{n} [\log(\frac{1}{\sqrt{2\pi}}) + \log(\frac{1}{x_i \sigma}) - \frac{1}{2}(\frac{y_i - \theta x_i^2}{x_i \sigma})^2]$$

Solving for theta, we have that

$$\frac{d}{d\theta} l(\theta \sigma; y) = -\frac{1}{2\sigma^2} \sum_{i=1}^{n} 2 \frac{1}{x_i^2} (y_i - \theta x_i^2)(-x_i^2) = \frac{1}{\sigma^2} \sum_{i=1}^{n} (y_i - \theta x_i^2) = 0$$

We can further simplify through

$$\sum_{i=1}^{n} y_i - \theta \sum_{i=1}^{n} x_i^2 = 0$$

$$\hat{\theta}_{MLE} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i^2}$$

Solving for the other unknown which is sigma, we have that

$$\frac{d}{d\sigma} l(\theta, \sigma; y) = -\frac{n}{\sigma} - \sum_{i=1}^{n} [\frac{y_i - \theta x_i^2}{x_i \sigma} \times \frac{y_i - \theta x_i^2}{x_i} \times (-\frac{1}{\sigma^2})] = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{n} (\frac{y_i - \theta x_i^2}{x_i})^2 = 0$$

We can simplify through

$$n\sigma^2 = \sum_{i=1}^{n} (\frac{y_i - \theta x_i^2}{x_i})^2$$

Where

$$\hat{\sigma}_{MLE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\frac{y_i - \hat{\theta}_{MLE} \times x_i^2}{x_i})^2}$$

8

```
knitr::include_graphics("/Users/takaooba/Desktop/Screen Shot 2022-11-28 at 4.41.25 PM.png")
```

8. An animal experiment was conducted to compare the efficacy of a new drug.
Let $X_1, \cdots, X_n$ be the response of mice receiving the drug and $Y_1, \cdots, Y_m$ be the response of mice receiving the placebo.
Consider a three-parameter model :

$X_i \sim N(\mu + \tau, \sigma^2)$ and $Y_j \sim N(\mu, \sigma^2)$, all random variables are independent.
Find the MLE for the unknown parameters μ, τ, σ .

Keyword is independent

Constructing the joint likelihood function we have that

$$L(\mu, \sigma, \tau; x, y) = \prod_{i=1}^{2} \frac{1}{\sqrt{2\pi}} \times \frac{1}{\sigma} \times \exp[-\frac{1}{2}(\frac{x_i - \mu - \tau}{\sigma})^2] \times \prod_{j=1}^{m} \frac{1}{\sqrt{2\pi}} \times \frac{1}{\sigma} \times \exp[-\frac{1}{2}(\frac{y_j - \mu}{\sigma})^2]$$

$$L(\mu,\sigma,\tau;x,y) = (\frac{1}{\sqrt{2\pi}})^{n+m} \times (\frac{1}{\sigma})^{n+m} \times \prod_{i=1}^{n} \exp[-\frac{1}{2}(\frac{x_i-\mu-\tau}{\sigma})^2] \times \prod_{j=1}^{m} \exp[-\frac{1}{2}(\frac{y_j-\mu}{\sigma})^2]$$

The joint log-likelihood function is given through

$$l(\mu,\sigma,\tau;x,y) = (n+m)\log(\frac{1}{\sqrt{2\pi}}) + (n+m)\log(\frac{1}{\sigma}) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu-\tau)^2 - \frac{1}{2\sigma^2}\sum_{j=1}^{m}(y_j-\mu)^2$$

The derivative with respect to tau is

$$\frac{d}{d\tau}l(\mu,\sigma,\tau;x,y) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu-\tau) = \frac{1}{\sigma^2}(-n\mu-n\tau+\sum_{i=1}^{n}x_i) = 0$$

$$-n\mu+\sum_{i=1}^{n}x_i = n\tau$$

$$\hat{\tau} = \bar{x}-\mu$$

Derivative with respect to mu

$$\frac{d}{d\mu}l(\mu,\sigma,\tau;x,y) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu-\tau) + \frac{1}{\sigma^2}\sum_{j=1}^{m}(y_j-\mu) = \frac{1}{\sigma^2}(\sum_{i=1}^{n}x_i - n\mu - n\tau + \sum_{j=1}^{m}y_j - m\mu) = 0$$

$$\sum_{i=1}^{n}x_i - n\tau + \sum_{j=1}^{m}y_j = (n+m)\mu$$

Replace tau back into the equation

$$\sum_{i=1}^{n}x_i - n(\bar{x}-\mu) + \sum_{j=1}^{m}y_j = (n+m)\mu$$

$$\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}x_i + n\mu + \sum_{j=1}^{m}y_j = (n+m)\mu$$

$$\sum_{j=1}^{m}y_j = m\mu$$

Estimation of tau and mu in terms of data:

$$\hat{\mu} = \bar{y}$$

$$\hat{\tau} = \bar{x}-\bar{y}$$

Derivative with respect to sigma

$$\frac{d}{d\sigma}l(\mu,\sigma,\tau;x,y) = -(n+m)\frac{1}{\sigma} + 2\frac{1}{\sigma^3}\sum_{i=1}^{n}\frac{1}{2}(x_i-\mu-\tau)^2 + 2\frac{1}{\sigma^3}\sum_{j=1}^{m}\frac{1}{2}(y_j-\mu)^2 = 0$$

$$-(n+m)\sigma^2 + \sum_{i=1}^{n}(x_i-\mu-\tau)^2 + \sum_{j=1}^{m}(y_j-\mu)^2 = 0$$

Estimation of sigma in terms of data:

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\hat{\mu}-\hat{\tau})^2 + \sum_{j=1}^{m}(y_j-\hat{\mu})^2}{n+m}} = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2 + \sum_{j=1}^{m}(y_j-\bar{y})^2}{n+m}}$$

9

```
knitr::include_graphics("/Users/takaooba/Desktop/Screen Shot 2022-11-28 at 4.41.30 PM.png")
```

9. A small survey of 100 families with three members shows that there are 35, 20, 15, 30 families with x persons (x=0, 1, 2, 3, 4 respectively) getting an infectious disease.
A model with two unknown parameters is proposed.

$$P(X=x\,|\,\theta_1,\theta_2) = p(x\,|\,\theta_1,\theta_2),\ x = 0,1,2,3$$
$$p(0\,|\,\theta_1,\theta_2) = P(X=0\,|\,\theta_1,\theta_2) = \theta_1,$$
$$p(1\,|\,\theta_1,\theta_2) = P(X=1\,|\,\theta_1,\theta_2) = \theta_2$$
$$p(2\,|\,\theta_1,\theta_2) = P(X=1\,|\,\theta_1,\theta_2) = 1-2\theta_1-\theta_2,$$
$$p(3\,|\,\theta_1,\theta_2) = P(X=2\,|\,\theta_1,\theta_2) = \theta_1$$

(i)Find MLE.
(ii) Plot the log joint likelihood surface.

i

Take the log:

$$\log(\theta_1), \log(\theta_2), \log(1-2\theta_1-\theta_2), \log(\theta_1)$$

The joint log-likelihood is

$$l(\theta_1,\theta_2) = t_1\log(\theta_1) + t_2\log(\theta_2) + t_3\log(1-2\theta_1-\theta_2) + t_4\log(\theta_1)$$

Taking the partial derivative with respect to the two thetas

$$\frac{d}{d\theta_1}l(\theta_1,\theta_2) = \frac{t_1+t_4}{\theta_1} - 2\frac{t_3}{1-2\theta_1-\theta_2} = 0$$

$$\frac{d}{d\theta_2} l(\theta_1, \theta_2) = \frac{t_2}{\theta_2} - \frac{t_3}{1 - 2\theta_1 - \theta_2} = 0$$

$$\theta_2 = \frac{2t_2}{t_1 + t_4}\theta_1$$

$$\hat{\theta}_1 = \frac{t_1 + t_4}{2(t_1 + t_4 + t_2 + t_3)}$$

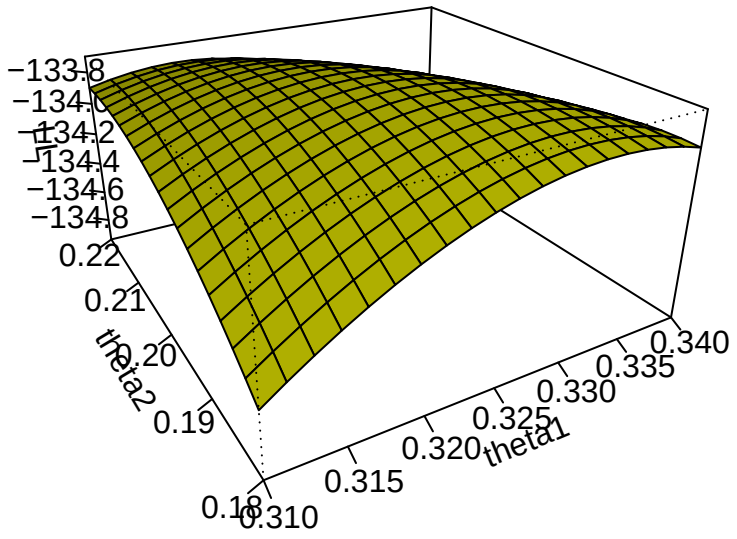
$$\hat{\theta}_2 = \frac{2t_2}{t_1 + t_4} \times \frac{t_1 + t_4}{2(t_1 + t_4 + t_2 + t_3)} = \frac{t_2}{t_1 + t_4 + t_2 + t_3}$$

ii

```
t_1 <- 35
t_2 <- 20
t_3 <- 15
t_4 <- 30
grid_size = 20
theta1 <- seq(0.31, 0.34, length.out = grid_size)
theta2 <- seq(0.18, 0.22, length.out = grid_size)

LL <- matrix(0, grid_size, grid_size)

ll <- function(theta1, theta2){
  t_1*log(theta1) + t_2*log(theta2) + t_3*log(1-2*theta1-theta2) + t_4*log(theta1)
}

for(i in 1:grid_size){
  for(j in 1:grid_size){
    LL[i, j] <- ll(theta1[i], theta2[j])
  }
}

ind <- which(LL == max(LL), arr.ind = TRUE); cat("MLE", theta1[ind[1]], theta2[ind[2]])
```

```
## MLE 0.3242105 0.2010526
```

```
persp(theta1,theta2, LL, theta = -30, phi = 25, shade = 0.75, col = "yellow", expand = 0.5, r = 2, lthe
```



10

```
knitr::include_graphics("/Users/takaooba/Desktop/Screen Shot 2022-11-28 at 4.41.32 PM.png")
```

10. Continue from problem 9.
(i) Find the Fisher information matrix
(ii) Construct 95% confidence intervals for θ_1 and θ_2.

i

Taking log:

$$\log(\theta_1), \log(\theta_2), \log(1-2\theta_1-\theta_2), \log(\theta_1)$$

The second derivative with respect to the theta1 is

$$-\frac{1}{\theta_1^2}, 0, -4(1-2\theta_1-\theta_2)^{-2}, -\frac{1}{\theta_1^2}$$

The second derivative with respect to theta2

$$0, -\frac{1}{\theta_2^2}, -(1-2\theta_1-\theta_2)^{-2}, 0$$

The partial derivative with respect to theta1 followed by partial derivative with respect to theta2

$$0, 0, -2(1-2\theta_1-\theta_2)^{-2}, 0$$

Finding the fisher information matrix:

$$-E[H(\theta_1,\theta_2)] = \frac{1}{\theta_1^2}\times\theta_1 + 4(1-2\theta_1-\theta_2)^{-2}\times(1-2\theta_1-\theta_2) + \frac{1}{\theta_1^2}\times\theta_1 = \frac{4}{1-2\theta_1-\theta_2} + \frac{2}{\theta_1}$$

The 2,2 entry of fisher information matrix is given through

$$-E[H(\theta_1,\theta_2)] = \frac{1}{\theta_2^2}\times\theta_2 + (1-2\theta_1-\theta_2)^{-2}\times(1-2\theta_1-\theta_2) = \frac{1}{1-2\theta_1-\theta_2} + \frac{1}{\theta_2}$$

1,2 and 2,1 entry of the fisher information matrix is given through

$$-E[H(\theta_1,\theta_2)] = 2(1-2\theta_1-\theta_2)^{-2}\times(1-2\theta_1-\theta_2) = \frac{2}{1-2\theta_1-\theta_2}$$

ii

```
t_1 <- 35
t_2 <- 20
t_3 <- 15
t_4 <- 30
n <- t_1 + t_2 + t_3 + t_4


theta1 <- (t_1 + t_4) / (2*(t_1 + t_4 + t_2 + t_3))
theta2 <- t_2/(t_1 + t_4 + t_2 + t_3)

Fisher_matrix <- matrix(0,2,2)
Fisher_matrix[1,1] <- 4/(1-2*theta1 - theta2) + 2/theta1
Fisher_matrix[2,2] <-1/(1-2*theta1 - theta2) + 1/theta2
Fisher_matrix[1,2] <- 2/(1-2*theta1 - theta2)
Fisher_matrix[2,1] <- 2/(1-2*theta1 - theta2)

asy_var1 <- solve(Fisher_matrix)[1,1] / n
asy_var2 <- solve(Fisher_matrix)[2,2] / n

cat("95% Confidence Interval for theta 1:", theta1 - qnorm(0.975)*sqrt(asy_var1), theta1 + qnorm(0.975)*
```

```
## 95% Confidence Interval for theta 1: 0.2782578 0.3717422
```

```
cat("95% Confidence Interval for theta 2:", theta2 - qnorm(0.975)*sqrt(asy_var1), theta2 + qnorm(0.975)*
```

```
## 95% Confidence Interval for theta 2: 0.1532578 0.2467422
```