STAT 640 Section 1 - Homework 2

Zhaoshan Duan

9/27/2021

Contents

19.

a.

First find the likelihood function and log likelihood functions:

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}}$$

Find log-likelihood function:

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2} \tag{1}$$

$$= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2} \tag{2}$$

$$= -\frac{n}{2}\log(2\pi) - n\log\sigma - \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2} \tag{3}$$

Since μ is known, we take the derivative with respect to σ:

$$\nabla\ell(\sigma) = \frac{\partial\ell(\mu, \sigma^2)}{\partial\sigma} \tag{4}$$

$$= -\frac{n}{\sigma} + \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^3} \tag{5}$$

Set gradient vector to 0 and solve for σ:

$$\nabla\ell(\sigma) = 0 \tag{6}$$

$$-\frac{n}{\sigma} + \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^3} = 0 \tag{7}$$

$$\hat{\sigma}_{mle} = \frac{\sqrt{\sum(x_i-\mu)^2}}{n} \tag{8}$$

b.

Using the same log-likelihood function, we take the derivative with respect to μ:

$$\nabla\ell(\mu) = \frac{\partial\ell(\mu, \sigma^2)}{\partial\mu} \tag{9}$$

$$= \frac{\sum_{i=1}^{n}(x_i-\mu)}{\sigma^2} \tag{10}$$

Set gradient vector to 0 and solve for μ:

$$\nabla\ell(\mu) = 0 \tag{11}$$

$$\frac{\sum_{i=1}^{n}(x_i-\mu)}{\sigma^2} = 0 \tag{12}$$

$$\hat{\mu}_{mle} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x} \tag{13}$$

c.

Since an unbiased estimator that satisfies the Cramer-Rao lower bound is automatically the best unbiased estimator,

Find the Fisher information first:

$$I(\mu) = -E[\frac{\partial^2}{\partial\theta^2} \log f(X_1; \mu)]$$

Find the second derivative: $\theta = \mu$

$$\begin{aligned}
\log f(X_1; \theta)] &= -\frac{1}{2}\log(2\pi\sigma^2) - \frac{(X_1 - \mu)^2}{2\sigma^2} && (14) \\
\frac{\partial}{\partial\theta} \log f(X_1; \theta)] &= 0 - (-1)\frac{2(X_1 - \mu)}{\sigma^2} = \frac{X_1 - \mu}{\sigma^2} && (15) \\
\frac{\partial^2}{\partial\theta^2} \log f(X_1; \theta)] &= -\frac{1}{\sigma^2} && (16) \\
\therefore I(\mu) &= \frac{1}{\sigma^2} && (17)
\end{aligned}$$

The Cramer-Rao lower bound is:

$$\frac{1}{n \cdot I(\mu)} = \frac{\sigma^2}{n}$$

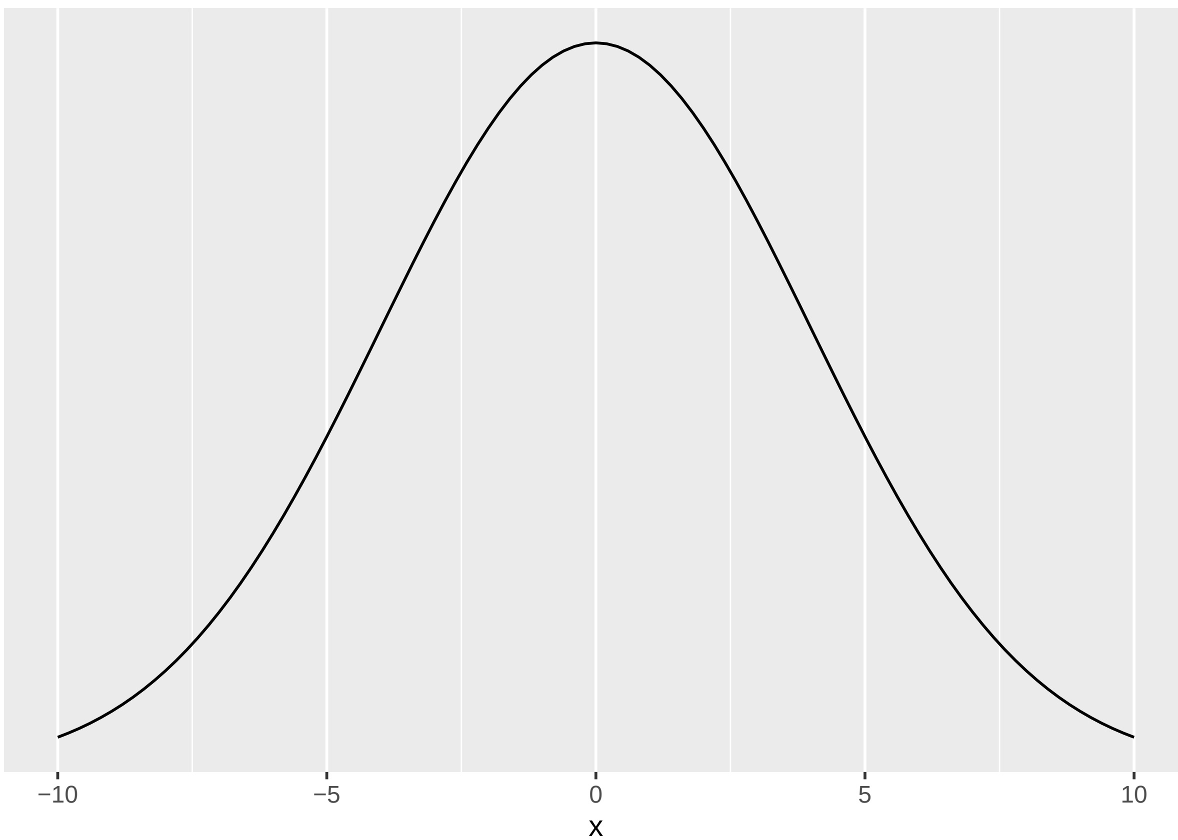
Since variance of $\bar{X}$ is $\frac{\sigma^2}{n}$, MLE estimator of μ is the best unbiased estimator. That is. no other unbiased estimator for μ can have lower variance than MLE estimator.

20.

Get expected value and variance of $\overline{X}$:

- expected value: $E(\overline{X}) = E(\sum_{i=1}^{25} \frac{x_i}{n}) = \frac{1}{n}\sum_{i=1}^{25} x_i = 0$
- variance: $\text{Var}(\overline{X}) = \text{Var}(\frac{1}{25}\sum_{i=1}^{25} X_i) = \frac{1}{25^2}\text{Var}\sum_{i=1}^{25}(X_i) = \frac{1}{25^2} \cdot 25 \cdot 100 = 4$

```
library(ggplot2)
p1 <- ggplot(data = data.frame(x = c(-10, 10)), aes(x)) +
  stat_function(fun = dnorm, n = 101, args = list(mean = 0, sd = 4)) + ylab("") +
  scale_y_continuous(breaks = NULL)
p1
```

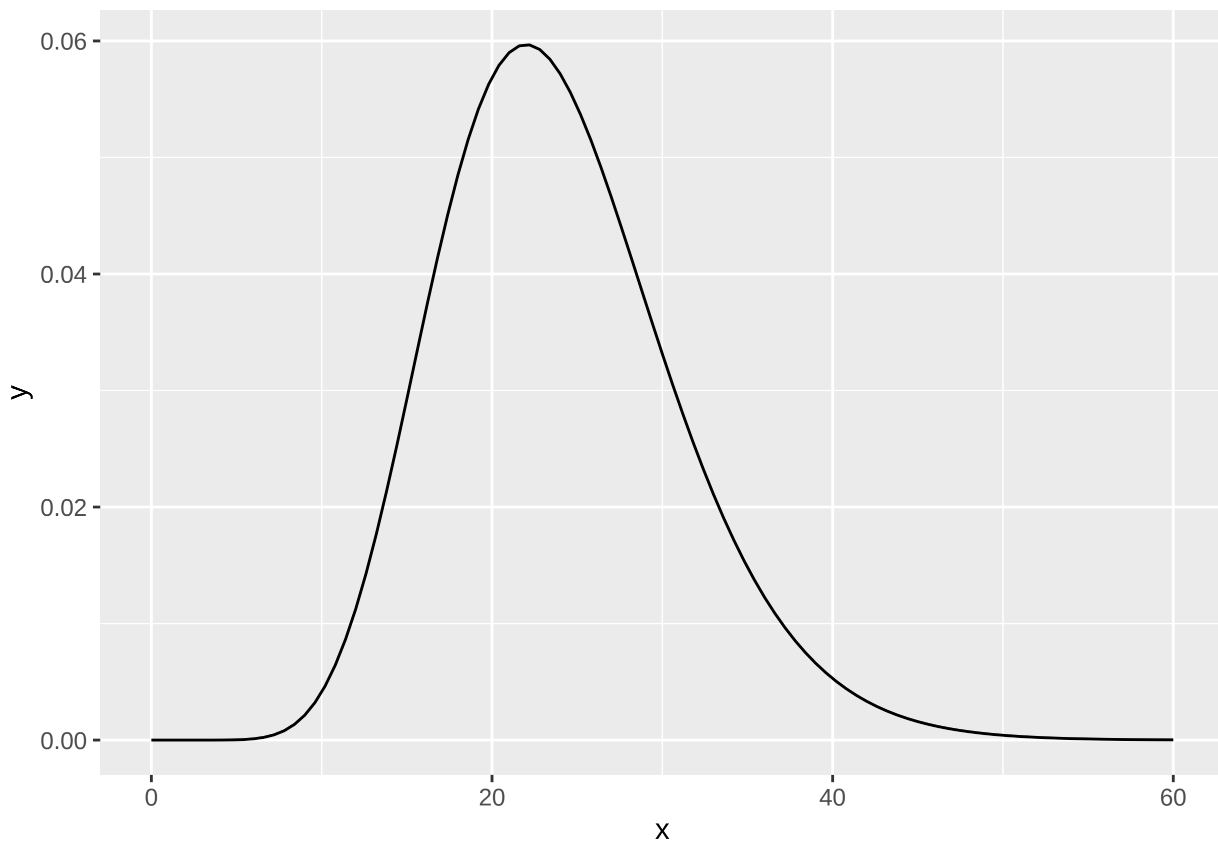


Since X follows a normal distribution, the variance follows chi-square distribution with $n-1$ degrees of freedom.

$$\begin{aligned} \hat{\sigma^2} &\sim \frac{\sigma^2}{n} \cdot \chi^2_{n-1} && (18) \\ &\sim \frac{100}{25} \cdot \chi^2_{24} && (19) \\ &\sim 4 \cdot \chi^2_{24} && (20) \end{aligned}$$

```
library(ggplot2)

ggplot(data.frame(x = c(0, 60)), aes(x = x)) +
     stat_function(fun = dchisq, args = list(df = 24))
```



Incorrect !

50.

a.

Find the first theoretical population moment $k = 1$:

$$\mu_1 = E(X_1) = \int_0^\infty x \frac{x}{\theta^2} e^{-x^2/(2\theta^2)} dx$$

Variable substitution $v = \frac{x^2}{2\theta^2}, dv = \frac{x}{\theta^2} dx, x = \theta\sqrt{2v}$

$$E(X_1) = \int_0^\infty \theta\sqrt{2v} e^{-v} dv = \theta\sqrt{2} \int_0^\infty v^{\frac{1}{2}} e^{-v} dv$$

Using the Gamma trick, $\Gamma(x = \frac{3}{2})$:

$$E(X_1) = \theta\sqrt{2}\Gamma(\frac{3}{2}) = \theta\sqrt{2}\Gamma(\frac{1}{2} + 1) = \theta\frac{\sqrt{2}}{2}\Gamma(\frac{1}{2}) = \theta\frac{\sqrt{2\pi}}{2}$$

Equate to the first sample moment:

$$\theta\frac{\sqrt{2\pi}}{2} = \overline{X} \tag{21}$$

$$\hat{\theta}_{mom} = \overline{X}\frac{\sqrt{2\pi}}{\pi} \tag{22}$$

b.

Find the likelihood function:

$$L(\theta) = \prod_{i=1}^n f(x|\theta) \tag{23}$$

$$= \prod_{i=1}^n \frac{x_i}{\theta^2} e^{-x_i^2/(2\theta^2)} \tag{24}$$

$$= \frac{\prod_{i=1}^n x_i}{\theta^{2n}} e^{\sum_{i=1}^n -x_i^2/2\theta^2} \tag{25}$$

Find the log likelihood function:

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

Take derivative:

$$\nabla\ell(\theta) = -2n\frac{1}{\theta} + \frac{1}{\theta^3}\sum_{i=1}^n x_i^2$$

Set to 0 and solve for θ:

$$\nabla \ell(\theta) = 0 \tag{26}$$
$$-2n\frac{1}{\theta} + \frac{1}{\theta^3}\sum_{i=1}^{n} x_i^2 = 0 \tag{27}$$
$$\frac{1}{\theta^3}\sum_{i=1}^{n} x_i^2 = 2n\frac{1}{\theta} \tag{28}$$
$$\frac{1}{\theta^2} = \frac{2n}{\sum_{i=1}^{n} x_i^2} \tag{29}$$
$$\hat{\theta}_{mle} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{2n}} \tag{30}$$

c.

Find the Fisher information of the sample. Because $X_1, .., X_n$ is an $i.i.d$ sample,

$$I_n(\theta_0) \stackrel{i.i.d}{=} nE\big[\frac{\partial}{\partial\theta}\log(f(X,\theta)\big]^2 \tag{31}$$
$$= -nE\big[\frac{\partial^2}{\partial\theta^2}\log[\frac{x}{\theta^2}e^{-x^2/(2\theta^2)}]\big] \tag{32}$$
$$= -nE\big[\frac{\partial^2}{\partial\theta^2}(\log x - \log\theta^2 - \frac{x^2}{2\theta^2})\big] \tag{33}$$
$$= -nE[\frac{\partial}{\partial\theta}(-\frac{2}{\theta} + x^2\theta^{-3})] \tag{34}$$
$$= -nE[\frac{2}{\theta^2} - 3\frac{x^2}{\theta^4}] \tag{35}$$
$$= -n\frac{2}{\theta^2} + n\frac{3}{\theta^4}E(x^2) \tag{36}$$

Find $E(X^2)$:

$$E(X^2) = \int_0^\infty x^2\frac{x}{\theta^2}e^{-x^2/(2\theta^2)}dx$$

Variable substitution $v = \frac{x^2}{2\theta^2}, dv = \frac{x}{\theta^2}dx, x = \theta\sqrt{2v}$

$$E(X^2) = \int_0^\infty \theta^2 2ve^{-v}dv = \theta^2 2\int_0^\infty ve^{-v}dv$$

Using the Gamma trick, $\Gamma(x = 2)$:

$$E(X^2) = 2\theta^2\Gamma(2) = 2\theta^2$$

Plug $E(X^2)$ back to $I(\theta)$:

$$I_n(\theta_0) = -n\frac{2}{\theta^2} + n\frac{3}{\theta^4}2\theta^2 = \frac{4n}{\theta^2}$$

Therefore, the asymptotic variance of MLE:

$$\text{Var} \approx \frac{1}{I_n(\theta)} \approx \frac{\theta^2}{4n}$$

55.

a.

Get likelihood and log-likelihood function, and take first derivative:

$$
\begin{aligned}
L(\theta) &= \prod f(\theta) && (37)\\
&= [0.25(2+\theta)]^{1997} \cdot [0.25(1-\theta)]^{906} \cdot [0.25(1-\theta)]^{904} \cdot [0.25\theta]^{32} && (38)\\
&= (0.5+0.25\theta)^{1997} \cdot (0.25-0.25\theta)^{906} \cdot (0.25-0.25\theta)^{904} \cdot (0.25\theta)^{32} && (39)\\
&= (0.5+0.25\theta)^{1997} \cdot (0.25-0.25\theta)^{1810} \cdot (0.25\theta)^{32} && (40)\\
\ell(\theta) &= 1997\log\left[0.5+0.25\theta\right)] + 1810\log[0.25-0.25\theta)] + 32\log[0.25\theta] && (41)\\
\nabla\ell(\theta) &= \frac{1997}{2+\theta} - \frac{1810}{1-\theta} + \frac{32}{\theta} = \frac{-3839\theta^2 - 1655\theta + 64}{(2+\theta)(1-\theta)\theta} && (42)
\end{aligned}
$$

Set the gradient to 0:

$$
\begin{aligned}
\nabla\ell(\theta) &= 0 && (43)\\
\frac{-3839\theta^2 - 1655\theta + 64}{(2+\theta)(1-\theta)\theta} &= 0 && (44)\\
-3839\theta^2 - 1655\theta + 64 &= 0 && (45)\\
& && (46)
\end{aligned}
$$

```r
result <- function(a,b,c){
  if(delta(a,b,c) > 0){
        x_1 = (-b+sqrt(delta(a,b,c)))/(2*a)
        x_2 = (-b-sqrt(delta(a,b,c)))/(2*a)
        result = c(x_1,x_2)
  }
  else if(delta(a,b,c) == 0){
        x = -b/(2*a)
  }
  else {"try again."}
}

delta<-function(a,b,c){
      b^2-4*a*c
}

a <- result(-3839, -1655, 64);a
```

```
## [1] -0.4668142  0.0357123
```

From using quadratic formal, and keep the positive root, we get $\hat{\theta}_{mle} \approx 0.0357$.

Next, find the fisher information by taking second derivative of the log likelihood function:

$$
\begin{aligned}
\nabla^2\ell(\theta) &= -\frac{1997}{(2+\theta)^2} - \frac{1810}{(1-\theta)^2} - \frac{32}{\theta^2} && (47)\\
\text{Var}(\hat{\theta}_{mle}) &= \frac{1}{I_n(\theta)} && (48)\\
&= -\frac{1}{E\big[-\frac{1997}{(2+\theta)^2} - \frac{1810}{(1-\theta)^2} - \frac{32}{\theta^2}\big]} && (49)\\
&= \frac{1}{\frac{1997}{(2+\theta)^2} + \frac{1810}{(1-\theta)^2} + \frac{32}{\theta^2}} && (50)
\end{aligned}
$$

```
var <-1/ (((1997/(2+0.0357)^2) + (1810/(1-0.0357)^2) + (32/0.0357^2) )) ;var
```

```
## [1] 3.631547e-05
```

```
sd <- var^{1/2}; sd
```

```
## [1] 0.006026232
```

Plug in the estimated θ, and we get asymptotic variance 0.00602.

b.

With 95% confidence interval, $\alpha = 0.05$, and

Using the asymptotic distribution of MLEs

$$(\hat{\theta_n} - z_{1-\alpha/2}\sqrt{\frac{1}{nI(\theta_0)}}, \hat{\theta_n} + z_{1-\alpha/2}\sqrt{\frac{1}{nI(\theta_0)}}) \tag{51}$$

$$(0.0357 - 1.96\sqrt{3.6315 \times 10^{-5}}, 0.0357 + 1.96\sqrt{3.6315 \times 10^{-5}}) \tag{52}$$

$$(0.0239, 0.0475) \tag{53}$$

```
0.0357 - 1.96 * 0.000036315^(1/2)
```

```
## [1] 0.02388866
```

```
0.0357 + 1.96 * 0.000036315^(1/2)
```

```
## [1] 0.04751134
```

c.

```
N <- 100000 # boostrap samples
thetaMLE <-  0.0357 # initial MLE estiamte
thetas <- array()

for (i in 1:N){
  data<- rmultinom(1, 3839 ,prob = c((2+thetaMLE)*.25, (1-thetaMLE)*.25, (1-thetaMLE)*.25, thetaMLE*.25
  n<- sum(data) #sample size

negative_loglik<- function(x){
  p<- c( (2+x)*.25, (1-x)*.25, (1-x)*.25, x*.25 ) #likelihood
  # log_likelihood
  ell<- -1 * sum(data*log(p))
  return(ell)}

res<- optim(0.5, negative_loglik, hessian = TRUE, method = "Brent",lower=0,upper=1)
thetas[i]<- res$par
}
sd(thetas)
```

```
## [1] 0.005856676
```

d.

```
quantile(thetas, c(0.025,0.975))
```

```
##       2.5%      97.5%
## 0.02464877 0.04760152
```