

Tutorial 9 Solutions

STAT 3013/4027/8027

1. Chapter 8 Question 55:

a. Let's write out the Likelihood. Note that we have

$$\begin{aligned} L(\theta) &= \frac{n!}{x_1! x_2! x_3! x_4!} p_1^{x_1} p_2^{x_2} p_3^{x_3} p_4^{x_4} \\ &\propto p_1^{x_1} p_2^{x_2} p_3^{x_3} p_4^{x_4} \\ &= [0.25(2 + \theta)]^{1997} [0.25(1 - \theta)]^{906} [0.25(1 - \theta)]^{904} [0.25\theta]^{32} \end{aligned}$$

$$\begin{aligned} \ell(\theta) &= 1997 \log[0.25(2 + \theta)] + 906 \log[0.25(1 - \theta)] \\ &\quad + 904 \log[0.25(1 - \theta)] + 32 \log[0.25\theta] \end{aligned}$$

$$\begin{aligned} \frac{d\ell(\theta)}{d\theta} &= \frac{1997 \times 0.25}{0.25(2 + \theta)} - \frac{906 \times 0.25}{0.25(1 - \theta)} \\ &\quad - \frac{904 \times 0.25}{0.25(1 - \theta)} + \frac{32 \times 0.25}{0.25\theta} \\ &\Rightarrow \frac{1997}{(2 + \theta)} - \frac{906 + 904}{(1 - \theta)} + \frac{32}{\theta} = 0 \tag{1} \\ &\Rightarrow 3839\theta^2 + 1655\theta - 64 = 0 \\ &\Rightarrow \dot{\theta} = \frac{-1655 \pm \sqrt{1655^2 - 4(3839)(-64)}}{2(3839)} \\ &\Rightarrow \dot{\theta} = -0.4668 \text{ or } 0.0357 \\ &\Rightarrow \hat{\theta} = 0.0357 \end{aligned}$$

Writing this in terms of the data, from Eqn. 1 we have:

$$\begin{aligned}
\frac{x_1}{(2+\theta)} - \frac{x_2+x_3}{(1-\theta)} + \frac{x_4}{\theta} &= 0 \\
\frac{x_1}{(2+\theta)} + \frac{x_4}{\theta} &= \frac{x_2+x_3}{(1-\theta)} \\
(x_1+x_2+x_3+x_4)\theta^2 + (x_4+2x_2+2x_3-x_1)\theta - 2x_4 &= 0 \\
a\theta^2 + b\theta + (-c) &= 0 \\
\dot{\theta} &= \frac{-b \pm \sqrt{b^2 - 4(a)(-c)}}{2(a)}
\end{aligned}$$

In order to get the asymptotic variance let's take the second derivative of the log likelihood:

$$\frac{d^2\ell(\theta)}{d\theta^2} = -\frac{x_1}{(2+\theta)^2} - \frac{x_2+x_3}{(1-\theta)^2} - \frac{x_4}{\theta^2}$$

Now let's get the Fisher Information:

$$\begin{aligned}
I(\theta) &= -E\left[\frac{d^2\ell(\theta)}{d\theta^2}\right] \\
&= -E\left[-\frac{x_1}{(2+\theta)^2} - \frac{x_2+x_3}{(1-\theta)^2} - \frac{x_4}{\theta^2}\right] \\
&= \frac{E[x_1]}{(2+\theta)^2} + \frac{E[x_2]+E[x_3]}{(1-\theta)^2} + \frac{E[x_4]}{\theta^2} \\
&= \frac{np_1}{(2+\theta)^2} + \frac{np_2+np_3}{(1-\theta)^2} + \frac{np_4}{\theta^2} \\
&= \frac{n[0.25(2+\theta)]}{(2+\theta)^2} + \frac{n[0.25(1-\theta)] + n[0.25(1-\theta)]}{(1-\theta)^2} + \frac{n[0.25\theta]}{\theta^2} \\
&= n0.25\left[\frac{1}{(2+\theta)} + \frac{2}{(1-\theta)} + \frac{1}{\theta}\right]
\end{aligned}$$

So the asymptotic standard error for $\hat{\theta} \approx \sqrt{1/I(\hat{\theta})} = 0.0058$.

b. To create a 95% interval asymptotic interval we have:

$$\begin{aligned} & \hat{\theta} - z_{1-\alpha/2} \sqrt{I(\hat{\theta})^{-1}} \quad , \quad \hat{\theta} + z_{1-\alpha/2} \sqrt{I(\hat{\theta})^{-1}} \\ & 0.0369 - 1.96(0.0058) \quad , \quad 0.0369 + 1.96(0.0058) \\ & 0.0255 \quad , \quad 0.0483 \end{aligned}$$

d. Let's write a function in R to solve a quadratic formula based on the data x_1, x_2, x_3, x_4 , as we will need this when we implement the bootstrap:

```
quad.form <- function(x1, x2, x3, x4){

  a <- x1 + x2 + x3 + x4
  b <- (-x1 + x4 + 2*x2 + 2*x3)
  c <- - 2*x4

  sol1 <- (-b + sqrt(b^2 - 4*a*c))/(2*a)
  sol2 <- (-b - sqrt(b^2 - 4*a*c))/(2*a)

  out <- c(sol1, sol2)
  return(out)
}

sol <- unname(quad.form(1997, 906, 904, 32))
sol
```

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

```
sol[sol>=0 & sol<=1]
```

```
## [1] 0.0357123
```

Let's create a single bootstrap sample:

```
X <- c(rep("starchy green", 1997), rep("starchy white", 906),
       rep("surgary green", 904), rep("surgary white", 32))

table(X)
```

```
## X
```

```
## starchy green starchy white surgary green surgary white
##          1997          906          904          32
```

```
n <- sum(table(X))

boot <- table(X[sample(1:n, n, replace=TRUE)])
x1 <- boot[1]
x2 <- boot[2]
x3 <- boot[3]
x4 <- boot[4]

sol <- unname(quad.form(x1, x2, x3, x4))
sol
```

```
## [1] 0.04250291 -0.47803299
```

```
sol[sol>=0 & sol<=1]
```

```
## [1] 0.04250291
```

Let's repeat this 10,000 times:

```
set.seed(2001)
S <- 10000
out <- rep(0,S)

for(s in 1:S){

boot <- table(X[sample(1:n, n, replace=TRUE)])
x1 <- boot[1]
x2 <- boot[2]
x3 <- boot[3]
x4 <- boot[4]

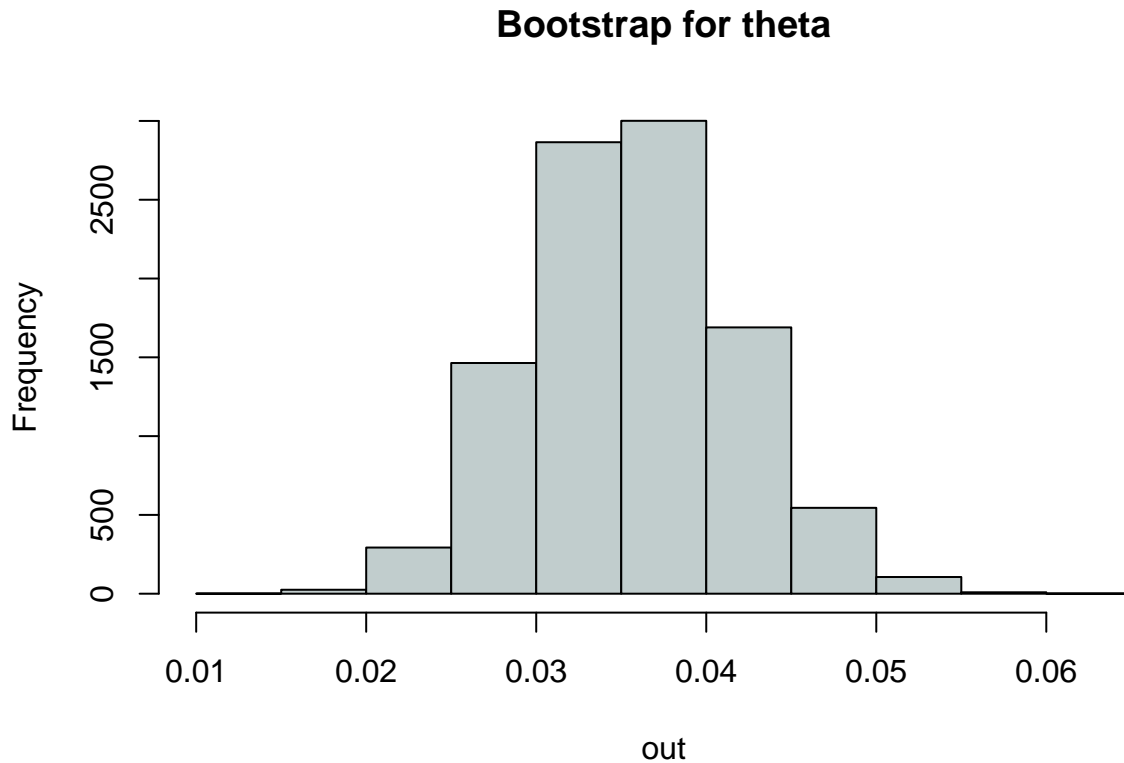
sol <- unname(quad.form(x1, x2, x3, x4))
```

```

out[s] <- sol[sol>=0 & sol<=1]
}

hist(out, col="azure3", main="Bootstrap for theta")

```



```
mean(out)
```

```
## [1] 0.03568094
```

```
sd(out)
```

```
## [1] 0.006032519
```

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

```
##      2.5%      97.5%
```

```
## 0.02438215 0.04793024
```

These are quite close to the asymptotic intervals.

2. Chapter 8 Question 56:

- a. Let's calculate the expected value of the first estimator:

$$\begin{aligned}
E[\tilde{\theta}_1] &= E\left[\frac{4X_1}{n} - 2\right] = \frac{4E[X_1]}{n} - 2 \\
&= \frac{4n[0.25(2 + \theta)]}{n} - 2 \\
&= 2 + \theta - 2 = \theta
\end{aligned}$$

So the estimator is unbiased. Now let's determine its variance:

$$\begin{aligned}
V\left[\frac{4X_1}{n} - 2\right] &= \frac{4^2 V[X_1]}{n^2} \\
&= \frac{4^2}{n^2} np_1(1 - p_1) \\
&= \frac{16}{n} [0.25(2 + \theta)](1 - [0.25(2 + \theta)]) \\
\Rightarrow SE(\tilde{\theta}_1) &= \sqrt{\frac{16}{n} [0.25(2 + \tilde{\theta}_1)](1 - [0.25(2 + \tilde{\theta}_1)])} \\
&= \sqrt{\frac{16}{3839} [0.25(2 + 0.0808)](1 - [0.25(2 + 0.0808)])} \\
&= 0.0323
\end{aligned}$$

Note: $\frac{4X_1}{n} - 2 = \frac{4(1997)}{3839} - 2 = 0.0808$.

b. Let's calculate the expected value of the second estimator:

$$\begin{aligned}
E[\tilde{\theta}_2] &= E\left[\frac{4X_4}{n}\right] = \frac{4E[X_4]}{n} \\
&= \frac{4n[0.25\theta]}{n} \\
&= \theta
\end{aligned}$$

So the estimator is unbiased. Now let's determine its variance:

$$\begin{aligned}
V\left[\frac{4X_4}{n}\right] &= \frac{4^2 V[X_4]}{n^2} \\
&= \frac{4^2}{n^2} np_4(1 - p_4) \\
&= \frac{16}{n} (0.25\theta)(1 - 0.25\theta) \\
\Rightarrow SE(\tilde{\theta}_2) &= \sqrt{\frac{16}{n} (0.25\tilde{\theta}_2)(1 - 0.25\tilde{\theta}_2)} \\
&= \sqrt{\frac{16}{3839} [0.25(0.0333)](1 - [0.25(0.0333)])} \\
&= 0.0059
\end{aligned}$$

Note: $\frac{4X_4}{n} = \frac{4(32)}{3839} = 0.0333$.

- c. We see that the standard error for the second estimator is smaller than the first. Now we need to compare it to the MLE. We already know that asymptotically the variance of the MLE achieves the Cramer-Rao lower bound, so we know that this may not be a fair comparison so let's look at both the bootstrap estimate and the asymptotic estimate:

$$SE(\hat{\theta})_{asym} = 0.0058 < SE(\tilde{\theta}_2) = 0.0059 < SE(\hat{\theta})_{boot} = 0.0060 < SE(\tilde{\theta}_1) = 0.0323.$$

With such little difference, we may wish to get more bootstrap samples just to make sure.

3. **Chapter 8 Question 60:** Here we have $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} f(x|\tau) = \frac{1}{\tau} \exp(-x/\tau)$.

- a. Let's get the likelihood, log likelihood, and solve for the MLE:

$$\begin{aligned}
L(\tau) &= \prod_{i=1}^n \frac{1}{\tau} \exp(-x_i/\tau) \\
&= \frac{1}{\tau^n} \exp\left(-\sum_{i=1}^n x_i/\tau\right) \\
\ell(\tau) &= -n \log(\tau) - \sum_{i=1}^n x_i/\tau \\
\ell'(\tau) &= -\frac{n}{\tau} + \sum_{i=1}^n x_i/\tau^2 = 0 \\
&\Rightarrow \hat{\tau} = \bar{X}
\end{aligned}$$

- b. Now we need to determine the distribution of $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{n} Y$. As we have the sum of independent random variables let's use the **moment generating function**:

$$\begin{aligned}
M_Y(t) &= \left(\frac{1}{1 - \tau t}\right)^n \\
M_{\frac{1}{n}Y}(t) = M_Y\left(\frac{t}{n}\right) &= \left(\frac{1}{1 - \frac{\tau t}{n}}\right)^n
\end{aligned}$$

This means that $\bar{X} \sim \text{gamma}(a = n, b = \tau/n)$. Where $f(x|a, b) = \frac{1}{\Gamma(a)b^a} x^{a-1} \exp(-x/b)$.

- c. Using the distributional properties of the MLE and the Central Limit theorem an approximation to the sampling distribution is:

$$\frac{\hat{\tau} - E[\hat{\tau}]}{SE(\hat{\tau})} \sim \text{normal}(0, 1) = Z$$

$$\frac{\hat{\tau} - ab}{\sqrt{ab^2}} \sim \text{normal}(0, 1) = Z$$

$$\frac{\hat{\tau} - \tau}{\sqrt{\hat{\tau}^2/n}} \sim \text{normal}(0, 1) = Z$$

- d. Unbiasedness and the variance was shown in (c).

- e. Asymptotically we know the MLE achieves the Cramer-Rao lower bound. But we should check the small sample case (i.e. not asymptotic case). Let's get the CRLB:

$$\begin{aligned}
I(\tau) &= -E[\ell''(\tau)] \\
&= -E\left[\frac{n}{\tau^2} - \frac{2n\bar{X}}{\tau^3}\right] \\
&= \frac{2nE[\bar{X}]}{\tau^3} - \frac{n}{\tau^2} \\
&= \frac{2n\tau}{\tau^3} - \frac{n}{\tau^2} \\
&= \frac{2n}{\tau^2} - \frac{n}{\tau^2} \\
&= n/\tau^2 \\
\Rightarrow \text{CRLB} &= I(\tau)^{-1} = \tau^2/n
\end{aligned}$$

We see that the $V(\hat{\tau})$ achieves the CRLB in small samples, so there is not another estimator with a smaller variance. We may have also considered the properties of **exponential families**.

- f. One possibility for an approximate confidence interval is to base it on the CLT:

$$\begin{aligned}
&\hat{\tau} - z_{1-\alpha/2}SE(\hat{\tau}) \quad , \quad \hat{\tau} + z_{1-\alpha/2}SE(\hat{\tau}) \\
&\hat{\tau} - z_{1-\alpha/2}\sqrt{ab^2} \quad , \quad \hat{\tau} + z_{1-\alpha/2}\sqrt{ab^2} \\
&\hat{\tau} - z_{1-\alpha/2}\sqrt{n\left(\frac{\hat{\tau}}{n}\right)^2} \quad , \quad \hat{\tau} + z_{1-\alpha/2}\sqrt{n\left(\frac{\hat{\tau}}{n}\right)^2} \\
&\hat{\tau} - z_{1-\alpha/2}\sqrt{\frac{\hat{\tau}^2}{n}} \quad , \quad \hat{\tau} + z_{1-\alpha/2}\sqrt{\frac{\hat{\tau}^2}{n}} \\
&\bar{X} - z_{1-\alpha/2}\sqrt{\frac{\bar{X}^2}{n}} \quad , \quad \bar{X} + z_{1-\alpha/2}\sqrt{\frac{\bar{X}^2}{n}}
\end{aligned}$$

We see that this small sample interval is the same as the asymptotic interval,

based on the CRLB derivation!

- Now suppose we want an interval for τ^2 , let's work out what this would be for an asymptotic interval:

$$V(\hat{\tau}^2) \approx \frac{\left[\frac{d}{d\tau}\tau^2\right]^2}{I(\tau)} = \frac{4\tau}{\frac{n}{\tau^2}} = \frac{4\tau^3}{n}$$

This suggests an α level interval should be:

$$\hat{\tau}^2 \pm \sqrt{\frac{4\hat{\tau}^3}{n}}$$

- g. We can find the form of an exact confidence interval for $\hat{\tau}$ (i.e. non-asymptotic) through using the fact that it has a gamma distribution.
- We know $\bar{X} \sim \text{gamma}(a = n, b = \tau/n)$.
 - We can simply use the quantiles for the gamma distribution (or convert the gamma to a χ^2 - can you show this transformation?).

$$\text{quantile gamma}(a,b)_{\alpha/2} < \bar{X} < \text{quantile gamma}(a,b)_{1-\alpha/2}$$

4. **Chapter 8 Question 68:** We have $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Poisson}(\lambda)$.

- a. First we need to show that $T = \sum_{i=1}^n X_i$ is a **sufficient statistic** this the brute force way:

$$\begin{aligned} P(X_1, \dots, X_n | T = t) &= \frac{P(X_1, \dots, X_n \cap T = t)}{P(T = t)} \\ &= \begin{cases} \frac{P(X_1, \dots, X_n)}{P(T=t)} & \text{if } \sum_{i=1}^n X_i = t \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Now we can show via moment generating functions (can you show?) that $T \sim \text{Poisson}(n\lambda)$. So we have:

$$\begin{aligned}
P(X_1, \dots, X_n | T = t) &= \frac{P(X_1, \dots, X_n)}{P(T = t)} \\
&= \frac{\frac{\exp(-n\lambda) \lambda^{\sum x_i}}{\prod x_i!}}{\frac{\exp(-n\lambda) (n\lambda)^t}{t!}} \\
&= \frac{t!}{n^t \prod x_i!}
\end{aligned}$$

We can see that this does not depend on λ , so T is a sufficient statistic!

b. Now we to check whether $T = X_1$ is a sufficient statistic.

$$\begin{aligned}
P(X_1, \dots, X_n | T = t) &= \frac{P(X_1, \dots, X_n \cap T = t)}{P(T = t)} \\
&= \frac{P(X_1, \dots, X_n \cap X_1 = t)}{P(X_1 = t)} \\
&= \frac{P(X_1 = t) P(X_2 = x_2) \times P(X_n = x_n)}{P(X_1 = t)} \quad \text{this is 0 if } X_1 \neq t \\
&= \frac{P(X_2 = x_2) \times P(X_n = x_n)}{1} = \frac{\exp(-(n-1)\lambda) \lambda^{\sum_{i=2}^n x_i}}{\prod_{i=2}^n x_i!}
\end{aligned}$$

We see that this quantity still has λ , thus X_1 is not a sufficient statistics for λ .

c. Let's use the factorization theorem:

$$\begin{aligned}
f(X_1, \dots, X_n) &= \frac{\exp(-n\lambda) \lambda^{\sum x_i}}{\prod x_i!} \\
&= [\exp(-n\lambda) \lambda^{\sum x_i}] \left[\frac{1}{\prod x_i!} \right] \\
&= [\exp(-n\lambda) \lambda^t] \left[\frac{1}{\prod x_i!} \right] \\
&= g[T(\mathbf{s}) | \theta] h(\mathbf{x})
\end{aligned}$$

Based on the factorization theorem, we see that $T = \sum X_i$ is a sufficient statistic for λ .