PS6

Monday, 16 November 2020 10:18

1a) Let $X \sim Exp(\lambda)$ for some $\lambda > 0$. We have that the median m satisfies

$$F(X \leq m) = 0.5 = F(X \geq m).$$

So

$$1 - e^{-\lambda m} = 0.5 \implies e^{-\lambda m} = 0.5 \implies m = -\frac{\ln(0.5)}{\lambda}.$$

Now the ML estimate of λ is $\hat{\lambda} = \frac{1}{\bar{x}}$. So by invariance under bijective reparametrisation

$$\hat{m} = -\ln(0.5) \cdot \bar{x}.$$

2a) We have $X_1, \ldots, X_n \overset{iid}{\sim} Cauchy(\theta)$ and $Y_i = \mathbb{I}(X_i \geq \theta)$ for $i \in \{1, \ldots, n\}$. We have that

$$\mathbb{P}(X_i \geq \theta) = 1 - F_{X_i}(\theta).$$

Now

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\pi(1+(x-\theta)^2)} dx = \frac{1}{\pi} \tan^{-1}(x-\theta) \Big|_{-\infty}^{x} = \frac{1}{\pi} \tan^{-1}(x-\theta) + \frac{1}{2}.$$

So $F_{X_i}(\theta) = \frac{1}{2}$. Furthermore $X_1, \ldots, X_n$ are independent so we must have that $Y_1, \ldots, Y_n \overset{iid}{\sim} Bernoulli(\frac{1}{2})$. Furthermore by the definition of the binomial distribution we have that $B = \sum_{i=1}^{n} Y_i \sim Binomial(n, \frac{1}{2})$.

b) Let $Y_{(i)} = \mathbb{I}(x_{(i)} \geq \theta)$ for $i \in \{1, \ldots, n\}$. Now

$$Y_{(i)} = 0 \iff (\forall j \leq i)[Y_{(j)} = 0]$$

$$\iff B = \sum_{i=1}^{n} Y_i \leq n - i.$$

Also

$$Y_{(i)} = 1 \iff (\forall j \geq i)[Y_{(j)} = 1]$$

$$\iff B = \sum_{i=1}^{n} Y_i \geq n - i + 1.$$

c) We have that

$$\mathbb{P}(X_{(k+1)} < \theta \leq X_{(n-k)}; \theta) = \mathbb{P}(\theta \leq X_{(n-k)}; \theta) - \mathbb{P}(\theta < X_{(k+1)})$$

$$= 1 - \mathbb{P}(X_{(n-k)} < \theta; \theta) - (1 - \mathbb{P}(X_{(k+1)} \leq \theta; \theta))$$

$$= -\mathbb{P}(Y_{(n-k)} = 0) + \mathbb{P}(Y_{(k+1)} = 0)$$

$$= \mathbb{P}(Y_{(k+1)} = 0) - \mathbb{P}(Y_{(n-k)} = 0)$$

$$= \mathbb{P}(B \leq n - (k+1)) - \mathbb{P}(B \leq n - (n-k))$$

$$= \mathbb{P}(B \leq n-k-1) - \mathbb{P}(B \leq k)$$

$$= \mathbb{P}(k < B \leq n-k-1)$$

$$= \mathbb{P}(k < B < n-k).$$

d) We want a $1-\alpha$ CI with $\alpha = 0.05$ and $n = 25$. So we require

$$\mathbb{P}(k < B < 25-k) = 0.95.$$

With R we see that

$$\mathbb{P}(3 < B < 25-3) = 0.926\ldots$$

and

$$\mathbb{P}(2 < B < 25-2) = 0.980\ldots.$$

So we need $k = 2$ to give us a 95% CI.

Problem Sheet 1

Problem Sheet 6

Question 1b

```
n <- 7
lambda <- 3.3
N <- 10000
estimates.t <- sapply(1:N, function(i) median(rexp(n, rate = lambda)))
mse <- mean(estimates.t + (log(0.5)/lambda))
mse
```

```
## [1] 0.02066255
```

```
bias <- mse - var(estimates.t)
bias
```

```
## [1] 0.006615086
```

Question 1c

```
estimates.ml <- sapply(1:N, function(i) -log(0.5)*mean(rexp(n, rate = lambda)))
mse <- mean(estimates.ml + (log(0.5)/lambda))
mse
```

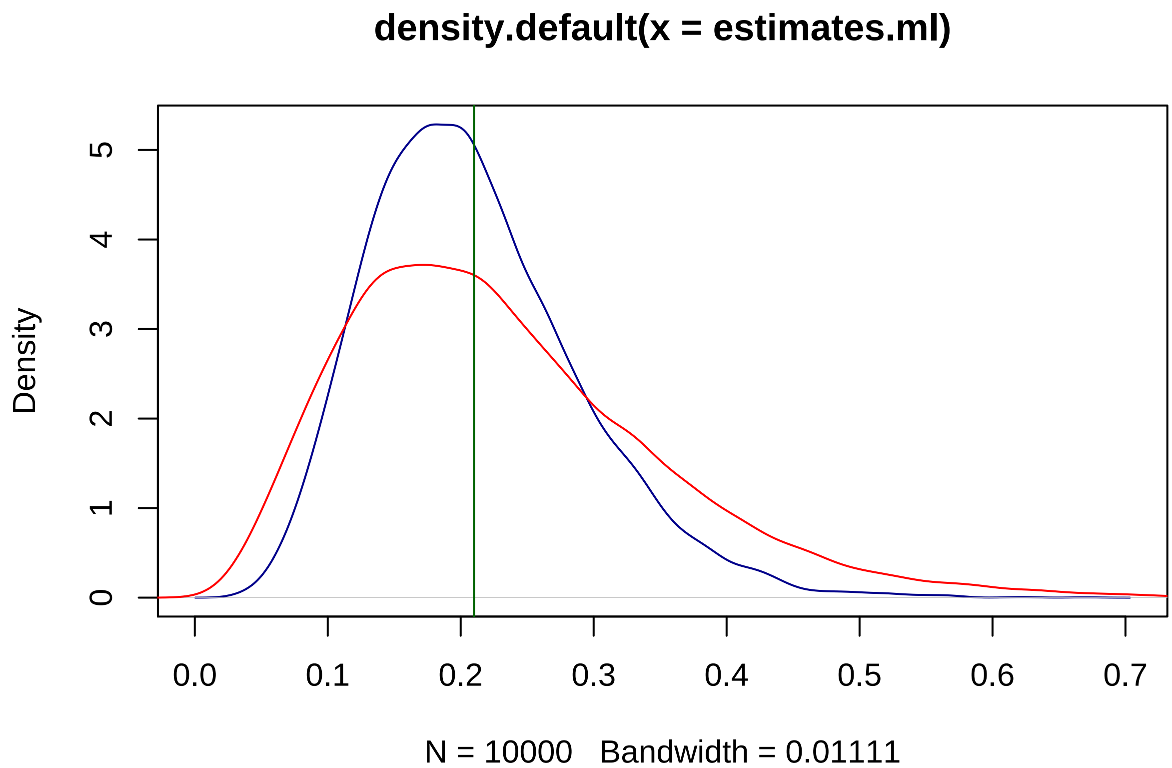
```
## [1] 1.525875e-05
```

```
bias <- mse - var(estimates.ml)
bias
```

```
## [1] -0.006298009
```

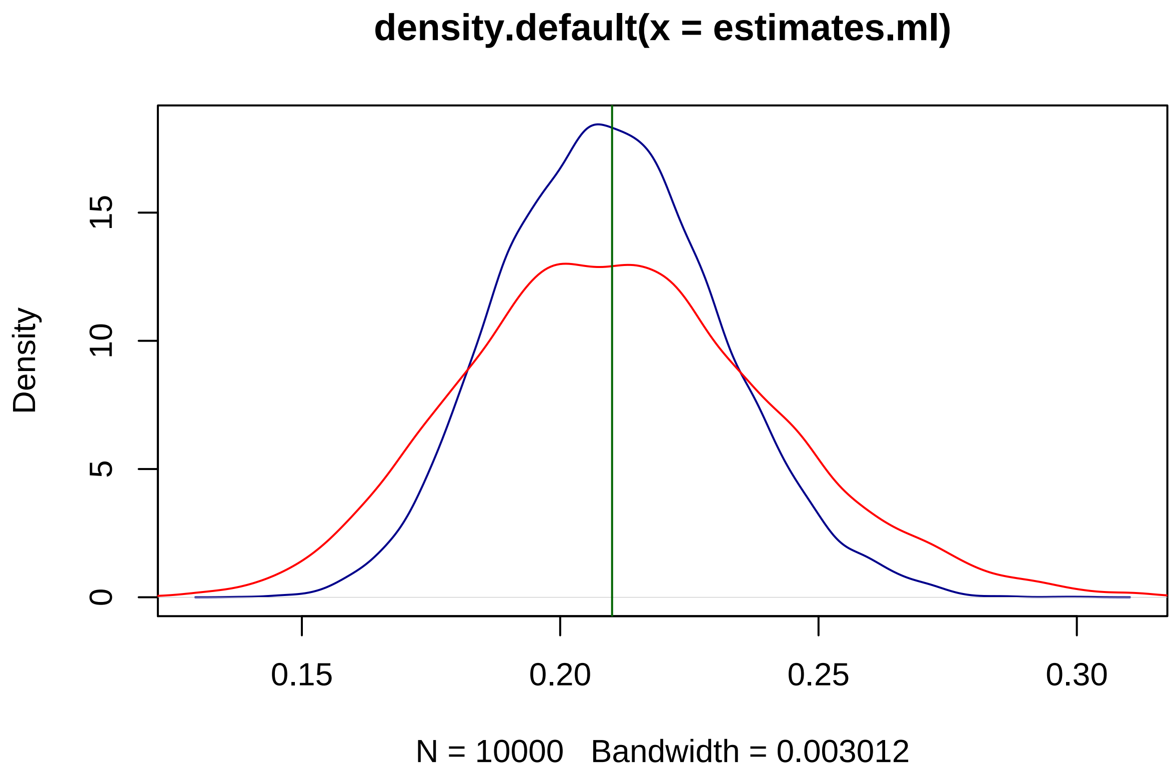
Question 1d

```
plot(density(estimates.ml), col = "darkblue")
lines(density(estimates.t), col = "red")

abline(v = -log(0.5)/lambda, col = "darkgreen")
```



Question 1e

```
n <- 100
lambda <- 3.3
N <- 10000
estimates.t <- sapply(1:N, function(i) median(rexp(n, rate = lambda)))
estimates.ml <- sapply(1:N, function(i) -log(0.5)*mean(rexp(n, rate = lambda)))

plot(density(estimates.ml), col = "darkblue")
lines(density(estimates.t), col = "red")
abline(v = -log(0.5)/lambda, col = "darkgreen")
```



When $n = 100$ the density peak is closer to the true value and the density is much tighter to the true value.