PS5

Saturday, 7 November 2020 14:05

1) Let $\underline{X} \overset{iid}{\sim} \text{Poisson}(\theta^*)$. We want to first find

$$J_n(\hat{\theta}_n) = \frac{1}{n}\sum_{i=1}^n \ell''(\hat{\theta}_n; X_i).$$

We have the liklihood function

$$\ell(\theta; x_i) = \log \frac{\theta^{x_i} e^{-\theta}}{x_i!} = x_i \log\theta - \theta - \log x_i!$$

$$\Rightarrow \ell'(\theta; x_i) = \frac{x_i}{\theta} - 1 \Rightarrow \ell''(\theta; x_i) = -\frac{x_i}{\theta^2}.$$

Hence

$$J_n(\hat{\theta}_n) = \frac{-1}{n}\sum_{i=1}^n -\frac{x_i}{\theta^2} = \bar{x}\,\hat{\theta}_n^{-2}.$$

Then as $\hat{\theta}_n = \bar{x}$ we have $J_n(\hat{\theta}_n) = \frac{1}{\bar{x}}$.

As $\{\hat{\theta}_n\}$ is a consistent sequence of estimators and the required conditions are satisfied we can say

$$\sqrt{nJ_n(\hat{\theta}_n)}\,(\hat{\theta}_n - \theta^*) \xrightarrow{\mathcal{D}(\cdot;\theta^*)} Z.$$

Hence an exact $1-\alpha$ CI for θ^* is defined by

$$L(\underline{x}) = \hat{\theta}_n - z_{\alpha/2}\Big/\sqrt{n J_n(\hat{\theta}_n)} = \bar{x} - z_{\alpha/2}\Big/\sqrt{n/\bar{x}}$$

and

$$U(\underline{x}) = \bar{x} + z_{\alpha/2}\Big/\sqrt{n/\bar{x}}.$$

So our CI is

$$\left[\bar{x} - z_{\alpha/2}\sqrt{\bar{x}/n},\ \bar{x} + z_{\alpha/2}\sqrt{\bar{x}/n}\right].$$

2a) We have $\underline{x} = (4, 23, 20, 11, 15, 29, 20, 16, 15, 14)$ from some iid random sample $\underline{X}$. By the CLT as $\mathbb{E}(X) < \infty$ and $\text{Var}(X) = \sigma^2 < \infty$ we have

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{\mathcal{D}} Z \sim N(0,1).$$

where $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. We have

$$\mathbb{P}\left(z_{0.025} \le \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \le z_{0.975}\right)$$

$$= \mathbb{P}\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \le z_{0.975}\right) - \mathbb{P}\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \le z_{0.025}\right)$$

$$\xrightarrow{n\to\infty} 0.975 - 0.025$$

$$= 0.95.$$

Also

$$\mathbb{P}\left(z_{0.025} \le \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \le z_{0.975}\right)$$

$$= \mathbb{P}\left(z_{0.025} \frac{\sigma}{\sqrt{n}} \le \bar{X}_n - \mu \le z_{0.975} \frac{\sigma}{\sqrt{n}} \right)$$

$$= \mathbb{P}\left(-\bar{X}_n + z_{0.025} \frac{\sigma}{\sqrt{n}} \le -\mu \le -\bar{X}_n + z_{0.975} \frac{\sigma}{\sqrt{n}} \right)$$

$$= \mathbb{P}\left(\bar{X}_n - z_{0.975} \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X}_n - z_{0.025} \frac{\sigma}{\sqrt{n}} \right)$$

$$= \mathbb{P}\left(\bar{X}_n - z_{0.975} \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X}_n + z_{0.975} \frac{\sigma}{\sqrt{n}} \right).$$

Now we have $\bar{X}_n = \frac{1}{10}(11 + 23 + \dots + 14) = 17.4$, $z_{0.975} = 1.96$ and $n = 10$. We can estimate σ^2 with the sample variance which gives us $\sigma^2 = \frac{1}{9}\left[3314 - \frac{174^2}{10}\right] = 31.82$ (2dp). So our CI is

$$[13.9, 20.9].$$

b) We have $\underline{X} \overset{iid}{\sim} \text{Poisson}(\theta)$. From question 1 we get the CI

$$\left[17.4 - 1.96\sqrt{\tfrac{17.4}{10}},\ 17.4 + 1.96\sqrt{\tfrac{17.4}{10}}\right] = [14.8, 20.0].$$

c) The second is a tighter interval so we "know" more about where μ is, however we had to assume the data was poisson.

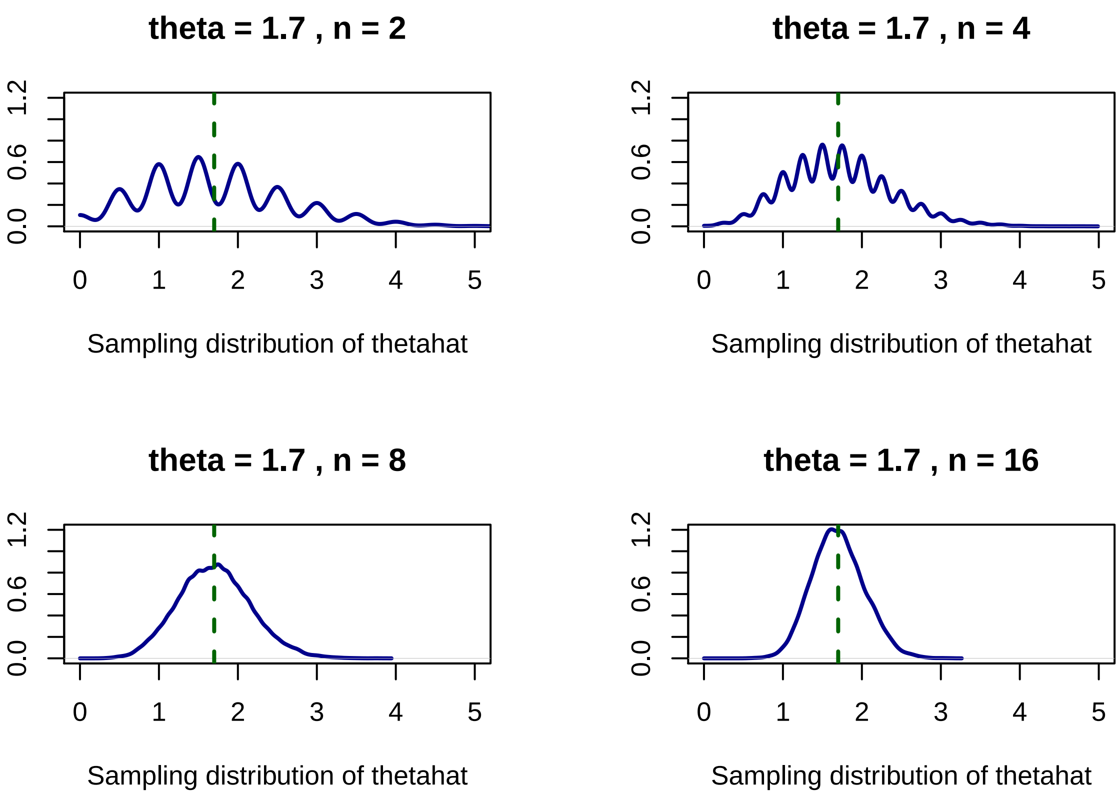
Problem Sheet 5

Problem Sheet 5

Question 3a

```
distribution <- function(n, theta) {
  N <- 10000
  that <- sapply(1:N, function(i) mean(rpois(n = n, lambda = theta)))
  plot(density(that, from = 0), xlim = c(0, 5), ylim = c(0,1.2),
  main = paste("theta =", theta, ", n =", n),
  xlab = "Sampling distribution of thetahat", ylab = "",
  col = "darkblue", lwd = 2)
  abline(v = theta, col = "darkgreen", lwd = 2, lty = 2)
}

par(mfrow = c(2, 2))
distribution(2, 1.7)
distribution(4, 1.7)
distribution(8, 1.7)
distribution(16, 1.7)
```

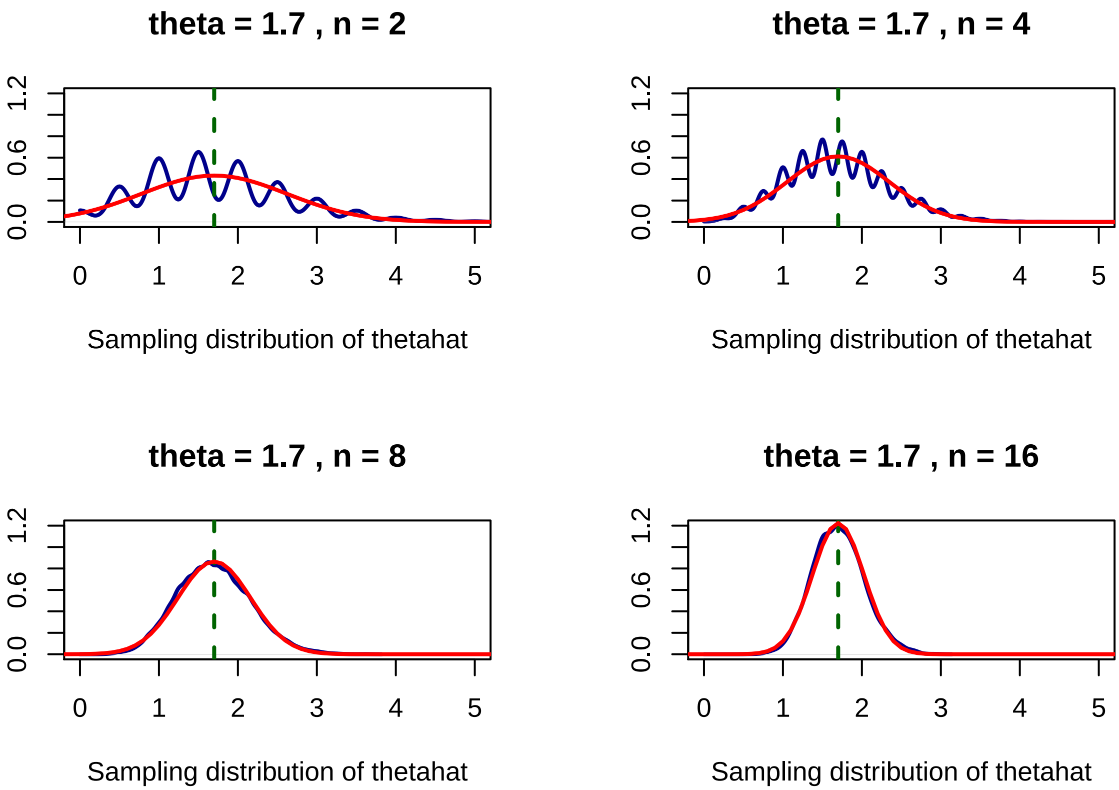


Question 3b

```r
distribution <- function(n, theta) {
  N <- 10000
  that <- sapply(1:N, function(i) mean(rpois(n = n, lambda = theta)))
  plot(density(that, from = 0), xlim = c(0, 5), ylim = c(0,1.2),
  main = paste("theta =", theta, ", n =", n),
  xlab = "Sampling distribution of thetahat", ylab = "",
  col = "darkblue", lwd = 2)
  abline(v = theta, col = "darkgreen", lwd = 2, lty = 2)

  x <- seq(-10, 10, by = .1)
  y <- dnorm(x, mean = theta, sd = sqrt(theta/n))
  lines(x, y, col = "red", lwd = 2)
}

par(mfrow = c(2, 2))
distribution(2, 1.7)
distribution(4, 1.7)
distribution(8, 1.7)
distribution(16, 1.7)
```


theta = 1.7 , n = 2
theta = 1.7 , n = 4
theta = 1.7 , n = 8
theta = 1.7 , n = 16
0.0
0.6
1.2
0
1
2
3
4
5
Sampling distribution of thetahat