Homework 4a

yi Chen

9/24/2018

homework 4a

yi chen (uni: yc3356)

yc3356@columbia.edu (mailto:yc3356@columbia.edu)

1. problem 4.5 of BDA

A) use the simulated data

```
N <- 100000000
x <- rnorm(N,mean=4,sd=1)
y <- rnorm(N,mean=3,sd=2)
results <- x/y
cat('mean is:',mean(results),'\n')
```

```
## mean is: 3.078484
```

```
cat('standard deviation is:',sd(results),'\n')
```

```
## standard deviation is: 25270.62
```

B) without simulation

for any f(x,y), the bivariate second order Taylor expansion about $\theta = (\mu_x, \mu_y)$ is:

$$f(x,y) \approx f(\boldsymbol{\theta}) + f_x^{'}(\boldsymbol{\theta})(x-\mu_x) + f_y^{'}(\boldsymbol{\theta})(y-\mu_y) + \frac{1}{2}\{f_{xx}^{''}(\boldsymbol{\theta})(x-\mu_x)^2 + f_{yy}^{''}(\boldsymbol{\theta})(y-\mu_y)^2 + 2f_{xy}^{''}(\boldsymbol{\theta})(x-\mu_x)\}(y-\mu_y)\}$$

Therefore:

$$E(f(X,Y)) \approx E(f(\boldsymbol{\theta})) + E(f_x^{'}(\boldsymbol{\theta})(x-\mu_x)) + E(f_y^{'}(\boldsymbol{\theta})(y-\mu_y)) + \frac{1}{2}E\{f_{xx}^{''}(\boldsymbol{\theta})(x-\mu_x)^2 + f_{yy}^{''}(\boldsymbol{\theta})(y-\mu_y)^2 + 2f_{xy}^{''}(\boldsymbol{\theta})(x-\mu_x)\}(y-\mu_y)\}$$

$$= E(f(\boldsymbol{\theta})) + f_x^{'}(\boldsymbol{\theta})E((x-\mu_x)) + f_y^{'}(\boldsymbol{\theta})E((y-\mu_y)) + \frac{1}{2}\{f_{xx}^{''}(\boldsymbol{\theta})E(x-\mu_x)^2 + f_{yy}^{''}(\boldsymbol{\theta})E(y-\mu_y)^2 + 2f_{xy}^{''}(\boldsymbol{\theta})E(x-\mu_x)\}(y-\mu_y)\}$$

$$= E(f(\boldsymbol{\theta})) + 0 + 0 = f(\mu_x, \mu_y) + \frac{1}{2}\{f_{xx}^{''}(\boldsymbol{\theta})V(X) + f_{yy}^{''}(\boldsymbol{\theta})V(Y) + 2f_{xy}^{''}(\boldsymbol{\theta})Cov(X,Y)\}$$

Here:$f(x,y) = y/x$. Thus:

$$f_{yy}^{''} = 0, f_{xy}^{''} = -y^{-2}, f_{xx}^{''} = \frac{2x}{y^3}$$

Thus, we have:

$$E(f(x,y)) = E(y/x) \approx \frac{\mu_y}{\mu_x} - \frac{Cov(Y,X)}{\mu_x^2} + \frac{V(X)\mu_y}{\mu_x^3}$$

Plug in the data we get:

$$E(f(x,y)) \approx \frac{3}{4} - \frac{0}{4^2} + \frac{1 \times 3}{4^3} = \frac{45}{64} = 0.703125$$

Using the same idea, for the variance we take the first order Taylor expansion :

$$V(y/x) = E\{[f(X,Y) - E(f(X,Y))]^2\} \approx E\{[f(\boldsymbol{\theta}) + f_x'(\boldsymbol{\theta})(x-\mu_x) + f_y'(\boldsymbol{\theta})(y-\mu_y) - E(f(X,Y))]^2\}$$

.

$$= E\{[f_x'(\boldsymbol{\theta})(x-\mu_x) + f_y'(\boldsymbol{\theta})(y-\mu_y)]^2\}$$

$$= E\{f_x'^2(\boldsymbol{\theta})(x-\mu_x)^2 + f_y'^2(\boldsymbol{\theta})(y-\mu_y)^2 + 2f_x'(\boldsymbol{\theta})(y-\mu_x)f_y'(\boldsymbol{\theta})(y-\mu_y)\}$$

$$= f_x'^2(\boldsymbol{\theta})V(X) + 2f_x'(\boldsymbol{\theta})f_y'(\boldsymbol{\theta})Cov(X,Y) + f_y'^2(\boldsymbol{\theta})V(Y)$$

Here again :$f(x,y) = y/x$. Thus:

$$f_{yy}'' = 0, f_{xy}'' = -y^{-2}, f_{xx}'' = \frac{2x}{y^3}$$

$$V(y/x) \approx \frac{1}{\mu_x^2} - 2\frac{\mu_y}{\mu_x^3}Cov(X,Y) + \frac{\mu_y^2}{\mu_x^4}V(x)$$

Plug in the data, we get:

$$V(y/x) \approx \frac{1}{4^2} - 2 \times \frac{3}{4^3} \times 0 + \frac{3^2}{4^4} \times 1 = \frac{25}{256} = 0.09765625$$

Thus, standard deviation is $\frac{5}{16} = 0.315$

C)

It should ensure the there at least would have very litter mass at 0.

2. problem 4.9 of BDA

For the maximum likelihood esitmation: since the esitmation need to be restricted to [0,1] which means all the MLE which is smaller than 0 would be estimated as 0 and all the MLE which is bigger than 1 would be estimated as 1. Besides these two points, all the other points are still follow the normal distritbuion with mean θ (sample mean is unbiased esimator) and variance $\frac{\sigma^2}{n}$ (according to the central limit theorem) Thus, the distribution of the MLE (here sample mean) is not a normal distribution. Thus:

$$P(\hat{\theta}_{MLE} = 0) = \phi(\frac{0-\theta}{\sqrt{\frac{\sigma^2}{n}}}) = \phi(\frac{-\sqrt{n}\theta}{\sigma})$$

$$P(\hat{\theta}_{MLE} = 1) = 1 - \phi(\frac{1-\theta}{\sqrt{\frac{\sigma^2}{n}}}) = 1 - \phi(\frac{\sqrt{n}(1-\theta)}{\sigma})$$

Here both n and the “true” value of θ are fixed. Thus, as the variance increasing, we have

$$\lim_{\sigma \to +\infty} P(\hat{\theta}_{MLE} = 0) = \lim_{\sigma \to +\infty} P(\hat{\theta}_{MLE} = 1) = \frac{1}{2}$$

That basically means, as the variance increasing, more and more probability has been shifted equally towards the point 1 and 0 and the probability of the value within 0 an 1 are decreasing. Thus, we can say:

$$\lim_{\sigma \to +\infty} MSE(\hat{\theta}_{MLE}) = \lim_{\sigma \to +\infty} E_{\hat{\theta}_{MLE}} (\hat{\theta}_{MLE} - \theta)^2 = \frac{1}{2}((1-\theta)^2 + \theta^2)$$

For the bayesian estimation: the posterior distribution of σ is:

$$p(\theta|y) = \frac{p(\theta)p(y|\theta)}{\int p(\theta)p(y|\theta)d\theta} \propto p(y|\theta)$$

As we know, since the uniform prior is a week informatative prior, thus the posterior distribution is based on the likelihood. As the σ increase, the posterior would be almost a constant over [0,1]. Thus, the postrior mean woule be 1/2. Thus:

$$\lim_{\sigma \to +\infty} MSE(\hat{\theta}_B) = (\theta - 1/2)^2$$

Thus, we can get:

$$\lim_{\sigma \to +\infty} (MSE(\hat{\theta}_{MLE}) - MSE(\hat{\theta}_B)) = \frac{1}{2}((1-\theta)^2 + \theta^2) - (\theta - 1/2)^2 = 1/2$$

Thus, we can say that the Bayesian posterior mean always has a smaller MSE than MLE.