Stat511(Section001): Homework #5

Due on Feb. 23, 2021 at 5:00pm

Instructor:Xuming He

Tiejin Chen **tiejin@umich.edu**

Problem 1

We can have the pmf of X_i is:

$$f(x) = \frac{1}{\sqrt{2\pi}\theta} exp(-\frac{(x-\theta)^2}{2\theta^2})$$

Then we can get the likehood function of the data is:

$$L(x_1, ..., x_n) = \frac{1}{2\pi^{n/2}\theta^n} exp(-\frac{1}{2\theta^2}\sum_{i=1}^{n}(x_i - \theta)^2)$$

Now we can take log of it to get:

$$l(x_1, ..., x_n) = log(\frac{1}{2\pi^{n/2}\theta^n}) - \frac{1}{2\theta^2}\sum_{i=1}^{n}(x_i - \theta)^2 = -log(2\pi^{n/2}) - log(\theta^n) - \frac{1}{2\theta^2}\sum_{i=1}^{n}(x_i - \theta)^2$$

Then take derivative of θ to get:

$$l'(x_1, ..., x_n) = -\frac{n}{\theta} + \frac{1}{\theta^3}\sum_{i=1}^{n}(x_i - \theta)^2 + \frac{1}{2\theta^2}(2n\bar{X} - 2n\theta) = 0$$

Then we can get:

$$-n\theta^2 + \sum_{i=1}^{n} x_i^2 - 2n\bar{X}\theta + n\theta^2 + n\theta\bar{X} - n\theta^2 = 0$$

And we have $\bar{X} = 0$, thus:

$$n\theta^2 = \sum_{i=1}^{m} x_i^2$$

Then we can get the MLE of θ is:

$$\hat{\theta} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}$$

Problem 2

(a)

For first moment we have:

$$\frac{1}{k}\sum_{i=1}^{k} X_i = E(X) = np$$

Then we have the second centered moment:

$$\frac{1}{k-1}\sum_{i=1}^{k}(X_i - \bar{X})^2 = Var(X) = np(1-p)$$

Then, we have:

$$1 - p = \frac{(k-1)\sum_{i=1}^{k} X_i}{k\sum_{i=1}^{k}(X_i - \bar{X})^2} \rightarrow \hat{p} = 1 - \frac{(k-1)\sum_{i=1}^{k} X_i}{k\sum_{i=1}^{k}(X_i - \bar{X})^2}$$

$$\hat{n} = \frac{\frac{1}{k-1}\sum_{i=1}^{k} X_i}{1 - \frac{(k-1)\sum_{i=1}^{k} X_i}{k\sum_{i=1}^{k}(X_i - \bar{X})^2}}$$

with methof of moments

Problem 2 continued on next page. . .

(b)
one data from binomial distribution, thus can been see as the sum of n Bernouli distribution with p. And we can write its pdf:

$$f(x) = I(x = 1)p + (1 - p)I(x = 0)$$

Hence the likelihood function of X is:

$$L(x) = (I(x = 1)p + (1 - p)I(x = 0))^n$$

First we assumed p is fixed, then we have $n \geq 8$. We can see that for every data in Bernouli, its likelihood value is whether p or (1-p). And they are all smaller than 1. Hence, the more data in Bernouli, the less likelihood we will get. And we have $L(x) \leq p^8$ since we know $X = 8$. And the equation holds only when $n = 8$ and all the Bernoulli trials turn out $x_i = 1$. If there exists any trial turns out $x_i = 0$. Then the likelihood will be $p^8(1 - p)^{n-8}$ which is smaller than p^8. Thus we can get, to make likelihood as larger as possible, n must be 8 for any p. Then, we have:

$$L(x) = p^8$$

This function is an incresing function with p. Hence it will reach the maximum value when $p = 1$. Then we can get MLE of n and p:

$$\hat{n} = 8$$

$$\hat{p} = 1$$

Problem 3

(a)
We know the joint pdf of them is:

$$f(x, y) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} exp(-\frac{1}{2(1-\rho^2)}\frac{x^2 - 2\rho xy + y^2}{\sigma^2})$$

Then we can get the likelihood function:

$$L(x_1, .., x_n, y_1, ..., y_n) = (2\pi\sigma^2\sqrt{1-\rho^2})^{-n} exp(-\frac{1}{2(1-\rho^2)}\frac{\sum_{i=1}^n x_i^2 - 2\rho x_i y_i + y_i^2}{\sigma^2})$$

Take log of it to get:

$$l(x_1, .., x_n, y_1, ..., y_n) = -nlog(2\pi) - 2nlog(\sigma) - nlog(\sqrt{1-\rho^2}) - \frac{1}{2(1-\rho^2)}\frac{\sum_{i=1}^n x_i^2 - 2\rho x_i y_i + y_i^2}{\sigma^2}$$

Then we have:

$$\frac{\partial l(x_1, .., x_n, y_1, ..., y_n)}{\partial \sigma} = -\frac{2n}{\sigma} + \frac{1}{1-\rho^2}\frac{\sum_{i=1}^n x_i^2 - 2\rho x_i y_i + y_i^2}{\sigma^3} = 0$$

$$\frac{\partial l(x_1, .., x_n, y_1, ..., y_n)}{\partial \rho} = \frac{n\rho}{1-\rho^2} - \frac{\rho}{(1-\rho^2)^2}\frac{\sum_{i=1}^n x_i^2 - 2\rho x_i y_i + y_i^2}{\sigma^2} + \frac{1}{1-\rho^2}\frac{\sum_{i=1}^n x_i y_i}{\sigma^2} = 0$$

From the first equation, we know:

$$2n = \frac{1}{1-\rho^2}\frac{\sum_{i=1}^n x_i^2 - 2\rho x_i y_i + y_i^2}{\sigma^2}$$

Plug this in the second equation to get:

$$\frac{n\rho}{1-\rho^2} = \frac{1}{1-\rho^2}\frac{\sum_{i=1}^n x_i y_i}{\sigma^2}$$

Problem 3 continued on next page. . .

$$\sigma^2 = \frac{\frac{1}{n}\sum_{i=1}^n x_i y_i}{\rho}$$

Thus plug this in the first equation to get:

$$\frac{2\sum_{i=1}^n x_i y_i}{\rho} = \frac{\sum_{i=1}^n x_i^2 - 2\rho x_i y_i + y_i^2}{1-\rho^2}$$

$$\hat{\rho}_{MLE} = \frac{2\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n (x_i^2 + y_i^2)}$$

$$\hat{\sigma^2}_{MLE} = \frac{\frac{1}{n}\sum_{i=1}^n x_i y_i}{\hat{\rho}_{MLE}}$$

(b)
We know that:

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y) = 2\sigma^2 + 2\rho\sigma^2$$

$$Var(X-Y) = Var(X) + Var(Y) - 2Cov(X,Y) = 2\sigma^2 - 2\rho\sigma^2$$

Then we can compute the sample variance which is the second centered moment of $(X+Y, X-Y)$ to get the method-of-moments estimator. And let $\mu_1 = \bar{X+Y}$ and $\mu_2 = \bar{X-Y}$. Then we have:

$$\frac{1}{n-1}\sum_{i=1}^n (x_i + y_i - \mu_1)^2 = Var(X+Y) = 2\sigma^2 + 2\rho\sigma^2$$

$$\frac{1}{n-1}\sum_{i=1}^n (x_i - y_i - \mu_2)^2 = Var(X-Y) = 2\sigma^2 - 2\rho\sigma^2$$

Then we can get:

$$\hat{\sigma^2} = \frac{1}{4(n-1)}[n(\mu_1^2 + \mu_2^2) + \sum_{i=1}^n x_i^2 + y_i^2 - 2(x_i + y_i)(\mu_1 + \mu_2)]$$

$$\hat{\rho} = \frac{1}{2}\frac{\sum_{i=1}^n (2x_i - \mu_1 - \mu_2)(2y_i - \mu_1 + \mu_2)}{\sum_{i=1}^n (x_i + y_i - \mu_1)^2 + (x_i - y_i - \mu_2)^2}$$

Problem 4

(a)
We have:

$$E(\tilde{\lambda}) = \frac{1}{n}\sum_{i=1}^n E(\frac{X_i}{c_i}) = \frac{1}{n}n\lambda = \lambda$$

Thus it is a unbiased estimator. Now consider the variance of it:

$$Var(\tilde{\lambda}) = \frac{1}{n^2}\sum_{i=1}^n Var(\frac{X_i}{c_i}) = \frac{1}{n^2}\sum_{i=1}^n \frac{c_i\lambda}{c_i^2} = \frac{1}{n^2}\sum_{i=1}^n \frac{\lambda}{c_i}$$

Hence the MSE of $\tilde{\lambda}$ is $\frac{1}{n^2}\sum_{i=1}^n \frac{\lambda}{c_i}$.
(b)
For every i, we have the pmf of X_i is:

$$p(X_i = x_i) = \frac{(c_i\lambda)^{x_i} e^{-c_i\lambda}}{x_i!}$$

Problem 4 continued on next page. . .

Thus the likelihood function would be:

$$L(x_1, ..., x_n) = \frac{\lambda^{\sum_{i=1}^n x_i} e^{-(\sum_{i=1}^n c_i)\lambda} \prod_{i=1}^n c_i^{x_i}}{\prod_{i=1}^n x_i!}$$

Thus the kernel of log likelihood function is:

$$l(x_1, ..., x_n) = \sum_{i=1}^n x_i log(\lambda) - \sum_{i=1}^n c_i \lambda$$

$$l'(x_1, ..., x_n) = \sum_{i=1}^n x_i \frac{1}{\lambda} - \sum_{i=1}^n c_i = 0$$

$$\hat{\lambda} = \frac{\sum_{i=1}^n x_i}{\sum_{i=1}^n c_i}$$

This is MLE estimator of λ. And

$$E(\hat{\lambda}) = \frac{1}{\sum_{i=1}^n c_i} E(\sum_{i=1}^n x_i) = \frac{\sum_{i=1}^n c_i \lambda}{\sum_{i=1}^n c_i} = \lambda$$

Thus it is an unbiased estimator. Then we consider its variance:

$$Var(\hat{\lambda}) = \frac{1}{(\sum_{i=1}^n c_i)^2} Var(\sum_{i=1}^n x_i) = \frac{\lambda}{\sum_{i=1}^n c_i}$$

Thus its MSE is $\frac{\lambda}{\sum_{i=1}^n c_i}$. Now compare $\frac{1}{\sum_{i=1}^n c_i}$ and $\frac{1}{n^2}\sum_{i=1}^n \frac{1}{c_i}$. By HM-GM-AM-QM inequalities, we can get:

$$\frac{n}{\sum_{i=1}^n c_i} = \frac{n}{\sum_{i=1}^n \frac{1}{\frac{1}{c_i}}} \leq \frac{\sum_{i=1}^n \frac{1}{c_i}}{n}$$

Thus:

$$\frac{1}{\sum_{i=1}^n c_i} \leq \frac{1}{n^2} \sum_{i=1}^n \frac{1}{c_i}$$

Therefore, $\hat{\lambda}$ is a better estimator compared to $\tilde{\lambda}$ under MSE.