HOMEWORK 6

2021/11/5

Assignment 1. Suppose $X_1, \ldots, X_n$ $i.i.d. \sim Bernoulli(\theta)$, $\quad \theta \in (0,1)$. With prior $\theta \sim Beta(a,b)$, $\hat{\theta}_B = \frac{\sum_{i=1}^n X_i + a}{n+a+b}$.

(1) Solve $MSE(\hat{\theta}_B)$.

Since $\sum_{i=1}^n X_i \sim B(n,\theta)$, we have

$$E\hat{\theta}_B = \frac{n\theta + a}{n+a+b}, \quad Var(\hat{\theta}_B) = \frac{n\theta(1-\theta)}{(n+a+b)^2}.$$

Thus, with the definition of $MSE = Var + Bias^2$, we conclude that

$$MSE(\hat{\theta}_B) = Var(\hat{\theta}_B) + \left(E\hat{\theta}_B - \theta\right)^2 = \frac{\left[(a+b)^2 - n\right]\theta^2 + \left[n - 2a(a+b)\right]\theta + a^2}{(n+a+b)^2}. \quad \square$$

(2) Solve $\hat{\theta}_{MoM}$, $\hat{\theta}_{MLE}$ and their MSE.

Since $EX_1 = \theta$, we have $\hat{\theta}_{MoM} = \bar{X}$, where $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ is the sample mean.

The likelihood function is $lik(\theta) = \theta^{\sum_i X_i}(1-\theta)^{n-\sum_i X_i}$. To maximize $lik(\theta)$, consider $l(\theta) = \log(lik(\theta)) = \sum X_i \log(\theta) + (n - \sum X_i)\log(1-\theta)$ and calculate

$$\frac{\partial l}{\partial \theta} = \sum X_i \frac{1}{\theta} - (n - \sum X_i)\frac{1}{1-\theta}.$$

Notice that $\frac{\partial l}{\partial \theta} = 0$ iff $\theta^* = \bar{X}$. Moreover, $\frac{\partial^2 l}{\partial \theta^2} = -\frac{\sum X_i}{\theta^2} - \frac{(n-\sum X_i}{(1-\theta)^2} \leq 0$. Thus, $\hat{\theta}_{MLE} = \bar{X}$. (For X_i are all 1 or all 0, some arguments are required since θ^* is not a inner point in $(0,1)$, which is similar to our homework before.)

For their MSE, notice that $n\bar{X} \sim B(n,\theta)$. Therefore,

$$\begin{aligned} MSE(\hat{\theta}_{MoM}) &= MSE(\hat{\theta}_{MLE}) = MSE(\bar{X}) \\ &= Var(\bar{X}) + \left(E\bar{X} - \theta\right)^2 = \frac{\theta(1-\theta)}{n}. \end{aligned}$$

$\square$

Remark: 本次参考答案中略去了关于 MLE 在样本上取值情况的讨论。

Date: 2021/11/24.

Thanks for Weiyu Li who is with the School of the Gifted Young, University of Science and Technology of China. Corresponding Email: liweiyu@mail.ustc.edu.cn.

Assignment 2. Suppose $X_1, \dots, X_n$ *i.i.d.* $\sim U(0,\theta), \quad \theta > 0$.

(1) Solve $\hat{\theta}_{MoM}$, $\hat{\theta}_{MLE}$ and their MSE.

Since $EX_1 = \frac{\theta}{2}$, we have $\hat{\theta}_{MoM} = 2\bar{X}$, where $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ is the sample mean. Moreover, with $Var(X_1) = \frac{\theta^2}{12}$, we can solve that

$$MSE(\hat{\theta}_{MoM}) = Var(2\bar{X}) + \left(E(2\bar{X}) - \theta\right)^2 = \frac{\theta^2}{3n}.$$

For MLE, the likelihood function is $lik(\theta) = \theta^{-n} I_{(X_{(n)},+\infty)}(\theta) I_{(0,+\infty)}(X_{(1)})$, which reaches its maximum as $\theta^* \downarrow X_{(n)}$. Extend the likelihood to $X_{(n)}$, thus$\hat{\theta}_{MLE} = X_{(n)}$.

One can write the $p.d.f$ of $X_{(n)}$ as

$$f_{X_{(n)}}(x) = n\left(P(X < x)\right)^{n-1} f(x) = \frac{nx^{n-1}}{\theta^n} I_{(0,\theta)}(x).$$

Therefore,

$$\begin{aligned}
EX_{(n)} &= \int_0^\theta x \frac{nx^{n-1}}{\theta^n} dx = \frac{n}{n+1}\theta, \\
EX_{(n)}^2 &= \int_0^\theta x^2 \frac{nx^{n-1}}{\theta^n} dx = \frac{n}{n+2}\theta^2, \\
MSE(\hat{\theta}_{MLE}) &= EX_{(n)}^2 - \left(EX_{(n)}\right)^2 + \left((E\bar{X}) - \theta\right)^2 \\
&= \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\theta\right)^2 + \left(\frac{n}{n+1}\theta - \theta\right)^2 \\
&= \frac{2}{(n+1)(n+2)}\theta^2.
\end{aligned}$$

□

(2) With prior $\theta \sim U(0,1)$, solve $\hat{\theta}_B$ and its MSE.

Consider the kernels

$$\begin{aligned}
f(\mathbf{x};\theta) &\propto \theta^{-n} I_{(X_{(n)},\infty)}(\theta), \\
\pi(\theta) &\propto I_{(0,1)}(\theta), \\
\pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x};\theta) \propto \theta^{-n} I_{(X_{(n)},1)}(\theta).
\end{aligned}$$

Since $\int \pi(\theta|\mathbf{x}) d\theta = 1$, we have to solve the problem case by case.

- **case 1:** $n = 1$**.**
  In this case, $\pi(\theta|X_1) = -\frac{1}{\theta \log X_1} I_{(X_1,1)}(\theta)$. Thus,
  $$\hat{\theta}_B = E(\theta|X_1) = \int_{X_1}^{1} -\theta \frac{1}{\theta \log X_1} d\theta = \frac{1 - X_1}{\log X_1}.$$
- **case 2:** $n = 2$**.**
  In this case, $\pi(\theta|\mathbf{X}) = \frac{X_{(2)}}{\theta^2(1-X_{(2)})} I_{(X_{(2)},1)}(\theta)$. Thus,
  $$\hat{\theta}_B = E(\theta|\mathbf{X}) = \int_{X_{(2)}}^{1} \theta \frac{X_{(2)}}{\theta^2(1 - X_{(2)})} d\theta = -\frac{X_{(2)} \log X_{(2)}}{1 - X_{(2)}}.$$
- **case 3:** $n > 2$**.**
  In this case, $\pi(\theta|\mathbf{X}) = \frac{1-n}{1-X_{(n)}^{1-n}} \theta^{-n} I_{(X_{(n)},1)}(\theta)$. Thus,
  $$\hat{\theta}_B = E(\theta|\mathbf{X}) = \frac{1-n}{2-n} \cdot \frac{1 - X_{(n)}^{2-n}}{1 - X_{(n)}^{1-n}}.$$

In all the cases above, I cannot solve the MSE but remain it as an integral from
$$MSE(\hat{\theta}_B) = E(\hat{\theta}_B - \theta)^2 = \int_{X_{(n)}}^{1} (\hat{\theta}_B - \theta)^2 \pi(\theta|\mathbf{X}) d\theta.$$
□

(3) With respect to the MSEs, which estimator is the best?

Since I cannot give an explicit form of $MSE(\hat{\theta}_B)$, just compare $MSE(\hat{\theta}_{MoM}) = \frac{\theta^2}{3n}$ with $MSE(\hat{\theta}_{MLE}) = \frac{2}{(n+1)(n+2)}\theta^2$.

Notice that $\frac{\theta^2}{3n} - \frac{2}{(n+1)(n+2)}\theta^2 = \frac{(n-1)(n-2)}{3n(n+1)(n+2)}\theta^2$. When $n \leq 2$, they are the same. When $n > 2$, the latter one is smaller. Thus, the maximum likelihood estimator is better if $n > 2$ and not worse than $\hat{\theta}_{MoM}$ otherwise.

2021/11/12

[Wei] 3.22.

(1) The likelihood function is $lik(\mu) = e^{n\mu - \sum_i x_i} I_{(-\infty, x_{(1)})}(\mu)$. To maximize the likelihood, μ should be as large as possible. Extend lik to $x_{(1)}$ thus

$$\hat{\mu}^* = x_{(1)}.$$

Since for any $x \geq \mu$,

$$P(X_1 > x) = \int_x^\infty f(t, \mu) dt = e^{\mu - x},$$

and $P(X_{(1)} > x) = (P(X_1 > x))^n$, the $p.d.f.$ of $x_{(1)}$ is

$$f_{X_{(1)}}(x) = -\frac{d}{dx} P(X_{(1)} > x) = ne^{n\mu - nx} I_{(\mu, +\infty)}(x).$$

Then we have

$$E\hat{\mu}^* = EX_{(1)} = \int_\mu^{+\infty} x \cdot ne^{n\mu - nx} dx = \mu + \frac{1}{n}.$$

So $\hat{\mu}^*$ is biased, while $\hat{\mu}^{**} = x_{(1)} - \frac{1}{n}$ is unbiased. $\square$

(2) Since $EX_1 = \mu + 1$, a moment estimator is

$$\hat{\mu} = \bar{X} - 1.$$

From $E\hat{\mu} = E\bar{X} - 1 = EX_1 - 1 = \mu$, we know that $\hat{\mu}$ is unbiased. $\square$

(3) Both $\hat{\mu}$ and μ^{**} are unbiased, so we only need to compare the variances of the two estimators, or equivalently, compare $Var(X_{(1)})$ with $Var(\bar{X}) = \frac{1}{n} Var X_1$. To simplify our calculation, assume $\mu = 0$ without loss of generality. (Since $X - \mu \sim Exp(1), Var(X_{(1)}) = Var(X_{(1)} - \mu), Var(X) = Var(X - \mu)$.) Then $X_1 \sim Exp(1)$ and

$$Var(\bar{X}) = \frac{1}{n}.$$

Since $EX_{(1)}^2 = \int_0^{+\infty} x^2 \cdot ne^{-nx} dx = \frac{2}{n^2}$,

$$Var(X_{(1)}) = EX_{(1)}^2 - (EX_{(1)})^2 = \frac{1}{n^2}.$$

Therefore, $\hat{\mu}^{**}$ is more efficient than $\hat{\mu}$ if $n \geq 2$, while they have the same efficiency if $n = 1$. $\square$

Remark: One could also prove that $X_{(1)}$ is sufficient and complete. Then from Lehmann-Scheff, $\hat{\mu}^{**}$ is the unique UMVUE and cannot be less efficient than $\hat{\mu}$.

[Wei] 3.41. From the property of Gamma distributions (you can use characteristic function to prove this), $\sum_{i=1}^{n} X_i \sim \Gamma(n\alpha, \lambda)$, $\frac{\bar{X}}{\alpha} \sim \Gamma(n\alpha, n\alpha\lambda)$. Therefore,

$$E\left(\frac{\bar{X}}{\alpha}\right) = \frac{n\alpha}{n\alpha\lambda} = \frac{1}{\lambda}, \quad Var\left(\frac{\bar{X}}{\alpha}\right) = \frac{n\alpha}{(n\alpha\lambda)^2} = \frac{1}{n\alpha\lambda^2}.$$

Then, check that the efficiency of $\frac{\bar{X}}{\alpha}$ is 1. The $p.d.f$ is $f(x|\lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$, $x > 0$. Since Gamma distributions are in the exponential distribution family, we can calculate the Fisher information by

$$I(\lambda) = -E_\lambda\left[\frac{\partial^2}{\partial\lambda^2} \log f(x|\lambda)\right] = \frac{\alpha}{\lambda^2}.$$

Thus the efficiency

$$e(\lambda) = \frac{(g'(\lambda))^2}{nI(\lambda)Var\left(\frac{\bar{X}}{\alpha}\right)} = 1,$$

which implies that $\frac{\bar{X}}{\alpha}$ is an effective estimation. □

Remark: 记由样本计算得到的 Fisher 信息为 $I_{\boldsymbol{x}}$, 由总体计算得到的 Fisher 信息为 I_X. 分布族为指数族时，才能用 $I_{\boldsymbol{x}} = nI_X$.

[Wei] 3.43. Let $W = \frac{1}{n}\sqrt{\frac{\pi}{2}} \sum_i |X_i - a|$, then (from Ex.3.8) $EW = \sigma$ (unbiased!), $Var(W) = \frac{\pi-2}{2n}\sigma^2$. Because $N(a, \sigma^2)$ is in exponential family, $I(\sigma) = -E_\sigma\left[\frac{\partial^2}{\partial\sigma^2} \log f(x|\sigma)\right] = \frac{2}{\sigma^2}$. (After time-consuming calculation......) $e(\sigma) = \frac{(\sigma')^2}{nI(\sigma)Var(W)} = \frac{1}{\pi-2}$. □

[Wei] 3.45. Let $W_1 = \bar{X}$, $W_2 = \frac{1}{2}(\max_i X_i + \min_i X_i)$. Notice that $Y_i = X_i - \theta \sim U[-\frac{1}{2}, \frac{1}{2}]$, then we have

$$EW_1 = EX_1 = \theta \text{ (unbiased!)}, \quad Var(W_1) = \frac{Var(X_1)}{n} = \frac{1}{12n}.$$

From symmetry,

$$\begin{aligned} EW_2 = \theta \text{ (unbiased!)}, \quad Var(W_2) &= \frac{1}{4} Var(Y_{(1)} + Y_{(n)}) \\ &= \frac{1}{4} E(Y_{(1)} + Y_{(n)})^2 \\ &= \frac{1}{4} \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{y} (x+y)^2 n(n-1)(y-x)^{n-2} dx dy \\ &= \frac{1}{2(n+1)(n+2)}. \end{aligned}$$

Thus, W_2 is more efficient if $n > 2$, and $Var(W_1) = Var(W_2)$ (same efficiency) if $n = 1$ or $n = 2$. □

[Wei] 3.46. Hints: write $p.d.f.$s of $X_{(1)}$ and $X_{(3)}$ and calculate their expectations you can find both $4X_{(1)}$ and $\frac{4}{3}X_{(3)}$ are unbiased.

Besides, an interesting observation is that $X_{(1)} \stackrel{d}{=} \theta - X_{(3)}$ (Calculate the p.d.f. of $\theta - X_{(3)}$ using density transformation formula from the p.d.f. of $X_{(3)}$ thus you can find $X_{(1)} \stackrel{d}{=} \theta - X_{(3)}$), which immediately implies that $Var(X_{(1)}) = Var(X_{(3)})$, thus $\frac{4}{3}X_{(3)}$ is more efficient.