数理统计作业

习题 1. Let $X_1, \cdots, X_n$ be a random sample from

$$f(x;\theta) = e^{-(x-\theta)} I_{[\theta,\infty)}(x) \quad \text{for } -\infty < \theta < \infty.$$

(a) Find a sufficient statistic.

(b) Find a maximum-likelihood estimator of θ.

(c) Find a method-of-moments estimator of θ.

(d) Is there a complete sufficient statistic? If so, find it.

(e) Find the UMVUE of theta if one exists.

(f) Find the Pitman estimator for the location parameter θ.

(g) Using the prior density $g(\theta) = e^{-\theta} I_{(0,\infty)}(\theta)$, find the posterior Bayes estimator of θ.

解. (a)

$$\begin{aligned} L(\theta) &= e^{n\theta} e^{-\sum_{i=1}^n x_i} \prod_{i=1}^n I_{[\theta,\infty)}(x_i) \\ &= e^{n\theta} e^{-\sum_{i=1}^n x_i} I_{(-\infty,y_1]}(x_i), \end{aligned}$$

Y_1 是充分统计量.

(b) 当 $x_i \geqslant \theta$ 时, $L(\theta) = e^{n\theta - \sum_{i=1}^n x_i}$, 则 θ 的 MLE 是 $\hat{\theta}_1 = Y_1$.

(c) 令

$$\begin{aligned} \bar{X} = \mathbb{E}(X) &= \int_\theta^\infty e^{-(x-\theta)} x \,\mathrm{d}x \\ &= \theta + 1, \end{aligned}$$

θ 的矩估计 $\hat{\theta}_2 = \bar{X} - 1$.

(d) 若对任意的统计量 $z(Y_1), \mathbb{E}[z(Y_1)] = 0$, 即

$$\int_\theta n e^{n\theta - ny_1} z(y_1) \,\mathrm{d}y_1 = 0,$$

关于 θ 求导, 得

$$-nz(\theta) + \int_\theta^\infty n^2 e^{n\theta - ny_1} z(y_1) \,\mathrm{d}y_1 = 0,$$

则对 $\forall\theta, z(\theta)=0$, 那么 $\mathbb{P}(z(Y_1)=0)=1$, Y_1 是完备充分统计量.

(e)

$$\mathbb{E}(Y_1)=\int_0^\infty ne^{n\theta-ny_1}y_1\,\mathrm{d}y_1=\frac{1}{n}+\theta,$$

$Y_1-\frac{1}{n}$ 是 θ 的无偏估计, 那么 $Y_1-\frac{1}{n}$ 是 θ 的 UMVUE.

(f) θ 的 Pitman 估计是

$$\begin{aligned}T_1(X_1,\cdots,X_n)&=\frac{\int\theta L(\theta)\,\mathrm{d}\theta}{\int L(\theta)\,\mathrm{d}\theta}\\&=\frac{\int_{-\infty}^{y_1}e^{n\theta-\sum_{i=1}^nX_i}\theta\,\mathrm{d}\theta}{\int_{-\infty}^{y_1}e^{n\theta-\sum_{i=1}^nX_i}\,\mathrm{d}\theta}\\&=Y_1-\frac{1}{n}.\end{aligned}$$

(g) θ 的 Bayes 后验估计是

$$\begin{aligned}T_2(X_1,\cdots,X_n)&=\frac{\int\theta L(\theta)g(\theta)\,\mathrm{d}\theta}{\int L(\theta)g(\theta)\,\mathrm{d}\theta}\\&=\frac{\int_0^{y_1}e^{(n-1)\theta-\sum_{i=1}^nX_i}\theta\,\mathrm{d}\theta}{\int_0^{y_1}e^{(n-1)\theta-\sum_{i=1}^nX_i}\,\mathrm{d}\theta}\\&=\frac{((n-1)Y_1-1)e^{(n-1)Y_1}+1}{(n-1)(e^{(n-1)Y_1}-1)}.\end{aligned}$$

习题 2. Let $X_1,\cdots,X_n$ be a random sample of size n from the following discrete density:

$$f(x;\theta)=\binom{2}{x}\theta^x(1-\theta)^{2-x}I_{\{0,1,2\}}(x),$$

where $\theta>0$.

(a) Is there a unidimensional sufficient statistic? If so, is it complete?

(b) Find a maximum-likelihood estimator of $\theta^2=\mathbb{P}(X_l=2)$. Is it unbiased?

(c) Find an unbiased estimator of θ whose variance coincides with the corresponding Cramér-Rao lower bound if such exists. If such an estimate does not exist, prove that it does not.

(d) Find a uniformly minimum-variance unbiased estimator of θ^2 if such exists.

(e) Using the squared-error loss function find a Bayes estimator of θ with respect to the beta prior distribution

$$g(\theta)=\frac{1}{\mathrm{Beta}(a,b)}\theta^{a-1}(1-\theta)^{b-1}I_{(0,1)}(\theta).$$

(f) Using the squared-error loss function, find a minimax estimator of θ.

(g) Find a mean-squared error consistent estimator of θ^2.

解. (a)

$$f(x;\theta)=(1-\theta)^2 I_{\{0,1,2\}}(x)\exp\left(x\log\frac{\theta}{1-\theta}\right),$$

则该分布属于指数族, 那么有完备充分统计量 $\sum_{i=1}^n X_i$, 令 $n=1$, 有完备充分统计量 X_1.

(b)

$$\begin{aligned}L(\theta)&=\prod_{i=1}^n\binom{2}{x_i}\theta^{x_i}(1-\theta)^{2-x_i}I_{\{0,1,2\}}(x_i)\\&\propto\theta^{\sum_{i=1}^n x_i}(1-\theta)^{2n-\sum_{i=1}^n x_i}.\end{aligned}$$

则令

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

得 $\hat\theta=\frac{\sum_{i=1}^n x_i}{2n}$, 又 $\theta>0$,

$$\widehat{\theta^2}=\frac{\left(\sum_{i=1}^n x_i\right)^2}{4n^2}.$$

因为 $\mathbb{E}\left(\sum_{i=1}^n X_i\right)^2=2n\theta+(4n^2-2n)\theta^2$, $\widehat{\theta^2}$ 不是无偏的.

(c) 首先

$$l(\theta)=\log\left[\prod_{i=1}^n\binom{2}{x_i}I_{\{0,1,2\}}(x_i)\right]+\left(\sum_{i=1}^n x_i\right)\log\theta+(2n-\sum_{i=1}^n x_i)\log(1-\theta),$$

则

$$\frac{\partial l}{\partial\theta}=\frac{\sum_{i=1}^n x_i}{\theta}-\frac{2n-\sum_{i=1}^n x_i}{1-\theta}=\frac{\left(\sum_{i=1}^n x_i\right)/2n-\theta}{\theta(1-\theta)/2n}.$$

则 $\frac{\sum_{i=1}^n X_i}{2n}$ 是达到 C-R 下界的无偏估计.

(d) 因为 $\mathbb{E}\left(\sum_{i=1}^n X_i\right)=2n\theta,\mathbb{E}\left(\sum_{i=1}^n X_i\right)^2=2n\theta+(4n^2-2n)\theta^2$, 那么

$$\mathbb{E}\left(\frac{\left(\sum_{i=1}^n X_i\right)^2-\sum_{i=1}^n X_i}{4n^2-2n}\right)=\theta^2,$$

即 $\frac{\left(\sum_{i=1}^n X_i\right)^2-\sum_{i=1}^n X_i}{4n^2-2n}$ 是 θ^2 的一个无偏估计, 又因为 $\sum_{i=1}^n X_i$ 是完备充分统计量, 则 $\frac{\left(\sum_{i=1}^n X_i\right)^2-\sum_{i=1}^n X_i}{4n^2-2n}$ 是 θ^2 是 UMVUE.

(e) 在均方误差损失函数下, θ 的 Bayes 估计是

$$\mathbb{E}(\theta \mid X_1=x_1,\cdots,X_n=x_n)=\frac{\int_0^\infty \theta L(\theta)g(\theta)\,\mathrm{d}\theta}{\int_0^\infty L(\theta)g(\theta)\,\mathrm{d}\theta}. \tag{1}$$

首先计算式 (1) 的分子,

$$\begin{aligned}
\int_0^\infty \theta L(\theta)g(\theta)\,\mathrm{d}\theta &= \prod_{i=1}^n \binom{2}{x_i} I_{\{0,1,2\}}(x_i)\int_0^\infty \theta\cdot\theta^{\sum_{i=1}^n x_i}(1-\theta)^{2n-\sum_{i=1}^n x_i}\\
&\quad\cdot\frac{1}{\mathrm{Beta}(a,b)}\theta^{a-1}(1-\theta)^{b-1}I_{(0,1)}(\theta)\,\mathrm{d}\theta\\
&=\prod_{i=1}^n \binom{2}{x_i} I_{\{0,1,2\}}(x_i)\cdot\frac{1}{\mathrm{Beta}(a,b)}\int_0^1 \theta^{a+\sum_{i=1}^n x_i+1-1}\\
&\quad\cdot(1-\theta)^{2n-\sum_{i=1}^n x_i+b-1}\,\mathrm{d}\theta\\
&=\prod_{i=1}^n \binom{2}{x_i} I_{\{0,1,2\}}(x_i)\cdot\frac{\mathrm{Beta}\left(a+\sum_{i=1}^n x_i+1,2n-\sum_{i=1}^n x_i+b\right)}{\mathrm{Beta}(a,b)}.
\end{aligned}$$

类似可得分母为

$$\int_0^\infty L(\theta)g(\theta)\,\mathrm{d}\theta=\prod_{i=1}^n \binom{2}{x_i} I_{\{0,1,2\}}(x_i)\frac{\mathrm{Beta}\left(\sum_{i=1}^n x_i+a,2n-\sum_{i=1}^n x_i+b\right)}{\mathrm{Beta}(a,b)}.$$

代入到式 (1), 得到 Bayes 估计为

$$T(X_1,\cdots,X_n)=\frac{\mathrm{Beta}\left(a+\sum_{i=1}^n X_i+1,2n-\sum_{i=1}^n X_i+b\right)}{\mathrm{Beta}\left(\sum_{i=1}^n X_i+a,2n-\sum_{i=1}^n X_i+b\right)}. \tag{2}$$

又因为

$$\begin{aligned}
\mathrm{Beta}\left(a+\sum_{i=1}^n X_i+1,2n-\sum_{i=1}^n X_i+b\right)&=\frac{\Gamma(a+\sum_{i=1}^n X_i+1)\Gamma(2n-\sum_{i=1}^n X_i+b)}{\Gamma(2n+a+b+1)},\\
\mathrm{Beta}\left(\sum_{i=1}^n X_i+a,2n-\sum_{i=1}^n X_i+b\right)&=\frac{\Gamma(\sum_{i=1}^n X_i+a)\Gamma(2n-\sum_{i=1}^n X_i+b)}{\Gamma(2n+a+b)},\\
\Gamma\left(a+\sum_{i=1}^n X_i+1\right)&=\left(a+\sum_{i=1}^n X_i\right)\Gamma\left(a+\sum_{i=1}^n X_i\right),\\
\Gamma(2n+a+b+1)&=(2n+a+b)\Gamma(2n+a+b),
\end{aligned}$$

将上式全部代入式 (2), 得到 θ 的 Bayes 估计为

$$T(X_1,\cdots,X_n)=\frac{a+\sum_{i=1}^n X_i}{2n+a+b}. \tag{3}$$

(f) 利用 (e) 的结果, 其中 a,b 是待定的. 首先计算 $\sum_{i=1}^n X_i$ 的分布, 因为对任意的 i, X_i 独立且 $X_i\sim B(2,\theta)$, 那么 $\sum_{i=1}^n X_i\sim B(2n,\theta)$, 根据式 (3),

$$\begin{aligned}\mathcal{R}_T(\theta)&=\mathbb{E}[(T-\theta)^2]\\&=\frac{1}{(2n+a+b)^2}\mathbb{E}\left(a-(a+b)\theta+\sum_{i=1}^n X_i-2n\theta\right)\\&=\frac{\operatorname{Var}\left(\sum_{i=1}^n X_i\right)+(a-(a+b)\theta)^2}{(2n+a^b)^2}\\&=\frac{2(n-a(a+b))\theta+((a+b)^2-2n)\theta^2+a^2}{(2n+a+b)^2},\end{aligned}$$

为了使上式等于常数, 令 $(a+b)^2=2n, a(a+b)=n$, 即 $a=b=\sqrt{\frac{n}{2}}$, 上式为常数, 即此时 T 是 minimax 估计.

(g) 因为 $X_i\sim B(2,\theta)$, $\mathbb{E}(X_i)=2\theta, \mathbb{E}(X_i^2)=2\theta+2\theta^2$, 则

$$\mathbb{E}\left(\frac{X_i^2-X_i}{2}\right)=\theta^2,$$

根据 Khinchin 大数定律,

$$\frac{\sum_{i=1}^n X_i^2-\sum_{i=1}^n X_i}{2n}\xrightarrow{p}\theta^2,$$

即 $\mathbb{P}\left(\left|\frac{\sum_{i=1}^n X_i^2-\sum_{i=1}^n X_i}{2n}-\theta^2\right|>\varepsilon\right)\to 0, n\to\infty$, 又 $0\leqslant\frac{\sum_{i=1}^n X_i^2-\sum_{i=1}^n X_i}{2n}\leqslant 1$, 则

$$\lim_{n\to\infty}\mathbb{E}\left(\frac{\sum_{i=1}^n X_i^2-\sum_{i=1}^n X_i}{2n}-\theta^2\right)^2=0,$$

即 $\frac{\sum_{i=1}^n X_i^2-\sum_{i=1}^n X_i}{2n}$ 是 θ^2 的均方相合估计量.