数理统计作业

习题 1 设 $X_1, \cdots .X_n$ 是来自 $N(0,1)$ 的随机样本, 令

$$\bar{X}_k = \frac{1}{k}\sum_{i=1}^{k} X_i, \quad \bar{X}_{n-k} = \frac{1}{n-k}\sum_{i=k+1}^{n} X_i.$$

求 (1)-(4) 的分布.(1) $\frac{1}{2}(\bar{X}_k + \bar{X}_{n-k})$.(2) $k\bar{X}_k^2 + (n-k)\bar{X}_{n-k}^2$.(3) X_1^2/X_2^2.(4) X_1/X_n.

解. (1) 因为 $X_i \sim N(0,1)$ 且相互独立, 则 $\bar{X}_k \sim N(0, 1/k), \bar{X}_{n-k} \sim N(0, 1/(n-k))$, 那么 $\frac{1}{2}(\bar{X}_k + \bar{X}_{n-k}) \sim N\left(0, \frac{n}{4k(n-k)}\right)$.

(2) $(n-k)\bar{X}_{n-k}^2 = \left(\sqrt{n-k}\bar{X}_{n-k}\right)^2$, 因为 $\sqrt{n-k}\bar{X}_{n-k} \sim N(0,1)$, 那么

$$\left(\sqrt{n-k}\bar{X}_{n-k}\right)^2 \sim \chi^2(1).$$

同理,$k\bar{X}_k^2 \sim \chi^2(1)$, 那么 $k\bar{X}_k^2 + (n-k)\bar{X}_{n-k}^2 \sim \chi^2(2)$.

(3) $X_1^2 \sim \chi^2(1)$ 且 $X_2^2 \sim \chi^2(1)$, 则

$$\frac{X_1^2}{X_2^2} \sim F(1,1).$$

(4) 设 X_1/X_n 的分布函数为 $F(x)$,X_1, X_n 的联合概率密度函数为 $p(x_1, x_n)$, 那么

$$\begin{aligned} F(x) &= \mathbb{P}\left(\frac{X_1}{X_n} < x\right) = \mathbb{P}\left(\frac{X_1}{X_n} < x, X_n > 0\right) + \mathbb{P}\left(\frac{X_1}{X_n} < x, X_n < 0\right) \\ &= \int_0^{+\infty} \mathrm{d}x_n \int_{-\infty}^{xx_n} p(x_1, x_n)\mathrm{d}x_1 + \int_{-\infty}^{0} \mathrm{d}x_n \int_{xx_n}^{+\infty} p(x_1, x_n)\mathrm{d}x_1. \end{aligned}$$

X_1, X_n 为独立的标准正态分布随机变量, 那么 X_1/X_n 的概率密度函数为

$$f(x) = \frac{\mathrm{d}F(x)}{\mathrm{d}x} = \frac{1}{x^2}\sqrt{\frac{2}{\pi}}.$$

习题 2 设 $X_1, \cdots, X_n$ 是来自双参数指数分布

$$f(x, \mu, \sigma) = \frac{1}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\}, \quad x \geqslant \mu$$

的简单随机样本, 其中 $-\infty < \mu < \infty, \sigma > 0$, 求 μ, σ 和 $\mathbb{P}(X_1 \geqslant t)(t > \mu)$ 的矩估计和极大似然估计.

解. 令 r 是正整数, 那么

$$
\begin{aligned}
\mathbb{E}(X^r) &= \int_{\mu}^{+\infty} x^r \frac{1}{\sigma} \exp\left\{-\frac{x-\mu}{\sigma}\right\} \\
&= \int_{0}^{+\infty} (x+\mu)^r \frac{1}{\sigma} \exp\left\{-\frac{x}{\sigma}\right\} \\
&= \sum_{k=0}^{r} \begin{pmatrix} r \\ k \end{pmatrix} \mu^{r-k}\sigma^k \int_0^{+\infty} \left(\frac{x}{\sigma}\right)^k \exp\left(-\frac{x}{\sigma}\right) \mathrm{d}\frac{x}{\sigma} \\
&= \sum_{k=0}^{r} \begin{pmatrix} r \\ k \end{pmatrix} \mu^{r-k}\sigma^k \Gamma(k+1) \\
&= r! \sum_{k=0}^{r} \frac{\mu^{r-k}\sigma^k}{(r-k)!}.
\end{aligned}
$$

那么要求 μ, σ^2 的矩估计, 即求

$$
\begin{cases}
\frac{X_1+\cdots+X_n}{n} = \mu + \sigma \\
\frac{X_1^2+\cdots+X_n^2}{n} = \mu^2 + 2\mu\sigma + 2\sigma^2
\end{cases}
$$

的解, 解为

$$
\begin{cases}
\hat{\sigma} = \sqrt{\frac{X_1^2+\cdots+X_n^2}{n} - \frac{(X_1+\cdots+X_n^2)^2}{n^2}} \\
\hat{\mu} = \frac{X_1+\cdots+X_n}{n} - \sqrt{\frac{X_1^2+\cdots+X_n^2}{n} - \frac{(X_1+\cdots+X_n^2)^2}{n^2}}
\end{cases}
$$

又因为

$$
\begin{aligned}
\mathbb{P}(X_1 \geqslant t) &= \int_t^{+\infty} \frac{1}{\sigma} \exp\left(-\frac{x-\mu}{\sigma}\right) \mathrm{d}x \\
&= \exp\left(-\frac{t-\mu}{\sigma}\right).
\end{aligned}
$$

那么 $\mathbb{P}(X_1 \geqslant t)$ 的矩估计为

$$
\mathbb{P}(\widehat{X_1 \geqslant} t) = \exp\left(-\frac{t - \frac{X_1+\cdots+X_n}{n} + \sqrt{\frac{X_1^2+\cdots+X_n^2}{n} - \frac{(X_1+\cdots+X_n^2)^2}{n^2}}}{\sqrt{\frac{X_1^2+\cdots+X_n^2}{n} - \frac{(X_1+\cdots+X_n^2)^2}{n^2}}}\right)
$$

$X_1, \cdots, X_n$ 的联合概率密度为

$$
L(\mu, \sigma) = \prod_{k=1}^{n} f(x_k, \mu, \sigma) = \prod_{k=1}^{n} \left(\frac{1}{\sigma} \exp\left(-\frac{x-\mu}{\sigma}\right)\right).
$$

取对数, 令

$$
I(\mu, \sigma) = \log L(\mu, \sigma) = \sum_{k=1}^{n} \left(-\log\sigma - \frac{x_k - \mu}{\sigma}\right).
$$

求一阶导和二阶导,

$$\nabla I = \left(\frac{n}{\sigma}, -\frac{n}{\sigma} + \frac{\sum_{k=1}^{n} x_k}{\sigma^2} + \frac{n\mu}{\sigma}\right),$$

$$\nabla^2 I = \begin{bmatrix} 0 & -\frac{n}{\sigma^2} \\ -\frac{n}{\sigma^2} & \frac{n}{\sigma^2} - \frac{2\sum_{k=1}^{n} x_k}{\sigma^3} + \frac{n\mu}{\sigma^3}. \end{bmatrix}$$

得到 $\sigma = +\infty$ 时取最大. 极大似然估计:

$$\hat{\mu} \in \mathbb{R}, \quad \hat{\sigma} = +\infty, \quad \mathbb{P}(\widehat{X_1 \geqslant t}) = 1.$$

习题 3 设 $X_1, \cdots, X_n$ 和 $Y_1, \cdots, Y_n$ 分别为来自正态分布总体 $N(\mu_1, \sigma^2)$ 和 $N(\mu_2, \sigma^2)$ 的随机样本, 求 μ_1, μ_2, σ^2 的 MLE.

解. $X_1, \cdots, X_n$ 的累计概率密度为

$$\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_i - \mu_1)^2}{2\sigma^2}\right).$$

取对数, 关于 μ, σ 求偏导, 令偏导为 0 得到方程

$$\begin{cases} n\mu_1 = \sum_{i=1}^{n} x_i \\ -n\sigma^2 + \sum_{i=1}^{n} (x_i - \mu_1)^2 = 0. \end{cases}$$

解, 得

$$\begin{cases} \hat{\mu}_1 = \frac{1}{n} \sum_{i=1}^{n} x_i \\ \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(x_i - \frac{1}{n} \sum_{i=1}^{n} x_i\right)^2. \end{cases}$$

同理可得

$$\begin{cases} \hat{\mu}_2 = \frac{1}{n} \sum_{i=1}^{n} y_i \\ \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(y_i - \frac{1}{n} \sum_{i=1}^{n} y_i\right)^2. \end{cases}$$

习题 4 设 $X_1, \cdots, X_n$ 来自均匀分布 $U(\theta, 2\theta)$ 的随机样本, 其中 $\theta \in (0, \infty)$, 求: (1) θ 的极大似然估计 $\hat{\theta}$;(2) 判断 $\hat{\theta}$ 是否为无偏估计, 如果不是无偏估计, 基于 $\hat{\theta}$ 构造一个无偏估计.

解. (1) X_i 的概率密度为

$$p(x) = \begin{cases} \frac{1}{\theta}, & \theta < x < \theta \\ 0, & x \leqslant \theta \text{或} x \geqslant \theta. \end{cases}$$

那么当 $(x_1,\cdots,x_n)\in(\theta,2\theta)^n$ 时,$X_1,\cdots,X_n$ 的联合密度函数为

$$f(x_1,\cdots,x_n,\theta)=\frac{1}{\theta^n},$$

其余情况下 $f(x_1,\cdots,x_n,\theta)=0$. 当 $(x_1,\cdots,x_n)\in(\theta,2\theta)$ 时,$\frac{1}{2}\max_i x_i\leqslant\theta\leqslant\min_i x_i$. 那么 θ 的极大似然估计为 $\hat{\theta}=\frac{1}{2}X_{(n)}$.

(2) $X_{(n)}$ 的分布函数为

$$\mathbb{P}(X_{(n)}<x)=\prod_{i=1}^{n}\mathbb{P}(X_i<x)=F(x)^n$$

其中 $F(x)$ 是 X_i 的分布函数, 那么可以得到

$$\begin{aligned}\mathbb{E}(\hat{\theta})&=\frac{1}{2}\int_{\theta}^{2\theta}\frac{nx(x-\theta)^{n-1}}{\theta^n}\mathrm{d}x\\&=\frac{2n+1}{2n+2}\theta\neq\theta.\end{aligned}$$

则它不是无偏估计.

根据上面的期望, 可知 $\frac{n+1}{2n+1}X_{(n)}$ 是一个无偏估计.

习题 5 设 X 服从对数正态分布, 即 $\log X\sim N(\mu,\sigma^2)$, 其中 $\mu\in\mathbb{R},0<\sigma<+\infty$. 设 $X_1,\cdots,X_n$ 来自总体 X 的随机样本, 求:(1) μ,σ^2 的极大似然估计 $\hat{\theta}$;(2) 求 $\mathbb{E}(X)$ 的极大似然估计.

解. (1) $\log X_1,\cdots,\log X_n$ 是 $\log X$ 的随机样本, 根据习题 3的结果, 可得极大似然估计

$$\begin{cases}\hat{\mu}=\frac{1}{n}\sum_{i=1}^{n}\log X_i\\ \hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}\left(\log X_i-\sum_{i=1}^{n}\log X_i\right)^2.\end{cases}$$

(2) 注意到 $X=e^{\log X}$, 那么

$$\begin{aligned}\mathbb{E}(X)&=\mathbb{E}(e^{\log X})\\&=\int_{-\infty}^{\infty}\exp(x)\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\mathrm{d}x\\&=\exp\left(\frac{\sigma^2+2\mu}{2}\right)\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-(\sigma^2+\mu))^2}{2\sigma^2}\right)\mathrm{d}x\\&=\exp\left(\frac{\sigma^2+2\mu}{2}\right).\end{aligned}$$

代入第 (1) 问的极大似然估计, 得到

$$\hat{\mathbb{E}}(X)=\exp\left(\frac{1}{2n}\sum_{i=1}^{n}\left(\log X_i-\sum_{i=1}^{n}\log X_i\right)^2+\frac{1}{n}\sum_{i=1}^{n}\log X_i\right).$$