# HOMEWORK 4

2021/10/20

**[Wei] 3.6.** Since $\alpha_1 = EX = kp$, $\mu_2 = VarX = kp(1-p)$, we derive

$$\begin{cases} p = 1 - \frac{\mu_2}{\alpha_1}, \\ k = \frac{\alpha_1^2}{\alpha_1 - \mu_2}. \end{cases}$$

Therefore, the MoMs of $k$ and $p$ are

$$\begin{cases} \hat{p}_{MoM} = 1 - \frac{S_n}{\bar{X}}, \\ \hat{k}_{MoM} = \frac{\bar{X}}{\hat{p}_{MoM}} = \frac{\bar{X}^2}{\bar{X} - S_n}, \end{cases}$$

where $\bar{X} = \frac{1}{n}\sum_i X_i$, $S_n = \frac{1}{n}\sum_i (X_i - \bar{X})^2$. $\square$

**[Wei] 3.7.** We calculate two moments of $X$ to give a simple expression of $p$.

$$\begin{aligned} \alpha_1 = EX &= \sum_{k=1}^{\infty} kP(X=k) \\ &= \sum_{k=1}^{\infty} -\frac{1}{\ln(1-p)} p^k \\ &= -\frac{p}{(1-p)\ln(1-p)}, \\ \alpha_2 = EX^2 &= \sum_{k=1}^{\infty} k^2 P(X=k) \\ &= \sum_{k=1}^{\infty} -\frac{1}{\ln(1-p)} kp^k \\ &= -\frac{p}{(1-p)^2 \ln(1-p)}. \end{aligned}$$

It is easy to observe that $\frac{\alpha_1}{\alpha_2} = 1 - p$, or equivalently,

$$p = 1 - \frac{\alpha_1}{\alpha_2}.$$

---

*Date*: 2021/11/3.

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.

Thus we can derive an MoM of $p$ as

$$\hat{p}_{MoM} = 1 - \frac{\hat{\alpha}_1}{\hat{\alpha}_2} = 1 - \frac{\sum_i X_i}{\sum_i X_i^2}.$$

□

**[Wei] 3.8.** (1) $\hat{\sigma}^{(1)}_{MoM} = \sqrt{\frac{\pi}{2}} \frac{1}{n} \sum_i |X_i|$.

(2) $\hat{\sigma}^{(2)}_{MoM} = \sqrt{\frac{1}{n} \sum_i (X_i - \bar{X})^2}$, where $\bar{X} = \frac{1}{n} \sum_i X_i$. □

**[Wei] 3.9.** Since $EX_1 = a$ and $Var(X_1) = \sigma^2$, we have MoMs of $a$ and $\sigma$ that

$$\begin{cases} \hat{a}_{MoM} = \bar{X}, \\ \hat{\sigma}_{MoM} = \sqrt{S_n}, \end{cases}$$

where $\bar{X} = \frac{1}{n} \sum_i X_i$, $S_n = \frac{1}{n} \sum_i (X_i - \bar{X})^2$.

Notice that $P(X > 1) = P\left(\frac{X-a}{\sigma} > \frac{1-a}{\sigma}\right) = \Phi\left(\frac{a-1}{\sigma}\right)$, we derive an MoM of $P(X > 1)$

$$\widehat{P(X > 1)}_{MoM} = \Phi\left(\frac{\hat{a}_{MoM} - 1}{\hat{\sigma}_{MoM}}\right) = \Phi\left(\frac{\bar{X} - 1}{\sqrt{S_n}}\right).$$

□

**[Wei] 3.10.** From $EX_1 = \frac{r}{\lambda}$, we derive an MoM of $\lambda$ that

$$\hat{\lambda}_{MoM} = \frac{r}{\bar{X}},$$

where $\bar{X} = \frac{1}{n} \sum_i X_i$.

To calculate $E\hat{\lambda}_{MoM}$, notice that $\bar{X} \sim \Gamma(nr, n\lambda)$, then

$$\begin{aligned} E\hat{\lambda}_{MoM} &= \int_0^{+\infty} \frac{r}{t} \frac{(n\lambda)^{nr}}{\Gamma(nr)} t^{nr-1} e^{-n\lambda t} dt \\ &= \frac{nr}{nr-1} \lambda \neq \lambda, \end{aligned}$$

which implies that $\hat{\lambda}_{MoM}$ is biased. □

Remark1: With correction, $\hat{\lambda}^*_{MoM} = \frac{nr-1}{nr} \hat{\lambda}_{MoM} = \frac{nr-1}{\sum_i X_i}$ is unbiased, and $\hat{\lambda}_{MoM}$ is asymptotically unbiased.

Remark2:Y$\sim \Gamma(\alpha, \beta)$, then $kY \sim \Gamma(\alpha, \beta/k), \forall k > 0$;$X_1, \ldots, X_n \sim \Gamma(\alpha, \beta)$ i.i.d., then $\sum_{i=1}^n X_i \sim \Gamma(n\alpha, \beta)$. You can check this using character function.

2021/10/22

**MLE.** 求解 MLE 的书写要求：

1. 可微情形：

(1) 对对数似然求导，令一阶导为 0 解得 $\theta^*$，并验证二阶导 (Hessian 阵)在 $\theta^*$ 处小于 0(负定)，并讨论 $\theta^*$ 在观测数据上的取值可能不是参数空间内点的情形，才能说是 MLE.

(2) 如果似然函数是自然参数形式的指数族，也可直接对对数似然求导，令一阶导为 0，解方程得到 MLE（解是自然参数空间内点时）.

2. 不可微情形：利用定义，使得 (对数) 似然最大。

3. 讨论 $\theta^*$ 在观测数据上的取值可能不是参数空间内点的情况：

记 $\theta^*$ 在观测数据上的取值为 $\theta_0$。课本做法是找一列样本 $X_n$,$\theta^*$ 在其上的取值 $\theta_n^* \in \Theta^0$，且 $lim_{n\to\infty}\theta_n^* = \theta_0$. 以 $P(\lambda)$ 为例，若在一组样本上得到 $\lambda_{MLE}$ 的值为 0，则找一列样本使得 $\lambda_n \to 0$。其想法可以理解为靠近 0 的估计值在 n 充分大时有可能会一直出现，从而说明真实情况是 $\lambda$ 很小。那么 $\lambda_{MLE}$ 在观测值上的取值为 0，对应的情况是 $\lambda$ 的真实值很小，而 $\lambda_{MLE}$ 的表达式仍是正确的。

**[Wei] 3.11.** (1)$X$ 的密度函数为

$$f_X(x) \overset{\xi:=\ln x\sim N(a,\sigma^2)}{=} f_\xi(\ln x)\frac{d\xi}{dx} = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(\ln x-a)^2}{2\sigma^2}}\frac{1}{x}, \quad x>0.$$

求得 X 的一、二阶矩如下

$$\begin{aligned}
\alpha_1 = EX &= Ee^\xi \\
&= \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{t-\frac{(t-a)^2}{2\sigma^2}}dt \\
&= \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(t-(a+\sigma^2))^2}{2\sigma^2}}e^{\frac{((a+\sigma^2)^2-a^2)}{2\sigma^2}}dt \\
&= e^{a+\frac{\sigma^2}{2}}, \\
\alpha_2 = EX^2 &= Ee^{2\xi} \\
&= \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{2t-\frac{(t-a)^2}{2\sigma^2}}dt \\
&= \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(t-(a+2\sigma^2))^2}{2\sigma^2}}e^{\frac{((a+2\sigma^2)^2-a^2)}{2\sigma^2}}dt \\
&= e^{2a+2\sigma^2}.
\end{aligned}$$

那么，

$$\begin{cases} a = 2\ln\alpha_1 - \frac{1}{2}\ln\alpha_2, \\ \sigma^2 = \ln\alpha_2 - 2\ln\alpha_1. \end{cases}$$

则 $a$ 和 $\sigma^2$ 的矩估计为：

$$\begin{cases} \hat{a}_{MoM} = 2\ln\bar{X} - \frac{1}{2}\ln\overline{X^2}, \\ \hat{\sigma^2}_{MoM} = \ln\overline{X^2} - 2\ln\bar{X}, \end{cases}$$

这里 $\bar{X} = \frac{1}{n}\sum_i X_i$, $\overline{X^2} = \frac{1}{n}\sum_i X_i^2$.

(2) 对数似然函数为：

$$l(a,\sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{\sum_i(\ln x_i - a)^2}{2\sigma^2} - \ln\left(\frac{1}{\prod x_i}\right), \quad x_i > 0, 1 \le i \le n.$$

由 $\frac{\partial l}{\partial a} = 0$, $\frac{\partial l}{\partial \sigma^2} = 0$, 得

$$a* = \frac{1}{n}\sum_i \ln X_i$$

$$(\sigma^2)^* = \frac{1}{n}\sum_i (\ln X_i - \hat{a}_{MLE})^2 = \frac{1}{n}\left(\sum_i (\ln X_i)^2 - \frac{1}{n}(\sum_i \ln X_i)^2\right)$$

此时 Hessian 阵

$$H|_{th\vec{e}ta=(a^*,(\sigma^2)^*)} = \begin{pmatrix} -\frac{n}{(\sigma^2)^*} & 0 \\ 0 & -\frac{2}{2((\sigma^2)^*)^2} \end{pmatrix} < 0$$

若 $(\sigma^2)^*(x_1,\dots,x_n) = 0$, 取值不在参数空间内。则考虑 $\vec{X}_m = (1,1,\dots,1,e)$(即 m 个观测值，前 m-1 个值为 1，第 m 个值为 e), 则 $(\sigma^2)^*_m > 0$, 且

$$\lim_{m\to\infty}(\sigma^2)^*_m = \lim_{m\to\infty}\frac{n-1}{n^2} = 0$$

综上，$\hat{a}_{MLE} = \frac{1}{n}\sum_i \ln X_i, \hat{\sigma^2}_{MLE} = \frac{1}{n}\left(\sum_i (\ln X_i)^2 - \frac{1}{n}(\sum_i \ln X_i)^2\right)$. □

**[Wei] 3.13.** (1) 总体的一二阶矩为

$$\begin{aligned}\alpha_1 = EX &= \int_{-\infty}^{+\infty} \frac{t}{2\sigma} e^{-\frac{|t-a|}{\sigma}} dt \\ &= \int_{-\infty}^{+\infty} \left(\frac{t-a}{2\sigma} + \frac{a}{2\sigma}\right) e^{-\frac{|t-a|}{\sigma}} dt \\ &= a, \\ \alpha_2 = EX^2 &= \int_{-\infty}^{+\infty} \frac{t^2}{2\sigma} e^{-\frac{|t-a|}{\sigma}} dt \\ &= \int_{-\infty}^{+\infty} \frac{(t-a+a)^2}{2\sigma} e^{-\frac{|t-a|}{\sigma}} dt \\ &= \int_{-\infty}^{+\infty} \left(\frac{(t-a)^2}{2\sigma} + \frac{a^2}{2\sigma}\right) e^{-\frac{|t-a|}{\sigma}} dt \\ &= a^2 + 2\sigma^2.\end{aligned}$$

那么

$$\begin{cases} a = \alpha_1, \\ \sigma = \sqrt{\frac{1}{2}(\alpha_2 - \alpha_1^2)} = \sqrt{\frac{1}{2}\mu_2}. \end{cases}$$

因此 $a$ 和 $\sigma$ 的矩估计为

$$\begin{cases} \hat{a}_{MoM} = \bar{X}, \\ \hat{\sigma}_{MoM} = \sqrt{\frac{1}{2}S_n}, \end{cases}$$

这里 $\bar{X} = \frac{1}{n}\sum_i X_i$, $S_n = \frac{1}{n}\sum_i (X_i - \bar{X})^2$.

(2) 对数似然为

$$l(a, \sigma) = -n\ln 2 - n\ln\sigma - \frac{\sum_i |x_i - a|}{\sigma}.$$

任意固定 $\sigma$, 最大化 $l(a) = l(a, \sigma)$ 等价于最小化 $\sum_{i=1}^n |x_i - a|$, 则 $\hat{a} = m_n$, $m_n$ 代表样本中位数，且 $n = 2k+1$ 时 $m_n = X_{(k+1)}$, $n = 2k$ 时 $m_n \in [X_{(k)}, X_{(k+1)}]$

(注：利用中位数的性质:以 $F_n$ 表示样本分布，$m_n$ 表示样本中位数，则 $F_n(m) \geq 1/2, \quad 1 - F_n(m-) \geq 1/2$. 容易验证 $m_n = argmin_a \sum_{i=1}^n |x_i - a|$)

令 $\frac{\partial l}{\partial \sigma} = 0$, 则 $\hat{\sigma} = \frac{1}{n}\sum_{i=1}^n |X_i - \hat{a}| = \frac{1}{n}\sum_{i=1}^n |X_i - m_n|$.

此时 $\frac{\partial^2 l}{\partial \sigma^2}|_{\hat{\sigma}} = -\frac{n}{\hat{\sigma}^2} < 0$.

若 $X_1, \ldots, X_n$ 不全相等，则 $\hat{\sigma} > 0$, 在参数空间内。否则考虑 $\vec{X}_{2k-1} = (1, \ldots, 1 + 1/k, \ldots, 1)$ (第 k 位取 1+1/k, 其余取 1), $\vec{X}_{2k} = (1, \ldots, 1+1/k, 1+1/k, \ldots, 1)$ (第 k 位和第 k+1 位取 1+1/k, 其余取 1) ,$k \geq 1$。则 $\sigma_{2k-1}, \sigma_{2k} > 0$, 且 $\lim_{k\to\infty} \hat{\sigma}_{2k-1} = \lim_{k\to\infty} \hat{\sigma}_{2k} = 0$

综上，极大似然估计为

$$\begin{cases} \hat{a}_{MLE} = m_n, \\ \hat{\sigma}_{MLE} = \frac{1}{n}\sum_{i=1}^n |X_i - m_n|, \end{cases}$$

$m_n$ 代表样本中位数。 □

**[Wei] 3.15.** (1) 总体的一二阶矩为

$$\begin{aligned} \alpha_1 = EX &= \int_\mu^{+\infty} \frac{t}{\sigma} e^{-\frac{t-\mu}{\sigma}} dt \\ &\overset{s:=\frac{t-\mu}{\sigma}}{=} \int_0^{+\infty} \left(s + \frac{\mu}{\sigma}\right) e^{-s} \sigma ds \\ &= \mu + \sigma, \\ \alpha_2 = EX^2 &= \int_\mu^{+\infty} \frac{t^2}{\sigma} e^{-\frac{t-\mu}{\sigma}} dt \\ &= \int_0^{+\infty} (\sigma s + \mu)^2 e^{-s} ds \\ &= 2\sigma^2 + 2\sigma\mu + \mu^2, \\ \mu_2 = Var(X) &= \alpha_2 - \alpha_1^2 = \sigma^2 \end{aligned}$$

那么

$$\begin{cases} \mu = \alpha_1 - \sigma, \\ \sigma = \sqrt{\mu_2}. \end{cases}$$

因此 $\mu$ 和 $\sigma$ 的矩估计为

$$\begin{cases} \hat{\mu}_{MoM} = \bar{X} - \sqrt{S_n}, \\ \hat{\sigma}_{MoM} = \sqrt{S_n}, \end{cases}$$

这里 $\bar{X} = \frac{1}{n}\sum_i X_i$, $S_n = \frac{1}{n}\sum_i (X_i - \bar{X})^2$.

(2) 对数似然函数为

$$l(\theta) = -n \ln \sigma - \frac{\sum_i x_i - n\mu}{\sigma}, \quad x_{(1)} \geq \mu.$$

固定 $\sigma$, 最大化 $l(\mu)$, 由定义取 $\hat{\mu}_{MLE} = X_{(1)}$. 由 $\frac{\partial l}{\partial \sigma} = 0$, 得 $\hat{\sigma} = \bar{X} - \hat{\mu}_{MLE} = \bar{X} - X_{(1)}$. 且 $\frac{\partial^2 l}{\partial \sigma^2}|_{\hat{\sigma}} = -\frac{n}{\hat{\sigma}^2} < 0$. （若 $X_i$ 全相等，可采取与之前类似的讨论，此处略过）. 故 $\hat{\sigma}_{MLE} = \bar{X} - X_{(1)}$.

(3) 由于 $P(X_1 \geq t) = e^{-\frac{t-\mu}{\sigma}}$, 则其矩估计和极大似然估计为

$$\widehat{P(X_1 \geq t)}_{MoM} = e^{-\frac{t-\hat{\mu}_{MoM}}{\hat{\sigma}_{MoM}}}.$$

$$\widehat{P(X_1 \geq t)}_{MLE} = e^{-\frac{t-\hat{\mu}_{MLE}}{\hat{\sigma}_{MLE}}}.$$

□

**[Wei] 3.21.** Suppose the ratio of black and white balls is $\theta \in [0,1]$. The likelihood function is

$$lik(\theta) = \binom{n}{k}\left(\frac{\theta}{\theta+1}\right)^{n-k}\left(\frac{1}{\theta+1}\right)^{k} = \binom{n}{k}\frac{\theta^{n-k}}{(\theta+1)^n}.$$

Thus the log-likelihood function is

$$l(\theta) = \ln\binom{n}{k} + (n-k)\ln\theta - n\ln(\theta+1).$$

If $k \neq 0$ or $n$, from $\frac{\partial l}{\partial\theta} = 0$, we obtain $\hat{\theta} = \frac{n-k}{k}$. Because $\frac{\partial^2 l}{\partial\theta^2}|_{\hat{\theta}} = k^2(\frac{1}{n} - \frac{1}{n-k}) < 0$, we have $\hat{\theta}_{MLE} = \frac{n-k}{k}$. If $k = 0$ or $n$, observe that $l$ reaches its maximum at $\theta = +\infty$ or $0$ respectively. In summary, if we denote $\frac{n}{0} := +\infty$, we have that $\hat{\theta}_{MLE} = \frac{n-k}{k}$ for all $k$. □