基科 32 曾柯又 2013012266

7.1 当 $x=0 \quad \hat{\theta}=1$

$x=1 \quad \hat{\theta}=1$

$x=2 \quad \hat{\theta}=2$或3

$x=3 \quad \hat{\theta}=3$

$x=4 \quad \hat{\theta}=3$

7.3 因为 $\log x$ 是严格单调的，且由 $L(\theta|\mathbf{x})$ 的定义可知其大于零，因此

$$L(\theta_1|\mathbf{x}) > L(\theta_2|\mathbf{x}) \Leftrightarrow \log L(\theta_1|\mathbf{x}) > \log L(\theta_2|\mathbf{x})$$

即

$$\forall \theta \neq \theta_0 \quad L(\theta|\mathbf{x}) < L(\theta_0|\mathbf{x}) \Leftrightarrow \forall \theta \neq \theta_0 \quad \log L(\theta|\mathbf{x}) < \log L(\theta_0|\mathbf{x})$$

7.10

(a) $f(x|\alpha,\beta) = \alpha \dfrac{x^\alpha}{\beta^\alpha} I(0 \leq x \leq \beta)$

因此

$$\begin{aligned} f(\mathbf{x}|\alpha,\beta) &= (\frac{\alpha}{\beta})^n (\prod_{i=1}^n x_i)^{\alpha-1} \prod_{i=1}^n I(0 \leq x_i \leq \beta) \\ &= (\frac{\alpha}{\beta})^n (\prod_{i=1}^n x_i)^{\alpha-1} I(x_{(n)} < \beta) I(x_{(1)} > 0) \end{aligned}$$

记 $t_1 = \displaystyle\prod_{i=1}^n x_i \quad t_2 = x_{(n)} \quad f(\mathbf{x}|\alpha,\beta) = (\frac{\alpha}{\beta})^n t_1^{\alpha-1} I(t_2 < \beta) I(x_{(1)} > 0)$

由因子分解定理 $T(\mathbf{X}) = (t_1, t_2)$, 为一个充分统计量。

(b) 记函数 $g(\alpha,\beta)=(\frac{\alpha}{\beta^{\alpha}})^n t_1^{\alpha-1} I(t_2\leq\beta)$ 首先对任意 α, 易看出 β 越小，g 越大，而 β 小于 t_2 时变为 0, 因此 $\hat{\beta}=t_2$

而对 $g(\alpha,\hat{\beta})=\dfrac{1}{\hat{\beta}^n}\alpha^n(\dfrac{t_1}{\hat{\beta}^n})^{\alpha-1}$

$\dfrac{\partial}{\partial\alpha}\log g(\alpha,\hat{\beta})=\dfrac{n}{\alpha}+\log(\dfrac{t_1}{\hat{\beta}}^n)$

$\because t_1<t_2^n=\hat{\beta}^n \ \Rightarrow\ \dfrac{t_1}{\hat{\beta}^n}<1$

且$\dfrac{n}{\alpha}+\log(\dfrac{t_1}{\hat{\beta}})^n$随$\alpha$单减, 故对应极值为极大值

$\therefore \hat{\alpha}=\dfrac{n}{\log(\frac{\hat{\beta}^n}{t_1})}=\dfrac{n}{\log(\frac{t_2^n}{t_1})}$

(c) $n=14 \quad x_{(n)}=25$ 即$\hat{\beta}=25$

$\hat{\alpha}=\dfrac{n}{n\log x_{(n)}-\log t_1}=\dfrac{1}{\log x_{(n)}-\frac{1}{n}\sum_{i=1}^n\log x_i}\approx 12.595$

7.11

(a) $L(\theta|\mathbf{x})=\theta^n(\prod_{i=1}^n x_i)^{\theta-1}$ 记 $t=\displaystyle\prod_{i=1}^n x_i$

$L(\theta|\mathbf{x})=\theta^n t^{\theta-1}$

$\dfrac{\partial\log L}{\partial\theta}=\dfrac{n}{\theta}+\log t$

$\because t<1$ 且$\dfrac{n}{\theta}+\log t$单调递减, 故极值为极大值，因此$\hat{\theta}=\dfrac{n}{\log(\frac{1}{t})}$

容易证明 $-\log X_i\sim Exp(\dfrac{1}{\theta})$ 故 $-\displaystyle\sum_{i=1}^n\log X_i\sim Gamma(n,\dfrac{1}{\theta})$

因此$\dfrac{\hat{\theta}}{n}\sim\dfrac{1}{x^2}\dfrac{\theta^n}{(n-1)!}(\dfrac{1}{x})^{n-1}e^{-\frac{\theta}{x}}$

$E\dfrac{\hat{\theta}}{n}=\dfrac{\theta}{n-1}$

$Var\dfrac{\hat{\theta}}{n}=E\Big(\dfrac{\hat{\theta}}{n}\Big)^2-\Big(E\dfrac{\hat{\theta}}{n}\Big)^2=\dfrac{\theta^2}{(n-1)(n-2)}-\dfrac{\theta^2}{(n-1)^2}=\dfrac{\theta^2}{(n-1)^2(n-2)}$

可得$Var\hat{\theta}=\dfrac{n^2\theta^2}{(n-1)^2(n-2)}\Rightarrow\ n\to\infty$ 时$Var\hat{\theta}\to 0$

(b) $EX_i = \dfrac{\theta}{\theta+1} \Rightarrow \dfrac{1}{n}\sum_{i=1}^{n} X_i = \dfrac{\tilde{\theta}}{\tilde{\theta}+1}$

$\tilde{\theta} = \dfrac{\bar{x}}{1-\bar{x}}$ 因为 $X_i \le 1 \ \Rightarrow \bar{x} \le 1$

又$P(\bar{x}=1)=0,\ \tilde{\theta} = \dfrac{\bar{x}}{1-\bar{x}} > 0$

7.12

(a) $EX=\theta$ 用矩估计的方法可得 $\tilde{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i$

用 MLE 估计:

$L(\theta|\mathbf{x}) = \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{n-\sum_{i=1}^{n} x_i}$

记 $t = \frac{1}{n}\sum_{i=1}^{n} x_i$

则 $L(\theta|\mathbf{x}) = (\theta^t(1-\theta)^{1-t})^n$

$\Rightarrow \dfrac{1}{n}\dfrac{\partial L(\theta|\mathbf{x})}{\partial\theta} = \dfrac{t}{\theta} - \dfrac{1-t}{1-\theta}$

该式随 θ 单减, 于是 $L(\theta|\mathbf{x})$ 在 $\theta=t$ 时取得极大值。但是考虑到 $\theta < \frac{1}{2}$, 故当 $t > \frac{1}{2}$ 时 $L(\theta|\mathbf{x})$ 在区间 $[0,\frac{1}{2}]$ 上单调递增,θ 只能取 $\frac{1}{2}$，$L(\theta|\mathbf{x})$ 最大，因此

$$\hat{\theta} = \begin{cases} \frac{1}{2} & t > \frac{1}{2} \\ t & t \le \frac{1}{2} \end{cases}$$

(b) 对于矩估计 $\tilde{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i$, 而 $\sum_{i=1}^{n} X_i \sim B(n,\theta)$, 因此

$E\tilde{\theta} = \frac{1}{n} \times n\theta = \theta$

$MSE(\tilde{\theta}) = Var\tilde{\theta} = \dfrac{1}{n^2}Var\sum_{i=1}^{n} X_i = \dfrac{\theta(1-\theta)}{n}$

对于 MLE 估计：

$$
\begin{aligned}
MSE(\hat{\theta}) &= E(\hat{\theta}-\theta)^2 \\
&= \sum_{k=1}^{n}(\hat{\theta}-\theta)^2\binom{n}{k}\theta^k(1-\theta)^{1-k} \\
&= \sum_{k=1}^{[\frac{n}{2}]}(\frac{k}{n}-\theta)^2\binom{n}{k}\theta^k(1-\theta)^{n-k} + \sum_{k=[\frac{n}{2}]+1}^{n}(\frac{1}{2}-\theta)^2\binom{n}{k}\theta^k(1-\theta)^{n-k}
\end{aligned}
$$

(c) $MSE(\tilde{\theta}) = \sum_{k=1}^{n}(\frac{k}{n}-\theta)^2\binom{n}{k}\theta^k(1-\theta)^{n-k}$ 因此

$$
\begin{aligned}
&MSE(\tilde{\theta}) - MSE(\hat{\theta}) = \\
&\qquad \sum_{k=[\frac{n}{2}]+1}^{n}\left((\frac{k}{n}-\theta)^2-(\frac{1}{2}-\theta)^2\right)\binom{n}{k}\theta^k(1-\theta)^{n-k} \\
&\qquad = \sum_{k=[\frac{n}{2}]+1}^{n}(\frac{k}{n}+\frac{1}{2}-2\theta)(\frac{k}{n}-\frac{1}{2})\binom{n}{k}\theta^k(1-\theta)^{n-k}
\end{aligned}
$$

$\because \theta < \dfrac{1}{2},\ \dfrac{k}{n} > \dfrac{1}{2}$

$\therefore MSE(\tilde{\theta}) - MSE(\hat{\theta}) > 0$

即最似然估计更好

7.13 $L(\theta|\mathbf{x}) = \frac{1}{2}e^{-\sum_{i=1}^{n}|x_i-\theta|}$

对于 $L' = \sum_{i=1}^{n}|x_i-\theta| = \sum_{i=1}^{n}|x_{(i)}-\theta|$

易知 $|x_{(i)}-\theta| + |x_{(j)}-\theta| \geq |x_{(i)}-x_{(j)}|$, 仅当 θ在$x_{(i)}, x_{(j)}$ 之间等号成立

因此当 n 为奇数时，容易知道当 $\hat{\theta} = m = x_{(\frac{n+1}{2})}$ 时，L' 取得最小，对应 L 取得最大当 n 为奇数时，有 $\hat{\theta} \in (x_{(\frac{n}{2})}, x_{(\frac{n}{2}+1)})$ 时,L' 取得最小，对应 L 取得极大值。

7.14

$$
\begin{aligned}
P(Z<z,W=1) &= P(X\le z, X\le Y)\\
&= \int_0^z \mathrm{d}x \int_x^\infty \mathrm{d}y \frac{1}{\lambda\mu} e^{-(\frac{x}{\lambda}+\frac{y}{\mu})}\\
&= \int_0^z \mathrm{d}x \frac{1}{\lambda} e^{-\frac{x}{\lambda}-\frac{x}{\mu}}\\
&= \frac{1}{1+\frac{\lambda}{\mu}}(1-e^{-(\frac{1}{\lambda}+\frac{1}{\mu})z})
\end{aligned}
$$

同理 $P(Z<z,W=0) = \frac{1}{1+\frac{\mu}{\lambda}}(1-e^{-(\frac{1}{\lambda}+\frac{1}{\mu})z})$

$f_{Z,W}(z,1) = \frac{\partial}{\partial z} P(Z<z,W=1) = \frac{1}{\lambda} e^{-(\frac{1}{\lambda}+\frac{1}{\mu})z}$

$f_{Z,W}(z,0) = \frac{\partial}{\partial z} P(Z<z,W=0) = \frac{1}{\mu} e^{-(\frac{1}{\lambda}+\frac{1}{\mu})z}$

$f_{Z,W}(z,w) = (\frac{1}{\lambda} e^{-(\frac{1}{\lambda}+\frac{1}{\mu})z})^w (\frac{1}{\mu} e^{-(\frac{1}{\lambda}+\frac{1}{\mu})z})^{1-w} = (\frac{1}{\lambda})^w (\frac{1}{\mu})^{1-w} e^{-(\frac{1}{\lambda}+\frac{1}{\mu})z} \Rightarrow$

$L(\lambda,\mu|(\mathbf{z},\mathbf{w})) = (\frac{1}{\lambda})^{\sum_{i=1}^n w_i} (\frac{1}{\mu})^{n-\sum_{i=1}^n w_i} e^{-(\frac{1}{\lambda}+\frac{1}{\mu})\sum_{i=1}^n z_i}$

记 $t_1 = \sum_{i=1}^n w_i,\ t_2 = \sum_{i=1}^n z_i,\ \theta_1 = \frac{1}{\lambda},\ \theta_2 = \frac{1}{\mu}$

则 $L(\lambda,\mu|(\mathbf{z},\mathbf{w})) = (\theta_1)^{t_1}(\theta_2)^{n-t_1} e^{-(\theta_1+\theta_2)t_2}$

$\log L = t_1 \log\theta_1 + (n-t_1)\log\theta_2 - (\theta_1+\theta_2)t_2$

$\frac{\partial \log L}{\partial \theta_1} = \frac{t_1}{\theta_1} - t_2$

$\frac{\partial \log L}{\partial \theta_2} = \frac{n-t_1}{\theta_2} - t_2$ 且两个偏倒数都随 θ_1，θ_2 单点递减，因此对应极值为极大值，即:

$\theta_1 = \frac{t_1}{t_2}\ \theta_2 = \frac{n-t_1}{t_2}$

$\Rightarrow \hat{\lambda} = \frac{\sum_{i=1}^n z_i}{\sum_{i=1}^n w_i} \quad \hat{\mu} = \frac{\sum_{i=1}^n z_i}{n-\sum_{i=1}^n w_i}$

7.18

(a) 为简化记号，设 $\mathbf{Z}=(X,Y),\mathbf{z}=(x,y),\mu_{\mathbf{Z}}=(\mu_X,\mu_Y)$

则 x,y 的分布为

$$f_{\mathbf{Z}}(\mathbf{z})=\frac{1}{2\pi\sqrt{|\Sigma|}}e^{-\frac{1}{2}((\mathbf{z}-\mu_{\mathbf{Z}})\Sigma^{-1}(\mathbf{z}-\mu_{\mathbf{Z}})^T)}$$

其中协方差矩阵为：$\begin{bmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y\\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{bmatrix}$

对 y 积分可以得到 x 的分布 $f_X(x)=\dfrac{1}{\sqrt{2\pi}\sigma_X}e^{-\frac{(x-\mu_X)^2}{\sigma_X^2}}$

于是 $EX=\mu_X,\ EX^2=VarX+(EX)^2=\sigma_X^2+\mu_X^2$

$\Rightarrow \tilde{\mu}_X=\bar{x},\ \tilde{\mu}_X+\tilde{\sigma}_X^2=\bar{x^2}$

$\Rightarrow \tilde{\sigma}_X^2=\bar{x^2}-\bar{x}^2=\frac{1}{n}\sum_{i=1}^n(x_i-\bar{x})^2$

同理对于 y, 有：

$\tilde{\mu}_Y=\bar{y},\ \tilde{\sigma}_Y^2=\frac{1}{n}\sum_{i=1}^n(y_i-\bar{y})^2$

而 $\rho=\dfrac{Cov(X,Y)}{\sqrt{VarX\,VarY}}=\dfrac{EXY-EXEY}{\sqrt{VarX\,VarY}}$, 得到:

$EXY=\sigma_X\sigma_Y\rho+\mu_X\mu_Y$

$\Rightarrow \tilde{\rho}\tilde{\sigma}_X\tilde{\sigma}_Y+\tilde{\mu}_X\tilde{\mu}_Y=\dfrac{1}{n}\displaystyle\sum_{i=1}^n x_iy_i=\dfrac{1}{n}\sum_{i=1}^n(x-\bar{x})(y-\bar{y})+\bar{x}\bar{y}$

$\Rightarrow \tilde{\rho}=\dfrac{\frac{1}{n}\sum_{i=1}^n(x_i-\bar{x})(y_i-\bar{y})}{\tilde{\sigma}_X\tilde{\sigma}_Y}$

(b) 不采用书中的办法，由于 Σ 是半正定对称矩阵, 于是存在一个正交矩阵 A 可将其对角化, 即 $A\Sigma A^T=\begin{pmatrix}\sigma_u^2 & 0\\ 0 & \sigma_v^2\end{pmatrix}$

记 $\mathbf{w}=(u,v)=\mathbf{z}A^T$, $\mu_{\mathbf{w}}=(\mu_u,\mu_v)=\mu_{\mathbf{Z}}A^T$ 于是 pdf 可写为

$$
\begin{aligned}
f_{\mathbf{Z}}(\mathbf{z}) &= \frac{1}{2\pi\sqrt{|\Sigma|}}e^{-\frac{1}{2}((\mathbf{w}-\mu_{\mathbf{w}})A\Sigma^{-1}A^T(\mathbf{w}-\mu_{\mathbf{w}})^T)}\\
&= \frac{1}{2\pi\sqrt{|\Sigma|}}e^{-\frac{1}{2}((\mathbf{w}-\mu_{\mathbf{w}})(A\Sigma A^T)^{-1}(\mathbf{w}-\mu_{\mathbf{w}})^T)}\\
&= \frac{1}{2\pi\sqrt{|\Sigma|}}e^{-\frac{1}{2}(\frac{(u-\mu_u)^2}{\sigma_u^2}+\frac{(v-\mu_v)^2}{\sigma_v^2})}
\end{aligned}
$$

一系列样本值 (x_i,y_i)，对应着一系列的 (u_i,v_i), 似然函数:

$$
\begin{aligned}
L &= \prod_{i=1}^{n}\frac{1}{2\pi\sqrt{|\Sigma|}}e^{-\frac{1}{2}(\frac{(u_i-\mu_u)^2}{\sigma_u^2}+\frac{(v_i-\mu_v)^2}{\sigma_v^2})}\\
&= (\frac{1}{2\pi\sqrt{|\Sigma|}})^n e^{-\frac{1}{2}\sum_{i=1}^{n}(\frac{(u_i-\mu_u)^2}{\sigma_u^2}+\frac{(v_i-\mu_v)^2}{\sigma_v^2})}\\
&= (\frac{1}{2\pi\sqrt{|\Sigma|}})^n e^{-\frac{1}{2}\sum_{i=1}^{n}(\frac{(u_i-\bar{u})^2}{\sigma_u^2}+\frac{(v_i-\bar{v})^2}{\sigma_v^2}+\frac{(\bar{u}-\mu_u)^2}{\sigma_u^2}+\frac{(\bar{v}-\mu_v)^2}{\sigma_v^2})}\\
&= (\frac{1}{2\pi\sqrt{|\Sigma|}})^n e^{-\frac{n}{2}(\bar{\mathbf{z}}-\mu_{\mathbf{z}})\Sigma^{-1}(\bar{\mathbf{z}}-\mu_{\mathbf{z}})^T}\exp(-\frac{1}{2}\sum_{i=1}^{n}tr(\begin{pmatrix}\frac{1}{\sigma_u^2} & 0\\ 0 & \frac{1}{\sigma_v^2}\end{pmatrix}(\mathbf{w}_i-\bar{\mathbf{w}})^T(\mathbf{w}_i-\bar{\mathbf{w}}))\\
&= (\frac{1}{2\pi\sqrt{|\Sigma|}})^n e^{-\frac{n}{2}(\bar{\mathbf{z}}-\mu_{\mathbf{z}})\Sigma^{-1}(\bar{\mathbf{z}}-\mu_{\mathbf{z}})^T}\exp(-\frac{1}{2}\sum_{i=1}^{n}tr(\Sigma^{-1}(\mathbf{z}_i-\bar{\mathbf{z}})^T(\mathbf{z}_i-\bar{\mathbf{z}}))
\end{aligned}
$$

设 $\mathbf{S}=\sum_{i=1}^{n}(\mathbf{z}_i-\bar{\mathbf{z}})^T(\mathbf{z}_i-\bar{\mathbf{z}})$

有

$$L=(\frac{1}{2\pi\sqrt{|\Sigma|}})^n e^{-\frac{n}{2}(\bar{\mathbf{z}}-\mu_{\mathbf{z}})\Sigma^{-1}(\bar{\mathbf{z}}-\mu_{\mathbf{z}})^T}\exp(-\frac{1}{2}tr(\Sigma^{-1}\mathbf{S}))$$

考虑到协方差矩阵的性质，对指数上第一项 $(\bar{\mathbf{z}}-\mu_{\mathbf{z}})\Sigma^{-1}(\bar{\mathbf{z}}-\mu_{\mathbf{z}})^T$, 仅当 $\hat{\mu}_{\mathbf{Z}}=\bar{\mathbf{z}}$ 时取得 0，对应 L 为极大。

再考虑 $L'=(\frac{1}{\sqrt{|\Sigma|}})^n\exp(-\frac{1}{2}tr(\Sigma^{-1}\mathbf{S}))$

$\log L'=-\frac{n}{2}\log|\Sigma|-\frac{1}{2}tr(\Sigma^{-1}\mathbf{S})=\frac{n}{2}\log|\Sigma^{-1}|-\frac{1}{2}tr(\Sigma^{-1}\mathbf{S})$

对Σ^{-1}的每一分量求导，结果写为矩阵形式，有

$\frac{\partial}{\partial\Sigma^{-1}}\log L'=\frac{n}{2}\Sigma-\frac{1}{2}\mathbf{S}$

该方程只有一个解，表明极值只有一个，并且$|\Sigma|\to\infty$时$L'\to 0$,而当$\Sigma=$

$\frac{1}{n}\mathbf{S}$时，L'为正,故该极值为极大值因此$\hat{\Sigma} = \frac{1}{n}\mathbf{S}$

若将矩阵写为分量形式即与 (a) 问中得到的结果相同。