1 第四次作业

10.29(b)

由式 (10.2.9) 我们知道, 一个 $M-$ 估计量绝不可能比极大似然估计量更有效. 但是我们知道什么时候它同样有效.

(a) 证明: $\psi(x-\theta)=cl'(\theta\mid x) \iff M-$ 估计量是渐近有效的. 其中 l 是对数似然, 而 c 是常数.

(b) 对于下列分布, 验证相应的 ψ 函数给出渐近有效的 M$-$ 估计量.

(1) 正态: $f(x)=\mathrm{e}^{-x^2/2}/(\sqrt{2\pi}), \psi(x)=x$

(2) 罗吉斯蒂克: $f(x)=\mathrm{e}^{-x}/\left(1+\mathrm{e}^{-x}\right)^2, \psi(x)=\tanh(x)$, 其中 $\tanh(x)$ 是双曲正切

(3) Cauchy: $f(x)=\left[\pi\left(1+x^2\right)\right]^{-1}, \psi(x)=2x/\left(1+x^2\right)$

(4) 最小信息分布:

$$f(x)=\begin{cases}\mathrm{Ce}^{-x^2/2} & |x|\le c\\ \mathrm{Ce}^{-c|x|+c^2/2} & |x|>c\end{cases}$$

$\psi(x)=\max\{-c,\min(c,x)\}, C$ 和 c 都是常数.

解:

注意到 (a) 问的结论，所以只需要验证 $\psi(x-\theta)=cl'(\theta\mid x)$ 这个条件即可.

由计算可知，(1)(3)(4) 均满足.

对于 (2)，满足 $\psi(\frac{x-\theta}{2})=l'(\theta\mid x)$，所以，确实如群友所说，题目打错了，应该是 $\psi(x)=\tanh\frac{x}{2}$.

注记 1: 有些同学直接算的渐近方差与 Fisher 信息量也是可以的（这也怪我，是我在群里群里说的思路，我的锅），不过有些同学在计算过程中发现 (2) 的积分不好算，事实上只需要注意换元 $t=exp(x-\theta)$，就化成分式的无穷积分了，然后用一下“部分分式”的技巧就可以算出来了。

注记 2: 关于 (a) 问，书上并没有说明条件的“充要”性，我这里改了一下。关于证明也不难，注意到《统计推断》中文版的 10.2 节的最后，452-453 页，关于“M-估计不可能比 MLE 估计更有效”的证明过程，最后

证明 ARE≤ 1 的时候，用的是柯西不等式，

$$\operatorname{ARE}\left(\hat{\theta}_{\mathrm{M}}, \hat{\theta}\right)=\frac{\left[\mathrm{E}_\theta \psi\left(X-\theta_0\right) l^{\prime}(\theta \mid X)\right]^2}{\mathrm{E}_\theta \psi(X-\theta)^2 \mathrm{E}_\theta l^{\prime}(\theta \mid X)^2} \leq 1$$

而柯西不等式的取等条件就是线性相关，也就是差一个乘积常数 c.

10.31

从多个总体收集到的二项数据常常用列联表表现出来. 在两个总体的情形, 列联表形如

总体

	1	总和	
成功	S_1	S_2	$S=S_1+S_2$
失败	F_1	F_2	$F=F_1+F_2$
总和	n_1	n_2	$n=n_1+n_2$

其中总体 1 是 binomial (n_1, p_1), 有 S_1 个成功, F_1 个失败; 总体 2 是 binomial (n_2, p_2), 有 S_2 个成功, F_2 个失败. 经常感兴趣的一个假设是

$$H_0: p_1=p_2 \text{ 对} H_1: p_1 \neq p_2 .$$

(a) 说明可以基于统计量

$$T=\frac{\left(\hat{p}_1-\hat{p}_2\right)^2}{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)(\hat{p}(1-\hat{p}))}$$

进行检验, 其中 $\hat{p}_1=S_1 / n_1, \hat{p}_2=S_2 / n_2, \hat{p}=\left(S_1+S_2\right) /\left(n_1+n_2\right)$. 另外, 证明当 n_1, $n_2 \rightarrow \infty$ 时, T 的分布趋于 χ_1^2.
(b) 测量与 H_0 相违背的另一个方法是计算期望频数表. 这个表的构造方法是, 在给定边缘总和的条件下, 根据 $H_0: p_1=p_2$ 填充表格, 即期望频数用这个期望频数表中的所有格子, 计算统计量 T^* :

$$\begin{aligned}
T^* & =\sum \frac{(\text{观测频数 }- \text{ 期盟频数})^2}{\text{期望频数}} \\
& =\frac{\left(S_1-\frac{n_1 S}{n_1+n_2}\right)^2}{\frac{n_1 S}{n_1+n_2}}+\cdots+\frac{\left(F_2-\frac{n_2 F}{n_1+n_2}\right)^2}{\frac{n_2 F}{n_1+n_2}}
\end{aligned}$$

用代数运算证明 $T^*=T$, 因此 T^* 是渐近 $\mathcal{X}^2$ 的.
(c) 检验 p_1 和 p_2 相等的另一个统计量是

$$T^{**}=\frac{\hat{p}_1-\hat{p}_2}{\sqrt{\frac{\hat{p}_1\left(1-\hat{p}_1\right)}{n_1}+\frac{\hat{p}_2\left(1-\hat{p}_2\right)}{n_2}}},$$

证明, 在 H_0 下, T^{**} 是渐近 n(0,1) 的, 因此, 其平方渐近于 $\mathcal{X}_1^2$. 进一步, 证明 $(T^{**})^2 \neq T^*$.
(d) 在什么情况下一个统计量比另一个更好?
(e) 检验石炭酸的使用是否与患者死亡有关.

解:

(a) 在 H_0 成立时, 由 CLT, $\sqrt{n_1}(\hat{p}_1 - p) \xrightarrow{d} N(0, p(1-p))$, $\sqrt{n_2}(\hat{p}_2 - p) \xrightarrow{d} N(0, p(1-p))$, 又由于 S_1 与 S_2 独立, (虽然这个条件没有明确指出来, 但是我们在两组不同的试验时, 总是认为这两组试验之间是独立的.)
所以由正态的线性性以及正态的依分布收敛对于两组独立的随机变量序列具有可加性 (注意, 假如不独立是很难证明的!), 所以

$$\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})p(1-p)}} \xrightarrow{d} N(0, p(1-p)) \xrightarrow{d} N(0,1)$$

再由弱大数律以及 Slutsky 定理, 命题得证.
至于"可以基于 T 进行检验", 首先它不是似然比检验统计量, 也不太像贝叶斯检验统计量, 所以言之有理即可. 比如, T 对于 $(p_1 - p_2)^2$ 这个量还算是一个比较"好"的估计, 因此可以用 T 来检验假设 H_0.
(b) 计算即可, 略.
(c) 注意到: 当 H_0 成立时, 由弱大数律, $\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2} \xrightarrow{P} \left(\frac{1}{n_1} + \frac{1}{n_2}\right) p(1-p)$, 再由 $f(x) = \sqrt{x}$ 的连续性, 以及 (a) 与 Slutsky 定理即可.
对于后面一部分, 只需证明, $\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2} \neq \left(\frac{1}{n_1} + \frac{1}{n_2}\right) \hat{p}(1-\hat{p})$. 所以只需要带入计算即可.
(d) 本题为开放题, 言之有理即可. 一般还是倾向于认为, T 更好.
(e) 用 (a) 的统计量进行检验, 算出 $T \approx 8.495$, 在检验水平 $\alpha = 0.05$ 时, 拒绝域为 $(3.841, \infty)$, 所以拒绝 H_0, 故有充分理由表明: 石炭酸的使用与患者死亡有关.

注: 本题的检验问题, 背景可以参考一下教材《数理统计》中 6.5.2 节"列联表中的齐一性检验".

10.33

修改并完成定理 10.3.1 的证明. 用定理 10.1.12 和 Slutsky 定理（定理 5.5.17）证明 $(\theta-\hat{\theta})\sqrt{-l''(\hat{\theta}\mid x)} \rightarrow N(0,1)$, 从而 $-2\log\lambda(X) \rightarrow \chi_1^2$.

解:

题目确实打错了，应该是乘、不是除。至于错误的起源，是定理 10.3.1 的证明过程:

证明: 首先在 $\hat{\theta}$ 的邻域展开 $\log L(\theta\mid x)=l(\theta\mid x)$ 为 Taylor 级数, 有

$$l(\theta\mid x)=l(\hat{\theta}\mid x)+l'(\hat{\theta}\mid x)(\theta-\hat{\theta})+l''(\hat{\theta}\mid x)\frac{(\theta-\hat{\theta})^2}{2}+\cdots.$$

现在把 $l\left(\theta_0\mid x\right)$ 的展开式代人 $-2\log\lambda(x)=-2l\left(\theta_0\mid x\right)+2l(\hat{\theta}\mid x)$ 中, 得到

$$-2\log\lambda(x)\approx\frac{(\theta-\hat{\theta})^2}{-l''(\hat{\theta}\mid x)},$$

显然，上面的等式应该是 $-2\log\lambda(x)\approx -l''(\hat{\theta}\mid x)(\theta-\hat{\theta})^2$.

至于如何证明 $(\theta-\hat{\theta})\sqrt{-l''(\hat{\theta}\mid x)} \rightarrow N(0,1)$，这是不难的:

首先，我们回忆《统计推断》定理 10.1.12 的证明过程:

回忆 $l(\theta\mid\boldsymbol{x})=\sum\log f\left(x_i\mid\theta\right)$ 是对数似然函数. 把其导数（关于 θ 的）记作 $l', l'', \cdots$. 现在真值 θ_0 的周围展开对数似然的一阶导数, (10.1.4) $l'(\theta\mid\boldsymbol{x})=l'\left(\theta_0\mid\boldsymbol{x}\right)+\left(\theta-\theta_0\right)l''\left(\theta_0\mid\boldsymbol{x}\right)+\cdots$, 这里, 我们将忽略其高阶项 (在正则性条件下这个手法是正当的). 现在用 $\hat{\theta}$ 替换 θ, 并看到等式 (10.1.4) 的左边是 0 . 重新整理此式并且乘以 $\sqrt{n}$, 就给出 (10.1.5)

$$\sqrt{n}\left(\hat{\theta}-\theta_0\right)=\sqrt{n}\frac{-l'\left(\theta_0\mid\boldsymbol{x}\right)}{l''\left(\theta_0\mid\boldsymbol{x}\right)}=\frac{-\frac{1}{\sqrt{n}}l'\left(\theta_0\mid\boldsymbol{x}\right)}{\frac{1}{n}l''\left(\theta_0\mid\boldsymbol{x}\right)}.$$

如果我们用 $I\left(\theta_0\right)=E\left[l'\left(\theta_0\mid\boldsymbol{X}\right)\right]^2=1/u(\theta)$ 来记关于一个观测的信息数, 应用中心极限定理和弱大数定律 (细节见习题 10.8）就将证明出 (10.1.6)

$$-\frac{1}{\sqrt{n}}l'\left(\theta_0\mid\boldsymbol{X}\right)\xrightarrow{L}\mathrm{n}\left(0, I\left(\theta_0\right)\right),$$
$$\frac{1}{n}l''\left(\theta_0\mid\boldsymbol{X}\right)\xrightarrow{p}I\left(\theta_0\right).$$

这样, 如果我们设 $W\sim N\left[0, I\left(\theta_0\right)\right]$, 则 $\sqrt{n}\left(\hat{\theta}-\theta_0\right)$ 依分布收敛到 $W/I\left(\theta_0\right)\sim N\left[0,1/I\left(\theta_0\right)\right]$, 定理证毕.

由上面的证明，我们可以知道，$(\theta-\hat{\theta})\sqrt{-l''(\theta\mid \boldsymbol{x})}\to N(0,1)$

又因为 MLE 的相合性，所以 $\hat{\theta}\xrightarrow{P}\theta$，且 $\sqrt{-l''(\theta\mid \boldsymbol{x})}$ 关于 θ 在一定范围内是连续函数，因此 $\sqrt{-l''(\hat{\theta}\mid \boldsymbol{x})}\xrightarrow{P}\sqrt{-l''(\theta\mid \boldsymbol{x})}$.

再由 Slutsky 定理，命题得证.

注：这个题的证明思路比较简单，所以请不要跳步，一步步的把证明过程写清楚。