数理统计 week 10

学业辅导中心

4.7.4 考察源自豌豆的两种交叉类型的遗传学问题. 孟德尔理论认为,(a)圆粒且黄色的;(b)皱纹且黄色的;(c)圆粒且绿色的;(d)皱纹且绿色的豌豆的概率分别为 9/16, 3/16, 3/16 以及 1/16. 若对 160 个独立观察体进行观测,各类发生的频率分别是 86, 35, 26 以及 13. 这些数据与孟德尔理论一致吗? 也就是在 $\alpha=0.01$ 时,对四组概率分别是 9/16, 3/16, 3/16 以及 1/16 的假设进行检验.

注记

- 说明了高中阶段的遗传推断并不严谨.
- 有人怀疑 Mendal 的实验数据过于完美.Weeden N. F. (2016). Are Mendel's Data Reliable? The Perspective of a Pea Geneticist. The Journal of heredity, 107(7), 635–646. https://doi.org/10.1093/jhered/esw058

习题 4.7.4

- 原假设: $H_0: p_1 = \frac{9}{16}, p_2 = \frac{3}{16}, p_3 = \frac{3}{16}, p_4 = \frac{1}{16}$
- 备择假设: 原假设中的概率不成立.
- 自由度: df = 4-1 = 3.

$$
\begin{aligned}
Q_3 &= \sum_{i=1}^{3} \frac{(X_i - np_i)^2}{np_i} \\
&= \left(\left\{ \frac{\left\{6 - 160\left(\frac{9}{16}\right)\right\}^2}{160\left(\frac{9}{16}\right)} + \frac{\left\{35 - 160\left(\frac{3}{16}\right)\right\}^2}{160\left(\frac{3}{16}\right)} \right\} + \left\{ \frac{\left\{26 - 160\left(\frac{3}{16}\right)\right\}^2}{160\left(\frac{9}{16}\right)} + \frac{\left\{13 - 160\left(\frac{1}{16}\right)\right\}^2}{160\left(\frac{9}{16}\right)} \right\} \right) \\
&= \frac{(86-90)^2}{90} + \frac{(35-30)^2}{30} + \frac{(26-30)^2}{30} + \frac{(13-10)^2}{10} \\
&= 0.178 + 0.833 + 0.533 + 0.900 \\
&= 2.444
\end{aligned}
$$

4.7.6 设随机实验结果可划分为一种互斥且穷尽状态 A_1，A_2，A_3，而且也可分成另一种互斥且穷尽状态 B_1，B_2，B_3，B_4．200 次独立实验结果的试验导致下述数据：

	B_1	B_2	B_3	B_4
A_1	10	21	15	6
A_2	11	27	21	13
A_3	6	19	27	24

在 5% 显著性水平上，对属性 A 与属性 B 是独立的假设进行检验，也就是 H_0 ：$P(A_i \cap B_j) = P(A_i)P(B_j)$，$i=1$，2，3 与 $j=1$，2，3，4，而备择假设是 A 与 B 不是相互独立的.

习题 4.7.6

独立性卡方检验.

- $Q = 12.942$
- $Q \sim \chi^2(6)$
- $\chi^2_{0.95}(6) = 12.592$
- 拒绝

4.7.7 某种遗传模型表明，特定三项分布的概率分别是 $p_1=p^2$，$p_2=2p(1-p)$，$p_3=(1-p)^2$，其中 $0<p<1$. 若 X_1，X_2，X_3 表示 n 次独立试验的不同属性的频率，请解释如何检证此遗传模型是合适的

- 方法: 卡方检验.
- 需要注意的细节: 估计参数 p, 调整检验的自由度.

4.7.9 有人建议用下面的数据拟合泊松分布：

x	0	1	2	3	$3<x$
频率	20	40	16	18	6

(a) 计算相应的卡方拟合度统计量.

提示：计算均值时，将 $3<x$ 处理成 $x=4$.

(b) 与这个卡方有关的自由度是多少呢？

(c) 在 $\alpha=0.05$ 显著性水平上，这些数据会导致对泊松模型的否认吗？

习题 4.7.9

先用极大似然法估计参数.

$$
\begin{aligned}
\bar{x} &= \frac{1}{n}\sum_{x=0}^{4} x f_x \\
&= \frac{(0)(20)+(1)(40)+(2)(16)+(3)(18)+(4)(6)}{20+40+16+18+6} \\
&= \frac{1}{100}(0+40+32+54+24) \\
&= 1.5
\end{aligned}
$$

习题 4.7.9

$X = x$	$P(X = x) = \frac{e^{-1.5}(1.5)^x}{x!}$
$X = 0$	0.22313
$X = 1$	0.334695
$X = 2$	0.251021
$X = 3$	0.125511
$P(X \geq 4)$	0.065642
Total	1

$$\chi^2 = \sum_{0}^{x} \frac{(f_0 - f_e)^2}{f_e}$$

习题 4.7.9

$$\chi^2 = \frac{(20-22.31302)^2}{22.31302} + \frac{(40-33.46952)^2}{33.46952} + \cdots + \frac{(6-6.564245)^2}{6.564245}$$
$$= 7.2286$$
$$\approx 7.229$$

自由度

$$\begin{aligned} df &= k-p-1 \\ &= 5-1-1 \\ &= 3 \end{aligned}$$

5.1.5 设 X_1，…，X_n 是 iid 随机变量，具有共同 pdf

$$f(x)=\begin{cases}e^{-(x-\theta)}, & x>\theta-\infty<\theta<\infty \\ 0, & 其他\end{cases} \tag{5.1.3}$$

这个 pdf 称为**位移指数**(shifted exponential)．设 $Y_n=\min\{X_1, \cdots, X_n\}$．通过获得 Y_n 的 cdf 与 pdf，证明 Y_n 依概率收敛 $Y_n \to \theta$.

习题 5.1.5

$$
\begin{aligned}
F_X(x) &= P(X \leq x) \\
&= \int_\theta^x f(x)dx \\
&= \int_\theta^x e^{-(x-\theta)}dx \\
&= 1 - e^{-(x-\theta)}
\end{aligned}
$$

根据位移指数分布的定义,

$$Y_n = \min\{X_1, X_2, \ldots, X_n\} > \theta$$

于是

$$
\begin{aligned}
\Pr(|Y_n - \theta| \leq \varepsilon) &= \Pr(Y_n - \theta \leq \varepsilon) \\
&= \Pr(Y_n \leq \varepsilon + \theta) \\
&= F_{Y_n}(\varepsilon + \theta) \\
&= 1 - e^{-n(\varepsilon+\theta-\theta)}
\end{aligned}
$$

5.1.7 对于习题 5.1.5，求 Y_n 的均值．Y_n 是 θ 的无偏估计量吗？以 Y_n 为基础，求 θ 的无偏估计量．

- Y_n 的分布函数:$F_{Y_n}(t) = 1 - \left(e^{\theta - t}\right)^n, t > \theta$
- Y_n 的密度函数:

$$\begin{aligned} f_{Y_n}(t) &= [F_{Y_n}(t)]' \\ &= \frac{\partial}{\partial t}\left\{1 - \left(e^{\theta - t}\right)^n\right\} \\ &= -e^{n\theta}(-n)e^{-nt} \\ &= ne^{n(\theta - t)}, \quad t > \theta \end{aligned}$$

习题 5.1.7

$$
\begin{aligned}
E\left(Y_n\right) &= \int_{\theta}^{\infty} t f_{Y_n}(t) dt \\
&= \int_{\theta}^{\infty} t n e^{n(\theta - t)} dt \\
&= n e^{n\theta} \left[-\frac{t}{n} e^{-nt} \Big|_{\theta}^{\infty} + \frac{1}{n} \int_{\theta}^{\infty} e^{-nt} dt \right] \\
&= n e^{n\theta} \left[\left(\frac{\theta}{n} e^{-n\theta} \right) - \frac{1}{n^2} \left(-e^{-n\theta} \right) \right] \\
&= \frac{1 + n\theta}{n}
\end{aligned}
$$

一个无偏估计量是 $Y_n - \dfrac{1}{n}$.

5.2.3 设 Y_n 表示来自下面连续型分布的随机样本 n 的最大值，此分布具有 cdf $F(x)$ 与 pdf $f(x)=F'(x)$. 求 $Z_n=n[1-F(Y_n)]$的极限分布.

提示

类似书上例 5.2.4, 直接用分布函数求出极限分布.

习题 5.2.3

$$
\begin{aligned}
G_{z_n}(t) &= \Pr\left(n\left(1 - F\left(Y_n\right)\right) \le t\right) \\
&= \Pr\left(1 - F\left(Y_n\right) \le \frac{t}{n}\right) \\
&= 1 - \Pr\left(F\left(Y_n\right) \le 1 - \frac{t}{n}\right) \\
&= 1 - \Pr\left(Y_n \le F^{-1}\left(1 - \frac{t}{n}\right)\right) \\
&= 1 - \left[F\left(F - \left(1 - \frac{t}{n}\right)\right)\right] \\
&= 1 - \left(1 - \frac{t}{n}\right)^n, 0 \le t \le n
\end{aligned}
$$

极限分布,

$$
G(t) = \lim_{\mathrm{a} \to \infty} G_{\mathrm{Z}}(t) = \begin{cases} 1 - e^{-t}, & 0 < t < \infty \\ 0 & , \text{ elsewhere} \end{cases}
$$

5.2.14 设 $\overline{X}_n$ 表示来自样本量为 n 的参数 $\mu=1$ 泊松分布随机样本均值.

(a) 证明：$Y_n=\sqrt{n}(\overline{X}_n-\mu)/\sigma=\sqrt{n}(\overline{X}_n-1)$的 mgf 是由 $\exp[-t/\sqrt{n}+n(e^{t/\sqrt{n}}-1)]$给出的.

(b) 研究当 $n\to\infty$时 Y_n 的极限分布.

提示：用它的麦克劳林级数代替表达式 $e^{t/\sqrt{n}}$，它是 Y_n的 mgf 指数.

5.2.15 利用习题 5.2.14，求$\sqrt{n}(\sqrt{\overline{X}_n}-1)$的极限分布.

注记

用矩母函数, 证明了题目中所述场合下的中心极限定理.

习题 5.2.14

先求矩母函数.

$$\begin{aligned}
M_{Y_n}(t) &= E\left(e^{tY_n}\right) \\
&= E\left(\exp\left(t\sqrt{n}\left(\bar{X}_{\mathrm{n}} - 1\right)\right)\right) \\
&= e^{-t\sqrt{n}} M_{X_n}(t\sqrt{n})
\end{aligned}$$

根据

$$\begin{aligned}
M_{\bar{X}_n}(t) &= E\left(e^{\overline{\bar{X}_n}}\right) \\
&= E\left(\exp\left(\frac{t}{n}(X_1 + X_2 + \cdots + X_n)\right)\right) \\
&= M_{X_1}\left(\frac{t}{n}\right) M_{X_2}\left(\frac{t}{n}\right) \cdots M_{X_n}\left(\frac{t}{n}\right) \\
&= \left[\exp\left[\mu\left(e^{t/n} - 1\right)\right]\right]^n \\
&= \exp\left[n\left(e^{t/n} - 1\right)\right]
\end{aligned}$$

于是,

$$M_Y(t) = e^{-t\sqrt{n}} \exp\left[n\left(e^{\sqrt{n}in} - 1\right)\right]$$
$$= \exp\left[-t\sqrt{n} + n\left(e^{t\sqrt{n}} - 1\right)\right]$$

取极限,

$$\lim_{n\to\infty} M_{Y_n}(t) = \lim_{n\to\infty} \exp\left[-t\sqrt{n} + n\left(e^{t/\sqrt{n}} - 1\right)\right]$$
$$= \lim_{n\to\infty} \exp\left[-t\sqrt{n} + n\left(e^{t/\sqrt{n}} - 1\right)\right]$$
$$= \lim_{n\to\infty} \exp\left[-t\sqrt{n} + n\left(\frac{t}{\sqrt{n}} + \frac{t^2}{2!n} + \frac{t^3}{3!n^{3/2}} + \cdots\right)\right]$$
$$= \lim_{n\to\infty} \exp\left[\frac{t^2}{2!} + \frac{t^3}{3!n^{1/2}} + \cdots\right]$$
$$= \exp\left(\frac{t^2}{2}\right)$$

习题 5.2.15

用 δ 方法, 取

$$g(t) = \sqrt{t}$$

g 在 $\theta = 1$ 处可导且导数值非零. 根据 δ 方法,

$$\sqrt{n}\left(\sqrt{\bar{X}_n} - 1\right) \xrightarrow{D} N\left(0, [g'(1)]^2\right)$$

$$\Rightarrow \sqrt{n}\left(\sqrt{\bar{X}_n} - 1\right) \xrightarrow{D} N\left(0, \left[\frac{1}{2\sqrt{1}}\right]^2\right)$$

于是,

$$\sqrt{n}\left(\sqrt{\bar{X}_n} - 1\right) \xrightarrow{D} N\left(0, \frac{1}{4}\right)$$

5.3.11 我们知道，对于很大的 n，$\overline{X}$ 近似服从 $N(\mu, \sigma^2/n)$. 求 $u(\overline{X})=\overline{X}^3$，$u\neq 0$ 的近似分布.

注记

使用 δ 方法.

习题 5.3.11

取 $u(t) = t^3$, u 在 $\mu \neq 0$ 时, 导数值非零. 于是

$$\sqrt{n}\left(\bar{X}^3 - \mu^3\right) \sim N\left(0, 9\mu^4\sigma^2\right)$$

从而,

$$\bar{X}^3 \overset{D}{\sim} N\left(\mu^3, \frac{9\mu^4\sigma^2}{n}\right)$$

6.1.2 设 X_1，X_2，…，X_n 表示来自于具有下面 pdf 或 pmf 分布的随机样本：

(a) $f(x;\theta)=\theta x^{\theta-1}$，$0<x<1$，$0<\theta<\infty$，其他为 0.

(b) $f(x;\theta)=e^{-(x-\theta)}$，$\theta\leqslant x<\infty$，$-\infty<\theta<\infty$，其他为 0.

在每一种情况下，求 θ 的极大似然估计量 $\hat{\theta}$.

括号 1

似然函数:

$$
\begin{aligned}
L(\theta) &= \prod_{i=1}^{n} f\left(x_i\right) \\
&= \prod_{i=1}^{n} \theta x_i^{\theta-1} \\
&= \theta^n \prod_{i=1}^{n} x_i^{\theta-1}
\end{aligned}
$$

解对数似然方程:

$$\frac{\partial \log(L(\theta))}{\partial\theta} = 0$$

$$\frac{\partial}{\partial\theta}\left(n\log\theta + (\theta-1)\sum_{i=1}^{n}\log(x_i)\right) = 0$$

$$\frac{n}{\theta} + \sum_{i=1}^{n}\log(x_i) = 0$$

$$\theta = \frac{-n}{\sum_{i=1}^{n}\log(x_i)}$$

于是, 极大似然估计量就是

$$\hat{\theta} = \frac{-n}{\sum_{i=1}^{n}\log(x_i)}$$

括号 2

似然函数:

$$L(\theta)=\prod_{i=1}^{n} f\left(x_i\right)=\prod_{i=1}^{n} e^{-\left(x_i-\theta\right)}=e^{-\sum_{i=1}^{n} x_i+n\theta}$$

对对数似然函数求导:

$$\frac{\partial}{\partial\theta}\left(-\sum_{i=1}^{n} x_i+n\theta\right)=n>0$$

似然函数是一个增函数, 由于 $\theta\leq\min\{X_i\}_{i=1}^{n}$, 因此, 极大似然估计是

$$\begin{aligned}\hat{\theta}&=Y_{(1)}\\&=\min\left(X_1,\cdots,X_n\right)\end{aligned}$$

6.1.4 假定 X_1，…，X_n 是 iid 的，具有 pdf $f(x;\theta)=2x/\theta^2$，$0<x\leqslant\theta$，其他为 0，求：

(a) θ 的极大似然估计量 $\hat{\theta}$.

(b) 常数 c 以使 $E(c\hat{\theta})=\theta$.

(c) 此分布中位数的极大似然估计量.

$$L(\theta)=\frac{2^n}{\theta^{2n}}\prod_{i=1}^{n}X_i\cdot 1(X_{(n)}\leq\theta)$$

$$l(\theta)=n\ln 2-2n\ln\theta+\sum_{i=1}^{n}\ln X_i$$

极大似然估计: $Y_n = \max\{X_1, \cdots, X_n\}$

$$P(Y_n \le y) = [P(X_i \le y)]^n$$

$$\mathbb{E}\hat{\theta}_n = \int_0^{\theta} \frac{2n}{\theta^{2n}} y^{2n+1}\, dy = \frac{2n}{2n+1}\theta$$

$$\int_0^{M} \frac{2x}{\theta^2} dx = \frac{1}{2} \Rightarrow M = \frac{\theta}{\sqrt{2}}$$

6.1.8 设下表

x	0	1	2	3	4	5
频率	7	14	12	13	6	3

表示来自于泊松分布样本量为 55 的随机样本总结情况. 求 $P(X=2)$的极大似然估计值.

$$\bar{x}=2.109;\ \bar{x}^2\mathrm{e}^{-\bar{x}}/2.$$

6.1.12 设随机样本 $X_1, X_2, \cdots, X_n$ 具有的分布有如下两种 pdf 形式. 若 $\theta=1$, 则 $f(x;\theta=1)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, $-\infty<x<\infty$. 若 $\theta=2$, 则 $f(x;\theta=2)=1/[\pi(1+x^2)]$, $-\infty<x<\infty$. 求 θ 的极大似然估计量.

$$f(x,\theta=1)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}, -\infty<x<\infty$$

$$f(x,\theta=2)=\frac{1}{[\pi(1+x^2)]}, -\infty<x<\infty$$

习题 6.1.12

$\theta = 1$ 时的似然函数是

$$L(X;\theta=1) = \prod_{i=1}^{n} f(x_i,\theta=1)$$
$$= (2\pi)^{-n/2} e^{\frac{-\sum_{i=1}^{n} x_i^2}{2}}$$

$\theta = 2$ 时的似然函数是

$$L(X;\theta=2) = \prod_{i=1}^{n} f(x_i,\theta=2)$$
$$= (\pi)^{-n} \times \frac{1}{\prod_{i=1}^{n}\left(1+x_i^2\right)}$$

哪个大选哪个.

6.2.2 已知 $f(x;\theta)=1/\theta$，$0<x<\theta$，其他为 0，正式计算

$$nE\left\{\left[\frac{\partial\log f(X;\theta)}{\partial\theta}\right]^2\right\}$$

的倒数，并将该值与 $(n+1)Y_n/n$ 的方差进行比较，其中 Y_n 表示来自这一分布样本量为 n 的随机样本的最大观测值. 请给予评论.

$$f_{Y_n}(y_n)=\frac{n{y_n}^{n-1}}{\theta^n}$$

思考

为什么会出现这样的结果？(super-efficiency)

习题 6.2.2

$$nE\left\{\left[\frac{\partial \ln f(X,\theta)}{\partial\theta}\right]^2\right\} = nE\left(\frac{1}{\theta}\right)^2 = \frac{n}{\theta^2}$$

对比 Y_n,

$$EY_n = \int_0^\theta n{y_n}^n/\theta^n dy_n = \frac{n}{n+1}\theta$$

$$EY_n^2 = \int_0^\theta n y_n^{n+1}/\theta^n dy_n = \frac{n}{n+2}\theta^2$$

$$\mathrm{Var}\,(Y_n) = \frac{n}{(n+2)(n+1)^2}\theta^2$$

因此,

$$\mathrm{Var}\left(\frac{n+1}{n}Y_n\right) = \left(\frac{n+1}{n}\right)^2\frac{n}{(n+2)(n+1)^2}\theta^2$$

$$= \frac{1}{(n+2)n}\theta^2$$

6.2.9 如果 X_1，X_2，…，X_n 是来自下述分布的随机样本，此分布具有 pdf

$$f(x;\theta)=\begin{cases}\dfrac{3\theta^3}{(x+\theta)^4}, & 0<x<\infty, 0<\theta<\infty \\ 0, & \text{其他}\end{cases}$$

证明 $Y=2\overline{X}$ 是 θ 的无偏估计量，并且确定它的有效性.

习题 6.2.9: 无偏性

$$
\begin{aligned}
E(X) &= 3\theta^3 \int_\theta^\infty \frac{t-\theta}{t^4} dt \\
&= 3\theta^3 \int_\theta^\infty \left\{t^{-3} - \theta t^{-4}\right\} dt \\
&= 3\theta^3 \int_\theta^\infty t^{-3} dt - 3\theta^4 \int_\theta^\infty t^{-4} d \\
&= 3\theta^3 \left(\frac{t^{-2}}{-2}\right)_\theta^\infty - 3\theta^4 \left(\frac{t^{-3}}{-3}\right)_\theta^\infty \\
&= 3\theta^3 \left(\frac{0-\theta^{-2}}{-2}\right) - 3\theta^4 \left(\frac{0-\theta^{-3}}{-3}\right) \\
&= \frac{3}{2}\theta - \theta \\
&= \frac{\theta}{2}
\end{aligned}
$$

习题 6.2.9: 有效性

$$l(\theta) = \frac{\partial L}{\partial \theta} = \frac{3}{\theta} - \frac{4}{x+\theta}$$

先计算 CR 下界.

$$\begin{aligned} E\left(l(\theta)^2\right) &= E\left(\frac{3}{\theta} - \frac{4}{x+\theta}\right)^2 \\ &= E\left(\frac{9}{\theta^2} + \frac{16}{(X+\theta)^2} - \frac{24}{\theta(X+\theta)}\right) \\ &= \frac{9}{\theta^2} + 16E\left(\frac{1}{(X+\theta)^2}\right) - \frac{24}{\theta}E\left(\frac{1}{X+\theta}\right) \end{aligned}$$

于是

$$\begin{aligned} I_n(\theta) &= n \times E\left(l(\theta)^2\right) \\ &= n\left(\frac{9}{\theta^2} + \frac{48}{5\theta^2} - \frac{18}{\theta^2}\right) = \frac{3n}{5\theta^2} \end{aligned}$$

再算估计量的方差.

$$
\begin{aligned}
E\left(X^2\right) &= 3\theta^3 \int_\theta^\infty \frac{(t-\theta)^2}{t^4} dt \\
&= 3\theta^3 \int_\theta^\infty \frac{t^2+\theta^2-2\theta t}{t^4} dt \\
&= 3\theta^3 \int_\theta^\infty \left\{t^{-2}+\theta^2 t^{-4}-2\theta t^{-3}\right\} dt \\
&= 3\theta^3 \int_\theta^\infty t^{-2} dt + 3\theta^5 \int_\theta^\infty t^{-4} dt - 6\theta^4 \int_\theta^\infty t^{-3} dt \\
&= 3\theta^3 \left(\frac{t^{-1}}{-1}\right)_\theta^\infty + 3\theta^5 \left(\frac{t^{-3}}{-3}\right)_\theta^\infty - 6\theta^4 \left(\frac{t^{-2}}{-2}\right)_\theta^\infty \\
&= 3\theta^3 \left(\frac{0-\theta^{-1}}{-1}\right) + 3\theta^5 \left(\frac{0-\theta^{-3}}{-3}\right) - 6\theta^4 \left(\frac{0-\theta^{-2}}{-2}\right) \\
&= 3\theta^2 + \theta^2 - 3\theta^2 \\
&= \theta^2
\end{aligned}
$$

习题 6.2.9: 有效性

$$\begin{aligned}\mathrm{Var}(Y) &= \mathrm{Var}(2\bar{X}) \\ &= (2)^2\,\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\ &= \frac{3\theta^2}{n}\end{aligned}$$

相对有效性,

$$\begin{aligned}e &= \frac{1/I_n(\theta)}{\mathrm{Var}(Y)} \\ &= \frac{5\theta^2}{3n} \times \frac{1}{3\theta^2/n} \\ &= \frac{5}{9} \approx 0.5556\end{aligned}$$

6.2.10 设 $X_1, X_2, \cdots, X_n$ 是来自分布 $N(0, \theta)$ 的随机样本. 我们希望估计标准差 $\sqrt{\theta}$. 求常数 c, 以使 $Y = c\sum_{i=1}^{n}|X_i|$ 成为 $\sqrt{\theta}$ 的无偏估计量, 并确定它的有效性.

习题 6.2.10

$$
\begin{aligned}
E\left[|X_i|\right] &= 2\sqrt{\theta}\int_0^{\infty}\frac{1}{\sqrt{2\pi\theta}}\exp\left\{-\frac{x^2}{2\theta}\right\}xdx \\
&= 2\sqrt{\theta}\int_0^{\infty}\frac{1}{\sqrt{2\pi\theta}}\exp(-z)dz\theta \\
&= 2\sqrt{\theta}\int_0^{\infty}\frac{1}{\sqrt{2\pi}}\exp(-z)dz \\
&= \frac{2\sqrt{\theta}}{\sqrt{2\pi}}\left(\int_0^{\infty}\exp(-z)dz\right) \\
&= \frac{2\sqrt{\theta}(\sqrt{2})}{\sqrt{2}(\sqrt{\pi})(\sqrt{2})} \\
&= \frac{\sqrt{2}}{\sqrt{\pi}}\sqrt{\theta}
\end{aligned}
$$

于是 $c = \sqrt{\frac{2}{\pi}}$.

习题 6.2.10

$$
\begin{aligned}
Var\left[\sqrt{\frac{\pi}{2}}|X_i|\right] &= E\left\{\sqrt{\frac{\pi}{2}}|X_i|\right\}^2 - \left\{E\left(\sqrt{\frac{\pi}{2}}|X_i|\right)\right\}^2 \\
&= E\left\{\frac{\pi}{2}X_i^2\right\} - \left\{\frac{\pi}{2}\left(E\left(|X_i|\right)\right)^2\right\} \\
&= \frac{\pi}{2}\left\{E\left(X_i^2\right) - \left(E\left(|X_i|\right)\right)^2\right\} \\
&= \frac{\pi}{2}\left\{\theta - \frac{2}{\pi}\theta\right\} \\
&= \frac{\pi}{2}\left\{\theta\left(1 - \frac{2}{\pi}\right)\right\} \\
&= \frac{\pi}{2} \times \frac{2}{\pi}\left\{\theta\left[\frac{\pi}{2} - 1\right]\right\} \\
&= \theta\left[\frac{\pi}{2} - 1\right]
\end{aligned}
$$

习题 6.2.10

$$
\begin{aligned}
Var(Y) &= \frac{1}{n}\left\{\theta\left[\frac{\pi}{2}-1\right]\right\} \\
&= \frac{\theta}{n}\left[\frac{\pi}{2}-1\right]
\end{aligned}
$$

$$
\begin{aligned}
I(\sqrt{\theta}) &= -E\left[\frac{\partial^2[f(x;\theta)]}{\partial\beta^2}\right] \\
&= E\left[\frac{1}{\theta}-3\frac{X^2}{\theta^2}\right] \\
&= -E\left[\frac{1}{\theta}\right]+E\left[3\frac{X^2}{\theta^2}\right] \\
&= -\frac{1}{\theta}+3\frac{1}{\theta^2}E\left(X^2\right) \\
&= -\frac{1}{\theta}+\frac{3\theta}{\theta^2}=\frac{2}{\theta}
\end{aligned}
$$

习题 6.2.10

$$
\begin{aligned}
e(Y) &= \frac{\frac{\theta}{2n}}{\frac{\theta}{n}\left[\frac{\pi}{2}-1\right]} \\
&= \frac{\frac{\theta}{2}}{2\left(\frac{\theta}{2}\right)\left[\frac{\pi}{2}-1\right]} \\
&= \frac{1}{2\left[\frac{\pi}{2}-1\right]} \\
&= \frac{1}{\pi-2}
\end{aligned}
$$

6.2.11 设 $\overline{X}$ 是来自分布 $N(\theta, \sigma^2)$的样本量为 n 的随机样本均值，$-\infty<\theta<\infty$，$\sigma^2>0$. 假定 σ^2 是已知的. 证明，$\overline{X}^2-\frac{\sigma^2}{n}$是 θ^2 的无偏估计量，并求它的有效性.

习题 6.2.11

$$
\begin{aligned}
&E\left(\bar{X}^2 - \frac{\sigma^2}{n}\right)\\
&= E\left(\bar{X}^2\right) - E\left(\frac{\sigma^2}{n}\right)\\
&= \mathrm{Var}(\bar{X}) + E(\bar{X})^2 - E\left(\frac{\sigma^2}{n}\right)\\
&= \frac{\sigma^2}{n} + (\theta)^2 - \frac{\sigma^2}{n}\\
&= \theta^2
\end{aligned}
$$

习题 6.2.11

$$
\begin{aligned}
I(\theta) &= -E\left[\frac{\partial^2 \log L}{\partial t^2}\right] \\
&= \frac{n}{4\sigma^2\theta^2}
\end{aligned}
$$

$$
e = \frac{4\sigma^2\theta^2}{n}\left(\frac{1}{Var\left(\bar{X}^2\right)}\right)
$$

注记

唯一需要注意的是: 考虑的是谁的 Fisher 信息.

6.2.14 设 S^2 是来自 $N(\mu, \theta)$ 的样本量为 $n>1$ 随机样本的样本方差 S^2，其中 μ 是已知的. 我们知道 $E(S^2)=\theta$.

(a) S^2 的有效性如何？

(b) 在这些条件下，θ 的似然估计量 $\hat{\theta}$ 是什么？

(c) $\sqrt{n}(\hat{\theta}-\theta)$ 的渐近分布是什么？

习题 6.2.14

$$\mathrm{Var}\left[\frac{(n-1)S^2}{\theta}\right] = 2(n-1)$$

$$\Rightarrow \mathrm{Var}\left(S^2\right) = \frac{2\theta^2}{n-1}$$

$$\begin{aligned}
I(\theta) &= E\left(l(\theta)^2\right) \\
&= \frac{1}{4\theta^2} + \frac{E\left(X^4\right)}{4\theta^4} - \frac{E\left(X^2\right)}{2\theta^3} \\
&= \frac{1}{4\theta^2} + \frac{3\theta^2}{4\theta^4} - \frac{\theta}{2\theta^3} \\
&= \frac{1}{2\theta^2}
\end{aligned}$$

因此, 相对有效性是 $\frac{n-1}{n}$.

习题 6.2.14

极大似然估计,

$$f = (2\pi\theta)^{-n/2} \exp\left[-\frac{1}{2\theta}(x_i-\mu)^2\right]$$

$$\log f = -\frac{n}{2}\log(2\pi\theta) - \frac{1}{2\theta}\sum_1^n (x_i-\mu)^2$$

$$l = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_1^n (x_i-\mu)^{2\pi k} = 0$$

$$\hat{\theta} = \sum_1^n (x_i-\mu)^2 / n$$

习题 6.2.14

根据

$$\sqrt{n}\left(\hat{\theta}-\theta_0\right) \to N\left(0, \frac{1}{I\left(\theta_0\right)}\right)$$

于是

$$\sqrt{n}(\hat{\theta}-\theta) \to N\left(0, 2\theta^2\right)$$