HW1 第一次作业解答

更新：2026 年 1 月 9 日

Exercise 1

利用切比雪夫不等式求一枚均匀硬币需抛掷多少次才能使样本均值 $\bar{X}$ 落在 0.4 和 0.6 之间的概率至少为 0.9（此处 $X_i = 1$ 表示抛掷硬币出现正面，否则 $X_i = 0, i = 1, 2, \cdots, n$）？若用中心极限定理计算这个问题，需抛掷的次数又是多少?

定理 0.1 (Chebyshev 不等式) 设随机变量 X 有有限的期望 $\mu = \mathbb{E}(X) < \infty$ 和有限方差 $\sigma^2 = Var(X) < \infty$，则对任意 $\epsilon > 0$，均有下列的不等式成立

$$P(|X-\mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

证明 设 $p(x)$ 为 X 的概率密度函数（非负），于是有

$$P(|X-\mu| > \epsilon) = \int_{|x-\mu|>\epsilon} p(x)\, dx$$

因为 $|X-\mu| \geq \epsilon$ 所以 $\frac{(x-\mu)^2}{\epsilon^2} \leq 1$，故

$$\int_{|x-\mu|>\epsilon} p(x)\, dx \leq \int_{|x-\mu|>\epsilon} \frac{(x-\mu)^2}{\epsilon^2} p(x)\, dx$$

扩大积分区域，有

$$\int_{|x-\mu|>\epsilon} \frac{(x-\mu)^2}{\epsilon^2} p(x)\, dx \leq \frac{1}{\epsilon^2} \int_{-\infty}^{\infty} (x-\mu)^2 p(x)\, dx$$

而根据方差的定义，得到

$$P(|X-\mu| > \epsilon) \leq \frac{1}{\epsilon^2} \int_{-\infty}^{\infty} (x-\mu)^2 p(x)\, dx = \frac{\sigma^2}{\epsilon^2}$$

定理证毕。

解答 设总共投掷了 n 次硬币，其结果分别为 $X_1, X_2, \cdots, X_n$，设每次硬币正面朝上的概率为 p 显然 $X_i \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(p)$。于是均值 $\bar{X} = \sum_{i=1}^{n} X_i/n$ 的期望和方差分别为

$$\mathbb{E}(\bar{X}) = \frac{1}{n} \cdot n\mathbb{E}(X_1) = p \quad Var(\bar{X}) = \frac{1}{n^2} \cdot nVar(X_1) = \frac{p(1-p)}{n}$$

(1) 由 Chebyshev 不等式 0.1 可知

$$P(|\bar{X}-0.5| \geq 0.1) \leq \frac{Var(\bar{X})}{0.1^2} = \frac{p(1-p)/n}{0.1^2}$$

若硬币均匀，则 $p=0.5$，所以 $P(|\bar{X}-0.5| \geq 0.1) = (0.5^2/n)/0.1^2 = 25/n$，于是

$$P(|\bar{X}-0.5| < 0.1) = 1 - \frac{25}{n} \geq 0.9$$

解得 $n_{\min} = 250$。

(2) 由中心极限定理，当 n 较大时

$$\frac{\bar{X}-0.5}{\sqrt{Var(\bar{X})}} = \frac{\bar{X}-0.5}{\sqrt{0.5^2/n}} \dot{\sim} N(0,\,1)$$

所以

$$P(|\bar{X}-0.5| < 0.1) \approx 2\Phi\left(\frac{0.1}{\sqrt{0.5^2/n}}\right) - 1 = 2\Phi(0.2\sqrt{n}) - 1$$

其中 $\Phi(\cdot)$ 为标准正态分布的分布函数。至少为 0.9，故

$$2\Phi(0.2\sqrt{n}) - 1 \geq 0.9$$

解得 $n \geq 67.64$，即 $n_{\min} = 68$。

Exercise 2

设 $X_1, X_2, \cdots, X_n$ 为从下列总体中抽取的简单样本，利用特征函数试求样本均值 $\bar{X}$ 的分布：

(1) 参数为 λ 的 Poisson 总体 $P(\lambda)$
(2) 参数为 λ 的指数分布 $Exp(\lambda)$

解答 (1) Poisson 分布 $P(\lambda)$ 的特征函数为

$$\varphi(t) = \exp(\lambda(e^{it}-1))$$

而 $X_1, X_2, \cdots, X_n$ 是简单样本，故它们互相独立，于是有 $n\bar{X} = \sum_{k=1}^{n} X_k$ 的特征函数

$$\varphi_{n\bar{X}} = \prod_{k=1}^{n} \varphi_{X_i}(t) = \prod_{k=1}^{n} \exp(\lambda(e^{it}-1)) = \exp\left(n\lambda(e^{it}-1)\right)$$

恰为 $P(n\lambda)$ 的特征函数。而随机变量的分布由特征函数唯一确定，所以 $n\bar{X} \sim P(n\lambda)$，其分布列为

$$P(\bar{X} = \frac{k}{n}) = \frac{(n\lambda)^k}{k!} e^{-n\lambda}, \quad k = 0, 1, \cdots$$

解答 (2) 指数分布 $Exp(\lambda)$ 的特征函数为

$$\varphi(t) = \left(1 - \frac{it}{\lambda}\right)^{-1}$$

与问题 (1) 类似，且由特征函数的性质，我们有 $\bar{X} = \sum_{k=1}^{n} X_k/n$ 的特征函数

$$\varphi_{\bar{X}} = \prod_{k=1}^{n} \varphi_{X_i}(t/n) = \prod_{k=1}^{n} \left(1 - \frac{it/n}{\lambda}\right)^{-1} = \left(1 - \frac{it}{n\lambda}\right)^{-n}$$

注意到 $\Gamma(\alpha, \lambda)$ 分布的特征函数为

$$\varphi_{\Gamma}(t) = \left(1 - \frac{it}{\lambda}\right)^{-\alpha}$$

所以 $\bar{X}$ 的分布为 $\Gamma(n, n\lambda)$。

注 也可以通过特征函数求出概率密度函数：当 $\int_{-\infty}^{\infty} |\varphi(t)|\, dt < \infty$ 时，概率密度函数为

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t)\, dt$$

注 解答 (2) 也可以直接通过 $Exp(\lambda) = \Gamma(1, \lambda)$ 快速发现结果。

Exercise 3

设 r.v. X 服从参数为 α, p 的 Gamma 分布 $\Gamma(p, \alpha)$，求证

(1) $\Gamma(p, \alpha)$ 的特征函数为 $\varphi(t) = \left(\frac{\alpha}{\alpha - it}\right)^p$

(2) $\mathbb{E}(X) = p/\alpha, \quad D(X) = p/\alpha^2$

(3) 若 $X_i \sim \Gamma(p_i, \alpha), i = 1, 2, \cdots, k$ 且 $X_1, X_2, \cdots, X_k$ 相互独立，记 $p = \sum_{i=1}^{k} p_i$，则 $\sum_{i=1}^{k} X_i \sim \Gamma(p, \alpha)$

(4) 若取 $\alpha = 1/2$, $p = n/2$，则 $\Gamma(n/2, 1/2)$ 是 χ_n^2 分布。

证明 (1) 由于 X 服从参数为 α, p 的 Gamma 分布 $\Gamma(p, \alpha)$，故其概率密度函数为

$$f_X(x) = \begin{cases} \dfrac{\alpha^p}{\Gamma(p)} x^{p-1} e^{-\alpha x}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

所以，根据特征函数的定义，可知

$$\begin{aligned} \varphi_X(t) = \mathbb{E}(e^{itX}) &= \int_0^{\infty} e^{itx} \frac{\alpha^p}{\Gamma(p)} x^{p-1} e^{-\alpha x}\, dx \\ &= \frac{\alpha^p}{\Gamma(p)} \int_0^{\infty} x^{p-1} e^{-(\alpha - it)x}\, dx \\ &= \frac{\alpha^p}{\Gamma(p)} \frac{1}{(\alpha - it)^p} \int_0^{\infty} ((\alpha - it)x)^{p-1} e^{-(\alpha - it)x}\, d(\alpha - it)x \end{aligned}$$

再由 Γ 函数的积分表达式

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t}\, dt$$

可知，原式可写作

$$\varphi_X(t) = \left(\frac{\alpha}{\alpha - it}\right)^p \frac{1}{\Gamma(p)} \Gamma(p)$$

所以，X 的特征函数为 $\varphi_X(t) = \left(\frac{\alpha}{\alpha - it}\right)^p$ 命题证毕。

注 上面用到了下面这个结论：

对于含有复参数的积分 $I(\alpha, \beta)$

$$I(\alpha, \beta) = \int_0^{\infty} x^{\alpha-1} e^{-\beta x}\, dx$$

积分 $I(\alpha, \beta)$ 收敛，当且仅当 α, β 的实部大于 0，即 $\Re(\alpha) > 0$, $\Re(\beta) > 0$。且收敛时有

$$I(\alpha, \beta) = \frac{\Gamma(\beta)}{\alpha^\beta}$$

证明 (2) 使用概率密度函数积分可求一阶矩和二阶矩

$$\begin{aligned}
\mathbb{E}(X) &= \frac{\alpha^p}{\Gamma(p)} \int_0^{\infty} x^p e^{-\alpha x}\, dx = \frac{\alpha^p}{\alpha^{p+1}\Gamma(p)} \int_0^{\infty} (\alpha x)^{(p+1)-1} e^{-\alpha x}\, d\alpha x \\
&= \frac{1}{\alpha} \frac{\Gamma(p+1)}{\Gamma(p)} = \frac{p}{\alpha}
\end{aligned}$$

同理，二阶矩为

$$\begin{aligned}
\mathbb{E}(X^2) &= \frac{\alpha^p}{\Gamma(p)} \int_0^{\infty} x^{p+1} e^{-\alpha x}\, dx = \frac{\alpha^p}{\alpha^{p+2}\Gamma(p)} \int_0^{\infty} (\alpha x)^{(p+2)-1} e^{-\alpha x}\, d\alpha x \\
&= \frac{1}{\alpha^2} \frac{\Gamma(p+2)}{\Gamma(p)} = \frac{p(p+1)}{\alpha^2}
\end{aligned}$$

于是，方差为

$$D(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \frac{p(p+1)}{\alpha^2} - \left(\frac{p}{\alpha}\right)^2 = \frac{p}{\alpha^2}$$

综上，命题证毕。

注 也可以使用矩母函数计算。已知特征函数 $\varphi_X(t)$，所以可求矩母函数为 $M_X(t) = \varphi_X(-it)$。由矩母函数在 0 的 n 阶导 $M_X^{(n)}(0)$ 和 $M_X(0) = 1$，可求得 X 的 n 阶矩 $\mathbb{E}(X^n)$。

证明 (3) 由 $X_i \sim \Gamma(p_i, \alpha), i = 1, 2, \cdots, k$ 且 $X_1, X_2, \cdots, X_k$ 相互独立，可得独立和的特征函数

$$\varphi_{\sum_{i=1}^k X_i}(t) = \prod_{i=1}^{k} \varphi_{X_i}(t) = \prod_{i=1}^{k} \left(\frac{\alpha}{\alpha - it}\right)^{p_i} = \left(\frac{\alpha}{\alpha - it}\right)^{\sum_{i=1}^k p_i} = \left(\frac{\alpha}{\alpha - it}\right)^p$$

恰为 $\Gamma(p, \alpha)$ 的特征函数，所以 $\sum_{i=1}^k X_i \sim \Gamma(p, \alpha)$，命题证毕。

证明 (4) 取 $\alpha = 1/2$，$p = n/2$，则 $\Gamma(n/2, 1/2)$ 的概率密度函数为

$$f_X(x) = \begin{cases} \dfrac{(1/2)^{(n/2)}}{\Gamma(n/2)} x^{n/2-1} e^{-(1/2)x}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

恰为 χ_n^2 分布的概率密度函数。所以 $\Gamma(n/2, 1/2)$ 是 χ_n^2 分布，命题证毕。

Exercise 4

设 $X_1 \sim N(\mu_1, \sigma_1^2)$，$X_2 \sim N(\mu_2, \sigma_2^2)$ 且 X_1 与 X_2 独立，用卷积公式证明 $Y = X_1 + X_2 \sim N(\mu, \sigma^2)$ 其中 $\mu = \mu_1 + \mu_2$，$\sigma^2 = \sigma_1^2 + \sigma_2^2$。

我们先证明一个引理

引理 0.2 对任意正实数 $a \in \mathbb{R}^+$，我们有下面的积分结果

$$\int_{-\infty}^{\infty} e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a}}$$

证明 设 $y = \sqrt{a}x$，则 $dx = 1/\sqrt{a}dy$，那么原式变成

$$\int_{-\infty}^{\infty} e^{-ax^2}\, dx = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} e^{-y^2}\, dy$$

而通过简单的极坐标变换即可得到 $\int_{-\infty}^{\infty} e^{-y^2}\, dy = \sqrt{\pi}$，代入即证。

回到原问题

证明 因为 $X_1 \sim N(\mu_1, \sigma_1^2)$，$X_2 \sim N(\mu_2, \sigma_2^2)$，所以它们的密度函数为

$$f_{X_1}(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right), \qquad f_{X_2}(x) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{(x-\mu_2)^2}{2\sigma_2^2}\right).$$

且 X_1 和 X_2 相互独立，所以它们的联合概率密度为

$$f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{(x_1-\mu_1)^2}{2\sigma_1^2} - \frac{(x_2-\mu_2)^2}{2\sigma_2^2}\right)$$

由于 $Y = X_1, X_2$，记 $Z = X_2$，那么有 $X_1 = Y - Z$，$X_2 = Z$，该变换的 Jacobi 行列式值为

$$|\det J| = \left|\det \begin{pmatrix} \frac{\partial X_1}{\partial Y} & \frac{\partial X_2}{\partial Y} \\ \frac{\partial X_1}{\partial Z} & \frac{\partial X_2}{\partial Z} \end{pmatrix}\right| = \left|\det \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\right| = 1$$

所以 (Y, Z) 的联合概率密度函数为

$$f_{Y,Z}(y, z) = f_{X_1}(y - z) f_{X_2}(z)$$

从而 Y 的概率密度函数为联合密度函数的边际分布，即

$$f_Y(y) = \int_{-\infty}^{\infty} f_{Y,Z}(y, z)\, dz = \int_{-\infty}^{\infty} f_{X_1}(y - z) f_{X_2}(z)\, dz$$

代入 $f_{X_1}(\cdot)$ 和 $f_{X_2}(\cdot)$ 的函数表达式，可得

$$f_Y(y) = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{((y-z)-\mu_1)^2}{2\sigma_1^2} - \frac{(z-\mu_2)^2}{2\sigma_2^2}\right) dz$$

化简为

$$f_Y(y) = \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{\infty} \exp\left[-\left(\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2}\right) \cdot z^2 + \left(\frac{y-\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}\right) \cdot z - \left(\frac{(y-\mu_1)^2}{2\sigma_1^2} + \frac{\mu_2^2}{2\sigma_2^2}\right)\right] dz$$

记

$$\begin{cases} A = \dfrac{1}{2\sigma_1^2} + \dfrac{1}{2\sigma_2^2} \\ B = \dfrac{y-\mu_1}{\sigma_1^2} + \dfrac{\mu_2}{\sigma_2^2} \\ C = \dfrac{(y-\mu_1)^2}{2\sigma_1^2} + \dfrac{\mu_2^2}{2\sigma_2^2} \end{cases}$$

被积的指数部分变为

$$-Az^2 + Bz - C = -A\left(z - \frac{B}{2A}\right)^2 - \left(\frac{B^2}{4A} + C\right)$$

注意到 $A > 0$，而根据引理 0.2

$$\int_{-\infty}^{\infty} \exp\left(-A\left(z - \frac{B}{2A}\right)^2\right) dz = \sqrt{\frac{\pi}{A}}$$

所以原式转变为

$$\begin{aligned} f_Y(y) &= \frac{1}{2\pi\sigma_1\sigma_2}\sqrt{\frac{\pi}{A}} \exp\left(-\frac{B^2}{4A} - C\right) \\ &= \frac{1}{2\pi\sigma_1\sigma_2}\sqrt{\frac{\pi}{\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2}}} \exp\left(-\frac{\left(\frac{y-\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}\right)^2}{4\left(\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2}\right)} - \frac{(y-\mu_1)^2}{2\sigma_1^2} - \frac{\mu_2^2}{2\sigma_2^2}\right) \\ &= \frac{1}{2\pi\sigma_1\sigma_2} \cdot \frac{\sqrt{2\pi}\sigma_1\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}} \exp\left(\left(\frac{\frac{\sigma_2^2}{\sigma_1^2}(y-\mu_1)^2 + 2(y-\mu_1)\mu_2 + \frac{\sigma_1^2}{\sigma_2^2}\mu_2^2}{2(\sigma_1^2+\sigma_2^2)}\right) - \left(\frac{(y-\mu_1)^2}{2\sigma_1^2} + \frac{\mu_2^2}{2\sigma_2^2}\right)\right) \end{aligned}$$

其中指数部分可化简为

$$\begin{aligned} &\frac{\sigma_2^4(y-\mu_1)^2 + 2\sigma_1^2\sigma_2^2(y-\mu_1)\mu_2 + \sigma_1^4\mu_2^2 - \sigma_2^2(\sigma_1^2+\sigma_2^2)(y-\mu_1)^2 - \sigma_1^2(\sigma_1^2+\sigma_2^2)\mu_2^2}{2\sigma_1^2\sigma_2^2(\sigma_1^2+\sigma_2^2)} \\ &= \frac{[\sigma_2^4 - \sigma_2^2(\sigma_1^2+\sigma_2^2)] \cdot (y-\mu_1)^2 + \sigma_1^2\sigma_2^2 \cdot 2(y-\mu_1)\mu_2 + [\sigma_1^4 - \sigma_1^2(\sigma_1^2+\sigma_2^2)] \cdot \mu_2^2}{2\sigma_1^2\sigma_2^2(\sigma_1^2+\sigma_2^2)} \\ &= \frac{-\sigma_1^2\sigma_2^2 \cdot [(y-\mu_1)^2 - 2(y-\mu_1)\mu_2 + \mu_2^2]}{2\sigma_1^2\sigma_2^2(\sigma_1^2+\sigma_2^2)} \\ &= -\frac{(y-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)} \end{aligned}$$

从而

$$
\begin{aligned}
f_Y(y) &= \frac{1}{2\pi\sigma_1\sigma_2} \cdot \frac{\sqrt{2\pi}\sigma_1\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}} \cdot \exp\left(-\frac{(y-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}\right) \\
&= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}} \cdot \exp\left(-\frac{(y-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}\right) \\
&= \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)
\end{aligned}
$$

恰为 $N(\mu, \sigma^2)$ 分布的概率密度函数，所以 $Y = X_1 + X_2 \sim N(\mu, \sigma^2)$，命题证毕。

HW2 第二次作业解答

更新：2026 年 1 月 9 日

Exercise 1

设 $X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} F$，而 $(X_{(1)}, X_{(2)}, \cdots, X_{(n)})$ 是其次序统计量，已知 $P(X_{(m)} < x) = \sum_{i=m}^{n} \binom{n}{i} (F(x))^i (1-F(x))^{n-i}$。证明如下恒等式：

$$\sum_{i=m}^{n} \binom{n}{i} (F(x))^i (1-F(x))^{n-i} = m\binom{n}{m} \int_0^{F(x)} t^{m-1}(1-t)^{n-m}\, dt$$

证明 记 $F(x) = p \in [0,1]$，则原问题变为去证

$$\sum_{i=m}^{n} \binom{n}{i} p^i (1-p)^{n-i} = m\binom{n}{m} \int_0^{p} t^{m-1}(1-t)^{n-m}\, dt$$

再记 $LHS = S(p), RHS = I(p)$，我们只需要证明 $S(p) \equiv I(p)$，$p \in [0,1]$。注意到

$$S'(p) = \sum_{i=m}^{n} \binom{n}{i} [ip^{i-1}(1-p)^{n-i} - (n-i)p^i(1-p)^{n-i-1}] \tag{1}$$

而由组合数的性质

$$\begin{aligned} i\binom{n}{i} &= n\binom{n-1}{i-1} \\ (n-i)\binom{n}{i} &= (n-i)\binom{n}{n-i} = n\binom{n-1}{n-i-1} = n\binom{n-1}{i} \end{aligned}$$

代回式 1 可得

$$\begin{aligned} S'(p) &= n\sum_{i=m}^{n} \left[\binom{n-1}{i-1} p^{i-1}(1-p)^{n-i} - \binom{n-1}{i} p^i (1-p)^{n-i-1}\right] \\ &= n\sum_{i=m}^{n} \binom{n-1}{i-1} p^{i-1}(1-p)^{n-i} - n\sum_{i=m+1}^{n} \binom{n-1}{i-1} p^{i-1}(1-p)^{n-i} \\ &= n\binom{n-1}{m-1} p^{m-1}(1-p)^{n-m} \end{aligned} \tag{2}$$

下面考虑 $I(p)$ 的导数 $I'(p)$

$$\begin{aligned} I'(p) &= \frac{d}{dp}\left[m\binom{n}{m}\int_0^p t^{m-1}(1-t)^{n-m}\,dt\right] \\ &= m\binom{n}{m}p^{m-1}(1-p)^{n-m} \end{aligned} \tag{3}$$

注意到

$$n\binom{n-1}{m-1} = n\frac{(n-1)!}{(m-1)!(n-m)!} = \frac{n!}{(m-1)!(n-m)!} = m\binom{n}{m}$$

结合式 2 和 3 有

$$S'(p) = n\binom{n-1}{m-1}p^{m-1}(1-p)^{n-m} = m\binom{n}{m}p^{m-1}(1-p)^{n-m} = I'(p), \quad \forall\, p \in [0,1]$$

记 $D(p) = S(p) - I(p)$, $p \in [0,1]$，则有

$$D'(p) = S'(p) - I'(p) = 0 \quad \forall\, p \in [0,1]$$

那么 $D(\cdot)$ 恒为一个常值函数，不妨设 $D(p) \equiv c$ 为常数。又注意到 $D(0) = S(0) - I(0) = 0$，故常值 $c = 0$，于是 $D(p) \equiv 0$，即

$$S(p) \equiv I(p), \quad \forall\, p \in [0,1]$$

问题证毕。

Exercise 2

设 $X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} N(a,\sigma^2)$，且 $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$，$S_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$，又设 $X_{n+1} \sim N(a,\sigma^2)$ 且与 $X_1, X_2, \cdots, X_n$ 独立，试求统计量

$$T = \frac{X_{n+1} - \bar{X}}{S_n}\sqrt{\frac{n-1}{n+1}}$$

的分布。

解答 首先根据 $X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} N(a,\sigma^2)$ 我们可以得到 $\bar{X}, S_n^2$ 的分布

$$\bar{X} \sim N(a, \frac{\sigma^2}{n}) \tag{4}$$

$$\frac{nS_n^2}{\sigma^2} \sim \chi^2_{n-1} \tag{5}$$

又因为 $X_{n+1} \sim N(a,\sigma^2)$ 且与 $(X_i)_{i<n+1}$ 独立，故 X_{n+1} 也与 $\bar{X}, S_n^2$ 独立。于是根据式 4 和 X_{n+1} 的分布可求

$$Z := \frac{X_{n+1} - \bar{X}}{\sigma\sqrt{(n+1)/n}} \sim N(0,1)$$

为简化式 5，记 $W = nS_n^2/\sigma^2 \sim \chi^2_{n-1}$，于是根据 t 分布的性质，有

$$T = \frac{X_{n+1} - \bar{X}}{S_n}\sqrt{\frac{n-1}{n+1}} = \frac{Z}{\sqrt{W/(n-1)}} \sim t_{n-1}$$

所以，统计量 $T \sim t_{n-1}$。

Exercise 3

设 $X_1, X_2, \cdots, X_n$ 独立，且 $X_i \sim N(0, \sigma_i^2)$, $i = 1, 2, \cdots, n$。定义

$$\xi = \sum_{i=1}^{n} \frac{(X_i - Z)^2}{\sigma_i^2} \quad \text{其中 } Z = \sum_{i=1}^{n} \frac{X_i}{\sigma_i^2} \Bigg/ \sum_{i=1}^{n} \frac{1}{\sigma_i^2}$$

求 ξ 的分布。

解答 记 $w_i = 1/\sigma_i^2$, $W = \sum_{i=1}^{n} w_i$，于是有

$$\xi = \sum_{i=1}^{n} w_i (X_i - Z)^2, \quad Z = \sum_{i=1}^{n} \frac{w_i X_i}{W}$$

化简 ξ 的表达式

$$\begin{aligned}
\xi &= \sum_{i=1}^{n} w_i (X_i - Z)^2 = \sum_{i=1}^{n} w_i (X_i^2 - 2ZX_i + Z^2) \\
&= \sum_{i=1}^{n} w_i X_i^2 - 2Z \sum_{i=1}^{n} w_i X_i + WZ^2 \\
&= \sum_{i=1}^{n} w_i X_i^2 - 2Z(ZW) + WZ^2 \\
&= \sum_{i=1}^{n} w_i X_i^2 - WZ^2
\end{aligned}$$

记 $Y_i = X_i/\sigma_i \Rightarrow Y_i \overset{\text{i.i.d.}}{\sim} N(0, 1)$，且

$$\sum_{i=1}^{n} w_i X_i^2 = \sum_{i=1}^{n} w_i \sigma_i^2 Y_i^2 = \sum_{i=1}^{n} Y_i^2$$

$$WZ = \left(\sum_{i=1}^{n} w_i \right) \cdot \left(\sum_{i=1}^{n} \frac{w_i X_i}{W} \right) = \sum_{i=1}^{n} w_i X_i = \sum_{i=1}^{n} \frac{Y_i}{\sigma_i}$$

所以，ξ 可以用 Y_i 表示为

$$\xi = \sum_{i=1}^{n} Y_i^2 - \frac{1}{W} \left(\sum_{i=1}^{n} \frac{Y_i}{\sigma_i} \right)^2 \tag{6}$$

如此 ξ 转变为标准正态的 Y_i 的表达，下面进行变换。

记向量 $Y = (Y_1, Y_2, \cdots, Y_n)^T \in \mathbb{R}^n$，那么根据 Y_i 相互独立且标准正态，有随机向量 Y 的分布为多元正态 $Y \sim N_n(0_n, I_n)$，其中 $0_n \in \mathbb{R}^n$, $I_n \in \mathbb{R}^{n \times n}$ 分别表示 n 维分量均为 0 的向量和 $n \times n$ 的单位阵。

作正交变换 $Z = AY \in \mathbb{R}^n$，其中 $A \in \mathbb{R}^{n \times n}$ 为正交阵。

$$A = \begin{pmatrix}
\dfrac{1}{\sigma_1\sqrt{W}} & \dfrac{1}{\sigma_2\sqrt{W}} & \cdots & \dfrac{1}{\sigma_n\sqrt{W}} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{pmatrix}_{n \times n}$$

A 可由 Schmidt 正交化构造。于是 $Z \sim N_n(A \cdot 0_n, AI_nA^T) = N_n(0_n, I_n)$。

由于

$$Z_1 = \begin{pmatrix} \frac{1}{\sigma_1\sqrt{W}} & \frac{1}{\sigma_2\sqrt{W}} & \cdots & \frac{1}{\sigma_n\sqrt{W}} \end{pmatrix} \cdot Y = \sum_{i=1}^{n} \frac{Y_i}{\sigma_i\sqrt{W}}$$

以及

$$\sum_{i=1}^{n} Z_i^2 = Z^TZ = (AY)^T(AY) = Y^TA^TAY = Y^TY = \sum_{i=1}^{n} Y_i^2$$

代回式 6 有

$$\xi = \sum_{i=1}^{n} Z_i^2 - Z_1^2 = \sum_{i=2}^{n} Z_i^2$$

而根据 $Z \sim N_n(0_n, I_n)$，知 $Z_i \sim N(0,1)$ 且相互独立。于是，由 χ^2 分布的定义知 $\xi \sim \chi^2_{n-1}$。

Exercise 4

设 $X_1, X_2, \cdots, X_n$ 为取自 Poisson 总体 $P(\lambda)$ 的样本。试证明：

$$\frac{\bar{X}_n - \lambda}{\sqrt{\bar{X}_n/n}}$$

的极限分布为 $N(0,1)$。其中 $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$。

证明 由于 $X_1, X_2, \cdots, X_n$ 为取自 Poisson 总体 $P(\lambda)$ 的样本，故

$$\mathbb{E}(X_i) = \mathbb{D}(X_i) = \lambda$$

于是，根据 Khinchin 大数定律 0.1 有

$$\bar{X}_n \xrightarrow{\mathcal{P}} \lambda$$

同理，根据 Lindeberg-Lévy 中心极限定理 0.2 有

$$\frac{\sqrt{n}(\bar{X}_n - \lambda)}{\sqrt{\lambda}} \xrightarrow{\mathcal{L}} N(0,1)$$

再根据 Slutsky 引理 0.3 有

$$\frac{\bar{X}_n - \lambda}{\sqrt{\bar{X}_n/n}} \xrightarrow{\mathcal{L}} \frac{\bar{X}_n - \lambda}{\sqrt{\lambda/n}} = \frac{\sqrt{n}(\bar{X}_n - \lambda)}{\sqrt{\lambda}} \xrightarrow{\mathcal{L}} N(0,1)$$

问题证毕。

定理 0.1 (Khinchin 大数定律) 考虑独立同分布的随机变量列 $(X_i)_{i\geq 1}$ 假设 $\mathbb{E}(|X_1|) < \infty$。记 $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$ 且 $\mu = \mathbb{E}(X_1)$。则

$$\bar{X}_n \xrightarrow{\mathcal{P}} \mu$$

定理 0.2 (Lindeberg-Lévy 中心极限定理) 考虑独立同分布的随机变量列 $(X_i)_{i\geq 1}$ 记 $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$ 且 $\mu = \mathbb{E}(X_1)$。若 $\sigma^2 = \mathbb{D}(X_1) \in (0, +\infty)$，则当 $n \to \infty$ 时有

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{\mathcal{L}} N(0,1)$$

引理 0.3 (Slutsky 引理) 令 $\{X_n\}$ 和 $\{Y_n\}$ 是两个随机变量的序列，满足当 $n \to \infty$ 时，

$$X_n \xrightarrow{\mathcal{L}} X, \quad Y_n \xrightarrow{\mathcal{P}} c$$

其中 $c \in \mathbb{R}$ 为常数。则有:

(1) $X_n \pm Y_n \xrightarrow{\mathcal{L}} X \pm c$

(2) $X_n Y_n \xrightarrow{\mathcal{L}} cX$

(3) $X_n / Y_n \xrightarrow{\mathcal{L}} X/c$

HW3 第三次作业解答

更新：2026 年 1 月 9 日

Exercise 1

设总体 X 服从双指数分布，其分布函数为

$$F(x) = \begin{cases} 1 - \exp\left(-\dfrac{x-\mu}{\sigma}\right), & x > \mu, \\ 0, & x \le \mu. \end{cases}$$

其中 $\mu \in \mathbb{R}$, $\sigma > 0$，我们有次序统计量 $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ 。试证明：

$$\frac{2(n-i+1)}{\sigma}(X_{(i)} - X_{(i-1)}) \sim \chi_2^2, \quad i = 2, 3, \cdots, n$$

证明 记总体 $Y = X - \mu$，则 Y 的概率密度为

$$F_Y(y) = 1 - \exp\left(-\frac{1}{\sigma}y\right)\mathbb{I}_{\{y>0\}}$$

为指数分布 Exp$(1/\sigma)$。记 $1/\sigma = \lambda$ 即 $Y \sim$ Exp(λ)。记 Y 的次序统计量为 $Y_{(1)} \le Y_{(2)} \le \cdots \le Y_{(n)}$，则有 $X_{(i)} - X_{(i-1)} = Y_{(i)} - Y_{(i-1)} := Z_i$, $i = 2, 3, \cdots, n$，特别地 $Z_1 = Y_{(1)}$。反解出 $Y_{(i)}$ 有

$$Y_{(i)} = \sum_{k=1}^{i} Z_k, \quad i = 1, 2, \cdots, n$$

容易得到 Jacobi 阵的行列式的绝对值为 1。

由 Y 的概率密度为 $f_Y(y) = \lambda e^{-\lambda y}$，以及次序统计量的联合分布，我们有

$$\begin{aligned} f_{Y_{(1)},Y_{(2)},\cdots,Y_{(n)}}(y_{(1)}, y_{(2)}, \cdots, y_{(n)}) &= n! \cdot f_Y(y_{(1)}) f_Y(y_{(2)}) \cdots f_Y(y_{(n)}) \mathbb{I}_{\{y_{(1)} \le y_{(2)} \le \cdots \le y_{(n)}\}} \\ &= n! \prod_{i=1}^{n} \lambda e^{-\lambda y_{(i)}} \mathbb{I}_{\{y_{(1)} \le y_{(2)} \le \cdots \le y_{(n)}\}} \\ &= n! \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} y_{(i)}\right) \mathbb{I}_{\{y_{(1)} \le y_{(2)} \le \cdots \le y_{(n)}\}} \end{aligned}$$

因为 Jacobi 阵的行列式的绝对值为 1 以及 $y_{(i)} = \sum_{k=1}^{i} z_k$，可以得到 $Z_1, Z_2, \cdots, Z_n$ 的联合分布为

$$\begin{aligned} f_{Z_1,Z_2,\cdots,Z_n}(z_1,z_2,\cdots,z_n) &= n!\lambda^n \exp\left(-\lambda \sum_{i=1}^{n}\sum_{k=1}^{i} z_k\right) \mathbb{I}_{\cap_k\{z_k\geq 0\}} \\ &= n!\lambda^n \exp\left(-\lambda \sum_{k=1}^{n}\sum_{i=k}^{n} 1\cdot z_k\right) \mathbb{I}_{\cap_k\{z_k\geq 0\}} \\ &= n!\lambda^n \exp\left(-\lambda \sum_{k=1}^{n}(n-k+1) z_k\right) \mathbb{I}_{\cap_k\{z_k\geq 0\}} \end{aligned}$$

注意到 $\prod_{k=1}^{n}(n-k+1) = n!$ 有

$$f_{Z_1,Z_2,\cdots,Z_n}(z_1,z_2,\cdots,z_n) = \prod_{k=1}^{n} \lambda(n-k+1)e^{-\lambda(n-k+1)z_k}$$

其中 $g(z_k) = \lambda(n-k+1)e^{-\lambda(n-k+1)z_k}$ 恰为 Exp$(\lambda(n-k+1))$ 的概率密度函数。所以

$$\int \lambda(n-k+1)e^{-\lambda(n-k+1)z_k}\,dz_k = 1$$

为求 Z_k 边际分布，将其他 $l \neq k$ 的 z_l 积分有

$$\begin{aligned} f_{Z_k}(z) &= g(z_k)\int\cdots\int_{z_l\geq 0,\ l\neq k} \prod_{l\neq k}^{n} \lambda(n-k+1)e^{-\lambda(n-k+1)z_l}\,dz_1\cdots dz_{k-1}dz_{k+1}\cdots dz_n \\ &= g(z_k)\prod_{l\neq k}^{n}\int_{z_l\geq 0} \lambda(n-k+1)e^{-\lambda(n-k+1)z_l}\,dz_l \\ &= g(z_k)\prod_{l\neq k}^{n} 1 = g(z_k) \end{aligned}$$

上面的第二个等式是因为 $g(z_l)\geq 0$ 且可积，使用 Fubini 定理交换积分和乘积符号。综上，我们有

$$Z_k \sim \text{Exp}(\lambda(n-k+1))$$

由引理：若 $W\sim \text{Exp}(\theta)$ 则 $2\theta W \sim \chi_2^2$，我们有

$$2\lambda(n-k+1)Z_k \sim \chi_2^2$$

又 $Z_i = X_{(i)} - X_{(i-1)}$ 和 $\lambda = 1/\sigma$ 所以

$$\frac{2(n-i+1)}{\sigma}(X_{(i)} - X_{(i-1)}) \sim \chi_2^2, \quad i = 2,3,\cdots,n$$

命题证毕。

Exercise 2

将指数分布写成指数族的自然形式，并求出其自然参数空间。

解答 指数分布 Exp(λ) 的联合密度函数为

$$f(x;\lambda)=\lambda e^{-\lambda x}\mathbb{I}_{(0,\infty)}(x)$$

根据指数族形式的定义，有 $C(\lambda)=\lambda$，$Q(\lambda)=-\lambda$，$T(x)=x$，$h(x)=\mathbb{I}_{(0,\infty)}(x)$。下面，令 $\varphi=Q(\lambda)$ 则有 $\lambda=-\varphi$，于是 $C(\lambda)=-\varphi=C^*(\varphi)$。则指数分布的指数族的自然形式为

$$f(x;\lambda)=-\varphi e^{\varphi x}\mathbb{I}_{(0,\infty)}(x)$$

自然参数空间为

$$\Theta^*=\{\varphi:\int_{\mathscr{X}}e^{\varphi x}\mathbb{I}_{(0,\infty)}(x)\,dx<\infty\}=\{\varphi:\int_{x>0}e^{\varphi x}\,dx<\infty\}$$

积分有限，故 $\varphi<0$，所以自然参数空间为 $\Theta^*=\{\varphi:\varphi<0\}$。

Exercise 3

证明指数族的自然参数空间为凸集。即对于自然参数空间

$$\Theta^*=\left\{\boldsymbol{\theta}=(\theta_1,\theta_2,\cdots,\theta_k):\int_{\mathscr{X}}\exp\left(\sum_{i=1}^k\theta_iT_i(x)\right)h(x)\,dx<\infty\right\}$$

其中 $h(x)>0$，而 $\mathscr{X}$ 为样本空间。结论是：对于任意 $\boldsymbol{\theta}_1,\boldsymbol{\theta}_2\in\Theta^*$，设 $0<\alpha<1$，记 $\boldsymbol{\theta}=\alpha\boldsymbol{\theta}_1+(1-\alpha)\boldsymbol{\theta}_2$ ，则 $\boldsymbol{\theta}\in\Theta^*$。

证明 记 $\boldsymbol{T}(x)=(T_1(x),T_2(x),\cdots,T_k(x))^T\in\mathbb{R}^k$，则我们需要证明：

$$\int_{\mathscr{X}}\exp\left(\boldsymbol{\theta}^T\boldsymbol{T}(x)\right)h(x)\,dx<\infty$$

其中 $\boldsymbol{\theta}=\alpha\boldsymbol{\theta}_1+(1-\alpha)\boldsymbol{\theta}_2$ 代入即

$$\exp\left(\boldsymbol{\theta}^T\boldsymbol{T}(x)\right)h(x)=\left[e^{\alpha\boldsymbol{\theta}_1^T\boldsymbol{T}(x)}h(x)^{\alpha}\right]\cdot\left[e^{(1-\alpha)\boldsymbol{\theta}_2^T\boldsymbol{T}(x)}h(x)^{1-\alpha}\right]$$

记 $p=1/\alpha$，$q=1/(1-\alpha)$，同时记 $\exp\left(\boldsymbol{\theta}^T\boldsymbol{T}(x)\right)h(x):=f(x)g(x)$，其中

$$f(x)=e^{\frac{1}{p}\boldsymbol{\theta}_1^T\boldsymbol{T}(x)}h(x)^{\frac{1}{p}}>0\quad g(x)=e^{\frac{1}{q}\boldsymbol{\theta}_2^T\boldsymbol{T}(x)}h(x)^{\frac{1}{q}}>0$$

由于 $\alpha\in(0,1)\Rightarrow p,q>1$ 且 $\frac{1}{p}+\frac{1}{q}=\alpha+(1-\alpha)=1$，所以由 Hölder 不等式 0.2 有

$$\int_{\mathscr{X}}f(x)g(x)\,dx\le\left(\int_{\mathscr{X}}f(x)^p\,dx\right)^{1/p}\left(\int_{\mathscr{X}}g(x)^q\,dx\right)^{1/q}\tag{1}$$

而

$$\begin{aligned}f(x)^p&=\left(e^{\frac{1}{p}\boldsymbol{\theta}_1^T\boldsymbol{T}(x)}h(x)^{\frac{1}{p}}\right)^p=e^{\boldsymbol{\theta}_1^T\boldsymbol{T}(x)}h(x)\\g(x)^q&=\left(e^{\frac{1}{q}\boldsymbol{\theta}_2^T\boldsymbol{T}(x)}h(x)^{\frac{1}{q}}\right)^q=e^{\boldsymbol{\theta}_2^T\boldsymbol{T}(x)}h(x)\end{aligned}$$

又因为 $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta^*$，所以

$$\int_{\mathscr{X}} f(x)^p \, dx = \int_{\mathscr{X}} e^{\boldsymbol{\theta}_1^T \boldsymbol{T}(x)} h(x) \, dx < \infty$$
$$\int_{\mathscr{X}} g(x)^q \, dx = \int_{\mathscr{X}} e^{\boldsymbol{\theta}_2^T \boldsymbol{T}(x)} h(x) \, dx < \infty$$

由不等式 1 可知

$$\int_{\mathscr{X}} f(x) g(x) \, dx = \int_{\mathscr{X}} \exp\left(\boldsymbol{\theta}^T \boldsymbol{T}(x)\right) h(x) \, dx < \infty$$

即有 $\boldsymbol{\theta} \in \Theta^*$ 证毕。

定理 0.1 (Young 不等式) 设 $a, b \geq 0$ $p, q > 1$ 满足

$$\frac{1}{p} + \frac{1}{q} = 1$$

则有

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

当且仅当 $a^p = b^q$ 时，等号成立。

证明 由 $\varphi(x) = e^x$ 为凸函数，由凸函数的性质或 Jensen 不等式

$$\varphi(\lambda x_1 + (1-\lambda) x_2) \leq \lambda \varphi(x_1) + (1-\lambda) \varphi(x_2)$$

令 $\lambda = 1/p$ 由条件 $1 - \lambda = 1/q$。下面，对任意 $a, b > 0$ 取 $x_1 = p \ln a$，$x_2 = q \ln b$，由上述的 Jensen 不等式

$$e^{\lambda p \ln a + (1-\lambda) q \ln b} \leq \lambda e^{p \ln a} + (1-\lambda) e^{q \ln b}$$

而左边

$$e^{\lambda p \ln a + (1-\lambda) q \ln b} = e^{\frac{1}{p} \cdot p \ln a} \cdot e^{\frac{1}{q} \cdot q \ln b} = ab$$

右边

$$\lambda e^{p \ln a} + (1-\lambda) e^{q \ln b} = \frac{1}{p} \cdot e^{\ln a^p} + \frac{1}{q} \cdot e^{\ln b^q} = \frac{a^p}{p} + \frac{b^q}{q}$$

综上

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

定理证毕。

定理 0.2 (Hölder 不等式) 设 $f(x), g(x)$ 为可测函数，$p, q > 1$ 满足

$$\frac{1}{p} + \frac{1}{q} = 1$$

则有

$$\int |f(x) g(x)| \, dx \leq \left(\int |f(x)|^p \, dx \right)^{1/p} \left(\int |g(x)|^q \, dx \right)^{1/q}$$

等号成立当且仅当存在常数 $C > 0$，使得对几乎处处的 x 有 $|f(x)|^p = C |g(x)|^q$ 成立。

证明 这里只证明不平凡的情形。记

$$A := \int |f(x)|^p \, dx \in (0, \ \infty), \quad B := \int |g(x)|^q \, dx \in (0, \ \infty)$$

设函数

$$u(x) = \frac{|f(x)|}{A^{1/p}} > 0, \quad v(x) = \frac{|g(x)|}{B^{1/q}} > 0$$

则容易得 $\int u(x)^p \, dx = \int v(x)^q \, dx = 1$。由 Young 不等式（定理 0.1），令 $a = u(x)$, $b = v(x)$，则

$$u(x)v(x) \leq \frac{u(x)^p}{p} + \frac{v(x)^q}{q}$$

由上式对任意的 x 都成立，且 $0 < u, v < \infty$，故积分不等式仍成立

$$\int u(x)v(x) \, dx \leq \int u(x)^p \, dx + \int v(x)^q \, dx = 1$$

即

$$\int \frac{|f(x)|}{A^{1/p}} \frac{|g(x)|}{B^{1/q}} \, dx = A^{-1/p} B^{-1/q} \int |f(x)g(x)| \, dx \leq 1$$

代入 A, B 即有

$$\int |f(x)g(x)| \, dx \leq \left(\int |f(x)|^p \, dx \right)^{1/p} \left(\int |g(x)|^q \, dx \right)^{1/q}$$

定理证毕。

Exercise 4

设 r.v. X 的分布函数为 $F(x)$，密度函数为 $f(x)$，且 $X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} f$。证明：

$$\underset{a<x_1<\cdots<x_n<b}{\int \cdots \int} f(x_1) \cdots f(x_n) \, dx_1 \cdots dx_n = \frac{1}{n!} [F(b) - F(a)]^n$$

证明 记事件 $A = \{\boldsymbol{X} \in \mathbb{R}^n : X_i \in (a, b) : i = 1, 2, \cdots, n\}$ 和事件 $B = \{\boldsymbol{X} \in \mathbb{R}^n : a < X_{(1)} < \cdots < X_{(n)} < b\}$，其中 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T$ 中的各个分量 X_i 为相互独立的随机变量。下面我们严格证明：$A = B$。

- 对 $\forall \boldsymbol{X} \in A$，显然有最小的 $X_{(1)} < X_i < b$ 和最大的 $X_{(n)} > X_i > a$，即有 $a < X_{(1)} < \cdots < X_{(n)} < b$，即 $\boldsymbol{X} \in B$，故 $A \subseteq B$。
- 对 $\forall \boldsymbol{X} \in B$，则对 $\forall i$ 有 $a < X_{(1)} < X_i < X_{(n)} < b$，所以 $\boldsymbol{X} \in A \Rightarrow B \subseteq A$。

综上，有 $A = B$，进一步 $P(A) = P(B)$。而 X_i 相互独立，所以

$$P(A) = P(\{X_i \in (a, b) : i = 1, 2, \cdots, n\}) = P(\{X_i \in (a, b)\})^n = [F(b) - F(a)]^n$$

注意到 n 个次序统计量相互独立，它们的联合概率密度为

$$f_{X_{(1)}, \cdots, X_{(n)}}(x_{(1)}, \cdots, x_{(n)}) = n! f(x_{(1)}) \cdots f(x_{(n)})$$

所以

$$
\begin{aligned}
P(B) &= \int\cdots\int_{a<x_{(i)}<b,\forall i} f_{X_{(1)},\cdots,X_{(n)}}(x_{(1)},\cdots,x_{(n)})\,dx_{(1)}\cdots dx_{(n)} \\
&= \int\cdots\int_{a<x_{(i)}<b,\forall i} n!f(x_{(1)})\cdots f(x_{(n)})\,dx_{(1)}\cdots dx_{(n)} \\
&= \int\cdots\int_{a<x_1<\cdots<x_n<b} n!f(x_1)\cdots f(x_n)\,dx_1\cdots dx_n
\end{aligned}
$$

由 $P(A)=P(B)$，我们有

$$
\begin{aligned}
[F(b)-F(a)]^n &= \int\cdots\int_{a<x_1<\cdots<x_n<b} n!f(x_1)\cdots f(x_n)\,dx_1\cdots dx_n \\
\Rightarrow \int\cdots\int_{a<x_1<\cdots<x_n<b} f(x_1)\cdots f(x_n)\,dx_1\cdots dx_n &= \frac{1}{n!}[F(b)-F(a)]^n
\end{aligned}
$$

命题证毕。

HW4 第四次作业解答

更新：2026 年 1 月 9 日

Exercise 1

设 $\boldsymbol{X}=(X_1,\cdots,X_n)^T$ 且 $X_i\overset{\text{i.i.d.}}{\sim}\text{Ge}(\theta)$ 几何分布，试用两种方法证明 $T(\boldsymbol{X})=\sum_{i=1}^n X_i$ 是充分统计量

- 充分统计量的定义
- 因子分解定理

证明 (利用充分统计量的定义) 由几何分布 $\text{Ge}(\theta)$ 的定义以及样本的独立性，可得分布族为

$$\mathscr{F}=\{f(\boldsymbol{x},\theta):\theta\in\Theta=(0,1)\}$$

注：样本空间为 $\mathscr{X}=\mathbb{N}^+\times\cdots\mathbb{N}^+$。其中概率函数为

$$f_{X_1,\cdots,X_n}(\boldsymbol{x},\theta)=\theta^n\prod_{i=1}^n(1-\theta)^{x_i-1}$$

那么条件概率为

$$P_\theta(X_1=x_1,\cdots,X_n=x_n|T=t)=\frac{P(X_1=x_1,\cdots,X_n=x_n,T=t)}{P(T=t)}$$

由几何分布的分布函数可知

$$\begin{aligned}\text{分子}&=P(X_1=x_1,\cdots,X_n=x_n,T=t)=P(X_1=x_1,\cdots,X_n=t-\sum\nolimits_{i=1}^{n-1}x_i)\\&=\left(\theta^{n-1}\prod\nolimits_{i=1}^{n-1}(1-\theta)^{x_i-1}\right)\cdot\left(\theta(1-\theta)^{t-\sum_{i=1}^{n-1}x_i}\right)\\&=\theta^n(1-\theta)^{\sum_{i=1}^{n-1}(x_i-1)+t-\sum_{i=1}^{n-1}x_i-1}=\theta^n(1-\theta)^{t-n}\end{aligned}$$

而

$$\begin{aligned}\text{分母}=P(T=t)&=\sum\nolimits_{\boldsymbol{x}\in\mathscr{X}}\prod\nolimits_{i=1}^n(1-\theta)^{x_i-1}\theta\cdot\mathbb{I}_{\{\boldsymbol{x}:T(\boldsymbol{x})=t\}}(\boldsymbol{x})\\&=\sum\nolimits_{\boldsymbol{x}\in\{\boldsymbol{x}:T(\boldsymbol{x})=t\}}\prod\nolimits_{i=1}^n(1-\theta)^{x_i-1}\theta=\sum\nolimits_{\boldsymbol{x}\in\{\boldsymbol{x}:T(\boldsymbol{x})=t\}}\theta^n(1-\theta)^{T(\boldsymbol{x})-n}\\&=\sum\nolimits_{\boldsymbol{x}\in\{\boldsymbol{x}:T(\boldsymbol{x})=t\}}\theta^n(1-\theta)^{t-n}\end{aligned}$$

将分子分母的结果代回条件概率得

$$
\begin{aligned}
P_\theta(X_1=x_1,\cdots,X_n=x_n|T=t) &= \frac{\theta^n(1-\theta)^{t-n}}{\sum_{\boldsymbol{x}\in\{\boldsymbol{x}:T(\boldsymbol{x})=t\}}\theta^n(1-\theta)^{t-n}} = \frac{1}{\sum_{\boldsymbol{x}\in\{\boldsymbol{x}:T(\boldsymbol{x})=t\}}1} \\
&= \frac{1}{\#\{\boldsymbol{X}:\sum_{i=1}^n X_i=t\}}
\end{aligned}
$$

其中 $\#\{\cdot\}$ 代表集合元素。综上可知，条件概率 $P_\theta(\boldsymbol{X}|T(\boldsymbol{X})=t)$ 与参数 θ 无关，故 $T(\boldsymbol{X})$ 是充分统计量。

证明 (利用因子分解定理) 注意到

$$
\begin{aligned}
f(\boldsymbol{x},\theta) &= \theta^n\prod\nolimits_{i=1}^n(1-\theta)^{x_i-1} = \theta^n(1-\theta)^{\sum_{i=1}^n(x_i-1)} \\
&= \theta^n(1-\theta)^{T(\boldsymbol{x})-n} := g(T(\boldsymbol{x}),\theta)h(x)
\end{aligned}
$$

其中 $h(x)\equiv 1$，故由因子分解定理，T 为充分统计量。

Exercise 2

设 $X=(X_1,\cdots,X_m)$ 为从正态总体 $N(a,\sigma^2)$ 中抽取的简单样本，$Y=(Y_1,\cdots,Y_n)$ 为从正态总体 $N(b,\sigma^2)$ 中抽取的简单样本，且两样本独立。记 $\bar{X}=\frac{1}{m}\sum_{i=1}^m X_i$，$\bar{Y}=\frac{1}{n}\sum_{j=1}^n Y_j$ 而

$$
S^2=\frac{1}{m+n-2}\left(\sum_{i=1}^m(X_i-\bar{X})^2+\sum_{j=1}^n(Y_j-\bar{Y})^2\right)
$$

证明：$\boldsymbol{T}(\boldsymbol{X},\boldsymbol{Y})=(\bar{X},\bar{Y},S^2)^T$ 为充分统计量。

证明 由正态分布的概率密度函数和样本的独立性，我们得到分布族的概率函数为

$$
\begin{aligned}
f(\boldsymbol{x},\boldsymbol{y};\boldsymbol{\theta}=(a,b,\sigma^2)^T) &= (2\pi\sigma^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^m(x_i-a)^2\right\}(2\pi\sigma^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{j=1}^n(y_j-b)^2\right\} \\
&= (2\pi\sigma^2)^{-n}\exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^m(x_i-a)^2+\sum_{j=1}^n(y_j-b)^2\right)\right\}
\end{aligned}
$$

注意到

$$
\begin{aligned}
\sum_{i=1}^m(x_i-a)^2 &= \sum_{i=1}^m(x_i-\bar{x}+\bar{x}-a)^2 \\
&= \sum_{i=1}^m(x_i-\bar{x})^2+2\sum_{i=1}^m(x_i-\bar{x})(\bar{x}-a)+\sum_{i=1}^m(\bar{x}-a)^2 \\
&= \sum_{i=1}^m(x_i-\bar{x})^2+2(\bar{x}-a)\sum_{i=1}^m(x_i-\bar{x})+m(\bar{x}-a)^2 \\
&= \sum_{i=1}^m(x_i-\bar{x})^2+m(\bar{x}-a)^2
\end{aligned}
$$

同理有 $\sum_{j=1}^n (y_j-b)^2 = \sum_{j=1}^n (y_j-\bar{y})^2 + n(\bar{y}-b)^2$。代回概率函数有

$$\begin{aligned} f(\boldsymbol{x},\boldsymbol{y};\boldsymbol{\theta}) &= (2\pi\sigma^2)^{-n}\exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^m (x_i-\bar{x})^2 + m(\bar{x}-a)^2 + \sum_{j=1}^n (y_j-\bar{y})^2 + n(\bar{y}-b)^2\right)\right\} \\ &= (2\pi\sigma^2)^{-n}\exp\left\{-\frac{1}{2\sigma^2}\left((m+n-2)s^2 + m(\bar{x}-a)^2 + n(\bar{y}-b)^2\right)\right\} \\ &= g(\bar{x},\bar{y},s^2;a,b,\sigma^2) \end{aligned}$$

故概率函数可分解为 $f(\boldsymbol{x},\boldsymbol{y};\boldsymbol{\theta}) = g(\boldsymbol{T}(\boldsymbol{x},\boldsymbol{y}),\boldsymbol{\theta})h(\boldsymbol{x},\boldsymbol{y})$，其中 $h(\boldsymbol{x},\boldsymbol{y})\equiv 1$。由因子分解定理可知 $\boldsymbol{T}$ 是充分统计量。

Exercise 3

在指数族的自然形式下，若自然参数空间有内点，求 $\mathrm{Cov}(T_i(X),T_j(X))$ 的表达式。

定理 0.1 设指数族的自然形式中，自然参数空间有内点，其内点集为 Θ_0。对 $\boldsymbol{\theta}\in\Theta_0$，有 $T_i(x)$ 的各阶矩均存在有限，可通过在积分号下求导算出。例如：

$$\begin{aligned} \mathbb{E}(T_i(x)) &= -\frac{\partial \log C^*(\boldsymbol{\theta})}{\partial\theta_i} \\ \mathrm{Cov}(T_i(x),T_j(x)) &= -\frac{\partial^2 \log C^*(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j} \end{aligned}$$

证明 由概率定义 $\int_{\mathscr{X}} f(x;\theta)=1$，有关系式

$$\frac{1}{C^*(\boldsymbol{\theta})} = \int_{\mathscr{X}} \exp\left\{\sum_{j=1}^k \theta_j T_j(x)\right\} h(x)\,dx \tag{1}$$

对式 1 两边对 θ_i 求偏导有

$$-\frac{1}{C^*(\boldsymbol{\theta})^2}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_i} = \int_{\mathscr{X}} T_i(x)\exp\left\{\sum_{j=1}^k \theta_j T_j(x)\right\} h(x)\,dx \tag{2}$$

注意到

$$\mathbb{E}(T_i(x)) = \int_{\mathscr{X}} T_i(x) C^*(\boldsymbol{\theta})\exp\left\{\sum_{j=1}^k \theta_j T_j(x)\right\} h(x)\,dx$$

所以

$$\mathbb{E}(T_i(x)) = -\frac{1}{C^*(\boldsymbol{\theta})^2}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_i}\cdot C^*(\boldsymbol{\theta}) = -\frac{\partial \log C^*(\boldsymbol{\theta})}{\partial\theta_i}$$

继续对式 2 关于 θ_j 求偏导有

$$\frac{2}{C^*(\boldsymbol{\theta})^3}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_i}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_j} - \frac{1}{C^*(\boldsymbol{\theta})^2}\frac{\partial^2 C^*(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j} = \int_{\mathscr{X}} T_i(x)T_j(x)\exp\left\{\sum_{l=1}^k \theta_l T_l(x)\right\} h(x)\,dx \tag{3}$$

注意到

$$
\begin{aligned}
\text{Cov}(T_i(x), T_j(x)) &= \mathbb{E}(T_i(x)T_j(x)) - \mathbb{E}(T_i(x))\mathbb{E}(T_j(x)) \\
&= \int_{\mathscr{X}} T_i(x)T_j(x)C^*(\boldsymbol{\theta}) \exp\left\{\sum_{l=1}^{k} \theta_l T_l(x)\right\} h(x)\, dx - \mathbb{E}(T_i(x))\mathbb{E}(T_j(x))
\end{aligned}
$$

代入式 2 和 $\mathbb{E}(T_i(x)), \mathbb{E}(T_j(x))$ 的表达式得

$$
\begin{aligned}
\text{Cov}(T_i(x), T_j(x)) = &\left(\frac{2}{C^*(\boldsymbol{\theta})^3}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_i}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_j} - \frac{1}{C^*(\boldsymbol{\theta})^2}\frac{\partial^2 C^*(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right)\cdot C^*(\boldsymbol{\theta}) \\
&- \left(-\frac{1}{C^*(\boldsymbol{\theta})}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_i}\right)\left(-\frac{1}{C^*(\boldsymbol{\theta})}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
=&\frac{1}{C^*(\boldsymbol{\theta})^2}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_i}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_j} - \frac{1}{C^*(\boldsymbol{\theta})}\frac{\partial^2 C^*(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j} \\
=&-\frac{\partial^2 \log C^*(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}
\end{aligned}
$$

命题证毕。

Exercise 4

设样本 $\boldsymbol{X} = (X_1, \cdots, X_n)^T \in \mathscr{X}$ 的概率函数 $f(\boldsymbol{x}, \theta) = f(x_1, \cdots, x_n; \theta)$ 依赖于参数 θ。如果 $T = T(\boldsymbol{X})$ 为 θ 的充分统计量，设 $S(\boldsymbol{X}) = G(T(\boldsymbol{X}))$，其中 G 为单值可逆函数，用因子分解定理证明 $S(\boldsymbol{X}) = G(T(\boldsymbol{X}))$ 也是 θ 的充分统计量。

证明 因为 $T = T(\boldsymbol{X})$ 为 θ 的充分统计量，所以概率函数 f 可分解

$$
f(\boldsymbol{x}, \theta) = g(T(\boldsymbol{x}), \theta)h(\boldsymbol{x})
$$

又因为 G 为单值可逆函数，故存在反函数 G^{-1}，由 $S(\boldsymbol{X}) = G(T(\boldsymbol{X}))$ 有

$$
T(\boldsymbol{x}) = G^{-1}(S(\boldsymbol{x}))
$$

代入 $f(\boldsymbol{x}, \theta)$ 有关系式

$$
f(\boldsymbol{x}, \theta) = g(G^{-1}(S(\boldsymbol{x})), \theta)h(\boldsymbol{x}) := \tilde{g}(S(\boldsymbol{x}), \theta)h(\boldsymbol{x})
$$

所以，由因子分解定理知，$S(\boldsymbol{X}) = G(T(\boldsymbol{X}))$ 也是 θ 的充分统计量。

HW5 第五次作业解答

更新：2026 年 1 月 9 日

Exercise 1

设 $\boldsymbol{X} = (X_1, \cdots, X_n) \overset{\text{i.i.d.}}{\sim} E(\theta, 1)$ 其中概率密度函数为

$$f(x; \theta) = e^{-(x-\theta)} \cdot \mathbb{I}(x > \theta)$$

其中 $\theta \in \mathbb{R}$ 为位置参数。证明：$T(\boldsymbol{X}) = X_{(1)}$ 为极小充分完全统计量。

证明 分 3 步证明：先证充分，再证极小，最后证完全。

(1) 充分统计量 样本 $(X_1, \cdots, X_n)$ 的联合密度为

$$\begin{aligned} f(\boldsymbol{x}; \theta) &= \prod_{i=1}^{n} e^{-(x_i-\theta)} \cdot \mathbb{I}(x_i > \theta) = \exp\left\{-\sum_{i=1}^{n} x_i + n\theta\right\} \cdot \mathbb{I}\left(\min_i x_i > \theta\right) \\ &= e^{n\theta}\mathbb{I}\left(x_{(1)} > \theta\right) \cdot \exp\left\{-\sum_{i=1}^{n} x_i\right\} := g(T(\boldsymbol{x}), \theta)h(\boldsymbol{x}) \end{aligned}$$

其中 $T(\boldsymbol{X}) = X_{(1)}$ 于是 $g(T(\boldsymbol{x}), \theta) = e^{n\theta}\mathbb{I}(x_{(1)} > \theta)$，而 $h(\boldsymbol{x}) = \exp\{-\sum_{i=1}^{n} x_i\}$。根据因子分解定理，$T(\boldsymbol{X}) = X_{(1)}$ 为充分统计量。

(2) 极小充分统计量 考虑样本 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 的联合密度的比为

$$\frac{f(\boldsymbol{x}; \theta)}{f(\boldsymbol{y}; \theta)} = \frac{e^{n\theta}\mathbb{I}\left(x_{(1)} > \theta\right) \cdot \exp\left\{-\sum_{i=1}^{n} x_i\right\}}{e^{n\theta}\mathbb{I}\left(y_{(1)} > \theta\right) \cdot \exp\left\{-\sum_{i=1}^{n} y_i\right\}} = \exp\left\{-\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i\right\} \frac{\mathbb{I}\left(x_{(1)} > \theta\right)}{\mathbb{I}\left(y_{(1)} > \theta\right)}$$

上式与 θ 无关，等价于 $X_{(1)} = Y_{(1)}$（否则取 θ 处于 $X_{(1)}$ 和 $Y_{(1)}$ 之间）。于是根据定理 0.1 可知 $T(\boldsymbol{X}) = X_{(1)}$ 是极小充分统计量。

(2) 极小充分完全统计量 首先 $X_{(1)}$ 为次序统计量，故其概率密度为

$$\begin{aligned} f_{X_{(1)}}(x) &= \binom{n}{1}\left[1 - \int_{\theta}^{x} e^{-(t-\theta)}\, dt\right]^{n-1} e^{-(x-\theta)} \cdot \mathbb{I}(x > \theta) \\ &= ne^{-n(x-\theta)}\mathbb{I}(x > \theta) \end{aligned}$$

于是对任一函数 $\varphi(\cdot)$ 有对于任意 θ 有

$$\mathbb{E}_\theta(\varphi(T(\boldsymbol{X}))) = \mathbb{E}_\theta(\varphi(X_{(1)})) = \int_\theta^\infty \varphi(x) n e^{-n(x-\theta)}\, dx = n e^{n\theta} \int_\theta^\infty \varphi(x) e^{-nx}\, dx$$

令 $\mathbb{E}_\theta(\varphi(T(\boldsymbol{X}))) = 0$，由于 $ne^{n\theta} > 0$ 有

$$ne^{n\theta} \int_\theta^\infty \varphi(x) e^{-nx}\, dx = 0 \Rightarrow \int_\theta^\infty \varphi(x) e^{-nx}\, dx = 0$$

等式对 θ 求导得 $0 - \varphi(\theta)e^{-n\theta} = 0$。于是 $\forall\, \theta$ 有 $\varphi(\theta) \equiv 0$。

综上有 $\mathbb{E}_\theta(\varphi(T(\boldsymbol{X}))) = 0$ 等价于 $\varphi(\cdot) \equiv 0$，即 $P_\theta(\varphi(T(\boldsymbol{X})) = 0) = 1$。由完全统计量的定义可知 $T(\boldsymbol{X}) = X_{(1)}$ 为极小充分完全统计量，命题证毕。

定理 0.1 设 $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} f(x,\theta)$, $\theta \in \Theta$，若下面等价条件成立：

$$\frac{f(x,\theta)}{f(y,\theta)} \text{不依赖于}\theta \Leftrightarrow T(x) = T(y)$$

则 T 为极小充分统计量。

Exercise 2

设 $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} U(\theta, 2\theta)$ 其中 $\theta > 0$。证明：$(X_{(1)}, X_{(n)})$ 是充分统计量，但不是完全的。

证明 $X_1, \cdots, X_n$ 的联合密度为

$$f(\boldsymbol{x};\theta) = \prod_{i=1}^n \frac{1}{\theta} \cdot \mathbb{I}_{(\theta,2\theta)}(x_i) = \theta^{-n}\mathbb{I}(\theta < x_{(1)} \le x_{(n)} < 2\theta) := g(T(\boldsymbol{x}),\theta)h(\boldsymbol{x})$$

其中 $g(T(\boldsymbol{x}),\theta) = \theta^{-n}\mathbb{I}(\theta < x_{(1)} \le x_{(n)} < 2\theta)$ 而 $h(\boldsymbol{x}) = 1$。由因子分解定理，$(X_{(1)}, X_{(n)})$ 是充分统计量。

下面，令 $Y_i = X_i/\theta$ 则易知 $Y_i \sim U(1,2)$ 与 θ 无关。于是 $Z = X_{(n)}/X_{(1)}$ 的分布也与 θ 无关，即 Z 的分布确定。所以，存在 $a < b$ 使得 $P(Z < a) = P(Z > b) > 0$，构造函数 $\varphi(\cdot)$ 为

$$\varphi(x_{(1)}, x_{(n)}) = \begin{cases} 1, & x_{(n)}/x_{(1)} < a \\ -1, & x_{(n)}/x_{(1)} > a \\ 0, & \text{其他} \end{cases}$$

于是 $\mathbb{E}_\theta(\varphi(T(\boldsymbol{X}))) = 0$，但 $\varphi(\cdot) \not\equiv 0$，即 $T(\boldsymbol{X}) = (X_{(1)}, X_{(n)})$ 不是完全统计量。

Exercise 3

设 $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$ 其中 $\mu \in \mathbb{R}$, $\sigma^2 \in \mathbb{R}^+$。证明：$\bar{X}$ 和 $(X_{(n)} - X_{(1)})$ 相互独立。

证明 首先，写出 $X_1,\cdots,X_n$ 的联合密度为

$$f(\boldsymbol{x};\boldsymbol{\theta})=(2\pi\sigma^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}$$

易知可以写成因子分解形式，故 $\bar{X}$ 为充分统计量。然后写出其自然形式

$$f(\boldsymbol{x};\boldsymbol{\theta})=(2\pi\sigma^2)^{-n/2}e^{n\mu^2/2\sigma^2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2+\frac{\mu}{\sigma^2}\sum_{i=1}^{n}x_i\right\}$$

注意到自然参数空间为

$$\Theta^*=\left\{\boldsymbol{\theta}=\left(-\frac{1}{2\sigma^2},\frac{\mu}{\sigma^2}\right):-\frac{1}{2\sigma^2}\in(-\infty,0),\ \frac{\mu}{\sigma^2}\in(-\infty,+\infty)\right\}$$

指数族中，由 Θ^* 有内点，故 $(\sum X_i,\sum X_i^2)$ 为完全统计量。

记 $Y_i=(X_i-\mu)/\sigma$，于是 $Y_i\sim N(0,1)$ 与 $\boldsymbol{\theta}$ 无关。注意到 $X_{(n)}-X_{(1)}=Y_{(n)}-Y_{(1)}$ 的分布，同理与 $\boldsymbol{\theta}=(\mu,\sigma^2)$ 无关。则由 Basu 定理，$X_{(n)}-X_{(1)}$ 和 $(\sum X_i,\sum X_i^2)$ 相互独立，即有 $\bar{X}\perp\!\!\!\perp(X_{(n)}-X_{(1)})$，命题证毕。

Exercise 4

设总体 X 服从均匀分布 $U(0,2\theta)$，令 $X_1,\cdots,X_n$ 为从总体 X 中抽取的简单随机样本。

a 证明 $\hat{\theta}_1=\bar{X}$ 和 $\hat{\theta}_2=\frac{(n+1)X_{(n)}}{2n}$ 为 θ 的无偏估计。
b 证明 $\hat{\theta}_1$ 为 θ 的强相合估计，$\theta_2^*=\frac{X_{(n)}}{2}$ 为 θ 的弱相合估计。
c 求 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 的方差，比较它们的有效性。

证明 (a) 由 $X\sim U(0,2\theta)$ 故其概率密度为

$$f(x;\theta)=\frac{1}{2\theta}\cdot\mathbb{I}_{(0,2\theta)}(x)$$

于是可求 $\hat{\theta}_1$ 的期望

$$\mathbb{E}(\hat{\theta}_1)=\mathbb{E}(\bar{X})=\mathbb{E}(X_1)=\int_0^{2\theta}\frac{x}{2\theta}\,dx=\frac{(2\theta)^2}{4\theta}-0=\theta$$

故 $\hat{\theta}_1=\bar{X}$ 为 θ 的无偏估计。同理求 $\hat{\theta}_2=\frac{(n+1)X_{(n)}}{2n}$ 的期望。注意到 $X_{(n)}$ 的密度函数为

$$f_{X_{(n)}}(x)=n\left(\frac{x}{2\theta}\right)^{n-1}\frac{1}{2\theta}$$

所以有

$$\begin{aligned}\mathbb{E}(\hat{\theta}_2)&=\mathbb{E}\left(\frac{(n+1)X_{(n)}}{2n}\right)=\frac{n+1}{2n}\int_0^{2\theta}n\left(\frac{x}{2\theta}\right)^{n-1}\frac{1}{2\theta}\cdot x\,dx\\&=\frac{n+1}{2}\int_0^{2\theta}\left(\frac{x}{2\theta}\right)^{n}dx=\frac{n+1}{2}\left[\frac{2\theta}{n+1}\left(\frac{2\theta}{2\theta}\right)^{n+1}-0\right]\\&=\frac{n+1}{2}\frac{2\theta}{n+1}=\theta\end{aligned}$$

于是 $\hat{\theta}_2=\frac{(n+1)X_{(n)}}{2n}$ 为 θ 的无偏估计，命题证毕。

证明 (b) 由 $X \sim U(0,2\theta)$, $\theta > 0$ 知我们只考虑 $X > 0$ 的分布，即有 $\mathbb{E}(|X|) = \mathbb{E}(X) = \theta < \infty$。故由强大数定律可知

$$\hat{\theta}_1 = \bar{X} \xrightarrow{\text{a.s.}} \mathbb{E}(X_1) = \theta$$

即 $\hat{\theta}_1$ 为 θ 的强相合估计。下面证：$\theta_2^* = \frac{X_{(n)}}{2}$ 为 θ 的弱相合估计。

注意到 $X_{(n)}$ 的分布函数为

$$F_{X_{(n)}}(x) = \int_0^{\min\{2\theta,x\}} n\left(\frac{t}{2\theta}\right)^{n-1} \frac{1}{2\theta}\, dt = \left(\frac{t}{2\theta}\right)^n \Big|_0^{\min\{2\theta,x\}}$$

考虑不平凡的情况 $F_{X_{(n)}}(x) = \left(\frac{t}{2\theta}\right)^n\Big|_0^x \cdot \mathbb{I}(x < 2\theta) = \left(\frac{x}{2\theta}\right)^n \cdot \mathbb{I}(x < 2\theta)$。所以对 $\forall\, \epsilon > 0$ 和不平凡的 $X_{(n)} < 2\theta$ 有

$$\begin{aligned} P(|X_{(n)} - 2\theta| > \epsilon) &= P(X_{(n)} - 2\theta < -\epsilon) = P(X_{(n)} < 2\theta - \epsilon) = F_{X_{(n)}}(2\theta - \epsilon) \\ &= \left(\frac{2\theta - \epsilon}{2\theta}\right)^n = (1 - \epsilon/2\theta)^n \to 0 \quad \text{as } n \to \infty \end{aligned}$$

当 $\epsilon < 2\theta$ 成立（平凡地，当 $\epsilon > 2\theta$，显然有 $P(|X_{(n)} - 2\theta| > \epsilon) = 0$ 同样成立)。于是有对 $\forall\, \epsilon > 0$ 有

$$\lim_{n\to\infty} P(|X_{(n)} - 2\theta| \le \epsilon) = 1 \Rightarrow \lim_{n\to\infty} P(|X_{(n)}/2 - \theta| \le \epsilon/2) = 1$$

即 $\theta_2^* = X_{(n)}/2 \xrightarrow{P} \theta$，为 θ 的弱相合估计。

解答 (c) 因为 $X \sim U(0,2\theta)$ 故有方差为 $\mathbb{D}(X) = \frac{(2\theta-0)^2}{12} = \frac{\theta^2}{3}$。所以样本均值的方差为

$$\mathbb{D}(\hat{\theta}_1) = \mathbb{D}(\bar{X}) = \frac{1}{n^2}\sum_{i=1}^n \mathbb{D}(X_i) = \frac{n}{n^2}\cdot\frac{\theta^2}{3} = \frac{\theta^2}{3n}$$

由问题 a 计算的结果 $\mathbb{E}(X_{(n)}) = 2n\theta/(n+1)$，下面计算 $\mathbb{E}(X_{(n)}^2)$ 有

$$\begin{aligned} \mathbb{E}(X_{(n)}^2) &= \int_0^{2\theta} n\left(\frac{x}{2\theta}\right)^{n-1}\frac{1}{2\theta}\cdot x^2\, dx = \int_0^{2\theta} n\left(\frac{x}{2\theta}\right)^n \cdot x\, dx \\ &= 4n\theta^2 \int_0^{2\theta}\left(\frac{x}{2\theta}\right)^{n+1} d\frac{x}{2\theta} = 4n\theta^2 \cdot \frac{1}{n+2}\cdot 1 = \frac{4n\theta^2}{n+2} \end{aligned}$$

于是有方差为

$$\mathbb{D}(X_{(n)}) = \mathbb{E}(X_{(n)}^2) - [\mathbb{E}(X_{(n)})]^2 = \frac{4n\theta^2}{n+2} - \left(\frac{2n\theta}{n+1}\right)^2 = \frac{4n\theta^2}{(n+1)^2(n+2)}$$

所以 $\hat{\theta}_2$ 的方差为

$$\mathbb{D}(\hat{\theta}_2) = \mathbb{D}\left(\frac{(n+1)X_{(n)}}{2n}\right) = \frac{(n+1)^2}{4n^2}\frac{4n\theta^2}{(n+1)^2(n+2)} = \frac{\theta^2}{n(n+2)}$$

因为二者都是无偏的，故比较二者方差，在任意 θ 下有

$$\mathbb{D}(\hat{\theta}_1) = \frac{\theta^2}{3n} \sim O(1/n) \qquad \mathbb{D}(\hat{\theta}_2) = \frac{\theta^2}{n(n+2)} \sim O(1/n^2)$$

当 $n > 1$ 时 $\mathbb{D}(\hat{\theta}_2) < \mathbb{D}(\hat{\theta}_1)$，故 $\hat{\theta}_2$ 比 $\hat{\theta}_1$ 更有效。

HW6 第六次作业解答

更新：2026 年 1 月 9 日

Exercise 1

设 $X_1, \cdots, X_n$ 为抽自对数级数分布的简单随机样本，

$$P(X=k) = -\frac{1}{\log(1-p)} \cdot \frac{p^k}{k}, \quad k = 1, 2, \cdots$$

其中参数 $p \in (0,1)$，求 p 的矩估计。

解答 记 k 阶矩为 $\alpha_k = \mathbb{E}\{X^k\}$，样本 k 阶矩为 $a_{nk} = \sum_{i=1}^n X_i^k / n$，于是有

$$\alpha_1 = \sum_{k=1}^{\infty} k \cdot P(X=k) = -\frac{1}{\log(1-p)} \sum_{k=1}^{\infty} k \cdot \frac{p^k}{k} = -\frac{1}{\log(1-p)} \sum_{k=1}^{\infty} p^k = -\frac{1}{\log(1-p)} \cdot \frac{p}{1-p}$$

于是使用一阶样本矩估计有

$$a_{n1} = -\frac{1}{\log(1-\hat{p})} \cdot \frac{\hat{p}}{1-\hat{p}} \tag{1}$$

同理求二阶矩

$$\alpha_2 = \sum_{k=1}^{\infty} k^2 \cdot P(X=k) = -\frac{1}{\log(1-p)} \sum_{k=1}^{\infty} k^2 \cdot \frac{p^k}{k} = -\frac{1}{\log(1-p)} \sum_{k=1}^{\infty} k \cdot p^k$$

记 $M_k = -\log(1-p)\alpha_k$，于是注意到

$$M_2 - p \cdot M_2 = \sum_{k=1}^{\infty} k \cdot p^k - \sum_{k=1}^{\infty} k \cdot p^{k+1} = \sum_{k=1}^{\infty} k \cdot p^k - \sum_{k=2}^{\infty} (k-1) \cdot p^k = p + \sum_{k=2}^{\infty} p^k = p + \frac{p^2}{1-p}$$

于是有

$$\alpha_2 = -\frac{1}{\log(1-p)} \cdot M_2 = -\frac{1}{\log(1-p)} \cdot \frac{p}{(1-p)^2}$$

所以使用二阶样本矩估计有

$$a_{n2} = -\frac{1}{\log(1-\hat{p})} \cdot \frac{\hat{p}}{(1-\hat{p})^2} \tag{2}$$

联合使用 a_{n1}, a_{n2} 进行估计，即求解方程式 1 和式 2

$$\begin{cases} a_{n1} = -\dfrac{1}{\log(1-\hat{p})} \cdot \dfrac{\hat{p}}{1-\hat{p}} \\ a_{n2} = -\dfrac{1}{\log(1-\hat{p})} \cdot \dfrac{\hat{p}}{(1-\hat{p})^2} \end{cases}$$

于是有 $a_{n1}/a_{n2}=1-\hat{p}$，即 p 的矩估计为

$$\hat{p}=1-\frac{a_{n1}}{a_{n2}}=1-\frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n X_i^2}$$

Exercise 2

设 $X_1,\cdots,X_n$ 是来自总体 $N(0,\sigma^2)$ 的简单随机样本，求 σ 的矩估计。

1. 利用 $\mathbb{E}\{|X_1|\}=\sigma\sqrt{2/\pi}$
2. 利用 $\sigma=\sqrt{\mathbb{D}\{X_1\}}$

解答 由 $\mathbb{E}\{|X_1|\}=\sigma\sqrt{2/\pi}$ 故使用有关 $|X_i|$ 的矩进行估计

$$\hat{\sigma}^2=\sqrt{\frac{\pi}{2}}\cdot\frac{1}{n}\sum_{i=1}^n|X_i|$$

类似地，$\sigma=\sqrt{\mathbb{D}\{X_1\}}$，利用 m_{n2} 估计 $\mathbb{D}\{X_1\}$ 得到 σ 的估计

$$\hat{\sigma}^2=\sqrt{\frac{1}{n}\sum_{i=1}^n(X_i-\bar{X})^2}$$

注 记 $Y_i=|X_i|$，于是有 Y_i 独立同分布，且

$$\begin{aligned}\mathbb{E}\{Y_1\}&=\int_{\{y=|x|:\ x\in\mathbb{R}\}}(2\pi\sigma^2)^{-1/2}e^{\frac{y^2}{2\sigma^2}}\,dy\\&=\int_{\{y=x:\ x>0\}}(2\pi\sigma^2)^{-1/2}e^{\frac{y^2}{2\sigma^2}}\,dy+\int_{\{y=-x:\ x<0\}}(2\pi\sigma^2)^{-1/2}e^{\frac{y^2}{2\sigma^2}}\,dy\\&=2\int_0^\infty(2\pi\sigma^2)^{-1/2}e^{\frac{y^2}{2\sigma^2}}\,dy=\sigma\sqrt{2/\pi}<\infty\end{aligned}$$

由强大数定理可知

$$\frac{1}{n}\sum_{i=1}^n|X_i|=\frac{1}{n}\sum_{i=1}^nY_i\xrightarrow{\text{a.s.}}\mathbb{E}\{Y_1\}=\sigma\sqrt{2/\pi}$$

所以使用 $|X_i|$ 相关矩的矩估计，其相合性得到保证。

Exercise 3

若 $X=e^\xi$，而 $\xi\sim N(\mu,\sigma^2)$，则 X 的分布称为对数正态分布。求 X 的概率密度函数。设 $X_1,\cdots,X_n$ 是从总体 X 中抽出的简单随机样本，求 μ 和 σ^2 的矩估计和 MLE。

解答 由 $\xi \sim N(\mu, \sigma^2)$ 可知 ξ 的概率密度函数为

$$f_\Xi(\xi; \mu, \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{(\xi-\mu)^2}{2\sigma^2}\right\}$$

而 $\xi = \log X$ 故有 X 的概率密度函数为

$$f_X(x; \mu, \sigma^2) = f_\Xi(\xi(x); \mu, \sigma^2) \cdot |d\xi/dx| = (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{(\log x-\mu)^2}{2\sigma^2}\right\} x^{-1}$$

(1) 求解矩估计。根据密度函数，可求 k 阶矩（注意 $X = e^\xi > 0$），例如

$$\begin{aligned}
\mathbb{E}\{X\} &= \int_0^\infty (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{(\log x-\mu)^2}{2\sigma^2}\right\} dx \\
&\xlongequal{\log x = t} \int_{-\infty}^{+\infty} (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{(t-\mu)^2}{2\sigma^2}\right\} e^t\, dt \\
&= \int_{-\infty}^{+\infty} (2\pi\sigma^2)^{-1/2} \exp\left\{\frac{2\sigma^2 t - (t^2 - 2\mu t + \mu^2)}{2\sigma^2}\right\} dt
\end{aligned}$$

注意到

$$2\sigma^2 t - (t^2 - 2\mu t + \mu^2) = -[t - (\mu + \sigma^2)]^2 + (\mu + \sigma^2)^2 - \mu^2 = -(t - \mu - \sigma^2)^2 + 2\mu\sigma^2 + \sigma^4$$

代回 $\mathbb{E}\{X\}$ 有

$$\mathbb{E}\{X\} = e^{\mu+\sigma^2/2} \int_{-\infty}^{+\infty} (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{(t-\mu-\sigma^2)^2}{2\sigma^2}\right\} dt$$

积分中函数可视为正态分布密度函数在 $\mathbb{R}$ 的积分，即为 1，所以 $\mathbb{E}\{X\} = e^{\mu+\sigma^2/2}$。类似地，二阶矩为

$$\begin{aligned}
\mathbb{E}\{X^2\} &= \int_0^\infty (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{(\log x-\mu)^2}{2\sigma^2}\right\} x\, dx \\
&\xlongequal{\log x = t} \int_{-\infty}^{+\infty} (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{(t-\mu)^2}{2\sigma^2}\right\} e^{2t}\, dt \\
&= e^{2\mu+2\sigma^2} \int_{-\infty}^{+\infty} (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{(t-\mu-2\sigma^2)^2}{2\sigma^2}\right\} dt
\end{aligned}$$

故 $\mathbb{E}\{X^2\} = e^{2\mu+2\sigma^2}$。于是解下列方程得到矩和参数的关系

$$\begin{cases}
\mu + \sigma^2/2 = \log \mathbb{E}\{X\} \leftarrow \log a_{n1} \\
2\mu + 2\sigma^2 = \log \mathbb{E}\{X^2\} \leftarrow \log a_{n2}
\end{cases}$$

于是得到矩估计为

$$\begin{cases}
\hat{\mu} = -\log \sqrt{a_{n2}} = -\log \sqrt{\dfrac{1}{n} \sum_{i=1}^n X_i^2} \\
\hat{\sigma}^2 = \log \dfrac{a_{n2}}{a_{n1}^2} = \log \dfrac{n \sum_{i=1}^n X_i^2}{(\sum_{i=1}^n X_i)^2}
\end{cases}$$

特别地，这里的矩估计不唯一。例如，如果使用 m_{nk} 估计则得到的表达式不同。

$$e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2} = \mathbb{E}\{X^2\} - [\mathbb{E}\{X\}]^2 = \mu_2 \leftarrow m_{n2}$$

于是有 $[\mathbb{E}\{X\}]^2(e^\sigma-1)=e^{2\mu+\sigma^2}(e^\sigma-1)=\mu_2$，于是有 $e^\sigma-1=\mu_2/\alpha_1^2$，那么另一种矩估计为

$$\begin{cases}\hat{\mu}=\log\alpha_1-\dfrac{1}{2}\hat{\sigma}^2=\log\bar{X}-\log\left(\sqrt{1+\dfrac{n\sum_{i=1}^n(X_i-\bar{X})^2}{(\sum_{i=1}^n X_i)^2}}\right)\\ \hat{\sigma}^2=\log\left(1+\dfrac{m_{n2}}{a_{n1}^2}\right)=\log\left(1+\dfrac{n\sum_{i=1}^n(X_i-\bar{X})^2}{(\sum_{i=1}^n X_i)^2}\right)\end{cases}$$

(2) 求解极大似然估计。已知概率密度，故似然函数为

$$L(\mu,\sigma^2;\boldsymbol{x})=(2\pi\sigma^2)^{-n/2}\exp\left\{\sum_{i=1}^n-\frac{(\log x_i-\mu)^2}{2\sigma^2}\right\}\prod_{i=1}^n x_i^{-1}$$

那么对数似然函数为

$$l(\mu,\sigma^2;\boldsymbol{x})=-\frac{n}{2}\log(2\pi\sigma^2)-\sum_{i=1}^n\frac{(\log x_i-\mu)^2}{2\sigma^2}-\sum_{i=1}^n\log x_i$$

分别求偏导有

$$\frac{\partial l(\mu,\sigma^2)}{\partial\mu}=\sum_{i=1}^n\frac{2(\log x_i-\mu)}{2\sigma^2}=0\quad\Leftrightarrow\quad\sum_{i=1}^n\log x_i=n\mu$$

所以有 μ 的 MLE 为 $\hat{\mu}_{\text{MLE}}=\frac{1}{n}\sum_{i=1}^n\log X_i$。代回对数似然，并对 σ^2 求偏导有

$$\frac{\partial l(\mu,\sigma^2)}{\partial\sigma^2}=-\frac{n}{2}\frac{2\pi}{2\pi\sigma^2}+\sum_{i=1}^n\frac{2(\log x_i-\hat{\mu}_{\text{MLE}})^2}{4\sigma^4}=0$$

于是有 σ^2 的 MLE 为 $\hat{\sigma}^2_{\text{MLE}}=\frac{1}{n}\sum_{i=1}^n(\log X_i-\hat{\mu}_{\text{MLE}})^2$。

Exercise 4

设 $X_1,\cdots,X_n$ 是来自双参数指数分布

$$f(x;\mu,\sigma)=\frac{1}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\}\cdot\mathbb{I}(x\geq\mu)$$

的简单随机样本，其中 $\mu\in\mathbb{R}$，$\sigma\in\mathbb{R}^+$。记 $\theta=(\mu,\sigma)$，求 μ,σ 和 $P_\theta(X_1\geq t)$（其中 $t>\mu$ 已知）的矩估计和 MLE。

解答 记 $t=(x-\mu)/\sigma$ 于是我们得到下列积分结果

$$\mathbb{E}\{X_1\}=\int_\mu^\infty\frac{x}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\}dx=\int_0^\infty(\mu+\sigma t)e^{-t}\,dt=\mu+\sigma$$
$$\mathbb{E}\{X_1^2\}=\int_\mu^\infty\frac{x^2}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\}dx=\int_0^\infty(\mu+\sigma t)^2e^{-t}\,dt=\mu^2+2\sigma\mu+2\sigma^2$$

(1) 求解矩估计：注意到 $\mathbb{E}\{X_1^2\}=\mu^2+2\sigma\mu+2\sigma^2=(\mu+\sigma)^2+\sigma^2=[\mathbb{E}\{X_1\}]^2+\sigma^2$，即 $\sigma^2=\mathbb{E}\{X_1^2\}-[\mathbb{E}\{X_1\}]^2=\mu_2$，于是可得 σ 的矩估计 $\hat{\sigma}_{\text{MM}}=\sqrt{m_{n2}}$。代回可得 μ 的矩估计 $\hat{\mu}_{\text{MM}}=a_{n1}-\sqrt{m_{n2}}$。下面计算 $P_\theta(X_1\geq t)$

$$P_\theta(X_1\geq t)=\int_t^\infty\frac{1}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\}dx=-\exp\left\{-\frac{x-\mu}{\sigma}\right\}\Bigg|_t^\infty=\exp\left\{-\frac{t-\mu}{\sigma}\right\}$$

结合矩估计的性质，可以得到以下矩估计

$$\begin{cases} \hat{\mu}_{\text{MM}} = a_{n1} - \sqrt{m_{n2}} = \bar{X} - \sqrt{\dfrac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2} \\ \hat{\sigma}_{\text{MM}} = \sqrt{m_{n2}} = \sqrt{\dfrac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2} \\ \hat{P}_{\text{MM}}(X_1 \geq t) = \exp\left\{-\dfrac{t - \hat{\mu}_{\text{MM}}}{\hat{\sigma}_{\text{MM}}}\right\} \end{cases}$$

(2) 求解似极大然估计：已知概率密度函数，可求似然函数

$$L(\mu, \sigma; \boldsymbol{x}) = \prod_{i=1}^{n} \frac{1}{\sigma} \exp\left\{-\frac{x_i - \mu}{\sigma}\right\} \cdot \mathbb{I}(x_i \geq \mu) = \frac{1}{\sigma^n} \exp\left\{\sum_{i=1}^{n} -\frac{x_i - \mu}{\sigma}\right\} \cdot \mathbb{I}(x_{(1)} \geq \mu)$$

于是对数似然为

$$\begin{aligned} l(\mu, \sigma; \boldsymbol{x}) &= \left(-n \log \sigma - \sum_{i=1}^{n} \frac{x_i - \mu}{\sigma}\right) \cdot \mathbb{I}(x_{(1)} \geq \mu) \\ &= \left(-n \log \sigma - \sigma^{-1} \sum_{i=1}^{n} x_i + n\mu\sigma^{-1}\right) \cdot \mathbb{I}(x_{(1)} \geq \mu) \end{aligned}$$

求解下列方程

$$\begin{cases} \dfrac{\partial l(\mu, \sigma)}{\partial \mu} = n\sigma^{-1} > 0 \\ \dfrac{\partial l(\mu, \sigma)}{\partial \sigma} = -n\sigma^{-1} + \sigma^{-2} \sum_{i=1}^{n} x_i - n\mu\sigma^{-2} = 0 \end{cases}$$

由 $\partial l / \partial \mu > 0$ 可知，极大化似然函数，需要 μ 尽可能大，但为了满足 $X_{(1)} \geq \mu$，故有 μ 的 MLE 为 $\hat{\mu}_{\text{MLE}} = X_{(1)}$。同时，求解第二个方程有 σ 的 MLE 为 $\hat{\sigma}_{\text{MLE}} = \bar{X} - \hat{\mu}_{\text{MLE}} = \bar{X} - X_{(1)}$。综上有极大似然估计为

$$\begin{cases} \hat{\mu}_{\text{MLE}} = X_{(1)} \\ \hat{\sigma}_{\text{MLE}} = \bar{X} - X_{(1)} \\ \hat{P}_{\text{MEL}}(X_1 \geq t) = \exp\left\{-\dfrac{t - \hat{\mu}_{\text{MLE}}}{\hat{\sigma}_{\text{MLE}}}\right\} \end{cases}$$

HW7 第七次作业解答

更新：2026 年 1 月 9 日

Exercise 1

设 $(X_1, \cdots, X_n)$ 是来自均匀分布 $U[\theta, 2\theta]$ 的简单随机样本，其中 $\theta \in (0, +\infty)$，求 θ 的 MLE。判断它是否为 θ 的无偏估计。如果不是，试对它略作修改，得到 θ 的一个无偏估计。

解答 由于 $X_i \sim U[\theta, 2\theta]$，其概率密度函数为

$$f(x;\theta) = \frac{1}{\theta} \cdot \mathbb{I}(\theta \le x \le 2\theta)$$

因此，样本的对数似然函数为

$$l(\theta) = \log L(\theta) = \log\left(\prod_{i=1}^{n} f(x_i;\theta)\right) = -n\log\theta \cdot \mathbb{I}(\theta \le x_{(1)} \le x_{(n)} \le 2\theta)$$

容易看出，当 $\theta \le x_{(1)} \le x_{(n)} \le 2\theta$ 时，$l(\theta)$ 随 θ 增大而减小，因此，$l(\theta)$ 在区间 $x_{(n)}/2$ 上取得最大值。故 θ 的 MLE 为 $\hat{\theta}_{\text{MLE}} = \frac{X_{(n)}}{2}$。

接下来计算 $\hat{\theta}_{\text{MLE}}$ 的期望，由次序统计量的性质可知，最大值 $X_{(n)}$ 的概率密度函数为

$$f_{X_{(n)}}(x) = \frac{n}{\theta^n}(x-\theta)^{n-1} \cdot \mathbb{I}(\theta \le x \le 2\theta)$$

因此，

$$\begin{aligned}
\mathbb{E}[\hat{\theta}_{\text{MLE}}] = \mathbb{E}\left[\frac{X_{(n)}}{2}\right] &= \int_{\theta}^{2\theta} \frac{x}{2} \cdot f_{X_{(n)}}(x)\,\mathrm{d}x = \int_{\theta}^{2\theta} \frac{x}{2} \cdot \frac{n}{\theta^n}(x-\theta)^{n-1}\,\mathrm{d}x \\
&\xlongequal{t=x-\theta} \int_0^{\theta} \frac{t+\theta}{2} \cdot \frac{n}{\theta^n} t^{n-1}\,\mathrm{d}t = \frac{n}{2\theta^n}\int_0^{\theta} t^n\,\mathrm{d}t + \frac{n\theta}{2\theta^n}\int_0^{\theta} t^{n-1}\,\mathrm{d}t \\
&= \frac{n}{2(n+1)}\theta + \frac{1}{2}\theta = \frac{2n+1}{2(n+1)}\theta
\end{aligned}$$

故 $\hat{\theta}_{\text{MLE}}$ 不是 θ 的无偏估计。但可以通过修正得到 θ 的一个无偏估计：

$$\hat{\theta}_{\text{U}} = \frac{2(n+1)}{2n+1} \cdot \hat{\theta}_{\text{MLE}} = \frac{n+1}{2n+1} X_{(n)}$$

Exercise 2

设 $X_1, \cdots, X_n$ 为抽自下列指数分布的简单随机样本，

$$f(x;\mu) = e^{-(x-\mu)} \cdot \mathbb{I}(x > \mu), \quad \mu \in \mathbb{R}$$

(1) 试求 μ 的极大似然估计 $\hat{\mu}_{\text{MLE}}$ 是 μ 的无偏估计吗？如果不是，试对它作修改，以得到 μ 的无偏估计 $\hat{\mu}_{\text{U}}$。

(2) 试求 μ 的矩估计 $\hat{\mu}_{\text{MM}}$，并证明它是 μ 的无偏估计。

(3) 试问 $\hat{\mu}_{\text{U}}$ 和 $\hat{\mu}_{\text{MM}}$ 哪一个更有效？

解答 (1) 样本的对数似然函数为

$$l(\mu) = \log L(\mu) = \log\left(\prod_{i=1}^{n} f(x_i;\mu)\right) = -\sum_{i=1}^{n}(x_i - \mu) \cdot \mathbb{I}(\mu < x_{(1)})$$

容易看出，当 $\mu < x_{(1)}$ 时，$l(\mu)$ 随 μ 增大而增大，因此，μ 的 MLE 为 $\hat{\mu}_{\text{MLE}} = X_{(1)}$。

记 $Y_i = X_i - \mu$，则 Y_i 的概率密度函数为 $f(y) = e^{-y} \cdot \mathbb{I}(y > 0) \sim \text{Exp}(1)$。因此，$Y_{(1)} = X_{(1)} - \mu$ 的概率密度函数为

$$f_{Y_{(1)}}(y) = ne^{-ny} \cdot \mathbb{I}(y > 0) \sim \text{Exp}(n)$$

故 $\mathbb{E}[Y_{(1)}] = 1/n$，所以

$$\mathbb{E}[\hat{\mu}_{\text{MLE}}] = \mathbb{E}[X_{(1)}] = \mathbb{E}[Y_{(1)}] + \mu = 1/n + \mu$$

故 $\hat{\mu}_{\text{MLE}}$ 不是 μ 的无偏估计。但可以通过修正得到 μ 的一个无偏估计 $\hat{\mu}_{\text{U}} = X_{(1)} - 1/n$。

解答 (2) 由 $Y_i \sim \text{Exp}(1)$，所以有 $\mathbb{E}[Y_i] = 1 = \mathbb{E}[X_i] - \mu$，因此，$\mu$ 的矩估计为 $\hat{\mu}_{\text{MM}} = \bar{X} - 1$ 而

$$\mathbb{E}[\hat{\mu}_{\text{MM}}] = \mathbb{E}[\bar{X}] - 1 = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_i] - 1 = \frac{1}{n}\sum_{i=1}^{n}(\mu + 1) - 1 = \mu$$

显然，$\hat{\mu}_{\text{MM}}$ 是 μ 的无偏估计。

解答 (3) 由于 $\hat{\mu}_{\text{U}}$ 和 $\hat{\mu}_{\text{MM}}$ 都是 μ 的无偏估计，因此，只需比较它们的方差即可。由 $Y_{(1)} \sim \text{Exp}(n)$ 可知 $\text{Var}(Y_{(1)}) = 1/n^2$，因此，

$$\text{Var}(\hat{\mu}_{\text{U}}) = \text{Var}(X_{(1)} - 1/n) = \text{Var}(Y_{(1)}) = 1/n^2$$

又因为 X_i 相互独立同分布，

$$\text{Var}(\hat{\mu}_{\text{MM}}) = \text{Var}(\bar{X} - 1) = \text{Var}(\bar{X}) = \frac{1}{n^2}\sum_{i=1}^{n}\text{Var}(X_i) = \frac{1}{n}\text{Var}(X_i)$$

而 $X_i = Y_i + \mu$，且 $Y_i \sim \text{Exp}(1)$，所以 $\text{Var}(X_i) = \text{Var}(Y_i) = 1$。因此，$\text{Var}(\hat{\mu}_{\text{MM}}) = 1/n$。故有

$$\text{Var}(\hat{\mu}_{\text{U}}) = \frac{1}{n^2} < \frac{1}{n} = \text{Var}(\hat{\mu}_{\text{MM}})$$

故 $\hat{\mu}_{\text{U}}$ 比 $\hat{\mu}_{\text{MM}}$ 更有效。

Exercise 3

设 $X_1,\cdots,X_n \overset{\text{i.i.d.}}{\sim} Ge(\theta)$ 分布族

$$P(X=i)=\theta(1-\theta)^{i-1},\quad i=1,2,\cdots,\ \theta\in(0,1)$$

(1) 试证明 $T=\sum_{i=1}^n X_i$ 是 θ 的充分完全统计量，且服从 Pascal（负二项）分布

$$P_\theta(T=t)=\binom{t-1}{n-1}\theta^n(1-\theta)^{t-n},\quad t=n,n+1,\cdots$$

(2) 计算 $\mathbb{E}_\theta(T)$, 并由此求 θ^{-1} 的 UMVUE。

(3) 试证明 $\psi(X_1)=\mathbb{I}(X_1=1)$ 是 θ 的无偏估计，计算 $\mathbb{E}_\theta(\psi(X_1)|T=t)$，并由此求得 θ 的 UMVUE。

证明 (1) 由 $X_1,\cdots,X_n \overset{\text{i.i.d.}}{\sim} Ge(\theta)$ 可知联合概率为

$$P(X_1=x_1,\cdots,X_n=x_n)=\prod_{i=1}^n\theta(1-\theta)^{x_i-1}=\theta^n(1-\theta)^{\sum_{i=1}^n x_i-n}$$

所以给定 $T=\sum_{i=1}^n X_i$ 的概率为

$$P(X_1=x_1,\cdots,X_n=x_n|T=t)=\frac{P(X_1=x_1,\cdots,X_n=x_n)}{P(T=t)}=\frac{\theta^n(1-\theta)^{t-n}}{P(T=t)}$$

而 $T=\sum_{i=1}^n X_i$ 其中 $X_i\geq 1$，由隔板法可知共 $\binom{t-1}{n-1}$ 个解满足条件。所以

$$P(T=t)=\sum_{\sum_{i=1}^n x_i=t}P(X_1=x_1,\cdots,X_n=x_n)=\binom{t-1}{n-1}\theta^n(1-\theta)^{t-n}$$

所以 T 服从负二项分布。同时

$$P(X_1=x_1,\cdots,X_n=x_n|T=t)=\frac{\theta^n(1-\theta)^{t-n}}{P(T=t)}=\frac{\theta^n(1-\theta)^{t-n}}{\binom{t-1}{n-1}\theta^n(1-\theta)^{t-n}}=\binom{t-1}{n-1}^{-1}$$

与 θ 无关，由充分统计量定义可知 $T=\sum_{i=1}^n X_i$ 是 θ 的充分统计量。

解答 (2) 由已知 T 的分布列，故有

$$\mathbb{E}_\theta(T)=\sum_{t=n}^\infty t\binom{t-1}{n-1}\theta^n(1-\theta)^{t-n}$$

由分布列的性质

$$\sum_{t=n}^\infty P(T=t)=1\quad\Rightarrow\quad\sum_{t=n}^\infty\binom{t-1}{n-1}\theta^n(1-\theta)^{t-n}=1$$

记 $S(\theta)=\sum_{t=n}^\infty\binom{t-1}{n-1}(1-\theta)^{t-n}=\theta^{-n}$，注意到

$$S'(\theta)=-\sum_{t=n}^\infty(t-n)\binom{t-1}{n-1}(1-\theta)^{t-n-1}=-\frac{n}{\theta^{n+1}}$$

于是有

$$\sum_{t=n}^{\infty} t\binom{t-1}{n-1}(1-\theta)^{t-n} = -S'(\theta)\cdot(1-\theta)+nS(\theta) = \frac{n(1-\theta)}{\theta^{n+1}}+\frac{n}{\theta^n}$$

所以

$$\mathbb{E}_\theta(T) = \theta^n\left(\frac{n(1-\theta)}{\theta^{n+1}}+\frac{n}{\theta^n}\right) = n(1-\theta)\theta^{-1}+n = n\theta^{-1}$$

同时，根据 Lehmann-Scheffe 定理，T 是充分完全统计量，而 $\mathbb{E}_\theta(T/n)=\theta^{-1}$ 是无偏的，所以 θ^{-1} 的 UMVUE 为 $T/n = \frac{1}{n}\sum_{i=1}^n X_i$。

证明 (3) 容易得

$$\mathbb{E}_\theta(\psi(X_1)) = \mathbb{E}_\theta(\mathbb{I}(X_1=1)) = P(X_1=1)=\theta$$

所以 $\psi(X_1)$ 是 θ 的无偏估计。而

$$\begin{aligned}\mathbb{E}_\theta(\psi(X_1)|T=t) = 1\cdot P(X_1=1|T=t) &= \frac{P(X_1=1, X_2+\cdots+X_n=t-1)}{P(T=t)}\\ &= \frac{P(X_1=1)P(T-X_1=t-1)}{P(T=t)}\end{aligned}$$

而 T 服从负二项分布，所以

$$\mathbb{E}_\theta(\psi(X_1)|T=t) = \frac{\theta\binom{t-2}{n-2}\theta^{n-1}(1-\theta)^{t-n}}{\binom{t-1}{n-1}\theta^n(1-\theta)^{t-n}} = \frac{n-1}{t-1},\quad t=n,n+1,\cdots$$

下面求 $\mathbb{E}_\theta(\psi(X_1)|T=t)$ 的期望

$$\mathbb{E}_\theta[\mathbb{E}_\theta(\psi(X_1)|T=t)] = \sum_{t=n}^{\infty}\frac{n-1}{t-1}\binom{t-1}{n-1}\theta^n(1-\theta)^{t-n} = \sum_{t=n}^{\infty}\binom{t-2}{n-2}\theta^n(1-\theta)^{t-n}$$

由广义二项式展开知

$$\sum_{t=n}^{\infty}\binom{t-2}{n-2}(1-\theta)^{t-n} = \sum_{t=0}^{\infty}\binom{t+n-2}{n-2}(1-\theta)^t = [1-(1-\theta)]^{-(n-1)}$$

代回则得

$$\mathbb{E}_\theta[\mathbb{E}_\theta(\psi(X_1)|T=t)] = \theta^n[1-(1-\theta)]^{-(n-1)} = \theta$$

所以 $g(T)=\mathbb{E}_\theta(\psi(X_1)|T=t)$ 是 θ 的无偏估计，而 T 为充分完全统计量，结合 Lehmann-Scheffe 定理，θ 的 UMVMU 为 $g(T)=\mathbb{E}_\theta(\psi(X_1)|T=t)$。

Exercise 4

设 $X_1,\cdots,X_n \overset{\text{i.i.d.}}{\sim} b(1,p)$ 两点分布，$0<p<1$ 是未知参数，试求：

(1) p^s 的 UMVUE。

(2) $p^s+(1-p)^{n-s}$ 的 UMVUE，其中 $s\in(0,n)$ 为整数。

解答 (1) 设 $T=\sum_{i=1}^n X_i$，于是 $T\sim B(n,p)$，那么

$$
\begin{aligned}
P(X_1=x_1,\cdots,X_n=x_n|T=t)&=\frac{P(X_1=x_1,\cdots,X_n=t-\sum_{i=1}^{n-1}x_i)}{P(T=t)}\\
&=\frac{\left(\prod_{i=1}^{n-1}p^{x_i}(1-p)^{1-x_i}\right)\cdot p^{t-\sum_{i=1}^{n-1}x_i}(1-p)^{1-t+\sum_{i=1}^{n-1}x_i}}{\binom{n}{t}p^t(1-p)^{n-t}}\\
&=\frac{p^t(1-p)^{n-t}}{\binom{n}{t}p^t(1-p)^{n-t}}=\binom{n}{t}^{-1}
\end{aligned}
$$

与 p 无关，故 T 为充分统计量。对任意函数 $\varphi(T)$ 的期望为

$$
\mathbb{E}_p(\varphi(T))=\sum_{t=0}^n\varphi(t)\binom{n}{t}p^t(1-p)^{n-t}=(1-p)^n\sum_{t=0}^n\varphi(t)\binom{n}{t}\left(\frac{p}{1-p}\right)^t
$$

当 $\mathbb{E}_p(\varphi(T))=0$ 对任意 p 成立时，由多项式的性质，系数 $\varphi(t)\binom{n}{t}=0$，于是 $\varphi(t)=0$ a.s. P，于是 T 是充分完全统计量。于是，由 Lehmann-Scheffe 定理，下面只需寻找 $g(T)$ 使得 $g(T)$ 是无偏估计即可。

对任意 $n\in\mathbb{N}^+$，定义

$$
n^{(s)}=\frac{n!}{(n-s)!}=n(n-1)(n-2)\cdots(n-s+1)
$$

当 $n<s$ 时 $n^{(s)}=0$。下面我们断言：$\mathbb{E}(T^{(s)})=n^{(s)}p^s$。

$$
\begin{aligned}
\mathbb{E}(T^{(s)})&=\sum_{t=0}^n t^{(s)}\binom{n}{t}p^t(1-p)^{n-t}=\sum_{t=s}^n\frac{t!}{(t-s)!}\cdot\frac{n!}{t!(n-t)!}p^t(1-p)^{n-t}\\
&=\sum_{t=s}^n\frac{n!}{(t-s)!(n-t)!}p^t(1-p)^{n-t}=\sum_{t=s}^n\frac{(n-s)!}{(t-s)!(n-t)!}\cdot\frac{n!}{(n-s)!}p^t(1-p)^{n-t}\\
&=n^{(s)}\sum_{t=s}^n\binom{n-s}{t-s}p^t(1-p)^{n-t}=n^{(s)}p^s\sum_{t=s}^n\binom{n-s}{t-s}p^{t-s}(1-p)^{(n-s)-(t-s)}\\
&=n^{(s)}p^s\sum_{t=0}^{n-s}\binom{n-s}{t}p^t(1-p)^{(n-s)-t}=n^{(s)}p^s[p+(1-p)]^{n-s}=n^{(s)}p^s
\end{aligned}
$$

断言证毕。于是我们得到 p^s 的 UMVUE 为 $g(T)=\frac{T^{(s)}}{n^{(s)}}=\frac{T!(n-s)!}{n!(T-s)!}$。

解答 (2) 注意到 $n-T\sim B(n,1-p)$，所以与上面同理 $(1-p)^{n-s}$ 的 UMVUE 为 $\frac{(n-T)^{(n-s)}}{n^{(n-s)}}$。于是 $p^s+(1-p)^{n-s}$ 的 UMVUE 为

$$
\begin{aligned}
\frac{T^{(s)}}{n^{(s)}}+\frac{(n-T)^{(n-s)}}{n^{(n-s)}}&=\frac{T!(n-s)!}{n!(T-s)!}+\frac{s!(n-T)!}{n!(s-T)!}\\
&=\frac{T!(n-s)!}{n!(T-s)!}\cdot\mathbb{I}(T\geq s)+\frac{s!(n-T)!}{n!(s-T)!}\cdot\mathbb{I}(T\leq s)
\end{aligned}
$$

HW8 第八次作业解答

更新：2026 年 1 月 9 日

Exercise 1

设 $X_1, \cdots, X_m \overset{\text{i.i.d.}}{\sim} N(a, \sigma^2)$ 以及 $Y_1, \cdots, Y_n \overset{\text{i.i.d.}}{\sim} N(a, 2\sigma^2)$ 且两样本独立。求 a 和 σ^2 的 UMVUE。

解答 由正态分布，可得概率函数

$$
\begin{aligned}
f(\boldsymbol{x}, \boldsymbol{y}; a, \sigma^2) &= (2\pi\sigma^2)^{-m/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^m (x_i - a)^2\right\} (4\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{4\sigma^2}\sum_{j=1}^n (y_j - a)^2\right\} \\
&= C(a, \sigma^2) \exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^m x_i^2 + \frac{1}{2}\sum_{j=1}^n y_j^2\right) + \frac{a}{\sigma^2}\left(\sum_{i=1}^m x_i + \frac{1}{2}\sum_{j=1}^n y_j\right)\right\}
\end{aligned}
$$

容易证明 $T_1 = \sum_{i=1}^m x_i^2 + \frac{1}{2}\sum_{j=1}^n y_j^2$ 和 $T_2 = \sum_{i=1}^m x_i + \frac{1}{2}\sum_{j=1}^n y_j$ 是 a, σ^2 的充分完全统计量（这是因为有内点）。于是

$$
\begin{aligned}
\mathbb{E}(T_2) &= \mathbb{E}\left(\sum_{i=1}^m x_i + \frac{1}{2}\sum_{j=1}^n y_j\right) = \left(m + \frac{n}{2}\right) a \\
\operatorname{Var}(T_2) &= \operatorname{Var}\left(\sum_{i=1}^m x_i + \frac{1}{2}\sum_{j=1}^n y_j\right) = \left(m + \frac{n}{2}\right) \sigma^2
\end{aligned}
$$

而

$$
\begin{aligned}
\mathbb{E}(T_1) &= \mathbb{E}\left(\sum_{i=1}^m x_i^2 + \frac{1}{2}\sum_{j=1}^n y_j^2\right) = \mathbb{E}\left(\sum_{i=1}^m (\operatorname{Var}(x_i) + [\mathbb{E}(x_i)]^2) + \frac{1}{2}\sum_{j=1}^n (\operatorname{Var}(y_j) + [\mathbb{E}(y_j)]^2)\right) \\
&= m(\sigma^2 + a^2) + \frac{n}{2}(2\sigma^2 + a^2) = (m+n)\sigma^2 + \left(m + \frac{n}{2}\right) a^2
\end{aligned}
$$

于是我们构造

$$
\begin{cases}
\hat{a} = \dfrac{2T_2}{2m+n} \\[2ex]
\hat{\sigma}^2 = \dfrac{T_1 - \dfrac{2T_2^2}{2m+n}}{m+n-1}
\end{cases}
$$

如此有 $\mathbb{E}(\hat{a}) = a$ 以及注意到

$$\begin{aligned}\mathbb{E}\left(T_1 - \frac{2T_2^2}{2m+n}\right) &= (m+n)\sigma^2 + \left(m+\frac{n}{2}\right)a^2 - \frac{\mathbb{E}(T_2^2)}{m+n/2} \\ &= (m+n)\sigma^2 + \left(m+\frac{n}{2}\right)a^2 - \frac{1}{m+n/2}\cdot\left[\left(m+\frac{n}{2}\right)\sigma^2 + \left(m+\frac{n}{2}\right)^2 a^2\right] \\ &= (m+n-1)\sigma^2\end{aligned}$$

于是有 $\mathbb{E}(\hat{\sigma}^2) = \sigma^2$。由 Lehmann–Scheffe 定理知，$\hat{a}, \hat{\sigma}^2$ 为 a, σ^2 的 UMVUE。

Exercise 2

设 $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$ 其中 $\sigma^2 > 0$ 为未知参数，试求

1. σ^2 的充分完全统计量
2. σ 和 $3\sigma^4$ 的 UMVUE。

解答 (1) 由 $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$ 可知概率函数为

$$f(\boldsymbol{x}; \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left\{\frac{1}{2\sigma^2}\sum_{i=1}^n x_i^2\right\}$$

自然参数空间在 $\mathbb{R}$ 有内点，故 $T = \sum_{i=1}^n x_i^2$ 是充分完全统计量。

解答 (2) 设 $Y_i = X_i/\sigma \sim N(0,1)$ 且相互独立，故 $T/\sigma^2 = \sum_{i=1}^n Y_i^2 \sim \chi_n^2$。所以

$$\mathbb{E}(T/\sigma^2) = n \quad \mathrm{Var}(T/\sigma^2) = 2n$$

于是

$$\mathbb{E}(T^2/\sigma^4) = [\mathbb{E}(T/\sigma^2)]^2 + \mathrm{Var}(T/\sigma^2) = n^2 + 2n$$

于是我们构造

$$\hat{\sigma}^4 = \frac{T^2}{n^2+2n}$$

有 $\mathbb{E}(3\hat{\sigma}^4) = 3\sigma^4$，由 Lehmann-Scheffe 定理知 $3\hat{\sigma}^4 = \frac{3T^2}{n^2+2n}$ 是 $3\sigma^4$ 的 UMVUE。

下面求 σ 的 UMVUE。由于 $Z = T/\sigma^2 \sim \chi_n^2$，于是 Z 的密度函数为

$$f_Z(z) = \frac{2^{-n/2}}{\Gamma(n/2)} z^{n/2-1} e^{-z/2} \cdot \mathbb{I}(z > 0)$$

于是 $\sqrt{Z}$ 的期望为

$$\begin{aligned}\mathbb{E}(\sqrt{Z}) &= \int_0^\infty \sqrt{z}\cdot\frac{2^{-n/2}}{\Gamma(n/2)} z^{n/2-1} e^{-z/2}\, dz = \int_0^\infty \frac{2^{-n/2}}{\Gamma(n/2)} z^{\frac{n+1}{2}-1} e^{-\frac{z}{2}}\, dz \\ &\xlongequal{u=z/2} \frac{2^{-1/2}}{\Gamma(n/2)}\int_0^\infty 2u^{\frac{n+1}{2}-1} e^{-u}\, dz = 2^{1/2}\Gamma\left(\frac{n+1}{2}\right)\Big/\Gamma\left(\frac{n}{2}\right)\end{aligned}$$

于是构造

$$\begin{cases} \hat{\sigma} = \dfrac{\Gamma\left(\frac{n}{2}\right)}{2^{1/2}\Gamma\left(\dfrac{n+1}{2}\right)}\sqrt{T} \\ 3\hat{\sigma}^4 = \dfrac{3T^2}{n^2+2n} \end{cases}$$

其中 $T = \sum_{i=1}^n X_i^2$ 就有 $\hat{\sigma}, 3\hat{\sigma}^4$ 为 $\sigma, 3\sigma^4$ 的 UMVUE。

Exercise 3

证明均匀分布族 $\mathscr{F} = \{U(0,\theta) : 0 < \theta < \infty\}$ 不是 C-R 正则分布族。

证明 设 $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} U(0,\theta)$ 均匀分布，则概率函数为

$$f(\boldsymbol{x};\theta) = \theta^{-n} \cdot \mathbb{I}(0 < x_{(1)} \le x_{(n)} < \theta)$$

对参数 $\theta \in \Theta$，支撑集为

$$\mathrm{Supp}_\theta(f) = \{\boldsymbol{x} : f(\boldsymbol{x};\theta) > 0\} = \{\boldsymbol{x} : \boldsymbol{x} \in (0,\theta)^n = (0,\theta)\times(0,\theta)\times\cdots\times(0,\theta)\}$$

即支撑集 $\mathrm{Supp}_\theta(f)$ 随参数 θ 变化，不是共同的支撑集，与 C-R 正则要求相悖。

Exercise 4

设 $X_1, \cdots, X_n$ 为自下列总体中抽取的简单样本

$$f(x;\theta) = \theta^{-1}e^{-\theta^{-1}x} \cdot \mathbb{I}(x > 0)$$

其中 $\theta > 0$ 为未知参数, 求 θ 的无偏估计的方差下限和 θ 的 UMVUE，并比较 θ 的 UMVUE 方差与 C-R 下界。

解答 $X_1, \cdots, X_n$ 的似然函数为

$$L(\theta;\boldsymbol{x}) = \prod_{i=1}^n f(x;\theta) = \prod_{i=1}^n \theta^{-1}e^{-\theta^{-1}x_i} \cdot \mathbb{I}(x_i > 0) = \theta^{-n}\exp\left\{-\theta^{-1}\sum_{i=1}^n x_i\right\} \cdot \mathbb{I}(x_{(1)} > 0)$$

于是对数似然为

$$l(\theta) = -n\log\theta - \theta^{-1}\sum_{i=1}^n x_i, \quad x_{(1)} > 0$$

求导

$$\frac{dl(\theta)}{d\theta} = -n\theta^{-1} + \theta^{-2}\sum_{i=1}^n x_i$$
$$\frac{d^2l(\theta)}{d\theta^2} = n\theta^{-2} - 2\theta^{-3}\sum_{i=1}^n x_i$$

注意到 X_i 相互独立，且 X_i 服从 $\mathrm{Exp}(\theta^{-1})$ 的期望为 θ。从而 Fisher 信息量为

$$I_1(\theta) = -\mathbb{E}\left(\frac{d^2 l(\theta)}{d\theta^2}\right) = -\mathbb{E}\left(n\theta^{-2} - 2\theta^{-3}\sum_{i=1}^{n} x_i\right) = -n\theta^{-2} + 2n\theta^{-2} = n\theta^{-2}$$

于是根据 Cramer-Rao 不等式有

$$\mathrm{Var}(\hat{\theta}) \geq \frac{1}{I_1(\theta)} = \frac{\theta^2}{n}$$

又注意到，设 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 有

$$\begin{cases} \mathbb{E}(\bar{X}) = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \mathbb{E}(X_i) = \dfrac{1}{n} n\theta = \theta \\ \mathrm{Var}(\bar{X}) = \dfrac{1}{n^2}\displaystyle\sum_{i=1}^{n} \mathrm{Var}(X_i) = \dfrac{1}{n^2} n\theta^2 = \dfrac{\theta^2}{n} \end{cases}$$

故 $\bar{X}$ 是 θ 的无偏估计，且其方差达到 Cramer-Rao 不等式估计的无偏估计的方差下界，所以 $\bar{X}$ 是 θ 的 UMVUE。

HW9 第九次作业解答

更新：2026 年 1 月 9 日

Exercise 1

设 $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$，证明

$$\hat{\sigma} = \frac{\Gamma(n/2)}{\sqrt{2}\Gamma((n+1)/2)} \sqrt{\sum_{i=1}^n X_i^2}$$

是 σ 的 UMVUE，并求其效率。

证明 记 $\hat{\sigma} = \hat{g}(T)$ 而 $T = T(\boldsymbol{X}) = \sum_{i=1}^n X_i^2$。容易知 $\boldsymbol{X}$ 的概率函数满足指数族，故 C-R 条件成立，下面求 C-R 下界。

$$g'(\theta) = \frac{1}{2\sqrt{\theta}} \quad \Rightarrow \quad g'(\sigma^2) = \frac{1}{2\sqrt{\sigma^2}} = \frac{1}{2\sigma}$$

而

$$\begin{aligned} I(\sigma^2) &= \mathbb{E}\left(\frac{\partial \log f_n(\boldsymbol{x}; \sigma^2)}{\partial \sigma^2}\right)^2 = \mathbb{E}\left[\frac{\partial}{\partial \sigma^2}\left(-\frac{n}{2}\log(2\pi\sigma^2) - \sum_{i=1}^n \frac{x_i^2}{2\sigma^2}\right)\right]^2 \\ &= \mathbb{E}\left(-\frac{n}{2\sigma^2} + \frac{T}{2\sigma^4}\right)^2 = \mathbb{E}\left(\frac{n^2}{4\sigma^4} + \frac{T^2}{4\sigma^8} - \frac{nT}{2\sigma^6}\right) \\ &= \frac{n^2}{4\sigma^4} + \frac{\mathbb{E}(T^2)}{4\sigma^8} - \frac{n\mathbb{E}(T)}{2\sigma^6} \end{aligned}$$

注意到 $T/\sigma^2 \sim \chi_n^2$，于是有 $\mathbb{E}(T) = n\sigma^2$，$\mathrm{Var}(T) = \mathbb{E}(T^2) - [\mathbb{E}(T)]^2 = 2n\sigma^4$。于是 Fisher 信息量为

$$I(\sigma^2) = \frac{n^2}{4\sigma^4} + \frac{2n\sigma^4 + (n\sigma^2)^2}{4\sigma^8} - \frac{n^2\sigma^2}{2\sigma^6} = \frac{n^2 + 2n + n^2 - 2n^2}{4\sigma^4} = \frac{n}{2\sigma^4}$$

于是根据 C-R 不等式得到 C-R 下界为

$$\mathrm{Var}_{\sigma^2}[\hat{g}(T)] \geq \frac{[g'(\sigma^2)]^2}{nI(\sigma^2)} = \frac{1/(4\sigma^2)}{n/(2\sigma^4)} = \frac{\sigma^2}{2n}$$

下面分析 $\hat{\sigma} = \hat{g}(T)$ 的无偏性和有效性。注意到 $T/\sigma^2 \sim \chi_n^2$，故有

$$\begin{aligned} \mathbb{E}(\sqrt{T}/\sigma) &= \int_0^\infty t^{\frac{1}{2}} \cdot \frac{1}{2^{n/2}\Gamma(n/2)} t^{\frac{n}{2}-1} e^{-\frac{t}{2}}\, dt = \frac{1}{\Gamma(n/2)} \int_0^\infty \frac{t^{\frac{n-1}{2}}}{2^{n/2}} e^{-\frac{t}{2}}\, dt \\ &= \frac{\sqrt{2}}{\Gamma(n/2)} \int_0^\infty (t/2)^{\frac{n+1}{2}-1} e^{-\frac{t}{2}}\, dt/2 = \frac{\sqrt{2}\Gamma((n+1)/2)}{\Gamma(n/2)} \end{aligned}$$

于是有

$$\mathbb{E}(\hat{\sigma})=\frac{\Gamma(n/2)}{\sqrt{2}\Gamma((n+1)/2)}\cdot\mathbb{E}(\sqrt{T})=\frac{\Gamma(n/2)}{\sqrt{2}\Gamma((n+1)/2)}\cdot\frac{\sqrt{2}\Gamma((n+1)/2)}{\Gamma(n/2)}\sigma=\sigma$$

所以 $\hat{g}(T)=\hat{\sigma}$ 是 σ 的无偏估计，又容易得到 T 为充分完全统计量，故 $\hat{g}(T)$ 是 $g(\sigma^2)=\sqrt{\sigma^2}$ 的 UMVUE。其方差为

$$\begin{aligned}\operatorname{Var}(\hat{\sigma})&=\left(\frac{\Gamma(n/2)}{\sqrt{2}\Gamma((n+1)/2)}\right)^2\operatorname{Var}(\sqrt{T})=\left(\frac{\Gamma(n/2)}{\sqrt{2}\Gamma((n+1)/2)}\right)^2\{\mathbb{E}(T)-[\mathbb{E}(\sqrt{T})]^2\}\\&=\left(\frac{\Gamma(n/2)}{\sqrt{2}\Gamma((n+1)/2)}\right)^2\left[n\sigma^2-\left(\frac{\sqrt{2}\Gamma((n+1)/2)}{\Gamma(n/2)}\sigma\right)^2\right]\\&=\left[n\left(\frac{\Gamma(n/2)}{\sqrt{2}\Gamma((n+1)/2)}\right)^2-1\right]\sigma^2\end{aligned}$$

于是，结合之前得到的 C-R 下界，我们有效率为

$$e_{\hat{\sigma}}(\sigma)=\left[n^2\left(\frac{\Gamma(n/2)}{\Gamma((n+1)/2)}\right)^2-2n\right]^{-1}$$

注 此处虽然 $\boldsymbol{X}$ 分布族为指数族，且 $\hat{g}(T)$ 是 UMVUE，但 $\hat{g}(T)=C\sqrt{T}$ 形式不是关于充分完全统计量 T 的线性形式，故不满足 C-R 不等式取等条件，效率 $e_{\hat{\sigma}}(\sigma)<1$。

Exercise 2

设 $X_1,\cdots,X_m\overset{\text{i.i.d.}}{\sim}N(a,\sigma_1^2)$ 和 $Y_1,\cdots,Y_n\overset{\text{i.i.d.}}{\sim}N(ca,\sigma_2^2)$，其中 $c\neq0,\sigma_1^2,\sigma_2^2$ 已知，a 未知，且 $X_1,\cdots,X_m$ 与 $Y_1,\cdots,Y_n$ 独立。

1. 试求 a 的 UMVUE。
2. 基于此 UMVUE 构造 a 的一个置信系数为 $1-\alpha$ 的置信区间。

解答 (1) $\boldsymbol{X},\boldsymbol{Y}$ 的概率函数为

$$\begin{aligned}f_n(\boldsymbol{X},\boldsymbol{Y};a)&=(2\pi\sigma_1^2)^{-m/2}(2\pi\sigma_2^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma_1^2}\sum_{i=1}^m(X_i-a)^2-\frac{1}{2\sigma_2^2}\sum_{j=1}^n(Y_j-ca)^2\right\}\\&:=C\exp\left\{-\frac{ma^2}{2\sigma_1^2}-\frac{nc^2a^2}{2\sigma_2^2}+\frac{a}{\sigma_1^2}\sum_{i=1}^mX_i+\frac{ca}{\sigma_2^2}\sum_{j=1}^nY_j-\frac{1}{2\sigma_1^2}\sum_{i=1}^mX_i^2-\frac{1}{2\sigma_2^2}\sum_{j=1}^nY_j^2\right\}\\&=C(a)\exp\left\{a\cdot\left(\frac{m\bar{X}}{\sigma_1^2}+\frac{cn\bar{Y}}{\sigma_2^2}\right)\right\}h(\boldsymbol{X},\boldsymbol{Y})\end{aligned}$$

其中

$$\begin{cases}C(a)=(2\pi\sigma_1^2)^{-m/2}(2\pi\sigma_2^2)^{-n/2}\exp\left\{-\dfrac{ma^2}{2\sigma_1^2}-\dfrac{nc^2a^2}{2\sigma_2^2}\right\}\\h(\boldsymbol{X},\boldsymbol{Y})=\exp\left\{-\dfrac{1}{2\sigma_1^2}\displaystyle\sum_{i=1}^mX_i^2-\dfrac{1}{2\sigma_2^2}\sum_{j=1}^nY_j^2\right\}\end{cases}$$

故 $\boldsymbol{X}, \boldsymbol{Y}$ 样本分布族为指数族，从而容易得到

$$T(\boldsymbol{X}, \boldsymbol{Y}) = \frac{m\bar{X}}{\sigma_1^2} + \frac{cn\bar{Y}}{\sigma_2^2}$$

为 a 的充分完全统计量（指数族形式 $\Rightarrow$ 充分，自然参数空间有内点 $\Rightarrow$ 完全）。注意到 $c \neq 0, \sigma_1^2, \sigma_2^2$ 已知，且 $\mathbb{E}(\bar{X}) = a, \mathbb{E}(\bar{Y}) = ca$，有

$$\mathbb{E}[T(\boldsymbol{X}, \boldsymbol{Y})] = \frac{m\mathbb{E}(\bar{X})}{\sigma_1^2} + \frac{cn\mathbb{E}(\bar{Y})}{\sigma_2^2} = \left(\frac{m}{\sigma_1^2} + \frac{c^2 n}{\sigma_2^2}\right) a$$

于是有 a 的无偏估计为

$$\hat{g}(T) = \frac{T}{\dfrac{m}{\sigma_1^2} + \dfrac{c^2 n}{\sigma_2^2}} = \frac{\dfrac{m\bar{X}}{\sigma_1^2} + \dfrac{cn\bar{Y}}{\sigma_2^2}}{\dfrac{m}{\sigma_1^2} + \dfrac{c^2 n}{\sigma_2^2}}$$

由 T 是充分完全统计量，以及 Lehmann-Scheffe 定理可知，$\hat{g}(T)$ 是 a 的 UMVUE。

解答 (2) 注意到 $X_i \overset{\text{i.i.d.}}{\sim} N(a, \sigma_1^2), Y_j \overset{\text{i.i.d.}}{\sim} N(ca, \sigma_2^2)$ 且相互独立，其中 $c \neq 0, \sigma_1^2, \sigma_2^2$ 已知，故有

$$\bar{X} \sim N\left(a, \frac{\sigma_1^2}{m}\right) \qquad \bar{Y} \sim N\left(ca, \frac{\sigma_2^2}{n}\right)$$

从而 $\hat{g}(T)$ 也服从正态分布，其中

$$\begin{aligned}
\operatorname{Var}[\hat{g}(T)] &= \operatorname{Var}\left[\left(\frac{m\bar{X}}{\sigma_1^2} + \frac{cn\bar{Y}}{\sigma_2^2}\right) \Big/ \left(\frac{m}{\sigma_1^2} + \frac{c^2 n}{\sigma_2^2}\right)\right] \\
&= \left(\frac{m}{\sigma_1^2} + \frac{c^2 n}{\sigma_2^2}\right)^{-2} \left[\left(\frac{m}{\sigma_1^2}\right)^2 \operatorname{Var}(\bar{X}) + \left(\frac{cn}{\sigma_2^2}\right)^2 \operatorname{Var}(\bar{Y})\right] \\
&= \left(\frac{m}{\sigma_1^2} + \frac{c^2 n}{\sigma_2^2}\right)^{-2} \left(\frac{m^2}{\sigma_1^4} \cdot \frac{\sigma_1^2}{m} + \frac{c^2 n^2}{\sigma_2^4} \cdot \frac{\sigma_2^2}{n}\right) \\
&= \left(\frac{m}{\sigma_1^2} + \frac{c^2 n}{\sigma_2^2}\right)^{-2} \left(\frac{m}{\sigma_1^2} + \frac{c^2 n}{\sigma_2^2}\right) = \left(\frac{m}{\sigma_1^2} + \frac{c^2 n}{\sigma_2^2}\right)^{-1}
\end{aligned}$$

结合 $\hat{g}(T)$ 服从正态分布，于是 a 的 $1-\alpha$ 置信区间为

$$\mathrm{CI}_a = \left[\hat{g}(T) - z_{1-\alpha/2}\left(\frac{m}{\sigma_1^2} + \frac{c^2 n}{\sigma_2^2}\right)^{-1/2}, \hat{g}(T) + z_{1-\alpha/2}\left(\frac{m}{\sigma_1^2} + \frac{c^2 n}{\sigma_2^2}\right)^{-1/2}\right]$$

其中 $z_{1-\alpha/2}$ 为标准正态分布的上 $1-\alpha/2$ 分位数点。

Exercise 3

设 $X_1, \cdots, X_n$ 是来自均匀分布 $U(\theta - 1/2, \theta + 1/2)$ 的简单随机样本，求 θ 的置信系数 $1-\alpha$ 的置信区间。

解答 设 $Y_i = X_i - (\theta - 1/2)$，于是 $Y_i \overset{\text{i.i.d.}}{\sim} U(0,1)$。从而有 $Y_{(n)}$ 的概率函数为

$$f_{Y_{(n)}}(y) = ny^{n-1} \cdot \mathbb{I}(0 < y < 1)$$

于是有

$$\mathbb{P}(a < Y_{(n)} < b) = \int_a^b ny^{n-1} = b^n - a^n = 1 - \alpha$$

当 $0 < a < b < 1$ 时成立。下面优化区间长度

$$\min_{a,b} |b - a| \qquad \text{s.t.} \quad b^n - a^n = 1 - \alpha$$

不难得到 $a^* = \sqrt[n]{\alpha}, b^* = 1$，于是有

$$\mathbb{P}(\sqrt[n]{\alpha} < Y_{(n)} < 1) = \mathbb{P}(\sqrt[n]{\alpha} < X_{(n)} - (\theta - 1/2) < 1) = 1 - \alpha$$

故 θ 的 $1 - \alpha$ 置信区间为

$$\mathrm{CI}_\theta = [X_{(n)} - 1/2,\ X_{(n)} - \sqrt[n]{\alpha} + 1/2]$$

其中 $X_{(n)}$ 为 $\boldsymbol{X}$ 的第 n 个次序统计量。

Exercise 4

设 $X \sim B(m, p_1)$, $Y \sim B(n, p_2)$，且 m, n 充分大，p_1, p_2 是未知参数，X, Y 独立。基于 X, Y，求 $p_2 - p_1$ 的置信系数近似为 $1 - \alpha$ 的置信区间。

解答 设 $U_i \overset{\text{i.i.d.}}{\sim} B(1, p_1), i = 1, \cdots, m$ 和 $V_j \overset{\text{i.i.d.}}{\sim} B(1, p_2), j = 1, \cdots, n$，且 U_i 和 V_j 相互独立。因为 m, n 充分大，故由中心极限定理

$$\frac{\bar{U} - \mathbb{E}(\bar{U})}{\sqrt{\mathrm{Var}(\bar{U})}} \xrightarrow{\mathscr{L}} N(0,1) \quad \frac{\bar{V} - \mathbb{E}(\bar{V})}{\sqrt{\mathrm{Var}(\bar{V})}} \xrightarrow{\mathscr{L}} N(0,1)$$

其中 $\mathbb{E}(U_i) = p_1$, $\mathbb{E}(V_j) = p_2$, $\mathrm{Var}(\bar{U}) = p_1(1-p_1)/m$, $\mathrm{Var}(\bar{V}) = p_2(1-p_2)/n$，故有

$$\frac{(\bar{V} - \bar{U}) - (p_2 - p_1)}{\sqrt{\dfrac{p_1(1-p_1)}{m} + \dfrac{p_2(1-p_2)}{n}}} \xrightarrow{\mathscr{L}} N(0,1)$$

又因为 $\bar{U}, \bar{V}$ 为求和形式，由大数定律

$$\bar{U} \xrightarrow{P} \mathbb{E}(U_i) = p_1 \quad \bar{V} \xrightarrow{P} \mathbb{E}(V_j) = p_2$$

其中 p_1, p_2 是常数，不随机。所以有

$$\sqrt{\frac{\bar{U}(1-\bar{U})}{m} + \frac{\bar{V}(1-\bar{V})}{n}} \Bigg/ \sqrt{\frac{p_1(1-p_1)}{m} + \frac{p_2(1-p_2)}{n}} \xrightarrow{P} 1$$

由 Slutsky 定理知

$$\frac{(\bar{V} - \bar{U}) - (p_2 - p_1)}{\sqrt{\dfrac{\bar{U}(1-\bar{U})}{m} + \dfrac{\bar{V}(1-\bar{V})}{n}}} \xrightarrow{\mathscr{L}} N(0,1)$$

注意到 $\bar{U} = X/m, \bar{V} = Y/n$，从而有 $p_2 - p_1$ 的 $1 - \alpha$ 近似置信区间为

$$\mathrm{CI}_{(p_2 - p_1)} = \left[\left(\frac{Y}{n} - \frac{X}{m}\right) \pm z_{1-\alpha/2}\sqrt{\frac{X(m-X)}{m^3} + \frac{Y(n-Y)}{n^3}}\right]$$

其中 $z_{1-\alpha/2}$ 为标准正态分布的上 $1 - \alpha/2$ 分位数点。

HW10 第十次作业解答

更新：2026 年 1 月 9 日

Exercise 1

设 $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \Gamma(\alpha, \lambda)$，其中 α 已知，$\lambda > 0$。已知总体的密度函数为

$$f(x; \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$$

试证明：$\bar{X}/\alpha$ 是 $g(\lambda) = 1/\lambda$ 的有效估计。

证明 先求 $g(\lambda)$ 的 C-R 下界。注意到 $g'(\lambda) = -\lambda^{-2}$; 又 $X \sim \Gamma(\alpha, \lambda)$, 即有 $\mathbb{E}(X) = \alpha/\lambda$, $\mathrm{Var}(X) = \alpha/\lambda^2$。从而 Fisher 信息量为

$$\begin{aligned} I(\lambda) &= \mathbb{E}_\lambda \left(\frac{\partial \log f(X; \lambda)}{\partial \lambda} \right)^2 = \mathbb{E}_\lambda \left(\frac{\alpha}{\lambda} - X \right)^2 = \frac{\alpha^2}{\lambda^2} - \frac{2\alpha}{\lambda} \mathbb{E}(X) + \mathbb{E}(X^2) \\ &= \frac{\alpha^2}{\lambda^2} - \frac{2\alpha}{\lambda} \cdot \frac{\alpha}{\lambda} + \left[\frac{\alpha}{\lambda^2} + \left(\frac{\alpha}{\lambda} \right)^2 \right] = \frac{\alpha^2 - 2\alpha^2 + \alpha + \alpha^2}{\lambda^2} = \frac{\alpha}{\lambda^2} \end{aligned}$$

如此，C-R 下界为

$$\mathrm{Var}_\lambda(\hat{g}(\boldsymbol{X})) \geq \frac{[g'(\lambda)]^2}{nI(\lambda)} = \frac{\lambda^{-4}}{n\alpha/\lambda^2} = \frac{1}{n\alpha\lambda^2}$$

其中 $\hat{g}(\boldsymbol{X}) = \bar{X}/\alpha$。而

$$\mathrm{Var}_\lambda(\bar{X}/\alpha) = \frac{1}{n^2\alpha^2} \mathrm{Var}_\lambda \left(\sum_{i=1}^n X_i \right) = \frac{1}{n^2\alpha^2} \cdot \frac{n\alpha}{\lambda^2} = \frac{1}{n\alpha\lambda^2}$$

即有 $\mathrm{Var}_\lambda(\hat{g}(\boldsymbol{X}))$ 达到了 C-R 下界，为 UMVUE。效率为 $e_{\hat{g}}(\lambda) = 1$ 是有效估计。

Exercise 2

设 $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} N(a, \sigma^2)$，其中 a 已知。证明：

$$\frac{1}{n} \sqrt{\frac{\pi}{2}} \sum_{i=1}^n |X_i - a|$$

为 σ 的无偏估计，且效率为 $1/(\pi - 2)$。

证明 记 $Y_i=(X_i-a)/\sigma$，于是有 $Y_i \stackrel{\text{i.i.d.}}{\sim} N(0,1)$。从而

$$\mathbb{E}(|Y|)=\int_{-\infty}^{\infty}|y|\frac{1}{\sqrt{2\pi}}e^{-y^2/2}\,dy=2\int_0^{\infty}y\frac{1}{\sqrt{2\pi}}e^{-y^2/2}\,dy.$$

令 $t=\frac{y^2}{2}$，则 $dt=y\,dy$，因此

$$\mathbb{E}(|Y|)=\frac{2}{\sqrt{2\pi}}\int_0^{\infty}e^{-t}\,dt=\frac{2}{\sqrt{2\pi}}=\sqrt{\frac{2}{\pi}}$$

而有 $\mathbb{E}(Y^2)=1$，因此

$$\operatorname{Var}(|Y|)=\mathbb{E}(|Y|^2)-(\mathbb{E}|Y|)^2=1-\frac{2}{\pi}$$

那么有

$$\begin{aligned}\mathbb{E}(\hat{\sigma})&=\mathbb{E}\left(\frac{1}{n}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{n}|X_i-a|\right)=\mathbb{E}\left(\frac{1}{n}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{n}\sigma|Y_i|\right)\\&=\sigma\cdot\frac{1}{n}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{n}\mathbb{E}(|Y_i|)=\sigma\cdot\frac{1}{n}\sqrt{\frac{\pi}{2}}\cdot n\sqrt{\frac{2}{\pi}}=\sigma\end{aligned}$$

故 $\hat{\sigma}$ 是 σ 的无偏估计。同理，可求 $\hat{\sigma}$ 的方差

$$\operatorname{Var}(\hat{\sigma})=\frac{1}{n^2}\frac{\pi}{2}\sum_{i=1}^{n}\sigma^2\operatorname{Var}(|Y_i|)=\frac{\pi}{2n}\left(1-\frac{2}{\pi}\right)\sigma^2=\frac{\pi-2}{2n}\sigma^2$$

而 C-R 下界为 $[g'(\sigma^2)]^2[nI(\sigma^2)]^{-1}$，其中 $g'(\sigma^2)=(\sqrt{\sigma^2})'=1/(2\sigma)$，而 Fisher 信息量为

$$\begin{aligned}I(\sigma^2)&=\mathbb{E}\left[\frac{\partial}{\partial\sigma^2}\left(-\frac{1}{2}\log(2\pi\sigma^2)-\frac{1}{2\sigma^2}(X-a)^2\right)\right]^2=\mathbb{E}\left(-\frac{1}{2\sigma^2}+\frac{1}{2\sigma^4}(X-a)^2\right)^2\\&=\frac{1}{4\sigma^4}-\frac{1}{2\sigma^6}\mathbb{E}(X-a)^2+\frac{1}{4\sigma^8}\mathbb{E}(X-a)^4=\frac{1}{4\sigma^4}-\frac{1\sigma^2}{2\sigma^6}+\frac{3\sigma^4}{4\sigma^8}=\frac{1}{2\sigma^4}\end{aligned}$$

于是效率为

$$e=\frac{[g'(\sigma^2)]^2[nI(\sigma^2)]^{-1}}{\operatorname{Var}(\hat{\sigma})}=\frac{\left(\frac{1}{2\sigma}\right)^2\left(\frac{n}{2\sigma^4}\right)^{-1}}{\frac{\pi-2}{2n}\sigma^2}=\frac{1}{\pi-2}$$ □

Exercise 3

设 $X_1,\cdots,X_n \stackrel{\text{i.i.d.}}{\sim} U(0,\theta)$，其中 $\theta>0$ 为未知参数。记 $X_{(n)}=\max\{X_1,\cdots,X_n\}$，下面进行假设检验

$$H_0:\theta\geq 2\quad\leftrightarrow\quad H_1:\theta<2$$

取拒绝域为

$$W=\{(X_1,\cdots,X_n):X_{(n)}\leq 1.5\}$$

试求此检验犯第一类错误的概率的最大值。

解答 对于 $X_1,\ldots,X_n \overset{\text{i.i.d.}}{\sim} U(0,\theta)$，$X_{(n)}=\max\{X_1,\ldots,X_n\}$ 的分布函数为

$$F_{X_{(n)}}(x)=\mathbb{P}(X_{(n)}\le x)=\left(\frac{x}{\theta}\right)^n,\qquad 0\le x\le\theta$$

在原假设 $H_0:\theta\ge 2$ 下，此检验的第一类错误概率为

$$\alpha(\theta)=\mathbb{P}_\theta(X_{(n)}\le 1.5)=\left(\frac{1.5}{\theta}\right)^n,\qquad \theta\ge 2$$

由于函数 $(1.5/\theta)^n$ 在 $\theta\ge 2$ 上严格单调递减，因此其最大值在 $\theta=2$ 处取得，于是

$$\sup_{\theta\in\Theta_0}\alpha(\theta)=\left(\frac{1.5}{2}\right)^n=\left(\frac{3}{4}\right)^n$$

因此，此检验第一类错误概率的最大值为 $(3/4)^n$。

Exercise 4

设 $X_1,\cdots,X_m \overset{\text{i.i.d.}}{\sim} \text{Poi}(\lambda_1)$ 和 $Y_1,\cdots,Y_n \overset{\text{i.i.d.}}{\sim} \text{Poi}(\lambda_2)$，且样本 $X_1,\cdots,X_m$ 和 $Y_1,\cdots,Y_n$ 相互独立。用大样本方法检验

$$H_0:\lambda_2-\lambda_1=0\quad\leftrightarrow\quad H_1:\lambda_2-\lambda_1\ne 0$$

检验水平 α 给定。

解答 令样本均值为

$$\bar{X}=\frac{1}{m}\sum_{i=1}^m X_i,\qquad \bar{Y}=\frac{1}{n}\sum_{j=1}^n Y_j.$$

由中心极限定理，当 $m,n\to\infty$ 时

$$\frac{\bar{X}-\lambda_1}{\sqrt{\lambda_1/m}}\xrightarrow{\mathscr{L}}N(0,1)\qquad \frac{\bar{Y}-\lambda_2}{\sqrt{\lambda_2/n}}\xrightarrow{\mathscr{L}}N(0,1)$$

且两者渐近独立。于是有

$$\frac{(\bar{Y}-\bar{X})-(\lambda_2-\lambda_1)}{\sqrt{\dfrac{\lambda_1}{m}+\dfrac{\lambda_2}{n}}}\xrightarrow{\mathscr{L}}N(0,1)$$

构造检验统计量

$$T=\frac{\bar{Y}-\bar{X}}{\sqrt{\dfrac{\bar{X}}{m}+\dfrac{\bar{Y}}{n}}}$$

由大数定理 $\bar{X}\xrightarrow{P}\mathbb{E}(X_i)=\lambda_1$，$\bar{Y}\xrightarrow{P}\mathbb{E}(Y_j)=\lambda_2$,。当 $H_0:\lambda_2-\lambda_1=0$ 成立时，结合 Slutsky 定理，有 $T\xrightarrow{\mathscr{L}}N(0,1)$。于是置信水平为 $1-\alpha$ 的拒绝域为

$$W=\{(\boldsymbol{X},\boldsymbol{Y}):|T|\ge z_{\alpha/2}\}$$

其中 $z_{\alpha/2}$ 为标准正态分布 $N(0,1)$ 的上 $\alpha/2$ 分位数点。检验的 p 值为 $\Pr(Z>|T|)$，其中 $Z\sim N(0,1)$。

HW11 第十一次作业解答

更新：2026 年 1 月 9 日

Exercise 1

设 $X_1,\cdots,X_m \overset{\text{i.i.d.}}{\sim} N(\mu_1,\sigma^2)$，$Y_1,\cdots,Y_n \overset{\text{i.i.d.}}{\sim} N(\mu_2,\sigma^2)$，其中 σ^2 未知。且样本 $X_1,\cdots,X_m$ 和 $Y_1,\cdots,Y_n$ 相互独立。求下列检验问题水平为 α 的似然比检验

$$H_0:\mu_2-\mu_1=0 \quad \leftrightarrow \quad H_1:\mu_2-\mu_1\neq 0$$

解答 首先写出似然函数

$$\begin{aligned}L&=(2\pi\sigma^2)^{-m/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^m(X_i-\mu_1)^2\right\}(2\pi\sigma^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{j=1}^n(Y_j-\mu_2)^2\right\}\\&=(2\pi\sigma^2)^{-\frac{m+n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^m(X_i-\mu_1)^2+\sum_{j=1}^n(Y_j-\mu_2)^2\right]\right\}\end{aligned}$$

容易得到在全集 Θ 下的 MLE 为

$$\hat{\mu}_1^{\Theta}=\bar{X}\quad \hat{\mu}_2^{\Theta}=\bar{Y}\quad \hat{\sigma}^2_{\Theta}=\frac{1}{m+n}\left[\sum_{i=1}^m(X_i-\bar{X})^2+\sum_{j=1}^n(Y_j-\bar{Y})^2\right]$$

于是有对应的似然函数为

$$\sup_{\Theta}L=(2\pi\hat{\sigma}^2_{\Theta})^{-\frac{m+n}{2}}\exp\left\{-\frac{1}{2\hat{\sigma}^2_{\Theta}}\cdot(m+n)\hat{\sigma}^2_{\Theta}\right\}=(2\pi\hat{\sigma}^2_{\Theta})^{-\frac{m+n}{2}}e^{-\frac{m+n}{2}}=(2\pi e\hat{\sigma}^2_{\Theta})^{-\frac{m+n}{2}}$$

在 H_0 对应的 Θ_0 下，有 $\mu_1=\mu_2$。于是 MLE 为

$$\hat{\mu}^{\Theta_0}=\frac{1}{m+n}\left[\sum_{i=1}^m X_i+\sum_{j=1}^n Y_j\right]\quad \hat{\sigma}^2_{\Theta_0}=\frac{1}{m+n}\left[\sum_{i=1}^m(X_i-\hat{\mu}^{\Theta_0})^2+\sum_{j=1}^n(Y_j-\hat{\mu}^{\Theta_0})^2\right]$$

于是有对应的似然函数为

$$\sup_{\Theta_0}L=(2\pi\hat{\sigma}^2_{\Theta_0})^{-\frac{m+n}{2}}\exp\left\{-\frac{1}{2\hat{\sigma}^2_{\Theta_0}}\cdot(m+n)\hat{\sigma}^2_{\Theta_0}\right\}=(2\pi\hat{\sigma}^2_{\Theta_0})^{-\frac{m+n}{2}}e^{-\frac{m+n}{2}}=(2\pi e\hat{\sigma}^2_{\Theta_0})^{-\frac{m+n}{2}}$$

综上，可有似然比为

$$\lambda = \frac{\sup_{\Theta} L}{\sup_{\Theta_0} L} = \left(\frac{\hat{\sigma}^2_{\Theta_0}}{\hat{\sigma}^2_{\Theta}}\right)^{\frac{m+n}{2}} = \left(\frac{\sum_{i=1}^m (X_i - \hat{\mu}^{\Theta_0})^2 + \sum_{j=1}^n (Y_j - \hat{\mu}^{\Theta_0})^2}{\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{j=1}^n (Y_j - \bar{Y})^2}\right)^{\frac{m+n}{2}}$$

记 $\hat{\mu}^{\Theta_0}$ 为 $\bar{Z}$ 则有

$$\bar{Z} = \frac{m\bar{X} + n\bar{Y}}{m+n}$$

似然比转化为

$$\lambda = \left(\frac{\sum_{i=1}^m (X_i - \bar{X} + \bar{X} - \bar{Z})^2 + \sum_{j=1}^n (Y_j - \bar{Y} + \bar{Y} - \bar{Z})^2}{\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{j=1}^n (Y_j - \bar{Y})^2}\right)^{\frac{m+n}{2}}$$

记 $(m+n-2)S_p^2 = \sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{j=1}^n (Y_j - \bar{Y})^2$，于是有

$$\lambda = \left(\frac{(m+n-2)S_p^2 + m(\bar{X} - \bar{Z})^2 + n(\bar{Y} - \bar{Z})^2}{(m+n-2)S_p^2}\right)^{\frac{m+n}{2}}$$

又注意到 $\bar{Z}$ 的定义，故有

$$m(\bar{X} - \bar{Z})^2 + n(\bar{Y} - \bar{Z})^2 = m\left(\frac{n\bar{X} - n\bar{Y}}{m+n}\right)^2 + n\left(\frac{m\bar{Y} - m\bar{X}}{m+n}\right)^2 = \frac{mn}{(m+n)^2}(\bar{Y} - \bar{X})^2$$

代入 λ 有似然比为

$$\lambda = \left(\frac{(m+n-2)S_p^2 + \frac{mn}{(m+n)^2}(\bar{Y} - \bar{X})^2}{(m+n-2)S_p^2}\right)^{\frac{m+n}{2}} = \left(1 + \frac{mn(\bar{Y} - \bar{X})^2}{(m+n-2)(m+n)^2 S_p^2}\right)^{\frac{m+n}{2}}$$

记检验统计量

$$T = \frac{\bar{Y} - \bar{X}}{S_p\sqrt{\frac{1}{m} + \frac{1}{n}}}$$

于是似然比化为

$$\lambda = \left(1 + \frac{T^2}{(m+n-2)(m+n)}\right)^{\frac{m+n}{2}}$$

于是，似然比 λ 关于 T 单增。则待定拒绝域形如

$$D = \{(\boldsymbol{X}, \boldsymbol{Y}) : |T| = |T(\boldsymbol{X}, \boldsymbol{Y})| > c\}$$

又因为 T 在 $H_0: \mu_2 - \mu_1 = 0$ 下服从 t_{df}，其中自由度 $\mathrm{df} = m+n-2$。于是 α 水平的拒绝域为

$$D = \{(\boldsymbol{X}, \boldsymbol{Y}) : |T| = |T(\boldsymbol{X}, \boldsymbol{Y})| > t_{m+n-2}(\alpha/2)\}$$

其中 $t_{m+n-2}(\alpha/2)$ 为 t_{m+n-2} 的上 $\alpha/2$ 分位数点。

Exercise 2

设 $X_1,\cdots,X_m \overset{\text{i.i.d.}}{\sim} N(\mu_1,\sigma_1^2)$，$Y_1,\cdots,Y_n \overset{\text{i.i.d.}}{\sim} N(\mu_2,\sigma_1^2)$，且样本 $X_1,\cdots,X_m$ 和 $Y_1,\cdots,Y_n$ 相互独立。求下列检验问题水平为 α 的似然比检验

$$H_0:\sigma_1^2=\sigma_2^2 \quad \leftrightarrow \quad H_1:\sigma_1^2\neq\sigma_2^2$$

解答 首先写出似然函数

$$L=(2\pi\sigma_1^2)^{-m/2}\exp\left\{-\frac{1}{2\sigma_1^2}\sum_{i=1}^m(X_i-\mu_1)^2\right\}(2\pi\sigma_2^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma_2^2}\sum_{j=1}^n(Y_j-\mu_2)^2\right\}$$

在全集 Θ 下的 MLE 为

$$\begin{aligned}&\hat\mu_1^{\Theta}=\bar X \qquad \hat\mu_2^{\Theta}=\bar Y\\&\hat\sigma_1^2=\frac{1}{m}\sum_{i=1}^m(X_i-\bar X)^2:=S_X^2 \qquad \hat\sigma_2^2=\frac{1}{n}\sum_{j=1}^n(Y_j-\bar Y)^2:=S_Y^2\end{aligned}$$

对应的似然函数为

$$\sup_{\Theta}L=(2\pi\hat\sigma_1^2)^{-\frac{m}{2}}\exp\left\{-\frac{1}{2\hat\sigma_1^2}(m\hat\sigma_1^2)\right\}(2\pi\hat\sigma_2^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\hat\sigma_2^2}(n\hat\sigma_2^2)\right\}=(2\pi e\hat\sigma_1^2)^{-\frac{m}{2}}(2\pi e\hat\sigma_2^2)^{-\frac{n}{2}}$$

而在 $H_0:\sigma_1^2=\sigma_2^2=\sigma^2$ 对应的 Θ_0 下的 MLE 为

$$\begin{aligned}&\hat\mu_1^{\Theta_0}=\bar X \qquad \hat\mu_2^{\Theta_0}=\bar Y\\&\hat\sigma^2=\frac{1}{m+n}\left[\sum_{i=1}^m(X_i-\bar X)^2+\sum_{j=1}^n(Y_j-\bar Y)^2\right]=\frac{mS_X^2+nS_Y^2}{m+n}\end{aligned}$$

对应的似然函数为

$$\sup_{\Theta_0}L=(2\pi\hat\sigma^2)^{-\frac{m}{2}}(2\pi\hat\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\hat\sigma^2}(m+n)\hat\sigma^2\right\}=(2\pi e\hat\sigma^2)^{-\frac{m+n}{2}}$$

综上，似然比为

$$\begin{aligned}\lambda&=\frac{\sup_\Theta L}{\sup_{\Theta_0}L}=\frac{(2\pi e\hat\sigma_1^2)^{-\frac{m}{2}}(2\pi e\hat\sigma_2^2)^{-\frac{n}{2}}}{(2\pi e\hat\sigma^2)^{-\frac{m+n}{2}}}=\frac{(\hat\sigma_1^2)^{-\frac{m}{2}}(\hat\sigma_2^2)^{-\frac{n}{2}}}{(\hat\sigma^2)^{-\frac{m+n}{2}}}=\frac{(\hat\sigma^2)^{\frac{m+n}{2}}}{(\hat\sigma_1^2)^{\frac{m}{2}}(\hat\sigma_2^2)^{\frac{n}{2}}}\\&=\frac{\left(\dfrac{mS_X^2+nS_Y^2}{m+n}\right)^{\frac{m+n}{2}}}{(S_X^2)^{\frac{m}{2}}(S_Y^2)^{\frac{n}{2}}}=\frac{\left(\dfrac{mS_X^2+nS_Y^2}{m+n}\right)^{\frac{m+n}{2}}(S_Y^2)^{-\frac{m+n}{2}}}{(S_X^2)^{\frac{m}{2}}(S_Y^2)^{\frac{n}{2}}(S_Y^2)^{-\frac{m+n}{2}}}\\&=\frac{\left(\dfrac{mS_X^2/S_Y^2+n}{m+n}\right)^{\frac{m+n}{2}}}{(S_X^2/S_Y^2)^{\frac{m}{2}}}=\frac{\left(\dfrac{(mS_X^2)/(nS_Y^2)+1}{m/n+1}\right)^{\frac{m+n}{2}}}{[(mS_X^2)/(nS_Y^2)]^{\frac{m}{2}}(m/n)^{-\frac{m}{2}}}\end{aligned}$$

记 $F=(mS_X^2)/(nS_Y^2)$ 则有

$$\lambda=\frac{\left(\dfrac{F+1}{m/n+1}\right)^{\frac{m+n}{2}}}{F^{\frac{m}{2}}(m/n)^{-\frac{m}{2}}}=\frac{(m/n)^{\frac{m}{2}}}{(m/n+1)^{\frac{m+n}{2}}}\frac{(F+1)^{\frac{m+n}{2}}}{F^{\frac{m}{2}}}$$

对 λ 的核部分 λ^* 取对数，得到

$$\log \lambda^* = \frac{m+n}{2}\log(1+F) - \frac{m}{2}\log F$$

于是有关于 F 的导数为

$$\frac{\partial \log \lambda^*}{\partial F} = \frac{m+n}{2}\frac{1}{1+F} - \frac{m}{2}\frac{1}{F} = \frac{(m+n)F - m(1+F)}{2F(1+F)} = \frac{nF-m}{2F(1+F)}$$

于是有，当 $F \leq m/n$ 时，λ 随 F 单减；而当 $F > m/n$ 时，λ 随 F 单增。于是，检验的拒绝域形如

$$D = \{(\boldsymbol{X}, \boldsymbol{Y}) : (F < k_1) \cup (F > k_2)\}$$

其中 $k_1 < k_2$ 待定。不难发现，在 $H_0 : \sigma_1^2 = \sigma_2^2 = \sigma^2$ 下，有

$$F = \frac{mS_X^2/\sigma^2}{nS_Y^2/\sigma^2} \sim F_{m-1,n-1}$$

简单起见，可以取 F 分布的上下 $\alpha/2$ 分位数点。于是检验的 α 水平的拒绝域为

$$D = \{(\boldsymbol{X}, \boldsymbol{Y}) : [F < F_{m-1,n-1}(1-\alpha/2)] \cup [F > F_{m-1,n-1}(\alpha/2)]\}$$

Exercise 3

设 $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \text{Exp}(\lambda)$，求

1. $H_0 : \lambda = \lambda_0 \leftrightarrow H_1 : \lambda \neq \lambda_0$
2. $H_0 : \lambda \leq \lambda_0 \leftrightarrow H_1 : \lambda > \lambda_0$

的水平为 α 的似然比检验。

解答 (1) 由指数函数的密度函数 $f(X;\lambda) = \lambda e^{-\lambda X} \cdot \mathbb{I}(X \geq 0)$ ，可知似然函数为

$$L(\boldsymbol{X};\lambda) = \prod_{i=1}^{n} f(X_i;\lambda) = \lambda^n \exp\left\{-\lambda \sum_{i=1}^{n} X_i\right\} \cdot \mathbb{I}(X_{(1)} \geq 0)$$

容易得到 λ 的 MLE 为 $\bar{X}^{-1}$。记 $T = \sum_{i=1}^n X_i = n\bar{X}$，于是有 $T \sim \Gamma(n,\lambda)$。针对检验问题 1，有似然比为

$$\Lambda = \frac{\sup_{\Theta} L(\boldsymbol{X};\lambda)}{\sup_{\Theta_0} L(\boldsymbol{X};\lambda)} = \frac{(n/T)^n \exp\{-(n/T)T\} \cdot \mathbb{I}(X_{(1)} \geq 0)}{\lambda_0^n \exp\{-\lambda_0 T\} \cdot \mathbb{I}(X_{(1)} \geq 0)} = \left(\frac{n}{e\lambda_0 T}\right)^n e^{\lambda_0 T}$$

容易得似然函数核部分 Λ^* 的对数为

$$\log \Lambda^* = -n\log T + \lambda_0 T \quad \Rightarrow \quad \frac{\partial \log \Lambda^*}{\partial T} = -\frac{n}{T} + \lambda_0$$

由此可知，当 $T \leq n/\lambda_0$ 时，似然比 Λ 关于 T 单减；而当 $T > n/\lambda_0$ 时，似然比 Λ 关于 T 单增。所以拒绝域形如

$$D = \{\boldsymbol{X} : (T < k_1) \cup (T > k_2)\}$$

其中 $k_1 < k_2$ 待定。又因为当 $H_0 : \lambda = \lambda_0$ 成立时，有 $T \sim \Gamma(n, \lambda_0)$，即有 $2\lambda_0 T \sim \chi^2_{2n}$。于是简单起见，可取水平为 α 的拒绝域为

$$D = \left\{ \boldsymbol{X} : \left[T < \frac{\chi^2_{2n}(1-\alpha/2)}{2\lambda_0} \right] \cup \left[T > \frac{\chi^2_{2n}(\alpha/2)}{2\lambda_0} \right] \right\}$$

解答 (2) 针对 $H_0 : \lambda \leq \lambda_0$，原假设下的极大对数似然变为

$$\sup_{\Theta_0} L(\boldsymbol{X}; \lambda) = \begin{cases} \lambda_0^n \exp\{-\lambda_0 T\} \cdot \mathbb{I}(X_{(1)} \geq 0), & \lambda_0 \leq \dfrac{n}{T} \\ \sup\limits_{\Theta} L(\boldsymbol{X}; \lambda), & \lambda_0 > \dfrac{n}{T} \end{cases}$$

于是，似然比为

$$\Lambda = \begin{cases} \left(\dfrac{n}{e\lambda_0 T} \right)^n e^{\lambda_0 T}, & \lambda_0 \leq \dfrac{n}{T} \\ 1, & \lambda_0 > \dfrac{n}{T} \end{cases}$$

而由之前的推导，当 $\lambda_0 \leq n/T$ 即 $T \leq n/\lambda_0$ 时，似然比 Λ 关于 T 单减。于是 Λ 是关于 T 不增的。从而有效用函数为

$$\varphi(\boldsymbol{X}) = \begin{cases} 1, & \Lambda > c \\ 0, & \Lambda \leq c \end{cases} = \begin{cases} 1, & T < k \\ 0, & T \geq k \end{cases}$$

其中 c, k 待定。注意到 $H_0 : \lambda \leq \lambda_0 \Rightarrow \lambda/\lambda_0 \leq 1$，于是有

$$\mathbb{P}_{H_0}\left(T < \frac{\chi^2_{2n}(1-\alpha)}{2\lambda_0} \right) = \mathbb{P}_{H_0}\left(T < \frac{\chi^2_{2n}(1-\alpha)}{2\lambda} \frac{\lambda}{\lambda_0} \right) \leq \mathbb{P}_{H_0}\left(T < \frac{\chi^2_{2n}(1-\alpha)}{2\lambda} \right) = \alpha$$

其中 $\chi^2_{2n}(1-\alpha)$ 为 χ^2_{2n} 的上 $1-\alpha$ 分位数点。于是我们有该检验的 $1-\alpha$ 水平的拒绝域为

$$D = \left\{ \boldsymbol{X} : T = T(\boldsymbol{X}) = \sum_{i=1}^n X_i < \frac{\chi^2_{2n}(1-\alpha)}{2\lambda_0} \right\}$$

Exercise 4

设 $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} B(1, p)$，用大样本方法检验

$$H_0 : p = p_0 \quad \leftrightarrow \quad H_1 : p \neq p_0$$

的水平为 α 的似然比检验。

解答 对于 Bernoulli 实验，其似然函数为

$$L(\boldsymbol{X}; p) = \prod_{i=1}^n p^{X_i}(1-p)^{1-X_i} = p^{n\bar{X}}(1-p)^{n-n\bar{X}} \qquad X_i \in \{0, 1\}$$

在全集 Θ 下，p 的 MLE 为 $\bar{X}$。于是，得似然比为

$$\lambda = \frac{\sup_{\Theta} L(\boldsymbol{X};p)}{\sup_{\Theta_0} L(\boldsymbol{X};p)} = \frac{\bar{X}^{n\bar{X}}(1-\bar{X})^{n-n\bar{X}}}{p_0^{n\bar{X}}(1-p_0)^{n-n\bar{X}}}$$

取对数有

$$\log\lambda = n\bar{X}\log\bar{X} + (n-n\bar{X})\log(1-\bar{X}) - n\bar{X}\log p_0 - (n-n\bar{X})\log(1-p_0)$$

关于统计量 $\bar{X}$ 的导数为

$$\begin{aligned}\frac{\partial\log\lambda}{\partial\bar{X}} &= n\log\bar{X} + n - n\log(1-\bar{X}) - n - n\log p_0 + n\log(1-p_0)\\ &= n\left(\log\frac{\bar{X}}{1-\bar{X}} - \log\frac{p_0}{1-p_0}\right)\end{aligned}$$

故有，当 $\mathrm{logit}\bar{X} \le \mathrm{logit}p_0$ 时，λ 关于 $\bar{X}$ 单减；而当 $\mathrm{logit}\bar{X} > \mathrm{logit}p_0$ 时，λ 关于 $\bar{X}$ 单增。特别地，logit 函数为单增函数，故上述可等价于：当 $\bar{X} \le p_0$ 时，λ 关于 $\bar{X}$ 单减；而当 $\bar{X} > p_0$ 时，λ 关于 $\bar{X}$ 单增。所以，容易得到拒绝域形如

$$D = \{\boldsymbol{X} : (\bar{X} < k_1) \cup (\bar{X} > k_2)\}$$

其中 $k_1 < k_2$ 待定。又由中心极限定理，当 n 充分大时，有

$$\frac{\bar{X} - \mathbb{E}(X_i)}{\sqrt{\mathrm{Var}(X_i)}} \xrightarrow{\mathscr{L}} N(0,1) \quad \Rightarrow \quad \frac{\sqrt{n}(\bar{X}-p)}{\sqrt{p(1-p)}} \xrightarrow{\mathscr{L}} N(0,1)$$

于是有

$$\mathbb{P}_{H_0}\left(\left|\frac{\sqrt{n}(\bar{X}-p_0)}{\sqrt{p_0(1-p_0)}}\right| > z_{\alpha/2}\right) = \alpha$$

其中 $z_{\alpha/2}$ 为标准正态分布的上 $\alpha/2$ 分位数点。综上可知，水平为 α 的拒绝域为

$$D = \left\{\boldsymbol{X} : \left[\bar{X} < p_0 - z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}\right] \cup \left[\bar{X} > p_0 + z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}\right]\right\}$$