期末考试 (2025 春)

Problem 1

试证明:

$$\mathrm{Cov}(X,Y)=\mathbb{E}[\mathrm{Cov}(X,Y|T)]+\mathrm{Cov}(\mathbb{E}[X|T],\mathbb{E}[Y|T])$$

Solution:
首先我们有:

$$\begin{aligned}\mathbb{E}[\mathrm{Cov}(X,Y|T)]&=\mathbb{E}[\mathbb{E}[XY|T]-\mathbb{E}[X|T]\cdot\mathbb{E}[Y|T]]\\&=\mathbb{E}[XY]-\mathbb{E}[\mathbb{E}[X|T]\cdot\mathbb{E}[Y|T]]\end{aligned}$$

其次我们有:

$$\begin{aligned}\mathrm{Cov}(\mathbb{E}[X|T],\mathbb{E}[Y|T])&=\mathbb{E}[\mathbb{E}[X|T]\mathbb{E}[Y|T]]-\mathbb{E}[\mathbb{E}[X|T]]\cdot\mathbb{E}[\mathbb{E}[X|T]]\\&=\mathbb{E}[\mathbb{E}[X|T]\mathbb{E}[Y|T]]-\mathbb{E}[X]\cdot\mathbb{E}[Y]\end{aligned}$$

于是我们有:

$$\begin{aligned}&\mathbb{E}[\mathrm{Cov}(X,Y|T)]+\mathrm{Cov}(\mathbb{E}[X|T],\mathbb{E}[Y|T])\\&=\mathbb{E}[XY]-\mathbb{E}[\mathbb{E}[X|T]\cdot\mathbb{E}[Y|T]]+\mathbb{E}[\mathbb{E}[X|T]\cdot\mathbb{E}[Y|T]]-\mathbb{E}[X]\cdot\mathbb{E}[Y]\\&=\mathbb{E}[XY]-\mathbb{E}[X]\cdot\mathbb{E}[Y]\\&=\mathrm{Cov}(X,Y)\end{aligned}$$

Problem 2

设简单随机样本 $X=(X_1,\dots,X_n)$ 取自以下分布族:

$$p(x;\theta)=\frac{1}{\theta_2-\theta_1}I_{(\theta_1,\theta_2)}(x)$$

(1) 求 $\theta=(\theta_1,\theta_2)$ 充分统计量

- **定理 $1.3.5$: (因子化定理, 数理统计讲义 命题 $1.5.6$)**
  设样本的可能分布族为 $\mathscr{F}_X=\{f_X(x;\theta):\theta\in\Theta\}$
  其中 $f_X(x;\theta)$ 为分布密度或离散的概率分布,
  则统计量 $T=T(X)$ 为分布族 $\mathscr{F}_X$ 参数 θ 的**充分统计量**的**充要条件**是:
  对于任意 $\theta\in\Theta$, $f_X(x;\theta)$ 都可分解为 $g(T(x);\theta)\cdot h(x)$,
  其中 $h(x)$ 是与 θ 无关的**非负函数**.

Solution:

$$\begin{aligned}\mathrm{P}\{X=x\}&=\mathrm{P}\{X_1=x_1,\dots,X_n=x_n\}\\&=\prod_{i=1}^n p(x_i;\theta)\\&=\prod_{i=1}^n\frac{1}{\theta_2-\theta_1}I_{(\theta_1,\theta_2)}(x_i)\\&=\frac{1}{(\theta_2-\theta_1)^n}I(\min_{1\le i\le n}X_i>\theta_1)I(\max_{1\le i\le n}X_i<\theta_2)\\&=g(X_{(1)},X_{(n)};\theta)h(x)\end{aligned}$$

其中 $\begin{cases} X_{(1)} := \min_{1\le i\le n} X_i \\ X_{(2)} := \max_{1\le i\le n} X_i \\ g(X_{(1)}, X_{(n)}; \theta) = \frac{1}{(\theta_2-\theta_1)^n} I(X_{(1)} > \theta_1) I(X_{(n)} < \theta_2) \\ h(x) \equiv 1 \end{cases}$
根据因子化定理可知 $T(X) := (X_{(1)}, X_{(n)})$ 为 $\theta = (\theta_1, \theta_2)$ 的充分统计量.

---

(2) 求 $\theta = (\theta_1, \theta_2)$ 的矩估计量
Solution:
一阶与二阶总体原点矩和二阶总体中心矩为:

$$\begin{aligned} \alpha_1 &= \mathbb{E}[X] = \int_{\theta_1}^{\theta_2} x \cdot \frac{1}{\theta_2-\theta_1} \mathrm{d}x = \frac{\theta_1+\theta_2}{2} \\ \alpha_2 &= \mathbb{E}[X^2] = \int_{\theta_1}^{\theta_2} x^2 \cdot \frac{1}{\theta_2-\theta_1} \mathrm{d}x = \frac{\theta_1^2+\theta_1\theta_2+\theta_2^2}{3} \\ \beta_2 &= \mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{(\theta_2-\theta_1)^2}{12} \end{aligned}$$

将总体原点矩替换为样本原点矩可得:

$$\begin{aligned} \hat{\alpha}_1 &= \overline{X} = \frac{1}{n}\sum_{i=1}^n X_i = \frac{\hat{\theta}_1+\hat{\theta}_2}{2} \\ \hat{\alpha}_2 &= \frac{1}{n}\sum_{i=1}^n X_i^2 = \frac{\hat{\theta}_1^2+\hat{\theta}_1\hat{\theta}_2+\hat{\theta}_2^2}{3} \\ \hat{\beta}_2 &= S^2 = \frac{1}{n}\sum_{i=1}^n (X_i-\overline{X})^2 = \frac{(\hat{\theta}_2-\hat{\theta}_1)^2}{12} \end{aligned}$$

解得矩估计量为:

$$\begin{aligned} \hat{\theta}_1 &= \overline{X} - \sqrt{3S^2} \\ \hat{\theta}_2 &= \overline{X} + \sqrt{3S^2} \end{aligned}$$

Problem 3

设简单随机样本 $X = (X_1, \ldots, X_n)$ 取自以下分布族:

$$p(x;\theta) = \begin{cases} (1-\theta)/2, & \text{if } x = -1 \\ 1/2, & \text{if } x = 0 \\ \theta/2, & \text{if } x = 1 \end{cases}$$

(1) 求 θ 的极大似然估计，判断是否无偏.
Solution:
对数似然函数为:

$$\begin{aligned} l(x;\theta) &= \log\left(\prod_{i=1}^n p(x_i;\theta)\right) \\ &= \log\left(\left(\frac{1-\theta}{2}\right)^{n_{-1}}\left(\frac{1}{2}\right)^{n_0}\left(\frac{\theta}{2}\right)^{n_1}\right) \\ &= n_{-1}\log(1-\theta) + n_1\log(\theta) + \mathrm{const} \end{aligned}$$

其中 n_{-1}, n_0, n_1 分别是 $x_1, \ldots, x_n$ 中取值 $-1, 0, 1$ 的数量.
取 $\frac{\partial}{\partial\theta} l(x;\theta) = -\frac{n_{-1}}{1-\theta} + \frac{n_1}{\theta} = 0$

- 若 $n_{-1} = n_1 = 0$，则对数似然函数与 θ 无关，MLE 不唯一.

- 若 $n_{-1}+n_1>0$，则 MLE 为 $\hat{\theta}=n_1/(n_{-1}+n_1)$

综上所述，MLE 的定义为:

$$\hat{\theta}=\begin{cases}\text{undefined}, & \text{if } N_{-1}+N_1=0\\ N_1/(N_{-1}+N_1), & \text{if } N_{-1}+N_1>0\end{cases}$$

但为讨论无偏性，当 $n_{-1}+n_1=0$ 时，我们为 $\hat{\theta}$ 指定一个值 c:

$$\hat{\theta}=\begin{cases}c, & \text{if } N_{-1}+N_1=0\\ N_1/(N_{-1}+N_1), & \text{if } N_{-1}+N_1>0\end{cases}$$

考虑 $S=N_{-1}+N_1$
记总体为 ξ，则有 $\mathrm{P}\{\xi\neq 0\}=\mathrm{P}\{\xi\in\{-1,1\}\}=p(-1;\theta)+p(1;\theta)=\frac{1-\theta}{2}+\frac{\theta}{2}=\frac{1}{2}$，
因此 S 服从二项分布 $B(n,1/2)$.

注意到:

$$\mathrm{P}\{\xi=1|\xi\neq 0\}=\frac{\mathrm{P}\{\xi=1\}}{\mathrm{P}\{\xi\neq 0\}}=\frac{\theta/2}{1/2}=\theta$$

因此在给定 $S=s>0$ 的条件下，$N_1\sim B(s,\theta)$.
于是对于 $s=1,2,\ldots,n$，我们有:

$$\begin{aligned}\mathbb{E}[\hat{\theta}|S=s]&=\mathbb{E}\left[\frac{N_1}{S}\Big|S=s\right]\\&=\frac{1}{s}\mathbb{E}[B(s,\theta)]\\&=\frac{1}{s}\cdot s\theta\\&=\theta\end{aligned}$$

计算 $\hat{\theta}$ 的期望如下:

$$\begin{aligned}\mathbb{E}[\hat{\theta}]&=\sum_{s=0}^{n}\mathbb{E}[\hat{\theta}|S=s]\cdot\mathrm{P}\{S=s\}\\&=\mathbb{E}[\hat{\theta}|S=0]\cdot\mathrm{P}\{S=0\}+\sum_{s=1}^{n}\mathbb{E}[\hat{\theta}|S=s]\cdot\mathrm{P}\{S=s\}\\&=c\cdot\binom{n}{0}\left(1-\frac{1}{2}\right)^{n}+\sum_{s=1}^{n}\theta\cdot\binom{n}{s}\left(\frac{1}{2}\right)^{s}\left(1-\frac{1}{2}\right)^{n-s}\\&=\frac{c}{2^n}+\theta\sum_{s=1}^{n}\binom{n}{s}\left(\frac{1}{2}\right)^{s}\left(1-\frac{1}{2}\right)^{n-s}\\&=\frac{c}{2^n}+\theta\left(1-\binom{n}{0}\left(1-\frac{1}{2}\right)^{n}\right)\\&=\frac{c}{2^n}+\theta\left(1-\frac{1}{2^n}\right)\end{aligned}$$

令 $\mathbb{E}[\hat{\theta}]=\theta\ (\forall\,\theta\in(0,1))$，则我们有:

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

当且仅当 $c=\theta\ (\forall\,\theta\in(0,1))$ 时上述等式成立.
但 c 必须为常数 (不依赖于未知参数 θ),
因此不存在常数 c 使得 $\mathbb{E}[\hat{\theta}]=\theta$ 对所有 θ 成立.
故 MLE $\hat{\theta}$ 不是无偏估计量.

(2) 求 θ 无偏估计方差的 C-R 下界.
Solution:
对于总体 ξ 我们有:

$$\frac{\partial}{\partial\theta}\log p(x;\theta)=\begin{cases}\frac{\partial}{\partial\theta}\log\left((1-\theta)/2\right)=-1/(1-\theta), & \text{if } x=-1\\ \frac{\partial}{\partial\theta}\log\left(1/2\right)=0, & \text{if } x=0\\ \frac{\partial}{\partial\theta}\log\left(\theta/2\right)=1/\theta, & \text{if } x=1\end{cases}$$

总体 ξ 的 Fisher 信息量为:

$$\begin{aligned}I_\xi(\theta)&=\mathbb{E}_\theta\left[\left(\frac{\partial}{\partial\theta}l(\theta|\xi)\right)^2\right]\\&=\left(-\frac{1}{1-\theta}\right)^2\cdot\mathrm{P}\{\xi=-1\}+0^2\cdot\mathrm{P}\{\xi=0\}+\left(\frac{1}{\theta}\right)^2\cdot\mathrm{P}\{\xi=1\}\\&=\frac{1}{(1-\theta)^2}\cdot\frac{1-\theta}{2}+0\cdot\frac{1}{2}+\frac{1}{\theta^2}\cdot\frac{\theta}{2}\\&=\frac{1}{2(1-\theta)}+0+\frac{1}{2\theta}\\&=\frac{1}{2\theta(1-\theta)}\end{aligned}$$

因此 C-R 下界为:

$$\frac{(\frac{\mathrm{d}}{\mathrm{d}\theta}\theta)^2}{I_X(\theta)}=\frac{1^2}{nI_\xi(\theta)}=\frac{1}{n\cdot\frac{1}{2\theta(1-\theta)}}=\frac{2\theta(1-\theta)}{n}$$

Problem 4

已知 $\hat{\theta}_n$ 是 θ 满足渐近正态性的估计量.
试证明 $\hat{\theta}_n$ 是 θ 的相合估计量.

Solution:
已知 $\hat{\theta}_n$ 是 θ 满足渐近正态性的估计量，即存在 $v(\theta)$ 使得:

$$\frac{\sqrt{n}(\hat{\theta}_n-\theta)}{\sqrt{v(\theta)}}\overset{\mathrm{d}}{\to}N(0,1)$$

对于任意 $\varepsilon>0$ 我们都有:

$$\begin{aligned}\lim_{n\to\infty}\mathrm{P}_\theta\{|\hat{\theta}_n-\theta|>\varepsilon\}&=\lim_{n\to\infty}\mathrm{P}_\theta\left\{\left|\frac{\sqrt{n}(\hat{\theta}_n-\theta)}{\sqrt{v(\theta)}}\right|>\frac{\sqrt{n}\varepsilon}{\sqrt{v(\theta)}}\right\}\\&=\lim_{n\to\infty}2\left(1-\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{v(\theta)}}\right)\right)\\&=2(1-1)\\&=0\end{aligned}$$

其中 $\Phi(\cdot)$ 为标准正态分布的累积分布函数.
命题得证.

Problem 5

设简单随机样本 $X=(X_1,\dots,X_n)$ 取自 $N(\mu,1)$.
求 μ^2 的一致最小方差无偏估计量，并判断其是否渐近有效.

- **Lemma 1 (正态总体的样本均值与样本方差的联合分布, S. Ross 命题 2.5)**
  若 $X=(X_1,\dots,X_n)$ 为取自 $\mathrm{N}(\mu,\sigma^2)$ 的简单随机样本，样本量为 n，
  定义样本均值 $\overline{X}=\frac{1}{n}\sum_{i=1}^n x_i$ 和已修偏样本方差 $S_n^{*2}=\frac{1}{n-1}\sum_{i=1}^n(x_i-\overline{X})^2$，
  则有 $\begin{cases}\overline{X}\perp S_n^{*2}\\ \overline{X}\sim\mathrm{N}(\mu,\frac{\sigma^2}{n})\\ S_n^{*2}\sim\sigma^2\frac{\chi^2(n-1)}{n-1}\end{cases}$ 成立.
- **Lemma 2:**
  我们证明 $(\overline{X},S_n^2)$ 是该正态分布族参数 (μ,σ^2) 的**充分完备统计量**:
  - 首先我们利用因子化定理证明统计量 $(\overline{X},S_n^2)$ 的**充分性**:
    记 $x=(x_1,\dots,x_n)$，则我们有:

$$\begin{aligned}\mathrm{P}\{X=x\}&=\mathrm{P}\{X_1=x_1,\dots,X_n=x_n\}\\&=\prod_{i=1}^n\mathrm{P}\{N(\mu,\sigma^2)=x_i\}\\&=\prod_{i=1}^n(2\pi\sigma^2)^{-\frac{1}{2}}\exp\left\{-\frac{1}{2\sigma^2}(x_i-\mu)^2\right\}\\&=(2\pi\sigma^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-\mu)^2\right\}\\&=(2\pi\sigma^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^n(x_i-\bar{x})^2+n(\bar{x}-\mu)^2\right]\right\}\\&=(2\pi\sigma^2)^{-n/2}\exp\left\{-\frac{n}{2\sigma^2}s^2\right\}\exp\left\{-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2\right\}\end{aligned}$$

    其中 $\begin{cases}s^2=\frac{1}{n}\sum_{i=1}^n(x_i-\bar{x})^2\\ \bar{x}=\frac{1}{n}\sum_{i=1}^n x_i\end{cases}$
    考虑统计量 $T=(\overline{X},S_n^2)$，
    记
    $\begin{cases}g(T(x);\mu,\sigma^2)=g(\bar{x},s^2;\mu,\sigma^2)=(2\pi\sigma^2)^{-n/2}\exp\{-\frac{n-1}{2\sigma^2}s^2\}\exp\{-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2\}\\ h(x)\equiv 1\end{cases}$
    根据**因子化定理**我们知道，$T=(\overline{X},S_n^2)$ 是参数 (μ,σ^2)的充分统计量.
  - 下面我们证明统计量 $(\overline{X},S_n^2)$ 的**完备性**:
    根据 **Lemma** 1 我们知道 $\begin{cases}\overline{X}\perp S_n^2\\ \overline{X}\sim N(\mu,\frac{\sigma^2}{n})\\ S_n^2\sim\sigma^2\frac{\chi^2(n-1)}{n}\end{cases}$
    因此 $(\overline{X},S_n^2)$ 的联合分布 $p_{(\overline{X},S_n^2)}(\bar{x},s^2;\mu,\sigma^2)=p_{\overline{X}}(\bar{x};\mu,\sigma^2)\cdot p_{S_n^2}(s^2;\mu,\sigma^2)$
    任意给定 $(\overline{X},S_n^2)$ 的函数 $\phi(\overline{X},S_n^2)$，我们有:

$$\begin{aligned}\mathrm{E}[\phi(\overline{X},S_n^2)]&=\mathrm{E}[\mathrm{E}[\phi(\overline{X},S_n^2)|S_n^2]]\\&=\int_0^\infty\mathrm{E}[\phi(\overline{X},S_n^2)|S_n^2=s^2]\cdot p_{S_n^2}(s^2;\mu,\sigma^2)\mathrm{d}s^2\\&=\int_0^\infty\mathrm{E}[\phi(\overline{X},s^2)]\cdot p_{S_n^2}(s^2;\mu,\sigma^2)\mathrm{d}s^2\end{aligned}$$

假设对于任意 (μ,σ^2) 我们都有 $\mathrm{E}[\phi(\overline{X},S_n^2)]=0$ 成立，
则我们知道 $\mathrm{E}[\phi(\overline{X},s^2)]$ 作为 s^2 的函数，在 $s^2\in(0,\infty)$ 上几乎处处为 0

这个结论表明，对于任意给定 $s^2>0$，
我们都有 $\mathrm{E}[\phi(\overline{X},s^2)]=\int_{-\infty}^{\infty}\phi(\bar{x},s^2)p_{\overline{X}}(\bar{x};\mu,\sigma^2)\mathrm{d}\bar{x}=0\ \ (\forall\, u\in\mathbb{R},\sigma^2>0)$ 成立，
这说明对于任意 $\begin{cases}\bar{x}\in\mathbb{R}\\ s^2>0\end{cases}$ 我们都有 $\phi(\bar{x},s^2)=0$ 成立.

上述推理表明统计量 $(\overline{X},S_n^2)$ 是**完备的**.

综上所述，$(\overline{X},S_n^2)$ 是该正态分布族参数 (μ,σ^2) 的**充分完备统计量**.

- **定理** $2.2.4$**: (Lehmann-Scheffé 数理统计讲义 命题** $2.2.30$**)**
  若:
  - $T(X)$ 是样本分布族 $\mathscr{F}_X=\{F_X(\theta):\theta\in\Theta\}$ 的参数 θ 的**充分完备统计量**.
  - $\hat{g}(X)$ 为参数函数 $g(\theta)$ 的方差有限的无偏估计量.

  则 $h(T)=\mathbb{E}[\hat{g}(X)|T]$ 为 $g(\theta)$ 的**一致最小方差无偏估计量** UMVUE
- **Lemma** 3**:**
  对于正态随机变量 $Y\sim N(\mu,\sigma^2)$ 我们有 $\mathrm{Var}(Y^2)=4\mu^2\sigma^2+2\sigma^4$.

  证明:
  注意到 $\mathrm{Var}(Y^2)=\mathbb{E}[Y^4]-(\mathbb{E}[Y^2])^2$
  二阶矩:

$$\mathbb{E}[Y^2]=\mathrm{Var}(Y)+(\mathbb{E}[Y])^2=\sigma^2+\mu^2$$

  记 $Z=Y-\mu\sim N(0,\sigma^2)$，则四阶矩计算如下:

$$\begin{aligned}
Y^4=(Z+\mu)^4=\mu^4+4\mu^3Z+6\mu^2Z^2+4\mu Z^3+Z^4\\
\Downarrow\\
\mathbb{E}[Y^4]=\mu^4+4\mu^3\mathbb{E}[Z]+6\mu^2\mathbb{E}[Z^2]+4\mu\mathbb{E}[Z^3]+\mathbb{E}[Z^4]\\
=\mu^4+4\mu^3\cdot 0+6\mu^2\cdot\sigma^2+4\mu\cdot 0+3\sigma^4\\
=\mu^4+6\mu^2\sigma^2+3\sigma^4
\end{aligned}$$

  于是我们有:

$$\begin{aligned}
\mathrm{Var}(Y^2)&=\mathbb{E}[Y^4]-(\mathbb{E}[Y^2])^2\\
&=\mu^4+6\mu^2\sigma^2+3\sigma^4-(\sigma^2+\mu^2)^2\\
&=4\mu^2\sigma^2+2\sigma^4
\end{aligned}$$

Solution:
根据总体分布 $\xi\sim N(\mu,1)$ 可知 $\overline{X}\sim N(\mu,1/n)$.
于是我们有:

$$\mathbb{E}[\overline{X}^2]=\mathrm{Var}(\overline{X})+(\mathbb{E}[\overline{X}])^2=\frac{1}{n}+\mu^2$$

因此 μ^2 的一个无偏估计量为 $\hat{\alpha}=\overline{X}^2-\frac{1}{n}$.

根据 **Lemma** 2 可知 $\overline{X}$ 是 μ 的充分完备统计量.
由 **Lehmann-Scheffé 定理**可知 μ^2 的无偏估计量 $\hat{\alpha}=\overline{X}^2-\frac{1}{n}$ 作为 $\overline{X}$ 的函数，
一定是 μ^2 的 UMVUE (一致最小方差无偏估计量).

下面我们将 $\hat{\alpha}$ 的方差与 C-R 下界比较，以判断其是否渐近有效.
根据 **Lemma** 3 可知 $\hat{\alpha}$ 的方差为:

$$
\begin{aligned}
\mathrm{Var}(\hat{\alpha}) &= \mathrm{Var}(\overline{X}^2 - \frac{1}{n}) \\
&= \mathrm{Var}(\overline{X}^2) \\
&= 4\mu^2 \cdot \frac{1}{n} + 2 \cdot \left(\frac{1}{n}\right)^2 \\
&= \frac{4\mu^2}{n} + \frac{2}{n^2}
\end{aligned}
$$

现在计算 C-R 下界.
总体 ξ 的对数似然函数为:

$$
\begin{aligned}
l(\mu; x) &= \log\left(\mathrm{P}\{N(\mu, 1) = x\}\right) \\
&= \log\left\{\frac{1}{\sqrt{2\pi}} \exp\left(\frac{1}{2}(x-\mu)^2\right)\right\} \\
&= -\frac{1}{2}\log(2\pi) - \frac{1}{2}(x-\mu)^2
\end{aligned}
$$

我们有 $\begin{cases} \frac{\partial}{\partial \mu} l(\mu; x) = x - \mu \\ \frac{\partial^2}{\partial \mu^2} l(\mu; x) = -1 \end{cases}$

Fisher 信息量为 (第二步的转化基于 "$\int p(x;\mu)\mathrm{d}x = 1$ 关于 μ 可在积分号下微分两次" 的条件):

$$
\begin{aligned}
I_\xi(\mu) &= \mathbb{E}_\mu\left[\left(\frac{\partial}{\partial \mu} l(\mu|\xi)\right)^2\right] \\
&= -\mathbb{E}_\mu\left[\frac{\partial^2}{\partial \mu^2} l(\mu|\xi)\right] \\
&= -\mathbb{E}_\mu[-1] \\
&= 1
\end{aligned}
$$

因此 C-R 下界为 $\mathrm{CRLB} = \frac{(\frac{\mathrm{d}}{\mathrm{d}\mu}\mu^2)^2}{I_X(\mu)} = \frac{(2\mu)^2}{nI_\xi(\mu)} = \frac{4\mu^2}{n \cdot 1} = \frac{4\mu^2}{n}$
比较 $\hat{\alpha}$ 的方差与 C-R 下界可得:

$$
\mathrm{Var}(\hat{\alpha}) - \mathrm{CRLB} = \left(\frac{4\mu^2}{n} + \frac{2}{n^2}\right) - \frac{4\mu^2}{n} = \frac{2}{n^2} \to 0 \quad (n \to \infty)
$$

这表明 $\hat{\alpha}$ 是渐近有效的.

Problem 6

设 x_0 取自密度函数为 $p(x)$ 的总体.
试构造 α 水平概率比检验:

$$
\begin{gathered}
H_0 : p(x) = 4xI_{[0,1/2)}(x) + 4(1-x)I_{[1/2,1]}(x) \\
\Updownarrow \\
H_1 : p(x) = 2xI_{(0,1)}(x)
\end{gathered}
$$

- **定理 $3.2.1$: (Neyman-Pearson 引理, 数理统计讲义 命题 $3.2.2$)**
  设参数空间为 $\Theta = \{\theta_0, \theta_1\}$
  样本 X 的分布具有分布密度 (或离散的概率) $p_0(\cdot) := p_{\theta_0}(\cdot)$ 和 $p_1(\cdot) := p_{\theta_1}(\cdot)$
  对于假设检验问题 $H_0 : \theta = \theta_0 \leftrightarrow H_1 : \theta = \theta_1$ 和显著水平 $\alpha \in (0, 1)$

- **存在性:**
  存在非负常数 k 以及**概率比检验** $\phi_0(x) = \begin{cases} 1, & \text{if } \frac{p_1(x)}{p_0(x)} > k \\ 0, & \text{if } \frac{p_1(x)}{p_0(x)} < k \end{cases}$
  满足 $\gamma_{\phi_0}(\theta_0) = \mathbb{E}_{\theta_0}[\phi_0(X)] = \int \phi_0(x) p_{\theta_0}(x) \mathrm{d}x = \alpha$
- **充分性:**
  存在性中描述的 ϕ_0 是显著水平 α 的**最有效检验法** (MP, most powerful)
  即对于显著水平 α 的任意检验函数 $\phi \in \mathit{\Phi}_\alpha$ 都有:

$$\gamma_{\phi_0}(\theta_1) = \mathbb{E}_{\theta_1}[\phi_0(X)] \geq \mathbb{E}_{\theta_1}[\phi(X)] = \gamma_\phi(\theta_1)$$

- **必要性:**
  若 $\phi^\star$ 为显著水平 α 的**最有效检验法**，则 $\phi^\star$ 必定是概率比检验.

Solution:
我们记:

$$\begin{aligned} p_0(x) &:= 4x I_{[0,1/2)}(x) + 4(1-x) I_{[1/2,1]}(x) \\ p_1(x) &:= 2x I_{(0,1)}(x) \end{aligned}$$

概率比:

$$\Lambda(x) = \frac{p_1(x)}{p_0(x)} = \begin{cases} \frac{2x}{4x} = \frac{1}{2}, & \text{if } x \in [0, \frac{1}{2}) \\ \frac{2x}{4(1-x)} = \frac{x}{2(1-x)}, & \text{if } x \in [\frac{1}{2}, 1] \end{cases}$$

Neyman-Pearson 引理给出最有效的 α 水平检验法为:

$$\phi_\alpha(x) = \mathbb{1}\{\Lambda(x) > c_\alpha\}$$

注意到 $\mathrm{P}_{H_0}\{X \geq \frac{1}{2}\} = \int_{1/2}^1 4(1-x) \mathrm{d}x = \frac{1}{2}$
因此我们需要对 $0 < \alpha \leq 1/2$ 和 $1/2 < \alpha < 1$ 两种情况进行分类讨论.

- 当 $0 < \alpha \leq 1/2$ 时,
  注意到 $\Lambda(x) \geq \frac{1}{2}$ $(\forall\, x \in (0,1])$，取 $c_\alpha \in [\frac{1}{2}, 1]$ 使得:

$$\begin{aligned} \alpha &= \mathbb{E}[\phi_\alpha(X)] \\ &= \mathrm{P}_{H_0}\{\Lambda(x) > c_\alpha\} \\ &= \int_{\{x:\Lambda(x)>c_\alpha\}} p_0(x) \mathrm{d}x \\ &= \int_{2c_\alpha/(1+2c_\alpha)}^1 4(1-x) \mathrm{d}x \\ &= (4x - 2x^2)\Big|_{2c_\alpha/(1+2c_\alpha)}^1 \\ &= \frac{2}{(1+2c_\alpha)^2} \end{aligned}$$

  解得:

$$c_\alpha = \frac{1}{2}\left(\sqrt{\frac{2}{\alpha}} - 1\right)$$

  于是最有效的 α 水平检验法为:

$$\phi_\alpha(x) = \mathbb{1}\left\{\Lambda(x) > \frac{1}{2}\left(\sqrt{\frac{2}{\alpha}} - 1\right)\right\}$$

- 当 $1/2<\alpha<1$ 时,
  我们需要将整个右半区间 $[1/2,1]$ 都纳入拒绝域 (其在 H_0 下概率正好是 $1/2$),
  剩余的 $\alpha-1/2$ 需要从 $[0,1/2)$ 进行随机化补齐.
  定义随机化检验函数:

$$\phi_\alpha(x)=\begin{cases}1, & \text{if } \Lambda(x)>\frac{1}{2} \text{ (i.e. } x\in(\frac{1}{2},1])\\ \gamma, & \text{if } \Lambda(x)=\frac{1}{2} \text{ (i.e. } x\in[0,\frac{1}{2}])\end{cases}$$

  取 γ 使得:

$$\begin{aligned}\alpha&=\mathbb{E}[\phi_\alpha(X)]\\&=\mathrm{P}_{H_0}\{X\in(1/2,1]\}+\gamma\cdot\mathrm{P}_{H_0}\{X\in[0,1/2]\}\\&=\frac{1}{2}+\gamma\cdot\frac{1}{2}\end{aligned}$$

  解得:

$$\gamma=2\alpha-1$$

  于是我们有:

$$\phi_\alpha(x)=\begin{cases}1, & x\in(\frac{1}{2},1]\\ 2\alpha-1, & x\in[0,\frac{1}{2}]\end{cases}$$

Problem 7

设简单随机样本 $X=(X_1,\ldots,X_n)$ 取自 $\exp(\lambda)$.
试证明:

$$\sqrt{n}\left(\frac{1}{\overline{X}\lambda}-1\right)\xrightarrow{\text{d}}N(0,1)\quad(n\to\infty)$$

- **(Delta 方法, 数理统计讲义 定理** $2.4.24$**)**
  设 k 维随机向量序列 $\{T_n\}$ 满足渐近正态性 $\sqrt{n}(T_n-\theta)\xrightarrow{\text{d}}N(\mu,\Sigma)$
  假设向量值函数 $\phi(t):\mathbb{R}^k\to\mathbb{R}^m$ 在 $t=\theta$ 处可微,
  记其梯度 $\nabla\phi(\theta)=\begin{bmatrix}\frac{\partial}{\partial t_1}\phi_1(\theta) & \cdots & \frac{\partial}{\partial t_1}\phi_m(\theta)\\ \vdots & & \vdots\\ \frac{\partial}{\partial t_k}\phi_1(\theta) & \cdots & \frac{\partial}{\partial t_k}\phi_m(\theta)\end{bmatrix}\in\mathbb{R}^{k\times m}$
  则我们有 $\sqrt{n}(\phi(T_n)-\phi(\theta))\xrightarrow{\text{d}}N(\nabla\phi(\theta)^{\mathrm{T}}\mu,\nabla\phi(\theta)^{\mathrm{T}}\Sigma\nabla\phi(\theta))$
  - 特别地，当 $\begin{cases}k=1\\m=1\end{cases}$ 时，上述结论可写为:
    若 $\sqrt{n}(T_n-\theta)\to N(\mu,\sigma^2)$ 且 $\phi(t)$ 在 $t=\theta$ 处可微,
    则 $\sqrt{n}(\phi(T_n)-\phi(\theta))\xrightarrow{\text{d}}N(\phi'(\theta)\mu,(\phi'(\theta))^2\sigma^2)$
    (二阶 Delta 方法)
    此时如果 $\phi'(\theta)=0$，则得到的极限分布是退化分布,
    我们需要考虑更高阶的 Taylor 展开.
    具体来说 (在 $\phi'(\theta)=0$ 的条件下) 我们有 $\phi(t)=\phi(\theta)+\frac{1}{2}\phi''(\theta)(t-\theta)^2+o((t-\theta)^2)$
    为简单起见，考虑 $\sqrt{n}(T_n-\theta)\xrightarrow{\text{d}}N(0,\sigma^2)$ 我们有:

$$
\begin{aligned}
n(\phi(T_n)-\phi(\theta)) &\approx \frac{1}{2}\phi''(\theta)\{\sqrt{n}(T_n-\theta)\}^2 \\
&\xrightarrow{\mathrm{d}} \frac{1}{2}\phi''(\theta)\{N(0,\sigma^2)\}^2 \\
&= \frac{1}{2}\phi''(\theta)\sigma^2\chi^2(1)
\end{aligned}
$$

Solution:
总体 $\xi\sim\exp(\lambda)$ 有:

$$
\begin{aligned}
\mathbb{E}[\xi] &= \frac{1}{\lambda} \\
\mathrm{Var}(\xi) &= \frac{1}{\lambda^2}
\end{aligned}
$$

由中心极限定理有:

$$
\sqrt{n}\left(\overline{X}-\frac{1}{\lambda}\right) \xrightarrow{\mathrm{d}} N\left(0,\frac{1}{\lambda^2}\right) \quad (n\to\infty)
$$

记 $\phi(t)=1/t$，则我们有 $\phi'(t)=-1/t^2$.
根据 Delta 方法可知，当 $n\to\infty$ 时我们有:

$$
\begin{gathered}
\sqrt{n}\left(\phi(\overline{X})-\phi\left(\frac{1}{\lambda}\right)\right) \xrightarrow{\mathrm{d}} N\left(\phi'\left(\frac{1}{\lambda}\right)\cdot 0,\left(\phi'\left(\frac{1}{\lambda}\right)\right)^2\cdot\frac{1}{\lambda^2}\right) \\
\Updownarrow \\
\sqrt{n}\left(\frac{1}{\overline{X}}-\lambda\right) \xrightarrow{\mathrm{d}} N(0,\lambda^2) \\
\Updownarrow \\
\sqrt{n}\left(\frac{1}{\overline{X}\lambda}-1\right) \xrightarrow{\mathrm{d}} N(0,1)
\end{gathered}
$$

Problem 8

设简单随机样本 $E=(E_1,\dots,E_n)$，$W=(W_1,\dots,W_n)$ 取自 $\exp(\lambda)$，且两者独立.
设 $(E_{n,(1)},\dots,E_{n,(n)})$ 为 $(E_1,\dots,E_n)$ 的次序统计量.

- **定理 $1.3.4$: (次序统计量的联合概率密度函数, 数理统计讲义 命题 $1.4.7$)**
  设 $X_{(1)},X_{(2)},\dots,X_{(n)}$ 为对应于简单随机样本 $X=(X_1,X_2,\dots,X_n)$ 的次序统计量，
  (我们可以看作存在映射关系 $T(X_1,X_2,\dots,X_n)=(X_{(1)},X_{(2)},\dots,X_{(n)})$)
  总体分布具有分布函数 F 和概率密度函数 f.
  则对于任意 $\begin{cases}1\le r\le n\\ 1\le j_1<j_2<\cdots<j_r\le n\end{cases}$
  $(X_{(j_1)},X_{(j_2)},\dots,X_{(j_r)})$ 具有联合概率密度函数:

  $$
  \begin{aligned}
  &f_{X_{(j_1)},X_{(j_2)},\dots,X_{(j_r)}}(y_{j_1},y_{j_2},\dots,y_{j_r}) \\
  &=\frac{n!}{(j_1-1)!(j_2-j_1-1)!\dots(j_r-j_{r-1}-1)!(n-j_r)!} \\
  &\quad\times[F(y_{j_1})]^{j_1-1}[F(y_{j_2})-F(y_{j_1})]^{j_2-j_1-1}\cdots[F(y_{j_r})-F(y_{j_{r-1}})]^{j_r-j_{r-1}-1}[1-F(y_{j_r})]^{n-j_r} \\
  &\quad\times f(y_{j_1})f(y_{j_2})\cdots f(y_{j_r}) \\
  &\quad\times I(y_{j_1}<y_{j_2}<\cdots<y_{j_r})
  \end{aligned}
  $$

- **Lemma:**
  设 $X\sim\exp(\lambda)$，记 $Y=X/k$，其概率密度函数为:

$$
\begin{aligned}
f_Y(y) &= f_X(x)\cdot\left|\frac{\partial x}{\partial y}\right| \\
&= \lambda e^{-\lambda x}\cdot k \\
&= k\lambda\exp(-k\lambda y) \\
&= \mathrm{P}\{\exp(k\lambda)=y\}
\end{aligned}
$$

于是我们有 $Y=X/k\sim\exp(k\lambda)$

(1) 试证明:

$$
(E_{n,(1)},E_{n,(2)},\dots,E_{n,(n)})\overset{\mathrm{d}}{=}\left(\frac{W_1}{n},\frac{W_1}{n}+\frac{W_2}{n-1},\dots,\sum_{j=1}^n\frac{W_j}{n-j+1}\right)
$$

Solution:
$E_{n,(1)},\dots,E_{n,(n)}$ 的联合概率密度函数为:

$$
\begin{aligned}
f_{E_{n,(1)},\dots,E_{n,(n)}}(x_1,\dots,x_n) &= n!\prod_{i=1}^n\lambda e^{-\lambda x_i}I(0<x_1<\cdots<x_n) \\
&= n!\lambda^n\exp\left(-\lambda\sum_{i=1}^n x_i\right)I(0<x_1<\cdots<x_n)
\end{aligned}
$$

记 $S_1=E_{n,(1)},\ S_k=E_{n,(k)}-E_{n,(k-1)}\ (k=2,\dots,n)$
则我们有 $E_{n,(k)}=\sum_{j=1}^k S_j\ (k=1,\dots,n)$

$$
\begin{bmatrix}E_{n,(1)}\\E_{n,(2)}\\\vdots\\E_{n,(n)}\end{bmatrix}=\begin{bmatrix}1&&&\\1&1&&\\\vdots&\vdots&\ddots&\\1&1&\cdots&1\end{bmatrix}\begin{bmatrix}S_1\\S_2\\\vdots\\S_n\end{bmatrix}
$$

$S_1,\dots,S_n$ 的联合概率密度函数为:

$$
\begin{aligned}
f_{S_1,\dots,S_n}(s_1,\dots,s_n) &= f_{E_{n,(1)},\dots,E_{n,(n)}}(x_1,\dots,x_n)\cdot\left|\frac{\partial(E_{n,(1)},\dots,E_{n,(n)})}{\partial(S_1,\dots,S_n)}\right| \\
&= n!\lambda^n\exp\left(-\lambda\sum_{i=1}^n x_i\right)\cdot\left|\det\begin{bmatrix}1&&&\\1&1&&\\\vdots&\vdots&\ddots&\\1&1&\cdots&1\end{bmatrix}\right| \\
&= n!\lambda^n\exp\left(-\lambda\sum_{k=1}^n\sum_{j=1}^k s_j\right) \\
&= n!\lambda^n\exp\left(-\lambda\sum_{j=1}^n(n-j+1)s_j\right)\quad\left(\text{note that }\prod_{j=1}^n(n-j+1)=n!\right) \\
&= \left(\prod_{j=1}^n(n-j+1)\lambda\right)\exp\left(-\sum_{j=1}^n(n-j+1)\lambda s_j\right) \\
&= \prod_{j=1}^n(n-j+1)\lambda\exp(-(n-j+1)\lambda s_j) \\
&= \prod_{j=1}^n\mathrm{P}\{\exp((n-j+1)\lambda)=s_j\}
\end{aligned}
$$

注意到 $W_i\sim\exp(\lambda)$，根据 **Lemma** 我们有 $W_i/(n-j+1)\sim\exp((n-j+1)\lambda)$.
于是我们有:

$$(S_1, S_2, \ldots, S_n) \overset{\mathrm{d}}{=} \left(\frac{W_1}{n}, \frac{W_2}{n-1}, \ldots, \frac{W_n}{1}\right)$$

最终有:

$$(E_{n,(1)}, E_{n,(2)}, \ldots, E_{n,(n)}) = \left(S_1, S_1 + S_2, \ldots, \sum_{j=1}^{n} S_j\right) \overset{\mathrm{d}}{=} \left(\frac{W_1}{n}, \frac{W_1}{n} + \frac{W_2}{n-1}, \ldots, \sum_{j=1}^{n} \frac{W_j}{n-j+1}\right)$$

---

(2) 试证明:

$$(E_{n,(n)} - E_{n,(n-k)}, E_{n,(n-1)} - E_{n,(n-k)}, \ldots, E_{n,(n-k+1)} - E_{n,(n-k)}) \overset{\mathrm{d}}{=} (E_{k,(k)}, E_{k,(k-1)}, \ldots, E_{k,(1)})$$

- 值得注意的是，样本个数对次序统计量是有影响的，注意区别 $E_{n,(k)}$ 和 $E_{k,(k)}$.

Solution:
根据 (1) 的结论我们有:

$$(E_{k,(1)}, E_{k,(2)}, \ldots, E_{k,(k)}) \overset{\mathrm{d}}{=} \left(\frac{W_1}{k}, \frac{W_1}{k} + \frac{W_2}{k-1}, \ldots, \sum_{j=1}^{k} \frac{W_j}{k-j+1}\right)$$

于是我们有:

$$\begin{aligned}
E_{n,(n-k+m)} - E_{n,(n-k)} &= \sum_{j=n-k+1}^{n-k+m} S_j \\
&= \sum_{j=1}^{m} S_{n-k+j} \\
&\overset{\mathrm{d}}{=} \sum_{j=1}^{m} \frac{W_{n-k+j}}{k-j+1} \\
&\overset{\mathrm{d}}{=} \sum_{j=1}^{m} \frac{W_j}{k-j+1} \\
&\overset{\mathrm{d}}{=} E_{k,(m)}
\end{aligned}$$

其中 $S_{n-k+j} \sim \exp\left((k-j+1)\lambda\right)$
综上所述，我们有:

$$(E_{n,(n)} - E_{n,(n-k)}, E_{n,(n-1)} - E_{n,(n-k)}, \ldots, E_{n,(n-k+1)} - E_{n,(n-k)}) \overset{\mathrm{d}}{=} (E_{k,(k)}, E_{k,(k-1)}, \ldots, E_{k,(1)})$$

The End