统计学基础 I : 数理统计 Assignment 8

姓名: 雍崔扬
学号: 21307140051
习题: E2.35, E2.36, E2.39, E2.42, E2.43(2), E3.1

Problem 1 (习题 2.35)

设 $X=(X_1,\ldots,X_n)$ 为取自 Bernoulli 分布族 $\{B(1,p):p\in(0,1)\}$ 的简单随机样本.
证明: $\frac{\sqrt{n}(\overline{X}-p)}{\sqrt{\overline{X}(1-\overline{X})}}$ 以 $N(0,1)$ 为极限分布.

- **Lemma 1: (Khinchin 弱大数定律)**
  设 $X_1,X_2,\ldots$ 是一列独立同分布的随机变量,
  当且仅当对于任意 $i=1,2,\ldots$ 都有 $\mathrm{E}[X_i]=\mu<\infty$ 时，有 $\overline{X}\xrightarrow{\mathrm{p}}\mu$ 成立.
- **Lemma 2: (Feller-Lévy 中心极限定理)**
  设 $X_1,X_2,\ldots$ 是一列独立同分布的随机变量,
  当且仅当对于任意 $i=1,2,\ldots$ 都有 $\begin{cases}0<\mathrm{Var}[X_i]=\sigma^2<\infty\\-\infty<\mathrm{E}[X_i]=\mu<\infty\end{cases}$ 时，有 $\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma}\xrightarrow{\mathrm{d}}\mathrm{N}(0,1)$
  成立.
- **Lemma 3: (连续映射定理)**
  若随机变量序列 $X_n\xrightarrow{\mathrm{p}}X$，且 $f(\cdot)$ 是在 X 的所有可能取值处连续的函数,
  则我们有 $f(X_n)\xrightarrow{\mathrm{p}}f(X)$ 成立.
  (这个定理也可以使用**几乎处处收敛** $\xrightarrow{\mathrm{a.s.}}$ 或**依分布收敛** $\xrightarrow{\mathrm{d}}$ 来叙述)
- **Lemma 4: 随机变量序列收敛性的变换规则**
  设 $\{X_n\},\{Y_n\}$ 为随机变量序列，X,Y 为随机变量.
  - **① 加减:**
    - 若 $\begin{cases}X_n\xrightarrow{\mathrm{p}}X\\Y_n\xrightarrow{\mathrm{p}}Y\end{cases}$，则 $X_n\pm Y_n\xrightarrow{\mathrm{p}}X\pm Y$
    - 若 $\begin{cases}X_n\xrightarrow{\mathrm{d}}X\\Y_n\xrightarrow{\mathrm{d}}c\quad(\text{i.e. }Y_n\xrightarrow{\mathrm{p}}c)\end{cases}$，则 $X_n\pm Y_n\xrightarrow{\mathrm{d}}X\pm c\ \ (\text{Slutzky})$
      一般来说，对于依分布收敛，$\begin{cases}X_n\xrightarrow{\mathrm{d}}X\\Y_n\xrightarrow{\mathrm{d}}Y\end{cases}\nRightarrow\quad X_n\pm Y_n\xrightarrow{\mathrm{d}}X\pm Y$
  - **② 乘除:**
    - 若 $\begin{cases}X_n\xrightarrow{\mathrm{p}}X\\Y_n\xrightarrow{\mathrm{p}}Y\end{cases}$，则 $X_nY_n\xrightarrow{\mathrm{p}}XY$
      当 $Y\neq0$ 时，进一步成立 $\frac{X_n}{Y_n}\xrightarrow{\mathrm{p}}\frac{X}{Y}$
    - 若 $\begin{cases}X_n\xrightarrow{\mathrm{d}}X\\Y_n\xrightarrow{\mathrm{d}}c\quad(\text{i.e. }Y_n\xrightarrow{\mathrm{p}}c)\end{cases}$，则 $X_nY_n\xrightarrow{\mathrm{d}}cX\ \ (\text{Slutzky})$
      当 $c\neq0$ 时，进一步成立 $\frac{X_n}{Y_n}\xrightarrow{\mathrm{d}}\frac{X}{c}$
      一般来说，对于依分布收敛，$\begin{cases}X_n\xrightarrow{\mathrm{d}}X\\Y_n\xrightarrow{\mathrm{d}}Y\end{cases}\nRightarrow\quad X_nY_n\xrightarrow{\mathrm{d}}XY$
  - **③ 组合:**

- 若 $\begin{cases} X_n \xrightarrow{\text{p}} X \\ Y_n \xrightarrow{\text{p}} Y \end{cases}$，则 $\begin{bmatrix} X_n \\ Y_n \end{bmatrix} \xrightarrow{\text{p}} \begin{bmatrix} X \\ Y \end{bmatrix}$
- 若 $\begin{cases} X_n \xrightarrow{\text{d}} X \\ Y_n \xrightarrow{\text{d}} c \quad (\text{i.e. } Y_n \xrightarrow{\text{p}} c) \end{cases}$，则 $\begin{bmatrix} X_n \\ Y_n \end{bmatrix} \xrightarrow{\text{d}} \begin{bmatrix} X \\ c \end{bmatrix}$

  一般来说，对于依分布收敛，$\begin{cases} X_n \xrightarrow{\text{d}} X \\ Y_n \xrightarrow{\text{d}} Y \end{cases} \nRightarrow \begin{bmatrix} X_n \\ Y_n \end{bmatrix} \xrightarrow{\text{d}} \begin{bmatrix} X \\ Y \end{bmatrix}$

Solution:
根据 **Khinchin 弱大数定律**我们有 $\overline{X} \xrightarrow{\text{p}} p$ 成立，
进而根据**连续映射定理**有 $\sqrt{\overline{X}(1-\overline{X})} \xrightarrow{\text{p}} \sqrt{p(1-p)}$ 成立，即有:

$$\sqrt{\frac{\overline{X}(1-\overline{X})}{p(1-p)}} \xrightarrow{\text{p}} 1$$

根据 **Feller-Lévy 中心极限定理**我们有:

$$\frac{\sqrt{n}(\overline{X}-p)}{\sqrt{p(1-p)}} \xrightarrow{\text{d}} N(0,1)$$

综合上述结论，根据**随机变量序列收敛性的变换规则**的 **Slutzky 定理**可知:

$$\frac{\sqrt{n}(\overline{X}-p)}{\sqrt{\overline{X}(1-\overline{X})}} = \frac{\frac{\sqrt{n}(\overline{X}-p)}{\sqrt{p(1-p)}}}{\sqrt{\frac{\overline{X}(1-\overline{X})}{p(1-p)}}} \xrightarrow{\text{d}} N(0,1)$$

命题得证.

Problem 2 (习题 2.36)

设 $(X_1,\dots,X_m)$ 和 $(Y_1,\dots,Y_n)$ 为分别取自 $N(\mu,\sigma_1^2)$ 和 $N(\mu,\sigma_2^2)$ 的独立样本.
证明: 当 m,n 趋于无穷时，统计量 $\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{S_m^2(X)}{m}+\frac{S_n^2(Y)}{n}}}$ 以 $N(0,1)$ 为极限分布.

Lemma 5: (样本矩是总体矩的相合估计量, 数理统计讲义 $2.4.7$)
记总体分布的累积分布函数为 $F(x;\theta)$.

- 样本 k 阶原点矩是总体 k 阶原点矩的相合估计量.
  即对于任意 $\theta\in\Theta$ 都有 $A_k = \frac{1}{n}\sum_{i=1}^n X_i^k \xrightarrow{\text{p}} \alpha_k = \int x^k \mathrm{d}F(x;\theta)$
  (对随机变量序列 $\{X_i^k\}$ 应用 **Khinchin 弱大数定律**即得)
- 样本 k 阶中心矩是总体 k 阶中心矩的相合估计量.
  即对于任意 $\theta\in\Theta$ 都有 $M_k = \frac{1}{n}\sum_{i=1}^n (X_i-\overline{X})^k \xrightarrow{\text{p}} \mu_k = \int (x-\mu)^k \mathrm{d}F(x;\theta)$
  其中 $\mu=\alpha_1$ 表示总体分布的均值.

Solution:
我们有 $\begin{cases} \overline{X} \sim N(\mu,\frac{\sigma_1^2}{m}) \\ \overline{Y} \sim N(\mu,\frac{\sigma_2^2}{n}) \end{cases}$
根据 $(X_1,\dots,X_m)$ 和 $(Y_1,\dots,Y_n)$ 的独立性，我们知道 $\overline{X}\perp\overline{Y}$
于是我们有 $\overline{X}-\overline{Y} \sim N(0,\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n})$，表明:

$$\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}} \stackrel{\text{d}}{=} N(0,1)$$

根据 **Lemma 5** 可知 $\begin{cases} S_m^2(X) \xrightarrow{\text{p}} \sigma_1^2 \\ S_n^2(Y) \xrightarrow{\text{p}} \sigma_2^2 \end{cases}$

因此我们有:

$$\frac{\sqrt{\frac{S_m^2(X)}{m}+\frac{S_n^2(Y)}{n}}}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}} \xrightarrow{\text{p}} 1$$

综合上述结论，根据**随机变量序列收敛性的变换规则**的 **Slutzky 定理**可知:

$$\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{S_m^2(X)}{m}+\frac{S_n^2(Y)}{n}}} = \frac{\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}}{\frac{\sqrt{\frac{S_m^2(X)}{m}+\frac{S_n^2(Y)}{n}}}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}} \xrightarrow{\text{d}} N(0,1)$$

命题得证.

Problem 3 (习题 2.39)

设 k 维随机变量序列 $X_n \xrightarrow{\text{d}} N(\mu,\Sigma)$ (其中 $\det(\Sigma) \neq 0$)
证明: $(X_n-\mu)^{\text{T}}\Sigma^{-1}(X_n-\mu) \xrightarrow{\text{d}} \chi^2(k)$

Solution:
根据 $X_n \xrightarrow{\text{d}} N(\mu,\Sigma)$ 可知 $\Sigma^{-\frac{1}{2}}(X_n-\mu) \xrightarrow{\text{d}} N(\Sigma^{-\frac{1}{2}}(\mu-\mu),\Sigma^{-\frac{1}{2}}\Sigma(\Sigma^{-\frac{1}{2}})^{\text{T}}) = N(0_k,I_k)$

记 $Y_n = \Sigma^{-\frac{1}{2}}(X_n-\mu) \xrightarrow{\text{d}} N(0_k,I_k)$，其 k 个分量 $Y_n^{(1)},\dots,Y_n^{(k)}$.
根据渐近正态性可知存在独立同分布的 $N(0,1)$ 随机变量 $Y^{(1)},\dots,Y^{(k)}$ 使得:

$$Y_n^{(i)} \xrightarrow{\text{d}} Y^{(i)} \quad (i=1,\dots,k)$$

于是我们有:

$$(X_n-\mu)^{\text{T}}\Sigma^{-1}(X_n-\mu) = Y_n^{\text{T}}Y_n = \sum_{i=1}^k (Y_n^{(i)})^2 \xrightarrow{\text{d}} \sum_{i=1}^k (Y^{(i)})^2 \stackrel{\text{d}}{=} \chi^2(k)$$

命题得证.

Problem 4 (习题 2.42)

设 $X_1,\dots,X_n$ 为取自分布族 $\{B(1,p): p\in(0,1)\}$ 的简单随机样本.
记 $Y_n = \frac{1}{n}\sum_{i=1}^n X_i$
确定 α 及 $f(p)$ 使 $n^\alpha(Y_n(1-Y_n)-f(p))$ 具有非退化的极限分布，
并写出这一极限分布.

- **Lemma 6: (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:

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

Solution:

根据 **Feller-Lévy 中心极限定理**我们有 $\sqrt{n}(Y_n-p)\xrightarrow{\text{d}}N(0,p(1-p))$ 成立.

定义 $\phi(y)=y(1-y)$，其导函数为 $\phi'(y)=1-2y$
我们发现 $\phi'(p)\begin{cases}=0 & \text{if } p=\frac{1}{2}\\ \neq 0 & \text{if } p\neq\frac{1}{2}\end{cases}$

- 首先考虑 $p\neq\frac{1}{2}$ 的情况:
  根据 **Delta 方法**我们有:

$$\begin{aligned}\sqrt{n}(Y_n(1-Y_n)-p(1-p))&=\sqrt{n}(\phi(Y_n)-\phi(p))\\&\xrightarrow{\text{d}}N(\phi'(p)\cdot 0,(\phi'(p))^2\cdot p(1-p))\\&=N((1-2p)\cdot 0,(1-2p)^2\cdot p(1-p))\\&=N(0,p(1-p)(1-2p)^2)\end{aligned}$$

  对应 $\begin{cases}\alpha=\frac{1}{2}\\f(p)=p(1-p)\end{cases}$
- 其次考虑 $p=\frac{1}{2}$ 的情况:
  此时 $\phi'(\frac{1}{2})=0$，表明 $\phi(p)$ 的一阶 Taylor 近似在 $p=\frac{1}{2}$ 处是失效的.

  考虑**二阶 Delta 方法 (Lemma 6)**
  二阶导函数 $\phi''(\frac{1}{2})\equiv-2$，于是我们有:

$$
\begin{aligned}
n(Y_n(1-Y_n)-\frac{1}{4}) &= n\left(\phi(T_n)-\phi\left(\frac{1}{2}\right)\right) \\
&\approx \frac{1}{2}\phi''\left(\frac{1}{2}\right)\left\{\sqrt{n}\left(T_n-\frac{1}{2}\right)\right\}^2 \\
&\xrightarrow{\mathrm{d}} \frac{1}{2}(-2)\left\{N\left(0,\frac{1}{4}\right)\right\}^2 \\
&= -\frac{1}{4}\chi^2(1)
\end{aligned}
$$

对应 $\begin{cases}\alpha=1 \\ f(p)=p(1-p)=\frac{1}{4}\end{cases}$

Problem 5 (习题 2.43(2))

若总体存在四阶矩:
(2) 写出样本变异系数 $\mathrm{CV}=\frac{S_n}{\overline{X}}$ 的渐近分布.

Lemma 7: (矩估计量渐近正态性的保证, 数理统计讲义 命题 $2.4.26$)
设总体分布存在 $2k$ 阶矩，$X=(X_1,\dots,X_n)$ 为取自总体的简单随机样本.
记 $\begin{cases}\mu=\alpha_1=\mathrm{E}[\xi] \\ \alpha_k=\mathrm{E}[\xi^k] \\ A_k(n)=\frac{1}{n}\sum_{i=1}^n X_i^k \quad (A_1(n)=\overline{X}) \\ \mu_k=\mathrm{E}[(\xi-\mu)^k] \quad (\mu_0=\mu_1=0) \\ M_k(n)=\frac{1}{n}\sum_{i=1}^n(X_i-\overline{X})^k\end{cases}$
则我们有:

- 当 $n\to\infty$ 时 $\sqrt{n}\begin{bmatrix}A_1(n)-\alpha_1 \\ \vdots \\ A_k(n)-\alpha_k\end{bmatrix}\xrightarrow{\mathrm{d}} N(0_k,\Sigma_{k\times k}^{(1)})$
  其中协方差矩阵元素的通式为 $\Sigma_{ij}^{(1)}=\alpha_{i+j}-\alpha_i\alpha_j$
- 当 $n\to\infty$ 时 $\sqrt{n}\begin{bmatrix}\overline{X}-\mu \\ M_2(n)-\mu_2 \\ \vdots \\ M_k(n)-\mu_k\end{bmatrix}\xrightarrow{\mathrm{d}} N(0_k,\Sigma_{k\times k}^{(2)})$
  其中协方差矩阵元素的通式为
  $\Sigma_{ij}^{(2)}=\mu_{i+j}-\mu_i\mu_j-i\mu_{i-1}\mu_{j+1}-j\mu_{i+1}\mu_{j-1}+ij\mu_{i-1}\mu_{j-1}\mu_2$
  - 特殊地，对于 $k=2$ 的情况，我们可将其表述为:
    $\sqrt{n}\begin{bmatrix}\overline{X}-\mu \\ S_n^2-\sigma^2\end{bmatrix}\xrightarrow{\mathrm{d}} N(\begin{bmatrix}0 \\ 0\end{bmatrix},\begin{bmatrix}\sigma^2 & \mu_3 \\ \mu_3 & \mu_4-\sigma^4\end{bmatrix})$
  - 对于正态随机变量 $\xi\sim N(\mu,\sigma^2)$，我们知道:
    $$\mathrm{E}\left[\left(\frac{\xi-\mu}{\sigma}\right)^m\right]=\begin{cases}0 & m=2k-1 \\ (m-1)!! & m=2k\end{cases}$$
    表明 $\mu_3=0$ 且 $\mu_4=(3)!!\cdot\sigma^4=3\sigma^4$.
    因此对于正态总体有:
    $$\sqrt{n}\begin{bmatrix}\overline{X}-\mu \\ S_n^2-\sigma^2\end{bmatrix}\xrightarrow{\mathrm{d}} N\left(\begin{bmatrix}0 \\ 0\end{bmatrix},\begin{bmatrix}\sigma^2 & \mu_3 \\ \mu_3 & \mu_4-\sigma^4\end{bmatrix}\right)=N\left(\begin{bmatrix}0 \\ 0\end{bmatrix},\begin{bmatrix}\sigma^2 & 0 \\ 0 & 2\sigma^4\end{bmatrix}\right)$$

上述命题表明:
只要待估计的参数函数可以表示为**总体矩**的可微函数，
那么**矩估计量**就是渐近正态的.

Solution:
记总体的均值为 μ，方差为 σ^2，三阶中心矩为 μ_3，四阶中心矩为 μ_4
设 $X=(X_1,\dots,X_n)$ 为取自总体的简单随机样本.
定义**总体变异系数**为 $v=\frac{\sigma}{\mu}$ 和**样本变异系数** $\mathrm{CV}=\frac{S_n}{\overline{X}}$
根据 **Lemma 7** 可知:

$$\sqrt{n}\begin{bmatrix}\overline{X}-\mu\\ S_n^2-\sigma^2\end{bmatrix}\stackrel{\mathrm{d}}{\to}N\left(\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}\sigma^2 & \mu_3\\ \mu_3 & \mu_4-\sigma^4\end{bmatrix}\right)$$

定义函数 $\phi(x,y)=\frac{\sqrt{y}}{x}$，我们有 $\begin{cases}\phi(\mu,\sigma^2)=\frac{\sqrt{\sigma^2}}{\mu}=\frac{\sigma^2}{\mu}=v\\ \phi(\overline{X},S_n^2)=\frac{\sqrt{S_n^2}}{\overline{X}}=\frac{S_n}{\overline{X}}=\mathrm{CV}\end{cases}$
其在 (μ,σ^2) 处的梯度 $\nabla\phi(\mu,\sigma^2)=\begin{bmatrix}\frac{\partial}{\partial\mu}\phi(\mu,\sigma^2)\\ \frac{\partial}{\partial\sigma^2}\phi(\mu,\sigma^2)\end{bmatrix}=\begin{bmatrix}-\frac{\sigma}{\mu^2}\\ \frac{1}{2\mu\sigma}\end{bmatrix}$
利用**Lemma 6 (Delta 方法)** 我们有:

$$\begin{aligned}\sqrt{n}(\mathrm{CV}-v)&=\sqrt{n}(\phi(\overline{X},S_n^2)-\phi(\mu,\sigma^2))\\ &\stackrel{\mathrm{d}}{\to}N\left(\nabla\phi(\mu,\sigma^2)^{\mathrm{T}}\begin{bmatrix}0\\0\end{bmatrix},\nabla\phi(\mu,\sigma^2)^{\mathrm{T}}\begin{bmatrix}\sigma^2 & \mu_3\\ \mu_3 & \mu_4-\sigma^4\end{bmatrix}\nabla\phi(\mu,\sigma^2)\right)\\ &=N\left(\begin{bmatrix}-\frac{\sigma}{\mu^2}\\ \frac{1}{2\mu\sigma}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}-\frac{\sigma}{\mu^2}\\ \frac{1}{2\mu\sigma}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\sigma^2 & \mu_3\\ \mu_3 & \mu_4-\sigma^4\end{bmatrix}\begin{bmatrix}-\frac{\sigma}{\mu^2}\\ \frac{1}{2\mu\sigma}\end{bmatrix}\right)\\ &=N\left(0,\frac{\sigma^4}{\mu^4}-\frac{\mu_3}{\mu^3}+\frac{\mu_4-\sigma^4}{4\mu^2\sigma^2}\right)\end{aligned}$$

Problem 6 (习题 3.1)

设总体分布为 $N(\mu,\sigma^2)$，其中 μ,σ^2 未知，$\mu\in\mathbb{R}$ 且 $\sigma^2>0$
指出下列假设中，哪些是简单假设，哪些是复合假设.

(1) $H_0:\mu=0$

- 这是复合假设，方差 $\sigma^2>0$

(2) $H_0:\mu=0,\sigma^2=1$

- 这是简单假设，总体分布唯一确定.

(3) $H_0:\mu\le 2,\sigma^2=1$

- 这是复合假设.

(4) $H_0:\mu=0,\sigma^2\ge 1$

- 这是复合假设.

(5) $H_0:-1\le\mu\le 1$

- 这是复合假设.

(6) $H_0:\sigma^2=1$

- 这是复合假设，均值 $\mu\in\mathbb{R}$

The End