HOMEWORK 1

2021/09/15

1.5 (1) 样本空间 $\mathcal{X}=\{(x_1,\cdots,x_5)\mid x_i=0$或$1,\ i=1,\cdots,5\}$ (简记作 $\{0,1\}^5$)

X 的概率分布为

$$P(\vec{X}=\vec{x})=\prod_{i=1}^{5}p^{x_i}(1-p)^{1-x_i}, \quad \vec{x}\in\mathcal{X}$$

rmk: ① 题中所问是向量 $\vec{X}=(X_1,\cdots,X_5)$ 的分布，简单样本，由独立性：$P(\vec{X}=\vec{x})=\prod_{i=1}^{5}P(X_i=x_i)$

② 记号 $b(n,p)$ 表示二项分布：如 n次独立试验，每次成功概率为p，成功次数服从$b(n,p)$. 本题中 $\sum_{i=1}^{5}X_i\sim b(5,p)$，$\vec{X}=(X_1,\cdots,X_5)$ 不服从 $b(5,p)$.

(2) X_1+X_2，$\min_{1\le i\le 5}X_i$ 是，其它不是

- $E(X_1)=p$，$D(X_1)=p(1-p)$，与未知参数p有关，不是统计量

[Wei] 1.9.

$\because \quad \bar{X}-50\sim N(0,1)$.

$\therefore \quad P\left(\bar{X}\in(50.6,51.8)\right)=P\left(\bar{X}-50\in(0.6,1.8)\right)=\Phi(1.8)-\Phi(0.6)=0.2383.$

□

[Wei] 1.10. Consider $c>0$.

$\because \quad \sqrt{n}(\bar{X}-\mu)=10(\bar{X}-\mu)\sim N(0,1)$.

$$\therefore \quad P\left(|\bar{X}|<c\right)=P\left(\bar{X}\in(-c,c)\right)=P\Big(10(\bar{X}-\mu)\in(10(-c-\mu),10(c-\mu))\Big)$$
$$=\Phi\big(10(c-\mu)\big)-\Phi\big(10(-c-\mu)\big)\le 0.05,\quad \forall\mu>0.$$

Date: 2021/09/25.

Thanks for Weiyu Li who is with the School of the Gifted Young, University of Science and Technology of China. Corresponding Email: liweiyu@mail.ustc.edu.cn.

Since the p.d.f. of $N(0,1)$ is unimodal, $P\left(|\bar{X}| < c\right)$ reaches its maximum as $\mu \to 0+$, that is,

$$\begin{aligned} &\sup_{\mu>0} P\left(|\bar{X}| < c\right) = \Phi(10c) - \Phi(-10c) = 2\Phi(10c) - 1 \le 0.05. \\ &\therefore \quad c \le \frac{z_{0.475}}{10} = 0.0063. \end{aligned}$$

Further, when $c \le 0$, the probability is 0. In summary, any $c \le 0.0063$ satisfies. □

[Wei] 1.11.

$$\because \quad \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim N(0,1).$$

$$\therefore \quad P\left(|\bar{X}-\mu| < 0.1\right) = P\left(\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \in \left(-\frac{0.1\sqrt{n}}{\sigma}, \frac{0.1\sqrt{n}}{\sigma}\right)\right) = 2\Phi(\sqrt{\frac{n}{50}}) - 1 \ge 99.7\%.$$

$$\therefore \quad n \ge 50 z_{0.0015}^2 = 440.3734.$$

$\therefore$ n should be at least 441. □

[Wei] 2.3.(2).

$$\begin{aligned} \because \quad & \Phi_X(t) = \mathbb{E}e^{itX} = \exp\{\lambda\left(e^{it} - 1\right)\}. \\ \therefore \quad & \Phi_{\bar{X}}(t) = \mathbb{E}e^{it\bar{X}} \\ & \overset{i.i.d.}{=} \prod_{k=1}^{n} \mathbb{E}e^{i\frac{t}{n}X} = \exp\{n\lambda\left(e^{\frac{it}{n}} - 1\right)\}. \\ \therefore \quad & \bar{X} \sim \frac{1}{n} Poisson(n\lambda). \end{aligned}$$

□

[Wei] 2.1. (1)

$$\begin{aligned} &\because \quad \bar{X} \sim N(20, 0.9), \quad \bar{Y} \sim N(20, 0.6), \quad \bar{X},\ \bar{Y} \text{ independent} \\ &\therefore \quad \frac{\bar{X}-\bar{Y}}{\sqrt{1.5}} \sim N(0,1) \\ &\therefore \quad P\left(|\bar{X}-\bar{Y}| > 0.3\right) = P\left(\left|\frac{\bar{X}-\bar{Y}}{\sqrt{1.5}}\right| > \frac{\sqrt{6}}{10}\right) = 2\left(1 - \Phi\left(\frac{\sqrt{6}}{10}\right)\right) = 0.8065. \end{aligned}$$

(2)

$$
\begin{aligned}
&\because \quad \frac{n-1}{\sigma^2}S^2 \sim \chi^2_{n-1}\\
&\therefore \quad \frac{9}{9}S_X^2 \sim \chi^2_9, \quad \frac{14}{9}S_Y^2 \sim \chi^2_{14}, \quad S_X^2,\ S_Y^2 \text{ independent}\\
&\therefore \quad S_X^2 + \frac{14}{9}S_Y^2 \sim \chi^2_{23}\\
&\therefore \quad P\left(9S_X^2 + 14S_Y^2 > 164\right) = P\left(\chi^2_{23} > 164/9\right) = 0.7453.
\end{aligned}
$$

□

[Wei] 2.2.

$$
\begin{aligned}
&\because \quad \bar{X} \sim N(\mu, \frac{\sigma^2}{n}), \quad \bar{Y} \sim N(\mu, \frac{\sigma^2}{n}), \quad \bar{X},\ \bar{Y} \text{ independent}\\
&\therefore \quad \frac{\bar{X}-\bar{Y}}{\sqrt{\frac{2}{n}}\sigma} \sim N(0,1)\\
&\therefore \quad P\left(|\bar{X}-\bar{Y}| > \sigma\right) = P\left(\left|\frac{\bar{X}-\bar{Y}}{\sqrt{\frac{2}{n}}\sigma}\right| > \sqrt{\frac{n}{2}}\right) = 2\left(1-\Phi\left(\sqrt{\frac{n}{2}}\right)\right) \approx 0.01\\
&\therefore \quad \sqrt{\frac{n}{2}} \approx z_{0.005} = 2.5758\\
&\therefore \quad n \approx 13.
\end{aligned}
$$

□

1.9 rmk: 正态分布

① $X \sim N(\mu, \sigma^2)$，特征函数：

$$\phi_X(t) = \mathbb{E}e^{itX} = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{itx}\cdot e^{-\frac{1}{2\sigma^2}(x-\mu)^2}dx = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(x^2-2it\sigma^2x-2\mu x+\mu^2)}dx$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}[(x-it\sigma^2-\mu)^2-2\mu\sigma^2it+t^2\sigma^4]}dx = e^{i\mu t-\frac{1}{2}\sigma^2t^2}$$

• 本课中常见的分布（正态、Poission、Gamma等）都与特征函数有1-1对应关系.

② $Y = \frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)$

因 $\phi_Y(t) = \mathbb{E}e^{i\frac{t}{\sqrt{\sigma^2/n}}(\bar{X}-\mu)} = \mathbb{E}e^{i\frac{t}{\sqrt{\sigma^2/n}}\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)} \overset{X_i\text{ 独立}}{=\!=} \prod_{i=1}^{n}\mathbb{E}e^{i\frac{t}{\sqrt{n\sigma^2}}(X_i-\mu)}$

$$= e^{-\frac{it\cdot n\mu}{\sqrt{n\sigma^2}}}\left(\phi_X\left(\frac{t}{\sqrt{n\sigma^2}}\right)\right)^n = e^{-\frac{i\mu t\sqrt{n}}{\sigma}}\cdot e^{n\left(\frac{i\mu t}{\sqrt{n}\sigma}-\frac{1}{2}\sigma^2\frac{t^2}{n\sigma^2}\right)} = e^{-\frac{1}{2}t^2}$$

③ $X_1, \cdots, X_k$ 独立，$Z = \sum_{i=1}^{k}a_iX_i$，$a_i \in \mathbb{R}$，则

$$\phi_Z(t) = \mathbb{E}e^{it\cdot\sum_{i=1}^{k}a_iX_i} = \prod_{i=1}^{k}\mathbb{E}e^{i(a_it)X_i} = \prod_{i=1}^{k}\phi_{X_i}(a_it)$$

2.1 (2) rmk: ① $X_1, \cdots, X_n$ i.i.d. $\sim N(\mu, \sigma^2)$, $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$, $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

② $Y_1, \cdots, Y_m$ 独立, $Y_i \sim \chi^2_{n_i}$, 则 $\sum_{i=1}^m Y_i \sim \chi^2_{\sum_{i=1}^m n_i}$

实际上, 卡方分布 χ^2_n 对应 Gamma 分布 $\Gamma(\frac{n}{2}, \frac{1}{2})$, 其 p.d.f. 定义域为 $\{x>0\}$

(Gamma 分布 $\Gamma(\alpha, \beta)$ 密度函数 $f(x|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$, $x>0$; 特征函数 $\phi(t) = (\frac{\beta}{\beta - it})^\alpha$)

χ^2_n 有特征函数 $\phi(t) = (1-2it)^{-\frac{n}{2}}$

记 $Z = \sum_{i=1}^m Y_i$, $\phi_Z(t) = \prod_{i=1}^m \phi_{Y_i}(t) = (1-2it)^{-\frac{1}{2}\sum_{i=1}^m n_i} \Rightarrow Z \sim \chi^2_{\sum_{i=1}^m n_i}$

2.5 思路: 利用密度函数换元 说明独立

- 这里密度函数换元可以当作积分换元理解. 若 $\varphi: U \to V$, $(x,y) \mapsto (u,v)$ 是 1-1 的, 可微的, 则 $\int_U f(x,y)\,dx\,dy = \int_V f(u(x,y), v(x,y))\,|J|\,du\,dv$, $|J| = \left|\frac{\partial(x,y)}{\partial(u,v)}\right|$

$\Rightarrow f_{U,V}(u,v) = f_{X,Y}(x,y)\left|\frac{\partial(x,y)}{\partial(u,v)}\right|$

方法 1 (推荐)

X_1、X_2 独立, 服从 $N(0,\sigma^2)$, 其联合 p.d.f. 为:

$f(x_1, x_2) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}(x_1^2 + x_2^2)}$ $\quad x_1, x_2 \in \mathbb{R}$.

令 $\begin{cases} x_1 = r\cos\theta \\ x_2 = r\sin\theta \end{cases}$ $\quad r \geq 0$, $\theta \in [0, 2\pi)$ (注意说明取值范围)

$\left|\frac{\partial(x_1,x_2)}{\partial(r,\theta)}\right| = r$, 则 $f(r,\theta) = \frac{1}{2\pi\sigma^2} r e^{-\frac{1}{2\sigma^2}r^2}$, $\quad r \geq 0$, $\theta \in [0, 2\pi)$

$\Rightarrow f_R(r) = \frac{1}{\sigma^2} r e^{-\frac{1}{2\sigma^2}r^2} \cdot 1(r \geq 0)$ $\quad f_\Theta(\theta) = \frac{1}{2\pi\sigma^2}\int_0^\infty r e^{-\frac{1}{2\sigma^2}r^2} dr = \frac{1}{2\pi} \cdot 1(0 \leq \theta < 2\pi)$

可验证 $\int_0^\infty f_R(r)dr = \int_0^{2\pi} f_\Theta(\theta)d\theta = 1$, f_R、f_Θ 是密度, 对应的 r.v. 为 R、Θ

(R, Θ) 联合 p.d.f. $f(r,\theta) = f_R(r) f_\Theta(\theta)$, 故 R 与 Θ 独立

又 $R = \sqrt{X_1^2 + X_2^2}$, $X_1/X_2 = \tan\Theta$, 故 $\sqrt{X_1^2+X_2^2}$ 与 X_1/X_2 独立. $\square$

方法 2:

令 $Y_1 = \frac{X_1}{X_2}$, $Y_2 = \sqrt{X_1^2+X_2^2}$

这里有一点问题，它不是 1-1 的，

(y_1,y_2) 所在区域恰被"对称地"覆盖了两次；又 $P(X_2=0)=0$，故 $\{X_2=0\}$ 没有影响

故换元时应为

$$f_{Y_1,Y_2}(x_1,x_2) = 2 f_{X_1,X_2}(x_1,x_2)\left|\frac{\partial(x_1,x_2)}{\partial(y_1,y_2)}\right| = \frac{y_2}{\pi\sigma^2} e^{-\frac{1}{2\sigma^2}y_2^2}\cdot\frac{1}{1+y_1^2}$$

之后与方法 1 类似求出边缘密度，$f_{Y_1,Y_2} = f_{Y_1}\cdot f_{Y_2}$ 即得证 □

2021/09/17

[Wei] 2.16.

$$\because \quad n\bar{X} = \sum_{i=1}^{n} X_i \sim \chi^2_{mn}$$

$$\therefore \quad f_{\bar{X}}(x) = f_{n\bar{X}}(nx)\frac{d(nx)}{dx} = \frac{1}{\Gamma(\frac{mn}{2})2^{\frac{mn}{2}}}(nx)^{\frac{mn}{2}-1}e^{-\frac{nx}{2}}I_{(0,\infty)}(x)n = \frac{(\frac{n}{2})^{\frac{mn}{2}}}{\Gamma(\frac{mn}{2})}x^{\frac{mn}{2}-1}e^{-\frac{nx}{2}}I_{(0,\infty)}(x).$$

□

2.17 $E[X-EX]^4 = EX^4 - 4EX^3EX + 6EX^2(EX)^2 - 4(EX)^4 + (EX)^4$

- 利用 $EX^j = \frac{1}{i^j}\phi_X^{(j)}(0)$, $\phi_X(t) = (1-2it)^{-\frac{n}{2}}$ $(X\sim\chi^2_n)$

可求 $EX = n$, $EX^2 = n(n+2)$, $\cdots$, $EX^k = n(n+2)\cdots(n+2k-2)$

$\Rightarrow$ 峰度 $\beta = \frac{E[X-EX]^4}{(Var(X))^2} - 3 = \frac{12}{n}$

变异系数 $v = \frac{\sqrt{Var(X)}}{EX} = \sqrt{\frac{2}{n}}$

□

[Wei] 2.18. Hint for method 1: Notice that $\frac{X}{2} \sim \Gamma(n, 1)$ and integral by parts.

Hint for method 2: Notice that when $\lambda = 0$ the equality holds. It suffices to prove that

$$\frac{d}{d\lambda}\int_0^{2\lambda} \frac{1}{2^n\Gamma(n)} x^{n-1} e^{-\frac{x}{2}} dx = \frac{d}{d\lambda}\sum_{k=n}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!}.$$

[Wei] 2.19.

$$\because \quad f_{\chi_n^2}(x) = \frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{1}{2}x}, \quad x > 0$$

$$\therefore \quad f'_{\chi_n^2}(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} x^{-\frac{3}{2}} e^{-\frac{1}{2}x}\left(-\frac{1}{2}x - \frac{1}{2}\right), & n = 1 \\ -\frac{1}{4}e^{-\frac{1}{2}x}, & n = 2 \\ \frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}} x^{\frac{n}{2}-2} e^{-\frac{1}{2}x}\left(\frac{n}{2} - 1 - \frac{1}{2}x\right), & n \geq 3 \end{cases}.$$

Therefore, when $n \leq 2$, the maximum of density cannot be reached in the support set of ξ ($\min_{x>0} f_\xi(x) = \lim_{x\to 0+} f_\xi(x)$); when $n \geq 3$, the maximum is reached at $\xi = n - 2$.

$$\mathbb{E}(\xi^r) = \int_{(0,+\infty)} x^r \frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{1}{2}x} dx = \frac{2^r\Gamma(\frac{n}{2} + r)}{\Gamma(\frac{n}{2})},$$

where the last equality exists when $r > -\frac{n}{2}$.

$$\begin{aligned} Var(\xi^r) = \mathbb{E}(\xi^{2r}) - (\mathbb{E}(\xi^r))^2 &= \frac{2^{2r}\Gamma(\frac{n}{2} + 2r)}{\Gamma(\frac{n}{2})} - \left(\frac{2^r\Gamma(\frac{n}{2} + r)}{\Gamma(\frac{n}{2})}\right)^2 \\ &= \frac{2^{2r}}{\Gamma(\frac{n}{2})^2}\left[\Gamma(\frac{n}{2} + 2r)\Gamma(\frac{n}{2}) - \Gamma(\frac{n}{2} + r)^2\right], \end{aligned}$$

where the second equality exists when $r > -\frac{n}{4}$. □

[Wei] 2.11.

$$\because \quad \bar{X} \sim N(a, \frac{\sigma^2}{n}), X_{n+1} \sim N(a, \sigma^2) \text{ and } \frac{nS_n^2}{\sigma^2} \sim \chi_{n-1}^2 \text{ are independent distributed}$$

$$\therefore \quad X_{n+1} - \bar{X} \sim N\left(0, \left(1 + \frac{1}{n}\right)\sigma^2\right)$$

$$\therefore \quad \frac{X_{n+1} - \bar{X}}{S_n}\sqrt{\frac{n-1}{n+1}} = \frac{\frac{X_{n+1} - \bar{X}}{\sqrt{1+\frac{1}{n}}\sigma}}{\sqrt{\frac{nS_n^2/\sigma^2}{n-1}}} \sim t_{n-1}.$$

□

[Wei] 2.12.

$$\because \quad \bar{X} \sim N(\mu_1, \frac{\sigma^2}{m}), \bar{Y} \sim N(\mu_2, \frac{\sigma^2}{n}) \text{ and they are independent}$$

$$\therefore \quad \alpha(\bar{X} - \mu_1) + \beta(\bar{Y} - \mu_2) \sim N\left(0, \left(\frac{\alpha^2}{m} + \frac{\beta^2}{n}\right)\sigma^2\right)$$

$$\because \quad \frac{mS_{1m}^2}{\sigma^2} \sim \chi_{m-1}^2, \frac{nS_{2n}^2}{\sigma^2} \sim \chi_{n-1}^2 \text{ and they are independent}$$

$$\therefore \quad \frac{mS_{1m}^2 + nS_{2n}^2}{\sigma^2} \sim \chi_{m+n-2}^2 \text{ and is independent with terms above}$$

$$\therefore \quad T = \frac{\frac{\alpha(\bar{X}-\mu_1)+\beta(\bar{Y}-\mu_2)}{\sqrt{\frac{\alpha^2}{m}+\frac{\beta^2}{n}}\sigma}}{\sqrt{\frac{mS_{1m}^2+nS_{2n}^2/\sigma^2}{m+n-2}}} \sim t_{m+n-2}.$$

□

[Wei] 2.13. Without loss of generality, consider *i.i.d.* $X_i \sim N(0,1)$. (Notice that $\xi = \frac{(X_1-\bar{X})/\sigma}{S/\sigma}$.) Or equivalently, $\mathbf{X} \sim N(\mathbf{0}, I_n)$.

Consider an orthonomal basis $\{e_i\}_{i=1}^n$, where $e_1 = \frac{1}{\sqrt{n}}(1,\dots,1)^T$ and $e_2 = \sqrt{\frac{n}{n-1}}\left(1-\frac{1}{n}, -\frac{1}{n}, \dots, -\frac{1}{n}\right)^T$. (Check that e_1 and e_2 are orthogonal.) Let $E = (e_1, e_2, \dots, e_n)^T$, then we have $Y := E\mathbf{X} \sim N(\mathbf{0}, I_n)$, that is, $Y_i := e_i^T \mathbf{X} \sim N(0,1)$, *i.i.d.*.

$$\therefore \quad \sqrt{\frac{n}{n-1}}\left(X_1 - \bar{X}\right) = Y_2$$

Further, $(n-1)S^2 = \sum_{i=2}^n Y_i^2$ from the proof of the distribution of $\frac{(n-1)S^2}{\sigma^2}$ and now that $\sigma = 1$.

$$\therefore \quad \xi = \frac{\sqrt{\frac{n-1}{n}}Y_2}{\sqrt{\frac{1}{n-1}\sum_{i=2}^n Y_i^2}} = \frac{n-1}{\sqrt{n}} \frac{1}{\sqrt{1+\frac{1}{Y_2^2/\sum_{i=3}^n Y_i^2}}}$$

$$\because \quad \frac{Y_2}{\sqrt{\sum_{i=3}^n Y_i^2/(n-2)}} \sim t_{n-2}$$

$$\therefore \quad \xi \sim \frac{(n-1)t_{n-2}}{\sqrt{n\left(t_{n-2}^2 + n - 2\right)}}.$$

□

Remark:X $\sim$ N(0,1), $S^2 \sim \mathcal{X}_n^2$, we say $\frac{X}{\sqrt{S^2/n}} \sim t_n$, only if X and S are independent!