Statistics Note

by xianfei

Last math class, my younth ends!

先看附录!

先看附录!

先看附录!

目录

1 特征函数作业

Formula Or Theorem:

- **定义 4.2.1(P193)** 设 X 是一个随机变量，称
$$\varphi(t)=E(e^{itX}),\qquad -\infty<t<+\infty$$
为 X 的**特征函数**
- **性质 4.2.3(P194)** 若 $Y=aX+b$，其中 a，b 是常数，则
$$\varphi_Y(t)=e^{ibt}\varphi_X(at)$$
- **性质 4.2.4(P194)** 独立随机变量和的特征函数为每个随机变量的特征函数的积，即设 X 与 Y 相互独立，则
$$\varphi_{X+Y}(t)=\varphi_X(t)\varphi_Y(t)$$
- **性质 4.2.5(P194)** 若 $E(X^l)$ 存在，则 X 的特征函数 $\varphi(t)$ 可 l 次求导，且对 $1\le k\le l$，有
$$\varphi^{(k)}(0)=i^kE(X^k)$$
特别地
$$E(X)=\frac{\varphi'(0)}{i}\quad,\quad Var(X)=-\varphi''(0)+(\varphi'(0))^2$$

1.1
由定义 4.2.1 得

$$\begin{aligned}\varphi(t)=E(e^{itx_k})&=\sum_{k=0}^{3}e^{itk}p_k\\&=0.1\cdot e^{3it}+0.2\cdot e^{2it}+0.3\cdot e^{it}+0.4\end{aligned}$$

1.2
由定义 4.2.1 得

$$\varphi(t)=E(e^{itx_k})=\sum_{k=1}^{\infty}e^{itk}\cdot p(1-p)^{k-1}=pe^{it}\sum_{k=1}^{\infty}\left[e^{it}(1-p)\right]^{k-1}\tag{$*$}$$

式 $(*)$ 中，$|e^{it}(1-p)|\le|e^{it}|\cdot|(1-p)|<1$ 得

$$\varphi(t)=\frac{pe^{it}}{1-e^{it}(1-p)}$$

即得

$$\varphi^{'}(t)=\frac{ipe^{it}}{(1-e^{it}(1-p))^2}\qquad\qquad \varphi^{''}(t)=\frac{-pe^{it}(1+e^{it}(1-p))}{(1-e^{it}(1-p))^3}$$

由性质 4.2.5 得数学期望、方差

$$E(X) = \frac{\varphi^{'}(0)}{i} = \frac{1}{p}$$
$$Var(X) = -\varphi^{''}(0) + (\varphi^{'}(0))^2 = \frac{1-p}{p^2}$$

2 大数定律与中心极限定理作业

Formula Or Theorem:

- **定义 4.3.1 (P206)** 设有一随机变量序列 $\{X_n\}$，如果它对任意的 $\varepsilon > 0$，满足
$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)\right| < \varepsilon\right) = 1 \tag{4.3.1}$$
形式，则称该随机变量序列 $\{X_n\}$ 服从**大数定律**

- **定理 2.3.1(Chebyshev 不等式 P80)** 设随机变量 X 的数学期望和方差都存在，则对任意常数 $\varepsilon > 0$，有
$$P(|X - E(X)| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$$
或
$$P(|X - E(X)| < \varepsilon) \geq 1 - \frac{Var(X)}{\varepsilon^2}$$

- **定理 4.3.2(Chebyshev 大数定律 P206)** 设 $\{X_n\}$ 为一列两两不相关的随机变量序列，若每个 X_i 的方差存在，且有共同上界，即 $Var(X_i) \leq c$，$i = 1, 2, \cdots$，则 $\{X_n\}$ 服从大数定律，即对任意的 $\varepsilon > 0$，**式 4.3.1** 成立 **(详见书)**

- **定理 4.3.4(Khinchin 大数定律 P207)** 设 $\{X_n\}$ 为一独立同分布的随机变量序列，若 X_i 的数学期望存在，则 $\{X_n\}$ 服从大数定律，即对任意的 $\varepsilon > 0$，**式 4.3.1** 成立 **(详见书)**

2.1 **D**

由 $X \sim P(2)$，所以 $E(X) = 2$，$Var(X) = 2$

由 $Chebyshev$ 不等式

$$P\left(|X - E(X)| \geq \varepsilon\right) \leq \frac{Var(X)}{\varepsilon^2}$$
$$P\left(|X - E(X)| < \varepsilon\right) \geq 1 - \frac{Var(X)}{\varepsilon^2}$$

令 $\varepsilon = 2$ 即可

2.2 **D**

先求 $E(X+Y)$，$Var(X+Y)$

$$E(X+Y) = E(X) + E(Y) = 0$$

$$\begin{cases} Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y) \\ Corr(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} \end{cases}$$

解得 $Var(X+Y) = 3$，再由 $Chebyshev$ 不等式即可得到下界

2.3 **D**

由题意知 $X_i \sim Exp(2)$，即得 $E(X_i) = 1/2$，$E(X_i^2) = 1/2$

X_n 是独立同分布，由 $Khinchin$ 大数定理

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\frac{1}{n}\sum_{i=1}^{n} X_i)^2 \xrightarrow{P} E(X_n^2) - (E(X_n))^2 = \frac{1}{4}$$

2.4 **A**

由 $Khinchin$ 大数定理

$$\frac{1}{n}\sum_{i=1}^{n} X_i(X_i - 1) = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} E(X_n^2) - E(X_n) = 1$$

Formula Or Theorem:

- **定理 4.4.1(Lindeberg-Levy 中心极限定理 P212)** 设 X_n 是独立同分布的随机变量序列，且 $E(X_i) = \mu$，$Var(X_i) = \sigma^2 > 0$ 存在，若记

$$Y_n^* = \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

则对任意实数 y，有

$$\lim_{n\to\infty} P(Y_n^* \le y) = \Phi(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-\frac{t^2}{2}}dt$$

- **定理 4.4.2(de Moivre-Laplace 中心极限定理 P214)** 设 n 重伯努利实验中，事件 A 在每次试验中出现的概率为 $p(0 < p < 1)$，记 S_n 为 n 次试验中事件 A 出现的次数，且记

$$Y_n^* = \frac{S_n - np}{\sqrt{npq}}$$

则对任意实数 y，有

$$\lim_{n\to\infty} P(Y_n^* \le y) = \Phi(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-\frac{t^2}{2}}dt$$

2.5 **B**

设最多可以装载 n 件产品，由题意

$$P\left(\sum_{i=1}^{n} X_i \le 5000\right) = 0.99 \tag{$*$}$$

利用 $Lindeberg-Levy$ 中心极限定理得

$$(*) = P\left(\frac{\sum\limits_{i=1}^{n} X_i - 50n}{\sigma\sqrt{n}} \leq \frac{5000-50n}{\sigma\sqrt{n}}\right) \approx \Phi\left(\frac{5000-50n}{5\sqrt{n}}\right) = 0.99$$

解得 $n_{max} = 95$

2.6 **C**

设第 i 次称重的重量为 X_i，易知 $\overline{X}_n = \left(\sum\limits_{i=1}^{n} X_i\right)/n$

由 $Lindeberg-Levy$ 中心极限定理

$$P\left(\left|\frac{\sum\limits_{i=1}^{n} X_i - na}{\sigma\sqrt{n}}\right| < \frac{0.1n}{\sigma\sqrt{n}}\right) \approx 2\Phi\left(\frac{0.1n}{\sigma\sqrt{n}}\right) - 1 \geq 0.95$$

解得 $n_{min} = 16$

2.7 **B**

由题意得 $X_i \sim P(1)$，则 $E(X_i) = 1$，$Var(X_i) = 1$

$$\begin{aligned}\lim_{n\to\infty} P\left(\sum_{i=1}^{n} X_i > n\right) &= 1 - \lim_{n\to\infty} P\left(\sum_{i=1}^{n} X_i \leq n\right) \\ &= 1 - \lim_{n\to\infty} P\left(\frac{\sum\limits_{i=1}^{n} X_i - n}{\sigma\sqrt{n}} \leq 0\right) = 1 - \Phi(0) = 0.5\end{aligned}$$

2.8 **C**

易得

$$\begin{aligned}Var(X_i - Y_i) = Var(X_i) + Var(Y_i) = \frac{13}{3} \\ E(X_i - Y_i) = E(X_i) - E(Y_i) = 0\end{aligned}$$

由 $Lindeberg-Levy$ 中心极限定理

$$\begin{aligned}\lim_{n\to\infty} P\left(\sum_{i=1}^{n}(X_i - Y_i) > 0\right) &= 1 - \lim_{n\to\infty} P\left(\sum_{i=1}^{n}(X_i - Y_i) \leq 0\right) \\ &= 1 - P\left(\frac{\sum\limits_{i=1}^{n}(X_i - Y_i)}{\sigma\sqrt{n}} \leq 0\right) = 1 - \Phi(0) = 0.5\end{aligned}$$

2.9 **C**

由 $deMoivre-Laplace$ 中心极限定理得

$$\begin{aligned}P(14 < X < 30) &\approx \Phi\left(\frac{30-20+0.5}{4}\right) - \Phi\left(\frac{14-20-0.5}{4}\right) \\ &= \Phi(2.63) + \Phi(1.63) - 1 \approx 0.944\end{aligned}$$

未修正的结果为

$$P(14 < X < 30) = \Phi(2.5) + \Phi(1.5) - 1 \approx 0.927$$

2.10

设每袋味精的质量为随机变量 $X_i(i = 1, .., 200)$，记随机变量序列为 $\{X_n\}$，即求

$$P\left(\sum_{i=1}^{200} X_i > 20500\right) \tag{$*$}$$

由 $Lindeberg - Levy$ 中心极限定理得

$$(*) = P\left(\frac{\sum\limits_{i=1}^{200} X_i - 100 \cdot 200}{10 \cdot \sqrt{200}} > \frac{500}{100 \cdot \sqrt{2}}\right) \approx 1 - \Phi\left(\frac{5}{\sqrt{2}}\right) \approx 0.0002326$$

2.11

设每次命中的环数为随机变量 X_i，记随机变量序列为 $\{X_n\}$，其中

$$E(X_i) = 10 \times 0.8 + 9 \times 0.1 + 8 \times 0.05 + 7 \times 0.02 + 6 \times 0.03 = 9.62$$

$$Var(X_i) = E(X_i^2) - (E(X_i))^2 = 0.82，\sigma = \sqrt{Var(X_i)} = 0.91$$

求 100 次射击命中环数在 900 环到 930 环之间的概率，即求

$$P\left(900 \leq \sum_{i=1}^{n} X_i \leq 930\right) \tag{$*$}$$

由 $Lindeberg - Levy$ 中心极限定理得

$$\begin{aligned}(*) &= P\left(\frac{900 - 100 \times 9.62}{0.91 \times 10} \leq \frac{\sum\limits_{i=1}^{n} X_i - 100 \times 9.62}{0.91 \times 10} \leq \frac{930 - 100 \times 9.62}{0.91 \times 10}\right) \\ &\approx \Phi\left(\frac{930 - 100 \times 9.62}{0.91 \times 10}\right) - \Phi\left(\frac{900 - 100 \times 9.62}{0.91 \times 10}\right) \\ &= \Phi(6.81) - \Phi(3.52) \approx 0.0002325\end{aligned}$$

3 数理统计基本概念作业

Formula Or Theorem:

- **定义 5.3.1(P232)** 设 $x_1, x_2, \cdots, x_n$ 为取自某总体的样本，若样本函数 $T = T(x_1, x_2, \cdots, x_n)$ 中不含有任何未知参数，则称 T 为**统计量**，统计量的分布称为**抽样分布**
- **定义 5.3.2(P233)** 设 $x_1, x_2, \cdots, x_n$ 为取自某总体的样本，其算数平均值称为样本均值，一般用 $\bar{x}$ 表示，即
$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$
- **定义 5.3.3(P236)** 设 $x_1, x_2, \cdots, x_n$ 为取自某总体的样本，则它关于样本均值 $\bar{x}$ 的平均偏差平方和
$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{1}$$
称为**样本方差 (无偏方差)**，其算数平方根 $s = \sqrt{s^2}$ 称为**样本标准差**
- **定理 5.3.2(P237)** 设总体 X 具有二阶矩，即 $E(X) = \mu$，$Var(X) = \sigma^2 < \infty$，$x_1, x_2, \cdots, x_n$ 为从该总体得到的样本，$\bar{x}$ 和 s^2 分别是样本均值和样本方差，则
$$E(\bar{x}) = \mu \quad , \quad Var(\bar{x}) = \sigma^2/n$$
$$E(s^2) = \sigma^2$$

3.1 A

由**定理 5.3.2** 可知 $E(\bar{x}) = \mu$、$E(s^2) = \sigma^2$

$$\begin{aligned}
&E(X) = \frac{1}{2}\int_{-\infty}^{+\infty} xe^{-|x|}dx = \frac{1}{2}\int_{-\infty}^{0} xe^{x}dx + \frac{1}{2}\int_{0}^{+\infty} xe^{-x}dx = 0\\
&E(X^2) = \frac{1}{2}\int_{-\infty}^{+\infty} x^2e^{-|x|}dx = \frac{1}{2}\int_{-\infty}^{0} x^2e^{x}dx + \frac{1}{2}\int_{0}^{+\infty} x^2e^{-x}dx = 2\\
&Var(X) = E(X^2) - E(X)^2 = 2S
\end{aligned}$$

即得 $E(\bar{x}) = 0$、$E(s^2) = 2$

3.2 D

将 Y 展开

$$
\begin{aligned}
Y &= \sum_{i=1}^{n}(x_i + x_{i+n} - 2\bar{x})^2 \\
&= \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} x_{i+n}^2 + 4\sum_{i=1}^{n}\bar{x}^2 + 2\sum_{i=1}^{n} x_i x_{i+n} - 4\sum_{i=1}^{n} x_i\bar{x} - 4\sum_{i=1}^{n} x_{i+n}\bar{x} \\
&= \sum_{i=1}^{2n} x_i^2 + 4n\bar{x}^2 + 2\sum_{i=1}^{n} x_i x_{i+n} - 4\bar{x}\sum_{i=1}^{2n} x_i \\
&= \sum_{i=1}^{2n} x_i^2 + 2\sum_{i=1}^{n} x_i x_{i+n} - 4n\bar{x}^2
\end{aligned}
$$

进而可以得到

$$
\begin{aligned}
E(Y) &= E\left(\sum_{i=1}^{2n} x_i^2\right) + 2E\left(\sum_{i=1}^{n} x_i x_{i+n}\right) - E(4n\bar{x}^2) \\
&= 2n(\mu^2+\sigma^2) + 2n\mu^2 - 4n(\mu^2 + \frac{\sigma^2}{2n}) \\
&= 2(n-1)\sigma^2
\end{aligned}
$$

3.3 **C**

由 $X \sim N(0,1)$ 可得 $\mu = 0$、$\sigma^2 = 1$，又

$$
\begin{aligned}
Cov(Y_1, Y_n) &= E(Y_1Y_n) - E(Y_1)E(Y_n) \\
&= E\left((x_1-\bar{x})(x_n-\bar{x})\right) - E(\bar{x})^2 \\
&= E(x_1)E(x_n) - E((x_1+x_n)\bar{x}) + E(\bar{x}^2) - E(\bar{x})^2 \\
&= Var(\bar{x}) - E(x_1\bar{x}) - E(x_n\bar{x})
\end{aligned}
$$

考虑 $E(x_t\bar{x})$

$$
E(x_t\bar{x}) = \frac{1}{n}E\left(\sum_{i=1}^{n} x_t x_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(x_t)E(x_i) = \frac{n\mu^2}{n} = \mu^2 \tag{*}
$$

故而 $Cov(Y_1, Y_n) = Var(\bar{x}) = \sigma^2/n$

注：$(*)$ 式想要说明，简单随机抽样下，样本均值与样本是独立的

3.4 **D**

根据**定义 5.3.1**

Formula Or Theorem:

- **定理 5.3.3(P241)** 设总体 X 的密度函数为 $p(x)$，分布函数为 $F(x)$，$x_1, x_2, \cdots, x_n$ 为样本，则第 k 个次序统计量 $x_{(k)}$ 的密度函数为

$$p_k(x) = \frac{n!}{(k-1)!(n-k)!}(F(x))^{k-1}(1-F(x))^{n-k}p(x)$$

特别地

$$\begin{aligned} p_1(x) &= n \cdot (1-F(x))^{n-1}p(x) \\ p_n(x) &= n \cdot (F(x))^{n-1}p(x) \end{aligned}$$

3.5 **D**

由 $X \sim U(0,1)$，写出分布函数以及密度函数

$$F(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}, \quad p(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & others \end{cases}$$

由**定理 5.3.3**

$$\begin{aligned} p_1(x) &= n(1-x)^{n-1} \qquad (0 \le x \le 1) \\ p_n(x) &= nx^{n-1} \qquad (0 \le x \le 1) \end{aligned}$$

分别求出 $E(x_{(1)})$、$E(x_{(n)})$

$$E(x_{(1)}) = \int_0^1 nx(1-x)^{n-1}dx = n\frac{\Gamma(2)\Gamma(n)}{\Gamma(n+2)} = \frac{1}{n+1}$$

$$E(x_{(n)}) = \int_0^1 nx^n dx = n\frac{\Gamma(n+1)\Gamma(1)}{\Gamma(n+2)} = \frac{n}{n+1}$$

因此

$$E(x_{(n)}) - E(x_{(1)}) = \frac{n-1}{n+1}$$

注：P242　例 5.3.8

4 三大抽样分布作业

Formula Or Theorem:

- **定义 5.4.1、5.4.2、5.4.3** χ^2 分布 (P250)、F 分布 (P252)、t 分布 (P254)
- **定理 5.4.1(P251)** 设 $x_1, x_2, \cdots, x_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的样本，其样本均值和样本方差分别为
$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} \quad , \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$
则有
(1) $\bar{x}$和s^2相互独立
(2) $\bar{x} \sim N(\mu, \sigma^2/n)$
(3) $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$

4.1 **C**
这里未指明 X、Y 相互独立

4.2 **A**
易知
$$X_1 - 2X_2 \sim N(0, 20)$$
$$3X_3 - 4X_4 \sim N(0, 100)$$
所以有
$$\frac{X_1 - 2X_2}{\sqrt{20}} \sim N(0,1) \quad \Rightarrow \quad \frac{(X_1 - 2X_2)^2}{20} \sim \chi^2(1)$$
$$\frac{3X_3 - 4X_4}{\sqrt{100}} \sim N(0,1) \quad \Rightarrow \quad \frac{(3X_3 - 4X_4)^2}{100} \sim \chi^2(1)$$
易得 $a = 1/20, b = 1/100, n = 2$

4.3 **B**
由**定理 5.3.2** 可知 $E(s^2) = \sigma^2 = Var(X)$
$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx = 0$$
$$E(X^2) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} x^4 e^{-\frac{x^2}{2}}dx = \sqrt{\frac{2}{\pi}}\int_{0}^{+\infty} x^4 e^{-\frac{x^2}{2}}dx$$
$$= \frac{4}{\sqrt{\pi}}\int_{0}^{+\infty} u^{\frac{3}{2}} e^{-u}du = \frac{4}{\sqrt{\pi}}\Gamma\left(\frac{5}{2}\right) = 3$$
$$E(s^2) = Var(X) = E(X^2) - E(X)^2 = 3$$

4.4 **D**

$$E\left(\bar{x}^2-\frac{1}{n}s^2\right)=E(\bar{x}^2)-\frac{1}{n}E(s^2)$$

由**定理 5.4.1**，可知 $\bar{x}\sim N(0,1/n)$，易得

$$E(\bar{x}^2)=Var(\bar{x})+E(\bar{x})^2=\frac{1}{n}$$

而 $E(s^2)=\sigma^2=1$，所以

$$E\left(\bar{x}^2-\frac{1}{n}s^2\right)=0$$

由**定理 5.4.1**，可知 $\bar{x}$ 与 s^2 相互独立，因此

$$\begin{aligned}Var\left(\bar{x}^2-\frac{1}{n}s^2\right)&=Var(\bar{x}^2)+\frac{1}{n^2}Var(s^2)\\&=\frac{1}{n^2}Var(n\bar{x}^2)+\frac{1}{n^2(n-1)^2}Var((n-1)s^2)\\&=\frac{1}{n^2}Var\left(\left(\frac{\bar{x}}{1/\sqrt{n}}\right)^2\right)+\frac{1}{n^2(n-1)^2}Var\left(\frac{(n-1)s^2}{1}\right)\end{aligned}\tag{$*$}$$

由于

$$\frac{\bar{x}}{1/\sqrt{n}}\sim N(0,1)\qquad\Rightarrow\qquad\left(\frac{\bar{x}}{1/\sqrt{n}}\right)^2\sim\chi^2(1)$$

由**定理 5.4.1** 可以知道

$$\frac{(n-1)s^2}{1}\sim\chi^2(n-1)$$

因而有

$$Var\left(\left(\frac{\bar{x}}{1/\sqrt{n}}\right)^2\right)=2\quad,\quad Var\left(\frac{(n-1)s^2}{1}\right)=2(n-1)$$

带入 $(*)$ 即得

$$Var\left(\bar{x}^2-\frac{1}{n}s^2\right)=\frac{2}{n(n-1)}$$

4.5 **C**

(法 1) 因为 $X\sim F(n,n)$，所以 $1/X\sim F(n,n)$ 所以有

$$P(X<1)=P\left(\frac{1}{X}<1\right)=P(X>1)$$

当然又有

$$P(X<1)+P(X>1)=1$$

因此

$$P(X<1)=\frac{1}{2}$$

(法 2) 因为 $X \sim F(n,n)$，记 X 的密度函数为 $p(x)$，则

$$p(x) = \frac{\Gamma(n)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1}(1+x)^{-n}$$

所以

$$P(X<1) = \frac{\Gamma(n)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}\right)} \int_0^1 x^{\frac{n}{2}-1}(1+x)^{-n}dx \tag{$*$}$$

考虑 $(*)$ 式的积分部分

$$\int_0^1 x^{\frac{n}{2}-1}(1+x)^{-n}dx = \int_0^1 (\sqrt{x})^{n-2}(1+x)^{-n}dx \tag{I}$$

令 $t=\sqrt{x}$，则 $dx = 2tdt$，带入 (I) 式

$$(I) = \int_0^1 \frac{t^{n-1}}{(1+t^2)^n}dt \tag{II}$$

这里令 $t = tanu$，则 $dt = sec^2udu$，带入 (II) 式

$$\begin{aligned}(II) &= 2\int_0^{\frac{\pi}{4}} \frac{tan^{n-1}u}{sec^{2n}u}\cdot sec^2udu = 2\int_0^{\frac{\pi}{4}} sin^{n-1}u\cdot cos^{n-1}udu \\ &= \frac{1}{2^{n-2}}\int_0^{\frac{\pi}{4}}(2\cdot sinu\cdot cosu)^{n-1}du = \frac{1}{2^{n-2}}\int_0^{\frac{\pi}{4}} sin^{n-1}2udu \\ &= \frac{1}{2^{n-1}}\int_0^{\frac{\pi}{2}} sin^{n-1}vdv\end{aligned}$$

当 $2\nmid n$ **时**

$$\int_0^{\frac{\pi}{2}} sin^{n-1}vdv = \frac{(n-2)!!}{(n-1)!!}\frac{\pi}{2}$$

带入 $(*)$ 式得

$$(*) = \frac{\Gamma(n)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}\right)}\cdot\frac{1}{2^{n-1}}\cdot\frac{(n-2)!!}{(n-1)!!}\frac{\pi}{2} = \frac{2^{n-1}\cdot(n-1)!}{((n-2)!!)^2\cdot\pi}\cdot\frac{1}{2^{n-1}}\cdot\frac{(n-2)!!}{(n-1)!!}\frac{\pi}{2} = \frac{1}{2}$$

当 $2\mid n$ **时**，同理可得 $(*) = 1/2$

注

$$I_m = \int_0^{\frac{\pi}{2}} cos^m xdx = \int_0^{\frac{\pi}{2}} sin^m xdx$$

则有

$$I_m = \begin{cases} \frac{(m-1)!!}{m!!}\cdot\frac{\pi}{2}, & 2\mid m \\ \frac{(m-1)!!}{m!!}, & 2\nmid m \end{cases}$$

5 三大抽样分布第二次作业

Formula Or Theorem:

- **推论 5.4.1(P254)** 设 $x_1, x_2, \cdots, x_m$ 是来自 $N(\mu_1, \sigma_1^2)$ 的样本，$y_1, y_2, \cdots, y_n$ 是来自 $N(\mu_2, \sigma_2^2)$ 的样本，且两样本相互独立，记
$$s_x^2 = \frac{1}{m-1}\sum_{i=1}^{m}(x_i - \bar{x})^2 \quad , \quad s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$$
则有
$$F = \frac{s_x^2/\sigma_1^2}{s_y^2/\sigma_2^2} \sim F(m-1, n-1)$$
特别，若 $\sigma_1^2 = \sigma_2^2$，则 $F = s_x^2/s_y^2 \sim F(m-1, n-1)$

- **推论 5.4.2(P256)** 设 $x_1, x_2, \cdots, x_n$ 是来自正态分布 $N(\mu, \sigma^2)$ 的一个样本，$\bar{x}$ 与 s^2 分别是该样本的样本均值与样本方差，则有
$$t = \frac{\sqrt{n}(\bar{x} - \mu)}{s} \sim t(n-1)$$

5.1 **A**

容易知道 X、Y 相互独立 (P133 $\rho = 0$)

$$f(x, y) = \frac{1}{12\pi}e^{-\frac{x^2}{8}} \cdot e^{-\frac{(y-1)^2}{18}} = \frac{1}{\sqrt{2\pi}\cdot 2}e^{-\frac{x^2}{2\cdot 2^2}} \cdot \frac{1}{\sqrt{2\pi}\cdot 3}e^{-\frac{(y-1)^2}{2\cdot 3^2}}$$

即可得到

$$X \sim N(0, 4) \qquad and \qquad Y \sim N(1, 9)$$

则

$$\left(\frac{X}{2}\right)^2 \sim \chi^2(1) \quad and \quad \left(\frac{Y-1}{3}\right)^2 \sim \chi^2(1)$$

即

$$\frac{9}{4} \cdot \frac{X^2}{(Y-1)^2} \sim F(1, 1)$$

5.2 **D**

这里仅证明 D 选项

$$\begin{aligned}\frac{1}{2}\sum_{i=1}^{2n}X_i^2 + \sum_{i=1}^{n}X_{2i-1}X_{2i} &= \frac{1}{2}\left(\sum_{i=1}^{2n}X_i^2 + 2\sum_{i=1}^{n}X_{2i-1}X_{2i}\right) \\ &= \frac{1}{2}\sum_{i=1}^{n}\left(X_{2i-1}^2 + 2X_{2i-1}X_{2i} + X_{2i}^2\right) \qquad (*) \\ &= \sum_{i=1}^{n}\left(\frac{X_{2i-1} + X_{2i}}{\sqrt{2}}\right)^2\end{aligned}$$

由于 $X_{2i-1} + X_{2i} \sim N(0, 2)$，所以 $(*) \sim \chi^2(n)$

5.3 **D**

因为 $X \sim N(0,\sigma^2)$，所以 $\bar{x}^2$、s^2 相互独立

$$\begin{aligned} Var(\hat{\sigma}^2) &= Var(cn\bar{x}^2) + Var((1-c)s^2) \\ &= c^2 Var(n\bar{x}^2) + (1-c)^2 Var(s^2) \\ &= \sigma^4 c^2 Var\left(\frac{n\bar{x}^2}{\sigma^2}\right) + \frac{(1-c)^2\sigma^4}{(n-1)^2} Var\left(\frac{(n-1)s^2}{\sigma^2}\right) \end{aligned} \tag{$*$}$$

由**定理 5.4.1**

$$\bar{x} \sim N\left(0, \frac{\sigma^2}{n}\right) \quad and \quad \frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$

即

$$\left(\frac{\sqrt{n}\bar{x}}{\sigma}\right)^2 \sim \chi^2(1)$$

带入 $(*)$ 式即得

$$(*) = 2\sigma^4 c^2 + \frac{2(1-c)^2\sigma^4}{(n-1)} = \frac{2\sigma^4}{n-1}(nc^2 - 2c + 1)$$

易知，当 $c = 1/n$ 时，$(*)$ 式取最小值

5.4 **A**

易得

$$\frac{\sum\limits_{k=1}^{i} X_k^2/(\sigma^2 \cdot i)}{\sum\limits_{k=i+1}^{10} X_k^2/(\sigma^2 \cdot (10-i))} = \frac{\sum\limits_{k=1}^{i} X_k^2}{\sum\limits_{k=i+1}^{10} X_k^2} \cdot \frac{10-i}{i}$$

所以

$$\frac{10-i}{i} = 4 \quad \Rightarrow \quad i = 2$$

5.5 **C**

$$\begin{aligned} &P(|X| < x) = 1 - P(|X| \geq x) \\ \Rightarrow &P(|X| \geq x) = 1 - \alpha \\ = &P(X \leq -x) + P(X \geq x) = 2P(X \geq x) \end{aligned}$$

$$\Rightarrow P(X \geq x) = \frac{1-\alpha}{2}$$

即 $x = U_{(1-\alpha)/2}$

6 矩估计作业

Formula Or Theorem:

- **替换原理与矩法估计 (P272)**
  (1) 用样本矩去替换总体矩，这里的矩可以是原点矩也可以是中心矩
  (2) 用样本矩的函数去替换相应的总体矩的函数

对于一个待估计参数 θ，需要用样本均值替换总体均值；对于两个待估计参数 θ、λ，需要用样本均值替换总体均值，样本方差替换总体方差

6.1 D

首先求出总体均值 $E(X)$

$$E(X) = -1\cdot 2\theta + 0\cdot\theta + 1\cdot(1-3\theta) = 1-5\theta$$

用样本均值 $\bar{x}$ 替换总体均值

$$1-5\hat{\theta} = \bar{x} \quad \Rightarrow \quad \hat{\theta} = \frac{1-\bar{x}}{5}$$

6.2 D

请读者验证

6.3 C

求总体均值 $E(X)$

$$\begin{aligned} E(X) &= \int_0^\theta \frac{6x^2}{\theta^3}(\theta-x)dx = 6\theta\int_0^\theta \left(\frac{x}{\theta}\right)^2\cdot\left(1-\frac{x}{\theta}\right)d\left(\frac{x}{\theta}\right) \\ &= 6\theta\int_0^1 t^2(1-t)dt = 6\theta\cdot\frac{\Gamma(3)\Gamma(2)}{\Gamma(5)} = \frac{\theta}{2} \qquad \textbf{[附录 I]} \end{aligned}$$

用样本均值 $\bar{x}$ 替换总体均值

$$\frac{\hat{\theta}}{2} = \bar{x} \quad \Rightarrow \quad \hat{\theta} = 2\bar{x}$$

那么

$$Var(\hat{\theta}) = 4Var(\bar{x}) = \frac{4\sigma^2}{n}$$

需要求出 σ^2，即 $Var(X)$

$$\begin{aligned} E(X^2) &= \int_0^\theta \frac{6x^3}{\theta^3}(\theta-x)dx = 6\theta^2\int_0^\theta \left(\frac{x}{\theta}\right)^3\cdot\left(1-\frac{x}{\theta}\right)d\left(\frac{x}{\theta}\right) \\ &= 6\theta^2\int_0^1 t^3(1-t)^3dt = 6\theta^2\cdot\frac{\Gamma(4)\Gamma(2)}{\Gamma(6)} = \frac{3\theta^2}{10} \end{aligned}$$

$$Var(X) = E(X^2) - E(X)^2 = \frac{\theta^2}{20}$$

因此 $Var(\hat{\theta})=\theta^2/5n$

6.4 C

利用积分求 $E(X)$、$Var(X)$

$$\begin{aligned}E(X)&=\frac{1}{\lambda}e^{\frac{\theta}{\lambda}}\int_\theta^{+\infty}xe^{-\frac{x}{\lambda}}dx=\frac{1}{\lambda}e^{\frac{\theta}{\lambda}}\cdot(-\lambda)\int_\theta^{+\infty}xde^{-\frac{x}{\lambda}}\\&=-e^{\frac{\theta}{\lambda}}\left(xe^{-\frac{x}{\lambda}}\Big|_\theta^{+\infty}-\int_\theta^{+\infty}e^{-\frac{x}{\lambda}}dx\right)=-e^{\frac{\theta}{\lambda}}\left(-\theta e^{-\frac{\theta}{\lambda}}+\lambda e^{-\frac{x}{\lambda}}\Big|_\theta^{+\infty}\right)\\&=\lambda+\theta\end{aligned}$$

$$\begin{aligned}E(X^2)&=\frac{1}{\lambda}e^{\frac{\theta}{\lambda}}\int_\theta^{+\infty}x^2e^{-\frac{x}{\lambda}}dx=\frac{1}{\lambda}e^{\frac{\theta}{\lambda}}\cdot(-\lambda)\int_\theta^{+\infty}x^2de^{-\frac{x}{\lambda}}\\&=-e^{\frac{\theta}{\lambda}}\left(x^2e^{-\frac{x}{\lambda}}\Big|_\theta^{+\infty}-2\int_\theta^{+\infty}xe^{-\frac{x}{\lambda}}dx\right)=-e^{\frac{\theta}{\lambda}}\left(-\theta^2e^{-\frac{x}{\lambda}}-2I\right)\end{aligned}$$

由上可知

$$\frac{1}{\lambda}e^{\frac{\theta}{\lambda}}\cdot I=\frac{1}{\lambda}e^{\frac{\theta}{\lambda}}\int_\theta^{+\infty}x^2e^{-\frac{x}{\lambda}}dx=\lambda+\theta$$

即得到

$$I=\lambda(\lambda+\theta)e^{\frac{\theta}{\lambda}}$$

带入得

$$E(X^2)=-e^{\frac{\theta}{\lambda}}\left(-\theta^2e^{-\frac{x}{\lambda}}-2\lambda(\lambda+\theta)e^{-\frac{\theta}{\lambda}}\right)=\theta^2+2\lambda^2+2\lambda\theta$$

即得

$$Var(X)=E(X^2)-E(X)^2=\lambda^2$$

令 $\bar{x}=\lambda+\theta \quad and \quad s^2=\lambda^2$ 即可

这里复习一下 4.2 特征函数求均值和方差

求 $p(x)$ 特征函数

$$\begin{aligned}\varphi(t)&=\int_\theta^{+\infty}e^{itx}\cdot\frac{1}{\lambda}e^{-\frac{1}{\lambda}(x-\theta)}dx=\frac{1}{\lambda}e^{\frac{\theta}{\lambda}}\int_\theta^{+\infty}e^{(-\frac{1}{\lambda}+it)x}dx\\&=\frac{1}{\lambda}e^{-\frac{1}{\lambda}}\cdot\frac{1}{it-1/\lambda}\cdot e^{(-\frac{1}{\lambda}+it)x}\Big|_\theta^{+\infty}=\frac{e^{i\theta t}}{1-i\lambda t}\end{aligned}$$

$$\varphi'(t)=\frac{i\theta e^{i\theta t}(1-i\lambda t)+i\lambda e^{i\theta t}}{(1-i\lambda t)^2}=e^{i\theta t}\cdot\frac{i(\lambda+\theta)+\lambda\theta t}{(1-i\lambda t)^2}$$

$$\varphi''(t)=i\theta e^{i\theta t}\cdot\frac{i(\lambda+\theta)+\lambda\theta t}{(1-i\lambda t)^2}+e^{i\theta t}\cdot\frac{-2\lambda^2-\lambda\theta+\lambda^2\theta ti}{(1-i\lambda t)^3}$$

即有

$$\varphi'(0) = i(\lambda+\theta) \quad and \quad \varphi''(0) = i\theta \cdot i(\lambda+\theta) + (-2\lambda^2 - \lambda\theta)$$
$$= -(\theta^2 + 2\lambda^2 + 2\lambda\theta)$$

由性质 **4.2.5** 得

$$E(X) = \frac{\varphi'(0)}{i} = \lambda + \theta$$

$$Var(X) = -\varphi''(0) + (\varphi'(0))^2 = \lambda^2$$

令 $\bar{x} = \lambda + \theta \quad and \quad s^2 = \lambda^2$ 即可

6.5 **B**

请读者验证

7 极大似然估计作业

Formula Or Theorem:

- **定义 6.3.1(P278)** 设总体的概率函数为 $p(x|\theta)$，$\theta \in \Theta$，其中 θ 是一个未知参数或几个未知参数组成的参数向量，Θ 是参数空间，$x_1, x_2, \cdots, x_n$ 是来自该总体的样本，将样本的联合概率函数看成 θ 的函数，用 $L(\theta|x_1, x_2, \cdots, x_n)$ 表示，简记为 $L(\theta)$，

$$L(\theta) = L(\theta|x_1, x_2, \cdots, x_n) = p(x_1|\theta)P(x_2|\theta)\cdots p(x_n|\theta)$$

$L(\theta)$ 称为**似然函数**. 如果某统计量 $\hat{\theta} = \hat{\theta}(x_1, x_2, \cdots, x_n)$ 满足

$$L(\hat{\theta}) = \max_{\theta \in \Theta} L(\theta)$$

则称 $\hat{\theta}$ 是 θ 的**最大似然估计**

- In brief, a statistic $\hat{\theta}(x_1, x_2, \cdots, x_n)$ is a **maximum likelihood estimator** of θ if, for each sample $x_1, x_1, \cdots, x_n, \hat{\theta}(x_1, x_2, \cdots, x_n)$ is a value for the parameter that maximizes the likelihood function $L(\theta|x_1, x_2, \cdots, x_n)$

7.1 **C**

对 $F(x|\theta)$ 求导得到密度函数

$$p(x|\theta)=\begin{cases}\frac{2\theta}{x^3}, & x\geq\theta(\theta>1)\\ 0, & others\end{cases}$$

写出似然函数

$$L(\theta)=\prod_{i=1}^{n}p(x|\theta)=\prod_{i=1}^{n}\frac{2\theta}{x_i^3}\qquad(x_i\geq\theta)$$

两边取对数得

$$lnL(\theta)=nln2\theta-\sum_{i=1}^{n}lnx_i^3$$

显然 $L(\theta)$ 单调递增，又有 $x_i\geq\theta$，所以 $\hat{\theta}=min\{x_i\}=x_{(1)}$

7.2 **B**

写出似然函数

$$L(\theta)=\prod_{i=1}^{n}(1+\theta)x_i^{\theta}$$

两边取对数得

$$lnL(\theta)=nln(1+\theta)+\theta\sum_{i=1}^{n}lnx_i$$

对 θ 求导

$$\frac{d}{d\theta}lnL(\theta)=\frac{n}{1+\theta}+\sum_{i=1}^{n}lnx_i$$

令导数等于 0，求导极大值点为

$$\hat{\theta}=-\frac{n}{\sum\limits_{i=1}^{n}lnx_i}-1$$

7.3 **B**

请读者验证

7.4 **C**

求出总体均值

$$E(X)=-\theta^2+(1-\theta)^2=1-2\theta$$

用样本均值替换总体均值

$$1-2\theta=\bar{x}\quad\Rightarrow\quad\bar{\theta}_M=\frac{1-\bar{x}}{2}$$

写出似然函数

$$\begin{aligned}L(\theta)&=\theta^{2N_1}\cdot(2\theta(1-\theta))^{N_2}\cdot(1-\theta)^{2(n-N_1-N_2)}\\&=2^{N_2}\cdot\theta^{2N_1+N_2}\cdot(1-\theta)^{2n-2N_1-N_2}\end{aligned}$$

令 $s=2N_1+N_2$，$t=2n-2N_1-N_2$ ，对 θ 求导

$$L(\theta)=2^{N_2}\cdot\theta^{s}\cdot(1-\theta)^{t}$$

$$\frac{d}{d\theta}L(\theta) = 2^{N_2}\cdot\theta^{s-1}(1-\theta)^{t-1}(s-(s+t)\theta)$$

求出极大值点为

$$\hat{\theta}_L = \frac{s}{s+t} = \frac{2N_1+N_2}{2n}$$

7.5 **A**

求出总体均值

$$E(X) = \sum_{t=1}^{N}\frac{t}{N} = \frac{1+N}{2}$$

用样本均值替换总体均值

$$\frac{1+N}{2} = \bar{x} \quad\Rightarrow\quad \hat{N}_M = 2\bar{x}-1$$

写出似然函数

$$L(N) = \prod_{i=1}^{N}\frac{1}{N} = \left(\frac{1}{N}\right)^N \qquad (x_i\leq N)$$

显然 $L(N)$ 单调递减，又有 $x_i\leq N$，所以 $\hat{N}_L = max\{x_i\} = x_{(n)}$

8 估计量的评判标准作业

Formula Or Theorem:

- **定义 6.2.1(P267)** 设 $\hat{\theta}=\hat{\theta}(x_1,x_2,\cdots,x_n)$ 是 θ 的一个估计，θ 的参数空间为 Θ，若对任意的 $\theta\in\Theta$，有

  $$E(\hat{\theta}) = \theta$$

  则称 $\hat{\theta}$ 是 θ 的**无偏估计**，否则称为**有偏估计**

- A statistic $\hat{\theta}$ is said to be an **unbiased estimator**, or its value an unbiased estimate, if and only if the mean of the sampling distribution of the estimator $E(\hat{\theta})=\theta$, whatever the value of θ

8.1 **C**

因为 $x_i\sim N(\mu,\sigma^2)$，容易知道

$$x_{i+1}-x_i\sim N(0,2\sigma^2)$$

标准化后得

$$\frac{x_{i+1}-x_i}{\sqrt{2}\sigma}\sim N(0,1) \quad\Rightarrow\quad \left(\frac{x_{i+1}-x_i}{\sqrt{2}\sigma}\right)^2\sim\chi^2(1)$$

求和得

$$\sum_{i=1}^{n-1}\left(\frac{x_{i+1}-x_i}{\sqrt{2}\sigma}\right)^2\sim\chi^2(n-1)$$

求数学期望得

$$\frac{1}{2\sigma^2}E\left(\sum_{i=1}^{n-1}(x_{i+1}-x_i)^2\right)=n-1$$

即得

$$cE\left(\sum_{i=1}^{n-1}(x_{i+1}-x_i)^2\right)=2c\sigma^2(n-1)$$

因为目标统计量是 σ^2 的无偏估计，可知

$$2c\sigma^2(n-1)=\sigma^2 \quad\Rightarrow\quad c=\frac{1}{2(n-1)}$$

8.2 **A**

求数学期望

$$E(\bar{x}+cs^2)=E(\bar{x})+cE(s^2)=np(1+c-cp)$$

由**定义 6.2.1**

$$1+c-cp=p \quad\Rightarrow\quad c=-1$$

8.3 **D**

这里仅证明 D 选项先求分布函数

$$F(x)=\begin{cases}1-e^{-(x-\theta)}, & x\geq\theta\\ 0, & x<\theta\end{cases}$$

根据**定理 5.3.3**

$$p_1(x)=n\cdot(1-F(x))^{n-1}p(x)$$

带入得到

$$p_1(x)=ne^{-n(x-\theta)}$$

求 $x_{(1)}$ 的数学期望

$$\begin{aligned}E(x_{(1)})&=\int_\theta^{+\infty}nxe^{-n(x-\theta)}dx=\int_0^{+\infty}n(t+\theta)e^{-nt}dt\\&=\frac{1}{n}\int_0^{+\infty}nte^{-nt}d(nt)+n\theta\int_0^{+\infty}e^{-nt}dt\\&=\frac{1}{n}\Gamma(2)+\theta=\frac{1}{n}+\theta\end{aligned}$$

所以 $E(x_{(1)}-1/n)=\theta$

8.4 **C**

直接求 $E(\hat{p})$

$$\begin{aligned}E(\hat{p})&=cE\left(\sum_{i=1}^{n}x_i(x_i-1)\right)=c\sum_{i=1}^{n}\big(E(x_i^2)-E(x_i)\big)\\&=c\sum_{i=1}^{n}\big(Var(x_i)+E(x_i)^2-E(x_i)\big)=c\sum_{i=1}^{n}(mp(1-p)+m^2p^2-mp)\\&=cnp^2m(m-1)\end{aligned}$$

由于 $\hat{p}$ 是 p^2 的无偏估计，根据**定义 6.2.1**

$$cnp^2m(m-1) = p^2 \quad \Rightarrow \quad c = \frac{1}{mn(m-1)}$$

Formula Or Theorem:

- **定义 6.1.3** 设 $\hat{\theta}_1$，$\hat{\theta}_2$ 是 θ 的两个无偏估计，如果对任意的 $\theta \in \Theta$ 有

$$Var(\hat{\theta}_1) \leq Var(\hat{\theta}_2)$$

  且至少有一个 $\theta \in \Theta$ 使得上述不等号严格成立，则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效

- A statistic $\hat{\theta}_1$ is said to be a **more efficient unbiased estimator** of the parameter θ than the statistic $\hat{\theta}_2$ if
  (1) $\hat{\theta}_1$ and $\hat{\theta}_2$ are both unbiased estimators of θ
  (2) the variance of the sampling distribution of the first estimator is no larger than that of the second and is smaller for at least one value of θ

8.5 **C**

因为 $X \sim U(0, \theta)$，写出总体的分布函数以及密度函数

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{\theta}, & 0 \leq x < \theta \\ 1, & x \geq \theta \end{cases} \quad , \quad p(x) = \begin{cases} \frac{1}{\theta}, & 0 \leq x \leq \theta \\ 0, & others \end{cases}$$

求出 $x_{(1)}$ 的密度函数，根据**定理 5.3.3**

$$p_1(x) = n \cdot (1 - F(x))^{n-1} p(x) = n\left(1 - \frac{x}{\theta}\right)^{n-1} \frac{1}{\theta}$$

求出 $E(x_{(1)})$

$$\begin{aligned} E(x_{(1)}) &= \frac{n}{\theta}\int_0^\theta x\left(1-\frac{x}{\theta}\right)^{n-1} dx = \theta n \int_0^\theta \frac{x}{\theta}\cdot\left(1-\frac{x}{\theta}\right)^{n-1} d\left(\frac{x}{\theta}\right) \\ &= \theta n \int_0^1 t(1-t)^{n-1} dt = \theta n \frac{\Gamma(2)\Gamma(n)}{\Gamma(n+2)} = \frac{\theta}{n+1} \qquad \text{[附录 I]} \end{aligned}$$

求出 $Var(x_{(1)})$

$$\begin{aligned} E(x_{(1)}^2) &= \frac{n}{\theta}\int_0^\theta x^2\left(1-\frac{x}{\theta}\right)^{n-1} dx = \theta^2 n \int_0^\theta \left(\frac{x}{\theta}\right)^2\cdot\left(1-\frac{x}{\theta}\right)^{n-1} d\left(\frac{x}{\theta}\right) \\ &= \theta^2 n \int_0^1 t^2(1-t)^{n-1} dt = \theta^2 n \frac{\Gamma(3)\Gamma(n)}{\Gamma(n+3)} = \frac{2\theta^2}{(n+2)(n+1)} \end{aligned}$$

$$Var(x_{(1)}) = E(x_{(1)}^2) - E(x_{(1)})^2 = \frac{n\theta^2}{(n+2)(n+1)^2}$$

求出 $x_{(n)}$ 的密度函数，根据**定理 5.3.3**

$$p_n(x) = n \cdot (F(x))^{n-1} p(x) = n \left(\frac{x}{\theta}\right)^{n-1} \frac{1}{\theta}$$

求出 $E(x_{(n)})$

$$\begin{aligned} E(x_{(n)}) &= \frac{n}{\theta} \int_0^\theta x \left(\frac{x}{\theta}\right)^{n-1} dx = \theta n \int_0^\theta \left(\frac{x}{\theta}\right)^n d\left(\frac{x}{\theta}\right) \\ &= \theta n \int_0^1 t^n (1-t)^0 dt = \theta n \frac{\Gamma(n+1)\Gamma(1)}{\Gamma(n+2)} = \frac{n\theta}{n+1} \end{aligned}$$

求出 $Var(x_{(n)})$

$$\begin{aligned} E(x_{(n)}^2) &= \frac{n}{\theta} \int_0^\theta x^2 \left(\frac{x}{\theta}\right)^{n-1} dx = \theta^2 n \int_0^\theta \left(\frac{x}{\theta}\right)^{n+1} d\left(\frac{x}{\theta}\right) \\ &= \theta^2 n \int_0^1 t^{n+1} (1-t)^0 dt = \theta^2 n \frac{\Gamma(n+2)\Gamma(1)}{\Gamma(n+3)} = \frac{\theta^2 n}{n+2} \end{aligned}$$

$$Var(x_{(n)}) = E(x_{(n)}^2) - E(x_n)^2 = \frac{n\theta^2}{(n+2)(n+1)^2}$$

求 $E(\hat{\theta_1})$ 与 $E(\hat{\theta_2})$

$$E(\hat{\theta_1}) = \frac{n+1}{n} \cdot \frac{n\theta}{n+1} = \theta$$

$$E(\hat{\theta_2}) = (n+1) \cdot \frac{\theta}{n+1} = \theta$$

所以 $\hat{\theta_1}$ 与 $\hat{\theta_2}$ 都是 θ 的无偏估计，再求 $Var(\hat{\theta_1})$ 和 $Var(\hat{\theta_2})$

$$Var(\hat{\theta_1}) = \left(\frac{n+1}{n}\right)^2 \cdot \frac{n\theta^2}{(n+2)(n+1)^2} = \frac{\theta^2}{n+2}$$

$$Var(\hat{\theta_2}) = (n+1)^2 \cdot \frac{n\theta^2}{(n+2)(n+1)^2} = \frac{n\theta^2}{n+2}$$

显然 $Var(\hat{\theta_1}) < Var(\hat{\theta_2})$，所以 $\hat{\theta_1}$ 更有效

8.6 **D**

由**定理 5.4.1**

$$\bar{x} \sim N(0, \sigma^2/n) \quad \Rightarrow \quad \frac{\bar{x}}{\sigma/\sqrt{n}} \sim N(0,1)$$

即有

$$\left(\frac{\bar{x}}{\sigma^2/\sqrt{n}}\right)^2 \sim \chi^2(1)$$

所以

$$E(\bar{x}^2) = \frac{n}{\sigma^2}$$

由**定理 5.4.1**

$$\frac{x_n}{\sigma} \sim N(0,1) \quad \Rightarrow \quad \left(\frac{x_n}{\sigma}\right)^2 \sim \chi^2(1)$$

所以

$$E(x_n) = \sigma^2 \quad , \quad Var(x_n) = 2\sigma^4$$

由**定理 5.4.1**

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$

所以

$$E(s^2) = \sigma^2 \quad , \quad Var(s^2) = \frac{2\sigma^4}{n-1}$$

由**定理 5.4.1**

$$\frac{x_i}{\sigma} \sim N(0,1) \quad \Rightarrow \quad \left(\frac{x_i}{\sigma}\right)^2 \sim \chi^2(1)$$

即有

$$\sum_{i=1}^{n} \left(\frac{x_i}{\sigma}\right)^2 \sim \chi^2(n)$$

所以

$$E\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) = \sigma^2 \quad , \quad Var\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) = \frac{2\sigma^4}{n}$$

因此 $\sum\limits_{i=1}^{n} x_i^2/n$ 是最有效的无偏估计量

8.7 **D**

容易知道

$$T \sim N((4a+9b)\mu, 4a^2+9b^2)$$

因为 T 是 μ 的无偏估计，所以

$$4a+9b=1$$

当 $4a^2+9b^2$ 最小时，T 最有效，带入即可知道 $a=4/25, b=1/25$

8.8 **B**

请读者验证

9 区间估计作业

Formula Or Theorem:

- **定义 6.6.1(P300)** 设 θ 是总体的一个参数，其参数空间为 Θ，$x_1, x_2, \cdots, x_n$ 是来自该总体的样本，对给定的一个 $\alpha(0 < \alpha < 1)$，假设有两个统计量 $\hat{\theta_L} = \hat{\theta_L}(x_1, x_2, \cdots, x_n)$ 和 $\hat{\theta_U} = \hat{\theta_U}(x_1, x_2, \cdots, x_n)$，若对任意的 $\theta \in \Theta$，有

$$P_\theta(\hat{\theta_L} \leq \theta \leq \hat{\theta_U}) \geq 1 - \alpha$$

则称随机区间 $[\hat{\theta_L}, \hat{\theta_U}]$ 为 θ 的**置信水平**为 $1 - \alpha$ 的**置信区间**，$\hat{\theta_L}$ 和 $\hat{\theta_U}$ 分别称为 θ 的**置信下限和置信上限**

- **Large sample confidence interval for μ (σ konwn)**

$$\bar{x} - u_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + u_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

9.1 **A**

由题意得

$$\begin{cases} \bar{x} - u_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} & = 43.88 \\ \bar{x} + u_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} & = 46.52 \end{cases}$$

消去 $\bar{x}$

$$1.6u_{1-\alpha/2} = 2.64 \quad \Rightarrow \quad u_{1-\alpha/2} = 1.65$$

即得

$$\Phi(1.65) = 0.9505 = 1 - \frac{\alpha}{2} \quad \Rightarrow \quad \alpha = 0.10$$

置信水平 $1 - \alpha = 0.90$

9.2 **C**

由题意得

$$E = u_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = u_{0.975} \cdot \frac{4}{\sqrt{n}} < 2$$

即得

$$\sqrt{n} \geq 2u_{0.975} \quad \Rightarrow \quad n \geq 16$$

9.3 **B**

由题意得

$$E = u_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = u_{0.975} \frac{0.3}{\sqrt{15}}$$

置信区间为

$$I = [\bar{x} - E, \bar{x} + E] = [2.88 - E, 2.88 + E]$$

9.4 **B**

请读者验证

10 区间估计第二次作业

Formula Or Theorem:

- **Small sample confidence interval for μ of normal population(σ unknown)**

$$\bar{x} - t_{1-\alpha/2} \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{1-\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

- **Confidence interval of σ^2 when μ unknown(P305)**

$$\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\alpha/2}}$$

10.1 **D**

σ 未知，考虑使用 $\sqrt{n}(\bar{x}-\mu)/s$ 作为枢轴量，μ 的置信区间为

$$[\bar{x} - t_{1-\alpha/2} \cdot \frac{s}{\sqrt{n}}, \quad \bar{x} + t_{1-\alpha/2} \cdot \frac{s}{\sqrt{n}}]$$

所以 e^{μ} 的置信区间为

$$\left[e^{\bar{x} - t_{1-\alpha/2} \cdot \frac{s}{\sqrt{n}}}, \quad e^{\bar{x} + t_{1-\alpha/2} \cdot \frac{s}{\sqrt{n}}}\right]$$

10.2 **A**

已知 $lnX \sim N(\mu, 1)$，**断言 (*)**$E(X) = e^{\mu+\frac{1}{2}}$，所以可以考虑 μ 的置信区间 (σ 已知)

记 $Y = lnX \sim N(\mu, 1)$，则

$$P\left(\left|\frac{\bar{y}-\mu}{\sigma/\sqrt{n}}\right| < u_{1-\alpha/2}\right) = 1-\alpha$$

即 μ 的置信区间为

$$\bar{y} - u_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{y} + u_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

其中 $\bar{y} = 0$，$u_{0.975} = 1.96$，$\sigma = 1$，$n = 4$，则

$$-0.98 < \mu < 0.98$$

所以

$$e^{-0.48} < E(X) < e^{1.48}$$

证明 (*) 处的一个结论 (P115 习题 2.6.13) 设随机变量 $X \sim N(\mu, \sigma^2)$，求 $Y = e^X$ 的数学期望与方差

求出 Y 的分布 (Y > 0)

$$F_Y(y) = P(Y \le y) = P(e^X \le y) = P(X \le lny) = F_X(lny)$$

所以

$$p_y(y) = \frac{d}{dy}F_Y(y) = \frac{1}{y} \cdot \frac{d}{dy}F_X(lny) = \frac{1}{\sqrt{2\pi}\sigma} = \frac{1}{\sqrt{2\pi}\sigma y}e^{-\frac{(lny-\mu)^2}{2\sigma^2}} \qquad (y > 0)$$

求 $E(Y)$

$$E(Y)=\frac{1}{\sqrt{2\pi}\sigma}\int_0^{+\infty}e^{-\frac{(lny-\mu)^2}{2\sigma^2}}dy$$

令 $(lny-\mu)/\sigma=t$，得到

$$\begin{aligned}E(Y)&=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}\cdot e^{\sigma t+\mu}\sigma dt=\frac{e^{\mu}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{1}{2}(t^2-2\sigma t)}dt\\&=e^{\mu+\frac{1}{2}\sigma^2}\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{(t-\sigma)^2}{2}}dt=e^{\mu+\frac{1}{2}\sigma^2}\end{aligned}$$

求 $Var(X)$

$$\begin{aligned}E(Y^2)&=\frac{1}{\sqrt{2\pi}\sigma}\int_0^{+\infty}ye^{-\frac{(lny-\mu)^2}{2\sigma^2}}dy=e^{2\mu+2\sigma^2}\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{(t-2\sigma)^2}{2}}dt\\&=e^{2\mu+2\sigma^2}\end{aligned}$$

$$Var(Y)=E(Y^2)-E(Y)^2=e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$$

题目中 $Y=lnX\sim N(\mu,1)$，所以 $X=e^Y$，则 $E(X)=e^{\mu+\frac{1}{2}}$

10.3 **C**

显然，读者自证

10.4 **B**

μ 未知，考虑使用 $(n-1)s/\sigma^2$ 作为枢轴量，σ^2 的置信区间为

$$\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}<\sigma^2<\frac{(n-1)s^2}{\chi^2_{\alpha/2}}$$

带入 $n=9$，$s^2=11$，$\alpha=0.5$ 即可

11 单侧置信区间作业

11.1 **A**

选定枢轴量

$$\frac{\bar{x}-\mu}{s/\sqrt{n}}\sim t(n-1)$$

则

$$P\left(\frac{\bar{x}-\mu}{s/\sqrt{n}}\le t_{1-\alpha}(n-1)\right)=1-\alpha$$

即

$$P\left(\mu\ge\bar{x}-t_{1-\alpha}(n-1)\frac{s}{\sqrt{n}}\right)=1-\alpha$$

所以 μ 的置信下限为 $\bar{x}-t_{1-\alpha}(n-1)\cdot s/\sqrt{n}$

11.2 **D**

选定枢轴量

$$\frac{\bar{x}-\mu}{s/\sqrt{n}}\sim t(n-1)$$

则

$$P\left(\frac{\bar{x}-\mu}{s/\sqrt{n}}\geq t_{\alpha}(n-1)\right)=1-\alpha$$

即

$$P\left(\mu\leq\bar{x}-t_{\alpha}(n-1)\frac{s}{\sqrt{n}}\right)=1-\alpha$$

所以 μ 的置信上限为 $\bar{x}-t_{\alpha}(n-1)\cdot s/\sqrt{n}$

注：$t_{\alpha}=-t_{1-\alpha}$　　读者自证不难

11.3 **D**

易知

$$\bar{x}_1-\bar{x}_2\sim N\left(\mu_1-\mu_2,\frac{\sigma_1^2+\sigma_2^2}{n}\right)$$

则选定枢轴量为

$$\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{(\sigma_1^2+\sigma_2^2)/n}}\sim N(0,1)$$

则有

$$P\left(\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{(\sigma_1^2+\sigma_2^2)/n}}\geq u_{\alpha}\right)=1-\alpha$$

即

$$P\left(\mu_1-\mu_2\leq(\bar{x}_1-\bar{x}_2)-u_{\alpha}\cdot\sqrt{\frac{\sigma_1^2+\sigma_2^2}{n}}\right)=1-\alpha$$

所以 $\mu_1-\mu_2$ 的置信上限为 $(\bar{x}_1-\bar{x}_2)-u_{\alpha}\cdot\sqrt{(\sigma_1^2+\sigma_2^2)/n}$

注：$u_{\alpha}=-u_{1-\alpha}$　　读者自证不难

11.4 **B**

显然 $\bar{x}\pm u_{1-\alpha/2}\cdot\sigma/\sqrt{n}$

12 假设检验的思想及单正态总体检验

Formula Or Theorem:

- **假设检验的基本步骤 (P315)**
  (1) 建立假设 (H_0 : vs H_1 :)
  (2) 选择检验统计量 (t or u)
  (3) 选择显著性水平 $\alpha =$
  (4) 给出拒绝域 $W =$
  (5) 做出判断 (reject or not reject)

- **some definitions**
  H_0: Null hypothesis
  H_1: Alternative hypothesis
  Type I error: Rejection of H_0 when H_0 is true
  Type II error: Nonrejection of H_0 when H_1 is true
  α: probability of making a Type I error(also called the level of significance)
  β: probability of making a Type II error

12.1 **C**

definition of Type II error

12.2 **D**

1. Null hypothesis: $\sigma \le \sigma_0 = 0.50$
Alternative hypothesis: $\sigma > \sigma_0 = 0.50$
2. Level of significance: $\alpha = 0.05$
3. Criterion:
test statistic:

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2(n-1)$$

rejection regions:

$$W = \{\chi^2 > \chi^2_{1-\alpha}(n-1)\}$$

4. Calculations:

$$\chi^2 = \frac{(25-1)\cdot 0.58^2}{0.50^2} = 32.294 < \chi^2_{0.95}(24) = 36.415$$

5. Decision:
$\chi^2 \notin W$, the null hypothesis cannot be rejected

12.3 **B**

test statistic:

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2(n-1)$$

12.4 **D**

rejection regions:

$$W = \left\{ \left| \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \right| > t_{1-\alpha/2}(n-1) \right\}$$

12.5 **D**

test statistic:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

and

$$s = \sqrt{\frac{Q}{n-1}} \quad , \quad \mu = 4$$

substitute

$$t = \sqrt{n(n-1)} \cdot \frac{\bar{x} - 4}{\sqrt{Q}}$$

12.6 **B**

1. Null hypothesis: $\mu \geq 1000$

Alternative hypothesis: $\mu < 1000$

2. Level of significance: $\alpha = 0.05$

3. Criterion:

test statistic:

$$u = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

rejection regions:

$$W = \{u < u_\alpha\}$$

4. Calculations:

$$u = \frac{950 - 1000}{100/\sqrt{25}} = -2.5 < u_{0.05} = -1.6$$

5. Decision:

$u \in W$, the null hypothesis must be rejected

12.7 **B**

1. Null hypothesis: $\mu = 2.64$

Alternative hypothesis: $\mu \neq 2.64$

2. Level of significance: $\alpha = 0.01$

3. Criterion:

test statistic:

$$u = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

rejection regions:

$$W = \{u < u_{\alpha/2} \quad or \quad u > u_{1-\alpha/2}\}$$

4. Calculations:

$$u = \frac{2.61 - 2.64}{0.06/\sqrt{36}} = -3 < u_{0.005} = -2.58$$

5. Decision:

$u \in W$, the null hypothesis must be rejected

12.8 **C**

I don't know !

12.9 **A**

obviously

12.10 **A**

obviously

12.11 **C**

检验是否有显著增加，所以是单边假设

12.12 **C**

1. 建立假设

$$H_0: \quad \mu \leq 0.8 \quad vs \quad H_1: \quad \mu > 0.8$$

2. 检验统计量

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t(n-1)$$

3. 显著性水平

$$\alpha = 0.05$$

4. 拒绝域

$$W = \{t > t_{1-\alpha}(n-1) = t_{0.95}(15)\}$$

5. 做出判断因为

$$t = \frac{0.92 - 0.8}{0.32/\sqrt{16}} = 1.5 < t_{0.95}(15) = 1.753$$

即 $t \notin W$，所以不能够拒绝原假设，也就是说厂方断言正确

13 双正态总体分布的假设检验作业

13.1 **D**

calculate s^2

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n}x_i^2+\sum_{i=1}^{n}\bar{x}^2-2\bar{x}\sum_{i=1}^{n}x_i\right)$$
$$= \frac{1}{n-1}\left(\sum_{i=1}^{n}x_i^2-n\bar{x}\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n}x_i^2-\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2\right)$$

calculate s_X and s_Y

$$s_X^2 = \frac{1}{5}\left(6978.93-\frac{1}{6}\cdot(204.6)^2\right) = 0.414$$

$$s_Y^2 = \frac{1}{8}\left(15280.173-\frac{1}{9}\cdot(370.8)^2\right) = 0.402$$

1. Null hypothesis: $\sigma_X^2 = \sigma_Y^2$

Alternative: $\sigma_X^2 \neq \sigma_Y^2$

2. Level of significance: $\alpha = 0.05$

3. Criterion:

test statistic:

$$F = \frac{s_X^2}{s_Y^2} \sim F(5,8)$$

rejection regions:

$$W = \{F < F_{\alpha/2}(5,8) \quad or \quad F > F_{1-\alpha/2}(5,8)\}$$

Calculations:

$$F = \frac{s_X^2}{s_Y^2} = 1.030 \in (0.207, 6.67)$$

Decision:

$F \notin W$, the null hypothesis cannot be rejected

13.2 **A**

obviously

13.3 **B**

1. Null hypothesis: $\mu_1 \geq \mu_2$

Alternative hypothesis: $\mu_1 < \mu_2$

2. Level of significance: $\alpha = 0.01$

3. Criterion:

test statistic:

$$t = \frac{\bar{x}-\bar{y}}{\sqrt{\frac{s_1^2+s_2^2}{n}}} \sim t(2n-2)$$

rejection regions:

$$W = \{t < t_\alpha(2n-2)\}$$

4. Calculations:

$$t = -4.6468 \text{ (by Excel) } < -2.5524$$

5. Decision:
$t \in W$, the null hypothesis must be rejected

13.4 **A**

I test $\sigma_1^2 = \sigma_2^2$
1. Null hypothesis: $\sigma_1^2 = \sigma_2^2$
Alternative hypothesis: $\sigma_1^2 \neq \sigma_2^2$
2. Level of significance: $\alpha = 0.05$
3. Criterion:
test statistic:

$$F = \frac{s_1^2}{s_2^2} \sim F(109, 99)$$

rejection regions:

$$W = \{F < F_{\alpha/2} \quad or \quad F > F_{1-\alpha/2}\}$$

4. Calculations:

$$F = \frac{120.96^2}{105.53^2} = 1.314 < 1.475$$

5. Decision:
$F \notin W$, the null hypothesis cannot be rejected

II test $\mu_1 = \mu_2$
1. Null hypothesis: $\mu_1 = \mu_2$
Althernative hypothesis: $\mu_1 \neq \mu_2$
2. Level of significance: $\alpha = 0.05$
3. Criterion:
test statistic:

$$u = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}} \dot\sim N(0,1)$$

note: because of large sample
rejection regions:

$$W = \{|u| > u_{1-\alpha/2}\}$$

4. Calculations:

$$u = \frac{2805 - 2680}{\sqrt{\frac{120.96^2}{110} + \frac{105.53^2}{100}}} = 7.996 > 1.96$$

5. Decision:
$u \in W$, the null hypothesis must be rejected

13.5 **B**

1. 建立假设

$$H_0: \quad \sigma_1^2 \leq \sigma_2^2 \quad vs \quad H_1: \quad \sigma_1^2 > \sigma_2^2$$

2. 检验统计量

$$F = \frac{s_1^2}{s_2^2} \sim F(9, 11)$$

3. 显著性水平

$$\alpha = 0.005$$

4. 拒绝域

$$W = \{F > F_{0.995}(9, 11)\}$$

5. 做出判断，因为

$$F = \frac{50^2}{20^2} = 6.25 > 5.536$$

即 $F \in W$，所以拒绝原假设，也就是说新生产工艺下灯管寿命的稳定性显著提高

14 非参数假设检验

14.1 B

obviously

14.2 D

公司	A	B	others
理论频率	45%	40%	15%
实测频数	102	82	16

1. 提出假设

$$H_0: \quad \pi_A = 0.45 \land \pi_B = 0.40 \land \pi_O = 0.15$$
$$H_1: \quad \pi_A \neq 0.45 \lor \pi_B \neq 0.40 \lor \pi_O \neq 0.15$$

注：仅记 A 公司的市场占有率为 π_A

2. 显著性水平

$$\alpha = 0.05$$

3. 标准

检验统计量

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - np_i)^2}{np_i} \sim \chi^2(2)$$

拒绝域

$$W = \{\chi^2 > \chi_{0.95}(2)\}$$

4. 计算

$$\chi^2 = \sum_{i=1}^{3} \frac{(f_i - np_i)^2}{np_i} = 8.183 > 5.991$$

5. 决定

$\chi^2 \in W$，拒绝原假设，广告战后各公司市场占有率有显著变化

14.3 B

1. 提出假设

$$H_0: \quad \text{骰子是均匀的} \quad vs \quad H_1: \quad \text{骰子是不均匀的}$$

2. 显著性水平

$$\alpha = 0.05$$

3. 标准

检验统计量

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - np_i)^2}{np_i} \sim \chi^2(5)$$

拒绝域

$$W = \{\chi^2 > \chi_{0.95}(5)\}$$

4. 计算

$$\chi^2 = \sum_{i=1}^{6} \frac{(f_i - np_i)^2}{np_i} = 18.7 > 11.07$$

注：$np_i = 120 \times 1/6 = 20$

5. 决定

$\chi^2 \in W$，拒绝原假设，骰子是不均匀的

15 附录

I. Γ 函数和 B 函数 (P102 - P105)

$$\Gamma(\alpha)=\int_0^{+\infty}x^{\alpha-1}e^{-x}dx$$

$$B(a,b)=\int_0^1 x^{a-1}(1-x)^{b-1}dx=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

II. 区间估计

<table>
<tr><td rowspan="2">点估计</td><td rowspan="2">总体分布</td><td rowspan="2">样本容量</td><td colspan="2">σ 未知</td><td colspan="2">σ 已知</td></tr>
<tr><td>统计量</td><td>区间</td><td>统计量</td><td>区间</td></tr>
<tr><td rowspan="4">总体均值 μ</td><td rowspan="2">正态分布</td><td>小样本 $(n<30)$</td><td>$t=\frac{\bar{x}-\mu}{s/\sqrt{n}}$</td><td>$\bar{x}\pm t_{\alpha/2}\cdot\frac{s}{\sqrt{n}}$</td><td rowspan="3">$z=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$</td><td rowspan="3">$\bar{x}\pm z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}$</td></tr>
<tr><td>大样本 $(n\geq 30)$</td><td rowspan="2">$z=\frac{\bar{x}-\mu}{s/\sqrt{n}}$</td><td rowspan="2">$\bar{x}\pm z_{\alpha/2}\cdot\frac{s}{\sqrt{n}}$</td></tr>
<tr><td rowspan="2">非正态分布</td><td>大样本 $(n\geq 30)$</td></tr>
<tr><td>小样本 $(n<30)$</td><td colspan="4">具体分布未知</td></tr>
<tr><td rowspan="2">总体方差 σ^2</td><td rowspan="2">正态分布</td><td rowspan="2">无要求</td><td>$\chi^2=\frac{(n-1)s^2}{\sigma^2}$</td><td colspan="3">$\left[\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)},\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}\right]$ (μ 未知)</td></tr>
<tr><td>$z=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$</td><td colspan="3">$\sigma>\frac{(\bar{x}-\mu)\sqrt{n}}{z_{\alpha/2}}$ (μ 已知)</td></tr>
</table>

III. F、t、u 分布分位数

$$F_\alpha(n,m)=\frac{1}{F_{1-\alpha}(m,n)}$$

$$u_{1-\alpha}=-u_\alpha$$

$$t_{1-\alpha}=-t_\alpha$$

IV. 勘误

P6 2.5

$$\Phi\left(\frac{1000-10n}{\sqrt{n}}\right)\geq 0.99$$

查表得

$$\Phi(2.33)=0.9901$$

那么

$$\frac{1000-10n}{\sqrt{n}}\geq 2.33\quad\Rightarrow\quad n=98$$