概率论 第6讲 · 数理统计 · 解析

> **用法**：先做完 02-试题 再对答案。每题给「**答案 / 思路 / 易错点**」，标「真题骨架」的题注明对应真题。

A 组 · 基础

第 1 题 · 统计量与常用统计量

（Ⅰ）答案：$\bar{X}=5$，$S^2=7.5$，$B_2=6$。

思路：

$$\bar{X}=\frac{2+3+5+6+9}{5}=\frac{25}{5}=5.$$

偏差平方和：$(2-5)^2+(3-5)^2+0^2+(6-5)^2+(9-5)^2=9+4+0+1+16=30$。

$$S^2=\frac{30}{5-1}=7.5,\qquad B_2=\frac{30}{5}=6.$$

（Ⅱ）答案：是统计量的有 $\bar{X}+\mu$ 中的 S^2 与 $\max\{X_1,\cdots,X_5\}$ —— 即 S^2、$\max$ **是统计量**；其余三个（$\bar{X}+\mu$、$X_1+X_2-\mu$、$\sum(X_i-\mu)^2$）**不是**。

易错点：

- **判断标准只有一个**：表达式里**有没有未知参数**。μ、σ 均未知 $\Longrightarrow$ 凡是出现 μ 的都**不是**统计量。
- S^2 虽然含 $\bar{X}$，但 $\bar{X}$ 由样本算出，**不含未知参数**，所以是统计量。
- 别把 $\sum(X_i-\bar{X})^2$ 当作"能不能算"的判断对象——它是统计量；而 $\sum(X_i-\mu)^2$ 不是。

第 2 题 · 三大分布辨析与构造

（Ⅰ）答案：$Y_1\sim\chi^2(6)$。

思路：$X_i\sim N(0,4)\Longrightarrow\dfrac{X_i}{2}\sim N(0,1)\Longrightarrow\left(\dfrac{X_i}{2}\right)^2=\dfrac{X_i^2}{4}\sim\chi^2(1)$。六个独立相加：

$$Y_1=\sum_{i=1}^{6}\left(\frac{X_i}{2}\right)^2\sim\chi^2(6).$$

（Ⅱ）答案：$Y_2\sim\chi^2(1)$。

思路：$X_1+X_2+X_3\sim N(0,\,4+4+4)=N(0,12)$，标准化后平方：

$$Y_2=\frac{(X_1+X_2+X_3)^2}{12}=\left(\frac{X_1+X_2+X_3}{\sqrt{12}}\right)^2\sim\chi^2(1).$$

（Ⅲ）答案：$Y_3\sim t(3)$。

思路：**拆成"标准正态 ÷ 根号下 χ^2 除以自由度"**。

- 分子：$\dfrac{X_1}{2}\sim N(0,1)$；
- 分母：$\dfrac{X_2^2+X_3^2+X_4^2}{4}\sim\chi^2(3)$；

$$\frac{X_1/2}{\sqrt{\dfrac{1}{3}\cdot\dfrac{X_2^2+X_3^2+X_4^2}{4}}}=\frac{X_1/2}{\sqrt{X_2^2+X_3^2+X_4^2/\sqrt{12}}}=\frac{\sqrt{3}\,X_1}{\sqrt{X_2^2+X_3^2+X_4^2}}\sim t(3).$$

（Ⅳ）答案：**不服从三大分布。**

思路：$X_1-X_2\sim N(0,8)\Longrightarrow\dfrac{(X_1-X_2)^2}{8}\sim\chi^2(1)$，但

$$2(X_1-X_2)^2=16\cdot\frac{(X_1-X_2)^2}{8}=16\chi^2(1).$$

χ^2 **分布要求"平方项之和，每项都是标准正态的平方"**，也就是**系数必须恰好把方差消掉。**这里系数 2 只消掉了一半（正确系数是 $\dfrac{1}{8}$）。

易错点：

- ⚠️ **系数错一点就不是 χ^2。**判断口诀：**先看它能不能写成 $\sum\left(\dfrac{\cdot}{\sigma}\right)^2$ 的形式。**
- **（Ⅳ）就是 🟧2017•8 的 B 选项**（真题是 $2(X_n-X_1)^2$，同样不服从）。
- ⚠️ **（Ⅲ）里分子分母共用 X 吗？** 本题分子是 X_1、分母是 X_2,X_3,X_4，**没有共用**，所以**独立**，t 分布成立。若分母含 X_1，就不独立，不能用。

第 3 题 • 样本均值与样本方差的数字特征

（Ⅰ）答案：$E\bar{X}=\mu$，$D\bar{X}=\dfrac{\sigma^2}{n}$，$ES^2=\sigma^2$。

思路：

$$E\bar{X}=\frac{1}{n}\sum EX_i=\frac{1}{n}\cdot n\mu=\mu;$$

$$D\bar{X}=\frac{1}{n^2}\sum DX_i=\frac{1}{n^2}\cdot n\sigma^2=\frac{\sigma^2}{n}\quad(\text{独立才可加});$$

ES^2 的关键是**先化简再取期望**：

$$\sum_{i=1}^{n}(X_i-\bar{X})^2=\sum X_i^2-n\bar{X}^2,$$

$$E\left[\sum X_i^2\right]=n(\sigma^2+\mu^2),\qquad E\bar{X}^2=D\bar{X}+(E\bar{X})^2=\frac{\sigma^2}{n}+\mu^2,$$

$$E\left[\sum(X_i-\bar{X})^2\right]=n(\sigma^2+\mu^2)-n\left(\frac{\sigma^2}{n}+\mu^2\right)=(n-1)\sigma^2\Longrightarrow ES^2=\sigma^2.$$

（Ⅱ）答案：因为 $\dfrac{1}{n}\sum(X_i-\bar{X})^2$ 是**有偏**的。

思路：由上一步 $E\left[\frac{1}{n}\sum(X_i-\bar{X})^2\right]=\frac{n-1}{n}\sigma^2\neq\sigma^2$。原因是用 $\bar{X}$ 代替 μ 后，"独立信息"损失了 1 份（自由度 $n\to n-1$）。除以 $n-1$ 恰好把 $\frac{n-1}{n}$ 补回 1。

易错点：

- $D\bar{X}$ 是 $\frac{\sigma^2}{\boldsymbol{n}}$，**不是** σ^2——这是第 5 讲切比雪夫不等式反复用到的结论。
- $\sum(X_i-\bar{X})^2$ 的两个恒等式：$=\sum X_i^2-n\bar{X}^2$（求 E 时用）。
- **"$n-1$ 是为了无偏"**这句话要能自己推（面试式追问）。

第 4 题 • 矩估计

答案：$\hat{\theta}=\frac{1-\bar{X}}{5}$。

思路：先求**总体一阶矩**（含 θ）：

$$EX=(-1)\cdot 2\theta+0\cdot\theta+1\cdot(1-3\theta)=-2\theta+1-3\theta=1-5\theta.$$

用样本一阶矩替换：$1-5\theta=\bar{X}$，解得

$$\hat{\theta}=\frac{1-\bar{X}}{5}.$$

易错点：

- 先**验证概率和为 1**：$2\theta+\theta+(1-3\theta)=1$ ✓（题目给了 $0<\theta<1/3$ 保证非负）。
- **只用一个参数 → 只用一阶矩**，别多此一举用二阶矩。
- 最后写成 $\hat{\theta}$ 是**估计量**（含大写 X_i），$\hat{\theta}(x_1,\cdots,x_n)$ 才是估计值。

第 5 题 • 极大似然估计

答案：$\hat{\theta}=-\frac{n}{\sum\limits_{i=1}^{n}\ln X_i}$。

思路：

$$L(\theta)=\prod_{i=1}^{n}\theta x_i^{\theta-1}=\theta^n\Big(\prod x_i\Big)^{\theta-1},\qquad \ln L(\theta)=n\ln\theta+(\theta-1)\sum_{i=1}^{n}\ln x_i.$$

$$\frac{\mathrm{d}\ln L}{\mathrm{d}\theta}=\frac{n}{\theta}+\sum_{i=1}^{n}\ln x_i=0\implies\hat{\theta}=-\frac{n}{\sum\ln X_i}.$$

易错点：

- ⚠️ $0<x_i<1\Longrightarrow\ln x_i<0\Longrightarrow\hat{\theta}>0$ ✓，解是合法的。可以顺手做这个**正性检验**。
- 支撑集 $0<x<1$ **不含** θ，所以属于**可微情形**，直接求导即可（对比第 8、17 题的"取边界"）。

- 另一种常考载体 $f(x;\theta)=(\theta+1)x^{\theta}\ (0<x<1,\theta>-1)$ 得 $\hat\theta=-\dfrac{n}{\sum\ln X_i}-1$——**结构相同，常数差 1**，注意区分。

第 6 题 • 无偏性定常数 🟩

答案：$k=\dfrac{1}{2(n-1)}$。

思路：关键是算 $E(X_{i+1}-X_i)^2$。用"**方差 + 均值差平方**"：

$$E(X_{i+1}-X_i)^2=D(X_{i+1}-X_i)+\left[E(X_{i+1}-X_i)\right]^2=\left(\sigma^2+\sigma^2\right)+0=2\sigma^2.$$

（X_{i+1} 与 X_i 独立，方差可加；同分布 $\Longrightarrow$ 均值之差为 0。）

于是

$$E\Big[k\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2\Big]=k\cdot(n-1)\cdot 2\sigma^2=2k(n-1)\sigma^2\overset{!}{=}\sigma^2\ \Longrightarrow\ k=\frac{1}{2(n-1)}.$$

易错点：

- ⚠️ 不能只写 $E(X_{i+1}-X_i)^2=\sigma^2$——**漏掉一项**：两个独立同分布变量**差的方差是** $2\sigma^2$。
- 共有 $n-1$ 对（相邻对），**不是** n **对**，求和上限别写错。

B 组 • 中等

第 7 题 • 拉普拉斯分布 • MLE + 数字特征　🟧 真题骨架 2018•23

（Ⅰ）答案：$\hat\lambda=\dfrac{1}{n}\sum_{i=1}^{n}|X_i|$。

思路：

$$L(\lambda)=\prod_{i=1}^{n}\frac{1}{2\lambda}\mathrm{e}^{-\frac{|x_i|}{\lambda}}=\Big(\frac{1}{2\lambda}\Big)^n\exp\Big(-\frac{1}{\lambda}\sum_{i=1}^{n}|x_i|\Big),$$

$$\ln L(\lambda)=-n\ln 2-n\ln\lambda-\frac{1}{\lambda}\sum_{i=1}^{n}|x_i|,\qquad \frac{\mathrm{d}\ln L}{\mathrm{d}\lambda}=-\frac{n}{\lambda}+\frac{1}{\lambda^2}\sum_{i=1}^{n}|x_i|=0.$$

解得 $\hat\lambda=\dfrac{1}{n}\sum_{i=1}^{n}|X_i|$。

（Ⅱ）答案：$E\hat\lambda=\lambda$，$D\hat\lambda=\dfrac{\lambda^2}{n}$。

思路：**先认清** $|X_i|$ **的分布**——这是全题的关键。$X\sim$ 拉普拉斯$(0,\lambda)$ 且密度只依赖于 $|x|$，由对称性可得

$$|X_i|\ \sim\ \mathrm{Exp}\Big(\frac{1}{\lambda}\Big)\quad(\text{密度 }\tfrac{1}{\lambda}\mathrm{e}^{-x/\lambda},\ x>0),$$

即**参数为** $1/\lambda$ **的指数分布**，故 $E|X_i|=\lambda$、$D|X_i|=\lambda^2$。于是

$$E\hat{\lambda} = \frac{1}{n}\cdot n\lambda = \lambda \quad (\text{无偏}), \qquad D\hat{\lambda} = \frac{1}{n^2}\cdot n\lambda^2 = \frac{\lambda^2}{n}.$$

易错点：

- ⚠️ **不要直接去算 $\int x\cdot\frac{1}{2\lambda}e^{-|x|/\lambda}\mathrm{d}x$ 的期望**——那样得到的是 $EX=0$，与本题无关。要求的是**绝对值的期望**。
- 因为 $\hat{\lambda}$ 无偏，$E\hat{\lambda}$ 直接可用；若要算 $D\hat{\lambda}$，**记得乘 $\frac{1}{n^2}$**。
- 本题与 **2018•23 结构完全一致**（真题用 $f=\frac{1}{2\sigma}e^{-|x|/\sigma}$），可直接回刷真题。

第 8 题 • 次序统计量 • 最大值分布 + 无偏　🟧真题骨架 2016•23

（Ⅰ）答案：$f_T(x)=\dfrac{6x^5}{\theta^6}$，$0<x<\theta$（其他为 0）。

思路：先求**分布函数**。

$$F(x)=\int_0^x\frac{2t}{\theta^2}\mathrm{d}t=\frac{x^2}{\theta^2}\quad(0<x<\theta).$$

由最大值公式 $F_T(x)=[F(x)]^3=\dfrac{x^6}{\theta^6}$，求导：

$$f_T(x)=\frac{\mathrm{d}}{\mathrm{d}x}\frac{x^6}{\theta^6}=\frac{6x^5}{\theta^6},\qquad 0<x<\theta.$$

（Ⅱ）答案：$a=\dfrac{7}{6}$。

思路：

$$ET=\int_0^\theta x\cdot\frac{6x^5}{\theta^6}\mathrm{d}x=\frac{6}{\theta^6}\cdot\frac{\theta^7}{7}=\frac{6}{7}\theta.$$

令 $E(aT)=a\cdot\dfrac{6}{7}\theta=\theta$，得 $a=\dfrac{7}{6}$。

易错点：

- ⚠️ **不要漏掉 $[F(x)]^{n-1}$ 这个因子**：$f_{(n)}=n[F]^{n-1}f$，本题 $n=3$，$f=2x/\theta^2$，$[F]^2=x^4/\theta^4$，乘起来 $3\cdot\frac{x^4}{\theta^4}\cdot\frac{2x}{\theta^2}=\frac{6x^5}{\theta^6}$ ✓。
- 若总体的 $F(x)=(x/\theta)^k$，则 $T=\max$ 的期望是 $\dfrac{kn}{kn+1}\theta$。本题 $k=2,n=3\Longrightarrow\dfrac{6}{7}\theta$ ✓（**用这个速算公式可直接验算**）。

第 9 题 • 区间估计　🟧真题骨架 2016•14

（Ⅰ）答案：双侧区间 $\left[11.32,\ 13.28\right]$；单侧置信上限 13.1225。

思路：$\sigma=2$ 已知 $\Longrightarrow$ 用 U 检验的枢轴量，且 $\dfrac{\sigma}{\sqrt{n}}=\dfrac{2}{4}=0.5$。

双侧：半宽 $=u_{\alpha/2}\cdot\dfrac{\sigma}{\sqrt{n}}=1.96\times0.5=0.98$。

$$\left[12.3-0.98,\ 12.3+0.98\right]=[11.32,\ 13.28].$$

单侧上限：**改用 u_α（不是 $u_{\alpha/2}$）**，半宽 $= 1.645 \times 0.5 = 0.8225$，

$$\bar{\mu} = 12.3 + 0.8225 = 13.1225.$$

（II）答案：$\left[11.2345,\ 13.3655\right]$。

思路：σ **未知** $\Longrightarrow$ 用 S 代替 σ，枢轴量变 $t(n-1)$：

$$t_{0.025}(15) \cdot \frac{S}{\sqrt{n}} = 2.131 \times 0.5 = 1.0655,$$

$$\left[12.3 - 1.0655,\ 12.3 + 1.0655\right] = [11.2345,\ 13.3655].$$

易错点：

- ⚠️ **双侧 vs 单侧的分位数不同**：双侧用 $u_{\alpha/2}$（本例 1.96），单侧用 u_α（1.645）。**单侧区间更短**是必然的（全部 α 集中一侧）。
- ⚠️ σ **已知用 U、未知用 t**——审题第一件事就是看 σ 给没给。
- 💡 **对称性技巧（真题 2016•14 的考点）**：均值的双侧区间关于 $\bar{X}$ 对称，所以

$$\text{下限} + \text{上限} = 2\bar{X}.$$

  真题只给了 $\bar{X} = 9.5$ 和上限 10.8，让你反求下限：$2 \times 9.5 - 10.8 = 8.2$，答案 $[8.2, 10.8]$。**区间估计里唯一"送分"的公式，务必记住。**

第 10 题 • 假设检验拒绝域　🟧真题骨架 2025•10

（I）答案：$\bar{X} \geq 2 + Z_\alpha\sqrt{\dfrac{3}{n}}$。

思路：$\sigma^2 = 3$ 已知 $\Longrightarrow$ 取检验统计量

$$U = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{X} - 2}{\sqrt{3}/\sqrt{n}} \sim N(0,1) \quad (H_0 \text{ 边界处}).$$

$H_1 : \mu > 2 \Longrightarrow$ **右尾检验** $\Longrightarrow$ 拒绝域为 $U \geq Z_\alpha$，即

$$\frac{\bar{X} - 2}{\sqrt{3/n}} \geq Z_\alpha \iff \bar{X} \geq 2 + Z_\alpha\sqrt{\frac{3}{n}}.$$

（II）答案：$c = 2.8225$。

思路：$n = 12$ 时 $\sqrt{\dfrac{3}{12}} = \sqrt{0.25} = 0.5$，

$$c = 2 + 1.645 \times 0.5 = 2 + 0.8225 = 2.8225.$$

易错点：

- ⚠️ $\sigma^2 = 3 \Longrightarrow \sigma = \sqrt{3}$。**不要把方差当标准差代进分母**——这是本题最大的坑。
- **拒绝域方向跟 H_1 走**：H_1 是"$>$"就放右尾（$\geq$）；H_0 里写 $\leq$ 不影响方向，只影响"边界取 μ_0"。
- **单侧用 Z_α**（不是 $Z_{\alpha/2}$）。若改成 $H_1 : \mu \neq 2$，则拒绝域为 $|U| \geq Z_{\alpha/2}$。

第 11 题 • 两类错误　🟧真题骨架 2021•10

（Ⅰ）答案：$\alpha = 1 - \Phi\left(\frac{5}{3}\right)$。

思路：$\bar{X} \sim N\left(\mu, \frac{9}{25}\right)$，即标准差 $\frac{3}{5} = 0.6$。第一类错误在 H_0 的**边界** $\mu = 5$ 处计算（这是 α 取到最大的地方）：

$$\alpha = P\{\bar{X} \geq 6 \mid \mu = 5\} = 1 - \Phi\left(\frac{6-5}{0.6}\right) = 1 - \Phi\left(\frac{5}{3}\right) \approx 1 - \Phi(1.667).$$

（Ⅱ）答案：$\beta = 1 - \Phi(1)$。

思路：第二类错误在 H_1 下的**具体真值** $\mu = 6.6$ 处计算，用**接受域**（W 的补集）：

$$\beta = P\{\bar{X} < 6 \mid \mu = 6.6\} = \Phi\left(\frac{6-6.6}{0.6}\right) = \Phi(-1) = 1 - \Phi(1).$$

易错点：

- ⚠️ α **与** β **用的分布不同**：α 用 $\mu = \mu_0$（边界），β 用 $\mu = \mu_1$（真值）。这是两类错误的**根本区别**。
- ⚠️ β **用"接受域"**：拒绝域是 $\bar{X} \geq 6$，所以 β 是 $\bar{X} < 6$ 的概率（方向别反）。
- 本题数字与 **2021•10 同构**（真题 $n = 16, \sigma^2 = 4, W = \{\bar{X} \geq 11\}, \mu = 11.5$，得 $\beta = 1 - \Phi(1)$），可直接回刷。

第 12 题 • 无偏性与有效性

（Ⅰ）答案：三者都是 μ 的无偏估计（系数和均为 1）。

思路：$E\hat{\theta} = \sum a_i EX_i = \mu \sum a_i$，**只要系数和** $\sum a_i = 1$ **就无偏**：

$$\frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1, \qquad \frac{1}{2} + \frac{1}{3} + \frac{1}{6} = \frac{3+2+1}{6} = 1, \qquad \frac{1}{2} + \frac{1}{2} = 1. \quad \checkmark$$

（Ⅱ）答案：$\hat{\theta}_1$ 最有效。

思路：三者都无偏，比较方差（独立性使方差可加）：

$$D\hat{\theta}_1 = \left(\tfrac{1}{9} + \tfrac{1}{9} + \tfrac{1}{9}\right)\sigma^2 = \frac{\sigma^2}{3} \approx 0.333\,\sigma^2;$$

$$D\hat{\theta}_2 = \left(\tfrac{1}{4} + \tfrac{1}{9} + \tfrac{1}{36}\right)\sigma^2 = \frac{9+4+1}{36}\sigma^2 = \frac{14}{36}\sigma^2 = \frac{7}{18}\sigma^2 \approx 0.389\,\sigma^2;$$

$$D\hat{\theta}_3 = \left(\tfrac{1}{4} + \tfrac{1}{4}\right)\sigma^2 = \frac{\sigma^2}{2} = 0.5\,\sigma^2.$$

$$\boxed{\hat{\theta}_1 \text{ 最有效}}$$

易错点：

- ⚠️ **必须先验证都无偏，才能比有效性**——这是"两级筛选"，直接比方差是逻辑错误。

- ⚠️ $\hat{\theta}_2$ 的方差要**逐项平方系数**再相加，别把 $\left(\frac{1}{2}+\frac{1}{3}+\frac{1}{6}\right)^2$ 当成答案。
- 直觉：**系数越"平均"（越接近 $\frac{1}{n}$），方差越小。**$\hat{\theta}_1$ 是全平均，$\hat{\theta}_2$ 向 X_1 偏重，$\hat{\theta}_3$ 干脆丢了一个样本，故 $\hat{\theta}_3$ 最差。

第 13 题 • 两样本泊松 • MLE　🟧真题骨架 2022•22

（Ⅰ）答案：$\hat{\lambda}=\dfrac{\sum x_i+\sum y_j}{n+2m}=\dfrac{39}{13}=3$。

思路：**两个样本要放在同一个似然函数里。**

$$L(\lambda)=\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}\mathrm{e}^{-\lambda}\cdot\prod_{j=1}^{m}\frac{(2\lambda)^{y_j}}{y_j!}\mathrm{e}^{-2\lambda},$$

$$\ln L(\lambda)=\left(\sum x_i\right)\ln\lambda-n\lambda+\left(\sum y_j\right)\ln(2\lambda)-2m\lambda+\text{常数}.$$

$$\frac{\mathrm{d}\ln L}{\mathrm{d}\lambda}=\frac{\sum x_i+\sum y_j}{\lambda}-(n+2m)=0\Longrightarrow\hat{\lambda}=\frac{\sum X_i+\sum Y_j}{n+2m}.$$

代入：$n=5$，$m=4$，$\sum x_i=3+2+4+1+5=15$，$\sum y_j=5+3+7+9=24$，

$$\hat{\lambda}=\frac{15+24}{5+2\times4}=\frac{39}{13}=3.$$

（Ⅱ）答案：$D(\hat{\lambda})=\dfrac{\lambda}{n+2m}=\dfrac{\lambda}{13}$。

思路：$\sum X_i\sim P(n\lambda)$ 方差 $n\lambda$；$\sum Y_j\sim P(2m\lambda)$ 方差 $2m\lambda$；两者独立故方差可加：

$$D\left(\sum X_i+\sum Y_j\right)=n\lambda+2m\lambda=\lambda(n+2m),$$

$$D\hat{\lambda}=\frac{\lambda(n+2m)}{(n+2m)^2}=\frac{\lambda}{n+2m}.$$

易错点：

- ⚠️ **不要把两组的似然分开求导**——它们是**同一个** λ 的两个来源，必须合并。
- ⚠️ Y **组的分母是** $n+2m$（不是 $n+m$）：因为 Y 的均值是 2λ，$2m$ 个"等效 λ 单位"。记忆：**"权重 = 样本量 × 该组的参数倍数"。**
- 若把 Y 的分布参数换成 $k\lambda$，则 $\hat{\lambda}=\dfrac{\sum X_i+\sum Y_j}{n+km}$，$D\hat{\lambda}=\dfrac{\lambda}{n+km}$——**这是本题的一般式。**

第 14 题 • 两正态总体方差比 • F 分布　🟧真题骨架 2023•9

答案：$c=3$，自由度为 $(m-1,\ n-1)$。

思路：**"拼 F"的两步——先各去方差，再除自由度。**

$$\frac{(m-1)S_1^2}{\sigma^2}\sim\chi^2(m-1),\qquad\frac{(n-1)S_2^2}{3\sigma^2}\sim\chi^2(n-1),$$

且两者独立。按 F 的定义（各自除以自由度后相除）：

$$F=\frac{\dfrac{S_1^2}{\sigma^2}}{\dfrac{S_2^2}{3\sigma^2}}=\frac{3S_1^2}{S_2^2}\ \sim\ F(m-1,\,n-1).$$

所以 $c=3$。

易错点：

- ⚠️ **两个总体的方差不同**（σ^2 与 $3\sigma^2$）时，"消方差"的系数就不同——这正是真题设陷阱的地方。若两方差相同，则 $c=1$。
- ⚠️ **自由度顺序不能反**：分子是 X 组的 $m-1$，分母是 Y 组的 $n-1$。
- 若题目问 $\dfrac{S_2^2}{S_1^2}\cdot c'$ 服从 $F(n-1,m-1)$，则由倒数性质 $c'=\dfrac{1}{3}$。

C 组 • 拔高

第 15 题 • 半正态分布 • 求常数 + MLE　🟧真题骨架 2019•23

（Ⅰ）答案：$A=\sqrt{\dfrac{2}{\pi}}$。

思路：**归一化条件** $\displaystyle\int_{-\infty}^{+\infty}f(x;\sigma)\mathrm{d}x=1$。令 $t=\dfrac{x-\mu}{\sigma}$（$\mathrm{d}x=\sigma\,\mathrm{d}t$）：

$$\int_{\mu}^{+\infty}\frac{A}{\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x=A\int_0^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t=A\cdot\sqrt{\frac{\pi}{2}}\overset{!}{=}1\ \Longrightarrow\ A=\sqrt{\frac{2}{\pi}}.$$

（Ⅱ）答案：$\hat{\sigma^2}=\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\mu)^2$。

思路：似然函数（注意支撑集 $x_i\geq\mu$）

$$L(\sigma)=\prod_{i=1}^{n}\frac{A}{\sigma}\mathrm{e}^{-\frac{(x_i-\mu)^2}{2\sigma^2}}=\left(\frac{A}{\sigma}\right)^n\exp\Big(-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2\Big)\prod_{i=1}^{n}\mathbf{1}_{\{x_i\geq\mu\}},$$

$$\ln L=-n\ln\sigma-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2+\text{常数}.$$

令 $v=\sigma^2$（对 σ 求导会出现平方根，**换元更省事**）：

$$\ln L=-\frac{n}{2}\ln v-\frac{1}{2v}\sum_{i=1}^{n}(x_i-\mu)^2,\qquad\frac{\mathrm{d}\ln L}{\mathrm{d}v}=-\frac{n}{2v}+\frac{1}{2v^2}\sum_{i=1}^{n}(x_i-\mu)^2=0.$$

解得 $v=\dfrac{1}{n}\sum(X_i-\mu)^2$，即 $\hat{\sigma^2}=\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\mu)^2$。

易错点：

- ⚠️ μ **已知**，所以支撑集 $x_i\geq\mu$ **不含未知参数**，不影响求导（若 μ 未知，就必须用 $\min X_i$ 约束，$\hat{\mu}=X_{(1)}$）。
- ⚠️ **对 σ^2 求导（而非对 σ）能避开根号运算——**这是本类题的一个通用技巧。

- 结果里是 n **而不是** $n-1$（MLE 一般有偏），与 S^2 的 $n-1$ 形成对比。

第 16 题 • 威布尔分布 • 无记忆性 + MLE　🟧真题骨架 2020•23

（Ⅰ）答案：$P\{T>t\}=\mathrm{e}^{-\left(\frac{t}{\theta}\right)^m}$；$P\{T>s+t\mid T>s\}=\exp\left\{-\dfrac{(s+t)^m-s^m}{\theta^m}\right\}$。

思路：由 $F(t)=1-\mathrm{e}^{-(t/\theta)^m}$ 直接得 $P\{T>t\}=1-F(t)=\mathrm{e}^{-(t/\theta)^m}$。再用条件概率：

$$P\{T>s+t\mid T>s\}=\frac{P\{T>s+t\}}{P\{T>s\}}=\frac{\mathrm{e}^{-\frac{(s+t)^m}{\theta^m}}}{\mathrm{e}^{-\frac{s^m}{\theta^m}}}=\exp\left\{-\frac{(s+t)^m-s^m}{\theta^m}\right\}.$$

（Ⅱ）答案：$\hat\theta=\left(\dfrac{1}{n}\sum_{i=1}^{n}t_i^m\right)^{1/m}$。

思路：先求密度（对 F 求导）：

$$f(t)=F'(t)=\frac{m}{\theta^m}t^{m-1}\mathrm{e}^{-t^m/\theta^m}\quad(t>0).$$

$$\ln L(\theta)=n\ln m-nm\ln\theta+(m-1)\sum\ln t_i-\frac{1}{\theta^m}\sum_{i=1}^{n}t_i^m,$$

$$\frac{\mathrm{d}\ln L}{\mathrm{d}\theta}=-\frac{nm}{\theta}+\frac{m}{\theta^{m+1}}\sum_{i=1}^{n}t_i^m=0\Longrightarrow\theta^m=\frac{1}{n}\sum_{i=1}^{n}t_i^m\Longrightarrow\hat\theta=\left(\frac{1}{n}\sum_{i=1}^{n}T_i^m\right)^{1/m}.$$

易错点：

- ⚠️ **别忘了"先由 F 求 f"这一步**——题目给的是分布函数，似然函数要用密度。
- ⚠️ **只当 $m=1$ 时才有"无记忆性"**（$P\{T>s+t\mid T>s\}=\mathrm{e}^{-t/\theta}=P\{T>t\}$，退化为指数分布）；一般 $m\neq1$ 时**不成立**——这句话是本题的"概念考点"。
- 对 θ 求导时 $-\dfrac{1}{\theta^m}\sum t_i^m$ 的导数是 $+\dfrac{m}{\theta^{m+1}}\sum t_i^m$，**符号与幂次都容易错。**
- 若题目改为"m 也未知"，则需**两个方程**联立（$\hat m$ 需要用 $\sum\ln t_i$ 的方程），近年未考。

第 17 题 • 无偏 + 均方误差最小　🟧真题骨架 2024•22

（Ⅰ）答案：$c=\dfrac{n+1}{n}$。

（Ⅱ）答案：$c=\dfrac{n+2}{n+1}$。

思路：先备好 $X_{(n)}$ 的两个矩。$X\sim U(0,\theta)\Longrightarrow F(x)=\dfrac{x}{\theta}$，$F_{(n)}(x)=\left(\dfrac{x}{\theta}\right)^n$，

$$f_{(n)}(x)=\frac{nx^{n-1}}{\theta^n},\qquad EX_{(n)}=\frac{n}{n+1}\theta,\qquad EX_{(n)}^2=\frac{n}{n+2}\theta^2.$$

（Ⅰ）无偏：

$$E(T_c)=c\cdot\frac{n}{n+1}\theta\overset{!}{=}\theta\Longrightarrow c=\frac{n+1}{n}.$$

（Ⅱ）最小 MSE：把 $h(c)$ 展开成 c 的**二次函数**：

$$h(c)=E(T_c-\theta)^2=c^2EX_{(n)}^2-2c\theta\,EX_{(n)}+\theta^2=c^2\cdot\frac{n\theta^2}{n+2}-2c\cdot\frac{n\theta^2}{n+1}+\theta^2.$$

对 c 求导：

$$h'(c)=2c\cdot\frac{n\theta^2}{n+2}-2\cdot\frac{n\theta^2}{n+1}=0\implies c=\frac{n+2}{n+1}.$$

易错点：

- ⚠️ **两个 c 不相等，且最小 MSE 的 c 更小。**比较大小（交叉相乘，两边同乘正数 $n(n+1)$ 不变号）：

$$n(n+2)=n^2+2n,\qquad (n+1)^2=n^2+2n+1\implies n(n+2)<(n+1)^2\implies\frac{n+2}{n+1}<\frac{n+1}{n}.$$

所以 $c_{\text{MSE}}=\dfrac{n+2}{n+1}<\dfrac{n+1}{n}=c_{\text{无偏}}$。

这不矛盾：**有偏估计只要偏差足够小，就能把方差压得更低，使总 MSE 更小**——这就是"**无偏 ≠ 最优**"的经典例证（$c=\frac{n+2}{n+1}$ 时估计量是有偏的，但 MSE 反而比无偏的那一个更小）。

- ⚠️ $EX_{(n)}^2$ **要单独算**（$\dfrac{n}{n+2}\theta^2$），**不能**用 $(EX_{(n)})^2$。
- 通用模板：$h'(c)=0\implies c=\dfrac{\theta\,EX_{(n)}}{EX_{(n)}^2}$（本题代入即 $\dfrac{\theta\cdot\frac{n}{n+1}\theta}{\frac{n}{n+2}\theta^2}=\dfrac{n+2}{n+1}$ ✓）。

第 18 题 • 假设检验概念综合　🟧真题骨架 2018•8

（Ⅰ）答案：**（D）**。

思路：核心是"α **越小，拒绝域越小**"。

- 拒绝域 $|\bar X-\mu_0|\geq c$，其中 $c=u_{\alpha/2}\dfrac{\sigma}{\sqrt n}$；
- α 减小 $\implies u_{\alpha/2}$ **增大** $\implies c$ **增大** $\implies$ **拒绝域变小、接受域变大。**

于是：

条件	**在 $\alpha=0.01$ 下**
$\alpha=0.05$ 时**拒绝**（$\lvert\bar X-\mu_0\rvert\geq c_{0.05}$）	因 $c_{0.01}>c_{0.05}$，**未必**达到 $\implies$ A、B 都错
$\alpha=0.05$ 时**接受**（$\lvert\bar X-\mu_0\rvert<c_{0.05}$）	因 $c_{0.05}<c_{0.01}$，**必然**也 $<c_{0.01}\implies$ **接受** $\implies$ **D 对**，C 错

> **记忆**："**小 α 拒绝 $\implies$ 大 α 必拒绝；大 α 接受 $\implies$ 小 α 必接受。**"

（Ⅱ）答案：$W=\{|\bar X-\mu_0|\geq 0.98\}$；$\beta=\Phi(-0.04)-\Phi(-3.96)\approx\Phi(-0.04)\approx 0.484$。

思路：$\dfrac{\sigma}{\sqrt n}=\dfrac{2}{4}=0.5$，

$$c=u_{0.025}\cdot\frac{\sigma}{\sqrt n}=1.96\times 0.5=0.98\implies W=\{|\bar X-\mu_0|\geq 0.98\}.$$

求 β：在 $\mu=\mu_0+1$ 下 $\bar X\sim N(\mu_0+1,\,0.25)$，标准差 0.5。**用接受域**：

$$\beta = P\{|\bar{X} - \mu_0| < 0.98 \mid \mu = \mu_0 + 1\}.$$

令 $\dfrac{\bar{X} - \mu_0 - 1}{0.5} = Z \sim N(0,1)$，则 $\bar{X} - \mu_0 = 1 + 0.5Z$，条件化为

$$-0.98 < 1 + 0.5Z < 0.98 \iff -3.96 < Z < -0.04,$$

$$\beta = \Phi(-0.04) - \Phi(-3.96) \approx 0.4840 - 0.0000 \approx 0.484.$$

易错点：

- ⚠️ β **必须在** H_1 **的具体值下算**（这里 $\mu = \mu_0 + 1$），且**用接受域**——$|\bar{X} - \mu_0| < 0.98$ 而不是 $\geq$。
- ⚠️ 计算时**把** $\bar{X} - \mu_0$ **替换成** $1 + 0.5Z$（含真值的偏移 1），这是最容易漏的一步。
- 结果 $\beta \approx 0.484$ 很大，说明**这个检验对** $\mu = \mu_0 + 1$ **的偏离"很不敏感"**——这正是 $\alpha = 0.05$ 时 β 随之增大的体现（α 小时 β 大）。

命题质量报告

一、考点覆盖表（对照第 20/21/22 章 15 道真题）

真题	考点	对应题	覆盖
2017·8	χ^2 分布辨析	A2（含同款陷阱 IV）	✅
2023·9	F 分布（方差比）	B14	✅
2016·14	区间估计（对称性）	B9（含对称性技巧）	✅
2016·23	次序统计量 $X_{(n)}$ + 无偏	B8	✅
2017·23	绝对误差 $Z = \lvert X - \mu \rvert$（变体）	B7（拉普拉斯同构）	◐ 部分（矩估计分支未另设）
2018·23	拉普拉斯 MLE + E/D	B7	✅
2019·23	半正态：求 A + MLE	C15	✅
2020·23	威布尔：无记忆性 + MLE	C16	✅
2021·9	无偏 + $D(\bar{X} - \bar{Y})$	B12（无偏/有效分支）	◐ 部分（两样本方差未单设）
2022·22	泊松两样本 MLE + D	B13	✅
2023·10	无偏（$a\lvert X_1 - X_2 \rvert$）	A6（同类"定常数"）	✅
2024·22	$U(0,\theta)$：无偏 + MSE	C17	✅
2018·8	显著性水平与拒绝/接受	C18（Ⅰ）	✅
2021·10	第二类错误	B11	✅
2025·10	单侧检验拒绝域	B10	✅

覆盖结论：15 道真题中 **13 题完全覆盖** ✅、**2 题部分覆盖** ◐。

二、难度分布与用时

组	题量	建议用时	难度
A 基础	6（1–6）	22 分钟	★☆☆
B 中等	8（7–14）	34 分钟	★★☆
C 拔高	4（15–18）	24 分钟	★★★
合计	**18**	**约 80 分钟**	可拆两天（第 1 天 1–9，第 2 天 10–18）

三、真题使用与防过拟合

- **本题卷 18 题全部为原创命题**，数字、样本量、分布参数均与真题不同；
- 标「**真题骨架**」的 10 道题**只对齐考法结构**（如"两样本不同参数泊松 MLE"），**不做数字复用**，做完可直接回刷真题原题而不会"背答案"；
- **参照了 🟩 参考仓库**的 4 道例题风格（拉普拉斯 MLE、均匀分布矩估计、两正态方差比、无偏性定 k），并在此基础上把数字重设。

四、仍未覆盖的方向（诚实清单）

1. **区间估计的 σ^2 分支**（含 χ^2 分布的两端分位点"反着放"）——真题 2016·14 只考均值，本卷也未设（§4.3 已在手册中写明，可自行练手）；
2. **两样本 t 检验**（$\sigma_1^2=\sigma_2^2$ 时的合并方差 S_w）——真题十年未考，数一考纲内但低频；
3. **分布函数的拟合优度 χ^2 检验**——超出考研范围（了解即可）；
4. **2017•23 的"矩估计分支"**——该真题要求先用一阶矩求 σ 的矩估计量再求 MLE，本卷 B7 只做了 MLE 分支，矩估计的"绝对误差一阶矩"未单设；
5. **2021•9 的 $D(\bar{X}-\bar{Y})$ 两样本方差计算**——考纲内、涉及二元正态的 ρ，本卷以 B12 的"单样本有效性"替代。

> 上述 1、4、5 三项可在二轮复习时用真题原题补齐；2、3 属边缘，暂不安排。

五、与既有材料的衔接

- 本讲**上游**：第 4 讲（数字特征，E, D, Cov——**尚未建文件夹**）、第 5 讲（大数定律与中心极限定理，切比雪夫不等式）。
- 本讲**核心依赖**：正态分布、$\chi^2/t/F$ 三大分布、期望方差的性质。
- 建议顺序：**第 5 讲 → 第 6 讲**（因为区间估计与假设检验反复用到第 5 讲的标准化与分位数）。