数理统计期末练习题

Chapter 1

每章练习题

1.1　充分统计量

费希尔-内曼分解定理/因子分解定理

求解充分统计量的核心工具是因子分解定理。求解充分统计量的“四步法”总结如下：

第一步：写出联合分布

首先，根据题目给出的总体分布，写出样本 $x_1, x_2, \ldots, x_n$ 的联合分布。$L(\theta; x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i; \theta)$

注：如果分布的取值范围依赖于参数 θ（比如均匀分布 $U(0, \theta)$），一定要使用示性函数 $I(\cdot)$ 把范围写进表达式里。

第二步：尝试“拆分”表达式

观察写出的 $L(\theta; \mathbf{x})$，尝试把它拆成两个部分的乘积：$L(\theta; \mathbf{x}) = g(T(\mathbf{x}), \theta) \cdot h(\mathbf{x})$

$g(T(\mathbf{x}), \theta)$ 部分：这一部分是关于参数 θ 和随机变量 x 的函数，$T(\mathbf{x})$ 是仅关于随机变量的函数。

$h(\mathbf{x})$ 部分：这一部分与参数 θ 无关，是只含随机变量 x_i 的函数。

第三步：识别 $T(\mathbf{x})$

在 g 部分中，那个把所有样本信息“打包”起来的表达式 $T(\mathbf{x})$，就是我们要找的充分统计量。

第四步：给出结论，写出 $h(x), g(T(x), \theta), T(x)$

问题 1. 设 $x_1, x_2, \cdots, x_n$ 是来自几何分布总体

$$P\{X = x\} = p(1-p)^{x-1}, \quad x = 1, 2, \cdots$$

的简单样本，其中 $0 < p < 1$。试求参数 p 的充分统计量。

解析： 我们利用因子分解定理。首先写出样本的联合分布。

1 样本的联合分布为：

$$L(p)=\prod_{i=1}^{n}p(1-p)^{x_i-1}=p^n(1-p)^{\sum_{i=1}^{n}x_i-n}$$

2 根据因子分解定理，如果似然函数可以写成 $L(p)=g(T(x),p)\cdot h(x)$ 的形式，则 $T(x)$ 是充分统计量。在本题中，我们可以令：

$$g(T(x),p)=p^n(1-p)^{\sum x_i-n},\quad h(x_1,\dots,x_n)=1$$

其中 $T(x_1,\dots,x_n)=\sum_{i=1}^{n}x_i$。

结论： $\sum_{i=1}^{n}x_i$ 是参数 p 的充分统计量（同样地，样本均值 $\bar{x}$ 也是充分统计量）。

问题 2. 设 $x_1,x_2,\cdots,x_n$ 是来自离散总体 X 的样本，X 的分布列为

$$P\{X=v_i\}=p_i,\quad i=1,2,\cdots,m$$

其中 $p_1,p_2,\cdots,p_m$ 是参数。用 n_i 表示 $x_1,x_2,\cdots,x_n$ 中等于 v_i 的个数,证明:$(n_1,n_2,\cdots,n_m)$ 是充分统计量。

证明： *1* 样本的联合分布可以表示为：

$$L(x_1,\dots,x_n;p_1,\dots,p_m)=\prod_{j=1}^{n}P(X=x_j)$$

2 由于在 n 个观测值中，v_1 出现了 n_1 次，v_2 出现了 n_2 次，…，v_m 出现了 n_m 次，因此联合概率可以写成：

$$L=p_1^{n_1}p_2^{n_2}\cdots p_m^{n_m}$$

显然，这个表达式只依赖于统计量 $T=(n_1,n_2,\cdots,n_m)$ 和参数 $(p_1,\dots,p_m)$。
3 令 $h(x_1,\dots,x_n)=1$，根据因子分解定理，$T(x_1,\dots,x_n)=(n_1,n_2,\cdots,n_m)$ 是参数的充分统计量。

问题 3. 设 $x_1,x_2,\cdots,x_n$ 是来自均匀分布 $U(\theta,2\theta)(\theta>0)$ 总体 X 的样本，求充分统计量。

解析： 对于均匀分布，关键在于使用指示函数 $I(\cdot)$ 来表示范围。
1 总体密度函数为：$f(x;\theta)=\dfrac{1}{2\theta-\theta}=\dfrac{1}{\theta},\quad \theta<x_i<2\theta$。
2 样本的联合密度函数为：

$$L(\theta)=\prod_{i=1}^{n}\frac{1}{\theta}I_{(\theta,2\theta)}(x_i)=\frac{1}{\theta^n}I(\theta<x_i<2\theta)=\frac{1}{\theta^n}I(\theta<x_{(1)})I(x_{(n)}<2\theta)$$

结论： 充分统计量为 $T(x_1,\dots,x_n)=(x_{(1)},x_{(n)})$。

问题 4. 设总体 X 的分布是双参数指数分布，其密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{1}{\sigma}\exp\left\{-\dfrac{x-\mu}{\sigma}\right\}, & x>\mu\\ 0, & x\le\mu\end{cases}$$

其中 $\sigma(\sigma>0)$ 及 μ 是常数。$x_1,x_2,\cdots,x_n$ 是来自总体 X 的样本。

1. 证明：当 σ 已知，μ 未知时，$x_{(1)} = \min\limits_{1\le i\le n}\{x_i\}$ 是 μ 的充分统计量；

2. 当 μ 已知，σ 未知时，求 σ 的充分统计量。

3. 求参数 $\theta = (\mu, \sigma)$ 的充分统计量。

【解析】

准备工作：首先，我们要写出样本 $x_1, \ldots, x_n$ 的联合分布。注意，这里有一个隐含的条件 $x_i > \mu$，我们用指示函数 $I(x > \mu)$ 来表示。

$$L(x_1,\ldots,x_n;\mu,\sigma) = \prod_{i=1}^{n} \frac{1}{\sigma}\exp\left\{-\frac{x_i-\mu}{\sigma}\right\} I(x_i > \mu) = \frac{1}{\sigma^n}\exp\left\{\frac{-\sum_{i=1}^{n} x_i + n\mu}{\sigma}\right\} I(x_{(1)} > \mu)$$

1σ 已知，证明 $x_{(1)}$ 是 μ 的充分统计量

证明：当 σ 已知时，参数只有 μ。我们将上面的 L 进行整理，尝试分离出只含 x 的部分和含 μ 的部分：

$$L(\mu) = \underbrace{\exp\left\{\frac{n\mu}{\sigma}\right\} I(x_{(1)} > \mu)}_{g(x_{(1)};\mu)} \cdot \underbrace{\frac{1}{\sigma^n}\exp\left\{-\frac{\sum x_i}{\sigma}\right\}}_{h(x_1,\ldots,x_n)}$$

根据分解定理，$x_{(1)}$ 是 μ 的充分统计量。

2μ 已知，求 σ 的充分统计量

当 μ 已知时，参数只有 σ。我们重新观察联合密度：

$$L(\sigma) = \frac{1}{\sigma^n}\exp\left\{-\frac{\sum(x_i-\mu)}{\sigma}\right\} I(x_{(1)} > \mu)$$

由于 μ 已知，指示函数 $I(x_{(1)} > \mu)$ 的值要么是 0 要么是 1，它不依赖于未知参数 σ，可以归入 $h(x)$。剩下的部分 $\frac{1}{\sigma^n}\exp\left\{-\frac{\sum(x_i-\mu)}{\sigma}\right\}$ 中，样本 x_i 是通过 $\sum_{i=1}^{n}(x_i-\mu)$ 这个整体与 σ 发生关系的。

因此，σ 的充分统计量为：

$$T = \sum_{i=1}^{n}(x_i - \mu)$$

（注：由于 μ 已知，写成 $\sum x_i$ 或样本均值 $\bar{x}$ 也是可以的）。

3 求参数 $\theta = (\mu, \sigma)$ 的充分统计量

解：现在 μ 和 σ 都是未知的。我们需要找到一组统计量，使得整个似然函数能被分解。回到联合密度函数：

$$L(\mu,\sigma) = \frac{1}{\sigma^n}\exp\left\{\frac{n\mu}{\sigma}\right\}\exp\left\{-\frac{\sum x_i}{\sigma}\right\} I(x_{(1)} > \mu)$$

观察发现，为了涵盖所有的未知参数，我们需要同时知道 $x_{(1)}$ 和 $\sum x_i$。

我们可以令:

$$T = (x_{(1)}, \sum_{i=1}^{n} x_i)$$

此时:

$$g(T;\mu,\sigma) = \frac{1}{\sigma^n}\exp\left\{\frac{n\mu - T_2}{\sigma}\right\} I(T_1 > \mu)$$
$$h(x) = 1$$

根据分解定理，联合充分统计量为:

$$T = (x_{(1)}, \sum_{i=1}^{n} x_i)$$

或者也可以写成 $(x_{(1)}, \bar{x})$。

1.2 参数估计

1.2.1 频率替换估计法

这类题目总结为以下三个固定步骤:

1 找概率模型: 根据题目给出的分布 (如正态分布、指数分布), 计算出某个特定事件 (如 $X \leq 0$ 或 $X < 2$) 发生的理论概率 P。这个 P 通常是关于待求参数的函数。

2 找样本频率: 题目通常会告诉你 "n 个样本中有 m 个满足条件", 那么样本频率就是 m/n。

3 解方程: 令理论概率 = 样本频率, 这就像是我们在用现实中观察到的比例去 "倒推" 模型中的参数。

问题 5. 设 $x_1, x_2, \ldots, x_n$ 是来自正态总体 $N(\mu, 1)$ 的简单样本, 在这 n 个样本观测值中有 m 个取值不大于零, 求 μ 的频率替换估计。

【解析】

设总体为 $X \sim N(\mu, 1)$。

1 确定理论概率: 考虑事件 $A = \{X \leq 0\}$。我们要计算 $P(X \leq 0)$, 但我们手头只有标准正态分布表 (即 $N(0,1)$ 的 Φ 函数)。通过公式 $Z = \dfrac{X - E(X)}{\sqrt{Var(X)}}$, 我们将 X 转化为标准正态分布, 该事件发生的理论概率 p 为:

$$p = P(X \leq 0) = P\left(\frac{X-\mu}{1} \leq \frac{0-\mu}{1}\right) = \Phi(-\mu)$$

其中 $\Phi(\cdot)$ 是标准正态分布的累积分布函数。

2 确定样本频率: 根据题意, 在 n 个样本观测值中有 m 个取值不大于零, 因此事件 A 发生的样本频率为:

$$\hat{p} = \frac{m}{n}$$

3 频率替换建立方程：利用频率替换原理，令理论概率等于样本频率：

$$\Phi(-\hat{\mu}) = \frac{m}{n}$$

4 解方程求估计量：对等式两边同时作用标准正态分布函数的反函数 Φ^{-1}，得：

$$-\hat{\mu} = \Phi^{-1}\left(\frac{m}{n}\right)$$

即：

$$\hat{\mu} = -\Phi^{-1}\left(\frac{m}{n}\right)$$

问题 6 (★). *设总体 X 服从两点分布 $B(1,p)$，其中参数 $p \in (0,1)$，$x_1, x_2, \cdots, x_n$ 是来自总体 X 的简单样本。(1) 求 $q(p) = Var_p(\bar{x})$；(2) 求 $q(p)$ 的频率替换估计。*

【解析】

第一问：计算理论方差

已知 $X \sim B(1,p)$，则总体的期望 $E(X) = p$，方差 $\mathrm{Var}(X) = p(1-p)$。样本均值 $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$，其中 $x_1, x_2, \ldots, x_n$ 是独立同分布于 X 的样本。根据方差的性质：对于独立随机变量 $x_1, \ldots, x_n$，有

$$\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(x_i)$$

由于每个 $x_i \sim B(1,p)$，故 $\mathrm{Var}(x_i) = p(1-p)$，代入得：

$$\mathrm{Var}_p(\bar{x}) = \frac{1}{n^2} \cdot n \cdot p(1-p) = \frac{p(1-p)}{n}$$

因此，$q(p) = \boxed{\frac{p(1-p)}{n}}$。

第二问：频率替换估计

在两点分布中，参数 p 的含义是事件 $\{X = 1\}$ 发生的概率。根据频率替换原则，我们用样本中事件发生的频率来估计概率。对于 $B(1,p)$，因为 x_i 只取 0 或 1，**样本均值 $\bar{x}$ 恰好就是事件 $\{X = 1\}$ 出现的频率**。因此，频率估计 $\hat{p} = \bar{x}$。将 $\hat{p} = \bar{x}$ 代入 $q(p)$ 的表达式中，得到 $q(p)$ 的频率替换估计量为：

$$\hat{q}(p) = \frac{\bar{x}(1-\bar{x})}{n}$$

问题 7. *设总体 X 服从参数为 $\lambda(\lambda > 0)$ 的***指数分布***，其密度函数为*

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

$x_1, x_2, \cdots, x_n$ 是来自总体 X 的简单样本，且这 n 个观察值中有 m 个取值小于 2，求 λ 的频率替换估计。

【解析】

第一步：计算理论概率

指数分布是连续型随机变量，X 的取值为连续非负实数（$[0,+\infty)$）。设事件 $A=\{X<2\}$。计算 X 落在该区间的概率：

$$P(A)=\int_0^2 \lambda e^{-\lambda x}dx=\left[-e^{-\lambda x}\right]_0^2=1-e^{-2\lambda}$$

第二步：建立等式

样本中事件 A（取值小于 2）发生的频率为 $\frac{m}{n}$。令理论概率等于样本频率：

$$1-e^{-2\lambda}=\frac{m}{n}$$

第三步：解出参数

整理等式：

$$e^{-2\lambda}=1-\frac{m}{n}=\frac{n-m}{n}$$

两边取自然对数：

$$-2\lambda=\ln\left(1-\frac{m}{n}\right)$$

解得 λ 的频率替换估计量为：

$$\hat{\lambda}=-\frac{1}{2}\ln\left(1-\frac{m}{n}\right)$$

1.2.2 矩估计和极大似然估计

问题 8. 设总体 X 的概率密度函数为

$$f(x,\theta)=\begin{cases}\frac{1}{2\theta}, & 0\le x<\theta\\ \frac{1}{2(1-\theta)}, & \theta\le x\le 1\\ 0, & \text{其他}\end{cases}$$

其中参数 $\theta\in(0,1)$，$x_1,x_2,\cdots,x_n$ 是来自总体 X 的简单样本，求参数 θ 的矩估计 $\hat{\theta}$。

【解析】

解题思路：矩估计法的核心是利用“样本矩等于总体矩”。这里只有一个未知参数 θ，我们只需要用到样本一阶原点矩，即期望 $E(X)$。

1 计算总体的期望 $E(X)$：根据期望的定义 $E(X)=\int_{-\infty}^{+\infty}xf(x,\theta)dx$：

$$\begin{aligned}E(X)&=\int_0^\theta x\cdot\frac{1}{2\theta}dx+\int_\theta^1 x\cdot\frac{1}{2(1-\theta)}dx\\&=\frac{1}{2\theta}\left[\frac{1}{2}x^2\right]_0^\theta+\frac{1}{2(1-\theta)}\left[\frac{1}{2}x^2\right]_\theta^1\\&=\frac{1}{2\theta}\cdot\frac{\theta^2}{2}+\frac{1}{2(1-\theta)}\cdot\frac{1-\theta^2}{2}\\&=\frac{\theta}{4}+\frac{(1-\theta)(1+\theta)}{4(1-\theta)}\\&=\frac{\theta}{4}+\frac{1+\theta}{4}=\frac{2\theta+1}{4}\end{aligned}$$

2 样本矩等于总体矩，$E(X)=\bar{x}$，反解出 θ：

$$2\theta+1=4\bar{x}\implies 2\theta=4\bar{x}-1\implies\theta=2\bar{x}-\frac{1}{2}$$

结论：参数 θ 的矩估计量为 $\hat{\theta}=2\bar{\boldsymbol{x}}-\frac{1}{2}$。

问题 9. 设总体 X 服从对数正态分布 $\mathrm{LN}(\mu,\sigma^2)$，其密度函数为

$$f(x;\mu,\sigma^2)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\sigma x}\exp\left\{-\dfrac{(\ln x-\mu)^2}{2\sigma^2}\right\}, & x>0,\\ 0, & x\le 0,\end{cases}$$

其中 μ 和 σ^2 均未知，$x_1,x_2,\cdots,x_n$ 是来自总体 X 的简单样本，求 μ 及 σ^2 的矩估计和极大似然估计。

【解析】

这道题目是对数正态分布参数估计的经典题型。对数正态分布与正态分布有着紧密的联系：如果 X 服从对数正态分布，那么 $Y=\ln X$ 就服从正态分布。考试现场推导不太现实，所以给出如下通用结论，当做结论记：

$$E(X^k)=\int_0^{+\infty}x^kf(x;\mu,\sigma^2)dx=\exp\left\{\mu k+\frac{1}{2}\sigma^2k^2\right\}$$

所以，在解答之前，我们有对数正态分布 $X\sim\mathrm{LN}(\mu,\sigma^2)$ 的前两阶原点矩：$E(X)=\exp\{\mu+\frac{1}{2}\sigma^2\}$, $E(X^2)=\exp\{2\mu+2\sigma^2\}$。

一、矩估计

矩估计法的核心是令总体矩等于样本矩。由于有两个未知参数，我们需要用到样本一阶和二阶原点矩。

1 建立方程组：设总体的一阶原点矩为 $E(X)$，二阶原点矩为 $E(X^2)$。样本的一阶原点矩为 $\bar{x}=\frac{1}{n}\sum x_i$，二阶原点矩为 $A_2=\frac{1}{n}\sum x_i^2$。我们有：

$$\begin{cases}e^{\mu+\frac{1}{2}\sigma^2}=\bar{x} & (1)\\ e^{2\mu+2\sigma^2}=A_2 & (2)\end{cases}$$

2 求解方程组：对两边取自然对数：

$$\begin{cases} \mu + \frac{1}{2}\sigma^2 = \ln \bar{x} & (3) \\ 2\mu + 2\sigma^2 = \ln A_2 & (4) \end{cases}$$

解得：

$$\hat{\sigma}^2 = \ln A_2 - 2\ln \bar{x} = \ln\left(\frac{A_2}{\bar{x}^2}\right)$$

$$\hat{\mu} = \ln \bar{x} - \frac{1}{2}\ln\left(\frac{A_2}{\bar{x}^2}\right) = \ln \bar{x} - \ln\sqrt{\frac{A_2}{\bar{x}^2}} = \ln\left(\frac{\bar{x}^2}{\sqrt{A_2}}\right)$$

结论： μ 和 σ^2 的矩估计量分别为：

$$\hat{\mu} = 2\ln \bar{x} - \frac{1}{2}\ln\left(\frac{1}{n}\sum x_i^2\right), \quad \hat{\sigma}^2 = \ln\left(\frac{1}{n}\sum x_i^2\right) - 2\ln \bar{x}$$

二、**极大似然估计**

1 写出似然函数：

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} f(x_i; \mu, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\left(\prod_{i=1}^{n}\frac{1}{x_i}\right)\exp\left\{-\sum_{i=1}^{n}\frac{(\ln x_i - \mu)^2}{2\sigma^2}\right\}$$

2 取对数似然函数：

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

3 求偏导并令其为零：

对 μ 求**偏导**：

$$\frac{\partial \ln L}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(\ln x_i - \mu) = 0 \implies \sum_{i=1}^{n}\ln x_i - n\mu = 0$$

解得：$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\ln x_i$。

对 σ^2 求**偏导**：

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(\ln x_i - \mu)^2 = 0$$

将 $\hat{\mu}$ 代入，解得：$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(\ln x_i - \hat{\mu})^2$。

结论： μ 和 σ^2 的极大似然估计量分别为：

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\ln x_i, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(\ln x_i - \hat{\mu})^2$$

问题 10. 设总体 X 的**分布函数**为

$$F(x;\beta)=\begin{cases}1-\dfrac{1}{x^{\beta}}, & x>1\\ 0, & x\le 1\end{cases}$$

其中 $\beta(\beta>1)$ 是未知参数，$x_1,x_2,\ldots,x_n$ 是来自总体 X 的简单样本。试求 β 的矩估计和极大似然估计。

【解析】

首先要看清题目给的是分布函数，分布函数与概率密度函数关系总结如下：
分布函数 $F(x)$：是“到此为止的总概率”，范围在 0 到 1 之间，是单调不减的。
概率密度 $f(x)$：是“概率在某处的分布强度”，它本身不是概率，它在某个区间上的积分才是概率。
关系：$F(x)\xrightarrow{\text{求导}}f(x)$；$f(x)\xrightarrow{\text{积分}}F(x)$。

因此，在进行估计之前，我们先求出总体的概率密度函数 $f(x;\beta)$。对分布函数 $F(x;\beta)$ 求导得：

$$f(x;\beta)=\frac{d}{dx}F(x;\beta)=\begin{cases}\beta x^{-(\beta+1)}, & x>1\\ 0, & \text{其他}\end{cases}$$

1 矩估计法

第一步：计算总体均值（一阶原点矩）

$$E(X)=\int_1^{+\infty}x\cdot f(x;\beta)dx=\int_1^{+\infty}x\cdot\beta x^{-\beta-1}dx=\beta\int_1^{+\infty}x^{-\beta}dx$$

由于题目给定 $\beta>1$，积分收敛：

$$E(X)=\beta\left[\frac{x^{-\beta+1}}{-\beta+1}\right]_1^{+\infty}=\beta\left(0-\frac{1}{1-\beta}\right)=\frac{\beta}{\beta-1}$$

第二步：令样本均值等于总体均值，令 $E(X)=\bar{x}$，即：

$$\frac{\beta}{\beta-1}=\bar{x}$$

第三步：解出 β

$$\beta=\bar{x}(\beta-1)\implies\beta=\bar{x}\beta-\bar{x}\implies\bar{x}=\beta(\bar{x}-1)$$

解得 β 的矩估计量为：

$$\hat{\beta}_M=\frac{\bar{x}}{\bar{x}-1}$$

2 极大似然估计法

第一步：写出似然函数 $L(\beta)$

$$L(\beta)=\prod_{i=1}^{n}f(x_i;\beta)=\prod_{i=1}^{n}\beta x_i^{-(\beta+1)}=\boxed{\beta^n\left(\prod_{i=1}^{n}x_i\right)^{-(\beta+1)}}$$

第二步：取对数似然函数 $\ln L(\beta)$，为了方便求导，我们取自然对数：

$$\ln L(\beta) = \ln\left[\beta^n\left(\prod_{i=1}^n x_i\right)^{-(\beta+1)}\right] = n\ln\beta - (\beta+1)\sum_{i=1}^n \ln x_i$$

第三步：对 β 求导并令其为 0

$$\frac{d}{d\beta}\ln L(\beta) = \frac{n}{\beta} - \sum_{i=1}^n \ln x_i = 0$$

第四步：解出 β

$$\frac{n}{\beta} = \sum_{i=1}^n \ln x_i \implies \hat{\beta}_L = \frac{n}{\sum_{i=1}^n \ln x_i}$$

由于 β 的二阶导数小于 0，该点确实是极大值点。所以 β 的极大似然估计量为：

$$\hat{\beta}_L = \frac{n}{\sum_{i=1}^n \ln x_i}$$

问题 11. 设 $x_1, x_2, \cdots, x_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的简单样本，μ 和 σ^2 均为未知，试求 $P\{X \le t\}$ 的极大似然估计。

【解析】

解题思路：本题考察极大似然估计的不变性原理。如果 $\hat{\theta}$ 是 θ 的极大似然估计，那么对于任何函数 $g(\theta)$，$g(\hat{\theta})$ 就是 $g(\theta)$ 的极大似然估计。

1 确定参数的极大似然估计：对于正态总体 $N(\mu, \sigma^2)$，我们已知其参数的极大似然估计分别为：

$$\hat{\mu} = \bar{x}, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})^2$$

2 写出待求目标的表达式：令 $g(\mu, \sigma^2) = P\{X \le t\}$。由于 $X \sim N(\mu, \sigma^2)$，进行标准化处理：

$$P\{X \le t\} = P\left\{\frac{X-\mu}{\sigma} \le \frac{t-\mu}{\sigma}\right\} = \Phi\left(\frac{t-\mu}{\sigma}\right)$$

其中 $\Phi(\cdot)$ 是标准正态分布的累积分布函数。

3 利用不变性原理：将 $\hat{\mu}$ 和 $\hat{\sigma}$ 代入上式即可得到 $P\{X \le t\}$ 的极大似然估计：

$$\hat{P}\{X \le t\} = \Phi\left(\frac{t - \bar{x}}{\sqrt{\frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})^2}}\right)$$

问题 12 (离散型随机变量极大似然估计)**.** 设总体 X 的概率分布列为

X	0	1	2	3
p	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中 $\theta\left(0 < \theta < \frac{1}{2}\right)$ 是未知参数，试求样本值为 $3, 1, 3, 0, 3, 1, 2, 3$ 时 θ 的极大似然估计

【解析】

这道题目是离散型随机变量的极大似然估计问题。处理这类题目的关键在于：先统计样本中各个数值出现的频数，再根据概率分布列写出似然函数。

第一步：统计样本频数

$x=0$ 出现了 1 次；$x=1$ 出现了 2 次；$x=2$ 出现了 1 次；$x=3$ 出现了 4 次。样本总数 $n=1+2+1+4=8$。

第二步：写出似然函数 $L(\theta)$

根据离散型随机变量极大似然估计的定义，似然函数为各样本值对应概率的乘积：

$$\begin{aligned}
L(\theta) &= P(X=0)^1 \cdot P(X=1)^2 \cdot P(X=2)^1 \cdot P(X=3)^4 \\
&= (\theta^2)^1 \cdot [2\theta(1-\theta)]^2 \cdot (\theta^2)^1 \cdot (1-2\theta)^4 \\
&= \theta^2 \cdot 4\theta^2(1-\theta)^2 \cdot \theta^2 \cdot (1-2\theta)^4 \\
&= 4\theta^6(1-\theta)^2(1-2\theta)^4
\end{aligned}$$

第三步：取对数似然函数 $\ln L(\theta)$

$$\ln L(\theta) = \ln 4 + 6\ln\theta + 2\ln(1-\theta) + 4\ln(1-2\theta)$$

第四步：对 θ 求导并解似然方程

$$\frac{d}{d\theta}\ln L(\theta) = \frac{6}{\theta} - \frac{2}{1-\theta} - \frac{4\cdot 2}{1-2\theta} = \frac{6}{\theta} - \frac{2}{1-\theta} - \frac{8}{1-2\theta}$$

令导数为 0：

$$\frac{6}{\theta} - \frac{2}{1-\theta} - \frac{8}{1-2\theta} = 0$$

化简得：

$$12\theta^2 - 14\theta + 3 = 0$$

第五步：确定估计值，利用求根公式解方程：

$$\theta = \frac{14 \pm \sqrt{14^2 - 4\times 12\times 3}}{2\times 12} = \frac{7\pm\sqrt{13}}{12}$$

判断是否符合约束条件 $0<\theta<\frac{1}{2}$：

$\sqrt{13}\approx 361$，$\theta_1 = \frac{7+361}{12} \approx 088$（不符合 $\theta < 05$，舍去），$\theta_2 = \frac{7-361}{12} \approx 028$（符合 $0<\theta<05$）

因此，θ 的极大似然估计量为：

$$\hat{\theta} = \frac{7-\sqrt{13}}{12}$$

问题 13. 设 X 为某电子元件的失效时间，其密度函数为

$$f(x;\alpha,\beta) = \begin{cases} \beta e^{-\beta(x-\alpha)}, & x>\alpha>0 \\ 0, & x\le\alpha \end{cases}$$

其中 $\beta > 0$。抽取 n 个该电子元件，独立地进行测试，失效时间分别为 $x_1, x_2, \cdots, x_n$。*(1)* 当 α 已知时，求 β 的极大似然估计；*(2)* 当 β 已知时，求 α 的极大似然估计；*(3)* 当 α 和 β 都是未知参数时，求 α 和 β 的极大似然估计。

【解析】

解题思路：极大似然估计的第一步是写出似然函数 L。注意密度函数中的范围限制 $x > \alpha$，这在求 α 的估计时至关重要。

似然函数：

$$L(\alpha,\beta) = \prod_{i=1}^{n} f(x_i;\alpha,\beta) = \begin{cases} \boxed{\beta^n \exp\{-\beta \sum_{i=1}^{n}(x_i-\alpha)\}}, & x_{(1)} > \alpha > 0 \\ 0, & \text{其他} \end{cases}$$

(1) α 已知，求 β：

取对数似然函数：$\ln L = n\ln\beta - \beta\sum_{i=1}^{n}(x_i-\alpha)$。对 β 求导并令其为 0：

$$\frac{d\ln L}{d\beta} = \frac{n}{\beta} - \sum_{i=1}^{n}(x_i-\alpha) = 0 \implies \hat{\beta} = \frac{n}{\sum_{i=1}^{n}(x_i-\alpha)}$$

(2) β 已知，求 α：

观察 $L(\alpha,\beta) = \beta^n e^{-\beta\sum x_i} e^{n\beta\alpha}$。因为 $\beta > 0$，所以 L 是关于 α 的增函数。为了使 L 最大，α 应该尽可能大。但受到条件 $x_{(1)} > \alpha$ 的限制，α 不能超过样本中的最小值。故 $\hat{\alpha} = x_{(1)}$。

(3) α, β 均未知：根据 (2)，无论 β 取何值，使 L 最大的 α 总是 $\hat{\alpha} = x_{(1)}$。将 $\hat{\alpha}$ 代入 (1) 的结果中，得到：

$$\hat{\alpha} = x_{(1)}, \quad \hat{\beta} = \frac{n}{\sum_{i=1}^{n}(x_i - x_{(1)})}$$

问题 14. 设总体 X 的概率密度函数为

$$f(x;\theta) = \begin{cases} \theta, & 0 \le x < 1 \\ 1-\theta, & 1 \le x \le 2 \\ 0, & \text{其他} \end{cases}$$

其中 $\theta (0 < \theta < 1)$ 是未知参数，$x_1, x_2, \cdots, x_n$ 是来自总体 X 的简单样本。求参数 θ 的极大似然估计 $\hat{\theta}$。

【解析】

解题思路：这是一个**离散型思维的连续分布问题**。我们需要统计样本落入不同区间的个数。

1 统计样本分布：设样本中有 k 个观测值落入区间 $[0,1)$，其余 $n-k$ 个观测值落入区间 $[1,2]$。这里 $k = \sum_{i=1}^{n} I(0 \le x_i < 1)$

2 写出似然函数：

$$L(\theta)=\prod_{i=1}^{n} f(x_i;\theta)=\theta^k(1-\theta)^{n-k}$$

3 取对数并求导：$\ln L=k\ln\theta+(n-k)\ln(1-\theta)$

$$\frac{d\ln L}{d\theta}=\frac{k}{\theta}-\frac{n-k}{1-\theta}=0$$

$$k(1-\theta)=(n-k)\theta \implies k-k\theta=n\theta-k\theta \implies n\theta=k$$

4 结论：$\hat{\theta}=\frac{k}{n}$，即样本落入 $[0,1)$ 区间的频率。

问题 15 (参数空间有限制条件)**.** 设 $x_1,x_2,\cdots,x_n$ 是来自正态总体 $N(\mu,\sigma^2)$ 的简单样本，μ 和 σ^2 都是未知参数，且 $\mu\le\mu_0$，其中 μ_0 是已知常数。求 μ 和 σ^2 的极大似然估计。

【解析】

解题思路：先求出无约束下的极值点，再判断这个点是否在限制范围内；如果不在，则根据函数的单调性在边界上寻找最值。

第一步：写出似然函数 $L(\mu,\sigma^2)$

由于样本相互独立，似然函数等于各样本概率密度函数的乘积：

$$L(\mu,\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

整理得：

$$L(\mu,\sigma^2)=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

第二步：写出对数似然函数 $\ln L(\mu,\sigma^2)$，取对数以便于求导计算：

$$\ell(\mu,\sigma^2)=\ln L(\mu,\sigma^2)=-\frac{n}{2}\ln(2\pi)-\frac{n}{2}\ln(\sigma^2)-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

第三步：分析无约束情况下的极值点，如果不考虑约束条件 $\mu\le\mu_0$，我们对 μ 和 σ^2 分别求偏导并令其为 0：

$\frac{\partial\ell}{\partial\mu}=\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)=0 \implies \hat{\mu}_{un}=\bar{x}$

$\frac{\partial\ell}{\partial\sigma^2}=-\frac{n}{2\sigma^2}+\frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2=0 \implies \hat{\sigma}^2_{un}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$

这说明，对数似然函数在 $\mu=\bar{x}$ 处取得关于 μ 的全局最大值。

第四步：结合约束条件 $\mu\le\mu_0$ 进行讨论，我们需要在区域 $\mu\in(-\infty,\mu_0]$ 上寻找使 $\ell(\mu,\sigma^2)$ 最大的参数。

情况 1：当 $\bar{x}\le\mu_0$，此时，无约束的最大值点 $\bar{x}$ 已经落在限制区域内。因此，μ 的极大似然估计就是样本均值：$\hat{\mu}=\bar{x}$。将 $\hat{\mu}$ 代入关于 σ^2 的估计式中，得到：$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$

情况 2：当 $\bar{x} > \mu_0$ ，此时，无约束的最大值点 $\bar{x}$ 在限制区域外面。**因为 μ 有约束，所以先讨论与 μ 相关的部分。进而确定无约束的参数**。观察对数似然函数 $\ell(\mu, \sigma^2)$ 中关于 μ 的部分：$-\dfrac{1}{2\sigma^2}\sum(x_i-\mu)^2$。这是一个开口向下的二次函数，其对称轴为 $\mu=\bar{x}$。当 $\mu<\bar{x}$ 时，该函数是单调递增的。因为我们的约束范围是 $\mu \le \mu_0 < \bar{x}$，所以在整个约束区间内，函数值随着 μ 的增大而增大。因此，最大值只能在边界 $\mu=\mu_0$ 处取得：$\hat{\mu}=\mu_0$。此时，再将 $\hat{\mu}=\mu_0$ 代入对 σ^2 求偏导的方程，得到：$\hat{\sigma}^2=\dfrac{1}{n}\sum_{i=1}^n(x_i-\mu_0)^2$

第五步：综合结论综上所述，μ 和 σ^2 的极大似然估计为：

$$\hat{\mu}=\min(\bar{x}, \mu_0)$$

$$\hat{\sigma}^2=\begin{cases}\dfrac{1}{n}\sum_{i=1}^n(x_i-\bar{x})^2, & \bar{x}\le\mu_0\\ \dfrac{1}{n}\sum_{i=1}^n(x_i-\mu_0)^2, & \bar{x}>\mu_0\end{cases}$$

问题 16 (双样本问题，仅做了解)**.** 设 $x_1, x_2, \cdots, x_{n_1}$ 是来自正态总体 $N(\mu_1, \sigma^2)$ 的简单样本，$y_1, y_2, \cdots, y_{n_2}$ 是来自正态总体 $N(\mu_1-\mu, \sigma^2)$ 的简单样本，且两样本相互独立，其中 $\mu \le 0$。求参数 μ, μ_1, σ^2 的极大似然估计。

【解析】

解题思路：这是一个双样本问题。**令 $\mu_2=\mu_1-\mu$，注意 $\mu\le 0$ 意味着 $\mu_1-\mu\ge\mu_1$，则约束条件为 $\mu_2\ge\mu_1$。**

1 写出联合似然函数：由于两个样本 X 和 Y 是相互独立的，且样本内部也是独立同分布的，所以联合似然函数等于所有观测值概率密度的乘积：

$$L(\mu_1, \mu_2, \sigma^2)=\left[\prod_{i=1}^{n_1} f(x_i;\mu_1,\sigma^2)\right]\cdot\left[\prod_{j=1}^{n_2} f(y_j;\mu_2,\sigma^2)\right]$$

代入正态分布公式：

$$L=\left[\prod_{i=1}^{n_1}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\dfrac{(x_i-\mu_1)^2}{2\sigma^2}}\right]\cdot\left[\prod_{j=1}^{n_2}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\dfrac{(y_j-\mu_2)^2}{2\sigma^2}}\right]$$

合并指数项和常数项：

$$L=(2\pi\sigma^2)^{-\dfrac{n_1+n_2}{2}}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n_1}(x_i-\mu_1)^2+\sum_{j=1}^{n_2}(y_j-\mu_2)^2\right]\right\}$$

取对数似然函数：

$$\ln L=-\frac{n_1+n_2}{2}\ln(2\pi)-\frac{n_1+n_2}{2}\ln(\sigma^2)-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n_1}(x_i-\mu_1)^2+\sum_{j=1}^{n_2}(y_j-\mu_2)^2\right]$$

2 均值的估计：

无约束时，对 μ_1, μ_2 求偏导，最大值点显然在 $\hat{\mu}_1 = \bar{x}$ 和 $\hat{\mu}_2 = \bar{y}$。

若 $\bar{y} \geq \bar{x}$：满足约束 $\mu_2 \geq \mu_1$。则 $\hat{\mu}_1 = \bar{x}, \hat{\mu}_2 = \bar{y}$。从而 $\hat{\mu} = \hat{\mu}_1 - \hat{\mu}_2 = \bar{x} - \bar{y}$。

若 $\bar{y} < \bar{x}$：不满足约束。函数值在边界 $\mu_1 = \mu_2$ 处达到最大，当我们在边界 $\mu_1 = \mu_2$ 寻找最值时，我们可以令 $\mu_1 = \mu_2 = \mu^*$（即假设两个样本来自同一个均值的总体）。此时，对数似然函数中与均值相关的部分变为：

$$Q(\mu^*) = -\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n_1}(x_i - \mu^*)^2 + \sum_{j=1}^{n_2}(y_j - \mu^*)^2\right]$$

对 μ^* 求导并令其为 0：

$$\frac{dQ}{d\mu^*} = \frac{1}{\sigma^2}\left[\sum_{i=1}^{n_1}(x_i - \mu^*) + \sum_{j=1}^{n_2}(y_j - \mu^*)\right] = 0$$

展开求和：

$$\left(\sum x_i - n_1\mu^*\right) + \left(\sum y_j - n_2\mu^*\right) = 0$$

$$(n_1 + n_2)\mu^* = \sum x_i + \sum y_j$$

由于 $\sum x_i = n_1\bar{x}$ 且 $\sum y_j = n_2\bar{y}$，解得：

$$\hat{\mu}^* = \frac{n_1\bar{x} + n_2\bar{y}}{n_1 + n_2}$$

这就是所谓的“合并均值”，即把所有数据看作一伙的，求总平均。从而 $\hat{\mu} = 0$。

3 方差的估计：无论 μ_1, μ_2 取什么值，我们对 σ^2 求偏导的逻辑是一样的。令 $S = \sum(x_i - \mu_1)^2 + \sum(y_j - \mu_2)^2$，对数似然函数简化为：

$$\ell = C - \frac{n_1 + n_2}{2}\ln(\sigma^2) - \frac{S}{2\sigma^2}$$

对 σ^2 求偏导：

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n_1 + n_2}{2\sigma^2} + \frac{S}{2(\sigma^2)^2} = 0$$

两边同乘 $2(\sigma^2)^2$：

$$-(n_1 + n_2)\sigma^2 + S = 0 \implies \hat{\sigma}^2 = \frac{S}{n_1 + n_2}$$

$\hat{\sigma}^2 = \dfrac{1}{n_1 + n_2}\left[\sum(x_i - \hat{\mu}_1)^2 + \sum(y_j - \hat{\mu}_2)^2\right]$。

总结：若 $\bar{x} \leq \bar{y}$，$\hat{\mu}_1 = \bar{x}, \hat{\mu} = \bar{x} - \bar{y}$；若 $\bar{x} > \bar{y}$，$\hat{\mu}_1 = \dfrac{n_1\bar{x} + n_2\bar{y}}{n_1 + n_2}, \hat{\mu} = 0$。

问题 17 (点估计应用题)**.** 某家具厂需设计家用门高度，要求人的头顶碰到上门框的机会不超过 0001。假设人的身高服从正态分布 $N(\mu, \sigma^2)$，$x_1, x_2, \cdots, x_n$ 是简单样本，试给出门框最低高度的估计？

【解析】

第一步：建立数学模型

设人的身高为随机变量 X，已知 $X \sim N(\mu, \sigma^2)$。设门框的高度为 h。题目要求 “头顶碰到上门框的机会不超过 0001” ，即身高 X 大于高度 h 的概率要小于等于 0001：$P(X > h) \leq 0001$ 为了求出 “最低” 高度，我们取临界情况，即令 $P(X > h) = 0001$。

第二步：利用正态分布标准化寻找 h 的表达式

由于 $X \sim N(\mu, \sigma^2)$，我们将其标准化：$\frac{X-\mu}{\sigma} \sim N(0,1)$ 于是概率式可以写为：$P\left(\frac{X-\mu}{\sigma} > \frac{h-\mu}{\sigma}\right) = 0001$ 根据标准正态分布的分位数定义，若 $Z \sim N(0,1)$，使得 $P(Z > z_\alpha) = \alpha$ 的 z_α 称为上侧 α 分位数。这里 $\alpha = 0001$，所以有：$\frac{h-\mu}{\sigma} = z_{0001}$ 解得门框高度 h 的理论值为：$h = \mu + z_{0001}\sigma$ 查标准正态分布表可知 $z_{0001} \approx 301$

第三步：给出参数的估计量

在实际问题中，总体的均值 μ 和标准差 σ 是未知的。我们手头有样本 $x_1, x_2, \cdots, x_n$，需要用样本指标去估计它们。

均值 μ 的估计：样本均值 $\bar{x}$ 是 μ 的无偏估计：$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^n x_i$

标准差 σ 的估计：样本标准差 s 是 σ 的无偏估计：$\hat{\sigma} = s = \sqrt{\frac{1}{n-1}\sum_{i=1}^n (x_i - \bar{x})^2}$

结论：根据参数估计的替换原理，将上述估计量代入 h 的表达式中，得到门框最低高度的估计量 $\hat{h}$ 为：$\hat{h} = \bar{x} + 301s$。

1.3 均方误差和无偏估计

均方误差（Mean Squared Error, MSE）定义为估计值与真实值之差的平方的期望值，

$$MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$$

偏差 - 方差分解公式：

$$\boxed{MSE(\hat{\theta}) = Var(\hat{\theta}) + [Bias(\hat{\theta})]^2 = Var(\hat{\theta}) + (E(\hat{\theta}) - \theta)^2},$$

其中 $Bias(\hat{\theta}) = E(\hat{\theta}) - \theta$ 是**偏差**。

设 $\theta \in \mathbb{R}$ 是总体的未知参数（例如：总体均值 μ、方差 σ^2、比例 p 等），$\hat{\theta} = g(X_1, X_2, \ldots, X_n)$ 是基于样本 $X_1, \ldots, X_n$ 构造的一个统计量。
若对所有可能的 θ 值，都有：

$$\boxed{\mathbb{E}_\theta[\hat{\theta}] = \theta}$$

则称 $\hat{\theta}$ 是 θ 的无偏估计量，下标 θ 表示期望是在参数真值为 θ 的分布下计算的。
经典例子：$\mathbb{E}[\bar{x} = \frac{1}{n}\sum x_i] = \mu$
$\mathbb{E}[S^2 = \frac{1}{n-1}\sum(x_i - \bar{x})^2] = \sigma^2$（用 n 代替 $n-1$ 则有偏）
无偏性不具有函数不变性：$\sqrt{S^2}$ 不是 σ 的无偏估计。

问题 18. 设总体 X 的数学期望 μ 和方差 σ^2 都存在，$x_1, x_2, \cdots, x_n$ 是来自总体 X 的简单样本，统计量 $\bar{x}$ 和 $a\bar{x}$ 是 μ 的两个估计，其中 $a(0 < a < 1)$ 是已知常数。

1. 求均方误差 $E_\theta(\bar{x} - \mu)^2$ 和 $E_\theta(a\bar{x} - \mu)^2$，其中 $\theta = (\mu, \sigma^2)$；
2. 问不等式 $E_\theta(a\bar{x} - \mu)^2 \leq E_\theta(\bar{x} - \mu)^2$ 是否对一切 $\theta \in \Theta = \{(\mu, \sigma^2) : -\infty < \mu < +\infty, \sigma^2 > 0\}$ 都成立？若不成立，给出使不等式成立的参数 θ 的范围。

【解析】

(1) 对于 $\bar{x}$：我们知道 $E(\bar{x}) = \mu$，$Var(\bar{x}) = \frac{\sigma^2}{n}$。推导如下：

$$Var(\bar{x}) = Var\left(\frac{1}{n}\sum_{i=1}^n x_i\right) = \frac{1}{n^2}Var\left(\sum_{i=1}^n x_i\right) = \frac{1}{n^2}\underbrace{(\sigma^2 + \sigma^2 + \cdots + \sigma^2)}_{n\text{ 个}} = \frac{1}{n^2}\cdot(n\sigma^2) = \frac{\sigma^2}{n}$$

所以，$E_\theta(\bar{x} - \mu)^2 = Var(\bar{x}) + [E(\bar{x}) - \mu]^2 = \frac{\sigma^2}{n}$。

对于 $a\bar{x}$：$E(a\bar{x}) = aE(\bar{x}) = a\mu$，$Bias(a\bar{x}) = E(a\bar{x}) - \mu = a\mu - \mu = (a-1)\mu$。

$Var(a\bar{x}) = a^2 Var(\bar{x}) = \frac{a^2\sigma^2}{n}$。

所以，$E_\theta(a\bar{x} - \mu)^2 = Var(a\bar{x}) + [Bias(a\bar{x})]^2 = \frac{a^2\sigma^2}{n} + (a-1)^2\mu^2$。

(2) 考察不等式：$\frac{a^2\sigma^2}{n} + (a-1)^2\mu^2 \leq \frac{\sigma^2}{n}$，移项得：$(a-1)^2\mu^2 \leq \frac{\sigma^2}{n} - \frac{a^2\sigma^2}{n} = \frac{(1-a^2)\sigma^2}{n}$
由于 $0 < a < 1$，则 $1 - a > 0$，且 $1 - a^2 = (1-a)(1+a)$。不等式两边同时除以 $(1-a)^2$：
$\mu^2 \leq \frac{(1-a)(1+a)\sigma^2}{n(1-a)^2} = \frac{(1+a)\sigma^2}{n(1-a)}$

由此可见，不等式并非对一切 μ 都成立。**成立范围：** $|\mu| \leq \sqrt{\frac{1+a}{n(1-a)}}\sigma$。

问题 19. 设 $x_1, x_2, \cdots, x_n$ 是来自正态总体 $N(\mu_0, \sigma^2)$ 的简单样本，其中 μ_0 是已知常数，σ^2 是未知参数。试在**均方误差标准下**，比较未知参数 σ^2 的三个估计量

$$S^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \bar{x})^2, \quad \hat{\sigma}_2^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \mu_0)^2, \quad \hat{\sigma}_3^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})^2$$

哪个较优？

【解析】

知识点拨： 若 $X \sim N(\mu, \sigma^2)$，则 $\frac{\sum(x_i-\mu)^2}{\sigma^2} \sim \chi^2(n)$，$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$。卡方分布 $\chi^2(k)$ 的期望为 k，方差为 $2k$。

1 对于 S^2：它是 σ^2 的无偏估计，偏差为 0。
$Var(S^2) = \frac{\sigma^4}{(n-1)^2}Var(\chi^2(n-1)) = \frac{\sigma^4}{(n-1)^2} \cdot 2(n-1) = \frac{2\sigma^4}{n-1}$。
故 $MSE(S^2) = \frac{2\sigma^4}{n-1}$。
2 对于 $\hat{\sigma}_2^2$：由于 μ_0 已知，$\frac{n\hat{\sigma}_2^2}{\sigma^2} \sim \chi^2(n)$。
$E(\hat{\sigma}_2^2) = \sigma^2$，偏差为 0
$Var(\hat{\sigma}_2^2) = \frac{\sigma^4}{n^2} \cdot 2n = \frac{2\sigma^4}{n}$。
故 $MSE(\hat{\sigma}_2^2) = \frac{2\sigma^4}{n}$。
3 对于 $\hat{\sigma}_3^2$：注意 $\hat{\sigma}_3^2 = \frac{n-1}{n}S^2$。
$E(\hat{\sigma}_3^2) = \frac{n-1}{n}\sigma^2$，
$Bias(\hat{\sigma}_3^2)^2 = (\frac{n-1}{n}\sigma^2 - \sigma^2)^2 = \frac{\sigma^4}{n^2}$。
$Var(\hat{\sigma}_3^2) = (\frac{n-1}{n})^2 Var(S^2) = \frac{(n-1)^2}{n^2}\frac{2\sigma^4}{n-1} = \frac{2(n-1)\sigma^4}{n^2}$。
$MSE(\hat{\sigma}_3^2) = \frac{2n-2}{n^2}\sigma^4 + \frac{1}{n^2}\sigma^4 = \frac{2n-1}{n^2}\sigma^4$。
比较： $MSE(S^2) = \frac{2}{n-1}\sigma^4 > MSE(\hat{\sigma}_2^2) = \frac{2}{n}\sigma^4$。
$MSE(\hat{\sigma}_2^2) - MSE(\hat{\sigma}_3^2) = (\frac{2}{n} - \frac{2n-1}{n^2})\sigma^4 = \frac{2n-(2n-1)}{n^2}\sigma^4 = \frac{1}{n^2}\sigma^4 > 0$。
因此，$MSE(\hat{\sigma}_3^2) < MSE(\hat{\sigma}_2^2) < MSE(S^2)$，$\hat{\sigma}_3^2$ 最优，这说明有时放弃无偏性可以换取更小的均方误差。

问题 20. *设总体 X 服从两点分布 $B(1,p)$，其中 $p(0<p<1)$ 是未知参数。*

1. *x_1 是来自总体 X 的简单样本，试问对 p 及 p^2 是否存在无偏估计？*
2. *$x_1, x_2, \cdots, x_n$ 是来自总体 X 的简单样本，求参数 p^2 的无偏估计 $(n \geq 2)$。*

【解析】

(1) 对于 p：令 $\hat{p} = x_1$，则 $E(x_1) = 1 \cdot p + 0 \cdot (1-p) = p$。所以 x_1 是 p 的无偏估计。

对于 p^2：设存在无偏估计 $g(x_1)$，则 $E[g(x_1)] = g(1)p + g(0)(1-p) = p^2$。整理得：$(g(1)-g(0))p + g(0) = p^2$。这是一个关于 p 的恒等式，但左边是 p 的一次多项式，右边是二次。多项式恒等定理告诉我们：如果两个多项式在某个区间内处处相等，那么它们对应的各项系数必须完全相同。故不存在 p^2 的无偏估计。

(2) 当 $n \geq 2$ 时，我们需要构造一个统计量使其期望为 p^2。

考虑样本均值 $\bar{x}$，我们知道 $E(\bar{x})=p$，$Var(\bar{x})=\dfrac{p(1-p)}{n}$。

则 $E(\bar{x}^2)=Var(\bar{x})+[E(\bar{x})]^2=\dfrac{p-p^2}{n}+p^2=\dfrac{n-1}{n}p^2+\dfrac{1}{n}p$。

这还不是 p^2。但我们注意到 $E(\bar{x})=p$，可以利用它消去 $\dfrac{1}{n}p$ 项。

令 $E[\bar{x}^2-\dfrac{1}{n}\bar{x}]=\dfrac{n-1}{n}p^2+\dfrac{1}{n}p-\dfrac{1}{n}p=\dfrac{n-1}{n}p^2$。

所以，$\hat{p^2}=\dfrac{n}{n-1}(\bar{x}^2-\dfrac{1}{n}\bar{x})=\dfrac{n\bar{x}^2-\bar{x}}{n-1}$ 是 p^2 的无偏估计。。

问题 21. 设 x_1,x_2,x_3 是来自总体 X 的简单样本（相互独立），试证明下列估计量

$$\hat{\mu}_1=\frac{1}{5}x_1+\frac{3}{10}x_2+\frac{1}{2}x_3,\quad \hat{\mu}_2=\frac{1}{3}x_1+\frac{1}{4}x_2+\frac{5}{12}x_3,\quad \hat{\mu}_3=\frac{1}{3}x_1+\frac{3}{4}x_2-\frac{1}{12}x_3$$

都是总体均值 $E(X)=\mu$ 的无偏估计，并求出每个估计的方差，问哪个估计较优？

【解析】

证明无偏性： 对于线性组合 $\sum c_ix_i$，其期望为 $\mu\sum c_i$。只需检查系数之和是否为 1。

1 $\hat{\mu}_1:\dfrac{1}{5}+\dfrac{3}{10}+\dfrac{1}{2}=\dfrac{2+3+5}{10}=1$。

2 $\hat{\mu}_2:\dfrac{1}{3}+\dfrac{1}{4}+\dfrac{5}{12}=\dfrac{4+3+5}{12}=1$。

3 $\hat{\mu}_3:\dfrac{1}{3}+\dfrac{3}{4}-\dfrac{1}{12}=\dfrac{4+9-1}{12}=1$。

所以，三者均为无偏估计。

计算方差： $Var(\sum c_ix_i)=\sigma^2\sum c_i^2$。

1 $Var(\hat{\mu}_1)=\sigma^2[(\dfrac{2}{10})^2+(\dfrac{3}{10})^2+(\dfrac{5}{10})^2]=\dfrac{4+9+25}{100}\sigma^2=038\sigma^2$。

2 $Var(\hat{\mu}_2)=\sigma^2[(\dfrac{4}{12})^2+(\dfrac{3}{12})^2+(\dfrac{5}{12})^2]=\dfrac{16+9+25}{144}\sigma^2=\dfrac{50}{144}\sigma^2\approx 0347\sigma^2$。

3 $Var(\hat{\mu}_3)=\sigma^2[(\dfrac{4}{12})^2+(\dfrac{9}{12})^2+(-\dfrac{1}{12})^2]=\dfrac{16+81+1}{144}\sigma^2=\dfrac{98}{144}\sigma^2\approx 0681\sigma^2$。

结论： 因为 $Var(\hat{\mu}_2)$ 最小，所以在无偏估计类中 $\hat{\mu}_2$ 最优。

问题 22. 设总体 $X\sim U[0,\theta](\theta>0)$，$x_1,x_2,\cdots,x_n$ 是来自总体 X 的简单样本，$x_{(n)}$ 是参数 θ 的极大似然估计，令 $\hat{\theta}(c)=cx_{(n)}$，其中 c 是常数。

1. 求均方误差 $MSE_\theta(\hat{\theta}(c))$；

2. 求常数 c，使得均方误差 $MSE_\theta(\hat{\theta}(c))$ 最小；

3. 记 (2) 中使 MSE 最小的估计为 $\hat{\theta}_1$，证明 $\hat{\theta}_1$ 是 θ 的有偏估计。这说明在均方误差标准下，θ 的有偏估计 $\hat{\theta}_1$ 要比其无偏估计 $\hat{\theta}_2=\dfrac{n+1}{n}x_{(n)}$ 优。

均匀分布 $U[0,\theta]$ 的最小次序统计量 $x_{(1)}$ 和最大次序统计量 $x_{(n)}$ 的期望分别为 $E(x_{(1)})=\dfrac{\theta}{n+1}$，$E(x_{(n)})=\dfrac{n}{n+1}\theta$。对于 $U[0,\theta]$ 分布的最大值 $x_{(n)}$，它的 k 阶原点矩（即 $x_{(n)}^k$ 的期望）公式为：$E(x_{(n)}^k)=\dfrac{n}{n+k}\theta^k$

直观理解：n 个点把 $[0,\theta]$ 分成 $n+1$ 段，平均每段长度就是 $\dfrac{\theta}{n+1}$。最小值的期望就是第一段的长度，最大值的期望是前 n 段的总长。

【解析】

(1) 已知 $E(x_{(n)}) = \frac{n}{n+1}\theta$，$E(x_{(n)}^2) = \frac{n}{n+2}\theta^2$。

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

$MSE(cx_{(n)}) = Var(cx_{(n)}) + [E(cx_{(n)}) - \theta]^2 = c^2\frac{n\theta^2}{(n+2)(n+1)^2} + (\frac{cn}{n+1} - 1)^2\theta^2$。

(2) 对 c 求导并令其为 0：

$$\frac{d}{dc}MSE = \theta^2\left[\frac{2cn}{(n+2)(n+1)^2} + 2(\frac{cn}{n+1} - 1)\frac{n}{n+1}\right] = 0$$

化简得：$\frac{c}{(n+2)(n+1)} + \frac{cn}{n+1} - 1 = 0 \implies c\left[\frac{1+n(n+2)}{(n+2)(n+1)}\right] = 1$

$c\frac{(n+1)^2}{(n+2)(n+1)} = 1 \implies c\frac{n+1}{n+2} = 1 \implies c = \frac{n+2}{n+1}$。

(3) $\hat{\theta}_1 = \frac{n+2}{n+1}x_{(n)}$。其期望 $E(\hat{\theta}_1) = \frac{n+2}{n+1}\cdot\frac{n}{n+1}\theta = \frac{n(n+2)}{(n+1)^2}\theta \neq \theta$，故是有偏的。

由于 $c = \frac{n+2}{n+1}$ 是使 MSE 最小的唯一解，而无偏估计 $\hat{\theta}_2$ 对应的 $c = \frac{n+1}{n}$，故 $\hat{\theta}_1$ 的 MSE 必然小于 $\hat{\theta}_2$ 的 MSE。

1.4 一致最小方差无偏估计 UMVUE

Rao-Blackwell 定理

定理内容： 设 T 是参数 θ 的充分统计量，$\hat{\theta}$ 是 θ 的一个无偏估计量。定义一个新的估计量：

$$\eta(T) = E[\hat{\theta}|T]$$

则有：

1. $\eta(T)$ 也是 θ 的无偏估计量，即 $E[\eta(T)] = \theta$；
2. $\eta(T)$ 的方差小于或等于 $\hat{\theta}$ 的方差，即 $Var(\eta(T)) \leq Var(\hat{\theta})$。

Lehmann-Scheffé 定理

定理内容：若统计量 $T(X)$ 是参数 θ 的完备充分统计量，且 $\varphi(X)$ 是 $q(\theta)$ 的一个无偏估计，则条件期望 $g(T)=E[\varphi(X)\mid T(X)]$ 是 $q(\theta)$ 的唯一的一致最小方差无偏估计（UMVUE）。

应用说明：寻找 UMVUE 的标准方案：

1 找 $q(\theta)$ 的完备充分统计量 $T(X)$，大多是情况就是找充分统计量，基本上不会出非完备情况。

2 找到一个 $T(X)$ 的函数 $\varphi(X)$ 是无偏估计，也就是 $E[\varphi(X)]=q(\theta)$

3 $\varphi(X)$ 就是 $q(\theta)$ 的一致最小方差无偏估计（UMVUE）。

注：指数族在自然参数空间内通常是完备的（如正态、泊松、伽马、二项、伯努利、均匀 U(0,θ) 的倒数形式等）。不需要每次都验证完备性，记住："常见单参数指数族 → 充分统计量自动完备"。

问题 23. 设 $x_1,x_2,\cdots,x_n$ 是来自 *Poisson* 总体 $P(\lambda)$ 的简单样本，λ 是正的未知参数。对于已知的正整数 k，试求 $q(\lambda)=\dfrac{\lambda^k}{k!}e^{-\lambda}$ 的一致最小方差无偏估计（*UMVUE*）。

【解析】

第一步：寻找参数 λ 的完备充分统计量

Poisson 分布概率质量函数为：

$$P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!},\quad k=0,1,2,\dots$$

样本 $x_1,x_2,\cdots,x_n$ 的联合分布为：

$$L(\lambda;x_1,\dots,x_n)=\prod_{i=1}^{n}\frac{\lambda^{x_i}e^{-\lambda}}{x_i!}=\frac{\lambda^{\sum_{i=1}^n x_i}e^{-n\lambda}}{\prod_{i=1}^n x_i!}$$

根据**因子分解定理**，联合分布可分解为：

$$L(\lambda;\mathbf{x})=\underbrace{\left(\lambda^{\sum x_i}e^{-n\lambda}\right)}_{g(T(\mathbf{x}),\lambda)}\cdot\underbrace{\left(\frac{1}{\prod x_i!}\right)}_{h(\mathbf{x})}$$

其中 $T(x)=\sum_{i=1}^n x_i$。因此，$T=\sum_{i=1}^n x_i$ 是参数 λ 的充分统计量。又因为 Poisson 分布属于指数族，且参数空间包含内点，故 $T=\sum_{i=1}^n x_i$ 是 λ 的**完备充分统计量**。

第二步：构造待估参数 $q(\lambda)$ 的无偏估计量

待估参数为 $q(\lambda)=\dfrac{\lambda^k}{k!}e^{-\lambda}$，这恰好是总体中单个随机变量取值为 k 的概率，即 $q(\lambda)=P(X_1=k)$。定义观测值 x_1 的函数 $u(x_1)$ 如下：

$$u(x_1)=I_{\{x_1=k\}}=\begin{cases}1, & x_1=k\\ 0, & x_1\neq k\end{cases}$$

验证其无偏性：$E[u(x_1)] = 1 \cdot P(x_1 = k) + 0 \cdot P(x_1 \neq k) = \dfrac{\lambda^k e^{-\lambda}}{k!} = q(\lambda)$

故 $u(x_1)$ 是 $q(\lambda)$ 的一个无偏估计量。

第三步：利用 Lehmann-Scheffé 定理求 UMVUE

根据 Lehmann-Scheffé 定理，由于 T 是完备充分统计量，且 $u(x_1)$ 是无偏估计量，则 $q(\lambda)$ 的 UMVUE 为：

$$\hat{q}(\lambda) = E[u(x_1)|T = t] = 1 \cdot P(u(x_1) = 1|T = t) + 0 \cdot P(u(x_1) = 0|T = t) = P(x_1 = k|\sum_{i=1}^{n} x_i = t)$$

到这一步都应该能写出来

第四步：详细计算条件概率

根据条件概率定义及样本的独立性：

$$P(x_1 = k|\sum_{i=1}^{n} x_i = t) = \frac{P(x_1 = k, \sum_{i=1}^{n} x_i = t)}{P(\sum_{i=1}^{n} x_i = t)} = \frac{P(x_1 = k, \sum_{i=2}^{n} x_i = t - k)}{P(\sum_{i=1}^{n} x_i = t)}$$

由于 x_i 独立同分布于 $P(\lambda)$，根据 Poisson 分布的可加性：$\sum_{i=2}^{n} x_i \sim P((n-1)\lambda)$，且 $\sum_{i=1}^{n} x_i \sim P(n\lambda)$。代入概率质量函数公式：

$$\hat{q} = \frac{\left(\dfrac{\lambda^k e^{-\lambda}}{k!}\right) \cdot \left(\dfrac{[(n-1)\lambda]^{t-k} e^{-(n-1)\lambda}}{(t-k)!}\right)}{\dfrac{(n\lambda)^t e^{-n\lambda}}{t!}}$$

消去分子分母中的 $e^{-n\lambda}$ 和 λ^t：

$$\hat{q} = \frac{t!}{k!(t-k)!} \cdot \frac{(n-1)^{t-k}}{n^t}$$

进一步整理：

$$\hat{q} = \binom{t}{k} \frac{(n-1)^{t-k}}{n^{t-k} \cdot n^k} = \binom{t}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{t-k}$$

结论： $q(\lambda)$ 的一致最小方差无偏估计为：

$$\hat{q}(\lambda) = \begin{cases} \dbinom{\sum_{i=1}^{n} x_i}{k} \left(\dfrac{1}{n}\right)^k \left(1 - \dfrac{1}{n}\right)^{\sum_{i=1}^{n} x_i - k}, & \sum x_i \geq k \\ 0, & \sum x_i < k \end{cases}$$

问题 24 (重要结论). 设 $x_1, x_2, \cdots, x_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的简单样本。*(1)* 当 σ^2 已知时，求 μ 的 *UMVUE*；*(2)* 当 μ 已知时，求 σ^2 的 *UMVUE*。

【解析】

(1) 当 σ^2 已知时，求 μ 的 UMVUE

第一步：写出样本的联合概率密度函数

总体密度为 $f(x; \mu) = \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\dfrac{(x-\mu)^2}{2\sigma^2}\right\}$。

样本的联合密度函数为：$L(\boldsymbol{x};\mu)=\prod_{i=1}^{n}f(x_i;\mu)=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}$

展开指数部分：$\sum(x_i-\mu)^2=\sum x_i^2-2\mu\sum x_i+n\mu^2$。代入整理得：

$$L(\boldsymbol{x};\mu)=\underbrace{(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{\sum x_i^2}{2\sigma^2}\right\}}_{h(\boldsymbol{x})}\cdot\underbrace{\exp\left\{-\frac{n\mu^2}{2\sigma^2}\right\}\exp\left\{\frac{\mu}{\sigma^2}\sum_{i=1}^{n}x_i\right\}}_{g(T(\boldsymbol{x}),\mu)}$$

第二步：寻找完备充分统计量

根据因子分解定理，统计量 $T=\sum_{i=1}^{n}x_i$ 是参数 μ 的充分统计量。

由于正态分布属于指数族分布，且**参数空间** $\Theta=(-\infty,+\infty)$ **包含内点**，因此，$T=\sum_{i=1}^{n}x_i$ 是 μ 的完备充分统计量。

第三步：构造无偏估计

我们需要寻找一个基于 T 的函数 $\phi(T)$，使其满足 $E[\phi(T)]=\mu$。考察统计量 $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{T}{n}$。由于 $x_i\sim N(\mu,\sigma^2)$，根据期望的线性性质：$E[\bar{x}]=E\left[\frac{1}{n}\sum_{i=1}^{n}x_i\right]=\frac{1}{n}\sum_{i=1}^{n}E[x_i]=\frac{1}{n}\cdot n\mu=\mu$，即 $\bar{x}$ 是 μ 的无偏估计。

第四步：利用 Lehmann-Scheffé 定理得出结论

因为 $\bar{x}$ 是完备充分统计量 T 的函数，且是 μ 的无偏估计。根据 Lehmann-Scheffé 定理，$\boxed{\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i}$ 是 μ 的一致最小方差无偏估计 (UMVUE)。

(2) 当 μ 已知时，求 σ^2 的 UMVUE

第一步：写出样本的联合概率密度函数

$L(\boldsymbol{x};\sigma^2)=(2\pi)^{-\frac{n}{2}}(\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}$

第二步：寻找完备充分统计量

根据因子分解定理，可以看出统计量 $S=\sum_{i=1}^{n}(x_i-\mu)^2$ 是 σ^2 的充分统计量。同样地，将密度函数写成指数族形式：$L(\boldsymbol{x};\sigma^2)=(2\pi)^{-\frac{n}{2}}\exp\left\{-\frac{n}{2}\ln(\sigma^2)-\frac{1}{2\sigma^2}S\right\}$ 这构成了关于参数 $\theta=-\frac{1}{2\sigma^2}$ 的指数族分布。由于**参数空间包含内点**，故统计量 $S=\sum_{i=1}^{n}(x_i-\mu)^2$ 是 σ^2 的完备充分统计量。

第三步：构造无偏估计

我们需要寻找一个基于 S 的函数，使其期望等于 σ^2。首先我们需要计算 S 的期望，根据结果考虑如何修正。

已知 $\frac{x_i-\mu}{\sigma}\sim N(0,1)$，则 $\left(\frac{x_i-\mu}{\sigma}\right)^2\sim\chi^2(1)$。根据卡方分布的可加性：$\frac{S}{\sigma^2}=\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2\sim\chi^2(n)$

利用卡方分布的期望性质 $E[\chi^2(n)]=n$，有：

$$E\left[\frac{S}{\sigma^2}\right]=n\implies\frac{1}{\sigma^2}E[S]=n\implies E[S]=n\sigma^2$$

为了得到 σ^2，我们将等式两边除以 n：$E\left[\frac{S}{n}\right]=\sigma^2$

因此，$\hat{\sigma}^2=\frac{S}{n}=\frac{1}{n}\sum_{i=1}^n(x_i-\mu)^2$ 是 σ^2 的无偏估计。

第四步：利用 Lehmann-Scheffé 定理得出结论

因为 $\boxed{\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2}$ 是完备充分统计量 S 的函数，且是 σ^2 的无偏估计。根据 Lehmann-Scheffé 定理，它是 σ^2 的 UMVUE。

问题 25. *设 $x_1,x_2,\cdots,x_n$ 是来自正态总体 $N(\mu,\sigma^2)$ 的简单样本，其中 μ,σ^2 是未知参数。若用 $\hat{\mu}^2=\bar{x}^2-\frac{1}{n}S^2$ 作为 μ^2 的估计，证明 $\hat{\mu}^2$ 是 μ^2 的 UMVUE。*

【解析】

1 **完备充分性**：在正态总体参数均未知时，$(\bar{x},S^2)$ 是 (μ,σ^2) 的完备充分统计量。$\hat{\mu}^2$ 是这两个统计量的函数。

由正态分布的指数族形式，联合密度可分解为：

$$f(x_1,...,x_n|\mu,\sigma^2)=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\exp\left(-\frac{\sum x_i^2}{2\sigma^2}+\frac{\mu}{\sigma^2}\sum x_i-\frac{n\mu^2}{2\sigma^2}\right)$$

μ,σ^2 是未知参数，根据因子分解定理，充分统计量为：

$$\left(\sum_{i=1}^{n}X_i,\sum_{i=1}^{n}X_i^2\right)$$

等价于常用形式：

$$(\bar{X},S^2)=\left(\frac{1}{n}\sum_{i=1}^{n}X_i,\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)$$

注：在正态总体下，$(\bar{x},S^2)$ 的充分完全性通常作为已知结论直接引用即可，不需要重新证明。

2 **无偏性**：我们需要证明 $E[\hat{\mu}^2]=\mu^2$。已知 $E[\bar{x}]=\mu$，$Var(\bar{x})=\frac{\sigma^2}{n}$。利用公式 $E[\bar{x}^2]=Var(\bar{x})+(E[\bar{x}])^2=\frac{\sigma^2}{n}+\mu^2$。又已知 $E[S^2]=\sigma^2$。则：

$$E[\hat{\mu}^2]=E\left[\bar{x}^2-\frac{1}{n}S^2\right]=E[\bar{x}^2]-\frac{1}{n}E[S^2]=\left(\mu^2+\frac{\sigma^2}{n}\right)-\frac{\sigma^2}{n}=\mu^2$$

3 **结论**：由于 $\hat{\mu}^2$ 是完备充分统计量的函数，且是 μ^2 的无偏估计，根据 Lehmann-Scheffé 定理，$\hat{\mu}^2$ 是 μ^2 的 UMVUE。

问题 26. *设总体 $X\sim B(1,p)$，$x_1,\ldots,x_n$ 是简单样本，求 p 的 UMVUE。*

【解析】

一、核心定理回顾 Lehmann-Scheffé 定理：若 T 是 θ 的充分完备统计量，且 $g(T)$ 是 θ 的无偏估计，则 $g(T)$ 是 θ 的 UMVUE。

1. 寻找充分统计量

$X \sim B(1,p)$，联合分布列：

$$L(x_1,\ldots,x_n;p) = p^{\sum x_i}(1-p)^{n-\sum x_i}$$

令 $T = \sum_{i=1}^n x_i$，则：

$$f = p^T(1-p)^{n-T} \cdot 1$$

故 T 是 p 的充分统计量。

3. 完备性验证

由于伯努利分布属于**指数族分布**，且**参数空间包含内点**，所以充分统计量 $T = \sum_{i=1}^n x_i$ 也是完备的。因此，$T = \sum_{i=1}^n x_i$ 是 p 的完备充分统计量。

4. 无偏估计构造

$E[T] = np \to E\left[\dfrac{T}{n}\right] = p$，即 $\dfrac{T}{n}$ 是无偏估计。

5. 由 Lehmann-Scheffé 定理，$\hat{p} = \dfrac{1}{n}\sum_{i=1}^n x_i$ 是 p 的 UMVUE。

问题 27 (了解)**.** *设 $x_1,\ldots,x_m \sim N(\mu_1,\sigma^2)$，$y_1,\ldots,y_n \sim N(\mu_2,\sigma^2)$，两样本独立，参数均未知。求 μ_1,μ_2,σ^2 的 UMVUE。*

【解析】

第一步：写出联合概率密度函数并整理

由于样本独立，联合概率密度函数（似然函数）为两组样本密度的乘积：

$$L(\mu_1,\mu_2,\sigma^2) = \left(\prod_{i=1}^m \frac{1}{\sqrt{2\pi}\sigma} e^{-\dfrac{(x_i-\mu_1)^2}{2\sigma^2}}\right) \cdot \left(\prod_{j=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\dfrac{(y_j-\mu_2)^2}{2\sigma^2}}\right)$$
$$= (2\pi\sigma^2)^{-\dfrac{m+n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^m (x_i-\mu_1)^2 + \sum_{j=1}^n (y_j-\mu_2)^2\right]\right\}$$

展开指数部分的平方项：

$\sum_{i=1}^m (x_i-\mu_1)^2 = \sum_{i=1}^m x_i^2 - 2\mu_1\sum_{i=1}^m x_i + m\mu_1^2$

$\sum_{j=1}^n (y_j-\mu_2)^2 = \sum_{j=1}^n y_j^2 - 2\mu_2\sum_{j=1}^n y_j + n\mu_2^2$

代回似然函数并整理成指数族形式：

$$L = C(\mu_1,\mu_2,\sigma^2)\exp\left\{\frac{\mu_1}{\sigma^2}\sum_{i=1}^m x_i + \frac{\mu_2}{\sigma^2}\sum_{j=1}^n y_j - \frac{1}{2\sigma^2}\left(\sum_{i=1}^m x_i^2 + \sum_{j=1}^n y_j^2\right)\right\}$$

其中 $C(\cdot)$ 是仅包含参数的项。

第二步：确定完备充分统计量

根据因子分解定理,联合充分统计量为:$T = \left(\sum_{i=1}^m x_i,\quad \sum_{j=1}^n y_j,\quad \sum_{i=1}^m x_i^2 + \sum_{j=1}^n y_j^2\right)$

由于正态分布属于指数族分布，且参数空间 $\Theta = \{(\mu_1,\mu_2,\sigma^2) : \mu_1 \in \mathbb{R}, \mu_2 \in \mathbb{R}, \sigma^2 > 0\}$ 包含内点，因此统计量 T 是完备充分统计量。

由于完备充分统计量是一一变换不变量，我们可以将 T 转换为更常用的形式。令：$\bar{x} = \frac{1}{m}\sum x_i$, $\bar{y} = \frac{1}{n}\sum y_j$, $S^2_{pooled} = \sum_{i=1}^m (x_i - \bar{x})^2 + \sum_{j=1}^n (y_j - \bar{y})^2$

利用恒等式 $\sum x_i^2 = \sum (x_i - \bar{x})^2 + m\bar{x}^2$，可知 $(\bar{x}, \bar{y}, S^2_{pooled})$ 与原统计量 T 是一一对应的。因此，$T^* = (\bar{x}, \bar{y}, S^2_{pooled})$ 也是完备充分统计量。

第三步：构造无偏估计并应用 Lehmann-Scheffé 定理

根据 Lehmann-Scheffé 定理，若 $\hat{g}$ 是参数 $g(\theta)$ 的无偏估计，且 $\hat{g}$ 是完备充分统计量的函数，则 $\hat{g}$ 是 $g(\theta)$ 的 UMVUE。

对于 μ_1**：**我们知道 $E[\bar{x}] = \mu_1$。且 $\bar{x}$ 是完备充分统计量 T^* 的函数。

$\therefore \hat{\mu}_1 = \bar{x}$ 是 μ_1 的 UMVUE。

对于 μ_2**：**同理，$E[\bar{y}] = \mu_2$。且 $\bar{y}$ 是完备充分统计量 T^* 的函数。

$\therefore \hat{\mu}_2 = \bar{y}$ 是 μ_2 的 UMVUE。

对于 σ^2**：**我们需要利用样本方差的性质。对于第一个样本：$\frac{1}{\sigma^2}\sum_{i=1}^m (x_i - \bar{x})^2 \sim \chi^2(m-1)$。对于第二个样本：$\frac{1}{\sigma^2}\sum_{j=1}^n (y_j - \bar{y})^2 \sim \chi^2(n-1)$。由于两样本独立，卡方分布具有可加性：$\frac{1}{\sigma^2}\left[\sum_{i=1}^m (x_i - \bar{x})^2 + \sum_{j=1}^n (y_j - \bar{y})^2\right] \sim \chi^2(m-1+n-1) = \chi^2(m+n-2)$ 已知卡方分布 $\chi^2(k)$ 的期望为 k，故：

$$E\left[\frac{S^2_{pooled}}{\sigma^2}\right] = m+n-2, \quad E\left[\frac{S^2_{pooled}}{m+n-2}\right] = \sigma^2$$

构造统计量 $\hat{\sigma}^2 = \frac{\sum_{i=1}^m (x_i - \bar{x})^2 + \sum_{j=1}^n (y_j - \bar{y})^2}{m+n-2}$。它是 σ^2 的无偏估计，且仅依赖于完备充分统计量 S^2_{pooled}。

$\therefore \hat{\sigma}^2 = \frac{\sum_{i=1}^m (x_i - \bar{x})^2 + \sum_{j=1}^n (y_j - \bar{y})^2}{m+n-2}$ 是 σ^2 的 UMVUE。

综上所述：

- μ_1 的 UMVUE 为 $\bar{x} = \frac{1}{m}\sum_{i=1}^m x_i$
- μ_2 的 UMVUE 为 $\bar{y} = \frac{1}{n}\sum_{j=1}^n y_j$
- σ^2 的 UMVUE 为 $S^2 = \frac{\sum_{i=1}^m (x_i - \bar{x})^2 + \sum_{j=1}^n (y_j - \bar{y})^2}{m+n-2}$

1.5 假设检验

1.5.1 基本概念

在假设检验问题 $H_0 : \theta \in \Theta_0, H_1 : \theta \in \Theta_1$ 中，我们需要根据样本 X 决定接受 H_0 还是拒绝 H_0。

定义 1.5.1 (检验函数与拒绝域). 检验通常由拒绝域 W 或检验函数 $\phi(x)$ 表示:

$$\phi(x) = \begin{cases} 1, & x \in W \quad (\text{拒绝}H_0) \\ 0, & x \notin W \quad (\text{接受}H_0) \end{cases}$$

定义 1.5.2 (势函数与两类错误). 势函数 $g(\theta)$ 定义为**拒绝原假设的概率**:

$$g(\theta) = E_\theta[\phi(X)] = P_\theta(X \in W)$$

- **第一类错误概率**(拒真):当 H_0 实际成立时,却错误地拒绝了 H_0

$$\alpha(\theta) = g(\theta), \quad \theta \in \Theta_0$$

- **显著性水平 α 为 H_0 为真时拒绝 H_0 的概率**:通常要求第一类错误概率最大不超过预设的显著性水平 α,$\sup_{\theta \in \Theta_0} g(\theta) \le \alpha$。

- **第二类错误概率**(取伪):当 H_1 实际成立时,却错误地接受了 H_0。

$$\beta(\theta) = 1 - g(\theta), \quad \theta \in \Theta_1$$

- **检验的势指当备择假设 H_1 为真时,正确拒绝原假设 H_0 的概率**。:

$$Power = g(\theta), \quad \theta \in \Theta_1$$

1.5.2 常见统计量

设 $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$。

1. σ^2 **已知,检验** μ:统计量 $Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1)$。

2. σ^2 **未知,检验** μ:统计量 $t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$。

3. μ **已知,检验** σ^2:统计量 $\chi^2 = \dfrac{\sum(X_i - \mu)^2}{\sigma_0^2} \sim \chi^2(n)$。

4. μ **未知,检验** σ^2:统计量 $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$。

1.6 似然比检验

1.6.1 定义与步骤

似然比检验是一种通用的构造检验的方法。sup 指‘上确界’，是所有上界中最小的那个。

定义 1.6.1 (似然比统计量)**.**

$$\lambda(x) = \frac{\sup_{\theta \in \Theta_0} L(\theta; x)}{\sup_{\theta \in \Theta} L(\theta; x)}$$

拒绝域形式为 $W = \{x : \lambda(x) \leq c\}$，其中 c 由显著性水平 α 确定：

$$\sup_{\theta \in \Theta_0} P_\theta(\lambda(X) \leq c) = \alpha$$

解题步骤：

1. 写出似然函数 $L(\theta)$。
2. 在全参数空间 Θ 上求极大似然估计 $\hat{\theta}$ 及最大值 $L(\hat{\theta})$。
3. 在原假设空间 Θ_0 上求受限极大似然估计 $\hat{\theta}_0$ 及最大值 $L(\hat{\theta}_0)$。
4. 计算 $\lambda(x)$ 并化简不等式 $\lambda(x) \leq c$。通常能化简为关于充分统计量的单调函数，从而确定等价的拒绝域形式。

1.6.2 一致最大势检验（UMPT）

引理 1 (Neyman-Pearson 引理)**.** 处理简单假设 $H_0 : \theta = \theta_0$ *vs* $H_1 : \theta = \theta_1$。拒绝域由似然比 $\dfrac{L(\theta_1)}{L(\theta_0)} > k$ 确定。

定义 1.6.2 (单调似然比分布族)**.** 设单参数分布族 $\{f(x|\theta), \theta \in \Theta\}$，若存在统计量 $T(X)$，使得对任意 $\theta_1 < \theta_2$，似然比 $\dfrac{f(x|\theta_2)}{f(x|\theta_1)}$ 是 $T(X)$ 的非递减函数（或非递增函数），则称该分布族对 $T(X)$ 具有单调似然比（*monotone likelihood ratio*，*MLR*）。

定理 1 (Karlin-Rubin 定理)**.** 条件：单参数分布族 $\{f(x|\theta)\}$ 对统计量 $T(X)$ 具有单调似然比；

检验单侧假设：$H_0 : \theta \leq \theta_0$ *vs* $H_1 : \theta > \theta_0$（右侧检验），或 $H_0 : \theta \geq \theta_0$ *vs* $H_1 : \theta < \theta_0$（左侧检验）。

结论：存在基于 $T(X)$ 的一致最优势检验（*UMPT*），其拒绝域为：

右侧检验：$T(X) > c$；左侧检验：$T(X) < c$。

其中，c 由显著性水平 α 确定，c 满足 $P_{\theta_0}(T(X) > c) = \alpha$（右侧）或 $P_{\theta_0}(T(X) < c) = \alpha$（左侧）。

1.6.3 一致最大势无偏检验 (UMPUT)

用于处理双边检验问题，如 $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$，或者区间假设。

无偏性

检验 ϕ 是无偏的，如果：

$$\sup_{\theta \in \Theta_0} g(\theta) \leq \alpha \leq \inf_{\theta \in \Theta_1} g(\theta)$$

即犯第一类错误的概率不超过 α，且犯第二类错误的概率不超过 $1-\alpha$（势不低于 α）。

指数族的 UMPUT

对于单参数指数族分布，考虑双边检验 $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$。UMPUT 的拒绝域形式通常为两边拒绝：

$$W = \{T(X) < C_1 \text{ 或} T(X) > C_2\}$$

确定常数 C_1, C_2 需联立以下两个方程：

1. 显著性水平条件

在原假设下，拒绝原假设的概率必须等于显著性水平 α：

$$P_{\theta_0}(T(X) < C_1) + P_{\theta_0}(T(X) > C_2) = \alpha$$

该条件保证了第一类错误率严格控制在 α 以内。

2. 无偏性条件

无偏性要求势函数 $g(\theta) = P_\theta(X \in W)$ 在 $\theta = \theta_0$ 处达到最小值 α，因此势函数在 θ_0 处的导数为 0：$g'(\theta_0) = 0$。对势函数求导可得：

$$\frac{d}{d\theta} \left[P_\theta(T < C_1) + P_\theta(T > C_2) \right] \bigg|_{\theta = \theta_0} = 0$$

结合指数族的得分函数性质，该条件等价于：

$$E_{\theta_0} \left[T(X) \cdot I(T < C_1) \right] + E_{\theta_0} \left[T(X) \cdot I(T > C_2) \right] = \alpha \cdot E_{\theta_0}[T(X)]$$

即拒绝域中统计量的加权均值等于原假设下统计量均值的 α 倍。

1.6.4 练习题

问题 28. 设总体 X 服从两点分布 $B(1,p)$，x_1, x_2, x_3 是来自总体 X 的简单样本，考虑假设检验问题 $H_0 : p = \frac{1}{2}, H_1 : p = \frac{3}{4}$。若取拒绝域

$$W = \{(x_1, x_2, x_3) : x_1 + x_2 + x_3 \geq 1\}$$

求势函数，以及犯第一类错误和第二类错误的概率及势。

【解析】

1 确定检验统计量及其分布

设 $T=\sum_{i=1}^{3} X_i$ 为样本中取值为 1 的个数。由于 X_1, X_2, X_3 独立同分布于 $B(1,p)$，根据二项分布的可加性，统计量 T 服从二项分布：

$$T \sim B(3,p)$$

其概率分布律为：

$$P(T=k)=\binom{3}{k}p^k(1-p)^{3-k}, \quad k=0,1,2,3$$

2 确定拒绝域

题目给出的拒绝域为 $W=\{(x_1,x_2,x_3):x_1+x_2+x_3\geq 1\}$。用统计量 T 表示，拒绝域即为：

$$W=\{T\geq 1\}$$

这意味着，只要样本中至少有一个 1，我们就拒绝原假设 H_0。

3 求势函数 $g(p)$

势函数 $g(p)$ 定义为在参数为 p 时，样本落入拒绝域（即拒绝 H_0）的概率：

$$g(p)=P_p(X\in W)=P_p(T\geq 1)$$

直接计算 $P(T\geq 1)$ 需要求 $P(T=1)+P(T=2)+P(T=3)$，利用对立事件计算更为简便：

$$g(p)=1-P_p(T<1)=1-P_p(T=0)$$

代入二项分布公式，当 $k=0$ 时：

$$P_p(T=0)=\binom{3}{0}p^0(1-p)^3=(1-p)^3$$

因此，势函数为：

$$g(p)=1-(1-p)^3, \quad 0<p<1$$

4 求犯第一类错误的概率 α

犯第一类错误是指：原假设 H_0 为真，但我们拒绝了它。即当 $p=\frac{1}{2}$ 时，落入拒绝域的概率：

$$\alpha=g\left(\frac{1}{2}\right)=1-\left(1-\frac{1}{2}\right)^3=1-\left(\frac{1}{2}\right)^3=1-\frac{1}{8}=\frac{7}{8}$$

5 求犯第二类错误的概率 β 及检验的势

犯第二类错误是指：备择假设 H_1 为真，但我们接受了 H_0。即当 $p=\frac{3}{4}$ 时，**未**落入拒绝域的概率：

$$\beta=P_{p=3/4}(X\notin W)=1-g\left(\frac{3}{4}\right)=\frac{1}{64}$$

检验的势是指在 H_1 为真时，正确拒绝 H_0 的概率：

$$势 = g\left(\frac{3}{4}\right) = \frac{63}{64}$$

综上所述:

- 势函数：$g(p) = 1-(1-p)^3$
- 犯第一类错误的概率：$\alpha = \dfrac{7}{8}$
- 犯第二类错误的概率：$\beta = \dfrac{1}{64}$
- 检验的势：$\dfrac{63}{64}$

问题 29. 设 $x_1, x_2, \ldots, x_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的简单样本，其中方差 σ^2 已知，考虑假设检验问题 $H_0 : \mu = \mu_0, H_1 : \mu = \mu_1 > \mu_0$。取拒绝域 $W = \{(x_1, x_2, \ldots, x_n) : \bar{x} \geq c\}$，其中 $\mu_0 < c < \mu_1$。

(1) 求犯第一类错误的概率 $\alpha(\mu_0)$；

(2) 求犯第二类错误的概率 $\beta(\mu_1)$；

(3) 证明：对固定的 n，当 $\alpha(\mu_0)$ 减小时，$\beta(\mu_1)$ 增大，而当 $\beta(\mu_1)$ 减小时，$\alpha(\mu_0)$ 增大；

(4) 证明：对固定的 c，当 n 充分大时，可使 $\alpha(\mu_0)$ 和 $\beta(\mu_1)$ 任意小。

【解析】

已知 $x_1, x_2, \ldots, x_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的简单样本，且 σ^2 已知。样本均值 $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ 服从正态分布：$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$，即标准化变量 $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$。记 $\Phi(x)$ 为标准正态分布的**分布函数**。

(1) 求犯第一类错误的概率 $\alpha(\mu_0)$

犯第一类错误是指：**原假设** H_0 **为真，但拒绝了** H_0。当 H_0 为真时，$\mu = \mu_0$，此时 $\bar{X} \sim N(\mu_0, \sigma^2/n)$。拒绝域为 $W = \{\bar{x} \geq c\}$。

$$\begin{aligned}
\alpha(\mu_0) &= P_{\mu_0}(\bar{X} \geq c) \\
&= P_{\mu_0}\left(\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \geq \frac{c-\mu_0}{\sigma/\sqrt{n}}\right) \\
&= P\left(Z \geq \frac{c-\mu_0}{\sigma/\sqrt{n}}\right) \\
&= 1 - \Phi\left(\frac{c-\mu_0}{\sigma/\sqrt{n}}\right)
\end{aligned}$$

(2) 求犯第二类错误的概率 $\beta(\mu_1)$

犯第二类错误是指：**备择假设 H_1 为真，但接受了** H_0。当 H_1 为真时，$\mu=\mu_1$，此时 $\bar{X}\sim N(\mu_1,\sigma^2/n)$。接受域为 $\bar{W}=\{\bar{x}<c\}$。

$$\begin{aligned}\beta(\mu_1)&=P_{\mu_1}(\bar{X}<c)\\&=P_{\mu_1}\left(\frac{\bar{X}-\mu_1}{\sigma/\sqrt{n}}<\frac{c-\mu_1}{\sigma/\sqrt{n}}\right)\\&=P\left(Z<\frac{c-\mu_1}{\sigma/\sqrt{n}}\right)\\&=\Phi\left(\frac{c-\mu_1}{\sigma/\sqrt{n}}\right)\end{aligned}$$

(3) 证明

对于固定的样本量 n 和已知参数 σ,μ_0,μ_1，错误概率 $\alpha(\mu_0)$ 和 $\beta(\mu_1)$ 仅依赖于临界值 c。由 (1) 和 (2) 的结果：$\alpha(c)=1-\Phi\left(\frac{c-\mu_0}{\sigma/\sqrt{n}}\right)$， $\beta(c)=\Phi\left(\frac{c-\mu_1}{\sigma/\sqrt{n}}\right)$

由于标准正态分布函数 $\Phi(x)$ 是单调严格递增函数：

- 对于 $\alpha(c)$：当 c 增大时，$\frac{c-\mu_0}{\sigma/\sqrt{n}}$ 增大，$\Phi(\cdot)$ 增大，故 $1-\Phi(\cdot)$ 减小。即 $\alpha(c)$ 关于 c 单调递减。
- 对于 $\beta(c)$：当 c 增大时，$\frac{c-\mu_1}{\sigma/\sqrt{n}}$ 增大，$\Phi(\cdot)$ 增大。即 $\beta(c)$ 关于 c 单调递增。

结论：当 $\alpha(\mu_0)$ 减小时，意味着临界值 c 必须增大；而 c 的增大必然导致 $\beta(\mu_1)$ 增大。反之，当 $\beta(\mu_1)$ 减小时，意味着临界值 c 必须减小；而 c 的减小必然导致 $\alpha(\mu_0)$ 增大。证毕。

(4) 证明：对固定的 c，当 n 充分大时，可使 $\alpha(\mu_0)$ 和 $\beta(\mu_1)$ 任意小

题目给定条件 $\mu_0<c<\mu_1$。

- **对于 $\alpha(\mu_0)$：**考察极限 $\lim_{n\to\infty}\alpha(\mu_0)$。令 $u_n=\frac{c-\mu_0}{\sigma/\sqrt{n}}=\frac{\sqrt{n}(c-\mu_0)}{\sigma}$。因为 $c>\mu_0$，所以 $c-\mu_0>0$。当 $n\to\infty$ 时，$u_n\to+\infty$。

  故：$\lim_{n\to\infty}\alpha(\mu_0)=\lim_{n\to\infty}[1-\Phi(u_n)]=1-\Phi(+\infty)=1-1=0$

- **对于 $\beta(\mu_1)$：**考察极限 $\lim_{n\to\infty}\beta(\mu_1)$。令 $v_n=\frac{c-\mu_1}{\sigma/\sqrt{n}}=\frac{\sqrt{n}(c-\mu_1)}{\sigma}$。因为 $c<\mu_1$，所以 $c-\mu_1<0$。当 $n\to\infty$ 时，$v_n\to-\infty$。

  故：$\lim_{n\to\infty}\beta(\mu_1)=\lim_{n\to\infty}\Phi(v_n)=\Phi(-\infty)=0$

综上所述，当 $n\to\infty$ 时，$\alpha(\mu_0)\to 0$ 且 $\beta(\mu_1)\to 0$。即当样本量 n 充分大时，两类错误的概率均可任意小。证毕。

问题 30 (单边检验). 某厂生产某种元件，其使用寿命不低于 1000 h 才合格。现从这批元件中随机抽取 25 件，其寿命均值 950 h。已知元件寿命服从标准差为 100 h 的正态分布。在显著性水平 $\alpha=0.05$ 下，问这批元件是否合格？

【解析】

1 建立假设

根据题意，元件合格的标准是寿命不低于 1000 h。我们要检验这批元件是否合格，通常将"合格"（即我们希望维持的现状或标准）设为原假设，或者将"不合格"（我们需要发现的问题）设为备择假设。设元件寿命 $X \sim N(\mu, \sigma^2)$，其中 $\sigma = 100$。建立单侧假设检验：$H_0 : \mu \geq \mu_0 \quad \text{vs} \quad H_1 : \mu < \mu_0$，$\mu_0 = 1000$

2 选取检验统计量

由于总体方差 σ^2 已知，应选用 Z 检验。在 H_0 成立的边界条件（$\mu = 1000$）下，统计量为：$Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1)$

3 确定拒绝域

这是一个左侧检验（因为 H_1 是 $\mu < 1000$）。对于显著性水平 $\alpha = 0.05$，查标准正态分布表可得临界值 $z_\alpha = z_{0.05} = 1.645$。拒绝域为：$W = \{Z < -z_{0.05}\} = \{Z < -1.645\}$ 即：如果计算出的 Z 值小于 -1.645，则拒绝原假设。

4 计算统计量的值

已知 $n = 25$，$\bar{x} = 950$，$\mu_0 = 1000$，$\sigma = 100$。代入公式计算：$Z = \dfrac{950 - 1000}{100/\sqrt{25}} = \dfrac{-50}{100/5} = \dfrac{-50}{20} = -2.5 < -1.645$，统计量落入拒绝域 W。所以在显著性水平 $\alpha = 0.05$ 下，拒绝原假设 H_0。

结论：有充分的理由认为这批元件的平均寿命显著低于 1000 h，即认为这批元件不合格。

问题 31 (双边检验)**.** *设总体 $X \sim N(\mu, \sigma_0^2)$，其中 σ_0^2 已知，μ 为未知参数，$x_1, x_2, \ldots, x_n$ 为来自总体 X 的简单样本，考虑假设检验问题*

$$H_0 : \mu = \mu_0, \quad H_1 : \mu \neq \mu_0$$

在显著性水平 α 下，试求似然比检验。

【解析】

第一步：写出似然函数

设总体 $X \sim N(\mu, \sigma_0^2)$，其中 σ_0^2 已知。样本 $x_1, x_2, \ldots, x_n$ 的似然函数为：

$$L(\mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left\{-\frac{(x_i - \mu)^2}{2\sigma_0^2}\right\} = (2\pi\sigma_0^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i - \mu)^2\right\}$$

参数空间为 $\Omega = \{\mu \mid -\infty < \mu < +\infty\}$。原假设对应的参数空间为 $\Omega_0 = \{\mu_0\}$

第二步：求解全参数空间的极大似然估计

在全参数空间 Ω 中，我们需要找到使 $L(\mu)$ 达到最大的 $\hat{\mu}$。观察似然函数 $L(\mu)$ 的表达式，要使 $L(\mu)$ 最大，等价于使指数部分的 $\sum_{i=1}^{n}(x_i - \mu)^2$ 最小。直接求导，我们知道当 μ 取样本均值时，该式取最小值。即全参数空间的极大似然估计为：

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

此时似然函数的最大值为：

$$\sup_{\mu\in\Omega} L(\mu) = L(\bar{x}) = (2\pi\sigma_0^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\bar{x})^2\right\}$$

第三步：求解 H_0 成立时的极大似然估计

在原假设 $H_0: \mu=\mu_0$ 成立的条件下，参数空间 Ω_0 只有唯一的值 μ_0。因此，受限极大似然估计就是直接代入 μ_0：

$$\sup_{\mu\in\Omega_0} L(\mu) = L(\mu_0) = (2\pi\sigma_0^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\mu_0)^2\right\}$$

第四步：构造似然比统计量 λ

似然比统计量定义为 $\lambda = \dfrac{\sup_{\mu\in\Omega_0} L(\mu)}{\sup_{\mu\in\Omega} L(\mu)}$。代入上述结果：

$$\lambda = \frac{(2\pi\sigma_0^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\mu_0)^2\right\}}{(2\pi\sigma_0^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\bar{x})^2\right\}}$$

化简表达式：

$$\lambda = \exp\left\{-\frac{1}{2\sigma_0^2}\left[\sum_{i=1}^{n}(x_i-\mu_0)^2 - \sum_{i=1}^{n}(x_i-\bar{x})^2\right]\right\}$$

利用恒等式 $\boxed{\sum_{i=1}^{n}(x_i-\mu_0)^2 = \sum_{i=1}^{n}(x_i-\bar{x}+\bar{x}-\mu_0)^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu_0)^2}$（交叉项为 0），指数部分的方括号内可以化简为：

$$\sum_{i=1}^{n}(x_i-\mu_0)^2 - \sum_{i=1}^{n}(x_i-\bar{x})^2 = n(\bar{x}-\mu_0)^2$$

因此，似然比统计量为：

$$\lambda = \exp\left\{-\frac{n(\bar{x}-\mu_0)^2}{2\sigma_0^2}\right\}$$

第五步：确定拒绝域

似然比检验法则为：当 $\lambda \le c$ 时拒绝 H_0，其中 c $(0<c<1)$ 是由显著性水平 α 确定的临界值。不等式 $\lambda \le c$ 等价于：

$$\exp\left\{-\frac{n(\bar{x}-\mu_0)^2}{2\sigma_0^2}\right\} \le c$$

两边取自然对数：

$$-\frac{n(\bar{x}-\mu_0)^2}{2\sigma_0^2} \le \ln c$$

整理得：

$$\frac{n(\bar{x}-\mu_0)^2}{\sigma_0^2} \ge -2\ln c$$

两边开方（注意绝对值）：

$$\left|\frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma_0}\right| \geq \sqrt{-2\ln c}$$

令 $k=\sqrt{-2\ln c}$，则拒绝域形式为：

$$\left|\frac{\bar{x}-\mu_0}{\sigma_0/\sqrt{n}}\right| \geq k$$

第六步：利用统计量的分布确定临界值

在 H_0 成立的条件下，统计量 $Z=\dfrac{\bar{x}-\mu_0}{\sigma_0/\sqrt{n}}$ 服从标准正态分布 $N(0,1)$。根据给定的显著性水平 α，我们需要满足：

$$P_{H_0}(\text{拒绝}H_0)=P\left(\left|\frac{\bar{x}-\mu_0}{\sigma_0/\sqrt{n}}\right| \geq k\right)=\alpha$$

由于 $Z\sim N(0,1)$，上述概率即为标准正态分布的双侧尾部概率。因此，临界值 k 应为标准正态分布的上侧 $\alpha/2$ 分位数 $z_{\alpha/2}$（即满足 $P(Z>z_{\alpha/2})=\alpha/2$）。

结论

该问题的似然比检验的拒绝域为：

$$W=\left\{(x_1,\ldots,x_n):\left|\frac{\bar{x}-\mu_0}{\sigma_0/\sqrt{n}}\right| \geq z_{\alpha/2}\right\}$$

其中 $z_{\alpha/2}$ 是标准正态分布的上侧 $\alpha/2$ 分位数。

问题 32 (了解)**.** *设 $x_1,x_2,\ldots,x_n$ 是来自正态总体 $N(\mu_0,\sigma^2)$ 的简单样本，其中 μ_0 已知，σ^2 是未知参数，对给定的水平 α，试分别求下列假设检验问题的一致最大势检验 UMPT:*

(1) $H_0:\sigma^2\geq\sigma_0^2,\quad H_1:\sigma^2<\sigma_0^2$;

(2) $H_0:\sigma^2\leq\sigma_1^2$ 或$\sigma^2\geq\sigma_2^2,\quad H_1:\sigma_1^2<\sigma^2<\sigma_2^2$。

【解析】

一、一致最大势检验 (UMPT) 的定义

设样本 X 的分布族为 $\{f(x;\theta):\theta\in\Theta\}$，考虑假设检验问题：

$$H_0:\theta\in\Theta_0 \quad \text{vs} \quad H_1:\theta\in\Theta_1$$

若存在一个检验 $\phi^*(x)$，满足以下两个条件：

1 水平条件：对任意 $\theta\in\Theta_0$，有 $E_\theta[\phi^*(X)]\leq\alpha$;

2 最大势条件：对于任意其他满足水平 α 的检验 $\phi(x)$，以及任意 $\theta\in\Theta_1$，都有

$$E_\theta[\phi^*(X)]\geq E_\theta[\phi(X)]$$

则称 $\phi^*(x)$ 为该假设检验问题在显著性水平 α 下的一致最大势检验 (UMPT)。

二、准备工作：写出似然函数与充分统计量

样本 $x_1, \ldots, x_n$ 来自 $N(\mu_0, \sigma^2)$，且 μ_0 已知。样本的联合概率密度函数为：

$$
\begin{aligned}
L(\sigma^2; x) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x_i-\mu_0)^2}{2\sigma^2}\right\} \\
&= (2\pi)^{-\dfrac{n}{2}} (\sigma^2)^{-\dfrac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_0)^2\right\}
\end{aligned}
$$

这是属于单参数指数族分布。令 $T(x) = \sum_{i=1}^{n}(x_i-\mu_0)^2$，$\theta = \sigma^2$。显然 $T(x)$ 是 σ^2 的充分统计量。我们知道 $\dfrac{x_i-\mu_0}{\sigma} \sim N(0,1)$，故

$$
\frac{T(x)}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{x_i-\mu_0}{\sigma}\right)^2 \sim \chi^2(n)
$$

三、检验问题 (1) 的求解

1 问题描述

$$
H_0 : \sigma^2 \geq \sigma_0^2 \quad \text{vs} \quad H_1 : \sigma^2 < \sigma_0^2
$$

2 验证单调似然比 (MLR) 性质

考虑似然函数关于统计量 $T(x)$ 的结构。对于任意 $\sigma_1^2 < \sigma_2^2$，考察似然比：

$$
\frac{L(\sigma_2^2; x)}{L(\sigma_1^2; x)} = \frac{(\sigma_2^2)^{-n/2}\exp\left\{-\dfrac{1}{2\sigma_2^2}T(x)\right\}}{(\sigma_1^2)^{-n/2}\exp\left\{-\dfrac{1}{2\sigma_1^2}T(x)\right\}} = \left(\frac{\sigma_2^2}{\sigma_1^2}\right)^{-n/2}\exp\left\{\left(\frac{1}{2\sigma_1^2}-\frac{1}{2\sigma_2^2}\right)T(x)\right\}
$$

由于 $\sigma_1^2 < \sigma_2^2$，则 $\dfrac{1}{2\sigma_1^2}-\dfrac{1}{2\sigma_2^2} > 0$。因此，该似然比是统计量 $T(x)$ 的严格单调递增函数。即分布族 $\{P_{\sigma^2}\}$ 具有关于统计量 $T(x)$ 的单调似然比性质（MLR）。

3 利用 Karlin-Rubin 定理构造 UMPT

根据 Karlin-Rubin 定理，对于单边检验问题 $H_0 : \theta \geq \theta_0$ vs $H_1 : \theta < \theta_0$，若似然比关于 $T(x)$ 单调递增，则 UMPT 的拒绝域形式为 $T(x) \leq C$。

具体地，检验函数为：

$$
\phi(x) = \begin{cases} 1, & T(x) < C \\ \gamma, & T(x) = C \\ 0, & T(x) > C \end{cases}
$$

由于 $T(x)$ 是连续型随机变量，边界处概率为 0，可取 $\gamma = 0$。拒绝域形式为 $W = \{x : T(x) \leq C\}$。

4 确定临界值 C

临界值 C 由显著性水平 α 在边界点 σ_0^2 处确定：

$$
P_{\sigma_0^2}(T(x) \leq C) = \alpha
$$

在 H_0 的边界 $\sigma^2=\sigma_0^2$ 处，$\dfrac{T(x)}{\sigma_0^2}\sim\chi^2(n)$。因此：

$$P\left(\frac{T(x)}{\sigma_0^2}\le\frac{C}{\sigma_0^2}\right)=\alpha$$

这意味着 $\dfrac{C}{\sigma_0^2}$ 应当是 $\chi^2(n)$ 分布的下侧 α 分位数。记 $\chi^2_\alpha(n)$ 为 $\chi^2(n)$ 的下侧 α 分位数（即累积概率为 α 的点）。则 $\dfrac{C}{\sigma_0^2}=\chi^2_\alpha(n)\implies C=\sigma_0^2\chi^2_\alpha(n)$。

5 结论 (1)

该假设检验问题的一致最大势检验（UMPT）拒绝域为：

$$W=\left\{(x_1,\ldots,x_n):\sum_{i=1}^n(x_i-\mu_0)^2\le\sigma_0^2\chi^2_\alpha(n)\right\}$$

四、检验问题 (2) 的求解

1 问题描述

$$H_0:\sigma^2\le\sigma_1^2\text{ 或}\sigma^2\ge\sigma_2^2\quad\text{vs}\quad H_1:\sigma_1^2<\sigma^2<\sigma_2^2$$

其中 $\sigma_1^2<\sigma_2^2$。这是一个双边假设检验问题（中间区域为备择假设）。

2 验证 UMPT 存在的条件

这是关于单参数指数族分布的双边检验。指数族分布密度形式为 $f(x;\theta)=C(\theta)e^{Q(\theta)T(x)}h(x)$。在本题中，令 $\theta=-\dfrac{1}{2\sigma^2}$（自然参数），则 $Q(\theta)=\theta$，$T(x)=\sum(x_i-\mu_0)^2$。由于 $Q(\theta)$ 是严格单调递增函数，且分布族具有 MLR 性质。对于形如 $H_0:\theta\le\theta_1$ 或$\theta\ge\theta_2$ vs $H_1:\theta_1<\theta<\theta_2$ 的检验问题，存在 UMPT，其拒绝域形式为：

$$C_1\le T(x)\le C_2$$

3 构造 UMPT

检验函数的形式为：

$$\phi(x)=\begin{cases}1, & C_1\le T(x)\le C_2\\ 0, & \text{其他}\end{cases}$$

即当统计量 $T(x)$ 落在区间 $[C_1,C_2]$ 内时拒绝 H_0。

4 确定临界值 C_1,C_2

临界值 C_1,C_2 由水平 α 在两个边界点 σ_1^2 和 σ_2^2 处确定：

$$\begin{cases}E_{\sigma_1^2}[\phi(X)]=P_{\sigma_1^2}(C_1\le T(X)\le C_2)=\alpha\\ E_{\sigma_2^2}[\phi(X)]=P_{\sigma_2^2}(C_1\le T(X)\le C_2)=\alpha\end{cases}$$

利用 $Y=\dfrac{T(X)}{\sigma^2}\sim\chi^2(n)$ 进行标准化：在 $\sigma^2=\sigma_1^2$ 时，概率为：

$$P\left(\frac{C_1}{\sigma_1^2}\le\frac{T(X)}{\sigma_1^2}\le\frac{C_2}{\sigma_1^2}\right)=\alpha$$

令 $F_{\chi^2(n)}(y)$ 为自由度为 n 的卡方分布的累积分布函数，则方程组为：

$$\begin{cases} F_{\chi^2(n)}\left(\dfrac{C_2}{\sigma_1^2}\right) - F_{\chi^2(n)}\left(\dfrac{C_1}{\sigma_1^2}\right) = \alpha \\ F_{\chi^2(n)}\left(\dfrac{C_2}{\sigma_2^2}\right) - F_{\chi^2(n)}\left(\dfrac{C_1}{\sigma_2^2}\right) = \alpha \end{cases}$$

5 结论 (2)

该假设检验问题的一致最大势检验（UMPT）拒绝域为：

$$W = \left\{(x_1, \dots, x_n) : C_1 \le \sum_{i=1}^{n}(x_i - \mu_0)^2 \le C_2\right\}$$

其中常数 C_1, C_2 由方程组 $\begin{cases} P_{\sigma_1^2}(C_1 \le T \le C_2) = \alpha \\ P_{\sigma_2^2}(C_1 \le T \le C_2) = \alpha \end{cases}$ 唯一确定。

问题 33 (了解). 设总体 $X \sim B(1,p), 0 < p < 1$，$x_1, x_2, \dots, x_n$ 为来自总体 X 的简单样本。对给定的水平 α，试求假设检验问题 $H_0 : p = p_0, H_1 : p \neq p_0$ 的一致最大势无偏检验 $UMPUT$。

【解析】

一、一致最大势无偏检验 (UMPUT) 的定义

对于参数 θ 的假设检验问题 $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_1$，若检验 ϕ^* 满足以下条件：

1 无偏性：

$$\begin{cases} E_\theta[\phi^*(X)] \le \alpha, & \forall\theta \in \Theta_0 \\ E_\theta[\phi^*(X)] \ge \alpha, & \forall\theta \in \Theta_1 \end{cases}$$

2 一致最大势：在所有满足无偏性的检验类 $\mathcal{C}_\alpha$ 中，ϕ^* 具有最大的势，即对于任意 $\phi \in \mathcal{C}_\alpha$，有：

$$E_\theta[\phi^*(X)] \ge E_\theta[\phi(X)], \quad \forall\theta \in \Theta_1$$

则称 ϕ^* 为水平 α 的一致最大势无偏检验 (UMPUT)。

二、识别指数族分布与充分统计量

总体 $X \sim B(1,p)$，其概率分布律为 $P(X = x) = p^x(1-p)^{1-x}$ $(x = 0, 1)$。样本 $x_1, \dots, x_n$ 的联合分布律为：

$$\begin{aligned} L(p; x) &= \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\ &= p^{\sum x_i}(1-p)^{n-\sum x_i} \\ &= (1-p)^n \left(\frac{p}{1-p}\right)^{\sum x_i} \\ &= (1-p)^n \exp\left\{\left(\ln\frac{p}{1-p}\right)\sum_{i=1}^{n} x_i\right\} \end{aligned}$$

令 $\theta = \ln \frac{p}{1-p}$（自然参数），$T(x) = \sum_{i=1}^n x_i$。由于 θ 是 p 的严格单调递增函数，且 $L(p;x)$ 具有单参数指数族的形式，故 $T(x) = \sum_{i=1}^n x_i$ 是 p 的充分统计量。统计量 T 服从二项分布 $B(n,p)$。

三、构造 UMPUT

对于单参数指数族分布的双侧检验问题：

$$H_0 : p = p_0 \quad \text{vs} \quad H_1 : p \neq p_0$$

根据 UMPUT 的构造定理，检验函数 $\phi(t)$ 应当具有如下形式（“双边拒绝，中间接受”）：

$$\phi(t) = \begin{cases} 1, & t < C_1 \text{ 或} t > C_2 \\ \gamma_1, & t = C_1 \\ \gamma_2, & t = C_2 \\ 0, & C_1 < t < C_2 \end{cases}$$

其中 C_1, C_2 是整数临界值，$0 \leq \gamma_1, \gamma_2 \leq 1$ 是边界随机化概率。

四、确定常数的方程组

常数 $C_1, C_2, \gamma_1, \gamma_2$ 由无偏性的边界条件（即在 $p = p_0$ 处的性质）确定。根据定理，UMPUT 必须满足以下两个条件（类似于在 p_0 处的一阶导数和函数值条件）：

1 水平条件：在 H_0 成立时，拒绝概率为 α。

$$E_{p_0}[\phi(T)] = \alpha$$

2 无偏性引出的导数条件：在 p_0 处势函数的导数为 0（意味着势函数在 p_0 处取最小值 α）。

$$E_{p_0}[T\phi(T)] = \alpha E_{p_0}[T] = np_0\alpha$$

将 $\phi(t)$ 的具体形式代入上述两个条件，得到方程组：

$$\begin{cases} \sum_{k=0}^n \phi(k) P_{p_0}(T = k) = \alpha \\ \sum_{k=0}^n k\phi(k) P_{p_0}(T = k) = np_0\alpha \end{cases}$$

展开为具体表达式（记 $b(k;n,p_0) = \binom{n}{k} p_0^k (1-p_0)^{n-k}$）：

$$\begin{cases} P(T < C_1) + \gamma_1 P(T = C_1) + \gamma_2 P(T = C_2) + P(T > C_2) = \alpha \\ \sum_{t \in R} t \cdot b(t;n,p_0) + C_1\gamma_1 b(C_1;n,p_0) + C_2\gamma_2 b(C_2;n,p_0) = np_0\alpha \end{cases}$$

其中第一式中的概率 P 均基于 $B(n,p_0)$ 计算，R 表示确定的拒绝域 $\{t < C_1\} \cup \{t > C_2\}$。

五、结论

该假设检验问题的一致最大势无偏检验（UMPUT）为：

$$\phi(x)=\begin{cases}1, & \sum x_i<C_1 \text{ 或 } \sum x_i>C_2\\ \gamma_1, & \sum x_i=C_1\\ \gamma_2, & \sum x_i=C_2\\ 0, & C_1<\sum x_i<C_2\end{cases}$$

其中常数 $C_1,C_2,\gamma_1,\gamma_2$ 由方程组

$$\begin{cases}E_{p_0}[\phi(T)]=\alpha\\ E_{p_0}[T\phi(T)]=np_0\alpha\end{cases}$$

唯一确定。

问题 34 (了解)**.** 设总体 $X\sim B(1,p)$ $(0<p<1)$，$x_1,x_2,\ldots,x_n$ 为来自总体 X 的简单样本。*(1)* 对假设检验问题：$H_0:p\le 01 \quad vs \quad H_1:p>01$，给定显著性水平 $\alpha=0.05$，构造该问题的一致最优势检验（*UMPT*）；*(2)* 当 $n=10$ 时，给出具体的检验函数；*(3)* 根据 *(2)* 中的检验函数，判断以下两种情况是否拒绝原假设：*(a)* $\sum_{i=1}^{10}x_i\ge 2$；*(b)* $\sum_{i=1}^{10}x_i<2$。

【解析】

一、构建一致最优势检验 (UMPT) 的理论推导

1 识别指数族分布与充分统计量

设总体 $X\sim B(1,p)$，样本 $x_1,\ldots,x_n$ 的联合概率函数（似然函数）为：

$$\begin{aligned}L(p;x)&=\prod_{i=1}^{n}p^{x_i}(1-p)^{1-x_i}\\&=p^{\sum x_i}(1-p)^{n-\sum x_i}\\&=(1-p)^n\left(\frac{p}{1-p}\right)^{\sum x_i}\\&=(1-p)^n\exp\left\{\ln\left(\frac{p}{1-p}\right)\sum_{i=1}^{n}x_i\right\}\end{aligned}$$

这是一个单参数指数族分布。令 $T(x)=\sum_{i=1}^{n}x_i$，它是参数 p 的充分统计量。令 $Q(p)=\ln\left(\dfrac{p}{1-p}\right)$，因为 $Q(p)$ 是关于 p 的严格单调递增函数，所以该分布族关于统计量 $T(x)$ 具有单调似然比（MLR）性质。

2 根据 Karlin-Rubin 定理构造 UMPT

假设检验问题为 $H_0:p\le 01 \quad \text{vs} \quad H_1:p>01$。根据 Karlin-Rubin 定理，对于 $H_1:p>p_0$ 的单边检验，存在一致最优势检验（UMPT），其拒绝域形式为 $T(x)>C$。

由于 $T(x) \sim B(n,p)$ 是离散型随机变量，为了使检验的实际水平严格等于 α，我们需要引入随机化检验函数 $\phi(x)$：

$$\phi(x) = \begin{cases} 1, & T(x) > C \\ \gamma, & T(x) = C \\ 0, & T(x) < C \end{cases}$$

其中，临界值 C（整数）和随机化概率 $\gamma \in [0,1)$ 由显著性水平 α 在边界点 $p_0 = 01$ 处确定：

$$E_{p_0}[\phi(X)] = P_{p_0}(T > C) + \gamma P_{p_0}(T = C) = \alpha$$

二、当 $n = 10, \alpha = 0.05$ 时的具体计算

当 $n = 10, p_0 = 01$ 时，统计量 $T \sim B(10, 01)$。我们需要确定 C 和 γ 使得 $E_{01}[\phi] = 0.05$。

1 计算 $B(10, 01)$ 的概率分布

概率质量函数为 $P(T = k) = \binom{10}{k}(01)^k(09)^{10-k}$。

- $P(T = 0) = \binom{10}{0}(09)^{10} \approx 03487$
- $P(T = 1) = \binom{10}{1}(01)(09)^9 \approx 03874$
- $P(T = 2) = \binom{10}{2}(01)^2(09)^8 = 45 \times 001 \times 0430467 \approx 01937$
- $P(T = 3) = \binom{10}{3}(01)^3(09)^7 = 120 \times 0001 \times 0478297 \approx 0.0574$
- $P(T = 4) = \binom{10}{4}(01)^4(09)^6 = 210 \times 00001 \times 0531441 \approx 00112$

2 确定临界值 C

我们需要找到满足 $P(T > C) \leq 0.05$ 的最小整数 C。计算上侧尾部概率：

- $P(T \geq 4) = 1 - P(T \leq 3) = 1 - (03487 + 03874 + 01937 + 0.0574) = 1 - 09872 = 00128$
- $P(T \geq 3) = P(T \geq 4) + P(T = 3) = 00128 + 0.0574 = 00702$

观察发现：$P(T > 2) = P(T \geq 3) = 00702 > 0.05$ $P(T > 3) = P(T \geq 4) = 00128 < 0.05$ 因此，临界值应取 $C = 3$。

3 确定随机化概率 γ

代入方程 $P(T > 3) + \gamma P(T = 3) = 0.05$：

$$00128 + \gamma \times 0.0574 = 0.05$$

$$\gamma = \frac{0.05 - 00128}{0.0574} = \frac{00372}{0.0574} \approx 0648$$

4 写出检验函数

该问题的一致最优势检验函数为：

$$\phi(x_1,\ldots,x_{10})=\begin{cases}1, & \sum x_i>3 \quad (\text{即}\sum x_i\geq 4)\\ 0648, & \sum x_i=3\\ 0, & \sum x_i<3\end{cases}$$

三、判断是否拒绝原假设

根据上述检验函数 $\phi(x)$，对以下情况进行判断：

(a) 当 $\sum_{i=1}^{10}x_i\geq 2$ 时：这里需要根据具体的观测值来判断。

- 假如观测值 $\sum x_i=2$：由于 $2<3$，根据检验函数 $\phi(2)=0$，此时接受原假设（不拒绝 H_0）。
- 即使题目意指“只要落在 ≥ 2 的区域”，我们发现 $P_{H_0}(T\geq 2)\approx 026>0.05$，该区域不能作为显著性水平为 0.05 的拒绝域。
- **结论**：对于观测值等于 2 的情况，不拒绝原假设。

(b) 当 $\sum_{i=1}^{10}x_i<2$ 时：这意味着观测值 $\sum x_i$ 为 0 或 1。

- 由于 $0<3$ 且 $1<3$，根据检验函数 $\phi(x)=0$，此时接受原假设（不拒绝 H_0）。
- **结论**：不拒绝原假设。

1.7 贝叶斯分析

贝叶斯推断的核心是利用样本信息修正先验信息。贝叶斯学派认为，统计推断的过程实际上是信息更新的过程。**随着数据量的增加，先验的影响力会越来越小**

1. 初始信息：先验分布 $\pi(\theta)$（试验前的认识）。
2. 新信息：样本信息 x，通过似然函数 $L(\theta|x)=f(x|\theta)$ 体现。
3. 综合信息：后验分布 $\pi(\theta|x)$（试验后的认识）。

根据贝叶斯公式，三者的关系为：$\pi(\theta|x)=\dfrac{f(x|\theta)\pi(\theta)}{\int_\Theta f(x|\theta)\pi(\theta)d\theta}\propto L(\theta|x)\pi(\theta)$，我们将 $L(\theta|x)\pi(\theta)$ 称为后验分布的核 (Kernel)，分母 $\int_\Theta f(x|\theta)\pi(\theta)d\theta$ 称为边缘概率。

连续型参数：

$$\pi(\theta|x)=\frac{f(x|\theta)\pi(\theta)}{\int_\Theta f(x|\theta)\pi(\theta)d\theta}\propto L(\theta;x)\pi(\theta) \tag{1.1}$$

离散型参数: 若参数空间 $\Theta=\{\theta_1,\theta_2,\ldots,\theta_k\}$ 是离散的，先验概率为 $P(\theta_i)$，则后验概率为

$$P(\theta_i|x)=\frac{P(x|\theta_i)P(\theta_i)}{\sum_{j=1}^k P(x|\theta_j)P(\theta_j)} \tag{1.2}$$

其中 $P(x|\theta_i)$ 是在参数取 θ_i 时观察到样本 x 的概率。

贝叶斯估计基本步骤

1. **确定参数 θ 的先验分布 $p(\theta)$**

2. **由样本集 $D = \{x_1, x_2, \ldots, x_N\}$ 求出样本联合分布 $p(D|\theta)$：**

$$p(D|\theta) = \prod_{n=1}^{N} p(x_n|\theta)$$

3. **利用贝叶斯公式，求 θ 的后验分布：**

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{\int_\phi p(D|\theta)p(\theta)d\theta}$$

4. **求出贝叶斯估计：**

$$\theta^* = \int_\phi \theta p(\theta|D)d\theta$$

1.7.1 常见先验分布

1. 贝塔分布 (Beta Distribution)

   典型场景：常作为二项分布或伯努利分布中成功概率 p 的先验分布。

   定义：若 $X \sim Beta(\alpha, \beta)$，其概率密度函数为：

$$f(x;\alpha,\beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1$$

   其中 $\alpha, \beta > 0$ 是超参数，$B(\alpha,\beta) = \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$。对 $n \in N^*, \Gamma(n) = (n-1)!$

   统计量：期望 $E[X] = \dfrac{\alpha}{\alpha+\beta}$，方差 $Var(X) = \dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

   特例：当 $\alpha = 1, \beta = 1$ 时，即为均匀分布 $U(0,1)$，表示无信息先验。此时，概率密度函数为 $f(x;\alpha,\beta) = 1$

2. 伯努利分布 (Bernoulli Distribution)

   定义：若随机变量 X 只取 0 或 1 两个值，且取 1 的概率为 p，取 0 的概率为 $1-p$，则称 X 服从参数为 p 的伯努利分布（或 0-1 分布），记作 $X \sim Bern(p)$。

   概率质量函数：$f(x;p) = p^x(1-p)^{1-x}, \quad x \in \{0,1\}$

   统计量：期望 $E[X] = p$，方差 $Var(X) = p(1-p)$

3. 二项分布 (Binomial Distribution)

定义：做 n 次独立的伯努利试验，设每次试验成功的概率为 p，则 n 次试验中成功的总次数 X 服从二项分布，记作 $X \sim Bin(n,p)$。

概率质量函数：$f(k;n,p) = P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k=0,1,\dots,n$，其中 $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$。

统计量：期望 $E[X] = np$，方差 $Var(X) = np(1-p)$

4. 正态分布 (Normal Distribution)

   典型场景：当方差 σ^2 已知时，常作为正态分布均值 μ 的先验分布。

   定义：若 $X \sim N(\mu_0, \tau^2)$，其概率密度函数为：$f(x;\mu_0,\tau^2) = \dfrac{1}{\sqrt{2\pi}\tau} \exp\left\{-\dfrac{(x-\mu_0)^2}{2\tau^2}\right\}$，$-\infty < x < \infty$

   共轭性：已知 σ^2，正态总体的均值 μ 在正态先验下，其后验分布仍为正态分布。

5. 几何分布 (Geometric Distribution)

   进行一系列独立重复的伯努利试验，每次试验成功概率为 p，X 表示首次成功前的失败次数，X 服从几何分布：$P(X=k|p) = (1-p)^k p, \quad k=0,1,2,\dots$

   统计量：期望 $E[X] = \dfrac{1-p}{p}$，方差 $Var(X) = \dfrac{1-p}{p^2}$

1.7.2 损失函数与贝叶斯估计

贝叶斯估计量 $\hat{\theta}_B$ 取决于所选用的损失函数 $L(\theta,\delta)$，目标是最小化后验期望损失（贝叶斯风险）。

平方损失函数

- **损失函数**：$L(\theta,\delta) = (\theta-\delta)^2$。
- **贝叶斯估计**：后验分布的均值。

$$\hat{\theta}_B = E(\theta|X) \tag{1.3}$$

0-1 损失函数

- **损失函数**：$L(\theta,\delta) = \begin{cases} 0, & |\theta-\delta| \le \epsilon \\ 1, & |\theta-\delta| > \epsilon \end{cases}$ （极限情况）。
- **贝叶斯估计**：最大后验估计。

$$\hat{\theta}_B = \arg\max_{\theta} \pi(\theta|X) \tag{1.4}$$

对于离散型参数，即选取后验概率最大的那个 θ 值。

绝对损失函数

- **损失函数**：$L(\theta,\delta)=|\theta-\delta|$。
- **贝叶斯估计**：后验分布的中位数。

1.7.3 贝叶斯区间估计与估计评价

贝叶斯可信区间

对于给定的置信水平 $1-\alpha$，贝叶斯可信区间 $[a,b]$ 满足：

$$P(a\leq\theta\leq b|x)\geq 1-\alpha \tag{1.5}$$

通常采用**等尾可信区间**：

$$\int_{-\infty}^{a}\pi(\theta|x)d\theta=\alpha/2,\quad \int_{b}^{+\infty}\pi(\theta|x)d\theta=\alpha/2$$

即利用后验分布的分位数。对于正态后验 $N(\mu_n,\sigma_n^2)$，区间为 $[\mu_n-z_{\alpha/2}\sigma_n,\mu_n+z_{\alpha/2}\sigma_n]$。

估计量的评价

- **均方误差 (MSE)**：$MSE(\hat{\theta})=E[(\hat{\theta}-\theta)^2]=Var(\hat{\theta})+[Bias(\hat{\theta})]^2$。比较两个估计量的 MSE，MSE 越小越优。
- **费雪信息量与 CR 下界**：

  $$I(\theta)=E\left[\left(\frac{\partial\ln f(X;\theta)}{\partial\theta}\right)^2\right]=-E\left[\frac{\partial^2\ln f(X;\theta)}{\partial\theta^2}\right]$$

  有效估计量的方差应达到 Cramer-Rao 下界 $\dfrac{1}{nI(\theta)}$。

1.7.4 练习题

问题 35. 设总体 $X\sim N(\mu,1)$，参数 $\mu\sim N(0,1)$，$X_1,\ldots,X_n$ 是来自于总体 X 的样本，并且损失函数为 $\lambda(\mu,d)=(\mu-d)^2$，求参数 μ 的贝叶斯估计量 $\hat{\mu}$。

【解析】

这是一个标准的正态-正态贝叶斯估计问题。我们需要先求出后验分布，然后根据平方损失函数的性质求出贝叶斯估计。

第一步：写出先验分布

已知参数 $\mu\sim N(0,1)$，其先验概率密度函数为：

$$\pi(\mu)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{\mu^2}{2}\right\}\propto\exp\left\{-\frac{\mu^2}{2}\right\}$$

第二步：写出似然函数

由于总体 $X \sim N(\mu, 1)$，其概率密度函数为 $f(x|\mu) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2}\right\}$。样本的似然函数为：

$$L(\mu; x) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(x_i-\mu)^2}{2}\right\} \propto \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}$$

利用恒等式 $\boldsymbol{\sum(x_i-\mu)^2 = \sum(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2}$，关于 μ 的核心部分为：

$$L(\mu; x) \propto \exp\left\{-\frac{n(\bar{x}-\mu)^2}{2}\right\}$$

第三步：求解后验分布

根据贝叶斯公式 $\pi(\mu|x) \propto L(\mu; x)\pi(\mu)$，将似然函数与先验密度相乘：

$$\begin{aligned}\pi(\mu|x) &\propto \exp\left\{-\frac{n(\bar{x}-\mu)^2}{2}\right\} \cdot \exp\left\{-\frac{\mu^2}{2}\right\} \\ &= \exp\left\{-\frac{1}{2}\left[n(\mu^2 - 2\bar{x}\mu + \bar{x}^2) + \mu^2\right]\right\} \\ &= \exp\left\{-\frac{1}{2}\left[(n+1)\mu^2 - 2n\bar{x}\mu + n\bar{x}^2\right]\right\}\end{aligned}$$

我们需要配方成关于 μ 的正态分布形式 $-\frac{(\mu-\mu_n)^2}{2\sigma_n^2}$。指数部分关于 μ 的项为 $(n+1)\mu^2 - 2n\bar{x}\mu$。

$$(n+1)(\mu^2 - \frac{2n\bar{x}}{n+1}\mu) = (n+1)\left(\mu - \frac{n\bar{x}}{n+1}\right)^2 + C$$

忽略常数 C，我们得到后验分布的核：

$$\pi(\mu|x) \propto \exp\left\{-\frac{n+1}{2}\left(\mu - \frac{n\bar{x}}{n+1}\right)^2\right\}$$

这正是正态分布 $N(\mu_n, \sigma_n^2)$ 的形式，正态分布标准核为 $\pi(\mu|x) = \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$，因此：

$$\text{后验均值}\mu_n = \frac{n\bar{x}}{n+1}, \quad \text{后验方差}\sigma_n^2 = \frac{1}{n+1}$$

即 $\mu|x \sim N\left(\frac{n\bar{x}}{n+1}, \frac{1}{n+1}\right)$。

第四步：根据损失函数求贝叶斯估计

在平方损失下，贝叶斯估计量 $\hat{\mu}$ 为后验分布的期望（均值）。因此：

$$\hat{\mu} = E(\mu|x) = \frac{n\bar{X}}{n+1}$$

问题 36. 设 $X_1, \ldots, X_n$ 为总体 $X \sim B(10, p)$ 的样本，$0 < p < 1$。

(1) 试证 $T_1 = \frac{1}{10}\bar{X}$ 是参数 p 的有效估计量，并求相应的信息量 $I(p)$；

(2) 如果 $p \sim U[0,1]$，在平方损失下，求参数 p 的贝叶斯估计量 T_2；

(3) 试比较两个估计量 T_1 和 T_2 的优劣。

【解析】

(1) 证明 T_1 是有效估计量并求信息量 $I(p)$。

必须先验证无偏性，再验证方差达到 CRLB。

第一步：计算 T_1 的期望与方差。已知总体 $X \sim B(10,p)$，则二项分布单个样本 X_i 的期望与方差为：

$$E(X_i) = 10p, \quad Var(X_i) = 10p(1-p)$$

样本均值 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 的期望与方差为：

$$E(\bar{X}) = 10p, \quad Var(\bar{X}) = \frac{10p(1-p)}{n}$$

对于估计量 $T_1 = \frac{1}{10}\bar{X}$：

- **首先证明无偏性**：

$$E(T_1) = E\left(\frac{1}{10}\bar{X}\right) = \frac{1}{10}E(\bar{X}) = \frac{1}{10}\cdot 10p = p$$

  所以 T_1 是参数 p 的无偏估计量。

- **方差**：

$$Var(T_1) = Var\left(\frac{1}{10}\bar{X}\right) = \frac{1}{100}Var(\bar{X}) = \frac{1}{100}\cdot\frac{10p(1-p)}{n} = \frac{p(1-p)}{10n}$$

第二步：计算 Fisher 信息量。设 X 为总体的一个观测值，其概率质量函数为：

$$f(x;p) = \binom{10}{x} p^x (1-p)^{10-x}, \quad x = 0,1,\ldots,10$$

对数似然函数为：

$$\ln f(x;p) = \ln\binom{10}{x} + x\ln p + (10-x)\ln(1-p)$$

求一阶导数：

$$\frac{\partial \ln f}{\partial p} = \frac{x}{p} - \frac{10-x}{1-p}$$

求二阶导数：

$$\frac{\partial^2 \ln f}{\partial p^2} = -\frac{x}{p^2} - \frac{10-x}{(1-p)^2}$$

单个样本的费雪信息量 $I_1(p)$ 为:

$$\begin{aligned} I_1(p) &= -E\left[\frac{\partial^2 \ln f}{\partial p^2}\right] = E\left[\frac{X}{p^2} + \frac{10-X}{(1-p)^2}\right] \\ &= \frac{10p}{p^2} + \frac{10-10p}{(1-p)^2} = \frac{10}{p} + \frac{10}{1-p} \\ &= \frac{10}{p(1-p)} \end{aligned}$$

样本容量为 n 的样本总费雪信息量 $I(p)$ 为:

$$I(p) = nI_1(p) = \frac{10n}{p(1-p)}$$

第三步:比较 $Var(\hat{\theta})$ 是否等于 CRLB。根据 Cramér-Rao 不等式,参数 p 的任何无偏估计量 $\hat{p}$ 的方差满足下界:

$$Var(\hat{p}) \geq \frac{1}{I(p)} = \frac{1}{\frac{10n}{p(1-p)}} = \frac{p(1-p)}{10n}$$

对比 T_1 的方差:

$$Var(T_1) = \frac{p(1-p)}{10n}$$

由于 $Var(T_1)$ 达到了 Cramér-Rao 下界,因此 T_1 是参数 p 的**有效估计量**。

(2) 求参数 p 的贝叶斯估计量 T_2

第一步:写出似然函数和先验分布。样本 $x = (x_1, \ldots, x_n)$ 的似然函数为:

$$\begin{aligned} L(p; x) &= \prod_{i=1}^{n} \binom{10}{x_i} p^{x_i} (1-p)^{10-x_i} \\ &\propto p^{\sum_{i=1}^n x_i} (1-p)^{10n - \sum_{i=1}^n x_i} \end{aligned}$$

先验分布 $p \sim U[0,1]$,即 $Beta(1,1)$ 分布,密度函数 $\boldsymbol{\pi(p) = 1}$**,** $\boldsymbol{0 < p < 1}$。

第二步:确定后验分布。后验分布密度函数为:

$$\begin{aligned} \pi(p|x) &\propto L(p;x)\pi(p) \\ &\propto p^{\sum x_i} (1-p)^{10n - \sum x_i} \cdot 1 \\ &= p^{(\sum x_i + 1) - 1} (1-p)^{(10n - \sum x_i + 1) - 1} \end{aligned}$$

这正是 Beta 分布的核。因此,后验分布为:

$$p|x \sim Beta\left(\sum_{i=1}^n x_i + 1,\ 10n - \sum_{i=1}^n x_i + 1\right)$$

第三步:求贝叶斯估计。在平方损失函数下,贝叶斯估计量 T_2 为后验分布的期望。若 $Y \sim Beta(\alpha, \beta)$,则 $E(Y) = \dfrac{\alpha}{\alpha+\beta}$。此处 $\alpha = \sum x_i + 1$,$\beta = 10n - \sum x_i + 1$。

$$T_2 = E(p|x) = \frac{\sum_{i=1}^n X_i + 1}{10n + 2}$$

(3) 比较 T_1 和 T_2 的优劣

我们将通过比较两个估计量的**方差**来评价其优劣。

比较估计量 T_1, T_2, 有: $Var(T_1) = Var(\frac{1}{n}\overline{X}) = \frac{1}{n^2}Var(\overline{X})$

$$Var(T_2) = Var\left(\frac{n\overline{X}+1}{10n+2}\right) = \left(\frac{n}{10n+2}\right)^2 Var(\overline{X}) > Var(T_1)$$

所以, T_1 优于 T_2

问题 37. 设随机变量 X 服从几何分布，即 $P(X=k)=\theta(1-\theta)^k, k=0,1,2,\ldots$，其中参数 θ 的先验分布为均匀分布 $U(0,1)$。

(1) 若只对 X 作一次观察，观察值为 3，在平方损失函数下求 θ 的 $Bayes$ 估计；

(2) 若对 X 三次观察，观察值为 $2,3,5$，在平方损失函数下求 θ 的 $Bayes$ 估计。

【解析】

解题准备：确定后验分布的一般形式

设总体 X 服从几何分布，其概率分布列为:

$$P(X=x)=\theta(1-\theta)^x, \quad x=0,1,2,\ldots$$

参数 θ 的先验分布为均匀分布 $U(0,1)$，即 $Beta(1,1)$ 分布，其密度函数为:

$$\pi(\theta)=1, \quad 0<\theta<1$$

在平方损失函数下，参数 θ 的贝叶斯估计量 $\hat{\theta}$ 为后验分布的期望 $E(\theta|X)$。

(1) 只作一次观察，观察值 $x=3$

1 写出似然函数

$$L(\theta;x)=P(X=3)=\theta(1-\theta)^3$$

2 求后验分布。根据贝叶斯公式，后验密度 $\pi(\theta|x)$ 正比于似然函数与先验密度的乘积:

$$\begin{aligned}\pi(\theta|x) &\propto L(\theta;x)\pi(\theta)\\ &= \theta(1-\theta)^3\cdot 1\\ &= \theta^{2-1}(1-\theta)^{4-1}\end{aligned}$$

观察该式，它是 $Beta(2,4)$ 分布的核。因此，后验分布为 $\theta|x\sim Beta(2,4)$。

3 求贝叶斯估计。后验分布的期望即为贝叶斯估计:

$$\hat{\theta}=E(\theta|x)=\frac{2}{2+4}=\frac{2}{6}=\frac{1}{3}$$

(2) 作三次观察，观察值 $x_1=2, x_2=3, x_3=5$

1 写出似然函数。样本 $x = (2, 3, 5)$ 的似然函数为：

$$\begin{aligned} L(\theta; x) &= \prod_{i=1}^{3} P(X = x_i) \\ &= \left[\theta(1-\theta)^2\right] \cdot \left[\theta(1-\theta)^3\right] \cdot \left[\theta(1-\theta)^5\right] \\ &= \theta^3(1-\theta)^{2+3+5} \\ &= \theta^3(1-\theta)^{10} \end{aligned}$$

2 求后验分布。同样利用贝叶斯公式：

$$\begin{aligned} \pi(\theta|x) &\propto L(\theta; x)\pi(\theta) \\ &= \theta^3(1-\theta)^{10} \cdot 1 \\ &= \theta^{4-1}(1-\theta)^{11-1} \end{aligned}$$

这对应于 $Beta(4, 11)$ 分布的核。因此，后验分布为 $\theta|x \sim Beta(4, 11)$。

3 求贝叶斯估计。计算后验分布的期望：

$$\hat{\theta} = E(\theta|x) = \frac{4}{4+11} = \frac{4}{15}$$

问题 38. 对正态分布 $N(\theta, 1)$ 观察，获得三个独立观察值 $X_1 = 2, X_2 = 3, X_3 = 4$，若 θ 的先验分布为 $N(3, 1)$，求 θ 的 0.95 可信区间。

【解析】

解题思路：

1. 确定模型结构：这是典型的正态似然函数配合正态先验分布，属于共轭先验情形。

2. 求解后验分布：利用**正态-正态模型的后验参数更新公式**求出 θ 的后验分布 $\theta|x \sim N(\mu_n, \sigma_n^2)$。设 $X_1, X_2, \ldots, X_n$ 是来自 $N(\mu, \sigma^2)$ 的独立同分布样本，其中 σ^2 已知。假设参数 μ 的先验分布为正态分布：$\mu \sim N(\mu_0, \tau_0^2)$。以下是需要牢记的核心公式：

   (1) 后验精度公式，在贝叶斯统计中，方差的倒数 $1/\sigma^2$ 称为精度。

   $$\frac{1}{\sigma_n^2} = \frac{1}{\tau_0^2} + \frac{\boldsymbol{n}}{\sigma^2}$$

   解释：后验精度 = 先验精度 + 样本精度。这体现了信息的叠加。

   (2) 后验均值公式，后验均值是先验均值和样本均值的加权平均：

   $$\mu_n = \frac{\frac{1}{\tau_0^2}}{\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}}\mu_0 + \frac{\frac{n}{\sigma^2}}{\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}}\bar{x}$$

   也可以写成更常见的形式：$\mu_n = \sigma_n^2 \left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_0}{\tau_0^2}\right)$

3. 构造可信区间：基于后验正态分布，利用标准正态分位数确定 95% 等尾可信区间。

第一步：整理样本信息。已知样本观测值为 $x_1 = 2, x_2 = 3, x_3 = 4$，样本容量 $n = 3$。计算样本均值：

$$\bar{x} = \frac{2+3+4}{3} = 3$$

总体方差已知：$\sigma^2 = 1$。

第二步：确定后验分布。模型设定如下：

- 总体分布：$X \sim N(\theta, \sigma^2) = N(\theta, 1)$
- 先验分布：$\theta \sim N(\mu_0, \tau^2) = N(3, 1)$，即 $\mu_0 = 3, \tau^2 = 1$。

对于正态均值的贝叶斯推断，后验分布 $\theta|x$ 依然服从正态分布 $N(\mu_n, \sigma_n^2)$。利用参数更新公式：

1 计算后验精度（方差的倒数）：

$$\frac{1}{\sigma_n^2} = \frac{n}{\sigma^2} + \frac{1}{\tau^2} = \frac{3}{1} + \frac{1}{1} = 4$$

所以，后验方差 $\sigma_n^2 = \dfrac{1}{4} = 0.25$，后验标准差 $\sigma_n = \sqrt{0.25} = 0.5$。

2 计算后验均值：

$$\mu_n = \sigma_n^2 \left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_0}{\tau^2} \right)$$

代入数值：

$$\mu_n = \frac{1}{4} \left(\frac{3 \times 3}{1} + \frac{3}{1} \right) = \frac{1}{4}(9+3) = \frac{12}{4} = 3$$

综上，θ 的后验分布为：

$$\theta|x \sim N(3, 0.5^2)$$

第三步：求 0.95 可信区间。贝叶斯可信区间基于后验分布计算。对于正态分布 $N(\mu_n, \sigma_n^2)$，其 $1-\alpha$ 等尾可信区间为：

$$[\mu_n - z_{\alpha/2}\sigma_n, \quad \mu_n + z_{\alpha/2}\sigma_n]$$

这里 $1-\alpha = 0.95$，即 $\alpha = 0.05$，$\alpha/2 = 0.0.25$。查标准正态分布表可知，上侧 0.0.25 分位数 $z_{0.0.25} = 1.96$，在标准正态分布曲线下，右侧尾部面积恰好等于 0.0.25 的那个横坐标值是 1.96。

代入数值：

- 下界：$3 - 1.96 \times 0.5 = 3 - 0.98 = 2.02$
- 上界：$3 + 1.96 \times 0.5 = 3 + 0.98 = 3.98$

所以，参数 θ 的 0.95 贝叶斯可信区间为 $[2.02, 3.98]$。

问题 39. 设随机变量 X 服从几何分布,其分布列为 $P(X=x)=\theta(1-\theta)^x, x=0,1,2,\ldots,$ 其中参数 θ 只能取 1/4,2/4 和 3/4 三个值，并以相同概率取这三个值。如今只获得一个观察值 $X=2$，要在 $0-1$ 损失函数下寻求 θ 的贝叶斯估计。

【解析】

第一步：整理已知信息。

1. 似然函数：总体 X 服从几何分布，当观测值 $X=2$ 时，似然函数为：

$$P(X=2|\theta)=\theta(1-\theta)^2$$

2. 先验分布：参数 θ 的取值空间为 $\Theta=\{\theta_1=0.25, \theta_2=05, \theta_3=075\}$。先验概率为均匀分布：

$$\pi(\theta_1)=\pi(\theta_2)=\pi(\theta_3)=\frac{1}{3}$$

3. 损失函数：采用 $0-1$ 损失函数。在该损失下，贝叶斯估计量 $\hat{\theta}_B$ 是使后验概率最大的那个 θ 值。

第二步：计算联合概率。

我们需要计算每个 θ_i 对应的联合概率 $P(X=2,\theta_i)=P(X=2|\theta_i)\pi(\theta_i)$。

1. 当 $\theta=1/4=0.25$ 时：

$$P(X=2|\theta=0.25)=0.25\times(1-0.25)^2=0.25\times0.75^2=0.25\times0.5625=0.140625$$

联合概率：

$$P(X=2,\theta=0.25)=0.140625\times\frac{1}{3}$$

2. 当 $\theta=2/4=0.5$ 时：

$$P(X=2|\theta=0.5)=0.5\times(1-0.5)^2=0.5\times0.5^2=0.5\times0.25=0.125$$

联合概率：

$$P(X=2,\theta=0.5)=0.125\times\frac{1}{3}$$

3. 当 $\theta=3/4=0.75$ 时：

$$P(X=2|\theta=0.75)=0.75\times(1-0.75)^2=0.75\times0.25^2=0.75\times0.0625=0.046875$$

联合概率：

$$P(X=2,\theta=0.75)=0.046875\times\frac{1}{3}$$

第三步：比较后验概率

根据贝叶斯公式,后验概率 $P(\theta_i|X=2)=\dfrac{P(X=2|\theta_i)\pi(\theta_i)}{\sum P(X=2|\theta_j)\pi(\theta_j)}\propto P(X=2|\theta_i)\pi(\theta_i)$。由于先验概率 $\pi(\theta_i)$ 都相同，因此后验概率的大小正比于似然函数 $P(X=2|\theta_i)$ 的大小。比较计算出的似然函数值，显然，$0.140625>0.125>0.046875$。

第四步：得出结论

在 $0-1$ 损失函数下，贝叶斯估计量是后验概率最大的值。

$$\hat{\theta}_B=\frac{1}{4}$$

1.8 多元统计基础

定义 1.8.1 (p 维正态分布). 设 $X = (X_1, X_2, \ldots, X_p)^T$ 为 p 维随机向量。如果对于任意常向量 $\mathbf{a} \in \mathbb{R}^p$，标量随机变量 $Y = \mathbf{a}^T X = \sum_{i=1}^p a_i X_i$ 都服从一元正态分布（包括方差为 *0* 的退化正态分布，即常数)，则称 X 服从 p 维正态分布。

1.8.1 多元正态分布的基本性质

设 $\mathbf{x} \sim N_p(\mu, \Sigma)$。

- $\mathbf{x}$: 是一个 $p \times 1$ 的列向量，包含 p 个随机变量，即 $\mathbf{x} = (X_1, X_2, \ldots, X_p)^T$。
- $\boldsymbol{\mu}$: 是均值向量，维度为 $p \times 1$。$\boldsymbol{\mu} = E[\mathbf{x}] = (E[X_1], E[X_2], \ldots, E[X_p])^T$
- $\boldsymbol{\Sigma}$: 是协方差矩阵，维度为 $p \times p$，必须是对称的（因为 $\sigma_{ij} = \sigma_{ji}$）且是半正定的。$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \ldots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \ldots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \ldots & \sigma_p^2 \end{pmatrix}$，对角线元素 σ_i^2 是第 i 个分量 X_i 的方差。非对角线元素 σ_{ij} 是 X_i 与 X_j 的协方差。
- 边缘分布：$\mathbf{x}$ 的任意子向量（或分量）仍然服从正态分布。
  - **操作方法**：直接从 μ 和 Σ 中截取对应的元素即可。例如 $x_1 \sim N(\mu_1, \sigma_{11})$。
- 独立性与不相关：
  - 对于正态分布，不相关 $\Leftrightarrow$ 独立。
  - 若要证明两个分量（或线性组合）独立，只需证明它们的协方差为 0。
- 联合概率密度函数：

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)\right\}$$

1.8.2 多元正态分布线性变换的分布

这是计算题的核心。若 $\mathbf{x} \sim N_p(\mu, \Sigma)$，令 $\mathbf{y} = A\mathbf{x} + \mathbf{b}$，则 $\mathbf{y}$ 也服从正态分布。
必须熟练掌握的三个数字特征公式：

1. **期望**：

$$E(A\mathbf{x} + \mathbf{b}) = A\mu + \mathbf{b}$$

2. **方差（协方差矩阵)**：

$$Var(A\mathbf{x} + \mathbf{b}) = A\Sigma A^T$$

3. **两组线性组合的协方差**：若 $\mathbf{y}_1 = A\mathbf{x}$，$\mathbf{y}_2 = B\mathbf{x}$，则：

$$Cov(\mathbf{y}_1, \mathbf{y}_2) = A\Sigma B^T$$

考点：如果算出来 $A\Sigma B^T = \mathbf{0}$，则 $\mathbf{y}_1$ 与 $\mathbf{y}_2$ 独立。

4. 协方差运算律：$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + \mathrm{Cov}(X,Y) + \mathrm{Cov}(Y,X)$

   $\mathrm{Cov}(AX, BY) = A\mathrm{Cov}(X,Y)B^T$

   $\mathrm{Cov}(aX+b, cY+d) = ac\mathrm{Cov}(X,Y)$，$\mathrm{Cov}(X_1+X_2, Y) = \mathrm{Cov}(X_1,Y) + \mathrm{Cov}(X_2,Y)$

1.8.3 条件分布与回归函数

这是大题中最容易丢分的地方，公式较繁琐，需死记硬背。
设 $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)$。
考点：给定 $\mathbf{x}_2$ 时，$\mathbf{x}_1$ 的条件分布

$$\mathbf{x}_1|\mathbf{x}_2 \sim N(\mu_{1|2}, \Sigma_{1|2})$$

- 条件均值 (回归函数)：

$$E(\mathbf{x}_1|\mathbf{x}_2) = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2)$$

  注：题目若问"求 x_1 对 x_2 的回归函数"，算的就是这个条件期望！

- 条件方差：

$$Var(\mathbf{x}_1|\mathbf{x}_2) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

- 2×2 矩阵求逆：主对调，副变号，除行列式

解题三步走：

1. 分块：明确谁是 $\mathbf{x}_1$ (未知的)，谁是 $\mathbf{x}_2$ (已知的/条件的)。
2. 求逆：计算条件变量的方差矩阵 Σ_{22} 的逆矩阵 Σ_{22}^{-1}。
3. 代入：套用上述两个公式。

1.8.4 练习题

问题 40. 判断题：X 为 p 维正态随机向量的充要条件为：对任一 p 维向量 $\mathbf{c}$，$\mathbf{c}^T X$ 是一维正态随机变量。(✓)

问题 41. 设 $\mathbf{x} \sim N_3(\mu, \Sigma)$，其中 $\mu = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$，$\Sigma = \begin{pmatrix} 16 & -4 & 2 \\ -4 & 4 & -1 \\ 2 & -1 & 4 \end{pmatrix}$。在给定 $x_1 + 2x_3$ 时，求 $\begin{pmatrix} x_2 - x_3 \\ x_1 \end{pmatrix}$ 的条件分布。

【解析】

这是一道关于多元正态分布中线性变换及条件分布的典型计算题。

解题思路：

1. 定义新变量：将题目中涉及的所有线性组合定义为新的随机向量。设 $Y = \begin{pmatrix} x_2 - x_3 \\ x_1 \end{pmatrix}$ 为我们关心的变量，设 $Z = x_1 + 2x_3$ 为条件变量。
2. 构造联合分布：将 Y 和 Z 组合成一个新的随机向量 $\mathbf{W} = \begin{pmatrix} Y \\ Z \end{pmatrix} = \begin{pmatrix} x_2 - x_3 \\ x_1 \\ x_1 + 2x_3 \end{pmatrix}$。由于 $\mathbf{x}$ 服从正态分布，$\mathbf{W}$ 作为 $\mathbf{x}$ 的线性变换，也服从多元正态分布。
3. 计算联合分布的参数：利用矩阵乘法计算 $\mathbf{W}$ 的均值向量和协方差矩阵。
4. 利用条件分布公式：对于多元正态分布 $\begin{pmatrix} Y \\ Z \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_Y \\ \mu_Z \end{pmatrix}, \begin{pmatrix} \Sigma_{YY} & \Sigma_{YZ} \\ \Sigma_{ZY} & \Sigma_{ZZ} \end{pmatrix}\right)$，给定 Z 时 Y 的条件分布 $Y|Z$ 依然是正态分布，其均值和协方差分别为：$E(Y|Z) = \mu_Y + \Sigma_{YZ}\Sigma_{ZZ}^{-1}(Z - \mu_Z)$，$Var(Y|Z) = \Sigma_{YY} - \Sigma_{YZ}\Sigma_{ZZ}^{-1}\Sigma_{ZY}$

步骤 1：构造线性变换矩阵

设 $\mathbf{W} = \begin{pmatrix} Y \\ Z \end{pmatrix} = \begin{pmatrix} x_2 - x_3 \\ x_1 \\ x_1 + 2x_3 \end{pmatrix} = A\mathbf{x}$。根据定义，变换矩阵 A 为：

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & 2 \end{pmatrix}$$

其中前两行对应 Y，第三行对应 Z。

步骤 2：计算 W 的均值向量 $E(\mathbf{W})$

$$E(\mathbf{W}) = A\mu = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix}$$

所以:

$$\mu_Y = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mu_Z = -3$$

步骤 3:计算 W 的协方差矩阵 $Var(\mathbf{W})$

$Var(\mathbf{W}) = A\Sigma A^T$。已知 $\Sigma = \begin{pmatrix} 16 & -4 & 2 \\ -4 & 4 & -1 \\ 2 & -1 & 4 \end{pmatrix}$。

先计算 ΣA^T:

$$\Sigma A^T = \begin{pmatrix} 16 & -4 & 2 \\ -4 & 4 & -1 \\ 2 & -1 & 4 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ -1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -6 & 16 & 20 \\ 5 & -4 & -6 \\ -5 & 2 & 10 \end{pmatrix}$$

再计算 $A(\Sigma A^T)$:

$$A\Sigma A^T = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} -6 & 16 & 20 \\ 5 & -4 & -6 \\ -5 & 2 & 10 \end{pmatrix} = \begin{pmatrix} 10 & -6 & -16 \\ -6 & 16 & 20 \\ -16 & 20 & 40 \end{pmatrix}$$

我们将此协方差矩阵分块:

$$Var(\mathbf{W}) = \begin{pmatrix} \Sigma_{YY} & \Sigma_{YZ} \\ \Sigma_{ZY} & \Sigma_{ZZ} \end{pmatrix}$$

其中:

$$\Sigma_{YY} = \begin{pmatrix} 10 & -6 \\ -6 & 16 \end{pmatrix}, \quad \Sigma_{YZ} = \begin{pmatrix} -16 \\ 20 \end{pmatrix}, \quad \Sigma_{ZY} = \begin{pmatrix} -16 & 20 \end{pmatrix}, \quad \Sigma_{ZZ} = 40$$

步骤 4:计算条件分布的参数

1 条件协方差:

$$\begin{aligned} Var(Y|Z) &= \Sigma_{YY} - \Sigma_{YZ}\Sigma_{ZZ}^{-1}\Sigma_{ZY} \\ &= \begin{pmatrix} 10 & -6 \\ -6 & 16 \end{pmatrix} - \begin{pmatrix} -16 \\ 20 \end{pmatrix} \frac{1}{40} \begin{pmatrix} -16 & 20 \end{pmatrix} \\ &= \begin{pmatrix} 3.6 & 2 \\ 2 & 6 \end{pmatrix} \end{aligned}$$

2 条件均值:令 z 为给定值 $x_1 + 2x_3$。

$$\begin{aligned} E(Y|Z=z) &= \mu_Y + \Sigma_{YZ}\Sigma_{ZZ}^{-1}(z - \mu_Z) \\ &= \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} -16 \\ 20 \end{pmatrix} \frac{1}{40}(z - (-3)) \\ &= \begin{pmatrix} 0.8 - 0.4z \\ 2.5 + 0.5z \end{pmatrix} \end{aligned}$$

结论

在给定 $Z = x_1 + 2x_3$ 的条件下，$Y = \begin{pmatrix} x_2 - x_3 \\ x_1 \end{pmatrix}$ 服从二元正态分布：

$$\begin{pmatrix} x_2 - x_3 \\ x_1 \end{pmatrix} \Bigg| (x_1 + 2x_3) \sim N_2 \left(\begin{pmatrix} 0.8 - 0.4(x_1 + 2x_3) \\ 2.5 + 0.5(x_1 + 2x_3) \end{pmatrix}, \begin{pmatrix} 3.6 & 2 \\ 2 & 6 \end{pmatrix} \right)$$

问题 42. 设 $\mathbf{x} = (x_1, x_2, x_3)' \sim N_3(\mu, \Sigma)$，其中 $\mu = \begin{pmatrix} 10 \\ 4 \\ 7 \end{pmatrix}$，$\Sigma = \begin{pmatrix} 9 & -3 & -3 \\ -3 & 5 & 1 \\ -3 & 1 & 5 \end{pmatrix}$。求 x_1 对 $(x_2, x_3)'$ 的回归函数。

【解析】

在多元正态分布中，x_1 对 $(x_2, x_3)'$ 的回归函数定义为条件期望 $E(x_1|x_2, x_3)$。对于正态分布而言，这个回归函数是线性的。

1 标记分块矩阵

我们将随机向量 $\mathbf{x}$ 分为两部分：

$$Y = x_1, \quad \mathbf{Z} = \begin{pmatrix} x_2 \\ x_3 \end{pmatrix}$$

根据题目给出的均值向量 μ 和协方差矩阵 Σ，我们将它们对应分块：

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} = \begin{pmatrix} 10 \\ 4 \\ 7 \end{pmatrix} \implies \mu_Y = 10, \quad \mu_{\mathbf{Z}} = \begin{pmatrix} 4 \\ 7 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} = \begin{pmatrix} 9 & -3 & -3 \\ -3 & 5 & 1 \\ -3 & 1 & 5 \end{pmatrix}$$

分块协方差矩阵为：

$$\Sigma_{YY} = 9$$

$$\Sigma_{Y\mathbf{Z}} = \begin{pmatrix} -3 & -3 \end{pmatrix}$$

$$\Sigma_{\mathbf{Z}Y} = \begin{pmatrix} -3 \\ -3 \end{pmatrix}$$

$$\Sigma_{\mathbf{ZZ}} = \begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix}$$

2 引用回归函数公式

多元正态分布的条件期望公式为:

$$E(Y|\mathbf{Z}) = \mu_Y + \Sigma_{Y\mathbf{Z}}\Sigma_{\mathbf{ZZ}}^{-1}(\mathbf{Z} - \mu_{\mathbf{Z}})$$

3 计算 $\Sigma_{\mathbf{ZZ}}$ 的逆矩阵

$$\Sigma_{\mathbf{ZZ}} = \begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix}$$

其行列式为:

$$|\Sigma_{\mathbf{ZZ}}| = 5 \times 5 - 1 \times 1 = 25 - 1 = 24$$

其逆矩阵为:

$$\Sigma_{\mathbf{ZZ}}^{-1} = \frac{1}{24}\begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}$$

4 计算回归系数 $\Sigma_{Y\mathbf{Z}}\Sigma_{\mathbf{ZZ}}^{-1}$

$$\begin{aligned}\Sigma_{Y\mathbf{Z}}\Sigma_{\mathbf{ZZ}}^{-1} &= \begin{pmatrix} -3 & -3 \end{pmatrix} \cdot \frac{1}{24}\begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix} \\ &= \begin{pmatrix} -0.5 & -0.5 \end{pmatrix}\end{aligned}$$

5 代入公式计算回归函数

$$\begin{aligned}E(x_1|x_2, x_3) &= 10 + \begin{pmatrix} -0.5 & -0.5 \end{pmatrix}\begin{pmatrix} x_2 - 4 \\ x_3 - 7 \end{pmatrix} \\ &= 15.5 - 0.5x_2 - 0.5x_3\end{aligned}$$

结论

x_1 对 $(x_2, x_3)'$ 的回归函数为:

$$x_1 = 15.5 - 0.5x_2 - 0.5x_3$$

问题 43. 设 $\mathbf{x} \sim N_{2p}(\mu, \Sigma)$,令 $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}$,其中 $\mathbf{x}_1$ 是 $\mathbf{x}$ 的前 p 个分量,$\mathbf{x}_2$ 是 $\mathbf{x}$ 的后 p 个分量,同时将 μ 和 Σ 进行相应分块,记为 $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ 和 $\Sigma = \begin{pmatrix} \Sigma_1 & \Sigma_2 \\ \Sigma_2 & \Sigma_1 \end{pmatrix}$。

(1) 证明 $\mathbf{x}_1 + \mathbf{x}_2$ 和 $\mathbf{x}_1 - \mathbf{x}_2$ 相互独立;

(2) 求 $\mathbf{x}_1 + \mathbf{x}_2$ 和 $\mathbf{x}_1 - \mathbf{x}_2$ 的分布。

【解析】

前置准备:符号与定义

已知 $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \sim N_{2p}(\mu, \Sigma)$，其中：

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} = \begin{pmatrix} \Sigma_1 & \Sigma_2 \\ \Sigma_2 & \Sigma_1 \end{pmatrix}$$

定义新的随机向量 $\mathbf{y}_1$ 和 $\mathbf{y}_2$：

$$\mathbf{y}_1 = \mathbf{x}_1 + \mathbf{x}_2, \quad \mathbf{y}_2 = \mathbf{x}_1 - \mathbf{x}_2$$

(1) 证明 $\mathbf{x}_1 + \mathbf{x}_2$ 和 $\mathbf{x}_1 - \mathbf{x}_2$ 相互独立

由于 $\mathbf{x}$ 服从多元正态分布，$\mathbf{y}_1$ 和 $\mathbf{y}_2$ 都是 $\mathbf{x}$ 的线性组合，因此它们联合服从多元正态分布。对于联合正态分布的随机向量，独立性等价于不相关，即协方差为 0。

我们需要计算 $\mathrm{Cov}(\mathbf{y}_1, \mathbf{y}_2)$：

$$\begin{aligned} \mathrm{Cov}(\mathbf{y}_1, \mathbf{y}_2) &= \mathrm{Cov}(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{x}_1 - \mathbf{x}_2) \\ &= \mathrm{Cov}(\mathbf{x}_1, \mathbf{x}_1) - \mathrm{Cov}(\mathbf{x}_1, \mathbf{x}_2) + \mathrm{Cov}(\mathbf{x}_2, \mathbf{x}_1) - \mathrm{Cov}(\mathbf{x}_2, \mathbf{x}_2) \end{aligned}$$

代入题目给出的协方差分块：

- $\mathrm{Cov}(\mathbf{x}_1, \mathbf{x}_1) = \Sigma_{11} = \Sigma_1$
- $\mathrm{Cov}(\mathbf{x}_1, \mathbf{x}_2) = \Sigma_{12} = \Sigma_2$
- $\mathrm{Cov}(\mathbf{x}_2, \mathbf{x}_1) = \Sigma_{21} = \Sigma_2$
- $\mathrm{Cov}(\mathbf{x}_2, \mathbf{x}_2) = \Sigma_{22} = \Sigma_1$

代入计算：

$$\begin{aligned} \mathrm{Cov}(\mathbf{y}_1, \mathbf{y}_2) &= \Sigma_1 - \Sigma_2 + \Sigma_2 - \Sigma_1 \\ &= O \quad (\text{零矩阵}) \end{aligned}$$

因为 $\mathbf{y}_1$ 和 $\mathbf{y}_2$ 联合服从正态分布且协方差矩阵为零，所以 $\mathbf{x}_1 + \mathbf{x}_2$ 与 $\mathbf{x}_1 - \mathbf{x}_2$ 相互独立。

(2) 求 $\mathbf{x}_1 + \mathbf{x}_2$ 和 $\mathbf{x}_1 - \mathbf{x}_2$ 的分布

由于它们是 $\mathbf{x}$ 的线性变换，结果必然服从 p 维正态分布 $N_p(\mu', \Sigma')$。我们需要分别求出期望向量和方差（协方差）矩阵。

1 求 $\mathbf{y}_1 = \mathbf{x}_1 + \mathbf{x}_2$ 的分布

均值：

$$E(\mathbf{y}_1) = E(\mathbf{x}_1) + E(\mathbf{x}_2) = \mu_1 + \mu_2$$

方差：

$$\begin{aligned} \mathrm{Var}(\mathbf{y}_1) &= \mathrm{Cov}(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{x}_1 + \mathbf{x}_2) \\ &= \mathrm{Var}(\mathbf{x}_1) + \mathrm{Cov}(\mathbf{x}_1, \mathbf{x}_2) + \mathrm{Cov}(\mathbf{x}_2, \mathbf{x}_1) + \mathrm{Var}(\mathbf{x}_2) \\ &= \Sigma_1 + \Sigma_2 + \Sigma_2 + \Sigma_1 \\ &= 2(\Sigma_1 + \Sigma_2) \end{aligned}$$

所以：

$$\mathbf{x}_1+\mathbf{x}_2 \sim N_p(\mu_1+\mu_2,\ 2(\Sigma_1+\Sigma_2))$$

2 求 $\mathbf{y}_2=\mathbf{x}_1-\mathbf{x}_2$ 的分布

均值：

$$E(\mathbf{y}_2)=E(\mathbf{x}_1)-E(\mathbf{x}_2)=\mu_1-\mu_2$$

方差：

$$\begin{aligned}\mathrm{Var}(\mathbf{y}_2)&=\mathrm{Cov}(\mathbf{x}_1-\mathbf{x}_2,\mathbf{x}_1-\mathbf{x}_2)\\&=\mathrm{Var}(\mathbf{x}_1)-\mathrm{Cov}(\mathbf{x}_1,\mathbf{x}_2)-\mathrm{Cov}(\mathbf{x}_2,\mathbf{x}_1)+\mathrm{Var}(\mathbf{x}_2)\\&=\Sigma_1-\Sigma_2-\Sigma_2+\Sigma_1\\&=2(\Sigma_1-\Sigma_2)\end{aligned}$$

所以：

$$\mathbf{x}_1-\mathbf{x}_2 \sim N_p(\mu_1-\mu_2,\ 2(\Sigma_1-\Sigma_2))$$

问题 44. *设 $\mathbf{x}\sim N_3(\mu,\Sigma)$，其中 $\mu=\begin{pmatrix}2\\-3\\1\end{pmatrix}$，$\Sigma=\begin{pmatrix}1&1&1\\1&2&2\\1&2&3\end{pmatrix}$。*

(1) 求 $2x_1-x_2+3x_3$ 的分布。

(2) 求常数 a 与 b，使 x_3 与 $x_3+ax_1+bx_2$ 相互独立。

【解析】

解题思路：

1 第一问：求线性组合 $Y=\mathbf{c}^T\mathbf{x}$ 的分布。由于 $\mathbf{x}$ 是多元正态分布，其线性组合服从一元正态分布 $N(\mathbf{c}^T\mu,\mathbf{c}^T\Sigma\mathbf{c})$。

2 第二问：利用正态分布的性质，两变量独立等价于协方差为 0。我们需要构造关于 a 和 b 的方程 $\mathrm{Cov}(x_3,x_3+ax_1+bx_2)=0$，解方程组即可。

(1) 求 $2x_1-x_2+3x_3$ 的分布

令 $Y=2x_1-x_2+3x_3$。这可以写成向量形式 $Y=\mathbf{c}^T\mathbf{x}$，其中 $\mathbf{c}=\begin{pmatrix}2\\-1\\3\end{pmatrix}$。由于 $\mathbf{x}\sim N_3(\mu,\Sigma)$，根据多元正态分布的性质，线性组合 Y 服从一元正态分布 $N(E(Y),D(Y))$。

1 计算期望 $E(Y)$：

$$E(Y)=\mathbf{c}^T\mu=\begin{pmatrix}2&-1&3\end{pmatrix}\begin{pmatrix}2\\-3\\1\end{pmatrix}=10$$

2 计算方差 $D(Y)$：

$$D(Y) = \mathbf{c}^T\Sigma\mathbf{c} = \begin{pmatrix} 2 & -1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = 29$$

结论：$2x_1 - x_2 + 3x_3 \sim N(10, 29)$。

(2) 求常数 a 与 b，使 x_3 与 $x_3 + ax_1 + bx_2$ 相互独立

设 $Z_1 = x_3$，$Z_2 = x_3 + ax_1 + bx_2$。由于 $\mathbf{x}$ 服从多元正态分布，$\mathbf{Z} = (Z_1, Z_2)^T$ 也服从多元正态分布。对于联合正态分布的随机变量，独立性等价于不相关，即要求 $\mathrm{Cov}(Z_1, Z_2) = 0$。

计算协方差：

$$\begin{aligned} \mathrm{Cov}(Z_1, Z_2) &= \mathrm{Cov}(x_3,\ x_3 + ax_1 + bx_2) \\ &= \mathrm{Cov}(x_3, x_3) + a\mathrm{Cov}(x_3, x_1) + b\mathrm{Cov}(x_3, x_2) \\ &= \mathrm{Var}(x_3) + a\mathrm{Cov}(x_1, x_3) + b\mathrm{Cov}(x_2, x_3) \end{aligned}$$

根据协方差矩阵 $\Sigma = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$，我们可以直接读取所需的方差和协方差值：

- $\mathrm{Var}(x_3) = \sigma_{33} = 3$
- $\mathrm{Cov}(x_1, x_3) = \sigma_{13} = 1$
- $\mathrm{Cov}(x_2, x_3) = \sigma_{23} = 2$

将这些值代入协方差公式并令其为 0：

$$3 + a(1) + b(2) = 0$$

即：

$$a + 2b = -3$$

这是一个关于 a 和 b 的不定方程。只要 a, b 满足该线性关系，x_3 与 $x_3 + ax_1 + bx_2$ 就相互独立。例如：令 $b = 0$，则 $a = -3$；或令 $a = 1$，则 $b = -2$。

结论：常数 a, b 需满足关系式 $a + 2b = -3$。

问题 45. 设 $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N_2(\mu, \Sigma)$，记 $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \triangleq \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} > 0$（即 $\sigma_1 > 0, \sigma_2 > 0, |\rho| < 1$）。

(1) 试写出 X 的联合密度函数和边缘密度函数；

(2) 试说明 ρ 的统计意义。

【解析】

(1) 写出 X 的联合密度函数和边缘密度函数

1 联合密度函数

首先写出二元正态分布的密度函数公式为:

$$f(x_1,x_2)=\frac{1}{2\pi|\Sigma|^{1/2}}\exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\}$$

设 $X=\begin{bmatrix}X_1\\X_2\end{bmatrix}$，其均值向量为 $\mu=\begin{bmatrix}\mu_1\\\mu_2\end{bmatrix}$，协方差矩阵为 $\Sigma=\begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\\rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix}$。

为了写出密度函数，我们需要计算 Σ 的行列式 $|\Sigma|$ 和逆矩阵 Σ^{-1}。

计算行列式:

$$|\Sigma|=\sigma_1^2\sigma_2^2-(\rho\sigma_1\sigma_2)^2=\sigma_1^2\sigma_2^2(1-\rho^2)$$

计算逆矩阵:

$$\Sigma^{-1}=\frac{1}{|\Sigma|}\begin{bmatrix}\sigma_2^2 & -\rho\sigma_1\sigma_2\\-\rho\sigma_1\sigma_2 & \sigma_1^2\end{bmatrix}=\frac{1}{1-\rho^2}\begin{bmatrix}\frac{1}{\sigma_1^2} & -\frac{\rho}{\sigma_1\sigma_2}\\-\frac{\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2}\end{bmatrix}$$

展开二次型 $(x-\mu)^T\Sigma^{-1}(x-\mu)$：令 $Q=(x-\mu)^T\Sigma^{-1}(x-\mu)$，其中 $x=\begin{bmatrix}x_1\\x_2\end{bmatrix}$。

$$\begin{aligned}Q&=\frac{1}{1-\rho^2}\begin{bmatrix}x_1-\mu_1 & x_2-\mu_2\end{bmatrix}\begin{bmatrix}\frac{1}{\sigma_1^2} & -\frac{\rho}{\sigma_1\sigma_2}\\-\frac{\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2}\end{bmatrix}\begin{bmatrix}x_1-\mu_1\\x_2-\mu_2\end{bmatrix}\\&=\frac{1}{1-\rho^2}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]\end{aligned}$$

代入上述结果得:

$$f(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

其中 $-\infty<x_1,x_2<+\infty$。

2 边缘密度函数

根据多元正态分布的性质，若 $X\sim N_2(\mu,\Sigma)$，则其任意分量（或子向量）也服从正态分布，其均值和方差对应于 μ 和 Σ 中的相应元素。

X_1 的边缘分布：$X_1\sim N(\mu_1,\sigma_{11})=N(\mu_1,\sigma_1^2)$。其边缘密度函数为:

$$f_1(x_1)=\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}\right\},\quad -\infty<x_1<+\infty$$

X_2 的边缘分布：$X_2\sim N(\mu_2,\sigma_{22})=N(\mu_2,\sigma_2^2)$。其边缘密度函数为:

$$f_2(x_2)=\frac{1}{\sqrt{2\pi}\sigma_2}\exp\left\{-\frac{(x_2-\mu_2)^2}{2\sigma_2^2}\right\},\quad -\infty<x_2<+\infty$$

(2) 说明 ρ 的统计意义

参数 ρ 称为相关系数。

1. 定义: 从协方差矩阵的形式 $\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$ 可以看出, $\mathrm{Cov}(X_1, X_2) = \sigma_{12} = \rho\sigma_1\sigma_2$。因此，$\rho = \dfrac{\mathrm{Cov}(X_1, X_2)}{\sqrt{\mathrm{Var}(X_1)}\sqrt{\mathrm{Var}(X_2)}}$。

2. 统计意义：ρ 刻画了随机变量 X_1 和 X_2 之间线性相关关系的密切程度和方向。

- 正负性：
  - 若 $\rho > 0$，表示 X_1 与 X_2 正相关（变化趋势相同）；
  - 若 $\rho < 0$，表示 X_1 与 X_2 负相关（变化趋势相反）。
- 程度（大小）：
  - $|\rho|$ 越接近 1，表示 X_1 与 X_2 的线性关系越强。
  - $|\rho| = 1$ 时，表示 X_1 与 X_2 存在严格的线性关系（$X_2 = aX_1 + b$，概率为 1）。
  - $\rho = 0$ 时，表示 X_1 与 X_2 不相关（无线性关系）。特别地，在正态分布中，$\rho = 0$ 等价于 X_1 与 X_2 相互独立。

Chapter 2

期末模拟卷

一、判断题（每题 2 分，共 8 分）

(1) 对于总体分布族 P_θ，已知 $T = T(\boldsymbol{X})$ 是参数 θ 的充分统计量，又 $S(\boldsymbol{X}) = G(T(\boldsymbol{X}))$，且函数 $G(T)$ 是可逆的，则 $S(\boldsymbol{X})$ 也是参数 θ 的充分统计量。(　)

【解析】

正确 (✓)

解析： 根据充分统计量的性质，如果 T 是充分统计量，且 $S = G(T)$ 是一一映射（即 G 可逆），则 S 也是充分统计量。这是因为样本 X 的条件分布 $P(X|S) = P(X|G(T))$，由于 S 和 T 包含的信息量完全相同（可以互相推导），给定 T 已知等价于给定 S 已知，故条件分布不依赖于 θ。

(2) 设 $\hat{\theta}$ 是 θ 的无偏估计，且有 $D(\hat{\theta}) > 0$，则 $\hat{\theta}^2$ 是 θ^2 的无偏估计。(　)

【解析】

错误 (×)

解析： 根据方差的定义 $D(\hat{\theta}) = E(\hat{\theta}^2) - [E(\hat{\theta})]^2$。因为 $\hat{\theta}$ 是 θ 的无偏估计，所以 $E(\hat{\theta}) = \theta$。故 $E(\hat{\theta}^2) = D(\hat{\theta}) + \theta^2$。题目已知 $D(\hat{\theta}) > 0$，所以 $E(\hat{\theta}^2) > \theta^2$。这意味着 $\hat{\theta}^2$ 是 θ^2 的有偏估计（且为高估）。

(3) 在贝叶斯估计中，随着样本数据量的增加，先验的影响力会越来越小。(　)

【解析】

正确 (✓)

解析： 贝叶斯后验分布由先验分布和似然函数共同决定。随着样本量 $n \to \infty$，似然函数在对数似然函数中占据主导地位（其权重与 n 成正比），而先验分布通常是固定的常数项或相对于 n 可以忽略。因此，数据越多，估计结果越依赖于数据本身，先验的影响逐渐减弱。

(4) $\boldsymbol{X}$ 为 p 维正态随机向量的充要条件为：对任一 p 维向量 $\boldsymbol{c}$，$\boldsymbol{c}^T\boldsymbol{X}$ 是一维正态随机变量。(　)

【解析】

正确 (✓)

解析： 这是多维正态分布的定义之一。如果一个随机向量的任意线性组合都服从一维正态分布，则该随机向量服从多维正态分布。

二、填空题（每题 4 分，共 16 分）

(1) 设 $x_1, x_2, ..., x_n$ 是来自均匀分布 $U(\theta, 2\theta)(\theta > 0)$ 总体 X 的样本,那么 $T(x_1, x_2, ..., x_n) =$ ________________ 是参数 θ 的充分统计量。

【解析】

$(x_{(1)}, x_{(n)})$

解析： 总体密度函数为 $f(x;\theta) = \dfrac{1}{\theta} I_{\{\theta \le x \le 2\theta\}}$。似然函数为：

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} I_{\{\theta \le x_i \le 2\theta\}} = \frac{1}{\theta^n} I_{\{\theta \le x_{(1)}\}} \cdot I_{\{x_{(n)} \le 2\theta\}}$$

根据因子分解定理,似然函数依赖于样本仅通过 $(x_{(1)}, x_{(n)})$,因此充分统计量是 $(x_{(1)}, x_{(n)})$（即最小次序统计量和最大次序统计量）。

(2) 设总体 X 的概率密度为 $f(x,\theta) = \begin{cases} \dfrac{1}{\theta} x^{\frac{1-\theta}{\theta}}, & 0 < x < 1 \\ 0, & \text{其它} \end{cases}$,$(\theta > 0)$，$x_1, x_2, ..., x_n$ 为来自总体 X 的样本值，则参数 θ 的极大似然估计为 ________________。

【解析】

$-\dfrac{1}{n} \sum_{i=1}^{n} \ln x_i$ （或 $-\overline{\ln x}$）

解析： 似然函数为：

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} x_i^{\frac{1}{\theta}-1} = \theta^{-n} \left(\prod_{i=1}^{n} x_i \right)^{\frac{1}{\theta}-1}$$

取对数似然函数：

$$\ln L(\theta) = -n \ln \theta + \left(\frac{1}{\theta} - 1 \right) \sum_{i=1}^{n} \ln x_i$$

对 θ 求导并令其为 0：

$$\frac{d \ln L}{d\theta} = -\frac{n}{\theta} - \frac{1}{\theta^2} \sum_{i=1}^{n} \ln x_i = 0$$

解得：

$$-\frac{n}{\theta} = \frac{1}{\theta^2} \sum_{i=1}^{n} \ln x_i \implies \hat{\theta} = -\frac{1}{n} \sum_{i=1}^{n} \ln x_i$$

由于 $0 < x_i < 1$，$\ln x_i < 0$，故 $\hat{\theta} > 0$。

(3) 设随机变量 X 服从几何分布，其分布列为 $P(X = x) = \theta(1-\theta)^x, x = 0, 1, 2, \cdots$，其中参数 θ 只能取 $\dfrac{1}{4}, \dfrac{1}{2}, \dfrac{3}{4}$ 三个值，并以相同概率取这三个值。如今只能获得一个观察值 $X = 2$，求后验概率分布 $P(\theta|X = 2)$ 的期望 $E(\theta|X = 2) =$ ________________。

【解析】

第一步：确定先验分布

根据题意，参数 θ 是离散型随机变量，其可能的取值为 $\Theta=\{\theta_1,\theta_2,\theta_3\}=\{\frac{1}{4},\frac{1}{2},\frac{3}{4}\}$。由于 θ 以相同概率取这三个值，故其先验分布列为：$P(\theta=\theta_i)=\frac{1}{3},\quad i=1,2,3$

第二步：计算似然函数

已知样本观测值为 $X=2$。对于给定的 θ，样本的条件概率为：$P(X=2|\theta)=\theta(1-\theta)^2$，分别计算 θ 取不同值时的似然：

- 当 $\theta=\frac{1}{4}$ 时，$P(X=2|\theta=\frac{1}{4})=\frac{1}{4}\times(1-\frac{1}{4})^2=\frac{1}{4}\times\frac{9}{16}=\frac{9}{64}$
- 当 $\theta=\frac{1}{2}$ 时，$P(X=2|\theta=\frac{1}{2})=\frac{1}{2}\times(1-\frac{1}{2})^2=\frac{1}{2}\times\frac{1}{4}=\frac{1}{8}=\frac{8}{64}$
- 当 $\theta=\frac{3}{4}$ 时，$P(X=2|\theta=\frac{3}{4})=\frac{3}{4}\times(1-\frac{3}{4})^2=\frac{3}{4}\times\frac{1}{16}=\frac{3}{64}$

第三步：计算全概率

根据全概率公式，观测到 $X=2$ 的边缘概率为：

$P(X=2)=\sum_{i=1}^{3}P(X=2|\theta_i)P(\theta_i)=\left(\frac{9}{64}\times\frac{1}{3}\right)+\left(\frac{8}{64}\times\frac{1}{3}\right)+\left(\frac{3}{64}\times\frac{1}{3}\right)=\frac{5}{48}$

第四步：求后验概率分布

根据贝叶斯公式 $P(\theta_i|X=2)=\dfrac{P(X=2|\theta_i)P(\theta_i)}{P(X=2)}$:

- $P(\theta=\frac{1}{4}|X=2)=\dfrac{\frac{9}{64}\times\frac{1}{3}}{20/192}=\dfrac{9/192}{20/192}=\dfrac{9}{20}$
- $P(\theta=\frac{1}{2}|X=2)=\dfrac{\frac{8}{64}\times\frac{1}{3}}{20/192}=\dfrac{8/192}{20/192}=\dfrac{8}{20}$
- $P(\theta=\frac{3}{4}|X=2)=\dfrac{\frac{3}{64}\times\frac{1}{3}}{20/192}=\dfrac{3/192}{20/192}=\dfrac{3}{20}$

第五步：计算后验期望

后验期望定义为 θ 在后验分布下的均值：$E(\theta|X=2)=\sum_{i=1}^{3}\theta_iP(\theta_i|X=2)$
$E(\theta|X=2)=\left(\frac{1}{4}\times\frac{9}{20}\right)+\left(\frac{1}{2}\times\frac{8}{20}\right)+\left(\frac{3}{4}\times\frac{3}{20}\right)=\frac{17}{40}=0.425$

(4) 某一产品有两种品牌，请用户进行评判其优劣，现选了 8 个用户，对这两种品牌的产品分别评分，结果如下：符号秩和检验用的统计量是 $W^+=$ ____________。

用户 i	1	2	3	4	5	6	7	8
品牌 X	55	32	43	48	45	60	50	45
品牌 Y	45	37	41	50	46	54	52	47

符号秩检验标准解题四步法

步骤 1：计算配对差值。对每对数据计算差值 $d_i = X_i - Y_i$，剔除 $d_i = 0$ 的样本。

步骤 2：编秩并处理结。

按 $|d_i|$ 从小到大排序。对相同绝对值的“结”取平均秩。给秩赋予原差值的正负号，得到符号秩。

步骤 3：计算检验统计量。

正秩和 W^+：所有正符号秩的和。负秩和 W^-：所有负符号秩绝对值的和。性质：$W^+ + W^- = \dfrac{n(n+1)}{2}$（$n$ 为非零差值的样本量）。

步骤 4：决策规则

小样本（$n \le 50$）：查 Wilcoxon 符号秩检验界值表，若 W^+ 位于界值范围内则不拒绝 H_0，否则拒绝。

大样本（$n > 50$）：使用正态近似 $Z = \dfrac{|W^+ - \frac{n(n+1)}{4}| - 0.5}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{\sum(t_j^3 - t_j)}{48}}}$，其中 t_j 为第 j 个结的相同数据个数

【解析】

步骤 1：构造秩表并计算平均秩

首先，对每对数据计算差值 $d_i = X_i - Y_i$，剔除所有 $d_i = 0$ 的样本。将绝对值 $|d_i|$ **从小到大排序**，对存在相同绝对值的“结”计算平均秩，最终得到完整秩表如下：

用户 i	X_i	Y_i	d_i	$\|d_i\|$	秩（平均秩）	符号秩
5	45	46	-1	1	1	-1
3	43	41	2	2	3.5	+3.5
4	48	50	-2	2	3.5	-3.5
7	50	52	-2	2	3.5	-3.5
8	45	47	-2	2	3.5	-3.5
2	32	37	-5	5	6	-6
6	60	54	6	6	7	+7
1	55	45	10	10	8	+8

步骤 2：计算正秩和 W^+，W^+ 为所有正符号秩的和：

$$W^+ = 8 + 3.5 + 7 = 18.5$$

扩展：确定显著性水平 α。本题属于配对样本的符号秩和检验，检验目的是判断两种品牌评分是否存在显著差异：

H_0：两种品牌评分无差异，即差值的中位数 $M_d = 0$，H_1：两种品牌评分存在差异，即 $M_d \neq 0$（双侧检验）

决策规则：若 $W_{\alpha/2}(n) < W^+ < \frac{n(n+1)}{2} - W_{\alpha/2}(n)$，则接受 H_0，n 为非零差值的个数

三、计算题（每题 12 分，共 36 分）

(1) 设 $x_1, x_2, ..., x_n$ 是来自正态总体 $N(0, \sigma^2)$ 的简单样本，①求参数 σ^2 的一致最小方差无偏估计 $\hat{\sigma}^2$；②样本 $x_1, x_2, ..., x_n$ 所包含的参数 σ^2 的 Fisher 信息量；③ 证明 $\hat{\sigma}^2$ 是参数 σ^2 的有效估计。(12 分)

【解析】

(1) 求 σ^2 的 UMVUE

第一步：找充分完备统计量

样本的联合概率密度函数为：

$$f(x_1, \ldots, x_n; \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x_i^2}{2\sigma^2}\right) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{\sum_{i=1}^{n} x_i^2}{2\sigma^2}\right)$$

根据因子分解定理，令 $T = \sum_{i=1}^{n} x_i^2$，则联合密度可写为 $g(T, \sigma^2)h(\mathbf{x})$ 的形式。因此，$T = \sum_{i=1}^{n} x_i^2$ 是 σ^2 的充分统计量。由于该分布属于指数型分布族，且参数空间包含内点，故 T 是**完备充分统计量**。

第二步：构造无偏估计

已知 $E[x_i] = 0$，$Var(x_i) = \sigma^2$，则 $E[x_i^2] = Var(x_i) + (E[x_i])^2 = \sigma^2$。考虑统计量 $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$，其期望为：

$$E[\hat{\sigma}^2] = E\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] = \frac{1}{n}\sum_{i=1}^{n} E[x_i^2] = \frac{1}{n} \cdot n\sigma^2 = \sigma^2$$

故 $\hat{\sigma}^2$ 是 σ^2 的无偏估计。

第三步：结论

根据 Lehmann-Scheffé 定理，由于 $\hat{\sigma}^2$ 是完备充分统计量 T 的函数且是无偏的，故 $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$ 是 σ^2 的 **UMVUE**。

(2) 计算 Fisher 信息量

设总体密度函数为 $f(x; \theta)$，其中 $\theta = \sigma^2$。单样本的 Fisher 信息量定义为：

$$I(\theta) = E\left[\left(\frac{\partial \ln f(X; \theta)}{\partial \theta}\right)^2\right]$$

在满足正则条件（通常默认满足）时，通常使用二阶导数形式计算更为简便：

$$\boxed{I(\theta) = -E\left[\frac{\partial^2 \ln f(X; \theta)}{\partial \theta^2}\right]}$$

对于容量为 n 的独立同分布样本，样本全信息量为：$I_n(\theta) = nI(\theta)$

设单个观测值的对数似然函数为 $l(\sigma^2; x) = \ln f(x; \sigma^2)$。令 $\theta = \sigma^2$：

$$l(\theta; x) = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\theta - \frac{x^2}{2\theta}$$

对 θ 求一阶导数：

$$\frac{\partial l}{\partial \theta} = -\frac{1}{2\theta} + \frac{x^2}{2\theta^2}$$

对 θ 求二阶导数：

$$\frac{\partial^2 l}{\partial \theta^2} = \frac{1}{2\theta^2} - \frac{x^2}{\theta^3}$$

单个样本的 Fisher 信息量为：

$$I(\theta) = -E\left[\frac{\partial^2 l}{\partial \theta^2}\right] = -\left(\frac{1}{2\theta^2} - \frac{E[x^2]}{\theta^3}\right) = -\left(\frac{1}{2\theta^2} - \frac{\theta}{\theta^3}\right) = \frac{1}{2\theta^2} = \frac{1}{2\sigma^4}$$

由于样本独立同分布，样本全信息量为：$I_n(\sigma^2) = nI(\sigma^2) = \dfrac{n}{2\sigma^4}$

(3) 证明有效性

第 0 步：陈述 Cramer-Rao 不等式（正式定义）

设 $x_1, x_2, \ldots, x_n$ 是来自总体 $f(x;\theta)$ 的样本，$\hat{\theta}$ 是参数 $g(\theta)$ 的一个无偏估计量。在满足正则性条件下，$\hat{\theta}$ 的方差满足：

$$Var(\hat{\theta}) \geq \frac{[g'(\theta)]^2}{I_n(\theta)}$$

其中 $I_n(\theta)$ 为样本全信息量。此下界称为 Cramer-Rao 下界 (CRLB)。若一个无偏估计量的方差等于该下界，则称该估计量为 $g(\theta)$ 的**有效估计**。

第一步：计算 Cramer-Rao 下界

在本题中，待估参数为 $\theta = \sigma^2$，待估函数为 $g(\theta) = \theta$。因此，$g'(\theta) = \dfrac{d(\sigma^2)}{d(\sigma^2)} = 1$。

由第 (2) 问知 $I_n(\sigma^2) = \dfrac{n}{2\sigma^4}$，则 CRLB 为：$CRLB = \dfrac{1}{I_n(\sigma^2)} = \dfrac{2\sigma^4}{n}$

第二步：计算估计量的方差

由于 $x_i \sim N(0,\sigma^2)$，则 $\dfrac{x_i}{\sigma} \sim N(0,1)$，从而 $\dfrac{x_i^2}{\sigma^2} \sim \chi^2(1)$。根据卡方分布的可加性，$\dfrac{\sum x_i^2}{\sigma^2} \sim \chi^2(n)$。已知 $\chi^2(n)$ 分布的方差为 $2n$，则：

$$Var\left(\frac{\sum x_i^2}{\sigma^2}\right) = 2n \implies \frac{1}{\sigma^4}Var\left(\sum x_i^2\right) = 2n \implies Var\left(\sum x_i^2\right) = 2n\sigma^4$$

因此，$\hat{\sigma}^2$ 的方差为：

$$Var(\hat{\sigma}^2) = Var\left(\frac{1}{n}\sum x_i^2\right) = \frac{1}{n^2}Var\left(\sum x_i^2\right) = \frac{2n\sigma^4}{n^2} = \frac{2\sigma^4}{n}$$

结论：

因为 $Var(\hat{\sigma}^2) = CRLB$，即该无偏估计量的方差达到了 Cramer-Rao 下界，所以 $\hat{\sigma}^2$ 是 σ^2 的**有效估计**。

(2) 设 $x_1, x_2, ..., x_n$ 是来自正态总体 $N(0,\sigma^2)$ 的简单样本，在显著性水平 α 下，① 给出下面假设检验问题的似然比检验；②给出该似然比检验的势函数 $g(\sigma^2)$；③ 判断该检验是否为无偏检验。

$$H_0: \sigma^2 \leq \sigma_0^2; \quad H_1: \sigma^2 > \sigma_0^2$$

对于自由度为 n 的卡方分布（记为 $\chi^2(n)$），其概率密度函数曲线下的总面积为 1。上 α 分位数，通常记为 $\chi^2_\alpha(n)$，是指满足以下条件的数值：在该数值右侧（即大于该数值）的区域面积（概率）恰好等于 α。
用数学公式表示，如果随机变量 $X \sim \chi^2(n)$，那么 $\chi^2_\alpha(n)$ 满足：$P(X > \chi^2_\alpha(n)) = \alpha$
或者用累积分布函数 $F(x)$ 表示（即左侧面积）：$P(X \le \chi^2_\alpha(n)) = 1 - \alpha$

【解析】

(1) 似然比检验

基础定义与公式：

似然比检验统计量定义为：

$$\lambda(\mathbf{x}) = \frac{\sup_{\sigma^2 \in \Theta_0} L(\sigma^2)}{\sup_{\sigma^2 \in \Theta} L(\sigma^2)}$$

其中 $\Theta_0 = \{\sigma^2 : \sigma^2 \le \sigma_0^2\}$，$\Theta = \{\sigma^2 : \sigma^2 > 0\}$。拒绝域的形式为 $\{\lambda(\mathbf{x}) \le c\}$，$c$ 为常数，且 $0 < c < 1$。

第一步：写出似然函数并求极大似然估计

似然函数为：$L(\sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{\sum x_i^2}{2\sigma^2}\right)$

由于对数函数是单调递增函数，似然函数 $L(\sigma^2)$ 和对数似然函数 $l(\sigma^2) = \ln L(\sigma^2)$ 的单调性完全一致，因此我们先计算对数似然：

$$\begin{aligned} l(\sigma^2) &= \ln\left[(2\pi\sigma^2)^{-n/2}\right] + \ln\left[\exp\left(-\frac{\sum x_i^2}{2\sigma^2}\right)\right] \\ &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{\sum x_i^2}{2\sigma^2} \end{aligned}$$

令 $t = \sigma^2$，其中 $t > 0$，则对数似然函数变为：

$$l(t) = C - \frac{n}{2}\ln t - \frac{S}{2t}$$

其中 $C = -\frac{n}{2}\ln(2\pi)$ 是常数，$S = \sum x_i^2$ 是样本平方和（非负）。对 $l(t)$ 求一阶导数：

$$l'(t) = -\frac{n}{2t} + \frac{S}{2t^2} = \frac{S - nt}{2t^2}$$

当 $t < \frac{S}{n}$ 时：$S - nt > 0$，$l'(t) > 0$，因此 $l(t)$ 单调递增，对应 $L(\sigma^2)$ 单调递增；
当 $t = \frac{S}{n}$ 时：$l'(t) = 0$，达到极大值点；
当 $t > \frac{S}{n}$ 时：$S - nt < 0$，$l'(t) < 0$，因此 $l(t)$ 单调递减，对应 $L(\sigma^2)$ 单调递减。

全参数空间 Θ 上的极大似然估计为 $\hat{\sigma}^2 = \dfrac{1}{n}\sum_{i=1}^n x_i^2$。$L(\sigma^2)$ 在 $(0, \hat{\sigma}^2)$ 上单调递增，在 $(\hat{\sigma}^2, +\infty)$ 上单调递减。

第二步：计算分子与分母的极大值

- **分母（全空间最大值）：** $\sup_{\sigma^2 \in \Theta} L(\sigma^2) = L(\hat{\sigma}^2)$。

- **分子（H_0 空间最大值）：** $H_0 : \sigma^2 \leq \sigma_0^2$。
  - 若 $\hat{\sigma}^2 \leq \sigma_0^2$，则 $\hat{\sigma}^2 \in \Theta_0$，此时最大值在 $\hat{\sigma}^2$ 处取得，即 $\sup_{\Theta_0} L = L(\hat{\sigma}^2)$。
  - 若 $\hat{\sigma}^2 > \sigma_0^2$，由于 $L(\sigma^2)$ 在 $\hat{\sigma}^2$ 左侧单调递增，故在原假设区间 $(0, \sigma_0^2]$ 上的最大值在右边界 σ_0^2 处取得，即 $\sup_{\Theta_0} L = L(\sigma_0^2)$。

第四步：构造似然比统计量

$$\lambda(x) = \begin{cases} \dfrac{L(\hat{\sigma}^2)}{L(\hat{\sigma}^2)} = 1, & \hat{\sigma}^2 \leq \sigma_0^2 \\ \dfrac{L(\sigma_0^2)}{L(\hat{\sigma}^2)}, & \hat{\sigma}^2 > \sigma_0^2 \end{cases}$$

拒绝域的形式为 $\lambda(x) \leq c < 1$，这等价于只考虑 $\hat{\sigma}^2 > \sigma_0^2$ 的情况。在该范围内，$\lambda(x)$ 是关于 $\hat{\sigma}^2$ 的单调减函数（因为 $\hat{\sigma}^2$ 越偏离 σ_0^2，似然比越小）。**因此，**$\lambda(x) \leq c$ **等价于** $\hat{\sigma}^2 \geq k$。即拒绝域形式为：$W = \{\sum_{i=1}^n x_i^2 \geq C\}$

第五步：确定临界值

在 H_0 的边界 $\sigma^2 = \sigma_0^2$ 处，统计量 $\dfrac{\sum x_i^2}{\sigma_0^2} \sim \chi^2(n)$。

根据显著性水平 α 确定临界值：

$$P_{\sigma_0^2}\left(\sum_{i=1}^n x_i^2 \geq C\right) = P\left(\frac{\sum x_i^2}{\sigma_0^2} \geq \frac{C}{\sigma_0^2}\right) = \alpha$$

令 $\dfrac{C}{\sigma_0^2} = \chi_\alpha^2(n)$，$\chi_\alpha^2(n)$ 为自由度为 n 的卡方分布的**上** α **分位数**。解得 $C = \sigma_0^2 \chi_\alpha^2(n)$。

结论：该检验的拒绝域为 $W = \{(x_1, \ldots, x_n) : \sum_{i=1}^n x_i^2 > \sigma_0^2 \chi_\alpha^2(n)\}$。

(2) 势函数

基础定义与公式：

势函数 $g(\sigma^2)$ 定义为在参数为 σ^2 时拒绝原假设 H_0 的概率，也就是样本落入拒绝域的概率：$g(\sigma^2) = P_{\sigma^2}(\mathbf{X} \in W)$

具体推导：

$g(\sigma^2) = P_{\sigma^2}\left(\sum_{i=1}^n x_i^2 > \sigma_0^2 \chi_\alpha^2(n)\right)$

为了计算该概率，我们需要将不等式左边标准化为服从 $\chi^2(n)$ 的变量。不等式两边同时除以真实的参数 σ^2：

$$g(\sigma^2) = P_{\sigma^2}\left(\frac{\sum x_i^2}{\sigma^2} > \frac{\sigma_0^2}{\sigma^2}\chi_\alpha^2(n)\right)$$

已知 $Y = \dfrac{\sum x_i^2}{\sigma^2} \sim \chi^2(n)$，设 $F_n(y)$ 为 $\chi^2(n)$ 分布的累积分布函数，则：$g(\sigma^2) = 1 - F_n\left(\dfrac{\sigma_0^2}{\sigma^2}\chi_\alpha^2(n)\right)$

(3) 无偏检验的判断

基础定义与公式：对于显著性水平为 α 的检验，如果其势函数 $g(\sigma^2)$ 满足：

- $\forall \sigma^2 \in \Theta_0, \quad g(\sigma^2) \le \alpha$
- $\forall \sigma^2 \in \Theta_1, \quad g(\sigma^2) \ge \alpha$

则称该检验是无偏检验。

证明过程：

考察势函数 $g(\sigma^2) = P\left(Y > \dfrac{\sigma_0^2}{\sigma^2}\chi_\alpha^2(n)\right)$，其中 $Y \sim \chi^2(n)$。

- 当 σ^2 增加时，比值 $\dfrac{\sigma_0^2}{\sigma^2}\chi_\alpha^2(n)$ 单调递减。
- 对于卡方分布，下限越小，右侧尾部面积（概率）越大。
- 因此，$g(\sigma^2)$ 是关于 σ^2 的**单调递增函数**。

验证边界条件： 当 $\sigma^2 = \sigma_0^2$ 时，$g(\sigma_0^2) = P\left(Y > \dfrac{\sigma_0^2}{\sigma_0^2}\chi_\alpha^2(n)\right) = P(Y > \chi_\alpha^2(n)) = \alpha$

结合单调性得出结论：

- 对于 $H_0 : \sigma^2 \le \sigma_0^2$，有 $g(\sigma^2) \le g(\sigma_0^2) = \alpha$
- 对于 $H_1 : \sigma^2 > \sigma_0^2$，有 $g(\sigma^2) > g(\sigma_0^2) = \alpha$

综上所述，该检验是**无偏检验**。

(3) 设 $\boldsymbol{x} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$，其中 $\boldsymbol{\mu} = \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}$，$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$。① 求 $2x_1 - x_2 + 3x_3$ 的分布；② 求常数 a 与 b，使 x_3 与 $x_3 + ax_1 + bx_2$ 相互独立。(12 分)

【解析】

(1) 求 $Y = 2x_1 - x_2 + 3x_3$ 的分布

基础定理： 若 $\boldsymbol{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$，令 $Y = \boldsymbol{c}^T\boldsymbol{x}$（其中 $\boldsymbol{c}$ 为常数向量），则 Y 服从一元正态分布 $\boxed{N(\boldsymbol{c}^T\boldsymbol{\mu}, \boldsymbol{c}^T\boldsymbol{\Sigma}\boldsymbol{c})}$。在本题中，系数向量 $\boldsymbol{c} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$。

第一步：计算期望 $E[Y]$

$E[Y] = \boldsymbol{c}^T\boldsymbol{\mu} = 2E[x_1] - 1E[x_2] + 3E[x_3]$

代入 $\boldsymbol{\mu} = (2, -3, 1)^T$：$E[Y] = 2(2) - 1(-3) + 3(1) = 4 + 3 + 3 = 10$

第二步：计算方差 $Var(Y)$

方差公式为 $Var(Y) = \boldsymbol{c}^T\boldsymbol{\Sigma}\boldsymbol{c}$。

我们可以先计算 $\boldsymbol{\Sigma c}$：$\boldsymbol{\Sigma c} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1(2) + 1(-1) + 1(3) \\ 1(2) + 2(-1) + 2(3) \\ 1(2) + 2(-1) + 3(3) \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ 9 \end{pmatrix}$

再计算 $\boldsymbol{c}^T(\boldsymbol{\Sigma c})$：$Var(Y) = \begin{pmatrix} 2 & -1 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 6 \\ 9 \end{pmatrix} = 2(4)+(-1)(6)+3(9) = 8-6+27 = 29$

结论： $2x_1 - x_2 + 3x_3 \sim N(10, 29)$

(2) 求常数 a, b 使变量独立

基础定理：

对于多元正态分布，两个分量（或其线性组合）相互独立的充要条件是它们的协方差为 0。令 $U = x_3$，$V = x_3 + ax_1 + bx_2$。我们需要满足 $Cov(U, V) = 0$。

计算协方差：

利用协方差的线性性质展开：

$$
\begin{aligned}
Cov(x_3, x_3 + ax_1 + bx_2) &= Cov(x_3, x_3) + aCov(x_3, x_1) + bCov(x_3, x_2) \\
&= Var(x_3) + aCov(x_1, x_3) + bCov(x_2, x_3)
\end{aligned}
$$

从协方差矩阵 $\boldsymbol{\Sigma}$ 中读取对应数值：

- $Var(x_3) = \sigma_{33} = 3$
- $Cov(x_1, x_3) = \sigma_{13} = 1$
- $Cov(x_2, x_3) = \sigma_{23} = 2$

代入方程：$3 + a(1) + b(2) = 0$ 即：$a + 2b = -3$

结论：

只要常数 a, b 满足关系式 $a + 2b = -3$，变量 x_3 与 $x_3 + ax_1 + bx_2$ 即相互独立。

助教笔记：易错点提醒

- 矩阵乘法的顺序：在计算方差 $Var(Y) = \boldsymbol{c}^T\boldsymbol{\Sigma c}$ 时，建议先算右边的矩阵向量乘法 $\boldsymbol{\Sigma c}$，得到一个列向量后，再与左边的行向量 $\boldsymbol{c}^T$ 相乘。这样计算量最小且不易出错。
- 协方差矩阵的读取：$\boldsymbol{\Sigma}$ 是对称矩阵，$\Sigma_{ij} = Cov(x_i, x_j)$。主对角线是方差：$Var(x_1) = 1, Var(x_2) = 2, Var(x_3) = 3$。非对角线是协方差：$Cov(x_1, x_3)$ 对应第一行第三列的元素，即 1。