# 第 9 次作业参考解答

**3.** 设总体的大小为 $N$，总体均值和方差分别为 $\mu, \sigma^2$，$X_i$ $(i = 1, \cdots, n)$ 为无放回抽取的简单随机样本。

- **(1) 证明：$\mathbb{E}(X_i) = \mu$, $\mathbf{Var}(X_i) = \sigma^2$.**

  – **期望**：样本 $X_i$ 是从总体中简单随机抽样得到的，即每个样本来自总体中的任意个体，且被抽取的概率相等。因此，单个样本的期望为：

$$\mathbb{E}(X_i) = \mathbb{E}\mathbb{E}[X_i | i^{th} \ is \ chosen] = \sum \mathbb{P}(i^{th} \ is \ chosen) \cdot X_i = \frac{1}{N} \sum_{i=1}^{N} X_i = \mu.$$

  – **方差**：样本 $X_i$ 的方差为：

$$\mathrm{Var}(X_i) = \mathbb{E}\mathrm{Var}(X_i | i^{th} \ is \ chosen) + \mathrm{Var}\mathbb{E}(X_i | i^{th} \ is \ chosen) = 0 + \left( \frac{1}{N} \sum_{i=1}^{N} X_i^2 - \left( \frac{1}{N} \sum_{i=1}^{N} X_i \right)^2 \right) = \sigma^2.$$

- **(2) 证明：$\mathbb{E}(\overline{X}) = \mu$, $\mathbf{Var}(\overline{X}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$.**

样本均值为：

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

  – **期望**：由于期望的线性性：

$$\mathbb{E}(\overline{X}) = \mathbb{E}\left( \frac{1}{n} \sum_{i=1}^{n} X_i \right) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(X_i).$$

  结合 (1) 中 $\mathbb{E}(X_i) = \mu$，得到：

$$\mathbb{E}(\overline{X}) = \frac{1}{n} \cdot n \cdot \mu = \mu.$$

  – **方差**：

$$\mathrm{Var}(\overline{X}) = \mathrm{Var}\left( \frac{1}{n} \sum_{i=1}^{n} X_i \right) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathrm{Cov}(X_i, X_j).$$

  在无放回抽样的情况下，样本之间存在协方差：

$$\mathrm{Cov}(X_i, X_j) = \begin{cases} \sigma^2, & i = j, \\ EX_iX_j - EX_iEX_j = -\frac{\sigma^2}{N-1}, & i \neq j. \end{cases}$$

  因为:

$$
\begin{aligned}
EX_iX_j &= \sum_{i\neq j} X_iX_j \cdot P(X_i, X_j) = \sum_{i\neq j} X_iX_j \cdot \frac{1}{N(N-1)} \\
&= \frac{1}{N(N-1)}(\sum_{i,j} X_iX_j - \sum_{i=j} X_iX_j) \\
&= \frac{1}{N(N-1)}((\sum_{i=1}^{N} X_i)^2 - \sum_{i=1}^{N} X_i^2) \\
&= \frac{1}{N(N-1)}(N^2\mu^2 - N(\sigma^2+\mu^2)) \\
&= \mu^2 - \frac{1}{N-1}\sigma^2
\end{aligned}
$$

因此:

$$
\mathrm{Var}(\overline{X}) = \frac{1}{n^2}\left(n\cdot\sigma^2 + n(n-1)\cdot\left(-\frac{\sigma^2}{N-1}\right)\right) = \frac{\sigma^2}{n}\left(1-\frac{n-1}{N-1}\right) = \frac{\sigma^2}{n}\cdot\frac{N-n}{N-1}
$$

**4.** 设随机样本 $X_i$ $(i=1,\cdots,n)$ 来自二项总体 $B(k,p)$。

(1) 给出参数 $k$ 和 $p$ 的矩估计。对于二项分布 $B(k,p)$，其期望和方差分别为：

$$
\mathbb{E}(X) = kp, \quad m_2(X) = kp(1-p).
$$

令 $a_1 = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，$m_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$，所以：

$$
\widehat{p} = 1 - \frac{m_2}{a_1}, \quad \widehat{k} = \frac{a_1}{\widehat{p}}.
$$

(2) 尝试讨论上述估计的不足之处。

- 当样本均值 $\overline{X}$ 接近 0 或者存在离群值（例如 $n=10, X_1...X_9=0, X_10=10$），$\widehat{p}$ 为负值。
- 无法保证 $(k > max_iX_i)$。
- 无法保证 k 为整数
- 并非无偏，样本量小时偏差较大。

**5.** 设随机样本 $X_i$ $(i=1,\cdots,n)$ 来自均匀分布 $U(\theta,2\theta)$，求 $\theta$ 的矩估计和极大似然估计。

(1) **矩估计**：对于均匀分布 $U(\theta,2\theta)$，其期望为：

$$
\mathbb{E}(X) = \frac{\theta+2\theta}{2} = \frac{3\theta}{2}.
$$

设样本均值为 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，令

$$
\frac{3\theta}{2} = \overline{X}.
$$

解得：

$$
\widehat{\theta} = \frac{2}{3}\overline{X}.
$$

(2) **极大似然估计**：

对于均匀分布 $U(\theta, 2\theta)$，样本的联合概率密度函数（pdf）为：

$$f(X_1, \cdots, X_n \mid \theta) = \left(\frac{1}{\theta}\right)^n I_{[\theta \le X_{\min}, 2\theta \ge X_{\max}]}.$$

其中 $X_{\min} = \min\{X_1, \cdots, X_n\}$，$X_{\max} = \max\{X_1, \cdots, X_n\}$。

$$\theta \le X_{\min}, \quad 2\theta \ge X_{\max}.$$

要使似然函数最大，$\theta$ 应尽可能小，因此：

$$\widehat{\theta}_{\mathrm{MLE}} = \frac{X_{\max}}{2}.$$

**Tips** 很多同学得到 $min\{X_{\min}, \frac{X_{\max}}{2}\}$ 结果当然是一样的，但是不需要比较（而且我不太明白为什么要这样写，$X_{\min}$ 是上界），而且在给定题设分布下，$\frac{X_{\max}}{2}$ 必然更小

**6.** 设函数

$$f(x; a, \sigma) = (\sqrt{2\pi}\sigma^3)^{-1}(x-a)^2 \exp\left(-\frac{1}{2\sigma^2}(x-a)^2\right), \quad x \in \mathbb{R},$$

其中 $a \in \mathbb{R}$，$\sigma > 0$ 为参数。

(1) 验证 $f(x; a, \sigma)$ 作为 $x$ 的函数满足概率密度的归一化要求。需要验证：

$$\int_{-\infty}^{\infty} f(x; a, \sigma)\, dx = 1.$$

令 $y = \frac{x-a}{\sigma}$，则 $dy = \frac{dx}{\sigma}$，并且有：

$$f(x; a, \sigma) = \frac{1}{\sqrt{2\pi}\sigma^3}(x-a)^2 \exp\left(-\frac{1}{2\sigma^2}(x-a)^2\right) = \frac{\sigma^2 y^2}{\sqrt{2\pi}\sigma^3} \exp\left(-\frac{1}{2}y^2\right).$$

积分变为：

$$\int_{-\infty}^{\infty} f(x; a, \sigma) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{1}{2}y^2} dy.$$

对 $\int_{-\infty}^{\infty} y^2 e^{-\frac{1}{2}y^2} dy$ 分部积分：

$$\int y^2 e^{-\frac{1}{2}y^2} dy = \left[-y e^{-\frac{1}{2}y^2}\right]_{-\infty}^{\infty} + \int e^{-\frac{1}{2}y^2} dy.$$

1.

$$\left[-y e^{-\frac{1}{2}y^2}\right]_{-\infty}^{\infty} = 0,$$

因为当 $y \to \pm\infty$，$y e^{-\frac{1}{2}y^2} \to 0$。

2. 由正态分布的归一性：

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2} dy = \sqrt{2\pi}.$$

因此：

$$\int_{-\infty}^{\infty} y^2 e^{-\frac{1}{2}y^2} dy = \sqrt{2\pi}.$$

代回原式得：

$$\int_{-\infty}^{\infty} f(x;a,\sigma)dx = \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi} \cdot = 1.$$

**Tips: 有同学发现 $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{1}{2}y^2} dy$ 是标准正态的二阶矩，即可直接得出是 1)**

(2) 设随机样本 $X_i$ $(i = 1, \cdots, n)$ 来自此总体，求 $a$ 和 $\sigma^2$ 的矩估计。

**期望：**

令 $y = \frac{x-a}{\sigma}$，则 $x = a + \sigma y$，$dx = \sigma\, dy$，积分变为：

$$\begin{aligned}
\mathbb{E}(X) &= \int_{-\infty}^{\infty} (a + \sigma y) \cdot \frac{y^2}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\, dy \\
&= a \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{1}{2}y^2}\, dy + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^3 e^{-\frac{1}{2}y^2}\, dy \\
&= a \text{ (第一部分利用 pdf 归一性，第二部分为奇函数，在对称区间上的积分为 0)}
\end{aligned}$$

**方差：**

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f(x;a,\sigma)dx.$$

同样令 $y = \frac{x-a}{\sigma}$，积分变为：

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} \left(a^2 + 2a\sigma y + \sigma^2 y^2\right) \cdot \frac{y^2}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} \cdot\, dy.$$

$$\mathbb{E}(X^2) = a^2 \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{1}{2}y^2} dy + 2a\sigma \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^3 e^{-\frac{1}{2}y^2} dy + \sigma^2 \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^4 e^{-\frac{1}{2}y^2} dy.$$

计算各项:- 第一项:$\int_{-\infty}^{\infty} y^2 e^{-\frac{1}{2}y^2} dy = \sqrt{2\pi}$;- 第二项:$\int_{-\infty}^{\infty} y^3 e^{-\frac{1}{2}y^2} dy = 0$;- 第三项:$\int_{-\infty}^{\infty} y^4 e^{-\frac{1}{2}y^2} dy = 3\sqrt{2\pi}$（通过分部积分可递归减少 y 的次数）。

因此：

$$\mathbb{E}(X^2) = a^2 + 3\sigma^2.$$

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = (a^2 + 3\sigma^2) - a^2 = 3\sigma^2.$$

**矩估计：**

$$\widehat{a} = a_1 = \overline{X}, \quad \widehat{\sigma^2} = \frac{m_2}{3}.$$

**Tips** 先计算 $y = \frac{x-a}{\sigma}$ 的方差再变换回 X，这样计算会更简单

(3) 列出 $a, \sigma^2$ 的极大似然估计所满足的方程，并指出一种迭代求解的方法。

样本的对数似然函数为：

$$\ell(a,\sigma^2) = \sum_{i=1}^{n} \ln f(X_i;a,\sigma) = -n\ln(\sqrt{2\pi}\sigma^3) + \sum_{i=1}^{n} \ln(X_i - a)^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (X_i - a)^2.$$

对 $a$ 和 $\sigma^2$ 求偏导，令偏导为 0，得到方程组：

$$\frac{\partial \ell}{\partial a} = -2\sum_{i=1}^{n} \frac{X_i - a}{(X_i - a)^2} + \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - a) = 0,$$

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{3n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(X_i - a)^2 = 0.$$

迭代方法，例如：

(a) 1. 对 $-\ell(a, \sigma^2)$ Gradient Descent(梯度下降)：

- 目标是最小化负对数似然函数 $-\ell(a, \sigma^2)$。
- 更新规则：

$$a^{(t+1)} = a^{(t)} - \eta \frac{\partial(-\ell(a, \sigma^2))}{\partial a}, \quad \sigma^2_{(t+1)} = \sigma^2_{(t)} - \eta \frac{\partial(-\ell(a, \sigma^2))}{\partial \sigma^2},$$

  其中 $\eta$ 是学习率，可以设定为自适应减小。

(b) 2. Coordinate Descent 坐标下降：

- 首先利用初始值（如样本均值 $\overline{X}$ 和样本方差 $S^2$）对 $a$ 和 $\sigma^2$ 进行初始化。
- 固定 $a$，求解 $\sigma^2$ 的更新值：

$$\sigma^2 = \frac{1}{3n}\sum_{i=1}^{n}(X_i - a)^2.$$

- 固定 $\sigma^2$，迭代求解 $a$。

**7.** 设随机样本 $X_i$ $(i = 1, \cdots, n)$ 来自 Bernoulli 总体 $B(p)$，请给出参数 $p$ 的矩估计和极大似然估计。

(1) **矩估计**：对于 Bernoulli 分布 $B(p)$，其期望为：

$$\mathbb{E}(X) = p.$$

设样本均值为 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，令：

$$p = \overline{X}.$$

因此：

$$\widehat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

(2) **MLE**： Bernoulli 分布的 pmf 为：

$$f(X_i; p) = p^{X_i}(1-p)^{1-X_i}, \quad X_i \in \{0, 1\}.$$

对于 $X_1, \cdots, X_n$，似然函数为：

$$L(p) = \prod_{i=1}^{n} f(X_i; p) = \prod_{i=1}^{n} p^{X_i}(1-p)^{1-X_i}.$$

取对数：

$$\ell(p) = \ln L(p) = \sum_{i=1}^{n} \left[X_i \ln p + (1 - X_i) \ln(1 - p)\right].$$

对 $p$ 求导并令其为零：

$$\frac{\partial \ell(p)}{\partial p} = \frac{\sum_{i=1}^{n} X_i}{p} - \frac{n - \sum_{i=1}^{n} X_i}{1 - p} = 0.$$

解得：

$$\widehat{p} = \frac{\sum_{i=1}^{n} X_i}{n}.$$

**8.** 设总体是总数为 $n$，单元概率分别为 $p_1, \cdots, p_m$（这里 $p_1 + \cdots + p_m = 1$）的多项分布，$X_i\ (i = 1, \cdots, m)$ 分别为 $m$ 个单元的观测频数（$X_1 + \cdots + X_m = n$）。求参数 $p_i\ (i = 1, \cdots, m)$ 的极大似然估计。

**MLE**：多项分布的 pmf 为：

$$f(X_1, \cdots, X_m; p_1, \cdots, p_m) = \frac{n!}{X_1! X_2! \cdots X_m!} \prod_{i=1}^{m} p_i^{X_i}.$$

对于样本 $X_1, \cdots, X_m$，似然函数为：

$$L(p_1, \cdots, p_m) = \frac{n!}{X_1! X_2! \cdots X_m!} \prod_{i=1}^{m} p_i^{X_i}.$$

取对数：

$$\ell(p_1, \cdots, p_m) = \ln L = \ln\left(\frac{n!}{X_1! X_2! \cdots X_m!}\right) + \sum_{i=1}^{m} X_i \ln p_i.$$

由于第一项不依赖于 $p_i$，忽略常数部分，优化目标为：

$$\ell(p_1, \cdots, p_m) = \sum_{i=1}^{m} X_i \ln p_i.$$

对每个 $p_i$ 求导并加上约束条件 $\sum_{i=1}^{m} p_i = 1$，使用拉格朗日乘子法：

$$\mathcal{L}(p_1, \cdots, p_m, \lambda) = \sum_{i=1}^{m} X_i \ln p_i + \lambda \left(1 - \sum_{i=1}^{m} p_i\right).$$

对 $p_i$ 求偏导数并令其为零：

$$\frac{\partial \mathcal{L}}{\partial p_i} = \frac{X_i}{p_i} - \lambda = 0, \quad i = 1, \cdots, m.$$

解得：

$$p_i = \frac{X_i}{n}, \quad i = 1, \cdots, m.$$

因此：

$$\widehat{p}_i^{MLE} = \frac{X_i}{n}, \quad i = 1, \cdots, m.$$

**9.** 设总体 $X$ 的分布如下：

| $X$ | 1 | 2 | 3 |
|---|---|---|---|
| $P$ | $\theta^2$ | $2\theta(1-\theta)$ | $(1-\theta)^2$ |

其中 $0<\theta<1$ 是未知参数。已知样本值 $x_1=1, x_2=2, x_3=1$，求 $\theta$ 的矩估计值和极大似然估计值。

(1) **矩估计：**

$$
\begin{aligned}
\mathbb{E}(X) &= 1\cdot\theta^2+2\cdot 2\theta(1-\theta)+3\cdot(1-\theta)^2 \\
&= \theta^2+4\theta(1-\theta)+3(1-\theta)^2 \\
&= \theta^2+4\theta-4\theta^2+3-6\theta+3\theta^2 \\
&= -2\theta+3
\end{aligned}
$$

样本均值为：

$$
\overline{X}=\frac{1+2+1}{3}=\frac{4}{3}.
$$

令 $\mathbb{E}(X)=\overline{X}$，得到：

$$
-2\theta+3=\frac{4}{3}.
$$

因此，$\theta$ 的矩估计值为：

$$
\widehat{\theta}=\frac{5}{6}.
$$

(2) **MLE：**

样本的联合概率为：

$$
\begin{aligned}
L(\theta) &= P(X=x_1)\cdot P(X=x_2)\cdot P(X=x_3) \\
&= \theta^2\cdot 2\theta(1-\theta)\cdot\theta^2 \\
&= 2\theta^5(1-\theta).
\end{aligned}
$$

对数似然函数为：

$$
\ell(\theta)=\ln L(\theta)=\ln 2+5\ln\theta+\ln(1-\theta).
$$

对 $\theta$ 求导：

$$
\frac{d\ell(\theta)}{d\theta}=\frac{5}{\theta}-\frac{1}{1-\theta}.
$$

令导数为零：

$$\frac{5}{\theta} - \frac{1}{1-\theta} = 0,$$

$$\widehat{\theta}_{\mathrm{MLE}} = \frac{5}{6}.$$

**10.** 设随机样本 $X_1, \cdots, X_n$ 来自概率密度函数：

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

其中 $\theta > 0$ 是未知参数。

(1) **矩估计**：

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

设样本均值为 $\overline{X}$，令 $\mathbb{E}(X) = \overline{X}$，则：

$$\frac{\theta}{\theta+1} = \overline{X}.$$

因此：

$$\widehat{\theta} = \frac{\overline{X}}{1-\overline{X}}.$$

(2) **MLE**：

似然函数为：

$$L(\theta) = \prod_{i=1}^{n} f(X_i) = \prod_{i=1}^{n} \theta X_i^{\theta-1} = \theta^n \prod_{i=1}^{n} X_i^{\theta-1}.$$

取对数似然函数：

$$\ell(\theta) = n \ln\theta + (\theta-1) \sum_{i=1}^{n} \ln X_i.$$

对 $\theta$ 求导：

$$\frac{d\ell(\theta)}{d\theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \ln X_i.$$

令导数为零：

$$\frac{n}{\theta} + \sum_{i=1}^{n} \ln X_i = 0,$$

$$\theta = -\frac{n}{\sum_{i=1}^{n} \ln X_i}.$$

因此：

$$\widehat{\theta}_{\mathrm{MLE}} = -\frac{n}{\sum_{i=1}^{n} \ln X_i}.$$

**11.** Python 示例

```python
import numpy as np

def generate_binomial_samples(k, p, n, num_trials=1000):
    """
    Generate binomial samples and calculate the estimated k and p using method of moments.

    Parameters:
    - k (int): Number of trials for the binomial distribution.
    - p (float): Probability of success.
    - n (int): Number of samples in each experiment.
    - num_trials (int): Number of repeated trials to observe results.

    Returns:
    - results (list): List of estimated k and p for each trial.
    """
    results = []
    for _ in range(num_trials):
        # Generate n samples from B(k, p)
        samples = np.random.binomial(k, p, n)

        # Method of moments estimation
        sample_mean = np.mean(samples)
        sample_var = np.var(samples, ddof=1)  # Use ddof=1 for unbiased variance

        # Estimate p and k
        if sample_var != 0:
            p_hat = 1 - sample_var / sample_mean
            k_hat = sample_mean / p_hat
        else:
            p_hat = np.nan  # Avoid division by zero
            k_hat = np.nan
```

```
        results.append((k_hat, p_hat))
    return results

# Example usage with different scenarios:
scenarios = [(10, 0.01, 10), (10, 0.5, 10), (10, 0.01, 1000), (10, 0.5, 1000)]
experiment_results = {}

for k, p, n in scenarios:
    experiment_results[(k, p, n)] = generate_binomial_samples(k, p, n)

experiment_results
```

你可以看到有些情况下会产生负数，例如 (k, p, n)=(10, 0.01, 10), 得到 (k, p) 部分结果：

```
    ...
  (-450359962737049.6, -2.220446049250313e-16),
  (nan, nan),
  (nan, nan),
  (1.8000000000000047, 0.11111111111111083),
  (nan, nan),
  (nan, nan),
  (nan, nan),
  (nan, nan),
  (1.8000000000000047, 0.11111111111111083),
  (-0.19999999999999993, -1.0000000000000004),
  (nan, nan),
  (-450359962737049.6, -2.220446049250313e-16),
  (-450359962737049.6, -2.220446049250313e-16),
  (-450359962737049.6, -2.220446049250313e-16),
  (-450359962737049.6, -2.220446049250313e-16),
  (-0.19999999999999993, -1.0000000000000004),
  (1.8000000000000047, 0.11111111111111083),
  (-450359962737049.6, -2.220446049250313e-16),
  (nan, nan),
  (nan, nan),
  (-450359962737049.6, -2.220446049250313e-16),
  (-450359962737049.6, -2.220446049250313e-16),
  (1.8000000000000047, 0.11111111111111083),
    ...
```

**总结：**

- 矩估计能用低阶矩的时候优先用低阶
- 一般认为矩估计用的 $m_2$ 是未修正的 $\frac{\sum_{i=1}^{n}(X_i-\bar{X})}{n}$
- 一个比较普遍的问题，求估计量得到的结果应当是样本的函数，而不是总体参数。例如不应该是 $\hat{\mu}=EX_i$，而是 $\hat{\mu}=\bar{X}$
- 关于 MLE，严格上是需要验证二阶导小于 0（Hessian matrix 负定）, 但是有如下定理，保证对于指数族分布的极值点（且是参数空间内点）就是 MLE(感兴趣可以了解指数族，目前遇到的分布基本都属于指数族)，以下定理来源于 陈希孺. 数理统计引论：

  **引理 2.6.2.** 设 $\Theta$ 为 $\mathbb{R}^k$ 中的一个开凸集，$f(\theta)=f(\theta_1,\dots,\theta_k)$ 为定义在其上的实函数。对 $\Theta$ 中任一点 $\theta=(\theta_1,\dots,\theta_k)$，$\frac{\partial f}{\partial\theta_i}$ 和 $\frac{\partial^2 f}{\partial\theta_i\partial\theta_j}$ $(i,j=1,\dots,k)$ 都存在有限且满足

  $$-Q(\theta)=\left(-\frac{\partial^2 f}{\partial\theta_i\partial\theta_j}\right)_{k\times k}>0, \tag{2.6.32}$$

  其中 $Q(\theta)$ 为 $k$ 阶方阵，其 $(i,j)$ 元为 $\frac{\partial^2 f}{\partial\theta_i\partial\theta_j}$。则：

  $1^*$ 方程组

  $$\frac{\partial f}{\partial\theta_i}=0,\quad i=1,\dots,k \tag{2.6.33}$$

  在 $\Theta$ 内至多只有一组解 $\theta_0$。

  $2^*$ 如方程组 (2.6.33) 有解 $\theta_0$，则

  $$f(\theta_0)=\sup_{\theta\in\Theta}f(\theta). \tag{2.6.34}$$

  其中 (2.6.32) 是指数族所满足的