

# 2023 春：非参数统计期末考试（回忆版）

## 第一题

- (1) 什么是非参数统计问题？给出一个具体例子。
- (2) 相比参数统计，非参数统计有哪些优势？举出至少两个。

## 第二题

- (1) 写出单样本  $U$  统计量公式，并证明其是无偏估计。
- (2) 对于  $\theta$  分别为均值和方差的两种情况，写出  $h, U$  的具体形式和自由度。
- (3) 写出两样本  $U$  统计量公式。

## 第三题（此题考察次序统计量相关知识。）

- (1) 写出  $R_1, R_2, \dots, R_n$  的联合分布。（不需要证明）
- (2) 写出  $R_i$  的边缘分布。（不需要证明）
- (3) 计算  $R_i$  的均值和方差。参考公式：

$$\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

## 第四题

考虑线性秩统计量：

$$L_n = \sum_{i=1}^n c_{ni} a_n(R_i)$$

其中，

$$a_n(R_i) = b_n \varphi\left(\frac{R_i}{n+1}\right), \quad \varphi: (0, 1) \rightarrow \mathbb{R}$$

- (1) 写出 N 条件。
- (2)  $\varphi$  满足什么条件可以推出  $L_n$  渐近正态？
- (3) 对于  $\varphi(\frac{R_i}{n+1})$ ，写出其 Hájek representation。
- (4) 给出一个线性秩统计量的具体例子，使其满足条件 (1) 和 (2)。

**第五题**

- (1) 写出：通过交叉验证选择带宽的公式。
- (2) 最近邻密度估计的定义为？
- (3) 在非参数回归中，局部多项式回归的定义为？

**第六题**（此题考察统计泛函与影响函数。）

记  $m_p$  为  $p$  分位数，写出统计泛函和影响函数。

**第七题**（此题考察极大经验似然。）

考虑优化问题：

$$\begin{aligned} \max \quad & \prod_{i=1}^n p_i \\ \text{s.t.} \quad & p_i > 0, i = 1, 2, \dots, n \\ & \sum_{i=1}^n p_i = 1 \\ & \sum_{i=1}^n p_i x_i^2 = \mu \end{aligned}$$

记  $x_{(1)}^2 = \min_{i=1, \dots, n} x_i^2$ ,  $x_{(n)}^2 = \max_{i=1, \dots, n} x_i^2$ 。对于  $\mu \in (x_{(1)}^2, x_{(n)}^2)$ ，证明此优化问题的解为：

$$p_i = \frac{1}{n - \lambda(x_i^2 - \mu)}, \quad 1 \leq i \leq n$$

其中  $\lambda$  满足

$$\sum_{i=1}^n \frac{x_i^2 - \mu}{n - \lambda(x_i^2 - \mu)} = 0$$

同时，请证明上述关于  $\lambda$  的方程在区间  $(\frac{n}{x_{(1)}^2 - \mu}, \frac{n}{x_{(n)}^2 - \mu})$  内有唯一解。

**第八题**

在  $x$  接近 0 的时候，如果对其进行正常的核密度估计，会因为负半轴没有样本点而造成估计的偏差。假设  $x = ph$ ,  $0 < p < 1$ ，给出一种修正方法，并计算偏差（在  $h = 0$  处展开到  $h$  阶）。

**注意：**后三题为课外补充内容。