

《概率论与数理统计》习题

第十八讲 频率学派常见点估计方法

1. 设总体分布列/密度函数如下, x_1, x_2, \dots, x_n 是样本, 试求未知参数的矩估计:

(a) $P(X = x) = (x - 1)\theta^2(1 - \theta)^{x-2}, x = 2, 3, \dots, 0 < \theta < 1.$

(b) $p(x; \theta) = (\theta + 1)x^\theta, 0 < x < 1, \theta > 0.$

2. 设总体为 $N(\mu, 1)$, 现对该总体观测 n 次, 发现有 k 次观测值为正, 使用频率替换方法求 μ 的估计。

3. 设总体的概率密度函数如下, x_1, x_2, \dots, x_n 是样本, 试求未知参数的最大似然估计:

(a) $p(x; \theta) = \frac{1}{2\theta}e^{-|x|/\theta}, \theta > 0;$

(b) $p(x; \theta_1, \theta_2) = \frac{1}{\theta_2 - \theta_1}, \theta_1 < x < \theta_2;$

4. 设 x_1, x_2, \dots, x_n 是来自密度函数为 $p(x; \theta) = e^{-(x-\theta)}, x > \theta$ 的总体的样本。

(a) 求 θ 的最大似然估计 $\hat{\theta}_1$, 它是否是相合估计? 是否是无偏估计?

(b) 求 θ 的矩估计 $\hat{\theta}_2$, 它是否是相合估计? 它是否是无偏估计?

5. 众所周知, 双胞胎可分为同卵双胞胎与异卵双胞胎。在一项针对双胞胎的研究中, 研究者关心的是一对双胞胎是同卵双胞胎的概率, 记为 p , 而且研究者也关心一个孩子是男孩的概率, 记为 q . 在这项研究中, 研究者招募到了 n 对双胞胎 (包括龙凤胎), 其中 n_1 是两个男孩的双胞胎, n_2 是两个女孩的双胞胎, $n_3 = n - (n_1 + n_2)$ 是龙凤胎 (一个男孩一个女孩)。另外, 研究者虽然知道不同性别的双胞胎一定不是同卵双胞胎, 但并不知道其中哪些相同性别的双胞胎为同卵双胞胎。

(a) 采用 EM 算法, 写出 $\theta = (p, q)$ 的最大似然估计的形式。(提示: 构造合适的潜变量, 定义似然函数, 写出 EM 算法中的 E 步和 M 步)

- (b) (编程题) 在 $n = 1000, n_1 = 432, n_2 = 232, n_3 = 336$ 时, 编程计算出 θ 的最大似然估计。