

《概率论与数理统计》第四周要点

复习要点：离散型随机变量及其分布

1. 随机变量

- 定义：随机变量是定义在样本空间上的实值函数，用来刻画随机现象的数值特征。
- 分类：
 1. **离散型**：可能取有限或可列无穷多个值。
 2. **连续型**：可能取某一区间内的所有实数。
- 本节重点讨论离散型随机变量及其概率分布。

—

2. 两点分布 (0-1 分布/伯努利分布)

- 定义：若随机变量 X 只可能取 0 与 1 两个值，且

$$P\{X = 1\} = p, \quad P\{X = 0\} = 1 - p, \quad 0 < p < 1,$$

则称 X 服从两点分布 (或伯努利分布)，记作 $X \sim \text{Bernoulli}(p)$ 。

- 应用场景：一次试验是否成功、是否合格、是否患病。

典型应用 投一枚均匀硬币，令 $X = 1$ 表示正面， $X = 0$ 表示反面，则 $X \sim \text{Bernoulli}(0.5)$ 。

—

3. 二项分布

- 定义：在 n 次独立重复试验中，每次事件 A 发生的概率为 p ，记 $X =$ “事件 A 发生的次数”，则

$$P\{X = k\} = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, \dots, n.$$

记作 $X \sim B(n, p)$ 。

- 特殊情形： $n = 1$ 时退化为两点分布。

典型应用 产品合格率 $p = 0.95$, 抽 $n = 10$ 件, $X =$ 合格品数。

$$P\{X = 8\} = \binom{10}{8} (0.95)^8 (0.05)^2.$$

4. 泊松分布

- 定义: 若 X 取 $0, 1, 2, \dots$, 且

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

则 X 服从参数 λ 的泊松分布, 记作 $X \sim \text{Poisson}(\lambda)$ 。

- 应用: 单位时间事故数、顾客到达数、书页排版错误数、放射粒子计数等。

典型应用 某机器平均每小时出现 2 次故障, 问 1 小时内无故障的概率:

$$P\{X = 0\} = e^{-2}.$$

5. 二项分布与泊松分布的关系

- 当 n 很大、 p 很小, 且 $np = \lambda$ 有限时, 有近似关系:

$$B(n, p) \approx \text{Poisson}(\lambda).$$

- 意义: 泊松分布可作为稀有事件的近似模型。

典型应用 零件不合格率 $p = 0.001$, 检验 $n = 2000$ 件。 $X \sim B(2000, 0.001)$, 近似 $X \sim \text{Poisson}(2)$ 。

6. 几何分布

- 定义: 在一系列独立重复试验中, 设事件 A 每次发生概率为 p 。令 $X =$ “第一次发生 A 所需的试验次数”, 则

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

记作 $X \sim \text{Geom}(p)$ 。

- 记忆无关性: 几何分布是唯一具有“无记忆性”的离散分布。

典型应用 掷骰子直到第一次出现 6 点, $p = 1/6$, 则 $X \sim \text{Geom}(1/6)$ 。

7. 分布函数 (Distribution Function)

- 定义: 随机变量 X 的分布函数为

$$F(x) = P\{X \leq x\}, \quad -\infty < x < \infty.$$

- 基本性质:

1. 单调非减;
2. 右连续;
3. $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow +\infty} F(x) = 1$ 。

- 离散型随机变量: 分布函数呈阶梯状。
- 连续型随机变量: 分布函数连续可导, 其导数为概率密度函数。

例 若 $X \sim B(2, 0.5)$, 则 $P(X = 0) = 0.25, P(X = 1) = 0.5, P(X = 2) = 0.25$ 。分布函数为:

$$F(x) = \begin{cases} 0, & x < 0, \\ 0.25, & 0 \leq x < 1, \\ 0.75, & 1 \leq x < 2, \\ 1, & x \geq 2. \end{cases}$$

8. 小结

- 两点分布: 一次试验成败。
- 二项分布: 多次独立重复试验中的成功次数。
- 泊松分布: 单位时间或空间稀有事件次数; 是二项分布的极限。
- 几何分布: 第一次成功所需试验次数, 具有无记忆性。
- 分布函数: 统一描述随机变量分布规律的重要工具。