Part III – Ch03: 离散型随机变量

试验结果数值化

- 有些随机试验的结果本身就是数值
  - 抛一枚骰子的点数: $1, 2, \ldots, 6$
  - 国家一年出生的婴儿数: $1, 2, \ldots, n$
- 有些试验结果看起来与数值无关, 但可以用数值来表示
  - 抛硬币: 用 1 表示正面朝上, 用 0 表示正面朝下
  - 针对一般事件 A, 可以定义

$$X = \begin{cases} 1, & \text{事件 } A \text{ 发生}, \\ 0, & \text{事件 } A \text{ 不发生}. \end{cases}$$

- 事件驱动的 (event-driven)

随机变量

试验结果用数值表示, 引入变量来表示随机事件——随机变量。

定义 0.15 设 Ω 是一个样本空间, 如果对每个基本事件 $\omega \in \Omega$, 都对应于一个实数 $X(\omega)$, 称这样的单射实值函数 $X : \Omega \to \mathbb{R}$ 为 随机变量 (random variable), 一般简写为 X.

随机变量与普通函数存在着本质的不同:

- $X(\omega)$ 随样本点 ω 的不同而取不同的值, 具有一定的随机性
- 各试验结果的出现具有一定的概率, X 的取值具有统计规律性

通过随机变量来描述随机现象或随机事件, 可以利用各种数学分析工具, 通过对随机变量的研究来揭示随机现象.

随机变量

根据取值类型, 可以将随机变量进行分类:

- 离散型随机变量: X 的取值是有限的、无限可列的
  - 抛一枚骰子的点数: $1, 2, \ldots, 6$ ——有限的

X	1	2	3	4	5	6

  - 国家一年出生的婴儿数: $1, 2, \ldots, n$ ——有限的或者无限可列的

X	1	2	$\ldots$	n	$\ldots$

- 非离散型随机变量: X 的取值是无限不可列的

Review of Discrete Random Variable - 1

October 6, 2023

离散型随机变量 – 分布列

分布列包含随机变量的取值和概率, 可以完全刻画其概率属性

X	x_1	x_2	$\dots$	x_n	$\dots$
P	p_1	p_2	$\dots$	p_n	$\dots$

- 分布列: 对于 X 所有的取值 x_k, 其概率为 $p_k = P(X = x_k)$
- 性质:
  - 对于任意的 k 都有 $p_k \geq 0$
  - $\sum_k p_k = 1$

分布列：例 0.50

例 0.50 设随机变量 X 的分布列为

$$P(X=k)=\frac{c}{4^k} \quad 且 \quad k \geq 0\,.$$

求概率 $P(X=1)$.

解答：例 0.50

题目: 设随机变量 X 的分布列为

$$P(X=k)=\frac{c}{4^k} \quad \text{且} \quad k\geq 0\,.$$

求概率 $P(X=1)$.

解答:

- 根据概率的完备性有

  $$1=\sum_{k=0}^{\infty}P(X=k)=\sum_{k=0}^{\infty}\frac{c}{4^k}=\frac{4}{3}c,$$

  可得 $c=3/4$, 从而得到 $P(X=1)=3/16$.

分布列：例 0.51

例 0.51 给定常数 $\lambda > 0$, 随机变量 X 的分布列为

$$P(X = k) = \frac{c\lambda^k}{k!} \quad 且 \quad k \geq 0\ .$$

- 求概率 $P(X = 1)$;
- 求概率 $P(X > 2)$.

解答：例 0.51

题目: 给定常数 $\lambda > 0$, 随机变量 X 的分布列为

$$P(X=k) = \frac{c\lambda^k}{k!} \quad \text{且} \quad k \geq 0\ .$$

求概率 $P(X=1)$ 和 $P(X>2)$.

解答:

- 根据概率的完备性有

  $$1 = \sum_{k=0}^{\infty} P(X=k) = c\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = c\cdot e^{\lambda},$$

  可得 $c = e^{-\lambda}$.
- 进一步得到 $P(X=1) = \lambda e^{-\lambda}$, 及

$$P(X>2) = 1 - P(X\leq 2) = 1 - p_0 - p_1 - p_2 = 1 - e^{-\lambda}(1+\lambda+\lambda^2/2)$$

离散型随机变量 – 数字特征

分布列 能一目了然的看出离散型随机变量得取值及相应的概率.

- 例 0.50 及例 0.51

针对一个具体的问题, 完全求解出概率分布列可能不是一件容易的事.

- 统计某个区域的工资水平, 包括不同职业不同人群得收入——不容易完全求解
- 在例 0.51中, 求解 $P(X > 2)$ 不需要依赖于 $P(X = 3), P(X = 4), \ldots$ ——不需要完全求解

通过一些指标来描述离散随机变量的总体特征, 或称为统计量.

- 期望、方差, 相关系数, 矩等.

离散型随机变量 – 期望

定义 0.16 设随机变量 X 的分布列为 $P(X = x_k) = p_k (k \geq 1)$. 若级数 $\sum_k p_k x_k$ 绝对收敛, 则称级数 $\sum_k x_k p_k$ 为随机变量 X 的期望, 记为 $\mathbb{E}(X)$

$$\mathbb{E}(X) = \sum_k x_k p_k$$

备注:

- 期望是常量不是变量.
- 期望反映随机变量 X 的平均值.
- 期望的性质:
  1. 设 $c \in \mathbb{R}$, 若随机变量 $X \equiv c$, 则 $\mathbb{E}(X) = c$, 不考虑其分布列.
  2. 线性: 对随机变量 X 及常数 $a, b \in \mathbb{R}$, $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$
- 除非特别说明, 通常直接利用定义计算期望, 不考虑绝对收敛性。

分布列：例 0.52

例 0.52 随意抛掷一枚骰子, X 为掷出的点数. 试求 $\mathbb{E}(X)$.

X	1	2	3	4	5	6
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

解答：例 0.52

题目: 随意抛掷一枚骰子, X 为掷出的点数. 试求 $\mathbb{E}(X)$.

X	1	2	3	4	5	6
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

解答:

- 根据期望的定义有

$$\mathbb{E}(X) = \frac{1+2+3+4+5+6}{6} = \frac{7}{2}\,.$$

期望：例 0.53

例 0.53 设随机变量 X 的分布列为

$$P\left(X=\frac{(-2)^k}{k}\right)=\frac{1}{2^k},\quad k\in\mathbb{N}^+,$$

求期望 $\mathbb{E}(X)$.

解答：例 0.53

题目: 设随机变量 X 的分布列为 $P\left(X=\frac{(-2)^k}{k}\right)=\frac{1}{2^k}, k\in\mathbb{N}^+$, 求期望 $\mathbb{E}(X)$.

解答:

- 根据期望的定义有

$$\mathbb{E}(X)=\sum_{k=1}^{+\infty}P\left(X=\frac{(-2)^k}{k}\right)\frac{(-2)^k}{k}=\sum_{k=1}^{+\infty}\frac{(-1)^k}{k}=-\ln 2.$$

- 但是

$$\sum_{k=1}^{+\infty}P\left(X=\frac{(-2)^k}{k}\right)\left|\frac{(-2)^k}{k}\right|=\sum_{k=1}^{+\infty}\frac{1}{k}\to+\infty.$$

  该级数并非绝对收敛, 其级数和可能随着求和顺序的改变而改变, 级数和并非唯一的数值, 故该随机变量的期望 $\mathbb{E}(X)$ 不存在.

期望：例 0.54

例 0.54 有 n 把钥匙只有一把能打开门。随机选取一把试开门, 若打不开则除去, 求打开门次数的平均数.

解答：例 0.54

题目: 有 n 把钥匙只有一把能打开门。随机选取一把试开门, 若打不开则除去, 求打开门次数的平均数.

解答:

- 设随机变量 X 表示尝试开门的次数, 其分布列为

$$P(X=k)=\frac{\binom{n-1}{k-1}}{\binom{n}{k-1}}\cdot\frac{1}{n-k+1}=\frac{1}{n}\qquad k\in[n].$$

- 根据期望的定义有

$$\mathbb{E}(X)=\sum_{k=1}^{n}\frac{(k)}{n}=\frac{(1+n)n}{2n}=\frac{n+1}{2}.$$

期望：例 0.55

例 0.55 有 4 个盒子编号分别为 1, 2, 3, 4。将 3 个不同的球随机放入 4 个盒子, 用 X 表示有球盒子的最小号码, 求 $\mathbb{E}(X)$.

解答：例 0.55

题目: 有 4 个盒子编号分别为 1, 2, 3, 4。将 3 个不同的球随机放入 4 个盒子, 用 X 表示有球盒子的最小号码, 求 $\mathbb{E}(X)$.

解答:

- 分布列:
  - $X=1$ 时, 分 1 号盒子中有 1 个球、2 个球、3 个球三种情况讨论 $C_3^1 3^2+C_3^2 C_3^1+1$
  - $X=2$ 时, 分 2 号盒子中有 1 个球、2 个球、3 个球三种情况讨论 $C_3^1 2^2+C_3^2 C_2^1+1$
  - $X=3$ 时, 分 3 号盒子中有 1 个球、2 个球、3 个球三种情况讨论 $C_3^1 1^2+C_3^2 C_1^1+1$
  - $X=4$ 时, 化简为子问题 “三个球都放在编号为 4 的盒子里”

X	1	2	3	4
P	$\frac{37}{4^3}$	$\frac{19}{4^3}$	$\frac{7}{4^3}$	$\frac{1}{4^3}$

- 期望: $\mathbb{E}(X)=\frac{100}{64}=\frac{25}{16}$

Appendix: 作业

一个口袋中有 5 个球, 在这 5 个球上分别标有数字 $1, 2, 3, 4, 5$. 从袋中不放回任取 3 个球, 每个球被取到的可能性相同, 求取到的球上标明的最大数字 X 的分布列.

计算: 函数期望

定理 0.6 设离散型随机变量 X 的分布列是 $P(X = x_k) = p_k > 0(k \geq 1)$. 对任意的实值函数 $g : \mathbb{R} \to \mathbb{R}$. 若级数 $\sum_k g(x_k)p_k$ 绝对收敛, 则有

$$\mathbb{E}[g(X)] = \sum_k g(x_k)p_k$$

备注: 求 $Y = g(X)$ 的期望, 不需 Y 分布列. 使用 X 的分布列即计算 $\mathbb{E}(Y)$.

计算: 函数期望

设离散型随机变量 X 的分布列是 $P(X = x_k) = p_k > 0(k \geq 1)$, 以及 $g_i : \mathbb{R} \to \mathbb{R}$ 是实值函数 ($i \in [n]$). 若级数 $\sum_k g_i(x_k)p_k$ 绝对收敛, 则对任意常数 $c_1, c_2, \ldots, c_n$, 有

$$\mathbb{E}\left[\sum_{i=1} c_i g_i(X)\right] = \sum_i c_i \mathbb{E}[g(x_k)] = \sum_i \sum_k c_i g_i(x_k) p_k$$

备注:

- 定理 0.6 及期望的线性性
- 例题: $\mathbb{E}\left[X^2 + X + \sin X + 4\right]$

凸函数和凹函数

定义 0.17 (凸函数) 设函数 $g:[a,b]\to\mathbb{R}$, 对任意 $x_1,x_2\in[a,b]$ 和 $\lambda\in[0,1]$. 若

$$g\left(\lambda x_1+(1-\lambda)x_2\right)\leq\lambda g(x_1)+(1-\lambda)g(x_2),$$

则称 g 是定义在 $[a,b]$ 的凸函数或者下凸函数.

定义 0.18 (凹函数) 设函数 $g:[a,b]\to\mathbb{R}$, 对任意 $x_1,x_2\in[a,b]$ 和 $\lambda\in[0,1]$. 若

$$g\left(\lambda x_1+(1-\lambda)x_2\right)\geq\lambda g(x_1)+(1-\lambda)g(x_2),$$

则称 g 是定义在 $[a,b]$ 的凹函数或者上凸函数.

Jensen 不等式

定理 0.7 对随机变量 $X \in [a,b]$ 和若实值函数 $f:[a,b] \to \mathbb{R}$ 是凸函数, 实值函数 $g:[a,b] \to \mathbb{R}$ 是凹函数, 则有

$$f\left[\mathbb{E}(X)\right] \leq \mathbb{E}\left[f(X)\right]$$

和

$$g\left[\mathbb{E}(X)\right] \geq \mathbb{E}\left[g(X)\right]$$

结论:

$$\left(\mathbb{E}(X)\right)^2 \leq \mathbb{E}(X^2)$$
$$e^{\mathbb{E}(X)} \leq \mathbb{E}\left[e^X\right]$$
$$\ln \mathbb{E}(X) \geq \mathbb{E}\left[\ln X\right]$$

离散型随机变量 – 方差

定义 0.19 设随机变量 X 的分布列为 $P(X=x_k)=p_k>0(k\geq 1)$. 若期望 $\mathbb{E}(X)$ 和

$$\mathbb{E}\left[(X-\mathbb{E}(X))^2\right]=\sum_{k\geq 1}p_k\left(x_k-\sum_k p_kx_k\right)^2,$$

存在, 则称 $\mathbb{E}\,(X-\mathbb{E}(X))^2$ 为随机变量 X 的方差, 记为 $\mathbb{VAR}(X)$. 称

$$\sigma(X)=\sqrt{\mathbb{VAR}(X)} \quad \text{或} \quad \mathbb{D}(X)=\sqrt{\mathbb{VAR}(X)}$$

为标准差.

备注:

- 方差是常量而不是变量.
- 方差反映随机变量 X 与其期望 $\mathbb{E}(X)$ 的偏离程度.

离散型随机变量方差的性质

1. $\mathbb{VAR}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$
2. 设 $c \in \mathbb{R}$ 是常数, 若随机变量 $X \equiv c$, 则 $\mathbb{VAR}(X) = 0$
3. 对随机变量 X 和常数 $a, b \in \mathbb{R}$, 有 $\mathbb{VAR}(aX + b) = a^2\mathbb{VAR}(X)$
4. 方差不具有线性性:

$$\mathbb{VAR}(f(X) + g(X)) \neq \mathbb{VAR}(f(X)) + \mathbb{VAR}(g(X)) \text{ , for } f, g \in \mathcal{C}$$

5. 对随机变量 X 和常数 a, 有 $\mathbb{VAR}(X) = \mathbb{E}\left(X - \mathbb{E}(X)\right)^2 \leq \mathbb{E}(X - a)^2$
6. Bhatia-Davis 不等式: 对随机变量 $X \in [a, b]$, 有

$$\mathbb{VAR}(X) \leq (b - \mathbb{E}(X))(\mathbb{E}(X) - a) \leq \frac{(b-a)^2}{4}$$