Ch 3 离散型随机变量

回顾前一次课

独立性：两个事件、多个事件相互独立性

相互独立性　　两两独立性

事件的独立　　事件的互斥（互不相容）

独立性的性质，以及如何判断独立性

小概率原理： 若事件A在一次试验中发生的概率非常小, 但经过多次独立地重复试验, 事件A的发生是必然的

应用案例：判断大型矩阵$AB = C$?

随机变量

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：$1, 2, \ldots, n, \ldots$

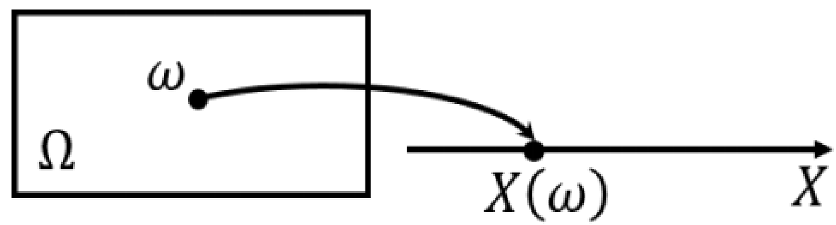
有些试验结果可能与数值无关, 但结果可以用数值进行表示

- 抛一枚硬币, 正面朝上用0表示, 正面朝下用1表示
- 流星坠落地球的落脚点用坐标纬度表示

试验结果用数值表示, 引入变量来表示随机事件 —— 随机变量

随机变量形式化表述

将样本空间Ω中每个样本点ω与一个实数$X(\omega)$ 相对应, $X(\omega)$ 是ω的实值函数, 称实值函数 $X(\omega): \Omega \rightarrow \mathrm{R}$为随机变量(random variable), 简写为 r.v.



一般用大写字母X,Y,Z表示

$X(\omega)$ 随样本点ω的不同而取不同的值, 例如

- 抛一枚骰子, 用随机变量X表示出现的点数, 则随机变量$X \in [6]$.
  - 出现的点数不超过 4 的事件可表示为 $\{X \leq 4\}$
  - 出现偶数点的事件可表示为 $\{X为偶数\}$
- 用随机变量X表示一盏电灯的寿命, 其取值为 $[0, +\infty)$, 电灯寿命不超过500的事件可表示为 $\{X \leq 500\}$

离散型随机变量

离散型随机变量：随机变量的取值是有限的、或无限可列的

非离散型随机变量：随机变量的取值无限不可列的

连续型随机变量：后面讲

离散型随机变量X的取值是有限、或可列的, 不妨假设其取值为$x_1, x_2, \cdots, x_n, \cdots$, 事件$X = x_k$的概率记为

$$p_k = P(X = x_k) \qquad k = 1, 2, \cdots$$

称之为**随机变量X的分布列**,

分布列包含随机变量的取值和概率, 完全刻画其概率属性

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

分布列的性质及其应用

性质： 随机变量X的分布列$P(X=x_k)=p_k\ \ (k\geq 1)$满足$p_k\geq 0$且$\sum_k p_k=1$

例:若随机变量X的分布列$P(X=k)=c/4^k\ \ (k\geq 0)$, 求$P(X=1)$

例题

给定常数$\lambda > 0$, 随机变量X的分布列$p_i = \ c\lambda^k/k! \ \ (k \geq 0)$, 求$P(X > 2)$

从 $1,2,\ldots,10$中不放回随机任意取5个数, 令随机变量X表示所取5个数中的最大值, 求X的分布列

离散变量的期望

随机变量具有一定的随机性, 我们希望研究随机变量的不变性, 以刻画随机变量的特征, 最常见的特征是期望与方差

定义： 离散随机变量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, 称级数和为随机变量X的期望 (expectation), 又被称为**均值(mean)**, 记为E(X)

$$E(X) = \sum_k p_k x_k$$

期望$E(X)$ 反映随机变量的平均值, 由随机变量的分布列决定, 是常量而不是变量, 根据随机变量的值x_i根据概率p_i加权所得

级数的绝对收敛保证了级数和不随级数各项次序的改变而改变, 期望$E(X)$反映X可能值的平均值, 不会随次序改变而改变

例题

随机掷一枚骰子, X表示观察到的点数, 求$E[X]$

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$E(X)$

例题

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.

期望的性质

若随机变量$X \equiv c$, 则$E(c) = c$

对随机变量X和常数$a, b \in \mathrm{R}$, 有$E(aX + b) = aE(X) + b$

函数期望的计算

定理: 设X为离散型随机变量, 以及$g: R \to R$是连续函数, 若X的分布列为$P(X = x_k) = p_k \ (k \geq 1)$, 且$\sum_{k\geq 1} g(x_k)p_k$绝对收敛, 则有

$$E[g(X)] = \sum_{k\geq 1} g(x_k)p_k$$

求$\boldsymbol{Y = g(X)}$的期望, 不需$\boldsymbol{Y}$分布列, 而$\boldsymbol{X}$的分布列可计算$\mathbf{E}[\boldsymbol{Y}]$

函数期望的计算

设X为离散型随机变量, 以及函数$g_i: R \to \mathrm{R}(\mathrm{i} \in [\mathrm{n}])$是连续函数,
且$E\left(g_i(X)\right)$存在. 对任意常数$c_1, c_2, \ldots, c_n$有

$$E\left(c_1 g_1(X) + c_2 g_2(X) + \cdots + c_n g_n(X)\right) = \sum_{i=1}^{n} c_i E\left(g_i(X)\right)$$

$E(X^2 + X + sinX + 4)$=

凸函数

设函数$g: [a,b] \to \mathrm{R}$, 若对任意$x_1, x_2 \in [a,b]$ 和$\lambda \in [0,1]$, 有

$$g(\lambda x_1 + (1-\lambda)x_2) \le \lambda g(x_1) + (1-\lambda) g(x_2)$$

成立, 称函数$g(x)$是定义在 $[a,b]$ 的凸函数

若对任意$x_1, x_2 \in [a,b]$和$\lambda \in [0,1]$, 有

$$g(\lambda\, x_1 \; + \; (1-\lambda)x_2) \ge \lambda\, g(x_1) + (1-\lambda)\, g(x_2)$$

成立, 称函数$g(x)$是定义在$[a,b]$的凹函数

Jensen不等式

X为$[a,b]$的离散型随机变量, $g:[a,b]\rightarrow \mathrm{R}$是连续的凸函数, 有

$$g\big(E(X)\big)\leq\ E\big(g(X)\big)$$

Jensen不等式

设X为$[a,b]$的离散型随机变量,若$g:[a,b] \to \mathrm{R}$是连续的凹函数,则有

$$g\big(E(X)\big) \geq E\big(g(X)\big)$$

对任意离散型随机变量X, 根据Jensen不等式有

$$\big(E(X)\big)^2 \leq E(X^2)$$

$$e^{E(X)} \leq E(e^X)$$

$$\ln E(X) \geq E(\ln X)$$