

概率论与数理统计



第

5

讲

随机变量、离散型随机变量及其分布律

为了研究随机现象的统计规律性，学习了如下基本概念：

E ：随机试验

S ：样本空间

A ：随机事件

\mathcal{F} ：将随机事件作为元素，组成一个大的集合，称其为事件域。

$$\begin{aligned} P: \mathcal{F} &\rightarrow R^1 \\ A &\rightarrow P(A) \end{aligned}$$

满足三条公理

非负性
规范性
可列可加性

此时求 $P(A)$ ，只是孤立的研究一个个事件，对 S 的全貌不了解。

同时, A 是集合, $P(A)$ 是数, 无法用图形和其他数学工具, 对其研究受到限制。

因此, 为了深入研究随机现象, 认识随机现象的整体性质, 需要全面地研究随机试验 E 中事件的概率。



那么, 如何系统而全面地描述 E 的随机事件呢?

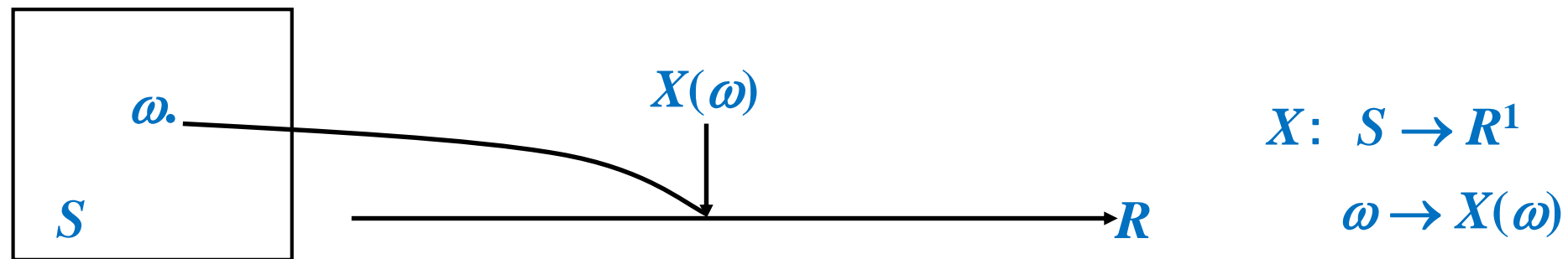


能否引入一个变量（取值为数），当它取不同的值时，可以表示不同的随机事件？

S 的某些样本点组成的集合

因此，只要将随机试验的样本点数量化，就可以将任意随机试验的结果数量化。

为此，只需要引入一个映射关系，将样本空间中的每个样本点与实数对应起来。



随机试验的结果有两类：

定量的

定性的

试验结果是定量的，表示试验结果本身就是数值

如：手机上每天收到的微信条数；在任取的 n 件产品中，合格品的个数；每天观看手机的时长；任意取到的一件电子产品的寿命。

试验结果是定性的，表示该结果本身不是数值

如：抛掷一枚有正反两个面的硬币，其结果是正面朝上或者反面朝上；任取一件产品，结果是合格品或者不合格品。

例1. 掷一枚硬币，观察其面朝上的情况(E) 样本空间: $S=\{\text{正面}, \text{反面}\}$

定义映射: $X: S \rightarrow R^1$ 满足: $X(\text{正面})=1, X(\text{反面})=0$

其中 $\{\omega: X(\omega)=1\}=\{\text{正面}\}, \{\omega: X(\omega)=0\}=\{\text{反面}\}$

X 的取值是随机的，所有可能的取值为 $\{0, 1\}$

X 为掷一枚硬币，出现正面的次数

例2. 从某型电子元件中任抽一件，观测其寿命(E) 样本空间, $S = \{t : t \geq 0\}$

定义映射: $X: S \rightarrow R$
 $t \rightarrow t$

X 为任抽一电子元件的寿命

其可能取值的范围为 $[0, +\infty)$

X 在某一范围内的取值可以表达 E 中的事件。如: $\{\omega: X(\omega) \in [a, b]\} = \{t: t \in [a, b]\}$

定义 设 (S, \mathcal{F}, P) 是一概率空间, 若 X 为样本空间 S 到实数域 R^1 上的映射:

$$\begin{aligned} X: S &\rightarrow R^1 \\ \omega &\rightarrow X(\omega) \end{aligned}$$

满足: $\forall x \in R^1$, 有 $\{\omega: X(\omega) \leq x\} \in \mathcal{F}$

则称 $X(\omega)$ 为 (S, \mathcal{F}, P) 上的一个随机变量。常常将 $\{\omega: X(\omega) \leq x\}$ 简记为 $\{X \leq x\}$ 。

随机变量通常用大写字母 X, Y, Z 或希腊字母 ξ, η 等表示, 而随机变量所取的值, 一般用小写字母 x, y, z 等表示。

注:

01

随机变量 X 随试验结果的不同而取不同的值，因而在试验之前只知道它可能取值的范围，而不能预先肯定它将取哪个值；

02

随机变量 X 的自变量是样本点，可以是数，也可以不是数，定义域是样本空间，与一般实函数的定义域不同；

03

由于试验结果的出现具有一定的概率，于是这种实值函数取每个值和每个确定范围内的值也有一定的概率。

引入随机变量 X 以后，就可以用 X 来描述事件。一般地，设 L 是实数域上一集合，将 X 在 L 上的取值写成 $\{X \in L\}$ ，它表示事件

$$\text{即 } \{X \in L\} = \{\omega: X(\omega) \in L\}$$

随机事件这个概念实际上是包容在随机变量这个更广的概念之内，也可以说随机事件是从静态的观点来研究随机现象，而随机变量则是一种动态的观点，因此，引入随机变量后，对随机现象统计规律性的研究，就从研究一些孤立的事件及其概率扩大为研究随机变量及其取值规律，从而发展成为了一个更高的理论体系。



在随机变量中有一类简单的随机变量，它们的可能的取值只有有限个，如被访问者的性别、年龄、职业，掷一枚骰子掷出的点数，一批产品中次品的个数等；

还有一些量，理论上讲能取无限个值，但是这些值都可以毫无遗漏的一个接一个的排列出来。例如，手机一天中收到的微信条数，某网站一天中被点击的次数，一个医学试样中白细胞的个数，打靶时首次命中目标时的射击次数等。

这种类型的随机变量我们称之为**离散型随机变量**

定义

若随机变量 X 所有可能的取值为有限个或可数个，则称 X 为离散型随机变量，简称为离散随机变量。

定义：若随机变量 X 所有可能的取值为 $x_1, x_2, \dots, x_n, \dots$ ，且 X 取这些值的概率为

$$P\{X=x_i\}=p_i, i=1, 2, \dots \quad (1)$$

则称(1)式为离散型随机变量 X 的概率分布，或称为 X 的分布律。

也可以用表格的形式表示

X	x_1	x_2	\dots	x_i	\dots
P	p_1	p_2	\dots	p_i	\dots

已知离散型随机变量 X 的分布律后，可以计算 X 取某值或落入某实数集合内的概率。它完全描述了 X 取值的概率规律。

“分布” 可以理解为全部概率1是如何分布在(分配到)随机变量 X 各个可能值 x_i 上的。

性质: (1) $p_i \geq 0, i=1, 2, \dots$ (2) $\sum_i p_i = 1$

$$X: S \rightarrow R^1 \\ \omega \rightarrow X(\omega)$$

证明: 令 $A_i = \{X = x_i\} = \{\omega : X(\omega) = x_i\}, i = 1, 2, \dots$

易知 $A_i \cap A_j = \emptyset, i \neq j$ 且 $\bigcup_i A_i = S$

$$\sum_i p_i = \sum_i P(A_i) = P\left(\bigcup_i A_i\right) = P(S) = 1$$

注: $\{X=x_i\}, i=1, 2, \dots$ 构成了 S 的一个划分。

例3.一汽车沿一街道行驶，需要通过四个均设有信号灯的路口（忽略黄灯），设汽车在每个路口遇到红灯的概率为 p ，各信号灯的工作是相互独立的。以 X 表示该汽车首次停下时，它已通过的路口的个数，求 X 的分布律。

解：易知 X 所有可能的取值为：0, 1, 2, 3, 4

设 A_k 表示汽车在第 k 个路口遇到绿灯， $k=1,2,3,4$ ，则它们互相独立

于是： $P\{X = 0\} = P(\bar{A}_1) = p$

$$P\{X = 1\} = P(A_1\bar{A}_2) = P(A_1)P(\bar{A}_2) = (1-p)p$$

$$P\{X = 2\} = P(A_1A_2\bar{A}_3) = P(A_1)P(A_2)P(\bar{A}_3) = (1-p)^2 p$$

例3. 一汽车沿一街道行驶，需要通过四个均设有信号灯的路口（忽略黄灯），设汽车在每个路口遇到红灯的概率为 p ，各信号灯的工作是相互独立的。以 X 表示该汽车首次停下时，它已通过的路口的个数，求 X 的分布律。

$$P\{X = 3\} = P(A_1 A_2 A_3 \bar{A}_4) = P(A_1)P(A_2)P(A_3)P(\bar{A}_4) = (1-p)^3 p$$

$$P\{X = 4\} = P(A_1 A_2 A_3 A_4) = P(A_1)P(A_2)P(A_3)P(A_4) = (1-p)^4$$

所以 X 的分布律为

X	0	1	2	3	4
P	p	$(1-p)p$	$(1-p)^2 p$	$(1-p)^3 p$	$(1-p)^4$

例4. 设一袋子中装有10个球，其中编号为0的球4个，编号为1的球2个，编号为2的球4个，且球的形状和大小相同。设 X 表示任取一球的号码数。求
(1) X 的分布律；(2) $P\{X \leq 1/2\}$, $P\{1 \leq X \leq 3\}$.

解: (1) X 可能的取值为0, 1, 2

$$\text{且 } P\{X=0\}=4/10=2/5 \quad P\{X=1\}=2/10=1/5 \quad P\{X=2\}=4/10=2/5$$

所以 X 的分布律为

X	0	1	2
P	$2/5$	$1/5$	$2/5$

二. 离散型随机变量及其分布律

X 表示任取一球的号码数。

(2) $P\{X \leq 1/2\}$, $P\{1 \leq X \leq 3\}$

X	0	1	2
P	2/5	1/5	2/5

解: (2) $P\{X \leq 1/2\} = P\{X=0\} = 2/5$ $P\{1 \leq X \leq 3\} = P\{X=1 \cup X=2\} = P\{X=1\} + P\{X=2\} = 3/5$

号码数 $\leq 1/2$

号码数在1和3之间

一般地, 设 L 是实数域上一集合, 则有 $P\{X \in L\} = \sum_{x_i \in L} P\{X = x_i\} = \sum_{x_i \in L} p_i$

例5 设离散型随机变量 X 的分布律为 $P\{X = i\} = \frac{a}{i(i+1)}, i = 1, 2, \dots$

求 (1) 常数 a 的值; (2) $P\{X < 5\}$.

解: (1) 由分布律的归一性知

$$\sum_{i=1}^{+\infty} P\{X = i\} = \sum_{i=1}^{+\infty} \frac{a}{i(i+1)} = a \sum_{i=1}^{+\infty} \frac{1}{i(i+1)} = 1$$

令

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

例5 设离散型随机变量 X 的分布律为 $P\{X = i\} = \frac{a}{i(i+1)}, i = 1, 2, \dots$

易知
$$S_n = \sum_{i=1}^n \left(\frac{1}{i} - \frac{1}{i+1} \right) = \left(1 - \frac{1}{2} \right) + \left(\frac{1}{2} - \frac{1}{3} \right) + \dots + \left(\frac{1}{n} - \frac{1}{n+1} \right) = 1 - \frac{1}{n+1}$$

所以有
$$a \sum_{i=1}^{+\infty} \frac{1}{i(i+1)} = a \lim_{n \rightarrow \infty} S_n = a \lim_{n \rightarrow \infty} \left(1 - \frac{1}{n+1} \right) = a$$

即有
$$a = 1$$

(2)
$$P\{X < 5\} = \sum_{i=1}^4 P\{X = i\} = \sum_{i=1}^4 \frac{1}{i(i+1)} = 1 - \frac{1}{5} = \frac{4}{5}$$

1. 单点分布 (One-point Distribution)

若随机变量 X 只取一个常数值 c , 即 $P\{X=c\}=1$, 则称 X 服从单点分布, 也称为退化分布。

2. 0-1分布(0-1 Distribution)

若随机变量 X 只可能取0和1两个值, 其分布律为

X	0	1
P	q	p

其中 $0 < p < 1, q = 1 - p$, 或记为 $P\{X=k\} = p^k q^{1-k}, k = 0, 1$

则称 X 服从参数为 p 的0-1分布或两点分布。

0-1分布是除单点分布外，最简单的一个分布类型，它虽然简单但实际中很常见，如新生婴儿是男是女，明天是否下雨，某一股票明天是涨还是跌，种子是否发芽等。

任何一个只有两种可能结果的随机现象，都可以用一个服从两点分布的随机变量来描述。

两点分布又称为伯努利(Bernoulli)分布。

a. 伯努利试验

若随机试验 E 只有两个可能结果 A 和 \bar{A} ， $P(A)=p(0<p<1)$ ，则称 E 为伯努利试验。

b. n 重伯努利试验

设 E 为一伯努利试验，它的两个可能的结果为 A 和 \bar{A} ， $P(A)=p$. 把 E 在相同条件下，独立地重复进行 n 次，作为一个试验，称这个试验为 n 重伯努利试验，记为 E^n 。

注 意



(1) n 重伯努利试验 E^n 的可能结果可用长为 n 的 A 与 \bar{A} 的序列表示。

如3重伯努利试验的所有可能结果为：

$$AAA, \bar{A}AA, A\bar{A}A, AA\bar{A}, \bar{A}\bar{A}A, \bar{A}A\bar{A}, A\bar{A}\bar{A}, \bar{A}\bar{A}\bar{A}$$

注意



(2) “重复” 是指：在每次试验中， $P(A)=p$ 保持不变。

(3) “独立” 是指：各次试验的结果是相互独立的，即若以 C_i 记第 i 次试验的结果， C_i 为 A 或 \bar{A} , $i=1, \dots, n$. 则： C_1, C_2, \dots, C_n 相互独立。

$$\text{即： } P(C_1 C_2 \dots C_n) = P(C_1) P(C_2) \dots P(C_n)$$

c. 一个问题

设 E 为伯努利试验，且 $P(A)=p (0 < p < 1)$ ，对于 n 重伯努利试验 E^n ，事件 A 恰好发生 $k (k=0, 1, \dots, n)$ 次的概率是多少？

例5. 设某人打靶单发命中率为0.7，现独立重复射击4发，求恰好命中3发的概率。

解：以 A_i 表示“第 i 发命中” $i=1, 2, 3, 4$ ， B 表示“恰好命中3发” 则有

$$\begin{aligned} P(B) &= P(A_1 A_2 A_3 \bar{A}_4 \cup A_1 A_2 \bar{A}_3 A_4 \cup A_1 \bar{A}_2 A_3 A_4 \cup \bar{A}_1 A_2 A_3 A_4) \\ &= P(A_1 A_2 A_3 \bar{A}_4) + P(A_1 A_2 \bar{A}_3 A_4) + P(A_1 \bar{A}_2 A_3 A_4) + P(\bar{A}_1 A_2 A_3 A_4) \\ &= P(A_1)P(A_2)P(A_3)P(\bar{A}_4) + P(A_1)P(A_2)P(\bar{A}_3)P(A_4) + P(A_1)P(\bar{A}_2)P(A_3)P(A_4) \\ &\quad + P(\bar{A}_1)P(A_2)P(A_3)P(A_4) = 0.7^3 \times 0.3 + 0.7^3 \times 0.3 + 0.7^3 \times 0.3 + 0.7^3 \times 0.3 \\ &= C_4^3 \times 0.7^3 \times 0.3 = C_4^3 \times 0.7^3 \times 0.3^{4-3} = C_n^k \times p^k \times (1-p)^{n-k} \end{aligned}$$

定理： 设 E 为伯努利试验，且 $P(A)=p$ ($0<p<1$)，记 $q=1-p$ ，对于 n 重伯努利试验 E^n ，事件 A 恰好发生 k ($0\leq k\leq n$)次的概率为

$$P(\text{“}A\text{恰好发生}k\text{次”}) = C_n^k p^k q^{n-k}$$

3.二项分布 (Binomial Distribution)

设在一次伯努利试验中有两个可能的结果, A 与 $S-A$, 且有 $P(A)=p$ 。则在 n 重伯努利试验中事件 A 发生的次数 X 是一个离散型随机变量, 其分布律为

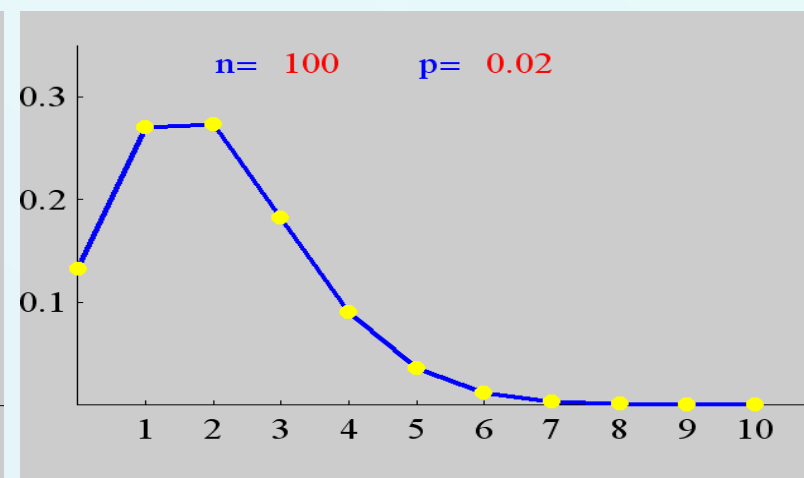
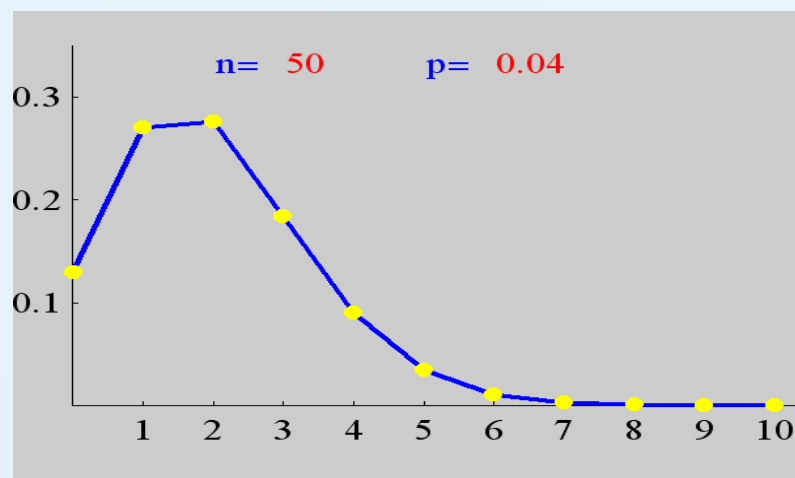
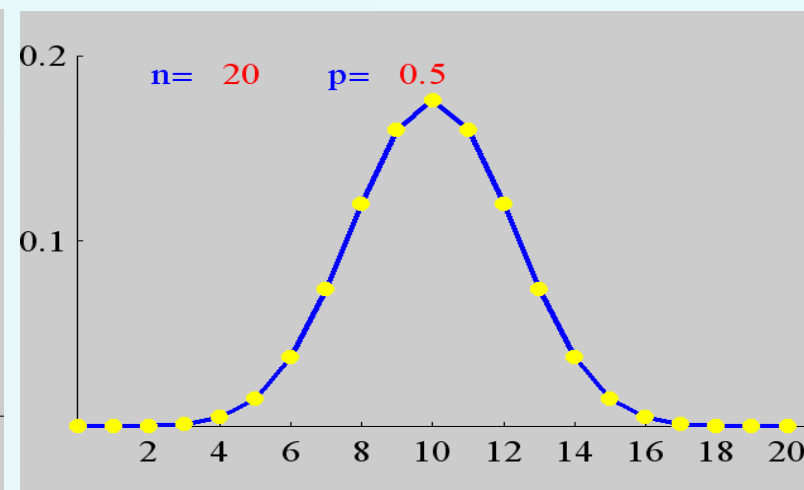
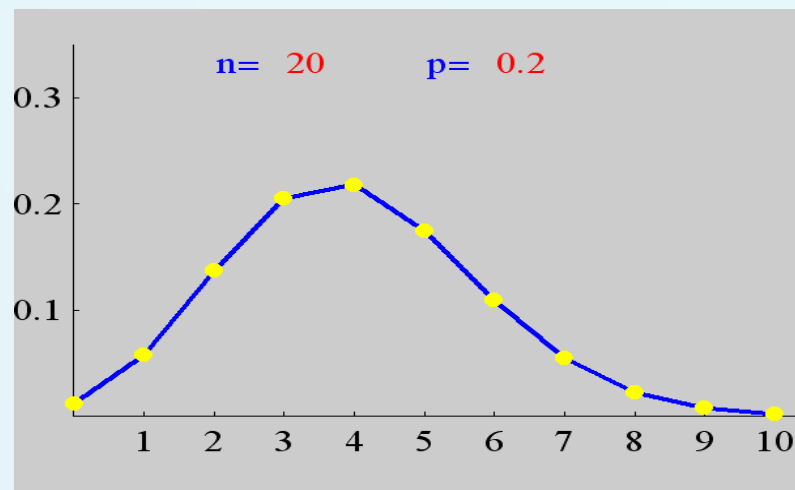
$$P\{X = k\} = C_n^k p^k q^{n-k} \quad k=0, 1, 2, \dots, n$$

称 X 服从参数为 n 和 p 的二项分布, 记为 $X \sim b(n, p)$ 。

由于 $C_n^k p^k q^{n-k}$ 是 $(p+q)^n$ 按二项式展开其中含因子 p^k 的那一项, 分布得名为二项分布。

三. 几种重要的离散型随机变量

二项分布 的图形



例6.为保证设备正常工作，需要配备适量维修工人，现有同类型设备80台，各台工作相互独立，且各台发生故障的概率都是0.01。

在通常情况下，一台设备的故障可由一人来处理。考虑两种配备维修工人的方法，其一是由4人维护，每人负责20台。

其二是由3人共同维护80台。试比较这两种方法在设备发生故障时不能及时维修的概率的大小。

解： 设 X 为第1人维护的20台中同一时刻发生故障的台数。

则有 $X \sim b(20, 0.01)$

设 A_i 表示 “第 i 人维护的20台机器发生故障不能及时维修” , $i=1, 2, 3, 4$ 。

易知, A_i 互相独立。则80台设备发生故障不能及时维修的概率为

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3 \cup A_4) &= 1 - P(\bar{A}_1)P(\bar{A}_2)P(\bar{A}_3)P(\bar{A}_4) = 1 - [P(\bar{A}_1)]^4 \\ &= 1 - [P\{X \leq 1\}]^4 = 1 - [P\{X = 0\} + P\{X = 1\}]^4 \\ &= 1 - [C_{20}^0 (0.01)^0 (0.99)^{20} + C_{20}^1 (0.01)^1 (0.99)^{19}]^4 = 0.0659 \end{aligned}$$

由3人共同维护80台

解： 设 Y 表示80台机器同一时刻发生故障的台数，则有 $Y \sim b(80, 0.01)$

80台机器发生故障不能及时维修的概率为

$$P\{Y \geq 4\} = 1 - P\{Y \leq 3\} = 1 - \sum_{i=0}^3 C_{80}^i (0.01)^i (0.99)^{80-i} = 0.0087$$

第二种方法发生故障而不能及时维修的概率小，且维修工人少一人。
运用概率论讨论国民经济问题，可以有效地使用人力、物力资源。

4. 几何分布

设离散型随机变量 X 可能的取值为 $1, 2, \dots, k, \dots$, 且取各个值的概率为

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

称 X 服从参数为 p 的几何分布, 记为 $X \sim G(p)$ 。

因为概率的和组成一个几何级数, 分布因此得名。

几何分布的概率背景, 在一个伯努利试验中, 事件 A 出现的概率为 p , 试验一个接着一个独立的进行, 用 X 表示事件 A 第一次发生时所进行的试验次数, 则 X 服从几何分布。

性质：几何分布的无记忆性。若 X 服从参数为 p 的几何分布， n, m 为任意两个正整数，则

$$P\{X > n+m | X > n\} = P\{X > m\}$$

可以理解为：若已经进行了 n 次试验，事件 A 没有发生，则又进行 m 次试验 A 依然没有发生的概率与已知的信息（前 n 次试验 A 没有发生）无关，这就是说，并不因为已经进行了 n 次试验 A 没有发生，而会使得在第 $n+1, n+2, \dots, n+m$ 次试验中 A 首次发生的概率提高。

例7. 同时掷两枚骰子，直到一枚骰子出现6点为止，试求抛掷次数 X 的分布律。

解：设 A 表示“第一颗骰子出现6点”， B 表示“第二颗骰子出现6点”， C 表示出现6点。

$$p = P(C) = P(A \cup B) = P(A) + P(B) - P(AB) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}$$

易知 X 服从几何分布，其分布律为

$$P\{X = n\} = (1 - p)^{n-1} p = \left(1 - \frac{11}{36}\right)^{n-1} \frac{11}{36} = \frac{11 \times 25^{n-1}}{36^n}, n = 1, 2, \dots$$

5. 超几何分布

设 n, M, N 为已知的常数, $n \leq M, M \leq N$, 若离散型随机变量 X 的分布律为

$$P\{X = m\} = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n} \quad m = 0, 1, \dots, l, \quad l = \min(M, n)$$

称 X 服从超几何分布, 记作 $X \sim H(n, M, N)$ 。

超几何分布对应于不放回的抽样模型。设一批同类型的产品共有 N 件, 其中次品有 M 件。今从中任取 n , 则这 n 件中所含的次品数 X 服从超几何分布。

应用:

1. 产品的抽样检验;
2. 估计湖中的鱼数问题 (Capture-Recapture问题) ;
3. 玩扑克牌游戏中, 检验牌洗的是否彻底 (皮尔逊做过检验) 。



作业： 3,5,8,9,10,16

第 5 讲

谢谢观看