第二章　随机变量及概率分布

一、随机变量的基本概念

1. 为什么引入随机变量?

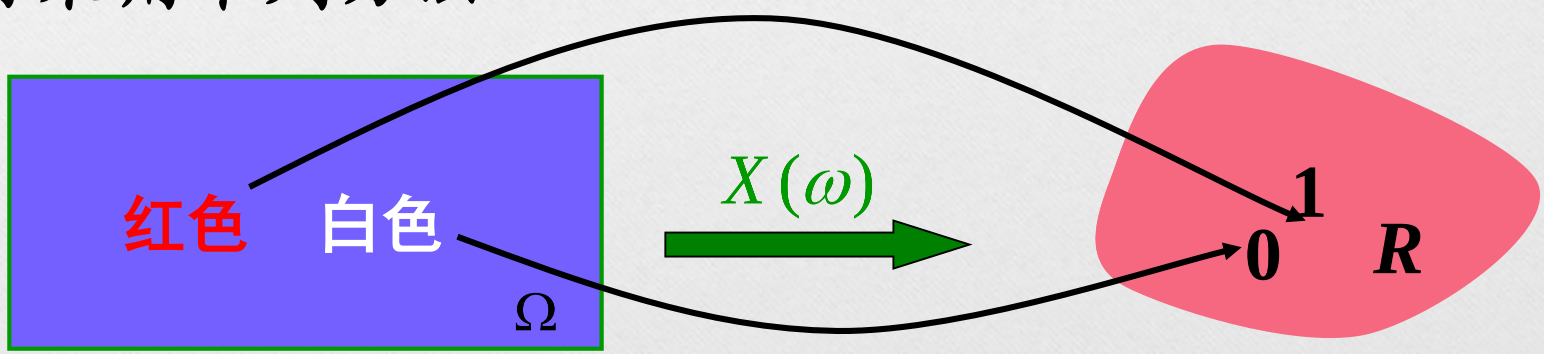
概率论是从数量上来研究随机现象内在规律性的，为了更方便有力的研究随机现象，就要用数学分析的方法来研究， 因此为了便于数学上的推导和计算，就需将任意的随机事件数量化．当把一些非数量表示的随机事件用数字来表示时，就建立起了随机变量的概念．

2. 随机变量的引入

实例1： 在一装有红球、白球的袋中任摸一个球，观察摸出球的颜色.

$$\Omega=\{\underbrace{\text{红色、白色}}_{\text{非数量}}\}\xrightarrow{?}\text{将 }\Omega\text{ 数量化}$$

可采用下列方法



即有　　X(红色)=1，　　X(白色)=0.

$$X(\omega)=\begin{cases}1, & \omega=\text{红色},\\ 0, & \omega=\text{白色}.\end{cases}$$

这样便将非数量的 Ω={红色，白色} 数量化了.

$$\mathrm{P}\{\mathrm{X}=0\}=\frac{1}{2},\ \mathrm{P}\{\mathrm{X}=1\}=\frac{1}{2}$$

实例2 抛掷骰子,观察出现的点数.

Ω={1，2，3，4，5，6}

样本点本身就是数量

$X(\omega)=\omega$ 恒等变换

$$X(1)=1,\ X(2)=2,\ X(3)=3,\ X(4)=4,\ X(5)=5,\ X(6)=6,$$

且有 $P\{X=i\}=\frac{1}{6},\ (i=1,2,3,4,5,6).$

3、随机变量的定义

设 S是随机试验，它的样本空间是Ω. 如果对于每一个样本点$\omega \in \Omega$, 有一个实数$X(\omega)$与之对应，这样就得到一个定义在Ω上的单值实值函数$X(\omega)$, 称$X(\omega)$为随机变量.

随机变量通常用$X, Y, Z, \ldots$来表示.

2.说明

(1) 随机变量与普通的函数不同

随机变量是一个函数，但它与普通的函数有着本质的差别，普通函数是定义在实数轴上的，而随机变量是定义在样本空间上的(样本空间的元素不一定是实数).

(2) 随机变量的取值具有一定的概率规律

随机变量随着试验的结果不同而取不同的值，由于试验的各个结果的出现具有一定的概率，因此随机变量的取值也有一定的概率规律.

实例1 设盒中有5个球 (2白3黑), 从中任抽3个,则

$$X(\omega)=\text{抽得的白球数},$$

是一个随机变量. $X(w)$ 的所有可能取值为:

$$0, \quad 1, \quad 2.$$

实例2 设某射手每次射击打中目标的概率是0.8, 现该射手不断向目标射击 , 直到击中目标为止,则

$$X(\omega)=\text{所需射击次数},$$

是一个随机变量. $X(w)$ 的所有可能取值为:

$$1, \quad 2, \quad 3, \quad \cdots \quad .$$

实例3 某公共汽车站每隔 5 分钟有一辆汽车通过, 如果某人到达该车站的时刻是随机的, 则

$$X(\omega) = \text{此人的等车时间},$$

是一个随机变量. $X(w)$ 的所有可能取值为: **[0,5].**

4.随机变量的分类

随机变量

离散型　　连续型

(1) 离散型　随机变量所取的可能值是有限多个或无限可列个，叫做离散型随机变量.

例1：观察掷一个骰子出现的点数.

例2：随机变量 X 为 “连续射击, 直至命中时的射击次数”.

(2) 连续型 随机变量所有可能取值连续地充满某个区间(并有密度函数)，叫连续型随机变量.

例1 随机变量 X 为“灯泡的寿命”.

X 的取值范围为 $[0,+\infty)$.

例2 某公共汽车站每隔 5 分钟有一辆汽车通过,如果某人到达该车站的时刻是随机的, 则若随机变量 X 为“此人的等车时间”. X 的取值范围为 $[0, 5]$.

二、离散型随机变量及其分布律

1.离散型随机变量的分布律

2.常见离散型随机变量的概率分布

1.离散型随机变量的分布律

定义 设离散型随机变量 X 所有可能取的值为 $x_k\ (k=1,2,\cdots)$，X 取各个可能值的概率，即事件 $\{X=x_k\}$ 的概率，为

$$P\{X=x_k\}=p_k,\quad k=1,2,\cdots.$$

则称$\{p_k\}$为离散型随机变量X的分布律。

说明 (1) $p_k \geq 0,\ k=1,2,\cdots$

(2) $\sum_{k=1}^{\infty} p_k = 1$

离散型随机变量的分布律也可表示为

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_n & \cdots \\ p_1 & p_2 & \cdots & p_n & \cdots \end{pmatrix}$$

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

例2.1 设离散随机变量X的分布律为

X	0	1	2
p	0.2	a	0.5

求常数 a

解 $a=0.3.$

例2.2 袋子中有5个同样大小的球，编号为1~5, 从中同时取出3个球，记 X 为取出的球的最大编号，求 X 的分布率.

解 X 的可能取值为 3, 4, 5

$$P\{X=3\}=\frac{1}{C_5^3}=\frac{1}{10},\quad P\{X=4\}=\frac{C_3^2}{C_5^3}=\frac{3}{10},\quad P\{X=5\}=\frac{C_4^2}{C_5^3}=\frac{6}{10}$$

则 X 的分布律为

X	3	4	5
P	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{6}{10}$

例2.3 已知一批零件共10个，其中3个不合格，现任取一件使用，若取到不合格零件，则丢弃，再重新抽取一件，如此下去，试求取到合格零件之前取出的不合格零件个数X的分布率.

解 X的可能取值为0,1,2,3.

设$A_i\ (i=0,1,2,3)$表示“第i次取出的零件不合格”.

利用概率的乘法公式可得:

$$P\{X=0\}=P(\bar{A}_1)=\frac{7}{10},$$

$$P\{X=1\}=P(A_1\bar{A}_2)=P(A_1)P(\bar{A}_2|A_1)=\frac{3}{10}\cdot\frac{7}{9}=\frac{7}{30}$$

$$P\{X=2\}=P(A_1A_2\bar{A}_3)=P(A_1)P(A_2|A_1)P(\bar{A}_3|A_1A_2)=\frac{3}{10}\cdot\frac{2}{9}\cdot\frac{7}{8}=\frac{7}{120}$$

$$P\{X=3\}=P(A_1A_2A_3\bar{A}_4)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)P(\bar{A}_4|A_1\,A_2A_3)$$

$$=\frac{3}{10}\times\frac{2}{9}\times\frac{1}{8}\times\frac{7}{7}=\frac{1}{120}$$

故 X 的分布律为

X	0	1	2	3
P	$\frac{7}{10}$	$\frac{7}{30}$	$\frac{7}{120}$	$\frac{1}{120}$

例2.4 已知 X 的分布律为

X	0	1	2	3
P	$\frac{7}{10}$	$\frac{7}{30}$	$\frac{7}{120}$	$\frac{1}{120}$

求$P\{X\le 1\}$、 $P\{X>1\}$、$P\{1 \le X < 2.5\}$.

解:

$P\{X\le 1\}=P\{X=0\}+P\{X=1\}=7/10+7/30=14/15$

$P\{X>1\}=P\{X=2\}+P\{X=3\}=7/120+1/120=1/15$

$P\{1 \le X < 2.5\}= P\{X=1\}+P\{X=2\}=7/30+7/120=7/24$

例2.5 对一目标连续进行射击，直到击中目标为止。如果每次射击的命中率为p，求射击次数X的分布律。

解　X的可能取值为1, 2, 3,….

设 A_i $(i = 1, 2, 3, 4)$表示“第i次未射中”, 事件$\{X=k\}$表示“前k-1次射击未中，第k次命中”

$$P\{X = k\} = A_1 A_2 \ldots A_{k-1} \overline{A_k}$$

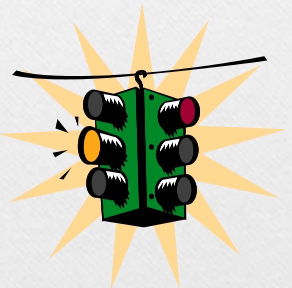
又每次射击命中与否是相互独立的, 则X的分布律为:

$$p_k = P\{X = k\} = P(A_1 A_2 \ldots A_{k-1} \overline{A_k})$$

$$= P(A_1)P(A_2)\ldots P(A_{k-1})P(\overline{A_k}) = (1-p)^{k-1} p, k = 1, 2, \ldots$$

练习

设一汽车在开往目的地的道路上需经过四组信号灯，每组信号灯以 1/2 的概率允许或禁止汽车通过. 以 X表示汽车首次停下时，它已通过的信号灯的组数(设各组信号灯的工作是相互独立的)，求X的分布律.



2.常见离散型随机变量的概率分布

1. 两点分布

设随机变量 X 只可能取0与1两个值，它的分布律为

X	0	1
p_k	$1-p$	p

则称 X 服从 **(0—1) 分布**或**两点分布**.

例1 “抛硬币”试验，观察正、反两面情况.

$$X=X(\omega)=\begin{cases}0, & \text{当}\omega=\text{正面},\\ 1 & \text{当}\omega=\text{反面}.\end{cases}$$

随机变量 X 服从 (0-1) 分布.

其分布律为

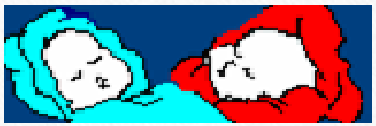
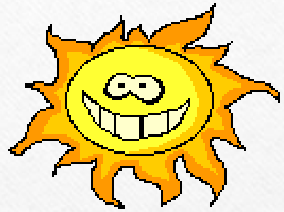
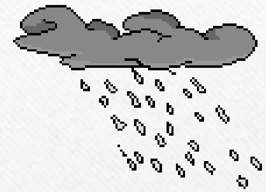
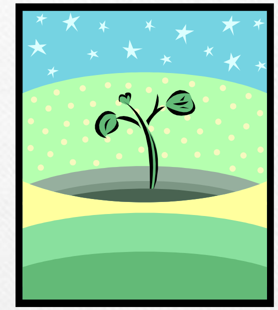
X	0	1
p_k	$\frac{1}{2}$	$\frac{1}{2}$

例2　200件产品中,有190件合格品,10件不合格品,现从中随机抽取一件, 若规定

$$X=\begin{cases}1, & \text{取得不合格品},\\ 0, & \text{取得合格品}.\end{cases}$$

X	0	1
p_k	$\frac{190}{200}$	$\frac{10}{200}$

则随机变量 X 服从(0-1)分布.

说明

两点分布是最简单的一种分布,任何一个只有两种可能结果的随机现象,比如新生婴儿是男还是女、明天是否下雨、种籽是否发芽等,都属于两点分布.

2. 二项分布

(1) 重复独立试验

将试验 S 重复进行 n 次, 若各次试验的结果互不影响 , 即每次试验结果出现的概率都不依赖于其它各次试验的结果, 则称这 n 次试验是相互独立的, 或称为 n 次重复独立试验.

(2) n 重伯努利试验

设试验S只有两个可能结果: A 及 $\overline{A}$,则称 E为伯努利试验.设 $P(A)=p\ (0<p<1)$,此时$P(\overline{A})=1-p$.

将S独立地重复地进行n次,则称这一串重复的独立试验为 n**重伯努利试验** .

例1 抛一枚硬币观察得到正面或反面. 若将硬币抛 n 次,就是n重伯努利试验.

例2 抛一颗骰子n次, 观察是否 “出现 1 点”, 就是n重伯努利试验.

(3) 二项分布

若 X 表示 n 重伯努利试验中事件 A 发生的次数，则 X 所有可能取的值为 **0, 1, 2, $\cdots$, n.**

当 $X=k\ (0\le k\le n)$ 时，即 A 在 n 次试验中发生了 k 次.

因此 A 在 n 次试验中发生 k 次的概率为

$$C_n^k p^k (1-p)^{n-k} \xrightarrow{\text{记 } q=1-p} C_n^k p^k q^{n-k}$$

得 X 的分布律为

X	0	1	$\cdots$	k	$\cdots$	n
p_k	q^n	$C_n^1 pq^{n-1}$	$\cdots$	$C_n^k p^k q^{n-k}$	$\cdots$	p^n

称这样的分布为二项分布. 称X服从参数为n, p的二项分布，记为 $X \sim B(n, p)$.

二项分布 $\xrightarrow{n=1}$ 两点分布

例2.5 在相同条件下相互独立地进行 5 次射击,每次射击时击中目标的概率为 0.6 ,则击中目标的次数 X 服从 $B(5, 0.6)$ 的二项分布.

例2.6 某特效药的临床有效率为0.95，现有10人服用，问至少有8人治愈的概率是多少？

解：设X表示10人中被治愈的人数，则 $X \sim B(10, 0.95)$

而所求的概率为 $P\{X \geq 8\} = 0.9885.$

例2.7 $X \sim B(2, p), Y \sim B(3, p)$, 设$P\{X \geq 1\} = \dfrac{5}{9}$

求$P\{Y \geq 1\}$

解 由$P\{X \geq 1\} = 1 - P\{X = 0\}$

$$= 1 - C_2^0 p^0 (1-p)^2 = \frac{5}{9}$$

得$p = \dfrac{1}{3}$

又由$P\{Y \geq 1\} = 1 - P\{Y = 0\}$

$$= 1 - C_3^0 p^0 (1-p)^3 = \frac{19}{27}$$

例2.8 某人进行射击,设每次射击的命中率为0.02,独立射击400 次,试求至少击中两次的概率.

解 设击中的次数为 X,则

$$X \sim b(400, 0.02).$$

X 的分布律为

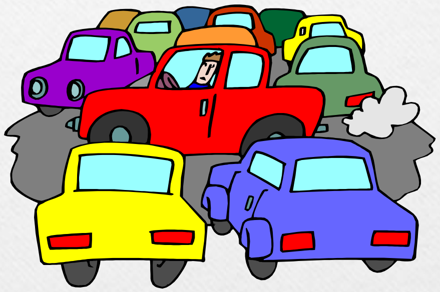
$$P\{X=k\}=C_{400}^{k}(0.02)^{k}(0.98)^{400-k},$$

$$k=0,1,\cdots,400.$$

因此 $P\{X \geq 2\}=1-P\{X=0\}-P\{X=1\}$

$$=1-(0.98)^{400}-400(0.02)(0.98)^{399}=0.9972.$$

例2.9 有一繁忙的汽车站,每天有大量汽车通过,设每辆汽车在一天的某段时间内,出事故的概率为0.0001,在每天的该段时间内有1000 辆汽车通过, 问出事故的次数不小于2的概率是多少?



解　设 1000 辆车通过,出事故的次数为 X, 则 $X \sim B(1000, 0.0001)$,

故所求概率为

$$P\{X \geq 2\} = 1 - P\{X = 0\} - P\{X = 1\}$$

$$= 1 - 0.9999^{1000} - C_{1000}^{1} \cdot 0.0001 \cdot 0.9999^{999} \ ??$$

二项分布 $\xrightarrow{np \to \lambda (n \to +\infty)}$ 泊松分布

二项分布的泊松逼近

泊松定理 设λ>0是常数，**n**是任意正整数，且λ= **np**，则对于任意取定的非负整数**k**，有

$$\lim_{n\to\infty} C_n^k p^k (1-p)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}$$

由**泊松定理** 若 **n** 很大 **p** 很小时，且 λ= **np**，则有

$$P\{X = k\} = C_n^k p^k (1-p)^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda}$$

$$\boxed{n \geq 20, p \leq 0.05}$$

接上例:

$$P\{X \geq 2\} = 1 - P\{X = 0\} - P\{X = 1\}$$

$$= 1 - 0.9999^{1000} - \binom{1000}{1} \cdot 0.0001 \cdot 0.9999^{999}$$

可利用泊松定理 $\lambda = 1000 \times 0.0001 = 0.1,$

$$P\{X \geq 2\} \approx 1 - \frac{e^{-0.1}}{0!} - \frac{0.1 \cdot e^{-0.1}}{1!} = 0.0047$$

$e \approx 2.71828\ldots$

3. 泊松分布

设随机变量X所有可能取的值为$0, 1, 2, \cdots,$而取各个值的概率为

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

其中$\lambda > 0$是常数.则称X服从参数为λ的泊松分布,记为$X \sim P(\lambda)$.

例2.11 一个工厂生产的产品中废品率为0.005，任取1000件，计算（1）其中至少有两件是废品的概率；（2）其中不超过5件废品的概率。

解 设X为任取的1000件产品中的废品数，则

$$X \sim b(1000, 0.005)$$

$$\lambda = 1000 \times 0.005 = 5$$

(1) $P\{X \geq 2\} = 1 - P\{X = 0\} - P\{X = 1\}$

$$\approx 1 - \mathrm{e}^{-5} - 5\mathrm{e}^{-5} = 0.9596$$

(2) $P\{X \leq 5\} = \sum_{k=0}^{5} P\{X = k\} = 1 - \sum_{k=6}^{\infty} P\{X = k\}$

$$\simeq 1 - \sum_{k=6}^{+\infty} \frac{5^k}{k!} e^{-5} = 0.616$$

例2.12 设随机变量X服从参数为5的泊松分布，求 $P\{X \le 10\}$.

解 $P\{X \le 10\} = 1 - P\{X \ge 11\}$

$$= 1 - \sum_{k=11}^{\infty} \frac{5^k}{k!} e^{-5}$$

$$= 0.986305$$

例2.13 设随机变量X服从泊松分布，且已知 $P\{X=1\}=P\{X=2\}$ 求 $P\{X=4\}$

解 设X服从参数为λ的泊松分布，则

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

由已知得 $\frac{\lambda^1}{1!}e^{-\lambda}=\frac{\lambda^2}{2!}e^{-\lambda}$ 解得 $\lambda=2$

则 $P\{X=4\}=\frac{2^4}{4!}e^{-2}=\frac{2}{3}e^{-2}$

4.几何分布

若随机变量X的概率分布为：

$$P\{X=k\}=q^{k-1}p,\quad k=1, 2, 3,\ldots$$

则称X服从参数为p的几何分布,其中q=1－p.

记为：$X \sim G(p)$

二项分布与 (0-1) 分布、泊松分布之间的关系.

二项分布是 (0-1) 分布的推广, 对于 n 次独立重复伯努里试验, 每次试验成功的概率为 p, 设

$$X_i = \begin{cases} 1, & \text{若第 } i \text{ 次试验成功} \\ 0, & \text{若第 } i \text{ 次试验失败} \end{cases}, \quad (i = 1, 2, \cdots, n)$$

它们都服从 (0-1) 分布并且相互独立, 那么

$X = X_1 + X_2 + \cdots + X_n$ 服从二项分布 $B(n, p)$.

二项分布的泊松逼近

二项分布 $\xrightarrow{np \to \lambda (n \to +\infty)}$ 泊松分布

$$P\{X = k\} = C_n^k p^k (1-p)^{n-k} \approx \frac{(np)^k}{k!} e^{-np} = \frac{\lambda^k}{k!} e^{-\lambda},$$

$$(k = 0, 1, 2, \cdots, n).$$