§2.2　离散型随机变量

一、离散型随机变量的分布律

定义：如果随机变量X至多取可列无穷个数值：$x_1, x_2, \dots,$ 记 $p_i = P\{X = x_i\}$, 且满足

$$\begin{cases} (1) \quad p_i \geq 0, \ \forall i; \\ (2) \quad \sum_{i=1}^{\infty} p_i = 1. \end{cases}$$

称X是离散型随机变量.

称 $$p_i = P\{X = x_i\},\quad i = 1,2,\ldots$$
为X 的分布律．表示为

X	x_1	x_2	$\cdots$	x_i	$\cdots$
$P\{X = x_i\}$	p_1	p_2	$\cdots$	p_i	$\cdots$

上节例1中彩金Y是离散型随机变量，其分布律为：

Y	0	0.05	0.2	2
$P\{Y = y_i\}$	0.5001	0.3589	0.1282	0.0128

其它例子 → 产品检验试验

对于离散型随机变量X，因

$$\{X \le x\} = \bigcup_{x_i \le x} \{X = x_i\}$$

故

$$F(x) = P\{X \le x\} = P[\bigcup_{x_i \le x} \{X = x_i\}]$$

由概率可加性

$$= \sum_{x_i \le x} P\{X = x_i\}$$

二、贝努里试验和二项分布

E1：抛一枚硬币出现正反面

E2：检查一件产品是否合格

E3：射击，观察是否命中

E4：考一门课，是否通过

贝努里试验

特点：关注试验的两个结果：A 和 $\overline{A}$

实际结果可能不止两个

贝努里试验仅有两个基本事件：A 和 $\overline{A}$，

记 $P(A)=p$

令随机变量

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

则X的分布律为

X	0	1
$P\{X=x_i\}$	$1-p$	p

称X服从(0—1)分布

思考: 怎样求X的分布函数?

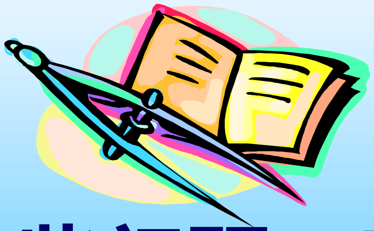
定义： 将试验E 按下述条件重复进行n次

(1) 每次试验的条件不变；

(2) 各次试验的结果互不影响.

称这n次试验为n次重复独立试验.

当试验E 是贝努里试验，称这n次独立试验为n重贝努里试验，或称贝努里概型.



对于n重贝努里试验，可考察哪些问题，考虑哪些变量 ？

(1) n次试验中事件A 发生的总次数X；

(2) 事件A 首次发生时的试验次数Y;

(3) 事件A 发生k 次时的试验次数Z；……

练习 尝试写出随机变量 Y、Z的分布律.

问题（1）

n重贝努里试验中，事件A 发生的总次数X 可能取数值： 0， 1， 2， …， n.

定理： 在n重贝努里试验中,事件A发生概率为$P(A) = p$，$0 < p < 1$，则事件A发生的次数X的分布律为

$$P\{X = k\} = C_n^k p^k (1-p)^{n-k},\quad k = 0,1,2,\cdots,n.$$

证 在n重贝努里试验中,事件A在指定的k次试验中出现的概率为

$$p^k (1-p)^{n-k}$$

从n次试验中选出k次试验有C_n^k种不同的方式.

且各种方式的事件互不相容，则：

$$P(X=k)=C_n^k p^k(1-p)^{n-k} \quad k=0,1,2,\cdots,n.$$

结论成立.

称随机变量 X 服从二项分布，记为 $X\sim B(n,p)$.

(0—1)分布可以看作 $X\sim B(1,p)$.

例子 ⇨ 产品抽检试验

强弱对抗试验

设备排障试验

三、泊松分布

定义：若随机变量X的分布律为

$$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda},\quad k=0,1,2,\ldots;\quad (\lambda>0)$$

称X服从参数为λ的泊松分布. 记为 $X\sim P(\lambda)$.

泊松分布的重要性在于:

(1) 现实中大量随机变量服从泊松分布;

(2) 泊松分布可视为二项分布的极限分布.

定理 设随机变量序列 $X_n \sim B(n, p_n)$, $n = 1, 2, \ldots,$ 即有

$$P\{X = k\} = C_n^k p_n^k (1 - p_n)^{n-k}, \qquad k = 0, 1, 2 \cdots n$$

若 $\lim\limits_{n\to\infty} np_n = \lambda \quad (\lambda > 0)$, 则有

$$\lim_{n\to\infty} P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \cdots$$

思考：你能从条件 $\lim\limits_{n\to\infty} np_n = \lambda > 0$, 中分析出什么结论吗?

注: $\lim\limits_{n\to\infty} np_n = \lambda \iff \lim\limits_{n\to\infty} \dfrac{p_n}{\frac{1}{n}} = \lambda$

即数列$\{p_n\}$与$\{\frac{1}{n}\}$是同阶的无穷小.故

(1) 当n够大, p较小时有

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

其中$\lambda = np$. → 设备排障试验

(2) 实际问题中, n次独立重复试验中,"稀有事件"出现的次数可认为服从泊松分布. #

例1 某种产品在生产过程中的废品率为p（$0<p<1$），对产品逐个检查，直到检查出5个不合格品为止，试写出停止检查时已检查的产品个数X的分布律.

解：关键是分析随机事件 $\{X=k\}$，

事件$\{X=k\}$相当于第k次检查到的产品必为不合格品，而前$k-1$次检查中查出4件不合格品.

如指定前四次：

$$\underbrace{1\quad 2\quad 3\quad 4}_{\text{不合格}}\quad \underbrace{\cdots\quad k-1}_{\text{合格}}\quad \underset{\text{不合格}}{\underset{\uparrow}{k}}$$

进行 k 次检查，指定的5次检查出现不合格品的概率为 $p^5(1-p)^{k-5}$.

从前$k-1$次检查中选出4 次出现不合格产品共有 C_{k-1}^{4} 种不同的方式.

故分布律为

$$P\{X=k\}=C_{k-1}^{4}\,p^5(1-p)^{k-5},$$

$$(k=5, 6, \ldots)$$

#

例2 设有一批同类产品共有N个，其中次品有M个，现从中逐个有放回地取出n个，试求取出n件中所含的次品件数X的分布律.

解 产品是逐件有放回取出，各次抽到次品是相互独立的，抽n件产品相当于做n重贝努里试验.

故 $X \sim \mathrm{B}(n\,,\,\frac{M}{N}).$

所以, X 的分布律为

$$P\{X = k\} = C_n^k \left(\frac{M}{N}\right)^k \left(1 - \frac{M}{N}\right)^{n-k}$$

其中 $k = 0,1,2,\ldots,n.$

思考：将抽取方式改为无放回抽取,试写出X 的分布律.

例3 强弱两队进行乒乓球对抗赛，得胜人数多的一方获胜，已知强队每个队员获胜的概率为0.6，下面两个方案中哪一个对弱队有利?

（1）双方各出3人；（2）双方各出7人.

解：设$A = \{$弱队获胜$\}$，弱队获胜的人数为X.

双方逐对较量独立进行，故为独立重复试验.

（1）当双方各出3人时，$X \sim B(3, 0.4)$

$$P(A) = P\{X \geq 2\} = \sum_{k=2}^{3} C_3^k (0.4)^k (0.6)^{3-k}$$

$$\approx 0.352$$

(2) 当双方各出7人时， $X \sim B(7, 0.4)$.

$$P(A) = P\{X \geq 4\} = \sum_{k=4}^{7} C_7^k (0.4)^k (0.6)^{7-k}$$

$$\approx 0.290$$

故第一种方案对弱队更有利一些.

#

例4 有300台独立运转的同类机床，每台机床发生故障的概率都是0.01，若一人排除一台的故障. 问至少需要多少名工人，才能保证不能及时排除故障的概率小于0.01.

解：设X 表示同一时刻发生故障的机床数,则

$$X \sim B(300, 0.01).$$

若 配N 个工人，应使

$$0.01 > P\{X > N\} = 1 - P\{X \le N\}$$

$$= 1 - \sum_{k=0}^{N} C_{300}^{k} (0.01)^{k} (1-0.01)^{300-k}$$

即求使上述不等式成立的最小N 值.

#

续例4

因为$300 \times 0.01 = 3$ (n较大，p较小),故可认为X近似服从 $\lambda = 3$ 的泊松分布，即 $X \sim P(3)$.

于是 $$0.01 > P\{X > N\} = \sum_{k=N+1}^{\infty} \frac{3^k}{k!} e^{-3}$$

查P251 的附表1 可得

$$P\{ X > 7 \} = 0.11905 > 0.01$$

$$P\{ X > 8 \} = 0.003803 < 0.01$$

所以，至少需要配备8 个修理工人.

#