一类离散随机变量分布

教学设计与制作：李曼曼

引例

1. 离散型随机变量

- **定义** 如果随机变量X取有限多个或可列个不同的值，则称X为离散型随机变量。

- **分布律或概率分布**：设离散型随机变量X的所有可能取值为 $a_1, a_2, a_3, \cdots$，则其分布律为

$$p_i = P\{X = a_i\}, i = 1, 2, 3, \cdots$$

非负性：$p_i \geq 0, i = 1, 2, 3, \cdots$ 规范性：$\sum_{i=1}^{+\infty} p_i = 1$

2. Bernoulli试验

Bernoulli试验：若某种试验只有两个结果(成功、失败；正面、反面)；

n重Bernoulli试验：n次独立重复的Bernoulli试验.

3.二项分布 $X \sim B(n,p)$

定义　在n重Bernoulli试验中，记成功的次数为X，其可能取值为：0,1,……,n. 则X 取值为k 的概率为：

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

称随机变量 X 服从参数为(n,p)的二项分布.

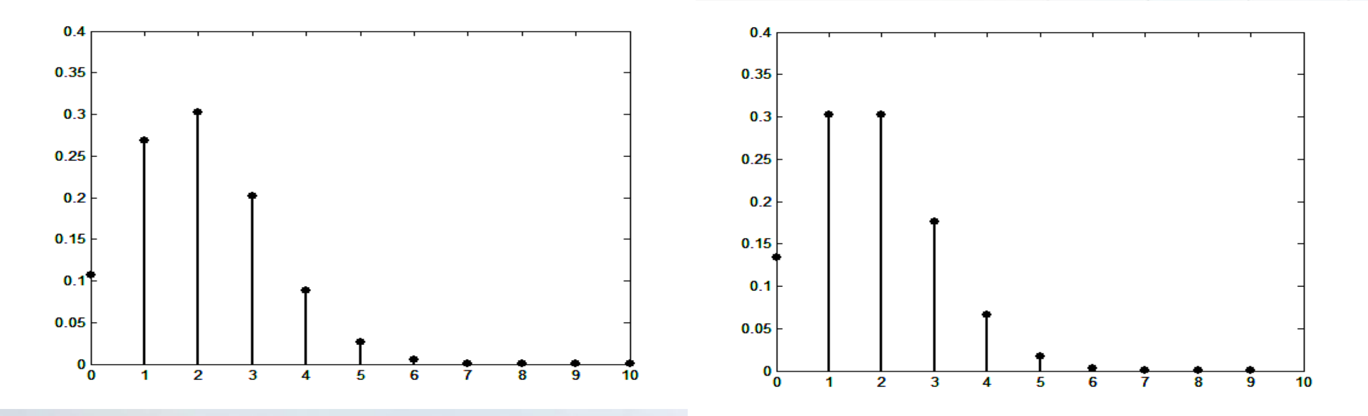
$$\sum_{k=0}^{n} P\{X=k\}=\sum_{k=0}^{n} C_n^p p^k q^{n-k}=(p+q)^n=1$$

特别的，$n=1$时，称其为二点分布或0-1分布，

记为 $X \sim B(1, p)$.

推论：二项随机变量是独立 0-1 随机变量之和.

$P\{X=i\}$



- $X \sim B(10,0.2)$

$X \sim B(9,0.2)$

图1 二项分布模拟

例1. 保险公司为了估计企业的利润，需要计算各种各样的概率，下面是典型问题之一. 若一年中某类保险者中个体死亡的概率等于0.003，现有10000个这类人参加人寿保险，试求这些保险者里面在未来一年中，

(1)有40个人死亡的概率；

(2)死亡人数不超过70个的概率.

解.利用n重Bernoulli概型，设 X 为未来一年中这些人里面死亡的人数，则

$$X \sim B(10000, 0.003)$$

$$P\{X = 40\} = C_{10000}^{40}(0.003)^{40}(0.997)^{9960}$$

$$P\{X \le 70\} = \sum_{k=0}^{70} C_{10000}^{k}(0.003)^{k}(0.997)^{10000-k}$$

4. 几何分布 $X \sim Ge(p)$

定义 X 为独立重复的Bernoulli试验中，“首次成功”时的试验次数,记为 $X \sim Ge(p)$.

$$P\{X = k\} = pq^{k-1}, k = 1, 2, \ldots$$

性质 几何分布具有**无记忆性**，即

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

几何分布应用

1.某产品的不合格率为0.05，那么首次检查到不合格品的检测次数 X 服从几何分布 $Ge(0.05)$；

2.某射击运动员命中9环以上的概率为0.8，则首次击中9环以上区域的射击次数 X 服从几何分布 $Ge(0.8)$.

5.负二项分布

满足以下条件的称为负二项分布

1. 实验包含一系列独立的试验。
2. 每次试验都有成功、失败两种结果。
3. 成功的概率p是恒定的。
4. 实验持续到r次成功，r可以为任意正数。

当r是整数时，负二项分布又称帕斯卡分布（巴斯卡分布）.

5.负二项分布

定义　X 为独立重复的Bernoulli试验中，“第 r 次成功”时的试验次数.记为 $X \sim NB(r, p)$.

$$P\{X = k\} = C_{k-1}^{r-1} p^r q^{k-r}, k = r, r+1, \ldots$$

特别的，$r = 1$时，等价于$X \sim Ge(p)$.

推论　负二项随机变量是独立几何随机变量之和.

例2.（负二项抽样） 为研究果蝇群中翅退化果蝇的比例，现对该种群进行抽样，直到取得100只翅退化的果蝇，则至少抽取了N只果蝇的概率为

$$P\{X \geq N\} = \sum_{k=N}^{+\infty} C_{k-1}^{99} p^{100} (1-p)^{k-100}$$

$$= 1 - \sum_{k=100}^{N-1} C_{k-1}^{99} p^{100} (1-p)^{k-100}$$

小结

Bernoulli试验

- n次独立重复，成功次数X
  - 二点分布 $n=1$
  - 二项分布 $n>1$
- X次独立重复，第r次成功恰好发生在X次试验
  - 几何分布 $r=1$
  - 负二项分布 $r>1$