概率论与数理统计

第九讲

几种常用的离散型随机变量

1. 0-1分布

又称两点分布或伯努利（Bernoulli）分布. 设随机变量X的分布律为

X	0	1
p_k	$1-p$	p

则称X服从参数为p（$0<p<1$）的0-1分布.

其分布律又可写成 $P\{X=k\}=p^k(1-p)^{1-k}$, $k=0, 1$.

常用它来表示两个状态的问题（即随机试验的结果只有两个，称为伯努利试验）：

$$S=\{s_1,s_2\},\qquad X(s_1)=1,\qquad X(s_2)=0.$$

第9讲 几种常用的离散型随机变量

0-1分布虽然简单, 却能很好地刻画实际问题中常见的对立现象：

- 试验的成功与失败
- 产品的合格与不合格
- 疾病的治愈与未愈
- 股市的“牛市”与“熊市”
- 生男孩与生女孩
- ……

例 (收入分布, 续)我国2012年家庭人均收入R(千元)分布如下：

R (千元)	1	2	4.5	9	15.9	25.8	34.3
收入低于R的家庭比例	0.05	0.1	0.25	0.5	0.75	0.9	0.95

若规定家庭人均收入2千元为贫困线. 又令

$$X=\begin{cases}1, & R\le 2\\ 0, & R>2\end{cases}$$

则X服从参数为0.1的两点分布, X可以看成是某家庭是否是贫困家庭的示性函数.

2. 二项分布

将上述伯努利试验独立地做n次，称为n重伯努利试验，设状态s_1, s_2在每次试验中出现的概率是不变的。

考虑X为n次试验中状态s_1出现的次数，则X的取值为$0, 1, \ldots, n$，其分布律如何？

例 (射击评估, 续) 设中靶率为p，射击5发子弹，发射击之间均独立，考虑X=3的一种情形

{击中，没中，击中，没中，击中} $\Longrightarrow$ $p^3(1-p)^2$

但中3枪的情形共有 $\binom{5}{3}$, 因此 $P\{X=3\}=\binom{5}{3}p^3(1-p)^2.$

一般地, 考虑n重伯努利试验.

两个要素：

① 成功率 p 为常数; ② 多次伯努利试验之间独立.

研究事件“恰有某特定的 k 次试验成功”的概率.

$$\underbrace{\circ\ \bullet\ \bullet\ \circ\ \circ\ \bullet\ \circ\ \circ\ \cdots\ \circ\ \bullet\ \circ\ \bullet}_{n\text{次试验}}$$

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

而“成功次数 $=k$”的不同情况共有$\binom{n}{k}$种,

它们是互斥的.

于是

$$P\{X=k\}=\binom{n}{k}p^k(1-p)^{n-k},\ \ k=0,1,\cdots,n$$

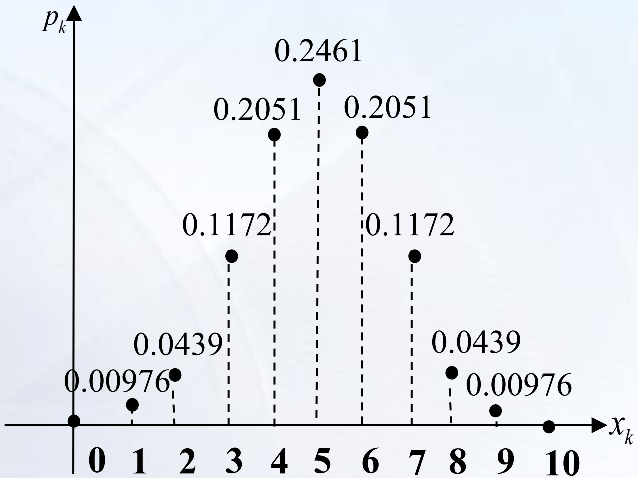
称具有上述分布律的随机变量为服从参数为n,p的二项分布, 记为：$X\sim B(n,p)$

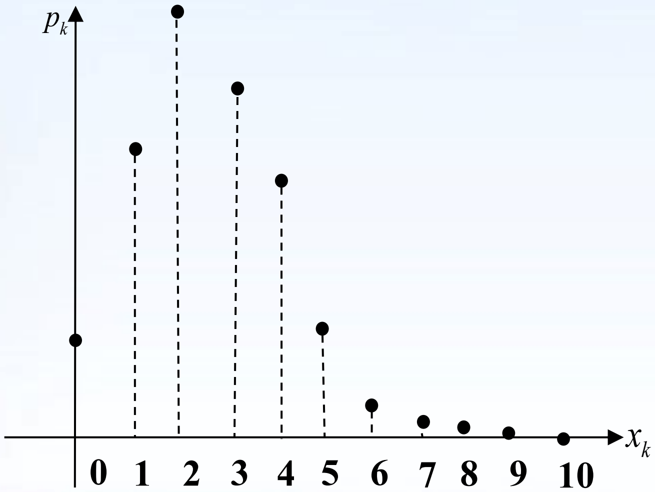
特别地, 0-1分布即为$B(1,p)$.

问题 它否是分布律？需要验证什么？

二项分布分布律的图示法表示

例如 $n=10,\ p=\dfrac{1}{2}$ $n=10,\ p=\dfrac{1}{4}$





例 (电子元件)一大批电子元件有10%已损坏, 若从这批元件中随机选取20只来组成一个线路, 问这线路能正常工作的概率是多少？

解： 因为元件的数量很大，所以取20只元件可看作是有放回抽样, 记X表示20只元件中好品的数量，

则 $X \sim B(20, 0.9).$ 于是

$$P\{\text{线路正常}\} = P\{X = 20\} = \binom{20}{20} \times 0.9^{20} \times 0.1^{20-20} \approx 0.1216.$$

- **注** 二项分布理论上只适用于有放回抽样, 但当n很大时, 也可近似用于无放回抽样.

例 (人寿保险)若一年中某类保险者里面每个人死亡的概率等于0.005，现有10000个人参加这类人寿保险，试求在未来一年中在这些保险者里面，(1) 有40个人死亡的概率; (2) 死亡人数不超过70个的概率.

解： 记X为未来一年中在这些人中死亡的人数，则

$$X \sim B(10000, 0.005)$$

$$(1)\quad P\{X=40\} = \binom{10000}{40} \times 0.005^{40} \times 0.995^{9960} \approx 0.0214$$

$$(2)\quad P\{X \le 70\} = \sum_{k=0}^{70} \binom{10000}{k} \times 0.005^{k} \times 0.995^{10000-k} \approx 0.997?$$

- 小概率事件原理：在利用概率模型进行推断问题中, 常认为小概率事件是不发生的, 如果发生了, 则说明原概率模型不真实.

例 (加州枪劫案)据报道, 加州某地发生一起枪劫案, 目击嫌疑人有两个：一个男的理平头黑人, 一个女的黑发梳马尾型. 不久抓到一对具上述特征的夫妇(情侣), 能否判他们有罪?

数学家通过计算机模拟, 得出一对夫妇具有上述特征的概率为 $p = 8.3\times10^{-8}$. 这是一个小概率事件.
陪审团在无其它证据的情况下, 裁决他们有罪.
而加州高院推翻了该裁决. 高院认为犯罪的认定应有唯一性. 用下列算式：设X为具上述特征的夫妇数, 则 $X\sim B(n, p)$. 要计算

$$P\{X>1\,|\,X\ge1\}=\frac{P\{X>1\}}{P\{X\ge1\}}=\frac{P\{X\ge1\}-P\{X=1\}}{P\{X\ge1\}}=\frac{1-(1-p)^n-np(1-p)^{n-1}}{1-(1-p)^n}$$

n估计为 8.3×10^6，而上式 ≈ 0.2966, 不算小.

3. 泊松分布

设随机变量 X 取值为0, 1, … ，分布律为

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

称X服从参数为 λ 的泊松（Possion）分布，记为$X\sim P(\lambda)$.

问： 它是否是分布律？

应用：

- 一段时间内物理试验仪器捕获的粒子数；
- 一段时间内计算机病毒入侵数；
- 一本书中的错字数；
- ……

例 （排队等候问题）某服务机构有两个服务窗口. 设一段时间内前来访问的人数$X \sim P(1)$. 问在这段时间内, 出现排队等候的概率为多少？

解： $P\{X > 2\} = 1 - P\{X \leq 2\}$

$$= 1 - e^{-1}(1 + 1 + \frac{1}{2}) = 1 - \frac{5}{2e}.$$

例 （疾病分布律）设某地区患某种疾病的人数$X \sim P(\lambda)$，λ未知，若已知患此病的概率为0.001，求X的分布律.

解： $1-e^{-\lambda}=0.001$，$e^{\lambda}=1000/999$，$\lambda = \ln\frac{1000}{999}$

故分布律为

$$P\{X=k\}=0.999\frac{(\ln\frac{1000}{999})^k}{k!},\quad k=0,1,\cdots.$$