信息学中的概率统计

王若松

前沿计算研究中心
北京大学

离散随机变量

1. 随机变量的定义
2. 离散随机变量的概率分布列
3. 离散随机变量函数的分布
4. 数学期望
5. 方差
6. 常用离散分布

1. 随机变量的定义

- 随机变量
  - 勺园餐厅一天用餐的人数
  - 测量误差
  - 掷三枚硬币，正面朝上的硬币数量
  - 考察一批产品，不合格的产品数量

- 定义：对于样本空间$S = \{e\}$，**随机变量** $X: S \to \mathbb{R}$为定义在 S上的实值函数
  - 常用大写字母 X, Y, Z表示
  - 分为**离散**和**连续**随机变量

- 对于取值 x，定义$A = \{e: X(e) = x\}$，则有$P(X = x) = P(A)$
- 随机变量的取值提供了样本空间的一个**划分**

2. 离散随机变量的概率分布列

- 离散随机变量：取**有限个**或**可列个**值的随机变量
  - 掷三枚硬币，正面朝上的硬币数量
  - 勺园餐厅一天用餐的人数

- 若离散随机变量 X全部可能取值为 $x_1, x_2, x_3, \ldots$， $p_i = P(X = x_i)$为X的**概率分布列**，简称**分布列**
- $\sum p_i = 1$且 $p_i \geq 0$

- 例：投掷两枚硬币，两枚硬币正面朝上概率均为p，求正面朝上的硬币数量的分布列
  - $P(X = 0) = (1 - p)^2, P(X = 1) = 2p(1 - p), P(X = 2) = p^2$

2. 离散随机变量的概率分布列

- 某闯关游戏，初始奖金为1元。每多闯一关，则多得一元。通过每一关的概率均为50%。若某一关失败则游戏结束。求最终奖金的分布列？
  - $P(X=1)=\frac{1}{2}, P(X=2)=\frac{1}{4}, P(X=3)=\frac{1}{8}, \ldots$

- 某闯关游戏，初始奖金为0.5元。每多闯一关，则奖金翻倍。通过每一关的概率均为50%。若某一关失败则游戏结束。求最终奖金的分布列？
  - $P(X=0.5)=\frac{1}{2}, P(X=1)=\frac{1}{4}, P(X=2)=\frac{1}{8}, \ldots$

2. 离散随机变量的概率分布列

- 回顾：如果某个随机试验只有两个可能的结果A和$\overline{A}$，且$P(A) = p$，将试验独立地重复进行n次，称为**n重伯努利试验**

- 设随机变量X为n重伯努利试验中A发生的次数
  - 若某个样本点 $e = (e_1, e_2, \ldots, e_n)$ 在事件 $\{X = k\}$ 中，则$P(\{e\}) = p^k(1-p)^{n-k}$
  - 共有$\binom{n}{k}$个样本点在在事件 $\{X = k\}$ 中
  - $P(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}, 0 \leq k \leq n$
  - 是否有$\sum_k P(X = k) = 1$?

- 当 $n = 1$
  - $P(X = 1) = p, P(X = 0) = 1 - p$

3. 随机变量函数的分布

- 给定离散随机变量X和函数g，求$Y = g(X)$的分布列
- 例： $P(X = 1) = 0.1$， $P(X = -1) = 0.1, P(X = 0) = 0.8$，求$Y = X^2$的分布列

- 若 $P(X = x_i) = p_i$，则对于$Y = g(X)$，有$P(Y = g(x_i)) = p_i$
- 若某些$g(x_i)$相等，将概率相加

4. 数学期望

- 如果为了最大化**平均情况**的奖金，两个闯关游戏中应该选择哪一个?
- 如何定义**平均情况**的奖金?

- **数学期望**：对于离散随机变量X，若$\sum_i |x_i p_i| < \infty$，则$X$的数学期望为 $E(X) = \sum_i x_i p_i$

- 例：投掷两枚硬币，两枚硬币正面朝上概率均为p，求正面朝上的硬币数量的数学期望
  - $P(X = 0) = (1 - p)^2, P(X = 1) = 2p(1 - p), P(X = 2) = p^2$
  - $E(X) = 2p(1 - p) \cdot 1 + p^2 \cdot 2 = 2p$

4. 数学期望

- 某闯关游戏，初始奖金为1元。每多闯一关，则多得一元。通过每一关的概率均为50%。若某一关失败则游戏结束。求最终奖金的期望？
  - $P(X=1)=\frac{1}{2}, P(X=2)=\frac{1}{4}, P(X=3)=\frac{1}{8}, \ldots$
  - $E(X)=\frac{1}{2}\cdot 1+\frac{1}{4}\cdot 2+\frac{1}{8}\cdot 3+\cdots=2$

- 某闯关游戏，初始奖金为0.5元。每多闯一关，则奖金翻倍。通过每一关的概率均为50%。若某一关失败则游戏结束。求最终奖金的期望？
  - $P(X=0.5)=\frac{1}{2}, P(X=1)=\frac{1}{4}, P(X=2)=\frac{1}{8}, \ldots$
  - $E(X)=\frac{1}{2}\cdot 0.5+\frac{1}{4}\cdot 1+\frac{1}{8}\cdot 2+\cdots=+\infty$

4. 数学期望

- n重伯努利试验中A发生的次数的数学期望
- $E(X)=\sum_{k=0}^{n}\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}\cdot k=\sum_{k=1}^{n}\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}\cdot k$
- $\binom{n}{k}k=\binom{n-1}{k-1}n$
- $E(X)=n\sum_{k=1}^{n}\binom{n-1}{k-1}\cdot p^k\cdot(1-p)^{n-k}$
- $E(X)=np\sum_{k=1}^{n}\binom{n-1}{k-1}\cdot p^{k-1}\cdot(1-p)^{(n-1)-(k-1)}$
- $E(X)=np\sum_{l=0}^{n-1}\binom{n-1}{l}\cdot p^l\cdot(1-p)^{(n-1)-l}=np\cdot\big(p+(1-p)\big)^{n-1}=np$

4. 数学期望

- 给定离散随机变量X和函数g，求$Y = g(X)$的数学期望
- 例：$P(X = 1) = 0.1$， $P(X = -1) = 0.1, P(X = 0) = 0.8$，求$Y = X^2$的数学期望

- 方法1： $P(Y = 1) = 0.2$， $P(Y = 0) = 0.8$，因此$E(Y) = 0.2$
- 方法2：$E\left(X^2\right) = 0.1 \cdot (-1)^2 + 0.1 \cdot 1^2 + 0.8 \cdot 0 = 0.2$

- **一般结论**：$E(g(X)) = \sum_i p_i g(x_i)$

北京大学 PEKING UNIVERSITY

4. 数学期望

- 期望的性质

- 对于$P(X=c)=1$的**单点分布（退化分布）**，$E(X)=c$

- 对于常数a和b，$E(aX+b)=a\cdot E(X)+b$
  - $E(aX+b)=\sum_i p_i(a\cdot x_i+b)=a\cdot\sum_i p_i x_i+\sum_i p_i\cdot b=a\cdot E(X)+b$

- 对于两个函数g_1和g_2，$E(g_1(X)\pm g_2(X))=E(g_1(X))\pm E(g_2(X))$
  - 例：$E\left(aX-bX^2\right)=a\cdot E(X)-b\cdot E(X^2)$

- 对于事件A，用1_A表示事件A的**示性函数。** A发生时取1，否则取0。则有$P(A)=E(1_A)$

4. 数学期望

- n重伯努利试验中A发生的次数的平方数学期望
- $E\left(X^2\right)=\sum_{k=0}^{n}\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}\cdot k^2=\sum_{k=1}^{n}\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}\cdot k^2$
- $E\left(X^2\right)=\sum_{k=1}^{n}\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}\cdot(k(k-1)+k)$
- $E\left(X^2\right)-E(X)=\sum_{k=1}^{n}\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}\cdot k(k-1)$
- $E\left(X^2\right)-E(X)=\sum_{k=2}^{n}\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}\cdot k(k-1)$
- $\binom{n}{k}k(k-1)=\binom{n-2}{k-2}n(n-1)$
- $E\left(X^2\right)-E(X)=n(n-1)\sum_{k=2}^{n}\binom{n-2}{k-2}\cdot p^k\cdot(1-p)^{n-k}$
- $E\left(X^2\right)-E(X)=n(n-1)p^2\sum_{k=2}^{n}\binom{n-2}{k-2}\cdot p^{k-2}\cdot(1-p)^{(n-2)-(k-2)}$
- $E\left(X^2\right)-E(X)=n(n-1)p^2\sum_{l=0}^{n-2}\binom{n-2}{l}\cdot p^l\cdot(1-p)^{(n-2)-l}=n(n-1)p^2$
- $E\left(X^2\right)=n(n-1)p^2+np$

4. 数学期望

- $P\big(X \geq E(X)\big) > 0$
- $P\big(X \leq E(X)\big) > 0$
  - 常用于概率证法
  - 证明：如果$P\big(X \geq E(X)\big) = 0$，则$E(X) = \sum x_i p_i < \sum E(X) \cdot p_i = E(X)$

- 有n个人，其中有m对人认识。
- 证明：总可以将n个人分为两组，且有至少$\frac{m}{2}$对属于不同组的人认识。

- 将每个人**等概率**地加入第一组和第二组， 用随机变量X表示属于不同组且认识的人的对数。
- 如何求$E(X)$?

4. 数学期望

- $E(X) = \sum_k P(X = k) \cdot k$
- $P(X = k) = \frac{1}{2^n} |\{e : X(e) = k\}|$
- $E(X) = \sum_{e \in S} \frac{X(e)}{2^n}$
- 对于每一对认识的人，有多少个样本点（分组方式）满足两个人属于不同组?

- $\sum_{e \in S} X(e) = m \cdot 2^{n-1}$
- $E(X) = \frac{m}{2}$
- $P\left(X \geq \frac{m}{2}\right) > 0 \Rightarrow$存在 $e \in S$，$X(e) \geq \frac{m}{2}$

4. 数学期望

- 尾不等式/集中不等式：随机变量与其期望的偏离

- **马尔可夫不等式**：若X为**非负**随机变量，若$E(X) > 0$，对于$a > 0$，有

$$P(X \geq a \cdot E(X)) \leq \frac{1}{a}$$

- 证明

$$E(X) = \sum x_i \cdot p_i \geq \sum_{x_i \geq a \cdot E(x)} x_i p_i \geq \sum_{x_i \geq a \cdot E(X)} a \cdot E(X) \cdot p_i$$

$$= a \cdot E(X) \cdot \sum_{x_i \geq a \cdot E(X)} p_i = a \cdot E(X) \cdot P(X \geq a \cdot E(X))$$

4. 数学期望

- **马尔可夫不等式**：若X为**非负**随机变量，若$E(X)>0$，对于$a>0$，有
- $P(X\geq a\cdot E(X))\leq \frac{1}{a}$
- 例： 投掷n枚硬币，有超过$3n/4$枚朝上的概率?
  - $E(X)=n/2$，$P\left(X\geq \frac{3n}{4}\right)=P\left(X\geq \frac{3}{2}\cdot E(x)\right)\leq \frac{2}{3}$
  - 是否可以证明更强的结果?
- 作业：对于任意$a\geq 1$，均可找到非负随机变量X满足$P\left(X\geq a\cdot E(X)\right)=\frac{1}{a}$

4. 数学期望

- 数学期望描述了随机变量在**平均情况**下的取值
- 随机变量在其数学期望周围**波动**
- 如何定义波动的大小?

- $E(X-E(X))$?
  - $E(X-E(X))=E(X)-E(X)=0$
- $E(|X-E(X)|)$?
  - 难以计算

5. 方差

- 给定随机变量X，若$E\left[\left(X-E(X)\right)^2\right]$存在，定义$\mathrm{Var}(X)=E\left[\left(X-E(X)\right)^2\right]$为$X$的**方差**

- 例：$P(X=1)=p$，$P(X=0)=1-p$，求$\mathrm{Var}(X)$
  - $\mathrm{Var}(X)=p\cdot(1-p)^2+(1-p)\cdot(0-p)^2=p(1-p)$

- 例：某闯关游戏，初始奖金为1元。每多闯一关，则多得一元。若某一关失败则游戏结束。求最终奖金的方差?
  - $P(X=1)=\frac{1}{2},P(X=2)=\frac{1}{4},P(X=3)=\frac{1}{8},\ldots$
  - $E(X)=2,\mathrm{Var}(X)=\frac{1}{2}\cdot(1-2)^2+\frac{1}{4}\cdot(2-2)^2+\frac{1}{8}\cdot(3-2)^2+\cdots=2$

PEKING UNIVERSITY

5. 方差

- 例：某闯关游戏，初始奖金为10角。每多闯一关，则多得10角。若某一关失败则游戏结束。求最终奖金的方差?
  - $P(X=10)=\frac{1}{2}, P(X=20)=\frac{1}{4}, P(X=30)=\frac{1}{8}, \ldots$
  - $E(X)=20$
  - $\mathrm{Var}(X)=E\left[\left(X-E(X)\right)^2\right]$
  - $\mathrm{Var}(X)=\frac{1}{2}\cdot(10-20)^2+\frac{1}{4}\cdot(20-20)^2+\frac{1}{8}\cdot(30-20)^2+\cdots=200$

- 方差的单位?

5. 方差

- 给定随机变量X，定义$\sigma(X)=\sqrt{\mathrm{Var}(X)}$为$X$的**标准差**
- 标准差与数学期望**量纲**相同

- 例：某闯关游戏，初始奖金为10角。每多闯一关，则多得10角。若某一关失败则游戏结束。求最终奖金的标准差?
  - $P(X=10)=\frac{1}{2}, P(X=20)=\frac{1}{4}, P(X=30)=\frac{1}{8}, \ldots$
  - $E(X)=20$
  - $\mathrm{Var}(X)=\frac{1}{2}\cdot(10-20)^2+\frac{1}{4}\cdot(20-20)^2+\frac{1}{8}\cdot(30-20)^2+\cdots=200$
  - $\sigma(X)=\sqrt{\mathrm{Var}(X)}=10\sqrt{2}$

PEKING UNIVERSITY

5. 方差

- 方差的性质

- 对于常数a和b，$\mathrm{Var}(aX+b)=a^2\,\mathrm{Var}(X)$，$\sigma(aX+b)=|a|\cdot\sigma(X)$
  - $\mathrm{Var}(aX+b)=E\left(\left(aX+b-E(aX+b)\right)^2\right)=E\left(\left(aX-aE(X)\right)^2\right)=a^2\mathrm{Var}(X)$

- 对于$P(X=c)=1$的单点分布，方差和标准差?
  - $\mathrm{Var}(X)=E\left((X-c)^2\right)=0$
  - $\sigma(X)=0$
  - 若随机变量X满足$\sigma(X)=0$，是否一定服从单点分布?

PEKING UNIVERSITY

5. 方差

- 方差的性质

- $\mathrm{Var}(X)=E\left(X^2\right)-(E(X))^2$

- **证明：** $\mathrm{Var}(X)=E\left(\left(X-E(X)\right)^2\right)=E\left(X^2-2X\cdot E(X)+\left(E(X)\right)^2\right)$
- $E\left(X^2-2X\cdot E(X)+\left(E(X)\right)^2\right)=E\left(X^2-2X\cdot E(X)\right)+\left(E(X)\right)^2$
- $E\left(X^2-2X\cdot E(X)\right)=E\left(X^2\right)-2E(X)\cdot E(X)=E\left(X^2\right)-2\cdot\left(E(X)\right)^2$

- **推论**： $E\left(X^2\right)\geq(E(X))^2$

5. 方差

- 设随机变量X为 n重伯努利试验中A发生的次数，求$\mathrm{Var}(X)$和$\sigma(X)$

- $\mathrm{Var}(X) = E\left(X^2\right) - \left(E(X)\right)^2$
- $E(X) = np$
- $E\left(X^2\right) = n(n-1)p^2 + np$
- $\mathrm{Var}(X) = np(1-p)$
- $\sigma(X) = \sqrt{np(1-p)}$

5. 方差

- **切比雪夫不等式**：若 $\sigma(X) > 0$，对于任意$c > 0$，

$$P(|X - E(X)| \geq c \cdot \sigma(X)) \leq 1/c^2$$

- 证明：考虑随机变量 $Y = \left(X - E(X)\right)^2$
- 对 Y使用马尔可夫不等式：对于任意常数$c > 0$，$P\left(Y \geq c^2 \cdot E(Y)\right) \leq 1/c^2$
- $P\left(\left(X - E(X)\right)^2 \geq c^2 \cdot \mathrm{Var}(X)\right) \leq 1/c^2$
- $P(|X - E(X)| \geq c \cdot \sigma(X)) \leq 1/c^2$

5. 方差

- **切比雪夫不等式**：若 $\sigma(X) > 0$，对于任意$c > 0$，

$$P(|X - E(X)| \geq c \cdot \sigma(X)) \leq 1/c^2$$

- 例：投掷n枚硬币，有超过 $3n/4$枚朝上的概率?
  - $E(X) = n/2, \mathrm{Var}(X) = n/4, \sigma(X) = \sqrt{n}/2$
  - $P\left(|X - E(X)| \geq \frac{4}{n}\right) \leq \frac{4}{n}$
  - $P\left(X - E(X) \geq \frac{4}{n}\right) = P\left(X - E(X) \geq -\frac{4}{n}\right)$
  - $P\left(X - E(X) \geq \frac{4}{n}\right) \leq \frac{2}{n}$
  - 马尔可夫不等式：$P\left(X \geq \frac{3n}{4}\right) = P\left(X \geq \frac{3}{2} \cdot E(x)\right) \leq \frac{2}{3}$

5. 方差

- 例：某课程平均分为70分，90分及以上为优秀。

- 利用马尔可夫不等式给出优秀率的上界?
  - 随机变量 X表示一名随机的学生的成绩，则$E(X) = 70$
  - $P(X \geq 90) = P\left(X \geq \frac{9}{7} \cdot E(X)\right) \leq \frac{7}{9}$

- 已知成绩的标准差为5分，利用切比雪夫不等式给出优秀率的上界?
  - $P(X \geq 90) \leq P(|X - E(X)| \geq 20) = P(|X - E(X)| \geq 4 \cdot \sigma(X)) \leq 1/16$

PEKING UNIVERSITY 北京大学

6. 常用离散分布

- 回顾：如果某个随机试验只有两个可能的结果A和$\overline{A}$，且$P(A) = p$，将试验独立地重复进行n次，称为**n重伯努利试验**

- 伯努利分布（0-1分布、两点分布）：单次伯努利试验的结果
- 随机变量X取0或1， $p = P(X = 1)$，称X服从参数为p的**伯努利分布（0-1分布、两点分布）**
  - 例：投硬币的结果、CPU测试检测结果

6. 常用离散分布

- 回顾：如果某个随机试验只有两个可能的结果A和$\overline{A}$，且$P(A) = p$，将试验独立地重复进行n次，称为**n重伯努利试验**

- 二项分布：n重伯努利试验中A发生的次数
- 随机变量X为n重伯努利试验中A发生的次数，则$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}, 0 \leq k \leq n$。称$X$服从**参数为$n, p$的二项分布**，记为$X \sim B(n, p)$

- $B(1, p)$：参数为p的伯努利分布
- 若$X \sim B(n, p)$，$E(X) = np$，$\mathrm{Var}(X) = np(1 - p)$

- 例：投n次硬币，正面朝上的次数

6. 常用离散分布

- 球与桶模型：有n个球，每个球都等可能被放到m个桶中的任一个。
- 某个特定桶中球（如第i个）的数量X_i服从何种分布？
  - 参数为$n, 1/m$的二项分布

- 考虑$n = \lambda m$（λ为常数），且$m \to \infty$的情况

- $P(X_i = k) = \binom{n}{k} \cdot \left(\frac{1}{m}\right)^k \cdot \left(1 - \frac{1}{m}\right)^{n-k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \cdot \left(\frac{1}{m}\right)^k \cdot \left(1 - \frac{1}{m}\right)^{n-k}$
- $= \frac{1}{k!} \cdot \frac{n}{m} \cdot \frac{n-1}{m} \cdot \ldots \cdot \frac{n-k+1}{m} \cdot \left(1 - \frac{1}{m}\right)^{n-k} \approx \frac{1}{k!} \cdot \lambda^k \cdot e^{-\lambda}$

6. 常用离散分布

- 泊松分布：二项分布的极限
- 随机变量X取非负整数，且$P(X=k)=\frac{1}{k!}\cdot\lambda^k\cdot e^{-\lambda}$，其中$\lambda>0$，则称$X$服从**参数为$\lambda$的泊松分布**，记为$X\sim\pi(\lambda)$

- 若$X\sim\pi(\lambda)$，是否满足$\sum_k P(X=k)=1$?
  - $\sum_{k\geq 0}\frac{1}{k!}\cdot\lambda^k\cdot e^{-\lambda}=e^{-\lambda}\cdot\sum_{k\geq 0}\frac{\lambda^k}{k!}=e^{-\lambda}\cdot e^{\lambda}=1$
- 若$X\sim\pi(\lambda)$，如何计算$E(X)$?
  - $\sum_{k\geq 0}\frac{1}{k!}\cdot\lambda^k\cdot e^{-\lambda}\cdot k=e^{-\lambda}\cdot\sum_{k\geq 1}\frac{1}{(k-1)!}\cdot\lambda^k=e^{-\lambda}\cdot\lambda\cdot\sum_{k\geq 1}\frac{1}{(k-1)!}\cdot\lambda^{k-1}=\lambda$
- 若$X\sim\pi(\lambda)$，如何计算$\mathrm{Var}(X)$?

6. 常用离散分布

- 若$X \sim \pi(\lambda)$，如何计算$\mathrm{Var}(X)$?
- 首先计算 $E\left(X^2\right)$
  - $\sum_{k\geq 0}\frac{1}{k!}\cdot\lambda^k\cdot e^{-\lambda}\cdot k^2 = \sum_{k\geq 0}\frac{1}{k!}\cdot\lambda^k\cdot e^{-\lambda}\cdot k(k-1) + \sum_{k\geq 0}\frac{1}{k!}\cdot\lambda^k\cdot e^{-\lambda}\cdot k$
  - $\sum_{k\geq 0}\frac{1}{k!}\cdot\lambda^k\cdot e^{-\lambda}\cdot k(k-1) = e^{-\lambda}\cdot\lambda^2\cdot\sum_{k\geq 2}\frac{1}{(k-2)!}\cdot\lambda^{k-2} = \lambda^2$

- $\mathrm{Var}(X) = E\left(X^2\right) - \left(E(X)\right)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$

北京大学 PEKING UNIVERSITY

6. 常用离散分布

- 二项分布与泊松分布的关系
- **泊松定理**：给定 $p_1, p_2, \ldots$ 满足 $\lim_{n\to+\infty} n\cdot p_n = \lambda$，则对于$k>0$，
- $\lim_{n\to+\infty}\binom{n}{k}\cdot p_n^k\cdot(1-p_n)^{n-k}=\frac{1}{k!}\cdot\lambda^k\cdot e^{-\lambda}$

- 证明：令$\lambda_n = n\cdot p_n$，则有
- $\binom{n}{k}\cdot p_n^k\cdot(1-p_n)^{n-k}=\frac{n(n-1)\cdots(n-k+1)}{k!}\cdot\left(\frac{\lambda_n}{n}\right)^k\left(1-\frac{\lambda_n}{n}\right)^{n-k}$
- $=\frac{\lambda_n^k}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda_n}{n}\right)^{n-k}$
- 其中$\lim_{n\to+\infty}\lambda_n=\lambda$，所以$\lim_{n\to+\infty}\left(1-\frac{\lambda_n}{n}\right)^{n-k}=e^{-\lambda}$

PEKING UNIVERSITY

6. 常用离散分布

- 将单位时间分为n段，每一段时间内事件发生的概率为p_n，假设$np_n \to \lambda$
- 单位时间内事件发生的次数服从$B(n, p_n)$
- 对n取极限，则事件发生的次数服从$\pi(\lambda)$

- 例子：
  - 每一瞬间CPU发生故障的概率均相等，则在一年内CPU发生故障的次数服从泊松分布
  - 每时每刻看到流星的概率是均匀随机，在一个晚上看到的流星数量服从泊松分布

6. 常用离散分布

- 回顾：如果某个随机试验只有两个可能的结果A和$\overline{A}$，且$P(A) = p$，将试验独立地重复进行n次，称为**n重伯努利试验**

- 几何分布：伯努利试验中首次发生结果A时的重复次数
- 在伯努利试验中，令随机变量X为结果A首次出现时的试验次数。对于$k \geq 1$，$P(X = k) = p(1-p)^{k-1}$。称X服从**参数为 p的几何分布**，记为$X \sim G(p)$

- 若$X \sim G(p)$，是否满足$\sum_k P(X = k) = 1$?
  - $\sum_{k \geq 1} p(1-p)^{k-1} = 1$
- 若$X \sim G(p)$，如何计算$E(X)$?

PEKING UNIVERSITY

6. 常用离散分布

- 若$X \sim G(p)$，如何计算$E(X)$?
  - $E(X) = \sum_{k\geq 1} p(1-p)^{k-1} \cdot k$
  - 令$f(p) = \sum_{k\geq 1}(1-p)^k$，则$E(X) = -p \cdot f'(p)$
  - 而$f(p) = \frac{1}{p}$，所以$E(X) = \frac{1}{p}$
- 若$X \sim G(p)$，如何计算$\mathrm{Var}(X)$?
  - $E(X^2) = \sum_{k\geq 1} p(1-p)^{k-1} \cdot k^2 = \sum_{k\geq 1} p(1-p)^{k-1} \cdot k(k-1) + \frac{1}{p}$
  - $\sum_{k\geq 1} p(1-p)^{k-1} \cdot k(k-1) = p(1-p)\sum_{k\geq 1}(1-p)^{k-2} \cdot k(k-1)$
  - 令$f(p) = \sum_{k\geq 1}(1-p)^k$，则$E(X^2) = p(1-p)f''(p) + \frac{1}{p} = \frac{2(1-p)}{p^2} + \frac{1}{p}$
  - $\mathrm{Var}(X) = E(X^2) - (E(X))^2 = \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}$

6. 常用离散分布

- 无记忆性：若$X \sim G(p)$，对于任意正整数m和n
- $P(X > m + n \mid X > m) = P(X > n)$

- 证明：注意到 $P(X > n) = \sum_{k>n} p(1-p)^{k-1} = (1-p)^n$

6. 常用离散分布

- 负二项分布：伯努利试验中结果A发生r次时的重复次数
- 在伯努利试验中，令随机变量X为结果A第 r次出现时的试验次数。对于$k \geq r$，$P(X = k) = \binom{k-1}{r-1}p^r(1-p)^{k-r}$。称$X$服从**参数为 $\boldsymbol{r,p}$的负二项分布**，记为$X \sim NB(r,p)$

- $NB(1,p)$：参数为p的几何分布
- 若$X \sim NB(r,p)$，是否满足$\sum_k P(X = k) = 1$?

6. 常用离散分布

- 若$X \sim NB(r,p)$，是否满足$\sum_k P(X=k)=1$?

- $\sum_{k\geq r}\binom{k-1}{r-1}p^r(1-p)^{k-r} = \sum_{l\geq 0}\binom{l+r-1}{r-1}p^r(1-p)^l$

- $p^{-r} = \big(1-(1-p)\big)^{-r} = 1+(-r)\big(-(1-p)\big)+\frac{(-r)(-r-1)}{2!}\big(-(1-p)\big)^2+\cdots$
- $= 1+r(1-p)+\frac{r(r+1)}{2!}(1-p)^2+\frac{r(r+1)(r+2)}{3!}(1-p)^3+\cdots$
- $= \sum_{l\geq 0}\binom{l+r-1}{l}(1-p)^l$
- $= \sum_{l\geq 0}\binom{l+r-1}{r-1}(1-p)^l$

- $\sum_{k\geq r}\binom{k-1}{r-1}p^r(1-p)^{k-r} = p^r\cdot\ p^{-r} = 1$

6. 常用离散分布

- 在伯努利试验中，令随机变量X为结果A第 r次出现时的试验次数。
- 令X_1为结果A第一次出现时的试验次数，X_2为结果A第二次出现时的试验次数（从结果A第一次出现之后算起），X_3为结果A第三次出现时的试验次数（从结果A第二次出现之后算起），以此类推。
- $X = X_1 + X_2 + \cdots + X_r$
- 若$X \sim NB(r, p)$，则X可表示成 r个服从参数为p的几何分布的随机变量之和。