概率论与数理统计

主讲教师：朱丽娜　讲师

研究方向：智能交通，车联网与智能驾驶

电子邮件：lnzhu@xidian.edu.cn

个人主页：http://web.xidian.edu.cn/lnzhu/

第二章 随机变量及其分布

- § 2.1 随机变量
- § 2.2 离散型随机变量及其概率分布
- § 2.3 随机变量的分布函数
- § 2.4 连续型随机变量及其概率密度
- § 2.5 随机变量的函数的分布

第二章 随机变量及其分布

- § 2.1 随机变量
- § 2.2 离散型随机变量及其概率分布
- § 2.3 随机变量的分布函数
- § 2.4 连续型随机变量及其概率密度
- § 2.5 随机变量的函数的分布

§ 2.1 随机变量

- 对于错综复杂的随机现象，如信道噪声、随机信号、测量误差等实际问题难以简单处理。为了更好的用数学方法来分析随机现象地统计规律性，人们将随机试验的结果与实数对应起来(结果数量化)，从而引入随机变量的概念

- 例1：将一枚硬币抛掷三次，观察正反面出现的情况，并考虑每次试验当中正面出现的次数，样本空间是
  - S＝{HHH，HHT，HTH，HTT，THH，THT，TTH，TTT}
  - 以X记三次投掷得到正面H的总数，那么对于样本空间S中的每一个样本点*e*，X都有一个数与之对应
  - X是定义在样本空间上的单值实值函数，定义域是样本空间，值域为{0,1,2,3}.使用函数的符号可将X写成

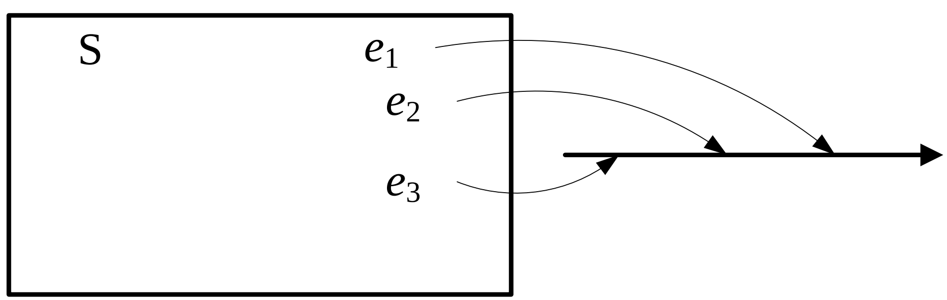
$$X = X(e) = \begin{cases} 3, e = HHH, \\ 2, e = HHT, HTH, THH, \\ 1, e = HTT, THT, TTH, \\ 0, e = TTT \end{cases}$$

§2.1 随机变量

- **例2：随机试验：袋中装有编号分别为1，2，3的3只球，在袋中任取一只，放回，然后再取一只球，记录他们的编号**
  - **样本点：是一个编号的数对：(i，j)，i，j=1,2,3**
  - **样本空间：S={e}={(i，j)|i，j=1,2,3}**
  - **现在关心的是两个球的号码之和，记做X，则对于每一个样本点e，X都有一个值与之对应**
  - **即从样本空间到实数集合上的一个映射 X：S→R**
    - **X是一个定义在样本空间S上的单值实值函数，**
    - **其定义域是样本空间；值域是实数集合{2,3,4,5,6}**
    - **因此X可写成 $X=X(e)=X((i,j))=i+j$，i，j＝1，2，3.**

§ 2.1 随机变量

- **定义：**
  - **设随机试验的样本空间为S＝{e}。X=X{e}是定义在样本空间S上的实值单值函数。称X=X{e}为随机变量**



§2.1 随机变量

- **关于随机变量需要注意的几个方面**
  - **(1) 如果随机试验结果本身是一个数，则直接令X＝X(e)=e，X是一个随机变量**
    - **如灯泡的寿命T，某学校学生的体重W，掷骰子的点数等等**
    - **随机变量的取值可以是有限的，可列的，或不可列的**
  - **(2) X是一个单值实值函数，定义域是样本空间S，值域是X的所有可能取值R_X。注意随机变量的值域不同于样本空间**
  - **(3) 随机变量的取值随试验结果而定，是样本点的函数(不同于普通函数)，在试验之前不能预知其具体取值，因此随机变量X(e)的取值是随机出现的，有一定的概率**

§ 2.1 随机变量

- **随机事件的描述**
  - 根据随机变量X的前述映射关系，可用X的取值集合来描述随机事件
- **例如**
  - 例1中X的取值为2，记做{X=2}，对应的样本点集合为A={HHT,HTH,THH}这是一个随机事件
  - 当且仅当A发生时有{X=2}，我们称概率P(A)为{X=2}的概率，即$P\{X=2\}=P(\mathrm{A})=3/8$
- **当然关心的X取值也可能有多个，用关于X的表达式来表示**
  - 如{X≤1}表示随机事件B＝{TTT,TTH,THT,HTT}
  - $P\{X\le 1\}=P(\mathrm{B})=4/8$
- **一般的，若L是一个实数集合，将X在L上取值写成{X∈L}，则{X∈L}表示事件B={e| X(e)∈L },此时有$P\{X\in L\}=P(\mathrm{B})$**

第二章 随机变量及其分布

§ 2.2 离散型随机变量及其概率分布

- 离散型随机变量
- 例子:
  - 将一枚硬币抛掷三次，观察出现正面的次数:
    - 随机变量的全部可能取值仅有4个：0，1，2，3
  - 某电话交换台一分钟内收到的呼叫次数:
    - 可列无限多个（理想状态下）
  - 某城市120急救电话台一昼夜收到的呼叫次数
    - 可列无限多个（理想状态下）

离散型随机变量X： 它全部可能取到的不相同的值是有限个或可列无限多个，称为离散型随机变量

§ 2.2 离散型随机变量及其概率分布

- 要掌握离散随机变量X的统计规律，需且只需知道X的两个问题
  - (1) X的所有可能的取值，我们可以通过随机试验的样本空间S来得到
  - (2) 每一个可能取值的概率，它们构成分布律的概念
  - 在后面我们会进一步学习X的数字特征等概念
- 分布律：
  - 设离散型随机变量X的所有可能取值为$x_k(k=1,2,\ldots)$，X取各个可能值的概率，即事件$\{X=x_k\}$的概率为
  - $$P\{X=x_k\}=p_k，k=1, 2, \ldots$$
  - 由概率的定义，p_k满足如下两个条件：
  - (1) 非负性：$p_k\geq 0$;　　(2) 规范性：$\sum_{k=1}^{\infty} p_k=1$
  - 则称$P\{X=x_k\}=p_k$，$k=1,2,\ldots$为离散型随机变量X的分布律

§ 2.2 离散型随机变量及其概率分布

- **分布律也可表示为表格的形式**

X	x_1	x_2	…	x_n	…
p_k	p_1	p_2	…	p_n	…

- **分布律含义：由分布律定义中的条件，概率1以一定的规律分布在各个可能值上**
- **规范性的证明：**
  - $\{X=x_1\}\cup\{X=x_2\}\cup\ldots$是必然事件，与样本空间相对应，且 $\{X=x_k\}\cap\{X=x_j\}=\Phi$，$k\neq j$
- **所以，$1=P(\{X=x_1\}\cup\{X=x_2\}\cup\ldots)$**
- $$=P[\bigcup_{k=1}^{\infty}\{X=x_k\}]=\sum_{k=1}^{\infty}P\{X=x_k\}\text{即}\sum_{k=1}^{\infty}p_k=$$

§ 2.2 离散型随机变量及其概率分布

例 1设随机变量 ***X*** 的分布律为：

$\boldsymbol{P(X=n)=c/4^n}$ $(n=1,2,\ldots,)$,

求常数c.

解：由规范性

$$1=\sum_{k=1}^{\infty}P\{X=x_k\}=\sum_{k=1}^{\infty}c/4^k=(c/4)/(1-1/4)=c/3$$

$\therefore\ c=3$

§ 2.2 离散型随机变量及其概率分布

- 几种重要的离散型随机变量的分布律

（一）(0－1)分布

- 设随机变量X只可能取两个值0与1，它的分布律是
  - $P\{X=k\}=p^k(1-p)^{1-k}$，$k=0,1$，$0<p<1$，

X	0	1
p_k	$1-p$	p

- 则称X服从(0－1)分布或两点分布，记作 $b(1,p)$
- 应用：(0－1)分布是经常遇到的一种分布
  - 一般的对于一个随机试验，如果它的样本空间只包含两个元素，即 $S=\{e_1,e_2\}$，总能在S上定义一个服从(0－1)分布的随机变量描述这个随机试验结果
  - $$X=X(e)=\begin{cases}0, & 当e=e_1\\ 1, & 当e=e_2\end{cases}。$$
- 例如：对新生儿的性别进行登记
  - 抛一枚硬币，观察其正面和反面出现的情况

§ 2.2 离散型随机变量及其概率分布

- （二）伯努利试验、二项分布
- 伯努利试验Bernouli
  - 设试验E只有两个可能结果：A及$\overline{A}$，则称E为伯努利试验。设$P(A)=p$ $(0<p<1)$，则$P(\overline{A})=1-p$。将试验E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
  - 其中，“重复”是指每次试验都有$P(A)=p$
  - “独立”是指各次试验的结果A互不影响，相互独立
- n重伯努利试验应用非常广泛，是研究最多的模型之一
  - (0-1)分布的模型就是1重伯努利实验
  - E :抛一枚硬币，观察H和T出现情况，将E独立地进行n次，观察H次数
  - 而一些不放回抽样的随机试验则可能不满足这种独立性，也就不是n重伯努利试验。

§2.2 离散型随机变量及其概率分布

- **伯努利试验的分布律**
  - **以X表示n重伯努利试验中事件A发生的次数（比如n次抛币试验中正面出现的频数），求X的分布律，即求**
  - **1）X的所有可能取的值为0，1，2，...，n**
  - **2）X的分布律，即求对任意的$k(0\le k\le n)$，概率$P\{X=k\}$**
- **解：$P\{X=k\}$即相当于求在n次试验中，有k次试验事件A发生，另外$n-k$次试验事件A不发生的概率**

§ 2.2 离散型随机变量及其概率分布

- 首先看一下事件A在某指定k次试验中发生，比如在第i_1，i_2，…，i_k次试验中A发生，而其余n－k次试验中不发生的概率。由于这n次试验是相互独立的，则相应概率为
  - $p \bullet p \bullet \ldots \bullet p \bullet (1-p) \bullet (1-p) \bullet \ldots \bullet (1-p) = p^k(1-p)^{n-k}$
  - $\uparrow \ \uparrow \quad \uparrow$
  - $i_1 \ i_2 \ldots \ i_k$
- 这样的指定方式共有C_n^k种两两互不相容的方式（两两不同的方式）
- 因此n次试验中事件A发生k次的概率为$C_n^k p^k(1-p)^{n-k}$，令$q=1-p$有
  - $P\{X=k\}= C_n^k p^k q^{n-k}$，$k=0$，1，…，n

§2.2 离散型随机变量及其概率分布

- **分布律的两个条件的验证：**
  - $P\{X=k\}=C_n^k p^k q^{n-k}\geq 0$，$k=0$，1，2，…，n
  - 而 $\sum_{k=0}^{n} P\{X=k\}=\sum_{k=0}^{n} C_n^k p^k q^{n-k}=(p+q)^n=1$
  - 即非负性和归一性均满足
    - 所以$P\{X=k\}=C_n^k p^k q^{n-k}$，$k=0$，1，…，n即为所求分布
  - 由于$p^k q^{n-k}$恰好是$(p+q)^n$的二项展开式中出现p^k的那一项，故称随机变量X服从参数为n，p的二项分布，记为$X\sim b(n, p)$，n＝1时化为(0-1)分布

§ 2.2 离散型随机变量及其概率分布

- 二项分布中的概率最大项:

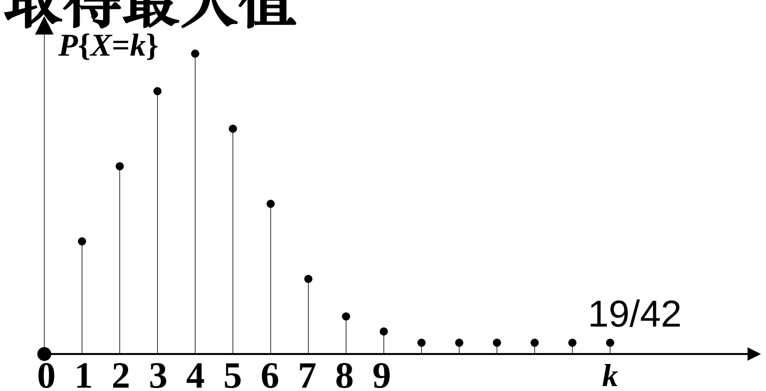
先求分布律中的极大值点，它应该满足如下不等式组

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

2）$P\{X=k\}=C_n^k p^k q^{n-k} \geq C_n^{k+1} p^{k+1} q^{n-k-1}=P\{X=k+1\}$

- 由1)得$k\leq(n+1)p$　由2)得$k\geq(n+1)p-1$
- 联立1)和2) 有$(n+1)p-1\leq k\leq(n+1)p$
- 当$(n+1)p$为整数时，$k=(n+1)p$和$(n+1)p-1$时取得最大值
- 当$(n+1)p$为非整数，$k=[(n+1)p]$时取得最大值

- 二项分布的一般图示:



例2：近似二项分布

- 按规定，某种型号的电子元件的使用寿命超过1500小时的为一级品。已知某一大批产品的一级品率为0.2，现在随机抽取20只，问20只元件中恰有k只为一级品的概率是多少?

解:

- 本题是不放回抽样问题，由于元件总数很大，而**抽查元件数量远远小于元件总数**，可以近似当作放回抽样来处理。
- A：抽取一只元件是一级品　$P(A)=0.2$
- 如果看作放回抽样，每次检查元件是相互独立的，则检查20只元件相当于做20重伯努利试验，以X记20只元件中一级品的只数，那么随机变量X服从二项分布，即$X\sim b(20, 0.2)$即有
- $P\{X=k\}=C_{20}^{k}\,0.2^{k}0.8^{20-k}$，$k=0，1，2，\ldots，20.$

§ 2.2 离散型随机变量及其概率分布

- **例3 大量试验条件下的小概率事件问题**
  - 某人进行射击，设每次击中的命中率为0.02，独立射击400次，试求至少击中两次的概率

一次击中目标是个小概率事件

- **解:**
  - 将一次射击看成是一次试验，设射击次数为400，则X～$b(400, 0.02)$
  - X的分布律为$P\{X=k\}=C_{400}^{k}0.02^{k}0.98^{400-k}$，$k=0$，1，2，…，400.
  - 所求的概率为$P\{X\geq 2\}=1-P\{X=0\}-P\{X=1\}$

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

$$=0.9972$$

结果接近于1

大量独立试验条件下，小概率发生几乎是肯定的，如抽奖问题

若400次中击中目标次数竟真的达不到两次，而由于这一事件的概率$P\{X<2\}=0.003$是小概率事件，由实际推断原理，有理由怀疑命中率达不到0.02，可能更低

§ 2.2 离散型随机变量及其概率分布

例4 维修工分配问题(用工优化)

- 设有80台同类型设备，各台工作独立，发生故障的概率都是0.01，且一台设备的故障仅能由一个人处理，考虑两种配备维修工人的方法

  (1)由4人维护，每人20台

  (2)由3人共同维护80台

  试比较发生故障时不能及时维修的概率大小?

解：按方案(1) 不考虑维修时间长短问题

- 记X为 “第一人维护的20台中同一时刻发生故障的台数“，
- 则$X \sim b(20, 0.01)$
- 用事件A_i表示：“第i人维护的20台中发生故障不能及时维修”这一事件
- 则80台中发生故障不能及时维修的概率为

$$P(A_1 \cup A_2 \cup A_3 \cup A_4) \geq P(A_1) = P(X \geq 2)$$

§ 2.2 离散型随机变量及其概率分布

- 而X～$b(20, 0.01)$，故有

$P(X\geq 2)= 1-P(X=0)-P(X=1)=1-0.99^{20}-20(0.01)0.99^{19}=0.0169$

即有$P(A_1\cup A_2\cup A_3\cup A_4)\geq 0.0169$

按方案(2)

- 记Y为80台中同一时刻发生故障的台数，则Y～$b(80, 0.01)$
- 发生故障不能及时维修的概率
- $P(A)=1-P(Y=0)-P(Y=1)-P(Y=2)-P(Y=3)$
- $=1-0.99^{80}-C_{80}^{1}(0.01)0.99^{79}-C_{80}^{2}(0.01)^{2}0.99^{78}-C_{80}^{3}(0.01)^{3}0.99^{77}$
- $=0.0087<<0.0169$
- 方案(2)不能及时维修的概率远远小于方案(1)
- 因此方案(2)的用工方案更优，而且节省人力

§ 2.2 离散型随机变量及其概率分布

（三）泊松分布

- 泊松分布是概率论中比较重要的一种分布，描述一系列在给定时间区间内到达某区域的随机质点数的分布规律，在排队论、路由优化等方面有广泛的应用...
- 如：到达某十字路口的汽车数、电话交换台接到的要求通话的呼唤数、纺织机上出现的断点数、到达某区域的放射性粒子数、人类的生育胎数、到达某铁路售票处要求售票的顾客数、某一秒种内路由器等待转发的数据包的个数(队列长度)
- 主要描述那些具有无穷可列个取值的离散型随机变量
- 这类问题的显著特点是可列个取值的总概率为1，服从泊松分布的数学模型在第十章，在学习随机过程时会接触到，这里直接给出泊松分布的定义形式

§ 2.2 离散型随机变量及其概率分布

- **泊松分布:**
  - 设随机变量X的所有可能取值为0，1，2，…，
  - 而取各值的概率为 $P\{X=k\}=\dfrac{\lambda^k e^{-\lambda}}{k!}$，$k=0$，1，2，…
  - 其中$\lambda>0$为常数(也叫泊松强度，具体含义在第四章介绍)，则称X服从参数为λ的泊松分布，记为X～$\pi(\lambda)$

- **泊松分布关于分布律的两个条件**
  - 非负性是显然的
  - 规范性: $\sum\limits_{k=0}^{\infty} P\{X=k\} = \sum\limits_{k=0}^{\infty}\dfrac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda}\sum\limits_{k=0}^{\infty}\dfrac{\lambda^k}{k!}$
  -
  - 考虑麦克劳林级数展开 $e^x=1+x+\dfrac{x^2}{2!}+\dfrac{x^3}{3!}+\cdots=\sum\limits_{k=0}^{\infty}\dfrac{x^k}{k!}$
  - 所以 $\sum\limits_{k=0}^{\infty} P\{X=k\} = e^{-\lambda}\sum\limits_{k=0}^{\infty}\dfrac{\lambda^k}{k!}=e^{-\lambda}e^{\lambda}=1$

§ 2.2 离散型随机变量及其概率分布

- **泊松分布的概率最大值问题**
  - 与二项分布的求解方法类似，先找极大值点，它应满足如下不等式组
  - 1）$P\{X=k\}=\dfrac{\lambda^k e^{-\lambda}}{k!} \geq \dfrac{\lambda^{k-1} e^{-\lambda}}{(k-1)!}=P\{X=k-1\}$
  - 2）$P\{X=k\}=\dfrac{\lambda^k e^{-\lambda}}{k!} \geq \dfrac{\lambda^{k+1} e^{-\lambda}}{(k+1)!}=P\{X=k+1\}$
  - 由1）得$k\leq\lambda$　由2得$k\geq\lambda-1$
  - 联立1）和2）有$\lambda-1\leq k\leq\lambda$
- **当λ为整数时，**
  - $k=\lambda$和$\lambda-1$时取得最大值
- **当λ为非整数，**
  - $k=[\lambda]$时取得最大值

三种分布之间的关系：
(1) 0－1分布考察一次试验中两个可能结果的概率
(2) 二项分布是把以上试验独立重复进行n次，考察出现某一个结果的次数的概率，n＝1时即是结果0－1分布
(3) 泊松分布考察一段时间内某一事件发生的次数的概率

§ 2.2 离散型随机变量及其概率分布

- **泊松定理**
  - (用泊松分布来逼近二项分布的定理)
- 设$\lambda>0$是一个常数，n是任意正整数，设$np_n=\lambda$，则对于任意固定的非负整数k，有

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

当试验次数n很大时，稀有事件A发生的次数可以近似用泊松分布来描述，而$\lambda=np$为n次中A发生的平均次数

- 证 由$p_n=\lambda/n$，有
- $$\binom{n}{k}p_n^k(1-p_n)^{n-k}=\frac{n(n-1)\cdots(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$$
- $$=\frac{\lambda^k}{k!}[1\cdot(1-\frac{1}{n})\cdots(1-\frac{k-1}{n})]\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$
  - 对任意固定的k，当n趋近于无穷大时，上式第二项极限为1，第三项是自然对数极限为$e^{-\lambda}$，最后一项极限为1，所以定理成立。
  - $np_n=\lambda$是常数，当n很大时p_n必定很小，所以当n很大，p_n很小时，可用泊松分布来近似二项分布。

§ 2.2 离散型随机变量及其概率分布

- 历史上Poisson分布是作为二项分布的一种近似，而二项分布的真正极限分布是正态分布(见中心极限定理)
- 于1837年由法国数学家S.D.Poisson（1781 ～ 1840年）引入。
- 近些年来，人们发现该分布在物理学及社会生活中对服务的各种要求等方面有愈来愈多的应用

§ 2.2 离散型随机变量及其概率分布

- **(四) 超几何分布，抽样检查，无放回抽样**
  - 产品抽样检查中，假定在N件产品中有D件不合格品，即不合格率p=D/N。随机抽n件做检查，发现k件不合格品的概率为超几何分布

$$P(X=k)=C^{k}_{D}C^{n-k}_{N-D}/C^{n}_{N},(k=0,1,\ldots \min\{n, D\})$$

  - 通常称这个随机变量X服从超几何分布。这种抽样检查方法等于无放回抽样。
  - 数学上不难证明，N趋近无穷时
  - $P(X=k)=C^{k}_{D}C^{n-k}_{N-D}/C^{n}_{N}$ 近似为$b(n,p)$ (二项分布)
  - 因此，在实际应用时，当$N\geq 10n$时，可用二项分布近似描述不合格品个数，当实验次数足够多的时候，不放回抽样可近似为放回抽样，而如果放回抽样则刚好满足二项分布，p=废品率D/N

§ 2.2 离散型随机变量及其概率分布

- **证明**

$$\frac{C_D^k C_{N-D}^{n-k}}{C_N^n} = \frac{D!}{k!\,(D-k)!} \times \frac{(N-D)!}{(n-k)!\,\big(N-D-(n-k)\big)!} \times \frac{n!\,(N-n)!}{N!}$$

$$= C_n^k \frac{D!}{(D-k)!} \times \frac{(N-D)!}{\big(N-D-(n-k)\big)!} \times \frac{(N-n)!}{N!}$$

$$= C_n^k \frac{D(D-1)\cdots(D-k+1)}{N^k}$$

$$\times \frac{(N-D)(N-D-1)\cdots(N-D-(n-k)-1)}{N^{n-k}} \times \frac{N^n}{N(N-1)\cdots(N-n+1)}$$

- **由p=D/N,知**

$$\lim_{N\to\infty} \frac{D(D-1)\cdots(D-k+1)}{N^k} = p\left(p-\frac{1}{N}\right)\cdots\left(p-\frac{k-1}{N}\right) = p^k$$

$$\lim_{N\to\infty} \frac{(N-D)(N-D-1)\cdots(N-D-(n-k)-1)}{N^{n-k}} = (1-p)^{n-k}$$

$$\lim_{N\to\infty} \frac{N^n}{N(N-1)\cdots(N-n+1)} = 1$$

- **带回可得。**

第二章 随机变量及其分布

- § 2.1 随机变量
- § 2.2 离散型随机变量及其概率分布
- § 2.3 随机变量的分布函数
- § 2.4 连续型随机变量及其概率密度
- § 2.5 随机变量的函数的分布

§ 2.3 随机变量的分布函数

- 非离散型随机变量X，由于其可能取值不能一个一个的列举而无法用分布律来描述。而且所遇到的非离散型随机变量通常取任一指定的实数值的概率都等于0
  - 实际问题中有很多此类变量：误差ε，元件寿命T等.
- 对于此类问题，主要研究随机变量取值落在某一个区间的概率，即：
  - $P\{x_1<X\le x_2\}= P\{X\le x_2\}-P\{X\le x_1\}$ 物理意义
  - 问题变为计算$P\{X\le x_2\}$及$P\{X\le x_1\}$，于是引入分布函数的概念。

§2.3 随机变量的分布函数

- **分布函数定义：**

  设X是一个随机变量，x是任意实数，函数

  $F(x)=P\{X\leq x\}$，$-\infty<x<\infty$

  称为X的分布函数

- **左开右闭区间上的概率表示：**
  - 对于任意实数x_1，x_2 $(x_1<x_2)$，有
  - $P\{x_1<X\leq x_2\}= P\{ X\leq x_2\}-P\{X\leq x_1\}=F(x_2)-F(x_1)$

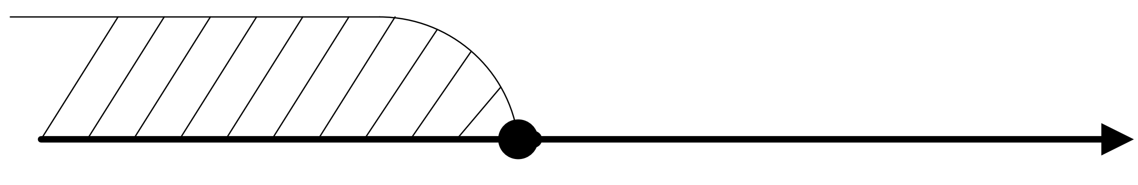
∴已知分布函数，就可以计算出x落在任意区间$(x_1，x_2]$上的概率

- 分布函数完整地描述了随机变量的统计规律性。

§2.3 随机变量的分布函数

➲ **分布函数的物理意义：**

- **将X看成是数轴上的随机点的坐标，那么，分布函数$F(x)$在x处的函数值就表示X落在$(-\infty, x]$上的概率。**



§ 2.3 随机变量的分布函数

- **分布函数具有以下基本性质**

1° $F(x)$是一个不减函数

- **对任意实数x_1，x_2 ($x_1<x_2$)，有$F(x_2)-F(x_1)=P\{x_1<X\leq x_2\}\geq 0$**

2° $0\leq F(x)\leq 1$且$F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0$，$F(\infty)=\lim\limits_{x\to\infty}F(x)=1$

- **仅从物理意义上加以说明：当x趋于$-\infty$时，即区间端点沿x轴无限左移，{X<x}逐渐趋于不可能事件，从而其概率趋于0，即有$F(-\infty)=0$；反之，若点x无限右移，则{X<x}趋于必然事件，从而概率趋于1，即有$F(\infty)=1$**

3° $F(x+0)=F(x)$ 即$F(x)$为右连续的

- **$F(x)$在点x处的右极限等于点x处的函数值，即$F(x)$是右连续的**

§ 2.3 随机变量的分布函数

例1：随机变量的分布律用分布函数来表示

- 设随机变量X的分布律为

X	-1	2	3
p_k	1/4	1/2	1/4

- 求X的分布函数，并求$P\{X\le 1/2\}$，$P\{3/2<X\le 5/2\}$，$P\{2\le X\le 3\}$

➲ 解：随机变量落在−1，2，3三个点上的概率为非0，落在其它点上为不可能事件，因此有

$$\begin{cases} P\{X<-1\}=0 \\ P\{X=-1\}=1/4 \\ P\{-1<X<2\}=0 \\ P\{X=2\}=1/2 \\ P\{2<X<3\}=0 \\ P\{X=3\}=1/4 \\ P\{X>3\}=0 \end{cases} \xRightarrow{\text{由概率的有限可加性}} F(x)=\begin{cases} 0, & x<-1 \\ P\{X=-1\}, & -1\le x<2 \\ P\{X=-1\}+P\{X=2\}, & 2\le x<3 \\ 1, & x\ge 3 \end{cases}$$

$$F(x)=\begin{cases} 0, & x<-1 \\ 1/4, & -1\le x<2 \\ 1/4+1/2 & ,2\le x<3 \\ 1, & x\ge 3 \end{cases}$$

§ 2.3 随机变量的分布函数

- $P\{X \le 1/2\} =$
- $P\{3/2 < X \le 5/2\} =$
- $P\{2 \le X \le 3\} =$

$$F(x)=\begin{cases} 0, & x<-1 \\ 1/4, & -1\le x<2 \\ 1/4+1/2 & ,2\le x<3 \\ 1, & x\ge 3 \end{cases}$$

- **注意区间左端的等号**

1° 概率函数与普通的代数函数不同之处在于物理意义不同，前者是随机变量X的取值落在(－∞，*x*]区间上的概率

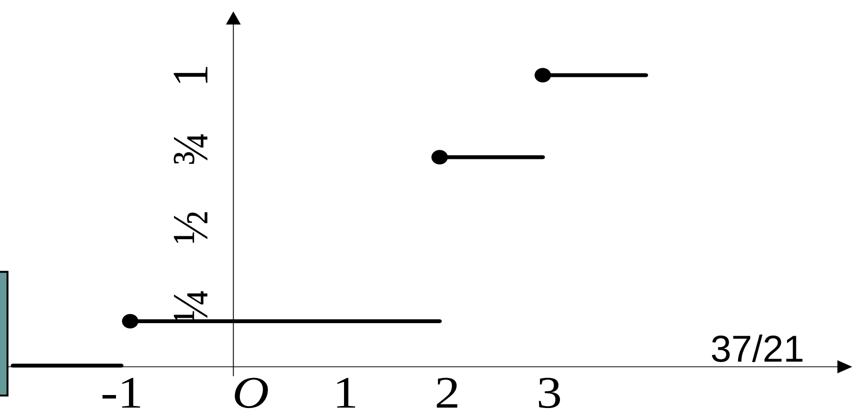
2° 求分布函数，一定要讨论区间(－∞，∞)上的所有情况

3° $F(x)$的值即为所有≤x的X取值中落在x_k处的概率p_k之和

4° $F(x)$的图形是一条阶梯形曲线，

- **在x＝－1，2，3处有阶跃点**
- **阶跃值分别为1/4，1/2，1/4**

曲线的阶跃点处，上为实心点，下为空心点，体现右连续,小于-1时F(x)=0也应画一个粗实线



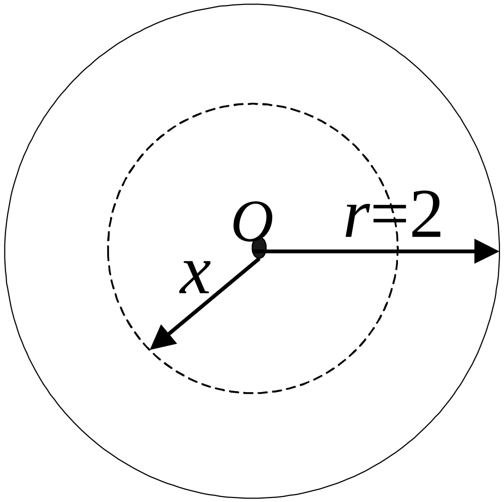
§ 2.3 随机变量的分布函数

- 分布律用分布函数表示的一般方法：
  - 一般的，设离散型随机变量X的分布律为
  - $P\{X=x_k\}=p_k$，$k=1,2,\ldots$
  - 由概率的可列可加性得X的分布函数为
  - $F(x)=P\{X\le x\}=\sum_{x_k\le x}P\{X=x_k\}$
  - 即$F(x)=\sum_{x_k\le x}p_k$
  - 这里和式是对所有满足$x_k\le x$的k求和的，分布函数在点$x=x_k$（$k=1$，2，…）处有跳跃，其跳跃值为$p_k=P\{X=x_k\}$

分布函数能表达所有随机变量的概率问题

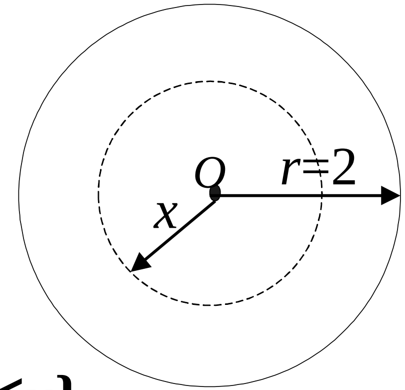
§ 2.3 随机变量的分布函数

- **例2 一般随机变量的分布函数的求解方法**
  - **一个靶子是半径2米的圆盘，设击中靶上任一同心圆盘上的点的概率与该同心圆盘的面积成正比，并设射击都能中靶，以X表示弹着点与圆心的距离，试求随机变量X的分布函数$F(x)=P\{X\leq x\}$**



§ 2.3 随机变量的分布函数

- **解：分段讨论：**
  - 1° 当$x<0$时，$\{X\le x\}$是不可能事件，$F(x)=0$
  - 2° 在圆盘上，$0\le x\le 2$，由题意，$P\{0\le X\le x\}=k\pi x^2$
  - 当x＝2时由于假定每次射击都能中靶，
  - 有$P\{0\le X\le 2\}=1$
  - $\therefore 1=k\pi 2^2$
  - 即$k\pi=1/4$代入得$F(x)=P\{X<0\}+P\{0\le X\le x\}$
  - $=0+(1/4)x^2=(1/4)x^2$
  - 3° 当$x>2$时，由题意$\{X\le x\}$是必然事件，$F(x)=1$

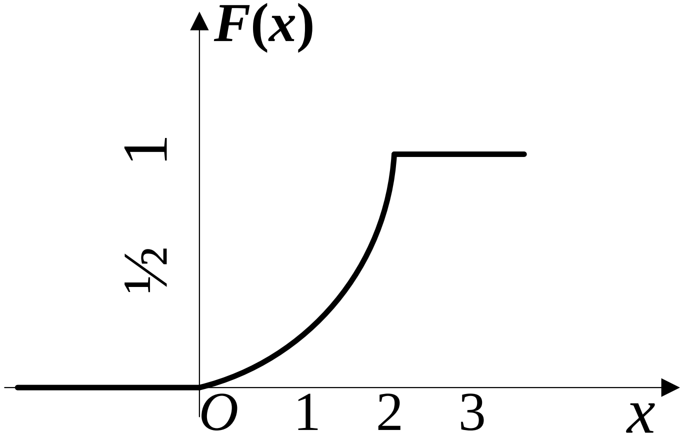


§ 2.3 随机变量的分布函数

- 综上，$F(x)$是一个分段函数，具体为

$$F(x)=\begin{cases} 0, & x<0 \\ \dfrac{x^2}{4}, & 0\le x<2 \\ 1, & x\ge 2 \end{cases}$$

- 它的图形是一条连续曲线

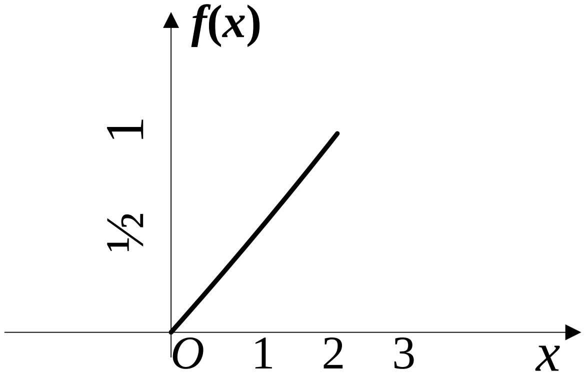


- ※注意区间端点的取值
- ※注意对x的讨论一定要全，在（$-\infty$，∞）上都要考虑
- ※注意x是值，X是随机变量

§ 2.3 随机变量的分布函数

➲ **现在对$F(x)$求导数**

- **（不可导的点的导数设为0）**
- **有** $$\frac{dF(x)}{dx}=f(x)=\begin{cases}\frac{1}{2}x, & 0<x<2\\ 0, & \text{其它}\end{cases}$$



- **这样有**$F(x)=\int_{-\infty}^{x} f(t)dt$
- **即$F(x)$是非负函数$f(t)$在区间$(-\infty，x]$上的积分，$f(t)$即为概率密度函数，其中X是连续型。**

本章作业

- P_{55}：2，4，5，8
- P_{56}：10，12，13，17，21，24，26，29