本次课的重点(两个函数)：
理解什么是随机变量.掌握分布函数的定义及判断性质，离散型随机变量的分布律及分布函数.

下次课内容：
要讲到第二章的 § 2.2 离散型随机变量的中的泊松分布，包括离散随机变量的分布律定义，二项分布及泊松分布.

随机变量

(1)抽象**函数**：

$$\Omega \ni \omega \xrightarrow{\text{单值实函数}} X(\omega) \in R^n$$

(2)随机：“其值随机会而定”.

试验前，不能预知将取何值，这要凭机会，一旦试验后，取值就确定.

思考：在实际中如何使用随机变量?

随机试验的语言进行直观表述:

某次试验结果的<u>数值性描述</u>, 称为随机变量(简化定义).

思考：随机变量与普通函数的区别?

随机变量： $X(\omega)$: $\Omega \longrightarrow R^n$ 是试验结果的抽象函数.

对任一实数 $x \rightarrow P\{\ \omega |\ X(\omega) \le x\ \}$ 存在

分布函数： $F(x)=P(X\le x)$

$F(x)$: $R \longrightarrow [0,1]$ 为一普通实函数.

注: (1) $F(x)$ 实质就是随机事件 $\{\ \omega:\ X(\omega) \le x\ \}$ 的概率；

$F(x)$ 值表示事件"随机点 X 落在 $(-\infty, x\ ]$ 内"的概率.

(2) 分布函数的三条性质：

单调不降，极限，右连续(不证明)

这三个性质可以判断一个一元函数是否为一个分布函数；

(任何类型的分布函数都要满足这三个性质)

常用的结论: $P\{a < X\le b\} = P\{X\le b\} - P\{X\le a\} = F(b) - F(a)$

研究一个随机变量 (random variable) $\begin{cases} \text{(1)首先关注取值情况, 即取哪些值.} \\ \text{(2)更关注它取各种值的概率情况, 即“概率的分布”.} \end{cases}$

E1: 从某厂大批产品 (M) 中随机地抽出100个, 其中所含废品数?

第一章核心模型: 抽球模型得出的超几何分布

有放回地抽取, 其中所含废品数?

第二章核心模型: 伯努利概型得出的二项分布

伯努利概型的应用: “在同样条件下进行重复试验”的一种数学模型, 特别在讨论某事件出现的频率时常用这种模型.

两种分布的关系: 不放回抽出: M很大, 其中所含废品数?

$p_i = P\{X = x_i\}$, 且满足

$$\begin{cases} (1)\ p_i \geqslant 0, \\ (2)\ \sum_{i=1}^{\infty} p_i = 1. \text{(“归一性”)} \end{cases}$$

离散型随机变量的标志: 分布律或阶梯型分布函数.

离散型随机变量的分布函数特点：

(1) 右连续的单调不降阶梯函数;(离散型随机变量的分布函数的标志)

(2) 在不连续点处的跳跃高度恰好为分布律p_i, i=1,2,…

离散型随机变量X 分布律与分布函数的关系

1.已知分布律求它的分布函数

分布函数$$F(x)=P\{X\le x\}=\sum_{x_i\le x}P\{X=x_i\}=\sum_{x_i\le x}p_i$$

2.已知分布函数求分布律

$$P\{X=x\}=F(x)-F(x-0),\forall x\in(-\infty,+\infty)$$

或在$F(x)$不连续点处的跳跃高度恰好为分布律P_i, i=1,2,…

x_i, i=1,2,…,($x_i< x_{i+1}$)为$F(x)$的不连续点

$$P\{X\le x_{i+1}\}=P\{(X\le x_i)\bigcup(X=x_{i+1})\}$$

$$\Rightarrow P\{X=x_{i+1}\}=F(x_{i+1})-F(x_i)$$

离散型随机变量X分布律与分布函数的关系

1.已知分布律求它的分布函数

分布函数$F(x)=P\{X\le x\}=\sum\limits_{x_i\le x} P\{X=x_i\}=\sum\limits_{x_i\le x} p_i$

注: (1) 离散型随机变量的分布函数是右连续的单调不降阶梯函数(离散型随机变量的分布函数的标志);

(2) 在不连续点处的跳跃高度恰好为分布律p_i, $i=1,2,\ldots$

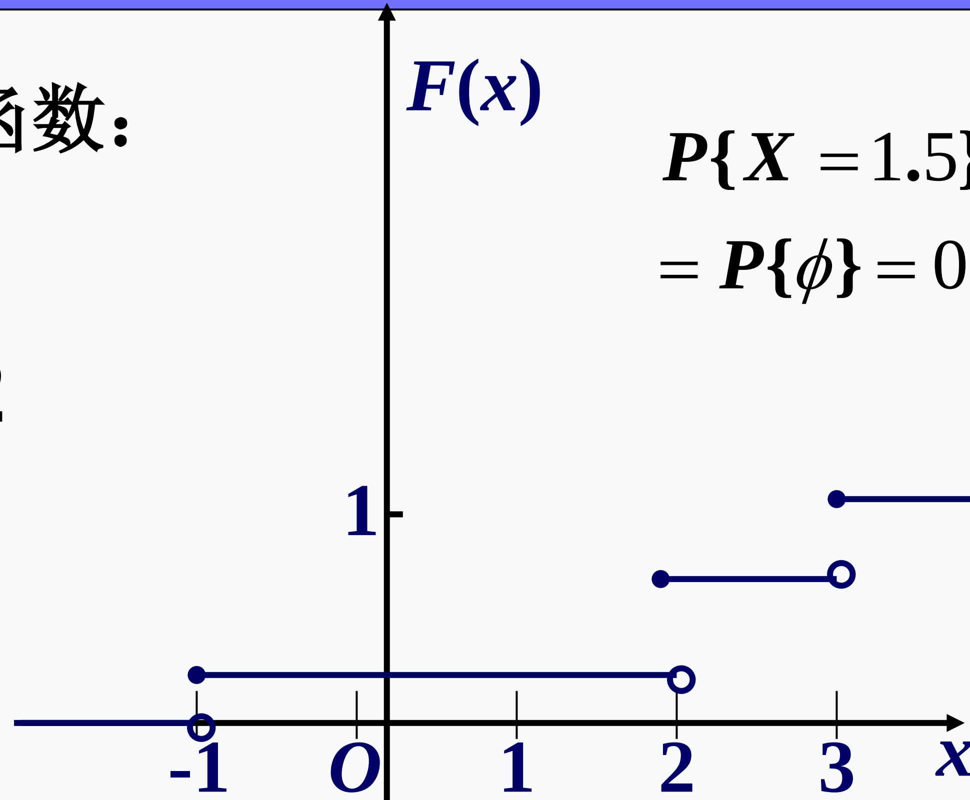
注:区别离散型随机变量的分布律与分布函数

分布函数: $\forall x\in R,\ F(x)=P\{X\le x\}$(分段形式函数)

分布律: $p_i=P\{X=x_i\}$, $i=1,2,\ldots$为X的分布律.

已知随机变量X的分布函数：

$$F(x)=\begin{cases} 0 & x<-1 \\ \dfrac{1}{6} & -1\le x<2 \\ \dfrac{2}{3} & 2\le x<3 \\ 1 & x\ge 3 \end{cases}$$



$$P\{X=1.5\}$$
$$=P\{\phi\}=0$$

求: (1) $p_i=P\{X=x_i\}$, i=1,2,… ;(2)$P\{X=1.5\}$=?

$$P\{X=-1\}=\frac{1}{6}-0,\qquad P\{X=2\}=\frac{2}{3}-\frac{1}{6}=\frac{1}{2}$$

$$P\{X=3\}=1-\frac{2}{3}=\frac{1}{3}\qquad F(1.5)=P\{X\le 1.5\}=P\{X=-1\}=\frac{1}{6}$$

证明:$(1)P\{X=x\}=F(x)-F(x-0),\forall x\in(-\infty,+\infty)$

$(2)P\{X<x\}=F(x-0),\forall x\in(-\infty,+\infty)$

证明:$(1)P\{X=x\}=\lim\limits_{\Delta x\to 0+}P\{x-\Delta x<X\leq x\}$

$$=\lim_{\Delta x\to 0+}[P\{X\leq x\}-P\{X\leq x-\Delta x\}]$$

$$=F(x)-F(x-0),\forall x\in(-\infty,+\infty)$$

$(2)P\{X<x\}=\lim\limits_{\Delta x\to 0+}P\{X\leq x-\Delta x\}$

$$=\lim_{\Delta x\to 0+}F(x-\Delta x)=F(x-0),$$

$$\forall x\in(-\infty,+\infty)$$

$(3)P\{X\geq x\}=?;(4)P\{X>x\}=?$

3) 概率的连续性

1)若$A_1 \supset A_2 \supset \cdots$,且$\bigcap\limits_{n=1}^{\infty} A_n = A(\lim\limits_{n\to\infty} A_n = A)$,则

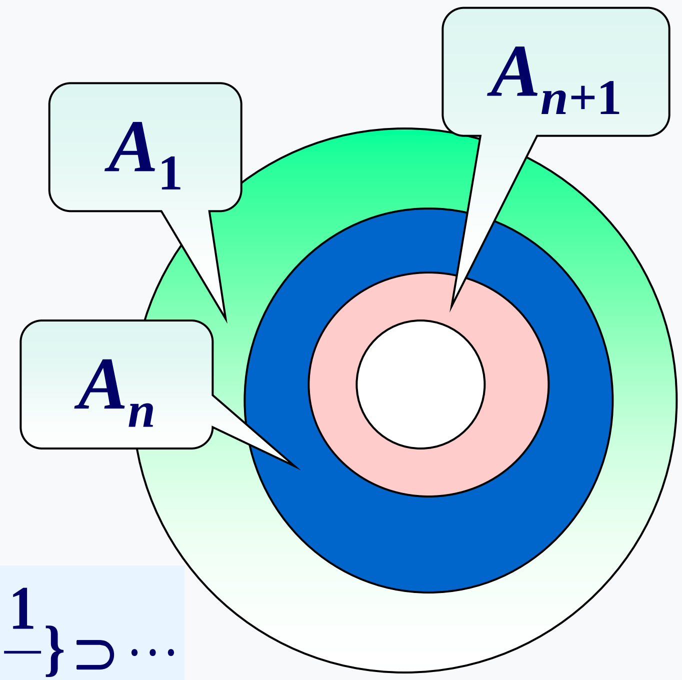
$$\lim_{n\to\infty} P(A_n) = P(A).$$

应用到分布函数:

$\forall x \in R, F(x+0) = F(x)\quad (F(x) \overset{\Delta}{=} P\{X \le x\})$

$$\lim_{n\to\infty} F\left(x+\frac{1}{n}\right) = F(x+(= F(x)$$

且$\{X \le x+1\} \supset \{X \le x+\frac{1}{2}\} \supset \cdots \supset \{X \le x+\frac{1}{n}\} \supset \cdots$



2)若$A_1 \subset A_2 \subset \cdots$,且$\bigcup\limits_{n=1}^{\infty} A_n = A(\lim\limits_{n\to\infty} A_n = A)$,则

$$\lim_{n\to\infty} P(A_n) = P(A).$$

二：伯努利试验和二项分布

E1：抛一枚硬币出现正反面.

E2：检查一件产品是否合格.

E3：股票市场中关心是涨还是跌.

E4：考一门课，是否通过.

特点：关注试验的两个结果：A和$\overline{A}$

注意实际结果可能不止两个

称这类试验为伯努利试验.

伯努利试验是一种非常重要的概率模型，它是“在同样条件下进行重复试验”的一种数学模型,特别在讨论某事件出现的频率时常用这种模型.

例如:在工业产品质量检查中,在当代遗传学中,机票超售等问题.

概率论：研究随机现象的数量规律的数学分支。

机票超售：航空公司电脑订座系统的普遍采用给旅客和公司都带来极大的方便，但是也对管理工作提出更高的要求，例如一架200座的飞机应出售多少座位?

简单而常用的方法是限定出售200座。不过，这并不是一个很好的答案，因为常有订了座位的旅客临时不来上机，出现空位，造成浪费。于是就实行超售，即在飞机起飞前出售的座位超过实有的座位。

据统计，国内航班中订座而到时不来上机的旅客超过5%。因此，超售是正确的选择，但是超售会造成拒登机。超售问题是很典型的概率论问题：假定每个订座旅客准时上机的可能性95%，发生拒登机的可能性P：

N	201	202	203	204	205	206	207
P	0.000	0.002	0.007	0.015	0.032	0.062	0.109

伯努利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?

$$\therefore F(x)=\begin{cases}0, & x<0\\ 1-p, & 0\le x<1\\ 1, & x\ge 1\end{cases}$$

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.

则称这n次试验为n次重复独立试验.

当试验E 是伯努利试验，称这n次独立试验为n重伯努利试验，或称n重伯努利概型.

伯努利概型的重要性

概率论是一门研究随机现象数量规律的学科. 它起源于17世纪中叶;

其后, 在对伯努利概型的深入研究中, 发现了两种形式的极限定理——大数定律和中心极限定理(第五章), 奠定了概率论在数学中的理论地位….

对于一个伯努利试验，我们常考察如下问题：

(1) 事件A 首次发生时的试验次数Z；

(2) 事件A 发生k 次时的试验次数Y；

(3) n 次试验中事件A 发生的总次数X.

(1)无限重复伯努利试验：

$$P\{Z=k\}=p(1-p)^{k-1}, k=1,2,\ldots,$$ 几何分布或首次分布。

n重伯努利试验： $A_i=\{第i次事件发生\}, \ i=1,\ldots,n$

射击气球(略)

$$\{Z=k\}=\overline{A_1}\cap\overline{A_2}\cap\cdots\cap\overline{A_{k-1}}\cap A_k, \ k=1,\cdots,n-1$$

$$\{Z=n\}=\overline{A_1}\,\overline{A_2}\cdots\overline{A_{n-1}}A_n\cup\overline{A_1}\,\overline{A_2}\cdots\overline{A_n}$$

$$P\{Z=k\}=p(1-p)^{k-1} \quad k=1,\cdots,n-1 \qquad P\{Z=n\}=(1-p)^{n-1}$$

射击气球

例4:某人有3发子弹，他射击空中汽球的命中率为0.9, 命中则停止射击, 他用去的子弹个数X的分布律为?

解：A={射击命中}　X就表示：A首次发生时试验次数

A_i={第i次射击命中}，i=1, 2, 3

$$P\{X=1\}=P(A_1)=0.9$$

$$P\{X=2\}=P(\overline{A_1}A_2)=0.1\cdot 0.9=0.09$$

$$P\{X=3\}=P(\overline{A_1}\,\overline{A_2}A_3\cup\overline{A_1}\,\overline{A_2}\,\overline{A_3})$$

$$=0.1^2 0.9+0.1^3=0.1^2$$

由分布律“归一性”：$P\{X=3\}=1-0.9-0.09=0.01$

1) 设事件A 首次发生的试验次数为 ξ,

则 $\{\xi=k\}$ 表示首次试验成功在第k次。$k=1,2,..,n$

ξ 的分布律为：$P\{\xi=k\}=pq^{k-1}$； $(q=1-p)$

称ξ服从参数为p的几何分布。

几何分布的一个重要性质：无后效性（无记忆性）

即$P\{\xi=n+m|\xi>n\}=P\{\xi=m\}$成立。

证明：$P\{\xi=n+m|\xi>n\}=\frac{1}{P\{\xi>n\}}*P\{\xi=n+m,\xi>n\}$

$$=\frac{1}{\sum_{k=n+1}^{+\infty}pq^{k-1}}*P\{\xi=n+m\}=\frac{1}{q^n}*pq^{n+m-1}$$

$$=pq^{m-1}$$

思考：证明几何分布的无记忆性。

对于一个伯努利试验，我们常考察如下问题：

(1) 事件A 首次发生时的试验次数Z；

$P\{Z=k\}=p(1-p)^{k-1}, k=1,2,\ldots,$几何分布或首次分布.

(2) 事件A 发生k 次时的试验次数Y；

(3) n 次试验中事件A 发生的总次数X.

(2)无限重复伯努利试验：

$$P\{Y=t\}=C_{t-1}^{k-1}p^k q^{t-k}, \qquad t=k,k+1,\ldots.$$

称Y 服从负二项分布(Negative binomial)(或帕斯卡(Pascal)分布)

负二项分布可看作几何分布的更一般情况.

产品检查试验(回顾): 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.

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.

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

$$P\{X=k\}=C_{k-1}^{5-1}p^5(1-p)^{k-5},\ (k=5,6,\ldots)$$

则X服从负二项分布(或帕斯卡(Pascal)分布)

对于一个伯努利试验，我们常考察如下问题：

(1) 事件A 首次发生时的试验次数Z;

$P\{Z=k\}=p(1-p)^{k-1}, k=1,2,\ldots,$ 几何分布或首次分布.

(2) 事件A 发生k 次时的试验次数Y;

(3) n 次试验中事件A 发生的总次数X.

(2)无限重复贝努力试验：

$$P\{Y=t\}=C_{t-1}^{k-1}p^{k}q^{t-k}, \qquad t=k,k+1,\ldots.$$

称Y 服从负二项分布(或帕斯卡(Pascal)分布)

负二项分布可看作几何分布的更一般情况。

(3) n 次试验中事件A 发生的总次数X.

定理： 在n重伯努利试验中，事件A发生概率为$P(A)=p$，$0<p<1$，则事件A发生的总次数X的分布律为

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

证： 在n重伯努利试验中，事件A在指定的k次试验中出现的概率为$p^k(1-p)^{n-k}$.

在n次试验中，选出k次试验来有C_n^k种不同的方式.

且各种方式的事件互不相容，由概率的有限可加性可得

$$P_n(k)=C_n^k\,p^k(1-p)^{n-k}$$

所以结论成立.

定理： 在n重伯努利试验中，事件A发生概率为$P(A)=p$，$0<p<1$，则事件A发生的总次数X的分布律为

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

名称: 由于$C_n^k p^k(1-p)^{n-k}$ 正是 $[p+(1-p)]^n$ 展开式中的通项公式,

称随机变量X 服从二项分布(Binomial distribution),
记为$X\sim B(n,p)$.

特别地, (0—1)分布, 可以看作$X\sim B(1,p)$.

注: 随机变量X服从二项分布要满足两个条件(判定条件)

(1) 满足n次贝努里试验条件：独立和重复性;

(2) 求事件A发生的总次数X.

但在现实中也可作近似处理,不一定豪厘不差.

对于一个伯努利试验(关注每次试验两个结果)：

(1) 事件A 首次发生时的试验次数Z;

$$P\{Z=k\}=p(1-p)^{k-1}, k=1,2,\ldots,$$几何分布或首次分布.

n次伯努利试验： $A_i=\{$第i次事件发生$\}$，$i=1,\ldots,n$

$$\{Z=k\}=\overline{A_1}\cap\overline{A_2}\cap\cdots\cap\overline{A_{k-1}}\cap A_k,\quad P\{Z=k\}=p(1-p)^{k-1}$$

$$\{Z=n\}=\overline{A_1}\,\overline{A_2}\cdots\overline{A_{n-1}}A_n\cup\overline{A_1}\,\overline{A_2}\cdots\overline{A_n}\quad P\{Z=n\}=(1-p)^{n-1}$$

反其道而行之

(2) 事件A 发生k 次时的试验次数Y; 记$Y\sim NB(k;p)$

$$P\{Y=t\}=C_{t-1}^{k-1}p^kq^{t-k},\qquad t=k,k+1,\ldots.$$

称Y 服从负二项分布(或帕斯卡(Pascal)分布)

(3) n 次试验中事件A 发生的总次数X, 服从二项分布.

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

记$X\sim B(n,p)$

(2) 事件A 发生k 次时的试验次数Y; $P\{Y=t\}=C_{t-1}^{k-1}p^kq^{t-k}, t=k,k+1,\ldots$

为了检查某厂产品的废品率p大小,有两个试验方案可采取:

一是从该厂产品中抽出若干个,检查其中的废品数X,这一方案导致二项分布,已于前述.

另一个方案是先指定一个自然数k.一个个地从该厂产品中抽样检查, 直到发现第k个废品为止.以W记到当时为止已检出的合格品个数.

需要的条件?

$$(1-x)^{-r}=\sum_{i=0}^{\infty}C_{-r}^{i}(-x)^i=\sum_{i=0}^{\infty}C_{i+r-1}^{i}x^i=\sum_{i=0}^{\infty}C_{i+r-1}^{r-1}x^i$$

$$P\{W=w\}=C_{w+k-1}^{k-1}p^k(1-p)^w,\ w=0,1,2,\ldots.$$

令$x=1-p$,并两边乘以p^r,得$1=p^r[1-(1-p)]^{-r}=\sum_{i=0}^{\infty}C_{i+r-1}^{r-1}p^r(1-p)^i$

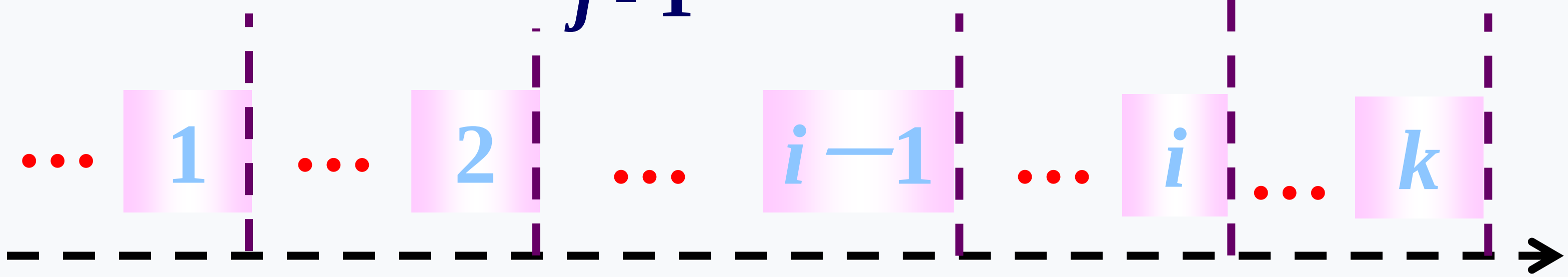
注:由于例中所描述的试验方式,它与二项分布比是“反其道而行之”:二项分布是定下总抽样个数n而把废品个数X作为变量;负二项分布则相反,它定下废品个数r而把总抽样次数减去r作为变量.

负二项分布的名称来由:由于“负指数二项展开式”.

例：向某一目标进行射击, 直至命中k次为止, 已知命中率为$p>0$. 求射击次数X 的分布律.

分析：X的分布律为： 负二项分布?

$$P\{X=j\}=C_{j-1}^{k-1}p^{k}(1-p)^{j-k} \quad j=k,k+1,\ldots$$



X_i表示第i—1次命中以后, 到第 i 次命中的射击次数.

X_i的分布律为:

$$X=X_1+X_2+\ldots+X_k$$

X_i	1	2	…..	m	……
$P\{X_i=m\}$	p	$(1-p)p$	…..	$(1-p)^{m-1}p$	……

产品抽检试验

例5：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 $$X \sim B(n, \frac{M}{N}).$$

所以, X的分布律为 $$P\{X = k\} = C_n^k (\frac{M}{N})^k (1 - \frac{M}{N})^{n-k},$$

其中 $k = 0,1,2,\dots,n.$

思考: 将抽取方式改为无放回抽取, 试写出X的分布律.

若N很大，X的分布律?N很大, M很小的情况？

例(超售问题): 航空公司电脑订座系统的普遍采用给旅客和公司都带来极大的方便，但是也对管理工作提出更高的要求, 例如一架200座的飞机应出售多少座位?

每个订座旅客当作一次试验，则他或到时不登机，记为A，或到时登机, 记为 $\bar{A}$, P(A)按过去统计资料取为5%。

主要的难点在于能否把旅客是否登机看作是独立的，显然对于购买团体票的旅客作此假定是不合适的，此外，大型的交通堵塞等偶然事件也会使这个假定偏离，不过在一般场合作此假定还是合适的. 全体订座旅客数*n*作为试验总数, 这便构成伯努利概型.

假定超售*m*个座位, 则共售出200+*m*个座位, 这时要求登机的旅客数X服从二项分布B(200+*m*, 0.95), 发生拒登机的可能性：$P=P\{X>200\}$

N	201	202	203	204	205	206	207
P	0.000	0.002	0.007	0.015	0.032	0.062	0.109

强弱对抗试验

例7：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：设双方逐对较量独立进行，故为独立重复试验.

设$A = \{$弱队每个队员获胜$\}$，弱队获胜的人数为X

$C = \{$弱队获胜$\}$

（1）当双方各出3人时，$X \sim B(3, 0.4)$.

所以 $P(C) = P\{X \geq 2\}$

$$= \sum_{k=2}^{3} C_3^k (0.4)^k (0.6)^{3-k} \approx 0.352$$

例7：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）当双方各出3人时， $X \sim B(3, 0.4)$.

$$P(C) = P\{X \geq 2\} = \sum_{k=2}^{3} C_3^k (0.4)^k (0.6)^{3-k} \approx 0.352$$

（2）当双方各出7人时， $X \sim B(7, 0.4)$.

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

$$\approx 0.290$$

故我们得到: 第一种方案对弱队更有利一些.

设备排障试验

例9：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

若 配N 个工人，应使$\{X > N\}$

A={一台机床发生故障}，$X \sim B(300, 0.01)$，则

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

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

即求使上述不等式成立的最小N 值.

例子 ⟶

产品抽检试验

设备排障试验

强弱对抗试验　人寿保险

注意: 在伯努利试验中出现的随机变量不是都服从二项分布.

对于一个伯努利试验,可以考察如下问题:

(1) 事件A 首次发生时的试验次数Z;

(2) 事件A 发生k 次时的试验次数Y;

(3) n 次试验中事件A 发生的总次数X.

泊松(Poisson, Simeon-Denis)(1781—1840)法国数学家. 1781 年6月21日生于法国卢瓦雷省的皮蒂维耶, 1840年4月25日卒于法国索镇. 泊松是法国数学家,物理学家和力学家.

1798年入巴黎综合工科学校深造. 在毕业时，因优秀的研究论文而被指定为讲师. 受到P.-S.拉普拉斯、J.-L.拉格朗日的赏识. 1800年毕业后留校任教, 1802年任副教授, 1806年接替J.-B.-J.傅里叶任该校教授. 1808年任法国经度局天文学家, 1809年任巴黎理学院力学教授. 1812年当选为巴黎科学院院士.

泊松的科学生涯开始于研究微分方程及其在摆的运动和声学理论中的应用. 他工作的特色是应用数学方法研究各类力学和物理问题, 并由此得到数学上的发现.他发表过300多篇论文，所著两卷《力学教程》在很长的时期内被认为是标准的教科书. 泊松还将数学应用于物理学，涉及电，磁，热，声，光等许多方面.

泊松在数学方面贡献很多. 最突出的是1837年在《关于刑事案件和民事案件审判概率的研究 》中提出描述随机现象的一种概率分布—泊松分布. 这一分布在公用事业、放射性现象等许多方面都有应用. 他还研究过定积分、傅里叶级数、数学物理方程等. 除泊松分布外, 还有许多数学名词是以他名字命名的, 如泊松积分、泊松求和公式、泊松方程、泊松定理等等.

法国力学家、物理学家和数学家S.D.泊松(1781-1840)

考虑下列现象：

1. 每小时服务台访客的人数;
2. 每天家中电话的通数;
3. 一本书中每页的错字数;
4. 某条道路上每月发生车祸的次数;
5. 生产线上的瑕品数，
6. 学生到办公室找老师的次数……
7. 放射性分裂落到某区域的质点数;
8. 热电子的发射;
9. 显微镜下落在某区域中的血球或微生物的数目等等.

大致上都有一些共同的特征:在某时间区段内,平均会发生若干次「事件」,但是有时候很少,有时又异常地多,因此事件发生的次数是一个随机变数，它所对应的概率函数称为 **Poisson分布**.

试验: 某电话总台一天接到的呼叫次数, 网站访问数, 公共汽车站来到的乘客数.

三、泊松分布(Poisson)(1837年)

若随机变量 X 的分布律为 $P\{X=k\}=\dfrac{\lambda^k}{k!}e^{-\lambda}$, $k=0,1,2,\ldots,$

$\lambda>0$, 则称随机变量 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:

它多是出现在当 X 表示在一定的时间或空间内出现的事件个数这种场合:主要在社会生活, 及物理科学.

(2) 泊松分布可视为二项分布的近似分布.

许多应用问题中，常遇到这样的伯努利试验, n很大, p很小, 而np大小适中. 泊松找到一个便于使用的近似公式:

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

即$P\{X_n = k\} = C_n^k (p_n)^k (1 - p_n)^{n-k}$, $k = 0, 2, \ldots, n$

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

$$\lim_{n\to\infty} P\{X_n = k\} = \frac{\lambda^k}{k!} e^{-\lambda}$$

证明: 参见教材(略).

注：也称Poisson逼近定理. 这个定理告诉我们可将较难计算的二项分布转化为泊松分布去计算.

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

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

(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, 如不幸事件, 意外事故、故障等.

宇宙粒子

发病率试验

产品销售试验

设备排障试验

例4：有300台独立运转的同类机床，每台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, $X \sim B(300, 0.01)$.

配N 个工人, 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.

因为300×0.01=3(n较大, p较小),

故由*Poisson*逼近定理可知:

认为X近似服从 $\lambda=3$ 的泊松分布, 即 $X \sim P(3)$.

例4：有300台独立运转的同类机床，每台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.

认为X近似服从 $\lambda=3$ 的泊松分布，即 $X \sim P(3)$.

于是 $0.01 > P\{X > N\} = 1 - F(N+1-1)$

$$\approx \sum_{k=N+1}^{+\infty} \frac{3^k}{k!} e^{-3}$$

查附表1 可得

$P\{X > 7\} = 1-F(7) = 1-F(8-1) = 0.011905 > 0.01$

$P\{X > 8\} = 1-F(8) = 1-F(9-1) = 0.003803 < 0.01$

所以，至少需要配备8 个修理工人.

$X(t)$(随机变量) ~ $P(\lambda_0 t)$:书上等给出泊松分布的参数$\lambda=\lambda_0 t$

参数λ_0是单位时间$\lambda_0 t$内事件出现的平均次数(平均到达率或强度).

例：设小轿车、大型客车，运输车三类车独立到达大型收费站该, 这三类车在$[0, t)$时间内的到达数是强度分别为λ_1, λ_2及λ_3的泊松分布.

(1)汽车在$[0, t)$时间内的到达数的分布?

(2) 非小轿车在$[0, t)$时间内的到达数的分布?

宇 宙 粒 子

例10:已知运载火箭在飞行中,进入它的仪器舱的宇宙粒子数服从参数为λ的泊松分布, 而进入仪器舱的每个粒子落到仪器的重要部位的概率等于p,试求恰有k个粒子落到仪器重要部位的概率.

分析：粒子落到仪器重要部位的试验是由前后相关联的两个试验所组成：

第一个试验是宇宙粒子进入仪器舱，进入的粒子数服从泊松分布；

再是进入仪器舱的这些粒子落到仪器舱重要部位，其数目服从什么分布？二项分布

这类问题可用全概率公式求解。

已知运载火箭在飞行中,进入它的仪器舱的宇宙粒子数服从参数为λ的泊松分布.而进入仪器舱的每个粒子落到仪器的重要部位的概率等于p,试求恰有k个粒子落到仪器重要部位的概率.

解：

设 X表示宇宙粒子进入仪器舱的个数。

由题设X~P(λ) 即 $P\{X=m\}=\frac{\lambda^m}{m!}e^{-\lambda}\quad, m=0,1,2,\ldots.$

显然 {X=m} ,(m=0,1,2,..)构成样本空间的一个划分

（回忆：样本空间的有限划分如何定义？）

设 Y表示落到重要部位的粒子数，由题意知

$$P\{Y=k|X=m\}=C_m^k p^k(1-p)^{m-k},\quad k=0,1,2,\ldots,m$$

由全概率公式得所求概率为

$$P\{Y=k\}=\sum_{m=o}^{+\infty}P\{X=m\}P\{Y=k|X=m\}$$

已知运载火箭在飞行中,进入它的仪器舱的宇宙粒子数服从参数为λ的泊松分布.而进入仪器舱的每个粒子落到仪器的重要部位的概率等于p,试求恰有k个粒子落到仪器重要部位的概率.

$$P\{Y=k\}=\sum_{m=o}^{+\infty}P\{X=m\}P\{Y=k|X=m\}$$

$$=\sum_{m=o}^{+\infty}\frac{\lambda^m e^{-\lambda}}{m!}*C_m^k p^k(1-p)^{m-k} \qquad (m\ge k)$$

$$=\frac{(\lambda p)^k}{k!}e^{-\lambda}\sum_{m=k}^{+\infty}\frac{\lambda^{m-k}(1-p)^{m-k}}{(m-k)!}$$

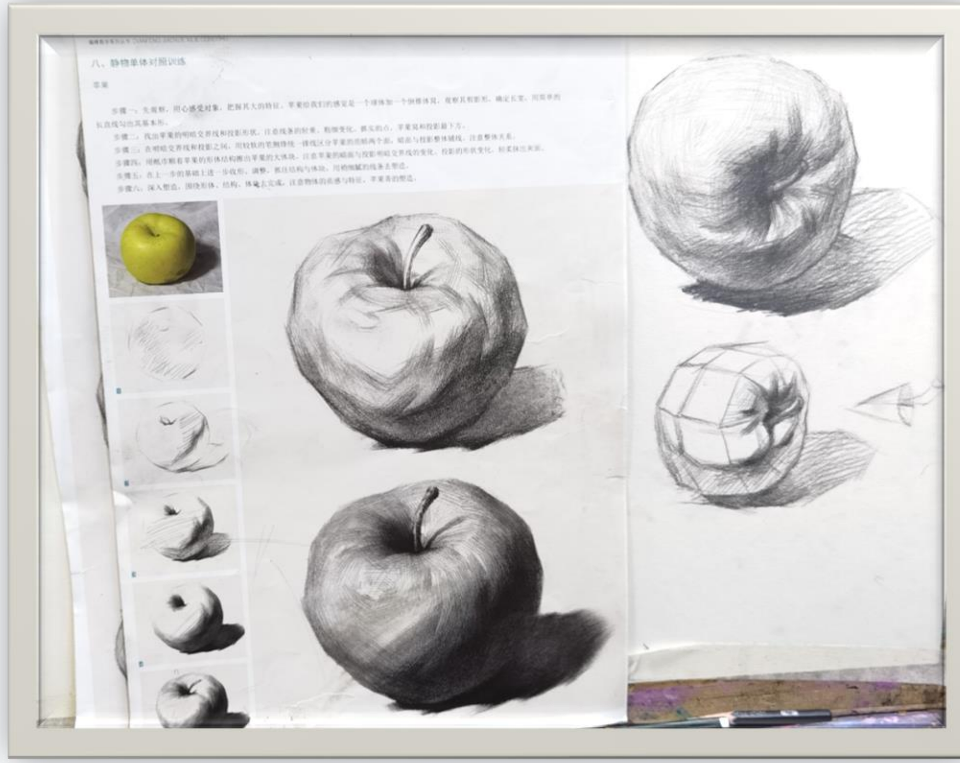
$$\underline{\underline{n=m-k}}\ \frac{(\lambda p)^k}{k!}e^{-\lambda}\sum_{n=0}^{+\infty}\frac{[\lambda(1-p)]^n}{n!}$$

$$=\frac{(\lambda p)^k}{k!}e^{-\lambda}*e^{\lambda(1-p)}=\frac{(\lambda p)^k}{k!}e^{-\lambda p} \qquad k=0,1,2,\ldots.$$

落到仪器重要部位的粒子数服从参数为λp 的泊松分布。

本次课的重点是：

离散型随机变量的分布律及与分布函数的区别，二项分布,泊松分布的分布律及相关计算



下次课内容：

§2.3 连续型随机变量的中的指数分布，包括连续型随机变量的密度函数, 均匀分布及指数分布.