概率论与随机过程

唐碧华 *bhtang@bupt.edu.cn*

- 学时数：48
- 教材：刘次华编《随机过程》，华中科技大学出版社第四版
- 参考书：

1. 《概率论与数理统计》，浙江大学或同济大学出版社
2. 《随机过程及其应用》，陆大琻，清华大学出版社
3. 《工程系统中的随机过程》，陆传赉，北京邮电大学出版社
4. **mathworld.wolfram.com**

教学安排

- 先修课程：高等数学、线性代数、复变函数、初等概率论与随机过程
- 考核方式：随堂练习**30%**、考勤**10%**，期末考试**60%**
- 电子讲稿：登陆**prof_tang@163.com**

  密码：**tang123456**
- 考试时间：见研究生院发布的考表
- 电子邮件：**bhtang@bupt.edu.cn**
- 上课时间：周二晚上**(3-233)**

概率论简史

17、18世纪，数学获得了巨大的进步。数学家们冲破了古希腊的演绎框架，向自然界和社会生活的多方面汲取灵感，数学领域出现了众多崭新的生长点，而后都发展成了完整的数学分支。除了分析数学这一大系统之外，概率论正是这一时期“使欧几里得几何相形见绌”的若干重大成就之一。

概率论简史

概率论起源于赌博问题的研究，有史记载**15**世纪上半叶，便有数学家在考虑这类问题，如意大利数学家帕乔利在**1494**年出版的《算术》里描述了这样的问题：两人进行赌博，规定谁先获胜**6**场谁为胜者。当甲获胜**5**场，乙获胜**2**场，比赛因故中断，问如何分配赌注？帕乔利的答案是赌注分成**7**份，按**5:2**分给甲乙两人。意大利学者卡丹认为不妥，应按**10:1**来分配（现在看来也是错的），卡丹与塔塔里亚等人就数学角度研究过赌博问题，他们的研究除了赌博外还与当时的人口、保险业等有关，但因为赌博的意味太浓，未能引起数学家的重视，概率概念的要旨也不明确，于是很快就被人淡忘了。

概率论简史

伽利略(**Galileo**，近代自然科学创始人之一，**1564-1642**) 解决了以下问题：同时投下三颗骰子，点数和为**9**的情形有**6** 种：**(1,2,6)**、**(1,3,5)**、**(1,4,4)**、**(2,2,5)**、**(2,3,4)** 和**(3,3,3)**。 点数和为**10** 的情形也有**6** 种：**(1,3,6)**、**(1,4,5)**、**(2,2,6)**、**(2,3,5)** **(2,4,4)** 和**(3,3,4)**，那么出现点数和为**9**与**10**的机会应相同，而经验告知，出现**10** 的机会比出现**9** 的机会要多，原因何在？伽利略利用列举法得出同时掷三颗骰子出现点数和为**9**的情形有**25**种，而出现点数和为**10**的情形却有**27**种。可见，已经产生了概率论的某些萌芽。

概率论简史

概率概念的要旨在17世纪中叶法国数学家帕斯卡与费马的讨论中才比较明确。

概率论简史

1653年的夏天，法国著名的数学家、物理学家帕斯卡（Blaise Pascal，1623——1662）前往浦埃托镇度假，旅途中，他遇到了“赌坛老手”梅累。为了消除旅途的寂寞，梅累向帕斯卡提出了一个十分有趣的“分赌注”的问题。问题如下：梅累与其赌友赌掷骰子，每人押了32个金币，并事先约定：如果梅累先掷出三个6点，或其赌友先掷出三个4点，便算赢家。遗憾的是，这场赌注不算小的赌博并未能顺利结束。当梅累掷出两次6点，其赌友掷出一次4点时，梅累接到通知，要他马上陪同国王接见外宾。君命难违，但就此收回各自的赌注又不甘心，他们只好按照已有的成绩分了这64个金币。这下可把他难住了。所以，当他碰到大名鼎鼎的帕斯卡，就迫不及待地向他请教了。然而，梅累的貌似简单的问题，却真正难住他了。虽然经过了长时间的探索，但他还是无法解决这个问题。

概率论简史

1654年左右，帕斯卡与费马在一系列通信中讨论了类似的“合理分配赌金”的问题。

该问题简化为：甲、乙两人同掷一枚硬币，规定：正面朝上，甲得一点；若反面朝上，乙得一点，先积满3点者赢取全部赌注。假定在甲得2点、乙得1点时，赌局由于某种原因中止了，问应该怎样分配赌注才算公平合理？

帕斯卡：若再掷一次，甲胜，甲获全部赌注，乙胜则甲、乙平分赌注。这两种情况可能性相同，所以这两种情况平均一下，甲应得赌金的3/4，乙得赌金的1/4。

费马：结束赌局至多还要2局，结果为四种等可能情况：

情况	1	2	3	4
结果	甲甲	甲乙	乙甲	乙乙

前3种情况，甲获全部赌金，仅第四种情况，乙获全部赌注。所以甲分得赌金的3/4，乙得赌金的1/4。

概率论简史

该问题一般情形为："两个赌徒约定赌若干局，且谁先赢 s 局便算赢家，若在一赌徒胜 a 局（$a<s$），另一赌徒胜b局($b<s$)时便终止赌博，问应如何分赌本？

费马：由甲已胜a 局，乙已胜b 局，要结束这场赌博最多还需要赌 $(s-a)+(s-b)-1$ 局。

帕斯卡：给出了以上问题的通解，令：

$m=(s-a)$，$n=(s-b)$，则甲乙应得赌金之比为：

$$\frac{C_{m+n-1}^{0}+C_{m+n-1}^{1}+\ldots+C_{m+n-1}^{n-1}}{C_{m+n-1}^{0}+C_{m+n-1}^{1}+\ldots+C_{m+n-1}^{m-1}}$$

概率论简史

费马和帕斯卡虽然没有明确定义概率的概念，但是他们定义了使某赌徒获胜的机遇，也就是赢的情况数与所有可能情况数的比，这实际上就是概率，所以概率的发展被认为是从帕斯卡和费马开始的。

正如对概率论有卓越贡献的法国数学家泊松(Poisson，1781-1840) 后来所说：由一位广有交游的人向一位严肃的冉森派所提出的一个关于机会游戏的问题乃是概率演算的起源。

概率论简史

在这期间，荷兰数学家惠更斯(C.Huygens，1629-1695)恰好在巴黎，也参与过他俩的讨论。后来，在1657年，他把讨论结果写成了一本书《论赌博中的计算》，这是概率论发展史上的第一本著作，书首次给出了概率论的一些概念和定理（如加法定理、乘法定理），关于数学期望是这样描述的：在赌局开始之前，每个赌徒就赌博的结局有一种（期望），如果共有N种可能，其中n种可能使他获得赌金a元，其余结果使他获得赌金b元，则他的期望是：

$$\frac{na+(N-n)b}{N}\ldots$$这就是数学期望的概念雏形

概率论简史

在概率论的现代表述中，概率是基本概念，数学期望是第二级的概念，但在历史上顺序却相反，先有数学期望的概念，而古典概型的概率定义，完全可以从期望概念中导出来。因此，可以认为概率论从此诞生了。

使概率论成为数学一个分支的真正奠基人是瑞士数学家雅各布·伯努利，他的重要贡献是建立了概率论中的第一个极限定理，即伯努利大数定律，发表在1713出版的遗著《猜度术》中。伯努利大数定律刻画了大量经观测中呈现的稳定性，美国概率史专家海金（Hacking）称此书标志着“概率漫长的形成过程的终结与数学概率论的开端”。

概率论简史

到1730年，法国数学家棣莫弗出版其著作《分析杂论》，当中包含了著名的“棣莫弗—拉普拉斯定理”。这就是概率论中第二个基本极限定理（中心极限定理）的原始雏形；棣莫弗于1733年、高斯于1809年各自独立引入了正态分布；蒲丰于1777年提出了投针问题的几何概率；泊松于1837年发表了泊松大数定律；拉普拉斯于1812年出版的《概率的分析理论》中，首先明确地对概率做了古典的定义。拉普拉斯以强有力的分析工具处理了概率论的基本内容，实现了从组合技巧向分析方法的过渡，使以往零散的结果系统化，开辟了概率论发展的新时期。

概率论简史

另一在概率论发展史上的代表人物是法国的泊松。他推广了伯努利形式下的大数定律，研究得出了一种新的分布，就是泊松分布。概率论继他们之后，在19世纪后期，其中心研究课题则集中在推广和改进伯努利大数定律及中心极限定理。

俄国数学家切比雪夫对此做出了重要贡献。他于1866年建立了关于独立随机变量序列的大数定律，使伯努利大数定律和泊松大数定律成为其特例，将棣莫弗—拉普拉斯的极限定理推广为更一般的中心极限定理。切比雪夫的成果后被其学生马尔可夫发扬光大，影响了20世纪概率论发展的进程。

概率论简史

19世纪末，一方面概率论在统计物理等领域的应用提出了对概率论基本概念与原理进行解释的需要，另一方面，科学家们在这一时期发现的一些概率论悖论也揭示出古典概率论中基本概念存在的矛盾与含糊之处。这些问题却强烈要求对概率论的逻辑基础做出更加严格的考察。也就是建立概率论的公理化体系。

概率论简史

俄国数学家伯恩斯坦和奥地利数学家冯•米西斯(R.von Mises，1883-1953)对概率论的严格化做了最早的尝试。但他们提出的公理理论并不完善。事实上，真正严格的公理化概率论只有在测度论和实变函数理论的基础才可能建立。测度论的奠基人，法国数学家博雷尔(E.Borel，1781-1956)首先将测度论方法引入概率论重要问题的研究，并且他的工作激起了数学家们沿这一崭新方向的一系列搜索。特别是原苏联数学家科尔莫戈罗夫的工作最为卓著。

概率论简史

1933年，科尔莫戈罗夫出版了他的著作《概率论基础》，这是概率论的一部经典性著作。其中，科尔莫戈罗夫给出了公理化概率论的一系列基本概念，提出了六条公理，整个概率论大厦可以从这六条公理出发建筑起来。科尔莫戈罗夫的公理体系逐渐得到数学家们的普遍认可。由于公理化，概率论成为一门严格的演绎科学，并通过集合论与其它数学分支密切地联系着。

概率论简史

在公理化基础上，现代概率论取得了一系列理论突破。公理化概率论首先使随机过程的研究获得了新的起点。1931年，科尔莫戈罗夫用分析的方法奠定了一类普通的随机过程——马尔可夫过程的理论基础。

科尔莫戈罗夫之后，对随机过程的研究做出重大贡献而影响着整个现代概率论的重要代表人物有莱维(P.Levy，1886-1971)、辛钦、杜布(J.L.Dob)和伊藤清等。

概率论简史

1948年莱维出版的著作《随机过程与布朗运动》提出了独立增量过程的一般理论，并以此为基础极大地推进了作为一类特殊马尔可夫过程的布朗运动的研究。1934年，辛钦提出平稳过程的相关理论。1939年，维尔(J.Ville)引进"鞅"的概念，1950年起，杜布对鞅概念进行了系统的研究而使鞅论成为一门独立的分支。从1942年开始，日本数学家伊藤清引进了随机积分与随机微分方程，不仅开辟了随机过程研究的新道路，而且为随机分析这门数学新分支的创立和发展奠定了基础。

概率论与随机过程

- **知识从哪里来?**

  必然性、偶然性

- **知识是什么?**

  概率论与随机过程：随机性、变化过程

- **知识到哪里去?**

  如何运用概率论与随机过程的理论知识解决通信中的实际问题？

第一章 概率论

1.1 预备知识（集合论）

1.2 概率空间

1.3 随机变量及其分布

1.4 随机变量的数字特征

1.5 随机变量的特征函数

概率空间

定义　若$(\Omega,\mathcal{F})$称为可测空间， P(A)是定义在$\mathcal{F}$上的实函数，如果：

1. $\forall \mathbf{A} \in \mathcal{F}, 0 \le P(\mathbf{A}) \le 1$（非负性）

2. $P(\Omega) = 1$ （归一性）

3. 若$\mathbf{A}_i \in \mathcal{F} (i = 1,2,\cdots)$，且$\mathbf{A}_i \bigcap \mathbf{A}_j = \Phi (i \ne j)$，则：

$$P\left(\bigcup_{k=1}^{\infty} \mathbf{A}_k\right) = \sum_{k=1}^{\infty} P(\mathbf{A}_k)$$（可列可加性）

称$P(\mathbf{A})$为事件**A**的概率，**(**$\Omega,\mathcal{F}$**,P)**称为概率空间。

第一章 概率论

1.1 预备知识（集合论）

1.1.1 集合的基本概念

集合、元素、集合的表示法、子集、包含、集合相等、真子集、空集、幂集、全集

1.1.2 集合的基本运算

并集、交集、相对补集、绝对补集、对称差、文氏图、算律、

1.1.3 集合中元素的计数

基数、有（无）穷集、包含排斥原理

集合的基本概念

✓把具有共同性质的一些东西，汇集成一个整体，就形成一个集合(***set***) 。

✓ 由确定的相互区别的一些对象组成的整体称为集合。

✓ 确定的可分辨的事物构成的整体。

例：北京邮电大学的学生、图书馆的藏书、全国的高等学校、自然数的全体、直线上的点、**26**个英文字母

集合的元素(*element*)

- ✓集合内的对象或单元称为元素。
- ✓集合通常用大写英文字母表示。例如，**N**代表自然数集合（包括**0**），**Z**代表整数集合，**Q**代表有理数集合，**R**代表实数集合，**C**代表复数集合。

集合的表示法

列举法 将集合中的元素一一列举，或列出足够多的元素以反映集合中元素的特征。

例如：**V=**$\{a,e,i,o,u\}$ 或**B={1,4,9,16,25,36……}**。

- ✓显式表示法
- ✓优点：具有透明性。
- ✓缺点：在表示具有某种特性的集合或元素过多时受到一定的局限。而且，从计算机的角度看，显式法是一种“静态”表示法，若一下子将这么多数据“输入”计算机，将占据大量内存。

描述法 通过描述集合中元素的共同特征来表示集合。例如：

V= $\{x|\ x$是元音字母$\}$

B= $\{x|\ x=a^2,\ a$是自然数$\}$

C= $\{x|\ x\in \mathbf{Z} \wedge 3<x\leq 6\}$，即C=$\{4,5,6\}$

✓突出优点：原则上不要求列出集合中全部元素，而只要给出该集合中元素的特性即可。

1.1.1 集合的基本概念

归纳法 通过归纳定义集合，主要由下面三部分构成：

基础：指出某些最基本的元素属于某集合；

归纳：指出由某些元素造出新元素的方法；

极小性：指出给集合的界限。

例如：命题公式的定义：

（1）单个命题常项或变项$p, q, r, \ldots p_i, q_i, r_i, \ldots, 0, 1$是命题公式；

（2）若A、B是合式公式，则$\neg A$、$A \wedge B$、$A \vee B$、$A \rightarrow B$、$A \leftrightarrow B$也是命题公式；

（3）只有有限次地应用（1）—（3）组成的符号串才是命题公式。

递归法 通过计算法则来定义集合中的元素。

例如：

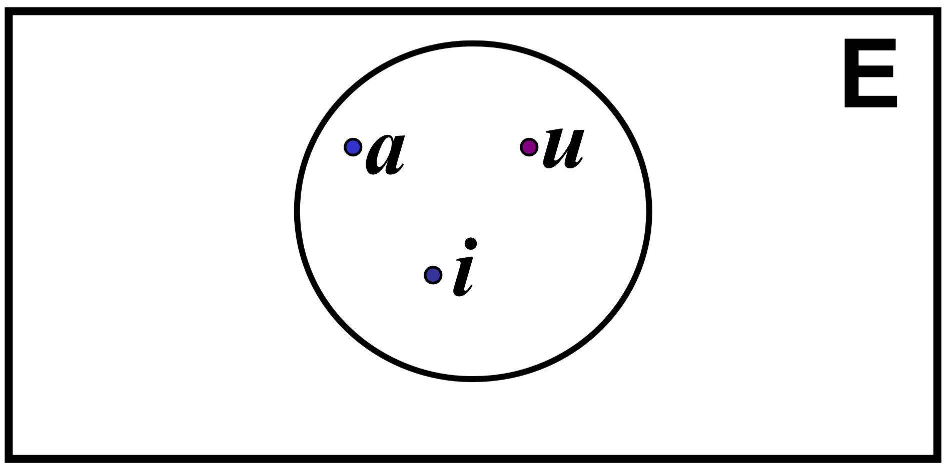
$$V=\{x_n \mid x_0=1,\ x_{n+1}=2x_n+1\}$$

BNF范式法(***Backus Normal Form***)：常常用来定义高级程序设计语言的标识符或表达式集合。

集合的表示法——文氏图(*Venn Diagram*)

用一个大的矩形表示全集(最大的集合——所涉及的全部集合都是这个集合的子集)，在矩形内画一些圆或其它的几何图形来表示集合，有时也用一些点来表示集合中的特定元素。

例如：集合V={a,e,i,o,u}，用文氏图表示如下：



集合中元素的表示

- ✓ 元素a属于集合**A**，记作$a \in$**A**。
- ✓ 元素a不属于集合**A**，记作$a \notin$ **A**

集合的特征

✓确定性：任何一个对象，或者是这个集合的元素，或者不是，二者必居其一。

例如：A={x| $x \in$ N $\wedge$ x<100}

C={x | x是北京邮电大学电子工程学院11级本科生}

✓互异性：集合中任何两个元素都是不同的，即集合中不允许出现重复的元素。

例如： 集合A={a,b,c,c,b,d}，应该是A={a,b,c,d}

集合的特征

✓无序性：集合与其中的元素的顺序无关。

例如： 集合$\{a,b,c,d,e\}$，$\{d,c,e,a,b\}$，$\{e,c,d,b,a\}$都是表示同一个集合。

✓多样性：集合中的元素可以是任意的对象，相互独立，不要求一定要具备明显的共同特征。

例如：$A=\{a,\{a\},\{\{a\},b\},\{\{a\}\}, 1\}$

$A=\{1,a,*,-3,\{a,b\},\{x|x\text{是汽车}\},\text{地球}\}$

集合之间的关系

定义1.1 子集(*subset*) 设A，B是两个集合，若A的元素都是B的元素，则称A是B的子集，也称B包含A，或A包含于B，记作A $\subseteq$ B，或B $\supseteq$ A （否则记为B $\not\supseteq$ A ）

例：A={0,1,2}，B={0,1}，C={1,2}，则B $\subseteq$ A， C$\subseteq$ A

子集的符号化形式为： $A \subseteq B \Leftrightarrow \forall x\,(x \in A \rightarrow x \in B)$

特别地，对任意集合S，均有：S $\subseteq$ S

集合之间的关系

定义1.2 集合相等 设A和B为两个集合，若A $\subseteq$ B且B $\subseteq$ A，则称A与B相等，记作A = B；若A与B不相等，记做A $\neq$ B。

集合A与B相等的符号化形式为：

$$A=B \Leftrightarrow A \subseteq B \wedge B \subseteq A$$

例1.1 设A=$\{x|\ x$是偶数，且$0<x<10\}$，B=$\{2,4,6,8\}$，则A=B。

实例

判断A＝B？

1. $A=\{x \mid x$ 是小于等于3的素数$\}$，
   $B=\{x \mid x=2 \vee x=3\}$， ✓

2. {1,2,4}和{1,2,2,4} ✓

3. {1,2,4}和{1,4,2} ✓

4. {{1,2},4}和{1,4,2} ✗

5. {1,3,5,…}和{x|x是正奇数} ✓

集合之间的关系

定义1.3 若$A \subseteq B$，且$A \neq B$，则称A是B的真子集(*proper subset*)，也称B真包含A，或A真包含于B，记为$A \subset B$，或$B \supset A$。

事实上，$A=\{0,1,2\}$，$B=\{0,1\}$，则：$B \subset A$

✓ ✗ ✗

判断：{0,1}、{1,3}、{0,1,2}是｛0,1,2}的真子集吗？

定义1.4 空集 不含任何元素的集合，称为空集（*empty set*），记作∅。

空集的符号化形式为： $\varnothing = \{x \mid x \neq x\}$

如：$A = \{x \mid x \in R \wedge x^2+1=0\} = \varnothing$

定理1.1空集是一切集合的子集。

推论 空集是唯一的。

例1.2 确定下列命题是否为真

(1) $\varnothing \subseteq \varnothing$；(2) $\varnothing \in \varnothing$；(3) $\varnothing \subseteq \{\varnothing\}$；(4) $\varnothing \in \{\varnothing\}$。

✓ ✗ ✓ ✓

集合之间的关系*

定义1.5 幂集(*power set*) 设A是集合，A的所有子集构成的集合称为A的幂集，记以P(A)或 2^A。

例1.3 A={a,b,c}，则：

$$P(A)=\{\varnothing,\{a\},\{b\},\{c\},\{a,b\},\{a,c\},\{b,c\},\{a,b,c\}\}$$

注：“*”内容为精讲。

集合A的幂集的符号化形式为：

$$P(A)=\{B \mid B \subseteq A\}$$

幂集的性质：

若A为有穷集，A有n个元素，则A的幂集2^A所含元素的个数为： $C_n^0 + C_n^1 + \dots + C_n^n = 2^n$ 。

B$\in$P(A)当且仅当B $\subseteq$A。

设A、B是两个集合，A$\subseteq$B当且仅当 P(A)$\subseteq$P(B)。

P(A)实际上是一个集合簇，或集合类。

实　例

例3.3　计算以下幂集：

(1) $P(\varnothing)$;

(2) $P(\{\varnothing\})$;

(3) $P(\{\varnothing,\{\varnothing\}\})$;

(4) $P(\{1,\{2,3\}\})$.

解　(1) $P(\varnothing)=\{\varnothing\}$;

(2) $P(\{\varnothing\})=\{\varnothing,\{\varnothing\}\}$;

(3) $P(\{\varnothing,\{\varnothing\}\})=\{\varnothing,\{\varnothing\},\{\{\varnothing\}\},\{\varnothing,\{\varnothing\}\}\}$;

(4) $P(\{1,\{2,3\}\})=\{\varnothing,\{1\},\{\{2,3\}\},\{1,\{2,3\}\}\}$.

1.1.1 集合的基本概念

定义1.6 在一个具体问题中，如果所涉及的集合都是某个集合的子集，则称这个集合为全集，记作**E(或U)**

例1.4 著名的罗素悖论：在一个很僻静的孤岛上，住着一些人家，岛上只有一个理发师，该理发师专给那些并且只给那些自己不刮脸的人刮脸。那么，谁给这位理发师刮脸？

解：设C={x| x为那些自己不刮脸的人}

b是这位理发师。

若：b∈C⇒b是自己不刮脸的人，那么谁给他刮脸？因岛上只有一个理发师，所以只能是他自己给自己刮脸⇒ b∉C

若：b∉C⇒ b是自己刮脸的人，而理发师只有一个，只要是被理发师刮脸的人一定是自己不刮脸的⇒ b∈C

罗素悖论

将集合分为两类：

✓集合A本身是集合A的元素，即$A \in A$

✓集合A本身不是集合A的元素，即$A \notin A$

构造集合$S= \{A，A \notin A\}$

问题：S是不是它自身的元素？

✓若$S \in S$，因集合S中中每个元素A均不属于本身，因此必须有$S \notin S$

✓若$S \notin S$，因集合S中中每个元素A均不属于本身，因此必须有$S \in S$

罗素悖论

举例说明A∈A

如：**A={x| x为所有不是苹果的东西}**

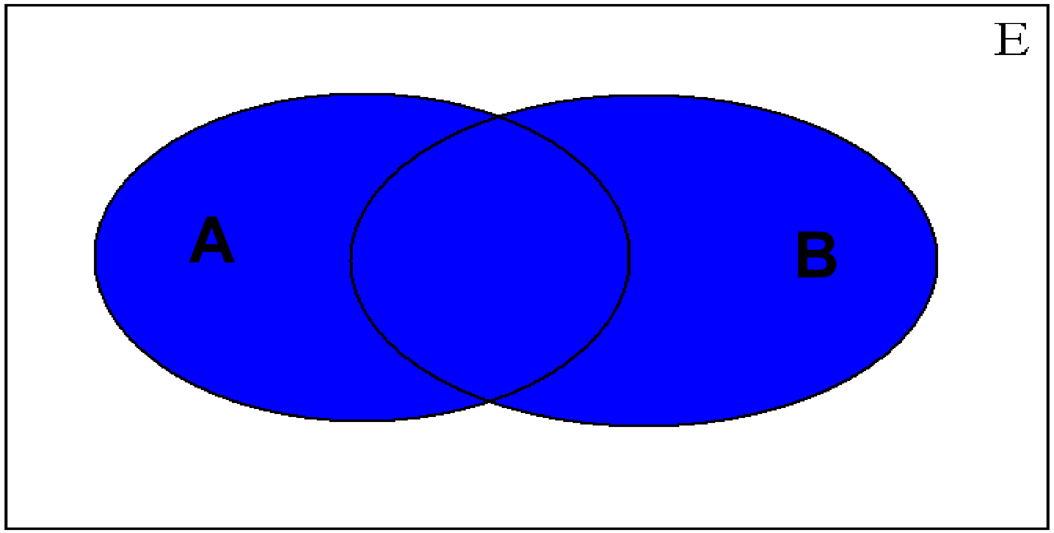
显然有：**A∈A**

定义1.7集合的并集(*Union*)

A∪B定义为： $A \cup B=\{x \mid x \in A \vee x \in B\}$

例如，令$A=\{a, b, c, d\}$，$B=\{c, d, e, f\}$，于是$A \cup B=\{a, b, c, d, e, f\}$。

并集的文氏图



A∪B

定义1.8 集合的交集(*Intersection*)

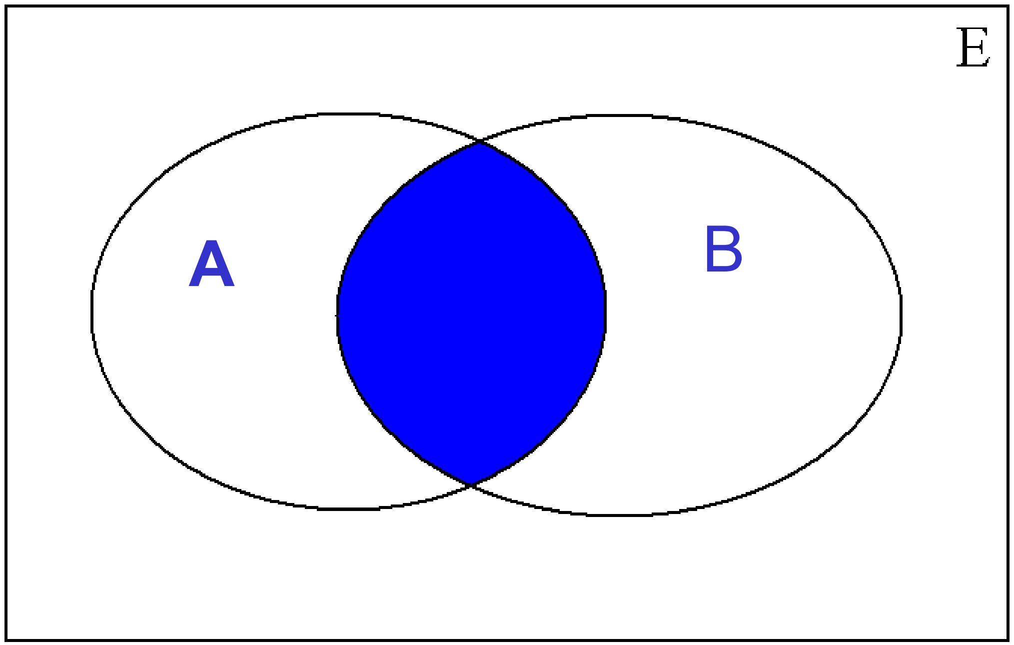
A∩B定义为：**A∩B={*x*| *x*∈A∧*x*∈B}**

例如，令**A={*a, b, c, d*}**，**B={*c, d, e, f*}**，于是

A∩B={*c, d* }。

交集的

文氏图



A∩B

并集和交集的推广：

设A_1，A_2，…，A_n …是可数*多个集合，则：

$A_1 \cup A_2 \cup \ldots \cup A_n \ldots$，简记为 $\bigcup_{i=1}^{\infty} A_i$

$A_1 \cap A_2 \cap \ldots \cap A_n \ldots$，简记为 $\bigcap_{i=1}^{\infty} A_i$

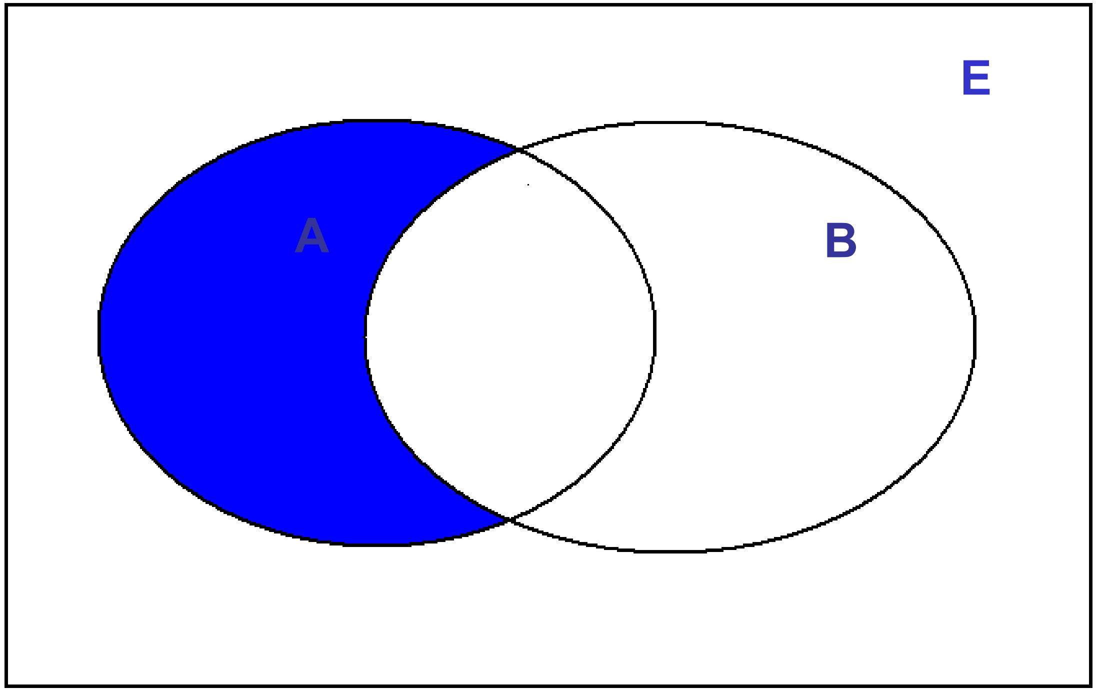
定义1.9 集合的差集(*Subtraction*) 定义为：

$$A\text{-}B=\{x|x\in A \wedge x\notin B\}$$

例：A={*a*, *b*, *c*, *d*}，B={*c*, *d*, *e*, *f*}，则：

$$A\text{-}B=\{a, b\}$$

差集的文氏图

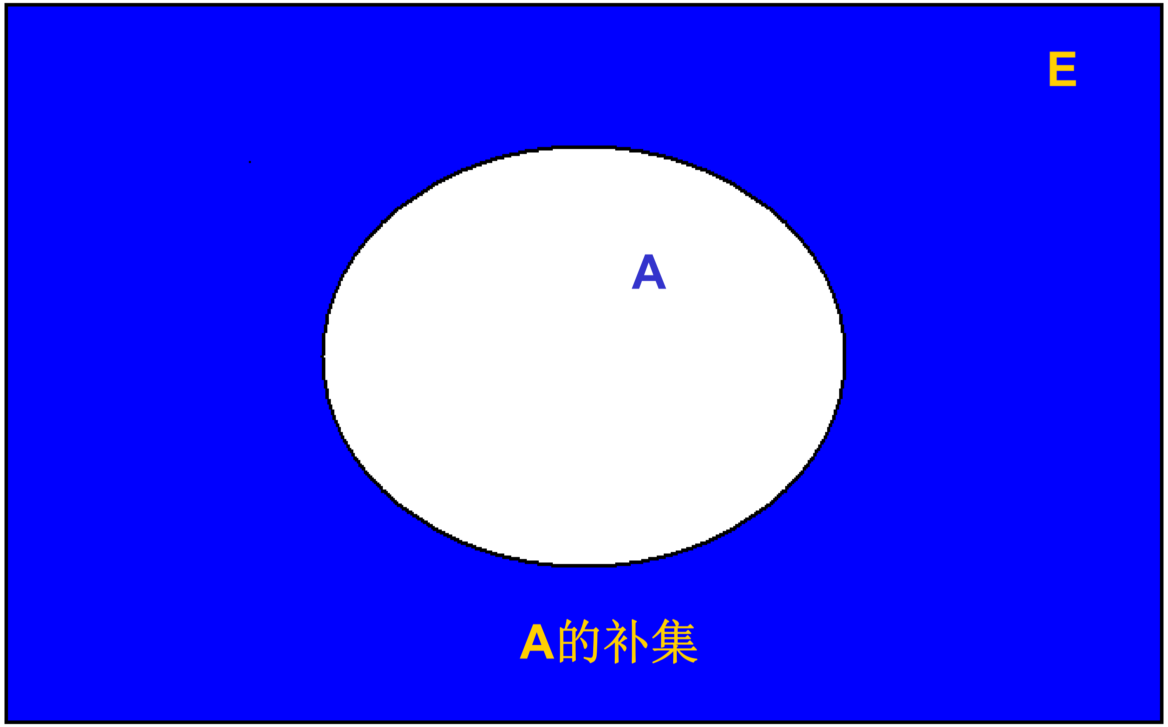


A-B

定义1.10 集合的补集(*Complement*) 全集**E**与**A**的差集称为**A**的余集或补集：

$$\overline{A}=E-A=\{x \mid x\in E\wedge x\notin A\}，\text{特别地}:\overline{\Phi}=\Omega,\overline{\Omega}=\Phi$$

补集的文氏图

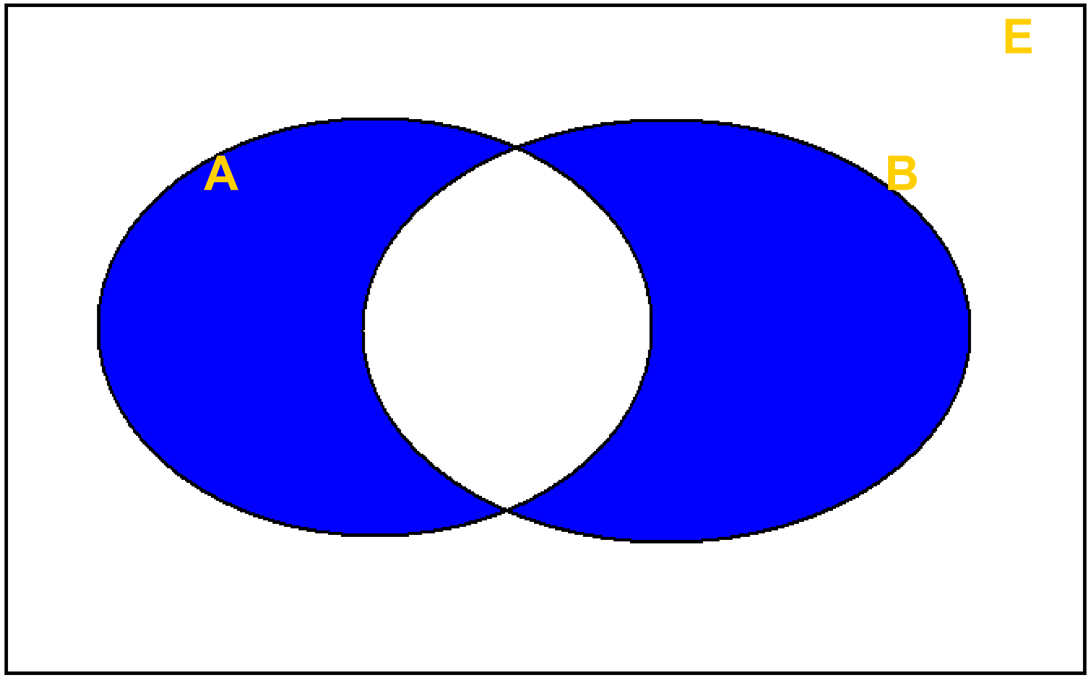


定义1.11 集合的对称差(*Symmetric Subtraction*) A与B的对称差定义为：

$$A\oplus B＝(A-B)\cup(B-A)= =\{x|x\in A \forall\ x\in B\}$$

等价定义：即 $A\oplus B＝(A\cup B)-(A\cap B)$

对称差的文氏图



$A\oplus B$

集合的对称差运算实例

例1.5 A＝{0,1,2}，B＝{2,3}

则有 A⊕B＝(A-B)∪(B-A)＝{0,1}∪{3}＝{0,1,3}

或A⊕B＝(A∪B)-(A∩B) ＝{0,1,2,3}－{2} ＝ {0,1,3}

例1.6 A给定全集U，描述A⊕A、U⊕A、∅⊕A、$A \oplus \overline{A}$

A⊕A=∅

U⊕A= $\overline{A}$

∅⊕A=A　　$A \oplus \overline{A} = U$

集合的基本运算定律：

等幂律：$A \cup A = A$

$A \cap A = A$

结合律：$(A \cup B) \cup C = A \cup (B \cup C)$

$(A \cap B) \cap C = A \cap (B \cap C)$

交换律：$A \cup B = B \cup A$

$A \cap B = B \cap A$

分配律：$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$

$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$

同一律：$A \cup \varnothing = A$

$A \cap \varnothing = \varnothing$

1.1.2 集合的基本运算

零　律：$A\cup E=E$　　　$A\cap\varnothing=\varnothing$

排中律：$A\cup\sim A=E$

矛盾律：$A\cap\sim A=\varnothing$

吸收律：$A\cup(A\cap B)=A$　　　$A\cap(B\cup A)=A$

德摩根律：$A-(B\cup C)=(A-B)\cap(A-C)$

$A-(B\cap C)=(A-B)\cup(A-C)$

$\sim(B\cup C)=\sim B\cap\sim C$

$\sim(B\cap C)=\sim B\cup\sim C$

$\sim\varnothing=E$

$\sim E=\varnothing$

双重否定律：$\sim(\sim A)=A$

例1.7 证明：**A-(B∩C)=(A-B)∪(A-C)**

证明：对$\forall x \in$A-(B∩C)

$$\Rightarrow x \in A \wedge x \notin B\cap C$$

$$\Rightarrow x \in A \wedge (x \notin B \vee x \notin C)$$

$$\Rightarrow (x \in A \wedge x \notin B) \vee (x \in A \wedge x \notin C)$$

$$\Rightarrow x \in A-B \vee x \in A-C$$

$$\Rightarrow x \in (A-B)\cup(A-C)$$

则：A-(B∩C) $\subseteq$ (A-B)∪(A-C)

同理可证：(A-B)∪(A-C) $\subseteq$ A-(B∩C)

则：A-(B∩C)=(A-B)∪(A-C)

集合的其它运算律：

$A \cap B \subseteq A$　　　　$A \cap B \subseteq B$

$A \subseteq A \cup B$　　　　$B \subseteq A \cup B$

$A - B \subseteq A$

$A - B = A \cap \sim B$

$A \cup B = B \Leftrightarrow A \subseteq B \Leftrightarrow A \cap B = A \Leftrightarrow A - B = \varnothing$

$A \oplus B = B \oplus A$

$(A \oplus B) \oplus C = A \oplus (B \oplus C)$

$A \oplus \varnothing = \varnothing \oplus A$

$A \oplus A = \varnothing$

$A \oplus B = A \oplus C \Rightarrow B = C$

例1.8 证明：$(A \oplus B) \oplus C = A \oplus (B \oplus C)$

证明:$(A \oplus B) \oplus C$

$$= ((A \oplus B) \cap \overline{C}) \cup ((\overline{A \oplus B}) \cap C)$$

$$= [((A \cap \overline{B}) \cup (\overline{A} \cap B)) \cap \overline{C}] \cup [\overline{((A \cap \overline{B}) \cup (\overline{A} \cap B))} \cap C]$$

$$= (A \cap \overline{B} \cap \overline{C}) \cup (\overline{A} \cap B \cap \overline{C}) \cup [((\overline{A} \cup B) \cap (A \cup \overline{B})) \cap C]$$

但$((\overline{A} \cup B) \cap (A \cup \overline{B})) \cap C$

$$=[((\overline{A} \cup B) \cap A) \cup ((\overline{A} \cup B) \cap \overline{B}] \cap C$$

$$= ((\overline{A} \cap A) \cup (A \cap B) \cup (\overline{A} \cap \overline{B}) \cup (B \cap \overline{B})) \cap C$$

$$= ((A \cap B) \cup (\overline{A} \cap \overline{B})) \cap C$$

$$= (A \cap B \cap C) \cup (\overline{A} \cap \overline{B} \cap C)$$

故$(A \oplus B) \oplus C$

$$=(A \cap \overline{B} \cap \overline{C}) \cup (\overline{A} \cap B \cap \overline{C}) \cup (A \cap B \cap C) \cup (\overline{A} \cap \overline{B} \cap C)$$

$$A\oplus(B\oplus C)$$

$$=[A\cap(\overline{B\oplus C})]\cup[\overline{A}\cap(B\oplus C)]$$

$$=[A\cap\overline{((B\cap\overline{C})\cup(\overline{B}\cap C))}]\cup[\overline{A}\cap((B\cap\overline{C})\cup(\overline{B}\cap C))]$$

$$=[A\cap((\overline{B}\cup C)\cap(B\cup\overline{C}))]\cup[(\overline{A}\cap B\cap\overline{C})\cup(\overline{A}\cap\overline{B}\cap C)]$$

因为

$$A\cap((\overline{B}\cup C)\cap(B\cup\overline{C}))$$

$$=A\cap[(\overline{B}\cap B)\cup(\overline{B}\cap\overline{C})\cup(C\cap B)\cup(C\cap\overline{C})]$$

$$=A\cap[(\overline{B}\cap\overline{C})\cup(C\cap B)]$$

$$=(A\cap\overline{B}\cap\overline{C})\cup(A\cap B\cap C)$$

故$A\oplus(B\oplus C)$

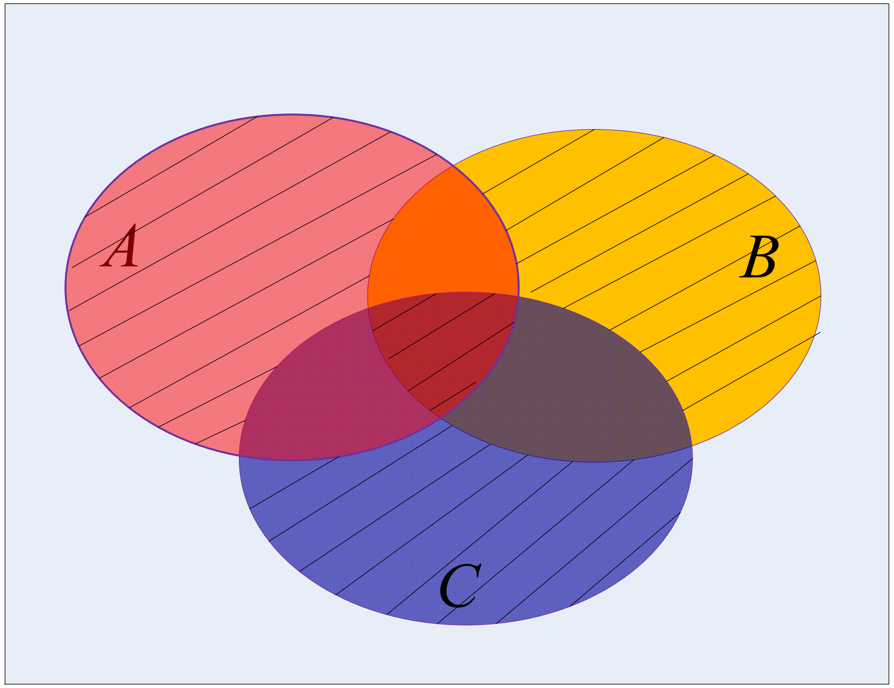
$$=(A\cap\overline{B}\cap\overline{C})\cup(A\cap B\cap C)\cup(\overline{A}\cap B\cap\overline{C})\cup(\overline{A}\cap\overline{B}\cap C)$$

因此:$(A\oplus B)\oplus C=A\oplus(B\oplus C)$

根据前面的证明结果，有如下图结果：

$$(A \oplus B) \oplus C = A \oplus (B \oplus C)$$

$$=(A \cap \overline{B} \cap \overline{C}) \cup (A \cap B \cap C) \cup (\overline{A} \cap B \cap \overline{C}) \cup (\overline{A} \cap \overline{B} \cap C)$$



例1.9 证明$A\oplus B= A\oplus C \Rightarrow B=C$

证明：由$A\oplus B = A\oplus C$，则：

$$A\oplus(A\oplus B) = A\oplus(A\oplus C)$$

$$\Rightarrow (A\oplus A)\oplus B= (A\oplus A)\oplus C$$

$$\Rightarrow \varnothing\oplus B= \varnothing\oplus C$$

$$\Rightarrow B=C$$

基数(*Cardinal Number*)的概念

定义1.12 设A={a_1, a_2,…, a_n}，即A中含n个元素，称集合A的基数是n，记为：$card$ A= n，或|A|= n。

定义1.13 设A为集合，若|A|= n，称A为有穷集合；否则成为无穷集合。

定义1.14 凡能与自然数集建立1-1对应关系的集合（对等集合）称为可数集合，记为：

$$A=\{a_1, a_2, \dots, a_n, \dots\}$$

1.1.3 集合中元素的计数

定理1.2 任何无穷集合**A**都包含一个可数子集。

证明：因**A**是无穷集合，必有：$a_1 \in A$

又因**A**是无穷集合，必有：$a_2 \in A - \{a_1\}$

.......同理，即有：$a_n \in A - \{a_1, a_2, \cdots a_{n-1}\}$

.......则，$\{a_1, a_2, \cdots a_{n-1}, a_n \cdots\} \subseteq A$ 证毕。

上述定理说明可数集合的基数是无穷集合中的最小者。

1.1.3 集合中元素的计数

定理1.3 可数子集**A**的子集如果不是有限集合，自然一定仍然是可数集合。

显然，不必证明。

定理1.4 若集合**A**可数，**B**有限，则**A**∪**B**仍然可数。

无穷集合**A**的基数称为势，无穷集合可以被定义为：

能与它本身的真子集对等的集合。

证明：有理数集Q为可数集合。

证明：设$A_i=\left\{\frac{1}{i},\frac{2}{i},\frac{3}{i},\cdots\right\}(i=1,2,3,\cdots)$，则$A_i$是可数集

$$Q=\left\{0;\frac{1}{1},\frac{-1}{1},\frac{2}{1},\frac{-2}{1},\cdots;\frac{1}{-1},\frac{-1}{-1},\frac{2}{-1},\frac{-2}{-1},\cdots\frac{1}{2},\frac{-1}{2},\frac{2}{2},\frac{-2}{2},\cdots;\cdots\right\}$$

在上述的集合中，去掉重复的元素后一定可以与自然数集N一一对应，因此Q为可数集合。

1.1.3 集合中元素的计数

定理1.6 设$A_i(i=1,2,\cdots)$是可数集合，且A_iA_j=$\Phi(i\neq j)$，则：

$\sum_{i=1}^{\infty}A_i$也是可数的。

证明：因$A_i(i=1,2,\cdots)$是可数集合，设：

$$A_1=\{a_{11},\rightarrow a_{12},a_{13},\rightarrow\cdots\}$$

$$A_2=\{a_{21},a_{22},a_{13},\cdots\}$$

$$A_3=\{a_{31},a_{32},a_{33},\cdots\}$$

………………………………

则:$\sum_{i=1}^{\infty}A_i=\{a_{11},a_{12},a_{21},a_{31},a_{22},a_{13},\cdots\}$，则$\sum_{i=1}^{\infty}A_i$是可数集。

1.1.3 集合中元素的计数

*Cantor*定理 **[0,1]**区间上所有的点构成的集合与**N**不可能对等。

无穷集合的基数、势等问题太复杂，因此我们在此重点讨论有穷集合的计数问题。

集合的分类：

- 集合
  - 无穷集合
    - 不可数集合，势的概念
    - 可数集合，基数的概念
  - 有穷集合

文氏图方法

- ✓ 根据已知条件画出对应的文氏图。

  注意：一般情况下任意两个集合应该是相交的。

- ✓ 将已知的集合的基数填入该集合的相应区域。一般地，从几个集合的交集开始填，逐步到其它空白区域，直到所有区域都填好为止。

例1.10 有**100**名程序员，其中**47**名熟悉**FORTRAN**语言，**35**名熟悉**PASCAl**语言，**23**名熟悉这两种语言。问有多少人对这两种语言都不熟悉？

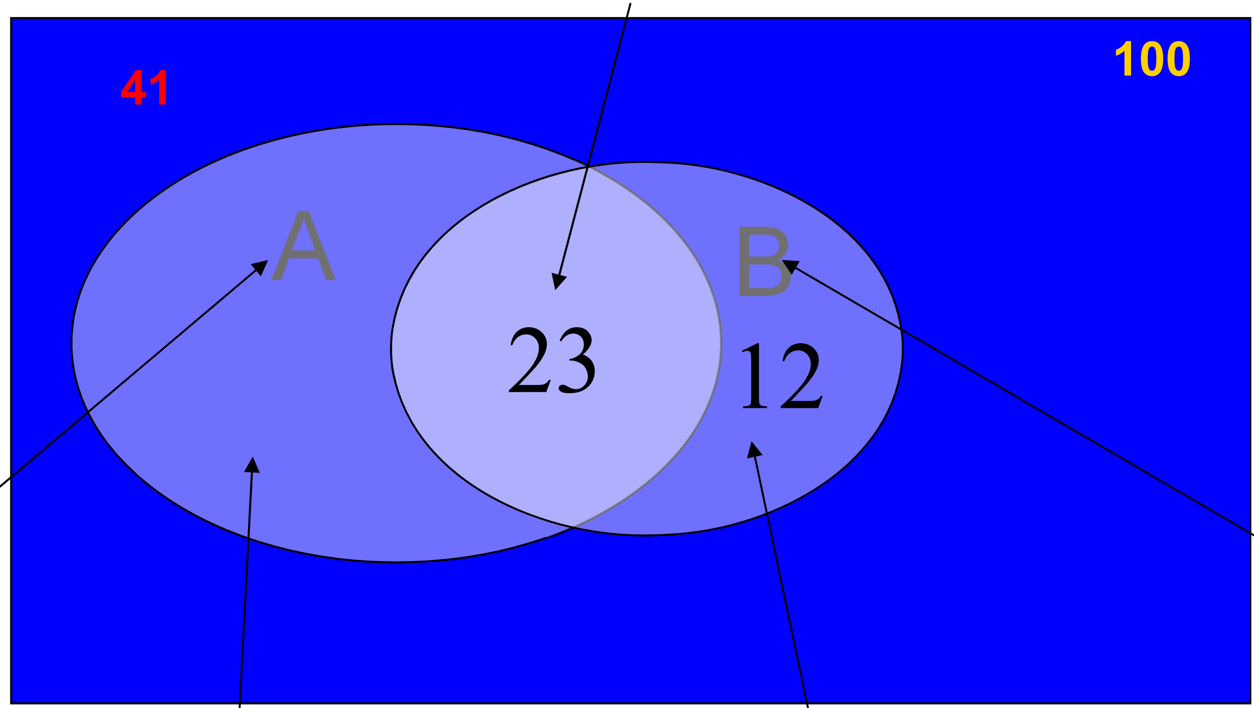
解 设**A**，**B**分别表示熟悉**FORTRAN** 和**PASCAL**语言的程序员的集合，则：

✓首先画出文氏图

✓将已知的集合的基数填入该集合的相应区域

解:

$|A \cap B| = 23$



$|A| = 47$

$|B| = 35$

$|A-B| = 24$ $|B-A| = 12$

$$|A \cup B| = |A-B| + |A \cap B| + |B-A| = 24+23+12=59$$

则： $|\sim(A \cup B)| = 100-59=41$

例1.11 对24名会外语的科技人员进行掌握外语情况的调查，其统计结果如下：会英、日、德和法语的人分别为13、5、10和9人，其中同时会英语和日语的有2人，会英、德和法语任两种语言的都是4人。已知会日语的人既不懂法语也不懂德语，分别求只会一种语言的人数和会三种语言的人数。

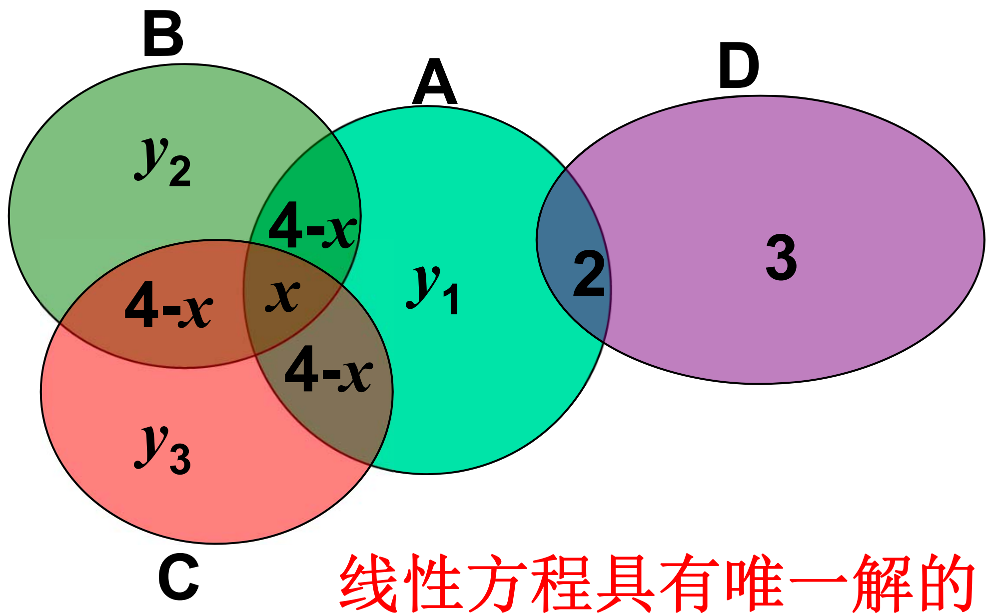
解：A：会英语； B：会法语；C：会德语；D：会日语。

x： 会三种语言；

y_1： 只会英语；

y_2： 只会法语；

y_3： 只会德语。



$$\begin{cases} y_1+2(4-x)+x+2=13 \\ y_2+2(4-x)+x=9 \\ y_3+2(4-x)+x=10 \\ y_1+y_2+y_3+3(4-x)+x=19 \end{cases}$$

线性方程具有唯一解的！

$$x=1, y_1=4, y_2=2, y_3=3$$

例1.12 求1到1000之间（包含1和1000在内）既不能被5和6，也不能被8整除的数有多少。

解：设$S=\{x|x\in Z \wedge 1\le x \le 1000\}$　$A=\{x| x \in S \wedge x$可被5整除$\}$

$B=\{x| x\in S \wedge x$可被6整除$\}$　$C=\{x | x \in S \wedge x$可被8整除$\}$

$|T|$表示有穷集T的元素数，$int(x)$表示小于或等于x的最大整数，$lcm(x_1, x_2,\ldots, x_n)$表示$x_1, x_2,\ldots, x_n$的最小公倍数，则：

$|A|= int(1000/5)=200$　$|B|=int(1000/6)=166$

$|C|= int(1000/8)=125$

$|A\cap B|= int(1000/lcm(5,6))=33$

$|A\cap C|= int(1000/lcm(5,8))=25$

$|B\cap C|= int(1000/lcm(6,8))=41$

$|A\cap B\cap C|= int(1000/lcm(5,6,8))=8$

|A|= int(1000/5)=200

|B|= int(1000/6)=166

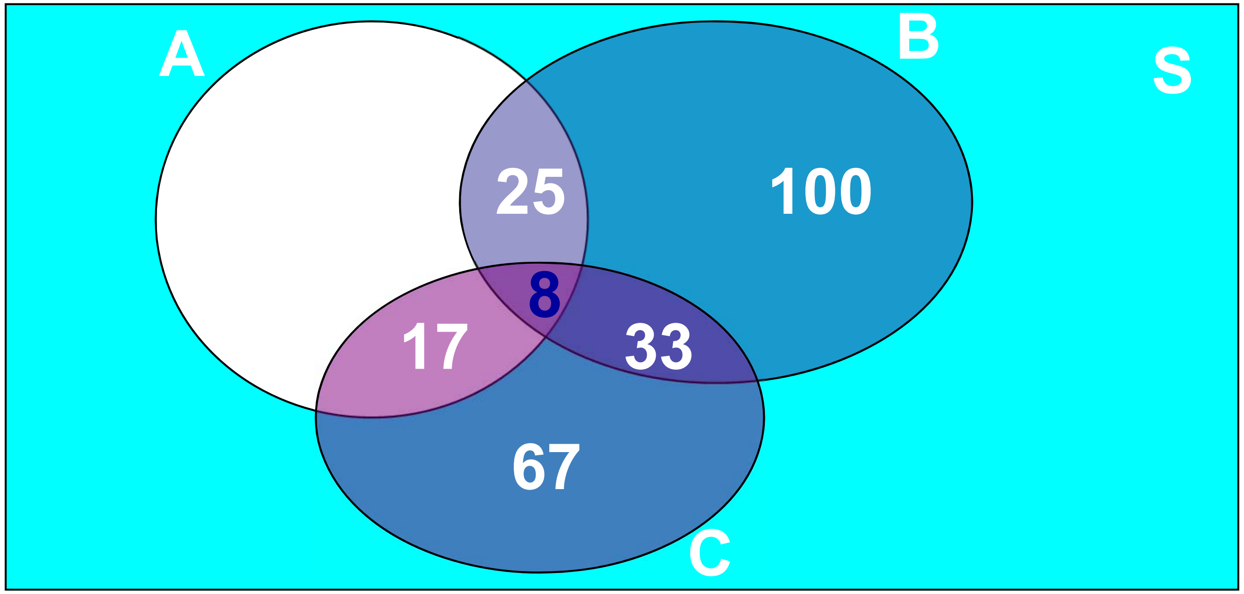
|C|= int(1000/8)=125

|A∩B|= int(1000/lcm(5,6))=33

|A∩C|= int(1000/lcm(5,8))=25

|B∩C|= int(1000/lcm(6,8))=41

|A∩B∩C|= int(1000/lcm(5,6,8))=8



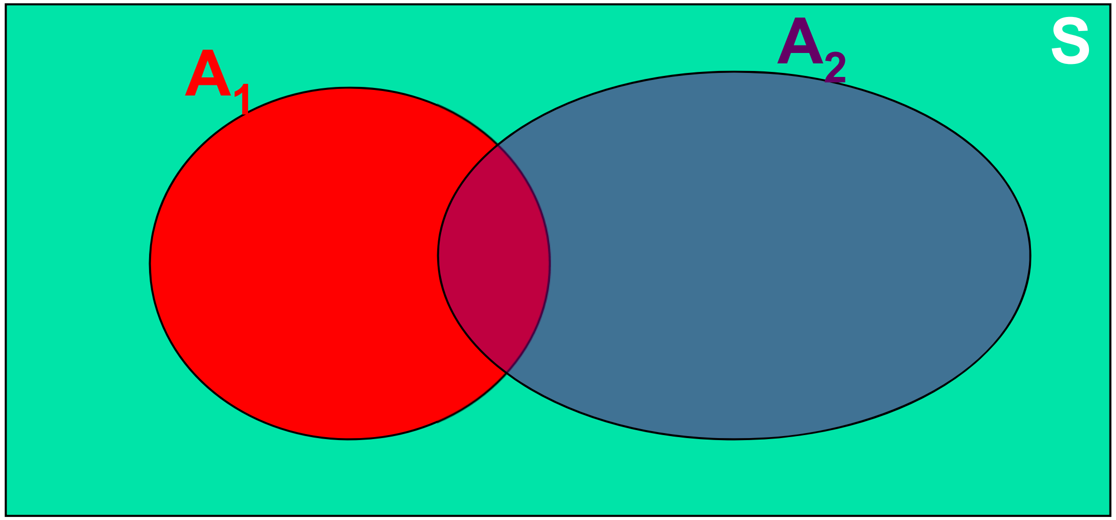
则：不能被5、6和8整除的数有：

1000-(200+100+33+67)=600个

包含排斥原理

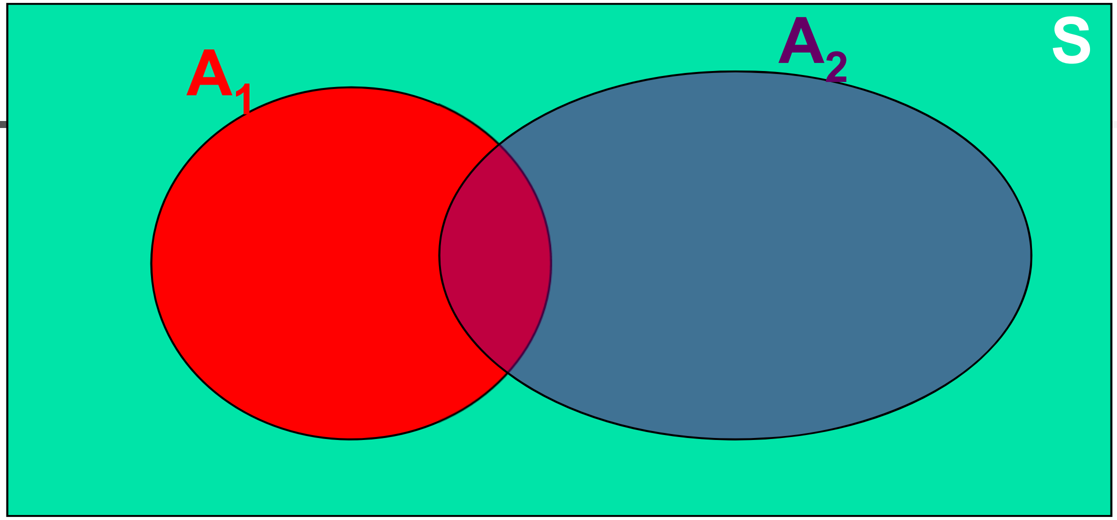
定理1.5 设A_1和A_2为有限集合，其元素个数分别为：

$|A_1|,|A_2|$，则$|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$



证明：若A_1与A_2不相交，即$A_1 \cap A_2 = \Phi$，则$|A_1 \cup A_2| = |A_1| + |A_2|$

若$A_1 \cap A_2 \neq \Phi$，则:$|A_1| = |A_1 \cap \overline{A_2}| + |A_1 \cap A_2|$



同理：$|A_2| = |A_2 \cap \overline{A_1}| + |A_2 \cap A_1|$，且 $|A_1| = |A_1 \cap \overline{A_2}| + |A_1 \cap A_2|$

则：$|A_1| + |A_2| = |A_1 \cap \overline{A_2}| + |A_2 \cap \overline{A_1}| + 2|A_1 \cap A_2|$

但：$|A_1 \cup A_2| = |A_1 \cap \overline{A_2}| + |A_2 \cap \overline{A_1}| + |A_1 \cap A_2|$

所以：$|A_1| + |A_2| = |A_1 \cup A_2| + |A_1 \cap A_2|$，则：$|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$

这就是两个集合并集基数的：包含排斥原理。

包含排斥原理可推广到三个集合的情况，其结果为：

$$|A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3|$$

例1.13在某工厂装配30辆汽车，可供选择的设备是CD机、导航仪、倒车雷达。已知15辆有CD机，8辆有导航仪，6辆有倒车雷达，而且其中3辆汽车这三种设备都有。问：希望知道至少有多少辆汽车没有提供任何设备。

解：设A_1，A_2，A_3分别表示配有CD机、导航、倒车雷达的汽车集合，则根据题意，有：

$|A_1|=15, |A_2|=8, |A_3|=6$，且$|A_1 \cap A_2 \cap A_3| = 3$

$$而|A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3|$$

$$+ |A_1 \cap A_2 \cap A_3|$$

$$= 15 + 8 + 6 - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + 3$$

$$= 32 - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3|$$

$$而: |A_1 \cap A_2| \geq |A_1 \cap A_2 \cap A_3|$$

$$|A_1 \cap A_3| \geq |A_1 \cap A_2 \cap A_3|$$

$$|A_2 \cap A_3| \geq |A_1 \cap A_2 \cap A_3|$$

$$则: |A_1 \cup A_2 \cup A_3| \leq 32\text{-}3\text{-}3\text{-}3\text{=}23$$

即最多有**23**辆汽车至少装配有一个设备，因此至少有**7**辆车没有提供任何设备。

对于包含排斥原理，可推广到n维的情况。

解：设A_1，$A_2,\cdots A_n$为有限集合，其元素个数分别为：

$|A_1|,|A_2|,\cdots|A_n|$，则：

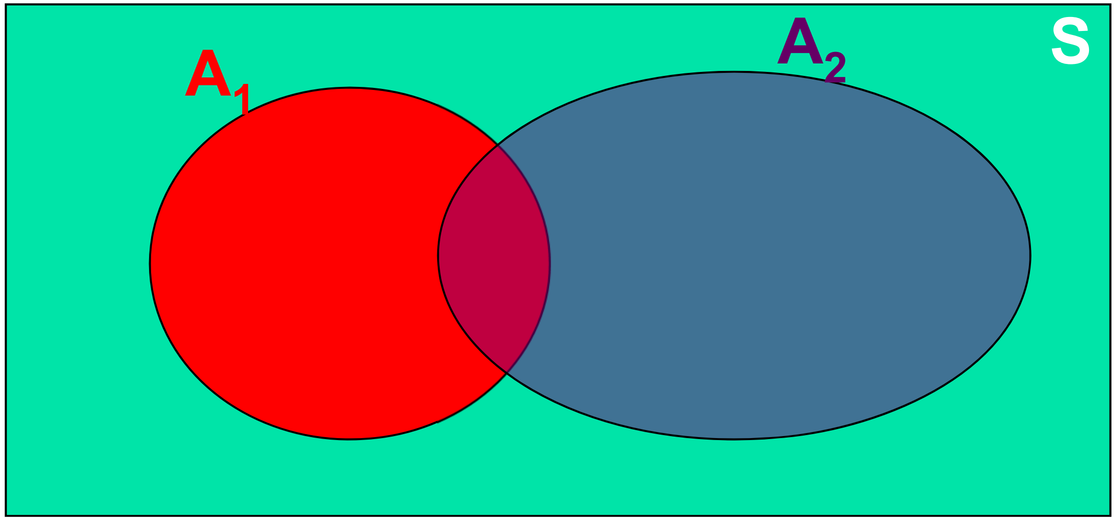
$$|A_1\cup A_2\cup\cdots\cup A_n|=\sum_{i=1}^{n}|A_i|-\sum_{1\le i<j\le n}|A_i\cap A_j|$$

$$+\sum_{1\le i<j<k\le n}|A_i\cap A_j\cap A_k|$$

$$+\ldots+(-1)^{n-1}|A_1\cap A_2\cap\cdots\cap A_n|$$

包含排斥原理

等价形式：若S是有穷集，P_1和P_2分别表示两种性质，对于S中的任何一个元素x，只能有下列四种情况之一：只具有性质P_1，只具有性质P_2，同时具有P_1和P_2两种性质，两种性质都不具有。令：A_1和A_2分别表示具有P_1和P_2元素的集合，用文氏图不难得出：



$$\left|\overline{A_1} \cap \overline{A_2}\right| = |S| - (|A_1| + |A_2|) + |A_1 \cap A_2|$$

定理1.6 S中不具有性质P_1，$P_2, \cdots P_m$的元素是：

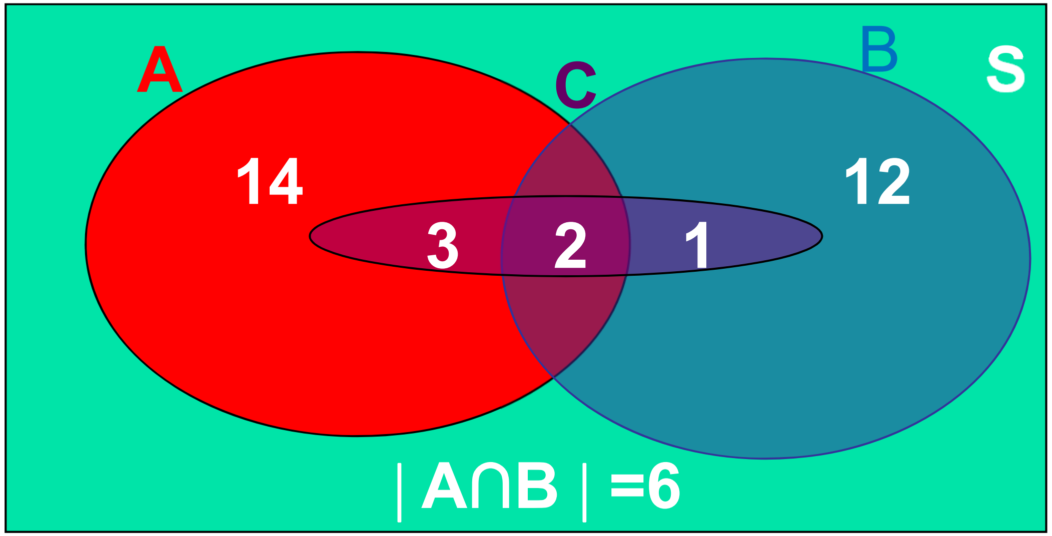
$$\left|\overline{A_1} \cap \overline{A_1} \cap \cdots \cap \overline{A_m}\right| = |S| - \sum_{i=1}^{m} |A_i| + \sum_{1 \le i < j \le m} |A_i \cap A_j|$$

$$- \sum_{1 \le i < j < k \le m} |A_i \cap A_j \cap A_k| + \cdots + (-1)^m |A_1 \cap A_2 \cap \cdots \cap A_m|$$

利用$m = 2$的结果，可用归纳法进行证明。

例1.14某班有25个学生，其中有14人会打篮球，12人会打排球，6人会打篮球和排球，5人会打篮球和网球，还有2人会打这三种球，已知6个会打网球的人都会打篮球或排球，求不会打球的人数。

解：A：会打篮球；B：会打排球；C：会打网球。



$$|\bar{A}\cap\bar{B}\cap\bar{C}| = |S| - (|A|+|B|+|C|)$$

$$+|A\cap B|+|A\cap C|+|B\cap C|$$

$$-|A\cap B\cap C|$$

$$=25-(14+12+6)+(6+5+3)-2$$

$$=5$$

1.2 概率空间

首先，回顾初等概率论的一些基本概念：

- 随机试验 $\tilde{E}$ ，满足如下条件：

1. 在相同条件下可重复进行；
2. 一次试验结果的随机性——不可预知性；
3. 全体可能结果的可知性。

1.2 概率空间

- 样本空间Ω——随机试验所有可能出现的结果组成的集合。
- 样本点ω——Ω中的元素。
- 随机事件——样本空间Ω 的子集合，称为事件？。
- 基本事件——Ω中每个样本点所构成的单点集。
- 必然事件——Ω本身。
- 不可能事件——不包含任何元素的空集合Φ。

跳转

1.2 概率空间

在初等概率论中，我们定义随机事件**A**为样本空间Ω的子集，即 $A \subset \Omega$，但事实上是不是任何一个样本空间Ω的子集合都是一个随机事件？（举例说明）

例1.16 **Ω={1,2,3,4,5,6}** 取$\mathcal{F}$= $\{\Phi, A, \overline{A}, \Omega\}$

容易验证，$\mathcal{F}$为随机事件域，$\because \Phi, A, \overline{A}, \Omega$ 均为随机事件

其中：不妨取**A={1,2}**，$\overline{\mathbf{A}}$ **={3,4,5,6}**

考虑Ω的基本事件**B**={1}，虽然**B**$\subseteq$Ω，但此时**B**并不是我们所讨论问题的随机事件（见张朝金著《概率中的反例》）<u>返回</u>

1.2 概率空间

若把**P(A)**看作集合**A**的函数，那么像高等数学里的普通函数一样，我们必须考虑**A**在何范围内，$\mathbf{A} \to \mathbf{P(A)}$才有定义？这是初等概率论的遗留问题。为此，我们考虑以事件**A**为元素的集合类或事件域$\mathcal{F}$。

$\mathcal{F}$的结构？在$\mathcal{F}$上的概率如何构造？

1.2 概率空间

问题 究竟如何定义随机事件？如果任意的$B \subset \Omega$ ，**B**即为随机事件，概率的定义存在什么问题（第一讲已经提到过，下面举例说明）？

3. 若$\mathbf{A}_i \in \mathcal{F}(i=1,2,\cdots)$，且$\mathbf{A}_i \bigcap \mathbf{A}_j = \Phi(i \neq j)$，则：

$$P\left(\bigcup_{k=1}^{\infty} \mathbf{A}_k\right) = \sum_{k=1}^{\infty} P(\mathbf{A}_k)$$（可列可加性）

称$P(\mathbf{A})$为事件**A**的概率，**(Ω,$\mathcal{F}$,P)**称为概率空间。

问题1：可列求和后不一定封闭！

问题2：求和与求P函数，不一定可以交换顺序。

反例1：可列求和后不一定封闭！

假定集合类：

$$\mathcal{F}=\left\{A_n=\left(a+\frac{1}{n},b\right]\middle|n=1,2,\cdots\right\}$$

$\forall n$，有：$A_n\in\mathcal{F}$

但：$\bigcup_{n=1}^{\infty}A_n=\lim_{N\to\infty}\bigcup_{n=1}^{N}A_n=\lim_{N\to\infty}A_N=\lim_{N\to\infty}\left(a+\frac{1}{N},b\right]$

而：$A_N=\left(a+\frac{1}{N},b\right]\uparrow$（单调上升）

因此，$\bigcup_{n=1}^{\infty}A_n=\lim_{N\to\infty}\left(a+\frac{1}{N},b\right]=(a,b]\notin\mathcal{F}$

解决办法：$\mathcal{F}$ 必须是σ-代数！

1.2 概率空间

反例2:求和与求P函数，不一定可以交换顺序。

事件域：$\mathcal{F}=\left\{A_n=\left(2,3-\frac{1}{n}\right]\middle|n=1,2,\cdots;(2,3)\right\}$

$\because A_1=\Phi$，$A_n=\left(2,3-\frac{1}{n}\right]\Big|n=2,3,\cdots$ 且$A_n\uparrow$

令：$B_1=\Phi$，$B_2=A_2-A_1=A_2, B_n=A_n-A_{n-1}=\left(3-\frac{1}{n-1},3-\frac{1}{n}\right]\Big|n=3,4,\cdots$

$\therefore B_1,B_2,\cdots B_n,\cdots$ 两两互不相容

且:$\bigcup\limits_{n=1}^{\infty}B_n=\bigcup\limits_{n=1}^{\infty}A_n=(2,3)$ 把集合序列的P函数值定义为:

$\forall n$，有$A_n\in\mathcal{F}$，$P(A_n)$ 集合构成区间长度的向上取整函数

1.2 概率空间

$$\therefore P\left(\bigcup_{n=1}^{\infty} B_n\right) = P\left(\bigcup_{n=1}^{\infty} A_n\right) = P\left(\lim_{N\to\infty}\bigcup_{n=1}^{N} A_N\right) = P\left\{\lim_{N\to\infty}\left(2,3-\frac{1}{N}\right]\right\}$$

$$= P\{(2,3)\} = 1$$

另：$P(B_1) = 0,$

$$P(B_n)=P(A_n \setminus A_{n-1})=P\left\{\left(3-\frac{1}{n-1},3-\frac{1}{n}\right]\right\}=\mathrm{P}\left(\frac{1}{n-1}-\frac{1}{n}\right)=1(n\geq 2)$$

$$\sum_{n=1}^{\infty} P(B_n) = \sum_{n=2}^{\infty} 1 = +\infty$$

即 $:1 = P\left(\bigcup_{k=1}^{\infty} B_k\right) \neq \sum_{k=1}^{\infty} P(B_k) = +\infty$ σ-可加性不满足！！

解决办法： P必须具有σ-可加性，为测度！

1.2 概率空间

定义1.16 设Ω是任一非空集合，$\mathcal{F}$是由Ω的一些子集组成的非空集合类，若$\mathcal{F}$满足：

1. $\Omega \in \mathcal{F}$
2. 若$\mathbf{A} \in \mathcal{F}$，有 $\overline{\mathbf{A}} \in \mathcal{F}$（补运算封闭）；
3. 若 $\mathbf{A}_k \in \mathcal{F}$ $(k = 1,2,\cdots)$ ，有$\bigcup_{k=1}^{\infty} \mathbf{A}_k \in \mathcal{F}$（可列并运算封闭）

则称$\mathcal{F}$是Ω上的一个σ-代数，$(\Omega, \mathcal{F})$称为可测空间。若Ω 是一个随机试验所对应的样本空间，则$\mathcal{F}$称为随机事件域，$\mathcal{F}$中的元素A称为随机事件。

显然：最简单的σ-代数为：$\mathcal{F} = \{\Phi, \Omega\}$

次简单的σ-代数为：$\mathcal{F} = \{\Phi, A, \overline{A}, \Omega\}$

1.2 概率空间

显然，有：

4. $\Phi \in \mathscr{F}$; $\left(\because \Omega \in \mathscr{F} \therefore \Phi = \overline{\Omega} \in \mathscr{F}\right)$

5. 若A,B∈ $\mathscr{F}$，有A\B∈ $\mathscr{F}$（差运算封闭）；

$$\left(\because A \cap B = \overline{\overline{A \cap B}} = \overline{\overline{A} \cup \overline{B}} \in \mathscr{F} \therefore A\text{-}B = A\overline{B} \in \mathscr{F}\right)$$

6. 若 $\mathbf{A}_i \in \mathscr{F}(i = 1,2,\cdots)$，有 $\bigcup_{i=1}^{n}\mathbf{A}_i, \bigcap_{i=1}^{n}\mathbf{A}_i, \bigcap_{i=1}^{\infty}\mathbf{A}_i \in \mathscr{F}$

（有限并、有限交、可列交运算封闭）

$$\left(\text{令} A_{n+1} = A_{n+2} = \cdots = \Phi, \therefore \bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{\infty} A_i \in \mathscr{F}\right)$$

1.2 概率空间

定义 设$(\Omega,\mathscr{F})$为可测空间，称υ是$\mathscr{F}$上的测度，则称$(\Omega,\mathscr{F},\upsilon)$为测度空间，若$\upsilon$满足：

(1)$\upsilon \geq 0$

(2)υ满足σ-可加性，即：$A_1, A_2 \cdots \in \mathscr{F}$，且$A_i \cap A_j = \Phi$，$(i \neq j$,其中$i, j = 1,2,\cdots)$，即：$\upsilon\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \upsilon(A_i)$

定义1.18(概率空间) 设$(\Omega,\mathscr{F})$为可测空间，P是$\mathscr{F}$上的概率测度，则称$(\Omega,\mathscr{F},P)$为概率空间；若$\mathbf{A} \in \mathscr{F}$，则称$\mathbf{A}$是$(\Omega,\mathscr{F},P)$为的随机事件或事件。

1.2 概率空间

定义1.17 若$(\Omega,\mathscr{F})$称为可测空间， P(A)是定义在$\mathscr{F}$上的实函数，如果：

1. $\forall \mathbf{A} \in \mathscr{F}, 0 \le P(\mathbf{A}) \le 1$（非负性）

2. $P(\Omega) = 1$ （归一性）

3. 若$\mathbf{A}_i \in \mathscr{F}(i = 1,2,\cdots)$，且$\mathbf{A}_i \bigcap \mathbf{A}_j = \Phi(i \ne j)$，则：

$$P\left(\bigcup_{k=1}^{\infty} \mathbf{A}_k\right) = \sum_{k=1}^{\infty} P(\mathbf{A}_k)$$（ σ-可加性）

称$P(\mathbf{A})$为事件A的概率，$(\Omega,\mathscr{F},\mathrm{P})$称为概率空间。

1.2 概率空间

显然，有：

4. P(Φ)=0；$\left(\because \Phi=\Phi\cup\Phi \therefore P(\Phi)=P(\Phi)+P(\Phi) \therefore P(\Phi)=0\right)$

5. 若$\mathbf{A}_i \in \mathcal{F}\,(i=1,\cdots n)$，且$\mathbf{A}_i \bigcap \mathbf{A}_j = \Phi\,(i \neq j)$，则：

$$P\left(\bigcup_{k=1}^{n} \mathbf{A}_k\right) = \sum_{k=1}^{n} P(\mathbf{A}_k)$$ （有限可加性）

6. 若A,B∈ $\mathcal{F}$，A⊂B, 则：P(B\A)= P(B)-P(A)（概率的可减性）

$\left(\because B=A\cup(B/A) \therefore P(B)=P(A)+P(B/A)\right)$

7. P(A∪B)= P(A)+P(B)- P(AB)（加法定理，待证明）

8. 一般地，有多除少补原理：若$\mathbf{A}_i \in \mathcal{F}(i=1,2,\cdots n)$，则：

$$P\left(\bigcup_{i=1}^{n}\mathbf{A}_i\right)=S_1-S_2+S_3-S_4+\cdots+(-1)^{n+1}S_n$$

其中$S_1=\sum_{i=1}^{n}P(\mathbf{A}_i)$

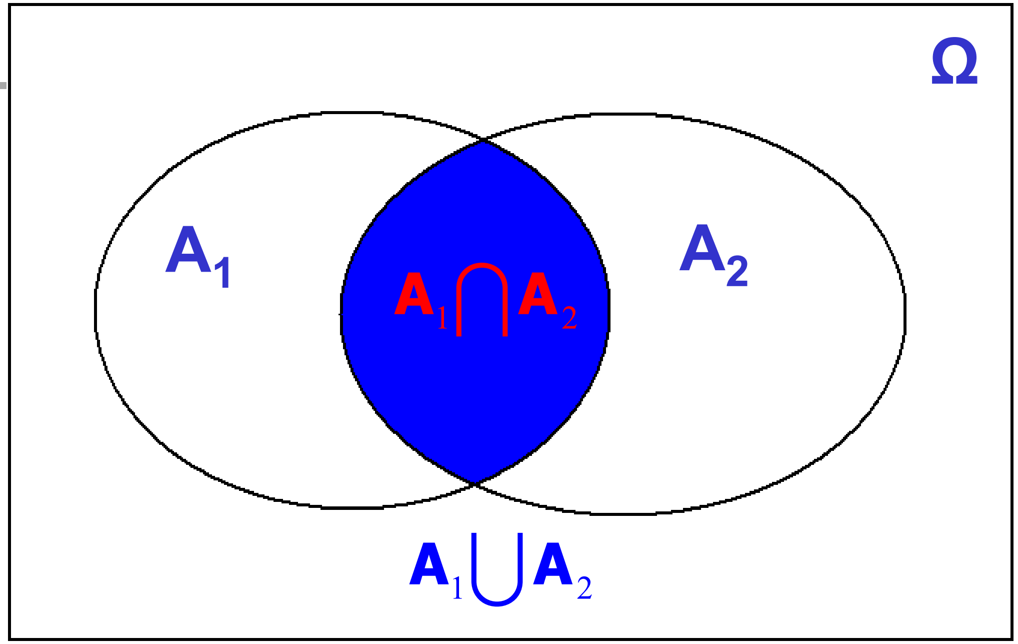
$$S_2=\sum_{1\le i<j\le n}P\left(\mathbf{A}_i\cap\mathbf{A}_j\right)$$

$$S_3=\sum_{1\le i<j<k\le n}P\left(\mathbf{A}_i\cap\mathbf{A}_j\cap\mathbf{A}_k\right)$$

....

$$S_n=P\left(\mathbf{A}_1\cap\mathbf{A}_2\cap\cdots\mathbf{A}_n\right)$$

多除少补原理



证明：用数学归纳法，若$\mathbf{A}_1$，$\mathbf{A}_2 \in \mathcal{F}$

$$\mathbf{A}_1 \bigcup \mathbf{A}_2 = \mathbf{A}_1 \bigcup \left[\mathbf{A}_2 - \left(\mathbf{A}_1 \bigcap \mathbf{A}_2 \right) \right],$$ 且两两互不相容，则：

$$P\left(\mathbf{A}_1 \bigcup \mathbf{A}_2\right) = P\left(\mathbf{A}_1\right) + P\left[\mathbf{A}_2 - \left(\mathbf{A}_1 \bigcap \mathbf{A}_2\right)\right]$$

$$= P\left(\mathbf{A}_1\right) + P\left(\mathbf{A}_2\right) - P\left(\mathbf{A}_1 \bigcap \mathbf{A}_2\right)$$

n=2情形的多除少补原理成立。

设$n=k$成立，即：$P\left(\bigcup_{i=1}^{k+1}\mathbf{A}_i\right)=P\left[\left(\bigcup_{i=1}^{k}\mathbf{A}_i\right)\bigcup\mathbf{A}_{k+1}\right]$

由加法定理，有：$P\left(\bigcup_{i=1}^{k+1}\mathbf{A}_i\right)=P\left(\bigcup_{i=1}^{k}\mathbf{A}_i\right)+P\left(\mathbf{A}_{k+1}\right)-P\left[\left(\bigcup_{i=1}^{k}\mathbf{A}_i\right)\bigcap\mathbf{A}_{k+1}\right]$

$$P\left[\left(\bigcup_{i=1}^{k}\mathbf{A}_i\right)\bigcap\mathbf{A}_{k+1}\right]=P\left[\left(\bigcup_{i=1}^{k}\left(\mathbf{A}_i\bigcap\mathbf{A}_{k+1}\right)\right)\right]=S_1'-S_2'+S_3'-S_4'+\cdots+(-1)^{k+1}S_k'$$

因设$n=k$有：$P\left(\bigcup_{i=1}^{k}\mathbf{A}_i\right)=S_1-S_2+S_3-S_4+\cdots+(-1)^{k+1}S_k$

其中$S_1=\sum_{i=1}^{k}P\left(\mathbf{A}_i\right)$

$$S_2=\sum_{1\le i<j\le k}P\left(\mathbf{A}_i\cap\mathbf{A}_j\right)$$

$$S_3=\sum_{1\le i<j<k\le k}P\left(\mathbf{A}_i\cap\mathbf{A}_j\cap\mathbf{A}_k\right)$$

……………………………

$$S_k=P\left(\mathbf{A}_1\cap\mathbf{A}_2\cap\cdots\mathbf{A}_k\right)$$

$$S_1'=\sum_{i=1}^{k}P\left(\mathbf{A}_i\cap\mathbf{A}_{k+1}\right)$$

$$S_2'=\sum_{1\le i<j\le k}P\left[\left(\mathbf{A}_i\cap\mathbf{A}_{k+1}\right)\cap\left(\mathbf{A}_j\cap\mathbf{A}_{k+1}\right)\right]$$

$$=\sum_{1\le i<j\le k}P\left(\mathbf{A}_i\cap\mathbf{A}_j\cap\mathbf{A}_{k+1}\right)$$

……………………………

$$S_{k-1}'=\sum_{1\le i_1<i_2\cdots<i_{k-1}\le k}P\left(\mathbf{A}_{i_1}\cap\mathbf{A}_{i_2}\cap\cdots\cap\mathbf{A}_{i_{k-1}}\cap\mathbf{A}_{k+1}\right)$$

$$S_k'=P\left(\mathbf{A}_1\cap\mathbf{A}_2\cap\cdots\cap\mathbf{A}_k\cap\mathbf{A}_{k+1}\right)$$

多除少补定理

$$\text{而}S_1 + P\left(\mathbf{A}_{k+1}\right) = \sum_{i=1}^{k+1} P\left(\mathbf{A}_i\right)$$

$$S_2 + S_1^{'} = \sum_{1\le i<j\le k+1} P\left(\mathbf{A}_i \cap \mathbf{A}_j\right)$$

$$S_3 + S_2^{'} = \sum_{1\le i<j<s\le k+1} P\left(\mathbf{A}_i \cap \mathbf{A}_j \cap \mathbf{A}_s\right)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$S_k^{'} = P\left(\mathbf{A}_1 \cap \mathbf{A}_2 \cap \cdots \mathbf{A}_k \cap \mathbf{A}_{k+1}\right)$$

由以上结果可知，结论对$n=k+1$也成立。

9. 半可加性成立：若$\mathbf{A}_i \in \mathcal{F}(i=1,2,\cdots n)$，则：

$$P\left(\bigcup_{i=1}^{n}\mathbf{A}_i\right) \le \sum_{i=1}^{n} P(\mathbf{A}_i)$$

证明：用数学归纳法，$P(\mathbf{A}_1 \bigcup \mathbf{A}_2) = P(\mathbf{A}_1) + P(\mathbf{A}_2) - P(\mathbf{A}_1 \bigcap \mathbf{A}_2)$

$$\le P(\mathbf{A}_1) + P(\mathbf{A}_2)$$

假设$n=k$成立，有：$P\left(\bigcup_{i=1}^{k}\mathbf{A}_i\right) \le \sum_{i=1}^{k} P(\mathbf{A}_i)$

当$n=k+1$，由：$P\left(\bigcup_{i=1}^{k+1}\mathbf{A}_i\right) = P\left(\bigcup_{i=1}^{k}\mathbf{A}_i\right) + P(\mathbf{A}_{k+1}) - P\left[\left(\bigcup_{i=1}^{k}\mathbf{A}_i\right)\bigcap \mathbf{A}_{k+1}\right]$

$$\le P\left(\bigcup_{i=1}^{k}\mathbf{A}_i\right) + P(\mathbf{A}_{k+1}) \le \sum_{i=1}^{k} P(\mathbf{A}_k) + P(\mathbf{A}_{k+1}) = \sum_{i=1}^{k+1} P(\mathbf{A}_k)$$

类似可得次可加性

1.2 概率空间

10. （次可加性）$\mathscr{F}$为集代数， **P**是概率测度， **A**$_i \in \mathscr{F}$

$$i=1,2,\cdots, A\in\mathscr{F}, A\subset\bigcup_{i=1}^{\infty}A_i\in\mathscr{F}，有：$$

$$P(A)\le\sum_{i=1}^{\infty}P(A_i)$$

特别地，当 $A=\bigcup_{i=1}^{\infty}A_i$，有

$$P\left(\bigcup_{i=1}^{\infty}A_i\right)\le\sum_{i=1}^{\infty}P(A_i)$$

1.2 概率空间

11. 若**A⊂B**, 则：　**P(A)⩽P(B)**（概率的单调性）

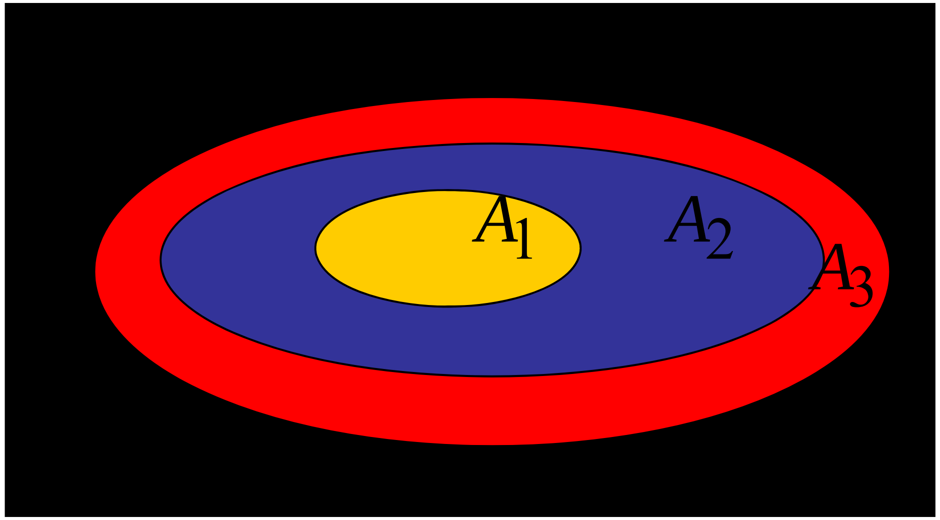
12. 若$A_n \in \mathcal{F}(n=1,2,\cdots)$，则：

$$\lim_{n\to\infty} P(A_n) = P\left(\lim_{n\to\infty} A_n\right) = \begin{cases} P\left(\bigcup_{n=1}^{\infty} A_n\right) & 若A_1 \subset A_2 \subset \cdots \\ P\left(\bigcap_{n=1}^{\infty} A_n\right) & 若A_1 \supset A_2 \supset \cdots \end{cases}$$

概率具有连续性——简单证明！

下面证明P的连续性

证明：假设，若$\mathbf{A}_n \in \mathcal{F}$，有$\mathbf{A}=\bigcup_{n=1}^{\infty}\mathbf{A}_n \in \mathcal{F}$，且$P\left(\mathbf{A}_n\right)$有限如下图：



$$P\left(\mathbf{A}\right)=P\left(\bigcup_{n=1}^{\infty}\mathbf{A}_n\right)=P\left\lfloor \mathbf{A}_1+\left(\mathbf{A}_2\setminus\mathbf{A}_1\right)+\left(\mathbf{A}_3\setminus\mathbf{A}_2\right)+\cdots+\left(\mathbf{A}_n\setminus\mathbf{A}_{n-1}\right)+\cdots\right\rfloor$$

$$=P\left(\mathbf{A}_1\right)+P\left(\mathbf{A}_2\right)-P\left(\mathbf{A}_1\right)+P\left(\mathbf{A}_3\right)-P\left(\mathbf{A}_2\right)+\cdots+P\left(\mathbf{A}_n\right)-P\left(\mathbf{A}_{n-1}\right)+\cdots$$

$$=\lim_{n\leftarrow\infty}P\left(\mathbf{A}_n\right)$$

1.2 概率空间

定义1.19（古典概型） 若随机试验**E**具有下列特征：

1.有限性：样本空间含有限个样本点$\Omega=\{e_1,e_2\cdots e_n\}$

2.等可能性：每一个样本点出现的概率相同，即

$$P(e_1)=P(e_2)=\cdots P(e_n)=\frac{1}{n}$$

则称**E**为古典概型。

设事件**A**包含Ω中的k个样本点，则：

$$P(\mathbf{A})=\frac{k}{n}=\frac{\text{事件}\mathbf{A}\text{包含的样本点数}}{\text{样本点总数}}$$

1.2 概率空间

定理1.7（条件概率） 若$(\Omega,\mathcal{F},P)$为一已知的概率空间，$\mathbf{B}\in\mathcal{F}$，$P(\mathbf{B})>0$，若定义$\mathcal{F}$上的集合函数：

$$P(\mathbf{A}|\mathbf{B})=\frac{P(\mathbf{AB})}{P(\mathbf{B})},\quad \mathbf{A}\in\mathcal{F}，则称$$

$P(\mathbf{A}|\mathbf{B})$是$\mathcal{F}$上的概率测度，且$P(\mathbf{A}|\mathbf{B})$满足：

(1) $0\le P(\mathbf{A}|\mathbf{B})\le 1,\quad \mathbf{A}\in\mathcal{F}$

(2) $P(\Omega|\mathbf{B})=1$

(3) 若$\mathbf{A}_i\in\mathcal{F}(i=1,2,\cdots)$，且$\mathbf{A}_i\bigcap\mathbf{A}_j=\Phi(i\neq j)$，则：

$$P\left(\bigcup_{k=1}^{\infty}\mathbf{A}_k|\mathbf{B}\right)=\sum_{k=1}^{\infty}P(\mathbf{A}_k|\mathbf{B})$$

1.2 概率空间

定理1.8（乘法公式） 设$P(\mathbf{B})>0$，则：

$$P(\mathbf{AB})=P(\mathbf{B})\times P(\mathbf{A}|\mathbf{B})$$

一般地，设$\mathbf{A}_i\in\mathcal{F}(i=1,\cdots n,\ n\geq 2)$，且：

$P(\mathbf{A}_1\cdots\mathbf{A}_{n-1})>0$

则：$P(\mathbf{A}_1\cdots\mathbf{A}_n)$

$=P(\mathbf{A}_1)\times P(\mathbf{A}_2|\mathbf{A}_1)\times P(\mathbf{A}_3|\mathbf{A}_1\cap\mathbf{A}_2)\cdots P(\mathbf{A}_n|\mathbf{A}_1\cdots\mathbf{A}_{n-1})$

用数学归纳法显然容易得证。

1.2 概率空间

定理1.9(全概率公式) 设随机实验**E**的样本空间**Ω**，$\mathbf{A}\in\mathcal{F}$，$\mathbf{B}_1,\mathbf{B}_2,\cdots\mathbf{B}_n$为**Ω**的一个划分，即$\mathbf{B}_i\in\mathcal{F}$，$\mathbf{B}_i\cap\mathbf{B}_j=\mathbf{\Phi}$，$i\neq j$，$i,j=1,\cdots n$，$P(\mathbf{B}_i)>0$，且$\bigcup_{i=1}^{n}\mathbf{B}_i=\Omega$，则有全概率公式：

$$P(\mathbf{A})=\sum_{i=1}^{n}P(\mathbf{B}_i)\times P(\mathbf{A}|\mathbf{B}_i)$$

证明：$P(\mathbf{A})=P(\mathbf{A\Omega})=P\left(\mathbf{A}\bigcup_{i=1}^{n}\mathbf{B}_i\right)=P\left(\bigcup_{i=1}^{n}\mathbf{AB}_i\right)$

$$=\sum_{i=1}^{n}P(\mathbf{AB}_i)=\sum_{i=1}^{n}P(\mathbf{B}_i)\times P(\mathbf{A}|\mathbf{B}_i)$$

1.2 概率空间

定理1.10（*Bays*公式） 若$(\Omega,\mathcal{F},P)$为一已知的概率空间，$\mathbf{B}_i\in\mathcal{F}$且$P(\mathbf{B}_i)>0\ (i,j=1,\cdots n)$，则对任意的$\mathbf{A}\in\mathcal{F}$，且$P(\mathbf{A})>0$，有：

$$P(\mathbf{B}_i|\mathbf{A})=\frac{P(\mathbf{AB}_i)}{P(\mathbf{A})}=\frac{P(\mathbf{B}_i)\times P(\mathbf{A}|\mathbf{B}_i)}{\sum_{i=1}^{n}P(\mathbf{B}_i)\times P(\mathbf{A}|\mathbf{B}_i)}(i,j=1,\cdots n)$$

1.2 概率空间

定义1.20(事件的独立性) 若$(\Omega,\mathscr{F},P)$为一已知的概率空间，对任意的**A**，$\mathbf{B}\in\mathscr{F}$且$P(\mathbf{A})>0$，若：

$$P(\mathbf{B}|\mathbf{A})=P(\mathbf{B})$$，则称两个随机事件**A**，**B**相互独立。

可以证明：随机事件**A**，**B**相互独立 $\Leftrightarrow P(\mathbf{AB})=P(\mathbf{A})P(\mathbf{B})$

定理1.11(独立性的性质) 若随机事件**A**，**B**相互独立，则

$\mathbf{A}$与$\overline{\mathbf{B}}$，$\mathbf{B}$与$\overline{\mathbf{A}}$，$\overline{\mathbf{A}}$与$\overline{\mathbf{B}}$也相互独立。

$$\mathbf{A}=\mathbf{A}\cap\boldsymbol{\Omega}=\mathbf{A}\cap\left(\mathbf{B}\cup\overline{\mathbf{B}}\right)=(\mathbf{A}\cap\mathbf{B})\cup\left(\mathbf{A}\cap\overline{\mathbf{B}}\right)$$

$$P(\mathbf{A})=P\left[(\mathbf{A}\cap\mathbf{B})\cup\left(\mathbf{A}\cap\overline{\mathbf{B}}\right)\right]=P(\mathbf{A}\cap\mathbf{B})+P\left(\mathbf{A}\cap\overline{\mathbf{B}}\right)$$，则：

$$P\left(\mathbf{A}\cap\overline{\mathbf{B}}\right)=P(\mathbf{A})-P(\mathbf{A}\cap\mathbf{B})=P(\mathbf{A})-P(\mathbf{A})P(\mathbf{B})=P(\mathbf{A})P\left(\overline{\mathbf{B}}\right)$$

1.2 概率空间

定义1.21（n个事件的独立性） 设$(\Omega,\mathcal{F},P)$为一已知的概率空间，若$\mathbf{A}_1$，$\mathbf{A}_2,\cdots\mathbf{A}_n\in\mathcal{F}$且对任意的$S\,(1<S\le n)$及任意$i_k$，$k$=1,2,$\cdots n$，$1<i_1<i_2<\cdots<i_s\le n$，有：

$$P\left(\mathbf{A}_{i_1}\cap\mathbf{A}_{i_2}\cap\cdots\mathbf{A}_{i_s}\right)=P\left(\mathbf{A}_{i_1}\right)P\left(\mathbf{A}_{i_2}\right)\cdots P\left(\mathbf{A}_{i_s}\right)$$

，则称随机事件$\mathbf{A}_1,\mathbf{A}_2,\cdots\mathbf{A}_n$相互独立。

以上定义包含2^n-1-n个恒等式：

$$C_n^2+C_n^3+\cdots C_n^n=\sum_{i=0}^{n}C_n^i-C_n^0-C_n^1$$

$$=(1+1)^n-1-n=2^n-1-n$$

1.2 概率空间

- 增加一个实际的例子：见王梓坤《概率论基础及其应用》，**P30**例**5**，小概率事件绝不可忽视！

1.3 随机变量及其概率分布

1.3.1随机变量的概念

1、一维随机变量

定义1.22设(Ω,$\mathcal{F}$,P)是概率空间，**X**=**X**(e)是定义在Ω上的实函数，如果对任意实数x，$\{e: X(e)\le x\}\in\mathcal{F}$，则称**X**($e$)是$\mathcal{F}$上的随机变量，简记为随机变量**X**。

注意上述定义式可改写为：$\{e: X(e)\in(-\infty, x]\}$

对随机变量主要需要关注两个问题：

- **X=X(e)**的值域
- 取各个值的概率规律

1.3 随机变量及其概率分布

2、二维随机变量

定义1.23设$(\Omega,\mathcal{F},P)$是概率空间，$(\mathbf{X},\mathbf{Y})=(\mathbf{X}(e),\mathbf{Y}(e))$是定义在Ω上的二维空间中取值的实向量函数，如果对任意实数，$(x,y)\in R^2,\{e:\mathbf{X}(e)\le x,\mathbf{Y}(e)\le y\}\in\mathcal{F}$则$(\mathbf{X}(e),\mathbf{Y}(e))$是$\mathcal{F}$上的二维随机变量。

1.3 随机变量及其概率分布

1.3.2离散型随机变量的概率分布

设离散型随机变量**X**所有可能的取值为 $x_k\,(k=1,2,\cdots)$ 且取各可能值的概率为：

$$P(\mathbf{X}=x_k)=p_k\,(k=1,2,\cdots) \qquad (1.2.1)$$

其中 p_k 满足：

（1） $p_k \geq 0(k=1,2,\cdots)$

（2） $\sum_{k=1}^{\infty} p_k = 1$

则称（1.2.1）式为离散型随机变量**X**的概率分布或分布律。

1.3 随机变量及其概率分布

离散型随机变量X的分布律也可记为下列形式：

X	x_1	x_2		x_k	
P	p_1	p_2		p_k	

1.3 随机变量及其概率分布

1.3.3常见离散型随机变量的概率分布

1、贝努利试验，二项分布

试验E重复进行n次，若每次试验结果相互独立，且只有两种结果A及$\overline{A}$，且$P(A)=p$，$P(\overline{A})=1-p=q(0\le p\le 1)$则称这一串独立重复试验E为$n$重贝努利试验，以X表示$n$重贝努利试验中事件A发生的次数，则X可能的取值为$0,1,2,\cdots n$，它的分布律为：

$$P(\mathbf{X}=k)=C_n^k p^k q^{n-k}\left(k=0,1,\cdots n\right)$$

称X服从参数为n，p的二项分布，记为：$\mathbf{X}\sim B(n,p)$

1.3 随机变量及其概率分布

2、泊松分布和泊松定理

在实际问题的分析中，在$[t_0, t_0+t)$某电话交换台来到的呼唤次数，某商店进入的顾客数，放射性物质不断放出的粒子数等，都满足如下条件：

1.平稳性 在$[t_0, t_0+t)$来到的电话呼唤次数k的概率p，只与时间区间的长度t有关，与起点t_0无关，即

$$p\{\mathbf{X}(t_0+t)-\mathbf{X}(t_0)=k\}=p_k(t),\quad k=0,1,2,\cdots$$

2.独立增强性 在n个不相重叠的区间$t_0\le t_1\le t_2\le\cdots t_{n-1}\le t_n$，各自来到的呼唤次数是相互独立的，即$\mathbf{X}(t_1)-\mathbf{X}(t_0)$，$\mathbf{X}(t_2)-\mathbf{X}(t_1),\cdots\mathbf{X}(t_n)-\mathbf{X}(t_{n-1})$是相互独立的。

1.3 随机变量及其概率分布

3.普通性 在足够小的时间间隔内最多来到一个呼唤。若记

$$\psi(t)=1-p_0(t)-p_1(t),$$

则：$\lim\limits_{t\to 0}\dfrac{\psi(t)}{t}=0$或$\psi(t)=o(t)$

4.非平凡性 $p_0(t)$不恒等于1，$\sum\limits_{k=0}^{\infty}p_k(t)=1$

满足以上条件的物理现象称为泊松过程。

单位时间内接收到得呼叫次数**X**服从于参数为λ的泊松分布，其中的λ为单位时间内呼叫强度或到达率，则：$P(\mathbf{X}=k)=\dfrac{e^{-\lambda}\lambda^k}{k!}(k=0,1,2,\cdots)(\lambda>0)$

$[t_0,t_0+t)$内接收到得呼叫次数**X**服从于参数为λt的泊松分布，

$$P(\mathbf{X}=k)=\frac{e^{-\lambda t}(\lambda t)^k}{k!}(k=0,1,2,\cdots)(\lambda>0)$$

泊松定理设λ为常数，n为正整数，记λ=np_n，则对任意固定的非负整数k，有：

$$\lim_{n\to\infty} C_n^k p_n^k \left(1-p_n^k\right)^{n-k} = \frac{e^{-\lambda}\lambda^k}{k!}(k=0,1,2,\cdots)(\lambda>0)$$

当n很大p_n很小时，二项分布的极限分布为泊松分布。

利用泊松定理，既可以用二项分布逼近泊松分布，又可用泊松分布近似具有较大n的二项分布。

如：当n很大，对给定的p，λ=np，要精确地算出**B(***n, p***)**是较困难的；

因此根据泊松定理，可以取值$\dfrac{e^{-np}(np)^k}{k!}$，而泊松分布的值有表可查。

如：二项分布p_3**(**800,0.005**)=**$C_{800}^3 0.005^3(1-0.005)^{800-3}$，其精确值为0.1945；因$np$**=**$800\times0.005$**=**4，得近似值为：$\dfrac{e^{-4}4^3}{3\,!}$**=**0.1954。

泊松分布常常用于研究稀有事件的概率，这类事件p很小，但n很大

1.3 随机变量及其概率分布

1.3.4随机变量的分布函数

定义1.24设$(\Omega,\mathscr{F},P)$是概率空间，$\mathbf{X}=\mathbf{X}(e)$是定义在$\mathbf{\Omega}$上的随机变量$\left(\forall\text{实数}x，\text{有}\{e:\ \mathbf{X}(e)\le x\}\in\mathscr{F}\right)$，称函数：

$$F(x)=P\{\mathbf{X}(e)\le x\}$$ 是随机变量$\mathbf{X}$的分布函数。

分布函数有如下性质：

1. $F(x)$单调不降
2. $F(x)$右连续，即$F(x+0)=F(x)$
3. $0\le F(x)\le 1$且$F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0$，$F(\infty)=\lim\limits_{x\to\infty}F(x)=1$

特别地：对离散型随机变量，有：$F(x)=\sum\limits_{x_k\le x}P(\mathbf{X}=x_k)$

仅证明性质2：欲证 $\lim\limits_{n\to\infty} F(x+\frac{1}{n}) = F(x+0) = F(x)$，其余略

证明：$\because \left\{e:\mathbf{X}(e)\le x+\frac{1}{n}\right\} \supseteq \left\{e:\mathbf{X}(e)\le x+\frac{1}{n+1}\right\} \therefore \left\{e:\mathbf{X}(e)\le x+\frac{1}{n}\right\}\downarrow$

且 $\bigcap\limits_{n\to\infty}\left\{e:\mathbf{X}(e)\le x+\frac{1}{n}\right\} = \lim\limits_{n\to\infty}\left\{e:\mathbf{X}(e)\le x+\frac{1}{n}\right\} = \{e:\mathbf{X}(e)\le x\}$

另，由概率的单调性有：$P\left\{e:\mathbf{X}(e)\le x+\frac{1}{n}\right\} \ge P\left\{e:\mathbf{X}(e)\le x+\frac{1}{n+1}\right\}$

且由概率的连续性有：$\lim\limits_{n\to\infty} P\left\{e:\mathbf{X}(e)\le x+\frac{1}{n}\right\} = P\left\{\lim\limits_{n\to\infty}\left[e:\mathbf{X}(e)\le x+\frac{1}{n}\right]\right\}$

$$= P\{e:\mathbf{X}(e)\le x\} = F(x)$$

即：$\lim\limits_{n\to\infty} F(x+\frac{1}{n}) = F(x+0) = F(x)$

1.3 随机变量及其概率分布

1.3.5连续型随机变量的概率密度

定义1.25对随机变量**X**的分布函数$F(x)$，若存在非负函数$f(x)$，使得对任意的$x \in \mathbf{R}$，有：

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

称**X**为连续型随机变量，$f(x)$为**X**的概率密度。

1.3 随机变量及其概率分布

$f(x)$有如下性质：

1. $f(x) \geq 0$

2. $\int_{-\infty}^{+\infty} f(x)dx = 1$

3. $P(x_1 < x \leq x_2) = \int_{x_1}^{x_2} f(t)dt \,(x_1 \leq x_2)$

4. 当x为$f(x)$的连续点，有：$F'(x) = f(x)$

1.3 随机变量及其概率分布

1.3.6常见连续型随机变量的概率密度

1.均匀分布：$f(x)=\begin{cases}\frac{1}{b-a} & a<x\le b\\ 0 & \text{其它}\end{cases}$

2.正态分布

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}\quad(-\infty<x<+\infty)$$

则称**X**服从参数为$a,\ \sigma^2$的正态分布，记：$\mathbf{X}\sim N\left(a,\ \sigma^2\right)$

当a=0，σ=1，称**X**服从标准正态分布，记：$\mathbf{X}\sim N(0,\ 1)$

2.正态分布

(1)产生正态分布的实际背景

中间大两头小的一些变量，具有特点：可以看成为许多微小的，独立的随机因素的总和!(数学描述由中心极限定理完成!)

(2)定义：已知两个常数a, σ，定义：$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}\quad(-\infty < x < +\infty)$

则称**X**服从参数为a, σ^2的正态分布，记：$\mathbf{X} \sim N\left(a, \sigma^2\right)$

特别地，当$a=0$，$\sigma=1$，称**X**服从标准正态分布，记：$\mathbf{X} \sim N(0, 1)$

(3)正态分布的分布密度函数$f(x)$具有如下性质：

1°处处大于0，且具有各级连续的导函数。

2° $f(x)$在$(-\infty, a)$严格上升，在a处取极大值$\frac{1}{\sigma\sqrt{2\pi}}$，在$(a,+\infty)$严格下降；

当$x \to +\infty$或$x \to -\infty$时，$f(x) \to 0$。

3° $f(x)$关于a点对称。因此：a表示大部分值聚集的点，σ^2表示聚集的程度。

1.3 随机变量及其概率分布

1.3.7二维离散型随机变量的联合分布和边缘分布

定义1.26设二维离散型随机变量$(\mathbf{X},\mathbf{Y})$所有可能的取值为$(x_i, y_j)(i,j=1,2,\cdots)$，且：

$$P(\mathbf{X}=x_i, \mathbf{Y}=y_j)=p_{ij}(i,j=1,2,\cdots)$$

其中$p_{ij}\geq 0$且$\sum\limits_{i,j} p_{ij}=1(i,j=1,2,\cdots)$,称上式为二维离散型随机变量$(\mathbf{X},\mathbf{Y})$的概率分布，或随机变量**X**与随机变量**Y**的联合分布律。

1.3 随机变量及其概率分布

1.3.7二维离散型随机变量的联合分布和边缘分布

定义1.27设二维离散型随机变量$(\mathbf{X},\mathbf{Y})$的联合分布律为：

$$P(\mathbf{X}=x_i,\mathbf{Y}=y_j)=p_{ij}(i,j=1,2,\cdots)$$

其中$p_{ij}\geq 0$且$\sum\limits_{i,j}p_{ij}=1(i,j=1,2,\cdots)$,则分别称：

$$p(i,\cdot)=\sum_{j=1}^{\infty}p_{ij}=P(\mathbf{X}=x_i)(i=1,2,\cdots)$$

$$p(\cdot,j)=\sum_{i=1}^{\infty}p_{ij}=P(\mathbf{Y}=y_j)(j=1,2,\cdots)$$

为随机变量$(\mathbf{X},\mathbf{Y})$关于$\mathbf{X}$和关于$\mathbf{Y}$的边缘分布律。

1.3 随机变量及其概率分布

1.3.7二维离散型随机变量的联合分布和边缘分布

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$	$p(i,\cdot)$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$	$\sum_{j=1}^{\infty} p_{1j}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$	$\sum_{j=1}^{\infty} p_{2j}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$	$\sum_{j=1}^{\infty} p_{ij}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$p(\cdot,j)$	$\sum_{i=1}^{\infty} p_{i1}$	$\sum_{i=1}^{\infty} p_{i2}$	$\cdots$	$\sum_{i=1}^{\infty} p_{ij}$	$\cdots$	$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} p_{ij}=1$

1.3.8联合分布函数和边缘分布函数

定义1.28称二元函数$F(x,y)=P\{\mathbf{X}\le x,\mathbf{Y}\le y\}$为二维随机变量$(\mathbf{X},\mathbf{Y})$的分布函数，或随机变量**X**与随机变量**Y**的联合分布分布函数。

$F(x,y)$有如下性质：

1. $F(x,y)$分别关于x，y单调不减

2. $0\le F(x,y)\le 1$，且：

$$F(-\infty,y)=P\{\mathbf{X}\le -\infty,\mathbf{Y}\le y\}=P\left\{\lim_{x\to-\infty}[\mathbf{X}\le x,\mathbf{Y}\le y]\right\}$$

$$=0=\lim_{x\to-\infty}P\{\mathbf{X}\le x,\mathbf{Y}\le y\}=\lim_{x\to-\infty}F(x,y)$$

$$F(x,-\infty)=\lim_{y\to-\infty}F(x,y)=0$$

$$F(-\infty,-\infty)=\lim_{x,y\to-\infty}F(x,y)=0$$

$$F(+\infty,+\infty)=\underset{x,y\to+\infty}{lm}F(x,y)=1$$

1.3 随机变量及其概率分布

1.3.8联合分布函数和边缘分布函数

3.$F(x,y)$分别关于x，y右连续，即：

$$F(x+0,y)=F(x,y),F(x,y+0)=F(x,y)$$

4.$a_1<b_1$，$a_2<b_2$时，有：

$$\Delta F=F(b_1,b_2)-F(b_1,a_2)-F(a_1,b_2)+F(a_1,a_2)\geq 0$$

证明：$F(b_1,b_2)-F(b_1,a_2)-F(a_1,b_2)+F(a_1,a_2)$

$$=\left[P\{\mathbf{X}\leq b_1,\mathbf{Y}\leq b_2\}-P\{\mathbf{X}\leq b_1,\mathbf{Y}\leq a_2\}\right]-\left[P\{\mathbf{X}\leq a_1,\mathbf{Y}\leq b_2\}-P\{\mathbf{X}\leq a_1,\mathbf{Y}\leq a_2\}\right]$$

$$=P\{\mathbf{X}\leq b_1,a_2\leq\mathbf{Y}\leq b_2\}-P\{\mathbf{X}\leq a_1,a_2\leq\mathbf{Y}\leq b_2\}$$

$$=P\{a_1\leq\mathbf{X}\leq b_1,a_2\leq\mathbf{Y}\leq b_2\}\geq 0,$$ 证毕。

1.3 随机变量及其概率分布

1.3.8联合分布函数和边缘分布函数

定义1.28设二维离散型随机变量$(\mathbf{X},\mathbf{Y})$的分布函数为：

$$F_{\mathbf{X}}(x)=\lim_{y\to+\infty}F(x,y)=F(x,+\infty),\quad F_{\mathbf{Y}}(y)=\lim_{x\to+\infty}F(x,y)=F(+\infty,y)$$

分别为$(\mathbf{X},\mathbf{Y})$关于$\mathbf{X}$和关于$\mathbf{Y}$的边缘分布函数。

简单分析：

$$F_{\mathbf{X}}(x)=\mathbf{P}\left(X\le x\right)=\mathbf{P}\left(X\le x,Y\le+\infty\right)=\mathbf{P}\left\{\lim_{y\to+\infty}\left(X\le x,Y\le y\right)\right\}$$

$$=\lim_{y\to+\infty}\mathbf{P}\left(X\le x,Y\le y\right)=\lim_{y\to+\infty}F(x,y)=F(x,+\infty)$$

该结论可被描述为：边缘分布函数是联合分布函数的极限。

甚至被推广为：低维分布函数是高维分布函数的极限！！

1.2 随机变量及其概率分布

1.2.9联合概率密度和边缘概率密度

定义1.29对随机变量$(\mathbf{X},\mathbf{Y})$的分布函数$F(x,y)$，存在非负函数$f(x,y)$，使得对任意的$x,y\in\mathbf{R}$，有：

$$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x}f(u,v)dudv$$

称$(\mathbf{X},\mathbf{Y})$为二维连续型随机变量，$f(x,y)$为$(\mathbf{X},\mathbf{Y})$的概率密度，或随机变量$\mathbf{X}$和随机变量$\mathbf{Y}$的联合概率密度。

1.2 随机变量及其概率分布

1.2.9联合概率密度和边缘概率密度

$f(x,y)$有如下性质：

$$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)dudv$$

1. $f(x,y)\geq 0$

2. $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = F(+\infty,+\infty)=1$

3. $f(x,y)$在(x,y)连续，则：$\dfrac{\partial^2 F(x,y)}{\partial x\partial y} = f(x,y)$

4. 随机点(x,y)落在平面区域G内的概率

$$P\{(\mathbf{X},\mathbf{Y})\in \mathbf{G}\} = \iint_G f(x,y)dxdy$$

1.3.9联合概率密度和边缘概率密度

定义1.30函数$f(x,y)$为$(\mathbf{X},\mathbf{Y})$的概率密度，则称：

$$f_{\mathbf{X}}(x)=\int_{-\infty}^{+\infty} f(x,y)\,dy,\quad f_{\mathbf{Y}}(y)=\int_{-\infty}^{+\infty} f(x,y)\,dx$$

分别成为$(\mathbf{X},\mathbf{Y})$关于$\mathbf{X}$和$\mathbf{Y}$的边缘概率密度。

简单分析：

$$\mathbf{F}_{\mathbf{X}}(x)=P(X\le x)=P(X\le x,Y\le +\infty)=\int_{-\infty}^{x}\left[\int_{-\infty}^{+\infty} f(x,y)\,dy\right]dx$$

因此：$f_{\mathbf{X}}(x)=\mathbf{F}'_{\mathbf{X}}(x)=\int_{-\infty}^{+\infty} f(x,y)\,dy$

1.3 随机变量及其概率分布

1.3.10随机变量的独立性和条件分布

定义1.30设$F(x,y)$，$F_{\mathbf{X}}(x), F_{\mathbf{Y}}(y)$分别为$(\mathbf{X},\mathbf{Y})$的联合分布函数和边缘分布函数，若对任意的$x, y\in\mathbf{R}$，有：

$$F(x,y)=F_{\mathbf{X}}(x)F_{\mathbf{Y}}(y)$$

则称随机变量**X**和随机变量**Y**相互独立。

特别地：若$(\mathbf{X},\mathbf{Y})$为二维离散型随机变量，若有：

$p_{ij}=p(i,\cdot)p(\cdot,j)$则称随机变量**X**和**Y**相互独立；

若$(\mathbf{X},\mathbf{Y})$为二维连续型随机变量，若有：$f(x,y)$
$=f_{\mathbf{X}}(x)f_{\mathbf{Y}}(y)$，也称随机变量**X**和**Y**相互独立。

1.3 随机变量及其概率分布

定理1.12若随机变量**X**和**Y**相互独立，$f(x)$、$g(y)$分别是x和y的连续函数，则：$f(\mathbf{X})$、$g(\mathbf{Y})$不仅是随机变量，且相互独立。

定义1.31设**X**和**Y**为离散型随机变量，p_{ij}、$p(i,\cdot)$和$p(\cdot,j)$分别为$(\mathbf{X},\mathbf{Y})$、**X**和**Y**的概率分布，设$p(i,\cdot)$、$p(\cdot,j)>0$，则称：

$P\left(j/i\right)=P\left\{\mathbf{Y}=y_j/\mathbf{X}=x_i\right\}=\dfrac{p_{ij}}{p(i,\cdot)}$为$\mathbf{X}=x_i$条件下随机变量**Y**的条件分布；同理也称$P\left(i/j\right)=P\left\{\mathbf{X}=x_i/\mathbf{Y}=y_j\right\}=\dfrac{p_{ij}}{p(\cdot,j)}$为$\mathbf{Y}=y_j$条件下随机变量**X**的条件分布。

1.3 随机变量及其概率分布

定义1.32设**X**和**Y**为离散型随机变量，p_{ij}、$p(i,\cdot)$和$p(\cdot,j)$分别为$(\mathbf{X},\mathbf{Y})$、**X**和**Y**的概率分布，设$p(i,\cdot)$、$p(\cdot,j)>0$，则称：

$$F\left(y/x_i\right)=P\left\{\mathbf{Y}\le y/\mathbf{X}=x_i\right\}=\frac{\sum\limits_{y_j\le y}p_{ij}}{p(i,\cdot)}$$
为$\mathbf{X}=x_i$条件下随机变量

Y的分布函数；同理也称$F\left(x/y_j\right)=P\left\{\mathbf{X}\le x/\mathbf{Y}=y_j\right\}=\dfrac{\sum\limits_{x_i\le x}p_{ij}}{p(\cdot,j)}$

为$\mathbf{Y}=y_j$条件下随机变量**X**的分布函数。

1.3 随机变量及其概率分布

当$\mathbf{X},\mathbf{Y}$为连续型随机变量，对任意的x，$P\{\mathbf{X}=x\}=0$，用上面的方法无法定义$F_{\mathbf{Y}/\mathbf{X}}(y/x)$。为此，我们考虑极限：

$$\lim_{\Delta x\to 0^+} P\{\mathbf{Y}\le y / x-\Delta x<\mathbf{X}\le x+\Delta x\}$$

$$=\lim_{\Delta x\to 0^+}\frac{P\{\mathbf{Y}\le y, x-\Delta x<\mathbf{X}\le x+\Delta x\}}{P\{x-\Delta x<\mathbf{X}\le x+\Delta x\}}$$

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

其中$P\{x-\Delta x<\mathbf{X}\le x+\Delta x\}>0$，因此可给出下面的定义。

定义1.33设$(\mathbf{X},\mathbf{Y})$的分布函数为$F(x,y)$，若对任意的$\Delta x>0$，$P\{x-\Delta x<\mathbf{X}\le x+\Delta x\}>0$，且

$$\lim_{\Delta x\to 0^+}\frac{F(x+\Delta x,y)-F(x-\Delta x,y)}{F(x+\Delta x,+\infty)-F(x-\Delta x,+\infty)}$$

存在，则将$F_{\mathbf{Y}/\mathbf{X}}(y/x)$

定义为 $\lim\limits_{\Delta x\to 0^+}\dfrac{F(x+\Delta x,y)-F(x-\Delta x,y)}{F(x+\Delta x,+\infty)-F(x-\Delta x,+\infty)}$称其为$\mathbf{X}=x$条件下$\mathbf{Y}$的条件分布函数(比初等概率论更完善)

当$(\mathbf{X},\mathbf{Y})$的分布密度为$f(x,y)$，上式可表示为：

$$F_{\mathbf{Y}/\mathbf{X}}(y/x)=\lim_{\Delta x\to 0^+}\frac{\int_{x-\Delta x}^{x+\Delta x}\int_{-\infty}^{y}f(u,v)\,dudv}{\int_{x-\Delta x}^{x+\Delta x}\int_{-\infty}^{+\infty}f(u,v)\,dudv}$$

利用中值定理，有：$F_{\mathbf{Y}/\mathbf{X}}(y/x)=\dfrac{\int_{-\infty}^{y} f(x,v)dv}{\int_{-\infty}^{+\infty} f(x,v)dv}$

定义1.34若$f(x,y)$是$(\mathbf{X},\mathbf{Y})$的分布密度，且$\int_{-\infty}^{+\infty} f(x,y)dy>0$，

$f(x,y)$在(x,y)处连续，定义：

$$F_{\mathbf{Y}/\mathbf{X}}(y/x)=\frac{\int_{-\infty}^{y} f(x,v)dv}{\int_{-\infty}^{+\infty} f(x,v)dv}=\frac{\int_{-\infty}^{y} f(x,v)dv}{f_{\mathbf{X}}(x)}=\int_{-\infty}^{y}\frac{f(x,v)}{f_{\mathbf{X}}(x)}dv$$

为$\mathbf{X}=x$条件下$\mathbf{Y}$的条件分布函数。并称$f_{\mathbf{Y}/\mathbf{X}}(y/x)=\dfrac{f(x,v)}{f_{\mathbf{X}}(x)}$

为$\mathbf{X}=x$条件下$\mathbf{Y}$的条件分布密度。

例1.17 设二维随机变量的联合密度为：

$$f(x,y)=\begin{cases} e^{-y} & 0<x<y \\ 0 & \text{其它} \end{cases}$$

(1) $(\mathbf{X},\mathbf{Y})$的边缘概率密度 $f_{\mathbf{X}}(x)$，$f_{\mathbf{Y}}(y)$

(2) $\mathbf{X},\mathbf{Y}$是否相互独立（说明理由）

(3) $P\{\mathbf{X}+\mathbf{Y}\le 1\}$

解： (1) $$f_{\mathbf{X}}(x)=\int_{-\infty}^{+\infty} f(x,y)\,dy=\begin{cases} \int_{x}^{+\infty} e^{-y}dy=e^{-x} & x\ge 0 \\ 0 & x<0 \end{cases}$$

$$f_{\mathbf{Y}}(y)=\int_{-\infty}^{+\infty} f(x,y)\,dx=\begin{cases} \int_{0}^{y} e^{-y}dx=ye^{-y} & y\ge 0 \\ 0 & y<0 \end{cases}$$

(2)由于$f(x,y) \neq f_{\mathbf{X}}(x) \cdot f_{\mathbf{Y}}(y)$

所以**X**，**Y**不独立

（3）$P\{\mathbf{X}+\mathbf{Y} \leq 1\} = \iint\limits_{\mathbf{X}+\mathbf{Y}\leq 1} f(x,y)\,dxdy$

$$= \int_0^{\frac{1}{2}} dx \int_x^{1-x} e^{-y} dy$$

$$= 1 + \frac{1}{e} - 2e^{-\frac{1}{2}}$$

1.3.11二维均匀分布和二维正态分布（略）

1.3.12随机变量函数的分布（仅讨论连续型情形）

（一）一维随机变量函数的分布

定理1.13设随机变量**X**的概率密度为$f(x)$，函数$y=g(x)$处处可导，对$\forall x$有：$g'(x)>0$或$g'(x)<0$，则随机变量$\mathbf{Y}=g(\mathbf{X})$的概率密度为:$\varphi(y)=\begin{cases} f\left[h(y)\right]\left|h'(y)\right| & \alpha<y<\beta \\ 0 & 其它 \end{cases}$ 其中$h(y)$为$g(x)$的反函数，$\alpha=min\left(g(-\infty,+\infty)\right)$，$\beta=max\left(g(-\infty,+\infty)\right)$。

1.3.12随机变量函数的分布

（二）多维随机变量函数的分布

1、二维随机变量函数的分布

定理1.14设二维随机变量$(\mathbf{X},\mathbf{Y})$的联合概率密度为$f(x,y)$，函数$\mathbf{Z}=g(\mathbf{X},\mathbf{Y})$为连续函数，则：

$$F_{\mathbf{Z}}(z)=p\{\mathbf{Z}\le z\}=\iint_{g(x,y)\le z} f(x,y)\,dxdy$$

1.3.12随机变量函数的分布

例1.18设随机变量**X**，**Y**相互独立，其概率分布分别为：

$$f_{\mathbf{X}}(x)=\begin{cases}1 & 0\le x<1\\ 0 & \text{其它}\end{cases},\quad f_{\mathbf{Y}}(y)=\begin{cases}e^{-y} & y\ge 0\\ 0 & y<0\end{cases}$$

求随机变量**Z**=2**X**＋**Y**的概率密度。

解法一随机变量**Z**=2**X**＋**Y**的分布函数为：

$$F_{\mathbf{Z}}(z)=P\{2\mathbf{X}+\mathbf{Y}\le z\}=\iint\limits_{2x+y\le z} f(x,y)\,dxdy$$

由于**X**，**Y**相互独立，则：

$$f(x,y)=f_{\mathbf{X}}(x)f_{\mathbf{Y}}(y)=\begin{cases}e^{-y} & 0\le x<1, y>0\\ 0 & \text{其它}\end{cases}$$

因此随机变量Z=2X + Y的分布函数为：

(1)当$\frac{z}{2} < 0$，$F_Z(z) = 0$

(2)当$0 \le \frac{z}{2} < 1$，$F_Z(z) = \int_0^{\frac{z}{2}} dx \int_0^{z-2x} e^{-y} dy = \frac{1}{2}\left(z - 1 + e^{-z}\right)$

(3)当$\frac{z}{2} \ge 1$，$F_Z(z) = \int_0^{1} dx \int_0^{z-2x} e^{-y} dy = 1 - \frac{1}{2}\left(e^{2-z} - e^{-z}\right)$

从而Z的概率密度为：

$$f_Z(z) = \begin{cases} 0 & z < 0 \\ \frac{1}{2}\left(1 - e^{-z}\right) & 0 \le z < 2 \\ \frac{1}{2}\left(e^{2} - 1\right)e^{-z} & z \ge 2 \end{cases}$$

解法二也可以用卷积公式

$$f_Z(z)=\int_{-\infty}^{+\infty} f_X(2x)\cdot f_Y(z-2x)\,dx = f_X(2x) * f_Y(y)$$

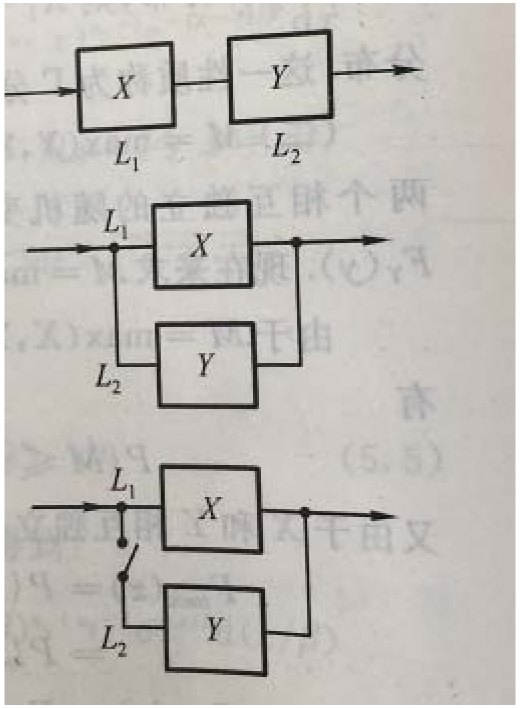
若$z-2x\ge 0$，有：$f_Y(z-2x)=0$且由$x<0$，有：$f_X(2x)=0$

$$f_Z(z)=\begin{cases} 0 & z<0 \\ \frac{1}{2}\left(1-e^{-z}\right) & 0\le z<2 \\ \frac{1}{2}\left(e^{2}-1\right)e^{-z} & z\ge 2 \end{cases}$$

显然用卷积公式更简捷，但对二次积分定限的要求更灵活。

应用实例

例　设系统L由两个相互独立的子系统L_1、L_2联接而成，联接的方式分别是(1)串联,(2)并联,(3)备用（当系统L_1损坏时，系统L_2开始工作），如右图所示。设L_1、L_2的寿命分别为X、Y，已知它们的概率密度分别为：



$$f_X(x)=\begin{cases}\alpha e^{-\alpha x} & x>0\\ 0 & x\le 0\end{cases}$$

$$f_Y(y)=\begin{cases}\beta e^{-\beta y} & y>0\\ 0 & y\le 0\end{cases}$$

其中$\alpha>0$，$\beta>0$且$\alpha\neq\beta$，试分别以上述三种联接方式写出L的寿命Z的概率密度。

解：(1)串联，当L_1、L_2中有一个损坏时，系统L就停止工作，因此寿命$Z=\min(X,Y)$。

$$F_X(x)=\begin{cases}1-e^{-\alpha x} & x>0\\ 0 & x\le 0\end{cases}\qquad F_Y(y)=\begin{cases}1-e^{-\beta y} & y>0\\ 0 & y\le 0\end{cases}$$

由X、Y彼此相互独立，则$Z=\min(X,Y)$的分布函数$F_Z(z)$为：

$$F_Z(z)=P\{Z\le z\}=1\text{-}P\{Z>z\}=1-P\{X>z,Y>z\}$$

$$=1-P\{X>z\}\cdot P\{Y>z\}=1\text{-}\left[1-F_X(z)\right]\cdot\left[1-F_Y(z)\right]$$

$$=\begin{cases}1-e^{-(\alpha+\beta)x} & z>0\\ 0 & z\le 0\end{cases}$$

则Z的概率密度为：$f_{\min(X,Y)}(z)=f_Z(z)=\begin{cases}(\alpha+\beta)e^{-(\alpha+\beta)x} & z>0\\ 0 & z\le 0\end{cases}$

解：(2)并联，当且仅当L_1、L_2都损坏时，系统L才停止工作，因此寿命$Z=\max(X,Y)$。

$$\because F_X(x)=\begin{cases}1-e^{-\alpha x} & x>0\\ 0 & x\le 0\end{cases} \qquad F_Y(y)=\begin{cases}1-e^{-\beta y} & y>0\\ 0 & y\le 0\end{cases}$$

由X、Y彼此相互独立，则$Z=\max(X,Y)$的分布函数$F_Z(z)$为：

$$F_Z(z)=P\{Z\le z\}=P\{X\le z,Y\le z\}=P\{X\le z\}\cdot P\{Y\le z\}$$

$$=F_X(z)\cdot F_Y(z)=\begin{cases}\left(1-e^{-\alpha z}\right)\left(1-e^{-\beta z}\right) & z>0\\ 0 & z\le 0\end{cases}$$

则Z的概率密度为：

$$f_{\max(X,Y)}(z)=f_Z(z)=\begin{cases}\alpha e^{-\alpha z}+\beta e^{-\beta z}-(\alpha+\beta)e^{-(\alpha+\beta)x} & z>0\\ 0 & z\le 0\end{cases}$$

解：(3)备用的情况，当L_1损坏时L_2才开始工作，因此寿命$Z=X+Y$，因此，当$z>0$时：

$$\because f_X(x)=\begin{cases}\alpha e^{-\alpha x} & x>0\\ 0 & x\le 0\end{cases} \qquad f_Y(y)=\begin{cases}\beta e^{-\beta y} & y>0\\ 0 & y\le 0\end{cases}$$

X、Y彼此相互独立，当$z>0$时$Z=X+Y$的概率密度$f_Z(z)$为：

$$f_Z(z)=\int_{-\infty}^{+\infty} f_X(z-y)\cdot f_Y(y)dy=\int_0^z \alpha e^{-\alpha(z-y)}\cdot \beta e^{-\beta y}dy$$

$$=\alpha\beta e^{-\alpha z}\int_0^z e^{-(\beta-\alpha)y}dy=\frac{\alpha\beta}{\beta-\alpha}\left[e^{-\alpha x}-e^{-\beta x}\right]$$

当$z\le 0$时$Z=X+Y$的概率密度$f_Z(z)=0$，则$Z=X+Y$的概率密度为

$$f_Z(z)=\begin{cases}\dfrac{\alpha\beta}{\beta-\alpha}\left[e^{-\alpha x}-e^{-\beta x}\right] & z>0\\ 0 & z\le 0\end{cases}$$

定理1.15设n维随机变量$\mathbf{X}=(\mathbf{X}_1,\cdots,\mathbf{X}_n)$的分布密度为$f_{\mathbf{X}}(x_1,\cdots,x_n)$，$n$元函数$g_j(x_1,\cdots,x_n)(j=1,\cdots,n)$满足条件：

（1）存在惟一的反函数，即方程组

$$y_j=g_j(x_1,\cdots,x_n)(j=1,\cdots,n) \qquad (*)$$

如果有解就存在惟一的实数解$x_j=x_j(y_1,\cdots,y_n)(j=1,\cdots,n)$;

(2)$g_j(x_1,\cdots,x_n)$和$x_j(y_1,\cdots,y_n)$都是连续函数；

（3）存在连续偏导数$\dfrac{\partial x_j}{\partial y_i},\dfrac{\partial y_j}{\partial x_i}$,若以$J$表示$Jacobi$行列式

$$J=\begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \cdots & \dfrac{\partial x_1}{\partial y_n} \\ & \cdots & \\ \dfrac{\partial x_n}{\partial y_1} & & \dfrac{\partial x_n}{\partial y_n} \end{vmatrix} \neq 0$$

则由$\eta_j = g_j(\xi_1,\cdots,\xi_n)(j=1,\cdots,n)$构成的$n$维随机变量$\mathbf{Y}=(\mathbf{Y}_1,\cdots,\mathbf{Y}_n)$的分布密度$f_{\mathbf{Y}}(y_1,\cdots,y_n)=$

$$\begin{cases} f[x_1(y_1,\cdots,y_n),\cdots,x_n(y_1,\cdots,y_n)]\,|J| & \text{当}(y_1,\cdots,y_n)\text{使(*)式有解} \\ 0 & \text{其他} \end{cases}$$

例1.19 设$\xi=(\xi_1,\xi_2)$的分布密度为$f(x_1,x_2)$，且有

$$\begin{cases} y_1 = ax_1 + bx_2 \\ y_2 = cx_1 + dx_2 \end{cases},\text{其中}\Delta = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0,\text{令}$$

$$\begin{cases} \eta_1 = a\xi_1 + b\xi_2 \\ \eta_2 = c\xi_1 + d\xi_2 \end{cases}$$

求$\eta=(\eta_1,\eta_2)$的分布密度$f_\eta(y_1,y_2)$.

解 方程组$\begin{cases} y_1 = ax_1 + bx_2 \\ y_2 = cx_1 + dx_2 \end{cases}$有唯一解

$$\begin{cases} x_1 = \dfrac{d}{\Delta} y_1 - \dfrac{b}{\Delta} y_2 \\ x_2 = -\dfrac{c}{\Delta} y_1 + \dfrac{a}{\Delta} y_2 \end{cases},$$

且$J = \begin{vmatrix} \dfrac{d}{\Delta} & -\dfrac{b}{\Delta} \\ -\dfrac{c}{\Delta} & \dfrac{a}{\Delta} \end{vmatrix} = \dfrac{1}{\Delta} \neq 0,$

于是得$\eta = (\eta_1, \eta_2)$的分布密度

$$f_\eta(y_1, y_2) = f\left(\frac{d}{\Delta} y_1 - \frac{b}{\Delta} y_2, -\frac{c}{\Delta} y_1 + \frac{a}{\Delta} y_2\right) \frac{1}{|\Delta|}$$

推论　设n维随机变量$\mathbf{Y}=(\mathbf{Y}_1,\cdots,\mathbf{Y}_n)$的分布密度为$f(x_1,\cdots,x_n)$，$n$元函数$g_j(x_1,\cdots,x_n)(j=1,\cdots,n)$满足条件：

（1）存在m支的反函数，即方程组

$$y_j=g_j(x_1,\cdots,x_n)(j=1,\cdots,n) \qquad (*)$$

有m组实数解$x_j=x_j^{(k)}(y_1,\cdots,y_n)(j=1,\cdots,n;k=1,\cdots,m)$;

(2)$g_j(x_1,\cdots,x_n)$和$x_j^{(k)}(y_1,\cdots,y_n)$都是连续函数;

（3）存在连续偏导数$\dfrac{\partial x_j^{(k)}}{\partial y_i},\dfrac{\partial y_j}{\partial x_i}$，若以$J$表示Jacobi行列式

$$J=\begin{vmatrix} \dfrac{\partial x_1^{(k)}}{\partial y_1} & \cdots & \dfrac{\partial x_1^{(k)}}{\partial y_n} \\ & \cdots & \\ \dfrac{\partial x_n^{(k)}}{\partial y_1} & & \dfrac{\partial x_n^{(k)}}{\partial y_n} \end{vmatrix} \neq 0$$

则n维随机变量$\mathbf{Y}=\left(\mathbf{Y}_1,\cdots,\mathbf{Y}_n\right)$的分布密度

$$f_{\mathbf{Y}}(y_1,\cdots,y_n)=$$

$$\begin{cases}\sum_{k=1}^{m} f[x_1^{(k)}(y_1,\cdots,y_n),\cdots,x_n^{(k)}(y_1,\cdots,y_n)]\,|J| & \text{当}(y_1,\cdots,y_n)\text{使(*)式有解} \\ 0 & \text{其他}\end{cases}$$

例1.20 设(ξ_1,ξ_2)的分布密度为:$f(x_1,x_2)=\dfrac{1}{2\pi\sigma^2}e^{-\frac{x_1^2+x_2^2}{2\sigma^2}}$

求$U=\sqrt{\xi_1^2+\xi_2^2}$的分布密度。

解法1 $\because(\xi_1,\xi_2)$的分布密度为:$f(x_1,x_2)=\dfrac{1}{2\pi\sigma^2}e^{-\frac{x_1^2+x_2^2}{2\sigma^2}}$

$$F_U(u)=P\{U\le u\}=\iint\limits_{\sqrt{x_1^2+x_2^2}\le z} f(x,y)\,dxdy$$

解法2: 令$\eta_1=\sqrt{\xi_1^{\ 2}+\xi_2^{\ 2}}$，$\eta_2=\xi_2$，求解方程组

$$\begin{cases} y_1=\sqrt{x_1^{\ 2}+x_2^{\ 2}} \\ y_2=x_2 \end{cases} \text{得:} \begin{cases} x_1=\pm\sqrt{y_1^{\ 2}-y_2^{\ 2}} \\ x_2=y_2 \end{cases}, |y_2|<y_1$$

$$\text{且} J=\begin{vmatrix} \pm\dfrac{y_1}{\sqrt{y_1^{\ 2}-y_2^{\ 2}}} & \mp\dfrac{y_2}{\sqrt{y_1^{\ 2}-y_2^{\ 2}}} \\ 0 & 1 \end{vmatrix}=\pm\frac{y_1}{\sqrt{y_1^{\ 2}-y_2^{\ 2}}}\neq 0$$

则$\eta=(\eta_1,\ \eta_2)$的分布密度

$$f_\eta(y_1,y_2)=\sum_{k=1}^{2} f[x_1^{(k)}(y_1,\cdots,y_n),x_2^{(k)}(y_1,\cdots,y_n)]|J|$$

$$=\frac{1}{\pi\sigma^2}e^{-\frac{y_1^{\ 2}}{2\sigma^2}}\frac{y_1}{\sqrt{y_1^{\ 2}-y_2^{\ 2}}}, |y_2|<y_1$$

$$\text{则} f_{\eta_1}(y_1)=\frac{1}{\pi\sigma^2}e^{-\frac{y_1^{\ 2}}{2\sigma^2}}\int_{-y_1}^{y_1}\frac{y_1}{\sqrt{y_1^{\ 2}-y_2^{\ 2}}}dy_2=\frac{y_1}{\sigma^2}e^{-\frac{y_1^{\ 2}}{2\sigma^2}}(y_1\geq 0)$$

1.4 随机变量的数字特征

1.4.1随机变量的数学期望和方差

定义1.35设随机变量**X**的分布律为：$P\{X=x_k\}=p_k(k=1,2,\cdots)$，若级数$\sum_k x_k p_k$绝对收敛，则称它为**X**的数学期望（或均值），记作：$E(\mathbf{X})=\sum_k x_k p_k$；若级数$\sum_k (x_k-E(\mathbf{X}))^2 p_k$收敛，则称它为**X**的方差，记作：$D(X)=\sum_k (x_k-E(\mathbf{X}))^2 p_k$。

方差描述的是随机变量**X**围绕“平均值”的离散程度，若把x_k想象为第k个质点所处位置的横坐标，p_k表示第k个质点的质量，则均值表示质点系的重心横坐标，而方差表示该质点系相对于通过重心的纵轴的转动惯量，且可以表示成如下矩阵形式：

$$\begin{pmatrix} x_0 & x_1 & x_2 & x_3 & \cdots & x_n & \cdots \\ p_0 & p_1 & p_2 & p_3 & \cdots & p_n & \cdots \end{pmatrix}$$

定义1.36设随机变量**X**的概率密度为$f(x)$，若$\int_{-\infty}^{+\infty} xf(x)dx$绝对收敛，则称其值为**X**的数学期望，记作：$E(\mathbf{X}) = \int_{-\infty}^{+\infty} xf(x)dx$

若积分$\int_{-\infty}^{+\infty} (x - E(\mathbf{X}))^2 f(x)dx$收敛，则称其值为**X**的方差，记作：$D(\mathbf{X}) = \int_{-\infty}^{+\infty} (x - E(\mathbf{X}))^2 f(x)dx$

1.4.1随机变量的数学期望和方差

测度论的观点，随机变量的数字特征可以表示为可测函数在概率空间$(\Omega,\mathcal{F},P)$是的积分：

$$E(\xi)=\int_{\Omega}\xi dP \qquad D(\xi)=\int_{-\infty}^{+\infty}\left[\xi-E(\xi)\right]^2 dP$$

概率空间$(\Omega,\mathcal{F},P)$是的积分可以变换到测度空间$\left(\mathbf{R}^{(1)},\mathcal{B}^{(1)},P_F\right)$，

因此：

$$E(\xi)=\int_{\Omega}\xi dP=\int_{-\infty}^{+\infty}xdF(x)\overset{\text{连续型}}{=}\int_{-\infty}^{+\infty}xf(x)dx$$

$$D(\xi)=\int_{-\infty}^{+\infty}\left[\xi-E(\xi)\right]^2 dP=\int_{-\infty}^{+\infty}\left[x-E(\xi)\right]^2 dF(x)\overset{\text{连续型}}{=}\int_{-\infty}^{+\infty}\left[x-E(\xi)\right]^2 f(x)dx$$

且：$D(\xi)=E\left[\xi-E(\xi)\right]^2=E\left(\xi^2\right)-E^2(\xi)$

1.4.2数学期望和方差的性质

(1) $E(c)=c$， $D(c)=0$

(2) $E(a\mathbf{X}+b)=aE(\mathbf{X})+b$,

$D(a\mathbf{X}+b)=a^2D(\mathbf{X})\left(a,b\text{为常数}\right)$

(3) 若$\mathbf{X}$，$\mathbf{Y}$独立，则$E(\mathbf{XY})=E(\mathbf{X})E(\mathbf{Y})$,

$D(\mathbf{X}\pm\mathbf{Y})=D(\mathbf{X})+D(\mathbf{Y})$

$$E\left(c\xi\right)=\int_{\Omega}c\xi dP=\int_{-\infty}^{+\infty}cxdF\left(x\right)=c\int_{-\infty}^{+\infty}xdF\left(x\right)=cE\left(\xi\right)$$

1.4.3常见随机变量的数学期望和方差

(1) $\mathbf{X} \sim B(n, p)$ $\quad E(\mathbf{X}) = np \quad D(\mathbf{X}) = npq = np(1-p)$

(2) 若$\mathbf{X}$服从参数为λ的泊松分布，$E(\mathbf{X}) = \lambda \quad D(\mathbf{X}) = \lambda$

(3) 若$\mathbf{X}$服从$N(a,\sigma^2)$，$E(\mathbf{X}) = a \quad D(\mathbf{X}) = \sigma^2$

(4) 若$\mathbf{X}$在(a,b)服从均匀分布，$E(\mathbf{X}) = \dfrac{a+b}{2}, D(\mathbf{X}) = \dfrac{(b-a)^2}{12}$

1.4.4随机变量函数的数学期望（仅讨论连续型）

定理1.16设**X**是概率空间$(\Omega, \mathcal{F}, P)$上的随机变量，其分布函数为$F(x)$，g是$\mathbf{R}^{(1)}$上的连续函数，则**Y**=g(**X**)的数学期望存在$\Leftrightarrow g(x)$在$\mathbf{R}^{(1)}$上关于$F(x)$的积分存在，且：

$$E(\mathbf{Y})=E\left[g(\mathbf{X})\right]=\int_{-\infty}^{+\infty}g(x)dF(x)$$

简单分析：$\because E(\mathbf{X})=\int_{-\infty}^{+\infty}xdF(x)$

$$\therefore E(\mathbf{Y})=E\left[g(\mathbf{X})\right]=\int_{-\infty}^{+\infty}g(x)dF(x)$$

例1.21设随机变量 X 的概率密度 $f(x)=\dfrac{1}{\pi(1+x^2)}$,

求 $E[\min(|X|,1)]$.

解 $$E[\min(|X|,1)]=\int_{-\infty}^{+\infty}\min(|x|,1)f(x)\mathrm{d}x$$

$$=\int_{|x|<1}|x|f(x)\mathrm{d}x+\int_{|x|\ge 1}f(x)\mathrm{d}x$$

$$=\frac{1}{\pi}\int_{-1}^{1}\frac{|x|}{1+x^2}\mathrm{d}x+\frac{1}{\pi}\int_{|x|\ge 1}\frac{1}{1+x^2}\mathrm{d}x$$

$$=\frac{2}{\pi}\int_{0}^{1}\frac{x}{1+x^2}\mathrm{d}x+\frac{2}{\pi}\int_{1}^{+\infty}\frac{1}{1+x^2}\mathrm{d}x=\frac{1}{\pi}\ln 2+\frac{1}{2}.$$

例1.22 设X，Y为相互独立的随机变量，且均服从$N(0,1)$，求：$E[\min(X,Y)]$。

解：由X，Y相互独立知X和Y的联合概率密度为：

$$f(x,y)=f_X(x)f_Y(y)=\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}，\ x,y\in R$$

则：$$E[\min(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\min(x,y)f(x,y)dxdy=\int_0^{2\pi}d\theta\int_0^{+\infty}\min(r\cos\theta,r\sin\theta)\frac{1}{2\pi}e^{-\frac{r^2}{2}}rdr$$

$$=\int_{-\frac{3\pi}{4}}^{\frac{\pi}{4}}d\theta\int_0^{+\infty}r\sin\theta\frac{1}{2\pi}e^{-\frac{r^2}{2}}rdr+\int_{\frac{\pi}{4}}^{\frac{5\pi}{4}}d\theta\int_0^{+\infty}r\cos\theta\frac{1}{2\pi}e^{-\frac{r^2}{2}}rdr$$

$$=\frac{-\sqrt{2}}{2\pi}\int_0^{+\infty}r^2e^{-\frac{r^2}{2}}dr+\frac{-\sqrt{2}}{2\pi}\int_0^{+\infty}r^2e^{-\frac{r^2}{2}}dr=\frac{-\sqrt{2}}{\pi}\int_0^{+\infty}r^2e^{-\frac{r^2}{2}}dr=\frac{-\sqrt{2}}{\pi}\int_0^{+\infty}(-r)de^{-\frac{r^2}{2}}$$

$$=\frac{-\sqrt{2}}{\pi}(-r)e^{-\frac{r^2}{2}}\Big|_0^{+\infty}-\frac{-\sqrt{2}}{\pi}\int_0^{+\infty}e^{-\frac{r^2}{2}}dr=-\frac{-\sqrt{2}}{\pi}\int_0^{+\infty}e^{-\frac{r^2}{2}}dr=-\frac{-\sqrt{2}}{\pi}\cdot\frac{\sqrt{2\pi}}{2}=-\frac{1}{\sqrt{\pi}}$$

1.4.5协方差与相关系数

对于二维随机变量(X,Y)，称$E(X-E(x))(Y-E(Y))$为 X 与 Y 的协方差，记作：

$$\begin{aligned}\operatorname{cov}(X,Y) &= E(X-E(Y))(Y-E(Y)) \\ &= E(XY-XE(Y)-YE(X)+E(X)E(Y)) \\ &= E(XY)-E(X)E(Y)\end{aligned}$$

若$\operatorname{cov}(X,Y)=0$，则称 X，Y 不相关，即：$E(XY)=E(X)E(Y)$

对于二维随机变量(X,Y)，若$D(X),D(Y)\neq 0$，则称$\dfrac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$为 X 与 Y 的相关系数，

记作r_{xy}，即：$r_{xy}=\dfrac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$

1.4.5协方差与相关系数

1．$0 \le \left| r_{xy} \right| \le 1$

2．若 X，Y 不相关，则：$r_{xy} = 0$

3．若 X 和 Y 有线性关系，即：$Y = aX + b \ (a, b = Const. \ \ a \ne 0)$，则：$\left| r_{xy} \right| = 1$

4．对于二维正态分布 $r_{xy} = 0 \Leftrightarrow$ X 和 Y 相互独立

（一般地，若 $r_{xy} = 0$ 不能推出 X 和 Y 独立，而独立一定能推出 $r_{xy} = 0$）

1.4.6几个重要的不等式

Chebyshev不等式 若$E(\mathbf{X})$有限，$D(\mathbf{X})$存在，则对$\forall \varepsilon > 0$，有：

$$P\left\{\left|\mathbf{X} - E(\mathbf{X})\right| \geq \varepsilon\right\} \leq \frac{D(\mathbf{X})}{\varepsilon^2}$$

Cauchy−Schwarz不等式 $\left|E(\mathbf{XY})\right|^2 \leq E\left(\left|\mathbf{X}\right|^2\right)E\left(\left|\mathbf{Y}\right|^2\right)$

证明：请参见扩展阅读的论文。

1.4.7条件数学期望

一般地，$E(\mathbf{X})=\int_{-\infty}^{+\infty} x dF(x), D(\mathbf{X})=\int_{-\infty}^{+\infty}(x-E(\mathbf{X}))^2 dF(x)$

定义1.37对于条件分布函数$F(y/x), F(x/y)$,若:

$\int_{-\infty}^{+\infty}|y| dF(y/x)<\infty, \int_{-\infty}^{+\infty}|x| dF(x/y)<\infty$，则称:

$E(\mathbf{Y}/x)=\int_{-\infty}^{+\infty} y dF(y/x)<\infty$为**X**=$x$条件下**Y**的条件数学期望,

$E(\mathbf{X}/y)=\int_{-\infty}^{+\infty} x dF(x/y)<\infty$为**Y**=$y$条件下**X**的条件数学期望。

例1.21随机变量X、Y的取值为1,2，…，n，其概率分布为：

$$P(X=i,Y=j)=\frac{1}{n^2},i,j=1,2,\cdots,n$$

求$E(Y|i)$，$E(X|j)$。

解：首先求出条件分布律为：

$$p(i,j)=\frac{1}{n^2},\quad p(i,\cdot)=\frac{1}{n},\qquad p(\cdot,j)=\frac{1}{n},$$

$$p(j\mid i)=\frac{p_{ij}}{p(i,\cdot)}=\frac{1}{n},\ p(i\mid j)=\frac{p_{ij}}{p(\cdot,j)}=\frac{1}{n},\ (i,j=1,2,\cdots n)$$

Y	y_1	y_2	$\cdots$	y_j
$P\left(Y/_{X=i}\right)$	$p\left(1/_i\right)$	$p\left(2/_i\right)$	$\cdots$	$p\left(j/_i\right)$

那么：

$$E(Y\mid i)=\sum_{j=1}^{n}jp(j\mid i)=\frac{n+1}{2}\qquad 同理：E(X\mid j)=\sum_{i=1}^{n}ip(i\mid j)=\frac{n+1}{2}$$

例1.22设二维正态分布服从 $N(0,1;0,1;r)$,试求$f(y|x), f(x|y)$, $E(Y|x), E(X|y)$。

解: 已知

$$f(x,y)=\frac{1}{2\pi\sqrt{1-r^2}}\exp\left\{-\frac{1}{2(1-r^2)}(x^2-2rxy+y^2)\right\}$$

$$则: f(x)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{x^2}{2}\right\}, f(y)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{y^2}{2}\right\}$$

$$于是: \quad f(y|x)=\frac{f(x,y)}{f_X(x)}=\frac{1}{\sqrt{2\pi(1-r^2)}}\exp\left\{-\frac{1}{2}\left(\frac{y-rx}{\sqrt{1-r^2}}\right)^2\right\}$$

同理可得：

$$f(x\mid y)=\frac{f(x,y)}{f_Y(y)}=\frac{1}{\sqrt{2\pi(1-r^2)}}\exp\left\{-\frac{1}{2}\left(\frac{x-ry}{\sqrt{1-r^2}}\right)^2\right\}$$

而：

$$E(\mathbf{Y}\mid x)=\int_{-\infty}^{+\infty} yf(y\mid x)dy=$$

$$\int_{-\infty}^{+\infty} y\frac{1}{\sqrt{2\pi(1-r^2)}}\exp\left\{-\frac{1}{2}\left(\frac{y-rx}{\sqrt{1-r^2}}\right)^2\right\}dy=rx$$

同理可得： $E(\mathbf{X}\mid y)=ry$

从此例可以看出，$E(Y|x)$，$E(X|y)$分别是x和y的函数。而$E(Y|X)=rX$； $E(X|y)=rY$。

1.4.8随机变量函数的条件数学期望

设 $g(x)$ 在 R 上连续，若 $\int_{-\infty}^{+\infty}|g(x)|dF(x/y)<\infty$ $\int_{-\infty}^{+\infty}|g(y)|dF(y/x)<\infty$

则有：

$$E(g(X)/y)=\int_{-\infty}^{+\infty}g(x)dF(x/y)<\infty \quad E(g(Y)/x)=\int_{-\infty}^{+\infty}g(y)dF(y/x)<\infty$$

例**1.23**已知ξ,η在$[\mathbf{0,1}]$上服从均匀分布且相互独立，求$\boldsymbol{D}(\xi|\xi+\eta)$.

解：易知ξ和η的联合分布密度为：

$$f(x,y)=\begin{cases}1 & 0<x<1,0<y<1\\ 0 & \text{其它}\end{cases}$$

令$U=\xi$，$V=\xi+\eta$

则U和V的联合概率密度为

$$\varphi(u,v)=\begin{cases}1 & 0<u<1,0<v-u<1\\ 0 & \text{其它}\end{cases}$$

则关于V的边沿概率密度为

$$f_V(v)=\int_{-\infty}^{+\infty}\varphi(u,v)\,du=\begin{cases}\int_0^v 1du=v & 0<v\le 1\\ \int_{v-1}^1 1du=2-v & 1<v<2\\ 0 & \text{其它}\end{cases}$$

当$0<v<1$时 $f_{U|V}(u|v)=\dfrac{\varphi(u,v)}{f_V(v)}=\begin{cases}\frac{1}{v} & 0<u<v\\ 0 & \text{其它}\end{cases}$

当$1\le v<2$时 $f_{U|V}(u|v)=\begin{cases}\frac{1}{2-v} & v-1<u<1\\ 0 & \text{其它}\end{cases}$

于是，当$0<v<1$时 $E(U|V=v)=\int_{-\infty}^{+\infty}uf(u|v)du=\int_0^v u\dfrac{1}{v}du=\dfrac{v}{2}$

$$E(U^2|V=v)=\int_{-\infty}^{+\infty}u^2f(u|v)du=\int_0^v u^2\frac{1}{v}du=\frac{v^2}{3}$$

当$1\le v<2$时

$$E(U|V=v)=\int_{-\infty}^{+\infty}uf(u|v)du=\int_{v-1}^{1}u\frac{1}{2-v}du=\frac{v}{2}$$

$$E(U^2|V=v)=\int_{-\infty}^{+\infty}u^2f(u|v)du=\int_{v-1}^{1}u^2\frac{1}{2-v}du=\frac{1-v+v^2}{3}$$

$$\therefore E(\xi|\xi+\eta)=\frac{\xi+\eta}{2}$$

由于当$0<v<1$时

$$E\left(U\mid V=v\right)=\int_{-\infty}^{+\infty}uf(u\mid v)du=\int_0^v u\frac{1}{v}du=\frac{v}{2}$$

$$E\left(U^2\mid V=v\right)=\int_{-\infty}^{+\infty}u^2 f(u\mid v)du=\int_0^v u^2\frac{1}{v}du=\frac{v^2}{3}$$

当$0<\xi+\eta\le 1$，$D\left(\xi\mid\xi+\eta\right)=E\left(\xi\mid\xi+\eta\right)^2-E^2\left(\xi\mid\xi+\eta\right)$

$$=\frac{v^2}{3}-\left(\frac{v}{2}\right)^2=\frac{v^2}{12}$$

同理，当$1<\xi+\eta\le 2$，$D\left(\xi\mid\xi+\eta\right)=\dfrac{(2-\xi-\eta)^2}{12}$

1.4.8随机变量函数的条件数学期望

重要结论设**X**，**Y**，**Z**为随机变量，$g(x)$，$g(y)$在R上连续，且$E(\mathbf{X})$，$E(\mathbf{Y})$，$E(\mathbf{Z})$，$E\left(g(y)\cdot\mathbf{X}\right)$均存在，则：

1.**X**，**Y**相互独立，有：$E\left(\mathbf{X}/y\right)=E(\mathbf{X})$

证明：由**X**，**Y**相互独立，有：$F\left(x/y\right)=F_{\mathbf{X}}(x)$

则：$E\left(\mathbf{X}/y\right)=\int_{-\infty}^{+\infty} x\,dF\left(x/y\right)=\int_{-\infty}^{+\infty} x\,dF_{\mathbf{X}}(x)=E(\mathbf{X})$

$2.\ E\left[E\left(X/_{y}\right)\right]=E(X)$

证明：仅讨论(X, Y)为连续型情形

$$E\left[E\left(X/_{y}\right)\right]=\int_{-\infty}^{+\infty} E\left(X/_{y}\right) dF_Y(y)=\int_{-\infty}^{+\infty} E\left(X/_{y}\right) f_Y(y)\,dy$$

$$=\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} xf\left(x/_{y}\right)dx\right) f_Y(y)dy=\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} xf\left(x/_{y}\right) f_Y(y)\,dx\right)dy$$

$$=\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} xf(x, y)\,dx\right)dy=\int_{-\infty}^{+\infty} x\left(\int_{-\infty}^{+\infty} f(x, y)\,dy\right)dx$$

$$=\int_{-\infty}^{+\infty} xf_X(x, y)dx=E(X)$$

$$3.\ E\left[g(\mathbf{Y})\cdot \mathbf{X}/y\right]=g(y)E\left(\mathbf{X}/y\right)$$

证明：只需证明对任意固定的y，有：

$$E\left[g(\mathbf{Y})\cdot \mathbf{X}/y\right]=g(y)E\left(\mathbf{X}/y\right)$$

而：$E\left[g(\mathbf{Y})\cdot \mathbf{X}/y\right]=\int_{-\infty}^{+\infty} g(y)\cdot x dF\left(x/y\right)$

$$=g(y)\int_{-\infty}^{+\infty} x dF\left(x/y\right)=g(y)E\left(X/y\right)$$

$$4.\ E[g(\mathbf{Y})\cdot \mathbf{X}] = E[g(\mathbf{Y})\cdot E(\mathbf{X}/y)]$$

证明：利用3的结果，有：

$$E\left[g(\mathbf{Y})\cdot \mathbf{X}/y\right] = g(y)E\left(\mathbf{X}/y\right)$$

$$E\left[g(\mathbf{Y})\cdot E\left(\mathbf{X}/y\right)\right] = E\left\{E\left[g(\mathbf{Y})\cdot \mathbf{X}/y\right]\right\}$$

$$= E\left(g(\mathbf{Y})\cdot \mathbf{X}\right)$$

$$5.\ E\left(c/y\right) = c$$

证明：$E\left(c/y\right) = \int_{-\infty}^{+\infty} c\,dF\left(x/y\right) = c$

$$6.E\left(g(\mathbf{Y})/y\right)=g(y)$$

证明：$E\left(g(\mathbf{Y})/y\right)=g(y)E\left(1/y\right)=g(y)$

特别地，有：$E\left(\mathbf{Y}/\mathbf{Y}\right)=\mathbf{Y}$

$$7.E\left[(a\mathbf{X}+b\mathbf{Y})/z\right]=aE\left(\mathbf{X}/z\right)+bE\left(\mathbf{Y}/z\right)$$

$$8.E\left[\mathbf{X}-E\left(\mathbf{X}/y\right)\right]=E\left[\mathbf{X}-g(\mathbf{Y})\right]^2$$

1.4.8随机变量函数的条件数学期望

由2. $E\left[E\left(X/y\right)\right]=E(X)$

若Y只取有限个值$y_i\ (i=1,2,\ldots n)$，且:$\sum_{i=1}^{n}P\{Y=y_i\}=1$

若记：$A_i=\{Y=y_i\}$，$E\left(X/A_i\right)=E\left(X/y\right)$，则有:

$$E(X)=\sum_{i=1}^{n}E\left(X/A_i\right)P(A_i)$$——全数学期望公式

1.5随机变量的特征函数

为什么要引入特征函数？

数学期望的$L-S$积分表示：$E(\mathbf{X})=\int_{-\infty}^{+\infty} x\,dF(x)$

积分的逆运算是微分（或求导），而求导运算要比积分运算简单得多，因此是否可将求数学期望的积分问题简化为求导问题？

$\mathbf{Z}=\mathbf{X}+\mathbf{Y}$，$\mathbf{X},\mathbf{Y}$相互独立，则：$f_{\mathbf{Z}}(z)=f_{\mathbf{X}}(x)*f_{\mathbf{Y}}(y)$

用分布函数求独立随机变量和的分布需要用卷积，是否可将卷积运算简化为乘积运算？

1.5随机变量的特征函数

1.5.1随机变量的特征函数

定义1.38设**X**是概率空间$(\Omega,\mathcal{F},P)$上的随机变量，分布函数是$F(x)$，称t的函数：

$$\varphi(t)\overset{\Delta}{=}E\left(e^{itX}\right)=\int_{-\infty}^{+\infty}e^{itx}dF(x) \qquad (1.4.1)$$

为**X**的特征函数。

由$\left|E\left(e^{itX}\right)\right|\leq E(1)=1$，知特征函数的定义有意义。

若X为离散型随机变量，其分布律为：

$$\begin{bmatrix} a_1 & a_2 & \cdots & a_n & \cdots \\ p_1 & p_2 & \cdots & p_n & \cdots \end{bmatrix}$$

则X的特征函数为：

$$\varphi(t) = E\left(e^{it\mathbf{X}}\right) = \sum_{k=1}^{\infty} e^{ita_k} p_k \qquad (1.4.2)$$

若X为连续型随机变量，分布密度函数为*f*(*x*)，则X的特征函数为：

$$\varphi(t) = E\left(e^{it\mathbf{X}}\right) = \int_{-\infty}^{+\infty} e^{itx} f(x)\, dx \qquad (1.4.3)$$

例1.24 X服从两点分布，求X的特征函数

$$\begin{bmatrix} 1 & 0 \\ p & q \end{bmatrix} \quad (\text{其中} p+q=1)$$

解： $\varphi(t) = E\left(e^{it\mathbf{X}}\right) = e^{it1}p + e^{it0}q = q + pe^{it}$

例1.25 $\mathbf{X} \sim B(n,p)$，求X的特征函数

$$p_k = C_n^k p^k q^{n-k}, (k = 0,1,2,\cdots n)$$

解： $\varphi(t) = E\left(e^{it\mathbf{X}}\right) = \sum_{k=0}^{n} e^{itk} p_k$

$$= \sum_{k=0}^{n} e^{itk} C_n^k p^k q^{n-k} = \left(q + pe^{it}\right)^n$$

例1.26 $X \sim \pi(\lambda)$，$p_k = \frac{\lambda^k}{k!}e^{-\lambda}(\lambda > 0)$，求X的特征函数

$$\varphi(t) = E\left(e^{itX}\right) = \sum_{k=0}^{\infty} e^{itk}\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}e^{\lambda e^{it}} = e^{\lambda\left(e^{it}-1\right)}$$

例1.27 $X \sim U(a-h, a+h)$，求X的特征函数

$$f(x) = \begin{cases} \frac{1}{2h} & a-h < x < a+h \\ 0 & \text{其它} \end{cases}$$

解：$\varphi(t) = E\left(e^{itX}\right) = \int_{-\infty}^{+\infty} e^{itx} f(x)\,dx$

$$= \int_{a-h}^{a+h} e^{itx}\frac{1}{2h}dx = \frac{\sinh t}{ht}e^{ita}(t \neq 0)$$

当$t=0$时，$\varphi(t)=0$.

例1.28 X ~ $N(0,1)$，求X的特征函数

$$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

解：$\varphi(t)=\int_{-\infty}^{+\infty}e^{itx}f(x)dx=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{itx}e^{-\frac{x^2}{2}}dx=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\int_{-\infty}^{+\infty}e^{-\frac{(x-it)^2}{2}}dx$

利用复变函数中的闭路 积分定理，有：

$$\int_{-\infty}^{+\infty}e^{-\frac{(x-it)^2}{2}}dx=\int_{-\infty}^{+\infty}e^{-\frac{x^2}{2}}dx=\sqrt{2\pi}$$

得：$\varphi(t)=e^{-\frac{t^2}{2}}$

另解：

解：$\varphi(t)=\int_{-\infty}^{+\infty} e^{itx} f(x)\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \cos tx e^{-\frac{x^2}{2}} dx$

$$\varphi'(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} [-x\sin tx]e^{-\frac{x^2}{2}} dx$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \sin tx de^{-\frac{x^2}{2}}$$

$$=\frac{1}{\sqrt{2\pi}}\left[\sin tx e^{-\frac{x^2}{2}}\Big|_{-\infty}^{+\infty}-\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} t\cos tx dx\right]=-t\varphi(t)$$

于是$\varphi(t)=ce^{-\frac{t^2}{2}}$,由$\varphi(0)=1$，知$\varphi(t)=e^{-\frac{t^2}{2}}$

1.5.2 特征函数的性质

（1） $\varphi(t)$在$R^{(1)}$上一致连续，且 $|\varphi(t)| \le \varphi(0) = 1$， $\varphi(-t) = \overline{\varphi(t)}$

简单回顾高等数学中函数连续性的概念：

$\forall \varepsilon > 0, \exists \delta(t,\varepsilon) > 0$， 当$h \to 0$时， 有：$|\varphi(t+h) - \varphi(t)| < \varepsilon$

一致连续性概念的定义：

$\forall \varepsilon > 0, \exists \delta(\varepsilon) > 0$， 当$h \to 0$时， 一致地有：$|\varphi(t+h) - \varphi(t)| < \varepsilon$

1.5.2 特征函数的性质

（1） $\varphi(t)$在$R^{(1)}$上一致连续，且 $|\varphi(t)| \le \varphi(0) = 1$，$\varphi(-t) = \overline{\varphi(t)}$

证明:$$|\varphi(t+h)-\varphi(t)| = \left|E\left(e^{iX(t+h)} - e^{itX}\right)\right| = \left|\int_{-\infty}^{+\infty}\left(e^{ix(t+h)} - e^{itx}\right)dF(x)\right|$$

$$\le \int_{-\infty}^{+\infty}\left|e^{ihx} - 1\right|dF(x) = \int_{|x|\le A}\left|e^{ihx} - 1\right|dF(x) + \int_{|x|>A}\left|e^{ihx} - 1\right|dF(x)$$

$$\le \int_{-A}^{A} 2\left|sin\frac{hx}{2}\right|dF(x) + 2\int_{|x|>A} dF(x)$$

对$\forall\varepsilon > 0$，取充分大的A，使得 $\int_{|x|>A} dF(x) < \frac{\varepsilon}{4}$

再取$\delta > 0$， 当$|h| < \delta$,对一切 $|x| \le$ A一致地有$\left|sin\frac{hx}{2}\right| < \frac{\varepsilon}{4}$

则对t一致地有$|\varphi(t+h) - \varphi(t)| < \varepsilon$

（2） $\varphi(t)$非负定，即对任意的正 整数n及任意$t_k \in R^{(1)}$

和复数$\lambda_k(k=1,2,\cdots n)$，均有：$\sum_{l=1}^{n}\sum_{k=1}^{n}\varphi(t_l-t_k)\lambda_l\overline{\lambda_k}\geq 0$

证明：$\sum_{l=1}^{n}\sum_{k=1}^{n}\varphi(t_l-t_k)\lambda_l\overline{\lambda_k}=\sum_{l,k=1}^{n}\left(\int_{-\infty}^{+\infty}e^{i(t_l-t_k)x}dF(x)\right)\lambda_l\overline{\lambda_k}$

$$=\int_{-\infty}^{+\infty}\left(\sum_{l,k=1}^{n}e^{it_lx}e^{-it_kx}\lambda_l\overline{\lambda_k}\right)dF(x)=\int_{-\infty}^{+\infty}\left(\sum_{l=1}^{n}e^{it_lx}\lambda_l\overline{\sum_{k=1}^{n}e^{it_kx}\lambda_k}\right)dF(x)$$

$$=\int_{-\infty}^{+\infty}\left(\sum_{l=1}^{n}e^{it_lx}\lambda_l\overline{\sum_{l=1}^{n}e^{it_lx}\lambda_l}\right)dF(x)=\int_{-\infty}^{+\infty}\left(\left|\sum_{l=1}^{n}\lambda_le^{it_lx}\right|^2\right)dF(x)\geq 0$$

(3) 设$\mathbf{X}$的特征函数为$\varphi_{\mathbf{X}}(t)$，则$\mathbf{Y}=a\mathbf{X}+b$的特征函数为：$\varphi_{\mathbf{Y}}(t)=e^{itb}\varphi_{\mathbf{X}}(at)$

证明：$\varphi_{\mathbf{Y}}(t)=E\left(e^{it\mathbf{Y}}\right)=E\left[e^{it(a\mathbf{X}+b)}\right]=e^{itb}\varphi_{\mathbf{X}}(at)$

而：标准正态分布$\mathbf{X}\sim N(0,1)$，有：$\varphi_{\mathbf{X}}(t)=e^{-\frac{t^2}{2}}$

若正态分布$\mathbf{Y}\sim N\left(a,\sigma^2\right)$，有$\mathbf{Y}=\sigma\mathbf{X}+a$

根据性质(3)得：$\varphi_{\mathbf{Y}}(t)=e^{ita}\varphi_{\mathbf{X}}(\sigma t)=e^{ita}e^{-\frac{\sigma^2t^2}{2}}$

(4) 若随机变量$\mathbf{X}_1,\cdots\mathbf{X}_n$相互独立，其特征函数分

别为$\varphi_{\mathbf{X}_1}(t),\cdots\varphi_{\mathbf{X}_n}(t)$，令$\mathbf{Y}=\sum_{k=1}^{n}\mathbf{X}_k$，则$\mathbf{Y}$的特征函数为$\varphi_{\mathbf{Y}}(t)=\prod_{k=1}^{n}\varphi_{\mathbf{X}_k}(t)$

证明：当$n=2$时证明结论成立

$$\varphi_{\mathbf{Y}}(t)=E\left(e^{it\mathbf{Y}}\right)=E\left[e^{it(\mathbf{X}_1+\mathbf{X}_2)}\right]=E\left(e^{it\mathbf{X}_1}e^{it\mathbf{X}_2}\right)$$

$$=E\left(e^{it\mathbf{X}_1}\right)\cdot E\left(e^{it\mathbf{X}_2}\right)=\varphi_{\mathbf{X}_1}(t)\cdot\varphi_{\mathbf{X}_2}(t)$$

一般情况可用数学归纳 法证明，略

(5) 若随机变量$\mathbf{X}$的n阶绝对矩$E\left(|\mathbf{X}|^n\right) < +\infty$，则有：

$$E\left(\mathbf{X}^k\right) = (-i)^k \varphi_{\mathbf{X}}^{(k)}(0), 1 \le k \le n$$

证明：$\varphi(t) = E\left(e^{it\mathbf{X}}\right) = \int_{-\infty}^{+\infty} e^{itx} dF(x)$

$$\varphi^{(k)}(t) = \frac{d^k}{dt^k}\left[\int_{-\infty}^{+\infty} e^{itx} dF(x)\right] = \int_{-\infty}^{+\infty} \frac{\partial^k}{\partial t^k} e^{itx} dF(x) = i^k \int_{-\infty}^{+\infty} x^k e^{itx} dF(x)$$

则：$\varphi^{(k)}(0) = i^k \int_{-\infty}^{+\infty} x^k dF(x) = i^k E\left(\xi^k\right)$，即得结论。

例1.29 设$\mathbf{X}_1$，$\mathbf{X}_2$独立且同为二项分布，即：

$$P(\mathbf{X}_1 = k) = C_n^k p^k q^{n-k} \quad (k = 0,1,2,\cdots n)$$

$$P(\mathbf{X}_2 = k) = C_m^k p^k q^{m-k} \quad (k = 0,1,2,\cdots m)$$

求$\mathbf{X}_1 + \mathbf{X}_2$的特征函数

解：由特征函数的性质（4），有：

$$\varphi_{\mathbf{X}_1+\mathbf{X}_2}(t) = \varphi_{\mathbf{X}_1}(t) \cdot \varphi_{\mathbf{X}_2}(t)$$

$$= \left(q + pe^{it}\right)^n \cdot \left(q + pe^{it}\right)^m = \left(q + pe^{it}\right)^{m+n}$$

1.5.3 逆转公式

由定义知：随机变量的特征函数被表示为分布分数的**F-S**变换，即随机变量函数的分布。那么，如何根据特征函数去求分布函数？

定理1.17（逆转公式） 设随机变量ξ的分布函数和特征函数分别为$F(x)$和$\varphi(t)$，则对任意的x_1，$x_2 \in R^{(1)}$，有：

（1）
$$\frac{F(x_2+0)+F(x_2-0)}{2}-\frac{F(x_1+0)+F(x_1-0)}{2}$$
$$=\frac{1}{2\pi}\lim_{l\to\infty}\int_{-l}^{l}\frac{e^{-itx_1}-e^{-itx_2}}{it}\varphi(t)dt$$

（2） 若x_1，x_2是$F(x)$的连续点，有：
$$F(x_2)-F(x_1)=\frac{1}{2\pi}\lim_{l\to\infty}\int_{-l}^{l}\frac{e^{-itx_1}-e^{-itx_2}}{it}\varphi(t)dt$$

证明：因涉及到 *Fubini* 定理等知识，因此略

由逆转公式，在 $F(x)$ 的连续点上（令 $x_2 = x$），当 x_1 沿 $F(x)$ 的连续点 $\to -\infty$，有：

$$F(x) = \frac{1}{2\pi} \lim_{x_1 \to -\infty} \lim_{l \to \infty} \int_{-l}^{l} \frac{e^{-itx_1} - e^{-itx}}{it} \varphi(t) dt$$

定理1.18 设 $F_1(x)$，$F_2(x)$是两个分布函数，$\varphi_1(t)$，$\varphi_2(t)$是其相应的特征函数，则：

$$F_1(x)=F_2(x) \Leftrightarrow \varphi_1(t)=\varphi_2(t)$$

证明："$\Rightarrow$" 设$F_1(x)=F_2(x)$，则有：

$$\varphi_1(t)=\int_{-\infty}^{+\infty} e^{itx}dF_1(x)=\int_{-\infty}^{+\infty} e^{itx}dF_2(x)=\varphi_2(t)$$

"$\Leftarrow$" 设$\varphi_1(t)=\varphi_2(t)$，则有：

$$\frac{F_1(x_2+0)+F_1(x_2-0)}{2}-\frac{F_1(x_1+0)+F_1(x_1-0)}{2}$$

$$=\frac{F_2(x_2+0)+F_2(x_2-0)}{2}-\frac{F_2(x_1+0)+F_2(x_1-0)}{2}$$

若x_1，x_2是$F_1(x)$，$F_2(x)$的连续点，则：

$$F_1(x_2)-F_1(x_1)=F_2(x_2)-F_2(x_1)$$

则：$$F_1(x_2)=\lim_{x_1\to-\infty}\left[F_1(x_2)-F_1(x_1)\right]$$

$$=\lim_{x_1\to-\infty}\left[F_2(x_2)-F_2(x_1)\right]=F_2(x_2)$$

若x^*是$F_1(x)$，$F_2(x)$的不连续点，取 $x_n\downarrow x^*$，使得x_n是$F_1(x)$，$F_2(x)$的连续点，则：

$$F_1(x^*)=F_1(x^*+0)=\lim_{n\to\infty}F_1(x_n)=\lim_{n\to\infty}F_2(x_n)$$

$$=F_2(x^*+0)=F_2(x^*)$$，即：$F_1(x)\equiv F_2(x)$

推论　若 $\varphi(t)$是特征函数，则有唯一的分布函数 $F(x)$存在，使得 $\varphi(t)$是$F(x)$的特征函数。

综上所述，我们知道分布函数$F(x)\underset{\longleftrightarrow}{1-1}\varphi(t)$，因此对分布函数的研究可以转化为对特征函数的研究。

定理1.19 若 $\varphi(t)$是一特征函数，且 $\int_{-\infty}^{+\infty}|\varphi(t)|dt < +\infty$，则由 $\varphi(t)$所决定的分布函数 $F(x)$是一连续型的分布函数，且 $F'(x)$在$R^{(1)}$上处处存在、有界且连续，并有：

$$F'(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx}\varphi(t)dt$$

（$F-S$变换与逆变换的关系，证明如下）

证明：令：$G(x)=\frac{1}{2}[F(x)+F(x-0)]$

（1） 首先证明 $G'(x)$存在，且 $G'(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}\varphi(t)dt$

先证 $G^{+}(x)$存在，且 $G^{+}(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}\varphi(t)dt$

$$\frac{G(x+2\Delta x)-G(x)}{2\Delta x}$$

$$=\frac{[F(x+2\Delta x)+F(x+2\Delta x-0)]-[F(x)+F(x-0)]}{2\cdot 2\Delta x}$$

$$=\frac{1}{2\pi}\lim_{l\to\infty}\int_{-l}^{+l}\frac{e^{-itx}-e^{-it(x+2\Delta x)}}{2it\Delta x}\varphi(t)dt=\frac{1}{2\pi}\lim_{l\to\infty}\int_{-l}^{+l}e^{-it(x+\Delta x)}\frac{\sin t\Delta x}{t\Delta x}\varphi(t)dt$$

$$=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-it(x+\Delta x)}\frac{\sin t\Delta x}{t\Delta x}\varphi(t)dt$$

由上页：$\dfrac{G(x+2\Delta x)-G(x)}{2\Delta x}=\dfrac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-it(x+\Delta x)}\dfrac{\sin t\Delta x}{t\Delta x}\varphi(t)dt$

根据控制收敛定理，有：

$$G^{+}(x)=\lim_{\Delta x\to 0^{+}}\frac{G(x+2\Delta x)-G(x)}{2\Delta x}$$

$$=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\lim_{\Delta x\to 0^{+}}e^{-it(x+\Delta x)}\frac{\sin t\Delta x}{t\Delta x}\varphi(t)dt=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}\varphi(t)dt$$

同理可证：$G^{-}(x)=\dfrac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}\varphi(t)dt$

则：$G^{+}(x)=G^{-}(x)$，则$G'(x)=\dfrac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}\varphi(t)dt$

（2） 证明$G'(x)$连续并且有界

$\because\ \forall x \in R^{(1)}$，有：$\left|e^{-itx}\varphi(t)\right| \le \left|\varphi(t)\right|$

而$\left|G'(x)\right| = \left|\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx}\varphi(t)dt\right| \le \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|\varphi(t)\right|dt < +\infty$

则$G'(x)$有界。因此对 $x_0 \in R^{(1)}$，有：

$$\lim_{x\to x_0} G'(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\lim_{x\to x_0} e^{-itx}\varphi(t)dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx_0}\varphi(t)dt = G'(x_0)$$

则$G'(x)$在$R^{(1)}$是连续的。

（3） 再证$F(x)\equiv G(x)$

事实上，由 $G(x)=\frac{1}{2}[F(x)+F(x-0)]$，知：

对$F(x)$的连续点，有： $F(x)=G(x)$；

对$F(x)$的不连续点x^*，取$F(x)$的连续点$x_n\downarrow x^*$，则：

$$F(x^*)=\lim_{n\to\infty}F(x_n)=\lim_{n\to\infty}G(x_n)=G(x^*)$$

定理证毕

定理1.20 设ξ为只取整数的随机变量，$P\{\xi = k\} = p_k$，$k = 0, \pm 1, \pm 2, \cdots$，特征函数 $\varphi(t) = \sum_{k=-\infty}^{+\infty} e^{itk} p_k$，则：

$$p_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-itk}\varphi(t)dt$$

证明：对任意的整数 s，有：$\varphi(t)e^{-its} = \sum_{\substack{k=-\infty \\ k\neq s}}^{+\infty} e^{-it(s-k)} p_k + p_s$

而 $\int_{-\pi}^{\pi} e^{-it(s-k)}dt = 0 (s \neq k)$，则：

$$\int_{-\pi}^{\pi} e^{-its}\varphi(t)dt = \sum_{\substack{k=-\infty \\ k\neq s}}^{+\infty} \int_{-\pi}^{\pi} e^{-it(s-k)} p_k dt + \int_{-\pi}^{\pi} p_s dt = 2\pi p_s$$

则：$p_s = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-its}\varphi(t)dt$ 即：$p_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-itk}\varphi(t)dt$

例1.30 设ξ_1，ξ_2的联合概率密度为：

$$f(x,y)=\begin{cases}\frac{1}{4}\left[1+xy\left(x^2-y^2\right)\right] & |x|\le 1, |y|\le 1 \\ 0 & \text{其它}\end{cases}$$

以$\varphi_i(t)$表ξ_i的特征函数，$i=1,2$，且$\eta=\xi_1+\xi_2$的特征函数为$\varphi(t)$，证明：$\varphi(t)=\varphi_1(t)\varphi_2(t)$，但$\xi_1$，$\xi_2$不独立。

该例子说明特征函数的性质（4）的逆不成立。

证明：ξ_1，ξ_2的边沿分布密度为：

$$f_{\xi_1}(x)=\begin{cases}\int_{-1}^{1}\frac{1}{4}\left[1+xy\left(x^2-y^2\right)\right]dy=\frac{1}{2} & |x|\le 1\\ 0 & \text{其它}\end{cases}$$

$$f_{\xi_2}(y)=\begin{cases}\int_{-1}^{1}\frac{1}{4}\left[1+xy\left(x^2-y^2\right)\right]dy=\frac{1}{2} & |y|\le 1\\ 0 & \text{其它}\end{cases}$$

即ξ_1，ξ_2均在$[-1,1]$均匀分布，但：

$f_{\xi_1}(x)f_{\xi_2}(y)=\frac{1}{4}\ne f(x,y)$，则$\xi_1$，$\xi_2$不独立

下面说明$\varphi(t)=\varphi_1(t)\varphi_2(t)$

显然$\eta=\xi_1+\xi_2$的分布密度函数为：

$$f_\eta(z)=\int_{-\infty}^{+\infty}f(x,z-x)dx \qquad (*)$$

由于$f(x,y)$在$|x|\le 1,|y|\le 1$外为零，则对固定的 z,$(*)$中的积分只有当$|x|\le 1$和$|z-x|\le 1$同时满足时非零，即：

$$\text{由}f_\eta(z)=\begin{cases}\int_{-1}^{z+1}\frac{1}{4}\left(1+3z^2x^2-2zx^3-z^3x\right)dx=\frac{1}{4}(2+z) & -2\le z\le 0\\ \int_{z-1}^{1}\frac{1}{4}\left(1+3z^2x^2-2zx^3-z^3x\right)dx=\frac{1}{4}(2-z) & 0<z\le 2\\ 0 & |z|>2\end{cases}$$

$f_\eta(z)$属于三角分布的密度函 数。

$$\varphi(t)=\int_{-\infty}^{+\infty}e^{itz}f_\eta(z)dz=\frac{1}{4}\left[\int_{-2}^{0}(2+z)e^{itz}dz+\int_{0}^{2}(2-z)e^{itz}dz\right]$$

$$=\frac{1}{4}\left[\frac{2-e^{2it}-e^{-2it}}{t^2}\right]=\left(\frac{sint}{t}\right)^2$$

同理可得： $\varphi_1(t)=\varphi_2(t)=\dfrac{sint}{t}$

则： $\varphi(t)=\varphi_1(t)\varphi_2(t)$

随机变量**X**的分布函数$F(x)$与特征函数$\varphi(t)$一一对应，研究随机变量**X**既可通过研究分布函数，又可通过研究特征函数来进行。但是一个函数应具备什么条件才能构成一个随机变量的特征函数？

（6）（***Bochner-Kh intchine***）**波赫纳－辛钦定理：函数$\varphi(t)$是特征函数 $\Leftrightarrow \varphi(t)$非负定、连续，且 $\varphi(0)=1$**

1.5.4 母函数

（1）母函数的定义

在离散型随机变量中，非负整数的随机变量具有重要的地位，如二项分布、泊松分布、几何分布等都是取非负整数值的，在研究这类随机变量时，可以采用母函数法。

定义1.39 设**X**是非负整数的随机变量，分布列:

$$p_k = P\{\mathbf{X} = k\}，\quad k = 0,1,2,\ldots$$

则称：$P(s) \stackrel{\Delta}{=} E\left(\mathbf{X}^k\right) = \sum_{k=0}^{\infty} p_k s^k$为**X**的母函数。

事实上，若设$s=e^{it}$，则：

$$P(s)=P\left(e^{it}\right)=\sum_{k=0}^{\infty}p_k e^{itk}=E\left(e^{it\mathbf{x}}\right)=\varphi(t)$$

所以母函数本质上也是特征函数，不过母函数在十九世纪初被拉普拉斯引入，称为概率论中第一个被系统应用的变换方法，因比较简单，在处理非负整数值随机变量的场合更方便。

因为$\sum_{k=0}^{\infty}p_k=1$，由幂级数的收敛性知道$P(s)$至少在$|s|\le 1$时一致收敛且绝对收敛。因此，母函数对任意非负整数值随机变量均存在。

注意：对于任意一数列$\{a_n\}$，也可以定义$\sum_{k=0}^{\infty} a_k s^k$为其母函数，但这里只讨论概率分布的母函数。

（2）母函数的计算

1. 二项分布 **X~**$B(n,p)$，$P(s)=\sum_{k=0}^{n} C_n^k p^k q^{n-k} s^k = (q+ps)^n$

2. 泊松分布 **X~**$\pi(\lambda)$，$P(s)=\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} s^k = e^{-\lambda} e^{\lambda s} = e^{\lambda(s-1)}$

3. 几何分布 **X~**$Geo(p)$，

$$P(s)=\sum_{k=1}^{\infty} q^{k-1} p s^k = ps\sum_{k=1}^{\infty}(qs)^{k-1} = \frac{ps}{1-qs}$$

（3）母函数的性质

1.唯一性 非负整数值随机变量的分布律由其母函数唯一确定。

证明：$P(s)=\sum_{k=0}^{\infty} p_k s^k = \sum_{k=0}^{n} p_k s^k + \sum_{k=n+1}^{\infty} p_k s^k$，$n=0,1,\ldots,$

上式左右两边对s求n阶导数，得：

$$P^{(n)}(s)=n!\cdot p_n + \sum_{k=n+1}^{\infty} k(k-1)\cdots(k-n+1)\, p_k s^{k-n}$$

令$s=0$，则$P^{(n)}(0)=n!\cdot p_n$, 所以：$p_n = P^{(n)}(0)\big/ n!$，$n=0,1,\ldots$

若概率分布$\{p_k\}$和$\{q_k\}$分别具有母函数$P(s)$和$H(s)$，且$P(s)$=$H(s)$，则根据以上推导结果，必有：$p_k=q_k$。反之亦然。因此，概率分布与母函数一一对应。

2.母函数与数字特征 设$P(s)$是$\mathbf{X}$的母函数，若：$E(\mathbf{X})$存在，则：$E(\mathbf{X})=P'(1)$，若$D(\mathbf{X})$存在，则：$D(\mathbf{X})=P''(1)+P'(1)-\left[P'(1)\right]^2$。

证明：由$P(s)=\sum_{k=0}^{\infty}p_k s^k$，所以$P'(s)=\sum_{k=1}^{\infty}kp_k s^{k-1}$，

$P''(s)=\sum_{k=2}^{\infty}k(k-1)p_k s^{k-2}$，上述两个级数至少在$|s|<1$收敛。

若$\mathbf{X}$的数学期望存在，即$E(\mathbf{X})=\sum_{k=1}^{\infty}kp_k$存在时，$P'(1)=E(\mathbf{X})$

若$\mathbf{X}$的数学期望不存在，即$E(\mathbf{X})=\sum_{k=1}^{\infty}kp_k=\infty$，

$\lim_{s\uparrow 1}P'(s)=\sum_{k=1}^{\infty}kp_k=\infty$，所以$E(\mathbf{X})=P'(1)$。

2. 母函数与数字特征 设$P(s)$是$\mathbf{X}$的母函数，若：$E(\mathbf{X})$存在，则：$E(\mathbf{X})=P'(1)$，若$D(\mathbf{X})$存在，则：$D(\mathbf{X})=P''(1)+P'(1)-\left[P'(1)\right]^2$。

证明（续）：同理，当$\mathbf{X}$的数学期望存在时，因为：

$$E\left(\mathbf{X}(\mathbf{X}-\mathbf{1})\right)=\sum_{k=1}^{\infty}k(k-1)p_k=P''(1)$$

则：$D(\mathbf{X})=E\left(\mathbf{X}^2\right)-E^2(\mathbf{X})=P''(1)+P'(1)-\left[P'(1)\right]^2$

于是得到了利用母函数求数学期望和方差的方法。

3.独立随机变量和的母函数 独立随机变量之和的母函数等于母函数之积。

证明：设**X**与**Y**为两个相互独立的非负整数值随机变量，其概率分布分别为$\{p_k\}$和$\{q_k\}$，相应的母函数分别为$P(s)$和$Q(s)$，令**Z**=**X**+**Y**，求**Z**的概率分布。

显然，**Z**也是非负整数值随机变量，若记：$r_k = P\{Z = k\}$

则：$r_k = p_0 q_k + p_1 q_{k-1} + \ldots + p_k q_0 = \sum_{l=1}^{k} p_l q_{k-l}$，

这便是离散卷积公式。

设$R(s)$为随机变量**Z**的母函数，$R(s)=\sum_{k=0}^{\infty} r_k s^k$

利用母函数在$|s|<1$的一致收敛性和绝对收敛性做下列运算：

$$P(s)Q(s)=\sum_{i=0}^{\infty} p_i s^i \sum_{j=0}^{\infty} q_j s^j=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} p_i q_j s^{i+j}$$

$$=\sum_{k=0}^{\infty}\left[\sum_{i=0}^{k} p_i q_{k-i}\right] s^k \left(\text{设} i+j=k\right)$$

$$=\sum_{k=0}^{\infty} r_k s^k = R(s)$$

因此：$R(s)=P(s)Q(s)$

即相互独立的非负整数值随机变量之和的母函数是这两个相应随机变量母函数的乘积。

4. 随机个独立同分布的非负整数值随机变量之和的母函数

若$\mathbf{X}_1$，$\mathbf{X}_2$,…是相互独立且同分布的非负整数值随机变量，N是与$\mathbf{X}_1$，$\mathbf{X}_2$,…独立的非负整数值随机变量，则：

$Y=\sum_{k=1}^{N}\mathbf{X}_k$的母函数：$H(s)=G(P(s))$，其中$G(s)$、$P(s)$分别是$N$、$\mathbf{X}_1$的母函数。

证明：$H(s)=\sum_{k=0}^{\infty}P\{Y=k\}s^k$

$$=\sum_{k=0}^{\infty}P\left\{Y=k,\bigcup_{l=0}^{\infty}(N=l)\right\}s^k=\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}P\{N=l,\ Y=k\}s^k$$

$$=\sum_{l=0}^{\infty}P\{N=l\}\sum_{k=0}^{\infty}P\{Y=k\}s^k=\sum_{l=0}^{\infty}P\{N=l\}[P(s)]^l=G(P(s))$$

例1.31 五颗骰子任意投掷，问总数为15的概率？

解：投掷每颗骰子可出现1, 2, 3, 4, 5, 6六种可能，它们出现的概率均为$1/6$。假设出现某数$\mathbf{X}_i$是随机变量，其相应的母函数为$\frac{1}{6}\left(s^1+s^2+\ldots+s^6\right)=\frac{s\left(1-s^6\right)}{6(1-s)}$

一次投掷五颗骰子时，每颗骰子出现的数是相互独立的，$\{\mathbf{X}_i\}$，$i=1,2,3,4,5$，五颗骰子出现的总和为：

$$\mathbf{Y}=\mathbf{X}_1+\mathbf{X}_2+\mathbf{X}_3+\mathbf{X}_4+\mathbf{X}_5$$

则$\mathbf{Y}$的母函数为：$H(s)=\left[\frac{s\left(1-s^6\right)}{6(1-s)}\right]^5$

展开$H(s)$，取s^{15}的系数得：

$$P\{\mathbf{Y}=15\}=\frac{1}{6}\left[C_{14}^{10}\times 1+C_5^1\times C_8^4\right]=0.084$$

（具体计算过程略）

由此看来，利用母函数解决一些古典概型中的问题很方便。