概率与测度入门

K.R.Parthasarathy 著

朱宗仁 译

武汉师范学院科研生产处

一九八三年三月

译 者 的 话

本书是应我系七七届开设选修课之急而于一九八一年下半年译出，并作为七七〇三班《概率与测度》选修课教材试用了一次。稍后又由七七届毕业生高付清、游春林、李柏林、金雁鸣、刘鸿翔等同志分别阅读了油印译本的有关章节，提出了有益的建议。在这次铅印之前，又特地请武汉大学数学系胡迪鹤教授对译文中某些部分进行了审阅，并汲取了他中肯宝贵的修改意见。我系二位主任陈星垣付教授和肖兢择付教授，科研处处长张志强同志特别是李成文院长对译作给予了充分的支持和热情的鼓励，在此谨向各位领导、教授和同志表示衷心的谢意。由于本人外语和专业水平有限，错误之处实恐难免，诚恳欢迎大家批评指正。

一九八二年十二月于武汉师范学院

序　言

1902年法国数学家H勒贝格写了著名论文“积分，长度和面积”。从1914年以来勒贝格测度的理论已经成为世界一切科学技术进步国家大学分析课程的一部分。1933年俄罗斯数学家A.N.柯尔莫哥洛夫写了名著“概率论基础”。在该书中他给出了概率论的基本公理。1950年P.R.哈尔莫斯的“测度论”和W.费勒的“概率论及其应用”的问世，使得世界各处从事数学研究的大学生和研究生都能接受这两门学科。

业经写成的此书希望能够在大学生与研究生的教学计划中把测度论与概率论二者作一学年的课程来开设。因为在概率论学习的高级阶段依赖于测度论的知识。所以我们试图把这两门学科结合起来成为单独的一册。

本书的材料是由作者向理科硕士讲授用的讲义发展而成的。他们是孟买大学数学系高级研究中心的研究生。统计学硕士、德里统计学院的研究生和理科硕士。德里印度工业学院的研究生。

本书分成八章。第一章讨论组合概率和Poisson与Loplace Demoivre古典极限定理。它们为把测度由布尔代数扩张到σ—代数提供了动力，第二章专门致力于布尔半—代数和拓扑意义下重要的子集类到σ—代数的测度扩张，第三章从事于由测度空间到可分距离空间中的波雷尔映象性质的讨论。特别是证明了Lusin定理和同构定理，波雷尔空间投影极限的测度扩张也作了研究，第四章讨论了积分，讨论了其积分为优良函数空间上唯一线性运算的Reisz表示定理以及由测度空间所产生的函数空间的性质，第五章讨论了乘积空间上的测度及转移测度，RK中的勒贝格测度，勒贝格积分的变量替换公式以及无穷可微函数的构造也都作了讨论，作为本书最长的第六章用正交投影法而不是习惯上所用的Radon-Nikodym定理法引进了条件期望的概念。Rodon-Nikodym定理和Lebesgue分解当作Von Neumeann的更一般的分解定理的推论而导出。各种意义下的条件期望的收敛性，正则条件概率的概念，遍历性定理和遍历性分解在这一章也都作了处理。第七章对于概率测度的弱收敛和特征函数作了简单的介绍，最后一章引进了局部紧群上Haar测度的构造，和齐性空间上的不变测度与拟不变测度。对具有拟不变测度的群上的Mackey-Weil定理也作了证明。为学者着想书中还包括有一些练习。一方面是测度论与概率论之间的联系，另一方面是测度与概率论同各种课目诸如泛函分析、统计，遍历性理论等之间的联系，通过附注、例子和练习都作了简要的说明。

K.R.P.　1977于新德里

目　　录（上册）

第一章　布尔代数上的概率

§1　集合与事件

在概率论中着眼于一个统计试验的全部可能的基本的结果，并假定它们组成一个称为样本空间的集合X，X的点或元素称为基本结果，我们用一些例子来加以说明。

例1.1.　具有二个基本结果的统计试验是最简单的统计试验，例如，掷一枚结果为正面或反面的硬币，观察一个结果为男或女的新生小孩的性别，检验一个加工的零件是否是次品，等等。

在这些场合中我们用0和1表示其结果，通常分别称之为“失败”和“成功”。这种样本空间X恰好包含二个点，即0和1。

例1.2.　抛掷一颗骰子并观察其点数，骰子有六面而其可能的点数构成集合 $X=\{1, 2, 3, 4, 5, 6\}$。

例1.3.　不断抛掷一枚硬币直到第一次出现正面为止，并观察每一阶段的结果，若以H表示正面，T表示反面，上述试验的每一个基本结果就是一个形如TTT…TH的有限序列，样本空间由一切这样的序列组成。

例1.4.　随意地分一付纸牌并观察其从头到尾的次序，空间X由52！个排列组成。

例1.5.　观察某地的气温，其基本结果恰好是实数，因而样本空间是实直线。

例1.6.　观察盒内气体的压力和温度，这里的X可以假定为平面R^2。

例1.7　在固定的一个小时的时间内，观察大气的温度曲线，样本空间X可视作区间[0、1]上的一切连续曲线的集合。

设$A\subset X$是某个统计试验的样本空间X的任意一个子集，试验的实现导致对一个基本结果x——X的一个元素的观察，若$x\in A$，我们说事件A发生了，若$x\overline{\in} A$，我们说事件没有发生或等价地说X—A（A的余）已经发生，从实际的观点出发，并非对任何事件都感到兴趣的，如在上例1.5中，考虑“测得温度是一个超越数”这一事件是没有任何实际意义的，但是，“实测的温度在某个区间[ab]内”这类事件是有意义的。我们可把以上的讨论概括如下：有一个由样本空间的一些子集所成的类**F**，对应于**F**的那些无素的事件是有“实际意义”的，我们假定这样一个由样本空间X的一些子集或一些事件所成的类**F**已被明确规定，我们简单地说**F**是与样本空间为X的某个统计试验有关的一切事件所成的类，所谓一个事件，我们理解为**F**的一个元素。

现在我们来分析一下由一切事件所成的类或族**F**应该满足什么样的条件。设$A\subset B\subset X$且使A、$B\in$**F**。若$x\in A$则$x\in B$，换言之，当A发生时，B也发生，这样，集合论的包含等价于蕴含这个逻辑的概念。

设A、$B\in$**F**，考虑集合$A\cup B$，$A\cap B$和X—A，提到事件A，B之一发生，等价地说，就

是试验产生一个属于$A\cup B$的观察结果，这自然要求$A\cup B$也是一个事件，A与B二者同时发生，理解为试验的观察结果x属于$A\cap B$，A不发生理解为x属于X—A。因此，自然地要求**F**对于有限并，有限交和取余的运算封闭，再假定全空间X而因此它的余即空集ϕ也属于**F**，那就没有什么遗漏了，以上所述使我们引入下面的定义。

定义1.8　称X的一些子集所成的类**F**为布尔代数，如果它满足下列条件：

① 若A、$B\in$**F**，则$A\cup B\in$**F**，$A\cap B\in$**F**；

② 若$A\in$**F**，则余$X-A\in$**F**；

③ 空集ϕ和全空间X属于**F**。

附注、1.9　今后在本书中对A的余X—A将记为A'，对X的任意两个子集A、B的交集$A\cap B$将记为AB。AB'记为A—B而对称差$(A-B)\cup(B-A)$记为$A\Delta B$。

例1.10　设X为任意一个非空集，**F**是由X的一切子集所成的类，那么**F**是一个布尔代数。

例1.11　设X＝R是实直线，族g定义为

$g=\{$形如$(-\infty,+\infty)$，$(-\infty,a)$，(a,∞)，(a,b)的一切区间，其中$a,b\in R\}$，

那么类

$$\mathbf{F}=\{A:A\subset R,\ A=\bigcup_{i=1}^{n}A_i,\ A_i\in g,\ A_i\cap A_j=\phi\quad \text{当}i\neq j\text{，对某个正整数}n\}$$

是一个布尔代数。（这里将空集视作当$b\leqslant a$时的区间(a,b)）

例1.12　设γ是任意一个集，X为元素取自γ的一切序列所成的空间，即任意$x\in X$均可写成$x=(y_1,y_2,\cdots)$，其中$y_i\in\gamma$对每一个$i=1,2,\cdots$，令A为k重笛卡尔乘积$\gamma\times\gamma\times\cdots\times\gamma$的任意一个子集，那么形如

$$C=\{x=(y_1,y_2\cdots);\ (y_{i_1},y_{i_2},\cdots,y_{i_k})\in A\}$$

的一个子集$C\subset X$称为K维柱集，（其中$i_1<i_2<\cdots<i_k$是一些固定的正整数）那么由一切有限维柱集所成的类**F**是一个布尔代数。

追溯集合论的术语和事件的术语之间的联系，把我们的讨论概括成一张表的形式。令**F**是某个统计试验的样本空间的一些子集组成的一个布尔代数，使得**F**是由一切事件所成的类，那么我们有以下的对应关系。

事件的术语	集合论的术语
A是一个事件	$A\in$**F**
事件A蕴含事件B	$A\subset B$
事件A不发生	A'
事件A，B中有一个发生	$A\cup B$
事件A和B两者都发生	AB
必然发生的事件	X
不可能发生的事件	ϕ
事件A和B不能同时发生	$A\cap B=\phi$

§2 布尔代数上的概率

考虑一个统计试验其基本结果由样本空间X以及X的一些子集所成的布尔代数**F**所描述。若将试验重复n次所得到的基本结果为$x_1, x_2, \cdots, x_n \in X$，设$A \subset X$是**F**的一个元素，令

$$p_n(A) = m(A)/n$$

其中$m(A)$是属于集合A的基本结果x_i的个数。数$p_n(A)$称为在已知的n次试验中事件A发生的频率，首先我们看到$A \to p_n(A)$是一个由**F**到单位区间$[0, 1]$中的映象，其次显然有

（Ⅰ）$p_n(A \cup B) = p_n(A) + p_n(B)$，若$A \cap B = \phi$；A，$B \in$ **F**

（Ⅱ）$p_n(X) = 1$。

由性质(Ⅰ)可推得，

$p_n(A_1 \cup A_2 \cup \cdots \cup A_k) = \sum_{i=1}^{n} p_n(A)$，若$A_i \cap A_j = \phi$，对一切$i \neq j$和$A_1, A_2, \cdots, A_k \in$ **F**，我们说p_n是一个在**F**上使$p_n(X) = 1$的非负有限可加函数，如果在观察结果的$x_1, x_2, \cdots$的发生中存在“统计规律性”，那么我们可以指望，对于$A \in$ **F**，$p_n(A)$将稳定于一个数$P(A)$。若果真如此，那么映象$A \to P(A)$将同样具有性质(Ⅰ)和(Ⅱ)。基于这些考虑，诱导我们引进下面的定义。

定义2.1　设$\{A\alpha, \alpha \in I\}$是集X的一些子集所成的类，其中I是某个指标集，如果当$\alpha, \beta \in I$，$\alpha \neq \beta$时，$A_\alpha \cap A_\beta = \phi$，那么称这样的一个类是两两不相交的。

定义2.2　设**F**是集X的一些子集所成的布尔代数，一个映象m：**F**$\to [0, 1]$如果当A，$B \in$ **F**，$A \cap B = \phi$时，有$m(A \cup B) = m(A) + m(B)$，则称m为有限可加的，如果对任何属于**F**的两两不相交的序列$\{A_n\}$，若$\bigcup_{n=1}^{\infty} A_n \in$ **F**，有

$$m\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} m(A_n)$$

则称m为可列可加的。如果映象p：**F**$\to [0, 1]$是有限可加的，并且$p(X) = 1$，则称p为**F**上的概率分布。

我们现在介绍一些例子。

例2.3　设X是一个有限或可列集，**F**是由X的这一切子集所成的布尔代数，将X的一切点列举成$\{x_1, x_2 \cdots\}$，又设$\{p_1, p_2 \cdots\}$为一个非负数的序列，对任何A，令

$$m(A) = \sum_{j: x_j \in A} p_j$$

那么显然可见m是**F**上的一个可数可加的函数，如果$\sum_{j: x_j \in X} p_j = 1$，则m是**F**上的一个概率分布。

例2.4　设F是定义在直线R上的单调增加函数，令

$$m((a, b]) = F(b) - F(a)，若a < b，a、b \in R,$$

记 $F(+\infty)=\lim_{a\to+\infty}F(a)$，$F(-\infty)=\lim_{a\to-\infty}F(a)$，规定

$$m((-\infty,\ a])=F(a)-F(-\infty),$$
$$m((b+\infty))=F(\infty)-F(b),$$
$$m((-\infty,\ +\infty))=F(+\infty)-F(-\infty)$$

那么m是一个定义在由一些区间所成的族g上的有限可加集函数（例1.11），即

$$m\left(\bigcup_{j=1}^{k}I_j\right)=\sum_{j=1}^{k}m(I_j)$$

其中I_1，I_2，…，I_k和$\bigcup_{j=1}^{k}I_j$属于g并且族$\{I_j,\ I\leqslant j\leqslant k\}$是两两不相交的。今设A是形如

$$A=\bigcup_{r=1}^{k}I_r$$

的任意一个集，其中I_1，I_2，…I_k属于g并且两两不相交，定义

$$\widetilde{m}(A)=\sum_{r=1}^{k}m(I_r)。$$

现在产生一个$\widetilde{m}$是否已被妥贴的定义的问题，因A完全可能还表成

$$A=\bigcup_{S=1}^{l}F_S \tag{2.2}$$

其中F_1，F_2，…F_l属于g并且两两不相交。这样A有两个表达式（2.1）和（2.2），但仍有

$$\sum_{r=1}^{k}m(I_r)=\sum_{S=1}^{l}m(F_s) \tag{2.3}$$

事实上，注意族g对有限交运算的封闭性，我们有

$$I_r=I_r\cap A=\bigcup_{s=1}^{l}(I_r\cap F_s),$$

$$F_s=F_s\cap A=\bigcup_{r=1}^{k}(F_s\cap I_r)。$$

因为在g上m是可加的，由此推得

$$m(I_r)=\sum_{s=1}^{l}m(I_r\cap F_s),$$

$$m(F_s)=\sum_{r=1}^{k}m(F_s\cap I_r),$$

而（2.3）就是以上两个等式的直接结果。上述讨论意味着，在可表成g中区间的并的有限个不相交的一切子集所成的布尔代数F上，$\widetilde{m}$是一个被妥贴定义了的有限可加的映象，换言之，

对应于R上的每一个单调增加函数F，可以在象例1.11中的布尔代数**F**上构造唯一的一个非负有限可加函数。如果

$$\lim_{\substack{b\to+\infty\\ a\to-\infty}}[F(b)-F(a)]=1$$

它就成为一个概率分布。

命题2.5 设m是一个在由X的一些子集所成的布尔代数**F**上的非负有限可加函数。如果$A\subset B$，A，$B\in$**F**，那么$m(A)\leqslant m(B)$。如果A_1，A_2，…，$A_k\in$**F**那么

$$m\ (\sum_{i=1}^{k}A_i)\leqslant\sum_{i=1}^{k}m(A_i)\qquad(2.4)$$

证明 为了证明第一部分，我们注意

$$B=A\cup(BA')\ \text{若}A\subset B。$$

因为**F**是一个布尔代数，A与BA'是属于**F**的不相交的子集，所以

$$m(B)=m(A)+m(BA')\geqslant m(A)。$$

为了证明第二部分，我们注意

$$\bigcup_{i=1}^{k}A_i=\bigcup_{i=1}^{k}B_i,$$

其中$B_1=A_1$，$B_2=A_1'$，…，$B_i=A_iA'_{i-1}\cdots A_1'\cdots$，$B_k=A_kA_{k-1}'\cdots A_1'$这样$B_1$，$B_2$，…，$B_k$是属于**F**的两两不相交的一些集合，并且$B_i\subset A_i(i=1,\ 2,\ \cdots,\ k)$，因此

$$m(\bigcup_{i=1}^{k}A_i)=m(\bigcup_{i=1}^{k}B_i)=\sum_{i=1}^{k}m(B_i)\leqslant\sum_{i=1}^{k}m(A_i)。$$

证毕。

附注、2.6 性质（2.4）是所谓有限次可加性，假定$\bigcup_{i=1}^{\infty}A_i$属于**F**，若m是可列可加的，则（2.4）对$k=\infty$仍然成立，其证明的方法相同。当$k=\infty$时，性质（2.4）称之为可列次可加性。

§3 概率分布与初等随机变量

考察一个统计试验，它的实现导致样本空间X的一个观察结果x，人们通常对确切的结果不感兴趣，感兴趣的是观察结果的一个函数。我们用一些例子来加以阐明。

例3.1 考察一个具有称之为“成功”和“失败”两个基本结果的特殊试验（参看例1.1），假定当成功出现时他得到一个卢比，当失败出现时他就失去一个卢比，于是他的增益可用定义为

$$f(0)=-1,\ f(1)=+1$$

的函数f表示出来，其中0和1表示结果为失败和成功。

例3.2 设想r个物品分布在n个房间中，假定物品彼此之间可以识别而来考察物品的分

配法，应该注意，可能不止一个物品占有一个房间，那么全部可能的分配法组成的样本空间X包含有n^r个点，对于每一个分配法$x\in X$，令$f(x)$为空房的数目。

例3.3 假定用枪发射一颗子弹，并设试验由观察子弹的弹道构成，对于每一个弹道x，令$f(x)$为子弹着落点的坐标。

从以上例子我们看到，这种函数的值依赖于随机的结果，因而这种函数的值在随机的意义上变化，在进一步讨论这方面的课题之前，我们将考虑在X上仅取有限个值的函数。

设$f: X\longrightarrow\gamma$是一个由样本空间到集合$\gamma$中的映象，**F**是我们将要在它上面考虑概率分布的，由X的一些子集所成的布尔代数，假定我们提出下面的问题：试验发生一个使得函数$f(x)$取已知值$y\in\gamma$的基本结果的概率是什么？考虑集合

$$\{x: f(x)=y\}=f^{-1}(\{y\})$$

倘若我们希望求得以上事件的概率，则必须使$f^{-1}(\{y\})\in\mathbf{F}$，由于这个原因，我们引进以下定义。

定义、3.4 设X是一个具有以其一些子集组成布尔代数**F**的样本空间，映象$f: X\to\gamma$称为一个γ—赋值的简单随机变量，如果f只取有限个值，并且对每一个$y\in\gamma$

$$f^{-1}(\{y\})\in\mathbf{F}$$

若γ是实直线，则称f为一个简单随机变量，我们用$S(X,\mathbf{F})$表示由一切简单随机变量所组成的集合。

对任何$A\subset X$，设

$$X_A(x)=1 \quad 若x\in A,$$
$$=0 \quad 若x\overline{\in} A。$$

那么X_A称为集合A的特征函数或示性函数，若$A\in\mathbf{F}$，X_A就是一个取0和1两个值的简单随机变量，若$a_1, a_2, \cdots, a_k$均为实数，$A_1, A_2, \cdots, A_k\in\mathbf{F}$，那么$\sum_{i=1}^{k}a_iX_{A_i}$是一个简单随机变量，反之，每一个简单随机变量能够表成使$A_i's$两两不相交的上述形式。显然，对一切A，$B\subset X$，有

$$\chi_{A\cup B}=\chi_A+\chi_B-\chi_{AB},$$
$$\chi_A\cdot\chi_B=\chi_{AB},$$
$$|\chi_A-\chi_B|=\chi_{A\Delta B}。$$

特别地，由此可推出一切简单随机变量所成的集$S(X, \mathbf{F})$，关于通常的加法，乘法运算和纯量乘法是一个代数。

定义、3.5 所谓布尔空间我们指的是一个二元对（X，**F**），其中X是一个集合，而**F**是由X的一些子集所成的布尔代数，所谓布尔概率空间我们指的是一个三元总体（X，**F**，P），其中（X，**F**）是一个布尔空间，而P是**F**上的一个概率分布，若S是（X，**F**）上的一个简单随机变量，P是**F**上的一个概率分布，我们用下面的数定义为S关于P的积分

$$\sum_i\alpha i\cdot P\ (S^{-1}\ (\{\alpha i\}))\ ,$$

其中求和遍历S可能取到的一切a_i，当P固定时，我们用符号$\int sdp$或简单地用E_s表示这个数，E_s也称s关于p的期望。

命题、3.6　若$s=\sum_{i=1}^{k}a_i\chi_{A_i}$，其中$A_1$，$A_2$，…$A_k$为**F**中的不相交集，$a_1$，$a_2$，…，$a_k$是实数，那么

$$\int s\,dp=\sum_{i=1}^{k}a_i\,p(A_i), \tag{3.1}$$

而且

（Ⅰ）对任意两个随机变量s_1和s_2以及任意两个实常数a和b，有

$$\int(as_1+bs_2)dp=a\int s_1dp+b\int s_2dp,$$

（Ⅱ）在**F**上按下式定义的函数

$$\mathbf{Q}(F)=\int s\chi_F dp,\ F\in\mathbf{F}$$

是有限可加的，即

$$\mathbf{Q}\left(\bigcup_{i=1}^{l}F_i\right)=\sum_{i=1}^{l}\mathbf{Q}(F_i)$$

其中F_1，F_2，…，F_l是**F**中的两两不相交的元素；

（Ⅲ）$\int s\,dp\geqslant 0$，若$p(\{x: s(x)<0\})=0$；

（Ⅳ）$\inf_{x\in X}s(x)\leqslant\int s\,dp\leqslant\sup_{x\in X}s(x)$。

证明，不失一般性，我们可设a_i'，s是各不相同的，而且$\bigcup_{i=1}^{k}A_i=X$，因而$s^{-1}(\{a_i\})=A_i$，s 的值域是集$\{a_1, a_2, \cdots, a_k\}$，于是等式（3.1）可由积分的定义直接得到。为了证明性质（Ⅰ），我们不妨设

$$s_1=\sum_{i=1}^{k}\alpha_i\chi_{A_i},\quad s_2=\sum_{j=1}^{l}\beta_j\chi_{B_j}$$

其中$A_1A_2\cdots$，A_k和B_1，B_2，…，B_l是属于**F**的不相交集，构成X的两个划分，那么

$$as_1+bs_2=\sum_{i=1}^{k}\sum_{j=1}^{l}(a\alpha_i+b\beta_j)\chi_{A_i}\chi_{B_j}$$

其中集A_iB_j构成X的另一个划分，于是

$$\int(as_1+bs_2)dp=\sum_i\sum_j(a\alpha_i+b\beta_j)p(A_iB_j)$$

$$=a\sum_i\alpha_i\{\sum_j P(A_iB_j)\}+b\sum_j\beta_j\{\sum_i P(A_iB_j)\}$$

$$=a\sum_i\alpha_i P(A_i)+b\sum_j\beta_j P(B_j)$$

$$=a\int s_1 dp+b\int s_2 dp。$$

这里我们应用了p的有限可加性以及

$$\bigcup_j A_i B_j=A_i(\bigcup_j B_j)=A_i X=A_i,$$

$$\bigcup_i A_i B_j=(\bigcup_i A_i)B_j=XB_j=B_j。$$

性质（Ⅱ）由性质（Ⅰ）推出，性质（Ⅲ）和（Ⅳ）直接由等式（3.1）推出。

附注、3.7　应该注意，性质（Ⅱ）指出了一种借助积分由一个已知的有限可加 函 数，在F上构造新的有限可加函数的方法。

我们现在应用期望概念及其在命题3.6中所叙述的性质来证明某些基本结果。

命题3.8　设A_1，A_2，…，A_n是X的一些子集，并且

$$B=\bigcup_{i=1}^{n}A_i，那么$$

$$\chi_B=\sum_{i=1}^{n}\chi_{A_i}-\sum_{1\leqslant i<j\leqslant n}\chi_{A_iA_j}+\cdots$$

$$+(-1)^{r-1}\sum_{1\leqslant i_1<i_2<\cdots<i_r\leqslant n}\chi_{A_{i_1}A_{i_2},\cdots,A_{i_r}}+\cdots$$

$$+(-1)^{n-1}\chi_{A_1A_2\cdots A_n},\tag{3.2}$$

其中$A_{i_1}A_{i_2}\cdots\cdots A_{i_r}$代表乘积$\bigcap_{j=1}^{r}A_{i_j}$

证明：若x不属于任何一个A_i，那么$\chi_B(x)=0$，并且等式（3.2）右边的每一项为零。若x恰好属于A_1，A_2……A_n中的k个集合；那么右边的第r项为（$-1)^{r-1}\binom{k}{r}$，若$r\leqq k$。而在其它情况下为0于是等式（3.2）右边为

$$\sum_{i=1}^{K}(-1)^{r-1}(r)=1-(1-1)^K=1$$

又$\chi_B(x)=1$，从而等式（3.2）始终成立，证毕

推论3.9　若P是(X，F)上的一个概率分布，那么对任何A_1，A_2……$A_n\in$F。

$$P(\bigcup_{i=1}^{n}A_i)=\sum_{r=1}^{n}(-1)^{r-1}s_r\tag{3.3}$$

其中 $$S_r=\sum_{1\leqslant i_1<i_2<\cdots<i_r\leqslant n}p(A_{i_1}A_{i_2}\cdots A_{i_r})\tag{3.4}$$

证明：将等式（3.2）两边积分，再由命题3.6即可推出等式（3.3）

命题3.10　设A_1，A_2……A_n是X的任意子集，B是恰好属于A_1，A_2……A_n中k_n个集的一切点所组成的子集，那么

$$\chi_B=\sum_{r=k}^{n}(-1)^{r-k}\binom{r}{k}\left[\sum_{1\leqslant i_1<i_2<\cdots<i_r\leqslant n}\chi_{A_{i_1}A_{i_2}\cdots A_{i_r}}\right] \tag{3.5}$$

证明：令x表示恰好属于A_1，A_2，…A_n中m个集的一个点，若m＜k，则$\chi_B(x)=0$，而且等式（3.5）右边的每一项都等于0，若m＝k。$\chi_B(x)=1$。当r＝k时在等式（3.5）右边方括号内的某一项为1，其余各项为0，因而右边也等于1。今设m＞k，则$\chi_B(x)=0$，右边为

$$\sum_{r=k}^{m}(-1)^{r-k}\binom{r}{k}\binom{m}{r}=\binom{m}{k}(1-1)^{m-k}=0$$

于是等式（3.5）始终成立，证毕。

推论3.11，设P是（X.F）上的一个概率分布，A_1，A_2，…A_n是属于F的n个事件，并设P_K是$A_1A_2\cdots A_n$中恰好有K个事件发生的概率，那么

$$P_k=\sum_{r=k}^{n}(-1)^{r-k}\binom{r}{k}s_r, \tag{3.6}$$

其中s_r由等式（3.4）所定义。

证明：以上结果可以通过将等式（3.5）两边积分并应用命题（3.6）而得到。

例3.12，设X是由整数1、2、…，N的一切排列组成的集合，F是X的一切子集所成的布尔代数，而P是对每一个排列赋予相同概率$\frac{1}{N!}$的概率分布（参看例2.3）令A_i表示让i固定的一切排列所成的集。在这种情况下，

$$P(A_{i_1}A_{i_2}\cdots A_{i_r})=\frac{(N-r)}{N!}，若i_1<i_2<\cdots<i_r$$

$$因而s_r=\binom{N}{r}\frac{(N-r)!}{N!}=\frac{1}{r!}$$

由推论3.9，说明至少让i's中一个固定的“随机”排列的概率可用表达式

$1-\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{(-1)^{N-1}}{N!}$给出，应用推论3.11，可以证明1、2、…N中恰有m个元素固定的随机排列的概率可用表达式

$$P_m=\sum_{r=m}^{N}(-1)^{r-m}\binom{r}{m}\frac{1}{r!}$$

$$=\frac{1}{m!}\left[1-\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{(-1)^{N-m}}{(N-m)!}\right] \tag{3.7}$$

给出，我们还要看到当N→∞时，P_m收敛于$\frac{e^{-1}}{m!}$。对每一个固定的N，考虑从1，2，…N的一切排列组成的集合X中选取一个随机排列的试验，令f（x)表示集合(1、2、…、N)中借排列x使之固定的元素数目，那么f就是一个取0、1、2、N诸值的简单随机变量，于是对m⩽N，f(x)＝m 的概率由等式（3、7）给出。当N→∞时，我们得到一个在一切非

负整数集上使得取m的概率为$\frac{e^{-1}}{m!}$的分布，这是列为我们今后研究对象的许多极限定理的一个特例，这个特殊的极限分布又是泊松分布的一个特殊情况。

例3.13，假定r个物品它们彼此可以识别，随机地分配到n个房间，试验由观察其分配法构成，因为任意一个物品可以占有n个房间中的任何一个房间，故有n^r个种可能的分配法，于是样本空间X由n^r个点组成，设F是X的一切子集所成的类，A_i是第i个房间空着的一切分配法组成的集合，所谓物品的随机分布我们理解为一切分配法有相等的概率n^{-r}，现在我们问这样的问题：恰有K个房间空着的概率是什么？记这个概率为P_k。若

$$i_1<i_2<\cdots<i_j，那么$$

$$P(A_{i_1}A_{i_2}\cdots A_{i_j})=\frac{(n-j)^r}{n^r}$$

按等式（3.4）的记法

$$S_j=\binom{n}{j}\left(1-\frac{j}{n}\right)^r$$

由推论3.11

$$P_k=\sum_{j=k}^{n}(-1)^{j-k}\binom{j}{k}\binom{n}{j}\left(1-\frac{j}{n}\right)^r \tag{3.8}$$

如前例就任何一次观察到的分配法x而言，空房的数目可设想为一个取值0，1，2，…，（n－1）的简单随机变量，这个随机变数取K值的概率等于P_k，由等式（3.8）给出，它的期望为$\sum_{k=0}^{n-1}KP_k$

可以看到，P_k是r和n即物品和房间数的函数，人们会问这样的问题：当r和n按某种适当的方式趋于∞时会发生什么情况？在这种情况下对于一个固定的K，P_k是否收敛于一个极限？

在等式（3.8）中令j－k＝s，用一种初等的计算即可证明，

$$P_k=\frac{1}{k!}\sum_{s=0}^{n-k}\frac{(-1)^s}{s!}\Big[\left(1-\frac{s+k}{n}\right)^r n^{s+k}$$

$$\left(1-\frac{s+k-1}{n}\right)\left(1-\frac{s+k-2}{n}\right)\cdots 1\Big] \tag{3.9}$$

因为当$0\leqslant x<1$时，$1-x\leqslant e^{-x}$，故有

$$\left(1-\frac{s+k}{n}\right)^r n^{s+k}\leqslant(ne^{-r/n})^{s+k}。$$

因而等式(3.9)中方括号内的项被$(ne^{-r/n})^{s+k}$所控制。现在令r和n按$ne^{-r/n}\to\lambda>0$这种方式趋于∞，于是对一切充分大的r和n（3.9）总和中的每一项的绝对值不大于$\frac{(\lambda+1)^{s+k}}{s!}$，

因为级数$\sum_{s=0}^{\infty}\frac{(\lambda+1)^s}{s!}$收敛，我们可以在等式（3.9）的求和号内取极限，又因

$$\lim_{n\to\infty}[(1-\frac{x}{n})^n e^x]^{\log n}=1,\quad 对一切x\geqslant a$$

和

$$\lim_{r,n\to\infty}(\log n-\frac{r}{n})=\log\lambda,$$

故

$$\lim_{r,n\to\infty}(1-\frac{s+k}{n})^r n^{s+k}$$

$$=\lim_{r,n\to\infty}[(1-\frac{S+k}{n})^n e^{s+k}]^{r/n}\ e^{(s+k)}(\log n-\frac{r}{n})$$

$$=\lambda^{s+k}$$

于是

$$\lim_{r,n\to\infty}P_k=e^{-\lambda}\frac{\lambda^k}{K!}$$

这个在一切非负整数集上对单点集$\{k\}$具 有 概 率$e^{-\lambda}\frac{\lambda^k}{K!}$的分布称为具有参数$\lambda$的泊松分布。

例3.14，假定现在把r个不可识别的物品随机地放置到n个房间里，在这种情形一个分配法对应一个矢量（x_1、x_2、$\cdots x_n$）其中x_i表示i号房间中物品的个数，假设X是一 切 可能的分配法组成的样本空间，现在要求出x中点的个数。我们可把一个分配法看作 如 下的一个图

$$\underbrace{0\ 0\cdots 0}_{x_1}\ |\ \underbrace{0\ 0\cdots 0}_{x_2}\ |\ \underbrace{\cdots\cdots 0\ 0\cdots 0}_{x_n}$$

这里 0 表示物品而 | 表示房间的墙，因为有n个房间，故上图中的垂直线段有$n-1$条，被物品或线段占据位置的总数有$n-1+r$个，此外，物品可以遍取r个位置。于是分配 法 的总数为 $\binom{n-1+r}{r}$ ，若一切分配法是等可能的，那么X的每一个点有概 率$\frac{1}{\binom{n-1+r}{r}}$。这个分布以波译——爱因斯坦统计学而著称。上述分布由物理学家S、N、波译从量子统计力 学 的角度所发现，它是波译——爱因斯坦统计学理论中的一个基本分布。

在上面的问题中，如设A_i表示i号房间空着这一事件，那么对$i_1<i_2<\cdots<i_j$，$A_{i_1}A_{i_2}\cdots A_{i_j}$就表示$i_1$，$i_2$，$\cdots$，$i_j$这j个房间空着的事件，这种分配法的数目等于r个相同物品可能被分布到$n-j$个房间里的方法数，

于是

$$P(A_{i_1}A_{i_2}\cdots A_{i_j})=\binom{n-j+r-1}{r}/\binom{n+r-1}{r}$$

若以P_k表示恰好K个房间空着的概率，则

$$P_k=\sum_{j=k}^{n}(-1)^{j-k}\binom{j}{k}\binom{n}{j}\binom{n-j+r-1}{r}\Big/\binom{n+r-1}{r} \tag{3.10}$$

练习3.15，在(3.10)中，若n，$r\to\infty$且$\frac{n^2}{r}\to\lambda$ 则 $P_k\to e^{-\lambda}\frac{\lambda^k}{k!}$

（$k=0，1，2，\cdots\cdots$）

例3.16，考虑一个装有m个白球和n个黑球一个坛子，随机不返回地选取K个球的一个样本，随机选取的意义是从全部m＋n个球中选取的任何K个球的一个样本的概率是相同的，而因此等于$\frac{1}{\binom{m+n}{k}}$

现在问这样的问题，在样本中出现r个白球的概率是什么？如果在某个样本中有r个白球，那么在同一样本中有k－r个黑球，这样的选法有$\binom{m}{r}\binom{n}{k-r}$种，若用$P_r$表示所求的概率则

$$P_r=\frac{\binom{m}{r}\binom{n}{k-r}}{\binom{m+n}{k}},\quad 0\leqslant r\leqslant \min(k、m)$$

这个在由整数 0、1、2、…… min（k、m）组成的集合上的分布，即所谓超几何分布 。

练习3.17，考虑一个装有r种不同颜色球的坛子。设第一种颜色的球有n_1个，第二种颜色的球有n_2个，等等，如果随机不返回地抽抽取K个球，那么得到m_1个第一种颜色 的 球，m_2个第二种颜色的球，……以及m_r个第r种颜色的球的概率等于

$$P_{m_1、m_2\cdots m_r}=\frac{\binom{n_1}{m_1}\binom{n_2}{m_2}\cdots\binom{n_r}{m_r}}{\binom{n_1+n_2+\cdots+n_r}{m_1+m_2+\cdots+m_r}}\qquad \begin{array}{l}0\leqslant m_j\leqslant\min(k,n_j)\\ m_1+m_2+\cdots+m_r=k\end{array}$$

练习3.18，在上面的练习中，设n_1、$n_2\cdots n_r\to\infty$，且

$$\lim\frac{n_i}{n_1+n_2+\cdots+n_r}=p_i\ (i=1、2、\cdots r)\ 则$$

$$\lim P_{m_1,m_2,\cdots m_r}=\frac{k!}{m_1!\,m_2!\cdots m_r!}P_1^{m_1}\ P_2^{m_1}\ P_r^{m_r} \tag{3.11}$$

其中$K=m_1+m_2+\cdots+m_r$（由等式（3.11）右边所列举的概率决定一个多项分布）。

附注3.19。例3.12和3.13以及练习3.15和和3.18给出了概率论中极限定理的特殊情况，这些极限定理的作用在于它们可使实际的计算更加容易。往后我们将用一些例子加以解释 。

定义3.20，设S_1和S_2是布尔概率空间（X・**F**，P）上的两个简单随机变量，则称数E $E(S_1-ES_1)(S_2-ES_2)$为S_1和S_2的协方差，并表示为Cov（S_1，S_2）若$S_1=S_2=S$，则cov（S,S）称为S的方差，并表示为V（S），这样 V(S)≧o，而$V(S)^{\frac{1}{2}}$称为S的标准差并表示为σ(S)。

练习3.21 V(S)＝0当且仅当在概率为1的集合上S等于 常 数E_s。cov（s_1，s_2）＝$Es_1s_2-Es_1\cdot Es_2$