第一章 概率论的基本概念

自然界发生的现象有两类：确定性现象、随机现象。

在大量重复试验下，随机事件的结果呈现出某种规律性。这种在大量重复试验中所呈现出的固有规律性，就是统计规律性。

随机现象：在个别试验中其结果呈现出不确定性，在大量重复试验中其结果又具有统计规律性的现象。概率论与数理统计是研究和揭示随机现象统计规律性的一门数学学科。

随机试验：满足以下特点的试验

（1）可以在相同的条件下重复地进行

（2）每次试验的可能结果不止一个，并且能事先明确试验的所有可能结果

（3）进行一次试验之前不能确定哪一个结果会出现

样本空间：随机试验E的所有可能结果组成的集合。记为S。

样本点：E的每个结果，S的每个元素

随机事件：S的子集，E的随机事件，简称事件

基本事件：由一个样本点组成的单点集

必然事件：S

不可能事件：

事件间的关系与事件的运算：事件是集合，可用集合论去理解。P3

（1）A包含于B，A与B相等

（2）和事件 A+B

（3）积事件 AB

（4）差事件 A-B

（5）相交为空 互不相容 互斥

（6）逆事件 对立事件

事件运算时，经常要用到下述定律： P4

（1）交换律 和、积

（2）结合律 A、B、C的和或积

（3）分配率 先和后积、先积后和

（4）德摩根律

对于一个随机试验，某个随机事件可能发生，也可能不发生，我们常常希望知道它发生的可能性的大小，称为概率。

频率：在相同条件下，进行了n次试验，事件A，频数，频率。

由定义知，频率具有下述特点：

（1）[ 0, 1 ]

（2）S事件的频率为1

（3）两两互不相容 和事件的频率等于事件频率的和

直观的想法是用频率来表示事件A在一次试验中发生的可能性大小。

是否可行呢？抛硬币试验 P6

大量试验证实，当重复次数足够多时，频率呈现出稳定性，逐渐稳定于某个常数。频率稳定性，统计规律性。我们可以用大量试验次数下的频率来作为概率。但在实际中，我们不可能对每一个试验都要大量重复，然后求得相关试验的频率。我们需要有一种数学方法来定义和计算概率。

概率：设E为随机试验，S为其样本空间，对于E的每一个事件A赋予一个实数，记为P(A)，称为事件A的概率，如果集合函数P(.)满足下列条件：

（1）非负性 P(A) >= 0

（2）规范性 P(S)=1

（3）可列可加性 两两互不相容 和事件的概率等于事件概率的和

第5章中，将证明当n趋于无穷大时，频率在一定意义下接近于概率P（A）

则可用P(A)来表征事件A在一次试验中发生的可能性大小（概率）

频率的性质

（1）P（）=0

（2）有限可加性 两两互不相容 和的概率等于概率的和

（3）事件A、B，差的概率的概率的差

（4）P(A) <= 1

（5）逆事件的概率

（6）加法公式 推广到多个事件的情况 P9

等可能概型（古典概型）：具有两个特点的试验

（1）样本空间只含有限个元素

（2）每个基本事件发生的可能性相同

由等可能概性的概念和概率的性质，易证，P（基本事件）=1/n

若事件A包含k个基本事件，P（A）=k / n

随机选取n个人，n<=365，假设每人的生日在一年365天中的任一天是等可能的，即都等于1/365，那么他们的生日各不相同的概率 P13

超几何分布 P12

袋中有a只白球、b只红球，k个人依次在袋中取一只球，不放回抽样，求第i个人取到白球的概率？ k<=a+b

与福利彩票类似，虽然取球的先后次序不同，但各人取到白球的概率是一样的。

概率很小的事，在一次试验中实际上几乎是不发生的，称之为：实际推断原理。 P14

条件概率 P（B | A）： 事件A已发生的情况下，事件B发生的概率。

对于一般的古典概型问题，易证：P(B|A) = P(AB)/P(A)

在一般场合，我们将上述公式作为条件概率的定义。易证，条件概率符合概率定义的三个条件：非负性、规范性、可列可加性。

由条件概率的定义式，可得乘法公式：P(AB)=P(B|A)P(A)

将其推广到多个事件的积事件。P16 例3、例4

样本空间的划分：设 为E的一组事件，若其满足 (i), (ii)，则称其为样本空间S的一个划分。

全概率公式： P(A)=… P18

很多实际问题中，P(A)不易求得，但却容易找到S的一个划分，且P(Bi)、P(A|Bi) 易求。

贝叶斯公式：P(Bi|A) = … P18

贝叶斯公式的理解：A事件的发生可由B1, … , Bn事件导致，现已知A事件发生，欲求其发生是由Bi导致的概率。或者理解为贡献。

关于全概率公式和贝叶斯公式的应用：P19 例5

P20 例8

独立性：设A，B是两事件，如果满足P(AB)=P(A)P(B)，则称事件A，B相互独立。

见P20 例1

定理1：若A，B独立，P(A)>0，则P(B|A)=P(B)

定理2：若A，B独立，则A与B非，A非与B，A非与B非，也相互独立

可将独立性的概念推广到3个事件、n个事件。

推论1：若A1, … , An相互独立，则其中任意k个事件也是相互独立的。

推论2：若n个事件相互独立，则将他们中任意多个时间换成各自的对立事件，所得的n个事件仍相互独立。

两事件相互独立的含义是它们中一个已发生，不影响另一个发生的概率。在实际应用中，对于事件的独立性，常常是根据事件的实际意义去判断。感冒的例子。 P22 例2

第二章 随机变量及其分布

随机试验，将样本空间的每个元素和数对应起来，映入随机变量。

定义：设随机试验的样本空间为S={e}。X=X(e) 是定义在样本空间S上的单值函数，称X=X(e)为随机变量。

随机变量的引入，使我们能用随机变量来描述各种随机现象，并能用数学分析的方法对其进行研究。

离散型随机变量：取值是有限个或可列无限多个。

非离散型随机变量，包含连续随机变量和其它我们不研究的变量。

离散型随机变量X的所有可能取值为（k=1,2,3…），X取各个可能值的概率为：

P{X=}=，k=1,2,3…

由概率的定义，满足如下两个条件：

（1）非负性

（2）规范性

称此公式为离散型随机变量X的分布律。分布律也可以用表格的形式来表示。P32 （2.4）

可以想象成：概率1以一定的规律分布在各个可能值上，这就是（2.4）被称为分布律的原因。

三种重要的离散型随机变量

（1） （0-1）分布 / 两点分布

分布律：P{X=k}=，k=0,1 ( 0<p<1 )

写成表格形式： P33

新生婴儿的性别统计、某车间的电力消耗是否超过负荷。

（2）二项分布、伯努利试验

试验E有两个结果，称E为伯努利试验，P(A)=p，P()=1-p，将E独立重复地进行n次，称这一串重复的独立试验为n重伯努利试验。

有着广泛的应用，是研究最多的模型之一。

以X表示n重伯努利试验发生的次数，X是一个随机变量，求分布律，

P{X=k}=，k=0,1,2…,n

显然，符合非负性、规范性。

称X服从参数为n，p的二项分布，并记为X~b(n, p)

特别地，当n为1时，就是（0-1）分布。

二项分布的实例 P35 例2,3,4

（3）泊松分布

P37

对于连续型随机变量，其可能取的值不能一一列举出来，不能像离散型随机变量那样用分布律去描述它。连续性随机变量，取任一指定值的概率都等于0，而且在实际中，我们并不关心它取某个值的概率，而是它的取值落在一个区间的概率。

随机变量X的分布函数：

分布函数完整地描述了随机变量的统计规律性。它使得我们能用数学分析的方法来研究随机变量。具有两个性质：

（1）不减函数

（2），且，

（3）F(x+0)=F(x)，即F(x)是右连续的

反之，符合性质（1）（2）（3）的函数F(x)必是某个随机变量的分布函数。

对于随机变量X的分布函数F(x)，存在非负可积函数f(x)，使对于任意实数x，有

则称X为连续型随机变量，f(x)称为X的概率密度函数，简称概率密度。

连续型随机变量的分布函数是连续函数。

在实际应用中遇到的随机变量基本上是离散型或连续型，本书只讨论这两种随机变量。

概率密度具有以下性质：

（1）非负性 （2）规范性 全区间积分为1 （3）P{x1<x<=x2}=…

（4）若f(x)在点x处连续，则有

反之，若f(x)具备性质（1）（2），引入G(x)，它是某一随机变量X的分布函数，f(x)是X的概率密度。

概率密度的定义与物理学中的线密度相类似，这就是称f(x)为概率密度的原因。

X落在小区间（x，x+]上的概率近似地等于.

一个事件的概率为0，并不意味着它是不可能事件。

概率分布指的是：分布函数，或者分布律（离散型）、概率密度（连续型）。

三种重要的连续型随机变量

（1）均匀分布 X~U(a, b) 概率密度函数图形为一个矩形，分布函数图形为…

（2）指数分布 X~E( 概率密度为

称X服从参数为的指数分布。

分布函数为：…

无记忆性：P45 底部

指数分布在可靠性理论与排队论中有广泛的应用。

（3）正态分布 X~

P46

已知随机变量X的分布函数，求g(X)的分布函数。 P50

第三章 多维随机变量及其分布 P60

第四章 随机变量的数字特征

概率就像演绎，通过已知的先验(幂律分布、正态分布、厄兰分布)，推导出结果.  
统计就像归纳，通过小样本数据发现特征和规律，然后放在更大的样本里验证.