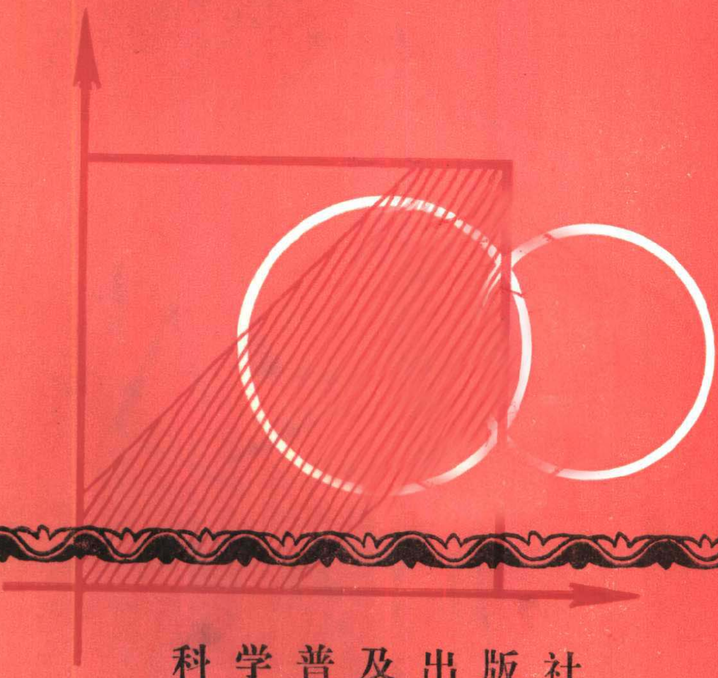
概率

鲜于方圣　编著



科学普及出版社

概　率

鲜于方圣　编著

科学普及出版社

内 容 提 要

概率论是数学的一个十分重要的分支，它在理论物理学、测地学、天文学、误差论、空间技术和自动控制等方面，都有广泛的应用。在概率论基础上建立起来的数理统计学，正直接地为生产计划和组织、工艺过程分析、质量检查和控制，以及许多其它目的服务。本书就是介绍有关概率论最基础知识的普及读物。

为便于读者理解和应用，书中举有大量的实例，题目均有答案或提示。

概 率

鲜于方圣 编著

责任编辑：吴之静

封面设计：王序德

•

科学普及出版社出版（北京海淀区白石桥路32号）

新华书店北京发行所发行 各地新华书店经售

保定科技印刷厂印刷

•

开本：787×1092毫米 1/32 印张：2 1/8 字数：44 千字

1984年6月第1版 1984年6月第1次印刷

印数：1—27，200册 定价：0.23元

统一书号：13051·1391 本社书号：0808

目　录

引　言

在现实世界中，有这样一类事件，它们在一组确定的条件下，一定会出现。比如，当我们用力向空中抛出一个球以后，球上升到最高点后落回地面的事件就一定会出现。又如，在标准大气压和温度15°C时，容器里的水就一定处于液体状态。象这样在一组确定条件实现之下必然出现的事件，我们称它为**必然事件**。

显然，现实中也有着这样一类事件，它们在一定条件下必然不会出现。比如，在标准大气压和温度15°C时，容器里的水处于固体状态的事件就一定不会出现。象这样一类事件就称为**不可能事件**。

除了必然事件和不可能事件以外，还存在着第三类事件，它们在一组确定的条件下，可能出现，也可能不出现。这样一类事件就称为**随机事件**。

比如，投掷一枚硬币，正面向上的事件可能出现，也可能不出现。因为投掷结果和许多因素有关，如硬币的初速度、硬币的形式、桌面的光洁度等等，而我们预先又不能估计到所有这些因素对结果的影响，因此，就无法预言单独一次的投掷结果。

下面就是随机事件的例子：

1．射击中靶。

2. 北京市年雨量大于900毫米。

3. 某电话总机在1分钟内接到呼唤不多于5次。

4. 从全副扑克牌中抽出黑桃3。

你也许会产生这样的印象：以为对于随机事件而言，并没有什么规律性存在。事实并不是这样的。在自然科学、生产实践和日常生活中，有着许多这样的类似现象：在同样一组条件大量重复之下，某一随机事件出现的次数往往有某种非常明确的规律性，即事件出现的百分率与某个常数很接近。比如，我们虽然无法预言单独一次投掷硬币的结果，但是，多次重复进行试验就会发现，硬币出现正面向上的事件的次数，和总投掷次数的$\frac{1}{2}$非常接近。

又如，从同一批产品中随意抽查1,000件产品，其中合格品个数往往总是与抽出产品总数的某一百分率很接近。

关于气体的分子运动，也有同样的情形。虽然在密闭容器中单个分子的运动方式是随机的，但是，就大量总体来说，每单位时间内撞到容器壁上的分子数目几乎保持一定的比例，因而气体对容器的压力是一个常数。

事实证明，这样一种大量现象所呈现的规律性，具有极其广泛的适用范围。概率论就是从事于确定这种规律性的研究。

也就是说，概率论是研究大量随机事件规律性的科学。

概率论是数学中一个十分重要的分支。概率论的理论与方法在科学技术的许多方面，如理论物理学、测地学、天文学、误差论、空间技术和自动控制等方面，都有着广泛的应用。在概率论基础上建立起来的数理统计学，正直接地为生

产计划和组织、工艺过程分析、质量检查和控制以及许多其它目的而服务。概率论思想渗透到科学技术的各个领域，已成为当代科学发展的一个特点。作为科学技术的重要工具，概率论正在蓬勃发展之中。

第一章 概率的概念

1.概率的古典定义

我们知道，在确定的一组条件实现时可能出现也可能不出现的事件称为随机事件，今后我们将把一组条件实现简单说成“进行试验”。这样，我们将把事件看作试验结果来研究。例如，射手向一个靶子进行射击时，射击就是一个试验，中靶和不中靶就是事件。

让我们考察这样一个例子：袋中有10个外形一样的球，其中有 9 个黑球、1 个白球。显然，从袋中随意取出一个球是黑球的事件是一个随机事件，取出一个球是白球的事件也是一个随机事件。我们称前一事件为事件A，后一事件为事件B。容易确信，事件A出现的可能性比事件B出现的可能性大，真可说是“十拿九稳”了。现在的问题是，我们应该取怎样一个数来刻画每一事件出现可能性的大小呢？

显然，在10种可能情形中，事件A有 9 种机会，事件B有 1 种机会。因此，很自然地，我们取对事件A有利的结果数（9），对于总结果数(10)的比$\frac{9}{10}$作为刻画事件A出现可能性大小的数，并称它为事件A的概率，而记作$P(A)=\frac{9}{10}$。

类似地，取 $P(B)=\frac{1}{10}$。

注意，在确定上述概率的时候，我们实际上应用了这样一个十分重要的假定：抽出10个球中任一个球的机会是相等的（等可能的）；并且，显然还认为，每次必定抽出10个球中的一个球，而抽出的球只能是黑球或白球，不能同时既是黑球又是白球。

现在，我们应用上述原则，解决下列几个例题。

例1 投掷一次硬币时，出现正面向上事件的概率是多少？

显然，出现正面和出现反面是等可能的，因而，所求概率应该是 $\frac{1}{2}$。

例2 一批产品中，甲工厂生产的有250件，乙工厂生产的有350件，丙工厂生产的有100件，求任意取出的一件产品是丙工厂产品的概率。

这里，总可能结果数是 $250+350+100=700$，每一结果是等可能的。对所研究事件（取出产品是丙厂产品事件）有利的结果数是100。因而，所求概率为 $\frac{100}{700}=\frac{1}{7}$。

我们已经掌握了确定最简单的一些随机事件的概率的方法。为了把这种最简单的情形确切规定起来，引入概率的（古典）定义，我们要先引入以下几个概念。

如果在同一试验中，两个事件不能同时出现，这两个事件就称为**互斥事件**（或**不相容事件**）。反之，能够同时出现的事件，称为**相容事件**。

例如，射击时，中靶与不中靶是互斥事件。但是，中靶

与命中环数不少于9环事件是相容事件。因为，命中9环（或10环）时这两个事件就同时出现。

如果一些事件中，无论哪一个出现的可能性不比另一事件出现的可能性大，这些事件就称为**等可能事件**。

例如，袋中有5个外形完全一样的球，颜色分别为红、黄、蓝、白、黑。从袋中摸出一球，此球为红球、黄球、蓝球、白球或黑球的事件为等可能事件。

如果诸事件中，必定有一个要在试验结果中出现，就称这一组事件构成**完备群**。

例如，射击三次，中靶1次、2次、3次和一次也不中靶四个事件构成完备群。

我们看到，在掷硬币时，出现正面向上事件与反面向上事件是互斥事件，也是等可能事件，并且这两个事件构成完备群。这样的事件，称为基本事件。

一般地，如果一组事件是等可能事件，其中任何两个事件是互斥事件，并且它们构成完备群，则称其中每一事件为**基本事件**。如果试验时，某一基本事件的出现导致随机事件A出现，则称此事件是对A有利的基本事件。

例如，从0、1、2、3、4、5、6、7、8、9十个数字中任意抽取一个数字，则出现以上各数字的每一事件都是基本事件。若我们关心的是出现数字表示大于2的偶数，则对这一事件有利的基本事件是抽出数字为4、6、8各个基本事件。

现在，我们来给出概率的（古典）定义。

设试验的基本事件总数为n，对事件A有利的基本事件数为m，则我们用数$\frac{m}{n}$来表达事件A出现可能性大小，并

称它为事件A的**概率**。即

$$P(A)=\frac{m}{n}.$$

容易验证，前面对于概率的计算都符合于上述概率定义。

从概率的定义，导出它的下列性质：

(1) 必然事件的概率等于1。

事实上，如果事件A是必然事件，那么，试验的每一个基本事件都是对A有利的基本事件。这时，$m=n$，因而

$$P(A)=\frac{m}{n}=\frac{n}{n}=1.$$

(2) 不可能事件的概率等于零。

事实上，如果事件A是不可能事件，那么，无论哪一个基本事件都不是有利于它的事件。这时，$m=0$，因而

$$P(A)=\frac{m}{n}=\frac{0}{n}=0.$$

(3) 随机事件的概率是零和1之间的正数。

事实上，对随机事件A有利的基本事件数只是试验基本事件总数的一部分。在这种情形，$0<m<n$，这表明$0<\frac{m}{n}<1$，因而

$$0<P(A)<1.$$

总的说来，任一事件的概率适合不等式

$$0\leqslant P(A)\leqslant 1.$$

现在，我们来研究直接利用概率的古典定义计算概率的一些例子。

例1 箱子里有40件正品，10件次品。从箱子里随意取出10件产品作质量检查。求抽出产品中：1)没有一件次品

（事件A）；2）恰有两件次品（事件B）的概率。

【解】 从50件产品中取出10件产品的每种不同取法看作一个基本事件。基本事件总数为C_{50}^{10}。对事件A有利的基本事件数是从40件正品中取出10件正品的方法数，即C_{40}^{10}。

因而，事件A的概率是

$$P(A)=\frac{C_{40}^{10}}{C_{50}^{10}}\approx 0.8$$

对事件B有利的基本事件数是$C_{10}^{2}\cdot C_{40}^{10-2}$，因而，事件$B$的概率是

$$P(B)=\frac{C_{10}^{2}\cdot C_{40}^{8}}{C_{50}^{10}}\approx 0.32.$$

例2 几个细胞被有害放射线照射后，染色体破裂成含"中心体"一边（称为长段）与另一边（称为短段）。这$2n$个小段随后又随机地结合成新的几对。若长段与短段相结合，细胞就成活；若两个长段相结合或两个短段相结合，细胞就会死亡。试求：

1） 这些细胞仍按原样结合（事件A）的概率；

2） 有n个细胞接活（事件B）的概率。

【解】 1）在这$2n$个小段中，任意取出的第一个小段和其余$2n-1$个小段结合时有$2n-1$种结合法，第二个小段和剩下的$2n-3$个小段结合时有$2n-3$种结合法，……，第n个小段和剩下一个小段有一种结合法。因此，$2n$个小段共有$(2n-1)(2n-3)\cdots 5\cdot 3\cdot 1$种结合法。这就是基本事件总数。其中，只有一种结合法（一个基本事件）是对事件A有利的。因此，事件A的概率是

$$P(A)=\frac{1}{1\cdot 3\cdot 5\cdots(2n-3)(2n-1)}=\frac{2^n\cdot n!}{(2n)!}$$

2） 可能结合数（基本事件总数）和1）相同。任取的第一个长段（或短段）和n个短段（或长段）有n种结合法，第二对长段和短段有$n-1$种结合法，……，第n对长段和短段有一种结合法。因此，对事件B有利的基本事件数等于$n(n-1)\cdots 2\cdot 1=n!$。所以，事件B的概率是

$$P(B)=\frac{n!}{1\cdot 3\cdot 5\cdots(2n-3)(2n-1)}=\frac{2^n}{C_{2n}^n}.$$

以下，我们引入一个现代统计物理学中的问题。

例3 （波尔兹曼统计学）设有n个不同的质点，每个都以同样的概率$\frac{1}{N}$落在$N(N>n)$个格子中的每一格（设每个格子能容纳的质点数没有限制），试求：

1） 某指定的n个格子中各有一个质点（事件A）的概率；

2） 任意n个格子中各有一个质点（事件B）的概率。

【解】

1）先计算基本事件总数。我们分n个步骤把n个质点分配到N个格子中去，每一步都有N种方法，共有不同方法数（基本事件总数）为$\underbrace{N\cdot N\cdots N}_{n\text{个}}=N^n$。

再计算有利事件个数。我们分n个步骤把n个质点分配到n个指定格子中去，第一步有n种方法，第二步有$n-1$种方法，……，第n步有一种方法。共有$n!$种方法。所求概率

$$P(A)=\frac{n!}{N^n}.$$

2） 从N个格子中，选取n个格子的方法数为C_N^n。这时，

有利事件个数为$C_N^n \cdot n!$，所求概率

$$P(B)=\frac{C_N^n \cdot n!}{N^n}=\frac{N!}{N^n(N-n)!}.$$

【注】 本题中，如果把n个质点看作是不可辨别的，则属于波泽-爱因斯坦统计学；如果假定每个质点不可辨别的同时，还假定每个格子只能容纳一个质点，则属于费米-狄拉克统计学。

以下问题显示概率对于统计的密切关系。

例4 设从某一湖里捕得1,000条鱼，涂以红点后放回。经过适当时间以后，再捕1,000条鱼，发现其中有100条是涂了红点的鱼。试估计湖中共有多少条鱼。

【解】 设湖中有n条鱼，第一次捕得n_1条鱼，都做上记号(本例中，$n_1=1000$)，第二次捕得r条鱼($r=1000$)，其中有记号的有k条($k=100$)。则第二次捕鱼中，出现恰有k条有记号鱼这一事件的概率为

$$p_k(n)=\frac{C_{n_1}^k \cdot C_{n-n_1}^{r-k}}{C_n^r}$$

我们求n，使这一概率为最大，并把这时的n（记为$\hat{n}$）作为估计数(称为n的最大似然率估值)。

由于

$$\frac{p_k(n)}{p_k(n-1)}=\frac{C_{n_1}^k \cdot C_{n-n_1}^{r-k}}{C_n^r}\cdot\frac{C_{n-1}^r}{C_{n_1}^k \cdot C_{n-1-n_1}^{r-k}}$$

$$=\frac{C_{n-n_1}^{r-k}}{C_{n-n_1-1}^{r-k}}\cdot\frac{C_{n-1}^r}{C_n^r}=\frac{n-n_1}{(n-n_1-r+k)}$$

$$\cdot\frac{n-r}{n}\left(\text{应用：}\ C_n^k=\frac{n}{n-k}C_{n-1}^k\right)$$

$$=\frac{n^2-nn_1-nr+n_1r}{n^2-nn_1-nr+nk},$$

因此，当$nk<n_1r$时这个比值大于1；当$nk>n_1r$时，比值小于1。这就是说，$p_k(n)$当$n<\frac{n_1r}{k}$时是n的增函数；而当$n>\frac{n_1r}{k}$时是n的减函数。所以，当n为不超过$\frac{n_1r}{k}$的最大整数时，它达到最大值。

在本例中：

$$\hat{n}=\frac{1000\times1000}{100}=10000.$$

进一步的研究，还可以合理地估计出n的真值范围。

2.概率的统计定义

概率的古典定义虽然可以用来解决许多简单的概率问题，但是，更多的事实却表明了这一概念的局限性。比如，在概率的古典定义里，假定了试验的基本事件个数是有限的，在实际上，我们却常常遇到这样一些试验，它的可能结果是无限的。在这种情形，概率的古典定义就不能应用。更重要的是，古典定义是从等可能性的假设出发的，而等可能性的假定往往以对称的想法为依据。然而，能够从等可能性或对称想法出发的问题，在实际上是很难遇见的。例如，要从等可能性出发推求射击中靶或生产出现废品的事件就不大可能。对于古典定义不能应用的场合，我们就有在大量试验基础上引入新的概率定义的必要。

为此，我们先引入事件频率的概念。

如果在确定条件下对某一事件A原则上可作无穷多次

重复试验，而在n次试验中这一事件出现了m次，则称$\frac{m}{n}$为事件A出现的**频率**。比如，如果在20次射击中击中目标13次，则击中目标的频率是$\frac{13}{20}$。

长期经验表明，许多随机事件的频率在大量重复试验之下，具有确定的稳定性。

例1 有人进行过多次投掷硬币试验，计算出现正面向上次数，得出结果如下：

试验者	掷钱次数	掷出正面向上次数	频率
蒲丰	4040	2048	0.5080
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

可见，出现正面向上事件的频率总在0.5这个数附近摆动。

例2 据瑞典官方统计资料，1935年全年各月出生女孩频率如下：

月份	1	2	3	4	5	6
总数	7280	6957	7883	7884	7892	7609
女孩	3537	3407	3866	3711	3775	3665
生女频率	0.486	0.489	0.490	0.471	0.478	0.482