概率与统计预报

及在地震与气象中的应用

南开大学数学系统计预报组

科学出版社

概率与统计预报
及在地震与气象中的应用

南开大学数学系统计预报组

科 学 出 版 社

1 9 7 8

内 容 简 介

本书前二章概要地叙述了概率论的基本知识及其在计算方法中的一些应用；后二章介绍了地震和气象的统计预报的若干方法，这些方法大都是在地震和气象工作的实践中总结出来的.

本书可供数学、计算数学及地震、气象等方面的同志参考。

概率与统计预报
及在地震与气象中的应用

南开大学数学系统计预报组

*

科学出版社出版
北京朝阳門内大街137号

上海商务印刷厂印刷

新华书店北京发行所发行　各地新华书店经售

*

1978年10月第一版　开本：787×1092 1/32
1978年10月第一次印刷　印张：6 7/8
印数：0001—17,130　字数：154,000

统一书号：13031·838
本社书号：1195·13—1

定价：0.85元

序　　言

本书的目的是简要地叙述概率论的基本概念及其在地震与气象统计预报等问题中的一些应用.

第一章简略地叙述概率论及数理统计中的基本概念与内容,介绍一些基本思想,并为以后的应用做些准备.

第二章讲概率论在计算方法中的应用,叙述了蒙特卡罗(Monte Carlo)方法的一些思路,可供计算数学专业或搞计算方法的同志参考.

第三章叙述我国在地震统计预报方面的部分研究成果.在党的领导和关怀下,我国地震预报及预防工作取得了很大成绩,在统计预报地震方面,也做了很多工作,限于我们的水平和能力,不可能作全面介绍.这里所叙述的一些内容,大多是我组在国家地震局的帮助下,在天津地震队、中国科学院地质研究所等单位的协作下,所得到的一些结果.

第四章概要地介绍天气预报中的几种统计方法.天气的统计预报内容非常丰富,这里只选述了几种既适用于天气预报、又对地震统计预报有参考意义的方法.本章涉及的应用实例是在天津市气象局的帮助下,由天津市气象台预报组和我组共同协作完成的.

我们的工作是在党组织的正确领导和广大群众的热情支持下进行的;此外,我们还不断地得到上述各单位及中国科学院地球物理研究所、国家海洋局地震预报组等的帮助,特此敬致谢意.

由于我们实践经验不足,理论水平有限,特别是第三、四

目　　录

第一章　概率论的基本概念

§1.1　事件的概率

(一)研究的对象　概率论与数理统计研究的对象是偶然事件(或称为随机事件)的数量规律性. 自然界许多事件是具有必然性的,例如,"在标准大气压下,水到100℃时会沸腾",这件事是必然会发生的,而"同性的电互相吸引"则必然不会发生. 但自然界也有许多事件是偶然性的,例如"投掷五分硬币出现正面"、"抽查10件产品时发现有一件次品"、"电话交换台在一小时内得到50次呼唤"、"7月间某河流的最高水位不超过6米"等等. 这两类事件有很大的差别,前一类必然事件在固定的条件下必然会发生(或必然不会发生),而后一类偶然事件即使在同样的条件下, 却可能发生, 也可能不发生. 当我们不断扔硬币时, 不管条件如何固定(同一硬币,同样的扔法,同时、同地等等), 总有时会得正面, 而另些时候却得反面, 谁也不能事先肯定下次会得那一面. 所以说这些事件是随机的.

在数学里, 必然现象的因果关系, 大都通过代数方程、微分方程、函数论等来研究, 而探讨随机事件的数量关系, 则主要是概率论与数理统计的任务.

人们通过对客观事物的观察或试验,获得大量的数据,然后运用数理统计来处理这些数据, 从中建立描述客观事物的数学模型,这时所用的方法主要是归纳法;概率论则主要从所得的数学模型出发, 运用演绎法, 进一步讨论客观事物的性质,揭示它们的规律性,从而达到改造世界的目的.

在一次试验里,虽然不能预言某随机事件是否必定出现,

但如果把这个试验重复做许多次，就可从中找出规律性来．例如，把硬币扔一万次，那么大概有5000次、也就是半数左右得正面；于是，人们利用这个多次的经验，来预言下面一次的结果，说：下次扔出正面的可能性是1/2．同样，根据过去长期的气象资料，我们可以预报明天的天气．

（二）**随机事件**　任一随机事件（简称事件），我们总可把它和某个随机试验联系起来．一个试验，如果它的结果有许多个，而且可以在相同的条件下不断地重复做下去，就称为随机试验，记为 E；它的任一可能出现的结果称为基本事件，记为 ω；它的全体基本事件构成一个集合 $\Omega=(\omega)$，称为该试验 E 的基本事件空间．

例1　E——观察所扔硬币的正反面；这里共有两个基本事件：ω_1——正面；ω_2——反面，$\Omega=(\omega_1,\ \omega_2)$．

例2　E——检查10件产品中的次品个数；这里共有11个基本事件：ω_i——发现 i 个次品，$i=0,\ 1,\ 2,\ \cdots,\ 10$；

$$\Omega=(\omega_0,\ \omega_1,\ \omega_2,\ \cdots,\ \omega_{10}).$$

例3　E——记录电话交换台在一小时内所得呼唤次数；ω_i——计有 i 次呼唤，$i=0,\ 1,\ 2,\ \cdots$，$\Omega=(\omega_0,\ \omega_1,\ \omega_2,\ \cdots)$．

例4　E——观察某河流在7月间的最高水位；ω_a——最高水位为 a 米，$0\leqslant a<\infty$；$\Omega=(\omega_a;\ 0\leqslant a<\infty)$．上面所说的"最高水位不超过6米"也是一事件，它是由许多基本事件 $\omega_a(0\leqslant a\leqslant 6)$ 所构成的，因而是一复合事件，我们把它记为 A，显然，$A=(\omega_a,\ 0\leqslant a\leqslant 6)$，它是 Ω 中的子集合，即 $A\subset\Omega$．

于是，我们从一个随机试验 E 出发，得到它的基本事件空间 $\Omega=(\omega)$；这个试验可能出现的事件（包括基本事件及复

合事件)，不是别的，无非是 Ω 的一些子集合.

任何随机试验里都有下列两个事件：一是必然事件，记为 Ω；另一是不可能事件，记为 ϕ. 例如，"扔硬币出现正面或反面"、"10 件产品中次品个数不超过 10"等都是必然事件；而"扔硬币既不出现正面又不出现反面""10 件产品中次品个数超过 10"等都是不可能事件.

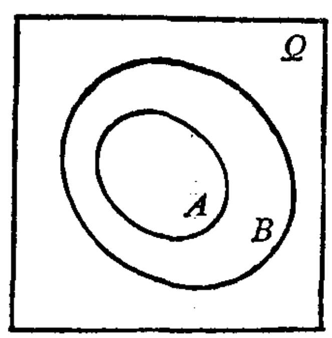
现在来讨论事件间的关系和运算. 当我们谈到许多事件时，如果没有特别声明，说的都是同一个随机试验的事件，也就是说，它们都是同一个基本事件空间的子集合. 以下设 A, B, $C\cdots$都是事件.

(1) 包含与相等：如果 A 发生必导致 B 发生，就说 B 包含 A，记作

$$A\subset B.$$

例如：(某河最高水位为 5 米)$\subset$(某河最高水位不超过 7 米). 如果 $A\subset B$, $B\subset A$，就说 A, B 相等，记为

$$A=B$$

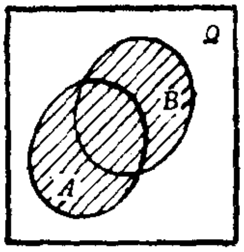


($A\subset B$)

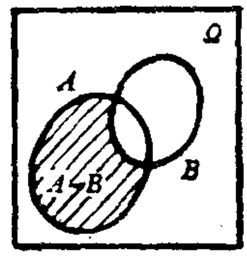
例如：(电话交换台所得到的呼唤次数不超过 5 次)=(所得呼唤次数或为 0 次，或为 1 次、或为 2 次、…、或为 5 次).

(2) 事件的和："A, B 中至少有一出现"也是一个事件，称为 A, B 的和，记为

$$A \cup B \quad (\text{或} A+B).$$



($A \cup B$)



($A-B$)

例如：(次品个数不超过 1)=(没有次品)∪(1 件次品)．
类似可以定义多个事件的和

$$A \cup B \cup C \cup \cdots.$$

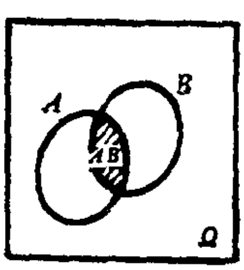
(3) 事件的差："A 出现而 B 不出现"也是一事件，称为 A 与 B 的差，记为

$$A-B.$$

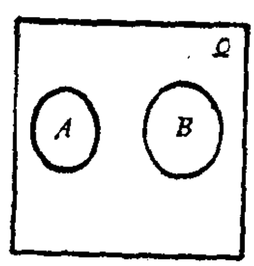
例如：(呼唤次数不超过 7)－(呼唤次数不超过 6)=(呼唤次数为 7)．

(4) 事件的交："A 与 B 同时出现"，也是一事件，称为 A, B 的交，记为

$$A \cap B \quad (\text{或} AB).$$



($A \cap B$)



(A 与 B 互斥)

例如：(呼唤次数为偶数) ∩ (呼唤次数不超过 3)=(呼唤次数为 2)．类似定义 $A \cap B \cap C \cap \cdots$.

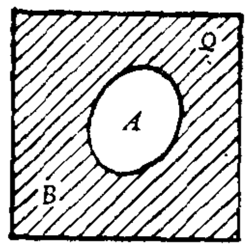
(5) 互斥：如果 A 与 B 不可能同时出现，也就是说，如果 A, B 的交是不可能事件，即 $AB=\phi$，就称 A 与 B 是互斥的(或互不相容的)．例如，(呼唤次数为 3)与(呼唤次数为偶数)是互斥的．

(6) 互逆：如果 A 与 B 互斥，但 A 与 B 必出现一个亦即

$$A\cap B=\phi,\ A\cup B=\Omega,$$

就说 A 与 B 互逆，或者说 A 是 B(或 B 是 A)的对立事件；记为

$$A=\overline{B}.$$



(A 与 B 互逆)

(三)**概率**　我们虽然不能预言某事件在一次试验中是否必定出现，但从经验中，常常可以发现一些事件出现的可能性比另一些大．例如，“天津冬季最低气温不低于 -30℃”的可能性比“低于 -30℃”的可能性大．这种可能性是客观存在的，自然应该用一个数字来表示，这样，我们就得到下列概念：表示事件 A 出现可能性大小的数字 $P(A)$ 称为 A 的概率．

这个定义是直观的、定性的，它只说明概率的作用，没有给出 $P(A)$ 的具体数值，那么怎样合理地求出 $P(A)$ 的值呢？这需要根据实验的条件作具体的分析．以下由浅入深，逐步讨论．

(A)　**等可能型**　设随机试验 E 只有有限多个(n 个)基本事件：

$$\Omega=(\omega_1,\ \omega_2,\ \cdots,\ \omega_n),$$

而且每个基本事件都处于平等的地位，任何 ω_i 都不比别的 ω_j 特别些，因而一切基本事件出现的可能性都是一样的，称这种随机试验为等可能型的．这时，自然合理地定义

$$P(\omega_i)=\frac{1}{n},\tag{1}$$

或者更一般地，定义

$$P(A)=\frac{k}{n},\tag{2}$$

这里 k 是事件 A 所含基本事件的个数($k\leqslant n$)．

例 1 (续)　扔硬币时出现正面(ω_1)与反面(ω_2)是等可能的，这时

$$P(\omega_i)=\frac{1}{2}\qquad(i=1,2).$$

例 6 (**随机取数**)　在 10 张同样的卡片上，分别写上 0, 1, 2, …, 9 十个数字，一张上写一个数，然后把卡片搅混，再任意取出一张，并以 x 表示上面的数字．由于卡片是同样的，抽取方法又是没有倾向性的，故可认为抽得每个数字的可能性是相等的，因而

$$P(x=i)=\frac{1}{10}\quad(i=0,1,\cdots,9);$$

$$P(x\text{ 是奇数})=\frac{5}{10}=\frac{1}{2},$$

因为这里只有 5 个奇数．

关于概率的性质有：

定理 1　(i)　$0\leqslant P(A)\leqslant 1$;

(ii)　$P(\Omega)=1$;

(iii)　设 $A_1, A_2, \cdots$是两两互斥的事件，则

$$P\left(\bigcup_{i=1} A_i\right)=\sum_{i=1} P(A_i).$$

证　由(2)直接得到(i)及(ii)，设 A_i 共含 a_i 个基本事件，

由互斥性 A_1，A_2，…不含公共的基本事件，因之 $\bigcup_{i=1} A_i$ 共含 $\sum_{i=1} a_i$ 个基本事件，由(2)得

$$P(\bigcup_{i=1} A_i)=\frac{\sum_{i=1} a_i}{n}=\sum_{i=1}\frac{a_i}{n}=\sum_{i=1} P(A_i)$$

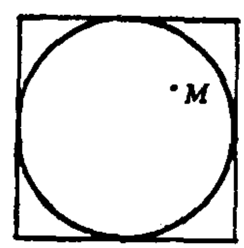
(B) **几何型** 设 Ω 是 k 维空间中的集，它具有有限的体积 $L(\Omega)>0$（一维时 L 是长度，二维时 L 是面积），今向 Ω 中投掷一质点 M，如果 M 在 Ω 中均匀分布，就说这个试验（掷点）是几何型的． 所谓"M 在 Ω 中均匀分布"的意义是："M 必须落在 Ω 中，而且落在某集 A 中的可能性大小与 A 的体积成正比，而与 A 的位置及形状无关"．

仍然用字母 A 表示事件："点落在 A 中"，则由均匀分布性，应定义

$$P(A)=\frac{L(A)}{L(\Omega)}. \tag{3}$$

根据(3)，容易证明定理 1 中的结论对几何型情况也正确．

例 7 在边长为 1 的正方形 Ω 中作一内切圆 A，并向 Ω 中掷点 M，设这个试验是几何型的，试求 M 落在 A 中的概率 $P(A)$．



解 $L(\Omega)=1$，A 的半径为 $1/2$，故 $L(A)=\pi\cdot(1/2)^2=\pi/4$．由(3)

$$P(A)=\frac{\pi}{4}\# \tag{4}$$

这个例子虽然非常简单，却启示我们一种利用实验来计算 π 的近似值的方法，如果能先求出 $P(A)$ 的值，那么根据(4)就有 $\pi=4P(A)$，而 $P(A)$ 却可通过多次掷点，用掷中的

频率来近似求出，即向 Ω 重复掷 n 次点，其中掷入 A 的次数设为 m，则当 n 很大时，频率 $\frac{m}{n}\approx P(A)$，因之

$$\pi\approx 4\cdot\frac{m}{n}.$$

"$\approx$"表示近似，这种通过随机试验来作近似计算的方法叫蒙特卡罗(Monte Carlo)方法或叫统计试验计算方法.

(C) **一般情形** 对一般的随机试验 E，我们将它重复作 n 次(可重复性是随机试验的定义中规定了的)，设事件 A 在其中出现了 m 次，称比值 m/n 为 A 出现的频率，而 m 为出现的频数. 显然，$m\leqslant n$，通常，如果 n 充分大，频率 m/n 稳定在某个数 $P(A)$ 的周围，即有

$$\frac{m}{n}\approx P(A),$$

这时我们称 $P(A)$ 为 A 的概率.

这样，我们通过频率来找出客观存在的概率. 事件 A 的频率依赖于试验次数 n，故把它记为 $f_n(A)$.

关于频率的性质有:

定理 2 (i) $0\leqslant f_n(A)\leqslant 1$;

(ii) $f_n(\Omega)=1$;

(iii) 设 $A_1, A_2, \cdots$ 是两两互斥的事件，则

$$f_n\Big(\bigcup_{i=1} A_i\Big)=\sum_{i=1} f_n(A_i). \tag{5}$$

证 根据频率的定义 $f_n(A)=\frac{m}{n}$，(i) (ii)是显然的，设 $A_1, A_2, \cdots$ 的频数分别为 $m_1, m_2, \cdots$，由于互斥的假定，这些事件不能两两同时出现，故事件 $\bigcup\limits_{i=1} A_i$ 的频数 k 满足

$$k=\sum_{i=1} m_i,$$

两边除以 n 得

$$\frac{k}{n}=\sum_{i=1}\frac{m_i}{n},$$

这就是(5)式#

定理2与定理1是类似的，因此容易想象，在一般情况，概率仍然具有定理1中的性质.

(D) **概率的抽象定义** 以上我们都是从随机试验这个具体事物出发，根据具体情况规定了概率 $P(A)$ 的数值；而且说明了，不论在那种情况，概率都具有定理1中的三个性质，这些性质是从实际中抽象出来的，是来自实际的. 然而，感性认识有待于提高到理性，概率的抽象定义，就是从这三个性质出发的.

设 $\Omega=(\omega)$ 是"点" ω 的集合，每个 ω 称为基本事件，因而称 Ω 为基本事件空间. Ω 中的一些子集 A 称为事件. 每一事件 A 对应于一个实数 $P(A)$，如果 $P(A)$ 满足下列三个条件，就称 $P(A)$ 为 A 的概率

(i) $0\leqslant P(A)\leqslant 1$;

(ii) $P(\Omega)=1$;

(iii) 设条件 $A_1, A_2, \cdots$ 两两互斥，则

$$P(\bigcup_{i=1} A_i)=\sum_{i=1} P(A_i).$$

历史上概率的概念经过了漫长的演变时间，也引起过不少争论，目前上述的几种定义是比较妥当的，但事物总是向前发展的，今后也许有更反映实际的理论来代替它.

由概率的定义可以推出：

(a) $P(\overline{A})=1-P(A)$.

证 因 A 与它的对立事件的和是必然事件，即 $A\cup\overline{A}=\Omega$，又因 $A, \overline{A}$ 互斥，故由(ii)及(iii)

$$1=P(\Omega)=P(A\cup\overline{A})=P(A)+P(\overline{A}).$$

(b) $P(\phi)=0$.

证 因不可能事件的对立事件是必然事件，故由(a)

$$P(\phi)=1-P(\Omega)=1-1=0.$$

(c) 如 $A\supset B$，则

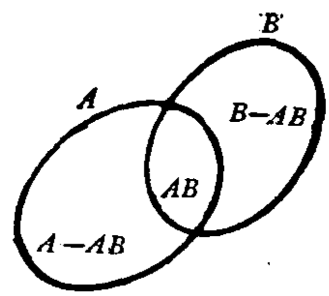
$$P(A-B)=P(A)-P(B).$$

证 由 $A=(A-B)\cup B$，及(iii)

$$P(A)=P(A-B)+P(B).$$

(d) 对任意二事件 A, B，有

$$P(A\cup B)=P(A)+P(B)-P(AB). \qquad (6)$$



证 由

$$A\cup B=(A-AB)\cup AB\cup(B-AB),$$

得 $$P(A\cup B)=P(A-AB)+P(AB)+P(B-AB),$$

再利用(c)，即得

$$P(A\cup B)=P(A)-P(AB)+P(AB)+P(B)-P(AB)$$
$$=P(A)+P(B)-P(AB).$$

例 8 地震预报需要预报三个要素，即发震的地点(经纬度)、时间及震级(它是反映地震大小的数字，可以自 0 级到 9 级)，今设某地在一年内发生 k 次五级以上地震的概率为

$$p_k=e^{-\lambda}\frac{\lambda^k}{k!} \qquad (k=0,\ 1,\ 2,\ \cdots), \qquad (7)$$

其中 $\lambda>0$ 是一个常数，试求该地一年内最多发生三次地震的概率 c_3.

解 以 A_k 表事件“恰好发生 k 次地震”，则 $A_0, A_1, A_2, \cdots$ 两两互斥，而且 $P(A_k)=p_k$. 由于事件

(最多发生三次地震)$=A_0\cup A_1\cup A_2\cup A_3$,
故由(iii)得

$$c_3=p_0+p_1+p_2+p_3=\sum_{k=0}^{3}e^{-\lambda}\frac{\lambda^k}{k!}.$$

类似可得:最少发生三次地震的概率为

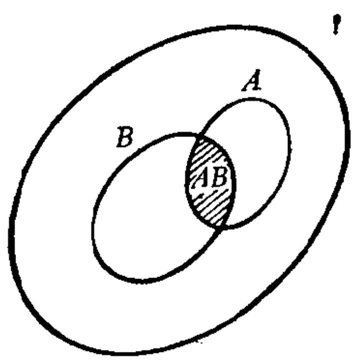
$$d_3=\sum_{k=3}^{\infty}e^{-\lambda}\frac{\lambda^k}{k!}.$$

(7) 中右方称为泊松分布,以后还要讲到,它的值有表可查.

§1.2 计算概率的一些方法

(一)条件概率 先考虑下列问题:设有质点(例如地震震中)在区域 Ω 中几何型地出现,则它出现在 A 中的(无条件)概率为

$$P(A)=\frac{L(A)}{L(\Omega)}, \tag{1}$$



"L"表示"体积". 现在假设通过某种手段,已经知道此质点只可能出现在 B 中,那么在这条件下,质点将出现在 A 中的条件概率 $P(A|B)$ 容易想象为

$$P(A|B)=\frac{L(AB)}{L(B)}. \tag{2}$$

由(1)(2)可见概率 $P(A)$ 一般与条件概率应是不同的. 为求两者间的关系,改写(2)为