

高等学校教材

# 概率论与 数理统计 教程

第二版

魏宗舒等  
编

概率论是研究随机现象数量规律的数学分支。概率论的肇始是在17世纪中叶，但它的起源之——解决与赌博有关的问题——可追溯到15世纪末。使概率论成为数学一个分支的奠基人是瑞士数学家雅各布·伯努利，他建立了概率论中的第一个极限定理，即伯努利大数定理……



高等教育出版社

021/226=2

2008

高等学校教材

# 概率论与数理统计教程

## 第二版

魏宗舒等 编

高等教育出版社

## 内容提要

本书包括事件与概率、离散型随机变量、连续型随机变量、大数定律与中心极限定理、数理统计的基本概念、点估计、假设检验、方差分析和回归分析、Excel 在统计分析中的应用等九章,可供高等师范院校与高等师范专科学校数学系作为教材使用。

### 图书在版编目(CIP)数据

概率论与数理统计教程/魏宗舒等编.—2版.—北京:高等教育出版社,2008.4  
ISBN 978-7-04-023571-5

I. 概… II. 魏… III. ①概率论-师范大学-教材②数理统计-师范大学-教材 IV. O21

中国版本图书馆 CIP 数据核字(2008)第 015456 号

策划编辑 李蕊 责任编辑 兰莹莹 封面设计 张申申  
责任绘图 吴文信 版式设计 王艳红 责任校对 杨凤珍  
责任印制 韩刚

---

出版发行	高等教育出版社	购书热线	010-58581118
社 址	北京市西城区德外大街 4 号	免费咨询	800-810-0598
邮政编码	100011	网 址	<a href="http://www.hep.edu.cn">http://www.hep.edu.cn</a>
总 机	010-58581000		<a href="http://www.hep.com.cn">http://www.hep.com.cn</a>
		网上订购	<a href="http://www.landaco.com">http://www.landaco.com</a>
经 销	蓝色畅想图书发行有限公司		<a href="http://www.landaco.com.cn">http://www.landaco.com.cn</a>
印 刷	北京民族印刷厂	畅想教育	<a href="http://www.widedu.com">http://www.widedu.com</a>
		版 次	1983 年 10 月第 1 版 2008 年 4 月第 2 版
开 本	850×1168 1/32	印 次	2008 年 4 月第 1 次印刷
印 张	15.375	定 价	21.10 元
字 数	390 000		

---

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 23571-00

## 第二版序

本书自1983年出版以来共印刷三十余次,受到广大读者的一致好评,同时也收到许多读者有益的意见和建议,积累至今,我们觉得有必要对本书进行一些修改。本次修订一方面是概率统计这门学科自身发展的需要,另一方面也是为了回报广大读者的厚爱。

本次修订的宗旨是:在保留第一版精华的同时,为适应新技术的发展修补错漏。在概率部分,改写了一部分实例与习题;引入了许多有时代气息的问题,如彩票问题、保险问题、投资组合问题和司法证明问题等等;补充了二维随机向量函数分布的变量变换定理;为了与后续课程相适应,将第一版中分布函数的左连续改为如今一贯通用的右连续。在统计部分,改正了一些编辑错误,使表述更准确。此外,将原书第九章改编为“Excel在统计分析中的应用”,以适应时代的需要。

在本书修改之时,原主编魏宗舒教授与部分原编写人员已经离我们而去,我们对他们表示深深的悼念,并向他们在本书中表现出的卓越的专业底蕴而致敬。

本次修改由汪荣明教授负责前四章,周纪芾教授负责后四章及第九章的编写。汪荣明教授通校了全书。此外,博士生胡少勇也参与了全书的修订,提出了许多衷恳的建议,并提供了不少例题与习题。

由于编者水平有限,书中纰漏在所难免,恳请广大读者批评指正。

汪荣明 周纪芾

2007年7月

## 第一版编者的话

这本概率论与数理统计教程是按照教育部 1980 年颁发的高等师范院校数学专业概率论与数理统计教学大纲编写的。把离散型随机变量和连续型随机变量分开讲授是这份大纲的一个重要特点,这无疑对掌握随机变量概念和今后在中学从事教学都是有好处的。我们在编写过程中尽力实现这一个想法。考虑到这一课程是高等师范院校数学专业教学计划唯一的一门处理随机现象的必修课程,我们致力于讲清其最基本的概念和方法,并用大量的例题来说明其应用的广泛性。

概率论是对随机现象统计规律演绎的研究,而数理统计是对随机现象统计规律归纳的研究。虽然两者在方法上是如此明显的不同,但作为一门学科,它们却是互相渗透、互相联系的。本书大体上也按照这样安排分为两部分。第一部分为概率论,包括一、二、三、四章,其中也掺杂一些数理统计的例子。第二部分为数理统计,包括第五、六、七、八、九章。在后一部分里用直观分析的地方较多,其中也有一小部分是概率论中的内容,在第九章(这是打星号(\*)的一章)中我们讲了数理统计在实际中的一些应用,希望读者能自己阅读至少其中一、二节,以便了解一下数理统计是怎样用来解决实际问题的。如果能亲临实际做一、二次数据分析,那对数理统计的领会就会更深了。为了帮助读者掌握书中的概念,除第九章外,每章末附有习题,其中有难有易。读者可酌量做一部分以开拓思路,加深理解。

按教学大纲的精神,本课程只要求有高等代数和数学分析的准备知识。凡是在这个水平以内的证明,我们尽可能较详尽地给

出。对于一些重要的定理,其证明不是最基本的,我们仍然给出,但打上星号(\*),读者可斟酌情况,决定阅读与否。还有一些比较冗长的证明或证明中需要用到的工具超出本书范围的,只好从略。§3.7是打星号(\*)的节,但由于其中所讲的特征函数是研究随机变量和的一个非常有力而又十分简洁的工具,并且在以后的章节中要几次用到它,因此建议读者了解一下特征函数的定义及其一些性质而跳过其证明,以便阅读以后的章节。

根据我们几次试用这本教材的情况,本书基本内容(除打星号(\*)的内容外)可以在大纲规定的76学时讲完。假如只有60学时,我们建议如下方案:

第一章:略去§1.5,估计12学时;

第二章:估计10学时;

第三章:在阅读本章时,请与第二章对照以掌握连续型变量的特性。略去一般性证明,估计10学时;

第四章:读§4.1及§4.3,略去这二节中定理的证明,估计4学时。

第五章:略去§5.3,估计4学时;

第六章:读§6.1及§6.2,估计4学时;

第七章:估计10学时;

第八章:略去§8.2中 $p$ 元的情形,估计6学时。

在本书编写过程中我们始终得到兄弟师范院校同行们的鼓励、支持和帮助,特别是评审本书的严士健教授,提了许多中肯而宝贵的意见,在此谨表感谢。本书的完成和本教研室全体同志的努力是分不开的,在此一并致谢。由于编者的水平,书中一定还存在许多错误,今后一定在工作中予以纠正。同时还盼读者能惠予批评指正。

编者

1982年3月

# 目 录

引言 .....	1
<b>第一章 事件与概率</b> .....	3
§ 1.1 随机事件和样本空间 .....	3
§ 1.2 概率和频率 .....	12
§ 1.3 古典概型 .....	15
§ 1.4 概率的公理化定义及概率的性质 .....	23
§ 1.5 条件概率、全概率公式和贝叶斯公式 .....	33
§ 1.6 独立性 .....	42
§ 1.7 伯努利概型 .....	49
习题 .....	52
<b>第二章 离散型随机变量</b> .....	60
§ 2.1 一维随机变量及分布列 .....	60
§ 2.2 多维随机变量、联合分布列和边际分布列 .....	71
§ 2.3 随机变量函数的分布列 .....	77
§ 2.4 数学期望的定义及性质 .....	80
§ 2.5 方差的定义及性质 .....	89
§ 2.6 条件分布与条件数学期望 .....	94
习题 .....	97
<b>第三章 连续型随机变量</b> .....	104
§ 3.1 随机变量及分布函数 .....	104
§ 3.2 连续型随机变量 .....	112
§ 3.3 多维随机变量及其分布 .....	119
§ 3.4 随机变量函数的分布 .....	130
§ 3.5 随机变量的数字特征、切比雪夫不等式 .....	146
§ 3.6 条件分布与条件期望、回归与第二类回归 .....	167

*§ 3.7	特征函数	176
习题		186
<b>第四章</b>	<b>大数定律与中心极限定理</b>	201
§ 4.1	大数定律	201
§ 4.2	随机变量序列的两种收敛性	207
§ 4.3	中心极限定理	215
*§ 4.4	中心极限定理(续)	221
习题		226
<b>第五章</b>	<b>数理统计的基本概念</b>	232
§ 5.1	母体与子样、经验分布函数	233
§ 5.2	统计量及其分布	237
§ 5.3	次序统计量及其分布	247
习题		256
<b>第六章</b>	<b>点估计</b>	262
§ 6.1	矩法估计	263
§ 6.2	极大似然估计	269
§ 6.3	克拉默 - 拉奥(Cramer - Rao)不等式	279
§ 6.4	充分统计量	287
§ 6.5	拉奥 - 勃拉克维尔(Rao - Blackwell)定理和一致 最小方差无偏估计	298
习题		304
<b>第七章</b>	<b>假设检验</b>	313
§ 7.1	假设检验的基本思想和概念	313
§ 7.2	参数假设检验	320
§ 7.3	正态母体参数的置信区间	334
§ 7.4	非参数假设检验	342
*§ 7.5	奈曼 - 皮尔逊基本引理和一致最优势检验	367
习题		375
<b>第八章</b>	<b>方差分析和回归分析</b>	382
§ 8.1	方差分析	382
§ 8.2	回归分析	404



习题 .....	427
<b>第九章 Excel 在统计分析中的应用 .....</b>	<b>434</b>
<b>附表 .....</b>	<b>448</b>
表 1 二项分布 $P(\xi \leq x) = \sum_{k=0}^x \binom{n}{k} p^k (1-p)^{n-k}$ 的数值表 .....	448
表 2 泊松分布 $P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ 的数值表 .....	458
表 3 正态分布函数 $N(0,1)$ 的数值表 .....	460
表 4 $\chi^2$ 检验的临界值表 .....	462
表 5 $F$ 检验的临界值表 .....	464
表 6 $t$ 检验的临界值表 .....	476
表 7 $D_n$ 的极限分布函数数值表 .....	477
表 8 科尔莫戈罗夫 (Колмогоров) 检验的临界值 ( $D_{n\alpha}$ ) 表 .....	478
表 9 两子样秩和检验的临界值表 .....	479
<b>参考书目 .....</b>	<b>480</b>

# 引 言

概率论与数理统计是研究随机现象统计规律的一门数学分支学科. 什么是随机现象呢? 让我们先做两个简单的试验.

试验 I: 一个盒子中有十个完全相同的白球, 搅匀后从中任意摸取一球.

试验 II: 一个盒子中有十个相同的球, 但五个是白色的, 另外五个是黑色的, 搅匀后从中任意摸取一球.

对于试验 I, 在球没有取出之前, 我们就能确定取出的必定是白球. 这种试验, 根据试验开始时的条件, 就可以确定试验的结果. 而对于试验 II 来说, 在球没有取出以前, 我们从试验开始时的条件, 不能确定试验的结果(即取出的球)是白的还是黑的. 也就是说一次试验的结果, 出现白球还是出现黑球, 在试验之前是无法确定的. 对于这一类试验, 骤然一看, 似乎没有什么规律可言. 但是, 实践告诉我们, 如果我们从盒子中反复多次取球(每次取出一球, 记录球的颜色后仍把球放回盒子中并且搅匀), 那么总可以观察到这样的事实: 当试验次数  $n$  相当大时, 出现白球的次数  $n_{\text{白}}$  和出现黑球的次数  $n_{\text{黑}}$  是很接近的, 比值  $\frac{n_{\text{白}}}{n}$  (或  $\frac{n_{\text{黑}}}{n}$ ) 会逐渐稳定于  $\frac{1}{2}$ . 出现这个事实是完全可以理解的, 因为在盒子中白球数等于黑球数, 从中任意摸取一个球, 取得白球或黑球的“机会”应该是平等的.

于是, 我们面对着两种类型的试验. 试验 I 所代表的类型, 在试验之前就能断定它有一个确定的结果, 这种类型的试验所对应的现象, 称为**确定性现象**. 确定性现象非常广泛, 例如:

“早晨, 太阳必然从东方升起.”

“苹果,不抓住必然往下掉。”

“边长为  $a, b$  的矩形,其面积必为  $a \cdot b$ 。”

“一个质点在  $t$  秒钟沿直线移动的距离为  $s(t)$ ,则此质点移动的速度必定是  $v(t) = \frac{ds(t)}{dt}$ 。”

.....

过去我们所学的各项数学课程基本上都是用来处理和研究这类确定性现象的。

试验 II 所代表的类型,它有多于一种可能的试验结果,但是在一次试验之前不能肯定试验会出现哪一个结果. 就一次试验而言,看不出有什么规律. 但是,“多次”地重复这个试验,试验结果又遵循某些规律,这种规律称之为“统计规律”,这一类试验称为随机试验,试验所代表的现象称为随机现象. 在客观世界中随机现象是极为普遍的,例如:

“某地区的年降雨量。”

“检查流水生产线上的一件产品,是合格品还是不合格品?”

“某电话交换站在单位时间内收到用户呼唤的次数。”

“打靶射击时,弹着点离靶心的距离。”

.....

概率论与数理统计就是研究随机现象的统计规律的数学学科,由于随机现象的普遍性,使得概率论与数理统计具有极其广泛的应用. 近年来,一方面它为科学技术、工农业生产等的现代化作出了重要的贡献;另一方面,广泛的应用也促进概率论与数理统计有了极大的发展.

# 第一章

## 事件与概率

### § 1.1 随机事件和样本空间

我们在引言中已经介绍了随机试验,现在进一步明确它的含义.一个试验如果满足下述条件:

- (1) 试验可以在相同的情形下重复进行;
- (2) 试验的所有可能结果是明确可知道的,并且不止一个;
- (3) 每次试验总是恰好出现这些可能结果中的一个,但在一次试验之前却不能肯定这次试验会出现哪一个结果.

就称这样的试验是一个**随机试验**,为方便起见,也简称为**试验**.

随机试验的每一个可能的结果,称为**基本事件**.因为随机试验的所有结果是明确的,从而所有的基本事件也是明确的,它们的全体,称作**样本空间**,通常用字母 $\Omega$ 表示. $\Omega$ 中的点,即基本事件,有时也称作**样本点**,常用 $\omega$ 表示.

**例 1.1** 在前述试验 II 中,令

$$\omega_1 = \{\text{取得白球}\}, \omega_2 = \{\text{取得黑球}\}$$

则

$$\Omega = \{\omega_1, \omega_2\}$$

**例 1.2** 一个盒子中有十个完全相同的球,分别标以号码 1, 2,  $\dots$ , 10, 从中任取一球,令

$$i = \{ \text{取得球的标号为 } i \}$$

则

$$\Omega = \{1, 2, \dots, 10\}$$

**例 1.3** 讨论某电话交换台在单位时间内收到的呼唤次数,

令

$$i = \{ \text{收到的呼唤次数为 } i \}$$

则

$$\Omega = \{0, 1, 2, \dots\}$$

**例 1.4** 测量某地水温, 令

$$t = \{ \text{测得的水温为 } t \text{ }^\circ\text{C} \}$$

则

$$\Omega = [0, 100]$$

在随机试验中, 有时关心的是带有某些特征的基本事件是否发生. 如在例 1.2 中, 我们可以研究

$$A = \{ \text{球的标号为 } 6 \}$$

$$B = \{ \text{球的标号是偶数} \}$$

$$C = \{ \text{球的标号小于等于 } 5 \}$$

这些结果是否发生? 其中  $A$  是一个基本事件, 而  $B$  与  $C$  则由多个基本事件所组成, 相对于基本事件, 就称它们是复杂事件. 无论是基本事件还是复杂事件, 它们在试验中发生与否, 都带有随机性, 所以都叫做随机事件, 或简称为事件. 习惯上人们常用大写字母  $A, B, C$  等表示事件. 在试验中, 如果出现  $A$  中所包含的某一个基本事件  $\omega$ , 则称作  $A$  发生, 并记作  $\omega \in A$ .

我们已经知道样本空间  $\Omega$  包含了全体基本事件, 而随机事件不过是有某些特征的基本事件所组成, 所以从集合论的观点来看, 一个随机事件不过是样本空间  $\Omega$  的一个子集而已. 如在例 1.2 中,  $\Omega = \{1, 2, \dots, 10\}$ , 显然, 前述的随机事件  $A, B, C$  都是  $\Omega$  的子集, 它们可以简单地表示为  $A = \{6\}$ ,  $B = \{2, 4, 6, 8, 10\}$ ,  $C = \{1, 2, 3, 4, 5\}$ .

又因为  $\Omega$  是所有基本事件所组成,因而在任一次试验中,必然要出现  $\Omega$  中的某一基本事件  $\omega$ ,即  $\omega \in \Omega$ .也就是在试验中, $\Omega$  必然会发生,所以今后又用  $\Omega$  来代表一个必然事件.相应地,空集  $\emptyset$  可以看成是  $\Omega$  的子集,在任意一次试验中,不可能有  $\omega \in \emptyset$ ,也就是说  $\emptyset$  永远不可能发生,所以  $\emptyset$  是不可能事件.必然事件和不可能事件的发生与否,已经失去了“不确定性”,因而本质上它们不是随机事件.但是为了方便起见,我们还是把它们看作随机事件,稍后我们会理解,它们不过是随机事件的两个极端情形而已.

一个样本空间  $\Omega$  中,可以有很多的随机事件.概率论的任务之一,是研究随机事件的规律,通过对较简单事件规律的研究去掌握更复杂事件的规律.为此,需要研究事件之间的关系和事件之间的一些运算.

如果没有特别的声明,在以下的叙述中总认为样本空间  $\Omega$  已经给定,并且还给出了  $\Omega$  中的一些事件,如  $A, B, A_i (i=1, 2, \dots)$  等等.

1. 如果事件  $A$  发生必然导致事件  $B$  发生,则称事件  $B$  包含事件  $A$ ,或称  $A$  是  $B$  的特款,并记作  $A \subset B$  或  $B \supset A$ . 比如在前面提到过的  $A = \{\text{球的标号为 } 6\}$  这一事件发生就导致事件  $B = \{\text{球的标号是偶数}\}$  的发生.因为摸到标号为 6 的球意味着标号为偶数的球出现了,所以前者是后者的特款,也就是后者包含了前者.

可以给上述的含义以一个直观的几何解释.设样本空间  $\Omega$  是一个正方形,  $A$  与  $B$  是两个事件,也就是  $\Omega$  的某两个子集. “ $A$  发生必然导致  $B$  发生”意味着“属于  $A$  的  $\omega$  必然属于  $B$ ”,即  $A$  中的点全在  $B$  中,其几何图形如图 1.1 所示:

由此可知,事件  $A \subset B$  的含义与集合论中的意义是一致的.

因为不可能事件  $\emptyset$  不含有任何  $\omega$ ,所

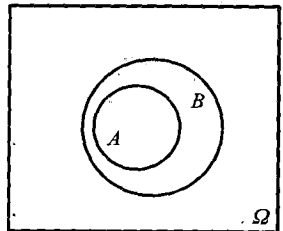


图 1.1

以对任一事件  $A$ , 我们约定

$$\emptyset \subset A$$

2. 如果有  $A \subset B, B \subset A$  同时成立, 则称事件  $A$  与  $B$  相等, 记作  $A = B$ . 易知, 相等的两个事件  $A, B$  总是同时发生或同时不发生.

如在例 1.2 中, 若

$$A = \{\text{球的标号为偶数}\}$$

$$B = \{\text{球的标号为 } 2, 4, 6, 8, 10\}$$

则显然有  $A = B$ . 直截了当地说, 所谓  $A = B$ , 就是  $A, B$  中含有相同的样本点. 因而, 上述关于  $A = B$  的说法, 看起来有点“绕圈子”, 但是稍后就会看到, 这种“绕圈子”的说法, 在验证两个事件是否相等时是非常有用的, 在许多情形中可以说是唯一的一种方法.

3. “事件  $A$  与  $B$  中至少有一个发生”, 这样的—个事件称作事件  $A$  与  $B$  的并(或和), 并记作  $A \cup B$ . 它的几何表示如图 1.2, 图中的阴影部分是事件“ $A \cup B$ ”. 如在例 1.2 中, 若

$$A = \{\text{球的标号为偶数}\}$$

$$B = \{\text{球的标号小于等于 } 3\}$$

则

$$A \cup B = \{\text{球的标号为 } 1, 2, 3, 4, 6, 8, 10\}$$

4. “事件  $A$  与  $B$  同时发生”, 这样的事件称作事件  $A$  与  $B$  的交(或积), 记作  $A \cap B$  (或  $AB$ ), 它对应图 1.3 中的阴影部分:

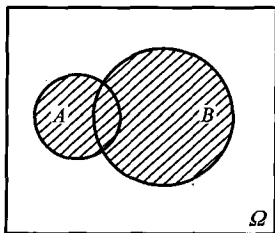


图 1.2

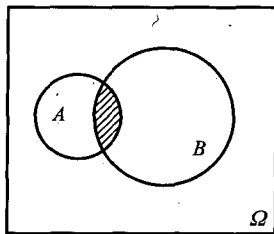


图 1.3

如在例 1.2 中,若  $A, B$  同上,则

$$A \cap B = \{\text{球的标号为 } 2\}$$

5. “事件  $A$  发生而  $B$  不发生”,这样的事件称为事件  $A$  与  $B$  的差,记作  $A - B$ ,它表示了图 1.4 中的阴影部分:

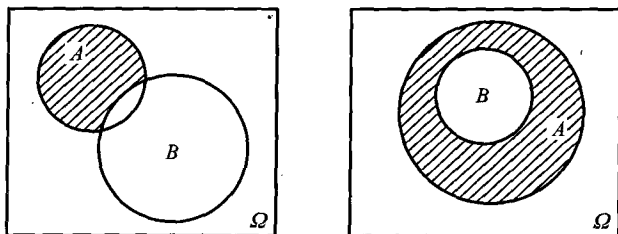


图 1.4

如在例 1.2 中,若  $A, B$  同上,则

$$A - B = \{\text{球的标号为 } 4, 6, 8, 10\}$$

我们还可以发现,  $A - B = A - AB = A \bar{B}$  ( $\bar{B}$  的定义见后面的第 7 条).

6. 若事件  $A$  与  $B$  不能同时发生,也就是说  $AB$  是一个不可能事件,即  $AB = \emptyset$ ,则称事件  $A$  与  $B$  互不相容. 下面的图 1.5 表示了这一情形.

如在例 1.2 中,若

$$A = \{\text{球的标号为偶数}\}$$

$$B = \{\text{球的标号为 } 3\}$$

则显然  $A$  与  $B$  不可能同时发生,即有  $AB = \emptyset$ ,也就是说  $A$  与  $B$  是互不相容的.

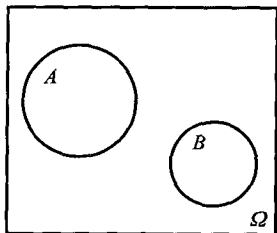


图 1.5

7. 若  $A$  是一个事件,令  $\bar{A} = \Omega - A$ ,称

$\bar{A}$  是  $A$  的对立事件或逆事件. 容易知道,在一次试验中,若  $A$  发生,则  $\bar{A}$  必不发生(反之亦然),即  $A$  与  $\bar{A}$  二者只能发生其中之一,并且也必然发生其中之一. 因而有



$$A\bar{A} = \emptyset$$

$$A \cup \bar{A} = \Omega$$

此外显然有

$$\overline{\bar{A}} = A$$

如在例 1.2 中,若

$$A = \{\text{球的标号为偶数}\}$$

则

$$\bar{A} = \{\text{球的标号为奇数}\}$$

8. 若有  $n$  个事件:  $A_1, A_2, \dots, A_n$ , 则“ $A_1, A_2, \dots, A_n$  中至少发生其中的一个”这样的事件称作  $A_1, A_2, \dots, A_n$  的并, 并记作  $A_1 \cup A_2 \cup$

$\dots \cup A_n$  或  $\bigcup_{i=1}^n A_i$ ; 若“ $A_1, A_2, \dots, A_n$  同时发生”, 这样的事件称作

$A_1, A_2, \dots, A_n$  的交, 记作  $A_1 A_2 \dots A_n$  或  $\bigcap_{i=1}^n A_i$ .

如果读者已经有了一定的集合论知识, 一定会发现事件间的关系及运算与布尔(Boole)代数中集合间的关系及运算是完全可以互相类比的. 下面给出这种类比的对应关系:

概率论	集合论
样本空间	$\Omega = \{\omega\}$
事件	子集
事件 $A$ 发生	$\omega \in A$
事件 $A$ 不发生	$\omega \notin A$
必然事件	$\Omega$
不可能事件	$\emptyset$
事件 $A$ 发生导致事件 $B$ 发生	$A \subset B$
事件 $A$ 与 $B$ 至少有一个发生	$A \cup B$
事件 $A$ 与 $B$ 同时发生	$A \cap B$ (或 $AB$ )
事件 $A$ 发生而 $B$ 不发生	$A - B$