高等学校创新教材

概率论与数理统计

苏保河 ◎ 编著

厦门大学出版社 XIAMEN UNIVERSITY PRESS
国家一级出版社
全国百佳图书出版单位

高等学校创新教材

概率论与数理统计

苏保河◎编著

厦门大学出版社
XIAMEN UNIVERSITY PRESS
国家一级出版社
全国百佳图书出版单位

图书在版编目(CIP)数据

概率论与数理统计/苏保河编著.—厦门:厦门大学出版社,2015.1
高等学校通识课程教材系列
ISBN 978-7-5615-5351-0

Ⅰ.①概… Ⅱ.①苏… Ⅲ.①概率论-高等学校-教材②数理统计-高等学校-教材
Ⅳ.①O21

中国版本图书馆 CIP 数据核字(2014)第 309020 号

官方合作网络销售商:

厦门大学出版社出版发行
(地址:厦门市软件园二期望海路 39 号 邮编:361008)
总 编 办 电 话:0592-2182177 传真:0592-2181253
营销中心电话:0592-2184458 传真:0592-2181365
网址:http://www.xmupress.com
邮箱:xmup@xmupress.com
沙县四通彩印有限公司印刷
2015 年 1 月第 1 版 2015 年 1 月第 1 次印刷
开本:720×970 1/16 印张:19 插页:2
字数:332 千字 印数:1～3 000 册
定价:37.00 元
本书如有印装质量问题请直接寄承印厂调换

高等学校通识课程教材系列

总 主 编　曹利军

执行主编　钟瑞栋

国务院侨务办公室立项

彭磷基外招生人才培养改革基金资助

内容提要

本书共分为八章，包括随机事件及其概率、一维随机变量及其数字特征、多维随机变量及其数字特征、大数定律和中心极限定理、数理统计的基本概念、参数估计、假设检验、方差分析和线性回归分析等内容．各章配有习题，并附有习题参考答案．本书的最大特点是将功能强大的数学软件 Mathematica 和 SPSS 融入教学之中，力图降低学生的学习负担，帮助他们更好地掌握概率论与数理统计知识．参加本书审稿的有：吴广庆、杜萍、刘中学、王为民、洪莉、张越等．

本书可作为高等院校“概率论与数理统计”课程的教材或教学参考书，适用于理工医和经济管理类各专业本科教学．本书也适用于社科类大学生作为通识课或选修课开设的“概率论与数理统计”课程，教师可根据学生基础和教学学时适当减少理论证明和公式推导．本书可供各类管理人员和工程技术人员参考．

讲授本教材内容，可根据学时多少、生源特征和教学目标作适当取舍，讲授学时可在 32～64 学时．与传统教学内容相比，建议增加应用的讲授和训练，减少计算技巧的讲解和练习，让学生把数学的思想和方法留给自己，把枯燥繁杂的计算交给计算机．

使用本教材，预计可以减轻学生负担 25%，有效激发学生的学习兴趣，彻底解决相关计算问题，有助于学生终身掌握概率论与数理统计知识，提高将理论方法应用于实际的能力．

前 言

“概率论与数理统计”是高等院校理工医和经管类各专业的重要基础课程，由于其逻辑性强，计算量大，应用性广，使我们在教学实践中遇到极大挑战：

首先，教学课时不断压缩与教学内容有增无减的矛盾难以调和. 2000 年以前，“概率论与数理统计”课程的学时普遍为 64 学时以上，近年来，受教学改革、节假日增多等影响，学时普遍被压缩为 48 学时以下，而且很多院校将每学时由 50 分钟压缩为 45 分钟. 与此同时，人们越来越认识到概率论与数理统计的重要作用，教学内容并没有减少.

其次，传统教学内容和方法难以适应高等教育从精英教育向大众教育的转变. 近年来，随着高校的大面积扩招，大学生不再都是百里挑一的高才生，对概率论与数理统计的理解能力整体上有所下降，不少学生学习比较吃力，学习效果不甚理想.

另外，学生培养目标与概率论与数理统计教学内容的矛盾日益突出. 近年来，高等院校越来越注重培养受社会欢迎的“应用型人才”，而传统教学内容偏重理论推导和习题演算，与实际应用严重脱节，难以培养学生利用概率论与数理统计知识解决实际问题的能力.

我们用什么来应对挑战？

荀子在《劝学》中说：“假舆马者，非利足也，而致千里；假舟楫者，非能水也，而绝江河. 君子生非异也，善假于物也.”今天，功能最强大、使用最普及的“物”或许就是计算机了. 我们认为，将功能强大的计算机和数学软件融入概率论与数理统计教学之中，能够有效解决学生数学基础薄弱和保证教学质量的矛盾，彻底破解棘手的计算

问题，大大激发学生的学习兴趣，显著提高概率论与数理统计的教学质量，为高等院校的教学改革和发展提供正能量.

将计算机和数学软件（本书采用 Mathematica 和 SPSS）融入概率论与数理统计教学中，其优越性至少有以下几方面：

第一，Mathematica 和 SPSS 能涵盖概率论与数理统计的所有内容，不存在教学内容上的盲区，不用担心内容的衔接与断档问题.

第二，Mathematica 和 SPSS 不仅功能极其强大，而且具有良好的人机界面，操作方便，简单易学，学生很容易掌握.

第三，对于基础较差的学生，即使他们没有很好地掌握概率论与数理统计的理论和技巧，也可以利用 Mathematica 和 SPSS 实现相关运算，使他们从抽象乏味的计算中解脱出来，有助于激发他们的学习兴趣和潜能，培养他们运用概率论与数理统计方法分析和解决实际问题的能力.

基于上述认识，我们将教学内容与数学软件有机融合起来，在系统讲述概率论与数理统计知识的基础上，增加了利用 Mathematica 和 SPSS 实现相关运算的内容. 此外，我们还增加了若干实际问题，通过建立数学模型和利用数学软件运算，培养学生分析和解决实际问题的能力.

实践是检验真理的唯一标准，本教材的创新之处是否有生命力，能否有效改善概率论与数理统计的教学效果，还有待于实践的检验. 欢迎各位同行集思广益，推动本教材的质量再上新台阶.

限于编者水平，教材中一定存在错误和不妥之处，敬请广大读者批评指正，谨表示诚挚的谢意！

苏保河

2014 年 10 月

目 录

第一章 随机事件及其概率 …… 1
第一节 随机试验与样本空间 …… 1
第二节 概率的定义 …… 6
第三节 计算概率的常用方法 …… 14
第四节 Mathematica 的简单应用 …… 25
习题一 …… 29

第二章 随机变量 …… 32
第一节 随机变量及其分布函数 …… 32
第二节 离散型随机变量 …… 34
第三节 连续型随机变量 …… 42
第四节 随机变量的函数的分布 …… 51
第五节 随机变量的数字特征 …… 57
第六节 Mathematica 在随机变量中的应用 …… 71
习题二 …… 77

第三章 多维随机变量 …… 81
第一节 多维随机变量及其联合分布函数 …… 81
第二节 二维离散型随机变量 …… 83
第三节 二维连续型随机变量 …… 91
第四节 多维随机变量的数字特征 …… 109
第五节 二维正态分布 …… 119
第六节 Mathematica 在多维随机变量中的应用 …… 122
习题三 …… 126

第四章 大数定律和中心极限定理 …… 132
第一节 大数定律 …… 132

第二节　中心极限定理 …… 137
第三节　Mathematica 的应用 …… 141
习题四 …… 143

第五章　数理统计的基本概念 …… 144
第一节　总体与样本 …… 144
第二节　常用抽样分布 …… 152
第三节　数学软件在数理统计中的简单应用 …… 160
习题五 …… 165

第六章　参数估计 …… 167
第一节　参数的点估计 …… 167
第二节　参数的区间估计 …… 177
第三节　数学软件在参数估计中的应用 …… 189
习题六 …… 195

第七章　假设检验 …… 198
第一节　假设检验的概念 …… 198
第二节　正态总体参数的假设检验 …… 203
第三节　两个正态总体参数的假设检验 …… 213
第四节　分布拟合检验 …… 223
第五节　应用实例 …… 227
第六节　数学软件在假设检验中的应用 …… 235
习题七 …… 247

第八章　方差分析和线性回归分析 …… 252
第一节　单因素方差分析 …… 252
第二节　一元线性回归分析 …… 258
第三节　SPSS 的应用 …… 265
习题八 …… 270

附表 1　泊松分布表 …… 272
附表 2　标准正态分布函数表 …… 274

附表 3 χ^2 分布的上 α 分位数表 …… 275
附表 4 t 分布的上 α 分位数表 …… 277
附表 5 F 分布的上 α 分位数表 …… 278

习题参考答案 …… 284
参考文献 …… 295

第一章 随机事件及其概率

在自然界和人类社会活动中，经常会看到两类现象：第一类是必然现象，例如，两个物体相互之间必然会有吸引力，同性电荷必然相互排斥，异性电荷必然相互吸引，等等；第二类是随机现象，例如，抛掷一枚质地均匀的硬币，可能出现正面，也可能出现反面，在抛掷之前不知道出现哪一结果，但是，如果我们重复抛掷同一硬币很多次，其结果就会呈现出规律性：出现正面的次数约为抛掷总次数的一半．概率论与数理统计就是研究随机现象规律的一门学科．

第一节 随机试验与样本空间

一、随机试验

我们把人们对随机现象的观察称为**试验**．例如，抛掷一枚硬币观察出现正面还是反面，观察一批电子产品的寿命等．概率论与数理统计研究的是**随机试验**．

定义 1 满足下述条件的试验称为随机试验：

(1) 试验可以在相同条件下重复进行；

(2) 每次试验的结果不止一个，在试验之前能够明确试验的所有结果；

(3) 每次试验之前无法预知会出现哪一结果．

随机试验可以简称为**试验**，经常用 E 表示．

例 1 E_1：掷一枚骰子，观察出现的点数．

例 2 E_2：记录某电话总机在某段时间接到的呼叫次数．

例 3 E_3：在一批电脑中任取一台，测试它的寿命(以小时计)．

二、随机事件

试验的结果不止一个，每次试验可能出现这一结果，也可能出现那一结果．

定义 2 随机试验的直接结果称为样本点，所有样本点的集合称为样本空间，样本空间的子集称为随机事件.

样本空间常用 S 表示，**随机事件**可以简称为**事件**，经常用大写字母 A,B,C 等表示.

由一个**样本点**组成的事件称为**基本事件**.

在一次试验中，当且仅当事件中的一个样本点出现时，称这一事件**发生**. 显然，样本空间 S 也是事件，它在每一次试验中必然发生，称为**必然事件**. 空集不包含样本点，它在每一次试验中都不可能发生，称为**不可能事件**，经常将其记为 $\varnothing$.

例 4 例 1 中 E_1 的样本空间为 $S_1=\{1,2,3,4,5,6\}$，“出现的点数大于 3”是一个事件，这一事件可记为

$$A_1=\{4,5,6\},$$

“出现的点数大于 0”是必然事件，“出现的点数大于 6”是不可能事件.

例 5 例 2 中 E_2 的样本空间为 $S_2=\{0,1,2,\cdots\}$，“接到的呼叫次数不超过 5”是一个事件，这一事件可记为

$$A_2=\{0,1,2,3,4,5\}.$$

例 6 例 3 中 E_3 的样本空间为 $S_3=\{t\,|\,t\geqslant 0\}$，“这台电脑的寿命大于 1000(小时)”是一个事件，这一事件可记为

$$\{t\,|\,t>1000\},$$

“这台电脑的寿命不小于 0”是必然事件，“这台电脑的寿命小于 0”是不可能事件.

三、随机事件的关系与运算

根据上述定义，事件也是集合，因此，事件的关系与运算的实质就是集合的关系与运算. 如果我们用平面内的矩形表示样本空间 S，用矩形内的平面图形表示事件 A，则可以将事件的关系和运算通过图形表示出来，这种直观的表示方法称为**文图**，如图 1-1 所示.

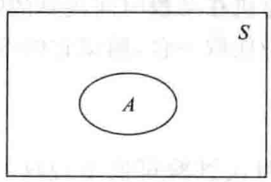

图 1-1

1. 包含和相等

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，记为 $B \supset A$ 或 $A \subset B$，如图 1-2 所示. 如果 $A \supset B$ 且 $B \supset A$，则称事件 A 与事件 B **相等**，记为 $A = B$ 或 $B = A$.

2. 互不相容(互斥)

如果事件 A 与事件 B 不能同时发生，则称事件 A 与事件 B **互不相容**或**互斥**，如图 1-3 所示. 如果 n 个事件 $A_1, A_2, \cdots, A_n$ 中任意两个事件都是互不相容的，则称这 n 个事件 $A_1, A_2, \cdots, A_n$ 是**两两互不相容**的.

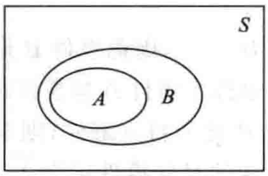

图 1-2

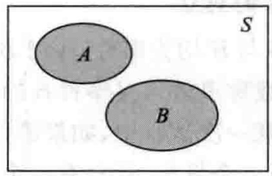

图 1-3

3. 事件的和

如果 A 与 B 均为事件，则事件 $A \cup B$ 表示"事件 A 与 B 中至少有一个发生"，称为事件 A 与事件 B 的**和事件**，如图 1-4 所示. $A \cup B$ 也可以记为 $A + B$.

类似地，事件 $\bigcup\limits_{i=1}^{n} A_i$ 称为 n 个事件 $A_1, A_2, \cdots, A_n$ 的和事件，表示 n 个事件 $A_1, A_2, \cdots, A_n$ 中至少有一个发生；事件 $\bigcup\limits_{i=1}^{\infty} A_i$ 称为可列个事件 $A_1, A_2, \cdots$ 的和事件，表示可列个事件 $A_1, A_2, \cdots$ 中至少有一个发生.

4. 事件的积

如果 A 与 B 均为事件，则事件 $A \cap B$ 表示"事件 A 与 B 都发生"，称为事件 A 与事件 B 的**积事件**，如图 1-5 所示. $A \cap B$ 也可以记为 AB.

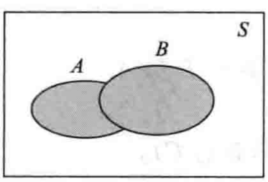

图 1-4

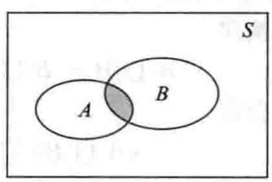

图 1-5

显然，事件 A 与事件 B 互不相容的充分必要条件为 $A \cap B = \varnothing$.

类似地，事件 $\bigcap_{i=1}^{n} A_i$ 称为 n 个事件 $A_1, A_2, \cdots, A_n$ 的积事件，表示 n 个事件 $A_1, A_2, \cdots, A_n$ 都发生；事件 $\bigcap_{i=1}^{\infty} A_i$ 称为可列个事件 $A_1, A_2, \cdots$ 的积事件，表示可列个事件 $A_1, A_2, \cdots$ 都发生.

5. 事件的差

如果 A 与 B 均为事件，则事件 $A-B$ 表示"事件 A 发生且事件 B 不发生"，称为事件 A 与事件 B 的**差事件**，如图 1-6 所示. 显然有

$$A-A=\varnothing,\quad A-S=\varnothing,\quad A-\varnothing=A.$$

6. 事件的对立

如果 A 与 B 均为事件，$A\cup B=S$，$A\cap B=\varnothing$，则称事件 B 是事件 A 的**对立事件**，或称事件 A 是事件 B 的对立事件，也称为事件 A 与事件 B 互为对立事件. 在任意一次试验中，如果事件 A 与事件 B 互为对立事件，则事件 A 与事件 B 中必有一个发生，且只有一个发生. 事件 A 的对立事件记为 $\overline{A}$，表示"事件 A 不发生"，如图 1-7 所示. A 的对立事件 $\overline{A}$ 也称为 A 的**逆事件**或**余事件**. 显然，

$$\overline{A}=S-A,\quad \overline{\overline{A}}=A$$
$$A\cup\overline{A}=S,\quad A\cap\overline{A}=\varnothing,$$
$$A-B=A\overline{B}=A-AB.$$

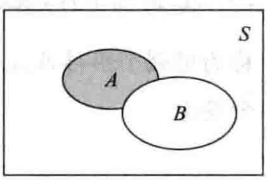

图 1-6

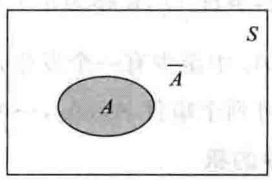

图 1-7

事件的运算有如下规律：

(1) 交换律

$$A\cup B=B\cup A,\quad A\cap B=B\cap A.$$

(2) 结合律

$$(A\cup B)\cup C=A\cup(B\cup C),$$
$$(A\cap B)\cap C=A\cap(B\cap C).$$

(3) 分配律

$$A\cup(B\cap C)=(A\cup B)\cap(A\cup C),$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$

$$A \cup (\bigcap_{i=1}^{n} B_i) = \bigcap_{i=1}^{n} (A \cup B_i), \quad A(\bigcup_{i=1}^{n} B_i) = \bigcup_{i=1}^{n} AB_i.$$

(4) 德摩根律

$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \quad \overline{A \cap B} = \overline{A} \cup \overline{B},$$

$$\overline{\bigcup_{i=1}^{n} A_i} = \bigcap_{i=1}^{n} \overline{A_i}, \quad \overline{\bigcap_{i=1}^{n} A_i} = \bigcup_{i=1}^{n} \overline{A_i}.$$

例 7 掷一枚骰子,观察出现的点数. A 表示"出现偶数点",B 表示"出现的点数大于 2",C 表示"出现的点数是小于 4 的奇数". 试表示下列事件:

$$A, \quad B, \quad C, \quad A \cup B, \quad AB, \quad A-B, \quad B\overline{C}, \quad \overline{A} \cup C.$$

解 $A = \{2,4,6\}, \quad B = \{3,4,5,6\}, \quad C = \{1,3\},$

$A \cup B = \{2,3,4,5,6\}, \quad AB = \{4,6\}, \quad A - B = A - AB = \{2\},$

$B\overline{C} = B - C = \{4,5,6\}, \quad \overline{A} \cup C = \{1,3,5\} \cup \{1,3\} = \{1,3,5\}.$

例 8 证明:(1)$(A \cup B)(A \cup \overline{B}) = A$; (2)$(A \cup B) - A = B - A$.

证 (1)$(A \cup B)(A \cup \overline{B}) = A \cup (B\overline{B}) = A \cup \varnothing = A$.

$$(2)(A \cup B) - A = (A \cup B)\overline{A} = A\overline{A} \cup B\overline{A} = \varnothing \cup B\overline{A} = B\overline{A} = B - A.$$

例 9 设 $A_i (i = 1,2,3)$ 是事件,试用 $A_i (i = 1,2,3)$ 表示下列事件:

(1)A_1 不发生,A_2,A_3 都发生.

(2)A_1,A_2,A_3 中恰有一个发生.

(3)A_1,A_2,A_3 中不多于两个发生.

(4)A_1,A_2,A_3 中至少有两个发生.

解 (1) 用 B_1 表示事件"A_1 不发生,A_2,A_3 都发生". 注意到 A_1 不发生,即$\overline{A_1}$ 发生,因此,$B_1 = \overline{A_1}A_2A_3$.

(2) 用 B_2 表示事件"A_1,A_2,A_3 中恰有一个发生". 注意到 B_2 包含:A_1 发生且 A_2,A_3 均不发生、A_2 发生且 A_1,A_3 均不发生、A_3 发生且 A_1,A_2 均不发生,因此,

$$B_2 = A_1\overline{A_2}\,\overline{A_3} \cup \overline{A_1}A_2\overline{A_3} \cup \overline{A_1}\,\overline{A_2}A_3.$$

(3) 用 B_3 表示事件"A_1,A_2,A_3 中不多于两个发生",其含义为"A_1,A_2,A_3 中至少有一个未发生",因此,

$$B_3 = \overline{A_1} \cup \overline{A_2} \cup \overline{A_3} = \overline{A_1A_2A_3}.$$

(4) 用 B_4 表示事件"A_1,A_2,A_3 中至少有两个发生",其含义为"A_1,A_2,A_3 中有且只有两个发生,或者三个都发生",因此

$$B_4 = A_1A_2\overline{A_3} \cup A_1\overline{A_2}A_3 \cup \overline{A_1}A_2A_3 \cup A_1A_2A_3.$$

例 10 抛 3 次硬币，事件 $A_i(i=1,2,3)$ 表示“第 i 次出现正面”，试用 $A_i(i=1,2,3)$ 表示下列事件：

(1) 仅第一次出现正面.

(2) 只出现一次正面.

(3) 三次都没有出现正面.

(4) 至少出现一次正面.

解 (1) 用 B_1 表示事件“仅第一次出现正面”，这意味着“第一次出现正面，且第二次和第三次均未出现正面”，因此，$B_1 = A_1\overline{A_2}\,\overline{A_3}$.

(2) 用 B_2 表示事件“只出现一次正面”，这表示“第一次出现正面且第二次和第三次均未出现正面，或第二次出现正面且第一次和第三次均未出现正面，或第三次出现正面且第一次和第二次均未出现正面”，因此，

$$B_2 = A_1\overline{A_2}\,\overline{A_3} \cup \overline{A_1}A_2\overline{A_3} \cup \overline{A_1}\,\overline{A_2}A_3.$$

(3) 用 B_3 表示事件“三次都没有出现正面”，这意味着“第一次没有出现正面，第二次没有出现正面，第三次也没有出现正面”，因此，$B_3 = \overline{A_1}\,\overline{A_2}\,\overline{A_3}$.

(4) 用 B_4 表示事件“至少出现一次正面”，这表示“或第一次出现正面，或第二次出现正面，或第三次出现正面”，因此，$B_4 = A_1 \cup A_2 \cup A_3$. 根据对立事件的概念和运算规律，

$$B_4 = \overline{\overline{A_1 \cup A_2 \cup A_3}} = \overline{\overline{A_1}\,\overline{A_2}\,\overline{A_3}}.$$

B_4 也可以表示为互不相容的事件的和事件

$$B_4 = A_1\overline{A_2}\,\overline{A_3} \cup \overline{A_1}A_2\overline{A_3} \cup \overline{A_1}\,\overline{A_2}A_3 \cup A_1A_2\overline{A_3} \cup A_1\overline{A_2}A_3 \cup \overline{A_1}A_2A_3 \cup A_1A_2A_3.$$

第二节　概率的定义

随机事件在一次试验中可能发生，也可能不发生. 虽然在试验之前不能确定事件是否发生，但是，我们希望能够知道事件发生的可能性的大小，并且希望用数来表示这个可能性. 用来表示某事件发生的可能性大小的数，我们称其为该事件的**概率**，我们要解决的问题是：怎样定义事件的概率？

一、概率的统计定义

我们知道，随机事件在一次试验中是否发生并不确定，但是，经过大量试

验之后，事件的发生却有统计规律性.

如果在相同条件下将试验重复进行 n 次，事件 A 发生了 k 次，则称比值 $\frac{k}{n}$ 为事件 A 在这 n 次试验中发生的**频率**，记为 $Q_n(A)$. 例如，在抛掷硬币中，将"出现正面"记为事件 A，如果抛掷硬币 10 次，发现正面出现了 6 次，我们就说事件 A 在这 10 次试验中发生的频率 $Q_{10}(A)=\frac{6}{10}$. 为了观察统计规律性，前人做过大量此类试验，表 1-1 给出了部分试验结果：

表 1-1

抛掷硬币者	抛掷次数 n	正面出现次数 k	正面出现频率 $\frac{k}{n}$
浦丰	4040	2048	0.5069
皮尔逊	24000	12012	0.5005
维尼	30000	14994	0.4998

由表 1-1 可以看出，随着抛掷次数增大，出现正面的频率越来越接近于 0.5. 大量试验证实，当试验次数充分增大时，随机事件发生的频率总会趋近于某一个常数，随机现象的这一性质称为**频率的稳定性**. 基于频率的稳定性，我们可以这样给出事件概率的定义：

定义 1 如果在相同条件下进行 n 次试验，事件 A 发生的频率在某一常数附近摆动，并且随着试验次数 n 的充分增大，频率总会趋近于这个常数，则称这个常数为事件 A 的概率，记为 $P(A)$.

定义 1 称为**概率的统计定义**. 由此定义可以看出，事件 A 的概率 $P(A)$ 就是在一次试验中，对事件 A 发生的可能性大小的描述. 例如，我们用 0.5 来描述在抛掷均匀硬币时出现正面的可能性.

二、概率的古典定义

古典概型是概率论发展初期研究的主要概率模型之一，概率的定义这样给出：

定义 2 设随机试验 E 满足下列两个条件：

(1) E 的样本空间 $S=\{e_1,e_2,\cdots,e_n\}$ 只有 n(有限数) 个样本点，

(2) 每个基本事件发生的可能性相同，

则称随机试验 E 为古典概型或等可能概型. 如果事件 $A=\{e_{i_1},e_{i_2},\cdots,e_{i_k}\}\subset S$，即 A