复旦卓越·经济学系列

概率论与数理统计

王学民 编著

复旦大学出版社

复旦卓越·经济学系列

本书得到上海财经大学“211 工程”三期重点学科建设项目和上海市
重点学科建设项目（B803）资助

概率论与数理统计

王学民　编著

復旦大學出版社

图书在版编目(CIP)数据

概率论与数理统计/王学民编著. —上海:复旦大学出版社,2011.6
(复旦卓越·经济学系列)
ISBN 978-7-309-08094-0

Ⅰ. 概…　Ⅱ. 王…　Ⅲ. ①概率论②数理统计　Ⅳ. O21

中国版本图书馆 CIP 数据核字(2011)第 074379 号

概率论与数理统计
王学民　编著
责任编辑/王联合

复旦大学出版社有限公司出版发行
上海市国权路 579 号　邮编:200433
网址:fupnet@fudanpress.com　http://www.fudanpress.com
门市零售:86-21-65642857　团体订购:86-21-65118853
外埠邮购:86-21-65109143
大丰市科星印刷有限责任公司

开本 787×960　1/16　印张 17.25　字数 303 千
2011 年 6 月第 1 版第 1 次印刷
印数 1—4 100

ISBN 978-7-309-08094-0/O·469
定价: 29.00 元

内容提要

全书共分八章，第一章～第四章是概率论部分，内容有随机事件与概率、随机变量及其分布、多维随机变量及其分布、随机变量的数字特征。第五章～第八章是数理统计部分，内容有样本及抽样分布、参数估计、假设检验、方差分析与回归分析。对于打“*”的章节（或段落），读者可将其略过或作为选读内容。在数理统计部分，本书给出了一些例题的SAS软件菜单系统的输出结果。

本书是按照教育部新颁发的“概率论与数理统计的基本要求”进行编写的，可作为经济管理各专业概率论与数理统计课程的教材。读者可从作者的主页（http://bb.shufe.edu.cn/webapps/silkII/indexpage/teacherweb/teacherweb_3300_1.html）下载如下资料：（1）本书的全套PPT课件；（2）书中例题和习题的数据。此外，全书习题解答的电子版，需要的老师也可直接发送到unionw@sina.com免费索取。

前言

foreword

本书是按照教育部新颁发的“概率论与数理统计的基本要求”进行编写的，可作为高校经济管理各专业概率论与数理统计课程的教材，也可供其他各专业人员参考。

全书共分八章，第一章～第四章是概率论部分，内容有随机事件与概率、随机变量及其分布、多维随机变量及其分布、随机变量的数字特征。第五章～第八章是数理统计部分，内容有样本及抽样分布、参数估计、假设检验、方差分析与回归分析。

全书对概率统计的基础理论、思想方法进行了严谨的论述，同时又突出和展现了概率统计的应用，使读者既能为今后进一步学习数理统计的后继课程打好理论基础，又能切实地学会概率统计的初步应用。对于打“*”的章节（或段落），读者可将其略过或作为选读内容。在数理统计部分，本书给出了一些例题的SAS软件菜单系统的输出结果。

读者可从作者的主页（http://bb.shufe.edu.cn/webapps/silkII/indexpage/teacherweb/teacherweb_3300_1.html）下载如下资料：(1)本书的全套PPT课件；(2)书中例题和习题的数据；(3)作者编写的《应用概率统计》一书中SAS菜单系统部分的电子版。此外，全书习题解答的电子版，需要的老师可直接发送到unionw@sina.com免费索取。

因编者水平有限、经验不足，书中错误、不足之处在所难免，敬请读者批评指正。

王学民

2011年5月

目 录
contents

第一章

随机事件与概率

在日常生活中，我们经常会遇到不确定性的问题，并且希望了解这种不确定性的程度。以下是这方面的一些例子。

(1) 明天下雨的“可能性”有多大?

(2) 股价在未来一周内上涨的“可能”有多大?

(3) 如果提高产品的价格，则销售量下降的“机率”有多少?

(4) 一个家庭中的所有四个孩子都是女孩的“机会”是多少?

(5) 某候选人被当选的“概率”是多大?

所有上述问题中的“可能性”、“机率”和“机会”等都是一个意思，今后我们一般就用规范的术语“概率”来表述。

概率简单地说就是一个数，它总是从 0 到 1 之间取值。小概率(接近 0)的事件很少发生，而大概率(接近 1)的事件则经常发生。例如，某一城市一年内发生七级以上地震的概率很小，而整个地球一年内发生七级以上地震的概率就很大了。

早在 17 世纪，就有零星的有关概率的文章出现。当时的绅士赌徒们试图确定牌和骰子赌博中的赔率，因而对概率发生了兴趣。赌博者希望小概率事件能有较高的赔率，而较高概率的事件则应有较小的赔率。赔率应与事件的概率相对应，这样才能做到公平，也就是投注者既不应最后轻易破产，也不应经常性地有过多的收益。这个问题被提到了当时的数学家面前，这就是概率论的起源。

本章中，我们将介绍随机事件和概率的一些基本概念，这部分内容是后面章节的基础，“概率”的说法将贯穿全书。

§1.1 随机事件及其运算

一、随机试验与随机事件

在一定条件下，至少有两个可能结果，但不能预先断定出现哪一个结果的现

象，称为**随机现象**。例如，抛一枚硬币、观察某一地区一天内的交通事故数和随机抽取一批产品所出现的废品率等，出现的结果都具有随机性，都是随机现象。

很多随机现象是可以大量重复的，对这种可重复的随机现象的观察称为**随机试验**，简称**试验**①，记为 E。并规定随机试验必须符合以下条件：

(1) 它可以在相同的条件下重复进行；

(2) 试验的所有可能的结果是事先已知的，并且不止一个；

(3) 每次试验只出现这些可能结果中的一个，但不能预先断定会出现哪个结果。

试验 E 的每一个可能的基本结果称为**样本点**，记为 ω。所有样本点的集合称为 E 的**样本空间**，记为 Ω。从集合论的角度来看，ω 是 Ω 的一个元素，即 $\omega \in \Omega$。

例 1.1.1 掷一颗骰子，试验结果定义为骰子朝上一面的数字。1, 2, …, 6 为样本点，样本空间 $\Omega = \{1, 2, \cdots, 6\}$。

例 1.1.2 抽检一件产品，如果把产品分为合格与不合格两类，则可取样本空间 $\Omega_1 = \{$合格品，不合格品$\}$；如果把产品分为四个等级，则可取样本空间 $\Omega_2 = \{$一等品，二等品，三等品，废品$\}$；如果把所有的产品都区别看待，则样本空间 $\Omega_3 = \{$全体产品$\}$。

例 1.1.3 一天内进入某超市的顾客数，其样本空间 $\Omega = \{0, 1, 2, \cdots\}$②。

例 1.1.4 观察电视机的寿命，其样本空间 $\Omega = \{t : t \geqslant 0\}$。

若干个样本点的集合称为**随机事件**，简称**事件**，记为大写字母 A, B, C 等。可见，事件是样本空间 Ω 的子集。特别地，由一个样本点组成的单点集称为**基本事件**。

在试验中，如果出现 A 中所包含的一个样本点，则称**事件 A 发生**。随机事件有两种极端的情况，一种是必然会出现的结果，称为**必然事件**；另一种是不可能出现的结果，称为**不可能事件**。从样本空间来看，必然事件是由其全部样本点组成的，故可记为 Ω；而不可能事件则不含有任何样本点，故记为空集 $\varnothing$。必然事件和不可能事件实质上都是确定性现象，已经失去了“随机性”。但为方便起见，仍将它们视作随机事件。

例 1.1.5 在例 1.1.1 中，我们研究事件

---

① 统计学中试验的概念与物理学中试验的概念是稍有不同的。当物理试验在相同的条件下重复进行时，会产生相同的试验结果。而在统计学中，即使试验在相同的条件下重复，结果也是随机确定的，可以得到完全不同的结果。

② 实际上，该例中的顾客数总有一个上限，不可能无限多。之所以取成无限是基于这样几点考虑：(1)顾客数的上限很难确定；(2)数学处理方便；(3)顾客数超过一个足够大数目的概率非常小，甚至为零，因而几乎不影响所作的分析。

$$A = \{出现的点数不超过 4\}$$
$$B = \{出现的点数是偶数\}$$
$$C = \{出现的点数不超过 6\}$$
$$D = \{出现的点数为 7\}$$

这些用文字表述的事件也可以表示为样本点的集合，即 $A = \{1, 2, 3, 4\}$，$B = \{2, 4, 6\}$，$C = \{1, 2, 3, 4, 5, 6\} = \Omega$，$D = \varnothing$。$C$ 是必然事件，D 是不可能事件。若在试验中出现的结果是 3，则事件 A 发生，事件 B 不发生。

二、事件间的关系及运算

由于事件可看作是一种集合，因此事件间的关系及运算与集合间的关系及运算是一致的。

随机事件和它们之间的关系可以用直观的几何图形来表示，这类图形称为**维恩(Venn)图**。

(1) 若事件 A 的每一个样本点都包含在事件 B 中，则称事件 B 包含事件 A，或称事件 A 包含于事件 B，记为 $B \supset A$ 或 $A \subset B$，如图 1.1.1 所示。$A \subset B$ 表明事件 A 发生必导致事件 B 发生。如果 $A \supset B$ 且 $A \subset B$，则称事件 A 与 B **相等**，记为 $A = B$，表明事件 A 与 B 是同一个事件。例如，掷一颗骰子，事件"出现奇数点"与事件"不出现偶数点"是相同的。

(2) 由或属于事件 A，或属于事件 B 的一切样本点组成的事件称为事件 A 与 B 的**并事件**或**和事件**，记为 $A \cup B$，如图 1.1.2 所示。$A \cup B$ 表明事件 A 与 B 中至少发生一个。在例 1.1.5 中，$A \cup B = \{1, 2, 3, 4, 6\}$。

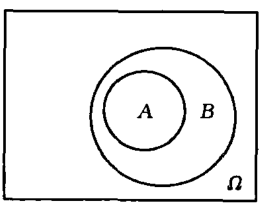

图 1.1.1　$A \subset B$

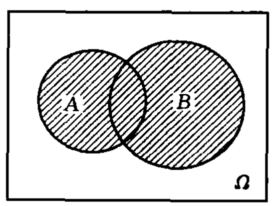

图 1.1.2　$A \cup B$

(3) 由事件 A 与事件 B 中共同的样本点组成的事件称为事件 A 与 B 的**交事件**或**积事件**，记为 $A \cap B$ 或 AB，如图 1.1.3 所示。$A \cap B$ 表明事件 A 与 B 同时发生。在例 1.1.5 中，$A \cap B = \{2, 4\}$。

(4) 由一切属于事件 A，但不属于事件 B 的样本点组成的事件，称为事件 A 与 B 的**差事件**，记为 $A - B$，如图 1.1.4 所示。$A - B$ 表明事件 A 发生而事件 B 不

发生。在例 1.1.5 中，$A-B=\{1, 3\}$。

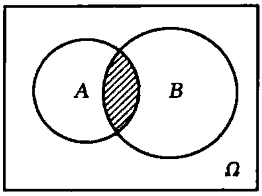

图 1.1.3　AB

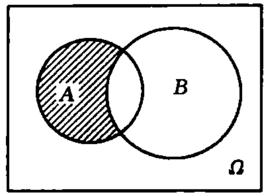

图 1.1.4　$A-B$

(5) 由样本空间中一切不属于事件 A 的样本点所组成的事件称为事件 A 的**对立事件**，记为 $\bar{A}$，如图 1.1.5 所示。$\bar{A}$ 表明事件 A 不发生。在例 1.1.5 中，$\bar{A}=\{5, 6\}$，$\bar{B}=\{1, 3, 5\}$。$\bar{A}$ 的对立事件就是原来的事件 A，即 $\bar{\bar{A}}=A$。所以对立事件是相互的，必然事件 Ω 与不可能事件 $\varnothing$ 互为对立事件。

(6) 若事件 A 与事件 B 没有共同的样本点，即 $AB=\varnothing$，则称事件 A 与 B 是**互不相容的**，如图 1.1.6 所示。$AB=\varnothing$ 表明事件 A 与 B 不可能同时发生。

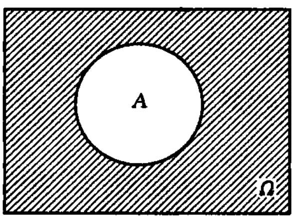

图 1.1.5　$\bar{A}$

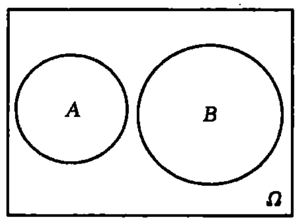

图 1.1.6　$AB=\varnothing$

事件并、交的概念可以推广到 n 个事件的情形。称 $\bigcup_{i=1}^{n} A_i$ 为 $A_1, A_2, \cdots, A_n$ 的并事件，表明事件 $A_1, A_2, \cdots, A_n$ 至少有一个发生。称 $\bigcap_{i=1}^{n} A_i$ 为 $A_1, A_2, \cdots, A_n$ 的交事件，表明事件 $A_1, A_2, \cdots, A_n$ 同时发生。这些概念还可进一步推广到可列个事件的情形。

例 1.1.6　设 A, B, C 是 Ω 中的随机事件，则

(1) “A 与 B 不发生，C 发生”可以表示成 $\bar{A}\bar{B}C$；

(2) “A, B, C 恰好发生一个”可以表示成 $A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C$；

(3) “A, B, C 中至少发生两个”可以表示成 $AB\cup AC\cup BC$；

(4) “A, B, C 中至多两个事件发生”可以表示成 $\bar{A}\bar{B}\bar{C}\cup A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C\cup AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC$，这个事件的对立事件是“$A, B, C$ 三个事件皆发生”，因此原事件亦可以表示成 $\Omega-ABC$。

例 1.1.7 事件 $A\cup B$ 可以写成互不相容事件的并，即有

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

事件的运算满足下述规则：

(1) 交换律：$A\cup B=B\cup A$，$AB=BA$；

(2) 结合律：$(A\cup B)\cup C=A\cup(B\cup C)$，$(AB)C=A(BC)$；

(3) 分配律：$(A\cup B)\cap C=AC\cup BC$，$(A\cap B)\cup C=(A\cup C)\cap(B\cup C)$；

(4) 德摩根(De Morgan)定理(对偶原则)：

$$\overline{A\cup B}=\bar{A}\bar{B},\quad \overline{AB}=\bar{A}\cup\bar{B}$$

德摩根定理的推广：

$$\overline{\bigcup_{i=1}^{n}A_i}=\bigcap_{i=1}^{n}\bar{A}_i,\ \overline{\bigcap_{i=1}^{n}A_i}=\bigcup_{i=1}^{n}\bar{A}_i$$

§1.2 事件的概率

随机事件 A 发生的可能性大小称为事件 A 发生的**概率**，记为 $P(A)$。在实践中，常用来确定概率的方法有三种：古典方法、频率方法和主观方法。

一、古典方法

确定概率的古典方法起源于 17 世纪很流行的赌博输赢的估计。例如，掷一枚均匀的硬币，出现正面与出现反面的可能性相同，都是 1/2；掷一颗完全均匀的骰子，出现每一个点数的可能性都是 1/6。这两个例子中的随机试验都具有这样两个特点：

(1) 样本空间 Ω 中的样本点只有有限个；

(2) 每个样本点出现的可能性相同。

具有上述两个特点的随机试验是概率论发展早期的主要研究对象，称为**古典型随机试验**，简称**古典概型**。相应地，其确定概率的方法称为**古典方法**，所得的概率称为**古典概率**。

在古典概型中，设 A 是样本空间 Ω 中的一个随机事件，则事件 A 的(古典)概率为

$$P(A)=\frac{A\text{ 中所含的样本点个数}}{\Omega\text{ 中的样本点总数}}\tag{1.2.1}$$

例 1.2.1 投掷两颗骰子，试求两颗骰子出现的点数之和为 7 的概率。

解 将两颗骰子分别记作骰子 1 和骰子 2，并记 (i,j) 为骰子 1 出现 i 点，骰子 2 出现 j 点。取样本空间 $\Omega=\{(i,j):i,j=1,2,\cdots,6\}$，则 Ω 中的样本点只

有有限个，且各样本点出现的概率相同，故这是一个古典概型。令

$$A=\{\text{两个骰子出现的点数之和为 }7\}$$
$$=\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$$

所以事件 A 发生的概率为

$$P(A)=\frac{A\text{ 中所含的样本点个数}}{\Omega\text{ 中的样本点总数}}=\frac{6}{36}=\frac{1}{6}$$

例 1.2.2 一批产品共有 N 件，其中不合格品有 M 件，现从中随机抽取 $n(\leqslant N-M)$ 件，试求恰好有 k 件不合格品的概率，这里 $k\leqslant\min(n,M)$。

解 将 n 件产品看成是从 N 件产品中同时抽取出来的，于是它们无次序之分，抽取的结果共有 $\binom{N}{n}$ 个。本例中“随机抽取”的意思是指所有的结果都是等可能出现的，故这是一个古典概型。令

$$A=\{\text{恰好有 }k\text{ 件不合格品}\}$$
$$=\{M\text{ 件不合格品中有 }k\text{ 件被抽中},N-M\text{ 件合格品中有 }n-k\text{ 件被抽中}\}$$

从 M 件不合格品中抽取 k 件有 $\binom{M}{k}$ 个结果，从 $N-M$ 件合格品中抽取 $n-k$ 件有 $\binom{N-M}{n-k}$ 个结果，从而事件 A 中含有 $\binom{M}{k}\binom{N-M}{n-k}$ 个样本点，因此事件 A 发生的概率为

$$P(A)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

在实践中，例 1.2.2 中的抽样结果通常是按不放回抽样的方式来实现的，即每次抽取一件产品，取后不放回。在对该例计算古典概率时，抽取的结果可以讲究次序，也可以不讲究次序。前者的样本空间 Ω_1 共含有 A_N^n（从 N 个不同元素中取 n 个元素的排列数）个样本点，后者的样本空间 Ω_2 共含有 $\binom{N}{n}$ 个样本点，两者都满足古典概型的条件。显然，Ω_1 比 Ω_2 大得多。取不讲究次序的较小样本空间 Ω_2 似乎更便于样本点个数的计算。

例 1.2.3 在例 1.2.2 中采用放回抽样的方式，即每次随机抽取一件产品，取后放回，同样求恰好有 k 件不合格品的概率。

解 需对结果讲究次序①，每次抽取时都有 N 种可能，于是 n 次放回抽取共有 N^n 个结果，显然所有结果出现的概率相同，故此为一个古典概型。令

$$\begin{aligned} A &= \{恰有 k 件不合格品\} \\ &= \{有 k 次抽到不合格品，有 n-k 次抽到合格品\} \end{aligned}$$

抽到的 k 件不合格品具体发生在 n 次抽取中的哪 k 次，共有 $\binom{n}{k}$ 种可能，对这每一种可能，"k 次抽到不合格品"有 M^k 个可能结果，"$n-k$ 次抽到合格品"有 $(N-M)^{n-k}$ 个可能结果，故事件 A 中含有的样本点个数为 $\binom{n}{k}M^k(N-M)^{n-k}$。因此，事件 A 发生的概率为

$$\begin{aligned} P(A) &= \frac{\binom{n}{k}M^k(N-M)^{n-k}}{N^n} \\ &= \binom{n}{k}\left(\frac{M}{N}\right)^k\left(1-\frac{M}{N}\right)^{n-k} \end{aligned}$$

例 1.2.4 箱中有 n 根签（完全相同，只是标记不同），其中有 m 根中奖签，现有 n 个人依次不放回地去抽签，试求第 k 个人抽到中奖签的概率。

解 记 $A_k=\{第 k 个人抽到中奖签\}$。

方法一 取样本空间

$$\Omega_1 = \{n 根签的所有排列\}$$

Ω_1 中共含有 $P_n=n!$ 个样本点，所有样本点出现的可能性相同，故这是一个古典概型。事件 A_k 可写成

$$A_k = \{第 k 个人抽到中奖签，其余 n-1 个人抽到的签任意排列\}$$

第 k 个人抽到中奖签有 m 种可能，对这每一种可能，其余 $n-1$ 个人抽到签的排列总数为 $P_{n-1}=(n-1)!$，从而事件 A_k 中含有的样本点个数为 $m(n-1)!$，故

$$P(A_k) = \frac{m(n-1)!}{n!} = \frac{m}{n}$$

方法二 取样本空间

① 若不讲究次序，则不满足古典概型的等可能性条件。

$$\Omega_2 = \{\text{前 } k \text{ 个人抽到签的所有排列}\}$$

共含有 A_n^k 个样本点，也是一个古典概型。事件 A_k 可写成

$$A_k = \{\text{第 } k \text{ 个人抽到中奖签，前 } k-1 \text{ 个人抽到的签任意排列}\}$$

A_k 中含有的样本点个数为 mA_{n-1}^{k-1}，故

$$P(A_k) = \frac{mA_{n-1}^{k-1}}{A_n^k} = \frac{m}{n}$$

方法三　简单地取样本空间

$$\Omega_3 = \{\text{第 } k \text{ 个人抽到签的所有结果}\}$$

共含有 n 个样本点，由于所有 n 个签在抽签中的地位是完全相同的，没有理由认为第 k 个人抽取时，一些签比另一些签有更大的可能性被抽中(这在概率论中称为“对称性”)，故可认为所有的样本点出现的可能性相同。事件 A_k 中含有的样本点个数为 m，因此

$$P(A_k) = \frac{m}{n}$$

上述例子清楚地表明，计算古典概率时，样本空间 Ω 取得越小，解法一般就越简单，当然必须保证 Ω 中的所有样本点是等概率出现的。从计算结果来看，不论先抽还是后抽，抽到中奖签的概率都是一样的，说明该抽签方式是公平的。

古典方法无需任何统计试验，可以由演绎方法准确计算得到。但是古典方法有其较大的局限性，如果试验的结果不是有限个的或不是等可能发生的，就无法用该方法来确定概率。例如，某公司正在销售一种新产品，欲估计顾客购买此产品的概率；保险公司要了解投保客户发生某种事故的概率等。类似这样的问题，使用古典方法是无法获得概率的，接下来介绍的频率方法将能够解决诸如此类的问题。

二、频率方法

频率方法是最常用，也是最基本的获得概率的方法。

设在相同的条件下重复 n 次试验，事件 A 发生的次数 n_A 称为事件 A 发生的**频数**，比值 n_A/n 则称为事件 A 发生的**频率**，记作 $f_n(A)$。虽然每重复 n 次试验得到的频率会有波动，但随着试验次数 n 的无限增大，频率 $f_n(A)/n$ 的波动会越来越小，并逐渐地稳定于某个数 p。这个数 p 就是用频率方法确定的概率，称为**统计概率或试验概率**。

易见，频率具有下述基本性质：

(1) $0 \leqslant f_n(A) \leqslant 1$；

(2) $f_n(\Omega) = 1$；

(3) 若事件 $A_1, A_2, \cdots, A_k$ 两两互不相容，则

$$f_n(A_1 \cup A_2 \cup \cdots \cup A_k) = f_n(A_1) + f_n(A_2) + \cdots + f_n(A_k)$$

例 1.2.5 历史上曾有人进行过"掷一枚均匀硬币"的试验，用来观察"出现正面"这一事件发生的规律，试验结果见表 1.2.1。

表 1.2.1 掷一枚硬币正面出现的频率

试验者	投掷次数(n)	出现正面次数(m)	频率(m/n)
德摩根	2 048	1 061	0.518 1
蒲丰(Buffon)	4 040	2 048	0.506 9
皮尔逊(K. Pearson)	12 000	6 019	0.501 6
皮尔逊(K. Pearson)	24 000	12 012	0.500 5

统计概率虽然用频率来解释，但频率和概率毕竟是有区别的。频率是个试验值，只能近似地反映事件发生的可能性大小；而概率是个理论值，它能准确地反映事件发生的可能性大小，它的值是唯一确定的。统计概率比较容易理解，但也应注意到，重复 $n+1$ 次试验的频率并非一定比重复 n 次试验的频率更接近真实的概率，而且现实世界中的试验也不可能无限制地进行下去。还有，在实际应用中，往往很难保证每次试验都是在完全相同的条件下重复进行的。不过，只要每次试验的条件近似相同，我们仍然可以用所得到的频率去估计概率。

*三、主观方法

在实际问题中，有些试验是无法在相同或近似相同的条件下重复进行的。例如，天空看上去阴沉沉的，估计下雨的可能性有多大；股价指数在未来一周内上升的可能性是多少；彗星碎片在未来 50 年内撞击地球的可能性有多大；明天的一场足球比赛甲队获胜的概率有多大；等等。这些事件都无法进行统计试验，也就不可能计算出事件发生的频率，因而只能通过直觉、个人信念和其他直接信息对事件发生的可能性大小作出主观的估计。这种估计概率的方法称为**主观方法**，估计的概率称为**主观概率**。主观概率依人的看法和经验而定，对同一事件不同的人会给出不同的概率。例如，对某种商品采用新的销售措施，销售经理认为能起到促销作用的概率为 0.8，而某位销售人员却认为这个概率只有 0.5。相对而言，前者为乐观的估计，后者的估计较为悲观。