成人教育/网络教育系列规划教材

概率论与数理统计

Probability and Mathematical Statistics

顾凤岐◎主　编

顾海燕　王　爽◎副主编

赵连盛◎主　审

人民交通出版社股份有限公司

China Communications Press Co.,Ltd.

成人教育/网络教育系列规划教材

概率论与数理统计

Probability and Mathematical Statistics

顾凤岐◎主　编

顾海燕　王　爽◎副主编

赵连盛◎主　审

人民交通出版社股份有限公司

China Communications Press Co.,Ltd.

内容提要

本书分上下篇，共10章。上篇（第1章至第5章）介绍了概率统计的基本概念，并对随机变量的概念、分布、性质及其数学特征进行了阐述。下篇（第6章至第10章）介绍了参数估计、假设检验、方差分析和回归分析等常用统计方法的理论与应用。此外，本书在每章中都编写了学习导读、学习目标、学习重点、学习难点和学习计划，有利于读者自学。

本书针对从事交通、土建生产与科研人员的自学而编写，也可作为本科、成人教育和高职院校教材使用。

图书在版编目（CIP）数据

概率论与数理统计/顾凤岐主编. —北京：人民交通出版社股份有限公司，2015.5

ISBN 978-7-114-12068-8

Ⅰ.①概… Ⅱ.①顾… Ⅲ.①概率论②数理统计 Ⅳ.①O21

中国版本图书馆CIP数据核字（2015）第032254号

书　　名：概率论与数理统计
著 作 者：顾凤岐
责任编辑：王　霞　陈力维　富砚博
出版发行：人民交通出版社股份有限公司
地　　址：（100011）北京市朝阳区安定门外外馆斜街3号
网　　址：http://www.ccpress.com.cn
销售电话：（010）59757973
总 经 销：人民交通出版社股份有限公司发行部
经　　销：各地新华书店
印　　刷：北京鑫正大印刷有限公司
开　　本：787×1092　1/16
印　　张：16.25
字　　数：370千
版　　次：2015年5月　第1版
印　　次：2015年5月　第1次印刷
书　　号：ISBN 978-7-114-12068-8
定　　价：32.00元

前　言

概率论与数理统计是研究随机现象的统计规律及其应用的学科，在交通、土建的生产和科研实践中有广泛的应用，是交通、土建类各相关专业不可缺少的一门基础课。

目前出版的概率论与数理统计教材很多，各有特色，质量也很高。本书在以往经典教材的基础上，根据成人教育的特点，以适应学生自学为目标而编写，每章增加了以下内容。

学习导读：统领全章，提纲挈领地介绍知识结构、体系和脉络。

学习目标：明确了对相关知识点的学习程度的原则要求。

学习重点：明确了需重点学习和掌握的内容。

学习难点：指出学习难点。

学习计划：建议相关知识点学习时间和学习进度。

本书由东北林业大学顾凤岐教授主编，顾海燕、王爽副主编。本书主审赵连盛认真审阅了全部书稿，并提出了许多宝贵修改意见。本书的编写过程中得到人民交通出版社股份有限公司的大力支持，并提出了很多很好的建议，在此一并表示感谢！

由于编者水平有限，加之时间仓促，如有错编欠妥之处，恳请读者不吝指正。

编　者

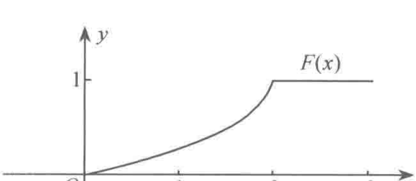
2015 年 1 月

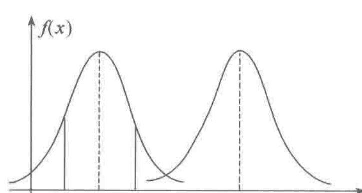
目　录

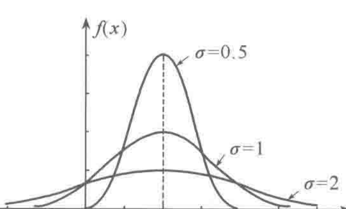
上篇　概　率　论

下篇 数理统计

上篇　概率论







第1章
概率论的基本概念

学习导读

本章主要介绍了样本空间的概念以及在概率论中的含义，描述了概率与频率的区别和联系，并给出了概率的一些基本性质，重点介绍了古典概型、条件概率、独立性、n 重伯努利试验的定义及其应用.

学习目标

了解 样本空间和概率的基本性质，古典概型、条件概率、独立性的基本概念.
掌握 古典概型、条件概率、独立性、n 重伯努利试验的定义.
理解 古典概型、条件概率、独立性、n 重伯努利试验的应用.

学习重点

利用古典概型、条件概率的公式解题.

学习难点

古典概型应用的条件；条件概率的性质；多重伯努利的应用条件.

学习计划

章　　节	建议自学（学时）	学 习 建 议	学习记录
1.1　样本空间、随机事件及运算	1.5	理解样本空间和随机事件的概念；掌握事件间的关系及其运算	
1.2　频率、概率、古典概型和几何概型	1.5	理解频率和概率之间的关系；掌握古典概型和几何概型的算法	
1.3　条件概率、乘法公式、全概率公式及贝叶斯公式	1.5	理解公式并会应用	
1.4　独立性	1.5	掌握相互独立事件概率的求法	

现实世界中的各种现象概括起来无非是两类现象:确定性现象和随机现象. 例如:水在通常条件下,温度达到 100℃时必然沸腾,温度为 0℃时必然结冰;同性电荷相互排斥,异性电荷相互吸引等,这类现象称为确定性现象,它们在一定的条件下一定会发生. 另有一类现象,在一定条件下,往往有多种可能的结果,事先并不能预测是哪一种结果,此类现象称为随机现象. 例如:测量一个物体的长度,其测量误差的大小;从一批电视机中随机取一台,电视机的寿命长短等都是随机现象. 概率论与数理统计,就是研究和揭示随机现象统计规律性的一门基础学科.

这里我们注意到,随机现象是与一定的条件密切联系的. 例如:在城市交通的某一路口,指定的 1 小时内,汽车的流量多少就是一个随机现象,而“指定的 1 小时内”就是条件,若换成 2 小时内、5 小时内,流量就会不同. 如将汽车的流量换成自行车流量,差别就会更大,故随机现象与一定的条件密切联系.

概率论与数理统计的应用是很广泛的,几乎渗透到所有科学技术领域,如工业、农业、国防与国民经济的各个部门. 例如,工业生产中,可以应用概率统计方法进行质量控制,工业试验设计,产品的抽样检查等. 还可使用概率统计方法进行气象预报、水文预报和地震预报等. 另外,概率统计的理论与方法也在向各基础学科、工程学科、经济学科渗透,产生了各种边缘性的应用学科,如排队论、计量经济学、信息论、控制论、时间序列分析等.

1.1 样本空间、随机事件及运算

1.1.1 随机试验

人们是通过试验去研究随机现象的,为对随机现象加以研究所进行的观察或实验,称为试验。若一个试验具有下列三个特点:

(1)可以在相同的条件下重复地进行;

(2)每次试验的可能结果不止一个,并且事先可以明确试验所有可能出现的结果;

(3)进行一次试验之前不能确定哪一个结果会出现.

则称这一试验为随机试验(Random Trial),记为 E.

下面举一些随机试验的例子.

E_1:抛一枚硬币,观察正面 H 和反面 T 出现的情况.

E_2:掷两颗骰子,观察出现的点数.

E_3:在一批电视机中任意抽取一台,测试它的寿命.

E_4:城市某一交通路口,统计一小时内的汽车流量.

E_5:记录某一地区一昼夜的最高温度和最低温度.

1.1.2 样本空间与随机事件

在一个试验中，不论可能的结果有多少，总可以从中找出一组基本结果，满足：

(1)每进行一次试验，必然出现且只能出现其中的一个基本结果.

(2)任何结果，都是由其中的一些基本结果所组成.

随机试验 E 的所有基本结果组成的集合称为样本空间(Sample Space)，记为 Ω. 样本空间的元素，即 E 的每个基本结果，称为样本点. 下面写出前面提到的试验 $E_k(k=1,2,3,4,5)$ 的样本空间 Ω_k：

$\Omega_1:\{H,T\}$；$\Omega_2:\{(i,j)\mid i,j=1,2,3,4,5,6\}$；$\Omega_3:\{t\mid t\geqslant 0\}$；$\Omega_4:\{0,1,2,3,\cdots\}$；

$\Omega_5:\{(x,y)\mid T_0\leqslant x\leqslant y\leqslant T_1\}$，这里 x 表示最低温度，y 表示最高温度，并设这一地区温度不会小于 T_0 也不会大于 T_1.

随机试验 E 的样本空间 Ω 的子集称为 E 的随机事件(Random Event)，简称事件[1]，通常用大写字母 $A,B,C,\cdots$表示. 在每次试验中，当且仅当这一子集中的一个样本点出现时，称这一事件发生. 例如，在掷骰子的试验中，可以用 A 表示“出现点数为偶数”这个事件，若试验结果是“出现 6 点”，就称事件 A 发生.

特别地，由一个样本点组成的单点集，称为基本事件. 例如，试验 E_1 有两个基本事件$\{H\}$、$\{T\}$；试验 E_2 有 36 个基本事件$\{(1,1)\},\{(1,2)\},\cdots,\{(6,6)\}$.

每次试验中都必然发生的事件，称为必然事件. 样本空间 Ω 包含所有的样本点，它是 Ω 自身的子集，每次试验中都必然发生，故它就是一个必然事件. 因而必然事件我们也用 Ω 表示. 在每次试验中不可能发生的事件称为不可能事件. 空集$\varnothing$不包含任何样本点，它作为样本空间的子集，在每次试验中都不可能发生，故它就是一个不可能事件. 因而不可能事件也用$\varnothing$表示.

1.1.3 事件之间的关系及其运算

事件是一个集合，因而事件间的关系与事件的运算可以用集合之间的关系与集合的运算来处理. 下面就讨论事件之间的关系及运算.

(1)如果事件 A 发生必然导致事件 B 发生. 则称事件 A 包含于事件 B(或称事件 B 包含事件 A)，记作 $A\subset B$(或 $B\Leftrightarrow A$).

$A\subset B$ 的一个等价说法是，如果事件 B 不发生，则事件 A 必然不发生.

若 $A\subset B$ 且 $B\subset A$，则称事件 A 与 B 相等(或等价)，记为 $A=B$.

为了方便起见，规定对于任一事件 A，有$\varnothing\subset A$. 显然，对于任一事件 A，有 $A\subset\Omega$.

---

[1] 严格地说，事件是指 Ω 中满足某些条件的子集. 当 Ω 是由有限个元素或由无穷可列个元素组成时，每个子集都可作为一个事件. 若 Ω 是由不可列无限个元素组成时，某些子集必须排除在外. 幸而这种不可容许的子集在实际应用中几乎不会遇到. 今后，我们讲的事件都是指它是容许考虑的那种子集.

(2)“事件 A 与 B 中至少有一个发生”的事件称为 A 与 B 的并(和)，记为 $A\cup B$.

由事件并的定义，立即得到：

对任一事件 A，有

$A\cup\Omega=\Omega$；$A\cup\varnothing=A$.

$A=\bigcup\limits_{i=1}^{n}A_i$ 表示“$A_1,A_2,\cdots,A_n$ 中至少有一个事件发生”这一事件.

$A=\bigcup\limits_{i=1}^{\infty}A_i$ 表示“可列无穷多个事件 A_i 中至少有一个发生”这一事件.

(3)“事件 A 与 B 同时发生”的事件称为 A 与 B 的交(积)，记为 $A\cap B$ 或(AB).

由事件交的定义，立即得到：

对任一事件 A，有

$A\cap\Omega=A$；$A\cap\varnothing=\varnothing$.

$B=\bigcap\limits_{i=1}^{n}B_i$ 表示“$B_1,\cdots,B_n$ 这 n 个事件同时发生”这一事件.

$B=\bigcap\limits_{i=1}^{\infty}B_i$ 表示“可列无穷多个事件 B_i 同时发生”这一事件.

(4)“事件 A 发生而事件 B 不发生”的事件称为 A 与 B 的差，记为 $A-B$.

由事件差的定义，立即得到：

对任一事件 A，有

$A-A=\varnothing$；$A-\varnothing=A$；$A-\Omega=\varnothing$.

(5)如果两个事件 A 与 B 不可能同时发生，则称事件 A 与 B 为互不相容(互斥)，记作 $A\cap B=\varnothing$.

基本事件都是两两互不相容的.

(6)若 $A\cup B=\Omega$ 且 $A\cap B=\varnothing$，则称事件 A 与事件 B 互为逆事件(对立事件). A 的对立事件记为 $\overline{A}$，$\overline{A}$ 是由所有不属于 A 的样本点组成的事件，它表示“事件 A 不发生”这样一个事件. 显然 $\overline{A}=\Omega-A$.

在一次试验中，若事件 A 发生，则 $\overline{A}$ 必不发生(反之亦然)，即在一次试验中，A 与 $\overline{A}$ 二者只能发生其中之一，并且也必然发生其中之一. 显然有 $\overline{\overline{A}}=A$.

对立事件必为互不相容事件；反之，互不相容事件未必为对立事件.

以上事件之间的关系及运算可以用文氏(Venn)图来直观地描述. 若用平面上一个矩形表示样本空间 Ω，矩形内的点表示样本点，圆 A 与圆 B 分别表示事件 A 与事件 B，则 A 与 B 的各种关系及运算如图 1-1～图 1-6 所示.

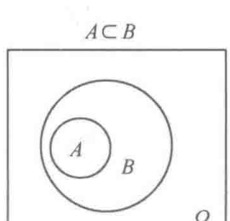


图　1-1

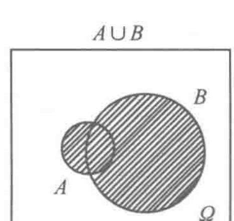


图　1-2

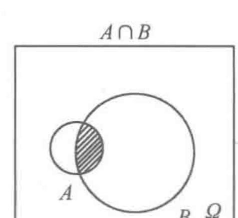


图　1-3

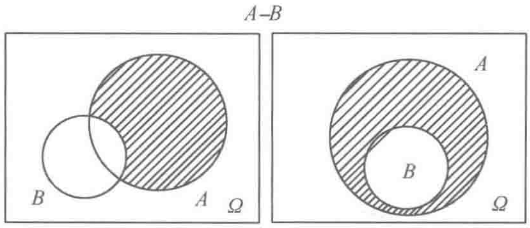


图 1-4

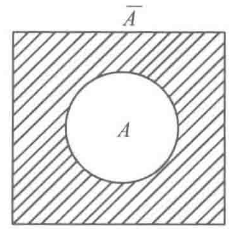


图 1-5

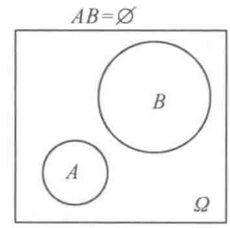


图 1-6

可以验证，一般事件的运算满足如下关系：

(1)交换律　$A\cup B=B\cup A, A\cap B=B\cap A$；

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

分配律可以推广到有穷或可列无穷的情形，即

$$A\cap(\bigcup_{i=1}^{n}A_i)=\bigcup_{i=1}^{n}(A\cap A_i), A\cup(\bigcap_{i=1}^{n}A_i)=\bigcap_{i=1}^{n}(A\cup A_i);$$

$$A\cap(\bigcup_{i=1}^{\infty}A_i)=\bigcup_{i=1}^{\infty}(A\cap A_i), A\cup(\bigcap_{i=1}^{\infty}A_i)=\bigcap_{i=1}^{\infty}(A\cup A_i).$$

(4)$A-B=A\overline{B}=A-AB$；

(5)对有穷个或可列无穷个 A_i，恒有

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

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

例 1-1　设 A、B、C 为三个事件，则可用 A、B、C 的运算式表示下列事件：

(1)A 发生而 B 与 C 都不发生：$A\overline{B}\overline{C}$，或 $A-B-C$，或 $A-(B\cup C)$.

(2)A、B 都发生而 C 不发生：$AB\overline{C}$ 或 $AB-C$.

(3)A、B、C 至少有一个事件发生：$A\cup B\cup C$.

(4)A、B、C 至少有两个事件发生：$(AB)\cup(AC)\cup(BC)$.

(5)A、B、C 恰好有两个事件发生：$(AB\overline{C})\cup(AC\overline{B})\cup(BC\overline{A})$.

(6) A、B、C 恰好有一个事件发生：$(A\overline{B}\overline{C})\cup(B\overline{A}\overline{C})\cup(B\overline{C}\overline{A})$.

(7)A、B 至少有一个发生而 C 不发生：$(A\cup B)\cap\overline{C}$.

(8)A、B、C 都不发生：$\overline{A\cup B\cup C}$ 或 $\overline{ABC}$.

例 1-2　在交通学院的学生中任选一名学生. 若事件 A 表示被选学生是男生，事件 B 表示该生是一年级学生，事件 C 表示该生是运动员.

(1)叙述 $AB\overline{C}$的意义.

(2)在什么条件下 $ABC=C$ 成立？

(3)在什么条件下 $\overline{A}\subset B$ 成立？

解：(1)该生是一年级男生，但不是运动员.

(2)全学院运动员都是一年级男生.

(3)全学院女生都在一年级.

例 1-3 设事件 A 表示“甲种产品畅销，乙种产品滞销”，求其对立事件$\overline{A}$.

解:设 B=“甲种产品畅销”，C=“乙种产品滞销”，则 $A=BC$，故

$$\overline{A}=\overline{BC}=\overline{B}\cup\overline{C}=\text{“甲种产品滞销或乙种产品畅销”}.$$

1.2 频率、概率、古典概型和几何概型

除必然事件与不可能事件外，任一随机事件在一次试验中都有可能发生，也有可能不发生. 人们常常希望了解某些事件在一次试验中发生的可能性的大小. 为此，我们首先引入频率的概念，它描述了事件发生的频繁程度，进而再引出表示事件在一次试验中发生的可能性大小的概率.

1.2.1 频率

定义 1.1 设在相同的条件下，进行了 n 次试验. 若随机事件 A 在 n 次试验中发生了 k 次，则比值$\frac{k}{n}$称为事件 A 在这 n 次试验中发生的频率(Frequency)，记为 $f_n(A)=\frac{k}{n}$.

由定义 1.1 容易推知，频率具有以下性质：

(1)对任一事件 A，有 $0\leqslant f_n(A)\leqslant 1$；

(2)对必然事件 Ω，有 $f_n(\Omega)=1$；

(3)若事件 A、B 互不相容，则

$$f_n(A\cup B)=f_n(A)+f_n(B).$$

一般地，若事件 $A_1,A_2,\cdots,A_m$ 两两互不相容，则

$$f_n\left(\bigcup_{i=1}^{m}A_i\right)=\sum_{i=1}^{m}f_n(A_i).$$

事件 A 发生的频率 $f_n(A)$表示 A 发生的频繁程度，频率大，事件 A 发生就频繁，在一次试验中，A 发生的可能性也就大，反之亦然. 因而，直观的想法是用 $f_n(A)$表示 A 在一次试验中发生可能性的大小. 但是，由于试验的随机性，即使同样是进行 n 次试验，$f_n(A)$的值也不一定相同。但大量实验证实，随着重复试验次数 n 的增加，频率 $f_n(A)$会逐渐稳定于某个常数附近，而偏离的可能性很小. 频率具有“稳定性”这一事实，说明了刻画事件 A 发生可能性大小的数——概率具有一定的客观存在性.(严格地说，这是一个理想的模型，因为我们在实际上并不能绝对保证在每次试验时，条件都保持完全一样，这只是一个理想的假设).

历史上有一些著名的试验，德·摩根(De Morgan)、蒲丰(Buffon)和皮尔逊(Pearson)曾进行过大量掷硬币试验，所得结果如表 1-1 所示.