高等学校教材

■ 陈家鼎　刘婉如　汪仁官　编著

概率统计讲义（第三版）

Probability and Statistics (Third Edition)

高等教育出版社

Probability and Statistics (Third Edition)

ISBN 7-04-014404-2

9 787040 144048 >

定价 18.20 元

高等学校教材

概率统计讲义

(第三版)

陈家鼎　刘婉如　汪仁官　编著

高等教育出版社

内容提要

本书是在第二版的基础上修订和扩充而成的，系统介绍了概率统计的基础理论和实用方法。内容简明扼要，文字通俗易懂。既注意对基本概念和定理论述准确，又注意介绍各方面的应用例子。只要求读者具有普通微积分知识和一些线性代数知识。本书可作为高等学校各类专业的教材，也可供有关人员参考。

图书在版编目(CIP)数据

概率统计讲义/陈家鼎，刘婉如，汪仁官编著．—3版．北京：高等教育出版社，2004.5

ISBN 7-04-014404-2

Ⅰ．概… Ⅱ．①陈…②刘…③汪… Ⅲ．①概率论－高等学校－教学参考资料②数理统计－高等学校－教学参考资料 Ⅳ．O21

中国版本图书馆CIP数据核字(2004)第019991号

出版发行	高等教育出版社	购书热线	010-64054588
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	http://www.hep.edu.cn
总　　机	010-82028899		http://www.hep.com.cn
经　　销	新华书店北京发行所		
印　　刷	北京凌奇印刷有限责任公司		
		版　　次	1980年7月第1版
开　　本	850×1168　1/32		2004年5月第3版
印　　张	14.5	印　　次	2004年7月第2次印刷
字　　数	370 000	定　　价	18.20元

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。

第一版序言

革命导师恩格斯说:“在表面上是偶然性在起作用的地方,这种偶然性始终是受内部的隐蔽着的规律支配的,而问题只是在于发现这些规律。”(见《马克思恩格斯选集》第四卷,第 243 页,1972 年版。)偶然事件的概率(即发生的可能性的大小)就是该偶然事件隐蔽着的特性,概率论与数理统计就是研究这种内在特性的一门数学学科。随着现代科学技术的迅速发展,这门数学学科也得到了蓬勃的发展。它不仅形成了结构宏大的理论,而且在很多科学研究、工程技术和经济管理的领域里有愈来愈多的应用。由于应用的广泛性,许多理工科专业(以及经济系科)都把“概率统计”列为学习课程,培养学生处理随机现象的能力。

去年 8 月,中央广播电视大学要北大数学系承担“概率统计”课的教学任务,加上校内一些系也需开设这门课,这促使我们考虑教材问题。这本“概率统计讲义”就是在这种形势的推动下,在我们以前编写的同名讲义的基础上,经过较大的改写、扩充而成的。

在这次编写工作中,我们注意了下列几点:

(1) 本书是针对 50~70 学时的讲课需要而编写的,只能讲解概率统计的一些基本内容与某些实用范围较广的方法,不能求多求全。但讲解详细,便于自学。凡小字排印部分都可略去不讲。在学时较紧的情况下,除前三章必须有足够时间教学外,其他各章都可以略去一部分。

(2) 努力贯彻理论联系实际的原则,对基本概念、重要公式和定理的实际意义多加解释,多举各方面的例子,力求通俗易懂,便于读者把所学的内容和实际工作结合起来。考虑到回归分析方法与正交试验法应用广泛,所以把它们分列专章,供读者选学。

(3) 虽然概率统计的严密的深入的数学理论不能离开实变函数论与测度论,但在目前情形下作为非数学专业用的概率统计教材,应该尽量少用专门的数学知识。这本讲义只用到普通的微积分知识,正文里基本上不用线性代数知识,有些结论不给出严密的数学证明。

一般说来,作为"概率统计"课的教材,应该有一章的篇幅介绍随机过程的最基本知识。但这次编写时间太紧,加上考虑到当前这门课教学时数的限制,故本书未涉及随机过程内容。

在这次编写过程中,我们参考了许多概率统计书籍和教材,特别是在例题和习题的选配方面,吸取了它们中的不少材料。我们还得到中国科学院系统科学研究所研究员张里千同志的帮助,谨在此致谢。

由于我们水平有限,加上编写时间仓促,书中的缺点、错误一定不少,欢迎读者批评指正。

编　　者

1980 年 1 月于北京大学数学系概率统计教研室

第三版序言

本版保留了第二版的绝大部分内容和优点，同时进行了较大的扩充，以便在内容上和编排上更好地适应高等学校各类专业“概率统计”课程的教学需要及概率统计这门学科的应用需要。本版的特点可概述如下：

（一）增添了许多重要内容（其中一些采用小字排印）。例如：

① 在“概率”的定义里，除了保留基本的“频率定义”外，还介绍了概率的“主观定义”和公理化定义；

② 对“条件分布”和“条件期望”作了较细致的介绍；

③ 对寻找置信区间的三种一般性方法进行了全面叙述；

④ 对“假设检验”中的 p 值方法进行了全面论述；

⑤ 介绍了有关两个正态总体的 Behrens－Fisher 问题的解；

⑥ 介绍了比率的检验方法（包括一个总体和两个总体的情形），特别是 Fisher 精确检验法；

⑦ 对逻辑斯谛回归作了初步介绍；

⑧ 叙述了统计决策和贝叶斯统计的大意；

⑨ 对随机过程的某些预测问题和统计问题作了初步介绍，等等。

（二）增加了许多实际应用例子。主要是增加了在日常生活、社会调查、商务管理、医学试验等方面的例子（第二版里的例子主要是工程方面的）。例如，在讲全概公式时介绍了敏感性社会调查的例子；在讲逆概公式时介绍了艾滋病的检查问题；在讲比率的两样本检验时介绍了两种药物疗效的比较；在讲回归分析时介绍了广告策略的制定，等等。

（三）内容编排上注意重点与非重点、难点与非难点、基本内

容与进一步内容的界限，做到层次分明、要求明确、方便教学。本版仍坚持第二版的编写原则，力求做到：内容简明扼要准确，文字通俗易懂流畅。虽然增加了许多内容（这是某些大学的课程所需要的），但新增内容的大部分或者用小字排印，或者打 * 号作为标志，用以表明这些内容不是"概率统计"课程的最基本内容。是否要求学生对这些内容了解或掌握，要根据课程的教学时数和学生的数学基础而定。

本版在编写过程中，吸取了北京大学概率统计系许多老师提出的宝贵意见，同时吸取了国内外近几年出版的多部概率统计优秀教材（见参考书目）的一些内容与讲法。在此，向所有帮助过我们的老师和一些教材的作者表示感谢。编者力图与时俱进，写出反映时代精神的合适教材。但限于水平，书中的缺点、谬误一定不少，欢迎读者批评指正。

编　者

2003 年 9 月

目　　录

第一章　随机事件与概率

§1　随机事件及其概率

粗略地说，在一定的条件下，可能发生也可能不发生的事件，称为**随机事件**(更确切的叙述见下面的定义).

例 1.1　投掷一枚分币，"正面朝上"这个事件(记作 A)，是一个随机事件.在该试验中，"正面朝下"(记作 B)，也是随机事件.(我们常把有币值的一面称为正面.)

例 1.2　投掷两枚分币，则

A＝"两个都是正面朝上"

B＝"两个都是正面朝下"

C＝"一个正面朝上，一个正面朝下"

都是随机事件.不难看出

D＝"至少有一个正面朝上"

也是随机事件.

例 1.3　从十个同类产品(其中有 8 个正品，2 个次品)中，任意抽取三个.那么，

A＝"三个都是正品"

B＝"至少一个是次品"

均为随机事件，而

"三个都是次品"和"至少一个是正品"

这两个事件呢，前者是不可能发生的；后者是必定要发生的.我们称不可能发生的事件为**不可能事件**，记作 V；称必定要发生的事件为**必然事件**，记作 U.为讨论问题方便起见，将不可能事件 V

和必然事件 U 也当作随机事件.

对于随机事件,在一次试验中是否发生,我们虽然不能预先知道,但是它们在一次试验中发生的可能性是有大小之分的.比如,在例 1.1 中,如果投掷的分币是匀称的,那么,随机事件 A(="正面朝上")和随机事件 B(="正面朝下")发生的可能性是一样的;在例 1.2 中,如果两个分币都是匀称的,那么随机事件 A(="两个都是正面朝上")和随机事件 B(="两个都是正面朝下")发生的可能性也是一样的,并且它们比随机事件 C(="一个朝上,一个朝下")发生的可能性要小.不仅如此,由我们的直觉还可以说,发生例 1.1 中随机事件 A(=投掷一枚分币出现"正面朝上")的可能性,比发生例 1.2 中随机事件 A(=投掷两枚分币,"两个都是正面朝上")的可能性要大.然而,对事件发生的可能性只停留在基本上是定性的了解与描述上,实在太不够了.我们希望对它给出客观的定量的描述.

怎样给出随机事件发生可能性大小的定量描述呢?用一个数——概率.随机事件 A 的概率用 $P(A)$ 表示.该数越大表明 A 发生的可能性越大."可能性大小"是人们凭直觉可以理解的观念,但怎样定义刻画"可能性大小"的概率呢?这就不是一个简单的问题.我们首先介绍概率的频率定义,然后介绍概率的主观定义(注意,是"主观定义",不是主观主义!),在 §4 中还要介绍概率的公理化定义.

回到例 1.1 中投掷一枚分币的试验,这种试验是在一定条件下作的.比如说,我们规定:"分币是匀称的,放在手心上,用一定的动作向上抛,让分币自由落在具有弹性的桌面上,等等."称这些条件为条件组 S.于是,在条件组 S 的一次实现下,事件 A("正面朝上")是否发生是不确定的.然而这只是问题的一方面.当条件组 S 大量重复实现时,事件 A 发生的次数,也称为**频数**,能体现出一定的规律性,约占总试验次数的一半.这也可以写成

$$A\text{ 发生的}\textbf{频率}=\frac{\text{频数}}{\text{试验次数}},\text{接近于}\frac{1}{2}$$

在我们的心目中，由长期经验积累所得的、所谓某事件发生的可能性的大小，不就是这个“频率的稳定值”吗？

历史上，有些人作过成千上万次投掷钱币的试验.下表列出他们的试验记录：

实验者	投掷次数 n	出现“正面朝上”的次数 μ(即频数)	频率 $=\mu/n$
DeMorgan	2 048	1 061	0.518
Buffon	4 040	2 048	0.506 9
Pearson	12 000	6 019	0.501 6
Pearson	24 000	12 012	0.500 5

容易看出，投掷次数越多，频率越接近 0.5.

定义 1.1 在不变的一组条件 S 下，重复做 n 次试验.记 μ 是 n 次试验中事件 A 发生的次数.当试验的次数 n 很大时，如果频率 μ/n 稳定地在某一数值 p 的附近摆动；而且一般说来随着试验次数的增多，这种摆动的幅度愈变愈小，则称 A 为随机事件，并称数值 p 为随机事件 A 在条件组 S 下发生的概率，记作

$$P(A)=p$$

显然，数值 p 就成为 A 在 S 下发生的可能性大小的数量刻画.例如 0.5 就成为掷一枚分币出现“正面朝上”的可能性的数量刻画.

上述定义也可简单地说成：“频率具有稳定性的事件叫做随机事件，频率的稳定值叫做该随机事件的概率.”

我们强调指出，人类的大量实践证明，在实际中遇到的事件一般都是随机事件，也就是说都是有确定的概率的.以后我们常简称随机事件为事件.

由于频率 $\frac{\mu}{n}$ 总介于 0,1 之间，因而由概率的定义知，对任何随机事件 A，有

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

而对必然事件 U 及不可能事件 V,显然有

$$P(U)=1, P(V)=0$$

定义 1.1 是概率的频率定义(又叫概率的统计定义).至于概率 $P(A)$的实际计算法,定义本身也给出了一种近似求法,即作大量的试验,计算事件 A 发生的频率.虽然得到的是近似值,但我们相信读者不至于因为现实生活中某一数值的获得只是些近似值而感到不实在.事实上,我们周围许多量的测量完全是近似的,如长度的概念并不会因为每次实测数值都是近似值而建立不起来,也不会因为温度计读数都是近似值而怀疑起“温度”的客观存在性.

以下介绍概率的主观定义.在现实世界里,有一些事件是不能重复或不能大量重复的,这时无法用上述定义 1.1 来定义概率.怎么办?一些统计学家认为,这样的事件不能定义概率,另一些统计学家(主要是贝叶斯(Bayes)学派的学者)则认为可以定义概率,他们认为应采用以下定义:

定义 1.2 一个事件的概率是人们根据已有的知识和经验对该事件发生可能性所给出的个人信念,这种信念用[0,1]中的一个数来表示,可能性大的对应较大的数.

定义 1.2 就是概率的主观定义,所定义的概率又叫做主观概率.粗一看,概率的主观定义很不科学,“个人信念”的主观色彩太浓.但仔细一想,现实世界中却有一些“可能性大小”是由个人信念来确定的,而且这样确定的概率合乎实际,对于人们的决策和行动有重要的指导作用.例如,一个企业家在某年某月某日说“此项产品在未来市场上畅销的概率是 0.8”.这里的 0.8 是根据他自己多年的经验和当时的一些市场信息综合而成的个人信念.如果这位企业家经验丰富,又有多次成功的业绩,我们就可以相信“畅销的概率是 0.8”.

又如一位外科医生要对一位心脏病患者做手术,他认为成功的概率是 0.9,这是他根据手术的难易程度、该病人的身体状况以

及自己的手术经验综合而成的个人信念.如果这位医生经验丰富，人们就会相信:手术成功的概率是0.9.

这样的例子很多.可见“主观概率”在一些情况下不可或缺,它是当事人对事件作了详细考察并充分利用个人已有的经验形成的“个人信念”,而不是没有根据的乱说一通.当然,“个人信念”毕竟是个人主观的东西,应该谨慎对待.我们的态度是,在事件不能重复或不便多次重复的情形下,采用概率的主观定义(定义1.2).采用“主观概率”时,“个人信念”中的“个人”应是有经验的人、专家或专家组.概率的主观定义乃是前面的频率定义(定义1.1)的一种补充[①].

§2 古典概型

上面介绍了概率的定义.定义1.1既是概念,同时又提供了近似计算概率的一般方法.但是在某些特殊情况下,并不需要临时做多次试验,也就是说临时多次实现条件组 S,从而求得概率的近似值,而是根据问题本身所具有的某种“对称性”,充分利用人类长期积累的关于“对称性”的实际经验,分析事件的本质,就可以直接计算其概率(采用定义1.2可得到相同结果).

例如上节的例1.1,即使我们不临时作大量的投掷试验,我们也会想到,“正面朝上”与“正面朝下”出现的机会相等.因此,可以推测在大量试验中“正面朝上”这件事发生的频率在0.5左右,即

---

① 概率既是可能性大小的度量,它不仅在自然科学、技术科学、社会科学中应用广泛,在思维科学中也起着重要的作用.大家知道,演绎法和归纳法是最重要的两种推理方法,二者相互补充、缺一不可.演绎推理的特点是,前提 A 与结论 B 间有必然关系:若 A 成立,则 B 一定成立;归纳推理的特点是,前提 A 与结论 B 间有或然关系:若 A 成立,则 B 可能成立.对于归纳推理(日常生活和科学研究中的大量推理属于归纳推理)来讲,“B 成立的可能性有多大”十分重要.在 A 成立的条件下 B 成立的概率就是所谓从 A 到 B 的“归纳强度”.对归纳法的深入研究离不开概率论.本书后面要讲的“统计推断”就是一种归纳推理.

它的概率为 0.5. 为什么“正面朝上”与“正面朝下”机会均等呢？这是因为问题本身有一种对称性(匀称的分币)，如果“朝上”与“朝下”出现的机会不相等，那反倒与我们长期形成的“对称”的经验不相符了.

例 2.1 盒中装有五个球(三个白球，二个黑球)从中任取一个，问：取到白球的概率是多少?

既然是“任取”，那么五个球被取到的机会一样，而白球有三个，因此，取到白球的概率应该是 3/5. 说得更清楚些，我们把五个球编上号如下(其中白球为 1,2,3 号；黑球为 4,5 号)：

①　②　③　④　⑤

因为是随便取一个，所以

“取到 1 号球”，“取到 2 号球”，“取到 3 号球”

“取到 4 号球”，“取到 5 号球”

这些结果发生的机会一样，而且是互相排斥的，以及除此以外不可能有别的结果. 注意到 1,2,3 号球是白球，所以“取到白球”这个事件发生的频率会稳定在 3/5 左右，因此按概率定义，它的概率是 3/5.

例 2.2 盒中装有球的情况如上例，现从中任取两个，问两个球全是白球的概率是多少?

这个问题较为复杂，不过仍可按上例的方法进行分析. 还是把五个球同样编号，因为是随便取两个，所以下列这些结果

“①,②”[①],　“①,③”

“①,④”,　“①,⑤”

“②,③”,　“②,④”

“②,⑤”,　“③,④”

“③,⑤”,　“④,⑤”

发生的机会一样，而且是互相排斥的，以及除此之外不可能有别的

---

① “①,②”是“取到 1,2 号球”的缩写. 下同

结果.再注意到,上列十种情况中,有且仅有三种,即"①,②","①,③","②,③"为全白.因此"全白"发生的频率会稳定在 3/10 左右.于是,它的概率是 3/10.

推而广之,对上面几个例子所讨论的问题及解决问题的办法进行归纳,可得出一般规律.

定义 2.1 称一个事件组 $A_1, A_2, \cdots, A_n$ 为一个**等概完备事件组**,如果它具有下列三条性质:

(1) $A_1, A_2, \cdots, A_n$ 发生的机会相同(等可能性);

(2) 在任一次试验中,$A_1, A_2, \cdots, A_n$ 至少有一个发生(也就是所谓"除此之外,不可能有别的结果")(完备性);

(3) 在任一次试验中,$A_1, A_2, \cdots, A_n$ 至多有一个发生(也就是所谓"它们是互相排斥的")(互不相容性).

等概完备事件组在这里也称为**等概基本事件组**;其中任一事件 A_i $(i=1,2,\cdots,n)$ **称为基本事件**.

(在例 1.1 中,等概基本事件组的 $n=2$,它的两个基本事件是"正面朝上"与"正面朝下".读者可对例 2.1 和例 2.2 分别找出等概基本事件组.)

若 $A_1, \cdots, A_n$ 是一个等概基本事件组,而事件 B 由其中的某 m 个基本事件所构成[①].大量实践经验表明,事件 B 的概率应由下列公式来计算[②]:

$$P(B) = m/n \tag{2.1}$$

---

① 更确切地说,所谓事件 B 由事件 $A_{i_1}, A_{i_2}, \cdots, A_{i_m}$ 构成,是指当且仅当这 m 个事件中有一个发生时事件 B 才发生.

② 通常,如果试验只可能有有限个不同的试验结果 $A_1, A_2, \cdots, A_n$;而且它们发生的机会相同,则不难看出,$A_1, A_2, \cdots, A_n$ 就是一个等概基本事件组.(因此,解决这类问题主要是把 n 和 m 数出来.)

"只可能有有限个不同的试验结果"中的"试验结果"一词,是比较朴素的、直观的、方便的,一般而言,也是不会引起混淆的(今后我们有时也用这个词).然而,毕竟不够准确,因此我们引进了等概完备事件组的概念

所谓古典概型就是利用关系(2.1)来讨论事件的概率的模型.

现在通过(2.1)式来讨论例2.2.考虑从三个白球两个黑球中任取两球,我们知道共有 $C_5^2=\frac{5\times4}{1\times2}=10$ 种不同的取法,它们出现的机会相同.每一种取法对应一个基本事件,所以等概基本事件组共含 $n=10$ 个事件(读者不难验证它的"完备性"和"互不相容性").而取得两球均为白球,共有 $m=C_3^2=\frac{3\times2}{2\times1}=3$ 种取法(即由三个基本事件构成),由(2.1)有

$$P(\text{取得两个白球})=\frac{m}{n}=\frac{3}{10}$$

下面再看几个例子.

例 2.3 设有一批产品共100件,其中有5件次品,现从中任取50件,问:无次品的概率是多少?

解 首先,从100件产品中任取50件,我们知道共有 C_{100}^{50} 个不同的结果,每一个结果就是一个事件.容易验证这些事件是一个等概基本事件组(是否等可能?是否完备?是否互不相容?读者自己想一想).

现在来看 $B=$"任取50件其中无次品",它由哪些基本事件所构成?多少个?很明显,要所取的50件中无次品,必须是从那95件正品中取来的.可见这种无次品的取法共有 C_{95}^{50} 种(即事件 B 含 C_{95}^{50} 个基本事件).

由关系式(2.1)得

$$\begin{aligned}P(B)&=C_{95}^{50}/C_{100}^{50}\\&=\frac{95!/(50!45!)}{100!/(50!50!)}\\&=\frac{50\cdot49\cdot48\cdot47\cdot46}{100\cdot99\cdot98\cdot97\cdot96}\\&=\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{47}{99}\cdot\frac{46}{97}\\&=\frac{1\,081}{38\,412}=2.8\%\end{aligned}$$

（请读者将本例跟例 2.2 进行比较.）

再考虑较为复杂的情形.

例 2.4 条件组 S 跟例 2.3 相同(即还是 100 件产品,其中有 5 件次品,从中任取 50 件),问:恰有两件次品的概率是多少?

解 等概基本事件组同例 2.3,总数 $n = C_{100}^{50}$. 现在,问题的关键在于,计算出事件 $A =$“恰有两件次品”所包含的基本事件数.

取出的 50 件中,恰有两件次品,即有 48 件正品,2 件次品. 这 48 件正品必是从 95 件正品中取出的,共有 C_{95}^{48} 种;而 2 件次品必是从 5 件次品中取出的,共有 C_5^2 种. 因此,“恰有两件次品”共包含 $C_{95}^{48} \cdot C_5^2$ 个基本事件.

于是,据(2.1)得

$$
\begin{aligned}
P(A) &= C_{95}^{48} \cdot C_5^2 / C_{100}^{50} \\
&= \frac{95!}{48!\,47!} \cdot \frac{5!}{2!\,3!} \Big/ \frac{100!}{50!\,50!} \\
&\approx 0.32
\end{aligned}
$$

就是说,任取 50 件,恰有两件次品的概率是 0.32.

例 2.5 设一批产品共 N 个,其中次品共 M 个(其他是正品). 现从中任取 n 个,问:恰好出现 m 个次品的概率是多少($n - m \leqslant N - M \quad 0 \leqslant m \leqslant n, m \leqslant M$)?

这是比例 2.4 更普遍的问题. 经过与例 2.4 同样的推理,可以知道

$$
P(\text{恰好出现 } m \text{ 个次品}) = \frac{C_{N-M}^{n-m} \cdot C_M^m}{C_N^n} \tag{2.2}
$$

现在来讨论比例 2.5 更一般的情形,我们可以证明一条在计算概率时十分有用的定理.

定理 2.1 设有 N 个东西分成 k 类,其中第 i 类有 N_i 个东西 $(i = 1, \cdots, k)$, $N_1 + \cdots + N_k = N$,从这 N 个东西中任取 n 个,而 $n = m_1 + m_2 + \cdots + m_k (0 \leqslant m_i \leqslant N_i, i = 1, \cdots, k)$,则事件 $A =$“恰有 m_1 个属于第 1 类,恰有 m_2 个属于第 2 类,…,恰有 m_k 个属于

第 k 类"的概率为

$$P(A)=\frac{C_{N_1}^{m_1}\cdot C_{N_2}^{m_2}\cdot\cdots\cdot C_{N_k}^{m_k}}{C_N^n} \tag{2.3}$$

证 我们可以用符号 $g_1,g_2,\cdots,g_N$ 表示这 N 个帆西. 任取 n 个,所有可能的结果共有 C_N^n 种. 每一种结果都是 n 个东西的组合. 每一种结果(即每一个组合)都看成一个基本事件,故共有 C_N^n 个基本事件,它们是等概的、互不相容的、完备的.

事件 A 包含多少个这样的基本事件呢? 这是关键问题. 根据事件 A 的定义,为使一个基本事件(即一个"组合")包含在 A 里,必须且只需这个基本事件(即这个"组合")里恰有 m_1 个来自第 1 类,m_2 个来自第 2 类,$\cdots$,m_k 个来自第 k 类. 从第 i 类任取 m_i 个东西,共有 $C_{N_i}^{m_i}$ 种结果($i=1,\cdots,k$). 把各类中取出的一种结果并在一起,所得到的 n 个东西正是包含在 A 中的基本事件. 按照乘法原理知 A 所含的基本事件数为 $C_{N_1}^{m_1}\cdot C_{N_2}^{m_2}\cdot\cdots\cdot C_{N_k}^{m_k}$. 利用公式(2.1)即知公式(2.3)成立. 定理 2.1 证毕.

习 题 一

1. 求例 1.2 及例 1.3 中的 $P(A)$,$P(B)$.

2. 袋中有红、黄、白色球各一个,每次任取一个,有放回地抽三次,求下列事件的概率:

A = "三个都是红的" = "全红",B = "全黄",C = "全白",D = "颜色全同",E = "全不同",F = "不全同",G = "无红",H = "无黄",I = "无白",J = "无红且无黄",K = "全红或全黄"

3. 从一副扑克的 52 张牌中,任意抽取两张,问都是黑桃的概率有多大?

4. 在例 2.4 中,求至少有两件次品的概率.

5. 五人排队抓阄,决定谁取得一物(即五个阄中有四个是白阄,只有一个是有物之阄).(1) 问第三人抓到有物之阄的概率是多少? (2) 前三人之一抓到有物之阄的概率是多少? (3) 如果有两物(即五个阄中有两个是有物之阄),问后两个人都抓不到有物之阄的概率是多少?

§3 事件的运算及概率的加法公式

我们常常看到，在一组条件之下，有多个随机事件．其中有些是比较简单的，也有比较复杂的．分析事件之间的关系，从而找到它们的概率以及概率之间的关系，这自然是必要的．而其基本点还是要搞清楚事件间的关系．

1．事件的包含与相等

设有事件 A 及 B．如果 A 发生，那么 B 必发生，就称事件 B 包含事件 A，并记作

$$A \subset B \text{ 或 } B \supset A$$

例如投掷两枚匀称的分币，令 A 表示“正好一个正面朝上”，B 表示“至少一个正面朝上”，显然有 $A \subset B$．

如果事件 A 包含事件 B，同时事件 B 也包含事件 A，那么就称事件 A 与 B 相等（或称等价），并记作

$$A = B$$

2．事件的并与交

定义 3.1 事件“A 或 B”称为事件 A 与事件 B 的**并**，记作 $A \cup B$ 或 $A + B$；某次试验中 $A \cup B$ 发生，即“A 或 B”发生，它意味着 A，B 中至少有一个发生．事件“A 且 B”称为事件 A 与 B 的**交**，记作 $A \cap B$ 或 AB 或 $A \cdot B$；$A \cap B$ 发生，即“A 且 B”发生，它意味着 A，B 都发生．

例如，投掷两枚匀称的分币，A 表示“正好一个正面朝上”的事件，B 表示“正好两个正面朝上”的事件，C 表示“至少一个正面朝上”的事件．于是有

$$A \cup B = C, \quad AC = A$$
$$BC = B, \quad AB = V\text{（不可能事件）}$$

把事件的并与交的概念推广到多于两个事件的情形是不困难的，请读者自行完成．

3. 对立事件及事件的差

定义 3.2 事件“非 A”称为 A 的**对立事件**，记作$\overline{A}$.

例如，投掷两枚分币，事件“至少一个正面朝上”是事件“两个都是正面朝下”的对立事件.

由该定义可知

$$(\overline{\overline{A}}) = A$$

即 A 也是$\overline{A}$的对立事件. 我们看到：

在一次试验中，A 和$\overline{A}$不会同时发生(即它们互相排斥)而且 A，$\overline{A}$至少有一个发生. 就是说，A 和$\overline{A}$满足：

$$A \cap \overline{A} = V$$

$$A \cup \overline{A} = U \tag{3.1}$$

定义 3.3 事件 A 同B 的**差**表示A 发生而B 不发生的事件，记作 $A \setminus B$.

由上述定义可知

$$A \setminus B = A \cap \overline{B} \tag{3.2}$$

再举一个打靶的例子. 事件 A 代表命中图 1.1(a)的小圆内，事件 B代表命中图 1.1(b)的大圆内. 则 $A \cup B$ 代表命中图 1.1(c)

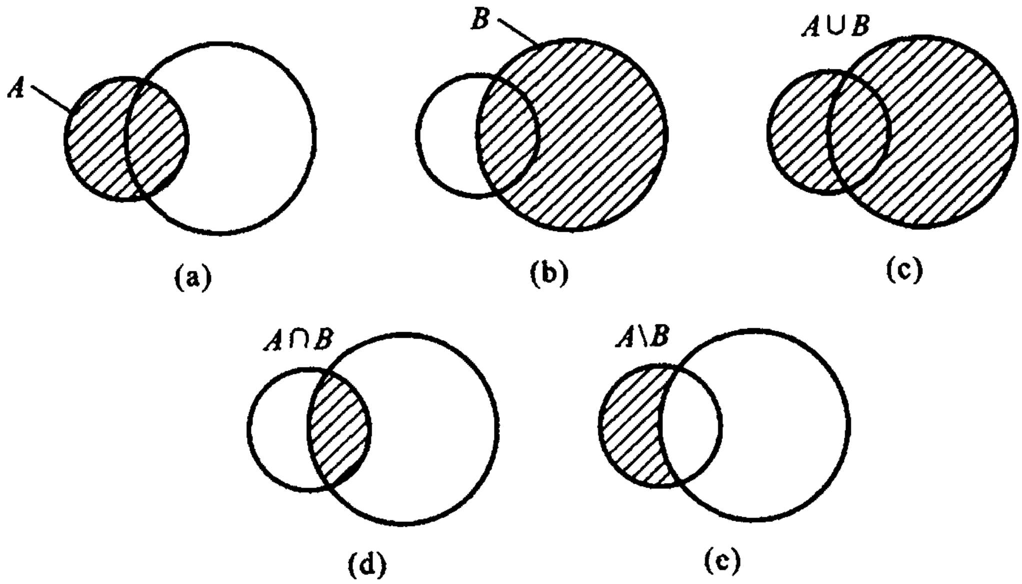


图 1.1

的阴影. $A\cap B$代表命中图 1.1(d)的阴影, $A\setminus B$ 代表命中图 1.1(e)的阴影.

4. 事件的运算规律

由定义不难验证,事件的运算满足以下规律:

(1) $A\cup B=B\cup A$ （“并”有交换律）

(2) $A\cup(B\cup C)=(A\cup B)\cup C$ （“并”有结合律）

(3) $A\cup A=A$

(4) $A\cup\overline{A}=U$

(5) $A\cup U=U$

(6) $A\cup V=A$

(7) $A\cap B=B\cap A$ （“交”有交换律）

(8) $(AB)C=A(BC)$ （“交”有结合律）

(9) $A\cap A=A$

(10) $A\cap\overline{A}=V$

(11) $A\cap U=A$

(12) $A\cap V=V$

(13) $A(B\cup C)=AB\cup AC$ （分配律）

(14) $A\cup(BC)=(A\cup B)(A\cup C)$ （分配律）

(15) $\overline{A\cup B}=\overline{A}\cap\overline{B}$

(16) $\overline{A\cap B}=\overline{A}\cup\overline{B}$

5. 事件的互不相容性

定义 3.4 如果事件 A 与事件 B 不能都发生,即

$$AB=V(\text{不可能事件})$$

那么,称 A 与 B 是互不相容的事件.

例如,投掷两枚分币,事件“正好一个正面朝上”和“两个都是正面朝上”就是互不相容的事件.不难看出, A 与$\overline{A}$也是互不相容的.

(称 n 个事件 $A_1,A_2,\cdots,A_n$ 是互不相容的,如果它们两两互

不相容.这个定义显然跟等概完备事件组中的“互相排斥”是一致的.)

6. 概率的加法公式

概率的加法公式(1):

如果事件 A,B 互不相容,则

$$P(A\cup B)=P(A)+P(B) \tag{3.3}$$

公式(3.3)表达了概率的最重要的特性:可加性.它是从大量的实践经验中概括出来的,成为我们研究概率的基础与出发点.从概率的定义来看,这个公式的成立是很自然的.设想把条件 S 重复实现了 n 次(n 充分大),其中事件 A 发生了 μ_1次,事件 B 发生了 μ_2次,由于 A 与 B 互不相容,故 $A\cup B$ 发生了 $\mu_1+\mu_2$次.但根据概率的定义,$\frac{\mu_1}{n}$应该与 $P(A)$ 很接近,$\frac{\mu_2}{n}$应该与 $P(B)$ 很接近.但是 $\frac{\mu_1}{n}+\frac{\mu_2}{n}=\frac{\mu_1+\mu_2}{n}$,于是 $\frac{\mu_1+\mu_2}{n}$ 自然应该与数值 $P(A)+P(B)$ 很接近.然而 $\frac{\mu_1+\mu_2}{n}$ 恰好是事件 $A\cup B$ 发生的频率,既然 n 充分大,所以 $\frac{\mu_1+\mu_2}{n}$ 与 $P(A\cup B)$ 很接近.因而 $P(A\cup B)$ 应该与 $P(A)+P(B)$ 相等.

由于 $A\cup\overline{A}=U$,A 与 $\overline{A}$ 互不相容,由加法公式(3.3)有

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

从而得

$$P(A)=1-P(\overline{A}) \tag{3.4}$$

此式虽然简单,却很有用.比如在习题一第 2 题中,$F=$“不全同”与 $D=$“全同”是对立的,即 $\overline{F}=D$,由于 $P(D)=\frac{3}{27}$(此处 $n=27$, $m=3$),故有 $P(F)=1-P(D)=1-\frac{3}{27}=\frac{24}{27}=\frac{8}{9}$.

公式(3.3)不难推广到 n 个事件的情形. 设 n 个事件 $A_1, A_2, \cdots, A_n$ 互不相容,则①

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n) \tag{3.5}$$

式(3.5)称为概率的有限可加性. 它可从(3.3)推导出来,证明留给读者.

概率的加法公式(2):

对任意两个事件 A, B,有

$$P(A \cup B) = P(A) + P(B) - P(AB) \tag{3.6}$$

此公式可利用加法公式(1)即(3.3)来证.

证 不难看出 $A \cup B = A \cup B\overline{A}$,由于 A 和 $B\overline{A}$ 互不相容,于是按(3.3)式有

$$P(A \cup B) = P(A) + P(B\overline{A}) \tag{3.7}$$

又由于

$$B = BA \cup B\overline{A}$$

且 BA 和 $B\overline{A}$ 也是互不相容的,由(3.3)式,有

$$P(B) = P(BA) + P(B\overline{A})$$

将 $P(B\overline{A}) = P(B) - P(BA)$ 代入(3.7),即得(3.6).

从以上证明过程看出:证明加法公式(2)时用到公式(1);且当 A 同 B 互不相容时,加法公式(1)是公式(2)的特殊情形.

---

① 作为公式(3.5)的应用,我们来推导出公式(2.1). 设 $A_1, A_2, \cdots, A_n$ 是一个等概基本事件组,据(3.5)知 $P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n) = nP(A_1)$,但 $A_1 \cup A_2 \cup \cdots \cup A_n$ 是必然事件,$P(A_1 \cup A_2 \cup \cdots \cup A_n) = 1$,于是 $P(A_1) = P(A_2) = \cdots = P(A_n) = \frac{1}{n}$,若 B 由 m 个基本事件组成,不妨设 $B = A_{i_1} \cup A_{i_2} \cup \cdots \cup A_{i_m}$ $(1 \leqslant i_1 < i_2 < \cdots < i_m \leqslant n)$. 据(3.5)知 $P(B) = P(A_{i_1}) + P(A_{i_2}) + \cdots + P(A_{i_m}) = \frac{m}{n}$. 这就证明了公式(2.1).

例 3.1 袋中有红、黄、白色球各一个,每次任取一个,有放回地抽三次.求"取到的三球里没有红球或没有黄球"的概率(参看习题一的第 2 题).

记 G="三球都不是红球",H="三球都不是黄球",我们所求的概率正好是 $P(G\cup H)$.容易看出 $P(G)=\frac{8}{27}$,$P(H)=\frac{8}{27}$,$P(GH)=P\{三球都是白球\}=\frac{1}{27}$,据公式(3.6)知

$$P(G\cup H)=P(G)+P(H)-P(GH)=\frac{15}{27}=\frac{5}{9}$$

作为本节的末尾,我们还要介绍一个重要概念:无穷多个事件的并.

定义 3.5 设 $A_1,A_2,\cdots,A_n,\cdots$是一列事件,$B$ 为这样的事件:它的发生当且仅当诸 $A_k(k=1,2,\cdots)$中至少一个发生.这个 B 就称为诸 A_k 的并(或叫做和),记作 $\bigcup\limits_{k=1}^{\infty}A_k\left(或\sum\limits_{k=1}^{\infty}A_k\right)$,有时也记作 $A_1\cup A_2\cup A_3\cup\cdots$.

类似地可定义无穷个事件的交.

例 3.2 一射手向某目标连续射击,决心射中为止.设 $A_1=\{第1次射击,命中\},\cdots,A_k=\{头\ k-1\ 次射击都未中,第\ k\ 次射击命中\}(k=2,3,\cdots)$,$B=\{终于命中\}$,从定义看出 $B=\bigcup\limits_{k=1}^{\infty}A_k$.

有一个比(3.5)更一般的公式,也是从实践经验中概括抽象出来的.

设 $A_1,A_2,\cdots,A_k,\cdots$是一列事件,如果 $A_k(k=1,2,\cdots)$两两互不相容,则有公式:

$$P\left(\bigcup_{k=1}^{\infty}A_k\right)=\sum_{k=1}^{\infty}P(A_k) \tag{3.8}$$

这个公式称为概率的"完全可加性".

习 题 二

1. 某产品40件,其中有次品3件.现从中任取两件,求其中至少有一件次品的概率.

2. 对立与互不相容有何异同?试举例说明.

3. A,B,C 三事件互不相容与 $ABC=V$ 是否是一回事?为什么?

4. 从一副扑克牌的13张黑桃中,一张接一张地有放回的抽取3次,问没有同号的概率(假定在13张中抽到任何一张的机会是一样的)?

5. 同第4题,问抽到有同号的概率?

6. 同第4题,问抽到的3张中最多只有两张同号的概率?

7. 将习题一第2题中的条件改为:盒中有四个球,其中两个红球、一个黄球、一个白球,其他不变,求 $A,B,\cdots,K$ 的概率,并求事件 $L=$"无红或无黄"的概率.

8. 利用加法公式(2)导出三个事件的概率加法公式.

9*. 利用加法公式(2)和数学归纳法证明下列若尔当(Jordan)公式:设 $A_1,A_2,\cdots,A_n(n\geqslant 2)$ 是 n 个事件,则

$$P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{k=1}^{n}\left((-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n} P\left(\bigcap_{l=1}^{k} A_{i_l}\right)\right)$$

$\left(\text{这里}\bigcap_{l=1}^{k} A_{i_l}\text{乃是事件}A_{i_1},A_{i_2},\cdots,A_{i_k}\text{之交}\right)$.

§4 集合与事件、*概率的公理化定义

为了更准确地理解和掌握事件的运算及其规律,我们需要介绍关于集合的一些知识以及集合运算与事件运算之间的联系.集合是全部现代数学里最基本最重要的概念之一.

定义4.1 一个集合是指具有确切含义的若干个东西的全体.

这些东西中的每一个东西称为这个集合的元素.通常用大写拉丁字母 $A,B,C,\cdots$ 来表示集合;而用小写拉丁字母 $a,b,c,\cdots$

表示元素. 如果 a 是 A 的一个元素,则称 a 属于 A,用记号“$a\in A$”表示. 如果 a 不是 A 的元素,则用记号“$a \overline{\in} A$”表示. 为了方便,数学上称没有元素的集合为空集合,简称空集,记作 $\varnothing$.

例 4.1 全体正整数组成一个集合 A,可表示为:

$A=\{1,2,3,\cdots\}$

例 4.2 不大于 10 的正整数的全体:

$B=\{1,2,\cdots,10\}$

例 4.3 圆心在原点的单位圆内点的全体:

$A=\{(x,y):x^2+y^2<1\}$

例 4.4 平面上面积为 1 个单位的矩形的全体.

从上面几个例子中看出,集合是一个非常一般的概念. 它的元素可以是数、平面上的点、几何图形等等.

例 4.5 §2 习题一第 2 题中的全体可能结果也是一个集合(记作 Ω),具体写出来是:

$\Omega=\{$(红,红,红),(红,红,黄),(红,红,白),
(红,黄,红),(红,黄,黄),(红,黄,白),
(红,白,红),(红,白,黄),(红,白,白),
(黄,红,红),(黄,红,黄),(黄,红,白),
(黄,黄,红),(黄,黄,黄),(黄,黄,白),
(黄,白,红),(黄,白,黄),(黄,白,白),
(白,红,红),(白,红,黄),(白,红,白),
(白,黄,红),(白,黄,黄),(白,黄,白),
(白,白,红),(白,白,黄),(白,白,白)$\}$

它共有 27 个元素.

$A=\{$(红,红,红)$\}$

也是一个集合,它仅含一个元素(红,红,红).

定义 4.2 如集合 A,B 的元素全同,则称 A,B 相等,用记号 $A=B$ 表示. 如 A 的元素都是 B 的元素,则称 B 包含 A(或称 A 包含在 B 中),也称 A 是 B 的一个子集合(简称子集),用记号

“$A\subset B$”或“$B\supset A$”表示. 我们规定空集合是任何集合的子集合.

由定义,显然有:

$$A=B\Longleftrightarrow A\subset B \text{ 且 } A\supset B$$

这里记号“$\Longleftrightarrow$”的意义是“当且仅当”.

定义 4.3 属于 A 或属于 B 的元素全体,称为 A,B 的并(和)集,记为 $A\cup B$,即

$$A\cup B\triangleq\{x:x\in A \text{ 或 } x\in B\}$$

这里,记号“$\triangleq$”的意义是按定义相等,下同.

定义 4.4 属于 A 且属于 B 的元素全体,称为 A,B 的交(积)集,记为 $A\cap B$,即

$$A\cap B\triangleq\{x:x\in A \text{ 且 } x\in B\}$$

我们不泛泛讨论集合间的关系,在这里只讨论某个非空集合 Ω 的若干个子集间的关系. 如§2 习题一第 2 题中的 11 个集合都是例 4.5 中的 Ω 的子集.

定义 4.5 设 A 是 Ω 的一个子集,Ω 中不属于 A 的元素全体称为 A 的余集,记为 A^c,即

$$A^c=\{x:x\in\Omega \text{ 但 } x\overline{\in} A\}$$

由定义,显然有:

$$(A^c)^c=A \tag{4.0}$$

当 Ω 是平面上一些点组成之集时,$A\cup B$,$A\cap B$,A^c 如图 1.2 阴影部分所示:

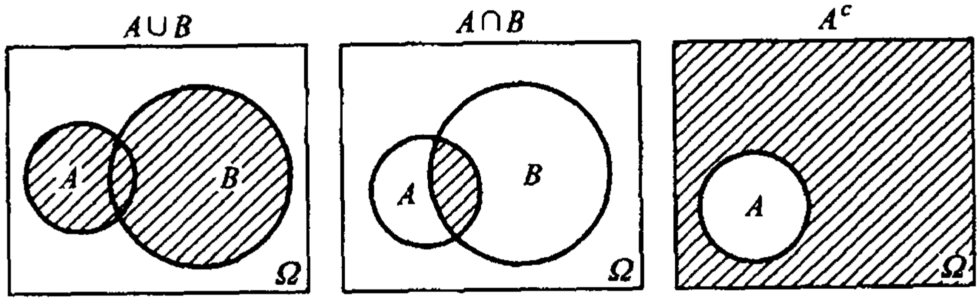


图 1.2

容易证明,集合的三个基本运算(并、交、余)[①]具有下列基本规则[②],当 Ω 是平面上的点集时,这些规则都具有非常明显的几何意义.

Ⅰ. 关于并

$$A \cup B = B \cup A \quad (\text{交换律}) \tag{4.1}$$

$$(A \cup B) \cup C = A \cup (B \cup C) \quad (\text{结合律}) \tag{4.2}$$

$$A \cup A = A \tag{4.3}$$

$$A \cup A' = \Omega \tag{4.4}$$

$$A \cup \Omega = \Omega \tag{4.5}$$

$$A \cup \varnothing = A \quad (\varnothing\text{是空集}) \tag{4.6}$$

Ⅱ. 关于交

$$A \cap B = B \cap A \quad (\text{交换律}) \tag{4.7}$$

$$(A \cap B) \cap C = A \cap (B \cap C) \quad (\text{结合律}) \tag{4.8}$$

$$A \cap A = A \tag{4.9}$$

$$A \cap A' = \varnothing \tag{4.10}$$

$$A \cap \Omega = A \tag{4.11}$$

$$A \cap \varnothing = \varnothing \tag{4.12}$$

Ⅲ. 关于并同交的分配律

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \quad (\text{第一分配律}) \tag{4.13}$$

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

---

① 还有所谓“差”运算,$A \setminus B \triangleq \{x : x \in A \text{ 但 } x \overline{\in} B\}$. 显然,$A \setminus B = A \cap B'$,它由交、余两个基本运算组合而得.

② 很多重要的关系式,可由这些基本规则推导而得. 比如对分解式 $B = (B \cap A) \cup (B \cap A')$ 可证明如下:

$$B \xlongequal{(4.11)} B \cap \Omega \xlongequal{(4.4)} B \cap (A \cup A') \xlongequal{(4.13)} (B \cap A) \cup (B \cap A')$$

又如对关系式 $A \cup B = A \cup (B \cap A')$,可证明如下:

$$A \cup (B \cap A') \xlongequal{(4.14)} (A \cup B) \cap (A \cup A') \xlongequal{(4.4)} (A \cup B) \cap \Omega \xlongequal{(4.11)} A \cup B$$

（第二分配律） (4.14)

Ⅳ. 关于并、交、余的对偶律

$$(A\cup B)^c = A^c\cap B^c \tag{4.15}$$

$$(A\cap B)^c = A^c\cup B^c \tag{4.16}$$

读者容易看出，我们在§3中定义的事件运算与这里定义的集合运算之间有很多相似之处. 事件的并运算与集合的“并”运算很相似，例如事件的并适合交换律，即事件 $A\cup$ 事件 $B=$ 事件 $B\cup$ 事件 A，集合的“并”也适合交换律，即集合 $A\cup$ 集合 $B=$ 集合 $B\cup$ 集合 A. 事件的交运算与集合的“交”运算也很相似，例如都适合交换律. 还可进一步看出，事件运算也具有与公式(4.0)至(4.16)相仿的法则. 在形式上，只需在这些公式里将拉丁字母 A, B, C 等理解为事件，A^c 理解为 A 的对立事件，Ω 理解为必然事件，$\varnothing$ 理解为不可能事件，这些公式就变成了事件运算的法则了.

我们说，集合的概念比起事件的概念来更为基本，更为单纯. 有了前者，后者可以定义得更明确，而且可以把事件的运算归结为集合的运算. 下面就是现在世界上流行的观点.

为了研究条件组 S 下的各种随机事件，我们把 S 下所有可能的不同结果（每个结果看成一个“基本事件”）的全体记作 Ω，S 下的随机事件就是若干个结果的集合（即 Ω 之子集），所谓观察到事件 A 发生就是指 S 下出现的结果属于 A. 显然，在这样的规定之下，Ω 本身就是一个必然事件，空集 $\varnothing$ 就是不可能事件. 如果 $A\subset\Omega$，则 $\Omega\setminus A=A^c$ 就是 A 的对立事件 $\overline{A}$. 所谓事件 A，B 互不相容，就是指 $A\cap B=\varnothing$. 而事件的“交”与“并”的运算与集合的交与并运算分别对应. 于是，事件运算完全归结为集合运算了.

例 4.6 投掷两枚分币（条件 S），所有可能的结果为：

$\omega_1=$“上，下”（第一枚分币正面朝上，第二枚正面朝下）

$\omega_2=$“上，上”，$\omega_3=$“下，上”，$\omega_4=$“下，下”

共有四个结果，$\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\}$.

事件 B（“恰有一个正面朝上”）正好由 ω_1 及 ω_3 组成，即 $B=$

$\{\omega_1, \omega_3\}$.

事件 C（"至少一个正面朝上"）正好由 $\omega_1, \omega_2, \omega_3$ 组成，即 $C = \{\omega_1, \omega_2, \omega_3\}$.

事件 A（"两个正面都朝上"）$= \{\omega_2\}$.

容易看出：$C = B \cup A$.

为了准确理解与深入研究随机现象，我们不能满足于从直觉出发形成的概率定义（频率的稳定值或可能性大小的"个人信念"），必须把概率论建立在坚实的数学基础上．苏联数学家科尔莫戈罗夫（Kolmogorov A N，1903—1987）于 1933 年在《概率论基本概念》一书中用集合论观点和公理化方法成功地解决了这个问题，得到了举世公认．简单介绍如下．

刚才说过，条件组 S 下所有可能的不同结果的全体记作 Ω．在科氏的公理系统里，Ω 是任何一个非空集合，称之为基本事件空间（有时叫做样本空间），其背景是条件组 S 下所有可能的结果的全体．

Ω 的一些子集组成的集合 $\mathcal{F}$ 叫做 σ 代数，若满足以下三条：

(1) $\Omega \in \mathcal{F}$

(2) 若 $A \in \mathcal{F}$，则 $A^c \triangleq \Omega - A \in \mathcal{F}$

(3) 若 $A_n \in \mathcal{F} (n = 1, 2, \cdots)$，则 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

$\mathcal{F}$ 上有定义的函数 $P = P(\cdot)$ 叫做概率测度（简称概率），若它满足下列三条：

(1) $P(A) \geqslant 0 \quad (A \in \mathcal{F})$ (4.17)

(2) $P(\Omega) = 1$ (4.18)

(3) 若 $A_n \in \mathcal{F} (n = 1, 2, \cdots)$，且两两不相交，则

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) \tag{4.19}$$

$\Omega, \mathcal{F}, P$ 乃是研究随机现象的三个要素．附有 $\mathcal{F}$ 和 P 的 Ω 叫做概率空间（或概率场）．若 $A \in \mathcal{F}$，则称 A 为随机事件，$P(A)$ 叫做 A 的概率（或称 A 发生的概率）．以上就是科尔莫戈罗夫的公理系统．$\mathcal{F}$ 的直观背景是：可以合理地定义概率的事件的全体．从 σ 代数的定义知，$\mathcal{F}$ 关于一些基本的集合运算是"封闭的"，即有下列结论：

① 若 $A_i\in\mathscr{F}(i=1,\cdots,n)$，则 $\bigcup\limits_{i=1}^{n}A_i\in\mathscr{F}$，$\bigcap\limits_{i=1}^{n}A_i\in\mathscr{F}$

② 若 $A_i\in\mathscr{F}(i=1,2,\cdots)$，则 $\bigcap\limits_{i=1}^{\infty}A_i\in\mathscr{F}$

③ 若 $A\in\mathscr{F},B\in\mathscr{F}$，则 $A\setminus B\in\mathscr{F}$.

这些结论是很容易证明的.因为 $\Omega\in\mathscr{F}$，故 $\varnothing=\Omega^c\in\mathscr{F}$，若 $A_i\in\mathscr{F}(i=1,2,\cdots,n)$，令 $A_i=\varnothing(i>n)$.则 $\bigcup\limits_{i=1}^{n}A_i=\bigcup\limits_{i=1}^{\infty}A_i\in\mathscr{F}$.又 $\left(\bigcap\limits_{i=1}^{n}A_i\right)^c=\bigcup\limits_{i=1}^{n}A_i^c$，故当 $A_i\in\mathscr{F}(i=1,\cdots,n)$ 时，$\left(\bigcap\limits_{i=1}^{n}A_i\right)^c\in\mathscr{F}$，从而 $\bigcap\limits_{i=1}^{n}A_i\in\mathscr{F}$. 类似地可证明上面的②成立.由于 $A\setminus B=A\cap B^c$ 故知③也成立.

可以证明，概率 $P(\cdot)$ 具有下列性质：

① $P(\varnothing)=0$

② 若 $A\in\mathscr{F}$，则 $P(A^c)=1-P(A)$

③ 若 $A_1,\cdots,A_n$ 都属于 $\mathscr{F}$ 且两两不相交，则

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)\tag{4.20}$$

④ 若 $A\subset B,A\in\mathscr{F},B\in\mathscr{F}$，则 $P(A)\leqslant P(B)$ 且

$$P(B\setminus A)=P(B)-P(A)\tag{4.21}$$

⑤ 若 $A_n\subset A_{n+1},A_n\in\mathscr{F}(n=1,2,\cdots)$，则

$$P\left(\bigcup_{n=1}^{\infty}A_n\right)=\lim_{n\to\infty}P(A_n)\tag{4.22}$$

⑥ 若 $A_n\supset A_{n+1},A_n\in\mathscr{F}(n=1,2,\cdots)$，则

$$P\left(\bigcap_{n=1}^{\infty}A_n\right)=\lim_{n\to\infty}P(A_n)\tag{4.23}$$

实际上，令 $A_1=\Omega,A_n=\varnothing(n=2,3,\cdots)$，从(4.19)和(4.18)知 $1=P(\Omega)=\sum\limits_{n=1}^{\infty}P(A_n)=1+\sum\limits_{n=2}^{\infty}P(\varnothing)$，故 $P(\varnothing)=0$. 若 $A\in\mathscr{F}$，令 $A_1=A$，$A_2=A^c,A_n=\varnothing(n\geqslant3)$，则 $\Omega=\bigcup\limits_{n=1}^{\infty}A_n$. 从(4.19)知 $1=\sum\limits_{n=1}^{\infty}P(A_n)=P(A)+P(A^c)+0$，从而 $P(A^c)=1-P(A)$；若 $A_1,A_2,\cdots,A_n$ 两两不交且

都属于 $\mathscr{F}$，令 $A_i=\varnothing(i>n)$，则从(4.19)直接推出(4.20)．由于 $B=A\cup(B-A)$，从(4.20)推知(4.21)成立．

当 $A_n\subset A_{n+1}(n\geqslant1)$时，令 $B_n=A_n-A_{n-1}(n\geqslant2)$，$B_1=A_1$，则 $B_i\in\mathscr{F}(i\geqslant1)$且它们两两不交，$\bigcup\limits_{n=1}^{\infty}A_n=\bigcup\limits_{n=1}^{\infty}B_n$，于是 $P\left(\bigcup\limits_{n=1}^{\infty}A_n\right)=P\left(\bigcup\limits_{n=1}^{\infty}B_n\right)=\sum\limits_{n=1}^{\infty}P(B_n)=\lim\limits_{n\to\infty}\sum\limits_{i=1}^{n}P(B_i)=\lim\limits_{n\to\infty}P\left(\bigcup\limits_{i=1}^{n}B_i\right)=\lim\limits_{n\to\infty}P(A_n)$，故(4.22)成立．当 $A_n\supset A_{n+1}(n\geqslant1)$时，则 $A_n^c\subset A_{n+1}^c$，利用关系式 $\left(\bigcap\limits_{n=1}^{\infty}A_n\right)^c=\bigcup\limits_{n=1}^{\infty}A_n^c$ 和(4.22)知 $P\left(\left(\bigcap\limits_{n=1}^{\infty}A_n\right)^c\right)=P\left(\bigcup\limits_{n=1}^{\infty}A_n^c\right)=\lim\limits_{n\to\infty}P(A_n^c)=1-\lim\limits_{n\to\infty}P(A_n)$，所以(4.23)成立．

要注意的是，当 $A\subset\Omega$ 且 $A\overline{\in}\mathscr{F}$时 $P(A)$没有定义，即对这样的 A 不能谈概率．人们自然想到，如果把 $\mathscr{F}$ 定义为 Ω 的所有子集组成的集合(它当然是一个 σ 代数)，则 Ω 的所有子集都有概率．为什么一般情况下不这样定义 $\mathscr{F}$ 呢？原因是，对任意的 Ω，如果用 $\mathscr{F}$ 表示 Ω 的所有子集组成的集合，则 $\mathscr{F}$ 上符合某些合理要求的概率测度可能不存在①．因而在一般情形下，$\mathscr{F}$ 是由一些子集组成的 σ 代数，而不一定是所有子集组成的 σ 代数．但应指出，当 Ω 是有限集或可数无穷集(即 Ω 的元素可排成一个序列)时，通常取 $\mathscr{F}$ 为 Ω 的所有子集组成．

---

① 在平面坐标系 xOy 里，设 $\Omega=\left\{(x,y):x^2+y^2=\left(\frac{1}{2\pi}\right)^2\right\}$(以原点为圆心，周长为 1 的圆周)，$\mathscr{F}$由 Ω 的一切子集组成．用 T_φ表示绕原点反时针方向旋转角为 φ 的旋转变换：即 $x=\gamma\cos\theta$，$y=\gamma\sin\theta\left(\gamma=\frac{1}{2\pi}\right)$时，$T_\varphi(x,y)=(x',y')$，其中 $x'=\gamma\cos(\theta+\varphi)$，$y'=\gamma\sin(\theta+\varphi)$．对任何 $A\subset\Omega$，令 $T_\varphi A=\{T_\varphi(x,y):(x,y)\in A\}$．可以证明，在 $\mathscr{F}$上不存在概率测度 $P(\cdot)$满足下列要求：

$$P(A)=P(T_\varphi A) \tag{注 4.1}$$

(对一切 $\varphi\in[0,2\pi)$及一切 $A\in\mathscr{F}$)．

注意，(注 4 1)乃是所谓"旋转不变性"

§5 条件概率、乘法公式、独立性

1. 条件概率

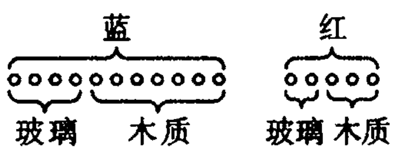
直到现在，我们对 $P(A)$ 的讨论都是相对于某组确定的条件 S 而言的. $P(A)$ 就是在条件组 S 实现之下，事件 A 发生的概率（为简略起见，"条件组 S"通常不再提及）. 除了这组基本条件"S"之外，有时我们还要提出附加的限制条件；也就是要求"在事件 B 已经发生的前提下"事件 A 发生的概率. 这就是条件概率的问题.

例 5.1 盒中装有 16 个球，其中 6 个是玻璃球，另外 10 个是木质球. 而玻璃球中有 2 个是红色的，4 个是蓝色的；木质球中有 3 个是红色的，7 个是蓝色的. 现从中任取一个.（这些就是所谓"条件组 S".）

记 $A=$"取到蓝球"，$B=$"取到玻璃球"

那么 $P(A)$，$P(B)$ 都是容易求得的. 但是如果已知取到的是蓝球，那么该球是玻璃球的概率是多少呢？也就是求在事件 A 已经发生的前提下事件 B 发生的概率（此概率记为 $P(B|A)$）. 将盒中球的分配情况列表并图示如下：

	玻璃	木质	
红	2	3	5
蓝	4	7	11
	6	10	16



由古典概型的(2.1)知

$$P(A)=11/16,P(B)=6/16$$

至于 $P(B|A)$，也可用古典概型来计算. 因取到的是蓝球，我们知道蓝球共有 11 个而其中有 4 个是玻璃球，所以

$$P(B|A)=4/11$$

定义 5.1 如果 A,B 是条件 S 下的两个随机事件，$P(A)\neq 0$[①]，则称在 A 发生的前提下 B 发生的概率为**条件概率**. 记作 $P(B|A)$.

注意 $P(B|A)$ 还是在一定条件下，事件 B 发生的概率：只是它的条件除原条件 S 外，又附加了一个条件（A 已发生）. 为区别这两者，后者就称为**条件概率**.

例 5.2 五个乒乓球（三个新、两个旧），每次取一个，无放回地取两次. 求：第一次取到新球的概率；第二次取到新球的概率；在第一次取到新球的条件下第二次取到新球的概率.

解 记

A = "第一次取到新球"

B = "第二次取到新球"

显然，

$$P(A)=3/5$$

但 $P(B)=?$（注意：B = "第二次取到新球"，它对第一次取到什么球没有限制或假定. 因此，回答 $P(B)=2/4$，或回答 $P(B)=3/4$ 都是没有根据的.）

这个问题可以用古典概型来解. 步骤是：先将五个球编号，然后数出两次抽取的全部可能结果的总数；再数出第二次取到新球所含的基本事件数（或者干脆将两次抽取的全部可能结果都排出来）. 其实，凭直观，$P(B)$ 应等于 3/5；否则"抽签"这个公认为公平的方法，就不公平了.

至于 $P(B|A)$，由条件概率的概念来看，反倒十分容易求得：

既然 A 已发生，那么第二次取时，盒中共有四个球，其中有两个新球. 因此按古典概型，此时，B 发生的概率应是 2/4. 这就是所要求的 $P(B|A)$.

---

① 本章仅讨论 $P(A)\neq 0$ 情况下的 $P(B|A)$.

2. 乘法公式

条件概率 $P(B|A)$跟事件的原概率有如下的一般关系(自然要求 $P(A)\neq 0$):

$$P(B|A)=\frac{P(AB)}{P(A)} \tag{5.1}$$

公式(5.1)是从大量社会实践中总结出来的普遍规律,不是用纯数学推导出来的,但在古典概型的情形可用数学方法给出证明.设条件组 S 下的一个等概完备事件组有 n 个基本事件,A 由其中的 m 个组成($m\geqslant 1$),B 由其中的 l 个组成,AB 由 k 个基本事件组成.按条件概率的概念,

$$P(B|A)=\frac{\text{在 } A \text{ 发生的前提下 } B \text{ 中含的基本事件数}}{\text{在 } A \text{ 发生的前提下基本事件总数}}$$

$$=\frac{k}{m}=\frac{k/n}{m/n}$$

但 $P(A)=\dfrac{m}{n}$,$P(AB)=\dfrac{k}{n}$,于是(5.1)成立.

(5.1)式揭露了事件的原概率 $P(A)$,$P(AB)$跟条件概率 $P(B|A)$这三个量之间的关系.通常,从两个方面来利用这个关系.一个方面是,已知 $P(A)$,$P(AB)$来求得 $P(B|A)$;另一个方面是,已知 $P(A)$与 $P(B|A)$来求得 $P(AB)$,在这种情况下,为了方便将(5.1)改写为

$$P(AB)=P(A)P(B|A) \tag{5.1$'$}$$

(5.1′)式称为概率的**乘法公式**.

3. 独立性

例 5.3 五个乒乓球(三个新,两个旧),每次取一个,**有放回**地取两次.

记

A=“第一次取到新球”

B=“第二次取到新球”

那么，显然有

$$P(B|A)=P(B)$$

即在 A 发生的条件下 B 的条件概率就等于 B 的原概率．它表示 A 发生并不影响 B 发生的概率．

由(5.1)式不难证明(请读者自行给出证明)，当 $P(A)\neq 0$ 时，

$$P(B|A)=P(B)\Longleftrightarrow P(AB)=P(A)P(B)$$

定义 5.2 称两个随机事件 A,B 是**相互独立的**，如果

$$P(AB)=P(A)P(B) \tag{5.2}$$

从以上对例 5.3 的讨论看出，所谓 A,B 相互独立，就是一个事件的发生并不影响另一事件发生的概率(细心的读者会发现，当 $P(B)\neq 0$ 时，(5.2)式也等价于 $P(A|B)=P(A)$)．在实际应用时，也正利用了这个事实．

例 5.4 甲、乙同时向一敌机炮击，已知甲击中敌机的概率为 0.6，乙击中敌机的概率为 0.5，求敌机被击中的概率．

解 记

$A=$“甲击中”

$B=$“乙击中”

$C=$“敌机被击中”

由加法公式，

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

显然可以认为甲击中(乙击中)并不影响乙击中(甲击中)的概率；亦即由对实际问题的分析，可以认为 A,B 相互独立．因此有

$$P(AB)=P(A)P(B)=0.6\times 0.5=0.3$$

于是，

$$P(C)=0.6+0.5-0.3=0.8$$

定理 5.1 若四对事件 A,B；$A,\overline{B}$；$\overline{A},B$；$\overline{A},\overline{B}$ 中有一对独立，则另外三对也独立(即这四对事件或者都独立，或者都不独立)．

证 这里仅证明“A,B 独立$\Longrightarrow A,\overline{B}$ 独立”,其余读者自行证明.

因为 $A=AU=A(B\cup\overline{B})=AB\cup A\overline{B}$,又知 $AB,A\overline{B}$互不相容.则有

$$P(A\overline{B})=P(A)-P(AB)$$

再因 A,B 独立,所以

$$\begin{aligned}P(A\overline{B})&=P(A)-P(A)P(B)\\&=P(A)[1-P(B)]\\&=P(A)P(\overline{B})\end{aligned}$$

这表明 $A,\overline{B}$ 独立.

利用这个结果,对例 5.4 还可以有另一解法.

解 我们先求 $P(\overline{C})$.注意到有$\overline{A\cup B}=\overline{A}\cdot\overline{B}$,且由 A,B 独立可知,$\overline{A},\overline{B}$ 也独立,所以

$$P(\overline{C})=P(\overline{A\cup B})=P(\overline{A}\,\overline{B})=P(\overline{A})P(\overline{B})$$

由(3.4)得

$$P(C)=1-P(\overline{C})=1-(1-0.6)(1-0.5)=0.8$$

这种解法的特点是通过对偶公式,把求事件并的概率问题转化为求事件交的概率.这在某些问题中是特别有效的(见例 5.5).

定义 5.3 称 A,B,C 是相互独立的,如果有

$$\begin{gathered}P(AB)=P(A)P(B)\\P(AC)=P(A)P(C)\\P(BC)=P(B)P(C)\\P(ABC)=P(A)P(B)P(C)\end{gathered}\tag{5.3}$$

类似于三个事件的相互独立性,对于 n 个事件$A_1,A_2,\cdots,A_n$ 的相互独立性有如下的定义:

定义 5.4 称 $A_1,A_2,\cdots,A_n$是相互独立的.如果对任何整数 $k(2\leqslant k\leqslant n)$有

$$P(A_{i_1}A_{i_2}\cdots A_{i_k})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k})\tag{5.4}$$

其中 $i_1, i_2, \cdots, i_k$ 是满足下面不等式的任何 k 个正整数：$1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$.

显然，当 $A_1, A_2, \cdots, A_n$ 相互独立时有

$$P(A_1 A_2 \cdots A_n) = P(A_1) P(A_2) \cdots P(A_n) \tag{5.5}$$

怎样判断一些事件是相互独立的呢？在很多情况下，根据对事件本质的分析就可以知道，并不需要复杂的计算.

例 5.5 设某型号的高射炮，每一门炮(发射一发)击中飞机的概率为 0.6. 现若干门炮同时发射(每炮射一发). 问欲以 99% 的把握击中来犯的一架敌机，至少需配置几门高射炮？

解 设 n 是以 99% 的概率击中敌机需配置的高射炮门数；并记

A_i = “第 i 门炮击中敌机”$(i = 1, 2, \cdots, n)$

A = “敌机被击中”

注意到

$$A = A_1 \cup A_2 \cup \cdots \cup A_n$$

于是要找 n，使

$$P(A) = P(A_1 \cup A_2 \cup \cdots \cup A_n) \geqslant 0.99 \tag{5.6}$$

由于 $\overline{A_1 \cup A_2 \cup \cdots \cup A_n} = \overline{A}_1 \cdot \overline{A}_2 \cdots \overline{A}_n$，且 $\overline{A}_1, \overline{A}_2, \cdots, \overline{A}_n$ 是相互独立的，所以

$$\begin{aligned} P(A) &= 1 - P(\overline{A}) = 1 - P(\overline{A}_1 \overline{A}_2 \cdots \overline{A}_n) \\ &= 1 - P(\overline{A}_1) P(\overline{A}_2) \cdots P(\overline{A}_n) = 1 - (0.4)^n \end{aligned}$$

因此，不等式(5.6)化为

$$1 - (0.4)^n \geqslant 0.99$$

即

$$(0.4)^n \leqslant 0.01$$

$$n \geqslant \frac{\lg 0.01}{\lg 0.4} = \frac{2}{0.397\,9} = 5.026$$

故至少需配置六门高射炮方能以 99% 以上的把握击中来犯的一

架敌机.

作为本节末尾,我们指出,一组事件两两独立并不能保证这组事件相互独立.请看下列有名的例子.

例 5.6 (Bernstein S N,1917).一质地均匀的正四面体,第一面染红色,第二面染黄色,第三面染蓝色,第四面染红、黄、蓝色(各占一部分).在桌上将此四面体任意掷一次,考察和桌面接触的那一面出现什么颜色.设 A =“红色出现”,B =“黄色出现”,C =“蓝色出现”.我们指出,这三个事件两两独立,但不相互独立.实际上,这里有四个基本事件:

$$A_i = \text{“第 } i \text{ 面接触桌面”} \quad (i=1,2,3,4).$$

既然是正四面体,当然 $P(A_i)=\frac{1}{4}(i=1,\cdots,4)$.显然,$A=A_1\cup A_4$,$B=A_2\cup A_4$,$C=A_3\cup A_4$,故 $P(A)=P(B)=P(C)=\frac{1}{2}$,因为 $AB=BC=CA=A_4$,故 $P(AB)=P(BC)=P(CA)=\frac{1}{4}$,可见三个事件两两独立.显然,$P(ABC)=\frac{1}{4}$,但是 $P(A)P(B)P(C)=\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}=\frac{1}{8}$,这表明,$A,B,C$ 不相互独立.

习 题 三

1. 一个工人看管三台机床,在一小时内机床不需要工人照管的概率:第一台等于0.9,第二台等于0.8,第三台等于0.7.求在一小时内三台机床中最多有一台需要工人照管的概率.(各机床是否需要照管是相互独立的)

2. 电路由电池 A 与两个并联的电池 B 及 C 串联而成.设电池 A,B,C 损坏的概率分别是 0.3,0.2,0.2.求电路发生断电的概率.(各电池是否损坏是互不影响的)

3. 某机械零件的加工由两道工序组成.第一道工序的废品率为 0.015,第二道工序的废品率为 0.02,假定两工序出废品是彼此无关的,求产品的合格率.

4. 在 1,2,…,100 中任取一数,问它既能被 2 整除又能被 5 整除的概率是多少? 又它能被 2 整除或能被 5 整除的概率是多少?

5. 加工某一零件共需经过四道工序. 设第一,二,三,四道工序的次品率分别是 2%,3%,5%,3%,假定各道工序是互不影响的,求加工出来的零件的次品率.

6. 当掷五枚分币时,已知至少出现两个正面,问正面数刚好是三个的条件概率是什么?

§6 全概公式与逆概公式

1. 全概公式

例 6.1 (继续讨论例 5.2)五个乒乓球(三个新、两个旧),每次取一个,无放回地取两次,求第二次取到新球的概率.

解 记

$A=$"第一次取到新球"

$B=$"第二次取到新球"

由于

$$B=BA\cup B\overline{A} \tag{6.1}$$

且 BA,$B\overline{A}$互不相容,则有

$$P(B)=P(BA)+P(B\overline{A})$$

再用乘法公式得

$$\begin{aligned}P(B)&=P(A)P(B\mid A)+P(\overline{A})P(B\mid\overline{A})\\&=\frac{3}{5}\cdot\frac{2}{4}+\frac{2}{5}\cdot\frac{3}{4}=\frac{3}{5}\end{aligned}$$

从形式上看,分解式(6.1)似乎将 B 复杂化了;但从实质上看,(6.1)将复杂的事件 B 分解为较简单的事件了. 把这个想法一般化,得

定理 6.1(全概公式)①　如果事件组 $A_1, A_2, \cdots, A_n$ 满足：

(1) $A_1, A_2, \cdots, A_n$ 互不相容而且 $P(A_i)>0(i=1,\cdots,n)$；

(2) $A_1 \cup A_2 \cup \cdots \cup A_n = U$(完备性)，

则对任一事件 B 皆有

$$P(B) = \sum_{i=1}^{n} P(A_i) P(B \mid A_i) \tag{6.2}$$

证　$B = BU = BA_1 \cup BA_2 \cup \cdots \cup BA_n$

注意到上式右边 n 个事件是互不相容的，于是有

$$\begin{aligned} P(B) &= P(BA_1) + P(BA_2) + \cdots + P(BA_n) \\ &= P(A_1)P(B|A_1) + P(A_2)P(B|A_2) \\ &\quad + \cdots + P(A_n)P(B|A_n) \end{aligned}$$

事实上，例 6.1 就是利用了 $n=2$ 时的全概公式计算出来的，其中 $A_1 = A, A_2 = \overline{A}$. 由于 A 和 $\overline{A}$ 满足定理的条件(1)和(2)，对于它们可以使用全概公式. 满足定理条件(1)和(2)的事件组 A_1, $A_2, \cdots, A_n$ 叫做**完备事件组**. 运用全概公式的关键往往在于找出一个完备事件组. 再看一个例子.

例 6.2　甲、乙、丙三人向同一飞机射击. 设甲、乙、丙射中的概率分别为 0.4,0.5,0.7. 又设若只有一人射中，飞机坠毁的概率为 0.2；若二人射中，飞机坠毁的概率为 0.6；若三人射中，飞机必坠毁. 求飞机坠毁的概率.

解　记　B = "飞机坠毁"

A_0 = "三人皆射不中"

A_1 = "只一人射中"

---

① 更一般的全概公式是：如果一列事件 $A_1, A_2, \cdots$，两两不相容，且 $\bigcup_{n=1}^{\infty} A_n = U$(必然事件)，$P(A_n)>0$(一切 $n \geqslant 1$)，则对任一事件 B 皆有

$$P(B) = \sum_{n=1}^{\infty} P(A_n) P(B|A_n).$$

$$A_2 = \text{“恰二人射中”}$$
$$A_3 = \text{“三人皆射中”}$$

显然，A_0, A_1, A_2, A_3是完备事件组. 而按加法与乘法公式有

$$P(A_0) = 0.6 \times 0.5 \times 0.3 = 0.09$$
$$\begin{aligned} P(A_1) &= 0.4 \times 0.5 \times 0.3 + 0.6 \times 0.5 \times 0.3 \\ &\quad + 0.6 \times 0.5 \times 0.7 \\ &= 0.36 \end{aligned}$$
$$\begin{aligned} P(A_2) &= 0.6 \times 0.5 \times 0.7 + 0.4 \times 0.5 \times 0.7 \\ &\quad + 0.4 \times 0.5 \times 0.3 \\ &= 0.41 \end{aligned}$$
$$P(A_3) = 0.4 \times 0.5 \times 0.7 = 0.14$$

再由题设有

$$P(B \mid A_0) = 0, \qquad P(B \mid A_1) = 0.2$$
$$P(B \mid A_2) = 0.6, \quad P(B \mid A_3) = 1$$

利用全概公式就得

$$\begin{aligned} P(B) &= \sum_{i=0}^{3} P(A_i) P(B \mid A_i) \\ &= 0.09 \times 0 + 0.36 \times 0.2 + 0.41 \times 0.6 + 0.14 \times 1 \\ &= 0.458 \end{aligned}$$

例 6.3(赌徒输光问题) 设甲有赌本 M 元，乙有赌本 N 元(M 和 N 都是正整数). 每一局若甲胜则乙给甲 1 元；若乙胜则甲给乙 1 元(没有和局). 设每局里甲胜的概率是 $p(0<p<1)$. 问：如果一局一局地赌博下去，甲输光的概率是多少？

解 记 $L = M + N$. L 是固定的正整数，$L \geqslant 2$. 当 $L = 2$ 时，显然甲输光的概率是 $1-p$，以下设 $L \geqslant 3$. 我们来研究更一般的问题：若甲有赌本 i 元，乙有赌本 $L-i$ 元，则甲输光的概率 p_i 是多少？(例 6.3 中要求的是 p_M).

问题扩大了，反而有利于寻找计算公式.

令 $A_i=\{$甲有赌本 i 元而最后输光$\}$$(i=1,2,\cdots,L-1)$，$B=\{$甲在第一局取胜$\}$. 当 $2\leqslant i\leqslant L-2$ 时，利用全概公式和条件概率的含义有下列递推公式：

$$\begin{aligned}p_i&=P(B)P(A_i|B)+P(B^c)P(A_i|B^c)\\&=pp_{i+1}+qp_{i-1}\end{aligned}\tag{6.3}$$

（这里 $q=1-p$，下同.）

容易知道 $p_1=pp_2+q$，$p_{L-1}=qp_{L-2}$. 可见，若令 $p_0=1$，$p_L=0$，则(6.3)式对一切 $1\leqslant i\leqslant L-1$ 均成立.

从(6.3)知 $p_{i+1}-p_i=\frac{q}{p}(p_i-p_{i-1})$，于是

$$p_{i+1}-p_i=\left(\frac{q}{p}\right)^i(p_1-p_0)=\left(\frac{q}{p}\right)^i(p_1-1)，所以$$

$$\begin{aligned}p_{i+1}-p_1&=\sum_{k=1}^{i}(p_{k+1}-p_k)\\&=\sum_{k=1}^{i}\left(\frac{q}{p}\right)^k(p_1-1)\end{aligned}\tag{6.4}$$

由于 $p_L=0$，在(6.4)中令 $i=L-1$ 得

$$p_1=\sum_{k=1}^{L-1}\left(\frac{q}{p}\right)^k(1-p_1)\tag{6.5}$$

当 $p\neq\frac{1}{2}$时，$p_1=\dfrac{\frac{q}{p}-\left(\frac{q}{p}\right)^L}{1-\left(\frac{q}{p}\right)^L}$

从(6.4)知

$$\begin{aligned}p_i&=p_1+\frac{\frac{q}{p}-\left(\frac{q}{p}\right)^i}{1-\frac{q}{p}}(p_1-1)\\&=\frac{\left(\frac{q}{p}\right)^i-\left(\frac{q}{p}\right)^L}{1-\left(\frac{q}{p}\right)^L}\quad(2\leqslant i\leqslant L-1).\end{aligned}$$

当 $p=\frac{1}{2}$ 时,从(6.5)知 $p_1=\frac{L-1}{L}$,利用(6.4)知

$$\begin{aligned} p_i &= p_1+\sum_{k=1}^{i-1}(p_1-1) \\ &= \frac{L-i}{L} \quad (2\leqslant i\leqslant L-1) \end{aligned}$$

于是

$$p_M=\begin{cases} \dfrac{\left(\dfrac{q}{p}\right)^M-\left(\dfrac{q}{p}\right)^{M+N}}{1-\left(\dfrac{q}{p}\right)^{M+N}} & \left(p\neq\dfrac{1}{2}\right) \\ \dfrac{N}{M+N} & \left(p=\dfrac{1}{2}\right). \end{cases}$$

这就是甲输光的概率.这个计算过程中关键的一步是,根据第一局的输赢结果建立方程(6.3).这种方法可以叫做“首步(局)分析法”,这种方法在许多问题中都得到应用.

例 6.4(敏感性问题的调查)　在竞技性体育运动项目中,为保证竞争的公平性,禁止运动员服用兴奋剂类药品.但服用违禁药品的行为属于隐私行为.因此,要调查运动员服用违禁药品的情况,是一件难事.[①] 这里的关键是要设计一个调查方案,使被调查者愿意作出真实回答,又能保守个人秘密.经过多年研究与实践,西方的一些心理学家和统计学家设计了一个调查方案,这个方案的核心是如下两个问题:

问题 A:你的生日是否在 7 月 1 日之前(不含 7 月 1 日)?

问题 B:你是否在比赛前服用过违禁药品?

被调查者只需回答其中一个问题,至于回答哪一个问题由被调查者事先从一个罐中随机抽取一个球的颜色而定.只抽一个球,

---

① 这是指对一个国家或地区的运动员进行宏观调查而言.对于一些特殊的运动员(例如比赛中成绩居前几位的)可以强制性的进行医学检查以查验其是否服用违禁药品。

看过颜色之后再放回，若抽出白球则回答问题 A；若抽出红球则回答问题 B. 罐中只有白球与红球，且红球的比率 π 是已知的，即 P(抽到红球) $= \pi$，P(抽到白球) $= 1 - \pi$. 被调查者无论回答问题 A 或问题 B，只需在下面答卷上认可的方框内打勾，然后把答卷放入投票箱内.

上述抽球与答卷都在一间无人的房间内进行，任何外人都不知道被调查者抽到什么颜色的球和在什么地方打勾. 如果向被调查者讲清这个方案的做法并严格执行，那么就很容易使被调查者确信他(她)参加这次调查不会泄露个人秘密，从而愿意参与调查.

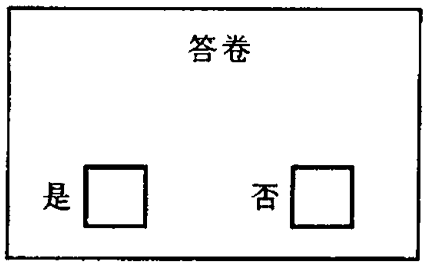


图 1.3

当有较多的人(例如 200 多人)参加调查后，就可打开投票箱进行统计. 设有 n 张答卷，其中 k 张答“是”，于是回答“是”的比率 φ 是 k/n，这可作为回答“是”的概率. 这里答“是”有两种情况：一是抽到白球后对问题 A 答“是”，这是一个条件概率，它等于生日在 7 月 1 日以前的概率，一般认为是 0.5，即 P(回答“是”| 抽到白球) $= 0.5$；另一种是抽到红球后对问题 B 回答“是”，这也是一个条件概率，它不是别的，正是赛前服用违禁药品的运动员在全体被调查的运动员中的比率 p，即可认为

$$P(\text{回答“是”}\mid\text{抽到红球}) = p$$

利用全概公式

P(回答“是”) $= P$(抽到白球)P(回答“是”| 抽到白球) $+$ P(抽到红球)P(回答“是”| 抽到红球)知有下列近似式：

$$\frac{k}{n} = 0.5(1 - \pi) + p \cdot \pi$$

由此知

$$p = \left(\frac{k}{n} - \frac{1 - \pi}{2}\right) \Big/ \pi$$

假如在此项调查中，罐里有 50 个球，其中红球 30 个，即 $\pi = 0.6$. 某国家的体育部门在五天内安排 15 个项目的运动员共 246 名参加调查，最后开箱统计，答卷全部有效(即没有一张卷上打两个勾者)，其中回答"是"的有 54 张，据此可算得

$$p = \left(\frac{54}{246} - \frac{0.4}{2}\right)\Big/ 0.6 = 0.032\ 5$$

这表明该国约有 3.25%的运动员赛前服用过违禁药品.

2. 逆概公式

例 6.5 发报台分别以概率 0.6 和 0.4 发出信号"·"和"－". 由于通信系统受到干扰，当发出信号为"·"时，收报台未必收到信号"·"，而是分别以概率 0.8 和 0.2 收到信号"·"和"－". 又若，当发出信号为"－"时，收报台分别以概率 0.9 和 0.1 收到信号"－"和"·"(见图 1.4). 求当收报台收到信号"·"时，发报台确实发出信号"·"的概率.

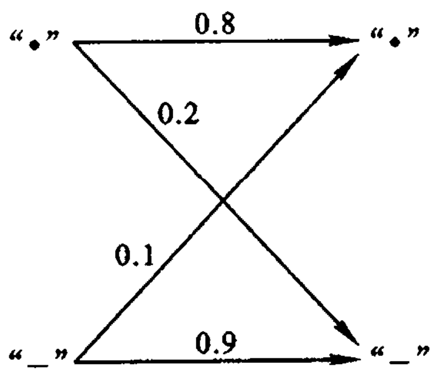


图 1.4

解 设 A ="发出信号'·'"

B ="收到信号'·'"

则所求的是 $P(A|B)$. 由条件概率公式

$$P(A|B) = \frac{P(AB)}{P(B)}$$

再由乘法公式与全概公式，有

$$P(AB) = P(A)P(B|A)$$

$$P(B) = P(A)P(B|A) + P(\overline{A})P(B|\overline{A})$$

而已知 $P(A) = 0.6, P(B|A) = 0.8, P(B|\overline{A}) = 0.1$. 故得

$$P(A|B) = \frac{0.6 \times 0.8}{0.6 \times 0.8 + 0.4 \times 0.1} = \frac{0.48}{0.52} = 0.923$$

定理 6.2(逆概公式)① 设 $A_1, A_2, \cdots, A_n$ 为一完备事件组，则对任一事件 B(自然要求 $P(B)\neq 0$)有

$$P(A_j \mid B) = \frac{P(A_j)P(B \mid A_j)}{\sum_{i=1}^{n} P(A_i)P(B \mid A_i)} \quad (j=1,\cdots,n) \qquad (6.6)$$

证 $$P(A_j \mid B) = \frac{P(A_j B)}{P(B)} = \frac{P(A_j)P(B \mid A_j)}{\sum_{i=1}^{n} P(A_i)P(B \mid A_i)}$$

逆概公式也称为贝叶斯(Bayes)公式，它在理论上与应用上都十分重要.

例 6.6[8](艾滋病检测)

艾滋病(AIDS)是一种可怕的接触性传染病. 为了防止其传播，我们要识别艾滋病病毒的携带者. 目前有一种血液试验检测法用于检测身体中是否有艾滋病病毒. 尽管这种检测法相当精确，但也可能带来两种误诊. 首先，它可能会对某些真有艾滋病的人作出没有艾滋病的诊断，这就是所谓假阴性；其次，它也可能对某些没有艾滋病的人作出有艾滋病的诊断，这就是所谓假阳性.

根据现有的统计资料，我们可以认为上述血液试验检测法的灵敏度如下：假阴性的概率是 0.05(即真有病的人的试验结果呈阴性的概率)，假阳性的概率是 0.01(即没有病的人的试验结果呈阳性的概率).

美国是艾滋病较为流行的国家之一，据保守估计大约 1 000 人中就有一人得这种病. 为了能有效地控制和减缓这种病的传播速度. 几年前，有人就提议应在申请结婚登记的新婚夫妇中进行有

---

① 更一般的逆概公式如下：设 $A_1, A_2, \cdots$ 是一列事件，两两不相容，$P(A_j)>0$ $(j\geqslant 1)$，而且 $\bigcup_{j=1}^{\infty} A_j$ 是必然事件，则对任何事件 B，只要 $P(B)>0$，则一定有

$$P(A_j \mid B) = \frac{P(A_j)P(B \mid A_j)}{\sum_{i=1}^{\infty} P(A_i)P(B \mid A_i)} \quad (j=1,2,\cdots)$$

无艾滋病病毒的血液试验.该项普查计划提出后,立刻就遭到许多专家学者的反对,他们认为这是一项既费钱又费力同时收效不大的计划,最终此项计划未被通过.那么,到底专家学者的意见对不对? 该普查计划该不该执行呢?

假如该计划得以实施,而你又做了血液试验,结果呈阳性,那么你真正得了艾滋病的可能性有多大呢? 我们定义事件:

$$A = \{\text{被检测人带有艾滋病病毒}\}$$

$$T = \{\text{试验结果呈阳性}\}.$$

我们关心的是条件概率 $P(A|T)$.由逆概公式(6.6)知

$$P(A|T) = \frac{P(T|A)P(A)}{P(A)P(T|A) + P(A^c)P(T|A^c)} \tag{6.7}$$

注意 $P(A)=0.001, P(T|A)=1-0.05, P(A^c)=0.999, P(T|A^c)=0.01$.于是 $P(A|T)\approx 0.087$.这个条件概率是相当小的.因此,你即使检测结果呈阳性,也不必太紧张.可见在这种血液检测法下,费钱费力但结果很不可靠.该普查计划缺乏执行的理由,究其原因,从 $P(A|T)$的表达式看出就是因为 $P(A)$太小.从(6.7)知

$$P(A|T) = \frac{0.95P(A)}{0.95P(A)+(1-P(A))0.01}$$

$$= \frac{0.95}{0.95-0.01+0.01\dfrac{1}{P(A)}}$$

可见 $P(A|T)$是 $P(A)$的严格增函数.而且 $P(A)$接近 1 时 $P(A|T)$也接近 1.对处于感染艾滋病的"高危"人群,用上述血液检测法进行普查倒是很有效的.

习　题　四

1. 两台机床加工同样的零件,第一台出现废品的概率是 0.03,第二台出现废品的概率是 0.02.加工出来的零件放在一起,并且已知第一台加工的零

件比第二台加工的零件多一倍，求任意取出的零件是合格品的概率；又：如果任意取出的零件经检查是废品，求它是由第二台机床加工的概率.

2. 盒中放有 12 个乒乓球，其中有 9 个是新的.第一次比赛时从中任取 3 个来用，比赛后仍放回盒中.第二次比赛时再从盒中任取 3 个，求第二次取出的球都是新球的概率.又：已知第二次取出的球都是新球，求第一次取到都是新球的概率.

3. 有三只盒子，在甲盒中装有 2 枝红芯圆珠笔、4 枝蓝芯圆珠笔，乙盒中装有 4 枝红的、2 枝蓝的，丙盒中装有 3 枝红的、3 枝蓝的.今从其中任取一枝.设到三只盒子中取物的机会相同，它是红芯圆珠笔的概率为多少？又若已知取得的是红的，它是从甲盒中取出的概率为多少？

§7 独立试验序列概型

本节再介绍一种概率模型.在这个模型中，基本事件的概率可以直接计算出来；但它与古典概型不同，这些基本事件不一定是等概的.

先看一个最简单的例子，它可用古典概型的方法来解决.

例 7.1 掷一枚匀称的分币，独立重复地掷五次，求其中恰有两次正面朝上的概率.

解 用古典概型很容易求出.

掷一枚分币，每次可出“上”，“下”(“上”，“下”分别表示正面朝上与正面朝下)两种可能性，独立掷五次，共有 $n=2^5=32$ 个基本事件(它们是上，上，上，上，上；上，上，上，上，下；…；下，下，下，下，下).而“恰有两次正面朝上”共占 $m=C_5^2=10$ 个基本事件，故所求概率 $p=\frac{m}{n}=\frac{10}{32}\left(=\frac{C_5^2}{2^5}\right)$.

现将上面计算的概率 p 改写为如下形式：

$$p=C_5^2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^3 \tag{7.1}$$

如何解释(7.1)式呢？$C_5^2=10$ 是“恰有两次正面朝上”的基本事件

数，而这 10 个事件中每一个的概率皆为$\left(\frac{1}{2}\right)^2\cdot\left(\frac{1}{2}\right)^3$（前两个$\frac{1}{2}$是发生两次"上"的概率，后三个$\frac{1}{2}$是发生三次"下"的概率. 由于是五次独立试验，所以乘积的概率等于概率的乘积，为$\left(\frac{1}{2}\right)^5$），显然这 10 个基本事件是互不相容的，由加法公式(3.5)就可推得(7.1)式.

如果我们掷的分币不是匀称的，出现"上"的概率为$\frac{2}{3}$，出现"下"的概率为 1/3，则

$$P(\text{“恰有两次正面朝上”}) = C_5^2\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^3$$

再看一个类似的例子.

例 7.2 设某人打靶，命中率为 0.7. 现独立地重复射击 5 次，求恰好命中两次的概率.

解 和上面同样的分析可以算出：

$$P(\text{“恰命中 2 次”}) = C_5^2(0.7)^2(0.3)^3$$

类似地有

$$P(\text{“恰命中 3 次”}) = C_5^3(0.7)^3(0.3)^2$$

$$P(\text{“恰命中 4 次”}) = C_5^4(0.7)^4(0.3)^1$$

$$P(\text{“恰命中 1 次”}) = C_5^1(0.7)^1(0.3)^4$$

也有

$$P(\text{“恰命中 0 次”}) = C_5^0(0.7)^0(0.3)^5$$

本例用古典概型是不易解决的.

像例 7.2 这样的问题是广泛存在的，其一般提法是：设单次试验中，某事件 A 发生的概率为 $p(0<p<1)$，现将此试验重复进行 n 次，求 A 发生 k 次的概率($k=0,1,2,\cdots,n$).

上两例中的 A,p,n,k 如下：

	例 7.1	例 7.2
A	正面朝上	命中
p	0.5	0.7
n	5	5
k	2	2

从例子已看出问题的答案,为明确起见,特别写成定理并给予一般的证明.

定理(独立试验序列概型计算公式)

设单次试验中,事件 A 发生的概率为 $p(0<p<1)$,则在 n 次重复试验中,

$$P(\text{“}A\text{ 发生 }k\text{ 次”}) = C_n^k p^k q^{n-k} \quad (q=1-p)$$

$$(k=0,1,2,\cdots,n)$$

证 在 n 次重复试验中,记 $B_1,B_2,\cdots,B_m$ 为构成事件“A 发生 k 次”的那些试验结果,于是有:

(1) “A 发生 k 次” $= B_1 \cup B_2 \cup \cdots \cup B_m$;$B_1,\cdots,B_m$ 互不相容;

(2) $P(B_1)=P(B_2)=\cdots=P(B_m)=p^k q^{n-k}$;

(3) $m=C_n^k$.

因此得

$$\begin{aligned} P(\text{“}A\text{ 发生 }k\text{ 次”}) &= P(B_1 \cup B_2 \cup \cdots \cup B_m) \\ &= P(B_1)+P(B_2)+\cdots+P(B_m) \\ &= mp^k q^{n-k} = C_n^k p^k q^{n-k} \end{aligned}$$

注意,“n 次重复试验”中的“重复”二字,是指这 n 次试验中各次试验的条件组是相同的.因此,这不仅意味着在各次试验中 A 发生的概率都是 p(于是 $\overline{A}$ 发生的概率也都是 q),而且还有各次试验的结果间是互相独立的含义.定理证明过程中的(2),就是基于这两个含义而得出的.在具体应用时也要注意这“重复”二字,

请看下列问题：

已知 80 个产品中有 5 个次品，现从中每次任取一个，无放回地取 20 次，求在所取 20 个中恰有 2 个次品的概率.

表面上看该问题的提法跟独立试验序列概型是相同的，其实不然. 因为它是“无放回”抽取，因此它各次试验的条件是有差别的，也就不能用上面的概型求解(怎么求？留给读者自行解决). 如果将问题中的“无放回”改为“有放回”，那么这 20 次试验的条件就完全相同，于是可用上面的概型求解(作为练习，留给读者).

在客观世界中，真正的完全重复的现象是不多见的. 比如对于例 7.2，一般来讲，这 5 次射击的条件不可能完全一样，只是近似于完全重复，可用独立试验序列概型来近似处理而已. 还拿抽样问题来讲，当原产品的批量相当大时，“无放回”就可以近似地当作“有放回”来处理，因此，也就可用独立试验序列概型来计算取到的产品中含 k 个次品的概率.

例 7.3 设每次射击打中目标的概率等于 0.001，如果射击 5 000 次，试求至少两次打中目标的概率.

容易看出，

$$P\{\text{至少两次打中目标}\} = \sum_{k=2}^{5\,000} P\{\text{恰有 } k \text{ 次打中目标}\}$$

$$= 1 - P\{\text{每次都未打中目标}\} - P\{\text{恰有 1 次打中目标}\}$$

$$= 1 - (1-0.001)^{5\,000} - C_{5\,000}^{1} \times (0.001)^{1} \times (1-0.001)^{5\,000-1}$$

$$= 1 - (0.999)^{5\,000} - 5\,000 \times 0.001 \times (0.999)^{4\,999}$$

$$\approx 1 - 0.006\,7 - 0.033\,5 = 0.959\,8$$

我们看到，计算还是很麻烦的. 当 p 很小 n 很大时，可用下列第一近似公式，避免这种麻烦.

第一近似公式：

$$P\{A \text{ 发生 } k \text{ 次}\} \approx \frac{(np)^k}{k!} e^{-np} \tag{7.2}$$

用在例 7.3 上去，$P\{\text{每次都未打中}\} \approx e^{-5\,000 \times 0.001} = e^{-5}$；

$$P\{恰有 1 次打中\} = \frac{5\,000 \times 0.001}{1!} e^{-5\,000 \times 0.001}$$
$$= 5e^{-5}$$

于是 $P\{至少两次打中目标\} \approx 1 - 6e^{-5} \approx 0.959\,6$.

第一近似公式的理论根据是下列事实：如果 $\lim\limits_{n \to \infty} np_n = \lambda > 0$，则

$$\lim_{n \to \infty} C_n^k p_n^k (1 - p_n)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}$$

这个关系式的证明不难，可参阅第二章 §2.

当 p 不是很小，而 n 很大时，可用第二近似公式：

$$P\{A 发生 k 次\} \approx \frac{1}{\sqrt{np(1-p)}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_k^2}$$

这里 $x_k = \dfrac{k - np}{\sqrt{np(1-p)}}$.

第二近似公式也是有理论根据的，数学证明较长，从略.

例 7.4 设每次射击打中目标的概率等于 $\frac{1}{6}$. 如果射击 6 000 次，问：射中次数在 900 至 1 100 之间的概率等于多少？

这个问题理论上好回答，所求概率 $= \sum\limits_{k=900}^{1\,100} C_{6\,000}^k \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{6\,000-k}$ 项数太多了，怎么算？以后我们知道，利用著名的中心极限定理(见第三章)，可容易算得这概率的近似值是 0.999 46.

在应用中，独立试验序列概型有时以所谓“随机游动”的方式出现.

例 7.5 (自由随机游动) 假设一质点在某直线(数轴)上运动，在时刻 0 从原点出发. 每隔一个单位时间位置向右或向左移动一个单位，而向右移动的概率总是 p，向左移动的概率总是 $q(p + q = 1)$. 问：在时刻 n 质点位于 K 的概率是多少(n 是正整数，K 是整数)？

解 不失一般性，我们只考虑 K 为正整数的情形($K \leqslant 0$ 时

可类似地讨论). 为了质点在时刻 n 位于 K,必须且只需在头 n 次游动时向右移动的次数比向左移动的次数多 K 次. 若以 x 表示向右移动的次数, y 表示向左移动的次数,则

$$x+y=n, \quad x-y=K$$

于是 $x=\dfrac{n+K}{2}$,因 x 是整数,故 K 与 n 具有相同的奇偶性. 可见, n, K 的奇偶性相反时,所求的概率为 0. 当 n, K 的奇偶性相同时, P(质点在时刻 n 位于 K) $=P$(质点在头 n 次游动时有 $\dfrac{n+K}{2}$次向右,有$\dfrac{n-K}{2}$次向左) $=C_n^{\frac{n+K}{2}} p^{\frac{n+K}{2}} q^{\frac{n-K}{2}}$

习 题 五

1. 设某种型号的电阻的次品率为 0.01,现在从产品中抽取 4 个,分别求出没有次品、有 1 个次品、有 2 个次品、有 3 个次品、全是次品的概率.

2. 某类电灯泡使用时数在 1 000 h 以上的概率为 0.2,求三个灯泡在使用 1 000 h 以后最多只有一个坏的概率.

3. 有 6 个元件,它们断电的概率第一个为 $p_1=0.6$,第二个为 $p_2=0.2$,其余四个都为 $p_3=0.3$,求线路断电的概率,若

(1) 所有的元件串联;

(2) 元件按图 1.5 连接.

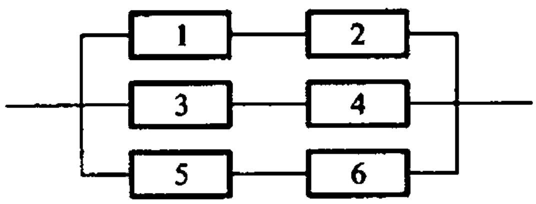


图 1.5

4. 设昆虫生产 k 个卵的概率 $p_k=\dfrac{\lambda^k}{k!}e^{-\lambda}(k=0,1,2,\cdots)$,又设一个虫卵能孵化为昆虫的概率等于 p. 若卵的孵化是互相独立的,问此昆虫的下一代有 l 条的概率是多少?

第二章　随机变量与概率分布

§1　随 机 变 量

上一章中，我们讨论了随机事件及其概率.为了进一步研究随机现象，我们将引进随机变量的概念.

例 1.1　设有产品 100 件，其中有 5 件次品、95 件正品.现从中随便抽取 20 件，问“抽得的次品件数”（以下简称“次品数”）是多少?

我们说，“次品数”可能是 1，也可能是 2，3，4，5，甚至可能是 0（即抽得的 20 件中无次品）.它随着不同的抽样批数而可能不同，就是说，“次品数”的值无法在抽样前给出确定性的答案.然而，“次品数”的不确定性只是问题的一个方面；另一方面，作为任何一批抽样的具体结果，即在 100 件产品中随机抽取了 20 件，在该 20 件产品中，“次品数”又是完全确定的.因而，“次品数”是个变量，它是随着抽样结果而变的变量，称为随机变量.下面给出定义.

定义 1.1　对于条件组 S 下的每一个可能结果 ω 都惟一地对应到一个实数值 $X(\omega)$，则称实值变量 $X(\omega)$ 为一个随机变量.简记为 X①.

例 1.2　设盒中 5 个球，其中 2 个白球、3 个黑球，从中随便抽取 3 个球.则“抽得的白球数”X 是一个随机变量.抽取结果 ω 跟“白球数”$X(\omega)$ 的对应关系见下表（①②③为黑球，④⑤为白球）：

① 常用大写拉丁字母 X, Y, Z 等（或希腊字母 ξ, η, ζ 等）表示随机变量.

ω			白球数 $X(\omega)$
①	②	③	0
①	②	④	1
①	②	⑤	1
①	③	④	1
①	③	⑤	1
①	④	⑤	2
②	③	④	1
②	③	⑤	1
②	④	⑤	2
③	④	⑤	2

我们看到，X 只可能取 0，1，2 这三个实值，而 $\{X=0\}$，$\{X=1\}$，$\{X=2\}$ 都是随机事件. 不仅如此，我们运用上一章的知识，还可求得：

$$P\{X=0\}=\frac{C_3^3}{C_5^3}=\frac{1}{10}$$

$$P\{X=1\}=\frac{C_3^2C_2^1}{C_5^3}=\frac{6}{10}$$

$$P\{X=2\}=\frac{C_3^1C_2^2}{C_5^3}=\frac{3}{10}$$

例 1.3 设某射手每次射击打中目标的概率是 0.8，现在连续射击 30 次. 则"击中目标的次数"X 是一个随机变量. 它只可能取到 0，1，2，…，30 这 31 个实值. 显然 $\{X=0\}$，$\{X=1\}$，$\{X=2\}$，…，$\{X=30\}$ 都是随机事件.（读者不难算出概率 $P\{X=k\}$ 是多少.）

例 1.4 某射手每次射击打中目标的概率是 0.8，现在连续向一个目标射击，直到第一次击中目标时为止. 则"射击次数"X 是一个随机变量. X 可能取到一切自然数.（请想一想，为什么？）$\{X=k\}$（$k=1,2,3,\cdots$）都是随机事件.

例 1.5 某出租汽车公司共有出租车 400 辆. 设每天每辆出租车出现故障的概率为 0.02，则一天内出现故障的出租车的总数

X 便是一个随机变量，X 可能取的值是 $0,1,2,\cdots,400$.

例 1.6 某公共汽车站每隔 5 分钟有一辆汽车通过；一位乘客对于汽车通过该站的时间完全不知道，他在任一时刻到达车站都是等可能的. 那么，他的候车时间 X 是一个随机变量，显然 $0\leqslant X<5$，又 $\{X>2\}$，$\{X\leqslant 3\}$ 都是随机事件.

和前面五个例子不同，例 1.6 的随机变量 X 所取的值不一定是整数，而且不能一一列举出来，它的取值是"连续的". 下面的例 1.7也是这样.

例 1.7 一门大炮在一定的条件下向某个地面目标瞄准射击，用 ρ 表示弹着点与目标之间的距离(图 2.1). 则 ρ 就是一个随机变量. 又若在地平面上取直角坐标系：原点在目标处，将大炮所在地点与目标地点连线的方向为 y 轴方向，与之垂直的方向为 x 轴方向(见图 2.1). 则弹着点 $M(X,Y)$ 的两个分量 X,Y 也都是随机变量.

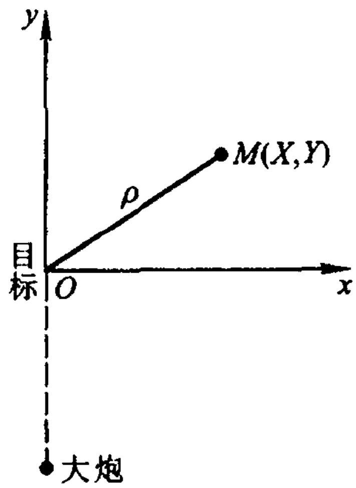


图 2.1

本例中随机变量 ρ,X,Y 的取值都是"连续的"，X,Y 还可能取到负值.

随机变量的概念在概率论和数理统计中既基本又重要. 在实际问题中广泛存在着随机变量. 比如在工业生产中，随便取一件产品，问它的质量指标(强度、硬度、光洁度、粘合力、纤度，……)是多少，这个质量指标就可以看作一个随机变量. 我们要学会把随机变量的概念与实际工作中的具体问题联系起来.

对于随机变量，通常分两类进行讨论. 如果 X 所可能取的值能够一一列举出来，则称 X 为离散型随机变量. 例1.1～1.5 中的随机变量都是离散型的. 如果 X 所可能取的值不能一一列举，则称 X 是非离散型的. 非离散型的随机变量范围很广，而其中最重要的也是实际工作经常遇到的是所谓连续型的随机变量(见 §3).

下面我们先讨论离散型的随机变量.

§2 离散型随机变量

1. 概率分布

离散型随机变量 X 只可能取有限个或者一串值. 设 X 可能取的值是 $x_1, x_2, \cdots, x_k, \cdots$.

为了完全描述随机变量 X,只知道它可能取的值是远远不够的,更重要的是要知道它取各个值的概率. 也就是说,要知道下列一串概率的值:

$$P\{X=x_1\}, P\{X=x_2\}, \cdots, P\{X=x_k\}, \cdots$$

记 $p_k = P\{X=x_k\}(k=1,2,\cdots)$,将 X 可能取的值及相应的概率列成下表:

X	x_1	x_2	x_3	$\cdots$	x_k	$\cdots$
p	p_1	p_2	p_3	$\cdots$	p_k	$\cdots$

这个表称为 X 的概率分布表. 它清楚而完整地表示了 X 取值的概率的分布情况. 为简单计,概率的分布情况也可直接用一系列等式

$$p_k = P\{X=x_k\} \quad (k=1,2,\cdots) \tag{2.1}$$

来表示. (2.1)称为 X 的**概率分布**.

关于 $p_k(k=1,2,\cdots)$,显然有:

(1) $p_k \geqslant 0 \quad (k=1,2,\cdots)$;

(2) $\sum\limits_k p_k = 1$.

(请读者想一想,为什么? 请参看第一章的公式(3.8).)

作为概率分布的一个例子,我们回头看例 1.2 中的 X("抽得的白球数"). 它的概率分布表如下:

X	0	1	2
p	0.1	0.6	0.3

它的概率分布,由下列一组等式表示:

$$P\{X=0\}=0.1$$
$$P\{X=1\}=0.6$$
$$P\{X=2\}=0.3$$

下面我们介绍几类常见的概率分布(简称"分布").

2. 二点分布

如果随机变量 X 的分布如下:

$$P\{X=1\}=p \quad (0<p<1)$$
$$P\{X=0\}=q=1-p \tag{2.2}$$

则称 X 服从二点分布(p 为参数).二点分布也叫伯努利分布.

例 2.1 100 件产品中,有 95 件正品、5 件次品,现从中随机抽取一件,假如抽得每件的机会都相同,那么:

抽得正品的概率 $=0.95$,抽得次品的概率 $=0.05$.

现在定义随机变量 X 如下:

$$X=\begin{cases}1 & \text{当取得正品} \\ 0 & \text{当取得次品}\end{cases}$$

则有

$$P(X=1)=0.95$$
$$P(X=0)=0.05$$

即 X 服从二点分布.

二点分布虽很简单,但有用.当一组条件下只有两个可能结果,且都有正概率时,能确定一个服从二点分布的随机变量.

3. 二项分布

如果随机变量 X 的分布如下:

$$P\{X=k\}=C_n^k p^k q^{n-k} \quad (k=0,1,2,\cdots,n) \tag{2.3}$$
$$(0<p<1, q=1-p)$$

则称 X 服从二项分布(参数为 n, p),或用记号

$$X \sim B(n, p)$$

来表示.

利用二项式定理,不难证明按(2.3)式给出的概率值满足:

$$\sum_{k=0}^{n} p_k = \sum_{k=0}^{n} \mathrm{C}_n^k p^k q^{n-k} = (p+q)^n = 1^n = 1$$

下面谈谈二项分布的实际背景.第一章的§7 讨论了独立试验序列,其中有这样一条定理:

"设单次试验中,事件 A 发生的概率为 $p(0<p<1)$,则在 n 次独立试验中,

$$P(\text{“}A\text{ 发生 }k\text{ 次”}) = \mathrm{C}_n^k p^k q^{n-k} \qquad (q = 1-p)$$

$$(k = 0, 1, 2, \cdots, n)\text{”}$$

由此可见,在 n 次独立试验中,"A 发生的次数"X 这个随机变量服从二项分布(所谓 A 发生 k 次就是"$X=k$").顺便提一句,由于

$$\sum_{k=0}^{n} \text{“}A\text{ 发生 }k\text{ 次”} = U$$

则有

$$\sum_{k=0}^{n} \mathrm{C}_n^k p^k q^{n-k} = \sum_{k=0}^{n} P(\text{“}A\text{ 发生 }k\text{ 次”})$$

$$= P\left(\sum_{k=0}^{n} \text{“}A\text{ 发生 }k\text{ 次”}\right) = P(U) = 1$$

这是从另一个途径来说明 $\sum\limits_{k=0}^{n} p_k = 1$.

最后,读者不难发现,$n=1$ 时的二项分布就是二点分布.

4. 泊松(Poisson)分布

如果随机变量 X 的概率分布如下:

$$P\{X=k\} = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} \quad (k=0,1,2,\cdots;\lambda>0) \tag{2.4}$$

则称 X 服从泊松分布(参数为 λ).

服从泊松分布的随机变量是很不少的，下面举一个例子.

例 2.2 放射性物质在某一段时间内放射的粒子数 X 是服从泊松分布的. Rutherford 和 Geiger 观察了放射性物质放出的 α 粒子个数的情况，一共做了 2 608 次观察，每次观察时间是 7.5 s，总共观察到 10 094 个 α 粒子，列表如下面所示. 从表中，我们看到按(2.4)算出的 $P\{X=k\}$ 跟 $\{X=k\}$ 的频率相当接近.

放射粒子数 X	观察到次数 μ_k	频　率 $p_k^* = \frac{\mu_k}{N}$	按泊松分布 $\left(\lambda = \frac{10\ 094}{2\ 608} = 3.87\right)$ 计算之概率 p_k
0	57	0.022	0.021
1	203	0.078	0.081
2	383	0.147	0.156
3	525	0.201	0.201
4	532	0.204	0.195
5	408	0.156	0.151
6	273	0.105	0.097
7	139	0.053	0.054
8	45	0.017	0.026
9	27	0.010	0.011
$\geqslant 10$	16	0.006	0.007
总　计	2 608	0.999	1.000

此外，在生物学、医学、工业及公用事业的排队等问题中，泊松分布是常见的. 例如，容器内的细菌数，铸件的疵点数（布的疵点

数),交换台的电话呼唤次数等等,一般都服从泊松分布.

这里,我们来分析推导放射的粒子数 X 为何服从泊松分布.

首先把体积为 V 的某放射性物质设想分割为 n 份相同体积 $\Delta V\left(\Delta V=\dfrac{V}{n}\right)$的小块,并假定:

(1) 对于每个特定的小块而言,在 7.5 s 内放出两个以上 α 粒子的概率为 0(实际上,是放射两个以上的概率很小很小,可以忽略);而放出一个 α 粒子的概率为 p_n:

$$p_n=\mu\Delta V$$

(即 p_n 只跟体积 ΔV 的大小成正比,而跟哪一个小块 ΔV 无关,比例系数为 μ.)

(2) 各小块放出粒子否,是相互独立的.

在这两条假定下,7.5 s 内体积为 V 的某放射性物质放出 k 个粒子,可近似地看作在 V 的 n 个独立的小块中,恰有 k 块放出粒子($n-k$ 块不放出粒子).于是,放出 k 个粒子的概率,就可按独立试验序列来近似计算:

$$P\{X=k\}\approx \mathrm{C}_n^k p_n^k q_n^{n-k}\quad (q_n=1-p_n)$$

然而,上式只是个近似式.容易理解,把 V 无限细分,就能得到 $P\{X=k\}$的精确式.也就是说,

$$P\{X=k\}=\lim_{n\to\infty}\mathrm{C}_n^k p_n^k q_n^{n-k}$$

下面,我们来求出这个极限值.记 $\lambda=\mu V$,将 $p_n=\mu\Delta V=\dfrac{\mu V}{n}=\dfrac{\lambda}{n}$代入,得

$$\begin{aligned}\mathrm{C}_n^k p_n^k q_n^{n-k}&=\frac{n!}{k!\ (n-k)!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}\\&=\frac{1}{k!}\cdot\frac{n(n-1)\cdots(n-k+1)}{n^k}\lambda^k\frac{\left(1-\dfrac{\lambda}{n}\right)^n}{\left(1-\dfrac{\lambda}{n}\right)^k}\end{aligned}$$

当 $n\to\infty$ 时，因 $\frac{n(n-1)\cdots(n-k+1)}{n^k}=\frac{n}{n}\cdot\frac{n-1}{n}\cdots\frac{n-k+1}{n}\to 1$，$\left(1-\frac{\lambda}{n}\right)^k\to 1$，$\left(1-\frac{\lambda}{n}\right)^n\to e^{-\lambda}$，故得

$$C_n^k p_n^k q_n^{n-k}\to\frac{\lambda^k}{k!}e^{-\lambda}，当\ n\to\infty \qquad (2.5)$$

即

$$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda}\quad (k=0,1,2,\cdots)$$

从以上的分析推导过程看出，某一具体问题，只要它符合类似于(1)，(2)的条件，那么就会出现服从泊松分布的随机变量.因此，有很多具体问题，它们的性质虽然各不相同，但它们的随机变量都服从泊松分布.

另外，从上面后半部分的推导可进一步看出：如果 $np\to\lambda>0$，(当 $n\to\infty$)，则

$$C_n^k p^k q^{n-k}\to\frac{\lambda^k}{k!}e^{-\lambda}，(当\ n\to\infty)$$

即泊松分布是二项分布当 $np\to\lambda(n\to\infty)$ 情况下的极限分布.因此，当 n 很大且 p 很小时，可用泊松分布来作二项分布的近似计算.这在第一章的末尾已经提到过.

5. 超几何分布

设一堆同类产品共 N 个，其中有 M 个次品.现从中任取 n 个(假定 $n\leqslant N-M$)，则这 n 个中所含的次品数 X 是一个离散型随机变量.我们知道(参见第一章例 2.5)，X 的概率分布如下：

$$P\{X=m\}=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}\quad (m=0,1,2,\cdots,l) \qquad (2.6)$$

这里 $l=\min(M,n)$.这个概率分布称为超几何分布.

下面来讨论超几何分布与二项分布的关系.我们来证明

若当 $N\to\infty$ 时，$M/N\to p$(n,m 不变)，则

$$\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}\to C_n^m p^m q^{n-m}\quad (N\to\infty) \qquad (2.7)$$

证

$$\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}=\frac{M!}{(M-m)!\,m!}\frac{(N-M)!}{[N-M-(n-m)]!\,(n-m)!}\cdot\frac{(N-n)!\,n!}{N!}$$

$$=\frac{n!}{m!\,(n-m)!}\left(\frac{M(M-1)\cdots(M-m+1)}{\underbrace{N\cdot N\cdot\cdots\cdot N}_{m\text{个}}}\right)\cdot\left(\frac{(N-M)(N-M-1)\cdots[N-M-(n-m)+1]}{\underbrace{N\cdot N\cdot\cdots\cdot N}_{(n-m)\text{个}}}\right)\cdot\left(\frac{\overbrace{NN\cdots N}^{n\text{个}}}{N(N-1)\cdots(N-n+1)}\right)$$

当 $N\to\infty$ 时,不难看出:第一个括号 $\to p^m$,

第二个括号 $\to(1-p)^{n-m}$,

第三个括号 $\to 1$,

因此命题得证.

习 题 六

1. 求例 1.1 中"抽得的次品件数" X 的概率分布.

2. 求例 1.3 中"击中目标次数" X 的概率分布.

3. 求例 1.4 中"所需射击次数" X 的概率分布.

4. 一批零件中有九个正品与三个废品.安装机器时,从这批零件中任取一个.如果每次取出的废品不再放回,而再取一个零件,直到取得正品时为止.求在取得正品以前已取出废品数的概率分布.

5. 抛掷一枚分币,直到出现"正面朝上"时为止.求抛掷次数的概率分布.

6. 从一副扑克牌中发出五张,求其中黑桃张数的概率分布.

7. 设 X 服从泊松分布,且已知 $P(X=1)=P(X=2)$,求 $P(X=4)$.

8. 已知一电话交换台每分钟的呼唤次数服从参数为 4 的泊松分布.求:(1) 每分钟恰有 8 次呼唤的概率;(2) 每分钟呼唤次数大于 8 的概率.

9. 设 X 服从泊松分布,分布律为

$$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda},k=0,1,2,\cdots$$

问当 k 取何值时 $P\{X=k\}$ 最大?

10. 验证等式

$$\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}=1\quad(\lambda>0)$$

(泊松分布的"总概率"为 1.)

11. 利用恒等式

$$(1+x)^N=(1+x)^M\cdot(1+x)^{N-M}$$

两边 x^n 的系数相等,验证等式

$$\sum_{m=0}^{l}\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}=1$$

(超几何分布的"总概率"为 1),其中 N,M,n 是正整数,$N\geqslant M,N\geqslant n$,$l=\min(M,n)$.(当 $i>k$ 时规定 $C_k^i=0$.)

§3 连续型随机变量

1. 概率密度函数

在本章一开始,我们已经遇到过非离散型随机变量的例子,弹着点与目标间的距离就是一个非离散型的随机变量,某人在车站等车的时间也是一个非离散型的随机变量.这两个随机变量的特点是,它们可能取某一区间内所有的值.例如等车时间就可以是区间[0,5)中的任一值.我们知道随机变量 X 取值虽然是"不确定的",但是它具有一定的"概率分布".比如,对于离散型的随机变量 X,对于任何常数 a,b,事件 $\{a<X<b\}$ 也有确定的概率.(为什么?请读者想一想!)对于非离散型的随机变量,考察事件 $\{X=a\}$ 发生的概率往往意义不大,我们干脆直接考察事件 $\{a<X<b\}$ 的概率.为此引进定义:

定义 3.1 对于随机变量 X,如果存在非负可积函数 $p(x)$ $(-\infty<x<+\infty)$,使对任意 $a,b(a<b)$都有

$$P\{a < X < b\} = \int_a^b p(x)\mathrm{d}x \tag{3.1}$$

则称 X 为**连续型随机变量**;称 $p(x)$为 X 的**概率密度函数**(简称概率密度或密度)[①].

这时从(3.1)不难看出,对任何实数 a,$P\{X = a\} = 0$[②],从而有

$$P\{a<X<b\} = P\{a\leqslant X<b\} = P\{a<X\leqslant b\}$$
$$= P\{a\leqslant X\leqslant b\} = \int_a^b p(x)\mathrm{d}x \tag{3.2}$$

作为概率密度 $p(x)$,不难推知有下列性质:

$$\int_{-\infty}^{+\infty} p(x)\mathrm{d}x = 1\text{[③]} \tag{3.3}$$

在实际工作中遇到的非离散型随机变量大多是连续型的,而且其密度函数$p(x)$至多有有限多个间断点(在非离散型的随机

---

① 由(3.1)式,当 $p(x)$在 $x = x_0$连续时,利用定积分的性质容易推知

$$\lim_{\Delta x\to 0+}\frac{P\left\{x_0-\dfrac{\Delta x}{2}<X<x_0+\dfrac{\Delta x}{2}\right\}}{\Delta x} = p(x_0)$$

由此可见,当 $p(x_0)$大时,X 在 x_0附近取值的概率也就较大.概率密度的"密度"一词,跟物理学中质量线密度的"密度"有相似之处.

② 实际上,对任何正整数 n,有$\{X = a\}\subset\left\{a-\dfrac{1}{n}<X<a+\dfrac{1}{n}\right\}$,于是 $P\{X = a\}\leqslant P\left\{a-\dfrac{1}{n}<X<a+\dfrac{1}{n}\right\} = \int_{a-\frac{1}{n}}^{a+\frac{1}{n}} p(x)\mathrm{d}x$,但 $\lim\limits_{n\to\infty}\int_{a-\frac{1}{n}}^{a+\frac{1}{n}} p(x)\mathrm{d}x = 0$,所以 $P\{X = a\} = 0$.

③ 可如下证明:令 $A_n = \{n\leqslant X<n+1\}(n = \cdots,-1,0,1,\cdots)$,易知 $\bigcup\limits_{n=-\infty}^{+\infty} A_n$是必然事件,故其概率为 1.另一方面根据概率的"完全可加性"知,

$$P\left(\bigcup_{n=-\infty}^{+\infty} A_n\right) = \sum_{n=-\infty}^{+\infty} P(A_n) = \lim_{k\to\infty}\sum_{n=-k}^{k}P(A_n)$$
$$= \lim_{k\to\infty}\sum_{n=-k}^{k}\int_n^{n+1} p(x)\mathrm{d}x = \lim_{k\to\infty}\int_{-k}^{k+1} p(x)\mathrm{d}x = \int_{-\infty}^{+\infty} p(x)\mathrm{d}x$$

这就证明了(3.3).

变量中，本讲义只讨论这种连续型的随机变量).

我们强调指出，若将 X 的密度函数在个别点上的值加以改变，得到的仍是 X 的密度函数. 换句话说，密度函数按定义并不是惟一确定的. 但容易看出，若 $p(x)$和 $q(x)$都是 X 的密度函数，又 $p(x)$，$q(x)$都在 $x=x_0$处连续，则一定有$p(x_0)=q(x_0)$(参见上页注①).

下面介绍几个实际工作中常见的连续型随机变量.

2. 均匀分布

如果随机变量 X 的概率密度为

$$p(x)=\begin{cases}\lambda & \text{当} a\leqslant x\leqslant b\\ 0 & \text{其他}\end{cases}\qquad (a<b)$$

则称 X 服从$[a,b]$区间上的**均匀分布**.

由 $\int_{-\infty}^{+\infty} p(x)\mathrm{d}x=1$ 可知 $\lambda=\dfrac{1}{b-a}$.

如果 X 在$[a,b]$服从均匀分布，则对于任意满足 $a\leqslant c<d\leqslant b$ 的c，d，按概率密度的定义(即(3.1)式)有：

$$P\{c<X<d\}=\int_c^d p(x)\mathrm{d}x=\lambda(d-c)$$

上式表明，X 取值于$[a,b]$中任一小区间的概率与该小区间的长度成正比，而跟该小区间的具体位置无关. 这就是均匀分布的概率意义.

§1 的例 1.6 中，候车时间 X 是服从均匀分布的(为什么?). 请读者自己写出它的密度函数.

3. 指数分布

如果随机变量 X 的概率密度为

$$p(x)=\begin{cases}\lambda e^{-\lambda x} & \text{当 } x\geqslant 0\\ 0 & \text{当 } x<0\end{cases}\qquad (\lambda>0)$$

则称 X 服从**指数分布**(参数为 λ).

(指数分布的实际背景在下一节中讨论.)

若 X 服从指数分布(参数为 λ),则按(3.1),对任何 $0 \leqslant a < b$,有

$$P\{a < X < b\} = \lambda \int_a^b e^{-\lambda x} dx = \int_{\lambda a}^{\lambda b} e^{-t} dt$$
$$= e^{-\lambda a} - e^{-\lambda b}$$

由此不难看出 $\int_{-\infty}^{+\infty} p(x) dx = \int_0^{+\infty} p(x) dx = 1$.

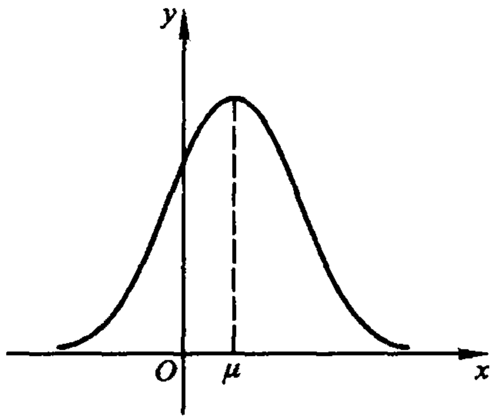


图 2.2

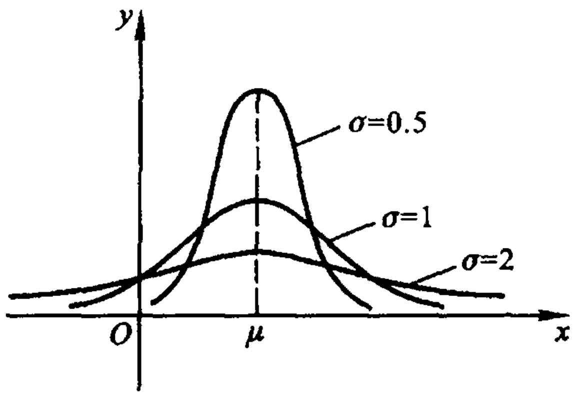


图 2.3

4. 正态分布

如果随机变量 X 的概率密度为

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad (-\infty < x < +\infty)(\sigma > 0) \qquad (3.4)$$

则称 X 服从正态分布 $N(\mu, \sigma^2)$;简记为 $X \sim N(\mu, \sigma^2)$.

$p(x)$在直角坐标系内的图形(见图 2.2)呈钟形,最大值点在 $x = \mu$,曲线相对于直线 $x = \mu$ 对称;在 $x = \mu \pm \sigma$ 处有拐点;当 $x \to \pm\infty$时,曲线以 x 轴为其渐近线.当 σ 大时,曲线平缓;当 σ 小时,曲线陡峭(见图 2.3).

参数 $\mu = 0$ 而 $\sigma^2 = 1$ 的正态分布,即 $N(0,1)$,称为标准正态分布;它的密度函数为

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

读者利用微积分的知识容易验证下式成立①：

$$\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx=1 \tag{3.5}$$

由(3.5)式，不难验证一般的正态密度，即(3.4)式也满足(3.3)式. 只须设$\frac{x-\mu}{\sigma}=t$，有

$$\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=1$$

表达式(3.4)看起来有些复杂，但大量实际经验与理论分析表明，测量误差及很多质量指标，如一批产品的长度、强度等，都可看作或近似看作是服从正态分布的. 正态分布在概率统计的理论与应用中占有特别重要的地位. 这一点我们以后还要谈到.

下面来计算服从正态分布 $N(\mu,\sigma^2)$的随机变量 X 落在区间(a,b)中的概率. 先介绍标准正态分布的概率计算，然后再谈一般的正态随机变量.

例 3.1 设 $X\sim N(0,1)$，求 $P\{1<X<2\}$，$P\{-1<X<1\}$.

解 由定义知道

$$P\{1<X<2\}=\int_1^2\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}dt=\int_{-\infty}^2\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}dt-\int_{-\infty}^1\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}dt$$

---

① 可用下法证明(3.5).

先计算重积分 $I=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}e^{-\frac{x^2+y^2}{2}}dxdy$，作变数替换：$x=r\cos\varphi$，$y=r\sin\varphi$，$(0\leqslant r<+\infty,0\leqslant\varphi<2\pi)$. 于是 $I=\int_0^{+\infty}\int_0^{2\pi}e^{-\frac{r^2}{2}}rdrd\varphi=\int_0^{+\infty}e^{-\frac{r^2}{2}}r\left[\int_0^{2\pi}d\varphi\right]dr=2\pi\int_0^{+\infty}e^{-\frac{1}{2}r^2}rdr=2\pi\int_0^{+\infty}e^{-u}du=2\pi$，但是 $I=\int_{-\infty}^{+\infty}e^{-\frac{x^2}{2}}dx\cdot\int_{-\infty}^{+\infty}e^{-\frac{y^2}{2}}dy=\left[\int_{\infty}^{+\infty}e^{-\frac{x^2}{2}}dx\right]^2$，于是$\int_{-\infty}^{+\infty}e^{-\frac{x^2}{2}}dx=\sqrt{I}=\sqrt{2\pi}$，就证明了(3.5).

设 $\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}dt$，$\Phi(x)$的数值已经算好，列在本书的附表 1 中. 于是 $P\{1<X<2\}=\Phi(2)-\Phi(1)=0.977\,3-0.841\,3=0.136\,0$.

同样，$P\{-1<X<1\}=\Phi(1)-\Phi(-1)$，由 $\Phi(x)$的定义可知 $\Phi(-x)=1-\Phi(x)$[①]，因此 $P\{-1<X<1\}=\Phi(1)-[1-\Phi(1)]=2\Phi(1)-1=2\times0.841\,3-1=0.682\,6$.

例 3.2 设 $X\sim N(\mu,\sigma^2)$，求 $P\{a<X<b\}$.

解 $P\{a<X<b\}=\int_a^b\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}dx$

设$\frac{x-\mu}{\sigma}=t$，则有

$$P\{a<X<b\}=\int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}dt$$
$$=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)$$

查附表 1，就可求出这个概率的值.

特别，

$P\{\mu-\sigma<X<\mu+\sigma\}=\Phi(1)-\Phi(-1)=0.682\,6$

$P\{\mu-2\sigma<X<\mu+2\sigma\}=\Phi(2)-\Phi(-2)=0.954\,4$

---

① $\Phi(-x)=\int_{-\infty}^{-x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$

$\xlongequal{\text{设 } t=-w}-\int_{+\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}w^2}dw$

$=\int_x^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}dt=1-\Phi(x)$

因此，附表 1 只对 x 的正值列出函数 $\Phi(x)$的值.

这个公式的几何意义见图 2.4.

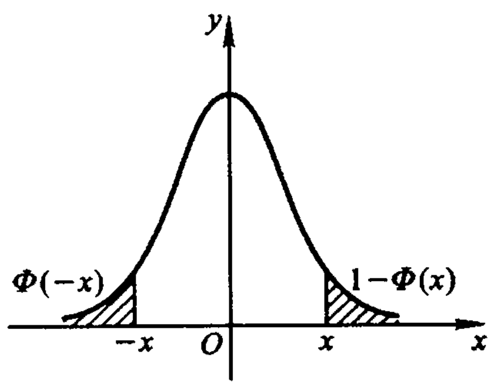


图 2.4

$P\{\mu-3\sigma<X<\mu+3\sigma\}=\Phi(3)-\Phi(-3)=0.997\ 4$[①]

由上面可以看出，服从正态分布 $N(\mu,\sigma^2)$ 的随机变量 X 之值基本上落在区间 $(\mu-2\sigma,\mu+2\sigma)$ 之内，而 X 几乎不在 $(\mu-3\sigma,\mu+3\sigma)$ 之外取值．

例 3.3 设 $X\sim N(2,0.3^2)$，求 $P\{X>2.4\}$．

解 $P\{X>2.4\}=\int_{2.4}^{+\infty}\frac{1}{\sqrt{2\pi}\times 0.3}e^{-\frac{1}{2\times 0.3^2}(x-2)^2}dx$

$=\int_{\frac{2.4-2}{0.3}}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=1-\Phi\left(\frac{2.4-2}{0.3}\right)=1-\Phi(1.33)$

$=0.091\ 8$（这里对附表 1 用了插补法）．

5．Γ 分布

如果随机变量 X 的概率密度为

$$p(x)=\begin{cases}\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x} & x>0\\ 0 & x\leqslant 0\end{cases}\quad(\alpha>0,\beta>0)\qquad(3.6)$$

其中 $\Gamma(\alpha)=\int_0^{\infty}x^{\alpha-1}e^{-x}dx$[②]，则称 X 服从 Γ 分布．简记为 $X\sim\Gamma(\alpha,\beta)$．

Γ 分布含 α,β 两个参数，很多常见的分布都是它的特殊情形：

$\Gamma(1,\beta)$ 就是前面介绍的指数分布；

$\Gamma\left(\frac{n}{2},\frac{1}{2}\right)$ 就是所谓自由度为 n 的卡方分布 $\chi^2(n)$．

6．韦布尔(Weibull)分布

如果随机变量 X 的概率密度为

$$p(x)=\begin{cases}m\frac{x^{m-1}}{\eta^m}\exp\left\{-\left(\frac{x}{\eta}\right)^m\right\} & ,x>0\\ 0 & ,x\leqslant 0\end{cases}$$

---

① 如利用 $\Phi(x)$ 的更精确的表，这三个概率分别是 0.682 7，0.954 5，0.997 3．

② 不难验证 $\Gamma(1)=1,\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi},\Gamma(\alpha)=(\alpha-1)\Gamma(\alpha-1)(\alpha>1)$．

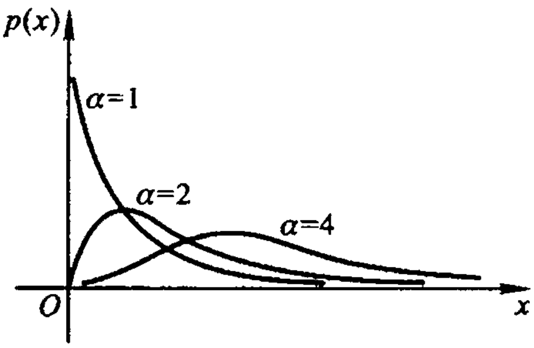


图 2.5　Γ分布密度函数示意图

$(\beta=1;\alpha=1,2,4)$

则称 X 服从韦布尔分布，简记为 $X\sim W(m,\eta)$，其中 $m>0,\eta>0$，m 称为形状参数，η 称为尺度参数. 许多机电产品(如轴承)的寿命服从韦布尔分布. $W(1,\eta)$便是前面介绍过的指数分布. 韦布尔分布在工业产品的寿命与可靠性研究中有广泛应用.

还有一些在理论上或实际中有用的分布，在此就不一一列举了. 可参看下一章的附表.

§4　分布函数与随机变量函数的分布

本节主要讨论随机变量函数的分布，而分布函数这个概念是处理这个问题的重要工具.

1. 分布函数

定义 4.1　设 X 是一随机变量(可以是连续型的，也可以是离散型的，甚至更一般的)，称函数

$$F(x)=P(X\leqslant x)\quad(-\infty<x<+\infty) \tag{4.1}$$

为 X 的**分布函数**[①].

① 有的书上称 $G(x)=P\{X<x\}$ 为 X 的分布函数；近来，越来越多的书上采用(4.1)式的定义. 另外，有的书上称分布函数为“累积分布函数”.

$F(x)$有下列几条一般的性质：

(1) $0 \leqslant F(x) \leqslant 1 \quad (-\infty < x < +\infty)$；

(2) $F(x)$是 x 的不减函数；

(3) $\lim\limits_{x \to -\infty} F(x) = 0, \quad \lim\limits_{x \to +\infty} F(x) = 1$.

(4) $F(x)$是 x 的右连续函数.

由(4.1)及概率的性质可立即证明性质(1)和(2). 至于性质(3)和(4)，由(4.1)式来看也是明显的；而严格的数学证明要用到概率的完全可加性及其推论(见第一章公式(3.8)、(4.22)及(4.23))，此处从略.

为了区别不同随机变量的分布函数，有时将随机变量 X 的分布函数记作 $F_X(x)$.

例 4.1 设 X 服从二点分布

$$P\{X=1\}=p, P\{X=0\}=q(0<p<1, q=1-p)$$

则有

$$F(x)=\begin{cases}0 & x<0 \\ q & 0 \leqslant x<1 \\ 1 & x \geqslant 1\end{cases}$$

$F(x)$的图形见图 2.6，为阶梯形，显然满足性质(1)，(2)，(3)，(4).

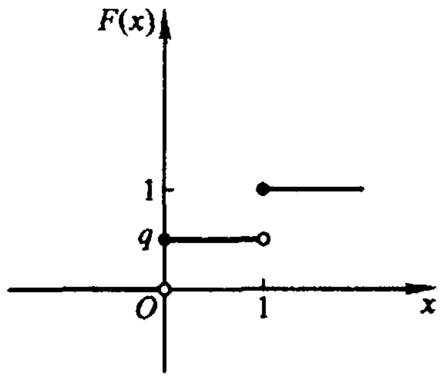


图 2.6

对于连续型的随机变量 X，由(3.1)与(4.1)知道分布函数 $F(x)$跟密度函数 $p(x)$有如下的关系：

$$F(x)=\int_{-\infty}^{x} p(t) \mathrm{d} t \tag{4.2}$$

上式所揭示的 $p(x)$和 $F(x)$这一对函数间的关系，在微积分中是常见的. 用那里的术语来说，$F(x)$是 $p(x)$的可变上限的定积分；因而有

(1) $F(x)$是 x 的连续函数(在整个实轴上)；

(2) 对于 $p(x)$的连续点 x_0而言，有

$$F'(x_0)=p(x_0)$$

在前面曾经提到过，本书只讨论 $p(x)$ 最多有有限多个间断点的情形. 因此，从上面的(2)得知，(除 $p(x)$ 的间断点外)有

$$p(x)=F'(x) \tag{4.3}$$

(4.2)和(4.3)无非表示了 $p(x)$，$F(x)$ 这两者中，由一个可决定另一个. 这本来也是自然的，因为从它们的定义((3.1)和(4.1))来看，它们都是用来刻画随机变量 X 的概率分布的情况的.

作为分布函数的直接应用，我们来看一个例子(指数分布的一个实际背景).

例 4.2 若已使用了 t(h)的电子管在以后的 Δt(h)内损坏的概率为 $\lambda\Delta t+o(\Delta t)$，其中 λ 是不依赖于 t 的数. 假定电子管寿命为零的概率是零. 求电子管在 T(h)内损坏的概率.

设 X 为电子管的寿命，显然对于成批的电子管而言，X 是一个随机变量. 按题意要求 $P\{X\leqslant T\}$；即要求 X 的分布函数 $F(T)$.

对于题设中的"已使用了 t 小时的电子管在以后的 Δt(h)内损坏的概率"，是一个条件概率，用记号来表示就是

$$P\{t<X\leqslant t+\Delta t \mid X>t\}$$

于是按题设有

$$P\{t<X\leqslant t+\Delta t \mid X>t\}=\lambda\Delta t+o(\Delta t)$$

再由条件概率公式，上式左边等于

$$\begin{aligned}\frac{P\{\text{“}t<X\leqslant t+\Delta t\text{”}\cdot\text{“}X>t\text{”}\}}{P\{X>t\}}&=\frac{P\{t<X\leqslant t+\Delta t\}}{P\{X>t\}}\\&=\frac{F(t+\Delta t)-F(t)}{1-F(t)}\end{aligned}$$

上式最后一个等号中用到 $F(t)=P(X\leqslant t)$. 因而有

$$\frac{F(t+\Delta t)-F(t)}{1-F(t)}=\lambda\Delta t+o(\Delta t)$$

即

$$\frac{F(t+\Delta t)-F(t)}{\Delta t}=[1-F(t)]\left[\lambda+\frac{o(\Delta t)}{\Delta t}\right]$$

让 $\Delta t\to 0$,得

$$F'(t)=\lambda[1-F(t)]$$

这是一个关于 $F(t)$的微分方程(一阶线性).联系初始条件

$$F(0)=0$$

解出

$$F(t)=1-e^{-\lambda t}$$

所以电子管在 T(h)内损坏的概率

$$P\{X\leqslant T\}=F(T)=1-e^{-\lambda T}$$

不难看出,X 的分布密度为

$$p(t)=\begin{cases}\lambda e^{-\lambda t} & 当\ t>0\\ 0 & 其他\end{cases}$$

这表明 X 服从参数为λ 的指数分布.

2. 随机变量函数的分布

设 $f(x)$是一个函数,所谓随机变量 X 的函数$f(X)$就是这样的一个随机变量 Y:当 X 取x 时,它取值 $y=f(x)$;记作

$$Y=f(X)$$

例如,设 X 是分子的速率,而 Y 是分子的动能,则 Y 是X 的函数:$Y=\frac{1}{2}mX^2$(m 是分子的质量).

我们的任务是,根据已知的 X 的分布来寻求 $Y=f(X)$的分布.

(1) X 是离散型.

对于 X 是离散型的情形,$f(X)$的分布是不难直接得到的(Y当然也是一个离散型随机变量).

设 X 的概率分布为:

X	x_1	x_2	$\cdots$	x_k	$\cdots$
$P(X=x_i)$	p_1	p_2	$\cdots$	p_k	$\cdots$

记 $y_i=f(x_i)(i=1,2,\cdots)$,如果诸 y_i的值也互不相等,则 Y 的概率分布为

Y	y_1	y_2	$\cdots$	y_k	$\cdots$
$P(Y=y_i)$	p_1	p_2	$\cdots$	p_k	$\cdots$

这是因为 $P(Y=y_i)=P(X=x_i)(i=1,2,\cdots)$.

例 4.3 已知 X 的概率分布为：

X	0	1	2	3	4	5
$P(X=x_i)$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{2}{9}$	$\frac{1}{9}$

则 $Y=2X+1$ 的概率分布为：

Y	1	3	5	7	9	11
$P(Y=y_i)$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{2}{9}$	$\frac{1}{9}$

在 $f(x_1),f(x_2),\cdots,f(x_k),\cdots$不是互不相等的情形，则应把那些相等的值分别合并，并根据概率加法公式把相应的 p_i 相加，就得到 Y 的概率分布.

例 4.4 X 的概率分布同例 4.3.求 $Y=(X-2)^2$的概率分布.

这时 $f(0),f(1),f(2),f(3),f(4),f(5)$分别为 4,1,0,1,4,9.不难看出，$Y$ 的概率分布为：

Y	4	1	0	9
$P(Y=y_i)$	$\frac{1}{12}+\frac{2}{9}$	$\frac{1}{6}+\frac{1}{12}$	$\frac{1}{3}$	$\frac{1}{9}$

(2) X 是连续型.

X 为连续型时，如何找出 $Y=f(X)$的分布呢？

例 4.5 已知 $X\sim N(\mu,\sigma^2)$，求 $Y=\dfrac{X-\mu}{\sigma}$的概率密度.

解 设 Y 的分布函数为$F_Y(y)$，于是

$$F_Y(y)=P(Y\leqslant y)\quad\text{（根据分布函数的定义）}$$

$$=P\left(\frac{X-\mu}{\sigma}\leqslant y\right)\quad\left(\text{因此 } Y=\frac{X-\mu}{\sigma}\right)$$

$$= P(X \leqslant \sigma y + \mu) \quad \text{(不等式变形)}$$
$$= F_X(\sigma y + \mu) \quad \text{(根据分布函数定义)}$$

其中 $F_X(x)$为 X 的分布函数.那么,我们有

$$F_Y(y) = F_X(\sigma y + \mu)$$

将上式两边对 y 求微商,利用密度函数是分布函数导数的关系,我们得到

$$p_Y(y) = p_X(\sigma y + \mu)\sigma$$

上式右端的 σ 是由复合函数求导数得来的.再将

$$p_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

代入,有

$$p_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{[(\sigma y + \mu) - \mu]^2}{2\sigma^2}} \times \sigma = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{y^2}{2}}$$

这表明 $Y \sim N(0,1)$.

请读者注意,在以上推导过程中,除去用到分布函数的定义以及分布函数和密度函数的关系之外,还用到这样一个等式

$$P\left(\frac{X-\mu}{\sigma} \leqslant y\right) = P(X \leqslant \sigma y + \mu)$$

表面上看,只是把不等式“$\frac{X-\mu}{\sigma} \leqslant y$”变形为“$X \leqslant \sigma y + \mu$”,实质上,它们是同一个随机事件,因而概率相等.这里的关键性作用在于把 $Y = \frac{X-\mu}{\sigma}$的分布函数在y之值$F_Y(y)$转化为 X 的分布函数在 $\sigma y + \mu$ 之值$F_X(\sigma y + \mu)$.这样就建立了分布函数之间的关系,然后通过求导可得到 Y 的密度函数.这种方法对于求随机变量的分布是很有用的,姑且叫它“分布函数法”吧.下面几个例子都是用的这个方法.

本例的证明过程中,还用到以下事实:

“如果随机变量 X 的分布函数 $F(x)$具有连续的导函数