

水利科学丛书 / SHUILIKEXUECONGSHU

● 金光炎 著

水文水资源应用统计计算



S

huiwen Shuiziyuan
Yingyong Tongji Jisuan

东南大学出版社

水文水资源应用统计计算

金光炎 著

东南大学出版社

• 南京 •

内 容 提 要

本书为水文水资源分析计算中应用概率统计方法的初级读物和实际操作的工具书。主要内容为:概率统计在水文水资源领域应用的基本知识,统计参数的估计方法和应用,资料插补延长和系列间关系分析的回归和相关计算,提供洪枯水设计数据的频率分析,抽样误差问题等。本书内容通俗易懂,深入浅出。对各种水文水资源的计算,不仅介绍了手算方法,还介绍了利用 Excel 软件在计算机上进行计算,不需编程,计算简便。附录中列出了对专门问题的说明和应用表格。

本书读者对象为中专以上的水文水资源工作者,适用于水利、气象、地质、地理、交通、农业等部门的有关人员,也可作为大专院校师生的参考书或辅导材料。

图书在版编目(CIP)数据

水文水资源应用统计计算 / 金光炎著. —南京:
东南大学出版社, 2011. 5

ISBN 978 - 7 - 5641 - 2730 - 5

I. ①水… II. ①金… III. ①水文计算②水资源-水利计算 IV. ①P333②TV214

中国版本图书馆 CIP 数据核字(2011)第 075451 号

水文水资源应用统计计算

出版发行 东南大学出版社
出 版 人 江建中
社 址 南京市四牌楼 2 号
邮 编 210096
网 址 <http://www.seupress.com>

经 销 全国各地新华书店
印 刷 南京新洲印刷有限公司
开 本 700 mm×1000 mm 1/16
印 张 9.25
字 数 179 千字
版 次 2011 年 5 月第 1 版
印 次 2011 年 5 月第 1 次印刷
书 号 ISBN 978 - 7 - 5641 - 2730 - 5
定 价 32.00 元

(本社图书若有质量问题,请与读者服务部联系。电话:025-83792328)

前 言

各类水文水资源问题可用不同的途径和方法加以解决。本书以概率论和数理统计学的方法为工具,对水文水资源数据进行处理和分析。

本书为水文水资源分析计算中应用概率统计方法的初级读物和实际操作的工具书。主要内容为:概率统计在水文水资源领域内应用的基础知识,统计参数的估计方法和应用,资料插补延长和系列间关系分析的回归和相关计算,提供洪枯水设计数据的频率分析,抽样误差问题等。附录中列出了对专门问题的说明和应用表格。

本书是在《水文统计计算》一书的基础上修订补充而成。该书自 1980 年发行以来,已有 30 多年,在这段时间内,计算机技术的普及和应用为水文水资源分析计算提供了有利条件。书中的计算,用两种方式进行叙述:一种为手算法,使读者了解计算的全过程;另一种为应用 Excel 软件的有关函数直接计算,可以快速得到计算结果。

Excel 软件的内容十分丰富。由于作者对其功能不够熟悉,其中所介绍的方法是在边算边学中所得,可能并非是最简便的,供参照应用。

本书对许多概率统计方面的理论问题只略作说明,不进行详细推导和论证。同时,对水文水资源中存在的争议,也未提及,以免混淆思路,影响初学者的学习。本书主要按照 SL 278—2002《水利水电工程水文计算规范》和 SL 44—2006《水利水电工程设计洪水计算规范》中所述的方法进行介绍,叙述浅近,对各种方法均有在计算机上利用 Excel 软件进行操作的具体介绍,不需编程,计算简便。

本书承尚新红、柏菊等同志的帮助,深表谢意!

本书读者对象是中专以上的水文水资源工作者,适用于水利、气象、地质、地理、交通、农业等部门的有关人员,也可作为大专院校师生的参考书和辅导材料。

限于作者的水平,会有不妥之处,请读者不吝指正。

金光炎

2011 年 2 月于蚌埠

目 录

1	随机事件和概率计算	(1)
1.1	概率论的研究对象	(1)
1.2	水文现象的必然性和随机性	(2)
1.3	事件和事件的种类	(2)
1.4	概率的意义	(4)
1.5	概率的直接计算	(5)
1.6	事件的频率	(6)
1.7	概率乘法定理	(7)
1.8	概率加法定理	(8)
1.9	二项概率定理	(9)
2	随机变数和概率分布	(12)
2.1	随机变数	(12)
2.2	总体和样本	(12)
2.3	频率分布	(14)
2.4	概率分布和概率密度	(15)
3	统计参数	(17)
3.1	常用的统计参数	(17)
3.2	统计参数的算例——手算法	(23)
3.3	统计参数的算例——Excel 算法 1	(24)
3.4	统计参数的算例——Excel 算法 2	(27)
3.5	参数的简捷换算法	(28)
4	正态分布	(30)
4.1	正态分布的形式	(30)
4.2	应用的表格	(31)
4.3	正态分布的特性	(32)

4.4	标准误和机误	(33)
4.5	概率格纸	(34)
5	Γ 分布	(39)
5.1	Γ 分布的形式	(39)
5.2	Γ 分布的主要特性	(40)
5.3	离均系数	(41)
6	简单的相关计算	(44)
6.1	相关的意义和作用	(44)
6.2	两变数的直线相关	(45)
6.3	相关系数和回归线的误差	(48)
6.4	相关计算算例——手算法	(50)
6.5	相关计算算例——Excel 算法 1	(51)
6.6	相关计算算例——Excel 算法 2	(54)
7	图解相关和经验方程	(57)
7.1	直线拟合	(57)
7.2	幂函数拟合	(59)
7.3	指数函数拟合	(61)
7.4	多项式拟合——手算法	(62)
7.5	多项式拟合——Excel 算法	(63)
8	复相关	(67)
8.1	三变数相关	(67)
8.2	多变数相关	(70)
8.3	复相关系数	(72)
8.4	多变数相关举例	(72)
9	频率计算	(74)
9.1	概述	(74)
9.2	经验频率	(75)
9.3	目估适线法	(76)
9.4	三点适线法	(78)
9.5	有特大值时的频率计算	(80)
10	抽样误差	(84)

10.1 概述	(84)
10.2 统计参数的抽样误差	(84)
10.3 设计值的抽样误差	(85)
10.4 经验频率的抽样误差	(85)
10.5 相关系数的抽样误差	(85)
 附录	(88)
附录 A 排列和组合	(88)
附录 B Γ 函数和 B 函数	(93)
附录 C 优化适线法	(99)
附录 D 有关统计检验等问题的说明	(107)
附录 E 线性矩	(113)
附录 F Γ 分布离均系数 Φ 值表	(125)
附录 G Γ 分布模比系数 K 值表	(128)
附录 H Γ 分布三点适线法用表	(135)
 主要参考文献	(137)

1 随机事件和概率计算

1.1 概率论的研究对象

人们在实践活动和生活过程中,经常会碰到各种各样的随机事件(偶然事件),而这些事件的发生机遇常与一定的可能性相联系。当可能性用数字来表达时,称为可能率,即概率,这是概率论中最基本的概念。

概率论是一门研究随机现象规律性的数学学科。早在三百多年前,概率论已被物理学家和数学家所注意。当初,概率论只用于博弈、保险事业和测量误差估计中。随着科学技术的发展和实际工作的需要,特别是到了近代,概率论在各类学科和生产实践中的应用日益增多。现在,几乎没有一门学科不在某种形式下应用概率论的原理和方法了。

先举两个例子来说说随机现象。

(1) 同一距离用同一皮尺量测多次,所得的结果彼此略有差异。这些差异是由于量测过程中受许多次要因素的影响而产生的。例如,拉皮尺时的松紧不一、风吹皮尺的影响、起点和终点的位置略有不同以及视觉上的偏差等。

(2) 给定相同的降雨强度和降雨时间,在同一块场地上进行多次人工降雨试验,每次所得的结果总有些不同:最大流量数值不同,集流时间不同,径流量也不相同。这些差异是属于随机性的,它们受许多因素的影响。例如,由于下渗的不均匀、蒸发的不同、喷雨装置的摆动、风的干扰和测量误差等。

上述例子说明,在基本条件保持不变的情况下,多次试验会得到不同的结果,其原因是在基本条件之外尚存在着许多次要因素,而这些次要因素或多或少地影响到试验的最终结果。我们说这种结果上的差异,是次要因素引起的随机性差异。这种具有随机性差异的现象称为随机现象。

显然,宇宙中任何实际事物不可能不带有某些随机成分。就是十分精确地固定试验条件和仔细地观测试验数据,也不可能做到反复试验结果的完全一致。因此,与水文现象有关的各种因素都带有一定的随机成分。通常,把这种含有随机成分的实测系列称为随机系列。

实践表明,由少数试验所得的随机系列,往往是杂乱无章的,随着试验次数的增多,这种随机系列会呈现出一定的规律性。此类规律称为统计规律。

与概率论相提并论的是数理统计学。数理统计学的内容,简单地说,是以概

率论为基础,用数学的理论和方方法对随机系列进行分析和计算。同时,概率论往往把由数理统计学所揭露的事实,提高到理论,从而丰富自己的内容。概率论和数理统计学是密切联系的,通常总是把它们称呼在一起。

概率论和数理统计学已单独成为一门数学分支,其理论是严谨的,它们的方法能解决许多实际问题。因此,目前得到了广泛的应用。

研究水文现象和水资源方面的问题,有不同的途径,主要是物理途径和统计途径。物理途径是根据所研究事物的成因和发生、发展规律进行分析计算;统计途径是按照它们的发生机遇与统计规律来做出评估。后者是本书的主要内容。

1.2 水文现象的必然性和随机性

水文现象与其他自然现象一样,在它的发生、发展和演变过程中,包含有必然性的一面,也有随机性的一面。促使水文现象发生的根本原因(成因)规定着它的规律性,这种规律性按照一定的秩序贯串在全部发展过程中,致使水文现象具有必然性的一面。例如,大气运行的结果,必然会引起降雨而产生径流,以及水文情势以年为周期的循环性和明显的季节性。这充分表明了水文现象产生的必然性和不可避免性。可是,水文现象的发展过程,不仅由其根本原因所决定,另外还要受到周围许多因素或多或少的影影响。这些影响的无限复杂性和多样性,使得水文现象在演变过程中不断地发生各种程度的非根本性的偏差,导致其固定秩序不能以纯粹的形式出现,而是伴随着无数可有可无、可以这样那样的情况出现。例如,一次暴雨后,必然会产生径流,但径流的形成受到气象和自然等因素的影响,致使无法用其固有的规律推知其实际出现的数量以及在时间和空间上的确切分布。这说明了水文现象的随机性和不确定性的一面。当然,对于任何自然现象,起主导和决定性作用的是必然性规律,但因伴随有随机性的因素,因此需要认真研究和处理它们。这说明了必然性和随机性是既相互联系又相互有区别的一对范畴,不能把它们看做是绝对对立的。如果认为水文现象只能受必然性规律所支配,那就会忽视其随机性的一面,而排除概率论和数理统计学这一有效分析工具。另一方面,如果只看到水文现象随机性的一面,忽视物理成因上的分析,必然会得到不切实际的结果。因此,在水文分析和水资源计算中,必须对各种有关因素进行详细研究,把必然性规律和随机性规律的分析密切结合起来,相辅相成地解决实际问题。

1.3 事件和事件的种类

“事件”是概率论中最基本的概念,它是指在一定的条件组合下,在试验结果

中所有可能出现或可能不出现的事例。事件分 3 类,分述如下。

(1) 必然事件

如果在条件组每次实现之下,某一事件在试验中一定会发生,称此事件为此试验的必然事件。例如,水在 760 mmHg 毫米汞柱($1 \text{ mmHg} = 133.3224 \text{ Pa}$)大气压力之下,加热到 100°C 时,则化为蒸汽。也就是说,在这 2 个条件(条件组)每次实现之下,水必然化为蒸汽。

(2) 不可能事件

如果在条件组每次实现之下,某一事件在试验中永远不会发生,称此事件为此试验的不可能事件。例如,天然河流如上游无阻水和蓄水建筑物,则洪水来临时必然涨水,发生断流是不可能事件。

(3) 随机事件

如果在条件组每次实现之下,若某一事件在试验中可以发生也可以不发生,称此事件为随机事件。在实际问题中,往往在一定条件组的每次实现之下,可能出现的事件不只有一种,而是有好几种,这种情况也是随机事件。举一个简单的例子,设盒中有红白黄三色的圆球,任取一个,可能是红的,可能是白的,也可能是黄的,即取出任何一色的圆球都是随机事件。又如,每年汛期,河流中必然会出现一次最大流量,对这种现象的本身来说完全不是随机的。但在每个年份内,这个最大流量可能大于指定的流量,也可能小于或等于指定的流量。因此,可以说,此河流中每年大于指定流量的最大流量的出现是随机事件。

为了分析计算的方便,用 A 、 B 、 C 等字母作为事件的代号。同时,在讨论 2 个或多个事件时,要表达事件之间的关系,若用文字来说明,可能很繁冗,因而需引入一些记号来表达它们之间的关系。下面是几类事件的定义和说明。

(1) 事件 $A+B$

表示事件 A 和事件 B 中至少发生其中的一个事件。同样,事件 $A_1 + A_2 + \cdots + A_n$ 表示在事件 $A_1, A_2, \cdots, A_n$ 中至少发生其中的一个事件。

(2) 事件 AB

表示事件 A 和事件 B 同时发生的事件。如果事件 A 和事件 B 永远不能同时发生,则称事件 A 和事件 B 为互斥(互相排斥)或互不相容。同样,事件 $A_1 A_2 \cdots A_n$ 表示事件 $A_1, A_2, \cdots, A_n$ 同时发生的事件。如果事件 $A_1, A_2, \cdots, A_n$ 中任意两个均为互斥,则称它们为彼此互斥。

下面举例说明。钱币有正反两面,在一次投掷中,出现正面设为事件 A ,出现反面设为事件 B 。再设第 1 次投掷时出现正面和反面的事件分别记为 A_1 和 B_1 ,第 2 次投掷分别记为 A_2 和 B_2 ,第 3 次投掷分别记为 A_3 和 B_3 。事件 $A_1 + A_2 + A_3$ 表示在连续 3 次投掷中,至少有 1 次出现正面的事件。事件 $B_1 B_2 B_3$ 表示在连续 3

次投掷中都出现反面的事件。显然,事件 $A_1 + B_1, A_2 + B_2, A_3 + B_3$ 都是必然事件。事件 $A_1 B_1$ 表示在第 1 次投掷中,既要它出现正面、同时又要它出现反面,当然是不可能事件。因此事件 A_1 和 B_1 互斥。

再引进一个概念。如果 $A+B$ 为必然事件,而 A 和 B 互斥,则称 B 为 A 的补事件,用 $\bar{A}$ 来代表 B 。同样, A 为 B 的补事件,用 $\bar{B}$ 来代表 A 。例如,投掷钱币出现正面和出现反面互为补事件。

1.4 概率的意义

每一事件的发生都有某种程度的可能性,有的可能性大一些,有的可能性小一些。例如,一枚均质钱币投掷一次有两种可能的结果,或者是出现正面,或者是出现反面,故可以直观地断定出现正面或反面的可能性是相等的。由于一枚钱币投掷一次的全部出现情况有两种,而实际出现的只能是其中的一种,因此可以说出现正面或反面的可能性都是 $1/2$ 。又如,一年之中,降小雨的次数较多,降大雨的次数较少,而降特大暴雨的次数更少。由此可推知,每次降中小雨的可能性大,降大雨的可能性较小,降特大暴雨的可能性更小。如果把这种可能性的大小用数量来表示,称这一数量为出现所指事件的概率。

概率分两类:先验概率和后验概率。如果某种事件出现和不出现的种种情况的可能性都非常清楚地知道,像上述投币那样,则称为先验概率。若对某一类事件不能预知其出现和不出现种种情况的可能性,像上述降雨问题那样,要估算它的概率,就只能通过多次试验来求得,则称为后验概率。

后验概率亦称经验概率,它在水文水资源计算中叫做频率。水文现象中,如某一大小的降雨量或洪峰流量的发生率,其先验概率无法知道,故仅能借助已有的实测资料用数理统计方法来估计它们的频率。由概率论中的大数定理得知,当观测试验资料(即实测资料)很多时,频率就比较稳定;当实测资料足够多时,频率会接近于一个常数。这就是说,观测试验次数越多,其频率会越接近于客观存在的值。因此,在统计分析时,要求有足够多的和可靠的资料。

需要说明,后验概率是从实践而得,但不能认为先验概率是脱离实际而未卜先知的。先验概率的获得也来源于实践,如投币问题,人们通过多次试验,发现正反面出现的概率几乎相等。经过了多次实践,才认识到它们在理论上的概率应相等,即各为 $1/2$ 。

不论在科学研究中或日常生活中,预先估计做某件事的成败概率,具有明显的意义。如果能较好地估计,常可在一定程度上作为我们行动的指导。

1.5 概率的直接计算

有许多这样的试验,它的各种结果具有对称性。由于这种对称性,可知各种结果在客观上具有同等的可能性,或者说各种结果的概率相同。例如,投掷一枚均质的钱币,没有理由说出现正面的可能性会比出现反面的可能性大一些或小一些,所以它们的概率相等。对于结果为对称且具有同等可能性的试验,可用以下方法进行计算。

设某一试验共有 n 种不同的可能结果,其中各个结果均具有对称性,即等可能性。以 m 表示有利于出现事件 A 的可能结果数,则出现事件 A 的概率为:

$$P(A) = \frac{m}{n} \quad (1.1)$$

式(1.1)称为概率直接计算公式,是概率计算中最基本的公式。概率论中常称它为古典概率公式。下面列举两个简例。

【例 1.1】 盒中装有形状和大小完全一样的 10 个圆球,其中 4 个为红球、6 个为白球,试问任意取出 1 个红球的概率是多少?

【解】 因为全部可能结果数 $n=10$,取出红球的可能结果数 $m=4$,故取出 1 个红球的概率为:

$$P(1 \text{ 个红球}) = \frac{4}{10} = 0.4$$

【例 1.2】 在 52 张扑克牌中,任意抽取 1 张,试问抽得 K 的概率是多少?

【解】 此题中 $n=52$,因 K 有 4 张,故抽得 1 张 K 的概率为:

$$P(1 \text{ 张 K}) = \frac{4}{52} = \frac{1}{13}$$

在式(1.1)中,如令 $m=n$,即有利场合的可能结果数就是试验的全部可能情况,故

$$P(A) = \frac{n}{n} = 1$$

概率等于 1 表示不论在哪次试验中,欲求的事件总是出现的,此即必然事件。

例 1 中,如果取出 1 个红球也行、取出 1 个白球也行,这就是必然事件。

在式(1.1)中,如令 $m=0$,也就是全为不利事件,有

$$P(A) = 0$$

概率为 0,表示在每次试验中都不会出现欲求的事件,此即不可能事件。例 1 中,如果要求取出 1 个黑球,当然是不可能的。

一般,每次试验的结果,可能出现这种情况,也可能出现那种情况,如在例 1

中,每次取出的球,可为红球(出现概率为 0.4),也可为白球(出现概率为 0.6),这类事件通称为随机事件。随机事件的概率总是介于 0 与 1 之间,不可能是负数,也不可能大于 1,这是概率的一个重要特性。

试验结果的对称性,通常只能在组织得很巧妙的试验中见到,所以上述的计算方法,只能适用于一些很简单的问题,至于较为复杂的问题如何计算,下面还要介绍。

1.6 事件的频率

频率的概念在前面已有介绍,这里再做进一步叙述。

如上所述,概率直接计算公式(1.1)只适用于这样一类试验,就是在试验中每种可能结果的发生都必须具备对称性的条件。但在许多实际问题中,远不是所有试验都严格符合这个条件,当然不能用式(1.1)了。例如,一个制作均匀的六面体,投掷一次,任何一面向上的概率都等于 $1/6$ 。如果这个六面体制作得不大均匀,那么某一面向上的概率可能比 $1/6$ 大,也可能比 $1/6$ 小,但它必然有一个客观存在的概率。这个概率可以通过大量试验来计算,也就是说,可用频率来估算概率。

频率的计算与概率的直接计算相似。假设做了一系列的试验,每次试验结果,事件 A 可以出现或不出现。我们把事件 A 出现的次数和试验的总次数相比,此比值即为事件 A 在这一系列试验中出现的频率。以 n 表示试验的总次数、 m 表示事件 A 出现的次数,则事件 A 出现的频率为:

$$P^*(A) = \frac{m}{n} \quad (1.2)$$

式中, P^* 表示与式(1.1)中的 P 有所区别,在不致混淆时,以后将统一使用符号 P 。

实验证明,对于次数不多的试验,事件的频率有着明显的随机性。例如,把一个钱币投掷 10 次,出现正面的次数不可能正好等于 5 次,也许是 6 次或更多,也许是 4 次或更少。另外 10 次投掷中,又可能与前 10 次的出现次数不一样。在试验次数增多之后,事件的频率就逐渐趋于稳定,而在某一较小范围内摆动。试验次数越多,频率就越趋近于一个常量。投掷钱币的试验中,从理论上讲,出现正面的频率应为 0.5,以前有人做过投掷 24 000 次的试验,结果有 12 012 次出现正面,其频率为 0.5005,这就与理论值十分接近了。

概率论中的大数定理,严格地证明了在试验次数无限多时频率接近于概率的事实。同时,试验也验证了大数定理的正确性,这给实际工作带来了很大的方便。水文现象中,多数事件(如某站日降雨量大于 100 mm,洪峰流量超过 $10\,000\text{ m}^3/\text{s}$ 等)的先验概率未知,就可以通过逐年积累的观测资料,用频率来估

算概率。

在理论上和实践上给出了频率与概率之间有机联系,这一点具有很大的实用意义。当无法求得复杂事件的概率时,可以做多次试验,把事件出现的频率作为其概率的近似值。应当记住频率与概率之间有联系的概念,也要识别它们之间的区别。

总之,概率是随机事件在客观上出现可能性的数字表达,是一个常量。频率是一个经验值,随着试验次数的增多而趋近于概率值。所以,复杂事件的概率是可以被认识的和可以设法估计的,不可知论和抱消极态度者都是不对的。水文事件同上述投掷钱币和抽扑克牌等简单事件不一样,因为水文资料不可能人为地在短时间内像简单事件那样重复进行试验获得,而必须通过各种手段年复一年地观测得到。这些资料是很宝贵的,有的地方还有历史资料(如档案中的洪水记录、洪水痕迹和当地居民的回忆等),都可用来作为估算水文事件频率的依据或参考。

1.7 概率乘法定理

这里只介绍独立事件同时出现的概率的计算方法。在多个事件中,若任何一个事件的出现不影响其他事件的出现,而其他事件的出现也不影响这一事件的出现,这种性质称为独立或相互独立。具有独立性质的事件叫做独立事件。例如,投掷二次钱币,第一次出现正面后,并不影响第二次再出现正面。这就是独立事件。又如,同一测流断面上的水位和流量,当出现高水位时必然通过大流量,而低水位与小流量相对应,所以水位与流量不是独立事件。

独立事件在概率论和数理统计学中占有很重要的地位,因为在具备独立的条件时,可以使计算方法大为简化,并便于分析研究。水文现象中有不少资料是相互独立或近于独立的,例如,同一站上各年年雨量是相互独立的。但是,也有不少水文资料相互之间并不独立,如年内相邻的日流量就是如此。

【定理 1.1】 2 个独立事件同时出现的概率等于这些事件各自出现概率的乘积。

设甲种试验共有 n_1 种结果,其中有利于 A 事件者有 m_1 种。另外,又有与甲完全无关(即相互独立)的乙种试验,它共有 n_2 种结果,其中有利于 B 事件的有 m_2 种。若甲、乙同时试验一次,总的试验共有 $n_1 n_2$ 种,同时出现 A 和 B 的结果共有 $m_1 m_2$ 种。因此,出现 A 同时也出现 B 的概率为:

$$P(AB) = \frac{m_1 m_2}{n_1 n_2} = \frac{m_1}{n_1} \frac{m_2}{n_2}$$

显然,

$$\frac{m_1}{n_1} = P(A)$$

$$\frac{m_2}{n_2} = P(B)$$

故得:

$$P(AB) = P(A)P(B) \quad (1.3)$$

这就是 2 个独立事件概率乘法定理的计算公式。同理,如果独立事件多于 2 个,其同时出现的概率可用下列定理。

【定理 1.2】 k 个独立事件同时出现的概率等于 k 个事件各自出现概率的乘积。

设 k 个独立事件为 $A_1, A_2, \dots, A_k$, 比照式(1.3), 得到该定理的表达式为:

$$P(A_1 A_2 \cdots A_k) = P(A_1)P(A_2) \cdots P(A_k) \quad (1.4)$$

【例 1.3】 两枚钱币同时投掷 1 次, 试问同时出现正面的概率是多少?

【解】 一枚钱币投掷 1 次得正面的概率是 $1/2$, 另一枚钱币投掷 1 次得正面的概率也是 $1/2$, 按独立事件概率乘法定理, 其同时出现正面的概率为 $1/2 \times 1/2 = 1/4$ 。

【例 1.4】 某测站的水位每年出现高于 2 m 的概率为 $1/2$ 和出现高于 3 m 的概率为 $1/5$, 试问在连续两年中, 第一年出现高于 2 m 和第二年出现高于 3 m 的概率是多少?

【解】 按独立事件概率乘法定理, 得欲求的概率为 $1/2 \times 1/5 = 1/10$ 。

1.8 概率加法定理

本节介绍互斥事件的概率加法定理。

【定理 1.3】 若干互斥事件中任何几个事件出现的概率等于这几个事件各自出现的概率之和。

设某种试验有 n 种不同的可能结果, 其中有利于 A_1 事件的有 m_1 种, 有利于 A_2 事件的有 m_2 种, $\dots$, 有利于 A_k 事件的有 m_k 种。这 k 个事件 $A_1, A_2, \dots, A_k$ 在一次试验中仅能出现 1 个(即为互斥), 则有利于 A_1 、或 A_2 、 $\dots$ 、或 A_k 的数目共有 $m_1 + m_2 + \dots + m_k$ 个, 故 k 个事件中任何一个出现的概率为:

$$P(A_1 + A_2 + \dots + A_k) = \frac{m_1 + m_2 + \dots + m_k}{n}$$

显然,

$$\frac{m_1}{n} = P(A_1)$$

$$\frac{m_2}{n} = P(A_2)$$

$$\vdots$$

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

故得:

$$P(A_1 + A_2 + \cdots + A_k) = P(A_1) + P(A_2) + \cdots + P(A_k) \quad (1.5)$$

这就是互斥事件概率加法定理的公式。

【例 1.5】 盒中有 5 个红球、10 个白球和 15 个黑球,试问任意摸取 1 个球得红球或白球的概率是多少?

【解】 盒中共有 30 个球,即 $n=30$ 。取出 1 个红球的概率为 $5/30$,取出 1 个白球的概率为 $10/30$,根据互斥事件的概率加法定理,得

$$P(1 \text{ 个红球或 } 1 \text{ 个白球}) = \frac{5}{30} + \frac{10}{30} = \frac{1}{2}$$

【例 1.6】 例 1.5 中,任取 2 个球,试问取出 2 个白球或 2 个黑球的概率是多少?

【解】 在 30 个球中任取 2 个球共有 $\binom{30}{2}$ 种可能情况,即 $n = \binom{30}{2} = 435$ 。10 个白球中任取 2 个球共有 $\binom{10}{2}$ 种可能情况,即 $m_1 = \binom{10}{2} = 45$;15 个黑球中任取 2 个球共有 $\binom{15}{2}$ 种可能情况,即 $m_2 = \binom{15}{2} = 105$ 。故得:

$$P(2 \text{ 个白球或 } 2 \text{ 个黑球}) = \frac{45 + 105}{435} = \frac{10}{29}$$

【例 1.7】 某流域上游建有甲、乙、丙 3 个水库,各水库互不相容。水库的破坏概率:甲库为 0.010,乙库为 0.005,丙库为 0.001。试求至少有 1 个水库破坏的概率是多少?

【解】 至少有 1 个水库破坏的概率为 $0.010 + 0.005 + 0.001 = 0.016$ 。

需要说明,例 1.6 及以后的叙述中,要用到组合 $\binom{n}{m}$ 所表达的组数计算问题。一般,当 n 较大时,计算比较麻烦,现可以利用 Excel 软件在计算机上进行计算,就方便得多。详见附录 A。

1.9 二项概率定理

独立事件进行多次重复试验,对于事件 A ,每次试验中只有两种可能的结果:发生与不发生。设发生的概率为 p ,不发生的概率为 q ,显然 $p + q = 1$ 。当 n 次重复试验后,现来计算恰好有 m 次发生事件 A 的概率。

在 n 次重复试验中,有 m 次发生,则其余的 $n-m$ 次一定是不发生。因此,在全部试验中,有 m 次是 p 的概率及 $n-m$ 次是 q 的概率。根据概率乘法定理得到它们同时出现的概率为 $p^m q^{n-m}$ 。但在 n 次重复试验中, p 和 q 出现的次序不只有 1 种。也就是说,它们的出现次序可以是 $p, p, q, \dots$,也可以是 $p, q, p, \dots$,等等,故共有 $\binom{n}{m}$ 种组合。由此得到 n 次重复试验中有 m 次发生事件 A 的概率为:

$$B(m) = \binom{n}{m} p^m q^{n-m} = \binom{n}{m} p^m (1-p)^{n-m} \quad (1.6)$$

因为式(1.6)为二项式 $(p+q)^n$ 中的通项(即第 $m+1$ 项),故定名为二项概率公式,由此引出的定理称为二项概率定理。

进而, n 次重复试验后,至少 m 次发生(即发生 m 次,或 $m+1$ 次, ..., 或 n 次)的概率 P 可由下式计算得到:

$$P = \binom{n}{m} p^m q^{n-m} + \binom{n}{m+1} p^{m+1} q^{n-m-1} + \dots + \binom{n}{n-1} p^{n-1} q + p^n \quad (1.7)$$

当 m 较小时,应用式(1.7),因项数较多,计算比较麻烦,可用二项展开式的对称性原理简化,即

$$P = 1 - \sum_{r=0}^{m-1} \binom{n}{r} p^r q^{n-r} \quad (1.8)$$

【例 1.8】 投掷钱币 10 次,试问出现 6 次正面的概率是多少?

【解】 此例中, $n=10$, $m=6$, $p=q=1/2$,故得欲求的概率为:

$$B(6) = \binom{10}{6} \left(\frac{1}{2}\right)^6 \left(\frac{1}{2}\right)^4 = 0.205$$

【例 1.9】 某河段,每年超过最大洪峰流量 Q 的概率为 10%,试问在最近 10 年中均不超过的概率是多少? 又超过 1 次的概率是多少?

【解】 此例中, $n=10$, $p=0.1$, $q=0.9$, 10 年中均不超过的概率为(此时 $m=0$):

$$B(0) = (0.9)^{10} = 0.349$$

超过 1 次的概率(此时 $m=1$)为:

$$B(1) = \binom{10}{1} (0.1)(0.9)^9 = 0.387$$

在水利工程的规划中,类似例 1.6 的问题较多。这种例子说明,如果用 Q 去设计该河段的河槽断面以排泄内涝,则在未来 10 年中该河段附近不受内涝的概率为 0.349,而要遭内涝灾害(即至少有 1 次内涝)的概率为 $1-0.349=0.651$ 。这就是说,受灾的可能性比不受灾的可能性几乎大 1 倍。假如允许它在 10 年中可内