# 水文统计原理与方法

金光炎著

# 再 版 序 言

这一版跟1959年的初版有很大的不同：內容增加了、体系調整了和把一些无发展前途的方法刪节了。

在修訂中，作者力图做到：詳細闡述水文分析中常用的机率論和数理統計法的基本概念和理論；充分介紹国內外在水文統計研究上的成就。

主要的修改情况如下：

緒論是全部重写的。其中叙述了用机率論与数理統計法来处理水文問題的理論依据和現实意义，幷簡要地介紹了水文統計的发展过程。

机率論方面的知識，从原有的一章分成了两章。在第二章中，只叙述机率的一般概念和定理；对于一些专門性的理論和方法，都放在第四章中。考虑到第四章中的內容，要用到各种参数，故把这一章置于第三章統計参数之后。

頻率分析为本书的重点，由原有的一章扩展至五章。第五章中对頻率分析作概括的介紹，其他几章分别叙述几类专門性的問題：第六章是頻率曲綫綫型，其中对克一閔曲綫、对数正态曲綫和耿貝尔曲綫都作了进一步的分析，刪节了无发展前途的曲綫綫型和增加了各綫型的比較和国內外的現况。第七章是頻率曲綫的繪制方法，亦即統計参数的估計方法。这里增添了近几年才发展起来的三点法和加求矩差法，幷充实了其他各种方法的內容，章末也列举了各方法間的比較結果和現况。第八章是特殊資料的頻率处理，除特大値处理的一节外，还加上零項和缺項处理方法等两节。第十章是組合頻率，这一章沒有多大的改动，但因其中要用到相关方面的知識，故把它放在第九章相关分析之后。

其他三章改动較少，但也加了一些重要的內容，如最小二乘法的原理、相关系数和經驗頻率的抽样誤差。另外，附表相应地增加了幷編制了索引。

由于水文統計的內容很多，近几年內发展較快，因此不可能把所有方法都列到本书中去。有的同志希望能增加經驗公式的选配方法、时序系列和随机过程等方面的知識，由于我对这些尙缺乏研究，故还不能系統介紹出来。

在本书修訂时，曾得到水文研究所領导上的大力支持和同志們的热情鼓励与帮助。在再版初稿完成后，承鉄道科学研究院徐在庸同志詳細审閱，使本书减少了不少缺点；水利电力部办公厅图书編輯部为修訂再版征求意見和精心編輯及描繪插图，他們都付出了不少的劳动。其他許多水利机关、高等院校中的同志提来很多有益的建議和意見，我都尽量地考虑了，恕不能一一提出他們的名字。这些都是完成本书的重要因素，在此表示深切的謝意。

本书的主要对象为具有大专水平的水文水利計算工作者，也可以供大专院校水文专业学生和其他数理統計工作者作为参考。

限于我的理論与实踐水平，其中一定会有許多缺点和錯誤，悬切地希望同志們的批評和指正。

金光炎写于水利水电科学研究院

1963年2月，北京

# 目　录

# 第一章 緒　論

## §1-1 机率論和数理統計概說

**(一)机率論的对象**

机率論是一門研究偶然現象的規律性的数学学科。早在三百多年前，机率論已被物理学家和数学家們所注意。当初，机率論被用在賭博、保險事业和測量誤差等的問題中。随着科学技术的发展，特别在近几十年来，机率論在自然科学和生产实踐中的地位日益提高；現在，几乎沒有一門自然科学不在某种形式下应用机率論的方法了。

对于偶然現象的詳細說明，我們把它放到后面的章节中去。这里先举两个例子：

(1)同一距离用同一皮尺量测多次，所得結果彼此略有差异。这些差异是由于測量过程中受許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、起点和終点位置的不同以及視覚上的偏差等。

(2)給定相同的降雨强度和降雨时間，在同一块場地上进行多次人工降雨試驗，每次所得結果，彼此总有些不同：最大流量数値不同，集流时間不同，徑流总量也不同。这些差异是偶然性的，它受着許多因素的影响，例如由于下滲的不均匀、蒸发的不一样、噴雨装置的擺动、风的干扰和測量誤差的不同等。

上述例子說明在基本条件保持不变的条件下，多次实驗会得出不一致的結果，其原因是在基本条件之外还存在着許多次要因素，而这些因素或多或少地影响到实驗的最后成果。我們說这种結果上的差异是偶然性的差异。并称这种具有偶然性的差异的現象为偶然現象。

十分明显，宇宙中的任一实际事件都不可能不带有某些偶然成分。那末，无論如何精确地来固定实驗的条件和万分仔細地来观測实驗的結果，也不可能作到反复实驗的結果全部准确一致。因此，气象和水文現象的实测值也都带有偶然成分，习慣上我們常称这种实測値的系列为偶然系列。

經驗表明：由少数实驗所得的偶然系列往往是杂乱无章的，但随着实驗次数的加多，这时偶然系列就会呈現出一定的規律性。我們常称此种規律为統計規律。机率論就是研究偶然現象和偶然系列在大量观測条件下的規律性的科学。

目前，机率論已单独成为一門数学分支，它的数学定理能充分反映客观存在于宇宙間的大量偶然現象的規律性，并且在邏輯上是正确的和严謹的。

**(二)数理統計的概念**

"統計"和"数理統計"两詞通常所代表的意义不尽相同。

統計一語系由拉丁文中的"Status"导来，原本的含义是"情况"。最初出現这个名詞是

在十八世紀初期，直到十八世紀末叶，才开始流傳起来。当时，它所表达的意义十分狭窄，只是国家政治情况的一些說明。目前統計一語已有了广泛的含义。其內容一般可分为：資料的收集，所得資料的統計归納，統計观察方法的拟定和統計資料的分析。

数理統計的含义仅代表上述統計內容中的一部分，即用数学的理論和方法来对統計資料进行分析和計算。如是，数理統計必須主要以机率論为基础。同时，机率論也往往把由数理統計所揭露的事实提高到理論从而丰富自己的內容。

因此，机率論和数理統計是密切联系着的，通常我們总是把它們称呼在一起。但是它們所研究的內容，又不尽相同，其区别将在后面叙述。

在水文学中，常常把数理統計簡称为統計以及把水文学中的数理統計法称为水文統計法。

## §1-2 水文統計的理論性和現实性

### (一)水文現象的必然性和偶然性

水文現象和其他的自然現象一样，在它本身的发生、发展和演变过程中，包含着必然性的一面，也包含着偶然性的一面。促使水文現象发生的根本原因，規定着它的規律性，这种規律性按照一定不移的秩序貫串在全部发展过程中，致使水文現象具有了必然性的一面。例如大气运行的結果，必然会引起降水而产生徑流，以及水文情势以年为周期的循环性和明显的季节性等，这都充分表明了水文現象产生的必然性和不可避免性。可是，水文現象的发展过程，决不只由其根本原因所規定，另外还要受到周圍許許多多因素的或大或小的影响。这些影响的无限复杂性和多样性促使水文現象在演变过程中不断地发生各种程度的非根本性的偏差，使得水文現象的演变的固定秩序不能以純粹的形式出現，而是伴随着无数可有可无、可以这样也可以那样的情况而出現。例如一次暴雨后，必然会产生徑流，但徑流的形成，受着許多气象和自然地理因子的影响，致使我們无法用其固有的規律推知其实际出現的数量以及其在时間上和地区上的分布。这就說明了水文現象的偶然性和不确定性的一面。当然，在任何自然現象中，起主要和决定性作用的乃是必然性規律，偶然性只是起着从属的作用。

必然性和偶然性在水文現象的演变过程中不但始終同时存在着，而且还相互密切联系着。

必然性和偶然性是相互联系而相互又有区别的一对范疇，不能把它們看作是絕对对立的。如果，认为水文現象只能受必然性的規律所支配，那就会忽視其偶然性变化的这一面，而排除数理統計分析这一有效工具。另一方面，如果只看到水文現象偶然性的方面，忽視規律性的分析，必然会导致迷信数理統計和其相应的数字結果，而得出不切实际的結論。所以在水文分析中，必須作到把研究必然性規律的物理成因分析和研究偶然性規律的統計分析密切結合起来，相輔相成地解决实际問題。

### (二)或然推理的作用

或然推理也叫概然推理。水文分析中，凭借已有的实测資料来推估水文情势的变化規律，这是属于或然推理的范疇。

或然推理的作用，是能使某一推测在一定的証据下获得某种程度的可靠性或可能性。因其有一定的証据为前提，故不同于主观的和抽象的設想。所以說，根据已有的实测資料和已掌握的水文規律来推算未来的水文情势，虽則不能获得完全可靠的結論，但能得到有一定程度可靠性的結果。这是因为通过統計分析，可以从实际資料中分析出水文情势有限部分(样本)的变化規律，并用这些变化規律来推测整个演变过程(总体，包括未来的过程)的趋势。这种推测，是建立在样本和总体具有相互联系的基础上。但总体和样本又有所区別，故用样本不能完全推断出总体的眞实情况，只能推出其具有一定可靠程度的模态。这种模态的获得，就是利用了或然推理这一工具。在数理統計中，专門有一个分支来研究此类模态的形式和規律，即抽样誤差的理論和应用。

另一方面，或然推理还可以帮助我們获得新的知識。每当我們研究一个新的問題时，特别象水文学这門年青的学科，其中許多問題在开始探索时或在探索过程中，总是缺乏完备的理論和方法，又加实际資料的短少和实踐經驗的不足，因此必然要求我們去利用一切可能的方法和途徑来获得对新問題的认識。或然推理的应用，能帮助我們发現統計規律，从而逐步揭露出隐藏在它背面的必然性規律。当然，由于或然推理所获得的知識是推测性的，故用于水文分析中必須經过多方面的驗証。由或然推理所得的結果，在水文統計中往往是以机率、频率、保証率和破坏率等概念来表达的。

**(三)水文統計中引用假定的作用**

在进行或然推理时，大都要用到假定的形式。在水文統計中引用了許多假定，例如假定样本的規律服从于某种綫型分布和某种过程用某些机率定理等。

大家知道，人类的认識是一个不断发展的过程。人們对客观世界的认識，常常受到科学水平、实踐水平和思維能力水平的限制，不可能一下子达到全面的和正确的认識。因此，在一定阶段中，科学理論往往要从各方面去探索，而在探索的进程中，又往往要采取假定的形式。但假定的目的是为了寻求規律，因此这种假定不能是毫无根据的揣想，而必須以一定的事实为根据，然后再拿到实踐中去檢驗。水文統計中的假定，也是如此。例如别的一些学科中用正态分布来估算誤差甚为有效，而我們假定水文特征的誤差也服从于正态分布；又如有些部門的統計資料多系鈴形分布而选用了皮尔逊曲綫族，而水文系列也有符合鈴形分布的，故亦假定采用这种綫型，等等。此后，再把这些假定置于实踐中，經过多次分析比較，如发現所作假定大都能符合实际情况，則保留这个假定或再加以发展，直至假定成为事实。反之，如果假定与事实不符，則应抛弃原来的假定和重新作出假定，再反复进行檢驗。

显然，人們在掌握事实还不够充分或对这些事实的可靠性还不很清楚时，根据有限的材料，有时可以作出好几种假定，以致形成多种假定并存的局面。例如，經驗频率的选取方法，过去曾有过五六种之多；频率曲綫的綫型，也曾多至十几种，等等。在科学研究中，多种假定并存的局面，对发展科学是有利的，它可以通过不断的討論和不断的实踐，逐步加以淘汰和澄清。結果总会发現其中某一种假定較为合理而实用；或者取某几种假定之长而融合成一。水文統計方法的愈来愈精炼，正是充分地表明了这一点。

**(四)水利規划实施的可能性与現实性**

目前，水利規划中大都采用統計分析的数据。用統計分析来回答水文問題，都带有可能性的含义，例如均冠以频率(或机率)为百分之几或几年一遇等。这种带有可能性含义的水文数据，今后究竟会不会出現，若会出現則又在什么时候出現以及将来是否还会出現比它更大的，等等，就是牵涉到如何来理解可能性与現实性的問題了。

可能性与現实性是一对相互密切联系而又相互有区别的范疇。可能性是反映事物发展过程中这样的一种情况，即它在发展的一定条件下可以成为現实，但在目前却还沒有成为現实。如果它在发展的一定条件下，已成为現实，这就是可能性轉为現实性了。唯物辯証法告訴我們：人們的主观能动性，对于可能性轉变为現实性的过程具有重大的意义。

可能性必須建立在可靠的基础上，不能凭借任何无客观依据的主观設想，这样才能成为与現实性有联系的且具有轉化为現实性的客观根据的可能性。并不是任何抽象的和虚伪的可能性都有轉变为現实的可能。因此水文統計分析，必須建立在有代表性的和可靠的資料基础上，以及密切与成因分析和实际情况相結合。

但可能性还不就是現实性，要使事物由可能性变为現实，必須經过人的努力，有时甚至还要經过严重的斗爭。例如，用百年一遇的标准来設計某一水庫，意思是說这个标准能保証安全的可能性每年为99%或破坏的可能性每年为1%。那末它能否保証安全呢？我們可以回答說，它能保証安全的把握性甚大，因为安全与破坏的比率已达到99比1，即安全率远远地超过了破坏率。但这不是說我們就可倚凭如此优越的安全率而高枕无忧了。恰恰相反，要确保安全，必須提高警惕，防止一切可能出現的不利偶然事件的发生。故而在汛期来到之前，必須作好水利工程的汛前檢查和准备防汛力量，一方面防患于未然，另一方面也积极籌划抵御超过設計标准洪水的措施。不但在水利工程完成之后要如此，就是在規划設計时也应作出全面的考虑和安排。例如应預先考虑到超过設計标准的洪水来临时的措施，或安排有計划地引洪、滞洪，或加强各樞紐間的配合等，給管理运用或防汛机关作为指导。这种前前后后的籌划和措施，都是把确保安全的可能性轉变为現实性的强有力的条件。如果忽視了这一点，会使本来只要經过努力和斗爭就可以实現的事情而遭到失败。

因此，水利工作者必須正确理解可能性与現实性的辯証关系。如果只看到可能性的一面或只强調現实性的一面，就会把可能性与現实性絕对化起来，或者把設計标准的水文数据神秘化，或者片面地否定它，这些都是不科学的。

**(五)水文分析中应用数理統計法是現实可行的**

随着国民經济的日益发展，水利工程和其他許多建設(如桥涵、都市給水等)的規划設計和管理运用部門都迫切要求提供各种标准的水文数据。这种水文数据的获得，可以有不同的途徑，例如：用长期預报的方法，用“极限暴雨”或“极限洪水”的方法，也可以用数理統計的方法等。

大家知道，在目前的科学技术水平下，我們还不能对未来的降水或徑流作出长期定量的預报，甚至在定性預报上其可靠性和把握性也不大，故用长期預报的方法是比較困难的。

对于从气象成因来推求极限暴雨或再轉换成极限洪水，理論上是可行的，但問題是在极限暴雨的推求中，主要也需根据已有的短期实測資料。这样，用短期的数据来确定各項

要素指标的最大值，会具有很大的誤差。又加各項要素之間的相互关系及遭遇机会还不清楚，形成暴雨的一些理論也未完善。因而用极限暴雨的方法任意性較大，应用它也不合理想。

因此，充分分析現有实測水文資料的規律，正确地运用数理統計法，可以基本滿足水利工程和其他各种建設的需要。同时，数理統計法的分析計算工作比較簡单，能为广大水文和水利工作者所掌握，所以数理統計法是一个現实可行的方法。

但是，这不是說要解决复杂的水文問題，单凭数理統計法就可以了。数理統計法只有在正确运用时才有效。为了使成果更为合理起見，还必須作到四檢查三結合。

所謂四檢查，就是：

（1）檢查資料的一致性：数理統計法要求同一計算系列中的所有水文資料是屬于同一类型的和在同一条件下产生的，不能选取不同性质的水文資料。例如：雨成洪水和融雪洪水，它們的成因不相同；又如水庫的兴建、河道的整治以及各种水利和农田工程的修筑，使工程完成前后的水文資料也是不同性质的；故都不能把它們統計在一起进行統計分析。

（2）檢查資料的代表性：統計分析的目的是要把过去資料的变化規律推論于今后的情况，因此对資料代表性的审查必須相当細致。現在我国水文紀录一般仅有十多年至数十年，而最长的也仅有百年之譜，有的地方甚至只有几年的資料。这些短期的記載，很可能都是丰水年或枯水年的，其代表性不足，因此根据它們所推得的結果，就会偏大或偏小。代表性好的系列，必須包含各种丰水、平水和枯水年的資料。在实际工作中，往往要去实地調查、考查历史文献和比照邻近地区的水文資料，这样，才能得到富有代表性的系列。

（3）檢查資料的可靠性：每一水文数据本身可靠程度的檢查也相当重要，如系列中有几个数据不甚可靠，则必将影响最后結果的客观性和准确性。檢查資料可靠性的目的，就是必須把精度不高的和写錯伪造的水文数据全部找出来并加以修正。

（4）檢查資料的独立性：統計分析中要求同一系列中的資料是相互独立的，因此不能把彼此有关系的資料統計在一起。例如前后几天間的日流量都是由同一場暴雨所造成，彼此并不独立，故不能用連續的日流量来作为一个統計系列。又如相邻几站間的暴雨，也是由同一次暴雨所形成的，相互間有关系，故也不能把它們統計在一起。

所謂三結合，就是把数理統計法、物理成因分析法和地理分布法三者密切結合起来。

物理成因分析法是根据气象和水文現象所固有的成因規律来推求其演变过程和結果。但在成因分析中，由于影响因素十分复杂，相互間的关系又不甚稳定，最后結果往往不能以純粹的和理論的形式出現，所以单靠物理成因分析法是不可能很好解决水文問題的。同时，在物理成因分析法中，常常要用到数理統計中的相关分析法。因此它必須与其他两种方法相結合。

地理分布法是把各种水文特征表示为分区图和等值綫图的形式。这种方法中各种特征值的計算和分析也离不开数理統計法。

因此，三种方法不是相互对立而是相輔相成的。只有在三种方法密切結合之下，才有可能使水文計算的結果更为客观和合理。

## §1-3 水文統計的沿革

### (一)水文統計的雛形和初期的应用

水文統計法确切始于何时,現已难以查考。据福斯特說[4-13],約始于1880～1890年,由美国的赫斯切尔及富里曼首先应用历时曲綫(即目前称謂的頻率曲綫)于实际工作中。另据郝頓說[4-3],他在1896年将机率方法用于徑流研究中系由于腊富特的建議。但由于当时实測資料缺少,分析比較困难,他所采用的方法,大都为正态分布的运用和在对数紙上的图解配綫。

自此以后,机率方法应用愈来愈多。直至1913年,美国的富勒[4-1]和海森[4-4]相继发表了系統性的論文,闡述了机率方法的应用。富勒的論文中只利用重現期$T$和徑流$Q$在半对数紙上作点繪配綫,得到一些河流的綜合結果为$Q=1+a\lg T$,其中$a$为常数。海森则在正态分布的机率格紙(纵坐标为均匀分格)上配綫。当时海森发現[4-2],就是用这种机率格紙,配綫仍不滿意,因此他提出想用纵坐标为对数分格的正态分布的机率格紙,但那一年他还沒有給出实际的算例。

至1921年,海森才正式提出了用纵坐标为对数分格的正态分布机率格紙的图解配綫方法[4-6],这是对数正态分布的最早应用。同是这一年霍尔提到想用皮尔逊曲綫族来配合任何分布型式的水文資料[4-5]。不过,当时霍尔认为用这类曲綫族要解出其中的参数是十分麻煩的。

1924年,福斯特完整地介紹了皮尔逊Ⅲ型曲綫的分析方法[4-7],并制成离均系数$\Phi$值表,給实际工作带来許多方便。他的方法水文界一直沿用至今。同个时期,霍尔还建議使用两端有限的皮尔逊Ⅰ型曲綫[4-9]。

到了1935年,苏联学者克里茨基和閔凱里首先提出了組合頻率的近似分析法[2-15]。自此在水文統計中有了多元分布的內容。

从此以后,水文統計的內容慢慢地丰富和完整起来,逐漸地形成了水文水利計算中的一个独立体系。

### (二)水文統計应用前后的概况

在水文統計法提出之前和在应用水文統計法初期,水利工程的設計数据,大都以实測資料和調查洪水为依据。当时,水利措施多以防洪为目的,对于防洪标准的选取,有的采用調查的历史洪水,有的采用实測的最大洪水。但是,对关系到人民生命財产安全的防洪工程仅应用已出現过的洪水作为設計标准,犹恐不够安全,于是又提出在历史洪水或实測最大洪水上加一个安全系数。这样就有一連串的問題发生了,即对資料多的和資料少的系列、变化幅度大的和变化幅度小的系列、研究比較充分的和研究不甚充分的系列,应如何分别地加安全系数,要加多少,加了安全系数后的設計数据将来出現的可能性若何,等等。为了要解决这些問題,水文統計法也应运而生以及日益得到发展。

水文統計法首先以实际資料为基础,应用了或然推理的原理,充分利用了假定的作用,基本上解决了上述所不能回答的問題。它是根据实际資料的变化規律(这些規律通常是以經驗頻率、均值、离差系数和偏差系数等特征来概括和表达的),利用頻率曲綫的模

型和考虑系列长短可能产生的抽样誤差，并結合成因分析和地区上的綜合平衡，以推测不同机遇的安全系数。这样，水文統計法就会自动反映出：系列水平高的、变化幅度大的、資料項数少的，安全系数就大一些；反之，安全系数就小一些。如此，采取安全系数的大小，就有了一个較为客观的尺度，而且也为統一标准創造了有利的条件。

这里所談到的安全系数，应从广义方面来理解，也就是說加了抽样誤差后的結果和頻率曲綫"外延"后的値，都可以看作是加了安全系数后的結果。由此可知，用水文統計法来推求設計数据，是建筑在一定科学基础上的，只要人們客观地去分析水文現象的規律，是能获得客观的結果的。

**(三)我国水文統計研究的簡况**

我国学者对水文統計进行的最早研究，并有正式文献可以查考者，首推1933年周镇倫先生的著作[1-1]。周先生原拟收集我国各处的雨量紀录以进行研究，但因收集資料困难，雨量紀录年份不长，而借用了美国的雨量資料进行研究。当时他只应用了正态分布和皮尔逊Ⅲ型曲綫。

1937年，陈椿庭先生著述論文[1-2]，把我国五大河流(长江、黄河、永定河、涇河和淮河)的洪水流量用对数正态曲綫和皮尔逊Ⅲ型曲綫进行頻率計算。之后，在我国的早期水利书籍(如張含英：防洪工程学，等)中就出現了頻率的概念，并轉載了陈先生的論文。

解放以后，我国水利事业得到蓬勃的发展，在学习苏联經驗的同时，水文統計法也得到了广泛的应用。各大河流和各地区的水利机构，在規划設計时，均有效地运用了水文統計法，并在实际工作中丰富了和发展了水文統計的方法和內容。

1955年3月以后，北京水利科学研究院水文研究所集合了某些水利单位的人員，組織了洪水頻率研究組，对頻率方面的理論和方法进行了学习和研究，并搜集了全国各地的水文資料进行分析驗証，于1956年提出了初步总結报告，并于1957年修訂成为《暴雨及洪水頻率計算方法的研究(初稿)》[1-10]，这是我国第一篇比較全面的叙述頻率問題的报告。

自1956年冬我国第一屆全国水文計算学术討論会关于頻率問題的討論和1958～1959年由长江流域規划办公室組織的頻率問題的討論以来，我們对水文統計的认識更有所提高。可以預料，随着水利事业的发展，水文統計在我国将会有新的发展。

# 第二章　机率論的基本知識

## §2-1　机率的概念

**(一)事件，事件的种类**

“事件”是机率論中最基本的概念，它是指在一定的条件組合下，在試驗的結果中所有可能出現或可能不出現的事情。

事件可以分为三类：

1.必然事件

如果在条件組每次实現之下，某一事件在試驗中不可避免地要发生，我們称此事件为此試驗的必然事件。

〔例1〕 水在760毫米大气压力之下，加热到100°C时，則化为蒸汽。也就是說，在这两个条件(条件組)每次实現之下，水化为蒸汽是必然的。

2.不可能事件

如果在条件組每次实現之下，某一事件在試驗中永远也不会发生，則称此事件为此試驗的不可能事件。

〔例2〕 天然河流如上游无人为阻水，則当洪水来临时，必然涨水，发生断流是不可能事件。

3.偶然事件(隨机事件、或然事件)

如果在条件組每次实現之下，若某一事件在試驗中可以发生也可以不发生，則称此事件为偶然事件。在实际問題中，往往在一定条件組每次实現之下，出現的事件不只有一种，而是或出現事件$A$，或出現事件$B$，或出現事件$C$等。这种情形均称为偶然事件。

〔例3〕 一粒骰子投擲一次，或出現1点，或出現2点，……，或出現6点，共有六种可能。此时，出現某一点数都是偶然事件。

〔例4〕 汛期內，每年在河中必然会发生一次最大的流量，对这种現象的本身来說完全不是偶然的。但在每个年份里，这个最大流量可能大于所指定的流量$Q$，也可能小于这个流量。因此，我們說該河中每年最大流量在数量上的出現是偶然事件。

水文現象在量的出現方面都属于偶然事件，故都可用偶然事件的方法来处理。

为了簡便起見，下面用字母$A$、$B$、$C$、……来作为事件的代号。由于在討論两个或更多个事件时，要表达这些事件間的关系若用文字来說明可能很繁长，因此我們引入一些記号来表示它們之間的关系。

事件“$A+B$”是表示在事件$A$和事件$B$中至少发生其中一个的事件。同样地，事件“$A_1+A_2+\cdots\cdots+A_n$”表示在事件$A_1$、$A_2$、……、$A_n$中至少发生其中一个的事件。

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$B$永远不可能同时发生，則称事件$A$和事件$B$互斥(互相排斥)或互不相容。同样，事件“$A_1A_2\cdots\cdots A_n$”表示

事件$A_1$、$A_2$、……、$A_n$均同时发生的事件。如果事件$A_1$、$A_2$、……、$A_n$中任意两个均为互斥，则称事件$A_1$、$A_2$、……、$A_n$彼此互斥。

〔例5〕 錢币有正反两面，在一次投擲中出現正面設为事件$A$及出現反面为事件$B$。再把在第一次投擲时出現正面或反面的事件，写成$A_1$及$B_1$，第二次投擲为$A_2$及$B_2$，第三次投擲为$A_3$及$B_3$。則事件"$A_1+A_2+A_3$"表示在連續三次投擲中，至少有一次出現正面的事件。事件"$B_1B_2B_3$"表示在連續三次投擲中都出現反面的事件。显然，事件"$A_1+B_1$"、"$A_2+B_2$"和"$A_3+B_3$"都是必然事件。

事件"$A_1B_1$"表示在第一次投擲中，既要它出現正面，而同时又要它出現反面，显然这是一个不可能事件。因此$A_1$和$B_1$为互斥；同样，$A_2$和$B_2$、$A_3$和$B_3$也均为互斥。

互斥这个概念，在机率論中甚为重要，希望讀者能淸楚掌握它的意义。

最后，我們再引进两个概念：

如果事件"$A+B$"为必然事件，而且事件$A$、$B$互斥，那末我們称$B$为$A$的补事件（或$A$为$B$的补事件）。此时，我們用$\bar{A}$来代表$B$（或用$\bar{B}$来代表$A$）。

$n$个事件$C_1$、$C_2$、……、$C_n$如果滿足下列两个条件，則称此$n$个事件为完全事件系：

（1）$n$个事件中任何两个事件是互斥的；

（2）事件"$C_1+C_2+\cdots\cdots+C_n$"是必然事件。

〔例6〕 在例5中擲出正面$A$和反面$B$互为补事件。

〔例7〕 投擲骰子，出現1点記为$C_1$，出現2点記为$C_2$，……，出現6点記为$C_6$。則此六个事件为完全事件系，而$C_1$是事件"$C_2+C_3+C_4+C_5+C_6$"的补事件。

**（二）机率的意义**

机率这个术語，也有称为：槪率、或然率和可能率等，它表示偶然事件可能发生的机遇。

每一事件都有某种程度的可能性：有的可能性大些，有的可能性小些。如果用数值来表示事件可能性的大小，則称該事件出現可能性的数值測度为事件的机率。在机率論中，我們不用"可能性大小"这句話，而是用"机率"这个术語来表示的。

大家知道，任一事件发生的可能性，最大也不会大到超过必然事件这一界限；因此，我們把必然事件作为机率測度上的上限，并有必然事件的机率等于1。同样，任一事件发生的可能性，以不可能事件为下限，故不可能事件的机率等于零。这样，我們就有了事件机率測度的范圍，即自0至1；而偶然事件的机率，总是处在0与1的范圍內。

在科学研究、日常工作和生活中，了解机率的大小，具有实际的意义。例如，甲乙两厂都生产同样的工业品，如果甲厂的次品率約为1%，而乙厂的次品率約为3%；那末我們可以說甲厂的技术水平一定比乙厂高，当你想买这种产品时，自然而然地会去向次品率低的甲厂訂购。

有了衡量机率大小的尺度，不但利于各个事件間可能性大小的比較，而且也利于单独或复合事件机率的計算；并且它在一定程度上可以作为我們行动的指南。

**（三）机率的直接計算，古典机率公式**

有許許多多的試驗，它的各种結果是具有对称性的，而由于这种对称性，可知各种結果在客观上具有同等的可能性。例如，擲一个质量均匀的錢币，我們沒有理由說，出現正

面的可能会比出現反面的可能大一些或者小一些。由于必然事件的机率等于1，我們就有权假定，擲錢币出現正面或出現反面的机率都等于$\frac{1}{2}$。再如擲一顆骰子，这顆骰子是一个結构勻称的立方体，各面刻着1至6点；因为这个立方体具有对称性，所以試驗的六种結果都是同等可能的，因而在投擲中每种結果的机率均等于$\frac{1}{6}$。

对于結果为对称且同等可能的一切試驗，都可以应用这样的方法来处理，这种方法叫做机率的直接計算。根据这种基本的概型，我們可以列出計算等可能事件的公式。

設某一試驗共有$n$种不同的可能結果，其中各个結果都是具有对称性，即等可能的；如以$m$代表事件$A$出現的次数，則出現事件$A$的机率为：

$$P\{A\}=\frac{m}{n} \tag{2-1}$$

这个公式叫做"古典机率公式"。当$m=n$时，得$P\{A\}=1$，此时$A$即为必然事件。当$m=0$时，得$P\{A\}=0$，此时$A$是不可能事件。由于$m$总是介于0与$n$之間，故$P\{A\}$一定介于0与1之間。这就是上面我們把机率测度范圍定在0与1之間的依据。

〔例8〕 罐中有白球六枚、紅球四枚、黑球二枚，如果这十二个球摸起来使人感到完全相同，其差别只在顏色，問任取四球为全白的机率若何？

〔解〕 此題应先算出在全部球(即12枚)中取4球的一切可能情形，即有$C_4^{12}$种；其次計算6个白球中选取4个的組合数，即$C_4^6$。則出現4个白球的机率为：

$$P\{\text{出現 4 白球}\}=\frac{C_4^6}{C_4^{12}}=\frac{1}{33}$$

〔例9〕 在52張扑克牌中，任取13張，其得不同号碼的机率是多少？

〔解〕 52張中任取13張的方式有$C_{13}^{52}$种，由于扑克牌中每种号碼有4張，故取得13張为不同号碼的方式有$4^{13}$种，則有

$$P\{\text{得13張不同号碼的牌}\}=\frac{4^{13}}{C_{13}^{52}}=0.000106$$

意思是說，多次玩牌游戏中，这种結果平均1万次中約可有1次。据玩牌者的經驗，13張不同号碼的牌是偶而能获得的，故万分之一机率的事件也偶而能出現。

試驗結果的对称性，通常只能在組織得很巧妙的試驗中見到，故式(2-1)只能用在一些比較簡单的問題中。至于对不呈对称結果的試驗或者比較复杂的問題，我們有时要設法把它們变到可以适用于这种方法的型式。

**(四)机率的性质**

从机率的意义出发，我們可以为机率归結出下面四个性质：

(1)必然事件的机率等于1，或

$$P\{\text{必然事件}\}=1 \tag{2-2}$$

(2)不可能事件的机率等于零，或

$$P\{\text{不可能事件}\}=0 \tag{2-3}$$

(3)机率为非負数但不大于1，即

$$0\leqslant P\{A\}\leqslant 1 \tag{2-4}$$

(4)如果事件$A$、$B$互斥，則

$$P\{A+B\}=P\{A\}+P\{B\} \tag{2-5}$$

头三个性质在前面都已經讲过了，現在只来証明第四个性质。

設一試驗共有 $n$ 种可能的結果，其中有利于事件$A$的有 $m_1$ 种及有利于事件 $B$ 的有$m_2$种，显然：

$$P\{A\}=\frac{m_1}{n},\qquad P\{B\}=\frac{m_2}{n}$$

因$A$、$B$互斥，故有利于$A$或有利于$B$的結果共有 $m_1+m_2$ 种，則

$$P\{A+B\}=\frac{m_1+m_2}{n}$$

由于

$$\frac{m_1+m_2}{n}=\frac{m_1}{n}+\frac{m_2}{n}$$

故公式(2-5)得以証实。

同理，可以推論到 $n$ 个事件的情况，即

$$P\{A_1+A_2+\cdots\cdots+A_n\}=P\{A_1\}+P\{A_2\}+\cdots\cdots+P\{A_n\} \tag{2-6}$$

如果$A_1$、$A_2$、……、$A_n$是一个完全事件系，則有

$$P\{A_1\}+P\{A_2\}+\cdots\cdots+P\{A_n\}=1$$

再如这 $n$ 个事件出現的机率均相等，即

$$P\{A_1\}=P\{A_2\}=\cdots\cdots=P\{A_n\}$$

就有

$$nP\{A_n\}=1 \quad 或 \quad P\{A_n\}=\frac{1}{n}$$

**(五)事件的频率**

上面所列出的古典机率公式(2-1)，只能适用于一种試驗，就是在試驗中每种可能結果的发生都具备对称性的条件。但在許多的实际問題中，远不是所有試驗都严格地符合这个条件，这时机率当然不能再用公式(2-1)来計算了。例如一顆制作得不甚对称的骰子，某一面的出現机率不再是$\frac{1}{6}$；但是，这一面的出現仍然是有某种机率的，其机率可能比$\frac{1}{6}$多一些，也可能比$\frac{1}{6}$少一些，这个机率表現了在多次投擲中平均出現的頻繁程度。水文現象中这种例子很多，例如日降雨量大于 100 毫米、汛期最高水位超过警戒水位、河中断流等事件都不能用古典机率公式来計算。但是，可以肯定这些事件都有它們本身客观的可能性程度。这种可能性程度原則上都可用数量来度量，它在同类的試驗中，可以通过它所对应的頻率反映出来。

設做了一系列的試驗，每次試驗的結果，事件$A$都可以出現或不出現。我們把 $A$ 出現的次数和試驗总次数的比叫做事件$A$在这系列試驗上的頻率。如以$P^*\{A\}$表示 $A$ 的頻率，若总共試驗了$n$次，而$A$出現了$m$次，則可按下式計算事件$A$的頻率：

$$P^*\{A\}=\frac{m}{n} \tag{2-7}$$

經驗証明：对于次数不多的試驗，事件的頻率有着明显的偶然性。例如把一个錢币擲五次，完全有可能只出現正面一次(其頻率为0.2)；在另擲五次中，又有可能出現正面四次(其頻率为0.8)。但在試驗次数加多之后，事件的頻率就大大失去了它的偶然性，而在某一定范圍之内摆动；試驗的次数愈多，頻率也就愈接近于一个常量。对于擲錢試驗，在理

論上讲出現正面和反面的頻率应各为0.5，不妨看看前人的实际試驗，其結果如表2-1。可見頻率在0.5左右作微小的摆动，这和理論上的頻率值十分接近。

表2-1

| 試驗者 | 擲錢币次数 | 出現正面次数 | 頻率 |
|---|---|---|---|
| 蒲丰 | 4040 | 2048 | 0.5080 |
| 皮尔逊 | 12000 | 6019 | 0.5016 |
| 皮尔逊 | 24000 | 12012 | 0.5005 |

此种在試驗次数增多后頻率趋于稳定性的性质，不断地为試驗和人类的实踐活动所証实。显然，这是表达了偶然現象的一种規律性。这个規律性的严格的理論証明首先由貝努里完成，即貝努里定理(見§4-9)。貝努里証明了，在試驗次数很多很多时，事件的頻率和其机率是十分接近的。

在理論上和实踐上給出頻率和机率間有机联系这一点，具有巨大的实际意义。当我們无法求得复杂事件的机率时，就可以通过多次試驗，把事件的頻率作为事件的机率的近似值。

我們也常常这样来理解机率和頻率的含义：机率是偶然事件在客观上可能出現的程度，是一个常量；頻率是經驗值，随着試驗次数的增多而接近于机率值。在許多实际問題中，頻率总是一个具体的数字；而机率似乎是抽象的理論值，它只能凭頻率值来估計。

数理統計中，还有称为事先机率和經驗机率的。对于一些簡单事件，可以事先根据事件的对称性，不必經过試驗，即能得到机率值的称为事先机率，如擲錢得正面的机率为$\frac{1}{2}$者就是。对于不可能事先求得机率的事件，把用多次試驗的頻率推估出的机率称为經驗机率或事后机率。水文分析中，无法知道事先机率，都是凭借过去已发生的事件用經驗机率或頻率去推测未来可能发生的情况。

頻率与机率的含义已如上述，它們之間有接近的地方，但又有严格的区别，一般在机率論中研究的对象为机率，而在数理統計学中討論的对象为頻率。

## §2-2 机率的主要定理

**(一)条件机率**

在讲条件机率之前，有必要先談一談独立和相关的概念。

如果事件$A$的发生不会影响事件$B$出現的机率，则称事件$A$对事件$B$独立。

〔**例1**〕 擲两顆骰子，設事件$A$为第一顆骰子出現1点，事件$B$为第二顆骰子出現1点。显然事件$A$的出現与事件$B$的出現与否无关，事件$B$的出現与事件$A$的出現也无关。在这时，我們称它們为彼此独立。

有时，要知道事件間是否独立，只需用經驗来判断就行了。例如：珠江流域的降雨和同一天黑龙江上的徑流；第一次擲錢得正面和第二次擲錢得正面，等等。

如果事件$A$的发生影响到事件$B$出現的机率，或者說事件$B$的机率将随事件$A$的出現