高等学校教材

随机水文学

四川联合大学 丁 晶
河 海 大 学 刘权授
合编

高 等 学 校 教 材

随 机 水 文 学

四川联合大学 丁　晶
河 海 大 学 刘权授
合编

中国水利水电出版社

内 容 提 要

本书较系统地介绍了随机水文学的基本原理、分析方法和计算模型，重点讲述如何应用随机模型和模拟技术解决工程水文计算和水文预测问题。

本书对水文、水资源专业的工程技术人员及相关专业的科技工作者有较大的参考价值。同时，本书可作为水文、水资源专业的研究生教材，其中的基本内容可供水文、水资源和水环境等专业的本科生使用。

图书在版编目（CIP）数据

随机水文学/丁晶，刘权授合编．—北京：中国水利水电出版社，1997

高等学校教材

ISBN 7-80124-356-0

Ⅰ．随…　Ⅱ．①丁…②刘…　Ⅲ．随机水文学-高等学校-教材　Ⅳ．P333.6

中国版本图书馆 CIP 数据核字（97）第 01251 号

书　名	高等学校教材·随机水文学
作　者	四川联合大学 丁 晶 河 海 大 学 刘权授　合编
出　版	中国水利水电出版社（北京市三里河路 6 号　100044）
发　行	新华书店北京发行所
经　售	全国各地新华书店
排　版	中国水利水电出版社微机排版中心
印　刷	北京市朝阳区小红门印刷厂
规　格	787×1092 毫米　16 开本　9.75 印张　224 千字
版　次	1997 年 10 月第一版　1997 年 10 月北京第一次印刷
印　数	0001—1000 册
定　价	9.80 元

编 者 的 话

本书是按照《一九九〇～一九九五年高等学校水利水电类专业本科生、研究生教材选题和编审出版规划》的规定进行编写的。

随机水文学是把随机过程的理论与方法引入水文学而逐渐形成的一门新学科。它以现实水文过程为研究对象，建立能够反映水文现象随机变化特性的数学模型，并通过模型模拟出的大量水文序列来满足水利工程规划、设计和运行的各种需要。本书共分八章，系统地介绍随机水文学的基本原理、分析方法和计算模型，重点讲述随机模型和模拟技术在水利工程中的实际应用。各章均列有一定数量的习题，书末列有主要的参考文献。

本书第一章由刘权授、丁晶共同编写，第二、五、六章由刘权授编写，第三、四、七、八章由丁晶编写。全书由长江水利委员会水文局季学武高级工程师主审，由丁晶修改定稿。

本书有些材料引自有关院校、生产和科研单位编写的教材和技术资料以及个人发表的论文，编者在此谨致以诚挚的谢意。

由于编者水平所限，书中定有不少的缺点和错误，敬请读者批评指正。

编 者

1997 年 1 月

目　录

第一章 绪　　论

本章介绍随机水文学的概念、基本方法、一般用途和发展趋势。

第一节 随机水文学

现实世界存在着两类现象：确定性现象与随机性现象。相应地有两大学科：确定性学科与随机性学科。随机水文学是从随机性角度研究水文现象随机变化规律的学科。

水文过程显示着水文现象随时间而变化的特性。由于水文现象受众多因素的影响，水文过程呈现出随机性，常称为随机水文过程。随着各种随机过程理论和时间序列分析技术不断被引入水文学中，研究随机水文过程的领域逐渐形成了一门新的学科，这就是随机水文学。随机水文学是指以随机水文过程为主要研究对象的一门学科。与随机水文过程相应的是随机水文序列，前者指的是连续过程，后者指的是离散过程。它们在本质上没有什么区别，而在实用上，通常只研究随机水文序列。一个水文序列是依一定的规律排列的。因此，在随机水文学中，水文现象发生的时间先后次序至关重要。

在全面随机分析的基础上对随机水文序列建立起反映水文现象主要变化特性的随机水文模型。根据建立的模型，既可模拟大量水文序列，也可作统计预测，以满足水利工程规划、设计、运行及各种研究的需要。这就是随机水文学的主要任务。本课程重点讲述如何依据观测到的样本序列建立随机水文模型，如何由模型模拟大量序列，如何将模拟序列应用于水利工程的规划和设计。说得具体一些，就是对随机水文现象建立合适的随机水文模型，模拟出序列（例如：年、月、日径流序列，洪水序列等），依据这些序列并结合工程特性和要求进行水利计算（径流调节、洪水调节等），最后计算出指定频率下的各种水利、水能特征值：保证出力、年电能、水库供水流量、水库坝前洪水位等，供全面分析之用。

由此可以看出，随机水文学的主要任务和水文计算的任务在很大程度上是一样的，但它们之间也有明显的差异。为了加深对随机水文学特点的理解，下面围绕随机水文学方法和水文计算方法的对比来做全面分析。

第二节 随机水文学的方法

现行水文计算提供工程规划、设计所需要的水文序列，其计算方法有两种。

一、实测序列法

以实测年、月径流序列作为计算序列，通过径流调节，直接得到水利指标（供水流量、保证出力，年电能……）的频率曲线，然后取其平均值（多年平均电能）或指定频率的特征值（保证出力、保证流量……）作为工程设计的依据。

这种方法的优点是：能广泛适应各种工程（单一或系统工程、发电、灌溉、供水工

程）兴利计算需要，而且由水文序列估计各种水利指标时概念明确、计算直观。该法的缺点是：几十年的实测序列只是一个历史样本，未来出现的序列不可能与过去实测序列一样；而且，现有科学水平也不能确定性地预测将会出现怎样的具体序列；只能进行的概率性预测，其可能性是多种多样的。因此，依据已有的实测序列计算出的水文特征或水利指标只是未来可能出现的一种情况，以此进行工程设计会有抽样误差，有时误差可能很大，而实测序列法无法回答这个误差问题。

以实测序列方法进行设计的新安江水库所出现的问题生动地说明了这一点。新安江水库是一座大型水库，初设时的水文计算是依据罗桐埠站1930～1956年实测与插补的年、月径流序列进行的，这27年的多年平均流量为360 m^3/s。1958年以后，新安江流域气候条件异常，出现了枯水年组。1958～1968年的多年平均流量为260 m^3/s，较原设计值低27.8%，发电量则较原设计值低得很多。地处我国北方的黄河三门峡段和松花江哈尔滨段也发生过这种现象。这说明实测序列法作为工程规划、设计的依据是存在严重缺点的。

实测序列法除存在上述严重缺点外，它的另一个缺点是难以用于防洪计算。实测序列一般较短，由实测洪水序列通过洪水调节得到的坝前洪水水位序列也相应较短，以较短序列大幅度外延推求工程设计所需要的、稀遇的设计最高洪水位是很困难的（会出现很大的抽样误差）。为了克服这一缺点，现行水文计算在作防洪计算时，以设计洪水过程线法代替实测序列法。

二、设计洪水过程线法

这种方法是依据观测的洪水资料，通过分析计算提供设计所需要的一套洪水序列，习惯上称之为设计洪水过程线法，其要点是：

（1）根据各种时段洪量频率曲线估计指定设计频率的时段洪量。

（2）选取洪水典型过程线与空间分配形式，用同频率或同倍比放大法将对前述设计频率洪量进行缩放，得到各种设计洪水过程线。

（3）利用上述设计洪水过程线进行水库防洪调节计算，其最高洪水位即为指定设计频率的水库设计洪水位。

从上述计算过程可以看出，这个方法有一个重要假定，那就是最高洪水位的频率等同于洪量的频率。

这个假定颇有问题，同样的时段洪量、时间过程与空间分配可以相差很大，调节计算后的水库最高洪水位也可相差很大，而且选取时空分布典型的主观任意性也颇大。此外，这种方法也不能给出设计水利指标（如水库设计洪水位）的抽样误差。

综上所述，不管是实测序列法还是设计洪水过程线法都存在不少缺点，而且工程情况的愈复杂，这种缺点就会愈严重。为了弥补现行水文计算的缺陷，克服它存在的严重缺点，适应当代复杂工程规划、设计的要求，提高设计可靠性，人们提供工程需要的水文序列时采用随机水文学方法。这种方法的要点为：

（1）应用实测水文序列建立具体反映时间（年、月、洪水）过程与空间（一个或多个工程地点）过程的水文随机模型。

（2）应用所建立的随机模型模拟一个或多个工程地点未来可能出现的各种水文序列（模拟序列）。

（3）应用大量的模拟序列，可适应工程需要进行径流或洪水调节，得到水利指标的频率分布曲线或特征值供设计之用。

随机水文学方法和现行水文计算法相比，有以下显著特点：

（1）随机水文模型全面表征水文现象统计变化的特性，不同的模型表征水文现象变化特性的重点有所差异。有重点表征水文现象随时间变化的模型（时间模型），有重点表征水文现象随空间变化的模型（空间模型），有既表征时间变化又表征空间变化的模型（时空模型）。一般而言，随机水文模型将水文现象在时间和空间上的变化有机地结合在一起．即它综合表征水文现象的时空统计变化特性。模拟序列出自模型。大量的模拟序列以直观的方式并和模型等价地表征水文现象的时空统计变化特性。因此，根据模拟序列计算各种水利指标时，就科学地统一考虑水文现象在时空上的变化特性。这样，随机水文学的方法克服了设计洪水过程线法将完整的洪水现象人为分割成洪量、时间过程和空间分配三个方面而分别孤立考虑和分析计算的弊病。特别值得一提的是，该法克服了“最高洪水位的频率等同于洪量频率”这个不符合实际的假定所造成的严重缺点。

（2）由随机水文学方法能模拟出大量的洪水序列。根据工程特性，模拟序列通过调洪计算即可得到相应的大量水利指标序列（如坝前洪水位等）。依据长指标序列可以既方便又合理地获得水利指标频率曲线（例如坝前洪水位频率线）和各种特征值，以用于工程的规划和设计。

（3）大量的模拟序列表征着未来水文现象可能出现的各种情况。正如前面提出的，实测序列法以短期实测序列为依据，而短期的实测序列只能表征未来水文现象的一种可能情况。显然，在工程设计时不能只考虑一种情况，而必须考虑工程运行期内可能出现的各种情况，并据此对水利指标的抽样误差作出估计，使设计更加合理、可靠。因此，随机水文学方法是对实测序列法的重大改进。在实际应用中，它有许多优越之处。

总之，随机水文学方法是在水文计算法基础上发展起来的一种先进方法。它一方面对现行水文计算法所存在的问题进行了重大的改进；另一方面，它更全面、客观，适应性也更强。所谓全面是指这种方法不仅可以提供单站（单点）各种特征水文量的模拟序列（年、月、旬、径流、洪水、雨量等），而且可以提供多站（多点）的序列；所谓客观是说该方法有一定的原则和规则可供遵循，大大减少了计算过程中的主观性；所谓适应性强是指该法能适应各种工程的规划、设计和管理的需要。

随机水文学方法的关键在于如何根据样本序列建立一个适用的随机模型。这就是本书重点讲述的内容。

第三节　随机水文学的应用

从上节介绍的随机水文学的主要任务可以看出，建立随机水文模型和随机模拟是关键之所在。随机水文模型及其随机模拟在水文水利计算、水文预报、水文测验及其他方面均得到广泛的应用。本节将简略介绍这些方面的应用。

一、在水文水利计算方面

1．在系统分析计算中的应用

水文要素（如年径流量等）可按其统计特性建立相应的随机模型；通过随机模型，借助统计试验方法可获得大量的随机模拟序列（以下简称为模拟序列）；在系统分析中，以模拟序列作为输入，根据系统的特性和设计要求进行各种计算，从而得出系统响应，即输出。显然，作为输入的模拟序列（如年径流量）是系统分析的基础，而合理可靠的模拟序列必须建立在合理可靠的随机模型的基础之上。

因此，随机模型在各种系统分析中占有重要的地位。例如，黄河上游建有刘家峡、盐锅峡、八盘峡、青铜峡和龙羊峡等水利工程。对于这样复杂的系统，在进行各种分析时（如水库调度分析），作为系统输入的多站模拟序列是必不可少的。较短的实测序列用于这样的系统分析存在着一些难以解决的问题，比如连续枯水段考虑的程度、防洪和兴利库容的有效结合、调度图的合理绘制、破坏深度及其相应概率的估计等。大量模拟序列的应用在一定程度上可以解决这一类问题。又如，在对四川省沱江进行规划，特别是进行沱江的环境保护系统分析时，要求以枯水流量序列作为系统输入。对枯水流量建立适当的随机模型，进而获得大量模拟序列，可以满足系统分析的要求。

为了研究三峡防洪系统对长江中、下游防洪效益的影响，需要预估进库的洪水过程以及中、下游各防洪控测点的洪水过程。如何预估未来可能出现的各种洪水过程是一个难题。建立多站洪水随机模型模拟出各站的洪水过程可以满足三峡防洪系统的研究需要。

2. 在处理一些特殊问题中的应用

在生产实际中，常常要研究某些水文特征量的统计特性，而这些统计特性很难用概率理论通过分析的方法来获得。在这种情况下，利用随机模型并借助统计试验法加以探讨，虽然得不到精确解，但求得的近似解可以满足生产实际的要求。例如，在研究洪水地区组成时，各控制断面洪水特征量的统计特性及其相互关系可利用随机模型给出的各断面模拟序列加以推求。随机解集模型已被用于嘉陵江北碚站洪水地区组成的研究中，并已得到现行的典型法和同频率法难以得到的成果。又如，通过随机模型得到的模拟序列来研究干旱持续历时特性。通过随机模型还可以研究人类活动对径流的影响。以上仅为几个例子，实际上随机模型结合统计试验法在解决一些特殊问题上已有多方面的应用。

二、在水文预报方面

1. 预报误差的处理

用传统预报模型作出的预报被认为是预报的第一步。第二步则是对第一步的预报结果进行调整，从而获得第二步的预报结果。构造第二步预报的依据在于预报误差的统计结构特征。若预报误差序列是相依的，那么就可利用误差的自相关特性建立合适的随机模型来预报将来的误差，从而提供第二步预报。因此，处理预报误差的实质是寻求合适的误差随机模型。一旦获得这种随机模型，即可在作业预报中处理误差，以提高预报精度。

2. 随机模型预报

对预报变量直接建立随机模型，即可进行预报。这种类型的统计预报一般给出条件期望值，另附以一定显著性水平的置信限。例如，长江汉口站的各月水位曾用季节性自回归滑动平均求和模型进行预报。又如，用季节性自回归模型预报黄河主要站的过渡期径流量。

3. 卡尔曼滤波实时预报

自 1960 年卡尔曼（Kalman）创立了卡尔曼滤波理论以后，这一理论便在自动控制方面

得到应用，并于70年代初被引入水文学。最近10年，卡尔曼滤波理论在水文预报上的应用有所发展，预报精度有了一定程度的提高。卡尔曼滤波实时预报是基于系统所建立的数学模型。就广义而言，这些模型属于随机模型的范畴。显然，只有系统受到随机因素的影响才需应用滤波作实时预报。因此，随机理论的应用和随机模型的建立是卡尔曼滤 波实时预报的关键之一。从这个意义上说，随机模型在近代预报中同样占有重要的地位。

三、在水文实验和站网布设方面

近年来，在考虑测验误差和评判测验精度时，用到了随机过程理论。例如，流速及含沙量等要素的脉动现象可被看作是一种随机过程。在水文测验中，要连续估计几次测量值的均值误差。为了正确地估计均值误差，必须考虑单项测次的误差过程，即要考虑各测次误差之间的相关性。又如，在一个流域中，点雨量的观测可被看作是随机场上的抽样。各点的测量值和流域平均值的误差之间可能存在着相关关系。在计算流域平均值时，必须计及这种相关关系。有人利用随机过程理论推导出测验过程中出现故障的次数和资料缺测长度的随机模型，并利用这一模型来改进测验方法。最近，随机模拟技术被用来测算地区上的某种水文要素。例如，为了测算森林区平均积雪深度，可通过小面积的抽样测量来建立一种适用的随机模型，并借助随机模拟估算出大面积上的平均积雪深度。

近10年来，随机分析技术日益应用于站网规划与布设方面。例如，利用雨量的空间相关结构设计雨量站网。布若斯（Bras）等人把流域上的点雨量看作多维随机过程，提出对雨量站布设的要求和方法。莫斯（Moss）等人将回归分析模拟用于地表水站网的设计。鲁海宁（Rouhani）建立了一种所谓降低方差分析的技术，用在随机场内设计最佳搜集资料的方案。

由以上可以看出，随机分析技术和随机模型在测验和站网规划方面的应用也是不可忽视的，并且随着资料的积累，先进技术的引入，加之计算机的普及，这种应用会愈来愈受到重视。

第四节 随机水文学的发展

随机水文学作为一门比较完整的学科出现，虽是最近30多年的事情，但它的基本概念早在本世纪初就被引入水文学中。海森和塞德劳的早期研究就表明，可以利用概率和统计理论来分析径流序列。50年代初期，赫斯特研究了径流和其他地球物理现象的长期实测序列。他的研究对随机水文学的发展产生了很大的影响。在海森和塞德劳早期经验性研究的基础上，巴尔纳于1954年应用正态随机数表来模拟径流。然而，随机水文学的迅速发展是从塞姆斯、费营 和叶菲耶维奇考虑了水文现象时序上客观存在着的相依性，引用马尔柯夫模型（自回归模型）才开始的。尽管这一模型在60年代初期才用来模拟年、月径流序列，但近30年来发展较快。目前，此类模型广泛用于各种水文序列的预报和模拟。从60年代末期以来，产生了新的模型。这类模型和上述马尔柯夫模型有别，故叫做非马尔柯夫模型，其中主要有自回归滑动平均求和模型、解集模型、散粒噪声模型、分数高斯噪声模型、快速分数高斯噪声模型以及折线模型等。自回归滑动平均求和模型以及解集模型常用于生产实际，其他几种模型由于复杂，耗费计算机时间较长，尚未广泛应用。

在建立以上一些模型时，通常是将数学上的理论直接用到随机水文过程中来，而没有全面分析水文过程的物理特性。为了克服这一缺点，不少水文学者致力于对水文过程进行概化，或者提出了一些有一定物理概念的随机模型，或者对已有模型作出物理解释。这一类模型尽管不太成熟，但加深了人们对随机模型的认识，并有助于对模型进行合理性分析，它们的出现促进了随机水文学的发展。

水资源系统规划、设计和运行常常涉及到若干个水文过程（如若干个站的径流过程），于是从单站的随机模型发展到多站的随机模型。现在已有不少多站模型，例如，多站自回归模型、多站分数高斯噪声模型以及多站折线模型等。

为了尽可能多地利用各种信息来提高模型的可靠性，从70年代起贝叶斯方法和卡尔曼滤波方法被用来估计模型参数。由于信息的不断增多，模型中的参数亦随之变化，所以特别称这样的模型为变参数模型。

近来，为了使模型除反映水文过程的相依特性外，还能反映水文变量边际分布的统计特性（如皮尔逊Ⅲ型分布），在广泛应用的线性正态模型基础上提出了非线性偏态模型。这一进展大大拓宽了随机水文的实用领域。

加拿大水文学者黑勒尔1985年在水资源通报上编辑出版了《水资源时间序列分析》专集。该书收集了世界知名学者撰写的论文19篇。它的问世标志着随机水文学发展到了一个新阶段——随机模型和随机模拟技术全面用于复杂的水资源系统。

在我国，统计理论和分析技术很早就引入水文学，但以随机过程理论和时间序列分析技术探索水文过程的变化规律和随机模拟水文现象却是在10多年前才开始的。对随机水文学作较系统的研究并在生产单位尝试应用模拟序列始于80年代初。80年代中、后期，有关三峡工程洪水随机模拟的研究大大促进了我国随机水文学的发展。我国水利水电工程设计洪水计算规范已引入洪水随机模拟的内容。这表明随机水文学在中国已开始用于工程设计。随着随机水文学在各方面的应用和理论研究的不断深入，它必将有更大、更快的发展。

习　　题

1. 严格说来，随机水文学的研究对象与传统水文频率分析的对象有什么不同？
2. 试论述随机水文学方法和传统水文计算方法的异同点。
3. 试叙述随机水文学的应用。
4. 对随机水文学的发展趋势有何看法？

第二章　随机过程的基本知识

随机水文学以随机水文过程作为研究对象。随机过程理论及分析方法是随机水文学的基本理论和方法，因此在讲述随机水文学的主要内容之前先介绍随机过程的有关基本知识。

第一节　随机过程的概念

我们从以下两个水文过程的实例引出随机过程的基本概念。

【例1】　某水文站每年由自记水位计连续记录每个时刻的水位形成每一年的瞬时水位过程。如果水文站以上流域影响水位的诸因素各年不变（或相对稳定），可把各年观测到的水位过程作为相同条件下进行随机试验的结果。一次（一年）试验的结果是时间 t 的某种函数（并非某一确定的数），而且事先无法确切地进行预测。由于水位变化的随机性，每次（每年）试验的结果是不相同的。每年观测的水位过程将反映出水位与时间 t 的不同函数形式。显然，可以用观测到的一族水位与时间的函数来描述和研究多年水位变化过程，如图2-1所示。

【例2】　某水文站一年的日流量过程也为一次随机试验的结果，n 年试验的结果就是 n 条日流量过程。一年试验的结果就是日流量随时间 t 而变化的某种函数。由这一族函数就可以反映 n 年日流量过程的变化规律。

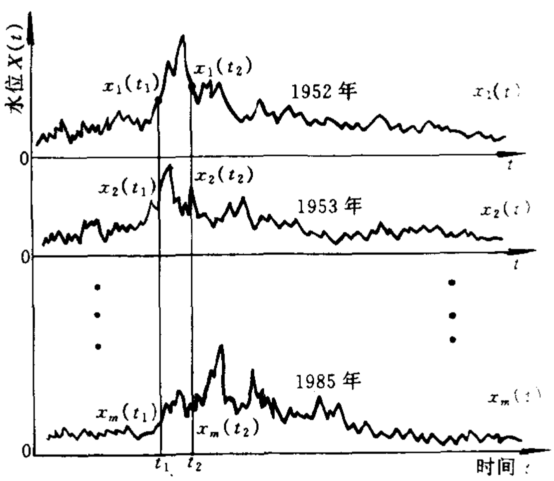

图 2-1　某水文站多年瞬时水位过程

一般来说，在相同的试验条件下，独立地重复多次随机试验，每一次试验结果是时间 t 的某种函数，其函数形式各次不同，且事先无法确定。我们称这族时间 t 的函数为随机函数。每次试验结果，即族中的每一个函数称为随机函数的一个现实或样本函数。可见，随机函数就是所有现实或样本函数的集合。

当随机函数随时间 t 连续地取有限区间内的值时，即称之为随机过程。而随机函数随时间 t 取离散值，则称为随机序列或时间序列。

按照上述定义，［例1］中的年内瞬时水位过程是一个随机水文过程，特定年水位与时间 t 的函数就是随机过程的一个现实或一个样本函数，它是一个普通的函数。［例2］中的年内日流量过程是一个时间序列，特定年日流量与时间 t 的关系表示成一串普通数列。

随机过程用 $X(t)$ 表示，各个现实用 $x_i(t)$（$i=1, 2, \cdots$）表示。随机序列用 $X(t$

$=1, 2, \cdots$）表示，各个现实用 x_{it}（$i=1, 2, \cdots$）表示。图 2-1 表示的是一个水位随机过程 $X(t)$，第 i 年观测的瞬时水位随时间 t 而变化的函数即为第 i 个现实 $x_i(t)$。

对于每一个固定时刻，如 $t=t_1$，$X(t_1)$ 是一个随机变量。图 2-1 中的 $x_1(t_1)$，$x_2(t_1)$，…，$x_m(t_1)$ 是随机变量 $X(t_1)$ 在 m 次试验中的取值。当 t 取不同值时，就有一串随机变量 $X(t_1)$，$X(t_2)$，…。习惯上称随机变量 $X(t_1)$ 是随机过程 $X(t)$ 在时刻 $t=t_1$ 的截口或称为随机过程在 $t=t_1$ 时的状态（在图 2-1 中，t_1 为 6 月 15 日 8 时。这一截口的特性则表示该时刻历年水位随机变化的特性）。因而，随机过程是随机变量 $X(t_1)$，$X(t_2)$，…所构成的。或者说，随机过程 $X(t)$ 是依赖于时间 t 的一族随机变量。

必须说明的是：一个随机变量 X 与它的一个样本值 x 是意义不同的量。前者是对这一具体随机现象（或随机试验结果）的总称；后者仅是对这个随机现象所取得的一个观测值（或试验值）。随机过程 $X(t)$ 是对随机试验得到的所有试验结果随时间 t 变化的函数的总称，即对任意时间 t 是一个随机变量；样本函数 $x_i(t)$ 则是随机试验的某一具体的观测值（或试验值）随时间 t 变化的函数。对任意固定 t，随机试验的观测值（或试验值），即为随机变量取得的一个观测值（或试验值）。区别了这些概念，自然会分清 X、x、$X(t)$、$x(t)$ 代表的意义。

第二节　随机过程的分布函数

随机过程在任意一时刻的状态是随机变量，由此可利用随机变量的统计描述方法来描述随机过程的统计特性。

$X(t)$是随机过程，对任一固定 $t_1\in T$，$X(t_1)$是一个随机变量，其分布函数为$F_1(x_1, t_1)=P[X(t_1)\leqslant x_1]$。若 $\partial F_1(x_1, t_1)/\partial x_1$ 存在，则有 $f_1(x_1, t_1)=\partial F(x_1, t)/\partial x_1$。$F_1(x_1, t_1)$ 与 $f_1(x_1, t_1)$ 分别叫做随机过程 $X(t)$ 的一维分布函数与一维概率密度。一维分布函数描述了随机过程在各个孤立时刻的统计特性。

随机过程 $X(t)$ 在任意两时刻 t_1、t_2 的两个随机变量（$X(t_1)$，$X(t_2)$）被称为 $X(t)$ 的二维随机变量。其二维分布函数为 $F_2(x_1, x_2; t_1, t_2)$，其二维概率密度为

$$f_2(x_1,x_2;t_1,t_2)=\frac{\partial^2 F_2(x_1,x_2;t_1,t_2)}{\partial x_1\partial x_2}$$

二维分布函数比一维分布函数包含了更多的信息，它反映了任意二时刻 t_1、t_2 状态间的统计关系。

同样可引入随机过程 $X(t)$ 的 n 维分布函数和 n 维概率密度

$$F_n(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots,t_n)=p[X(t_1)\leqslant x_1,X(t_2)\leqslant x_2,\cdots,X(t_n)\leqslant x_n] \quad (2\text{-}1)$$

$$f_n(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots,t_n)=\frac{\partial^n F_n(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots,t_n)}{\partial x_1\partial x_2\cdots\partial x_n} \quad (2\text{-}2)$$

显然，n 愈大，则 n 维分布函数描述随机过程的统计特性也愈趋完善。一般分布函数族（F_1，F_2，…）或概率密度族（f_1，f_2，…）完全地确定了随机过程的全部统计特性。

例如，要描述年内日流量序列的统计特性，它的一维分布函数（或概率密度）描述了 365 日各日流量截口的统计特性；二维分布函数（或概率密度）描述了任意两个日流量截口

之间的联系；n（$n \leqslant 365$）维分布函数（或概率密度），则描述了任意 n 个日流量截口之间的联系。一维、二维、…、n 维日流量分布函数族（或概率密度族）就完全描述了年内日流量序列的全部统计特性。

第三节　随机过程的数字特征

虽然随机过程的分布函数族（F_1，F_2，…）能完善地描述随机过程的统计特性，但要具体分析确定它，往往既复杂又困难，使用也不方便。因而，像在随机变量中引入一些数字特征一样，在随机过程中也引入了一些数字特征。这些数字特征既便于刻画随机过程的重要统计特征，又便于进行实际计算。

随机过程的主要数字特征为数学期望函数、方差函数和相关函数等。

1. 数学期望函数（均值函数）

对应于某固定时刻 t，随机过程成为一个随机变量，因此可以按通常随机变量一样的方法定义过程的数学期望。数学期望在一般情况下依赖于 t，是 t 的确定函数。此函数称为随机过程的数学期望函数，即

$$m_x(t) = E[X(t)] = \int_{-\infty}^{\infty} x f(x,t)\mathrm{d}t \tag{2-3}$$

$m_x(t)$ 是一个平均函数，随机过程在它附近起伏变化，如图 2-2 所示，图中细线表示随机过程的各个样本函数，粗线表示数学期望。

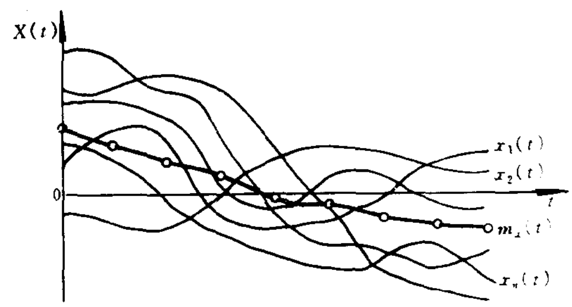

图 2-2　随机过程围绕平均函数起伏变化的情况

需要指出：这里的 $m_x(t)$ 是随机过程 $X(t)$ 的所有样本函数在时刻 t 的平均值。这是统计平均（又称集平均），应与后面将引入的时间平均概念相区别。

2. 方差函数（标准差函数）

随机变量 $X(t)$ 的二阶中心矩 $DX(t)$ 被称为随机过程 $X(t)$ 的方差函数，即

$$DX(t) = E[X(t) - m_x(t)]^2 = \int_{-\infty}^{\infty} [x(t) - m_x(t)]^2 f(x)\mathrm{d}x \tag{2-4}$$

方差的平方根 $\sqrt{DX(t)}$，即

$$\sigma_x(t) = \sqrt{DX(t)} \tag{2-5}$$

称为随机过程的标准差函数。$DX(t)$ 和 $\sigma_x(t)$ 是 t 的普通函数，描述了随机过程对于数学期望 $m_x(t)$ 的偏离程度。

3. 自协方差函数

若 $X(t_1)$ 和 $X(t_2)$ 为随机过程 $X(t)$ 在任意两个时刻 t_1，t_2 的两个截口，$f_2(x_1, x_2; t_1, t_2)$ 是相应的二维概率密度，则称二阶中心相关矩

$$\begin{aligned} COV(t_1,t_2) &= E\{[X(t_1) - m(t_1)][X(t_2) - m(t_2)]\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [x_1 - m(t_1)][x_2 - m(t_2)] f_2(x_1,x_2;t_1,t_2)\mathrm{d}x_1\mathrm{d}x_2 \end{aligned} \tag{2-6}$$

为随机过程 $X(t)$ 的自协方差函数（或称协方差函数）。它刻画了随机过程 $X(t)$ 在时刻 t_1 与 t_2 之间的统计联系。

4．自相关函数

随机过程 $X(t)$ 的自相关函数

$$\rho(t_1,t_2)=\frac{COV(t_1,t_2)}{\sigma(t_1)\sigma(t_2)} \tag{2-7}$$

也称为随机过程 $X(t)$ 的标准化协方差函数或自相关函数，简称为相关函数。

自协方差函数和自相关函数刻画了随机过程在任意两个不同截口之间的线性相关程度。

上述基本数字特征已能刻画随机过程的主要统计特性，而且易于观测和实际计算。但在随机水文学中，随机过程各截口的标准化三阶中心矩，即偏态系数 $C_S(t)$ 也是一个重要的统计特征。对于偏态系数及其它数字特征，在熟悉基本数字特征后就不难理解了．此处不再赘述。

第四节 平稳随机过程

实际中最常用的、在数学上也最成熟的随机过程是平稳随机过程。为了易于理解平稳的概念，先做如下说明。

在河流断面上，年径流量的总体若以 n 年为一组，客观上存在着多组样本，如表 2-1 所示。所有可能出现的样本 $x_1(t)$，$x_2(t)$，$x_3(t)$ …的集合构成了随机过程 $X(t)$。对于特定的时间 t_1，$X_1=X(t_1)$ 是一个随机变量。同样，$X_2=X(t_2)$，$X_3=X(t_3)$，…，$X_n=X(t_n)$ 都是随机变量。

表 2-1　　　　随机序列及其数字特征

t	t_1	t_2	t_3	…	t_n	t_{n+1}	t_{n+2}
第一个样本 $x_1(t)$	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	…	$x_{1,n}$	$x_{1,n+1}$	$x_{1,n+2}$
第二个样本 $x_2(t)$	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	…	$x_{2,n}$	$x_{2,n+1}$	$x_{2,n+2}$
第三个样本 $x_3(t)$	$x_{3,1}$	$x_{3,2}$	$x_{3,3}$	…	$x_{3,n}$	$x_{3,n+1}$	$x_{3,n+2}$
随机变量	X_1	X_2	X_3	…	X_n	X_{n+1}	X_{n+2}
数学期望	m_1	m_2	m_3	…	m_n	m_{n+1}	m_{n+2}
方　差	D_1	D_2	D_3	…	D_n	D_{n+1}	D_{n+2}

从 t_1 时刻开始，有 n 个随机变量组成 n 维随机变量，其 n 维分布函数为

$$F_n=(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots t_n)$$

从 t_2 时刻开始的 n 个随机变量也可组成 n 维随机变量，也有 n 维分布函数

$$F_n(x_1,x_2,\cdots,x_n;t_2,t_3,\cdots,t_n,t_{n+1})$$

同样，从 t_3 时刻开始的 n 个随机变量组成 n 维随机变量，其 n 维分布函数为

$$F_n(x_1,x_2,\cdots,x_n;t_3,t_4,\cdots,t_n,t_{n+1},t_{n+2})$$

从 t_4 时刻开始的 n 维分布函数依次类推，就不一一列举了。

一般情况下，同样是 n 个随机变量，但开始时刻不同，它们的 n 维分布函数是不相等的，即

$$
\begin{aligned}
&F_n(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots,t_n)\\
\neq\ &F_n(x_1,x_2,\cdots,x_n;t_2,t_3,\cdots,t_n,t_{n+1})\\
\neq\ &F_n(x_1,x_2,\cdots,x_n;t_3,t_4,\cdots,t_n,t_{n+1},t_{n+2})\\
\neq\ &\cdots
\end{aligned}
$$

但对一类特定的随机过程，即平稳随机过程而言，这些仅开始时刻不同而维数相同的 n 维随机变量的 n 维分布函数，都是相等的。由此，我们引出平稳过程的定义如下。

一个随机过程 $X(t)$，如果其 n 维分布函数不因所选开始时刻的改变而不同，即对任何 n 与 k，$X(t)$ 的 n 维分布函数满足

$$F_n(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots,t_n)=F_n(x_1,x_2,\cdots,x_n;t_k,t_{k+1},\cdots,t_{k+n}) \tag{2-8}$$

则 $X(t)$ 被称为平稳过程，否则即被称为非平稳过程。

这说明平稳随机过程的统计特性与所选取的时间起点无关，或者说统计特性不随时间的推移而变化。例如，在同一断面上利用1900年开始的相当长的年径流系列计算得到的年径流量 n 维分布函数与利用任一年，比如1913年，开始的相当长的年径流系列计算得出的 n 维分布函数是相等的。这种现象的解释是：若产生随机过程的主要物理条件在时间进程中没有变化，则该随机过程的统计特性也不会随时间而变化。例如，如果气候条件与下垫面条件都没有重大变化，则年径流的统计特性也不会随时间而变化，因而不同开始时刻的年径流 n 维分布函数也不会有变化。由于平稳随机过程的这个特点，可使问题的分析计算大为简化。它具有一系列简单的特性，因而这个随机过程在生产上得到了广泛的应用。

平稳过程的数字特征具有以下特点：

(1) 若 $X(t)$ 为平稳过程，则它的一维分布函数及一维概率密度都与时间无关。

根据平稳过程的定义，当 $n=1$ 时，对任意 τ 有

$$F_1(x_1,t_1)=F_1(x_1,t_1+\tau)$$

当 $\tau=-t_1$ 时，则有

$$F_1(x_1,t_1)=F_1(x_1;t_1-t_1)=F_1(x_1;0)=F_1(x_1)$$

同样有

$$f_1(x_1;t_1)=f_1(x_1;t_1+\tau)=f_1(x_1)$$

这样 $X(t)$的均值（数学期望）也与时间无关，即其均值 m_x 为常数。

$$E[X(t)]=\int_{-\infty}^{\infty}x_1f_1(x_1,t_1)\mathrm{d}x_1=\int_{-\infty}^{\infty}x_1f_1(x_1)\mathrm{d}x_1=m_x \tag{2-9}$$

同样 $X(t)$ 的均方值 ψ_x^2 与方差 σ_x^2 也是常数

$$E[X^2(t)]=\int_{-\infty}^{\infty}x_1^2f_1(x_1)\mathrm{d}x_1=\psi_x^2 \tag{2-10}$$

$$D[X(t)]=\int_{-\infty}^{\infty}(x_1-m_x)^2f_1(x_1)\mathrm{d}x_1=\sigma_x^2 \tag{2-11}$$

于是，平稳随机过程的所有样本函数曲线都将在水平直线 $m_x(t)=m_x$ 周围波动，其平稳偏离度是 σ_x。