

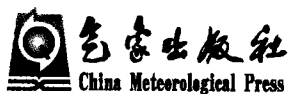
Tongji Tianqi
Yubao Yuanli
yu Fangfa

统计天气预报原理与方法

-----◎ 孔玉寿 钱建明 臧增亮 编著

统计天气预报原理与方法

孔玉寿 钱建明 臧增亮 编著



内 容 简 介

本书介绍了天气分析预报的统计学原理与方法,内容包括:概率统计中的几个基本概念、因子筛选、回归分析、判别分析、聚类分析、时间序列及其谱分析、主分量分析、统计天气预报的几种新方法以及预报的综合集成和质量评定等。

本书既注重基本原理的阐述,更侧重于实际应用方法的介绍,不但可作为高等院校大气科学专业及相关专业的教材,也可供广大从事天气预报及相关领域的科研和业务人员参考。

图书在版编目(CIP)数据

统计天气预报原理与方法/孔玉寿,钱建明,臧增亮编著.

北京:气象出版社,2010.11

ISBN 978-7-5029-5079-8

I. ①统… II. ①孔… ②钱… ③臧… III. ①天气预报-气象资料-统计分析 IV. ①P468.0

中国版本图书馆 CIP 数据核字(2010)第 214714 号

出版发行:气象出版社

地 址:北京市海淀区中关村南大街 46 号

总 编 室:010-68407112

网 址:<http://www.cmp.cma.gov.cn>

责任编辑:林雨晨

封面设计:博雅思企划

责任校对:永 通

印 刷:北京京科印刷有限公司

开 本:720 mm×960 mm 1/16

字 数:715 千字

印 次:2010 年 12 月第 1 次印刷

邮政编码:100081

发 行 部:010-68409198

E-mail: qxcbs@cma.gov.cn

终 审:章澄昌

责任技编:吴庭芳

印 张:36.5

版 次:2010 年 12 月第 1 版

定 价:60.00 元

本书如存在文字不清、漏印以及缺页、倒页、脱页等,请与本社发行部联系调换

前 言

统计天气预报亦称概率统计天气预报,是运用概率统计理论和方法,从大量的历史气象资料中寻找出气象要素或物理量(预报因子)与未来天气(预报量)之间的统计关系,建立预报方程,作预报时用当时可获得的预报因子值代入预报方程,经计算作出天气预报的方法。

一、统计天气预报在大气科学中的作用和地位

大气科学是以全球大气为对象,研究发生在大气中的各种现象及其演变规律,以及如何利用这些规律为人类服务的一门学科。它的主要目的是掌握大气状态的变化规律,预测天气、气候变化,最终达到控制和人工影响天气、气候。大气科学的研究内容十分广泛,它依据物理学和化学的基本原理,运用各种技术手段和数学工具,研究大气的物理和化学特性、大气运动的各种能量及其转换过程、各种天气气候现象及其演变过程、天气气候现象的预报方法、影响天气气候过程的技术措施、大气现象各种信息的观测和获取以及传递的方法和手段等。其中统计天气预报所依据的方法是揭示大气运动规律、预测天气和气候变化的主要方法之一。因此,它在大气科学中的作用和地位,也主要表现在以下两个方面:

1. 统计天气预报所依据的方法是揭示大气运动规律的一种有效工具

分析研究大气运动规律的方法,大体上可分为三大类:理论分析法、统计分析法和实验分析法。

应用理论分析法研究问题的代表学科是天气学和动力气象学。它依据天气观测的事实,应用天气分析方法和热力学、流体力学的概念与原理,分析天气系统的分布和空间结构、演变过程及其与天气变化的关系,从而揭示大气热力、动力过程的基本规律,为制作天气预报提供了基本的理论依据和定性的预报方法。

实验分析法包括实验室模拟实验(也叫物理模式法)和数值模拟实验。理论研究为复杂的大气现象建立了符合物理定律的数学模式,因而可以通过实验室实验,或在计算机上对模式进行时间积分,来仿真地研究其变化情况,这就是实验室“物理模拟”和计算机“数值模拟”。尤其是数值模拟,已经成为当今揭示大气运动规律的主要方法。在此基础上发展起来的数值天气预报,已成为当今天气预报的主要技术。

但是,理论分析法和实验分析法(包括数值模拟)都有一定的局限性,它们难以担当揭示大气运动规律的全部责任。统计天气预报所依据的统计分析原理和方法,则从另一条途径加入揭示大气运动规律的行列。

统计分析法运用概率统计理论和回归分析、判别分析、聚类分析、时间序列分析、主分量分析、谱分析等方法,从大量的历史气象资料中,寻找大气现象间的统计规律或相互关系。比如,平均的气候状况,前期的气象要素或物理量与未来天气的关系,等等,都可以用统计分析法得到。统计分析法还能从统计的意义上,找出影响某种天气或天气过程的主要因素,甚至可以发现一些理论分析中难以推断出的事实和规律。统计分析法还可以为检验数值模拟与数值预报模式的性能,或调整模式中的有关参数提供依据,卡尔曼滤波、动态相似统计等方法还可用于资料的变分同化与补缺,改善数值预报初始场,因此也是改进数值模拟、数值预报所不可缺少的分析研究方法。所以,随着概率统计理论和方法的发展,可以通过对大气运动的统计规律的了解和分析,进一步加深对大气动力学、热力学规律的认识,促进数值预报和天气学预报方法的发展。因此我们说,统计天气预报所依据的方法是揭示大气运动规律的一种有效工具。

2. 统计天气预报是预测天气和气候变化的一种主要方法

天气预报技术发展到今天,逐渐形成了三类基本模式:根据天气学原理和经验建立的、以物理定性关系描述为主的天气学模式;由当前气象要素与未来天气之间的统计关系建立的统计学模式;利用当前气象要素的分布与未来天气之间的物理定量关系建立的动力学(数值)模式。与这三类预报模式相应的基本预报方法也有三种:天气图方法、统计预报方法和数值预报方法。

现在,天气形势的预报,尤其是短期形势预报,基本上可以依赖数值预报了;中期形势预报也已经从梦想变为现实;部分气象要素预报,也可以直接由数值预报制作出来了。数值预报已成为制作天气预报的主要方法甚至是基础。但是,毕竟数值预报的优势是形势预报。以定性推理判断为主要特征的天气图预报,虽有其难以替代的作用,但其主观定性且强烈依赖经验的特点大大限制了它的发展。

统计天气预报却具有其他预报方法无法比拟的优越性。尤其在制作客观定量的要素预报和中、长期天气预报方面,有着独特的优势。它与动力学方法结合形成的统计-动力预报方法,更为数值天气预报的发展和数值预报产品的应用开辟了一条新的途径。统计天气预报的基本原理和方法,也是作为现代天气预报技术的一个重点发展方向——数值预报产品解释应用技术的主要基础。因此,天气图方法、数值预报方法和统计预报方法三者互相配合、取长补短,是使天气预报的准确率不断提高的主要技术途径。

二、统计天气预报的发展

首先值得指出,迄今为止,还没有看到一本比较权威的著作对统计天气预报的发展历史做过系统的介绍。也许,这是因为统计天气预报的发展是渐进的,没有出现过像其他学科那样的有公认里程碑意义的成就;或许,因它本身就是气象学、概率论和数理统计等多学科知识融合的产物,人们对它的历史反而熟视无睹。因此,我们也不可能对此进行详细而系统的论述。

实际上可以认为,统计天气预报的历史能追溯到遥远的古代。我们知道,自古以来,人类就有对天气现象和物候的观测,并在此基础上总结出诸如“朝霞不出门,晚霞行千里”、“八月十五云遮月,正月十五雪打灯”之类的大量天气谚语,人们根据这些天气谚语制作天气预报。这种天气谚语通常认为是经验性的,但从统计学的观点看,实际上都是统计规律的总结。因为,任何经得起实践检验的经验,都是以大量的观测事实为依据而总结出来的,也许这种“总结”只是无意识地运用了统计学原理。

自从气象学成为一门学科以来,概率论和统计学的应用更是始终伴随着大气科学的发展。1820年德国气象学家布兰德斯(H. W. Brands)绘制了世界上第一张(地面)天气图,人们“统计”了大量地面天气图上气压系统与天气分布的关系,得出“高压多晴天,低压多雨天”的粗浅认识,开始采用非常简单的外推法预报高、低压系统的移动来制作天气预报。到了20世纪20—30年代,中外气象学家开始更广泛地应用相关和回归分析方法来研究涛动、季风及早涝等大气现象的演变规律,并取得了许多有意义的成果。有人将20世纪50年代后期Ponotsky的专著《The Application of Statistic in Meteorology》的出版,看做天气预报的概率统计方法成为一门学科的标志,可惜该书现已难以查证。其后,统计预报逐渐成为制作天气预报的一种主要方法。我国自20世纪60年代开始使用和发展统计天气预报技术,在70年代前期达到了高潮,并逐渐发展,至今仍是广大预报员很熟悉也很常用的一种天气预报方法。

在统计天气预报技术的发展过程中,有一个因素起着重要的促进作用,那就是数值预报产品的统计解释应用。自1950年数值天气预报获得成功以后,如何弥补数值预报在制作气象要素预报方面的缺陷而更好发挥其效益,成为亟待解决的现实问题。以1959年克莱因(W. H. Klein)提出的完全预报(Perfect Prediction, PP)法和1972年格莱恩(H. R. Glahn)和劳里(D. A. Lowry)提出的模式输出统计(Model Output Statistics, MOS)预报法为代表,统计学方法成为数值预报产品定量解释应用的基本工具,并形成了动力-统计这一统计天气预报新的学科方向。卡尔曼滤波等新方法的应用不但使数值预报产品定量解释应用有了更多的方法,而且也促进了统计天气预报技术本身的发展。

经过长期的预报实践,人们发现了传统统计学方法的若干不足,于是借鉴现代控

制论等新理论中的有关观点和方法不断加以改进和完善,使其预报能力更强,形成了统计学预报技术发展的两个最显著的特点:一是从线性统计向非线性统计发展,适应了解决天气预报中大量非线性问题的需要;二是从静态、非时变、固定参数的预报模型向动态、时变、具有自适应能力的预报模型发展,为解决大气系统这一高度动态、时变的系统的预测问题提供了技术途径。多层递阶预报、卡尔曼滤波方法的应用、灰色预测、动态相似统计预报等基于统计学理论的新预报方法相继开发出来,统计天气预报的技术方法不断进步,预报能力不断提高。

三、统计天气预报的特点

从科学内涵、学习研究和技术应用等方面看,统计天气预报有如下一些显著的特点:

1. 统计天气预报处理的对象是气象中的随机现象和随机过程

我们知道,对客观世界的认识存在着两种截然不同的观点,这就是确定论的观点和随机论(不确定论)的观点。在气象领域也是如此。确定论的观点认为,各种物理系统的变化遵循一定的自然法则,系统的未来状态就由这些法则和初始状态所确定。作为宇宙间一个物理系统的大气系统,它的状态可用一组四个变量的函数(包括三个空间变量,一个时间变量)来描述。因此,对大气运动在空间和时间上的变化的研究,就转化为求解方程组在给定的边界条件和初始条件下的解的问题。也就是说,大气在未来任何时刻的状态,完全是由其现在的状态和方程组所描述的法则决定的。数值天气预报就是确定论观点的典型例子。

随机论的观点则认为,大气运动是随机的,不确定的,但具有统计规律,即可用统计学方法揭示其演变的一般规律。这就是统计天气预报的理论依据。

实际上,大气运动既有随机性的特征,也有确定性的规律(数值天气预报的成功就是例证)。因此,把两种观点有机地结合起来,是全面深刻掌握大气运动规律的科学的方法。天气-统计和动力-统计方法的出现,就是这种结合的很好体现。从大气运动的可预报性来看,统计学方法的运用,将可大大延长其可预报期限。

2. 统计天气预报依赖于多学科知识的融合

概率论和数理统计是统计天气预报的基本工具和基础,气象中的随机现象和随机过程则是其处理的对象,而电子计算机的应用是有效处理海量气象资料、实现统计天气预报业务化的重要条件,所以说统计天气预报涉及众多学科领域知识的应用。这一特点反映在教科书上,表现为概念多、符号多、公式多、知识点多、公式推导复杂、理论联系实际突出、学生掌握困难等显著特点。因此,学习和研究统计天气预报技术,不但需要扎实的数理基础(包括概率论、数理统计和线性代数等知识)和一定的计算机应用能力,还必须掌握相应的天气学和大气动力学知识。只有与天气学和大气

动力学分析结合起来,才能将概率统计方法正确地运用到天气预报中去,这就是所谓的统计预报物理化问题。预报因子的选取和统计预报方程的建立,都必须具有明确的物理意义,这一统计天气预报的基本要求,强调了气象专业知识在统计天气预报中的作用。

3. 统计分析得到的结果是统计对象的期望值

由于基于不确定论观点的统计天气预报处理的对象是气象中的随机现象和随机过程,因此其统计分析得到的结果只是统计对象的统计规律——期望值,或说平均状况,它难以有效地揭示那些偏离期望值较远的、特殊的小概率事件以及转折性天气过程这些统计特性不稳定的大气现象。

为此,在充分理解统计天气预报在大气科学中的作用和地位,从而重视其研究和应用的基础上,强调一下其局限性是必要的。事实表明,对用概率统计方法难以解决的小概率或转折性天气的预报,结合天气学和动力学(数值预报)方法进行分析判断,是提高预报准确率的有效途径。所以说,综合应用多种预报方法和气象信息,而不片面强调某种方法的作用,是解决天气预报问题的科学态度。

四、本书内容简介

本书是作者在多年教学实践和从事相关科学研究的基础上,汲取了国内外同类教材的优点和相关科研成果编写而成的。全书共分十章,大致内容如下:

第1章介绍了概率统计中的几个基本概念,包括基本统计量、数理统计中的几个常用分布以及统计假设检验与应用等。主要为了解决“概率论与数理统计”与“统计天气预报”这两门课程的衔接问题,便于读者对后续内容的学习和掌握。

第2章介绍了统计预报制作的一般过程,使读者对使用统计学方法制作天气预报有一个初步的认识。在此基础上,着重介绍作为建立统计预报模型基础工作的因子分析和筛选的一般思路、原则与方法。

第3章至第5章介绍的是在气象要素场分析和预报上常用的几种主要多元分析方法,包括回归分析、判别分析和聚类分析等。

第6章与第7章均属时间序列分析的内容,主要介绍气象要素随时间变化的规律性的分析和预报方法。

第8章介绍的主分量分析与要素场的正交函数展开方法,作为气象信息分析和提取的重要手段,可与其他方法结合用来分析气象要素场的变化规律和制作预报。

第9章和第10章主要为反映学科的发展,着重介绍统计天气预报技术发展出现的几种有代表性的新方法,并针对日益增多的预报方法,介绍了如何评估它们的预报质量以及如何通过集成预报来提高它们的综合预报能力的方法。

本书前言及第9、10章由孔玉寿编写,第1—3和第6—8章由钱建明编写,第4、

第5章及计算程序的编写和实例的计算由臧增亮完成。

本书介绍的统计预报的原理和方法虽然主要是针对天气预报这个专业方向的,但在大气科学的其他方向也可广泛应用。甚至,其基本的技术原理和方法,还可用于如股市预测等凡是涉及信息分析和预测的其他领域。因此,本书主要作为高等院校大气科学本科专业的教材,也可供从事天气预报及相关领域的科研和业务人员参考。

由于笔者学识有限,相关资料收集也不够齐全,加之篇幅所限,难以将不断发展中的统计天气预报技术给予更全面的概括和介绍,不当之处敬请读者批评指正。

本书的出版得到解放军理工大学气象学院各级领导的大力支持,得到黄思训教授、费建芳教授、章东华教授和南京大学陈星教授等的热情关心和指导,得到气象出版社陶国庆主任的全力帮助,在此表示衷心的感谢。并对所有为本书的编写、出版提供过帮助的单位和个人,一并致以深切的谢意!

编著者

二〇〇九年十二月于南京

目 录

第 1 章 概率统计中的几个基本概念	(1)
1.1 基本统计量	(1)
1.1.1 平均值	(1)
1.1.2 标准差与方差	(4)
1.1.3 协方差与相关系数	(5)
1.1.4 分级相关系数与 χ^2 统计量	(6)
1.1.5 自协方差与自相关系数	(11)
1.1.6 落后交叉协方差与落后交叉相关系数	(12)
1.1.7 峰度系数与偏度系数	(13)
1.2 数理统计中的几种常用分布	(13)
1.2.1 χ^2 分布	(13)
1.2.2 t 分布	(14)
1.2.3 F 分布	(15)
1.3 统计量的检验及其应用	(16)
1.3.1 平均值的显著性检验及应用	(16)
1.3.2 两组样本平均值差异的显著性检验及应用	(18)
1.3.3 方差的显著性检验及应用	(20)
1.3.4 相关系数的检验	(22)
1.3.5 变量分布的检验	(23)
参考文献	(23)
第 2 章 因子筛选	(25)
2.1 制作统计天气预报的一般过程	(25)
2.1.1 确定预报对象	(25)
2.1.2 提供、筛选预报因子	(25)
2.1.3 建立预报模型(方程)	(32)

2.1.4 检验、修改预报模型	(33)
2.1.5 输入因子资料得到预报结果	(33)
2.2 预报因子的筛选方法	(34)
2.2.1 因子与预报对象都分两级时的因子筛选	(34)
2.2.2 因子值离散,预报量连续时的因子筛选	(44)
2.2.3 因子连续,预报量离散时的因子筛选	(47)
2.2.4 因子和预报量值都连续时的因子筛选	(48)
2.2.5 精选多个因子的方法	(52)
2.3 因子分析	(65)
2.3.1 因子分析的一般数学模型	(65)
2.3.2 主因子分析	(66)
2.3.3 特殊因子方差阵的迭代估计	(78)
2.3.4 因子轴的转动	(79)
2.3.5 因子分析法在天气预报中的应用举例	(86)
参考文献	(88)

第3章 回归分析

3.1 回归函数	(89)
3.2 一元线性回归	(91)
3.2.1 一元线性回归模型	(91)
3.2.2 回归计算举例	(94)
3.2.3 回归效果分析	(95)
3.3 多元线性回归方程	(105)
3.3.1 多元线性回归模型	(105)
3.3.2 偏回归系数的最小二乘估计	(107)
3.3.3 回归效果分析	(112)
3.3.4 预报计算举例	(121)
3.4 逐步回归	(124)
3.4.1 “最优”回归方程的选择	(124)
3.4.2 逐步回归的基本运算工具	(127)
3.4.3 逐步回归的计算公式与步骤	(133)
3.5 非线性回归	(143)
3.5.1 函数变换法	(143)
3.5.2 多项式展开法	(145)

3.5.3 一般的非线性回归模型	(148)
3.5.4 Logit 模型	(149)
3.6 事件概率回归	(151)
3.6.1 预报模型	(151)
3.6.2 概率回归系数的最小二乘估计	(151)
3.6.3 确定临界值作预报	(153)
3.6.4 预报计算举例	(153)
参考文献	(157)

第 4 章 判别分析

(160)

4.1 二级判别	(161)
4.1.1 问题的提出	(161)
4.1.2 判别函数	(161)
4.1.3 Fisher 判别准则及判别系数的确定	(162)
4.1.4 求判别临界值 y_c	(166)
4.1.5 应用举例	(167)
4.1.6 判别函数的显著性检验	(172)
4.1.7 判别因子的筛选	(173)
4.2 多级判别	(174)
4.2.1 判别函数离差平方和的分解	(174)
4.2.2 多级判别的 Fisher 准则	(177)
4.2.3 判别函数的性质	(178)
4.2.4 判别函数的显著性检验	(179)
4.2.5 决策规则	(182)
4.2.6 多级判别计算举例	(187)
4.3 逐步判别	(192)
4.3.1 单因子的判别能力及其显著性检验	(192)
4.3.2 逐步判别的基本思想及其计算过程	(198)
4.3.3 逐步判别中一些问题的讨论	(206)
4.3.4 非线性逐步判别	(207)
4.3.5 回归逐步判别	(207)
参考文献	(210)

第 5 章 聚类分析	(212)
5.1 聚类分析中的统计量	(212)
5.1.1 样品与指标	(212)
5.1.2 相似系数和距离	(213)
5.2 系统聚类法	(217)
5.2.1 系统聚类法的基本思想	(217)
5.2.2 最短距离法	(218)
5.2.3 最长距离法和中间距离法	(219)
5.2.4 重心法和类平均法	(220)
5.2.5 离差平方和法	(222)
5.3 逐步修改聚类法	(226)
5.3.1 整批修改聚类法	(227)
5.3.2 逐个修改聚类法	(230)
5.4 最优分割法	(234)
5.4.1 极差分割法	(235)
5.4.2 方差(变差)分割法	(237)
5.4.3 用最优分割法作预报	(245)
参考文献	(251)
第 6 章 时间序列分析	(253)
6.1 随机过程及其基本概念	(253)
6.1.1 随机过程的定义	(253)
6.1.2 随机过程的统计描述	(254)
6.1.3 平稳随机过程及其遍历性	(257)
6.1.4 平稳时间序列	(263)
6.2 平稳时间序列的预报	(265)
6.2.1 一维平稳时间序列的预报	(265)
6.2.2 多维平稳时间序列的预报	(268)
6.3 实际气象时间序列的分析	(272)
6.3.1 气候趋势的分析	(274)
6.3.2 周期振动的分析	(278)
6.4 平稳随机序列的线性模型	(295)
6.4.1 自回归模型(AR)	(295)

6.4.2 自回归-滑动平均模型($ARMA(p,q)$)	(302)
参考文献	(319)
第7章 时间序列谱分析	(321)
7.1 谱的基本概念	(321)
7.2 功率谱与相关函数	(325)
7.2.1 功率谱的概念	(325)
7.2.2 功率谱与相关函数	(327)
7.3 滤波	(338)
7.3.1 滤波的概念	(338)
7.3.2 几种常见的滤波方法	(341)
7.4 交叉谱分析	(351)
7.4.1 交叉谱的概念	(351)
7.4.2 交叉谱的估计	(355)
7.4.3 凝聚谱与位相谱	(358)
7.5 时-空谱分析	(361)
7.5.1 时-空谱的基本原理	(361)
7.5.2 时-空谱的计算方法	(364)
7.5.3 时-空谱的计算举例	(370)
参考文献	(373)
第8章 气象要素场的主分量分析	(375)
8.1 主分量	(375)
8.1.1 两个变量的主分量	(375)
8.1.2 多变量的主分量	(381)
8.2 气象要素场的切比雪夫多项式展开	(385)
8.2.1 切比雪夫多项式	(385)
8.2.2 切比雪夫多项式展开要素场的方法	(389)
8.2.3 展开要素场的精确度指标	(395)
8.3 气象要素场的自然正交函数展开	(397)
8.3.1 自然正交函数展开的原理	(398)
8.3.2 自然正交函数展开的计算方法	(404)
8.3.3 自然正交函数展开在天气气候分析、预报中的应用	(421)
参考文献	(427)

第 9 章 统计天气预报的几种新方法	(429)
9.1 动力-统计模型	(429)
9.1.1 PP 法	(430)
9.1.2 MOS 法	(431)
9.1.3 MED 法	(432)
9.2 动态时变统计模型	(433)
9.2.1 多层递阶预报方法	(434)
9.2.2 卡尔曼滤波方法的应用	(441)
9.3 非线性统计模型	(449)
9.3.1 非线性统计的基本概念	(450)
9.3.2 几种常用非线性模型的处理和应用方法	(457)
参考文献	(481)
第 10 章 预报的综合集成和质量评定	(484)
10.1 综合集成预报	(484)
10.1.1 综合集成预报的基本思路	(484)
10.1.2 综合集成预报的技术方法	(486)
10.2 预报质量的评定	(491)
10.2.1 离散型预报量的预报质量评定	(492)
10.2.2 连续型预报量的预报质量评定	(495)
10.2.3 形态型预报量的预报质量评定	(496)
10.2.4 概率预报的质量评定	(498)
10.2.5 预报质量的模糊评定方法	(499)
参考文献	(500)
附录 A 矩阵和向量的微分	(502)
附录 B 求函数的条件极值	(503)
附录 C 求矩阵的特征值及其特征向量	(504)
附录 D 统计预报常用 Fortran 程序	(516)
1. 逐个引入因子	(516)

2. 一元线性回归	(518)
3. 多元线性回归	(519)
4. 逐步回归	(522)
5. 判别分析	(528)
6. 整批修改聚类	(530)
7. 逐个修改聚类(K-means).....	(536)
8. 一维平稳时间序列分析	(543)
9. 谐波分析	(546)
10. 方差分析	(547)
 附录 E 标准正态分布的分布函数	 (552)
 附录 F t 分布表	 (554)
 附录 G χ^2 分布表	 (556)
 附录 H F 分布表	 (558)
 附录 I 相关系数检验表	 (566)

第1章 概率统计中的几个基本概念

由于大气运动存在非常复杂的过程,人们对大气运动规律的认识,一直存在两种看起来截然相反的基本观点:一是确定论的观点,认为只要有足够多的资料,根据流体力学和热力学原理,借助于大型计算机,就可以对大气的运动进行较为完整的数学描述;另一种是随机论(不确定论)的观点,认为大气是一个开放系统,大气运动是随机的、不确定的,但具有统计规律,即可用概率统计的方法揭示其演变的一般规律,得出未来天气出现的可能性预报。这两种观点,实际上是从不同的角度,道出了大气运动本身内在的不同特征,即大气运动既有随机性的一面,也具有确定性的一面。为了对大气运动的这种随机性(不确定性)作出科学而又客观的判断,我们需要求助于统计学方法来解决。

用统计学方法作天气的分析和预报,要依据大量的气象观测资料来进行。我们把某个气象要素及其变化称为一个变量(或随机变量),它的全体称为总体,而把收集到的该要素的实际资料称为样本。从统计学的观点看,统计天气预报就是通过对样本的分析来估计和推断总体的规律性。因此,在学习具体的统计预报技术之前,本章首先介绍概率统计中涉及的几个基本概念,包括常用的一些统计量及其检验和应用等。

1.1 基本统计量

统计量是统计学中用来表征样本资料统计特性(分布特点)的参量。基本的统计量主要包括平均值、标准差、方差、协方差、相关系数、峰度系数、偏度系数等。

1.1.1 平均值

平均值是统计学中最常用的一个重要统计量。如月平均气温、年平均降水量等都是这种统计量。平均值作为总体数学期望的一个估计值,其计算是针对具体样本的,描述的是样本资料数值变化的平均状态,是样本的函数。

对某个具体气象要素 x 而言,若在某个时间段的 n 次观测样本资料为 $\{x_i, i=1, n\}$, 则在该时间段内要素 x 的平均值计算式为: