

农业气象统计

魏淑秋 編著

福建科学技术出版社

农业气象统计

魏淑秋 編著

福建科学技术出版社

一九八五年三月

内 容 简 介

本书较系统地介绍了概率统计学原理、方法在农业气象上的应用。并用很大篇幅介绍了近代多元分析方法的应用。如用多元回归、逐步回归、回归积分、主因素分析等研究作物与气象环境之间关系，影响的关键时期、主导因子等问题；应用贝叶斯判别准则等研究作物生态适应性；应用判别分析、聚类分析等方法研究各种问题的客观分类、判别与区划；此外，应用韵律分析、谐波分析、马尔柯夫链、谱分析等时间序列分析方法进行周期性变化等研究；应用多项式、自回归模型等研究产量时间趋势项。应用各种统计方法建立数学模式、控制与预报方程。也较详细地介绍了一些非线性模型和非参数统计等相关。线性规划简介编在附录部份，由讲授运筹学的陈伟侯同志撰写。

本书内容丰富、新颖，有大量实例，文字深入浅出，便于自学。

本书可供农业气象及其有关学科如农学类(农学、植保、园艺、畜牧、土化等)以及生物、地学、林学等专业科技工作者及大专院校师生、研究生等参考。

农业气象统计

北京农业大学魏淑秋编著

福建科学技术出版社出版

(福州得贵巷27号)

北京农业大学农业资源环境及遥感研究所代理发行

北京航空学院印刷厂印刷

开本787×1092毫米1/16 印张24 61.28千字

1985年3月第1版

1985年3月第1次印刷

印数1—8000

书号13211.25 定价：4.60元

序

統計学是农业气象科学中最有用的工具。有了这一工具，农业气象便由描述阶段上升到精确科学的范畴。正如馬克思所說过的：“科学只有当它成功地运用数学的时候才能达到完善的地步”（《馬克思恩格斯全集》俄文第二版，第33卷，第72頁）。

如果說二十年代J.W.史密斯（Smith, 1920）著的《农业气象学》和四十年代美国气象局出版的《气候和人》算是定性的、描述性的农业气象学的话，那么七十年代由美、英、法、苏、日等九个国家的权威集体编写（1976年）的《植物和大气》，以及由联合国农业气象委员会委托J.西曼（Seemann, 1979）等四人联合编写的《农业气象学》，则可算是定量的、数值的农业气象学。天气预报中有数值预报、統計预报；农业气象中为什么不能有統計农业气象学呢？如果再有人說統計学是唯心的、是数学游戏的话，那就只会妨碍学术的进步！

早在半个世紀前（1924年），世界統計学权威費希尔就已經用回归积分（Regression integral）研究雨量对小麦的影响，奠定了統計农业气象学的基础。三十五年后，日本松田昭美于1959年把費希尔方法成功地应用到日本福岡的水稻生产上。再过二十年（1979年）卢其堯、王世耆、魏淑秋等同志又把它应用到北京和河北樂城冬小麦产量与气候条件的統計分析上，也获得成功。魏淑秋等同志在我国推广了这种方法也取得良好效果。

由于农业生产結構，不只是农、林、牧、副、漁五业結構，而是由农业經濟結構、农业技术結構、农业生态結構交織而成的多維結構，而且結構还是多层次的，必須用系統分析的方法来解决。試看西曼等人编写的《农业气象学》，除了应用到常用的統計方法外，还用綫性规划、决策論、庫存論、动态规划等，这已經是运筹学的范畴。讀者若不懂得这些內容，就休想赶上当代先进水平！本书作者针对这个道理，也介绍了綫性规划和决策論等运筹学內容，显得非常实用。

本书文字深入浅出，一般大专学生都能接受；若用作教材，定能获得良好的教学效果。

书中內容丰富而又頗有針對性。例如，针对农业院校和农业研究单位，有“田間試驗設計”，针对气象科学，有“時間序列”。M.G.肯特爾（Kendall）的《高級統計学理論》（The Advanced Theory of Statistics）是权威的統計书籍，但书中既无“田間試驗設計”，又无“時間序列”，更无“綫性规划”和“决策論”的內容。相反的，气候統計的权威书，如V.康拉特（Conrad）等著的《气候学方法》（Methods in Climatology）、英国气象局C.E.P.白罗克斯（Brooks）等著的《气象統計方法手冊》（Handbook of Statistical Methods in Meteorology）等书，都把“時間序列”放到重要的位置。所以該书具有“全”而有“針對性”的优点。

作为一个老教师，我有責任將中、青年教师的优秀著作向讀者介紹。

杨昌业

1984年9月于北京

前 言

农业气象统计是概率论数理统计在农业气象上的应用。

农业气象是农业与气象的一门边缘学科，研究范围是极其广阔的，其应用也是十分广泛的。

本书参考国内外有关气象统计、生物统计、医用、水文等应用统计及近代多元分析等书籍，较系统地介绍了概率统计的基本原理、方法在农业气象上的应用，并根据作者的教学与科研实践，列举一百多个典型实例，说明其计算过程及实际应用。

本书共分十章，前四章特征数，随机变量及其分布函数的应用，统计假设检验的应用，回归分析及其应用等，大都为基础概率论的应用；第五章为时间序列分析的应用；第六、七、八章为近代多元分析方法的应用；第九章为田间试验设计及其数据处理方法；第十章介绍了一些统计分析方法在农业气候资源考察与区划中的应用。为适应近代对系统分析的需要，特请讲授“运筹学”的陈维候先生撰写了“线性规划简介”搜集在附录中，供入门学习。

本书可供农业气象及有关学科，如农学、生物、气象、地学、林学等专业技术工作者和大专院校师生参考。

本书在编写与出版过程中，曾得到有多年气候统计教学经验的杨昌业教授、廖荣孙等先生的指导与审阅；得到石元春先生的多方支持与帮助；得到陈遵肅、韩湘玲等许多同行的支持与帮助；收到过许多听过作者讲课的同学们的来信，表达了他们对本书的热切希望并提出了宝贵意见。本书附录部份由讲授运筹学的陈伟候同志撰写，魏荣桦、王道龙、耿雯、杜荣、庞青玉等研究生参加本书全部或部份复算工作。在此表示衷心的感谢。

限于作者水平，缺点、错误在所难免，敬请读者批评指正。

1984年9月

目 录

概 述

- 第一节 统计学的任务和工作步骤..... (1)
- 第二节 农业气象统计的特点..... (5)

第一章 特征数

- 第一节 代表数的计算及其应用..... (6)
- 第二节 变异数的计算及其应用..... (11)
- 第三节 大样本平均数及方差的计算..... (16)

第二章 随机变量及其分布函数

- 第一节 随机变量..... (18)
- 第二节 分布函数..... (20)
- 第三节 几种主要的概率分布简介..... (22)
- 第四节 分布特征参数的计算与应用..... (32)

第三章 统计假设检验的应用

- 第一节 假设检验..... (37)
- 第二节 假设检验的应用..... (39)
- 第三节 分布函数的应用..... (77)

第四章 回归分析及其应用

- 第一节 相关关系..... (81)
- 第二节 直线回归及简单相关..... (82)
- 第三节 简化回归计算..... (88)
- 第四节 非参数统计相关..... (91)
- 第五节 多元回归方程的建立..... (98)
- 第六节 复相关及其检验..... (106)
- 第七节 偏相关及其检验..... (108)
- 第八节 回归效果的分析..... (110)
- 第九节 回归的误差问题..... (113)
- 第十节 各自变量的重要性..... (114)
- 第十一节 多元回归示例..... (118)
- 第十二节 逐步回归..... (121)
- 第十三节 线型选配及非线性回归..... (129)

第十四节 回归积分计算及其举例.....	(157)
第十五节 最接近于准确平均值的加权平均求法.....	(162)
第十六节 0、1 回归.....	(165)

第五章 时间序列分析的应用

第一节 韵律分析.....	(173)
第二节 谐波分析.....	(176)
第三节 方差分析.....	(187)
第四节 平稳时间序列分析.....	(194)
第五节 转移概率.....	(198)
第六节 谱分析.....	(203)

第六章 判别分析及其应用

第一节 二级分辨.....	(220)
第二节 多级判别及实例.....	(227)
第三节 应用贝叶斯准则进行判别分析和研究作物生态适应性.....	(231)

第七章 聚类分析

第一节 概 述.....	(239)
第二节 聚类分析的统计量及其均一化.....	(239)
第三节 分类系统的形成.....	(243)
附 模糊聚类.....	(251)

第八章 经验正交函数分析及其应用

第一节 经验正交函数分析的特点.....	(254)
第二节 经验正交函数分解的基本原理.....	(254)
第三节 经验正交函数的分解.....	(256)
第四节 经验正交函数分解示例.....	(259)
第五节 经验正交函数分解的应用.....	(265)

第九章 田间试验数据的处理及正交试验法

第一节 田间试验数据的处理.....	(270)
第二节 正交试验法.....	(279)

第十章 统计分析在农业气候资源考察与区划中的应用

第一节 概 述.....	(294)
第二节 作物产量材料的处理.....	(295)
第三节 统计分析方法的应用.....	(298)
第四节 统计决策的应用.....	(299)

附 录 线性规划简介

第一节	什么是线性规划.....	(305)
第二节	线性规划模型的一般形式.....	(310)
第三节	二维线性规划问题的图解方法.....	(311)
第四节	单纯形法的初步了解.....	(314)
第五节	归结成规范形式.....	(319)
第六节	用单纯形方法求解示例.....	(324)
第七节	单纯形法的一般步骤.....	(327)
第八节	从阶段 I 变到阶段 II	(332)
第九节	变量有界的线性规划.....	(337)
第十节	应用举例.....	(342)
附 表		(349)
参考文献		(346)

概 述

第一节 统计学的任务和工作步骤

一、统计学的任务

统计学主要解决以下几方面的问题：

(一) 表达某一事物或某种现象的基本特征

用几个简单特征数如平均数、均方差、变异系数等来描述某地气候基本特征，这是统计学的最基本任务。

(二) 进行统计假设检验

比较两个事物、两种现象的差异，或者用样本去估计总体（母体）情况，必须通过统计假设检验的方法去判定。例如：两个事物的差异，是由于试验条件的不同引起的，还是其它的偶然因素引起的；是本质区别，还是非本质的抽样波动；样本与总体是否有本质差别等问题，均属于统计假设检验问题。

(三) 表达两个事物或几个事物之间的关系

如农业气象中通常研究的光、热、水等气象因素与作物生长、发育及产量形成之间的关系和规律，以及根据这些关系和规律进行控制和预报未来状况等，这就是统计上的相关及回归分析等问题。它也是农业气象研究中的重要而基本的分析方法。

(四) 寻找引起事物发生变化的主要因素

引起一些事物和现象的变化，可能是由于一个或几个因素的变化而引起的。如何分析各种因素对某一事物变化所起作用的大小，如何抓住主要特征、主要因素去分析和认识某种事物或现象，这要用到方差分析、主因素（主成分）分析等分析方法。

(五) 研究各种事物或现象的自身变化

如气象要素随着时间的变化，气象要素的时间序列；又如，逐年产量的波动，产量时间序列，以及某一种作物由出苗到收获的生育过程序列等。研究这些时间序列的变化，统计上常用马尔科夫链、谐波分析、韵律分析、谱分析等方法，寻找其变化的周期、韵律，以及逐渐演变的转移概率等。

(六) 研究客观分类问题

一切事物或现象虽然没有完全相同的，但相似的情况很多。根据内因和外部形态、功能和性质等某一角度，将相似的情况归为一类，然后分门别类地去深入研究一些问题的规律，这是人们观察、了解实际问题的重要方法，统计上常用聚类分析、判别分析等方法研究分类问题。

(七) 研究取样和试验方法

如抽样的代表性问题，测量的精度、确度问题，试验方法、试验设计的合理性等问题。所谓合理，就是根据一定的试验目的，做到既能节约人力、物力，又能得到具有一定精度的可靠结果。目前除一般单因素田间试验设计以外，尚有多因素优选法——正交试验设计。

二、统计工作的步骤

统计工作大体上可分为：统计归纳、统计分析和统计推断三个步骤。下面将简述各步骤工作内容，以便对统计工作有个概括的了解。各步骤的具体计算方法、分析方法及判断方法将在以后各章节中逐一加以介绍。

(一) 统计归纳阶段

搜集并初步整理资料。首先可以用列表和作图的方法，把所搜集的资料加以归纳、整理，并计算出平均值、标准差等特征数，以揭示事物的基本特点，进行一般的特征描述。

例如：为了探讨伏、春雨对旱地冬小麦产量的影响，我们可以把逐年伏雨、春雨及与之对应的小麦产量列成一张表，如下表（1）所示。

表（1） 北京冬小麦产量与伏、春雨间的关系

年 份	(下/7至下/8) 伏雨 (mm)	(3~5月) 春雨 (mm)	冬小麦产量 (斤/亩)
1949~1950	404.1	186.2	83
1950~1951	340.2	148.2	86
1951~1952	166.5	57.4	89
1952~1953	307.6	101.1	124
1953~1954	309.7	56.0	109
1954~1955	515.8	122.1	143
1955~1956	406.6	46.1	113
1956~1957	545.9	38.8	73
1957~1958	237.6	58.3	100
1958~1959	208.8	34.3	80
1959~1960	1007.8	19.1	75
1960~1961	152.9	39.8	80
1961~1962	283.9	41.6	124
1962~1963	148.8	97.0	50

由上表可以初步看出，伏雨多产量高；伏雨少产量低，它们之间呈一定相关关系。由图（1）更清楚地看出它们之间的正相关关系，也说明了伏雨对冬小麦产量是有作用的。

由表（1）还可以进一步计算出伏旱、伏涝或春旱的指标。设伏雨旱涝指标如下：

$$V = \frac{R_i - \bar{R}}{\bar{R}} \times 100\%$$

V 为伏雨相对平均变率；

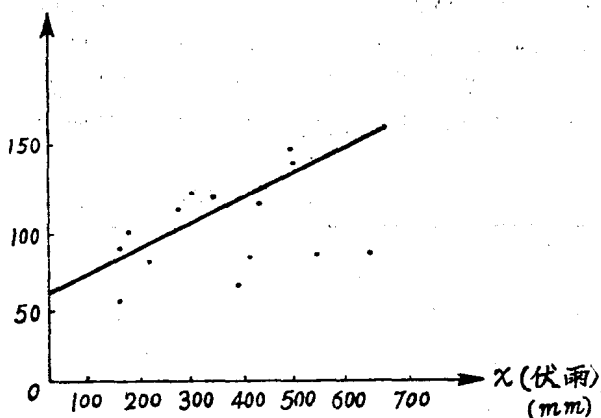
R_i 为某一年伏雨量；

$\bar{R}$ 为多年伏雨量平均。

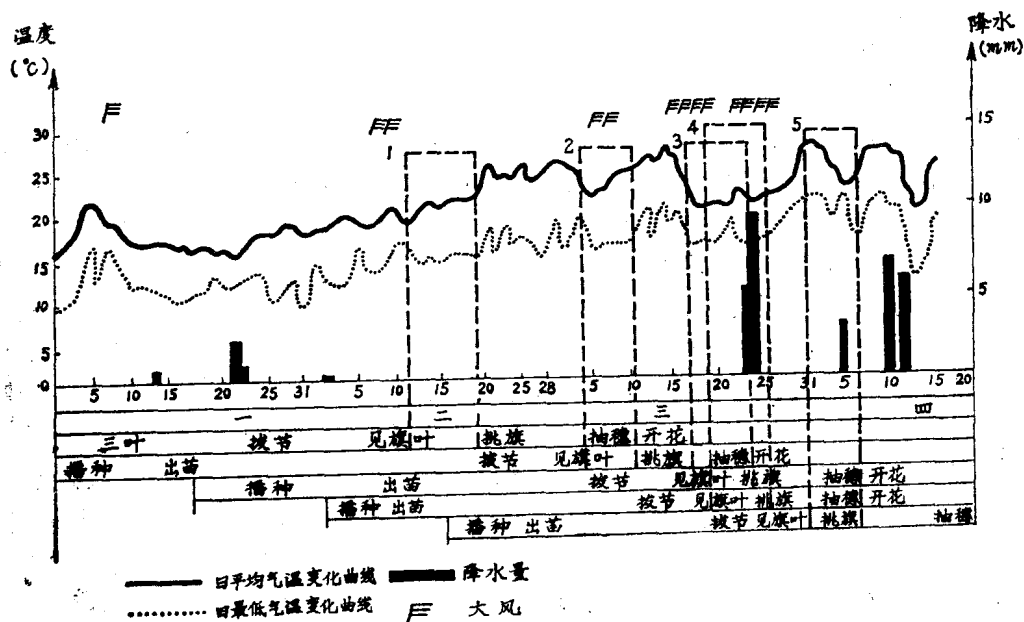
以相对平均变率 V 做为确定其伏旱、伏涝指标。如相对变率 V 小于 -2 时为伏旱，大于 2 时为伏涝， $-1 \sim -2$ 时为偏旱， $1 \sim 2$ 时为偏涝； $-1 \sim 1$ 时为正常。由此还可进一步确定必须进行冬灌的指标（春雨也可以进行同样的分析）这就是统计归纳的过程，通过统计归纳也可以得出初步结果服务于生产。

又如：在云南元江县进行种植杂交高粱的示范，开始时苗全且壮，长势喜人，估计亩产可过千斤。但实际亩产仅收83斤，对这个问题产生的原因，说法不一，有的归结为氮肥过多，有的认为拿错了种子，也有人认为是在花粉形成时期遇低温造成结实率低的原因。究竟是什么原因？北京农业大学赵明斋先生利用杂交高粱分期播种试验资料，用作图分析方法回答了这个问题。他将所要分析的项目及数据抄录在一张表上，根据资料进行绘图[图(2)]。

产量(斤/亩)



图(1) 北京地区冬小麦产量与伏雨关系



图(2) 气象要素变化与杂交高粱生育期图(赵明斋, 1970)

注：图中数字1、2、3、4、5分别为第一期、第二期、第三期、第四期、第五期播种，见旗叶至挑旗期间的温度状况。

从图(2)上明显看出,凡是在旗叶可见至挑旗这段时间内,若有寒潮大风天气,随之降温,其结实率就低。因为高粱该时段正值花粉形成阶段,对温度极为敏感,低温造成结实率降低,因而前期可望产量过千斤的降至亩产仅83斤的惨状,从而找到了失败的主要原因。

由图(2)还可找出初步指标:在旗叶可见至挑旗阶段平均气温低于23℃,极端最低气温低于14℃时,严重影响了花粉的形成,结实率一般在50%以下,如表(2)。

表(2) 杂交高粱旗叶可见至挑旗期间温度对结实率的影响

播种期(日/月)	20/11	1/1	16/1	1/2	15/2
旗叶可见至挑旗时间	12/2~19/2	5/3~9/3	17/3~24/3	20/3~26/3	1/4~6/4
旗叶可见至挑旗期间 温度变化(℃)	18.3~21.8	20.7~23.8	19.2~22.2	18.7~22.2	20.8~26.5
旗叶可见至挑旗期间 平均气温(℃)	20.7	23.6	20.4	20.5	23.8
旗叶可见至挑旗期间 极端最低气温(℃)	13.4	14.8	13.8	13.8	14.8
结实率(%)	11	50	0	15	90

当然这仅是一年材料,还可以多作几年,进一步揭示气温与结实率之间的关系和规律。

列表和作图方法,简单易行,一目了然,不需作大量繁琐计算,对于日常服务工作和初步探索一些问题起着重要作用。同时,列表和作图也是作进一步统计分析工作的基础,是一种常用的方法。

作图与列表时应注意:

- (1) 表及图的名称、时间、地点、单位、坐标、图例等需注明清楚。
- (2) 列表与作图项目应掌握在一定数目内。过多、过繁都不便分析;但亦应尽力充分利用图面。几个要素可以同时绘上,进行综合分析,便于相互比较,从而寻找主要原因。
- (3) 列表与作图之前,对所搜集的资料,必须进行审查去掉或纠正不合理资料,且应注意复查,注意不要抄错,或出现字迹不清楚等毛病。

(二) 统计分析阶段

对经过统计归纳初步整理的资料,作进一步分析,找出这些数字资料之间的关系及所遵循的规律。如前面二个例子都可以作进一步分析,分别计算伏、春雨对冬小麦产量形成的作用大小;伏、春雨与小麦产量之间的相关程度,以及所遵循的规律,拟合出经验公式等。前述第二例中,也可以进一步分析低温与结实率之间的关系,写出它们之间的相关式等等。此外,诸如小麦分蘖数与叶龄的关系,有人经过分析得出它们之间关系呈指数函数关系,即 $y = 10^{bx} + c$, y 为分蘖数, x 为叶龄数, b 、 c 为常数。根据实际资料拟合经验公式,求其常数 b 、 c 等。通过一定数理统计方法如:回归分析、方差分析、判别分析、聚类分析、时间序列分析以及主因素分析等分析工作阶段,称统计分析阶段。

(三) 统计推断阶段

由统计分析得出的规律和结论是否正确,其可靠程度如何?要加以判断。我们用来寻找规律的资料往往是样本,如何用样本去估计总体情况,必须经过推断性的假设检验工作,才可能找到真正规律性东西,去伪存真,得到可靠的结论,但任何一个统计结果,即便是通过

检验得出的结论，也不是绝对正确的，而是统计上的正确，是在一定程度上的正确。因此还必须估计出可靠的程度。这就是统计推断阶段。

第二节 农业气象统计的特点

农业气象是农业与气象的一门边缘学科。通俗地说，它是研究农作物、林果、蔬菜、畜牧、植物病虫害等与气象环境条件之间的相互关系及其规律性的科学。它研究农田小气候，温室、大棚、风障等保护地小气候，畜舍小气候以及营造防护林等森林小气候；对天气、气候灾害进行观测和预报；研究各种作物生态适应性及其地理分布；研究如何充分利用和开发农业气候资源及其进行农业气候区划、作物气候区划等等。农业气象工作的对象及研究范围是十分广泛的，其研究手段和工具也是多种多样，就统计工具来说，它包括农业气象及农业气候统计分析、统计模拟、统计预报、时间序列分析及多元分析的应用、田间试验设计及其数据处理等方面；此外农业气象还需要涉及部分运筹学内容。因此，农业气象统计比一般生物统计和气象统计有着更广泛的内容和特点。

时间序列分析在农业气象中有着特殊的地位，因为我们不仅要研究气象要素的时间序列，而且要研究农作物及果蔬等产量时间序列，还要研究生命过程序列。相关回归分析在农业气象中也十分重要，应用它可以分析作物生育及产量形成过程与气象环境条件之间的关系，寻找出影响的关键时期及主导因子；研究不同地方作物生态适应性，建立各种统计模型，以便于控制和预报。多元分析在农业气象中的应用也十分广泛，如应用经验正交分解确定其主要影响因子；应用判别分析、聚类分析进行农业气候区划的研究，进行生态类型的分析等。各种分布函数及假设检验，可用来寻找适宜的播种期和收获期，检定各种农业气象指标及其农业气象措施的效应，以及分析各种试验的误差等。农业气象也是一门实验科学，十分重视各种试验研究，尤其田间试验。因此，本书对田间试验设计、数据处理及多因素优选法等，也作了简要介绍。

农业气象统计既不同于一般气象统计，也区别于生物统计。但农业气象工作本身既要了解气象、气候规律，也要了解生物规律。在此基础上才能研究两者之间的相互关系，相互影响的效应及其变化规律。在研究手段上，也要掌握气象及生物研究的手段。因此，农业气象统计内容，也包含着气象统计和生物统计的有关内容。

农业气象统计，同许多学科的应用统计一样，是概率论数理统计在本学科中的应用。但又有其各自的特点和侧重面，是一般概率统计所无法代替的。农业气象统计涉及面较广，为避免重复和保持一定的体系，以便循序渐进，因此本书的章节大体按一般统计学的体系安排，并以大量实例阐明统计学方法在农业气象领域中的应用。

第一章 特征数

表达某种事物或现象基本特征的数称为特征数。特征数主要包括代表数和变异数。

第一节 代表数的计算及其应用

代表数是描述一个系列的“水平”、“大多数”、“大概”、“一般”、“中常”等情况。代表数包括平均数、中位数、众数。

一、平均数($\bar{x}$)

平均数是表达一个系列的“水平”、平均位置的量(以 $\bar{x}$ 表示)。平均数 $\bar{x}$ 分为算术平均数、几何平均数、调和平均数等。最常用的是算术平均数。它又分为简单算术平均数和加权算术平均数。

(一) 简单算术平均数的计算

1. 按定义直接计算： N 个样本的平均值 $\bar{x}_N$ 的计算公式为：

$$\bar{x}_N = \frac{1}{N}(x_1 + x_2 + \dots + x_N) = \frac{1}{N} \sum_{i=1}^N x_i = \frac{1}{N} \sum_i x_i \quad (1.1)$$

($i=1, 2, \dots, N$)

其式中 x_i 为 N 个样本中第 i 个样本

$\bar{x}_N$ 为 N 个样本的简单算术平均值。

例1—1 某地 1971~1980 年逐年小麦产量(斤/亩)为：320, 345, 336, 347, 382, 296, 343, 352, 369, 372, 求该地近十年小麦产量水平如何?

由公式(1.1)得简单算术平均值为

$$\bar{x}_N = \frac{1}{N} \sum_{i=1}^{10} x_i = \frac{1}{10}(320 + 345 + \dots + 372) = 346.2 \text{ (斤/亩)}$$

如果已知前 N 个样本算术平均值, 另外再加上一或二个样本, 是否需要将前 N 个样本中每一个数都重新一一加起来求和呢? 当然不必要, 可用添加变量时的简单算术平均值公式进行计算。

添加一个样本时

$$\bar{x}_{N+1} = \frac{1}{N+1}(N\bar{x}_N + x_{N+1}) \quad (1.2)$$

添加二个样本时

$$\bar{x}_{N+1} = \frac{1}{N+2}(N\bar{x}_N + x_{N+1} + x_{N+2}) \quad \text{其余类推。}$$

例1—2 例1—1中又增加1981年小麦产量412斤/亩, 求最近十一年小麦产量水平如何? 由公式(1.2)得

$$\bar{x}_{N+1} = \bar{x}_{10+1} = \frac{1}{10+1} (10 \times 346.2 + 412) = 352.2 \text{ (斤/亩)}$$

如果已知 N 个样本平均值,若去掉 N 个样本中一个或两个数,是否需要重新一一计算?当然也是不必要的,可用减少一个或两个变量时的算术平均值的公式计算。

减少一个样本时

$$\bar{x}_{N-1} = \frac{1}{N-1} (N\bar{x}_N - x_i) \quad (1.3)$$

减少两个样本时

$$\bar{x}_{N-2} = \frac{1}{N-2} (N\bar{x}_N - x_i - x_j)$$

其中, x_i 、 x_j 为 N 个样本中第 i 、第 j 个数。

例1—3 例1—1中1971~1980年十年小麦产量中,1975年由于记录错误,需要删去一个数据($x_{75}=382$ 斤/亩)求其平均数。

根据公式(1.3)

$$\bar{x}_{10-1} = \frac{1}{10-1} (10 \times 346.2 - 382) = 342.2 \text{ (斤/亩)}$$

2. 递推法:在应用电子计算机进行大量计算时,为了便于随时得到中间结果,可采用递推的方法,其表达式如下

令 $\bar{x}_0 = 0$, 对 $i=1, 2, \dots, N$

$$\bar{x}_i = \frac{i-1}{i} \bar{x}_{i-1} + \frac{1}{i} x_i = \bar{x}_{i-1} + \frac{1}{i} (x_i - \bar{x}_{i-1}) \quad (1.4)$$

例1—4 例1—1中欲分别得到前4年的产量平均值。

根据递推公式(1.4)得

$$\bar{x}_1 = \bar{x}_{1-1} + \frac{1}{1} (x_1 - \bar{x}_{1-1}) = 320;$$

$$\bar{x}_2 = \bar{x}_{2-1} + \frac{1}{2} (x_2 - \bar{x}_{2-1}) = 320 + \frac{1}{2} (345 - 320) = 332.5;$$

$$\bar{x}_3 = \bar{x}_{3-1} + \frac{1}{3} (x_3 - \bar{x}_{3-1}) = 332.5 + \frac{1}{3} (336 - 332.5) = 333.7;$$

$$\bar{x}_4 = \bar{x}_{4-1} + \frac{1}{4} (x_4 - \bar{x}_{4-1}) = 333.7 + \frac{1}{4} (347 - 333.7) = 337.0;$$

以下类推。

3. 简化算术平均值计算:若数字较大,位数较多,计算繁杂,为简化计算,可设一接近均值的任一常数 c ,按(1.5)式计算算术平均值。

$$\bar{x} = c + \frac{1}{N} \sum (x_i - c) \quad (1.5)$$

例1—5 简化计算例1—1中十年小麦产量平均值。

设 $c=300$,则 (x_i-300) 为20, 45, 36, 47, 82, -4, 43, 52, 69, 72

根据公式 (1.5) 得

$$\bar{x} = 300 + \frac{1}{10}(20 + 45 + 36 + \dots + 72) = 346.2 (\text{斤/亩})$$

与例 1—1 结果完全一样，而计算简便些。

4. 滑动平均数：表达一个系列的变化趋势时，为了突出主要的变化趋势，将次要的、小的波动平滑掉；或者，对气象序列，产量序列作韵律分析时，为了突出明显、主要的波动，将一些非周期性的小波动滤掉；或者，求算某界限温度出现的平均日期时，事先将序列进行平滑。近年来，又用它描述产量时间趋势项（描述由于人为因素影响产量的部分）。

滑动平均数的公式为

$$\bar{x}_{mi} = \frac{1}{m} \sum_{k=i}^{i+(m-1)} x_k \quad (1.6)$$

其中：\$i=1, 2, \dots, N\$（样本数序号），

\$m=2, 3, 4, \dots\$，（例三年滑动平均，则 \$m=3\$）。

例1—6 某地欲进行9月份降水趋势预报，试用滑动平均对降水变化趋势进行分析。某地降水资料如表1—1第一行所示。按公式 (1.6) 进行三年滑动平均 (\$m=3\$)。

$$\bar{x}_{3i} = \frac{1}{3} \sum_{k=i}^{i+(3-1)} x_k \quad (i=1, 2, \dots, N)$$

计算的三年滑动平均结果，列于表1—1的第二行；五年滑动平均 (\$m=5\$)，按公式 (1.6)

$$\bar{x}_{5i} = \frac{1}{5} \sum_{k=i}^{i+(5-1)} x_k \quad (i=1, 2, \dots, N)$$

进行计算，其结果列于表1—1中第三行。具体计算如下：

三年滑动平均：

$$\bar{x}_{31} = \frac{1}{3}(58 + 69 + 62) = 63.0;$$

$$\bar{x}_{32} = \frac{1}{3}(69 + 62 + 70) = 67.0;$$

$$\bar{x}_{33} = \frac{1}{3}(62 + 70 + 65) = 65.7;$$

\$\vdots\$

五年滑动平均：

$$\bar{x}_{51} = \frac{1}{5}(58 + 69 + 62 + 70 + 65) = 64.8;$$

$$\bar{x}_{52} = \frac{1}{5}(69 + 62 + 70 + 65 + 76) = 68.4;$$

\$\vdots\$

表1—1 某地1961~1980年降水量及滑动平均

降水量 (mm)	58	69	62	70	65	76	70	78	72	80	70
三年滑动平均		63.0	67.0	65.7	70.3	70.3	74.7	73.3	76.7	74.0	75.0
五年滑动平均			64.8	68.4	68.6	71.8	72.2	75.2	74.0	75.0	71.4

降水量 (mm)	75	60	68	59	66	57	62	55	67
三年滑动平均	68.3	67.7	62.3	64.3	60.7	61.7	58.0	61.3	
五年滑动平均	70.6	66.4	65.6	62.0	62.4	59.8	61.4		

由图1—1可以看出无论三年或五年滑动平均, 均从1976年开始有回升趋势, 但由于已是降至最低点后刚刚回升, 不可能回升很快, 故可估计1981年降水量为正常趋势。

由图1—1还可以看出间隔数 m 越大, 曲线越平滑, 主要趋势看得越清楚, 但也不能无限增大, 因滑动平均结果总要损失一些样本, m 年滑动平均要损失 $m-1$ 个样本。

5. 二次均值计算: 在数字计算机大量处理数据时, 为克服舍、入误差所造成的影响, 提高均值精度, 可进行二次均值计算:

$$\text{二次均值} \quad \bar{x}^{(2)} = \bar{x}^{(1)} + \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x}^{(1)}) \quad \text{当} \sum_{i=1}^N (x_i - \bar{x}^{(1)}) \neq 0 \text{时} \quad (1.7)$$

$$\bar{x}^{(1)} = \frac{1}{N} \sum_{i=1}^N x_i$$

(二) 加权算术平均数的计算

简单算术平均数, 是将每个变数看作等权的情形; 如果每个变数权重不同时, 则要用加权法计算平均数。其公式为:

$$\bar{x} = \frac{1}{\sum f_i} \sum_{i=1}^h f_i x_i \quad (1.8)$$

f_i 为第 i 个样本的权重 (频数); x_i 为第 i 个随机变数。

例1—7 某地100年年降水量记录如表1—2, 求其平均值。

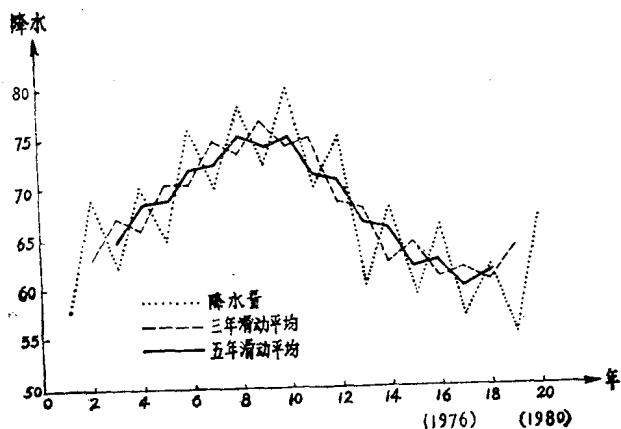


图1—1 三年和五年滑动平均