

中长期水文预报与

SPSS 应用

旦木仁加甫 编著

ZHONGCHANGQI SHUIWEN YUBAO YU
SPSS YINGYONG



黄河水利出版社

中长期水文预报与 SPSS 应用

旦木仁加甫 编著

黄河水利出版社

· 郑州 ·

内 容 提 要

本书系统地介绍了基于 SPSS Statistics 17.0 中文版环境平台的有关中长期水文预报的方法、统计检验及分析计算过程,并通过 SPSS 应用实例,详细介绍了每一种中长期水文预报方法的 SPSS 实现过程和输出结果。主要内容有:SPSS 简介与数据文件的建立,周期均值叠加,一元线性回归分析,多元线性回归分析,逐步回归分析,后向逐步剔除回归分析,前向逐步引入回归分析,强迫剔除回归分析,加权最小二乘回归分析,曲线参数估计法,平稳时间序列分析,非平稳序列逐步回归趋势分析,非平稳序列逐步回归周期分析,枯季退水曲线分析,含分类自变量的回归分析,二值 Logistic 回归分析,多项 Logistic 回归分析,有序回归分析,非线性回归分析等。本书是一本实用性、可操作性和可读性紧密结合的、探索和研究中长期水文预报方法与 SPSS 应用技术的参考书和工具书。

本书可供从事中长期水文预报与 SPSS 应用的工程技术、教学和研究人員阅读、使用,也可作为高等学校水文学及水资源专业的上机实习教材,还可供其他领域从事中长期预报和 SPSS 应用的人员参考。

图书在版编目(CIP)数据

中长期水文预报与 SPSS 应用/旦木仁加甫编著. —郑州:黄河水利出版社,2011.6

ISBN 978-7-5509-0069-1

I. ①中… II. ①旦… III. ①水文预报:中期预报—统计分析—软件包,SPSS ②水文预报:长期预报—统计分析—软件包,SPSS IV. ①P338

中国版本图书馆 CIP 数据核字(2011)第 119294 号

组稿编辑:王路平 电话:0371-66022212 E-mail: hhslwlp@126.com

出版社:黄河水利出版社

地址:河南省郑州市顺河路黄委会综合楼 14 层 邮政编码:450003

发行单位:黄河水利出版社

发行部电话:0371-66026940,66020550,66028024,66022620(传真)

E-mail: hhslcbs@126.com

承印单位:河南地质彩色印刷厂

开本:787 mm × 1 092 mm 1/16

印张:13.5

字数:310 千字

版次:2011 年 6 月第 1 版

印数:1—1 700

印次:2011 年 6 月第 1 次印刷

定价:39.00 元

前 言

中长期水文预报是根据前期水文气象要素,用天气学、数理统计、宇宙-地球物理分析等方法,对未来较长时间的水文要素进行科学预测的一门边缘学科。近几年来,中长期水文预报在理论研究和生产实践方面有了长足的进展,但由于影响因素的复杂性和科学技术水平的限制,目前仍处于探索、发展阶段。通常把预见期为3~15 d的称为中期预报,15 d至1年以内的称为长期预报,1年以上的称为超长期预报。对径流预报而言,预见期超过流域最大汇流时间的即为中长期预报。

SPSS是目前世界上最流行的著名统计软件之一,被广泛应用于社会科学和自然科学的各个领域。在国际学术界有条不紊的规定,即在国际学术交流中,凡是用SPSS软件完成的计算和统计分析,可以不必说明算法,由此可见其影响之大和信誉之高。

SPSS的基本功能包括数据管理、统计分析、图表分析和输出管理,可为中长期水文预报提供崭新的操作环境平台和丰富的技术方法。SPSS基本内容包括描述性统计等十几个大类,每个类包括多种统计过程(即模块),每个统计过程又允许用户选择不同的方法及参数。SPSS可根据数据绘制各种图形,还以图表和相关信息的形式输出统计分析结果。

目前,在国内市面上还未见到结合中长期水文预报方法、生产实践及现行《水文情报预报规范》(GB/T 22482—2008)而编著成书的SPSS应用专著,鉴于此,作者结合自己25年的基层水文水资源工作实践,精心编著了本书。本书系统地介绍了基于SPSS Statistics 17.0中文版环境平台的有关中长期水文预报的方法、统计检验、分析计算过程及SPSS应用实例。全书共分19章,主要内容有:SPSS简介与数据文件的建立,周期均值叠加,一元线性回归分析,多元线性回归分析,逐步回归分析,后向逐步剔除回归分析,前向逐步引入回归分析,强迫剔除回归分析,加权最小二乘回归分析,曲线参数估计法,平稳时间序列分析,非平稳序列逐步回归趋势分析,非平稳序列逐步回归周期分析,枯季退水曲线分析,含分类自变量的回归分析,二值Logistic回归分析,多项Logistic回归分析,有序回归分析,非线性回归分析等。

全书主要特点:

(1)第1章是SPSS简介与数据文件的建立,是与中长期水文预报有关的SPSS应用的预备知识。

(2)第2~19章分别介绍了18种中长期水文预报方法与SPSS应用,每章只介绍一种方法,基本按照中长期水文预报方法(基本思路、计算公式或回归方程)、统计检验、分析计算过程、SPSS应用实例等步骤来介绍。

(3)上述步骤中,分析计算过程是按照我国现行《水文情报预报规范》(GB/T 22482—2008)的要求编写的,而SPSS应用实例是按照分析计算过程来介绍的。在每个应用实例中,还详细介绍了相应中长期水文预报方法的SPSS实现过程和输出结果,有些实

例中说明了注意事项,有些还提出了建设性的讨论意见。

(4)一部分中长期水文预报方法有可能在不同章节中重复应用,遇此情形时除内容有所变动外,只指明了首次介绍过该方法的章节以供参阅。

(5)应用实例介绍与启用了打开数据、新建查询、单因素 ANOVA、计算变量、线图、线性回归、散点/点状、曲线估计、自相关、创建模型、均值、排序个案、描述、重新编码为不同变量、个案汇总、频率、二元 Logistic 回归、多项 Logistic 回归、有序回归、非线性回归等 20 个 SPSS 模块,其中一部分模块在不同章节中可能会重复启用,此时除需要选择不同的方法及参数外,也只指明了首次介绍过该模块的章节以供参阅。

(6)一种中长期水文预报方法可由不同的 SPSS 模块来实现,书中除枯季退水曲线分析外,其他方法一般以一个 SPSS 模块为主,另外几个 SPSS 模块为辅。

(7)中长期水文预报技术方法有所创新,如作者首次提出了有关枯季退水曲线分析的历年开始消退值三分位数分组法,经 SPSS 应用,效果很好;再如利用序列平均值和距平值构建分类变量的做法,将会启迪和拓展中长期水文预报的建模思路。

(8)SPSS 结果输出和结果分析中的有关表格均用中英文予以对照,以方便 SPSS Statistics 英文版本的用户。

(9)每一种中长期水文预报方法既可由 SPSS 来建立相应的中长期水文预报方案,又可由 SPSS 来验证相应方案。

(10)SPSS 的应用为中长期水文预报提供了崭新的操作环境平台,即从数据管理、统计分析、图表分析到输出管理自成系统。

(11)SPSS 的应用为中长期水文预报提供了丰富的技术方法,如后向逐步剔除回归分析、前向逐步引入回归分析、强迫剔除回归分析、加权最小二乘回归分析、含分类自变量的回归分析、二值 Logistic 回归分析、多项 Logistic 回归分析、有序回归分析等。

(12)SPSS 是国际上公认的著名统计软件之一,所以用 SPSS 实现中长期水文预报,有利于与国际通用统计学方法的接轨。

总之,本书是一本实用性、可操作性和可读性紧密结合的,探索和研究中长期水文预报方法与 SPSS 应用技术的参考书和工具书,可供从事中长期水文预报与 SPSS 应用的工程技术、教学和研究人员阅读、使用,也可作为高等学校水文学及水资源专业的上机实习教材,还可供其他领域从事中长期预报和 SPSS 应用的人员参考。

由于中长期水文预报及 SPSS 应用技术发展迅速,加上作者水平有限,书中难免有不足之处,敬请广大读者批评指正。反馈意见请发电子邮件至:dmrjf@sina.com。

且木仁加甫

2011 年 3 月 18 日于新疆库尔勒

目 录

前 言

第 1 章 SPSS 简介与数据文件的建立	(1)
1.1 SPSS 简介	(1)
1.2 数据文件的建立	(4)
第 2 章 周期均值叠加	(9)
2.1 基本思路、 F 检验与分析计算过程	(9)
2.2 SPSS 应用实例	(10)
第 3 章 一元线性回归分析	(25)
3.1 回归方程、统计检验与分析计算过程	(25)
3.2 SPSS 应用实例	(26)
第 4 章 多元线性回归分析	(35)
4.1 回归方程、统计检验与分析计算过程	(35)
4.2 SPSS 应用实例	(37)
第 5 章 逐步回归分析	(43)
5.1 基本思路、计算公式、统计检验与分析计算过程	(43)
5.2 SPSS 应用实例	(45)
第 6 章 后向逐步剔除回归分析	(52)
6.1 基本思路、计算公式、统计检验与分析计算过程	(52)
6.2 SPSS 应用实例	(53)
第 7 章 前向逐步引入回归分析	(59)
7.1 基本思路、计算公式、统计检验与分析计算过程	(59)
7.2 SPSS 应用实例	(60)
第 8 章 强迫剔除回归分析	(67)
8.1 基本思路、计算公式、统计检验与分析计算过程	(67)
8.2 SPSS 应用实例	(68)
第 9 章 加权最小二乘回归分析	(76)
9.1 回归方程、统计检验与分析计算过程	(76)
9.2 SPSS 应用实例	(77)
第 10 章 曲线参数估计法	(87)
10.1 基本思路、统计检验与分析计算过程	(87)
10.2 SPSS 应用实例	(89)
第 11 章 平稳时间序列分析	(98)
11.1 建立自回归移动平均方程与分析计算过程	(98)

11.2	SPSS 应用实例	(99)
第 12 章	非平稳序列逐步回归趋势分析	(110)
12.1	基本思路、计算公式、统计检验与分析计算过程	(110)
12.2	SPSS 应用实例	(111)
第 13 章	非平稳序列逐步回归周期分析	(121)
13.1	基本思路、计算公式、统计检验与分析计算过程	(121)
13.2	SPSS 应用实例	(122)
第 14 章	枯季退水曲线分析	(135)
14.1	基本思路与分析计算过程	(135)
14.2	SPSS 应用实例	(136)
第 15 章	含分类自变量的回归分析	(151)
15.1	回归方程、统计检验与分析计算过程	(151)
15.2	SPSS 应用实例	(152)
第 16 章	二值 Logistic 回归分析	(163)
16.1	回归方程、统计检验与分析计算过程	(163)
16.2	SPSS 应用实例	(166)
第 17 章	多项 Logistic 回归分析	(177)
17.1	回归方程、统计检验与分析计算过程	(177)
17.2	SPSS 应用实例	(178)
第 18 章	有序回归分析	(187)
18.1	回归方程、统计检验与分析计算过程	(187)
18.2	SPSS 应用实例	(188)
第 19 章	非线性回归分析	(196)
19.1	基本思路、统计检验与分析计算过程	(196)
19.2	SPSS 应用实例	(197)
参考文献		(210)

第 1 章 SPSS 简介与数据文件的建立

1.1 SPSS 简介

1.1.1 SPSS Statistics 概述

SPSS 是“Statistical Package for Social Sciences”(社会科学统计软件包)的缩写。1968 年,3 位美国斯坦福大学的研究生开发了最早的 SPSS 统计软件系统,1975 年,在此基础上于芝加哥联合成立了 SPSS 公司。1984 年,SPSS 总部推出了世界上第一个统计分析软件微机版本 SPSS/PC+,极大地扩充了它的应用范围。随着 SPSS 产品服务领域的扩大和服务程度的加深,SPSS 公司于 2000 年正式将英文全称更改为“Statistical Product and Service Solutions”,意思是“统计产品与服务解决方案”。目前,该系统已经发展成为世界上最流行的著名统计软件之一,被广泛应用于社会科学和自然科学的各个领域。在国际学术界有条不成文的规定,即在国际学术交流中,凡是用 SPSS 软件完成的计算和统计分析,可以不必说明算法,由此可见其影响之大和信誉之高。

SPSS 的基本功能包括数据管理、统计分析、图表分析和输出管理。SPSS 基本内容主要有描述性统计、均值比较、相关分析、回归分析、聚类分析、判别分析、因子分析、主成分分析、可靠性分析、时间序列分析、生存分析等十几个大类,每个类包括多种统计过程(即模块),例如回归分析中又分线性回归分析、曲线估计、Logistic 回归等统计过程,每个统计过程又允许用户选择不同的方法及参数。SPSS 专有的绘图系统,可以根据数据绘制各种图形。SPSS 还以图表和相关信息的形式输出统计分析结果。

本书以 SPSS Statistics 17.0 中文版为环境平台,结合中长期水文预报与 SPSS 应用而撰写,书中的 SPSS 结果输出和结果分析中的有关表格均用中英文予以对照,以方便 SPSS Statistics 英文版本的用户。

1.1.2 SPSS 的启动、主界面和退出

1.1.2.1 SPSS 启动

单击 Windows 的“开始”按钮,在“所有程序”菜单项“SPSS Statistics”中找到子菜单“SPSS Statistics”并单击,或双击 Windows 桌面上“SPSS Statistics”快捷方式,即可启动 SPSS。

1.1.2.2 SPSS 数据编辑窗口

SPSS 主界面主要有两个,一个是 SPSS 数据编辑窗口,另一个是 SPSS 结果输出窗口。本节先介绍数据编辑窗口。

SPSS 数据编辑窗口与微软公司的 Excel 比较相似,但是 SPSS 的数据统计功能比

Excel 强很多。

数据编辑窗口由标题栏、菜单栏、工具栏、变量名栏、数据输入区、数据编辑区、标尺栏、显示区滚动条、窗口切换标签和状态栏组成,见图 1-1 和图 1-2。

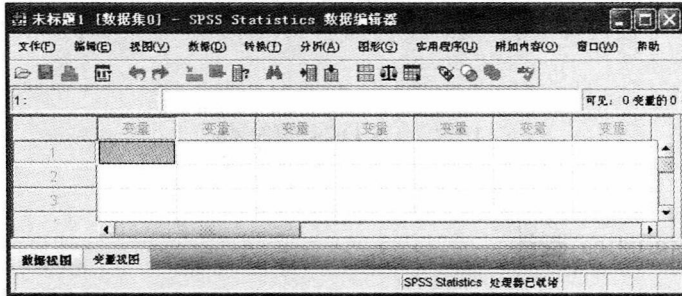


图 1-1 数据编辑窗口(数据视图)

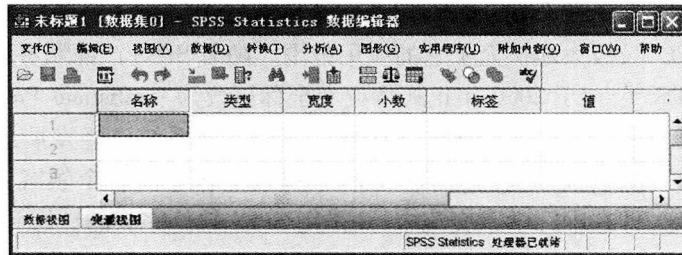


图 1-2 数据编辑窗口(变量视图)

本节主要介绍数据编辑窗口的菜单栏和窗口切换标签,其他内容请参阅 SPSS 专业书籍。

菜单栏列出了 SPSS 的命令菜单,每个菜单对应一组相应的功能。例如,“文件”菜单用来实现有关文件的调入、存储、显示和打印等;“编辑”菜单用来实现有关文件内容的选择、复制、剪贴、寻找和替换等;“视图”菜单用来对数据编辑窗口的各栏目是否显示进行选择;“数据”菜单用来实现数据变量定义,数据格式选定,观察对象的选择、排序、加权,数据文件的转换、连接、汇总等;“转换”菜单用来实现有关数据的计算、重新赋值、缺失值替代等;“分析”菜单用一系列的统计方法输出分析数据、图表等;“图形”菜单用来绘制相关统计图;“实用程序”菜单供用户进行命令解释、字体选择、文件信息、定义输出标题、窗口设计等;“附加内容”菜单供用户应用 SPSS 的辅助软件进行深入分析;“窗口”菜单用来实现对窗口的管理;“帮助”菜单帮助用户调用、查询、显示帮助文件等。

在数据编辑窗口下方有两个标签:“数据视图”和“变量视图”,即窗口切换标签。“数据视图”对应的表格用于查看、录入和修改数据,见图 1-1;“变量视图”对应的表格用于输入和修改变量的定义,见图 1-2。

在数据编辑窗口中完成变量定义、数据输入后,单击某个统计功能菜单,SPSS 会自动完成统计分析,并弹出结果输出窗口,其中存放了数据统计的结果。

1.1.2.3 SPSS 结果输出窗口

SPSS 结果输出窗口是显示和管理 SPSS 统计分析结果、报表、图形及相关信息的窗

口,见图 1-3。

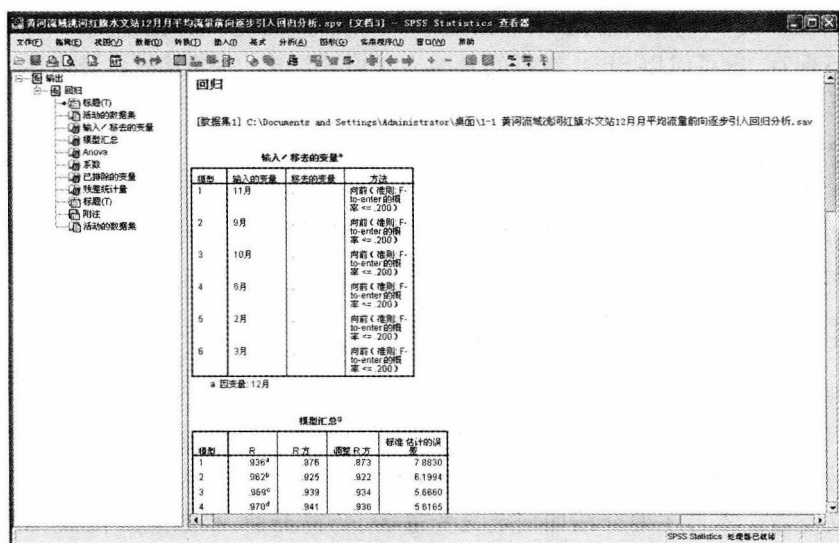


图 1-3 结果输出窗口

结果输出窗口左边部分是索引输出区,右边部分是详解输出区。可以对详解输出区中的统计表格进行编辑等操作。

结果输出窗口中的内容可以以文件类型 *. spo 的形式保存,也可以以 *. doc 的形式导出。如果用户要单独保存统计表格或图形,应选中目标并单击鼠标右键,选择复制命令,粘贴到用户所需的文件中即可(表格或图形的输出格式必须是 SPSS 所支持的)。

1.1.2.4 退出 SPSS

依次单击菜单“文件→退出”,或单击标题栏上关闭按钮即可退出 SPSS。

1.1.3 SPSS 统计分析基本步骤

SPSS 统计功能强大,但操作简单,其基本步骤归纳如下。

1.1.3.1 数据输入

将数据以电子表格的方式输入到 SPSS 中,也可以从其他可转换的数据文件中导入数据。数据输入主要有两项工作,一是定义变量,二是输入变量值。

1.1.3.2 数据预分析

输入完数据后,根据统计分析工作的需要,用户要对数据进行必要的预分析,包括数据的编辑、加工、基本统计特性描述等,以保证后续统计分析工作的有效性。

1.1.3.3 统计分析

按照统计分析工作要求和数据基本特点确定统计分析方法,并进行相应的统计分析。

1.1.3.4 统计结果可视化

统计分析完成后,SPSS 会自动生成一系列可视化的数据、统计图表及相关信息,用户可根据统计分析工作和数据基本特点的需求来选择使用。

1.1.3.5 保存和导出统计分析结果

将数据编辑窗口中的数据、结果输出窗口中的统计图表及相关信息以 SPSS 自带的文件格式进行保存,也可以以 SPSS 许可的其他文件格式导出,以供其他系统使用。

1.2 数据文件的建立

在使用 SPSS 软件进行数据分析时,首先要建立数据文件。一个数据文件的建立主要包括定义变量、数据输入、数据文件的保存等内容。

1.2.1 定义变量

输入数据前要定义变量,包括定义变量名、变量类型、变量宽度、小数点后位数、变量名标签、变量值标签、变量缺失值、变量列宽、变量对齐格式及度量标准等信息。

单击数据编辑窗口下方的“变量视图”标签,见图 1-2,在此界面即可按照 SPSS 的定义规则定义变量。本节主要介绍如何定义变量名、变量类型、小数点后位数、变量名标签及变量度量标准的属性,其他内容请参阅 SPSS 专业书籍。

1.2.1.1 变量名

(1)变量名总长度不能超过 64 个字符,即最多可容纳 32 个汉字或 64 个英文字母,长变量名将在结果输出时被分为多行显示。

(2)变量名必须以字母、汉字或@ 开头,英文字母不区分大小写,不能以圆点结尾。

(3)ALL、AND、BY、EQ、GE、GT、LE、LT、NE、NOT、OR、TO、WITH 等 SPSS 内部特有的字符不能作为变量名。

(4)SPSS 默认的变量名是以 VAR 开头,后面补足 5 位数字,如 VAR00001。

(5)变量名必须是唯一的,不允许与其他变量名重名。

(6)变量名避免使用下划线结尾,以免与 SPSS 某些程序自动生成的变量名发生冲突。

如果变量名不符合其命名规则,SPSS 会自动给出错误提示信息。

1.2.1.2 变量类型

SPSS 中,变量类型有 8 种,见图 1-4,其中数值型、日期型、字符串型是常用的基本变量类型。

数值型是系统默认的标准数据类型,默认的最大数值显示宽度是 8 位,小数点后位数是 2 位,见图 1-4。用户可以对数值显示宽度和小数点后位数进行修改,这不会影响实际数据的存储,也不影响数据的计量。

日期型数据主要用来表示日期或不同格式的时间,有很多显示格式,见图 1-5,用户可以根据需要进行选择。

字符串型数据的默认最大显示宽度是 8 个字符位,见图 1-6,它不能进行算术运算,但区分英文字

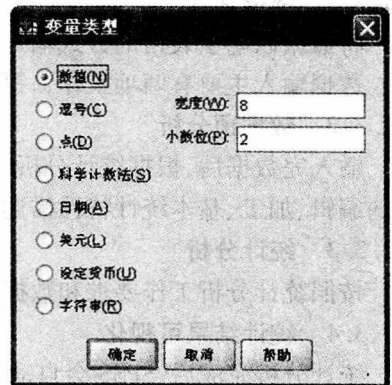


图 1-4 变量类型对话框(数值型)

母大小写,也可支持文字数字混排。



图 1-5 变量类型对话框(日期型)

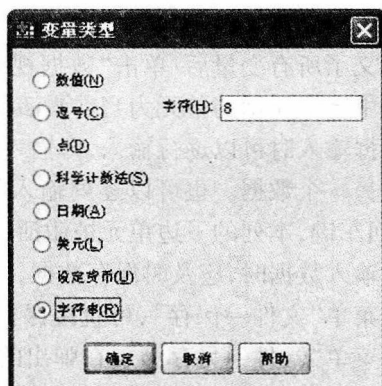


图 1-6 变量类型对话框(字符串型)

1.2.1.3 小数点后位数

小数点后位数用来确定类似数值型变量的小数点后的位数,见图 1-7,可在对应的单元格内输入正整数或单击上下箭头来确定小数点后位数。

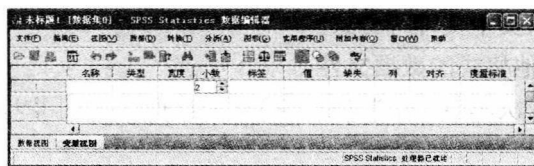


图 1-7 小数点后位数对话框

1.2.1.4 变量名标签

通过在变量名标签内输入文字,可以对变量含义进行详细的说明,从而提高变量名的可视性和分析结果的可读性。

1.2.1.5 变量度量标准

SPSS 给出的变量度量标准有度量尺度、有序尺度和名义尺度 3 种,见图 1-8。度量尺度是定距或定比的数值型变量的度量标准。有序尺度是包含一定次序的描述性分类变量的度量标准,有序变量可为字符串型或数值型,并进行明确的分类。名义尺度是无序分类变量的度量标准,名义变量可为字符串型或有明确注解的数值变量。

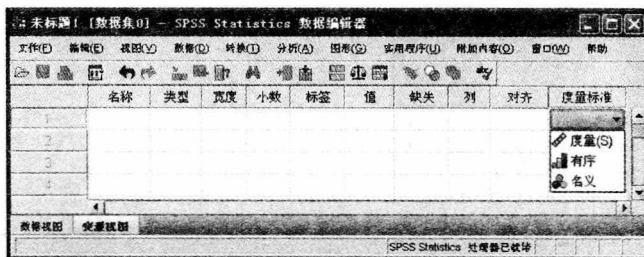


图 1-8 变量度量标准对话框

1.2.2 数据输入与保存

定义了所有变量后,单击“数据视图”标签,就可以在数据视图中输入数据。数据编辑窗口中亮光显示的单元为当前激活的单元,意味着在该单元中可以输入或修改数据。

数据输入时可以逐行输入,即输入完一个数据后,按 Tab 键,本行的右边单元被激活,再输入另一个数据。也可以逐列输入数据,也就是按照变量输入数据,输入完一个数据后,按回车键,本列的下边单元被激活,再输入另一个数据。

在输入数据时,应及时保存数据,防止数据的丢失。保存步骤为:在数据编辑窗口依次单击菜单“文件→保存”,可直接保存为 SPSS 默认的数据文件格式(*.spo);或者,依次单击菜单“文件→另存为”,在弹出的数据保存对话框中,确定保存数据文件的路径、文件名以及文件格式后单击“保存”按钮即可。SPSS 支持的数据文件保存格式很多,用户可以根据自己的需要进行选择。

1.2.3 建立数据文件的其他方式

建立数据文件的方式除前面介绍的人工输入数据外,还有打开其他格式的数据文件、使用数据库查询和导入文本文件等方式。

1.2.3.1 直接打开其他格式的数据文件

在数据编辑窗口依次单击菜单“文件→打开→数据”,见图 1-9,弹出如图 1-10 所示的打开数据对话框。

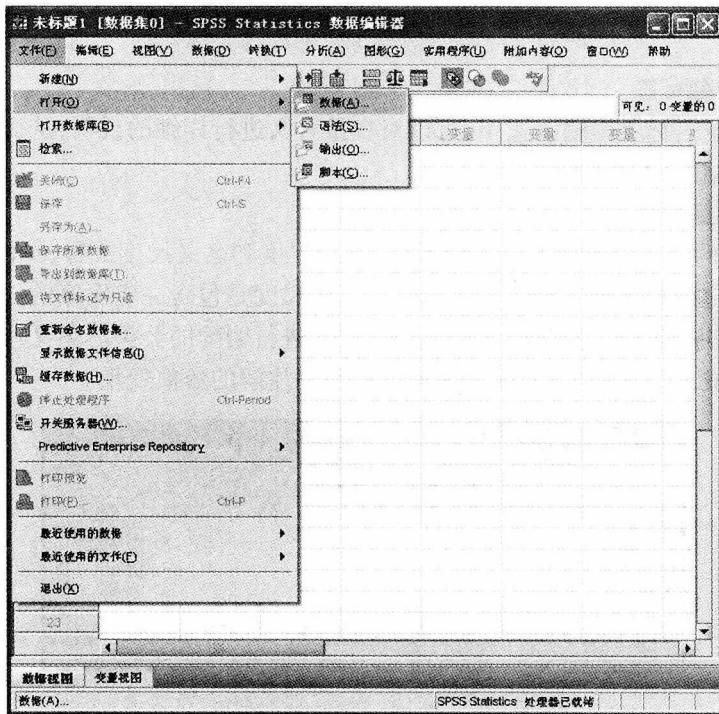


图 1-9 SPSS 打开数据模块



图 1-10 打开数据对话框

在图 1-10 中可以打开不同格式的数据文件,SPSS 可以打开的数据文件类型有:SPSS (.sav、.sys、.syd、.por)、Excel(.xls)、Lotus(.w)、Text(.txt)等。

1.2.3.2 使用数据库查询建立数据文件

在数据编辑窗口依次单击菜单“文件→打开数据库→新建查询”,见图 1-11,弹出如图 1-12所示的数据库向导对话框。



图 1-11 SPSS 新建查询模块

在图 1-12 中选中所需的数据源,根据向导的提示即可将数据导入 SPSS。

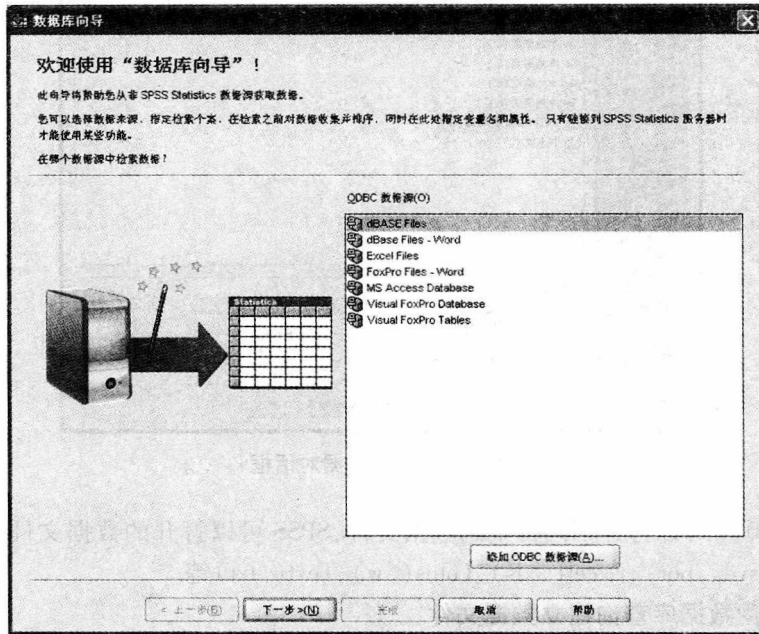


图 1-12 数据库向导对话框

1.2.3.3 导入文本文件建立数据文件

在数据编辑窗口依次单击菜单“文件→检索”,弹出与图 1-10 一样的打开数据对话框,从中可以将选中的文本文件(.txt)导入 SPSS。

第2章 周期均值叠加

2.1 基本思路、 F 检验与分析计算过程

2.1.1 基本思路

一个随时间变化的等时距水文要素观测样本,可以看成是有限个不同周期波叠加而成的过程。从样本序列中识别周期时,可以将序列分成若干组,当分组组数等于客观存在的周期长度时,组内各个数据的差异小,而组间各个数据的差异大;反之,如果组间差异显著大于组内差异,序列就存在周期,其长度就是组间差异最大而组内差异最小的分组组数。通常一个序列的总体差异是固定的,组间差异增大,组内差异则减小。那么,组内差异比组间差异小到什么程度才算是显著呢?通常是用 F 检验来进行判断。

2.1.2 F 检验

设某水文要素随时间变化的等时距样本序列为 X_1, X_2, \dots, X_n , 排成表 2-1 的形式,其中 $j=1, 2, \dots, b$, 表示分为 b 组, $b=2, 3, \dots, m$ (当样本数 n 为偶数时, $m=n/2$; 当 n 为奇数时, $m=(n-1)/2$); 就是说,可能存在的周期数 b 为 $2, 3, \dots, m$ 。 i 为每组含有的项数, $i=1, 2, \dots, a$, 表示每组有 a 个数据。 T_j 为每组的合计数, \bar{X}_j 为每组的组平均值。对于不同的 b , 可算得相应的方差比 F :

表 2-1 试验周期分组排列

i		试验周期分组(j)			
		1	2	...	b
每组项数	1	X_{11}	X_{12}	...	X_{1b}
	2	X_{21}	X_{22}	...	X_{2b}
	\vdots	\vdots	\vdots		\vdots
	a	X_{a1}	X_{a2}	...	X_{ab}
	T_j	T_1	T_2	...	T_b
	\bar{X}_j	\bar{X}_1	\bar{X}_2	...	\bar{X}_b
	$T_j \times T_j$	$T_1 \times T_1$	$T_2 \times T_2$...	$T_b \times T_b$
	$T_j \times T_j / a_j$	$T_1 \times T_1 / a_j$	$T_2 \times T_2 / a_j$...	$T_b \times T_b / a_j$

$$F = (S_1/f_1)/(S_2/f_2)$$

式中: S_1 为组间离差平方和, $S_1 = \sum (T_j \times T_j/a - T \times T/n)$, $T = \sum T_j$; S_2 为组内离差平方和, $S_2 = \sum \sum (X_{ij} \times X_{ij}) - \sum (T_j \times T_j/a)$; $f_1 = b - 1$, 为对应于 S_1 的自由度; $f_2 = n - b$, 为对应于 S_2 的自由度。

当 b 分别取 2, 3, \dots , m 时, 可计算得 $m - 1$ 个不同的 F 值, SPSS 会自动给出其相伴概率 ρ 值, 用户可挑选最小的 ρ 值, 与给定的显著性水平 α (即信度) 值相比较。

$\rho < \alpha$, 则表明在这一信度水平上, 差异显著, 有周期存在, 所对应的分组组数 b 即为周期长度, 各组平均值即为对应于 b 的周期波振幅; $\rho \geq \alpha$, 则表明在这一信度水平上, 差异不显著, 不存在周期。

2.1.3 分析计算过程

(1) 建立由等时距水文观测变量序列组成的 SPSS 数据文件并保存。

(2) 打开 SPSS 数据文件, 在数据编辑窗口构建分组变量: 确定等时距水文观测变量分组组数 b ($b = 2, 3, \dots, m$), 每组定义一个相应变量; 变量取值很简单, 如 $b = 3$ 时, 依序取值为 1, 2, 3, 1, 2, 3 \dots , 排列至观测变量序列终止时刻 (或时段)。

(3) 接着在数据编辑窗口识别周期。

从观测变量序列识别第一周期波, 若通过 F 检验, 按周期长度 b 将观测变量各组平均值从序列开始时刻 (或时段) 排列至终止时刻 (或时段), 构成第一周期波序列。从观测变量序列中剔除第一周期波序列, 生成新序列。

从新序列识别第二周期波, 若通过 F 检验, 将新序列各组平均值按上述方法排序, 构成第二周期波序列。从新序列中剔除第二周期波序列, 生成另一个新序列。

其余周期波的识别也以此类推, 直到不能识别或者不想识别周期为止。

(4) 最后在数据编辑窗口对所识别的各周期波序列进行叠加, 计算观测变量与相应周期波叠加值之间的相对拟合误差, 生成相应的历史拟合曲线图, 保存数据编辑窗口中的数据和结果输出窗口中的统计结果、相关信息等。

(5) 进行预报: 外延叠加值即为预报值。

2.2 SPSS 应用实例

本次选用笔者编著的《常用中长期水文预报 Visual Basic 6.0 应用程序及实例》中列举的浙江省钱塘江流域金华江金华水文站 1952 ~ 1973 年年最高水位序列, 用 SPSS 进行了周期均值叠加分析和预报, 两者结果完全一致。

2.2.1 构建分组变量

金华江金华水文站 1952 ~ 1973 年年最高水位序列见图 2-1, 可见, 其样本容量 $n = 22$, 为偶数, 所以 $m = n/2 = 11$, 可以生成分组组数 b 为 2, 3, \dots , 11 的 10 个分组变量。

打开 SPSS 数据文件, 在数据编辑窗口依序定义分组变量名为“分组 2”, “分组 3”, \dots , “分组 11”, 再按照 2.1.3 部分的方法给各变量输入值, 构建的分组变量见图 2-2。