

IBM SPSS Forecasting 26

The IBM logo, consisting of the letters 'IBM' in a bold, black, sans-serif font. Each letter is composed of eight horizontal bars of equal height and width, stacked vertically to form the characters.

注释

使用本信息及其支持的产品前，请先阅读 第 7 页的『通知』中的信息。

产品信息

本版本适用于 IBM® SPSS Statistics V26.0.0 及所有后续发行版和修订版，直到新版本中另有声明为止。

目录

预测(T)	1	谱图	4
时间序列简介	1	SPECTRA 命令其他功能	5
时间序列数据	1	通知	7
数据转换.	2	商标	8
估计期和验证期	2	索引	11
季节性分解	2		
季节性分解: 保存.	3		
SEASON 命令其他功能.	3		

预测(T)

SPSS® Statistics Professional Edition 或"预测"选项中包含以下预测功能。

时间序列简介

时间序列是可通过定期度量时间段中某个变量获得的一组观察值。例如，在一系列库存数据中，观察值可能表示若干个月的日常库存水平。显示产品市场份额的序列可能由过去几年来的每周市场份额组成。总销售额序列可能包含许多年来每个月的一个观察值。这些示例中的每个示例都有一个共同点，即都曾在一定时间长度内的已知间隔定期观察某个变量。因此，典型时间序列的数据形式是表示定期间隔内获得的测量值的单个观察值序列或观察值列表。

表 1. 每日库存数据序列

时间	星期	日期	库存水平
t ₁	1	星期一	160
t ₂	1	星期二	135
t ₃	1	星期三	129
t ₄	1	星期四	122
t ₅	1	星期五	108
t ₆	2	星期一	150
		...	
t ₆₀	12	星期五	120

进行时间序列分析的最重要原因之一是尝试预测序列的未来值。解释过去值的序列的模型还可以预测下几个值是增加还是减少，以及增减的幅度有多大。显然，成功进行此类预测的能力对于任何公司或科学领域来说都非常重要。

时间序列数据

基于列的数据

每个时间序列字段都包含单个时间序列的数据。此结构是"季节性分解"过程和"谱分析"过程所使用的传统时间序列数据结构。例如，要在数据编辑器中定义时间序列，请展开变量列表，然后在任意空白行输入变量名。时间序列中的每个观察值都对应一个个案（数据编辑器中的一行）。

如果您打开一个包含时间序列数据的电子表格，那么每个序列都应该安排在该电子表格的一列中。如果已将时间序列安排在电子表格的行中，则可以采用任意方式打开该电子表格，并使用数据 > 变换...，将行变换成列。

多维数据

对于多维数据，每个时间序列字段都包含多个时间序列的数据。这样，特定字段内的单独时间序列将由类别字段（称为维度字段）的一组值标识。

例如，不同区域和品牌的销售数据可能存储在一个 *sales* 字段中，因此本示例中的维度为 *region* 和 *brand*。*region* 和 *brand* 的每个组合都标识 *sales* 的一个特定时间序列。例如，在下表中，*region* 为 "north" 且 *brand* 为 "brandX" 的记录定义单个时间序列。

表 2. 多维数据

date	region	brand	sales
01/01/2014	north	brandX	82350
01/01/2014	north	brandY	86380
01/01/2014	south	brandX	91375
01/01/2014	south	brandY	70320
01/02/2014	north	brandX	83275
01/02/2014	north	brandY	85260
01/02/2014	south	brandX	94760
01/02/2014	south	brandY	69870

数据转换

Core System 中提供的许多数据转换过程在进行时间序列分析时非常有用。这些转换仅适用于基于列的数据，这些数据中的每个时间序列字段都包含单个时间序列的数据。

- 定义日期过程（位于"数据"菜单中）生成日期变量，这些变量用于确定季节性并区分历史期、验证期和预测期。"预测"可以与由"创建日期"过程创建的变量一起使用。
- 创建时间序列过程（在"转换"菜单上）创建新的时间序列变量作为现有时间序列变量的函数。它包括使用相邻观察值来求平滑值、平均值或差分值的函数。
- 替换缺失值过程（在"转换"菜单上）将系统缺失值和用户缺失值替换为根据多种方法之一估计的值。序列开头和末尾的缺失数据不会引发特殊的问题，只会缩短序列的有效长度。序列中间的缺口（内嵌缺失数据）是更为严重的问题。

有关时间序列的数据转换的详细信息，请参阅《Core System 用户指南》。

估计期和验证期

通常，将时间序列划分成估计或历史期和验证期非常有用。在估计（历史）期的观察值的基础上建立一个模型，然后在验证期对其进行测试，以了解其性能。通过使用该模型对已知的点（验证期中的点）进行预测，即可了解该模型在预测方面的性能如何。

验证期的个案通常指保持个案，因为它们是从模型构建过程中保留下来的。在对该模型满意之后，就可以进行充分的预测作业了，您可以重新定义估计期以包含保持个案，然后建立最终模型。

季节性分解

"季节性分解"过程可将一个序列分解成一个季节性成分、一个组合趋势和循环的成分和一个"误差"成分。此过程是对统计方法 I（也称为比率与移动平均值方法）的实现。

示例 科学家想要对特定气象站的臭氧层每月测量结果进行分析。目标是确定数据中是否存在任何趋势。为了揭示真实趋势，由于季节性影响，科学家首先需要考虑所读取资料中的变异。可使用"季节性分解"过程来删除任何系统性的季节性变化。然后对季节性调整序列执行趋势分析。

统计 一组季节性因素

季节性分解数据注意事项

数据 变量应为数值型。

假定 变量不应包含任何内嵌的缺失数据。至少必须定义一个季节性日期成分。

估算季节性因素

1. 从菜单中选择：

分析 > 预测 > 季节性分解

2. 从可用变量列表中选择一或多个变量，并将其移至选定变量列表中。注意，该列表只包含数值变量。

模型类型

"季节性分解"过程提供两种方法来对季节性因素进行建模：乘法或加法。

- 乘法。季节性成分是一个因子，用来与经过季节性调整的序列相乘以得到原始序列。实际上，"趋势"会评估与序列的总体水平成正比的季节性成分。无季节性变动的观测值的季节性成分为 1。
- 加法。将季节性调整添加到季节调整序列中即可得到初始序列。此调整尝试从序列中移去季节性影响，以查看可能被季节性成分"掩盖"的其他兴趣特征。实际上，"趋势"会评估不依赖于序列的总体水平的季节性成分。无季节性变动的观测值的季节性成分为 0。

移动平均值权重

"移动平均值权重"选项允许您指定在计算移动平均数时如何处理序列。这些选项仅在序列的季节性为偶数时才可用。如果季节性为奇数，则所有点的权重都相等。

- 所有点相等 (*All points equal*)。使用等于季节性的跨度以及所有权重相等的点来计算移动平均值。如果季节性是奇数，那么始终使用此方法。
- 按 0.5 对端点加权 (*Endpoints weighted by .5*)。使用等于季节性加 1 的跨度以及以 0.5 加权的跨度的端点计算具有偶数季节性的序列的移动平均值。

根据需要，您可以：

- 单击保存指定如何保存新变量。

季节性分解：保存

创建变量

允许您选择如何处理新变量。

- 添加至文件 (*Add to file*)。将季节性分解所创建的新序列作为常规变量保存在活动数据集中。变量名由三字母前缀、下划线和数字组成。
- 替换现有 (*Replace existing*)。由季节性分解创建的新序列在活动数据集中保存为临时变量。同时，将删除由"预测"过程创建的任何现有的临时变量。变量名由一个三个字母的前缀、一个井字符 (#) 和一个数字组成。
- 不要创建 (*Do not create*)。新序列不添加到活动数据集中。

新变量名称

"季节性分解"过程创建了四个新变量（序列），并且每个指定的序列都带有以下三个字母的前缀：

SAF 季节性调整因子。这些值指示每个周期对序列水平的影响。

SAS 季节性调整序列。这些值是在删除序列的季节性变化之后获得的。

STC 平滑的趋势循环成分。这些值显示序列中出现的趋势和循环行为。

ERR 残差或"误差"值。这些值是在从序列中删除季节性、趋势和循环成分之后保留的。

SEASON 命令其他功能

使用命令语法语言还可以：

- 在 SEASON 命令中指定任意季节性，而不是在"定义日期"过程所提供的选项选择一个周期。

请参阅命令语法参考以获取完整的语法信息。

谱图

"谱图"过程用于标识时间序列中的周期行为。它不需要分析一个时间点与下一个时间点之间的变异，只要按不同频率的季节性成分分析整体序列的变异。平滑序列在低频率具有更强的季节性成分；而随机变异（"白噪声"）将成分强度分布到所有频率。

不能使用该过程分析包含缺失数据的序列。

示例 建造新住房的比率是一个国家/地区经济的重要晴雨表。有关住房的数据开始时通常会表现出一个较强的季节性成分。但在估计当前数字时，分析人员需要注意数据中是否呈现了较长的周期。

统计 正弦和余弦变换、周期图值和每个频率或周期成分的谱密度估计。启用双变量分析后，将统计交叉周期图的实部和虚部、余弦密度、正交谱、增益、平方一致性和每个频率或周期成分的相位谱。

图 对于单变量和双变量分析为周期图和频谱密度。对于双变量分析为平方一致性、正交谱、交叉振幅、余弦密度、相位谱和增益。

谱图数据注意事项

数据 变量应为数值型。

假定 变量不应包含任何内嵌的缺失数据。要分析的时间序列应该是平稳的，任何非零平均值应该从序列中删除。

- 固定。要用 ARIMA 模型进行拟合的时间序列所必须满足的条件。纯的 MA 序列是平稳的，但 AR 和 ARMA 序列可能不是。平稳序列的平均值和方差不随时间改变。

获取谱分析

1. 从菜单中选择：

分析 > 预测 > 谱分析

2. 从可用变量列表中选择一或多个变量，并将其移至选定变量列表中。注意，该列表只包含数值变量。

3. 选择其中一个"频谱窗口"选项来选择如何平滑周期图，以便获得谱密度估计值。可用的平滑选项有"Tukey-Hamming"、"Tukey"、"Parzen"、"Bartlett"、"Daniell (单元格)"和"无"。

- *Tukey-Hamming*。权重为 $W_k = .54D_p(2 \pi f_k) + .23D_p(2 \pi f_k + \pi/p) + .23D_p(2 \pi f_k - \pi/p)$ ， $k = 0, \dots, p$ ，其中 p 是一半跨度的整数部分， D_p 是阶数 p 的 Dirichlet 内核。
- *Tukey*。权重为 $W_k = 0.5D_p(2 \pi f_k) + 0.25D_p(2 \pi f_k + \pi/p) + 0.25D_p(2 \pi f_k - \pi/p)$ ， $k = 0, \dots, p$ ，其中 p 是一半跨度的整数部分， D_p 是阶数 p 的 Dirichlet 内核。
- *Parzen*。权重为 $W_k = 1/p(2 + \cos(2 \pi f_k)) (F[p/2](2 \pi f_k))^{**2}$ ， $k = 0, \dots, p$ ，其中 p 是一半跨度的整数部分，而 $F[p/2]$ 是阶数 $p/2$ 的 Fejer 内核。
- *Bartlett*。谱窗口的形状，窗口上半部分的权重按如下公式计算： $W_k = F_p(2 * \pi * f_k)$ ， $k = 0, \dots, p$ ，其中 p 是半跨度的整数部分， F_p 是阶数 p 的 Fejer 内核。下半部分与上半部分对称。
- *Daniell (单元格)*。所有权重均等于 1 的频谱窗口形状。
- 无。无平滑。如果选择了此选项，那么谱密度估计与周期图相同。

范围。一个连续值范围，在该范围上将执行平滑。通常使用奇数。较大的跨度对谱密度图进行的平滑比较小的跨度程度大。

变量中心化 (Center variables)。调整序列以使在计算谱之前其平均值为 0，并且移去可能与序列平均值关联的较大项。

双变量分析 - 每个双变量中的第一个变量。 如果选择两个或多个变量，则可以选择此选项来请求双变量谱分析。

- "变量"列表中的第一个变量将被视为自变量，而其他所有变量将被视为因变量。
- 系统将使用独立于其他序列命名的第一个序列对第一个序列之后的每个序列进行分析。还会对每个序列执行单变量分析。

图 周期图和谱密度对单变量分析和双变量分析均可用。其他所有选项仅对双变量分析可用。

- 周期图 (*Periodogram*). 针对频率或周期绘制的未平滑谱振幅图 (绘制在对数刻度中)。低频率变动是平滑序列的特征。均匀地分布在所有频率上的变动则表示"白噪音"。
- 平方一致性 (*Squared coherency*). 两个序列的增益的乘积。
- 正交谱 (*Quadrature spectrum*). 交叉周期图的虚部，是两个时间序列的异相频率成分的相关性测量。成分的异相为 $\pi/2$ 弧度。
- 交叉振幅 (*Cross amplitude*). 余谱密度平方和正交谱平方之和的平方根。
- 谱密度。已进行平滑而移去了不规则变动的周期图。
- 余谱密度 (*Cospectral density*). 交叉周期图的实部，是两个时间序列的同相频率分量的相关性测量。
- 相位谱 (*Phase spectrum*). 一个序列的每个频率成分提前或延迟另一个序列的程度的测量。
- 增益 (*Gain*). 用一个序列的谱密度除以跨振幅的商。这两个序列都有自己的获得值。

按频率 (*By frequency*). 所有图都是根据频率生成的，范围从频率 0 (常数或中项) 到频率 0.5 (两个观测值的循环项)。

按周期 (*By period*). 所有图都是按周期生成的，范围从 2 (两个观测值的循环项) 到等于观测值数量的期间 (常数或中项)。周期显示在对数刻度上。

SPECTRA 命令其他功能

使用命令语法语言还可以：

- 将计算的谱分析变量保存到活动数据集中以供将来使用。
- 为谱窗指定定制权重。
- 按频率和周期生成图。
- 打印图中显示的每个值的完整列表。

请参阅命令语法参考以获取完整的语法信息。

通知

本信息是为在美国提供的产品和服务编写的。本资料的其他语言版本可以从 IBM 获取。但是，您可能需要拥有该语言的产品副本或产品版本才能访问这些资料。

IBM 可能在其他国家或地区不提供本文中讨论的产品、服务或功能特性。有关您当前所在区域的产品和服务的信息，请向您当地的 IBM 代表咨询。任何对 IBM 产品、程序或服务的引用并非意在明示或暗示只能使用 IBM 的产品、程序或服务。只要不侵犯 IBM 的知识产权，任何同等功能的产品、程序或服务，都可以代替 IBM 产品、程序或服务。不过，用户应自行负责评估和验证任何非 IBM 产品、程序或服务的工作情况。

IBM 公司可能已拥有或正在申请与本文档内容有关的各项专利。您获得该文档并不意味着授予您任何这些专利许可。您可以将书面许可查询函件发送至：

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US*

有关双字节 (DBCS) 信息的许可查询，请与您所在国家或地区的 IBM 知识产权部门联系，或用书面方式将查询寄往：

*Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan*

International Business Machines Corporation"按现状"提供本出版物，不附有任何种类的（无论是明示的还是暗含的）保证，包括但不限于暗含的有关非侵权、适销和适用于某种特定用途的保证。某些管辖区域在某些交易中不允许免除明示或暗含的保证。因此本条款可能不适用于您。

本信息可能含有技术误差或排版错误。此处的信息会定期进行更改；这些更改会体现在本出版物的新版本中。IBM 可以随时对本资料中描述的产品和/或程序进行改进和/或更改，而不另行通知。

本信息中对任何非 IBM Web 站点的引用都只是为了方便起见才提供的，不以任何方式充当对那些 Web 站点的保证。那些 Web 站点中的资料不是 IBM 产品资料的一部分，使用那些 Web 站点带来的风险将由您自行承担。

IBM 可以按它认为适当的任何方式使用或分发您所提供的任何信息而无须对您承担任何责任。

本程序的被许可方如果要了解有关程序的信息以达到如下目的：(i) 允许在独立创建的程序和其他程序（包括本程序）之间进行信息交换，以及 (ii) 允许对已经交换的信息进行相互使用，请与下列地址联系：

*IBM Director of Licensing
IBM Corporation*

North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

此类信息的提供应遵照相关条款和条件，其中包括在某些情况下支付适当费用。

本资料中描述的许可程序及其所有可用的许可资料均由 IBM 依据 IBM 客户协议、IBM 国际软件许可协议或任何同等协议中的条款提供。

所引用的性能数据和客户示例只用于阐述说明。根据具体配置和操作条件，实际性能结果可能有所不同。

涉及非 IBM 产品的信息可从这些产品的供应商、其出版说明或其他可公开获得的资料中获取。IBM 没有对这些产品进行测试，也无法确认其性能的精确性、兼容性或任何其他关于非 IBM 产品的声明。有关非 IBM 产品性能的问题应当向这些产品的供应商提出。

有关 IBM 未来方向或意向的声明均可能未经通知即变更或撤销，并且仅代表目标和目的。

本信息包含日常业务运营中使用的数据和报告的示例。为了尽可能详尽地对其进行说明，示例中包含了人员的姓名、公司、品牌和产品的名称。所有这些名称都是虚构的，如果与实际人员或公司企业有任何类似则纯属巧合。

版权许可：

本信息包括源语言形式的样本应用程序，这些样本说明不同操作平台上的编程方法。如果是为按照在编写样本程序的操作平台上的应用程序编程接口 (API) 进行应用程序的开发、使用、经销或分发为目的，您可以任何形式对这些样本程序进行复制、修改、分发，而无须向 IBM 付费。这些示例并未在所有条件下作全面测试。因此，IBM 不能担保或暗示这些程序的可靠性、可维护性或功能。本样本程序仍然是“按现状”提供的，不附有任何种类的保证。对于因使用样本程序所引起的任何损害，IBM 概不负责。

凡这些实例程序的每份拷贝或其任何部分或任何衍生产品，都必须包括如下版权声明：

© IBM 2019. 此部分代码是根据 IBM Corp. 公司的样本程序衍生出来的。

© Copyright IBM Corp. 1989 - 20019. All rights reserved.

商标

IBM、IBM 徽标和 ibm.com 是 International Business Machines Corp., 在全球许多管辖区域注册的商标或注册商标。其他产品和服务名称可能是 IBM 或其他公司的商标。Web 站点 www.ibm.com/legal/copytrade.shtml 上的“Copyright and trademark information”部分中提供了 IBM 商标的最新列表。

Adobe、Adobe 徽标、PostScript 和 PostScript 徽标是 Adobe Systems Incorporated 在美国和/或其他国家或地区的注册商标或商标。

Intel、Intel 徽标、Intel Inside、Intel Inside 徽标、Intel Centrino、Intel Centrino 徽标、Celeron、Intel Xeon、Intel SpeedStep、Itanium 和 Pentium 是 Intel Corporation 或其子公司在美国和其他国家或地区的商标或注册商标。

Linux 是 Linus Torvalds 在美国、其他国家或地区或两者的注册商标。

Microsoft、Windows、Windows NT 和 Windows 徽标是 Microsoft Corporation 在美国、其他国家或地区或两者的商标。

UNIX 是 The Open Group 在美国和其他国家或地区的注册商标。

Java 和所有基于 Java 的商标和徽标是 Oracle 及/或其分支机构的商标和注册商标。

索引

[B]

保持个案 2

[D]

调和分析 4

[G]

估计期 2

[J]

季节性分解 2, 3
 保存新变量 3
 创建变量 3
 计算移动平均值 2
 假设 2
 模型 2

[L]

历史期 2

[P]

谱图 4, 5
 假设 4
 居中转换 4
 频谱窗口 4
 双变量谱分析 4

[Y]

验证期 2



Printed in China