

IBM SPSS Forecasting 23

IBM

注释

使用本信息及其支持的产品前，请先阅读 第 51 页的『声明』中的信息。

产品信息

本版本适用于 IBM SPSS Statistics V23.0.0 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

第 1 章 时间序列简介	1	要建模的时间序列	24
时间序列数据	1	选择维度值	25
数据转换	2	观测值	25
估计期和验证期	2	分析时间间隔	26
建立模型并生成预测值	2	聚集和分布	26
第 2 章 时间序列建模器	3	缺失值	27
为专家建模器指定选项	5	常规数据选项	27
模型选择和事件说明	5	常规构建选项	27
使用专家建模器处理离群值	5	要显示的序列	28
自定义的指数平滑法模型	5	输出选项	29
自定义的 ARIMA 模型	6	估计期	30
自定义的 ARIMA 模型的模型规范	6	预测	31
自定义的 ARIMA 模型中的转换函数	7	保存	31
自定义的 ARIMA 模型中的离群值	8	交互式输出	31
输出	8	第 7 章 应用时间因果模型	33
统计和预测表	9	应用时间因果模型	33
图	9	时间因果模型预测	33
将输出限制为最佳或最差拟合模型	10	使用时间因果模型预测	33
保存模型预测和模型规范	10	模型参数和预测	34
选项	11	一般选项	34
TSMODEL 命令的附加功能	12	要显示的序列	35
第 3 章 应用时间序列模型	13	输出选项	35
输出	14	保存	37
统计和预测表	14	时间因果模型方案	38
图	15	运行时间因果模型方案	38
将输出限制为最佳或最差拟合模型	16	定义方案期	38
保存模型预测和模型规范	16	添加方案和方案组	39
选项	17	选项	42
TSAPPLY 命令的其他功能	17	第 8 章 拟合优度测量	43
第 4 章 周期性分解	19	第 9 章 离群值类型	45
季节性分解保存	19	第 10 章 ACF/PACF 图指南	47
SEASON 命令的其他功能	20	声明	51
第 5 章 谱图	21	商标	52
SPECTRA 命令的其他功能	22	索引	55
第 6 章 时间因果模型	23		
获取时间因果模型	23		

第 1 章 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“时间序列建模器”过程、“周期性分解”过程和“谱图”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，请单击**变量视图**选项卡，然后在任意空白行中输入变量名。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”菜单中的“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

注: 从 OLAP 立方体导入 (例如从 IBM® Cognos® TM1® 导入) 的数据表示为多维数据。

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，这些变量用于确定周期性并区分历史期、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。

建立模型并生成预测值

“预测”附加模块提供了两个过程，用于完成创建模型和生成预测值的任务。

- 第 3 页的第 2 章，『时间序列建模器』过程为时间序列创建模型并生成预测值。它包括专家建模器，用于自动确定每个时间序列的最佳模型。对于希望进行高级控制的熟练分析人员，还提供了定制模型构建的工具。
- 第 13 页的第 3 章，『应用时间序列模型』过程将时间序列建模器创建的现有时间序列模型应用于活动数据集。使用此过程，可以在不重新建立模型的情况下获得其新数据或修订数据可用的序列的预测。如果确认模型已发生更改，则可使用时间序列建模器重新建立该模型。

第 2 章 时间序列建模器

“时间序列建模器”过程会估计时间序列的指数平滑法模型、单变量自回归积分移动平均值 (ARIMA) 模型和多变量 ARIMA (或转换函数模型) 模型, 并生成预测值。此过程包含一个专家建模器, 后者尝试自动确定并估计一个或多个因变量序列的最佳拟合 ARIMA 或指数平滑法模型, 这样就不必通过反复试验来确定适当的模型。另外, 可以指定定制的 ARIMA 模型或指数平滑法模型。

示例。假定您是产品经理, 负责预测 100 种独立产品中每一种产品的下个月的单位销售额和收入, 但是您对建模时间序列又知之甚少。所有 100 种产品的单位销售额的历史数据都存储在一个 Excel 电子表格中。在 IBM SPSS® Statistics 中打开电子表格后, 您使用专家建模器申请了未来一个月的预测。专家建模器会为每一种产品查找最合适的单位销售额模型, 并使用这些模型生成预测。由于专家建模器可以处理多个输入序列, 因此, 您只需要运行该过程一次即可获得所有产品的预测。通过选择将预测保存到活动数据集, 您可以轻松地将结果导回 Excel 中。

统计。拟合优度量: 平稳 R 方、 R 方 (R^2)、均方根误差 (RMSE)、平均绝对误差 (MAE)、平均绝对误差百分比 (MAPE)、最大绝对误差 (MaxAE)、最大绝对误差百分比 (MaxAPE)、标准化 BIC 准则。残差: 自相关函数、偏自相关函数、Ljung-Box Q 。ARIMA 模型: 从属变量的 ARIMA 顺序、独立变量的传输函数顺序和离群值估算。此外, 还有指数平滑法模型的平滑法参数估计值。

图。所有模型的摘要图: 平稳 R 方、 R 方 (R^2)、均方根误差 (RMSE)、平均绝对误差 (MAE)、平均绝对误差百分比 (MAPE)、最大绝对误差 (MaxAE)、最大绝对误差百分比 (MaxAPE)、标准化 BIC 准则的直方图; 残差自相关和偏自相关的箱图。单模型结果: 预测值、拟合值、观察值、置信区间的上限和下限、残差自相关和偏自相关。

时间序列建模器的数据注意事项

数据。因变量和任何自变量必须是数值型。

假设。因变量和任何自变量都视为时间序列, 即每个个案代表一个时间点, 连续的个案之间由常数时间区间分隔。

- **平稳性。**对于自定义的 ARIMA 模型, 要建模的时间序列应该是平稳序列。将非平稳序列转换成平稳序列的最有效方法是使用差分变化 - 可从“创建时间序列”对话框执行此操作。
- **预测值。**为了使用带有自 (预测) 变量的模型生成预测值, 活动数据集应包含预测期中全部个案的这些变量的值。此外, 自变量不应包括估计期中的任何缺失值。

定义日期

尽管不是必需的, 但还是建议使用定义日期对话框指定与第一个个案相关的日期以及连续个案之间的时间间隔。此操作应在使用时间序列建模器之前执行, 并将导致一组变量上会标记每个个案的相关日期。此操作还会设置数据的假定周期性, 例如, 如果连续个案之间的时间区间为一个月, 则周期性为 12。如果您希望创建季节性模型, 则需要此周期性。如果不想创建季节性模型, 输出中也不需要日期标签, 则可以跳过“定义日期”对话框。这样, 与每个个案相关的标签将仅是个案号。

使用时间序列建模器

1. 从菜单中选择:

分析 > 预测 > 创建传统模型...

2. 在“变量”选项卡上，选择一个或多个要建模的因变量。
3. 从“方法”下拉框中，选择一种建模方法。如果要自动建模，请保留**专家建模器**的缺省方法。这样将调用专家建模器来为每个因变量确定最佳拟合模型。

要生成预测，请执行以下操作：

4. 单击**选项**选项卡。
5. 指定预测期。这将生成一个包括预测和观察值的图表。

根据需要，您可以：

- 选择一个或多个自变量。自变量的处理方法与回归分析中预测变量的处理方法十分类似，但自变量是可选的。它们可以包含在 ARIMA 模型中，但不能包含在指数平滑法模型中。如果您指定**专家建模器**作为建模方法并包括自变量，则将只考虑 ARIMA 模型。
- 单击**条件**，指定建模的详细信息。
- 保存预测值、置信区间和噪声残差。
- 以 XML 格式保存估计的模型。已保存模型可应用于新数据或修订的数据，以便在不重新构建模型的情况下获取更新后的预测。
- 获取所有估计模型的汇总统计。
- 为自定义的 ARIMA 模型中的自变量指定转换函数。
- 启用自动检测离群值的功能。
- 将特定时间点作为自定义的 ARIMA 模型的离群值来建模。

建模方法

可用的建模方法有：

专家建模器。专家建模器会自动查找每个相依序列的最佳拟合模型。如果指定了自变量（预测）变量，则专家建模器为 ARIMA 模型中的内容选择那些与该相依序列具有统计显著关系的模型。适当时，使用差分和/或平方根或自然对数转换对模型变量进行转换。缺省情况下，专家建模器既考虑指数平滑法模型也考虑 ARIMA 模型。但是，您可以将专家建模器限制为仅搜索 ARIMA 模型或仅搜索指数平滑法模型。还可以指定自动检测离群值。

指数平滑法。使用此选项可指定定制的指数平滑法模型。您可以从各种指数平滑法模型中进行选择，它们在处理趋势和季节性上有所不同。

ARIMA。使用此选项可指定定制的 ARIMA 模型。其中包含显式指定自回归的阶和移动平均值的阶，以及差分度。可以包含自变量（预测变量）并为它们当中的任何一个或全部定义转换函数。还可以指定自动检测离群值或指定显式离群值集合。

估计期和预测期

估计期。估计期定义用于确定模型的个案集。缺省情况下，估计期包含活动数据集中的所有个案。要设置估计期，请在选择个案对话框中选择**基于时间或个案范围**。取决于可用的数据，过程中使用的估计期可能会因因变量而异，因此会与显示的值不同。对于给定的因变量，真正的估计期是从该变量中消去指定估计期的开始或结束时发生的任何连续缺失值之后剩下的期间。

预测期。预测期从估计期后的第一个个案开始，缺省情况下，到活动数据集中的最后一个个案结束。您可以从“选项”选项卡设置预测期的结束时间。

为专家建模器指定选项

专家建模器提供用于约束备选模型集、指定离群值的处理方法以及包括事件变量在内的选项。

模型选择和事件说明

使用“模型”选项卡，可以指定专家建模器可以处理的模型类型并指定事件变量。

模型类型。 可用选项有：

- **所有模型。** 专家建模器既考虑 ARIMA 模型也考虑指数平滑法模型。
- **仅IP指数平滑法模型。** 专家建模器仅考虑指数平滑法模型。
- **仅限 ARIMA 模型。** 专家建模器仅考虑 ARIMA 模型。

专家建模器考虑季节性模型。 只有在为活动数据集定义了周期性时才启用此选项。选择（选中）此选项时，专家建模器既考虑季节性模型，又考虑非季节性模型。如果未选择此选项，则专家建模器仅考虑非季节性模型。

当前周期性。 指示当前为活动数据集定义的周期性（如果有）。当前周期性以整数形式给出，例如，12 表示年度周期性，每个个案代表一个月份。如果尚未设置周期性，则显示值无。季节性模型要求具有周期性。您可以从“定义日期”对话框中设置周期性。

事件。 选择要视为事件变量的任何自变量。对于事件变量，值为 1 的个案指示相依序列将受该事件影响的时间。1 以外的值指示无影响。

使用专家建模器处理离群值

使用“离群值”选项卡，可以选择自动检测离群值以及要检测的离群值的类型。

自动检测离群值。 缺省情况下，不自动检测离群值。选择（选中）此选项可执行自动检测离群值的操作，并可选择以下一种或多种离群值类型：

- 加性
- 移位水平
- 创新的
- 瞬时的
- 季节性加性
- 局部趋势
- 加性修补

自定义的指数平滑法模型

模型类型。 指数平滑法模型¹分为季节性模型和非季节性模型。季节性模型只有在为活动数据集定义了周期时才可用（请参见下文的“当前周期性”）。

- **简单 (Simple).** 此模型适用于没有趋势或季节性的序列。其唯一的平滑参数是水平。简单指数平滑法与 ARIMA 模型极为相似，包含零阶自回归、一阶差分、一阶移动平均值，并且没有常数。
- **Holt 线性趋势 (Holt's linear trend).** 该模型适用于具有线性趋势并没有季节性的序列。其平滑参数是水平和趋势，不受相互之间的值的约束。Holt 模型比 Brown 模型更通用，但在计算大序列时要花的时间更长。Holt 指数平滑法与 ARIMA 模型极为相似，包含零阶自回归、二阶差分以及二阶移动平均值。

1. Gardner, E. S. 1985. Exponential smoothing: The state of the art. *Journal of Forecasting*, 4, 1-28.

- **Brown 线性趋势 (Brown's linear trend)**. 该模型适用于具有线性趋势并没有季节性的序列。其平滑参数是水平和趋势，并假定二者等同。因此，Brown 模型是 Holt 模型的特例。Brown 指数平滑法与具有零阶自回归、二阶差分和二阶移动平均值的 ARIMA 模型极为相似，且移动平均值第二阶的系数等于第一阶的系数二分之一平方。
- **阻尼趋势 (Damped trend)**. 此模型适用于具有线性趋势的序列，且该线性趋势正逐渐消失并且没有季节性。其平滑参数是水平、趋势和阻尼趋势。阻尼指数平滑法与具有一阶自回归、一阶差分和二阶移动平均值的 ARIMA 模型极为相似。
- **简单季节性 (Simple seasonal)**. 此模型适用于没有趋势并且季节性影响随时间变动保持恒定的序列。其平滑参数是水平和季节。简单季节性指数平滑法与 ARIMA 模型极为相似，包含零阶自回归、一阶差分、一阶季节性差分和一阶、 p 阶和 $p + 1$ 阶移动平均值，其中 p 是季节性区间中的周期数（对于月数据， $p = 12$ ）。
- **Winters 可加性 (Winters' additive)**. 此模型适用于具有线性趋势和不依赖于序列水平的季节性效应的序列。其平滑参数是水平、趋势和季节。Winters 可加的指数平滑法与 ARIMA 模型极为相似，包含零阶自回归、一阶差分、一阶季节性差分和 $p + 1$ 阶移动平均值，其中 p 是季节性区间中的周期数（对于月数据， $p = 12$ ）。
- **Winters 相乘性 (Winters' multiplicative)**. 此模型适用于具有线性趋势和依赖于序列水平的季节性效应的序列。其平滑参数是水平、趋势和季节。Winters 的可乘指数平滑法与任何 ARIMA 模型都不相似。

当前周期性。 指示当前为活动数据集定义的周期性（如果有）。当前周期性以整数形式给出，例如，12 表示年度周期性，每个个案代表一个月份。如果尚未设置周期性，则显示值无。季节性模型要求具有周期性。您可以从“定义日期”对话框中设置周期性。

因变量转换。 可以指定在建模之前对每个因变量执行的转换。

- 无。没有执行转换。
- 平方根。平方根转换。
- 自然对数。自然对数转换。

自定义的 ARIMA 模型

使用时间序列建模器，可以构建定制的非周期性或周期性 ARIMA（自回归集成移动平均值）模型，也称为 Box-Jenkins²模型 - 带或不带固定预测变量集均可。您可以为任意或所有预测变量定义转换函数，指定自动检测离群值或指定一组确切的离群值。

- “变量”选项卡上指定的所有自变量（预测变量）都显式包括在该模型中。这点与使用专家建模器相反；使用专家建模器时，只有在自变量与因变量之间具有统计显著关系时才会包括自变量。

自定义的 ARIMA 模型的模型规范

“模型”选项卡用于指定定制的 ARIMA 模型的结构。

ARIMA 阶数。 在“结构”网格的相应单元格中，输入模型的各个 ARIMA 成分的值。所有值都必须为非负整数。对于自回归和移动平均值成分，该值表示最大阶。模型中将包含所有正的较低阶。例如，如果指定 2，则模型包括阶数 2 和 1。只有在为活动数据集定义了周期性时，才会启用“季节性”列中的各个单元格（请参见下面的“当前周期性”）。

- **自回归 (p)。** 模型中的自回归阶数。自回归阶指定要使用序列中以前的哪些值来预测当前值。例如，自回归阶为 2 时，指定序列中过去两个时段的价值用于预测当前值。

2. Box, G. E. P., G. M. Jenkins, and G. C. Reinsel. 1994. *Time series analysis: Forecasting and control*, 3rd ed. Englewood Cliffs, N.J.: Prentice Hall.

- **差分 (d)**。指定在估计模型之前应用于序列的差分的阶。在出现趋势（具有趋势的序列通常是不稳序列，而 ARIMA 建模假定其稳定）时需要差分，并将其用于去除其影响。差分的阶与序列趋势度相对应，一阶差分导致线性趋势，二阶差分导致二次趋势，等等。
- **移动平均值 (q)**。模型中的移动平均值的阶数。移动平均值的阶指定如何使用先前值的序列平均数的偏差来预测当前值。例如，如果移动平均值的阶为 1 和 2，则指定在预测序列的当前值时将考虑上两个时段的每个时段中的序列的平均值的偏差。

季节性阶。季节性自回归成分、移动平均值成分和差分成分与其非季节性对应成分起着相同的作用。但对于季节性的阶，当前序列值受以前的序列值的影响，序列值之间间隔一个或多个季节性周期。例如，对于月数据（季节性周期为 12），季节性 1 阶表示当前序列值受自当前周期起 12 个周期之前的序列值的影响。因此，对于月数据，指定季节性 1 阶等同于指定非季节性 12 阶。

当前周期性。指示当前为活动数据集定义的周期性（如果有）。当前周期性以整数形式给出，例如，12 表示年度周期性，每个个案代表一个月份。如果尚未设置周期性，则显示值无。季节性模型要求具有周期性。您可以从“定义日期”对话框中设置周期性。

因变量转换。可以指定在建模之前对每个因变量执行的转换。

- **无**。没有执行转换。
- **平方根**。平方根转换。
- **自然对数**。自然对数转换。

在模型中包含常数。除非您确定总平均值序列值为 0，否则包含常数是标准方法。当应用差分时，建议不包含常数。

自定义的 ARIMA 模型中的转换函数

使用“转换函数”选项卡（只有在指定了自变量时才显示），可以为“变量”选项卡上指定的任意或所有自变量定义转换函数。转换函数可用于指定如何使用自变量（预测变量）的过去值来预测相依序列的未来值。

转换函数的阶数。在“结构”网格的相应单元格中，输入转换函数的各个成分的值。所有值都必须为非负整数。对于分子和分母成分，该值代表最大阶。模型中将包含所有正的较低阶。此外，0 阶始终包括在分子成分中。例如，如果您指定 2 作为分子，则模型会包括阶数 2、1 和 0。如果指定 3 作为分母，则模型会包括阶数 3、2 和 1。只有在为活动数据集定义了周期时，“季节性”列中的各个单元格才会启用（请参见下面的“当前周期性”）。

- **分子**。转换函数的分子的阶。指定所选独立（预测）序列的哪些以前的值可用于预测相依序列的当前值。例如，分子的阶为 1 时，指定独立序列过去一个时间段的值（以及独立序列的当前值）用于预测每个相依序列的当前值。
- **分母**。转换函数的分母的阶。指定如何使用所选独立（预测）序列的以前值的序列平均数偏差来预测相依序列的当前值。例如，分母的阶为 1 时，指定在预测每个相依序列的当前值时考虑独立序列过去一个时间段的平均值偏差。
- **差分**。指定在估计模型之前应用于所选独立（预测）序列的差分的阶数。存在趋势时必须使用差分来去除其效果。

季节性阶。季节性分子、分母和差分成分与其非季节性对应成分起着相同的作用。但对于季节性的阶，当前序列值受以前的序列值的影响，序列值之间间隔一个或多个季节性周期。例如，对于月数据（季节性周期为 12），季节性 1 阶表示当前序列值受自当前周期起 12 个周期之前的序列值的影响。因此，对于月数据，指定季节性 1 阶等同于指定非季节性 12 阶。

当前周期性。 指示当前为活动数据集定义的周期性（如果有）。当前周期性以整数形式给出，例如，12 表示年度周期性，每个个案代表一个月份。如果尚未设置周期性，则显示值无。季节性模型要求具有周期性。您可以从“定义日期”对话框中设置周期性。

延迟。 设置延迟会导致指定的区间个数延迟自变量的影响。例如，如果延迟设置为 5，则时间 t 的自变量的值在五个周期过去 ($t + 5$) 之前不会影响预测。

转换。 为一组自变量指定转换函数也包括在这些变量上执行可选的转换。

- 无。没有执行转换。
- 平方根。平方根转换。
- 自然对数。自然对数转换。

自定义的 ARIMA 模型中的离群值

“离群值”选项卡提供了以下选项来处理离群值³：自动检测离群值、指定特定点作为离群值或者不检测离群值或不为其建模。

不检测离群值或为其建模。 缺省情况下，不检测离群值，也不为其建模。选择此选项可禁用任何检测离群值或为其建模的功能。

自动检测离群值。 选择此选项可执行自动检测离群值的操作，并可选择以下一种或多种离群值类型：

- 加性
- 移位水平
- 创新的
- 瞬时的
- 季节性加性
- 局部趋势
- 加性修补

请参阅 第 45 页的第 9 章，『离群值类型』主题以获取更多信息。

将特定时间点作为离群值来建模。 选择此选项可指定特定的时间点作为离群值。每个离群值在“离群值定义”网格中占单独的一行。在给定行中为所有单元格输入值。

- **类型。** 离群值类型。支持的类型有：加性（缺省）、水平移位、更新、瞬时、周期性加性或局部趋势。

注 1: 如果没有为活动数据及定义日期规范，那么“离群值定义”网格将只显示一列观察值。要指定离群值，请输入相关个案的行号（与数据编辑器中显示的相同）。

注 2: “离群值定义”网格中的周期列（如存在）指活动数据集中 *CYCLE_* 变量的值。

输出

可用输出包括单个模型的结果以及跨所有模型计算出的结果。单个模型的结果可以限制为基于用户特定标准的最佳或最差拟合模型的集合。

3. Pena, D., G. C. Tiao, and R. S. Tsay, eds. 2001. *A course in time series analysis*. New York: John Wiley and Sons.

统计和预测表

“统计”选项卡提供显示建模结果的表的选项。

按模型显示拟合测量、Ljung-Box 统计和离群值个数。选择（选中）此选项可显示包含每个估计模型的所选拟合测量、Ljung-Box 值以及离群值数目的表。

拟合测量。您可以将以下一项或多项纳入包含每个估计模型的拟合测量的表中：

- 固定的 R 方
- R 方
- 均方根误差
- 平均绝对误差百分比
- 平均绝对误差
- 最大绝对误差百分比
- 最大绝对误差
- 规范化的 BIC

比较模型的统计。这组选项控制如何显示包含跨所有估计模型计算出的统计信息的表。每个选项分别生成单独的表。可以选择以下选项中的一个或多个：

- 拟合优度。固定的 R 方、 R 方、均方根误差、平均绝对误差百分比、平均绝对误差、最大绝对误差百分比、最大绝对误差以及规范化的 BIC 准则的汇总统计和百分位数表。
- 残差自相关函数 (ACF)。所有估计模型中残差的自相关汇总统计和百分位表。
- 残差部分自相关函数 (PACF)。所有估计模型中残差的部分自相关汇总统计和百分位表。

个别模型的统计。这组选项控制如何显示包含每个估计模型的详细信息的表。每个选项分别生成单独的表。可以选择以下选项中的一个或多个：

- 参数估计。显示每个估计模型的参数估计值的表。为指数平滑法和 ARIMA 模型显示不同的表。如果存在离群值，则它们的参数估计值也将显示在单独的表中。
- 残差自相关函数 (ACF)。按每个估计模型的延迟显示残差自相关表。该表包含自相关的置信区间。
- 残差部分自相关函数 (PACF)。按每个估计模型的延迟显示残差部分自相关表。该表包含部分自相关的置信区间。

显示预测值。显示每个估计模型的模型预测和置信区间的表。预测期在“选项”选项卡中设置。

图

“图”选项卡提供显示建模结果的图的选项。

模型比较图

这组选项控制如何显示包含跨所有估计模型计算出的统计信息的图。每个选项分别生成单独的图。可以选择以下选项中的一个或多个：

- 固定的 R 方
- R 方
- 均方根误差
- 平均绝对误差百分比
- 平均绝对误差

- 最大绝对误差百分比
- 最大绝对误差
- 规范化的 BIC
- 残差自相关函数 (ACF)
- 残差部分自相关函数 (PACF)

单个模型图

序列。 选择（选中）此选项可获取每个估计模型的预测值图。可以选择在图中包含以下一项或多项：

- **观察值。** 相依序列的观察值。
- **预测值。** 预测期的模型预测值。
- **拟合值。** 估计期的模型预测值。
- **预测值的置信区间。** 预测期的置信区间。
- **拟合值的置信区间。** 估计期的置信区间。

残差自相关函数 (ACF)。 显示每个估计模型的残差自相关图。

残差部分自相关函数 (PACF)。 显示每个估计模型的残差部分自相关图。

将输出限制为最佳或最差拟合模型

“输出过滤器”选项卡提供的选项可用于将表格输出和图形输出限制为估计模型子集。可以选择根据提供的拟合标准将输出限制为最佳和/或最差拟合模型。缺省情况下，输出中包括所有估计模型。

最佳拟合模型。 选择（选中）此选项将在输出中包含最佳拟合模型。选择拟合优度测量并指定要包含的模型数。选择此选项不妨碍同时选择最差拟合模型。如果同时选择两者，则输出将由最差拟合模型和最佳拟合模型组成。

- **模型的固定数量。** 指定为 n 个最佳拟合模型显示结果。如果该数量超过估计模型的数量，则显示所有模型。
- **占模型总数的百分比。** 指定为其拟合优度值在所有估计模型的前 n 个百分比范围内的模型显示结果。

最差拟合模型。 选择（选中）此选项将在输出中包含最差拟合模型。选择拟合优度测量并指定要包含的模型数。选择此选项不妨碍同时选择最佳拟合模型。如果同时选择两者，输出将由最佳拟合模型和最差拟合模型组成。

- **模型的固定数量。** 指定为 n 个最差拟合模型显示结果。如果该数量超过估计模型的数量，则显示所有模型。
- **占模型总数的百分比。** 指定为其拟合优度值在所有估计模型的后 n 个百分比范围内的模型显示结果。

拟合优度。 选择用于过滤模型的拟合优度测量。缺省值为固定的 R 方。

保存模型预测和模型规范

使用“保存”选项卡可以将模型预测另存为活动数据集中的新变量，也可以将模型规范以 XML 格式保存到外部文件中。

保存变量。 可以将模型预测、置信区间和残差另存为活动数据集中的新变量。每个相依序列都会带来其自己的一组新变量，每个新变量都包含估计期和预测期的值。如果预测期超出了该相依变量序列的长度，则增加新个案。选中每一项所关联的“保存”复选框可以选择保存新变量。缺省情况下不保存新变量。

- **预测值。** 模型预测值。
- **置信区间的下限。** 预测值的置信区间下限。

- **置信区间的上限。** 预测值的置信区间上限。
- **噪声残值。** 模型残差。如果执行了因变量转换（例如自然对数），则为转换后的序列的残差。
- **变量名的前缀。** 指定要用于新变量名称的前缀或保留缺省的前缀。变量名由前缀、关联因变量的名称和模型标识组成。需要时可以扩展变量名以避免变量命名冲突。前缀必须符合有效变量名称规则。

导出模型文件。 所有估计模型的模型规范都将以 XML 格式导出到指定的文件中。已保存的模型。

- **XML 文件。** 以 XML 文件保存的模型规范可用于 IBM SPSS 应用程序。
- **PMML 文件。** 以 PMML 相容的 XML 文件保存的模型规范可用于 PMML 相容的应用程序，包含 IBM SPSS 应用程序。

选项

“选项”选项卡可用于设置预测期、指定缺失值的处理方法、设置置信区间宽度、指定模型标识的自定义前缀并设置为自相关显示的延迟个数。

预测期。 预测期始终以估计期（用于确定模型的一组个案）结束后的第一个个案开始，到活动数据集中的最后一个个案或用户指定的日期结束。缺省情况下，估计期的结束是活动数据集中的最后一个个案，但可以在以下位置更改它：选择个案对话框中的**基于时间或个案范围**。

- **估计期结束之后的第一个个案到活动数据集内的最后一个个案。** 如果估计期在活动数据集中的最后一个个案之前结束，而您需要直到最后一个个案的预测值，则请选择此选项。此选项通常用来生成保持期的预测，以便将模型预测与实际值子集进行比较。
- **估计期结束之后到指定日期之间的第一个个案。** 选择此选项可显式指定预测期的结束。此选项通常用于在实际序列结束后生成预测。在“日期”网格中为所有单元格输入值。

如果没有为活动数据集定义日期规范，则“日期”网格只显示观察值一列。要指定预测期的结束时间，请输入相关个案在数据编辑器中显示的行号。

“日期”网格中的循环列（如果存在）是活动数据集中的 *CYCLE_* 变量的值。

用户缺失值。 这些选项控制用户缺失值的处理。

- **视为无效。** 用户缺失值当作系统缺失值处理。
- **视为有效。** 用户缺失值当作有效数据处理。

缺失值策略。 下列规则适用于在建模过程期间处理缺失值（包括系统缺失值和视为无效的用户缺失值）：

- 模型中包含估计期发生的因变量存在缺失值的个案。缺失值的具体处理方式取决于估计方法。
- 如果自变量在估计期有缺失值，系统会发出警告。对于专家建模器，将在不使用变量的情况下对包含自变量的模型进行估计。对于自定义 ARIMA，不会估计涉及自变量的模型。
- 如果在预测期内任何自变量有缺失值，建模过程将发出警告并进行最大程度的预测。

置信区间宽度(%)。 将为模型预测和残差自相关计算置信区间。您可以指定任何小于 100 的正数值。缺省情况下，将使用 95% 置信区间。

输出中模型标识的前缀。 “变量”选项卡上指定的每个因变量都可带来一个单独的估计模型。模型都用唯一名称区别，名称由可自定义的前缀和整数后缀组成。您可以输入前缀，也可以保留模型的缺省值。

ACF 和 PACF 输出中显示的最大延迟数。 可以设置在自相关和偏自相关表和图中显示的最大延迟数。

TSMODEL 命令的附加功能

如果将选择内容粘贴到语法窗口并编辑随之产生的 TSMODEL 命令语法，则您可以自定义自己的时间序列建模。使用命令语法语言可以：

- 指定数据的季节性周期（使用 AUXILIARY 子命令的 SEASONLENGTH 关键字）。该设置的优先级高于活动数据集的当前周期性（如果有）。
- 指定自定义 ARIMA 和转换函数成分的不连续延迟（使用 ARIMA 和 TRANSFERFUNCTION 子命令）。例如，您可以指定自定义 ARIMA 模型的自回归延迟的阶数为 1、3 和 6；或转换函数的分子延迟阶数为 2、5 和 8。
- 为单次运行的数据序列建模器过程提供一组以上的建模规格，例如：建模方法、ARIMA 阶数、自变量等等（使用 MODEL 子命令）。

请参阅命令语法参考以获取完整的语法信息。

第 3 章 应用时间序列模型

“应用时间序列模型”过程从外部文件加载现有的时间序列模型，并将它们应用于活动数据集。使用此过程，可以在不重新建立模型的情况下获得其新数据或修订数据可用的序列的预测值。模型由时间序列建模器过程生成。

示例。假定您是一家大型零售店的库存经理，您负责管理 5,000 种产品。您曾使用专家建模器创建了一些模型，用来预测每种产品在未来三个月的销售情况。您的数据仓库每个月都会使用实际销售数据进行刷新，您希望使用这些数据来生成每月更新预测值。通过“应用时间序列模型”过程，您可以使用原有模型，然后只需重新估计模型参数以说明新数据即可实现此预测。

统计。拟合优度量：平稳 R 方、 R 方 (R^2)、均方根误差 (RMSE)、平均绝对误差 (MAE)、平均绝对误差百分比 (MAPE)、最大绝对误差 (MaxAE)、最大绝对误差百分比 (MaxAPE)、标准化 BIC 准则。残差：自相关函数、偏自相关函数、Ljung-Box Q 。

图。所有模型的摘要图：平稳 R 方、 R 方 (R^2)、均方根误差 (RMSE)、平均绝对误差 (MAE)、平均绝对误差百分比 (MAPE)、最大绝对误差 (MaxAE)、最大绝对误差百分比 (MaxAPE)、标准化 BIC 准则的直方图；残差自相关和偏自相关的箱图。单模型结果：预测值、拟合值、观察值、置信区间的上限和下限、残差自相关和偏自相关。

应用时间序列模型数据注意事项

数据。对其应用模型的变量（自变量和因变量）应为数值变量。

假设。将模型应用到活动数据集中与模型中指定的变量同名的变量。所有此类变量均视为时间序列，即每个个案代表一个时间点，连续的个案之间由常数时间区间分隔。

- **预测值。**为了使用带有自（预测）变量的模型生成预测值，活动数据集应包含预测期中全部个案的这些变量的值。如果重新估计模型参数，则自变量不应包含估计期中的任何缺失值。

定义日期

“应用时间序列模型”过程要求活动数据集的周期性（如果有）与要应用的模型的周期性相匹配。如果只是使用建立模型时使用的同一数据集（或许包含新的数据或修订的数据）进行预测，则需满足此条件。如果活动数据集没有周期性，那么可能会将您导航到定义日期对话框以创建一个周期。但是，如果创建模型时没有指定周期性，则活动数据集也应该没有周期性。

应用模型

1. 从菜单中选择:

分析 > 预测 > 应用传统模型...

2. 输入模型文件的文件规范，或者单击浏览并选择模型文件，模型文件通过时间序列建模器过程创建。

根据需要，您可以：

- 使用活动数据集中的数据重新估计模型参数。使用重新估计的参数创建预测值。
- 保存预测值、置信区间和噪声残差。
- 以 XML 格式保存重新估计的模型。

模型参数和适合度量的好处

从模型文件加载。 使用该模型文件中的模型参数（而不必重新估计这些参数）生成预测值。在输出中显示并用于过滤模型（最佳拟合或最差拟合）的拟合优度量取自模型文件，反映了建立（或最近一次更新）每个模型时使用的数据。使用此选项，预测值将不考虑活动数据集中因变量或自变量的历史数据。如果您希望历史数据影响预测值，则必须选择**根据数据重新估计**。此外，预测值也不考虑预测期中相依序列的值，但会考虑预测期中自变量的值。如果相依序列的值已更新，并且您希望将它们包含在预测值中，则需要重新估计，将估计期调整为包含这些值。

根据数据重新估计。 使用活动数据集中的数据重新估计模型参数。重新估计模型参数不会影响到模型结构。例如，ARIMA(1,0,1) 模型将保持原样，但会重新估计自回归参数和移动平均值参数。重新估计不会导致检测新的离群值。离群值（如果有）将始终从模型文件中获取。

- **估计期。** 估计期定义用于重新估计模型参数的个案集。缺省情况下，估计期包含活动数据集中的所有个案。要设置估计期，请在选择个案对话框中选择**基于时间或个案范围**。根据可用数据的不同，过程中使用的估计期可能会因模型而异，因此会与显示的值不同。对于给定的模型，真正的估计期是从模型的因变量中排除指定估计期的开始或结束时发生的所有连续缺失值之后剩下的时段。

预测期

每个模型的预测期始终从估计期结束后的第一个个案开始，到活动数据集中的最后一个个案或用户指定的日期结束。如果没有重新估计参数（这是缺省情况），则每个模型的估计期是建立（或上次更新）模型时使用的个案集。

- **估计期结束之后的第一个个案到活动数据集内的最后一个个案。** 如果估计期在活动数据集中的最后一个个案之前结束，而您需要直到最后一个个案的预测值，则请选择此选项。
- **估计期结束之后到指定日期之间的第一个个案。** 选择此选项可显式指定预测期的结束。在“日期”网格中为所有单元格输入值。

如果没有为活动数据集定义日期规范，则“日期”网格只显示观察值一列。要指定预测期的结束时间，请输入相关个案在数据编辑器中显示的行号。

“日期”网格中的**循环列**（如果存在）是活动数据集中的 *CYCLE_* 变量的值。

输出

可用输出包括单个模型的结果以及所有模型的结果。单个模型的结果可以限制为基于用户特定标准的最佳或最差拟合模型的集合。

统计和预测表

“统计”选项卡提供的选项用于显示模型拟合度统计、模型参数、自相关函数以及预测值的表。除非重新估计模型参数（“模型”选项卡上的**根据数据重新估计**），否则，显示的拟合测量值、Ljung-Box 值和模型参数都是模型文件中的相应值，并反映每个模型建立（或上次更新）时使用的数据。始终从该模型文件获取离群值信息。

按模型显示拟合测量、Ljung-Box 统计和离群值个数。 选择（选中）此选项可显示包含每个模型的所选拟合测量、Ljung-Box 值以及离群值个数的表。

拟合测量。 您可以将以下一项或多项纳入包含每个模型的拟合测量的表中：

- 固定的 R 方
- R 方
- 均方根误差

- 平均绝对误差百分比
- 平均绝对误差
- 最大绝对误差百分比
- 最大绝对误差
- 规范化的 BIC

比较模型的统计。 这组选项控制如何显示包含所有模型的统计的表。每个选项分别生成单独的表。可以选择以下选项中的一个或多个：

- **拟合优度。** 固定的 R 方、 R 方、均方根误差、平均绝对误差百分比、平均绝对误差、最大绝对误差百分比、最大绝对误差以及规范化的 BIC 准则的汇总统计和百分位数表。
- **残差自相关函数 (ACF)。** 所有估计模型中残差的自相关汇总统计和百分位表。此表只在重新估计模型参数时可用（“模型”选项卡上的**根据数据重新估计**）。
- **残差部分自相关函数 (PACF)。** 所有估计模型中残差的部分自相关汇总统计和百分位表。此表只在重新估计模型参数时可用（“模型”选项卡上的**根据数据重新估计**）。

个别模型的统计。 这组选项控制如何显示包含每个模型的详细信息的表。每个选项分别生成单独的表。可以选择以下选项中的一个或多个：

- **参数估计。** 显示每个模型的参数估计值的表。为指数平滑法和 ARIMA 模型显示不同的表。如果存在离群值，则它们的参数估计值也将在单独的表中显示。
- **残差自相关函数 (ACF)。** 按每个估计模型的延迟显示残差自相关表。该表包含自相关的置信区间。此表只在重新估计模型参数时可用（“模型”选项卡上的**根据数据重新估计**）。
- **残差部分自相关函数 (PACF)。** 按每个估计模型的延迟显示残差部分自相关表。该表包含部分自相关的置信区间。此表只在重新估计模型参数时可用（“模型”选项卡上的**根据数据重新估计**）。

显示预测值。 显示每个模型的模型预测值和置信区间的表。

图

“图”选项卡提供的选项可用于显示模型拟合度统计、自相关函数以及序列值（包括预测值）的图。

模型比较图

这组选项控制如何显示包含所有模型的统计值的图。除非重新估计模型参数（“模型”选项卡上的**根据数据重新估计**），否则，显示的值均为模型文件中的值，并反映每个模型建立（或上次更新）时使用的数据。此外，只有在重新估计模型参数时自相关图才可用。每个选项分别生成单独的图。可以选择以下选项中的一个或多个：

- 固定的 R 方
- R 方
- 均方根误差
- 平均绝对误差百分比
- 平均绝对误差
- 最大绝对误差百分比
- 最大绝对误差
- 规范化的 BIC
- 残差自相关函数 (ACF)
- 残差部分自相关函数 (PACF)

单个模型图

序列。选择（选中）此选项可获取每个模型的预测值的图。只有在重新估计模型参数时（“模型”选项卡上的**根据数据重新估计**），观察值、拟合值、拟合值的置信区间以及自相关才可用。可以选择在图中包含以下一项或多项：

- **观察值。** 相依序列的观察值。
- **预测值。** 预测期的模型预测值。
- **拟合值。** 估计期的模型预测值。
- **预测值的置信区间。** 预测期的置信区间。
- **拟合值的置信区间。** 估计期的置信区间。

残差自相关函数 (ACF)。 显示每个估计模型的残差自相关图。

残差部分自相关函数 (PACF)。 显示每个估计模型的残差部分自相关图。

将输出限制为最佳或最差拟合模型

“输出过滤器”选项卡提供的选项可用于将表格输出和图表输出限制为模型的子集。可以选择根据提供的拟合标准将输出限制为最佳和/或最差拟合模型。缺省情况下，输出中包含所有模型。除非重新估计模型参数（“模型”选项卡上的**根据数据重新估计**），否则，用于过滤模型的拟合测量的值均为模型文件中的那些值，并反映每个模型建立（或上次更新）时使用的数据。

最佳拟合模型。 选择（选中）此选项将在输出中包含最佳拟合模型。选择拟合优度测量并指定要包含的模型数。选择此选项不妨碍同时选择最差拟合模型。如果同时选择两者，则输出将由最差拟合模型和最佳拟合模型组成。

- **模型的固定数量。** 指定为 n 个最佳拟合模型显示结果。如果指定的数量超过模型的总数，则显示所有模型。
- **占模型总数的百分比。** 指定为其拟合优度值在所有模型的前 n 个百分比范围内的模型显示结果。

最差拟合模型。 选择（选中）此选项将在输出中包含最差拟合模型。选择拟合优度测量并指定要包含的模型数。选择此选项不妨碍同时选择最佳拟合模型。如果同时选择两者，输出将由最佳拟合模型和最差拟合模型组成。

- **模型的固定数量。** 指定为 n 个最差拟合模型显示结果。如果指定的数量超过模型的总数，则显示所有模型。
- **占模型总数的百分比。** 指定为其拟合优度值在所有模型的后 n 个百分比范围内的模型显示结果。

拟合优度。 选择用于过滤模型的拟合优度测量。缺省值为固定的 R 方。

保存模型预测和模型规范

使用“保存”选项卡可以将模型预测另存为活动数据集中的新变量，也可以将模型规范以 XML 格式保存到外部文件中。

保存变量。 可以将模型预测、置信区间和残差另存为活动数据集中的新变量。每个模型都会带来其自己的一组新变量。如果预测期超出与该模型关联的因变量序列的长度，则添加新个案。除非重新估计模型参数（“模型”选项卡上的**根据数据重新估计**），否则仅为该预测期创建预测值和置信区间上下限。选中每一项所关联的“保存”复选框可以选择保存新变量。缺省情况下不保存新变量。

- **预测值。** 模型预测值。
- **置信区间的下限。** 预测值的置信区间下限。
- **置信区间的上限。** 预测值的置信区间上限。

- **噪声残值。** 模型残差。如果执行了因变量转换（例如自然对数），则为转换后的序列的残差。此选项只在重新估计模型参数时可用（“模型”选项卡上的**根据数据重新估计**）。
- **变量名的前缀。** 指定要用于新变量名称的前缀或保留缺省的前缀。变量名由前缀、关联因变量的名称和模型标识组成。需要时可以扩展变量名以避免变量命名冲突。前缀必须符合有效变量名称规则。

导出模型文件包含重新估计的参数和拟合统计的模型规范将以 XML 格式导出到指定的文件中。此选项只在重新估计模型参数时可用（“模型”选项卡上的**根据数据重新估计**）。

- **XML 文件。** 以 XML 文件保存的模型规范可用于 IBM SPSS 应用程序。
- **PMML 文件。** 以 PMML 相容的 XML 文件保存的模型规范可用于 PMML 相容的应用程序，包含 IBM SPSS 应用程序。

选项

“选项”选项卡可用于指定缺失值的处理方法，设置置信区间宽度并设置为自相关显示的延迟个数。

用户缺失值。 这些选项控制用户缺失值的处理。

- **视为无效。** 用户缺失值当作系统缺失值处理。
- **视为有效。** 用户缺失值当作有效数据处理。

缺失值策略。 以下规则用于处理缺失值（包括视为无效的系统缺失值 and 用户缺失值）：

- 模型中包含估计期发生的因变量存在缺失值的个案。缺失值的具体处理方式取决于估计方法。
- 对于 ARIMA 模型，如果预测值在估计期内有任何缺失值，则将发出警告。不会重新估计任何包含该预测值的模型。
- 如果在预测期内任何自变量有缺失值，建模过程将发出警告并进行最大程度的预测。

置信区间宽度(%)。 将为模型预测和残差自相关计算置信区间。您可以指定任何小于 100 的正数值。缺省情况下，将使用 95% 置信区间。

ACF 和 PACF 输出中显示的最大延迟数。 可以设置在自相关和偏自相关表和图中显示的最大延迟数。此选项只在重新估计模型参数时可用（“模型”选项卡上的**根据数据重新估计**）。

TSAPPLY 命令的其他功能

如果将选择内容粘贴到语法窗口并编辑生成的 TSAPPLY 命令语法，则可以使用其他功能。使用命令语法语言可以：

- 指定仅将模型文件中的模型子集应用于活动数据集（使用 MODEL 子命令的 DROP 和 KEEP 关键字）。
- 将两个或多个模型文件中的模型应用于数据（使用 MODEL 子命令）。例如，一个模型文件可能包含表示单位销售额的序列的模型，而另一个模型文件则可能包含表示收入的序列的模型。

请参阅命令语法参考以获取完整的语法信息。

第 4 章 周期性分解

“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。

假设。变量不应包含任何内嵌的缺失数据。至少必须定义一个周期性日期成分。

估计季节性因素

1. 从菜单中选择:

分析 > 预测 > 周期性分解...

2. 从可用列表选择一个或多个变量，并将其移至“变量”列表中。注意，该列表只包含数值变量。

模型类型。“周期性分解”过程提供两种方法来对周期性因素进行建模：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。无季节性变动的观察值的季节性成分为 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。无季节性变动的观察值的季节性成分为 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周期性为偶数时才可用。如果周期性为奇数，则所有点的权重都相等。

- **所有点相等 (All points equal).** 使用等于周期性的跨度以及所有权重相等的点来计算移动平均值。如果周期性是奇数，那么始终使用此方法。
- **按 0.5 对端点加权 (Endpoints weighted by .5).** 使用等于周期性加 1 的跨度以及以 0.5 加权的跨度的端点计算具有偶数周期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 命令的其他功能

使用命令语法语言还可以：

- 在 SEASON 命令中指定任意周期性，而不是在“定义日期”过程所提供的选项中选择一个周期。

请参阅命令语法参考以获取完整的语法信息。

第 5 章 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，只要按不同频率的周期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周期性成分；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。

假设。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 > 时间序列 > 谱分析...

2. 从可用列表选择一个或多个变量，并将它们移至“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 (两个观察值的循环项)。

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

SPECTRA 命令的其他功能

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。

第 6 章 时间因果模型

时间因果建模尝试发现时间序列数据中的关键因果关系。在时间因果建模中，您指定一组目标序列以及这些目标的一组候选输入。然后，此过程将为每个目标构建自回归时间序列模型，并且仅包括那些与目标具有因果关系的输入。此方法不同于传统时间序列建模，在传统时间序列建模中，您必须为目标序列显式指定预测变量。由于时间因果建模通常涉及为多个相关的时间序列构建模型，因此结果称为模型系统。

在时间因果建模上下文中，术语因果指的是格兰杰因果关系。对于时间序列 X 和时间序列 Y，如果同时依据 X 和 Y 的过去值对 Y 进行回归所产生的 Y 模型优于仅对 Y 的过去值进行回归所产生的模型，那么将时间序列 X 称为时间序列 Y 的“格兰杰原因”。

示例

业务决策制定者可以使用时间因果建模来发现描述业务的大量基于时间的度量中的因果关系。分析可能会揭示几个可控输入，这些输入对关键绩效指标的影响最大。

大型 IT 系统的管理者可以使用时间因果建模在大量相互关联的操作度量中检测异常。这样，通过因果模型不仅可以检测异常，还能发现这些异常的最有可能的根本原因。

字段要求

必须至少有一个目标。缺省情况下，将不会使用预定义角色为“无”的字段。

数据结构

时间因果建模支持两种类型的数据结构。

基于列的数据

对于基于列的数据，每个时间序列字段都包含单个时间序列的数据。此结构是“时间序列建模器”所使用的传统时间序列数据结构。

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，两种不同销售渠道（零售和 Web）的销售数据可能存储在单个 *sales* 字段中。名为 *channel* 且值为“retail”和“web”的维度字段标识与这两种销售渠道中的每种渠道相关联的记录。

获取时间因果模型

此功能需要“统计预测”选项。

从菜单中选择：

分析 > 预测 > 创建时间因果模型...

1. 如果观测值由日期/时间字段定义，请指定该字段。
2. 对于多维数据，请指定标识时间序列的维度字段。
 - 指定维度字段的顺序定义了这些字段在所有后续对话框和输出中的显示顺序。请使用向上和向下箭头按钮对维度字段进行重新排序。

- 您可以指定维度值的子集，以便仅对维度字段的特定值进行分析。例如，如果存在表示地区和品牌的维度，那么可以将分析限制在特定地区。维度子集将应用于分析中使用的所有维度字段。
- 另外，还可以通过在度量字段级别按角色指定维度值来定制分析。例如，如果存在表示销售渠道的维度（其值为“retail”和“web”）以及针对这些渠道中的销售和宣传的度量，那么可以将 Web 宣传同时指定为零售和 Web 销售的输入。缺省情况下，这种类型的定制处于启用状态，并且限制为最多从每个维度字段的指定数目（缺省为 250 个）的相异值中进行选择。

3. 单击**继续**。

注： 如果活动数据集具有日期指定项，那么步骤 1、2 和 3 不适用。日期指定项通过“定义日期”对话框或 DATE 命令进行创建。

4. 单击**字段**以指定要包括在模型中的时间序列并指定定义观测值的方式。必须将至少一个字段指定为目标或者同时指定为输入和目标。
5. 单击**数据指定项**以指定可选设置，其中包括分析时间间隔、聚集和分布设置以及缺失值处理方式。
6. 单击**构建选项**以定义估计期、指定输出内容以及指定构建设置（例如每个目标的最大输入数目）。
7. 单击**模型选项**以请求进行预测、保存预测并将模型系统导出到外部文件。
8. 单击**运行**以运行此过程。

要建模的时间序列

在“字段”选项卡上，请使用**时间序列**设置来指定要包括在模型系统中的序列。

对于基于列的数据，术语**序列**的含义与术语**字段**相同。对于多维数据，包含时间序列的字段称为**度量字段**。对于多维数据，时间序列由度量字段以及每个维度字段的值定义。对于基于列的数据和 multidimensional data，注意事项如下所示。

- 将对指定为“候选输入”或“同时作为目标和输入”的序列加以考虑，以便将其包括在每个目标的模型中。每个目标的模型都始终包含该目标自身的延迟值。
- 指定为“强制输入”的序列始终包括在每个目标的模型中。
- 必须将至少一个序列指定为目标或者同时指定为目标和输入。
- 如果选中了**使用预定义角色**，那么角色为“输入”的字段将设置为“候选输入”。没有任何预定义角色映射到“强制输入”。

多维数据

对于多维数据，请在网格中指定度量字段及相关角色，网格中的每一行都指定单个度量和角色。缺省情况下，模型系统包含此网格中每一行的所有维度字段组合的序列。例如，如果存在 *region* 维度和 *brand* 维度，那么在缺省情况下，指定度量 *sales* 作为目标意味着对于 *region* 与 *brand* 的每个组合，都存在单独的 *sales* 目标序列。

对于网格中的每一行，您可以通过单击维度的省略号按钮来定制任何维度字段的值集合。此操作将打开“选择维度值”子对话框。另外，您还可以添加、删除或复制网格行。

序列计数列显示当前对相关度量指定的维度值集合的数目。显示的值可能大于序列的实际数目（每个集合各有一个对应的序列）。当指定的某些维度值组合未与相关度量所包含的序列相对应时，将发生这种情况。

选择维度值

对于多维数据，您可以通过指定哪些维度值将应用于具有特定角色的特定度量字段来定制分析。例如，如果 *sales* 是一个度量字段，*channel* 是值为“retail”和“web”的维度，那么您可以指定“web”销售是输入，“retail”销售是目标。另外，还可以指定将应用于分析中使用的所有度量字段的维度子集。例如，如果 *region* 是指示地理区域的维度字段，那么您可以将分析限制在特定区域。

所有值 指定包括当前维度字段的所有值。这是缺省选项。

选择要包括或排除的值

使用此选项可以指定当前维度字段的值集。如果对**方式**选择了**包括**时，那么将仅包括**选择的值**列表中指定的值。如果对**方式**选择了**排除**，那么将包括除**选择的值**列表中指定的值之外的所有值。

您可以对要从中进行选择的价值集进行过滤。满足过滤条件的值将显示在**匹配**选项卡中，不满足过滤条件的值将显示在**未选择的值**列表的**不匹配**选项卡中。**全部**选项卡列示所有未选择的值，而无论指定了什么过滤条件。

- 指定过滤器时，可以使用星号 (*) 来表示通配符。
- 要清除当前过滤器，请在“过滤显示的值”对话框中对搜索项指定空值。

观测值

在“字段”选项卡上，使用**观测值**设置可以指定用于定义观测值的字段。

注：如果活动数据集具有日期指定项，那么观测值由日期指定项定义，并且在时间因果建模过程中不可修改。日期指定项通过“定义日期”对话框或 `DATE` 命令进行创建。

由日期/时间定义的观测值

您可以指定观测值由具有日期、时间或日期时间格式的字段定义，或者由表示日期/时间的字符串字段定义。字符串字段可以表示 `YYYY-MM-DD` 格式的日期、`HH:MM:SS` 格式的时间或者 `YYYY-MM-DD HH:MM:SS` 格式的日期时间。在字符串表示中，可以省略前导零。例如，字符串 `2014-9-01` 相当于 `2014-09-01`。

除了用于定义观测值的字段以外，请选择用于描述观测值的适当时间间隔。根据指定的时间间隔不同，您还可以指定其他设置，例如观测值之间的间隔（增量）或者每周的天数。对于时间间隔，注意事项如下所示：

- 如果各个观测值之间的时间间距不规则（例如处理销售订单的时间），请使用**不规则值**。如果选择了**不规则**，那么必须通过“数据指定项”选项卡上的**时间间隔**设置指定用于分析的时间间隔。
- 如果观测值表示日期和时间，并且时间间隔为小时、分钟或秒，请使用**每天的小时数**、**每天的分钟数**或**每天的秒数**。如果观测值表示时间（持续时间）并且未引用日期，而时间间隔为小时、分钟或秒，请使用**小时数（非周期性）**、**分钟数（非周期性）**或**秒数（非周期性）**。
- 根据选择的时间间隔，此过程可以检测缺失的观测值。由于此过程假定所有观测值之间的时间间距相等，并假定未缺失观测值，因此有必要检测缺失的观测值。例如，如果时间间隔为“天”，并且日期 `2014-10-27` 后面跟着 `2014-10-29`，那么表明缺失 `2014-10-28` 的观测值。对于任何缺失的观测值，将插补值。您可以在“数据指定项”选项卡上指定用于处理缺失值的设置。
- 指定的时间间隔使此过程能够检测到同一时间间隔内的多个需要聚集到一起的观测值并使各个观测值在时间间隔边界（例如每个月的第一天）处对齐，以确保各个观测值之间的间距相等。例如，如果时间间隔为“月”，那么同一个月内的多个日期将聚集到一起。此类聚集称为**分组**。缺省情况下，进行分组时，将计算观测值的总和。通过“数据指定项”选项卡上的**聚集和分布**设置，您可以指定另一种分组方法，例如计算各个观测值的平均值。

- 对于某些时间间隔，附加设置可以定义正常等间距时间间隔中的中断。例如，如果时间间隔为“天”，但只有工作日有效，那么您可以指定每周有 5 天，并且每周从星期一开始。

由周期或循环周期定义的观测值

观测值可以由一个或多个表示周期或周期反复循环（直至达到任意数目的循环级别为止）的整数字段定义。借助此结构，您可以描述任何标准时间间隔都无法支持的观测值序列。例如，通过使用表示年的循环字段和表示月的周期字段，并将一个循环的长度设置为 10，可以描述仅由 10 个月组成的财年。

用于指定循环周期的字段定义了周期性级别的层次结构，在此层次结构中，最低级别由**周期**字段定义。次高级别由级别为 1 的循环字段指定，接着由级别为 2 的循环字段指定，依此类推。除最高级别以外，每个级别的字段值对于次高级别都必须具有周期性。最高级别的值不得具有周期性。例如，对于由 10 个月组成的财年，月在年中具有周期性，而年不具有周期性。

- 在特定级别，循环长度是次低级别的周期长度。在财年示例中，只有一个循环级别，并且循环长度为 10，这是因为次低级别表示月，而指定的财年包含 10 个月。
- 请指定任何并非起始于 1 的周期性字段的起始值。要检测缺失值，有必要进行此设置。例如，如果周期性字段起始于 2，但起始值指定为 1，那么此过程将假定该字段的每个循环中的第一个周期都有一个缺失值。

由记录顺序定义的观测值

对于基于列的数据，您可以指定观测值由记录顺序定义，使得第一个记录表示第一个观测值，第二个记录表示第二个观测值，依此类推。这样，就会假定各个记录表示时间间距相等的观测值。

分析时间间隔

用于分析的时间间隔可以与观测值的时间间隔不同。例如，观测值的时间间隔为“天”时，您可以选择“月”作为分析时间间隔。这样，系统先将每日数据聚集为每月数据，然后再构建模型。您还可以选择将时间间隔较长的数据分布到较短的时间间隔内。例如，如果观测值是每季度数据，那么您可以将每季度数据分布为每月数据。

执行分析所采用的时间间隔的可用选项取决于定义观测值的方式以及这些观测值的时间间隔。特别是，如果观测值由循环周期定义，或者您为活动数据集定义了日期指定项，那么仅支持聚集。在这种情况下，分析时间间隔必须大于或等于观测值的时间间隔。

通过“数据指定项”选项卡上的**时间间隔**指定分析的时间间隔。聚集或分布数据的方法在“规范”选项卡上的**聚集和分布**设置中指定。

聚集和分布

聚集函数

如果用于分析的时间间隔比观测值的时间间隔长，那么将对输入数据进行聚集。例如，当观测值的时间间隔为“天”并且分析时间间隔为“月”时，将执行聚集。可用的聚集函数如下所示：

mean、sum、mode、min 或 max。

分布函数

如果用于分析的时间间隔比观测值的时间间隔短，那么将对输入数据进行分布。例如，当观测值的时间间隔为“季度”并且分析时间间隔为“月”时，将执行分布。可用的分布函数如下所示：mean 或 sum。

分组函数

当观测值由日期/时间定义，并且同一个时间间隔内存在多个观测值时，将进行分组。例如，如果观测值的时间间隔为“月”，那么同一个月的多个日期将分组到一起，并与它们所在的月份相关联。可用的分组函数如下所示：mean、sum、mode、min 或 max。当观测值由日期/时间定义，并且观测值的时间间隔指定为“不规则”时，将始终执行分组。

注：尽管分组是一种聚集形式，但在对缺失值进行任何处理之前执行，而正式的聚集是在对所有缺失值进行处理之后执行。如果观测值的时间间隔指定为“不规则”，那么将仅使用分组函数来执行聚集。

将跨天观测值聚集到前一天

指定是否将时间跨天边界的观测值聚集到前一天的值。例如，如果每一天的时间范围是从 20:00 开始的 8 小时，那么对于每小时观测值，此设置指定是否将介于 00:00 与 04:00 之间的观测值包括在前一天的聚集结果中。仅当观测值的时间间隔为“每天的小时数”、“每天的分钟数”或“每天的秒数”，并且分析时间间隔为“天”时，此设置才适用。

所指定字段的定制设置

您可以对每个字段指定聚集函数、分布函数和分组函数。这些设置将覆盖聚集函数、分布函数和分组函数的缺省设置。

缺失值

输入数据中的缺失值将替换为插补值。提供了下列替换方法：

线性插值

使用线性插值法替换缺失值。缺失值之前的最后一个有效值以及之后的第一个有效值用于插值。如果序列中的第一个或最后一个观测值具有缺失值，那么将使用序列开头或末尾的两个最近邻非缺失值。

序列平均值

使用整个序列的平均值替换缺失值。

邻近点的平均值

使用有效周围值的平均值替换缺失值。邻近点的跨度为缺失值上下用于计算平均值的有效值数目。

邻近点的中位数

使用有效周围值的中位数替换缺失值。邻近点的跨度为缺失值上下用于计算中位数的有效值数目。

线性趋势

此选项使用序列中的所有非缺失观测值来拟合简单线性回归模型，然后使用该模型来插补缺失值。

其他设置：

缺失值的最大百分比 (%)

指定任何序列所允许的缺失值的最大百分比。在分析中，将排除缺失值数目超过指定最大值的序列。

用户缺失值

此选项指定是否将用户缺失值视为有效数据，并因此将其包括在序列中。缺省情况下，将排除用户缺失值并将其当作系统缺失值进行处理，随后再进行插补。

常规数据选项

每个维度字段的相异值的最大数目

此设置应用于多维数据，用于指定任意一个维度字段所允许的相异值的最大数目。缺省情况下，此限制设置为 10000，但可以增大为任意大的数字。

常规构建选项

置信区间宽度 (%)

此设置控制预测及模型参数的置信区间。您可以指定任何小于 100 的正数值。缺省情况下，将使用 95% 置信区间。

每个目标的最大输入数目

此设置指定每个目标的模型中允许的最大输入数目。您可以指定 1 到 20 范围内的整数。每个目标的模型都始终包含自身的延迟值，因此将此值设置为 1 表示唯一的输入是目标自身。

模型误差

此设置控制用于确定每个目标的最佳输入集合的迭代式过程。您可以指定大于 0 的任何值。缺省值为 0.001。

离群值阈值 (%)

如果根据模型计算而得的可能性指出观测值为超出此阈值的离群值，那么会将该观测值标记为离群值。您可以指定 50 到 100 范围内的值。

每个输入的最大延迟项数

此设置指定每个目标的模型中每个输入的延迟项数。缺省情况下，从用于分析的时间间隔自动确定延迟项数。例如，如果时间间隔为“月”（增量为一个月），那么延迟项数为 12。另外，您可以显式指定延迟项数。指定的值必须是 1 到 20 范围内的整数。

要显示的序列

这些选项指定要显示其输出的序列（目标或输入）。所指定序列的输出内容由**输出选项**设置确定。

显示与最佳拟合模型相关联的目标

缺省情况下，将显示与 10 个最佳拟合模型（由 R 平方值确定）相关联的目标的输出。您可以指定另一个固定的最佳拟合模型数，也可以指定最佳拟合模型百分比。另外，还可以从以下拟合优度量中选择：

R 方 线性模型的拟合优度，有时称为判定系数。它是目标变量的变动中，由模型解释的比例。其值的范围为 0 到 1。如果值较小，那么表示该模型与数据的拟合度不好。

根均方百分比误差

用于度量序列的模型预测值与实测值之间的差异程度。它与使用的单位无关，因此可用于比较具有不同单位的序列。

均方根误差

均方误差的平方根。度量因变量序列与其模型预测水平的相差程度，用和因变量序列相同的单位表示。

BIC BIC 准则。选择和比较基于 -2 对数似然估计的模型的度量。值越小，模型越好。BIC 也会“惩罚”过多参数模型（例如，具有大量输入的复杂模型），但比 AIC 更严格。

AIC Akaike 信息标准。选择和比较基于 -2 对数似然估计的模型的度量。值越小，模型越好。AIC“惩罚”过多参数模型（例如，具有大量输入的复杂模型）。

指定各个序列

您可以指定需要其输出的各个序列。

- 对于基于列的数据，请指定包含所需序列的字段。所指定字段的顺序定义了这些字段在输出中的显示顺序。
- 对于多维数据，请通过在网格中针对包含序列的度量字段添加条目来指定特定序列。然后，指定定义了该序列的维度字段的值。
 - 您可以直接在网格中输入每个维度字段的值，也可以从可用维度值的列表中选择。要从可用维度值的列表中选择，请单击所需维度的单元格中的省略号按钮。此操作将打开“选择维度值”子对话框。

- 您可以通过在“选择维度值”子对话框中单击望远镜图标并指定搜索项来搜索维度值列表。空格被视为搜索项的组成部分。搜索项中的星号 (*) 并不表示通配符。
- 网格中的序列顺序定义了这些序列在输出中的显示顺序。

对于基于列的数据和多维数据，输出都限制为 30 个序列。此限制包括所指定的各个序列（输入或目标）以及与最佳拟合模型相关联的目标。逐个指定的序列优先于与最佳拟合模型相关联的目标。

输出选项

这些选项指定输出内容。目标的输出组中的选项为**要显示的序列**设置所指定的最佳拟合模型的相关目标生成输出。序列的输出组中的选项为**要显示的序列**设置所指定的各个序列生成输出。

总体模型系统

显示模型系统中各个序列之间的因果关系的图形表示。所显示目标的模型拟合度统计表和离群值表都将包括在输出项中。如果在**序列的输出**组中选择了此选项，那么将为**要显示的序列**设置所指定的每个序列创建一个单独的输出项。

序列之间的因果关系具有相关的显著性水平，显著性水平较低表示关联较为重要。您可以选择隐藏显著性水平大于指定值的关联。

模型拟合度统计和离群值

这些表列出选择显示的目标序列的模型拟合度统计和离群值。这些表与“总体模型系统”显示内容中的表包含相同的信息。这些表支持所有的标准透视表功能和编辑表功能。

模型效应和模型参数

这些表列出选择显示的目标序列的模型效应检验和模型参数。模型效应检验包括模型中包含的每个输入的 F 统计和相关显著性值。

影响图

显示感兴趣的序列与其所影响或影响其的其他序列之间因果关系的图形表示。影响感兴趣序列的序列称为**原因**。选择**效应**将生成进行初始化以显示效应的影响图。选择**原因**将生成进行初始化以显示原因的影响图。选择**包括原因和效用**将生成两个单独的影响图，即，一个初始化为原因，另一个初始化为效应。您可以在显示影响图的输出项中来回切换原因和效应。

您可以指定要显示的原因或效应的级别数，其中，第一个级别正好是感兴趣的序列。另外的每个级别都显示感兴趣序列的更多间接原因或效应。例如，在显示的效应中，第三个级别由那些包含第二个级别中的序列作为直接输入的序列组成。第三个级别中的序列受感兴趣序列间接影响，这是因为，感兴趣序列是第二个级别中的序列的直接输入。

序列图 这些图列出选择显示的目标序列的观测值和预测值。请求进行预测时，此图还显示这些预测的预测值和置信区间。

残差图 这些图列出选择显示的目标序列的模型残差。

主要输入

这些图列出所显示的每个目标的长期状况以及该目标的 3 个主要输入。主要输入是显著性值最小的输入。为了支持输入和目标采用不同的刻度，y 轴表示每个序列的 z 得分。

预测表 这些表列出选择显示的目标序列的预测值及这些预测的置信区间。

离群值根本原因分析

确定哪些序列最有可能是感兴趣序列中各个离群值的原因。离群值根本原因分析针对**要显示的序列**设置中指定的各个序列的列表中包括的每个目标序列执行。

输出

交互离群值表和图表

这个表和图表列出每个感兴趣序列的离群值及离群值根本原因。对于每个离群值，表中都包含一行。图表是影响图。在表中选择一行将在影响图中突出显示一条路径，该路径从感兴趣序列通往最有可能引起相关联离群值的序列。

离群值的透视表

这个表列出每个感兴趣序列的离群值及离群值根本原因。这个表与交互式显示中的表包含相同的信息。这个表支持所有的标准透视表功能和编辑表功能。

因果级别

您可以指定搜索根本原因时要包括的级别数目。此处使用的级别概念与针对影响图描述的概念相同。

所有模型之间的模型拟合度

这个直方图根据所有模型及所选拟合统计来显示模型拟合度。提供了下列拟合统计：

R 方 线性模型的拟合优度，有时称为判定系数。它是目标变量的变动中，由模型解释的比例。其值的范围为 0 到 1。如果值较小，那么表示该模型与数据的拟合度不好。

根均方百分比误差

用于度量序列的模型预测值与实测值之间的差异程度。它与使用的单位无关，因此可用于比较具有不同单位的序列。

均方根误差

均方误差的平方根。度量因变量序列与其模型预测水平的相差程度，用和因变量序列相同的单位表示。

BIC BIC 准则。选择和比较基于 -2 对数似然估计的模型的度量。值越小，模型越好。BIC 也会“惩罚”过多参数模型（例如，具有大量输入的复杂模型），但比 AIC 更严格。

AIC Akaike 信息标准。选择和比较基于 -2 对数似然估计的模型的度量。值越小，模型越好。AIC“惩罚”过多参数模型（例如，具有大量输入的复杂模型）。

长期离群值

这个条形图显示估计期的每个时间间隔内所有目标的离群值数目。

序列转换

这个表列出已对模型系统中的序列应用的所有转换。可能的转换包括缺失值插补、聚集和分布。

估计期

缺省情况下，估计期从所有序列中的最早观测值的时间开始，并以最晚观测值的时间结束。

按开始时间和结束时间

您可以同时指定估计期的开始时间和结束时间，也可以仅指定开始时间或者仅指定结束时间。如果省略了估计期的开始时间或结束时间，那么将使用缺省值。

- 如果观测值由日期/时间字段定义，请以用于该日期/时间字段的格式输入开始时间值和结束时间值。
- 对于由循环周期定义的观测值，请对每个循环周期字段指定值。每个字段都将显示在单独的列中。
- 如果存在对活动数据集生效的日期指定项，那么您必须对该日期指定项的每个组成部分（例如月份）指定值。每个组成部分都将显示在单独的列中。
- 如果观测值由记录顺序定义，那么估计期的开始时间和结束时间由相关个案的行号（显示在数据编辑器中）定义。

按最晚或最早时间间隔

将估计期定义为指定数目的时间间隔，这些时间间隔从数据中的最早时间间隔开始或者以最晚时间间隔结束，并可以使用可选的偏移量。在此上下文中，时间间隔是指分析时间间隔。例如，假定观测值按月获取，但分析时间间隔为季度。指定**最晚**并对**时间间隔数目**指定值 24 表示最晚的 24 个季度。

您可以选择性地排除指定数目的时间间隔。例如，指定最晚的 24 个时间间隔并对排除数目指定 1 表示估计期由最后一个时间间隔之前的 24 个时间间隔组成。

预测

将**记录扩展至将来**选项用于设置时间间隔数目，以预测估计期结束之后的情况。在这种情况下，时间间隔为“数据指定项”选项卡上指定的分析时间间隔。请求进行预测时，将为所有并非同时作为目标的输入序列自动构建自回归模型。然后，使用这些模型生成这些输入序列在预测期内的值。

保存

目标选项

可以将数据转换（例如对缺失值进行的聚集或插补）以及对新变量（在**保存目标**设置中指定）进行的转换保存到 IBM SPSS Statistics 数据文件或者当前会话中的新数据集。保存的数据中的日期/时间值将与每个时间间隔的开始时间（例如当月的第一天）对齐，并表示模型系统的分析时间间隔。仅在观测值由日期指定项或记录顺序定义，并且数据未进行聚集时，您才可以将任何新变量保存到活动数据集中。

保存目标

可以将模型预测、置信区间和残差另存为新变量。要保存的每个指定目标都将生成自己的一组新变量，每个新变量都包含估计期和预测期的值。对于预测值、置信区间和噪声残值，您可以指定要用作新变量前缀的根名称。完整变量名称是根名称与目标序列所在字段的名称的并置。根名称必须符合有效变量名称的规则。需要时可以扩展变量名以避免变量命名冲突。

指示包含预测的个案数

创建一个变量，用于指示记录是否包含预测数据。您可以指定变量名称。缺省值为 *ForecastIndicator*。

要保存的目标

指定是为模型系统中的所有目标序列创建新变量，还是仅为**要显示的序列**设置所指定的目标序列创建新变量。

导出模型系统

将模型系统保存到压缩文件归档（.zip 文件）。“时间因果模型预测”过程可以使用此模型系统文件，以获取经过更新的预测或者生成任何可用输出。另外，此文件还可以由“时间因果模型方案”过程用于运行方案分析。

交互式输出

时间因果建模输出包括众多交互式输出对象。通过在输出查看器中激活（双击）对象，即可使用交互式功能。

总体模型系统

显示模型系统中各个序列之间的因果关系。将特定目标连接到其输入的所有线具有相同颜色。线的宽度指示因果关系的显著性，较粗的线表示关联更为显著。并非同时作为目标的输入以黑色方块指示。

- 可以针对主要模型、指定的序列、所有序列或者无输入模型来显示关系。主要模型是满足您通过**要显示的序列**设置对最佳拟合模型指定的条件的模型。
- 通过依次在图表中选择序列名称、右键单击并从上下文菜单中选择**创建影响图**，可以为一个或多个序列生成影响图。

- 您可以选择隐藏显著性水平大于指定值的因果关系。显著性水平较小表明因果关系较为显著。
- 通过依次在图表中选择序列名称、右键单击并从上下文菜单中选择**突出显示序列显著性**，可以显示一个特定序列的显著性。

影响图 显示感兴趣的序列与其所影响或影响其的其他序列之间因果关系的图形表示。影响感兴趣序列的序列称为**原因**。

- 您可以通过指定所需序列的名称来更改感兴趣序列。在影响图中双击任何节点都会将感兴趣序列更改为该节点的相关联序列。
- 可以在原因与效应之间切换显示，并可以更改要显示的原因或效应的级别数。
- 单击任何节点将打开与该节点相关联的序列的详细时序图。

离群值根本原因分析

确定哪些序列最有可能是感兴趣序列中各个离群值的原因。

- 通过在“离群值”表中选择离群值的行，可以显示任何离群值的根本原因。另外，还可以通过在序列图中单击离群值的图标来显示根本原因。
- 单击任何节点将打开与该节点相关联的序列的详细时序图。

总体模型质量

这个直方图根据所有模型及特定拟合统计来显示模型拟合度。单击条形图中的条形将对点图进行过滤，从而只显示与所选条形相关联的模型。通过指定序列名称，可以在点图中找到特定目标序列的模型。

离群值分布

这个条形图显示估计期的每个时间间隔内所有目标的离群值数目。单击条形图中的条形将对点图进行过滤，从而只显示与所选条形相关联的离群值。

注:

- 如果保存包含时间因果建模的交互式输出的输出文档，并且想要保留交互式功能，那么请确保在“将输出另存为”对话框上选中**随输出文档一起存储必需的模型信息**。
- 某些交互式功能要求活动数据集是用来构建时间因果模型系统的数据。

第 7 章 应用时间因果模型

应用时间因果模型

您可以通过两个过程来应用使用“时间因果建模”过程创建的模型。这两个过程都需要模型系统文件，您可以在“时间因果建模”过程中保存该文件。

时间因果模型预测

您可以通过此过程获得有更新数据的序列的预测值，而不必重建模型。另外，还可以生成任何可用于“时间因果建模”过程的输出。

时间因果模型方案

使用此过程可以调查模型系统中特定时间序列的特定值如何影响在因果上与其相关的时间序列的预测值。

时间因果模型预测

“时间因果模型预测”过程装入由“时间因果建模”过程创建的模型系统文件，并将模型应用于活动数据集。您可以通过此过程获得有更新数据的序列的预测值，而不必重建模型。另外，还可以生成任何可用于“时间因果建模”过程的输出。

假设

- 活动数据集中的数据结构（基于列或者多维）必须与构建模型系统时使用的结构相同。对于多维数据，维度字段必须与用于构建模型系统的字段相同。并且，用于构建模型系统的维度值在活动数据集中必须存在。
- 模型将应用于活动数据集中那些与模型系统中指定的字段同名的字段。
- 构建模型系统时用于定义观测值的一个或多个字段在活动数据集中必须存在。假定各个观测值之间的时间间隔与构建模型时相同。如果观测值由日期指定项定义，那么同一日期指定项在活动数据集中必须存在。日期指定项通过“定义日期”对话框或 DATE 命令进行创建。
- 分析时间间隔以及任何聚集设置、分布设置和缺失值设置与构建模型时的相应设置相同。

使用时间因果模型预测

此功能需要“统计预测”选项。

从菜单中选择:

分析 > 预测 > 应用时间因果模型...

1. 输入指定作为模型系统文件的文件，或者单击**浏览**并选择模型系统文件。模型系统文件由“时间因果建模”过程创建。
2. 单击用于重新估算模型、创建预测并生成输出的选项。
3. 单击**继续**。
4. 指定是要使用现有的模型参数，还是要根据活动数据集中的数据重新估算模型参数。
5. 指定要预测未来的多长时间，或者指定不进行预测。
6. 单击**选项**以指定输出内容。

- 单击**保存**以保存预测，并在重新估算模型参数时将更新后的模型系统导出到外部文件。
- 单击**运行**以运行此过程。

模型参数和预测

从模型文件装入

通过使用模型系统文件中的模型参数以及活动数据集中的数据生成预测，而无需重新估算模型参数。输出中显示的用来选择最佳拟合模型的拟合优度量从模型系统文件中获取。这样，拟合度量将反映开发（或最近一次更新）每个模型时使用的数据。此选项适用于根据先前用来构建模型系统的数据生成预测和输出。

从数据重新估算

通过使用活动数据集中的数据来重新估算模型参数。重新估算模型参数不会影响将哪些输入包括在每个目标的模型中。当存在原始估计期以外的新数据，并且您希望使用更新后的数据来生成预测或其他输出时，此选项适用。

所有观测值

指定估计期从所有序列中的最早观测值的时间开始，并以最晚观测值的时间结束。

按开始时间和结束时间

您可以同时指定估计期的开始时间和结束时间，也可以仅指定开始时间或者仅指定结束时间。如果省略了估计期的开始时间或结束时间，那么将使用缺省值。

- 如果观测值由日期/时间字段定义，请以用于该日期/时间字段的格式输入开始时间值和结束时间值。
- 对于由循环周期定义的观测值，请对每个循环周期字段指定值。每个字段都将显示在单独的列中。
- 如果存在对活动数据集生效的日期指定项，那么您必须对该日期指定项的每个组成部分（例如月份）指定值。每个组成部分都将显示在单独的列中。
- 如果观测值由记录顺序定义，那么估计期的开始时间和结束时间由相关个案的行号（显示在数据编辑器中）定义。

按最晚或最早时间间隔

将估计期定义为指定数目的时间间隔，这些时间间隔从数据中的最早时间间隔开始或者以最晚时间间隔结束，并可以使用可选的偏移量。在此上下文中，时间间隔是指分析时间间隔。例如，假定观测值按月获取，但分析时间间隔为季度。指定**最晚**并对**时间间隔数目**指定值 24 表示最晚的 24 个季度。

您可以选择性地排除指定数目的时间间隔。例如，指定最晚的 24 个时间间隔并对排除数目指定 1 表示估计期由最后一个时间间隔之前的 24 个时间间隔组成。

将记录扩展至将来

设置时间间隔数目，以预测估计期结束之后的情况。在这种情况下，时间间隔为分析时间间隔。请求进行预测时，将为所有并非同时作为目标的输入序列自动构建自回归模型。然后，使用这些模型生成这些输入序列在预测期内的值，以获取对这些输入的目标的预测。

一般选项

置信区间宽度 (%)

此设置控制预测及模型参数的置信区间。您可以指定任何小于 100 的正数值。缺省情况下，将使用 95% 置信区间。

离群值阈值 (%)

如果根据模型计算而得的可能性指出观测值为超出此阈值的离群值，那么会将该观测值标记为离群值。您可以指定 50 到 100 范围内的值。

要显示的序列

这些选项指定要显示其输出的序列（目标或输入）。所指定序列的输出内容由**输出选项**设置确定。

显示与最佳拟合模型相关联的目标

缺省情况下，将显示与 10 个最佳拟合模型（由 R 平方值确定）相关联的目标的输出。您可以指定另一个固定的最佳拟合模型数，也可以指定最佳拟合模型百分比。另外，还可以从以下拟合优度量中选择：

R 方 线性模型的拟合优度，有时称为判定系数。它是目标变量的变动中，由模型解释的比例。其值的范围为 0 到 1。如果值较小，那么表示该模型与数据的拟合度不好。

根均方百分比误差

用于度量序列的模型预测值与实测值之间的差异程度。它与使用的单位无关，因此可用于比较具有不同单位的序列。

均方根误差

均方误差的平方根。度量因变量序列与其模型预测水平的相差程度，用和因变量序列相同的单位表示。

BIC BIC 准则。选择和比较基于 -2 对数似然估计的模型的度量。值越小，模型越好。BIC 也会“惩罚”过多参数模型（例如，具有大量输入的复杂模型），但比 AIC 更严格。

AIC Akaike 信息标准。选择和比较基于 -2 对数似然估计的模型的度量。值越小，模型越好。AIC“惩罚”过多参数模型（例如，具有大量输入的复杂模型）。

指定各个序列

您可以指定需要其输出的各个序列。

- 对于基于列的数据，请指定包含所需序列的字段。所指定字段的顺序定义了这些字段在输出中的显示顺序。
- 对于多维数据，请通过在网格中针对包含序列的度量字段添加条目来指定特定序列。然后，指定定义了该序列的维度字段的值。
 - 您可以直接在网格中输入每个维度字段的值，也可以从可用维度值的列表中选择。要从可用维度值的列表中选择，请单击所需维度的单元格中的省略号按钮。此操作将打开“选择维度值”子对话框。
 - 您可以通过在“选择维度值”子对话框中单击望远镜图标并指定搜索项来搜索维度值列表。空格被视为搜索项的组成部分。搜索项中的星号 (*) 并不表示通配符。
 - 网格中的序列顺序定义了这些序列在输出中的显示顺序。

对于基于列的数据和多维数据，输出都限制为 30 个序列。此限制包括所指定的各个序列（输入或目标）以及与最佳拟合模型相关联的目标。逐个指定的序列优先于与最佳拟合模型相关联的目标。

输出选项

这些选项指定输出内容。**目标的输出组**中的选项为**要显示的序列**设置所指定的最佳拟合模型的相关目标生成输出。**序列的输出组**中的选项为**要显示的序列**设置所指定的各个序列生成输出。

总体模型系统

显示模型系统中各个序列之间的因果关系的图形表示。所显示目标的模型拟合度统计和离群值表都将包括在输出项中。如果在**序列的输出组**中选择了此选项，那么将为**要显示的序列**设置所指定的每个序列创建一个单独的输出项。

序列之间的因果关系具有相关的显著性水平，显著性水平较低表示关联较为重要。您可以选择隐藏显著性水平大于指定值的关联。

模型拟合度统计和离群值

这些表列出选择显示的目标序列的模型拟合度统计和离群值。这些表与“总体模型系统”显示内容中的表包含相同的信息。这些表支持所有的标准透视表功能和编辑表功能。

模型效应和模型参数

这些表列出选择显示的目标序列的模型效应检验和模型参数。模型效应检验包括模型中包含的每个输入的 F 统计和相关显著性值。

影响图

显示感兴趣的序列与其所影响或影响其的其他序列之间因果关系的图形表示。影响感兴趣序列的序列称为**原因**。选择**效应**将生成进行初始化以显示效应的影响图。选择**原因**将生成进行初始化以显示原因的影响图。选择**包括原因和效用**将生成两个单独的影响图，即，一个初始化为原因，另一个初始化为效应。您可以在显示影响图的输出项中来回切换原因和效应。

您可以指定要显示的原因或效应的级别数，其中，第一个级别正好是感兴趣的序列。另外的每个级别都显示感兴趣序列的更多间接原因或效应。例如，在显示的效应中，第三个级别由那些包含第二个级别中的序列作为直接输入的序列组成。第三个级别中的序列受感兴趣序列间接影响，这是因为，感兴趣序列是第二个级别中的序列的直接输入。

序列图 这些图列出选择显示的目标序列的观测值和预测值。请求进行预测时，此图还显示这些预测的预测值和置信区间。

残差图 这些图列出选择显示的目标序列的模型残差。

主要输入

这些图列出所显示的每个目标的长期状况以及该目标的 3 个主要输入。主要输入是显著性值最小的输入。为了支持输入和目标采用不同的刻度，y 轴表示每个序列的 z 得分。

预测表 这些表列出选择显示的目标序列的预测值及这些预测的置信区间。

离群值根本原因分析

确定哪些序列最有可能是感兴趣序列中各个离群值的原因。离群值根本原因分析针对**要显示的序列**设置中指定的各个序列的列表中包括的每个目标序列执行。

输出

交互离群值表和图表

这个表和图表列出每个感兴趣序列的离群值及离群值根本原因。对于每个离群值，表中都包含一行。图表是影响图。在表中选择一行将在影响图中突出显示一条路径，该路径从感兴趣序列通往最有可能引起相关联离群值的序列。

离群值的透视表

这个表列出每个感兴趣序列的离群值及离群值根本原因。这个表与交互式显示中的表包含相同的信息。这个表支持所有的标准透视表功能和编辑表功能。

因果级别

您可以指定搜索根本原因时要包括的级别数目。此处使用的级别概念与针对影响图描述的概念相同。

所有模型之间的模型拟合度

这个直方图根据所有模型及所选拟合统计来显示模型拟合度。提供了下列拟合统计：

R 方 线性模型的拟合优度，有时称为判定系数。它是目标变量的变动中，由模型解释的比例。其值的范围为 0 到 1。如果值较小，那么表示该模型与数据的拟合度不好。

根均方百分比误差

用于度量序列的模型预测值与实测值之间的差异程度。它与使用的单位无关，因此可用于比较具有不同单位的序列。

均方根误差

均方误差的平方根。度量因变量序列与其模型预测水平的相差程度，用和因变量序列相同的单位表示。

BIC BIC 准则。选择和比较基于 -2 对数似然估计的模型的度量。值越小，模型越好。BIC 也会“惩罚”过多参数模型（例如，具有大量输入的复杂模型），但比 AIC 更严格。

AIC Akaike 信息标准。选择和比较基于 -2 对数似然估计的模型的度量。值越小，模型越好。AIC“惩罚”过多参数模型（例如，具有大量输入的复杂模型）。

长期离群值

这个条形图显示估计期的每个时间间隔内所有目标的离群值数目。

序列转换

这个表列出已对模型系统中的序列应用的所有转换。可能的转换包括缺失值插补、聚集和分布。

保存

保存目标

可以将模型预测、置信区间和残差另存为新变量。要保存的每个指定目标都将生成自己的一组新变量，每个新变量都包含估计期和预测期的值。对于预测值、置信区间和噪声残值，您可以指定要用作新变量前缀的根名称。完整变量名称是根名称与目标序列所在字段的名称的并置。根名称必须符合有效变量名称的规则。需要时可以扩展变量名以避免变量命名冲突。

指示包含预测的个案数

创建一个变量，用于指示记录是否包含预测数据。您可以指定变量名称。缺省值为 *ForecastIndicator*。

要保存的目标

指定是为模型系统中的所有目标序列创建新变量，还是仅为要显示的序列设置所指定的目标序列创建新变量。

目标选项

可以将数据转换（例如对缺失值进行的聚集或插补）以及对新变量（在保存目标设置中指定）进行的转换保存到 IBM SPSS Statistics 数据文件或者当前会话中的新数据集。保存的数据中的日期/时间值将与每个时间间隔的开始时间（例如当月的第一天）对齐，并表示模型系统的分析时间间隔。仅在观测值由日期指定项或记录顺序定义，并且数据未进行聚集时，您才可以将任何新变量保存到活动数据集中。

导出模型系统

将模型系统保存到压缩文件归档（.zip 文件）。该模型系统文件可以由此过程复用。另外，此文件还可以由“时间因果模型方案”过程用于运行方案分析。仅当重新估算模型参数时，此选项才可用。

时间因果模型方案

“时间因果模型方案”过程使用活动数据集中的数据对时间因果模型系统运行用户定义的方案。方案由称为根序列的时间序列以及该序列在指定时间范围内的一组用户定义值定义。指定的值随后将用来生成对根序列所影响的时间序列的预测。此过程需要一个由“时间因果建模”过程创建的模型系统文件。假定活动数据集就是用来创建模型系统文件的那些数据。

示例

通过使用“时间因果建模”过程，业务决策制定者发现了影响许多重要绩效指标的关键度量。该度量可控制，因此决策制定者希望调查该度量的各组值对下一个季度的影响。通过将模型系统文件装入到“时间因果模型方案”过程中，并对该关键度量指定值集，很容易完成此调查。

运行时间因果模型方案

此功能需要“统计预测”选项。

从菜单中选择：

分析 > 预测 > 应用时间因果模型...

1. 输入指定作为模型系统文件的文件，或者单击**浏览**并选择模型系统文件。模型系统文件由“时间因果建模”过程创建。
2. 单击用于运行方案的选项。
3. 单击**继续**。
4. 在“时间因果模型方案”对话框中的“方案”选项卡上，单击**定义方案期**并指定方案期。
5. 对于基于列的数据，单击**添加方案**以定义每个方案。对于多维数据，请单击**添加方案**以定义每个方案，并单击**添加方案组**以定义每个方案组。
6. 单击**选项**以指定输出内容，并指定该方案所影响的序列广度。
7. 单击**运行**以运行此过程。

定义方案期

方案期是您指定用于运行方案的值所在的周期。方案期可以在估计期结束之前或之后开始。您可以选择性地指定预测到方案期结束之后。缺省情况下，将生成截至方案期结束为止的预测。所有方案将使用相同的方案期和指定项来确定预测时间范围。

注：预测从方案期开始后的第一个时间段开始。例如，如果方案期从 2014-11-01 开始并且时间间隔为月，那么第一个预测针对于 2014-12-01。

按开始时间、结束时间和预测时间范围来指定

- 如果观测值由日期/时间字段定义，请以用于该日期/时间字段的格式输入开始时间值、结束时间值和预测时间范围值。日期/时间字段的值与相关联时间间隔的开始时间一致。例如，如果分析时间间隔为“月”，那么值 10/10/2014 将调整为 10/01/2014，即当月的第一天。
- 对于由循环周期定义的观测值，请对每个循环周期字段指定值。每个字段都将显示在单独的列中。
- 如果存在对活动数据集生效的日期指定项，那么您必须对该日期指定项的每个组成部分（例如月份）指定值。每个组成部分都将显示在单独的列中。
- 如果观测值由记录顺序定义，那么开始时间、结束时间和预测时间范围由相关个案的行号（显示在数据编辑器中）定义。

按相对于估计期结束时间的的时间间隔数来指定

以相对于估计期结束时间的时间间隔数方式定义开始时间和结束时间，其中，时间间隔是指分析时间间隔。估计期结束时间定义为时间间隔 0。估计期结束时间之前的时间间隔的值为负，估计期结束时间之后的时间间隔的值为正。另外，您还可以指定方案期结束后要预测的时间间隔数。缺省值为 0。

例如，假定分析时间间隔为“月”，并且您指定了 1 作为开始时间间隔，指定了 3 作为结束时间间隔，并指定了 1 作为在此之后要预测的时间间隔数。这样，方案期是估计期结束后的 3 个月。因此，将为方案期的第 2 个和第 3 个月生成预测，并为方案期结束后的第 1 个月生成预测。

添加方案和方案组

“方案”选项卡指定要运行的方案。要定义方案，您必须先通过单击**定义方案期**来定义方案期。要创建方案和方案组（仅适用于多维数据），请单击相关联的**添加方案**或**添加方案组**按钮。通过在相关联的网格中选择特定方案或方案组，可以对其进行编辑、复制或删除。

基于列的数据

网格中的**根字段**列指定一个时间序列字段，此字段的值将替换为方案值。**方案值**列按最早到最晚顺序显示指定的方案值。如果方案值由表达式定义，那么此列将显示该表达式。

多维数据

各个方案

“各个方案”网格中的每一行都指定一个时间序列，该时间序列的值将替换为指定的方案值。此序列由**根度量**列中指定的字段与每个维度字段的指定值的组合定义。**方案值**列的内容与基于列的数据相同。

方案组 方案组定义一组方案，这些方案基于单个根度量字段和多个维度值集合。对于指定的度量字段，每个维度值集合（每个维度字段各有一个对应的值）都定义了一个时间序列。对于这样的每一个时间序列，都将生成一个单独的方案，然后，这些时间序列的值替换为方案值。方案组的方案值由一个表达式指定，该表达式将应用于该组中的每个时间序列。

序列计数列显示与方案组相关联的维度值集合的数目。显示的值可能大于与该方案组相关联的时间序列的实际数目（每个集合各有一个对应的序列）。当指定的某些维度值组合未与该组的根度量所包含的序列相对应时，将发生这种情况。

作为方案组示例，假定存在度量字段 *advertising* 以及两个维度字段 *region* 和 *brand*。您可以定义以 *advertising* 作为根度量并且包含 *region* 与 *brand* 的所有组合的方案组。然后，您可以指定 *advertising**1.2 作为表达式，以调查将 *advertising* 增大 20% 对每个与 *advertising* 字段相关联的时间序列的影响。如果 *region* 有 4 个值，并且 *brand* 有 2 个值，那么共有 8 个这样的时间序列，因此这个组将定义 8 个方案。

方案定义

用于定义方案的设置取决于您的数据是基于列的数据还是多维数据。

根序列 指定方案的根序列。每个方案都基于单个根序列。对于基于列的数据，请选择用于定义根序列的字段。对于多维数据，请通过在网格中针对包含序列的度量字段添加条目来指定根序列。然后，指定用于定义根序列的维度字段的值。在指定维度值时，下列操作适用：

- 您可以直接在网格中输入每个维度字段的值，也可以从可用维度值的列表中选择。要从可用维度值的列表中选择，请单击所需维度的单元格中的省略号按钮。此操作将打开“选择维度值”子对话框。
- 您可以通过在“选择维度值”子对话框中单击望远镜图标并指定搜索项来搜索维度值列表。空格被视为搜索项的组成部分。搜索项中的星号 (*) 并不表示通配符。

指定受影响的目标

如果您知道受根序列影响的特定目标，并且希望仅调查对这些目标的影响，请使用此选项。缺省情况下，将自动确定根序列所影响的目标。您可以使用“选项”选项卡上的设置来指定此方案所影响的序列的广度。

对于基于列的数据，请选择所需的目标。对于多维数据，请通过在网格中针对包含目标序列的目标度量字段添加条目来指定目标序列。缺省情况下，将包括所指定度量字段中包含的所有序列。通过对一个或多个维度字段定制所包括的值，可以对包括的序列的集合进行定制。要对包括的维度值进行定制，请单击所需维度的省略号按钮。此操作将打开“选择维度值”对话框。

序列计数列（对于多维数据）显示当前对相关目标度量指定的维度值集合的数目。显示的值可能大于受影响目标序列的实际数目（每个集合各有一个对应的序列）。当指定的某些维度值组合未与相关目标度量所包含的序列相对应时，将发生这种情况。

方案标识

每个方案都必须具有唯一标识。此标识将显示在与该方案相关联的输出中。对于此标识的值，除了唯一性以外，没有任何限制。

指定根序列的方案值

使用此选项可以指定根序列在方案期中具有的显式值。必须对网格中列出的每个时间间隔指定一个数字值。通过单击**读取**、**预测**或**读取@5424测**，可以获取方案期中每个时间间隔的根序列值（实际值或预测值）。

指定根序列的方案值的表达式

您可以定义表达式，以计算根序列在方案期中的值。您可以直接输入表达式，也可以单击计算器按钮并从方案值表达式构建器中创建表达式。

- 表达式可以包含模型系统中的任何目标或输入。
- 如果方案期延伸到现有数据的时间范围以外，那么此表达式将应用于此表达式中各个字段的预测值。
- 对于多维数据，此表达式中的每个字段都指定一个时间序列，该时间序列由该字段以及您对根度量指定的维度值定义。这些时间序列用于对此表达式进行求值。

例如，假定根字段为 *advertising* 并且表达式为 $advertising*1.2$ 。此方案中使用的各个 *advertising* 值表示在现有值的基础上增大 20%。

注：要创建方案，请在“方案”选项卡上单击**添加方案**。

方案值表达式构建器： 使用方案值表达式构建器可以创建一个表达式，此表达式用于计算单个方案或方案组的方案值。要构建表达式，请将组成部分粘贴到**表达式**字段中，或者直接在**表达式**字段中输入内容。

- 表达式可以包含模型系统中的任何目标或输入。
- 通过从“函数组”列表中选择组，然后在“函数”列表中双击函数（或者选择函数，然后单击“函数组”列表旁边的箭头），可以粘贴函数。请输入所有标注了问号的参数。标注了**全部**的函数组提供了所有可用函数的列表。对话框的保留区显示了有关当前所选函数的简要描述。
- 字符串常量必须括在引号内。
- 如果值包含小数，那么必须使用句点 (.) 作为小数指示符。

要访问“方案值表达式构建器”，请单击“方案定义”或“方案组定义”对话框旁边的计算器按钮。

选择维度值： 对于多维数据，您可以定制用于定义方案或方案组所影响的目标的维度值。另外，您还可以定制用于为方案组定义根序列集合的维度值。

所有值 指定包括当前维度字段的所有值。这是缺省选项。

选择值 使用此选项可以指定当前维度字段的值集。您可以对要从中进行选择的值集进行过滤。满足过滤条件的值将显示在**匹配**选项卡中，不满足过滤条件的值将显示在**未选择的值**列表的**不匹配**选项卡中。**全部**选项卡列示所有未选择的值，而无论指定了什么过滤条件。

- 指定过滤器时，可以使用星号 (*) 来表示通配符。
- 要清除当前过滤器，请在“过滤显示的值”对话框中对搜索项指定空值。

要定制受影响目标的维度值，请完成下列步骤：

1. 在“方案定义”或“方案组定义”对话框中，选择要为其定制维度值的目标度量。
2. 单击要定制的维度的相应列中的省略号按钮。

要对方案组的根序列的维度值进行定制，请完成下列步骤：

1. 在“方案组定义”对话框中，单击要定制的维度的省略号按钮（位于根序列网格中）。

方案组定义

根序列 指定方案组的根序列集合。对于此集中的每个时间序列，都将生成一个单独的方案。请通过在网格中针对包含所需序列的度量字段添加条目来指定根序列。然后，指定用于定义集合的维度字段的值。缺省情况下，将包括所指定根度量字段中包含的所有序列。通过对一个或多个维度字段定制所包括的值，可以对包括的序列的集合进行定制。要对包括的维度值进行定制，请单击某个维度的省略号按钮。此操作将打开“选择维度值”对话框。

序列计数列显示当前为相关根度量包括的维度值集合的数目。显示的值可能大于该方案组的根序列的实际数目（每个集合各有一个对应的序列）。当指定的某些维度值组合未与根度量所包含的序列相对应时，将发生这种情况。

指定受影响的目标序列

如果您知道受根序列集合影响的特定目标，并且希望仅调查对这些目标的影响，请使用此选项。缺省情况下，将自动确定每个根序列所影响的目标。您可以使用“选项”选项卡上的设置来指定每个方案所影响的序列的广度。

请通过在网格中针对包含序列的度量字段添加条目来指定目标序列。缺省情况下，将包括所指定度量字段中包含的所有序列。通过对一个或多个维度字段定制所包括的值，可以对包括的序列的集合进行定制。要对包括的维度值进行定制，请单击所需维度的省略号按钮。此操作将打开“选择维度值”对话框。

序列计数列显示当前对相关目标度量指定的维度值集合的数目。显示的值可能大于受影响目标序列的实际数目（每个集合各有一个对应的序列）。当指定的某些维度值组合未与相关目标度量所包含的序列相对应时，将发生这种情况。

方案标识前缀

每个方案组都必须具有唯一的前缀。此前缀用于构造标识，该标识将显示在此方案组中每个方案的相关输出中。对于一个方案，其标识依次由前缀、下划线以及每个用于标识根序列的维度字段的值组成。各个维度值之间以下划线分隔。对于前缀的值，除了唯一性以外，没有任何限制。

根序列的方案值的表达式

方案组的方案值由表达式指定，该表达式用于计算该组中各个根序列的值。您可以直接输入表达式，也可以单击计算器按钮并从方案值表达式构建器中创建表达式。

- 表达式可以包含模型系统中的任何目标或输入。
- 如果方案期延伸到现有数据的时间范围以外，那么此表达式将应用于此表达式中各个字段的预测值。

- 对于组中的每个根序列，此表达式中的字段指定了时间序列，这些时间序列由这些字段以及用于定义根序列的维度值定义。这些时间序列用于对此表达式进行求值。例如，如果根序列由 `region='north'` 和 `brand='X'` 定义，那么表达式中使用的时间序列由同样的维度值定义。

例如，假定根度量字段为 *advertising*，并假定存在两个度量字段 *region* 和 *brand*。并且，假定方案组包含维度字段值的所有组合。然后，您可以指定 `advertising*1.2` 作为表达式，以调查将 *advertising* 增大 20% 对每个与 *advertising* 字段相关联的时间序列的影响。

注： 方案组仅应用于多维数据，要创建方案组，请在“方案”选项卡上单击**添加方案组**。

选项

受影响的目标的最大级别

指定受影响目标的最大级别数。后续每个级别（最多 5 个级别）都包含更间接地受根序列影响的目标。具体而言，第一个级别包含将根序列用作直接输入的目标。第二个级别中的目标将第一个级别中的目标用作直接输入，依此类推。增大此设置的值将增加计算的复杂程度，并可能影响性能。

最大自动检测目标数

指定要为每个根序列自动检测的受影响目标的最大数目。增大此设置的值将增加计算的复杂程度，并可能影响性能。

影响图 显示每个方案的根序列与其影响的目标序列之间的因果关系的图形表示。受影响目标的方案值以及预测值表都将包括在输出项中。图形包含受影响目标的预测值的图。单击影响图中的任何节点都将打开与该节点相关联的序列的详细时序图。对于每个方案，将生成单独的影响图。

序列图 生成每个方案中每个受影响目标的预测值序列图。

预测表和方案表

这些表列出每个方案的预测值和方案值。这些表与影响图中的表包含相同的信息。这些表支持所有的标准透视表功能和编辑表功能。

将置信区间包括在图和表中

指定是否将方案预测置信区间同时包括在图表和表输出中。

置信区间宽度 (%)

此设置用于控制方案预测置信区间。您可以指定任何小于 100 的正数值。缺省情况下，将使用 95% 置信区间。

第 8 章 拟合优度测量

本节提供了时间序列建模中使用的拟合优度测量的定义。

- **固定的 R 方 (Stationary R-squared)**. 将模型的平稳部分与简单平均值模型相比较的测量。当具有趋势或季节性模式时，该度量适用于普通 R 方。固定的 R 方可以为负值，其范围为负无穷到 1。负值表示考虑中的模型比基线模型差。正值表示考虑中的模型比基线模型好。
- **R 方 (R-squared)**. 总变动在由模型解释的序列中的比例估计。当序列很平稳时，此度量最有用。R 方可以是负无穷大到 1 范围中的负值。负值表示考虑中的模型比基线模型差。正值表示考虑中的模型比基线模型好。
- **RMSE**. 均方根误差。均方误差的平方根。度量因变量序列与其模型预测水平的相差程度，用和因变量序列相同的单位表示。
- **MAPE**. 平均绝对误差百分比。度量因变量序列与其模型预测水平的相差程度。它与使用的单位无关，因此可用于比较具有不同单位的序列。
- **MAE**. 平均绝对误差。度量序列与其模型预测水平的差别程度。MAE 以原始序列单位报告。
- **MaxAPE**. 最大绝对误差百分比。最大的预测误差，以百分比表示。该度量对于想象预测的最坏情况方案很有用。
- **MaxAE**. 最大绝对误差。最大的预测误差，以和因变量序列相同的单位表示。与 MaxAPE 相同，它对于想象预测的最坏情况方案很有用。最大绝对误差和最大绝对误差百分比可能发生在不同的序列点上，例如，当较大序列的绝对误差比较小值的绝对误差稍微大一些时。在此情况下，最大绝对误差将发生在较大序列值处，而最大绝对误差百分比将发生在较小序列值处。
- **标准化 BIC (Normalized BIC)**. 规范化的 BIC (BIC 准则)。尝试代表模型复杂性的模型整体拟合的一般度量。它是基于均方误差的分数，包括模型中参数数量的罚分和序列长度。罚分去除了具有更多参数的模型优势，从而可以容易地比较相同序列的不同模型的统计。

第 9 章 离群值类型

本节提供了时间序列建模中使用的离群值类型的定义。

- *加法 (Additive)*. 影响单个观察值的离群值。例如，可能将数据编码错误标识为可加离群值。
- *移位水平 (Level shift)*. 从某个特定的序列点开始将所有观察值移动一个常数的离群值。移位水平可能由于策略的更改而造成的。
- *创新的 (Innovational)*. 在某个特定的序列点附加到噪声项的离群值。对于平稳的序列，创新离群值将影响多个观察值。对于不平稳的序列，它可能影响在某个特定的序列点开始的每个观察值。
- *瞬时 (Transient)*. 其影响按指数衰减到 0 的离群值。
- *季节性可加 (Seasonal additive)*. 一个离群值，它影响特定观察值以及通过一个或多个季节性期间与之分隔的所有后续观察值。该离群值对所有这些观察值具有同等的影响。如果从某年开始，每个一月的销售额都增加，那么可能发生季节性可加离群值。
- *局部趋势 (Local trend)*. 从某个特定的序列点开始局部趋势的离群值。
- *可加修补 (Additive patch)*. 由两个或更多连续可加离群值构成的组。选择此离群值类型将导致除了检测加性离群值的组以外，还检测各个加性离群值。

第 10 章 ACF/PACF 图指南

此处所示的图是纯 ARIMA 过程或理论 ARIMA 过程的图。以下为过程的一些常用标识方法：

- 不稳序列的 ACF 的显著性会保留 6 个或一打以上的延迟，而不是迅速降低为 0。您必须区分此类序列，直至其稳定后，才可标识该过程。
- 自回归过程有一个指数不断衰减的 ACF，而且 PACF 的第一个或更多延迟中存在花序。花序数表示自回归的阶数。
- 移动平均值过程的 ACF 的第一个或更多延迟中存在花序，且有一指数不断衰减的 PACF。花序数表示移动平均值的阶数。
- 混合 (ARMA) 过程通常显示 ACF 和 PACF 中的指数衰减。

在标识阶段，您无需担心 ACF 或 PACF 的符号，也无需担心 ACF 或 PACF 的指数衰减至 0 的速度。这些都取决于 AR 和 MA 系数的符号和实际值。在某些实例中，指数不断衰减的 ACF 会在正值和负值之间来回变换。

根据实际数据绘制的 ACF 和 PACF 图决不会像此处显示的图那样规则。您必须学会在任何给定的图中分辨出基本规律。始终检查残差的 ACF 和 PACF，以防标识有误。请记住：

- 季节性过程将显示这些季节性延迟（季节性周期的倍数）的模式。
- 您有权将不显著值视为 0。也就是说，您可以忽略图中置信区间内的值。但是，您不必忽略它们，尤其在它们延续统计显著值的模式时。
- 偶然自相关有时会在统计上具有显著性。如果统计显著性是独立的，经常出现在高延迟上，而且未出现在季节性延迟中，则可以忽略此统计显著性。

有关 ACF 和 PACF 图的更完整的讨论，请参见与 ARIMA 分析有关的所有文本。

表 3. ARIMA(0,0,1), $q>0$

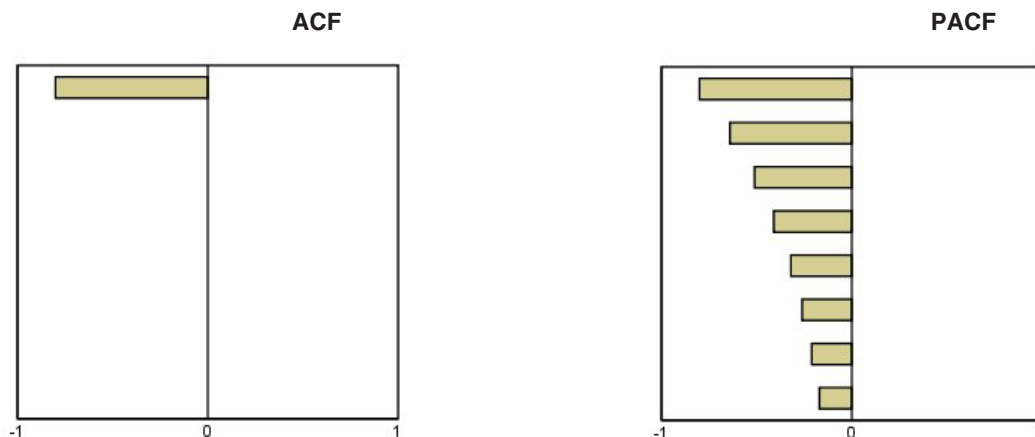
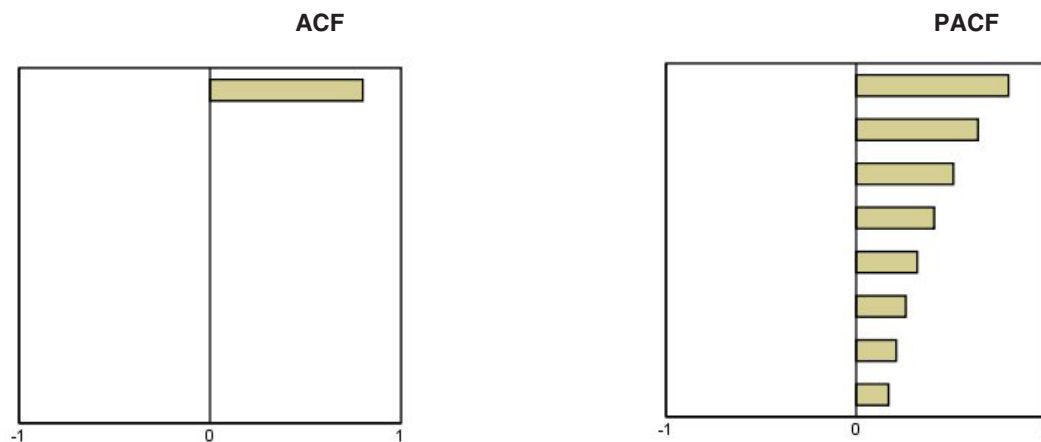


表 4. $ARIMA(0,0,1)$, $q < 0$



$ARIMA(0,0,2)$, $\theta_1 \theta_2 > 0$

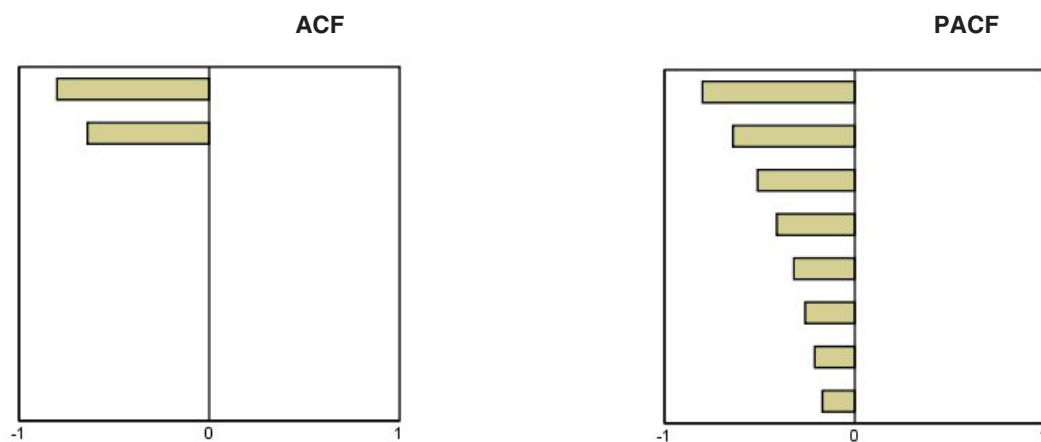


表 5. $ARIMA(1,0,0)$, $\phi > 0$

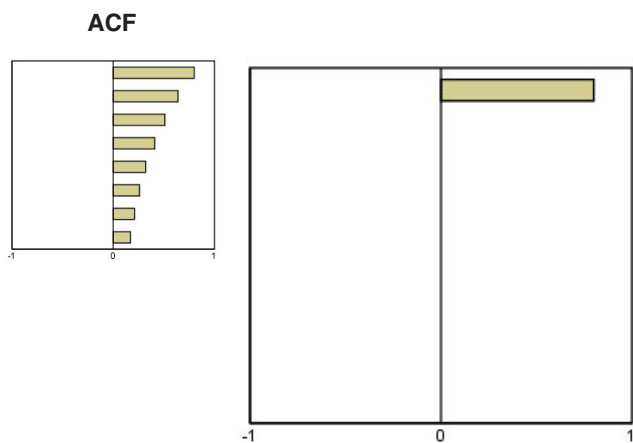
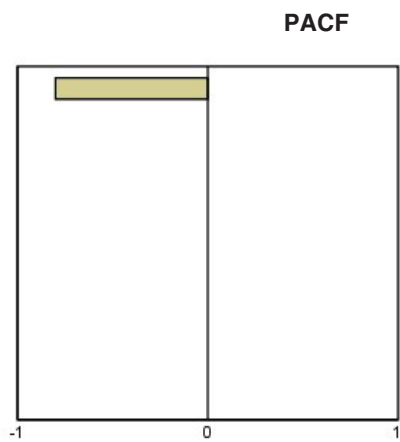
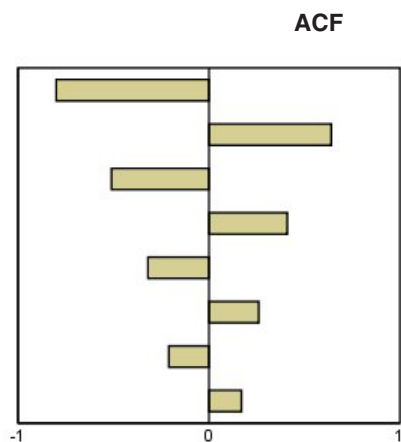
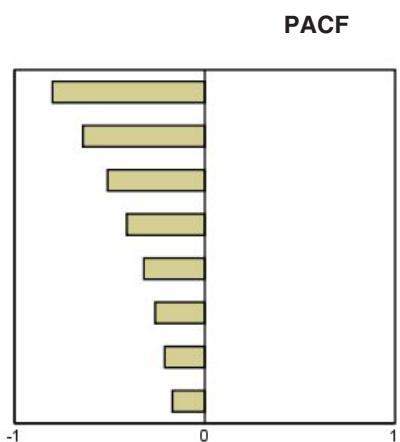
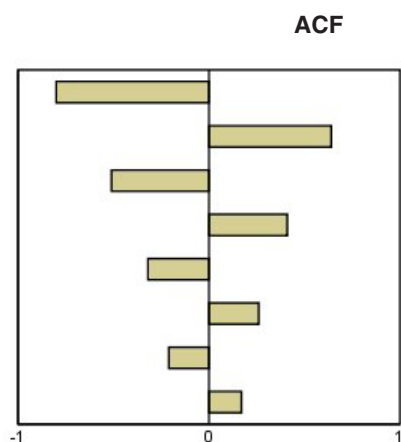


表 6. $ARIMA(1,0,0)$, $f < 0$



$ARIMA(1,0,1)$, $<0, >0$



$ARIMA(2,0,0)$, $\rho_1 \rho_2 > 0$

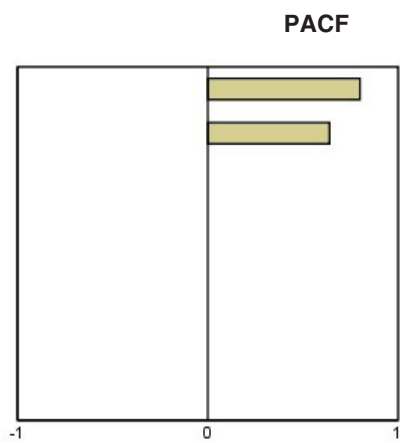
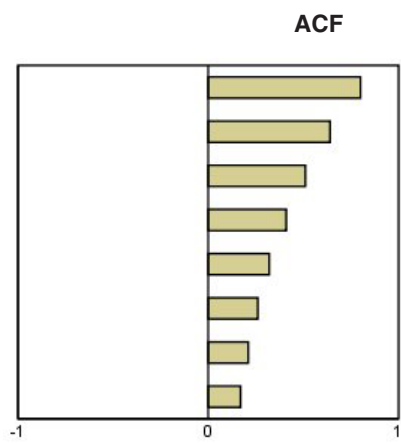
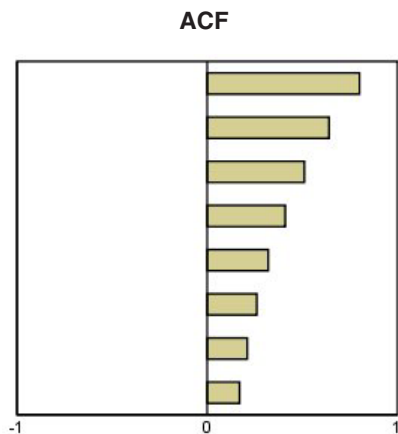


表 7. $ARIMA(0,1,0)$ (积分序列)



声明

本信息是为在美国提供的产品和服务编写的。

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
Armonk, NY 10504-1785
U.S.A.

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.
1623-14, Shimotsuruma, Yamato-shi
Kanagawa 242-8502 Japan

本条款不适用英国或任何这样的条款与当地法律不一致的国家或地区：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。有些州/省不允许针对某些交易的明示或暗示免责条款，因此本声明可能不适用于您。

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 Software Group
ATTN: Licensing

200 W. Madison St.
Chicago, IL; 60606
U.S.A.

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。

此处所含的性能数据均在受控环境下决定。因此，在其他操作环境中获得的结果可能差异较大。有些测量可能在开发级的系统中进行，不保证这些测量结果与常用系统上的测量结果相同。此外，有些测量结果可能通过推断来估计得出。实际结果可能有所差异。此文档的用户应针对其具体环境验证适用的数据。

非 IBM 产品的相关信息来自这些产品的供应商，及其发布的公告或其他公开来源。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 未来方向或意向的所有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。所有这些名称均为虚构，与真实商业企业使用的名称和地址的任何雷同纯属巧合。

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：

©（贵公司的名称）（年）。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. _（输入年份）_ . 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。当前的 IBM 商标列表，可从 Web 站点 www.ibm.com/legal/copytrade.shtml 上“版权和商标信息”部分获取。

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
- 保存
 - 模型预测 10, 16
 - 新变量名称 10, 16
 - XML 格式的重新估计的模型 16
 - XML 格式的模型规范 10
- 变量名称
 - 在“时间序列建模器”中 10
 - 在“应用时间序列模型”中 16
- 标准化 BIC (BIC 准则) 43
 - 在“时间序列建模器”中 9
 - 在“应用时间序列模型”中 14, 15

[C]

- 残差
 - 在“时间序列建模器”中 9
 - 在“应用时间序列模型”中 14, 15
- 重新估计模型参数
 - 在“应用时间序列模型”中 13
- 创新离群值 45
 - 在“时间序列建模器”中 5, 8

[D]

- 调和分析 21
- 对数转换
 - 在“时间序列建模器”中 5, 6, 7

[G]

- 估计期 2
 - 在“时间序列建模器”中 3
 - 在“应用时间序列模型”中 13

[J]

- 季节性可加离群值 45
 - 在“时间序列建模器”中 5, 8
- 简单的季节性指数平滑法模型 5
- 简单的指数平滑法模型 5
- 局部趋势离群值 45
 - 在“时间序列建模器”中 5, 8
- 均方根误差 43
 - 在“时间序列建模器”中 9
 - 在“应用时间序列模型”中 14, 15

[K]

- 可加的修补离群值 45
 - 在“时间序列建模器”中 5, 8
- 可加离群值 45
 - 在“时间序列建模器”中 5, 8

[L]

- 离群值
 - 定义 45
 - 专家建模器 5
 - ARIMA 模型 8
- 历史期 2
- 历史数据
 - 在“时间序列建模器”中 9
 - 在“应用时间序列模型”中 15

[M]

- 模型
 - 指数平滑法 3, 5
 - 专家建模器 3
 - ARIMA 3
- 模型参数
 - 在“时间序列建模器”中 9
 - 在“应用时间序列模型”中 14
- 模型名称
 - 在“时间序列建模器”中 11

[N]

- 拟合优度
 - 定义 43
 - 在“时间序列建模器”中 9
 - 在“应用时间序列模型”中 14, 15
- 拟合值
 - 在“时间序列建模器”中 9
 - 在“应用时间序列模型”中 15

[P]

- 偏自相关函数
 - 纯 ARIMA 过程图 47
 - 在“时间序列建模器”中 9
 - 在“应用时间序列模型”中 14, 15
- 平方根转换
 - 在“时间序列建模器”中 5, 6, 7
- 平均绝对误差 43
 - 在“时间序列建模器”中 9

- 平均绝对误差 (续)
 - 在“应用时间序列模型”中 14, 15
- 平均绝对误差百分比 43
 - 在“时间序列建模器”中 9
 - 在“应用时间序列模型”中 14, 15
- 平稳 R2 43
 - 在“时间序列建模器”中 9
 - 在“应用时间序列模型”中 14, 15
- 谱图 21, 22
 - 假设 21
 - 居中转换 21
 - 频谱窗口 21
 - 双变量谱分析 21

[Q]

- 缺失值
 - 在“时间序列建模器”中 11
 - 在“应用时间序列模型”中 17

[S]

- 时间序列分析
 - 时间因果模型 23
- 时间序列建模器 3
 - 保存预测值 10
 - 残差部分自相关函数 9
 - 残差自相关函数 9
 - 估计期 3
 - 跨所有模型的统计 9
 - 离群值 5, 8
 - 模型参数 9
 - 模型名称 11
 - 拟合优度统计 9
 - 拟合值 9
 - 缺失值 11
 - 事件 5
 - 新变量名称 10
 - 序列转换 5, 6, 7
 - 以 XML 格式保存模型规范 10
 - 预测期 3, 11
 - 预测值 9
 - 指数平滑法 3, 5
 - 置信区间 9, 11
 - 周期性 5, 6, 7
 - 专家建模器 3
 - 转换函数 7
 - 最佳和最差拟合模型 10
 - ARIMA 3, 6
 - Box-Ljung 统计 9

时间因果模型 23, 24, 25, 26, 27, 28, 29, 30, 31
时间因果模型方案 38, 39, 40, 41, 42
时间因果模型预测 33, 34, 35, 37
事件 5
 在“时间序列建模器”中 5
水平移位离群值 45
 在“时间序列建模器”中 5, 8
瞬时离群值 45
 在“时间序列建模器”中 5, 8

[Y]

验证期 2
应用时间序列模型 13
 保存预测值 16
 残差部分自相关函数 14, 15
 残差自相关函数 14, 15
 重新估计模型参数 13
 估计期 13
 跨所有模型的统计 14, 15
 模型参数 14
 拟合优度统计 14, 15
 拟合值 15
 缺失值 17
 新变量名称 16
 以 XML 格式保存重新估计的模型 16
 预测期 13
 预测值 14, 15
 置信区间 15, 17
 最佳和最差拟合模型 16
 Box-Ljung 统计 14
预测期
 在“时间序列建模器”中 3, 11
 在“应用时间序列模型”中 13
预测值
 在“时间序列建模器”中 9
 在“应用时间序列模型”中 14, 15

[Z]

指数平滑法模型 3, 5
置信区间
 在“时间序列建模器”中 9, 11
 在“应用时间序列模型”中 15, 17
周期性
 在“时间序列建模器”中 5, 6, 7
周期性分解 19, 20
 保存新变量 19
 创建变量 19
 计算移动平均值 19
 假设 19
 模型 19
专家建模器 3
 离群值 5

专家建模器 (续)
 限制模型空间 5
转换函数 7
 差分阶数 7
 分母的阶 7
 分子的阶 7
 季节性阶 7
 延迟 7
自然对数转换
 在“时间序列建模器”中 5, 6, 7
自相关函数
 纯 ARIMA 过程图 47
 在“时间序列建模器”中 9
 在“应用时间序列模型”中 14, 15
阻尼指数平滑法模型 5
最大绝对误差 43
 在“时间序列建模器”中 9
 在“应用时间序列模型”中 14, 15
最大绝对误差百分比 43
 在“时间序列建模器”中 9
 在“应用时间序列模型”中 14, 15

A

ACF
 纯 ARIMA 过程图 47
 在“时间序列建模器”中 9
 在“应用时间序列模型”中 14, 15
ARIMA 模型 3
 离群值 8
 转换函数 7

B

Box-Ljung 统计
 在“时间序列建模器”中 9
 在“应用时间序列模型”中 14
Brown 的指数平滑法模型 5

H

Holt 的指数平滑法模型 5

M

MAE 43
 在“时间序列建模器”中 9
 在“应用时间序列模型”中 14, 15
MAPE 43
 在“时间序列建模器”中 9
 在“应用时间序列模型”中 14, 15
MaxAE 43
 在“时间序列建模器”中 9
 在“应用时间序列模型”中 14, 15
MaxAPE 43

MaxAPE (续)
 在“时间序列建模器”中 9
 在“应用时间序列模型”中 14, 15

P

PACF
 纯 ARIMA 过程图 47
 在“时间序列建模器”中 9
 在“应用时间序列模型”中 14, 15

R

R2 43
 在“时间序列建模器”中 9
 在“应用时间序列模型”中 14, 15
RMSE 43
 在“时间序列建模器”中 9
 在“应用时间序列模型”中 14, 15

W

Winters 的指数平滑法模型
 加性 5
 可乘 5

X

XML
 以 XML 格式保存重新估计的模型 16
 以 XML 格式保存时间序列模型 10



Printed in China