

IBM SPSS Forecasting 30



참고

이 정보와 이 정보가 지원하는 제품을 사용하기 전에, 43 페이지의 『주의사항』에 있는 정보를 확인하십시오.

제품 정보

이 개정판은 새 개정판에 별도로 명시하지 않는 한 버전 30, 릴리스 0, IBM® SPSS® Statistics 수정 0 및 모든 후속 릴리스와 수정에 적용됩니다.

© Copyright International Business Machines Corporation .

목차

제 1 장 예측.....	1
시계열 소개.....	1
시계열 데이터.....	1
데이터 변환.....	2
추정 및 검증 기간.....	2
모형 작성 및 예측 생성.....	2
시계열 모델러.....	3
자동 모형 생성기 옵션 지정.....	4
사용자 정의 지수평활 모형.....	5
사용자 정의 ARIMA 모형.....	6
출력.....	8
모형 예측 및 모형 지정 사항 저장.....	10
옵션.....	10
TSMODEL 명령 추가 기능.....	11
시계열 모형 적용.....	11
결과.....	13
모형 예측 및 모형 지정 사항 저장.....	15
옵션.....	15
TSAPPLY 명령 추가 기능.....	16
계절 분해.....	16
계절 분해 저장.....	16
SEASON 명령 추가 기능.....	17
스펙트럼 도표.....	17
SPECTRA 명령 추가 기능.....	19
시간 인과 모형.....	19
시간 인과 모델 얻기.....	19
모형에 대한 시계열.....	20
관측값.....	21
분석에 대한 시간 간격.....	22
통합 및 분포.....	22
결측값.....	23
일반 데이터 옵션.....	23
일반 작성 옵션.....	23
표시할 계열.....	24
결과 옵션.....	24
추정 기간.....	26
예측값.....	26
저장.....	27
대화형 결과.....	27
시간 인과 모형 적용.....	28
시간 인과 모형 적용.....	28
임시 인과 모델 예측.....	28
시간 인과 모형 시나리오.....	33
적합도 측도.....	37
이상치 유형.....	38
ACF/PACF 도표.....	38
주의사항.....	43
상표.....	44

색인..... 45

제 1 장 예측

다음 예측 기능이 SPSS Statistics Professional Edition 또는 예측 옵션에 포함되어 있습니다.

시계열 소개

시계열은 일정한 시간 주기에서 단일 변수를 측정하여 얻은 관측값 세트입니다. 예를 들어, 재고 데이터 계열에서 관측값은 여러 달 동안의 일일 재고 수준을 나타낼 수 있습니다. 제품의 시장 점유율을 나타내는 계열은 몇 년 동안의 주별 시장 점유율로 구성될 수 있습니다. 전체 판매액 수치 계열은 여러 해 동안의 월별 관측값으로 구성될 수 있습니다. 이러한 각각의 예에서 공통적인 내용은 일부 변수가 특정 시간 길이 동안에 일정하고 알려진 구간에서 관측된다는 것입니다. 따라서 일반적인 시계열의 데이터 형식은 일정한 구간에서 측정된 측정값을 나타내는 단일 시퀀스 또는 목록의 관측값입니다.

표 1. 일일 재고 시계열

시간	주	일	재고 수준
t ₁	1	월요일	160
t ₂	1	화요일	135
t ₃	1	수요일	129
t ₄	1	목요일	122
t ₅	1	금요일	108
t ₆	2	월요일	150
		...	
t ₆₀	12	금요일	120

시계열 분석을 수행하는 가장 중요한 이유 중 하나는 계열의 미래 값 예측을 시도하는 것입니다. 이전 값을 설명하는 계열의 모형은 다음 몇몇 값이 증가 또는 감소할 것인지 여부 및 정도 또한 예측할 수 있습니다. 이러한 예측을 제대로 수행할 수 있는 기능이 모든 사업 또는 과학 분야에 중요합니다.

시계열 데이터

열 기반 데이터

각 시계열 필드에는 단일 시계열 데이터가 포함됩니다. 이 구조는 시계열 모형 생성기 프로시저, 계절분해 프로시저 및 스펙트럼 도표 프로시저에서 사용되는 경우와 같이 시계열 데이터의 기존 구조입니다. 예를 들어, 데이터 편집기에서 시계열을 정의하려면 변수 보기 탭을 클릭하고 임의의 빈 행에 변수 이름을 입력하십시오. 시계열의 각 관측값은 케이스(데이터 편집기의 행)에 해당합니다.

시계열 데이터를 포함하는 스프레드시트를 열면, 각 계열은 스프레드시트의 열에 배열되어야 합니다. 시계열이 포함된 스프레드시트를 이미 행에 배열한 경우에는 스프레드시트를 연 다음 데이터 메뉴에서 전치를 사용하여 행을 열로 이동할 수 있습니다.

다차원 데이터

다차원 데이터의 경우, 각 시계열 필드에는 여러 시계열에 대한 데이터가 포함됩니다. 특정 필드 내에서 별도의 시계열은 차원 필드라고 하는 범주형 필드의 값 세트로 식별됩니다.

예를 들어, 여러 지역 및 브랜드에 대한 판매 데이터는 해당 케이스의 차원이 *region* 및 *brand*가 되도록 단일 *sales* 필드에 저장될 수 있습니다. *region* 및 *brand*의 각 조합은 *sales*에 대한 특정 시계열을 식별합니다. 예를 들어 다음 표에서 *region* 값이 'north'고 *brand* 값이 'brandX'인 레코드는 단일 시계열을 정의합니다.

표 2. 다차원 데이터			
날짜	영역	브랜드	영업
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

참고: IBM Cognos® TM1®에서와 같이 OLAP 큐브에서 가져온 데이터는 다차원 데이터로 표시됩니다.

데이터 변환

Core System에서 제공하는 여러 데이터 변환 프로시저는 시계열 분석에 유용합니다. 이러한 변환은 각 시계열 필드에 단일 시계열 데이터가 포함된 열 기반 데이터에만 적용됩니다.

- 날짜 메뉴의 날짜 정의 프로시저는 주기성 설정 및 과거 주기, 검증 주기, 시계열 분석 기간 간 구별에 사용되는 날짜변수를 생성합니다. 예측은 날짜 정의 프로시저에서 생성된 변수를 사용하는 작업을 위해 고안되었습니다.
- 변환 메뉴의 시계열 변수 생성 프로시저는 기존 시계열 변수의 함수로 새로운 시계열 변수를 생성합니다. 평활, 평균 및 차이에 대해 가까운 관측값을 사용하는 함수가 포함되어 있습니다.
- 변환 메뉴의 결측값 대체 프로시저는 시스템 및 사용자 결측값을 여러 방법 중 하나에 따른 추정값으로 대체합니다. 계열 값의 시작 부분이나 종료 부분에 결측 데이터가 있는 경우 특별한 문제가 발생하지는 않으며 단지 계열 값의 가용 길이가 줄어든 뿐입니다. 그러나 계열 값 중간의 갭(삽입된 결측 데이터)은 훨씬 더 심각한 문제를 일으킬 수 있습니다.

시계열 데이터 변환에 대한 자세한 내용은 *Core System* 사용자 안내서를 참조하십시오.

추정 및 검증 기간

시계열을 추정 또는 과거 주기 및 검증 주기로 나누는 것이 유용한 경우가 많습니다. 추정(과거)주기의 관측값을 기초로 모형을 개발한 다음 검증하여 검증 주기에서 적절한 정도를 알아봅니다. 이미 알고 있는 점(검증 주기에 있는 점)을 모형에서 강제로 예측하게 하여 모형이 얼마나 잘 예측하는지 알아봅니다.

검증 주기에 있는 케이스는 모형 작성 프로세스에서 제외되므로 일반적으로 검증 케이스로 참조됩니다. 모형의 시계열 분석 작업이 충분하다고 만족하는 경우 추정 기간을 다시 정의하여 검증 케이스를 포함시킨 다음 최종 모형을 설정할 수 있습니다.

모형 작성 및 예측 생성

Forecasting 추가 기능 모듈에서 제공하는 다음 두 프로시저를 사용하여 모형을 작성하고 예측을 생성하는 작업을 수행할 수 있습니다.

- 3 페이지의 『시계열 모델러』 프로시저는 시계열 모형을 작성하고 예측을 생성합니다. 자동으로 각 시계열에 대해 최적의 모형을 결정하는 자동 모형 생성기가 포함되어 있습니다. 더 높은 수준의 제어를 필요로 하는 숙련된 분석가에게는 사용자 정의 모형 작성 도구도 제공합니다.
- 11 페이지의 『시계열 모형 적용』 프로시저는 시계열 모델러에서 작성한 기존 시계열 모형을 활성 데이터 세트에 적용합니다. 새로운 또는 변경된 데이터를 사용할 수 있는 계열에 대해 모형을 다시 설정하지 않고 예측을 생성할 수 있습니다. 모형을 바꿔야 한다고 생각할 이유가 있는 경우에는 시계열 모델러를 사용하여 모형을 다시 설정할 수 있습니다.

- 19 페이지의 『시간 인과 모형』 프로시저에서는 각 목표에 대한 자동 선형회귀 시계열 모형을 작성하고 목표와 인과 관계가 있는 최적의 입력을 자동으로 결정합니다. 이 프로시저는 인과 관계를 탐색하는 데 사용할 수 있는 대화식 출력을 생성합니다. 또한 예측값을 생성하고 이상값을 검출하며 이상값을 초래할 가능성이 가장 높은 계열을 판별할 수도 있습니다.
- 28 페이지의 『임시 인과 모델 예측』 프로시저는 활성 데이터 세트에 시간 인과 모형을 적용합니다. 이 프로시저를 사용하여 모형을 다시 작성하지 않고도 최신 데이터를 사용할 수 있는 계열의 예측값을 얻을 수 있습니다. 또한 시간 인과 모형 프로시저에 의해 검출된 이상값을 초래할 가능성이 가장 높은 계열을 판별할 수도 있습니다.

시계열 모델러

시계열 모델러 프로시저는 시계열에 대해 지수평활, 일변량 자동 선형회귀 집적 이동 평균(ARIMA) 및 다변량 ARIMA(또는 전이 함수 모형) 모형을 추정하고 예측을 생성합니다. 프로시저에는 자동으로 하나 이상의 종속변수 계열에 대해 최적 적합 ARIMA 또는 지수평활 모형을 식별하고 추정하는 자동 모형 생성기가 포함되어 있으므로 적절한 모형을 식별하기 위해 시행착오를 거칠 필요가 없습니다. 사용자 정의 ARIMA 또는 지수평활 모형을 지정할 수도 있습니다.

예. 제품 관리자로서 100가지 제품 각각에 대해 다음 달의 판매건수 및 매출을 예측해야 하는 책임이 있으나 시계열을 모델링해 본 경험이 거의 또는 전혀 없습니다. 100가지 제품 전체의 과거 판매건수 데이터는 단일 Excel 스프레드시트에 저장되어 있습니다. IBM SPSS Statistics에서 스프레드시트를 연 다음 자동 모형 생성기를 사용하여 다음 한 달에 대한 예측을 요청합니다. 자동 모형 생성기에서 제품 각각의 판매건수에 대한 최적의 모형을 찾아 해당 모형을 사용하여 예측을 생성합니다. 전문가 모형기는 다중 입력 계열을 처리할 수 있으므로 프로시저를 한 번만 실행하면 제품 전체에 대한 예측을 구할 수 있습니다. 예측을 활성 데이터 세트에 저장하도록 선택하면 쉽게 결과를 Excel로 다시 내보낼 수 있습니다.

통계. 적합도 측도: 정상 R -제곱, R -제곱(R^2), 평균 제곱 오차 제곱근(RMSE), 평균 절대 오차(MAE), 평균 절대 퍼센트 오차(MAPE), 절대 오차의 최대값(MaxAE), 절대 퍼센트 오차의 최대값(MaxAPE), 정규화된 베이지안 정보 기준(BIC)입니다. 잔차: 자기상관 함수, 편자기상관 함수, Ljung-Box Q 입니다. ARIMA 모형에 대해서는 종속변수의 ARIMA 차수, 독립변수의 전이 함수 차수 및 이상값 추정값입니다. 또한 지수평활 모형의 평활 모수 추정값입니다.

도표. 모든 모형 전체의 요약 도표: 정상 R -제곱의 히스토그램, R -제곱(R^2), 평균 제곱 오차 제곱근(RMSE), 평균 절대 오차(MAE), 평균 절대 퍼센트 오차(MAPE), 절대 오차의 최대값(MaxAE), 절대 퍼센트 오차의 최대값(MaxAPE), 정규화된 베이지안 정보 기준(BIC), 잔차 자기상관 및 편자기상관 상자도표입니다. 개별 모형의 결과: 예측 값, 적합값, 관측값, 신뢰 한계 상한 및 하한, 잔차 자기상관 및 편자기상관입니다.

시계열 모델러 데이터 고려 사항

데이터. 종속변수 및 독립변수는 숫자이어야 합니다.

가정. 종속변수 및 독립변수는 시계열로 처리됩니다. 즉, 각 케이스는 시점을 나타내고 연속 케이스는 일정한 시간 간격으로 구분됩니다.

- **정상성** 사용자 정의 ARIMA 모형의 경우 모델링되는 시계열은 정상이어야 합니다. 비정상 계열을 정상 계열로 변환하는 가장 효율적인 방법은 차분 변환을 통한 방법입니다. 차분 변환은 시계열 작성 대화 상자에서 사용 가능합니다.
- **예측** 독립변수(예측변수)가 포함된 모형을 사용하여 예측을 생성하는 경우 활성 데이터 세트에는 예측 기간의 모든 케이스에 대해 이러한 변수의 값이 있어야 합니다. 또한 독립변수는 추정 기간에 결측값이 없어야 합니다.

날짜 정의

필수는 아니지만 날짜 정의 대화 상자를 사용하여 첫 번째 케이스와 관련된 날짜 및 연속 케이스 간 시간 간격을 지정하는 것이 좋습니다. 시계열 모델러를 사용하기 전에 수행되며 각 케이스와 관련된 날짜를 설명하는 변수 세트가 생성됩니다. 또한 가정된 데이터 주기성을 설정합니다. 예를 들어, 연속 케이스 간 시간 간격이 한 달인 경우 주기성은 12입니다. 계절 모형을 생성하려는 경우 이러한 주기성은 필수입니다. 계절 모형을 생성하지 않으며 결과에 날짜 레이블이 필요하지 않은 경우 날짜 정의 대화 상자를 건너뛸 수 있습니다. 각 케이스와 관련된 레이블은 단순히 케이스 번호입니다.

시계열 모델러 사용 방법

1. 메뉴에서 다음을 선택합니다.

분석 > 예측 > 기존 모형 적용...

2. 변수 탭에서 모델링할 종속변수를 하나 이상 선택합니다.
3. 방법 드롭다운 상자에서 모델링 방법을 선택합니다. 자동 모델링의 경우 기본값 방법인 **전문가 모형기**를 선택합니다. 전문가 모형기를 호출하여 각 종속변수에 대해 최적 적합 모형을 결정합니다.

예측 생성:

4. 옵션 탭을 클릭합니다.
5. 예측 기간을 지정합니다. 예측 및 관측값이 포함된 도표를 생성합니다.

선택적으로 다음을 수행할 수 있습니다.

- 하나 이상의 독립변수를 선택합니다. 독립변수는 회귀분석의 예측변수처럼 처리되지만 선택적입니다. ARIMA 모형에는 포함될 수 있지만 지수평활 모형에는 포함될 수 없습니다. 모델링 방법으로 **자동 모형 생성기**를 지정하고 독립변수를 포함시키는 경우 ARIMA 모형만 간주합니다.
- **기준**을 클릭하여 모델링 세부사항을 지정합니다.
- **예측, 신뢰구간 및 잡음 잔차를 저장**합니다.
- **추정된 모형을 XML 형식으로 저장**합니다. 저장된 모형을 새로운 또는 변경된 데이터에 적용하여 모형을 다시 작성하지 않고 업데이트된 예측을 구할 수 있습니다. 프로시저를 사용하여 수행합니다.
- **모든 추정된 모형에 대한 요약 통계를 구**합니다.
- **사용자 정의 ARIMA 모형에서 독립변수의 전이 함수**를 지정합니다.
- **이상값 자동 발견**을 설정합니다.
- **사용자 정의 ARIMA 모형의 이상값으로 특정 시점**을 모델링합니다.

방법 모델링

사용 가능한 모델링 방법은 다음과 같습니다.

자동 모델 생성기. 전문가 모형기에서 자동으로 각 종속 계열에 대해 최적 적합 모형을 찾아냅니다. 독립변수(예측변수)가 지정되어 있는 경우 전문가 모형기에서 종속 계열과 통계적으로 유의 관계가 있는 변수만 ARIMA 모형에 포함되도록 선택합니다. 모델 변수는 차이 및/또는 제곱근이나 자연 로그 변환을 사용하여 적절한 위치에서 변환됩니다. 기본적으로 전문가 모형기는 지수평활 및 ARIMA 모형을 모두 고려합니다. 그러나 전문가 모형기에서 ARIMA 모형만 검색하거나 지수평활 모형만 검색하도록 제한할 수 있습니다. 또한 이상값 자동 발견을 지정할 수 있습니다.

지수평활. 이 옵션을 사용하여 사용자 정의 지수평활 모형을 지정합니다. 추세 및 계절성을 다르게 처리하는 다양한 지수평활 모델에서 선택할 수 있습니다.

ARIMA. 이 옵션을 사용하여 사용자 정의 ARIMA 모델을 지정합니다. 자동 선행회귀 및 이동 평균 차수와 차이의 수를 명시적으로 지정하는 작업이 포함됩니다. 독립변수(예측변수)를 포함하고 임의의 또는 모든 독립변수의 전이 함수를 정의할 수 있습니다. 또한 이상치 자동 발견을 지정하거나 이상치 세트를 명시적으로 지정할 수 있습니다.

추정 및 시계열 분석 기간

추정 기간 추정 기간은 모형 결정에 사용되는 케이스 세트를 정의합니다. 기본적으로 추정 기간은 활성 데이터 세트의 모든 케이스를 포함합니다. 추정 기간을 설정하려면 케이스 선택 대화 상자에서 **시간 또는 케이스 범위 기본**을 선택하십시오. 사용 가능한 데이터에 따라 프로시저에서 사용되는 추정 기간이 종속변수별로 다를 수 있으며 따라서 표시된 값과 다를 수 있습니다. 종속변수가 지정되어 있는 경우 실제 추정 기간은 지정된 추정 기간의 시작 또는 끝에서 발생하는 변수의 연속 결측값을 제거한 후 남은 주기입니다.

예측 기간 예측 기간은 추정 기간 다음의 첫 번째 케이스에서 시작하며 기본적으로 활성 데이터 세트의 마지막 케이스까지 계속됩니다. 옵션 탭에서 예측 기간의 종료를 설정할 수 있습니다.

자동 모형 생성기 옵션 지정

전문가 모형기 옵션을 사용하여 후보 모형 세트를 제한하고 이상치 처리를 지정하며 이벤트 변수를 지정할 수 있습니다.

모형 선택 및 이벤트 지정 사항

모형 탭을 사용하여 전문가 모형기에서 고려할 모형 유형을 지정하고 이벤트 변수를 지정할 수 있습니다.

모델 유형. 다음 옵션을 사용할 수 있습니다.

- **모든 모형.** 자동 모형 생성기에서 ARIMA 및 지수평활 모형을 모두 고려합니다.
- **지수평활 모델만.** 전문가 모형기에서 지수평활 모형만 고려합니다.
- **ARIMA 모형만** 자동 모형 생성기에서 ARIMA 모형만 고려합니다.

전문가 모형기에서 계절 모형 고려 이 옵션은 활성 데이터 세트에 대해 주기성이 정의된 경우에만 사용할 수 있습니다. 이 옵션이 선택되어 있는 경우 전문가 모형기에서 계절 및 비계절 모형을 모두 고려합니다. 이 옵션이 선택되어 있지 않은 경우 자동 모델 생성기는 비계절 모델만 고려합니다.

현재 주기성. 활성 데이터 세트에 대해 현재 정의되어 있는 주기성이 있는 경우 해당 주기성을 나타냅니다. 현재 주기성은 정수로 지정됩니다. 예를 들어, 연간 주기성 12는 각 케이스가 하나의 월을 나타냅니다. 주기성이 설정되어 있지 않은 경우 값 지정없음이 표시됩니다. 계절 모형에는 주기성이 필요합니다. 날짜 정의 대화 상자에서 주기성을 설정할 수 있습니다.

이벤트 이벤트 변수로 처리될 독립변수를 선택합니다. 이벤트 변수의 경우 값이 1인 케이스는 종속 계열이 이벤트의 영향을 받을 것으로 예상되는 시기를 나타냅니다. 1 이외의 값은 영향이 없다는 것을 나타냅니다.

전문가 모형기로 이상값 처리

이상값 탭을 사용하여 이상값 자동 발견 및 발견할 이상값 유형을 선택할 수 있습니다.

자동으로 이상치 검색. 기본적으로 이상값 자동 발견을 수행하지 않습니다. 이상값 자동 발견을 수행하려면 이 옵션을 선택한 다음 아래 이상값 유형 중 한 개 이상을 선택합니다.

- 가법
- 수준 이동
- 혁신적
- 일시적
- 계절 가법
- 로컬 추세
- 가법 패치

사용자 정의 지수평활 모형

모델 유형. 지수평활 모형¹계절 또는 비계절로 분류됩니다. 계절 모형은 활성 데이터 세트에 대해 주기성이 정의된 경우에만 사용할 수 있습니다(아래 "현재 주기성" 참조).

- **simple.** 이 모델은 추세 또는 계절성이 없는 계열에 적합합니다. 고유한 평활 모수는 수준입니다. 단순 지수평활은 자동 선형회귀 차수가 0, 차이 차수가 1, 이동 평균 차수가 1, 상수가 없는 ARIMA 모형과 가장 비슷합니다.
- **Holt 선형 추세(Holt's linear trend).** 이 모델은 선형 추세가 있고 계절성이 없는 계열에 적합합니다. 평활 모수는 서로 다른 값에 의해 제약을 받지 않는 수준과 추세입니다. Holt 모형은 Brown 모형보다 더 일반적이지만 대형 계열 계산에는 더 오래 걸립니다. Holt 지수평활은 자동 선형회귀 차수가 0, 차분계산 차수가 2, 이동 평균 차수가 2인 ARIMA 모형과 가장 비슷합니다.
- **Brown 선형 추세(Brown's linear trend).** 이 모델은 선형 추세가 있고 계절성이 없는 계열에 적합합니다. 평활 모수는 동일한 것으로 간주되는 수준과 추세입니다. 따라서 Brown의 모형은 Holt 모형의 특별한 케이스입니다. Brown의 지수평활은 자동 선형회귀 차수가 0, 차분계산 차수가 2, 이동 평균 차수가 2이며, 이동 평균의 두 번째 차수에 대한 계수가 첫 번째 차수에 대한 계수의 절반의 제곱과 같은 ARIMA 모형과 가장 비슷합니다.
- **진폭감소 추세(Damped trend).** 이 모델은 선형 추세가 소멸되고 계절성이 없는 계열에 적합합니다. 평활 모수는 수준, 추세 및 진폭감소 추세입니다. 진폭감소 지수평활은 자동 선형회귀 차수가 1, 차분계산 차수가 1, 이동 평균 차수가 2인 ARIMA 모형과 가장 비슷합니다.
- **단순 계절모형(Simple seasonal).** 이 모델은 추세가 없고 계절 효과가 시간에 따라 일정한 계열에 적합합니다. 평활 모수는 수준과 계절입니다. 단순 계절모형 지수평활은 자동 선형회귀 차수가 0, 차분계산 차수가 1, 계절

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

차분 차수가 1, 이동 평균 차수가 1, p 및 $p+1$ 인 ARIMA 모형과 가장 비슷합니다. 여기서 p 는 계절 구간에서의 주기 수(예를 들어 월별 데이터인 경우 p 는 12)입니다.

- **Winters 가법(Winters' additive)**. 이 모델은 선형 추세 및 계절 효과가 계절 수준에 종속되지 않은 계절에 적합합니다. 평활 모수는 수준, 추세 및 계절입니다. Winters의 가법 지수평활은 자동 선형회귀 차수가 0, 차분계산 차수가 1, 계절 차분 차수가 1, 이동 평균 차수가 $p+1$ 인 ARIMA 모형과 가장 비슷합니다. 여기서 p 는 계절 구간에서의 주기 수(예를 들어 월별 데이터인 경우 p 는 12)입니다.
- **Winters 승법(Winters' multiplicative)**. 이 모델은 선형 추세 및 계절 수준에 따른 계절 효과가 있는 계절에 적합합니다. 평활 모수는 수준, 추세 및 계절입니다. Winters의 승법 지수평활은 ARIMA 모형과 다릅니다.

현재 주기성. 활성 데이터 세트에 대해 현재 정의되어 있는 주기성이 있는 경우 해당 주기성을 나타냅니다. 현재 주기성은 정수로 지정됩니다. 예를 들어, 연간 주기성 12는 각 케이스가 하나의 월을 나타냅니다. 주기성이 설정되어 있지 않은 경우 값 지정없음이 표시됩니다. 계절 모형에는 주기성이 필요합니다. 날짜 정의 대화 상자에서 주기성을 설정할 수 있습니다.

종속 변수 변환 모델링 전에 각 종속변수에 대해 수행되는 변환을 지정할 수 있습니다.

- **지정없음.** 변환을 수행하지 않습니다.
- **제곱근.** 제곱근 변환을 수행합니다.
- **자연로그.** 자연로그 변환.

사용자 정의 ARIMA 모형

시계열 모델러를 사용하면 고정된 예측변수 세트가 있거나 없는 사용자 정의 비계절 또는 Box-Jenkins 모형이라고도 하는 계절 ARIMA(통합 자기회귀 이동 평균) 모형을²빌드할 수 있습니다. 임의의 또는 모든 예측변수의 전이 함수를 정의하고 이상치 자동 발견을 지정하거나 이상치 세트를 명시적으로 지정할 수 있습니다.

- 변수 탭에서 지정된 모든 독립변수(예측변수)는 명시적으로 모형에 포함됩니다. 이는 종속변수와 통계적으로 유의 관계가 있는 경우에만 독립변수가 포함되는 자동 모형 생성기 사용 시와 대비됩니다.

사용자 정의 ARIMA 모형의 모형 지정 사항

모델 탭을 사용하여 사용자 정의 ARIMA 모형의 구조를 지정할 수 있습니다.

ARIMA 차수. 구조 눈금의 해당 셀에 모형의 다양한 ARIMA 구성요소 값을 입력합니다. 모든 값은 0 또는 양의 정수이어야 합니다. 자기회귀 및 이동 평균 구성요소의 경우 값은 최대 차수를 나타냅니다. 더 낮은 양의 차수가 모두 모델에 포함됩니다. 예를 들어, 2를 지정하면 모델에는 2와 1 차수가 포함됩니다. 계절 열의 셀은 활성 데이터 세트에 대해 주기성이 정의된 경우에만 사용할 수 있습니다(아래의 "현재 주기성" 참조).

- **자동 선형회귀(p)** 모형의 자기회귀 차수 수입니다. 자기회귀 차수는 계절의 이전 값 중 현재 값 예측에 사용될 값을 지정합니다. 예를 들어, 자동 선형회귀 차수 2는 과거 2개 시간 주기의 계절 값을 현재 값 예측에 사용하도록 지정합니다.
- **차이(d).** 모형을 추정하기 전 계절에 적용할 차이 차수를 지정합니다. 추세가 존재하며(추세가 있는 계절은 일반적으로 비정상이며 ARIMA 모델링은 정상성을 가정) 해당 효과 제거를 위해 사용되는 경우 차이가 필요합니다. 차이 차수는 계절 추세 수준에 해당합니다. 1차 차이는 선형 추세, 2차 차이는 2차 추세 등을 나타냅니다.
- **이동 평균(q).** 모형의 이동 평균 차수 수입니다. 이동 평균 차수는 이전 값에 대한 계절 평균 편차를 사용하여 현재 값을 예측하는 방법을 지정합니다. 예를 들어, 이동 평균 차수 1과 2는 계절의 현재 값을 예측하는 경우 지난 2개 시간 주기 각각의 계절 평균값 편차를 고려하도록 지정합니다.

계절 차수. 계절 자기회귀, 이동 평균, 차이 구성요소는 해당 비계절 구성요소와 동일한 역할을 합니다. 그러나 계절 차수의 경우 현재 계절 값이 한 개 이상의 계절 주기에 의해 구분된 이전 계절 값의 영향을 받습니다. 예를 들어, 월별 데이터(계절 주기 12)의 경우 계절 차수 1은 현재 계절 값이 현재 계절 이전의 계절 값 12 주기의 영향을 받는다는 것을 의미합니다. 월별 데이터의 경우 계절 차수 1은 비계절 차수 12를 지정하는 것과 동일합니다.

현재 주기성. 활성 데이터 세트에 대해 현재 정의되어 있는 주기성이 있는 경우 해당 주기성을 나타냅니다. 현재 주기성은 정수로 지정됩니다. 예를 들어, 연간 주기성 12는 각 케이스가 하나의 월을 나타냅니다. 주기성이 설정되어 있지 않은 경우 값 지정없음이 표시됩니다. 계절 모형에는 주기성이 필요합니다. 날짜 정의 대화 상자에서 주기성을 설정할 수 있습니다.

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.

종속 변수 변환 모델링 전에 각 종속변수에 대해 수행되는 변환을 지정할 수 있습니다.

- **지정않음.** 변환을 수행하지 않습니다.
- **제공근.** 제공근 변환을 수행합니다.
- **자연로그.** 자연로그 변환.

모형에 상수항 포함. 전체 평균 계열 값이 0인지 확실하지 않은 경우 상수의 포함은 표준입니다. 차분이 적용되는 경우 상수를 제외하는 것이 좋습니다.

사용자 정의 ARIMA 모형의 전이 함수

전이 함수 탭(독립변수가 지정된 경우에만 있음)을 사용하여 변수 탭에서 지정된 임의의 또는 모든 독립변수에 대해 전이 함수를 정의할 수 있습니다. 전이 함수를 사용하여 독립 계열의 미래 값 예측을 위해 독립변수(예측 변수)의 이전 값 사용하는 방법을 지정할 수 있습니다.

전이 함수 차수 구조 눈금의 해당 셀에 전이 함수의 다양한 구성요소 값을 입력합니다. 모든 값은 0 또는 양의 정수 이어야 합니다. 분자 및 분모 구성요소의 경우 값은 최대 차수를 나타냅니다. 더 낮은 양의 차수가 모두 모델에 포함됩니다. 또한 분자 구성요소에 대해 차수 0은 항상 포함됩니다. 예를 들어, 분자에 2를 지정하는 경우 모델에 차수 2, 1 및 0이 포함됩니다. 분모에 3을 지정하면 모델에는 차수 3, 2 및 1이 포함됩니다. 계절 열의 셀은 활성 데이터 세트에 대해 주기성이 정의된 경우에만 사용할 수 있습니다(아래의 "현재 주기성" 참조).

- **분자.** 전이 함수의 분자 차수입니다. 선택된 독립(예측변수) 계열의 이전 값 중 종속 계열의 현재 값 예측에 사용할 값을 지정합니다. 예를 들어, 분자 차수 1은 각 종속 계열의 현재 값 예측에 과거 1개 시간 주기의 독립 계열 값과 독립 계열의 현재 값을 사용하도록 지정합니다.
- **분모.** 전이 함수의 분모 차수입니다. 선택된 독립(예측변수) 계열의 이전 값에 대해 계열 평균의 편차를 사용하여 종속 계열의 현재 값을 예측하는 방법을 지정합니다. 예를 들어, 분모 차수 1은 각 종속 계열의 현재 값을 예측하는 경우 과거 1개 시간 주기의 독립 계열 평균 값 편차를 고려하도록 지정합니다.
- **차이.** 모형을 추정하기 전 선택된 독립(예측변수) 계열에 적용할 차이 차수를 지정합니다. 추세가 있으며 해당 효과 제거를 위해 사용되는 경우 차이가 필요합니다.

계절 차수. 계절 분자, 분모 및 차이 구성요소는 해당 비계절 구성요소와 동일한 역할을 합니다. 그러나 계절 차수의 경우 현재 계열 값이 한 개 이상의 계절 주기에 의해 구분된 이전 계열 값의 영향을 받습니다. 예를 들어, 월별 데이터(계절 주기 12)의 경우 계절 차수 1은 현재 계열 값이 현재 계열 이전의 계열 값 12 주기의 영향을 받는다는 것을 의미합니다. 월별 데이터의 경우 계절 차수 1은 비계절 차수 12를 지정하는 것과 동일합니다.

현재 주기성. 활성 데이터 세트에 대해 현재 정의되어 있는 주기성이 있는 경우 해당 주기성을 나타냅니다. 현재 주기성은 정수로 지정됩니다. 예를 들어, 연간 주기성 12는 각 케이스가 하나의 월을 나타냅니다. 주기성이 설정되어 있지 않은 경우 값 지정않음이 표시됩니다. 계절 모형에는 주기성이 필요합니다. 날짜 정의 대화 상자에서 주기성을 설정할 수 있습니다.

지연. 지연을 설정하면 지정된 구간 수 만큼 독립변수의 영향력이 지연됩니다. 예를 들어, 지연이 5로 설정된 경우 시간 t 의 독립변수 값은 5개 주기가 경과할 때까지($t + 5$) 예측에 영향을 미치지 않습니다.

변환 독립변수 세트의 경우 전이 함수 지정 사항에는 해당 변수에 대해 수행될 선택적 변환도 포함됩니다.

- **지정않음.** 변환을 수행하지 않습니다.
- **제공근.** 제공근 변환을 수행합니다.
- **자연로그.** 자연로그 변환.

사용자 정의 ARIMA 모형의 이상치

이상값 탭에서는 이상값 처리에 대해 다음과 같은 선택사항을 제공합니다.³: 자동으로 검색하거나, 이상값으로 특정 지점을 지정하거나, 이를 발견하거나 모델링하지 않습니다.

이상값을 발견 또는 모델링하지 않음 기본적으로 이상값을 발견하거나 모델링하지 않습니다. 이상값 발견 또는 모델링을 사용하지 않으려면 이 옵션을 선택합니다.

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

자동으로 이상치 검색. 이상값 자동 발견을 수행하려면 이 옵션을 선택한 다음 아래 이상값 유형 중 한 개 이상을 선택합니다.

- 가법
- 수준 이동
- 혁신적
- 일시적
- 계절 가법
- 로컬 추세
- 가법 패치

특정 시점을 이상값으로 모델링 특정 시점을 이상치로 지정하려면 이 옵션을 선택합니다. 각 이상값에 대해 별도의 이상값 정의의 눈금 행을 사용합니다. 지정된 행의 모든 셀에 값을 입력합니다.

- **유형.** 이상값 유형입니다. 지원되는 유형은 가법(기본값), 수준 이동, 혁신, 일시적, 계절 가법 및 로컬 추세입니다.

참고 1: 활성 데이터 세트에 대해 날짜 지정 사항이 정의되어 있지 않은 경우 이상치 정의의 눈금에 단일 열 관측값이 표시됩니다. 이상값을 지정하려면 관련 케이스의 행 번호를 데이터 편집기에 표시된 대로 입력합니다.

참고 2: 이상치 정의의 눈금에 순환 열이 있는 경우 활성 데이터 세트의 *CYCLE_* 변수 값을 참조합니다.

출력

사용 가능한 결과로는 개별 모형의 결과 및 모든 모형에 대해 계산된 결과가 있습니다. 개별 모형의 결과를 사용자가 지정한 기준에 따라 가장 적합한 모형 또는 가장 적합하지 않은 모형의 세트로 제한할 수 있습니다.

통계 및 예측표

통계 탭은 모델링 결과 테이블 표시 옵션을 제공합니다.

모형별 적합도, Ljung-Box 통계 및 이상값 수 표시. 각 추정 모형에 대해 선택된 적합도, Ljung-Box 값 및 이상치 수가 포함된 테이블을 표시하려면 이 옵션을 선택합니다.

적합도. 추정된 각 모형의 적합도가 포함된 표에 다음 중 하나 이상을 선택하여 포함시킬 수 있습니다.

- 정상 *R*-제공
- *R*-제공
- 평균 제공근 오차
- 평균 절대 백분율 오차
- 평균 절대 오차
- 최대 절대 백분율 오차
- 최대 절대 오차
- 정규화된 BIC

모형 비교 통계 이 옵션 그룹은 추정된 모형 모두에 대해 계산된 통계를 포함하고 있는 테이블의 표시를 제어합니다. 각 옵션은 별도의 표를 생성합니다. 다음 옵션 중 한 개 이상을 선택할 수 있습니다.

- **적합도.** 고정 *R*-제공, *R*-제공, 루트 평균 제공 오차, 평균 절대 백분율 오차, 평균 절대 오차, 최대 절대 백분율 오차, 최대 절대 오차 및 정규화된 베이지안 정보 기준에 대한 요약 통계와 백분위수 표입니다.
- **잔차 자기상관 함수(ACF)** 추정된 모형 모두의 잔차 자기상관에 대한 요약 통계 및 백분위수 표입니다.
- **잔차 편자기상관 함수(PACF)** 추정된 모형 모두의 잔차 편자기상관에 대한 요약 통계 및 백분위수 표입니다.

개별 모형 통계 이 옵션 그룹은 각 추정된 모형의 상세한 정보를 포함하고 있는 표의 표시를 제어합니다. 각 옵션은 별도의 표를 생성합니다. 다음 옵션 중 한 개 이상을 선택할 수 있습니다.

- **모수 추정값.** 각 추정된 모형의 모수 추정값 표를 표시합니다. 지수평활 및 ARIMA 모형에 대해 별도의 테이블이 표시됩니다. 이상값이 존재하는 경우 해당 모수 추정값도 별도의 테이블에 표시됩니다.

- **잔차 자기상관 함수(ACF)** 각 추정된 모형의 시차별 잔차 자기상관 표를 표시합니다. 표에는 자기상관의 신뢰구간이 포함됩니다.
- **잔차 편자기상관 함수(PACF)** 각 추정된 모형의 시차별 잔차 편자기상관 표를 표시합니다. 테이블에는 편자기상관의 신뢰구간이 포함됩니다.

예측 표시 각 추정된 모형의 모형 예측 및 신뢰구간 테이블을 표시합니다. 예측 기간은 옵션 탭에서 설정됩니다.

플롯

도표 탭은 모델링 결과 도표 표시 옵션을 제공합니다.

모형 비교 도표

이 옵션 그룹은 추정된 모형 모두에 대해 계산된 통계를 포함하고 있는 도표의 표시를 제어합니다. 각 옵션은 별도의 도표를 생성합니다. 다음 옵션 중 한 개 이상을 선택할 수 있습니다.

- 정상 R -제곱
- R -제곱
- 평균 제곱근 오차
- 평균 절대 백분율 오차
- 평균 절대 오차
- 최대 절대 백분율 오차
- 최대 절대 오차
- 정규화된 BIC
- 잔차 자기상관 함수
- 잔차 편자기상관 함수(PACF)

개별 모형 도표

계열 각 추정된 모형의 예측값 도표를 구하려면 이 옵션을 선택합니다. 다음 중 한 개 이상을 선택하여 도표에 포함시킬 수 있습니다.

- 관측값 증속 계열의 관측값입니다.
- 예측 예측 기간에 대한 모형 예측값입니다.
- 적합값 추정 기간에 대한 모형 예측값입니다.
- 예측 신뢰구간 예측 기간에 대한 신뢰구간입니다.
- 적합값 신뢰구간 추정 기간에 대한 신뢰구간입니다.

잔차 자기상관 함수(ACF) 각 추정된 모형의 잔차 자기상관 도표를 표시합니다.

잔차 편자기상관 함수(PACF) 각 추정된 모형의 잔차 편자기상관 도표를 표시합니다.

가장 적합한 모형 또는 가장 적합하지 않은 모형으로 결과 제한

결과 필터 탭은 표 및 차트 결과 모두를 추정된 모형의 서브세트로 제한하는 옵션을 제공합니다. 입력한 적합 기준에 따라 결과를 최적 적합 모형 또는 가장 적합하지 않은 모형으로 제한하도록 선택할 수 있습니다. 기본적으로 모든 추정된 모형이 결과에 포함됩니다.

최적 적합 모형 최적 적합 모형을 결과에 포함시키려면 이 옵션을 선택합니다. 적합도 측도를 선택하고 포함시킬 모형의 수를 지정합니다. 이 옵션을 선택하는 경우에도 가장 적합하지 않은 모형을 선택할 수 있습니다. 이러한 케이스에서는 가장 적합하지 않은 모형 및 최적 적합 모형으로 결과가 구성됩니다.

- **고정된 모형의 수** n 개의 최적 적합 모형에 대해 결과를 표시하도록 지정합니다. 해당 숫자가 추정된 모형의 수를 초과하는 경우 모든 모형이 표시됩니다.
- **전체 모형 수의 퍼센트** 추정 모형 전체에서 적합도 값이 상위 n 퍼센트인 모형에 대해 결과를 표시하도록 지정합니다.

가장 적합하지 않은 모형 가장 적합하지 않은 모형을 결과에 포함시키려면 이 옵션을 선택합니다. 적합도 측도를 선택하고 포함시킬 모형의 수를 지정합니다. 이 옵션을 선택하는 경우에도 최적 적합 모형을 선택할 수 있습니다. 이러한 케이스에서는 최적 적합 모형 및 가장 적합하지 않은 모형으로 결과가 구성됩니다.

- **고정된 모형의 수** n 개의 가장 적합하지 않은 모형에 대해 결과를 표시하도록 지정합니다. 해당 숫자가 추정된 모형의 수를 초과하는 경우 모든 모형이 표시됩니다.
- **전체 모형 수의 퍼센트** 추정 모형 전체에서 적합도 값이 하위 n 퍼센트인 모형에 대해 결과를 표시하도록 지정합니다.

적합도 측도 모형 필터링에 사용할 적합도 측도를 선택합니다. 기본값은 정상 R 제공입니다.

모형 예측 및 모형 지정 사항 저장

저장 탭을 사용하여 활성 데이터 세트에서 모형 예측을 새로운 변수로 저장하고 모형 사양을 외부 파일에 XML 형식으로 저장할 수 있습니다.

저장할 변수. 모형 예측, 신뢰구간 및 잔차를 활성 데이터 세트에서 새로운 변수로 저장할 수 있습니다. 각각의 종속 계열에서 자체의 새로운 변수 세트가 발생하고 각각의 새로운 변수는 추정 및 예측 기간 모두에 대한 값을 포함합니다. 예측 기간이 종속변수 계열의 길이 이상으로 확장되는 경우 새로운 케이스가 추가됩니다. 각각의 새로운 변수에 대해 관련 저장 확인 상자를 선택하여 새로운 변수 저장을 선택합니다. 기본적으로 새로운 변수는 저장되지 않습니다.

- **예측값.** 모형 예측값입니다.
- **신뢰 한계 하한** 예측값의 신뢰 한계 하한입니다.
- **신뢰 한계 상한** 예측값의 신뢰 한계 상한입니다.
- **잡음 잔차** 모형 잔차입니다. 종속변수의 변환을 수행하는 경우에는(예: 자연 로그) 변환된 계열의 잔차입니다.
- **변수 이름 접두부** 새로운 변수 이름에 대해 사용할 접두부를 지정하거나 기본값 접두부를 유지합니다. 변수 이름은 접두부, 관련 종속변수 이름 및 모형 식별자로 구성됩니다. 필요한 경우 변수 이름이 확장되어 변수 이름 충돌 문제를 방지합니다. 접두부는 유효한 변수 이름 규칙을 준수해야 합니다.

모형 파일 내보내기. 모든 추정된 모형의 모형 사양을 XML 형식의 지정된 파일로 내보냅니다. 프로시저를 사용하여, 새로운 최신 데이터를 기반으로 저장된 모형을 사용하여 업데이트된 예측을 얻을 수 있습니다.

- **XML 파일.** 모형 사양을 IBM SPSS 애플리케이션과 사용할 수 있는 XML 파일로 저장합니다.
- **PMML 파일.** 모형 지정 사항을 IBM SPSS 애플리케이션을 포함하여 PMML 호환 애플리케이션과 사용할 수 있는 PMML 호환 XML 파일로 저장합니다.

옵션

옵션 탭을 사용하여 예측 기간 설정, 결측값 처리 지정, 신뢰구간 너비 설정, 모형 식별자에 대한 사용자 정의 접두부 지정 및 자기상관에 대해 표시되는 시차 수 설정을 할 수 있습니다.

예측 기간 예측 기간은 언제나 추정 기간(모형 결정에 사용되는 케이스 세트) 종료 다음의 첫 번째 케이스와 함께 시작하여 활성 데이터 세트의 마지막 케이스 또는 사용자가 지정할 날짜까지 계속됩니다. 기본적으로 추정 기간의 종료는 활성 데이터 세트의 마지막 케이스이지만 케이스 선택 대화 상자에서 **시간 또는 케이스 범위를 기준으로**를 선택함으로써 변경될 수 있습니다.

- **추정 기간 종료 다음의 첫 번째 케이스에서 활성 데이터 세트의 마지막 케이스까지** 추정 기간 종료가 활성 데이터 세트의 마지막 케이스 이전이고 마지막 케이스까지 예측을 생성하려는 경우 이 옵션을 선택합니다. 일반적으로 이 옵션을 사용하여 검증용 기간에 대해 예측을 생성하고 실제 값 서버세트가 포함된 모형 예측을 비교할 수 있습니다.
- **추정 기간 종료 다음의 첫 번째 케이스에서 지정한 날짜까지** 명시적으로 예측 기간의 종료를 지정하려면 이 옵션을 선택합니다. 일반적으로 이 옵션을 사용하여 실제 계열 종료 이후의 예측을 생성합니다. 날짜 눈금의 모든 셀에 값을 입력합니다.

활성 데이터 세트에 대해 날짜 지정 사항이 정의되어 있지 않은 경우 날짜 눈금에 단일 열 관측값이 표시됩니다. 예측 기간의 종료를 지정하려면 관련 케이스의 행 번호를 데이터 편집기에 표시된 대로 입력합니다.

날짜 눈금에 순환 열이 있는 경우 활성 데이터 세트의 **CYCLE_** 변수 값을 참조합니다.

사용자 결측값. 사용자 결측값 처리를 제어하는 옵션입니다.

- **유효하지 않은 값으로 처리** 사용자 결측값을 시스템 결측값처럼 처리합니다.
- **유효한 값으로 처리** 사용자 결측값을 유효한 데이터로 처리합니다.

결측값 정책 모델링 프로시저를 수행하는 동안 다음과 같은 규칙을 적용하여 결측값(시스템 결측값 및 유효하지 않은 값으로 처리되는 사용자 결측값 포함)을 처리합니다.

- 추정 기간 내에서 발생하는 종속변수의 결측값이 포함된 케이스가 모형에 있습니다. 결측값에 대한 특정 처리는 추정 방법에 따라 다릅니다.
- 추정 기간 내에서 독립변수에 결측값이 있는 경우 경고를 발급합니다. 자동 모형 생성기의 경우 독립변수와 관련된 모형은 변수 없이 추정됩니다. 사용자 정의 ARIMA의 경우 독립변수와 관련된 모형은 추정되지 않습니다.
- 예측 기간 내에서 독립변수에 결측값이 있는 경우 프로시저에서 경고를 발급하고 가능한 최대 범위까지 예측합니다.

신뢰구간 너비(%). 신뢰 구간은 모델 예측 및 잔차 자기상관에 대해 계산됩니다. 100보다 작은 양수 값을 지정할 수 있습니다. 기본적으로, 95% 신뢰구간이 사용됩니다.

결과에서 모형 식별자의 접두부 변수 탭에서 지정된 각 종속변수에서 별도의 추정된 모형이 발생합니다. 모형은 사용자 정의 접두부와 정수 접미문자로 구성된 고유한 이름으로 구별합니다. 접두부를 입력하거나 모형의 기본 값을 유지할 수 있습니다.

ACF 및 PACF 결과에 표시되는 최대 시차수 자기상관 및 편자기상관 표와 도표에 표시되는 최대 시차수를 설정할 수 있습니다.

TSMODEL 명령 추가 기능

선택을 구문 창에 붙여넣고 결과 TSMODEL 명령 구문을 편집하여 시계열 모델링을 사용자 정의할 수 있습니다. 명령 구문 언어를 사용하면 다음이 가능합니다:

- **AUXILIARY** 하위 명령에서 SEASONLENGTH 키워드를 사용하여 데이터의 계절 주기를 지정합니다. 활성 데이터 세트에 대해 현재 주기성이 있는 경우 현재 주기성을 바꿉니다.
- **ARIMA** 및 **TRANSFERFUNCTION** 부명령문을 사용하여 사용자 정의 ARIMA의 비연속 시차 및 전이 함수 성분을 지정합니다. 예를 들어, 자기회귀 시차 차수 1, 3 및 6을 사용하여 사용자 정의 ARIMA 모형을 지정하거나 분자 시차 차수 2, 5 및 8을 사용하여 전이 함수를 지정할 수 있습니다.
- **MODEL** 부명령문을 사용하여 시계열 모델러 프로시저 단일 실행에 대해 두 개 이상의 모델링 지정 사항 세트(예: 모델링 방법, ARIMA 차수, 독립변수 등)를 제공합니다.

명령 구문에 대한 자세한 내용은 *Command Syntax Reference*를 참조하십시오.

시계열 모형 적용

시계열 모형 적용 프로시저는 외부 파일로부터 기존의 시계열 모형을 불러와 활성 데이터 세트에 적용합니다. 이 프로시저를 사용하여 새로운 또는 변경된 데이터를 사용할 수 있는 계열에 대해 모형을 다시 설정하지 않고 예측을 생성할 수 있습니다. Models are generated using the Time Series Modeler procedure.

예. 5,000개 제품 각각에 대해 책임을 지고 있는 주요 소매업자의 재고 관리자가 전문가 모형을 사용하여 향후 3개월 간의 각 제품 판매액을 예측하는 모형을 작성하였습니다. 데이터 웨어하우스는 월별로 업데이트된 예측을 생성하는데 사용할 실제 데이터를 사용하여 매달 새로 고칩니다. 시계열 모형 적용 프로시저에서 원래 모형으로 새로운 데이터를 설명하는 모형 모수를 재추정하여 이 작업을 수행할 수 있습니다.

통계. 적합도 측도: 정상 R -제곱, R -제곱(R^2), 평균 제곱 오차 제곱근(RMSE), 평균 절대 오차(MAE), 평균 절대 퍼센트 오차(MAPE), 절대 오차의 최대값(MaxAE), 절대 퍼센트 오차의 최대값(MaxAPE), 정규화된 베이지안 정보 기준(BIC)입니다. 잔차: 자기상관 함수, 편자기상관 함수, Ljung-Box Q 입니다.

도표. 모든 모형 전체의 요약 도표: 정상 R -제곱의 히스토그램, R -제곱(R^2), 평균 제곱 오차 제곱근(RMSE), 평균 절대 오차(MAE), 평균 절대 퍼센트 오차(MAPE), 절대 오차의 최대값(MaxAE), 절대 퍼센트 오차의 최대값(MaxAPE), 정규화된 베이지안 정보 기준(BIC), 잔차 자기상관 및 편자기상관 상자도표입니다. 개별 모형의 결과: 예측 값, 적합값, 관측값, 신뢰 한계 상한 및 하한, 잔차 자기상관 및 편자기상관입니다.

시계열 모형 적용 데이터 고려 사항

데이터. 모형이 적용될 변수(종속 및 독립)는 숫자이어야 합니다.

가정. 모형은 모형에서 지정된 변수와 동일한 이름의 활성 데이터 세트 변수에 적용됩니다. 이러한 모든 변수는 시계열로 처리되며 각 케이스는 시점을 나타내고 연속 케이스가 일정한 시간 간격으로 구분된다는 것을 의미합니다.

- **예측 독립변수(예측변수)**가 포함된 모형을 사용하여 예측을 생성하는 경우 활성 데이터 세트에는 예측 기간의 모든 케이스에 대해 이러한 변수의 값이 있어야 합니다. 모형 모수를 재추정하는 경우 독립변수는 추정 기간에 결측값이 없어야 합니다.

날짜 정의

시계열 모형 적용 프로시저에서는 활성 데이터 세트에 주기성이 있는 경우 적용될 모형의 주기성과 일치해야 합니다. 모형 설정에 사용된 것과 동일한 데이터 세트(새로운 또는 변경된 데이터를 포함할 수 있음)를 사용하여 단순 예측을 수행하는 경우 이 조건을 만족합니다. 활성 데이터 세트에 대해 주기성이 존재하지 않는 경우 날짜 정의 대화 상자로 이동하여 하나를 작성할 수 있습니다. 그러나 주기성을 지정하지 않고 모형을 작성한 경우 활성 데이터 세트에도 주기성이 없어야 합니다.

모형 적용 방법

1. 메뉴에서 다음을 선택합니다.

분석 > 예측 > 기존 모델 적용...

2. Enter the file specification for a model file or click **Browse** and select a model file (model files are created with the Time Series Modeler procedure).

선택적으로 다음을 수행할 수 있습니다.

- 활성 데이터 세트의 데이터를 사용하여 모형 모수를 재추정합니다. 재추정된 모수를 사용하여 예측을 작성합니다.
- [예측, 신뢰구간 및 잡음 잔차를 저장합니다.](#)
- [재추정된 모형을 XML 형식으로 저장합니다.](#)

모형 모수 및 적합도 측도

모형 파일에서 불러오기 모형 파일의 모형 모수를 재추정하지 않은 채 이러한 모수를 사용하여 예측을 생성합니다. 결과에 표시되고 가장 적합한 또는 적합하지 않은 모형 필터링에 사용되는 적합도 측도는 모형 파일에서 선택하며 각 모형을 개발 또는 마지막 업데이트할 때 사용된 데이터를 반영합니다. 이 옵션을 사용하는 경우 예측에서 종속변수 또는 독립변수에 대해 활성 데이터 세트의 과거 데이터를 고려하지 않습니다. 과거 데이터가 예측에 영향을 미치게 하려는 경우 **데이터에서 재추정**을 선택해야 합니다. 또한 예측에서 예측 기간의 종속 계열 값을 고려하지 않지만 예측 기간의 독립변수 값은 고려합니다. 새로운 최신 종속 계열 값을 예측에 포함시키려는 경우 이러한 값을 포함시킬 추정 기간을 조정하여 재추정해야 합니다.

데이터에서 재추정 활성 데이터 세트의 데이터를 사용하여 모형 모수를 재추정합니다. 모형 모수의 재추정은 모형 구조에 영향을 주지 않습니다. 예를 들어, ARIMA(1,0,1) 모형이 그대로 유지되지만 자동 선형회귀 및 이동 평균 모수를 재추정합니다. 재추정은 새로운 이상값을 발견하지 않습니다. 이상값이 있는 경우 언제나 모형 파일에서 선택합니다.

- **추정 기간** 추정 기간은 모형 모수 재추정에 사용되는 케이스 세트를 정의합니다. 기본적으로 추정 기간은 활성 데이터 세트의 모든 케이스를 포함합니다. 추정 기간을 설정하려면 케이스 선택 대화 상자에서 **시간 또는 케이스 범위 기본**을 선택하십시오. 사용 가능한 데이터에 따라 프로시저에서 사용되는 추정 기간이 모형별로 다를 수 있으며 따라서 표시된 값과 다를 수 있습니다. 모형이 지정되어 있는 경우 실제 추정 기간은 지정된 추정 기간의 시작 또는 끝에서 발생하는 모형의 종속변수에서 연속 결측값을 제거한 후 남은 주기입니다.

예측 기간

각 모형의 예측 기간은 언제나 추정 기간 종료 다음의 첫 번째 케이스와 함께 시작하여 활성 데이터 세트의 마지막 케이스 또는 사용자가 지정한 날짜까지 계속됩니다. 모수를 재추정하지 않는 경우(기본값인 경우) 각 모형의 추정 기간은 모형을 개발 또는 마지막 업데이트할 때 사용되었던 케이스 세트입니다.

- **추정 기간 종료 다음의 첫 번째 케이스에서 활성 데이터 세트의 마지막 케이스까지** 추정 기간 종료가 활성 데이터 세트의 마지막 케이스 이전이고 마지막 케이스까지 예측을 생성하려는 경우 이 옵션을 선택합니다.
- **추정 기간 종료 다음의 첫 번째 케이스에서 지정한 날짜까지** 명시적으로 예측 기간의 종료를 지정하려면 이 옵션을 선택합니다. 날짜 눈금의 모든 셀에 값을 입력합니다.

활성 데이터 세트에 대해 날짜 지정 사항이 정의되어 있지 않은 경우 날짜 눈금에 단일 열 관측값이 표시됩니다. 예측 기간의 종료를 지정하려면 관련 케이스의 행 번호를 데이터 편집기에 표시된 대로 입력합니다.

날짜 눈금에 순환 열이 있는 경우 활성 데이터 세트의 `CYCLE_` 변수 값을 참조합니다.

결과

사용 가능한 결과로는 개별 모형의 결과 및 모든 모형에 대한 결과가 있습니다. 개별 모형의 결과를 사용자가 지정한 기준에 따라 가장 적합한 모형 또는 가장 적합하지 않은 모형의 세트로 제한할 수 있습니다.

통계 및 예측표

통계 탭은 모형 적합 통계, 모형 모수, 자기상관 함수 및 예측 표 표시 옵션을 제공합니다. 모형 모수를 재추정하지 않는 경우(모형 탭의 **데이터에서 재추정**) 표시된 적합도 값, Ljung-Box 값 및 모형 모수는 모형 파일에서 가져온 것이며 각 모형을 개발 또는 마지막 업데이트할 때 사용되었던 데이터를 반영합니다. 이상값 정보는 언제나 모형 파일에서 선택합니다.

모형별 적합도, Ljung-Box 통계 및 이상값 수 표시. 각 모형에 대해 선택된 적합도, Ljung-Box 값 및 이상값 수가 포함된 테이블을 표시하려면 이 옵션을 선택합니다.

적합도. 각 모형의 적합도가 포함된 표에 다음 중 하나 이상을 선택하여 포함시킬 수 있습니다.

- 정상 R -제곱
- R -제곱
- 평균 제곱근 오차
- 평균 절대 백분율 오차
- 평균 절대 오차
- 최대 절대 백분율 오차
- 최대 절대 오차
- 정규화된 BIC

모형 비교 통계 이 옵션 그룹은 모형 모두에 대한 통계를 포함하고 있는 테이블의 표시를 제어합니다. 각 옵션은 별도의 표를 생성합니다. 다음 옵션 중 한 개 이상을 선택할 수 있습니다.

- **적합도.** 고정 R -제곱, R -제곱, 루트 평균 제곱 오차, 평균 절대 백분율 오차, 평균 절대 오차, 최대 절대 백분율 오차, 최대 절대 오차 및 정규화된 베이저안 정보 기준에 대한 요약 통계와 백분위수 표입니다.
- **잔차 자기상관 함수(ACF)** 추정된 모형 모두의 잔차 자기상관에 대한 요약 통계 및 백분위수 표입니다. 이 테이블은 모형 모수를 재추정하는 경우에만 사용할 수 있습니다(모형 탭의 **데이터에서 재추정**).
- **잔차 편자기상관 함수(PACF)** 추정된 모형 모두의 잔차 편자기상관에 대한 요약 통계 및 백분위수 표입니다. 이 테이블은 모형 모수를 재추정하는 경우에만 사용할 수 있습니다(모형 탭의 **데이터에서 재추정**).

개별 모형 통계 이 옵션 그룹은 각 모형의 상세한 정보를 포함하고 있는 테이블의 표시를 제어합니다. 각 옵션은 별도의 표를 생성합니다. 다음 옵션 중 한 개 이상을 선택할 수 있습니다.

- **모수 추정값.** 각 모형의 모수 추정값 테이블을 표시합니다. 지수평활 및 ARIMA 모형에 대해 별도의 테이블이 표시됩니다. 이상값이 존재하는 경우 해당 모수 추정값도 별도의 테이블에 표시됩니다.
- **잔차 자기상관 함수(ACF)** 각 추정된 모형의 시차별 잔차 자기상관 표를 표시합니다. 표에는 자기상관의 신뢰구간이 포함됩니다. 이 테이블은 모형 모수를 재추정하는 경우에만 사용할 수 있습니다(모형 탭의 **데이터에서 재추정**).
- **잔차 편자기상관 함수(PACF)** 각 추정된 모형의 시차별 잔차 편자기상관 표를 표시합니다. 테이블에는 편자기상관의 신뢰구간이 포함됩니다. 이 테이블은 모형 모수를 재추정하는 경우에만 사용할 수 있습니다(모형 탭의 **데이터에서 재추정**).

예측 표시 각 모형의 모형 예측 및 신뢰구간 테이블을 표시합니다.

플롯

도표 탭은 모형 적합 통계, 자기상관 함수 및 계열 값(예측 포함) 도표 표시 옵션을 제공합니다.

모형 비교 도표

이 옵션 그룹은 모형 모두에 대한 통계를 포함하고 있는 도표의 표시를 제어합니다. 모형 모수를 재추정하지 않는 경우(모형 탭의 **데이터에서 재추정**) 표시된 값은 모형 파일에서 가져온 것이며 각 모형을 개발 또는 마지막 업데이트할 때 사용되었던 데이터를 반영합니다. 또한 자기상관 도표는 모형 모수를 재추정하는 경우에만 사용할 수 있습니다. 각 옵션은 별도의 도표를 생성합니다. 다음 옵션 중 한 개 이상을 선택할 수 있습니다.

- 정상 R-제공
- R-제공
- 평균 제공근 오차
- 평균 절대 백분율 오차
- 평균 절대 오차
- 최대 절대 백분율 오차
- 최대 절대 오차
- 정규화된 BIC
- 잔차 자기상관 함수
- 잔차 편자기상관 함수(PACF)

개별 모형 도표

계열 각 모형의 예측값 도표를 구하려면 이 옵션을 선택합니다. 관측값, 적합값, 적합값의 신뢰구간 및 자기상관은 모형 모수를 재추정하는 경우에만 사용할 수 있습니다(모형 탭의 **데이터에서 재추정**). 다음 중 한 개 이상을 선택하여 도표에 포함시킬 수 있습니다.

- 관측값 종속 계열의 관측값입니다.
- 예측 예측 기간에 대한 모형 예측값입니다.
- 적합값 추정 기간에 대한 모형 예측값입니다.
- 예측 신뢰구간 예측 기간에 대한 신뢰구간입니다.
- 적합값 신뢰구간 추정 기간에 대한 신뢰구간입니다.

잔차 자기상관 함수(ACF) 각 추정된 모형의 잔차 자기상관 도표를 표시합니다.

잔차 편자기상관 함수(PACF) 각 추정된 모형의 잔차 편자기상관 도표를 표시합니다.

가장 적합한 모형 또는 가장 적합하지 않은 모형으로 결과 제한

결과 필터 탭은 표 및 차트 결과 모두를 모형의 서브세트로 제한하는 옵션을 제공합니다. 입력한 적합 기준에 따라 결과를 최적 적합 모형 또는 가장 적합하지 않은 모형으로 제한하도록 선택할 수 있습니다. 기본적으로 모든 모형이 결과에 포함됩니다. 모형 모수를 재추정하지 않는 경우(모형 탭의 **데이터에서 재추정**) 모형 필터링에 사용되는 적합도 값은 모형 파일에서 가져온 것이며 각 모형을 개발 또는 마지막 업데이트할 때 사용되었던 데이터를 반영합니다.

최적 적합 모형 최적 적합 모형을 결과에 포함시키려면 이 옵션을 선택합니다. 적합도 측도를 선택하고 포함시킬 모형의 수를 지정합니다. 이 옵션을 선택하는 경우에도 가장 적합하지 않은 모형을 선택할 수 있습니다. 이러한 케이스에서는 가장 적합하지 않은 모형 및 최적 적합 모형으로 결과가 구성됩니다.

- **고정된 모형의 수** n 개의 최적 적합 모형에 대해 결과를 표시하도록 지정합니다. 해당 숫자가 전체 모형 수를 초과하는 경우 모든 모형이 표시됩니다.
- **전체 모형 수의 퍼센트** 모형 전체에서 적합도 값이 상위 n 퍼센트인 모형에 대해 결과를 표시하도록 지정합니다.

가장 적합하지 않은 모형 가장 적합하지 않은 모형을 결과에 포함시키려면 이 옵션을 선택합니다. 적합도 측도를 선택하고 포함시킬 모형의 수를 지정합니다. 이 옵션을 선택하는 경우에도 최적 적합 모형을 선택할 수 있습니다. 이러한 케이스에서는 최적 적합 모형 및 가장 적합하지 않은 모형으로 결과가 구성됩니다.

- **고정된 모형의 수** n 개의 가장 적합하지 않은 모형에 대해 결과를 표시하도록 지정합니다. 해당 숫자가 전체 모형 수를 초과하는 경우 모든 모형이 표시됩니다.

- 전체 모형 수의 퍼센트 모형 전체에서 적합도 값이 하위 n 퍼센트인 모형에 대해 결과를 표시하도록 지정합니다.

적합도 측도 모형 필터링에 사용할 적합도 측도를 선택합니다. 기본값은 정상 R -제공입니다.

모형 예측 및 모형 지정 사항 저장

저장 탭을 사용하여 활성 데이터 세트에서 모형 예측을 새로운 변수로 저장하고 모형 사양을 외부 파일에 XML 형식으로 저장할 수 있습니다.

저장할 변수. 모형 예측, 신뢰구간 및 잔차를 활성 데이터 세트에서 새로운 변수로 저장할 수 있습니다. 각 모형에서 자체의 새로운 변수 세트가 발생합니다. 예측 기간이 모형과 관련된 종속변수 계열의 길이 이상으로 확장되는 경우 새로운 케이스가 추가됩니다. 모형 모수를 재추정하지 않는 경우(모형 탭의 **데이터에서 재추정**) 예측값 및 신뢰 한계는 예측 기간에 대해서만 작성됩니다. 각각의 새로운 변수에 대해 관련 저장 확인 상자를 선택하여 새로운 변수 저장을 선택합니다. 기본적으로 새로운 변수는 저장되지 않습니다.

- **예측값.** 모형 예측값입니다.
- **신뢰 한계 하한** 예측값의 신뢰 한계 하한입니다.
- **신뢰 한계 상한** 예측값의 신뢰 한계 상한입니다.
- **잡음 잔차** 모형 잔차입니다. 종속변수의 변환을 수행하는 경우에는(예: 자연 로그) 변환된 계열의 잔차입니다. 이 선택 사항은 모형 모수를 재추정하는 경우에만 사용할 수 있습니다(모형 탭의 **데이터에서 재추정**).
- **변수 이름 접두부** 새로운 변수 이름에 대해 사용할 접두부를 지정하거나 기본값 접두부를 유지합니다. 변수 이름은 접두부, 관련 종속변수 이름 및 모형 식별자로 구성됩니다. 필요한 경우 변수 이름이 확장되어 변수 이름 충돌 문제를 방지합니다. 접두부는 유효한 변수 이름 규칙을 준수해야 합니다.

모형 파일 내보내기 재추정된 모수 및 적합 통계를 포함하여 모형 지정 사항을 XML 형식의 지정된 파일로 내보냅니다. 이 옵션은 모형 모수를 재추정하는 경우에만 사용할 수 있습니다(모형 탭의 **데이터에서 재추정**).

- **XML 파일.** 모형 지정 사항을 IBM SPSS 애플리케이션과 사용할 수 있는 XML 파일로 저장합니다.
- **PMML 파일.** 모형 지정 사항을 IBM SPSS 애플리케이션을 포함하여 PMML 호환 애플리케이션과 사용할 수 있는 PMML 호환 XML 파일로 저장합니다.

옵션

옵션 탭을 사용하여 결측값 처리를 지정하고 신뢰구간 너비를 설정하며 자기상관에 대해 표시되는 시차 수를 설정할 수 있습니다.

사용자 결측값. 사용자 결측값 처리를 제어하는 옵션입니다.

- **유효하지 않은 값으로 처리** 사용자 결측값을 시스템 결측값처럼 처리합니다.
- **유효한 값으로 처리** 사용자 결측값을 유효한 데이터로 처리합니다.

결측값 정책 다음과 같은 규칙을 적용하여 결측값(시스템-결측값 및 유효하지 않은 값으로 처리되는 사용자 결측값 포함)을 처리합니다.

- 추정 기간 내에서 발생하는 종속변수의 결측값이 포함된 케이스가 모형에 있습니다. 결측값에 대한 특정 처리는 추정 방법에 따라 다릅니다.
- ARIMA 모형의 경우 추정 기간 내에서 예측변수에 결측값이 있는 경우 경고를 발급합니다. 예측변수와 관련된 모형은 재추정되지 않습니다.
- 예측 기간 내에서 독립변수에 결측값이 있는 경우 프로시저에서 경고를 발급하고 가능한 최대 범위까지 예측합니다.

신뢰구간 너비(%). 신뢰 구간은 모델 예측 및 잔차 자기상관에 대해 계산됩니다. 100보다 작은 양수 값을 지정할 수 있습니다. 기본적으로, 95% 신뢰구간이 사용됩니다.

ACF 및 PACF 결과에 표시되는 최대 시차수 자기상관 및 편자기상관 표와 도표에 표시되는 최대 시차수를 설정할 수 있습니다. 이 옵션은 모형 모수를 재추정하는 경우에만 사용할 수 있습니다(모형 탭의 **데이터에서 재추정**).

TSAPPLY 명령 추가 기능

선택을 구문 창으로 붙여넣고 결과로 나타나는 TSAPPLY 명령 구문을 편집하는 경우 추가 기능을 사용할 수 있습니다. 명령 구문 언어를 사용하면 다음이 가능합니다:

- MODEL 부명령문에서 DROP 및 KEEP 키워드를 사용하여 모형 파일에서 모형의 서브세트만 활성 데이터 세트에 적용하도록 지정합니다.
- MODEL 부명령문을 사용하여 모형을 모형 파일 두 개 이상으로부터 데이터에 적용합니다. 예를 들어, 한 모형 파일에 판매 건수를 나타내는 계열의 모형이 있고 다른 모형 파일에 매출을 나타내는 계열의 모형이 있을 수 있습니다.

명령 구문에 대한 자세한 내용은 *Command Syntax Reference*를 참조하십시오.

계절 분해

계절 분해 프로시저는 계열을 계절 성분, 결합된 추세 및 순환 성분 및 "오차" 성분으로 분해합니다. 이 프로시저는 비율 이동 평균 방법이라고도 하는 Census Method I의 구현입니다.

예. 과학자가 특정 기상 관측소의 월별 오존 수준 측정값을 분석하려고 합니다. 데이터에 추세가 있는지 결정하는 것이 목표입니다. 실제 추세를 확인하려면 과학자는 먼저 계절 효과로 인한 기록의 변동을 설명해야 합니다. 계절 분해 프로시저를 사용하여 계통 계절 변동을 제거할 수 있습니다. 그런 다음 계절별 조정 계열에 대해 추세 분석을 수행합니다.

통계. 계절 요인 세트입니다.

계절 분해 데이터 고려 사항

데이터. 변수는 숫자이어야 합니다.

가정. 변수에는 포함된 결측 데이터가 없어야 합니다. 주기별 날짜 성분을 한 개 이상 정의해야 합니다.

계절 요인 추정

1. 메뉴에서 다음을 선택합니다.

분석 > 예측 > 계절 분해...

2. 사용 가능한 목록에서 하나 이상의 변수를 선택하여 변수 목록으로 이동합니다. 목록에는 숫자 변수만 있습니다.

모델 유형. 계절 분해 프로시저에는 계절 요인을 모델링하는 두 가지 다른 접근법(승법 또는 가법)이 있습니다.

- 승법. 계절 성분은 계절별 조정 계열에 곱하여 원래 계열을 산출하는 요인입니다. 실제로 계절 성분은 전체 계열 수준에 비례하는 계절 성분을 추정합니다. 계절 변동이 없는 관측값의 계절 성분은 1입니다.
- 가법. 계절 조정은 관측값을 얻기 위해 계절별 조정 계열에 추가됩니다. 이 조정에서는 계절 성분에서 "감출 수 있는" 다른 주요 공정특성을 보기 위해 계열에서 계절 효과를 제거합니다. 실제로 계절 성분은 전체 계열 수준에 영향을 받지 않습니다. 계절 변동이 없는 관측값의 계절 성분은 0입니다.

이동 평균가중치 이동 평균가중치 옵션을 사용하여 이동 평균을 계산할 때 해당 계열을 처리하는 방법을 지정할 수 있습니다. 이 옵션은 해당 계열의 주기성이 짝수인 경우에만 사용할 수 있습니다. 주기성이 홀수인 경우에는 모든 시점이 동일한 가중치를 갖게 됩니다.

- 모든 시점이 동일(*All points equal*). 이동 평균은 주기성과 동일한 범위와 모든 점의 가중치를 동일하게 하여 계산됩니다. 이 방법은 주기성이 홀수일 때 항상 사용됩니다.
- 엔드포인트에 0.5를 가중치 적용(*Endpoints weighted by .5*). 주기성이 짝수인 계열의 이동 평균은 주기성에 1을 더한 값과 같은 범위와 범위의 엔드포인트에 0.5 가중치를 부여한 값을 사용하여 계산됩니다.

선택적으로 다음을 수행할 수 있습니다.

- **저장**을 클릭하여 새 변수가 저장되는 방법을 지정합니다.

계절 분해 저장

변수 작성 새 변수 처리 방법을 선택할 수 있습니다.

- 파일에 추가(Add to file). 계절 분해로 작성된 새 계열은 활성 데이터 세트에 정규 변수로 저장됩니다. 변수 이름은 3자 접두부, 밑줄 및 숫자로 구성됩니다.
- 기존변수 바꾸기(Replace existing). 계절 분해에 의해 작성된 새 계열은 활성 데이터 세트에 임시 변수로 저장됩니다. 이 때 예측 프로시저로 작성된 기존의 임시 변수가 있으면 바뀝니다. 변수 이름은 3개의 접두부, 파운드 부호(#) 및 숫자로 구성됩니다.
- 만들지 않음(Do not create). 새로운 계열이 활성 데이터 세트에 추가되지 않습니다.

새로운 변수 이름

계절 분해 프로시저에서는 지정된 각 계열에 대해 다음에 나오는 세 글자로 된 접두부가 포함된 네 개의 새 변수(계열)를 작성합니다.

SAF 계절조정 요인입니다. 이러한 값은 각 주기의 계열 수준에 대한 영향을 나타냅니다.

SAS. 계절별 조정 계열 계열의 계절 변동을 제거한 다음에 구한 값입니다.

STC 평활 추세-순환 성분입니다. 이러한 값은 계열에 있는 추세 및 순환 동작을 나타냅니다.

ERR 잔차 또는 "오차" 값입니다. 계열에서 계절, 추세 및 순환 성분이 제거된 후 남아있는 값입니다.

SEASON 명령 추가 기능

명령 구문 언어를 사용하여 수행할 수 있는 추가 기능은 다음과 같습니다.

- 날짜 정의 프로시저에서 제공하는 항목 중 하나를 선택하기보다는 SEASON 명령 내에서 주기성을 지정할 수 있습니다.

명령 구문에 대한 자세한 내용은 *Command Syntax Reference*를 참조하십시오.

스펙트럼 도표

스펙트럼 도표 프로시저를 사용하여 시계열에서 주기 동작을 식별합니다. 한 시점에서 다음 시점에서의 변동을 분석하는 대신에 스펙트럼 분석은 계열 전체의 변동을 서로 다른 빈도의 주기 성분으로 분석합니다. 평활 계열은 낮은 빈도에서 더 강한 주기 성분을 갖게 됩니다. 임의 변동("백색 잡음")은 모든 빈도에 대하여 성분 강도를 갖게 됩니다.

결측 데이터를 포함하는 계열은 이 프로시저를 사용하여 분석할 수 없습니다.

예. 새 주택 건설 비율은 중요한 경제 상태 지표입니다. 주택 건설 시작 데이터는 일반적으로 강한 계절 성분을 나타냅니다. 데이터에 현재 상태를 평가할 때 분석가가 유의해야 하는 더 긴 순환이 있습니까?

통계. 각각의 빈도 또는 주기 성분에 대한 사인 및 코사인 변환, 주기도 값 그리고 스펙트럼 밀도 추정값입니다. 이 변량 분석이 선택된 경우에는 각각의 빈도 또는 주기 성분에 대한 교차 주기도, 공스펙트럼 밀도, 구적 스펙트럼, 증가, 제곱 일관도 및 위상 스펙트럼의 실제 또는 가상 부분입니다.

도표. 일변량 및 이변량 분석의 경우 주기도 및 스펙트럼 밀도입니다. 이변량 분석에 대해서는 제곱 일관도, 구적 스펙트럼, 교차진폭, 공스펙트럼 밀도, 위상 스펙트럼 및 증가입니다.

스펙트럼 도표 데이터 고려 사항

데이터. 변수는 숫자이어야 합니다.

가정. 변수에는 포함된 결측 데이터가 없어야 합니다. 분석할 시계열 데이터는 정상이어야 하며 0이 아닌 평균은 계열에서 빼야 합니다.

- 고정. ARIMA 모델에 적합한 시계열에 의해 충족되어야 하는 조건입니다. 순수 MA 계열은 정상적이지만 AR과 ARMA 계열은 아닐 수 있습니다. 정상 계열은 시간에 따라 일정한 평균과 분산을 가집니다.

스펙트럼 분석 구하기

1. 메뉴에서 다음을 선택합니다.

분석 > 시계열 > 스펙트럼 분석...

2. 사용 가능한 목록에서 하나 이상의 변수를 선택하여 변수 목록으로 이동합니다. 목록에는 숫자 변수만 있습니다.

3. 스펙트럼 밀도 추정값을 구하기 위해 주기를 평활하는 방법을 선택하려면 스펙트럼 창 옵션 중 하나를 선택합니다. 사용 가능한 평활 옵션은 Tukey-Hamming, Tukey, Parzen, Bartlett, Daniell(Unit) 및 지정없음입니다.

- **Tukey-Hamming (Tukey-Hamming).** 가중치는 $k = 0, \dots, p$ 의 경우 $W_k = .54D_p(2 \pi f k) + .23D_p(2 \pi f k + \pi/p) + .23D_p(2 \pi f k - \pi/p)$, 여기서, p 는 스펙의 절반의 정수 부분이고, D_p 는 차수 p 의 Dirichlet 커널입니다.
- **Tukey.** 가중치는 $k = 0, \dots, p$ 의 경우 $W_k = 0.5D_p(2 \pi f k) + 0.25D_p(2 \pi f k + \pi/p) + 0.25D_p(2 \pi f k - \pi/p)$, 여기서, p 는 스펙의 절반의 정수 부분이고, D_p 는 차수 p 의 Dirichlet 커널입니다.
- **Parzen.** 가중치는 $k = 0, \dots, p$ 의 경우 $W_k = 1/p(2 + \cos(2 \pi f k)) (F[p/2](2 \pi f k))^{**2}$, 여기서 p 는 스펙 절반의 정수 부분이고 $F[p/2]$ 는 차수 $p/2$ 의 Fejer 커널입니다.
- **Bartlett.** 창 상반부의 가중치가 $k = 0, \dots, p$ 인 경우 $W_k = F_p(2 \pi f k)$ 로 계산되는 스펙트럼 창, 여기서 p 는 스펙 절반의 정수 부분이고, F_p 는 차수 p 의 Fejer 커널입니다. 하반은 상반과 대칭입니다.
- **Daniell(Unit).** 가중치가 모두 1인 스펙트럼 창 형태입니다.
- 없음. 매끄럽게 하지 않음. 이 옵션을 선택하면 스펙트럼 밀도 추정값이 주기와 같아집니다.

범위. 평활이 수행되는 연속 값의 범위입니다. 일반적으로 홀수가 사용됩니다. 범위가 클수록 스펙트럼 밀도 도표는 더욱 평평해집니다.

계열을 평균 0으로 조정(Center variables). 스펙트럼 계산 이전에 계열 평균이 0이 되고, 계열 평균과 연관될 수 있는 큰 항이 제거되도록 계열을 조정합니다.

이변량 분석은 처음 변수와 기타 변수입니다. 둘 이상의 변수를 선택한 경우 이 옵션을 선택하여 이변량 스펙트럼 분석을 요청할 수 있습니다.

- 변수 목록의 첫 번째 변수를 독립변수로 취급하고 나머지 변수를 종속변수로 취급합니다.
- 다른 계열과는 독립적으로 이름 붙여진 첫 번째 계열을 이용하여 첫 번째 이후의 계열을 분석합니다. 각 계열의 이변량 분석도 수행됩니다.

도표. 주기도 및 스펙트럼 밀도를 이변량 및 이변량 분석에 사용할 수 있습니다. 다른 모든 선택 사항은 이변량 분석에 대해서만 사용할 수 있습니다.

- **주기도(Periodogram).** 빈도 또는 주기에 대한 스펙트럼 진폭 (로그 척도로 도표화됨)의 평활되지 않은 도표입니다. 낮은 빈도 변동은 평활 계열의 특징을 나타냅니다. 전체 빈도에 균일하게 펼쳐진 변동은 "백색 잡음"을 나타냅니다.
- **제곱 일관도(Squared coherency).** 두 계열의 gain 값의 곱입니다.
- **구적 스펙트럼(Quadrature spectrum).** 교차 주기의 허수 부분으로, 두 시계열의 비위상 빈도 성분 상관에 대한 측도입니다. 성분의 위상은 $\pi/2$ 라디안입니다.
- **교차진폭(Cross amplitude).** 공스펙트럼 밀도 제곱과 구적 스펙트럼 제곱을 더한 합의 제곱근입니다.
- **스펙트럼 밀도.** 불규칙한 변동을 제거하도록 평활화된 주기도입니다.
- **공스펙트럼 밀도(Cospectral density).** 두 시계열에 대한 동위상 빈도 성분의 상관 측도인 교차 주기의 실수 부분입니다.
- **거듭제곱 스펙트럼 밀도 (Power spectral density).** 신호의 전력 스펙트럼 밀도 (또는 전력 밀도 또는 전력 밀도 스펙트럼 또는 전력 스펙트럼)는 단위 주파수당 주파수의 함수로서 신호에 존재하는 전력을 설명합니다.
- **Coherence.** Coherence는 신호 처리에서 두 신호 또는 데이터 세트 간의 관계를 검사하는 데 사용할 수 있는 통계입니다. 일반적으로 선형 시스템의 입력 및 출력 사이의 전력 전송을 추정하는 데 사용됩니다.
- **위상 스펙트럼(Phase spectrum).** 한 계열의 각 빈도 성분이 다른 계열을 리드하거나 시차를 두는 범위에 대한 측도입니다.
- **이득(Gain).** 교차 진폭을 계열 중 하나의 스펙트럼 밀도로 나눈 몫입니다. 두 계열에는 각각 고유한 gain 값이 있습니다.

빈도 기준(By frequency). 모든 도표가 빈도를 기준으로 만들어지며 빈도의 범위는 0(상수항 또는 평균 항)에서 0.5(두 관측값 사이를 순환하는 항)까지입니다.

주기 기준(By period). 모든 도표는 주기별로 작성되며 범위는 2 (두 개의 관측값 주기에 대한 항)에서 관측값 수 (상수 또는 평균 항)와 동일한 주기까지입니다. 주기는 로그 척도에 표시됩니다.

SPECTRA 명령 추가 기능

명령 구문 언어를 사용하여 수행할 수 있는 추가 기능은 다음과 같습니다.

- 나중에 다시 사용하려면 계산된 스펙트럼 분석 변수를 활성 데이터 세트에 저장합니다.
- 스펙트럼 창에 대하여 사용자 정의 가중치를 지정합니다.
- 빈도와 주기를 동시에 나타내는 도표를 작성합니다.
- 도표에 나타나는 값에 대한 완전한 목록을 인쇄합니다.

명령 구문에 대한 자세한 내용은 *Command Syntax Reference*를 참조하십시오.

시간 인과 모형

시간 인과 모델링은 시계열 데이터에서 핵심 인과 관계를 검색하려고 시도합니다. 시간 인과 모형화에서는 대상 계열 세트 및 해당 목표에 대한 후보 입력 세트를 지정합니다. 그런 다음 프로시저에서 각 목표에 대한 자동 선형 회귀 시계열 모형을 작성하고 목표와 인과 관계가 있는 입력만 포함시킵니다. 이 접근 방식은 대상 계열에 대한 예측자를 명시적으로 지정해야 하는 기존 시계열과 다릅니다. 시간 인과 모델링에는 일반적으로 여러 관련 시계열에 대한 모델 작성이 포함되므로, 결과를 모델 시스템이라고 합니다.

시간 인과 모델링의 컨텍스트에서, 인과 용어는 그랜저 인과성을 가리킵니다. X 및 Y 둘 다의 지난 값 관점에서 Y 에 대한 회귀추정이 Y 의 지난 값에 대해서만 회귀추정하는 것보다 Y 에 대해 더 나은 모델이 생성되는 경우 시계열 X 는 다른 시계열 Y 의 "그랜저 인과" 관계라고 합니다.

예제

비즈니스 의사결정자는 비즈니스를 설명하는 시간 기반 메트릭의 큰 세트 내에서 인과 관계를 알아내기 위해 시간 인과 모델링을 사용할 수 있습니다. 분석은 몇 개의 제어 가능한 입력을 드러낼 수 있습니다. 이 입력은 핵심성과 지표(KPI)에 대한 가장 큰 영향을 가지고 있습니다.

대규모 IT 시스템의 관리자는 시간 인과 모델링을 사용하여 상호 관련된 작동 메트릭의 큰 세트에서 이상 항목을 발견할 수 있습니다. 인과 모델은 이상 항목 발견을 넘어서 이상 항목의 가장 근본적인 원인을 발견할 수 있도록 합니다.

필드 요구 사항

최소 하나의 목표가 있어야 합니다. 기본적으로, 사전 정의된 역할이 없음인 필드는 사용되지 않습니다.

데이터 구조

시간 인과 모델링은 두 가지 유형의 데이터 구조를 지원합니다.

열 기반 데이터

열 기반 데이터의 경우, 각 시계열 필드에는 단일 시계열에 대한 데이터가 포함됩니다. 이 구조는 시계열 Modeler에서 사용되는 전형적인 시계열 데이터 구조입니다.

다차원 데이터

다차원 데이터의 경우, 각 시계열 필드에는 여러 시계열에 대한 데이터가 포함됩니다. 특정 필드 내에서 별도의 시계열은 차원 필드라고 하는 범주형 필드의 값 세트로 식별됩니다. 예를 들어 두 개의 다른 판매 채널(소매 및 웹)의 판매 데이터가 단일 sales 필드에 저장될 수 있습니다. 이름이 channel이고 값이 'retail' 및 'web'인 차원 필드는 두 판매 채널 각각과 연관되는 레코드를 식별합니다.

시간 인과 모델 얻기

이 기능에는 통계량 예측 옵션이 필요합니다.

메뉴에서 다음을 선택합니다.

분석 > 예측 > 시간 인과 모형 작성하기 ...

1. 관측값이 날짜/시간 필드로 정의된 경우 해당 필드를 지정하십시오.
2. 데이터가 다차원이면 시계열을 식별하는 차원 필드를 지정하십시오.

- 차원 필드의 지정된 순서는 모두 연속 대화 상자 및 출력에 차원이 나타나는 순서를 정의합니다. 차원 필드를 다시 정렬하려면 위로 및 아래로 화살표 단추를 사용하십시오.
- 분석을 특히 차원 필드의 값으로 제한하는 차원 값의 부분 집합을 지정할 수 있습니다. 예를 들어, 지역 및 브랜드에 대한 차원이 있는 경우 특정 지역으로 분석으로 제한할 수 있습니다. 차원 서브세트는 분석에서 사용되는 모든 메트릭 필드에 적용됩니다.
- 또한 메트릭 필드 수준에서 역할별 차원 값을 지정하여 분석을 사용자 정의할 수도 있습니다. 예를 들어, 판매 채널('소매점' 및 '웹' 값을 가짐)에 대한 차원과 그러한 채널에서의 판매 및 광고에 대한 메트릭이 있는 경우, 웹 광고를 소매점 및 웹 판매 모두에 대한 입력으로 지정할 수 있습니다. 기본적으로 이러한 유형의 사용자 정의가 사용되고 선택 범위가 각 차원 필드의 지정된 고유 값 수(기본적으로 250개)로 제한됩니다.

3. 계속을 클릭합니다.

참고: 활성 데이터 세트에 날짜 지정이 있으면 1, 2 및 3단계가 적용되지 않습니다. 날짜 지정은 날짜 정의 대화 상자에서 또는 DATE 명령으로 작성됩니다.

4. 필드를 클릭하여 모델에 포함시킬 시계열을 지정하고 관측값을 정의하는 방법을 지정하십시오. 하나 이상의 필드가 대상 또는 입력과 대상 둘 다를 지정되어야 합니다.
5. **데이터 지정**을 클릭하여 결측값 처리와 통합 및 분포 설정, 그리고 분석을 위한 시간 구간을 포함하는 선택적 설정을 지정하십시오.
6. **작성 옵션**을 클릭하여 추정 기간을 정의하고 출력의 콘텐츠를 지정하며 목표당 입력 최대 수와 같은 작성 설정을 지정하십시오.
7. **모델 옵션**을 클릭하여 예측을 요청하고 예측값을 저장하며 모델 시스템을 외부 파일로 내보내십시오.
8. 프로시저를 실행하려면 **실행**을 클릭하십시오.

모형에 대한 시계열

필드 탭에서 **시계열** 설정을 사용하여 모형 시스템에 포함시킬 계열을 지정하십시오.

열 기반 데이터에 대해 계열 용어의 의미는 필드 용어 의미와 같습니다. 다차원 데이터의 경우, 시계열을 포함하는 필드는 메트릭 필드로 언급됩니다. 다차원 데이터에 대해 시계열은 차원 필드 각각에 대한 메트릭 필드 및 값으로 정의됩니다. 열 기반 및 다차원 데이터에 대해 다음 고려사항이 적용됩니다.

- 후보 입력으로, 또는 목표 및 입력 둘 다를 지정되는 계열은 각 목표에 대해 모델에서 포함을 위해 고려됩니다. 각 목표의 모형에는 항상 대상 자체의 지연된 값이 포함됩니다.
- 강제 입력으로 지정되는 계열은 항상 각 목표에 대해 모델에 포함됩니다.
- 하나 이상의 계열이 대상 또는 목표와 입력 둘 다를 지정되어야 합니다.
- **사전 정의된 역할 사용**이 선택되면, 입력 역할을 가지고 있는 필드가 후보 입력으로 설정됩니다. 사전 정의된 어떤 규칙도 강제 입력에 맵핑하지 않습니다.

다차원 데이터

다차원 데이터의 경우, 눈금에서 메트릭 필드 및 연관된 역할을 지정합니다. 눈금의 각 행은 단일 메트릭 및 역할을 지정합니다. 기본적으로 모형 시스템은 격자의 각 행에 대한 차원 필드의 모든 조합에 대한 계열을 포함합니다. 예를 들어, *region* 및 *brand*에 대한 차원이 있는 경우, 기본적으로 목표로 메트릭 *sales*를 지정하면, 이는 각각의 *region* 및 *brand* 조합에 대해 별도의 *sales* 목표 계열이 있음을 의미합니다.

격자의 각 행에 대해 차원의 생략 기호 단추를 클릭하여 차원 필드의 값 세트를 사용자 정의할 수 있습니다. 이 동작은 차원 값 선택 하위 대화 상자를 엽니다. 또한 눈금 행을 추가, 삭제 또는 복사할 수도 있습니다.

계열 수 열은 연관된 정량에 대해 현재 지정된 차원 값 세트 수를 표시합니다. 표시된 값은 실제 계열 수(세트당 하나의 계열)보다 클 수 있습니다. 이 조건은 지정된 차원 값 조합 중 일부가 연관 메트릭에 의해 포함된 계열에 해당되지 않는 경우에 발생합니다.

차원 값 선택을 선택하십시오

다차원 데이터의 경우, 특정 역할이 있는 특정 메트릭 필드에 적용되는 차원 값을 지정하여 분석을 사용자 정의할 수 있습니다. 예를 들어 *sales*가 메트릭 필드이고 *channel* 값이 'retail' 및 'web'인 경우 'web' 판매를 입력으로 하고 'retail' 판매를 목표로 지정할 수 있습니다. 또한 분석에 사용되는 모든 메트릭 필드에 적용되는 차원 서브세트

를 지정할 수도 있습니다. 예를 들어, *region*이 지리적 지역을 나타내는 차원 필드인 경우 분석을 특정 지역으로 제한할 수 있습니다.

모든 값

현재 차원 필드의 모든 값이 포함됨을 지정합니다. 이 옵션이 기본값입니다.

포함하거나 제외할 값 선택

현재 차원 필드의 값 세트를 지정하려면 이 옵션을 사용하십시오. **모드**에 대해 **포함**이 선택되는 경우, **선택된 값** 목록에 지정되는 값만 포함됩니다. **모드**에 대해 **제외**가 선택되는 경우, **선택된 값** 목록에 지정된 값이 아닌 다른 모든 값이 포함됩니다.

선택할 값 세트를 필터링할 수 있습니다. 필터 조건에 충족하는 값은 **매치됨** 탭에 나타나고, 필터 조건과 일치하지 않는 값은 **선택되지 않은 값** 목록의 **매치하지 않음** 탭에 나타납니다. **모두** 탭은 필터 조건과 관계없이 선택되지 않은 모든 값을 나열합니다.

- 필터를 지정할 때 별표(*)를 사용하여 와일드카드 문자를 표시할 수 있습니다.
- 현재 필터를 지우려면, 표시된 값 필터링 대화 상자에 검색어로 비어 있는 값을 지정하십시오.

관측값

필드 탭에서 **관측값** 설정을 사용하여 관측값을 정의하는 필드를 지정하십시오.

참고: 활성 데이터 세트에 날짜 지정이 있는 경우 관측값은 날짜 지정에 의해 정의되며 시간 인과 모델링 프로시저에서 수정할 수 없습니다. 날짜 지정은 날짜 정의 대화 상자에서 또는 DATE 명령으로 작성됩니다.

날짜/시간으로 정의된 관측값

관측값이 날짜, 시간 또는 날짜시간 형식의 필드에 의해, 또는 날짜/시간을 나타내는 문자열 필드에 의해 정의됨을 지정할 수 있습니다. 문자열 필드는 YYYY-MM-DD 형식의 날짜, HH:MM:SS 형식의 시간 또는 YYYY-MM-DD HH:MM:SS 형식의 날짜시간으로 표시될 수 있습니다. 앞에 있는 0은 문자열 표현에서 생략할 수 있습니다. 예를 들어, 문자열 2014-9-01은 2014-09-01과 같습니다.

관측값을 정의하는 필드 외에, 관측값을 설명하는 적절한 시간 구간을 선택하십시오. 지정된 시간 구간에 따라, 관측값(증분) 사이의 구간이나 주당 일 수와 같은 다른 설정을 지정할 수도 있습니다. 시간 간격에는 다음 고려사항이 적용됩니다.

- 관측값이 시간에서 비정규적으로 간격이 있는 경우(판매 순서가 처리되는 시간과 같이), **비정규** 값을 사용하십시오. **비정규**가 선택될 때, 데이터 지정 사항 탭의 **시간 간격** 설정에서 분석에 사용되는 시간 구간을 지정해야 합니다.
- 관측값이 날짜와 시간을 나타내고 시간 구간이 시, 분 또는 초인 경우 **하루 중 시간(시)**, **하루 중 시간(분)** 또는 **하루 중 시간(초)**을 사용하십시오. 관측값이 날짜에 대한 참조가 없는 시간(기간)을 나타내고 시간 간격이 시, 분 또는 초인 경우 **시간(비주기적)**, **분(비주기적)** 또는 **초(비주기적)**을 사용하십시오.
- 선택된 시간 간격을 기초로, 프로시저는 결측 관측값을 발견할 수 있습니다. 프로시저에서는 모든 관측값이 시간에서 동일하게 간격을 두고 결측 관측값이 없다고 가정하므로, 결측 관측값을 발견해야 합니다. 예를 들어, 시간 구간이 일(Days)이고 날짜 2014-10-27 뒤에 2014-10-29가 있는 경우, 2014-10-28에 대해 결측 관측값이 있습니다. 값은 결측 관측값에 대해 대체됩니다. 결측값 처리에 대한 설정은 데이터 지정 사항 탭으로부터 지정할 수 있습니다.
- 지정된 시간 간격은 프로시저가 함께 통합해야 하는 동일한 시간 간격의 여러 관측값을 발견하고 관측값에 동일하게 간격이 있도록 월의 첫 번째와 같은 구간 경계에 관측값을 맞출 수 있도록 합니다. 예를 들어 시간 간격이 월인 경우 동일한 월에서 여러 개의 날짜가 함께 통합됩니다. 이 유형의 통합을 그룹화라고 합니다. 기본적으로, 관측값은 그룹화될 때 합산됩니다. 데이터 지정 사항 탭의 **통합 및 분포** 설정에서, 그룹화에 다른 방법(예: 관측값의 평균)을 지정할 수 있습니다.
- 일부 시간 구간의 경우, 추가 설정은 동일하게 간격이 있는 정규 구간에서 중단을 설정할 수 있습니다. 예를 들어, 시간 구간이 일(Days)이지만 평일만 유효한 경우, 주에 5일이 있고 주는 월요일에 시작함을 지정할 수 있습니다.

관측값이 주기 또는 순환 주기로 정의됨

관측값은 임의의 순환 수준 수까지, 주기 또는 반복 주기 순환을 나타내는 하나 이상의 정수 필드로 정의할 수 있습니다. 이 구조에서, 표준 시간 구간 중 하나에 맞지 않는 관측값 계열을 설정할 수 있습니다. 예를 들어, 회계 연도에 10개월만 있는 경우 월을 나타내는 주기 필드와 연도를 나타내는 순환 필드로 설명할 수 있으며 한 순환 길이가 10입니다.

순환 주기를 지정하는 필드는 주기적 수준의 계층 구조를 정의합니다. 가장 낮은 수준은 주기 필드에 의해 정의됩니다. 다음 최상위 수준은 수준이 1인 순환 필드에 의해 지정되고, 그 다음은 수준 2의 순환 필드로 지정되며 뒤로도 마찬가지로입니다. 가장 높은 수준을 제외하고, 각 수준의 필드 값은 다음 최상위 수준에 관하여 주기적이어야 합니다. 최상위 수준의 값은 주기적이 될 수 없습니다. 예를 들어 10개월의 회계 연도의 경우 월은 연도에서 주기적이며 연도가 주기적이지 않습니다.

- 특정 수준의 순환 길이는 다음 최하위 수준의 주기입니다. 회계연도 예의 경우, 단 하나 순환 수준이 있고 순환 길이는 10입니다. 다음으로 가장 낮은 수준이 월을 나타내고 지정된 회계 연도에 10달이 있기 때문입니다.
- 1에서 시작하지 않는 주기적 필드의 시작 값을 지정하십시오. 이 설정은 누락된 값을 발견하는 데 필요합니다. 예를 들어, 주기적 필드는 2에서 시작하지만 시작 값은 1로 지정되는 경우, 프로시저는 해당 필드의 각 순환에 있는 첫 번째 주기에 대해 결측값이 있다고 가정합니다.

레코드 순서에 의해 정의되는 관측값

열 기반 데이터의 경우 첫 번째 레코드가 첫 번째 관측값을 나타내고 두 번째 레코드가 두 번째 관측값을 나타내는 방식으로 관측값을 레코드 순서로 정의하도록 지정할 수 있습니다. 그런 다음 레코드가 시간 내에서 동일한 간격인 관측값을 나타낸다고 가정합니다.

분석에 대한 시간 간격

분석에 사용되는 시간 간격은 관측값의 시간 간격과 다를 수 있습니다. 예를 들어 관측값의 시간 간격이 일인 경우 분석의 시간 간격에 월을 선택할 수 있습니다. 데이터는 모델이 작성되기 전에 매일 데이터에서 매월 데이터까지 통합됩니다. 또한 더 긴 구간에서 더 짧은 시간 간격으로 데이터를 분배하도록 선택할 수도 있습니다. 예를 들어 관측값의 시간 간격이 분기별인 경우 분기별에서 월별 데이터로 데이터를 분배할 수 있습니다.

분석이 행해지는 시간 간격에 대해 사용 가능한 선택은 해당 관측값 정의 방법과 관측값의 시간 간격에 따라 다릅니다. 특히, 관측값이 순환 주기로 정의되거나 날짜 지정 사항이 활성 데이터 세트에 대해 정의되는 경우, 통합만 지원됩니다. 그러한 경우, 분석의 시간 구간은 관측값의 시간 구간보다 크거나 같아야 합니다.

분석 시간 간격은 데이터 지정 사항 탭의 **시간 간격** 설정에서 지정됩니다. 데이터가 통합되거나 분포되는 방법은 데이터 지정 사항 탭의 **통합 및 분포** 설정에서 지정됩니다.

통합 및 분포

집계 함수

분석에 사용되는 시간 구간이 관측에 사용되는 시간 구간보다 길 경우, 입력 데이터는 통합됩니다. 예를 들어 관측값의 시간 간격이 일이고 분석의 시간 간격이 월인 경우 통합이 수행됩니다. mean, sum, mode, min 또는 max 집계 함수를 사용할 수 있습니다.

분포 함수

분석에 사용되는 시간 간격이 관측의 시간 간격보다 짧은 경우, 입력 데이터는 분포됩니다. 예를 들어, 관측값의 시간 간격이 분기이고 분석의 시간 간격이 월일 경우 분포가 수행됩니다. mean 또는 sum 분포 함수를 사용할 수 있습니다.

그룹화 함수

그룹화는 관측값이 날짜/시간에 의해 정의되고 여러 관측값이 동일 시간 간격에 발생하는 경우에 적용됩니다. 예를 들어, 관측값의 시간 간격이 월일 경우, 동일 월에 있는 여러 날짜가 그룹화되어 날짜가 발생하는 일과 연관됩니다. mean, sum, mode, min 또는 max와 같은 그룹화 함수를 사용할 수 있습니다. 다음 그룹화는 항상 관측값이 날짜/시간에 의해 정의되고 관측값의 시간 간격이 비정규로 지정된 경우에 수행됩니다

참고: 그룹화가 통합 양식이어도, 그룹화는 결측값 처리 이전에 수행됩니다(정상 통합은 결측값 처리 이후에 수행됩니다). 관측값의 시간 간격이 비정규로 지정되는 경우, 통합은 그룹화 함수로만 수행됩니다.

교차-일 관측값을 이전 일로 통합

1일 경계를 교차하는 시간을 사용하는 관측값이 전날의 값에 통합되는지 여부를 지정합니다. 예를 들어, 20:00시에 시작하는 8시간 노동의 시간별 관측값의 경우, 이 설정은 00:00 및 04:00 사이의 관측값이 전날 통합 결과에 포함되는지 여부를 지정합니다. 이 설정은 관측값의 시간 구간이 하루 중 시간(시), 하루 중 시간(분) 또는 하루 중 시간(초)이고 분석의 시간 구간이 일(Days)인 경우에만 적용됩니다.

지정된 필드에 대한 사용자 정의 설정

필드 기준으로 필드에 통합, 분포 및 그룹화 함수를 지정할 수 있습니다. 이 설정은 통합, 분포 및 그룹화 함수에 대한 기본 설정을 대체합니다.

결측값

입력 데이터의 결측값은 대체된 값으로 바뀝니다. 다음 방법으로 바꿀 수 있습니다.

선형 보간법

선형 보간법을 사용하여 결측값을 바꿉니다. 결측값 이전의 마지막 유효한 값과 결측값 이후의 첫 번째 유효한 값이 보간법에 사용됩니다. 계열에서 첫 번째 또는 마지막 관측값에 결측값이 있는 경우, 계열의 시작 또는 종료에서 두 개의 가장 근접한 비결측 값이 사용됩니다.

계열 평균

결측값을 전체 계열에 대한 평균으로 바꿉니다.

근접한 값들의 평균

결측값을 유효한 근접 값의 평균으로 바꿉니다. 근접한 값들의 계산너비는 평균을 계산하는데 사용되는 결측값 전후의 유효값 수입니다.

근접한 값들의 중앙값

결측값을 근접한 유효값의 중위수로 바꿉니다. 근접한 값들의 계산너비는 평균을 계산하는데 사용되는 결측값 전후의 유효값 수입니다.

선형 추세

이 옵션은 단순 선형 회귀분석 모델을 적합시키기 위해 계열에서 모든 비결측 관측값을 사용합니다. 이 모델은 결측값을 대체하기 위해 사용됩니다.

기타 설정:

결측값의 최대 퍼센트(%)

어떤 계열에 대해서도 허용되는 최대 결측값 퍼센트를 지정합니다. 지정된 최대값보다 결측값이 더 많은 계열은 분석에서 제외됩니다.

사용자 결측값

이 옵션은 사용자 결측값을 유효 데이터로 처리되어 계열에 포함되는지 여부를 지정합니다. 기본적으로, 사용자 결측값은 제외되고, 시스템 결측값과 같이 처리됩니다. 그리고 나서 이 값은 대체됩니다.

일반 데이터 옵션

차원 필드당 고유 값의 최대 수(M)

이 설정은 다차원 데이터에 적용되며 하나의 차원 필드에 대해 허용되는 최대 고유 값 수를 지정합니다. 기본적으로 이 한계는 10000으로 설정되지만 임의로 큰 수로 늘릴 수 있습니다.

일반 작성 옵션

신뢰구간 너비(%)

이 설정은 예측 및 모델 모수 둘 다의 신뢰구간을 제어합니다. 100보다 작은 양수 값을 지정할 수 있습니다. 기본적으로, 95% 신뢰구간이 사용됩니다.

각 목표에 대한 최대 입력 수

이 설정은 각 목표에 대한 모델에서 허용되는 최대 입력 수를 지정합니다. 1 - 20 사이의 정수를 지정할 수 있습니다. 각 목표의 모형은 항상 자체 지연 값을 포함하므로 이 값을 1로 설정하면 입력만이 목표가 됨을 지정합니다.

모델 허용 한도

이 설정은 각 목표에 대한 최상의 입력 세트를 판별하기 위해 사용되는 반복 프로세스를 제어합니다. 0값을 초과하는 값을 지정할 수 있습니다. 기본값은 0.001입니다.

이상치 임계치(%)

모델에서 계산된 확률(이상치)이 이 임계값을 초과하는 경우 관측값은 이상치로 플래그가 붙습니다. 50 - 100 범위의 값을 지정할 수 있습니다.

각 입력에 대한 시차 수

이 설정은 각 목표에 대한 모형에서 각 입력의 시차 항 수를 지정합니다. 기본적으로, 시차 항 수는 분석에서 사용되는 시간 간격에서 자동으로 결정됩니다. 예를 들어, 시간 간격이 개월(1개월씩 증가)인 경우 시차 수는 12입니다. 필요한 경우 시차 수를 명시적으로 지정할 수 있습니다. 지정된 값은 1 - 20 범위의 정수여야 합니다.

표시할 계열

이 옵션은 출력이 표시되는 계열(목표 또는 입력)을 지정합니다. 지정된 계열에 대한 출력의 콘텐츠는 **출력 옵션** 설정으로 판별됩니다.

최적 적합 모형과 연관된 목표 표시

기본적으로 결과는 R 제곱 값에 의해 결정된 대로 10개의 최적 적합 모형과 연관되는 목표에 대해 표시됩니다. 최적 적합 모형의 다른 고정 숫자를 지정하거나 최적 적합 모형의 백분율을 지정할 수 있습니다. 다음 적합도 측도에서 선택할 수도 있습니다.

R 제곱

선형 모형의 적합도 측도이며, 때때로 결정계수라고 합니다. 이 항목은 모형으로 설명한 목표변수의 변동 비율이 됩니다. 값의 범위는 0부터 1까지입니다. 작은 값은 모형이 데이터에 적합하지 않음을 나타냅니다.

제공된 평균 제공 퍼센트 오차(G)

모형 예측값과 계열 관측값의 차이에 대한 측도입니다. 사용된 단위와 상관이 없으므로 다른 단위의 계열을 비교하는 데 사용할 수 있습니다.

평균 제공된 오차

평균 제공 오차의 제공된 값입니다. 종속 계열이 모형 예측 수준과 얼마나 다른지에 대한 측도로서, 종속 계열과 같은 단위로 표시됩니다.

BIC

Bayesian 정보 기준. -2 감소된 로그 우도에 기반하여 모형을 선택하고 비교하기 위한 측도입니다. 값이 작을수록 모형이 우수함을 나타냅니다. BIC도 초과 모수화된 모형(예: 입력이 많은 복잡한 모형)에 "페널티"를 부여하지만 AIC보다 더 엄격하게 부여합니다.

AIC

Akaike 정보 기준. -2 감소된 로그 우도에 기반하여 모형을 선택하고 비교하기 위한 측도입니다. 값이 작을수록 모형이 우수함을 나타냅니다. AIC는 초과 모수화된 모형(예: 입력이 많은 복잡한 모형)에 "페널티"를 부여합니다.

개별 계열 지정

출력할 개별 계열을 지정할 수 있습니다.

- 열 기반 데이터의 경우 원하는 계열이 포함된 필드를 지정합니다. 지정된 필드의 순서는 출력에 필드가 나타나는 순서를 정의합니다.
- 다차원 데이터 경우 계열이 포함된 매트릭 필드의 격자에 항목을 추가하여 특정한 계열을 지정합니다. 그런 다음 계열을 정의하는 차원 필드의 값을 지정합니다.
 - 각 차원 필드의 값을 눈금에 직접 입력하거나 사용 가능한 차원 값 목록에서 선택할 수 있습니다. 사용 가능한 차원 값 목록에서 선택하려면 원하는 차원 셀에서 생략 기호를 클릭하십시오. 이 동작은 차원 값 선택 하위 대화 상자를 엽니다.
 - 차원 값 선택 하위 대화 상자에서 쌍안경 아이콘을 클릭하고 검색어를 지정하여 차원 값 목록을 검색할 수 있습니다. 공백은 검색어의 일부로 처리됩니다. 검색어의 별표(*)는 와일드카드 문자를 표시하지 않습니다.
 - 눈금에서 계열의 순서는 출력에 나타나는 순서를 정의합니다.

열 기반 데이터 및 다차원 데이터 둘 다에 대해, 출력은 30 계열로 제한됩니다. 이 한계에는 사용자가 지정하는 개별 계열(입력 또는 목표)과 최적 적합 모형과 연관되는 목표가 포함됩니다. 개별적으로 지정된 계열은 최적 적합 모형과 연관된 목표보다 우선순위가 높습니다.

결과 옵션

이러한 옵션은 출력의 콘텐츠를 지정합니다. **목표에 대한 출력결과** 그룹의 옵션은 **표시할 계열** 설정의 최적 적합 모형과 연관되는 목표에 대한 출력을 생성합니다. **계열에 대한 출력결과** 그룹의 옵션은 **표시할 계열** 설정에 지정된 개별 계열에 대한 출력을 생성합니다.

전체 모형 시스템

모델 시스템에서 계열 사이의 인과 관계에 대한 그래픽 표현을 표시합니다. 표시된 목표에 대한 모델 적합 통계 및 이상값 둘 다의 테이블이 출력 항목의 일부로 포함됩니다. **계열에 대한 출력결과** 그룹에서 이 옵션이 선택되면, **표시할 계열** 설정에 지정된 개별 계열마다 별도의 출력 항목이 작성됩니다.

계열 사이의 인과 관계에는 연관된 유의 수준이 있으며, 유의 수준이 적으면 한층 유의적 연결을 표시합니다. 유의수준이 지정된 값보다 더 큰 관계를 숨기도록 선택할 수 있습니다.

모형 적합 통계량 및 이상값

표시를 위해 선택한 목표 계열에 대한 이상치 및 모형 적합 통계의 테이블. 이러한 표에는 전체 모형 시스템 시각화의 표와 같은 정보가 포함되어 있습니다. 이 테이블은 테이블 피벗 및 편집에 대한 모든 표준 기능을 지원합니다.

모형 효과와 모형 모수(O)

표시하도록 선택한 대상 계열의 모형 효과 검정 및 모형 모수 표. 모델 효과 검정에는 모델에 포함된 각 입력에 대한 F 통계량 및 연관된 유의수준이 포함됩니다.

영향 다이어그램

관심이 있는 계열과 이 계열이 영향을 주거나 이 계열에 영향을 주는 다른 계열 사이의 인과 관계의 그래픽 표현을 표시합니다. 관심 계열에 영향을 주는 계열을 원인이라고 합니다. **효과**를 선택하면 효과를 표시하기 위해 초기화된 영향 다이어그램이 생성됩니다. **원인**을 선택하면 원인을 표시하도록 초기화되는 영향 다이어그램이 생성됩니다. **원인 및 효과 모두**를 선택하면 두 개의 개별 영향(으)로 생성됩니다(하나는 원인으로 초기화되고, 하나는 효과로 초기화됩니다). 영향 다이어그램을 표시하는 출력 항목에서 원인 및 효과 사이에 대화형으로 토글할 수 있습니다.

표시할 원인 또는 효과의 수준 수를 지정할 수 있습니다. 첫 번째 수준은 관심 계열입니다. 각각의 추가 수준은 관심 계열의 한층 간접적인 원인 또는 효과를 보여줍니다. 예를 들어 효과 표시의 세 번째 수준은 직접 입력으로 두 번째 수준에 계열이 포함되는 계열로 구성됩니다. 세 번째 수준에 있는 계열은 관심 계열에 의해 간접적으로 영향을 받습니다. 관심 계열은 두 번째 수준에 있는 계열의 직접 입력입니다.

계열 도표

표시를 위해 선택되는 목표 계열의 관측 및 예측 값 도표. 예측이 요청될 때, 도표는 또한 예측에 대한 예측된 값과 신뢰구간을 표시합니다.

잔차 도표

표시를 위해 선택되는 목표 계열에 대한 모형 잔차의 도표.

상위 입력

목표의 상위 3개의 입력과 함께 시간 경과에 따라 표시되는 각 목표의 도표. 상위 입력은 유의성 값이 가장 낮은 입력입니다. 입력 및 목표에 대해 다른 스케일을 사용하기 위해 y축이 각 계열의 z 점수를 나타냅니다.

예측표

표시를 위해 선택되는 목표 계열에 대한 예측된 값 및 해당 예측의 신뢰구간 테이블.

이상값 근본 원인 분석

관심 계열에서 각 이상값의 원인될 가능성이 가장 큰 계열을 판별합니다. **표시할 계열** 설정의 개별 계열 목록에 포함된 각 대상 계열에 대해 이상값 근본 원인 분석이 수행됩니다.

Output

대화형 이상값 테이블 및 차트

각 관심 계열에 대한 이상값과 해당 이상값의 근본 원인 표 및 도표. 테이블에는 이상값마다 하나의 행이 포함됩니다. 차트는 영향 다이어그램입니다. 테이블에서 행을 선택하면 영향 다이어그램에서, 관심 계열에서 연관된 가중값의 가장 큰 원인이 되는 계열까지의 경로가 강조 표시됩니다.

이상값의 피벗 테이블

각 관심 계열에 대한 이상값 및 해당 이상값의 근본 원인 테이블. 이 테이블에는 대화형 화면에 있는 테이블과 동일한 정보가 포함됩니다. 이 표는 피벗 및 편집 표에 대한 모든 표준 기능을 지원합니다.

인과 수준

근본 원인 검색에 포함시킬 수준 수를 지정할 수 있습니다. 여기에서 사용되는 수준의 개념은 영향 다이어그램에 대해 설명된 것과 동일합니다.

모든 모형에 대해 모형 적합

모든 모형 및 선택된 적합 통계량에 대한 모형 적합의 히스토그램. 다음 적합 통계량을 사용할 수 있습니다.

R 제곱

선형 모델의 적합도 측도이며, 때때로 결정계수라고 합니다. 이 항목은 모형으로 설명한 목표변수의 변동 비율이 됩니다. 값의 범위는 0부터 1까지입니다. 작은 값은 모델이 데이터에 적합하지 않음을 나타냅니다.

제공된 평균 제공 퍼센트 오차(G)

모델 예측값과 계열 관측값의 차이에 대한 측도입니다. 사용된 단위와 상관이 없으므로 다른 단위의 계열을 비교하는 데 사용할 수 있습니다.

평균 제공된 오차

평균 제공 오차의 제공된입니다. 종속 계열이 모델 예측 수준과 얼마나 다른지에 대한 측도로서, 종속 계열과 같은 단위로 표시됩니다.

BIC

Bayesian 정보 기준. -2 감소된 로그 우도에 기반하여 모델을 선택하고 비교하기 위한 측도입니다. 값이 작을수록 모형이 우수함을 나타냅니다. BIC도 초과 모수화된 모형(예: 입력이 많은 복잡한 모형)에 "페널티를 부여"하지만 AIC보다 더 엄격하게 부여합니다.

AIC

Akaike 정보 기준. -2 감소된 로그 우도에 기반하여 모델을 선택하고 비교하기 위한 측도입니다. 값이 작을수록 모형이 우수함을 나타냅니다. AIC는 초과 모수화된 모형(예: 입력이 많은 복잡한 모형)에 "페널티를 부여합니다".

시간 경과에 따른 이상값

추정 기간에서 각 시간 구간에 대한 모든 목표에서의 이상값 수 막대형 차트.

계열 변환

모델 시스템에서 계열에 적용된 변환의 테이블. 가능한 변환은 결측값 대체, 통합 및 분배입니다.

추정 기간

기본적으로 추정 기간은 모든 계열에 걸쳐 최초 관측값 시간에 시작되고 최근 관측값 시간에 종료됩니다.

시작 및 종료 시간 기준

추정 기간의 시작 및 종료 시간을 둘 다 지정하거나 시작만 또는 종료만 지정할 수 있습니다. 추정 기간의 시작 또는 종료를 생략하는 경우, 기본값이 사용됩니다.

- 날짜/시간 필드에 의해 관측값이 정의된 경우, 날짜/시간 필드에 사용되는 동일한 형식으로 시작 및 종료 값을 입력하십시오.
- 순환 주기에 의해 정의된 관측값의 경우, 순환 주기 필드마다 값을 지정하십시오. 각 필드는 별도의 열에 표시됩니다.
- 활성 데이터 세트에 대해 적용되는 날짜 지정 사항이 있는 경우, 날짜 지정 사항의 각 구성요소(예: 월)에 값을 지정해야 합니다. 각 성분은 별도의 열에 표시됩니다.
- 관측값이 레코드 순서에 의해 정의되는 경우, 추정 기간의 시작 및 종료는 관련 케이스의 행 번호(데이터 편집기에 표시된 대로)에 의해 정의됩니다.

최근이거나 최초의 시간 간격(L)

선택적 오프셋으로, 데이터의 최초 시간 간격에 시작하거나 최근 시간 간격에 종료하는, 지정된 시간 간격 수로 추정 기간을 정의합니다. 이 컨텍스트에서, 시간 간격은 분석의 시간 간격을 가리킵니다. 예를 들어, 관측값이 매월 단위이지만 분석의 시간 간격은 분기일 수 있습니다. 최근과 시간 구간 수로 24 값을 지정하면 최근 24개 분기를 의미합니다.

선택적으로, 지정된 시간 구간 수를 제외할 수 있습니다. 예를 들어 마지막 24개의 시간 간격을 지정하고 제외할 수를 1로 지정한 경우 이는 추정 기간이 마지막 1개 앞에 있는 24개의 구간으로 구성되었음을 의미합니다.

예측값

레코드를 미래로 확장 옵션은 추정 기간이 끝난 이후에 예측할 시간 구간 수를 설정합니다. 이 경우의 시간 간격은 데이터 지정 사항 탭에 지정된 분석의 시간 간격입니다. 예측이 요청되면 물론 목표가 아닌 입력 계열에서 자기회귀분석 모델이 자동으로 작성됩니다. 해당 모형을 사용하여 예측 기간에서 해당 입력 계열에 대한 값을 생성합니다.

저장

대상 옵션

데이터(결측값의 데이터 통합 또는 대체) 및 새 변수(목표 저장 설정에 지정됨) 둘 다에 대한 변환을 현재 세션의 IBM SPSS Statistics 데이터 파일 또는 새 데이터 세트로 저장할 수 있습니다. 저장된 데이터의 날짜/시간 값은 각 시간 간격의 시작(예: 월의 시작)에 맞추고 모형 시스템의 분석 시간 간격을 표시합니다. 관측값이 날짜 지정 사항 또는 레코드 순서로 정의되고 데이터가 통합되지 않는 경우에만 새 변수를 활성 데이터 세트에 저장할 수 있습니다.

목표 저장

모형 예측, 신뢰구간 및 잔차를 새로운 변수로 저장할 수 있습니다. 저장할 각 지정 목표는 새 변수 세트를 생성하고 각각의 새 변수에는 추정 및 시계열 분석 기간 둘 다에 대한 값이 포함됩니다. 예측값, 신뢰구간 및 잡음 잔차의 경우 새 변수의 접두문자로 사용할 루트 이름을 지정할 수 있습니다. 전체 변수 이름은 대상 계열이 포함된 필드 이름과 루트 이름이 결합된 이름입니다. 루트 이름은 유효한 변수 이름 규칙을 준수해야 합니다. 필요한 경우 변수 이름이 확장되어 변수 이름 충돌 문제를 방지합니다.

예측값을 포함하는 케이스 표시

레코드에 예측값 데이터를 포함할지 여부를 표시하는 변수를 작성합니다. 변수 이름을 지정할 수 있습니다. 기본값은 *ForecastIndicator*입니다.

저장할 목표

새 변수가 모형 시스템의 모든 대상 계열에 대해 작성되거나 표시할 계열 설정에 지정된 대상 계열에 대해서만 작성되는지 여부를 지정합니다.

모형 시스템 내보내기

모형 시스템을 압축된 파일 아카이브(.zip 파일)로 저장합니다. 모델 시스템 파일은 시간 인과 모델 예측 프로시저에서 업데이트된 예측값을 얻거나 사용 가능한 임의의 출력을 생성하는 데 사용될 수 있습니다. 또한 시간 인과 모형 시나리오 프로시저에서 모형 시스템 파일을 사용하여 시나리오 분석을 실행할 수도 있습니다.

대화형 결과

시간 인과 모형화의 결과는 수 많은 상호작용 결과 개체를 포함합니다. 상호작용 기능은 결과 뷰어에서 오브젝트를 활성화(두 번 클릭)하면 사용 가능합니다.

전체 모형 시스템

모형 시스템의 계열 사이에 인과 관계를 표시합니다. 특정 목표를 해당 입력에 연결하는 모든 선의 색상은 동일합니다. 이 선의 두께는 인과 연결의 유의성을 나타냅니다. 선이 두꺼울수록 연결 유의성이 큼니다. 또한 목표가 아닌 입력은 검은색 정사각형으로 표시됩니다.

- 어떤 입력 없이 상위 모형, 지정된 계열, 모든 계열 또는 모형에 대한 관계를 표시할 수 있습니다. 상위 모형은 **표시할 계열** 설정에서 가장 적합한 모형에 지정된 표준을 충족하는 모형입니다.
- 도표에서 계열 이름을 선택하고 이를 마우스 오른쪽 단추로 클릭한 다음 컨텍스트 메뉴에서 **영향 다이어그램 생성**을 선택하여 하나 이상의 계열에 대한 영향 다이어그램을 생성할 수 있습니다.
- 유의수준이 지정된 값보다 더 큰 인과 관계를 숨기도록 선택할 수 있습니다. 유의수준이 더 작은 경우 더 유의적인 인과 관계를 나타냅니다.
- 도표에서 계열 이름을 선택하고 이를 마우스 오른쪽 단추로 클릭한 다음 컨텍스트 메뉴에서 **계열의 관계 강조표시**를 선택하여 특정 계열에 대한 관계를 표시할 수 있습니다.

영향 다이어그램

관심이 있는 계열과 이 계열이 영향을 주거나 이 계열에 영향을 주는 다른 계열 사이의 인과 관계의 그래픽 표현을 표시합니다. 관심 계열에 영향을 주는 계열을 원인이라고 합니다.

- 원하는 계열의 이름을 지정하여 관심 있는 계열을 변경할 수 있습니다. 영향 다이어그램에서 노드는 두 번 클릭하면 관심 계열이 해당 노드와 연관되는 계열로 변경됩니다.
- 원인과 효과 사이를 전환하여 표시할 수 있으며 표시할 원인과 효과 레벨의 수를 변경할 수 있습니다.
- 노드를 한 번 클릭하면 해당 노드와 연관된 계열에 대한 세부 순차 다이어그램이 열립니다.

이상값 근본 원인 분석

관심 계열에서 각 이상값의 원인될 가능성이 가장 큰 계열을 판별합니다.

- 이상값 테이블에서 이상값에 대한 행을 선택하여 이상값에 대한 근본 원인을 표시할 수 있습니다. 순차도 표에 이상값의 아이콘을 클릭하여 근본 원인을 표시할 수도 있습니다.

- 노드를 한 번 클릭하면 해당 노드와 연관된 계열에 대한 세부 순차 다이어그램이 열립니다.

전체 모형 품질

특정 적합 통계량에 대한, 모든 모형의 모형 적합 히스토그램. 막대형 차트에서 막대를 클릭하면 선택된 막대와 연관되는 모형만 표시하도록 점도표를 필터링합니다. 계열 이름을 지정하여 점도표에서 특정 목표 계열에 대한 모형을 찾을 수 있습니다.

이상치 분포

추정 기간에서 각 시간 구간에 대한 모든 목표에서의 이상값 수 막대형 차트. 막대형 차트에서 막대를 클릭하면 선택된 막대와 연관되는 이상치만 표시하도록 점도표를 필터링합니다.

참고:

- 시간적 인과 모형화의 대화형 결과를 포함하는 출력 문서를 저장하고 대화형 변수를 유지하려면 출력을 다른 이름으로 저장 대화 상자에서 **출력 문서와 함께 필수 모형정보 저장**이 선택되었는지 확인하십시오.
- 일부 대화형 기능에서는 활성 데이터 세트가 시간 인과 모형 시스템을 작성하기 위해 사용된 데이터여야 합니다.

시간 인과 모형 적용

시간 인과 모형 적용

시간 인과 모형화 프로시저에서 작성된 모형을 적용하는 데 두 가지 프로시저를 사용할 수 있습니다. 두 프로시저 모두 시간 인과 모형화 프로시저의 일부로 저장될 수 있는 모형 시스템 파일을 필요로 합니다.

시간 인과 모형 분석

이 프로시저를 사용하여 모형을 다시 작성하지 않고도 최신 데이터를 사용할 수 있는 계열의 예측값을 얻을 수 있습니다. 또한 시간 인과 모형화 프로시저에서 사용할 수 있는 결과를 생성할 수도 있습니다.

시간 인과 모형 시나리오

이 프로시저를 사용하여 모형 시스템에서 특정 시계열의 특정 값이 이와 인과적으로 관련된 시계열의 예측값에 얼마나 많은 영향을 미치는지를 조사할 수 있습니다.

임시 인과 모형 예측

시간 인과 모형 시계열 분석 프로시저에서는 시간 인과 모형화 프로시저에서 작성된 모형 시스템 파일을 로드하고 이 모형을 활성 데이터 세트에 적용합니다. 이 프로시저를 사용하여 모형을 다시 작성하지 않고도 최신 데이터를 사용할 수 있는 계열의 예측값을 얻을 수 있습니다. 또한 시간 인과 모형화 프로시저에서 사용할 수 있는 결과를 생성할 수도 있습니다.

가정

- 활성 데이터 세트의 데이터 구조인 열 기반 또는 다차원은 모형 시스템이 작성될 때 사용된 것과 동일한 구조여야 합니다. 다차원 데이터의 경우 다차원 필드는 모형 시스템 작성에 사용된 것과 같아야 합니다. 또한 모형 시스템 작성에 사용된 차원 값이 활성 데이터 세트에도 있어야 합니다.
- 모형은 모형 시스템에서 지정된 필드와 동일한 이름의 활성 데이터 세트 필드에 적용됩니다.
- 모형 시스템을 작성할 때 관측값이 정의된 필드는 활성 데이터 세트에도 있어야 합니다. 관측값 사이의 시간 간격은 모형이 작성될 때와 같다고 가정합니다. 관측값이 날짜 지정에 의해 정의된 경우에는 활성 데이터 세트에 동일한 날짜 지정이 있어야 합니다. 날짜 지정은 날짜 정의 대화 상자에서 또는 DATE 명령으로 작성됩니다.
- 분석의 시간 간격과 통합, 분배 및 결측값에 대한 설정이 모형이 작성될 때와 같습니다.

시간 인과 모형 시계열 분석 사용

이 기능에는 통계량 예측 옵션이 필요합니다.

메뉴에서 다음을 선택합니다.

분석 > 예측 > 시간 인과 모형 적용...

1. 모델 시스템 파일의 파일 지정 사항을 입력하거나 **찾아보기**를 클릭한 후 모델 시스템 파일을 선택하십시오. 모형 시스템 파일은 시간 인과 모형화 프로시저에서 작성됩니다.
2. 이 옵션을 클릭하여 모형을 재추정하고, 예측값을 작성하고, 결과를 생성하십시오.
3. **계속**을 클릭합니다.
4. 기존 모형 모수를 사용하거나 활성 데이터 세트의 데이터에서 모형 모수를 재추정할지 여부를 지정하십시오.
5. 예측할 이후의 시간을 지정하거나 예측하지 않도록 지정하십시오.
6. **옵션**을 클릭하여 결과 내용을 지정하십시오.
7. 모형 모수를 재추정할 때 **저장**을 클릭하고 업데이트된 모형 시스템을 외부 파일로 내보내십시오.
8. 프로시저를 실행하려면 **실행**을 클릭하십시오.

모형 모수 및 시계열 분석

모형 파일에서 로드

예측값은 모형 시스템 파일의 모형 모수와 활성 데이터 세트의 데이터(모형 모수를 재추정하지 않음)를 사용하여 생성됩니다. 결과에 표시되고 최적 적합 모형 선택에 사용되는 적합도는 모형 시스템 파일에서 가져옵니다. 적합도는 각 모형이 개발될 때(또는 마지막으로 업데이트되었을 때) 사용된 데이터를 반영합니다. 이 옵션은 모형 시스템 작성에 사용된 데이터에서 예측값 및 결과를 생성하는 데 적합합니다.

데이터에서 재추정

모형 모수는 활성 데이터 세트의 데이터를 사용하여 재추정됩니다. 모형 모수의 재추정은 각 목표의 모형에 포함되는 입력에 영향을 미치지 않습니다. 이 옵션은 원래 추정 기간 이후의 새 데이터를 가지고 있고 업데이트된 데이터를 사용하여 예측값 및 기타 결과를 생성할 경우 적합합니다.

모든 관측값

추정 기간이 가장 빠른 관측값의 시간에 시작하고 모든 계열에서 가장 마지막 관측값의 시간에 종료하도록 지정합니다.

시작 및 종료 시간 기준

추정 기간의 시작 및 종료 시간을 둘 다 지정하거나 시작만 또는 종료만 지정할 수 있습니다. 추정 기간의 시작 또는 종료를 생략하는 경우, 기본값이 사용됩니다.

- 날짜/시간 필드에 의해 관측값이 정의된 경우, 날짜/시간 필드에 사용되는 동일한 형식으로 시작 및 종료 값을 입력하십시오.
- 순환 주기에 의해 정의된 관측값의 경우, 순환 주기 필드마다 값을 지정하십시오. 각 필드는 별도의 열에 표시됩니다.
- 활성 데이터 세트에 대해 적용되는 날짜 지정 사항이 있는 경우, 날짜 지정 사항의 각 구성요소(예: 월)에 값을 지정해야 합니다. 각 성분은 별도의 열에 표시됩니다.
- 관측값이 레코드 순서에 의해 정의되는 경우, 추정 기간의 시작 및 종료는 관련 케이스의 행 번호(데이터 편집기에 표시된 대로)에 의해 정의됩니다.

최근이거나 최초의 시간 간격(L)

선택적 오프셋으로, 데이터의 최초 시간 간격에 시작하거나 최근 시간 간격에 종료하는, 지정된 시간 간격 수로 추정 기간을 정의합니다. 이 컨텍스트에서, 시간 간격은 분석의 시간 간격을 가리킵니다. 예를 들어, 관측값이 매월 단위이지만 분석의 시간 간격은 분기일 수 있습니다. **최근과 시간 구간 수로 24** 값을 지정하면 최근 24개 분기를 의미합니다.

선택적으로, 지정된 시간 구간 수를 제외할 수 있습니다. 예를 들어 마지막 24개의 시간 간격을 지정하고 제외할 수를 1로 지정한 경우 이는 추정 기간이 마지막 1개 앞에 있는 24개의 구간으로 구성되었음을 의미합니다.

레코드를 미래로 확장(X)

예측값에 대한 시간 간격 수를 추정 기간 종료 이후까지 설정합니다. 이 경우 시간 간격은 분석의 시간 간격입니다. 예측이 요청되면 물론 목표가 아닌 입력 계열에서 자기회귀분석 모델이 자동으로 작성됩니다. 해당 모형을 사용하여 예측 기간에서 해당 입력 계열에 대한 값을 생성함으로써 해당 입력의 목표에 대한 예측값을 얻을 수 있습니다.

일반 옵션

신뢰구간 너비(%)

이 설정은 예측 및 모델 모수 둘 다의 신뢰구간을 제어합니다. 100보다 작은 양수 값을 지정할 수 있습니다. 기본적으로, 95% 신뢰구간이 사용됩니다.

이상치 임계치(%)

모델에서 계산된 확률(이상치)이 이 임계값을 초과하는 경우 관측값은 이상치로 플래그가 붙습니다. 50 - 100 범위의 값을 지정할 수 있습니다.

표시할 계열

이 옵션은 출력이 표시되는 계열(목표 또는 입력)을 지정합니다. 지정된 계열에 대한 출력의 콘텐츠는 **출력 옵션** 설정으로 판별됩니다.

최적 적합 모형과 연관된 목표 표시

기본적으로 결과는 R 제공 값에 의해 결정된 대로 10개의 최적 적합 모형과 연관되는 목표에 대해 표시됩니다. 최적 적합 모형의 다른 고정 숫자를 지정하거나 최적 적합 모형의 백분율을 지정할 수 있습니다. 다음 적합도 측도에서 선택할 수도 있습니다.

R 제공

선형 모형의 적합도 측도이며, 때때로 결정계수라고 합니다. 이 항목은 모형으로 설명한 목표변수의 변동 비율이 됩니다. 값의 범위는 0부터 1까지입니다. 작은 값은 모형이 데이터에 적합하지 않음을 나타냅니다.

제공된 평균 제공 퍼센트 오차(G)

모델 예측값과 계열 관측값의 차이에 대한 측도입니다. 사용된 단위와 상관이 없으므로 다른 단위의 계열을 비교하는 데 사용할 수 있습니다.

평균 제공된 오차

평균 제공 오차의 제공된 값입니다. 종속 계열이 모델 예측 수준과 얼마나 다른지에 대한 측도로서, 종속 계열과 같은 단위로 표시됩니다.

BIC

Bayesian 정보 기준. -2 감소된 로그 우도에 기반하여 모델을 선택하고 비교하기 위한 측도입니다. 값이 작을수록 모형이 우수함을 나타냅니다. BIC도 초과 모수화된 모형(예: 입력이 많은 복잡한 모형)에 "페널티를 부여"하지만 AIC보다 더 엄격하게 부여합니다.

AIC

Akaike 정보 기준. -2 감소된 로그 우도에 기반하여 모델을 선택하고 비교하기 위한 측도입니다. 값이 작을수록 모형이 우수함을 나타냅니다. AIC는 초과 모수화된 모형(예: 입력이 많은 복잡한 모형)에 "페널티를 부여합니다".

개별 계열 지정

출력할 개별 계열을 지정할 수 있습니다.

- 열 기반 데이터의 경우 원하는 계열이 포함된 필드를 지정합니다. 지정된 필드의 순서는 출력에 필드가 나타나는 순서를 정의합니다.
- 다차원 데이터 경우 계열이 포함된 매트릭 필드의 격자에 항목을 추가하여 특정한 계열을 지정합니다. 그런 다음 계열을 정의하는 차원 필드의 값을 지정합니다.
 - 각 차원 필드의 값을 눈금에 직접 입력하거나 사용 가능한 차원 값 목록에서 선택할 수 있습니다. 사용 가능한 차원 값 목록에서 선택하려면 원하는 차원 셀에서 생략 기호를 클릭하십시오. 이 동작은 차원 값 선택 하위 대화 상자를 엽니다.
 - 차원 값 선택 하위 대화 상자에서 쌍안경 아이콘을 클릭하고 검색어를 지정하여 차원 값 목록을 검색할 수 있습니다. 공백은 검색어의 일부로 처리됩니다. 검색어의 별표(*)는 와일드카드 문자를 표시하지 않습니다.
 - 눈금에서 계열의 순서는 출력에 나타나는 순서를 정의합니다.

열 기반 데이터 및 다차원 데이터 둘 다에 대해, 출력은 30 계열로 제한됩니다. 이 한계에는 사용자가 지정하는 개별 계열(입력 또는 목표)과 최적 적합 모형과 연관되는 목표가 포함됩니다. 개별적으로 지정된 계열은 최적 적합 모형과 연관된 목표보다 우선순위가 높습니다.

결과 옵션

이러한 옵션은 출력의 콘텐츠를 지정합니다. **목표에 대한 출력결과** 그룹의 옵션은 **표시할 계열** 설정의 최적 적합 모델과 연관되는 목표에 대한 출력을 생성합니다. **계열에 대한 출력결과** 그룹의 옵션은 **표시할 계열** 설정에 지정된 개별 계열에 대한 출력을 생성합니다.

전체 모형 시스템

모델 시스템에서 계열 사이의 인과 관계에 대한 그래픽 표현을 표시합니다. 표시된 목표에 대한 모델 적합 통계 및 이상값 둘 다의 테이블이 출력 항목의 일부로 포함됩니다. **계열에 대한 출력결과** 그룹에서 이 옵션이 선택되면, **표시할 계열** 설정에 지정된 개별 계열마다 별도의 출력 항목이 작성됩니다.

계열 사이의 인과 관계에는 연관된 유의 수준이 있으며, 유의 수준이 적으면 한층 유의적 연결을 표시합니다. 유의수준이 지정된 값보다 더 큰 관계를 숨기도록 선택할 수 있습니다.

모형 적합 통계량 및 이상값

표시를 위해 선택한 목표 계열에 대한 이상치 및 모형 적합 통계의 테이블. 이러한 표에는 전체 모형 시스템 시각화의 표와 같은 정보가 포함되어 있습니다. 이 테이블은 테이블 피벗 및 편집에 대한 모든 표준 기능을 지원합니다.

모형 효과와 모형 모수(0)

표시하도록 선택한 대상 계열의 모형 효과 검정 및 모형 모수 표. 모델 효과 검정에는 모델에 포함된 각 입력에 대한 F 통계량 및 연관된 유의수준이 포함됩니다.

영향 다이어그램

관심이 있는 계열과 이 계열이 영향을 주거나 이 계열에 영향을 주는 다른 계열 사이의 인과 관계의 그래픽 표현을 표시합니다. 관심 계열에 영향을 주는 계열을 원인이라고 합니다. **효과**를 선택하면 효과를 표시하기 위해 초기화된 영향 다이어그램이 생성됩니다. **원인**을 선택하면 원인을 표시하도록 초기화되는 영향 다이어그램이 생성됩니다. **원인 및 효과 모두**를 선택하면 두 개의 개별 영향(으)로 생성됩니다(하나는 원인으로 초기화되고, 하나는 효과로 초기화됩니다). 영향 다이어그램을 표시하는 출력 항목에서 원인 및 효과 사이에 대화형으로 토글할 수 있습니다.

표시할 원인 또는 효과의 수준 수를 지정할 수 있습니다. 첫 번째 수준은 관심 계열입니다. 각각의 추가 수준은 관심 계열의 한층 간접적인 원인 또는 효과를 보여줍니다. 예를 들어 효과 표시의 세 번째 수준은 직접 입력으로 두 번째 수준에 계열이 포함되는 계열로 구성됩니다. 세 번째 수준에 있는 계열은 관심 계열에 의해 간접적으로 영향을 받습니다. 관심 계열은 두 번째 수준에 있는 계열의 직접 입력입니다.

계열 도표

표시를 위해 선택되는 목표 계열의 관측 및 예측 값 도표. 예측이 요청될 때, 도표는 또한 예측에 대한 예측된 값과 신뢰구간을 표시합니다.

잔차 도표

표시를 위해 선택되는 목표 계열에 대한 모형 잔차의 도표.

상위 입력

목표의 상위 3개의 입력과 함께 시간 경과에 따라 표시되는 각 목표의 도표. 상위 입력은 유의성 값이 가장 낮은 입력입니다. 입력 및 목표에 대해 다른 스케일을 사용하기 위해 y축이 각 계열의 z 점수를 나타냅니다.

예측표

표시를 위해 선택되는 목표 계열에 대한 예측된 값 및 해당 예측의 신뢰구간 테이블.

이상값 근본 원인 분석

관심 계열에서 각 이상값의 원인될 가능성이 가장 큰 계열을 판별합니다. **표시할 계열** 설정의 개별 계열 목록에 포함된 각 대상 계열에 대해 이상값 근본 원인 분석이 수행됩니다.

Output

대화형 이상값 테이블 및 차트

각 관심 계열에 대한 이상값과 해당 이상값의 근본 원인 표 및 도표. 테이블에는 이상값마다 하나의 행이 포함됩니다. 차트는 영향 다이어그램입니다. 테이블에서 행을 선택하면 영향 다이어그램에서, 관심 계열에서 연관된 가중값의 가장 큰 원인이 되는 계열까지의 경로가 강조 표시됩니다.

이상값의 피벗 테이블

각 관심 계열에 대한 이상값 및 해당 이상값의 근본 원인 테이블. 이 테이블에는 대화형 화면에 있는 테이블과 동일한 정보가 포함됩니다. 이 표는 피벗 및 편집 표에 대한 모든 표준 기능을 지원합니다.

인과 수준

근본 원인 검색에 포함시킬 수준 수를 지정할 수 있습니다. 여기에서 사용되는 수준의 개념은 영향 다이어그램에 대해 설명된 것과 동일합니다.

모든 모형에 대해 모형 적합

모든 모형 및 선택된 적합 통계량에 대한 모형 적합의 히스토그램. 다음 적합 통계량을 사용할 수 있습니다.

R 제곱

선형 모형의 적합도 측도이며, 때때로 결정계수라고 합니다. 이 항목은 모형으로 설명한 목표변수의 변동 비율이 됩니다. 값의 범위는 0부터 1까지입니다. 작은 값은 모형이 데이터에 적합하지 않음을 나타냅니다.

제공된 평균 제공 퍼센트 오차(G)

모형 예측값과 계열 관측값의 차이에 대한 측도입니다. 사용된 단위와 상관이 없으므로 다른 단위의 계열을 비교하는 데 사용할 수 있습니다.

평균 제공된 오차

평균 제공 오차의 제공된 값입니다. 종속 계열이 모형 예측 수준과 얼마나 다른지에 대한 측도로서, 종속 계열과 같은 단위로 표시됩니다.

BIC

Bayesian 정보 기준. -2 감소된 로그 우도에 기반하여 모형을 선택하고 비교하기 위한 측도입니다. 값이 작을수록 모형이 우수함을 나타냅니다. BIC도 초과 모수화된 모형(예: 입력이 많은 복잡한 모형)에 "페널티를 부여"하지만 AIC보다 더 엄격하게 부여합니다.

AIC

Akaike 정보 기준. -2 감소된 로그 우도에 기반하여 모형을 선택하고 비교하기 위한 측도입니다. 값이 작을수록 모형이 우수함을 나타냅니다. AIC는 초과 모수화된 모형(예: 입력이 많은 복잡한 모형)에 "페널티를 부여합니다".

시간 경과에 따른 이상값

추정 기간에서 각 시간 구간에 대한 모든 목표에서의 이상값 수 막대형 차트.

계열 변환

모형 시스템에서 계열에 적용된 변환의 테이블. 가능한 변환은 결측값 대체, 통합 및 분배입니다.

저장

목표 저장

모형 예측, 신뢰구간 및 잔차를 새로운 변수로 저장할 수 있습니다. 저장할 각 지정 목표는 새 변수 세트를 생성하고 각각의 새 변수에는 추정 및 시계열 분석 기간 둘 다에 대한 값이 포함됩니다. 예측값, 신뢰구간 및 잔차의 경우 새 변수의 접두문자로 사용할 루트 이름을 지정할 수 있습니다. 전체 변수 이름은 대상 계열이 포함된 필드 이름과 루트 이름이 결합된 이름입니다. 루트 이름은 유효한 변수 이름 규칙을 준수해야 합니다. 필요한 경우 변수 이름이 확장되어 변수 이름 충돌 문제를 방지합니다.

예측값을 포함하는 케이스 표시

레코드에 예측값 데이터를 포함할지 여부를 표시하는 변수를 작성합니다. 변수 이름을 지정할 수 있습니다. 기본값은 *ForecastIndicator*입니다.

저장할 목표

새 변수가 모형 시스템의 모든 대상 계열에 대해 작성되거나 표시할 계열 설정에 지정된 대상 계열에 대해서만 작성되는지 여부를 지정합니다.

대상 옵션

데이터(결측값의 데이터 통합 또는 대체) 및 새 변수(목표 저장 설정에 지정됨) 둘 다에 대한 변환을 현재 세션의 IBM SPSS Statistics 데이터 파일 또는 새 데이터 세트로 저장할 수 있습니다. 저장된 데이터의 날짜/시간 값은 각 시간 간격의 시작(예: 월의 시작)에 맞추고 모형 시스템의 분석 시간 간격을 표시합니다. 관측값이 날짜 지정 사항 또는 레코드 순서로 정의되고 데이터가 통합되지 않는 경우에만 새 변수를 활성 데이터 세트에 저장할 수 있습니다.

모형 시스템 내보내기

모형 시스템을 압축된 파일 아카이브(.zip 파일)로 저장합니다. 이 프로시저에서는 모형 시스템 파일을 재사용할 수 있습니다. 또한 시간 인과 모형 시나리오 프로시저에서 시나리오 분석을 실행하는 데 사용될 수 있습니다. 이 옵션은 모형 모수가 재추정될 때만 사용할 수 있습니다.

시간 인과 모형 시나리오

시간 인과 모형 시나리오 프로시저는 활성 데이터 세트의 데이터를 사용하여 시간 인과 모형 시스템에 대해 사용자 정의 시나리오를 실행합니다. 시나리오는 원 계열이라는 시계열과 지정된 시간 경과에 따른 해당 계열의 사용 정의 값 세트를 사용하여 정의됩니다. 지정된 값은 루트 계열에 의해 영향을 받는 시계열에 대해 예측을 생성하기 위해 사용됩니다. 이 프로시저에서는 시간 인과 모형화 프로시저에서 작성된 모형 시스템 파일이 필요합니다. 여기서는 활성 데이터 세트가 모형 시스템 파일을 작성하는 데 사용된 것과 동일한 데이터라고 가정합니다.

예

비즈니스 의사결정자는 시간 인과 모형화 프로시저를 사용하여 중요한 성능지표 수에 영향을 미치는 핵심 정량을 발견했습니다. 메트릭은 제어 가능하므로, 의사 결정자는 다음 사분기에서 메트릭에 대한 다양한 값 세트의 효과를 조사하려고 합니다. 조사는 시간 인과 모형 시나리오 프로시저에 모형 시스템 파일을 로드하고 핵심 메트릭에 대한 값 세트를 지정하여 쉽게 수행할 수 있습니다.

시간 인과 모형 시나리오 실행

이 기능에는 통계량 예측 옵션이 필요합니다.

메뉴에서 다음을 선택합니다.

분석 > 예측 > 시간 인과 모형 적용...

1. 모형 시스템 파일의 파일 지정 사항을 입력하거나 **찾아보기**를 클릭한 후 모형 시스템 파일을 선택하십시오. 모형 시스템 파일은 시간 인과 모형화 프로시저에서 작성됩니다.
2. 시나리오를 실행하려면 옵션을 클릭하십시오.
3. **계속**을 클릭합니다.
4. 시나리오 탭(시간 인과 모형 시나리오 대화 상자에 있음)에서 **시나리오 주기 정의**를 클릭하고 시나리오 주기를 지정하십시오.
5. 열 기반 데이터의 경우 **시나리오 추가**를 클릭하여 각 시나리오를 정의하십시오. 다차원 데이터의 경우, **시나리오 추가**를 클릭하여 각각의 개별 시나리오를 정의한 후 **시나리오 그룹 추가**를 클릭하여 각각의 시나리오 그룹을 정의하십시오.
6. **옵션**을 클릭하여 결과 내용을 지정하고 시나리오의 영향을 받는 계열의 너비를 지정하십시오.
7. 프로시저를 실행하려면 **실행**을 클릭하십시오.

시나리오 주기 정의

이 시나리오 주기는 시나리오 실행에 사용되는 값을 지정하는 주기입니다. 추정 기간 종료 전 또는 후에 시작할 수 있습니다. 선택적으로 시나리오 주기의 종료 이후에도 예측하도록 지정할 수 있습니다. 기본적으로 예측값은 시나리오 주기의 종료 시 생성됩니다. 모든 시나리오는 동일한 시나리오 주기와 남은 예측에 대한 지정 사항을 사용합니다.

참고: 예측은 시나리오 주기 시작 후 첫 번째 시간 주기에 시작됩니다. 예를 들어, 시나리오 주기가 2014-11-01에 시작되고 시간 간격이 개월인 경우, 첫 번째 예측값은 2014-12-01입니다.

시작 시간, 종료 시간, 예측범위 지정 시간을 기준으로 지정(C)

- 관측값이 날짜/시간 필드로 정의된 경우 날짜/시간 필드에 사용되는 것과 동일한 형식의 시작, 종료 및 예측범위 값을 입력하십시오. 날짜/시간 필드의 값은 연관된 시간 간격의 시작에 맞춰집니다. 예를 들어 분석의 시간 간격이 월인 경우 10/10/2014 값은 월의 시작인 10/01/2014로 조정됩니다.
- 순환 주기로 정의되는 관측값의 경우 각 순환 주기 필드에 값을 지정하십시오. 각 필드는 별도의 열에 표시됩니다.
- 날짜 지정 사항이 활성 데이터 세트에 영향을 미치는 경우 날짜 지정 사항의 각 성분(예: 월)에 대한 값을 지정해야 합니다. 각 성분은 별도의 열에 표시됩니다.
- 관측값이 레코드 순서로 정의된 경우 시작, 종료 및 예측범위는 관련된 케이스의 행 번호(데이터 편집기에 표시된 대로)로 정의됩니다.

추정 기간 종료와 관련된 시간 간격으로 지정(F)

추정 기간의 종료와 관련된 시간 간격의 수에 대해 시작 및 종료를 정의합니다. 여기서 시간 간격은 분석의 시간 간격입니다. 추정 기간의 종료는 시간 간격 0으로 정의됩니다. 추정 기간 종료 이전의 시간 간격은 음수 값이며 추정 기간 종료 이후의 간격은 양수 값입니다. 또한 시나리오 주기의 종료 이후에도 예측하는 구간 수를 지정할 수도 있습니다. 기본값은 0입니다.

예를 들어 분석의 시간 간격이 월이고 시작 간격을 1로, 종료 간격을 3으로, 종료 이후에도 남아 있는 예측을 1이라고 가정합니다. 그러면 시나리오 기간은 추정 기간의 종료 후 3개월입니다. 예측값은 시나리오 주기의 두 번째 및 세 번째 달과 시나리오 주기 종료 이후의 추가 1개월에 대해 생성됩니다.

시나리오 및 시나리오 그룹 추가

시나리오 탭은 실행할 시나리오를 지정합니다. 시나리오를 정의하려면 **시나리오 기간 정의**를 클릭하여 먼저 시나리오 기간을 정의해야 합니다. 시나리오 및 시나리오 그룹(다차원 데이터에만 적용됨)은 연관된 **시나리오 추가** 또는 **시나리오 그룹 추가** 단추를 클릭하여 작성됩니다. 연관된 격자에서 특정한 시나리오 또는 시나리오 그룹을 선택하여 이를 편집하거나, 해당 사본을 작성하거나, 삭제할 수 있습니다.

열 기반 데이터

격자의 **루트 필드** 열은 시나리오 값으로 대체되는 값이 있는 시계열 필드를 지정합니다. **시나리오 값** 열은 가장 빠른 것으로 최근 순서로 지정된 시나리오 값을 표시합니다. 시나리오 값이 표현식으로 정의된 경우, 열은 표현식을 표시합니다.

다차원 데이터

개별 시나리오

개별 시나리오 눈금의 각 행은 값이 지정된 시나리오 값으로 바뀌는 시계열을 지정합니다. 이 계열은 **원 메트릭** 열에 지정된 필드와 각 차원 필드에 지정된 값의 조합으로 정의됩니다. **시나리오 값** 열의 내용은 열 기반 데이터와 같습니다.

시나리오 그룹

시나리오 그룹은 단일 루트 메트릭 필드를 기반으로 하는 하나의 시나리오 세트와 여러 차원 값 세트를 정의합니다. 지정된 메트릭 필드에 대해 각 차원 값 세트(차원 필드당 하나의 값)는 시계열을 정의합니다. 이러한 시계열 각각에 대해 개별 시나리오가 생성되고 해당 값은 시나리오 값으로 대체됩니다. 시나리오 그룹에 대한 시나리오 값은 표현식으로 지정되며, 표현식은 그룹의 각 시계열에 적용됩니다.

계열 개수 열은 시나리오 그룹과 연관되는 차원 값 세트 수를 표시합니다. 표시된 값은 시나리오 그룹과 연관되는 실제 시계열 수(세트당 하나의 계열)보다 클 수 있습니다. 이 조건은 지정된 차원 값 조합 중 일부가 그룹에 대한 루트 메트릭에 의해 포함된 계열에 해당되지 않는 경우에 발생합니다.

시나리오 그룹의 예로 메트릭 필드 *advertising*과 두 개의 차원 필드 *region* 및 *brand*를 고려하십시오. *region* 및 *brand*의 모든 조합을 포함하고 루트 메트릭으로 *advertising*을 기반으로 하는 시나리오 그룹을 정의할 수 있습니다. 표현식을 *advertising*1.2*로 지정하여 *advertising* 필드와 연관된 각 시계열에 대해 *advertising* 값을 20% 늘린 효과를 조사할 수 있습니다. 네 개의 *region* 값과 두 개의 *brand* 값이 있는 경우, 8개의 시계열이 있으므로 그룹에 의해 8개의 시나리오가 정의됩니다.

시나리오 정의

시나리오 정의를 위한 설정은 데이터가 열 기반 또는 다차원 데이터 여부에 따라 다릅니다.

루트 계열

시나리오에 대한 루트 계열을 지정합니다. 각 시나리오는 단일 루트 계열을 기반으로 합니다. 열 기반 데이터의 경우 원 계열을 정의하는 필드를 선택합니다. 다차원 데이터 경우 계열이 포함된 메트릭 필드의 격자에 항목을 추가하여 원 계열을 지정합니다. 그런 다음 루트 계열을 정의하는 차원 필드의 값을 지정합니다. 다음은 차원 값 지정에 적용됩니다.

- 각 차원 필드의 값을 눈금에 직접 입력하거나 사용 가능한 차원 값 목록에서 선택할 수 있습니다. 사용 가능한 차원 값 목록에서 선택하려면 원하는 차원 셀에서 생략 기호를 클릭하십시오. 이 동작은 차원 값 선택 하위 대화 상자를 엽니다.

- 차원 값 선택 하위 대화 상자에서 쌍안경 아이콘을 클릭하고 검색어를 지정하여 차원 값 목록을 검색할 수 있습니다. 공백은 검색어의 일부로 처리됩니다. 검색어의 별표(*)는 와일드카드 문자를 표시하지 않습니다.

영향 목표 지정

원 계열의 영향을 받는 특정 목표를 알고 해당 목표에 대한 효과만 조사하려는 경우 이 옵션을 사용하십시오. 기본적으로, 루트 계열에 의해 영향을 받는 목표는 자동으로 결정됩니다. 옵션 탭의 설정으로 시나리오에 의해 영향을 받는 계열의 폭을 지정할 수 있습니다.

열 기반 데이터의 경우 원하는 목표를 선택하십시오. 다차원 데이터의 경우, 계열을 포함하는 목표 메트릭 필드에 대한 눈금에 진입을 추가하여 목표 계열을 지정합니다. 기본적으로, 지정된 메트릭 필드에 포함되는 모든 계열이 포함됩니다. 하나 이상의 차원 필드에 대해 포함된 값을 사용자 정의하여 포함된 계열 세트를 사용자 정의할 수 있습니다. 포함된 차원 값을 사용자 정의하려면 원하는 차원에 대한 생략 기호를 클릭하십시오. 이 동작은 차원 값 선택 대화 상자를 엽니다.

계열 개수 열(다차원 데이터에 대한)은 현재 연관된 목표 메트릭에 대해 지정된 차원 값 세트 수를 표시합니다. 표시된 값은 실제 영향을 받는 대상 계열 수(세트당 하나의 계열)보다 더 클 수 있습니다. 이 조건은 지정된 차원 값 조합 중 일부가 연관 목표 메트릭에 의해 포함된 계열에 해당되지 않는 경우에 발생합니다.

시나리오 ID

각 시나리오에는 고유한 식별자가 있어야 합니다. 식별자는 시나리오와 연관된 결과에 표시됩니다. 식별자 값에는 고유성 이외의 제한사항이 없습니다.

원 계열에 대한 시나리오 값 지정(V)

시나리오 주기에서 루트 계열의 명시적 값을 지정하려면 이 옵션을 사용하십시오. 눈금에 나열되는 각 시간 구간에 대한 숫자 값을 지정해야 합니다. 읽기, 예측 또는 **Read\Forecast**를 클릭하여 시나리오 주기에서 각 구간에 대한 루트 계열(실제 또는 예측된)의 값을 확보할 수 있습니다.

원 계열에 대한 시나리오 값의 표현식 지정(Y)

시나리오 주기에서 원 계열의 값을 계산하기 위한 표현식을 정의할 수 있습니다. 직접적으로 표현식을 입력하거나 계산기 단추를 클릭하고 시나리오 값 표현식 작성기로부터 표현식을 작성할 수 있습니다.

- 표현식은 모델 시스템에 목표 또는 입력을 포함할 수 있습니다.
- 시나리오 주기가 기존 데이터를 넘는 경우 표현식에 있는 필드의 예측값에 표현식이 적용됩니다.
- 다차원 데이터의 경우 표현식의 각 필드는 원 메트릭에 대해 지정된 필드와 차원 값으로 정의된 시계열을 지정합니다. 이는 표현식을 평가하는 데 사용되는 시계열입니다.

예로서, 루트 필드가 *advertising*이고 표현식이 *advertising*1.2*인 것으로 가정해 보십시오. 시나리오에 사용되는 *advertising* 값은 기존 값의 20% 증가된 값을 나타냅니다.

참고: 시나리오는 시나리오 탭에서 **시나리오 추가**를 클릭하여 작성됩니다.

시나리오 값 표현식 작성기

시나리오 값 표현식 작성기를 사용하여 단일 시나리오 또는 시나리오 그룹에 대한 시나리오 값을 계산하는 데 사용되는 표현식을 작성합니다. 표현식을 작성하려면 **표현식** 필드에 구성요소를 붙여넣거나 직접 **표현식** 필드에 입력하십시오.

- 표현식은 모델 시스템에 목표 또는 입력을 포함할 수 있습니다.
- 함수 그룹 목록에서 그룹을 선택하고 함수 목록의 함수를 두 번 클릭하여 함수를 붙여넣거나 함수를 선택한 다음, 함수 그룹 목록 옆에 있는 화살표를 클릭할 수 있습니다. 물음표로 표시된 모수를 입력합니다. **모두**라는 레이블이 있는 함수 그룹에서는 사용 가능한 모든 함수 목록을 제공합니다. 대화 상자의 지정된 영역에는 현재 선택한 함수에 대한 간략한 설명이 표시됩니다.
- 문자 상수는 따옴표로 묶어야 합니다.
- 값에 소수점이하자리수가 포함된 경우 소수점 구분자로는 마침표(.)를 사용해야 합니다.

시나리오 값 표현식 작성기에 액세스하려면, 시나리오 정의 또는 시나리오 그룹 정의 대화 상자에서 계산기 단추를 클릭하십시오.

차원 값 선택

다차원 데이터의 경우, 시나리오 또는 시나리오 그룹에 의해 영향을 받는 목표를 정의하는 차원 값을 사용자 정의할 수 있습니다. 또한 시나리오 그룹의 원 계열 세트를 정의하는 차원 값을 사용자 정의할 수도 있습니다.

모든 값

현재 차원 필드의 모든 값이 포함됨을 지정합니다. 이 옵션이 기본값입니다.

값 선택

현재 차원 필드의 값 세트를 지정하려면 이 옵션을 사용하십시오. 선택할 값 세트를 필터링할 수 있습니다. 필터 조건에 충족하는 값은 **매치됨** 탭에 나타나고, 필터 조건과 일치하지 않는 값은 **선택되지 않은 값 목록의 매치하지 않음** 탭에 나타납니다. **모두** 탭은 필터 조건과 관계없이 선택되지 않은 모든 값을 나열합니다.

- 필터를 지정할 때 별표(*)를 사용하여 와일드카드 문자를 표시할 수 있습니다.
- 현재 필터를 지우려면, 표시된 값 필터링 대화 상자에 검색어로 비어 있는 값을 지정하십시오.

영향을 받는 목표의 차원 값을 사용자 정의하려면 다음을 수행하십시오.

1. 시나리오 정의 또는 시나리오 그룹 정의 대화 상자에서, 차원 값을 사용자 정의할 목표 메트릭을 선택하십시오.
2. 사용자 정의하려는 차원의 열에서 줄생략 기호 단추를 클릭하십시오.

시나리오 그룹의 루트 계열에 대한 차원 값을 사용자 정의하려면 다음을 수행하십시오.

1. 시나리오 그룹 정의 대화 상자에서 사용자 정의할 차원에 대한 생략 기호 단추(원 계열 격자에 있음)를 클릭하십시오.

시나리오 그룹 정의

원 계열

시나리오 그룹을 위한 루트 계열의 세트를 지정합니다. 개별 시나리오는 세트의 시계열마다 생성됩니다. 원하는 계열을 포함하는 메트릭 필드에 대한 눈금에 진입을 추가하여 루트 계열을 지정합니다. 그런 다음 세트를 정의하는 차원 필드의 값을 지정합니다. 기본적으로, 지정된 루트 메트릭 필드에 포함되는 모든 계열이 포함됩니다. 하나 이상의 차원 필드에 대해 포함된 값을 사용자 정의하여 포함된 계열 세트를 사용자 정의할 수 있습니다. 포함된 차원 값을 사용자 정의하려면 차원에 대한 생략 기호를 클릭하십시오. 이 동작으로 차원 값 선택 대화상자가 열립니다.

계열 개수 열은 현재 연관 루트 메트릭에 대해 포함되는 차원 값 세트 수를 표시합니다. 표시된 값은 시나리오 그룹의 실제 루트 계열 수(세트당 하나의 계열)보다 클 수 있습니다. 이 조건은 차원 값에 지정된 일부 조합이 원 메트릭에 의해 포함된 계열에 해당하지 않는 경우 발생합니다.

영향 목표 계열 지정

원 계열 세트의 영향을 받는 특정 목표를 알고 해당 목표에 대한 효과만 조사하려는 경우 이 옵션을 사용하십시오. 기본적으로 각 원 계열의 영향 목표는 자동으로 판별됩니다. 옵션 탭의 설정으로 각각의 개별 시나리오에 의해 영향을 받는 계열의 폭을 지정할 수 있습니다.

계열을 포함하는 메트릭 필드에 대한 눈금에 항목을 추가하여 대상 시리즈를 지정합니다. 기본적으로, 지정된 메트릭 필드에 포함되는 모든 계열이 포함됩니다. 하나 이상의 차원 필드에 대해 포함된 값을 사용자 정의하여 포함된 계열 세트를 사용자 정의할 수 있습니다. 포함된 차원 값을 사용자 정의하려면 원하는 차원에 대한 생략 기호를 클릭하십시오. 이 동작은 차원 값 선택 대화 상자를 엽니다.

계열 수 열은 연관된 대상 정량에 대해 현재 지정된 차원 값 세트 수를 표시합니다. 표시된 값은 실제 영향을 받는 대상 계열 수(세트당 하나의 계열)보다 더 클 수 있습니다. 이 조건은 지정된 차원 값 조합 중 일부가 연관 목표 메트릭에 의해 포함된 계열에 해당되지 않는 경우에 발생합니다.

시나리오 ID 접두문자

각 시나리오 그룹에는 고유한 접두문자가 있어야 합니다. 접두문자는 시나리오 그룹의 각 개별 시나리오와 연관된 결과에 표시되는 식별자를 작성하는 데 사용됩니다. 개별 시나리오에 대한 식별자는 접두부이며, 그 뒤에 밑줄, 그 뒤에 원 계열을 식별하는 각 차원 필드의 값이 있습니다. 차원 값은 밑줄로 구분됩니다. 접두문자의 값에 대해 유일성이 아닌 다른 제한사항은 없습니다.

루트 계열에 대한 시나리오 값의 표현식

시나리오 그룹에 대한 시나리오 값은 표현식으로 지정되며, 표현식은 그룹의 각 루트 계열의 값을 계산하기 위해 사용됩니다. 표현식을 직접 입력하거나 계산기 단추를 클릭하여 시나리오 값 표현식 작성기에서 표현식을 작성할 수 있습니다.

- 표현식은 모델 시스템에 목표 또는 입력을 포함할 수 있습니다.
- 시나리오 주기가 기존 데이터를 넘는 경우 표현식에 있는 필드의 예측값에 표현식이 적용됩니다.

- 그룹에서 각 루트 계열에 대해, 표현식의 필드는 루트 계열을 정의하는 해당 필드 및 차원 값에 의해 정의되는 시계열을 지정합니다. 이는 표현식을 평가하는 데 사용되는 시계열입니다. 예를 들어, 루트 계열이 `region='north'` 및 `brand='X'`에 의해 정의되는 경우, 표현식에 사용되는 시계열은 동일한 차원 값에 의해 정의됩니다.

예제로서, 루트 메트릭 필드가 `advertising`이고 두 개의 차원 필드 `region` 및 `brand`가 있다고 가정해 보십시오. 또한 시나리오 그룹에 차원 필드 값의 모든 조합이 있다고 가정합니다. 표현식을 `advertising*1.2`로 지정하여 `advertising` 필드와 연관된 각 시계열에 대해 `advertising` 값을 20% 늘린 효과를 조사할 수 있습니다.

참고: 시나리오 그룹은 다차원 데이터에만 적용되고, 시나리오 탭에서 **시나리오 그룹 추가**를 클릭하여 작성됩니다.

옵션

영향 목표의 최대 수준

영향 받는 목표의 최대 수준 수를 지정합니다. 최대 5개까지의 각 연속 수준은 원 계열에 의해 더 간접적인 영향 목표를 포함합니다. 특히, 첫 번째 수준에는 직접 입력으로 루트 계열을 가지고 있는 목표가 포함됩니다. 두 번째 수준의 목표에는 직접 입력 등으로 지정된 첫 번째 수준인 목표가 있습니다. 이 설정의 값을 증가시키면 계산의 복잡도가 증가하여 성능에 영향을 줄 수 있습니다.

자동 검색된 최대 목표

각 루트 계열에 대해 자동으로 발견되는 영향을 받는 최대 목표 수를 지정합니다. 이 설정의 값을 증가시키면 계산의 복잡도가 증가하여 성능에 영향을 줄 수 있습니다.

영향 다이어그램

각 시나리오의 루트 계열과 영향을 미치는 목표 계열 사이의 인과 관계에 대한 그래픽 표현을 표시합니다. 시나리오 값과 영향을 받은 목표의 예측값 둘 다에 대한 테이블이 출력 항의 일부로 포함됩니다. 이 그래프는 영향 목표의 예측값을 도표로 포함합니다. 영향 다이어그램에서 노드를 한 번 클릭하면 해당 노드와 연관된 계열에 대한 세부 순차 다이어그램이 열립니다. 시나리오마다 별도의 영향 다이어그램이 생성됩니다.

계열 도표

각 시나리오에서 영향을 받는 목표 각각에 대한 예측값 계열 도표가 생성됩니다.

예측 및 시나리오 테이블

각 시나리오에 대한 시나리오 값 및 예측값의 테이블. 이러한 표에는 영향 다이어그램의 표와 같은 정보가 포함되어 있습니다. 이 테이블은 테이블 피벗 및 편집에 대한 모든 표준 기능을 지원합니다.

도표 및 테이블에 신뢰구간 포함

시나리오 예측값에 대한 신뢰구간을 도표 및 표 결과 둘 다에 포함시킬지 여부를 지정합니다.

신뢰구간 너비(%)

이 설정은 시나리오 예측에 대한 신뢰구간을 제어합니다. 100보다 작은 양수 값을 지정할 수 있습니다. 기본적으로, 95% 신뢰구간이 사용됩니다.

적합도 측도

이 절에서는 시계열 모델링에 사용되는 적합도 측도의 정의를 제공합니다.

- **정상 R 제곱(Stationary R-squared).** 모델의 고정 부분을 단순 평균 모델과 비교하는 측도입니다. 추세나 계절 패턴이 있는 경우 보통 R-제곱보다 이 측도를 사용하는 것이 좋습니다. 정상 R-제곱은 음의 무한대에서 1까지의 음수일 수 있습니다. 음수 값은 고려 중인 모델이 기준 모델보다 더 나쁜 것을 의미합니다. 양수 값은 고려 중인 모형이 기준선 모형보다 좋음을 의미합니다.
- **R 제곱(R-squared).** 모델에서 설명하는 계열의 총 변동 비율에 대한 추정값입니다. 이 측도는 정상 계열에 가장 유용합니다. R 제곱은 음의 무한대에서 1까지의 음수일 수 있습니다. 음수 값은 고려 중인 모델이 기준 모델보다 더 나쁜 것을 의미합니다. 양수 값은 고려 중인 모형이 기준선 모형보다 좋음을 의미합니다.
- **RMSE.** 평균 제곱근 오차. 평균 제곱 오차의 제곱근입니다. 종속 계열이 모형 예측 수준과 얼마나 다른지에 대한 측도로서, 종속 계열과 같은 단위로 표시됩니다.
- **MAPE.** 절대 백분율 오차 평균. 종속 계열이 모형 예측 수준에서 얼마나 달라지는지에 대한 측도입니다. 사용된 단위와 상관이 없으므로 다른 단위의 계열을 비교하는 데 사용할 수 있습니다.
- **MAE.** 평균 절대 오차. 계열이 모형 예측 수준에서 얼마나 달라지는지에 대한 측도입니다. MAE는 원래의 계열 단위로 보고됩니다.

- **MaxAPE.** 최대 절대 백분율 오차. 퍼센트로 표현되는 최대 예측 오차입니다. 이 측도는 예측에 대한 최악의 시나리오를 예상하는 데 유용합니다.
- **MaxAE.** 최대 절대 오차. 종속 계열과 같은 단위로 표현되는 최대 예측 오차입니다. MaxAPE와 마찬가지로 예측에 대한 최악의 케이스 시나리오를 예상하는 데 유용합니다. 절대 오차의 최대값과 절대 퍼센트 오차의 최대값은 다른 계열 지점에서 발생할 수 있습니다. 예를 들어 대형 계열 값의 절대 오차는 소형 계열 값의 절대 오차보다 약간 큼니다. 그런 경우, 절대 오차의 최대값은 더 큰 계열 값에서 발생하며 최대 절대 퍼센트 오차는 더 작은 계열 값에서 발생합니다.
- 정규화된 **BIC(Normalized BIC).** 정규화된 BIC(Normalized Bayesian Information Criterion). 모형의 전반적인 적합도에 대한 일반적인 측도로서 모형 복잡성을 설명해 줍니다. 평균 제곱 오차를 기반으로 하는 스코어이며, 모형 내 모수의 수와 계열의 길이에 대한 페널티를 포함합니다. 페널티는 모수가 더 많은 모형의 이점을 제거하게 되므로 동일 계열의 경우 다른 모형끼리 통계를 쉽게 비교할 수 있습니다.

이상치 유형

이 절에서는 시계열 모델링에 사용되는 이상치 유형의 정의를 제공합니다.

- **가법(Additive).** 단일 관측값에 영향을 미치는 이상값입니다. 예를 들어 데이터 코딩 오류는 가법 이상값으로 식별될 수 있습니다.
- **수준 이동(Level shift).** 특정 계열 점에서 시작하여 모든 관측값을 상수만큼 이동하는 이상치입니다. 방법을 변경하면 수준 이동이 발생합니다.
- **혁신적(Innovational).** 특정 계열 지점에서 잡음 항에 추가되는 이상치입니다. 정상 계열의 경우 혁신적 이상값은 몇몇 관측값에 영향을 주는 반면 비정상 계열에서는 특정 계열 지점에서 시작되는 모든 관측값에 영향을 줄 수 있습니다.
- **일시적(Transient).** 영향력이 0까지 지수형으로 감소하는 이상치입니다.
- **계절 가법모형(Seasonal additive).** 특정 관측값 및 하나 이상의 계절 기간에 의해 분리된 모든 후속 관측값에 영향을 주는 이상치입니다. 이와 같은 관측값은 모두 동일한 영향을 받습니다. 계절 가법 이상값은 특정 해의 첫 달인 매1월에 판매가 더 높은 경우에 발생할 수 있습니다.
- **지역 추세(Local trend).** 특정 계열 지점에서 지역 추세를 시작하는 이상치입니다.
- **가법 패치(Additive patch).** 둘 이상의 연속된 가법 이상치의 그룹입니다. 이러한 이상값 유형을 선택하면 이상값 그룹 이외에 각각의 가법 이상값을 확인할 수 있습니다.

ACF/PACF 도표

여기 표시된 도표는 순수 또는 이론적 ARIMA 프로세스 도표입니다. 프로세스를 식별하려면 다음을 참조하십시오.

- 비고정 계열은 0으로 빠르게 감소하지 않고 6개 이상의 시차에 대해 유의미하게 유지되는 ACF가 있습니다. 프로세스를 식별하려면 먼저 이러한 계열이 고정될 때까지 구별해야 합니다.
- 자동 선형회귀 프로세스는 처음 한 개 이상의 PACF 시차에 지수형으로 감소하는 ACF 및 말뚝표시가 있습니다. 말뚝표시 수는 자동 선형회귀 차수를 나타냅니다.
- 이동 평균 프로세스는 처음 한 개 이상의 ACF 시차에 말뚝표시를 가지고 있고 지수형으로 감소하는 PACF를 가지고 있습니다. 말뚝표시 수는 이동 평균 차수를 나타냅니다.
- 혼합(ARMA) 프로세스는 일반적으로 ACF 및 PACF 모두에서 지수 감소를 나타냅니다.

식별 단계에서는 ACF 또는 PACF의 부호나 ACF 또는 PACF가 0으로 기하급수적으로 접근하는 감소 속도에 대해 걱정할 필요가 없습니다. 이러한 부분은 AR 및 MA 계수의 부호와 실제 값에 따라 달라집니다. 일부 경우에는 지수형으로 감소하는 ACF가 번갈아 양의 값과 음의 값을 가집니다.

실제 데이터의 ACF 및 PACF 도표는 여기에 표시된 도표처럼 분명하지 않습니다. 지정된 도표에서 중요한 것을 찾아내는 방법을 알아야 합니다. 식별이 잘못되는 경우에 대비하여 항상 잔차의 ACF 및 PACF를 확인합니다. 다음 내용을 알고 있어야 합니다.

- 계절 프로세스는 계절 시차(다중 계절 주기)에서 이러한 패턴을 표시합니다.
- 유의미하지 않은 값을 0으로 처리할 수 있습니다. 즉, 도표의 신뢰구간 내에 있는 값을 무시할 수 있습니다. 그러나 특히 통계적으로 유의한 값의 패턴이 계속되는 경우에는 이러한 값을 무시할 필요가 없습니다.

- 특별한 자기상관은 우연에 의해서만 통계적으로 유의합니다. 통계적으로 유의한 자기상관이 비교적 높은 시차에서 분리되어 있으며 계절 시차에서 발생하지 않는 경우 무시할 수 있습니다.

ACF 및 PACF 도표에 대한 더 자세한 내용은 ARIMA 분석에 대한 모든 텍스트를 참조하십시오.

표 3. $ARIMA(0,0,1), q>0$

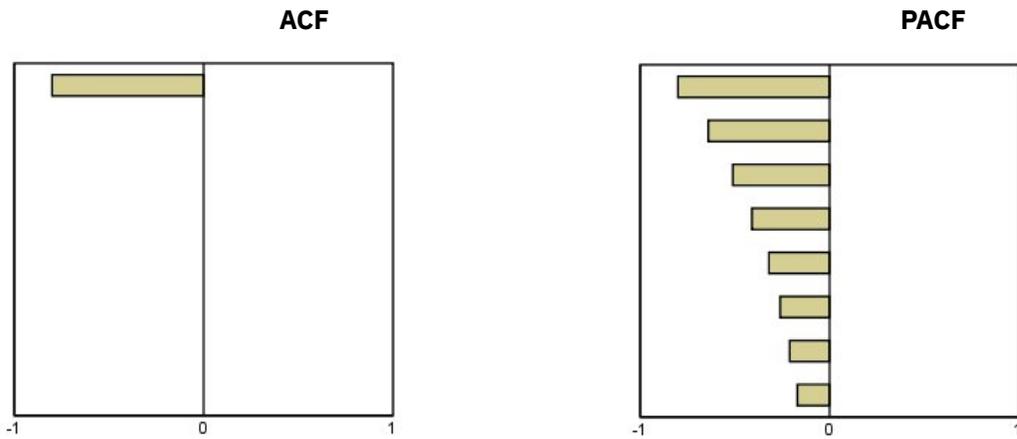
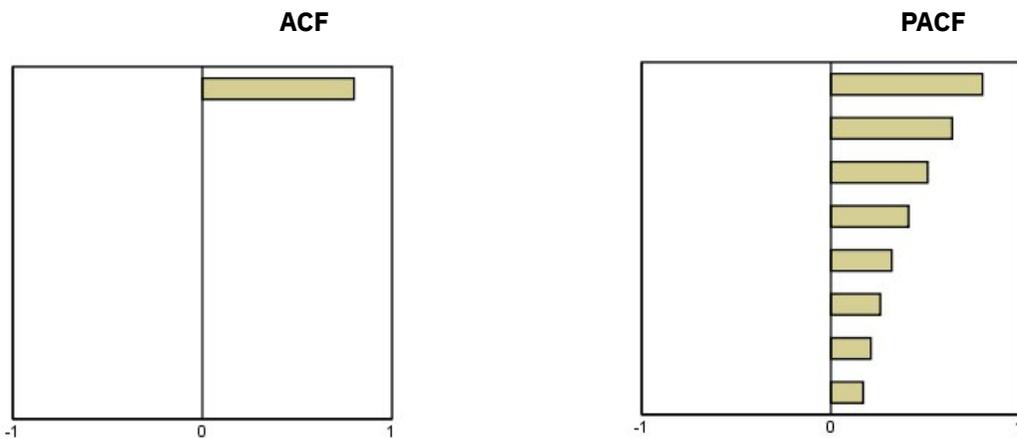
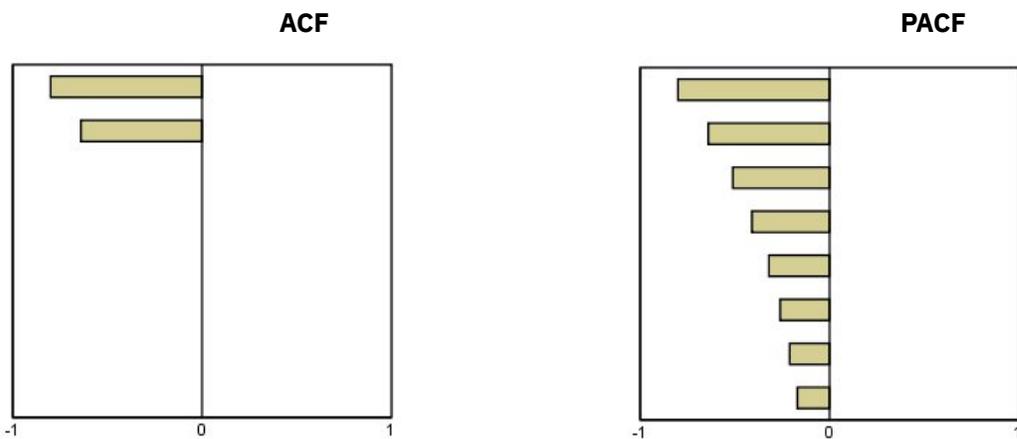


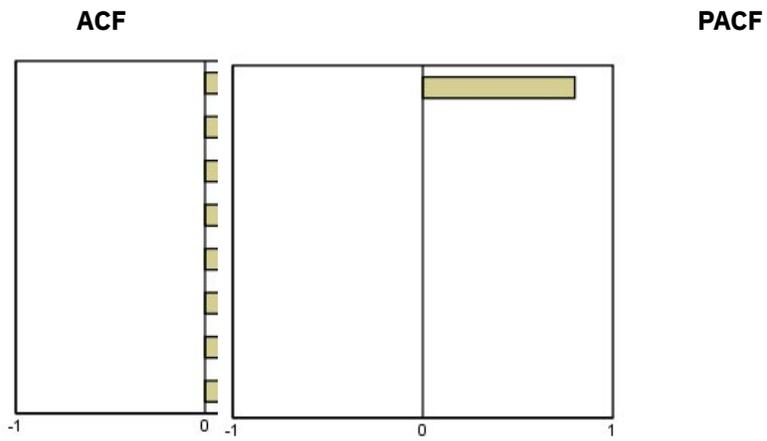
표 4. $ARIMA(0,0,1), q<0$



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



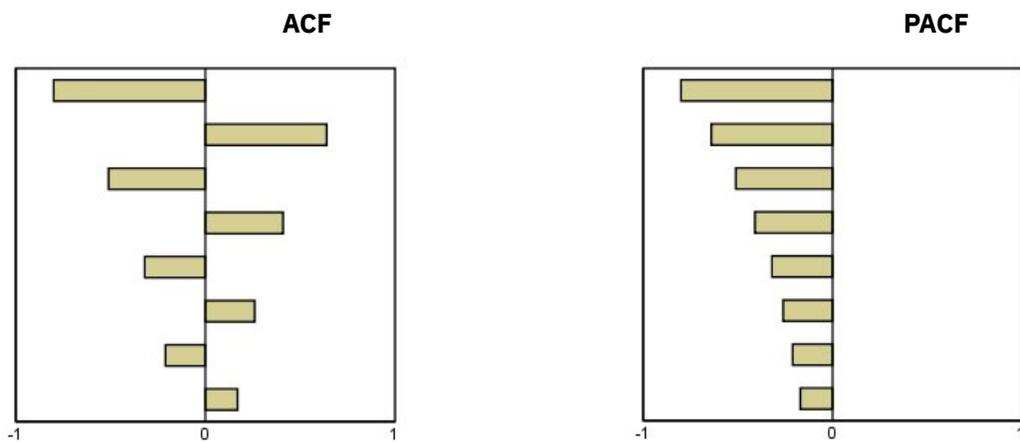
⌘ 5. $ARIMA(1,0,0), f > 0$



⌘ 6. $ARIMA(1,0,0), f < 0$



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



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

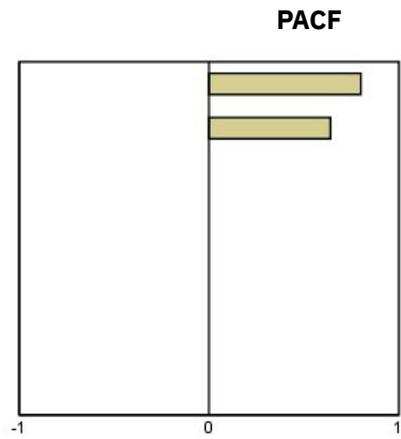
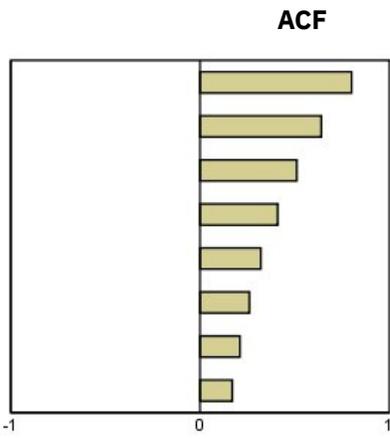
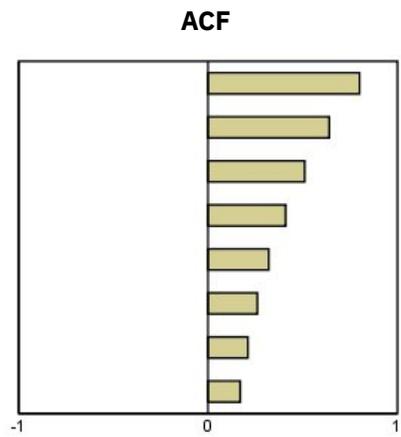


표 7. $ARIMA(0,1,0)$ (통합 계열)



주의사항

이 정보는 미국에서 제공되는 제품과 서비스를 위해 작성되었습니다. 본 자료는 IBM에서 다른 언어로도 제공될 수 있습니다. 그러나 자료에 접근하기 위해서는 해당 언어로 된 제품 또는 제품 버전의 사본이 필요할 수 있습니다.

IBM은 다른 국가에서 이 책에 기술된 제품, 서비스 또는 기능을 제공하지 않을 수도 있습니다. 현재 사용할 수 있는 제품 및 서비스에 대한 정보는 한국 IBM 담당자에게 문의하십시오. IBM 제품, 프로그램 또는 서비스를 언급했다고 해서 해당 IBM 제품, 프로그램 또는 서비스만을 사용할 수 있다는 것을 의미하지는 않습니다. IBM의 지적 재산권을 침해하지 않는 한, 기능상으로 동등한 제품, 프로그램 또는 서비스를 대신 사용할 수도 있습니다. 그러나 비IBM 제품, 프로그램 또는 서비스의 운영에 대한 평가 및 검증은 사용자의 책임입니다.

IBM은 이 책에서 다루고 있는 특정 내용에 대해 특허를 보유하고 있거나 현재 특허 출원 중일 수 있습니다. 이 책을 제공한다고 해서 특허에 대한 라이선스까지 부여하는 것은 아닙니다. 라이선스에 대한 의문사항은 다음으로 문의하십시오.

07326

서울특별시 영등포구 국제금융로 10, 31FC

한국 아이.비.엠 주식회사 한국 아이.비.엠 주식회사KO

2바이트(DBCS) 정보에 관한 라이선스 문의는 한국 IBM에 문의하거나 다음 주소로 서면 문의하시기 바랍니다.

Intellectual Property Licensing

Legal and Intellectual Property Law

IBM Japan Ltd.

19-21, Nihonbashi-Hakozakicho, Chuo-ku

Tokyo 103-8510, Japan

IBM은 타인의 권리 비침해, 상품성 및 특정 목적에의 적합성에 대한 묵시적 보증을 포함하여(단, 이에 한하지 않음) 묵시적이든 명시적이든 어떠한 종류의 보증 없이 이 책을 "현상태대로" 제공합니다. 일부 국가에서는 특정 거래에서 명시적 또는 묵시적 보증의 면책사항을 허용하지 않으므로, 이 사항이 적용되지 않을 수도 있습니다.

이 정보에는 기술적으로 부정확한 내용이나 인쇄상의 오류가 있을 수 있습니다. 이 정보는 주기적으로 변경되며, 변경된 사항은 최신판에 통합됩니다. IBM은 이 책에서 설명한 제품 및/또는 프로그램을 사전 통지 없이 언제든지 개선 및/또는 변경할 수 있습니다.

이 정보에서 언급되는 비IBM 웹 사이트는 단지 편의상 제공된 것으로, 어떤 방식으로든 이들 웹 사이트를 옹호하고자 하는 것은 아닙니다. 해당 웹 사이트의 자료는 본 IBM 제품 자료의 일부가 아니므로 해당 웹 사이트 사용으로 인한 위험은 사용자 본인이 감수해야 합니다.

IBM은 귀하의 권리를 침해하지 않는 범위 내에서 적절하다고 생각하는 방식으로 귀하가 제공한 정보를 사용하거나 배포할 수 있습니다.

(i) 독립적으로 작성된 프로그램과 기타 프로그램(본 프로그램 포함) 간의 정보 교환 및 (ii) 교환된 정보의 상호 이용을 목적으로 본 프로그램에 관한 정보를 얻고자 하는 라이선스 사용자는 다음 주소로 문의하십시오.

07326

서울특별시 영등포구 국제금융로 10, 31FC

서울국제금융센터(31FC) 한국 아이.비.엠 주식회사KO

이러한 정보는 해당 조건(예를 들면, 사용료 지불 등)하에서 사용될 수 있습니다.

이 정보에 기술된 라이선스가 부여된 프로그램 및 프로그램에 대해 사용 가능한 모든 라이선스가 부여된 자료는 IBM이 IBM 기본 계약, IBM 프로그램 라이선스 계약(IPLA) 또는 이와 동등한 계약에 따라 제공한 것입니다.

인용된 성능 데이터와 고객 예제는 예시 용도로만 제공됩니다. 실제 성능 결과는 특정 구성과 운영 조건에 따라 다를 수 있습니다.

비IBM 제품에 관한 정보는 해당 제품의 공급업체, 공개 자료 또는 기타 범용 소스로부터 얻은 것입니다. IBM에서는 이러한 제품들을 테스트하지 않았으므로, 비IBM 제품과 관련된 성능의 정확성, 호환성 또는 기타 청구에 대해서는 확신할 수 없습니다. 비IBM 제품의 성능에 대한 의문사항은 해당 제품의 공급업체에 문의하십시오.

IBM의 미래 방향 또는 의도에 관한 언급은 특별한 통지 없이 변경될 수 있습니다.

이 정보에는 일상의 비즈니스 운영에서 사용되는 자료 및 보고서에 대한 예제가 들어 있습니다. 이들 예제에는 개념을 가능한 완벽하게 설명하기 위하여 개인, 회사, 상표 및 제품의 이름이 사용될 수 있습니다. 이들 이름은 모두 가공의 것이며 실제 인물 또는 기업의 이름과 유사하더라도 이는 전적으로 우연입니다.

저작권 라이선스:

이 정보에는 여러 운영 플랫폼에서의 프로그래밍 기법을 보여주는 원어로 된 샘플 응용프로그램이 들어 있습니다. 귀하는 이러한 샘플 프로그램의 작성 기준이 된 운영 플랫폼의 애플리케이션 프로그래밍 인터페이스(API)에 부합하는 애플리케이션을 개발, 사용, 판매 또는 배포할 목적으로 IBM에 추가 비용을 지불하지 않고 이들 샘플 프로그램을 어떠한 형태로든 복사, 수정 및 배포할 수 있습니다. 이러한 샘플 프로그램은 모든 조건하에서 완전히 테스트된 것은 아닙니다. 따라서 IBM은 이러한 프로그램의 신뢰성, 서비스 가능성 또는 기능을 보증하거나 진술하지 않습니다. 본 샘플 프로그램은 일체의 보증 없이 "현상태대로" 제공됩니다. IBM은 귀하의 샘플 프로그램 사용과 관련되는 손해에 대해 책임을 지지 않습니다.

이러한 샘플 프로그램 또는 파생 제품의 각 사본이나 그 일부에는 반드시 다음과 같은 저작권 표시가 포함되어야 합니다.

© Copyright IBM Corp. 2021. 이 코드의 일부는 IBM Corp. 샘플 프로그램에서 파생되었습니다. 샘플 프로그램

© Copyright IBM Corp. 1989-2021. All rights reserved.

상표

IBM, IBM 로고 및 ibm.com은 전세계 여러 국가에 등록된 International Business Machines Corp.의 상표 또는 등록상표입니다. 기타 제품 및 서비스 이름은 IBM 또는 타사의 상표입니다. 현재 IBM 상표 목록은 웹 "저작권 및 상표 정보"(www.ibm.com/legal/copytrade.shtml)에 있습니다.

Adobe, Adobe 로고, PostScript 및 PostScript 로고는 미국 및/또는 기타 국가에서 사용되는 Adobe Systems Incorporated의 등록상표 또는 상표입니다.

Intel, Intel 로고, Intel Inside, Intel Inside 로고, Intel Centrino, Intel Centrino 로고, Celeron, Intel Xeon, Intel SpeedStep, Itanium 및 Pentium은 미국 또는 기타 국가에서 사용되는 Intel Corporation 또는 그 계열사의 상표 또는 등록상표입니다.

Linux는 미국 또는 기타 국가에서 사용되는 Linus Torvalds의 등록상표입니다.

Microsoft, Windows, Windows NT 및 Windows 로고는 미국 또는 기타 국가에서 사용되는 Microsoft Corporation의 상표입니다.

UNIX는 미국 및 기타 국가에서 사용되는 The Open Group의 등록상표입니다.

Java 및 모든 Java 기반 상표와 로고는 Oracle 및/또는 그 계열사의 상표 또는 등록상표입니다.

색인

A

ACF

- 순수 ARIMA 프로세스 도표 [38](#)
- 시계열 모델러 [8, 9](#)
- 시계열 모형 적용 [13](#)

ARIMA 모델

- 이상치 [7](#)
- 전이 함수 [7](#)

ARIMA 모형 [3](#)

B

Box-Ljung 통계

- 시계열 모델러 [8](#)
- 시계열 모형 적용 [13](#)

Brown's 지수평활 모형 [5](#)

H

Holt's 지수평활 모형 [5](#)

M

MAE

- 시계열 모델러 [8, 9](#)
- 시계열 모형 적용 [13](#)

MAPE

- 시계열 모델러 [8, 9](#)
- 시계열 모형 적용 [13](#)

P

PACF

- 순수 ARIMA 프로세스 도표 [38](#)
- 시계열 모델러 [8, 9](#)
- 시계열 모형 적용 [13](#)

R

R2

- 시계열 모델러 [8, 9](#)
- 시계열 모형 적용 [13](#)

RMSE

- 시계열 모형 적용 [13](#)

W

Winters' 지수평활 모형

- 가법 [5](#)
- 승법 [5](#)

X

XML

XML (계속)

- 재추정된 모형을 XML 형식으로 저장 [15](#)
- XML로 시계열 모형 저장 [10](#)

가

가법 이상치

- 시계열 모델러 [5, 7](#)

가법 패치 이상치

- 시계열 모델러 [5, 7](#)

검증 주기 [2](#)

검증용 케이스 [2](#)

결측값

- 시계열 모델러 [10](#)
- 시계열 모형 적용 [15](#)

계절 가법 이상치

- 시계열 모델러 [5, 7](#)

계절 분해

- 가정 [16](#)
- 모델 [16](#)

변수 작성 [16](#)

새 변수 저장 [16](#)

이동 평균 계산 [16](#)

과거 주기 [2](#)

기록 데이터

- 시계열 모델러 [9](#)
- 시계열 모형 적용 [13](#)

다

단순 계절 지수평활 모형 [5](#)

단순 지수평활 모형 [5](#)

라

로그 변환

- 시계열 모델러 [5-7](#)

로컬 추세 이상치

- 시계열 모델러 [5, 7](#)

마

모델

- 지수평활 [5](#)

모델 매개변수

- 시계열 모델러 [8](#)
- 시계열 모형 적용 [13](#)

모형

- 전문가 모델러 [3](#)
- 지수평활 [3](#)

ARIMA [3](#)

모형 모수 재추정

- 시계열 모형 적용 [11](#)

모형 이름

- 시계열 모델러 [10](#)

바

변수 이름

- 시계열 모델러 [10](#)
- 시계열 모형 적용 [15](#)

사

수준 이동 이상치

- 시계열 모델러 [5, 7](#)

스펙트럼 도표

- 가운데 변환 [17](#)
- 가정 [17](#)
- 스펙트럼 창 [17](#)
- 이변량 스펙트럼 분석 [17](#)

시간 인과 모델

시간 인과 모형

시간 인과 모형 시계열 분석

시간 인과 모형 시나리오

시계열 모델러

- 가장 적합한 또는 가장 적합하지 않은 모형 [9](#)
- 결측값 [10](#)
- 계열 변환 [5-7](#)
- 모델 매개변수 [8](#)
- 모형 이름 [10](#)
- 새로운 변수 이름 [10](#)
- 신뢰구간 [9, 10](#)
- 예측 [8, 9](#)
- 예측 기간 [3, 10](#)
- 예측 저장 [10](#)
- 이벤트 [4](#)
- 이상치 [5, 7](#)
- 잔차 자기상관 함수 [8, 9](#)
- 잔차 편자기상관 함수 [8, 9](#)
- 적합값 [9](#)
- 적합도 통계량 [8, 9](#)
- 전문가 모델러 [3](#)
- 전이 함수 [7](#)
- 전체 모형의 통계 [8, 9](#)
- 주기성 [4-7](#)
- 지수평활 [3, 5](#)
- 추정 기간 [3](#)
- ARIMA [3, 6](#)
- Box-Ljung 통계 [8](#)
- XML로 모형 지정 사항 저장 [10](#)

시계열 모형 적용

- 가장 적합한 또는 가장 적합하지 않은 모형 [14](#)
- 결측값 [15](#)
- 모델 매개변수 [13](#)
- 모형 모수 재추정 [11](#)
- 새로운 변수 이름 [15](#)
- 신뢰구간 [13, 15](#)
- 예측 [13](#)
- 예측 기간 [11](#)
- 예측 저장 [15](#)
- 잔차 자기상관 함수 [13](#)
- 잔차 편자기상관 함수 [13](#)
- 재추정된 모형을 XML 형식으로 저장 [15](#)
- 적합값 [13](#)
- 적합도 통계량 [13](#)
- 전체 모형의 통계 [13](#)
- 추정 기간 [11](#)
- Box-Ljung 통계 [13](#)

시계열 분석

시계열 분석 (계속)

시간 인과 모형

신뢰구간

- 시계열 모델러 [9, 10](#)
- 시계열 모형 적용 [13, 15](#)

아

예측

- 시계열 모델러 [8, 9](#)
- 시계열 모형 적용 [13](#)

예측 기간

- 시계열 모델러 [3, 10](#)
- 시계열 모형 적용 [11](#)

이벤트

- 시계열 모델러 [4](#)

이상치

- 자동 모델 편집기 [5](#)
- 정의 [38](#)
- ARIMA 모델 [7](#)
- 일시적 이상값
- 시계열 모델러 [5, 7](#)

자

자기상관 함수

- 순수 ARIMA 프로세스 도표 [38](#)
- 시계열 모델러 [8, 9](#)
- 시계열 모형 적용 [13](#)

자동 모델 편집기

- 모형 공간 제한 [4](#)
- 이상치 [5](#)

자연 로그 변환

- 시계열 모델러 [5-7](#)

잔차

- 시계열 모델러 [8, 9](#)
- 시계열 모형 적용 [13](#)

저장

- 모형 예측 [10, 15](#)
- 새로운 변수 이름 [10, 15](#)
- XML 형식의 재추정된 모형 [15](#)
- XML의 모형 스펙 [10](#)

적합값

- 시계열 모델러 [9](#)
- 시계열 모형 적용 [13](#)

적합도

- 시계열 모델러 [8, 9](#)
- 시계열 모형 적용 [13](#)
- 정의 [37](#)

전문가 모델러

전이 함수

- 계절 차수 [7](#)
- 분모 차수 [7](#)
- 분자 차수 [7](#)
- 지연 [7](#)
- 차이 차수 [7](#)

절대 오차의 최대값

- 시계열 모델러 [8, 9](#)
- 시계열 모형 적용 [13](#)

정규화된 BIC(베이지안 정보 기준)

- 시계열 모델러 [8, 9](#)
- 시계열 모형 적용 [13](#)

정상 R2

정상 R2 (계속)
시계열 모델러 8, 9
시계열 모형 적용 13
제공근 변환
시계열 모델러 5-7
제공근 평균제공오차(RMSE) 37
조화 분석 17
주기성
시계열 모델러 4-7
지수평활 모형 3, 5
진폭감소 지수평활 모형 5

차

최대 절대 백분율 오차
시계열 모델러 8, 9
시계열 모형 적용 13
최대 절대 오차(MaxAE)
시계열 모델러 8, 9
시계열 모형 적용 13
최대 절대 퍼센트 오차(MaxAPE)
시계열 모델러 8, 9
시계열 모형 적용 13
추정 기간
시계열 모델러 3
시계열 모형 적용 11

파

편자기상관 함수
순수 ARIMA 프로세스 도표 38
시계열 모델러 8, 9
시계열 모형 적용 13
평균 절대 백분율 오차
시계열 모델러 8, 9
시계열 모형 적용 13
평균 절대 오차
시계열 모델러 8, 9
시계열 모형 적용 13
평균 제공 오차 제공근
시계열 모델러 8, 9
시계열 모형 적용 13
평균 제공 오차 제공근(RMSE)
시계열 모형 적용 13
평균 제공근 오차
시계열 모델러 8, 9

하

혁신적 이상값
시계열 모델러 5, 7

