

***IBM SPSS Complex Samples  
24***

**IBM**

참고

이 정보와 이 정보가 지원하는 제품을 사용하기 전에, 반드시 59 페이지의 『주의사항』에 나오는 일반 정보를 읽으십시오.

제품 정보

이 개정판은 새 개정판에서 별도로 명시하지 않는 한, IBM SPSS Statistics의 버전 24, 릴리스 0. 수정사항 0 및 모든 후속 릴리스와 수정에 적용됩니다.

# 목차

제 1 장 복합 표본 프로시저 소개 . . . . .	1
복합 표본의 특성 . . . . .	1
복합 표본 프로시저 사용 . . . . .	2
계획 파일 . . . . .	2
추가 자료 . . . . .	3
제 2 장 복합 디자인에서 표본추출 . . . . .	5
새 표본 계획 작성 . . . . .	5
표본추출 마법사: 계획 변수 . . . . .	5
표본추출 마법사 탐색을 위한 나무 제어 . . . . .	6
표본추출 마법사: 표본추출 방법 . . . . .	6
표본추출 마법사: 표본 결과 . . . . .	7
불균등 크기 정의 . . . . .	8
표본추출 마법사: 결과변수 . . . . .	8
표본추출 마법사: 계획 요약 . . . . .	9
표본추출 마법사: 표본 작성 선택 옵션 . . . . .	9
표본추출 마법사: 표본 작성 결과 파일 . . . . .	9
표본추출 마법사: 완료 . . . . .	10
기존 표본 계획 수정 . . . . .	10
표본추출 마법사: 계획 요약 . . . . .	10
기존 표본 계획 실행 . . . . .	11
CSPLAN 및 CSSELECT 명령 추가 기능 . . . . .	11
제 3 장 분석용 복합 표본 준비 . . . . .	13
새 분석 계획 작성 . . . . .	13
분석 준비 마법사: 계획 변수 . . . . .	13
분석 마법사 탐색을 위한 트리 제어 . . . . .	14
분석 준비 마법사: 추정 방법 . . . . .	14
분석 준비 마법사: 크기 . . . . .	14
불균등 크기 정의 . . . . .	15
분석 준비 마법사: 계획 요약 . . . . .	15
분석 준비 마법사: 완료 . . . . .	16
기존 분석 계획 수정 . . . . .	16
분석 준비 마법사: 계획 요약 . . . . .	16
제 4 장 복합 표본 계획 . . . . .	17
제 5 장 복합 표본 빈도분석 . . . . .	19
복합 표본 빈도분석 통계 . . . . .	19
복합 표본 결측값 . . . . .	20
복합 표본 옵션 . . . . .	21

제 6 장 복합 표본 기술통계 . . . . .	23
복합 표본 기술통계 . . . . .	23
복합 표본 기술통계 결측값 . . . . .	24
복합 표본 옵션 . . . . .	24
제 7 장 복합 표본 교차 분석표 . . . . .	25
복합 표본 교차 분석표 통계 . . . . .	25
복합 표본 결측값 . . . . .	26
복합 표본 옵션 . . . . .	27
제 8 장 복합 표본 비율 . . . . .	29
복합 표본 비율 통계 . . . . .	29
복합 표본 비율 결측값 . . . . .	30
복합 표본 옵션 . . . . .	30
제 9 장 복합 표본 일반선형모형 . . . . .	31
복합 표본 일반선형모형 . . . . .	32
복합 표본 일반선형모형 통계 . . . . .	33
복합 표본 가설 검정 . . . . .	33
복합 표본 일반선형모형 평균 추정 . . . . .	34
복합 표본 일반선형모형 저장 . . . . .	34
복합 표본 일반선형모형 옵션 . . . . .	35
CSGLM 명령 추가 기능 . . . . .	35
제 10 장 복합 표본 로지스틱 회귀분석 . . . . .	37
복합 표본 로지스틱 회귀분석 참조 범주 . . . . .	37
복합 표본 로지스틱 회귀 모형 . . . . .	38
복합 표본 로지스틱 회귀분석 통계 . . . . .	38
복합 표본 가설 검정 . . . . .	39
복합 표본 로지스틱 회귀분석 승산비 . . . . .	40
복합 표본 로지스틱 회귀분석 저장 . . . . .	40
복합 표본 로지스틱 회귀분석 옵션 . . . . .	41
CSLOGISTIC 명령 추가 기능 . . . . .	41
제 11 장 복합 표본 순서 회귀분석 . . . . .	43
복합 표본 순서 회귀분석 반응 확률 . . . . .	44
복합 표본 순서 회귀 모형 . . . . .	44
복합 표본 순서 회귀분석 통계 . . . . .	45
복합 표본 가설 검정 . . . . .	45
복합 표본 순서 회귀분석 승산비 . . . . .	46
복합 표본 순서 회귀분석 저장 . . . . .	46
복합 표본 순서 회귀분석 옵션 . . . . .	47
CSORDINAL 명령 추가 기능 . . . . .	48

제 12 장 복합 표본 Cox 회귀분석 . . . . .	49	저장 . . . . .	54
이벤트 정의 . . . . .	50	내보내기 . . . . .	56
예측변수 . . . . .	51	옵션 . . . . .	57
시간 종속 예측변수 정의 . . . . .	51	CSCOXREG 명령 추가 기능 . . . . .	58
하위 그룹 . . . . .	51	주의사항 . . . . .	59
모형 . . . . .	52	상표 . . . . .	61
통계 . . . . .	53	색인 . . . . .	63
도표 . . . . .	54		
가설 검정 . . . . .	54		

---

## 제 1 장 복합 표본 프로시저 소개

기존 소프트웨어 패키지의 분석 프로시저에서는 데이터 파일의 관측값은 관련 모집단에서 단순 무작위 표본을 대표한다고 가정했습니다. 그러나 이러한 가정은 비용 효율성과 편리성을 고려한 보다 구조화된 표본추출 방법을 찾고 있는 다수의 회사 및 조사 담당자에게는 그리 적합한 방식은 아니었습니다.

복합 표본 옵션을 사용하면 복합 디자인에 따라 표본을 선택하고 디자인 사양을 데이터 분석에 통합시켜 유효한 결과를 얻을 수 있습니다.

---

### 복합 표본의 특성

복합 표본은 단순 무작위 표본과 여러 가지 면에서 다를 수 있습니다. 단순 무작위 표본에서 개별 표본추출 단위는 등확률과 WOR(비복원)로 전체 모집단에서 무작위로 선택됩니다. 반면에 특정 복합 표본에서는 다음 중 일부 또는 전체 기능을 사용할 수 있습니다.

**층화.** 표본추출을 층화하면 모집단의 겹치지 않는 하위 그룹 내에서 표본을 개별적으로 선택할 수 있습니다. 예를 들어, 사회 경제적인 그룹, 직업 범주, 연령 그룹 또는 인종 그룹 등으로 계층화할 수 있습니다. 층화를 사용하면 중요 하위 그룹의 적정 표본 결과를 유지하고 추정값에 대한 전반적인 정확도를 향상시키고 계층별로 다양한 표본추출 방법을 사용할 수 있습니다.

**군집화.** 표본추출을 군집화하면 표본추출 단위의 그룹 또는 군집을 선택할 수 있습니다. 예를 들어, 군집은 학교, 병원 또는 지리 영역일 수 있으며 표본추출 단위는 학생, 환자 또는 시민일 수 있습니다. 군집화는 다단계 디자인 및 지리 영역 표본에서 일반적으로 사용되는 기능입니다.

**다단계.** 다단계 표본추출에서 먼저 군집을 기준으로 첫 단계 표본을 선택한 다음 선택한 군집에서 부표본을 작성하여 두 번째 단계 표본을 만듭니다. 두 번째 단계 표본이 부군집을 기반으로 하는 경우 세 번째 단계 표본을 만들 수 있습니다. 예를 들어, 설문조사에서 먼저 도시의 표본을 작성한 다음, 선택한 도시에서 세대를 표본추출하고, 마지막으로, 선택한 세대에서 개인을 대상으로 설문조사를 실시할 수 있습니다. 표본추출 및 분석 준비 마법사를 사용하면 디자인의 세 단계를 지정할 수 있습니다.

**순차적 표본추출.** 무작위로 선택하기 어려운 경우 고정 구간에서 단위를 계통적으로 또는 순차적으로 표본추출할 수 있습니다.

**부등 선택 확률.** 단위가 균등하게 포함되어 있지 않은 군집을 표본추출하는 경우 PPS(크기 비례) 표본추출을 사용하면 군집의 선택 확률을 해당 단위 비율과 동일하게 만들 수 있습니다. 또한 PPS 표본추출에서 보다 일반적인 가중 방법을 사용하여 단위를 선택할 수 있습니다.

**무제한 표본추출.** 무제한 표본추출에서는 WR(복원)로 단위를 선택합니다. 따라서 특정 개별 단위가 표본에 대해 여러 차례 선택될 수 있습니다.

표본추출 가중치. 표본추출 가중치는 복합 표본을 작성하는 동안 자동으로 계산되며 각 표본추출 단위가 대상 모집단에서 나타내는 "빈도"에 이상적으로 대응합니다. 그러므로 표본에 대한 가중치 합계에 따라 모집단 크기를 추정해야 합니다. 복합 표본 분석 프로시저에서 복합 표본을 제대로 분석하려면 표본추출 가중치가 필요합니다. 이러한 가중치는 모든 복합 표본 옵션에서 사용되어야 하며 가중치를 케이스 복제로 처리하는 가중치 케이스 프로시저를 통해 다른 분석 프로시저와 함께 사용하면 안됩니다.

---

## 복합 표본 프로시저 사용

복합 표본 프로시저 사용은 특정 요구에 따라 달라집니다. 기본 사용자 유형은 다음을 수행합니다.

- 복합 디자인에 따라 설문조사를 계획하고 수행하며 표본은 나중에 분석합니다. 기본 설문조사 도구는 표본 추출 마법사입니다.
- 복합 디자인에 따라 이전에 구한 표본 데이터 파일을 분석합니다. 복합 표본 분석 프로시저를 사용하기 전에 분석 준비 마법사를 사용해야 할 경우도 있습니다.

사용자 유형에 상관 없이 복합 표본 프로시저에 디자인 정보를 제공해야 합니다. 손쉬운 재사용을 위해 이 정보는 계획 파일에 저장됩니다.

## 계획 파일

계획 파일에 복합 표본 지정 사항이 포함됩니다. 계획 파일의 유형은 두 가지입니다.

**표본추출 계획.** 표본추출 마법사에서 생성한 지정 사항에 따라 복합 표본 작성에 사용되는 표본 디자인이 정의됩니다. 표본추출 계획 파일에 이러한 지정 사항이 포함됩니다. 표본추출 계획 파일에는 지정된 표본 디자인에 적합한 추정 방법을 사용하는 기본값 분석 계획도 포함됩니다.

**분석 계획.** 이 계획 파일에는 복합 표본 분석 프로시저에서 복합 표본에 대한 적합한 분산 추정값을 계산하는데 필요한 정보가 들어 있습니다. 계획에는 표본 구조, 각 단계의 추정 방법 및 표본 가중치와 같은 필요한 변수에 대한 참조가 포함됩니다. 분석 준비 마법사를 사용하여 분석 계획을 만들고 편집할 수 있습니다.

지정 사항을 계획 파일에 저장하면 다음과 같은 몇 가지 이점이 있습니다.

- 설문조사자가 다단계 표본추출 계획의 첫 번째 단계를 지정하고 첫 번째 단계의 단위를 작성할 수 있습니다. 또한 두 번째 단계의 표본추출 단위에 대한 정보를 수집한 다음 표본추출 계획을 수정하여 두 번째 단계를 포함시킬 수 있습니다.
- 표본추출 계획 파일에 대한 액세스 권한이 없는 분석가는 분석 계획을 지정하고 각 복합 표본 분석 프로시저에서 해당 계획을 참조할 수 있습니다.
- 계획자가 대용량 공용 표본이 포함된 표본추출 계획 파일을 게시할 수 있으므로 분석가들은 더 간략해진 지침에 따라 작업할 수 있고 각 분석가가 자신만의 분석 계획을 직접 지정하지 않아도 됩니다.

---

## 추가 자료

표본추출 기법에 대한 자세한 내용은 다음 자료를 참조하십시오.

Cochran, W. G. 1977. *Sampling Techniques*, 3rd ed. New York: John Wiley and Sons.

Kish, L. 1965. *Survey Sampling*. New York: John Wiley and Sons.

Kish, L. 1987. *Statistical Design for Research*. New York: John Wiley and Sons.

Murthy, M. N. 1967. *Sampling Theory and Methods*. Calcutta, India: Statistical Publishing Society.

Särndal, C., B. Swensson, and J. Wretman. 1992. *Model Assisted Survey Sampling*. New York: Springer-Verlag.





---

## 제 2 장 복합 디자인에서 표본추출

표본추출 마법사는 표본추출 계획 파일을 작성하거나 수정, 실행하는 과정을 단계별로 안내합니다. 마법사를 사용하려면 잘 정의된 대상 모집단, 표본추출 단위 목록 및 적절한 표본 디자인이 있어야 합니다.

---

### 새 표본 계획 작성

1. 메뉴에서 다음을 선택합니다.

분석 > 복합 표본 > 표본 선택...

2. 표본 디자인을 선택하고 표본 계획을 저장할 계획 파일 이름을 선택합니다.

3. 마법사에서 작업을 계속하려면 다음을 클릭합니다.

4. 선택적으로 계획 변수 단계에서 계층, 군집 및 입력 표본 가중치를 정의할 수 있습니다. 이러한 내용을 정의한 후 다음을 클릭합니다.

5. 선택적으로 표본추출 방법 단계에서 항목 선택 방법을 선택할 수 있습니다.

**PPS Brewer** 또는 **PPS Murthy**를 선택하는 경우 마침을 클릭하여 표본을 작성합니다. 그 밖의 경우 다음을 클릭하고 다음 작업을 수행합니다.

6. 표본 결과 단계에서 표본추출할 단위의 수나 비율을 지정합니다.

7. 이제 마침을 클릭하여 표본을 작성할 수 있습니다.

또는 다음 단계를 추가로 수행할 수 있습니다.

- 저장할 결과변수를 선택합니다.
- 디자인에 두 번째 또는 세 번째 단계를 추가합니다.
- 표본 작성 단계, 난수 시드, 사용자 결측값을 계획 변수의 유효한 값으로 처리할지 여부를 비롯하여 다양한 선택 옵션을 설정합니다.
- 결과 데이터를 저장할 위치를 선택합니다.
- 선택을 명령 구문으로 붙여넣습니다.

---

### 표본추출 마법사: 계획 변수

이 단계에서는 층화변수와 군집 변수를 선택하고 입력 표본 가중치를 정의할 수 있습니다. 단계에 대한 레이블을 지정할 수도 있습니다.

**계층화 기준.** 층화변수의 교차 분류는 고유한 부모집단이나 계층을 정의합니다. 계층별로 별도의 표본이 작성됩니다. 추정값의 정확도를 높이려면 중요 특성에 대해 계층 내의 단위가 가능한 한 동일해야 합니다.

군집 군집 변수는 관측 단위 그룹 또는 군집을 정의합니다. 모집단에서 관측 단위를 직접 표본추출하는 작업이 비용이 많이 들거나 불가능한 경우 모집단에서 군집의 표본을 추출한 다음 선택한 군집에서 관측 단위의 표본을 추출할 수 있습니다. 그러나 군집을 사용하면 표본추출 단위 사이에 상관성이 생기므로 정확도가 떨어질 수 있습니다. 이러한 결과를 최소화하려면 중요 특성에 대해 군집 내의 단위가 가능한 한 동일해야 합니다. 다단계 디자인을 수립하려면 군집 변수를 하나 이상 정의해야 합니다. 여러 가지 다양한 표본추출 방법을 사용하는 경우에도 군집이 필요합니다. 자세한 정보는 『표본추출 마법사: 표본추출 방법』 주제를 참조하십시오.

**입력 표본 가중치.** 현재 표본 디자인이 더 큰 표본 디자인의 일부인 경우 더 큰 디자인에서 이전 단계의 표본 가중치를 사용할 수 있습니다. 현재 디자인의 첫 단계에서 이러한 가중치를 포함하는 숫자변수를 지정할 수 있습니다. 현재 디자인의 후속 단계에 대한 표본 가중치는 자동으로 계산됩니다.

**단계 레이블.** 각 단계에 대해 선택적으로 문자 레이블을 지정할 수 있습니다. 이러한 문자 설명은 단계별 정보 파악에 유용합니다.

**참고:** 소스 변수 목록의 내용은 마법사의 모든 단계에 동일하게 적용됩니다. 즉, 특정 단계의 소스 목록에서 제거된 변수는 모든 단계의 목록에서 제거됩니다. 소스 목록에 반환된 변수는 모든 단계의 목록에 표시됩니다.

## 표본추출 마법사 탐색을 위한 나무 제어

표본추출 마법사에서 각 단계의 왼쪽에는 모든 단계의 개요가 표시됩니다. 개요에서 사용할 수 있는 단계 이름을 클릭하여 마법사를 탐색할 수 있습니다. 이전 단계가 모두 유효하면 즉, 각 이전 단계에 필요한 최소 지정 사항을 입력했으면 해당 단계를 사용할 수 있습니다. 주어진 단계가 유효하지 않은 이유에 대한 자세한 내용은 단계별 도움말을 참조하십시오.

---

## 표본추출 마법사: 표본추출 방법

이 단계에서는 활성 데이터 세트에서 케이스를 선택하는 방법을 지정할 수 있습니다.

**방법.** 여기서 지정한 내용이 선택 방법을 선택하는 데 사용됩니다. 일부 표본추출 유형을 사용하여 WR(복원) 또는 WOR(비복원)로 표본추출할 것인지 선택할 수 있습니다. 자세한 내용은 유형 설명을 참조하십시오. 군집이 정의된 경우에만 일부 PPS(크기 비례) 유형을 사용할 수 있으며 디자인의 첫 단계에서만 모든 PPS 유형을 사용할 수 있습니다. WR 방법은 디자인의 마지막 단계에서만 사용할 수 있습니다.

- **단순 무작위 표본.** 단위가 등확률로 선택됩니다. 복원 또는 비복원으로 선택할 수 있습니다.
- **단순 계통.** 단위는 표본추출 프레임(또는 계층이 지정된 경우는 계층)을 통해 고정 구간에서 선택되며 비복원으로 추출됩니다. 처음 구간 내에서 무작위로 선택한 단위가 시작점으로 설정됩니다.
- **단순 순차.** 단위는 비복원 등확률로 순차적으로 선택됩니다.
- **PPS.** 크기 비례로 단위를 무작위로 선택하는 첫 단계 방법입니다. 복원으로 모든 단위를 선택할 수 있습니다. 군집만 비복원으로 표본추출할 수 있습니다.
- **PPS 계통.** 크기 비례로 단위를 구조적으로 선택하는 첫 단계 방법입니다. 단위는 비복원으로 선택됩니다.
- **PPS 순차.** 군집 크기 비례 및 비복원으로 단위를 순차적으로 선택하는 첫 단계 방법입니다.

- **PPS Brewer.** 군집 크기 비례 및 비복원으로 각 계층에서 군집을 두 개 선택하는 첫 단계 방법입니다. 군집 변수에서 이 방법을 사용하도록 지정해야 합니다.
- **PPS Murthy.** 군집 크기 비례 및 비복원으로 각 계층에서 군집을 두 개 선택하는 첫 단계 방법입니다. 군집 변수에서 이 방법을 사용하도록 지정해야 합니다.
- **PPS Sampford.** 군집 크기 비례 및 비복원으로 각 계층에서 군집을 세 개 이상 선택하는 첫 단계 방법입니다. 이는 더 확장된 기능의 Brewer 방법입니다. 군집 변수에서 이 방법을 사용하도록 지정해야 합니다.
- 분석 시 **WR 추정 사용.** 기본적으로 추정 방법은 선택한 표본추출 방법과 일치하는 계획 파일에서 지정합니다. 따라서 표본추출 방법에 WOR 추정이 있더라도 복원 추정을 사용할 수 있습니다. 이 옵션은 1단계에서만 사용할 수 있습니다.

**크기 측도(MOS).** PPS 방법을 선택한 경우 각 단위의 크기를 정의하는 크기 측도를 지정해야 합니다. 이러한 크기를 변수에서 명시적으로 정의하거나 데이터에서 계산할 수 있습니다. 또는 MOS에 하한과 상한을 설정하여 모든 MOS 변수 값이나 데이터에서 계산된 값을 바꿀 수 있습니다. 이러한 옵션은 1단계에서만 사용할 수 있습니다.

---

## 표본추출 마법사: 표본 결과

이 단계에서는 현재 단계 내에서 표본추출할 단위의 수나 비율을 지정할 수 있습니다. 표본 결과는 고정되거나 계층별로 다를 수 있습니다. 표본 결과 지정을 위해 이전 단계에서 선택한 군집을 계층 정의에 사용할 수 있습니다.

단위. 정확한 표본 결과나 표본추출 단위 비율을 지정할 수 있습니다.

- **값.** 단일값이 모든 계층에 적용됩니다. 단위 메트릭으로 개수를 선택한 경우 양의 정수를 입력해야 합니다. 비율을 선택한 경우에는 음이 아닌 값을 입력해야 합니다. 또한 복원 표본추출이 아닌 경우 비율 값이 1 이하이어야 합니다.
- **계층에 대해 불균등 값.** 불균등 크기 정의 대화 상자를 통해 계층별로 크기 값을 입력할 수 있습니다.
- **값을 읽을 변수.** 계층의 크기 값을 포함하는 숫자변수를 선택할 수 있습니다.

비율을 선택한 경우 표본추출된 단위 수에 대해 상한 및 하한을 설정하는 옵션이 표시됩니다.

## 불균등 크기 정의

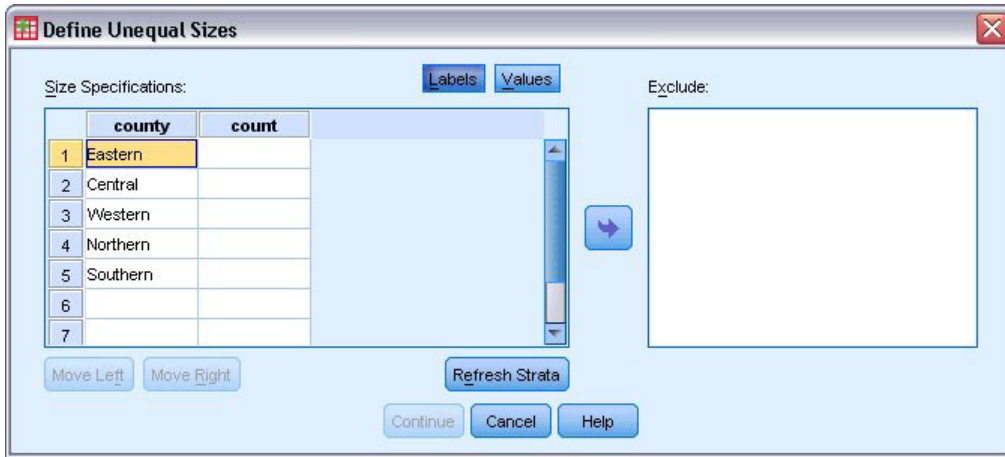


그림 1. 불균등 크기 정의 대화 상자

불균등 크기 정의 대화 상자에서는 계층별로 크기를 입력할 수 있습니다.

**크기 지정 사항 격자.** 이 격자에는 최대 다섯 개의 층화변수 또는 군집 변수, 행마다 계층/군집 조합 하나씩에 대한 교차 분류가 표시됩니다. 적합한 격자 변수에는 현재 및 이전 단계의 모든 층화변수는 물론 이전 단계의 모든 군집 변수가 포함되어 있습니다. 격자 내에서 변수 순서를 바꾸거나 제외 목록으로 이동할 수 있습니다. 맨 오른쪽 열에 크기를 입력합니다. 격자 셀에 층화변수와 군집 변수에 대한 값 레이블로 표시할지 데이터 값으로 표시할지를 전환하려면 레이블 또는 값을 클릭합니다. 레이블을 붙이지 않은 값을 포함하는 셀에는 항상 값이 표시됩니다. 격자를 해당 변수에 대해 레이블이 있는 데이터 값 조합으로 다시 채우려면 계층 최신 정보 표시를 클릭합니다.

**제외.** 계층/군집 조합의 서브세트 크기를 지정하려면 하나 이상의 변수를 제외 목록으로 옮깁니다. 이러한 변수는 표본 결과를 정의하는 데 사용되지 않습니다.

---

## 표본추출 마법사: 결과변수

이 단계에서는 표본 작성 시 저장할 변수를 선택할 수 있습니다.

**모집단 크기.** 주어진 단계의 모집단에서 추정된 단위 수입입니다. 저장된 변수의 루트 이름은 *PopulationSize\_*입니다.

**표본 비율.** 주어진 단계에서의 표본추출 비율입니다. 저장된 변수의 루트 이름은 *SamplingRate\_*입니다.

**표본 결과.** 주어진 단계에서 작성된 단위 수입입니다. 저장된 변수의 루트 이름은 *SampleSize\_*입니다.

**표본 가중치.** 포함 확률의 반대입니다. 저장된 변수의 루트 이름은 *SampleWeight\_*입니다.

일부 단계별 변수는 자동으로 생성됩니다. 이러한 변수는 다음과 같습니다.

**포함 확률.** 주어진 단계에서 작성된 단위 비율입니다. 저장된 변수의 루트 이름은 *InclusionProbability\_*입니다.

**누적 가중치.** 현재 단계를 비롯하여 이전 단계에 대한 누적 표본 가중치입니다. 저장된 변수의 루트 이름은 *SampleWeightCumulative\_*입니다.

**지수.** 주어진 단계 내에서 여러 번 선택한 단위를 식별합니다. 저장된 변수의 루트 이름은 *Index\_*입니다.

**참고:** 저장된 변수 루트 이름에는 단계 번호를 반영하는 정수 접미문자가 포함됩니다. 예를 들어, 1단계의 저장된 모집단 크기의 경우 루트 이름은 *PopulationSize\_1\_*입니다.

---

## 표본추출 마법사: 계획 요약

이 단계는 각 단계의 마지막 단계로 현재 단계를 통해 표본 디자인 사양에 대한 요약을 제공합니다. 여기에서 다음 단계로 진행하거나 필요에 따라 단계를 작성하거나 표본 작성 옵션을 설정할 수 있습니다.

---

## 표본추출 마법사: 표본 작성 선택 옵션

이 단계에서는 표본 작성 여부를 선택할 수 있습니다. 난수 시드 및 결측값 처리와 같은 다른 표본추출 옵션도 제어할 수 있습니다.

**표본 작성.** 표본 작성 여부뿐만 아니라 표본추출 디자인의 일부 실행 여부를 선택할 수 있습니다. 단계는 순서대로 작성되어야 합니다. 즉, 1단계를 작성하지 않은 경우 2단계를 작성할 수 없습니다. 계획을 편집하거나 실행하는 경우 잠긴 단계의 표본을 다시 추출할 수 없습니다.

**시드.** 난수 생성을 위한 시드 값을 선택할 수 있습니다.

**사용자 결측값 포함.** 사용자 결측값의 유효성 여부를 결정합니다. 이 경우 사용자 결측값은 개별 범주로 처리됩니다.

**데이터가 이미 정렬되어 있음.** 층화변수 값을 기준으로 표본 프레임을 미리 정렬한 경우 이 옵션을 사용하면 선택 과정을 보다 신속하게 진행할 수 있습니다.

---

## 표본추출 마법사: 표본 작성 결과 파일

이 단계에서 표본 케이스, 가중 변수, 결합 확률 및 케이스 선택 규칙을 지정할 위치를 선택할 수 있습니다.

**표본 데이터.** 이 옵션을 사용하면 표본 결과를 기록할 위치를 결정할 수 있습니다. 출력결과를 활성 데이터 세트에 추가하거나, 새 데이터 파일에 기록하거나, 외부 IBM® SPSS® Statistics 데이터 파일에 저장할 수 있습니다. 데이터 세트는 현재 세션을 진행하는 동안 사용할 수 있지만 이를 명시적인 데이터 파일로 저장하지 않으면 이후의 세션에서는 사용할 수 없습니다. 데이터 세트 이름은 변수 이름 규칙을 준수해야 합니다. 외부 파일이나 새 데이터 파일을 지정한 경우 선택한 케이스에 대한 활성 데이터 세트의 변수와 표본추출 결과변수가 기록됩니다.

**결합 확률.** 이러한 옵션을 사용하면 결합 확률을 기록할 위치를 결정할 수 있습니다. 이는 외부 IBM SPSS Statistics 데이터 파일에 저장됩니다. PPS WOR, PPS Brewer, PPS Sampford 또는 PSS Murthy 방법을 선택하고 WR 추정을 지정하지 않은 경우 결합 확률이 만들어집니다.

**케이스 선택 규칙.** 한 번에 한 단계씩 표본을 작성하는 경우 케이스 선택 규칙을 텍스트 파일에 저장하면 후속 단계의 하위프레임을 작성할 때 유용합니다.

---

## 표본추출 마법사: 완료

마지막 단계입니다. 계획 파일을 저장하고 표본을 작성하거나 선택을 구문 창에 붙여넣을 수 있습니다.

기존 계획 파일의 단계를 변경하는 경우 편집된 계획을 새 파일에 저장하거나 기존 파일을 덮어쓸 수 있습니다. 기존 단계를 변경하지 않고 단계를 새로 추가하는 경우 마법사를 통해 기존 계획 파일을 자동으로 덮어쓰게 됩니다. 계획을 새 파일에 저장하려면 마법사에서 생성한 명령을 구문 창에 붙여넣기를 선택하고 명령을 사용하여 파일 이름을 변경합니다.

---

## 기존 표본 계획 수정

1. 메뉴에서 다음을 선택합니다.

분석 > 복합 표본 > 표본 선택...

2. 표본 디자인 편집을 선택하고 편집할 계획 파일을 선택합니다.

3. 마법사에서 작업을 계속하려면 다음을 클릭합니다.

4. 계획 요약 단계에서 표본추출 계획을 검토하고 다음을 클릭합니다.

후속 단계는 대부분 새 디자인의 경우와 같습니다. 자세한 내용은 단계별 도움말을 참조하십시오.

5. 마침 단계로 이동하여 편집된 계획 파일에 대해 새 이름을 지정하거나 기존 계획 파일을 덮어쓰도록 선택합니다.

선택적으로 다음을 수행할 수 있습니다.

- 이미 표본추출한 단계를 지정합니다.
- 계획에서 단계를 제거합니다.

---

## 표본추출 마법사: 계획 요약

이 단계에서는 표본추출 계획을 검토하고 이미 표본추출한 단계를 표시할 수 있습니다. 계획을 편집하는 경우 계획에서 단계를 제거할 수도 있습니다.

**이전에 표본추출한 단계.** 확장된 표본추출 프레임을 사용할 수 없는 경우에는 한 번에 한 단계씩 다단계 표본 추출 디자인을 실행해야 합니다. 이미 표본추출한 단계를 드롭다운 목록에서 선택합니다. 실행된 단계는 모두 잠기므로 표본 작성 선택 옵션 단계에서 사용할 수 없으며 계획을 편집하는 경우 수정할 수 없습니다.



단계 제거. 다단계 디자인에서 2단계와 3단계를 제거할 수 있습니다.

---

## 기존 표본 계획 실행

1. 메뉴에서 다음을 선택합니다.

분석 > 복합 표본 > 표본 선택...

2. 표본 작성을 선택하고 실행할 계획 파일을 선택합니다.

3. 마법사에서 작업을 계속하려면 다음을 클릭합니다.

4. 계획 요약 단계에서 표본추출 계획을 검토하고 다음을 클릭합니다.

5. 표본 계획을 실행하는 경우 단계 정보를 포함하는 개별 단계를 건너뛸 수 있습니다. 이제 언제든지 마침 단계로 이동할 수 있습니다.

선택적으로 이미 표본추출한 단계를 지정할 수 있습니다.

---

## CSPLAN 및 CSSELECT 명령 추가 기능

명령 구문을 사용하여 수행할 수 있는 추가 기능은 다음과 같습니다.

- 결과변수의 사용자 정의 이름을 지정할 수 있습니다.
- 뷰어에서 결과를 관리할 수 있습니다. 예를 들어, 표본을 계획하거나 수정하는 경우 표시되는 계획의 단계별 요약은 표시하지 않거나 표본 디자인을 실행하는 경우 표시되는 계층별로 표본추출한 케이스의 분포 요약을 표시하지 않거나 케이스 처리 요약을 요청할 수 있습니다.
- 활성 데이터 세트에서 외부 표본 파일이나 다른 데이터 파일에 기록할 변수 서브세트를 선택할 수 있습니다.

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





---

## 제 3 장 분석용 복합 표본 준비

분석 준비 마법사는 다양한 복합 표본 분석 프로시저에서 사용할 분석 계획을 작성하거나 수정하는 과정을 단계별로 안내합니다. 마법사를 사용하려면 복합 디자인에 따라 작성된 표본이 있어야 합니다.

표본 작성에 사용된 표본추출 계획 파일에 대한 액세스(기본 분석 계획을 포함하는 표본추출 계획 재호출) 권한이 없는 경우 새 계획을 작성하는 것이 가장 좋습니다. 표본 작성에 사용된 표본추출 계획 파일에 대한 액세스 권한이 없으면 표본추출 계획 파일에 포함된 기본 분석 계획을 사용하거나 기본 분석 지정 사항을 비롯한 다음 변경 사항을 새 파일에 저장합니다.

---

### 새 분석 계획 작성

1. 메뉴에서 다음을 선택합니다.

분석 > 복합 표본 > 분석 준비...

2. 계획 파일 작성을 선택하고 분석 계획을 저장할 계획 파일 이름을 선택합니다.
3. 마법사에서 작업을 계속하려면 다음을 클릭합니다.
4. 계획 변수 단계에서 표본 가중치를 포함하는 변수를 지정하고 선택적으로 계층 또는 군집을 정의합니다.
5. 마침을 클릭하여 계획을 저장합니다.

또는 다음 단계를 추가로 수행할 수 있습니다.

- 추정 방법 단계에서 표준 오차 추정 방법을 선택합니다.
- 크기 단계에서 표본추출된 단위 수나 단위당 포함 확률을 지정합니다.
- 디자인에 두 번째 또는 세 번째 단계를 추가합니다.
- 선택을 명령 구문으로 붙여넣습니다.

---

### 분석 준비 마법사: 계획 변수

이 단계에서는 층화변수와 군집 변수를 식별하고 표본 가중치를 정의할 수 있습니다. 단계에 대한 레이블을 제공할 수도 있습니다.

**계층.** 층화변수의 교차 분류는 고유한 부모집단이나 계층을 정의합니다. 전체 표본은 각 계층의 독립 표본 조합을 나타냅니다.

**군집** 군집 변수는 관측 단위 그룹 또는 군집을 정의합니다. 여러 단계에서 작성된 표본은 초기 단계에서 군집을 선택한 다음 선택한 군집에서 부표본 단위를 선택합니다. 복원 군집 표본추출에 의해 수집된 데이터 파일을 분석하는 경우 중복 지수를 군집 변수로 포함해야 합니다.

**표본 가중치.** 첫 번째 단계에서 표본 가중치를 제공해야 합니다. 현재 디자인의 후속 단계에 대한 표본 가중치는 자동으로 계산됩니다.

**단계 레이블.** 각 단계에 대해 선택적으로 문자 레이블을 지정할 수 있습니다. 이러한 문자 설명은 단계별 정보 파악에 유용합니다.

**참고:** 소스 변수 목록에는 마법사 단계와 상관없이 모두 같은 내용이 포함됩니다. 즉, 특정 단계의 소스 목록에서 제거된 변수는 모든 단계의 목록에서 제거됩니다. 소스 목록으로 반환된 변수는 모든 단계를 나타냅니다.

## 분석 마법사 탐색을 위한 트리 제어

분석 마법사에서 각 단계의 왼쪽에는 모든 단계의 개요가 표시됩니다. 개요에서 사용할 수 있는 단계 이름을 클릭하여 마법사를 탐색할 수 있습니다. 모든 이전 단계가 유효하면 즉, 각 이전 단계에서 해당 단계별로 필요한 최소 지정 사항을 지정했으면, 현재 단계를 사용할 수 있습니다. 주어진 단계가 유효하지 않은 이유에 대한 자세한 내용은 단계별 도움말을 참조하십시오.

---

## 분석 준비 마법사: 추정 방법

이 단계에서는 각 단계의 추정 방법을 지정할 수 있습니다.

**WR(복원 표본추출)** 복합 표본추출 디자인에서 분산을 추정할 때는 WR 추정에 유한 모집단(FPC)에서의 표본추출에 대한 보정이 포함되지 않습니다. 단순 무작위 표본(SRS)에서 분산을 추정할 때는 FPC를 포함하거나 제외하도록 선택할 수 있습니다.

분석 가중치가 모집단 크기를 구성하지 않도록 척도화한 경우 SRS 분산 추정에 대해 FPC를 포함하지 않도록 선택하는 것이 좋습니다. SRS 분산 추정값은 디자인 효과와 같은 통계를 계산하는 데 사용됩니다. WR 추정은 디자인의 마지막 단계에서만 지정될 수 있습니다. WR 추정을 선택한 경우 마법사에서 다른 단계를 추가할 수 없습니다.

**등확률 WOR(등확률 비복원 표본추출).** 등확률 WOR 추정에는 유한 모집단 보정이 포함되며 단위가 등확률로 표본추출되었다고 가정합니다. 등확률 WOR은 디자인의 어느 단계에서나 지정할 수 있습니다.

**부등확률 WOR(부등확률 비복원 표본추출).** 부등확률 WOR은 유한 모집단 보정을 사용하는 것 이외에 부등확률로 선택한 표본추출 단위(일반적으로 군집)를 나타냅니다. 이 추정 방법은 첫 번째 단계에서만 사용할 수 있습니다.

---

## 분석 준비 마법사: 크기

이 단계는 현재 단계에 필요한 포함 확률이나 모집단 크기를 지정하는 데 사용됩니다. 크기는 계층별로 다르거나 고정될 수 있습니다. 크기 지정은 위해 이전 단계에서 지정한 군집을 사용하여 계층을 정의할 수 있습니다. 이 단계는 등확률 WOR이 추정 방법으로 선택되었을 때만 필요합니다.

**단위.** 정확한 모집단 크기나 단위 표본추출 시 사용된 확률을 지정할 수 있습니다.

- **값.** 단일값이 모든 계층에 적용됩니다. 모집단 크기를 메트릭 단위로 선택한 경우 음수가 아닌 정수를 입력해야 합니다. 포함 확률을 선택한 경우 0과 1사이(포함)의 값을 입력해야 합니다.
- 계층에 대해 불균등 값. 불균등 크기 정의 대화 상자를 통해 계층별로 크기 값을 입력할 수 있습니다.
- 값을 읽을 변수. 계층의 크기 값을 포함하는 숫자변수를 선택할 수 있습니다.

## 불균등 크기 정의

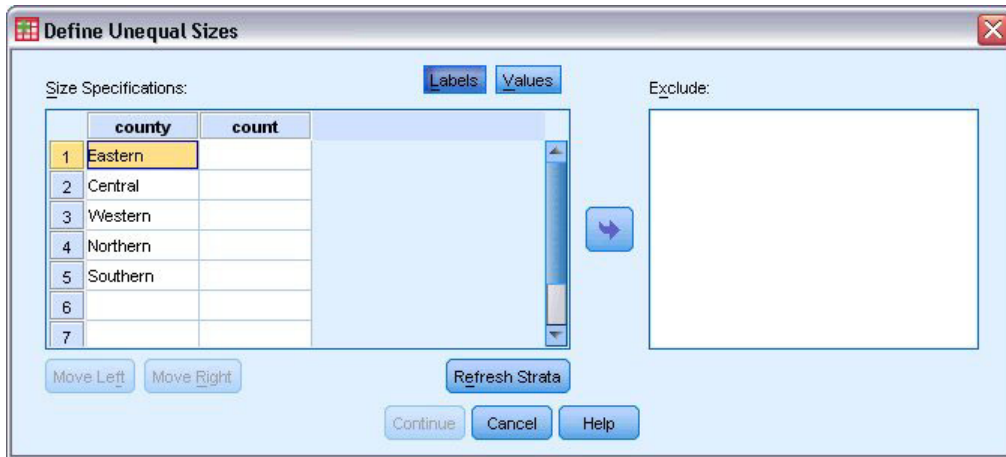


그림 2. 불균등 크기 정의 대화 상자

불균등 크기 정의 대화 상자에서는 계층별로 크기를 입력할 수 있습니다.

**크기 지정 사항 격자.** 이 격자에는 최대 다섯 개의 층화변수 또는 군집 변수, 행마다 계층/군집 조합 하나씩에 대한 교차 분류가 표시됩니다. 적합한 격자 변수에는 현재 및 이전 단계의 모든 층화변수는 물론 이전 단계의 모든 군집 변수가 포함되어 있습니다. 격자 내에서 변수 순서를 바꾸거나 제외 목록으로 이동할 수 있습니다. 맨 오른쪽 열에 크기를 입력합니다. 격자 셀에 층화변수와 군집 변수에 대한 값 레이블로 표시할지 데이터 값으로 표시할지를 전환하려면 레이블 또는 값을 클릭합니다. 레이블을 붙이지 않은 값을 포함하는 셀에는 항상 값이 표시됩니다. 격자를 해당 변수에 대해 레이블이 있는 데이터 값 조합으로 다시 채우려면 계층 최신 정보 표시를 클릭합니다.

**제외.** 계층/군집 조합의 서브세트 크기를 지정하려면 하나 이상의 변수를 제외 목록으로 옮깁니다. 이러한 변수는 표본 결과를 정의하는 데 사용되지 않습니다.

## 분석 준비 마법사: 계획 요약

이 단계는 각 단계의 마지막 단계로 현재 단계를 통해 분석 디자인 사양에 대한 요약물을 제공합니다. 여기에서 다음 단계로 진행하거나 필요에 따라 단계를 작성하거나 분석 지정 사항을 저장할 수 있습니다.

다음과 같은 경우 다른 단계를 추가할 수 없습니다.

- 계획 변수 단계에 군집 변수가 지정되어 있지 않습니다.
- 추정 방법 단계에서 WR 추정을 선택했습니다.

- 최대 세 단계를 지원하는 마법사의 세 번째 분석 단계입니다.

---

## 분석 준비 마법사: 완료

마지막 단계입니다. 지금 계획 파일을 저장하거나 선택을 구문 창에 붙여넣을 수 있습니다.

기존 계획 파일의 단계를 변경하는 경우 편집된 계획을 새 파일에 저장하거나 기존 파일을 덮어쓸 수 있습니다. 기존 단계를 변경하지 않고 단계를 새로 추가하는 경우 마법사를 통해 기존 계획 파일을 자동으로 덮어쓰게 됩니다. 계획을 새 파일에 저장하려면 마법사에서 생성한 명령을 구문 창에 붙여넣기를 선택하고 명령을 사용하여 파일 이름을 변경합니다.

---

## 기존 분석 계획 수정

1. 메뉴에서 다음을 선택합니다.

분석 > 복합 표본 > 분석 준비...

2. 계획 파일 편집을 선택하고 분석 계획을 저장할 계획 파일 이름을 선택합니다.
3. 마법사에서 작업을 계속하려면 다음을 클릭합니다.
4. 계획 요약 단계에서 분석 계획을 검토하고 다음을 클릭합니다.

후속 단계는 대부분 새 디자인의 경우와 같습니다. 자세한 내용은 각 단계의 도움말을 참조하십시오.

5. 마침 단계로 이동하여 편집된 계획 파일에 대해 새 이름을 지정하거나 기존 계획 파일을 덮어쓰도록 선택합니다.

선택적으로 계획에서 단계를 제거할 수 있습니다.

---

## 분석 준비 마법사: 계획 요약

이 단계에서는 분석 계획을 검토하여 계획에서 단계를 제거할 수 있습니다.

단계 제거. 다단계 디자인에서 2단계와 3단계를 제거할 수 있습니다. 디자인에는 최소한 한 단계가 있어야 하므로 계획에서 1단계를 편집할 수는 있지만 제거할 수는 없습니다.

---

## 제 4 장 복합 표본 계획

복합 표본 분석 프로시저에서 유효한 결과를 도출하려면 분석이나 표본 계획 파일의 분석 지정 사항이 필요합니다.

**계획.** 분석 또는 표본 계획 파일의 경로를 지정합니다.

**결합 확률.** PPS WOR 방법을 사용하여 도출된 군집의 부등확률 WOR 추정을 사용하려면 결합 확률을 포함하는 데이터 세트를 열거나 별도의 파일을 지정해야 합니다. 이 파일이나 데이터 세트는 표본을 추출할 때 표본추출 마법사에서 작성됩니다.



---

## 제 5 장 복합 표본 빈도분석

복합 표본 빈도분석 프로시저를 통해 선택한 변수에 대한 빈도표를 작성한 다음 일변량 통계량을 표시합니다. 선택적으로 하나 이상의 범주형 변수로 정의된 하위 그룹별 통계도 요청할 수 있습니다.

**예제.** 이 예제에서는 복합 표본 빈도분석 프로시저를 사용하여 NHIS(National Health Interview Survey)의 결과를 참조함은 물론 공용 데이터에 대한 적절한 분석 계획을 세워 미국 시민의 비타민 사용량에 따른 일변량 표 통계를 구해볼 수 있습니다.

**통계.** 이 프로시저에서는 셀 모집단 크기와 표 백분율의 추정값과 더불어 각 추정값의 표준 오차, 신뢰구간, 변동 계수, 디자인 효과, 디자인 효과 제공근, 누적값 및 가중치를 부여하지 않은 개수를 작성합니다. 또한 동등 셀 비율 검정을 위해 카이제곱 및 우도비 통계가 계산됩니다.

복합 표본 빈도분석 데이터 고려 사항

**데이터.** 빈도표가 작성될 변수는 범주형이어야 합니다. 부모집단 변수는 문자 또는 숫자일 수 있지만 범주형이어야 합니다.

**가정.** 데이터 파일의 케이스는 복합 표본 계획 대화 상자에서 선택한 파일의 지정 사항에 따라 분석해야 할 복합 디자인의 표본을 나타냅니다.

복합 표본 빈도분석 구하기

1. 메뉴에서 다음을 선택합니다.

분석 > 복합 표본 > 빈도...

2. 계획 파일 선택. 선택적으로 사용자 정의 결합 확률 파일을 선택합니다.

3. 계속을 클릭합니다.

4. 복합 표본 빈도 대화 상자에서, 최소 한 개의 빈도 변수를 선택합니다.

선택적으로 부모집단을 정의할 변수를 지정할 수 있습니다. 통계는 부모집단별로 계산됩니다.

---

### 복합 표본 빈도분석 통계

**셀.** 셀 모집단 크기와 표 백분율의 추정값을 요청할 수 있습니다.

**통계.** 모집단 크기 또는 표 백분율과 연결된 통계를 작성합니다.

- 표준 오차. 추정값의 표준 오차입니다.
- 신뢰구간. 지정한 수준을 사용하는 추정값의 신뢰구간입니다.
- 변동 계수. 추정값 간의 표준 오차 비율입니다.
- 가중치를 부여하지 않은 개수. 추정값을 계산하는 데 사용된 단위 수입니다.

- **디자인 효과.** 표본이 단순 무작위 표본이라고 가정하여 얻은 분산에 대한 추정값 분산 비율입니다. 복합 디자인 지정으로 인한 효과를 측정하는 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다.
- **디자인 효과 제공근.** 복합 디자인 지정으로 인한 효과를 측정하는 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다.
- **누적값.** 변수의 각 값을 누적 추정합니다.

동등 셀 비율 검정. 변수의 범주가 같은 빈도를 갖는다는 가설에 대한 카이제곱 검정과 우도비 검정을 수행합니다. 변수별로 검정이 수행됩니다.

## 복합 표본 결측값

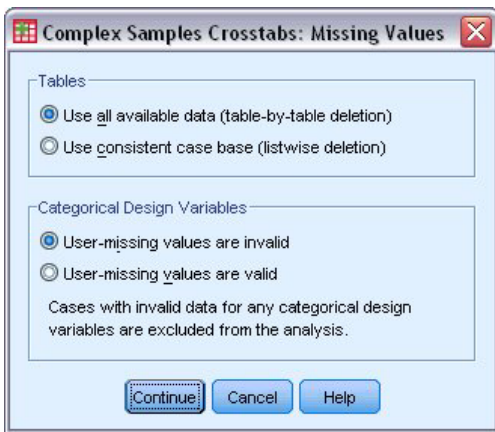


그림 3. 결측값 대화 상자

**테이블.** 분석에 사용되는 케이스를 결정합니다.

- **사용 가능한 모든 데이터 사용.** 결측값을 테이블 단위로 확인합니다. 따라서 통계를 계산하는 데 사용되는 케이스가 빈도표나 교차 분석표에 따라 달라질 수 있습니다.
- **일관된 케이스 기준 사용.** 모든 변수에서 결측값을 확인합니다. 따라서 통계를 계산하는 데 사용되는 케이스가 모든 테이블에서 일관됩니다.

**범주형 계획 변수.** 사용자 결측값의 유효 여부를 결정합니다.



## 복합 표본 옵션

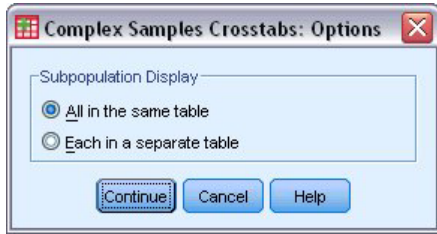


그림 4. 옵션 대화 상자

부모집단 출력. 부모집단을 같은 테이블에 표시할지, 아니면 별도의 테이블에 표시할지를 선택할 수 있습니다.



---

## 제 6 장 복합 표본 기술통계

복합 표본 기술통계 프로시저를 통해 여러 변수에 대한 일변량 요약 통계를 표시합니다. 선택적으로 하나 이상의 범주형 변수로 정의된 하위 그룹별 통계도 요청할 수 있습니다.

**예제.** 이 예제에서는 복합 표본 기술통계 프로시저를 사용하여 NHIS(National Health Interview Survey) 결과를 참조함은 물론 공용 데이터에 대한 적절한 분석 계획을 세워 미국 시민의 활동 수준에 대한 일변량 기술 통계를 구해볼 수 있습니다.

**통계.** 이 프로시저에서는 평균 및 합계와 더불어 각 추정값의  $T$  검정, 표준 오차, 신뢰구간, 변동 계수, 가중치를 부여하지 않은 개수, 모집단 크기, 디자인 효과 및 디자인 효과 제공근을 작성합니다.

복합 표본 기술통계 데이터 고려 사항

**데이터.** 측도는 척도변수이어야 합니다. 부모집단 변수는 문자 또는 숫자일 수 있지만 범주형이어야 합니다.

**가정.** 데이터 파일의 케이스는 복합 표본 계획 대화 상자에서 선택한 파일의 지정 사항에 따라 분석해야 할 복합 디자인의 표본을 나타냅니다.

복합 표본 기술통계 구하기

1. 메뉴에서 다음을 선택합니다.

    분석 > 복합 표본 > 기술통계...

2. 계획 파일 선택. 선택적으로 사용자 정의 결합 확률 파일을 선택합니다.

3. 계속을 클릭합니다.

4. 복합 표본 기술통계 대화 상자에서 한 개 이상의 측도 변수를 선택합니다.

선택적으로 부모집단을 정의할 변수를 지정할 수 있습니다. 통계는 부모집단별로 계산됩니다.

---

### 복합 표본 기술통계

**요약값.** 측도변수의 평균과 합계의 추정값을 요청할 수 있습니다. 또한 지정한 값에 대해 추정값의  $T$  검정을 요청할 수 있습니다.

**통계.** 평균 또는 합계와 연결된 통계를 작성합니다.

- **표준 오차.** 추정값의 표준 오차입니다.
- **신뢰구간.** 지정한 수준을 사용하는 추정값의 신뢰구간입니다.
- **변동 계수.** 추정값 간의 표준 오차 비율입니다.
- **가중치를 부여하지 않은 개수.** 추정값을 계산하는 데 사용된 단위 수입니다.
- **모집단 크기.** 모집단의 추정된 단위 수입니다.

- 디자인 효과. 표본이 단순 무작위 표본이라고 가정하여 얻은 분산에 대한 추정값 분산 비율입니다. 복합 디자인 지정으로 인한 효과를 측정한 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다.
- 디자인 효과 제공근. 복합 디자인 지정으로 인한 효과를 측정한 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다.

## 복합 표본 기술통계 결측값

측도변수 통계. 분석에 사용되는 케이스를 결정합니다.

- 사용 가능한 모든 데이터 사용. 결측값은 변수별로 결정되므로 통계 계산에 사용되는 케이스는 측도변수에 따라 달라질 수 있습니다.
- 일관된 케이스 기준 확보. 결측값은 모든 변수에서 결정되므로 통계 계산에 사용되는 케이스는 일관됩니다.

범주형 계획 변수. 사용자 결측값의 유효 여부를 결정합니다.

## 복합 표본 옵션

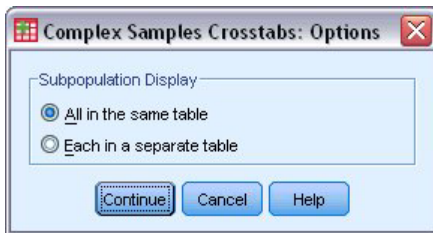


그림 5. 옵션 대화 상자

부모집단 출력. 부모집단을 같은 테이블에 표시할지, 아니면 별도의 테이블에 표시할지를 선택할 수 있습니다.

---

## 제 7 장 복합 표본 교차 분석표

복합 표본 교차 분석표 프로시저를 통해 선택한 대응변수에 대한 교차 분석표를 작성하고 이차원 통계를 표시합니다. 선택적으로 하나 이상의 범주형 변수로 정의된 하위 그룹별 통계도 요청할 수 있습니다.

**예제.** 이 예제에서는 복합 표본 교차 분석표 프로시저를 사용하여 NHIS(National Health Interview Survey) 결과를 참조함은 물론 공용 데이터에 대한 적절한 분석 계획을 세워 미국 시민의 비타민 사용량에 따른 흡연 빈도의 교차 분류 통계를 구해볼 수 있습니다.

**통계.** 이 프로시저에서는 셀 모집단 크기와 행, 열 및 표 백분율의 추정값과 더불어 각 추정값의 표준 오차, 신뢰구간, 변동 계수, 기대값, 디자인 효과, 디자인 효과 제공근, 잔차, 수정된 잔차 및 기중치를 부여하지 않은 개수를 작성합니다. 2X2 테이블에 대해 승산비, 상대적 위험 및 위험 차이를 계산합니다. 또한 행 변수와 열 변수의 독립성 검정을 위한 Pearson 통계와 우도비 통계를 계산합니다.

복합 표본 교차 분석표 데이터 고려 사항

**데이터.** 행 변수와 열 변수는 범주형이어야 합니다. 부모집단 변수는 문자 또는 숫자일 수 있지만 범주형이어야 합니다.

**가정.** 데이터 파일의 케이스는 복합 표본 계획 대화 상자에서 선택한 파일의 지정 사항에 따라 분석해야 할 복합 디자인의 표본을 나타냅니다.

복합 표본 교차 분석표 구하기

1. 메뉴에서 다음을 선택합니다.

분석 > 복합 표본 > 교차 분석표...

2. 계획 파일 선택. 선택적으로 사용자 정의 결합 확률 파일을 선택합니다.

3. 계속을 클릭합니다.

4. 복합 표본 교차 분석표 대화 상자에서 한 개 이상의 행 변수 및 열 변수를 선택합니다.

선택적으로 부모집단을 정의할 변수를 지정할 수 있습니다. 통계는 부모집단별로 계산됩니다.

---

### 복합 표본 교차 분석표 통계

**셀.** 셀 모집단 크기와 행, 열 및 표 백분율의 추정값을 요청할 수 있습니다.

**통계.** 모집단 크기 및 행, 열 및 표 백분율과 관련된 통계를 작성합니다.

- 표준 오차. 추정값의 표준 오차입니다.
- 신뢰구간. 지정한 수준을 사용하는 추정값의 신뢰구간입니다.
- 변동 계수. 추정값 간의 표준 오차 비율입니다.

- 기대값. 행 변수와 열 변수가 서로 독립이라는 가설하에 기대되는 추정값입니다.
- 가중치를 부여하지 않은 개수. 추정값을 계산하는 데 사용된 단위 수입입니다.
- 디자인 효과. 표본이 단순 무작위 표본이라고 가정하여 얻은 분산에 대한 추정값 분산 비율입니다. 복합 디자인 지정으로 인한 효과를 측정하는 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다.
- 디자인 효과 제공근. 복합 디자인 지정으로 인한 효과를 측정하는 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다.
- 잔차. 기대값은 두 변수 사이에 관계가 없을 경우 셀에서 예상되는 케이스의 수입입니다. 행 변수와 열 변수가 서로 독립적인 경우 잔차가 양수이면 셀에 예상보다 많은 케이스가 있음을 의미합니다.
- 수정 잔차. 셀의 잔차(관측값에서 기대값을 뺀 값)를 표준 오차의 추정값으로 나눈 값입니다. 결과로 나타나는 표준화 잔차는 평균보다 높거나 낮은 표준 편차 단위로 표현됩니다.

**2X2 테이블 요약.** 행 변수와 열 변수가 각각 두 개의 범주를 가진 교차 분석표에 대한 통계를 작성합니다. 각각은 요인의 존재와 이벤트 발생 간의 연관 강도에 대한 측도입니다.

- 승산비. 요인 발생이 드문 경우 승산비를 상대적 위험의 추정값으로 사용할 수 있습니다.
- 상대적 위험도. 요인 존재 시 이벤트 위험도 대 요인 부재 시 이벤트의 위험도 비율입니다.
- 위험차. 요인 존재 시 이벤트 위험과 요인 부재 시 이벤트 위험 간의 차이입니다.

**행 변수와 열 변수의 독립성 검정.** 행 변수와 열 변수가 독립이라는 가설에 대한 카이제곱 검정과 우도비 검정을 수행합니다. 각 대응변수별로 분리된 검정이 수행됩니다.

## 복합 표본 결측값

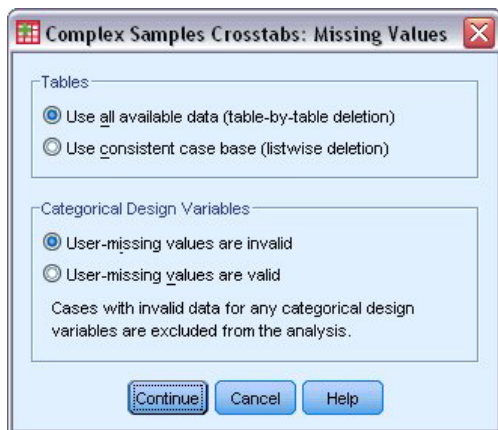


그림 6. 결측값 대화 상자

**테이블.** 분석에 사용되는 케이스를 결정합니다.

- 사용 가능한 모든 데이터 사용. 결측값을 테이블 단위로 확인합니다. 따라서 통계를 계산하는 데 사용되는 케이스가 빈도표나 교차 분석표에 따라 달라질 수 있습니다.

- 일관된 케이스 기준 사용. 모든 변수에서 결측값을 확인합니다. 따라서 통계를 계산하는 데 사용되는 케이스가 모든 테이블에서 일관됩니다.

범주형 계획 변수. 사용자 결측값의 유효 여부를 결정합니다.

---

## 복합 표본 옵션

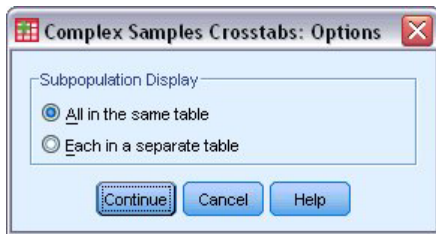


그림 7. 옵션 대화 상자

부모집단 출력. 부모집단을 같은 테이블에 표시할지, 아니면 별도의 테이블에 표시할지를 선택할 수 있습니다.





---

## 제 8 장 복합 표본 비율

복합 표본 비율 프로시저에서는 변수 비율에 대한 일변량 요약 통계를 표시합니다. 선택적으로 하나 이상의 범주형 변수로 정의된 하위 그룹별 통계도 요청할 수 있습니다.

예제. 복합 표본 비율 프로시저를 사용하여 복합 디자인에 따라 수행된 국가 차원의 시장설문조사 결과와 적절한 데이터 분석 계획을 통해 현재 특성 값 대 마지막 검사 값의 비율에 대한 기술통계를 구할 수 있습니다.

통계. 이 프로시저에서는 비율 추정값,  $T$  검정, 표준 오차, 신뢰구간, 변동 계수, 가중치를 부여하지 않은 개수, 모집단 크기, 디자인 효과 및 디자인 효과 제공근을 작성합니다.

복합 표본 비율 데이터 고려 사항

데이터. 분자와 분모는 양의 값을 갖는 척도변수이어야 합니다. 부모집단 변수는 문자 또는 숫자일 수 있지만 범주형이어야 합니다.

가정. 데이터 파일의 케이스는 복합 표본 계획 대화 상자에서 선택한 파일의 지정 사항에 따라 분석해야 할 복합 디자인의 표본을 나타냅니다.

복합 표본 비율 구하기

1. 메뉴에서 다음을 선택합니다.

    분석 > 복합 표본 > 비율...

2. 계획 파일 선택. 선택적으로 사용자 정의 결합 확률 파일을 선택합니다.

3. 계속을 클릭합니다.

4. 복합 표본 비율 대화 상자에서 한 개 이상의 분자변수 및 분모변수를 선택합니다.

선택적으로 통계를 산출한 하위 그룹을 정의하는 변수를 지정할 수 있습니다.

---

### 복합 표본 비율 통계

통계. 비율 추정값과 연결된 통계를 작성합니다.

- 표준 오차. 추정값의 표준 오차입니다.
- 신뢰구간. 지정한 수준을 사용하는 추정값의 신뢰구간입니다.
- 변동 계수. 추정값 간의 표준 오차 비율입니다.
- 가중치를 부여하지 않은 개수. 추정값을 계산하는 데 사용된 단위 수입니다.
- 모집단 크기. 모집단의 추정된 단위 수입니다.
- 디자인 효과. 표본이 단순 무작위 표본이라고 가정하여 얻은 분산에 대한 추정값 분산 비율입니다. 복합 디자인 지정으로 인한 효과를 측정하는 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다.

- 디자인 효과 제공근. 복합 디자인 지정으로 인한 효과를 측정한 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다.

**T 검정.** 지정한 값에 대해 추정값의 *T* 검정을 요청할 수 있습니다.

## 복합 표본 비율 결측값

**비율.** 분석에 사용되는 케이스를 결정합니다.

- 사용 가능한 모든 데이터 사용. 비율별로 결측값을 확인합니다. 따라서 통계를 계산하는 데 사용되는 케이스는 분자-분모 쌍에 따라 달라질 수 있습니다.
- 일관된 케이스 기준 확보. 모든 변수에서 결측값을 확인합니다. 따라서 통계를 계산하는 데 사용되는 케이스가 일관됩니다.

**범주형 계획 변수.** 사용자 결측값의 유효 여부를 결정합니다.

## 복합 표본 옵션

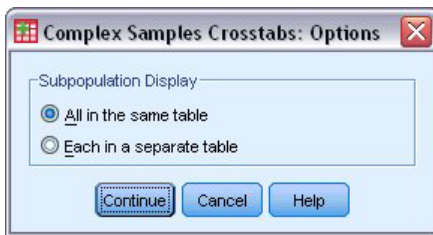


그림 8. 옵션 대화 상자

**부모집단 출력.** 부모집단을 같은 테이블에 표시할지, 아니면 별도의 테이블에 표시할지를 선택할 수 있습니다.

---

## 제 9 장 복합 표본 일반선형모형

복합 표본 일반선형모형(CSGLM) 프로시저에서는 복합 표본추출 방법으로 작성한 표본에 대한 분산 및 공분산 분석을 비롯하여 선형 회귀 분석을 수행합니다. 선택적으로 부모집단에 대한 분석을 요청할 수 있습니다.

**예제.** 식료품 체인점에서 복합 디자인에 따라 고객 그룹의 구매 성향을 설문조사했습니다. 설문조사 결과와 각 고객의 지난 달 구매 금액을 기반으로 이 상점에서는 표본추출 디자인을 통합하여 고객의 성별에 따라 한 달 간 고객의 상품 구매 횟수가 구매 금액과 관련이 있는지 파악하려 합니다.

**통계.** 이 프로시저에서는 추정값, 표준 오차, 신뢰구간,  $T$  검정, 디자인 효과 및 모형 모수에 대한 디자인 효과의 제곱근을 비롯하여 모수 추정값 사이의 상관 및 공분산을 작성합니다. 독립변수와 종속변수에 대한 기술통계 및 모형 적합의 측도값을 구할 수도 있습니다. 또한 모형 요인의 수준과 요인 상호작용에 대한 추정된 주변 평균을 요청할 수도 있습니다.

복합 표본 일반선형모형 데이터 고려 사항

**데이터.** 종속변수는 양적변수입니다. 요인은 범주형입니다. 공변량은 종속변수와 관련된 양적변수입니다. 부모 집단 변수는 문자 또는 숫자일 수 있지만 범주형이어야 합니다.

**가정.** 데이터 파일의 케이스는 복합 표본 계획 대화 상자에서 선택한 파일의 지정 사항에 따라 분석해야 할 복합 디자인의 표본을 나타냅니다.

복합 표본 일반선형모형 구하기

1. 메뉴에서 다음을 선택합니다.

분석 > 복합 표본 > 일반 선형 모형...

2. 계획 파일 선택. 선택적으로 사용자 정의 결합 확률 파일을 선택합니다.

3. 계속을 클릭합니다.

4. 복합 표본 일반선형모형 대화 상자에서 종속변수를 선택합니다.

선택적으로 다음을 수행할 수 있습니다.

- 데이터에 따라 적절한 요인 및 공변량의 변수를 선택합니다.
- 부모집단을 정의하는 데 사용할 변수를 지정합니다. 부모집단 변수의 선택된 범주에 대해서만 분석이 수행됩니다.

---

## 복합 표본 일반선형모형

**모형 효과 지정.** 기본적으로 이 프로시저에서는 기본 대화 상자에 지정된 요인 및 공변량을 사용하여 주효과 모형을 작성합니다. 그러나 상호작용 효과 및 중첩 항이 포함되도록 사용자 정의 모형을 작성할 수도 있습니다.

중첩되지 않은 항

선택된 요인과 공변량에 대해 다음과 같습니다.

**상호작용.** 선택한 모든 변수의 최고 수준 상호작용 항을 생성합니다.

**주효과.** 선택한 각 변수에 대한 주효과 항을 생성합니다.

**모든 2원배치.** 선택한 변수에 대해 가능한 2원배치 상호작용을 모두 생성합니다.

**모든 3원배치.** 선택한 변수에 대해 가능한 3원배치 상호작용을 모두 생성합니다.

**모든 4원배치.** 선택한 변수에 대해 가능한 4원배치 상호작용을 모두 생성합니다.

**모든 5원배치.** 선택한 변수에 대해 가능한 5원배치 상호작용을 모두 생성합니다.

중첩 항

이 프로시저에서는 모형에 대해 중첩 항을 설정할 수 있습니다. 중첩 항은 요인 또는 공변량의 효과를 모델링하는 데 유용합니다. 이들 값은 다른 요인 수준과 상호작용하지 않습니다. 예를 들어, 식료품 체인점은 여러 점포에서의 고객의 소비 성향을 살펴볼 수 있습니다. 각 고객은 체인점 중의 한 곳만 자주 가기 때문에 고객 효과는 점포 효과 내에 중첩되었다고 할 수 있습니다.

또한 같은 공변량을 포함하고 있는 다항 항 같이 상호작용 효과를 포함시키거나 여러 수준의 중첩을 중첩 항에 추가할 수 있습니다.

**제한.** 중첩 항에는 다음과 같은 제한이 있습니다.

- 상호작용 내의 모든 요인은 고유해야 합니다. 따라서  $A$ 가 요인이면  $A*A$  지정은 유효하지 않습니다.
- 중첩 효과 내의 모든 요인은 고유해야 합니다. 따라서  $A$ 가 요인이면  $A(A)$  지정은 유효하지 않습니다.
- 공변량 내에 효과를 중첩할 수 없습니다. 따라서  $A$ 가 요인이고  $X$ 가 공변량이면  $A$  지정은 유효하지 않습니다.

**절편.** 이 모형에는 대개 절편이 포함됩니다. 데이터가 원점을 전달한다고 가정할 경우에는 절편을 제외시킬 수 있습니다. 모형에 절편을 포함시키는 경우에도 그와 관련된 통계를 표시하지 않도록 선택할 수 있습니다.

---

## 복합 표본 일반선형모형 통계

모형 모수. 이 그룹을 사용하면 모형 모수와 관련된 통계의 표시를 제어할 수 있습니다.

- 추정값. 계수의 추정값을 표시합니다.
- 표준 오차. 각 계수 추정값에 대한 표준 오차를 표시합니다.
- 신뢰구간. 각 계수 추정값에 대한 신뢰구간을 표시합니다. 이 구간의 신뢰수준은 옵션 대화 상자에서 설정합니다.
- T 검정. 각 계수 추정값에 대한 T 검정을 표시합니다. 각 검정에 대한 귀무가설은 계수 값이 0임을 의미합니다.
- 모수 추정값의 공분산. 모형 계수에 대한 공분산 행렬의 추정값을 표시합니다.
- 모수 추정값의 상관. 모형 계수에 대한 상관 행렬의 추정값을 표시합니다.
- 디자인 효과. 표본이 단순 무작위 표본이라고 가정하여 얻은 분산에 대한 추정값 분산 비율입니다. 복합 디자인 지정으로 인한 효과를 측정하는 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다.
- 디자인 효과 제공근. 복합 디자인 지정으로 인한 효과를 측정하는 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다.

모형 적합.  $R^2$  및 평균 제곱 오차 제공근 통계를 표시합니다.

종속변수와 공변량의 모집단 평균. 종속변수, 공변량 및 요인에 대한 요약 정보를 표시합니다.

표본 디자인 정보. 가중치를 부여하지 않은 개수와 모집단 크기를 비롯하여 표본에 대한 요약 정보를 표시합니다.

---

## 복합 표본 가설 검정

검정 통계량. 이 그룹을 사용하면 가설 검정에 사용되는 통계의 유형을 선택할 수 있습니다. F, 수정된 F, 카이제곱 및 수정된 카이제곱 중에서 선택할 수 있습니다.

표본추출 자유도. 이 그룹을 사용하면 모든 검정 통계량에 대한 p 값을 계산하는 데 사용되는 표본추출 디자인 자유도를 제어할 수 있습니다. 표본추출 디자인을 기반으로 하는 경우 이 값은 표본추출의 첫 번째 단계의 계층 수 및 기본 표본추출 단위 수 사이의 차이를 가리킵니다. 또는 양의 정수를 지정하여 사용자 정의 자유도를 설정할 수도 있습니다.

다중 비교에 대한 조정. 다중 대비를 통해 가설 검정을 수행하는 경우 포함된 대비에 대한 유의 수준에서 전반적인 유의 수준을 조정할 수 있습니다. 이 그룹을 사용하면 조정 방법을 선택할 수 있습니다.

- 최소유의차. 이 방법은 특정 선형 대비가 귀무가설 값과 다르다는 가설을 거부할 전체 확률을 제어하지 않습니다.
- 순차 Sidak(Sequential Sidak). 순차 단계별로 낮아지는 거부 Sidak 프로시저로, 개별 가설은 거부하지만 동일한 전체 유의 수준을 유지한다는 점에서 훨씬 덜 보수적인 방법입니다.

- 순차 Bonferroni(Sequential Bonferroni). 순차 단계별로 낮아지는 거부 Bonferroni 프로시저로, 개별 가설은 거부하지만 동일한 전체 유의 수준을 유지한다는 점에서 훨씬 덜 보수적인 방법입니다.
- Sidak. 이 방법은 Bonferroni 접근법보다 더 엄격한 한계를 제공합니다.
- Bonferroni. 이 방법은 관측 유의 수준을 다중 대비가 검정되는 요인에 맞춥니다.

## 복합 표본 일반선형모형 평균 추정

평균 추정 대화 상자를 사용하면 모형 하위 대화 상자에서 지정한 요인 수준 및 요인 상호작용에 대한 모형 추정된 주변 평균을 표시할 수 있습니다. 전체 모집단 평균을 표시하도록 요청할 수도 있습니다.

**항.** 추정된 평균은 선택한 요인 및 요인 상호작용에 대해 계산됩니다.

**대비.** 대비는 추정된 평균을 비교하기 위해 가설 검정을 어떻게 설정할지 결정합니다.

- 단순. 각 수준의 평균을 지정한 수준의 평균과 비교합니다. 이 대비 유형은 대조군이 있는 경우 유용합니다.
- 편차. 각 수준의 평균(참조 범주 제외)을 모든 수준의 평균(총 평균)과 비교합니다. 요인의 수준들은 원하는 순서대로 지정할 수 있습니다.
- 차이. 각 수준의 평균(처음 제외)을 이전 수준의 평균과 비교합니다. 역 Helmert 대비라고도 합니다.
- Helmert. 요인의 각 수준 평균(마지막 제외)을 후속 수준의 평균과 비교합니다.
- 반복. 각 수준의 평균(마지막 제외)을 후속 수준의 평균과 비교합니다.
- 다항. 선형 효과, 2차 효과, 3차 효과 등을 비교합니다. 처음 자유도에는 모든 범주에서의 선형 효과가 포함됩니다. 그리고 두 번째 자유도는 2차 효과, 세 번째는 3차 효과의 순으로 계속 적용됩니다. 이들 대비는 다항 추세를 추정하는 데 사용됩니다.

**참조 범주.** 단순대비와 편차 대비에는 참조 범주 또는 다른 항을 비교하는 데 기준으로 삼을 요인 수준이 필요합니다.

## 복합 표본 일반선형모형 저장

**저장할 변수.** 이 그룹을 사용하면 모형 예측값과 잔차를 작업 파일에 새 변수로 저장할 수 있습니다.

**IBM SPSS Statistics 데이터로 모형 내보내기.** 모수 추정값, 표준 오차, 유의수준 및 자유도와 함께 모수 상관 행렬이나 공분산 행렬이 포함되어 있는 IBM SPSS Statistics 형식의 데이터 세트를 씁니다. 행렬 파일에서 변수의 순서는 다음과 같습니다.

- **rowtype\_.** 값(및 레이블), COV(공분산), CORR(상관), EST(모수 추정값), SE(표준 오차), SIG(유의 수준) 및 DF(표본추출 디자인 자유도)를 사용합니다. 각 모형 모수에 대해 행 유형 COV (또는 CORR)를 갖는 별도의 케이스는 물론 각기 다른 행 유형에 대해 별도의 케이스가 있습니다.
- **varname\_.** 모수 추정값 테이블에 표시된 모수 문자에 해당하는 값 레이블이 있는 행 유형 COV 또는 CORR에 대해 모든 모형 모수의 순서 목록에 해당하는 값 P1, P2, ...를 사용합니다. 다른 행 유형의 셀은 비어 있습니다.

- **P1, P2, ...** 이러한 변수는 모수 추정값 테이블에 표시된 모수 문자에 해당하는 변수 레이블과 함께 모든 모형 모수의 순서 목록에 해당하며 행 유형에 따라 값을 사용합니다. 중복 모수의 경우 모든 공분산은 0으로 설정되고 상관은 시스템 결측값으로 설정됩니다. 모든 모수 추정값은 0으로 설정되고 모든 표준 오차, 유의 수준 및 잔차의 자유도는 시스템 결측값으로 설정됩니다.

참고: 이러한 프로시저가 여기에 내보낸 모든 행 유형을 받아 들이지 않는 경우 이 파일은 행렬 파일을 읽는 다른 프로시저에서 추가 분석에 즉시 사용할 수 없습니다.

**XML로 모형 내보내기.** 모수 추정값과 모수 공분산 행렬(선택한 경우)을 XML(PMML) 형식으로 저장합니다. 스코어링을 목적으로 이 모형 파일을 사용하여 모형 정보를 다른 데이터 파일에 적용할 수 있습니다.

## 복합 표본 일반선형모형 옵션

**사용자 결측값.** 종속변수와 공변량을 비롯한 모든 계획 변수에는 유효한 데이터가 있어야 합니다. 이러한 변수에 대해 데이터가 하나라도 유효하지 않은 케이스는 분석 대상에서 제외됩니다. 이 제어를 사용하면 계층, 군집, 부모집단 및 요인변수 사이에서 사용자 결측값을 유효한 데이터로 취급할지 여부를 결정할 수 있습니다.

**신뢰구간.** 계수 추정값과 추정된 주변 평균에 대한 신뢰구간 수준입니다. 50보다 크거나 같고 100보다 작은 값을 지정합니다.

## CSGLM 명령 추가 기능

명령 구문을 사용하여 수행할 수 있는 추가 기능은 다음과 같습니다.

- CUSTOM 하위 명령을 사용하여 효과에 대한 효과 대 선형 조합의 검정이나 값을 지정할 수 있습니다.
- EMMEANS 하위 명령을 사용하여 추정된 주변 평균을 계산할 때 해당 평균 이외의 다른 값으로 공변량을 고정할 수 있습니다.
- EMMEANS 하위 명령을 사용하여 다항 대비에 대한 메트릭 방법을 지정할 수 있습니다.
- CRITERIA 하위 명령을 사용하여 비정칙성 확인을 위한 허용 오차 값을 지정할 수 있습니다.
- SAVE 하위 명령을 사용하여 저장된 변수에 대한 사용자 지정 이름을 만들 수 있습니다.
- PRINT 하위 명령을 사용하여 일반 추정가능 함수 테이블을 생성할 수 있습니다.

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





---

## 제 10 장 복합 표본 로지스틱 회귀분석

복합 표본 로지스틱 회귀분석 프로시저에서는 복합 표본추출 방법으로 구한 표본의 이분형 또는 다항 종속변수에 대한 로지스틱 회귀분석을 수행합니다. 선택적으로 부모집단에 대한 분석을 요청할 수 있습니다.

**예제.** 대출 담당자가 복합 디자인에 따라 은행의 각기 다른 여러 지점에서 대출을 받은 고객의 과거 레코드를 수집했습니다. 이 담당자는 표본 디자인을 통합하여 고객의 대출금 상환 연체와 연령, 고용 이력 및 신용 대출 금액 사이에 연관성이 있을 가능성이 있는지 확인하려고 합니다.

**통계.** 이 프로시저에서는 추정값, 지수 추정값, 표준 오차, 신뢰구간,  $T$  검정, 디자인 효과 및 모형 모수에 대한 디자인 효과의 제공근을 비롯하여 모수 추정값 사이의 상관 및 공분산을 작성합니다. 유사  $R^2$  통계, 분류표 및 종속변수와 독립변수에 대한 기술통계도 사용할 수 있습니다.

복합 표본 로지스틱 회귀분석 데이터 고려 사항

**데이터.** 종속변수는 범주형입니다. 요인은 범주형입니다. 공변량은 종속변수와 관련된 양적변수입니다. 부모집단 변수는 문자 또는 숫자일 수 있지만 범주형이어야 합니다.

**가정.** 데이터 파일의 케이스는 복합 표본 계획 대화 상자에서 선택한 파일의 지정 사항에 따라 분석해야 할 복합 디자인의 표본을 나타냅니다.

복합 표본 로지스틱 회귀분석 구하기

1. 메뉴에서 다음을 선택합니다.

분석 > 복합 표본 > 로지스틱 회귀분석...

2. 계획 파일 선택. 선택적으로 사용자 정의 결합 확률 파일을 선택합니다.

3. 계속을 클릭합니다.

4. 복합 표본 로지스틱 회귀분석 대화 상자에서 종속변수를 선택합니다.

선택적으로 다음을 수행할 수 있습니다.

- 데이터에 따라 적절한 요인 및 공변량의 변수를 선택합니다.
- 부모집단을 정의하는 데 사용할 변수를 지정합니다. 부모집단 변수의 선택된 범주에 대해서만 분석이 수행됩니다.

---

### 복합 표본 로지스틱 회귀분석 참조 범주

기본적으로 복합 표본 로지스틱 회귀분석 프로시저에서는 값이 가장 높은 범주를 참조 범주로 지정합니다. 이 대화 상자를 사용하면 가장 높은 값, 가장 낮은 값 또는 사용자 정의 범주를 참조 범주로 지정할 수 있습니다.

---

## 복합 표본 로지스틱 회귀 모형

**모형 효과 지정.** 기본적으로 이 프로시저에서는 기본 대화 상자에 지정된 요인 및 공변량을 사용하여 주효과 모형을 작성합니다. 그러나 상호작용 효과 및 중첩 항이 포함되도록 사용자 정의 모형을 작성할 수도 있습니다.

중첩되지 않은 항

선택된 요인과 공변량에 대해 다음과 같습니다.

**상호작용.** 선택한 모든 변수의 최고 수준 상호작용 항을 생성합니다.

**주효과.** 선택한 각 변수에 대한 주효과 항을 생성합니다.

**모든 2원배치.** 선택한 변수에 대해 가능한 2원배치 상호작용을 모두 생성합니다.

**모든 3원배치.** 선택한 변수에 대해 가능한 3원배치 상호작용을 모두 생성합니다.

**모든 4원배치.** 선택한 변수에 대해 가능한 4원배치 상호작용을 모두 생성합니다.

**모든 5원배치.** 선택한 변수에 대해 가능한 5원배치 상호작용을 모두 생성합니다.

중첩 항

이 프로시저에서는 모형에 대해 중첩 항을 설정할 수 있습니다. 중첩 항은 요인 또는 공변량의 효과를 모델링하는 데 유용합니다. 이들 값은 다른 요인 수준과 상호작용하지 않습니다. 예를 들어, 식료품 체인점은 여러 점포에서의 고객의 소비 성향을 살펴볼 수 있습니다. 각 고객은 체인점 중의 한 곳만 자주 가기 때문에 고객 효과는 점포 효과 내에 중첩되었다고 할 수 있습니다.

또한 같은 공변량을 포함하고 있는 다항 항 같이 상호작용 효과를 포함시키거나 여러 수준의 중첩을 중첩 항에 추가할 수 있습니다.

**제한.** 중첩 항에는 다음과 같은 제한이 있습니다.

- 상호작용 내의 모든 요인은 고유해야 합니다. 따라서  $A$ 가 요인이면  $A*A$  지정은 유효하지 않습니다.
- 중첩 효과 내의 모든 요인은 고유해야 합니다. 따라서  $A$ 가 요인이면  $A(A)$  지정은 유효하지 않습니다.
- 공변량 내에 효과를 중첩할 수 없습니다. 따라서  $A$ 가 요인이고  $X$ 가 공변량이면  $A$  지정은 유효하지 않습니다.

**절편.** 이 모형에는 대개 절편이 포함됩니다. 데이터가 원점을 전달한다고 가정할 경우에는 절편을 제외시킬 수 있습니다. 모형에 절편을 포함시키는 경우에도 그와 관련된 통계를 표시하지 않도록 선택할 수 있습니다.

---

## 복합 표본 로지스틱 회귀분석 통계

**모형 적합.** 전체 모형 성능을 측정하는 통계의 표시를 제어합니다.

- **유사 R-제곱.** 로지스틱 회귀 모형에는 선형 회귀분석의  $R^2$  통계에 정확하게 상응하는 항목이 없습니다. 대신 이 모형에는  $R^2$  통계의 특성을 유사하게 재현하는 여러 척도값이 있습니다.
- **분류표.** 종속변수에 대한 모형 예측된 범주를 기준으로 하여 관측 범주에 대한 교차 분류표를 표시합니다.

모수. 이 그룹을 사용하면 모형 모수와 관련된 통계의 표시를 제어할 수 있습니다.

- **추정값.** 계수의 추정값을 표시합니다.
- **지수 추정값.** 계수 추정값의 승수로 거듭제곱한 자연 로그의 밑을 표시합니다. 추정값에는 통계를 검정하기에 적합한 특성이 포함되어 있는 반면, 데이터를 해석하기는 지수 추정값인  $\exp$ 가 더 쉽습니다.
- **표준 오차.** 각 계수 추정값에 대한 표준 오차를 표시합니다.
- **신뢰구간.** 각 계수 추정값에 대한 신뢰구간을 표시합니다. 이 구간의 신뢰수준은 옵션 대화 상자에서 설정합니다.
- **T 검정.** 각 계수 추정값에 대한 T 검정을 표시합니다. 각 검정에 대한 귀무가설은 계수 값이 0임을 의미합니다.
- **모수 추정값의 공분산.** 모형 계수에 대한 공분산 행렬의 추정값을 표시합니다.
- **모수 추정값의 상관.** 모형 계수에 대한 상관 행렬의 추정값을 표시합니다.
- **디자인 효과.** 표본이 단순 무작위 표본이라고 가정하여 얻은 분산에 대한 추정값 분산 비율입니다. 복합 디자인 지정으로 인한 효과를 측정하는 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다.
- **디자인 효과 제공근.** 복합 디자인 지정으로 인한 효과를 측정하는 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다.

모형변수의 요약 통계. 종속변수, 공변량 및 요인에 대한 요약 정보를 표시합니다.

표본 디자인 정보. 가중치를 부여하지 않은 개수와 모집단 크기를 비롯하여 표본에 대한 요약 정보를 표시합니다.

---

## 복합 표본 가설 검정

**검정 통계량.** 이 그룹을 사용하면 가설 검정에 사용되는 통계의 유형을 선택할 수 있습니다.  $F$ , 수정된  $F$ , 카이제곱 및 수정된 카이제곱 중에서 선택할 수 있습니다.

**표본추출 자유도.** 이 그룹을 사용하면 모든 검정 통계량에 대한  $p$  값을 계산하는 데 사용되는 표본추출 디자인 자유도를 제어할 수 있습니다. 표본추출 디자인을 기반으로 하는 경우 이 값은 표본추출의 첫 번째 단계의 계층 수 및 기본 표본추출 단위 수 사이의 차이를 가리킵니다. 또는 양의 정수를 지정하여 사용자 정의 자유도를 설정할 수도 있습니다.

**다중 비교에 대한 조정.** 다중 대비를 통해 가설 검정을 수행하는 경우 포함된 대비에 대한 유의 수준에서 전반적인 유의 수준을 조정할 수 있습니다. 이 그룹을 사용하면 조정 방법을 선택할 수 있습니다.

- **최소유의차.** 이 방법은 특정 선형 대비가 귀무가설 값과 다르다는 가설을 거부할 전체 확률을 제어하지 않습니다.

- 순차 *Sidak*(*Sequential Sidak*). 순차 단계별로 낮아지는 거부 *Sidak* 프로시저로, 개별 가설은 거부하지만 동일한 전체 유의 수준을 유지한다는 점에서 훨씬 덜 보수적인 방법입니다.
- 순차 *Bonferroni*(*Sequential Bonferroni*). 순차 단계별로 낮아지는 거부 *Bonferroni* 프로시저로, 개별 가설은 거부하지만 동일한 전체 유의 수준을 유지한다는 점에서 훨씬 덜 보수적인 방법입니다.
- *Sidak*. 이 방법은 *Bonferroni* 접근법보다 더 엄격한 한계를 제공합니다.
- *Bonferroni*. 이 방법은 관측 유의 수준을 다중 대비가 검정되는 요인에 맞춥니다.

## 복합 표본 로지스틱 회귀분석 승산비

승산비 대화 상자를 사용하면 지정된 요인 및 공변량에 대한 모형 추정된 승산비를 표시할 수 있습니다. 참조 범주를 제외하고 종속변수의 각 범주에 대해 개별 승산비 세트가 계산됩니다.

**요인.** 선택된 각 요인에 대해 지정된 참조 범주의 승산을 기준으로 해당 요인의 각 범주에서의 승산비를 표시합니다.

**공변량.** 선택된 각 공변량에 대해 평균의 승산을 기준으로 해당 공변량의 평균과 지정된 단위의 변화량을 더한 값의 승산비를 표시합니다.

요인이나 공변량의 승산비를 계산할 때 이 프로시저에서는 다른 모든 요인을 가장 높은 수준으로 고정하고 다른 모든 공변량을 해당 평균으로 고정합니다. 요인이나 공변량이 모형의 다른 예측변수와 상호작용하는 경우 승산비는 지정된 변수의 변화량뿐만 아니라 해당 요인이나 공변량이 상호작용하는 변수의 값에도 영향을 받습니다. 예를 들어 *age\*age*에서와 같이 지정된 공변량이 모형에서 자체적으로 상호작용하는 경우 승산비는 공변량의 변화와 공변량의 값 모두에 영향을 받습니다.

## 복합 표본 로지스틱 회귀분석 저장

**저장할 변수.** 이 그룹을 사용하면 모형 예측된 범주와 예측된 확률을 활성 데이터 세트에 새 변수로 저장할 수 있습니다.

**IBM SPSS Statistics** 데이터로 모형 내보내기. 모수 추정값, 표준 오차, 유의수준 및 자유도와 함께 모수 상관 행렬이나 공분산 행렬이 포함되어 있는 IBM SPSS Statistics 형식의 데이터 세트를 씁니다. 행렬 파일에서 변수의 순서는 다음과 같습니다.

- **rowtype\_.** 값(및 레이블), COV(공분산), CORR(상관), EST(모수 추정값), SE(표준 오차), SIG(유의 수준) 및 DF(표본추출 디자인 자유도)를 사용합니다. 각 모형 모수에 대해 행 유형 COV (또는 CORR)를 갖는 별도의 케이스는 물론 각기 다른 행 유형에 대해 별도의 케이스가 있습니다.
- **varname\_.** 모수 추정값 테이블에 표시된 모수 문자에 해당하는 값 레이블이 있는 행 유형 COV 또는 CORR에 대해 모든 모형 모수의 순서 목록에 해당하는 값 P1, P2, ...를 사용합니다. 다른 행 유형의 셀은 비어 있습니다.
- **P1, P2, ...** 이러한 변수는 모수 추정값 테이블에 표시된 모수 문자에 해당하는 변수 레이블과 함께 모든 모형 모수의 순서 목록에 해당하며 행 유형에 따라 값을 사용합니다. 중복 모수의 경우 모든 공분산은 0으

로 설정되고 상관은 시스템 결측값으로 설정됩니다. 모든 모수 추정값은 0으로 설정되고 모든 표준 오차, 유의 수준 및 잔차의 자유도는 시스템 결측값으로 설정됩니다.

참고: 이러한 프로시저가 여기에 내보낸 모든 행 유형을 받아 들이지 않는 경우 이 파일은 행렬 파일을 읽는 다른 프로시저에서 추가 분석에 즉시 사용할 수 없습니다.

**XML로 모형 내보내기.** 모수 추정값과 모수 공분산 행렬(선택한 경우)을 XML(PMML) 형식으로 저장합니다. 스코어링을 목적으로 이 모형 파일을 사용하여 모형 정보를 다른 데이터 파일에 적용할 수 있습니다.

---

## 복합 표본 로지스틱 회귀분석 옵션

추정. 이 그룹을 사용하면 모형 추정에 사용되는 여러 기준을 제어할 수 있습니다.

- **최대반복수.** 알고리즘에서 실행할 반복의 최대 횟수입니다. 음수가 아닌 정수를 지정합니다.
- **최대 단계-반분.** 각 반복에서 단계 크기는 로그 우도 증가 또는 최대 단계 반분에 도달할 때까지 요인이 0.5씩 감소됩니다. 양의 정수를 지정합니다.
- **모수 추정값 변화량을 기준으로 반복수 제한.** 이 옵션을 선택하면 모수 추정값의 절대 변화량 또는 상대 변화량이 지정된 값보다 크거나 같아지는 시점에 반복을 멈춥니다. 지정된 값은 0 또는 양의 정수이어야 합니다.
- **로그-우도 변화량을 기준으로 반복수 제한.** 이 옵션을 선택하면 로그-우도 함수의 절대 변화량 또는 상대 변화량이 지정된 값보다 작아지는 시점의 계산 후에 반복을 멈춥니다. 지정된 값은 0 또는 양의 정수이어야 합니다.
- **데이터 점의 완전 분리 확인.** 이 옵션을 선택하면 모수 추정값이 중복되지 않았는지 확인하기 위한 검정이 수행됩니다. 모든 케이스를 올바르게 분류하는 모형을 프로시저에서 생성할 수 있는 경우에 분리가 발생합니다.
- **반복계산과정 표시.** 초기 추정값인 0 번째 반복에서 시작하여  $n$  번째 반복마다 모수 추정값과 통계를 표시합니다. 반복계산과정 인쇄를 선택하면  $n$  값에 관계 없이 항상 마지막 반복이 인쇄됩니다.

**사용자 결측값.** 종속변수와 공변량을 비롯한 모든 계획 변수에는 유효한 데이터가 있어야 합니다. 이러한 변수에 대해 데이터가 하나라도 유효하지 않은 케이스는 분석 대상에서 제외됩니다. 이 제어를 사용하면 계층, 군집, 부모집단 및 요인변수 사이에서 사용자 결측값을 유효한 데이터로 취급할지 여부를 결정할 수 있습니다.

**신뢰구간.** 계수 추정값, 지수화한 계수 추정값 및 승산비에 대한 신뢰구간 수준입니다. 50보다 크거나 같고 100보다 작은 값을 지정합니다.

---

## CSLOGISTIC 명령 추가 기능

명령 구문을 사용하여 수행할 수 있는 추가 기능은 다음과 같습니다.

- **CUSTOM** 하위 명령을 사용하여 효과에 대한 효과 대 선형 조합의 검정이나 값을 지정할 수 있습니다.
- **ODDSRATIOS** 하위 명령을 사용하여 요인 및 공변량에 대한 승산비를 계산할 때 다른 모형 변수의 값을 고정할 수 있습니다.

- CRITERIA 하위 명령을 사용하여 비정칙성 확인을 위한 허용 오차 값을 지정할 수 있습니다.
- SAVE 하위 명령을 사용하여 저장된 변수에 대한 사용자 지정 이름을 만들 수 있습니다.
- PRINT 하위 명령을 사용하여 일반 추정가능 함수 테이블을 생성할 수 있습니다.

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



---

## 제 11 장 복합 표본 순서 회귀분석

복합 표본 순서 회귀분석 프로시저에서는 복합 표본추출 방법으로 구한 표본의 이분형 또는 순서 종속변수에 대한 회귀분석을 수행합니다. 선택적으로 부모집단에 대한 분석을 요청할 수 있습니다.

**예제.** 입법화되기 전에 법안을 준비 중인 국회의원들이 법안에 대해 국민의 지지가 있는지 여부와 법안에 대한 지원이 유권자 수와 관련이 있는지 관심이 있습니다. 여론 설문조사 기관이 복합 표본추출 디자인에 따라 인터뷰를 디자인하고 수행합니다. 복합 표본 순서 회귀분석을 사용하여 유권자 수를 기준으로 지지 수준에 대한 모형을 적합시킬 수 있습니다.

복합 표본 순서 회귀분석 데이터 고려 사항

**데이터.** 종속변수는 순서입니다. 요인은 범주형입니다. 공변량은 종속변수와 관련된 양적변수입니다. 부모집단 변수는 문자 또는 숫자일 수 있지만 범주형이어야 합니다.

**가정.** 데이터 파일의 케이스는 복합 표본 계획 대화 상자에서 선택한 파일의 지정 사항에 따라 분석해야 할 복합 디자인의 표본을 나타냅니다.

복합 표본 순서 회귀분석 얻기

1. 메뉴에서 다음을 선택합니다.

분석 > 복합 표본 > 순서 회귀분석...

2. 계획 파일 선택. 선택적으로 사용자 정의 결합 확률 파일을 선택합니다.

3. 계속을 클릭합니다.

4. 복합 표본 순서 회귀분석 대화 상자에서 종속변수를 선택합니다.

선택적으로 다음을 수행할 수 있습니다.

- 데이터에 따라 적절한 요인 및 공변량의 변수를 선택합니다.
- 부모집단을 정의하는 데 사용할 변수를 지정합니다. 분산이 전체 데이터 세트를 기준으로 적절히 추정되었 다면 분석은 부모집단 변수의 선택된 범주에 대해서만 수행됩니다.
- 연결 함수를 선택합니다.

**연결 함수.** 연결 함수는 누적 확률의 변환으로 모형을 추정할 수 있습니다. 다음 다섯 개의 연결 함수를 사용할 수 있습니다.

- 로짓.  $f(x)=\log(x/(1-x))$ . 일반적으로 분배 범주에 사용됩니다.
- 보 로그-로그.  $f(x)=\log(-\log(1-x))$ . 일반적으로 더 높은 범주가 더 개연성이 있는 경우에 사용됩니다.
- 음수 로그-로그.  $f(x)=-\log(-\log(x))$ . 일반적으로 더 낮은 범주가 더 개연성이 있는 경우에 사용됩니다.
- 프로빗.  $f(x)=\Phi^{-1}(x)$ . 일반적으로 잠복 변수가 정규 분포일 경우 사용됩니다.

- **Cauchit(역 Cauchy)**.  $f(x)=\tan(\pi(x-0.5))$ . 일반적으로 잠복 변수에 많은 극단값이 있는 경우 사용됩니다.

## 복합 표본 순서 회귀분석 반응 확률

반응 확률 대화 상자를 사용하면 반응의 누적 확률(즉, 종속변수의 특정 범주를 포함하는 확률)이 종속변수 값의 증가 또는 감소에 따라 증가할지 여부를 지정할 수 있습니다.

## 복합 표본 순서 회귀 모형

모형 효과 지정. 기본적으로 이 프로시저에서는 기본 대화 상자에 지정된 요인 및 공변량을 사용하여 주효과 모형을 작성합니다. 그러나 상호작용 효과 및 중첩 항이 포함되도록 사용자 정의 모형을 작성할 수도 있습니다.

중첩되지 않은 항

선택된 요인과 공변량에 대해 다음과 같습니다.

**상호작용.** 선택한 모든 변수의 최고 수준 상호작용 항을 생성합니다.

**주효과.** 선택한 각 변수에 대한 주효과 항을 생성합니다.

**모든 2원배치.** 선택한 변수에 대해 가능한 2원배치 상호작용을 모두 생성합니다.

**모든 3원배치.** 선택한 변수에 대해 가능한 3원배치 상호작용을 모두 생성합니다.

**모든 4원배치.** 선택한 변수에 대해 가능한 4원배치 상호작용을 모두 생성합니다.

**모든 5원배치.** 선택한 변수에 대해 가능한 5원배치 상호작용을 모두 생성합니다.

중첩 항

이 프로시저에서는 모형에 대해 중첩 항을 설정할 수 있습니다. 중첩 항은 요인 또는 공변량의 효과를 모델링하는 데 유용합니다. 이들 값은 다른 요인 수준과 상호작용하지 않습니다. 예를 들어, 식료품 체인점은 여러 점포에서의 고객의 소비 성향을 살펴볼 수 있습니다. 각 고객은 체인점 중의 한 곳만 자주 가기 때문에 고객 효과는 점포 효과 내에 중첩되었다고 할 수 있습니다.

또한 같은 공변량을 포함하고 있는 다항 항 같이 상호작용 효과를 포함시키거나 여러 수준의 중첩을 중첩 항에 추가할 수 있습니다.

**제한.** 중첩 항에는 다음과 같은 제한이 있습니다.

- 상호작용 내의 모든 요인은 고유해야 합니다. 따라서 A가 요인이면 A\*A 지정은 유효하지 않습니다.
- 중첩 효과 내의 모든 요인은 고유해야 합니다. 따라서 A가 요인이면 A(A) 지정은 유효하지 않습니다.
- 공변량 내에 효과를 중첩할 수 없습니다. 따라서 A가 요인이고 X가 공변량이면 A 지정은 유효하지 않습니다.



---

## 복합 표본 순서 회귀분석 통계

모형 적합. 전체 모형 성능을 측정하는 통계의 표시를 제어합니다.

- 유사 R-제곱. 순서 회귀 모형에는 선형 회귀분석의  $R^2$  통계에 정확하게 상응하는 항목이 없습니다. 대신 이 모형에는  $R^2$  통계의 특성을 유사하게 재현하는 여러 측도값이 있습니다.
- 분류표. 종속변수에 대한 모형 예측된 범주를 기준으로 하여 관측 범주에 대한 교차 분류표를 표시합니다.

모수. 이 그룹을 사용하면 모형 모수와 관련된 통계의 표시를 제어할 수 있습니다.

- 추정값. 계수의 추정값을 표시합니다.
- 지수 추정값. 계수 추정값의 승수로 거듭제곱한 자연 로그의 밑을 표시합니다. 추정값에는 통계를 검정하기에 적합한 특성이 포함되어 있는 반면, 데이터를 해석하기는 지수 추정값인  $\exp$ 가 더 쉽습니다.
- 표준 오차. 각 계수 추정값에 대한 표준 오차를 표시합니다.
- 신뢰구간. 각 계수 추정값에 대한 신뢰구간을 표시합니다. 이 구간의 신뢰수준은 옵션 대화 상자에서 설정합니다.
- T 검정. 각 계수 추정값에 대한 T 검정을 표시합니다. 각 검정에 대한 귀무가설은 계수 값이 0임을 의미합니다.
- 모수 추정값의 공분산. 모형 계수에 대한 공분산 행렬의 추정값을 표시합니다.
- 모수 추정값의 상관. 모형 계수에 대한 상관 행렬의 추정값을 표시합니다.
- 디자인 효과. 표본이 단순 무작위 표본이라고 가정하여 얻은 분산에 대한 추정값 분산 비율입니다. 복합 디자인 지정으로 인한 효과를 측정하는 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다.
- 디자인 효과 제공근. 복합 디자인 지정으로 인한 효과를 측정하는 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다. 표준 오차의 값과 비교한 단위로 표현한 것입니다.

라인 평행성. 이 그룹을 사용하면 개별 회귀선이 각 반응 범주(마지막 범주는 제외)에 대해 적합된 평행하지 않은 선이 있는 모형과 관련된 통계를 요청할 수 있습니다.

- Wald 검정. 회귀모수가 모든 누적 반응에 대해 동일한 귀무가설의 검정을 생성합니다. 평행하지 않은 선이 있는 모형이 추정되며 같은 모수의 Wald 검정이 적용됩니다.
- 모수 추정값. 평행하지 않은 선이 있는 모형에 대한 계수와 표준 오차의 추정값을 표시합니다.
- 모수 추정값의 공분산. 평행하지 않은 선이 있는 모형의 계수에 대한 공분산 행렬의 추정값을 표시합니다.

모형변수의 요약 통계. 종속변수, 공변량 및 요인에 대한 요약 정보를 표시합니다.

표본 디자인 정보. 가중치를 부여하지 않은 개수와 모집단 크기를 비롯하여 표본에 대한 요약 정보를 표시합니다.

---

## 복합 표본 가설 검정

검정 통계량. 이 그룹을 사용하면 가설 검정에 사용되는 통계의 유형을 선택할 수 있습니다. F, 수정된 F, 카이제곱 및 수정된 카이제곱 중에서 선택할 수 있습니다.

표본추출 자유도. 이 그룹을 사용하면 모든 검정 통계량에 대한  $p$  값을 계산하는 데 사용되는 표본추출 디자인 자유도를 제어할 수 있습니다. 표본추출 디자인을 기반으로 하는 경우 이 값은 표본추출의 첫 번째 단계의 계층 수 및 기본 표본추출 단위 수 사이의 차이를 가리킵니다. 또는 양의 정수를 지정하여 사용자 정의 자유도를 설정할 수도 있습니다.

다중 비교에 대한 조정. 다중 대비를 통해 가설 검정을 수행하는 경우 포함된 대비에 대한 유의 수준에서 전반적인 유의 수준을 조정할 수 있습니다. 이 그룹을 사용하면 조정 방법을 선택할 수 있습니다.

- **최소유의차.** 이 방법은 특정 선형 대비가 귀무가설 값과 다르다는 가설을 거부할 전체 확률을 제어하지 않습니다.
- **순차 Sidak(Sequential Sidak).** 순차 단계별로 낮아지는 거부 Sidak 프로시저로, 개별 가설은 거부하지만 동일한 전체 유의 수준을 유지한다는 점에서 훨씬 덜 보수적인 방법입니다.
- **순차 Bonferroni(Sequential Bonferroni).** 순차 단계별로 낮아지는 거부 Bonferroni 프로시저로, 개별 가설은 거부하지만 동일한 전체 유의 수준을 유지한다는 점에서 훨씬 덜 보수적인 방법입니다.
- **Sidak.** 이 방법은 Bonferroni 접근법보다 더 엄격한 한계를 제공합니다.
- **Bonferroni.** 이 방법은 관측 유의 수준을 다중 대비가 검정되는 요인에 맞춥니다.

---

## 복합 표본 순서 회귀분석 승산비

승산비 대화 상자를 사용하면 지정된 요인 및 공변량에 대한 모형 추정된 누적 승산비를 표시할 수 있습니다. 이 기능은 로짓 연결 함수를 사용하는 모형에만 사용할 수 있습니다. 단일 누적 승산비는 마지막을 제외하고 종속변수의 모든 범주에 대해 계산됩니다. 비례 승산 모형은 모두 같은 것으로 가정합니다.

요인. 선택된 각 요인에 대해 지정된 참조 범주의 승산을 기준으로 해당 요인의 각 범주에서의 누적 승산비를 표시합니다.

공변량. 선택된 각 공변량에 대해 평균의 승산을 기준으로 해당 공변량의 평균과 지정된 단위의 변화량을 더한 값의 누적 승산비를 표시합니다.

요인이나 공변량의 승산비를 계산할 때 이 프로시저에서는 다른 모든 요인을 가장 높은 수준으로 고정하고 다른 모든 공변량을 해당 평균으로 고정합니다. 요인이나 공변량이 모형의 다른 예측변수와 상호작용하는 경우 승산비는 지정된 변수의 변화량뿐만 아니라 해당 요인이나 공변량이 상호작용하는 변수의 값에도 영향을 받습니다. 예를 들어  $age*age$ 에서와 같이 지정된 공변량이 모형에서 자체적으로 상호작용하는 경우 승산비는 공변량의 변화와 공변량의 값 모두에 영향을 받습니다.

---

## 복합 표본 순서 회귀분석 저장

저장할 변수. 이 그룹을 사용하면 모형 예측 범주, 예측 범주의 확률, 누적 확률 및 예측 확률을 활성 데이터 세트의 새 변수로 저장할 수 있습니다.

**IBM SPSS Statistics** 데이터로 모형 내보내기. 모수 추정값, 표준 오차, 유의수준 및 자유도와 함께 모수 상관 행렬이나 공분산 행렬이 포함되어 있는 IBM SPSS Statistics 형식의 데이터 세트를 씁니다. 행렬 파일에서 변수의 순서는 다음과 같습니다.

- **rowtype\_.** 값(및 레이블), COV(공분산), CORR(상관), EST(모수 추정값), SE(표준 오차), SIG(유의 수준) 및 DF(표본추출 디자인 자유도)를 사용합니다. 각 모형 모수에 대해 행 유형 COV (또는 CORR)를 갖는 별도의 케이스는 물론 각기 다른 행 유형에 대해 별도의 케이스가 있습니다.
- **varname\_.** 모수 추정값 테이블에 표시된 모수 문자에 해당하는 값 레이블이 있는 행 유형 COV 또는 CORR에 대해 모든 모형 모수의 순서 목록에 해당하는 값 P1, P2, ...를 사용합니다. 다른 행 유형의 셀은 비어 있습니다.
- **P1, P2, ...** 이러한 변수는 모수 추정값 테이블에 표시된 모수 문자에 해당하는 변수 레이블과 함께 모든 모형 모수의 순서 목록에 해당하며 행 유형에 따라 값을 사용합니다. 중복 모수의 경우 모든 공분산은 0으로 설정되고 상관은 시스템 결측값으로 설정됩니다. 모든 모수 추정값은 0으로 설정되고 모든 표준 오차, 유의 수준 및 잔차의 자유도는 시스템 결측값으로 설정됩니다.

참고: 이러한 프로시저가 여기에 내보낸 모든 행 유형을 받아 들이지 않는 경우 이 파일은 행렬 파일을 읽는 다른 프로시저에서 추가 분석에 즉시 사용할 수 없습니다.

**XML로 모형 내보내기.** 모수 추정값과 모수 공분산 행렬(선택한 경우)을 XML(PMML) 형식으로 저장합니다. 스코어링을 목적으로 이 모형 파일을 사용하여 모형 정보를 다른 데이터 파일에 적용할 수 있습니다.

## 복합 표본 순서 회귀분석 옵션

**추정 방법.** 모수 추정 방법을 선택할 수 있습니다. Newton-Raphson, Fisher 스코어링 또는 Fisher 스코어링 반복이 Newton-Raphson 방법으로 전환하기 전에 수행되는 하이브리드 방법 중에서 선택합니다. 하이브리드 방법의 Fisher 스코어링 단계 동안 Fisher 반복의 최대 수에 도달하기 전에 수렴이 얻어진 경우 알고리즘은 Newton-Raphson 방법으로 계속됩니다.

**추정.** 이 그룹을 사용하면 모형 추정에 사용되는 여러 기준을 제어할 수 있습니다.

- **최대반복수.** 알고리즘에서 실행할 반복의 최대 횟수입니다. 음수가 아닌 정수를 지정합니다.
- **최대 단계-반분.** 각 반복에서 단계 크기는 로그 우도 증가 또는 최대 단계 반분에 도달할 때까지 요인이 0.5씩 감소됩니다. 양의 정수를 지정합니다.
- **모수 추정값 변화량을 기준으로 반복수 제한.** 이 옵션을 선택하면 모수 추정값의 절대 변화량 또는 상대 변화량이 지정된 값보다 크거나 같아지는 시점에 반복을 멈춥니다. 지정된 값은 0 또는 양의 정수이어야 합니다.
- **로그-우도 변화량을 기준으로 반복수 제한.** 이 옵션을 선택하면 로그-우도 함수의 절대 변화량 또는 상대 변화량이 지정된 값보다 작아지는 시점의 계산 후에 반복을 멈춥니다. 지정된 값은 0 또는 양의 정수이어야 합니다.
- **데이터 점의 완전 분리 확인.** 이 옵션을 선택하면 모수 추정값이 중복되지 않았는지 확인하기 위한 검정이 수행됩니다. 모든 케이스를 올바르게 분류하는 모형을 프로시저에서 생성할 수 있는 경우에 분리가 발생합니다.

- 반복계산과정 표시. 초기 추정값인 0 번째 반복에서 시작하여  $n$  번째 반복마다 모수 추정값과 통계를 표시합니다. 반복계산과정 인쇄를 선택하면  $n$  값에 관계 없이 항상 마지막 반복이 인쇄됩니다.

**사용자 결측값.** 종속변수와 공변량을 비롯한 척도 계획 변수에는 유효한 데이터가 있어야 합니다. 이러한 변수에 대해 데이터가 하나라도 유효하지 않은 케이스는 분석 대상에서 제외됩니다. 이 제어를 사용하면 계층, 군집, 부모집단 및 요인변수 사이에서 사용자 결측값을 유효한 데이터로 취급할지 여부를 결정할 수 있습니다.

**신뢰구간.** 계수 추정값, 지수화한 계수 추정값 및 승산비에 대한 신뢰구간 수준입니다. 50보다 크거나 같고 100보다 작은 값을 지정합니다.

## CSORDINAL 명령 추가 기능

명령 구문을 사용하여 수행할 수 있는 추가 기능은 다음과 같습니다.

- CUSTOM 하위 명령을 사용하여 효과에 대한 효과 대 선형 조합의 검정이나 값을 지정할 수 있습니다.
- ODDS RATIOS 하위 명령을 사용하여 요인 및 공변량에 대한 누적 승산비를 계산할 때 평균 이외의 값에서 다른 모형 변수의 값을 고정할 수 있습니다.
- ODDS RATIOS 하위 명령을 사용하여 승산비를 요청할 때 요인에 대한 사용자 정의 참조 범주로 레이블을 붙이지 않은 값을 사용합니다.
- CRITERIA 하위 명령을 사용하여 비정칙성 확인을 위한 허용 오차 값을 지정할 수 있습니다.
- PRINT 하위 명령을 사용하여 일반 추정가능 함수 테이블을 생성할 수 있습니다.
- SAVE 부명령문을 사용하여 25개 이상의 확률 변수를 저장합니다.

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

---

## 제 12 장 복합 표본 Cox 회귀분석

복합 표본 Cox 회귀분석 프로시저는 복합 표본추출 방법으로 구한 생존 분석을 수행합니다. 선택적으로 부모 집단에 대한 분석을 요청할 수 있습니다.

**예제.** 정부 법률 집행 기관은 해당 관할 구역에서 재범 비율에 대해 관심을 갖고 있습니다. 재범을 측정하는 방법 중 하나는 범죄자가 두 번째 체포될 때까지의 시간입니다. 당국은 Cox 회귀분석을 사용하여 다시 체포되는 시간을 모델링하고 싶어하지만 비례 위험 가정이 연령 범주에서 잘못될지 모른다고 우려했습니다.

의료 연구자들은 허혈 후 발작 기능 회복 프로그램을 마친 환자의 생존 시간을 조사하고 있습니다. 상당한 비 사망 이벤트가 있고 이러한 이벤트 시간이 기록되면 환자 내역은 변하기 때문에 개체당 여러 케이스가 있을 가능성이 있습니다. 허혈 발작 시에 위험이 시작되고 기능 회복 훈련 프로그램을 거쳐 생존하는 환자만 표본에 있기 때문에 표본은 관측된 생존 시간이 기능 회복 훈련의 길이로 "부풀려" 진다는 의미에서 왼쪽으로 잘립니다.

복합 표본 Cox 회귀분석 데이터 고려 사항

**생존 시간.** 프로시저는 이벤트가 발생하기 전의 기간인 생존 시간 분석에 Cox 회귀분석을 적용합니다. 구간 시작 시간에 따라 생존 시간을 지정하는 방법은 두 가지가 있습니다.

- **시간=0.** 일반적으로 각 개체에 대한 구간이 시작될 때 완전한 정보를 갖게 되며 종료 시간을 포함하는 변수를 갖게 됩니다(또는 날짜 및 시간변수에서 종료 시간으로 단일 변수를 만듭니다).
- **개체별로 다릅니다.** 이것은 지연된 입력이라고도 하는 왼쪽 잘림이 있을 때 적절합니다. 발작 후 기능 회복 훈련 프로그램을 떠나는 환자에 대한 생존 시간을 분석하는 경우 위험이 발작할 때 시작된다는 것을 생각할 수 있습니다. 그러나 표본에 기능 회복 훈련 프로그램에서 생존한 환자만 포함되어 있는 경우 표본은 관측된 생존 시간이 기능 회복 훈련에 의해 "부풀려" 진다는 의미에서 왼쪽으로 잘립니다. 연구에 입력하는 시간으로 기능 회복 훈련을 종료한 시간을 지정하여 이것을 설명할 수 있습니다.

**날짜 및 시간변수.** 날짜 및 시간변수는 구간의 시작과 종료를 직접 정의하는 데 사용할 수 있습니다. 날짜 및 시간변수가 있는 경우 이것을 사용하여 생존 시간이 포함되어 있는 변수를 만들어야 합니다. 왼쪽 잘림이 없는 경우 연구에 입력하는 날짜와 관측 날짜 사이의 차이를 기반으로 종료 시간이 포함되어 있는 변수를 만듭니다. 왼쪽 잘림이 있는 경우 연구 시작 날짜와 입력 날짜 사이의 차이를 기준으로 시작 시간이 포함되어 있는 변수 및 연구 시작 날짜와 관측 날짜 사이의 차이를 기준으로 종료 시간이 포함되어 있는 변수를 만듭니다.

**이벤트 상태.** 개체가 구간 내에서 중요 이벤트를 경험했는지 여부를 기록하는 변수가 필요합니다. 이벤트가 발생하지 않은 개체는 오른쪽-중도절단됩니다.

**개체 식별자.** 여러 케이스에서 단일 개체에 대한 관측을 분할하여 조각별-상수, 시간 종속 예측변수를 쉽게 통합할 수 있습니다. 예를 들어, 환자 발작 후 생존 시간을 분석하는 경우 의료 기록을 나타내는 변수를 예측변수로 사용해야 합니다. 시간이 지남에 따라 환자들은 의료 기록 후에 주요 의료 이벤트를 경험할 수 있습니다.



다음 표에서는 다음과 같은 데이터 세트를 구조화하는 방법을 표시합니다. 환자 ID는 개체 식별자이고, 종료 시간은 관측된 구간을 정의하고, 상태는 주요 의료 이벤트를 기록하고 심장 마비 사전 기록과 출혈 이전 기록은 조각별-상수, 시간 종속 예측변수입니다.

표 1. 조각별 상수, 시간 종속 예측변수를 통합하기 위한 데이터 구조.

환자 ID	종료 시간	상태	심장 마비 사전 기록	출혈 사전 기록
1	5	심장 마비	아니요	아니요
1	7	출혈	예	아니요
1	8	사망	예	예
2	24	사망	아니요	아니요
3	8	심장 마비	아니요	아니요
3	15	사망	예	아니요

가정. 데이터 파일의 케이스는 복합 표본 계획 대화 상자에서 선택한 파일의 지정 사항에 따라 분석해야 할 복합 디자인의 표본을 나타냅니다.

일반적으로 Cox 회귀 모형은 비례 위험을 가정합니다. 즉 한 케이스에서 다른 케이스까지의 위험 비율은 시간에 따라 변하지 말아야 합니다. 이 가정을 하지 않을 경우 모형에 시간 종속 예측변수를 추가해야 할 수 있습니다.

**Kaplan-Meier 분석.** 예측변수를 선택하지 않고(또는 선택한 예측변수를 모형에 입력하지 않는 경고) 옵션 탭에서 기준선 생존 곡선을 계산하기 위한 곱 한계값 방법을 선택하는 경우 프로시저는 생존 분석의 Kaplan-Meier 유형을 수행합니다.

복합 표본 Cox 회귀분석을 얻으려면

1. 메뉴에서 다음을 선택합니다.

분석 > 복합 표본 > Cox 회귀분석...

2. 계획 파일 선택. 선택적으로 사용자 정의 결합 확률 파일을 선택합니다.

3. 계속을 클릭합니다.

4. 연구에서 시작과 종료 시간을 선택하여 생존 시간을 지정합니다.

5. 이벤트 상태변수를 선택합니다.

6. 이벤트 정의를 클릭하고 최소 하나의 이벤트 값을 정의합니다.

선택적으로 개체 식별자를 선택할 수 있습니다.

## 이벤트 정의

종료 이벤트가 발생했음을 나타내는 값을 지정합니다.

- 개별값. 격자에 입력하거나 값 레이블이 정의된 값 목록에서 선택하여 하나 이상의 값을 지정합니다.
- 값의 범위. 최소 및 최대값을 입력하거나 값 레이블이 정의된 목록에서 값을 선택하여 값의 범위를 지정합니다.

---

## 예측변수

예측변수 탭을 사용하면 모형 효과를 작성하는 데 사용되는 요인과 공변량을 지정할 수 있습니다.

**요인.** 요인은 범주형 예측변수로 숫자 또는 문자가 될 수 있습니다.

**공변량.** 공변량은 척도 예측변수로 숫자가 되어야 합니다.

**시간 종속 예측변수.** 비례 위험을 가정하지 않는 특정 상황이 있습니다. 즉, 위험 비율은 시간이 지남에 따라 변경되며 하나 또는 그 이상의 예측변수에 대한 여러 값이 서로 다른 시점에서 서로 다르게 나타납니다. 이런 경우 시간 종속 예측변수를 지정해야 합니다. 자세한 정보는 『시간 종속 예측변수 정의』 주제를 참조하십시오. 시간 종속 예측변수는 요인이나 공변량으로 선택할 수 있습니다.

### 시간 종속 예측변수 정의

시간 종속 예측변수 정의 대화 상자를 사용하면 내장된 시간변수,  $T_$ 에 종속되는 예측변수를 만들 수 있습니다. 이 변수를 사용하여 다음의 두 가지 일반적 방법으로 시간-종속 공변량을 정의할 수 있습니다.

- 비례적이지 않은 위험을 허용하는 확장된 Cox 회귀 모형을 추정하려는 경우 시간 종속 예측변수를 해당 시간변수  $T_$ 와 공변량의 함수로 정의하여 추정할 수 있습니다. 일반적 예는 시간변수와 예측변수의 단순 곱이지만 더 복잡한 함수도 지정할 수 있습니다.
- 일부 변수는 다른 시간 주기에서 다른 값을 가질 수 있으나 시간에 대해 체계적인 관계가 없습니다. 이러한 경우 논리 표현식을 사용하여 수행할 수 있는 구간별 시간 종속 예측변수를 정의해야 합니다. 논리 표현식은 참인 경우 1의 값을, 거짓인 경우 0의 값을 취합니다. 일련의 논리 표현식을 사용하면 여러 측정에서 시간 종속 예측변수를 작성할 수 있습니다. 예를 들어, 4주동안 일주일에 한 번씩 혈압을 측정할 경우( $BP1$ 에서  $BP4$ 로 나타냄) 시간 종속 예측변수를  $(T_ < 1) * BP1 + (T_ \geq 1 \ \& \ T_ < 2) * BP2 + (T_ \geq 2 \ \& \ T_ < 3) * BP3 + (T_ \geq 3 \ \& \ T_ < 4) * BP4$ 로 정의할 수 있습니다. 괄호로 묶인 항 중 정확히 한 개만 지정된 케이스에 대해 1과 같고 나머지는 0이 된다는 점에 유의하십시오. 이 함수는 시간이 1주일 미만이면  $BP1$ 을 사용하고, 1주일이 넘고 2주일 미만이면  $BP2$ 를 사용하고, 나머지도 마찬가지로 사용합니다.

참고: 구간별 시간 종속 예측변수가 세그먼트 내에서 일정한 경우 위에서 제공한 혈압 예제에서처럼 여러 케이스로 개체를 분할하여 조각별-상수, 시간 종속 예측변수를 지정하는 것이 쉬울 수 있습니다. 자세한 정보는 49 페이지의 제 12 장 『복합 표본 Cox 회귀분석』의 개체 식별자에 대한 설명을 참조하십시오.

시간 종속 예측변수 정의 대화 상자에서 함수-설정 제어를 사용하여 시간-종속 공변량의 표현식을 설정하거나 숫자 표현식 입력란에 직접 입력할 수 있습니다. 문자 상수는 따옴표나 어포스트로피로 묶어야 하며 숫자 상수는 점을 소수점 구분자로 사용하는 미국식으로 입력해야 합니다. 결과변수에는 사용자가 지정하는 이름이 지정되고 예측변수 탭에 요인 또는 공변량으로 포함되어야 합니다.

---

## 하위 그룹

**기준선 계층.** 별도의 기준선 위험과 생존 함수가 이 변수의 각 값에 대해 계산되지만 계층에서는 모형 계수의 단일 세트가 추정됩니다.

부모집단 변수. 부모집단을 정의하는 데 사용할 변수를 지정합니다. 부모집단 변수의 선택된 범주에 대해서만 분석이 수행됩니다.

---

## 모형

모형 효과 지정. 기본적으로 이 프로시저에서는 기본 대화 상자에 지정된 요인 및 공변량을 사용하여 주효과 모형을 작성합니다. 그러나 상호작용 효과 및 중첩 항이 포함되도록 사용자 정의 모형을 작성할 수도 있습니다.

중첩되지 않은 항

선택된 요인과 공변량에 대해 다음과 같습니다.

상호작용. 선택한 모든 변수의 최고 수준 상호작용 항을 생성합니다.

주효과. 선택한 각 변수에 대한 주효과 항을 생성합니다.

모든 2원배치. 선택한 변수에 대해 가능한 2원배치 상호작용을 모두 생성합니다.

모든 3원배치. 선택한 변수에 대해 가능한 3원배치 상호작용을 모두 생성합니다.

모든 4원배치. 선택한 변수에 대해 가능한 4원배치 상호작용을 모두 생성합니다.

모든 5원배치. 선택한 변수에 대해 가능한 5원배치 상호작용을 모두 생성합니다.

중첩 항

이 프로시저에서는 모형에 대해 중첩 항을 설정할 수 있습니다. 중첩 항은 요인 또는 공변량의 효과를 모델링하는 데 유용합니다. 이들 값은 다른 요인 수준과 상호작용하지 않습니다. 예를 들어, 식료품 체인점은 여러 점포에서의 고객의 소비 성향을 살펴볼 수 있습니다. 각 고객은 체인점 중의 한 곳만 자주 가기 때문에 고객 효과는 점포 효과 내에 중첩되었다고 할 수 있습니다.

또한 같은 공변량을 포함하고 있는 다항 항 같이 상호작용 효과를 포함시키거나 여러 수준의 중첩을 중첩 항에 추가할 수 있습니다.

제한. 중첩 항에는 다음과 같은 제한이 있습니다.

- 상호작용 내의 모든 요인은 고유해야 합니다. 따라서  $A$ 가 요인이면  $A*A$  지정은 유효하지 않습니다.
- 중첩 효과 내의 모든 요인은 고유해야 합니다. 따라서  $A$ 가 요인이면  $A(A)$  지정은 유효하지 않습니다.
- 공변량 내에 효과를 중첩할 수 없습니다. 따라서  $A$ 가 요인이고  $X$ 가 공변량이면  $A$  지정은 유효하지 않습니다.



---

## 통계

표본 디자인 정보. 가중치를 부여하지 않은 개수와 모집단 크기를 비롯하여 표본에 대한 요약 정보를 표시합니다.

이벤트 및 중도절단 요약. 중도절단 케이스의 수와 백분율에 대한 요약 정보를 표시합니다.

이벤트 시간의 위험 세트. 각 기준선 계층에서 각 이벤트 시간에 대한 이벤트 수와 위험 수를 표시합니다.

모수. 이 그룹을 사용하면 모형 모수와 관련된 통계의 표시를 제어할 수 있습니다.

- 추정값. 계수의 추정값을 표시합니다.
- 지수 추정값. 계수 추정값의 승수로 거듭제곱한 자연 로그의 밑을 표시합니다. 추정값에는 통계를 검정하기에 적합한 특성이 포함되어 있는 반면, 데이터를 해석하기는 지수 추정값인  $\exp$ 가 더 쉽습니다.
- 표준 오차. 각 계수 추정값에 대한 표준 오차를 표시합니다.
- 신뢰구간. 각 계수 추정값에 대한 신뢰구간을 표시합니다. 이 구간의 신뢰수준은 옵션 대화 상자에서 설정합니다.
- T 검정. 각 계수 추정값에 대한 T 검정을 표시합니다. 각 검정에 대한 귀무가설은 계수 값이 0임을 의미합니다.
- 모수 추정값의 공분산. 모형 계수에 대한 공분산 행렬의 추정값을 표시합니다.
- 모수 추정값의 상관. 모형 계수에 대한 상관 행렬의 추정값을 표시합니다.
- 디자인 효과. 표본이 단순 무작위 표본이라고 가정하여 얻은 분산에 대한 추정값 분산 비율입니다. 복합 디자인 지정으로 인한 효과를 측정하는 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다.
- 디자인 효과 제공근. 복합 디자인 지정으로 인한 효과를 측정하는 것으로 값이 1과 차이가 많이 날수록 효과가 큰 것을 의미합니다.

모형 가정. 이 그룹을 사용하면 비례 위험 가정의 검정을 생성할 수 있습니다. 검정은 적합 모형을 각 예측변수  $x$ 에 대해 시간 종속 예측변수  $x*_TF$ 를 포함하는 대체 모형과 비교합니다. 여기서  $_TF$ 는 지정된 시간 함수입니다.

- 시간 함수. 대체 모형에 대한  $_TF$ 의 형식을 지정합니다. 항등 함수의 경우  $_TF=T$ 입니다. 로그 함수의 경우  $_TF=\log(T)$ 입니다. **Kaplan-Meier**의 경우  $_TF=1-S_{KM}(T)$ 이며  $S_{KM}(\cdot)$ 은 생존 함수의 Kaplan-Meier 추정값입니다. 순위의 경우  $_TF$ 는 관측된 종료 시간 중  $T$ 의 순위입니다.
- 대체 모형에 대한 모수 추정값. 대체 모형에 있는 각 모수에 대한 추정값, 표준 오차 및 신뢰구간을 표시합니다.
- 대체 모형의 공분산 행렬. 대체 모형에 있는 모수 사이의 추정된 공분산 행렬을 표시합니다.

기준선 생존 및 누적 위험 함수. 기준선 생존 함수와 기준 누적 위험 함수를 표준 오차와 함께 표시합니다.

참고: 예측변수 탭에서 정의한 시간 종속 예측변수가 모형에 포함된 경우 이 옵션은 사용할 수 없습니다.

---

## 도표

도표 탭을 사용하면 위험 함수, 생존 함수, 생존 함수의 로그-로그 및 1-생존 함수의 도표를 요청할 수 있습니다. 지정된 함수와 함께 신뢰구간 도표를 선택할 수도 있습니다. 신뢰수준은 옵션 탭에서 설정합니다.

**예측변수 패턴.** 요청한 도표와 내보내기 탭에서 내보낸 생존 파일에 사용할 예측변수 값의 패턴을 지정할 수 있습니다. 예측변수 탭에 정의한 시간 종속 예측변수가 모형에 포함된 경우 이러한 옵션은 사용할 수 없습니다.

- **도표 요인.** 기본적으로 각 요인은 가장 높은 수준에서 계산됩니다. 원하는 경우 다른 수준을 입력하거나 선택합니다. 또는 해당 요인의 확인란을 선택하여 단일 요인의 각 수준마다 별도의 선을 도표로 그리도록 선택할 수 있습니다.
- **도표 공변량.** 각 공변량은 평균으로 계산됩니다. 원하는 경우 다른 값을 입력하거나 선택합니다.

---

## 가설 검정

**검정 통계량.** 이 그룹을 사용하면 가설 검정에 사용되는 통계의 유형을 선택할 수 있습니다.  $F$ , 수정된  $F$ , 카이제곱 및 수정된 카이제곱 중에서 선택할 수 있습니다.

**표본추출 자유도.** 이 그룹을 사용하면 모든 검정 통계량에 대한  $p$  값을 계산하는 데 사용되는 표본추출 디자인 자유도를 제어할 수 있습니다. 표본추출 디자인을 기반으로 하는 경우 이 값은 표본추출의 첫 번째 단계의 계층 수 및 기본 표본추출 단위 수 사이의 차이를 가리킵니다. 또는 양의 정수를 지정하여 사용자 정의 자유도를 설정할 수도 있습니다.

**다중 비교에 대한 조정.** 다중 대비를 통해 가설 검정을 수행하는 경우 포함된 대비에 대한 유의 수준에서 전반적인 유의 수준을 조정할 수 있습니다. 이 그룹을 사용하면 조정 방법을 선택할 수 있습니다.

- **최소유의차.** 이 방법은 특정 선형 대비가 귀무가설 값과 다르다는 가설을 거부할 전체 확률을 제어하지 않습니다.
- **순차 Sidak(Sequential Sidak).** 순차 단계별로 낮아지는 거부 Sidak 프로시저로, 개별 가설은 거부하지만 동일한 전체 유의 수준을 유지한다는 점에서 훨씬 덜 보수적인 방법입니다.
- **순차 Bonferroni(Sequential Bonferroni).** 순차 단계별로 낮아지는 거부 Bonferroni 프로시저로, 개별 가설은 거부하지만 동일한 전체 유의 수준을 유지한다는 점에서 훨씬 덜 보수적인 방법입니다.
- **Sidak.** 이 방법은 Bonferroni 접근법보다 더 엄격한 한계를 제공합니다.
- **Bonferroni.** 이 방법은 관측 유의 수준을 다중 대비가 검정되는 요인에 맞춥니다.

---

## 저장

**저장할 변수.** 이 그룹을 사용하면 결과의 진단과 보고에 사용하기 위해 모형 관련 변수를 활성 데이터 세트에 저장할 수 있습니다. 시간 종속 예측변수가 모형에 포함되어 있으면 아무 것도 사용할 수 없습니다.

- **생존 함수.** 각 케이스에 대한 관측된 시간과 예측변수 값에서 생존 확률(생존 함수의 값)을 저장합니다.

- 생존 함수에 대한 신뢰구간의 하한. 각 케이스에 대한 관측된 시간과 예측변수 값에서 생존 함수에 대한 신뢰구간의 하한을 저장합니다.
- 생존 함수에 대한 신뢰구간의 상한. 각 케이스에 대한 관측된 시간과 예측변수 값에서 생존 함수에 대한 신뢰구간의 상한을 저장합니다.
- 누적 위험 함수. 각 케이스에 대한 관측된 시간과 예측변수 값에서 누적 위험 또는  $-\ln(\text{생존})$ 을 저장합니다.
- 누적 위험 함수에 대한 신뢰구간의 하한. 각 케이스에 대한 관측된 시간과 예측변수 값에서 누적 위험 함수에 대한 신뢰구간의 하한을 저장합니다.
- 누적 위험 함수에 대한 신뢰구간의 상한. 각 케이스에 대한 관측된 시간과 예측변수 값에서 누적 위험 함수에 대한 신뢰구간의 상한을 저장합니다.
- 선형 예측변수의 예측값. 참조 값 수정 예측변수를 회귀계수와 곱한 선형결합을 저장합니다. 선형 예측변수는 위험 함수와 기준선 위험 함수의 비율입니다. 비례 위험 모형에서 이 값은 시간에 따라 일정합니다.
- Schoenfeld 잔차. 모형에서 각 중도절단되지 않은 케이스와 각 중복되지 않은 모수에 대해 Schoenfeld 잔차는 모형 모수와 관련된 예측변수의 관측된 값과 관측된 이벤트 시간에서 위험 세트에 있는 케이스에 대한 예측변수의 예상값 사이의 차이입니다. Schoenfeld 잔차는 비례 위험 가정을 평가하는 데 사용할 수 있습니다. 예를 들어, 예측변수  $x$ 의 경우 시간 종속 예측변수  $x \cdot \ln(T_)$  대 시간에 대한 Schoenfeld 잔차의 도표는 비례 위험이 있는 경우 0에서 수평선을 표시해야 합니다. 모형에 있는 각 중복되지 않은 모수에 대해 개별 변수가 저장됩니다. Schoenfeld 잔차는 중도절단되지 않은 케이스에 대해서만 계산됩니다.
- Martingale 잔차. 각 케이스에 대해 martingale 잔차는 관측된 중도절단(중도절단된 경우 0, 아닌 경우 1)과 관측 시간 동안 이벤트의 기대 사이의 차이입니다.
- 편차 잔차. 편차 잔차는 0에 대해 보다 대칭적으로 나타내기 위해 "조정된" martingale 잔차입니다. 예측변수와 비교하여 편차 잔차의 도표는 패턴 없이 표시되어야 합니다.
- Cox-Snell 잔차. 각 케이스에 대해 Cox-Snell 잔차는 관측 시간 동안의 이벤트 기대 또는 관측된 중도절단 - martingale 잔차입니다.
- 스코어 잔차. 모형에 있는 각 케이스와 각 중복되지 않은 모수에 대해 스코어 잔차는 유사-우도의 첫 도함수에 대한 케이스의 연속입니다. 모형에 있는 각 중복되지 않은 모수에 대해 개별 변수가 저장됩니다.
- DFBeta 잔차. 모형에 있는 각 케이스와 각 중복되지 않은 모수에 대해 DFBeta 잔차는 케이스를 모형에서 제거할 때 모수 추정값의 대략적인 변화입니다. 상대적으로 큰 DFBeta 잔차를 갖는 케이스는 분석에서 불필요한 영향력을 발휘할 수 있습니다. 모형에 있는 각 중복되지 않은 모수에 대해 개별 변수가 저장됩니다.
- 통합 잔차. 여러 케이스가 단일 개체를 나타낼 때 개체에 대한 통합 잔차는 같은 개체에 속하는 모든 케이스에서 대응하는 케이스 잔차의 합입니다. Schoenfeld 잔차의 경우 통합 버전은 Schoenfeld 잔차가 중도절단되지 않은 케이스에 대해서만 정의되기 때문에 통합되지 않은 버전과 동일합니다. 이러한 잔차는 개체 식별자가 시간 및 이벤트 탭에 지정되었을 때만 사용할 수 있습니다.

저장된 변수의 이름. 자동 이름 생성은 모든 작업을 유지할 수 있도록 해줍니다. 사용자 정의 이름을 사용하면 데이터 편집기에 저장된 변수를 먼저 삭제하지 않고 이전 실행의 결과를 취소/대체할 수 있습니다.

---

## 내보내기

**IBM SPSS Statistics** 데이터로 모형 내보내기. 모수 추정값, 표준 오차, 유의수준 및 자유도와 함께 모수 상관 행렬이나 공분산 행렬이 포함되어 있는 IBM SPSS Statistics 형식의 데이터 세트를 씁니다. 행렬 파일에서 변수의 순서는 다음과 같습니다.

- **rowtype\_.** 값(및 레이블), COV(공분산), CORR(상관), EST(모수 추정값), SE(표준 오차), SIG(유의 수준) 및 DF(표본추출 디자인 자유도)를 사용합니다. 각 모형 모수에 대해 행 유형 COV (또는 CORR)를 갖는 별도의 케이스는 물론 각기 다른 행 유형에 대해 별도의 케이스가 있습니다.
- **varname\_.** 모수 추정값 테이블에 표시된 모수 문자에 해당하는 값 레이블이 있는 행 유형 COV 또는 CORR에 대해 모든 모형 모수의 순서 목록에 해당하는 값 P1, P2, ...를 사용합니다. 다른 행 유형의 셀은 비어 있습니다.
- **P1, P2, ...** 이러한 변수는 모수 추정값 테이블에 표시된 모수 문자에 해당하는 변수 레이블과 함께 모든 모형 모수의 순서 목록에 해당하며 행 유형에 따라 값을 사용합니다. 중복 모수의 경우 모든 공분산은 0으로 설정되고 상관은 시스템 결측값으로 설정됩니다. 모든 모수 추정값은 0으로 설정되고 모든 표준 오차, 유의 수준 및 잔차의 자유도는 시스템 결측값으로 설정됩니다.

참고: 이러한 프로시저가 여기에 내보낸 모든 행 유형을 받아 들이지 않는 경우 이 파일은 행렬 파일을 읽는 다른 프로시저에서 추가 분석에 즉시 사용할 수 없습니다.

생존 함수를 **IBM SPSS Statistics** 데이터로 내보내기. 생존 함수, 생존 함수의 표준 오차, 생존 함수의 신뢰구간의 상한과 하한, 도표 탭에 지정된 예측변수 패턴과 기준에서 계산된 각 실패나 이벤트 시간에 대한 누적 위험 함수가 포함되어 있는 IBM SPSS Statistics 형식의 데이터 세트를 씁니다. 행렬 파일에서 변수의 순서는 다음과 같습니다.

- 기준선 총화변수. 별도의 생존표가 총화변수의 각 값에 대해 생성됩니다.
- 생존 시간변수. 이벤트 시간, 별도의 케이스가 각 고유 이벤트 시간에 대해 만들어집니다.
- **Sur\_0, LCL\_Sur\_0, UCL\_Sur\_0.** 기준선 생존 함수 및 신뢰구간의 상한과 하한.
- **Sur\_R, LCL\_Sur\_R, UCL\_Sur\_R.** "참조" 패턴(결과에서 패턴 값 테이블 참조)에서 계산된 생존 함수 및 신뢰구간의 상한과 하한.
- **Sur\_##, LCL\_Sur\_##, UCL\_Sur\_##, ...** 도표 탭에 지정된 각 예측변수 패턴에서 계산된 생존 함수 및 신뢰구간의 상한과 하한. 번호가 ##인 패턴과 일치하는 결과의 패턴 값 테이블을 참조하십시오.
- **Haz\_0, LCL\_Haz\_0, UCL\_Haz\_0.** 기준선 누적 위험 함수 및 신뢰구간의 상한과 하한.
- **Haz\_R, LCL\_Haz\_R, UCL\_Haz\_R.** "참조" 패턴(결과에서 패턴 값 테이블 참조)에서 계산된 누적 위험 함수 및 신뢰구간의 상한과 하한.
- **Haz\_##, LCL\_Haz\_##, UCL\_Haz\_##, ...** 도표 탭에 지정된 각 예측변수 패턴에서 계산된 누적 위험 함수 및 신뢰구간의 상한과 하한. 번호가 ##인 패턴과 일치하는 결과의 패턴 값 테이블을 참조하십시오.

**XML로 모형 내보내기** 모수 추정값과 기준선 생존 함수를 포함하여 생존 함수를 예측하는 데 필요한 모든 정보를 XML(PMML) 형식으로 내보냅니다. 스코어링을 목적으로 이 모형 파일을 사용하여 모형 정보를 다른 데이터 파일에 적용할 수 있습니다.

---

## 옵션

추정. 이러한 제어는 회귀계수의 추정 기준을 지정합니다.

- **최대반복수.** 알고리즘에서 실행할 반복의 최대 횟수입니다. 음수가 아닌 정수를 지정합니다.
- **최대 단계-반분.** 각 반복에서 단계 크기는 로그 우도 증가 또는 최대 단계 반분에 도달할 때까지 요인이 0.5씩 감소됩니다. 양의 정수를 지정합니다.
- **모수 추정값 변화량을 기준으로 반복수 제한.** 이 옵션을 선택하면 모수 추정값의 절대 변화량 또는 상대 변화량이 지정된 값보다 크거나 같아지는 시점에 반복을 멈춥니다. 지정된 값은 0 또는 양의 정수이어야 합니다.
- **로그-우도 변화량을 기준으로 반복수 제한.** 이 옵션을 선택하면 로그-우도 함수의 절대 변화량 또는 상대 변화량이 지정된 값보다 작아지는 시점의 계산 후에 반복을 멈춥니다. 지정된 값은 0 또는 양의 정수이어야 합니다.
- **반복계산과정 표시.** 모수 추정값과 유사 로그 우도에 대한 반복계산과정을 표시하고 모수 추정값과 유사 로그 우도에서 변화의 마지막 평가를 인쇄합니다. 반복계산과정 테이블은 0번째 반복(초기 추정값)에서 시작하여  $n$ 번째 반복마다 인쇄됩니다.  $n$ 은 증분 값입니다. 반복계산과정을 요청하는 경우  $n$ 에 관계 없이 마지막 반복은 항상 표시됩니다.
- **모수 추정에 대한 동물 차분.** 관측된 실패 시간이 연관되어 있을 때 이러한 방법 중 하나는 동물을 차분하는 데 사용됩니다. Efron 방법은 계산 비용이 더 많이 듭니다.

생존 함수. 이러한 제어는 생존 함수와 관련된 계산 기준을 지정합니다.

- **기준선 생존 함수를 추정하는 방법.** Breslow (또는 Nelson-Aalan 또는 경험) 방법은 관측된 실패 시간에 단계별로 감소하지 않는 단계 함수에 의해 기준 누적 위험을 추정한 다음 관계 생존= $\exp(-\text{누적 위험})$ 에 의해 기준선 생존을 계산합니다. Efron 방법은 계산 비용이 더 많이 들며 동물이 없을 때는 Breslow 방법으로 줄어듭니다. 곱 한계값 방법은 증가하지 않는 오른쪽 연속형 함수에 의해 기준선 생존을 추정하며 모형에 예측변수가 없을 때는 이 방법은 Kaplan-Meier 추정으로 줄어듭니다.
- **생존 함수의 신뢰구간.** 신뢰구간은 원래 단위, 로그 변환을 통해 또는 로그-로그 변환 등 세 가지 방법으로 계산됩니다. 로그-로그 변환만이 신뢰구간의 한계가 0과 1 사이에 있도록 보증하지만 로그 변환이 일반적으로 "가장 잘" 수행되는 것으로 간주되고 있습니다.

사용자 결측값. 케이스가 분석에 포함되려면 모든 변수는 유효한 값을 가져야 합니다. 이러한 제어를 사용하면 사용자 결측값이 범주형 모형(요인, 이벤트, 계층 및 부모집단 변수 포함) 및 표본추출 계획 변수 사이에서 유효한 것으로 간주할지 여부를 결정할 수 있습니다.

신뢰구간(%). 이것은 계수 추정값, 지수화한 계수 추정값, 생존 함수 추정값 및 누적 위험 함수 추정값에 사용되는 신뢰구간 수준입니다. 0보다 크거나 같고 100보다 작은 값을 지정합니다.

---

## CSCOXREG 명령 추가 기능

명령을 사용하여 수행할 수 있는 추가 기능은 다음과 같습니다.

- 사용자 정의 가설 검정(CUSTOM 하위 명령 및 /PRINT LMATRIX 사용)을 수행합니다.
- 허용 오차 지정(/CRITERIA SINGULAR 사용).
- 일반 추정가능 함수 테이블(/PRINT GEF 사용).
- 여러 예측변수 패턴(여러 PATTERN 하위 명령 사용).
- 루트 이름이 지정된 경우 저장된 변수의 최대 수(SAVE 하위 명령 사용). 대화 상자는 CSCOXREG 기본 25 변수를 표시합니다.

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



---

## 주의사항

이 정보는 미국에서 제공되는 제품 및 서비스용으로 작성된 것입니다. 본 자료는 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. 이 책에서 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.

150-945

서울특별시 영등포구 국제금융로 10, 31FC

한국 아이.비.엠 주식회사

대표전화서비스: 02-3781-7114

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 지역에서는 특정 상거래에 있어 명시적 또는 묵시적 보증 책임에 대한 불인정을 허용하지 않으므로, 이런 지역에서는 위의 사항이 적용되지 않습니다.

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 비IBM의 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.

150-945

서울특별시 영등포구 국제금융로 10, 3IFC

한국 아이.비.엠 주식회사

대표전화서비스: 02-3781-7114

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, 공개자료 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.

IBM이 제시하는 방향 또는 의도에 관한 모든 언급은 특별한 통지 없이 변경될 수 있습니다.

이 정보에는 일상의 비즈니스 운영에서 사용되는 자료 및 보고서에 대한 예제가 들어 있습니다. 이들 예제는 개념을 가능한 완벽하게 설명하기 위하여 개인, 회사, 상표 및 제품의 이름이 사용될 수 있습니다. 이들 이름은 모두 가공의 것이며 실제 기업 및 인물과 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 관련된 손해에 대해 책임을 지지 않습니다.



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

© (귀하의 회사명) (연도). 이 코드의 일부는 IBM Corp.의 샘플 프로그램에서 파생됩니다.

© Copyright IBM Corp. \_enter the year or years\_. All rights reserved.

---

## 상표

IBM, IBM 로고 및 [ibm.com](http://ibm.com)은 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 "저작권 및 상표 정보" 웹 페이지([www.ibm.com/legal/copytrade.shtml](http://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

## [가]

가중치를 부여하지 않은 개수

복합 표본 교차 분석표 25

복합 표본 기술통계 23

복합 표본 비율 29

복합 표본 빈도분석 19

결측값

복합 표본 20, 26

복합 표본 기술통계 24

복합 표본 로지스틱 회귀분석 41

복합 표본 비율 30

복합 표본 순서 회귀분석 47

복합 표본 일반선형모형 35

계통 표본추출

표본추출 마법사 6

계획 파일 2

군집

분석 준비 마법사 13

표본추출 마법사 5

기대값

복합 표본 교차 분석표 25

기준선 계층

복합 표본 Cox 회귀분석 51

## [나]

누적 확률

복합 표본 순서 회귀분석 46

누적값

복합 표본 빈도분석 19

## [다]

다항 대비

복합 표본 일반선형모형 34

단계 반복

복합 표본 로지스틱 회귀분석 41

복합 표본 순서 회귀분석 47

단순 대비

복합 표본 일반선형모형 34

단순 무작위 표본

표본추출 마법사 6

대비

복합 표본 일반선형모형 34

디자인 효과

복합 표본 교차 분석표 25

복합 표본 기술통계 23

복합 표본 로지스틱 회귀분석 38

복합 표본 비율 29

복합 표본 빈도분석 19

복합 표본 순서 회귀분석 45

복합 표본 일반선형모형 33

복합 표본 Cox 회귀분석 53

디자인 효과 제공근

복합 표본 교차 분석표 25

복합 표본 기술통계 23

복합 표본 로지스틱 회귀분석 38

복합 표본 비율 29

복합 표본 빈도분석 19

복합 표본 순서 회귀분석 45

복합 표본 일반선형모형 33

복합 표본 Cox 회귀분석 53

## [마]

모수 수렴

복합 표본 로지스틱 회귀분석 41

복합 표본 순서 회귀분석 47

모수 추정값

복합 표본 로지스틱 회귀분석 38

복합 표본 순서 회귀분석 45

복합 표본 일반선형모형 33

복합 표본 Cox 회귀분석 53

모수 추정값의 공분산

복합 표본 로지스틱 회귀분석 38

복합 표본 순서 회귀분석 45

복합 표본 일반선형모형 33

모수 추정값의 상관

복합 표본 로지스틱 회귀분석 38

복합 표본 순서 회귀분석 45

복합 표본 일반선형모형 33

모집단 크기

복합 표본 교차 분석표 25

복합 표본 기술통계 23

복합 표본 비율 29

복합 표본 빈도분석 19

모집단 크기 (계속)

표본추출 마법사 8

## [바]

반복

복합 표본 로지스틱 회귀분석 41

복합 표본 순서 회귀분석 47

반복 대비

복합 표본 일반선형모형 34

반복계산과정

복합 표본 로지스틱 회귀분석 41

복합 표본 순서 회귀분석 47

반응 확률

복합 표본 순서 회귀분석 44

복합 표본

가설 검정 33, 39, 45

결측값 20, 26

읍선 21, 24, 27, 30

복합 표본 교차 분석표 25

통계 25

복합 표본 기술통계 23

결측값 24

통계 23

복합 표본 로지스틱 회귀분석 37

명령 추가 기능 41

모형 38

승산비 40

읍선 41

저장할 변수 40

참조 범주 37

통계 38

복합 표본 비율 29

결측값 30

통계 29

복합 표본 빈도분석 19

통계 19

복합 표본 순서 회귀분석 43

모형 44

반응 확률 44

승산비 46

읍선 47

저장할 변수 46

통계 45

복합 표본 일반선형모형 31

명령 추가 기능 35

모형 32

옵션 35

저장할 변수 34

통계 33

평균 추정 34

복합 표본 Cox 회귀분석

가설 검정 54

날짜 및 시간변수 49

도표 54

모형 52

모형 내보내기 56

시간 종속 예측변수 51

예측변수 51

옵션 57

이벤트 정의 50

저장할 변수 54

통계 53

하위 그룹 51

Kaplan-Meier 분석 49

복합 표본추출

분석 계획 13

표본 계획 5

부모집단

복합 표본 Cox 회귀분석 51

분류표

복합 표본 로지스틱 회귀분석 38

복합 표본 순서 회귀분석 45

분리

복합 표본 로지스틱 회귀분석 41

복합 표본 순서 회귀분석 47

분석 계획 13

비례 위험의 검정

복합 표본 Cox 회귀분석 53

## [사]

상대적 위험도

복합 표본 교차 분석표 25

수정된 잔차

복합 표본 교차 분석표 25

수정된 카이제곱

복합 표본 33, 39, 45

복합 표본 Cox 회귀분석 54

순차 표본추출

표본추출 마법사 6

순차 Bonferroni 수정

복합 표본 33, 39, 45

복합 표본 Cox 회귀분석 54

순차 Sidak 수정

복합 표본 33, 39, 45

복합 표본 Cox 회귀분석 54

스코어 잔차

복합 표본 Cox 회귀분석 54

승산비

복합 표본 교차 분석표 25

복합 표본 로지스틱 회귀분석 40

복합 표본 순서 회귀분석 46

시간 종속 예측변수

복합 표본 Cox 회귀분석 51

신뢰구간

복합 표본 교차 분석표 25

복합 표본 기술통계 23

복합 표본 로지스틱 회귀분석 38

복합 표본 비율 29

복합 표본 빈도분석 19

복합 표본 순서 회귀분석 45

복합 표본 일반선형모형 33, 35

신뢰수준

복합 표본 로지스틱 회귀분석 41

복합 표본 순서 회귀분석 47

## [아]

열 퍼센트

복합 표본 교차 분석표 25

예측 범주

복합 표본 로지스틱 회귀분석 40

복합 표본 순서 회귀분석 46

예측 확률

복합 표본 로지스틱 회귀분석 40

복합 표본 순서 회귀분석 46

예측값

복합 표본 일반선형모형 34

우도 수렴

복합 표본 로지스틱 회귀분석 41

복합 표본 순서 회귀분석 47

위험차

복합 표본 교차 분석표 25

유사 R2 통계

복합 표본 로지스틱 회귀분석 38

복합 표본 순서 회귀분석 45

입력 표본 기중치

표본추출 마법사 5

## [자]

자유도

복합 표본 33, 39, 45

복합 표본 Cox 회귀분석 54

잔차

복합 표본 교차 분석표 25

복합 표본 일반선형모형 34

조정된 F 통계

복합 표본 33, 39, 45

복합 표본 Cox 회귀분석 54

주변 평균 추정값

복합 표본 일반선형모형 34

## [차]

차이 대비

복합 표본 일반선형모형 34

참조 범주

복합 표본 로지스틱 회귀분석 37

복합 표본 일반선형모형 34

최소유의차

복합 표본 33, 39, 45

복합 표본 Cox 회귀분석 54

중화

분석 준비 마법사 13

표본추출 마법사 5

## [카]

카이제곱

복합 표본 33, 39, 45

복합 표본 Cox 회귀분석 54

크기 측도

표본추출 마법사 6

## [타]

통합 잔차

복합 표본 Cox 회귀분석 54

## [파]

편차 대비

복합 표본 일반선형모형 34

편차 잔차

복합 표본 Cox 회귀분석 54

평균  
 복합 표본 기술통계 23

평행선 검정  
 복합 표본 순서 회귀분석 45

포함 확률  
 표본추출 마법사 8

표 퍼센트  
 복합 표본 교차 분석표 25  
 복합 표본 빈도분석 19

표본 가중치  
 분석 준비 마법사 13  
 표본추출 마법사 8

표본 결과  
 표본추출 마법사 7, 8

표본 계획 5

표본 디자인 정보  
 복합 표본 Cox 회귀분석 53

표본 비율  
 표본추출 마법사 8

표본추출  
 복합 디자인 5

표본추출 방법  
 표본추출 마법사 6

표본추출 추정  
 분석 준비 마법사 14

표준 오차  
 복합 표본 교차 분석표 25  
 복합 표본 기술통계 23  
 복합 표본 로지스틱 회귀분석 38  
 복합 표본 비율 29  
 복합 표본 빈도분석 19  
 복합 표본 순서 회귀분석 45  
 복합 표본 일반선형모형 33

## [ 하 ]

합계  
 복합 표본 기술통계 23

행 퍼센트  
 복합 표본 교차 분석표 25

## B

Bonferroni  
 복합 표본 33, 39, 45  
 복합 표본 Cox 회귀분석 54

Breslow 추정 방법  
 복합 표본 Cox 회귀분석 57

Brewer의 표본추출 방법  
 표본추출 마법사 6

## C

COV(변동 계수)  
 복합 표본 교차 분석표 25  
 복합 표본 기술통계 23  
 복합 표본 비율 29  
 복합 표본 빈도분석 19

Cox-Snell 잔차  
 복합 표본 Cox 회귀분석 54

## E

Efron 추정 방법  
 복합 표본 Cox 회귀분석 57

## F

F 통계  
 복합 표본 33, 39, 45  
 복합 표본 Cox 회귀분석 54

Fisher 스코어링  
 복합 표본 순서 회귀분석 47

## H

Helmert 대비  
 복합 표본 일반선형모형 34

## M

martingale 잔차  
 복합 표본 Cox 회귀분석 54

Murthy의 표본추출 방법  
 표본추출 마법사 6

## N

Newton-Raphson 방법  
 복합 표본 순서 회귀분석 47

## P

PPS 표본추출  
 표본추출 마법사 6

## R

R2 통계  
 복합 표본 일반선형모형 33

## S

Sampford의 표본추출 방법  
 표본추출 마법사 6

Schoenfeld의 편잔차  
 복합 표본 Cox 회귀분석 54

Sidak 수정  
 복합 표본 33, 39, 45  
 복합 표본 Cox 회귀분석 54

## T

T 검정  
 복합 표본 로지스틱 회귀분석 38  
 복합 표본 순서 회귀분석 45  
 복합 표본 일반선형모형 33





