

7. 신뢰도 검정

신뢰도 분석(Reliability Analysis)

<note>

1) 신뢰도란?

⇒ 반복된 측정에서 어느 정도 동일한 결과를 얻는가의 의미

2) 타당도란?

⇒ 연구자가 관찰하려고 하는 것을 어느 정도 제대로 관찰하였는가?

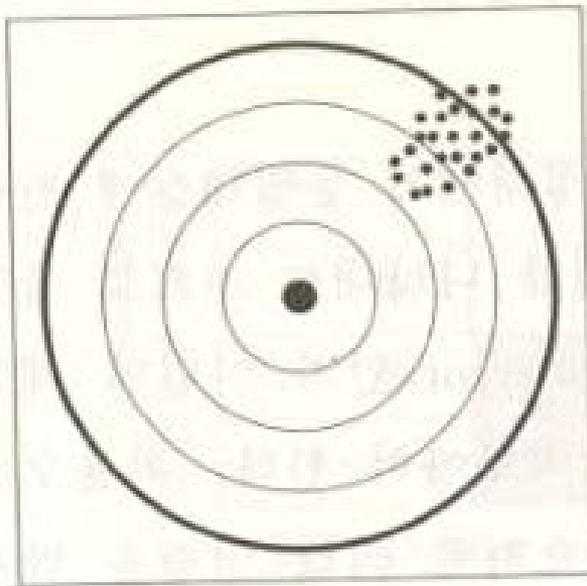
(예) 적성검사의 타당도 → 장래의 성공적인 직무 수행 여부를 제대로 측정해 내는가?

3) 크론바흐 알파(Cronbach α)의 이용

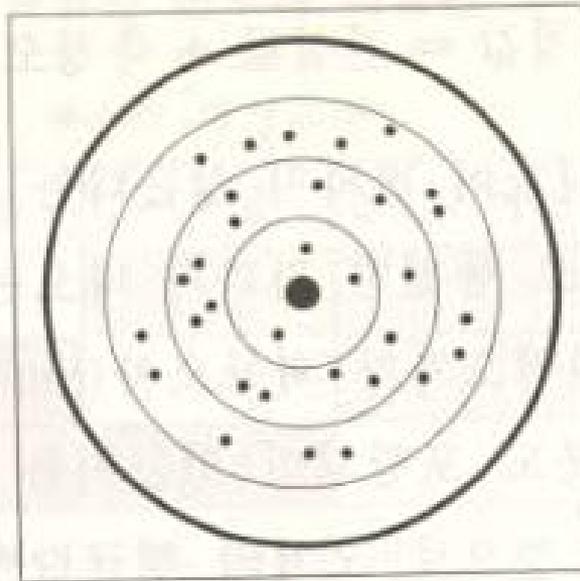
- ① 설문지의 여러 문항이 얻고자 하는 정보에 잘 접근하고 있다면 → 신뢰도가 높다
- ② 요인분석에서 같은 요인으로 묶인 변수들의 구성이 타당하다. → 신뢰도가 높다
- ③ 앞의 경우 → Cronbach α 값이 높게 나온다.

<note> 신뢰도와 타당도의 관계에 대한 비유
⇒ 과녁을 향해 사격을 한다고 했을 때 ...

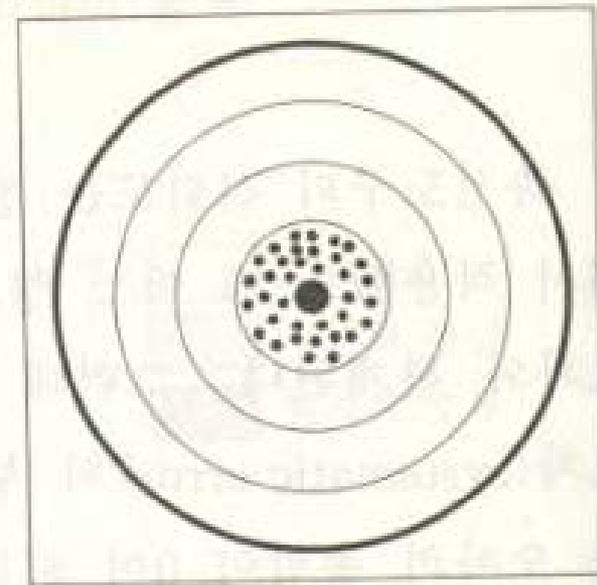
[그림 6-2] 신뢰도와 타당도의 관계에 대한 비유



(가) 신뢰할 수 있지만 타당하지 않다



(나) 타당하지만 신뢰할 수 없다



(다) 타당하고 신뢰할 수 있다

1. 신뢰도의 원리

⇒ 신뢰도의 정확성은 측정오차가 존재하는가의 여부와 관련이 있음.

즉, 관측 값이 $X^* = X + \varepsilon$ 이라 하자. 이 때,

X^* : 관측 값

X : 참 값(직접 측정할 수 없는)

ε : 측정오차이고, $E(\varepsilon) = 0$

이라고 하자.

이 때, X 의 관측 값으로서 X^* 의 신뢰도 계수 R_{tt} 는 다음과 같이 구한다. 즉, 관측 값의 분산에 대한 참 값의 분산의 비율

$$R_{tt} = \frac{\text{Var}(X)}{\text{Var}(X^*)} = 1 - \frac{\text{Var}(\varepsilon)}{\text{Var}(X^*)}$$

<note> 이 때,

① $0 \leq R_{tt} \leq 1$

② 측정값의 분산이 모두 측정오차에 의해 생겼다면
→ 참 값의 분산은 0가 되어, 신뢰도 값도 0이 된다.

즉,

$$\frac{\text{Var}(\varepsilon)}{\text{Var}(X^*)} = 1 \text{ 이므로, } R_{tt} = 1 - 1 = 0 \text{ 이다.}$$

③ 측정오차가 전혀 없다면, 즉 $\text{Var}(\varepsilon) = 0$ 이면

$$R_{tt} = 1$$

이다.

2. 신뢰도 평가방법

(1) 재측정법 \Rightarrow 동일한 개념 X를 한 시점과 또 다른 시점에 조사하여 이들 사이의 관계가 있는지를 조사.

< 단점 >

- ① 대부분의 경우 사회현상은 조사 시점마다 다르다.
- ② 두 번째 조사의 경우, 첫 번째 조사 내용을 기억하는 사람이 많다.

(2) Cronbach α

\Rightarrow 문항 상호 간에 어느 정도 일관성을 가지는가를 측정.

$$\text{① Cronbach } \alpha = \frac{k}{K - 1} \left(1 - \frac{\sigma_i^2}{\sigma_t^2} \right)$$

여기서, k : 항목 수, σ_i^2 : 개별항목의 분산

σ_t^2 : 전체항목의 총 분산

$$\textcircled{2} \text{ 표준화 된 Cronbach } \alpha = \frac{k \times r}{1 + (k - 1) \times r}$$

여기서, k : 항목 수, r : 항목에 대한 평균 상관 계수

< note > 신뢰도의 제고 방안

- 1) 설문지 문항이 명확해야 함 → 측정하고자 하는 것을 정확히 표현.
- 2) 문항 수의 증가 → 적절하게 통제해야 함.
- 3) 성의 없거나 일관성이 없는 문항은 제외
- 4) 중요한 질문의 경우 → 유사한 질문에 응답하게 하여 서로 응답의 신뢰성을 조사.
- 5) 기존의 신뢰성이 검증된 설문지를 이용.

< 참고 > 상관계수 r 과 문항수(k)의 증가에 따른 신뢰도의 변화

	r	0.0	0.2	0.4	0.6	0.8	1.0
k							
2		0.0	0.33	0.57	0.75	0.89	1.0
4		0.0	0.50	0.73	0.86	0.94	1.0
6		0.0	0.60	0.80	0.90	0.96	1.0
8		0.0	0.67	0.84	0.92	0.97	1.0
10		0.0	0.71	0.87	0.94	0.98	1.0

<Note> 신뢰도 검정

(1) 분석개요

- 변수의 일관성 측정
- 독립된 여러 측정 항목에 의해 대상을 측정하는 경우 결과가 서로 비슷하게 나타나는지를 검정
- 주요 통계량: Cronbach α 계수 계산

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum \sigma_i^2}{\sigma_y^2} \right)$$

k = 측정항목수

σ_i^2 = 개별측정항목의 분산

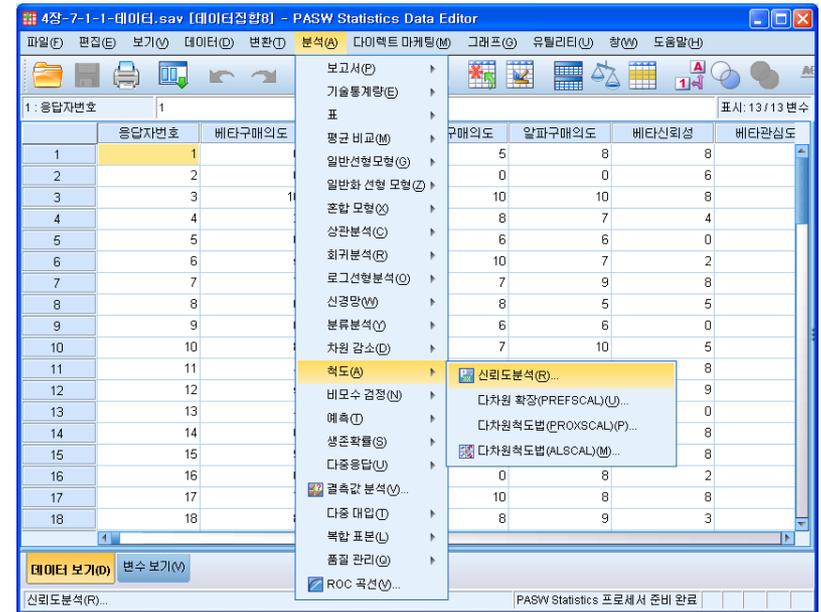
σ_y^2 = 전체측정항목의 분산

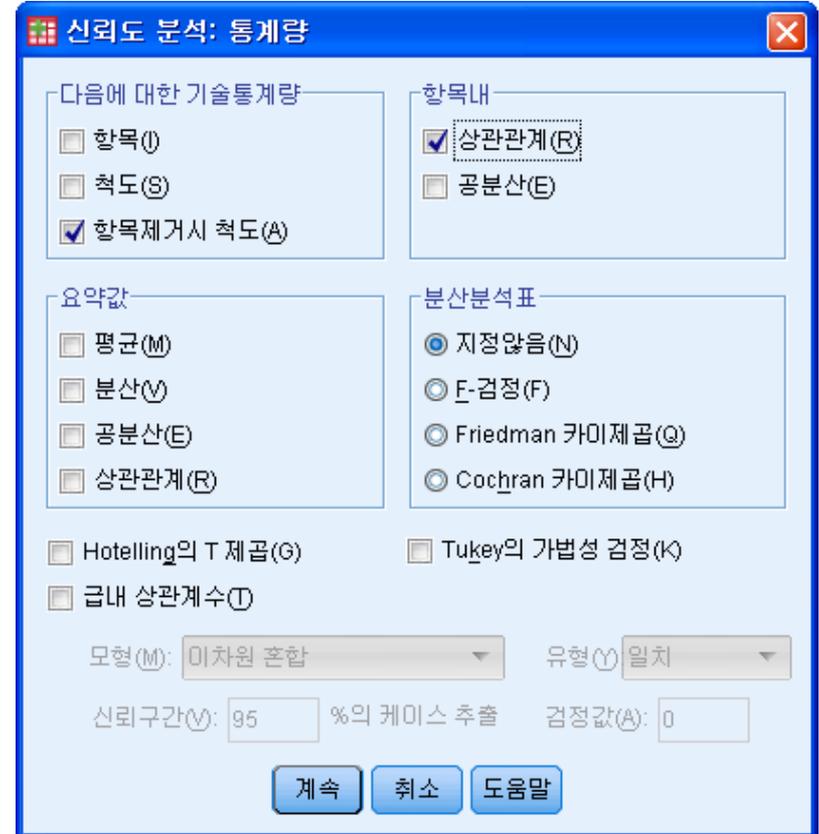
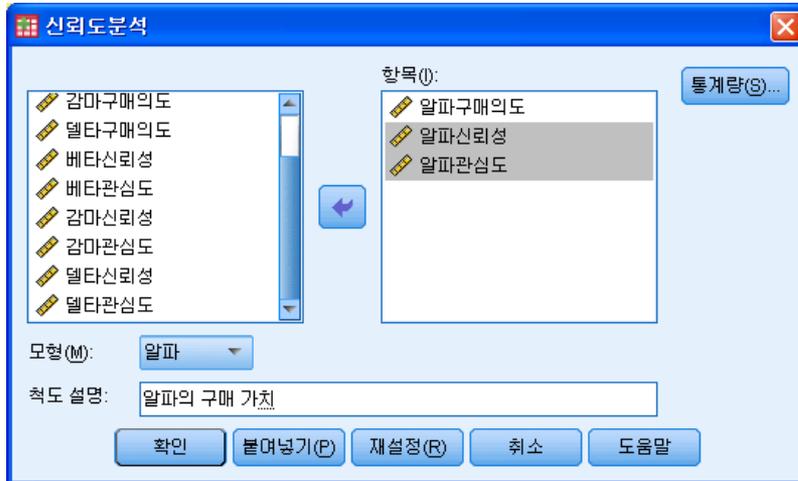
(2) 분석데이터(4장-7-1-1.sav)

- 소비자 60명의 타이어 구매의도, 신뢰도, 관심도 조사
- 알파의 구매가치 조사

(3) 분석과정

- STEP 01: [신뢰도분석] 메뉴 클릭
- STEP 02: 변수 지정
- STEP 03: 통계량 지정





(4) 결과해석

신뢰도 통계량

	Cronbach의 알파	Cronbach's Alpha Based on Standardized Items	항목 수			
항목 총계 통계량						
	항목이 삭제된 경우 척도 평균	항목이 삭제된 경우 척도 분산	수정된 항목-전체 상관관계	제곱 다중 상관관계	항목이 삭제된 경우 Cronbach 알파	
알파구매의도	13.52	18.186	.597	.368	.833	
알파신뢰성	12.08	25.908	.683	.556	.744	.601
알파관심도	12.77	19.436	.751	.621	.624	.742
						.000

1) 전체항목의 통계량

⇒ ① 각 항목에 대해 어떤 특정한 항목이 척도로부터 제외될 때 그 척도에 대한 평균, 분산, Cronbach α 값 등을 의미한다.

(예) 만일 알파구매의도가 제외되면 평균 값은 13.52가 된다.

항목 총계 통계량

	항목이 삭제된 경우 척도 평균	항목이 삭제된 경우 척도 분산	수정된 항목-전체 상관관계	제곱 다중 상관관계	항목이 삭제된 경우 Cronbach 알파
알파구매의도	13.52	18.186	.597	.368	.833
알파신뢰성	12.08	25.908	.683	.556	.744
알파관심도	12.77	19.436	.751	.621	.624

② 수정된 항목-전체 상관관계는 개별항목에 관한 점수와 잔존하는 나머지 항목의 전체 점수 사이의 피어슨 상관계수이다.

(예) 예를 들면, 변수 알파구매의도와 (알파신뢰성+ 알파관심도)까지의 전체 점수들 간의 상관계수는 0.597이다.

③ **제공다중상관관계**는 회귀방정식으로부터 구한 다중상관관계의 제곱 즉, R^2 이다. 즉, 특정변수로서의 종속변수와 나머지 변수들을 독립변수로 하였을 때의 회귀방정식에서 R^2 를 의미한다.

(예) 예를 들면, 알파구매의도 변수의 응답에서 다른 변수에 의해 약 36.8% 정도의 관측 변량을 설명할 수 있음을 의미한다.

항목 총계 통계량

	항목이 삭제된 경우 척도 평균	항목이 삭제된 경우 척도 분산	수정된 항목-전체 상관관계	제공 다중 상관관계	항목이 삭제된 경우 Cronbach 알파
알파구매의도	13.52	18.186	.597	.368	.833
알파신뢰성	12.08	25.908	.683	.556	.744
알파관심도	12.77	19.436	.751	.621	.624

④ 항목이 삭제된 경우의 Cronbach α
⇒ 해당 변수가 제외되었을 경우의 Cronbach α 값이다.
신뢰도 척도인 알파 값은 클수록 좋은 값이므로,

㉠ 그 변수를 제외함으로써 알파 값이 커진다면
→ 해당 변수는 전체 신뢰도 저하의 원인이 되므로 분석에서 제외하는 것이 바람직하다.

㉡ 그 변수를 제외함으로써 알파 값이 작아진다면
→ 해당 변수는 제외하지 말아야 한다.

(예)

◀ 알파구매의도를 제외하면 신뢰도는 원래 0.805에서 0.833로 높아진다.

◀ 알파관심도를 제외하면 신뢰도는 원래 0.805에서 0.624로 낮아진다. 따라서, X1은 중요한 변수이므로 분석에서 제외하지 말아야 한다.

2) 신뢰도 통계량

㉠ Alpha 값에 대하여

⇒ Cronbach α 값을 나타낸다. 이 예의 경우

◀ Alpha = .805

◀ Standardized item alpha = .827

이므로 신뢰도가 대체로 양호하다고 할 수 있다.

<note> 사회과학적 관점에서 볼 때,

◀ if Alpha \geq 0.5 이면, 신뢰도가 있다.

◀ if Alpha \geq 0.6 이면, 신뢰도가 높은 편이다.

◀ if Alpha \geq 0.7 이면, 신뢰도가 높다.

◀ if Alpha \geq 0.8 이면, 신뢰도가 아주 높다.

Chapter 5. 가설검정

1. 가설검정

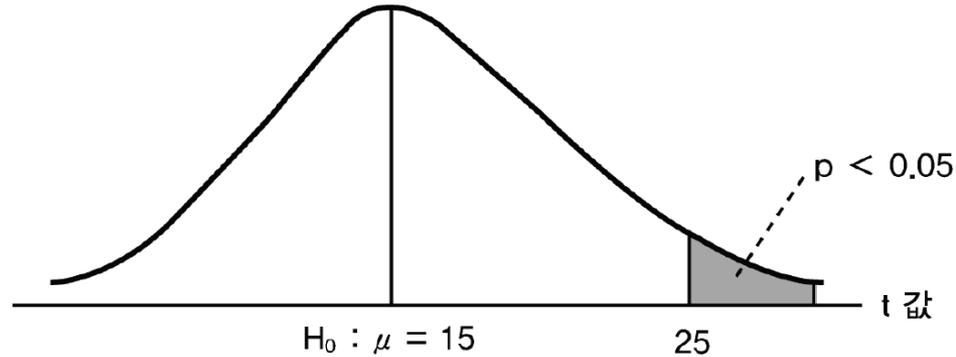
1.1. 가설이란

- 가설

- 표본으로부터 주어지는 정보를 이용하여, 모수에 대한 예상, 주장 또는 단순한 추측을 기술하는 것
- 대립가설(alternative hypothesis: H_1): 데이터로부터 얻은 강력한 증거에 의해 연구자가 입증하고자 하는 가설
- 귀무(영)가설(null hypothesis: H_0): 대립가설에 상반되는 가설로서 분석 이전에 자연현상에 가까운 사실에 대한 가설

1.2. 가설검정

- 가설검정
 - 설정된 가설 중에 어느 것이 맞는 지를 검정
 - 검정통계량을 활용
 - 분포: 정규분포, t -분포, χ^2 -분포, F -분포
- 기각역
 - 귀무가설을 기각하게 되는 검정통계량의 관측값 영역
- 유의수준
 - 5%, 1%를 주로 사용
- 오류
 - α 와 β 오류



	실제현상	귀무가설이 사실	대립가설이 사실
검정결과			
귀무가설을 채택		옳은 결정	β
귀무가설을 기각		α	옳은 결정

1.3. 가설검정의 종류

표본의 개수	검정 대상	모분산 파악여부	분석 구분
1개	평균	알고 있음	한 표본에서 평균에 대한 Z -검정
		모름	한 표본에서 평균에 대한 t -검정
	비율	관계없음	한 표본에서 비율에 대한 비율검정
	분산	관계없음	한 표본에서 분산에 대한 모분산검정
2개	평균	관계없음 독립된 표본	두 표본에서 평균에 대한 t -검정
		관계없음 쌍체 표본	두 표본에서 평균에 대한 쌍체 t -검정
	비율	관계없음	두 표본에서 비율에 대한 비율검정
	분산	관계없음	두 표본에서 분산에 대한 모분산검정

2. 한 표본에 대한 가설검정

2.1. Z-검정

(1) 분석개요

- 모집단에 대한 분산을 알고 있는 경우 가설 검정 방법
- 음료수 병의 함량에 대한 조사
- 분산은 1ml로 알려져 있음
 - H_0 : 음료수병 함량은 350ml이다($\mu_1 = \mu_0$).
 - H_1 : 음료수병 함량은 350ml이 아니다($\mu_1 \neq \mu_0$).

(2) 분석데이터

- 10개의 자료를 수집함

(3) 분석과정

- 기술통계 확인

- STEP 01: [평균비교] [일표본 T검정]
- STEP 02: 분석 변수 지정

The screenshot shows the IBM SPSS Statistics Data Editor interface. The title bar indicates the file is '5장-2-1-1-데이터.sav [데이터집합1]'. The menu bar includes '파일(F)', '편집(E)', '보기(V)', '데이터(D)', '변환(T)', '분석(A)', '다이렉트 마케팅(M)', '그래프(G)', '유틸리티(U)', '창(W)', and '도움'. The 'Analyze' menu is open, showing options like '보고서(P)', '기술통계량(E)', '표', '평균 비교(M)', '일반선형모형(G)', '일반화 선형 모형(Z)', '혼합 모형(X)', '상관분석(C)', '회귀분석(R)', '로그선형분석(O)', '신경망(W)', '분류분석(Y)', '차원 감소(D)', and '척도(A)'. The '평균 비교(M)' option is selected, and its submenu is open, showing '집단별 평균분석(M)...', '일표본 T검정(S)...', '독립표본 T검정(T)...', '대응표본 T검정(P)...', and '일원배치 분산분석(O)...'. The '일표본 T검정(S)...' option is highlighted. In the background, a data table is visible with columns '음료수함량', '변수', and '변수', and rows numbered 1 to 11.

	음료수함량	변수	변수
1	350.8		
2	352.2		
3	349.5		
4	350.3		
5	348.6		
6	348.3		
7	350.0		
8	351.5		
9	351.5		
10	350.1		
11			

검정변수 : 음료수합량

검정값 : 350 입력

일표본 T 검정

검정변수(T):

- 음료수합량

검정값(Y): 350

확인 붙여넣기(P) 재설정(R) 취소 도움말

옵션(O)... 붓스트랩(B)...

(4) 결과해석

일표본 통계량

	N	평균	표준편차	평균의 표준오차
음료수합량	10	350.280	1.2647	.3999

일표본 검정

	검정값 = 350					
	t	자유도	유의확률 (양쪽)	평균차	차이의 95% 신뢰구간	
					하한	상한
음료수합량	.700	9	.502	.2800	-.625	1.185

검정 통계량 0.700,
p-value=0.502 로서 유의수준 5% 하에서 귀무가설을 채택함.

2.2. t-검정

(1) 분석개요

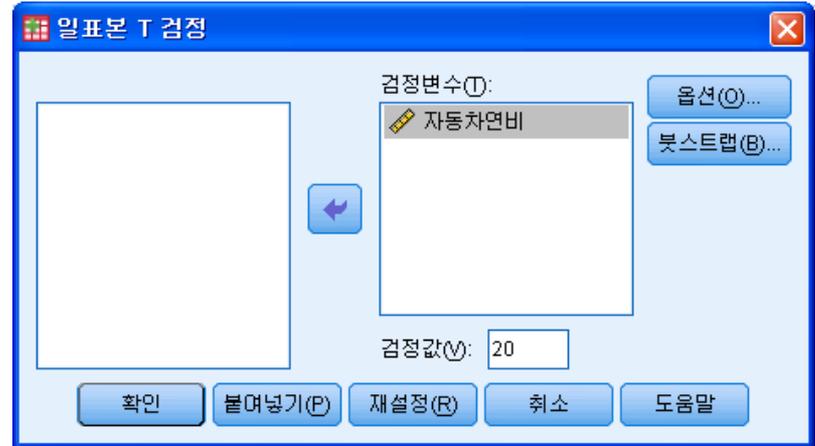
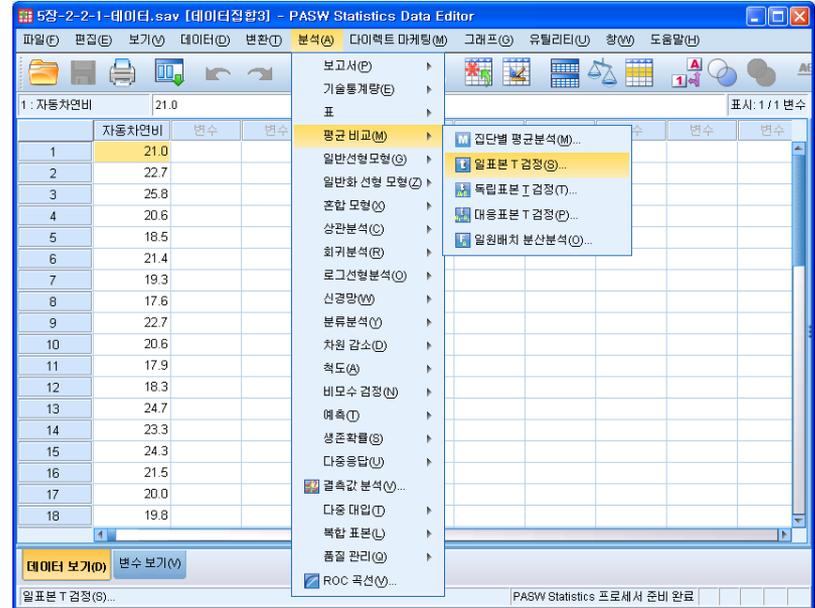
- 모집단에 대한 분산을 모르고 있는 경우 가설 검정 방법
- 자동차 연비에 대한 조사
 - H_0 : 자동차의 연비는 20km/l 미만이다 ($\mu_1 < \mu_0$).
 - H_1 : 자동차의 연비는 20km/l 이상이다 ($\mu_1 \geq \mu_0$).

(2) 분석데이터

- 20대 자동차 데이터를 수집함

(3) 분석과정

- STEP 01: [일표본 T-검정] 메뉴 클릭
- STEP 02: 검정변수 및 검정값 지정



(4) 결과해석

일표본 통계량

	N	평균	표준편차	평균의 표준오차
자동차연비	20	21.140	2.3383	.5229

일표본 검정

	검정값 = 20					
	t	자유도	유의확률 (양쪽)	평균차	차이의 95% 신뢰구간	
					하한	상한
자동차연비	2.180	19	.042	1.1400	.046	2.234

2.3. 비율검정

(1) 분석개요

- 한 표본에 대한 비율검정은 성공, 실패 또는 불량률과 같이 비율에 대한 검정
- 대학 졸업자의 전공분야 취직에 대한 조사
 - H_0 : 대학졸업자 중 자신의 전공을 살릴 수 있는 직장에 입사하는 비율이 20%이다 ($p_1 = p_0$).
 - H_1 : 대학졸업자 중 자신의 전공을 살릴 수 있는 직장에 입사하는 비율이 20%가 아니다 ($P_1 \neq p_0$).

(2) 분석데이터

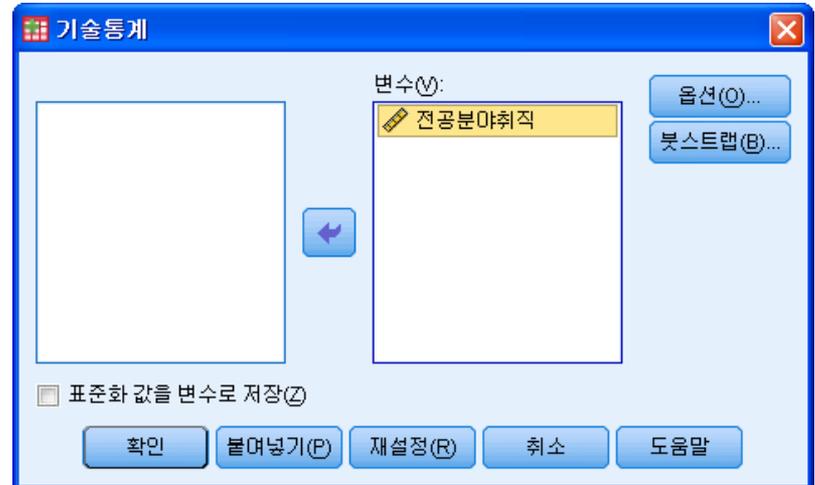
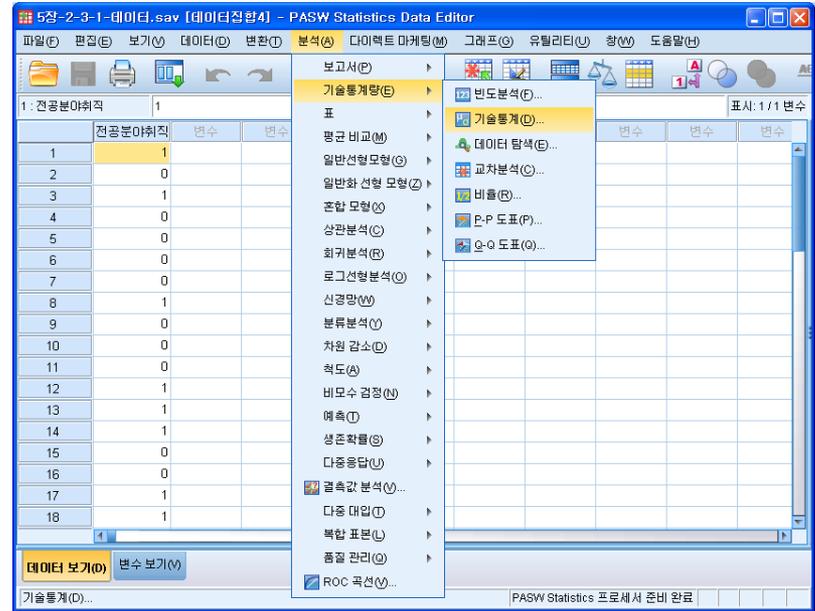
- 30명의 대학졸업자 데이터를 수집함

(3) 분석과정

(5장-2-3-1.sav)

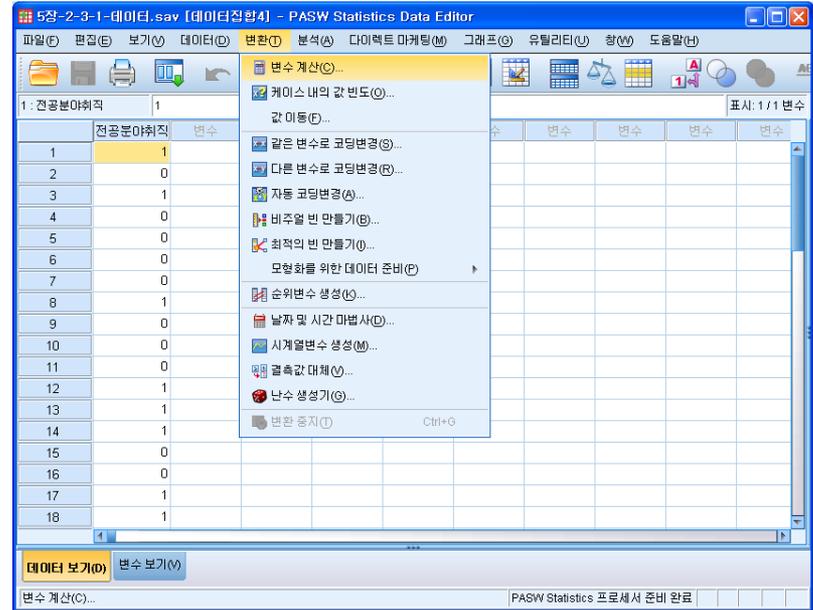
- 기술통계 확인

- STEP 01: [기술통계] 메뉴 클릭
- STEP 02: 분석 변수 지정



- Z값 계산

- STEP 01: [변수 계산] 메뉴 클릭
- STEP 02: Z값 계산식 입력
- STEP 03: 계산된 Z값 보기



대상변수(T):

z

=

숫자표현식(E):

$(0.43-0.20)/\text{sqrt}(0.2*(1-0.2)/30)$

유형 및 설명(L)...

전공분야취직



+	<	>	7	8	9
-	<=	>=	4	5	6
*	=	~=	1	2	3
/	&		0	.	
**	~	()	삭제		

함수 집단(G):

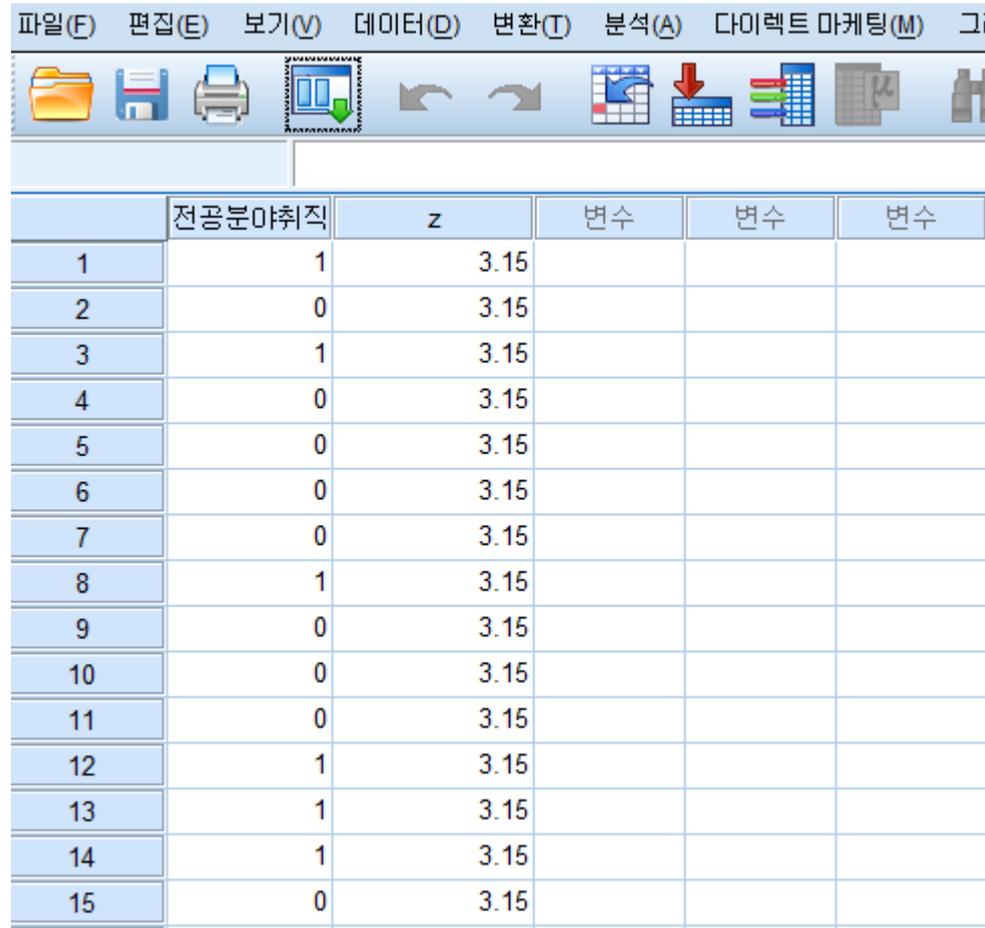
- 모두
- 산술
- CDF 및 비중심 CDF
- 변환
- 현재 날짜/시간
- 날짜 산술
- 날짜 작성
- ...

함수 및 특수변수(F):

조건(O)... (선택적 케이스 선택 조건)

확인 붙여넣기(P) 재설정(R) 취소 도움말

Z값 계산



The image shows a screenshot of a spreadsheet application. The menu bar includes '파일(F)', '편집(E)', '보기(V)', '데이터(D)', '변환(T)', '분석(A)', '다이렉트 마케팅(M)', and '그리'. The toolbar contains icons for file operations, printing, and data analysis. The main area displays a table with 15 rows and 6 columns. The columns are labeled '전공분야취직', 'z', and three '변수' (Variables) columns. The 'z' column contains the value 3.15 for all rows. The '전공분야취직' column contains values 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 1, 1, 0 for rows 1 through 15 respectively.

	전공분야취직	z	변수	변수	변수
1	1	3.15			
2	0	3.15			
3	1	3.15			
4	0	3.15			
5	0	3.15			
6	0	3.15			
7	0	3.15			
8	1	3.15			
9	0	3.15			
10	0	3.15			
11	0	3.15			
12	1	3.15			
13	1	3.15			
14	1	3.15			
15	0	3.15			

1) 유의확률 계산(p.216)

- STEP 04: 다시 [변수계산] 메뉴 클릭
- STEP 05: 계산식 입력
- STEP 06: 유의확률 확인

2) $\alpha = 0.05$ 일 경우 기각역 설정

*양측검정의 경우

$z < -1.96$ or $z > 1.96$ 이면 유의수준 5%하에서 귀무가설을 기각함.

(4) 결과해석

The top screenshot shows the '변수 계산' (Variable Calculation) dialog box. The '대상변수' (Target Variable) is '유의확률' (Significance Level) and the '숫자표현식' (Numeric Expression) is '(1-CDF.NORMAL(43,20,SQRT(.20*(1-.20)/30)))*2'. The bottom screenshot shows the 'PASW Statistics Data Editor' window with a data table. The table has columns for '전공분야취직' (Major Field Employment), 'z', and '유의확률' (Significance Level). The '유의확률' column contains values of .0016 for all rows.

전공분야취직	z	유의확률	변수	변수	변수	변수	변수	변수
1	3.15	.0016						
2	0	.0016						
3	1	.0016						
4	0	.0016						
5	0	.0016						
6	0	.0016						
7	0	.0016						
8	1	.0016						
9	0	.0016						
10	0	.0016						
11	0	.0016						
12	1	.0016						
13	1	.0016						
14	1	.0016						
15	0	.0016						
16	0	.0016						
17	1	.0016						
18	1	.0016						

2.4. 모분산검정

(1) 분석개요

- 한 표본에 대한 분산검정은 표본추출을 통해 나온 표본분산에 대해 기대하고 있는 모분산과 같은지를 검정
- 음료수 함량 조사의 모분산에 대한 조사
 - H_0 : 모분산은 1이다($\sigma_1^2 = \sigma_0^2$).
 - H_1 : 모분산은 1이 아니다 ($\sigma_1^2 \neq \sigma_0^2$).

(2) 분석데이터

- 10개의 음료수 병 함량 데이터를 수집함

(3) 분석과정

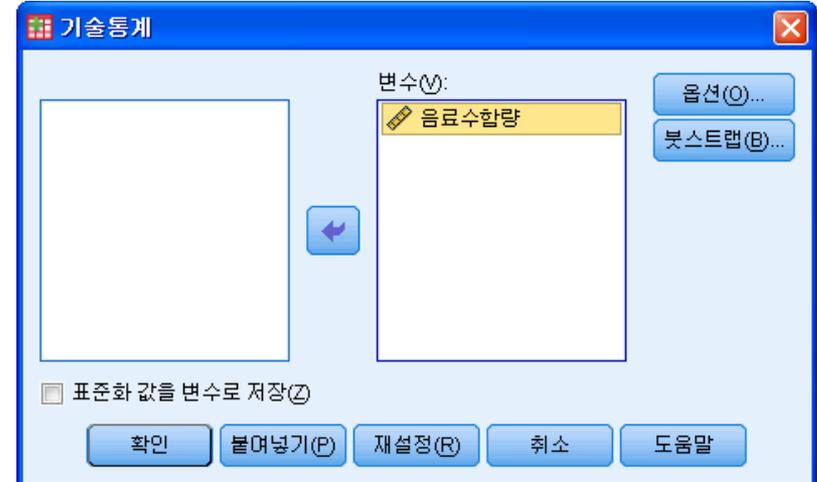
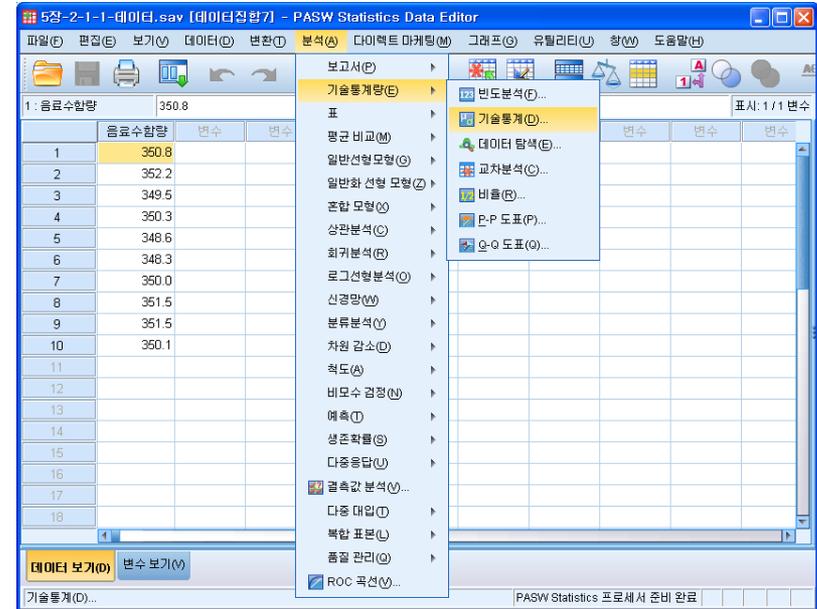
(5장-2-1-1.sav)

- 기술통계 확인

- STEP 01: [기술통계] 메뉴 클릭
- STEP 02: 분석 변수 지정

기술통계량

	N	최소값	최대값	평균	표준편차
음료수합량	10	348.3	352.2	350.280	1.2647
유효수 (목록별)	10				



- X² 값 계산

- STEP 01: [변수 계산] 메뉴 클릭
- STEP 02: X²값 계산 식 입력
- STEP 03: 계산된 X² 값 보기

$$X^2 = (n-1)S^2 / \sigma^2$$



