



제10장 상관관계분석을 통한 상호관계검증



| 도입 사례 |

미국에서 682명의 여성고객을 대상으로 슈퍼마켓에서의 행동을 조사한 자료에 의하면 그들 중 593명이 그들이 정상적으로 구입하는 상품에 대해 쿠폰을 사용했고 그들 중 95%가 최근 한 달 내에 쿠폰을 사용했으며 한 가구당 방문한 슈퍼마켓은 2.9개의 점포에 해당한다는 것이 밝혀졌습니다.

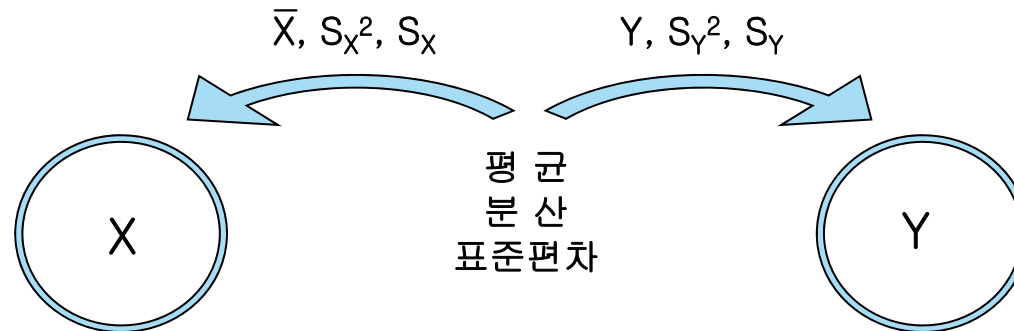
한편, 위에서 수집한 쿠폰사용과 소득간의 관계를 특정한 표로 동시에 정리(통계용어로는 cross-tabulation)하고 이들 현상간의 관계를 분석함으로써 (1) 모든 소득수준의 가구가 쿠폰을 사용하고 있음을 알 수 있었고, (2) 하위 소득수준의 가구는 중간이나 고소득수준의 가구보다 쿠폰을 훨씬 더 많이 사용하는 것을 밝힌바 있습니다. 그리고 이와 같은 소득수준과 쿠폰사용간의 관계는 통계적으로 유의한 것으로 나타났습니다. 슈퍼마켓에서 쿠폰과 관련된 전략을 수립할 경우에 그와 같은 발견은 유용하게 사용될 수 있었습니다.

생각해 볼 문제————

- ① 어떤 기준으로 사회현상들간에 **상호관계가 존재**한다고 말할 수 있을까요?
- ② 현상간에 존재하는 상호관계가 현상간의 인과관계 설명에 어떠한 역할을 할까요?

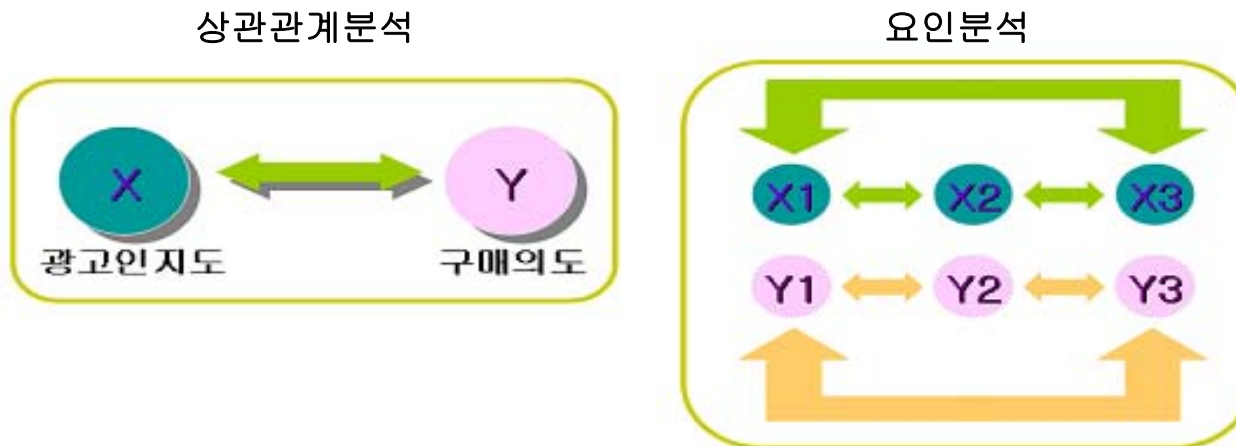
1. 상관관계분석이란?

[그림 10-1] 기술통계분석



◆ 기술통계분석에서는 사회현상 X와 Y에 대해 각각 설명

[그림 10-2] 상호관계분석





제1절 상관관계분석이란 무엇인가?

- ◆ 제10장부터는 현상들간의 **관계(relationship)**에 대해 설명
- ◆ 과학의 목표가 현상들간의 관계를 설명해 주는 것
- ◆ 즉 현상들간에 인과관계(causal relationship)가 존재하는지를 알아보아야 할 것
- ◆ 인과관계가 밝혀지기 위해서는 우선 두 개의 현상이 동시에 변하는 현상인지를 밝힐 필요
- ◆ **상관관계분석(analysis of association)**
 - 상호관계분석에 속하는 추리통계기법으로, 조사연구대상 변수들간의 관련성이 있는지를 분석하는 데 사용되는 기법
 - **하나의 변수가 다른 변수와 동시에 변하는지, 또 변한다면 어떤 방향(같은 방향 혹은 반대 방향)으로 변하는지, 그리고 그 변화 정도는 어떠한지 등을 분석할 때 사용하는 방법**
 - 상관관계분석에서는 두 현상의 동반변화에 대한 관심
 - 그러나 상관관계분석이 인과관계분석의 기초분석으로 사용될 수 있음
 - 예1) 한국고교 학생들의 영어점수와 수학점수간의 상관관계분석
 - 예2) 광고비와 매출간의 상관관계분석



2. 상관관계분석의 원리

- ⇒ 계량적 자료수집 시 사용된 근원척도의 유형에 따라 상관관계분석 기법이 상이

1) 등간/비율척도로 측정된 현상들간의 상관관계분석

- ◆ 기술통계분석의 산포도(dispersion measures) 측정도구
 - 범위(range), 편차(deviation), 분산(variance), 표준편차(standard deviation)
 - ☞ 하나의 현상(변수)의 변화폭을 측정해 주는 도구



제1절 상관관계분석이란 무엇인가?

◆ 공분산(covariance: COV_{xy})

- ▣ 두 가지 현상(변수)이 동시에 변화하는 정도를 측정해 주는 도구
- ▣ 하나의 현상의 변동폭을 알려주는 도구인 분산과 밀접한 관련이 있으면서, 두 현상들간의 관계분석기능을 수행하는 도구
- ▣ 하나의 변수(X_i)가 평균(\bar{X})을 기준으로 변함에 따라, 다른 변수(Y_i)가 평균(\bar{Y})을 기준으로 어떤 방향으로 얼마나 변화하는가를 보여주는 도구
- ▣ 표본크기가 증가할수록 $\sum(X_i - \bar{X})(Y_i - \bar{Y})$ 의 값도 커지므로 이 값을 표본의 크기 N (실제로는 $N-1$)으로 나누어서 표본의 크기에 영향을 받지 않는 두 변수의 동반변화폭을 알려주는 도구 (n-1로 나누어야 추정결과가 불편추정치가 되기 때문)
- ▣ 현상을 측정할 때 어떠한 측정단위(예: 섭씨 혹은 화씨로 온도 측정; kg 혹은 lb로 무게 측정; cm 혹은 inch로 키 측정 등)를 사용하는가에 따라 동일한 현상에 대한 공분산이 달라질 수 있음 → X, Y 단위에 영향을 받음,

상관관계 **정도** 알려주는 지표가 되지 못함

공분산(covariance)

(식 10-1)

$$COV_{xy} = \sum(X_i \text{의 변화폭})(Y_i \text{의 변화폭}) / (\text{표본크기} - 1)$$

$$= \sum(X_i - \bar{X})(Y_i - \bar{Y}) / (N - 1)$$



제1절 상관관계분석이란 무엇인가?

◆ Pearson 상관계수

- 공분산을 표준화한 도구
- 동일한 측정수준(예: 비율척도)이라도 상이한 측정단위(예: kg혹은 lb로 몸무게 측정)에 따라 변할 수 있는 공분산의 단점을 제거하면서 두 변수간의 동반변화 여부 및 정도를 판단할 수 있는 도구

☞ 표본의 크기가 30 이상인 경우 등간척도 이상으로 측정된 두 변수가 정규분포를 따르게 되므로 Pearson 상관계수를 구해서 두 변수간의 상관관계를 판단

☞ $-1 \leq \text{Pearson 상관계수} \leq 1$

Pearson 상관계수

(식 10-2)

$$\begin{aligned} r &= COV_{xy} / \text{표준화 도구} \\ &= COV_{xy} / (X \text{의 표준편차})(Y \text{의 표준편차}) \\ &= \mathbf{COV_{xy} / s_x s_y} \end{aligned}$$

(\because 표준편차: 측정단위와 관계없이 하나의 변수의 변화폭을 표시해 주는 도구)



제1절 상관관계분석이란 무엇인가?

◆ 적률상관계수(product-moment correlation coefficient)

▫ Pearson 상관계수

▫ 두 분포의 1차율의 적(積), 즉 곱

◆ 편차(deviation: $X_i - \bar{X}$)

▫ 분포의 1차율(first moments of a distribution), 편차의 제곱은 분포의 2차율, 편차의 세제곱은 분포의 3차율

적률상관계수(=Pearson 상관계수) (식 10-3)

$$\begin{aligned} r &= \sum (X_i \text{의 표준점수})(Y_i \text{의 표준점수}) / (N-1) \\ &= \sum Z_x Z_y / (N-1) \end{aligned}$$

☞ 실제로는 SPSS Statistics와 같은 통계분석 프로그램에 분석명령을 내림과 거의 동시에 분석결과가 계산됨

제1절 상관관계분석이란 무엇인가?

〈표 10-1〉 SUV의 선호도와 보유기간

응답자 번호	SUV 선호도(Y_i)	SUV 보유기간(X_i)
1	6	10
2	9	12
3	8	12
4	3	4
5	10	12
6	4	6
7	5	8
8	2	2
9	11	18
10	9	9
11	10	17
12	2	2

- SUV 선호도 : 1(매우 싫어함) ~ 11(매우 좋아함), 등간척도
- SUV 보유기간 : 개월 수, 비율척도



제1절 상관관계분석이란 무엇인가?

〈표 10-2〉 Pearson상관계수 계산결과

$$\bar{X} = \sum X_i / N = (10 + 12 + \cdots + 2) / 12 = 9.33$$

$$\bar{Y} = \sum Y_i / N = (6 + 9 + \cdots + 2) / 12 = 6.58$$

$$\begin{aligned} \sum (X_i - \bar{X})(Y_i - \bar{Y}) &= (10 - 9.33)(6 - 6.58) + (12 - 9.33)(9 - 6.58) + \cdots \\ &\quad + (2 - 9.33)(2 - 6.58) = 179.6668 \end{aligned}$$

$$\sum (X_i - \bar{X})^2 = (10 - 9.33)^2 + (12 - 9.33)^2 + \cdots + (2 - 9.33)^2 = 304.6668$$

$$\sum (Y_i - \bar{Y})^2 = (6 - 6.58)^2 + (9 - 6.58)^2 + \cdots + (2 - 6.58)^2 = 120.9168$$

$$r = 179.6668 / \sqrt{(304.6668)(120.9168)} = 0.9361$$



2) 서열척도로 측정된 현상들간의 상관관계분석

(1) Spearman 서열상관계수

- ◆ 서열척도로 측정된 수는 가감승제의 의미가 없음
 - ☞ Pearson 상관관계 적용 불가 (\because 평균/표준편차 구할 수 없음)
- ◆ Spearman 서열상관계수
 - 서열척도로 측정된 두 개의 변수들간의 상관관계
- ◆ Spearman 서열상관계수의 기본논리
 - 서열척도로 측정된 두 개의 변수들에 대한 응답자의 응답순위가 완전히 일치될 때에는 한 변수의 순위에서 다른 변수의 순위를 뺀 차이는 0이 될 것임
 - 두 변수들간의 서열이 다를수록 차이는 점차 커진다고 할 수 있음
 - ☞ 두 변수의 서열의 차이정보를 가공해서 서열상관계수를 구함

제1절 상관관계분석이란 무엇인가?

〈표 10-3〉 영어와 수학의 석차

시험자 번호	영어석차(X_i)	수학석차(Y_i)	$d_i = X_i - Y_i$	d_i^2
1	1	1	0	0
2	4	3	1	1
3	8	4	4	16
4	7	6	1	1
5	10	10	0	0
6	9	9	0	0
7	5	5	0	0
8	2	2	0	0
9	3	8	- 5	25
10	6	7	- 1	1
n=10	-	-	-	$\sum d_i^2 = 44$



제1절 상관관계분석이란 무엇인가?

Spearman 서열상관계수

(식 10-4)

$$r_s = 1 - 6 \sum d_i^2 / n (n^2 - 1)$$

<표 10-4> Spearman 서열상관계수 계산결과

$$\begin{aligned} r_s &= 1 - 6 \sum d_i^2 / n (n^2 - 1) \\ &= 1 - 6 \times 44 / 10 (100 - 1) \\ &= 1 - 264 / 990 \\ &= 0.73 \end{aligned}$$



제1절 상관관계분석이란 무엇인가?

(2) Kendall의 일치계수

◆ Kendall의 일치계수 (W, Kendall's coefficient of concordance)

- 서열변수로 측정된 **3개 이상의 변수들**, 즉 **k개의 변수들**간에 존재하는 상관관계를 **파악**하는 데 사용 (Spearman 서열상관계수는 두 개 변수들간 존재 상관관계분석 가능)

<표 10-5> 국어, 영어와 수학의 석차

시험자 번호	국어석차(X_i)	영어석차(Y_i)	수학석차(Z_i)	$R_i = X_i + Y_i + Z_i$
1	1	1	1	3
2	4	3	3	10
3	8	4	4	16
4	7	6	6	19
5	10	10	10	30
6	9	9	9	27
7	5	5	5	15
8	2	2	2	6
9	3	8	8	19
10	6	7	7	20
n=10	-	-	-	$\bar{R} = 16.5$



제1절 상관관계분석이란 무엇인가?

◆ Kendall의 일치계수의 기본논리

▣ 변수간의 **서열의 합**을 이용

☞ 만일 3개 변수의 서열의 합이 완벽히 일치한다면, 그 서열의 합은 3, 6, 9, ... 등으로 나타날 것이고, 3개 변수의 서열이 전혀 일치하지 않는다면 그 서열의 합은 거의 동일해질 것임

Kendall의 일치계수(= W)

(식 10-5)

W = 실제로 얻어진 R_i 의 변화폭/ R_i 의 최대의 변화폭

= R_i 편차 제곱의 합 / k 개 변수간 순위가 완전 일치할 때의 R_i 의 총합

$$= \frac{\sum (R_i - \bar{R})^2}{\frac{1}{12} k^2 (n^3 - n)}$$

단, R_i 은 서열순위의 합

\bar{R} 은 서열순위의 합의 평균

k 는 변수(혹은 평가자)의 수

n 은 표본의 크기

<표 10-6> Kendall의 일치계수 계산결과

$$\begin{aligned}
 W &= \sum (R_i - \bar{R})^2 / \frac{1}{12} k^2 (n^3 - n) \\
 &= [(3-16.5)^2 + (10-16.5)^2 + \dots + (20-16.5)^2] / \frac{1}{12} 3^2 (10^3 - 10) \\
 &= (182.25 + 42.25 + \dots + 12.25) / \frac{1}{12} 3^2 (10^3 - 10) \\
 &= 654.5 / 742.5 \\
 &= 0.88
 \end{aligned}$$

☞ Kendall의 일치계수는 부(-)의 값을 갖지 않음

☞ Pearson 상관계수와는 다른 급(class)의 상관관계 판단도구

※ Kendall의 서열상관계수

= Kendall의 τ

= 두 개의 서열변수간의 상관계수

≠ Kendall의 일치계수

3) 명목척도로 측정된 현상들간의 상관관계분석

- ◆ 명목척도로 측정된 수도 가감승제의 의미가 없음
 - ☞ Pearson 상관관계 적용 불가

- ◆ **유관표**(contingency table)
 - 두 변수를 교차(cross)하여 작성한 표

<표 10-7> 성별과 보유 차종 1

	남	여	소 계
Sonata	25(a)	25(c)	50
SM5	25(b)	25(d)	50
소 계	50	50	100

☞ 명목척도로 측정된 두 변수간에 **관련성이 거의 없는** 경우

<표 10-8> 성별과 보유 차종 2

	남	여	소 계
Sonata	50(e)	0(g)	50
SM5	0(f)	50(h)	50
소 계	50	50	100

☞ 명목척도로 측정된 두 변수간에 **관련성이 매우 큰** 경우

* 실제 표본조사 시행 시에는 표 10-7, 10-8의 중간에 해당하는 현상이 관찰됨



제1절 상관관계분석이란 무엇인가?

◆ χ^2 값

- 표본에서 관찰된 결과가 우연에 의한 결과와 의미 있게 다른지를 알려주는 도구
- χ^2 란 우연에 의해서 기대되는 빈도와 실제 관찰되는 빈도를 비교한 값

두 개의 명목변수 χ^2 값 (식 10-6)

$$\chi^2 = \sum \sum (O_{ij} - E_{ij})^2 / E_{ij}$$

단, O_{ij} 는 표본에서 실제 관찰한 빈도

E_{ij} 란 우연에 의해 기대되는 빈도

<표 10-9> χ^2 의 계산 예

<표 10-7>의 χ^2 값(예 a)

$$\begin{aligned}\chi^2 &= [(25-25)^2 + (25-25)^2 + (25-25)^2 + (25-25)^2] / 25 \\ &= (0 + 0 + 0 + 0) / 25 \\ &= 0\end{aligned}$$

<표 10-8>의 χ^2 값(예 b)

$$\begin{aligned}\chi^2 &= [(50-25)^2 + (0-25)^2 + (50-25)^2 + (0-25)^2] / 25 \\ &= (625 + 625 + 625 + 625) / 25 \\ &= 100\end{aligned}$$



제1절 상관관계분석이란 무엇인가?

- ◆ 일반적으로 χ^2 값이 클수록, 표본에서 관찰된 빈도와 우연에 의해 기대되는 빈도와 차이가 커짐
- ◆ χ^2 분포는 자유도(df)에 의해 분포의 구체적인 형태가 변화함
- ◆ χ^2 의 자유도는 <식 10-7>에 의해 결정

* χ^2 의 자유도

(식 10-7)

$$df = [\text{유관표의 행(row)수}-1] [\text{유관표의 열(column)수}-1] \\ = (r-1)(c-1)$$

예) 2×2 표에서는 $df = (2-1)(2-1) = 1$

3×3 표에서는 $df = (3-1)(3-1) = 4$

4×3 표에서는 $df = (4-1)(3-1) = 6$

☞ 실제 관찰빈도와 기대관찰빈도가 계산되고, 이에 따른 χ^2 값, 자유도가 결정되면, 가설설정시에 연구자가 주관적으로 정해 두었던 유의도 수준(그러나 관행적으로는 $p = .01$ 혹은 $.05$)에 비추어 두 변수간에 관계가 있는지 없는지를 판단하게 됨

◆ χ^2 검정(χ^2 test)

- 명목척도로 측정된 변수간에 관계가 있는지 여부를 χ^2 를 이용하여 판단하는 방법
- 관계의 여부만을 판단해 줄 뿐, 두 변수간의 관계의 크기를 분석해 주지는 못함, 즉 두 변수의 독립성 여부만을 판단해 주는 것



제1절 상관관계분석이란 무엇인가?

◆ 유관계수 혹은 분할계수(contingency coefficient: **C**)

- 두 변수간의 관계의 크기를 분석할 수 있는 상관관계분석방법
- 이 계수의 계산도 χ^2 값에 의존

유관계수/분할계수(contingency coefficient[=C]) (식 10-8)

$$C = \sqrt{\frac{\chi^2}{N + \chi^2}}$$

<표 10-10> 유관계수/분할계수 계산 예

$$C = \sqrt{\frac{100}{100 + 100}}$$

$$=.707$$

- * 정의상 유관계수는 Pearson 상관계수처럼 1이 될 수 없으므로, 두 변수간에 완벽한 상관관계가 있는 경우에도 크기가 실제보다 작게 계산될 수 있음,
이 계수에 부호도 없고 여러가지 측면에서 해석상의 어려움이 있음
=> 완벽하지는 않더라도 **카이제곱 (χ^2)검정과 유관계수(C)를 동시에 사용해서 자료를 해석하는 것이 바람직함**



제2절 상관관계분석 결과는 어떻게 해석할 것인가?

◆ 상관관계분석

- 표본을 기초로 한 기술통계분석 결과를 이용하여 모집단에서의 변수들간 관계를 추정하는 추리통계분석기법
- 표본을 대상으로 관찰한 변수들의 측정값으로 표본통계량을 계산하고, 표본통계량과 해당 표본통계량의 표본분포(sampling distribution)에 근거해서 모집단에서도 상관관계가 성립하는지에 대한 가설을 검정하게 되는 것

◆ 모든 상관관계분석시의 가설의 일반적인 형태(<표 10-11> 참조)

<표 10-11> 상관관계분석시의 가설

연구가설(H_1): 두 변수(현상)간에는 상관관계가 존재할 것이다.

귀무가설(H_0): 두 변수(현상)간에는 상관관계가 존재하지 않을 것이다.



1) 등간/비율척도로 측정된 현상들간의 상관관계분석결과의 해석

◆ 분석의 대상 변수는

- '현재급여'와 '최초급여'
- 급여이므로 비율척도의 수준으로 측정된 사회현상

◆ 연구가설(H_1)

- "현재급여와 최초급여간에는 상관관계가 존재할 것이다"

◆ 귀무가설(H_0)

- "현재급여와 최초급여간에는 상관관계가 존재하지 않을 것이다"

◆ 통계가설형태로 표현된 연구가설(H_1)과 귀무가설(H_0)

☞ <표 10-12>

☞ $H_0: \rho = 0$ 라는 가정하에서, 표본에서 얻은 r 이 모집단의 ρ 로부터 얼마나 멀리 떨어져 있는지를 판단하는 것이 상관관계분석시의 가설검정내용

☞ 구체적으로 r 을 이용한 검정통계량(test statistic)을 계산해서 가설을 검정



<표 10-12> Pearson 상관관계분석시 가설

$$H_1 : \rho \neq 0$$

$$H_0 : \rho = 0$$

Pearson 상관계수의 검정통계량 (식 10-9)

$$t = \frac{r \sqrt{n-2}}{\sqrt{1-r^2}}$$

단, n 은 표본의 크기

r 은 표본에서 계산한 Pearson 상관계수의 크기

t 는 $n-2$ 의 자유도를 가진 t 분포를 따름



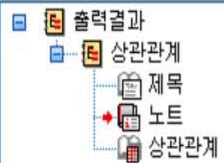
제2절 상관관계분석 결과는 어떻게 해석할 것인가?



[그림 10-3] Pearson 상관계수의 해석

*출력결과1 [문서1] - IBM SPSS Statistics Viewer

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 삽입(I) 형식(O) 분석(A) 다이렉트 마케팅(M) 그래프(G) 유틸리티(U) 확장(X) 창(W) 도움말(H)



상관관계

상관관계

		현재급여	최초급여
현재급여	Pearson 상관	1	.880**
	유의확률 (양측)		.000
	N	474	474
최초급여	Pearson 상관	.880**	1
	유의확률 (양측)	.000	
	N	474	474

** . 상관관계가 0.01 수준에서 유의합니다(양측).



2) 서열척도로 측정된 현상들간의 상관관계분석결과의 해석

- ◆ 분석의 대상 변수는
 - '직종구분'과 '피교육연수'
 - 서열척도로 가정
- ◆ 연구가설(H_1)
 - "직종구분과 피교육연수간에는 상관관계가 존재할 것이다"
- ◆ 귀무가설(H_0)
 - "직종구분과 피교육연수간에는 상관관계가 존재하지 않을 것이다"



(1) Spearman 서열상관계수의 해석

◆ 통계가설형태로 표현된 귀무가설(H_0)

□ " $H_0: \rho = 0$ "

☞ $H_0: \rho=0$ 라는 가정하에서, 표본에서 얻은 r_s 이 모집단의 ρ 로부터 얼마나 멀리 떨어져 있는지를 판단하는 것이 서열상관관계분석시의 가설검정내용

⇒ 구체적으로 r_s 을 이용한 검정통계량(test statistic)을 계산해서 가설을 검정
Spearman 서열상관계수 검정통계량 (식 10-10)

$$t = r_s \sqrt{\frac{N-2}{1-(r_s)^2}}$$

단, t 는 $n-2$ 의 자유도를 가진 t 분포를 따름



제2절 상관관계분석 결과는 어떻게 해석할 것인가?



[그림 10-4] 서열상관계수의 해석

*출력결과2 [문서2] - IBM SPSS Statistics Viewer

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 삽입(I) 형식(O) 분석(A) 다이렉트 마케팅(M) 그래프(G) 유틸리티(U) 확장(X) 창(W) 도움말(H)



출력결과
비모수 상관
제목
노트
상관관계

비모수 상관

상관관계

			피교육년수	직종구분
Kendall의 타우-b	피교육년수	상관계수	1.000	.415**
		유의확률 (양측)	.	.000
		N	474	474
	직종구분	상관계수	.415**	1.000
		유의확률 (양측)	.000	.
		N	474	474
Spearman의 rho	피교육년수	상관계수	1.000	.484**
		유의확률 (양측)	.	.000
		N	474	474
	직종구분	상관계수	.484**	1.000
		유의확률 (양측)	.000	.
		N	474	474

** 상관계수가 0.01 수준에서 유의합니다(양측).



(2) Kendall 서열상관계수의 해석

◆ 통계가설형태로 표현된 귀무가설(H_0)

□ " $H_0: \tau = 0$ "

☞ 즉, $H_0: \tau = 0$ 라는 가정하에서, 표본에서 얻은 Kendall 서열상관계수가 모집단의 τ 로부터 얼마나 멀리 떨어져 있는지를 판단하는 것이 Kendall 서열상관관계분석시의 가설검정내용

Kendall 서열상관계수 검정통계량

(식 10-11)

$$z = \frac{\tau - \mu_\tau}{\sigma_\tau} = \frac{\tau}{\sqrt{\frac{2(2N+5)}{9N(N-1)}}}$$

단, z 검정통계량은 평균 $\mu_\tau = 0$, 표준편차 $\sigma_\tau = \sqrt{\frac{2(2N+5)}{9N(N-1)}}$ 를 가지는 정규분포에 접근

⇒ 다시 [그림 10-4] 참조



3) 명목척도로 측정된 현상들간의 상관관계분석 결과의 해석

- ◆ 분석 대상 변수
 - '소수민족여부'와 '직종구분'
- ◆ 연구가설(H_1)
 - "소수민족여부와 직종구분간에는 상관관계가 존재할 것이다"
- ◆ 귀무가설(H_0)
 - "소수민족여부와 직종구분간에는 상관관계가 존재하지 않을 것이다"

(1) χ^2 검정

- ◆ 명목변수의 실제 관찰빈도
- ◆ 기대관찰빈도
 - 우연(chance)에 의해 기대할 수 있는 빈도
 - "두 변수는 서로 독립적이다(즉, 관계가 없다)"는 귀무가설(H_0)하에서 기대할 수 있는 빈도
- ◆ χ^2 값과 자유도 결정
 - ☞ χ^2 값이 클수록 실제관찰 빈도가 기대빈도와 차이가 많이 나게 되고, 따라서 이러한 결과는 우연(chance)에 의한 것이 아닐 가능성, 즉 귀무가설이 거짓일 가능성을 제고시키는 것



제2절 상관관계분석 결과는 어떻게 해석할 것인가?



[그림 10-5] 명목변수간의 상관관계의 해석 1

*출력결과1 [문서1] - IBM SPSS Statistics Viewer

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 삽입(I) 형식(O) 분석(A) 다이렉트 마케팅(M) 그래프(G) 유틸리티(U) 확장(X) 창(W) 도움말(H)



- 출력결과
 - 교차분석
 - 제목
 - 노트
 - 케이스 처리 요약
 - 직종구분 * 소수민족
 - 카이제곱 검정

교차분석

케이스 처리 요약

	유효		케이스 결측		전체	
	N	퍼센트	N	퍼센트	N	퍼센트
직종구분 * 소수민족 여부	474	100.0%	0	0.0%	474	100.0%

직종구분 * 소수민족 여부 교차표

빈도

		소수민족 여부		전체
		아니오	예	
직종구분	사무직	276	87	363
	관리직	14	13	27
	경영직	80	4	84
전체		370	104	474

카이제곱 검정

	값	자유도	근사 유의확률 (양측검정)
Pearson 카이제곱	26.172 ^a	2	.000



(2) 유관계수/분할계수 분석결과

◆ 유관계수

- 명목척도로 측정된 두 변수간의 상관관계의 크기까지는 분석

<표 10-13> 명목변수간의 상관관계의 해석 2

대칭적 측도

		값	근사 유의확률
명목척도 대 명목척도	파이	.235	.000
	Cramer의 V	.235	.000
	분할계수	.229	.000
유효 케이스 수		474	

제3절 상관관계분석은 어떻게 실행하는가?

- ◆ 모든 SPSS Statistics 분석을 위해서는 분석대상 자료파일이 열려져 있어야 함
- [그림 10-6] 분석대상 자료파일 열기**

Employee data.sav [데이터세트1] - IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A) 다이렉트 마케팅(M) 그래프(G) 유틸리티(U) 확장(X) 창(W) 도움말(H)

표시: 10 / 10 변수

	번호	성별	생년월일	피교육	직종	현재급여	최초급여	근무월수	경력	소수민족	변수	변수	변수	변수	변수
1	1 m		02/03/1952	15	3	\$57,000	\$27,000	98	144	0					
2	2 m		05/23/1958	16	1	\$40,200	\$18,750	98	36	0					
3	3 f		07/26/1929	12	1	\$21,450	\$12,000	98	381	0					
4	4 f		04/15/1947	8	1	\$21,900	\$13,200	98	190	0					
5	5 m		02/09/1955	15	1	\$45,000	\$21,000	98	138	0					
6	6 m		08/22/1958	15	1	\$32,100	\$13,500	98	67	0					
7	7 m		04/26/1956	15	1	\$36,000	\$18,750	98	114	0					
8	8 f		05/06/1966	12	1	\$21,900	\$9,750	98	0	0					
9	9 f		01/23/1946	15	1	\$27,900	\$12,750	98	115	0					
10	10 f		02/13/1946	12	1	\$24,000	\$13,500	98	244	0					
11	11 f		02/07/1950	16	1	\$30,300	\$16,500	98	143	0					
12	12 m		01/11/1966	8	1	\$28,350	\$12,000	98	26	1					
13	13 m		07/17/1960	15	1	\$27,750	\$14,250	98	34	1					
14	14 f		02/26/1949	15	1	\$35,100	\$16,800	98	137	1					
15	15 m		08/29/1962	12	1	\$27,300	\$13,500	97	66	0					
16	16 m		11/17/1964	12	1	\$40,800	\$15,000	97	24	0					
17	17 m		07/18/1962	15	1	\$46,000	\$14,250	97	48	0					
18	18 m		03/20/1956	16	3	\$103,750	\$27,510	97	70	0					
19	19 m		08/19/1962	12	1	\$42,300	\$14,250	97	103	0					
20	20 f		01/23/1940	12	1	\$26,250	\$11,550	97	48	0					

데이터 보기 변수 보기

IBM SPSS Statistics 프로세서 준비 완료

Unicode:ON



1) 등간/비율척도로 측정된 현상들간의 상관관계분석의 실행

- ◆ [그림 10-6]처럼 분석대상 자료파일을 열고, [그림 10-7]과 같이 **분석(A) → 상관분석(C) → 이변량 상관관계수(B)**를 차례로 선택
- ◆ [그림 10-7]에서와 같은 절차를 밟게 되면, [그림 10-8]에서와 같은 작은 창이 열리게 됨
 - ☞ 이 창에서는 상관관계분석대상 변수 두 개를 지정하는 창, 두 개 변수간의 상관계수(즉, 이변량 상관관계)를 분석하는 방법 3개, 그리고 가설의 유의성 검정시 단측검정(즉, 한쪽검정) 혹은 양측검정(즉, 양쪽검정) 여부를 선택할 수 있는 등의 선택을 수행할 수 있게 되어 있음
 - ☞ 이변량 상관관계수 분석 창에서 먼저, 분석대상 변수 두 개를 선택
 - ☞ [그림 10-8]에서의 분석대상 변수: '현재급여'와 '최초급여'



제3절 상관관계분석은 어떻게 실행하는가?

[그림 10-7] 등간/비율척도로 측정된 현상들간의 상관관계분석의 실행

Employee data.sav [데이터세트1] - IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A) 다이렉트 마케팅(M) 그래프(G) 유틸리티(U) 확장(X) 창(W) 도움말(H)

보고서(P) 기술통계량(E) 표(B) 평균 비교(M) 일반선형모형(G) 일반화 선형 모형(Z) 혼합 모형(X) 상관분석(C) 회귀분석(R) 로그선형분석(Q) 신경망(W) 분류분석(F) 차원 축소(D) 척도분석(A) 비모수검정(N) 시계열 분석(T) 생존분석(S) 다중반응(U) 결측값 분석(Y)... 다중대체(T) 복합 표본(L) 시뮬레이션(I)... 품질관리(Q) ROC 곡선(V)... 공간과 시간 모형화(S)...

표시: 10 / 10 변수

번호	성별	생년월일	현재급여	최초급여	근무월수	경력	소수민족	변수	변수	변수	변수	변수
1	m	02/03/19	\$57,000	\$27,000	98	144	0					
2	m	05/23/19	\$40,200	\$18,750	98	36	0					
3	f	07/26/19	\$21,450	\$12,000	98	381	0					
4	f	04/15/19	\$21,900	\$13,200	98	190	0					
5	m	02/09/19			98	138	0					
6	m	08/22/19			98	67	0					
7	m	04/26/19			98	114	0					
8	f	05/06/19			98	0	0					
9	f	01/23/19			98	115	0					
10	f	02/13/19	\$24,000	\$13,500	98	244	0					
11	f	02/07/19	\$30,300	\$16,500	98	143	0					
12	m	01/11/19	\$28,350	\$12,000	98	26	1					
13	m	07/17/19	\$27,750	\$14,250	98	34	1					
14	f	02/26/19	\$35,100	\$16,800	98	137	1					
15	m	08/29/19	\$27,300	\$13,500	97	66	0					
16	m	11/17/19	\$40,800	\$15,000	97	24	0					
17	m	07/18/19	\$46,000	\$14,250	97	48	0					
18	m	03/20/19	\$103,750	\$27,510	97	70	0					
19	m	08/19/19	\$42,300	\$14,250	97	103	0					
20	f	01/23/19	\$26,250	\$11,550	97	48	0					

데이터 보기 변수 보기

이변량 상관(B)...

IBM SPSS Statistics 프로세서 준비 완료

Unicode:ON



제3절 상관관계분석은 어떻게 실행하는가?



[그림 10-8] 이변량 상관계수분석 창

Employee data.sav [데이터세트1] - IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A) 다이렉트 마케팅(M) 그래프(G) 유틸리티(U) 확장(X) 창(W) 도움말(H)

표시: 10 / 10 변수

번호	성별	생년월일	피교육	직종	현재급여	최초급여	근무월수	경력	소수민족	변수	변수	변수	변수	변수
1	m	02/03/1952	15	3	\$57,000	\$27,000	98	144	0					
2	m	05/23/1958						36	0					
3	f	07/26/1929						381	0					
4	f	04/15/1947						190	0					
5	m	02/09/1955						138	0					
6	m	08/22/1958						67	0					
7	m	04/26/1956						114	0					
8	f	05/06/1966						0	0					
9	f	01/23/1946						115	0					
10	f	02/13/1946						244	0					
11	f	02/07/1950						143	0					
12	m	01/11/1966						26	1					
13	m	07/17/1960						34	1					
14	f	02/26/1949						137	1					
15	m	08/29/1962						66	0					
16	m	11/17/1964						24	0					
17	m	07/18/1962						48	0					
18	m	03/20/1956						70	0					
19	m	08/19/1962						103	0					
20	f	01/23/1940	12	1	\$26,250	\$11,550	97	48	0					

이변량 상관계수

변수(V):

- 사원코드 [번호]
- 생년월일 [생년월일]
- 피교육년수 [피교육]
- 직종구분 [직종]
- 현 근무월수 [근무월수]
- 입사전 타 근무월수 [...]
- 소수민족 여부 [소수...]

상관계수

☒ Pearson ☐ Kendall의 타우-b ☐ Spearman

유의성 검정

☒ 양측(T) ☐ 단측(L)

☒ 유의한 상관계수 플래그(E)

확인 불여넣기(P) 재설정(R) 취소 도움말

데이터 보기 변수 보기

IBM SPSS Statistics 프로세서 준비 완료 Unicode:ON



제3절 상관관계분석은 어떻게 실행하는가?

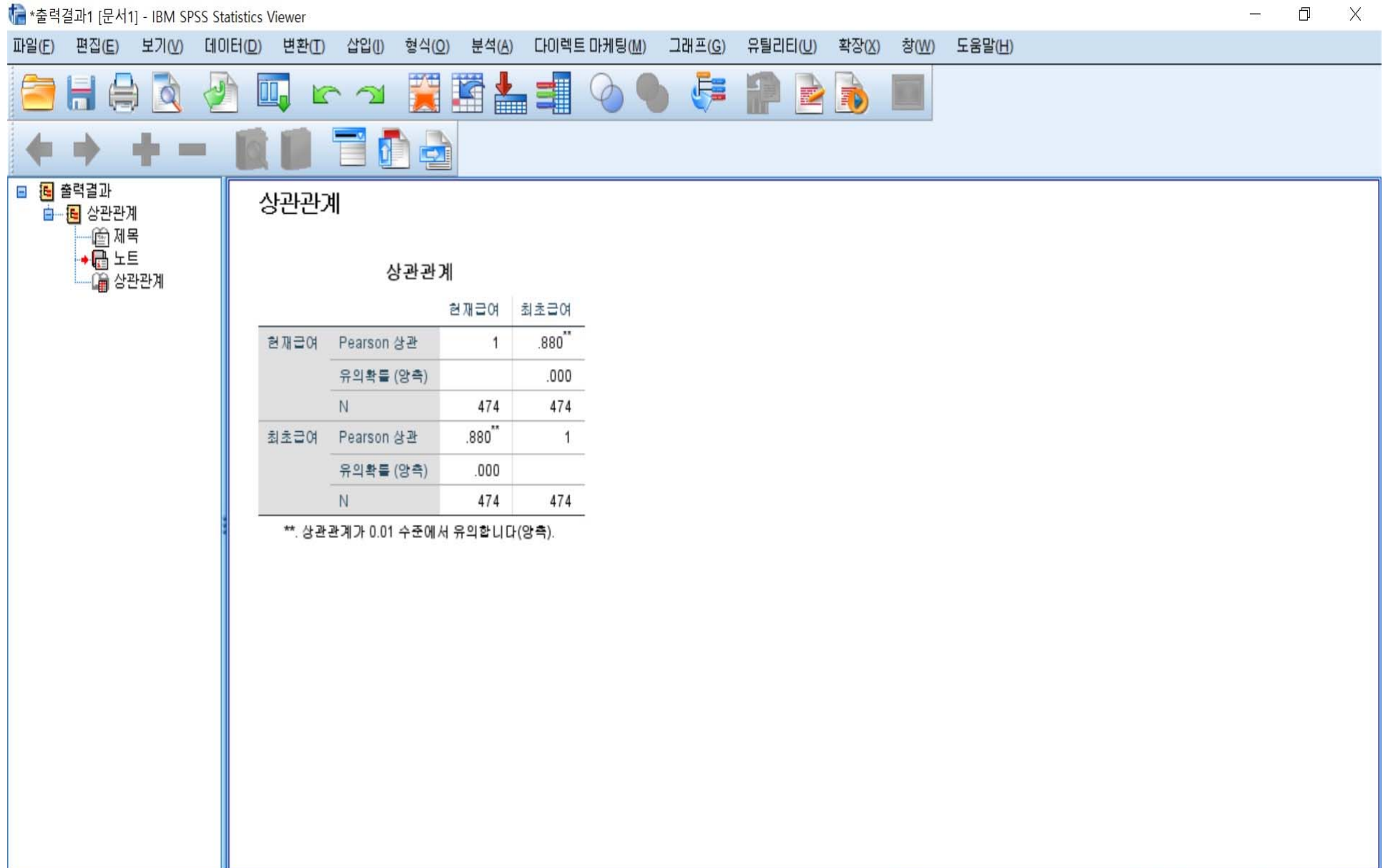
- ◆ 분석대상 변수 두 개를 선택한 다음에는, **Pearson(N)**을 클릭(체크)합니다. 즉, 3개의 상관계수기법 중 등간/비율척도 수준으로 측정된 변수들간의 상관계수를 분석해주는 Pearson 상관관계분석을 선택
- ◆ 상관계수분석시 가설은 "두 변수간에 상관관계가 있다"이므로 가설에는 + 혹은 -의 방향이 없음
 - ☞ 검정통계량의 유의성 검정은 **양쪽**을 선택합니다. 한편, 상관계수가 유의하게 판명되는 경우 별(*)표시를 하라는 옵션도 선택을 하는 것이 일반적
- ◆ 다음 **확인** 버튼을 클릭하게 되면, [그림 10-9]와 같은 결과 창을 얻게 됨



제3절 상관관계분석은 어떻게 실행하는가?



[그림 10-9] Pearson 상관관계분석 결과 창





2) 서열척도로 측정된 현상들간의 서열상관관계 분석의 실행

- ◆ 먼저 [그림 10-6]처럼 분석대상 자료파일을 열고,
- ◆ [그림 10-7]과 같이 **분석(A)** → **상관분석(C)** → **이변량 상관계수(B)**를 차례로 선택
- ◆ [그림 10-10]에서와 같은 절차를 밟게 되면, [그림 10-11]에서와 같은 작은 창이 열리게 됨
 - Pearson 상관관계를 선택할 때와 동일한 창
 - ☞ 이 창에서는 상관관계분석대상 변수 두 개를 지정하는 창, 두 개 변수간의 상관계수(즉, 이변량 상관관계)를 분석하는 방법 3개, 그리고 가설의 유의성 검정시 단측검정(즉, 한쪽검정) 혹은 양측검정(즉, 양쪽검정) 여부를 선택할 수 있는 등의 선택을 수행할 수 있음
 - ☞ 분석대상 변수 두 개를 선택



제3절 상관관계분석은 어떻게 실행하는가?



[그림 10-10] 서열적으로 측정된 현상들간의 상관관계분석의 실행

Employee data.sav [데이터세트1] - IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A) 다이렉트 마케팅(M) 그래프(G) 유틸리티(U) 확장(X) 창(W) 도움말(H)

보고서(P) 기술통계량(E) 표(B) 평균 비교(M) 일반선형모형(G) 일반화 선형 모형(Z) 혼합 모형(X) 상관분석(C) 회귀분석(R) 로그선형분석(Q) 신경망(W) 분류분석(F) 차원 축소(D) 척도분석(A) 비모수검정(N) 시계열 분석(T) 생존분석(S) 다중반응(U) 결측값 분석(Y) 다중대체(T) 복합 표본(L) 시뮬레이션(I) 품질관리(Q) ROC 곡선(V) 공간과 시간 모형화(S)

표시: 10 / 10 변수

번호	성별	생년월일	현재급여	최초급여	근무월수	경력	소수민족	변수	변수	변수	변수	변수
1	m	02/03/19	\$57,000	\$27,000	98	144	0					
2	m	05/23/19	\$40,200	\$18,750	98	36	0					
3	f	07/26/19	\$21,450	\$12,000	98	381	0					
4	f	04/15/19	\$21,900	\$13,200	98	190	0					
5	m	02/09/19			98	138	0					
6	m	08/22/19			98	67	0					
7	m	04/26/19			98	114	0					
8	f	05/06/19			98	0	0					
9	f	01/23/19			98	115	0					
10	f	02/13/19	\$24,000	\$13,500	98	244	0					
11	f	02/07/19	\$30,300	\$16,500	98	143	0					
12	m	01/11/19	\$28,350	\$12,000	98	26	1					
13	m	07/17/19	\$27,750	\$14,250	98	34	1					
14	f	02/26/19	\$35,100	\$16,800	98	137	1					
15	m	08/29/19	\$27,300	\$13,500	97	66	0					
16	m	11/17/19	\$40,800	\$15,000	97	24	0					
17	m	07/18/19	\$46,000	\$14,250	97	48	0					
18	m	03/20/19	\$103,750	\$27,510	97	70	0					
19	m	08/19/19	\$42,300	\$14,250	97	103	0					
20	f	01/23/19	\$26,250	\$11,550	97	48	0					

데이터 보기 변수 보기

이변량 상관(B)...

IBM SPSS Statistics 프로세서 준비 완료

Unicode:ON



제3절 상관관계분석은 어떻게 실행하는가?



[그림 10-11] 이변량 상관관계수 분석 창

Employee data.sav [데이터세트1] - IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A) 다이렉트 마케팅(M) 그래프(G) 유틸리티(U) 확장(X) 창(W) 도움말(H)



표시: 10 / 10 변수

	번호	성별	생년월일	피교육	직종	현재급여	최초급여	근무월수	경력	소수민족	변수	변수	변수	변수	변수
1	1	m	02/03/1952	15	3	\$57,000	\$27,000	98	144	0					
2	2	m	05/23/1958						36	0					
3	3	f	07/26/1929						381	0					
4	4	f	04/15/1947						190	0					
5	5	m	02/09/1955						138	0					
6	6	m	08/22/1958						67	0					
7	7	m	04/26/1956						114	0					
8	8	f	05/06/1966						0	0					
9	9	f	01/23/1946						115	0					
10	10	f	02/13/1946						244	0					
11	11	f	02/07/1950						143	0					
12	12	m	01/11/1966						26	1					
13	13	m	07/17/1960						34	1					
14	14	f	02/26/1949						137	1					
15	15	m	08/29/1962						66	0					
16	16	m	11/17/1964						24	0					
17	17	m	07/18/1962						48	0					
18	18	m	03/20/1956						70	0					
19	19	m	08/19/1962						103	0					
20	20	f	01/23/1940	12	1	\$26,250	\$11,550	97	48	0					

이변량 상관관계수

변수(V):

- 직종구분 [직종]
- 피교육년수 [피교육]

상관계수

☐ Pearson ☒ Kendall의 타우-b ☒ Spearman

유의성 검정

☒ 양측(T) ☐ 단측(L)

☒ 유의한 상관관계수 플래그(E)

확인 불여넣기(P) 재설정(R) 취소 도움말

데이터 보기 변수 보기



제3절 상관관계분석은 어떻게 실행하는가?



[그림 10-12] 서열상관계수 분석결과 창

*출력결과1 [문서1] - IBM SPSS Statistics Viewer

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 삽입(I) 형식(O) 분석(A) 다이렉트 마케팅(M) 그래프(G) 유틸리티(U) 확장(X) 창(W) 도움말(H)

출력결과
 비모수 상관
 제목
 노트
 상관관계

비모수 상관

상관관계

		직종구분	피교육년수
Kendall의 타우-b	직종구분	상관계수	1.000
		유의확률 (양측)	.000
		N	474
	피교육년수	상관계수	.415**
		유의확률 (양측)	.000
		N	474
Spearman의 rho	직종구분	상관계수	1.000
		유의확률 (양측)	.000
		N	474
	피교육년수	상관계수	.484**
		유의확률 (양측)	.000
		N	474

** . 상관관계가 0.01 수준에서 유의합니다(양측).



제3절 상관관계분석은 어떻게 실행하는가?

- ◆ 이어서 확인 버튼을 클릭

☞ [그림 10-12]와 같은 결과 창을 얻게 됨

- ◆ 분석결과를 해석

3) 명목척도로 측정된 현상들간의 상관관계분석의 실행

- ◆ 먼저 [그림 10-6]처럼 분석대상 자료파일을 열고,

- ◆ [그림 10-13]에서와 같이 **분석(A)** → **기술통계량(E)** → **교차분석(C)**을 차례로 선택



제3절 상관관계분석은 어떻게 실행하는가?



[그림 10-13] 명목척도로 측정된 현상들간의 상관관계분석의 실행

Employee data.sav [데이터세트1] - IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A) 다이렉트 마케팅(M) 그래프(G) 유틸리티(U) 확장(X) 창(W) 도움말(H)

보고서(P) 기술통계량(E) 표(B) 평균 비교(M) 일반선형모형(G) 일반화 선형 모형(Z) 혼합 모형(X) 상관분석(C) 회귀분석(R) 로그선형분석(Q) 신경망(W) 분류분석(F) 차원 축소(D) 척도분석(A) 비모수검정(N) 시계열 분석(T) 생존분석(S) 다중반응(U) 결측값 분석(Y)... 다중대체(T) 복합 표본(L) 시뮬레이션(I)... 품질관리(Q) ROC 곡선(V)... 공간과 시간 모형화(S)...

빈도분석(F)... 기술통계(D)... 데이터 탐색(E)... 교차분석(C)... TURF 분석 비율통계량(R)... P-P 도표... Q-Q 도표...

표시: 10 / 10 변수

번호	성별	생년월일	근무월수	경력	소수민족	변수	변수	변수	변수	변수
1	m	02/03/19	98	144	0					
2	m	05/23/19	98	36	0					
3	f	07/26/19	98	381	0					
4	f	04/15/19	98	190	0					
5	m	02/09/19	98	138	0					
6	m	08/22/19	98	67	0					
7	m	04/26/19	98	114	0					
8	f	05/06/19	\$21,900	\$9,750	98	0	0			
9	f	01/23/19	\$27,900	\$12,750	98	115	0			
10	f	02/13/19	\$24,000	\$13,500	98	244	0			
11	f	02/07/19	\$30,300	\$16,500	98	143	0			
12	m	01/11/19	\$28,350	\$12,000	98	26	1			
13	m	07/17/19	\$27,750	\$14,250	98	34	1			
14	f	02/26/19	\$35,100	\$16,800	98	137	1			
15	m	08/29/19	\$27,300	\$13,500	97	66	0			
16	m	11/17/19	\$40,800	\$15,000	97	24	0			
17	m	07/18/19	\$46,000	\$14,250	97	48	0			
18	m	03/20/19	\$103,750	\$27,510	97	70	0			
19	m	08/19/19	\$42,300	\$14,250	97	103	0			
20	f	01/23/19	\$26,250	\$11,550	97	48	0			

데이터 보기 변수 보기

교차분석(C)...

IBM SPSS Statistics 프로세서 준비 완료

Unicode:ON



제3절 상관관계분석은 어떻게 실행하는가?

- ◆ [그림 10-13]에서와 같은 절차를 밟게 되면, [그림 10-14]에서와 같은 교차분석 창이 열리게 됨
 - ☞ 이 창은 상관관계 분석대상 변수 두 개를 지정하는 창, 두 개 변수간의 상관관계의 크기를 분석하는 방법을 선택할 수 있는 '통계량'버튼 등의 선택을 수행할 수 있게 설계되어 있음

- ◆ 교차분석 창에서 분석대상 변수 두 개를 선택
 - ☞ 목표는 두 변수간에 존재하는 상관관계를 파악하는 것이기에, 행과 열에 들어가는 변수를 서로 바꾸어도 무방



제3절 상관관계분석은 어떻게 실행하는가?

[그림 10-14] 교차분석 창

Employee data.sav [데이터세트1] - IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A) 다이렉트 마케팅(M) 그래프(G) 유틸리티(U) 확장(X) 창(W) 도움말(H)

표시: 10 / 10 변수

번호	성별	생년월일	피교육	직종	현재급여	최초급여	근무월수	경력	소수민족	변수	변수	변수	변수	변수
1	1 m							144	0					
2	2 m							36	0					
3	3 f							381	0					
4	4 f							190	0					
5	5 m							138	0					
6	6 m							67	0					
7	7 m							114	0					
8	8 f							0	0					
9	9 f							115	0					
10	10 f							244	0					
11	11 f							143	0					
12	12 m							26	1					
13	13 m							34	1					
14	14 f							137	1					
15	15 m							66	0					
16	16 m							24	0					
17	17 m							48	0					
18	18 m							70	0					
19	19 m							103	0					
20	20 f							48	0					

교차분석

행(R):

열(C):

레이어 1/1

미전(N) 다음(N)

표 레이어에 레이어 변수 표시(L)

수평배열 막대도표 표시(B)

교차표를 출력하지 않음(I)

확인 붙여넣기(P) 재설정(R) 취소 도움말

데이터 보기 변수 보기

IBM SPSS Statistics 프로세서 준비 완료 Unicode:ON



제3절 상관관계분석은 어떻게 실행하는가?

[그림 10-15] 변수 선택

Employee data.sav [데이터세트1] - IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A) 다이렉트 마케팅(M) 그래프(G) 유틸리티(U) 확장(X) 창(W) 도움말(H)

표시: 10 / 10 변수

	번호	성별	생년월일	피교육	직종	현재급여	최초급여	근무월수	경력	소수민족	변수	변수	변수	변수	변수
1	1	m							144	0					
2	2	m							36	0					
3	3	f							381	0					
4	4	f							190	0					
5	5	m							138	0					
6	6	m							67	0					
7	7	m							114	0					
8	8	f							0	0					
9	9	f							115	0					
10	10	f							244	0					
11	11	f							143	0					
12	12	m							26	1					
13	13	m							34	1					
14	14	f							137	1					
15	15	m							66	0					
16	16	m							24	0					
17	17	m							48	0					
18	18	m							70	0					
19	19	m							103	0					
20	20	f							48	0					

교차분석

행(R): 직종구분 [직종]

열(C): 소수민족 여부 [소수민족]

레이어 1/1

표 레이어에 레이어 변수 표시(L)

수평배열 막대도표 표시(B)

교차표를 출력하지 않음(I)

확인 붙여넣기(P) 재설정(R) 취소 도움말

데이터 보기 변수 보기

IBM SPSS Statistics 프로세서 준비 완료 Unicode:ON



제3절 상관관계분석은 어떻게 실행하는가?

- ◆ 다음에는 교차분석 창의 **통계량** 버튼을 눌러서 명목척도로 측정된 두 변수간의 관계가 존재하는지를 검정으로 파악하고
 - ☞ 즉, [그림 10-16]의 통계량 창에서 **카이제곱(H)**과 **분할계수(O)**를 클릭함
- ◆ 교차분석 창에서 **계속** 버튼을 클릭하게 되면 [그림 10-17]과 같이 교차분석 창이 다시 열리게 됨
- ◆ 이어서 교차분석 창의 **확인**을 클릭하게 되면, [그림 10-18]과 같이 명목변수 수준에서 측정된 변수들간에 존재하는 상관관계 그리고 이 상관관계의 크기가 계산됨



제3절 상관관계분석은 어떻게 실행하는가?

[그림 10-16] 통계량 창에서 χ^2 검정 및 분할계수의 선택

Employee data.sav [데이터세트1] - IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A) 다이렉트 마케팅(M) 그래프(G) 유틸리티(U) 확장(X) 창(W) 도움말(H)

표시: 10 / 10 변수

번호	성별	생년월일	피교육	직종	현재급여	최초급여	근무월수	경력	소수민족	변수	변수	변수	변수	변수
1	1 m							144	0					
2	2 m							36	0					
3	3 f													
4	4 f													
5	5 m													
6	6 m													
7	7 m													
8	8 f													
9	9 f													
10	10 f													
11	11 f													
12	12 m													
13	13 m													
14	14 f													
15	15 m													
16	16 m													
17	17 m													
18	18 m													
19	19 m							103	0					
20	20 f							48	0					

교차분석

행(R): 직종구분 [직종]

열(C): 소수민족 여부 [소수민...]

레이어 1/1

표 레이어에 레이어 변수 표시(L)

수평배열 막대도표 표시(B)

교차표를 출력하지 않음(I)

확인 붙여넣기(P) 재설정(R) 취소 도움말

교차분석: 통계량

☒ 카이제곱(H) ☐ 상관관계(R)

명목

☒ 분할계수(O) ☐ 순서형

☒ 파이 및 크레이머의 V(P) ☐ 감마(G)

☒ 람다(L) ☐ Somers의 d

☒ 불확실성 계수(U) ☐ Kendall의 타우-b

☐ Kendall의 타우-c

명목 대 구간

☐ 에타(E) ☐ 카파(K)

☐ 위험도(I) ☐ McNemar

☐ Cochran 및 Mantel-Haenszel 통계량

검정할 공통오즈비(O): 1

계속(C) 취소 도움말

데이터 보기 변수 보기

IBM SPSS Statistics 프로세서 준비 완료

Unicode:ON



제3절 상관관계분석은 어떻게 실행하는가?

[그림 10-17] 교차분석 창에서 χ^2 검정 및 분할계수계산

Employee data.sav [데이터세트1] - IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A) 다이렉트 마케팅(M) 그래프(G) 유틸리티(U) 확장(X) 창(W) 도움말(H)

표시: 10 / 10 변수

번호	성별	생년월일	피교육	직종	현재급여	최초급여	근무월수	경력	소수민족	변수	변수	변수	변수	변수
1	1 m							144	0					
2	2 m							36	0					
3	3 f							381	0					
4	4 f							190	0					
5	5 m							138	0					
6	6 m							67	0					
7	7 m							114	0					
8	8 f							0	0					
9	9 f							115	0					
10	10 f							244	0					
11	11 f							143	0					
12	12 m							26	1					
13	13 m							34	1					
14	14 f							137	1					
15	15 m							66	0					
16	16 m							24	0					
17	17 m							48	0					
18	18 m							70	0					
19	19 m							103	0					
20	20 f							48	0					

교차분석

행(R): 직종구분 [직종]

열(C): 소수민족 여부 [소수민족]

레이어 1/1

이전(B) 다음(N)

☐ 표 레이어에 레이어 변수 표시(L)

☐ 수평배열 막대도표 표시(B)

☐ 교차표를 출력하지 않음(I)

확인 붙여넣기(P) 재설정(R) 취소 도움말

정확(X)... 통계량(S)... 셀(E)... 형식(F)... 유형(L)... 북스트랩(A)...

데이터 보기 변수 보기

IBM SPSS Statistics 프로세서 준비 완료 Unicode:ON



제3절 상관관계분석은 어떻게 실행하는가?



[그림 10-18] 교차분석결과: 명목변수간의 상관관계 분석 결과

*출력결과1 [문서1] - IBM SPSS Statistics Viewer

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 삽입(I) 형식(O) 분석(A) 다이렉트 마케팅(M) 그래프(G) 유틸리티(U) 확장(X) 창(W) 도움말(H)

출력결과
교차분석
제목
노트
케이스 처리 요약
직종구분 * 소수민족
카이제곱 검정
대형적 측도

교차분석

케이스 처리 요약

	유효		케이스 결측		전체	
	N	퍼센트	N	퍼센트	N	퍼센트
직종구분 * 소수민족 여부	474	100.0%	0	0.0%	474	100.0%

직종구분 * 소수민족 여부 교차표

빈도

		소수민족 여부		전체
		아니오	예	
직종구분	사무직	276	87	363
	관리직	14	13	27
	경영직	80	4	84
전체		370	104	474

카이제곱 검정

	값	자유도	근사 유의확률 (양측검정)
Pearson 카이제곱	26.172 ^a	2	.000
우도비	29.436	2	.000
선형 대 선형 연관	9.778	1	.002

IBM SPSS Statistics 프로세서 준비 완료 | Unicode:ON | H: 398, W: 756 pt