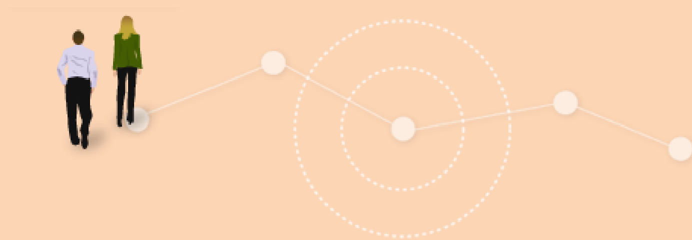
9 장

목 차

01 분산분석

:: Keywords 분산분석의 개념 | 분산분석의 구분 | 분산분석의 가정



분산분석의 개념

■ 분산분석(ANnalysis Of VAriance : ANOVA)

3개 이상의 집단에 대한 평균 차이를 검증하는 분석 방법

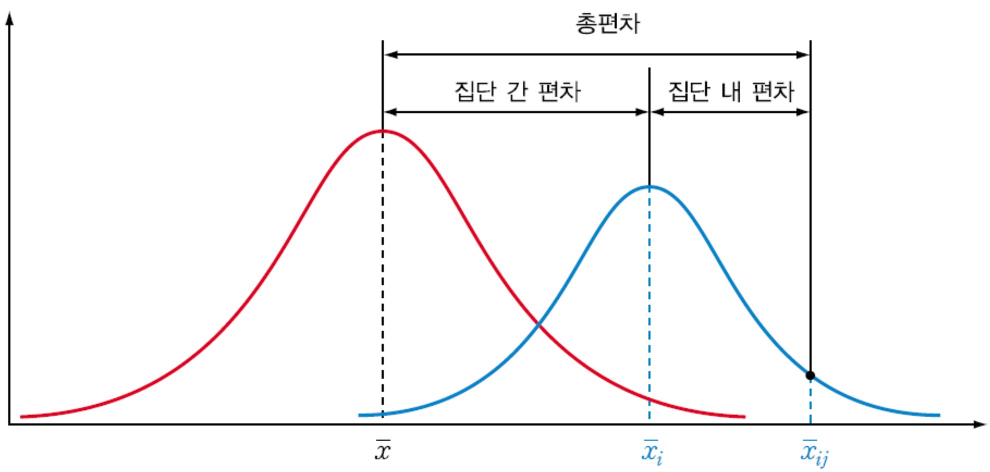
→ 특성에 대한 산포의 제곱합을 요인별 제곱합으로 분해한 후 영향 요인을 찾아냄.

→ 가설검정은 F분포를 이용

cf. t검정에서는 직접적으로 두 집단에 대한 차이를 비교했지만,
3개 이상의 집단을 직접 비교하는 방법은 상당히 복잡하므로
분산분석을 사용하는 것이 편리

분산분석의 개념

■ 분산분석에서의 편차



총편차$= \overline{x_{ij}} - \overline{x}$

집단 간 편차 $= \overline{x_i} - \overline{x}$

집단 내 편차 $= \overline{x_{ij}} - \overline{x_i}$

편의점 5개 브랜드 전체의 만족도 평균은 $\overline{x}$,
편의점 i의 만족도 평균은 $\overline{x_i}$, 편의점 i의 만족도에 대한 측정치 중 하나인 j를 $\overline{x_{ij}}$라고 하면,
총편차는 집단 간 편차와 집단 내 편차로 구성.

분산분석의 개념

■ 분산분석에서의 편차

집단 간의 분산이 크면 클수록,
집단 내의 분산이 작으면 작을수록 → 집단 간의 평균 차이가 커짐

이때 집단 간의 상대적인 비율을 확인한 것을 분산비율 F라 한다.

$$F = \frac{(\text{집단 간 변동})}{(\text{집단 내 변동})} = \frac{(\text{집단 간 평균제곱})}{(\text{집단 내 평균제곱})}$$

분산분석의 구분

■ 일원 분산분석(one-way ANOVA)

한 가지의 요인을 기준으로 집단 간의 차이를 조사하는 것

Ex. 편의점의 종류를 기준으로 고객의 만족도를 조사하는 경우

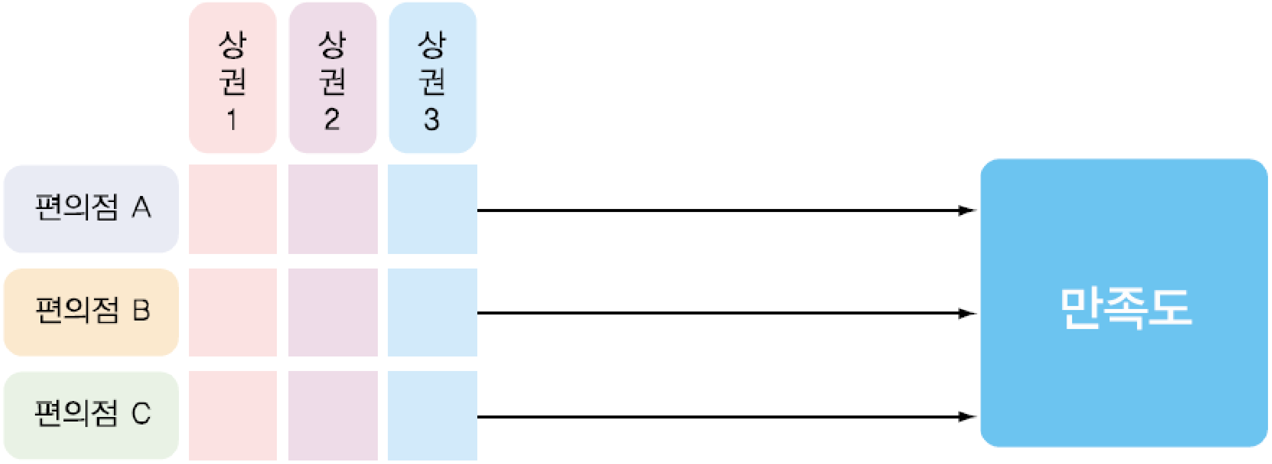
분산분석의 구분

■ 이원 분산분석(two-way ANOVA)

두 가지 요인을 기준으로 집단 간의 차이를 조사하는 것

Ex. 편의점을 종류와 위치를 기준으로 나누고,

편의점에 대한 고객의 만족도를 조사하는 경우

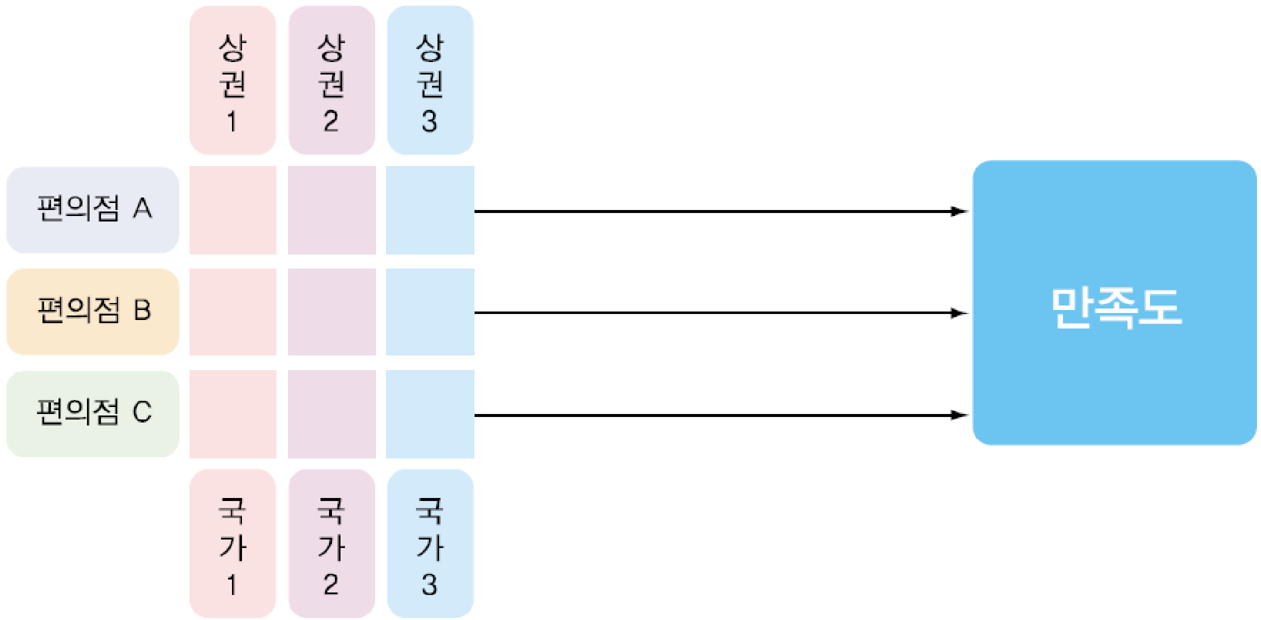


분산분석의 구분

■ 다원 분산분석(multi-way ANOVA)

세 가지 이상의 요인을 기준으로 집단 간의 차이를 조사하는 것

Ex. 편의점의 종류와 상권, 본사 국가를 기준으로
편의점에 대한 고객의 만족도를 조사하는 경우



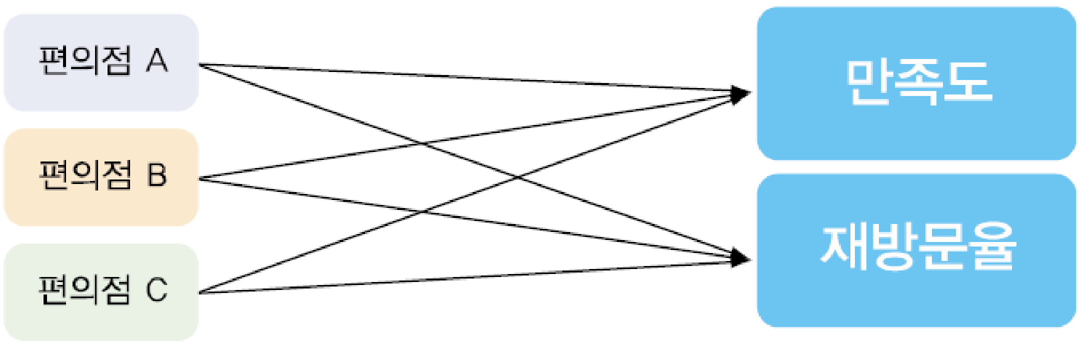
분산분석의 구분

■ 다변량 분산분석(multi-variate ANOVA)

독립변수 1개 이상에 대해 종속변수 2개 이상으로 조사하는 것

Ex. 편의점의 종류를 독립변수로 구성하고,

종속변수로 고객의 만족도와 재방문율 2개로 구성하여 조사를 하는 경우



분산분석의 구분

■ 분산분석의 구분

분산분석은 3개 이상의 집단에 대한 평균 차이를 알아보는 검정으로,
독립변수와 종속변수의 개수에 따라 다음과 같이 구분된다.

<table>
<tr><th>구분</th><th colspan="2">명칭</th><th>독립변수의 개수</th><th>종속변수의 개수</th></tr>
<tr><td rowspan="3">단일변량
분산분석</td><td>일원 분산분석</td><td>One-way ANOVA</td><td>1개</td><td rowspan="3">1개</td></tr>
<tr><td>이원 분산분석</td><td>Two-way ANOVA</td><td>2개</td></tr>
<tr><td>다원 분산분석</td><td>Multi-way ANOVA</td><td>3개 이상</td></tr>
<tr><td>다변량 분산분석</td><td>–</td><td>MANOVA</td><td>1개 이상</td><td>2개 이상</td></tr>
</table>

분산분석의 구분

■ 분산분석의 가정

- 각 모집단은 정규분포여야 하며, 집단 간 분산은 동일해야 한다.
- 각 표본들은 독립적으로 추출되어야 한다.
- 각 표본의 크기는 적절해야 한다.

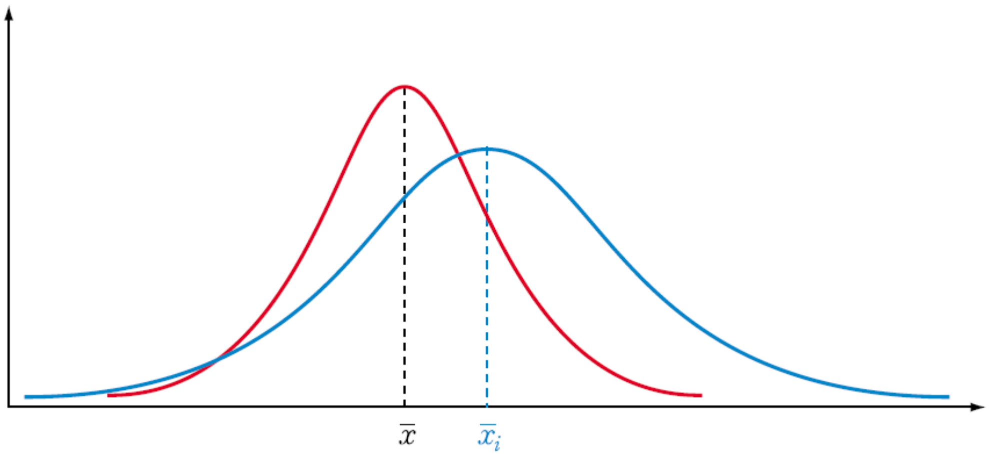
분산분석의 구분

- **각 모집단은 정규분포여야 하며, 집단 간 분산은 동일해야 한다.**

  모집단을 서로 비교하기 위해서는 각 모집단이 좌우대칭인 정규분포여야 하지만 집단 간의 평균은 서로 다를 수 있다.

  두 집단을 비교할 때 분산이 동일하지 않으면 집단 간의 평균 차이를 구별하기 쉽지 않다.

  분산분석은 집단이 3개 이상의 경우에 사용하는 분석 방법이므로 분산이 다르면 계산이 어려워진다.



[그림 9-7] 집단 간 분산이 동일하지 않은 경우

분산분석의 구분

- **각 표본들은 독립적으로 추출되어야 한다.**

  표본을 구성하는 과정에서 각각의 표본들은 모두 독립적으로 구성되어야 한다.
  → 표본을 구성하는 과정에서 어느 집단이 다른 집단에 영향을 주지 않아야 한다.

분산분석의 구분

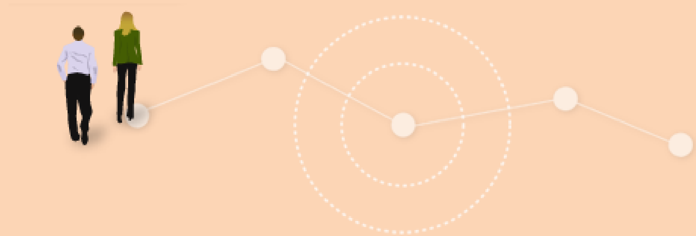
- **각 표본의 크기는 적절해야 한다.**

  분석을 진행하기 위해서는 표본의 크기가 충분해야 한다.

  이를 충족하면 분산분석을 실시할 때 표본의 개수에 상관없이 분석을 진행할 수 있다.

02 일원 분산분석

:: Keywords 총편차 | 집단 간 편차 | 집단 내 편차 | 평균제곱 | 결과 해석



일원 분산분석의 과정

■ 편의점별 만족도 측정 데이터

번호	편의점 A	편의점 B	편의점 C
1	1	4	4
2	4	4	3
3	3	3	4
4	3	4	3
5	3	4	4
6	3	5	4
7	3	4	3
8	–	4	3
9	–	–	3

각 편의점에 대한 소비자
만족도에 차이가 있는지를
α=0.05의 수준에서 알아보자.

일원 분산분석의 과정

■ 가설 수립

조사의 목적이 편의점별로 소비자의 만족도에 차이가 있는지 맞춰져 있으므로,

귀무가설 H_0는 편의점별로 소비자 만족도에 차이가 없다

대립가설 H_1은 편의점별로 소비자 만족도에 차이가 있다

편의점 A, B, C의 평균을 각각 $\overline{x_A}, \overline{x_B}, \overline{x_C}$ 라 하면

H_0 : $\overline{x_A} = \overline{x_B} = \overline{x_C}$ → 편의점별 소비자 만족도에 차이가 없다.

H_1: 편의점별 소비자 만족도에 차이가 있다.

일원 분산분석의 과정

■ 자료의 구성

기준으로 각 집단들의 평균을 $\overline{x_A}, \overline{x_B}, \overline{x_C}$, 전체 평균을 $\overline{x}$로 표시한 후,

각각의 평균을 도출하면

번호	편의점 A	편의점 B	편의점 C	전체
1	1	4	4	
2	4	4	3	
3	3	3	4	
4	3	4	3	
5	3	4	4	
6	3	5	4	
7	3	4	3	
8	–	4	3	
9	–	–	3	
표기	$\overline{x_A}$	$\overline{x_B}$	$\overline{x_C}$	$\overline{x}$
평균	2.857	4.000	3.444	3.434

일원 분산분석의 편차

■ 총편차(Sum of Squares Total : SST)

관측된 자료에 대한 전체 평균과의 차이의 제곱합

$$SST = \sum\sum(x_{ij} - \overline{x})^2$$

24개 측정치에 대해 총편차를 구하면

$$SST = (1-3.434)^2 + (4-3.434)^2 + (3-3.434)^2 + \cdots + (3-3.434)^2 = 13.973$$

일원 분산분석의 편차

■ 집단 간 편차(Sum of Squares Between samples : SSB)

해당 집단에서 측정된 평균에 대한 전체 평균 차이의 제곱합

→ 집단 간 편차를 확인하는 이유는

해당 집단의 대표값인 평균을 이용하여 전체의 평균과 비교하기 위해서

$$SSB = \sum n_i (\overline{x_j} - \overline{x})^2$$

편의점 A, B, C에 대한 집단 간 편차를 구하면

$$SSB = 7(2.857 - 3.434)^2 + 8(4.000 - 3.434)^2 + 9(3.444 - 3.434)^2 = 4.894$$

일원 분산분석의 편차

■ 집단 내 편차(Sum of Squares Within samples : SSW)

집단 내에서 측정된 자료에 대한 집단 평균 차이의 제곱합

→ 집단 내 편차를 확인하는 이유는

동일한 집단에서 발생한 측정치의 편차를 측정해 통제가 가능하지 않은 외부 변수를 고려하기 위한 것

$$SSW = \sum\sum(x_{ij} - \overline{x_i})^2$$

일원 분산분석의 편차

■ 집단 내 편차(Sum of Squares Within samples : SSW)

편의점 A, B, C 에 대한 각각의 집단 내 편차를 구하면

$$SSW_A = (1 - 2.857)^2 + (4 - 2.857)^2 + \cdots + (3 - 2.857)^2 = 4.857$$

$$SSW_B = (4 - 4.000)^2 + (4 - 4.000)^2 + \cdots + (4 - 4.000)^2 = 2.000$$

$$SSW_C = (4 - 3.444)^2 + (3 - 3.444)^2 + \cdots + (3 - 3.444)^2 = 2.222$$

최종적으로 계산된 집단 내 편차는 $SSW_A + SSW_B + SSW_C = 9.079$

일원 분산분석의 편차

(총편차)=(집단 간 편차) + (집단 내 편차)

$$SST = SSB + SSW$$

$$13.973 = 4.894 + 9.079$$

일원 분산분석의 결과 해석

■ 어떤 값을 비교할 때 가장 기본이 되는 값은 평균

→ 편차를 자유도로 나누면 편차의 평균이 되는데, 각 편차는 각기 다른 자유도를 가진다.

전체 자료 n개로 부터 SST는 자유도가 $(n-1)$,
표본이 i개일 때 SSB의 자유도는 $(i-1)$,
SSW의 자유도는 $(n-i)$

결국 집단 간/내의 편차(제곱합)을 자유도로 나누면 분산이 되므로, 이를 분산분석이라 한다.

일원 분산분석의 결과 해석

집단 간 평균제곱($Mean\ Squares\ Between\ samples$: MBS)은 $MSB = \dfrac{SSB}{i-1}$

집단 내 평균제곱($Mean\ Squares\ Within\ samples$: MSW)은 $MSW = \dfrac{SSW}{n-i}$

일원 분산분석의 결과 해석

■ 분산비율 F

집단 간 평균제곱과 집단 내 평균제곱을 구했다면,

이들 두 평균제곱의 비율을 확인해야 하는데, 이를 분산비율 F라 한다.

$$F = \frac{(\text{집단 간 변동})}{(\text{집단 내 변동})} = \frac{(\text{집단 간 평균제곱})}{(\text{집단 내 평균제곱})} = \frac{MSB}{MSW} = \frac{\frac{SSB}{i-1}}{\frac{SSW}{n-i}}$$

일원 분산분석의 결과 해석

■ 분산비율 F

분산비율을 구하기 위해 먼저 MSB와 MSW를 구하면

$$MSB = \frac{4.894}{2} = 2.447, \quad MSW = \frac{9.079}{21} = 0.432$$

따라서 표본비율 $F = \dfrac{2.447}{0.432} = 5.664$

일원 분산분석의 결과 해석

■ 일원 분산분석표

구분	편차(제곱합)	자유도	평균제곱	분산비율 F
집단 간 집단 내	$SSB = 4.894$ $SSW = 9.079$	$i-1=2$ $n-i=21$	$MSB = 2.447$ $MSW = 0.432$	$\frac{MSB}{MSW} = \frac{2.447}{0.432} = 5.664$
합계	$SST = 13.973$	$n-1=23$	−	−

분산비율이 5.664

이는 만족도가 내부적으로 느껴지는 것보다 편의점이라는 집단으로

보았을 때 약 5.664배의 차이가 난다는 의미

일원 분산분석의 결과 해석

Note **집단 내의 자유도가 $(n-i)$인 이유**

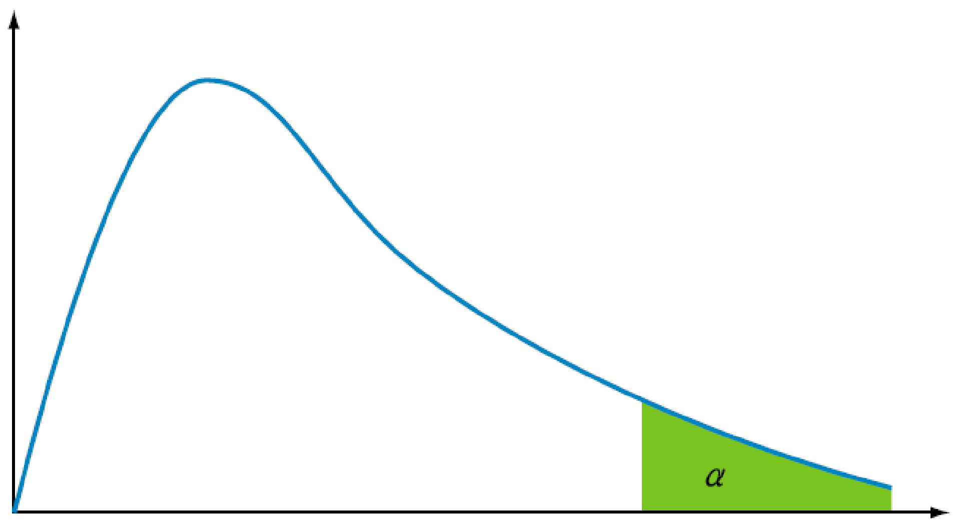
[표 9-3]의 편의점 A, B, C는 각각 7개, 8개, 9개의 표본으로 구성되어 있으므로, 각각에 대한 자유도를 구하면 편의점 A는 $7-1=6$(개), 편의점 B는 $8-1=7$(개), 편의점 C는 $9-1=8$(개)이다. 그러므로 집단 3개의 자유도를 모두 합하면 $(7-1)+(8-1)+(9-1)=(7+8+9)+(-3)=21$이며, 이는 전체 표본의 개수 24개에서 표본의 개수 3개를 뺀 21과 같음을 알 수 있다.

일원 분산분석의 결과 해석

■ 가설 채택

5.664라는 값은 표본을 모수를 통해 추정하는 것이므로,

이에 대한 확률적인 근거를 통해 H_0 혹은 H_1 중 어느 것을 채택할 지 판단해야 한다.



[그림 9-8] F분포도

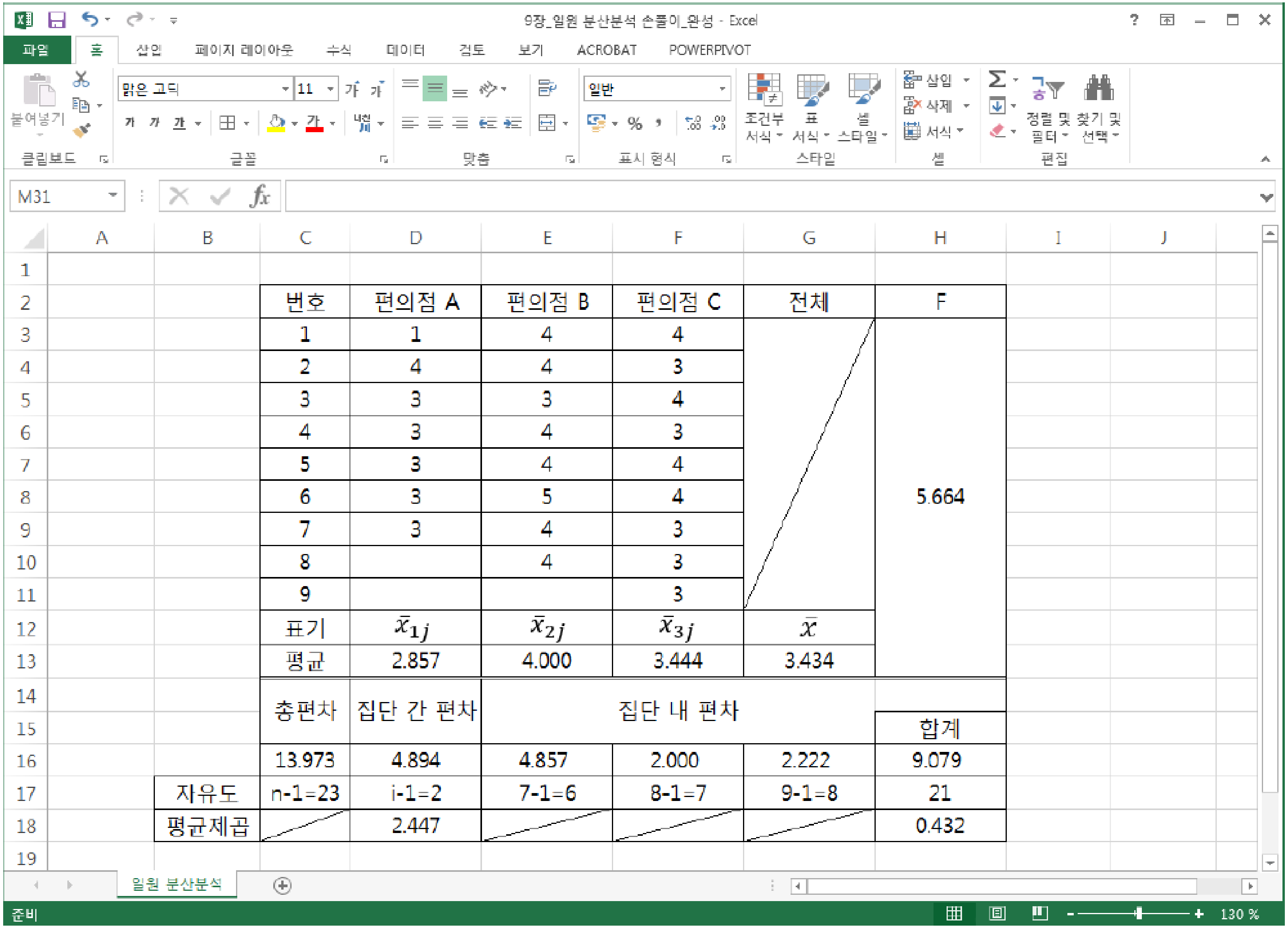
일원 분산분석의 결과 해석

[표 9-5] F분포표($\alpha = 0.05$)

분자 / 분모	1	2	3	4	5	6	7	8	9	10
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22

분자의 자유도가 2. 분모의 자유도는 21에 해당하는 F분포표의 값은 3.47
3.47은 5%에서의 분산비율이 3.47이라는 것을 의미
5.664는 3.47보다 크기 때문에, 5.664가 나올 확률은 5%보다 더 작아진다.
그러므로 귀무가설을 기각하고 대립가설을 채택한다.

일원 분산분석의 결과 해석 (Excel 계산 결과)



	번호	편의점 A	편의점 B	편의점 C	전체	F
	1	1	4	4		5.664
	2	4	4	3		
	3	3	3	4		
	4	3	4	3		
	5	3	4	4		
	6	3	5	4		
	7	3	4	3		
	8		4	3		
	9			3		
	표기	$\bar{x}_{1j}$	$\bar{x}_{2j}$	$\bar{x}_{3j}$	$\bar{x}$	
	평균	2.857	4.000	3.444	3.434	
	총편차	집단 간 편차	집단 내 편차			합계
	13.973	4.894	4.857	2.000	2.222	9.079
자유도	n-1=23	i-1=2	7-1=6	8-1=7	9-1=8	21
평균제곱		2.447				0.432

일원 분산분석의 결과 해석

참고 F분포표의 값보다 큰 값의 분산비율에 대한 확률이 작아지는 이유는?

[표 9–4]의 분산비율을 비교하면서 편의점의 만족도 차이가 5.664배라고 했다. 하지만 F분포표에서는 3.47배의 차이가 나야 5%의 확률이라 했다. H_0는 편의점별로 소비자의 만족도 차이가 없다는 것인데, 5%의 확률에서 3.47보다 훨씬 큰 값인 5.664로 조사되었다는 것은 만족도 차이에 대한 유의수준의 F값이 5%의 확률보다 더 적게 발생하리라는 것을 의미하므로 H_0를 기각하고 H_1를 채택해야 한다.

일원 분산분석

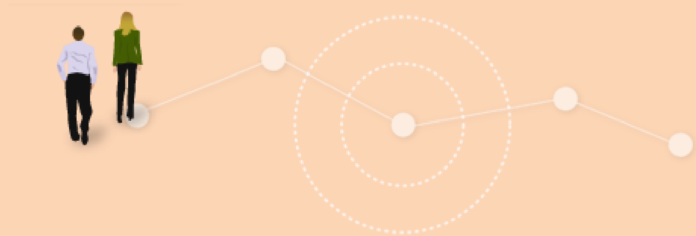
예제 9-1 일원 분산분석

준비파일 | 9장_일원 분산분석.xlsx

[표 9-3]에 대한 일원 분산분석을 Excel을 이용하여 확인하라.

03 이원 분산분석

:: Keywords 상호작용 효과 | 결과 해석



이원 분산분석

편의점별 상권의 차이가 이용자의 만족도에 영향을 미칠까?

[그림 9-14] **표본의 상권별 구분**

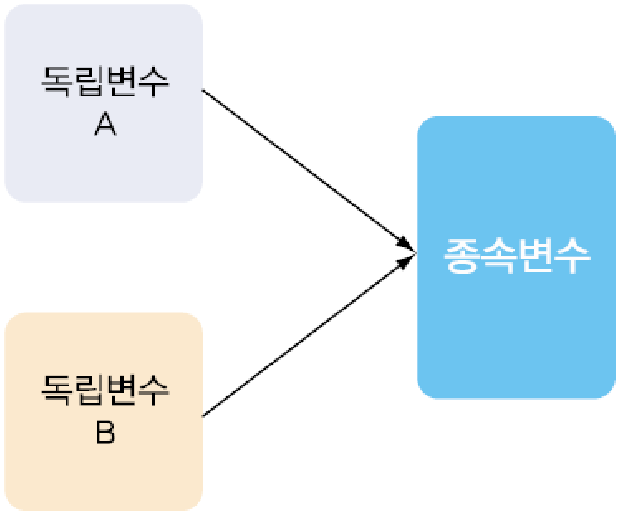
국내의 편의점 3개를 표본으로 하여, 이들의 만족도가 상권별로 차이가 있는지 확인할 수 있을까?

앞서 살펴본 편의점 브랜드에 따른 만족도 차이에 대한 조사와는 달리, 이 문제는 브랜드와 상권이라는 2개의 독립변수로 구분하여 만족도의 차이를 조사해야 한다.

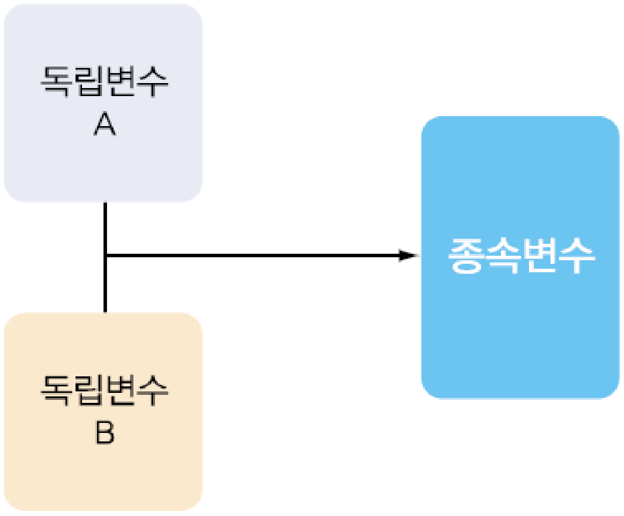
상호작용 효과

- **상호작용 효과(interaction effect)**

상호작용 효과는 두 개의 독립변수가 동시에 작용하여 종속변수에 미치게 되는 영향을 말한다.



(a) 주효과



(b) 상호작용 효과

이원 분산분석의 과정

원인이 되는 변수(독립변수)는 3개의 편의점
3개의 상권
결과가 되는 변수(종속변수)는 소비자의 만족도

두 독립변수에 대한 소비자 만족도에 차이가 있는지를 $\alpha = 0.05$의 수준에서 알아보자

[표 9-6] 편의점별-상권별 만족도 측정 데이터

구분	편의점 A	편의점 B	편의점 C
강남	1	4	4
	4	4	3
	1	3	4
홍대	2	3	4
	2	2	3
	3	3	3
종로	2	4	2
	3	2	4
	2	3	4

이원 분산분석의 과정

■ 가설 수립

조사의 목적이 편의점별, 상권별 두 독립변수에 대해 소비자 만족도에 차이가 있는지에 맞춰져 있으므로 가설은 기본적으로 2개이고 상호작용을 생각하면 총 3개의 가설을 수립할 수 있다.

$\overline{x}$: 전체 평균

$\overline{x_i}$: 첫 번째 독립변수(편의점명)에 해당하는 관측값 k개의 평균

$\overline{x_j}$: 두 번째 독립변수(편의점 위치)에 해당하는 관측값 k개의 평균

$\overline{x_{ij}}$: 첫 번째 독립변수 i과 두 번째 독립변수 j의 관측값 k개의 평균

$\overline{x_{ijk}}$: 첫 번째 독립변수 i번째와 두 번째 독립변수 j번째에 속하는 k번째 관측값

이와 같이 정의하였다면

이원 분산분석의 과정

가설을 다음과 같이 나타낼 수 있다.

- 편의점별
  - $H_0^i : \overline{x_{Ai}} = \overline{x_{Bi}} = \overline{x_{Ci}} \Rightarrow$ 편의점별 소비자의 만족도에 차이가 없다.
  - H_1^i : 편의점별 소비자의 만족도에 차이가 있다.
- 상권별
  - $H_0^j : \overline{x_{Aj}} = \overline{x_{Bj}} = \overline{x_{Cj}} \Rightarrow$ 상권별 소비자의 만족도에 차이가 없다.
  - H_1^j : 상권별 소비자 만족도에 차이가 있다.
- 상호작용
  - $H_0^{ij} : \overline{x_{Aij}} = \overline{x_{Bij}} = \overline{x_{Cij}} \Rightarrow$ 상호작용에 의한 소비자의 만족도에 차이가 없다.
  - H_1^{ij} : 상호작용에 의한 소비자의 만족도에 차이가 있다.

이원 분산분석의 과정

■ 자료 구성

[표 9-7] 이원 분산분석의 자료 구성

구분	편의점 A	편의점 B	편의점 C	평균
강남	1	4	4	$\overline{x_{i1k}} = 3.111$
	4	4	3	
	1	3	4	
$\overline{x_{ij}}$	$\overline{x_{11k}} = 2.000$	$\overline{x_{21k}} = 3.667$	$\overline{x_{31k}} = 3.667$	–
홍대	2	3	4	$\overline{x_{i2k}} = 2.778$
	2	2	3	
	3	3	3	
$\overline{x_{ij}}$	$\overline{x_{12k}} = 2.333$	$\overline{x_{22k}} = 2.667$	$\overline{x_{32k}} = 3.333$	–
종로	2	4	2	$\overline{x_{i3k}} = 2.889$
	3	2	4	
	2	3	4	
$\overline{x_{ij}}$	$\overline{x_{13k}} = 2.333$	$\overline{x_{23k}} = 3.000$	$\overline{x_{33k}} = 3.333$	–
평균	$\overline{x_{1jk}} = 2.222$	$\overline{x_{2jk}} = 3.111$	$\overline{x_{3jk}} = 3.444$	$\overline{x} = 2.926$

이원 분산분석의 과정

이원 분산분석에서도 평균 차이를 확인하기 위하여 총편차를 확인하면
이원 분산분석의 편차들 사이에는 다음이 성립한다.

$$\begin{array}{ccccccccc} (\text{총편차}) & = & (\text{독립변수 } i \text{ 의 편차}) & + & (\text{독립변수 } j \text{ 의 편차}) & + & (i \text{ 와 } j \text{ 의 상호작용}) & + & (\text{집단 내 편차}) \\ SST & = & SSB_i & + & SSB_j & + & SSB_{ij} & + & SSW \end{array}$$

각각의 편차는

$$SST = \sum\sum\sum(\overline{x_{ijk}} - \overline{x})^2$$

$$SSB_i = \sum k_i(\overline{x_{ik}} - \overline{x})^2$$

$$SSB_j = \sum k_j(\overline{x_{jk}} - \overline{x})^2$$

$$SSB_{ij} = \sum k(\overline{x_{ij}} - \overline{x_i} - \overline{x_j} + \overline{x})^2$$

$$SSW = \sum\sum\sum(x_{ijk} - \overline{x})$$

자유도는 $(i-1)+(j-1)+(i-1)(j-1)+ij(k-1) = i \cdot j \cdot k - 1$

이원 분산분석의 과정

편차를 구하면

$$SST = (1-2.926)^2 + (4-2.926)^2 + \cdots + (4-2.926)^2 = 23.852$$

$$SSB_i = 9(2.222-2.926)^2 + 9(3.111-2.926)^2 + 9(3.444-2.926)^2 = 7.183$$

$$SSB_j = 9(3.111-2.926)^2 + 9(2.778-2.926)^2 + 9(2.889-2.926)^2 = 0.517$$

$$SSB_{ij} = 3(2.000-2.222-3.111+2.926)^2 + \cdots + 3(3.333-3.444-2.889+2.926)^2$$
$$= 1.481$$

$$SSW = SST - SSB_i - SSB_j - SSB_{ij} = 14.671$$

이원 분산분석의 결과 해석

각 집단 간의 평균을 비교하기 위해 집단 간 평균제곱 MSB와
집단 내 평균제곱 MSW를 구해야 한다.

$$MSB_i = \frac{SSB_i}{i-1}\ ,\quad MSB_j = \frac{SSB_j}{j-1}$$

$$MSB_{ij} = \frac{SSB_{ij}}{(i-1)(j-1)}\ ,\quad MSW = \frac{SSW}{i \cdot j \cdot (k-1)}$$

이원 분산분석의 결과 해석

평균제곱을 구하면

$$MSB_i = \frac{7.183}{3-1} = 3.592\ , \quad MSB_j = \frac{0.517}{3-1} = 0.259$$

$$MSB_{ij} = \frac{1.481}{(3-1)(3-1)} = 0.370\ , \quad MSW = \frac{14.671}{3 \cdot 3 \cdot 2} = 0.815$$

구분	편차(제곱합)	자유도	평균제곱	분산비율 F
독립변수 i	$SSB_i = 7.183$	$i-1=2$	$MSB_i = 3.592$	$MSB_i / MSW = 4.407$
독립변수 j	$SSB_j = 0.517$	$j-1=2$	$MSB_j = 0.259$	$MSB_j / MSW = 0.318$
ij 상호작용	$SSB_{ij} = 1.481$	$(i-1)(j-1)=4$	$MSB_{ij} = 0.370$	$MSB_{ij} / MSW = 0.454$
집단 내	$SSW = 14.671$	$i \cdot j \cdot (k-1) = 18$	$MSW = 0.815$	
합계	$SST = 23.852$	$n-1=26$	–	–

이원 분산분석의 결과 해석

■ 가설 채택

$MSB_i/MSW = 4.407, MSB_j\ /\ MSW = 0.318, MSBij/MSW = 0.454$라는 값은

표본은 통해 모수를 추정하는 것이므로,

확률적인 근거를 통해 H_0 혹은 H_1 중 어느 것을 채택해야 할지에 대해

판단해야 한다.

MSB_i, MSB_i, MSB_{ii}의 자유도인 2,2,4를 기준으로

F분포표의 값은 $3.55, 3.55, 2.93$

이원 분산분석의 결과 해석

■ 가설 채택

MSB_i, MSB_i, MSB_{ii}의 자유도인 2,2,4를 기준으로

F분포표의 값은 $3.55, 3.55, 2.93$

$MSB_i \,/\, MSB = 4.407$이므로

귀무가설 H_0^i를 기각하고, 대립가설 H_1^i을 채택한다.

$H_0^i : \overline{x}_{Ai} = \overline{x}_{Bi} = \overline{x}_{Ci} \Longrightarrow$ 편의점별 소비자 만족도에 차이가 없다.

H_1^i : 편의점별 소비자 만족도에 차이가 있다.

이원 분산분석의 결과 해석

■ 가설 채택

MSB_i, MSB_i, MSB_{ii}의 자유도인 2,2,4를 기준으로
F분포표의 값은 3.55, 3.55, 2.93

MSB_j / MSW = 0.318이므로
귀무가설 H_0^i를 기각하지 못한다.

$H_0^j : \overline{x_{Aj}} = \overline{x_{Bj}} = \overline{x_{Cj}} \implies$ 상권별 소비자 만족도에 차이가 없다.

$H_1^j :$ 상권별 소비자 만족도에 차이가 있다.

이원 분산분석의 결과 해석

■ 가설 채택

MSB_i, MSB_i, MSB_{ii}의 자유도인 2,2,4를 기준으로

F분포표의 값은 $3.55, 3.55, 2.93$

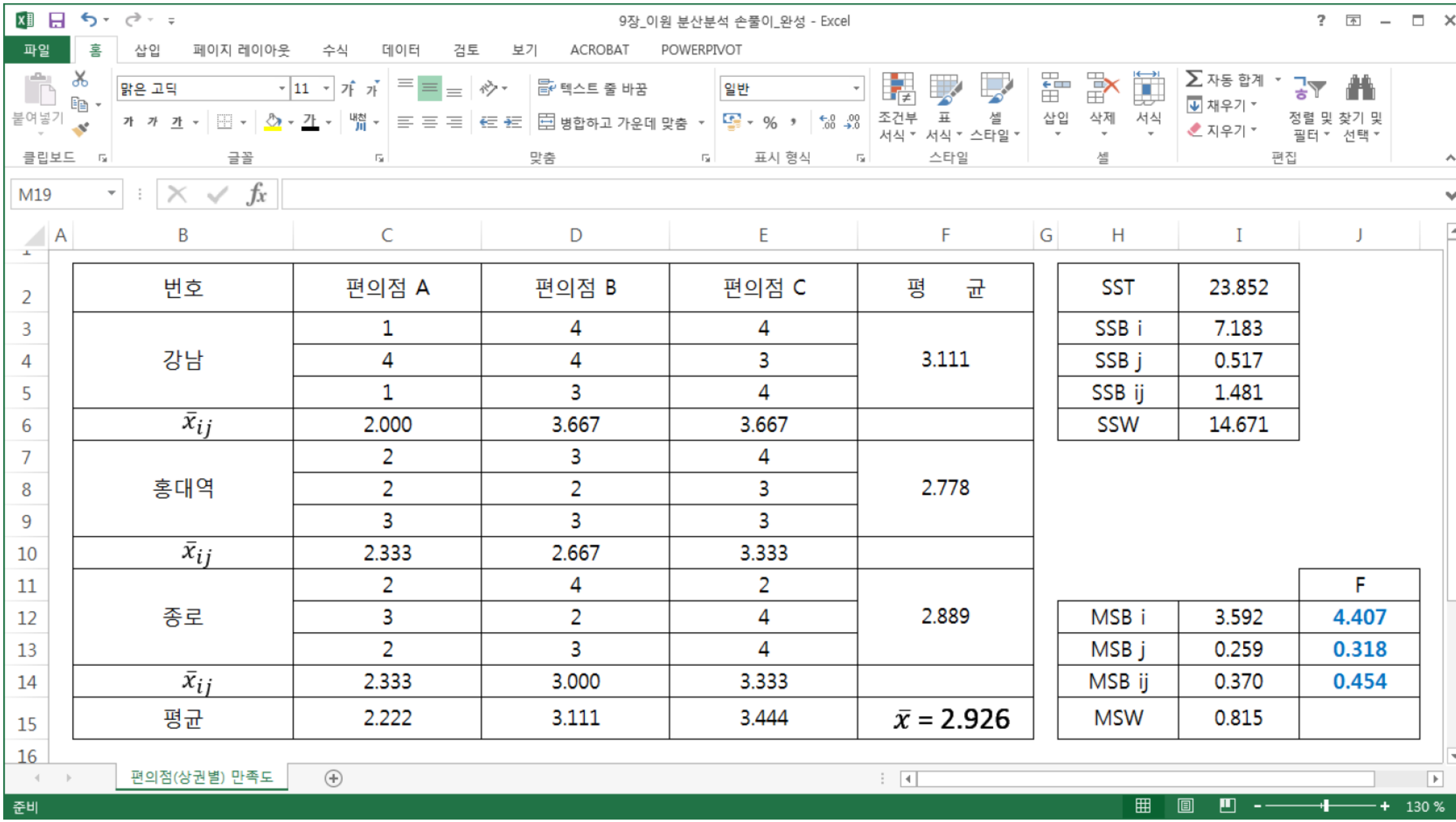
MSB_{ij} / MSW = 0.454이므로

귀무가설 H_0^i를 기각하지 못한다.

H_0^{ij} : $\overline{x}_{Aij} = \overline{x}_{Bij} = \overline{x}_{Cij}$ $\Longrightarrow$ 상호작용에 의한 소비자 만족도에 차이가 없다.

H_1^{ij} : 상호작용에 의한 소비자 만족도에 차이가 있다.

이원 분산분석의 결과 해석(Excel 계산 결과)



번호	편의점 A	편의점 B	편의점 C	평 균
강남	1	4	4	3.111
	4	4	3	
	1	3	4	
$\bar{x}_{ij}$	2.000	3.667	3.667	
홍대역	2	3	4	2.778
	2	2	3	
	3	3	3	
$\bar{x}_{ij}$	2.333	2.667	3.333	
종로	2	4	2	2.889
	3	2	4	
	2	3	4	
$\bar{x}_{ij}$	2.333	3.000	3.333	
평균	2.222	3.111	3.444	$\bar{x}$ = 2.926

SST	23.852
SSB i	7.183
SSB j	0.517
SSB ij	1.481
SSW	14.671

		F
MSB i	3.592	4.407
MSB j	0.259	0.318
MSB ij	0.370	0.454
MSW	0.815	

이원 분산분석

예제 9-2 이원 분산분석

준비파일 | 9장_이원 분산분석.xlsx

[표 9-6]에 대한 이원 분산분석을 Excel을 이용하여 확인하라.