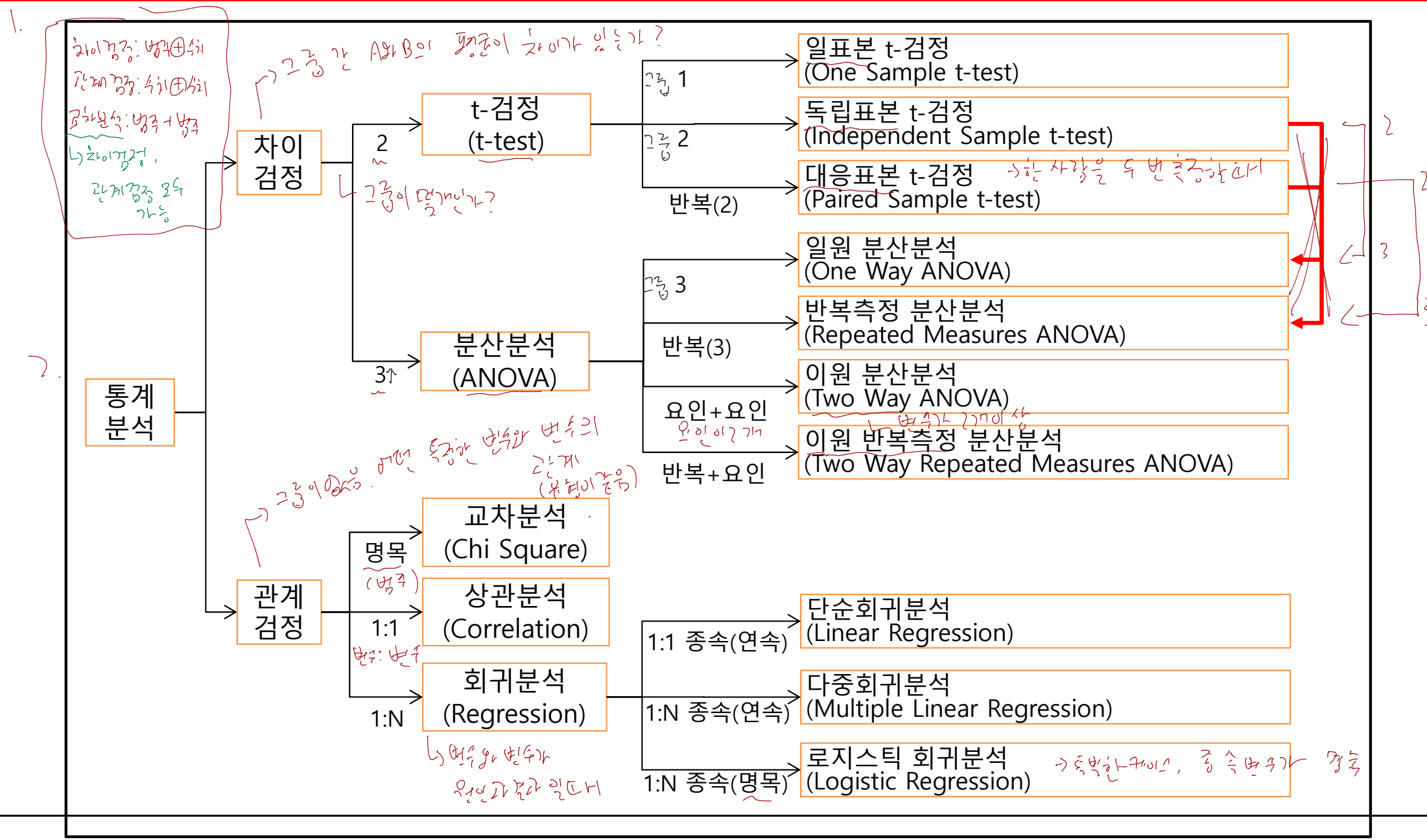


1. 통계분석방법

통계분석방법

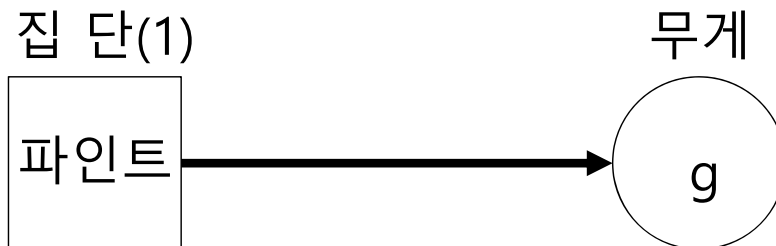


One Sample t-test

■ 문제의 정의

- B아이스크림회사에서 판매하는 아이스크림 중 파인트의 무게는 320g이다.
- 그러나 K대학 앞에 있는 점포에서 파는 아이스크림의 무게가 320g이 아니라는 소비자들의 불만이 있었다.
- 이에 따라 소비자단체에서는 B아이스크림회사에서 만든 아이스크림이 320g인지를 검사하고자 한다.

□ : 평균 (평균)
○ : 표본



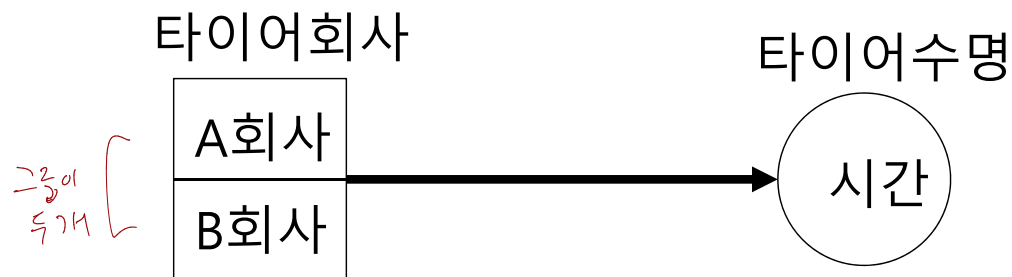
일표본 t-검정 (One Sample t-test)

weight
319
242
291
276
348
299
293
304
296
286
301
324
279
323

Independent Sample t-test

■ 문제의 정의

- 이교수는 이번에 자동차 타이어를 교체하려고 하는데 수명이 긴 타이어로 교체하려고 한다.
- 시중에는 A회사의 타이어와 B회사의 타이어가 있는데, 이 교수는 이 중에서 어느 타이어를 골라야 하는가?



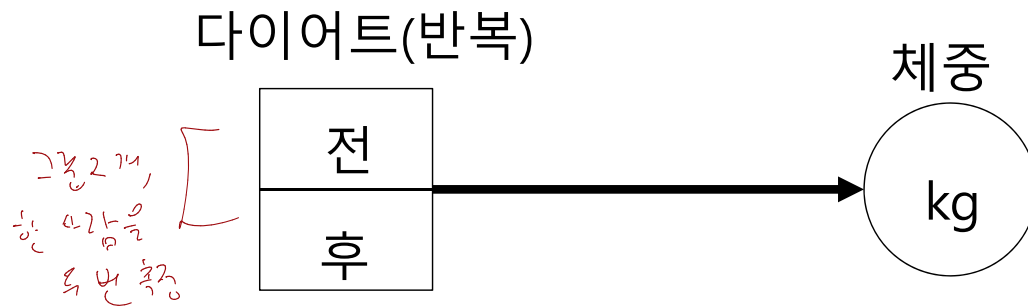
독립표본 t-검정(Independent Sample t-test)

타이어회사	타이어수명
1	48,187
1	47,245
1	51,020
1	50,732
1	52,416
1	49,278
1	38,214
1	46,742
1	48,706
1	54,280
1	50,635
1	51,052
1	51,568
1	51,569

Paired Sample t-test

■ 문제의 정의

- K제약회사의 신제품 개발부서에서는 3개월 안에 살이 빠지는 다이어트 약을 개발하였다.
- 회사 경영진에게 새롭게 개발한 다이어트약이 효과가 있는지를 보고하기 위하여 약의 효능을 검증하였다.
- 약을 먹기 전의 체중과 약을 먹은 후 3개월 후의 체중을 조사하였다.
- 과연 새로운 약은 다이어트에 효과가 있는가?



대응표본 t-검정 (Paired Sample t-test)

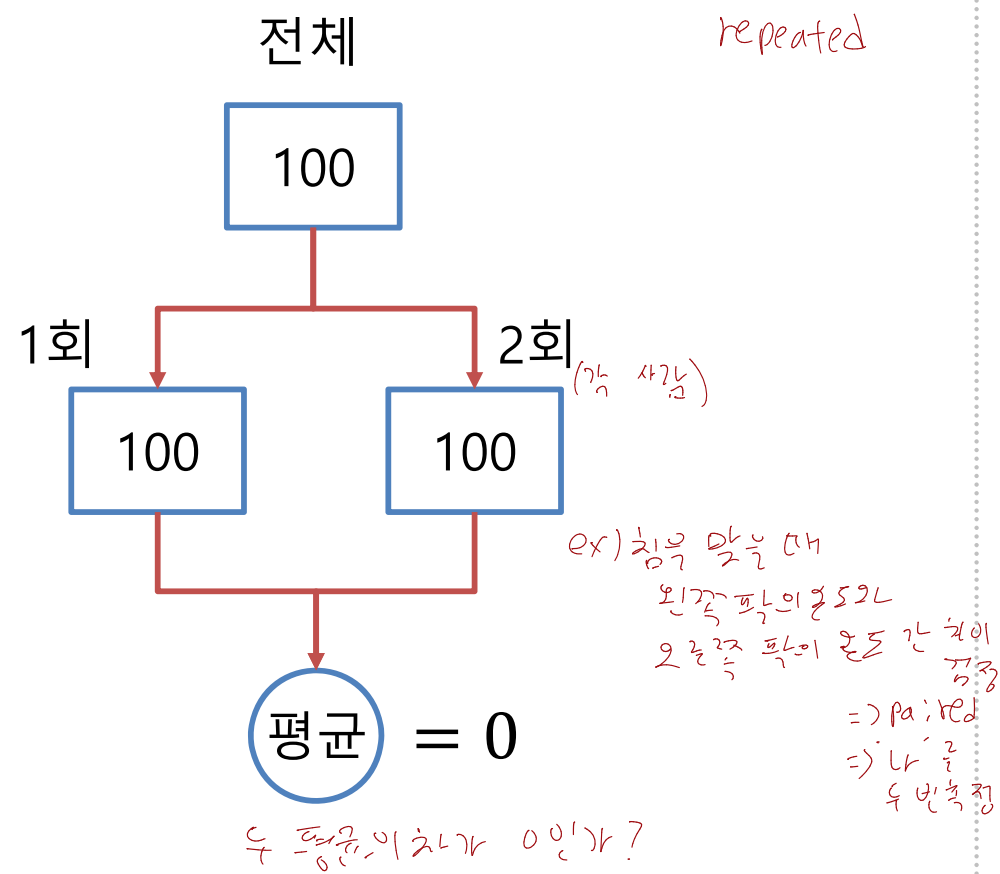
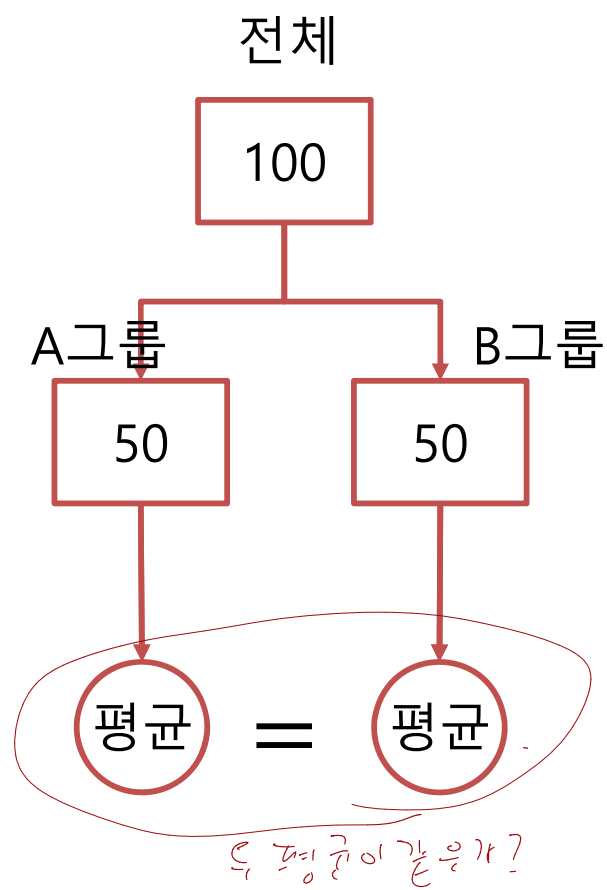
섭취전	섭취후
82	75
54	50
74	74
75	71
71	69
76	73
70	68
62	62
77	68
75	72
72	70
83	77
78	71
74	74

Paired Sample t-test

- 독립표본과 대응표본의 차이점 (시각의 의미라고 부르기 애매함 (시각의 관측해보라는 '나'를 몇번 관측가?))

- 독립표본 : 대상에서 1번만 측정
- 대응표본 : 동일 대상에서 반복해서 측정

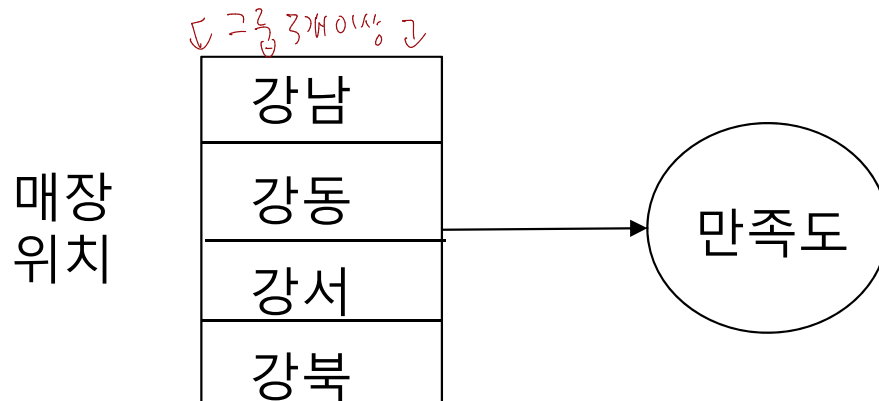
반복해서 측정할 때: Paired or repeated



One-Way ANOVA

■ 문제의 정의

- 별다방을 프랜차이즈 운영하는 K회사는 강남, 강동, 강서, 강북에 위치한 매장의 서비스에 대한 고객조사를 실시하였다.
- 과연 4곳 매장의 고객만족도는 차이가 있는지? 있다면 어느 레스토랑의 서비스 만족도가 가장 안 좋은가?



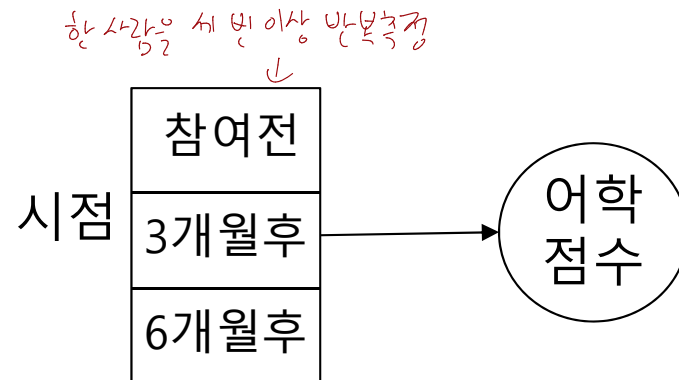
일원 분산분석(One Way ANOVA)

구분	점수
1	4
2	3
3	2
4	4
1	5
2	4
3	1
4	1
1	5
2	5
3	5
4	5
1	5
2	4

Repeated Measures ANOVA

■ 문제의 정의

- K대학에서 운영하는 6개월 어학 프로그램의 효과를 측정하려고 한다.
- 프로그램 참여전, 3개월 후, 6개월 후에 영어실력을 테스트하였다.
- 과련 어학프로그램은 효과가 있는가? 있다면 언제부터 효과가 나타났는가?



예) 오른쪽, 왼쪽, 오른쪽 반복

왼쪽 반복 한 번씩

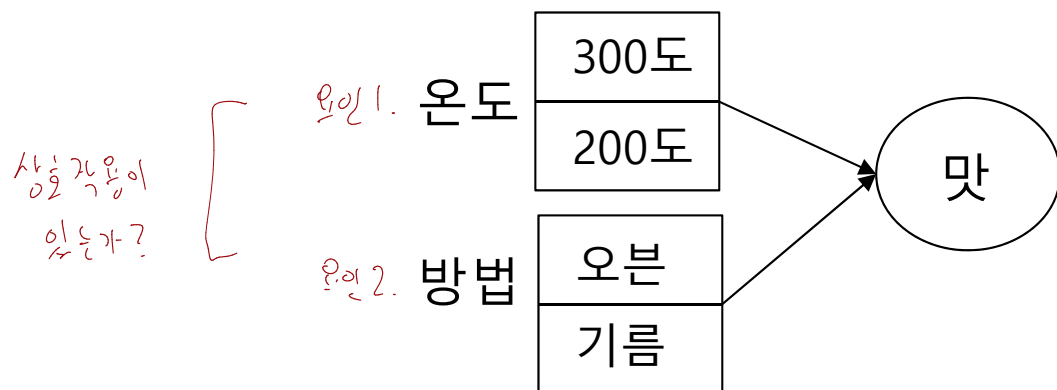
반복측정 분산분석 (Repeated Measures ANOVA)

영어1	영어3	영어6
10.00	16.00	25.00
11.00	17.00	26.00
9.00	15.00	21.00
9.00	14.00	20.00
9.00	20.00	22.00
11.00	16.00	26.00
9.00	21.00	49.00
9.00	9.00	16.00
9.00	12.00	29.00
13.00	22.00	40.00
13.00	22.00	32.00
23.00	47.00	49.00
9.00	18.00	28.00

Two-Way ANOVA

■ 문제의 정의

- 치킨의 맛을 결정하는 두 가지 요인은 튀길 때의 기름 온도와 튀기는 방법이다.
- 튀길 때의 온도를 200도와 300도로 하고,
- 튀기는 방법을 오븐과 기름으로 하여
- 치킨을 튀긴 후에 사람들에게 맛을 평가하도록 하였다.
- 과연 온도와 방법이 맛을 결정하는데 중요한 요인인가? 이 두 가지 요인들 간의 상호작용효과는 없었는가?



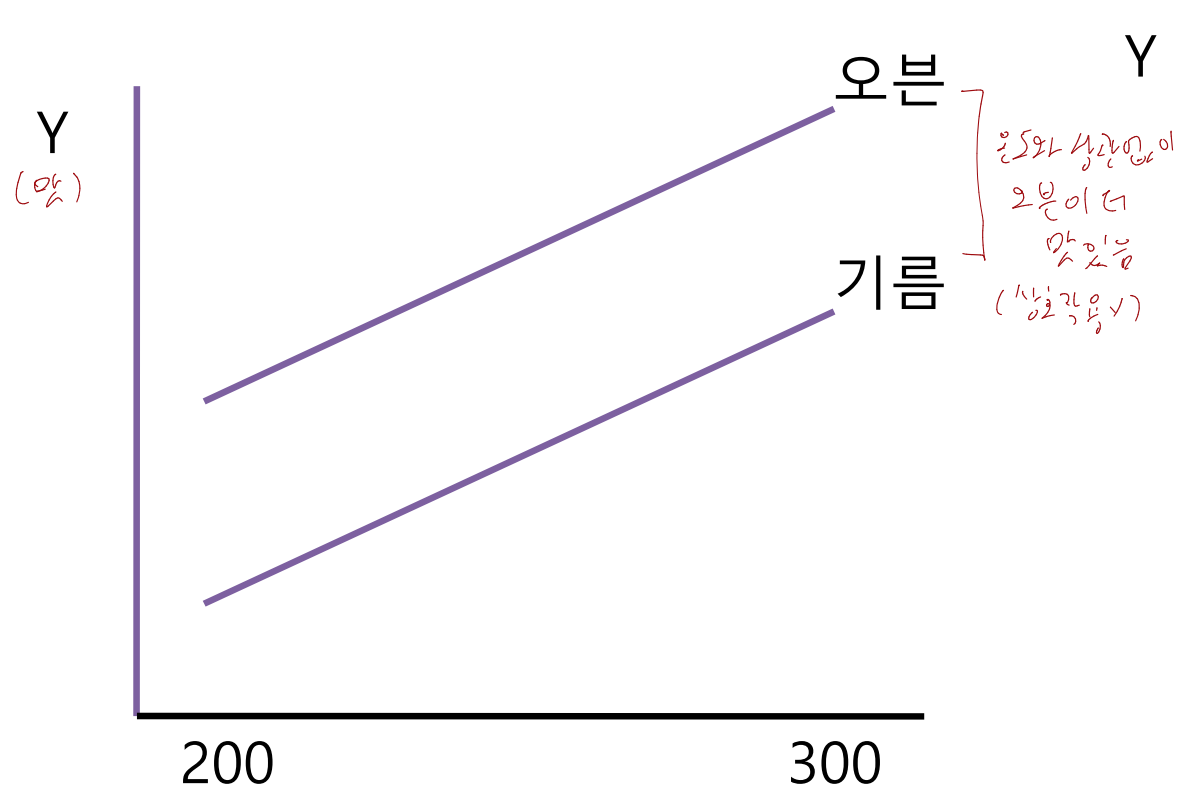
이원 분산분석 (Two Way ANOVA)

온도	방법	통닭맛
2	1	95
2	1	93
2	1	94
2	1	98
2	1	97
2	1	94
2	1	96
2	1	92
2	1	91
2	1	94
2	1	95
2	1	96
2	1	94
2	1	95

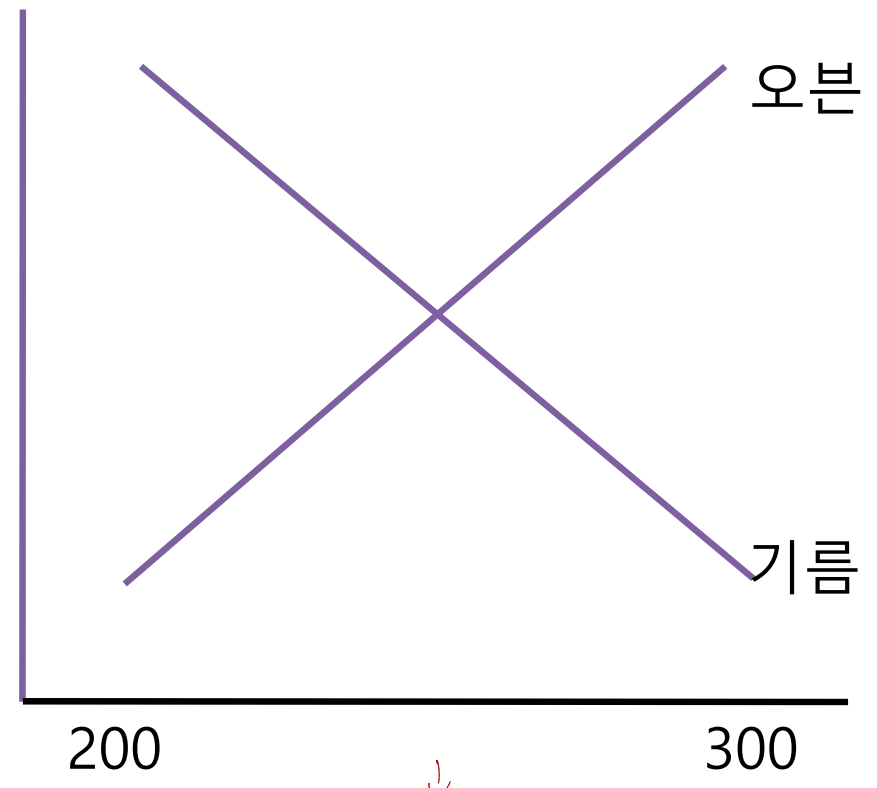
Two-Way ANOVA

■ 평균반응 프로파일(average response profile)

■ 상호작용 없는 경우

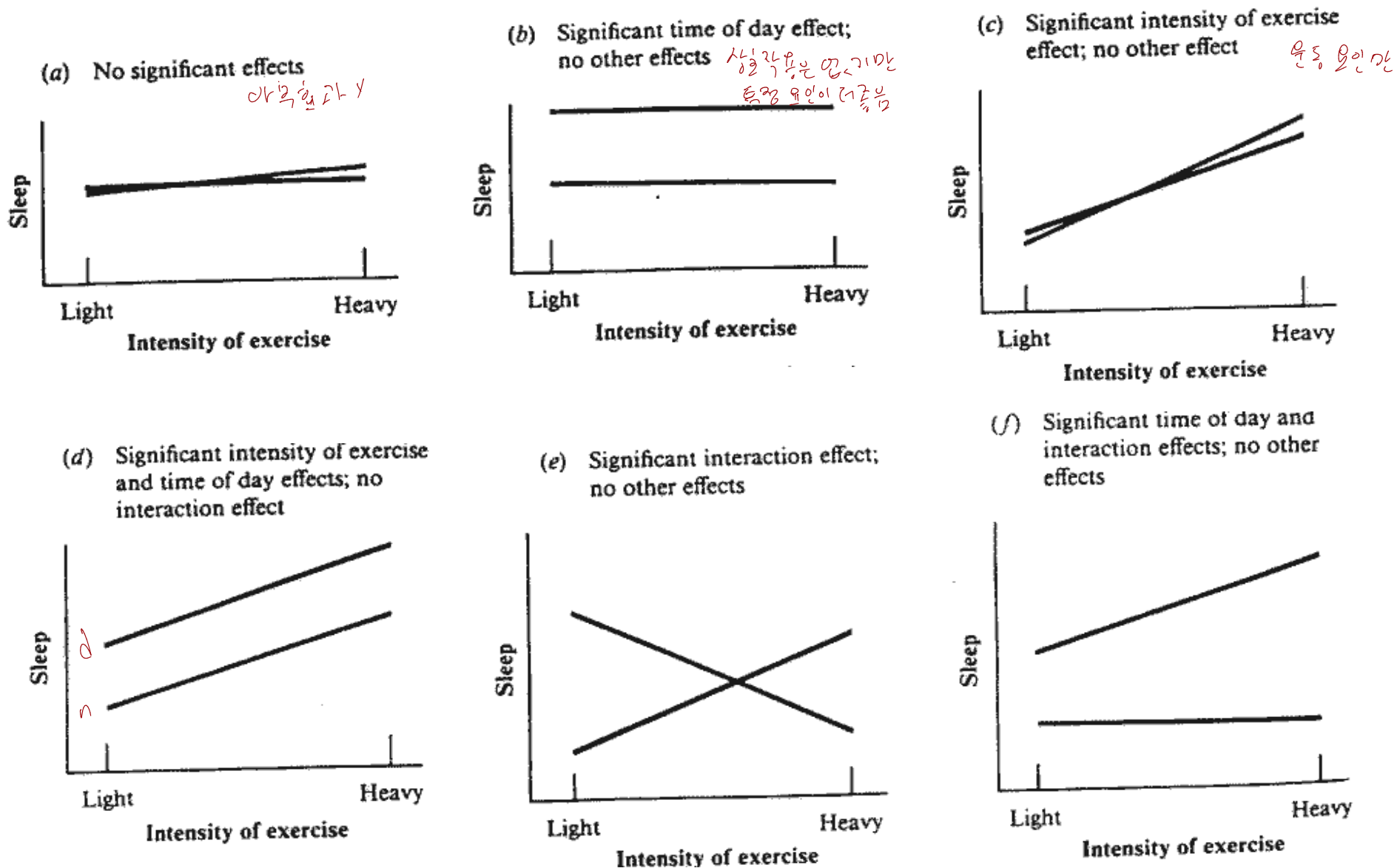


상호작용 있는 경우



↓
온도가 바뀌면서 특기하는 방향에 반대로 '맛'이

Two-Way ANOVA



Two-Way Repeated Measures ANOVA

■ 문제의 정의

- K 병원에서는 새롭게 아로마테라피 치료를 개발
- 새롭게 개발한 치료제로 향기요법을 처치 받는 실험군과
- 일반 향기치료제로 가짜 향기요법을 처치 받는 대조군을 나누고,
- 치료전과 후에 통증이 차이가 있는지를 검증하였다.
- 아로마테라피 치료제는 효과가 있었는가?



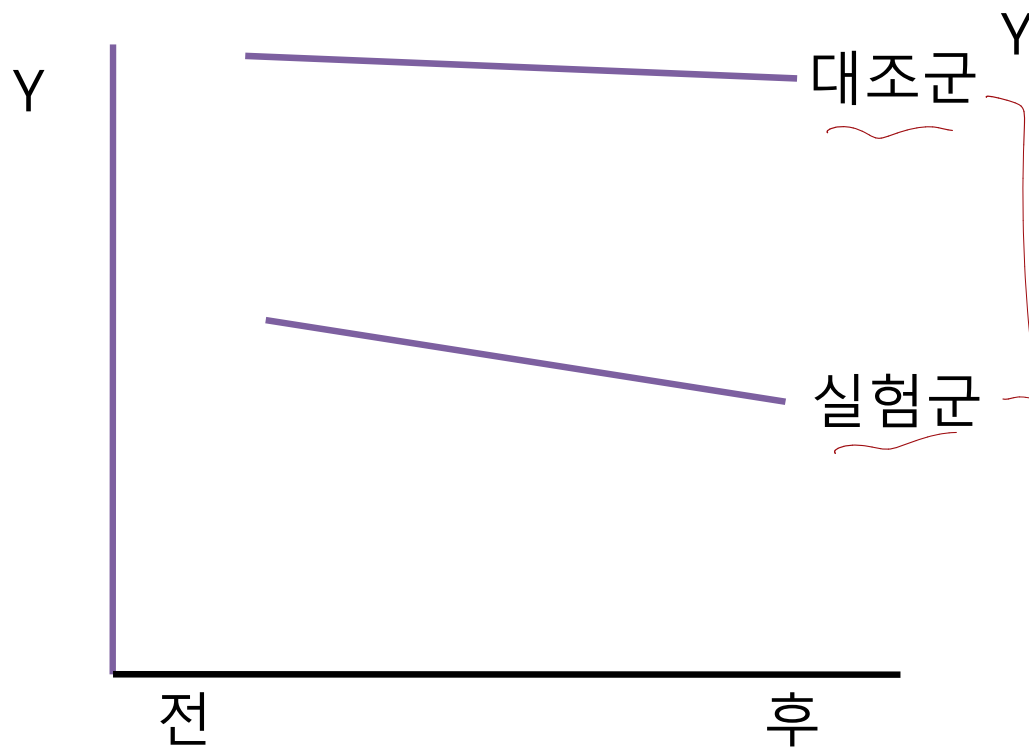
이원 반복측정 분산분석
(Two Way Repeated Measures ANOVA)

Group	Pre	Post
1	12.32	23.09
1	12.32	23.09
1	16.62	44.97
1	18.63	40.78
1	20.12	35.12
1	20.12	35.12
1	22.78	42.07
1	22.78	42.07
1	22.91	16.17
1	22.91	16.17
1	24.36	31.08
1	28.38	51.63
1	28.38	51.63
1	29.92	70.65

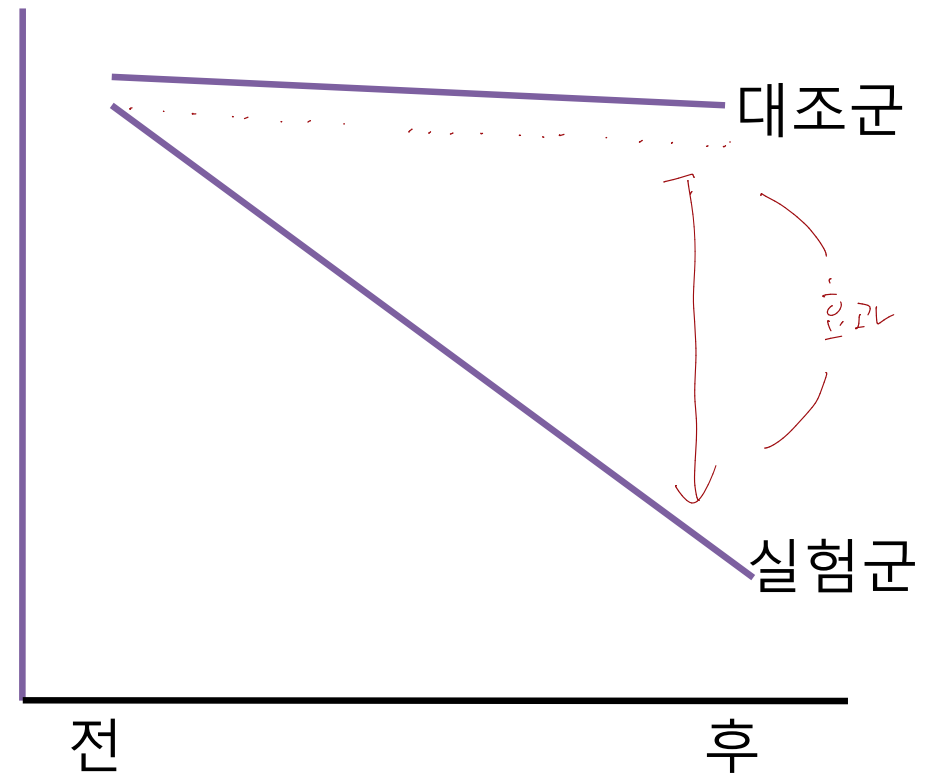
Two-Way Repeated Measures ANOVA

- 평균반응 프로파일(average response profile)

- 상호작용 없는 경우



- 상호작용 있는 경우



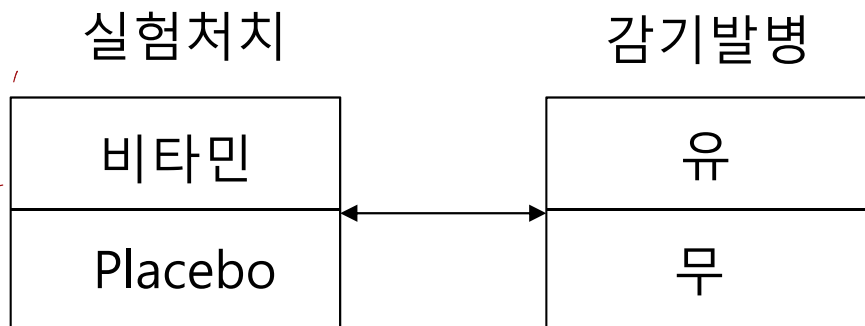
Chi-Square test

↳ 치사공정도 실험, 관계공정도 실험

■ 문제의 정의

- K 병원에서는 비타민과 감기와와의 관계를 연구하고자 한다.
- 비타민을 투여할 실험군과 가짜약을 투여할 대조군으로 구분하고 겨울동안 감기가 걸렸는지를 확인하였다.
- 과연 비타민이 감기에 효과가 있었는지 검증해 보라

독립변수와
종속변수의 관계,
빈도, 빈도 = 명목



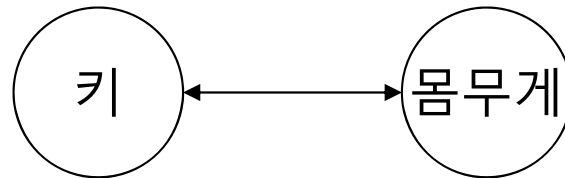
교차분석 (Chi Square)

실험처치	감기발병
1	1
1	1
1	1
1	1
1	1
1	1
1	1
1	1
1	1
1	2
1	2
1	1
1	1
1	1

Correlation Test

■ 문제의 정의

- K속옷회사는 몸무게와 키와의 관계를 조사하고자 한다.
- 몸무게와 키와는 어떤 관계가 있는가?



상관분석(Correlation)

상관분석: |r|, 원인과 결과 X

단순 회귀분석: |r|, 원인과 결과 O

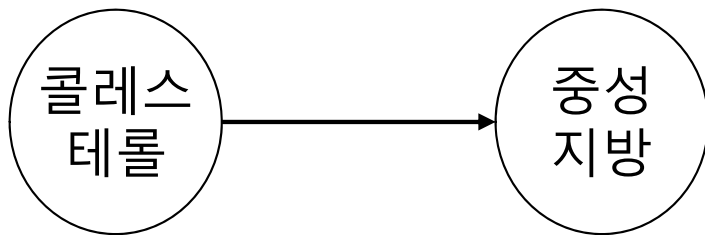
다중회귀: 다(多)인 |r|, 원인과 결과 O

몸무게	키
72	176
72	172
70	182
43	160
48	163
54	165
51	168
52	163
73	182
45	148
60	170
62	166
64	172
47	160

Regression

■ 문제의 정의

- 일개 기업체에서 근무하고 있는 직원(100명)들의 정기적인 건강검진 결과의 일부 자료이다.
- 콜레스테롤이 높으면 중성지방도 높다고 말할 수 있는가?
- 콜레스테롤 수치를 알면 중성지방 수치를 예측할 수 있는가?



단순 회귀분석(Regression)

상관분석: |대|, 원인과 결과 X

단순 회귀분석: |대|, 원인과 결과 O

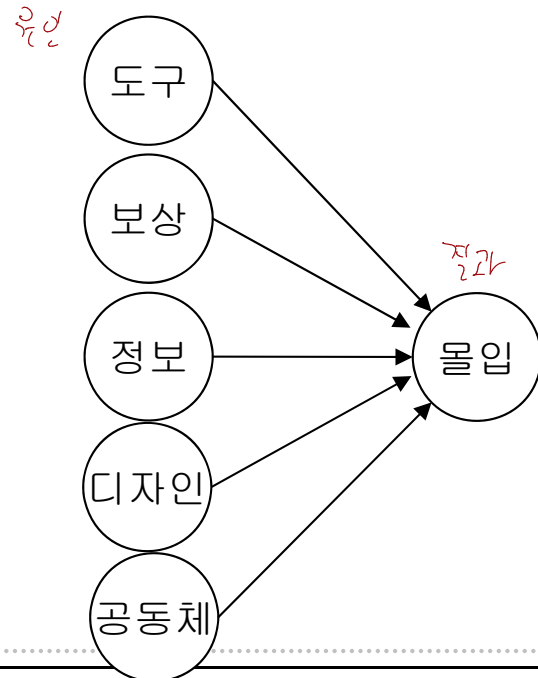
다중회귀: 다(多) |대|, 원인과 결과 O

Col	Fat
146.00	162.30
190.50	224.60
179.10	193.70
164.10	161.10
184.80	181.10
153.90	137.30
179.00	164.80
224.70	228.20
145.00	112.70
154.30	123.00
200.80	185.50
164.10	132.50
169.50	139.50
131.70	77.30

Multiple Regression

■ 문제의 정의

- 온라인게임의 몰입(즐거움)에 영향을 주는 요인이 무엇인지를 연구하고자 한다.
- 영향을 주는 변수로는 도구, 보상, 정보, 디자인, 공동체가 있다.
- 온라인게임 몰입에 영향을 주는 변수는 무엇이고, 어떤 변수가 온라인게임 몰입에 가장 큰 영향을 주는 변수인가?



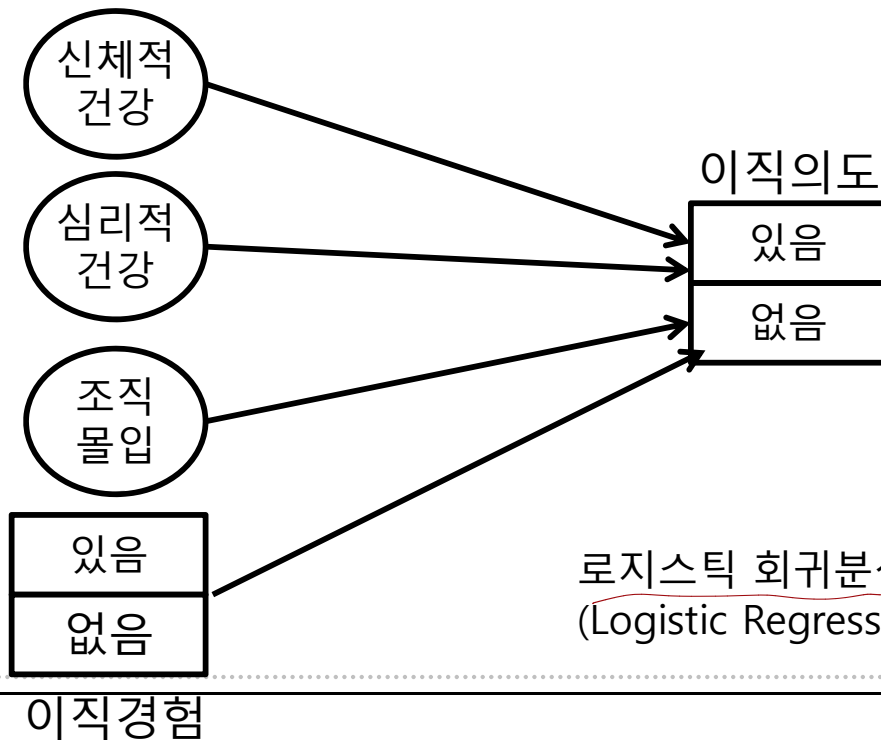
상관분석: |r|, 원인과 결과 X
 단순 회귀분석: |r|, 원인과 결과 0
 다중회귀: 다(多)개 |, 원인과 결과 0

id	도구	보상	정보	디자인	공동체	몰입
1	.75621	1.2236	-1.457	1.2315	1.6724	2.0639
2	-.3122	1.2874	-.3785	.70040	-.7892	-.3387
3	.06430	-.3206	.83555	-.1950	-1.120	.75093
4	1.8016	-.2745	-1.206	1.2018	-.4250	-2.202
5	.19992	.68203	-1.437	.14436	-1.481	-.1751
6	1.5472	-.4749	-1.450	-.8258	-.3604	-.3387
7	.62093	.56739	.29029	-.4474	.38722	.75093
8	-.9375	-.2066	-.9176	-1.099	-3.018	-3.292
9	.00033	.07652	-1.473	.33488	-.0849	-.3387
10	-1.987	1.4583	-1.241	-1.665	-.4744	-1.113
11	-1.213	-.4636	.40901	-1.677	-.2295	-.3387
12	1.5274	.75934	.29792	-.3436	.74100	.75093
13	.55031	-1.038	-.4389	1.2661	1.5892	.20029
14	-.1043	.95034	.90965	-.3209	.17023	-.3387

Logistic Regression

■ 문제의 정의

- K기업의 인사담당인 이부장은 이직의도에 영향을 주는 요인을 알고 싶다.
- 신체적 건강과 심리적건강, 조직몰입, 이직경험이 이직의도에 영향을 주는가?



로지스틱 회귀분석 : 종속변수를 그룹으로 분류될 때 (Yes or No)
(Logistic Regression) (이유는 0)

신체적건강	심리적건강	조직몰입	이직경험	이직생각
43	18	28	1	0
54	27	28	1	0
60	30	26	0	0
57	17	23	0	0
60	30	29	1	0
42	27	26	1	1
48	21	23	0	0
46	18	28	0	0
57	30	23	0	0
59	24	26	0	0
47	23	27	1	0
36	18	27	1	0
43	24	26	0	0
46	17	16	1	1