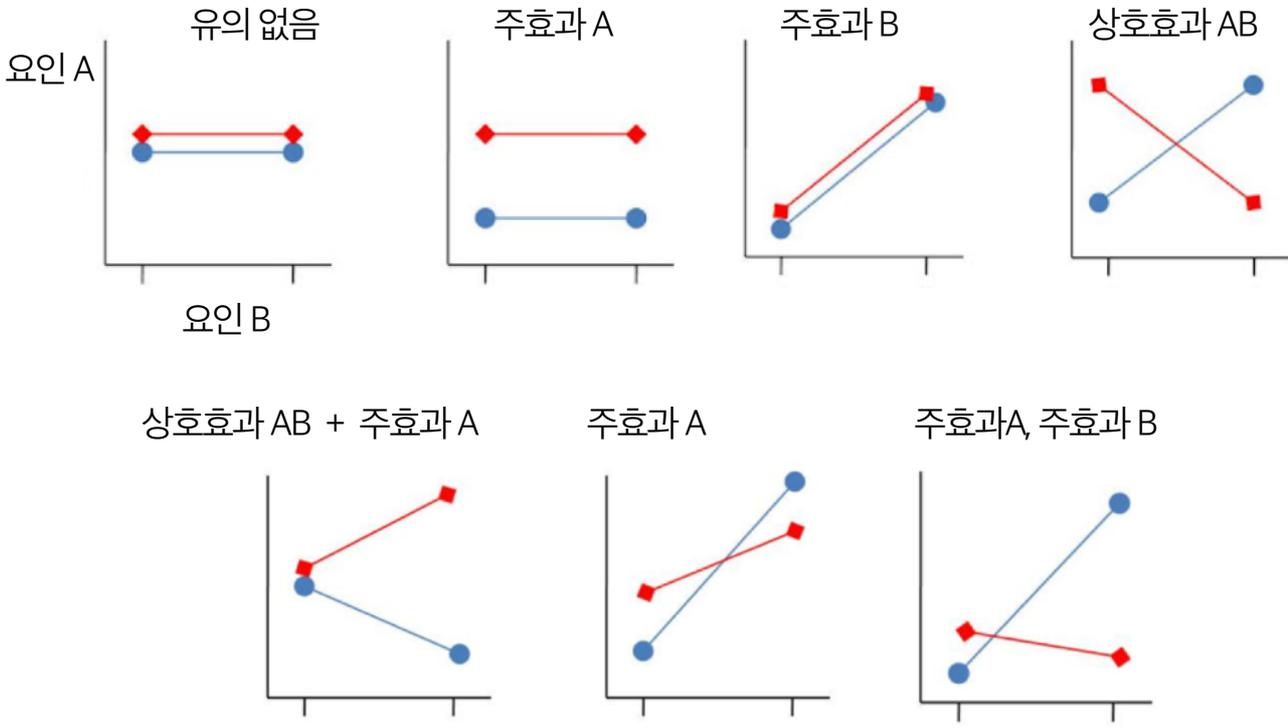


### 1. 이원분산분석 모형

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}$$

- ▶  $i = 1, 2, \dots, a, j = 1, 2, \dots, b, k(i, j)$ 내 조건 반복( $n_{ij}$ )
- ▶ 주효과  $\alpha_i$  (a개 수준),  $\beta_j$  (b개 수준)
- ▶ 상호 (교호) interaction 효과  $(\alpha\beta)_{ij}$ : 한 요인의 효과가 다른 요인의 수준에 의존하는 경우를 말합니다. 교호작용도를 사용하여 가능한 교호작용을 시각화할 수 있습니다.
- ▶ 오차항 가정  $e_{ij} \sim (iid)N(0, \sigma^2)$



### Confounding 교락

- \* 두개 이상 인자의 효과가 함께 나타나고 그 효과를 각각의 인자에 의한 효과로 분리해 낼 수 없을 경우 그 인자들은 교락(Confounding) 되어 있다고 한다.
- \* 구할 필요가 없는 2 인자 교호작용이나 고차의 교호작용(interaction effect)을 블록 또는 주효과와 교락시켜 실험의 효율을 높일 수 있다.

## 2. 모수 추정 및 총변동 분해

- ▶ 요인 A 주효과 :  $\hat{\alpha}_i = \bar{y}_{i..}$
- ▶ 요인 B 주효과 :  $\hat{\beta}_j = \bar{y}_{.j}$
- ▶ 상호 효과  $(\hat{\alpha\beta})_{ij} = \bar{y}_{ij}$ .
- ▶ SST (총변동) = SSA (주효과 A 변동) + SSB (주효과 B 변동) + SSAB (상호효과 변동) + SSE (오차변동)

$$\sum (y_{ijk} - \bar{y}_{...})^2 = \sum (\bar{y}_{i..} - \bar{y}_{...})^2 + \sum (\bar{y}_{.j} - \bar{y}_{...})^2 + \sum (\bar{y}_{ij} - \bar{y}_{i..} - \bar{y}_{.j} + \bar{y}_{...})^2 + \sum (y_{ijk} - \bar{y}_{i..} - \bar{y}_{.j} - \bar{y}_{ij})^2$$

## 3. 이원분산분석표

변동	자유도	자승합	평균자승합 (=자승합/자유도)	F
요인 A 주효과	a-1	SSA	MSA	MSA/MSE~F
요인 B 주효과	b-1	SSB	MSB	MSB/MSE~F
(AB) 교호효과	(a-1)(b-1)	SSAB	MSAB	MSAB/MSE~F
오차변동	차이	SSE	MSE	
Total 총변동	n-1	SST		

## 4. 사례연구 : JOB.csv

직장을 갖는데 걸리는 시간을 학력, 성별 2개 요인에 따라 차이가 있는지 분석하시오.

엑셀에는 각 요인의 반복 수가 같은 경우만 분석이 가능하며, 오른쪽 형식의 데이터 포맷으로 입력되어 있어야 한다.

	F	G	H	I	J
1		E1	E2	E3	E4
2	Male	10	12	15	8
3		9	11	8	9
4		12	9	7	5
5		16	14	7	11
6		14	12	7	13
7		17	16	9	8
8		13	10	14	7
9		9	10	15	11
10		11	5	11	10
11		15	11	13	8
12	Female	7	7	5	7
13		13	12	13	9
14		14	6	12	3
15		6	15	3	7
16		11	10	13	9
17		14	13	11	6
18		13	9	15	10
19		11	15	5	15
20		14	12	9	4
21		12	13	8	11

Anova: Two-Factor With Replication

Input

Input Range:

Rows per sample:

Alpha:

Output options

Output Range:

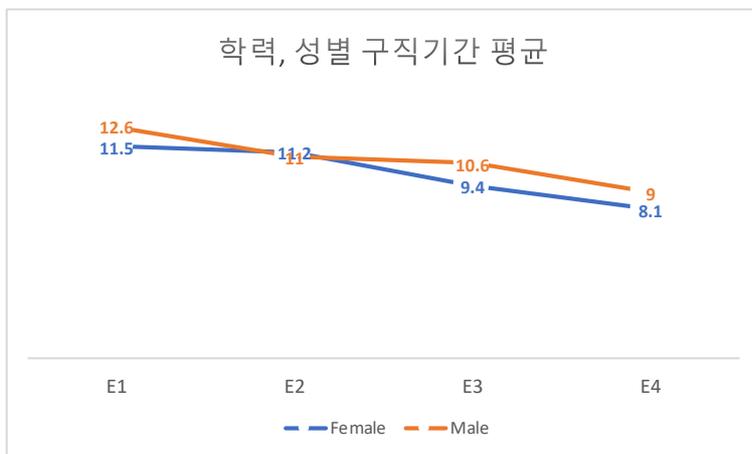
New Worksheet Ply:

성별 효과, 상호 효과 유의하지 않고 학력의 효과만 유의하다. 학력이 높아질수록 구직 기간이 짧아진다. 학력간

행 레이블	평균 : Length	표준 편차 : Length
E1	12.05	2.783433132
E2	11.1	2.87923601
E3	10	3.605551275
E4	8.55	2.854382595
(비어 있음)		
총합계	10.425	3.315776681

행 레이블	평균 : Length	표준 편차 : Length
Female	10.05	3.528101472
Male	10.8	3.043024811
(비어 있음)		
총합계	10.425	3.315776681

Source of Variation	자승합	자유도	평균자승합	F	P-value	F crit
성별	11.25	1	11.250	1.115	0.294	3.974
학력	135.85	3	45.283	4.490	0.006	2.732
상호효과	6.25	3	2.083	0.207	0.892	2.732
오차	726.2	72	10.086			
Total	879.55	79				



	E1	E2	E3	E4
행 레이블	평균 : Leng	표준 편차 : Leng	평균 : Leng	표준 편차 : Leng
Female	11.5	2.729	11.2	2.960
Male	12.6	2.728	11	2.793

## 5. 공분산 분석 Analysis of Covariance

$$\text{모형: } y_{ij} = \mu + \alpha_i + bx_{ij} + e_{ij}$$

공변량(covariate) -  $x_{ij}$

- \* 종속변수 값에 대한 요인들의 유의성 검정을 제대로 하기 위해 고려되는 변량
- \* 일반적으로 종속변수의 실험 전 값이다. (예) 교육효과에서의 사전점수
- \* 공변량은 관심의 대상이 아니라 요인의 유의성 검정을 정확하기 위하여 고려함

### 데이터 Heart.csv

스텝업 운동 후 힘든 정도를 파악하기 위하여 운동 후 맥박과 일반 상태에서 맥박을 측정하였다.

- ▶ 계단 높이 수준=2 : Height: 0 if step at the low (5.75") height, 1 if at the high (11.5") height
- ▶ 운동 빈도 수준=3 : Frequency: the rate of stepping. 0 if slow (14 steps/min), 1 if medium (21 steps/min), 2 if high (28 steps/min)
- ▶ 쉬는 상태의 맥박 : Rest\_HR: the resting heart rate of the subject before a trial, in beats per minute
- ▶ 운동 후 맥박 HR: the final heart rate of the subject after a trial, in beats per minute

쉬는 상태의 맥박은 공변량, 계단 높이 요인과 운동 빈도 요인이 운동 후 맥박에 영향을 주는지 공변량 이원분산분석 하시오.

$$\text{모형: 이원 교호작용 공분산분석 } y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + bx_{ijk} + e_{ijk}$$

- \*  $\alpha_i, \beta_j$ : 주효과
- \*  $(\alpha\beta)_{ij}$ : 교호 효과
- \*  $b$ : 회귀계수, 공변량(실험 이전 점수) 효과 제거 - 요인의 유의성을 보다 정확하게 분석하기 위하여 오차항 변동을 줄이는 역할을 함

**일반 선형 모형**

C1	Height	반응(E):	HR
C2	Frequency	모형(D):	Height Frequency Height*Frequency RestHR
C3	RestHR		
C4	HR		

**일반 선형 모형 - 공변량**

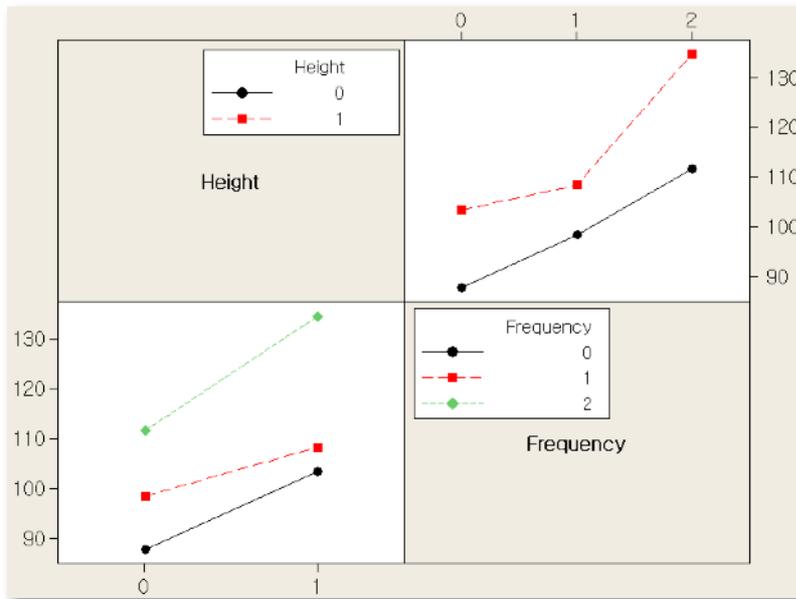
C1	Height	공변량(C):	
C2	Frequency		
C3	RestHR		RestHR
C4	HR		

**일반 선형 모형 - 요인 플롯**

C1	Height	주 효과 플롯	
C2	Frequency	요인(E):	Height Frequency
C3	RestHR	Y(반응) 척도에 대한 최소값(M):	[
C4	HR	Y(반응) 척도에 대한 최대값(X):	[
		제목(T):	
		상호작용 플롯	
		요인(A):	Height Frequency
		<input checked="" type="checkbox"/> 전체 상호작용 플롯 행렬 표시(D)	

**일반 선형 모형 - 비교**

<input checked="" type="radio"/> 쌍 비교(P)	
<input type="radio"/> 통제 수준에서의 비교(W)	
항(B):	Height Frequency



HR에 대한 분산 분석 (검정을 위해 수정된 제곱합을 사용)

출처	DF	Seq SS	Adj SS	Adj MS	F	P
Height	1	3499.2	1757.3	1757.3	14.98	0.001
Frequency	2	3727.8	4027.1	2013.6	17.16	0.000
Height*Frequency	2	210.6	213.7	106.9	0.91	0.416
RestHR	1	1985.0	1985.0	1985.0	16.92	0.000
오차	23	2698.6	2698.6	117.3		
전체	29	12121.2				

S = 10.8319 R-제곱 = 77.74% R-제곱(수정) = 71.93%

공변량을 고려하지 않는 경우

오차항의 자유도가 1개 많음 (공변량) - 오차항 평균(MSE)은 195.1 -> 공변량 있는 경우 117.3으로 낮음. 공변량 존재와 관계없이 요인들의 SS(변동제곱합)은 동일함 -> 그러므로 공변량은 주효과, 교호효과와 F-값은 커져 유의성 증가함.

HR에 대한 분산 분석 (검정을 위해 수정된 제곱합을 사용)

출처	DF	Seq SS	Adj SS	Adj MS	F	P
Height	1	3499.2	3499.2	3499.2	17.93	0.000
Frequency	2	3727.8	3727.8	1863.9	9.55	0.001
Height*Frequency	2	210.6	210.6	105.3	0.54	0.590
오차	24	4683.6	4683.6	195.1		
전체	29	12121.2				