

실험설계

개요

어떤 원인이 반응에 유의한 영향을 주고 있는가를 파악하고 그 영향이 양적으로 어느 정도 큰가를 알아내고자 실시함(추정과 검정)

적은 영향밖에 미치지 못하는 요인(오차)들은 전체적으로 어느 정도 영향을 주고 있으며, 측정오차는 어느 정도인가를 알아내고자 실시함(오차 항 추정)

유의한 영향을 미치는 원인들이 어떠한 조건을 가질 때 가장 바람직한 반응을 얻을 수 있는가를 알아내기 위해서 실시함(최적화)

용어

- **요인(Factor):** 요인은 제어 가능한 변수 중 하나로 반응에 미치는 영향을 의미하며, 실험공정에서 연구된다. 요인(인자)은 X 로 표현된다. 인자(요인)는 온도나 시간처럼 정량적일 수 있고 또한 다른 기계, 다른 작업자, 또는 환경처럼 정성적일 수 있다.
- **반응(Response):** 주어진 수준에서 인자들의 수준 조합에 의해 측정되는(생산되는) 특성치이다.
- **수준(Level):** 인자 수준이란 실험에서 조절될 수 있는 인자 값을 의미한다.
- **교호작용(Interaction):** 두 인자들 사이의 조합에서 일어나는 효과

기본원칙

Randomization 랜덤화

- 관심 요인 외에 기타 원인들의 영향이 실험결과에 미치지 않게 함
- 시간에 따라 변하는 인자의 효과나 경향을 실험을 시간대에 대해 균일하게 배치함으로써 약화시킬 수 있다.

Blocking 블록화

- 동일한 성질을 가진 단위들의 집합 (Block)
- 실험전체를 시간적 혹은 공간적으로 분할하여 Block을 만들어 주면 각 Block 내에서는 실험환경이 균일하게 되어 좀더 좋은 결과를 얻을 수 있다.



- Block 은 실험 계획 시 또 다른 독립변수(요인)로 취급해야 한다.
- 실험이 이틀에 걸쳐서 수행되었다고 하면 실험 일을 Block 이라 한다.

Replication 반복화

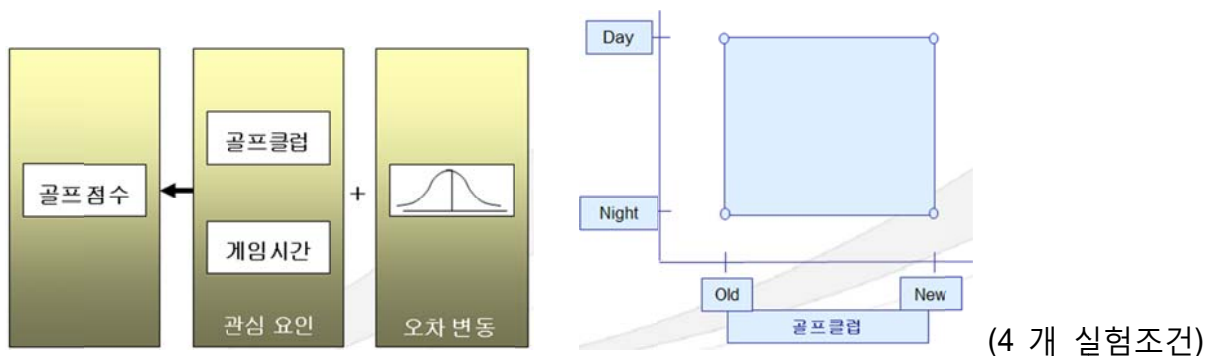
- 실험조건을 처음부터 다시 setup 하여 실험하는 것.
- 실험의 재현성을 알아보기 위한 방법 : 실험결과의 신뢰성을 높일 수 있다.
- 오차 변동을 계산할 수 있다.

(참고) Repetition (반복) : 동일실험 조건에서 즉시 반복 실험 => 평균 관측치를 하나의 관측치로 이용

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.

2²-요인 설계



- 요인 A 주효과 : 클럽효과 수준 = (New, Old) : 수준별 평균의 차이
- 요인 B 주효과 : 경기기각 수준 = (밤, 낮)
- (AB) 교호 효과 : 두 요인 수준의 결합 조건의 평균 차이



One-way ANOVA 일원 분산분석

개요

- 요인 (분류형, 질적 설명변수)이 하나
- 집단이 3 개 이상인 모집단 평균 차이 검정
- 실험설계에서는 CRD completely Randomized Design 완전 임의설계

데이터

| 요인 | 일원 배치 설계 | | | | 모집단 중앙 위치 비교 |
|-----|----------|----------|---------------|-------|---------------------------------|
| | 1 | 2 | ... | 반복 수 | |
| 1 | y_{11} | y_{12} | ... y_{1m} | n_1 | y_{11} y_{12} ... y_{1m} |
| 2 | y_{21} | y_{22} | ... y_{2n2} | n_2 | y_{21} y_{22} ... y_{2n2} |
| ... | ... | ... | | ... | ... |
| k | y_{k1} | y_{k2} | ... y_{knk} | n_k | y_{k1} y_{k2} ... y_{knk} |

기호

i = 처리변수 수준 첨자, j = 반복 첨자

n_i = 처리변수 i 수준의 반복 수, n = 실험 총반복 수

y_{ij} : 처리변수 i 번째 수준의 j 번째 반복의 종속변수 결과 값

\bar{y}_i : 처리변수 i 번째 수준의 종속변수 평균

$\bar{y}_..$: 종속변수 반응 값의 총평균

모형 model

$$Y_{ij} = \mu + \alpha_i + e_{ij}, \quad i = 1, 2, \dots, k, \quad j = 1, 2, \dots, n_i$$

- MVUE : $\hat{\mu} = \bar{y}_.., \hat{\alpha}_i = \bar{y}_i$

가정 assumption

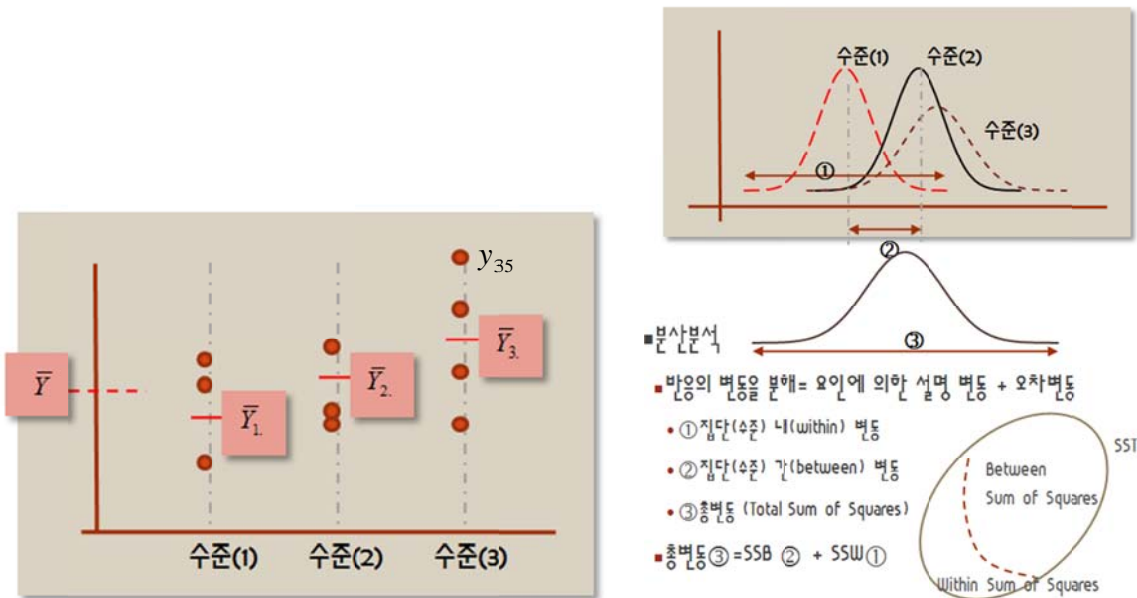
$e_{ij} \sim iidN(0, \sigma^2)$ 독립성 / 정규성 / 등분산성



집단 간 평균의 차이?

귀무가설 : $\mu_1 = \mu_2 = \dots = \mu_k$ (모든 집단 평균은 동일하다) \Leftrightarrow 모든 $\alpha_i = 0$ (요인의 주효과는 유의하지 않음) \Leftrightarrow 요인의 효과는 유의하지 않다.

대립가설 : 적어도 하나 이상의 집단 평균은 유의하다.



변동 분해 Variation Decomposition

총변동자승합 SST (Total Sum of Squares)
$$SST = \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{..})^2 \quad (3)$$

$$SST = \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{i.})^2 + \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{Y}_{i.} - \bar{Y}_{..})^2 = (1) + (2)$$

- (1) = SSE (Error SS 집단내 변동, 오차변동)
- (2) = SSA (Between SS / Treatment SS 집단간 변동, 처리변동, 요인 A 주효과)

가정 $e_{ij} \sim iidN(0, \sigma^2)$ 으로부터 $y_{ij} \sim iidN(\mu + \alpha_i, \sigma^2)$ 이므로

- $SSA \sim \chi^2(k - 1)$, $SSE \sim \chi^2(n - k - 1)$ 이고 서로 독립이다.
- 그러므로 $\frac{SSA/(k - 1)}{SSE/(n - k - 1)} \sim F(k - 1, n - k - 1)$



분산분석표 ANOVA table

| 변동 | 자유도 | 자승합 | 평균자승합 | F |
|------|-------|-----|-----------------------------|---|
| 요인변동 | k-1 | SSA | $MSA = \frac{SSA}{(k-1)}$ | $TS = \frac{MSA}{MSE} \sim F(k-1, n-k-1)$ |
| 오차변동 | n-k-1 | SSE | $MSE = \frac{SSE}{(n-k-1)}$ | |
| 총변동 | n-1 | SST | | |

비즈니스 표 작성

| 집단 | 평균(M) | 표준편차(SD) | F-검정통계량 |
|-----|-------|----------|---------|
| 1 | | | 유의확률 포함 |
| ... | | | |
| k | | | |

사후 검정 (Post-hoc test) 혹은 다중 비교 (multiple comparison)

분산 분석의 F-검정은 단지 귀무가설 $H_0 : u_1 = u_2 = \dots = u_t$ 즉 전체적인 차이를 검정하는 것이다. 그러므로 수준별 차이(pairwise: 예: $H_0 : u_1 = u_3$)가 있는지 혹은 수준의 선형 결합 대비(contrast: 예: $H_0 : u_1 = \frac{u_2 + u_3}{2}$)의 차이가 있는지 검정할 필요가 있는데 이를 사후 검정 혹은 다중 비교(대비 포함)라 한다. 사후 검정이므로 비록 F-검정 결과와 관계없이 (귀무가설을 채택하더라도) 시행하게 된다.

다중 비교에서는 여러 개의 가설을 동시에 검정하므로 유의수준을 조정해야 한다. 이를 조정된 실험 유의수준 (controlled experimental error rate)이라 하고 $1-(1-\alpha)^c$ 이다. 여기서 c는 가설 수를 의미한다. pairwise 다중 비교의 경우 $c = t(t-1)/2$ 가 된다.

Fisher's Least Significant Difference

- pairwise (두 수준별 평균 비교) 검정에 사용하나 이는 다중 비교에 해당되지는 않는다. 두 모집단 평균 차이 검정에서 통합 분산 (pooled variance) 대신 MSE 사용한다.

Tukey HSD(honestly significant difference) procedure

- $q = \frac{\bar{Y} - \bar{Y}_j}{MSE / \sqrt{n}}$ 분포에 대한 표 제공



- 가장 보수적인(귀무가설 기각하지 않음) 방법이다. (자연과학 분야)

Duncan Multiple range test

- Tukey 방법과 매우 유사하나 수준별 표본 평균을 크기 순으로 나열하여 차이가 가장 큰 것을 비교해 가면서 유의 수준을 $1-(1-\alpha)^r$ 으로 조정해 가면서 검정한다. r 은 검정 단계 순서이다.
- 귀무가설을 기각할 확률이 매우 높아 자주 사용하지 않는다.

Scheffe's S method

- 대비(contrast)까지 고려한 다중 비교 방법이다. (사회과학 분야)

Dunnett's procedure

- 처리 효과의 수준 하나가 control (실험 집단)인 경우 (예: placebo 집단, 교육을 하지 않는 집단, 이전 약 투여 집단) 이 집단과 다른 집단들을 pairwise 비교할 경우 사용된다.

예제데이터 SHIPS.xls

유람선 운영 S 사는 배의 크기에 따른 고객 만족도 점수의 차이가 있는지 알아보기 위한 조사한 것이다. 배 크기별로 8 개 유람선 임의 추출, 각 유람선 탑승 고객의 평가 점수를 평균한 측정치이다.

```

SAS | • One-way ANOVA
ods graphics on;
proc glm data=ship;
  class size;
  model score=size;
  means size / scheffe;
run;quit;

proc tabulate data=ship;
  class size;
  var score;
  table (size all), (score)*(mean std);
run;
    
```

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|-----------------|----|----------------|-------------|---------|--------|
| Model | 2 | 226.0933333 | 113.0466667 | 3.70 | 0.0419 |
| Error | 21 | 640.8400000 | 30.5161905 | | |
| Corrected Total | 23 | 866.9333333 | | | |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|--------|----|-------------|-------------|---------|--------|
| Size | 2 | 226.0933333 | 113.0466667 | 3.70 | 0.0419 |

Type III 는 다른 요인의 효과가

고정된 경우 대응 변인의 유의성 검증



Means with the same letter are not significantly different.

| Scheffe Grouping | Mean | N | Size |
|------------------|--------|---|--------|
| A | 92.200 | 8 | Small |
| A | | | |
| B | 89.650 | 8 | Medium |
| B | | | |
| B | 84.800 | 8 | Large |

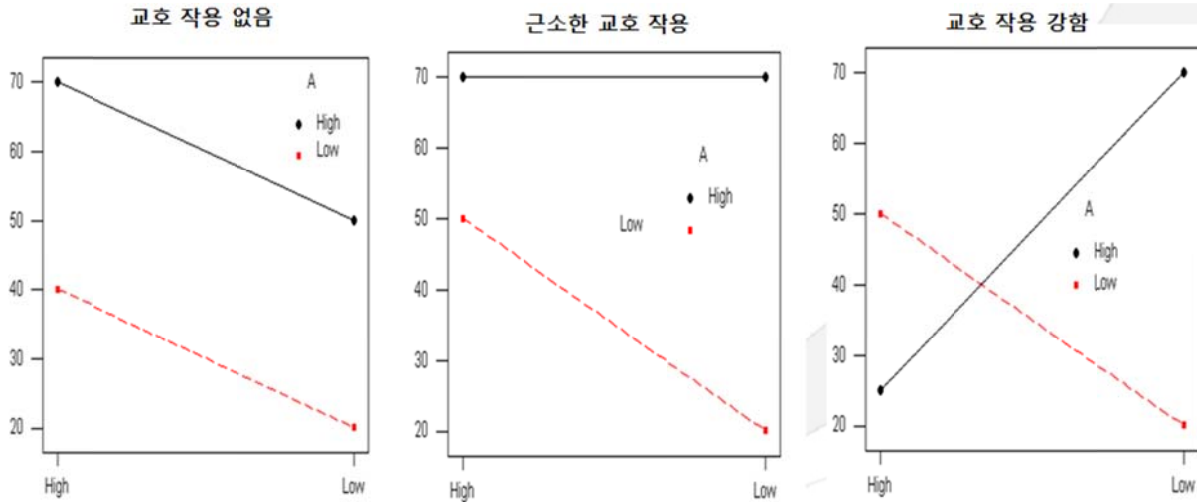
| 배크기 | 평균M | 표준편차SD | F-통계량 |
|----------------------|-------|--------|-----------------|
| Large ^a | 84.8 | 5.13 | 3.7 (0.0419) |
| Medium ^{ab} | 89.65 | 4.97 | |
| Small ^a | 92.2 | 6.37 | |
| 합계 | 88.88 | 6.14 | |

- F-검정 결과 배 크기에 따른 만족도의 차이는 유의함
- 다중비교 결과 (소형>중형) > (중형>대형) : 그러므로 소형과 대형 배의 승객의 만족도 차이만 유의하다.

Two-Way ANOVA

개요

- 요인 (분류형, 질적 설명변수)이 두 개
- 요인 개별효과를 주효과 (각 요인의 수준별 평균차이)
- 두 요인의 수준 결합에 의해 평균의 차이가 나는 효과, 사회과학분석에서는 일반적으로 교호효과를 사용하지 않는다 (이유는 설명이 용이하지 않음). 주효과를 보기 위한 변동분해로만 사용한다.



모형 model $Y_{ij} = \mu + \alpha_i + b_j + (ab)_{ij} + e_{ij}, i = 1, 2, \dots, a, j = 1, 2, \dots, b$

가정 assumption $e_{ij} \sim iidN(0, \sigma^2)$ 독립성 / 정규성 / 등분산성

통계적 가설

귀무가설 : 요인 A의 수준별 평균의 차이는 없다 ⇔ 요인 A 주효과 유의하지 않음

귀무가설 : 요인 B의 수준별 평균의 차이는 없다

귀무가설 : 모든 $(ab)_{ij} = 0$ ⇔ 요인 (A, B)의 교호작용은 유의하지 않음

변동 분해 Variation Decomposition

총변동자승합 SST (Total Sum of Squares) $SST = \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{..})^2$

$$SST = \sum_{i=1}^a \sum_{j=1}^b (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2 + \sum_{i=1}^a \sum_{j=1}^b (\bar{Y}_{i.} - \bar{Y}_{..})^2 + \sum_{i=1}^a \sum_{j=1}^b (\bar{Y}_{.j} - \bar{Y}_{..})^2 + \sum_{i=1}^a \sum_{j=1}^b (\bar{Y}_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2$$

= SSE + SSA + SSB + SSAB

분산분석표 ANOVA table

| 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 | | |

비즈니스 표 작성

| 요인 A | 1 | ... | A | F-검정통계량 |
|------|---|-------|---|------------|
| 요인 B | | | | |
| 1 | | 평균±SD | | 요인 A 주효과 |
| ... | | | | 요인 B 주효과 |
| b | | | | 요인 AB 교호효과 |



예제데이터 SHIPS.xls

- Height: 0 if step at the low (5.75") height, 1 if at the high (11.5") height
- Frequency: the rate of stepping. 0 if slow (14 steps/min), 1 if medium (21 steps/min), 2 if high (28 steps/min)
- Rest_HR: the resting heart rate of the subject before a trial, in beats per minute
- HR: the final heart rate of the subject after a trial, in beats per minute

요인 (계단 높이, 올라가는 속도)가 운동 후 심장박동(HR)에 영향을 미치나?

```

SAS • One-way ANOVA
ods graphics on;
proc glm data=heart;
  class frequency height;
  model hr=frequency|height;
  means frequency / tukey;
run;quit;

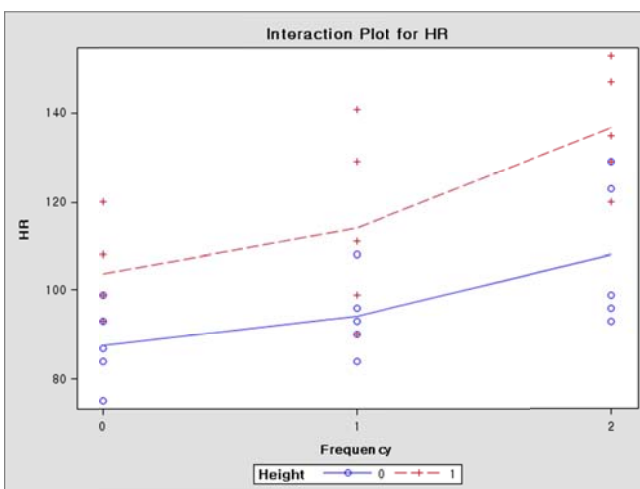
proc tabulate data=ship;
  class size;
  var score;
  table (size all), (score)*(mean std);
run;
    
```

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|-----------------|----|----------------|-------------|---------|--------|
| Model | 5 | 7437.60000 | 1487.52000 | 7.62 | 0.0002 |
| Error | 24 | 4683.60000 | 195.15000 | | |
| Corrected Total | 29 | 12121.20000 | | | |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|------------------|----|-------------|-------------|---------|--------|
| Frequency | 2 | 3727.800000 | 1863.900000 | 9.55 | 0.0009 |
| Height | 1 | 3499.200000 | 3499.200000 | 17.93 | 0.0003 |
| Frequency*Height | 2 | 210.600000 | 105.300000 | 0.54 | 0.5899 |

대체

- 빈도, 높이 모두 유의함, 교호작용 유의하지 않음



Means with the same letter are not significantly different.

| Tukey Grouping | Mean | N | Frequency |
|----------------|---------|----|-----------|
| A | 122.400 | 10 | 2 |
| B | 104.100 | 10 | 1 |
| B | 95.700 | 10 | 0 |



| 빈도 | 높이 | | | | 합계 | | F-통계량 |
|----------------|------|-------|-------|-------|-------|-------|-------------------|
| | 0 | | 1 | | 평균M | SD | |
| | 평균M | SD | 평균M | SD | | | |
| 0 ^a | 87.6 | 9.1 | 103.8 | 10.52 | 95.7 | 12.61 | 빈도 9.55(0.0009) |
| 1 ^a | 94.2 | 8.9 | 114 | 21 | 104.1 | 18.44 | 높이 17.93(0.0003) |
| 2 ^b | 108 | 16.7 | 136.8 | 13.35 | 122.4 | 20.82 | 교호효과 0.54(0.5899) |
| 합계 | 96.6 | 14.26 | 118.2 | 20.31 | 107.4 | 20.44 | |

- 높이가 높으면 심장박동이 높아진다
- 빈도가 많으면 심장박동이 높아진다. 빈도=2 만 다른 빈도와 유의한 차이 있음
- 교호작용은 유의하지 않음 => (높이=1, 빈도=2) 결합조건에서 심장박동이 가장 높아진다.

예제

다음은 가솔린 종류(A, B, C)와 첨가제(1, 2, 3, 4)에 따른 연비의 차이가 있는
알아보기 위하여 동일 차종 24 대를 이용하여 (물론 차에 의한 차이도 있을 것이라
생각되지만 무시한다) 측정한 결과이다.

| 가솔린 \ 첨가제 | 1 | | 2 | | 3 | | 4 | |
|-----------|------|------|------|------|------|------|------|------|
| | A | 27.4 | 28.6 | 33 | 32 | 33.5 | 32.3 | 30.8 |
| B | 33.3 | 34.5 | 35.6 | 34.4 | 33.4 | 33.1 | 29.6 | 30.6 |
| C | 33 | 33.5 | 34.7 | 33.3 | 33 | 32 | 28.6 | 29.8 |

공분산분석 ANCOVA (Analysis of Covariance)

공변량(covariate)

- 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

모형 $Y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + x_{ijk} + \epsilon_{ijk}$

```

SAS | • One-way ANOVA
ods graphics on;
proc glm data=heart;
class frequency height;
model hr=resthr frequency | height ;
means frequency / tukey;
run;quit;
    
```



| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|-----------------|----|----------------|-------------|---------|--------|
| Model | 6 | 9422.59415 | 1570.43236 | 13.38 | <.0001 |
| Error | 23 | 2698.60585 | 117.33069 | | |
| Corrected Total | 29 | 12121.20000 | | | |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|------------------|----|-------------|-------------|---------|--------|
| RestHR | 1 | 1984.994148 | 1984.994148 | 16.92 | 0.0004 |
| Frequency | 2 | 4027.137543 | 2013.568772 | 17.16 | <.0001 |
| Height | 1 | 1757.252838 | 1757.252838 | 14.98 | 0.0008 |
| Frequency*Height | 2 | 213.719291 | 106.859645 | 0.91 | 0.4162 |

- 공변량은 유의함
- 빈도, 높이 요인 유의함, 교호효과 유의하지 않음

