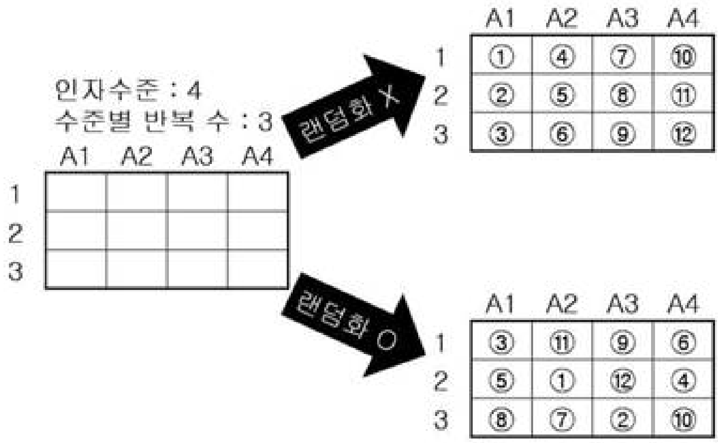
# 제 2 장 일원배치 분산분석

## 1 랜덤화 과정

1. 일원배치법의 특징
   - 수준수와 각 수준에서 취해지는 측정치의 반복수에는 제한이 없다.
   - 반복수가 모든 수준에 대하여 같지 않아도 된다.
   - 실험의 측정은 실험의 장 전체를 완전히 랜덤화하여 모든 특성치를 랜덤한 순서에 의해 측정




## 2 자료 구조와 모형

1. 자료의 구조와 모형

| | 처리 | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | $\cdots$ | $l$ | 합계 |
| | $Y_{11}$ | $Y_{21}$ | $\cdots$ | $Y_{l1}$ | $Y_{\cdot 1}$ |
| | $Y_{12}$ | $Y_{22}$ | $\cdots$ | $Y_{l2}$ | $Y_{\cdot 2}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| | $Y_{1m}$ | $Y_{2m}$ | $\cdots$ | $Y_{lm}$ | $Y_{\cdot m}$ |
| 합계 | $Y_{1\cdot}$ | $Y_{2\cdot}$ | $\cdots$ | $Y_{l\cdot}$ | $\boldsymbol{Y}_{\cdot\cdot}$ |
| 평균 | $\overline{Y}_{1\cdot}$ | $\overline{Y}_{2\cdot}$ | $\cdots$ | $\overline{Y}_{l\cdot}$ | $\overline{\boldsymbol{Y}}_{\cdot\cdot}$ |

① $Y_{ij} = \mu_i + \epsilon_{ij}\ (i=1,\ 2,\ \cdots,\ l,\ j=1,\ 2,\ \cdots,\ m),\ \epsilon_{ij} \overset{iid}{\sim} N(0,\ \sigma^2)$

② $Y_{ij} = \mu + \tau_i + \epsilon_{ij}\ (i=1,\ 2,\ \cdots,\ l,\ j=1,\ 2,\ \cdots,\ m),\ \epsilon_{ij} \overset{iid}{\sim} N(0,\ \sigma^2)$

$\mu = \frac{1}{l}\sum_{i=1}^{l}\mu_i$ (총 평균, overall mean)

$\tau_i = \mu_i - \mu (i=1,\ 2,\ \cdots,\ l)$ ($i$ 번째 처리 효과), $\sum_{i=1}^{l} \tau_i = 0$

$\Rightarrow\ \sum_{i=1}^{l} \tau_i = \sum_{i=1}^{l} (\mu_i - \mu) = \sum_{i=1}^{l} \mu_i - l\mu = l\left(\frac{1}{l}\sum_{i=1}^{l} \mu_i - \mu\right) = 0$

## 3 분산분석의 원리

1. 통계적 가설

1) $H_0 : \mu_1 = \mu_2 = \ \cdots \ = \mu_m$ vs $H_1$ : not $H_0$

2) $H_0 : \tau_1 = \tau_2 = \ \cdots \ = \tau_m = 0$ vs $H_1$ : not $H_0$

2. 모수의 직관적인 추정

1) 모수의 추정

① $\hat{\mu} = \sum \frac{Y_{ij}}{N} = \overline{Y}_{..}$

② $\hat{\mu_i} \ = \sum \frac{Y_{ij}}{l} = \overline{Y}_{i.}$

③ $\hat{\tau_i} = \hat{\mu_i} - \hat{\mu} = \overline{Y}_{i.} - \overline{Y}_{..}$

④ $\hat{\epsilon_i} = Y_{ij} - \overline{Y}_{i.}$

2) 최소제곱법(least squares models)에 의한 추정

; 오차들의 제곱합이 최소가 되도록 모수들의 추정량을 구하는 방법으로 최소제곱법으로 구한 추정량들은 일반적으로 매우 바람직한 성질을 가짐이 알려져 있다.

① 최소제곱법의 방법

; $Q \ = \sum\sum \epsilon_{ij}^2 = \sum\sum(Y_{ij} - \mu - \tau_i)^2$

$\Rightarrow$ $Q$를 최소로하는 $\mu$ 와 $\tau_i$ 들은 다음 연립방정식의 해로 결정된다.

$$\frac{\partial Q}{\partial \mu} = -2\sum\sum(Y_{ij} - \mu - \tau_i) = 0$$

$$\frac{\partial Q}{\partial \tau_i} = -2\sum_j(Y_{ij} - \mu - \tau_i) = 0 \ \ (i=1,\ 2,\ \cdots,\ l)$$

$$\therefore \ l m \mu + m \sum \tau_i = Y_{..}$$

$$m\mu + m\tau_i = Y_{i.} \ \ (i=1,\ 2,\ \cdots,\ l)$$

이 정규방정식의 해를 $\hat{\mu}$, $\hat{\tau_1}$, $\hat{\tau_2}$, $\cdots$, $\hat{\tau_t}$ 이라 하고 식을 정리하면 다음과 같다.

$$\begin{aligned}
&N\hat{\mu} + r\hat{\tau_1} + r\hat{\tau_2} + \ \cdots \ + r\hat{\tau_t} = N\overline{Y}_{..} \\
&r\hat{\mu} + r\hat{\tau_1} = r\overline{Y}_{1.} \\
&r\hat{\mu} + r\hat{\tau_2} = r\overline{Y}_{2.} \\
&r\hat{\mu} + r\hat{\tau_3} = r\overline{Y}_{2.} \\
&\vdots \qquad\qquad \vdots \\
&r\hat{\mu} + r\hat{\tau_l} = m\overline{Y}_{..}
\end{aligned}$$

일반적으로 조건 $\sum \tau_i = 0$을 삽입하여 해를 구하면

$\hat{\mu} = \overline{Y}_{..}$, $\hat{\tau_i} = \overline{Y}_{i\cdot} - \overline{Y}_{..}$, $\hat{\mu_i} = \hat{\mu} + \hat{\tau_i} = \overline{Y}_{i\cdot}$의 값을 구할 수 있다.

3. 제곱합과 자유도

1) 총편차의 분해식

① 관측값 = 총 평균 + 처리효과 편차 + 오차편차

$Y_{ij} = \overline{Y}_{..} + (\overline{Y}_{i\cdot} - \overline{Y}_{..}) + (Y_{ij} - \overline{Y}_{i\cdot})$

② 총편차 = 처리효과 편차 + 오차편차

$(Y_{ij} - \overline{Y}_{..}) = (\overline{Y}_{i\cdot} - \overline{Y}_{..}) + (Y_{ij} - \overline{Y}_{i\cdot})$

2) 제곱합의 분해

; $(Y_{ij} - \overline{Y}_{..}) = (\overline{Y}_{i\cdot} - \overline{Y}_{..}) + (Y_{ij} - \overline{Y}_{i\cdot})$의 양변을 제곱하고 모든 관측값들에 대하여 합을 취한다.

$$\Rightarrow \sum_{i=1}^{l}\sum_{j=1}^{m}(Y_{ij} - \overline{Y}_{..})^2$$

$$= \sum_{i=1}^{l}\sum_{j=1}^{m}\left[(\overline{Y}_{i\cdot} - \overline{Y}_{..}) + (Y_{ij} - \overline{Y}_{i\cdot})\right]^2$$

$$= \sum_{i=1}^{l}\sum_{j=1}^{m}(\overline{Y}_{i\cdot} - \overline{Y}_{..})^2 + \sum_{i=1}^{l}\sum_{j=1}^{m}(Y_{ij} - \overline{Y}_{i\cdot})^2 + 2\sum_{i=1}^{l}\sum_{j=1}^{m}(\overline{Y}_{i\cdot} - \overline{Y}_{..})(Y_{ij} - \overline{Y}_{i\cdot})$$

$$= \sum_{i=1}^{l}\sum_{j=1}^{m}(\overline{Y}_{i\cdot} - \overline{Y}_{..})^2 + \sum_{i=1}^{l}\sum_{j=1}^{m}(Y_{ij} - \overline{Y}_{i\cdot})^2 + 2\sum_{i=1}^{l}(\overline{Y}_{i\cdot} - \overline{Y}_{..})\sum_{j=1}^{m}(Y_{ij} - \overline{Y}_{i\cdot})$$

$$= \sum_{i=1}^{l}\sum_{j=1}^{m}(\overline{Y}_{i\cdot} - \overline{Y}_{..})^2 + \sum_{i=1}^{l}\sum_{j=1}^{m}(Y_{ij} - \overline{Y}_{i\cdot})^2 + 2\sum_{i=1}^{l}(\overline{Y}_{i\cdot} - \overline{Y}_{..})\left(\sum_{j=1}^{m}Y_{ij} - m\overline{Y}_{i\cdot}\right)$$

$$= \sum_{i=1}^{l}\sum_{j=1}^{m}(\overline{Y}_{i\cdot} - \overline{Y}_{..})^2 + \sum_{i=1}^{l}\sum_{j=1}^{m}(Y_{ij} - \overline{Y}_{i\cdot})^2 + 2\sum_{i=1}^{l}(\overline{Y}_{i\cdot} - \overline{Y}_{..})\left(\sum_{j=1}^{m}Y_{ij} - m\frac{1}{m}\sum_{j=1}^{m}Y_{ij}\right)$$

$$= \sum_{i=1}^{l}\sum_{j=1}^{m}(\overline{Y}_{i\cdot} - \overline{Y}_{..})^2 + \sum_{i=1}^{l}\sum_{j=1}^{m}(Y_{ij} - \overline{Y}_{i\cdot})^2 + 0$$

$$= \sum_{i=1}^{l}\sum_{j=1}^{m}(\overline{Y}_{i\cdot} - \overline{Y}_{..})^2 + \sum_{i=1}^{l}\sum_{j=1}^{m}(Y_{ij} - \overline{Y}_{i\cdot})^2$$

$$\therefore \sum_{i=1}^{l}\sum_{j=1}^{m}(Y_{ij} - \overline{Y}_{..})^2 = \sum_{i=1}^{l}\sum_{j=1}^{m}(\overline{Y}_{i\cdot} - \overline{Y}_{..})^2 + \sum_{i=1}^{l}\sum_{j=1}^{m}(Y_{ij} - \overline{Y}_{i\cdot})^2$$

3) 제곱합의 간단식

① $SS_T = \sum_{i=1}^{l}\sum_{j=1}^{m}(Y_{ij} - \overline{Y}_{..})^2$

$$= \sum_{i=1}^{l}\sum_{j=1}^{m}(Y_{ij}^2 - 2Y_{ij}\overline{Y}_{..} + \overline{Y}_{..}^2)$$

$$= \sum_{i=1}^{l}\sum_{j=1}^{m}Y_{ij}^2 - 2\overline{Y}_{..}\sum_{i=1}^{l}\sum_{j=1}^{m}Y_{ij} + lm\overline{Y}_{..}^2$$

$$= \sum_{i=1}^{l}\sum_{j=1}^{m}Y_{ij}^2 - 2\frac{Y_{..}}{N}Y_{..} + N\left(\frac{Y_{..}}{N}\right)^2$$

$$= \sum_{i=1}^{l}\sum_{j=1}^{m}Y_{ij}^2 - 2\frac{Y_{..}^2}{N} + \frac{Y_{..}^2}{N}$$

$$= \sum_{i=1}^{l}\sum_{j=1}^{m}Y_{ij}^2 - \frac{Y_{..}^2}{N}$$

$$= \sum_{i=1}^{l}\sum_{j=1}^{m}Y_{ij}^2 - CT$$

② $SS_{tr} = \sum_{i=1}^{l}\sum_{j=1}^{m}(\overline{Y}_{i\cdot} - \overline{Y}_{..})^2$

$$= m\sum_{i=1}^{l}\left(\overline{Y}_{i\cdot} - \overline{Y}_{..}\right)^2$$

$$= m\sum_{i=1}^{l}\left(\overline{Y}_{i\cdot}^2 - 2\overline{Y}_{..}\overline{Y}_{i\cdot} + \overline{Y}_{..}^2\right)$$

$$= m\sum_{i=1}^{l}\overline{Y}_{i\cdot}^2 - 2m\overline{Y}_{..}\sum_{i=1}^{l}\overline{Y}_{i\cdot} + lm\overline{Y}_{..}^2$$

$$\because\ l\times\frac{\sum_{i=1}^{l}\overline{Y}_{i\cdot}}{l} = l\times\frac{\overline{Y}_{1\cdot} + \overline{Y}_{2\cdot} + \cdots + \overline{Y}_{l\cdot}}{l} = l\times\overline{Y}_{..}$$

$$= m\left(\sum_{i=1}^{l}\overline{Y^2}_{i\cdot} - 2m\overline{Y}_{..}\times l\overline{Y}_{..} + N\overline{Y^2}_{..}\right)$$

$$= m\left(\sum_{i=1}^{l}\overline{Y^2}_{i\cdot} - N\overline{Y^2}_{..}\right)$$

$$= m\sum_{i=1}^{l}\overline{Y^2}_{i\cdot} - lm\left(\frac{Y_{..}}{N}\right)^2$$

$$= \sum_{i=1}^{m}Y_{i\cdot}^2 - N\left(\frac{Y_{..}}{N}\right)^2$$

$$= \sum_{i=1}^{m}Y_{i\cdot}^2 - \frac{Y_{..}^2}{N}$$

3) 제곱합의 자유도

; $SS_T$의 자유도 $= SS_{tr}$의 자유도 + $SS_E$의 자유도

$\Rightarrow \ (N-1) \ = \ (l-1) \ + \ (N-l)$

$cf.$ $SS_E$의 자유도가 왜 $l(m-1)$일까?

$$SS_E = \sum_{j=1}^{l}\sum_{i=1}^{m}(Y_{ij} - \overline{Y}_{i.})^2$$

$$= \sum_{j=1}^{m}(Y_{1j} - \overline{Y}_{1.})^2 + \sum_{j=1}^{m}(Y_{2j} - \overline{Y}_{2.})^2 + \cdots + \sum_{j=1}^{m}(Y_{lj} - \overline{Y}_{l.})^2$$

따라서 자유도 : $(m-1) + (m-1) + \cdots + (m-1) = l(m-1) = N - l$

4. 평균제곱

1) $MS_{tr} = \dfrac{SS_{tr}}{l-1}$ : 처리 평균제곱(treatment MS)

2) $MS_E = \dfrac{SS_E}{N-l}$ : 오차평균제곱(error MS)

$MS_E$ 는 $\sigma^2$의 비편향 추정량이다(즉, $E(MS_E) = \sigma^2$)

5. 추정량

; $i$ 번째 처리의 표본 : $Y_{i1}$, $Y_{i2}$, $\cdots$, $Y_{il} \sim N(\mu_{i,} \ \sigma^2)$

1) $\hat{\mu_i} = \overline{Y}_{i\cdot} = \sum_{j=1}^{l}\dfrac{Y_{ij}}{m}$ : $\mu_i$의 비편향 추정량

2) $\hat{\sigma}^2 = s_i^2 = \dfrac{\sum_{j=1}^{m}(Y_{ij} - \overline{Y}_{i.})^2}{m-1}$ : $\sigma^2$의 비편향 추정량

3) 공통분산의 추정

$$s^2 = \frac{(m-1)s_1^2 + (m-1)s_2^2 + \cdots + (m-1)s_l^2}{(m-1)+(m-1)+\cdots+(m-1)}$$

$$= \frac{\sum_{i=1}^{l}(m-1)s_i^2}{l(m-1)}$$

$$= \frac{\sum_{i=1}^{l}\sum_{j=1}^{m}(Y_{ij} - \overline{Y}_{i.})^2}{N-l}$$

$$= \frac{SS_E}{N-l}$$

$$= MS_E$$

4) $MS_E$와 $MS_{tr}$의 기댓값

① $\overline{Y}_{i\cdot} = \frac{1}{m}\sum_{j=1}^{m} Y_{ij} = \frac{1}{m}\sum_{j=1}^{m}(\mu+\tau_i+\epsilon_{ij}) = \frac{1}{m}(m\mu+m\tau_i+\sum_{j=1}^{m}\epsilon_{ij})$

$= \mu+\tau_i+\overline{\epsilon}_{i\cdot}$

② $\overline{Y}_{\cdot\cdot} = \frac{1}{lm}\sum_{i=1}^{l}\sum_{j=1}^{m} Y_{ij} = \frac{1}{lm}\sum_{i=1}^{l}\sum_{j=1}^{m}(\mu+\tau_i+\epsilon_{ij}) = \frac{1}{lm}(lm\mu+m\sum_{i=1}^{l}\tau_i+\sum_{i=1}^{l}\sum_{j=1}^{m}\epsilon_{ij})$

$= \mu+\overline{\epsilon}_{\cdot\cdot}$

③ $E(MS_E) = E\left(\frac{\sum_{i=1}^{l}\sum_{j=1}^{m}\left(Y_{ij}-\overline{Y}_{i\cdot}\right)^2}{N-l}\right)$

$= \frac{1}{N-l}E\left(\sum_{i=1}^{l}\sum_{j=1}^{m}\left(\mu+\tau_i+\epsilon_{ij}-\mu-\tau_i-\overline{\epsilon}_{i\cdot}\right)^2\right)$

$= E\left(\frac{\sum_{i=1}^{l}\sum_{j=1}^{m}\left(\epsilon_{ij}-\overline{\epsilon}_{i\cdot}\right)^2}{N-l}\right)$

$= \sigma^2$

④ $E(MS_{tr}) = E\left(\frac{m\sum_{i=1}^{l}\left(\overline{Y}_{i\cdot}-\overline{Y}_{\cdot\cdot}\right)^2}{l-1}\right)$

$= \frac{1}{l-1}E\left(m\sum_{i=1}^{l}\left(\mu+\tau_i+\overline{\epsilon}_{i\cdot}-(\mu+\overline{\epsilon}_{\cdot\cdot})\right)^2\right)$

$= \frac{1}{l-1}E\left(m\sum_{i=1}^{l}\left(\tau_i+(\overline{\epsilon}_{i\cdot}-\overline{\epsilon}_{\cdot\cdot})\right)^2\right)$

$= \frac{1}{l-1}E\left(m\sum_{i=1}^{l}\left(\tau_i^2+2\tau_i(\overline{\epsilon}_{i\cdot}-\overline{\epsilon}_{\cdot\cdot})+(\overline{\epsilon}_{i\cdot}-\overline{\epsilon}_{\cdot\cdot})^2\right)\right)$

$= \frac{1}{l-1}\left[mE\left(\sum_{i=1}^{l}\tau_i^2\right)+2mE\left(\sum_{i=1}^{l}\tau_i(\overline{\epsilon}_{i\cdot}-\overline{\epsilon}_{\cdot\cdot})\right)+mE\left(\sum_{i=1}^{l}(\overline{\epsilon}_{i\cdot}-\overline{\epsilon}_{\cdot\cdot})^2\right)\right]$

$= \frac{1}{l-1}\left[m\sum_{i=1}^{l}\tau_i^2+2m\sum_{i=1}^{l}\tau_i E(\overline{\epsilon}_{i\cdot}-\overline{\epsilon}_{\cdot\cdot})+mE\left(\sum_{i=1}^{l}(\overline{\epsilon}_{i\cdot}-\overline{\epsilon}_{\cdot\cdot})^2\right)\right]$

$= \frac{1}{l-1}\left[m\sum_{i=1}^{l}\tau_i^2+mE\left(\sum_{i=1}^{l}(\overline{\epsilon}_{i\cdot}-\overline{\epsilon}_{\cdot\cdot})^2\right)\right]$ $(\because\ \overline{\epsilon}_{ij}\sim N(0,\ \sigma^2))$

$= m\sum_{i=1}^{l}\frac{\tau_i^2}{l-1}+mE\left(\frac{\sum_{i=1}^{l}(\overline{\epsilon}_{i\cdot}-\overline{\epsilon}_{\cdot\cdot})^2}{l-1}\right)$

$= m\sigma_\tau^2+m\times\frac{\sigma^2}{m}$ $(\because\ \overline{\epsilon}_{i\cdot}\sim N\left(0,\ \frac{\sigma^2}{m}\right))$

$= m\sigma_\tau^2+\sigma^2$

6. 분산분석 (분산비 검정 (variance-ratio test), F-검정)

1) 분산분석

; $l$ 개의 처리효과 사이에 차이가 없다는 귀무가설 $H_0 : \tau_1 = \tau_2 = \cdots = \tau_l = 0$을 검정하고자 한다. 물론 대립가설은 $\mu_i$들이 모두 같지는 않다는 것이다. 귀무가설을 검정하기 위한 규칙을 찾기 위해서, 모든 처리의 모평균이 같으면 $(\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot})$가 작을 것으로 기대된다. 따라서 처리의 평균제곱 $m\sum_{i=1}^{l}(\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot})^2/(m-1)$이 작게된다. 반면에 모평균들의 차이가 현저하면 처리의 평균제곱도 크게 될 것이다. 오차의 분산 $\sigma^2$의 추정량으로 잔차의 평균제곱이 이용되며, 이것은 처리의 평균제곱이 유의적 차이를 나타내기 전에 그것이 어느 정도 큰지를 결정하는 척도로 사용될 수 있다.

**분산분석**

a) 모형 $Y_{ij} = \mu + \tau_i + \epsilon_{ij}$ , $\epsilon_{ij} \sim N(0,\ \sigma^2)$ $(i=1,\ \cdots,\ l,\ j=1,\ \cdots,\ m)$

b) 가설 $H_0 : \tau_1 = \tau_2 = \cdots = \tau_l = 0$ (처리 효과 차이가 없다)

$H_1$ : not $H_0$ (그렇지 않다.)

b) 검정 통계량 $F = \dfrac{MS_{tr}}{MS_E} \sim F_{l-1, N-l}$

c) 기각역 $F > F_{l-1, N-l, \alpha}$

2) 분산분석표

| source | $SS$ | $df$ | $MS$ | $F$ |
|---|---|---|---|---|
| 처리(Treatments) | $SS_{tr}$ | $l-1$ | $MS_{tr}$ | $MS_{Trt}/MS_E$ |
| 오차(Error) | $SS_E$ | $N-l$ | $MS_E$ | |
| 합계(Total) | $SS_T$ | $N-1$ | | |

3) 처리 평균에 대한 추론

① $\mu_i$에 대한 $100(1-\alpha)\%$ 신뢰구간

; $\overline{Y}_{i\cdot}$ : $\mu_i$의 점추정량, $\overline{Y}_{i\cdot} \sim N(\mu_i,\ \sigma^2/m)$, $MS_E = s^2$ : $\sigma^2$에 대한 추정량

$\Rightarrow \overline{Y}_{i\cdot} \pm t_{N-l,\alpha/2}\dfrac{\sqrt{MS_E}}{\sqrt{m}}$

②$(\mu_i - \mu_j)$에 대한 $100(1-\alpha)\%$신뢰구간

$\Rightarrow (\overline{Y}_{i\cdot} - \overline{Y}_{j\cdot}) \pm t_{N-l,\alpha/2}\dfrac{\sqrt{2MS_E}}{\sqrt{m}}$ $(i \neq j)$

**예제** **2-1**. 어느 공장에서 제품을 생산하는데 열처리 온도에 따라서 제품의 강도가 차이를 보이는지 조사하기 위해 열처리 온도를 $A_1 = 125℃$, $A_2 = 150℃$, $A_3 = 175℃$, $A_4 = 200℃$로 변화시키고, 각 열처리 온도를 16개의 제품을 표본으로 추출하여 강도를 측정한 결과 다음의 자료를 얻었다.

| 125℃ | 150℃ | 175℃ | 200℃ |
|---|---|---|---|
| 12 | 18 | 15 | 39 |
| 15 | 22 | 26 | 46 |
| 16 | 26 | 43 | 49 |
| 17 | 30 | 20 | 38 |

다음 물음에 답하여라.

(a) 주어진 데이터에서 모든 관측값을 분할하여 $SS_T$, $SS_{tr}$, $SS_E$를 구하여라.

sol1) 제곱합

관측값$(Y_{ij}) = \begin{bmatrix} \\ \\ \\ \end{bmatrix}$ $\overline{Y_{1\cdot}} =$ , $\overline{Y_{2\cdot}} =$ , $\overline{Y_{3\cdot}} =$ , $\overline{Y_{4\cdot}} =$ , $\overline{Y_{\cdot\cdot}} =$

총편차$(Y_{ij} - \overline{Y_{\cdot\cdot}})$ = 처리효과$(\overline{Y_{i\cdot}} - \overline{Y_{\cdot\cdot}})$ + 잔차$(Y_{ij} - \overline{Y_{i\cdot}})$

$$\begin{bmatrix} \\ \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \\ \\ \end{bmatrix} + \begin{bmatrix} \\ \\ \\ \end{bmatrix}$$

$$\therefore \sum_{i=1}^{l}\sum_{j=1}^{m}(Y_{ij} - \overline{Y_{\cdot\cdot}})^2 = \sum_{i=1}^{l}\sum_{j=1}^{m}(\overline{Y_{i\cdot}} - \overline{Y_{\cdot\cdot}})^2 + \sum_{i=1}^{l}\sum_{j=1}^{m}(Y_{ij} - \overline{Y_{i\cdot}})^2$$

$SS_T =$

$SS_{tr} =$

$SS_E =$

sol2) 간단식

| | 125℃ | 150℃ | 175℃ | 200℃ |
|---|---|---|---|---|
| | 12 | 18 | 15 | 39 |
| | 15 | 22 | 26 | 46 |
| | 16 | 26 | 43 | 49 |
| | 17 | 30 | 20 | 38 |
| 합계 | | | | |
| 평균 | | | | |

$$CT = \frac{Y_{..}^2}{lm} =$$

$$SS_T = \sum_{i=1}^{l}\sum_{j=1}^{m} Y_{ij}^2 - CT$$

$$SS_{tr} = \sum_{i=1}^{l} \frac{Y_{i\cdot}^2}{m} - CT$$

$$SS_E = SS_T - SS_{tr}$$

(b) 분산분석표를 작성하라

| source | $SS$ | $df$ | $MS$ | $F$ |
|---|---|---|---|---|
| 처리(Treatments) | | | | |
| 오차(Error) | | | | |
| 합계(Total) | | | | |

(c) 위 data에 의하면 열처리 온도에 따라서 제품의 강도 사이에 유의적 차이가 존재 하는지를 $\alpha = 0.05$에서 검정하여라.

(d) 각 수준의 모평균의 $95\%$ 신뢰구간을 구하여라.

$$\overline{Y}_{i\cdot} \pm t_{N-l,\alpha/2}\frac{\sqrt{MS_E}}{\sqrt{m}}$$

(e) 모든 두 수준간의 모평균 차의 $95\%$ 신뢰구간을 구하여라.

$$(\overline{Y}_{i\cdot} - \overline{Y}_{j\cdot}) \pm t_{N-l,\alpha/2}\frac{\sqrt{2MS_E}}{\sqrt{l}}$$

(f) SPSS 분석결과 : 일원배치 분산분석

① 분산분석

**기술통계**

강도

| | N | 평균 | 표준편차 | 표준오차 | 평균에 대한 95% 신뢰구간 | | 최소값 | 최대값 |
|---|---|---|---|---|---|---|---|---|
| | | | | | 하한값 | 상한값 | | |
| 125℃ | 4 | 15.00 | 2.160 | 1.080 | 11.56 | 18.44 | 12 | 17 |
| 150℃ | 4 | 24.00 | 5.164 | 2.582 | 15.78 | 32.22 | 18 | 30 |
| 175℃ | 4 | 26.00 | 12.193 | 6.096 | 6.60 | 45.40 | 15 | 43 |
| 200℃ | 4 | 43.00 | 5.354 | 2.677 | 34.48 | 51.52 | 38 | 49 |
| 합계 | 16 | 27.00 | 12.291 | 3.073 | 20.45 | 33.55 | 12 | 49 |

**분산의 동질성 검정**

강도

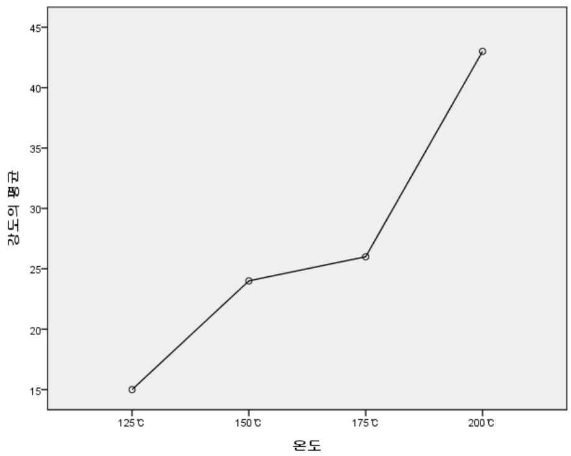
| Levene 통계량 | df1 | df2 | 유의확률 |
|---|---|---|---|
| 2.202 | 3 | 12 | .141 |

**분산분석**

강도

| | 제곱합 | df | 평균 제곱 | 거짓 | 유의확률 |
|---|---|---|---|---|---|
| 집단-간 | 1640.000 | 3 | 546.667 | 10.479 | .001 |
| 집단-내 | 626.000 | 12 | 52.167 | | |
| 합계 | 2266.000 | 15 | | | |

평균 도표



② 다중비교

**다중 비교**

종속 변수:강도

| | (I) 온도 | (J) 온도 | 평균차(I-J) | 표준오차 | 유의확률 | 95% 신뢰구간 하한값 | 95% 신뢰구간 상한값 |
|---|---|---|---|---|---|---|---|
| Tukey HSD | 125℃ | 150℃ | -9.00 | 5.107 | .337 | -24.16 | 6.16 |
| | | 175℃ | -11.00 | 5.107 | .192 | -26.16 | 4.16 |
| | | 200℃ | -28.00* | 5.107 | .001 | -43.16 | -12.84 |
| | 150℃ | 125℃ | 9.00 | 5.107 | .337 | -6.16 | 24.16 |
| | | 175℃ | -2.00 | 5.107 | .979 | -17.16 | 13.16 |
| | | 200℃ | -19.00* | 5.107 | .013 | -34.16 | -3.84 |
| | 175℃ | 125℃ | 11.00 | 5.107 | .192 | -4.16 | 26.16 |
| | | 150℃ | 2.00 | 5.107 | .979 | -13.16 | 17.16 |
| | | 200℃ | -17.00* | 5.107 | .027 | -32.16 | -1.84 |
| | 200℃ | 125℃ | 28.00* | 5.107 | .001 | 12.84 | 43.16 |
| | | 150℃ | 19.00* | 5.107 | .013 | 3.84 | 34.16 |
| | | 175℃ | 17.00* | 5.107 | .027 | 1.84 | 32.16 |
| LSD | 125℃ | 150℃ | -9.00 | 5.107 | .103 | -20.13 | 2.13 |
| | | 175℃ | -11.00 | 5.107 | .052 | -22.13 | .13 |
| | | 200℃ | -28.00* | 5.107 | .000 | -39.13 | -16.87 |
| | 150℃ | 125℃ | 9.00 | 5.107 | .103 | -2.13 | 20.13 |
| | | 175℃ | -2.00 | 5.107 | .702 | -13.13 | 9.13 |
| | | 200℃ | -19.00* | 5.107 | .003 | -30.13 | -7.87 |
| | 175℃ | 125℃ | 11.00 | 5.107 | .052 | -.13 | 22.13 |
| | | 150℃ | 2.00 | 5.107 | .702 | -9.13 | 13.13 |
| | | 200℃ | -17.00* | 5.107 | .006 | -28.13 | -5.87 |
| | 200℃ | 125℃ | 28.00* | 5.107 | .000 | 16.87 | 39.13 |
| | | 150℃ | 19.00* | 5.107 | .003 | 7.87 | 30.13 |
| | | 175℃ | 17.00* | 5.107 | .006 | 5.87 | 28.13 |

관측평균을 기준으로 합니다.
오류 조건은 평균 제곱(오류) = 52.167입니다.

*. 평균차는 0.05 수준에서 유의합니다.

**강도**

| | 온도 | N | 집단군 1 | 집단군 2 |
|---|---|---|---|---|
| Student-Newman-Keuls[a,b] | 125℃ | 4 | 15.00 | |
| | 150℃ | 4 | 24.00 | |
| | 175℃ | 4 | 26.00 | |
| | 200℃ | 4 | | 43.00 |
| | 유의확률 | | .120 | 1.000 |
| Tukey HSD[a,b] | 125℃ | 4 | 15.00 | |
| | 150℃ | 4 | 24.00 | |
| | 175℃ | 4 | 26.00 | |
| | 200℃ | 4 | | 43.00 |
| | 유의확률 | | .192 | 1.000 |

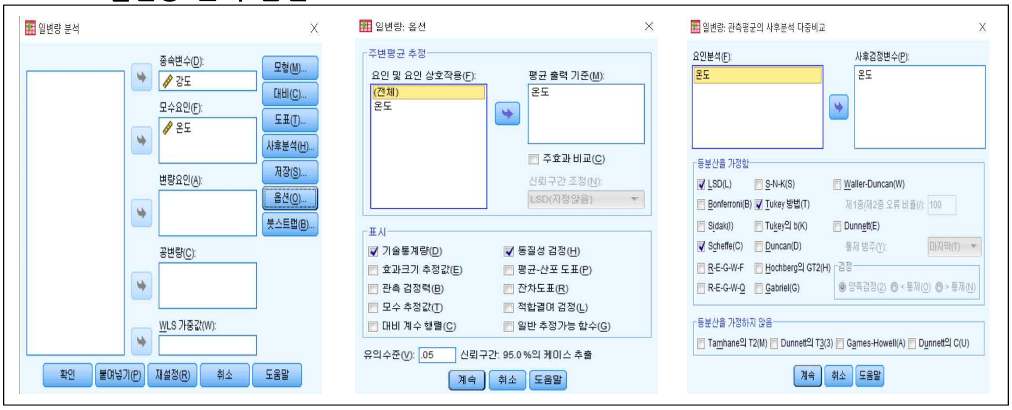
동일 집단군에 있는 집단에 대한 평균이 표시됩니다.
관측평균을 기준으로 합니다.
오류 조건은 평균 제곱(오류) = 52.167입니다.

a. 조화평균 표본 크기 4.000을(를) 사용합니다.

b. 유의수준 = 0.05.

(g) SPSS 분석결과 : 일변량분석

## SPSS 일변량 분석 옵션




① 분산분석

### 기술통계량

종속 변수:강도

| 온도 | 평균 | 표준편차 | N |
|---|---|---|---|
| 125℃ | 15.00 | 2.160 | 4 |
| 150℃ | 24.00 | 5.164 | 4 |
| 175℃ | 26.00 | 12.193 | 4 |
| 200℃ | 43.00 | 5.354 | 4 |
| 합계 | 27.00 | 12.291 | 16 |

### 오차 분산의 동일성에 대한 Levene의 검정[a]

종속 변수:강도

| F | df1 | df2 | 유의확률 |
|---|---|---|---|
| 2.202 | 3 | 12 | .141 |

여러 집단에서 종속변수의 오차 분산이 동일한 영가설을 검정합니다.

a. Design: 절편 + 온도

### 개체-간 효과 검정

종속 변수:강도

| 소스 | 제 III 유형 제곱합 | 자유도 | 평균 제곱 | F | 유의확률 |
|---|---|---|---|---|---|
| 수정 모형 | 1640.000[a] | 3 | 546.667 | 10.479 | .001 |
| 절편 | 11664.000 | 1 | 11664.000 | 223.591 | .000 |
| 온도 | 1640.000 | 3 | 546.667 | 10.479 | .001 |
| 오차 | 626.000 | 12 | 52.167 | | |
| 합계 | 13930.000 | 16 | | | |
| 수정 합계 | 2266.000 | 15 | | | |

a. R 제곱 = .724 (수정된 R 제곱 = .655)

**온도**

종속 변수:강도

| 온도 | 평균 | 표준오차 | 95% 신뢰구간 | |
|---|---|---|---|---|
| | | | 하한값 | 상한값 |
| 125℃ | 15.000 | 3.611 | 7.132 | 22.868 |
| 150℃ | 24.000 | 3.611 | 16.132 | 31.868 |
| 175℃ | 26.000 | 3.611 | 18.132 | 33.868 |
| 200℃ | 43.000 | 3.611 | 35.132 | 50.868 |

② 다중비교

**다중 비교**

종속 변수:강도

| | (I) 온도 | (J) 온도 | 평균차(I-J) | 표준오차 | 유의확률 | 95% 신뢰구간 | |
|---|---|---|---|---|---|---|---|
| | | | | | | 하한값 | 상한값 |
| Tukey HSD | 125℃ | 150℃ | -9.00 | 5.107 | .337 | -24.16 | 6.16 |
| | | 175℃ | -11.00 | 5.107 | .192 | -26.16 | 4.16 |
| | | 200℃ | -28.00* | 5.107 | .001 | -43.16 | -12.84 |
| | 150℃ | 125℃ | 9.00 | 5.107 | .337 | -6.16 | 24.16 |
| | | 175℃ | -2.00 | 5.107 | .979 | -17.16 | 13.16 |
| | | 200℃ | -19.00* | 5.107 | .013 | -34.16 | -3.84 |
| | 175℃ | 125℃ | 11.00 | 5.107 | .192 | -4.16 | 26.16 |
| | | 150℃ | 2.00 | 5.107 | .979 | -13.16 | 17.16 |
| | | 200℃ | -17.00* | 5.107 | .027 | -32.16 | -1.84 |
| | 200℃ | 125℃ | 28.00* | 5.107 | .001 | 12.84 | 43.16 |
| | | 150℃ | 19.00* | 5.107 | .013 | 3.84 | 34.16 |
| | | 175℃ | 17.00* | 5.107 | .027 | 1.84 | 32.16 |
| LSD | 125℃ | 150℃ | -9.00 | 5.107 | .103 | -20.13 | 2.13 |
| | | 175℃ | -11.00 | 5.107 | .052 | -22.13 | .13 |
| | | 200℃ | -28.00* | 5.107 | .000 | -39.13 | -16.87 |
| | 150℃ | 125℃ | 9.00 | 5.107 | .103 | -2.13 | 20.13 |
| | | 175℃ | -2.00 | 5.107 | .702 | -13.13 | 9.13 |
| | | 200℃ | -19.00* | 5.107 | .003 | -30.13 | -7.87 |
| | 175℃ | 125℃ | 11.00 | 5.107 | .052 | -.13 | 22.13 |
| | | 150℃ | 2.00 | 5.107 | .702 | -9.13 | 13.13 |
| | | 200℃ | -17.00* | 5.107 | .006 | -28.13 | -5.87 |
| | 200℃ | 125℃ | 28.00* | 5.107 | .000 | 16.87 | 39.13 |
| | | 150℃ | 19.00* | 5.107 | .003 | 7.87 | 30.13 |
| | | 175℃ | 17.00* | 5.107 | .006 | 5.87 | 28.13 |

관측평균을 기준으로 합니다.
오류 조건은 평균 제곱(오류) = 52.167입니다.

*. 평균차는 0.05 수준에서 유의합니다.

**강도**

| | 온도 | N | 집단군 1 | 집단군 2 |
|---|---|---|---|---|
| Student-Newman-Keuls[a,b] | 125℃ | 4 | 15.00 | |
| | 150℃ | 4 | 24.00 | |
| | 175℃ | 4 | 26.00 | |
| | 200℃ | 4 | | 43.00 |
| | 유의확률 | | .120 | 1.000 |
| Tukey HSD[a,b] | 125℃ | 4 | 15.00 | |
| | 150℃ | 4 | 24.00 | |
| | 175℃ | 4 | 26.00 | |
| | 200℃ | 4 | | 43.00 |
| | 유의확률 | | .192 | 1.000 |

동일 집단군에 있는 집단에 대한 평균이 표시됩니다.
관측평균을 기준으로 합니다.
오류 조건은 평균 제곱(오류) = 52.167입니다.

a. 조화평균 표본 크기 4.000을(를) 사용합니다.

b. 유의수준 = 0.05.

**예제** **2-2.** 국내 4개 회사에서 생산되는 무가당 오렌지 쥬스($A$, $B$, $C$, $D$)의 신 맛을 48명의 평가 요원이 쥬스마다 12명씩 랜덤 배치되어 9점 척도를 사용하여 평가한 결과이다.

| $A$ | $B$ | $C$ | $D$ |
|---|---|---|---|
| 8 8 | 6 5 | 8 8 | 9 4 |
| 7 9 | 7 7 | 8 5 | 6 5 |
| 8 7 | 6 7 | 6 7 | 4 4 |
| 9 8 | 6 8 | 6 7 | 5 6 |
| 6 8 | 7 6 | 7 6 | 7 5 |
| 7 6 | 8 7 | 7 8 | 8 4 |

(a) 분산분석표를 작성하라

| source | $SS$ | $df$ | $MS$ | $F$ |
|---|---|---|---|---|
| 처리(Treatments) | 24.896 | | | |
| 오차(Error) | | | | |
| 합계(Total) | 86.313 | | | |

(b) 위 data에 의하면 오렌지 쥬스에 따라서 신맛에 유의적 차이가 존재하는 지를 유의수준 5%에서 검정하여라.

(c) 각 수준의 모평균의 $95\%$ 신뢰구간을 구하여라.

(d) 모든 두 수준간의 모평균 차의 $90\%$ 신뢰구간을 구하여라.

**예제 2-3.** 주어진 자료는 돼지 체중증가에 가장 효율적인 사료 내 단백질 첨가 퍼센티지를 결정하려는 비교실험에서 얻어진 측정값들이다. 이 실험에서는 단백질 %가 처리이다. 동일 품종으로 나이와 무게가 비슷한 돼지 20마리를 동원하여, 각 처리별로 5마리씩 랜덤 배치한 다음 일정 기간 후에 돼지의 체중 증가분을 측정하였다.

| 10% | 20% | 30% | 40% |
|---|---|---|---|
| 60.8 | 87.9 | 102.6 | 68.7 |
| 57.0 | 84.2 | 102.1 | 67.7 |
| 65.0 | 83.1 | 100.2 | 74.0 |
| 58.6 | 85.7 | 96.5 | 66.3 |
| 61.7 | 90.3 | 99.9 | 69.8 |

(a) 다음은 SPSS 분석 결과이다. 단백질 %에 따른 체중 증가량에 차이가 있는 지를 유의수준 1%에서 검정하여라.

**기술통계량**

종속 변수:체중증가분

| 처리 | 평균 | 표준편차 | N |
|---|---|---|---|
| 10% | 60.620 | 3.0646 | 5 |
| 20% | 86.240 | 2.8962 | 5 |
| 30% | 100.260 | 2.4048 | 5 |
| 40% | 69.300 | 2.9266 | 5 |
| 합계 | 79.105 | 15.9108 | 20 |

**오차 분산의 동일성에 대한 Levene의 검정[a]**

종속 변수:체중증가분

| F | df1 | df2 | 유의확률 |
|---|---|---|---|
| .155 | 3 | 16 | .925 |

여러 집단에서 종속변수의 오차 분산이 동일한 영가설을 검정합니다.

a. Design: 절편 + 처리

**개체-간 효과 검정**

종속 변수:체중증가분

| 소스 | 제 III 유형 제곱합 | 자유도 | 평균 제곱 | F | 유의확률 |
|---|---|---|---|---|---|
| 수정 모형 | 4681.378[a] | 3 | 1560.459 | 194.280 | .000 |
| 절편 | 125152.020 | 1 | 125152.020 | 15581.676 | .000 |
| 처리 | 4681.377 | 3 | 1560.459 | 194.280 | .000 |
| 오차 | 128.512 | 16 | 8.032 | | |
| 합계 | 129961.910 | 20 | | | |
| 수정 합계 | 4809.889 | 19 | | | |

a. R 제곱 = .973 (수정된 R 제곱 = .968)

(b) 각 수준의 모평균의 $95\%$ 신뢰구간을 구하고 아래 SPSS 결과와 비교하여라.

**처리**

종속 변수:체중증가분

| 처리 | 평균 | 표준오차 | 95% 신뢰구간 하한값 | 95% 신뢰구간 상한값 |
|---|---|---|---|---|
| 10% | 60.620 | 1.267 | 57.933 | 63.307 |
| 20% | 86.240 | 1.267 | 83.553 | 88.927 |
| 30% | 100.260 | 1.267 | 97.573 | 102.947 |
| 40% | 69.300 | 1.267 | 66.613 | 71.987 |

예제 **2-4.** 다음 분석결과는 1990~1991 시즌의 잉글랜드, 이탈리아, 스페인, 독일, 프랑스의 경기당 평균 득점의 평균을 비교한 결과이다. 다음 물음에 답하여라.

경기당득점

| | N | 평균 | 표준편차 | 표준오차 | 평균에 대한 95% 신뢰구간 하한값 | 평균에 대한 95% 신뢰구간 상한값 | 최소값 | 최대값 |
|---|---|---|---|---|---|---|---|---|
| 잉글랜드 | 18 | 1,3545 | ,24796 | ,05844 | 1,2312 | 1,4778 | ,93 | 1,93 |
| 이탈리아 | 18 | 1,1356 | ,39255 | ,09252 | ,9404 | 1,3308 | ,71 | 2,18 |
| 스페인 | 18 | 1,2456 | ,41833 | ,09860 | 1,0376 | 1,4536 | ,79 | 2,29 |
| 독일 | 18 | 1,3304 | ,30745 | ,07247 | 1,1775 | 1,4833 | ,84 | 2,00 |
| 프랑스 | 18 | 1,0497 | ,28241 | ,06656 | ,9093 | 1,1901 | ,66 | 1,76 |
| 합계 | 90 | 1,2232 | ,34849 | ,03673 | 1,1502 | 1,2962 | ,66 | 2,29 |

경기당득점

| Levene 통계량 | 자유도1 | 자유도2 | 유의확률 |
|---|---|---|---|
| 1,107 | 4 | 85 | ,359 |

경기당득점

| | 제곱합 | 자유도 | 평균제곱 | F | 유의확률 |
|---|---|---|---|---|---|
| 집단-간 | 1,206 | 4 | ,302 | 2,669 | ,038 |
| 집단-내 | 9,603 | 85 | ,113 | | |
| 합계 | 10,809 | 89 | | | |

(a) 국가에 경기당 평균득점에 차이가 있는 지를 유의수준 5%에서 검정하여라.

(b) 결정계수를 구하고 설명하여라.

(b) 각 수준의 모평균의 $90\%$ 신뢰구간을 구하여라.

(c) 모든 두 수준간의 모평균 차의 $95\%$ 신뢰구간을 구하여라.

**예제** **2-5**. 어떤 직물의 가공시 처리액의 농도 $A$를 인자로 하여 $A_1 = 3.0\%$, $A_2 = 3.3\%$, $A_3 = 3.6\%$, $A_4 = 3.9\%$, $A_5 = 4.2\%$에서 반복 각 4회, 전체 20회를 랜덤하게 하여 처리한 후의 인장강도를 측정하여 다음 데이터를 얻었다.

| $A_1$ | $A_2$ | $A_3$ | $A_4$ | $A_5$ |
|---|---|---|---|---|
| 46.8 | 51.2 | 50.2 | 40.8 | 30.2 |
| 58.0 | 62.4 | 39.8 | 41.8 | 25.8 |
| 51.4 | 58.5 | 45.2 | 45.5 | 32.4 |
| 56.5 | 61.9 | 48.8 | 35.9 | 29.2 |

(a) 처리액의 농도에 따라 인장강도에 차이가 있는가를 유의수준 $1\%$에서 검정하여라.

| source | $SS$ | $df$ | $MS$ | $F$ |
|---|---|---|---|---|
| 처리(Treatments) | | | | |
| 오차(Error) | | | | |
| 합계(Total) | | | | |

(b) 결정계수를 구하고 설명하여라.

(c) $A$의 각 수준에서 인장강도의 모평균의 $95\%$ 신뢰구간을 구하여라.

**예제** **2-6**. 보일러 부식의 문제에 있어서는 연소생성가스 중의 $SO_3\%$가 문제가 된다. 유황분의 함유량에 차이가 있다고 믿어지는 4종류의 기름연료를 사용하였을 때에 연소생성가스 중의 $SO_3\%$를 측정한 결과 다음과 같은 데이터를 얻었다.

| $A_1$ | $A_2$ | $A_3$ | $A_4$ |
|---|---|---|---|
| 6.1 | 4.4 | 3.1 | 2.8 |
| 5.8 | 4.2 | 3.3 | 2.6 |
| 6.4 | 3.9 | 3.1 | 2.6 |
| | | 3.2 | 3.1 |
| | | 3.6 | 2.7 |
| | | 3.5 | 2.5 |

(a) 연료 기름의 종류가 다름에 따라 $SO_3\%$의 함량에 차이가 있다고 할 수 있는지 유의수준 1%에서 검정하여라.

| source | $SS$ | $df$ | $MS$ | $F$ |
|---|---|---|---|---|
| 처리(Treatments) | | | | |
| 오차(Error) | | | | |
| 합계(Total) | | | | |

(b) 각 연료기름의 종류에 따른 $SO_3$%의 95% 신뢰구간을 구하여라.

(c) 모든 두 수준간의 모평균 차의 95% 신뢰구간을 구하여라.

**예제** **2-7.** 사이프러스에서 발굴된 비잔틴시대에 주조된 동전의 은 함량(%)을 조사하여 주조시기별로 표와 같이 정리하였다. .

| 1기 | 2기 | 3기 | 4기 |
|---|---|---|---|
| 5.9 | 6.9 | 4.9 | 5.3 |
| 6.8 | 9.0 | 5.5 | 5.6 |
| 6.4 | 6.6 | 4.6 | 5.5 |
| 7.0 | 8.1 | 4.5 | 5.1 |
| 6.6 | 9.3 | | 6.2 |
| 7.7 | 9.2 | | 5.8 |
| 7.2 | 8.6 | | 5.8 |
| 6.9 | | | |
| 6.2 | | | |

(a) 다음은 SPSS 일원배치 분산분석 결과이다. 주조시기에 따른 은 함량(%)가 차이가 있는지를 유의수준 1%에서 검정하여라.

**기술통계량**

종속 변수:은함량

| 조사시기 | 평균 | 표준편차 | N |
|---|---|---|---|
| 1기 | 6.744 | .5434 | 9 |
| 2기 | 8.243 | 1.0998 | 7 |
| 3기 | 4.875 | .4500 | 4 |
| 4기 | 5.614 | .3625 | 7 |
| 합계 | 6.563 | 1.3695 | 27 |

**오차 분산의 동일성에 대한 Levene의 검정**[a]

종속 변수:은함량

| F | df1 | df2 | 유의확률 |
|---|---|---|---|
| 4.286 | 3 | 23 | .015 |

여러 집단에서 종속변수의 오차 분산이 동일한 영가설을 검정합니다.

a. Design: 절편 + 조사시기

**개체-간 효과 검정**

종속 변수:은함량

| 소스 | 제 III 유형 제곱합 | 자유도 | 평균 제곱 | F | 유의확률 |
|---|---|---|---|---|---|
| 수정 모형 | 37.748[a] | 3 | 12.583 | 26.272 | .000 |
| 절편 | 1003.449 | 1 | 1003.449 | 2095.181 | .000 |
| 조사시기 | 37.748 | 3 | 12.583 | 26.272 | .000 |
| 오차 | 11.015 | 23 | .479 | | |
| 합계 | 1211.720 | 27 | | | |
| 수정 합계 | 48.763 | 26 | | | |

(b) 결정계수를 구하고 설명하여라.

(c) 각 수준에서 은 함량의 모평균의 $99\%$ 신뢰구간을 구하고 아래 SPSS 결과와 비교하여라.

**조사시기**

종속 변수:은함량

| 조사시기 | 평균 | 표준오차 | 95% 신뢰구간 하한값 | 95% 신뢰구간 상한값 |
|---|---|---|---|---|
| 1기 | 6.744 | .231 | 6.267 | 7.222 |
| 2기 | 8.243 | .262 | 7.702 | 8.784 |
| 3기 | 4.875 | .346 | 4.159 | 5.591 |
| 4기 | 5.614 | .262 | 5.073 | 6.155 |