工業經營學會誌 第 20卷 第 44輯 1997年 11月 333

# 다중비교 절차를 이용한 제조공정의 분석

## Analysis of the Manufacturing Process using Multiple Comparison Procedure

최봉욱*
Bong-Wook Choi
김광섭**
Gwang-Sub Kim

**Abstract**

The purpose of this paper is to compare the manufacturing process with random covariate using multiple comparison procedure.

The methodology that compares each manufacturing process by inspecting the number of nonconforming items out of k-treatment, has serveral limitations and problems according to the method and contect of the analysis. The proper way of analysis, therefore, could be obtained by the multiple comparison procedure of simultaneous confidence region of variance components.

Effections that affect a manufactuing process may be predictive of responce to treatments are called covariates. In the study of comparing several treatments, prsense of covariate may bias the estimates of treatment effects.

## 1. 서론

일반적으로 비교하는 연구들은 실험으로부터 얻은 자료를 분석하는 방법이 집단들간의 유사성을 검정하는 것이다. 이러한 비교 연구들을 분산분석(ANOVA)이라 하는데, 이 방법은 k개의 처리(treatment)들의 유사성을 검정하기 때문에 집단들간의 다양하고 세밀한 비교(집단들간의 쌍별비교등)들에 관한 추론(inference)을 제공하지 못한다. 그러나 우리는 실제적으로 그런 차이를 보이는 처리들에 대해 알고 싶어하는 경우가 많다.

따라서 우리의 관심이 그런 k개의 처리들로부터 쌍별 비교에 관한 추론들이라면 그런 비교를 하기 위한 하나의 접근방법은 전통적으로 적절하다고 판단된 어떤 수준(level)에서의 적당한 과정(가설검정 혹은 신뢰구간 추정)에 의해 따로 따로 각 비교를 평가하는 것이었다.

그러나 이런 다중 t-검정은 다중효과(Tukey 1977)를 고려하지 않았다. 예를 들면 4개의 처리에 다중 t-검정을 시행하면 유의수준, 또는 제1종 오류를 0.05로 잡았다고 가정하면 이 경우 검정해야 할 가설들

* 아주대학교 산업공학과 박사과정
** 아주대학교 산업공학과 교수

은 4(4-1)/2=6개이다. 이 6개의 귀무가설에 대하여 각 검정마다 유의수준을 0.05로 놓는다면 이들 가설들이 모두 옳을 때 어느 하나의 귀무가설이라도 기각할 확률은 $1-(1-0.05)^6 \cong 0.14$가 되어 0.05보다 아주 큰 값이 되며, k가 커질수록 이런 확률은 점점 더 빨리 증가하여 1에 가깝게 된다. 오류율이 커진다는 것은 실험 데이터를 통하여 오판할 확률이 커짐을 뜻한다.

따라서 이러한 다중효과를 적절하게 조절하는 통계적인 절차들이 필요하게 된다. 이러한 통계적인 절차를 다중비교절차(MCP : multiple comparison procedure)라 부른다.

다중비교절차는 처리효과들 사이에 차이가 있는가를 비교하기 때문에, 처리효과를 추정하는 것이 중요하다. 일반적으로 반응변수의 평균을 가지고 처리효과를 추정한다. 그러나 어떤 원료에 대하여 같은 영향을 주었는데도 불구하고 다른 반응을 보일 수 있다. 이것은 어떤 다른 변수가 반응변수에 영향을 주었기 때문일 것이다. 이처럼 어떤 실험을 통하여 반응변수를 관측하다 보면, 하나 또는 그이상의 다른 변수들이 함께 관측된다. 이런 경우에 이 다른 변수가 반응변수에 어떤 교락효과를 작용하기 때문에, 이 다른 변수의 효과를 포함한 반응변수를 관측하게 된다. 따라서 이런 반응변수의 평균만을 가지고 추정을 한다면, 처리효과를 잘못 추정하게 될 것이다.

반응변수와 함께 관측되어, 반응변수에 영향을 주는 변수를 공변인(covariate)이라 한다. 공변인은 범주형 공변인(categorical covariate)과 연속형 공변인(quantitative covariate)의 형태로 나눌 수 있다. 범주형 공변인은 오차의 분산에 대한 추정량에 영향을 주고, 연속형 공변인은 처리효과에 대한 추정량과 오차의 분산에 대한 추정량에 영향을 준다. 따라서 반응변수에 작용하는 공변인의 교락효과를 통제 또는 조정하여야 한다.

이러한 목적을 달성하기 위한 방법으로써, 범주형 공변인은 실험단계에서 구획(blocking)기법으로 통제하고, 연속형 공변인은 공분산분석(analysis of covariance)을 사용해 조정할 수 있다.

본 연구에서는 자동차의 알루미늄 휠을 생산하는 제조기간의 비교를 위해서 다중비교 방법을 적용하였다. 다중비교 방법을 적용하기 전에 이러한 모형을 소개하고 공변인을 고려하지 않은 다중비교방법과 공변인을 고려한 다중비교방법을 비교하여 어떤 차이가 있는지를 분석하고자 한다.

## 2. 다중비교절차

### 2-1. 공변인을 고려한 모형

단순한 공분산 모형의 분석은 하나의 공변인을 갖는 일원 배치 설계법으로 구성되어 진다. Thigpen과 Paulson(1974)가 고려한 $X_{ij}$와 $Y_{ij}$는 공변인과 반응변수로 나타내고, $(X_{ij}, Y_{ij})'$는 평균벡터 $(\xi_i, \theta_i)$와 공분산 행렬

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \rho\sigma_{XY} \\ \rho\sigma_{XY} & \sigma_Y^2 \end{pmatrix}$$

을 갖는 2변량 정규분포를 따른다고 하자.

그리고 $\xi_1 = \xi_2 = \cdots = \xi_k = \xi$라 가정하면, 공변인 $X_{ij}$와 반응변수 $Y_{ij}$에 대한 조건부 모형은 다음과 같다.

$$Y_{ij} = \theta_i + \alpha_j + \beta(X_{ij} - \xi) + \varepsilon_{ij} \qquad (1 \le i \le k, 1 \le j \le n_i)$$

여기에서 $\theta_i$는 $i$번째 처리효과, $\alpha_j$는 $j$번째 구획효과, $\xi$는 공변인의 모평균이고,

$\beta$는 기울기이고, 그리고 $\varepsilon_{ij}$는 i.i.d. $N(0, \sigma^2)$이다.

Bryant와 Paulson(1976)의 일원배치 설계법보다 다른 하나 이상의 랜덤 공변인을 포함하는 더 일반적인 모형은 모든 랜덤 공변인들은 동일한 분포를 하고, 공변인

$\boldsymbol{X}_i = (X_{i1}, X_{i2}, \cdots, X_{iq})'$와 반응변수 $Y_i$는 벡터 $(\boldsymbol{X}_i', Y_i)'$로 측정되며, 평균베터 $(\boldsymbol{\xi}, \eta_i)$와 공분산 행렬

$$\Sigma = \begin{pmatrix} \Sigma_{XX} & \sigma'_{XY} \\ \sigma'_{XY} & \sigma^2_Y \end{pmatrix}$$

을 갖는 i.i.d.( $q$+1)변량 정규분포를 따른다.

여기서 $\xi$는 $q\times 1$, $\eta_i'$는 스칼라이며, $\Sigma_{XX}$는 $q\times q$, $\sigma_{XY}$는 $q\times 1$이다. 그 $\eta_i'$들은 $\dot{\mu}$: $r\times 1$이 미지의 매개변수이고, $A$: $N\times r$이 $i$번째열이 $a'_i$인 알고 있는 계획행렬인 선형 모형 $\eta = (\eta_1, \cdots, \eta_N)' = A\mu$으로 연계된다.

이와같이 $X_i$에 대한 조건적으로 $Y_i = a'_i \mu + (X_i - \xi')\beta + E_i$ $(1 \le i \le N)$로 쓸수 있다. 여기서 $\beta = \Sigma^{-1}{}_{XX}\ \sigma_{XY}$이고, $E_i'$는 i.i.d. $N(0, \sigma^2)$이다. 관심있는 매개변수는 $\theta_i = \dot{B}'_i \mu$이다.

2-2. 쌍비교를 위한 다중비교절차

2-2-1. 정확한 무조건부 절차

다음과 같이 $\widehat{\Xi} = (\widehat{\xi_1}, \widehat{\xi_2}, \cdots, \widehat{\xi_m})$이라 정의하자.

여기서

$$\widehat{\xi_j} = \frac{1}{\#_j} \sum_{i=1}^{N} \sigma_{ij} X_i \quad (1 \le j \le m) \tag{2.2.1}$$

$\#_j = \sum_{i=1}^{N} \delta_{ij}$ 이다.

따라서

$$\widehat{Z} = X - \Delta\widehat{\Xi}, \tag{2.2.2}$$

$$H = I - A(A'A)^{-}A', \tag{2.2.3}$$

$$S_{XX} = \widehat{Z}'H\widehat{Z}, \quad S_{XY} = \widehat{Z}'HY, \quad S_{YY} = \widehat{Y}'H\widehat{Y} \tag{2.2.4}$$

$(A'A)^{-}$는 $A'A$의 generalized inverse을 의미한다. 그리고

$$\nu = N - q - rank(A) \tag{2.2.5}$$

이며, 그때 $\beta$와 $\mu$의 LS 추정량은 각각 다음과 같이 주어진다.

$$\widehat{\beta} = S_{XX}{}^{-1}\ S_{XY}, \tag{2.2.6}$$

$$\widehat{\mu} = (A'A)^{-}A'(Y - \widehat{Z}\widehat{\beta}), \tag{2.2.7}$$

그리고 $\sigma^2$의 불편 추정량은

$$S^2 = \frac{1}{\nu}\{S_{YY} - S'_{XY}\ S^{-1}{}_{XX}\ S_{XY}\} \tag{2.2.8}$$

이 된다.

(2.2.7)을 이용하면 $\theta_i$의 LS 추정량은 다음과 같이 나타낼 수 있다.

$$\widehat{\theta_i} = b'_i \widehat{\mu} = b'_i (A'A)^{-}A'(Y - \widehat{Z}\widehat{\beta}) \quad (1 \le i \le k) \tag{2.2.9}$$

정확한 무조건부 절차의 적용을 위한 필수적인 균형조건은 다음과 같다.

(1) $B(A'A)^{-}B' = V = \{v_{ii'}\}$라 하자. 여기서 $B = (b_1, b_2, \cdots, b_k)'$이다.

첫 번째 균형조건은 그때

$$\nu_{ij} + \nu_{i'i'} - 2\mu_{ii'} = 2\nu_0 \quad (1 \le i \ne i' \le k) \tag{2.2.10}$$

이다.

(2) 두 번째 조건은

$$b'_i(A'A)^- A'\Delta = d' \quad (1 \le i \le k) \tag{2.2.11}$$

여기서 $d = (d_1, \cdots, d_m)'$는 상수이다.

(3) 세번째 조건은

$$\Delta = A\Gamma \tag{2.2.12}$$

인 상수 행렬 $\Gamma = \{\gamma_{ij}\}$이 존재하는 것이다.

이 세조건 하에 하나의 쌍비교에 대한 가설은 $H_{0j}: \theta_i = \theta_{i'}$ $vs$ $H_{1j}: \theta_i \neq \theta_{i'}$ $(1 \le i \neq i' \le k)$이다. 이러한 가설을 검정하기 위한 공변인 $X$의 무조건부 검정 통계량은 다음과 같다.

$$|T_{ii'}| = \frac{|\hat{\theta}_i - \hat{\theta}_{i'} - (\theta_i - \theta_{i'})|}{S\sqrt{v_0}} \quad (1 \le i \neq i' \le k) \tag{2.2.13}$$

여기서 $|T_{ii'}|$은 자유도가 $\nu$인 스튜던트 t-분포이다. 각각의 $H_{0j}$의 기각 영역은 $|T_{ii'}| > \xi_j$의 형태이다.

따라서 모든 쌍차이를 검정하기 위해 Union-Intersection방법을 사용하여 다음과 같은 가설을 쓸수 있다.

$$H_0 = \bigcap_{j=1}^{m} H_{0j} \ \ vs \ \ H_1 = \bigcup_{j=1}^{m} H_{1j} \quad m = \binom{k}{2} \tag{2.2.14}$$

임계값 $\xi_j$는 다음과 같이 결정된다.

$$\Pr\{|T_{ii'}| > \xi_j \ \text{for} \ some \ \ j = 1, 2, \cdots, m \mid H_0 \ ; \ true\} = \alpha \tag{2.2.15}$$

만약 $\xi_1 = \xi_2 = \cdots = \xi_k = \xi$이라면 (2.2.12)는 다음과 같다.

$$\Pr\{\max_{1 \le i \le i' \le k} |T_{ii'}| > \xi \mid H_0 \ ; \ true\} = \alpha \tag{2.2.16}$$

$H_0$하에서 모든 $|T_{ii'}|$ $(1 \le i \neq i' \le k)$들이 자유도 $\nu$인 $m$-변수 t-분포이다.

Bryant와 Bruvold(1980)은

$$\overline{Q}_{q,k,\nu} = \max_{1 \le i \le i' \le k} \frac{|\hat{\theta}_i - \hat{\theta}_{i'} - (\theta_i - \theta_{i'})|}{(v_0 S^2)^{1/2}}$$

의 $X_i$에 대한 무조건부 분포를 유도했다.

만약 $\overline{Q}^{(\alpha)}_{q,k,\nu}$를 $\overline{Q}_{q,k,\nu}$분포의 상위 $\alpha$점 이라 하면 (2.2.13)으로부터 $\theta_i$들 사이의 모든 쌍비교에 대한 $100(1-\alpha)\%$ 무조건부 동시신뢰구간은 다음과 같이 주어진다.

$$\theta_i - \theta_{i'} \in [\hat{\theta}_i - \hat{\theta}_{i'} \pm \overline{Q}^{(\alpha)}_{q,k,\nu} S\sqrt{v_0}] \quad (1 \le i \neq i' \le k) \tag{2.2.17}$$

2-2-2 근사적 조건부 절차

일반적으로 고정된 효과 선형 모형에서 쌍비교의 절차는 TK-절차가 이용된다. 사실 TK-절차의 응용 중 하나는 단일 고정된 공변인을 갖는 일원 배치 설계법 이었다. Hochberg와 Varon-Salomon(1984)는 랜덤 공변인의 관찰된 값들에 관한 조건부 공변인 설계분석에서 같은 절차를 이용하는 것을 제안했다.

(2.2.7)에 의해 주어진 추정량 $\hat{\mu}$는 평균벡터

$$E(\hat{\mu}|X) = (A'A)^- A'A\mu + (A'A)^- A'\Delta(\hat{\Xi} - \Xi)\beta \tag{2.2.18}$$

와 공분산 행렬

$$cov(\hat{\mu}|X) = \sigma^2\{(A'A)^- + (A'A)^- A'\hat{Z}(\hat{Z}'H\hat{Z})^- \hat{Z}'A(A'A)^-\}$$
$$= \sigma^2 R \tag{2.2.19}$$

를 갖는 다변량 정규이다.

이 추정량과 식 (2.2.9)와 (2.2.11)를 이용하면 $\hat{\theta}_i$들이

$$E(\hat{\theta}_i|X) = b_i'(A'A)^{-}A'A\mu + b_i'(A'A)^{-}A'\Delta(\hat{\Xi}-\Xi)\beta$$
$$= b_i'(A'A)^{-}A'A\mu + d'(\hat{\Xi}-\Xi)\beta \quad (1 \le i \ne i' \le k) \qquad (2.2.20)$$

그리고

$$cov(\hat{\theta}|X) = \sigma^2 BRB' = \sigma^2 W \qquad (2.2.21)$$

와 같다.

위와같이 조건부 $X$에 대하여 각

$$T_{ii'} = \frac{\hat{\theta}_i - \hat{\theta}_{i'} - (\theta_i - \theta_{i'})}{S\sqrt{w_{ii} + w_{i'i'} - 2w_{ii'}}} \quad (1 \le i \ne i' \le k) \qquad (2.2.22)$$

여기서 $T_{ii'}$는 자유도 $\nu$를 갖는 스튜던트 t-분포이며, $w_{ii'}$는 (2.2.21)에서 정의된 행렬 $W$의 원소이다.

각각의 $H_{0j}$의 기각 영역은 $|T_{ii'}| > \xi_j$의 형태이다. 따라서 모든 쌍차이를 검정하기 위해 Union-Intersection방법을 사용하여 다음과 같은 가설을 쓸수 있다.

$$H_0 = \bigcap_{j=1}^{m} H_{0j} \ vs \ H_1 = \bigcup_{j=1}^{m} H_{1j} \quad m = \binom{k}{2} \qquad (2.2.23)$$

임계값 $\xi_j$는 다음과 같이 결정된다.

$$\Pr\{|T_{ii'}| > \xi_j \ \text{for} \ some \ \ j = 1, 2, \cdots, m \mid H_0 \ ; \ true\} = \alpha \qquad (2.2.24)$$

만약 $\xi_1 = \xi_2 = \cdots = \xi_k = \xi$이라면 (2.2.12)는 다음과 같다.

$$\Pr\{\max_{1 \le i \le i' \le k} |T_{ii'}| > \xi \mid H_0 \ ; \ true\} = \alpha \qquad (2.2.25)$$

$H_0$하에서 모든 $|T_{ii'}|$ $(1 \le i \ne i' \le k)$들이 자유도 $\nu$인 $m$-변수 t-분포이다.

따라서 (2.2.25)을 이용하면 쌍비교 $\theta_i - \theta_{i'}$에 대한 $100(1-\alpha)\%$ 조건부 동시신뢰 구간은 다음과 같이 주어진다.

$$\theta_i - \theta_{i'} \in [\hat{\theta}_i - \hat{\theta}_{i'} \pm \ Q^{(\alpha)}_{k,\nu} S\sqrt{\frac{w_{ii} + w_{i'i'} - 2w_{ii'}}{2}}] \quad (1 \le i \ne i' \le k) \quad (2.2.26)$$

## 3. 수치예제

자동차 알루미늄 휠을 생산하는 제조공정에서, 제조기에서 동일한 규격의 제품을 생산하고 있다. 이 제조기에서는 용광로에서 녹인 원료를 온도에 의해서 압력을 높여서 주물하는 방식을 이용하고 있다. 여기에 반응변수에 작용하는 요인으로 용탕온도, 압력, 순도, 제조기의 불순물 함유량 그리고 총 시간 등이 있다. 반응변수는 가압의 정도에 관계가 있는 것으로 추측되며, 온도를 공변인으로 사용한다.

우리의 관심은 알루미늄 휠을 생산하는 제조기에 차이가 있는지를 알고 싶어하기 때문에 각각의 제조기가 처리가 되고, 또한 온도를 구획으로 간주한다. 알루미늄 휠을 생산하는 제조기는 20개 이지만 동일한 제품을 제조하며 온도들이 각 제조기에 다 사용되는 제조기들에 국한하여 다중비교를 실시하였다.

다음은 제조공정으로부터 얻은 자료이다.

표3-1 제조공정으로부터 얻은 공변인과 반응량

| 제조기 | 온도 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | A | | B | | C | | D | |
| | X | Y | X | Y | X | Y | X | Y |
| 1 | 1.18 | 3.3 | 1.14 | 4.8 | 1.12 | 3.4 | 1.16 | 3.8 |
| 2 | 1.14 | 3.0 | 1.15 | 4.1 | 1.17 | 1.2 | 1.16 | 5.6 |
| 3 | 1.16 | 0 | 1.19 | 1.4 | 1.16 | 1.5 | 1.15 | 0 |
| 4 | 1.22 | 0 | 1.18 | 1.1 | 1.21 | 1.0 | 1.21 | 0 |
| 5 | 1.10 | 0.5 | 1.12 | 1.4 | 1.14 | 1.8 | 1.11 | 0 |
| 6 | 1.17 | 8.6 | 1.20 | 2.8 | 1.21 | 4.3 | 1.20 | 3.4 |

위의 자료에 대하여 $S_{XX}=0.005929152$ $S_{XY}=-0.034407$ $S_{YY}=36.54667$이므로
$\hat{\beta}=S_{XY}/S_{XX}$ = -5.8030이다.

표3.2 보정된 처리들의 평균값( $\hat{\theta}_i = \overline{Y_i} - \hat{\beta}(\overline{X_i} - \overline{X_{..}})$)

| 제조기 | 1 | 2 | 3 | 4 | 5 | 6 | |
|---|---|---|---|---|---|---|---|
| $\overline{X_{i.}}$ | 1.15 | 1.155 | 1.165 | 1.205 | 1.1175 | 1.195 | $\overline{X_{..}}$ = 1.1646 |
| $\overline{Y_{i.}}$ | 3.825 | 3.475 | 0.725 | 0.525 | 0.925 | 4.775 | |
| $\hat{\theta}_i$ | 3.740 | 3.419 | 0.727 | 0.759 | 0.652 | 4.928 | |

먼저 공변인 $X$와 구획을 무시했을 때 반응변수 $Y$의 분산분석을 적용하면

표3.3 분산분석표(공변인 $X$와 구획을 무시했을 때)

| 요인 | 자유도 | 제곱합 | 평균제곱 |
|---|---|---|---|
| 처리 $\theta_i$ | 5 | 69.28 | 13.856 |
| 오차 $e_{ij}$ | 18 | 37.605 | 2.089167 |
| 총합 | 23 | 106.885 | |

위의 분산분석의 결과로부터 $F_{5,18}=6.63231 \geq F_{5,18}^{(0.05)}=2.77285$이므로 전체처리효과들이 같다는 가설은 기각이 된다.

표3.4 분산분석(공변인 $X$를 무시했을 때)

| 요인 | 자유도 | 제곱합 | 평균제곱 |
|---|---|---|---|
| 구획 $\alpha_i$ | 3 | 1.058333 | 0.352778 |
| 처리 $\theta_i$ | 5 | 69.28 | 13.856 |
| 오차 $e_{ij}$ | 15 | 36.54667 | 2.436444 |
| 총합 | 23 | 106.885 | |

위의 분산분석의 결과로부터 구획효과를 알아보기 위한 난괴법실험의 반응변수 $Y$만의 분산분석의 결과로부터 $F_{5,15}=5.681976 \geq F_{5,15}^{(0.05)}=2.901295$이므로 제조기들의 전체처리효과들이 같다는 가설은 기각이 된다.

따라서 모든 쌍차이 $\theta_i-\theta_{i'}$에 대한 공변인 $X$와 구획을 무시 했을 때와 공변인 $X$를 무시 했을때의 동시신뢰구간과 공변인 $X$와 구획을 고려했을 때 무조건부와 조건부의 95% 동시신뢰구간의 결과는 다음과 같다.

표3.5 공변인 $X$와 구획을 무시한 쌍차이의 동시신뢰구간

| $(i,i')$ | $\hat{\theta}_i - \hat{\theta}_{i'}$ ± 허용한계 | $(i,i')$ | $\hat{\theta}_i - \hat{\theta}_{i'}$ ± 허용한계 |
|---|---|---|---|
| (1,2) | 0.35 ± 3.53 | (2,6) | -1.3 ± 3.53 |
| (1,3) | 3.1 ± 3.53 | (3,4) | 0.2 ± 3.53 |
| (1,4) | 3.3 ± 3.53 | (3,5) | -0.2 ± 3.53 |
| (1,5) | 2.9 ± 3.53 | (3,6) | -4.05 ± 3.53 |
| (1,6) | -0.95 ± 3.53 | (4,5) | -0.4 ± 3.53 |
| (2,3) | 2.75 ± 3.53 | (4,6) | -4.25 ± 3.53 |
| (2,4) | 2.95 ± 3.53 | (5,6) | -3.85 ± 3.53 |
| (2,5) | 2.55 ± 3.53 | | |

표3.6 공변인 $X$을 무시한 쌍차이의 동시신뢰구간

| $(i,i')$ | $\hat{\theta}_i - \hat{\theta}_{i'}$ ± 허용한계 | $(i,i')$ | $\hat{\theta}_i - \hat{\theta}_{i'}$ ± 허용한계 |
|---|---|---|---|
| (1,2) | 0.35 ± 3.59 | (2,6) | -1.3 ± 3.59 |
| (1,3) | 3.1 ± 3.59 | (3,4) | 0.2 ± 3.59 |
| (1,4) | 3.3 ± 3.59 | (3,5) | -0.2 ± 3.59 |
| (1,5) | 2.9 ± 3.59 | (3,6) | -4.05 ± 3.59 |
| (1,6) | -0.95 ± 3.59 | (4,5) | -0.4 ± 3.59 |
| (2,3) | 2.75 ± 3.59 | (4,6) | -4.25 ± 3.59 |
| (2,4) | 2.95 ± 3.59 | (5,6) | -3.85 ± 3.59 |
| (2,5) | 2.55 ± 3.59 | | |

표3.7 공변인 $X$와 구획을 고려한 쌍차이의 무조건부 동시신뢰구간

| $(i,i')$ | $\hat{\theta}_i - \hat{\theta}_{i'}$ ± 허용한계 | $(i,i')$ | $\hat{\theta}_i - \hat{\theta}_{i'}$ ± 허용한계 |
|---|---|---|---|
| (1,2) | 0.32 ± 4.11 | (2,6) | -1.51 ± 4.11 |
| (1,3) | 3.01 ± 4.11 | (3,4) | -0.03 ± 4.11 |
| (1,4) | 2.98 ± 4.11 | (3,5) | 0.08 ± 4.11 |
| (1,5) | 3.09 ± 4.11 | (3,6) | -4.20 ± 4.11 |
| (1,6) | -1.19 ± 4.11 | (4,5) | 0.11 ± 4.11 |
| (2,3) | 2.69 ± 4.11 | (4,6) | -4.16 ± 4.11 |
| (2,4) | 2.66 ± 4.11 | (5,6) | -4.28 ± 4.11 |
| (2,5) | 2.77 ± 4.11 | | |

표3.8 공변인 $X$와 구획을 고려한 쌍차이의 조건부 동시신뢰구간

| $(i, i')$ $\hat{\theta}_i - \hat{\theta}_{i'}$ ± 허용한계 | $(i, i')$ $\hat{\theta}_i - \hat{\theta}_{i'}$ ± 허용한계 |
|---|---|
| (1,2) 0.32 ± 1.85 | (2,6) -1.51 ± 2.82 |
| (1,3) 3.01 ± 1.97 | (3,4) -0.03 ± 2.82 |
| (1,4) 2.98 ± 1.83 | (3,5) 0.08 ± 3.23 |
| (1,5) 3.09 ± 2.49 | (3,6) -4.20 ± 2.39 |
| (1,6) -1.19 ± 3.09 | (4,5) 0.11 ± 6.57 |
| (2,3) 2.69 ± 1.20 | (4,6) -4.16 ± 1.90 |
| (2,4) 2.66 ± 3.38 | (5,6) -4.28 ± 5.55 |
| (2,5) 2.77 ± 2.70 | |

위의 표3.5부터 표3.8까지는 95%의 유의수준에서 쌍별비교에 대한 다중비교절차의 신뢰구간을 구한것이며 임계값은 $Q_{k,\nu}^{(\alpha)}$를 적용하였다. 표3.5에서는 $Q_{6,18}^{(0.05)}$ = 4.50와 S=1.445, n=4이며, 허용한계는 $Q_{6,18}^{(0.05)}S/\sqrt{n}$ = 3.53이다. 표3.6에서는 $Q_{6,15}^{(0.05)}$ = 4.60, S = 1.5609, n=4이며, 허용한계는 $Q_{6,15}^{(0.05)}S/\sqrt{n}$ = 3.59이다. 표3.7에서는 임계값을 $\overline{Q_{1,6,14}^{(0.05)}}$을 사용하였다.

그리고 표3.8에서의 신뢰구간의 허용한계는

$$Q_{k,\nu}^{(\alpha)}S\sqrt{\frac{W_{ii}+W_{i'i'}-2W_{ii'}}{2}} = Q_{k,\nu}^{(\alpha)}S\sqrt{\frac{1}{2}\left[\frac{2}{n}+\frac{(\overline{X}_i-\overline{X}_{i'})^2}{S_{XX}}\right]} \quad (1 \le i < i' \le k)\text{이다.}$$

## 4. 결론

표3.5와 표3.6은 공변인을 고려하지 않은 경우에 다중비교방법을 적용하여 얻은 동시신뢰구간이며 표3.7과 표3.8은 공변인을 고려한 경우에 다중비교절차를 적용하여 얻은 동시신뢰구간이다. 표3.6과 표3.7을 비교해보면 먼저 처리효과의 차이에 대한 추정량에서 차이가 있다. 표3.6의 처리효과에 대한 추정량은 공변인의 교락효과를 조정하지 않았기 때문에 잘못된 추정량일 가능성이 높다. 표3.6의 결과로부터 같은 처리효과를 갖는 집단으로 (1, 2, 3, 4, 5), (6)로 분류할 수 있겠으나, 처리효과에 대한 추정값이 잘못되었을 가능성이 대단히 높기 때문에 무의미할 수 있을 것이다. 그러므로 표3.7로부터, 같은 처리효과를 갖는 집단을 분류하는 것이 옳은 판단이 될 가능성이 높다.

처리효과에 대한 다중비교 방법을 적용하는 이유는 분산분석에서 귀무가설이 기각되었을 때 처리들 간에 차이가 있으므로 차이가 있는 처리들간의 집단을 찾아주기 위해서 이다. 무조건부 다중비교절차와 조건부 다중비교절차를 보면 무조건부 절차는 일정한 동시신뢰구간을 갖는데 비해 조건부절차는 다른 동시신뢰구간을 갖는다. 표3.7과 표3.8에서보면 표3.8이 유의한 차이를 많이 찾아주고 신뢰구간의 길이도 대체로 짧은 구간을 얻었다. 이 예제에서는 처리효과가 다른 것을 많이 찾아주는 다중비교절차인 조건부 다중비교절차를 이용하는 것이 바람직하다.

## 참고문헌

[1] Bryant, J. L. and Bruvold, N. T. (1980), "Multiple comparison procedures in the analysis of covariance.", *J. Amer. Statist. Assoc.*, **75**, 874-880.

[2] Bryant, J. L. and Fox, G. E. 91985), "Some comments on a class of simultianeous procedures in ANCOVA", *Commun. Statist., Sre. A*, **14**, 2511-2530.

[3] Bryant, J. L. and Paulson, A. S. (1976), "An extension of Tukey's method of multiple comparisons to experimental designs with random concomitant variables", *Biometrika*, **63**, 631-638.

[4] Hochberg, Y. and Varon-Salomon, Y. (1984), "On simultaneous pairwise comparisons in analysis of covariance", *J. Amer. Statist. Assoc.*, **79**, 863-866.

[5] Kramer, C. Y. (1957), "Extension of multiple range tests to group correlated adjusted means", *Biometrics*, **13**, 13-18.

[6] Thigpen, C. C. and Paulson, A. S. (1974), "A multiple range test for analysis of covariance", *Biometrika*, **61**, 479-484.

[7] Tukey, J. L. (1953), *The Problem of multiple Comparison*, Mimeographed monograph.

[8] Tukey, J. W. (1977) , "Some thoughts on clinical trials, especially problems of multiplicity", *Science*, **198**, 679-684.