

다중 비교 방법

여러 표준 편차에 대한 그래픽 다중 비교 절차

Senin J. Banga & Gregory D. Fox

2013년 6월 18일

개요

k 개 표준 편차의 다중 비교를 위해 그래픽을 사용하는 새로운 절차를 제공합니다. 새 절차는 분산 동질성 검정으로서, Levene(1960) 검정의 Brown & Forsythe(1974) 버전인 W_{50} 과 제1종 및 제2종 오차 속성이 유사합니다. 그러나 다중 비교와 연관된 그래픽 표시는 표준 편차가 서로 다른 표본을 선별하는 데 유용한 시각적인 도구를 제공합니다.

색인 용어: 등분산, Levene의 검정, Brown-Forsythe 검정, Layard의 검정, 다중 비교

1. 소개

일반적으로 검정 W_{50} 이라 불리는 Brown & Forsythe(1974)가 수정한 Levene 검정(1960)은 분산 동질성 검정에 가장 널리 사용되는 절차 중 하나일 것입니다. 검정 W_{50} 의 인기가 높은 부분적인 이유는 로버스트하고 점근적 분포가 없기 때문입니다. 검정 W_{50} 은 다른 분산 동질성 검정에 비해 계산하기도 더 쉽습니다. (해당 검정의 비교에 대해서는 Conover 외(1981)를 참조하십시오.) 검정 W_{50} 은 또한 SAS, Minitab, R 및 JMP 같은 여러 통계 소프트웨어 패키지에서도 제공되기 때문에 쉽게 이용할 수 있습니다.

그러나 일부 분포에 대한 검정 W_{50} 의 검정력은 특히 표본이 작을 때 매우 낮을 수 있습니다. 예를 들어 Pan(1999)은 검정 W_{50} 의 검정이 정규 분포를 포함한 일부 분포에 대해 두 표준 편차 간 차이의 크기에 관계 없이 해당 차이를 탐지하기에 충분하지 않을 수 있음을 보여줍니다. Pan의 분석에서는 같은 한계가 다중 표본 설계에도 적용될지 분명히 알 수 없습니다. 이 한계는 표본이 3개 이상인 설계에 적용되지 않을 것이라고 기대할 수 있는데, 그 이유는 해당 설계에 두 표본 설계보다 더 많은 데이터가 포함될 가능성이 높기 때문입니다. 검정 W_{50} 은 대표본 속성이 우수한 것으로 알려져 있습니다(Miller, 1968; Brown & Forsythe, 1974; Conover 외, 1981).

유의 검정 W_{50} 에 이어 Bonferroni 다중성 수정에 기초한 동시 쌍별 비교 절차를 실시하는 방법은 일반적인 관행이 되었습니다. 그러나 Pan(1999)이 지적했듯이 이런 방법은 검정 W_{50} 의 검정력이 두 표본 설계에서 낮기 때문에 실패하거나 잘못된 결과를 가져올 가능성이 큽니다. Bonferroni 수정을 사용하면 특히 쌍별 비교의 수가 많을 때 보수적이기 때문에 문제가 악화됩니다. 반대로, 일원 분산 분석 후에 평균을 비교하기 위해 사용할 수 있는 여러

효과적인 다중 비교 절차가 있습니다. 예는 Tukey (1953), Hochberg 외(1982) 및 Stoline(1981)을 참조하십시오. 표본 분산의 상호 비교를 위한 유사한 사후 분석이 유용할 것입니다.

본 문서에서는 그래픽을 사용하여 다중 표본의 분산(또는 표준 편차)을 비교하는 방식을 제안합니다. 이 분석은 Hochberg 외(1982)에서 설명한 평균에 대한 불확실 구간과 유사한 분산에 대한 "불확실 구간"에 기반을 두고 있습니다. 첫째, 다중 쌍별 비교 절차는 Bonett(2006)이 수정한 Layard(1973)의 두 표본 설계 분산 동질성 검정에 기반을 두고 있습니다. 쌍별 비교에 사용된 다중성 수정은 Nakayama(2009)가 제안한 Tukey-Kramer 방법(Tukey, 1953; Kramer, 1956)의 대표본 일반화에 기반을 두고 있습니다. 당사에서 "다중 비교 구간" 또는 "MC 구간"이라고 지칭하는 불확실성 구간은 Hochberg 외(1982)에서 설명한 최적 근사 절차를 사용하는 쌍별 비교 절차로부터 도출되었습니다. 그 결과로 얻어지는 MC 검정은 MC 구간의 한 쌍 이상이 겹치지 않을 경우에만 귀무 가설을 기각합니다. 겹치지 않는 MC 구간은 분산(또는 표준 편차)이 유의하게 다른 표본을 식별합니다.

MC 검정의 소표본 속성을 평가하기 위한 시뮬레이션 연구를 실시합니다. 비교를 위해 W_{50} 검정도 시뮬레이션 연구에 포함시킵니다.

2. 그래픽 다중 비교 절차

$Y_{i1}, \dots, Y_{in_i}, \dots, Y_{k1}, \dots, Y_{kn_k}$ 는 k 개의 독립 표본이고, 각 표본은 독립적이고 평균이 $E(Y_{il}) = \mu_i$ 고 분산이 $\text{Var}(Y_{il}) = \sigma_i^2 > 0$ 인 동일한 분포를 따릅니다. 또한 표본이 공통 첨도가 $\gamma = E(Y - \mu)^4 / \sigma^4 < \infty$ 인 모집단에서 추출되었다고 가정합니다.

또한 \bar{Y}_i 와 S_i 는 각각 표본 i 의 평균과 표준 편차입니다. m_i 는 절사 비율이 $1/[2\sqrt{n_i - 4}]$ 인 표본 i 의 절사 평균이고, $\hat{\gamma}_{ij}$ 는 다음 식으로 표현되는 표본 (i, j) 의 합동 첨도 추정량입니다.

$$\begin{aligned} \hat{\gamma}_{ij} &= (n_i + n_j) \frac{\sum_{l=1}^{n_i} (Y_{il} - m_i)^4 + \sum_{l=1}^{n_j} (Y_{jl} - m_j)^4}{\left[\sum_{l=1}^{n_i} (Y_{il} - \bar{Y}_i)^2 + \sum_{l=1}^{n_j} (Y_{jl} - \bar{Y}_j)^2 \right]^2} \\ &= (n_i + n_j) \frac{\sum_{l=1}^{n_i} (Y_{il} - m_i)^4 + \sum_{l=1}^{n_j} (Y_{jl} - m_j)^4}{[(n_i - 1)S_i^2 + (n_j - 1)S_j^2]^2} \end{aligned}$$

$\hat{\gamma}_{ij}$ 는 표본 평균 \bar{Y}_i 를 절사 평균 m_i 로 대체한 Layard(1973)의 합동 첨도 추정량과 점근적으로 동일합니다. 따라서 $\hat{\gamma}_{ij}$ 는 모분산이 같을 경우 알 수 없는 공통 첨도 γ 의 일정한 추정량입니다. Bonett(2006)은 Layard의 합동 첨도 추정량 대신에 이 추정량을 제안하여 두 표본 문제에서 Layard 검정의 소표본 성능을 개선합니다. 본 문서에서는 Bonett(2006)이 수정한 Layard 검정을 간단히 Bonett 검정이라고 계속 지칭합니다.

비교할 독립 집단 또는 표본이 3개 이상 있다고 가정합니다($k > 2$). 여기서 제안하는 그래픽 다중 비교 절차는 Bonett 검정에 기초한 다중 쌍별 비교에서 도출되었습니다. 대신, 쌍별 비교를 검정 W_{50} 에 기초하는 방법도 있습니다. 그러나 두 표본 설계에서 검정 W_{50} 의 검정력 성능은 정규 분포를 포함한 일부 분포에 대해 문제가 됩니다(Pan, 1999). 또한 Banga & Fox(2013)에 따르면 Bonett 검정에 기초한 분산 비율에 대한 신뢰 구간이 일반적으로 검정 W_{50} 에 기초한 것보다 우월합니다.

주어진 표본 쌍 (i, j) 에 대해 유의 수준이 α' 인 양측 Bonett 검정은 다음과 같은 경우에만 분산의 동질성 귀무 가설을 기각합니다.

$$|\ln(c_i S_i^2) - \ln(c_j S_j^2)| > z_{\alpha'/2} \sqrt{\frac{\hat{y}_{ij} - k_i}{n_i - 1} + \frac{\hat{y}_{ij} - k_j}{n_j - 1}}$$

여기서 $z_{\alpha'/2}$ 는 표준 정규 분포의 $\alpha'/2 \times 100$ 번째 상위 백분위수입니다.

$$k_i = \frac{n_i - 3}{n_i}, k_j = \frac{n_j - 3}{n_j}, c_i = \frac{n_i}{n_i - z_{\alpha'/2}}, c_j = \frac{n_j}{n_j - z_{\alpha'/2}}$$

여러 쌍별 비교, 정확히는 $k(k-1)/2$ 개의 비교가 있으므로, 다중성 조정이 필요합니다. 예를 들어 전체 또는 모임별 목표 유의 수준 α 가 주어진 경우에 사용되는 일반적인 방법 중 하나인 Bonferroni 수정에서는 각 $k(k-1)/2$ 쌍별 비교의 유의 수준 $\alpha' = 2\alpha/(k(k-1))$ 을 선택합니다. 그러나 Bonferroni 수정은 비교할 표본 수가 증가할수록 점점 보수적인 쌍별 비교 절차를 생성하는 것으로 잘 알려져 있습니다. 더 나은 대안은 Tukey-Kramer 방법(Tukey, 1953; Kramer, 1956)의 대표본 근사를 토대로 Nakayama(2009)가 제안한 방법입니다. 구체적으로 말해, 전체 다중 쌍별 비교 검정은 일부 표본 쌍 (i, j) 에 대해 다음이 참일 경우에만 유의합니다.

$$|\ln(c_i S_i^2) - \ln(c_j S_j^2)| > \frac{q_{\alpha, k}}{\sqrt{2}} \sqrt{\frac{\hat{y}_{ij} - k_i}{n_i - 1} + \frac{\hat{y}_{ij} - k_j}{n_j - 1}}$$

여기서 $q_{\alpha, k}$ 는 동일한 분포를 따르는 k 개의 독립적인 표준 정규 랜덤 변수 범위의 상위 α 점입니다. 즉, $q_{\alpha, k}$ 는 다음을 충족합니다.

$$\Pr\left(\max_{1 \leq i < j \leq k} |Z_i - Z_j| \leq q_{\alpha, k}\right) = 1 - \alpha$$

여기서 Z_1, \dots, Z_k 는 동일한 분포를 따르는 독립적인 표준 정규 랜덤 변수입니다.

Barnard(1978)는 정규 범위의 분포 함수를 계산하는 16점 가우시안 구적법에 기초한 간단한 숫자 알고리즘을 제시합니다.

Hochberg 외(1982)의 주장처럼, 위에서 설명한 다중 쌍별 비교 절차를 근사하는 그래픽 다중 비교 절차에서는 다음과 같은 경우에만 귀무 가설이 기각될 것입니다.

$$|\ln(c_i S_i^2) - \ln(c_j S_j^2)| > q_{\alpha, k}(V_i + V_j)/\sqrt{2}$$

여기서 V_i 는 다음을 최소화하도록 선택합니다.

$$\sum_{i \neq j} \sum (V_i + V_j - b_{ij})^2$$

설명

$$b_{ij} = \sqrt{\frac{\hat{y}_{ij} - k_i}{n_i - 1} + \frac{\hat{y}_{ij} - k_j}{n_j - 1}}$$

이 문제에 대한 해법은 Hochberg 외(1982)의 설명대로 다음을 선택하는 것입니다.

$$V_i = \frac{(k-1) \sum_{j \neq i} b_{ij} - \sum \sum_{1 \leq j < l \leq k} b_{jl}}{(k-1)(k-2)}$$

따라서 이 근사 절차에 기초한 분산 동질성 검정에서는 아래 제시된 구간 중 한 쌍 이상이 겹치지 않을 경우에만 귀무 가설을 기각합니다.

$$\left[S_i \sqrt{c_i \exp(-q_{\alpha,k} V_i / \sqrt{2})}, S_i \sqrt{c_i \exp(q_{\alpha,k} V_i / \sqrt{2})} \right], i = 1, \dots, k$$

그래픽 MC 절차에서는 해당 구간을 그래프에 표시하여 구간이 겹치지 않는 표본을 시각적으로 식별합니다. 또한 전체 분산 동질성(또는 표준 편차) 검정의 p-값도 산정할 수 있습니다. p-값을 계산하는 자세한 알고리즘은 다음 항목에 나와 있습니다. 그러나 먼저 MC 절차에 대한 몇 가지 간단한 사실에 대해 알아봅니다.

설명

1. 표본 쌍 (i, j) 에 기초한 합동 침도 추정량 $\hat{\nu}_{ij}$ 는 모든 k 표본에 기초한 전체 합동 침도 추정량으로 대체할 수 있었습니다. 이 방법을 사용하면 계산이 어느 정도 단순화되지만, 여기에 나와 있지 않은 시뮬레이션 결과에서는 $\hat{\nu}_{ij}$ 를 사용하여 더 나은 결과를 얻을 수 있는 것으로 나타났습니다.
2. 표본 i 에 해당되는 구간은 표본 모집단의 표준 편차에 대한 신뢰 구간이 아닙니다. Hochberg 외(1982)에서는 해당 구간을 "불확실성 구간"이라고 지칭합니다. 당사는 이를 "다중 비교 구간" 또는 "MC 구간"이라고 합니다. MC 구간은 다중 표본 설계의 표준 편차 또는 분산을 비교하는 데만 유용합니다.
3. 본 문서에서 설명하는 MC 구간은 3개 이상의 표준 편차를 비교하는 데만 사용할 수 있습니다. 표본이 2개뿐인 경우에도 비교 구간을 구성할 수 있지만, 검정 결과를 통해 얻어지는 것과 동일한 정보만 얻어집니다. Banga & Fox(2013)가 설명하고 Minitab의 두 표본 분산 명령과 함께 제공되는 것과 같은 표준 편차 비율에 대한 신뢰 구간을 구성하면 훨씬 더 유용한 정보를 얻을 수 있습니다.

3. 그래픽 다중 비교 방법의 p-값

그래픽 (MC) 방법의 p-값을 계산하는 알고리즘에 대해 설명하기 전에, 먼저 Bonnett(2006)이 수정한 Layard 검정과 연관된 p-값을 두 표본 설계에서 도출합니다. 그런 다음, 두 표본 설계 결과를 다중 비교 절차에 적용하는 방법을 제시합니다.

3.1 두 표본 설계의 p-값

앞에서 언급했듯이 Bonnett(2006)이 조정한 Layard 검정은 두 표본 설계에서 다음과 같은 경우에만 분산의 동질성 귀무 가설을 기각합니다.

$$|\ln(c_1 S_1^2) - \ln(c_2 S_2^2)| > z_{\alpha/2} se$$

이는 다음과 동일합니다.

$$|\ln(c_{\alpha/2} S_1^2 / S_2^2)| > z_{\alpha/2} se$$

설명

$$se = \sqrt{\frac{\hat{\nu}_{12} - k_1}{n_1 - 1} + \frac{\hat{\nu}_{12} - k_2}{n_2 - 1}}$$

$$c_{\alpha/2} = \frac{c_1}{c_2} = \frac{n_1}{n_1 - z_{\alpha/2}} \frac{n_2 - z_{\alpha/2}}{n_2}$$

Bonett은 소표본 불균형 설계에서 비등 꼬리 오차 확률의 영향을 완화하기 위한 소규모 표본 조정으로서 상수 $c_{\alpha/2}$ 를 도입했습니다. 그러나 대표본 불균형 설계에서 상수의 영향은 미미하며, 균형 설계에는 상수가 영향을 미치지 않습니다.

따라서 균형 설계인 경우 분산 동질성에 대한 양측 검정의 p-값은 다음과 같이 계산됩니다.

$$P = 2 \Pr(Z > |Z_0|)$$

설명

$$Z_0 = \frac{\ln(S_1^2) - \ln(S_2^2)}{se}$$

설계가 불균형이면 $P = 2 \min(\alpha_L, \alpha_U)$ 입니다. 여기서 α_L 은 다음 등식에서 α 의 가장 작은 해입니다.

$$\exp[\ln(c_\alpha S_1^2/S_2^2) - z_\alpha se] = 1 \quad (1)$$

그리고 α_U 는 다음 등식에서 α 의 가장 작은 해입니다.

$$\exp[\ln(c_\alpha S_1^2/S_2^2) + z_\alpha se] = 1 \quad (2)$$

아래에는 α_L 및 α_U 를 구하는 알고리즘이 나와 있습니다. 알고리즘의 수학적 세부 사항은 부록 부분을 참조하십시오.

$$L(z, n_1, n_2, S_1, S_2) = \ln \frac{n_1}{n_2} + \ln \frac{n_2 - z}{n_1 - z} - z se + \ln \frac{S_1^2}{S_2^2}, z < \min(n_1, n_2)$$

이고

$$z_m = \frac{n_1 + n_2 - \sqrt{(n_1 - n_2)(n_1 - n_2 - \frac{4}{se})}}{2}$$

일 때 해 α_L 과 α_U 는 다음과 같은 절차에 따라 계산됩니다.

사례 1: $n_1 < n_2$

- 위에서 주어진 z_m 을 계산하고 $L(z_m, n_1, n_2, S_1, S_2)$ 를 평가합니다.
- $L(z_m) \leq 0$ 이면 $L(z, n_1, n_2, S_1, S_2)$ 의 근 z_L 을 구간 $(-\infty, z_m]$ 에서 찾고 $\alpha_L = \Pr(Z > z_L)$ 을 계산합니다.
- $L(z_m) > 0$ 이면 함수 $L(z, n_1, n_2, S_1, S_2)$ 에 근이 없습니다. $\alpha_L = 0.0$ 으로 설정합니다.

사례 2: $n_1 > n_2$

- $L(0, n_1, n_2, S_1, S_2) = \ln S_1^2/S_2^2$ 을 계산합니다.
- $L(0, n_1, n_2, S_1, S_2) \geq 0$ 이면 $L(z, n_1, n_2, S_1, S_2)$ 의 근 z_0 를 구간 $[0, n_2)$ 에서 찾고, 그렇지 않으면 근 z_L 을 구간 $(-\infty, 0)$ 에서 찾습니다.
- $\alpha_L = \Pr(Z > z_L)$ 을 계산합니다.

α_U 를 계산하려면 위의 절차를 적용하고 $L(z, n_1, n_2, S_1, S_2)$ 함수 대신 $L(z, n_2, n_1, S_2, S_1)$ 함수를 사용합니다.

3.2 그래픽 다중 비교의 p-값

설계에 k ($k > 2$)개의 표본이 있다고 가정하고, P_{ij} 는 표본 쌍 (i, j) 와 연관된 검정의 p-값이라고 합니다. 다중 비교 검정에서는 k 비교 구간 중 한 쌍 이상이 겹치지 않을 경우에만 분산의 동질성 귀무 가설을 기각한다는 것을 기억하십시오. 따라서 다중 비교 절차와 연관된 전체 p-값은 논리적으로 다음과 같습니다.

$$P = \min\{P_{ij}, 1 \leq i < j \leq k\}$$

P_{ij} 를 계산하기 위해 다음을 사용하여 두 표본 설계의 알고리즘을 실행합니다.

$$se = V_i + V_j$$

여기서 V_i 는 전과 같이 정의합니다.

$n_i \neq n_j$ 면

$$P_{ij} = \min(\alpha_L, \alpha_U)$$

입니다. 여기서 $\alpha_L = \Pr(Q > z_L\sqrt{2})$, $\alpha_U = \Pr(Q > z_U\sqrt{2})$, z_L 은 $L(z, n_i, n_j, S_i, S_j)$ 함수의 가장 작은 근이고, z_U 는 $L(z, n_j, n_i, S_j, S_i)$ 함수의 가장 작은 근이고, Q 는 앞에서 정의한 랜덤 변수입니다. 수량 z_L 및 z_U 는 앞에서 설명한 두 표본 설계 알고리즘을 표본 쌍 (i, j) 에 적용하여 구합니다.

$n_i = n_j$ 면 $P_{ij} = \Pr(Q > |z_o|\sqrt{2})$ 입니다. 설명:

$$z_o = \frac{\ln S_i^2 - \ln S_j^2}{V_i + V_j}$$

4. 시뮬레이션 연구 및 결과

전체적인 분산 동질성 검정으로서 MC 검정의 소표본 성능을 조사하기 위한 두 가지의 주요 시뮬레이션 연구를 실시합니다. 모든 시뮬레이션은 Mathematica 소프트웨어 패키지 버전 8을 사용하여 실시했습니다.

연구 1

첫 번째 연구는 MC 검정과 검정 W_{50} 의 제1종 오차 속성을 평가 및 비교하기 위해 고안되었습니다. 세 가지 설계, 즉 3-표본 설계, 4-표본 설계 및 6-표본 설계의 다양한 분포에서 생성된 표본을 사용한 두 검정의 성능을 비교합니다. 각 설계에서 표본 크기를 10개부터 50개까지 10개 단위로 변경합니다. 표본은 다음 모분포에서 추출합니다.

- 정규 분포
- 균등 분포와 모수가 (3,3)인 베타 분포로 대표되는 대칭 얇은 꼬리 분포
- 자유도가 5인 t 분포($t(5)$)와 Laplace 분포로 대표되는 대칭 두꺼운 꼬리 분포
- 지수 분포와 자유도가 1인 카이 제곱 분포($\chi^2(1)$) 및 자유도가 5인 카이 제곱 분포($\chi^2(5)$)로 대표되는 치우친 두꺼운 꼬리 분포
- 관측치의 90%가 표준 정규 분포에서 추출되고 나머지 10%는 평균이 0이고 표준 편차가 3인 정규 분포에서 추출된 오염 정규 분포(CN(0.9, 3)).

각 시뮬레이션은 10,000개의 표본 추출 반복실험으로 구성됩니다. 목표 명목 α 수준은 0.05입니다. 시뮬레이션 오차는 약 0.002입니다. 각 검정에 대한 모의 유의 수준은 표 1에 보고되어 있습니다.

표 1 모의 유의 수준 ($\alpha = 0.05$)의 비교

설명	분포 [첨도]	n_i	$k = 3$		$k = 4$		$k = 6$	
			MC	W_{50}	MC	W_{50}	MC	W_{50}
정규 분포	정규 분포 [3.0]	10	.038	.033	.038	.031	.036	.029
		20	.039	.038	.040	.038	.041	.033
		30	.043	.041	.044	.038	.046	.039
		40	.046	.043	.046	.041	.048	.041
		50	.046	.046	.046	.044	.052	.047
대칭 얇은 꼬리 분포	균등 분포 [1.8]	10	.029	.029	.025	.024	.023	.020
		20	.028	.026	.030	.026	.028	.023
		30	.037	.035	.034	.032	.034	.030
		40	.038	.037	.037	.037	.035	.033
		50	.041	.041	.036	.036	.036	.036
	베타(3, 3) [2.5]	10	.031	.032	.031	.029	.031	.025
		20	.035	.031	.036	.027	.037	.026
		30	.041	.035	.037	.034	.037	.032
		40	.040	.036	.039	.035	.040	.033
		50	.044	.039	.044	.037	.044	.035
대칭 두꺼운 꼬리 분포	Laplace 분포 [6.0]	10	.056	.038	.063	.041	.071	.039
		20	.054	.044	.058	.043	.059	.041
		30	.051	.042	.053	.043	.052	.044
		40	.048	.045	.048	.045	.048	.046
		50	.045	.045	.051	.046	.049	.047
	$t(5)$ [9.0]	10	.042	.032	.044	.031	.042	.031
		20	.043	.039	.045	.038	.045	.040
		30	.039	.040	.040	.040	.041	.040
		40	.041	.042	.040	.041	.039	.038
		50	.040	.050	.039	.046	.038	.046

설명	분포 [첨도]	n_i	$k = 3$		$k = 4$		$k = 6$	
			MC	W_{50}	MC	W_{50}	MC	W_{50}
치우친 두꺼운 꼬리 분포	$\chi^2(5)$ [5.4]	10	.040	.039	.046	.040	.048	.039
		20	.040	.043	.040	.040	.042	.039
		30	.039	.047	.042	.044	.043	.042
		40	.040	.046	.041	.044	.039	.042
		50	.037	.047	.038	.047	.040	.048
	지수 분포 [9.0]	10	.063	.051	.073	.049	.076	.048
		20	.051	.049	.053	.048	.057	.046
		30	.042	.048	.046	.051	.049	.049
		40	.034	.050	.038	.046	.037	.049
		50	.033	.045	.037	.047	.038	.046
	$\chi^2(1)$ [15.0]	10	.084	.048	.098	.050	.118	.050
		20	.053	.046	.060	.047	.068	.046
		30	.041	.041	.045	.045	.050	.047
		40	.044	.049	.046	.047	.045	.047
		50	.038	.050	.037	.049	.040	.049
오염 정규 분포	CN(0.9, 3) [8.3]	10	.020	.016	.018	.012	.016	.010
		20	.014	.015	.012	.013	.008	.007
		30	.012	.014	.010	.011	.007	.008
		40	.009	.017	.009	.014	.006	.008
		50	.009	.016	.007	.012	.006	.009

결과에 따르면 두 검정 모두 대부분의 분포에 효과적인 것으로 나타났습니다. 모의 유의 수준은 대부분 목표인 0.05에 가깝습니다. 그러나 모의 유의 수준은 두 검정에서 모두 정규 분포와 가벼운 꼬리 대칭 분포에서 작은 표본을 추출할 때 보수적(0.05 미만)인 경향이 있습니다. 이들 분포에서 MC 검정의 모의 유의 수준은 검정 W_{50} 보다 목표 유의 수준에 더 가깝습니다.

두꺼운 꼬리 분포에서 작은 표본을 추출할 때 검정 W_{50} 은 보수적인 경향이 있고 MC 검정은 개방적인 경향이 있습니다. MC 검정은 매우 치우친 분포에서 작은 표본을 추출할 때 더욱 개방적입니다. 예를 들어 자유도가 1인 카이 제곱 분포에서 크기가 10인 표본을 추출하면 MC 검정의 모의 유의 수준은 3, 4 및 6 표본 설계에 대해 각각 0.084, 0.098 및 0.118입니다.

두 검정은 모두 특이치에 영향을 받습니다. 오염 정규 분포에 대한 유의 수준은 표본 크기가 클 때(50개)도 매우 보수적입니다.

연구 2

두 번째 연구에서는 4표본 설계에서 2개 절차의 제2종 오차 속성(검정력)을 평가 및 비교합니다. 이 연구에서는 연구 1의 표본 크기 20과 $k = 4$ 조건에 사용된 것과 동일한 표본을 사용합니다. 관측치는 1, 2, 3 또는 4 배수로 확대합니다. 예를 들어 1:1:4:4로 표시된 조건에서 표본 1과 2는 연구 1에 사용된 것과 동일합니다. 표본 3과 4의 관측치는 4배수로 확대합니다.

비교를 위해 조건 1:1:1:1을 포함시킵니다. 이 조건에 대한 결과는 연구 1에서 표본 크기 20과 $k = 4$ 에 대해 보고된 결과와 동일합니다. 표본 크기를 20개로 선택한 이유는 연구 1의 결과에 따라 두 검정에서 모두 표본 크기가 20일 때 대부분의 분포에 대해 목표 수준에 가까운 유의 수준이 얻어진 것으로 나타났기 때문입니다.

이들 실험의 모의 검정력 수준은 분산의 동질성 귀무 가설의 기각으로 이어지는 표본 반복실험의 비율로 계산됩니다.

결과는 표 2에 보고되어 있습니다.

표 2 모의 검정력 수준 ($\alpha = 0.05$)의 비교

설명	분포	표준 편차 비율							
		1:1:1:1		1:1:2:2		1:2:3:4		1:1:4:4	
		MC	W_{50}	MC	W_{50}	MC	W_{50}	MC	W_{50}
	정규 분포	.040	.038	.846	.853	.998	.994	1.00	1.00
대칭 얇은 꼬리 분포	균등 분포	.030	.026	.985	.962	1.00	.999	1.00	1.00
	베타(3, 3)	.036	.027	.938	.916	1.00	.999	1.00	1.00
대칭 두꺼운 꼬리 분포	Laplace 분포	.058	.043	.597	.629	.931	.921	.996	.998
	$t(5)$.045	.038	.657	.703	.952	.949	.997	.998
치우친 두꺼운 꼬리 분포	$\chi^2(5)$.040	.040	.625	.704	.949	.949	.996	.999
	지수 분포	.053	.048	.431	.507	.804	.779	.963	.978
	$\chi^2(1)$.060	.047	.298	.291	.602	.504	.838	.824
오염 분포	CN(0.9, 3)	.012	.013	.499	.612	.889	.917	.989	.998

결과는 MC 검정과 검정 W_{50} 의 제2종 오차 속성(검정력)이 유사함을 시사합니다.

일반적으로, 두 검정을 사용하여 얻은 모의 검정력 수준은 자릿수(order of magnitude)가 같습니다. 두 검정의 검정력이 0.1보다 많이 차이나는 경우는 하나 뿐입니다.

MC 검정의 모의 검정력 수준은 꼬리가 얇거나 중간인 대칭 분포에서 표본을 추출할 때 검정 W_{50} 보다 약간 더 낮습니다. 반면에, 검정 W_{50} 은 표본을 두꺼운 꼬리 분포에서 추출할 때 MC 검정보다 약간 더 강력한 것으로 보입니다.

5. 예제

이어지는 내용에서는 그래픽 MC 절차와 검정 W_{50} 을 Ott 외(2010) 397페이지에서 얻은 데이터 집합에 적용합니다. 데이터에 대한 설명은 다음과 같습니다.

한 주조 회사에는 원료를 밀납형 안에 붓기 전에 가열하는 오븐이 여러 개 있습니다. 이런 금속을 편차가 거의 없는 정밀한 온도로 가열하는 것이 매우 중요합니다. 오븐 3개를 무작위로 선정하고, 10회 연속 가열하면서 각 오븐의 온도(°C)를 매우 정확하게 기록합니다. 수집된 데이터는 다음과 같습니다.

오븐 1	1670.87	1670.88	1671.51	1672.01	1669.63	1670.95	1668.70	1671.86	1669.12	1672.52
오븐 2	1669.16	1669.60	1669.76	1669.18	1671.92	1669.69	1669.45	1669.35	1671.89	1673.45
오븐 3	1673.08	1672.75	1675.14	1674.94	1671.33	1660.38	1679.94	1660.51	1668.78	1664.32

그림 1에는 각 오븐 온도의 상자 그림이 나와 있습니다. 상자 그림에서는 기록된 온도에 특이치가 없고 오븐 3의 온도 변동이 오븐 1 또는 오븐 2와 다름을 알 수 있습니다.

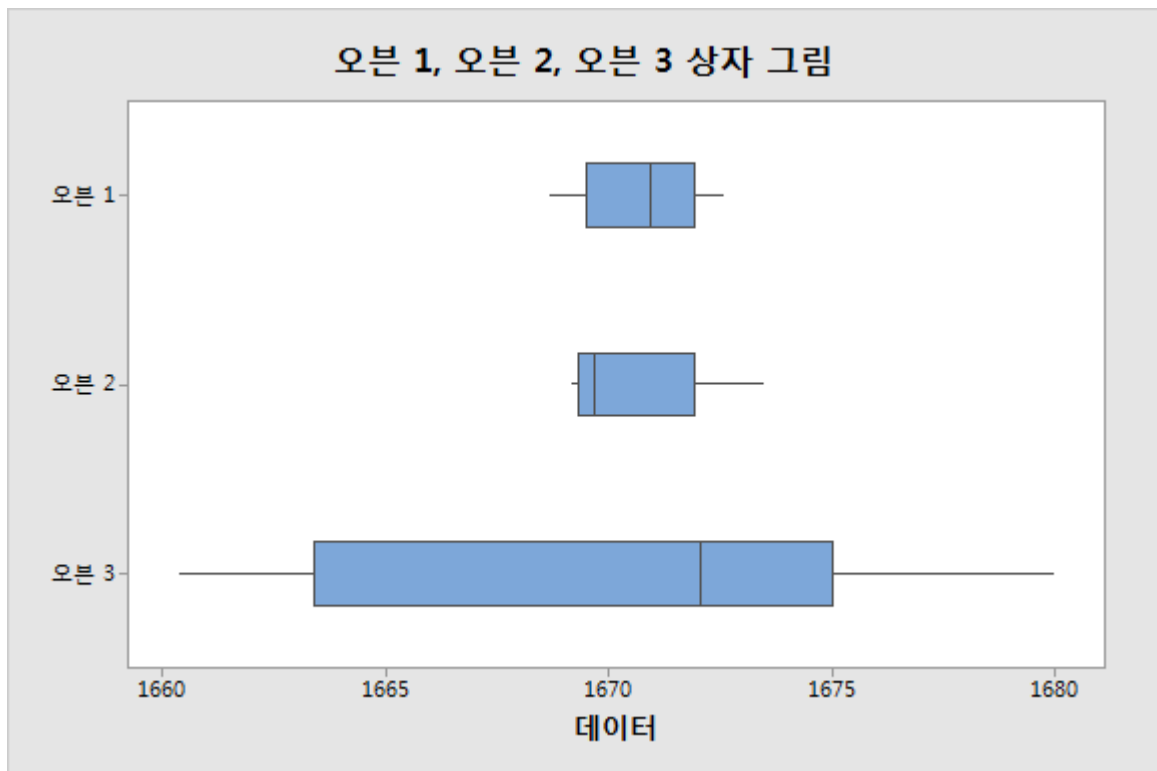


그림 1 오븐 온도(°C)의 상자 그림

그림 2에는 동일한 데이터에 대한 MC 구간과 범례에 Levene 검정이라고 나와 있는 전체 MC 검정 및 검정 W_{50} 의 결과가 나와 있습니다. 두 검정 모두에 대한 유의 p-값은 세 오븐의 온도 변동이 각각 다름을 나타냅니다. 겹치지 않는 MC 구간에 의해 오븐 3의 변동이 오븐 2 또는 오븐 1과 다름을 확인할 수 있습니다. MC 구간은 오븐 1, 2, 3에 대해 각각 (0.896, 2.378), (1.072, 2.760), (4.366, 12.787)입니다.

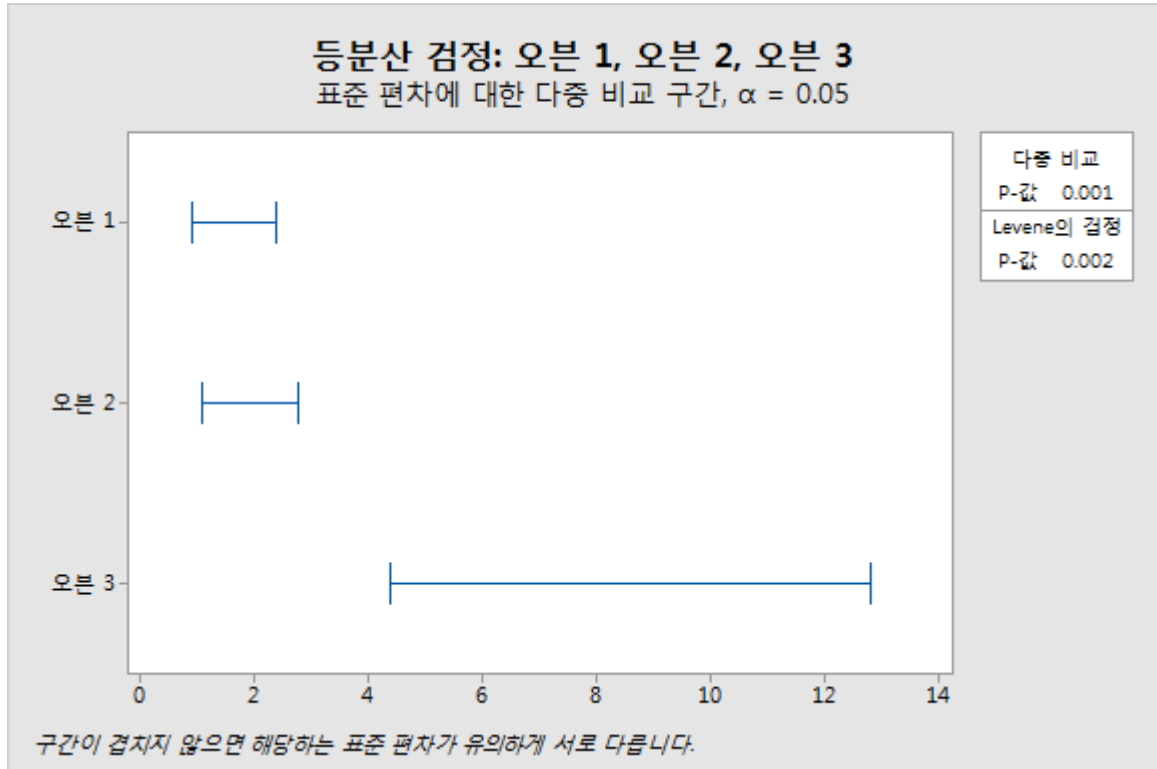


그림 2 MC 검정 및 검정 W_{50} (Levene 검정)에 대한 MC 구간과 p-값

6. 결론

전체적으로, 시뮬레이션 결과에서는 다중 소규모 표본 설계에서 MC 검정의 성능이 검정 W_{50} 과 유사한 것으로 나타났습니다. MC 검정은 꼬리가 얇거나 중간인 대칭 또는 근대칭 분포에 약간 더 적합한 반면, 검정 W_{50} 은 데이터를 매우 치우친 분포와 꼬리가 두꺼운 분포에서 추출할 때 선호될 수 있습니다. MC 절차의 분명한 장점 중 하나는 전체 표준 편차 동질성 검정이 유의할 때 표본 편차 또는 분산이 서로 다른 표본을 선별하는 강력한 시각적 도구를 제공한다는 데 있습니다. 그래픽 MC 절차는 Minitab 릴리스 17에서 제공됩니다.

7. 부록

Bonett(2006)이 조정한 Layard 검정은 두 표본 설계에서 다음과 같은 경우에만 분산의 동질성 귀무 가설을 기각합니다.

$$|\ln(c_1 S_1^2) - \ln(c_2 S_2^2)| > z_{\alpha/2} se$$

이는 다음과 동일합니다.

$$|\ln(c_{\alpha/2} S_1^2 / S_2^2)| > z_{\alpha/2} se$$

설명

$$se = \sqrt{\frac{\hat{\gamma}_{12} - k_1}{n_1 - 1} + \frac{\hat{\gamma}_{12} - k_2}{n_2 - 1}}$$

$$c_{\alpha/2} = \frac{c_1}{c_2} = \frac{n_1}{n_1 - z_{\alpha/2}} \frac{n_2 - z_{\alpha/2}}{n_2}$$

따라서 균형 설계인 경우 $c_{\alpha/2} = 1$ 이므로 검정의 p-값은 다음과 같습니다.

$$P = 2 \Pr(Z > |Z_0|)$$

설명

$$Z_0 = \frac{\ln(S_1^2) - \ln(S_2^2)}{se}$$

불균형 설계인 경우 $P = 2 \min(\alpha_L, \alpha_U)$ 입니다. 여기서,

α_L 은 다음 등식에서 α 의 가장 작은 해입니다.

$$\exp[\ln(c_{\alpha} S_1^2 / S_2^2) - z_{\alpha} se] = 1 \quad (1)$$

그리고 α_U 는 다음 등식의 가장 작은 해 α 입니다.

$$\exp[\ln(c_{\alpha} S_1^2 / S_2^2) + z_{\alpha} se] = 1 \quad (2)$$

이들 등식에서 α 의 해를 구하는 방법은 우선 $z \equiv z_{\alpha}$ 에 대해 등식의 해를 구한 다음 랜덤 변수 Z 가 표준 정규 분포를 따른다고 할 때 $\alpha = \Pr(Z > z)$ 를 구하는 것입니다. 위 등식의 해를 구하는 방법에 대해 설명하기 전에, (1)번 등식을 $L(z) = 0$ 등식으로 다시 표현할 수 있음을 주지하십시오. 여기서,

$$L(z, n_1, n_2, S_1, S_2) = \ln \frac{n_1}{n_2} + \ln \frac{n_2 - z}{n_1 - z} - z se + \ln \frac{S_1^2}{S_2^2}, z < \min(n_1, n_2)$$

입니다. 마찬가지로, (2)번 등식은 $U(z) = 0$ 등식과 동등합니다. 여기서,

$$U(z, n_1, n_2, S_1, S_2) = \ln \frac{n_1}{n_2} + \ln \frac{n_2 - z}{n_1 - z} + z se + \ln \frac{S_1^2}{S_2^2}, z < \min(n_1, n_2)$$

입니다. 그리고 $L(z, n_2, n_1, S_2, S_1) = -U(z, n_1, n_2, S_1, S_2)$ 입니다. 결과적으로 두 함수 중 하나의 근만 구해야 합니다.

(1)번 또는 (2)번 등식의 해를 구하는 알고리즘은 다음과 같은 결과로부터 도출됩니다.

결과

n_1, n_2, S_1 및 S_2 이 주어졌고 고정일 때, 불균형 설계의 경우 함수 $L(z, n_1, n_2, S_1, S_2)$ 의 근은 최대 2개입니다.

4. $n_1 < n_2$ 면 $L(z, n_1, n_2, S_1, S_2)$ 는 볼록 함수입니다. 이 함수는 $L(-\infty, n_1, n_2, S_1, S_2) = L(n_1, n_1, n_2, S_1, S_2) = +\infty$ 를 충족하며, 다음과 같을 때 최소값에 도달합니다.

$$z_m = \frac{n_1 + n_2 - \sqrt{(n_1 - n_2)(n_1 - n_2 - \frac{4}{se})}}{2}$$

따라서 $L(z_m, n_1, n_2, S_1, S_2) \leq 0$ 이면 근이 2개입니다. 하나는 $(-\infty, z_m]$ 구간 안에 있고, 다른 하나는 $[z_m, n_1)$ 구간 안에 있습니다. 반면에 $L(z_m, n_1, n_2, S_1, S_2) > 0$ 이면 $L(z, n_1, n_2, S_1, S_2)$ 함수에 근이 없습니다.

5. $n_1 > n_2$ 면 $L(z, n_1, n_2, S_1, S_2)$ 가 $+\infty$ 부터 $-\infty$ 까지 단조롭게 감소하므로 고유 근이 있습니다. $L(0, n_1, n_2, S_1, S_2) = \ln S_1^2/S_2^2 \geq 0$ 이면 근이 $[0, n_2)$ 구간 안에 있고 그렇지 않으면 $(-\infty, 0)$ 구간에 있습니다.

증명

아래에서 $L(z) \equiv L(z, n_1, n_2, S_1, S_2)$ 라고 합니다.

먼저 $n_1 < n_2$ 면 $L(z)$ 가 볼록 함수이고 다음과 같을 때 최소값에 도달함을 증명하려고 합니다.

$$z_m = \frac{n_1 + n_2 - \sqrt{(n_1 - n_2)(n_1 - n_2 - \frac{4}{se})}}{2}$$

앞에서 정의한 것처럼

$$L(z) = \ln \frac{n_1}{n_2} + \ln \frac{n_2 - z}{n_1 - z} - z se + \ln \frac{S_1^2}{S_2^2}, z < \min(n_1, n_2)$$

입니다. 그러면 $\lim_{z \rightarrow -\infty} L(z) = +\infty$ 이고

$$\lim_{z \rightarrow \min(n_1, n_2)} L(z) = \begin{cases} +\infty & (n_1 < n_2) \\ -\infty & (n_2 < n_1) \end{cases}$$

입니다. 또한 $L(z)$ 의 도함수는 다음을 충족합니다.

$$-\frac{(n_1 - z)(n_2 - z)}{se} L'(z) = z^2 - (n_1 + n_2)z + n_1 n_2 + \frac{n_1 - n_2}{se}$$

다음과 같다고 할 때,

$$Q(z) = -\frac{(n_1 - z)(n_2 - z)}{se} L'(z)$$

$n_1 < n_2$ 면 2차 방정식 $Q(z)$ 에는 근이 2개이고 다음 식으로 표현됩니다.

$$z_1 = \frac{n_1 + n_2 - \sqrt{(n_1 - n_2)(n_1 - n_2 - \frac{4}{se})}}{2}$$

그리고

$$z_2 = \frac{n_1 + n_2 + \sqrt{(n_1 - n_2)(n_1 - n_2 - \frac{4}{se})}}{2}$$

$Q(n_1) = \frac{n_1 - n_2}{se} < 0$ 이므로, $z_1 < n_1 = \min(n_1, n_2) < z_2$ 고 따라서 $(-\infty, z_1)$ 의 z 에 대해 $Q(z) > 0$ 이고 (z_1, n_1) 의 z 에 대해 $Q(z) < 0$ 입니다. 그러므로 $(-\infty, z_1)$ 의 z 에 대해 $L'(z) < 0$ 이고 (z_1, n_1) 의 z 에 대해 $L'(z) > 0$ 이라고 할 수 있습니다. 따라서 $L(z)$ 는 $(-\infty, \min(n_1, n_2))$ 변역에서 볼록 함수이고 $z_1 \equiv z_m$ 에서 최소값에 도달합니다.

$n_1 > n_2$ 면 두 가지 경우, 즉 $n_1 - n_2 > 4/se$ 인 경우와

$0 < n_1 - n_2 < 4/se$ 인 경우가 있습니다. 첫 번째 경우에서 z_1 과 z_2 는 $Q(z)$ 의 근이고 $n_2 = \min(n_1, n_2) < z_1 < z_2$ 입니다. (이유는 $n_2 - \frac{z_1 + z_2}{2} = \frac{n_2 - n_1}{2} < 0$ 이기 때문입니다.) 따라서 $(-\infty, \min(n_1, n_2))$ 변역의 z 에 대해 $Q(z) > 0$ 입니다. 두 번째 경우에는 $Q(z)$ 에 근이 없으므로 변역에서 $Q(z) > 0$ 입니다.

그러므로 $n_1 > n_2$ 면 $L'(z) < 0$ 이므로 $L(z)$ 는 $+\infty$ 부터 $-\infty$ 까지 단조롭게 감소합니다.

8. 참고 문헌

- Banga, S. J. & Fox, G. D. (2013). On Bonett's Robust Confidence Interval for a Ratio of Standard Deviations. *인쇄 중*.
- Barnard, J. (1978). Probability Integral of the Normal Range. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Vol. 27, 197-198.
- Bonett, D. G. (2006). Robust Confidence Interval for a Ratio of Standard Deviations. *Applied Psychological Measurements*, 30, 432-439.
- Brown, M. B. & Forsythe A. B. (1974). Robust Tests for the Equality of Variances. *Journal of the American Statistical Association*, 69, 364-367.
- Conover, W. J., Johnson, M. E. & Johnson, M. M. (1981). A Comparative Study of Tests for Homogeneity of Variances, with Applications to the Outer Continental Shelf Bidding Data. *Technometrics*, 23, 351-361.
- Hochberg, Y., Weiss, G. & Hart S. (1982). On Graphical Procedures for Multiple Comparisons. *Journal of the American Statistical Association*, 77, 767-772.
- Kramer, C. Y. (1956). Extension of Multiple Range Tests to Group Means with Unequal Numbers of Replications. *Biometrics*, 12, 307-310.
- Layard, M. W. J. (1973). Robust Large-Sample Tests for Homogeneity of Variances. *Journal of the American Statistical Association*, 68, 195-198.
- Levene, H. (1960). "Robust Tests for Equality of Variances," I. Olkin, ed., *Contributions to Probability and Statistics*, Palo Alto, CA: Stanford University Press, 278-292.
- Miller, R. G. (1968). Jackknifing Variances. *Annals of Mathematical Statistics*, 39, 567-582.
- Nakayama, M. K. (2009). Asymptotically Valid Single-Stage Multiple-Comparison Procedures. *Journal of Statistical Planning and Inference*, 139, 1348-1356.
- Ott, R. L. and Longnecker, M. (2010). *An introduction to Statistical Methods and Data Analysis, sixth edition*, Brooks/Cole, Cengage Learning.
- Pan, G. (1999). On a Levene Type Test for Equality of Two Variances. *Journal of Statistical Computation and Simulation*, 63, 59-71.
- Stoline, M. R. (1981). The Status of Multiple of Comparisons: Simultaneous Estimation of All Pairwise Comparisons in One-Way ANOVA Designs. *The American Statistician*, 35, 134-141.

Tukey, J. W. (1953). *The Problem of Multiple Comparisons*. Mimeographed monograph.

Wolfram, S. (1999). *The Mathematica Book*, 4th ed. Wolfram Media/Cambridge University Press.

© 2020 Minitab, LLC. All rights reserved. Minitab®, Minitab Workspace™, Companion by Minitab®, Salford Predictive Modeler®, SPM®, and the Minitab® logo are all registered trademarks of Minitab, LLC, in the United States and other countries. Additional trademarks of Minitab, LLC can be found at www.minitab.com. All other marks referenced remain the property of their respective owners.