

Bonett 방법

표준 편차의 비율에 대한 BONETT의 로버스트 신뢰 구간

Senin J. Banga & Gregory D. Fox

2013년 6월 18일

개요

당사는 Bonett(2006)방법으로 구한 두 표준 편차 비율의 신뢰 구간(CI)에 있는 작은 오류를 정정하기 위한 대체 방법을 제안합니다. Bonett 구간의 기초인 Layard(1973)의 검정 통계량에 대한 합동 척도 추정치는 모 분산이 같을 때만 일치성을 갖습니다. 당사는 모 분산이 같을 때와 같지 않을 때에 모두 일치성을 가지는 대체 추정량을 도출하고 새 추정량을 사용하여 올바른 CI를 계산합니다. 시뮬레이션 연구에서는 새 CI가 일반적으로 Levene/Browne-Forsythe 검정 W_{50} 과 Pan(1999)의 검정 L_{50} 에 기초한 CI보다 더 정확하고 정밀한 것으로 나타났습니다. 당사는 Pan과 마찬가지로 검정 W_{50} 에 기초한 CI가 표본이 작을 때 정밀성 손실을 보임으로 인해 종종 구간의 너비가 무한하게 된다는 사실을 관측했습니다. 검정 L_{50} 에 기초한 CI는 대칭 및 근대칭 분포에서 효과적이지만, 모집단이 치우쳤을 때는 효과적이지 않습니다.

색인 용어: 분산의 동질성, Levene의 검정, Brown-Forsythe 검정, Layard의 검정, 분산 비율에 대한 신뢰 구간(CI)

1. 소개

널리 알려져 있는 기존 F 검정의 신뢰 구간(CI)은 정규성 이탈에 매우 민감합니다. 사실, 기존 F 검정은 너무 민감해서 실제 적용하기에는 거의 적합하지 않습니다. 이런 이유로 많은 이들이 더 로버스트한 대안을 제안했습니다. 이 중에 "검정 W_{50} "이라는 검정은 제1종 오류

속성이 매우 우수하면서도, 계산이 간단하고, 해석하기 쉽기 때문에 선호되는 경우가 많습니다. (비교 분석에 대해서는 Conover 외 (1981), Balakrishnan & Ma(1990), Lim & Loh(1996)를 참조하십시오.) 검정 W_{50} 은 Levene(1960)이 처음 제안하고, 그 이후 Brown과 Forsythe(1974)이 개선시킨 방법에 기초한 것입니다. 검정 W_{50} 은 널리 채택되어 왔으며, Minitab Statistical Software, SAS, R 및 JMP 같은 유명 통계 소프트웨어 패키지에서 제공됩니다.

검정 W_{50} 의 제2종 오류 속성은 제1종 오류 속성보다 약간 덜 두드러집니다. Pan(1999)에 따르면, 정규 분포를 포함한 일부 분포의 경우 검정 W_{50} 의 2-표본 문제에 대한 검정력 상한은 1에 훨씬 못 미칩니다. 그리고 이 상한은 모분산 간의 차이가 어느 정도인 지에 영향을 받지 않습니다. 이 문제는 자연히 검정 W_{50} 에 기초한 CI까지 확대됩니다. Pan에 따르면, 검정 W_{50} 에 기초한 모 분산의 비율에 대한 CI가 무한($0, +\infty$)하기 때문에 유용하지 않을 가능성이 적지 않게 존재합니다. Pan의 관찰 결과는 본 문서의 뒷부분에 보고된 시뮬레이션 결과와 일치합니다.

Pan은 W_{50} 방법의 문제를 해결하기 위한 L_{50} 이라는 대체 방법을 제안합니다. 시뮬레이션 결과를 토대로, Pan은 검정 L_{50} 이 검정 W_{50} 보다 더 강력하면서도 두 검정이 똑같이 로버스트하고 똑같이 바람직한 접근 속성을 갖고 있다는 결론을 내립니다. 그러나 Pan의 시뮬레이션에 사용된 표본은 대칭 분포 또는 꼬리가 두껍거나 얇은 약간 치우친 분포에서 추출되었습니다. 소규모 표본을 사용한 경우 왜도가 L_{50} 검정의 성능에 미칠 수 있는 잠재적인 영향에 대한 구체적인 설명은 없습니다.

Pan은 L_{50} 방법이 수정된 Fligner-Killeen 순위 검정과 Hall-Padmanabhan 대응 검정 등과 같이 매우 로버스트한 다른 방법들만큼 강력하다고 주장하기도 합니다. 그러나 실제로는 Fligner-Killeen 순위 검정과 Hall-Padmanabhan 대응 검정은 힘든 연산을 많이 요구하기 때문에 L_{50} 및 W_{50} 검정보다 유용성이 다소 떨어집니다.

최근에 Bonett(2006)은 CI를 구하는 새로운 방법을 제안했는데, Layard(1973) 분산 동질성 검정의 2-표본 버전에 기초한 것입니다. Bonett은 Layard 방법의 소표본 성능을 개선하기 위해 몇 가지를 조정합니다. 예를 들어, Bonett은 Layard의 방법과 점근적으로 동등하지만 소표본 치우침을 덜 나타내는 합동 척도 추정량을 제안합니다.

그러나 Layard의 기본 합동 척도 추정량과 Bonett이 제안한 대체 추정량 모두 모분산이 같지 않은 경우 일치성을 갖지 않습니다. 따라서 Bonett(2006)이 제안한 구간은 CI가 아니라 등분산 검정의 합격 구간이라고 하는 것이 더 적절합니다. 그러므로, Bonett(2006)에서 보고된 모의 포함 확률을 1에서 빼면 등분산 검정의 제1종 오류율이 얻어집니다. 이렇게 구한 제1종 오류율과 Layard의 기본 검정의 제1종 오류율을 비교하면, Bonett의 조정이 소표본에

대한 Layard 검정의 성능을 개선하는 것을 확인할 수 있습니다. 그러나 Bonett에 의해 제안된 분산 비율은 다시 검토해야 합니다.

또한 Bonett은 제안된 구간을 Shoemaker(2003)의 근사 F 검정에 기초한 CI에도 비교합니다. 그러나 Shoemaker의 검정과 연관된 분산 비율의 CI도 (Shoemaker의 논문 106페이지에 간략히 설명되어 있는 것과 같이) Layard의 합동 침도 추정량에 기초합니다. 그러므로 Shoemaker의 논문 7장에 계산되어 있는 CI도 등분산 검정의 합격 구간이라고 설명하는 것이 가장 적절합니다. 이 같은 오류에도 불구하고, Bonett의 시뮬레이션 결과에서는 그의 조정에 의해 Layard 등분산 검정의 소표본 성능이 개선되었고 조정 후에 도출된 등분산 검정의 성능이 Shoemaker의 검정보다 낫다고 결론지을 수 있습니다.

본 문서에서는 Layard 검정의 2-표본 형식을 확장하여 분산 또는 표준 편차의 비율에 대한 귀무 가설을 검정함으로써 Bonett(2006)의 오류를 정정합니다. 이를 위해, 당사는 모든 주어진 가설 비율에 대해 일치성을 갖는 합동 침도 추정량을 제안합니다. 다음, 검정 통계량을 뒤집어서 비율에 대한 CI를 구합니다. 마지막으로, 소표본 설계에서 새 CI의 로버스트성 속성을 평가하기 위한 시뮬레이션 연구를 실시합니다. 또한 소표본일 때, 새 CI의 성능을 기존 F 검정, 검정 W_{50} 및 검정 L_{50} 에 대한 CI의 성능과 비교합니다.

2. Layard 검정과 몇 가지 확장

$Y_{i1}, \dots, Y_{in_i}, \dots, Y_{k1}, \dots, Y_{kn_k}$ 는 k 개의 독립 표본이고, 각 표본은 독립적이고 평균이 $E(Y_{ij}) = \mu_i$ 고 분산이 $\text{Var}(Y_{ij}) = \sigma_i^2 > 0$ 인 동일한 분포를 따릅니다. 또한 표본이 공통 침도가 $\gamma = E(Y - \mu)^4 / \sigma^4 < \infty$ 인 모집단에서 추출되었다고 가정합니다. Layard는 $\gamma_e = \gamma - 3$ 이상의 침도를 사용합니다.

\bar{Y}_i 와 S_i 는 각각 표본 i 의 평균과 표준 편차입니다. 또한 $\tau^2 = 2 + (1 - 1/\bar{n})\gamma_e = 2 + (1 - 1/\bar{n})(\gamma - 3)$ 입니다. 여기서 $\bar{n} = \sum n_i / k$ 입니다. Layard(1973)에 나타나 있는 것처럼, 대표본에 대해 $\tau^2 \cong \text{Var}((n_i - 1)^{1/2} \ln S_i^2)$ 입니다.

등분산 귀무 가설을 검정하기 위해 Layard는 성분 $Z_i = (n_i - 1)^{1/2} \ln S_i^2 / \tau$ 가 귀무 가설 하에서 표준 정규 분포에 근사하는 분포를 보이는 벡터에 대해 직교 변환을 수행합니다. 그 다음에, 직교 변환의 거리 보존 속성을 이용하여 검정 통계량 S' (아래 참조)의 분포는 등분산 귀무 가설 하에서 자유도가 $k - 1$ 인 카이 제곱 분포를 따르는 점근 분포를 나타냄을 입증합니다.

$$S' = \sum_{i=1}^k (n_i - 1) \left(\ln S_i^2 - \frac{\sum_{i=1}^k (n_i - 1) \ln S_i^2}{\sum_{i=1}^k (n_i - 1)} \right)^2 / \tau^2$$

일반적으로, $Z_i = (n_i - 1)^{\frac{1}{2}}(\ln S_i^2 - \ln \sigma_i^2)/\tau$ 의 분포는 표준 정규 분포에 근사합니다. 그러므로 Layard의 기법을 적용하여 보다 일반화된 검정 통계량인 T'_k 를 도출할 수 있습니다.

$$T'_k = \sum_{i=1}^k (n_i - 1) \frac{(\ln S_i^2 - \ln \sigma_i^2)^2}{\tau^2} - \left(\sum_{i=1}^k (n_i - 1) \frac{\ln S_i^2 - \ln \sigma_i^2}{\tau \sqrt{\sum_{i=1}^k (n_i - 1)}} \right)^2$$

T'_k 의 분포는 귀무 가설과 대립 가설 하에서 모두 자유도가 $k - 1$ 인 카이 제곱 분포에 근사합니다.

T'_k 의 식은 S' 의 식과 더 유사하게 표현할 수 있습니다. 제곱 항을 이중 합으로 표현하고 몇 가지 대수 연산을 수행하면 다음이 얻어집니다.

$$T'_k = \sum_{i=1}^k (n_i - 1) \left(\ln S_i^2 - \ln \sigma_i^2 - \frac{\sum_{i=1}^k (n_i - 1) (\ln S_i^2 - \ln \sigma_i^2)}{\sum_{i=1}^k (n_i - 1)} \right)^2 / \tau^2$$

모든 분산이 같으면 $T'_k = S'$ 입니다. 그러므로 S' 와 T'_k 는 등분산 귀무 가설을 검정할 때 같은 검정 통계량입니다. 그러나 T'_k 를 더 일반적으로 사용하여 분산의 함수로 표현되는 모든 가설을 검정할 수도 있습니다. 예를 들어 T'_k 를 사용하여 모든 주어진 $\sigma_{0i} > 0, i = 1, \dots, k$ 에 대해 $H_0: \sigma_i = \sigma_{0i}$ 형식의 귀무 가설을 검정할 수 있습니다.

$\tau^2 = 2 + (1 - 1/\bar{n})(\gamma - 3)$ 은 알 수 없기 때문에 S' 또는 T'_k 에 기초한 검정에서는 모집단의 공통 첨도 γ 에 대한 추정량이 필요합니다. 예를 들어 등분산 귀무 가설을 검정하기 위해 Layard는 다음과 같은 공통 첨도 합동 추정량을 제안합니다.

$$\hat{\gamma} = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^4}{\left[\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2 \right]^2} \sum_{i=1}^k n_i$$

그러나 Layard는 분산이 같지 않을 때 공통 첨도의 추정량 $\hat{\gamma}$ 이 일치성을 갖지 않을 수 있다고 지적합니다.

특히 2-표본 설계의 경우에는 귀무 가설 $H_0: \sigma_1/\sigma_2 = \rho_0$ 을 주어진 가설 비율 $\rho_0 > 0$ 에 대해 검정하여 표준 편차 간의 차이가 얼마나 큰 지를 평가할 수 있습니다. 그러나 표준 편차 비율에 대한 CI를 계산하여 해당 차이를 더 직접적으로 평가할 수 있습니다.

$\rho_0 = 1$ 이면 귀무 가설이 분산의 동질성 가설과 동일합니다. 그러므로 Layard의 첨도 추정량의 2-표본 버전을 $\tau^2 = 2 + (1 - 1/\bar{n})(\gamma - 3)$ 식의 γ 로 대체하여 $\hat{\tau}^2$ 을 구한 후에 검정을 $T'_2 = S'$ 에 기초하여 수행할 수 있습니다.

그러나 $\rho_0 \neq 1$ 이면 검정을 S' 가 아닌 T_2' 에 기초하여 수행해야 합니다. 또한 $\rho_0 \neq 1$ 이면 Layard의 합동 척도 추정량이 일치성을 갖지 않을 수 있으므로 모집단의 공통 척도를 추정하는 데 사용할 수 없습니다. 그러므로 모든 가설 비율 $\rho_0 > 0$ 에 대해 일치성을 갖는 대체 합동 척도 추정량이 필요합니다.

다음에는 해당 추정량을 도출합니다. 추정량은 ρ_0 의 함수이기 때문에 $\hat{\gamma}_P(\rho_0)$ 로 나타냅니다.

검정 통계량 $T_2 = \tau^2 T_2' / \hat{\tau}^2$ 도 정의합니다. 여기서 $\hat{\tau}^2 = 2 + (1 - 1/\bar{n})(\hat{\gamma}_P(\rho_0) - 3)$ 입니다.

Slutzky의 정리에 의해 T_2 는 자유도가 1인 카이 제곱 분포를 근사적으로 따릅니다.

마지막으로, T_2 를 뒤집어서 $\rho = \sigma_1/\sigma_2$ 에 대한 CI를 구합니다.

3. 표준 편차 비율의 CI

앞에서는 분산 또는 표준 편차의 비율로 표현되는 귀무 가설을 검정할 때 대체 척도 추정량이 필요한 이유에 대해 자세히 설명했습니다. 이어지는 결과에서는 해당 추정량을 제시합니다.

결과 1

주어진 $\rho = \sigma_1/\sigma_2 > 0$ 에 대해, 2-표본 모형에서 공통 모집단 척도의 합동 척도 일치 추정량은 다음과 같이 나타낼 수 있습니다.

$$\hat{\gamma}_P(\rho) = (n_1 + n_2) \frac{\sum_{j=1}^{n_1} (Y_{1j} - \bar{Y}_1)^4 + \rho^4 \sum_{j=1}^{n_2} (Y_{2j} - \bar{Y}_2)^4}{[(n_1 - 1)S_1^2 + \rho^2(n_2 - 1)S_2^2]^2}$$

이 결과의 증명은 부록 A에 나와 있습니다.

예상대로, $\sigma_1/\sigma_2 = 1$ 이면 표준 편차(또는 분산)가 같음을 의미하므로 $\hat{\gamma}_P(1)$ 은 Layard의 합동 척도 추정량 $\hat{\gamma}$ 와 같습니다.

일반화된 통계량 T_k' 의 2-표본 버전인 통계량 T_2' 은 다음 식으로 표현됩니다.

$$T_2' = \frac{(\ln \hat{\rho}^2 - \ln \rho^2)^2}{\left(\frac{1}{n_1 - 1} + \frac{1}{n_2 - 1}\right) \tau^2}$$

여기서 $\hat{\rho} = S_1/S_2$, $\rho = \sigma_1/\sigma_2$, 그리고 $\tau^2 = 2 + (1 - 1/\bar{n})\hat{\gamma}_e = 2 + (1 - 1/\bar{n})(\gamma - 3)$ 입니다.

Layard(1973)에 나와 있는 것처럼, 대표본에서 $\tau^2 \cong \text{Var}((n_i - 1)^{1/2} \ln S_i^2)$ 입니다.

Bonett(2006)에서는 Shoemaker(2003)에서도 사용된 대체 근사 $\text{Var}((n_i - 1)^{1/2} \ln S_i^2) \cong \gamma - (n_i - 3)/n_i$ 를 사용합니다. 대표본에서 이 두 가지 근사는 동등합니다. 그러나 Shoemaker는 그의 등분산 검정을 소표본에서 사용할 때 후자의 근사가 유익하다고 보고합니다. 이 조정을 사용하여 통계량 T_2' 를 다음과 같이 수정할 수 있습니다.

$$T_2' = \frac{(\ln \hat{\rho}^2 - \ln \rho^2)^2}{\frac{\gamma - g_1}{n_1 - 1} + \frac{\gamma - g_2}{n_2 - 1}}$$

여기서 $g_i = (n_i - 3)/n_i$ 입니다.

따라서 귀무 가설 $H_0: \rho = \rho_0$ 의 검정을 위한 검정 통계 $T_2 = \tau^2 T_2' / \hat{\tau}^2$ 는 다음 식으로 표현됩니다.

$$T_2 = \frac{(\ln \hat{\rho}^2 - \ln \rho_0^2)^2}{\frac{\hat{\gamma}_P(\rho_0) - g_1}{n_1 - 1} + \frac{\hat{\gamma}_P(\rho_0) - g_2}{n_2 - 1}}$$

이 T_2 식에서는 분모의 제곱근이 합동 척도에 대한 표준 오차의 대표본 추정치라고 생각할 수 있습니다.

나아가 $\hat{\gamma}_P(1) \equiv \hat{\gamma}$ 식에서 Bonett(2006)은 절사 비율이 $1/[2(n_i - 4)^{1/2}]$ 인 절사된 표본 평균을 사용합니다. 이에 따라 합동 척도 추정량에 동일한 조정을 적용합니다.

$$\hat{\gamma}_P(\rho) = (n_1 + n_2) \frac{\sum_{j=1}^{n_1} (Y_{1j} - m_1)^4 + \rho^4 \sum_{j=1}^{n_2} (Y_{2j} - m_2)^4}{[(n_1 - 1)S_1^2 + \rho^2(n_2 - 1)S_2^2]^2}$$

여기서 m_i 는 절사 비율이 $1/[2(n_i - 4)^{1/2}]$ 인 표본 i 의 절사 평균입니다. 절사 평균 m_i 는 모평균 μ_i 의 일치 추정량이므로 이 합동 척도 추정량 버전과 이전 버전은 점근적으로 동일합니다. 그러나 이 대체 버전은 T_2 에 기초한 검정의 소표본 성능을 개선할 수 있습니다.

이제 검정 통계량 T_2 을 뒤집어서 분산 또는 표준 편차의 비율에 대한 근사 CI를 도출할 수 있습니다. 그러나 우선 표준 편차의 비율에 대한 Bonett(2006) CI를 도출하는 데 수반된 오류에 대해 간략히 설명합니다.

3.1 Bonett 구간

T_2 를 뒤집어서 CI를 구하는 대신, Bonett(2006)은 다음 통계량을 역으로 바꿉니다.

$$T = \frac{(\ln \hat{\rho}^2 - \ln \rho^2)^2}{\frac{\hat{\gamma}_P(1) - g_1}{n_1 - 1} + \frac{\hat{\gamma}_P(1) - g_2}{n_2 - 1}}$$

따라서 그 결과로 얻어지는 구간은 단순히 등분산 검정의 합격 영역에 불과합니다. 그 이유는 합동 척도 추정량 $\hat{\rho}_P(1)$ 이 분산이 같거나 가설 비율이 1일 때만 일치성을 가지기 때문입니다. 그 결과로 얻어지는 구간은 Bonett(2006)에서 다음과 같이 보고되었습니다.

$$\exp[\ln(c S_1^2/S_2^2) \pm z_{\alpha/2} se]$$

여기서,

$$se^2 = \frac{\hat{\gamma}(1) - g_1}{n_1 - 1} + \frac{\hat{\gamma}(1) - g_2}{n_2 - 1}$$

상수 c 는 불균형 설계에서 비등 꼬리 오차 확률의 영향을 완화하기 위한 소규모 표본 조정으로서 포함됩니다. 이 상수는 다음 식으로 표현됩니다.

$$c = \frac{n_1}{n_1 - z_{\alpha/2}} \frac{n_2 - z_{\alpha/2}}{n_2}$$

상수는 설계가 균형일 때 없어지고, 상수의 영향은 표본 크기가 증가함에 따라 무시할 만한 수준으로 감소합니다.

표 1에는 위 구간을 CI으로 잘못 해석할 때 초래되는 결과가 나와 있습니다. 해당 결과는 Bonett(2006)의 구간을 토대로 모의 포함 확률을 계산하는 소규모 시뮬레이션 연구에 기초한 것입니다. 등분산의 경우(왼쪽 열), 표준 정규 분포에서 독립 표본 2개를 추출합니다. 이분산의 경우(오른쪽 열), 두 번째 표본의 관측치를 일정하게 4배로 확대합니다. 추정 포함 확률은 100,000 반복실험 기준입니다. 목표 명목 포함 확률은 0.95입니다.

표 1 서로 다른 모 분산이 Bonett(2006)의 CI에 미치는 영향($\alpha = 0.05$)

n_1, n_2	모의 포함 확률	
	등분산	이분산
10, 10	0.963	0.972
50, 50	0.952	0.991
100, 100	0.952	0.994

구간이 일정한 합동 척도 일치 추정량에 기초한 경우, 두 경우의 포함 확률이 동일하다고 예상할 수 있습니다. 그러나 분산이 같지 않을 때는 구간이 일관되게 더 보수적이라는 점에 주목하십시오. 나아가 표본 크기가 커질수록 포함 확률이 1에 근접합니다.

Shoemaker(2003)의 근사 CI를 사용할 때도 유사한 결과가 얻어집니다.

3.2 CI 계산

앞에서 주어진 검정 통계량 T_2 에 기초하여 $\rho = \sigma_1/\sigma_2$ 고 $\rho_0 > 0$ 일 때 귀무 가설 $H_0: \rho = \rho_0$ 을 대립 가설 $H_A: \rho \neq \rho_0$ 에 대해 검정하는 문제에 대해 생각해 보십시오. 귀무 가설 하에서, 검정 통계량

$$T_2 = \frac{(\ln \hat{\rho}^2 - \ln \rho_0^2)^2}{\frac{\hat{Y}_P(\rho_0) - g_1}{n_1 - 1} + \frac{\hat{Y}_P(\rho_0) - g_2}{n_2 - 1}}$$

은 자유도가 1인 카이 제곱 분포를 근사적으로 따릅니다. 따라서 검정은 다음과 같은 경우에만 귀무 가설을 α 유의 수준에서 기각합니다.

$$(\ln \hat{\rho}^2 - \ln \rho_0^2)^2 > z_{\alpha/2}^2 \left(\frac{\hat{Y}_P(\rho_0) - g_1}{n_1 - 1} + \frac{\hat{Y}_P(\rho_0) - g_2}{n_2 - 1} \right)$$

여기서 z_{α} 는 표준 정규 분포의 상위 $\alpha \times 100$ 번째 백분위수를 나타냅니다. 자유도가 1인 카이 제곱 분포의 $\alpha \times 100$ 번째 상위 백분위수 $\chi_{1,\alpha}^2$ 는 다음 조건을 충족합니다. $\chi_{1,\alpha}^2 = z_{\alpha/2}^2$.

Bonett(2006)의 시뮬레이션 결과는 불균형 설계에서 비등 꼬리 오차 확률의 영향을 줄이기 위한 소규모 표본 조정이 효과적이었음을 보여줍니다. 따라서 T_2 에 기초한 검정에 유사한 조정을 적용합니다. 이 조정을 적용하면 검정에서 다음과 같은 경우에만 귀무 가설이 기각됩니다.

$$(\ln \rho_0^2 - \ln(c\hat{\rho}^2))^2 > z_{\alpha/2}^2 \left(\frac{\hat{Y}_P(\rho_0) - g_1}{n_1 - 1} + \frac{\hat{Y}_P(\rho_0) - g_2}{n_2 - 1} \right)$$

여기서 c 는 다음 식으로 표현되는 Bonett의 조정 상수입니다.

$$c = \frac{n_1}{n_1 - z_{\alpha/2}} \frac{n_2 - z_{\alpha/2}}{n_2}$$

마찬가지로, T_2 에 기초하여 $\rho = \sigma_1/\sigma_2$ 에 대해 설정한 근사 $(1 - \alpha)100$ 퍼센트 신뢰 구간은 다음 식으로 계산됩니다.

$$\left\{ \rho \in (0, \infty) : (\ln \rho^2 - \ln(c\hat{\rho}^2))^2 - z_{\alpha/2}^2 \left(\frac{\hat{Y}_P(\rho) - g_1}{n_1 - 1} + \frac{\hat{Y}_P(\rho) - g_2}{n_2 - 1} \right) \leq 0 \right\}$$

c 는 균형 설계에서 어떤 영향도 미치지 않고 대표본 불균형 설계에서만 미미한 영향이 있습니다.

다음 결과에서는 신뢰 구간의 대체 식을 신뢰도의 특성을 설명하는 데 편리한 형식으로 제공합니다. 이 식에서 합동 척도 추정량은 다음 식으로 표현되는 개별 표본 척도에 기초하여 다시 작성됩니다.

$$\hat{\gamma}_i = n_i \frac{\sum_{j=1}^{n_i} (Y_{ij} - m_i)^4}{[(n_i - 1)S_i^2]^2}, i = 1, 2$$

결과 2

T_2 에 기초하여 $\rho = \sigma_1/\sigma_2$ 에 대해 설정한 근사 $(1 - \alpha)100$ 퍼센트 신뢰 구간은 다음 식으로 표현할 수 있습니다.

$$\hat{\rho}\sqrt{c} \left\{ r \in (0, \infty) : H(r^2) \leq 0 \right\}$$

또는 이와 동등하게 $\rho^2 = \sigma_1^2/\sigma_2^2$ 에 대한 신뢰 구간을 다음과 같이 표현할 수 있습니다.

$$c\hat{\rho}^2 \left\{ r \in (0, \infty) : H(r) \leq 0 \right\}$$

여기서,

$$H(x) = (\ln x)^2 - z_{\alpha/2}^2 se^2(cx), x > 0$$

$$se^2(x) = A \frac{\hat{\gamma}_1 K^2/n_1 + \hat{\gamma}_2 x^2/n_2}{(K+x)^2} - B$$

$$A = \frac{(n_1+n_2)(n_1+n_2-2)}{(n_1-1)(n_2-1)}, B = \frac{g_1}{n_1-1} + \frac{g_2}{n_2-1}, K = \frac{n_1-1}{n_2-1}$$

이 결과의 증명은 부록 B를 참조하십시오.

함수 $H(x)$ 가 양의 실선으로 연속적이고 $H(0) = H(+\infty) = +\infty$ 고 $H(1) < 0$ 임은 쉽게 확인됩니다. 그러므로 중간 값 정리에 의해 함수 $H(x)$ 는 구간 $(0, 1)$ 과 구간 $(0, +\infty)$ 에서 각각 근을 하나 이상 받아들입니다.

다음 결과에서는 신뢰 구간을 하나의 구간이나 분리 구간의 합집합으로 설명합니다.

결과 3

함수 $H(x)$ 에 정확히 2개의 근 x_L 및 x_U 가 있으면 $0 < x_L < 1 < x_U$ 이며, $\rho^2 = \sigma_1^2/\sigma_2^2$ 에 대한 신뢰 구간은 다음과 같습니다.

$$[c\hat{\rho}^2 x_L, c\hat{\rho}^2 x_U]$$

따라서 $\rho = \sigma_1/\sigma_2$ 에 대한 CI는 다음 구간입니다.

$$[\hat{\rho}\sqrt{cx_L}, \hat{\rho}\sqrt{cx_U}]$$

반면에 함수 $H(x)$ 에 근이 2개보다 많으면 $\rho^2 = \sigma_1^2/\sigma_2^2$ 에 대한 신뢰 구간은 겹치지 않는 구간의 합집합입니다. 각 구간의 끝점은 함수가 위쪽으로 열리는 연속적인 근입니다.

이 결과의 증명은 부록 C를 참조하십시오.

설명

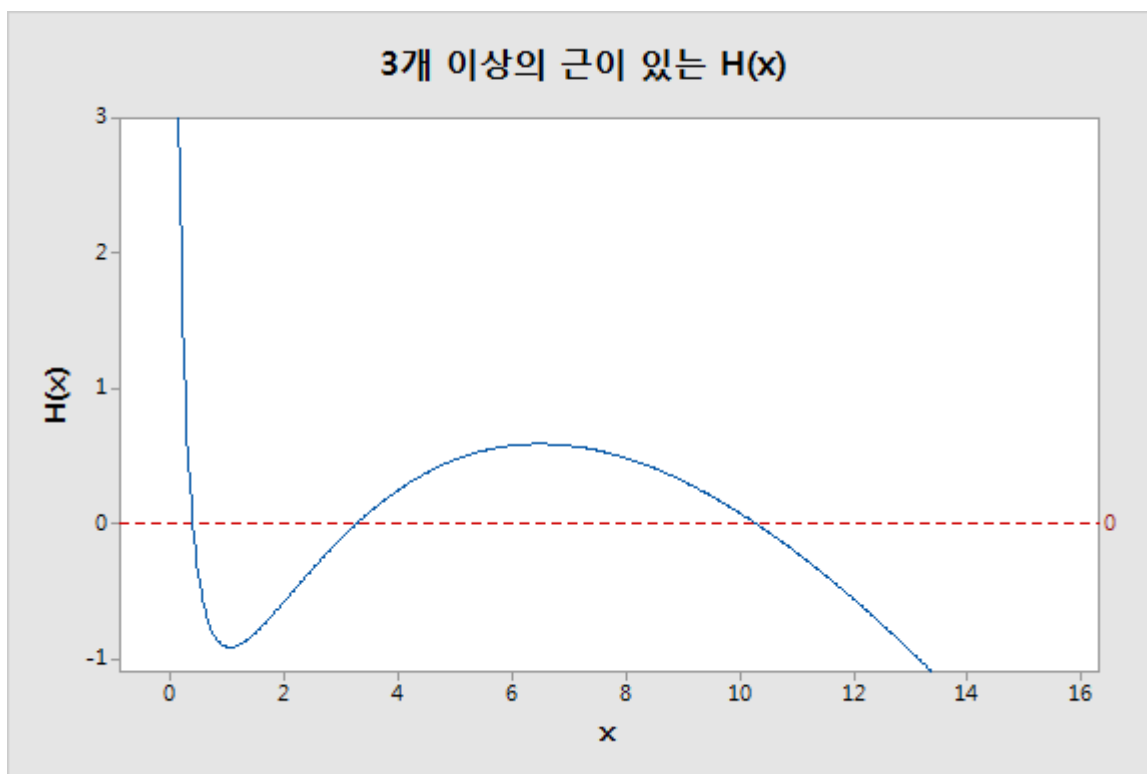
함수 $H(x)$ 가 2개 이상의 근을 가지는 것도 수학적으로 가능하지만, 표본 하나 이상이 너무 작거나 심하게 불균형인 극도로 비정상적이고 실용적으로 무의미한 설계에서만 가능한 것으로 관측되었습니다. $H(x)$ 에는 근이 2개 또는 4개라고 추측됩니다.

다음 예는 함수 $H(x)$ 에 근이 3개 이상 있도록 조작된 데이터에 기초하고 있습니다. 데이터는 다음과 같이 요약됩니다 $n_1 = 169$, $n_2 = 7$, $S_1 = 301.855$, $S_2 = 4606.170$, $\hat{\gamma}_1 = 1.877$, $\hat{\gamma}_2 = 6.761$, $c = 0.728$, $A = 30.381$, $B = 0.101$ 및 $K = 28.000$.

$\alpha = 0.05$ 일 때 함수 $H(x)$ 는 다음과 같이 표현됩니다.

$$H(x) = (\ln x)^2 - 1.960^2 \left(30.381 \frac{1.877 \times 28^2/169 + 6.761 \times (.728x)^2/7}{(28.000 + 0.728x)^2} - 0.101 \right)$$

이 경우에 함수 $H(x)$ 에는 근이 4개입니다. 아래에는 함수의 그래프가 나와 있습니다. 네 번째 근은 너무 크기 때문에 그래프에 표시되지 않았습니니다. 그러나 $H(+\infty) = +\infty$ 기 때문에 네 번째 근이 존재함을 알 수 있습니다.



4개 근의 수치는 $x_1 = 0.389$, $x_2 = 3.282$, $x_3 = 10.194$ 및 $x_4 = 39685.0$ 으로 계산됩니다. 추정 표준 편차 비율은 $\hat{\rho} = S_1/S_2 = 0.066$ 입니다. $\rho^2 = \sigma_1^2/\sigma_2^2$ 에 대한 신뢰 구간은 다음과 같이 표현할 수 있습니다.

$$[c \hat{\rho}^2 x_1, c \hat{\rho}^2 x_2] \cup [c \hat{\rho}^2 x_3, c \hat{\rho}^2 x_4] = [0.001, 0.010] \cup [0.032, 124.072]$$

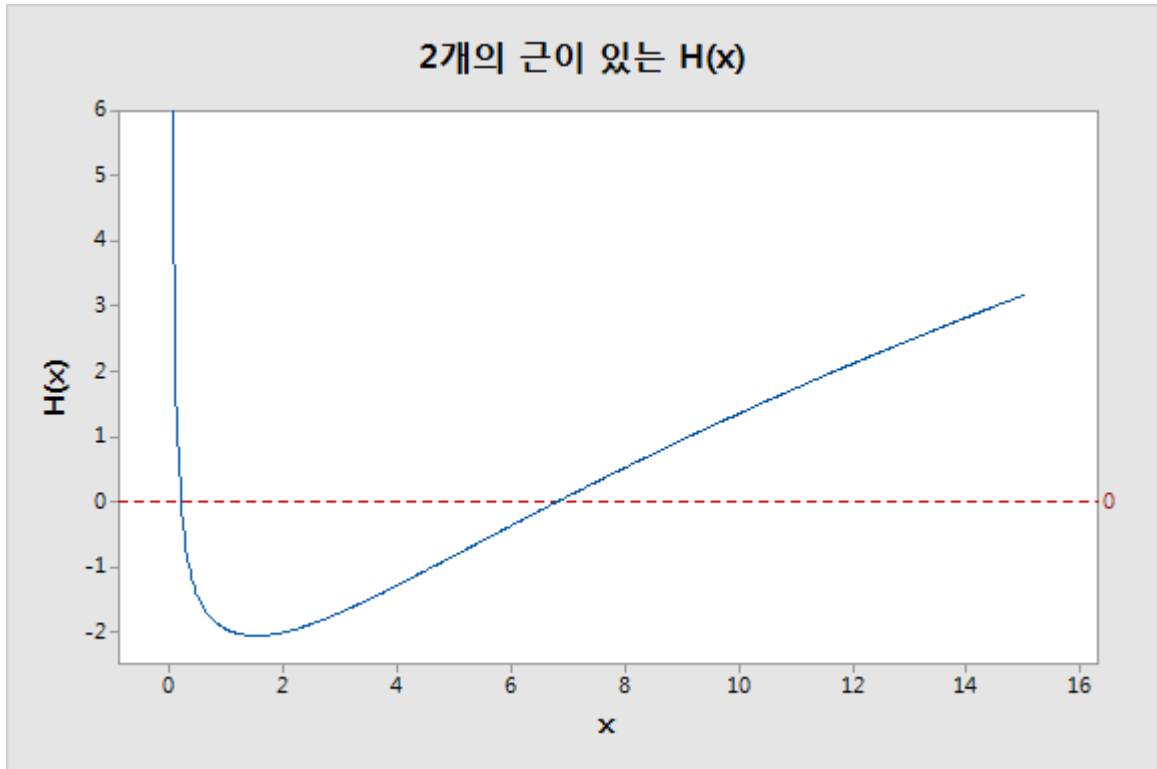
표준 편차 ρ 의 비율에 대한 신뢰 구간은 구간 끝점의 제곱근을 취하여 구합니다.

표본이 너무 작지 않고($n_i \geq 10$) 표본 크기의 차이가 크지 않으면 함수 $H(x)$ 는 일반적으로 2개의 근을 갖습니다. 결과 2에 설명되어 있는 것처럼 근 하나는 1 미만이고 다른 근은 1보다 큽니다. 아래에는 무작위로 생성된 데이터에 기초한 예가 나와 있습니다. 데이터는 다음과 같이 요약할 수 있습니다. $n_1 = 10$, $n_2 = 12$, $S_1 = 1.150$, $S_2 = 1.043$, $\hat{\gamma}_1 = 2.704$, $\hat{\gamma}_2 = 3.671$, $c = 1.041$, $A = 4.444$, $B = 0.146$ 및 $K = 0.818$

$\alpha = 0.05$ 일 때 함수 $H(x)$ 는 이 경우에 다음과 같이 표현됩니다.

$$H(x) = (\ln x)^2 - 1.960^2 \left(4.444 \frac{2.704 \times 0.818^2/10 + 3.671 \times (1.041x)^2/12}{(0.818 + 1.041x)^2} - 0.146 \right)$$

함수 $H(x)$ 에는 아래 그림과 같이 근이 2개 있습니다.



두 근의 수치는 $x_1 = 0.200$ 및 $x_2 = 6.824$ 로 계산됩니다. 추정 표준 편차 비율은 $\hat{\rho} = S_1/S_2 = 1.102$ 입니다. $\rho^2 = \sigma_1^2/\sigma_2^2$ 에 대한 신뢰 구간은 다음과 같습니다.

$$[c \hat{\rho}^2 x_1, c \hat{\rho}^2 x_2] = [0.253, 8.634]$$

표준 편차 ρ 의 비율에 대한 CI는 위 구간의 끝점의 제곱근을 취하여 구합니다.

이제 신뢰 한도를 구하는 알고리즘 2개에 대해 설명합니다.

첫 알고리즘에서는 근 수치 찾기 절차를 사용하여 함수 $H(x)$ 의 근을 구합니다. 분산의 비율에 대한 신뢰 하한에 해당되는 근은 $(0, 1)$ 구간으로 한정됩니다. 이 근을 x_L 으로 나타내면, 결과 3에 의해 분산 비율의 신뢰 하한은 $c\hat{\rho}^2 x_L$ 로 계산되고 표준 편차 비율의 신뢰 하한은 $\hat{\rho}\sqrt{cx_L}$ 로 구해집니다. 마찬가지로, 분산 비율의 신뢰 상한은 $c\hat{\rho}^2 x_U$ 고, 표준 편차 비율의 신뢰 상한은 $\hat{\rho}\sqrt{cx_U}$ 입니다. 여기서 $x_U > 1$ 은 $H(x)$ 의 나머지 근입니다. 신뢰 상한을 구하는 간단한 방법은 $1/\rho^2$ 의 하한이 ρ^2 의 상한이라는 사실을 이용하는 것입니다. 첫째, 함수 $H(x)$ 의 식에서는 $1/\rho^2 = \sigma_2^2/\sigma_1^2$ 비율에 대한 신뢰 한계를 계산할 때처럼 첫 표본과 둘째 표본의 역할이 서로 바뀝니다. 둘째, 하한을 구하는 알고리즘은 새 함수 $H(x)$ 에 적용됩니다. 마지막으로, 그 결과로 얻은 한계를 뒤집어서 원하는 신뢰 상한을 구합니다.

분산 비율의 신뢰 하한을 구하는 다른 방법은 다음의 식으로 표현되는 재귀 관계를 사용하여 귀납적으로 계산하는 것입니다.

$$\rho_0^2 = 1$$

$$\rho_{i+1}^2 = \exp \left[\ln(c \hat{\rho}^2) - z_{\alpha/2} \sqrt{\frac{\hat{Y}_P(\rho_i) - g_1}{n_1 - 1} + \frac{\hat{Y}_P(\rho_i) - g_2}{n_2 - 1}} \right], i = 0, 1, 2, \dots$$

분산 비율의 신뢰 하한은 ρ_{j+1}^2 이며, $|\rho_{j+1}^2 - \rho_j^2| < \varepsilon$ 입니다. 여기서 $j > 0$ 고 ε 는 작은 값으로 선택됩니다(예: $\varepsilon = 10^{-6}$). 신뢰 상한을 구하려면 간단히 위에서 $-z_{\alpha/2}$ 를 $+z_{\alpha/2}$ 로 대체합니다.

귀납적인 절차는 근본적으로 $H(\rho^2/(c\hat{\rho}^2)) = 0$ 등식을 ρ^2 에 대해 구하는 반복적인 절차이기 때문에 신뢰 상한을 계산하는 알고리즘 2개는 동일합니다. 귀납적인 알고리즘은 사용하기가 더 쉬우므로 근 찾기 절차를 사용할 수 없을 때 유용한 대안입니다.

4. 시뮬레이션 연구 및 결과

본 문서에서는 Layard의 2분산 동일성 검정을 확장하여 분산 비율을 검정하는 절차를 도출했습니다. 당사는 이 절차를 ELTR(Extended Layard's Test for the Ratio)이라고 부릅니다. 다음 항목에서는 ELTR 절차에 기초한 CI의 소표본 속성에 대해 알아봅니다. 이를 위해서는 Bonett(2006)이 취한 일반적인 접근방식을 따릅니다.

ELTR 절차에 기초한 CI를 검정 L_{50} (Pan, 1999) 및 검정 W_{50} (Levene/Brown-Forsythe 검정)에 기초한 CI에 비교합니다. 연구 1에는 비교를 위해 기존 F 검정에 기초한 CI도 포함되었습니다. 데이터가 정규 분포를 따를 때 기존 F 검정이 최적임은 잘 알려져 있습니다. 검정 W_{50} 및 L_{50} 에 기초한 CI 계산에 대해서는 Pan(1999)에 제시되어 있습니다. F 검정에 기초한 CI 계산에 대해서는 여러 통계학 입문 교과서에서 확인할 수 있습니다.

Bonett(2006)에도 나와 있습니다.

각각 100,000번의 표본 추출 반복실험을 사용한 시뮬레이션 연구를 3차례 실시했습니다. 각 반복실험은 크기가 중간 이하인 독립 표본 2개로 구성됩니다. 각 표본은 대칭, 비대칭, 두꺼운 꼬리 및 얇은 꼬리 모집단을 포함한 속성이 알려진 모분포에서 추출되었습니다. 각 시뮬레이션과 연관된 표준 오차는 명목 신뢰 수준이 90%, 95% 및 99%일 때 각각 0.0009, 0.0007 및 0.0003입니다.

각 절차의 성능을 평가하기 위해, 분산 비율에 대해 얻어진 포함 확률과 모의 구간의 평균 너비를 보고합니다. 검정 W_{50} 과 연관된 구간 중 일부는 너비가 무한했습니다(Pan(1999)이 밝혀낸 가능성). 이런 경우에는 유한 구간의 평균 너비와 무한 너비를 가지는 구간의 백분율을 보고합니다. 모든 시뮬레이션은 Mathematica 소프트웨어 패키지 버전 8을 사용하여 실시되었습니다.

연구 1: 정규 데이터 포함 확률 비교

첫 번째 연구에서는 다양한 크기의 랜덤 표본을 정규 분포에서 추출합니다. 결과는 표 2에 제시되어 있습니다.

표 2 포함 확률과 평균 구간 너비 비교

$1 - \alpha$	n_1, n_2	측정	절차			
			F	ELTR	L_{50}	W_{50}
0.90	10, 10	포함 확률	0.898	0.918	0.913	0.921
		너비	3.72	5.06	4.72	8.03 (0.01%)
	30, 10	포함 확률	0.90	0.909	0.897	0.911
		너비	2.42	3.01	3.58	3.17
	25, 25	포함 확률	0.902	0.907	0.914	0.916
		너비	1.61	1.73	1.85	1.938
	50, 50	포함 확률	0.90	0.901	0.906	0.907
		너비	1.03	1.06	1.13	1.15
0.95	10, 10	포함 확률	0.949	0.963	0.958	0.964
		너비	4.90	7.72	6.52	497.24 (0.22%)
	30, 10	포함 확률	0.95	0.957	0.945	0.959
		너비	2.98	4.91	4.67	4.07
	25, 25	포함 확률	0.951	0.955	0.958	0.961
		너비	1.99	2.24	2.31	2.49
	50, 50	포함 확률	0.951	0.952	0.953	0.954
		너비	1.25	1.31	1.38	1.41

$1 - \alpha$	n_1, n_2	측정	절차			
			F	ELTR	L_{50}	W_{50}
0.99	10, 10	포함 확률	0.989	0.993	0.992	0.994
		너비	8.29	17.76	12.52	$> 10^4$ (8.8%)
	30, 10	포함 확률	0.99	0.992	0.986	0.994
		너비	4.26	15.76	8.26	6.77
	25, 25	포함 확률	0.99	0.992	0.992	0.993
		너비	2.86	3.66	3.43	4.03
	50, 50	포함 확률	0.99	0.991	0.991	0.991
		너비	1.71	1.89	1.92	2.02

질은 회색 줄은 각 절차에 대해 각 신뢰 수준($1 - \alpha$)과 각 표본 크기 조합(n_1, n_2)에서 얻어진 포함 확률을 나타냅니다. 구간 너비(너비)의 평균은 각 포함 확률 밑에 표시되어 있습니다. 조건에 대한 구간 중에 무한 구간이 있을 경우에는 유한 구간의 평균과 무한 구간의 백분율을 모두 보고합니다.

예상한 대로 F 절차와 연관된 CI가 가장 정확하고 정밀하다는 결과가 나왔습니다. F 절차를 사용하여 얻어진 포함 확률은 다른 절차와 연관된 포함 확률보다 목표값에 더 가깝습니다. 그리고 F 절차와 연관된 구간의 평균 너비는 다른 절차와 연관된 너비보다 더 작습니다. 그러나 표에는 또한 ELTR 및 L_{50} 절차를 사용하여 구성된 CI가 F 절차에 기초한 CI만큼 정확하고 정밀한 것으로 나와 있습니다.

검정 W_{50} 에 기초한 구간도 꽤 정확합니다. 그러나 W_{50} 구간은 매우 넓을 수 있으며 표본 크기에 따라 너비가 무한할 수도 있습니다. 두 표본에 모두 관측치가 10개 뿐이면 W_{50} 절차에 의해 생성된 구간 중 0.01% 이상의 너비가 무한합니다. 그리고 무한한 간격의 백분율은 목표 범위가 커질수록 증가합니다. 대부분의 조건에서 ELTR 및 L_{50} 구간의 평균 너비는 W_{50} 구간의 평균 너비보다 작습니다.

연구 2: 비정규 데이터 포함 확률 비교

두 번째 연구는 모집단이 정규 분포를 따르지 않을 때 ELTR, L_{50} 및 W_{50} 절차의 성능을 평가 및 비교하기 위해 고안되었습니다. 특이치가 절차의 성능에 미치는 영향을 평가하기 위해 오염 정규 분포도 포함되었습니다. 이 오염 분포는 CN(0.1, 3)으로 표시하여 관측치의 90%가

표준 정규 분포에서 추출된 한편 나머지 10%는 평균이 0이고 표준 편차가 3인 정규 모집단에서 추출되었음을 나타냅니다. 결과는 표 3에 제시되어 있습니다.

표 3 일부 비정규 모형의 포함 확률과 평균 구간 너비 비교. 명목 신뢰 수준은 $1 - \alpha = 0.95$ 입니다

분포 [γ] n_1, n_2	ELTR	L_{50}	W_{50}	분포 [γ] n_1, n_2	ELTR	L_{50}	W_{50}
균등 분포				$\chi^2(5)$			
[1.8]				[5.4]			
10, 10	0.971	0.971	0.966	10, 10	0.956	0.938	0.956
	<i>5.27</i>	<i>4.87</i>	<i>42.08</i> <i>(0.1%)</i>		<i>11.61</i>	<i>8.78</i>	<i>> 10⁴</i> <i>(2.6%)</i>
10, 30	0.964	0.961	0.957	10, 30	0.959	0.923	0.956
	<i>2.51</i>	<i>2.4</i>	<i>2.89</i>		<i>6.25</i>	<i>4.14</i>	<i>190.645</i> <i>(0.3%)</i>
25, 25	0.967	0.972	0.968	25, 25	0.956	0.944	0.954
	<i>1.43</i>	<i>1.79</i>	<i>1.88</i>		<i>3.66</i>	<i>2.92</i>	<i>3.26</i>
50, 50	0.959	0.962	0.959	50, 50	0.959	0.946	0.952
	<i>0.83</i>	<i>1.06</i>	<i>1.08</i>		<i>2.07</i>	<i>1.7</i>	<i>1.77</i>
베타(3, 3)				지수			
[2.5]				[9]			
10, 10	0.968	0.966	0.966	10, 10	0.947	0.916	0.95
	<i>6.26</i>	<i>5.59</i>	<i>254.62</i> <i>(0.1%)</i>		<i>20.99</i>	<i>14.47</i>	<i>> 10⁴</i> <i>(9.1%)</i>
10, 30	0.96	0.954	0.96	10, 30	0.954	0.896	0.953
	<i>3.14</i>	<i>2.76</i>	<i>3.71</i>		<i>10.46</i>	<i>6.19</i>	<i>> 10⁴</i> <i>(4.1%)</i>
25, 25	0.959	0.966	0.965	25, 25	0.956	0.931	0.951
	<i>1.81</i>	<i>2.06</i>	<i>2.18</i>		<i>6.09</i>	<i>4.13</i>	<i>5.48</i> <i>(0.008%)</i>
50, 50	0.957	0.959	0.958	50, 50	0.962	0.942	0.952

분포 [γ] n_1, n_2	ELTR	L_{50}	W_{50}	분포 [γ] n_1, n_2	ELTR	L_{50}	W_{50}
	1.06	1.23	1.26		3.18	2.24	2.38
Laplace 분포				$\chi^2(1)$			
[6]				[15]			
10, 10	0.946	0.935	0.961	10, 10	0.928	0.889	0.947
	13.47	10.45	$> 10^4$ (3.0%)		55.09	37.4	$> 10^5$ (25.1%)
10, 30	0.947	0.919	0.957	10, 30	0.943	0.882	0.956
	6.78	4.82	$> 10^4$ (0.4%)		18.71	11.14	$> 10^6$ (25.7%)
25, 25	0.945	0.94	0.952	25, 25	0.952	0.925	0.954
	4.00	3.372	3.86		10.97	6.84	$> 10^4$ (0.4%)
50, 50	0.952	0.949	0.955	50, 50	0.958	0.936	0.951
	2.19	1.91	1.99		5.08	3.31	3.75 (0.001%)
$t(5)$				로그 정규 분포			
[9]				[113.9]			
10, 10	0.957	0.946	0.965	10, 10	0.923	0.876	0.955
	11.07	8.81	$> 10^3$ (2.0%)		59.22	46.15	$> 10^5$ (23.0%)
10, 30	0.957	0.93	0.959	10, 30	0.949	0.866	0.958
	6.06	4.24	$> 10^3$ (0.7%)		29.13	17.67	$> 10^6$ (31.6%)
25, 25	0.954	0.948	0.96	25, 25	0.947	0.917	0.965
	3.54	2.93	4.86 (0.01%)		16.21	8.73	$> 10^4$ (2.4%)
50, 50	0.954	0.947	0.954	50, 50	0.955	0.928	0.96
	2.10	1.71	1.77 (0.003%)		8.62	4.11	164.38 (0.2%)

분포 [γ] n_1, n_2	ELTR	L_{50}	W_{50}	분포 [γ] n_1, n_2	ELTR	L_{50}	W_{50}
1/2 정규 분포				CN(0.1, 3)			
[3.9]				[8.3]			
10, 10	0.956	0.942	0.954	10, 10	0.977	0.965	0.979
	10.41	7.89	$> 10^4$ (1.5%)		12.64	9.52	$> 10^4$ (4.9%)
10, 30	0.959	0.93	0.954	10, 30	0.981	0.952	0.979
	5.18	3.64	13.00 (0.02%)		7.82	4.71	944.68 (1.1%)
25, 25	0.959	0.952	0.959	25, 25	0.982	0.972	0.981
	3.01	2.62	2.88		4.63	3.22	3.71
50, 50	0.96	0.951	0.954	50, 50	0.983	0.972	0.978
	1.69	1.54	1.59		2.64	1.83	1.91

질은 회색 줄은 각 절차, 모분포 및 표본 크기 조합에 대해 얻어진 포함 확률을 나타냅니다. 구간 너비의 평균은 각 포함 확률 밑에 표시되어 있습니다. 조건에 대한 구간 중에 무한 구간이 있을 경우에는 유한 구간의 평균과 무한 구간의 백분율을 모두 보고합니다. 각 모분포의 첨도(γ)는 대괄호 안에 표시되었습니다.

꼬리가 얇은 대칭 분포에 대한 결과는 세 가지 방법에서 모두 유사하게 보수적인 포함 확률이 생성됨을 나타냅니다. 그러나 ELTR 및 L_{50} 구간은 표본이 작을 경우 W_{50} 구간보다 더 정밀합니다. 예를 들어 표본을 모수가 (3, 3)인 베타 분포에서 추출하면 ELTR 및 L_{50} 구간에 대해 얻어지는 포함 확률은 W_{50} 구간만큼 정확하지만, W_{50} 구간이 일관되게 더 넓습니다.

ELTR 및 W_{50} 구간도 꼬리가 두꺼운 대칭 분포치고는 약간 보수적이지만, L_{50} 구간은 개방적입니다. L_{50} 구간은 설계가 불균형일 때 더욱 더 개방적입니다. 예를 들어 Laplace 분포에서 크기가 10 및 30인 표본을 추출할 때 L_{50} 구간에 대해 얻어지는 포함 확률은 0.919입니다. 그리고 자유도가 5인 t 분포에서 크기가 같은 표본을 추출할 때 L_{50} 구간에 대해 얻어지는 포함 확률은 0.930입니다.

L_{50} 구간은 매우 치우친 두꺼운 꼬리 분포에서 작은 표본을 추출할 때도 상당히 개방적입니다. 예를 들어 표본을 대수 정규 분포에서 추출할 때 얻어지는 포함 확률은 0.866까지 떨어질 수 있습니다. 이런 분포에서는 W_{50} 방법이 세 가지 방법 중에서 가장 개방적이지 않은 방법입니다. 그러나 너비가 무한한 W_{50} 구간이 너무 많습니다. 예를 들어 자유도가 1인 카이

제곱 분포($\chi^2(1)$)에서 표본을 추출하면 25%가 넘는 W_{50} 구간의 너비가 무한할 수 있습니다. ELTR 구간은 약간 덜 정확하지만, 훨씬 더 좁기 때문에 W_{50} 구간보다 더 유용합니다.

마지막으로, 특이치는 세 절차에 모두 악영향을 미칩니다. L_{50} 방법이 악영향을 가장 적게 받는데, 그 이유는 L_{50} 방법이 특이치가 검정 W_{50} 에 미치는 영향을 줄이기 위해 고안되었기 때문일 수 있습니다(Pan, 1999). 표본을 오염 정규 분포 $CN(0.1, 3)$ 에서 추출할 때 ELTR 및 W_{50} 절차에 대해 얻어지는 포함 확률의 최소값은 0.977입니다. 해당 구간은 표본 크기가 커져도 개선 정도는 적다는 추가적인 시뮬레이션 결과(제시되지 않음)도 나타났습니다.

연구 3: 등첨도 가정에 대한 민감도

마지막 연구에서는 ELTR 절차를 도출하는 데 사용된 등첨도 가정에 대한 해당 절차의 민감도를 조사합니다. 모집단의 첨도가 같지 않을 때, 즉 $\gamma_1 \neq \gamma_2$ 일 때 ELTR 절차의 성능을 조사합니다. L_{50} 및 W_{50} 절차도 모집단이 유사하다는 가정 하에서 도출되었기 때문에 이 연구에 포함되었습니다. 모집단의 첨도가 같지 않다면 이 유사성 가정은 만족되지 않습니다. 결과는 표 4에 제시되어 있습니다.

표 4 등첨도 가정에 대한 ELTR 절차의 민감도. 명목 신뢰 수준은 $1 - \alpha = 0.95$ 입니다.

분포 1, 분포 2 [γ_1, γ_2] n_1, n_2	ELTR	L_{50}	W_{50}	분포 1, 분포 2 [γ_1, γ_2] n_1, n_2	ELTR	L_{50}	W_{50}
베타(3, 3), 정규 분포 [2.5, 3]				정규 분포, CN(0.9, 3) [3, 8.3]			
10, 10	0.964	0.961	0.964	10, 10	0.955	0.948	0.951
	0.27	0.23	204.50 (0.20%)		6.88	5.16	> 10 ⁴ (1.89%)
30, 10	0.946	0.939	0.946	30, 10	0.941	0.91	0.942
	0.16	0.17	0.15		5.26	3.77	3.20
10, 30	0.966	0.956	0.967	10, 30	0.961	0.95	0.958
	0.14	0.11	0.17		4.26	2.40	630.42 (1.10%)
50, 50	0.951	0.95	0.949	50, 50	0.936	0.91	0.907
	0.04	0.05	0.05		1.27	1.11	1.19

분포 1, 분포 2 [γ_1, γ_2] n_1, n_2	ELTR	L_{50}	W_{50}	분포 1, 분포 2 [γ_1, γ_2] n_1, n_2	ELTR	L_{50}	W_{50}
정규 분포, Laplace 분포 [3, 6]				1/2 정규 분포, $\chi^2(5)$ [3.9, 5.4]			
10, 10	0.941	0.935	0.947	10, 10	0.956	0.94	0.954
	6.67	5.17	$> 10^6$ (2.90%)		0.42	0.32	304.41 (2.60%)
30, 10	0.912	0.888	0.914	30, 10	0.954	0.918	0.949
	5.06	3.85	3.21		0.33	0.22	0.20
10, 30	0.963	0.943	0.955	10, 30	0.962	0.934	0.958
	3.33	2.25	$> 10^3$ (0.40%)		0.23	0.15	3.28 (0.30%)
50, 50	0.935	0.894	0.889	50, 50	0.955	0.941	0.945
	0.98	1.04	1.12		0.07	0.06	0.07
정규 분포, 1/2 정규 분포 [3, 3.9]				$\chi^2(5)$, 지수 [5.4, 9]			
10, 10	0.956	0.948	0.957	10, 10	0.938	0.914	0.94
	28.16	20.65	$> 10^4$ (1.50%)		211.17	137.88	$> 10^6$ (9.10%)
30, 10	0.946	0.924	0.947	30, 10	0.928	0.875	0.929
	20.59	14.83	12.78		194.70	93.02	83.02
10, 30	0.961	0.946	0.962	10, 30	0.968	0.93	0.954
	14.06	9.37	49.11 (0.02%)		102.35	55.29	$> 10^5$ (3.90%)
50, 50	0.953	0.95	0.952	50, 50	0.95	0.92	0.923
	4.32	4.16	4.33		29.64	23.37	25.54

질은 회색 줄은 각 절차, 모분포 조합(분포 1, 분포 2) 및 표본 크기 조합에 대해 얻어진 포함 확률을 나타냅니다. 구간 너비의 평균은 각 포함 확률 밑에 표시되어 있습니다. 조건에 대한

구간 중에 무한 구간이 있을 경우에는 유한 구간의 평균과 무한 구간의 백분율을 모두 보고합니다. 각 모분포의 첨도(γ_1, γ_2)는 대괄호 안에 표시되었습니다.

일반적으로 ELTR 절차는 표본이 충분히 클 때 같지 않은 첨도에 악영향을 받지 않는 것으로 보입니다. 그러나 설계가 불균형이고 꼬리가 더 두꺼운 분포에서 더 작은 표본을 추출할 때 얻어지는 포함 확률은 개방적입니다. 꼬리가 더 두꺼운 분포에서 더 큰 표본을 추출하면 더 나은 포함 확률이 얻어집니다.

표본 크기가 충분히 크면 L_{50} 및 W_{50} 구간도 같지 않은 첨도로 인해 초래되는 분포의 비유사성에 일반적으로 로버스트한 것으로 보입니다. 그러나 표본을 정규 분포와 Laplace 분포, 또는 정규 분포와 오염 정규 분포에서 추출할 때 L_{50} 및 W_{50} 구간의 포함 확률은 표본이 50개 정도로 커도 안정적이지 않습니다.

L_{50} 구간은 일반적으로 ELTR 및 W_{50} 구간보다 더 개방적입니다. 세 가지 경우에 L_{50} 구간에 대해 얻어진 포함 확률은 0.90 미만입니다. 반대로, W_{50} 구간에 대해 얻어진 포함 확률 중에는 하나만 0.90 미만입니다. ELTR 구간에 대해 얻어진 가장 낮은 포함 확률은 0.912입니다.

이전 연구(표 3)에서는 세 절차에서 모두 오염 정규 분포 CN(0.1, 3)에서 표본 2개를 모두 추출할 경우 훨씬 더 보수적인 구간이 생성되는 것으로 나타났습니다. 본 연구에서는 CN(0.1, 3)에서 표본을 하나만 추출할 때 세 절차가 모두 훨씬 더 효과적인 것으로 나타났습니다. 그러나 표본 크기가 50개로 증가하면 L_{50} 및 W_{50} 구간의 성능이 크게 저하되는 것으로 보입니다.

5. 예제

다음 항목에서는 네 가지 절차(F, ELTR, L_{50} 및 W_{50})를 모두 Pan(1999)에서 얻은 데이터 집합에 적용합니다. Ott(1993, 352페이지)에서는 데이터에 대해 다음과 같이 설명합니다.

철광산에서 일하는 화학자가 철광석 1파운드당 산화철 함유량(무게, 온스 단위)의 분산이 파운드당 산화철 함유량의 평균이 증가함에 따라 같이 증가하는 경향이 있다고 의심합니다. 이 이론을 검증하기 위해, 2개 위치에서 각각 1파운드의 철광석 표본 10개를 선택했습니다. 위치 1은 다른 위치(위치 2)보다 산화철 평균 함유량이 훨씬 더 많습니다. 철광석 표본에 함유된 산화철의 양은 아래에 나와 있습니다.

위치 1	8.1	7.4	9.3	7.5	7.1	8.7	9.1	7.9	8.4	8.8
위치 2	3.9	4.4	4.7	3.6	4.1	3.9	4.6	3.5	4.0	4.2

네 가지 방법을 사용하여 계산한 $\sigma_2/\sigma_1 = 1/\rho$ 에 대한 95% CI는 아래 표에 제시되어 있습니다.

절차	95% CI
F	(0.262, 1.055)
ELTR	(0.277, 0.924)
L_{50} (Pan)	(0.295, 0.938)
W_{50} (Levene/Brown-Forsythe)	(0.237, 0.908)

6. 결론

시뮬레이션에는 ELTR 절차에 기초한 CI가 일반적으로 검정 L_{50} 및 W_{50} 에서 도출된 CI만큼 정확한 것으로 나타났습니다. 그러나 ELTR 및 L_{50} 구간은 대부분의 분포에 대해 W_{50} 구간보다 더 정밀합니다. W_{50} 구간은 매우 치우친 두꺼운 꼬리 분포에서 작은 표본을 추출할 때 ELTR 구간과 L_{50} 구간보다 더 정확한 경향이 있습니다. 그러나 이 같은 장점은 일반적으로 상당한 정밀도 손실에 의해 상쇄됩니다. 그 결과로 얻어지는 W_{50} 구간은 일반적으로 너무 넓고 너비가 무한할 가능성이 큼니다.

설계 의도대로 L_{50} 구간은 W_{50} 구간의 정밀도를 개선합니다. 그러나 치우친 모집단에 대해 L_{50} 구간은 지나치게 개방적인 포함 확률을 생성할 만큼 짧습니다. 반대로 ELTR 구간은 일반적으로 더 안정적입니다. ELTR 구간은 일반적으로 너무 길거나 짧지 않으므로 포함 확률은 일반적으로 너무 보수적이거나 너무 개방적이지 않습니다. 그러므로 ELTR 절차는 대부분의 실제 용도에 가장 적합한 절차인 것으로 보입니다.

ELTR 구간은 검정 L_{50} 또는 검정 W_{50} 에 기초한 구간보다 계산하기가 좀 더 힘듭니다. 그러나 일반적으로 (W_{50} 구간에 비한) 정밀도 개선과 (L_{50} 구간에 비한) 정확도 개선은 추가적인 계산 노력보다 훨씬 더 가치가 큼니다. ELTR 절차는 Minitab Statistical Software 17 릴리스에 2-

표본 분산 분석의 일부로 구현되었으며, 이 소프트웨어에서는 해당 절차를 Bonett의 절차라고 합니다.

향후 연구에서는 Layard의 합동 첨도 추정량을 다음 식으로 표현되는 Bonett의 더 로버스트한 버전으로 대체할 때 다중 표본 설계에서 Layard 검정의 소표본 속성을 조사하는 방안을 고려할 수 있을 것입니다.

$$\hat{\gamma}_A = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - m_i)^4}{\left[\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - m_i)^2 \right]^2} \sum_{i=1}^k n_i$$

여기서 m_i 는 절사 비율이 $1/[2(n_i - 4)^{1/2}]$ 이고 $i = 1, \dots, k$ 인 표본 i 의 절사 평균입니다.

또한 로그 변환 표본 분산의 점근 분산에 대한 Shoemaker의 근사를 사용하면 유용할 수 있습니다.

마지막으로 Bonett(2006)이 제안한 구간은 CI로 적합하지 않지만, 두 분산의 동일성 검정에 대한 합격 영역이라고 해석하면 대부분의 분포에서 매우 정확하고 정밀합니다. 이런 합격 영역은 다중 분산을 비교하는 그래픽 절차의 기초로 사용하기에 매우 적합합니다. Hochberg, Weiss & Hart(1982)는 등평균 검정을 위한 유사한 절차를 제안했습니다. 해당 절차는 Minitab Statistical Software 17 릴리스의 등분산 검정 분석의 일부로 구현되었으며, 이 소프트웨어에서는 해당 절차를 다중 비교 절차라고 합니다.

7. 부록

부록 A: 결과 1의 증명

2절의 표기 및 가정을 사용하여 주어진 $\rho = \sigma_1/\sigma_2$ 에 대해 $X_j = \rho Y_{2j}$ 라고 하면

$$\text{Var}(X_j) = \rho^2 \text{Var}(Y_{2j}) = \rho^2 \sigma_2^2 = \sigma_1^2 = \text{Var}(Y_{1j})$$

이고

$$E \left(X_j - \mu_{X_j} \right)^4 / \sigma_{X_j}^4 = \rho^4 E(Y_{2j} - \mu_2) / (\rho^4 \sigma_2^4) = E(Y_{2j} - \mu_2) / \sigma_2^4 = \gamma$$

입니다. 가정에 의해 $E(Y_{1j} - \mu_1) / \sigma_1^4 = \gamma$ 이므로, 두 표본 Y_{1j} 및 $X_j = \rho Y_{2j}$ 의 모집단은 분산이 σ_1^2 으로 같고 첨도는 γ 로 같다고 유추할 수 있습니다. Layard(1973)에 의해 두 표본 Y_{1j} 및 X_j 에 기초한 γ 의 일치 합동 첨도 추정량은 필요에 따라

$$\begin{aligned} \hat{\gamma}' &= (n_1 + n_2) \frac{\sum_{j=1}^{n_1} (Y_{1j} - \bar{Y}_1)^4 + \sum_{j=1}^{n_2} (X_j - \bar{X})^4}{[(n_1 - 1)S_1^2 + (n_2 - 1)S_X^2]^2} \\ &= (n_1 + n_2) \frac{\sum_{j=1}^{n_1} (Y_{1j} - \bar{Y}_1)^4 + \rho^4 \sum_{j=1}^{n_2} (Y_{2j} - \bar{Y}_2)^4}{[(n_1 - 1)S_1^2 + \rho^2(n_2 - 1)S_2^2]^2} = \hat{\gamma}_P(\rho) \end{aligned}$$

으로 표현됩니다.

부록 B: 결과 2의 증명

이미 확인한 바와 같이, T_2 를 기준으로 $\rho = \sigma_1/\sigma_2$ 에 대해 설정한 근사 $(1 - \alpha)100$ 퍼센트 신뢰 구간은 다음 식으로 표현됩니다.

$$\left\{ \rho \in (0, \infty) : (\ln \rho^2 - \ln(c\hat{\rho}^2))^2 - z_{\alpha/2}^2 \left(\frac{\hat{Y}_P(\rho) - k_1}{n_1 - 1} + \frac{\hat{Y}_P(\rho) - k_2}{n_2 - 1} \right) \leq 0 \right\}$$

합동 척도 추정량은 다음 식으로 표현되는 개별 표본에 대한 척도 추정량을 기준으로 표현할 수 있습니다.

$$\hat{y}_i = n_i \frac{\sum_{j=1}^{n_i} (Y_{ij} - m_i)^4}{[(n_i - 1)S_i^2]^2}, i = 1, 2$$

더 구체적으로 말해, $t = \rho/\hat{\rho}$ 라고 하면 다음이 성립합니다.

$$\hat{Y}_P(\rho) = (n_1 + n_2) \frac{\sum_{j=1}^{n_1} (Y_{1j} - m_1)^4 + \rho^4 \sum_{j=1}^{n_2} (Y_{2j} - m_2)^4}{[(n_1 - 1)S_1^2 + \rho^2(n_2 - 1)S_2^2]^2} = (n_1 + n_2) \frac{\hat{y}_1 K^2/n_1 + \hat{y}_2 t^4/n_2}{(K + t^2)^2}$$

여기서 $K = (n_1 - 1)/(n_2 - 1)$ 입니다.

결과적으로 제곱 표준 오차 항은 다음 식으로 나타낼 수 있습니다.

$$\frac{\hat{Y}_P(\rho) - k_1}{n_1 - 1} + \frac{\hat{Y}_P(\rho) - k_2}{n_2 - 1} = A \frac{\hat{y}_1 K^2/n_1 + \hat{y}_2 t^4/n_2}{(K + t^2)^2} - B$$

여기서,

$$A = \frac{(n_1 + n_2)(n_1 + n_2 - 2)}{(n_1 - 1)(n_2 - 1)}, B = \frac{k_1}{n_1 - 1} + \frac{k_2}{n_2 - 1}$$

따라서 $r^2 = \rho^2/(c\hat{\rho}^2)$ 이라고 하면 다음이 성립함을 쉽게 알 수 있습니다.

$$\begin{aligned} & (\ln \rho^2 - \ln(c\hat{\rho}^2))^2 - z_{\alpha/2}^2 \left(\frac{\hat{Y}_P(\rho) - k_1}{n_1 - 1} + \frac{\hat{Y}_P(\rho) - k_2}{n_2 - 1} \right) \\ &= (\ln r^2)^2 - z_{\alpha/2}^2 \left(A \frac{\hat{y}_1 K^2/n_1 + \hat{y}_2 c^2 r^4/n_2}{(K + c r^2)^2} - B \right) \end{aligned}$$

논리적으로, T_2 에 기초한 $\rho = \sigma_1/\sigma_2$ 에 대한 $(1 - \alpha)100$ 퍼센트 신뢰 구간은 다음 식으로 나타낼 수 있습니다.

$$\hat{\rho}\sqrt{c} \left\{ r \in (0, \infty) : H(r^2) \leq 0 \right\}$$

또는 이와 동등하게 $\rho^2 = \sigma_1^2/\sigma_2^2$ 에 대한 신뢰 구간을 다음과 같이 표현할 수 있습니다.

$$c\hat{\rho}^2 \left\{ r \in (0, \infty) : H(r) \leq 0 \right\}$$

여기서,

$$H(x) = (\ln x)^2 - z_{\alpha/2}^2 se^2(cx), x > 0$$

이고, 다음이 성립합니다.

$$se^2(x) = A \frac{\hat{\gamma}_1 K^2/n_1 + \hat{\gamma}_2 x^2/n_2}{(K+x)^2} - B$$

부록 C: 결과 3의 증명

$H(x)$ 가 양의 실선에서 연속적이고 $H(0) = H(+\infty) = +\infty$ 고 $H(1) < 0$ 임은 쉽게 확인됩니다. 중간 값 정리에 의해, 함수 $H(x)$ 는 구간 $(0, 1)$ 과 구간 $(0, +\infty)$ 에서 각각 근을 하나 이상 가집니다. 따라서 함수 $H(x)$ 에 근이 정확히 2개 있으면 하나는 1 미만이고 다른 하나는 1을 초과합니다. 이 함수는 위로 열리므로 r 가 근 사이에 있을 경우 부등식 $H(r) \leq 0$ 가 충족됩니다. 이 2개의 근은 $\rho^2/(c\hat{\rho}^2)$ 에 대한 CI의 끝점을 정의합니다. 따라서 $x_L < 1 < x_U$ 가 2개의 근이라고 하면, 결과 2에 의해 분산 ρ^2 의 비율에 대한 신뢰 하한은 $c\hat{\rho}^2 x_L$ 로 계산되고 표준 편차 비율의 신뢰 하한은 $\hat{\rho}\sqrt{cx_L}$ 로 구해집니다. 마찬가지로, 분산 비율의 신뢰 상한은 $c\hat{\rho}^2 x_U$ 고, 표준 편차 비율의 신뢰 상한은 $\hat{\rho}\sqrt{cx_U}$ 입니다.

반면에 함수 $H(x)$ 에 근이 3개 이상이면 함수가 위로 열리는 보수적인 근 사이에 r 가 있을 경우에 부등식 $H(r) \leq 0$ 이 충족됩니다. 따라서 신뢰 구간은 겹치지 않는 구간의 합집합입니다.

8. 참고 문헌

- Balakrishnan, N. & Ma, C. W. (1990). A Comparative Study of Various Tests for the Equality of Two Population Variances. *Journal of Statistical Computation and Simulation*, 35, 41-89.
- Bonett D. G. (2006). Robust Confidence Interval for a Ratio of Standard Deviations. *Applied Psychological Measurements*, 30, 432-439.
- Boos, D. D. & Brownie, C. (1989). Bootstrap Methods for Testing Homogeneity of Variances. *Technometrics*, 31, 69-82.
- 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.

- Fligner, M. A. & Killeen, T. J. (1976). Distribution-Free Two-Sample Tests for Scale. *Journal of the American Statistical Association*, 71, 210-213.
- Hall, P. & Padmanabhan, A. R. (1997). Adaptive Inference for the Two-Sample Scale Problem. *Technometrics*, 39, 412-422.
- Hochberg, Y., Weiss, G. & Hart S.,(1982). On Graphical Procedures for Multiple Comparisons. *Journal of the American Statistical Association*, 77, 767-772.
- 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," in I. Olkin, ed., *Contributions to Probability and Statistics*, Palo Alto, CA: Stanford University Press, 278-292.
- Lim, T.-S. & Loh, W.-Y. (1996). A Comparison of Tests of Equality of Variances. *Computational Statistics and Data Analysis*, 22, 287-301.
- Ott, L. (1993). *An Introduction to Statistical Methods and Data Analysis*, Belmont, CA: Duxbury Press.
- Pan, G. (1999). On a Levene Type Test for Equality of Two Variances. *Journal of Statistical Computation and Simulation*, 63, 59-71.
- Shoemaker, L. H. (2003). Fixing the F Test for Equal Variances. *The American Statistician*, 57, 105-114.
- Wolfram, S. (1999). *The Mathematica Book*, 4th ed. Wolfram Media/Cambridge University Press.

© 2015, 2017 Minitab Inc. All rights reserved.

Minitab®, Quality. Analysis. Results.® and the Minitab® logo are all registered trademarks of Minitab, Inc., in the United States and other countries. See minitab.com/legal/trademarks for more information.