Levene's Test는 두 개 이상의 집단의 분산이 동일한지 여부를 검정하는 통계적 방법입니다. 이 검정은 특히 데이터가 정규분포를 따르지 않을 때, 또는 집단 간 분산의 동질성 가정이 불확실할 때 유용합니다. Levene's Test는 각 데이터 점에서 중앙값을 빼고 절대값을 취한 변환된 데이터의 분산을 비교하여 분산의 동질성을 검정합니다.

**Levene's Test 절차**

**1. 가설 설정**

* **귀무 가설 (H0)**: 모든 집단의 분산이 같다.
* **대립 가설 (H1)**: 적어도 하나의 집단의 분산이 다르다.

**2. 변환된 데이터 계산**

Levene's Test에서는 각 데이터 점에서 그 집단의 중앙값을 빼고 절대값을 취한 변환된 데이터를 사용합니다. 이를 통해 비정규 분포 데이터에도 적용 가능합니다.

1. 각 데이터 점 𝑌𝑖𝑗에서 해당 집단의 중앙값 𝑀𝑖를 뺍니다:

𝑍𝑖𝑗 = ∣𝑌𝑖𝑗−𝑀𝑖∣

여기서 𝑌𝑖𝑗는 𝑖번째 집단의 𝑗번째 관측값, 𝑀𝑖는 𝑖번째 집단의 중앙값입니다.

1. 변환된 데이터 𝑍𝑖𝑗의 평균을 구합니다:

=

여기서 는 𝑖번째 집단의 표본 크기입니다.

**3. ANOVA (분산 분석) 수행**

변환된 데이터 𝑍𝑖𝑗에 대해 ANOVA를 수행합니다.

1. **총 변동(Total Sum of Squares, SST)**:

𝑆𝑆𝑇 =

여기서 는 전체 평균입니다.

1. **집단 간 변동(Between Sum of Squares, SSB)**:

𝑆𝑆𝐵 =

1. **집단 내 변동(Within Sum of Squares, SSW)**:

𝑆𝑆𝑊 =

1. **F-통계량**:

𝐹 = ​

여기서 𝑘는 집단의 수, 𝑁은 전체 표본 크기입니다.

**4. p-값 계산 및 결론 도출**

F-통계량을 통해 p-값을 계산하고, 이를 사전에 정한 유의 수준 (예: 0.05)과 비교합니다. p-값이 유의 수준보다 작으면 귀무 가설을 기각하고, 적어도 하나의 집단의 분산이 다르다고 결론내립니다.

**Levene's Test의 장점**

* **비정규성에 강건**: 데이터가 정규분포를 따르지 않아도 사용 가능합니다.
* **유연성**: 여러 집단 간의 분산 비교에도 사용될 수 있습니다.

**요약**

Levene's Test는 각 집단의 분산이 동일한지 여부를 검정하는 데 유용한 방법으로, 특히 데이터가 정규분포를 따르지 않을 때 사용됩니다. 절차는 중앙값을 기준으로 변환된 데이터의 분산을 비교하고, ANOVA를 통해 분산의 동질성을 검정하는 것입니다. p-값을 통해 결론을 도출하며, 비정규 분포에서도 신뢰할 수 있는 결과를 제공합니다.

**예시 상황**

세 개의 학급이 있고, 각 학급의 학생들이 시험을 보았습니다. 우리는 각 학급의 시험 점수 분산이 동일한지 확인하고자 합니다. 다음은 각 학급의 시험 점수입니다:

* **학급 1**: 85, 90, 88, 92, 86
* **학급 2**: 78, 82, 84, 79, 81
* **학급 3**: 92, 95, 91, 89, 94

**1. 가설 설정**

* **귀무 가설 (H0)**: 모든 학급의 분산이 같다.
* **대립 가설 (H1)**: 적어도 한 학급의 분산이 다르다.

**2. 중앙값 계산 및 데이터 변환**

각 학급의 중앙값을 구한 후, 각 데이터 점에서 중앙값을 뺀 절대값을 계산합니다.

**중앙값 계산**

* 학급 1: 중앙값 = 88
* 학급 2: 중앙값 = 81
* 학급 3: 중앙값 = 92

**데이터 변환**

* 학급 1: ∣85−88∣,∣90−88∣,∣88−88∣,∣92−88∣,∣86−88∣=3,2,0,4,2
* 학급 2: ∣78−81∣,∣82−81∣,∣84−81∣,∣79−81∣,∣81−81∣=3,1,3,2,0
* 학급 3: ∣92−92∣,∣95−92∣,∣91−92∣,∣89−92∣,∣94−92∣=0,3,1,3,2

**3. 변환된 데이터의 평균 계산**

각 학급의 변환된 데이터의 평균을 계산합니다.

* 학급 1: 평균 = = = 2.2
* 학급 2: 평균 = = = 1.8
* 학급 3: 평균 = = = 1.8

**4. ANOVA (분산 분석) 수행**

변환된 데이터에 대해 ANOVA를 수행하여 F-통계량을 계산합니다.

**총 변동 (SST)**

𝑆𝑆𝑇 =

여기서 는 전체 변환된 데이터의 평균입니다.

전체 변환된 데이터:

* 학급 1: 3, 2, 0, 4, 2
* 학급 2: 3, 1, 3, 2, 0
* 학급 3: 0, 3, 1, 3, 2

전체 평균 :

= = = 1.

𝑆𝑆𝑇=++++++++++++++

이를 계산하면 𝑆𝑆𝑇=21.33입니다.

**집단 간 변동 (SSB)**

𝑆𝑆𝐵 =

𝑆𝑆𝐵 = 5+5+5

이를 계산하면 𝑆𝑆𝐵=0.8입니다.

**집단 내 변동 (SSW)**

𝑆𝑆𝑊 = 𝑆𝑆𝑇−𝑆𝑆𝐵 = 21.33−0.8 = 20.53

**F-통계량 계산**

𝐹=

여기서 𝑘=3이고 𝑁=15입니다.

𝐹==≈0.234

**5. p-값 계산 및 해석**

F-통계량과 자유도를 사용하여 p-값을 계산합니다. 이를 F-분포표를 통해 얻거나 소프트웨어를 사용하여 계산할 수 있습니다. 만약 p-값이 0.05보다 작으면 귀무 가설을 기각합니다.

이 예에서는 p-값이 0.05보다 클 가능성이 큽니다(정확한 값을 계산하려면 소프트웨어를 사용해야 함). 따라서, 우리는 귀무 가설을 기각하지 않고, 세 학급의 분산이 동일하다고 결론 내릴 수 있습니다.

**요약**

1. 각 학급의 중앙값을 구하고, 각 데이터 값에서 중앙값을 뺀 절대값을 계산합니다.
2. 변환된 데이터의 평균을 계산합니다.
3. 변환된 데이터를 사용하여 ANOVA를 수행하고 F-통계량을 계산합니다.
4. F-통계량과 p-값을 사용하여 귀무 가설을 기각할지 여부를 결정합니다.

이 과정을 통해 Levene's Test를 수행하고, 세 학급의 분산이 동일한지 여부를 검정할 수 있습니다.