**분산분석(ANOVA)**

**분산분석(ANalysis Of VAriance) : F-검정통계량**

**Theory**

앞서 학습한 T-검정이 두 집단 간의 평균 차이를 비교하는 통계분석 방법이라면, 분산분석은 두 개 이상의 집단에서 그룹 평균 간 차이를 그룹 내 변동에 비교하여 살펴보는 통계 분석 방법입니다. 즉, 두 개이상 집단들의 평균 간 차이에 대한 통계적 유의성을 검증하는 방법입니다. 분산분석은 독립변수의 개수에 따라 일원배치, 이원배치, 다원배치로 나누어집니다.

**일원배치 분산분석(One Way-ANOVA)**

분산분석에서 반응값에 대한 하나의 범주형 변수의 영향을 알아보기 위해 사용되는 검증 방법입니다. 각 표본의 수는 같지 않아도 되며, F-검정 통계량을 사용합니다. 일원배치 분산분석은 아래 내용 가정 하에 이루어집니다.

* 각 집단의 측정치는 서로 독립적이다. -> stats.chi2\_contingency
* 각 집단의 측정치는 정규분포를 따른다. -> stats.shapiro
* 각 집단의 측정치의 분산은 같다. -> stats.levene
* 귀무가설 : 각 집단 간의 모평균의 차이는 없다.
* 대립가설 : 각 집단 간의 모평균은 적어도 한 집단에서 통계적으로 유의한 차이가 존재한다.

**◦ 분산분석표**

테이블이(가) 표시된 사진

자동 생성된 설명

**◦ 사후검정**

사후검정이란 분산분석의 결과 귀무가설이 기각되어 적어도 한 집단에서 평균의 차이가 있음이 통계적으로 증명되었을 경우, 어떤 집단들에 대해 평균의 차이가 존재하는지를 알아보기 위해 실시하는 분석이다. 귀무가설을 아래와 같이 설정한 뒤, **두 집단씩 짝을 지어** 각각 다중비교를 수행한다.

* 귀무가설 : 각 집단 간의 평균은 같다.
* 대립가설 : 각 집단 간의 평균은 통계적으로 유의한 차이가 있다.

**이원배치 분산분석(Two Way-ANOVA)**

분산분석에서 반응값에 대해 두 개의 범주형 변수 A, B의 영향을 알아보기 위해 사용되는 검증 방법이다. 이원배치 분산분석을 수행하기 위해선 두 독립변수 A, B 사이에 상관관계가 있는지를 살펴보는 **교호작용**(두 독립변수의 범주들의 조합으로 인해 반응변수에 미치는 특별한 영향)에 대한 검증이 반드시 진행되어야 한다.

이원배치 분산분석을 수행하기 위해선 정규성, 등분산성이 보장되어야 하며, 교호작용효과가 없어야 한다.

이원배치 분산분석에서는 **두 개의 독립변수값에 따르는 데이터의 주효과와 상호작용효과에 대한 검정을 수행한다.** 주효과는 각각의 독립변수가 종속변수에 미치는 효과를 의미하며, 교호작용효과는 여러 독립변수들의 조합이 종속변수에 주는 영향을 의미한다. 즉, 교호작용효과 검정은 한 독립변수가 종속변수에 미치는 영향이 다른 독립변수의 수준에 따라서 달라지는지를 분석하는 것이다.

두 독립변수 간의 상관성이 있을 경우, 교호작용효과가 있는 것이다. 교호작용효과가 없을 경우, 주효과 검정을 진행한다. **교호작용효과가 있을 경우 애초에 검증이 무의미하다. 사실상 교호작용효과만 검정해도 된다.**

* 귀무가설 : A 변수에 따른 종속변수의 값에는 차이가 없다. (주효과 검정)

B 변수에 따른 종속변수의 값에는 차이가 없다. (주효과 검정)

A와 B 변수 간에 상호작용 효과가 없다. (교호작용효과 검정)

* 대립가설 : A 변수에 따른 종속변수의 차이가 있다.

B 변수에 따른 종속변수의 차이가 있다.

A와 B 변수 간에 상호작용 효과가 존재한다.

▷ 전략 : 정상성, 등분상성, 교호작용효과 검정 -> 주효과 검정

**◦ 분산분석표**

테이블이(가) 표시된 사진

자동 생성된 설명