

03

## 일원 배치 분산 분석 개념

[일원 배치 분산분석]

# 데이터에서 종속변수가 몇개인지, 독립변수가 범주형인지를 보고 어떠한 검정을 하는게 적합할지 판단하는 것이 중요합니다!

File Edit View Insert Cell Kernel Navigate Widgets Help Trusted Python 3 (ipykernel)

3 분산분석 (ANOVA)

3.1 일원배치 분산분석 (One-way ANOVA)

- T-검정이 두 집단 간의 평균 차이를 비교하는 통계분석 방법이라면 분산분석은 두 개 이상의 다수 집단 간 평균을 비교하는 통계분석 방법이다.
- 분산분석은 개수에 따라 일원배치, 이원배치, 다원배치 분산분석으로 나뉘어진다.

3.1.1 분산분석의 개념

- 분산분석은 두 개 이상의 집단에서 그룹 평균 간 차이를 그룹 내 변동에 비교하여 살펴보는 통계 분석 방법이다.
- 즉, 두 개 이상 집단들의 평균 간 차이에 대한 통계적 유의성을 검증하는 방법이다.
- 분산분석의 분류는 아래와 같다.

분석 구분	분석 명칭	독립변수 개수	종속변수 개수
단일변량 분산분석	일원배치 분산분석	1개	1개
	이원배치 분산분석	2개	
	다원배치 분산분석	3개 이상	
다변량 분산분석	MANOVA	1개 이상	2개 이상

### 일원배치 분산분석의 개념 - 분산분석에서 반응값에 대한 하나의 범주형 변수의 영향을 알아보기 위해 사용되는 검증 방법이다. - 모집단의 수에는 제한이 없으며, 각 표본의 수는 같지 않아도 된다. - F 검정 통계량을 이용한다.

3.1.2 일원배치 분산분석의 가정

- 각 집단의 측정치는 서로 독립적이며, 정규분포를 따른다.
- 각 집단 측정치의 분산은 같다. (등분산 가정)

[일원 배치 분산분석 flow]

06:17

File Edit View Insert Cell Kernel Navigate Widgets Help Trusted Python 3 (ipykernel)

3.1.7 Python을 활용한 일원배치 분산분석

- 함수 사용법

일원배치 분산분석

```
import scipy.stats as stats
import scikit_posthocs as sp
import pingouin as pg
from statsmodel.stats.multicomp import pairwise_tukeyhsd
from statsmodel.stats.multicomp import MultiComparison
```

정규성 검정

Yes → stats.kruskal()

NO → stats.kruskal()

등분산 검정

Yes → Pg.welch\_anova( dv = "양적 자료", Between = "집단 질적 자료", Data = )

NO → Pg.welch\_anova( dv = "양적 자료", Between = "집단 질적 자료", Data = )