

F검정 (분산분석)

통계분석의 주요 검정 기법

F검정의 이해

분산

F검정의 가설

#01. 작업 준비

패키지 가져오기

데이터 가져오기

#02. 분산분석 시행

해석

# F검정 (분산분석)

## 통계분석의 주요 검정 기법

기법	설명
독립표본 t검정	독립된 두 그룹의 평균비교(정규분포 가정이 가능할 때)
대응표본 t검정	쌍을 이룬 두 그룹의 평균비교(정규분포 가정이 가능할 때)
분산분석	세 그룹의 평균 비교 (정규분포 가정이 가능할 때)
상관분석	두 수치형 변수 사이의 상관관계 유무와 강도 계산
회귀분석	종속변수와 독립변수의 관계를 모델링

## F검정의 이해

분산을 비교할 때 사용

세 개 이상의 집단을 대상으로 두 집단씩 여러번 **t검정** 을 수행하게 되면 신뢰도가 급격히 감소한다.

분산 분석은 이러한 현상이 발생하지 않는다.

## 분산

- 데이터의 변동 정도를 의미하는 지표

## F검정 (분산분석)

## 통계분석의 주요 검정 기법

## F검정의 이해

## 분산

## F검정의 가설

## #01. 작업 준비

## 패키지 가져오기

## 데이터 가져오기

## #02. 분산분석 시행

## 해석

- 편차의 제공에 대한 평균
- 분산이 크다 --> 데이터의 산포도가 크다 (데이터가 넓게 퍼져있음)

## F검정의 가설

가설	설명	식
귀무가설	데이터간의 분산이 같다	$H_0 : \sigma_1 = \sigma_2$
대립가설	데이터간의 분산이 같지 않다	$H_0 : \sigma_1 \neq \sigma_2$

## #01. 작업 준비

## 패키지 가져오기

```
from pandas import read_excel
from scipy.stats import f_oneway
```

## 데이터 가져오기

어떤 지역에서 동일한 제품을 판매하고 있는 두 가게에 대한 일별 매출액

```
df = read_excel("https://data.hossam.kr/E02/store.xlsx")
df
```

	store1	store2
0	46	78

## F검정 (분산분석)

통계분석의 주요 검정 기법

F검정의 이해

분산

F검정의 가설

#01. 작업 준비

패키지 가져오기

데이터 가져오기

#02. 분산분석 시행

해석

	store1	store2
1	47	57
2	58	31
3	47	28
4	27	67
5	58	77
6	56	36
7	26	57
8	47	36
9	25	57

## #02. 분산분석 시행

```
f_oneway(df['store1'], df['store2'])
```

```
F_onewayResult(statistic=1.4591624718860445, pvalue=0.24269553293319623)
```

## 해석

p-value가 0.05보다 크므로 귀무가설을 기각할 수 없다.

## F검정 (분산분석)

즉, 두 가게의 일별 매출 분산은 같다.

통계분석의 주요 검정 기법

F검정의 이해

분산

F검정의 가설

#01. 작업 준비

패키지 가져오기

데이터 가져오기

#02. 분산분석 시행

해석