

# t-검정 실습 예제

## 1. 단일표본 t-검정 : 한 집단의 평균과 특정한 값을 비교하는 경우 사용

- 중학교 2학년 학생들을 대상으로 한 과학성취도 국제비교연구에서 국제평균은 470점이 었다. 우리나라 학생들의 과학 성취수준이 국제 성취수준과 같은가?
  - 귀무가설(영가설) : 우리나라 학생들의 과학성취도의 평균은 470점이다
  - 대립가설 : 우리나라 학생들의 과학성취도의 평균은 470점과 다를것이다.

```
In [ ]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy import stats
```

```
%precision 3
```

```
In [ ]: # "(9)일표본t검정.csv" 파일 읽어오기

df = pd.read_csv('../data/(9)일표본t검정.csv')
df
```

```
In [ ]: # 과학 점수에 결측치 있는지 확인하기. 원본파일에는 결측치 없음.

df[df['과학성취도'].isna()]
```

```
In [ ]: # 만약 결측치가 있다면...
# 결측치가 있는 데이터 읽어와서 결측치 처리해보기

df2 = pd.read_csv('../data/(9)일표본t검정_na.csv', encoding='cp949')
df2
```

```
In [ ]: df2[df2['과학성취도'].isna()]
```

```
In [ ]: # 결측치(NA)를 0으로 대체하기 --> .fillna() 메서드 사용.
df_filled = df2.fillna(0)
df_filled.loc[7]
```

```
In [ ]: # 결측치를 포함하고 있는 데이터 행을 삭제하기 위해서는 .dropna() 메서드 사용.

df_dropped = df2.dropna()
df_dropped.head(10)
```

```
In [ ]: # 원본 데이터로 기출통계

df['과학성취도'].describe()
```

```
In [ ]: # 원본 데이터로 단일표본 t-검정 수행.

t, p = stats.ttest_1samp(df['과학성취도'], 470)

t, p
```

```
In [ ]: print(f"p값 = {p:.110f}")
```

## 결론)

- 1000명의 학생을 대상으로 과학성취도를 평가한 결과, 평균은 486.06점, 표준편차는 20.11이다. t통계값은 25.26이고, 이에 따른 유의확률은 .000으로 유의수준 .05 미만으로 귀무가설이 기각되고, 대립가설이 채택된다.
- 즉, 우리나라 학생들의 과학성취도의 평균은 470점이 아니라고 결론 내릴 수 있다.

## 우리나라 중학생 과학성취도 점수를 표준화 하여 분석하기(Z점수, T점수)

```
In [ ]: # 점수를 넘파이 배열로 만들기
```

```
scores = np.array(df['과학성취도'])  
scores[:10]
```

```
In [ ]: # 데이터에서 평균을 빼고 표준편차로 나누는 표준화 작업하여 Z점수 구하기  
# --> 표준화된 데이터는 평균이 0, 표준편차가 1임
```

```
z = (scores - np.mean(scores)) / np.std(scores)  
z[:10]
```

```
In [ ]: # 평균이 50, 표준편차가 10이 되도록 한 T점수 구하기
```

```
t = 50 + 10 * z  
t[:10]
```

```
In [ ]: max(scores), max(z), max(t)
```

```
In [ ]: # scipy 라이브러리에 zscore()함수를 이용하여 Z점수 구하기
```

```
stats.zscore(scores)[:10]
```

```
In [ ]: df['Z점수'] = z  
df['T점수'] = t
```

```
df
```

## 결론)

- T점수가 50점이면 평균적인 결과이고, 60점이면 상위 결과라는 통일된 평가를 할 수 있다.
- T점수가 60점이라면 평균에서 표준편차가 한 단위 만큼 높은 점수에 해당한다.

```
In [ ]: # T점수로 계급수 25개, 계급폭 4로 하여 히스토그램 그리기
```

```
fig = plt.figure(figsize=(10, 6))  
ax = fig.add_subplot(111)  
  
freq, *_ = ax.hist(t, bins=25, range=(0, 100))  
ax.set_xlabel('score')  
ax.set_ylabel('person number')
```

```
ax.set_xticks(np.linspace(0, 100, 25+1))
ax.set_yticks(np.arange(0, freq.max()+1,50))
plt.show()
```

```
In [ ]: df['과학성취도'].plot(kind='hist', bins=10, range=(0,540))
```

```
In [ ]: np.histogram(t, bins=25, range=(0,100))
```

```
In [ ]: np.histogram(scores, bins=10, range=(0,540))
```

## 2. 대응표본 t-검정 : 같은 모집단에서 추출된 표본들의 평균비교(전-후)

- 체력증진을 위해 개발된 새로운 프로그램을 고등학교 2학년 학생 150명을 대상으로 한 학기동안 적용하였다. 체력증진 프로그램을 적용하기 전후 학생들의 체력에 차이가 있는가?
  - 귀무가설(영가설): 전후 차이가 없다. 즉, 프로그램의 효과가 없다.
  - 대립가설: 전후 차이가 있다. 프로그램의 효과가 있다.

```
In [ ]: df3 = pd.read_csv('../data/(10)대응표본t검정.csv')
df3
```

```
In [ ]: # 기술통계

df3.describe()
```

```
In [ ]: # 대응표본 t-검정 수행
# t통계값은 사후점수에서 사전점수를 뺀 평균 차이점수를 사용하여 계산되는데,
# 사후점수가 더 높기 때문에 양수로 제시된다.

t3, p3 = stats.ttest_rel(df3['사후체력'], df3['사전체력'])

t3, p3
```

## 결론)

- 체력증진 프로그램의 효과를 알아보기 위하여 사전체력과 사후체력의 기술통계와 대응표본 t-검정 결과는 다음과 같다.
- 사전체력의 평균은 64.07, 표준편차는 21.61이며, 사후체력의 평균은 68.25, 표준편차는 16.01이다.
- 사전체력과 사후체력의 차이에 대한 통계적 유의성을 검정한 결과 t 통계값은 2.94, 유의확률(p값)은 0.004로서 유의수준 0.05보다 미만이다.
- 그러므로 체력증진 프로그램에 의한 학생들의 사전과 사후 체력에 차이가 있는 것으로 분석되었다. ---> 귀무가설 기각. 대립가설 채택.

## 3. 독립표본 t-검정 : 다른 모집단에서 추출된 두 표본의 평균비교

- 성별에 따라 고등학교 2학년 학생들의 외국어 능력에 차이가 있는가?
  - 귀무가설(영가설): 성별에 따라 외국어 능력에 차이가 없다.
  - 대립가설: 차이가 있다.

```
In [ ]: df4 = pd.read_csv('../data/(11)독립표본t검정.csv')
df4
```

```
In [ ]: # 성별이 1(남성)인 학생의 외국어 점수들 저장하기

# df4.loc[df4['성별']==1,['외국어']]
m_scores = df4.loc[df4['성별']==1,['외국어']]
m_scores
```

```
In [ ]: # 성별이 2(여성)인 학생의 외국어 점수들 저장하기

# df4.loc[df4['성별']==2,['외국어']]
f_scores = df4.loc[df4['성별']==2,['외국어']]
f_scores
```

```
In [ ]: m_scores.describe()
```

```
In [ ]: f_scores.describe()
```

```
In [ ]: # 두 표본의 분산이 동일한지 아닌지를 먼저 판단하기
# Levene의 등분산 검정. scipy.stats.levene() 함수를 사용
# p-value가 유의수준(0.05)보다 작으면, 두 표본의 분산이 서로 다르다고 분석함.
# p값이 유의수준이상이면, 두 표본의 분산이 같다고 분석함.

m = np.array(m_scores['외국어']) # 표본은 1차원이어야 하므로 넘파이배열로 변환.
f = np.array(f_scores['외국어'])

stats.levene(m, f, center='mean') # p값이 0.28로 유의수준 0.05 이상이므로 두 표본의
```

---> 레벤의 등분산 검정을 통해 p값이 0.29로 유의수준 0.05 이상이므로 두 표본의 분산은 같다고 가정하고 독립표본 t-검정을 수행함.

```
In [ ]: # 독립표본 t-검정 수행
# 기본적으로 equal_var=True로 설정되어 있음. 디폴트.
# 즉, 두 표본의 분산이 동일하다고 가정하고 t-검정 수행

t4, p4 = stats.ttest_ind(m, f)

t4, p4
```

```
In [ ]: # equal_var=False를 지정하면 웰치의 방법으로 독립표본 t-검정이 수행됨.
# 웰치의 t검정은 두 표본의 분산이 다르다는 것을 가정하고 검정을 수행한다.
t4, p4 = stats.ttest_ind(m, f, equal_var=False)

t4, p4
```

## 결론)

- Levene의 등분산 검정결과, 유의확률(p값)이 0.282로 유의수준(0.05) 이상이므로 두 집단의 분산이 같다는 귀무가설이 채택된다. 그러므로 독립표본 t-검정시 equal\_var 인수에 'True' 또는 생략하여 등분산을 가정하는 p값을 구한다.

- 고등학교 2학년 학생들의 성별에 따른 외국어 점수의 기술통계와 차이를 알아보기 위하여 두 독립표본 t검정을 실시한 결과는 다음과 같다.
- 남학생들의 외국어 능력의 평균은 39.71, 표준편차는 10.15이며,
- 여학생들의 외국어 능력의 평균은 42.44, 표준편차는 9.12이다.
- 남녀 학생들의 외국어 능력에 차이가 있는지에 대한 t통계값은 -2.14, 유의확률(p값)은 0.034로서 유의수준 0.05에서 성별에 따라 외국어 능력에 유의한 차이가 있는 것으로 분석되었다.

In [ ]: