

## 단일 표본 T-Test

## #01. 작업준비

패키지 참조

## #02. 예시 1

가설 설정

데이터 가져오기

p-value의 기준

결과 해석

## 03. 예시2

가설 설정

결과 해석

## 예시 3

가설 설정

결과 해석

참고: t-statistic (T값)

# 단일 표본 T-Test

하나의 모집단 평균이 **이전보다** 커졌는지/작아졌는지/달라졌는지를 통계적으로 알아보기 위해 사용  
기준에 알려져 있던 사실이 현재는 어떻게 변화 했는지를 확인하는 방법

## #01. 작업준비

### 패키지 참조

`scipy` 패키지의 설치가 필요하다.

```
from scipy import stats
from pandas import read_excel
```

## #02. 예시 1

주어진 데이터는 강아지 25마리의 체온을 측정한 데이터이다. 과거의 자료에서 강아지의 평균 체온은 24.3 알려져 있다. 이번에 조사한 데이터에서도 강아지의 평균 체온이 24.3이라고 해도 좋은지 알고 싶다.

### 가설 설정

강아지의 평균 체온 :  $\mu$ 귀무가설( $H_0$ ) :  $\mu = 24.3$

## 단일 표본 T-Test

## #01. 작업준비

패키지 참조

## #02. 예시 1

가설 설정

데이터 가져오기

p-value의 기준

결과 해석

## 03. 예시2

가설 설정

결과 해석

## 예시 3

가설 설정

결과 해석

참고: t-statistic (T값)

대립가설( $H_1$ ):  $\mu \neq 24.3$ 

## 데이터 가져오기

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

	체온
0	25.8
1	24.6
2	26.1
3	22.9
4	25.1
5	27.3
6	24.0
7	24.5
8	23.9
9	26.2
10	24.3
11	24.6
12	23.3

## 단일 표본 T-Test

## #01. 작업준비

패키지 참조

## #02. 예시 1

가설 설정

데이터 가져오기

p-value의 기준

결과 해석

## 03. 예시2

가설 설정

결과 해석

## 예시 3

가설 설정

결과 해석

참고: t-statistic (T값)

	체온
13	25.5
14	28.1
15	24.8
16	23.5
17	26.3
18	25.4
19	25.5
20	23.9
21	27.0
22	24.8
23	22.9
24	25.4

```
result = stats.ttest_1samp(df['체온'], 24.3)
result
```

```
TtestResult(statistic=2.7127692953052716, pvalue=0.012145537241648427, c
```

## p-value의 기준

## 단일 표본 T-Test

## #01. 작업준비

패키지 참조

## #02. 예시 1

가설 설정

데이터 가져오기

p-value의 기준

결과 해석

## 03. 예시2

가설 설정

결과 해석

## 예시 3

가설 설정

결과 해석

참고: t-statistic (T값)

- $P\text{-value} < 0.01$  : 귀무가설이 옳을 확률이  $0.01$  이하 → 틀렸다(간간한 기준) --> 대립가설 채택
- $P\text{-value} < 0.05$  : 귀무가설이 옳을 확률이  $0.05$  이하 → 틀렸다(일반적인 기준) --> 대립가설 채택
- $0.05 < P\text{-value} < 0.1$  : 애매한 경우(샘플링을 다시한다)
- $0.1 < P\text{-value}$  : 귀무가설이 옳을 확률이 0.1 이상 → 틀리지 않았다(맞다와 다름)

## 결과 해석

출력에서 일표본 t검정에 대한 유의확률값이  $0.0121$ 로 통상적인 유의수준 0.05보다 작으므로 유의수준  $0.05$ 에서 강아지 25마라의 체온이 추출된 모집단의 평균인  $24.3$ 이라고 할 수 없다.

## 03. 예시2

## 가설 설정

강아지의 평균 체온이 과거의 자료인 24.3보다 작아졌는지 알고자 하는 경우

귀무가설( $H_0$ ):  $\mu = 24.3$

대립가설( $H_1$ ):  $\mu < 24.3$

```
result = stats.ttest_1samp(df['체온'], 24.3, alternative='less')
result
```

## 단일 표본 T-Test

## #01. 작업준비

패키지 참조

## #02. 예시 1

가설 설정

데이터 가져오기

p-value의 기준

결과 해석

## 03. 예시2

가설 설정

결과 해석

## 예시 3

가설 설정

결과 해석

참고: t-statistic (T값)

```
TtestResult(statistic=2.7127692953052716, pvalue=0.9939272313791758, df=
```

## 결과 해석

p-value 가 0.05 보다 크므로 귀무가설을 기각할 수 없다. 즉, 강아지의 평균 체온은 24.3 과 같다(작지 않다).

## 예시 3

## 가설 설정

강아지의 평균 체온이 과거의 자료인 24.3보다 커졌는지 알고자 하는 경우

귀무가설( $H_0$ ):  $\mu = 24.3$

대립가설( $H_1$ ):  $\mu > 24.3$

```
t, p = stats.ttest_1samp(df['체온'], 24.3, alternative='greater')
"t-statistic: {:.3f}, p-value: {:.3f}".format(t, p)
```

```
't-statistic: 2.713, p-value: 0.006'
```

## 결과 해석

p-value 가 0.05 보다 작으므로 귀무가설을 기각하고 대립가설을 채택한다. 즉, 강아지의 평균 체온은 24.3 보다 크다

## 단일 표본 T-Test

## #01. 작업준비

패키지 참조

## #02. 예시 1

가설 설정

데이터 가져오기

p-value의 기준

결과 해석

## 03. 예시2

가설 설정

결과 해석

## 예시 3

가설 설정

결과 해석

## 참고: t-statistic (T값)

## 참고: t-statistic ( T값 )

계산된 차이를 표준 오차 단위로 나타낸 것으로, T의 크기가 클수록 귀무 가설에 대한 증거가 큼니다. 즉, 유의한 차이가 있다는 증거가 더 명확한 것입니다. 반면 T가 0에 가까울수록 유의미한 차이가 없을 가능성이 커집니다.