

연습문제 풀이

패키지 참조

문제 1

데이터 가져오기

신뢰구간 구하기

모듈화 한 기능 사용하기

신뢰구간 시각화

빵의 무게가 500g보다 작은지 검
정하기

가설 설정

빵의 평균 무게가 500보다 큰
지에 대한 검정

결론

연습문제 풀이

패키지 참조

```
from scipy import stats
from scipy.stats import t
from pandas import read_excel
from math import sqrt
from matplotlib import pyplot as plt
import seaborn as sb
import sys
import os
```

```
sys.path.append(os.path.dirname(os.path.dirname(os.getcwd())))
from helper import get_confidence_interval
```

문제 1

데이터 가져오기

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

연습문제 폴이

패키지 참조

문제 1

데이터 가져오기

신뢰구간 구하기

모듈화 한 기능 사용하기

신뢰구간 시각화

빵의 무게가 500g보다 작은지 검
정하기

가설 설정

빵의 평균 무게가 500보다 큰
지에 대한 검정

결론

| | 무게 |
|----|-----|
| 0 | 495 |
| 1 | 496 |
| 2 | 502 |
| 3 | 494 |
| 4 | 499 |
| 5 | 501 |
| 6 | 493 |
| 7 | 494 |
| 8 | 495 |
| 9 | 501 |
| 10 | 493 |
| 11 | 498 |
| 12 | 500 |
| 13 | 492 |
| 14 | 493 |
| 15 | 494 |
| 16 | 500 |
| 17 | 492 |

연습문제 폴이

패키지 참조

문제 1

데이터 가져오기

신뢰구간 구하기

모듈화 한 기능 사용하기

신뢰구간 시각화

빵의 무게가 500g보다 작은지 검
정하기

가설 설정

빵의 평균 무게가 500보다 큰
지에 대한 검정

결론

| | 무게 |
|----|-----|
| 18 | 497 |
| 19 | 499 |
| 20 | 491 |
| 21 | 492 |
| 22 | 493 |
| 23 | 499 |
| 24 | 491 |
| 25 | 496 |
| 26 | 498 |
| 27 | 490 |
| 28 | 491 |
| 29 | 492 |

신뢰구간 구하기

```
# 신뢰수준
clevel = 0.95

# 샘플 사이즈
n = len(df['무게'])
```

연습문제 풀이

패키지 참조

문제 1

데이터 가져오기

신뢰구간 구하기

모듈화 한 기능 사용하기

신뢰구간 시각화

빵의 무게가 500g보다 작은지 검
정하기

가설 설정

빵의 평균 무게가 500보다 큰
지에 대한 검정

결론

```
# 자유도
dof = n - 1

# 평균
sample_mean = df['무게'].mean()

# 표본 표준 편차
sample_std = df['무게'].std(ddof=1)

# 표본 표준오차
sample_std_error = sample_std / sqrt(n)

# 신뢰구간
cmin, cmax = t.interval(clevel, dof, loc=sample_mean, scale=sample_std_e
cmin, cmax

(494.03836792675105, 496.6949654065823)
```

모듈화 한 기능 사용하기

```
cmin, cmax = get_confidence_interval(df['무게'])
cmin, cmax
```

```
(494.03836792675105, 496.6949654065823)
```

신뢰구간 시각화

연습문제 풀이

패키지 참조

문제 1

데이터 가져오기

신뢰구간 구하기

모듈화 한 기능 사용하기

신뢰구간 시각화

빵의 무게가 500g보다 작은지 검
정하기

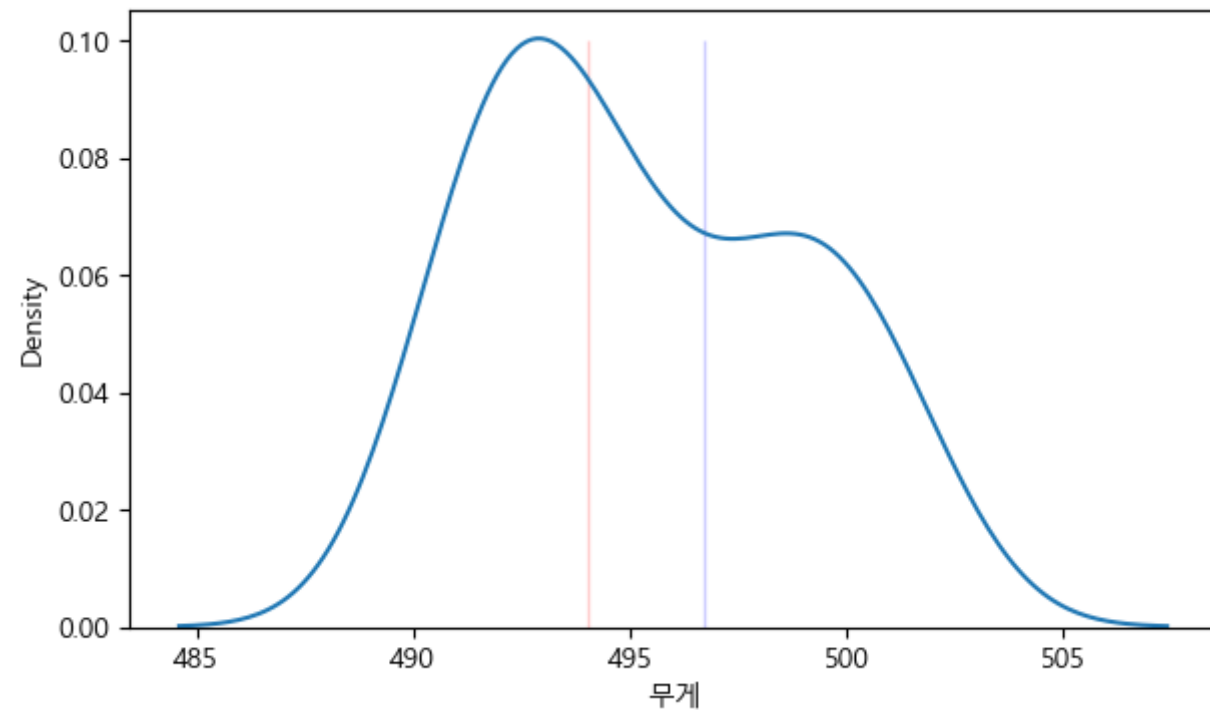
가설 설정

빵의 평균 무게가 500보다 큰
지에 대한 검정

결론

```
plt.rcParams["font.family"] = 'AppleGothic' if sys.platform == 'darwin'
plt.rcParams["font.size"] = 10
plt.rcParams["figure.figsize"] = (7, 4)
plt.rcParams["axes.unicode_minus"] = False

sb.kdeplot(x=df['무게'])
sb.lineplot(x=[cmin, cmin], y=[0, 0.1], color='red')
sb.lineplot(x=[cmax, cmax], y=[0, 0.1], color='blue')
plt.show()
plt.close()
```



연습문제 폴이

패키지 참조

문제 1

데이터 가져오기

신뢰구간 구하기

모듈화 한 기능 사용하기

신뢰구간 시각화

빵의 무게가 500g보다 작은지 검
정하기

가설 설정

빵의 평균 무게가 500보다 큰
지에 대한 검정

결론

빵의 무게가 500g보다 작은지 검정하기

가설 설정

 $\mu =$ 빵의 무게

| 가설 | 내용 |
|---------------|----------------|
| 귀무가설(H_0) | $\mu \geq 500$ |
| 대립가설(H_1) | $\mu < 500$ |

빵의 평균 무게가 500보다 크지에 대한 검정

```
# 귀무가설 : 빵의 무게가 500보다 크지 않다.
# -> p-value가 1인 경우는 귀무가설을 기각할 근거가 없다는 의미
t, p = stats.ttest_1samp(df['무게'], 500, alternative='greater')
"t-statistic: {:.f}, p-value: {:.f}".format(t, p)
```

```
't-statistic: -7.134111, p-value: 1.000000'
```

빵의 무게는 500보다 크지는 않다.

```
# 귀무가설 : 빵의 무게가 500과 차이가 없다.
# -> p-value가 0이라는 것은 주어진 데이터에서 검정하는 가설이 완전히 거부되어야 한
```

연습문제 풀이

패키지 참조

문제 1

데이터 가져오기

신뢰구간 구하기

모듈화 한 기능 사용하기

신뢰구간 시각화

빵의 무게가 500g보다 작은지 검
정하기

가설 설정

빵의 평균 무게가 500보다 큰
지에 대한 검정

결론

연습문제_풀이.ipynb

```
t, p = stats.ttest_1samp(df['무게'], 500, alternative='two-sided')
"t-statistic: {:.f}, p-value: {:.f}".format(t, p)
```

```
't-statistic: -7.134111, p-value: 0.000000'
```

무조건 500은 아님

```
# 귀무가설 : 빵의 무게가 500보다 작지 않다.
t, p = stats.ttest_1samp(df['무게'], 500, alternative='less')
"t-statistic: {:.f}, p-value: {:.f}".format(t, p)
```

```
't-statistic: -7.134111, p-value: 0.000000'
```

무조건 500보다 작다

결론

클래임 걸어야 함