

April 18, 2023

```
[7]: import scipy.stats as st
import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.stats.weightstats import ztest
from statsmodels.stats.proportion import proportions_ztest

dataAnggur = pd.read_csv('../data/anggur.csv')
```

Melakukan test hipotesis 1 sampel,

0.0.1 a. Nilai rata-rata pH di atas 3.29?

H_0 = Nilai rata-rata pH sama dengan 3.29 ($\mu = 3.29$)

H_1 = Nilai rata-rata pH lebih dari 3.29 ($\mu > 3.29$)

Tingkat Signifikan $\alpha = 0.05$

Lakukan uji statistik dengan one tailed test ke arah kanan (right tailed test) karena ($\mu > 3.29$).

Ambil daerah kritis ($z > z_\alpha$)

Hitung nilai z dengan rumus

$$z = \frac{(x - \mu_0)}{(\sigma/\sqrt{n})}$$

Pengambilan Keputusan

Tes Daerah Kritis

- Reject H_0 jika ($z > z_\alpha$)
- Fail to reject H_0 jika ($z \leq z_\alpha$)

Tes Signifikansi

- Reject H_0 jika $p < \alpha$
- Fail to reject H_0 jika $p \geq \alpha$

```
[8]: # Diketahui
rerata = 3.29
alpha = 0.05
```

```
# Menggunakan ztest module, menghitung z dan p
z, p = ztest(dataAnggur['pH'], value = rerata)

# Menghitung z_alpha
z_a = st.norm.ppf(1-alpha)

# Hasil
print(f"Nilai z: {round(z, 4)}")
print(f"Nilai z_alpha: {round(z_a, 4)}")
print(f"Nilai p: {round(p, 6)}")
```

Nilai z: 4.1038

Nilai z_alpha: 1.6449

Nilai p: 4.1e-05

Hasil Tes

Tes Daerah Kritis

Karena z lebih besar dibandingkan dengan z_α ($4.103 > 1.644$), reject H_0 .

Tes Signifikansi

Karena p lebih kecil dibandingkan α ($0.000041 < 0.05$), reject H_0 .

Kesimpulan

Dengan tingkat signifikansi sebesar 0.05, ada bukti yang cukup untuk menolak klaim bahwa nilai rata-rata pH adalah 3.29. Maka nilai rata-rata pH lebih dari 3.29

0.0.2 b. Nilai rata-rata Residual Sugar tidak sama dengan 2.50?

H_0 = Nilai rata-rata Residual Sugar sama dengan 2.50 ($\mu = 2.50$)

H_1 = Nilai rata-rata Residual Sugar lebih dari 2.50 ($\mu \neq 2.50$)

Tingkat Signifikan $\alpha = 0.05$

Lakukan uji statistik dengan two tailed test pada bagian kanan $\mu > 2.50$ dengan ($z > z_{\alpha/2}$) dan bagian kiri $\mu < 2.50$ dengan ($z < -z_{\alpha/2}$)

Hitung nilai z dengan rumus

$$z = \frac{(x - \mu_0)}{(\sigma/\sqrt{n})}$$

Pengambilan Keputusan

Tes Daerah Kritis

- Reject H_0 jika $z < -z_{\alpha/2}$ atau $z > z_{\alpha/2}$
- Fail to reject H_0 jika $-z_{\alpha/2} \leq z \leq z_{\alpha/2}$

Tes Signifikansi

- Reject H_0 jika $p < \alpha$
- Fail to reject H_0 jika $p \geq \alpha$

```
[9]: # Diketahui
rerata = 2.50
alpha = 0.05

# Menggunakan ztest module, menghitung z dan p
z, p = ztest(dataAnggur['residual sugar'], value = rerata)

# Menghitung z_alpha
z_a = st.norm.ppf(1-(alpha/2))

# Hasil
print(f"Nilai z: {round(z, 4)}")
print(f"Nilai z_alpha/2: {round(z_a, 4)}")
print(f"Nilai p: {round(p, 6)}")
```

Nilai z: 2.148

Nilai z_alpha/2: 1.96

Nilai p: 0.031717

Hasil Tes

Tes Daerah Kritis

Karena z lebih besar dibandingkan dengan z_α ($2.148 > 1.96$), reject H_0 .

Tes Signifikansi

Karena p lebih kecil dibandingkan α ($0.031 < 0.05$), reject H_0 .

Kesimpulan

Dengan tingkat signifikansi sebesar 0.05, ada bukti yang cukup untuk menolak klaim bahwa nilai rata-rata residual sugar sama dengan 2.5. Maka nilai rata-rata Residual Sugar lebih dari 2.5.

0.0.3 c. Nilai rata-rata 150 baris pertama kolom sulphates bukan 0.65?

H_0 = Nilai rata-rata 150 baris pertama kolom sulphates sama dengan 0.65 ($\mu = 0.65$)

H_1 = Nilai rata-rata 150 baris pertama kolom sulphates tidak sama dengan 0.65 ($\mu \neq 0.65$)

Tingkat Signifikan $\alpha = 0.05$

Lakukan uji statistik dengan two tailed test pada bagian kanan $\mu > 0.65$ dengan ($z > z_{\alpha/2}$) dan bagian kiri $\mu < 0.65$ dengan ($z < -z_{\alpha/2}$)

Hitung nilai z dengan rumus

$$z = \frac{(x - \mu_0)}{(\sigma/\sqrt{n})}$$

Pengambilan Keputusan

Tes Daerah Kritis

- Reject H_0 jika $z < -z_{\alpha/2}$ atau $z > z_{\alpha/2}$
- Fail to reject H_0 jika $-z_{\alpha/2} \leq z \leq z_{\alpha/2}$

Tes Signifikansi

- Reject H_0 jika $p < \alpha$
- Fail to reject H_0 jika $p \geq \alpha$

```
[10]: # Diketahui
rerata = 0.65
alpha = 0.05

# Menggunakan ztest module, menghitung z dan p
z, p = ztest(dataAnggur['sulphates'].head(150), value = rerata)

# Menghitung z_alpha
z_a = st.norm.ppf(1-(alpha/2))

# Hasil
print(f"Nilai z: {round(z, 4)}")
print(f"Nilai z_alpha/2: {round(z_a, 4)}")
print(f"Nilai p: {round(p, 6)}")
```

Nilai z: -4.9648

Nilai z_alpha/2: 1.96

Nilai p: 1e-06

Hasil Tes

Tes Daerah Kritis

Karena z lebih kecil dibandingkan dengan $-z_{\alpha/2}$ ($-4.9648 < -1.96$), reject H_0 .

Tes Signifikansi

Karena p lebih kecil dibandingkan α ($0.000001 < 0.05$), reject H_0 .

Kesimpulan

Dengan tingkat signifikansi sebesar 0.05, ada bukti yang cukup untuk menolak klaim bahwa nilai rata-rata 150 baris pertama kolom sulphates sama dengan 0.65.

0.0.4 d. Nilai rata-rata total sulfur dioxide di bawah 35?

H_0 = Nilai rata-rata total sulfur dioxide sama dengan 35 ($\mu = 35$)

H_1 = Nilai rata-rata total sulfur dioxide kurang dari 35 ($\mu < 35$)

Tingkat Signifikan $\alpha = 0.05$

Lakukan uji statistik dengan one tailed test ke arah kiri (left tailed test) karena ($\mu < 35$). Ambil daerah kritis ($z < -z_{\alpha}$)

Hitung nilai z dengan rumus

$$z = \frac{(x - \mu_0)}{(\sigma/\sqrt{n})}$$

Pengambilan Keputusan

Tes Daerah Kritis

- Reject H_0 jika $(z < -z_\alpha)$
- Fail to reject H_0 jika $(z \geq z_\alpha)$

Tes Signifikansi

- Reject H_0 jika $p < \alpha$
- Fail to reject H_0 jika $p \geq \alpha$

```
[11]: # Diketahui
rerata = 35
alpha = 0.05

# Menggunakan ztest module, menghitung z dan p
z, p = ztest(dataAnggur['total sulfur dioxide'], value = rerata)

# Menghitung z_alpha
z_a = st.norm.ppf(1-alpha)

# Hasil
print(f"Nilai z: {round(z, 4)}")
print(f"Nilai -z_alpha: -{round(z_a, 4)}")
# Karena merupakan one tailed test ke arah kiri, maka p-valuenya 1-p
print(f"Nilai p: {1-p}")
```

Nilai z : 16.7864

Nilai $-z_\alpha$: -1.6449

Nilai p : 1.0

Hasil Tes

Tes Daerah Kritis

Karena z lebih besar dibandingkan dengan $-z_\alpha$ ($16.78 > -1.6449$), fail to reject H_0 .

Tes Signifikansi

Karena p lebih kecil dibandingkan α ($1 > 0.05$), fail to reject H_0 .

Kesimpulan

Dengan tingkat signifikansi sebesar 0.05, tidak ada bukti yang cukup untuk menolak klaim bahwa nilai rata-rata total sulfur dioxide sama dengan 35.

0.0.5 e. Proporsi nilai total Sulfat Dioxide yang lebih dari 40, adalah tidak sama dengan 50% ?

H_0 = Proporsi nilai Conductivity yang lebih dari 40 sama dengan 50 ($p = 50\%$)

H_1 = Proporsi nilai Conductivity yang lebih dari 40 tidak sama dengan 50 ($p \neq 50\%$)

Tingkat Signifikan $\alpha = 0.05$

Lakukan uji statistik dengan two tailed test pada bagian kanan dengan ($z > z_{\alpha/2}$) dan bagian kiri dengan ($z < -z_{\alpha/2}$)

Hitung nilai z dengan rumus

$$z = \frac{(p - p_0)}{(p_0 q_0 / \sqrt{n})}$$

Pengambilan Keputusan

Tes Daerah Kritis

- Reject H_0 jika $z < -z_{\alpha/2}$ atau $z > z_{\alpha/2}$
- Fail to reject H_0 jika $-z_{\alpha/2} \leq z \leq z_{\alpha/2}$

Tes Signifikansi

- Reject H_0 jika $p < \alpha$
- Fail to reject H_0 jika $p \geq \alpha$

```
[12]: p_0 = 0.50
      alpha = 0.05

      # Test
      TotalSD_over_40 = len(dataAnggur[dataAnggur["total sulfur dioxide"] > 40])

      z, p = proportions_ztest(TotalSD_over_40,
                              len(dataAnggur), value=p_0, prop_var=p_0)
      z_a = st.norm.ppf(1-alpha/2)

      # Results
      print(f"Nilai z: {round(z, 4)}")
      print(f"Nilai z_alpha/2: {round(z_a, 4)}")
      print(f"Nilai p: {round(p, 4)}")
```

Nilai z: 0.7589

Nilai z_alpha/2: 1.96

Nilai p: 0.4479

Hasil Tes

Tes Daerah Kritis

Karena z memenuhi $-z_{\alpha} < z < z_{\alpha}$ ($-1.96 < 0.758 < 1.96$), fail to reject H_0 .

Tes Signifikansi

Karena p lebih besar dibandingkan α ($0.447 > 0.05$), fail to reject H_0 .

Kesimpulan

Dengan tingkat signifikansi sebesar 0.05, tidak ada bukti yang cukup untuk menolak klaim bahwa proporsi nilai sulfat dioxide yang lebih dari 40 adalah sama dengan 50%.