

Pengujian Hipotesis Terhadap Satu Sampel

```
In [ ]: # Import Libraries
import pandas as pd
import scipy.stats as s
from statsmodels.stats.weightstats import ztest
from statsmodels.stats.proportion import proportions_ztest

# Read csv file
df = pd.read_csv("../data/anggur.csv")

display(df)
```

	fixed acidity	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH	sulphates	alcohol	quality
0	5.90	0.4451	0.1813	2.049401	0.070574	16.593818	42.27	0.9982	3.27	0.71	8.64	7
1	8.40	0.5768	0.2099	3.109590	0.101681	22.555519	16.01	0.9960	3.35	0.57	10.03	8
2	7.54	0.5918	0.3248	3.673744	0.072416	9.316866	35.52	0.9990	3.31	0.64	9.23	8
3	5.39	0.4201	0.3131	3.371815	0.072755	18.212300	41.97	0.9945	3.34	0.55	14.07	9
4	6.51	0.5675	0.1940	4.404723	0.066379	9.360591	46.27	0.9925	3.27	0.45	11.49	8
...
995	7.96	0.6046	0.2662	1.592048	0.057555	14.892445	44.61	0.9975	3.35	0.54	10.41	8
996	8.48	0.4080	0.2227	0.681955	0.051627	23.548965	25.83	0.9972	3.41	0.46	9.91	8
997	6.11	0.4841	0.3720	2.377267	0.042806	21.624585	48.75	0.9928	3.23	0.55	9.94	7
998	7.76	0.3590	0.3208	4.294486	0.098276	12.746186	44.53	0.9952	3.30	0.66	9.76	8
999	5.87	0.5214	0.1883	2.179490	0.052923	16.203864	24.37	0.9983	3.29	0.70	10.17	7

1000 rows × 12 columns

Langkah-Langkah Pembuktian Hipotesis:

1. Tentukan hipotesis nol H_0 .
2. Tentukan hipotesis alternatif H_1 .
3. Tentukan tingkat signifikan α .
4. Tentukan uji statistik yang sesuai dan tentukan daerah kritis.
5. Hitung nilai uji statistik dari data sample. Hitung *p-value* sesuai dengan uji statistik yang digunakan.
6. Ambil keputusan "Tolak H_0 " jika nilai uji statistik terletak di daerah kritis, atau dengan tes signifikan, "Tolak H_0 " jika *p-value* lebih kecil dibanding tingkat signifikansi α yang diinginkan.

Q1: Nilai rata-rata pH di atas 3.29?

Langkah-langkah:

1. $H_0 : \mu = 3.29$
2. $H_1 : \mu > 3.29$
3. Significance Level : $\alpha = 0.05$
4. Uji Statistik: One-Tailed Test

Daerah Kritis: $1 - \alpha = 0.95$ dan $P(z < 1.645) = 0.95$ sehingga daerah kritisnya adalah $z > 1.645$.

Perhitungannya juga ada di kode di bawah ini.

5. Test Statistik:

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

Perhitungan z dan p-value ada pada kode di bawah ini.

6. Tolak H_0 jika nilai uji terletak di daerah kritis atau dengan dengan tes signifikan ($z > z_\alpha$) atau tolak H_0 jika p-value lebih kecil dibandingkan tingkat signifikansi α yang diinginkan. Jika di luar kondisi tersebut, terima H_0 .

Pengambilan keputusan tersebut ada pada kode di bawah ini.

```
In [ ]: df_pH = df["pH"]

# Significance Level
alpha = 0.05

# z value and p value
z_val_pH, p_val_pH = ztest(df_pH, value = 3.29, alternative = 'larger')
print("z =", z_val_pH)

# z-alpha value
z_alpha_val_pH = s.norm.ppf(1 - alpha)
print("z-alpha =", z_alpha_val_pH)
```

```
# Pengambilan Keputusan
if (z_val_pH > z_alpha_val_pH):
    print("Nilai z lebih besar dari z-alpha sehingga nilai uji terletak di daerah kritis.")
    print("Keputusan dari uji ini adalah tolak H0.\n")
else:
    print("Nilai z tidak lebih besar dari z-alpha sehingga nilai uji tidak terletak di daerah kritis.")
    print("Keputusan dari uji ini adalah tidak tolak H0.\n")

# p value
p_val_pH = s.norm.sf(z_val_pH)
print("p =", p_val_pH)

# Pengambilan Keputusan
if (p_val_pH < alpha):
    print("Nilai p lebih kecil dari tingkat signifikansi yang diinginkan")
    print("Keputusan dari uji ini adalah tolak H0")
else:
    print("Nilai p tidak lebih kecil dari tingkat signifikansi yang diinginkan")
    print("Keputusan dari uji ini adalah tidak tolak H0")
```

z = 4.1037807933651145

z-alpha = 1.6448536269514722

Nilai z lebih besar dari z-alpha sehingga nilai uji terletak di daerah kritis.

Keputusan dari uji ini adalah tolak H0.

p = 2.0322630043302333e-05

Nilai p lebih kecil dari tingkat signifikansi yang diinginkan

Keputusan dari uji ini adalah tolak H0

Kesimpulan:

Nilai rata-rata pH di atas 3.29.

Q2: Nilai rata-rata Residual Sugar tidak sama dengan 2.50?

Langkah-langkah:

1. $H_0 : \mu = 2.50$

2. $H_1 : \mu \neq 2.50$

3. Significance Level : $\alpha = 0.05$

4. Uji Statistik: Two-Tailed Test

Daerah Kritis: $z > z_{\alpha/2}$ atau $z < -z_{\alpha/2}$

Perhitungannya juga ada di kode di bawah ini.

5. Test Statistik:

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

Perhitungan z-value dan p-value ada pada kode di bawah ini.

6. Tolak H_0 jika $(z > z_{\alpha/2}$ atau $z < -z_{\alpha/2})$ atau tolak H_0 jika p-value lebih kecil dibandingkan tingkat signifikansi α yang diinginkan. Jika di luar kondisi tersebut, terima H_0 .

Pengambilan keputusan tersebut ada pada kode di bawah ini.

```
In [ ]: df_residual_sugar = df["residual sugar"]

# Significance Level
alpha = 0.05

# z value and p value
z_val_residual_sugar, p_val_residual_sugar = ztest(df_residual_sugar, value = 2.50, alternative = 'two-sided')

print("z =", z_val_residual_sugar)

# z-alpha value
z_alpha_val_residual_sugar = s.norm.ppf(1 - (alpha/2))
print("z-alpha =", z_alpha_val_residual_sugar)

# Pengambilan Keputusan
if (z_val_residual_sugar > z_alpha_val_residual_sugar):
    print("Nilai z lebih besar dari z-alpha/2 sehingga nilai uji terletak di daerah kritis.")
    print("Keputusan dari uji ini adalah tolak H0.\n")
elif (z_val_residual_sugar < z_alpha_val_residual_sugar*(-1)):
    print("Nilai z lebih kecil dari minus z-alpha/2 sehingga nilai uji terletak di daerah kritis.")
    print("Keputusan dari uji ini adalah tolak H0.\n")
else:
    print("Nilai z berada diantara dari minus z-alpha/2 dan z-alpha/2 sehingga nilai uji tidak terletak di daerah kritis.")
    print("Keputusan dari uji ini adalah tidak tolak H0.\n")

print("p =", p_val_residual_sugar)

# Pengambilan Keputusan
if (p_val_residual_sugar < alpha):
    print("Nilai p lebih kecil dari tingkat signifikansi yang diinginkan")
    print("Keputusan dari uji ini adalah tolak H0.")
else:
    print("Nilai p tidak lebih kecil dari tingkat signifikansi yang diinginkan")
    print("Keputusan dari uji ini adalah tidak tolak H0.")
```

$z = 2.1479619435539523$
 $z\text{-alpha} = 1.959963984540054$
Nilai z lebih besar dari $z\text{-alpha}/2$ sehingga nilai uji terletak di daerah kritis.
Keputusan dari uji ini adalah tolak H_0 .

$p = 0.031716778818727434$
Nilai p lebih kecil dari tingkat signifikansi yang diinginkan
Keputusan dari uji ini adalah tolak H_0 .

Kesimpulan:

Nilai rata-rata Residual Sugar tidak sama dengan 2.50.

Q3: Nilai rata-rata 150 baris pertama kolom sulphates bukan 0.65?

Langkah-langkah:

1. $H_0 : \mu = 0.65$
2. $H_1 : \mu \neq 0.65$
3. Significance Level : $\alpha = 0.05$
4. Uji Statistik: Two-Tailed Test

Daerah Kritis: $z > z_{\alpha/2}$ atau $z < -z_{\alpha/2}$

Perhitungannya juga ada di kode di bawah ini.

5. Test Statistik:

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

Perhitungan z -value dan p -value ada pada kode di bawah ini.

6. Tolak H_0 jika ($z > z_{\alpha/2}$ atau $z < -z_{\alpha/2}$) atau tolak H_0 jika p -value lebih kecil dibandingkan tingkat signifikansi α yang diinginkan. Jika di luar kondisi tersebut, terima H_0 .

Pengambilan keputusan tersebut ada pada kode di bawah ini.

```
In [ ]: df_sulphates = df["sulphates"].head(150)

# Significance Level
alpha = 0.05

# z value and p value
z_val_sulphates, p_val_sulphates = ztest(df_sulphates, value = 0.65, alternative = 'two-sided')

print("z =", z_val_sulphates)

# z-alpha value
z_alpha_sulphates = s.norm.ppf(1 - (alpha/2))
print("z-alpha =", z_alpha_sulphates)

# Pengambilan Keputusan
if (z_val_sulphates > z_alpha_sulphates):
    print("Nilai z lebih besar dari z-alpha/2 sehingga nilai uji terletak di daerah kritis.")
    print("Keputusan dari uji ini adalah tolak H0.\n")
elif (z_val_sulphates < z_alpha_sulphates*(-1)):
    print("Nilai z lebih kecil dari minus z-alpha/2 sehingga nilai uji terletak di daerah kritis.")
    print("Keputusan dari uji ini adalah tolak H0.\n")
else:
    print("Nilai z berada diantara dari minus z-alpha/2 dan z-alpha/2 sehingga nilai uji tidak terletak di daerah kritis.")
    print("Keputusan dari uji ini adalah tidak tolak H0.\n")

print("p =", p_val_sulphates)

# Pengambilan Keputusan
if (p_val_sulphates < alpha):
    print("Nilai p lebih kecil dari tingkat signifikansi yang diinginkan")
    print("Keputusan dari uji ini adalah tolak H0.")
else:
    print("Nilai p tidak lebih kecil dari tingkat signifikansi yang diinginkan")
    print("Keputusan dari uji ini adalah tidak tolak H0.")
```

$z = -4.964843393315918$
 $z\text{-alpha} = 1.959963984540054$
Nilai z lebih kecil dari minus $z\text{-alpha}/2$ sehingga nilai uji terletak di daerah kritis.
Keputusan dari uji ini adalah tolak H_0 .

$p = 6.875652918327359e-07$
Nilai p lebih kecil dari tingkat signifikansi yang diinginkan
Keputusan dari uji ini adalah tolak H_0 .

Kesimpulan:

Nilai rata-rata 150 baris pertama kolom sulphates bukan 0.65.

Q4: Nilai rata-rata total sulfur dioxide di bawah 35?

Langkah-langkah:

1. $H_0 : \mu = 35$

$$2. H_1 : \mu < 35$$

3. Significance Level : $\alpha = 0.05$

4. Uji Statistik: One-Tailed Test

Daerah Kritis: $1 - \alpha = 0.95$ dan $P(z > -1.645) = 0.95$ sehingga daerah kritisnya adalah $z < -1.645$.

Perhitungannya juga ada di kode di bawah ini.

5. Test Statistik:

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

Perhitungan z-value dan p-value ada pada kode di bawah ini.

6. Tolak H_0 jika nilai uji terletak di daerah kritis atau dengan tes signifikan ($z < -z_\alpha$) atau tolak H_0 jika p-value lebih kecil dibandingkan tingkat signifikansi α yang diinginkan. Jika di luar kondisi tersebut, terima H_0 .

Pengambilan keputusan tersebut ada pada kode di bawah ini.

```
In [ ]: df_total_sulfur_dioxide_1 = df["total sulfur dioxide"]

# Significance Level
alpha = 0.05

# z value and p value
z_val_total_sulfur_dioxide_1, p_val_total_sulfur_dioxide_1 = ztest(df_total_sulfur_dioxide_1, value = 35, alternative = 'smaller')
print("z =", z_val_total_sulfur_dioxide_1)

# z-alpha value
z_alpha_total_sulfur_dioxide_1 = -s.norm.ppf(1 - alpha)
print("z-alpha =", z_alpha_total_sulfur_dioxide_1)

# Pengambilan Keputusan
if (z_val_total_sulfur_dioxide_1 < z_alpha_total_sulfur_dioxide_1):
    print("Nilai z lebih kecil dari minus z-alpha sehingga nilai uji terletak di daerah kritis.")
    print("Keputusan dari uji ini adalah tolak H0.\n")
else:
    print("Nilai z tidak lebih kecil dari minus z-alpha sehingga nilai uji tidak terletak di daerah kritis.")
    print("Keputusan dari uji ini adalah tidak tolak H0.\n")

print("p =", p_val_total_sulfur_dioxide_1)

# Pengambilan Keputusan
if (p_val_total_sulfur_dioxide_1 < alpha):
    print("Nilai p lebih kecil dari tingkat signifikansi yang diinginkan")
    print("Keputusan dari uji ini adalah tolak H0")
else:
    print("Nilai p tidak lebih kecil dari tingkat signifikansi yang diinginkan")
    print("Keputusan dari uji ini adalah tidak tolak H0")
```

z = 16.786387372296744

z-alpha = -1.6448536269514722

Nilai z tidak lebih kecil dari minus z-alpha sehingga nilai uji tidak terletak di daerah kritis.

Keputusan dari uji ini adalah tidak tolak H_0 .

p = 1.0

Nilai p tidak lebih kecil dari tingkat signifikansi yang diinginkan

Keputusan dari uji ini adalah tidak tolak H_0

Kesimpulan:

Nilai rata-rata total sulfur dioxide tidak di bawah 35.

Q5: Proporsi nilai total Sulfat Dioxide yang lebih dari 40, adalah tidak sama dengan 50% ?

Langkah-langkah:

1. H_0 : Proporsi nilai total Sulfat Dioxide yang lebih dari 40 sama dengan 50% ($p = 0.5$)

2. H_1 : Proporsi nilai total Sulfat Dioxide yang lebih dari 40 tidak sama dengan 50% ($p \neq 0.5$)

3. Significance Level : $\alpha = 0.05$

4. Uji Statistik: Two-Tailed Test

Daerah Kritis: $1 - \alpha = 0.95$ dan $P(z < 1.645) = 0.95$ sehingga daerah kritisnya adalah $z > 1.645$.

Perhitungannya juga ada di kode di bawah ini.

5. Test Statistik:

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

Perhitungan z-value dan p-value ada pada kode di bawah ini.

6. Tolak H_0 jika ($z > z_{\alpha/2}$ atau $z < -z_{\alpha/2}$) atau tolak H_0 jika p-value lebih kecil dibandingkan tingkat signifikansi α yang diinginkan. Jika di luar kondisi tersebut, terima H_0 .

Pengambilan keputusan tersebut ada pada kode di bawah ini.

```
In [ ]: df_total_sulfur_dioxide_2 = df[df["total sulfur dioxide"] > 40]

# Significance Level
alpha = 0.05

# z value and p value
z_val_total_sulfur_dioxide_2, p_val_total_sulfur_dioxide_2 = proportions_ztest(len(df_total_sulfur_dioxide_2), len(df), value = 0.5, prop_var = 0.5)

print("z =", z_val_total_sulfur_dioxide_2)

# z-alpha value
z_alpha_total_sulfur_dioxide_2 = s.norm.ppf(1 - (alpha/2))
print("z-alpha =", z_alpha_total_sulfur_dioxide_2)

# Pengambilan Keputusan
if (z_val_total_sulfur_dioxide_2 > z_alpha_total_sulfur_dioxide_2):
    print("Nilai z lebih besar dari z-alpha/2 sehingga nilai uji terletak di daerah kritis.")
    print("Keputusan dari uji ini adalah tolak H0.\n")
elif (z_val_total_sulfur_dioxide_2 < z_alpha_total_sulfur_dioxide_2*(-1)):
    print("Nilai z lebih kecil dari minus z-alpha/2 sehingga nilai uji terletak di daerah kritis.")
    print("Keputusan dari uji ini adalah tolak H0.\n")
else:
    print("Nilai z berada diantara dari minus z-alpha/2 dan z-alpha/2 sehingga nilai uji tidak terletak di daerah kritis.")
    print("Keputusan dari uji ini adalah tidak tolak H0.\n")

print("p =", p_val_total_sulfur_dioxide_2)

# Pengambilan Keputusan
if (p_val_total_sulfur_dioxide_2 < alpha):
    print("Nilai p lebih kecil dari tingkat signifikansi yang diinginkan")
    print("Keputusan dari uji ini adalah tolak H0.")
else:
    print("Nilai p tidak lebih kecil dari tingkat signifikansi yang diinginkan")
    print("Keputusan dari uji ini adalah tidak tolak H0.")
```

z = 0.7589466384404118

z-alpha = 1.959963984540054

Nilai z berada diantara dari minus z-alpha/2 dan z-alpha/2 sehingga nilai uji tidak terletak di daerah kritis.
Keputusan dari uji ini adalah tidak tolak H0.

p = 0.4478844782641115

Nilai p tidak lebih kecil dari tingkat signifikansi yang diinginkan
Keputusan dari uji ini adalah tidak tolak H0.

Kesimpulan:

Proporsi nilai total Sulfat Dioxide yang lebih dari 40, adalah tidak berbeda dengan 50%.