Hypothesis Testing

Python Programming Lab

05506231 Statistics and Probability

Asst. Prof. Dr.Anantaporn Hanskunatai

Outline

- Z-test
- t-test on different mean(μ_x-μ_y)
- Paired t-test
- Chi-square test
- Case Study 1: HeartDisease Dataset
  - Z-test for the Difference in Two Proportions
  - Chi-square test
- Case Study 2: BloodPressure Dataset
  - Paired t-test

Z-test

- Test for p_1-p_2 (Example from lecture slide no.12)
- In a test, a new deodorant was preferred by 320 of 400 people asked in North & 300 of 425 people asked in South
- Is there a difference between two groups at α=5%?
  - Ho: p1 -p2 = 0
  - Ha: p1 - p2 ≠ 0
  - two-tailed test with α = 0.05

$$Z = \frac{(\hat{p}_1-\hat{p}_2)-0}{\sqrt{\hat{p}\,\hat{q}\,\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{(0.8-0.71) - 0}{\sqrt{0.7515x(1-0.7515)/(1/400+1/425)}} = 3.1264$$

p-value = 0.00177

Z-test

```
from statsmodels.stats.proportion import proportions_ztest
import numpy as np
```

Set alpha , sample size

```
significance = 0.05     #alpha value
successes = np.array([320, 300])
samples = np.array([400, 425])
```

Compute z-statistics and p-value

```
stat,p_value = proportions_ztest(count=successes,nobs=samples,alternative='two-sided')
```

Z-test

Show the outputs: z-statistics, p-value, and conclusion

```
print('z_stat: %0.5f, p_value: %0.5f' % (stat, p_value))
if p_value < significance:
   print ("Reject the null hypothesis")
else:
   print ("Accept the null hypothesis")
```

Result from program

```
z_stat: 3.12644, p_value: 0.00177
Reject the null hypothesis
```

T-test: Example 1

- Test on different means μ_x-μ_y (Example from lecture slide no.15)

Brand A	43	53	65	49	55	60	47	50	60	55
Brand B	62	43	54	67	59	45	46	63	65	45

  - Ho: $\mu_A - \mu_B = 0$
  - Ha: $\mu_A - \mu_B \neq 0$
  - two-tailed test with $\alpha = 0.05$

$$t = \frac{(\bar{X}_A - \bar{Y}_B) - (\mu_A - \mu_B)}{\sqrt{s_p^2\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}} = \frac{(53.7 - 54.9)}{\sqrt{66.95(2/10)}} = -0.328$$

p-value = 0.7467

T-test: Example 1

```
import numpy as np
import scipy.stats as stats
```

Set alpha , sample size

```
significance = 0.05
A = np.array([43, 53, 65, 49, 55, 60, 47, 50, 60, 55])
B = np.array([62, 43, 54, 67, 59, 45, 46, 63, 65, 45])
```

Compute t-statistics and p-value

```
stat, p_value = stats.ttest_ind(a = A, b = B, equal_var = True)
```

T-test: Example 1

Show the outputs: t-statistics, p-value, and conclusion

```
print('t_stat: %0.5f, p_value: %0.4f' % (stat, p_value))
if p_value < significance:
   print ("Reject the null hypothesis")
else:
   print ("Accpet the null hypothesis")
```

Result from program

```
t_stat: -0.32795, p_value: 0.7467
Accpet the null hypothesis
```

T-test: Example 2

- Test on different means μ_x-μ_y (Example from lecture slide no.16)
- $\sigma_x^2 \neq \sigma_y^2$ (unknown)

Brand A	43	53	65	49	55	60	**147**	50	60	55
Brand B	62	43	54	67	59	45	46	63	65	45

  - Ho: μ_A-μ_B = 0
  - Ha: μ_A-μ_B ≠ 0
  - two-tailed test with α = 0.05

$$t = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}} = \frac{(63.7 - 54.9)}{\sqrt{896.23/10+88.8/10}} = 0.887$$

p-value = 0.3946

T-test: Example 2

```
import numpy as np
import scipy.stats as stats
```

Set alpha , sample size

```
significance = 0.05
A = np.array([43, 53, 65, 49, 55, 60, 147, 50, 60, 55])
B = np.array([62, 43, 54, 67, 59, 45, 46, 63, 65, 45])
```

Compute t-statistics and p-value

```
stat, p_value = stats.ttest_ind(A, B, equal_var=False)
```

T-test: Example 2

Show the outputs: t-statistics, p-value, and conclusion

```
print('t_stat: %0.5f, p_value: %0.4f' % (stat, p_value))
if p_value < significance:
   print ("Reject the null hypothesis")
else:
   print ("Accpet the null hypothesis")
```

Result from program

```
t_stat: 0.88668, p_value: 0.3946
Accpet the null hypothesis
```

Paired t-Test: Example 3

- Paired t-Test (two samples are dependent) (Example from lecture slide no.17)
- Test on different in pair (μ_D)

Student	1	2	3	4	5	6	7	8	9
Pretest	60	45	80	87	79	75	60	30	45
Posttest	75	65	90	80	89	95	85	69	40

- Ho: $\mu_D = 0$
- Ha: $\mu_D \neq 0$
- two-tailed test with $\alpha = 0.01$

$$t = \frac{\bar{d} - \mu_D}{s_D/\sqrt{n}} = \frac{-14.1 - 0}{14.37/\sqrt{9}} = -2.945$$

p-value = 0.019

Paired t-Test: Example 3

```
import numpy as np
import scipy.stats as stats
```

Set alpha , sample size

```
significance = 0.01
group1 = np.array([60, 45, 80, 87, 79, 75, 60, 30, 45])
group2 = np.array([75, 65, 90, 80, 89, 95, 85, 69, 40])
```

Compute t-statistics and p-value

```
stat, p_value = stats.ttest_rel(group1,group2)
```

Paired t-Test: Example 3

Show the outputs: t-statistics, p-value, and conclusion

```
print('t_stat: %0.5f, p_value: %0.4f' % (stat, p_value))
if p_value < significance:
   print ("Reject the null hypothesis")
else:
   print ("Accpet the null hypothesis")
```

Result from program

```
t_stat: -2.94514, p_value: 0.0186
Accpet the null hypothesis
```

Chi-Square Test: Example 1

- Chi-Square Test (Example from lecture slide no.22)
- If there is a relationship between gender and married status at α=5%

	Male	Female	Total
Married	25 (24)	15 (16)	40
Single	35 (36)	25 (24)	60
Total	60	40	100

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$$= \frac{(25-24)^2}{24} + \frac{(15-16)^2}{16} + \frac{(35-36)^2}{36} + \frac{(25-24)^2}{24}$$

$$= 0.1736$$

p-value = 0.68

H_0: Gender & married status are independent

α < p-value (0.05 < 0.68) then cannot reject H_0

Chi-Square Test: Example 1

```
import pandas as pd
from scipy.stats import chi2_contingency
```

```
df = pd.DataFrame(index=["Married","Single"], data={'Male': [25, 35],'Female': [15,25]})
```

```
chi2, p, dof, expected = chi2_contingency(df,correction=False)
print(f"chi2 statistic:     {chi2:.5g}")
print(f"p-value:            {p:.5g}")
print(f"degrees of freedom: {dof}")
print("expected frequencies:")
print(expected)
```

```
df
```

	Male	Female
Married	25	15
Single	35	25

```
chi2 statistic:     0.17361
p-value:            0.67692
degrees of freedom: 1
expected frequencies:
[[24. 16.]
 [36. 24.]]
```

Chi-Square Test: Example 2

- Chi-Square Test (Example from lecture slide no.23)
- Is there a relationship between preference and group at α=1% ?

	Prefer X	Prefer Y	No Preference	Total
White	525 (502.2)	105 (149.8)	50 (28)	680
Black	400 (461.6)	200 (137.7)	25 (25.7)	625
Hispanic	600 (561.3)	150 (167.5)	10 (31.3)	760
Total	1525	455	85	2065

H_0: No relationship among group preferences

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}} = \frac{(525-502.2)^2}{502.2}+\frac{(105-149.8)^2}{149.8}+\frac{(50-28)^2}{28}+\dots+\frac{(10-31.3)^2}{31.3} = 87.1$$

p-value ≈ 0

p-value $< \alpha$ $(0 < 0.01)$ then reject H_0

Chi-Square Test: Example 2

```
row1 = [525,105,50]
row2 = [400,200,25]
row3 = [600,150,10]
data=[row1,row2,row3]
chi2, p, dof, expected = chi2_contingency(data)
print(f"chi2 statistic:     {chi2:.5g}")
print(f"p-value:            {p:.5g}")
print(f"degrees of freedom: {dof}")
print("expected frequencies:")
print(expected)
```

```
chi2 statistic:     87.136
p-value:            5.3406e-18
degrees of freedom: 4
expected frequencies:
[[502.17917676 149.83050847  27.99031477]
 [461.56174334 137.71186441  25.72639225]
 [561.2590799  167.45762712  31.28329298]]
```

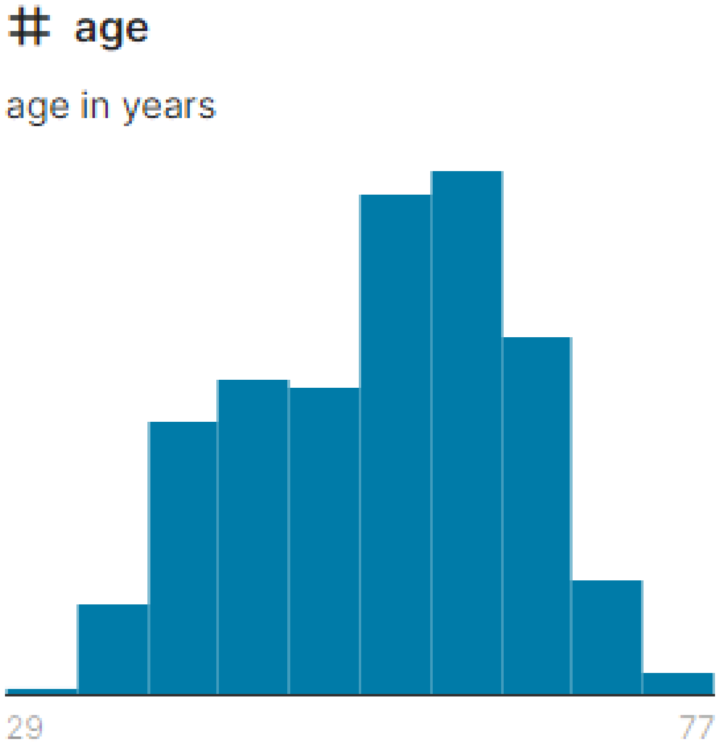
Case Study 1

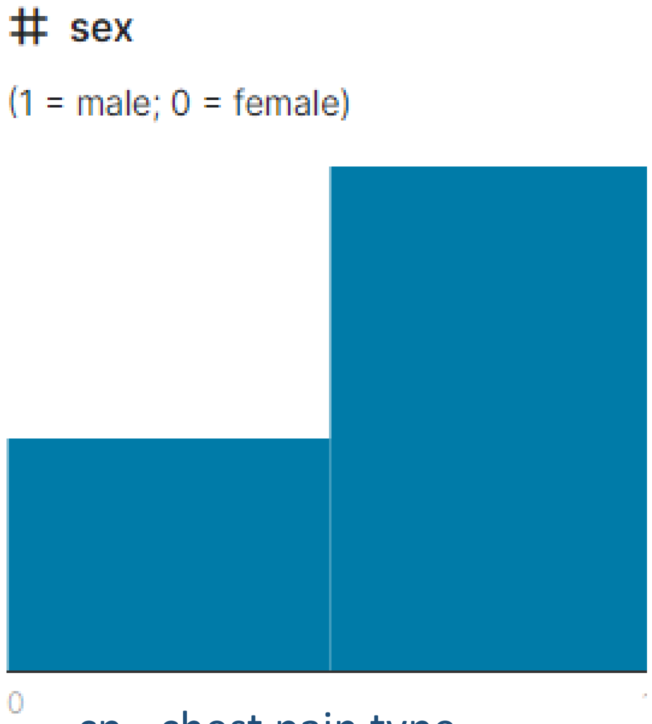
HeartDisease Dataset

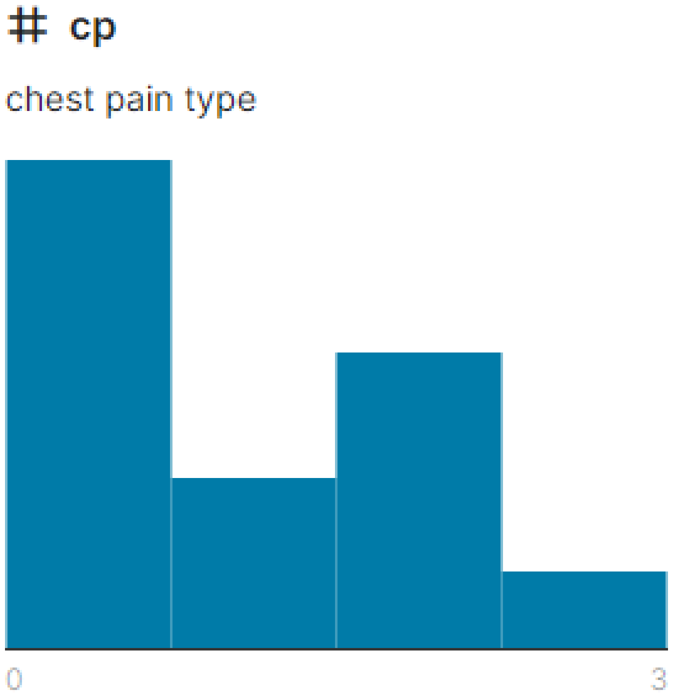
HeartDisease Dataset

1	age	sex	cp	trestbps	chol	fbs	restecg	target
2	63	1	3	145	233	1	0	1
3	37	1	2	130	250	0	1	1
4	41	0	1	130	204	0	0	1
5	56	1	1	120	236	0	1	1
6	57	0	0	120	354	0	1	1
7	57	1	0	140	192	0	1	1
8	56	0	1	140	294	0	0	1
9	44	1	1	120	263	0	1	1
10	52	1	2	172	199	1	1	1
11	57	1	2	150	168	0	1	1
12	54	1	0	140	239	0	1	1
13	48	0	2	130	275	0	1	1
14	49	1	1	130	266	0	1	1
15	64	1	3	110	211	0	0	1
16	58	0	3	150	283	1	0	1
17	50	0	2	120	219	0	1	1

Heart Dataset







cp - chest pain type

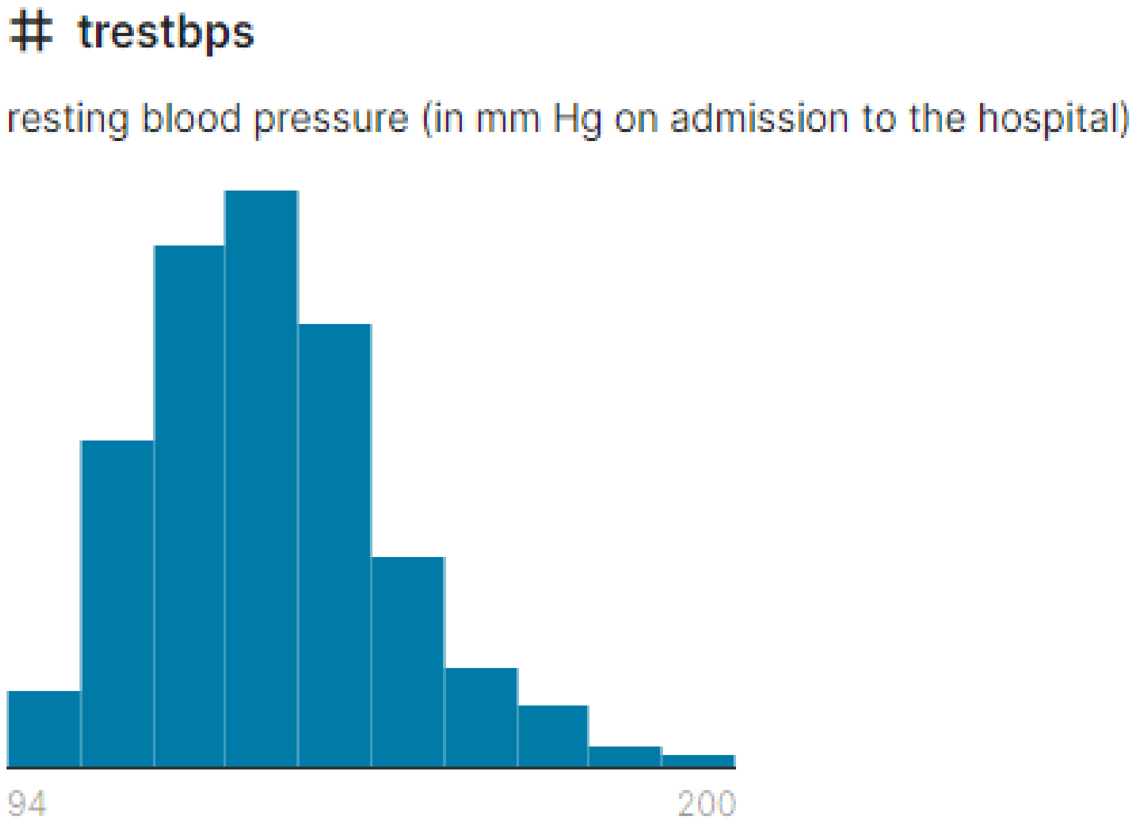
0: Typical angina: chest pain related decrease blood supply to the heart

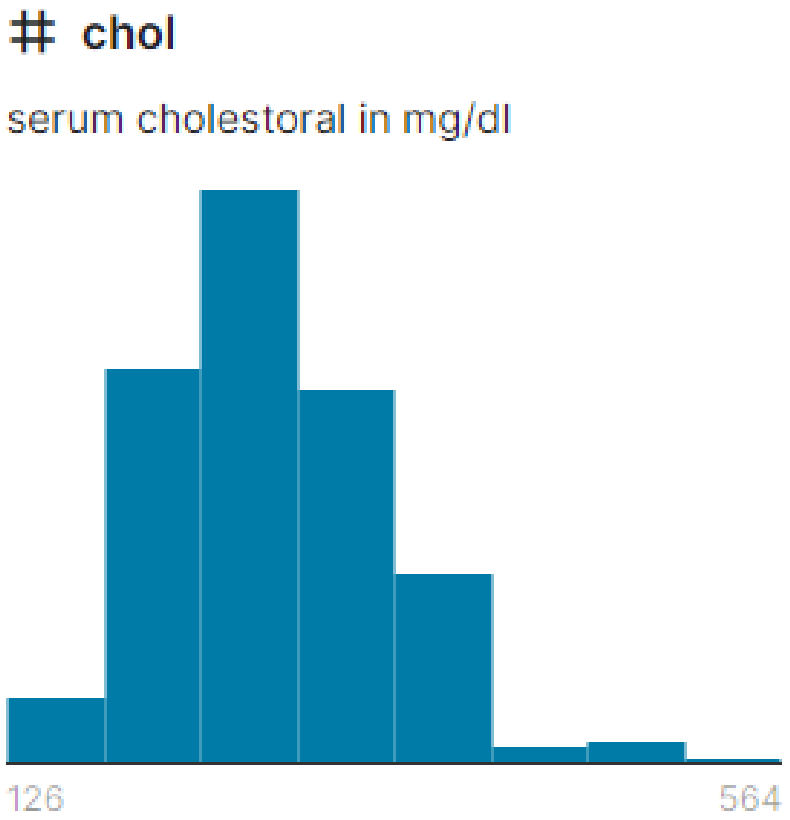
1: Atypical angina: chest pain not related to heart

2: Non-anginal pain: typically esophageal spasms (non heart related)

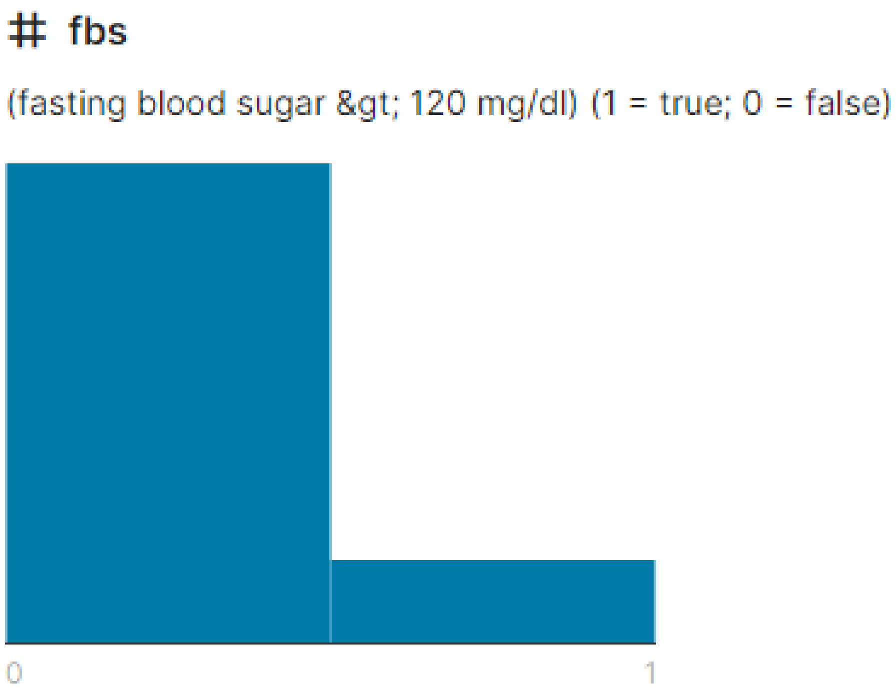
3: Asymptomatic: chest pain not showing signs of disease

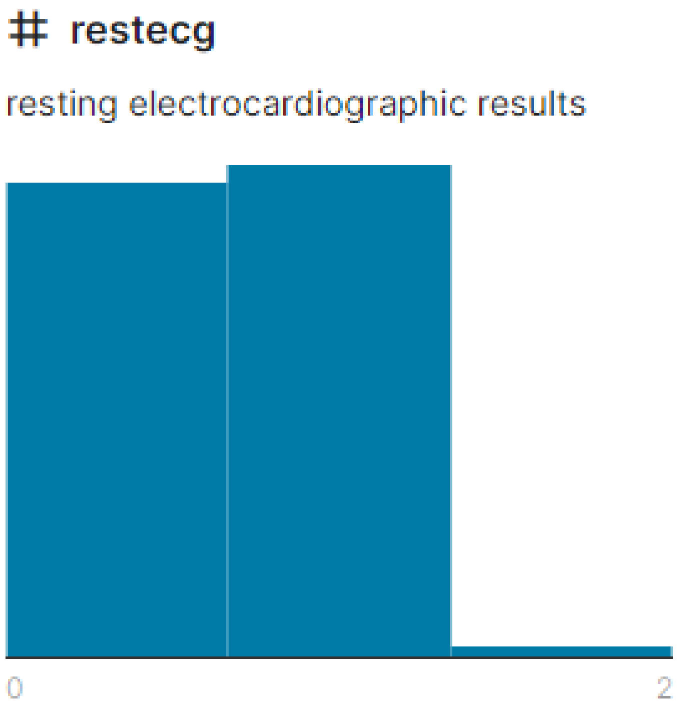
Heart Dataset

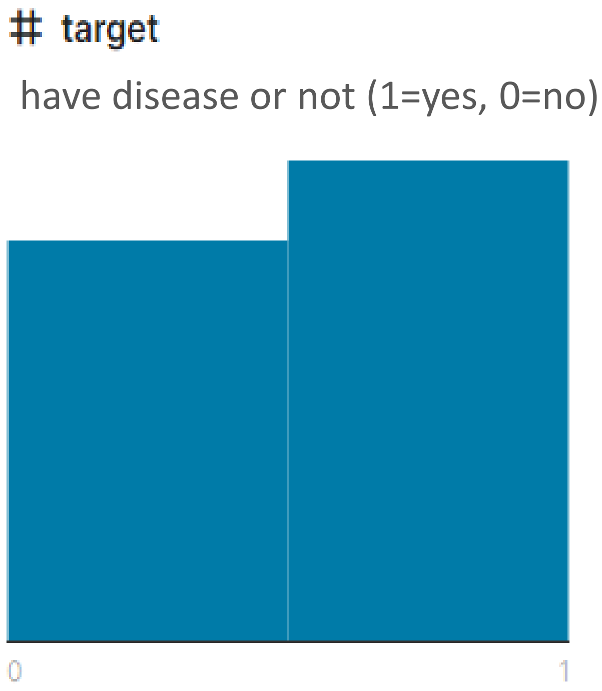




Heart Dataset







fbs - (fasting blood sugar > 120 mg/dl)
(1 = true; 0 = false)

restecg
0: Nothing to note
1: ST-T Wave abnormality
2: Possible or definite left ventricular hypertrophy

Z-Test for the Difference in Two Proportions: Heart Disease

- Is there a difference between the proportion of females with heart and the proportion of males with heart disease?
- Use z-test for p_1-p_2
  - Ho: p1 -p2 = 0
  - Ha: p1 - p2 ≠ 0
  - two-tailed test with $\alpha = 0.01$

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}\,\hat{q}\,\frac{1}{n_1} + \frac{1}{n_2}}}$$

- p1 is the proportion of females having heart disease
- p2 is the proportion of males having heart disease

Z-Test for the Difference in Two Proportions: Heart Disease

Upload dataset

```
from google.colab import files
uploaded = files.upload()
```

```
import numpy as np
import pandas as pd
df = pd.read_csv('HeartDisease.csv')
```

p dataframe

Gender	HeartDisease	Total
Female	72	96
Male	93	207

Generate Gender-HeartDisease Dataframe

```
df['Gender'] = df.sex.replace({1: 'Male', 0: 'Female'})
p = df.groupby('Gender')['target'].agg([lambda z: np.sum(z==1),'size'])
p.columns = ['HeartDisease', 'Total']
p
```

Z-Test for the Difference in Two Proportions: Heart Disease

Compute z-statistics and p-value

```
from statsmodels.stats.proportion import proportions_ztest
import numpy as np
```

```
significance = 0.01
successes = np.array([ p.HeartDisease.Female, p.HeartDisease.Male])
samples = np.array([p.Total.Female, p.Total.Male])
```

```
stat, p_value = proportions_ztest(count=successes, nobs=samples,  alternative='two-sided')
```

alternative : `str` `in` ['two-sided', 'smaller', 'larger']

The alternative hypothesis can be either two-sided or one of the one- sided tests, smaller means that the alternative hypothesis is `prop < value` and larger means `prop > value`. In the two sample test, smaller means that the alternative hypothesis is `p1 < p2` and larger means `p1 > p2` where `p1` is the proportion of the first sample and `p2` of the second one.

Z-Test for the Difference in Two Proportions: Heart Disease

Show the outputs: z-statistics, p-value, and conclusion

```
print('z_stat: %0.5f, p_value: %0.6f' % (stat, p_value))
if p_value < significance:
   print ("Reject the null hypothesis")
else:
   print ("Accept the null hypothesis")
```

Result from program

```
z_stat: 4.89023, p_value: 0.000001
Reject the null hypothesis
```

Chi-square Test

- If there is a relationship between sex and heart disease at $\alpha=1\%$
- H_0: sex and heart disease are independent

```
df['target'].replace({1:'Yes', 0:'No'},inplace=True)
```

	age	sex	cp	trestbps	chol	fbs	restecg	target	Gender
0	63	1	3	145	233	1	0	Yes	Male
1	37	1	2	130	250	0	1	Yes	Male
2	41	0	1	130	204	0	0	Yes	Female
3	56	1	1	120	236	0	1	Yes	Male
4	57	0	0	120	354	0	1	Yes	Female
...	...	...	...	...	...	...	...	...	...
298	57	0	0	140	241	0	1	No	Female
299	45	1	3	110	264	0	1	No	Male
300	68	1	0	144	193	1	1	No	Male
301	57	1	0	130	131	0	1	No	Male
302	57	0	1	130	236	0	0	No	Female

Chi-square Test

```
Table1 = pd.crosstab(df.Gender, df.target, margins=True)
```

```
table1
```

target	No	Yes	All
Gender			
Female	24	72	96
Male	114	93	207
All	138	165	303

```
Table1 = pd.crosstab(df.Gender, df.target)
```

```
Table1
```

target	No	Yes
Gender		
Female	24	72
Male	114	93

Chi-square Test

```python
from scipy.stats import chi2_contingency
chi2, p, dof, expected = chi2_contingency(Table1,correction=False)
print(f"chi2 statistic:     {chi2:.5g}")
print(f"p-value:            {p:.5g}")
print(f"degrees of freedom: {dof}")
print("expected frequencies:")
print(expected)
significance = 0.01
if p < significance:
   print ("sex and have heart disease are dependent")
else:
   print ("sex and have heart disease are independent")
```

```
chi2 statistic:     23.914
p-value:            1.0072e-06
degrees of freedom: 1
expected frequencies:
[[ 43.72277228  52.27722772]
 [ 94.27722772 112.72277228]]
sex and have heart disease are dependent
```

Case Study 2

BloodPressure Dataset

BloodPressure.csv

patient	sex	agegrp	bp_before	bp_after
1	Male	30-45	143	153
2	Male	30-45	163	170
3	Male	30-45	153	168
4	Male	30-45	153	142
5	Male	30-45	146	141
6	Male	30-45	150	147
7	Male	30-45	148	133
8	Male	30-45	153	141
9	Male	30-45	153	131
10	Male	30-45	158	125
11	Male	30-45	149	164
12	Male	30-45	173	159
13	Male	30-45	165	135
14	Male	30-45	145	159
15	Male	30-45	143	153

Paired t-Test: Blood Pressure Difference

- Measure effect on blood pressure medicine in same group of peoples
- Use paired t-test to analyze the blood pressure before and after treatment to test if the treatment (medicine) has a significant affect on the blood pressure
  - Ho: $\mu_D = 0$
  - Ha: $\mu_D \neq 0$
  - two-tailed test with $\alpha = 0.01$

Paired t-Test: Blood Pressure Difference

Upload dataset

```
from google.colab import files
uploaded = files.upload()
```

```
import numpy as np
import pandas as pd
import scipy.stats as stats
df = pd.read_csv('BloodPressure.csv')
```

```
df[['bp_before','bp_after']].describe()
```

	bp_before	bp_after
count	120.000000	120.000000
mean	156.450000	151.358333
std	11.389845	14.177622
min	138.000000	125.000000
25%	147.000000	140.750000
50%	154.500000	149.500000
75%	164.000000	161.000000
max	185.000000	185.000000

The blood pressure before the treatment was higher (156.45 ± 11.39) compared to the blood pressure after treatment (151.36 ± 14.18)

Paired t-Test: Blood Pressure Difference

Compute t-statistics and p-value

```
significance = 0.01
stat,p_value = stats.ttest_rel(df['bp_before'], df['bp_after'])
```

Show the outputs: t-statistics, p-value, and conclusion

```
print('t_stat: %0.5f, p_value: %0.4f' % (stat, p_value))
if p_value < significance:
   print ("Reject the null hypothesis")
else:
   print ("Accpet the null hypothesis")
```

Result from program

```
t_stat: 3.33719, p_value: 0.0011
Reject the null hypothesis
```

There is a statistically significant decrease in blood pressure