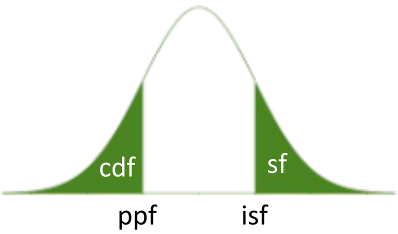
stats.<t/norm/chi2/f>.???



help(<ชื่อฟังก์ชัน>)

H_0 ใช้ $=, \leq, \geq$
H_1 ใช้ $\neq, >, <$

```
np.mean()
np.sqrt()
np.std(??? , ddof=???)
np.var(??? , ddof=???)
np.array([???]) #1d
```

```
np.array([[???],  [???]]) #2d
O = pd.crosstab(???, ???)
O = np.array(O)
n = df.shape[0]
```

การประมาณ μ แบบช่วง:

- ทราบ σ^2

1 $x - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leq \mu \leq x + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$

- ไม่ทราบ σ^2 แต่ $n \geq 30$

2 $x - Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} \leq \mu \leq x + Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$

- ไม่ทราบ σ^2, $n < 30$

3 $\boldsymbol{x - t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} \leq \mu \leq x + t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}}, v = n - 1$

การประมาณ P แบบช่วง:

4 $\hat{p} - Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}} \leq P \leq \hat{p} + Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$

การประมาณ σ^2 แบบช่วง:

5 $\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}}$

การทดสอบ μ 1 ประชากร:

- ทราบ σ^2

6 $Z = \frac{\overline{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$

- ไม่ทราบ σ^2 แต่ $n \geq 30$

7 $Z = \frac{\overline{x}-\mu_0}{\frac{s}{\sqrt{n}}}$

```
ztest(x1=???,
      value=???,
      alternative=<'larger'/'smaller'/'two-sided'>)
```

- ไม่ทราบ σ^2, $n < 30$

8 $t = \frac{\overline{x}-\mu_0}{\frac{s}{\sqrt{n}}}, v = n - 1$

```
stats.ttest_1samp(a=???,
                  popmean=???,
                  alternative=<'two-sided'/'greater'/'less'>)
```

การทดสอบ μ 2 ประชากรอิสระ:

- ทราบ σ_1^2 และ σ_2^2

$Z = \frac{(\overline{x}_1-\overline{x}_2) - (\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ 9

- ไม่ทราบ σ_1^2 และ σ_2^2 แต่ $n_1, n_2 \geq 30$

$Z = \frac{(\overline{x}_1-\overline{x}_2) - (\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ 10

- ไม่ทราบ σ_1^2 และ σ_2^2 $n_1, n_2 < 30$ รู้ว่า $\sigma_1^2 = \sigma_2^2$

$t = \frac{(\overline{x}_1-\overline{x}_2) - (\mu_1-\mu_2)}{\sqrt{s_p^2(\frac{1}{n_1} + \frac{1}{n_2})}}, v = n_1 + n_2 - 2$ 11

โดยที่ $s_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$

- ไม่ทราบ σ_1^2 และ σ_2^2 $n_1, n_2 < 30$ รู้ว่า $\sigma_1^2 \neq \sigma_2^2$

$t = \frac{(\overline{x}_1-\overline{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}, v = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}$ 12

```
stats.ttest_ind(a=???,
                b=???,
                equal_var=<True/False>,
                alternative=<'two-sided'/'greater'/'less'>)
```

การทดสอบ P 1 ประชากร:

13 $Z = \frac{\hat{p} - P_0}{\sqrt{\frac{P_0Q_0}{n}}}$

```
proportions_ztest(count=???,
                  nobs=???,
                  value=???,
                  alternative=<'two-sided'/'smaller'/'larger'>,
                  prop_var=???<1ประชาคือvalue/2ประชาไม่ต้องระบุ>)
```

การทดสอบ σ^2 1 ประชากร:

$\chi^2 = \frac{(n-1)s^2}{\sigma_0{}^2}$ $v = n - 1$ 16

การทดสอบ P 2 ประชากร:

14 $Z = \frac{(\hat{p}_1-\hat{p}_2) - P_0}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$, $\hat{p} = \frac{n_1\hat{p}_1+n_2\hat{p}_2}{n_1+n_2} = \frac{x_1+x_2}{n_1+n_2}$

การทดสอบ σ^2 2 ประชากร:

dfn dfd

$F = \frac{s_1{}^2/\sigma_1{}^2}{s_2{}^2/\sigma_2{}^2}$ $v_1 = n_1 - 1,$ $v_2 = n_2 - 1$ 17

การตั้งสมมติฐานใช้ $\sigma_1{}^2/\sigma_2{}^2$

การทดสอบ μ 2 ประชากรไม่อิสระ มีลักษณะเป็นคู่:

15 $t = \frac{d-d_0}{\sqrt{\frac{s_d{}^2}{n}}}, v = n - 1$

$\bar{d} = \frac{\sum(x_1 - x_2)}{n}$

$s_d{}^2 = \frac{\sum(x_1 - x_2)^2 - nd^2}{n - 1}$

```
stats.ttest_rel(a=???,
                b=???,
                alternative=<'two-sided'/'greater'/'less'>)
```

Chi-Square Test > ทดสอบอัตราส่วน:

18 $$\chi^2 = \sum_{i=1}^{k} \frac{(\mathrm{O}_i - \mathrm{E}_i)^2}{\mathrm{E}_i}, \nu = k - 1$$

```
stats.chisquare(f_obs=???, f_exp=???)
```

Chi-Square Test > ทดสอบความเป็นอิสระ:

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(\mathrm{O}_{ij} - \mathrm{E}_{ij})^2}{\mathrm{E}_{ij}}, \nu = (r-1)(c-1)$$ 19

$$\mathrm{E}_{ij} = \frac{\mathrm{R_i C_j}}{N}$$

```
chi2_cal, p_val, dof, E = stats.chi2_contingency(???,
                                  correction=False)
```

สัมประสิทธิ์สหสัมพันธ์:

20 $$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2}\sqrt{n\sum Y^2 - (\sum Y)^2}}$$

```
r, p_val = stats.pearsonr(x, y)
```

การทดสอบเกี่ยวกับสัมประสิทธิ์สหสัมพันธ์:

H_0: x กับ y ไม่มีความสัมพันธ์เชิงเส้น

2-tailed

$$t = \frac{r}{S_r}, S_r = \sqrt{\frac{1-r^2}{n-2}}, \nu = n - 2$$ 21

การวิเคราะห์การถดถอยอย่างง่าย:

$$\hat{y} = b_0 + b_1 x_i \qquad b_0 = \bar{y} - b_1\bar{x}$$

$$b_1 = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}$$

```
x = ???
y = ???
x_wconst = sm.add_constant(x)
lr = sm.OLS(y, x_wconst).fit()
print(lr.summary())
```

การทำนาย y:

```
y_hat = lr.predict(x)
```

การทดสอบสมมติฐานค่า $\boldsymbol{\beta_1}$ โดยการวิเคราะห์ความแปรปรวน:

22 $$\mathrm{F} = \frac{\mathrm{MSR}}{\mathrm{MSE}} \qquad \nu_1 = 1, \nu_2 = n - 2$$ right-tailed

$$\mathrm{MSR} = \frac{\mathrm{SSR}}{1}, \ \mathrm{SSR} = b_0\sum y + b_1\sum xy - n\bar{y}^2$$

$$\mathrm{MSE} = \frac{\mathrm{SSE}}{n-2}, \ \mathrm{SSE} = \mathrm{SST} - \mathrm{SSR}, \ \mathrm{SST} = \sum y^2 - n\bar{y}^2$$

สัมประสิทธิ์การตัดสินใจ:

23 $$r^2 = \frac{\mathrm{SSR}}{\mathrm{SST}} = 1 - \frac{\mathrm{SSE}}{\mathrm{SST}}$$

การเปิด Jupyter Lab :

เปิดโปรแกรม Anaconda Prompt แล้วพิมพ์คำสั่ง

1. เปิดโปรแกรม Anaconda Prompt (คลิกขวาที่ shortcut icon ที่ desktop เลือก run as administrator) แล้วพิมพ์คำสั่ง
3. conda activate ProgStat_Py3_11
4. jupyter lab
5. สร้าง notebook แล้วทำการ import library ตามด้านขวา

คำสั่งในการ import library:

```
import pandas as pd
import numpy as np
from scipy import stats
import statsmodels.api as sm
from statsmodels.stats.weightstats import ztest
from statsmodels.stats.proportion import proportions_ztest
```

*หากเกิด error ไม่พบ library statsmodels ให้พิมพ์คำสั่ง (ใน Notebook):

%pip install statsmodels