

Homework #4

Confidence Interval (CI) of Mean ปริมาณแอลกอฮอล์ของไวน์แดง

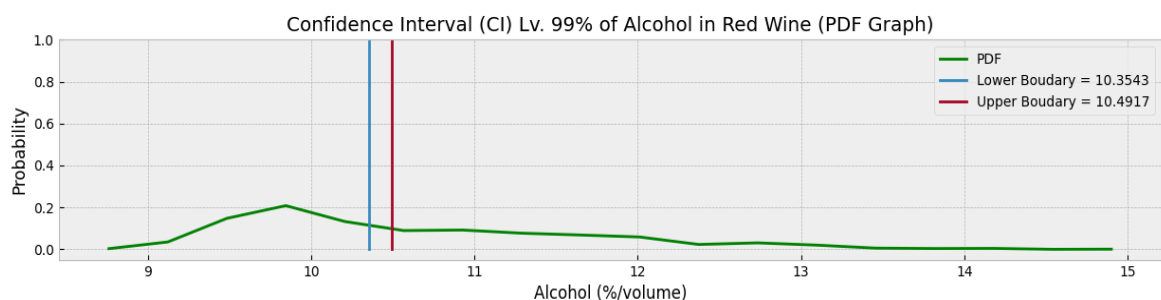
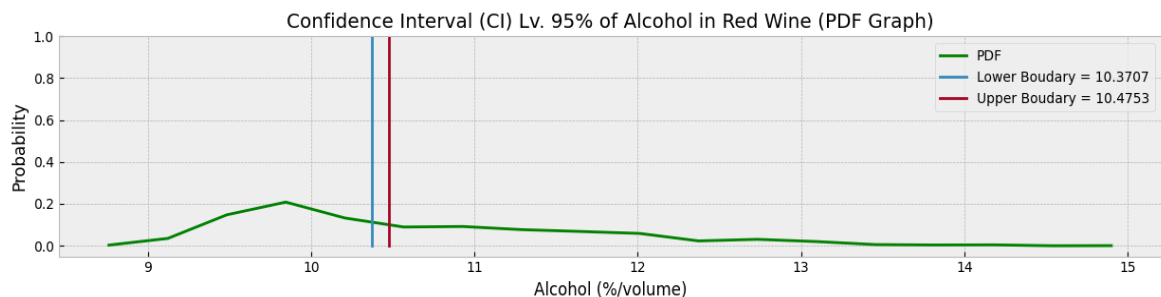
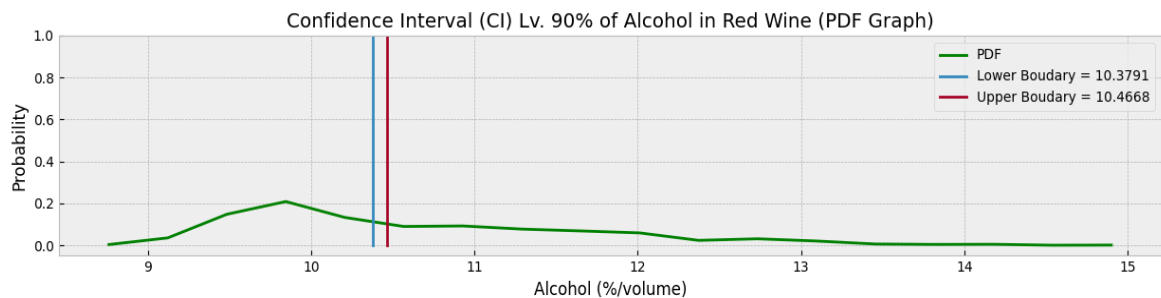
ผู้จัดทำเลือก 1 คอลัมน์ มีทั้งหมด 1599 แถว:

1. Alcohol (%/volume)

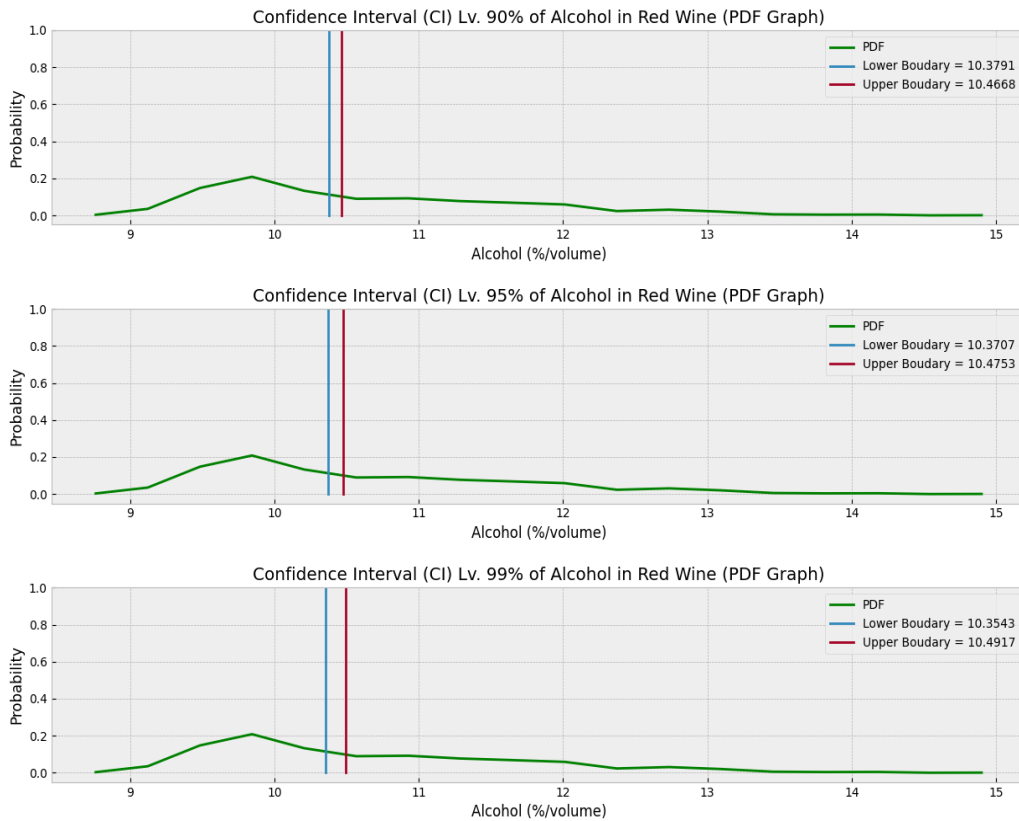
	CI Lv.	Lower Boundary	Mean	Upper Boundary	Margin Error	Z - Score	Standard Error	Standard Deviation
1.Alcohol (%/volume)	90%	10.3791	10.4230	10.4668	0.0439	1.6458	0.0267	1.0657
2.Alcohol (%/volume)	95%	10.3707	10.4230	10.4753	0.0523	1.9614	0.0267	1.0657
3.Alcohol (%/volume)	99%	10.3543	10.4230	10.4917	0.0687	2.5789	0.0267	1.0657

อ้างอิงจากการคำนวณในโปรแกรม WineGraph3.py

กราฟ Confidence Interval (CI) of Mean ปริมาณแอลกอฮอล์ของไวน์แดง



บทวิเคราะห์ข้อมูลจากกราฟทั้งหมด



Confidence Interval (CI) of Mean เป็นช่วงค่าเฉลี่ยที่บอกระดับความมั่นใจของข้อมูล โดยสามารถอิงจากกลุ่มตัวอย่าง (sample) และสามารถอิงไปถึงกลุ่มข้อมูลจริง (population) ทั้งหมดได้

สูตรการคำนวณหา Confidence Interval (CI) of Mean

$$\text{Confidence Interval (CI)} = \bar{x} \pm t_{n-1, \alpha/2} \left(\frac{s}{\sqrt{n}} \right)$$

Margin of error (ME)
 S = Standard Deviation

n = จำนวน Samples

t = t-score หรือ z-score

\bar{x} = ค่าเฉลี่ย

t-score at the given df and alpha level
Standard error (SE) of the mean

โดยจากกราฟ มีกลุ่มตัวอย่างปริมาณแอลกอฮอล์อยู่ทั้งหมด 1599 ข้อมูล เราจะสามารถสรุปได้ดังนี้

1. ช่วงระดับความเชื่อมั่น 90% จะมีปริมาณแอลกอฮอล์เฉลี่ยอยู่ที่ 10.3791 - 10.4668 %/volume
2. ช่วงระดับความเชื่อมั่น 95% จะมีปริมาณแอลกอฮอล์เฉลี่ยอยู่ที่ 10.3707 - 10.4753 %/volume
3. ช่วงระดับความเชื่อมั่น 99% จะมีปริมาณแอลกอฮอล์เฉลี่ยอยู่ที่ 10.3543 - 10.4917 %/volume

เมื่อดูจากภาพด้านบน และจากข้อมูลที่ได้ จะเห็นได้ว่า ทั้ง 3 ช่วง เป็นช่วงที่แคบมาก เพราะเนื่องจากมีปริมาณข้อมูลอยู่มากถึง 1599 ข้อมูล ทำให้เรามั่นใจได้ว่า ช่วงค่าเฉลี่ยที่ได้จะไม่ห่างกันมาก เนื่องจากข้อมูลชุดนี้ มีข้อมูลที่มีความคล้ายคลึงกัน และไปในทิศทางเดียวกัน ยังมีข้อมูลมาก ทำให้ช่วงค่าเฉลี่ยที่ได้ ยิ่งแคบและเล็กลง เนื่องจาก Standard Error มีค่าน้อยลง ทำให้ Margin Error ลดลงไปด้วย

“ เราจึงสามารถสรุปได้ว่า ช่วงทั้ง 3 ช่วงนี้ สามารถอิงไปถึงกลุ่มข้อมูลจริง (Population) ได้ว่า ”

1. ที่ระดับความเชื่อมั่น 90% มีช่วงปริมาณแอลกอฮอล์เฉลี่ยอยู่ที่ 10.3791 - 10.4668 %/volume จะสามารถครอบคลุมปริมาณแอลกอฮอล์เฉลี่ยจริงได้ ที่ระดับความเชื่อมั่นนี้
2. ที่ระดับความเชื่อมั่น 95% มีช่วงปริมาณแอลกอฮอล์เฉลี่ยอยู่ที่ 10.3707 - 10.4753 %/volume จะสามารถครอบคลุมปริมาณแอลกอฮอล์เฉลี่ยจริงได้ ที่ระดับความเชื่อมั่นนี้
3. ที่ระดับความเชื่อมั่น 99% มีช่วงปริมาณแอลกอฮอล์เฉลี่ยอยู่ที่ 10.3543 - 10.4917 %/volume จะสามารถครอบคลุมปริมาณแอลกอฮอล์เฉลี่ยจริงได้ ที่ระดับความเชื่อมั่นนี้

สรุปได้ว่า หากเราทำการผลิตไวน์ขึ้นมาโดยโรงงานผลิตไวน์นี้ สามารถมีความมั่นใจได้ว่า หากมีการผลิตเพิ่มเติม และมีการสุ่มเป็นกลุ่มตัวอย่างอีกครั้ง ค่าเฉลี่ยปริมาณแอลกอฮอล์ในไวน์แดงจะอยู่ในช่วงนี้ ซึ่งจะขึ้นกับระดับความเชื่อมั่นที่เรากำหนด โดยค่าเฉลี่ยปริมาณแอลกอฮอล์ในไวน์แดงจะอยู่ในช่วงราว ๆ 10.35 - 10.49 %/volume อย่างแน่นอน



รายละเอียด Source Code ของโปรแกรม WineGraph3.py

```

wineGraph3.py X testcode.py
wineGraph3.py > ...
You, seconds ago | 1 author (You)
1 # defining the libraries
2
3 import numpy as np
4 import matplotlib.pyplot as plt
5 import pandas # collection for data
6
7 import scipy.stats
8
9 # Init Style of Graph and Insert table of data in form of columns
10 plt.style.use('bmh')
11 columns = pandas.read_csv('testgraphredwine.csv')
12
13 # All columns
14 x = columns['alcohol'] # x (independent variable) = alcohol
15
16 # initial value
17 dataArray = 1.0 * np.array(x)
18 print('dataArray :', dataArray)
19 number = len(dataArray)
20 mean = np.mean(dataArray)
21 standardError = scipy.stats.sem(dataArray)
22 # standardError = standard deviation / samples 1.0656771926520383/ 39.98749804626440991456385162254 # from HW 1
23 print('standardError(hw1) :', 1.0656771926520383 / 39.98749804626440991456385162254 ) # from HW 1
24 print('standardError(hw4) :', standardError, '\n')
25
wineGraph3.py X testcode.py
wineGraph3.py > ...
26
27 def confidence_interval(confidence):
28     con = format(confidence, '.2f')
29     print(f'*** confidence of {con} % ***)
30
31     z_score = scipy.stats.t.ppf( (1 + confidence) / 2.0, number-1 )
32     print('z-score :', z_score)
33
34     marginError = standardError * z_score
35     # marginError = standardError * z-score
36
37     print('margin error :', marginError, '\n')
38
39     print('***SUMMARY***')
40     print('mean :', mean)
41     print('lower-upper boundary:', mean - marginError, mean + marginError)
42
43     print('*****\n\n')
44     return mean, mean - marginError, mean + marginError
45
46 mean1, lowerB1, upperB1 = confidence_interval(0.90)
47 mean2, lowerB2, upperB2 = confidence_interval(0.95)
48 mean3, lowerB3, upperB3 = confidence_interval(0.99)
49
50
51 # getting data of the histogram
52 al_count, al_bins_count = np.histogram(x, bins=18) # y (quantity) and x (value)
53
54
55 # finding the PDF of the histogram using count values
56 al_pdf = al_count / sum(al_count)

```

```

wineGraph3.py X testcode.py
wineGraph3.py > ...
58
59
60 figure, al_func = plt.subplots(3, 1, figsize=(8, 10))
61 plt.tight_layout(pad=5, h_pad=5.0)
62
63 y = np.linspace(0,1)
64
65 al_func[0].set_title('Confidence Interval (CI) Lv. 90% of Alcohol in Red Wine (PDF Graph)')
66 al_func[0].set_xlabel("Alcohol (%/volume)")
67 al_func[0].set_ylabel("Probability")
68 al_func[0].plot(al_bins_count[1:], al_pdf, color="green", label="PDF" )
69 x1 = np.linspace(lowerB1,lowerB1)
70 x2 = np.linspace(upperB1,upperB1)
71 al_func[0].plot(x1,y, label="Lower Boudary = {:.4f}".format(lowerB1))
72 al_func[0].plot(x2,y, label="Upper Boudary = {:.4f}".format(upperB1))
73 al_func[0].legend()
74 al_func[0].axis(ymax=1)
75
76 al_func[1].set_title('Confidence Interval (CI) Lv. 95% of Alcohol in Red Wine (PDF Graph)')
77 al_func[1].set_xlabel("Alcohol (%/volume)")
78 al_func[1].set_ylabel("Probability")
79 al_func[1].plot(al_bins_count[1:], al_pdf, color="green", label="PDF" )
80 x1 = np.linspace(lowerB2,lowerB2)
81 x2 = np.linspace(upperB2,upperB2)
82 al_func[1].plot(x1,y, label="Lower Boudary = {:.4f}".format(lowerB2))
83 al_func[1].plot(x2,y, label="Upper Boudary = {:.4f}".format(upperB2))
84 al_func[1].legend()
85 al_func[1].axis(ymax=1)
86
87 al_func[2].set_title('Confidence Interval (CI) Lv. 99% of Alcohol in Red Wine (PDF Graph)')
88 al_func[2].set_xlabel("Alcohol (%/volume)")
89 al_func[2].set_ylabel("Probability")
90 al_func[2].plot(al_bins_count[1:], al_pdf, color="green", label="PDF" )
91 x1 = np.linspace(lowerB3,lowerB3)
92 x2 = np.linspace(upperB3,upperB3)
93 al_func[2].plot(x1,y, label="Lower Boudary = {:.4f}".format(lowerB3),)
94 al_func[2].plot(x2,y, label="Upper Boudary = {:.4f}".format(upperB3))
95 al_func[2].legend()
96 al_func[2].axis(ymax=1)
97
98
99 plt.show()
100

```

OUTPUT ของโปรแกรม WineGraph3.py

```
[Running] python -u "c:\Users\ASUS\Desktop\Prob-stat\wineGraph3.py"
dataArray : [ 8.4  8.4  8.5 ... 14.  14.  14.9]
standardError(hw1) : 0.026650259324028716
standardError(hw4) : 0.026650259324028723

*** confidence of 0.90 % ***
z-score : 1.645807731000542
margin error : 0.04386120282865575

***SUMMARY***
mean : 10.422983114446529
lower-upper boundary: 10.379121911617874 10.466844317275184
*****

*** confidence of 0.95 % ***
z-score : 1.9614496156420809
margin error : 0.05227314090787792

***SUMMARY***
mean : 10.422983114446529
lower-upper boundary: 10.37070997353865 10.475256255354408
*****

*** confidence of 0.99 % ***
z-score : 2.5789094543589206
margin error : 0.06872860573185464

***SUMMARY***
mean : 10.422983114446529
lower-upper boundary: 10.354254508714675 10.491711720178383
*****
```

แหล่งที่มาของชุดข้อมูล (Reference/URL) :

- ที่มาของชุดข้อมูล Winequality-red.csv

<https://www.kaggle.com/uciml/red-wine-quality-cortez-et-al-2009>

- ที่มาคำอธิบายแต่ละส่วนประกอบของไวน์

<https://archive.ics.uci.edu/ml/datasets/wine+quality>

https://rstudio-pubs-static.s3.amazonaws.com/57835_c4ace81da9dc45438ad0c286bcbb4224.html

<https://waterlibrary.com/th-รู้ไหมว่า-ระดับปริมาณแ:/#:~:text=ปัจจุบันมีแอลกอฮอล์อยู่ใน,สูงขึ้นด้วยเช่นกัน>

- วิธีการทำไวน์

https://www.youtube.com/watch?v=7gguYRxLMFI&ab_channel=Insider

- ประเภทของไวน์

<https://www.unlockmen.com/terrazas-unlock-wine-101-1/>

<https://thewinelist.shop/blog/news/wine-101>

- รายละเอียดอื่นๆ เกี่ยวกับคุณภาพและวิธีรับรสที่ดีของไวน์

<https://www.blockdit.com/posts/5e5f68d77b00780ed6462939>

<https://www.dummies.com/food-drink/drinks/wine/the-special-technique-for-tasting-wine/>

<https://www.quickanddirtytips.com/house-home/entertaining/wine/4-ways-to-know-if-your-wine-is-good>