


```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
import scipy.stats as stats
```

✓ Sampling Distribution

Q1. Suppose an automobile battery manufacturer claims that the mean lifetime of their battery is 60 months with a standard deviation of 6 months. Suppose the distribution of battery life is approximately normal. Find the probability that the mean lifetime of 40 randomly sampled batteries will be less than 58 months.


```
from scipy.stats import norm
mu=60
sigma=6
n=40
x_bar=58
se=sigma/(n**0.5)
z=(x_bar-mu)/se
prob=norm.cdf(z)
print(prob)
```

 0.017507490509831244

✓ Interval Estimation

Q2. A random sample of 40 households was selected as part of a study on electricity usage, and the number of kilowatt-hours (kWh) was recorded for each household in the sample for the first quarter of 2020. The average usage was found to be 310 kWh. In a very large study in the first quarter of the previous year, it was found that the standard deviation of the usage was 89 kWh. Assuming the standard deviation is unchanged and that the usage is normally distributed, provide an expression for calculating a 95% confidence interval for the mean usage in the first quarter of 2019.

```
from scipy.stats import norm
x_bar=310
sigma=89
n=40
confidence_level=0.95
se=sigma/(n**0.5)
z=norm.ppf(1-(1-confidence_level)/2)
lower_bound=x_bar-z*se
upper_bound=x_bar+z*se
print(lower_bound)
print(upper_bound)
```

 282.419121062447
337.580878937553

✓ Hypothesis Testing

Q3. You are a manager of a Chinese restaurant. You want to determine whether the mean waiting time to place an order has changed in the past month from its previous population mean value of 4.5 minutes. State the null and alternative hypotheses.

```
from scipy.stats import t
pop_mean=4.5
samp_mean=4.8
sample_std=0.4
n=30
alp=0.05
se=sample_std/(n**0.5)
t_stat=(samp_mean-pop_mean)/se
p_val=2*(1-t.cdf(abs(t_stat),df=n-1))
if p_val<alp:
    print("Reject the null hypothesis")
else:
```

```
print("Fail to reject the null hypothesis")
print(t_stat)
```

```
↗ Reject the null hypothesis
4.107919181288744
```

Q4. What is the p-value in a two-tailed z-test for one sample, where the computed test statistic (z- stat) is equal to +2.00?

```
from scipy.stats import norm
z_stat=2.00
p_val=2*(1-norm.cdf(abs(z_stat)))
print(p_val)
```

```
↗ 0.04550026389635842
```

Q5. Samy, Product Manager of K2 Jeans, wants to launch a product line into a new market area. A Survey of a random sample of 400 households in that market showed a mean income per household of 30000 rupees. The standard deviation based on an earlier pilot study of households is 8000 rupees. Samy strongly believes the product line will be adequately profitable only in markets where the mean household income is greater than 29000 rupees. Samy wants our help in deciding whether the product line should be introduced in the new market. Perform statistical analysis and based on that draw a conclusion. Assume a level of significance (α) of 5% $1/2$

```
from scipy.stats import norm
samp_mean=30000
pop_mean=29000
sig=8000
n=400
alp=0.05
se=sig/(n**0.5)
z_score=(samp_mean-pop_mean)/se
p_val=1-norm.cdf(z_score)
if p_val<alp:
    print("Reject the null hypothesis")
else:
    print("Fail to reject the null hypothesis")
print(z_score)
print(p_val)
```

```
↗ Reject the null hypothesis
2.5
0.006209665325776159
```