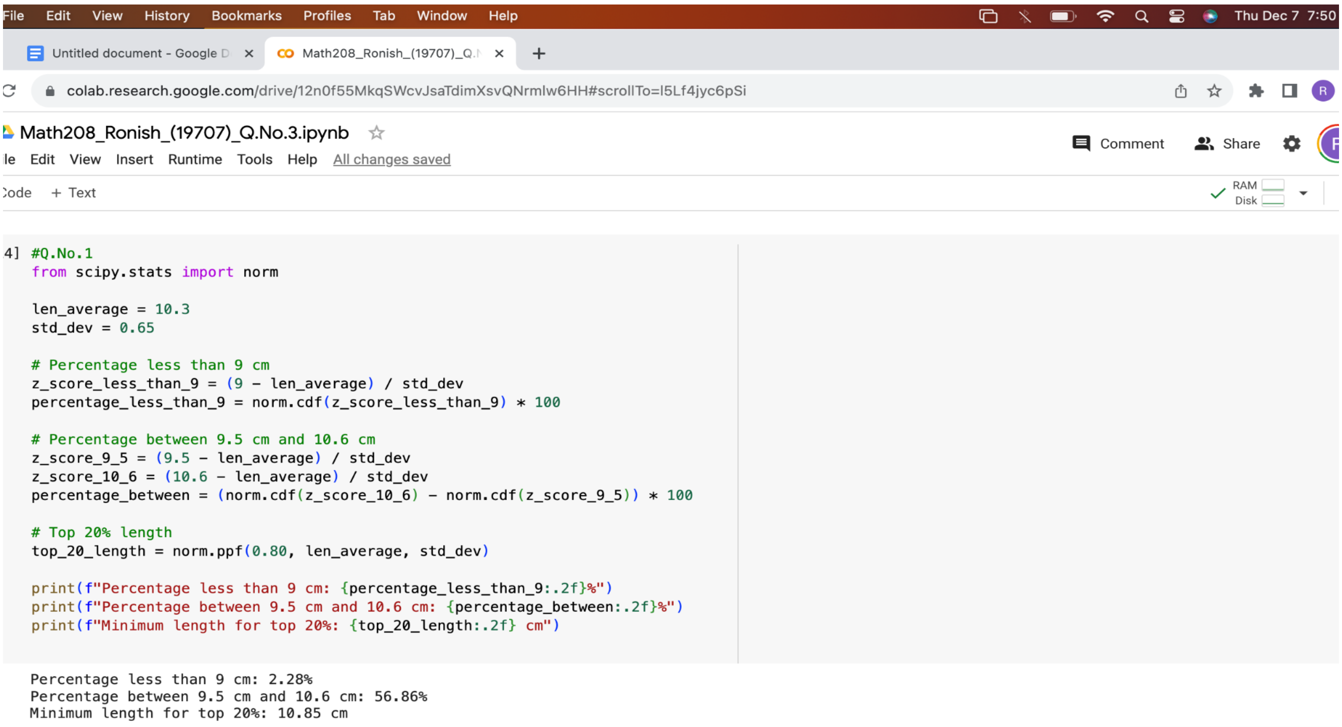
Ronish Shrestha (19707)
Assignment 4

Q.No.1) Assuming that the lengths of American anchovies appease the normal distribution with the mean μ = 10.3cm and standard deviation σ = 0.65cm, please find the percentages of the lengths in the population of American anchovies. a. Less than 9cm. b. Between 9.5cm and 10.6cm. c. What is the minimum length if a restaurant claimed that the lengths of the sold anchovies are in the top of 20%?

Solution:



```python
#Q.No.1
from scipy.stats import norm

len_average = 10.3
std_dev = 0.65

# Percentage less than 9 cm
z_score_less_than_9 = (9 - len_average) / std_dev
percentage_less_than_9 = norm.cdf(z_score_less_than_9) * 100

# Percentage between 9.5 cm and 10.6 cm
z_score_9_5 = (9.5 - len_average) / std_dev
z_score_10_6 = (10.6 - len_average) / std_dev
percentage_between = (norm.cdf(z_score_10_6) - norm.cdf(z_score_9_5)) * 100

# Top 20% length
top_20_length = norm.ppf(0.80, len_average, std_dev)

print(f"Percentage less than 9 cm: {percentage_less_than_9:.2f}%")
print(f"Percentage between 9.5 cm and 10.6 cm: {percentage_between:.2f}%")
print(f"Minimum length for top 20%: {top_20_length:.2f} cm")
```

```
Percentage less than 9 cm: 2.28%
Percentage between 9.5 cm and 10.6 cm: 56.86%
Minimum length for top 20%: 10.85 cm
```

a. Less than 9 cm = 2.28%
b. Between 9.5 and 10.6 cm = 56.86%
c. Min length for top 20% = 10.85cm

Q.No.2) If the random variables X and Y are normal distributions with μ = 10 & σ = 3 and μ = 15 & σ = 8, namely, X~N(10, 3) and Y~N(15, 8), and they are independent, what is the probability distribution and statistical parameters of (1) X + Y (2) X — Y (3) 3X (4) 4X + 5Y

Solution:

If X and Y are independent normal distributions with means (μ) and standard deviations (σ) as described, we can calculate the probability distribution and statistical parameters for the following combinations:

(1) X + Y:

Explanation:
The sum of two independent normal distributions is also a normal distribution. The mean of the sum is the sum of the means, and the variance of the sum is the sum of the variances.

Mean (μ_X+Y) = μ_X + μ_Y = 10 + 15 = 25
Variance (σ^2_X+Y) = σ^2_X + σ^2_Y = 3^2 + 8^2 = 9 +64 = 73
SD = $\sqrt{variance}$ = $\sqrt{73}$
So,the probability distribution for is (X + Y)~ N(25, √73).

(2) X - Y:
Similarly, the difference of two independent normal distributions is a normal distribution.

Mean (μ_X-Y) = μ_X - μ_Y = 10 - 15 = -5
Variance (σ^2_X-Y) = σ^2_X + σ^2_Y = 3^2 + 8^2 = 9 + 64 =73
SD = $\sqrt{variance}$ = $\sqrt{73}$
So, the probability distribution for is (X - Y) ~ N(-5, √73).

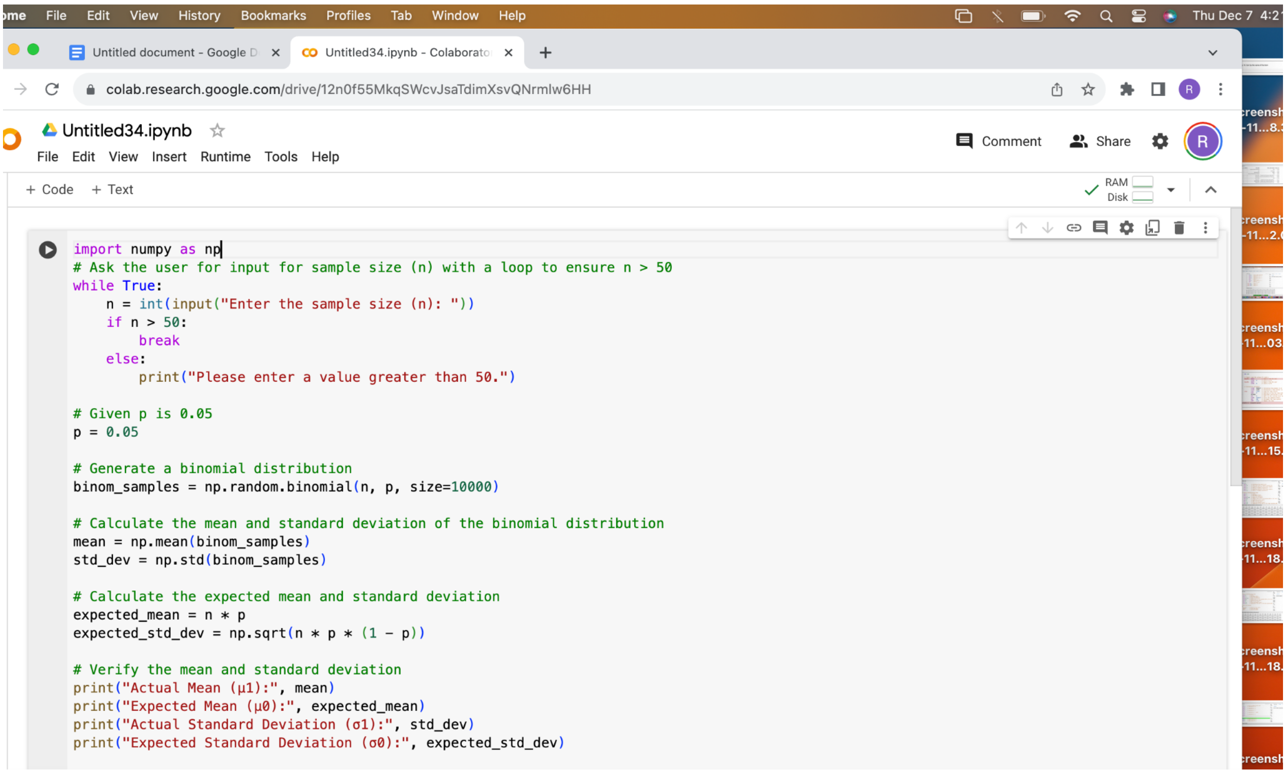
(3) 3X:
Multiplying a normal distribution by a constant scales both the mean and variance.
Mean (μ_3X) = 3μ_X = 3 * 10 = 30
Variance (σ^2_3X) = (3^2) σ^2_X = 81
SD = $\sqrt{81}$ = 9
So, the probability distribution for 3X is 3X ~ N(30, 9).

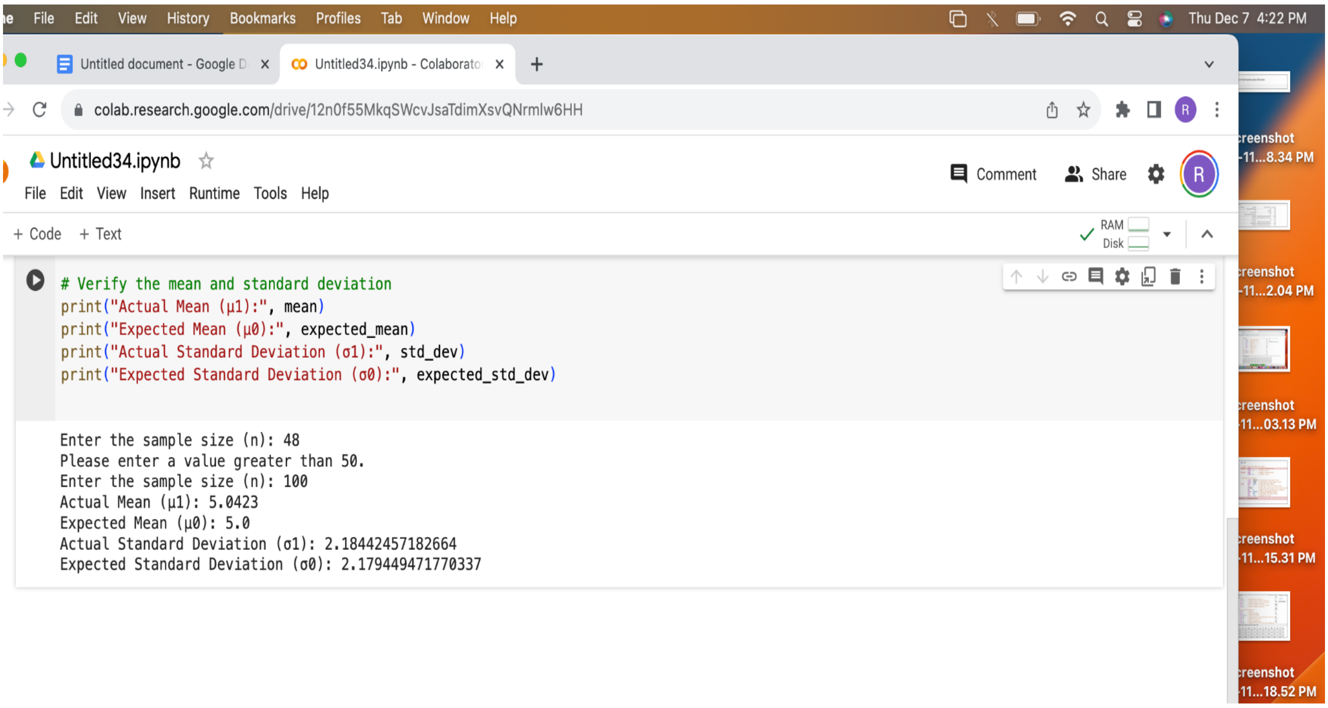
(4) 4X + 5Y:
The sum of scaled normal distributions is also a normal distribution.
Mean (μ_4X+5Y) = 4μ_X + 5μ_Y = 4*10 + 5*15= 40 + 75 = 115
Variance (σ^2_4X+5Y) = ((4.σx)^2 + (5.σ3)^2 = (4.3)^2 + (5.8)^2= 344
SD = √344
So, the probability distribution for 4X + 5Y is 4X + 5Y ~ N(115, √344).

These are the probability distributions and statistical parameters for the given combinations of random variables X and Y.

Q.No.3) For the students in Engineering School, please write Python program to verify the mean $\mu = np$ and standard deviation σ = npq for p=0.05 and selecting any n greater than 50 in binomial distribution.

Solution:





Q.No.4) In general, if np > 5 and nq > 5 in binomial distribution, binomial probabilities can be approximated using the normal distribution. Please select any big enough n and p's values to verify in Python program or Excel and plot the histogram.

Solution:

File Edit View History Bookmarks Profiles Tab Window Help Thu Dec 7 10:35 PM

MATH208_HW_assignment_4 × Math208_Ronish_(19707)_Q.N ×

colab.research.google.com/drive/12n0f55MkqSWcvJsaTdimXsvQNrmlw6HH#scrollTo=sn5N8hNVEvax

Math208_Ronish_(19707)_Q.No.3.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

Code + Text RAM Disk

28]
```
#Q.No.4
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import binom, norm

# Define the parameters
n = 60  # number of trials
p = 0.5  # probability of success

# Generate binomial distribution
binomial_dist = binom(n, p)
x = np.arange(0, n+1)
binomial_probs = binomial_dist.pmf(x)

# Approximate with normal distribution
mean, var = binom.stats(n, p, moments='mv')
normal_dist = norm(mean, np.sqrt(var))
normal_probs = normal_dist.pdf(x)

# Plotting
plt.figure(figsize=(12, 6))
plt.bar(x, binomial_probs, alpha=0.5, label='Binomial Distribution')
plt.plot(x, normal_probs, color='red', linewidth=2, label='Normal Approximation')
plt.xlabel('Number of Successes')
plt.ylabel('Probability')
plt.title('Binomial Distribution vs. Normal Approximation')
plt.legend()
plt.grid(True)
plt.show()
```

Edit View History Bookmarks Profiles Tab Window Help Thu Dec 7 10:35 PM

MATH208_HW_assignment_4 × Math208_Ronish_(19707)_Q.N ×

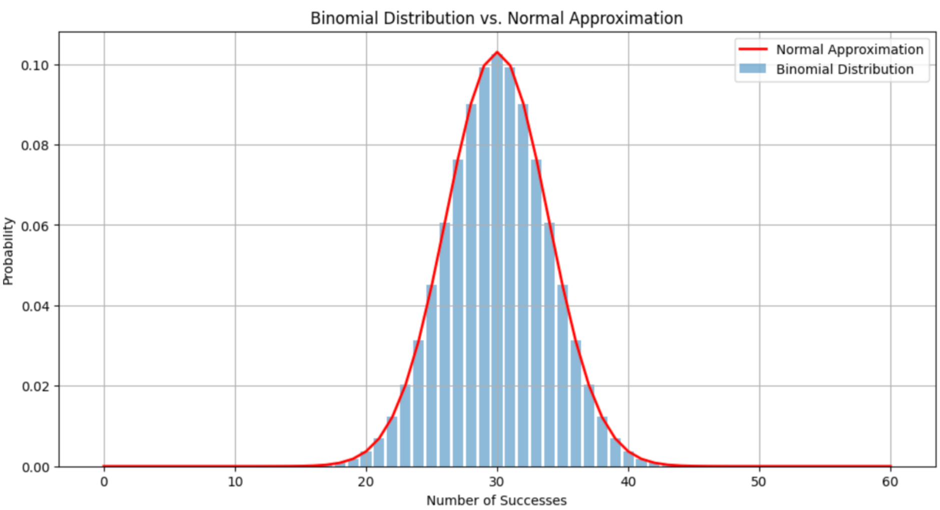
colab.research.google.com/drive/12n0f55MkqSWcvJsaTdimXsvQNrmlw6HH#scrollTo=sn5N8hNVEvax

ath208_Ronish_(19707)_Q.No.3.ipynb

dit View Insert Runtime Tools Help All changes saved

Comment Share

+ Text RAM Disk



Q.No.5) In coin tossing experiments, please find the probability of the exact 6 heads from 12 tossing by ONLY using the normal distribution method.

Solution:

We have,
n = 12 (no. of trials or tosses)
p = 0.5 (probability of getting head)
q= 0.5 (probability of failure(no head))

For a binomial distribution, the mean (μ) and standard deviation (σ) are given by:

Mean(μ) = n * p = 12 * 0.5 = 6
S.D (σ) = $\sqrt{n * p * q} = \sqrt{12 * 0.5 * 0.5} = \sqrt{3}$

The probability of getting exactly 6 heads can be found by finding the area under the normal distribution curve between 5.5 and 6.5 (since we want exactly 6 heads).

To find the probability of getting exactly 6 heads, we need to calculate the area under the normal curve between 5.5 and 6.5.
For exactly 6 heads, we consider the range from 5.5 to 6.5 (to include 6 heads but not 7).
Z-Score for 5.5 (Z1) = $\frac{(x - \mu)}{\sigma} = \frac{(5.5 - 6)}{\sqrt{3}}$ = -0.289

Z-Score for 6.5 (Z2) = $\frac{(x - \mu)}{\sigma} = \frac{(6.5 - 6)}{\sqrt{3}}$ = 0.289

Using Python function , The probabilities P(Z 1) and P(Z 2) are found using the cumulative distribution function (CDF) for a standard normal distribution.

```
from scipy.stats import norm

# Given values for the problem
mean = 6  # Mean (μ) from np = 12 * 0.5
standard_deviation = 3**(0.5)  # Standard Deviation (σ) from √(np(1-p))

# Calculate Z-Scores for 5.5 and 6.5
z1 = (5.5 - mean) / standard_deviation
z2 = (6.5 - mean) / standard_deviation

# Use the Z-Scores to find the probability for each
p_z1 = norm.cdf(z1)
p_z2 = norm.cdf(z2)

# The probability of exactly 6 heads is the difference between P(Z_2) and P(Z_1)
probability_exactly_6_heads = p_z2 - p_z1

p_z1, p_z2, probability_exactly_6_heads
```

```
(0.38641499634222376, 0.6135850036577762, 0.22717000731555248)
```

Here,
P(Z1) ~ 0.386
P(Z2) ~ 0.614 [Using CDF in python]

Now,
Probability of exactly six heads = P(Z2) - P(Z1)
= 0.614 - 0.386
= 0.227

Therefore, the probability of getting exactly 6 heads in 12 coin tosses is 22.7%

Q.No.6) Given that the defective rate of a product of the batteries in a manufacturing company is 6%, 150 batteries are randomly selected from the population. Please find the probability of 12 or more defective ones in them by ONLY using the normal distribution method.

Solution

Given:
n = 150
Probability of a defective battery (p) = 0.06.
Mean (μ) = np = 150 * 0.06 = 9
Standard Deviation (σ) = $\sqrt{n * p * q} = \sqrt{150 * 0.06 * 0.94} = 2.91$

The probability of 12 or more defective batteries can be calculated using the normal approximation to the binomial distribution.

$Z = \frac{(x - \mu)}{\sigma} = \frac{(12 - 9)}{2.91} \approx 1.031$

Now, the probability using the standard normal distribution table for z ≈ 1.031

$P(X \geq 12) \approx P(X > 11.5)$

$P(X > 11.5) \approx 1 - P(Z \leq 1.06)$

Using Python function for the standard normal distribution (Z), we get,

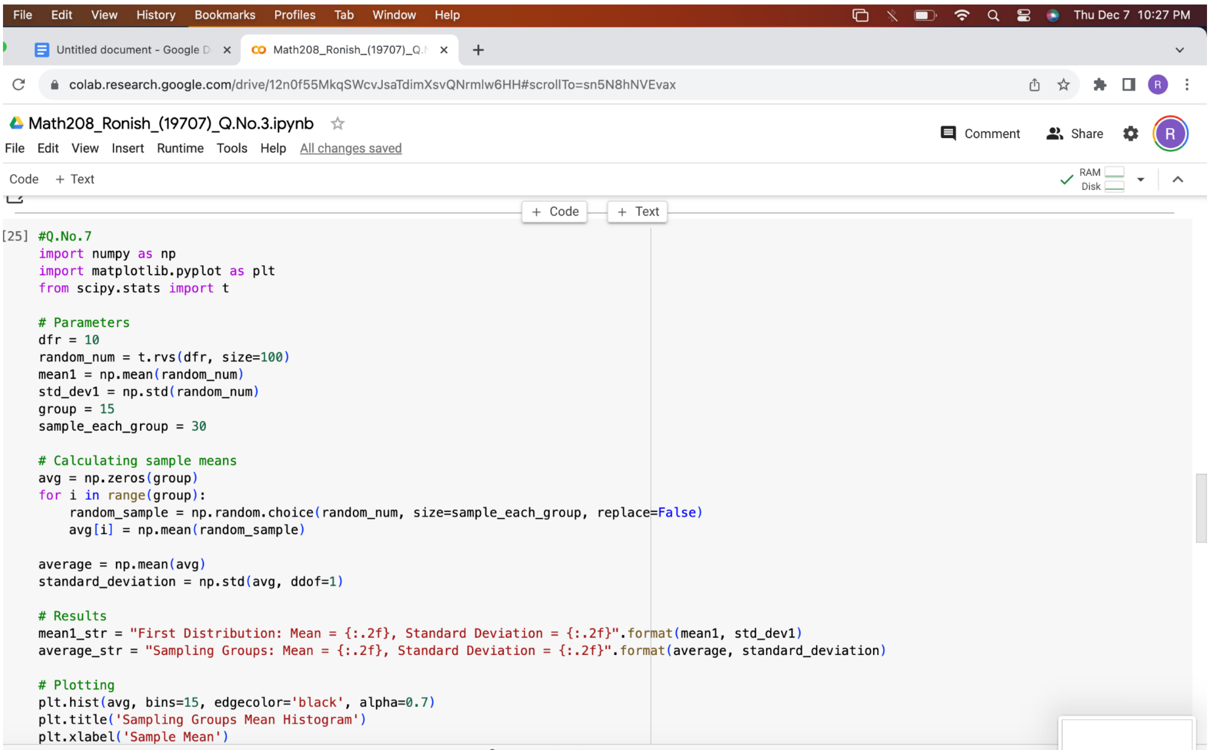
```
import scipy.stats as stats

# Define the value of Z
z = 1.031

# Calculate the probability P(Z ≤ z)
probability = stats.norm.cdf(z)

print(f"P(Z ≤ {z}) = {probability:.4f}")
```

```
P(Z ≤ 1.031) = 0.8487
```

Using standard normal distribution:

P (X ≥ 12) = 1 – P (Z $\leq$ 1.031)
 = 1 – 0.8487
 ~0.1513
Therefore, the probability of 12 or more defective batteries is 0.1513.[i.e. (15.13%)]

Q.No.7) Write a Python program by calling functions in the following link to create 100 random numbers in T distribution with df =10 (degree of freedom) and calculate the mean μ and standard deviation σ. After that, the 30 samples will be randomly selected from these random numbers in each sampling group. A total of 15 sampling groups should be created. Based on Central Limit Theorem (CLT), the mean value x in total 15 sampling group is roughly the mean μ of 100 random numbers and σx = σ/ n . Please verify it and plot the histogram, which should be normal distribution.



File Edit View History Bookmarks Profiles Tab Window Help Thu Dec 7 10:27 PM

Untitled document - Google D × Math208_Ronish_(19707)_Q.N ×

colab.research.google.com/drive/12n0f55MkqSWcvJsaTdimXsvQNrmlw6HH#scrollTo=sn5N8hNVEvax

Math208_Ronish_(19707)_Q.No.3.ipynb

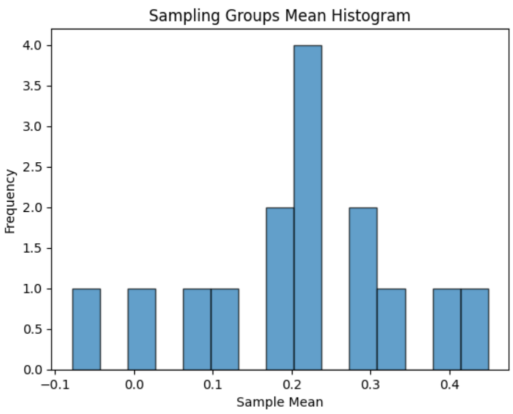
File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk

```
plt.show()

mean1_str, average_str
```




```
('First Distribution: Mean = 0.18, Standard Deviation = 1.04',
 'Sampling Groups: Mean = 0.21, Standard Deviation = 0.14')
```