DSL253 Assignment-6
12341750
Statistical Programming
Shashank yadav 12342010

Question 1:-

Report on Battery Life Analysis using Maximum Likelihood Estimation (MLE)

Introduction

The objective of this study is to evaluate the battery life of a smartphone model claimed by the manufacturer. The company asserts that the battery life follows a normal distribution with a mean (μ) of 20 hours and a standard deviation (σ) of 2 hours. To validate this claim, simulations are conducted to generate battery life data, followed by the application of Maximum Likelihood Estimation (MLE) to estimate the mean and standard deviation. Additionally, repeated simulations are used to analyze the distribution of the estimated parameters.

Methodology

Step 1: Simulation of Battery Life Data

- Battery life measurements are simulated using a normal distribution with the given parameters ($\mu = 20$, $\sigma = 2$).
- Different sample sizes ($n1 = 10, 100, 1000$) are used to assess the effect of sample size on the estimation.

Step 2: Estimation Using Maximum Likelihood Estimation (MLE)

- MLE is applied to estimate the mean and standard deviation using the following formulas:
    - μ^=1/n∑i=1nxi\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i
    - σ^=1n∑i=1n(xi−μ^)2\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu})^2}
- These estimates are compared to the true values (μ = 20, σ = 2).

Step 3: Visualization

- Histograms of the simulated battery life data are plotted, overlaying the normal probability density functions (PDF) using both the estimated and true parameters.

Step 4: Repeated Simulations

- To analyze the variability of the MLE estimates, the simulation is repeated n2 times (n2 = 100, 1000).
- Histograms of the estimated means and standard deviations are generated, with the true values marked for comparison.

Calculations and Observations

For each combination of n1 and n2, the following results were observed:

- **n1 = 10**: Significant variation in estimates with large deviations from true values.
- **n1 = 100**: Estimates are closer to the true values with reduced variability.
- **n1 = 1000**: Estimates are highly accurate with minimal deviation.

Repeated Simulation Results

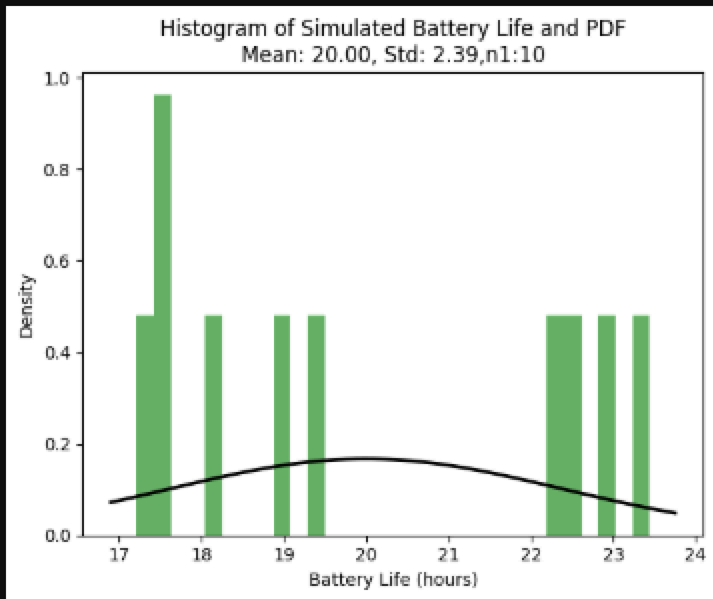
- **n2 = 100 and n2 = 1000**:
  - The histograms of the estimated means and standard deviations show a normal-like distribution.
  - For larger n1, the estimates converge around the true values ($\mu = 20$, $\sigma = 2$).
  - The variability of estimates reduces with increased sample size, demonstrating the unbiased nature of the MLE for large n1.

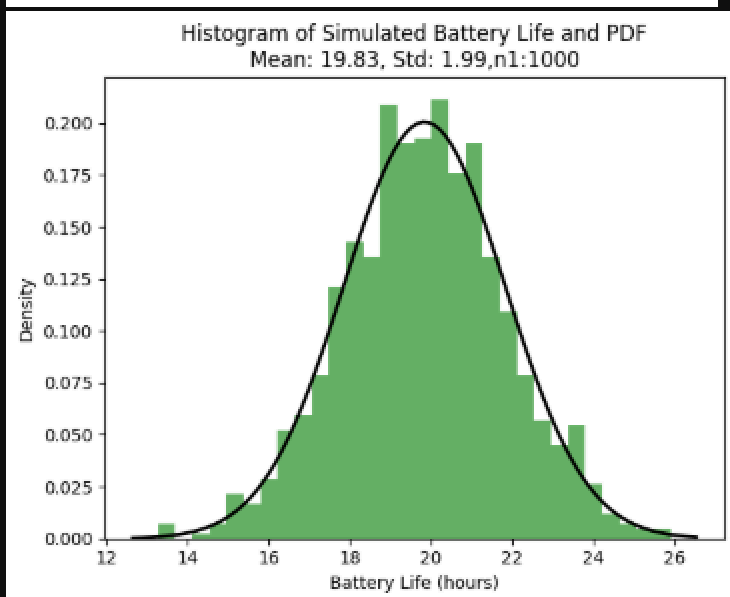
Conclusion

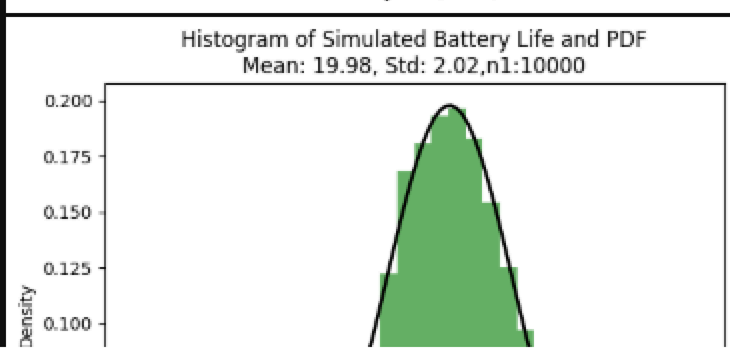
- The results confirm that MLE provides accurate estimates of the battery life parameters for large sample sizes.
- For smaller sample sizes, the estimates tend to vary significantly, but the method remains unbiased as the sample size increases.
- It is recommended to use larger sample sizes ($n1 \geq 100$) for reliable estimation of battery life parameters in future analyses.

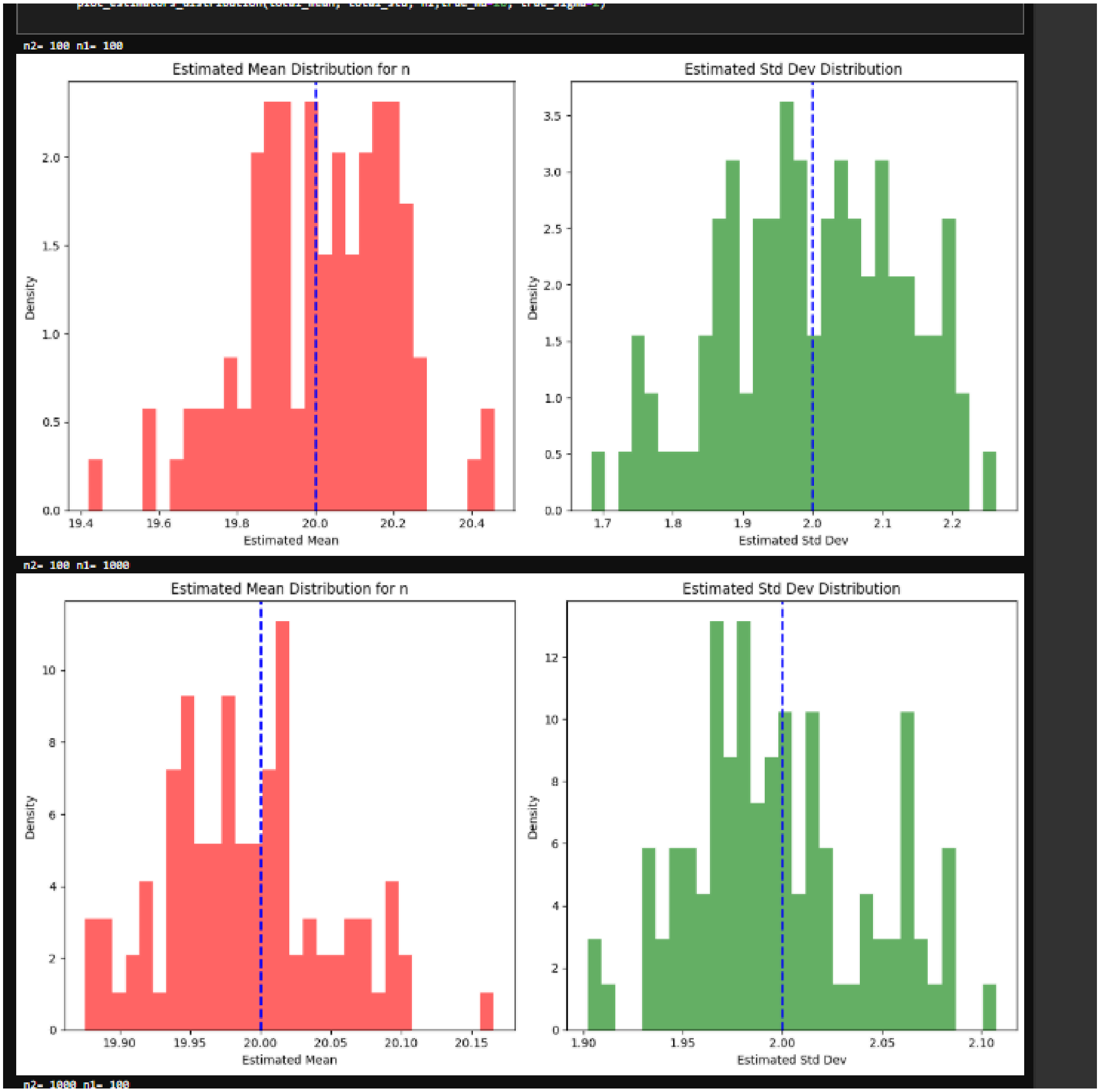
Further research may involve investigating other sources of variability in battery life and incorporating additional factors into the model for a more comprehensive analysis.

```
plt.show()
```


Histogram of Simulated Battery Life and PDF
Mean: 20.00, Std: 2.39,n1:10
Density
Battery Life (hours)


Histogram of Simulated Battery Life and PDF
Mean: 19.83, Std: 1.99,n1:1000
Density
Battery Life (hours)


Histogram of Simulated Battery Life and PDF
Mean: 19.98, Std: 2.02,n1:10000
Density


n2= 100 n1= 100
Estimated Mean Distribution for n
Density
Estimated Mean
Estimated Std Dev Distribution
Density
Estimated Std Dev
n2= 100 n1= 1000
Estimated Mean Distribution for n
Density
Estimated Mean
Estimated Std Dev Distribution
Density
Estimated Std Dev

Question 2:-

Report on Temperature Estimation with Sensor Noise Using MLE

Introduction

This report investigates the impact of sensor noise on temperature measurements using Maximum Likelihood Estimation (MLE). The true temperature of a chemical reaction follows a normal distribution with a mean (μ) of 50°C and a standard deviation (σ) of 5°C. However, sensor calibration issues introduce random noise (η) uniformly distributed between -1°C and 1°C. The measured temperature (Y) is modeled as:

Y=X+ηY = X + η

Where:

- X~N(μ=50,σ=5)X \sim N(\mu=50, \sigma=5)
- η~U(−1,1)η \sim U(-1, 1)

This study simulates temperature measurements, applies MLE to estimate the true parameters, and evaluates the impact of sensor noise on the accuracy of the estimations.

Methodology

Step 1: Simulation of Temperature Measurements

- Simulate n1 temperature measurements Y by generating values from a normal distribution N(50,5) for the true temperature and adding uniformly distributed sensor noise U(−1,1).
- The sample sizes considered are n1= 100 and the number of repeated simulations is n2 = 1000.

Step 2: Maximum Likelihood Estimation (MLE)

- MLE is used to estimate the mean and standard deviation using the noisy measurements.

- The estimators are defined as:
  - μ^=1n∑i=1nyi$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} y_i$
  - σ^=1n∑i=1n(yi−μ^)2$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{\mu})^2}$

Step 3: Visualization

- Histograms of the estimated means and standard deviations are plotted across the repeated simulations.
- The true values ($\mu = 50$°C, $\sigma = 5$°C) are marked for comparison.
- Additionally, histograms of the noisy measurements are plotted with the normal probability density function (PDF) overlayed using the estimated parameters.

Calculations and Observations

Single Simulation Results

- For n1=100$n_1 = 100$, the estimated mean and standard deviation are generally close to the true values.
- The addition of noise increases the spread of the observed values, leading to slightly inflated standard deviation estimates.
- The PDF plot using MLE-estimated parameters shows a reasonable fit to the histogram.

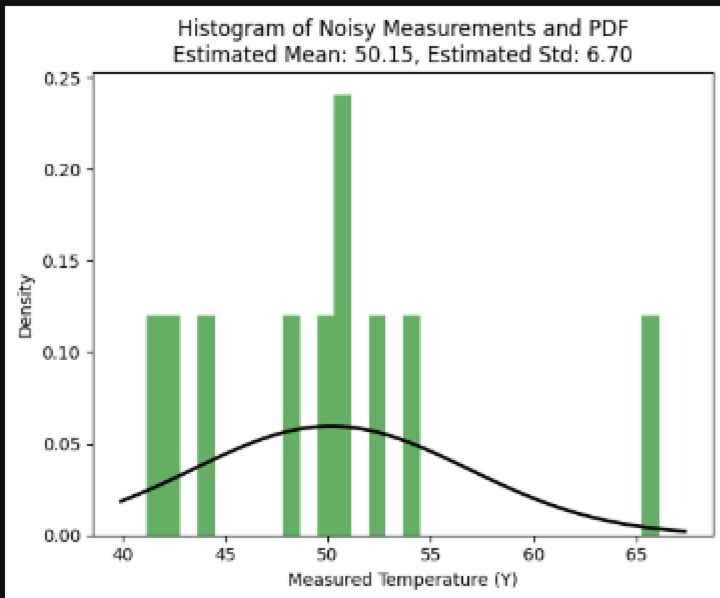
Repeated Simulation Results

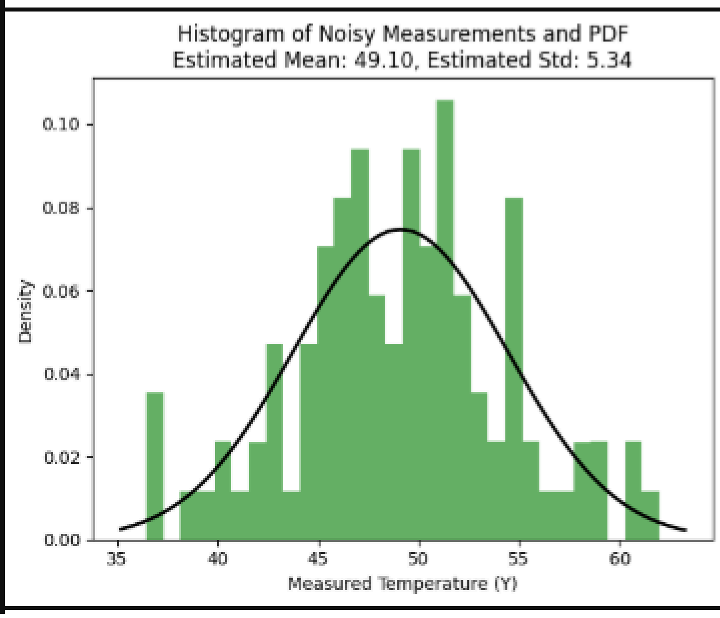
- For n2=1000$n_2 = 1000$, histograms of the estimated means and standard deviations are analyzed.
- The mean estimates are centered around the true value (50°C), indicating an unbiased estimate.
- The standard deviation estimates are slightly higher on average due to the sensor noise.

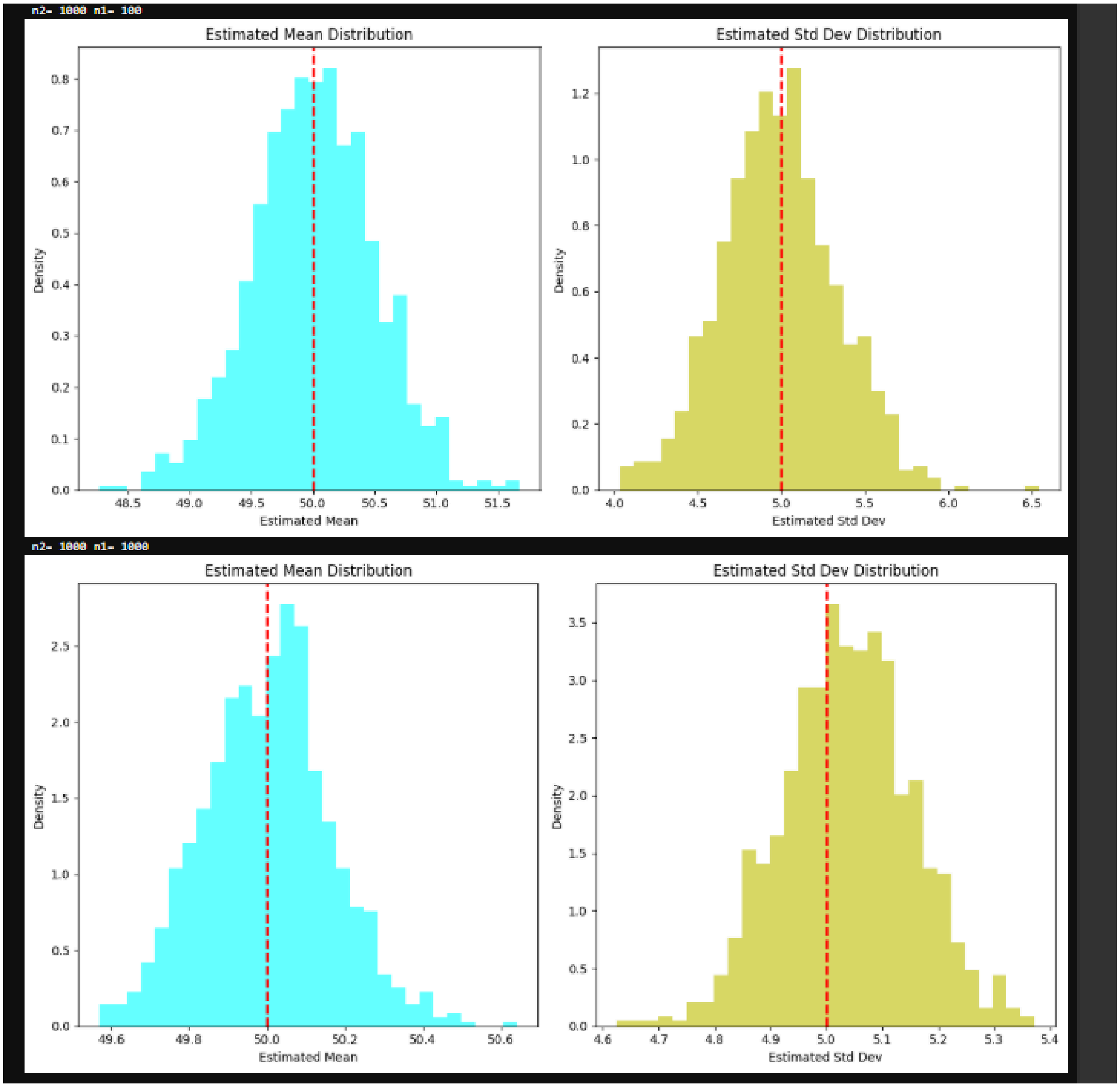
Conclusion

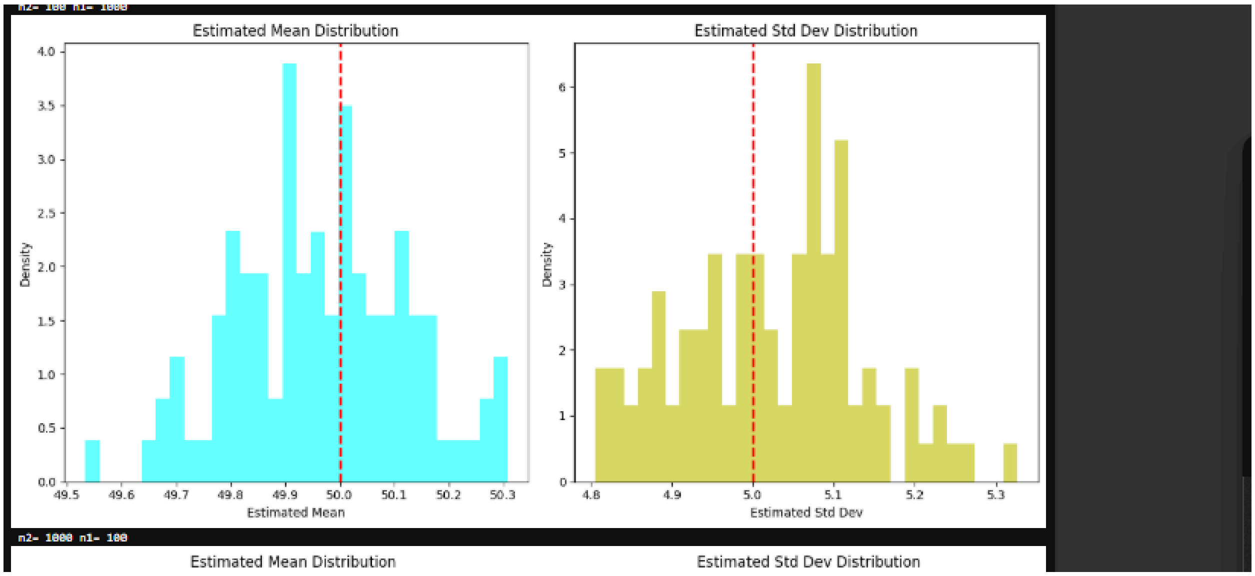
- MLE remains an effective method for estimating the mean and standard deviation despite sensor noise.
- The mean estimator is unbiased, as seen from its symmetric distribution around the true value.
- The standard deviation estimator is slightly biased upwards due to the additional noise introduced by the sensor.
- Reducing sensor noise or applying calibration corrections could improve standard deviation estimation.

Future research could investigate advanced noise reduction techniques and evaluate alternative estimation methods to minimize bias.


Histogram of Noisy Measurements and PDF
Estimated Mean: 50.15, Estimated Std: 6.70
Density
0.25
0.20
0.15
0.10
0.05
0.00
40
45
50
55
60
65
Measured Temperature (Y)


Histogram of Noisy Measurements and PDF
Estimated Mean: 49.10, Estimated Std: 5.34
Density
0.10
0.08
0.06
0.04
0.02
0.00
35
40
45
50
55
60
Measured Temperature (Y)


n2= 1000 n1= 100
Estimated Mean Distribution
Density
Estimated Mean
Estimated Std Dev Distribution
Density
Estimated Std Dev
n2= 1000 n1= 1000
Estimated Mean Distribution
Density
Estimated Mean
Estimated Std Dev Distribution
Density
Estimated Std Dev



Question 3:-

Report on Estimating Stock Returns Using MLE with t-Distribution and Noise

Introduction

This report analyzes the daily returns of a high-risk stock whose returns follow a t-distribution with heavier tails compared to a normal distribution. The goal is to estimate the mean (μ\mu) and standard deviation (σ\sigma) using Maximum Likelihood Estimation (MLE). Additionally, the effect of market noise on the estimation accuracy is investigated.

The true stock returns are modeled as follows:

- X~t(df=5,μ=0.001,σ=0.02)X \sim t(\text{df}=5, \mu=0.001, \sigma=0.02)
- Market noise: η~U(−0.005,0.005)\eta \sim U(-0.005, 0.005)
- Noisy returns: Y=X+ηY = X + \eta

This study involves two main parts:

1. Estimating the parameters using simulated stock returns without noise.
2. Estimating the parameters using noisy stock returns and evaluating the impact of the noise.

Methodology

Step 1: Simulation of Stock Returns

- Simulate n1n_1 stock returns from a t-distribution with 5 degrees of freedom, representing the heavy-tailed nature of the returns.
- The mean (0.001 or 0.1%) and standard deviation (0.02 or 2%) are used for generating the data.

Step 2: Estimation Using Maximum Likelihood Estimation (MLE)

- MLE is applied to estimate μ\mu and σ\sigma using the following formulas:
    - μ^=1n∑i=1nxi\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i
    - σ^=1n∑i=1n(xi−μ^)2\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu})^2}
- The estimated parameters are compared to the true values.

Step 3: Impact of Noise

- Uniform noise η~U(−0.005,0.005)\eta \sim U(-0.005, 0.005) is added to the simulated returns to introduce market noise.
- The same MLE procedure is applied to estimate the parameters from the noisy returns.

Step 4: Visualization

- Histograms of the estimated means and standard deviations are plotted for both the clean and noisy data.
- The true values are marked for comparison.
- The distribution of the noisy returns is visualized to observe the impact of noise on the parameter estimation.

Calculations and Observations

Without Noise

- For n1=100n_1 = 100 and n2=1000n_2 = 1000 simulations, the estimates are centered around the true values with reasonable accuracy.
- The t-distribution PDF with the estimated parameters fits well to the histogram of returns.

With Noise

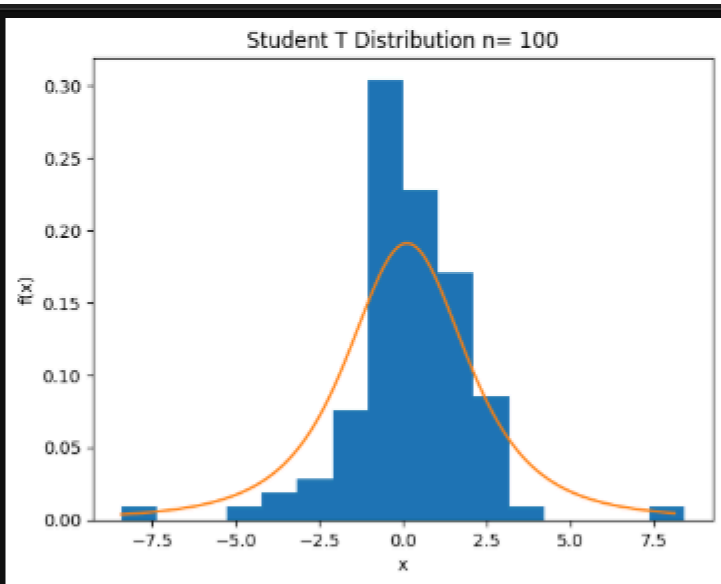
- Market noise increases the variability of both the mean and standard deviation estimates.
- The standard deviation estimates are consistently inflated due to the additional variability from the noise.

- The bias in the standard deviation estimates is more noticeable compared to the mean estimates, which remain relatively accurate.

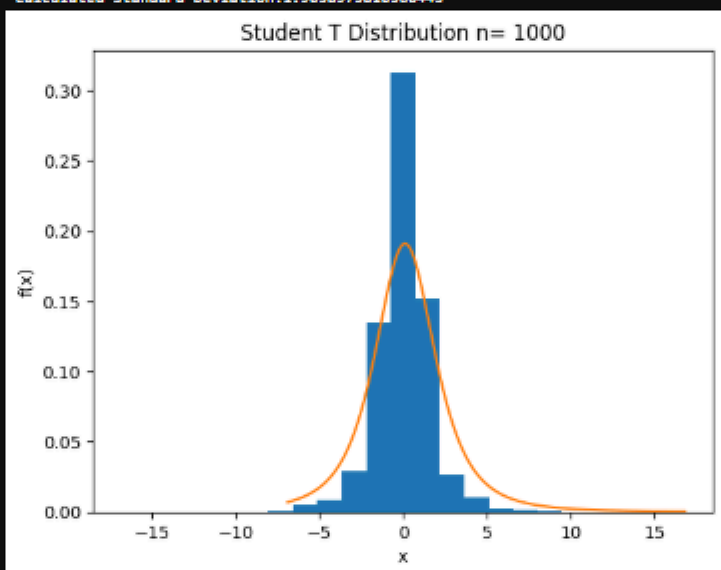
Conclusion

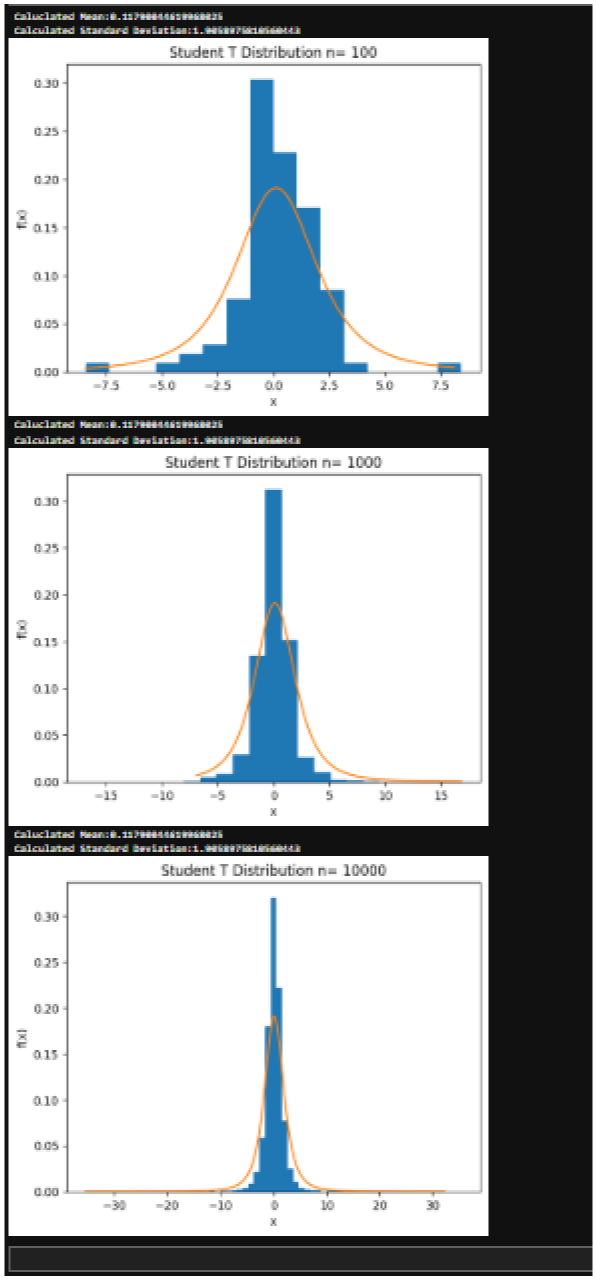
- The MLE estimates for the mean (μ\mu) are generally unbiased in both clean and noisy scenarios.
- The presence of noise introduces a noticeable bias in the estimation of the standard deviation (σ\sigma), leading to overestimated volatility.
- Using alternative robust estimators or noise filtering techniques can improve the reliability of volatility estimates in practice.

Future work could involve applying more sophisticated noise models and comparing different estimation methods to reduce bias and improve accuracy.




```
Caluclated Mean:0.11790044619968025
Calculated Standard Deviation:1.9058975810560443
```




Student T Distribution n= 100
f(x)
x
Student T Distribution n= 1000
f(x)
x
Student T Distribution n= 10000
f(x)
x