**E**stimation **A**nd **C**onfidence **I**ntervals

MULAGAPATI PADMAVATHI

DATA SCIENCE WEEKDAY BATCH

**Background**

In quality control processes, especially when dealing with high-value items, destructive sampling is a necessary but costly method to ensure product quality. The test to determine whether an item meets the quality standards destroys the item, leading to the requirement of small sample sizes due to cost constraints.

**Scenario**

A manufacturer of print-heads for personal computers is interested in estimating the mean durability of their print-heads in terms of the number of characters printed before failure. To assess this, the manufacturer conducts a study on a small sample of print-heads due to the destructive nature of the testing process.

**Data**

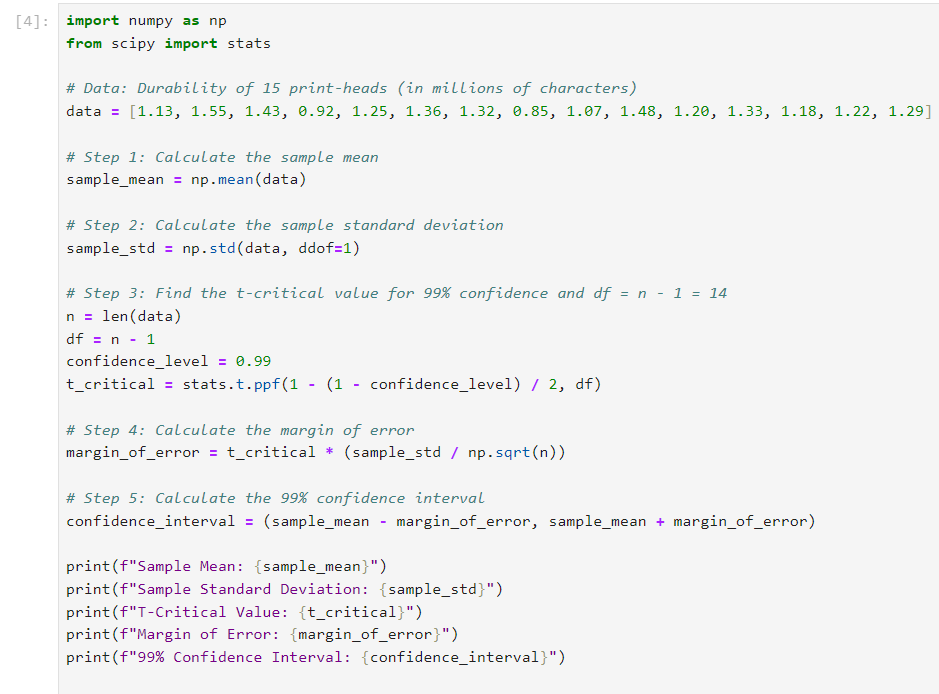
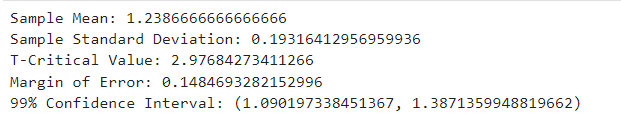
A total of 15 print-heads were randomly selected and tested until failure. The durability of each print-head (in millions of characters) was recorded as follows:

1.13, 1.55, 1.43, 0.92, 1.25, 1.36, 1.32, 0.85, 1.07, 1.48, 1.20, 1.33, 1.18, 1.22, 1.29

**Asignment Tasks**

1. **Build 99% Confidence Interval Using Sample Standard Deviation**

Assuming the sample is representative of the population, construct a 99% confidence interval for the mean number of characters printed before the print-head fails using the sample standard deviation. Explain the steps you take and the rationale behind using the t-distribution for this task.



**Steps for Constructing a 99% Confidence Interval**

1. **Calculate the Sample Mean**:
   * The first step is to compute the **sample mean**, which gives us the central tendency of the data. The sample mean represents the average durability (in millions of characters) of the print-heads in the sample.
2. **Calculate the Sample Standard Deviation**:
   * Next, we calculate the **sample standard deviation**, which measures the spread or dispersion of the durability values around the mean. Since we do not have the population standard deviation, we use the sample standard deviation.
3. **Determine the T-Critical Value (tα/2t)**:
   * To construct the confidence interval, we use the **t-distribution** because the population standard deviation is unknown and the sample size is small (n<30n < 30n<30). The **t-distribution** accounts for the increased uncertainty when estimating the population mean from a small sample.
   * The **t-critical value** is determined using the confidence level (99%) and the degrees of freedom (df=n−1=14).
4. **Calculate the Margin of Error**:
   * The margin of error quantifies the uncertainty in our estimate of the population mean. It is calculated using the formula:

Margin of Error=tα/2×(s/sqrt(n)).

* + This gives us the range of values on either side of the sample mean within which the true population mean is likely to fall with 99% confidence.

1. **Construct the Confidence Interval**:
   * Finally, the confidence interval is constructed as:

xˉ±Margin of Error

* + This interval gives us a range of values that we are 99% confident contains the true mean durability of the print-heads.

**b. Build 99% Confidence Interval Using Known Population Standard Deviation**

If it were known that the population standard deviation is 0.2 million characters, construct a 99% confidence interval for the mean number of characters printed before failure.