**Estimation and Confidence Intervals**

**a) 99% Confidence Interval using Sample Standard Deviation (t-distribution)**

* **Sample size (n): 15**
* **Sample mean (xˉ\bar{x}xˉ): 1.239 million characters**
* **Sample standard deviation (s): 0.193 million characters**
* **Why t-distribution?  
  Since the population standard deviation is not known and the sample size is small (n<30n < 30n<30), the appropriate distribution for constructing the confidence interval is the *t-distribution* with n−1=14n-1 = 14n−1=14 degrees of freedom.**
* **Critical value: t0.005,14≈2.9768t\_{0.005,14} \approx 2.9768t0.005,14​≈2.9768**
* **Margin of error:**

**E=t×sn=2.9768×0.19315≈0.149E = t \times \frac{s}{\sqrt{n}} = 2.9768 \times \frac{0.193}{\sqrt{15}} \approx 0.149E=t×n​s​=2.9768×15​0.193​≈0.149**

* **99% CI:**

**(1.090,  1.387)(1.090, \; 1.387)(1.090,1.387)**

**Interpretation: With 99% confidence, the true mean durability of print-heads lies between 1.09 and 1.39 million characters.**

**b) 99% Confidence Interval using Known Population Standard Deviation (z-distribution)**

* **Population standard deviation (σ\sigmaσ): 0.2 million characters**
* **Why z-distribution?  
  When σ\sigmaσ is known, the sampling distribution of the mean follows the *standard normal distribution* (z).**
* **Critical value: z0.005≈2.576z\_{0.005} \approx 2.576z0.005​≈2.576**
* **Margin of error:**

**E=z×σn=2.576×0.215≈0.133E = z \times \frac{\sigma}{\sqrt{n}} = 2.576 \times \frac{0.2}{\sqrt{15}} \approx 0.133E=z×n​σ​=2.576×15​0.2​≈0.133**

* **99% CI:**

**(1.106,  1.372)(1.106, \; 1.372)(1.106,1.372)**

**Interpretation: With 99% confidence, the true mean durability lies between 1.11 and 1.37 million characters.**

**Key Insight**

**Both approaches give very similar confidence intervals. The t-based interval is slightly wider because it accounts for the uncertainty of estimating the standard deviation from a small sample. When the population standard deviation is known, the z-based interval is slightly narrower.**

**Code used:**

**import numpy as np**

**data = np.array([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])**

**n = len(data)**

**mean = data.mean()**

**s = data.std(ddof=1)**

**alpha = 0.01**

**# critical values**

**t\_crit\_14 = 2.9768   # t\_{0.005, df=14}**

**z\_crit = 2.5758      # z\_{0.005}**

**me\_t = t\_crit\_14 \* (s / np.sqrt(n))**

**ci\_t = (mean - me\_t, mean + me\_t)**

**pop\_sigma = 0.2**

**me\_z = z\_crit \* (pop\_sigma / np.sqrt(n))**

**ci\_z = (mean - me\_z, mean + me\_z)**

**print(f"n = {n}")**

**print(f"sample mean = {mean:.6f}")**

**print(f"sample sd (s) = {s:.6f}")**

**print()**

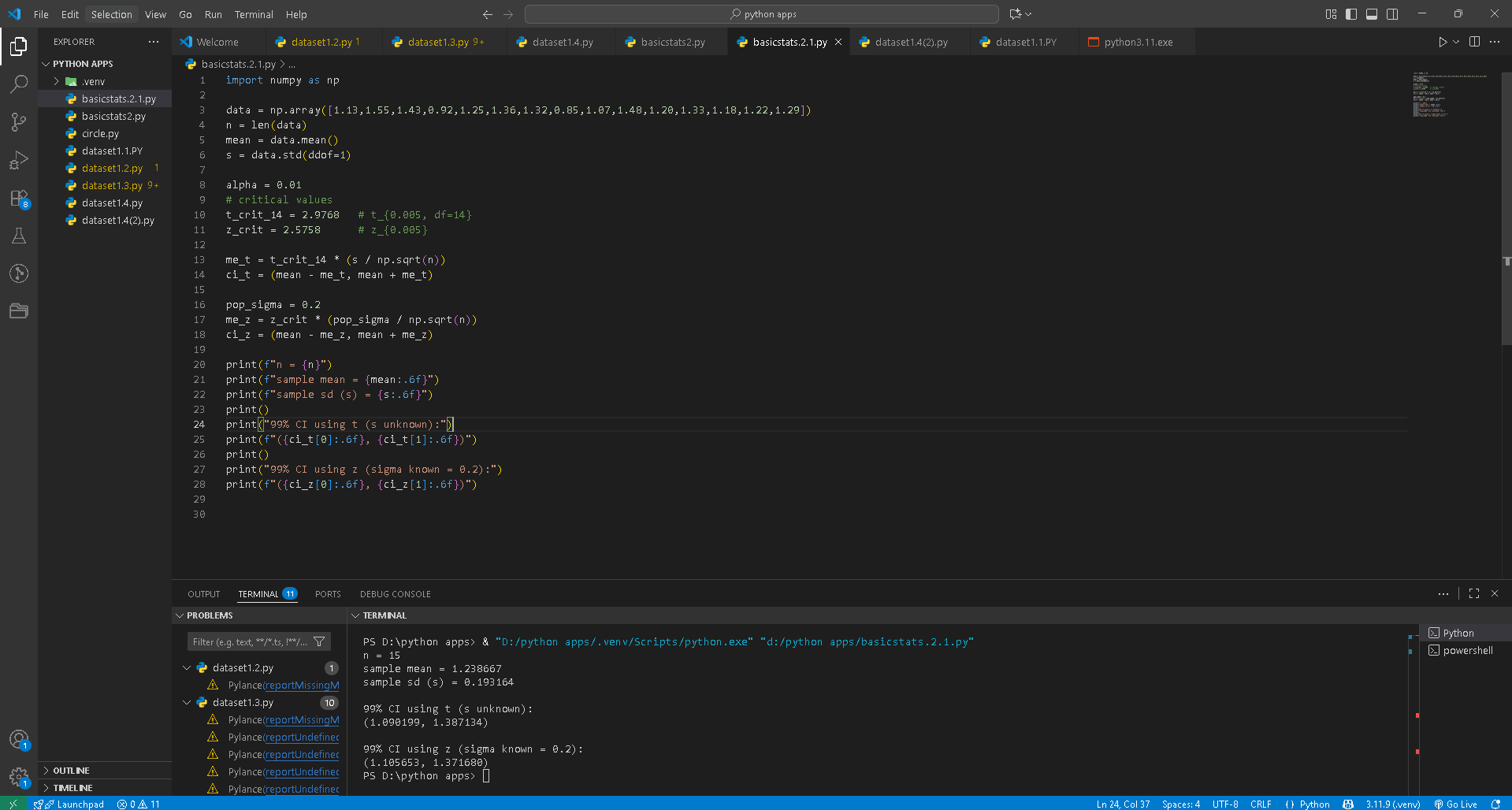
**print("99% CI using t (s unknown):")**

**print(f"({ci\_t[0]:.6f}, {ci\_t[1]:.6f})")**

**print()**

**print("99% CI using z (sigma known = 0.2):")**

**print(f"({ci\_z[0]:.6f}, {ci\_z[1]:.6f})")**

****