

Interpretation of Section 1, Part A and Section 1, Part B

Section-1, Part A: Natural Variation in Completion Times

In this section, the data were clean, meaning there were no extreme or unusual values. We compared three methods for estimating the average completion time: the mean, the median, and the trimmed mean (which calculates the average after removing the top and bottom 10% of values).

The findings showed that the **mean** provided the most accurate estimate of the true average. It had almost zero bias and showed very little variation between different samples. The **median** and **trimmed mean** produced slightly lower estimates and were a bit less accurate, but still consistent overall.

Therefore, when the data are normal and free from outliers, the **mean** is the most effective and reliable choice.

Section-1, Part B: Presence of Outliers or Errors

In this section, we introduced some extreme values into the data to represent possible errors or unusually long completion times.

These large outliers caused the **mean** to fluctuate noticeably between samples. The **median** and **trimmed mean**, on the other hand, remained stable and were not heavily influenced by these extreme values.

Even though the mean still reflected the overall contaminated average, it became less dependable because of its high variability. The **median** and **trimmed mean** gave smaller estimates compared to the contaminated mean, but they were much steadier and resistant to outliers.

Overall Summary

When the data are clean and consistent, the **mean** is the best option. However, when the data contain outliers or recording errors, the **median** or **trimmed mean** perform better since they provide more stable and reliable estimates.

In real-life training time data, where a few individuals may take significantly longer due to unusual circumstances or errors, using a **trimmed mean** is generally the most sensible and dependable approach.