Sharvari Lahane

Data Science - Batch May 2024 (Baner, Pune) - Assignment 2

Task a:

Build 99% Confidence Interval Using Sample Standard Deviation

To construct a 99% confidence interval for the mean number of characters printed before the print-head fails using the sample standard deviation.

Step 1:
Calculate the sample mean ($\bar{x}$) and sample standard deviation (s) from the given data.

Using the data, we calculate:

$\bar{x}$ = (1.13 + 1.55 + ... + 1.29) / 15

$\bar{x}$ = 1.24 million characters

s = $\sqrt{}$[(Σ(x - $\bar{x}$)^2) / (n - 1)]

s= 0.193 million characters

where n is the sample size (15).

Step 2:
Determine the degrees of freedom (df) for the t-distribution.

Since we're using the sample standard deviation,

we'll use the t-distribution with:-

df = n - 1

df= 15 - 1

df= 14

Step 3:

Look up the critical t-value ($t\alpha/2$) for a 99% confidence interval with 14 degrees of freedom.

Using a t-distribution table or calculator, we find:

$t\alpha/2 = t(0.005, 14)$

$t\alpha/2 \approx 2.977$

Step 4:

Calculate the margin of error (E) using the sample standard deviation and critical t-value.

$E = t\alpha/2 * (s / \sqrt{n})$

$E \approx 2.977 * (0.193 / \sqrt{15})$

$E \approx 0.144$ million characters

Step 5:

Construct the 99% confidence interval for the population mean (μ).

The confidence interval is:

$\bar{x} - E \leq \mu \leq \bar{x} + E$ 1.24 - 0.144 ≤ μ ≤ 1.24 + 0.144 1.096 million characters ≤ μ ≤ 1.384 million characters

Rationale for using the t-distribution:

Since the sample size is small (15), we can't assume normality of the population distribution.

The t-distribution is a more conservative and robust choice for small sample sizes, as it takes into account the uncertainty in the sample standard deviation.

The t-distribution is used to construct confidence intervals when the population standard deviation is unknown, which is the case here.

Task b:

Build 99% Confidence Interval Using Known Population Standard Deviation

If the population standard deviation (σ) is known to be 0.2 million characters, we can construct a 99% confidence interval for the mean number of characters printed before failure using the steps below:

Step 1:

Calculate the sample mean ($\bar{x}$) from the given data.

$\bar{x}$ = 1.24 million characters

Step 2:

Determine the critical z-value ($z\alpha/2$) for a 99% confidence interval.

Using a standard normal distribution table or calculator, we find:

$z\alpha/2 = z(0.005)$

$z\alpha/2 \approx 2.576$

Step 3:

Calculate the margin of error (E) using the known population standard deviation and critical z-value.

$E = z\alpha/2 * (\sigma / \sqrt{n})$

$E \approx 2.576 * (0.2 / \sqrt{15})$

E ≈ 0.104 million characters

Step 4:

Construct the 99% confidence interval for the population mean (μ).

The confidence interval is:

$\bar{x}$ - E ≤ μ ≤ $\bar{x}$ + E 1.24 - 0.104 ≤ μ ≤ 1.24 + 0.104 1.136 million characters ≤ μ ≤ 1.344 million characters