Semester V Subject Statistics for AI&DS Academic Year 2023 2024

## CONFIDENCE INTERVALS:

* In a frequency-statistics, a confidence interval(CI) is a range of estimates for an unknown parameter.
* A confidence interval is computed at a designated confidence level, the 95% confidence level is most common, but other levels, such as 90% or 99% are sometimes used.
* One way to think of a 90% confidence interval is as follows: It is the interval that encloses the central 90% of the bootstrap sampling distribution of a sample statistics.
* A large sample would produce a narrower confidence level.
* The greater variability in the sample produces a wider confidence interval, and a higher confidence level would demand a wider confidence interval.

Given a sample size n, and a sample statistic of interest, the algorithm for a bootstrap confidence interval is as follows:

(1) Draw a random sample of size n with replacement from the data.

(2) Record the statistics of interest for the resample.



Scanned by TapScanner

Semester V Subject Statistics for AI&DS Academic Year 2023 2024

(3) Repeat the steps 1-2 many times.

(4) Calculate the confidence interval using the formula $CI = \bar{x} \pm z \frac{s}{\sqrt{n}}$

where, CI → Confidence Interval
$\bar{x}$ → Sample Mean
z → Confidence level value
s → Sample standard Deviation
n → sample size.

| z-value |
|---|
| 80% - 1.282 |
| 85% - 1.440 |
| 90% - 1.645 |
| 95% - 1.960 |
| 99% - 2.576 |
| 99.5% - 2.807 |
| 99.9% - 3.291 |

(5) The trim points are the endpoints of a x% bootstrap confidence interval.

Example:
Calculate the range of heights (95% confidence level) for the given population. The mean value $\bar{x} = 175cm$, SD = 20cm, Sample size (n) = 40.

Solution:

$\bar{x} = 175cm$

$SD = 20cm$

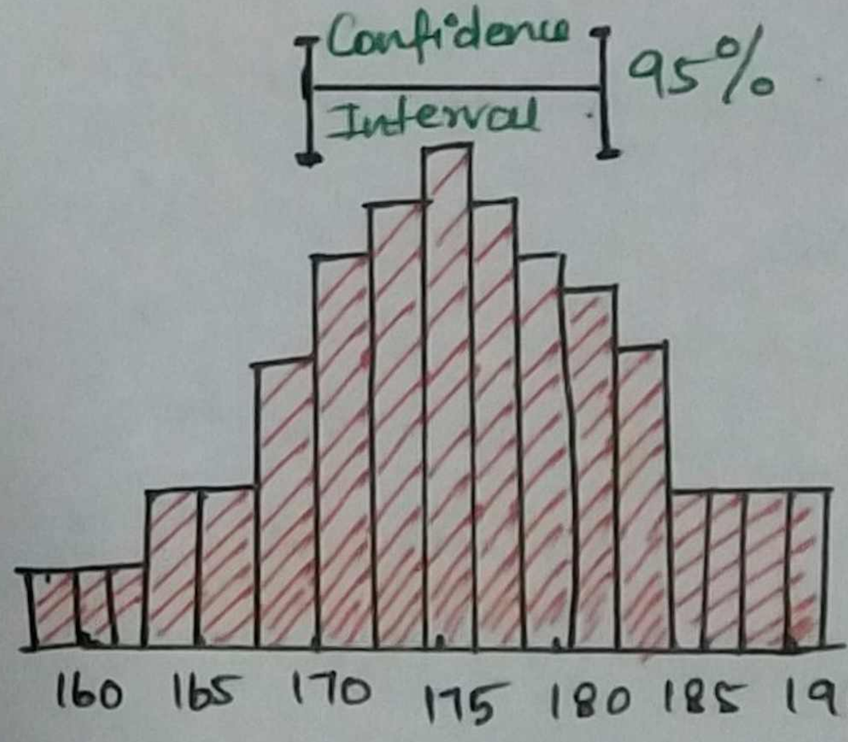
$n = 40$.

$$CI = 175 \pm 1.960 \times \frac{20}{\sqrt{40}}$$

$= 175 \pm 6.20cm$.

$(175 - 6.20, 175 + 6.20)$

168.8cm to 181.2cm → 95% confidence Interval.






Scanned by TapScanner