U3A - Maths

24-Jun-2022 Term 2 / Week 9

Central Limit Theorem - Examples (CLT)

- Central Limit Theorem enables analysis of population with unknown distribution.
- It is done by considering "Sample Mean" ($\bar{x}$) rather than individual values
- As per CLT, Sample Means are normally distributed with:

  Mean ($\mu_{\bar{x}}$) = μ

  Std Dev ($\sigma_{\bar{x}}$) = $\sigma / \sqrt{n}$

where μ – Population Mean

σ – Population Std. Dev.

n – Sample size.

- For Non-normal distribution, the sample size (n) is ≥ 30. (Greater the sample => Better the approx)
- CLT can also be used for population with normal distribution! In such a case, there is no restriction on sample size!

- Lastly, note that we can always calculate the population Mean (μ) and population Std. Dev (σ) for any given population.
- Quite often population μ & σ are obtained from historical data.

Ex-1 Mobile phone bills for residents in a city have a mean (μ) of \$63 and a Std. Dev (σ) of \$10. The data distribution is unknown. Calculate and plot the distribution of sample Means for following sample sizes.

(a) $n = 100$ & (b) $n = 36$

(a) We have $\mu = 63$ & $\sigma = 10$

$\therefore$ for sample means,

Mean of Sample Means $\mu_{\bar{x}} = \mu = 63$

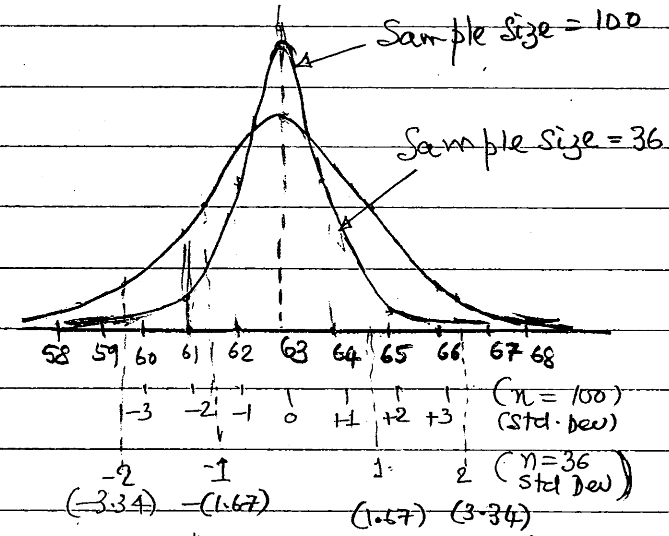
Std. Dev. of sample Means $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{10}{\sqrt{100}} = 1$

(Also called Standard Error of Means)

(b) Similarly, for $n = 36$

Mean of Means: $\mu_{\bar{x}} = \mu = 63$

Std. Dev of Means $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{10}{\sqrt{36}} = 1.67$



- Note that larger sample size gives smaller std. dev (less Std. Error!)

- In the above Example, a smaller sample size has its own its own advantages! It gives better information on the variability of the population data!

- A large sample size provides a more accurate population Mean, but hides the variability of the data!!

Ex. 2

In a regional town, it is given that the time spent driving has a mean (μ) of 25 min and a std. Dev (σ) of 1.5 min The distribution of data is unknown.
If a sample of 50 drivers are randomly selected, what is the probability that the mean driving time is between 24.7 and 25.5 min.

We have:

Population Mean (μ) = 25 min
Population Std. Dev (σ) = 1.5 min

Sample size (n) = 50
For sample Means ($\bar{x}$):

$$\mu_{\bar{x}} = 25 \text{ min}$$

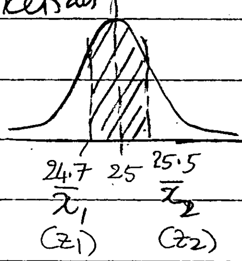
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.5}{\sqrt{50}}$$

$$= 0.212 \text{ min.}$$

Given values for calculation

$\bar{x}_1 = 24.7$ min

$\bar{x}_2 = 25.5$ min



$$\therefore Z_1 = \frac{\bar{x}_1 - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

$$= \frac{24.7 - 25}{0.212} = -1.415$$

Similarly $Z_2 = \frac{25.5 - 25}{0.212} = +2.358$

$$P[\bar{x}_1 : \bar{x}_2] = P[\bar{x}_2] - P[\bar{x}_1]$$

$$= 0.9909 - 0.0786$$

$$= 0.9123 \quad [91.2\%]$$

Home work

The mean boarding & room expenses per year at four-year colleges is \$7540. If a 9 four-year colleges are randomly selected, what is the probability that The mean board & room expenses is less than \$7800. It is also given that board & room expenses are normally distributed with a std. dev. of \$1245.