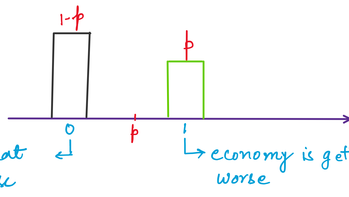
Since we have understood probability, random variable, central limit theorem, we can now solve the problem stated at the starting.

Problem: Assume that 40% population in India says that the economy is getting worse. If we take a sample of 250 people, how likely is that 50% or more of them will say they think the economy is getting worse. Calculate 99% confidence interval for proportion of the population who felt that the economy is getting worse.

Solution: We will not be able to survey the entire population to answer this question. But the entire population can be put in two buckets as follows



This is a binomial distribution.
Let we derive a sample of 250 people, in which 108 say the economy is getting worse.

We have 250 samples with in a sample

The sample mean $\bar{x} = \frac{1(108) + 0(142)}{250}$

$$\bar{x} = \frac{108}{250} = 0.432$$

Sample variance $\sigma_1^2 = \frac{108(1-0.432)^2 + 142(0-0.432)^2}{250-1}$

$$\sigma_1^2 = \frac{34.8434 + 26.5}{249} = \frac{61.344}{249} = 0.246$$

$$\sigma_1 = \sqrt{0.246} = 0.4963$$

Recall that the variance of the samples

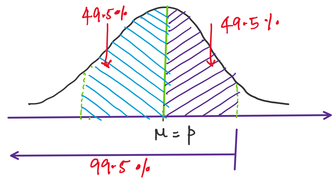
$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{\sigma^2}{250}$$

We do not know σ, so

$$\sigma_{\bar{x}} \approx \frac{\sigma_1}{\sqrt{250}} = \frac{0.4963}{\sqrt{250}} = 0.0314$$

We want a 99% confidence interval, which communicate how accurate our estimate is likely to be.

Using CLT, we can say that $\bar{x}$ is coming from a Normal distribution with mean $\mu = \bar{x}$ and $\sigma_1 = \sigma/\sqrt{n}$.



99% confidence interval means, how many standard deviation away from the mean we have to be so that we are 99% confident that any sample from this sampling distribution will be in that interval.

from the z-table, $0.5 + 0.495 = 0.995$ corresponds to $z = 2.58$. Thus, $2.58\,\sigma_{\bar{x}}$ away from the mean μ would give us the 99% confidence interval.

In other words, 99% chance that a random $\bar{x}$ is with $2.58\,\sigma_{\bar{x}}$ of $\mu = p$.

We are 99% confidence that p is with in
$2.58\,\sigma_{\bar{x}} = 2.58 \times 0.0314 = 0.0810$ of 0.432
i.e. 0.432 ± 0.08

upper $0.432 + 0.08 = 0.513$
$0.432 - 0.08 = 0.352$

We are 99% confident that true population proportion is within the range 35.2% to 51.3%.
or
the true % age of the people who think that the economy is getting worse is in the range 35.2% to 51.3%. There is 99% chance of this.