Learning Outcome

By the end of this lecture, you will be able to:

- Define the concept of sampling error
- Identify how to determine the mean and standard deviation for the sampling distribution of the sample mean

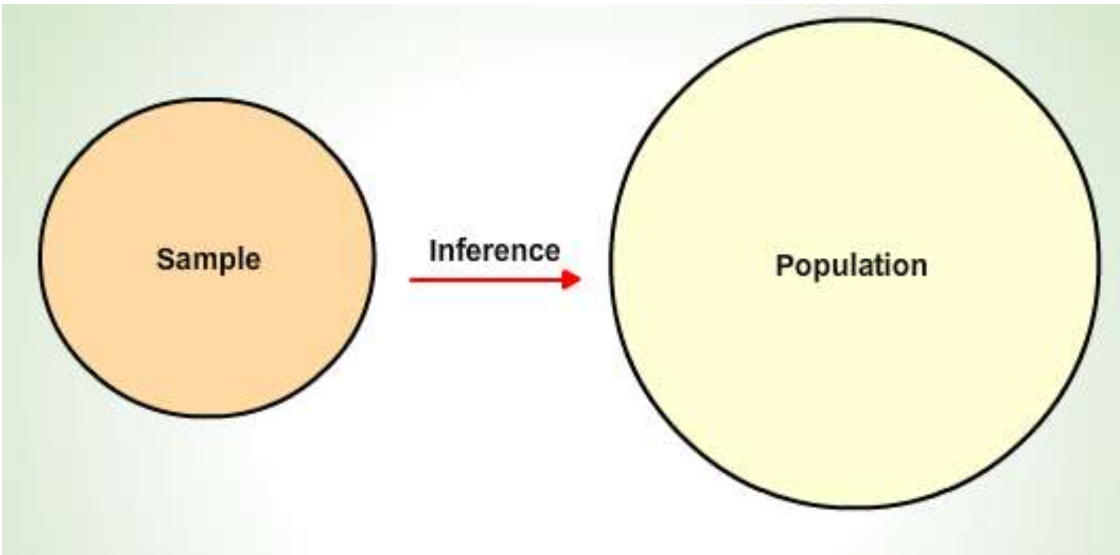
Introduction

Suppose, a researcher measured the heights of 100 men living in the city of Riyadh and found all of them are more than six feet in height. He concluded that all the men in the city are taller than six feet!

City of Riyadh

Will you agree with this conclusion?

Obviously not! The selected group of men may consist only the men, who are taller than six feet in height. This sample would be a highly unreliable sample and this sampling has an error.



City of Riyadh

Clearly, a part of the population can give the true picture of the population, provided sampling error is controlled properly. In this lecture, let's learn more about the sampling errors.

Sample Statistics

Sample statistics are used to estimate population parameters:

- The mean: $\bar{x} = \frac{\sum x}{n}$ is an estimate of the population mean (μ)
- The standard deviation: $S = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$ is an estimate of the population standard deviation $(\ \sigma\ )$
- The sample proportion

It is an estimate of the population proportion (P).

But different samples provide different estimates of the population parameter. So, sample results have potential variability, thus sampling error exist.

Sampling Error

Sampling error is the difference between a value (a statistic) computed from a sample and the corresponding value (a parameter) computed from a population.

$$\text{Sampling error for the mean} = \bar{x} - \mu \qquad (1)$$

Where:

$\bar{x}$ = Sample mean

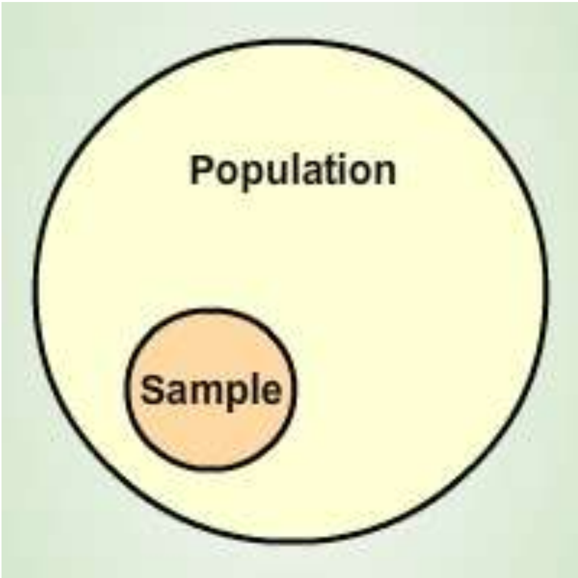
μ = Population mean

Notes:

- Different samples will yield different sampling errors
- The sampling error may be positive or negative (May be greater than or less than μ)
- The expected sampling error decreases as the sample size increases

Sampling Distribution

The sampling distribution is the distribution of all possible values of a statistic, for a given sample of size n, selected from a population.



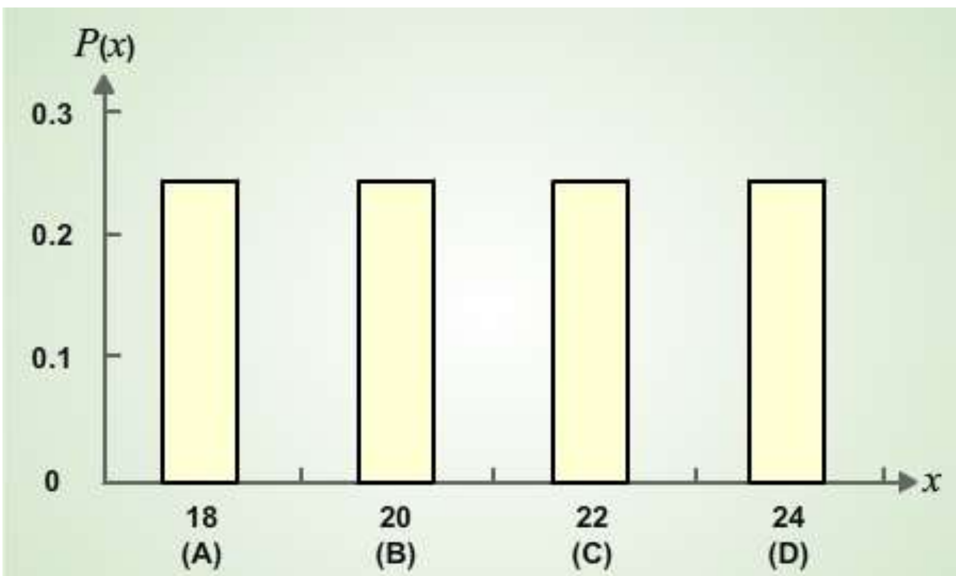
Population and Sample

Developing a Sampling Distribution > Example

Assume a class of 4 students as a population, $N = 4$ (A, B, C, D), let X be a random variable denoting the score of the students assume the values of x (score): 18 for A, 20 for B, 22 for C, 24 for D.

Summary Measures for the Population Distribution

$$\mu = \frac{\sum x_i}{N} = \frac{18+20+22+24}{4} = 21 \quad and \quad \sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}} = 2.236$$



Sampling distribution

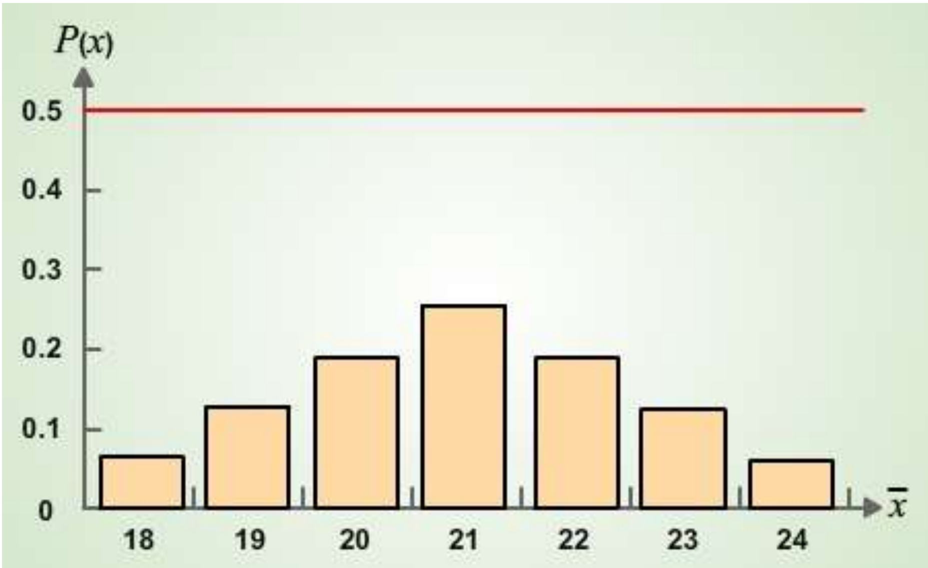
Calculate the summary measures for sample size

Want to draw a sample of size n =2 students. There are 16 possible samples (sampling with replacement):

1st observation	**2nd observation**			
	18	20	22	24
18	(18,18)	(18,20)	(18,22)	(18,24)
20	(20,18)	(20,20)	(20,22)	(20,24)
22	(22,18)	(22,20)	(22,22)	(22,24)
24	(24,18)	(24,20)	(24,22)	(24,24)

There are 16 Sample Means:

1st observation	2nd observation			
	18	20	22	24
18	18	19	20	21
20	19	20	21	22
22	20	21	22	23
24	21	22	23	24



Sampling distribution

The sampling distribution for all the sample means has bell shape.

And

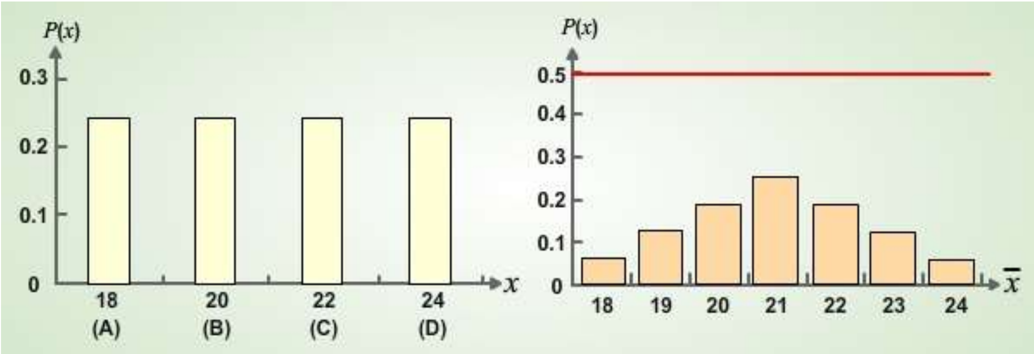
$$\mu_{\bar{x}} = \frac{\sum \bar{x}_i}{N} = \frac{18 + 19 + 21 + \cdots + 24}{16} = 21$$

$$\sigma_{\bar{x}} = \sqrt{\frac{\sum (x_i - \mu_{\bar{x}})^2}{N}}$$

$$= \sqrt{\frac{(18-21)^2 + (19-21)^2 + \cdots + (24-21)^2}{16}} = 1.58$$

Comparing the Population with its Sampling Distribution

Comparing the Population with its Sampling Distribution:



Population | Sample

Uniform Distribution | No longer uniform

$N = 4\,,\quad \mu = 21\,,\quad \sigma = 2.236$ | $n = 2\,,\quad \mu_{\bar{x}} = 21\,,\quad \sigma_{\bar{x}} = 1.58$

Theorem

If $X_1, X_2, \ldots X_n$ is a random sample of size n from a population that is normally distributed with mean (μ) and standard deviation (σ), regardless of the sample size (n) and the sampling distribution of $\bar{x}$ is also normal with:

$$Mean\quad \mu_{\bar{x}} = \mu\ \ and\ standard\ deviation\quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad (2)$$

Z -value for the sampling distribution of $\bar{x}$:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \qquad (3)$$

Apply the Finite Population Correction if:

- The sample is large, relative to the population (n is greater than 5% of N)
- The sampling is without replacement

Then:

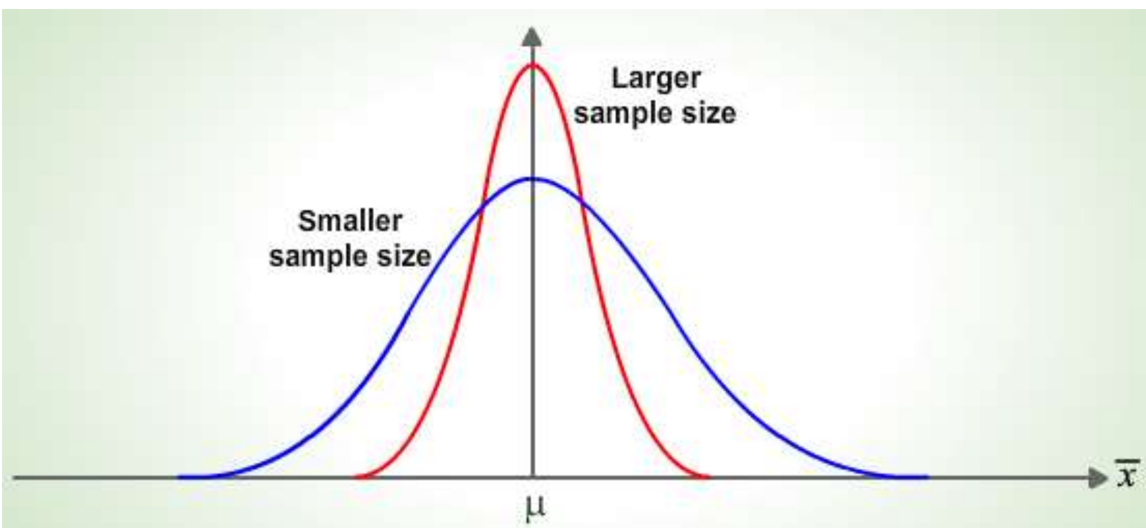
$$z = \frac{(\bar{x} - \mu)}{\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}} \qquad (4)$$

Sampling Distribution Properties

1. The Normal Population Distribution and Normal Sampling Distribution have the same mean:

$$\mu_{\bar{x}} = \mu \qquad (5)$$

2. For sampling with replacement, as n increases, $\sigma_{\bar{x}}$ decreases:



Normal population distribution

Recap

In this lecture, you have learned that:

- Sample statistics are used to estimate population parameters
- Sampling error is the difference between a value (a statistic) computed from a sample and the corresponding value (a parameter) computed from a population
- The sampling distribution is the distribution of all possible values of a statistic for a given sample of size n selected from a population
- The Normal Population Distribution and Normal Sampling Distribution have the same mean
- For sampling with replacement, the standard deviation for the sample mean decreases as n increases