Sampling Distributions
Chapter 7

207 QUA

King Saud University

December 15, 2022

Random Variables

- A random variable is a variable, the values of which occur according to some specified probability distribution.
- A *discrete* random variable is a random variable which takes only finitely many or countably infinitely many different values.
- A *continuous* random variable is a random variable that can take on a continuum of values.

Probability Distributions

- For a *discrete* random variable:
  The probability distribution is a mathematical formula that gives the probability of each value of the variable.
- For a *continuous* random variable:
  The probability distribution is a curve described by a mathematical formula that specifies, by way of areas under the curve, the probability that the variable falls within a particular interval.

Definition

- Sample statistics are statistics calculated from a random sample.
- Sample statistics are random variables.
- Random variables have probability distributions.
- The *sampling distribution* is the probability distribution of a statistic calculated from a random sample of a particular size

An Example

- Suppose we obtained the GPA from a random sample of 50 students from the college of business and computed the mean GPA for the sample.
- The mean GPA is a sample statistic and it's called the sample mean.
- If we obtained different random samples of size 50 you will compute a different mean for each sample.
- We are interested in the distribution of all potential mean GPAs we might calculate for samples of 50 students.

The Population

- Assume there is a population of four people ($N = 4$).
- The variable of interest, X, is the age of individuals.
- The values of X, in years, are $18, 20, 22$, and 24.

The Probability Distribution of X

X	$P(X = x_i)$
18	0.25
20	0.25
22	0.25
24	0.25

Table 1: The Probability Distribution of X

The Probability Distribution of X

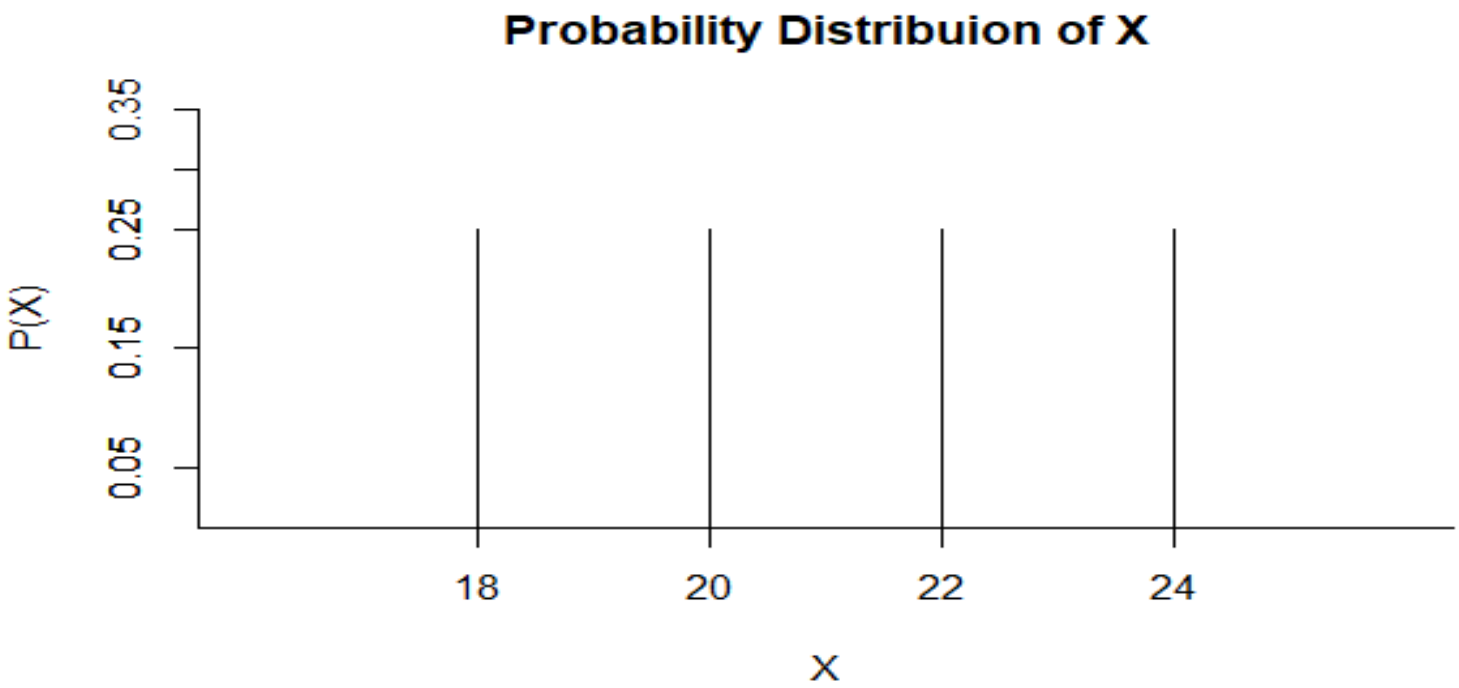

Figure 1: Graphical representation of Table 1

Summary Measures of The Probability distribution of X

- The population mean, μ_X, is:

$$\mu_X = \sum_{i=1}^{N} x_i P(X = x_i) = 21$$

- The population standard deviation, σ_X, is:

$$\sigma_X = \sqrt{\sum_{i=1}^{N} (x_i - \mu_X)^2 P(X = x_i)} = 2.236$$

All Possible Samples of size ($n = 2$)

1^{st} Obs	2^{nd} 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)

Table 2: All the 16 possible samples when sampling with replacement

All Possible Sample Means

1st Obs	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

Table 3: The means of the 16 possible samples

The Probability Distribution of $\bar{X}$

$\bar{X}$	$P(\bar{X} = \bar{x}_i)$
18	0.0625
19	0.125
20	0.1875
21	0.25
22	0.1875
23	0.125
24	0.0625

Table 4: The Probability Distribution of $\bar{X}$

The Probability Distribution of $\bar{X}$

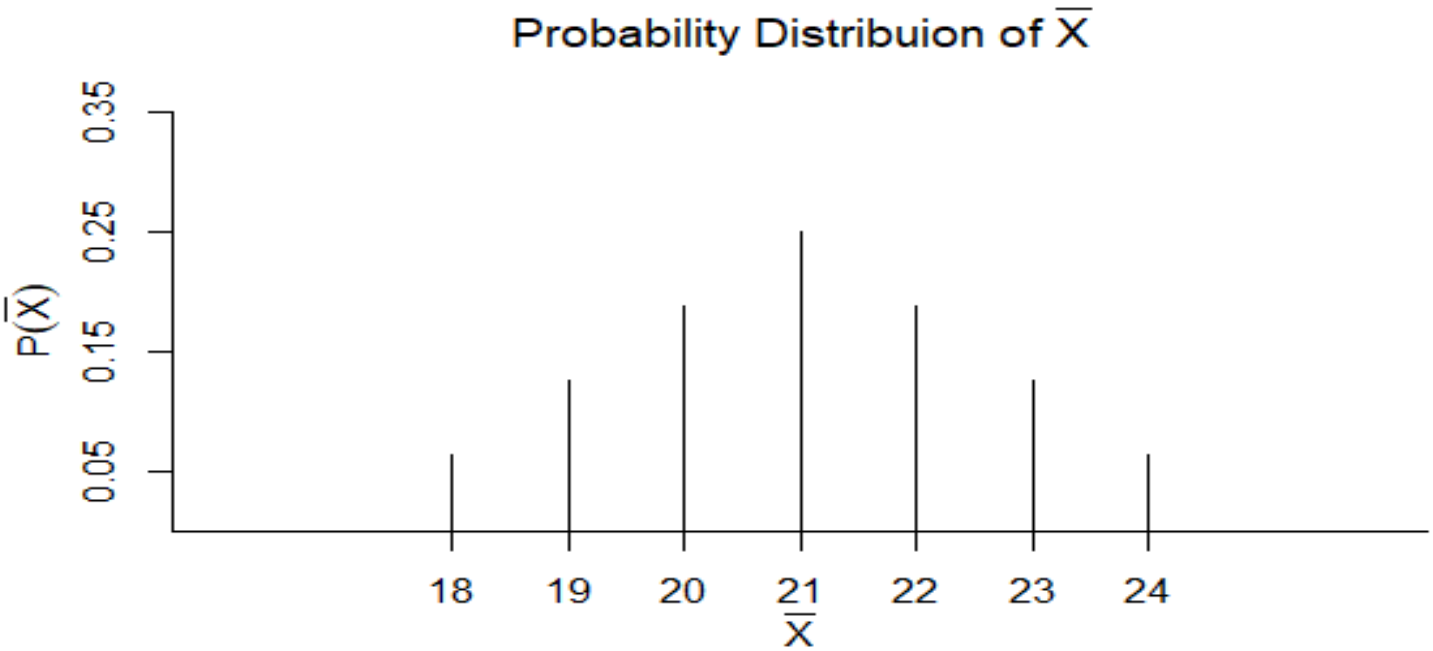

Figure 2: Graphical representation of Table 4

Summary Measures of The Probability distribution of $\bar{X}$

- The population mean, $\mu_{\bar{X}}$, is:

$$\mu_{\bar{X}} = \sum_{i=1}^{N} \bar{x}_i P(\bar{X} = \bar{x}_i) = 21$$

- The population standard deviation, $\sigma_{\bar{X}}$, is:

$$\sigma_{\bar{X}} = \sqrt{\sum_{i=1}^{N} (\bar{x}_i - \mu_{\bar{X}})^2 P(\bar{X} = \bar{x}_i)} = 2.236$$

- Let's practice!


1957