## Inferential Statistics and Probability
a Holistic Approach

Chapter 7
Central Limit Theorem



1

1

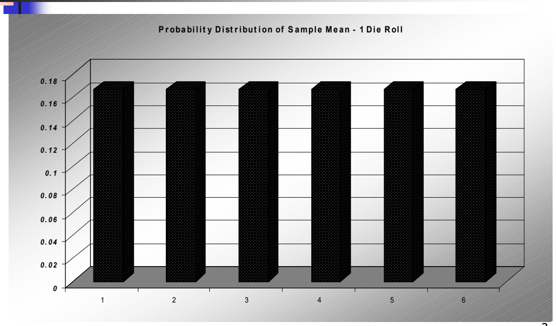
## Distribution of Sample Mean

- Random Sample: $X_1, X_2, X_3, \ldots, X_n$
  - Each $X_i$ is a Random Variable from the same population
  - All $X_i$'s are Mutually Independent
- $\overline{X}$ is a function of Random Variables, so $\overline{X}$ is itself Random Variable.
- In other words, the Sample Mean can change if the values of the Random Sample change.
- What is the Probability Distribution of $\overline{X}$ ?

2

2

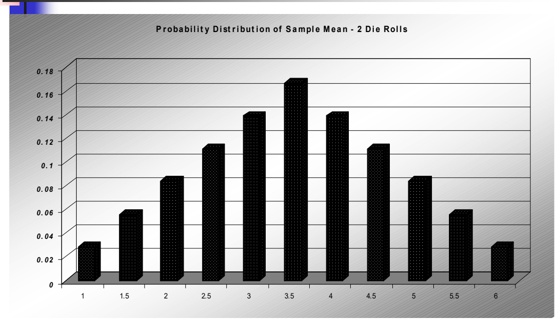
## Example – Roll 1 Die




3

3

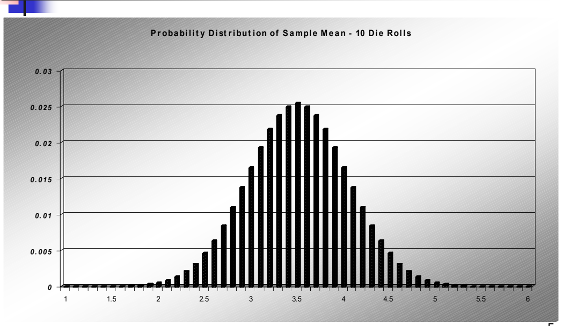
## Example – Roll 2 Dice




4

4

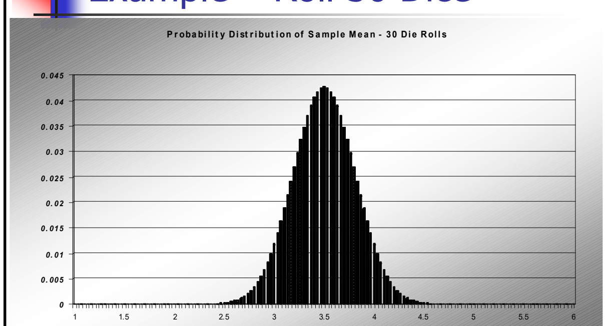
## Example – Roll 10 Dice




5

5

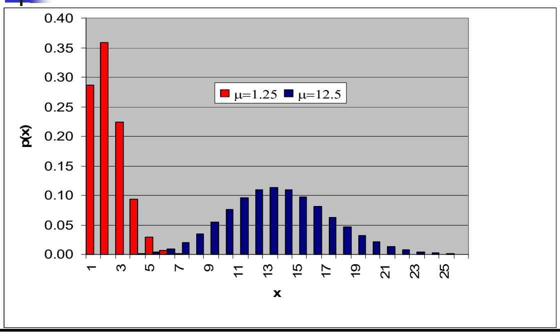
## Example – Roll 30 Dice




6

6

## Example - Poisson

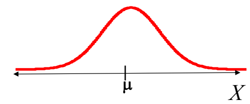


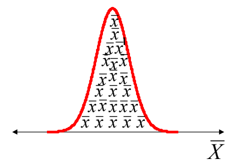

7

## Central Limit Theorem – Part 1

- IF a Random Sample of **any size** is taken from a population with a **Normal Distribution** with mean= μ and standard deviation = σ
- THEN the distribution of the sample mean has a Normal Distribution with:

$$\mu_{\overline{X}} = \mu \qquad \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$



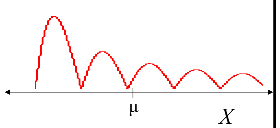

8

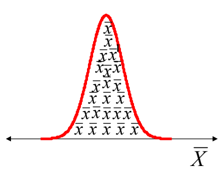
8

## Central Limit Theorem – Part 2

- IF a random sample of **sufficiently large size** is taken from a population with **any Distribution** with mean= μ and standard deviation = σ
- THEN the distribution of the sample mean has approximately a Normal Distribution with:

$$\mu_{\overline{X}} = \mu \qquad \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$




9

9

## Central Limit Theorem

3 important results for the distribution of $\overline{X}$

- Mean Stays the same

$$\mu_{\overline{X}} = \mu$$

- Standard Deviation Gets Smaller

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

- If n is sufficiently large, $\overline{X}$ has a Normal Distribution

10

10

## Example

The mean height of American men (ages 20-29) is μ = 69.2 inches. If a random sample of 60 men in this age group is selected, what is the probability the mean height for the sample is greater than 70 inches? Assume σ = 2.9".

$$P(\overline{X} > 70) = P\left( Z > \frac{(70 - 69.2)}{2.9/\sqrt{60}} \right)$$

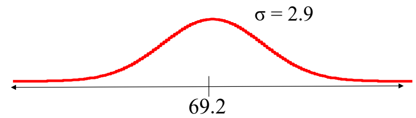
$$= P(Z > 2.14) = 0.0162$$

11

11

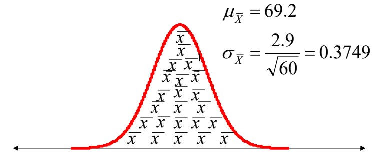
## Example (cont)

μ = 69.2
σ = 2.9




$\mu_{\overline{X}} = 69.2$

$\sigma_{\overline{X}} = \frac{2.9}{\sqrt{60}} = 0.3749$



12

12

## Example – Central Limit Theorem

The waiting time until receiving a text message follows an exponential (skewed) distribution with an expected waiting time of 1.5 minutes. Find the probability that the mean waiting time for the 50 text messages exceeds 1.6 minutes.

$\mu = 1.5 \qquad \sigma = 1.5 \qquad n = 50$

Use Normal Distribution (n>30)

$$P(\overline{X} > 1.6) = P\left(Z > \frac{(1.6-1.5)}{1.5/\sqrt{50}}\right) = P(Z > 0.47) = 0.3192$$

13

13

## Binomial Distribution and Sample Proportion

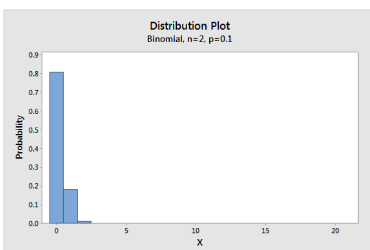
- Let X have a binomial distribution
  - n independent trials
  - p is the probability of success on a single trial
  - X is the number of successes in sample
- Sample proportion
  - $\hat{p}$ is the proportion of successes in sample

$$\hat{p} = \frac{X}{n}$$

14

14

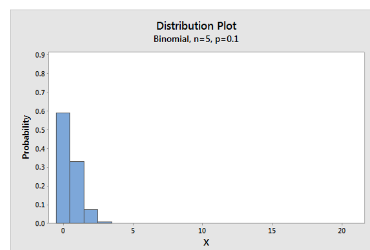
## Binomial np=0.2




15

15

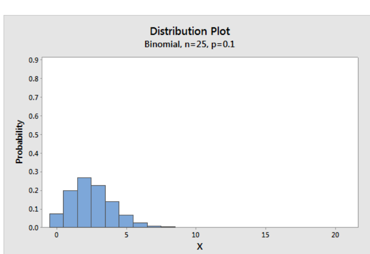
## Binomial np=0.5




16

16

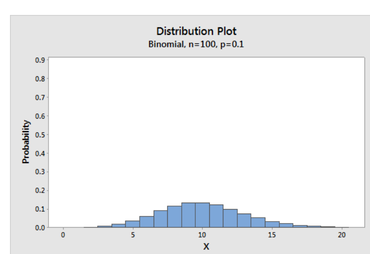
## Binomial np=2.5




17

17

## Binomial np=10




18

18

## Central Limit Theorem
### Sample Proportion

- The sample proportion of successes from a sample from a Binomial distribution is a random variable.
- If X is a random variable from a Binomial distribution with parameters n and p,
  and np ≥ 10 and n(1-p) ≥ 10,
  then the following is true for the Sample Proportion, $\hat{P}$ :
- $$\mu_{\hat{p}} = p \qquad \sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$
- The Distribution of $\hat{P}$ is approximately Normal.

19

19

## Example

- 45% of all community college students in California receive fee waivers.
- Suppose you randomly sample 1000 community college students to determine the proportion of students with fee waivers in the sample.
- 483 of the sampled students are receiving fee waivers.
- Determine $\hat{P}$ . Is the result unusual?

$$\hat{P} = \frac{483}{1000} = 0.483 \qquad \sigma_{\hat{p}} = \sqrt{\frac{0.45(1-0.45)}{1000}} = 0.0157 \qquad Z = \frac{\hat{p} - p}{\sigma_{\hat{p}}} = \frac{0.483 - 0.45}{0.0157}$$

$$Z = 2.10$$

- Result is unusual (more than 2 standard deviations from the expected value of the sample proportion).

20

20