Statistics for Data Science -1

Lecture 10.3: Distribution of a Binomial random variable

Usha Mohan

Indian Institute of Technology Madras

Learning objectives

Learning objectives

1. Derive the formula for the probability mass function for Binomial distribution.

Learning objectives

1. Derive the formula for the probability mass function for Binomial distribution.

Learning objectives

1. Derive the formula for the probability mass function for Binomial distribution.
2. Understand the effect of parameters n and p on the shape of the Binomial distribution.

Learning objectives

1. Derive the formula for the probability mass function for Binomial distribution.
2. Understand the effect of parameters n and p on the shape of the Binomial distribution.
3. Expectation and variance of the binomial distribution.

Learning objectives

1. Derive the formula for the probability mass function for Binomial distribution.
2. Understand the effect of parameters n and p on the shape of the Binomial distribution.
3. Expectation and variance of the binomial distribution.
4. To understand situations that can be modeled as a Binomial distribution.

Binomial random variable

Binomial random variable

Definition

X is a binomial random variable with parameters n and p that represents the number of successes in n independent Bernoulli trials, when each trial is a success with probability p. X takes values $0, 1, 2, \ldots, n$ with the probability

$$P(X = i) = \binom{n}{i} \times p^i \times (1 - p)^{(n-i)}$$

Example: Tossing a coin thrice

- $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$
- Success= head, Failure = tails

Example: Tossing a coin thrice

- $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$
- Success= head, Failure = tails
- X is the random variable which counts the number of heads in the tosses. $n = 3$ $p = 0.5$

Example: Tossing a coin thrice

- $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$
- Success= head, Failure = tails
- X is the random variable which counts the number of heads in the tosses. $n = 3$ $p = 0.5$
  - Probability mass function

X	0	1	2	3
$P(X = x_i)$	$\binom{3}{0}\frac{1}{2}^0\frac{1}{2}^3$	$\binom{3}{1}\frac{1}{2}^1\frac{1}{2}^2$	$\binom{3}{2}\frac{1}{2}^2\frac{1}{2}^1$	$\binom{3}{3}\frac{1}{2}^3\frac{1}{2}^0$

Example: Tossing a coin thrice

- $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$
- Success= head, Failure = tails
- X is the random variable which counts the number of heads in the tosses. $n = 3$ $p = 0.5$
  - Probability mass function

X	0	1	2	3
$P(X = x_i)$	$\binom{3}{0}\frac{1}{2}^0\frac{1}{2}^3$	$\binom{3}{1}\frac{1}{2}^1\frac{1}{2}^2$	$\binom{3}{2}\frac{1}{2}^2\frac{1}{2}^1$	$\binom{3}{3}\frac{1}{2}^3\frac{1}{2}^0$
$P(X = x_i)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Shape of the pmf for same n different p

A binomial distribution is

- right skewed if $p < 0.5$
- is symmetric if $p = 0.5$
- is left skewed if $p > 0.5$

. We demonstrate the same for $n = 4$ and different p

$n = 4, p = 0.3$, X = number of successes

$n = 4, p = 0.3,\ X = $ number of successes

- Let $n = 4$ independent Bernoulli trials.

$n = 4, p = 0.3,\ X\ =$ number of successes

- Let $n = 4$ independent Bernoulli trials.
- Let $p = 0.3$ is probability of success.

$n = 4, p = 0.3,\ X = $ number of successes

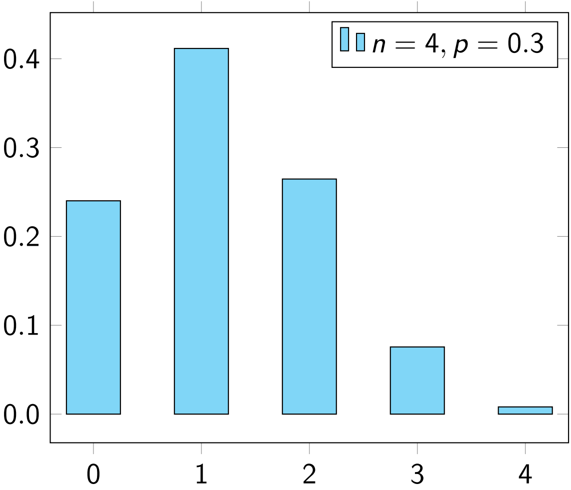
- Let $n = 4$ independent Bernoulli trials.
- Let $p = 0.3$ is probability of success.
- Let $X =$ number of successes in 4 independent trials.

$n = 4, p = 0.3,\ X = $ number of successes

- Let $n = 4$ independent Bernoulli trials.
- Let $p = 0.3$ is probability of success.
- Let $X =$ number of successes in 4 independent trials.
- The probability distribution of X

X	0	1	2	3	4
$P(X = i)$	0.2401	0.4116	0.2646	0.0756	0.0081

Graph of pmf of Binomial distribution- Right skewed



$n = 4, p = 0.5, X =$ number of successes

$n = 4, p = 0.5,\ X =$ number of successes

- Let $n = 4$ independent Bernoulli trials.

$n = 4, p = 0.5$, X = number of successes

- Let $n = 4$ independent Bernoulli trials.
- Let $p = 0.5$ is probability of success.

$n = 4, p = 0.5,\ X = $ number of successes

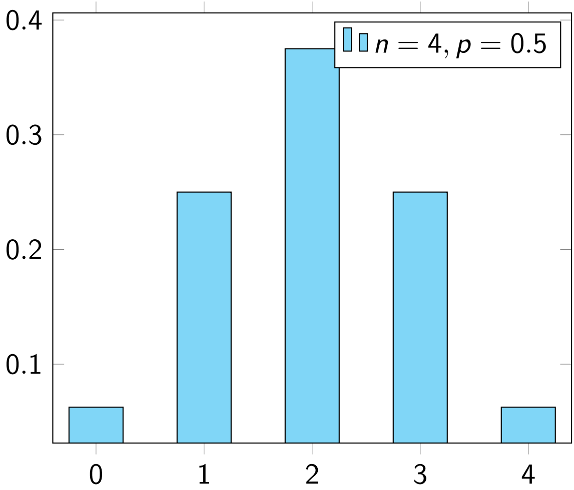
- Let $n = 4$ independent Bernoulli trials.
- Let $p = 0.5$ is probability of success.
- Let $X =$ number of successes in 4 independent trials.

$n = 4, p = 0.5, X =$ number of successes

- Let $n = 4$ independent Bernoulli trials.
- Let $p = 0.5$ is probability of success.
- Let $X =$ number of successes in 4 independent trials.
- The probability distribution of X

X	0	1	2	3	4
$P(X = i)$	0.0625	0.25	0.375	0.25	0.0625

Graph of pmf of Binomial distribution- symmetric



$n = 4, p = 0.8, X =$ number of successes

$n = 4, p = 0.8,$ X = number of successes

- Let $n = 4$ independent Bernoulli trials.

$n = 4, p = 0.8, X =$ number of successes

- Let $n = 4$ independent Bernoulli trials.
- Let $p = 0.8$ is probability of success.

$n = 4, p = 0.8,\ X\ =$ number of successes

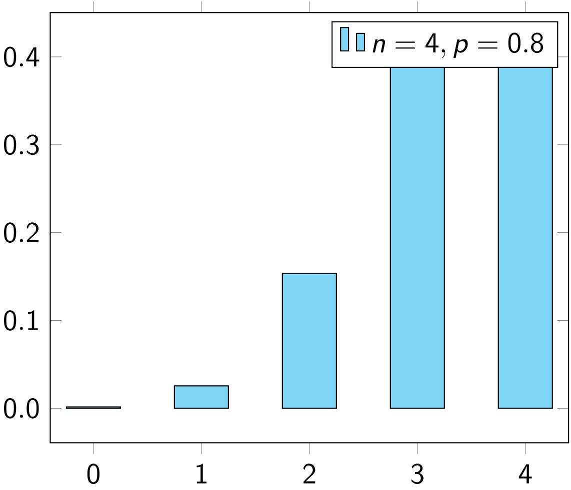
- Let $n = 4$ independent Bernoulli trials.
- Let $p = 0.8$ is probability of success.
- Let $X\ =$ number of successes in 4 independent trials.

$n = 4, p = 0.8, X$ = number of successes

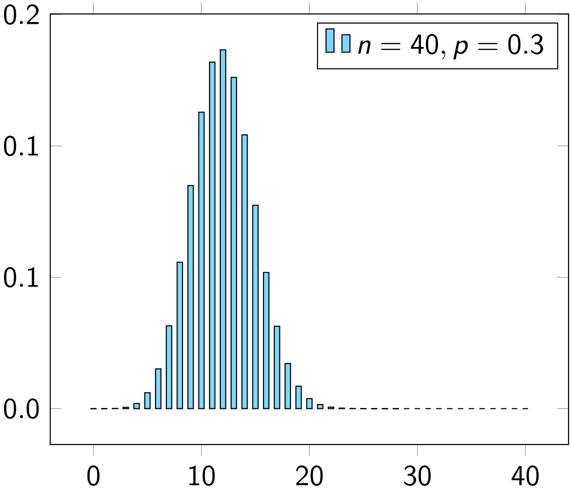
- Let $n = 4$ independent Bernoulli trials.
- Let $p = 0.8$ is probability of success.
- Let X = number of successes in 4 independent trials.
- The probability distribution of X

X	0	1	2	3	4
$P(X = i)$	0.0016	0.0256	0.1536	0.4096	0.4096

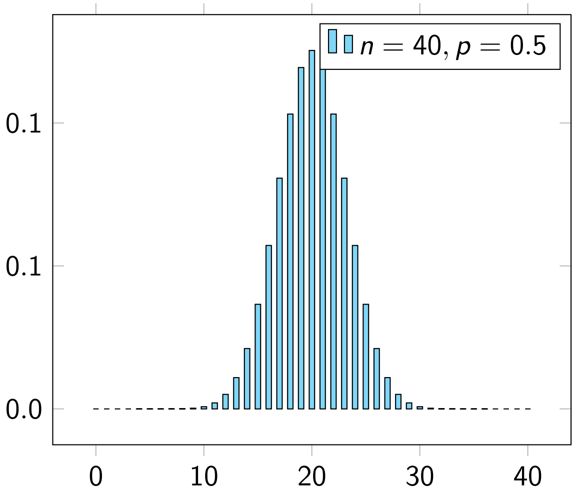
Graph of pmf of Binomial distribution- left skewed



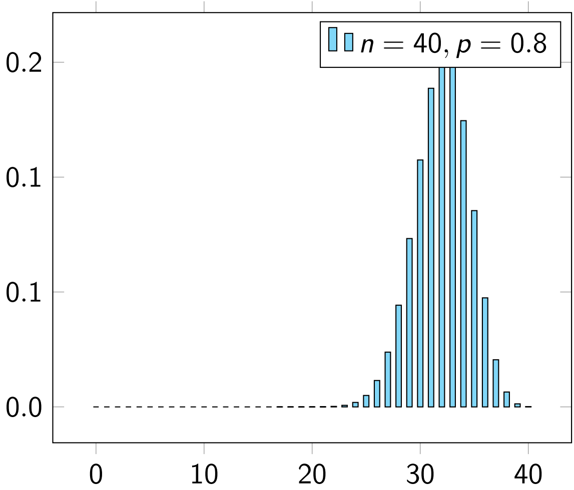
Graph of pmf of Binomial distribution- Right skewed- large n



Graph of pmf of Binomial distribution- symmetric-large n



Graph of pmf of Binomial distribution- left skewed- large n



Effect of n and p on shape of distribution

- small n, small p- right skewed
- small n, large p- left skewed
- small n $p = 0.5$- symmetric
- For large n, the binomial distribution approaches symmetry.

Section summary

- Introduced the Binomial random variable and its pmf.
- Studied effect of n and p on ths shape of the distribution.