Introduction to Statistics and Data Science using *eStat*

Chapter 5 Probability Distribution

5.3 Discrete Random Variable - Binomial Distribution -

Jung Jin Lee
Professor of Soongsil University, Korea
Visiting Professor of ADA University, Azerbaijan

5.3 Discrete Random Variable

5.3.1 Binomial Distribution

- Examples similar to tossing coins are observed frequently around us.
  - Products are inspected and classified as defective or good.
  - Ask one voter about the pros and cons of a particular candidate.

- There are two possible outcomes such as {defective, normal}, {pros, cons}.
  - probability of outcomes in each experiment is different.
  - Single experiment is called a Bernoulli trial
  - one outcome of two is referred to as 'success' and the other as 'failure'.

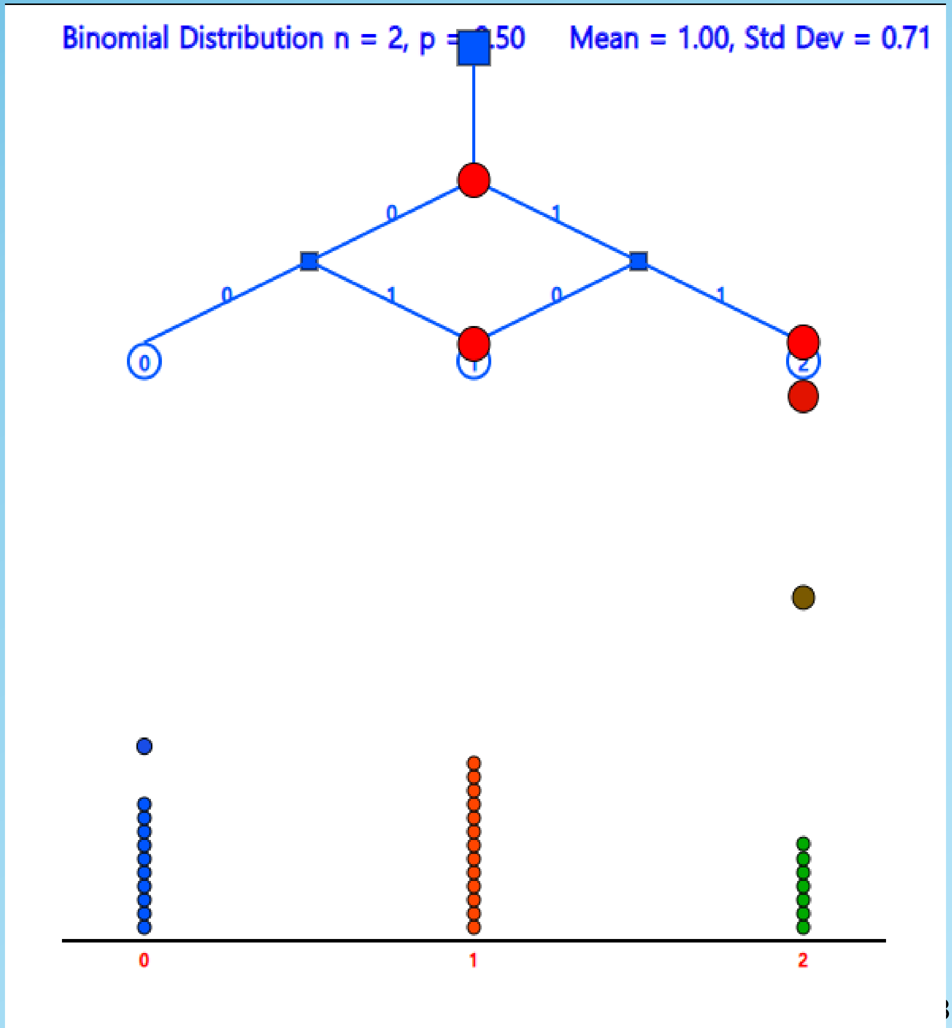
- Bernoulli trial is repeated and the number of 'success' is counted.
  - Throw a coin five times and examine the number of heads.
  - Inspect 100 products and count the number of defective products.
  - Number of voters in favor of a particular candidate among 50 voters.

  ⇨ Binomal random variable

5.3 Discrete Random Variable

5.4.1 Binomial Distribution - Simulation

- Drop a ball from the top and if it hit a bar, it has one-half chance to fall to the left (zero point) or right (one point).
- The dropped ball again falls to the left and right with a 1/2 chance.
- When you drop 100 balls, examine the sum of the total scores.



5.3 Discrete Random Variable

5.3.1 Binomial Distribution

- Counting of success for repeated Bernoulli trials trial with the same probability of success is called a **binomial random variable**,
  ⇨ **binomial distribution**.

[Ex 5.3.4] Four more baseball games will be played by Tiger this season. If Tiger has a 60% chance of winning every game, what is the probability of
1) losing all of them?
2) winning only once?
3) winning twice?
4) winning three times?
5) winning all four times?
6) Find the probability distribution function of the random variable
   X = ‘the number of games the tiger wins’.

5.3 Discrete Random Variable

<Answer of Ex 5.3.4>

- Sample space is all about winning or losing game and there are elements shown as follows by marking the winning in O and the losing in X.

S = {‘XXXX’, ‘OXXX’, ‘XOXX’, ‘XXOX’, ‘XXXO’, ‘OOXX’, ‘OXOX’, ‘OXXO’, ‘XOOX’, ‘XOXO’, ‘XXOO’, ‘OOOX’, ‘OOXO’, ‘OXOO’, ‘XOOO’, ‘OOOO’}

1) Tiger will loose all games is {'XXXX'} and the probability of this event is (0.4)×(0.4)×(0.4)×(0.4) $= 0.4^4$.

2) There are four events that the Tiger is winning once and losing three times such as {‘OXXX’, ‘XOXX’, ‘XXOX’, ‘XXXO’}. Since the probability of each event is (0.6)×(0.4)×(0.4)×(0.4), the probability of the Tiger winning once is $_4C_1\ 0.6 \times 0.4^3$

.

5.3 Discrete Random Variable

<Answer of Ex 5.3.4>

3) There are six events that the Tiger is winning two times and losing two times such as {'OOXX', 'OXOX', 'OXXO', 'XOOX', 'XOXO', 'XXOO'}. Since the probability of each event is (0.6)×(0.6)×(0.4)×(0.4), the probability of the Tiger winning twice is $_4C_2 \; 0.6^2 \times 0.4^2$.

4) There are four events that the Tiger is winning three times and losing one time such as {'OOOX', 'OOXO', 'OXOO', 'XOOO'}. Since the probability of each event is (0.6)×(0.6)×(0.6)×(0.4), the probability of the Tiger winning three times is $_4C_3 \; 0.6^3 \times 0.4^1$.

5) There is one event that the Tiger is winning four times such as {'OOOO'}. Since the probability of each event is (0.6)×(0.6)×(0.6)×(0.6), the probability of the Tiger winning all four times is $_4C_4 \; 0.6^4 \times 0.4^0$.

5.3 Discrete Random Variable

<Answer of Ex 5.3.4>

6) Probability distribution function of X = 'the number of games Tiger wins.

x	P(X=x)
0	${}_4C_0(0.4)^4 = 0.0256$
1	${}_4C_1(0.6)(0.4)^3 = 0.1536$
2	${}_4C_2(0.6)^2(0.4)^2 = 0.3456$
3	${}_4C_3(0.6)^3(0.4) = 0.3456$
4	${}_4C_4(0.6)^4 = 0.1296$

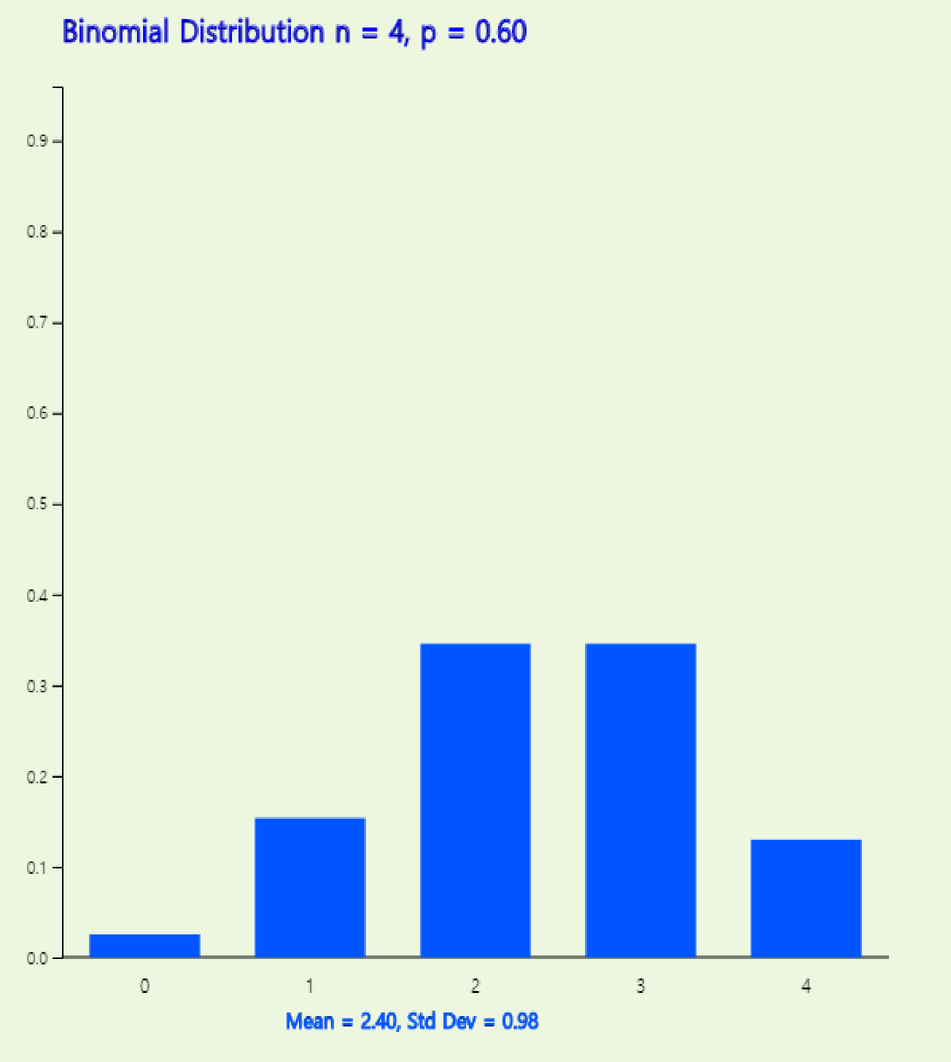
5.3 Discrete Random Variable

[Ex 5.3.5] By using『eStatU』, obtain the probability and probability distribution function of [Ex 5.3.4].

<Answer>

- Select a binomial distribution from 『eStatU』
- Enter n = 4, p = 0.6 and press the [Execute] button to display a binomial function graph.
- Click the 'Binary table' button.

n = 4	p = 0.600		
x	P(X = x)	P(X x)	P(X x)
0	0.0256	0.0256	1.0000
1	0.1536	0.1792	0.9744
2	0.3456	0.5248	0.8208
3	0.3456	0.8704	0.4752
4	0.1296	1.0000	0.1296



5.3 Discrete Random Variable

5.3.1 Binomial Distribution

If the probability of success is p in a Bernoulli trial and the trial is repeated n times independently, the probability distribution function that the random variable X = 'the number of success' is is as follows. It is called a binomial distribution and denoted as $B(n, p)$.

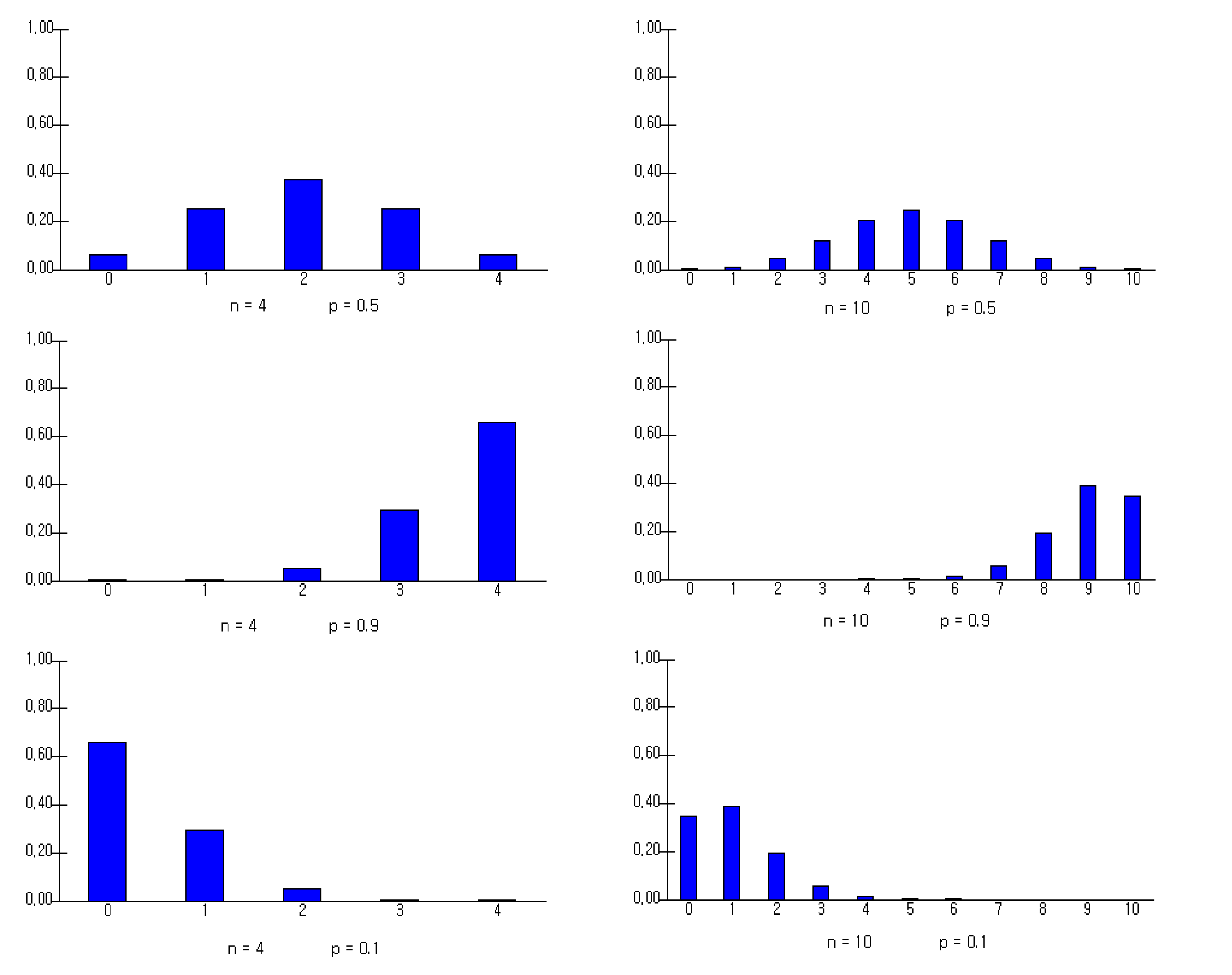
$$f(x) = {}_nC_x\, p^x \times (1-p)^{n-x}, \qquad x = 0, 1, \dots, n$$

The expectation and variance of the binomial distribution are as follows:

$$E(X) = np, \quad V(X) = np(1-p)$$

5.3 Discrete Random Variable

5.3.1 Binomial Distribution



5.3 Discrete Random Variable

5.3.1 Binomial Distribution

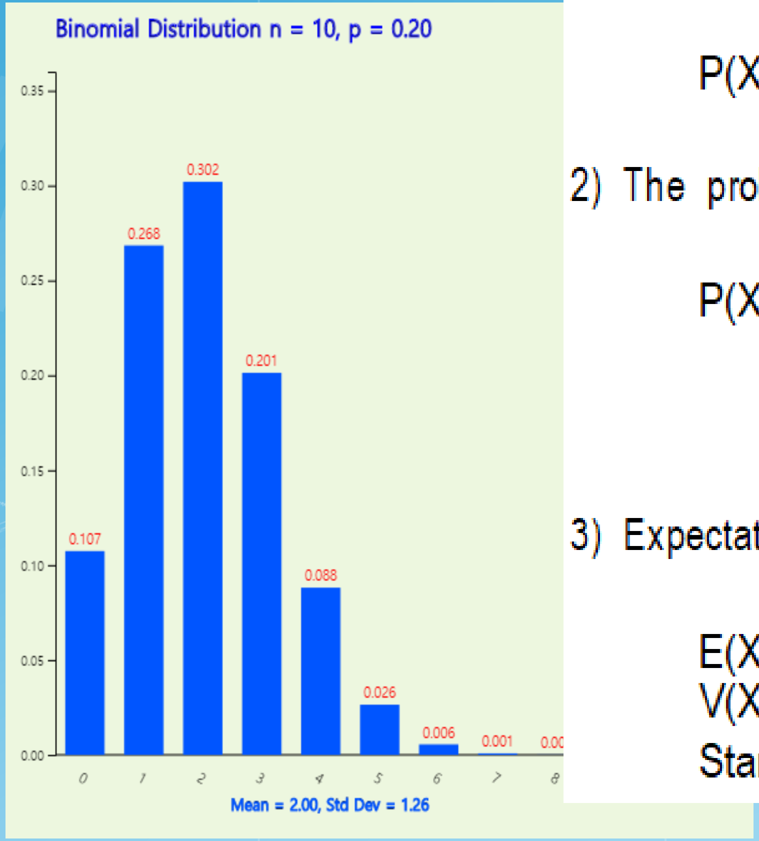
[Ex 5.3.6] Past experience shows that a salesperson from an insurance company has a 20% chance of meeting a customer and insuring that person. The salesperson is scheduled to meet 10 customers this morning. Calculate the following probabilities directly and check using 『eStatU』

1) What are the probability that three customers will get insurance?
2) What is the probability that two or more customers will get insurance?
3) How many people on average would sign up? Standard deviation?

5.3 Discrete Random Variable

5.3.1 Binomial Distribution

<Answer of 5.3.6>



- This is a Binomial distribution when n = 10, p = 0.2.

1) The probability that three customers will get insurance is as follows.

$$P(X=3) = {}_{10}C_3 (0.2)^3 (1-0.2)^{10-3} = 0.2013$$

2) The probability that two or more customers will get insurance may use the complement event as follows.

$$P(X \geq 2) = 1 - P(X=0) - P(X=1)$$
$$= 1 - {}_{10}C_0 (0.2)^0 (1-0.2)^{10} - {}_{10}C_1 (0.2)^1 (1-0.2)^{10-1}$$
$$= 1 - 0.1074 - 0.2684 = 0.6242$$

3) Expectation and standard deviation are as follows..

$$E(X) = np = 10 \times 0.2 = 2$$
$$V(X) = np(1-p) = 10 \times 0.2 \times 0.8 = 1.6$$
$$\text{Standard deviation} = \sqrt{1.6} = 1.265$$

Thank you