Introduction to Statistics and Data Science using *eStat*

Chapter 5 Probability Distribution

5.3 Discrete Random Variable

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

5.3 Discrete Random Variable

- Statistical experiment in which two coins are thrown.
  ⇨ sample space = {'Tail-Tail', 'Tail-Head', 'Head-Tail' and 'Head-Head'}.
  ⇨ probability of each element = 1/4.
- Interested in counting the number of heads.
  ⇨ X is 'number of heads',
  ⇨ possible values of X can be 0, 1, or 2.
- Random variable : function from sample space to a real number

Sample Space	X=Number of {Head}
'Tail-Tail'	0
'Head-Tail'	1
'Tail-Head'	1
'Head-Head'	2

- If the possible values of random variable are finite or countably infinite, its is called a discrete random variable.
- If the possible values of random variable are uncountably infinite, it is called a continuous random variable.

5.3 Discrete Random Variable

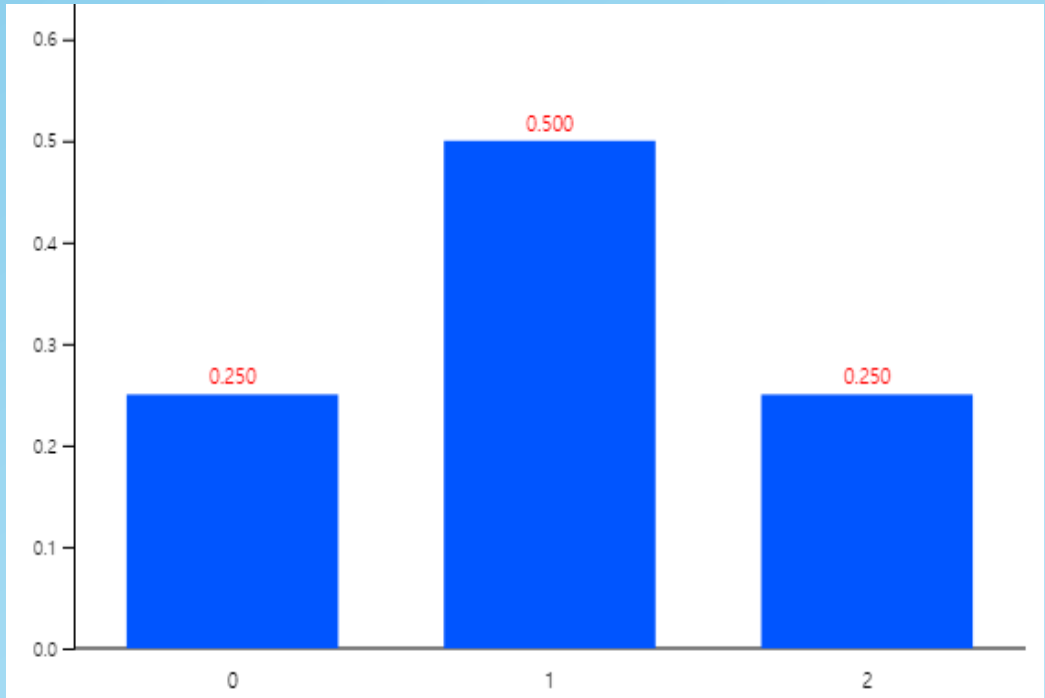
- Probability that X being zero is 1/4 because P{Tail-Tail} is 1/4,
- Probability that X being 1 is 2/4 because P({Tail-Head, Head-Tail}) is 2/4,
- Probability that X being 2 is 1/4 because P({Head-Head}) is 1/4.
- Summarized probabilities for value of X is **probability distribution function** denoted as f(x).

1) Table style

X = x	P(X=x)
0	1/4
1	2/4
2	1/4
계	1

2) Function style

$$f(x) = 1/4,\ x=0$$
$$= 2/4,\ x=1$$
$$= 1/4,\ x=2$$



5.3 Discrete Random Variable

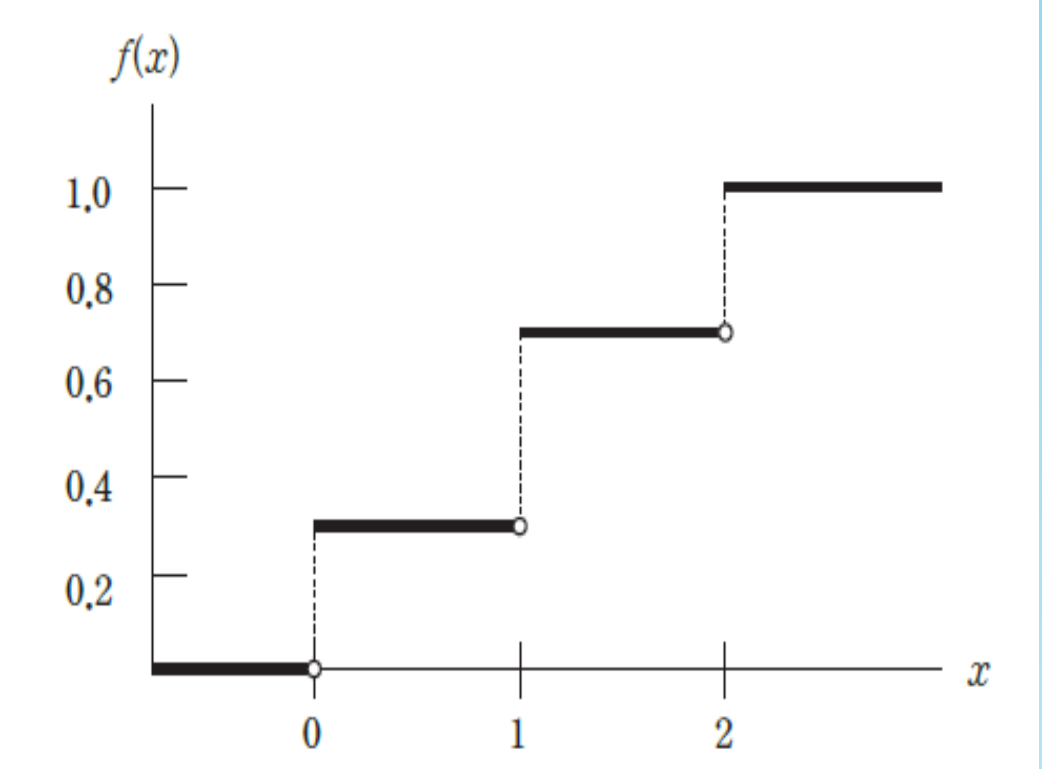
- Cumulative probability of $P(X \le x)$ as the value of random variable X increases is referred to as **cumulative distribution function,** F(x).

1) Table style

X = x	P(X ≤ x)
0	1/4
1	3/4
2	4/4

2) Function style

$$F(x) = 0 \quad , x<0$$
$$= 1/4, \ 0 \le x < 1$$
$$= 3/4, \ 1 \le x < 2$$
$$= 1 \quad , 2 \le x$$



5.3 Discrete Random Variable

[Example 5.3.1] There are 200 families living in a village. The number of visits to hospitals by each household over the past year is as follows. Obtain the probability distribution function and the cumulative distribution function of X = 'hospital visit'.

Hospital visit	0	1	2	3	4
Household	74	80	30	10	6

<Answer>

Probability distribution function

X = x	P(X=x)
0	0.37
1	0.40
2	0.15
3	0.05
4	0.03
Total	1.00

Cumulative distribution function

X = x	P(X≤x)
0	0.37
1	0.77
2	0.92
3	0.97
4	1.00

5.3 Discrete Random Variable

- If possible values of X are $x_1, x_2, \ldots, x_n$, a mean and variance of X are used as measures of central tendency and dispersion.

- Mean of X called an expectation of X, denoted E(X) or μ,
- Variance of X, denoted as V(X) or σ^2.
- Standard deviation of X, denoted σ, is the square root of the variance X.

$$E(X) = \mu = \sum_{i=1}^{n} x_i P(X = x_i)$$

$$V(X) = \sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 P(X = x_i) = \sum_{i=1}^{n} x_i^2 P(X = x_i) - \mu^2$$

5.3 Discrete Random Variable

[Example 5.3.2] Find the expected value and variance of the random variable X = 'Number of Heads' when tossing a coin twice' such as in Table 5.3.2.

<Answer>

$$E(X) = \mu = \sum_{i=1}^{n} x_i P(X = x_i) = 0 \times \frac{1}{4} + 1 \times \frac{2}{4} + 2 \times \frac{1}{4} = 1$$

$$V(X) = \sum_{i=1}^{n} x_i^2 P(X = x_i) - \mu^2 = 0^2 \times \frac{1}{4} + 1^2 \times \frac{2}{4} + 2^2 \times \frac{1}{4} - 1^2 = \frac{1}{2}$$

5.3 Discrete Random Variable

- **Expectation and variance of aX + b**

$$E(aX+b) = a\,E(X) + b$$
$$V(aX+b) = a^2 V(X)$$

5.3 Discrete Random Variable

[Example 5.3.3] The mean of a mid-term score on statistics was 60 points and the variance was 100. To adjust the score, the professor is thinking of the following alternative. Find the mean and variance of each alternative.

1) Add 20 points to each student's score.
2) Each student's score is multiplied by 1.4.
3) Multiply each student's score by 1.2 and add 10 points.

5.3 Discrete Random Variable

<Answer>

- X is the mid-term score and its mean and variance are E(X) = 60 and V(X) = 100.

1) Mean and variance of the new random variable X + 20 are as follows.
   E(X + 20) = E(X) + 20 = 60 + 20
   V(X + 20) = V(X) = 100
2) Mean and variance of the new random variable 1.4X are as follows.
   E(1.4X) = 1.4 E(X) = 1.4 × 60 = 84
   V(1.4X) = 1.4^2 V(X) = 1.96 × 100 = 196
3) Mean and variance of the new random variable 1.2X + 10 are as follows.e 1.4X.
   E(1.2X + 10) = 1.2 E(X) + 10 = 1.2 × 60 + 10 = 82
   V(1.2X + 10) = 1.2^2 V(X) = 1.44 × 100 = 144

5.3 Discrete Random Variable

- **Standardized random variable**

- If the mean of a random variable X is μ, and the standard deviation is σ, then $Z = \frac{X - \mu}{\sigma}$ is a new random variable with the mean of 0 and the variance of 1.

Thank you