Introduction to Statistics and Data Science using *eStat*

Chapter 5 Probability Distribution

5.1 Definition of Probability

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

5.1 Definition of Probability

- **Similar events occur frequently or carried out around our lives.**

  - **A machine is producing products repeatedly at a production plant.**
    **=> product is either a normal or a defective product,
    but not known what will be.**
  - **We order pizza at home every Sunday.**
    **=> It usually takes about 30 minutes for a pizza to be delivered to the
    house, but exact time is not known.**

- **What these examples have in common.**
  - ① **Repetition of similar events.**
  - ② **Various possible outcomes are known.**
  - ③ **There is no telling what exactly will happen.**
  - **=> Statistical experiment**

- **Statistical experiment.**
  => Examples of statistical experiment?
  => Examples of non-statistical experiment?

- **Sample space :** **Set of all possible outcomes from a statistical experiment**
  => denoted by S,   S = {normal, defective}
  - Discrete sample space: Number of elements in S is finite or countable
  - Continuous sample space: Number of elements in S is uncountable

- **Event :** **A subset of the sample space**
  => denoted in English capitals A, B, and C ...
  => A = {defective}

5.1 Definition of Probability

- **Probability is 'representation of likelihood of an event ' between 0 and 1.**

- If an event is likely to occur, the probability is close to 1.

  If it is unlikely to occur the probability is close to 0.

- Two definitions of probability,
  - classical definition
  - relative frequency definition

❖ **Classical definition of probability**

▪ **Assume that <u>all elements in the sample space are likely to occur equally</u>.**

▪ **The probability of an A will occur in case of discrete sample space is**

$$P(A) = \frac{\text{Number of elements belonging to event A}}{\text{Total number of elements in sample space}}$$

▪ **The probability of an A will occur in case of continuous sample space is**

$$P(A) = \frac{\text{Measurement of elements belonging to event A}}{\text{Measurement of elements in sample space}}$$

* Measurement can be length, area, volume etc.

[Ex 5.1.1] An office worker went on a business trip to a city and there are two restaurants (A and B) near his lodging.
  - He threw a dice to count the number of points that appear on the top.
  - If he had odd numbers, he would go to the restaurant A,
     if he had even numbers, he would go to the restaurant B.
  - What is the probability that the restaurant A would be picked?

<Answer>
• Sample space is {1, 2, 3, 4, 5, 6}.
• The number of odd events is {1, 3, 5}, so there are three elements.
• Probability that restaurant A will be selected is 3/6 = 1/2.

5.1 Definition of Probability

[Ex 5.1.2] The time it takes for a pizza to be delivered to the house has the same possibility for any time from 10 to 30 minutes .
 - What is the probability that a pizza delivery will be delivered between 20 and 25 minutes?

<Answer>
- Sample space is all values from 10 to 30 minutes {(10,30)}.
- Event where pizza is delivered between 20 and 25 minutes is { (20,25) }.
- Probability of this event is (25-20) / (30-10) = 0.25 by measuring distance.

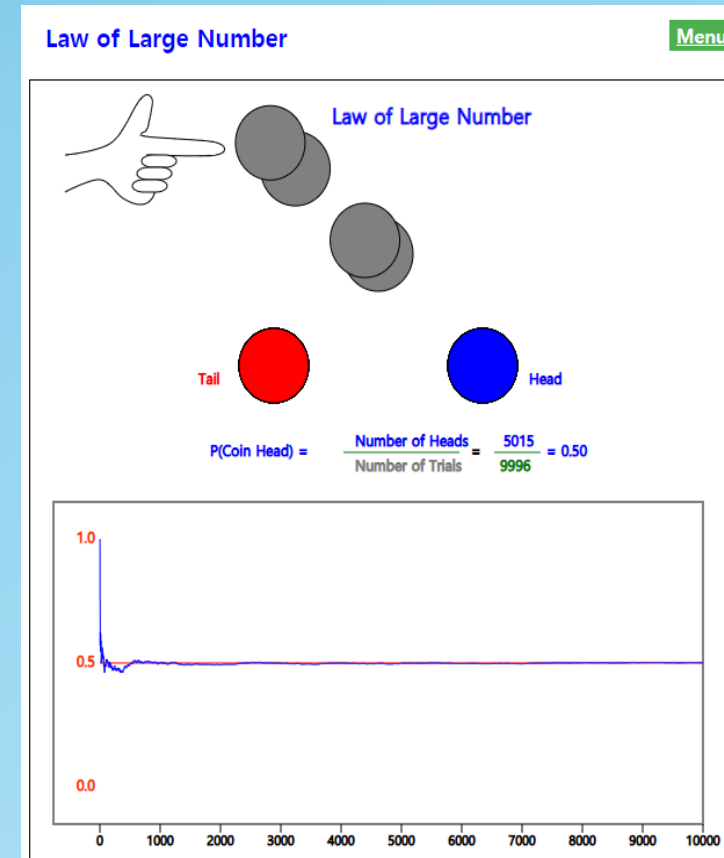# 5.1 Definition of Probability

❖ **Relative frequency definition of probability**

▪ **Probability that event A will occur is the rate at which event A occurs when the experiments are conducted under the same conditions repeatedly.**

❖ **Law of Large Number**

▪ **If a coin is thrown many times, the probability of {Head} event converges to half**

Law of Large Number

$$P(\text{Coin Head}) = \frac{\text{Number of Heads}}{\text{Number of Trials}} = \frac{5015}{9996} = 0.50$$

Thank you