SECTION 2: PROBABILITY

2.1 Introduction

The oldest way of defining probabilities are based on called the classical probability definition which is that if there are N equally possibilities, of which one must occur and m are regarded as favorable, or as “success”, then the probability of a success is given by the ratio $\frac{m}{N}$.

Probability is based on observations of certain events. **Probability of an event** is the ratio of the number of observations of the event to the total numbers of the observations. **An experiment** is a situation involving chance or probability that leads to results called outcomes. **An outcome** is the result of a single trial of an experiment. **The probability of an event** is the measure of the chance that the event will occur as a result of an experiment.

The probability of an event tells us that how likely the event will happen. Situations in which each outcome is equally likely, then we can find the probability using probability formula. Probability is a chance of prediction. If the probability that an event will occur is "x", then the probability that the event will not occur is "1 - x". If the probability that one event will occur is "a" and the independent probability that another event will occur is "b", then the probability that both events will occur is "ab". Probability of an event A can be written as:

P(A) =Number of favorable outcomes /Total number of possible outcomes.
(http://www.probabilityformula.org/)

Even the classic definition of probability is somehow useful for finding the events or occurrences probabilities but it has limited applicability. In many situations, outcomes are not equally likely to occur. It is widely used if we are concerned with the outcome of an election, if we are concerned with a person’s recovery from a disease or if we are concerned with sales of a product in a day.

For example, let assume that we are interested in the probability of a disease occurrences in a population (the people living in a region). In this population, there are N people and we know how many people having that disease in the population (assuming there are m people). So we can calculate the probability of the disease occurrences by the proportion of m to N (m/N). Many other examples can be given like that.

In the examples such as these the probability of an event (outcome or happening) is based on the ratio of same kind outcomes in total occurrences.

Beside the classic probability concept, Russian Mathematician Andrey Nikolaevich Kolmogorov (1950) introduced first a set of postulates (axioms) for probability and then the theory of probability was developed under the axiomatic probability concept.

Before giving these axioms, we need learn many definitions such as **sample space**, **random event.**

In the probability concept, our interest is based on random events or experiments. The word “random” or “randomness” is really important. Here we do not know which outcome can be occurred before an event happens but we know all possible outcomes with that event. In this explanation, we need a set of possibilities with related the event.

Definition 2.1. The set of all possible outcomes of an experiment is called the ***Sample Space*** for the experiment. It is usually denoted by the letter S (capital S).

Definition 2.2. Each outcome (point) or a set of outcomes of the sample space is referred as a ***random event*** briefly an event and denoted by letter A or B.

Example 2.1. Suppose a die is rolled and the number of dots on the upturned face is recorded.

Sample space for the example would be S={1, 2, 3, 4, 5, 6}

Example 2.2. Suppose a coin tossing experiment, the sample space S={T, H}

Example 2.3. Suppose rolling two dices, the sample space S={(1,1), (1,2),..,(1,6), (2,1),…, (2,6),…, (6,1), (6,2),…,(6,6)} so there are 6×6=36 points (outcomes) of the experiment and these consist of the sample space.

Example 2.4. A bowl contains 3 blue, 4 red and 5 black identical balls. 3 balls are chosen randomly from the bowl. How many points are found in the sample space?

In case of replacement, 3×3×3=27
In case of without replacement, 3×3×3=27

Example 2.5. Suppose that measuring the weight of fetus. The sample space for the example is S ={ x | x>0 }, where the number of elements in the set is uncountable.

Probabilities are values of a set function, also called a probability measure, for, as we shall see, this function assigns real numbers to the various subsets of a sample space S. As we shall formulate them here, the postulates (axioms) of probability apply only when the sample space S is discrete.

Kolmogorov's Axioms

Axiom 1. The probability of an event is a nonnegative real number; that is, P(A) ≥ 0 for subset A of S.
Axiom 2. P(S)=1.
Axiom 3. If $A_1, A_2, ...$ is a finite or infinite sequence of mutually exclusive events of S, then
$P(A_1 \cup A_2 \cup ...) = P(A_1) + P(A_2) + ...$ (∪:union sign)

Axioms per se (kendi başına, kendiliğinden) require no proof but the axioms and frequency (classic) probability concept are really related. Since proportions are always positive or zero, the first axiom is complete agreement with frequency interpretation. The second axiom states indirectly that certainty is identified with probability of 1; after all, it is always assume that one of the possibilities in S must occur, and it is to certain event that we assign a probability of 1. For example, a coin with two Heads on both sides, the sample space S={ "H" }, so when filliping the coin, the result Head always occurs it means that an event A=S={ "H" } is certain to occur and the probability of A is 1. As far as the frequency interpretation is concerned, a probability of 1 implies that the event in question will occur 100 percent of the time.

Theorem 2.1. If A is an event in a discrete sample space S, then P(A) equals the sum of the probabilities of the individual outcomes comprising A.

Proof: Let $B_1, B_2, \ldots,$ be the finite or infinite sequence of outcomes that comprise the event *A*. Thus,

$$A = B_1 \cup B_2 \cup$$

And since the individual outcomes, $B_1, B_2, \ldots,$ are mutually exclusive, the third axioms of probability yields

$$P(A) = P(B_1) + P(B_2) + \ldots$$

This completes the proof.

Theorem 2.2. If an experiment can result in any one of *N* different equally likely outcomes, and if *m* of these outcomes together constitute event *A*, then the probability of event *A* is

$$P(A) = \frac{m}{N}$$

Proof: Let $B_1, B_2, \ldots, B_N$, represent the individual outcomes in *S*, each with probability 1/N. If *A* is the union of *m* of these mutually exclusive outcomes, and it does matter which ones, then

$$\begin{aligned} P(A) &= P(B_1 \cup B_2 \cup \ldots \cup B_m) \\ &= P(B_1) + P(B_2) + \ldots + P(B_m) \\ &= \underbrace{\frac{1}{N} + \frac{1}{N} + \ldots + \frac{1}{N}}_{m} = \frac{m}{N} \end{aligned}$$

Theorem 2.3. If *A* and A' are complementary (tümleyen) events in a sample space *S*, then

$$P(A') = 1 - P(A)$$

Proof: By the second axiom, $P(S) = 1$ and $A \cup A' = S$ so *A* and A' are mutually exclusive then,

$$\begin{aligned} &P(S) = 1 \\ &P(A \cup A') = 1 \\ &P(A) + P(A') = 1 \\ &P(A') = 1 - P(A) \end{aligned}$$

the proof is completed.

Theorem 2.4. $P(\varnothing) = 0$ for any sample space S.

Proof: Since S and ∅ are mutually exclusive and S∪∅=S in accordance with the definition of empty set ∅, it follows that

$$\begin{aligned} &P(S) = P(S \cup \varnothing) \\ &P(S) = P(S) + P(\varnothing) \\ &P(\varnothing) = 0. \end{aligned}$$

Theorem 2.5. If A and B are events in a sample space S and $A \subset B$, then $P(A) \leq P(B)$.

Proof: Since $A \subset B$, we can write

$$B = A \cup (A' \cap B)$$

as can easily be verified by means of Venn diagram. Then, since A and $A' \cap B$ are mutually exclusive, we get

$$\begin{aligned} P(B) &= P(A) + P(A' \cap B) && \text{by Axiom 3} \\ &\geq P(A) && \text{by Axiom 1.} \end{aligned}$$

Theorem 2.6. $0 \leq P(A) \leq 1$ for any event A.

Proof: Using the Theorem 2.5 and the fact that $\varnothing \subset A \subset S$ for any event A in S, we have

$$P(\varnothing) \leq P(A) \leq P(S)$$

Then, $P(\varnothing) = 0$ and $P(S) = 1$ leads to the result that

$$0 \leq P(A) \leq 1$$

Theorem 2.7. If A and B are two events in a sample space S, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Proof: Assigning a, b, and c to the mutually exclusive events $A \cap B$, $A \cap B'$, and $A' \cap B$ as in the Venn diagram. We find that

$$\begin{aligned} P(A \cup B) &= \text{a} + \text{b} + \text{c} \\ &= (\text{a} + \text{b}) + (\text{c} + \text{a}) - \text{a} \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

Theorem 2.8. If A, B, and C are three events in a sample space S, then

$$\begin{aligned} P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - \\ - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) \end{aligned}$$

Proof: Writing $A \cup B \cup C$ as $A \cup (B \cup C)$ and using the formula of Theorem 2.7 twice, once for $P[A \cup (B \cup C)]$ and once for $P(B \cup C)$, we get

$$\begin{aligned} P(A \cup B \cup C) &= P[A \cup (B \cup C)] \\ &= P(A) + P(B \cup C) - P[A \cap (B \cup C)] \\ &= P(A) + P(B) + P(C) - P(B \cap C) - P[A \cap (B \cup C)] \end{aligned}$$

Then, using the distributive law that $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ and getting probability of this event

$$
\begin{aligned}
P[A\cap(B\cup C)] &= P[(A\cap B)\cup(A\cap C)]\\
&= P[(A\cap B)]+P[(A\cap C)]-P[(A\cap B)\cap(A\cap C)]\\
&= P[(A\cap B)]+P[(A\cap C)]-P(A\cap B\cap C)]
\end{aligned}
$$

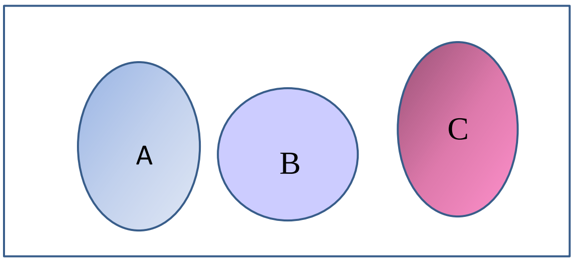
and hence that the result is substituted into the first Eq. The theorem is proofed now.

$$
\begin{aligned}
P(A\cup B\cup C) = P(A)+P(B)+P(C)-P(A\cap B)-\\
-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)
\end{aligned}
$$

Example 2.6. Suppose $P(A)=1/2$, $P(B)=1/8$, $P(C)=1/4$ where A, B, C are mutually exclusive. Determine the values of:

a) $P(A\cup B)$
b) $P(A\cup B\cup C)$
c) $P(A-B)$
d) $P(\bar{A}\cap\bar{B})$

Solution:



a) Since A and B are mutually exclusive $P(A\cup B)=P(A)+P(B)=5/8$
b) Since A, B and C are mutually exclusive $P(A\cup B\cup C)=P(A)+P(B)+P(C)=7/8$
c) $P(A-B)=P(A)=1/2$
d) $P(\bar{A}\cap\bar{B})=1-P(A\cup B)=1-P(A)-P(B)=1-5/8=3/8$

Example 2.7. Suppose you play a game over and over again, each time with chance 1/N of winning the game, no matter what the results of previous games. What is the probability of at least one win in the n games?

Solution:

A = {at least one win in the n games}

$$P(A)=1-\left(1-\frac{1}{N}\right)^n$$

Example 2.8. In a certain population, 10% of the people are rich, 5% are famous, and 3% are rich and famous. For a person picked at random from this population.

a) What is the chance that person is rich but not famous?
b) What is the chance that person is either rich or famous?

Solution:

a) P (rich but not famous) = P (rich) - P (rich and famous) = 10%-3%=7%
b) P (rich or famous) = P (rich) + P (famous)- P (rich and famous) = 10%+5%-3%=12%

Example 2.9. Event A, B, and C are such that
$P(A)=0.7$, $P(B)=0.6$ $P(C)=0.5$ $P(A\cap B)=0.4$, $P(A\cap C)=0.3$, $P(B\cap C)=0.2$, $P(A\cap B\cap C)=0.1$
Find:

a) $P(\bar{A}\cap B)$
b) $P(\bar{A}\cap \bar{B}\cup C)$

Solution:

a) $P(\bar{A}\cap B)=P(B-A)=P(B)-P(A\cap B)=0.6-0.4=0.2$
b) $P(\bar{A}\cap \bar{B}\cup C)=1-P(A\cup B)+\left[P(A\cap C)+P(B\cap C)-P(A\cap B\cap C)\right]$
$=1-0.9+0.3+0.2-0.1=0.5$

Example 2.10. (Moore, 1960) the chieftain (kabile reisi) of a primitive tribe (ilkel kabile) of Indians (kızıldereliler) had feeling that some of his men were cheating him of tax payments. According to the laws of tribe, a tax had to be paid by any man who owned three pieces of taxable property: teepees (kızıldereli çadırı), horses and squaws (kızıldereli kadın). The tribe practiced monogamy (tek eşlilik), and no man owned more than one horse or one teepee. The chieftain knew that none would admit guilt, and so he decided on trick scheme. He inquired (soruşturmak) of the 2300 men in his tribe how many owned: a horse, a teepee; a squaw; a squaw and a horse; a squaw and teepee; a horse and a teepee. He knew that he could expect honest answer to these questions; the result of which are given in the table below. It was observed that every man admitted ownership of at least one piece of property. How many men should be paying taxes? (Barr and Zehna, 1971, page, 19)

property	Number of men owning property
Horse	800
Teepee	1600
Squaw	1100
Squaw and horse	700
Squaw and teepee	500
Horse and teepee	200

(For student)