Introduction to Statistics (MAT 283)

Dipti Dubey

Department of Mathematics
Shiv Nadar University

Conditional Probability: Let S be a sample space associated with a random experiment. The conditional probability of an event A, given that event B has occurred, is defined by

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

provided $P(B) > 0$.

INDEPENDENT EVENTS: Two events A and B of a sample space S are called independent if and only if

$$P(A \cap B) = P(A)P(B).$$

• Two events A and B are independent if the occurrence or nonoccurrence of either one does not affect the probability of the occurrence of the other.

EXAMPLE: A coin is tossed three times and the eight possible outcomes, $HHH, HHT, HTH, THH, HTT, THT, TTH, \text{and} TTT$, are assumed to be equally likely. If A is the event that a head occurs on each of the first two tosses, B is the event that a tail occurs on the third toss, and C is the event that exactly two tails occur in the three tosses.
We have

$$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

$$A = \{HHH, HHT\}$$

$$B = \{HHT, HTT, THT, TTT\}$$

$$C = \{HTT, THT, TTH\}$$

In this example events A and B are independent; events B and C are dependent; events A and C are dependent;

In this example events A and B are independent; events B and C are dependent; events A and C are dependent;

EXAMPLE: Toss a coin and then independently roll a die. What is the probability of observing heads on the coin and a 2 or 3 on the die?

Solution: Let A denote the event of observing a head on the coin and let B be the event of observing a 2 or 3 on the die. Then
$P(A \cap B) = P(A)P(B) = (1/2)(2/6) = 1/6$

Theorem: If A and B are independent, then A and B^c are also independent.

Theorem: Two possible mutually exclusive (disjoint) events are always dependent.

Proof: Suppose not. Then

$$P(A \cap B) = P(A)P(B)$$

$$P(\emptyset) = P(A)P(B)$$

$$0 = P(A)P(B)$$

Hence, we get either $P(A) = 0$ or $P(B) = 0$. This is a contradiction to the fact that A and B are possible events. This completes the proof.

Theorem: Two possible independent events are not mutually exclusive.
Proof: Let A and B be two independent events and suppose A and B are mutually exclusive. Then

$$P(A)P(B) = P(A \cap B)$$

$$= P(\emptyset)$$

$$= 0.$$

Therefore, we get either $P(A) = 0$ or $P(B) = 0$. This is a contradiction to the fact that A and B are possible events.

There are many situations where the ultimate outcome of an experiment depends on what happens in various intermediate stages. This issue is resolved by the Bayes Theorem.

PARTITION: Let S be a set and let $\mathcal{P} = \{A_i\}_{i=1}^{m}$ be a collection of subsets of S. The collection $\mathcal{P}$ is called a partition of S if

(a) $S = \bigcup_{i=1}^{m} A_i$

(b) $Ai \cap Aj = \emptyset$ for $i \neq j$.

LAW OF TOTAL PROBABILITY: If the events $\{B_i\}_{i=1}^{m}$ constitute a partition of the sample space S and $P(B_i) \neq 0$ for $i = 1, 2, \ldots, m$, then for any event A in S

$$P(A) = \sum_{i=1}^{m} P(B_i) P(A/B_i).$$

LAW OF TOTAL PROBABILITY: If the events $\{B_i\}_{i=1}^m$ constitute a partition of the sample space S and $P(B_i) \neq 0$ for $i = 1, 2, \ldots, m$, then for any event A in S

$$P(A) = \sum_{i=1}^{m} P(B_i)P(A/B_i).$$

Proof: Let S be a sample space and A be an event in S. Let $\{B_i\}_{i=1}^m$ be any partition of S. Then $A = \bigcup_{i=1}^m (A \cap B_i)$. Thus

$$P(A) = P(\cup_{i=1}^m (A \cap B_i))$$

$$= \sum_{i=1}^{m} P(A \cap B_i)$$

$$= \sum_{i=1}^{m} P(B_i)P(A/B_i).$$

BAYES' THEOREM: If the events $\{B_i\}_{i=1}^{m}$ constitute a partition of the sample space S and $P(B_i) \neq 0$ for $i = 1, 2, \ldots, m$, then for any event A in S such that $P(A) \neq 0$,

$$P(B_k/A) = \frac{P(B_k)P(A/B_k)}{\sum_{i=1}^{m} P(B_i)P(A/B_i)}, \ k = 1, 2, \ldots, m.$$

BAYES' THEOREM: If the events $\{B_i\}_{i=1}^{m}$ constitute a partition of the sample space S and $P(B_i) \neq 0$ for $i = 1, 2, \ldots, m$, then for any event A in S such that $P(A) \neq 0$,

$$P(B_k/A) = \frac{P(B_k)P(A/B_k)}{\sum_{i=1}^{m} P(B_i)P(A/B_i)}, \; k = 1, 2, \ldots, m.$$

Proof: Using the definition of conditional probability, we get

$$P(B_k/A) = \frac{P(A \cap B_k)}{P(A)}.$$

Using above Theorem , we get

$$P(B_k/A) = \frac{P(A \cap B_k)}{\sum_{i=1}^{m} P(B_i)P(A/B_i)} = \frac{P(B_k)P(A/B_k)}{\sum_{i=1}^{m} P(B_i)P(A/B_i)}.$$

This completes the proof.

EXAMPLE: Two boxes containing marbles are placed on a table. The boxes are labeled B_1 and B_2.

	No. of green marbles	No. of white marbles
B_1	7	4
B_2	3	10

The boxes are arranged so that the probability of selecting box B_1 is $\frac{1}{3}$ and the probability of selecting box B_2 is $\frac{2}{3}$.

She will win a TV if she selects a green marble.

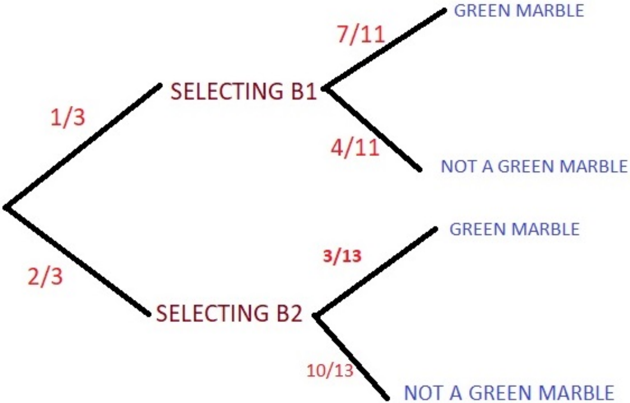
(a) What is the probability that Kathy will win the TV (that is, she will select a green marble)?
(b) If Kathy wins the color TV, what is the probability that the green marble was selected from the first box?

(a) Let A be the event of drawing a green marble (or winning a TV).
A will happen if she selects a green marble from either box B_1 or B_2. Therefore,

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

$$= P(A/B_1)P(B_1) + P(A/B_2)P(B_2)$$
$$= \left(\frac{7}{11}\right)\left(\frac{1}{3}\right) + \left(\frac{3}{11}\right)\left(\frac{2}{3}\right)$$
$$= \frac{157}{429}$$

(b) Given that Kathy won the TV, the probability that the green marble was selected from B_1 is



$$P(B_1/A) = \frac{P(B_1)P(A/B_1)}{P(B_1)P(A/B_1) + P(B_2)P(A/B_2)}$$

$$= \frac{(1/3)\,(7/11)}{(7/11)\,(1/3) + (3/13)\,(2/3)}$$

$$= \frac{91}{157}$$

RANDOM VARIABLE

RANDOM VARIABLE:Consider a random experiment whose sample space is S. A random variable X is a function from the sample space S into the set of real numbers $\mathbb{R}$ such that for each interval I in $\mathbb{R}$, the set $\{s \in S : |X(s) \in I\}$ is an event in S.

$$X : S \to \mathbb{R}$$

$$s \mapsto X(s)$$

RANDOM VARIABLE

RANDOM VARIABLE:Consider a random experiment whose sample space is S. A random variable X is a function from the sample space S into the set of real numbers $\mathbb{R}$ such that for each interval I in $\mathbb{R}$, the set $\{s \in S : |X(s) \in I\}$ is an event in S.

$$X : S \to \mathbb{R}$$
$$s \mapsto X(s)$$

SPACE OF THE VARIABLE: The set $R_X = \{x \in \mathbb{R} | x = X(s),\ s \in S\}$ is called the space of the random variable X.

EXAMPLE:1

RANDOM EXPERIMENT: Tossing a fair coin

$$S = \{H, T\}$$

Define

$$X : S \to \mathbb{R}$$

such that

$$X(H) = 0$$

$$X(T) = 1$$

Then $R_X = \{0, 1\}$.

EXAMPLE:2 Consider an experiment in which a coin is tossed ten times. What is the sample space of this experiment? Define a random variable for this sample space and then find the space of the random variable.

EXAMPLE:2 Consider an experiment in which a coin is tossed ten times. What is the sample space of this experiment? Define a random variable for this sample space and then find the space of the random variable.

The sample space of this experiment is given by

$$S = \{s | s \text{ is a sequence of 10 heads or tails}\}.$$

EXAMPLE:2 Consider an experiment in which a coin is tossed ten times. What is the sample space of this experiment? Define a random variable for this sample space and then find the space of the random variable.

The sample space of this experiment is given by

$$S = \{s | s \text{ is a sequence of 10 heads or tails}\}.$$

Let $X : S \to \mathbb{R}$ be a function from the sample space S into the set of reals $\mathbb{R}$ defined as follows:

EXAMPLE:2 Consider an experiment in which a coin is tossed ten times. What is the sample space of this experiment? Define a random variable for this sample space and then find the space of the random variable.

The sample space of this experiment is given by

$$S = \{s | s \text{ is a sequence of 10 heads or tails}\}.$$

Let $X : S \to \mathbb{R}$ be a function from the sample space S into the set of reals $\mathbb{R}$ defined as follows:

$X(s) =$number of heads in sequence s.

Then X is a random variable. This random variable, for example, maps the sequence HHTTTHTTHH to the real number 5, that is

$$X(HHTTTHTTHH) = 5.$$

The space of this random variable is $R_X = \{0, 1, 2, ..., 10\}$.

EXAMPLE

A balanced coin is tossed four times. List the elements of the sample space that are presumed to be equally likely, as this is what we mean by a coin being balanced, and the corresponding values x of the random variable X, the total number of heads.

Solution

If H and T stand for heads and tails, the results are as shown in the following table:

Element of sample space	*Probability*	x
HHHH	$\frac{1}{16}$	4
HHHT	$\frac{1}{16}$	3
HHTH	$\frac{1}{16}$	3
HTHH	$\frac{1}{16}$	3
THHH	$\frac{1}{16}$	3
HHTT	$\frac{1}{16}$	2
HTHT	$\frac{1}{16}$	2

Element of sample space	*Probability*	x
HTTH	$\frac{1}{16}$	2
THHT	$\frac{1}{16}$	2
THTH	$\frac{1}{16}$	2
TTHH	$\frac{1}{16}$	2
HTTT	$\frac{1}{16}$	1
THTT	$\frac{1}{16}$	1
TTHT	$\frac{1}{16}$	1
TTTH	$\frac{1}{16}$	1
TTTT	$\frac{1}{16}$	0

Thus, we can write $P(X = 3) = \frac{4}{16}$, for example, for the probability of the event that the random variable X will take on the value 3.

Given a random variable X and a set $A \subseteq \mathbb{R}$, define

$$X^{-1}(A) = \{s \in S | X(s) \in A\}$$

and

$$P(X = x) = P(X^{-1}(x)) = P(\{s \in S | X(s) = x\})$$

$$P(X \in A) = P(X^{-1}(A)) = P(\{s \in S | X(s) \in A\})$$

Discrete Random Variable: If the space of random variable X is countable, then X is called a *discrete random variable*.

Discrete Random Variable: If the space of random variable X is countable, then X is called a *discrete random variable*.

The probability measure defined over a discrete sample space automatically provides the probabilities that a random variable will take on any given value within its range.

Discrete Random Variable: If the space of random variable X is countable, then X is called a *discrete random variable*.

The probability measure defined over a discrete sample space automatically provides the probabilities that a random variable will take on any given value within its range.

Probability Mass Function: Let R_X be the space of the random variable X. The function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = P(X = x)$$

is called the probability mass function (pmf) of X.

Theorem: A function can serve as the probability mass function of a discrete random variable X if and only if its values, f_X, satisfy the conditions

1. $f_X(x) \geq 0$ for each value within its domain;

2. $\sum_x f_X(x) = 1$, where the summation extends over all the values within its domain.

EXAMPLE: In a class of 60 students, there are 15 from mathematics department, 25 from economic department, and 20 from computer science department. One student is selected at random. What is the sample space of this experiment? Construct a random variable X for this sample space and then find its space. Further, find the probability mass function of this random variable X.