Probability and Statistics (IT302)

11th Tuesday 2020 10:30AM-11:00AM

Class 5

Random Variable

Definition

A **Random Variable** is a function that assigns a real number to each outcome in the sample space of a random experiment. A **Random Variable** is denoted by an uppercase letter such as $\boldsymbol{X}$. After an experiment is conducted, the measured value of the random variable is denoted by a lowercase letter such as $\boldsymbol{x}$ = 70 milliamperes [1].

Alternate Definition

A **Random Variable** is a function that associates a real number with each element in the Sample Space. Capital letter $\boldsymbol{X}$ denotes a Random Variable and its corresponding small letter, $\boldsymbol{x}$ in this case, for one of its values [2].

Source :

1. Applied Statistics and Probability for Engineers by Douglas C. Montgomery and George C. Runger, Third Edition, PP. 54
2. Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall, PP. 81

Random Variable Contd.

Given an experiment and the corresponding set of possible outcomes (the Sample Space), a Random Variable associates a particular number with each outcome; see below Figure. Referred this number as the **numerical value** or the **experimental value** of the Random Variable. Mathematically, **a Random Variable is a real-valued function of the experimental outcome**.

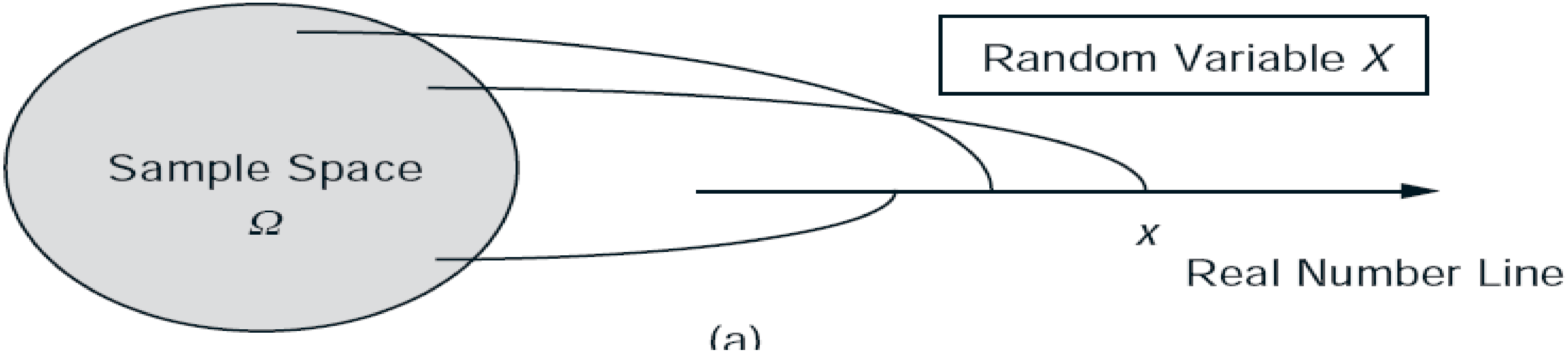


Figure : Visualization of a random variable. It is a function that assigns a numerical value to each possible outcome of the experiment.

Source : Introduction to Probability by Dimitri P. Bertsekas and John N, Chapter 2

Random Variable Example

Example 3.1: Two balls are drawn in succession without replacement from an urn containing 4 red balls and 3 black balls. The possible outcomes and the values y of the random variable Y , where Y is the number of red balls, are

Sample Space	**y**
RR	2
RB	1
BR	1
BB	0

Source :
Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall, PP. 82

Random Variable Example Contd.

Example 3.2: A stockroom clerk returns three safety helmets at random to three steel mill employees who had previously checked them. If Smith, Jones, and Brown, in that order, receive one of the three hats, list the sample points for the possible orders of returning the helmets, and find the value m of the random variable M that represents the number of correct matches.

Solution **:** If S, J, and B stand for Smith's, Jones's, and Brown's helmets, respectively, then the possible arrangements in which the helmets may be returned and the number of correct matches are

Sample Space	m
SJB	3
SBJ	1
BJS	1
JSB	1
JBS	0
BSJ	0

Random Variable Example Contd.

Example 3.3: Consider the simple condition in which components are arriving from the production line and they are stipulated to be defective or not defective. Define the random variable *X* by

$$X = \begin{cases} 1, \text{if the component is defectrive} \\ 0, \text{if the component is not defective} \end{cases}$$

Clearly the assignment of 1 or 0 is arbitrary though quite convenient. The Random Variable for which 0 and 1 are chosen to describe the two possible values is called a **Bernoulli Random Variable**.

Random Variable Example Contd.

Example 3.4: Statisticians use **sampling plans** to either accept or reject batches or lots of material. Suppose one of these sampling plans involves sampling independently 10 items from a lot of 100 items in which 12 are defective.

Let X be the Random Variable defined as the number of items found defective in the sample of 10. In this case, the random variable takes on the values 0, 1, 2, . . . , 9, 10.

Example 3.5: Suppose a sampling plan involves sampling items from a process until a defective is observed. The evaluation of the process will depend on how many consecutive items are observed.

In that regard, let X be a Random Variable defined by the number of items observed before a defective is found. With N a non-defective and D a defective, sample spaces are $S = \{D\}$ given $X = 1$, $S = \{ND\}$ given $X = 2$, $S = \{NND\}$ given $X = 3$, and so on.

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall, PP. 83

Random Variable Example Contd.

Example 3.6: Interest centers around the proportion of people who respond to a certain mail order solicitation. Let X be that proportion. X is a Random Variable that takes on all values x for which $0 \leq x \leq 1$.

Example 3.7: Let X be the Random Variable defined by the waiting time, in hours, between successive speeders spotted by a radar unit. The Random Variable X takes on all values x for which $x \geq 0$.

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall, PP. 83

Random Variable Example Contd.

Example : In an experiment involving a sequence of 5 tosses of a coin, **the number of heads in the sequence is a Random Variable**. However, the 5-long sequence of heads and tails is not considered a random variable because it does not have an explicit numerical value.

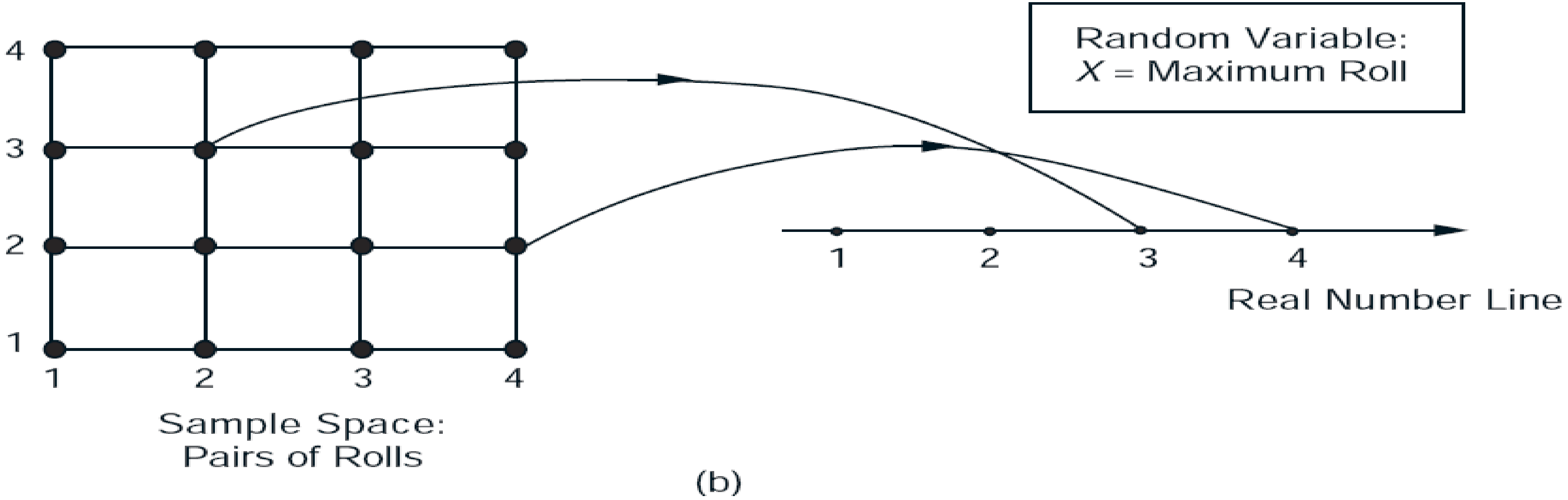
In an experiment involving two rolls of a die, the following are examples of Random Variables:
(1) The sum of the two rolls.
(2) The number of sixes in the two rolls.
(3) The second roll raised to the fifth power.

In an experiment involving the transmission of a message, the time needed to transmit the message, the number of symbols received in error, and the delay with which the message is received are all random variables.

Source : Introduction to Probability by Dimitri P. Bertsekas and John N, Chapter 2

Random Variable Example Contd.

The experiment consists of two rolls of a 4-sided die, and the random variable is the maximum of the two rolls. If the outcome of the experiment is (4, 2), the experimental value of this random variable is 4.



Source : Introduction to Probability by Dimitri P. Bertsekas and John N, Chapter 2

Main Concepts Related to Random Variable

Starting with a probabilistic model of an experiment:

- A **Random Variable** is a real-valued function of the outcome of the experiment.
- A **function of a Random Variable** defines another random variable.
- We can associate with each Random Variable certain “averages” of interest, such the **mean** and the **variance**.
- A **Random Variable** can be **conditioned** on an event or on another Random Variable.
- There is a notion of **independence** of a Random Variable from an event or from another Random Variable.

Source : Introduction to Probability by Dimitri P. Bertsekas and John N, Chapter 2

Continuous and Discrete Random Variable

If a sample space contains a finite number of possibilities or an unending sequence with as many elements as there are whole numbers, it is called a **Discrete Sample Space**.

If a sample space contains an infinite number of possibilities equal to the number of points on a line segment, it is called a **Continuous Sample Space**.

A random variable is called a **Discrete Random Variable** if its set of possible outcomes is countable. But a random variable whose set of possible values is an entire interval of numbers is not discrete.

When a random variable can take on values on a continuous scale, it is called a **Continuous Random Variable**. Often the possible values of a continuous random variable are precisely the same values that are contained in the continuous sample space.

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall, PP. 83-84

Continuous Random Variable

Sometimes a measurement (such as current in a copper wire or length of a machined part) can assume any value in an interval of real numbers (at least theoretically). Then arbitrary precision in the measurement is possible. Of course, in practice, we might round off to the nearest tenth or hundredth of a unit. The Random Variable that represents this measurement is said to be a **Continuous Random Variable**. The range of the Random Variable includes all values in an interval of real numbers; that is, the range can be thought of as a continuum.

A **Continuous Random Variable** is a Random Variable with an interval (either finite or infinite) of real numbers for its range.

Examples of **Continuous Random Variables**: Electrical Current, Length, Pressure, Temperature, Time, Voltage, Weight

Source : Applied Statistics and Probability for Engineers by Douglas C. Montgomery and George C. Runger, Third Edition, PP. 54

Discrete Random Variable

In other experiments, we might record a count such as the number of transmitted bits that are received in error. Then the measurement is limited to integers. Or we might record that a proportion such as 0.0042 of the 10,000 transmitted bits were received in error. Then the measurement is fractional, but it is still limited to discrete points on the real line. Whenever the measurement is limited to discrete points on the real line, the **Random Variable** is said to be a **Discrete Random Variable**.

A **Discrete Random Variable** is a **Random Variable** with a finite (or countably infinite) range.

Examples of **Discrete Random Variables**: Number of Scratches on a surface, Proportion of Defective parts among 1000 tested, Number of transmitted bits received in error

Source : Applied Statistics and Probability for Engineers by Douglas C. Montgomery and George C. Runger, Third Edition, PP. 54

Discrete Random Variable

The Random Variable X that counts the number of heads in three tosses of a coin.

$X(TTT) = 0$	$X(HHT) = 2$	$X(TTH) = 1$	$X(HTH) = 2$
$X(THT) = 1$	$X(THH) = 2$	$X(HTT) = 1$	$X(HHH) = 3$

Definition 3.1 *A random variable X is a real valued function defined on a sample space S. The value of the function at each sample point is denoted by* $X(s)$. *The set of values* $\{X(s) : s \in S\}$ *is called the* range *and is denoted* R_X.

The random variable X has range $R_X = \{0, 1, 2, 3\}$. It is an example of a *Discrete Random Variable*; so called because its range is a discrete set. In the general case we say that the Random Variable X is discrete if its range is the discrete set $\{x_1, x, \ldots, x_n, \ldots\}$. Because it is convenient to do so we shall assume that the numbers in the range appear in increasing order, so $R_X = \{x_1 < x_2 < \ldots < x_n < \ldots\}$

Source : Introduction to Probability and Statistics for Science, Engineering, and Finance by Walter A. Rosenkrantz, PP. 114

Exercises

Question 3.1 Classify the following Random Variables as discrete or continuous:

Question	Discreet/ Continuous
X: the number of automobile accidents per year in Virginia.	Discreet
Y: the length of time to play 18 holes of golf.	Continuous
M: the amount of milk produced yearly by a particular cow.	Continuous
N: the number of eggs laid each month by a hen.	Discreet
P: the number of building permits issued each month in a certain city.	Discreet
Q: the weight of grain produced per acre.	Continuous

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall, PP. 91

Exercises

Question No 3.2 : An overseas shipment of 5 foreign automobiles contains 2 that have slight paint blemishes. If an agency receives 3 of these automobiles at random, list the elements of the Sample Space S, using the letters B and N for blemished and non-blemished, respectively; then to each sample point assign a value x of the Random Variable X representing the number of automobiles with paint blemishes purchased by the agency.

Solution: A table of Sample Space and assigned values of the Random Variable is

Sample Space	x
NNN	0
NNB	1
NBN	1
BNN	1
NBB	2
BNB	2
BBN	2
BBB	3

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall, PP. 91

Exercises

Question No 3.3 : Let W be a random variable giving the number of heads minus the number of tails in three tosses of a coin. List the elements of the sample space S for the three tosses of the coin and to each sample point assign a value w of W.

Solution: A table of Sample Space and assigned values of the Random Variable is

Sample Space	w
HHH	3
HHT	1
HTH	1
THH	1
HTT	−1
THT	−1
TTH	−1
TTT	−3

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall, PP. 91

Exercises

Question No 3.4 : **A** coin is flipped until 3 heads in succession occur. List only those elements of the sample space that require 6 or less tosses. Is this a discrete sample space?

Solution: S = {HHH, THHH, HTHHH, TTHHH, TTTHHH, HTTHHH, THTHHH, HHTHHH, . . . }; The sample space is discrete containing as many elements as there are positive integers.

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall, PP. 91

Additional Slides for Random Variable

When a random experiment is performed, we are often not interested in all of the details of the experimental result but only in the value of some numerical quantity determined by the result. For instance, in tossing dice we are often interested in the sum of the two dice and are not really concerned about the values of the individual dice. That is, we may be interested in knowing that the sum is 7 and not be concerned over whether the actual outcome was (1, 6) or (2, 5) or (3, 4) or (4, 3) or (5, 2) or (6, 1).

Also, a civil engineer may not be directly concerned with the daily risings and declines of the water level of a reservoir (which we can take as the experimental result) but may only care about the level at the end of a rainy season. These quantities of interest that are determined by the result of the experiment are known as random variables.

Since the value of a random variable is determined by the outcome of the experiment, we may assign probabilities of its possible values.

Source : Introduction to Probability and Statistics for Engineers and Scientists by Sheldon M. Ross, Third Edition, PP. 89

Additional Slides for Random Variable

EXAMPLE 4.1a Letting X denote the random variable that is defined as the sum of two fair dice, then

$P\{X = 2\} = P\{(1, 1)\} = 1/36$ Equation 4.1.1

$P\{X = 3\} = P\{(1, 2), (2, 1)\} = 2/36$

$P\{X = 4\} = P\{(1, 3), (2, 2), (3, 1)\} = 3/36$

$P\{X = 5\} = P\{(1, 4), (2, 3), (3, 2), (4, 1)\} = 4/36$

$P\{X = 6\} = P\{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\} = 5/36$

$P\{X = 7\} = P\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\} = 6/36$

Source : Introduction to Probability and Statistics for Engineers and Scientists by Sheldon M. Ross, Third Edition, PP. 89

Additional Slides for Random Variable (Example 4.1a Contd.)

$P\{X = 8\} = P\{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\} = 5/36$

$P\{X = 9\} = P\{(3, 6), (4, 5), (5, 4), (6, 3)\} = 4/36$

$P\{X = 10\} = P\{(4, 6), (5, 5), (6, 4)\} = 3/36$

$P\{X = 11\} = P\{(5, 6), (6, 5)\} = 2/36$

$P\{X = 12\} = P\{(6, 6)\} = 1/36$

In other words, the random variable X can take on any integral value between 2 and 12 and the probability that it takes on each value is given by Equation 4.1.1.

Another random variable of possible interest in this experiment is the value of the first die. Letting Y denote this random variable, then Y is equally likely to take on any of the values 1 through 6. That is, $P\{Y = i\} = 1/6, \quad i = 1, 2, 3, 4, 5, 6$

Source : Introduction to Probability and Statistics for Engineers and Scientists by Sheldon M. Ross, Third Edition,. PP. 90