Chapter 1
Part 1

Random Variable
&
Special Probability Distribution

Illustration

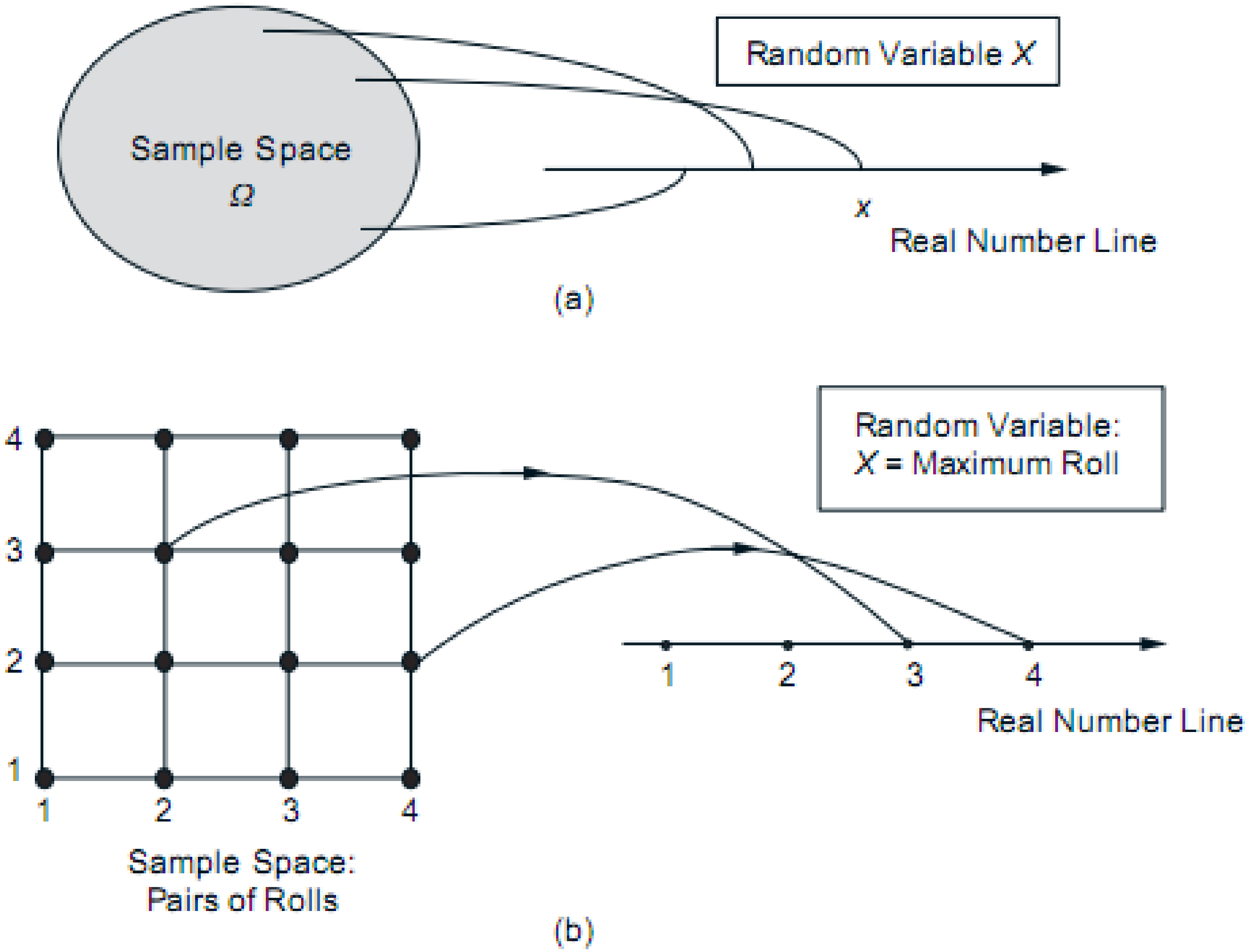


Figure 2.1: (a) Visualization of a random variable. It is a function that assigns a numerical value to each possible outcome of the experiment. (b) An example of a random variable. The experiment consists of two rolls of a 4-sided die, and the random variable is the maximum of the two rolls. If the outcome of the experiment is $(4, 2)$, the experimental value of this random variable is 4.

Definition

R.V say X is a function defined over a sample space S, that associates a real number, X(e)=x, with each possible outcome e in S

A ***random variable*** is a function or rule that assigns a number to each outcome of an experiment. Basically it is just a symbol that represents the outcome of an experiment

Example

$$\max(1,1) = 1, \max(2,2) = 2, \max(3,2) = 3, \max(4,3) = 4$$

each of the events B_1, B_2, B_3, B_4 of S contain

the pairs (i, j) other word X has value

$$x = 1 \text{ over } B_1, x = 2 \text{ over } B_2, x = 3 \text{ over } B_3, x = 4 \text{ over } B_4$$

example…

- An experiments involving a sequence of 5 tosses of a coin, the number of Heads in the sequence is a random variable
- What of X value, if X={sum of H}

- X = number of heads when the experiment is flipping a coin 20 times.
- C = the daily change in a stock price.
- R = the number of miles per gallon you get on your auto during a family vacation.
- Y = the amount of medication in a blood pressure pill.
- V = the speed of an auto registered on a radar detector used on I-20

Two Types of Random Variables…

- **Discrete Random Variable** – usually count data [Number of]

one that takes on a ***countable*** number of values – this means you can list **all** possible outcomes without missing any, although it might take you an infinite amount of time.

- X = values on the roll of two dice: X has to be either 2, 3, 4, …12.
- Y = number of accidents on the UNS campus during a week: Y has to be 0, 1, 2, 3, 4, 5, 6, 7, 8, ……………

- **Continuous Random Variable** – usually measurement data [time, weight, distance, etc] * one that takes on an uncountable number of values – this means you can never list all possible outcomes even if you had an infinite amount of time.
- X = time it takes you to drive home from class: X > 0, might be 30.1 minutes measured to the nearest tenth but in reality the actual time is 30.10000001…………………. minutes?)
- Exercise:

try to list all possible numbers between 0 and 1.

Probability Distributions…

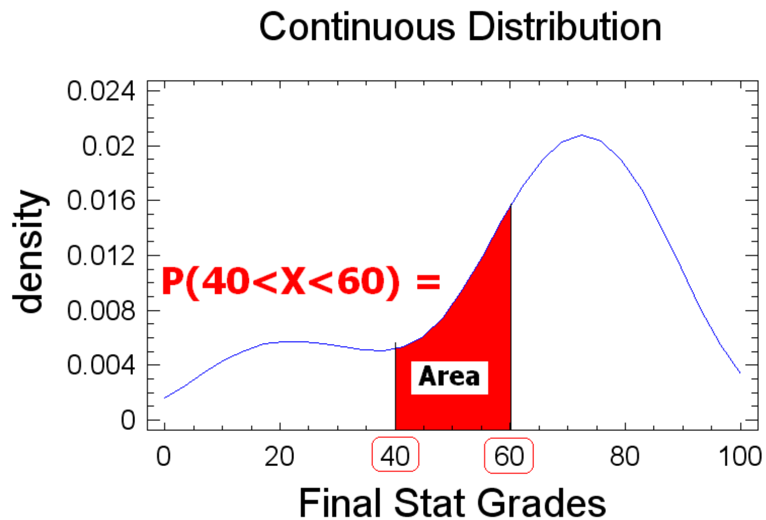
- A ***probability distribution (density function)*** is a table, formula, or graph that describes the values of a random variable and the probability associated with these values.

– Discrete Probability Distribution, notation P(X=x)

example, X = outcome of rolling one die

X	1	2	3	4	5	6
P(X)	1/6	1/6	1/6	1/6	1/6	1/6

– Continuous Probability Distribution



Def. Bain : 1997

If the set of all possible values of a random variable, X, is a countable set, x1, x2, x3,… then X is called a discrete random variable. The function :
f(x)=P[X=x], x=x1, x2,…
That assigns the probability to each possible value x will be called the discrete Probability density function (discrete pdf)

theorem

For a discrete random variable X with possible values $x_1, x_2, \ldots, x_n$, a **probability mass function** is a function such that

(1) $f(x_i) \geq 0$

(2) $\sum_{i=1}^{n} f(x_i) = 1$

(3) $f(x_i) = P(X = x_i)$ (3-1)

Example 1

1. The experiment consist of two independent tosses of a fair coin, let X be the number of heads obtained, then the pdf/pmf of X is :

$$f(x) = \begin{cases} \frac{1}{4}, \text{if } x = 0 \text{ or } x = 2 \\ \frac{1}{2}, \text{if } x = 1 \\ 0, \text{otherwise} \end{cases}$$

2. If $f(x) = c(2x-1), \ x = 1, 2, \ldots, 12$

   then find c !

Cummulative Density Function

Definition

The **cumulative distribution function** of a discrete random variable X, denoted as $F(x)$, is

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

For a discrete random variable X, $F(x)$ satisfies the following properties.

(1) $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$

(2) $0 \leq F(x) \leq 1$

(3) If $x \leq y$, then $F(x) \leq F(y)$ (3-2)

Theorem

A function F(x) is a CDF for some R.V X if and only if it satisfies the following properties :

1. $\lim_{x \to -\infty} F(x) = 0$
2. $\lim_{x \to \infty} F(x) = 1$
3. $\lim_{h \to 0^+} F\left(x+h\right) = F(x)$
4. $a < b$ implies $F(a) \leq F(b)$

Definition

A RV X is called a continuous RV if there is a function f(x) called the probability density function (pdf) of X such that the CDF can be represented as

$$F(x) = \int_{-\infty}^{x} f(t)\, dt$$

Properties :

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x)\, dx = 1$

Continuous Probability Distribution

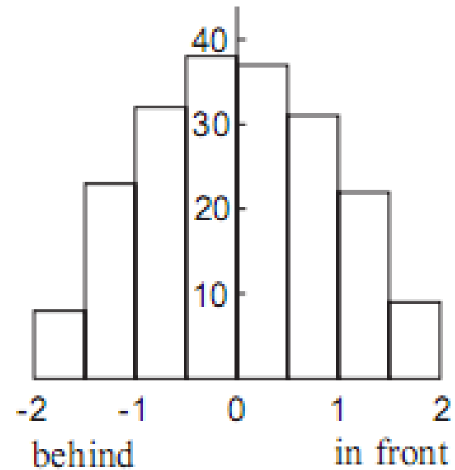
Ex. In a similar experiment shot-putter were asked to aim at a line 10 m away. They threw the shot 200 times and throws were measured within 2m either side of the line. The result were as shown below.

in front

metres	1.99-1.50	1.49-1.00	0.99-0.5	0.49-0
frequency	9	22	31	37

behind

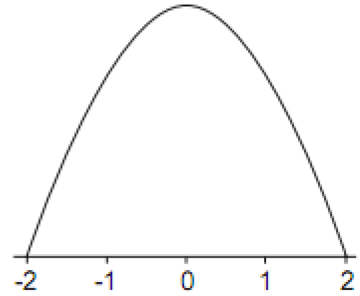
metres	0-0.49	0.50-0.99	1.00-1.49	1.50-1.99
frequency	38	32	23	8



Unless u know the total sample size, u cannt put a scale on the y-axis

We ‘ll find the function of x

Since it crosses the x-axis at ±2, these are roots of the equation. The function therefore of the form $f(x) = A\left(4 - x^2\right)$ with A=constant



a. Find A
b. Find f(x)

solution

a. $$\int_{-2}^{2} A\left(4-x^2\right)dx = A\left[4x-\frac{x^3}{3}\right]_{-2}^{2}$$

$$= A\left\{\left[8-\frac{8}{3}\right]-\left[-8-\frac{-8}{3}\right]\right\}$$

$$= \frac{32A}{3}.$$

This must equal 1, so A must take the value $\frac{3}{32}$.

b. Note that this only works if the range of answers is restricted to -2 to $+2$. This is usually made clear by defining a probability density function (p.d.f.) as follows:

$$f(x) = \begin{cases} \frac{3}{32}\left(4-x^2\right) & \text{for } -2 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

Example 1 Th. 2.2.2 Bain : 59

$$F(x)=\begin{cases}0, x<-2\\0.2, -2\le x<0\\0.7, 0\le x<2\\1, x\ge 2\end{cases}$$

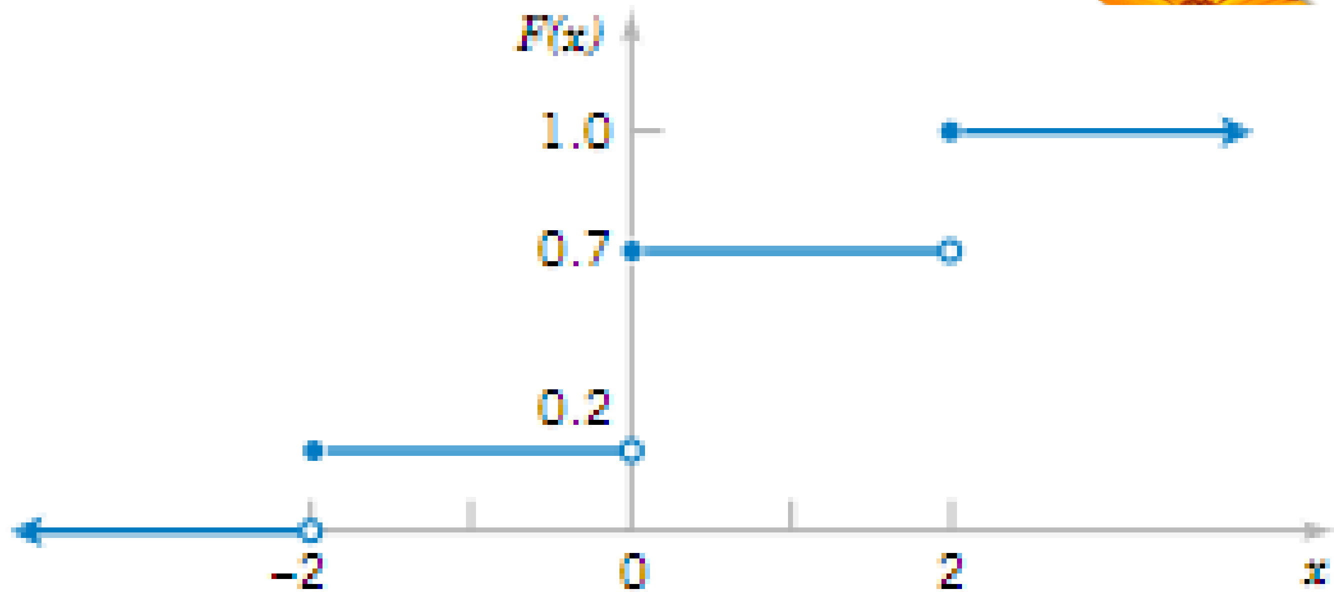


Figure 3-3 displays a plot of $F(x)$. From the plot, the only points that receive nonzero probability are −2, 0, and 2. The probability mass function at each point is the change in the cumulative distribution function at the point. Therefore,

$$f(-2)=0.2-0=0.2 \qquad f(0)=0.7-0.2=0.5 \qquad f(2)=1.0-0.7=0.3$$

EXERCISE

3-13. The sample space of a random experiment is $\{a, b, c, d, e, f\}$, and each outcome is equally likely. A random variable is defined as follows:

outcome	a	b	c	d	e	f
x	0	0	1.5	1.5	2	3

Determine the probability mass function of X.

3-14. Use the probability mass function in Exercise 3-11 to determine the following probabilities:

(a) $P(X = 1.5)$ (b) $P(0.5 < X < 2.7)$
(c) $P(X > 3)$ (d) $P(0 \le X < 2)$
(e) $P(X = 0 \quad \text{or} \quad X = 2)$

Verify that the following functions are probability mass functions, and determine the requested probabilities.

3-15.

x	-2	-1	0	1	2
$f(x)$	1/8	2/8	2/8	2/8	1/8

(a) $P(X \le 2)$ (b) $P(X > -2)$
(c) $P(-1 \le X \le 1)$ (d) $P(X \le -1 \quad \text{or} \quad X = 2)$

3-16. $f(x) = (8/7)(1/2)^x, \quad x = 1, 2, 3$

(a) $P(X \le 1)$ (b) $P(X > 1)$
(c) $P(2 < X < 6)$ (d) $P(X \le 1 \quad \text{or} \quad X > 1)$

Exercise

3-33.

$$F(x) = \begin{cases} 0 & x < 1 \\ 0.5 & 1 \le x < 3 \\ 1 & 3 \le x \end{cases}$$

(a) $P(X \le 3)$ (b) $P(X \le 2)$
(c) $P(1 \le X \le 2)$ (d) $P(X > 2)$

3-34. Errors in an experimental transmission channel are found when the transmission is checked by a certifier that detects missing pulses. The number of errors found in an eight-bit byte is a random variable with the following distribution:

$$F(x) = \begin{cases} 0 & x < 1 \\ 0.7 & 1 \le x < 4 \\ 0.9 & 4 \le x < 7 \\ 1 & 7 \le x \end{cases}$$

Determine each of the following probabilities:
(a) $P(X \le 4)$ (b) $P(X > 7)$
(c) $P(X \le 5)$ (d) $P(X > 4)$
(e) $P(X \le 2)$

3-35.

$$F(x) = \begin{cases} 0 & x < -10 \\ 0.25 & -10 \le x < 30 \\ 0.75 & 30 \le x < 50 \\ 1 & 50 \le x \end{cases}$$

(a) $P(X \le 50)$ (b) $P(X \le 40)$
(c) $P(40 \le X \le 60)$ (d) $P(X < 0)$
(e) $P(0 \le X < 10)$ (f) $P(-10 < X < 10)$

3-36. The thickness of wood paneling (in inches) that a customer orders is a random variable with the following cumulative distribution function:

$$F(x) = \begin{cases} 0 & x < 1/8 \\ 0.2 & 1/8 \le x < 1/4 \\ 0.9 & 1/4 \le x < 3/8 \\ 1 & 3/8 \le x \end{cases}$$

Determine the following probabilities:
(a) $P(X \le 1/18)$ (b) $P(X \le 1/4)$
(c) $P(X \le 5/16)$ (d) $P(X > 1/4)$
(e) $P(X \le 1/2)$