III. PROBABILITY: 2

CREDITS (30 HOURS); L, T, SPW

COMBINATORY ANALYSIS

SECOND SEMESTER: 2020

KEPHIPS

COURSE INSTRUCTOR:

JAFF LAWRENCE ASUIYI

678179539

COURSE OUTLINE

III.0 Basic Concepts

III.1 Calculation of probabilities

- **1.1 Kolmogorov axioms**
- **1.2 Conditional and independent probabilities**
- **1.3 Bayes theorem and axiom on total probability**

III. 2. Random variables

- **2.1. Definition**
- **2. 2. Moment of a random variable**
- **2. 3. Joint law and marginal laws of a pair**
- **2. 3. Bienaymé-Tchebychev Inequality**
- **2. 4. Basic laws on large numbers**
- **2. 5. TCL**

III. 3. Probability Laws

References:

1.Prem Mann, Introductory Statistics, 8/E. John Wiley & Sons 2013

2.Advanced Modern Engineering Mathematics Fourth Edition Glyn James 2011

3.Advanced Engineering Mathematics ErwinKreyszig. 10th Edition 2011

4. Schaum's Outline of Probability

III.0 BASIC CONCEPTS

Probability is a scientific measure of the likelihood or degree of certainty that something will happen by chance.

TRIAL OR EXPERIMENT, OUTCOME, SAMPLE SPACE

0.1: **A trial or an experiment** is an activity with observable results. E.g.

- Playing a lucky game.
- Rolling a die
- Observing the duration of fully charged battery
- Selecting a production component.

0.2. **An outcome** is the result of an experiment. E.g.

- The appearance of a 5 in the throw of a die.
- The sex of a new born baby
- A production component is defective or not defective.

0.3**. The sample space, possibility space S** is the set of all possible outcomes in a given experiment.

Table 0.1 Examples of Experiments, Outcomes, and Sample Spaces

Experiment	Outcomes	Sample Space
Toss a coin once	Head, Tail	$S = \{\text{Head, Tail}\}$
Roll a die once	1, 2, 3, 4, 5, 6	$S = \{1, 2, 3, 4, 5, 6\}$
Toss a coin twice	*HH*, *HT*, *TH*, *TT*	$S = \{HH, HT, TH, TT\}$
Play lottery	Win, Lose	$S = \{\text{Win, Lose}\}$
Take a test	Pass, Fail	$S = \{\text{Pass, Fail}\}$
Select a worker	Male, Female	$S = \{\text{Male, Female}\}$
Select a component	Defective, Good	$S = \{Defective, Good\}$

0.4 AN EVENT

An event, A is a subset of the sample space, S i.e. $A \subset S$. It is denoted by a capital letter. E.g. If E is the event of getting an even number when a die is rolled then $E = \{2, 4, 6\}$

Events obtained by using set operations on events in S

- ❖ If A and B are events then:

1. $A \cup B$ is the event "either A or B or both." $A \cup B$ is called the *union* of A and B.
2. $A \cap B$ is the event "both A and B." $A \cap B$ is called the *intersection* of A and B.
3. A' is the event "not A." A' is called the *complement* of A.
4. $A - B = A \cap B'$ is the event "A but not B." In particular, $A' = S - A$.

- ❖ If $A \cap B = \emptyset$ then A and B are disjoint or mutually exclusive events e.g. A and A'

Example: A coin is tossed twice, let A be the event "at least one head occurs" and B the event "the second toss is a tail". List the elements of

$$\text{S, A, B}, A \cup B, A \cap B, A', and\ A - B.$$

Solution: $S = \{TT, HT, TH, HH\}$, $A = \{TH, HT, HH\}$, $B = \{TT, HT\}$

$A \cup B = \{TT, TH, HT, HH\}$ $A \cap B = \{HT\}$ $A' = \{TT\}$

$$A - B = \{TH, HH\}$$

0.5 Random or fair experimen

A random or fair trial is an experiment whose outcome is determined by chance. Thus, all the outcomes are **equally likely or equiprobable** i.e. they have equal chances of occurring.

E.g.: tossing a fair coin or fair die

III.1 Calculation of probabilities

The probability of an event A is

$$P(A) = \frac{\text{Number of ways in which A can occur}}{\text{Number of ways in which the sample space, S can occur}}$$

$$\Rightarrow \mathrm{P}(\mathrm{A}) = \frac{n(A)}{n(S)} \quad \text{or} \quad \mathrm{P}(\mathrm{A}) = \frac{f_A}{\Sigma f}$$

$n(A) =$ Number of elements in A or number of ways in which A can occur.

$\mathrm{n}(\mathrm{S}) =$ Number of elements in S or number of ways in which S can occur.

$f_A =$ frequency of A, $\Sigma f =$ sum of all the frequencies

Example 0.1

Find the probability of getting an odd number when a die is rolled.

Solution

$$Sample\ space, S = \{1, 2, 3, 4, 5, 6\} \Longrightarrow n(S) = 6$$

$$O = Odd\ number \Longrightarrow O = \{1,3,5\} \Longrightarrow n(O) = 3$$

$$P(odd\ number) = P(O) = \frac{n(O)}{n(S)} = \frac{3}{6} = \frac{1}{2}$$

Exercise 0.1

A pen is drawn at random from a box which contains 3 blue, 2 black and 5 red identical pens.

a) Find the probability that it is blue.

b) Find the probability that it is black or red.

Solution

a) $$P(Blue) = \frac{n(Blue)}{n(S)} = \frac{3}{10}$$

b) $$P(Black\ or\ Red) = \frac{n(Black\ or\ red)}{n(S)} = \frac{7}{10}$$

Expectation of an event

The expectation of an event A happening in n trials is denoted by $E(A)$ and defined by

$$E(A) = n \times P(A)$$

Thus expectation is the average occurrence of an event.

Example

Find the expectation of obtaining a 4 when a fair die is rolled 3 times.

Solution

Probability of getting a 4 in one trial: $P(4) = \frac{1}{6}$

Number of trials: $n = 3$

Expected number of 4's: $E(4) = 3 \times P(4) = 3 \times \frac{1}{6} = \frac{1}{2}$

1.1.0 Probability scale

If $A \subset S$ then $0 \leq P(A) \leq 1$

Thus probability is a function

$$P: \{all\ events\} \rightarrow [0,1]$$

$P(A) = 0 \Rightarrow$ A is an impossible event i.e. $A = \phi$

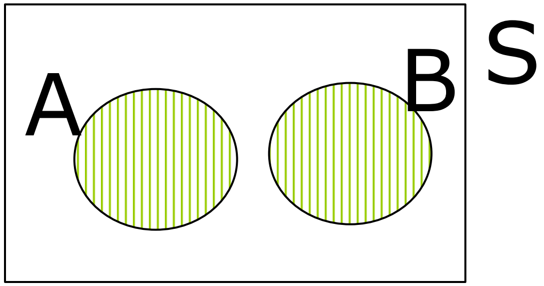
$P(A) = 1 \Rightarrow$ $A = S$ or A is certain to occur.

1.1.1 Kolmogorov's Axioms of probability

Axiom 1: For any $event\ A,\ 0 \le P(A) \le 1$

Axiom 2: $If\ S\ is\ the\ sample\ space\ then\ P(S) = 1$

Axiom 3: If A and B are disjoint or mutually exclusive,



$\mathrm{A} \cap \mathrm{B} = \emptyset \Longrightarrow P(A \cap B) = 0$ *then,*

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

Generally if $A_1, A_2, \ldots, A_n$ are disjoin events then

$$P(A_1 \cup A_2 \cup \ldots \cup\ A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

i.e. The probability of a union of events that cannot occur is the sum of their individual probabilities

Example 1.1

A fair die is thrown. If

E is the event of getting an even number and

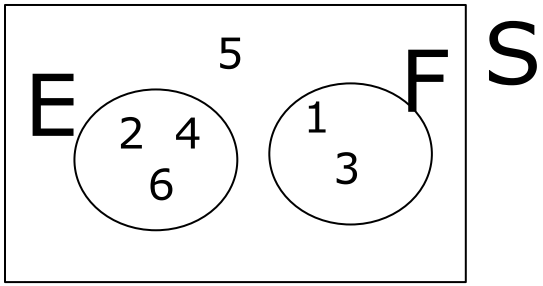
F is the event of getting a factor of 15

a) List the elements of E and F
b) Represent these events on a Venn diagram
c) Find P(E), P(F) and P($(E \cup F)$

Solution

Sample space $S = \{1, 2, 3, 4, 5, 6\}$

$a)\ E = \{2, , 4, 6\} \quad F = \{1, 3\}$



c) $P(E) = \frac{n(E)}{n(S)} = \frac{3}{6} = \frac{1}{2}$

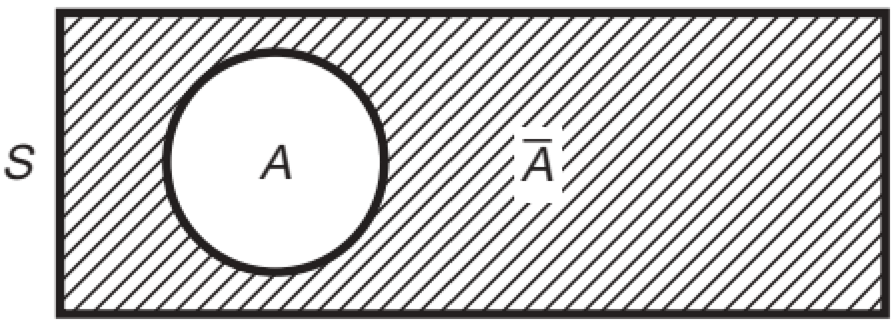
$P(F) = \frac{n(F)}{n(S)} = \frac{2}{6} = \frac{1}{3}$

$P(E \cup F) = P(E) + P(F)$ (disjoint events)

$= \quad \frac{1}{2} \quad + \quad \frac{1}{3} \quad = \frac{5}{6}$

1.1.2 Complementary Events

The complement of an even A denoted by A′ or A^c or $\bar{A}$ or $\underline{A}$ is the event that A does not occur.



$$\boxed{P\left(\overline{A}\right)=1-P\left(A\right) \text{ or } P\left(A\right)+P\left(\overline{A}\right)=1}$$

NB: The phrase "at least" is often used to introduce complementary events.

E.g. At least one red is the complement of no red

$$\Rightarrow P\left(\text{at least 1 red}\right)=1-P\left(\text{no red}\right)$$

Example

A sample of four electronic components is taken from the output of a production line. The probabilities of the various outcomes are calculated to be: Pr [0 defectives] = 0.6561, Pr [1 defective] = 0.2916, Pr [2 defectives] = 0.0486, Pr [3 defectives] = 0.0036, Pr [4 defectives] = 0.0001. What is the probability of at least one defective?

Solution

- It would be perfectly correct to calculate as follows:

$$\Pr[\text{at least one defective}] = \Pr[1\text{ defective}] + \Pr[2\text{ defectives}] + \Pr[3\text{ defectives}] + \Pr[4\text{ defectives}]$$

$$= 0.2916 + 0.0486 + 0.0036 + 0.0001 = 0.3439.$$

- But it would be easier to calculate instead:

$$\begin{aligned}\Pr[\text{at least one defective}] &= 1 - \Pr[0\text{ defectives}]\\ &= 1 - 0.6561\\ &= 0.3439 \text{ or } 0.344.\end{aligned}$$

Example

The probability that a software obtained from a particular technician will be compatible with Windows 10 is 13/16.

Find the probability that a software obtained from that technician will not be compatible with Windows 10.

Solution

Let C =Software is compatible.

$$P(C) = \frac{13}{16}$$

$$P(Software\ is\ not\ compatible)\ is$$

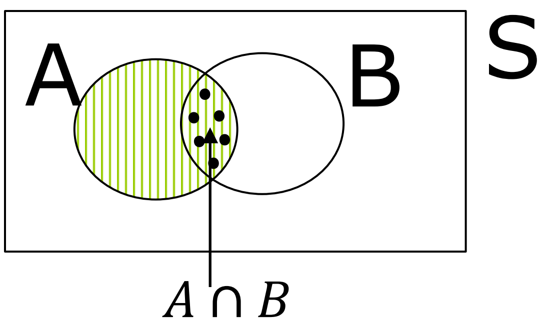
$$P(\bar{C}) = 1 - P(C) = 1 - \frac{13}{16} = \frac{3}{16}$$

1.2 Conditional and independent probabilities

1.21 **CONDITIONAL EVENTS**

Given two events A and B, the probability of A occurring when B has already occurred is called the conditional probability of A given B.

The probability of A given B is



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

Example 1.2

A fair die is thrown. If:

A is the event of getting an even number and

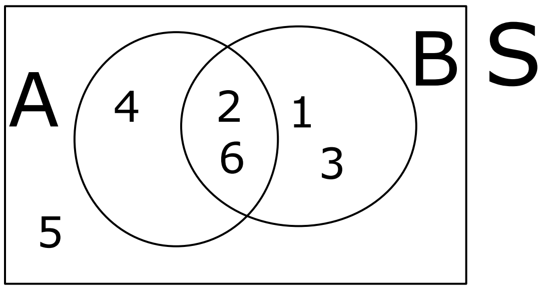
B is the event of getting a factor of 6

a) Represent these events A and B on a venn diagram
b) Find P(A), P(B) and P(A/B)

Solution

Sample space $S = \{1, 2, 3, 4, 5, 6\}$

a) A$= \{2, 4, 6\}$ B$= \{1, 2, 3, 6\}$ $A \cap B = \{2, 6\}$



c) $P(A) = \frac{n(A)}{n(S)} = \frac{3}{6} = \frac{1}{2}$

$$P(B) = \frac{n(B)}{n(S)} = \frac{4}{6} = \frac{2}{3}$$

$$P(A/B) = \frac{n(A \cap B)}{n(B)} = \frac{2}{4} = \frac{1}{2}$$

Using the formula: $P(A \cap B) = \frac{n(A \cap B)}{n(s)} = \frac{2}{6} = \frac{1}{3}$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{1/3}{2/3} = \frac{1}{2}$$

1.21 Tree diagram

A tree diagram is a diagrammatic application of conditional events.

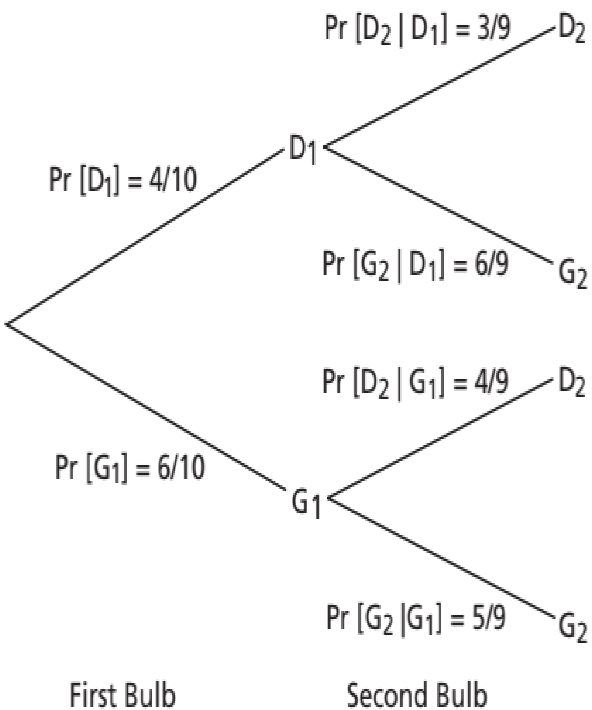
- Number of branches from each point equals the number of outcomes under consideration.
- Sum of probabilities on branches from a point equals 1
- Probabilities at the end of each branch is the product of probabilities on the branch.

Example

Four of the light bulbs in a box of ten bulbs are burnt out or otherwise defective. If two bulbs are selected at random without replacement and tested, (i) what is the probability that exactly one defective bulb is found? (ii) What is the probability that exactly two defective bulbs are found?

Solution Let us use the symbols D_1 for a defective first bulb, D_2 for a defective second bulb, G_1 for a good first bulb, and G_2 for a good second bulb.

TREE DIAGRAM



Event	Probability
2 defective bulbs	$\Pr[D_1 \cap D_2] = \Pr[D_1] \times \Pr[D_2 \vert D_1] = \left(\frac{4}{10}\right)\left(\frac{3}{9}\right) = \frac{12}{90}$
1 good, 1 defective	$\Pr[D_1 \cap G_2] = \Pr[D_1] \times \Pr[G_2 \vert D_1] = \left(\frac{4}{10}\right)\left(\frac{6}{9}\right) = \frac{24}{90}$
1 good, 1 defective	$\Pr[G_1 \cap D_2] = \Pr[G_1] \times \Pr[D_2 \vert G_1] = \left(\frac{6}{10}\right)\left(\frac{4}{9}\right) = \frac{24}{90}$
2 good bulbs	$\Pr[G_1 \cap G_2] = \Pr[G_1] \times \Pr[G_2 \vert G_1] = \left(\frac{6}{10}\right)\left(\frac{5}{9}\right) = \frac{30}{90}$

Solution continues

i) Pr [exactly one defective bulb is found] = Pr $[D_1 \cap G_2]$ + Pr $[G_1 \cap D_2]$

$$= \frac{24 + 24}{90} = \frac{48}{90} = 0.533.$$

The first term corresponds to getting first a defective bulb and then a good bulb, and the second term corresponds to getting first a good bulb and then a defective bulb.

ii) Pr [exactly two defective bulbs are found] = Pr $[D_1 \cap D_2] = \frac{12}{90} = 0.133$. There is only one path which will give this result.

Notice that testing could continue until either all 4 defective bulbs or all 6 good bulbs are found.

Examle 1.22

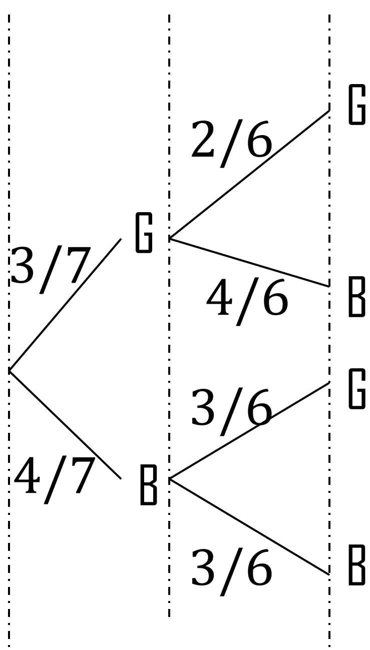
There are 3 girls and 4 boys in a class. If two students are selected from the class to attend two different meetings at the same time.

i. Draw a tree diagram to show all the possible outcomes and their respective probabilities.

ii. Calculate the probability that the students selected are of the same sex.

iii. If the selected students are of the same sex find the probability that they are both boys.

Solution



$$ii.\ P(same\ sex) = P(GG\ or\ BB)$$
$$= P(GG) + P(BB)$$
$$= \frac{3}{7} \times \frac{2}{6} + \frac{4}{7} \times \frac{3}{6}$$
$$= \frac{1}{7} + \frac{2}{7} = \frac{3}{7}$$

iii. $P(BB/same\ sex) = \frac{P(BB)}{P(GG\ or\ BB)}$

$$= \frac{2/7}{3/7} = \frac{2}{7} \times \frac{7}{3} = \frac{2}{3}$$

1.22 Independent events

Two events of separate experiments A and B are independent if the occurrence or non-occurrence of either event does not affect the occurrence or non-occurrence of the other. A and B are independent if and only if

$$\boxed{\mathrm{P}(A \cap B) = \mathrm{P}(A) \times \mathrm{P}(B)}$$

For independent events, $P(A/B) = P(A)$

If A and B are independent then $A\ and\ B', A'\ and\ B, A'\ and\ B'$ are pairs of independent events

Example 1.23

The probability that a bulb from a certain store will last for more than one year is 3/5 and the probability that a fully charged phone battery from the same store will last for more than one week is 2/7. A student bought the a bulb and a cell phone battery from that store.

Find the probability that the bulb will last for more than one year and the battery will lass for **less** than one week when fully charged.

Solution

Let A = Bulb will last for more than one year $\Longrightarrow P(A) = 3/5$

B = Battery will last for at least one week. $\Longrightarrow P(B) = 2/7$

$\bar{B}$ = Battery will last for less than one week

$$P(\bar{B}) = 1 - P(B) = 1 - \frac{2}{7} = \frac{5}{7}$$

Since A and B are independent A and $\bar{B}$ are also independent

$$P(A \cap \bar{B}) = P(A) \times P(\bar{B}) = \frac{3}{5} \times \frac{5}{7} = \frac{3}{7}$$

1.3 Bayes theorem and axiom on total probability

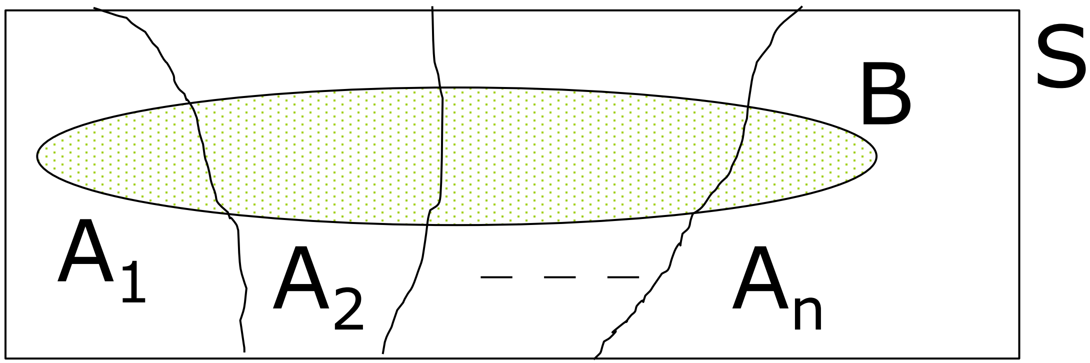
$$P(A/B) = \frac{p(A \cap B)}{P(B)} \Longrightarrow P(B) \times P(A/B) = P(A \cap B) \ldots .(1)$$

$$P(B/A) = \frac{p(A \cap B)}{P(A)} \Longrightarrow P(A) \times P(B/A) = P(A \cap B) \ldots .(2)$$

From (1) and (2): $\mathrm{P(B)} \times P(A/B) = P(A) \times P(B/A)$

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

If the sample space can be partition into disjoint events $A_1, A_2, \ldots, A_n$ then for $i = 1, 2, \ldots, n$



Law of Total Probability: $P(B) = P(B \cap A_1) + P(B \cap A_2) + \cdots + P(B \cap A_n)$

$$\Longrightarrow P(B) = P(A_1)P(B/A_1) + P(A_2)P(B/A_2) + \cdots + P(A_n)P(B/A_n)$$

Bayes' theorem: $P(A_i/B) = \dfrac{P(A_i)P(B/A_i)}{P(B)}$

$$P(A_i/B) = \frac{P(A_i)P(B/A_i)}{P(A_1)P(B/A_1) + P(A_2)P(B/A_2) + \cdots + P(A_n)P(B/A_n)}$$

Example1.31

A business man has three cartoons labelled A, B and C.

Cartoon A contains 7 Dell and 3 HP laptops.

Cartoon B contains 6 Dell and 4 HP laptops.

Cartoon C contains 2 Dell and 8 HP laptops.

A cartoon is chosen at random and a laptop is removed from it also at random.

a) Draw a tree diagram to show all the possible outcomes.

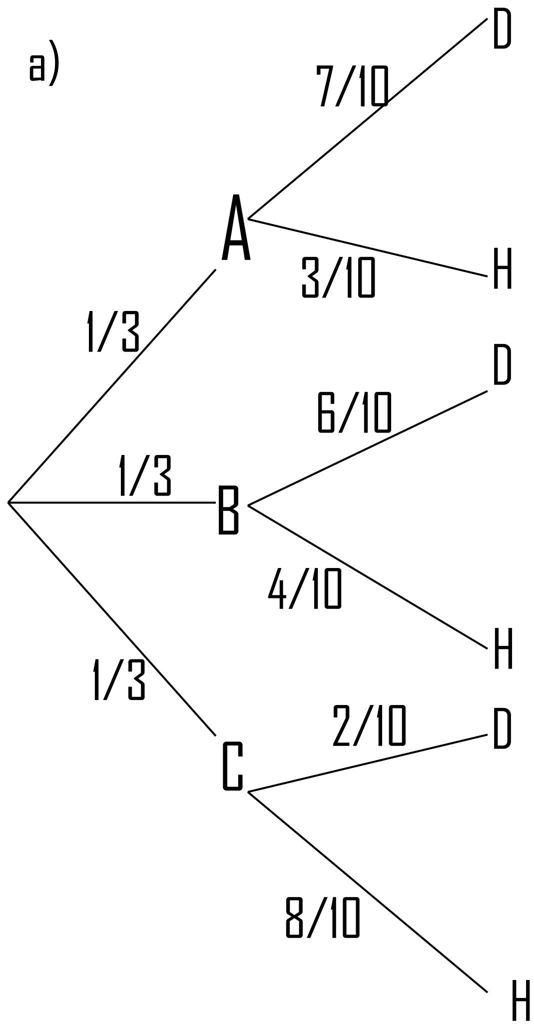
   Hence or otherwise

a) What is the probability of choosing Dell laptop.

b) Find the probability that a dell laptop chosen by a customer was from cartoon B.

Solution

a)



$$P(A)P(D/A) = \frac{1}{3} \times \frac{7}{10} = \frac{7}{30}$$

$$P(A)P(H/A) = \frac{1}{3} \times \frac{3}{10} = \frac{1}{10}$$

$$P(B)P(D/B) = \frac{1}{3} \times \frac{6}{10} = \frac{1}{5}$$

$$P(B)P(H/B) = \frac{1}{3} \times \frac{4}{10} = \frac{2}{15}$$

$$P(C)P(D/C) = \frac{1}{3} \times \frac{2}{10} = \frac{1}{15}$$

$$P(C)P(H/C) = \frac{1}{3} \times \frac{8}{10} = \frac{4}{15}$$

Solution continues

b) By the law of total probability

$$P(D) = P(A)P(D/A) + P(B)P(D/B) + P(C)P(D/C)$$

$$P(D) = \frac{7}{30} + \frac{1}{5} + \frac{1}{15} = \frac{1}{2}$$

c) probability that a dell laptop chosen by a customer was from cartoon B means P(laptop from B given that it is dell)

By Bayes' theorem

$$P(B/D) = \frac{P(B \cap D)}{P(D)} = \frac{P(B)P(D/B)}{P(D)} = \frac{1/5}{1/2} = \frac{2}{5}$$

Method 2

Cartoon	Dell, D	HP, H	Total
A	7	3	10
B	6	4	10
C	2	8	10
Total	15	15	30

b) $P(D) = \frac{n(D)}{n(S)} = \frac{15}{30} = \frac{1}{2}$

c) $P(B/D) = \frac{P(B \cap D)}{P(D)} = \frac{6/30}{1/2} = \frac{2}{5}$

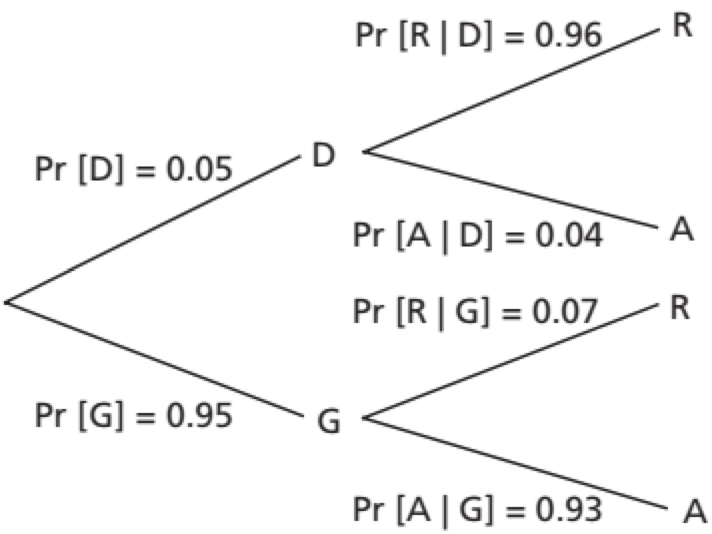
Exercise 1.31

A company produces machine components which pass through an automatic testing machine. 5% of the components entering the testing machine are defective. However, the machine is not entirely reliable. If a component is defective there is 4% probability that it will not be rejected. If a component is not defective there is 7% probability that it will be rejected.

a) What fraction of all the components are rejected?
b) What fraction of the components rejected are actually not defective?
c) What fraction of those not rejected are defective?

Solution

Let D represent a defective component, and G a good component.
Let R represent a rejected component, and A an accepted component.



$\Pr[D \cap R] =$	$\Pr[D] \times \Pr[R \mid D] =$	$(0.05)(0.96) =$	0.0480	Rejected
$\Pr[D \cap A] =$	$\Pr[D] \times \Pr[A \mid D] =$	$(0.05)(0.04) =$	0.0020	Accepted
$\Pr[G \cap R] =$	$\Pr[G] \times \Pr[R \mid G] =$	$(0.95)(0.07) =$	0.0665	Rejected
$\Pr[G \cap A] =$	$\Pr[G] \times \Pr[A \mid G] =$	$(0.95)(0.93) =$	0.8835	Accepted
		Total =	1.0000	(Check)

$$\Pr[R] = \Pr[D \cap R] + \Pr[G \cap R] = 0.0480 + 0.0665 = 0.1145$$

$$\Pr[A] = \Pr[D \cap A] + \Pr[G \cap A] = 0.0020 + 0.8835 = 0.8855$$

a) The answer to part (a) is that "in the long run" the fraction rejected will be the probability of rejection, 0.1145 or (with rounding) 0.114 or 11.4 %.

b) Fraction of components rejected which are not defective
= probability that a component is good, given that it was rejected

$$= \Pr[G \mid R] = \frac{\Pr[G \cap R]}{\Pr[R]} = \frac{0.0665}{0.1145} = 0.58 \text{ or } 58\ \%.$$

c) Fraction of components passed which are actually defective
= probability that a component is defective, given that it was passed

Using equation 2.4, this is $\Pr[D \mid A] = \dfrac{\Pr[D \cap A]}{\Pr[A]} = \dfrac{0.0020}{0.8855} = 0.0023$ or $0.23\ \%$.

Exercise 1.32

30% of computers in a certain store are from Dubai while the rest are from U.S. 2% of computers from Dubai and 0.1% of computers from U.S. are not up to standard.

A student bought a computer from this store.

a) Draw a tree diagram to show all the possible outcomes and their respective probabilities.

   Hence or otherwise

b) Find the probability that it is not up to standard (in 4 d.p.)

c) Find the probability that it is from Dubai given that it is not up to standard, giving your answer in 3 decimal places.

.

111.2 RANDOM VARIABLES

2.1 BASIC DEFINITIONS AND FORMULAS

***23 APRIL*, 2020**

JAFF LAWRENCE

OBJECTIVES

- Definitions, Notation
- Probability Distributions
- Probability density/mass function
- Expectation, Variance and Standard deviation
- Mode and Median
- Properties of E(X) and Var(X)
- Cumulative probability distribution
- Probability generating function

1 Definitions

1.0 A variable: A variable is a characteristic that can assume different values.

1.1 A random variable (r.v.): A r.v., X is a numerical-valued function which maps each point in the sample space *S* to exactly one numeric value determined by chance.

1.2 Notation

Random variables are denoted using capital letters such as $\boldsymbol{X}$ and the values that they can take by the corresponding lower case letters such as x

e.g $R = r \in \{0, 1, 2, 3\}$

1.3 TYPES OF RANDOM VARIABLES

- A quantitative or numerical variable is discrete (which take only whole numbers) or continuous (which take any value within an interval)
- A qualitative r.v. is descriptive e.g. colours, opinions, choice of food, religion, etc.

Examples of discrete r.v.s: number of heads obtained when n coins are tossed, number of children in a family, etc

Examples of continuous r.v.s: Measurements such as heights, weights, time, temperature, lifetime of electronic devices etc

1.4. Probability Distribution-D.R.V

A discrete probability distribution is a table which shows all the different values of the r.v. and their respective probabilities.

Example

Find the possible values and the probability distribution for the number of heads obtained when two coins are tossed.

Soution

		Second coin	
		T	**H**
First coin	T	TT	TH
	H	HT	HH

T= Tail H = Head

Sample space

$$S = \{TT, TH, HT, HH\}$$

$$X = number\ of\ heads$$

$$X(TT) = 0, X(TH) = 1, X(HT) = 1, X(HH) = 2$$

$$X = \{0, 1, 2\}$$

$$P(X = 0) = P(TT) = 1/2 \times 1/2 = 1/4$$

$$P(X = 1) = P(TH \text{ or HT}) = P(TH) + P(HT)$$

$$= 1/2 \times 1/2 + 1/2 \times 1/2 = 1/4 + 1/4 = 1/2$$

$$P(X = 2) = P(HH) = 1/2 \times 1/2 = 1/4$$

The probability distribution or Probability law of X is

x	0	1	2
$P(X = x)$	1/4	1/2	1/4

Observe that $\sum_{\text{all x}} P(X = x) = 1/4 + 1/2 + 1/4 = 1$

Properties of a probability distribution

If x is a d.r.v. then

$$1. \quad \sum_{all\ x} P(X = x) = 1$$

$$2. \quad 0 \le P(X = x) \le 1$$

1.5.0 Probability Density/Mass Function, p.d.f or p.m.f

The p.d.f or p.m.f is a non-negative function $f(x)$ that is used to alocate probabilities to the different values of the r.v. i.e.

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

Example

Show that $f(x) = {}^2C_x\left(\frac{1}{4}\right), x = 0, 1, 2$ is a p.d.f. for the r.v. in the example above.

Solution

x	0	1	2
$P(X = x)$	1/4	1/2	1/4

$$\boldsymbol{f(x) = 2_{C_x}\left(\frac{1}{4}\right), x = 0, 1, 2}$$

$$P(X = 0) = f(0) = {}^2C_0\left(\frac{1}{4}\right) = 1\left(\frac{1}{4}\right) = \frac{1}{4}$$

$$P(X = 1) = f(1) = {}^2C_1\left(\frac{1}{4}\right) = 2\left(\frac{1}{4}\right) = \frac{1}{2}$$

$$P(X = 2) = f(2) = {}^2C_2\left(\frac{1}{4}\right) = 1\left(\frac{1}{4}\right) = \frac{1}{4}$$

Hence $f(x) is\ a\ p.d.f\ for\ the\ randon\ variable\ in\ example\ ono$

1.5.1 Properties of a pdf, $f(x)$

- A discrete random variable then:

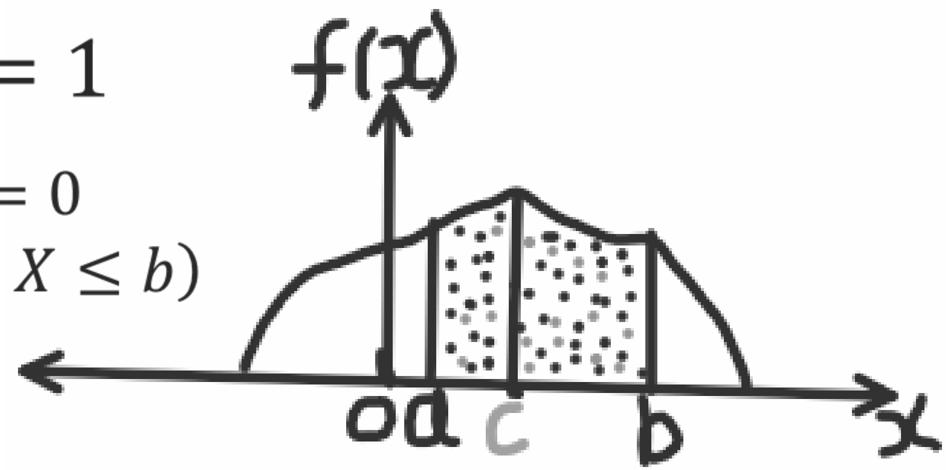
  $\mathrm{P(X = x) = f(x)}$ and $\sum f(x) = 1$
- A continuous random variable then

  $P(a \leq X \leq b) = \int_a^b f(x)dx = Shaded\ Area$

  $and\ \int_{-\infty}^{\infty} f(x)dx = 1$

NB: $\mathrm{P}(a) = \mathrm{P(c)} = \mathrm{P(b)} = 0$

$\mathrm{P}(a < X < b) = P(a \leq X \leq b)$



1.6. Cumulative or Distribution function

- The cumulative probability of x is $F(x) = P(X \leq x)$

  It is the sum of probabilities up to x.

- $F(x) = P(X \leq x) = \sum_{u \leq x} f(u)$ if X is a d.r.v.

  In this case P$(X = x) = F(x) - F(x-1)$

- $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du$ if X is a c.r.v.

  In this case $P(a \leq X \leq b) = F(b) - F(a) = \int_{a}^{b} f(x)dx$

- P.d.f. $f(x) = F'(x)$

Example: Find the distribution function for the r.v. below and represent it graphically

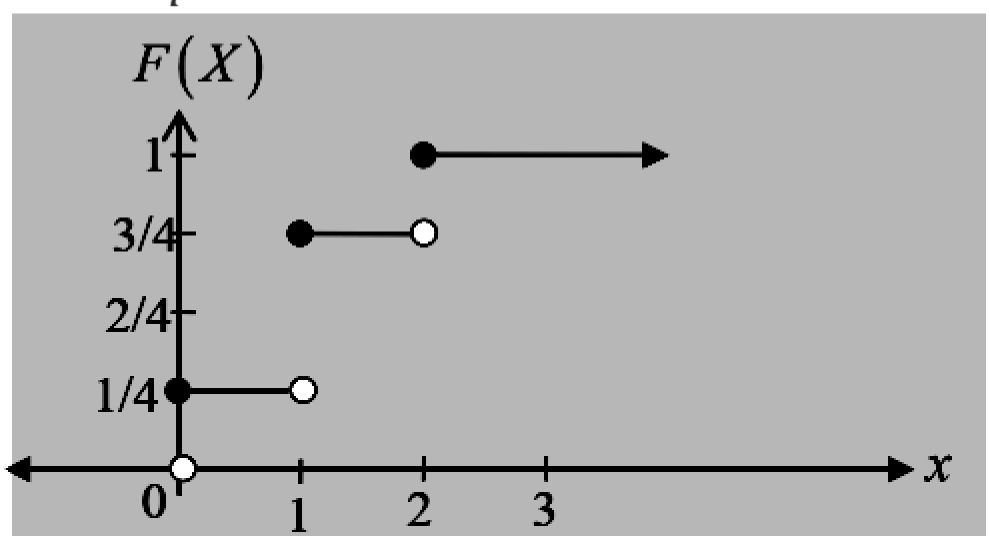
x	0	1	2
$P(X=x)$	1/4	1/2	1/4

Solution

Cumulative Distribution $F(x)=P(X\le x)$

$x\in$	$x\le 0$	$0\le x<1$	$1\le x<2$	$x\ge 2$
$F(x)$	0	1/4	3/4	1

Graphical representation



Example :

The p.d.f of a r.v. X is $f(x) = \begin{cases} cx^2, 0 \le x \le 3 \\ 0, \quad Otherwise \end{cases}$

Find: a) The value of c. b) the distribution function. c) $P(1 < X < 2)$ d) The median

Solution

a) Since X is a c.r.v. $\int_{-\infty}^{\infty} f(x)dx = 1 \Rightarrow \int_0^3 cx^2 dx = 1 \Rightarrow \left[\frac{cx^3}{3}\right]_0^3 = 1$

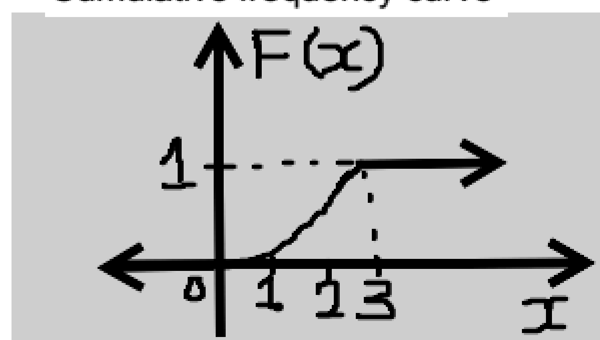
$$9c - 0 = 1 \Rightarrow c = \frac{1}{9} \Rightarrow f(x) = \frac{x^2}{9}, 0 \le x \le 3$$

b) $F(x) = P(X \le x) = \int_{-\infty}^{x} f(u)du$

$\frac{1}{9}\int_0^x x^2 du = \frac{1}{9}\left[\frac{x^3}{3}\right]_0^x = \frac{x^3}{27}$

$$F(x) = \begin{cases} 0, & x \le 0 \\ \frac{x^3}{27}, & 0 \le x < 3 \\ 1, & x \ge 3 \end{cases}$$

Cumulative frequency curve



c) $P(1 < X < 2) = F(2) - F(2) = \frac{2^3}{9} - \frac{1^3}{9} = \frac{7}{9}$

b) Median, m: $F(m) = \frac{1}{2} \Rightarrow \frac{m^3}{9} = \frac{1}{2} \Rightarrow m^3 = \frac{9}{2} \Rightarrow m = \sqrt[3]{9/2}$

2 MOMENTS OF A RANDOM VARIABLE

The kth moment of a r.v. X about the origin is:

$$E(X^k) = \begin{cases} \sum x^k P(x) \,, if\ X\ is\ discrete \\ \int_{-\infty}^{\infty} x^k f(x) dx \,, if\ X\ is\ continuous \end{cases}$$

Where $\mathrm{k} = 0, 1,2, \ldots \text{and}\ P(x)\ or\ f(x)$ is the p.d.f.

The 1st moment is called the expectation E(X) or mean $\mu\ or\ \mu_X$

The 2nd moment is $E(X^2)$

2.1 Expectation E(X) or Mean μ

1. Expectation or mean,

$$E(X) or\ \mu = \sum xP(x)\ , if\ X\ is\ discrete$$

$$E(X) or\ \mu = \int_{-\infty}^{\infty} xf(x)dx\ , if\ X\ is\ continuous$$

2. Expectation of a function, g(x) is the 1st moment of $g(x)$

$$E[g(x)] or\ \mu = \sum g(x)P(x)\ , if\ X\ is\ d.r.v.$$

$$E[g(x)] or\ \mu = \int_{-\infty}^{\infty} g(x)f(x)dx\ , if\ X\ is\ c.r.v.$$

2.2. Variance Var(X) or σ^2 or σ_X^2 and Standard deviation, σ or σ_X

1. Variance, $Var(X)$ or σ^2 is the 2nd moment of the function $X - \mu$. i.e.

$$Var(X) \text{ or } \sigma^2 = E[(X-\mu)^2]$$

It can be shown that

$$Var(X) \text{ or } \sigma^2 = E(X^2) - \mu^2$$

$Var(X)$ or $\sigma^2 = \sum x^2 P(x) - \mu^2$ if X is a d.r.v.

$Var(X)$ or $\sigma^2 = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2$ if x is a c.r.v.

2. Standard deviation, $\sigma = \sqrt{Var(X)}$

2.3. Mode and Median

Mode = Value with the highest probability or greatest value of the p.d.f.

Median = Middle value when the r.v. are in order.

Median, m: $P(X \leq m) = \frac{1}{2} \; \left(or \; \int_a^m f(x)dx = \frac{1}{2} \right)$

NB: The mean, mode and median are measures of central tendency while the variance and standard deviation are measures of spread or variation

Example

Find the mean, variance, standard deviation, mode and median of the data in example one above

Solution

x	0	1	2
$\mathrm{P}(X=x)$	1/4	1/2	1/4

$$\text{Mean, } \mu = \sum_{\text{all x}} x\mathrm{P}(X=x) = 0(1/4)+1(1/2)+2(1/4)=1$$

$$E(X^2) = \sum_{\text{all x}} x^2\mathrm{P}(X=x) = 0^2(1/4)+1^2(1/2)+2^2(1/4)=3/2$$

$$Var(X) = E(X^2)-\mu^2 = 3/2-1^2 = 1/2$$

$$\text{Standard deviation, } \sigma=\sqrt{Var(X)} = \sqrt{1/2} = 0.71$$

$$Mode = 1$$

$$Median = 1$$

NB: For a symmetrical distribution

$$Mean = Mode = Median$$

Example

A p.d.f of a r.v. X is defined by $f(x) = \begin{cases} \frac{3}{4}(1 - x^2), -1 \le x \le 1 \\ 0, otherwise \end{cases}$

i. Calculate

a) $P(0 \le X \le 0.5)$

b) $P(-0.3 \le X \le 0.7)$

c) $P(|X| < 0.5)$

d) $P(X > 0.5)$

e) $P(X \le 0.7)$

ii. Find :

a) The mean

b) The variance

c) The standard deviation

iii. Find the mode and the median

Solution

$$f(x) = \begin{cases} \frac{3}{4}(1 - x^2), & -1 \le x \le 1 \\ 0, & elsewhere \end{cases}$$

a) $P(0 \le x \le 0.5) = \int_0^{0.5} f(x)dx$

$$= \frac{3}{4}\int_0^{0.5}(1 - x^2)dx$$

$$\frac{3}{4}\left[x - \frac{x^3}{3}\right]_0^{0.5} = \frac{11}{32}$$

b) $P(-0.3 \le x \le 0.7) = \int_{-0.3}^{0.7} f(x)dx$

$$\frac{3}{4}\left[x - \frac{x^3}{3}\right]_{-0.3}^{0.7} = 0.6575$$

c) $P(|x| \le 0.5) = P(-0.5 \le x \le 0.5)$

$$= \int_{-0.5}^{0.5} f(x)dx$$

$$= \frac{3}{4}\int_{-0.5}^{0.5}(1 - x^2)dx$$

$$\frac{3}{4}\left[x - \frac{x^3}{3}\right]_{-0.5}^{0.5} = \frac{11}{16}$$

d) $P(x \ge 0.5) = P(0.5 \le x \le 1)$

$$\int_{0.5}^{1} f(x)dx$$

$$\frac{3}{4}\left[x - \frac{x^3}{3}\right]_{0.5}^{1} = \frac{5}{32}$$

e) $P(x \le 0.7) = P(-1 \le x \le 0.7)$

$$\int_{-1}^{0.7} f(x)dx$$

$$\frac{3}{4}\left[x - \frac{x^3}{3}\right]_{-1}^{0.7} = 0.9393$$

ii. Mean, $\mu = \frac{3}{2}\int_{-1}^{1} xf(x)dx$

$$= \frac{3}{4}\int_{-1}^{1}(x - x^3)dx$$

$$\frac{3}{4}\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_{-1}^{1} = 0$$

Variance, $Var(X)\ or\ \sigma^2 = E(X^2) - \mu^2$

$$Var(X)\ or\ \sigma^2 = \int_{-1}^{1} x^2 f(x)dx - (0)^2$$

$$= \frac{3}{4}\int_0^1 x^2(1 - x^2)dx$$

$$= \frac{3}{4}\int_{-1}^{1}(x^2 - x^4)dx = \frac{1}{5}$$

Standard deviation, $\sigma = \sqrt{Var(X)} = \sqrt{\frac{1}{5}} = 0.45$

2.4. Properties of Expectation and Variance

If a and b are real constants then

1. $E(a) = a$ and $Var(a) = 0$
2. $E(aX + b) = aE(X) + b$

   $Var(aX + b) = a^2 Var(X)$
3. If X and Y are independent

   $E(aX + bY) = aE(X) + bE(Y)$

   $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y)$
4. $\mathrm{E}(aX - bY) = aE(X) - bE(Y)$

   $Var(aX - bY) = a^2 Var(X) + b^2 Var(Y)$

NB:

$NB: Var(aX - bY) = Var[aX + (-b)Y]$

$= a^2 Var(X) + (-b)^2 Var(Y) = a^2 Var(X) + b^2 Var(Y)$

Example

Given that X and Y are independent random variables with

$$E(X)=2, E(Y)=3, \quad Var(X)=3, Var(Y)=4,$$

Find

$i.\ E\left(X+2Y^2+1\right)$

$ii.\ E\left(2X^2-3X+2\right)$

$iii.\ Var\left(2X-3Y-1\right)$

Solution

$$E(X)=2, E(Y)=3, \quad Var(X)=3, Var(Y)=4$$

$$E(X^2)-2^2=3 \Rightarrow E(X^2)=7$$

$$E(Y^2)-3^2=4 \Rightarrow E(Y^2)=13$$

$$i. E(X+2Y^2+1)=E(X)+2E(Y^2)+1$$

$$=2+2(13)+1=29$$

$$ii. E(2X^2-3X+2)=2E(X^2)-3E(X)+2$$

$$=2(7)-3(2)+2=10$$

$$iii. Var(2X-3Y-1)=2^2Var(X)+3^2Var(Y)$$

$$=4(3)+9(4)$$

$$=48$$

NB:

$$V(X)=E(X^2)-\mu^2$$
$$\Longrightarrow 3=E(X^2)-2^2$$
$$\Longrightarrow E(X^2)=3+4=7$$

$$V(Y)=E(Y^2)-\mu^2$$
$$\Longrightarrow 4=E(X^2)-3^2$$
$$\Longrightarrow E(X^2)=4+9=13$$

2.5 PROBABILITY GENERATING FUNCTION (p.g.f)

The p.g.f of a r.v. X is

- ❖ $\text{G}(t) = E(t^x) = \sum t^x P(x)\,, if\ X is\ a\ d.r.v.$
- ❖ $G(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx\,, if\ X\ is\ a\ c.r.v.$

Applications

- ❖ Mean $E(X) or\ \mu = G'(1)$
- ❖ Variance, $Var(X) or\ \sigma^2 = G''(1) + G'(1) - [G'(1)]^2$

2.3. JOINT AND MARGINAL DISTRIBUTIONS

The joint probability function of r.v.s X and Y is

$$P(X = x, Y = y) = f(x, y)$$

3.1. Discrete case

$$f(x, y) \geq 0 \; and \; \sum_x \sum_y f(x, y) = 1$$

i.e. sum of all probabilities is 1

3.1.1: Joint probability table

$X \backslash Y$	y_1	y_2	$\cdots$	y_n	Totals ↓
x_1	$f(x_1, y_1)$	$f(x_1, y_2)$	$\cdots$	$f(x_1, y_n)$	$f_1(x_1)$ ← $P(X=x_1)$
x_2	$f(x_2, y_1)$	$f(x_2, y_2)$	$\cdots$	$f(x_2, y_n)$	$f_1(x_2)$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
x_m	$f(x_m, y_1)$	$f(x_m, y_2)$	$\cdots$	$f(x_m, y_n)$	$f_1(x_m)$
Totals →	$f_2(y_1)$	$f_2(y_2)$	$\cdots$	$f_2(y_n)$	1 ← Grand Total

$P(Y=y_1)$

Marginal distributions = Totals

Marginal Probability Distribution of X

x	X_1	X_2	…	X_M	Total
$P(x) = f_1(x)$	$f_1(x_1)$	$f_1(x_2)$	…	$f_1(x_m)$	1

Marginal Probability Distribution of Y

y	y_1	y_2	…	y_n	Total
$P(y) = f_2(y)$	$f_2(y_1)$	$f_2(y_2)$	…	$f_2(y_n)$	1

3.1.2. Continuous case

$$f(x, y) \geq 0 \; and \; \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx \, dy = 1$$

$$P(a < X < b, c < Y < d) = \int_{x=a}^{b} \int_{y=c}^{d} f(x, y) \, dx \, dy$$

Example

The joint density function of two continuous random variables X and Y is

$$f(x, y) = \begin{cases} cxy & 0 < x < 4, 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the value of the constant c.

(b) Find $P(1 < X < 2, 2 < Y < 3)$.

(c) Find $P(X \geq 3, Y \leq 2)$.

Solution

$$a)\text{Total probability} = 1 \Rightarrow \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1 = \int_{0}^{4}\int_{1}^{5} cxy\ dydy = \int_{0}^{4}\frac{cx}{2}\left[y^2\right]_1^5 dx$$

$$= \int_{0}^{4}\frac{cx}{2}\left[5^2 - 1^2\right]dx = \int_{0}^{4} 12cx\ dx = \left[6cx^2\right]_0^4 = 6c\left[4^2 - 0^2\right] = 96c$$

$$\text{Hence, } 96c = 1 \Rightarrow c = 1/96$$

$$b)\text{ Hence } P\left(1 < X < 2, 2 < Y < 3\right) = \frac{1}{96}\int_{x=1}^{2}\int_{y=2}^{3} xy\ dydy = \frac{1}{96}\int_{1}^{2}\frac{x}{2}\left[y^2\right]_{y=2}^{3} dx$$

$$= \frac{1}{96}\int_{1}^{2}\frac{x}{2}\left[3^2 - 2^2\right]dx = \frac{1}{96}\int_{1}^{2}\frac{5}{2}x\ dx = \frac{1}{96}\left[\frac{5}{2}\frac{x^2}{2}\right]_1^2 = \frac{1}{96}\left[\frac{5}{4}\left(2^2 - 1^2\right)\right] = \frac{1}{96}\left(\frac{5}{4}\right)(3) = \frac{5}{128}$$

Solution continues

c) $$P\left(X \geq 3, Y \leq 2\right) = \frac{1}{96}\int_{x=3}^{4}\int_{y=1}^{2} xy\ dydy = \frac{1}{96}\int_{x=3}^{4}\frac{x}{2}\left[y^2\right]_{y=1}^{2} dx$$

$$= \frac{1}{96}\int_{x=3}^{4}\frac{x}{2}\left[2^2 - 1^2\right]dx = \frac{1}{96}\int_{3}^{4}\frac{3}{2}x\ dx = \frac{1}{32}\left[\frac{1}{2}\frac{x^2}{2}\right]_3^4$$

$$= \frac{1}{32}\left[\frac{1}{4}\left(4^2 - 3^2\right)\right] = \frac{1}{32}\left(\frac{1}{4}\right)(7) = \frac{7}{128}$$

Example:

A student of KEBHIP took a survey of 500 telephone subscribers to determine people's favourite choice.

	Male	Female	Total
MTN	80	120	
ORANGE	100	25	
NEXTTEL	50	125	
TOTAL			

i. Complete the table

ii. Hence, construct the corresponding joint probability distribution

Find the probability that a subscriber selected at random is:

iii. a) Female and prefers MTN b) Male c) Prefers MTN

iv. a) Male and prefers MTN b) Male or prefers MTN

v. Prefers MTN given that the subscriber is Male

v. Given that the subscriber's favourite choice is ORANGE, what is the probability that the subscriber is male.

solution:

i.

	Male	Female	Total
MTN	80	120	**200**
ORANGE	100	25	**125**
NEXTTEL	50	125	**175**
TOTAL	**230**	**270**	**500**

ii. Joint Probability distribution: Divide each value by 500

	Male	Female	Total
MTN	0.16	0.24	**0.40**
ORANGE	0.20	0.05	**0.25**
NEXTTEL	0.10	0.25	**0.35**
TOTAL	**0.46**	**0.54**	**1**

iii. a) $P(Female \cap MTN) = 0.24$

b) $P(Male) = 0.46$

c) $P(MTN) = 0.40$

iv. a) $P(Male \cap MTN) = 0.16$

b) $P(Male \cup MTN) = 0.16 + 0.20 + 0.10 + 0.24 + 0.40 = 0.7$

v. $P(MTN/Male) = \frac{P(MTN \cap Male)}{P(Male)} = \frac{0.16}{0.46} = \frac{8}{23}\ or\ 0.35$

vi. $P(Male/MTN) = \frac{P(MTN \cap Male)}{P(MTN)} = \frac{0.16}{0.40} = 0.4$

Expectation, Variance, Standardard deviation, Covariance, and Correlation coefficient.

$E(X), E(Y), var(X)\ and\ var(Y)$ can be obtained by using the marginal distributions of X and Y. $E(XY) = \sum_x \sum_y xyP(x, y)$

Covariance of x and y, $Cov(x, y)$ or $S_{XY} = E(XY) - \mu_X \mu_Y$

Product moment Linear correlation coefficient, r:

$$r = \frac{S_{XY}}{\sqrt{\sigma_X^2 \sigma_Y^2}} \text{ or } \frac{S_{XY}}{\sigma_X \sigma_Y}$$ where $-1 \le r \le 1$

r is a measure of the extent to which the points (x,y) approximate to a straight line.

Interpretation of r: $-1 \le r \le 1$

- $r = 1 \Longrightarrow perfect\ linear + ve\ correlation$: Points (x,y) on a straight line with positive gradient. i.e. y and x vary by the same ratio.
- $\mathrm{r} > 0 \Longrightarrow$Positive linear correlation between x and y. y increases as x decreases and vice versa.
- $r = -1 \Longrightarrow perfect\ linear - ve\ correlation$: Points (x,y) on a straight line with negative gradient. i.e. y and x vary by the same ratio but in opposite directions.
- $\mathrm{r} < 0 \Longrightarrow$ Negative linear correlation between x and y. y increases as x decreases and vice versa.

Independent events and regression lines

- ❖ X and Y are independent iff for all x and y:
  $\boldsymbol{P(X = x, Y = y) = P(X = x)P(Y = y)}$
- ❖ Regression line y on x: $y - \mu_Y = \frac{S_{XY}}{Var(X)}(x - \mu_X)$

  $\Longrightarrow y = a + bx$:

  It is used to estimate values of y.
- ❖ Regression line x on y: x$-\mu_X = \frac{S_{XY}}{Var(Y)}(y - \mu_Y)$

  $\Longrightarrow x = c + dy$

  It is used to estimate values of x

Example:

The joint probability distribution of 2 r.v.s X and Y is as follows:

X \ Y	0	1	2
1	0.0	0.06	0.12
3	0.50	0.24	0.08

Find:

i. The marginal distributions of X and Y.

ii. $E(X), Var(X),\ \sigma_X, E(Y), Var(Y), \sigma_Y, E(XY)$ *and* $Cov(X,Y)$

iii. The correlation coefficient and interpret your result.

iv. The regression lines: (a) y on x, (b) x on y

v. Hence, estimate the value of y corresponding to $x = 2$

vi. Determine whether X and Y are independent.

Solution

i.

x	1	3
P(x)	0.18	0.82

y	0	1	2
P(y)	0.50	0.30	0.20

ii. $E(X) = \sum xP(x) = 1(0.18) + 3(0.82) = 2.64$

$Var(X) = E(X^2) - \mu_X^2 = 1^2(0.18) + 3^2(0.82) - 2.64^2 = 0.59$

$\sigma_X = \sqrt{Var(X)} = \sqrt{0.59} = 0.77$

$$E(Y) = \sum yP(y) = 0(0.50) + 1(0.30) + 2(0.20) = 0.70$$

$Var(Y) = E(Y^2) - \mu_Y^2 = 0^2(0.50) + 1^2(0.30) + 2^2(0.20) - 0.70^2 = 0.61$

$\sigma_Y = \sqrt{Var(Y)} = \sqrt{0.61} = 0.78$

$$E(XY) = \sum_X \sum_Y xyP(x, y)$$

Solution continues

$$E(XY) = 1(0)(0.00) + 1(1)(0.06) + 1(2)(0.12)$$
$$+3(0)(0.50) + 3(1)(0.24) + 3(2)(0.08) = 1.5$$
$$Cov.(X, Y) \text{ or } S_{XY} = E(XY) - \mu_X \mu_Y$$
$$Cov.(X, Y) \text{ or } S_{XY} = 1.5 - (2.64)(0.70) = -0.35$$

iii. Correlation coefficient, $r = \frac{S_{XY}}{\sqrt{\sigma_X^2 \sigma_Y^2}} = \frac{-0.35}{\sqrt{0.59 \times 0.61}} = -0.58$

Hence, there is a moderate negative linear correlation between X and Y. Y tends to decrease as X increases.

Solution continuses

Regression line y on x: $y - \mu_Y = \frac{s_{XY}}{Var(X)}(x - \mu_X)$

$$y - 0.70 = \frac{-0.35}{0.59}(x - 2.64) \Longrightarrow y = 2.27 - 0.59x$$

The gradient is -0.59 indicating that y decreases by 0.59 for every unit increase in x.

Regression line x on y: $x - \mu_X = \frac{s_{XY}}{Var(Y)}(y - \mu_Y)$

$$x - 2.64 = \frac{-0.35}{0.61}(y - 0.70) \Longrightarrow x = 3.04 - 0.57y$$

The gradient is -0.57 indicating that x decreases by 0.57 for every unit increase in y.

v. When $x = 2, y = 2.27 - 0.59(2) = 1.09$

vi. X and Y are independent iff for all (x, y) $P(x, y) = P(x).P(y)$

$$P(1,0) = 0.00, \qquad P(X = 1) \times P(Y = 0) = 0.18 \times 0.5 = 0.09$$

$P(1,0) \neq P(X = 1) \times P(Y = 0)$ Hence by counter example X and Y are not independent..

Example

The joint probability distribution of X and Y is

$$P(x,y) = \begin{cases} 0.1 \; for \; (x,y) = (1,1) \\ a \; for \; (x,y) = (1,2) \\ 0.4 \; for \; (x,y) = (2,1) \\ 0.2 \; for \; (x,y) = (2,2) \end{cases}$$

i. Find the value of the constant a. Hence:

ii. Determine the marginal probability mass functions of X and Y.

iii. Find $P(X = 2) \; and \; P(2X \leq Y)$

Example

The joint probability function of two discrete random variables X and Y is given by $f(x, y) = c(2x + y)$, where x and y can assume all integers such that $0 \le x \le 2$, $0 \le y \le 3$, and $f(x, y) = 0$ otherwise.

(a) Find the value of the constant c.

(c) Find $P(X \ge 1, Y \le 2)$.

Find the marginal probability functions (a) of X and (b) of Y for the random variables of Problem 2.8.

2. 3. Chebyschev's Inequality

Suppose that X is a r.v. (discrete or continuous) having mean μ $and\ variance\ \sigma^2$. Then if k is a positive constant,

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Example

Let $k = 2$ then $P(|X-\mu| \geq 2\sigma) \leq \frac{1}{2^2}$

i.e. $P(|X-\mu| \geq 2\sigma) \leq 0.25$ or $P(|X-\mu| < 2\sigma) \geq 0.75$

The probability of a r.v. differing from its mean by more than 2 standard deviation is less or equal to o.25

2. 4. The law of large numbers

As the number of trials increase the observed probabilities approaches the theoretical probabilities.

e.g. The probability of getting a head when a coin is tossed is ½ meaning that if we toss a coin 10 times we expect that 50% of the trials will be heads. This will likely not be the case. However if we toss the same coin 10000 times we shall get close to 5000 heads

Example

Given that

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, where\ x = 0,1,2,3, \ldots$$

is the probability density function of a random variable, X.

Show that the mean and variance are each of value λ.

Hence find the variance of X.

Solution

- $P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0,1,2,3,4, \ldots$
- $G(t) = \sum_{x=0}^{\infty} t^x P(X = x) = \sum_{x=0}^{\infty} t^x \frac{e^{-\lambda}\lambda^x}{x!}$
- $G(t) = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!} =$

$$= e^{-\lambda}\left[1 + \lambda t + \frac{(\lambda t)^2}{2!} + \frac{(\lambda t)^3}{3!} + \cdots\right]$$

- $G(t) = e^{-\lambda}[e^{\lambda t}]$

Solution continues: $G(t) = e^{-\lambda}[e^{\lambda t}]$

- $G'(t) = \lambda e^{-\lambda}[e^{\lambda t}]$
- $G''(t) = \lambda^2 e^{-\lambda}[e^{\lambda t}]$
- $\boldsymbol{E(X) = G'(1) = \lambda e^{-\lambda}[e^{\lambda}] = \lambda}$
- $Var(X) = G''(1) + G'(1) - [G'(1)]^2$
- $= \lambda^2 e^{-\lambda}[e^{\lambda}] + \lambda - \lambda^2$
- $Var(X) = \lambda^2 + \lambda - \lambda^2$
- $\boldsymbol{Var(X), \sigma^2} = \lambda$ Standard deviation, $\sigma = \sqrt{\lambda}$

Exercise

1. The p.d.f. of a random variable X is

$$f(x) = \begin{cases} kx^2, & x = 1, 2, 3 \\ k(7-x)^2, & x = 4, 5, 6 \\ 0, & \text{otherwise} \end{cases}$$

Find

a) The value of k

b) The mean of X

c) The standard deviation of X in four significant figures

d) The mean and variance of $3X - 4$

Qquestion 2

2. A discrete random variable X has the following probability distribution\

$X = x$	0	1	2	3	4
$P(X = x)$	0.12	p	0.4	q	0.08

Given that $E(X) = 2.02$, find

a) The values of p and q.
b) $Var(X)$
c) The mean and variance of $Y = 5X - 2$
d) Find the cumulative probability distribution of X
e) Find the probability generating function of X.

Solution Qn. 1 (a)

$$f(x)=\begin{cases} kx^2, & x=1,2,3 \\ k(7-x)^2, & x=4,5,6 \\ 0, & \text{otherwise} \end{cases}$$

x	1	2	3	4	5	6
f(x)	k	4k	9k	9k	4k	k

Since X is a D.r.v. $\sum_{all\ x} P(X=x)=1$

$$k+4k+9k+9k+4k+k=1$$

$$28k=1 \Rightarrow k=\frac{1}{28}$$

Qn. 1(b) and (c)

b) Mean $\mu = \sum_{all\ x} xP(X = x)$

$$= 1(k) + 2(4k) + 3(9k) + 4(9k) + 5(4k) + 6(k)$$

Mean, $\mu = 98k = 98\left(\frac{1}{28}\right) = 3.5$

c) $E\left(X^2\right) = \sum_{\text{all } x} x^2 P\left(X = x\right)$

$$E(X^2) = 1^2(k) + 2^2(4k) + 3^2(9k) + 4^2(9k) + 5^2(4k) + 6^2(k)$$

$$E(X^2) = 378k = 378\left(\frac{1}{28}\right) = \frac{27}{2}$$

$$Var(X)\ or\ \sigma^2 = E(X^2) - \mu^2$$

$$Var(X)\ or\ \sigma^2 = \frac{27}{2} - 3.5^2 = 1.25$$

Standard deviation, $\sigma = \sqrt{Var(X)} = \sqrt{1.25} = 1.12$

Qn. 1 (d)

d) $E(3X-4)=3E(X)-4=3(3.5)-4=6.5$

$Var(3X-4)=3^2Var(X)=9(1.25)=11.25$

Qu. 2 (a)

x	0	1	2	3	4
$P(X=x)$	0.12	p	0.4	q	0.08

Since X is a D.r.v. $\sum_{all\ x} P(X=x)=1$

$$0.12+p+0.4+q+0.08=1$$

$$p+q=0.4 \ldots\ldots\ldots.(1)$$

Mean $\mu=\sum_{all\ x} xP(X=x)=2.02$

$$\Longrightarrow 0(0.12)+1(p)+2(0.4)+3(q)+4(0.08)=2.02$$

$$p+3q=0.9 \ldots\ldots\ldots(2)$$

$$(2)-(1): 2q=0.5 \Longrightarrow q=\frac{0.5}{2}=0.25$$

$$Sub.\,in\ (1):\ p=0.4-0.25=0.15$$

Qn. 2 (b) and (c)

b) $$E(X^2) = \sum_{\text{all } x} x^2 P(X = x)$$

$$E(X^2) = 0(0.12) + 1^2(0.15) + 2^2(0.4) + 3^2(0.25) + 4^2(0.0$$

$$E(X^2) = 5.28$$

$$Var(X) \text{ or } \sigma^2 = E(X^2) - \mu^2$$

$$Var(X) \text{ or } \sigma^2 = 5.28 - 2.02^2 = 1.20$$

c) $Y = 5X - 2$

$$E(Y) = 5E(X) - 2 = 5(2.02) - 2 = 8.1$$

$$Var(Y) = 5^2 Var(X) = 25(1.20) = 30$$

Qn. 2 (d) :

$X = x$	0	1	2	3	4
$P(X = x)$	0.12	$\boxed{0.15}$	0.4	$\boxed{0.25}$	0.08

Cumulative Distribution $F(x) = P(X \leq x)$

$x \in$	$]-\infty, 0[$	$[0,1[$	$[1,2[$	$[2,3[$	$[3,4[$	$[4, \infty[$
$F(x)$	0	0.12	0.27	0.67	0.92	1

OR

x	$x<0$	$0 \leq x < 1$	$1 \leq x < 2$	$2 \leq x < 3$	$3 \leq x < 4$	$x \geq 4$
$F(x)$	0	0.12	0.27	0.67	0.92	1

Qn. 2 (e) Probability generating function

$X = x$	0	1	2	3	4
$P(X = x)$	0.12	$\boxed{0.15}$	0.4	$\boxed{0.25}$	0.08

e) $G(t) = E(t^x) = \sum_{\text{all } x} t^x P(X = x)$

$$G(t) = t^0(0.12) + t^1(0.15) + t^2(0.4) + t^3(0.25) + t^4(0.08)$$

$$G(t) = 0.12 + 0.15t + 0.4t^2 + 0.25t^3 + 0.08t^4$$

III.3. PROBABILITY LAWS

PROBABILITY LAWS

1. Probability Scale: $0 \leq P \leq 1$

   The probability of any event lies between 0 and 1 inclusive.
2. The sum of the probabilities of all the outcomes in the sample space is 1.
3. Complementary events

$$P(A') = 1 - P(A) \; or \; P(A) + P(A') = 1$$

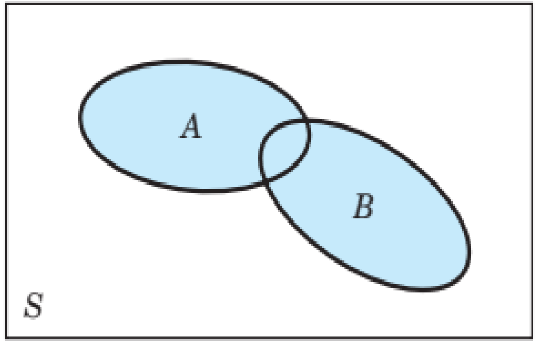
Example:

Five fair coins are tossed simultaneously. Find the probability of getting at least one head.

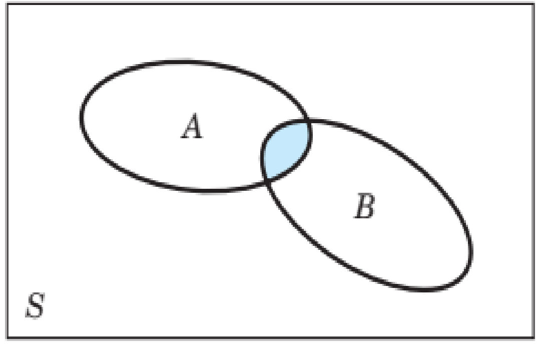
Solution: $P(at\ least\ 1\ head) = 1 - P(No\ head) = 1 - \left(\frac{1}{2}\right)^5 = 1 - \frac{1}{32} = \frac{31}{32}$

PROBABILITY LAWS

4. Addition Rule: If A and B are any two events then



Union $A \cup B$



Intersection $A \cap B$

$$\boldsymbol{P(A\ or\ B)}$$
$$\boldsymbol{= P(A) + P(B) - P(A\ and\ B)}$$

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

If A, B and C are any three events then

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

Example

In tossing a fair die, what is the probability of getting an odd number or a number less than 4?

Solution

Let A= Odd number and B =Number less than 4

$$S = \{1, 2, 3, 4,5\ 6\}, A = \{1. 3, 5\}, B = \{1, 2, 3\}$$

$$A \cap B = \{1, 3\}$$

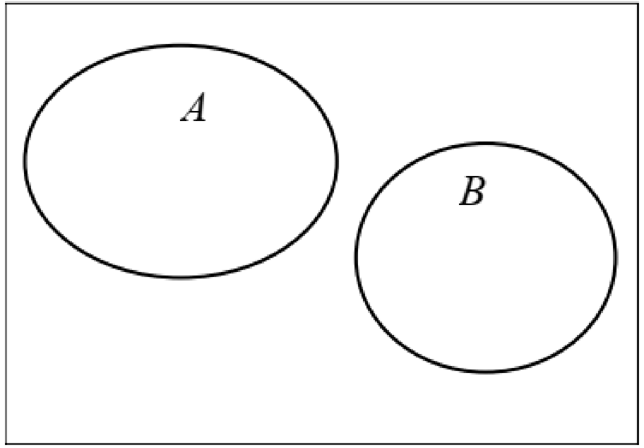
$$P(A\ or\ B) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$= \frac{3}{6} + \frac{3}{6} - \frac{2}{6}$$

$$= \frac{4}{6} = \frac{2}{3}$$

PROBABILITY LAWS

5. Addition Rule for disjoint or mutually exclusive events



A and B are Disjoint or Mutually Exclusive

$$A \cap B = \emptyset \; and \; P(A \cap) = 0$$

$$P(A \; or \; B) = P(A) + P(B)$$

$$\boldsymbol{P(A \cup B) = P(A) + P(B)}$$

If $A_1, A_2, A_3, \ldots$ *are disjoint or mutually exclusive events* then

$$\mathbf{P}(\boldsymbol{A_1 \cap A_2 \cap A_3 \cap \cdots}) = \mathbf{0} \quad \textbf{and}$$

$$\boldsymbol{P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots}$$

Example

If the probability that on any workday a workshop will get 10-20, 21-30, 31-40, over 40 customers to serve is 0.20, 0.35, 0.25, 0.12, respectively. What is the probability that on a given workday the workshop gets

a) At least 21 customers?
b) 0-9 customers?

Solution

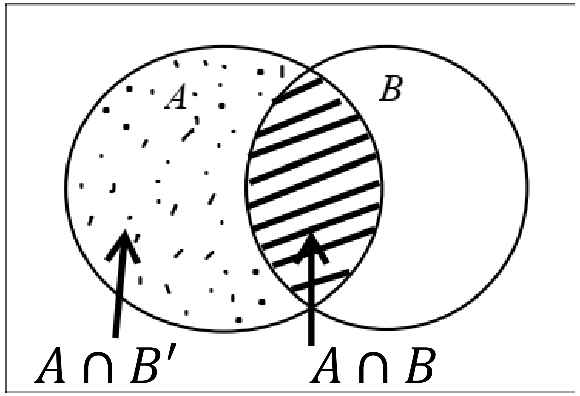
a) Since the events are mutually exclusive

$P(\geq 21\, customers) =$ 0.35+0.25+0.12=0.72

b) $P(0-9) = 1 - (0.20 +$ 0.35+0.25+0.12=0.72=1−0.92=0.08)

PROBABILITY LAWS

6. Total Probability Law: For any events A and B



$$\boldsymbol{P(A) = P(A \cap B) + P(A \cap B')}$$

Example

Given that $P(A \cap B) = \frac{1}{6}, P(A' \cap B) = \frac{2}{5}$

Find $P(B)\ and\ P(B')$

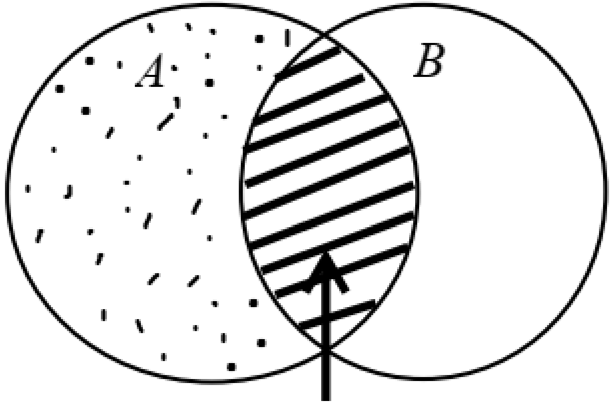
Solution

$$P(B) = P(A \cap B) + P(A' \cap B)$$
$$= \frac{1}{6} \qquad + \qquad \frac{2}{5} \qquad = \frac{17}{30}$$

$$P(B') = 1 - P(B) = 1 - \frac{17}{30} = \frac{13}{30}$$

PROBABILITY LAWS

7. Conditional Probability



$A \cap B$

$$P(A\ given\ B) = \frac{P(A\ and\ B)}{P(B)}\text{, where } P(B) \neq 0$$

$$P(B \mid A) \equiv \frac{P(A \cap B)}{P(A)}$$

$$P(A \cap B) \equiv P(A)\, P(B \mid A)$$

8. Multiplication rule for 2 events A and B

$$P(A\ and\ B) = P(A)P(B/A)$$

$$P(A\ and\ B) = P(B)P(A/B)$$

Probability Laws

Example

In producing screws, let A mean "screw too slim" and B "screw too short". Let $P(A) = 0.1$ and let the conditional probability that a slim screw is also too short be $P(B/A) = 0.2$. What is the probability that a screw that we pick randomly from the lot produced will be both too slim and too short?

Solution

$$P(A \cap B) = P(A)P(B/A) = 0.1 \times 0.2 = 0.02$$

PROBABILITY LAWS

9. Independent Events :

A and B are independent iff $\boldsymbol{P(A \cap B) = P(A) \times P(B)}$

or

$$\boldsymbol{P(A/B) = P(A)}$$

Example

Two screws are drawn at random from a box containing 10 screws, three of which are defective. Find the probability that neither of the two screws is defective if

a) The screws are drawn with replacement.

b) The screws are drawn without replacement

PROBABILITY LAWS

Solution

Let D=Defective and D′= Not defective.

a) For drawing with replacement, D_1 and D_2 are independent

$$P(D_1 \cap D_2) = P(D_1) \times P(D_2) = \frac{7}{10} \times \frac{7}{10} = \frac{49}{100}$$

b) For drawing without replacement, D_1 and D_2 are dependent

$$P(D_1 \cap D_2) = P(D_1)P(D_2/D_1) = \frac{7}{10} \times \frac{6}{9} = \frac{7}{15}$$

PROBABILITY LAWS

Combinatorial Analysis

Fundamental Principles of counting

If one trial can be accomplished in n_1 ways, the second trial in n_2 ways … and the kth trial in n_k ways then all the k trials can be accomplished in the order in

$$\boldsymbol{n_1 \times n_2 \times \cdots \times n_k\, ways}$$

Example

If a man has 3 pairs of shoes, 2 pairs of trousers and 5 shirts, find the number of ways in which he can choose a pair of shoes, a pair of trousers and a shirt to wear.

Solution

There are $3 \times 2 \times 5 = 30\, ways$

PROBABILITY LAWS

Number of arrangements in a line

n different items can be arranged in a line in

$$n!\, ways$$

Where $n! = product\ of\ all\ natural\ numbers\ from\ n\ down\ to\ 1\ i.e.$

$$n! = n \times (n-1) \ldots \times 3 \times 2 \times 1 \text{ and } 0! = 1$$

$$(n+1)! = (n+1)n!.$$

Example 1: Find the number of ways of arranging 5 laptops on a shelf.

Solution: $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$

Example 2; In how many ways can 3 teachers out 7 seat on 3 chairs arranged in a line?

Solution: There are 7 ways of occupying the 1st position, 6 ways of occupying the 2nd and 5 ways of occupying the 3rd.

Total number of ways$= 7 \times 6 \times 5 = 210\, ways$

Probability Laws

Arrangement of n different items in a circle

Fixed one and arrange the remaining $(n-1)$ items relative to it as in a line.

Thus n different objects can be arranged in a circle in

$$(n-1)! \quad ways$$

Example

In how many ways can 7 people be arranged at a round table.

Solution

$$(7-1)! = 6! = 720 \quad ways$$

Probability Laws

Permutation or Arrangement

The number of permutations of n different items taken r at a time or number of arrangements/permutations of r objects out of n different objects in a line

- **without repetitions is**

$$^{n}P_{r} = \frac{n!}{(n-r)!}, n \geq r$$

- **and with repetitions is** n^r

PROBABILITY LAWS

Example

Calculate the number of permutations there are of:

a) Five distinct objects taken two at a time. b) Seven distinct objects taken three at a time.

Solution

a) $^5P_2 = \frac{5!}{(5-2)!} = \frac{5!}{3!} = \frac{5\times4\times3!}{3!} = 5 \times 4 = 20$

b) $^7P_3 = \frac{7!}{(7-3)!} = \frac{7!}{4!} = \frac{7\times6\times5\times4!}{4!} = 7 \times 6 \times 5 = 210$

Example

Find the number of 3 digit numbers can be formed using the digits 2, 4, 5, 6

a) If repetitions are not allowed. b) If repetitions are allowed.

Solution

a) 3 digits are arranged out of 5 different digits in

$$^{5}P_3 = \frac{5!}{(5-3)!} = \frac{5!}{2!} = \frac{5 \times 4 \times 3 \times 2!}{2!} = 5 \times 4 \times 3 = 60\ ways$$

b) If repetitions are allowed each of the 3 digits can be obtained in 5 ways

$$\Longrightarrow 5^3 = 125\ numbers$$

PROBABILITY LAWS

Combination or Selection

The number of combinations or selections of n objects taken r at a time is

$$ {}^{n}\boldsymbol{C}_{\boldsymbol{r}} = \frac{\boldsymbol{n}!}{\boldsymbol{r}!\,(\boldsymbol{n}-\boldsymbol{r})!}\,, \qquad \boldsymbol{r} \leq \boldsymbol{n} $$

${}^{n}C_{r}$ is the number of ways in which r items can be selected from n different items

NB: ${}^{n}C_{0} = {}^{n}C_{n} = 1$ ${}^{n}C_{n} = {}^{n}C_{1} = 1$

Example

Find the number of samples of 3 sim cards that can be selected from 8 sim cards.

Solution:

$$ {}^{8}C_{3} = \frac{8!}{3!\,(8-3)!} = \frac{8!}{3!\,5!} = \frac{8 \times 7 \times 6 \times 5!}{3 \times 2 \times 1 \times 5!} = \frac{8 \times 7 \times 6}{3 \times 2 \times 1} = 56 $$

PROBABILITY LAWS

Example

Find the number of samples of 5 light bulbs that can be selected from 50 bulbs.

Solution: $^{50}C_5 = \frac{50!}{5!(50-5)!} = \frac{50!}{5!45!}$

$$= \frac{50 \times 49 \times 48 \times 47 \times 46 \times 45!}{5 \times 4 \times 3 \times 2 \times 1 \times 45!} = 2118760$$

Exercise

a) Find the number of ways of selecting a committee of 7 people from 10 engineers, 5 physicists and 6 computer scientists

b) How many of these committees consist of 3 engineers, 2 physicists and 2 computer scientists.

c) Find the probability that a committee selected from these committees at random consist of 3 engineers, 2 physicists and 2 computer scientists.

Solution

a) 7 people are selected from 21 in ${}^{21}C_7 = 116280$ ways

b) 3 engineers are selected from in ${}^{10}C_3 = 120\ ways$

2 physicists from 5 in ${}^{5}C_2 = 10\ ways$

2 computer scientists from 6 in ${}^{6}C_2 = 15\ ways$

Number of committees with 3 engineers, 2 physicists and 2 computer scientists$= 120 + 10 + 15 = 145\ committees$

c) P(3 engineers, 2 physicists, 2 computer scientists)$= \frac{145}{116280} = \frac{29}{23256}$

PROBABILITY LAWS

Number of arrangements of n items r, s and t of which are alike or the same

$$\frac{\boldsymbol{n}!}{\boldsymbol{r}! \times \boldsymbol{s}! \times \boldsymbol{t}!} \ \boldsymbol{ways}$$

Example:

Find the number of possible arrangements of the letters of the word ENGINEER

Solution

8 letters with 3 E's and 2 N arranged in $\frac{8!}{3!\times 2!} = \frac{8\times 7\times 6\times 5\times 4\times 3!}{3!\times 2\times 1} = 3360$

Example

A committee of 5 people is to be selected from a group of 4 men and 6 women. What is the probability that the committee contains

a) 2 men and 3 women. b) at least 1 man.

Solution

n(S)=Number of ways of selecting 5 people from 10$= {}^{10}C_5$ =252

a) n(2M and 3W)$= {}^{4}C_2 \times {}^{6}C_3 = 120$

$$P(2M3W)= \frac{n(2M \text{ and } 3W)}{n(S)} = \frac{120}{252} = \frac{10}{21}$$

b) P(at least 1M)$= 1 - P(No\ M) = 1 - P(5W) = 1 - \frac{n(5W)}{n(S)}$

$$n(5W) = {}^{6}C_5 = 6$$

Hence, P(at least 1M)$= 1 - \frac{6}{252} = \frac{41}{42}$

Example

Five laptops are labelled A, B , C, D and E. If these laptops are arranged on a shelf. Find the probability that A and B are together.

Solution

5 laptops can be arranged on a shelf in $n(S) = 5! = 120$ways

Numbers arrangements with A and B together $4! \times 2! = 48\ ways$

$$P(A\ and\ B\ together) = \frac{48}{120} = \frac{2}{5}$$

Example

The probability of a car repair being on time is 0.40. The probability of a car repair being satisfactory is 0.50. The probability that a car repair is neither satisfactory nor on time is 0.25. What is the probability of a repair being satisfactory and on time?

Solution

A=repair is on time P(0.40) B=the repair is satisfactory P(B)=0.50

$$P[(A \cup B)'] = 0.25 \Longrightarrow P(A \cup B) = 1 - 0.25 = 0.75$$

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.15$$

STUDY QUESTIONS

Of a lot of 10 items, 2 are defective. **(a)** Find the number of different samples of 4. Find the number of samples of 4 containing **(b)** no defectives, **(c)** 1 defective, **(d)** 2 defectives.

An urn contains 2 green, 3 yellow, and 5 red balls. We draw 1 ball at random and put it aside. Then we draw the next ball, and so on. Find the probability of drawing at first the 2 green balls, then the 3 yellow ones, and finally the red ones.

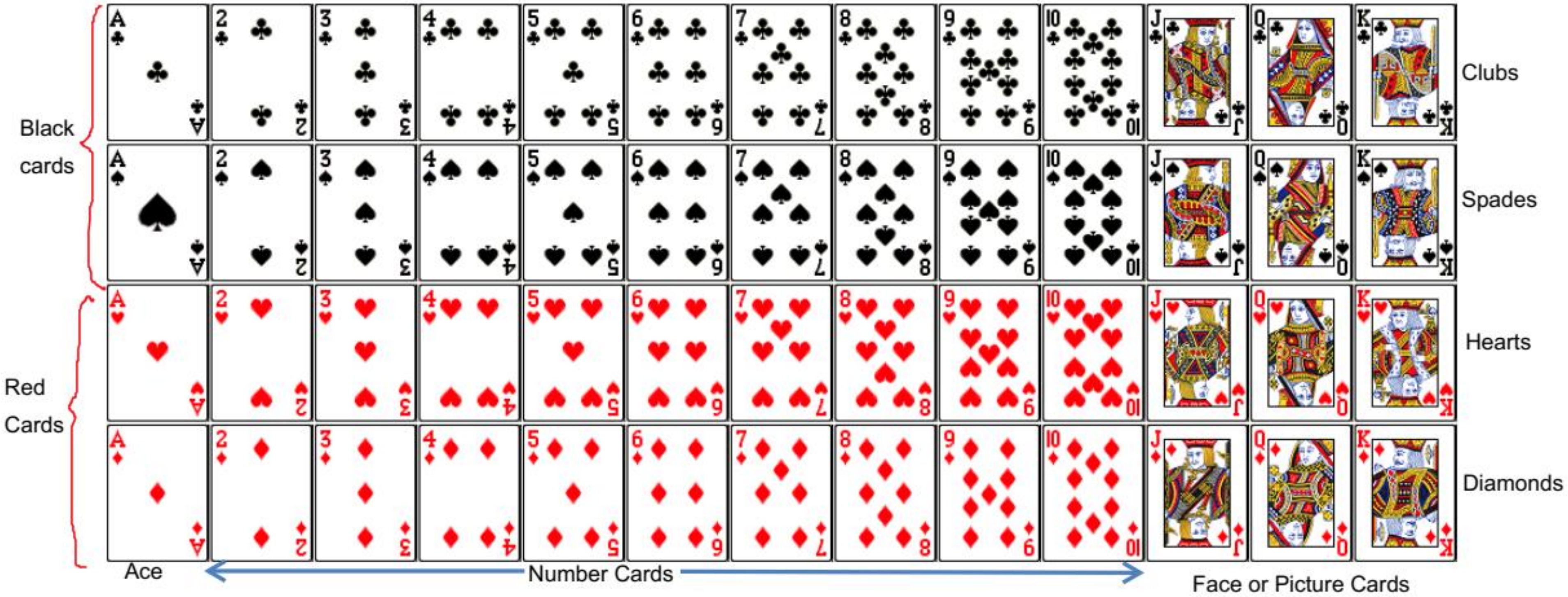
Problem 3. Calculate the probabilities of selecting at random:

(a) the winning horse in a race in which ten horses are running,

(b) the winning horses in both the first and second races if there are ten horses in each race.

Problem 4. The probability of a component failing in one year due to excessive temperature is $\frac{1}{20}$, due to excessive vibration is $\frac{1}{25}$ and due to excessive humidity is $\frac{1}{50}$. Determine the probabilities that during a one-year period a component: (a) fails due to excessive temperature and excessive vibration, (b) fails due to excessive vibration or excessive humidity, and (c) will not fail because of both excessive temperature and excessive humidity.

In a batch of 45 lamps there are ten faulty lamps. If one lamp is drawn at random, find the probability of it being (a) faulty and (b) satisfactory.



Two cards are drawn from a pack of playing cards without replacement. Find the probability that:

a) Both of them are face cards.

b) b) Both of them are spades given that they are face cards.

SPECIAL RANDOM VARIABLE

Given a question involving a r.v.:

- Define the random variable
- Identify the distribution and parameters
- Determine the mean, variance and standard deviation
- Reason out each question to identify the appropriate formula

SPECIAL DISCRETE DISTRIBUTIONS

A. THE DISCRETE UNIFORM OR RECTANGULAR DISTRIBUTION

If X is a d.r.v which is equally likely to take any of its possible n values then

$X \sim \operatorname{Rec}(n)$

P.d.f : $P(X=x)=\dfrac{1}{n},\ x=1,\ 2,\ 3,\ \ldots,n$

$E(X)=\dfrac{n+1}{2} \qquad Var(X)=\dfrac{n^2-1}{12}$

Example 1

X is the score on a fair 10 sided spinner.

a) Find $P(2<X\le 4)$
b) Define $Y=5X+3$. Find the mean and variance of Y

B. THE GEOMETRIC DISTRIBUTION

A r.v. X has a geometric distribution with parameter p i.e $X \sim Geo(p)$ iff X is the number of **independent trials** upto and including the first success, each trial being either a success with fixed probability p or failure with fixed probabiliy q where $q=1-p$

If $X \sim Geo(p),\ q=1-p$

$P.d.f:\ P(X=x)=q^{x-1}p$ for $x=1,2,3,\ldots$

$E(X)=\dfrac{1}{p} \qquad Var(X)=\dfrac{q}{p^2} \qquad$ **Mode = 1**

The distribution function:

$F(x)=P(X\le x)=1-q^x$ **Proof**

$\Rightarrow P(X>x)=q^x$ (complementrary events)

Example B1

A man and his wife decide to continue to have children until a boy is born. Given that the probability of having a boy for this family is 0.17. Find:

a) The probability that the first boy borne will be the third child.
b) The probability that they will have more than 2 children to have a boy.
c) The smallest value of n for which there is at least 95% chance of having a boy on or before the n^{th} birth.

Solution

Let X =No. of children up to and including the first boy.

$X{\sim}Geo(0.17) \quad p=0.17, q=1-p=0.83$

$P(X=x)=q^{x-1}p\ for\ x=1,2,3,\ldots$

a) $P(X=3)=0.83^2(0.17)=0.1171$
b) $P(X>x)=q^x$

   $P(X>2)=0.83^2=0.6889$
c) $P(X\le n)\ge 0.95$

   $1-q^n\ge 0.95$

   $0.83^n\ge 0.05$

   $n\ln(0.83)\ge \ln(0.05)$

   $n\ge \dfrac{\ln(0.05)}{\ln(0.83)}$

   $n\ge 16.08$

   $n\in\mathbb{N} \Longrightarrow Least\ value\ of\ n\ is\ 17$

Exercise

A student continues trying to recover a corrupt document from his laptop. If the probability of success in each trial is 0.75.

Find

a) the probability that he will succeed in the third attempt.
b) The mean and variance of number of trials needed for him to succeed.
c) The most likely number of attempts until he succeeds.

C. THE BINOMIAL DISTRIBUTION

A random variable X follows a binomial distribution with parameters n and p iff X is the number of successes in n independent trials, each trial being either a success with probability p or failure with probability q where $(q=1-p)$.

If $X \sim Bin(n,p)$ then:

$P.d.f:\ P(X=x)={}^nC_x q^{n-x}p^x,\ x=0,1,2,3,\ldots,n$

$E(X)=np \qquad Var(X)=npq$

Example C1

In a mass production of a certain type of device, it is found that 5% are defective. Devices are selected at random and packed in boxes, each containing 10 such devices.

i. If one box is selected at random, calculate to 3 decimal places, the probability that there will be

a) exactly 2 defectives,

b) more than one defective

ii. If 2 boxes are selected at random, find the probability that there are no defective devices in either box.

iii. If 3 boxes are selected, calculate the probability continuous random variable X has probability density function f given that exactly one box will contain just 2 defectives.

DI. THE POISSON DISTRIBUTION

A discrete r.v X follows a Poisson distribution with parameter λ i.e $X \sim Po(\lambda)$ iff X is the number of independent events occurring in a given interval with a constant mean of λ.

If $X \sim Po(\lambda)$ then

$$P.d.f: \mathrm{P}(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \ x = 0,1,2,...$$

$$E(X) = \lambda \qquad \mathrm{Var}(X) = \lambda$$

Mode

$$\begin{cases} \lambda \in \mathbb{Z} \Rightarrow Mode = \lambda - 1 \text{ and } \lambda \\ \lambda \notin \mathbb{Z} \Rightarrow Mode = \text{Greatest integer less than } \lambda \end{cases}$$

For n units of interval $X \sim Po(n\lambda)$

If $X \sim Po(\lambda)$ and $Y \sim Po(\mu)$ are independent then $X + Y \sim Po(\lambda + \mu)$

Example D1

A shop sells a particular make of cell phones at a rate of 4 per week on average. The number sold in a week has a Poisson distribution.

a) Fnd the probability that the shop sells exactly 2 in a week.

b) Find the probability that the shop sells at least 1 in a week.

c) Find the number of cell phones that the shop expects to sell in one month.

Exercise DI

The number of flaws in a fibre optic cable follows a Poisson distribution. The average number of flaws in 50m of cable is 1.2

(i) What is the probability of exactly three flaws in 150m of cable? 0.212

(ii) What is the probability of at least two flaws in 100m of cable? 0.691

(iii) What is the probability of exactly one flaw in the first 50m of cable and exactly one flaw in the second 50m of cable? (0.361) (0.361)=0.13

DII The Poisson Approximation to the Binomial

If $\boxed{np < 5}$ OR n is large and p is small (n >50 and p<0.15) then $X \sim \mathrm{Bin}(n, p) \approx X \sim Po(np)$

Exercise D2

It is known that 3% of the circuit boards from a production line are defective. If a random sample of 120 circuit boards is taken from this production line. Use the Poisson approximation to extimate the probability that the sample contains

(i) Exactly 2 defective boards 0.177

(ii) At least 2 defective boards 0.878

SPECIAL CONTINUOUS RANDOM VARIABLES

A. The Uniform Rectangular Distribution

If the random variable X is equally likely to take any value in the interval $a \le x \le b$ then X follows a uniform rectangular distribution with parameters a and b.

If $X \sim \mathrm{Re}c(a,b)$ then

$$p.d.f: f(x) = \frac{1}{b-a}, \ a \le x \le b$$

$$E(X) = \frac{a+b}{2} \qquad \mathrm{Var}(X) = \frac{(b-a)^2}{12}$$

Exercise

A random variable X is such that $X \sim Rec(3,6)$

Find a) the p.d.f. of X.

b) P(2<X<4)

c) P(X>5)

c) E(X) and Var(X)

B. THE EXPONENTIAL DISTRIBUTION

A continuous r.v X follows an exponential distribution with parameter λ iff X is the waiting time between events that follow a Poisson distribution with parameter λ.

If $X \sim Exp(\lambda)$ then

$$\text{P.d.f.} \quad f(x) = \begin{cases} \lambda e^{-\lambda x}, x \ge 0 \\ 0, \quad \text{otherwise} \end{cases}$$

$$\mathrm{E}(X) = \frac{1}{\lambda} \qquad \mathrm{Var}(X) = \frac{1}{\lambda^2}$$

$$Mode = 0$$

Example B1

The life time in years, of light bulbs of a certain make is a random variable T with probability density function f given by

$$f(x) = \begin{cases} Ae^{-\lambda t} & t \geq 0, \lambda > 0 \\ 0, & otherwise \end{cases}$$ where A and λ are constants

a) Find A in terms of λ
b) If is found that of 300 such bulbs, 45 failed within one year of use. Estimate to 4 decimal places, the value of λ.
c) Calculate, to 2 decimal places, the mean and standard deviation of T.

Exercise B

The time, t seconds, between the arrivals of successive vehicles at a particular point has p.d.f.

$$f(t) = \begin{cases} 0.025e^{-0.025t} & t \geq 0 \\ 0, & t < 0 \end{cases}$$

A woman who takes 20 seconds to cross the road, sets off as one vehicle passes.

Find the probability that she will complete the crossing before the next vehicle arrives.

C. THE NORMAL DISTRIBUTION

If a c.r.v. has p.d.f

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}}, x \in]-\infty, +\infty[$$

Then X follows a normal distribution with parameters μ and σ^2 i.e. $X \sim N(\mu, \sigma^2)$

The graph of the p.d.f. is called the **normal distribution curve** and it has the following characteristics:

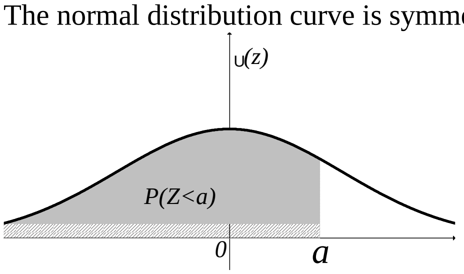
1) It is bell-shaped.
2) It is symmetrical about the mean.i.e. mode=median=mean
3) Total area under the normal curve is 1.
4) The mean divides the area into two equal parts.

The standard normal variable

The standard normal variable $z = \dfrac{x-\mu}{\sigma}$ has mean 0 and variance $1 \Rightarrow Z \sim N(0,1)$

$$P.d.f : \phi(z) = \frac{1}{2\pi} e^{-\frac{z^2}{2}}, z \in]-\infty, +\infty[$$

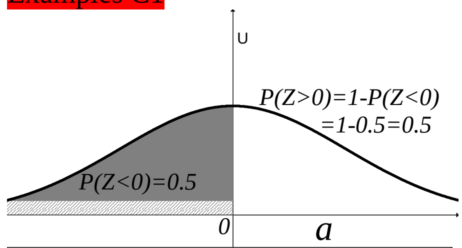
The normal distribution curve is symmetrical about mean 0

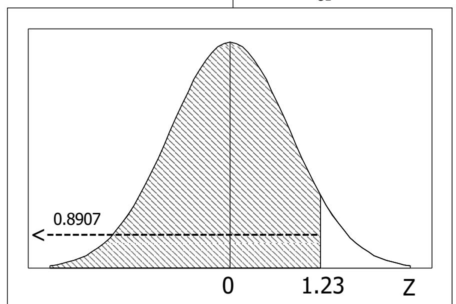


Shaded area is $P(Z<a) = \phi(a) = \int_{-\infty}^{a} \phi(z)\,dz$ which is obtained using the normal distribution table or the calculator.

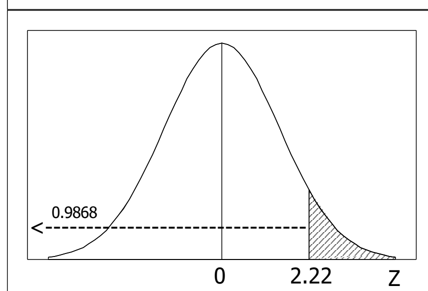
NB The total area under the normal curve is 1 and half of the area is 0.5.

Examples C1





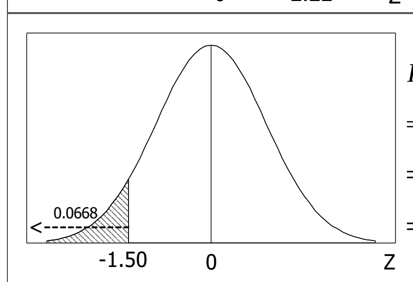
$P(z<1.23)$
$=\Phi(1.23)$
$=0.8907$



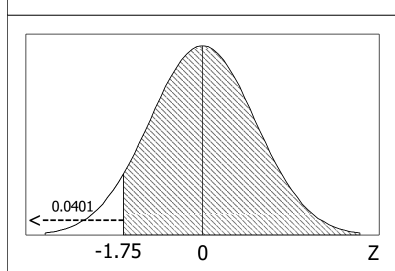
P(z>2.22)

= 1 – P(z<2.22) =1− 0.9868

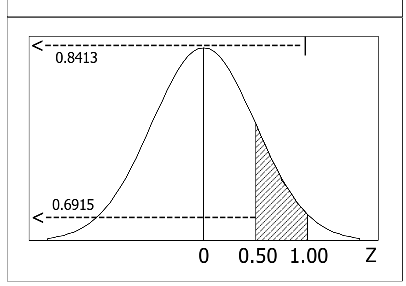
= 0.0132



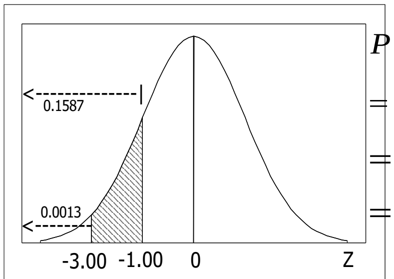
$P(z<-1.5)$
$=P(z>1.5)$
$=1-\Phi(1.5)$
$=1-0.9332=0.0668$



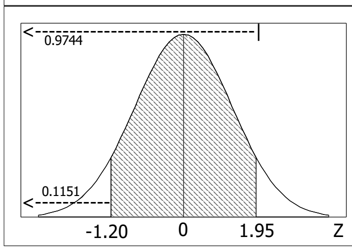
$P(z>-1.75)$
$=P(z<1.75)$
$=\Phi(1.75)$
$=0.0401$



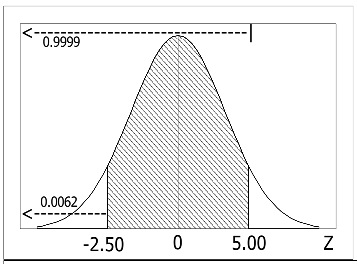
$P(0.50<z<1.00)$
$=\Phi(1.00)-\Phi(0.5)$
$=0.8413-0.6915$
$=0.1498$



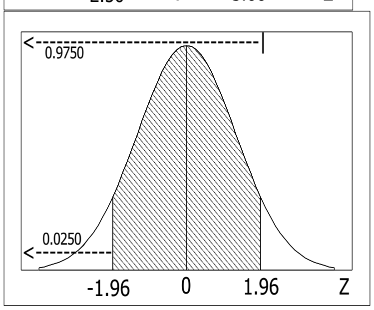
$$P(-3.00<z<-1.00)$$
$$=\Phi(3.00)-\Phi(1.00)$$
$$=0.9987-0.8413$$
$$=0.1574$$



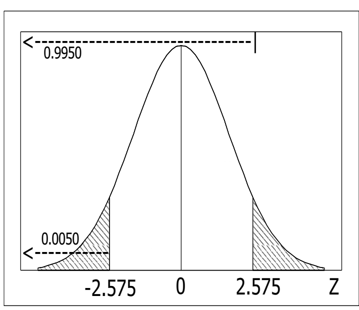
$$P(-1.20<z<-1.95)$$
$$=[\Phi(1.95)-0.5]$$
$$+[\Phi(1.20)-0.5]$$
$$=\Phi(1.95)+\Phi(1.20)-1$$
$$=0.9744+0.8849-1$$
$$=8593$$



$$P(-2.50<z<5.00)$$
$$=\Phi(5.00)+\Phi(2.5)-1$$
$$=0.9999+0.9938-1$$
$$=0.9937$$



$$P(|z|<1.96)$$
$$=P(-1.96<z<1.96)$$
$$=2[\Phi(1.96)-0.5]$$
$$=2\Phi(1.96)-1$$
$$=2(0.975)-1$$
$$=0.95$$



$$P(|z|>2.575)$$
$$=P(Z<-2.575)$$
$$+P(z<2.575)$$
$$=2[1-\Phi(2.575)]$$
$$=0.010$$

How to find normal probabilities

Given $X \sim N(\mu,\sigma^2)$ to find a probability:

1. Standardize X using $Z=\dfrac{x-\mu}{\sigma}$ in 3 d.p.

2. Use the normal table or calculator.

Example C2

The weight, X grams, of the contents of a tin of baked beans can be modelled by a normal random variable with mean of 421 and standard deviation of 2.3.

a) Find

i. $P(X=423)$ ii. $P(X<425)$ iii. $P(X<413)$

iv. $P(418<X<424)$ v. $P(X>425)$ vi. $P(|X|>427)$

b) Determine the value of X such that $P(X<x)=0.98$

c) The weight, Y grams, of the contents of a tin of milk can be modelled by a normal random variable with a mean of μ such that $P(Y<410)=0.01$.

Solution hint

$$X \sim N(421, 2.3^2) \Rightarrow \mu=421, \sigma=2.3$$

$$Z=\frac{X-\mu}{\sigma} \Rightarrow Z=\frac{X-421}{2.3}$$

Exercise C1

The marks X in a test are normally distributed with mean 40 and standard deviation 10. Find, to 4 decimal places, the probability that a randomly chosen candidate scored

i.	Less than 65 marks	3mks
ii.	Between 24 and 40 marks	3mks

CII Find the mean or standard deviation

Example C3

The lengths of cucumbers from a certain greenhouse are normally distributed. If 10% of the cucumbers are longer than 35cm and 10% are shorter than 25cm find the mean and standard deviation of the lengths

ANS: 30cm, 3.90cm

Example

The masses of particles produced in a particular factory are normally distributed with mean
μ and standard deviation σ. Given that 5% of the particles have a mass greater than 85g and 10% have a mass less than 25g.
Find the values of μ *and* σ

CIII The Normal Distribution of sums, differences and products

If X and Y are independent normal variables and a and b are real constants then

$X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$

1. X+Y $\sim N(\mu_X+\mu_Y, \sigma_X^2+\sigma_Y^2)$

2. $X-Y \sim N(\mu_X-\mu_Y, \sigma_X^2+\sigma_Y^2)$

3. X+X $\sim N(2\mu_X, 2\sigma_X^2)$

4. 2X $\sim N(2\mu_X, 2^2\sigma_X^2)$

5. aX+bY $\sim N(a\mu_X+b\mu_Y, a^2\sigma_X^2+b^2\sigma_Y^2)$

6.5. aX − bY $\sim N(a\mu_X-b\mu_Y, a^2\sigma_X^2+b^2\sigma_Y^2)$

CIV Normal Approximations to discrete distributions

Discrete	**Condition**	**Approximation**
$X \sim Bin(n, p)$ $\mu = np, \sigma^2 = npq$	n is large and p is small $n \geq 10 \wedge p \approx 0.5$ $OR\ n > 30$	$X \sim Bin(n, p)$ $\approx X \sim N(np, npq)$
$X \sim Po(\lambda)$ $\mu = \lambda, \quad \sigma^2 = \lambda$	$\lambda > 20$	$X \sim Po(\lambda)$ $\approx X \sim N(\lambda, \lambda)$

Continuity Correction

To compensate for the change from the discrete to the continuous scale the continuity correction is as follows

Discrete scale	Continuous scale
$x = 3$	$2.5 < x < 3.5$
$x < 3$	$x < 2.5$
$x \leq 3$	$x < 3.5$
$x \geq 3$	$x > 2.5$
$x > 3$	$x > 3.5$

Example C6

(i) 5% of the students in a certain school study on scholarship. 60 students are chosen at random from the school. Use Poison distribution to find to 3 significant figures, the probability that at least 2 of them study on scholarship.

(ii) 500 students are chosen at random from the school. Use the normal approximation to estimate to 3 significant figures the probability that between 20 and 30 students inclusive will be studying on scholarship.