Probability

This is a measure of how likely it is that something will occur. In talking about probability we need an experiment.

An **Experiment** is any action with outcomes that are recorded data. The number of times we do it is sample space. It is the set of all possible outcomes of an experiment. It is denoted by “S”. When a coin is tossed twice, first outcome is **{H,T}**, second outcome is **{HT,TH,TT,HH}**. The sample space is therefore four for the example given. We have ;

> **Finite Sample Space:** A sample space which takes integer values or has countable number.
>
> **Infinite Sample Space:** the ages of a class can range from 17 to 30 i.e. $17 \leq x \leq 30$. An individual can start from 17 and start counting 18, 19… Another person can use 17.01, 17.02, 17.03… etc. This makes it infinite.

Event

An event **‘A’** is an outcome or the set of outcomes that are of interest to the experimenter. The probability that an event A would occur is written as $\boldsymbol{P(A)}$ and is” read probability of A”. The probability of an event A, $\boldsymbol{P(A)}$ is a measure of the likelihood that an event A would occur i.e

$$\boldsymbol{P(A) = \frac{number\ of\ ways\ that\ A\ can\ occur}{Total\ number\ of\ possible\ outcomes}}$$

Question

An ordinary die is thrown. Find the probability that the number obtained

a. Is a multiple of 3
b. Less than 7
c. A factor of six

Solution

Sample space when die is thrown = {1, 2, 3, 4, 5, 6}

a. P(multiple of 3) = $\frac{2}{6} = \frac{1}{3}$
b. P(less than 7) = $\frac{6}{6} = 1$
c. P(factor of six) = $\frac{4}{6} = \frac{2}{3}$

Compliment of an Event

The compliment of an event A is denoted by $\boldsymbol{A' or \bar{A}}$. If the set of all outcomes in the sample space **"S"** , that do not correspond to an event **"A"** .

$$\boldsymbol{P(A) + P(A') = 1}$$

Example

Sample space S = {1, 2, 3, 4…10}

A = {even prime number}

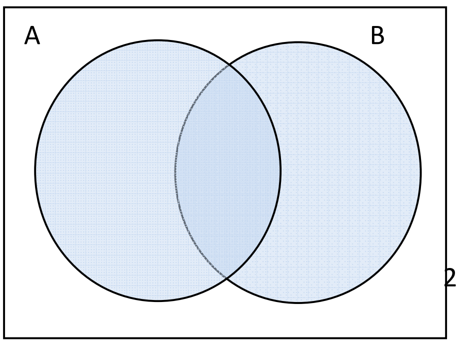
A'= {1,3,4,5,6,7,8,9,10}

Example

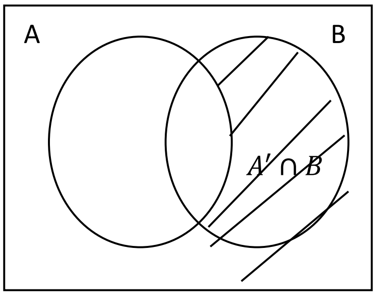
The probability of an event occurring is 0.27. What is the probability that it will not occur.

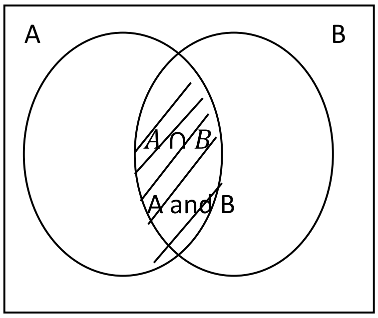
$$ans = 1 - 0.27 = 0.73$$

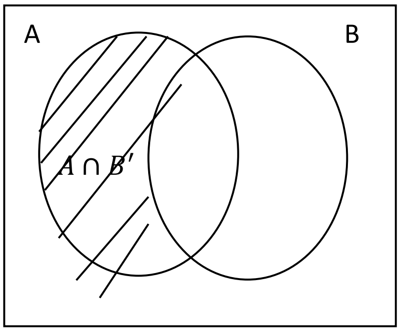
ILLUSTRATING TWO OR MORE EVENTS USING VENN DIAGRAMS



$(A \cup B)$







Probability Rule

$$\boldsymbol{n(A \cup B) = n(A) + n(B) - n(A \cap B) \ldots\ldots\ldots\ldots.(1)}$$

$$divide\ equation\ (1)\ by\ S\ (sample\ space)\quad i.e\ \frac{event}{sample\ space} = prob\ of\ that\ event$$

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

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

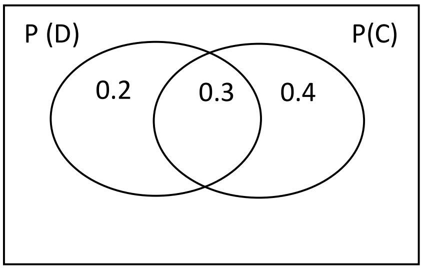
$$\boldsymbol{P(B) = P(A \cap B) + P(A' \cap B)}$$

Question

For events C and D $P(C) = 0.7$ $P(D \cup C) = 0.9$ $P(C \cap D) = 0.3$ find

a. $P(D)$ b. $P(D' \cap C)$ c. $P(D \cap C')$ d. $P(D' \cap C')$

Solution



$S = 1$

a. $0.9 = 0.7 + P(D) - 0.3$

   $P(D) = 0.5$

b. $P(C) = P(C \cap D) + P(C \cap D')$

   $0.7 = 0.3 + P(C \cap D')$ Thus $P(C \cap D') = 0.4$

c. $P(D) = P(C \cap D) + P(C' \cap D)$

   $0.5 = 0.3 + P(C' \cap D)$

   $P(C' \cap D) = 0.2$

d. $P(D' \cap C') = 1 - 0.9 = 0.1$

Question

Given that $P(A') = \frac{2}{3}$, probability of $B = \frac{1}{2}$ and $P(A \cap B) = \frac{1}{12}$. Find $P(A \cup B)$

Solution

$P(A \cup B) = [(1 - P(A)')] + P(B) - P(A \cap B)$

$$= \left[\left(1 - \frac{2}{3}\right)\right] + \frac{1}{2} - \frac{1}{12} = \frac{1}{3} + \frac{5}{12} = \frac{3}{4}$$

EXHAUSTIVE EVENTS

If two events A and B are such that between them, they make the whole of the possibility space, then A and B are said to be Exhaustive events and $P(A \cup B) = 1$

$\Longrightarrow 1 - P(A) + P(B) - P(A \cap B)$

Example

$S = \{1,2, \ldots ,10\}$

$A = \{2,4,6, \ldots ,10\}$

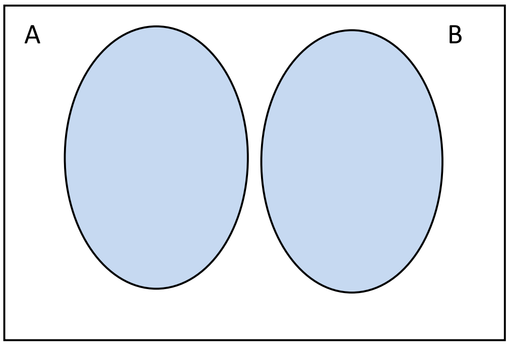
$B = \{1,3,5, \ldots 9\}$

$(A \cup B) = \{1,2,3,4, \ldots ,10\} = S$

EXCLUSIVE OR MUTUALLY EXCLUSIVE EVENTS

A and B are said to be exclusive or mutually exclusive events if there is no intersection between them; i.e. if they cannot occur at the same time. It is expressed mathematically as

$P(A \cup B) = P(A) + P(B)$ i.e P(A$\cap B) = 0$



Question

It is known that $P(X) = \frac{1}{2}$ and $P(Y) = \frac{1}{4}$, given that X and Y are mutually exclusive, find

a. $P(X \cup Y)$ $b. P(Y \cap X)$ $c. P(Y \cap X')$ $d. P(Y' \cap X')$

Solution

a. $P(X \cup Y) = P(X) + P(Y)$

$$= \frac{1}{2} + \frac{1}{4}$$

$$= \frac{2+1}{4} = \frac{3}{4}$$

b. $P(Y \cap X) = 0$

c. $P(Y \cap X') = \frac{1}{4}$

DEMORGAN'S RULE

$\boldsymbol{P(A \cup B)' = 1 - P(A \cup B) = P(A' \cap B')}$

$\boldsymbol{P(A \cap B)' = 1 - P(A \cap B) = P(A' \cap B')}$

Question

Given that $P(A') = \frac{2}{3}$ then $P(B) = \frac{1}{2}$ then $P(A \cap B) = \frac{1}{12}$ find $P(A \cap B)$

Solution

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

$$= \frac{1}{3} + \frac{1}{2} - \frac{1}{12}$$

$$= \frac{4+6-1}{12} = \frac{9}{12} = \frac{3}{4} = \frac{3}{4}$$

Question

Events A and B are mutually exclusive and exhaustive events. $P(A) = 0.4$. Find

a. $P(B)$
b. $P(A \cap B)$

Solution

a. $P(A) + P(B) = 1$

$P(B) = 1 - 0.4 = 0.6$

c. $P(A \cap B) = 0\ since\ they\ are\ mutually\ exclusive$

Question

A and B are two events such that $P(A) = \frac{8}{15}$ then $P(B) = \frac{2}{3}$ and $P(A \cap B) = \frac{1}{5}$. Are A and B exhaustive events.

Solution

For exhaustive events

$P(A \cup B) = 1$

$$\therefore P(A \cap B) = \frac{2}{3} + \frac{8}{15} - \frac{1}{5}$$

$$= \frac{10 + 8 - 3}{15} = \frac{15}{15} = 1$$

INDEPENDENT EVENTS

If either of the events A and B can occur without being affected by the order, then the two (2) events are independent.

For independent events;

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

$$\Longrightarrow \boldsymbol{P(A \cap B) = P(A) \cdot P(B)}$$

Question

If events A and B are such that they are independent and $P(A) = 0.3$ and $P(B) = 0.5$. Find

a. $P(A \cap B)$
b. $P(A \cup B)$
c. Are A and B mutually exclusive

Solution

a. $P(A \cap B) = P(A) \cdot P(B)$
$= 0.3 \times 0.5 = 0.15$
b. $P(A \cup B) = P(A) + P(B) - P(A) \cdot P(B) = 0.3 + 0.5 - 0.15$
$= 0.8 - 0.15$
$= 0.65$
c. No

Question

Event A and B are such that $P(A) = 0.4$ and $P(B) = 0.25$. If A and B are independent events. Find

a. $P(A \cap B)$
b. $P(A' \cap B')$
c. $P(A \cap B')$

Solution

a. $P(A \cap B) = P(A) \cdot P(B)$
$= 0.4 \times 0.25$
$= 0.100 = 0.1$

b. $P(A' \cap B') = 1 - P(A \cup B)$

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

$= 0.4 + 0.25 - 0.1$

$P(A \cup B) = 0.55$

$\therefore P(A' \cap B') = 1 - 0.55$

$= 0.45$

c. $P(A \cap B') = P(A) - P(A \cap B)$

$= 0.4 - 0.1$

$= 0.3$

Question

The probability that an event A occurs, $P(A) = 0.4$, B is an event independent of A and the probability of the union of A and B $P(A \cup B) = 0.7$. Find $P(B)$.

Solution

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

$= P(A) + P(B) - [P(A) \cdot P(B)]$

$0.7 = 0.4 + P(B) - [0.4 \times P(B)]$

$0.3 = P(B) - 0.4P(B)$

$= P(B)[1 - 0.4]$

$0.3 = 0.6 \times P(B)$

$$P(B) = \frac{0.3}{0.6} = \frac{3}{6} = \frac{1}{2} = 0.5$$

Question

The event A and B are independent and are such that $P(A) = X$ then $P(B) = X + 0.2$ and $P(A \cap B) = 0.15$.

a. Find the value of X
b. Find $P(A \cup B)$

Solution

For independent events

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

$$0.15 = X \times (X + 0.2)$$
$$0.15 = X^2 + 0.2X$$
$$X^2 + 0.2X - 0.15 = 0$$
$$X = 0.3 \; or - 0.5$$

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

$$= 0.3 + 0.5 - 0.15$$
$$= 0.8 - 0.15$$
$$= 0.65$$

CONDITIONAL PROBABILITY

If A and B are two events not from the same experiment then the conditional probability that A occurs if given B has already occurred is written as $P(A, given\ B) = P(A/B)$

$$\boldsymbol{P(A, given\ B) = \frac{P(A \cap B)}{P(B)}}$$

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

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

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

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

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

Question

X and Y are two events such that $P(X/Y) = 0.4$ $P(Y) = 0.25$ and $P(X) = 0.2$. Find

a. $P(Y/X)$ $\quad b. P(Y \cap X)$ $\quad c. P(Y \cup X)$

Solution

a. $P(X/Y) \cdot P(Y) = P(Y/X) \cdot P(X)$

$$0.4 \times 0.25 = P(Y/X) \cdot 0.2$$

$$0.1 = P(Y/X) \cdot 0.2$$

$$P(Y/X) = \frac{0.1}{0.2} = 0.5$$

b. $P(Y \cap X) = P(Y/X) \cdot P(X)$

$= 0.5 \times 0.2$

$= 0.1$

c. $P(Y \cup X) = P(Y) + P(X) - P(Y \cap X)$

$= 0.25 + 0.2 - 0.1$

$= 0.35$

Question

If $P(A/B) = \frac{2}{5}$ and $P(B) = \frac{1}{4}$ and $(A) = \frac{1}{3}$. Find

a. $P(B/A)$ $\quad b. P(A \cap B)$

Solution

a. $P(B/A) \cdot P(A) = P(A/B) \cdot P(B)$

$$P(B/A) \times \frac{1}{3} = \frac{2}{5} \times \frac{1}{4}$$

$$= \frac{2}{20} = \frac{1}{10}$$

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

b. $P(A \cap B) = P(A/B) \cdot P(B)$

$$= \frac{2}{5} \times \frac{1}{4} = \frac{2}{20} = \frac{1}{10}$$

INDEPENDENT (CONDITIONAL EVENTS)

If A and B are independent, then;

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

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

$P(A'/B') = P(A')$

$P(B'/A') = P(B')$

Question

A and B are two independent events such that $P(A) = 0.2$ and $P(B) = 0.15$. Evaluate the following probabilities;

a. $P(A/B)$ $b. P(A \cap B)$ $c. P(A \cup B)$

Solution

For independent events

a. $P(A/B) = P(A) = 0.2$
b. $P(A \cap B) = 0.2 \times 0.15$
c. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$$= 0.2 + 0.15 - 0.03 \quad = 0.32$$

Question

A and B are exhaustive events and it is known that $P(A/B) = \frac{1}{4}, P(B) = \frac{2}{3}$. Find $P(A)$

Solution

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

$\frac{1}{4} \times \frac{2}{3} = \frac{1}{6}$

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

$$1 = P(A) + \frac{2}{3} - \frac{1}{6}$$

$P(A) = 1 + \frac{1}{6} - \frac{2}{3} = \frac{1}{2} = 0.5$

Question

Two events A and B are such that $P(A) = \frac{8}{15}, P(B) = \frac{1}{3}$ and $P(A/B) = \frac{1}{5}$. Calculate the probability that

a. Both events occur
b. Only one of the two events occurs
c. Neither events occur

PERMUTATION AND COMBINATION

Theorem 1

The number of ways of arranging n unlike objects in a line is n!

$$n! = n(n-1)(n-2)(n-3) \ldots 3X2X1$$

Example

$4! = 4 \times 3 \times 2 \times 1$

Question

Arrange the letters a, b, c, d in a line

$4! = 24$

Theorem 2

The number of ways of arranging in line n objects of which p are alike is $\frac{n!}{p!}$

Question

In how many ways can the letters of the word STATISTICS be arranged

Solution

$$= \frac{10!}{3!\,3!\,2!} = 50400\ ways$$

Theorem 3

The number of ways of arranging n unlike objects in a ring (circle) is $(\ n - 1)!$

Question

How many ways can 4 people be arranged on a round table?

Solution

$(4 - 1)! = 3! = 3X2X1$

It the letters of the word $MINIMUM$ are arranged in a line at random. What is the probability that the arrangements begin with the letter MMM.

Solution

$\frac{7!}{3!\,2!} = 420 ways\ can\ MINIMUM$ be arranged

If MMM are fixed the arrangement of the rest $(I, I, U, N) = \frac{4!}{2!} = 12 ways$

$$Probability = \frac{12}{420} = \frac{1}{35}$$

Question

The letters of the word MATHEMATICS are written ,one on each 11 separate cards. The cards are laid out in a line.

a. Calculate the number of different arrangements of these letters.
b. Determine the probability that the vowels are all placed together.

Solution

a. Number of arrangements

$$= \frac{11!}{2!\,2!\,2!} = 4989600\ ways$$

b. $M\ , T\ , H\ , M\ , T\ , C\ , S \qquad Vowel = A\ A\ E\ I$

$$= \frac{8!}{2!\,2!} \times \frac{4!}{2!} = 120960$$

$$Probability\ = \frac{120960}{4989600}$$

Question

How many ways can the letters of the word FACETIOUS be arranged in a line. What is the probability that the arrangement begins with ‘F’ and ends with ‘S?

Solution

$No\ of\ arrangement\ =\ 9! = 362880$

The ‘F’ and ‘S’, the rest is ACETIOU = 7! = 5040

$Probability\ =\frac{7!}{9!}=\frac{1}{72}$

Question

a. In how many ways can seven people sit at a source table?
b. What is the probability that husband and wife sit together?

Solution

a. $(7-1)!=6!=720ways$
b. $Pro\ that\ husband\ and\ wife\ sit\ together\ =\frac{(7-2)!}{6!}=\frac{1}{6}$ and since the husband and wife can be arranged

$$\frac{5!\times 2!}{6!}=\frac{1}{6}\times 2!=\frac{2}{6}=\frac{1}{3}$$

Theorem 4

The number of permutations or ordered arrangements, of r objects taken from n unlike objects is written as $P_r^n=\frac{n!}{(n-r)!}$

Question

7 cards are labeled A,B,C,D,E,F,G are thoroughly shuffled and dealt out face upwards on a table. Find the probability as a fraction in its simplest form that the first three cards to appear are the cards labeled ABC in that order.

Solution

$$r=3\ \ n=7\ \ P_3^7=\frac{7!}{(7-3)!}=210ways$$

Probability that the first 3 cards is ABC $=\frac{1}{210}$

Combination

The number of combination of R objects from n unlike objects is $\frac{n!}{r!(n-r)!}$

Keyword is selected or chosen

Note

Arrangement is not important

Example

From a group of 10boys and 8girls, 2 students are chosen at random. If the students are girls, in how many ways can they be chosen?

Solution

$$n = 8 \;\; r = 2 \quad \frac{8!}{2!\,6!} = 28$$

$$\mathit{Probability\ that\ they\ are\ both\ girls} = \frac{28}{153}$$

Try

From a group of 6men and 8women, 5 people are chosen at random. Find the probability that there are more men chosen than women.

Probability distribution

Discrete variables

In this section we will learn:

- About probability distribution for discrete random variables
- How to calculate and use E(X), the expectation (mean)
- How to calculate and use E(g(X)), the expectation of a simple function of X.
- How to calculate and use Var(X), the variance of X.
- About the cumulative distribution function F(X).
- About the results relating to expectation algebra for random variables X and Y

For a discrete random variable, the sum of the probabilities is 1

i.e. $\sum_{all\ x} P(X = x) = 1 \ \ or \quad \sum p_i = 1 \ \ for\ i = 1,2, \dots, n$

Expectation of X

The expectation of X (expected value or mean),written E(X), is given by

$$E(X) = \sum_{all\ x} xP(X = x)$$

this can be expressed in the form

$E(X) = \sum x_i p_i \quad i = 1,2, \dots, n$
$\mu = E(X)$

VARIANCE OF X

For a discrete random variable X, with $\mu = E(X)$, the variance is defined as follows:

The variance of X is written as Var(X) which is given by

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

or

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

$= E(X^2 - 2\mu X + \mu^2)$

$= E(X^2) - 2\mu^2 + \mu^2$

$= E(X^2) - \mu^2$

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

You could write $Var(X) = E(X^2) - E^2(X)$

$\sigma = \sqrt{Var(X)} = standard\ deviation\ X$

Question

X is the random variable 'the number on a biased die', and the p.d.f of X is as shown.

x	1	2	3	4	5	6
$P(X=x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{5}$	y	$\frac{1}{5}$	$\frac{1}{6}$

Find (a) the value of y, (b) E(X) (c) $E(X^2)$ (d) Var(X)

e. P(X = 1) f. P(X > 2) g. P(X $\geq$ 3) h. $P(1 < X < 5)$

i. P $(1 \leq X \leq 5)$ j. $P(1 \leq X < 5)$ k. P$(1 < X \leq 5)$

Solution

a. $\sum_{all\ x} P(X = x) = 1 \ \ or \quad \sum p_i = 1 \ \ for\ i = 1{,}23{,}4{,}5{,}6$

$$\frac{1}{6} + \frac{1}{6} + \frac{1}{5} + y + \frac{1}{5} + \frac{1}{6} = 1$$

$$y = \frac{1}{10}$$

x	1	2	3	4	5	6
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{6}$

b. $E(X) = \sum x_i p_i$

$$1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{5}\right) + 4\left(\frac{1}{10}\right) + 5\left(\frac{1}{5}\right) + 6\left(\frac{1}{6}\right) = 3.5$$

c. $E(X^2) = 1^2\left(\frac{1}{6}\right) + 2^2\left(\frac{1}{6}\right) + 3^2\left(\frac{1}{5}\right) + 4^2\left(\frac{1}{10}\right) + 5^2\left(\frac{1}{5}\right) + 6^2\left(\frac{1}{6}\right) = 15\frac{7}{30}$

d. Var(X)= E(X^2)- E^2(X)

$$= 15\frac{7}{30} - 3.5^2$$

e. P(X = 1) = $\frac{1}{6}$

f. P(X > 2) = P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6)

$$= \frac{1}{5} + \frac{1}{10} + \frac{1}{5} + \frac{1}{6} = \frac{2}{3}$$

g. $P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6)$

$$= \frac{1}{5} + \frac{1}{10} + \frac{1}{5} + \frac{1}{6}$$

h. $P(1 < X < 5) = P(X = 2) + P(X = 3) + P(X = 4)$

$$= \frac{1}{6} + \frac{1}{5} + \frac{1}{10} = \frac{7}{15}$$

i. $P(1 \leq X < 5) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$

$$= \frac{1}{6} + \frac{1}{6} + \frac{1}{5} + \frac{1}{10} = \frac{19}{30}$$

j. $P(1 \leq X \leq 5) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$

$$= \frac{1}{6} + \frac{1}{6} + \frac{1}{5} + \frac{1}{10} + \frac{1}{5} = \frac{5}{6}$$

k. $P(1 < X \leq 5) = P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$

$$= \frac{1}{6} + \frac{1}{5} + \frac{1}{10} + \frac{1}{5} = \frac{2}{3}$$

Properties of independent random variables

1. $E(aX \pm bY) = aE(X) \pm bE(Y)$
2. $E(aX \pm b) = aE(X) \pm b$
3. $Var(aX \pm bY) = a^2Var(X) + b^2Var(Y)$
4. $Var(a) = 0$
5. $Var(aX) = aVar(X)$
6. $Var(aX + b) = a^2Var(X)$

Question

Independent random variables X and Y are such that E(X)= 4 , E(Y)= 5 , Var(X)= 1, Var(Y)= 2.

Find

a. E(4X + 2Y)
b. Var(3X + 2Y)

Solution

a. $E(4X\ +\ 2Y) = 4E(X) + 2E(Y) = 26$
b. $Var(3X\ +\ 2Y) = 3^2Var(X) + 2^2Var(Y) = 17$

Question

Independent random variables X and Y are such that E(X^2)= 14, E(Y^2)= 20, Var(X) = 10, Var(Y) = 11

Find the value of

a. E(3X - 2Y)

b. Var(5X – 2Y)

Solution

a. $E(3X\ -\ 2Y) = 3E(X) - 2E(Y)$
   But $E(X) = \sqrt{E(X^2) - Var(X)}$ and $E(Y) = \sqrt{E(Y^2) - Var(Y)}$
   $E(X) = \sqrt{14 - 10} = \pm 2 \quad and \quad E(Y) = \sqrt{20 - 11} = \pm 3$
   $\therefore E(3X\ -\ 2Y) = 3E(X) - 2E(Y) = 3(2) - 2(3) = 0$
   $or\ E(3X\ -\ 2Y) = 3E(X) - 2E(Y) = 3(-2) - 2(3) = -12$
   $or\ E(3X\ -\ 2Y) = 3E(X) - 2E(Y) = 3(2) - 2(-3) = 12$

b. $Var\big(5X - 2Y\big) = 5^2Var(X) + (-2)^2Var(Y) = 294$

Question

Independent random variables X and Y are such that E(X)=3 E(X^2)= 12, E(Y) = 4 E(Y^2)= 18

Find the value of

a. E(3X - 2Y)
b. E(2Y – 3X)
c. E(6X + 4Y)
d. Var(2X – Y)
e. Var(2X + 2Y)
f. Var(3Y + 2X)

Solution

a. $E(3X - 2Y) = 3E(X) - 2E(Y) = 1$
b. $E(2Y - 3X) = 2E(Y) - 3E(X) = -1$
c. $E(6X + 4Y) = 6E(X) + 4E(Y) = 34$
d. $Var(2X - Y) = 2^2 Var(X) + (-1)^2 Var(Y)$

$$Var(X) = E(X^2) - E^2(X)$$

$$= 12 - 3^2 = 3$$

$$Var(Y) = E(Y^2) - E^2(Y)$$

$$= 2$$

a. $Var(2X - Y) = 2^2 Var(X) + (-1)^2 Var(Y) = 14$

SOLVED QUESTIONS

Question 1

Suppose your vehicle is licensed in a state that issues license plates that consist of three digits (between 0 and 9) followed by three letters (between A and Z). If a license number is selected randomly, what is the probability that yours is the one selected?

Solution

3 digits between 0 and 9, so the probability of any three numbers is 1/(10*10*10);
3 letters A to Z, so the probability of any three numbers is 1/(26*26*26); The probability that your license plate
is chosen is then (1/103)*(1/263) = 5.7 x 10-8

Question 2

If P(A)= 0.3, P(B)=0.2, and $P(A \cap B) = 0.1$, determine the following probabilities:

a. P(A') b. $P(A \cup B)$ c. P($A' \cap B$) d. $P(A \cap B')$
e. $P[(A \cup B)']$ f. $P(A' \cup B)$

Solution

a) P (A') = 1- P (A) = 0.7
b) P (A ∪ B) = P (A) + P (B) - P (A ∩B) = 0.3+0.2 - 0.1 = 0.4
c) P (A′ ∩B) + P (A ∩ B) = P (B). Therefore, P (A′∩B) = 0.2 - 0.1 = 0.1

d) P (A) = P(A ∩B) + P(A ∩B′). Therefore, P(A ∩B′) = 0.3 - 0.1 = 0.2

e) P((A ∪B)') = 1 - P(A ∪B) = 1 - 0.4 = 0.6
f) P(A′ ∗B) = P(A') + P(B) - P(A′ ∩B) = 0.7 + 0.2 - 0.1 = 0.8

Question 3

A maintenance firm has gathered the following information regarding the failure mechanisms for air conditioning systems:

		evidence of gas leaks	
		yes	no
evidence of electrical failure	yes	55	17
	no	32	3

The units without evidence of gas leaks or electrical failure showed other types of failure. If this is a representative sample of AC failure, find the probability
(a) That failure involves a gas leak
(b) That there is evidence of electrical failure given that there was a gas leak
(c) That there is evidence of a gas leak given that there is evidence of electrical failure

Solution

a) P(gas leak) = (55 + 32)/107 = 0.813
b) P(electric failure|gas leak) = (55/107)/(87/102) = 0.632
c) P(gas leak| electric failure) = (55/107)/(72/107) = 0.764

Question 4

Suppose 2% of cotton fabric rolls and 3% of nylon fabric rolls contain flaws. Of the rolls used by a manufacturer, 70% are cotton and 30% are nylon. What is the probability? That a randomly selected roll used by the manufacturer contains flaws?

Solution

Let F denote the event that a roll contains a flaw.
Let C denote the event that a roll is cotton.
$P(F) = P(F|C)P(C) + P(F|C')P(C')$

$$= (0.02)(0.70) + (0.03)(0.30) = 0.023)$$

Question 5

A batch of 25 injection-molded parts contains 5 that have suffered excessive shrinkage.

(a) If two parts are selected at random, and without replacement, what is the probability that the second part selected is one with excessive shrinkage?

(b) If three parts are selected at random, and without replacement, what is the probability that the third part selected is one with excessive shrinkage?

Solution

Let A denote a event that the first part selected has excessive shrinkage.
Let B denote the event that the second part selected has excessive shrinkage.
a) P(B)= P(B|A)P(A) + P(B|A ')P(A')
= (4/24)(5/25) + (5/24)(20/25) = 0.20
b) Let C denote the event that the second part selected has excessive shrinkage.

$$P(C) = P(C|A \cap B)P(A \cap B) + P(C|A \cap B')P(A \cap B') + P(C|A' \cap B)P(A' \cap B) + P(C|A' \cap B')P(A' \cap B')$$

$$= \frac{3}{23}\left(\frac{2}{24}\right)\left(\frac{5}{25}\right) + \frac{4}{23}\left(\frac{20}{24}\right)\left(\frac{5}{25}\right) + \frac{4}{23}\left(\frac{5}{24}\right)\left(\frac{20}{25}\right) + \frac{5}{23}\left(\frac{19}{24}\right)\left(\frac{20}{25}\right) = 0.20$$

Question 6

The probability that a lab specimen contains high levels of contamination is 0.10. Five samples are checked, and the samples are independent.

(a) What is the probability that none contains high levels of contamination?

(b) What is the probability that exactly one contains high levels of contamination?

(c) What is the probability that at least one contains high levels of contamination?

Solution

It is useful to work one of these exercises with care to illustrate the laws of probability. Let H_i denote the event that the ith sample contains high levels of contamination.

a) $P(H_1' \cap H_2' \cap H_3' \cap H_4' \cap H_5') = P(H_1')P(H_2')P(H_3')P(H_4')P(H_5')$

by independence. Also, $P(H_i') = 0.9$. Therefore, the answer is $0.9^5 = 0.59$

b) $A_1 = (H_1 \cap H_2' \cap H_3' \cap H_4' \cap H_5')$

$A_2 = (H_1' \cap H_2 \cap H_3' \cap H_4' \cap H_5')$

$A_3 = (H_1' \cap H_2' \cap H_3 \cap H_4' \cap H_5')$

$A_4 = (H_1' \cap H_2' \cap H_3' \cap H_4 \cap H_5')$

$A_5 = (H_1' \cap H_2' \cap H_3' \cap H_4' \cap H_5)$

The requested probability is the probability of the union $A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5$ and these events are mutually exclusive. Also, by independence $P(A_i) = 0.9^4(0.1) = 0.0656$. Therefore, the answer is $5(0.0656) = 0.328$.

c) Let B denote the event that no sample contains high levels of contamination. The requested probability is $P(B') = 1 - P(B)$. From part (a), $P(B') = 1 - 0.59 = 0.41$.

Question 7

Eight cavities in an injection-molding tool produce plastic connectors that fall into a common stream. A sample is chosen every several minutes. Assume that the samples are independent.

(a) What is the probability that five successive samples were all produced in cavity one of the mold?

(b) What is the probability that five successive samples were all produced in the same cavity of the mold?

(c) What is the probability that four out of five successive samples were produced in cavity one of the mold?

Solution

Let A denote the event that a sample is produced in cavity one of the mold.

a) By independence, $P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) = (\frac{1}{8})^5 = 0.00003$

b) Let B_i be the event that all five samples are produced in cavity i. Because the B's are mutually exclusive, $P(B_1 \cup B_2 \cup ... \cup B_8) = P(B_1) + P(B_2) + ... + P(B_8)$

From part a., $P(B_i) = (\frac{1}{8})^5$. Therefore, the answer is $8(\frac{1}{8})^5 = 0.00024$

c) By independence, $P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5') = (\frac{1}{8})^4 (\frac{7}{8})$. The number of sequences in which four out of five samples are from cavity one is 5. Therefore, the answer is $5(\frac{1}{8})^4 (\frac{7}{8}) = 0.00107$.

Question 8

Customers are used to evaluate preliminary product designs. In the past, 95% of highly successful products received good reviews, 60% of moderately successful products received good reviews, and 10% of poor products received good reviews. In addition, 40% of products have been highly successful, 35% have been moderately successful, and 25% have been poor products.

(a) What is the probability that a product attains a good review?

(b) If a new design attains a good review, what is the probability that it will be a highly successful product?

(c) If a product does not attain a good review, what is the probability that it will be a highly successful product?

Solution

Let G denote a product that received a good review. Let H, M, and P denote products that were high, moderate, and poor performers, respectively.

a)

$$P(G) = P(G|H)P(H) + P(G|M)P(M) + P(G|P)P(P)$$
$$= 0.95(0.40) + 0.60(0.35) + 0.10(0.25)$$
$$= 0.615$$

b) Using the result from part a.,

$$P(H|G) = \frac{P(G|H)P(H)}{P(G)} = \frac{0.95(0.40)}{0.615} = 0.618$$

c) $$P(H|G') = \frac{P(G'|H)P(H)}{P(G')} = \frac{0.05(0.40)}{1-0.615} = 0.052$$

Question 9

A steel plate contains 20 bolts. Assume that 5 bolt are not torqued to the proper limit. Four bolts are selected at random, without replacement, to be checked for torque.

(a) What is the probability that all four of the selected bolts are torqued to the proper limit?

(b) What is the probability that at least one of the selected bolts is not torqued to the proper limit?

Solution

Let A_i denote the event that the ith bolt selected is not torqued to the proper limit.

a) Then,

$$P(A_1 \cap A_2 \cap A_3 \cap A_4) = P(A_4 | A_1 \cap A_2 \cap A_3)P(A_1 \cap A_2 \cap A_3)$$
$$= P(A_4 | A_1 \cap A_2 \cap A_3)P(A_3 | A_1 \cap A_2)P(A_2 | A_1)P(A_1)$$
$$= \left(\frac{2}{17}\right)\left(\frac{3}{18}\right)\left(\frac{4}{19}\right)\left(\frac{5}{20}\right) = 0.282$$

b) Let B denote the event that at least one of the selected bolts are not properly torqued. Thus, B' is the event that all bolts are properly torqued. Then,

$$P(B) = 1 - P(B') = 1 - \left(\frac{15}{20}\right)\left(\frac{14}{19}\right)\left(\frac{13}{18}\right)\left(\frac{12}{17}\right) = 0.718$$

Question 10

A lot of 50 spacing washers contains 30 washers that are thicker than the target dimension. Suppose that three washers are selected at random, without replacement, from the lot.

(a) What is the probability that all three washers are thicker than the target?

(b) What is the probability that the third washer selected is thicker than the target if the first two washers selected are thinner than the target?

(c) What is the probability that the third washer selected is thicker than the target?

Solution

Let A_i denote the event that the ith washer selected is thicker than target.

a) $\left(\frac{30}{50}\right)\left(\frac{29}{49}\right)\left(\frac{28}{8}\right) = 0.207$

b) $30/48 = 0.625$

c) The requested probability can be written in terms of whether or not the first and second washer selected are thicker than the target. That is,

$$P(A_3) = P(A_1A_2A_3 \text{ or } A_1A_2'A_3 \text{ or } A_1'A_2A_3 \text{ or } A_1'A_2'A_3)$$

$$= P(A_3|A_1A_2)P(A_1A_2) + P(A_3|A_1A_2')P(A_1A_2')$$

$$+P(A_3|A_1'A_2)P(A_1'A_2) + P(A_3|A_1'A_2')P(A_1'A_2')$$

$$= P(A_3|A_1A_2)P(A_2|A_1)P(A_1) + P(A_3|A_1A_2')P(A_2'|A_1)P(A_1)$$

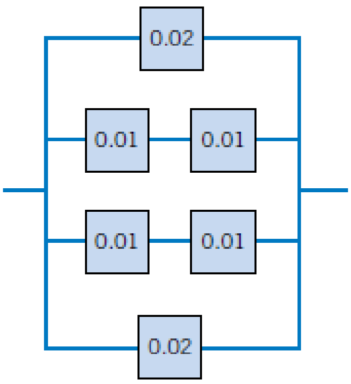
$$+P(A_3|A_1'A_2)P(A_2|A_1')P(A_1') + P(A_3|A_1'A_2')P(A_2'|A_1')P(A_1')$$

$$= \frac{28}{48}\left(\frac{30}{50}\frac{29}{49}\right) + \frac{29}{48}\left(\frac{20}{50}\frac{30}{49}\right) + \frac{29}{48}\left(\frac{20}{50}\frac{30}{49}\right) + \frac{30}{48}\left(\frac{20}{50}\frac{19}{49}\right)$$

$$= 0.60$$

Question 11

The following circuit operates if and only if there is a path of functional devices from left to right. Assume that devices fail independently and that the probability of *failure* of



Solution

Let A_i denote the event that the ith row operates. Then,

$P(A_1) = 0.98, P(A_2) = (0.99)(0.99) = 0.9801, P(A_3) = 0.9801, \ P(A_4) = 0.98.$

The probability the circuit does not operate is

$$P(A_1')P(A_2')P(A_3')P(A_4') = (0.02)(0.0199)(0.0199)(0.02) = 1.58 \times 10^{-7}$$

Discrete probability distribution

Bernoulli process

$$P(X) = p^x(1-p)^{1-x}, x = 0 \; or \; 1 \; and \; 0 < p < 1; q = 1 - p$$

Mean $\mu = E(x) = p \; and \; \sigma^2 = var(x) = p(1-p) = pq$

The random variable of this experiment is a binary variable which assumes the value;0 and 1.

If $x = 0 \Longrightarrow probabiliry \; of \; failure$

$$P(x = 0) = p^0(1-p)^{1-0} = q$$

If x=1 $\Longrightarrow probability \; of \; success$

$$P(x = 1) = p^1(1-p)^{1-1} = p$$

Binomial distribution

Binomial experiment is the generalization of the Bernoulli trials.

Conditions for a binomial distribution

1. A finite number, **n** trials are carried out.
2. The trials are independent and identical.
   An independent event is an event in which the occurrence of one does not affect the other.
   e.g.; tossing a coin.
3. The outcome of each trial is either a success or a failure.

4. The probability p of a successful outcome is the same for each trial.

$$X \sim B(n,p) \text{ or } X \sim Bin(n,p)$$

Note:

The number of trials, n and the probability of success, p are both needed to describe the distribution completely. They are known as the parameters of the binomial distribution.

If$X \sim B(n,p)$, the probability of obtaining r success in n trials is $P(X = r)$ where $P(X = r) = \binom{n}{r} p^r q^{n-r}\ for\ r = 0,1,2,3, \dots, n$

$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!} \quad ; n! = n(n-1)(n-2)(n-3)\dots 3 \cdot 2 \cdot 1$$

n = number of trials.

P = probability of success on a single trial

r = number of success in n trials.

Mean; $\mu = np$ Variance; $\sigma^2 = npq$ standard deviation; $\sigma = \sqrt{npq}$

Note

$P(X = 0) = \binom{n}{0} q^n p^0\ but\ p^0 = 1 \quad \binom{n}{0} = \frac{n!}{0!\,n!} = 1$

So $P(X = 0) = q^n = (1-p)^n$

$$P(X=n)=\binom{n}{n}q^{n-n}p^{n}\ but\ q^{0}=1 \quad \binom{n}{n}=\frac{n!}{n!\,0!}=1$$

So $P(X=n)=p^{n}$

Software

dbinom(r,n,p); pbinom(r,n,p)

Question 1

1. The random variable X is distributed $B(7,0.2)$. Find, correct to 3 decimal places
   a. P(X=3)
   b. $P(1<X\leq 4)$
   c. P(X>1)

Solution

P=0.2 q=0.8 n=7

a. $P(X=3)=\binom{7}{3}q^{4}\,p^{3}$
   $=0.115$

Software

```
> dbinom(3,7,0.2)
[1] 0.114688
```

b. $P(1 < X \leq 4) = P(X = 2) + P(X = 3) + P(X = 4)$

$$= \binom{7}{2} q^5 p^2 + \binom{7}{3} q^4 p^3 + \binom{7}{4} q^3 p^4$$

$$= 0.419$$

```
> dbinom(2,7,0.2)+dbinom(3,7,0.2)+dbinom(4,7,0.2)

[1] 0.4186112

> sum(dbinom(2:4,7,0.2))

[1] 0.4186112

>pbinom(7,7,0.2)-
(dbinom(0,7,0.2)+dbinom(1,7,0.2)+dbinom(5,7,0.2)+dbinom(6,7,0.2)
+dbinom(7,7,0.2))

[1] 0.4186112
```

c. $P(X > 1) = 1 - P(X \leq 1)$

$$= P(X = 2) + P(X = 3) + \cdots + P(X = 7)$$

$$= 1 - [P(X = 0) + P(X = 1)]$$

$$= 1 - (q^7 + \binom{7}{1} q^6 p^1)$$

$$= 0.423$$

Software

```
> 1-pbinom(1,7,0.2)

[1] 0.4232832

> sum(dbinom(2:7,7,0.2))

[1] 0.4232832
```

```
>
dbinom(2,7,0.2)+dbinom(3,7,0.2)+dbinom(4,7,0.2)+dbinom(5,7,0.2)+
dbinom(6,7,0.2)+dbinom(7,7,0.2)

[1] 0.4232832
```

Question 2

A box contains a large number of pens. The probability that a pen is faulty is 0.1. How many pens would you need to select to be more than 95% certain of picking at least one faulty one?

Solution

P=0.1   q=0.9   n=?

$P(X \geq 1) > 0.95$

$P(X \geq 1) = 1 - P(X < 1)$

$= 1 - P(X = 0)$

$= 1 - q^n$

$= 1 - 0.9^n$

$1 - 0.9^n > 0.95$

$0.05 > 0.9^n \;\; or \;\; 0.9^n < 0.05$

By trials take n=25, 26, 27,… (Long process)

Take log of both sides.

nlog0.9 < log0.05

n > 28.4

n ≈ 29

You need to select at least 29 pens.

Question 3

Suppose that a consignment of 300 electrical fuses contains 5% defectives. If a random sample of ten fuses is selected and tested, find the probability of observing at least three defectives.

Solution

X~B(10,0.05)

$P(X \geq 3) = 1 - \geq P(X < 3)\ or\ 1 - P(X \leq 2)$

$$= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$$

$$= 1 - \left[0.95^{10} + \binom{10}{1} \cdot 0.95^{9} \cdot 0.05 + \binom{10}{2} \cdot 0.95^{8} \cdot 0.05^{2}\right]$$

$$= 0.0115$$

Software

```
> 1-pbinom(2,10,0.05)

[1] 0.01150356

> sum(dbinom(3:10,10,0.05))

[1] 0.01150356

>
dbinom(3,10,0.05)+dbinom(4,10,0.05)+dbinom(5,10,0.05)+dbinom(6,10,0.05)+dbinom(7,10,0.05)+dbinom(8,10,0.05)+dbinom(9,10,0.05)+dbinom(10,10,0.05)
```

```
[1] 0.01150356
```

Question 4

30% of pupils in a school travel to school by bus. From a sample of 10 pupils chosen at random, find the probability that;

a. Only three travel by bus.
b. Less than half travel by bus.

Solution

P=0.3   q=0.7   n=10 $X \sim B(10{,}0.3)$

a. $P(X=3) = \binom{10}{3} q^7 p^3$

$= 0.267$

Software

```
> dbinom(3,10,0.3)
[1] 0.2668279
```

b. $P(X<5) = P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4)$

$$q^{10} + \binom{10}{1} q^9 p^1 + \binom{10}{2} q^8 p^2 + \binom{10}{3} q^7 p^3 + \binom{10}{4} q^6 p^4$$

$= 0.850$

Software

```
> pbinom(4,10,0.3)
[1] 0.8497317
```

```
> sum(dbinom(0:4,10,0.3))

[1] 0.8497317

> 1-sum(dbinom(5:10,10,0.3))

[1] 0.8497317

> dbinom(0,10,0.3)+dbinom(1,10,0.3)+dbinom(2,10,0.3)+dbinom(3,10,0.3)+dbinom(4,10,0.3)

[1] 0.8497317


> 1-(dbinom(5,10,0.3)+dbinom(6,10,0.3)+dbinom(7,10,0.3)+dbinom(8,10,0.3)+dbinom(9,10,0.3)+dbinom(10,10,0.3))

[1] 0.8497317
```

Question 5

In a survey on washing powder, it is found that the probability that a shopper chooses soapysuds is 0.25. Find the probability that in a random sample on 9 shoppers.

a. Exactly 3 choose soapysuds.
b. More than 7 choose soapysuds.

Solution

P=0.25   q=0.75 n=9

a. $P(X = 3) = \binom{9}{3} q^6 p^3$

$= 0.234$

Software

```
> dbinom(3,9,0.25)

[1] 0.2335968
```

b. $P(X > 7) = P(X = 8) + P(X = 9)$

$$= \binom{9}{8} q^1 p^8 + p^9$$

$$= 0.000107$$

Software

```
> dbinom(8,9,0.25)+dbinom(9,9,0.25)

[1] 0.0001068115

> sum(dbinom(8,9,0.25))

[1] 0.0001029968

> 1-sum(dbinom(0:7,9,0.25))

[1] 0.0001068115
```

Question 6

A bag contains counters of which 40% are red and the rest yellow. A counter is taken from the bag, its color noted and then replaced. This is performed eight times in all. Calculate the probability that;

a. Exactly three will be red.
b. At least one will be red.
c. More than four will be yellow.

Solution

n=8 p=0.4 q=0.6

a. $P(X = 3) = \binom{8}{3} q^5 p^3$
$= 0.279$

Software

```
> dbinom(3,8,0.4)

[1] 0.2786918
```

b. $P(X \geq 1) = 1 - P(X < 1)$
$= 1 - P(X = 0)$
$= 1 - q^8$
$= 0.983$

Software

```
> 1-dbinom(0,8,0.4)

[1] 0.9832038

> sum(dbinom(1:8,8,0.4))

[1] 0.9832038

> pbinom(8,8,0.4)-pbinom(0,8,0.4)
```

```
[1] 0.9832038

>
dbinom(2,8,0.4)+dbinom(1,8,0.4)+(dbinom(3,8,0.4)+dbinom(4,8,0.4)
+dbinom(5,8,0.4)+dbinom(6,8,0.4)+dbinom(7,8,0.4)+dbinom(8,8,0.4)
)

[1] 0.9832038
```

c. n=8   p=0.6 q=0.4

$P(X > 4) = 1 - P(X \leq 4)$

$= 1 - P(X < 5)$

$= P(X = 5) + P(X = 6) + P(X = 7) + P(X = 8)$

$= \binom{8}{5} q^3 p^5 + \binom{8}{6} q^2 p^6 + \binom{8}{7} q^1 p^7 + p^8$

$= 0.594$

Software

```
> sum(dbinom(5:8,8,0.6))

[1] 0.5940864

> pbinom(8,8,0.6)-pbinom(4,8,0.6)

[1] 0.5940864

> 1-sum(dbinom(0:4,8,0.6))

[1] 0.5940864

>
dbinom(5,8,0.6)+dbinom(6,8,0.6)+dbinom(7,8,0.6)+dbinom(8,8,0.6)
```

```
[1] 0.5940864
```

Question 7

$X\ is\ B\ (4, p)\ and\ P(X = 4) = 0.0256$

Find P(X=2)

Solution

$$P(X = 4) = \binom{n}{4} q^{n-4} p^4$$

$$0.0256 = \binom{4}{4} q^0 p^4 = p^4$$

$$0.0256 = p^4$$

$$p = 0.4$$

$$\Rightarrow P(X = 2) = \binom{4}{2} 0.6^2 0.4^2$$

software

```
> dbinom(2,4,0.4)

[1] 0.3456
```

Question 8

A multiple choice test contains 25 questions, each with four answers. Assume a student just guesses on each question.

a. What is the probability that the student answers more than 20 questions correctly?
b. What is the probability that the student answers less than 5 questions correctly?

Solution

n=25   p=0.25   q=0.75

a. $P(X > 20) = P(X = 21) + P(X = 22) + P(X = 23) + P(X = 24) + P(X = 25)$

$$= \binom{25}{21} q^4 p^{21} + \binom{25}{22} q^3 9^{22} + \binom{25}{23} q^2 p^{24} + \binom{25}{24} q^1 p^{24} + p^{25} = 0$$

Software

```
> sum(dbinom(21:25,25,0.25))

[1] 9.676935e-10

> 1-pbinom(20,25,0.25)

[1] 9.676935e-10
```

Question 9

Suppose having an experiment of two outcomes. The probability of success is $\frac{1}{3}$ and probability of failure is $\frac{2}{3}$. What is the probability of having one success obtained in two trials?

Solution

Outcome	Probability
SS	1/9
SF	2/9
FS	2/9
FF	4/9

$P[exactly\ 1\ success] = P[sf, fs] = {}^{2}/_{9} + {}^{2}/_{9} = {}^{4}/_{9}$

But using binomial theorem ${}^{n}C_r p^r q^{n-r}$ where $p =$ success, $q =$ failure and $p + q = 1$

From the example $n = 2, p = \frac{1}{3}, q = \frac{2}{3}$

$$ {}^{2}C_1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^1 = 2 \times \frac{1}{3} \times \frac{2}{3} = \frac{4}{9} $$

Software

```
> dbinom(1,2,1/3)

[1] 0.4444444
```

Question 10

In a game of chance, you play by rolling a fair die four times and you count the number of results which are 6s.

i. What is the probability that in one play of the game you obtain exactly three 6s?

solution

$P[X = 3] = \ {}^{4}C_3 \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^1 = \frac{20}{1296}$

Software

```
> dbinom(3,4,1/6)

[1] 0.0154321
```

ii. What is the probability that in one of the game you obtain exactly two 5s

solution

$P[X=2] = \ {}^{4}C_{2}\left(\frac{1}{6}\right)^{2}\left(\frac{5}{6}\right)^{2} = \frac{25}{216}$

Software

```
> dbinom(2,4,1/6)

[1] 0.1157407
```

Question 11

Of the telephone calls received by an airline reservation agent,60% requests for information and 40%are to make reservations. Assume the calls can be viewed as Bernoulli trials with success defines to be call for a reservation. Six calls were reserved.

1. What is the probability that exactly 2 calls are for reservation?

2. What is the probability that at least 4 are for information?

Solution

n=6   p=0.4 q=0.6

1. $P(X=2) = \ {}^{6}C_{2}(0.4)^{2}(0.6)^{4} = \frac{972}{3125}$

Software

```
> dbinom(2,6,0.4)

[1] 0.31104
```

n=6   p=0.6   q=0.4

2. $P(X \geq 4) = P(X = 4) + P(X = 5) + P(X = 6)$

$$= \; {}^{6}C_{4}(0.6)^{4}(0.4)^{2} + \; {}^{6}C_{5}(0.6)^{5}(0.4)^{1} + \; {}^{6}C_{6}(0.6)^{6}(0.4)^{0} = 0.1792$$

Software

```
> sum(dbinom(4:6,6,0.6))

[1] 0.54432
```

Question 12

The probability that it will be a fine day is 0.4. Find the expected number of fine days in week and also the standard deviation.

Solution

The expected number of five days $= E(X) \quad = np \; = 7 \times 0.4 = 2.8$

Standard deviation of $= \sqrt{Var(X)}$ $=\sqrt{7 \times 0.4 \times 0.6}$ = 1.3 days

Question 13

A used car sales woman estimates that each times she shows a customer a car, there is a probability 0.1 that the customer will buy the car. The sales woman would like to sell at least one car per week. If showing a car is a Bernoulli trial how many car would the saleswoman show per week so that the probability is 0.95 of at least one sale?

Ans 29

Question 14

The random variable x is $B(n, 0.3)$ and $E(x)$ =2.4 find n and standard deviation of x?

Ans n= 8, s.d=1.3

Question 15

In a group of people the expected number who wear glasses is 2 and the variance is 1.6 find the probability that

a. A person chosen at random from the group wear glasses.
b. 6 people in the group wear glasses.

Solution

a. E(X)= np = 2   Var(X) = 1.6

   np = 2 … (1)

   npq = 1.6 … (2)

   from (1) and (2)

   2q = 1.6

   q = 0.8 $\Longrightarrow$ $\therefore$ p = 0.2

b. n = 10

   $P(X = 6) = {}^{10}C_6(0.2)^6(0.8)^4 = 0.00551$

   Software

```
> dbinom(6,10,0.2)
[1] 0.005505024
```

Poisson distribution

Conditions for a Poisson model

1. Events occur singly and at random in a given interval of time or space.
2. The parameter λ; $\lambda > 0$ is the mean number of occurrences in the given interval, is known and is finite(i.e the occurrence rate per unit).

The variable X is the number of occurrences in the given interval.

$X \sim Po(\lambda)$ $\quad \mu = \lambda \quad$ $\sigma^2 = \lambda$

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

Typical examples of random variables for which the Poisson probability distribution provides a good model are.

1. The number of traffic accidents per month at a busy intersection.
2. The number of death claims received per day by an insurance company.
3. The number of unscheduled admissions per day to a hospital.

Poisson distribution is used to model the occurrence of a random event that happens in some time periods.

Question 1

A student finds that the average number of amoebas in 10ml of pond water from a particular pond is 4. Assuming that the number of amoebas follow a Poisson distribution, find the probability that in a 10ml sample

a. There are exactly 5 amoebas
b. There are no amoebas.
c. There are fewer than three amoebas.

Solution

X is the number of amoebas in 10ml if pond water, where $X \sim P_o(4)$
$\lambda = 4$

a. $P(X = 5) = \frac{e^{-4}4^5}{5!}$
$= 0.156$

Software

```
> dpois(5,4)

[1] 0.1562935
```

b. $P(X = 0) = \frac{e^{-4}4^0}{0!}$
$= 0.0183$

```
> dpois(0,4)

[1] 0.01831564
```

c. $P(X < 3) = 1 - P(X \geq 3)$
$= P(X = 0) + P(X = 1) + P(X = 2)$

$$= \frac{e^{-4}4^0}{0!} + \frac{e^{-4}4^1}{1!} + \frac{e^{-4}4^2}{2!}$$
$$= 0.238$$

Software

```
> sum(dpois(0:2,4))

[1] 0.2381033

> ppois(2,4)

[1] 0.2381033
```

Note

Unit interval

For this example, the mean number of amoebas in 10ml of pond water from a particular period is four so the number in 10ml is distributed $P_0(4)$. now suppose you want to find a probability relating to the number of amoebas in 5ml of water from the same pond. The mean number of amoebas in 5ml is two, so the number in 5ml is distributed $P_0(2)$.

Question 2

On average the school photocopier breaks down eight times during the school week(Mon-Fri). Assuming that the number of breakdowns can be modeled by a Poisson distribution. Find the probability that it breaks down.

a. Five times in a given week.

b. Once on Monday.
c. Eight times in a fortnight.

Solution

a. X is the number of breakdowns in a week, where X~$P_0(8)$

$$P(X=5)=\frac{e^{-8}8^5}{5!}=0.0916$$

Software

```
> dpois(5,8)

[1] 0.09160366
```

b. Let B be the number of breakdowns in a day. The mean number of breakdowns in a day is $\frac{8}{5}=1.6$ so $B{\sim}P_o(1.6)$.

$$P(Y=1)=\frac{e^{-1.6}1.6^1}{1!}=1.6e^{-1.6}=0.323$$

Software

```
> dpois(1,1.6)

[1] 0.3230344
```

c. Let Y be the number of breakdowns in a fortnight.

The mean number of breakdowns in a fortnight is 16 so $Y \sim P_o(16)$

$$P(Y = 8) = \frac{e^{-16}16^8}{8!} = 0.0120$$

Software

```
> dpois(8,16)
[1] 0.01198747
```

Question 3

X follows a Poisson distribution with standard deviation 1.5. Find $P(X \geq 3)$.

Solution

If $X \sim P_o(\lambda)$ $var(X) = \lambda$

$Var(X) = (standard\ deviation)^2 = 2.25$

$So\ \lambda = 2.25\ and\ X \sim P_o(2.25)$

$$P(X \geq 3) = 1 - P(X < 3)$$

$$= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$$

$$= 1 - \left[\frac{e^{-2.25}2.25^0}{0!} + \frac{e^{-2.25}2.25^1}{1!} + \frac{e^{-2.25}2.25^2}{2!}\right]$$

$$= 0.391$$

Software

```
> 1-ppois(2,2.25)
```

```
[1] 0.3906607

> 1-sum(dpois(0:2,2.25))

[1] 0.3906607

> 1-(dpois(0,2.25)+dpois(1,2.25)+dpois(2,2.25))

[1] 0.3906607
```

Question 4

An insurance company receives on an average two claims per week from a particular factory. Assuming that the number of claims can be modeled by a Poisson distribution, find the probability that it receives

a. 3 claims in a given week.
b. More than four claims in a given week.
c. Four claims in a given fortnight.
d. No claims on a given day, assuming that the factory operates on a five-day week.

Solution

Let X be the number of claims per week, $X \sim P_o(2)$

a. $P(X = 3) = \frac{e^{-2}2^3}{3!} = 0.180$

Software

```
> dpois(3,2)
```

```
[1] 0.1804470
```

b. $P(X > 4) = 1 - P(X \leq 4)$

$$= 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)]$$

$$= 1 - e^{-2}[1 + 2 + \frac{2^2}{2!} + \frac{2^3}{3!} + \frac{2^4}{4!}]$$

$$= 0.053$$

Software

```
> 1-sum(dpois(0:4,2))

[1] 0.05265302

> 1-ppois(4,2)

[1] 0.05265302

> 1-(dpois(0,2)+dpois(1,2)+dpois(2,2)+dpois(3,2)+dpois(4,2))

[1] 0.05265302
```

c. Let Y be the number of claims in a fortnight.

The mean number of claims in a fortnight is 4 so $Y \sim P_o(4)$

$$P(Y = 4) = \frac{e^{-4}4^4}{4!} = 0.195$$

Software

```
> dpois(4,4)
[1] 0.1953668
```

d. let F be the number of claims in a given day.

The mean number of claims in a day is 0.4. $F \sim P_o(0.4)$

$P(F=0)=\frac{e^{-0.4}0.4^0}{0!}=e^{-0.4}=0.670$

Software

```
> dpois(0,0.4)

[1] 0.67032
```

Question 5

A sales manager receives six telephone calls on average between 9:30am and 10:30am on a weekday. Find the probability that.

a. She will receive two or more calls between 9:30am and 10:30am on Tuesday.
b. She will receive exactly two calls between 9:30am and 9:40am on Wednesday.

Solution

Let X be the number of telephone calls received by the manager between 9:30am-10:30am. $X \sim P_o(6)$

a. $P(X \geq 2) = 1 - P(X < 2)$

$$= 1 - [P(X=0) + P(X=1)]$$
$$= 1 - [\frac{e^{-6}6^0}{0!} + \frac{e^{-6}6}{1!}]$$
$$= 1 - e^{-6}[1+6]$$
$$= 0.983$$

Software

```
> 1-(dpois(0,6)+dpois(1,6))
[1] 0.9826487
> 1-sum(dpois(0:1,6))
[1] 0.9826487
> 1-ppois(1,6)
[1] 0.9826487
```

b. Let Y be the number of telephone calls received by the manager between 9:30am to 9:40am on Wednesday.
The mean number of calls received between 9:30am on Wednesday is 1. $Y \sim P_o(1)$.

$$P(X=2) = \frac{e^{-1}1}{2!}$$
$$= 0.1839$$

Software

```
> dpois(2,1)
[1] 0.1839397
```

Question 6

The number of bacterial colonies on a petri dish can be modeled by a Poisson distribution with average number 2.5 per cm^2. Find the probability that

a. In $1cm^2$ there are no bacterial colonies.

b. In $2cm^2$ there are more than four bacterial colonies.

c. In $4cm^2$ there are six bacterial colonies.

Solution

Let X be the number of bacterial colonies on a petri dish. $X \sim P_o(2.5)$.

a. $P(X=0) = \frac{e^{-2.5}2.5^0}{0!} = e^{-2.5} = 0.082$

Software

```
> dpois(0,2.5)

[1] 0.082085
```

b. Let Y be the number of bacteria on a $2cm^2$ petri dish.
The mean number is $Y \sim P_o(5)$.

$$
\begin{aligned}
P(X > 4) &= 1 - P(X \leq 4) \\
&= 1 - [P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4)] \\
&= 1 - e^{-5}\left[1 + 5 + \frac{5^2}{2!} + \frac{5^3}{3!} + \frac{5^4}{4!}\right] \\
&= 0.559
\end{aligned}
$$

Software

```
> 1-sum(dpois(0:4,5))
[1] 0.5595067
> 1-ppois(4,5)
[1] 0.5595067
> 1-(dpois(0,5)+dpois(1,5)+dpois(2,5)+dpois(3,5)+dpois(4,5))
[1] 0.5595067
```

c. Let B be the number of bacteria on $4cm^2$ petri dish. $B \sim P_o(10)$

$$\begin{aligned} \mathrm{P}(\mathrm{X}=6) &= \frac{\mathrm{e}^{-10}10^6}{6!} \\ &= 0.063 \end{aligned}$$

Software

```
> dpois(6,10)
[1] 0.06305546
```

Question 7

Customers walk into a store at an average rate of 20 per hour. Find the probability that.

a. No customer have arrived at the store in 10min.
b. No more than 4 customers have arrived at the store in 30min.

Solution

a. Let X be the number of customers arriving at the store in 30min with mean 10. $Y \sim P_o(\frac{10}{3})$.

$$P(X=0) = \frac{e^{-\frac{10}{3}}(10/3)^0}{0!} = 0.0356$$

Software

```
> dpois(0,3.3333)
[1] 0.03567518
```

b. Let Y be the number of customers arriving at the store in 30min with mean 10. $Y \sim P_o(10)$

$$P(X \leq 4) = P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4)$$

$$= e^{-10}\left(1 + 10 + \frac{10^2}{2!} + \frac{10^3}{3!} + \frac{10^4}{4!}\right)$$

$$= 0.029$$

Software

```
> ppois(4,10)

[1] 0.02925269

> sum(dpois(0:4,10))

[1] 0.02925269
```

Question 8

The average number misprints on each page in the first draft of novel are four. Find the probability that on a randomly selected double page

a. There are three misprints on each page.

Solution

$\lambda = 4$

$$[P(X = 3)]^2 = [\frac{4^3 e^{-4}}{3!}]^2 = 0.0382$$

software

```
> (dpois(3,4))^2
[1] 0.03816819
```

b. There six misprints in total.

Solution

$\lambda = 4 \times 2$

$$P(X = 6) = \frac{8^6 e^{-8}}{6!} = 0.122$$

software

```
> dpois(6,8)
[1] 0.1221382
```

Using the Poisson distribution as an approximation to the binomial distribution

When n is large$(n > 50)$ and p is small $(p < 0.1)$ the distribution $X{\sim}B(n,p)$ can be approximated using a Poisson distribution with same mean ie $X{\sim}Po(np)$. The approximation gets better as n gets larger and p gets smaller.

Question 1

Eggs are packed into boxes of 500. On average 0.7% of the eggs are to be broken when the eggs are unpacked. Find, correct to 2 significant figures the probability that in a box 500 of eggs,

a. exactly three are broken
b. at least two are broken

Solution

$n = 500 \qquad p = 0.007 \quad E(x) = np = 500 \times 0.007 = 3.5$

Since n > 50 and $p < 0.1$, use a Poisson approximation.

Given $X{\sim}Po(3.5)$

a. $P(x = 3) = e^{-3.5}\frac{3.5^3}{3!} = 0.22$

Software

```
> dpois(3,3.5)

[1] 0.2157855
```

b. $P(X \geq 2) = 1 - (P(X = 0) + P(X = 1)$

$= 1 - (e^{-3.5} + 3e^{-3.5})$

$= 0.86$

Software

```
> 1-ppois(1,3.5)

[1] 0.8641118
```

Question 2

A Christmas draw aims to sell 5000 tickets, 50 of which will win a prize.

a. Calculate P(X ≤ 3)
b. Calculate how many tickets should be bought in order for there to be 90% probability of winning at least- one prize.

Solution

a. P(a ticket-wins a prize) = $\frac{50}{5000} = 0.01$

   $X \sim B(200\,, 0.01)$
   Ans= 0.86

(b) $X \sim B(n\,, 0.01) \quad n > 50 \quad p < \ 0.1 \quad P_0\ (0.01n)$

$P(X\ \geq 1\ ) = 0.9$

$P(X\ \geq 1) = 1 - P(X = 0\ )$

0.9 = 1 - $e^{-0.01n}$

$e^{-0.01n} = 0.1$

-0.01n = ln(0.1)

$n = \frac{\ln(0.1)}{0.01} = 230.25$

So the least integer value of n must be 231

$n = 230 \qquad np = 230 \times 0.01 = 2.3 \quad and \ 1 - e^{-2.3} = 0.8997 < 0.9$

$n = 231 \qquad np = 231 \times 0.01 = 2.31 \quad and \ 1 - e^{-2.31} = 0.9007 > 0.9$

Question 3

X is $B(250, p)$. The value of p is such that it is valid to apply a Poisson approximation. When this is done, it is found that P(X = 0) = 0.0235. Find the value of p.

Solution

E(X)= np

$$P(X = 0) = e^{-250p}\frac{(250p)^0}{0!} = 0.0235$$

$$e^{-250p} = 0.0235$$

$$p = \frac{\ln(0.0235)}{-250} = 0.0150$$

The sum of independent Poisson variables

For independent variables, X and Y , if $X \sim P_0(m) \quad Y \sim P_0\,(n)$ then $X + Y \sim P_o(m + n)$

Question

Two identical racing cars are being tested on a circuit. For each, the number of mechanical breakdowns can be modeled by a Poisson distribution with a mean of one breakdown in 100 lags. If a car breaks down it is attended and continues on the circuit. The first car is tested for 20 lags and second car for 40 lags. Find

The probability that the services team is called out to attend to the breakdowns

a. Once b. more than twice

Solution

$X \sim P_0(0.2) \quad Y \sim P_0(0.4) \quad T = X + Y \quad , \quad P_0(0.6)$

a. $P(T = 1) = 0.6e^{-0.6} = 0.329$

b. $P(T > 2) = 1 - \left(P(T = 0) + P(T = 1) + P(T = 2)\right)$
$= 0.023$

Geometric distribution

Conditions for a geometric model.

For a situation to be described using a geometric model.

1. Independent trials are carried out.
2. The outcome of each trial is deemed either a success or a failure.
3. The probability p of a successful outcome is the same for each trial.

The discrete random variable ,X, is the number of trials needed to obtain the first successful outcome.

$$\boldsymbol{X \sim Geo(p)}$$

The probability of success,p, is all that is needed to describe the distribution completely. It is known as the parameter of the distribution.

If $\sim Geo(p)$, the probability that the first success is obtained at the r^{th} attempt is $P(X = r)$ where

$$P(X = r) = q^{r-1}p \ for\, r = 1,2,3,4, \ldots$$

Expectation and Variance of the Geometric Distribution

If $X \sim Geo(p)$ then

$$E(X) = \frac{1}{p} \qquad Var(X) = \sigma^2 = \frac{q}{p^2} = \frac{1-p}{p^2}$$

Software

dgeom(r-1,p)   pgeom(r-1,p)

Note

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

$P(X \leq x) = 1 - q^x$

$P(X > x) = q^x$

The mode of any geometric distribution is 1

Question 1

Jack is playing a board game in which he needs to throw a six with an ordinary die in order to start the game. Find the probability that.

a. Exactly four attempts are needed to obtain a six.
b. At least two attempts are needed
c. He is successful in throwing a six in three or fewer attempts.

d. He needs more than three attempts to obtain a six.

Solution

X is the number of attempts up to and including the first occurrence of a six.

$p = \frac{1}{6}$ $q = \frac{5}{6}$ $, X \sim Geo(\frac{1}{6})$

a. $P(X = 4) = q^3 p$

$$= (\frac{5}{6})^3 \times (\frac{1}{6})$$

$$= 0.096$$

Software

```
> dgeom(3,1/6)
[1] 0.09645062
```

b. $P(X \geq 2) = 1 - P(X = 1)$

$$= 1 - p$$

$$= 1 - \frac{1}{6}$$

$$= \frac{5}{6}$$

Software

```
> 1-dgeom(0,1/6)
[1] 0.8333333
```

c. $P(X \leq 3) = P(X = 1) + P(X = 2) + P(X = 3)$

$$= p + qp + q^2p$$

$$= \frac{1}{6} + \left(\frac{5}{6} \times \frac{1}{6}\right) + \left(\frac{5}{6}\right)^2 \frac{1}{6}$$

$$= 0.42$$

Or

$P(X \leq 3) = 1 - q^3$

$$= 1 - \left(\frac{5}{6}\right)^3$$

$$= 0.42$$

Software

```
> sum(dgeom(0:2,1/6))

[1] 0.4212963

> pgeom(2,1/6)

[1] 0.4212963
```

d. $P(X > 3) = 1 - P(X \leq 3)$

$$= 1 - (1 - q^3)$$

$$= q^3$$

$$= (\frac{5}{6})^3 = 0.58$$

Software

```
> 1-sum(dgeom(0:2,1/6))

[1] 0.5787037

> 1-pgeom(2,1/6)
```

```
[1] 0.5787037
```

Question 2

On a particular production line the production that an item is faulty is 0.08.In a quality control test; items are selected at random from the production line. It is assumed that quality of an item is independent of that of other items.

a. Find the probability that the first faulty item
   i. Does not occur in the first six selected
   ii. Occurs in fewer than five selections.

Solution

X is the number of items picked until a faulty one is selected.

Using a geometric model with $p = 0.08, q = 0.92, X \sim Geo(0.08)$.

i. $P(X \leq 6) = q^6 = 0.92^6 = 0.61$

   Software

```
> 1-pgeom(5,0.08)
[1] 0.606355
> 1-sum(dgeom(0:5,0.08))
[1] 0.606355
> 1-
(dgeom(0,0.08)+dgeom(1,0.08)+dgeom(2,0.08)+dgeom(3,0.08)+
dgeom(4,0.08)+dgeom(5,0.08))
[1] 0.60635
```

ii. $P(X < 5) = P(X \leq 4) - 1 - q^4 = 1 - 0.92^4 = 0.28$

Software

```
> sum(dgeom(0:3,0.08))
[1] 0.2836070
> pgeom(3,0.08)
[1] 0.2836070
> dgeom(0,0.08)+dgeom(1,0.08)+dgeom(2,0.08)+dgeom(3,0.08)
[1] 0.2836070
```

Question 3

Identical independent trials of an experiment are carried out. The probability of a successful outcome is p. On average, five trials are required until a successful outcome occurs.

a. Find the value of p.
b. Find the probability that the first successful outcome occurs on the fifth trial.

Solution

X is the number of trials up to and including the first success.

$X \sim Geo(P)\ and\ E(X) = 5$

a. $E(X) = \frac{1}{p}$

$$5 = \frac{1}{p}$$

$$p = 0.2$$

b. $X \sim Geo(0.2)$ $i.e\ p = 0.2 \quad q = 0.8$

$$P(X = 5) = q^4 p$$
$$= 0.8^4 \times 0.2$$
$$= 0.08192$$

Software

```
> dgeom(4,0.2)

[1] 0.08192
```

Question 4

$X \sim Geo(p)$ and it is known that $P(X = 2) = 0.21$ and $p < 0.5$. Find $P(X = 1)$.

Solution

$P(X = 2) = qp \;\; where\ q = 1 - p$

$0.21 = (1 - p)p$

$p^2 - p + 0.21 = 0$

$(p - 0.3)(p - 0.7) = 0$

$p = 0.3 \;\; or \;\; p = 0.7$

$Since\ p < 0.5\ , p = 0.3$

$P(X = 1) = p = 0.3$

Software

```
> dgeom(0,0.3)

[1] 0.3
```

Question 5

$X \sim Geo(0.5)$ .find

a. The mode
b. The mean of X,
c. The standard deviation of X.

Solution

a. 1
b. E(X)=$\frac{1}{p} = \frac{1}{0.5} = 2$
c. Sd=$\sqrt{\frac{q}{p^2}} = \sqrt{\frac{0.5}{0.5^2}} = \sqrt{2}$

Question 6

The random variable X is$X \sim Geo(0.35)$. Calculate

a. $P(X = 4)$ b.$P(X > 4)$ c.$P(X \leq 3)$ d.$E(X)$

Solution

a. $P(X = 4) = q^3 p$

$$= 0.65^3 \times 0.35 = 0.096$$

software

```
> dgeom(3,0.35)
[1] 0.09611875
```

b. $P(X > 4) = q^4$

$$= 0.65^4$$
$$= 0.179$$

Software

```
> 1-pgeom(3,0.35)
[1] 0.1785062
> 1-sum(dgeom(0:3,0.35))
[1] 0.1785063
```

c. $P(X \leq 3) = 1 - q^3$

$$= 1 - 0.65^3$$
$$= 0.725$$

Software

```
> dgeom(0,0.35)+dgeom(1,0.35)+dgeom(2,0.35)
[1] 0.725375
> sum(dgeom(0:2,0.35))
[1] 0.725375
> pgeom(2,0.35)
```

[1] 0.725375

d. $E(X) = \frac{1}{p}$

$$= \frac{1}{0.35}$$

Question 7

$X \sim Geo(p)$ and the probability that the first success is obtained on the second attempt is 0.1275.If p > 0.5, find P(X>2)

Solution

$P(X = 2) = qp = 0.1275$

$(1 - p)p = 0.1275$

$p^2 - p + 0.1275 = 0$

$(p - 0.85)(p - 0.15) = 0$

$p = 0.85 \ or \ p = 0.15$

$since \ p > 0.5 \ ; then \ p = 0.85$

$P(X > 2) = q^2 = 0.15^2 = 0.0225$

Negative Binomial

A generalization of a geometric distribution in which the random variable is the number of Bernoulli trials required to obtain r success result in the negative binomial distribution. Let the random variable X denote the number of trials until r success occur. Then X is a negative binomial random variable with parameters 0 < p < 1 and r = 1, 2, 3, … and

$$P(X = x) = \binom{x-1}{r-1}(1-p)^{x-r}p^r \quad x = r, r+1, r+2, \ldots$$

If X is a negative binomial random variable with parameter p and

$$\mu = E(X) = \frac{r}{p} \quad \sigma^2 = V(X) = \frac{r(1-p)}{p^2}$$

Software

```
dnbinom(x-r,r,p)
```

Question 1

Suppose that X is a negative binomial random variable with $p = 0.2$ and $r = 4$. Determine the following

a. E(X) b.P(X=19) c. P(X=20) d.P(X=21)

Solution

$P = 0.2 \ r = 4 \ q = 0.8$

a. $Mean = E(X) = \frac{4}{0.2} = 20$

b. $\mathrm{P(X = 20)} = \binom{19}{3}(0.80)^{16}\, 0.2^4 = 0.0436$

Software

```
> dnbinom(16,4,0.2)

[1] 0.04363988
```

c. $\mathrm{P(X = 19)} = \binom{18}{3}(0.80)^{15}\, 0.2^4 = 0.0459$

Software

```
> dnbinom(15,4,0.2)

[1] 0.04593672
```

d. $\mathrm{P(X = 21)} = \binom{20}{3}(0.80)^{17}\, 0.2^4 = 0.0411$

Software

```
> dnbinom(17,4,0.2)

[1] 0.04107283
```

Hyper geometric distribution

A set of N objects contains
K objects classified as successes
N - K objects classified as failures
A sample of size n objects is selected randomly (without replacement) from the N objects, where $K \leq N$ and $n \leq N$. let the random variable X denote the number of successes in the sample. Then X is a hyper geometric random variable and

$$P(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \quad x = \max\{0, n + K - N\} \, to \, \min\{K, n\}$$

if X is a hyper geometric random variable with parameters N,K and n then

$$\mu = E(X) = np \; and \;\; \sigma^2 = V(X) = np(1-p)\left(\frac{N-n}{N-1}\right) \quad where \; p = \frac{k}{N}$$

Software
dhyper(prob,K,N-K,n)

CONTINUOUS DISTRIBUTION

Normal Distribution

$Z \sim N(0,1) \quad Z = \frac{X-\mu}{\sigma}$

a. $P(Z < 1.377)$

Software

```
> pnorm(1.377)
[1] 0.9157439
```

b. $P(Z > -1.377)$

Software

```
> 1-pnorm(-1.377)
[1] 0.9157439
```

c. $P(Z > 1.377)$

software

```
> 1-pnorm(1.377)
[1] 0.08425612
```

d. $P(Z < -1.377)$

software

```
> pnorm(-1.377)
[1] 0.08425612
```

e. $P(0.345 < Z < 1.751)$

software

```
> pnorm(1.751)-pnorm(0.345)
[1] 0.3250742
```

f. $P(-2.696 < Z < 1.865)$

Software

```
> pnorm(1.865)-pnorm(-2.696)
[1] 0.9654004
```

g. $P(-1.4 < Z < 0.6)$

Software

```
> pnorm(0.6)-pnorm(-1.4)
[1] 0.6449902
```

h. $P(|Z| < 1.433) = P(Z < -1.433) + P(Z > 1.433)$

Software

```
> pnorm(-1.433)+1-pnorm(1.433)
[1] 0.1518578
```

Try

If $Z \sim N(0,1)$ find;

Question 1

a. $P(Z < 0.874)$
b. $P(Z > -0.874)$
c. $P(Z > 0.874)$
d. $P(Z < -0.874)$

Question 2

If $Z \sim N(0,1)$ find;

a. $P(Z > 1.8)$
b. $P(Z < -0.65)$
c. $P(Z > -2.46)$
d. $P(Z < 1.36)$
e. $P(Z > 2.56)$

Question 3

If $Z \sim N(0,1)$ find;

a. $P(0.829 < z < 1.834)$
b. $P(-2.56 < z < 0.134)$
c. $P(-1.762 < z < -0.246$
d. $P(0 < z < 1.73)$
e. $P(-2.05 < z < 0)$
f. $P(Z < -1.97 \; or \; Z > 2.5)$
g. $P(|Z| > 0.754)$
h. $P(|Z| < 1.78)$

Question 4

The time taken by the milkman to deliver to high street is normally distributed with a mean of 12mins and standard deviation of 2mins. He delivers every day. Estimate the number of days during the year when he takes

a. Longer than 17mins
b. Less than 10mins
c. Between 9 and 13mins

Question 5

The life of a certain electric bulb is normally distributed with a mean life of 2000hs and standard deviation of 120hrs. Estimate the probability that the life of such a bulb will be

a. Greater than 2150hrs
b. Greater than 1910hrs
c. Within the range 1850hrs to 2090hrs.

Question 6

The masses of a certain type of cabbage are normally distributed with mean of 1000g and a standard deviation of 0.15kg.In a batch of 800 cabbages; estimate how many have a mass between 750g and 1290g.

Question 7

The weights of vegetable marrows supplied to retailers by a wholesaler have a normal distribution with mean 1.5kg and standard deviation 0.6kg. The wholesaler supplies three sizes of marrow

Size 1, under 0.9kg

Size 2, from 0.9kg to 2.4kg

Size 3, over 2.4kg

Question 8

$X \sim N(100,81). Find$

a. $P(|X-100|<18)$
b. $P(|X-100|>5)$
c. $P(12<X-100<15)$

Uniform Distribution

Software

punif(X,a,b)

Question 1

Suppose that buses arrive at bus stop every 15mins and that the waiting time for the bus to arrive has a uniform probability distribution on the interval from 0 to 15mins

a. What is the probability that X will exceed 10mins?
b. What is the probability that X will be at most 12mins?

Question 2

The probability density function of time X required to complete an assembly operation is uniformly distributed for $30 \le X \le 40$sec. Determine the proportion assemblies that require less than 35sec to complete.

Exponential Distribution

$$f(x) = \begin{cases} \theta e^{-\theta x}, & x \ge 0 \\ 0 & , elsewhere \end{cases}$$

Question 1

Suppose the time in days between service calls on a photocopier machine follows an exponential distribution with mean call of 0.02 per day.

a. What is the probability that the time until the machine again requires service exceeds 60days? ans (0.3011942)
b. What is the probability that the time until the machine again requires service is less than 20days. ans (0.32968)

Question 2

The lifetime of a mechanical assembly in a vibration test is exponentially distributed with mean of 400hrs. what is the probability that

a. An assembly on test fails is less 100hrs
b. An assembly operates more than 500hrs before life.
c. An assembly on test fails is at most 200hrs.

NORMAL DISTRIBUTION

1. Standardize a normal variable and use standard normal tables.
2. Use the normal distribution as model to solve problems
3. Use the normal distribution as an approximation to the binomial distribution and to the Poisson distribution.

The normal distribution is one of the most important distributions in statistics. Many measured quantities in the natural sciences follow a normal distribution and under certain circumstances it's also a useful approximation to the binomial distribution and to the Poisson distribution.

The normal variable X is continuous. Its probability density function f(x) depends on its mean μ and standard deviation σ,

where $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty$

It can be describe as $X \sim N(\mu, \sigma^2)$

Finding probabilities

The probability that X lies between a and b is written $P(a < X < b)$. To find this probability, you need to find the area under the normal curve between a and b.

One way of finding areas is to integrate, but since the normal function is complicated and very difficult to integrate, tables are used instead

The standard normal variable, Z

In order to use the same set of tables for all possible values of $\mu\, and\, \sigma^2$, the variable X is standardised so that the mean is 0 and the standard deviation is 1. Notice that since the variance is the square of the standard deviation, the variance is also 1. This standardized normal variable is called Z and $Z \sim N(0\,,1)$.

In general

To standardize X, where $X \sim N(\,\mu\,,\sigma^2)$

1. Subtract the mean μ
2. Then divide by the standard deviation σ

   Therefore $Z = \frac{X-\,\mu}{\sigma}\;\; where\, Z \sim N(0\,,1)$

Finding the Z- value from standard table

1) Find the value of Z using the standard tables.

i)P(Z < 0.85) ii)P(Z > 0.85) iii)P(Z < -1.38)   iv)P(Z > -1.38)

Solution

i) 0.8023 ii) 1 – 0.823 = 0.1977   iii) 0.0838   iv) 1 – 0.0838 = 0.9162

2) Find the following

a) P(0.35 < Z < 1.76) b) P(-2.70 < Z < 1.87)   c) P($|z|$ < 1.43)   d) P($|Z|$ > 1.433

Solution

$(a)\ \emptyset(1.76) - \emptyset(0.35) = 0.9608 - 0.6368 = 0.324$

$(b)\ \emptyset(1.87) - \emptyset(-2.70) = 0.9693 - 0.0035 = 0.9658$

$(c)\ P(|Z| < 1.43$

$P(-1.433 < Z < 1.433)$

$\emptyset(1.433) - \emptyset(-1.433) = 0.9236 - 0.0764 = 0.8472$

OR

$2\emptyset(1.43) - 1 = 2(0.9236) - 1 = 0.8472$

$d)\ P(|Z| > 1.433)$

P (Z < -1.433) + P (Z > 1.433)

$2\big(1 - \emptyset(1.433)\big)$

$2(1 - 0.9236) = 0.1528$

$Z \sim N(0\ ,1).\ show\ that$

$P(-1.96 < Z < 1.96) = 0.95$

$P(-2.58 < Z < 2.58) = 0.99$

Using standard normal tables for any normal variable X

X given as $X \sim N(\mu, \sigma^2)$ can be standard as $Z = \frac{X-\mu}{\sigma}$ where $Z \sim N(0,1)$

Question

Length of metal strips produced by a machine are normally distributed with mean length of 150cm and a standard deviation of 10cm. Find the probability that the length of a randomly selected strip is

a) Shorter than 165cm
b) Within 5cm of the mean

Solution

a. X is the length ; $\mu = 150cm$; $\sigma = 10$; $X \sim N(150, 10^2)$. To find prob that- the length is shorter than 165cm ie P(X < 165)

$$Z = \frac{X-150}{10} = \frac{165-150}{10} = 1.5$$

So $P(X < 165)$ *becomes* $(Z < 1.5)$

$P(X < 165) = P(Z < 1.65) = \emptyset(1.5) = 0.9332$

b. To find the probability that length is within 5cm of the mean, you need to find $P(|X - 150| < 5)$

$$P(-5 < X - 150 < 5)$$

$$P\left(\frac{-5}{10} < \frac{X-150}{10} < \frac{5}{10}\right)$$

$$P(-0.5 < Z < 0.5)$$

$P(|Z| < 0.5) = 2\emptyset(0.5) - 1 = 2 \times 0.6915 - 1 = 0.38$

The probability that the length is within 5cm of the mean is 0.38

Question

The time taken by milkman to deliver to the high street is normally distributed with a mean of 12 minutes and a standard deviation of 2 minutes. He delivers milk every day. Estimate the number of days during the year when he takes

a. Longer than 17 minutes
b. less than ten minutes
c. between nine and 13 minutes.

Solution

X is the time, in minutes, taken to deliver milk to the high street. $X \sim N(12, 2^2)$

Standardized X using $Z = \frac{X-\mu}{\sigma}$ ie $Z = \frac{x-12}{2}$

a) $P(X > 17) = P\left(Z > \frac{17-12}{2}\right)$

= P(Z > 2.5)

= $1 - \emptyset(2.5) = 1 - 0.9938$

= 0.0062

Find the number of days multiply by 365

$365 \times 0.0062 = 2.263 \approx 2$

On two days in the year he takes longer than 17 minutes.