



Free download & print from www.itute.com  
Do not reproduce by other means  
© Copyright itute.com 2005

## Probability

### Probability experiment, outcomes and sample space

An experiment is the process of collecting data that shows variation in its outcomes. The set of all possible distinct outcomes is called the sample space of the experiment.

### Probability experiment, discrete random variables and probability distributions in table form

A random variable  $X$  is a numerical valued function defined on a sample space, i.e. if  $e$  is an outcome of the sample space, then  $X(e)$  is a numerical value assigned to  $e$ .

Example 1 *Probability experiment*: A fair die is rolled and the uppermost number is recorded.

Examples of *discrete random variables* (denoted by capital letters):

$P$ : The number that appears uppermost.

$Q$ : Number of odd uppermost numbers.

$R$ : Number of uppermost numbers that are less than 3.

*Probability distributions* of  $P$ ,  $Q$  and  $R$  (each one is displayed as a table showing the values of the random variable and the corresponding probabilities):

$x$	1	2	3	4	5	6
$\Pr(P = x)$	1/6	1/6	1/6	1/6	1/6	1/6

$x$	0	1
$\Pr(Q = x)$	3/6	3/6

$x$	0	1
$\Pr(R = x)$	4/6	2/6

Notes: (1)  $0 \leq \Pr \leq 1$ .

(2) For a probability distribution,  $\sum \Pr = 1$ .

Example 2 *Probability experiment*: Two fair dice are rolled and the uppermost numbers are recorded.

Examples of discrete random variables:

$A$ : The difference of the two uppermost numbers.

$B$ : Number of even uppermost numbers.

$C$ : Number of equal uppermost numbers.

Probability distributions:

$a$	0	1	2	3	4	5
$\Pr(A = a)$	6/36	10/36	8/36	6/36	4/36	2/36

$b$	0	1	2
$\Pr(B = b)$	9/36	18/36	9/36

$c$	0	2
$\Pr(C = c)$	30/36	6/36

Example 3 *Probability experiment*: An unbiased coin is tossed three times and the result (H or T) of each toss is recorded.

Examples of random variables:

$X$ : Number of tails in the three tosses.

$Y$ : Difference between the number of heads and the number of tails in the three tosses.

$W$ : Square of the number of heads in the three tosses.

Probability distributions:

$x$	0	1	2	3
$\Pr(X = x)$	1/8	3/8	3/8	1/8

$x$	1	3
$\Pr(Y = x)$	6/8	2/8

$x$	0	1	4	9
$\Pr(W = x)$	1/8	3/8	3/8	1/8

In examples 1, 2 and 3, the outcomes in each experiment are **equally likely**,  $\therefore$  the probabilities can be calculated by

$$\text{probability} = \frac{\text{number of favourable outcomes}}{\text{total number of outcomes}}$$

Example 4 A probability experiment has the probability distribution of the discrete random variable  $V$  shown in the following table.

$x$	-1.3	1.2	10.3	11.1
$\Pr(V = x)$	0.31	0.22	0.4	0.07

(a) State all the possible values of  $V$ .

(b) State the mode of the distribution of  $V$ .

(c) Find the probability that  $V < 10.3$ .

(d) Find the probability that  $V < 1.3$ .

(e) Find the probability that  $0 < V < 15$ .

(a) Possible values of  $V$  are: -1.3, 1.2, 10.3 and 11.1. There are only four because  $\sum \Pr = 0.31 + 0.22 + 0.4 + 0.07 = 1$

(b) **Mode** is the value of a random variable that has the highest probability. In this example, the mode is 10.3.

(c)

$$\Pr(V < 10.3) = \Pr(V = -1.3) + \Pr(V = 1.2) = 0.31 + 0.22 = 0.53$$

$$(d) \Pr(V < 1.3) = \Pr(V = -1.3) + \Pr(V = 1.2) = 0.31 + 0.22 = 0.53$$

$$(e) \Pr(0 < V < 15) = \Pr(V = 1.2) + \Pr(V = 10.3) + \Pr(V = 11.1) = 0.22 + 0.4 + 0.07 = 0.69.$$

Alternatively,

$$\Pr(0 < V < 15) = 1 - \Pr(V = -1.3) = 1 - 0.31 = 0.69.$$

## Probability functions of discrete random variables

The probability distribution of a discrete random variable can be specified by a rule and some parameters in some cases instead of a table, e.g. for the discrete random variable  $X$  in example 3, the **rule** is  $p(x) = \Pr(X = x) = {}^n C_x (k)^x$  for  $x = 0, 1, 2, 3$ , where  $n = 3$  and  $k = \frac{1}{2}$  are the **parameters** of the probability distribution of  $X$ .

**Example 1** Given the rule of the probability function for discrete random variable  $H$ ,  $p(x) = \Pr(H = x) = 0.1(x^2 + 1)$  for  $x = -1, 0, 1, 2$ , set up a table for the probability distribution of  $H$ .

$x$	-1	0	1	2
$\Pr(H = x)$	0.2	0.1	0.2	0.5

Note: For  $p(x)$  to be a probability function,  $\sum p(x) = 1$ .

**Example 2** Given the rule of the probability function for discrete random variable  $K$ ,  $p(x) = \Pr(K = x) = \frac{e^{1-\frac{x^2}{2}}}{e+2\sqrt{e}}$  for  $x = -1, 0, 1$ , set up a table for the probability distribution of  $K$ . Show that  $\sum p(x) = 1$ .

$x$	-1	0	1
$\Pr(K = x)$	~0.274	~0.452	~0.274

$$\sum p(x) = \sum \Pr = \frac{\sqrt{e}}{e+2\sqrt{e}} + \frac{e}{e+2\sqrt{e}} + \frac{\sqrt{e}}{e+2\sqrt{e}} = \frac{e+2\sqrt{e}}{e+2\sqrt{e}} = 1$$

**Example 3** Five cards are drawn from a deck of 52 playing cards without replacement. Let the discrete random variable  $S$  be the number of diamond cards in the five drawn cards. The probability function for  $S$  is given by

$$p(x) = \Pr(S = x) = \frac{{}^{13}C_x \times {}^{39}C_{5-x}}{{}^{52}C_5} \text{ for } x = 0, 1, 2, 3, 4, 5.$$

The parameters are 52, 13 and 5.

Use the probability function to set up a table for the probability distribution of  $S$ .

$x$	0	1	2	3	4	5
$\Pr(S = x)$	0.222	0.411	0.274	0.082	0.011	0.005

Note: Due to rounding errors,  $\sum p(x) = 1.005 \neq 1$ .

## Expected value and variance of a discrete random variable

The commonly used measure of the centre of a probability distribution of a random variable  $X$  is called the **expected value** (the **mean**, or **expectation**) of  $X$ , and it is denoted as  $E(X)$ , or  $\mu_x$ . If the random experiment is repeated many times and the average value of the random variable is calculated, it approaches the expected value as the number of repetitions increases.

Other numerical measures of centre are **median** and **mode** of a random variable.

The median  $m$  of random variable  $X$  has the following properties:  $\Pr(X \leq m) \geq 0.5$  and  $\Pr(X \geq m) \geq 0.5$ .

The mode of  $X$  is the value of  $X$  that has the highest probability.

The spread of the probability distribution is measured by a number called the **variance** of  $X$ , and it is denoted as  $Var(X)$ . Another number called the **standard deviation** of  $X$  is often used instead of the variance to measure the spread. It is denoted by  $sd(X)$ , or  $\sigma_x$ , and  $sd(X) = \sqrt{Var(X)}$ . In many cases, if the random experiment is repeated many times, ~68% of the recorded values of  $X$  lie within one  $\sigma_x$  of  $\mu_x$ , and ~95% of the recorded values of  $X$  lies within  $2\sigma_x$  of  $\mu_x$ .

∴ in a single trial of the random experiment,

$$\Pr(\mu - \sigma \leq X \leq \mu + \sigma) \approx 0.68, \text{ and}$$

$$\Pr(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 0.95.$$

## Calculation of expected value of a discrete random variable

$$\mu_x = E(X) = \sum x \Pr(X = x)$$

**Example 1** Find the expected value of the following probability distribution.

$x$	-2	0	1	10
$\Pr(X = x)$	0.26	0.33	0.38	0.03

$$E(X) = -2(0.26) + 0(0.33) + 1(0.38) + 10(0.03) = 0.16$$

**Example 2** Find the expected value of  $X$  with probability

$$\text{function } \Pr(X = x) = {}^3C_x \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{3-x} \text{ for } x = 0, 1, 2, 3.$$

$$E(x) = 0 \left\{ {}^3C_0 \left(\frac{1}{4}\right)^0 \left(\frac{3}{4}\right)^3 \right\} + 1 \left\{ {}^3C_1 \left(\frac{1}{4}\right)^1 \left(\frac{3}{4}\right)^2 \right\} + 2 \left\{ {}^3C_2 \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^1 \right\} + 3 \left\{ {}^3C_3 \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^0 \right\} = 0.75.$$

Note: It would be easier to set up the probability distribution table and use it to calculate  $E(X)$ . See example 2 in the next section.

### Calculation of $Var$ and $\sigma$ of a discrete random variable

The variance of discrete random variable  $X$  is defined as

$$Var(X) = E((X - \mu_X)^2) = \sum (x - \mu_X)^2 \Pr(X = x).$$

It can be shown that the above definition is equivalent to

$$Var(X) = E(X^2) - (\mu_X)^2 = \sum x^2 \Pr(X = x) - (\mu_X)^2.$$

The last formula is easier to use than the first in calculating variance.

**Example 1** Calculate the variance and standard deviation of the following discrete probability distribution.

$x$	-2	0	1	10
$\Pr(X = x)$	0.26	0.33	0.38	0.03

Firstly calculate the mean of  $X$ ,  $\mu_X = E(X) = 0.16$  (See previous example 1).

$$Var(X) = (-2)^2(0.26) + 0^2(0.33) + 1^2(0.38) + 10^2(0.03) - 0.16^2 \approx 4.4$$

$$\sigma_X = \sqrt{Var(X)} = \sqrt{4.4} \approx 2.1.$$

**Example 2** Find the variance and standard deviation of  $X$  with

probability function  $\Pr(X = x) = {}^3C_x \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{3-x}$  for  $x = 0, 1, 2, 3$ .

Set up the probability distribution table for  $X$ .

$x$	0	1	2	3
$\Pr(X = x)$	$\frac{27}{64}$	$\frac{27}{64}$	$\frac{9}{64}$	$\frac{1}{64}$

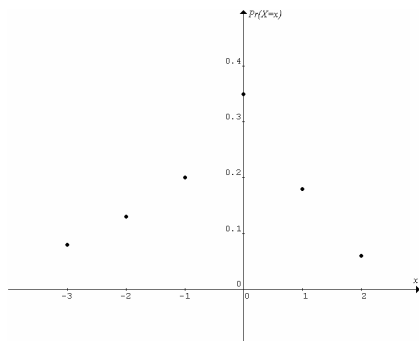
$$\mu_X = 0\left(\frac{27}{64}\right) + 1\left(\frac{27}{64}\right) + 2\left(\frac{9}{64}\right) + 3\left(\frac{1}{64}\right) = 0.75$$

$$Var(X) = 0^2\left(\frac{27}{64}\right) + 1^2\left(\frac{27}{64}\right) + 2^2\left(\frac{9}{64}\right) + 3^2\left(\frac{1}{64}\right) - 0.75^2 = 0.5625$$

$$\sigma_X = \sqrt{0.5625} = 0.75$$

**Example 3** Calculate the mean, variance and standard deviation of the following probability distribution.

Find  $\Pr(\mu - \sigma \leq X \leq \mu + \sigma)$  and  $\Pr(\mu - 2\sigma \leq X \leq \mu + 2\sigma)$ .



Set up the probability distribution table for  $X$ .

$x$	-3	-2	-1	0	1	2
$\Pr(X = x)$	0.08	0.13	0.2	0.35	0.18	0.06

$$\mu_X = -3(0.08) + (-2)(0.13) + (-1)(0.2) + 0(0.35) + 1(0.18) + 2(0.06) = -0.4$$

$$E(X) = (-3)^2(0.08) + (-2)^2(0.13) + (-1)^2(0.2) + 0^2(0.35) + 1^2(0.18) + 2^2(0.06) - (-0.4)^2 = 1.7.$$

$$\sigma_X = \sqrt{1.7} \approx 1.3$$

$$\Pr(\mu - \sigma \leq X \leq \mu + \sigma) = \Pr(-1.7 \leq X \leq 0.9) = 0.2 + 0.35 = 0.55$$

$$\Pr(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = \Pr(-3.0 \leq X \leq 2.2)$$

$$= 0.08 + 0.13 + 0.2 + 0.35 + 0.18 + 0.06 = 1.$$

**Expected value and variance of random variable  $Y = kX + b$**

$$E(Y) = E(kX + b) = kE(X) + b$$

$$Var(Y) = Var(kX + b) = k^2 Var(X)$$

$$sd(Y) = sd(kX + b) = k.sd(X)$$

**Example 1** Find the expected value, the variance and the standard deviation of  $Y = 2X - 3$ , where  $X$  is defined in the previous example 1.

$$E(X) = 0.16, Var(X) = 4.4, sd(X) = 2.1.$$

$$\therefore E(Y) = E(2X - 3) = 2E(X) - 3 = 2 \times 0.16 - 3 = -2.68$$

$$Var(Y) = Var(2X - 3) = 2^2 Var(X) = 4 \times 4.4 = 17.6$$

$$sd(Y) = sd(2X - 3) = 2.sd(X) = 2 \times 2.1 = 4.2 \text{ or}$$

$$sd(Y) = \sqrt{Var(Y)} = \sqrt{17.6} = 4.2$$

**Example 2** The probability distribution of the weekly sales,  $X$ , of a particular brand of iPod has  $\mu = 5.7$  and  $\sigma = 2.8$ . If the fixed cost in selling this brand of iPod is \$35 per week and the profit on each unit sold is \$38, write down the random variable  $Y$ , the weekly net profit, in terms of  $X$ . Hence find the expected weekly net profit and the standard deviation of the weekly net profit. Find the range of weekly net profit that the shopkeeper expects to get with 0.68 chances.

$$Y = 38X - 35$$

$$E(Y) = 38E(X) - 35 = 38 \times 5.7 - 35 = 181.60$$

$$sd(Y) = 38 \times sd(X) = 38 \times 2.8 = 106.40$$

The shopkeeper expects  $\Pr(\mu - \sigma \leq Y \leq \mu + \sigma) \approx 0.68$ ,  $(181.60 - 106.40) \leq Y \leq (181.60 + 106.40)$ , i.e.  $75.20 \leq Y \leq 288$ .

## Bernoulli trials and Markov chains

Here we consider experiments consisting of successive repetitions. Each repetition is called a trial. Furthermore, we assume that there are only two possible results (success  $S$  or failure  $F$ ) for each trial. The use of  $S$  and  $F$  is to emphasize the point that they are the only possible results and do not bear the same connotation as success or failure in real life, e.g. failing an exam may be labeled as a success. For an experiment consisting of 5 trials, a possible outcome as an example is  $SFFSF$ .

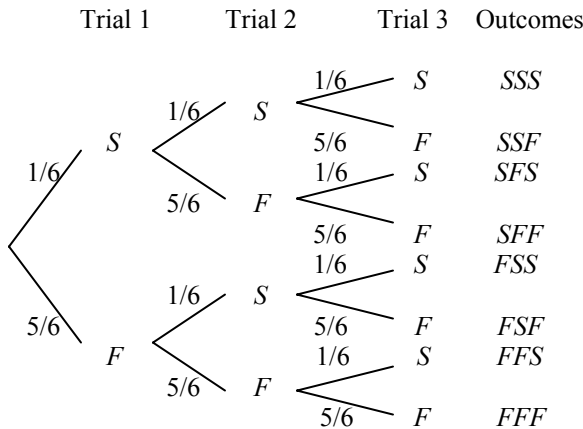
Let  $p$  be the probability of success and  $\therefore 1 - p = q$  is the probability of failure in a trial.

If  $p$  and hence  $q$  remain constant from trial to trial, then we say the experiment consists of **Bernoulli trials**.

If  $p$  and hence  $q$  depend only on the preceding trial result, then we say the experiment is a **Markov chain**.

The following tree diagrams serve to highlight the difference between the two experiments (consisting of 3 trials each).

*Bernoulli trials:* Roll a fair die 3 times. Success means getting 1, and failure means not getting 1.  $p = \frac{1}{6}$  and  $\therefore q = 1 - \frac{1}{6} = \frac{5}{6}$ .



Define random variable  $X$  on the sample space as the number of successes in an outcome.

Outcome with  $X = 0$  is:  $FFF$

Outcomes with  $X = 1$  are:  $SFF, FSF, FFS$

Outcomes with  $X = 2$  are:  $SSF, SFS, FSS$

Outcome with  $X = 3$  is:  $SSS$

Probability distribution of  $X$ :

$x$	0	1	2	3
$\Pr(X = x)$	$1\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^0$	$3\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^1$	$3\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^2$	$1\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^3$

Note:  $\sum \Pr(X = x) = 1$ .

The numbers 1, 3, 3, 1 in the probability calculations are the number of ways in arranging  $S$  and  $F$  in a sequence, e.g. for the sequence consisting of 2 indistinguishable  $S$ 's and 1  $F$ , there are  $\frac{3!}{1!2!} = 3$  arrangements. They are  $SSF, SFS$  and  $FSS$ .

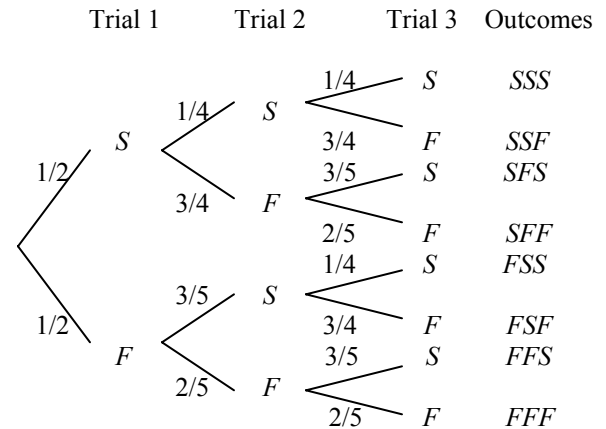
$\frac{3!}{1!2!}$  is the definition of  ${}^3C_1$  or  ${}^3C_2$ .

In general for a sequence consisting of  $x$   $S$ 's and  $(n - x)$   $F$ 's, the number of arrangements of  $S$  and  $F$  is  ${}^nC_x$  or  ${}^nC_{n-x}$ .

For an experiment consisting of  $n$  Bernoulli trials,

$\Pr(X = x) = {}^nC_x (p^x)(1 - p)^{n-x}$ ,  $x = 0, 1, 2, 3, \dots, n$ . This is known as the **binomial distribution** of  $X$ . The number of trials  $n$  and the probability of success  $p$  are the parameters of the distribution. A binomial distribution with parameters  $n$  and  $p$  is denoted as  $Bi(n, p)$ .

*Markov chain:* Consider an experiment consisting of 3 trials. In the first trial,  $p$  and  $q$  are equal. In the subsequent trials,  $p$  and  $q$  depend on the preceding trial result.



Define random variable  $X$  on the sample space as the number of successes in an outcome.

$$\Pr(X = 0) = \Pr(FFF) = \left(\frac{1}{2}\right)\left(\frac{2}{5}\right)\left(\frac{2}{5}\right) = 0.08$$

$$\begin{aligned} \Pr(X = 1) &= \Pr(SFF \cup FSF \cup FFS) \\ &= \Pr(SFF) + \Pr(FSF) + \Pr(FFS) \\ &= \left(\frac{1}{2}\right)\left(\frac{3}{4}\right)\left(\frac{2}{5}\right) + \left(\frac{1}{2}\right)\left(\frac{3}{5}\right)\left(\frac{3}{4}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{5}\right)\left(\frac{3}{5}\right) = 0.495 \end{aligned}$$

$$\begin{aligned} \Pr(X = 2) &= \Pr(SSF \cup SFS \cup FSS) \\ &= \Pr(SSF) + \Pr(SFS) + \Pr(FSS) \\ &= \left(\frac{1}{2}\right)\left(\frac{1}{4}\right)\left(\frac{3}{4}\right) + \left(\frac{1}{2}\right)\left(\frac{3}{4}\right)\left(\frac{3}{5}\right) + \left(\frac{1}{2}\right)\left(\frac{3}{5}\right)\left(\frac{1}{4}\right) = 0.39375 \end{aligned}$$

$$\Pr(X = 3) = \Pr(SSS) = \left(\frac{1}{2}\right)\left(\frac{1}{4}\right)\left(\frac{1}{4}\right) = 0.03125$$

Probability distribution of  $X$ :

$x$	0	1	2	3
$\Pr(X = x)$	0.08	0.495	0.39375	0.03125

Note:  $\sum \Pr(X = x) = 1$ .

The two distributions above are typical examples of discrete probability distributions.

## Expected value and variance of random variable $X$ with a binomial distribution

For a binomial random variable  $X$  with parameters  $n$  and  $p$ , it can be shown that  $E(X) = \sum x \Pr(X = x) = np$  and

$$\text{Var}(X) = E(X^2) - (\mu_x)^2 = np(1-p).$$

Instead of using the general methods for discrete random variables in calculating  $E(X)$  and  $\text{Var}(X)$ , a quicker way is to use  $E(X) = np$  and  $\text{Var}(X) = np(1-p)$ .

Note: The last two formulas are for binomial distributions only.

**Example 1** From past experience 5% of students of a large school are out of uniform on any school day. Five students are selected randomly from the school to check for uniform. (a) Set up the probability distribution of the number of students out of uniform. (b) Find the probability that at most one is out of uniform. (c) Calculate the mean and standard deviation of the distribution.

(a)

Probability experiment: Checking 5 students for school uniform.

Number of trials  $n = 5$  (5 students)

Success: Out of uniform

Probability of success  $p = 0.05$  (5%)

Random variable  $X$ : Number out of uniform.

The experiment can be considered as a sequence of 5 Bernoulli trials, and  $X$  has a binomial distribution.

$$\Pr(X = 0) = {}^5C_0 (0.05)^0 (0.95)^5 = 0.774$$

$$\Pr(X = 1) = {}^5C_1 (0.05)^1 (0.95)^4 = 0.204$$

$$\Pr(X = 2) = {}^5C_2 (0.05)^2 (0.95)^3 = 0.021$$

$$\Pr(X = 3) = {}^5C_3 (0.05)^3 (0.95)^2 = 0.001$$

$$\Pr(X = 4) = {}^5C_4 (0.05)^4 (0.95)^1 = 0.000$$

$$\Pr(X = 5) = {}^5C_5 (0.05)^5 (0.95)^0 = 0.000$$

These probabilities can be obtained by graphics calculator, 2nd DISTR 0: binompdf (5, 0.05) ENTER.

Binomial distribution of  $X$ :

$x$	0	1	2	3	4	5
$\Pr(X = x)$	0.774	0.204	0.021	0.001	0.000	0.000

$$\begin{aligned} \text{(b) } \Pr(\text{at most one}) &= \Pr(X \leq 1) = \Pr(X = 0) + \Pr(X = 1) \\ &= 0.774 + 0.204 = 0.978. \end{aligned}$$

This probability can be obtained by graphics calculator, 2nd DISTR A: binomcdf (5, 0.05, 1) ENTER. 0.977

The small discrepancy is due to rounding errors.

$$\text{(c) } \mu_x = np = 5 \times 0.05 = 0.25.$$

$$\text{Var}(X) = np(1-p) = 5 \times 0.05 \times 0.95 = 0.2375,$$

$$\therefore \sigma_x = \sqrt{0.2375} = 0.49.$$

**Example 2** In a science experiment, a student has a 40% chance of getting results. How many times the experiment should be performed if the student wishes to be 95% confident that (a) at least one experiment (b) at least two experiments, will be successful?

Assume the sequence of experiments forms a sequence of Bernoulli trials, and has a binomial distribution with  $p = 0.4$  and  $n$  is to be found.

$$\text{(a) } \Pr(X \geq 1) \geq 0.95, 1 - \Pr(X = 0) \geq 0.95,$$

$$\therefore \Pr(X = 0) \leq 0.05, {}^nC_0 (0.4)^0 (0.6)^n \leq 0.05,$$

$$\therefore 0.6^n \leq 0.05, n \log_e 0.6 \leq \log_e 0.05, n \geq \frac{\log_e 0.05}{\log_e 0.6} = 5.8645,$$

i.e.  $n \geq 6$ .

See part (b) for alternative methods in solving the inequality.

$$\text{(b) } \Pr(X \geq 2) \geq 0.95, 1 - \Pr(X = 0) - \Pr(X = 1) \geq 0.95,$$

$$\therefore \Pr(X = 0) + \Pr(X = 1) \leq 0.05,$$

$${}^nC_0 (0.4)^0 (0.6)^n + {}^nC_1 (0.4)^1 (0.6)^{n-1} \leq 0.05,$$

$$\therefore 0.6^n + 0.4n(0.6)^{n-1} \leq 0.05.$$

Graphics calculator: Enter equation  $y = 0.6^n + 0.4n(0.6)^{n-1}$ ,

2nd TBLSET, TblStart = 1,  $\Delta$ Tbl = 1, Indpnt: Auto Ask,

Depend: Auto Ask, 2<sup>nd</sup> TABLE.

Scroll down to where  $Y_1$  is less than 0.05, i.e.  $n \geq 10$ .

Alternatively, use graphics calculator to sketch

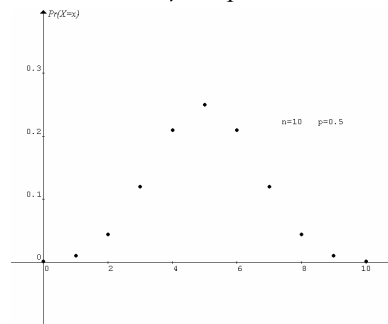
$$y = 0.6^n + 0.4n(0.6)^{n-1} \text{ and } y = 0.05.$$

Find the intersection at  $n = 9.82$ .

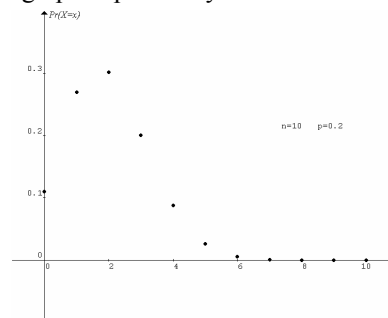
$$\text{For } 0.6^n + 0.4n(0.6)^{n-1} \leq 0.05, n \geq 10.$$

## Graphs of binomial distributions

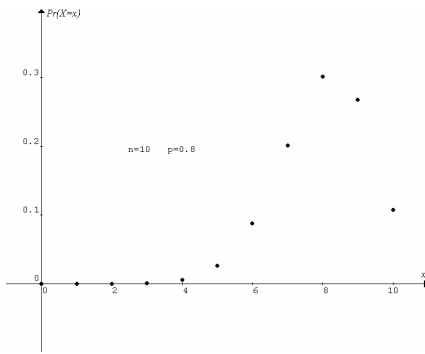
If  $p = 0.5$ , the graph of the distribution is bell-shaped and symmetrical about the mean  $\mu = np$ .



If  $p < 0.5$ , the graph is positively skewed.



If  $p > 0.5$ , the graph is negatively skewed.

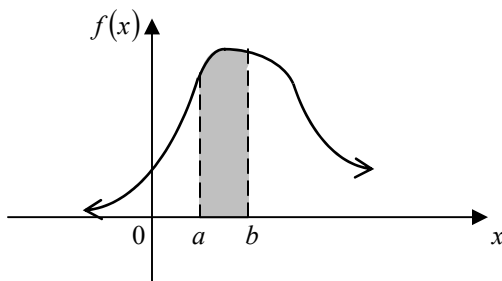


If  $n$  becomes large, all binomial distributions tend towards bell-shape irrespective of the value of  $p$ .

### Continuous random variables and probability density functions

In many experiments a random variable can take any real value, but its *exact* value in an attempt of the experiment cannot be determined, due to the limitations of measuring instrument and the person taking the measurement, e.g. measuring the height  $H$  of a randomly chosen person from a large number of people. It is more appropriate to consider a range of values (an interval) for the height rather than a single value. Hence the assigned probability for  $H = h$  cm is zero, i.e.  $\Pr(H = h) = 0$ , and  $0 \leq \Pr(a < H < b) \leq 1$ .

A random variable that fits the above descriptions is called a **continuous** random variable. The **probability distribution** of a continuous random variable is given by its **probability density function**  $f(x)$  such that  $\Pr(a < X < b) = \int_a^b f(x)dx$ , i.e. area under the curve  $f(x)$  in the interval  $(a, b)$



Notes:

- (1) For  $f(x)$  to be a probability density function of random variable  $X$ , it must satisfy the conditions:  $f(x) \geq 0$  for all real  $x$ , and  $\int_{-\infty}^{\infty} f(x)dx = 1$ .
- (2) Do not confuse probability density function  $f(x)$  with probability function  $p(x)$  discussed earlier.  $f(x)$  is for continuous random variable and the area under the graph of  $f(x)$  gives the probability.  $p(x)$  is for discrete random variable and direct substitution of a value of  $x$  gives the probability.

### Calculation of expected value, variance and standard deviation of a continuous random variable

For continuous random variable  $X$  with probability density function  $f(x)$ ,

$$\mu_X = E(X) = \int_{-\infty}^{\infty} xf(x)dx,$$

$$Var(X) = E((X - \mu_X)^2) = \int_{-\infty}^{\infty} (x - \mu_X)^2 f(x)dx \text{ or}$$

$$Var(X) = E(X^2) - (\mu_X)^2 = \int_{-\infty}^{\infty} x^2 f(x)dx - (\mu_X)^2,$$

$$\sigma_X = \sqrt{Var(X)}.$$

For many natural continuous probability distributions,  $\Pr(\mu - \sigma < X < \mu + \sigma) \approx 0.68$ , and  $\Pr(\mu - 2\sigma < X < \mu + 2\sigma) \approx 0.95$ .

### Calculation of median and mode of a continuous random variable

The median of continuous random variable  $X$  with probability density function  $f(x)$  is defined as the value  $m$  of  $X$  such that

$$\int_{-\infty}^m f(x)dx = 0.5. \text{ Hence } \int_m^{\infty} f(x)dx = 0.5.$$

The mode of continuous random variable  $X$  with probability density function  $f(x)$  is the value  $M$  of  $X$  such that

$f(M) \geq f(x)$  for all real  $x$ . The equal sign suggests that there may be more than one mode in the distribution. If  $f(x)$  is differentiable for all  $x$ , then it is possible in some cases to find the mode by differentiation (maximum of  $f(x)$ ).

If the probability density function is symmetrical about a mode  $M$ , then the median, the mean and  $M$  have the same value.

Example 1 Consider the probability density function

$$f(x) = \frac{2}{\pi} \sin^2 x \text{ for } 0 < x < \pi \text{ and } f(x) = 0 \text{ elsewhere.}$$

- (a) Show that  $f(x)$  satisfies the requirements of a probability density function.
- (b) Find  $\Pr(1 < X < 2)$ .
- (c) Find the mean, median and mode of  $X$ .
- (d) Find the variance of  $X$ .
- (e) Verify that  $\Pr(\mu - \sigma < X < \mu + \sigma) \approx 0.68$ .

- (a) Since  $\sin^2 x$  is always positive for all  $x \in (0, \pi)$ ,  $\therefore f(x) > 0$  for all  $x \in (0, \pi)$ . Elsewhere,  $f(x) = 0$ .  $\therefore f(x) \geq 0$  for all real  $x$ .

Use graphics calculator to graph  $y = \frac{2}{\pi} \sin^2 x$ ,  $0 \leq x \leq \pi$ .

2nd CALC 7:  $\int f(x)dx$ . Lower limit  $x = 0$ . Upper limit  $x = \pi$ .

$$\int_0^{\pi} \frac{2}{\pi} \sin^2 x dx = 1. \therefore \int_{-\infty}^{\infty} f(x)dx = 1.$$

$$(b) \Pr(1 < X < 2) = \int_1^2 \frac{2}{\pi} \sin^2 x dx = 0.5835.$$

(c) Since  $f(x) = \frac{2}{\pi} \sin^2 x$ ,  $0 \leq x \leq \pi$  is symmetrical about  $\frac{\pi}{2}$ ,  
the mean = median = mode of  $X = \frac{\pi}{2}$ .

$$(d) \text{Var}(X) = E(X^2) - (\mu_X)^2 = \int_{-\infty}^{\infty} x^2 f(x) dx - (\mu_X)^2$$

$$\text{Var}(X) = \int_0^{\pi} x^2 \left( \frac{2}{\pi} \sin^2 x \right) dx - \left( \frac{\pi}{2} \right)^2 = 2.7899 - 2.4674 = 0.3225$$

$$\sigma_X = \sqrt{0.3225} = 0.5679.$$

(e)

$$\mu - \sigma = \frac{\pi}{2} - 0.5679 = 1.0029, \quad \mu + \sigma = \frac{\pi}{2} + 0.5679 = 2.1387,$$

$$\Pr(1.0029 < X < 2.1387) = \int_{1.0029}^{2.1387} \frac{2}{\pi} \sin^2 x dx = 0.6502$$

Example 2 Consider the probability density function

$f(x) = 30x^4(1-x)$  for  $0 < x < 1$  and  $f(x) = 0$  elsewhere.

(a) Show that  $f(x)$  satisfies the requirements of a probability density function.

(b) Find  $\Pr(X < 0.5)$ .

(c) Find the mean of  $X$ .

(d) Find the variance of  $X$ .

(e) Use calculus to find the median of  $X$ .

(f) Find the mode of  $X$ .

(a) For  $0 < x < 1$ ,  $x^4 > 0$  and  $(1-x) > 0$ ,  $\therefore f(x) > 0$ .

Elsewhere,  $f(x) = 0$ .  $\therefore f(x) \geq 0$  for all real  $x$ .

Use graphics calculator to graph  $y = 30x^4(1-x)$ ,  $0 \leq x \leq 1$ .

2nd CALC 7:  $\int f(x) dx$ . Lower limit  $x = 0$ . Upper limit  $x = 1$ .

$$\int_0^1 30x^4(1-x) dx = 1. \quad \therefore \int_{-\infty}^{\infty} f(x) dx = 1.$$

$$\text{Or by calculus. } \int_0^1 30x^4(1-x) dx = \int_0^1 30(x^4 - x^5) dx$$

$$= \left[ 30 \left( \frac{x^5}{5} - \frac{x^6}{6} \right) \right]_0^1 = 30 \times \frac{1}{30} = 1.$$

$$(b) \Pr(X < 0.5) = \int_0^{0.5} 30x^4(1-x) dx = \int_0^{0.5} 30(x^4 - x^5) dx$$

$$= \left[ 30 \left( \frac{x^5}{5} - \frac{x^6}{6} \right) \right]_0^{0.5} = 30 \left( \frac{0.5^5}{5} - \frac{0.5^6}{6} \right) = 0.1094$$

$$(c) E(X) = \int_0^1 x 30(x^4 - x^5) dx = \int_0^1 30(x^5 - x^6) dx$$

$$= \left[ 30 \left( \frac{x^6}{6} - \frac{x^7}{7} \right) \right]_0^1 = 30 \times \frac{1}{42} = 0.7143$$

$$(d) \text{Var}(X) = \int_0^1 x^2 30(x^4 - x^5) dx - 0.7143^2$$

$$= \int_0^1 30(x^6 - x^7) dx - 0.7143^2$$

$$= \left[ 30 \left( \frac{x^7}{7} - \frac{x^8}{8} \right) \right]_0^1 - 0.7143^2 = 30 \times \frac{1}{56} - 0.7143^2 = 0.0255.$$

(e) Let  $0 < m < 1$  be the median of  $X$ .  $\therefore \Pr(X < m) = 0.5$ .

$$\therefore \int_0^m 30x^4(1-x) dx = \int_0^m 30(x^4 - x^5) dx = 0.5.$$

$$\therefore \left[ 30 \left( \frac{x^5}{5} - \frac{x^6}{6} \right) \right]_0^m = 30 \left( \frac{m^5}{5} - \frac{m^6}{6} \right) = 0.5, \quad \therefore 12m^5 - 10m^6 = 1.$$

Use graphics calculator to find  $m = 0.7356$ .

(f)  $f(x) = 30x^4(1-x)$  for  $0 < x < 1$ .

To find the maximum of  $f(x)$ , let  $f'(x) = 30(4x^3 - 5x^4) = 0$ ,

$\therefore x^3(4 - 5x) = 0$ ,  $\therefore x = 0.8$ , hence the mode is 0.8.

### Normal distributions – examples of continuous probability distributions

A normal distribution  $N(\mu, \sigma^2)$  has the following probability

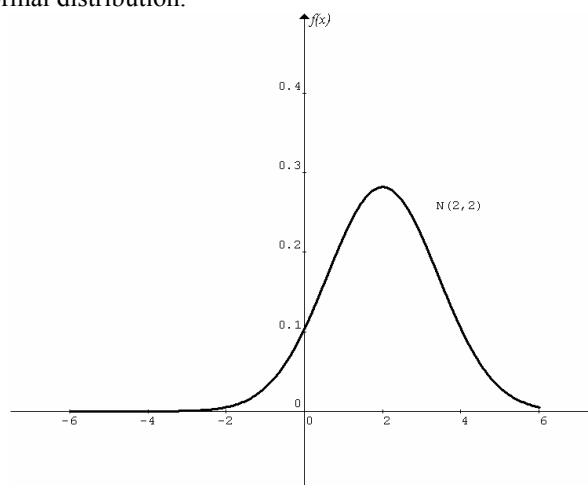
density function:  $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ , where the mean

$\mu \in R$  and the standard deviation  $\sigma > 0$  are the parameters of the normal distribution.

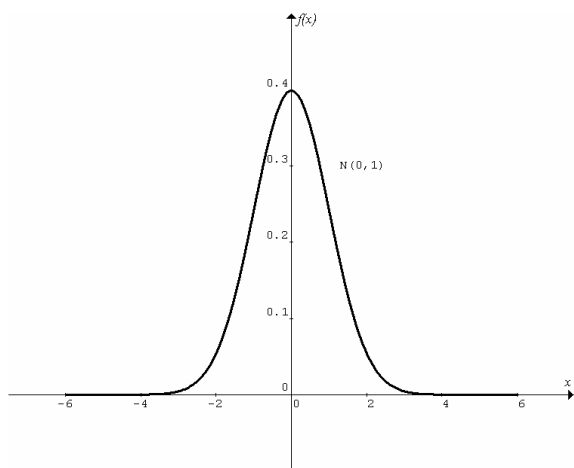
Note:  $N(\mu, \sigma^2)$  is the notation used in the study design (2006-2009),  $\sigma^2$  is the variance. You may find in some textbooks the notation  $N(\mu, \sigma)$  is used.

A normal distribution is bell-shaped and symmetrical about its mean  $\mu$ . At  $x = \mu$ ,  $f(\mu) = \frac{1}{\sigma\sqrt{2\pi}}$  is the maximum,  $\therefore \mu$  is also

the mode. Furthermore,  $\Pr(X < \mu) = 0.5$  and  $\Pr(X > \mu) = 0.5$  due to the symmetry about the mean.  $\therefore \mu$  is also the median of a normal distribution.

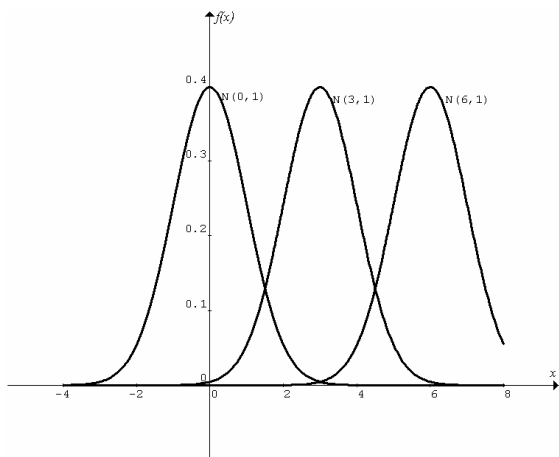


If  $\mu = 0$  and  $\sigma = 1$ , then  $N(0,1)$  is called standard normal distribution. Its random variable is usually denoted as  $Z$ .

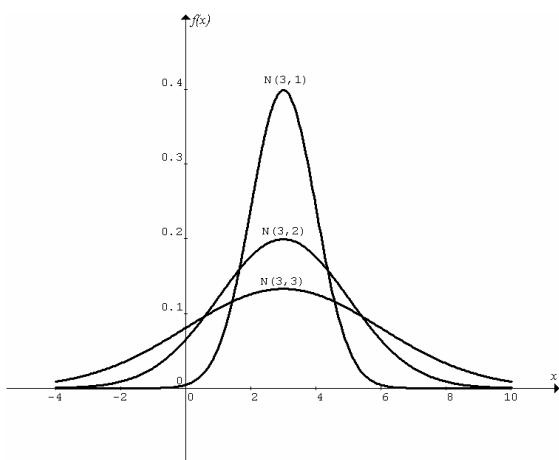


### Effects of changing $\mu$ and $\sigma$ on the graph of $N(\mu, \sigma^2)$

Increasing (decreasing) the value of  $\mu$  causes the graph to translate to the right (left).



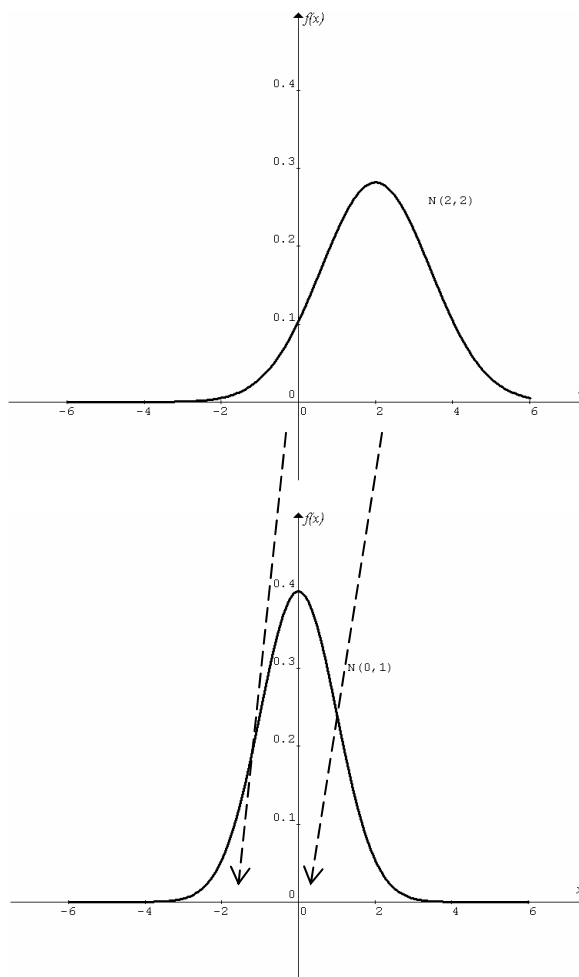
Increasing (decreasing) the value of  $\sigma$  causes the graph to spread out (contract) and lowers (raises) the maximum value of  $f(x)$ .



### Transforming a normal distribution to the standard normal distribution

Any value of the random variable  $X$  with the normal distribution  $N(\mu, \sigma^2)$  can be matched up with a value of the random variable  $Z$  with the standard normal distribution  $N(0,1)$  by the relationship  $Z = \frac{X - \mu}{\sigma}$ , such that  $\Pr(X < x) = \Pr\left(Z < \frac{x - \mu}{\sigma}\right)$ .

Example 1  $\mu = 2, \sigma^2 = 2, \therefore \sigma = \sqrt{2}$ .



For  $X = 2, Z = \frac{2-2}{\sqrt{2}} = 0 \therefore \Pr(X < 2) = \Pr(Z < 0) = 0.5$ .

For  $X = 0, Z = \frac{0-2}{\sqrt{2}} = -1.414$ .

$\therefore \Pr(X < 0) = \Pr(Z < -1.414) = 0.079$



Example 1 (i) Write down the probability density function  $f(x)$  for  $N(2.5,1.5)$ . (ii) Determine  $\Pr(X < 3)$  by (a) finding the area under the graph of  $f(x)$ , (b)  $normalcdf(-,-,-,-)$ , and (c) changing to the standard normal and then  $normalcdf(-,-)$ . (iii) Determine  $\Pr(1 < X < 3)$  by repeating (a), (b) and (c). (iv) Show that  $\Pr(\mu - 2\sigma < X < \mu + 2\sigma) \approx 0.95$ .

(i)  $f(x) = \frac{1}{1.5\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-2.5}{1.5}\right)^2}$ .

(ii)

(a)  $\Pr(X < 3) = \int_{-\infty}^3 f(x)dx = 0.63$ . Note: the writer used  $\mu - 5\sigma = -5$  as the lower limit to evaluate the definite integral.

(b)  $\Pr(X < 3) = normalcdf(-E99,3,2.5,1.5) = 0.63$ .

(c)  $Z = \frac{3-2.5}{1.5} = \frac{1}{3}$ ,

$\Pr(X < 3) = \Pr\left(Z < \frac{1}{3}\right) = normalcdf(-E99,1/3) = 0.63$ .

(iii)

(a)  $\Pr(1 < X < 3) = \int_1^3 f(x)dx = 0.47$ .

(b)  $\Pr(1 < X < 3) = normalcdf(1,3,2.5,1.5) = 0.47$ .

(c)  $Z = \frac{1-2.5}{1.5} = -1$ ,

$\Pr(1 < X < 3) = \Pr\left(-1 < Z < \frac{1}{3}\right) = normalcdf(-1,1/3) = 0.47$ .

(iv)  $\mu - 2\sigma = 2.5 - 2 \times 1.5 = -0.5$ ,  $\mu + 2\sigma = 2.5 + 2 \times 1.5 = 5.5$   
 $\Pr(-0.5 < X < 5.5) = normalcdf(-0.5,5.5,2.5,1.5) \approx 0.95$ .

Example 2  $X$  has a normally distribution with  $\mu = 80$  and  $\sigma^2 = 25$ . Find (a)  $x$  such that  $\Pr(X < x) = 0.8$ , and (b)  $c > 0$  such that  $\Pr(\mu - c < X < \mu + c) = 0.8$ .

$\sigma = 5$ .

(a)  $\Pr(X < x) = 0.8$ ,  $x = invNorm(0.8,80,5) = 84.2$ .

(b)  $\Pr(\mu - c < X < \mu + c) = 0.8$ ,  $\therefore \Pr(\mu - c < X < \mu) = 0.4$ ,  
 $\therefore \Pr(X < \mu - c) = 0.1$ ,  $\mu - c = invNorm(0.1,80,5) = 73.6$ ,  
 $\therefore c = \mu - 73.6 = 80 - 73.6 = 6.4$

Example 3  $X$  has a normally distribution with  $\mu = 80$  and  $\sigma^2 = 25$ . Without using graphics calculator, find (a)  $c > 0$  such that  $\Pr(\mu - c < X < \mu + c) = 0.95$ , and (b)  $x$  such that  $\Pr(X < x) = 0.975$ .

(a) Since  $\Pr(\mu - 2\sigma < X < \mu + 2\sigma) = 0.95$ ,  $\therefore c = 2\sigma = 10$ .

(b) Outside the interval  $\mu - 2\sigma < X < \mu + 2\sigma$ ,  $prob = 0.05$ ,  
 $\therefore \Pr(X > \mu + 2\sigma) = 0.025$ ,  
 $\therefore \Pr(X < \mu + 2\sigma) = 1 - 0.025 = 0.975$ , hence  $x = \mu + 2\sigma = 90$ .

Example 4 The height of a population is normally distributed with  $\mu = 175$  cm. If 2% of the population is over 190 cm, find the standard deviation of the random variable  $X$ , the height in cm.

Given  $\Pr(X > 190) = 0.02$ ,  $\therefore \Pr(X < 190) = 1 - 0.02 = 0.98$ ,

$\therefore \Pr\left(Z < \frac{190-175}{\sigma}\right) = 0.98$ .

Hence  $\frac{15}{\sigma} = invNorm(0.98) = 2.0537$ ,  $\therefore \sigma = 7.3$ .

Note: It is not necessary to enter 0 and 1, by default  $invNorm(\_)$  is the inverse of the standard normal distribution.

Example 5 The body-weight of an individual in a population is normally distributed. If 5% of the population is over 80 kg and 3% is less than 20 kg, find the mean and standard deviation of the weight of the population.

Given  $\Pr(X > 80) = 0.05$  and  $\Pr(X < 20) = 0.03$ ,

$\therefore \Pr(X < 80) = 1 - 0.05 = 0.95$ ,

$\therefore \Pr\left(Z < \frac{80-\mu}{\sigma}\right) = 0.95$  and  $\Pr\left(Z < \frac{20-\mu}{\sigma}\right) = 0.03$ .

Hence  $\frac{80-\mu}{\sigma} = invNorm(0.95)$  and  $\frac{20-\mu}{\sigma} = invNorm(0.03)$ .

$\therefore \frac{80-\mu}{\sigma} = 1.64485$  and  $\frac{20-\mu}{\sigma} = -1.88079$ .

Solve simultaneously,

$80 - \mu = 1.64485\sigma$  .....(1)

$20 - \mu = -1.88079\sigma$  .....(2)

(1) - (2),  $60 = 3.52564\sigma$ ,  $\therefore \sigma = 17$  and  $\mu = 52$

Example 6 The distance in metres that a javelin is thrown by a competitor follows a normal distribution with a mean of 80.80 and a standard deviation of 4.50. (a) Find the probability that a javelin is thrown more than 85.00 m. (b) In five throws what is the probability that the competitor throws more than 85.00 m in more than 2 occasions? (c) Find the expected number of throws by the competitor exceeding 85.00 m in 5 throws.

(a)  $N(80.80,4.50^2)$ ,

$\Pr(X > 85.00) = normalcdf(85.00,E99,80.80,4.50) = 0.175$ .

(b)  $Bi(5,0.175)$ ,

$\Pr(X > 2) = 1 - \Pr(X \leq 2) = 1 - binomcdf(5,0.175,2) = 1 - 0.96 = 0.04$ .

(c)  $Bi(5,0.175)$ ,  $\mu = E(X) = np = 5 \times 0.175 = 0.88$ .

## Conditional probability

The probability of an event  $A$  may or may not depend on the knowledge of the occurrence of another event  $B$ .

The probability of  $A$  given the knowledge of  $B$  occurring is denoted by  $\Pr(A|B)$ . This is called **conditional probability**.

If  $\Pr(A|B) = \Pr(A)$ , then events  $A$  and  $B$  are **independent**.

If  $\Pr(A|B) \neq \Pr(A)$ , then events  $A$  and  $B$  are dependent.

Conditional probability is defined as  $\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$ .

Hence  $\Pr(A \cap B) = \Pr(A|B) \times \Pr(B)$ . This is known as **the multiplication rule**.

If  $A$  and  $B$  are independent, the multiplication rule can be simplified to  $\Pr(A \cap B) = \Pr(A) \times \Pr(B)$ .

Conversely, if  $\Pr(A \cap B) = \Pr(A) \times \Pr(B)$ , then  $A$  and  $B$  are independent.

Note: Do not confuse independent events with **mutually exclusive** events.

If  $\Pr(A \cap B) = 0$ , then  $A$  and  $B$  are mutually exclusive events, and **the addition rule**  $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$  becomes  $\Pr(A \cup B) = \Pr(A) + \Pr(B)$ .

A **Karnaugh table** is a useful tool in calculating conditional probabilities.

	$A$	$A'$	
$B$	$\Pr(A \cap B)$	$\Pr(A' \cap B)$	$\Pr(B)$
$B'$	$\Pr(A \cap B')$	$\Pr(A' \cap B')$	$\Pr(B')$
	$\Pr(A)$	$\Pr(A')$	1

Example 1 A fair coin is tossed 3 times. Let  $A$  be the event that more than 2 heads occurring, and  $B$  the event that more than 1 head occurring. Find  $\Pr(A'|B)$  and  $\Pr(B|A')$ .

$$\Pr(A) = \Pr(X = 3) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}.$$

$$\Pr(B) = \Pr(X > 1) = \Pr(X = 2) + \Pr(X = 3) = \frac{3}{8} + \frac{1}{8} = \frac{1}{2}.$$

$$\Pr(A \cap B) = \Pr(X = 3) = \frac{1}{8}.$$

Set up the Karnaugh table:

	$A$	$A'$	
$B$	1/8	3/8	1/2
$B'$	0	1/2	1/2
	1/8	7/8	1

$$\Pr(A'|B) = \frac{\Pr(A' \cap B)}{\Pr(B)} = \frac{3/8}{1/2} = \frac{3}{4}.$$

$$\Pr(B|A') = \frac{\Pr(B \cap A')}{\Pr(A')} = \frac{3/8}{7/8} = \frac{3}{7}.$$

Example 2 The weight in grams of a potato in a crate is normally distributed with  $\mu = 220$  and  $\sigma = 10$ . Find (a) the probability that a potato weighs more than 215 grams given that it is not less than 210 grams, (b) the probability that a potato weighs less than 215 grams given that it is not less than 210 grams.

$$\begin{aligned} \text{(a)} \quad \Pr\left((X > 215) | (X < 210)'\right) &= \Pr((X > 215) | (X \geq 210)) \\ &= \frac{\Pr((X > 215) \cap (X \geq 210))}{\Pr(X \geq 210)} \\ &= \frac{\Pr(X > 215)}{\Pr(X \geq 210)} = \frac{\text{normalcdf}(215, E99, 220, 10)}{\text{normalcdf}(210, E99, 220, 10)} = 0.82. \end{aligned}$$

$$\begin{aligned} \text{(b)} \quad \Pr\left((X < 215) | (X < 210)'\right) &= \Pr((X < 215) | (X \geq 210)) \\ &= \frac{\Pr((X < 215) \cap (X \geq 210))}{\Pr(X \geq 210)} \\ &= \frac{\Pr(210 \leq X < 215)}{\Pr(X \geq 210)} = \frac{\text{normalcdf}(210, 215, 220, 10)}{\text{normalcdf}(210, E99, 220, 10)} = 0.18. \end{aligned}$$

Note: A quicker way,

$$\begin{aligned} \Pr\left((X < 215) | (X < 210)'\right) &= 1 - \Pr\left((X > 215) | (X < 210)'\right) \\ &= 1 - 0.82 = 0.18. \end{aligned}$$