LECTURE 1 OF 6

TOPIC : 10.0 SPECIAL PROBABILITY DISTRIBUTION

SUBTOPIC : 10.1 Binomial and Poisson Distributions

LEARNING OUTCOMES :

At the end of lesson students should be able to:

(a) identify the binomial distribution B(n, p)

(b) find the mean and variance of binomial distribution

(c) use the binomial distribution to solve related problem

The Binomial Distribution

Binomial distribution is used to model situations where there are just two possible outcomes, success and failure. Together with this property, the following conditions also need to be fulfilled.

(a) there must be a fixed number of trials (n trials),

(b) the trials are independent of each other, and

(c) the probability of success (p) must be the same for each trial.

Example

Determine whether the following situations can be modelled by a binomial distribution.

(a) The number of heads obtained when unbiased coin is tossed 10 times.

Answer : *This is a binomial situation. The 10 trials are independent of each other and the probability of obtaining a head is 0.5 for each trial.*

(b) The number of lettuce seeds that germinate from a pack of 30 seeds, given that the probability that a lettuce seed germinates is 0.8

Answer : *This a binomial situation, the 30 trials are independent and the probability that a seed germinates is constant at 0.8 for each trial.*

(c) The number of black pens obtained when four pens are picked at random one at a time, from a box containing 5 black pens and 15 red pens.

Answer : *This is not a binomial situation. The events are not independent since the outcome of the second trial depends on the outcome of the first.*

Definition

If X is a random variable of a binomial distribution, then the probability of x success in n trials is given by the following probability distribution function

$$P(X = x) = {}^{n}C_{x}\, p^{x} q^{n-x} \quad , \quad x = 0, 1, 2, \ldots, n$$

where

n	:	the number of trial
x	:	the number of success
p	:	the probability for a success
q	:	the probability for a failure ($q = 1 - p$)

We write $X \sim B(n , p)$ where n and p are known as the parameter of the binomial distribution.

Note : $${}^{n}C_{x} = \frac{n!}{x!(n-x)!}$$

Example 1

If $X \sim B(10 , 0.9)$, find

(a) $P(X = 6)$

(b) $P(X > 8)$

(c) $P(X \geq 2)$

Example 2

A four-sided dice is thrown five times. If X denotes the number of four obtained, describe the distribution of X and find $\mathrm{P}(X = 3)$.

Example 3

The probability of a Group B blood donor is 0.09. Calculate the least number of donors needed to ensure that the probability of obtaining at least one group B donor is greater than 0.95.

Example 4

In a biology examination 45 % of students passed. If a class of 20 students took the examination, find the probability that

(a) exactly 15 students passed.

(b) at least 2 students passed.

(c) more than 16 student passed.

(d) not more than 1 students passed.

Note :

As the number of trials (n) gets larger, the calculation becomes tedious. Therefore, binomial table is used instead. The tabulated value is for the probability in the form of $P(X \geq x)$ where X is the number of successes. The table below shows part of the probability values that will be used in this topic.

Binomial Table

n	**r** \ **p**	**0.10**	**…**	**…**	**0.35**	**…**	**0.50**
20	0	1.0000	…	…	1.0000	…	1.0000
	1	0.8784	…	…	0.9998	…	1.0000
	2	0.6083	…	…	0.9979	…	1.0000
	3	0.3231	…	…	0.9879	…	0.9998
	4	0.1330	…	…	0.9556	…	0.9987
	5	0.0432	…	…	0.8818	…	0.9941
	6	0.0113	…	…	0.7546	…	**0.9793**
	7	0.0024	…	…	**0.5834**	…	0.9423
	8	0.0004	…	…	0.3990	…	0.8684

For example,

for $n = 20$ and $p = 0.35$, $P(X \geq 7) = 0.5834$

for $n = 20$ and $p = 0.50$, $P(X \geq 6) = 0.9793$

Example 5

If $X \sim B(20, 0.35)$. By using the binomial table, find

(a) $P(X > 5)$

(b) $P(2 < X < 7)$

(c) $P(X < 4)$

Example 6

X is a random variable such that $X \sim B(5, 0.3)$. By using the binomial table, find

(a) $P(X \geq 3)$

(b) $P(X > 3)$

(c) $P(X \leq 3)$

(d) $P(X < 3)$

(e) $P(X = 3)$

The Expected Value (Mean) and Variance of The Binomial Distribution

Definition

If $X \sim B(n, p)$, then the expected value or mean of X written as $E(X) = np$, and the variance of X written as $Var(X) = npq$.

$$\boxed{\begin{array}{c} E(X) = np \\ \\ Var(X) = npq \end{array}}$$

Example 7

Let $X \sim B(n, 0.22)$ and $E(X) = 11$. Find

(a) the value of n

(b) $Var(X)$

(c) $P(X = 4)$

Example 8

A coin is tossed 40 times. Find the mean and the standard deviation of the number of heads appeared .

Example 9

All the pupils in a particular school, 45 % go to school by bus. In a random sample of 30 pupils, calculate

(a) the mean number of pupils that go to school by bus

(b) the probability that not more than 10 pupils go to school by bus.

Example 10

At the local swimming club, the expected number of members that come can swim a mile is 4.5 with variance is 3.15. Find the probability that at least three members can swim a mile.

Example 11

For a particular strain of seeds, only one out of five germinates. If a farmer plants 20 of these seeds, what is the probability that

(a) exactly five germinate

(b) more than five germinate

(c) not less than five germinate

Example 12

In a local election 40 % of voters support the Future Party. A random sample of 15 voters is taken. What is the probability that it contains :

(a) no voters of the Future Party

(b) at least two voters for the Future Party

(c) less than six voters for the Future Party

LECTURE 2 OF 6

TOPIC : 10.0 SPECIAL PROBABILITY DISTRIBUTIONS

SUBTOPIC : 10.1 Binomial and Poisson Distributions

LEARNING OUTCOMES :

At the end of lesson students should be able to:

(a) identify the Poisson distribution, $P_o(\lambda)$

(b) identify the mean and variance of Poison distribution.

CONTENTS :

The Poisson Distribution

If we are interested in the number of occurrences that take place in an interval (time/volume) then we use the Poisson Distribution. Situations to be modeled by Poisson distribution are,

(a) an experiment consists of counting the number of times a certain event occurs.

(b) the probability that an event occurs is the same for each interval.

(c) events occur independently

The variable X is the number of occurrences in the given interval is written as $X \sim P_o(\lambda)$ where λ is the parameter of the Poisson Distribution and $\lambda > 0$. When the Poisson Distribution is appropriate, the probability of exactly X occurs is given by the formula

$$\boxed{P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}}$$

Example 1

If $X \sim P_o(6)$, find the probability by using the Poisson formula

(a) $P(X = 3)$

(b) $P(X \geq 2)$

(c) $P(X < 2)$

(d) $P(3 < X \leq 5)$

Example 2

An average of three cars arrive at a highway tollgate every minute. If this rate is approximately Poisson distribution what is the probability that exactly,

(a) five cars will arrive in one minute

(b) seven cars will arrive in five minutes

Example 3

Emergency calls to an ambulance service are received at random times, at an average of 2 per hour. Calculate the probability that, in a randomly chosen one hour period,

(a) no emergency calls are received

(b) exactly one call is received in the first half hour and exactly one call is received in the second half.

Finding Probability Using The Poisson Table

Sometimes we need to calculate the probability such as P(X < 10). In this case, we have to calculate P(X = 0) + P(X = 1) + P(X = 2) + … + P(X = 9). This procedure is time consuming thus, it is easier to use the Poisson Distribution Table.

Example 4

If $X \sim P_o(5.5)$, find :

(a) $P(X \geq 5)$

(b) $P(X \leq 5)$

(c) $P(X = 5)$

(d) $P(3 \leq X < 8)$

Example 5

The number of breakdowns in a particular machine occurs at a rate of 2.5 per month. Assuming that the number of breakdowns follows Poisson distribution, find the probability that,

(a) more than three breakdowns occur in a particular month.

(b) less than ten breakdowns occur in three months period.

(c) exactly three breakdowns occur in two months.

Example 6

The Environmental Department is researching the pollution in the Juru river. Ten samples of 200 ml each of the water are taken and an average of 1.3 of foreign particles are found. Find the probability that,

(a) more than 3 particles are found in 1 litre of the Juru river water

(b) less than 4 particles are found in 0.5 litre of the Juru river water.

Mean and Variance of Poisson Distribution.

If X is a random variable following the Poisson distribution, $X \sim P_o(\lambda)$, then, the mean and variance of X is $E(X) = \lambda$ and $Var(X) = \lambda$.

Example 7

If $X \sim P_0(1.8)$, find the mean, $E(X)$ and the variance, $Var(X)$.

Example 8

Given that X is a Poisson distribution with mean, λ. If $P(X = 2) = 2P(X = 0)$, find the value of λ. Hence, calculate $P(X = 2)$.

LECTURE 3 OF 6

TOPIC : 10.0 SPECIAL PROBABILITY DISTRIBUTIONS

SUBTOPIC : 10.1 Binomial and Poisson Distributions

LEARNING OUTCOMES :

At the end of lesson students should be able to:

(c) use the Poisson distribution to solve related problem

CONTENTS

Example 1

An average number of books sold at a bookshop per week is 10, find

(a) the probability that 5 books are sold at that bookshop per week

(b) the probability that less than 4 books are sold

(c) the probability that more than 7 books are sold in 2 weeks

Example 2

A large number of 10 ml/samples are collected from a lake. The mean number of bacteria in a 10 ml/sample of liquid is 5. Find the probability that a sample taken has

(a) no bacteria

(b) one bacterium

(c) more than three bacteria

Example 3

A boy who fishes regularly in a lake in Petaling Jaya catches an average of 2.4 fish per hour. Assuming the number of fish he catches follows a Poisson distribution, find the probability that he catches

(a) two or more fish in half an hour.

(b) between 4 to 6 fish (inclusive) in 90 minutes.

Example 4

Customers enter an antique shop independently of one another and at random intervals of time at an average rate of four per hour throughout the five days of a week on which the shop is open. The owner has a coffee-break of fifteen minutes each morning; if one or more customers arrive during this period then his coffee goes cold, otherwise he drinks it while it is hot.

Let X be the random variables denoting the number of customers arriving during a Monday coffee – break , and let Y be the random variables denoting the number of days during a week on which the owner's coffee goes cold. Assuming that X has a Poisson distribution, determine (correct to 3 significant figures)

a) $P(X = 0)$

b) $P(X \geq 2)$

c) $E(Y)$

d) $P(Y = 2)$

LECTURE 4 OF 6

TOPIC : 10.0 SPECIAL PROBABILITY DISTRIBUTIONS

SUBTOPIC : 10.2 Normal Distribution

LEARNING OUTCOMES :

At the end of lesson students should be able to:

(a) identify the normal distribution $N\left(\mu,\sigma^2\right)$

(b) standardize the normal random variable

.

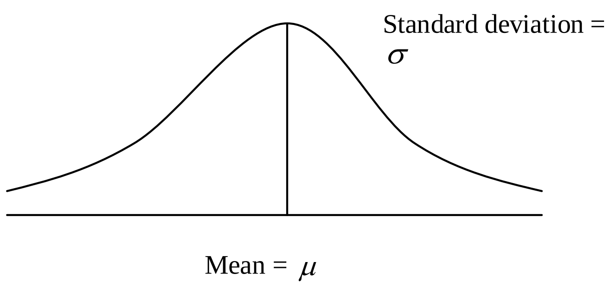
CONTENTS

The Normal Distribution

The normal distribution is one of the most important distribution in statistics. If a continuous random variable *X* has a normal distribution with mean μ and standard deviation σ, then we write,

$$\boxed{X \sim N(\mu,\sigma^2)}$$

The normal probability distribution or normal curve is given by a bell-shaped (symmetric) curve. Such a curve is shown in figure below. It has a mean of μ and a standard deviation of σ. A continuous random variable *X* that has a normal distribution is called a normal random variable.



A normal probability distribution, when plotted, gives a bell-shaped curve such that,

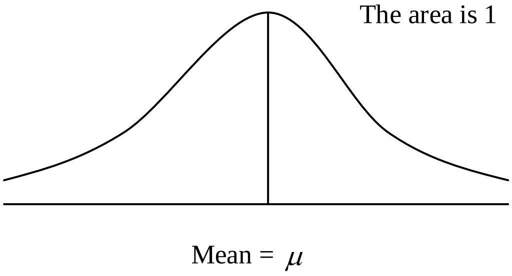
(a) The total area under the curve is 1.

(b) The curve is symmetric about the mean

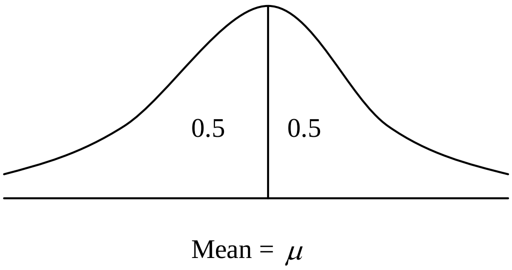
(c) The two tails of the curve extend indefinitely

A normal distribution possesses the following three properties:

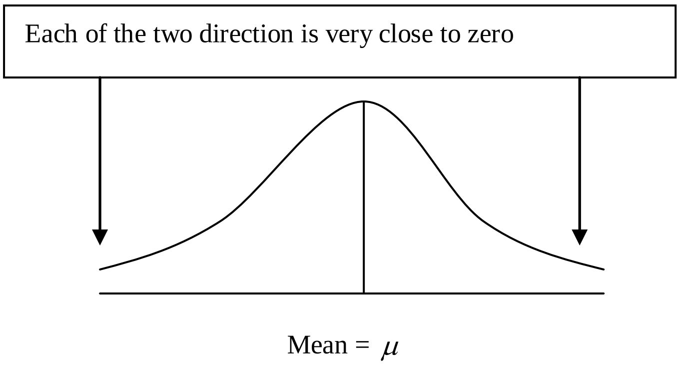
1. The total area under a normal curve is 1 or 100% as shown below,



2. A normal curve is symmetric about the mean, Consequently, $\frac{1}{2}$ of the total area under a normal curve lies on the left side of the mean and $\frac{1}{2}$ lies on the right side of the mean.

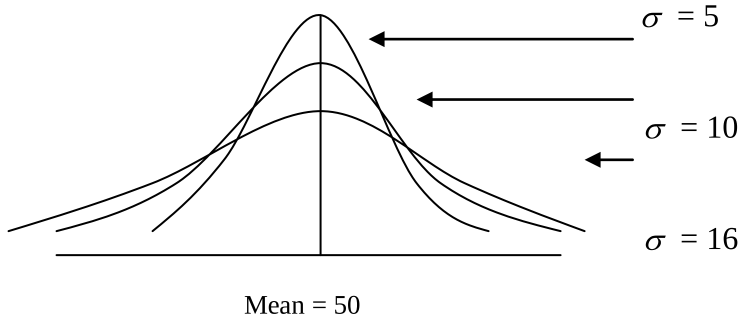


3. The tail of a normal distribution curve extend indefinitely in both direction without touching or crossing the horizontal axis. Although a normal curve never meets the horizontal axis it becomes so close to this axis that the area under the curve beyond these point in both direction can be taken as virtually zero. These areas are shown below,



The mean μ and the standard deviation σ are the *parameters* of the normal distribution. Given the values of these two parameters, we can find the area under a normal curve for any interval.

There are a family of normal distribution curves. Each different set of values of μ and σ gives a different normal curve. The value of μ determines the center of a normal distribution on the horizontal axis and the value of σ gives the spread of the normal distribution curve. As an illustration, the three normal distribution curves below have the same mean but different standard deviations.



The Standard Normal Distribution

The standard normal distribution is a special case of the normal distribution. For the standard normal distribution , the value of mean is equal to zero and the value of standard deviation is equal to 1. The random variable that possesses the standard normal distribution is denoted by *Z* and is called the the *Z* values or *Z* scores. They are also called standard units or standard scores.

The Standardizing A Normal Distribution

The procedure to convert the given normal distribution to the standard normal distribution is called standardizing a normal distribution. The units of a normal distribution are denoted by *x*. The standard normal distribution are denoted by *z*. For a normal random variable *x*, a particular value of *x* can be converted to a *z* value by using the formula,

$$\boxed{Z = \frac{X - \mu}{\sigma}}$$

where μ and σ are the mean and standard deviation of the normal distribution of x.Thus, to find the z value for any x value, we calculate the difference between the given x value and the mean μ and divide this difference by the standard deviation σ. If the value of x is equal to μ then z value is equal to zero. *Note that we always round z values to two decimal places.*

Example 1

Find the area under the standard normal curve between $z = 0$ and $z = 1.95$

Example 2

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

(a) $P(Z > 1.2)$

(b) $P(-2.0 < Z < 2.0)$

(c) $P(-1.2 < Z < 1.0)$

Example 3

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

(a) $P(Z < a) = 0.9$

(b) $P(Z > a) = 0.25$

Example 4

Let X be a continuous random variable that has a normal distribution with a mean of 50 and a standard deviation of 10. Convert the following x values to z values, if the given normal distribution $\mu = 50$ and $\sigma = 10$

(a) $X = 55$

(b) $X = 35$

Example 5

Let X be a continuous random variable that has a normal distribution with a mean of 25 and a standard deviation of 4. Find the area, if the given normal distribution $\mu = 25$ and $\sigma = 4$.

(a) between $X = 25$ and $X = 32$

(b) between $X = 18$ and $X = 34$

Example 6

Let X be a continuous random variable that has a normal distribution with a mean of 40 and a standard deviation of 5. Find the following probability for this normal distribution.

(a) $P(X > 55)$

(b) $P(X < 49)$

Example 7

The random variable X has a normal distribution with mean μ and variance σ^2. Given that $P(X > 25.512) = 0.3594$ and $P(X < 16.104) = 0.0301$. Find μ and σ^2.

LECTURE 5 OF 6

TOPIC : 10.0 SPECIAL PROBABILITY DISTRIBUTIONS

SUBTOPIC : 10.2 Normal Distribution

LEARNING OUTCOMES :

At the end of lesson students should be able to:

(c) determine the mean and variance of normal distribution function

(d) use the normal distribution to solve related problems.

CONTENT

Example 1

A random variable has a normal distribution with standard deviation $\sigma = 10$. Find its mean if the probability that it will take on a value more than 77.5 is 0.1736.

Example 2

$X \sim N(\mu, \sigma^2)$. If $P(X > 12) = 0.3$ and $P(X < 6) = 0.4$, find the value of μ and σ.

Example 3

In a company, the wages of a certain grade of staff are normally distributed with a standard deviation of RM 400. If 20.05% of staff earn less than RM 300 a week,

(a) What is the average wage?

(b) What the percentage of staff earns more than RM500 a week?

LECTURE 6 OF 6

TOPIC : 10.0 SPECIAL DISTRIBUTION FUNCTIONS

SUBTOPIC : 10.2 Continuous Probability Distribution

LEARNING OUTCOMES :

At the end of the lesson students are able to:

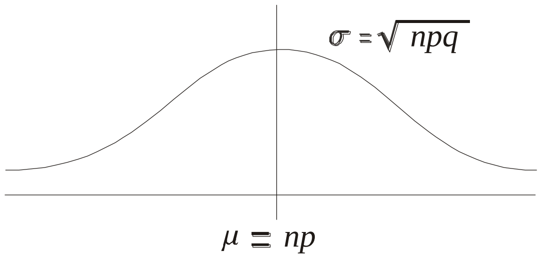
(a) Use the normal distribution to approximate the binomial distribution

CONTENT

The Normal Approximation to the Binomial Distribution.

If $X \sim B(n,p)$ then $\mu = E(X) = np$ and $\sigma^2 = \text{Var}(X) = \text{npq}$, where $q = 1\text{-}p$. If n is large such that $np >$ 5 and $nq > 5$, the following approximation can be used

$X \sim N(np,npq)$



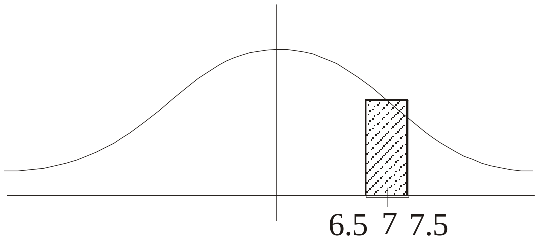
The closer p is to 0.5, the better the approximation.

Continuity corrections

When using this approximation, remember to use a **continuity correction**. This is necessary because the normal distribution is continuous and it is being used to model the binomial distribution which is discrete. Make sure that you understand the following examples.

(a) P(X = 7) becomes P(6.5 < X < 7.5)
and we write

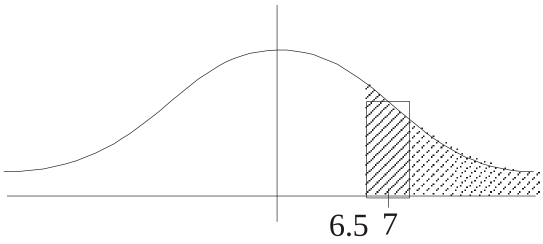
P(X = 7) $\rightarrow$ P(6.5 < X < 7.5)



(b) P(X is at least 7) = P($X \geq 7$)

and we write

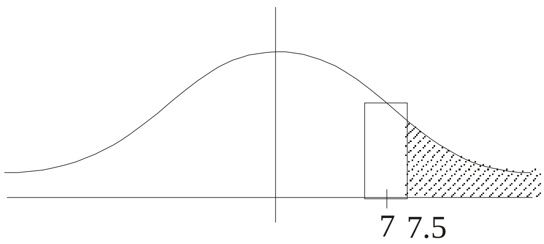
P($X \geq 7$) $\rightarrow$ P($X > 6.5$)



(c) P(X is more than 7) = P($X > 7$)

and we write

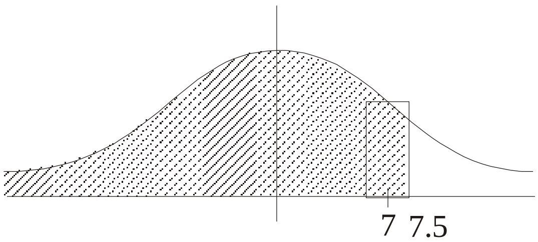
P($X > 7$) $\rightarrow$ P($X > 7.5$)



(d) P(X is 7 or fewer) = P($X \leq 7$)

and we write

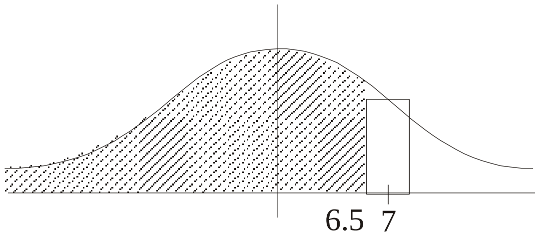
P($X \leq 7$) $\rightarrow$ P($X < 7.5$)



(e) P(X is less than 7) = P(X< 7)

and we write

$P(X < 7) \rightarrow P(X < 6.5)$



Also note that

(f) $P(3 \le X \le 5) \rightarrow P(2.5 < X < 5.5)$

(g) $P(3 \le X < 5) \rightarrow P(2.5 < X < 4.5)$

(h) $P(3 < X \le 5) \rightarrow P(3.5 < X < 5.5)$

(i) $P(3 < X < 5) \rightarrow P(3.5 < X < 4.5)$

Example 1

A traffic survey shows that 60% of cars turn left at a T-junction. Find the probability that out of 100 cars,

(a) more than 70 cars turn left

(b) between 61 and 70 cars inclusive, will turn left.

Example 2

Mountain guides at Mount Kinabalu estimate that 5% of the climbers experience sickness due to low air pressure.

(a) In a group of 100 climbers, what is the probability that 4 or more climbers will experience sickness due to low air pressure?

(b) In a group of 200 climbers, what is the probability that 8 or more climbers will experience sickness due to low air pressure?

Example 3

A manufacturer of chocolates produces 3 times as many soft centred chocolates as hard centred ones. Assuming that the chocolates are randomly distributed within boxes of chocolates,

(a) find the probability that in a box containing 20 chocolates there are

(i) an equal number of soft centred and hard centred chocolates,

(ii) fewer than 5 hard centred chocolates.

(b) A random sample of 5 boxes is taken. Find the probability that exactly 3 of them contain fewer than 5 hard centred chocolates.

(c) A large box of chocolates contains 100 chocolates. Using a suitable approximation estimate the probability that it contains fewer than 21 hard centred chocolates.

Example 4

A large box contains many plastic syringes, but previous experience indicates that 10% of the syringes in the box are defective. 5 syringes are taken at random from the box.

(a) Use a binomial model to calculate, giving your answer to three decimal places, the probability that

(i) none of the 5 syringes is defective

(ii) at least 2 syringes out of the 5 are defective.

(b) Discuss the validity of the binomial model in this context.

(c) Instead of removing 5 syringes, 100 syringes are picked at random and removed. A normal distribution may be used to estimate the probability that at least 15 out of the 100 syringes are defective. Give a reason why it may be convenient to use a normal distribution to do this, and calculate the required estimate.

Example 5

Autovend sells new cars and second-hand cars. The probability of selling new car is 0.4 and the probability of selling automatic cars is 0.01. The sales of new and second-hand cars are randomly distributed throughout the week,

(a) For a particular week, where the company sells 20 cars, find the probability that

(i) at most 5 new cars are sold.

(ii) more new cars are sold than second-hand cars.

(b) During a 3 month period 200 cars are sold. Using a suitable approximation find the probability that

(i) less than 85 new cars are sold

(ii) 12 cars with automatic gearboxes are sold.