Chapter 3 (Statistics): Binomial Distribution (Teacher's copy)

Objectives

At the end of the chapter, you should be able to

(a) understand the concepts of discrete random variable and binomial variable;

(b) understand that the binomial distribution B(n, p) is a probability distribution with mean np and variance $np\left(1-p\right)$;

- know the conditions under which the binomial distribution is a suitable model and able to comment on the appropriate use of the model and the assumptions made.

(c) understand the relationship between binomial probabilities and the binomial expansion of $\left(a+b\right)^n$ for positive integer n;

- use of the notations $n!$ and $\begin{pmatrix} n \\ r \end{pmatrix}$

(d) use of graphic calculator to calculate probabilities;

(e) use the binomial distributions to model practical situations.

Content

3.1 Random Variable
3.1.1 Discrete Random Variable
3.2 Motivation
3.3 Binomial Distribution
3.4 Conditions on Binomial Distribution
3.5 Use of Graphic Calculator for Binomial Distributions
3.6 Mean and Variance of a Binomial Distribution
3.7 Mode of a Binomial Distribution

3.1 Random Variable

From the previous chapter, we learnt to calculate probabilities for different outcomes of an experiment. A random variable is a qualitative variable that takes **different numerical values** according to the results of the experiment.

Random variables are represented by capital letters *X*, *Y*, … and the possible observed values of *X*, *Y*, ... are represented by lower case letters *x*, *y*, … .[avoid using Z, B, N, P and T]

Consider the experiment when a fair coin is tossed twice.
The sample space, $\text{S}=\{\text{HH, HT, TH, TT}\}$
We can let <u>*X* represents the number of heads obtained when a fair coin is tossed twice.</u>

We observe that
- the possible values of *X* are 0, 1, 2 (Numerical Values)
- the value it obtained is subjected to chance (Random)

Hence, *X* is a random variable.

The expression $X=1$ describes the event of obtaining 1 head when a fair coin is tossed twice.
The value of $\text{P}(X=1)$ is the probability of obtaining 1 head when a fair coin is tossed twice.
Therefore $\text{P}(X=1)$ = $\text{P(H,T)}+\text{P(T,H)}=\frac{1}{2}\times\frac{1}{2}+\frac{1}{2}\times\frac{1}{2}=\frac{1}{2}$

3.1.1 Discrete Random Variable

A discrete random variable is one which can only take a **finite or countably infinite** number of values**.**

Examples of discrete random variables:

1. A fair coin is tossed 4 times. Let *X* be the random variable denoting the number of heads obtained. Then *X* is a **discrete** random variable as *X* can only take exact values *x* = 0, 1, 2, 3, 4.

2. Let *X* be the random variable denoting the number of eggs that a hen lays in a day. This is a **discrete** random variable because its only possible values are *x* = 0, 1, 2,..., i.e. there is a countable number of values.

3. Let *Y* be the random variable denoting the number of accidents reported on an expressway in 4 hours. *Y* is a discrete random variable because it is possible to have infinite number of accidents in 4 hours but the values are countable.

Note: We usually denote a random variable by a capital letter e.g. *X*, *Y*, *T*, …,
the particular value it takes by a lower case letter e.g. *x*, *y*, *t*, …

Discuss

Which of the following is a discrete random variable and why?

1. 4 cards are drawn from a pack of playing cards. Let X be the random variable "the number of jacks obtained when 4 cards are drawn from a pack of playing cards ".

Yes. X can take values $x = 0,1,2,3,4$. (countable number of values)

2. Let Y be the random variable "The *height* of a randomly chosen 18-year-old student"

No. Y can any values in a given range.

3.2 Motivation

In the previous chapter, we learnt about calculating the probability of an event occurring in a given experiment. Let us consider the following:

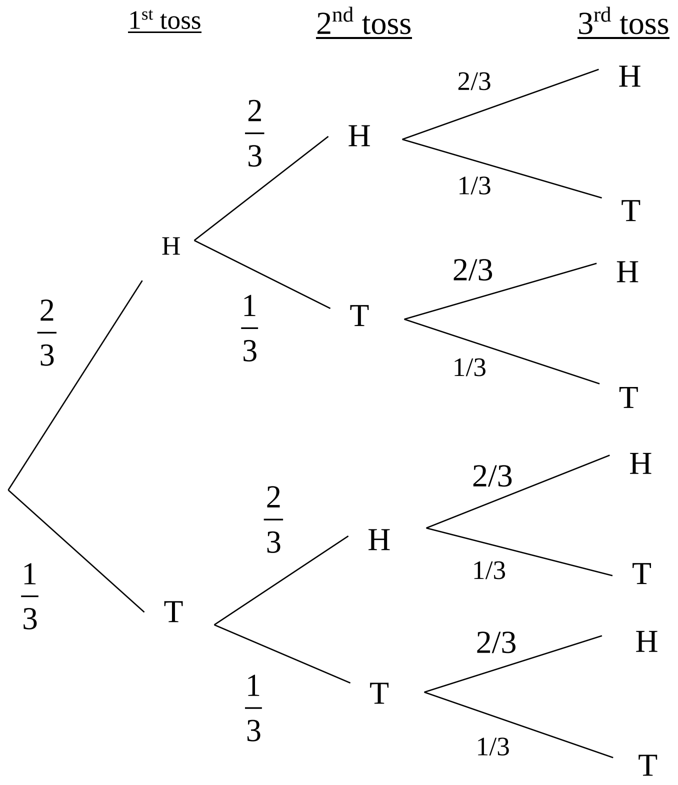
Scenario:

A biased coin is being tossed 3 times. The probability of getting heads in each toss is $\frac{2}{3}$.

Calculate the probability of getting:
(i) no heads
(ii) exactly two heads.

Method:

Since the probability of getting heads for each toss of the coin is known to be $\frac{2}{3}$, therefore the probability of getting tails for each toss of the coin is $\frac{1}{3}$. Let us then draw the probability tree for the experiment where the coin was tossed 3 times:



Using this probability tree, we could then calculate that:

(i) P(no heads was obtained) = P(TTT) = $\left(\frac{1}{3}\right)^3 = \frac{1}{27}$

(ii) P(Exactly two heads were obtained)

$= \mathrm{P}(\mathrm{HHT}) + \mathrm{P}(\mathrm{HTH}) + \mathrm{P}(\mathrm{THH})$

$= \left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right) + \left(\frac{2}{3}\right)\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) + \left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^2$

$= \frac{4}{9}$

Note that in both events, we observed the possible outcomes that would give us **no heads** or **exactly two heads** (**{TTT}** and **{HHT, HTH, THH}** respectively, and calculated the probability of each outcome separately.

Example 1
Based on the example in Section 3.2, what is the probability of getting
(i) no heads, (ii) exactly one head,
(iii) exactly two heads, (iv) all heads.

Solution:
Let X be the random variable denoting the number of heads obtained when the biased coin is tossed 3 times.

	$\text{P}(X=x)$	**Calculation**
(i)	$\text{P}(X=0)$	$\text{P}(\text{TTT})=\binom{3}{0}\left(\frac{1}{3}\right)^3=\frac{1}{27}$
(ii)	$\text{P}(X=1)$	$\text{P}(\text{HTT})+\text{P}(\text{THT})+\text{P}(\text{TTH})=\binom{3}{1}\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^2=\frac{2}{9}$
(iii)	$\text{P}(X=2)$	$\text{P}(\text{HHT})+\text{P}(\text{HTH})+\text{P}(\text{THH})=\binom{3}{2}\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)=\frac{4}{9}$
(iii)	$\text{P}(X=3)$	$\text{P}(\text{HHH})=\binom{3}{3}\left(\frac{2}{3}\right)^3=\frac{8}{27}$

We can therefore deduce that $\text{P}(X=x)=\binom{3}{x}\left(\frac{2}{3}\right)^x\left(\frac{1}{3}\right)^{3-x}$ where $x=0,1,2,3$.

For a scenario with 3 coin flips, it is still easy to manually calculate the probability of 0, 1, 2, or 3 heads. However, what would happen if we were asked to consider a scenario with 100 coin flips? 1000 coin flips? Such calculations would be too tedious.

This is where the generalisation above comes in (for $\text{P}(X=x)$). Using the above expression, we can easily calculate the probability of x heads, given that we know the number of coin flips and the probability of obtaining heads.

This idea of generalising for experiments with only **two** outcomes is the motivation behind this chapter, **Binomial Distribution**.

3.3 Binomial Distribution

Definition
Let X be the random variable denoting the number of successes in n trials of a binomial experiment, then X is said to follow a **Binomial distribution.** This is written as:

$$X \sim \mathrm{B}(n, p)$$

Note that the **number of trials**, n, and the **probability of success** of a trial, p, are required to describe the binomial distribution.
The values of n and p are known as the **parameters** of the distribution.

If $X \sim \mathrm{B}(n, p)$, the probability of obtaining exactly x successes in n trials is denoted by $\mathrm{P}(X = x)$, where

$$\mathrm{P}(X = x) = {}^{n}\mathrm{C}_{x}\, p^{x} q^{n-x} \text{ for } x = 0, 1, 2, \ldots, n. \quad \text{(given in MF26)}$$

(p is the probability of success and $q = 1 - p$ is the probability of failure)

<u>Note</u>:

1. $\mathrm{P}(X = 0) = {}^{n}\mathrm{C}_{0}\, p^{0} q^{n-0} = q^{n}$

2. $\mathrm{P}(X = n) = {}^{n}\mathrm{C}_{n}\, p^{n} q^{n-n} = p^{n}$

3. $$\begin{aligned}(q+p)^{n} &= {}^{n}C_{o}\, p^{0} q^{n} + {}^{n}C_{1}\, p q^{n-1} + {}^{n}C_{2}\, p^{2} q^{n-2} + \ldots + {}^{n}C_{n}\, p^{n} q^{0} \\ &= \mathrm{P}(X=0) + \mathrm{P}(X=1) + \mathrm{P}(X=2) + \ldots + \mathrm{P}(X=n) \\ &= \sum_{x=0}^{n} \mathrm{P}(X = x)\end{aligned}$$

$$\Rightarrow \quad \sum_{x=0}^{n} \mathrm{P}(X = x) = 1 \quad \text{since} \quad (q+p)^{n} = 1 \text{ as } p + q = 1$$

3.4 Conditions on Binomial Distribution

Not all experiments can be modelled with a binomial distribution. There are 4 conditions to be fulfilled for an experiment to be considered a **binomial model**:

1. The experiment consists of a **finite** number of repeated trials.
2. Each trial has **exactly two possible outcomes** (namely, ‘success’ or ‘failure’).
3. The probability of success, denoted by p, is **constant** for each trial.
4. The trials are **independent**.

   (<u>Recall</u>: Events are said to be **independent** if the occurrence of any one event does not affect the probability of occurrence of another.)

Let us now look at why the experiment in **Example 1** is a binomial model:

1. The coin is tossed **3** times (tossing the coin is a trial; **3** is **finite**).
2. Tossing a coin has **exactly 2 outcomes**, heads ('success') or tails ('failure'); we can also say not getting heads is a 'failure').
3. The probability of getting heads is **constant** at $\frac{2}{3}$.
4. The outcome of a coin toss (whether heads or tails is obtained) is **independent** of the outcome of a separate coin toss.

Exercise 1

For each of the following experiments described below, justify whether the random variables defined follows a binomial distribution.

(a) Mary shoots at a target 20 times. The probability that she will hit the target each time is 0.4. Let X be the random variable "number of times she hits the target".

Solution:

Conditions to satisfy a Binomial model	**Is (a) a Binomial Model?**
(i) The experiment consists of a finite number, n, of repeated trials. (ii) Each trial has two possible outcomes, namely, a 'success' or a 'failure'. (iii) The probability of success, denoted by p, is constant for each trial. (iv) The trials are independent. (Recall: Events are said to be independent if the occurrence of one of them does not affect the probability of occurrence of the other.)	(i) 20 tries to hit the target. (ii) There are only 2 outcomes, hit or no hit. (iii) The probability of Mary hitting the target is constant at 0.4 for every attempt. (iv) The results of a shot at the target is (assumed to be) independent of the results of another shot.

Hence X **does** follow a binomial distribution.

(b) A bag contains 10 red balls and 6 green balls. A ball is drawn at random 4 times in succession without replacement. Let Y be the random variable "number of red balls drawn".

Solution:

Conditions to satisfy a Binomial model	**Is (b) a Binomial Model?**
(i) The experiment consists of a finite number, n, of repeated trials. (ii) Each trial has two possible outcomes, namely, a 'success' or a 'failure'. (iii) The probability of success, denoted by p, is constant for each trial. (iv) The trials are independent. (Recall: Events are said to be independent if the occurrence of one of them does not affect the probability of occurrence of the other.)	(i) There are 4 draws (finite). (ii) There are only 2 outcomes, drawing a red ball, or drawing a non-red (green) ball. (iii) The probability of drawing a red ball is not **constant** for every draws. (iv) The outcomes of successive draws are **not independent**.

Hence Y ~~does~~ / **does not** follow a binomial distribution.

(c) In a Physics test which consists of 40 Multiple Choice Questions, there are 5 choices in each question, of which only one is correct. A student attempts to answer all the questions by guessing randomly. Let U be the random variable "number of questions that a student guesses correctly".

Solution:

Conditions to satisfy a Binomial model	**Is (c) a Binomial Model?**
(i) The experiment consists of a finite number, n, of repeated trials.	(i) There are 40 MCQ (finite).
(ii) Each trial has two possible outcomes, namely, a 'success' or a 'failure'.	(ii) There are 2 possible outcomes, each guess is correct or incorrect.
(iii) The probability of success, denoted by p, is constant for each trial.	(iii) The probability of the student guessing correctly is constant at 0.2 for every question.
(iv) The trials are independent. (Recall: Events are said to be independent if the occurrence of one of them does not affect the probability of occurrence of the other.)	(iv) The guess for each question is assumed to be independent of the guess of other questions (since the student attempts to answer all the questions by guessing randomly).

Hence U **does** / ~~does not~~ follow a binomial distribution.

Example 2

The probability that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted the disease, what is the probability that

(i) exactly 5 survive,
(ii) at least 2 survive,
(iii) 10 to 12 survive,
(iv) at most half of them survive?

Solution:

Let X be the random variable denoting the number of people who survive out of 15 people

$$X \sim \mathrm{B}(15,\ 0.4)$$

$$\mathrm{P}(X = x) = {}^{15}\mathrm{C}_x\,(0.4)^x\,(0.6)^{15-x}\ ,\ x = 0, 1, 2, \ldots, 15$$

(i) $\mathrm{P}(X = 5) = {}^{15}\mathrm{C}_5\,(0.4)^5\,(0.6)^{10}$
$= 0.186$

(ii) $\mathrm{P}(X \geq 2)$
$= 1 - P(X < 2)$
$= 1 - P(X \leq 1)$

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

$= 1 - (0.6)^{15} - {}^{15}C_1\,(0.4)^1\,(0.6)^{14}$

$= 0.995$

(iii) $P(10 \le X \le 12) = P(X = 10) + P(X = 11) + P(X = 12)$

$= {}^{15}C_{10}\,(0.4)^{10}\,(0.6)^5 + {}^{15}C_{11}\,(0.4)^{11}\,(0.6)^4 + {}^{15}C_{12}\,(0.4)^{12}\,(0.6)^3$

$= 0.0336$

OR

$P(10 \le X \le 12) = P(X \le 12) - P(X \le 9)$

(iv) $P(X \le 7.5) = P(X \le 7)$

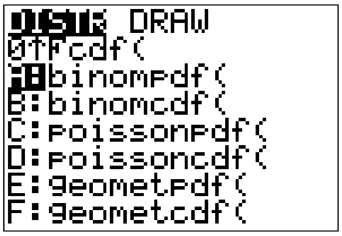
$= P(X = 0) + P(X = 1) + \ldots + P(X = 7)$

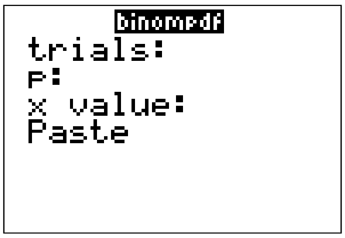
$= 0.787$

3.5 Use of Graphic Calculator for Binomial Distributions

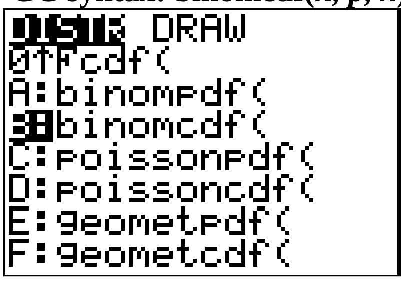
For $X \sim B(n, p)$, the function **binompdf(*n*, *p*, *x*)** is used to find **P(*X* = *x*)** .

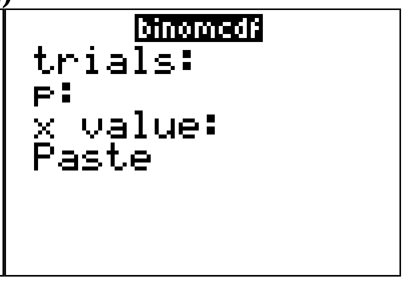
GC syntax: binompdf(*n*, *p*, *x*)





For $X \sim B(n, p)$, the function **binomcdf(*n*, *p*, *x*)** is used to find **P(*X* ≤ *x*)** .
(Note that the equality sign is included)
GC syntax: binomcdf(*n*, *p*, *x*)





Note:
The GC only allows us to determine $P(X \le x)$.

To find $P(X < x)$, we need to rewrite $P(X < x)$ as $P(X \le x - 1)$.
e.g. $P(X < 6) = P(X \le 6 - 1) = P(X \le 5)$

In **Example 2**, $X \sim \mathrm{B}(15, 0.4)$,

(i) $\mathrm{P}(X = 5) = 0.186$

Steps	Screenshot	Remarks
Press 2nd VARS ALPHA MATH You should see this screen.	NORMAL FLOAT AUTO REAL RADIAN MP binompdf trials: p: x value: Paste	
Enter the values of : trials (n), probability of success (p) and number of success (x). Press **ENTER**.	NORMAL FLOAT AUTO REAL RADIAN binompdf trials:15 p:0.4 x value:5 Paste	
	binompdf(15,.4,5) .1859378448	

(ii) $\mathrm{P}(X \geq 2) = 1 - \mathrm{P}(X < 2) = 1 - \mathrm{P}(X \leq 1) = 0.995$

Steps	Screenshot	Remarks
Press 2nd VARS ALPHA APPS You should see this screen.	NORMAL FLOAT AUTO REAL RADIAN binomcdf trials: p: x value: Paste	Note the heading is 'binomcdf'
Enter the values of : trials (n), probability of success (p) and number of success (x). Press **ENTER**, **ENTER**.	NORMAL FLOAT AUTO REAL R binomcdf trials:15 p:0.4 x value:1 Paste	
Press 1 – 2nd (-)	NORMAL FLOAT AUTO REAL RADIAN MP binomcdf(15,0.4,1) 0.005172 1-Ans 0.994827	

(iii) $P(10 \le X \le 12) = P(X \le 12) - P(X \le 9) = 0.0336$

OR $P(X = 10) + P(X = 11) + P(X = 12) = 0.0336$

(iv) $P(X \le 7) = 0.787$

In short,

Phrase	Notation	Use of GC to calculate probability
"…equals x…"	$P(X = x)$	**GC: binompdf (*n*, *p*, *x*)**
"…less than x…"	$P(X < x)$	$P(X < x) = P(X \le x-1)$ **GC: binomcdf (*n*, *p*, $x-1$)**
"…more than x…"	$P(X > x)$	$P(X > x) = 1 - P(X \le x)$ **GC: 1 – binomcdf (*n*, *p*, *x*)**
"….at most x…" or "…not more than x…"	$P(X \le x)$	**GC: binomcdf (*n*, *p*, *x*)**
"…at least x…" or "…not less than x…"	$P(X \ge x)$	$P(X \ge x) = 1 - P(X \le x-1)$ **GC: 1 – binomcdf (*n*, *p*, $x-1$)**

Example 3

Given that 20% of the population is left-handed,

(i) find the probability that a group of eight people contains not more than three left-handers.

(ii) find the least number of people that should be selected if the probability that there is at least one left-hander is greater than 0.95.

Solution:

(i) Let X be the random variable denoting the number of left-handers out of a group of 8 people

$X \sim B(8, 0.2)$

$P(X \le 3) = 0.944$ (to 3 s.f.)

(ii) Let Y be the random variable denoting the number of left-handers out of a group of n people" .

$Y \sim B(n, 0.2)$

$P(Y \ge 1) > 0.95$

$1 - P(Y = 0) > 0.95$

$P(Y = 0) < 0.05$

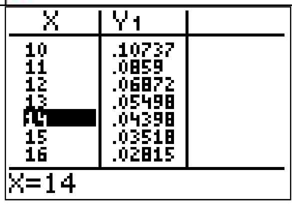
Aim: To find smallest n such that $P(Y = 0)$ is less than 0.05

Method 1: Analytical Method

$$^{n}C_0\,(0.2)^0\,(1 - 0.2)^n < 0.05$$
$$(0.8)^n < 0.05$$
$$\ln(0.8)^n < \ln(0.05)$$
$$n\ln(0.8) < \ln(0.05)$$
$$n > \frac{\ln(0.05)}{\ln(0.8)}$$ (note the change in sign as $\ln(0.8) < 0$)
$$\Rightarrow n > 13.4$$

$\therefore$ least n is 14.

Method 2: Using GC

Steps	**Screenshot**	**Remarks**
Press **Y**$_1$ = Then Press 2nd VARS ALPHA MATH Key in binompdf(X, 0.2, 0) You should see this screen.		
Press **2**$^{\text{nd}}$ **TABLE**. Scroll down the List Y$_1$ and select the least value of X with probability less than 0.05.		

Solution (for presentation):
Using GC,

n	$P(Y = 0)$
13	0.05498 > 0.05
14	0.04398 < 0.05
15	0.03518 < 0.05

Thus, least n is 14.

Exercise 2

1. Given that $X \sim B(20, 0.6)$, use GC to find

(i) $P(X = 7)$
(ii) $P(X \le 9)$
(iii) $P(X < 7)$
(iv) $P(X > 5)$
(v) $P(X \ge 10)$
(vi) $P(10 < X < 19)$
(vii) $P(6 < X \le 15)$
(viii) $P(4 \le X < 13)$
(ix) $P(5 \le X \le 16)$
(x) $P(X \ge 7 \mid X < 18)$.

[Ans: (i) 0.0146 (ii) 0.128 (iii) 0.00647 (iv) 0.998
(v) 0.872 (vi) 0.755 (vii) 0.943 (viii) 0.584
(ix) 0.984 (x) 0.994]

Solution:

(i) $\mathrm{P}(X=7)=0.0146$

(ii) $\mathrm{P}(X\le 9)=0.128$

(iii) $\mathrm{P}(X<7)=\mathrm{P}(X\le 6)=0.00647$

(iv) $\mathrm{P}(X>5)=1-\mathrm{P}(X\le 5)=0.998$

(v) $\mathrm{P}(X\ge 10)=1-\mathrm{P}(X<10)=1-\mathrm{P}(X\le 9)=0.872$

(vi) $\mathrm{P}(10<X<19)=\mathrm{P}(X\le 18)-\mathrm{P}(X\le 10)=0.755$

(vii) $\mathrm{P}(6<X\le 15)=\mathrm{P}(7\le X\le 15)=\mathrm{P}(X\le 15)-\mathrm{P}(X\le 6)=0.943$

(viii) $\mathrm{P}(4\le X<13)=\mathrm{P}(X\le 12)-\mathrm{P}(X\le 3)=0.584$

(ix) $\mathrm{P}(5\le X\le 16)=\mathrm{P}(X\le 16)-\mathrm{P}(X\le 4)=0.984$

(x) $\mathrm{P}(X\ge 7\,|\,X<18)=\dfrac{\mathrm{P}(7\le X<18)}{\mathrm{P}(X<18)}=\dfrac{\mathrm{P}(X\le 17)-\mathrm{P}(X\le 6)}{\mathrm{P}(X\le 17)}=\dfrac{0.98992}{0.99639}=0.994$

2. Given that $X\sim \mathrm{B}(15, 0.2)$, find the least integer r such that $\mathrm{P}(X<r)>0.85$. [Ans: 6]

Solution:

$\mathrm{P}(X<r)=\mathrm{P}(X\le r-1)>0.85$

We key in binomcdf(15, 0.2, X - 1)

Using GC,

r	$\mathrm{P}(X\le r-1)$
5	$0.83577<0.85$
6	$0.93895>0.85$
7	$0.98194>0.85$

Thus, least integer r is 6.

3. The probability that a particular make of light bulb is faulty is 0.2. The light bulbs are packed in boxes of 12.
State, in the context of this question, two assumptions needed for the number of faulty light bulbs to be well modelled by a binomial distribution.

(a) Find the probability there are
   (i) more than 3 faulty light bulbs in two boxes
   (ii) more than 3 faulty light bulbs in each of the 2 boxes

(b) A buyer accepts a consignment of boxes if when he chooses three boxes at random, he finds that they contain no more than one faulty light bulb altogether. Find the probability that he will accept the consignment.

[Ans: (ai) 0.736 (aii)0.0422 (b) 0.00325]

Solution:

Assumption:

1. A light bulb being faulty is independent of any other light bulb being faulty.

2. The probability of a light bulb being faulty is constant at 0.2 for every ligh bulb.

(a)(i) Let X be the random variable denoting the number of faulty light bulbs in two boxes.

$$X \sim \text{B}(24, 0.2)$$

$$\text{P}(X > 3) = 1 - \text{P}(X \le 3) = 0.736$$

(a)(ii) Let Y be the random variable denoting the number of faulty light bulbs in a box of 12.

$$Y \sim \text{B}(12, 0.2)$$

$$\text{P}(Y > 3) = 1 - \text{P}(Y \le 3) = 0.20543$$

P(more than 3 faulty light bulbs in each of the 2 boxes) $= \text{P}(Y > 3)\text{P}(Y > 3)$

$= 0.20543^2$

$= 0.0422$

(b) Let U be the random variable denoting the number of faulty light bulbs out of 36.

$$U \sim \text{B}(36, 0.2)$$

P(accept consignment) = $\text{P}(U \le 1) = 0.00325$.

4. In a certain country, it is known that 80% of the adult population are right-handed, 15% are left-handed and 5% are ambidextrous (i.e. able to use either hand equally well). 10 adults from this country are selected at random. State, in the context of this question, two assumptions needed for the number of ambidextrous adults to be well modelled by a binomial distribution.

(a) Find the probability that

(i) at least one of those chosen will be ambidextrous

(ii) more than six of those chosen will be left-handed

(iii) at most four of those chosen will be right-handed

(iv) less than three of those chosen will be right-handed.

(v) Given that there are more than one ambidextrous adult in the sample of 10, find the probability that there are at most four of those chosen who will be ambidextrous.

(b) Find the least value of n such that the probability of obtaining at least three left-handed adults in a random sample of n adults exceeds 0.06.

(c) Find the greatest value of m if, in a random sample of m adults, the probability of at least 1 right-handed is at most 0.96.

[Ans: (ai) 0.401 (aii) 0.000135 (aiii) 0.00637 (aiv) 0.0000779 (av) 0.999 (b) 7 (c) 2]

Solution:

Assumption:

1. An adult being ambidextrous is independent of any other adult being ambidextrous.
2. The probability of an adult being ambidextrous is constant at 0.05 for every adult.

(ai) Let X be the random variable denoting the number of ambidextrous adults out of 10 adults

$$X \sim \text{B}(10, 0.05)$$

$$\text{P}(X \ge 1) = 1 - \text{P}(X < 1) = 1 - \text{P}(X = 0) = 0.401$$

(aii) Let Y be the random variable denoting the number of left-handed adults out of 10 adults

$$Y \sim B(10,\ 0.15)$$

$P(Y > 6) = 1 - P(Y \leq 6) = 0.000135$

(aiii) Let T be the random variable denoting the number of right-handed adults out of 10 adults

$$T \sim B(10,\ 0.80)$$

$P(T \leq 4) = 0.00637$

(aiv) $P(T < 3) = P(T \leq 2) = 0.0000779$

(av) P(there are at most four of those chosen who are ambidextrous | there are more than one ambidextrous adult)

$= P(X \leq 4 \mid X > 1) = \dfrac{P(X \leq 4 \cap X > 1)}{P(X > 1)} = \dfrac{P(1 < X \leq 4)}{P(X > 1)} = \dfrac{P(2 \leq X \leq 4)}{P(X > 1)}$

$= \dfrac{P(X \leq 4) - P(X \leq 1)}{1 - P(X \leq 1)} = 0.999$

(b) Let U be the random variable denoting the number of left-handed adults out of n adults

$$U \sim B(n,\ 0.15)$$

P(at least three left-handed in a random sample of n adults) > 0.06

$P(U \geq 3) > 0.06$

$1 - P(U < 3) > 0.06$

$1 - P(U \leq 2) > 0.06$

$P(U \leq 2) < 0.94$

Using GC,

n	$P(U \leq 2)$
6	0.95266 > 0.94
7	0.92623 < 0.94
8	0.89479 < 0.94

Thus, least value of n is 7.

(c) Let W be the random variable denoting the number of right-handed adults out of m adults

$$W \sim B(m, 0.80)$$

$P(W \geq 1) \leq 0.96$

$1 - P(W = 0) \leq 0.96$

$P(W = 0) \geq 0.04$

Using GC,

m	$P(W = 0)$
1	0.2 ≥ 0.04
2	0.04 ≥ 0.04

3	$0.008 < 0.04$

Thus, greatest value of m is 2.

3.6 Mean and Variance of a Binomial Distribution

If discrete random variable X follows a Binomial distribution with parameters n and p,
i.e. $X \sim \mathrm{B}(n, p)$, then

$$\text{Mean of } X,\ \mathrm{E}(X) = np$$

$$\text{Variance of } X,\ \mathrm{Var}(X) = np(1-p)$$

Note:
1. The mean of a random variable provides the long run average of the variable on the expected average outcome over many observations.
2. Standard deviation of $X = \sqrt{\mathrm{Var}(X)} = \sqrt{np(1-p)}$.

Example 4

The probability that it will be a rainy day is 0.6. Find the standard deviation and expected number of rainy days in a week. State an assumption for your working.

Solution:

Assumption: The weather condition on any particular day is independent of the weather condition on any other days.

Let X be the random variable denoting the number of rainy days in a week

$$X \sim \mathrm{B}(7, 0.6)$$

Expected no. of rainy days in a week = $\mathrm{E}(X) = np = 7 \times 0.6 = 4.2$

Standard deviation = $\sqrt{\mathrm{Var}(X)} = \sqrt{np(1-p)} = \sqrt{7(0.6)(0.4)} = 1.30$ days

Example 5

A random variable X follows a binomial distribution such that $X \sim \mathrm{B}(n, p)$. Given that the mean of X is 2.6 and the variance of X is 2.08, find n and p.

Solution:

$\mathrm{E}(X) = 2.6$ $\qquad$ $\mathrm{Var}(X) = 2.08$

$np = 2.6 \cdots (1)$ $\qquad$ $np(1-p) = 2.08 \cdots (2)$

Sub (1) into (2), $2.6(1-p) = 2.08$

$$p = 0.2$$

$$n = 13$$

3.7 Mode of a Binomial Distribution

If $X \sim \text{B}(n, p)$, then Mode of X is the value of X that is most likely to occur.

Usually, mode is near the mean.

The mode can be found by calculating all the probabilities and finding the value of X with the highest probability.

Note that there can be more than 1 mode.

Example 6

Of the inhabitants of a certain African village, 80% are known to have an eye disorder. If 12 people are waiting to see the doctor, what is the most likely number of them to have the eye disorder?

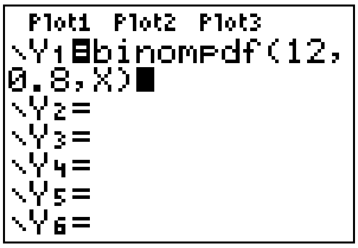
Solution:

Let X be the random variable denoting the number of people who have an eye disorder out of 12 people

$$X \sim \text{B}(12, 0.8)$$

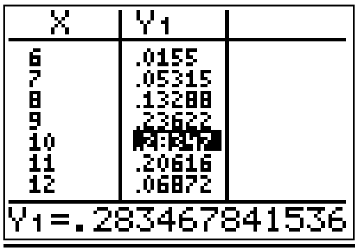
To find the most likely number or the most probable value of X, we list down the probabilities of X using a graphic calculator and find the value of x with the greatest value of $\text{P}(X = x)$.

1. Press **Y_1 =**
2. Press **2nd DISTR** to select **A:binompdf (**
3. Key in binompdf(12, 0.8, X)
4. Press **2nd TABLE** to get to the TABLE.
5. Scroll down the List Y_1 and select the value of x with the greatest probability.



From GC, $X = 10$ gives the highest probability.

$\text{P}(X = 9) = 0.23622$
$\text{P}(X = 10) = 0.28347$
$\text{P}(X = 11) = 0.20616$



Therefore the most likely number of people is 10.

Example 7

A random variable X follows a binomial distribution such that $X \sim \text{B}(5, 0.5)$. Find the modal number (modes) of the distribution.

Solution:

Using GC, $\text{P}(X = 2) = 0.3125$ and $\text{P}(X = 3) = 0.3125$, therefore the modes are 2 and 3.

Example 8

In a certain inspection scheme, a sample of ten items is selected at random from a very large batch and the number of defectives is recorded. If this number is more than two, the batch is rejected, and it is accepted if there are no defectives. Otherwise, a further sample of five items is selected, again at random. If there are any defectives this time, the batch is rejected, otherwise it is accepted. If the proportion of defectives is in fact 10%, find the probability that

(i) the batch is accepted as a result of the first inspection,
(ii) a further sample is taken,
(iii) the batch is accepted after a second sample is taken,
(iv) the batch is accepted given that a second sample is taken,
(v) the batch is accepted.

Solution:
Let X be the random variable denoting the number of defectives in the first sample of 10 items

$$X \sim \mathrm{B}(10, 0.1)$$

(i) P(batch is accepted as a result of the 1st inspection) $= \mathrm{P}(X = 0) = (0.9)^{10} = 0.349$

(ii) P(a further sample is taken) $= \mathrm{P}(X = 1 \text{ or } X = 2) = \mathrm{P}(X = 1) + \mathrm{P}(X = 2)$
$= 0.581$

(iii) Let Y be the random variable denoting the number of defectives in second sample of 5

$$Y \sim \mathrm{B}(5, 0.1)$$

P(batch is accepted after a 2nd sample is taken)
$= \mathrm{P}(X = 1 \text{ or } X = 2)\ \mathrm{P}(Y = 0)$
$= (0.58113)(0.9)^5 = 0.343$

(iv) $$\mathrm{P}(\text{batch is accepted} \mid 2^{\text{nd}} \text{ sample taken}) = \frac{\mathrm{P}(\text{2nd sample taken and batch accepted})}{\mathrm{P}(\text{2nd sample taken})}$$
$$= \frac{\mathrm{P}\left(X=1 \text{ or } X=2\right)\mathrm{P}(Y=0)}{\mathrm{P}(X=1 \text{ or } X=2)}$$
$$= \mathrm{P}(Y = 0) = (0.9)^5 \approx 0.590$$

(v) P(batch is accepted) $= \mathrm{P}(X = 0) + \mathrm{P}(X = 1 \text{ or } X = 2)\ \mathrm{P}(Y = 0)$
$= 0.34868 + 0.34315 \approx 0.692$

Example 9

In XYZ School, 1 out of 10 students are myopic. The school has 40 classes and each class has 30 students. State 2 assumptions needed for the myopic students to be well modelled by a Binomial Distribution.

(i) Find the probability that there are at least 4 myopic students in a class.

(ii) Find the probability that the school has at most 6 classes, each with at least 4 myopic students in the class.

In PQR School, 25% of the students are myopic. A random class with class size n is taken to study the problem of myopic in the school. Find the least value of n such that the probability of getting at least 2 myopic students in the sample is greater than 0.97.

Solution:

Assumption:

1. A student being myopic is independent of any other student being myopic.
2. The probability of a student being myopic is constant at 0.1 for every student.

(i) Let X be the random variable denoting the number of myopic students out of 30 students

$$X \sim \mathrm{B}(30, 0.1)$$

$$\begin{aligned}\text{P(at least 4 myopic students in a class)} &= \mathrm{P}(X \geq 4) \\ &= 1 - \mathrm{P}(X < 4) \\ &= 1 - \mathrm{P}(X \leq 3) \\ &= 0.353\end{aligned}$$

(ii) Let Y be the random variable denoting the number of classes that have at least 4 myopic students out of 40 classes

$$Y \sim \mathrm{B}(40, 0.35256)$$

$$\begin{aligned}\text{P(at most 6 classes that have at least 4 myopic students in the class)} &= \mathrm{P}(Y \leq 6) \\ &= 0.00397\end{aligned}$$

Let Q be the random variable denoting the number of myopic students out of n students

$$Q \sim \mathrm{B}(n, 0.25)$$

$$\mathrm{P}(\mathrm{Q} \geq 2) > 0.97$$

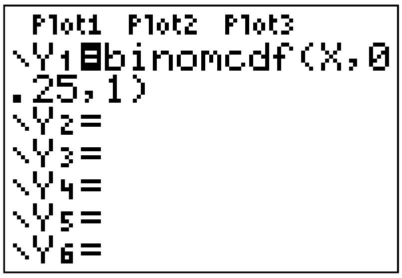
$$1 - \mathrm{P}(\mathrm{Q} < 2) > 0.97$$

$$1 - \mathrm{P}(\mathrm{Q} \leq 1) > 0.97$$

$$\mathrm{P}(\mathrm{Q} \leq 1) < 0.03$$

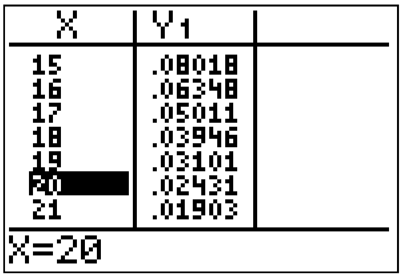
Use GC:

To find the least value of n, we list down the $\mathrm{P}(\mathrm{Q} \leq 1)$ using a graphic calculator and find the least value of n such that $\mathrm{P}(\mathrm{Q} \leq 1) < 0.03$.



1. Press **Y_1 =**
2. Press **2nd DISTR** to select **A:binomcdf(**
3. Key in **binomcdf(*X*, 0.25, 1)**
4. Press **2nd TABLE** to get to the TABLE.
5. Scroll down the List Y_1 and select the value of x with first Y_1 smaller than 0.03.

Using GC,



n	$\text{P}(Q \le 1)$
19	$0.03101 > 0.03$
20	$0.02431 < 0.03$
21	$0.01903 < 0.03$

Therefore the least value of n is 20.

Exercise 3

1. A box contains a large number of red and yellow tulip bulbs in the ratio 1 : 3. Bulbs are picked at random from the box. If X is the number of red tulip bulbs out of 12 bulbs, find
   (i) $\text{P}(4 \le X < 6)$
   (ii) $\text{E}(X)$
   (iii) $\text{P}(X < \text{Var}(X))$.

[Ans: (i) 0.297 (ii) 3 (iii) 0.391]

Solution:
Let X be the random variable denoting the number of red tulip bulbs out of 12 bulbs

$X \sim \text{B}(12, \frac{1}{4})$

(i) $\text{P}(4 \le X < 6) = \text{P}(4 \le X \le 5) = \text{P}(X \le 5) - \text{P}(X \le 3) = 0.297$

[binomcdf(12,0.25,5) − binomcdf(12,0.25,3)]

(ii) $\text{E}(X) = np = 12 \times \frac{1}{4} = 3$

(iii) $\text{Var}(X) = npq = 12 \times \frac{1}{4} \times \frac{3}{4} = \frac{9}{4} = 2\frac{1}{4}$

$\text{P}(X < \text{Var}(X)) = \text{P}(X < 2\frac{1}{4}) = \text{P}(X \le 2) = 0.391$

[binomcdf(12,0.25,2)]

2. The random variable X has a binomial distribution with mean 12 and variance 8. Find $\mathrm{P}(6 \le X < 16)$.

[Ans: 0.883]

Solution:

$np = 12$ and $np(1-p) = 8$

$$12(1-p) = 8$$
$$1 - p = \frac{8}{12}$$
$$p = 1 - \frac{8}{12}$$
$$= \frac{1}{3}$$
$$n\left(\frac{1}{3}\right) = 12$$
$$n = 36$$

$\therefore\ X \sim \mathrm{B}(36,\ \frac{1}{3})$

$$\mathrm{P}(6 \le X < 16) = \mathrm{P}(X \le 15) - \mathrm{P}(X \le 5)$$
$$= 0.883$$

Question Prompt

Would the answer be the same or different if the question is changed to find $\mathrm{P}(6 \le X < 16.5)$ instead?

Ans:

The answer would be different because

$\mathrm{P}(6 \le X < 16.5)$

$= \mathrm{P}(6 \le X \le 16)$

$= \mathrm{P}(X \le 16) - \mathrm{P}(X \le 5)$

$= 0.934$

3. Samples, each of 8 articles, are taken at random from a large consignment in which 20% of articles are defective. Find the number of defective articles which is most likely to occur in a single sample, and find the probability of obtaining the number.

If 100 samples of 8 articles are to be examined, calculate the expected number of samples in which you would find 3 or more defective articles.

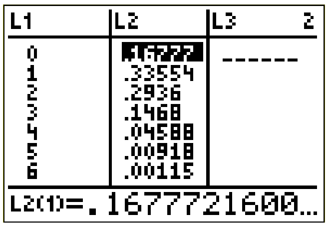
[Ans: 1, 0.336, 20.3]

Solution:

Let X be the random variable denoting the number of defective articles out of 8

$X \sim \mathrm{B}(8,\ 0.20)$

Using GC,
P(X = 0) = 0.168
P(X = 1) = 0.336
P(X = 2) = 0.294
Thus, the number of defective articles which is most likely to occur is 1.
P(X = 1) = 0.336



$P(X \geq 3) = 1 - P(X \leq 2)$
$= 0.20308.$

Let Y be the random variable denoting the number of samples which contain 3 or more defective articles out of 100 samples

$$Y \sim B(100, 0.20308)$$

Expected number of samples with 3 or more defective articles is 100 (0.20308) = 20.3

4. In a large batch of television sets, the proportion of defective television sets is 2%. A random sample of 20 television sets is to be tested. The batch of television sets is rejected if there are more than 2 defective television sets and it is accepted if less than 2 television sets are defective. Otherwise, another random sample of 10 television sets is taken. The batch is accepted if there is no defective television set, otherwise the batch is rejected. Find the probability that the batch of television sets is
(i) accepted as a result of inspection of the first sample.
(ii) accepted as a result of inspection of the second sample.
(iii) rejected.

[Ans: (i) 0.940 (ii) 0.0432 (iii) 0.0167]

Solution:
Let X be the random variable denoting the number of defective television sets out of 20

$$X \sim B(20, 0.02)$$

(i) P(batch is accepted as a result of inspection of the first sample) $= P(X < 2)$
$= P(X \leq 1)$
$= 0.94010$
$= 0.940$

(ii) Let Y be the random variable denoting the number of defective television sets out of 10

$$Y \sim B(10, 0.02)$$

P(batch is accepted as a result of inspection of the second sample)
$= P(X = 2)P(Y = 0)$
$= 0.043166$
$= 0.0432$

(iii) $\text{P(batch is rejected)} = \text{P}(X > 2) + \text{P}(X = 2)\text{P}(Y \geq 1)$

$= 1 - \text{P}(X \leq 2) + \text{P}(X = 2)(1 - \text{P}(Y = 0))$

$= 0.0167$

Alternative Method:

P(batch is rejected) $= 1 - 0.94010 - 0.043166$ $(1 -$ part (i) ans $-$ part (ii) ans$)$

$= 0.0167$

5. It may be assumed that the dates of birth in a large population are distributed throughout the year so that the probability of a randomly chosen person's date of birth being in any particular month is $\frac{1}{12}$.

(a) Find the probability that out of 10 people chosen at random, at least one will have a birthday in January.

(b) Suppose there are 5 samples (each containing 10 people) obtained from the population, find the probability that

(i) each sample has at least one person born in February.

(ii) the second sample is the only sample with at least one person born in February.

(iii) there is one sample with at least one person born in February.

(iv) there are at most 3 samples with at least one person born in February.

[(a) 0.581 (bi) 0.0663 (bii) 0.0179 (biii) 0.0895 (biv) 0.695]

Solution:

(a) Let Y be the random variable denoting the number of people whose birthdays are in January out of 10 people".

$$Y \sim \text{B}(10, \frac{1}{12})$$

$\text{P}(Y \geq 1) = 1 - \text{P}(Y = 0) = 0.58110 = 0.581$

(bi) P(each sample has at least one person born in February) = $(0.5811)^5 = 0.0663$

(bii) P(second sample is the only sample with at least one person born in February)

$= (1 - 0.58110)^4 (0.58110) = 0.0179$

(biii) Let Q be the random variable denoting the number of samples with at least one person born in February out of 5 samples

$Q \sim \text{B}(5, 0.58110)$

$\text{P}(Q = 1) = 0.0895$

Question Prompt
Is there any other method to solve (biii)?

Ans:
Yes. We can make use of the answer of (bii) and multiply it by 5, i.e.
P(there is one sample with at least one person born in February)
$= 5(0.58110)(1-0.58110)^4 = 0.0895$

(biv) $\mathrm{P}(Q \le 3) = 0.695$

Practice Questions

1. Given that $X \sim \mathrm{B}(n, 0.6)$ and $\mathrm{P}(X > 3) > 0.3$, find the least value of *n*.

[Ans: 5]

Solution:
$X \sim \mathrm{B}(n, 0.6)$

$$\begin{aligned} \mathrm{P}(X > 3) &> 0.3 \\ 1 - \mathrm{P}(X \le 3) &> 0.3 \\ \mathrm{P}(X \le 3) &< 1 - 0.3 \\ \mathrm{P}(X \le 3) &< 0.7 \end{aligned}$$

From G.C.,

When $n = 4$, $\mathrm{P}(X \le 3) = 0.8704 > 0.7$

When $n = 5$, $\mathrm{P}(X \le 3) = 0.66304 < 0.7$

When $n = 6$ $\mathrm{P}(X \le 3) = 0.45568 < 0.7$

Hence, least *n* is 5.

2. The probability that a marksman hits a target is 0.3.
   (i) If 24 shots are fired,
      (a) what is the probability that the marksman hits the target 6 times?
      (b) what is the expected number of hits?
   (ii) Find the least number of shots which should be fired if the probability that the target is hit at least once is greater than 0.9?

[Ans: (ia) 0.160 (ib) 7.2 (ii) 7]

Solution:
(i) (a) Let *X* be the random variable denoting the number of hits out of 24 shots
$X \sim \mathrm{B}(24, 0.3)$
$\mathrm{P}(X = 6) = 0.160$

(b) $\mathrm{E}(X) = np = 24 \times 0.3 = 7.2$

(ii) Let *Y* be the random variable denoting the number of hits out of *n* shots

$Y \sim B(n, 0.3)$

$P(Y \geq 1) > 0.9$

$1 - P(Y = 0) > 0.9$

$P(Y = 0) < 0.1$

From G.C, when $n = 6$, $P(Y = 0) = 0.11765 > 0.1$

when $n = 7$, $P(Y = 0) = 0.08235 < 0.1$

when $n = 8$, $P(Y = 0) = 0.05765 < 0.1$

$\therefore$ least n is 7.

3. 20% of the butterflies in a district are of type A. If a random sample of 10 is taken, find the probability that
   (i) there will be just 2 butterflies of type A in the sample
   (ii) the 10^{th} butterfly caught will be the 2^{nd} one of type A.

[Ans: (i) 0.302 (ii) 0.0604]

Solution:

(i) Let X be the random variable denoting the number of type A butterfly in a sample size of 10

$X \sim B(10, 0.2)$

$P(X = 2) = 0.302$

(ii) Let Y be the random variable denoting the number of type A butterfly in a sample size of 9

$Y \sim B(9, 0.2)$

P(the 10^{th} butterfly is the second type A)

$= P(Y = 1)\,P(10^{th} \text{ butterfly is type A})$

$= P(Y = 1) \times 0.2$

$= 0.0604$

4. [N95/P2/Q7b (modified)]
   When a machine is used to dig up potatoes, there is a probability of 0.1 for each individual potato that it will be damaged in the process.
   (i) Find, correct to 3 decimal places, the probability that a random selection of 12 potatoes dug up by machine will include at least 3 damaged ones.
   (ii) A random sample of n potatoes is selected, and the number of damaged potatoes in the sample is denoted by the random variable Y. Write down expressions, in terms of n, for the mean and standard deviation of Y.
   (iii) Find the most likely number of damaged potatoes in a sack of 50 potatoes.

[Ans: (i) 0.111 (ii) $0.1n$, $0.3\sqrt{n}$ (iii) 5]

Solution:

4(i) Let X be the random variable denoting the number of damaged potatoes out of 12 potatoes.

$X \sim B(12, 0.1)$

$P(X \geq 3) = 1 - P(X \leq 2)$
$= 0.11087$
$= 0.111$ (3 d.p.)

(ii) Let Y be the random variable denoting the number of damaged potatoes out of n potatoes.

$Y \sim B(n, 0.1)$

$E(Y) = 0.1n$

Standard deviation of $Y = \sqrt{\mathrm{Var}(Y)}$
$= \sqrt{n(0.1)(0.9)}$
$= 0.3\sqrt{n}$

(iii) Let W be the random variable denoting the number of damaged potatoes out of 50 potatoes.

$W \sim B(50, 0.1)$

From GC, $P(W = 4) = 0.1809$
$P(W = 5) = 0.18492$
$P(W = 6) = 0.1541$

Thus, the most likely number of damaged potatoes in a sack of 50 potatoes is 5.

5. (a) Suppose the random variable X has distribution $B(n, p)$.

(i) Given that $n = 12$, $p = 0.3$, find the mode of X.

(ii) Given that $p = \frac{1}{6}$, find the least value of n such that $P(X \geq 1) > 0.9$.

(iii) Given that $n = 15$, $p = 0.1$, find the greatest value of the integer r such that $P(X < r) \leq 0.95$.

(b) X is a binomial random variable, where the number of trials is 5 and the probability of success of each trial is p. Write down an expression for $P(X = 3)$ in terms of p. Hence find the values of p such that $P(X = 3) = 0.12$.

[Ans: (a) (i) 3 (ii) 13 (iii) 4 (b) 0.287, 0.863]

Solution:

5(a)(i) $X \sim B(12, 0.3)$

From GC, $P(X = 2) = 0.16779$
$P(X = 3) = 0.2397$
$P(X = 4) = 0.23114$

$\therefore$ mode of X is 3.

(ii) $X \sim B\left(n, \frac{1}{6}\right)$

$P(X \geq 1) > 0.9$

$1 - P(X = 0) > 0.9$

$\mathrm{P}(X=0)<0.1$

From GC,
when n = 12, $\mathrm{P}(X=0)=0.11216>0.1$
when n = 13, $\mathrm{P}(X=0)=0.09346<0.1$
when n = 14, $\mathrm{P}(X=0)=0.07789<0.1$
$\therefore$ least value of n is 13.

(iii) $X \sim \mathrm{B}(15, 0.1)$
$\mathrm{P}(X<r)\le 0.95$
$\mathrm{P}(X\le r-1)\le 0.95$
From GC,
$\mathrm{P}(X\le 2)=0.8159<0.95$
$\mathrm{P}(X\le 3)=0.94444<0.95$
$\mathrm{P}(X\le 4)=0.98728>0.95$
$\therefore$ greatest integer $r-1=3$
greatest integer $r=4$

(b) $X \sim \mathrm{B}(5, p)$

$\mathrm{P}(X=3)=\binom{5}{3}p^3(1-p)^2=10p^3(1-p)^2$

$\mathrm{P}(X = 3) = 0.12 \Rightarrow \quad 10p^3(1-p)^2=0.12$

From GC, $p = 0.287$ or $p = 0.863$

6. [N2009/II/Q11(part)]

A fixed number, n, of cars is observed and the number of those cars that are red is denoted by R.

(i) State, in context, two assumptions needed for R to be well modelled by a binomial distribution.

Assume now that R has the distribution $\mathrm{B}(n, p)$.

(ii) Given that $n = 20$ and $p = 0.15$, find $\mathrm{P}(4 \le R < 8)$.

(iii) Given that $n = 20$ and $\mathrm{P}(R = 0 \text{ or } 1) = 0.2$, write down an equation for the value of p, and find this value numerically.

[Ans: (ii) 0.346 (iii) 0.142]

Solution:

6(i) The two assumptions are:
A car being red is independent of any other car being red.
OR The colour of each car must be independent of the colour of any other car. (from exam report)
The probability of a car being red is constant for every car.
OR the probability that any one car is red is the same throughout the sample.

(from exam report)

(ii) $R \sim \mathrm{B}(20, 0.15)$

$\mathrm{P}(4 \le R < 8) = \mathrm{P}(4 \le R \le 7)$

$= \mathrm{P}(R \le 7) - \mathrm{P}(R \le 3)$

$= 0.34635$

$= 0.346$ (3 s.f.)

(iii) $R \sim \mathrm{B}(20, p)$

$\mathrm{P}(R = 0 \text{ or } 1) = 0.2$

$\mathrm{P}(R=0) + \mathrm{P}(R=1) = 0.2$

$\binom{20}{0} p^0 (1-p)^{20} + \binom{20}{1} p(1-p)^{19} = 0.2$

$(1-p)^{20} + 20p(1-p)^{19} = 0.2$

From GC, $p = 0.142$ (3 s.f.)

7. [N2011/II/7]

When I try to contact (by telephone) any of my friends in the evening, I know that on average the probability that I succeed is 0.7. On one evening I attempt to contact a fixed number, n, of different friends. If I do not succeed with a particular friend, I do not attempt to contact that friend again that evening. The number of friends whom I succeed in contacting is the random variable R.

(i) State, in the context of this question, two assumptions needed to model R by a binomial distribution.

(ii) Explain why one of the assumptions stated in part (i) may not hold in this context.

Assume now that these assumptions stated in part (i) do in fact hold.

(iii) Given that $n = 8$, find the probability that R is at least 6.

[Ans: (iii) 0.552]

Solution:

(i) The probabilty of a friend being successfully contacted is constant at 0.7 for every friend.

OR Each phone call was equally likely to be successful. (from exam report)

A friend being successfully contacted is independent of any other friend being successfully contacted.

Or The success of each phone call was independent of the success of others. (from exam report)

(ii) Each phone call may not be equally successful as friends may have different sleeping and phone answering habits, hence the first assumption might not hold in this context.

(iii)

$R \sim B(8, 0.7)$

$P(R \ge 6) = 1 - P(R \le 5) = 0.552$

8. [JJC/2011/Promo/Q11]

In a college with large student population, a proportion own an iPhone. A random sample of n students is taken from the population. The random variable X denotes the number of students in the sample who own an iPhone.

(i) State, in context, two assumptions needed for X to be well modelled by a binomial distribution.

Assume now that X has the distribution $\mathrm{B}(n, p)$.

(ii) Write down an expression for $\mathrm{P}(X = 1)$ in terms of n and p.

(iii) Given that $n = 10$, $\mathrm{P}(X = 1) = 0.2$ and $0.2 < p < 1$, find the value of p.

Given that $3\,\mathrm{E}(X) = 4\,\mathrm{Var}(X)$, show that $p = \frac{1}{4}$.

(iv) Given that the standard deviation of X is $\frac{3}{2}$, find the value of n.

(v) The mean and standard deviation of X are denoted by μ and σ respectively. Find $\mathrm{P}(\mu \le X \le \mu + \sigma)$.

[Ans: (iii) 0.242 (iv) 12 (v) 0.452]

Solution:

8 **(i)** The probability of a student owing an iPhone is constant for every student.
A student owning an iPhone is independent of any other student owning an iPhone

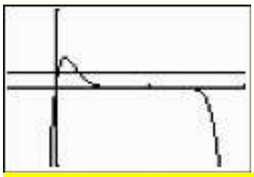
(ii) $\mathrm{P}(X = 1) = \binom{n}{1} p^1 (1-p)^{n-1} = \underline{\underline{np(1-p)^{n-1}}}$

(iii)

$$\mathrm{P}(X = 1) = 0.2$$
$$10p(1-p)^9 = 0.2$$

By GC,

$p = \underline{0.242}$ or 0.0252 (rej $\because 0.2 < p < 1$)



Window: x from –1 to 2
y from –1 to 1

$$3\,\mathrm{E}(X) = 4\,\mathrm{Var}(X)$$
$$3np = 4np(1-p)$$
$$3 = 4(1-p) \quad \text{(Dividing both sides by } np \text{ since } np \neq 0\text{)}$$
$$1 - p = \frac{3}{4}$$
$$p = \frac{1}{4} \text{ (Shown)}$$

(iv) $\sqrt{np(1-p)} = \frac{3}{2}$

$$np(1-p) = \frac{9}{4}$$

$$n\left(\frac{1}{4}\right)\left(\frac{3}{4}\right) = \frac{9}{4}$$
$$n = \underline{\underline{12}}$$

(v) $\mu = 12\left(\frac{1}{4}\right) = 3$, $\mu+\sigma = 3 + \frac{3}{2} = 4\tfrac{1}{2}$

$$\begin{aligned}\mathrm{P}\left(\mu \le X \le \mu+\sigma\right) &= \mathrm{P}\left(3 \le X \le 4\tfrac{1}{2}\right)\\ &= \mathrm{P}(X = 3) + \mathrm{P}(X = 4)\\ &= 0.25810 + 0.19358\\ &= \underline{\underline{0.452}}\end{aligned}$$

Summary

Binomial Distribution	
Conditions	(i) The experiment consists of n repeated trials. (ii) Each trial has two possible outcomes, namely a 'success' or a 'failure'. (iii) The probability of a success, denoted by p, is the same in each trial. (iv) The repeated trials are independent.
Define random variable	Let X be the random variable "number of trials with a success out of n trials". $X \sim \text{B}(n, p)$, $x = 0, 1, 2, \ldots, n$
Probability Distribution	$X \sim \text{B}(n, p)$ n : number of independent trials p : probability of success for each trial $\text{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, 2, \ldots, n$ (finite)
Expectation	$\text{E}(X) = np$
Variance	$\text{Var}(X) = np(1-p)$ Standard deviation = $\sqrt{Var(x)}$
Mode	Mode: the value of x that gives the highest probability. Use of GC to find mode: 1. Press **Y_1 =** 2. Press **2nd DISTR** to select **A:binompdf (** 3. Key in binompdf (n, p, x) 4. Press **2nd TABLE** 5. Scroll down the List Y_1 and select the value of x with the greatest probability

Checklist

I am able to:

- ☐ know the conditions under which a binomial distribution can be used.
- ☐ give examples of random variables which follow a binomial distribution.
- ☐ comment on the appropriateness of using Binomial distribution for a random variable.
- ☐ know the formula for $\text{P}(X = x)$ for a binomial random variable.
- ☐ find Binomial probabilities involving $\text{P}(X = x)$, $\text{P}(X < x)$, $\text{P}(X \le x)$, $\text{P}(X > x)$ and $\text{P}(X \ge x)$ using a GC.
- ☐ find the mean and variance of a binomial random variable.
- ☐ find the mode (the value of the random variable that is most likely to occur) using a GC.
- ☐ find unknowns (n, p) based on given probabilities where GC may be used.