A Level · Edexcel · Maths

3 hours 24 questions

4.4 Choosing Distributions (A Level only)

4.4.1 Modelling with Distributions / 4.4.2 Normal Approximation of Binomial

Easy (6 questions)	/40
Medium (6 questions)	/48
Hard (6 questions)	/54
Very Hard (6 questions)	/61
Total Marks	**/203**

Scan here to return to the course
or visit savemyexams.com

Easy Questions

1 (a) The table below shows five scenarios involving different random variables. Complete the table by placing a cross (×) in the correct box to indicate whether the random variable can be modelled by a binomial distribution, a normal distribution or neither. The first scenario is completed for you.

Scenario	Binomial	Normal	Neither
The digits 1 to 9 are written on individual counters and placed in a bag. A child randomly selects one of the nine counters. The random variable A represents the number that is written on the selected counter.			X
A farmer has many hens. The random variable B represents the mass of a randomly selected hen.			
A fair coin is flipped 100 times. The random variable C represents the number of times it lands on tails.			
A teacher has a 30-minute break for lunch. The random variable D represents the number of emails he receives during his lunch break.			
In a class of 30 students, each student rolls a fair six-sided dice with sides labelled 1 to 6. The random variable represents the number of students who roll a number less than 5.			

(4 marks)

(b) Write down the name of the probability distribution of A, the random variable described in part (a).

(1 mark)

2 (a) In an experiment there are a fixed number of trials and each trial results in a success or failure. Let X be the number of successful trials. Write down the two other conditions that would need to be present to make X follow a binomial distribution.

(2 marks)

(b) A fair spinner has 8 sectors labelled with the numbers 1 through 8. For each of the following cases, give a reason to explain why a binomial distribution would **not** be appropriate for modelling the specified random variable.

(i) The random variable A is the number of times the spinner is spun until it lands on '1' for the first time.

(ii) When the spinner is spun it rotates exactly 115°. The random variable is the number of times the spinner lands on '1' when the spinner is spun 20 times.

(iii) The random variable C is the sector number that the spinner lands on when it is spun once.

(3 marks)

(c) State which one of the random variables defined in part (b) follows a discrete uniform distribution.

(1 mark)

3 (a) For each of the following, state with a reason whether the random variable in question is a discrete random variable or a continuous random variable.

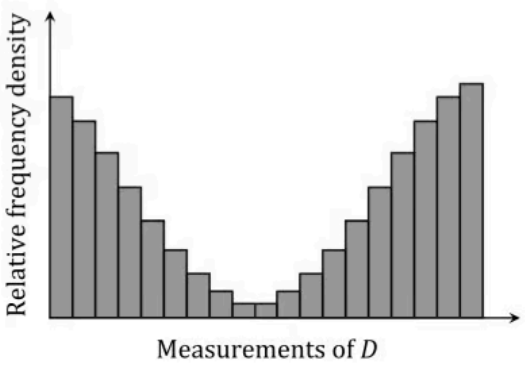
(i) 100 red squirrels from the wild are sampled. The random variable A is the tail length of a randomly selected red squirrel.

(ii) 100 students sit a test which is marked out of 50. The random variable B is the number of marks achieved by a randomly selected student.

(iii) 100 men are in a shoe shop. The random variable C is the shoe size of a randomly selected man.

(3 marks)

(b) The following histogram shows the distribution of results when a large number of measurements of the specified random variable D are made. State with a reason whether a normal distribution would be appropriate for modelling the random variable.



(1 mark)

4 (a) The random variable $X \sim B(n,p)$ can be approximated by $Y \sim N(\mu,\sigma^2)$ when certain conditions are fulfilled.

State the condition for n which is required to use this approximation.

(1 mark)

(b) (i) State the value of p that will give the most accurate estimate.

(ii) Give a reason to support your value.

(2 marks)

(c) For each of the following binomial random variables, X:

- state, with reasons, whether X can be approximated by a normal distribution
- if appropriate, write down the normal approximation to X in the form $N(\mu,\sigma^2)$, giving the values of μ and σ^2.

(i) $X \sim B(6,0.45)$

(ii) $X \sim B(60,0.05)$

(iii) $X \sim B(60,0.45)$

(5 marks)

5 (a) The random variable $X \sim B(100, 0.36)$ is approximated by $Y \sim N(\mu, \sigma^2)$.

Find the value of μ and show that $\sigma = 4.8$.

(3 marks)

(b) Explain why a continuity correction must be incorporated when using this approximation.

(1 mark)

(c) Use continuity corrections to find the value of k in each of the following approximations:

(i) $\mathrm{P}(X \leq 30) \approx \mathrm{P}(Y < k)$

(ii) $\mathrm{P}(X < 30) \approx \mathrm{P}(Y < k)$

(iii) $\mathrm{P}(X \geq 30) \approx \mathrm{P}(Y > k)$

(iv) $\mathrm{P}(X > 30) \approx \mathrm{P}(Y > k)$

(4 marks)

6 (a) $X \sim B(150, 0.6)$ is approximated by a normal distribution $Y \sim N(\mu, \sigma^2)$.

(i) State two reasons why a normal approximation is suitable.

(ii) Find the values of μ and σ.

(4 marks)

(b) (i) Calculate $\mathrm{P}(87 \leq X \leq 101)$.

(ii) Explain why $\mathrm{P}(87 \leq X \leq 101) \approx \mathrm{P}(86.5 < Y < 101.5)$.

(iii) Calculate $\mathrm{P}(86.5 < Y < 101.5)$.

(5 marks)

Medium Questions

1 (a) State the conditions that must be satisfied to be able to model a random variable X with a binomial distribution $\mathrm{B}(n,p)$.

(4 marks)

(b) A fair spinner has 8 sectors labelled with the numbers 1 through 8. For each of the following cases, state with a reason whether or not a binomial distribution would be appropriate for modelling the specified random variable.

(i) The random variable S is the number of the sector that the spinner lands on when it is spun.

(ii) The random variable W is the number of times the spinner is spun until it lands on '7' for the first time.

(iii) The random variable Y is the number of times the spinner lands on a prime number when it is spun twelve times.

(iv) On the first spin, it is a 'win' if the spinner lands on an even number. On subsequent spins it is a 'win' if the spinner lands *either* on the same number as the previous spin *or* on a factor of the number from the previous spin. The random variable L is the number of wins when the spinner is spun ten times.

(4 marks)

(c) For the random variable S defined in (b)(i) above, give the name of the probability distribution that *would* be appropriate for modelling S.

(1 mark)

2 (a) For each of the following, state with a reason whether the random variable in question is a discrete random variable or a continuous random variable.

(i) A student cuts a one-metre length of rope into two pieces at a random point. The random variable A is the length of the shorter of these two pieces.

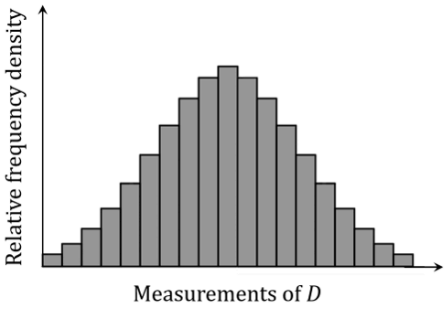
(ii) You ask a sample of students in your school about their preferences for after-school activities. The random variable B is the number of students who say they prefer participating in lawn bowling.

(iii) People are chosen at random from the UK population. The random variable C is the age of a randomly selected person, defined as the age which they turned on their most recent birthday.

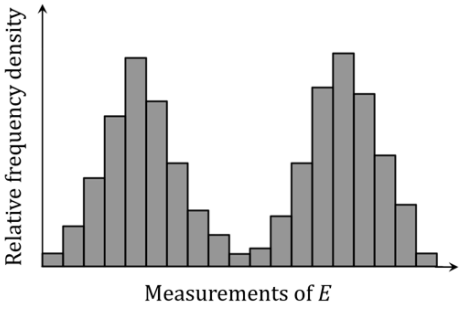
(3 marks)

(b) Each of the following histograms shows the distribution of results when a large number of measurements of the specified random variables – D, E or F – are made. In each case, state with a reason whether a normal distribution would be appropriate for modelling the random variable.

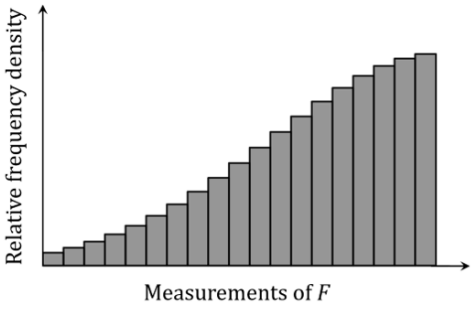
(i)



(ii)



(iii)



(3 marks)

3 For each of the following binomial random variables, X:

- state, with reasons, whether X can be approximated by a normal distribution
- if appropriate, write down the normal approximation to X in the form $N(\mu, \sigma^2)$, giving the values of μ and σ.

(i) $X \sim B(10, 0.5)$

(ii) $X \sim B(80, 0.54)$

(iii) $X \sim B(625, 0.45)$

(iv) $X \sim B(625, 0.45)$

(8 marks)

4 (a) The random variable $W \sim B(500, 0.4)$.

Give two reasons why a normal distribution can be used to approximate W.

(2 marks)

(b) Find, using the normal approximation:

(i) $\mathrm{P}(189 \leq W \leq 211)$

(ii) $\mathrm{P}(W > 220)$

(6 marks)

(c) Using the normal approximation, find the largest value of w such that $\mathrm{P}(W \leq w) < 0.1$.

(3 marks)

5 (a) Write down two conditions under which the normal distribution may be used as an approximation to the binomial distribution $\mathrm{B}(n,p)$.

(2 marks)

(b) On a casino roulette wheel, the probability of the ball landing on a black number is $\frac{9}{19}$.

The wheel is spun 30 times, and the ball lands on a black number X times.

Find $\mathrm{P}(X = 14)$.

(1 mark)

(c) In a separate experiment, the wheel is spun 1000 times and Y, the number of times the ball lands on a black number, is recorded.

(i) Explain why a normal approximation would be appropriate in this case.

(ii) Write down the normal distribution that could be used to approximate the distribution of Y.

(3 marks)

(d) Use the distribution from (c)(ii) to approximate the probability that in at least one half of the 1000 spins the ball lands on a black number.

(3 marks)

6 (a) As part of a marketing promotion, 47% of packets of a particular brand of crisps contain a zombie toy as a prize. A random sample of 100 packets is taken.

Find the exact value of the probability that exactly 49 of the packets contain a prize.

(1 mark)

(b) Write down the normal distribution that could be used to approximate the distribution for the number of the 100 packets that contain a prize.

(1 mark)

(c) By first working out the approximated probability, find the percentage error when using a normal approximation to calculate the probability that exactly 49 of the packets contain a prize. Give your answer correct to two decimal places.

(3 marks)

Hard Questions

1 (a) State the conditions that must be satisfied to be able to model a random variable X with a binomial distribution $\mathrm{B}(n,p)$.

(4 marks)

(b) A fair spinner has 5 sectors labelled with the numbers 1 through 5. The spinner is spun and a fair coin is flipped, and the number the spinner lands on along with the result of the coin flip (heads or tails) are recorded. For each of the following cases, state with a reason whether or not a binomial distribution would be appropriate for modelling the specified random variable.

(i) If the coin lands on heads, then the random variable S is the number of the sector that the spinner lands on times two. Otherwise S is the number of the sector that the spinner lands on plus 10.

(ii) The random variable W is the number of times the spinner is spun and the coin is flipped until an odd number on the spinner occurs together with tails on the coin.

(iii) The random variable Y is the number of times a prime number on the spinner occurs together with heads on the coin, when the spinner is spun and the coin is flipped 21 times.

(iv) Each time the spinner is spun and the coin is flipped, it is a 'win' if a square number on the spinner occurs together with heads on the coin, or it is a 'loss' if a non-square number on the spinner occurs together with tails on the coin. Any other outcome is a 'draw'. The random variable L is the number of losses when the spinner is spun and the coin is flipped twelve times.

(4 marks)

(c) For the random variable S defined in (b)(i) above, give the name of the probability distribution that *would* be appropriate for modelling *S*. Justify your answer.

(2 marks)

2 (a) For each of the following, state with a reason whether the random variable in question is a discrete random variable or a continuous random variable.

(i) A cake recipe calls for a certain amount of flour to be used. The random variable A is the number of cakes that can be made, following the recipe exactly each time, from a bag containing a random amount of flour.

(ii) A student cuts a one-metre length of rope into two pieces at a random point. The random variable B is the difference in length between the two pieces of rope that result.

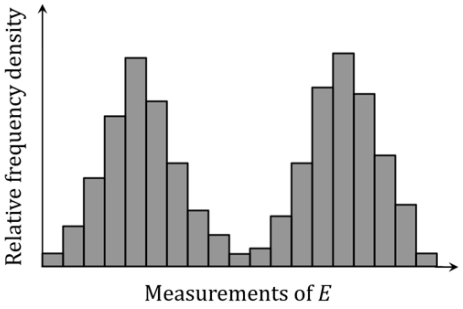
(iii) People are chosen at random from the UK population. The random variable C is the age of a randomly selected person, measured from their date and time of birth.

(3 marks)

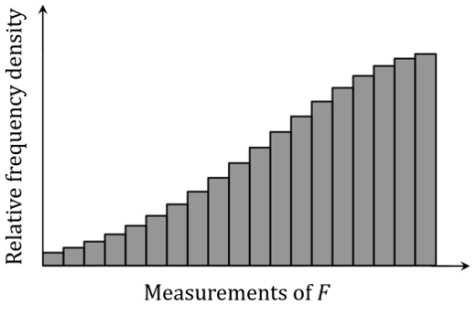
(b) Each of the following histograms shows the distribution of results when a large number of measurements of the random variables D, E or F are made. In each case, state with a reason whether a normal distribution would be appropriate for modelling the random variable. Where a normal model is appropriate, suggest a real-world variable that might show such a distribution.

(i)

(ii)



(iii)



(4 marks)

3 (a) State the conditions that must be met for the distribution of a binomial random variable to be able to be approximated by a normal random variable.

(2 marks)

(b) For each of the following binomial random variables, $X \sim B(n,p)$:

- if X can be approximated by a normal distribution, then write down the normal approximation to X in the form $\mathrm{N}(\mu, \sigma^2)$, giving the values of μ and σ
- if X cannot be approximated by a normal distribution, then give a reason why

(i) $X \sim B(15,0.5)$

(ii) $X \sim B(150,0.56)$

(iii) $X \sim B(1500,0.005)$

(4 marks)

(c) Given that the random variable $Y \sim B(60,0.55)$ is able to be approximated by a normal distribution,

(i) use the appropriate normal distribution to approximate the value of $\mathrm{P}(Y > 60)$

(ii) use your answer to (c)(i) to comment on the limitations of using normal approximations for binomial distributions.

(4 marks)

4 (a) The random variable $W \sim B(1200, 0.6)$.

Give two reasons why a normal distribution can be used to approximate W.

(2 marks)

(b) Find, using the appropriate normal approximation:

(i) $\mathrm{P}(700 < W \leq 791)$

(ii) $\mathrm{P}(W \geq 719)$

(6 marks)

(c) Using the normal approximation, find the largest value of k (where k is an integer) such that $\mathrm{P}((720 - k) < W < (720 + k)) < 0.5$.

(4 marks)

5 (a) Write down two conditions under which the normal distribution may be used as an approximation to the binomial distribution $B(n,p)$.

(2 marks)

(b) On a European-style casino roulette wheel, the probability of the ball landing on a red number is $\frac{18}{37}$.

The wheel is spun 36 times, and the ball lands on a red number X times.

Find $\mathrm{P}(17 < X \leq 18)$.

(1 mark)

(c) In a separate experiment, the wheel is spun 1000 times and Y, the number of times the ball lands on a red number, is recorded.

(i) Explain why a normal approximation would be appropriate in this case.

(ii) Write down the normal distribution that could be used to approximate the distribution of Y.

(3 marks)

(d) Use the distribution from (c)(ii) to approximate the probability that in 1000 spins the ball lands on a red number either less than 482 times or more than 491 times.

(3 marks)

6 (a) Due to a manufacturing irregularity, 41% of Adventure Dude action figures were produced with two left hands. Although not especially rare, and therefore not especially collectible, these so-called 'double left' figures are nonetheless considered to be collector's items by hard-core Adventure Dude fanatics.

A vintage toy shop has obtained 100 Adventure Dude action figures. These may be assumed to represent a random sample.

Find the exact value of the probability that exactly 45 of the 100 figures are 'double left' figures.

(1 mark)

(b) Use an appropriate normal approximation to approximate the probability that exactly 45 of the 100 figures are 'double left' figures.

(3 marks)

(c) Find the percentage error when using your normal approximation from part (b) to estimate the probability that exactly 45 of the 100 figures are 'double left' figures. Give your answer correct to two decimal places.

(2 marks)

Very Hard Questions

1 (a) State the conditions that must be satisfied to be able to model a random variable X with a binomial distribution $B(n,p)$.

(4 marks)

(b) A fair spinner has 5 sectors labelled with the numbers 2, 3, 5, 7 and 11. A fair dice has 6 sides labelled with the numbers 1 through 6. The spinner is spun and the dice is rolled, and the numbers that the spinner and the dice land on are recorded. For each of the following cases, state with a reason whether or not a binomial distribution would be appropriate for modelling the specified random variable.

(i) The random variable S is the square root of the number that the spinner lands on, times the number that the dice lands on.

(ii) The random variable W is the number of times that both the spinner and the dice land on a prime number, when the spinner is spun and the dice is rolled 47 times.

(iii) The random variable Y is the number of times the spinner is spun and the dice is rolled until the number the spinner lands on is a factor of the number the dice lands on.

(iv) On the first spin of the spinner and roll of the dice, it is a 'win' if the number on the spinner is greater than the number on the dice. On subsequent spins of the spinner and rolls of the dice, it is only a 'win' if the number the spinner lands on is higher than both the number the dice lands on *and* all the numbers that the dice has landed on previously. The random variable Z is the number of 'wins' when the spinner is spun and the dice is rolled 24 times.

(4 marks)

(c) For the random variable S defined in (b)(i) above, give the name of the probability distribution that *would* be appropriate for modelling S. You must justify your answer using mathematical reasoning.

(3 marks)

2 (a) For each of the following, state with a reason whether the random variable in question is a discrete random variable or a continuous random variable.

(i) A string collector decides to measure all the pieces of string in his collection. The random variable A is the length of a randomly chosen piece of string from the collection, rounded to the nearest centimetre.

(ii) A random sample of 1000 people is chosen from the US population, and the number of siblings each has is recorded. The random variable B is the mean number of siblings for the 1000 people in the sample.

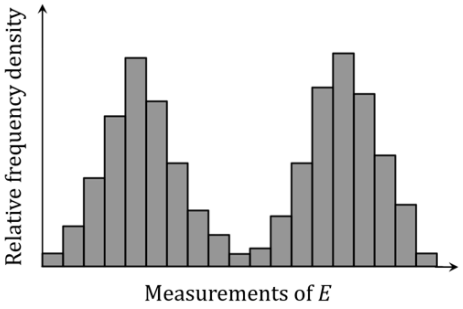
(iii) The masses of individual eggs in the nests of a particular species of bird are measured and recorded. The random variable C is the total mass of the eggs in a randomly chosen nest.

(3 marks)

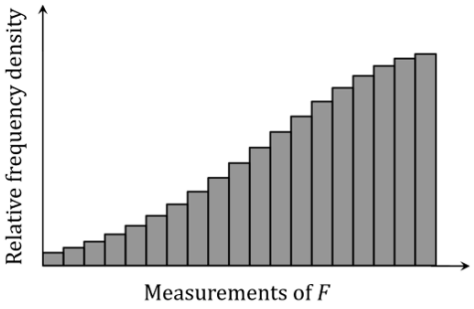
(b) Each of the following histograms shows the distribution of results when a large number of measurements of the specified random variables – D, E or F – are made. In each case, state with a reason whether a normal distribution would be appropriate for modelling the random variable, and suggest a real-world variable that might show such a distribution.

(i)

(ii)



(iii)



(6 marks)

3 (a) State the conditions that must be met for the distribution of a binomial random variable to be able to be approximated by a normal random variable.

(2 marks)

(b) For each of the following binomial random variables, $X \sim B(n,p)$:

if X can be approximated by a normal distribution, then write down the normal approximation to X in the form $N(\mu, \sigma^2)$, giving the values of μ and σ.

if X cannot be approximated by a normal distribution, then give a reason why

(i) $X \sim B(13,0.49)$

(ii) $X \sim B(130,0.53)$

(iii) $X \sim B(1300,0.08)$

(5 marks)

(c) Given that the random variable $Y \sim B(70,0.42)$ is able to be approximated by a normal distribution,

(i) use the appropriate normal distribution to approximate the value of $\mathrm{P}(Y < 0)$

(ii) use your answer to (c)(i) to comment on the limitations of using normal approximations for binomial distributions.

(4 marks)

4 (a) The random variable $W \sim B(980, 0.4)$.

Give two reasons why a normal distribution can be used to approximate W.

(2 marks)

(b) Find, using the appropriate normal approximation:

(i) $\mathrm{P}(386 < W < 398)$

(ii) $\mathrm{P}(W \geq 400)$

(6 marks)

(c) Using the appropriate normal approximation, find the smallest value of k (where $k \in Z$) such that $\mathrm{P}(k < W < (784 - k)) < 0.5$.

(4 marks)

5 (a) On the roulette wheel at the Dunes Oddstacker casino, the probability of the ball landing on a red number is $\frac{6}{13}$, and the probability of the ball landing on a black number is the same. In general, the majority of bettors will lose their bets if the ball lands neither on a red number nor on a black number.

The wheel is spun 50 times, and the ball lands on a number that is either red or black X times.

Find $\mathrm{P}(44 < X < 46)$.

(1 mark)

(b) In a separate experiment, the wheel is spun 1000 times and Y, the number of times the ball lands neither on a red number nor on a black number, is recorded.

Explain why a normal approximation would be appropriate in this case.

(2 marks)

(c) Use an appropriate normal distribution to approximate the probability that in 1000 spins the number of times the ball lands neither on a red number nor on a black number is neither 80 or more nor less than 75.

(4 marks)

6 (a) Due to a production error, 58% of Bobbie Sue dolls were manufactured with proportions that might be seen on a real human being. Plastic surgeons have been buying up these so-called 'realie' dolls, out of a concern that if too many of them are seen by the general public then numbers of people seeking plastic surgery will decrease.

A toy shop has received an order of 100 Bobbie Sue dolls. These may be assumed to represent a random sample.

Find the exact value of the probability that

(i) exactly 58

(ii) exactly 15

of the 100 dolls are 'realie' dolls.

(2 marks)

(b) Use an appropriate normal approximation to approximate the values of the probabilities found in (a)(i) and (ii).

(4 marks)

(c) Find the percentage errors when using your normal approximation from part (b) to estimate the two probabilities in part (a). Use your results to comment on the use of normal distributions to approximate binomial distributions.

(5 marks)