1 The random variable X has the distribution Po(1.3). The random variable Y is defined by $Y = 2X$.

(i) Find the mean and variance of Y. [3]

(ii) Give a reason why the variable Y does not have a Poisson distribution. [1]

2 An engineering test consists of 100 multiple-choice questions. Each question has 5 suggested answers, only one of which is correct. Ashok knows nothing about engineering, but he claims that his general knowledge enables him to get more questions correct than just by guessing. Ashok actually gets 27 answers correct. Use a suitable approximating distribution to test at the 5% significance level whether his claim is justified. [5]

3 Three coats of paint are sprayed onto a surface. The thicknesses, in millimetres, of the three coats have independent distributions $N(0.13, 0.02^2)$, $N(0.14, 0.03^2)$ and $N(0.10, 0.01^2)$. Find the probability that, at a randomly chosen place on the surface, the total thickness of the three coats of paint is less than 0.30 millimetres. [5]

4 The volumes of juice in bottles of Apricola are normally distributed. In a random sample of 8 bottles, the volumes of juice, in millilitres, were found to be as follows.

332 334 330 328 331 332 329 333

(i) Find unbiased estimates of the population mean and variance. [3]

A random sample of 50 bottles of Apricola gave unbiased estimates of 331 millilitres and 4.20 millilitres^2 for the population mean and variance respectively.

(ii) Use this sample of size 50 to calculate a 98% confidence interval for the population mean. [3]

(iii) The manufacturer claims that the mean volume of juice in all bottles is 333 millilitres. State, with a reason, whether your answer to part **(ii)** supports this claim. [1]

5 The management of a factory thinks that the mean time required to complete a particular task is 22 minutes. The times, in minutes, taken by employees to complete this task have a normal distribution with mean μ and standard deviation 3.5. An employee claims that 22 minutes is not long enough for the task. In order to investigate this claim, the times for a random sample of 12 employees are used to test the null hypothesis $\mu = 22$ against the alternative hypothesis $\mu > 22$ at the 5% significance level.

(i) Show that the null hypothesis is rejected in favour of the alternative hypothesis if $\bar{x} > 23.7$ (correct to 3 significant figures), where $\bar{x}$ is the sample mean. [3]

(ii) Find the probability of a Type II error given that the actual mean time is 25.8 minutes. [4]

6 Customers arrive at an enquiry desk at a constant average rate of 1 every 5 minutes.

(i) State one condition for the number of customers arriving in a given period to be modelled by a Poisson distribution. [1]

Assume now that a Poisson distribution is a suitable model.

(ii) Find the probability that exactly 5 customers will arrive during a randomly chosen 30-minute period. [2]

(iii) Find the probability that fewer than 3 customers will arrive during a randomly chosen 12-minute period. [3]

(iv) Find an estimate of the probability that fewer than 30 customers will arrive during a randomly chosen 2-hour period. [4]

7

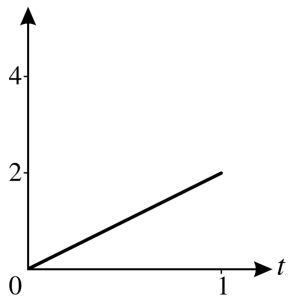


Fig. 1

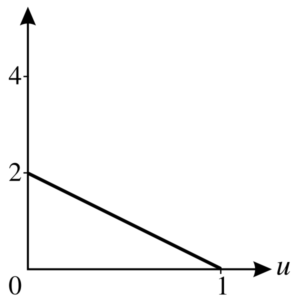


Fig. 2

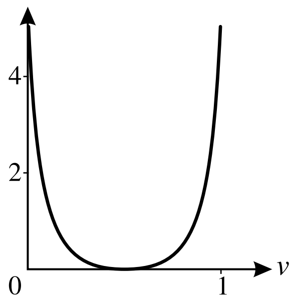


Fig. 3

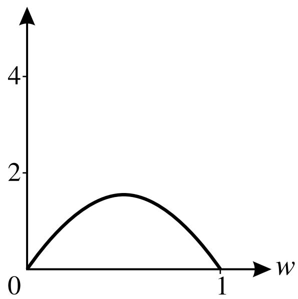


Fig. 4

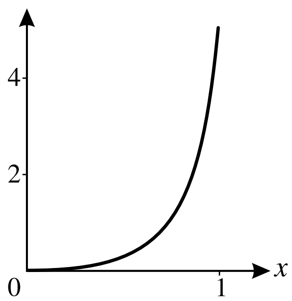


Fig. 5

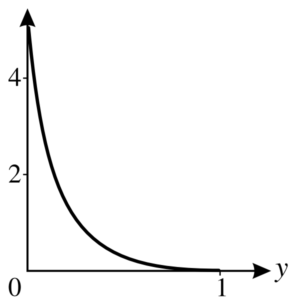


Fig. 6

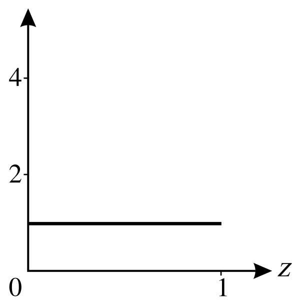


Fig. 7

Each of the random variables T, U, V, W, X, Y and Z takes values between 0 and 1 only. Their probability density functions are shown in Figs 1 to 7 respectively.

(i) **(a)** Which of these variables has the largest median? [1]

(b) Which of these variables has the largest standard deviation? Explain your answer. [2]

(ii) Use Fig. 2 to find $\mathrm{P}(U < 0.5)$. [2]

(iii) The probability density function of X is given by

$$\mathrm{f}(x) = \begin{cases} ax^n & 0 \leqslant x \leqslant 1, \\ 0 & \text{otherwise,} \end{cases}$$

where a and n are positive constants.

(a) Show that $a = n + 1$. [3]

(b) Given that $\mathrm{E}(X) = \frac{5}{6}$, find a and n. [4]

1 The random variable X has the distribution Po(1.3). The random variable Y is defined by $Y = 2X$.

(i) Find the mean and variance of Y. [3]

(ii) Give a reason why the variable Y does not have a Poisson distribution. [1]

2 An engineering test consists of 100 multiple-choice questions. Each question has 5 suggested answers, only one of which is correct. Ashok knows nothing about engineering, but he claims that his general knowledge enables him to get more questions correct than just by guessing. Ashok actually gets 27 answers correct. Use a suitable approximating distribution to test at the 5% significance level whether his claim is justified. [5]

3 Three coats of paint are sprayed onto a surface. The thicknesses, in millimetres, of the three coats have independent distributions $N(0.13, 0.02^2)$, $N(0.14, 0.03^2)$ and $N(0.10, 0.01^2)$. Find the probability that, at a randomly chosen place on the surface, the total thickness of the three coats of paint is less than 0.30 millimetres. [5]

4 The volumes of juice in bottles of Apricola are normally distributed. In a random sample of 8 bottles, the volumes of juice, in millilitres, were found to be as follows.

332 334 330 328 331 332 329 333

(i) Find unbiased estimates of the population mean and variance. [3]

A random sample of 50 bottles of Apricola gave unbiased estimates of 331 millilitres and 4.20 millilitres^2 for the population mean and variance respectively.

(ii) Use this sample of size 50 to calculate a 98% confidence interval for the population mean. [3]

(iii) The manufacturer claims that the mean volume of juice in all bottles is 333 millilitres. State, with a reason, whether your answer to part **(ii)** supports this claim. [1]

5 The management of a factory thinks that the mean time required to complete a particular task is 22 minutes. The times, in minutes, taken by employees to complete this task have a normal distribution with mean μ and standard deviation 3.5. An employee claims that 22 minutes is not long enough for the task. In order to investigate this claim, the times for a random sample of 12 employees are used to test the null hypothesis $\mu = 22$ against the alternative hypothesis $\mu > 22$ at the 5% significance level.

(i) Show that the null hypothesis is rejected in favour of the alternative hypothesis if $\bar{x} > 23.7$ (correct to 3 significant figures), where $\bar{x}$ is the sample mean. [3]

(ii) Find the probability of a Type II error given that the actual mean time is 25.8 minutes. [4]

6 Customers arrive at an enquiry desk at a constant average rate of 1 every 5 minutes.

(i) State one condition for the number of customers arriving in a given period to be modelled by a Poisson distribution. [1]

Assume now that a Poisson distribution is a suitable model.

(ii) Find the probability that exactly 5 customers will arrive during a randomly chosen 30-minute period. [2]

(iii) Find the probability that fewer than 3 customers will arrive during a randomly chosen 12-minute period. [3]

(iv) Find an estimate of the probability that fewer than 30 customers will arrive during a randomly chosen 2-hour period. [4]

7

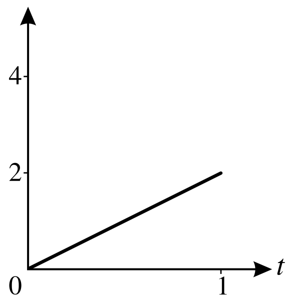


Fig. 1

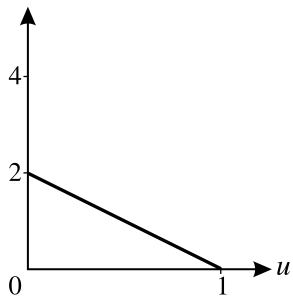


Fig. 2

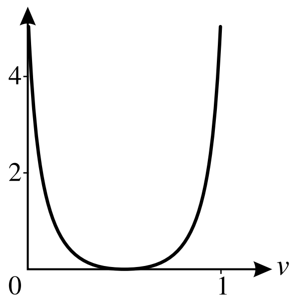


Fig. 3

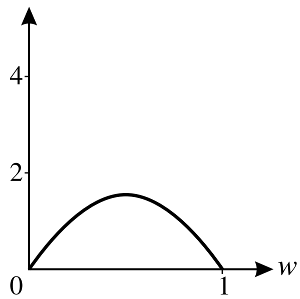


Fig. 4

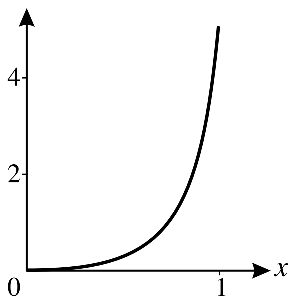


Fig. 5

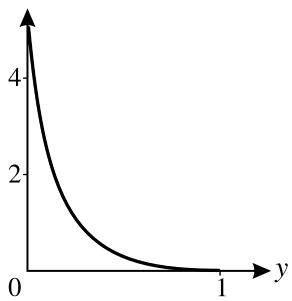


Fig. 6

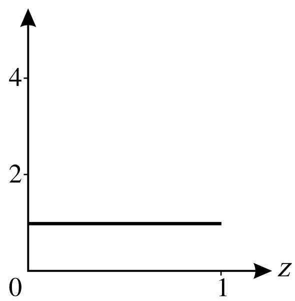


Fig. 7

Each of the random variables T, U, V, W, X, Y and Z takes values between 0 and 1 only. Their probability density functions are shown in Figs 1 to 7 respectively.

(i) **(a)** Which of these variables has the largest median? [1]

(b) Which of these variables has the largest standard deviation? Explain your answer. [2]

(ii) Use Fig. 2 to find $\mathrm{P}(U < 0.5)$. [2]

(iii) The probability density function of X is given by

$$\mathrm{f}(x) = \begin{cases} ax^n & 0 \leqslant x \leqslant 1, \\ 0 & \text{otherwise,} \end{cases}$$

where a and n are positive constants.

(a) Show that $a = n + 1$. [3]

(b) Given that $\mathrm{E}(X) = \frac{5}{6}$, find a and n. [4]

1 Test scores, X, have mean 54 and variance 144. The scores are scaled using the formula $Y = a + bX$, where a and b are constants and $b > 0$. The scaled scores, Y, have mean 50 and variance 100. Find the values of a and b. [4]

2 35% of a random sample of n students walk to college. This result is used to construct an approximate 98% confidence interval for the population proportion of students who walk to college. Given that the width of this confidence interval is 0.157, correct to 3 significant figures, find n. [5]

3 Jack has to choose a random sample of 8 people from the 750 members of a sports club.

(i) Explain fully how he can use random numbers to choose the sample. [3]

Jack asks each person in the sample how much they spent last week in the club café. The results, in dollars, were as follows.

15 25 30 8 12 18 27 25

(ii) Find unbiased estimates of the population mean and variance. [3]

(iii) Explain briefly what is meant by 'population' in this question. [1]

4 The random variable X has probability density function given by

$$f(x) = \begin{cases} ke^{-x} & 0 \leqslant x \leqslant 1, \\ 0 & \text{otherwise.} \end{cases}$$

(i) Show that $k = \dfrac{e}{e - 1}$. [3]

(ii) Find $E(X)$ in terms of e. [4]

5 Records show that the distance driven by a bus driver in a week is normally distributed with mean 1150 km and standard deviation 105 km. New driving regulations are introduced and in the next 20 weeks he drives a total of 21 800 km.

(i) Stating any assumption(s), test, at the 1% significance level, whether his mean weekly driving distance has decreased. [6]

(ii) A similar test at the 1% significance level was carried out using the data from another 20 weeks. State the probability of a Type I error and describe what is meant by a Type I error in this context. [2]

6 Ranjit goes to mathematics lectures and physics lectures. The length, in minutes, of a mathematics lecture is modelled by the variable X with distribution $N(36, 3.5^2)$. The length, in minutes, of a physics lecture is modelled by the independent variable Y with distribution $N(55, 5.2^2)$.

(i) Find the probability that the total length of two mathematics lectures and one physics lecture is less than 140 minutes. [4]

(ii) Ranjit calculates how long he will need to spend revising the content of each lecture as follows. Each minute of a mathematics lecture requires 1 minute of revision and each minute of a physics lecture requires $1\frac{1}{2}$ minutes of revision. Find the probability that the total revision time required for one mathematics lecture and one physics lecture is more than 100 minutes. [4]

7 The numbers of men and women who visit a clinic each hour are independent Poisson variables with means 2.4 and 2.8 respectively.

(i) Find the probability that, in a half-hour period,

(a) 2 or more men and 1 or more women will visit the clinic, [4]

(b) a total of 3 or more people will visit the clinic. [3]

(ii) Find the probability that, in a 10-hour period, a total of more than 60 people will visit the clinic. [4]

1. A large batch of tomatoes is delivered to a packing station. The weights of these tomatoes may be assumed to be independent and normally distributed with mean 106 grams and standard deviation 8 grams.

(a) Find the probability that the weight of a randomly selected tomato exceeds 120 grams. [3]

(b) A pack contains 10 randomly selected tomatoes. Find the probability that the total weight of these 10 tomatoes is less than 1 kilogram. [6]

2. The number of computer breakdowns per day at a large office may be assumed to follow a Poisson distribution with mean μ. The IT Manager believes that the value of μ should be 1·5 but he decides to check this. He therefore defines the following hypotheses.

$$H_0{:}\mu = 1{\cdot}5;\quad H_1{:}\mu \neq 1{\cdot}5$$

(a) For one test, he decides to count the number of breakdowns, x, in a 10-day period and to define the critical region as $x \leqslant 9$ or $x \geqslant 22$. Find the significance level of this test. [5]

(b) For another test, he decides to count the number of breakdowns occurring during a 100-day period. Given that 170 breakdowns occur, calculate the approximate p-value and state your conclusion. [7]

3. When a weighing machine is used to weigh an object, the reading obtained, in grams, is a normally distributed random variable with mean equal to the actual weight of the object and standard deviation 0·2. Successive weighings are independent.

(a) When an object A was weighed three times, the readings obtained were 11·5, 11·7 and 11·6. Calculate a 95% confidence interval for the weight of object A. [5]

(b) Before an object B was weighed, Graham believed that it would weigh 12 grams but Jim believed that it would weigh more than that.

(i) State suitable hypotheses to test their beliefs.

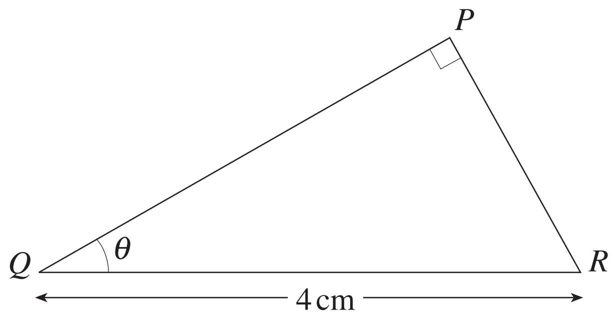
(ii) When the object B was weighed four times, the readings obtained were 12·1, 12·2, 12·4 and 12·1. Calculate the p-value of the four readings and state your conclusion. [7]

(c) Calculate a 90% confidence interval for the difference between the weights of objects A and B. [5]

4. *(a)* The random variable X has the binomial distribution $B(n, p)$. Given that $E(X) = 3$ and $E(X^2) = 11{\cdot}1$, find the values of n and p. [6]

(b) The independent random variable Y has the binomial distribution $B(15, 0{\cdot}4)$ and $U = XY$. Find the mean and variance of U. [8]

5.



The above diagram shows a right-angled triangle in which the hypotenuse $QR = 4$ cm and $P\hat{Q}R = \theta$ radians, where θ is a continuous random variable uniformly distributed between 0 and $\dfrac{\pi}{4}$.

(a) Show that the area, A cm^2, of the triangle PQR is given by

$$A = 4\sin 2\theta. \qquad [1]$$

(b) Calculate $P(A \leqslant 2)$. [5]

(c) Determine $E(A)$. [4]

6. Ann and Brenda buy a packet of seeds which states that, on average, 75% of the seeds will germinate. They believe, however, that the germination rate is less than this so they plant a certain number of seeds and count how many germinate.

(a) State suitable hypotheses. [1]

(b) Ann plants 50 seeds and decides to reject the statement on the packet if less than 30 germinate.

(i) Find the significance level of this procedure.

(ii) Find the probability of accepting the statement on the packet if the actual germination rate is 50%. [6]

(c) Brenda plants 200 seeds and finds that 140 germinate. Find the approximate p-value of this result and state your conclusion in context. [6]

1. Jamie is given a coin and he wishes to estimate p, the probability of its landing 'heads' when tossed. He therefore tosses the coin 250 times and obtains 140 'heads'.

(a) Calculate an unbiased estimate of p. [1]

(b) Calculate an approximate 99% confidence interval for p. [5]

(c) State, with a reason, whether or not your results suggest that the coin is biased. [1]

2. A grower sells melons and claims that their mean weight is 1 kg. A shopkeeper buys a large number of these melons and he believes that the mean weight is less than 1 kg. In order to investigate his belief, he selects a random sample of 100 melons and he determines the weight, x kg, of each one. He produces the following summary statistics.

$$\sum x = 99 \cdot 6, \ \sum x^2 = 99 \cdot 24$$

(a) State suitable hypotheses to test the shopkeeper's belief. [1]

(b) Calculate the p-value of these results and state your conclusion. [7]

(c) State what the Central Limit Theorem enabled you to assume in your solution to *(b)*. [1]

3. A bag contains six coins, of which one is a 20p coin, three are 10p coins and two are 5p coins. A random sample of three of these coins is taken **without replacement**. Determine the sampling distribution of the total value of the coins in the sample. [9]

4. A firm specialises in the manufacture of accurate watches. As part of a quality control procedure, 12 watches were selected and the number of seconds gained over a period of a week was recorded for each watch. The results were as follows.

$$6, \quad 8, \quad -5, \quad 3, \quad 4, \quad -2, \quad 6, \quad 5, \quad -8, \quad 1, \quad -4, \quad 4$$

You may assume that this is a random sample from the $N(\mu, \sigma^2)$ distribution.

(a) Calculate unbiased estimates of μ and σ^2. [4]

(b) Calculate a 95% confidence interval for μ. [5]

(c) The firm claims that 'on average, this type of watch is accurate to within 5 seconds after a week'. State, with a reason, whether or not your answer to *(b)* supports this claim. [1]

5. The director of a large chain of hotels wishes to compare the mean lifetimes of two types of electric light bulbs, Type A and Type B. He therefore determines the lifetime, x thousand hours, of each of 75 randomly selected bulbs of Type A and the lifetime, y thousand hours, of each of 75 randomly selected bulbs of Type B. He obtains the following results.

$$\sum x = 82{\cdot}6, \quad \sum x^2 = 92{\cdot}4, \quad \sum y = 86{\cdot}3, \quad \sum y^2 = 102{\cdot}2$$

(a) State suitable hypotheses for a two-sided test. [1]

(b) Calculate the p-value of these results. [10]

(c) Interpret your p-value in context. [1]

6. The probability distribution of the discrete random variable X is given in the following table, where $0 < \theta < \frac{1}{3}$.

x	-1	0	1
$P(X = x)$	θ	2θ	$1 - 3\theta$

(a) Obtain an expression for $E(X)$ and show that

$$\text{Var}(X) = 2\theta(3 - 8\theta).$$ [3]

In order to estimate θ, a random sample of n observations of X is taken.

(b) The mean of the observations in the sample is denoted by $\overline{X}$. Show that

$$U = \frac{1 - \overline{X}}{4}$$

is an unbiased estimator for θ and obtain an expression for the variance of U. [4]

(c) The number of observations in the sample equal to zero is denoted by N.
Show that

$$V = \frac{N}{2n}$$

is an unbiased estimator for θ and obtain an expression for the variance of V. [5]

(d) Show that

$$\text{Var}(V) - \text{Var}(U) > 0$$

State, with a reason, which is the better estimator, U or V. [3]

TURN OVER

7. The length, y metres, of an elastic string and its tension, x Newtons, are related by an equation of the form $y = \alpha + \beta x$. In order to estimate the values of α and β, the values of y were measured for six different values of x. The following results were obtained.

x	10	20	30	40	50	60
y	2·02	2·23	2·39	2·56	2·77	2·95

The values of x are exact but the values of y are subject to independent normally distributed measurement errors with mean zero and standard deviation 0·02 metres.

(a) Calculate least squares estimates for α and β. [8]

(b) Determine a 90% confidence interval for α. [5]

1. The time taken by Alan to drive to work may be assumed to be normally distributed with mean 28 minutes and standard deviation 2 minutes.

(a) Find the probability that,

(i) on a particular day, he takes more than 30 minutes to drive to work,

(ii) in a particular 5-day week, the mean time taken to drive to work is less than 30 minutes. [8]

(b) The time taken by Brenda to drive to work may be assumed to be normally distributed with mean 25 minutes and standard deviation 3 minutes. Find the probability that, on a particular day, Brenda takes longer to drive to work than Alan. [5]

2. The random variable X has a normal distribution with unknown mean μ and standard deviation 0·5.

(a) A random sample of 60 values of X was taken and it was found that $\sum x = 1290$. Calculate a 95% confidence interval for μ, giving the end-points of your interval correct to two decimal places. [5]

(b) Determine the minimum sample size required for the width of a 95% confidence interval for μ to be less than 0·1. [4]

3. A factory manufactures screws and packs them in large bags. The number of defective screws in a bag can be modelled by a Poisson distribution whose mean is known to have been 0·5. However, new equipment has been installed which, it is hoped, will decrease this mean. The Quality Controller plans to take samples of bags to investigate whether or not there is a reduction in the mean.

(a) State suitable hypotheses. [1]

(b) He takes a random sample of 30 bags and finds that they contain a total of 12 defective screws. Calculate the p-value and state your conclusion. [4]

(c) He then takes a random sample of 200 bags and finds that they contain a total of 80 defective screws. Calculate an approximate p-value and state your conclusion. [6]

4. A zoologist believes that the mean weights of the adult males and females of a certain species of animal are equal. In order to test this belief, she weighs random samples of males and females with the following results.

Weights of males (kg)	14·3	15·8	13·9	13·4	14·5	15·1	13·6	14·2
Weights of females (kg)	13·2	14·8	13·7	14·7	15·0	13·1	13·5	

You may assume that these are random samples from normal populations with a common standard deviation of 0·5.

(a) State suitable hypotheses for carrying out a two-sided test. [1]

(b) Determine the p-value of these results and state whether or not the zoologist's belief is supported at the 5% level of significance. [9]

5. *(a)* The continuous random variable U is uniformly distributed on [a, b].
Write down the probability density function of U and **hence** show that

$$E\left(U^2\right) = \frac{a^2 + ab + b^2}{3} \, .$$ [5]

(b) A piece of string of length 12 cm is cut at a random point. The length of the resulting shorter piece is denoted by X cm and the length of the longer piece by Y cm. You may assume that X is uniformly distributed on [0, 6].

(i) Find the mean and variance of X.

(ii) Express Y in terms of X and hence find the mean of XY.

(iii) Suppose now that 100 pieces of string of length 12 cm are each cut at a random point. Use the Central Limit Theorem to find, approximately, the probability that the total length of the 100 shorter pieces is greater than 280 cm. [11]

6. David is given a biased coin and is told that the probability of obtaining a head when the coin is tossed is either 0·3 or 0·6. To determine which, he defines the following hypotheses.

$$H_0 : p = 0{\cdot}3; \quad H_1 : p = 0{\cdot}6.$$

(a) He tosses the coin 20 times and denotes the number of heads obtained by x.
He will accept H_1 if $x \geqslant 9$ and he will accept H_0 if $x \leqslant 8$.
Calculate the probability of

(i) accepting H_1 when H_0 is true,

(ii) accepting H_0 when H_1 is true. [8]

(b) He now tosses the coin 80 times and denotes the number of heads obtained by y.
He will accept H_1 if $y \geqslant 36$ and he will accept H_0 if $y \leqslant 35$.
Using a normal approximation, calculate the probability of

(i) accepting H_1 when H_0 is true,

(ii) accepting H_0 when H_1 is true. [8]