1. A manager in a sweet factory believes that the machines are working incorrectly and the proportion p of underweight bags of sweets is more than 5%. He decides to test this by randomly selecting a sample of 5 bags and recording the number x that are underweight. The manager sets up the hypotheses H_0: $p = 0.05$ and H_1: $p > 0.05$ and rejects the null hypothesis if $x > 1$.

(a) Find the size of the test.

(2)

(b) Show that the power function of the test is

$$1 - (1 - p)^4(1 + 4p)$$

(3)

The manager goes on holiday and his deputy checks the production by randomly selecting a sample of 10 bags of sweets. He rejects the hypothesis that $p = 0.05$ if more than 2 underweight bags are found in the sample.

(c) Find the probability of a Type I error using the deputy's test.

(2)

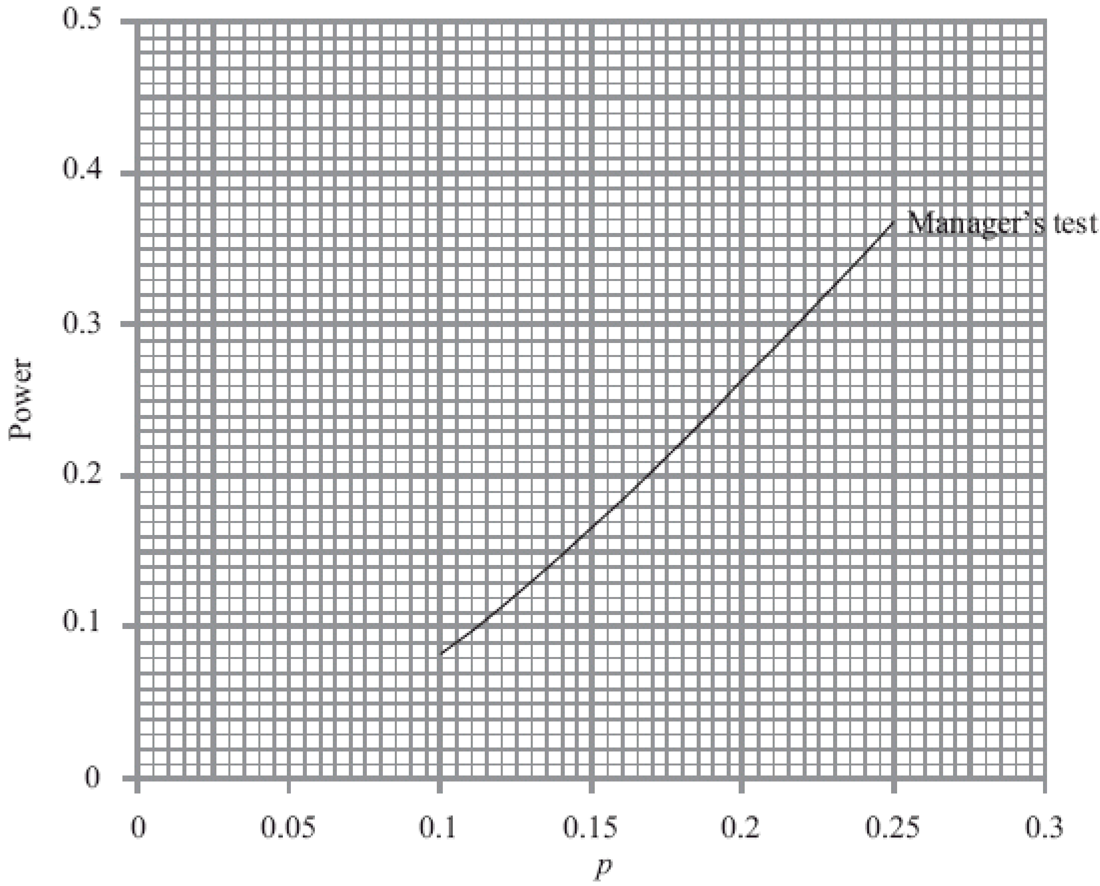
The table below gives some values, to 2 decimal places, of the power function for the deputy's test.

p	0.10	0.15	0.20	0.25
Power	0.07	s	0.32	0.47

(d) Find the value of s.

(1)

The graph of the power function for the manager's test is shown the diagram below.



(e) On the same axes, draw the graph of the power function for the deputy's test.

(1)

(f) (i) State the value of *p* where these graphs intersect.

(ii) Compare the effectiveness of the two tests if *p* is greater than this value.

(2)

The deputy suggests that they should use his sampling method rather than the manager's.

(g) Give a reason why the manager might not agree to this change.

(1)
(Total 12 marks)

2. Define, in terms of H_0 and/or H_1,

(a) the size of a hypothesis test,

(1)

(b) the power of a hypothesis test.

(1)

The probability of getting a head when a coin is tossed is denoted by p.

This coin is tossed 12 times in order to test the hypotheses H_0: $p = 0.5$ against H_1: $p \neq 0.5$, using a 5% level of significance.

(c) Find the largest critical region for this test, such that the probability in each tail is less than 2.5%.

(4)

(d) Given that $p = 0.4$

(i) find the probability of a type II error when using this test,

(ii) find the power of this test.

(4)

(e) Suggest two ways in which the power of the test can be increased.

(2)
(Total 12 marks)

3. (a) Define

(i) a Type I error,

(1)

(ii) a Type II error.

(1)

A manufacturer sells socks in boxes of 50.

The mean number of faulty socks per box is 7.5. In order to reduce the number of faulty socks a new machine is tried. A box of socks made on the new machine was tested and the number of faulty socks was 2.

(b) (i) Assuming that the number of faulty socks per box follows a binomial distribution derive a critical region needed to test whether or not there is evidence that the new machine has reduced the mean number of faulty socks per box. Use a 5% significance level.

(2)

(ii) Stating your hypotheses clearly, carry out the test in part (i).

(3)

(c) Find the probability of the Type I error for this test.

(2)

(d) Given that the true mean number of faulty socks per box on the new machine is 5, calculate the probability of a Type II error for this test.

(3)

(e) Explain what would have been the effect of changing the significance level for the test in part (b) to $2\frac{1}{2}$%.

(1)

(Total 13 marks)

1. (a) $X \sim B(5, p)$

Size = P(reject H_0/ p = 0.05)

= P(X > 1/ p = 0.05)

= 1 – 0.9774 M1

= 0.0226 A1 2

Note

M1 for finding P (X>1)
A1 awrt 0.0226

M1 for finding P(Y > 2)
A1 awrt0.0115

(b) Power = 1 – P(0) – P(1) M1

$= 1 - (1 - p)^5 - 5(1 - p)^4 p$ M1

$= 1 - (1 - p)^4 (1 - p + 5p)$

$= 1 - (1 - p)^4 (1 + 4p)$ A1cso 3

Note

M1 for 1-P(0) – P(1)
M1 for $1 - (1 - p)^5 - 5(1 - p)^4 p$
A1 cso

B1 0.18 cao

(c) $Y \sim B(10, p)$

P (Type I error) = P(Y > 2/p = 0.05) M1

= 1 – 0.9885

= 0.0115 A1 2

Note

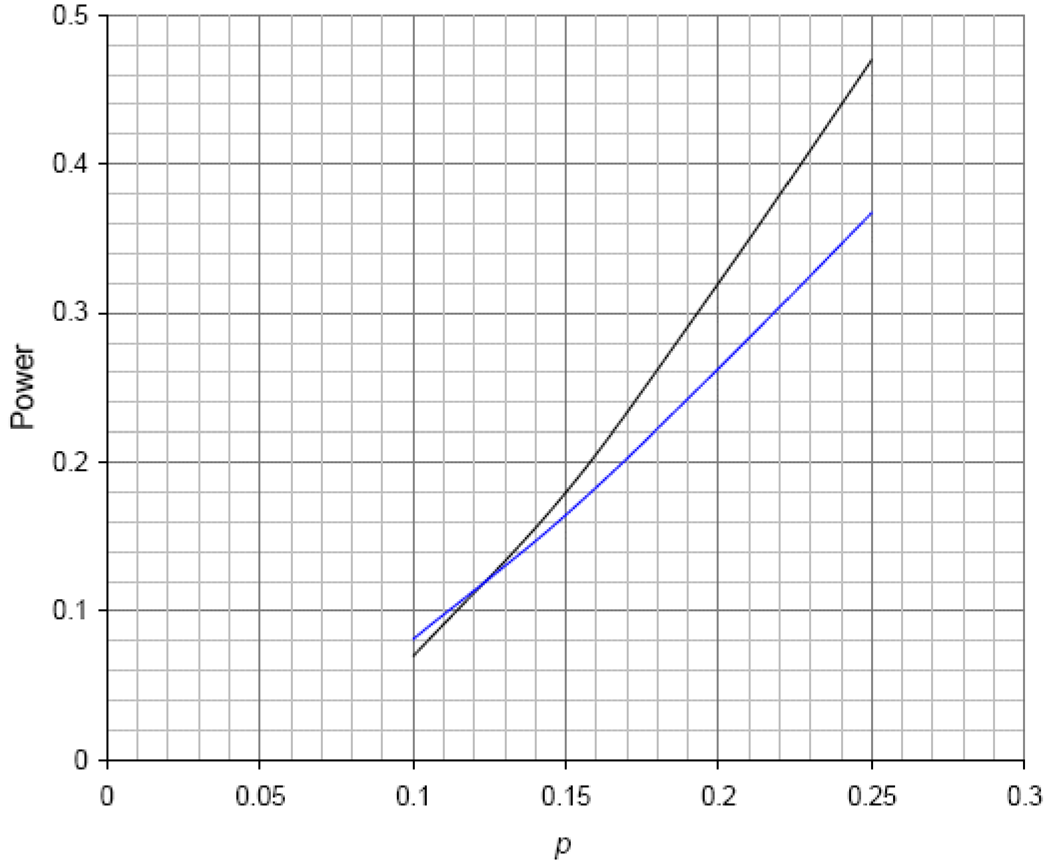
B1 graph. ft their value of s

(d) s = 0.18 B1 1

Note

B1 ft their intersection.
B1 deputy test more powerful o.e.

(e)



B1ft 1

Note

If give first statement they must suggest p unlikely to be above 0.12

(f) (i) intersection 0.12 – 0.13 "their graphs intersection" B1ft

(ii) if $p > 0.12$ the deputy's test is more powerful. B1 2

(g) More powerful for $p < 0.12$ and p unlikely to be above 0.12

Allow it would cost more/take longer/more to sample B1 1

[12]

2. (a) Size is the probability of H0 being rejected when it is in fact true. B1 1

or

P(reject H_0/ H_0 is true) oe

(b)	The power of the test is the probability of rejecting H_0 when H_1 is true.		B1	1
	or			
	P(rejecting H_0/H_1 is true) / P(rejecting H_0/H_0 is false) oe			
(c)	X ~B(12,0.5)		B1	
	$P(X \leq 2) = 0.0193$		M1	
	$P(X \geq 10) = 0.0193$			
	$\therefore$ critical region is $\{X \leq 2 \ \cup \ X \geq 10\}$		A1A1	4
(d)	(i)	P(Type II error) = $P(3 \leq X \leq 9 \mid p = 0.4)$	M1	
		$= P(X \leq 9) - P(X \leq 2)$	M1dep	
		$= 0.9972 - 0.0834$		
		$= 0.9138$	A1	
		<u>Note</u>		
		first M1 for either correct area or follow through from their critical region 2nd M1 dependent on them having the first M1. for finding their area correctly A1 cao		
	(ii)	Power = 1 – 0. 9138		
		= 0.0862	B1 ft	4
		<u>Note</u>		
		B1 follow through from their (i)		
(e)	Increase the sample size		B1	
	Increase the significance level/larger critical region		B1	2

[12]

3.	(a)	(i)	Type I : H_0 rejected when true	B1	
		(ii)	Type II : H_0 accepted when false	B1	2
	(b)	(i)	$P = \frac{7.5}{50} = 0.15$	B1	
			cr $X \leq 3$	B1	2

(ii)	$H_0 : p = 0.15$, $H_1 : p < 0.15$	both	B1	
	$x = 2$ in cr $X \leq 3$ so H_0 is rejected		M1	
	The new machine has reduced the mean number of faulty socks		A1	3
(c)	P(Type I error) = P(X ≤ 3 | p = 0.15) = 0.0460		M1 A1	2
(d)	P(Faulty) = $\frac{5}{50} = 0.1$		B1	
	P(Type II error) = P(X ≥ 4 | p = 0.1) = 1 – 0.2503		M1	
	= 0.7497	or awrt 0.750	A1	3
(e)	Critical region changes to X ≤ 2. H_0 still rejected		B1	1

[13]

1. Many candidates were able to gain full marks in this question and even those who were unable to answer parts (a) to (c) gained several marks in the latter parts.

In part (b) a complete solution was often seen although several candidates wrote Power = 1 – P(0) – P(1) and then concluded that Power = $1 - (1 - p)^4(1 + 4p)$ with no steps in between. This did not gain full marks.

In part (d) several candidates used the power function given in part (b) rather than find the power for the deputy's test using the tables.

2. Many candidates were able to write down the definitions in one form or another in part (a) and part (b). Only a few did not read the question and write it in terms of H_0 and/or H_1.

Part (c) was not well done. Many candidates recognised the correct distribution but were unable to gain the correct critical regions. In part (d)(i) whilst many candidates were able to find the P(Type II error) using their answer to part (c) a sizeable number simply worked out the significance level used. Nearly all candidates knew how to calculate the power of the test correctly.

In part (e) the better candidates had learnt this and gained both marks.

3. Many candidates started well with correct definitions, but subsequently made lots of careless errors with hypotheses, inequalities and critical regions and lost accuracy marks as a result.