Quality of Hypothesis Tests (From Edexcel 6686)

Q1, (Jun 2008, Q6)

A drug is claimed to produce a cure to a certain disease in 35% of people who have the disease. To test this claim a sample of 20 people having this disease is chosen at random and given the drug. If the number of people cured is between 4 and 10 inclusive the claim will be accepted. Otherwise the claim will not be accepted.

(a) Write down suitable hypotheses to carry out this test.

(2)

(b) Find the probability of making a Type I error.

(3)

The table below gives the value of the probability of the Type II error, to 4 decimal places, for different values of p where p is the probability of the drug curing a person with the disease.

P(cure)	0.2	0.3	0.4	0.5
P(Type II error)	0.5880	r	0.8565	s

(c) Calculate the value of r and the value of s.

(3)

(d) Calculate the power of the test for $p = 0.2$ and $p = 0.4$

(2)

(e) Comment, giving your reasons, on the suitability of this test procedure.

(2)

Q2, (Jun 2009, Q3)

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)

Q3, (Jun 2010, Q3)

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 in Figure 1.

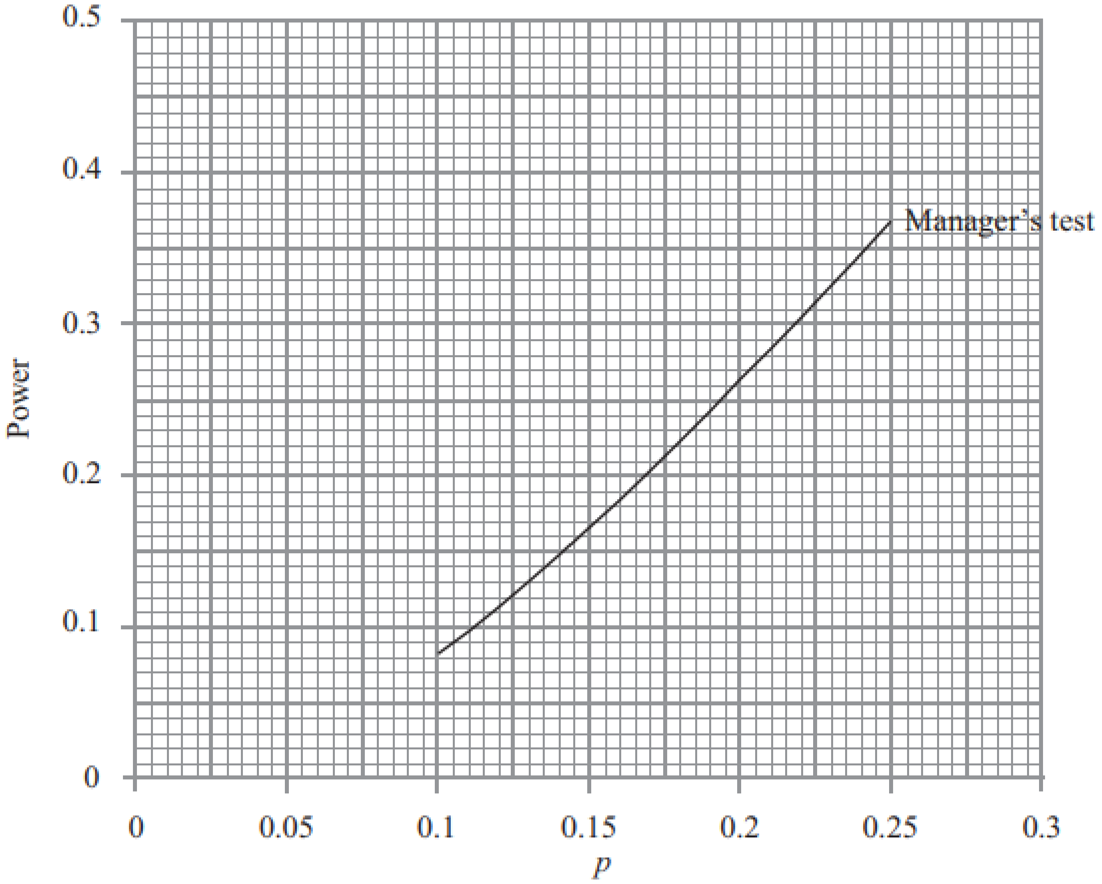


Figure 1

(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)

Q4, (Jun 2011, Q4)

A proportion p of letters sent by a company are incorrectly addressed and if p is thought to be greater than 0.05 then action is taken.
Using H_0: $p = 0.05$ and H_1: $p > 0.05$, a manager from the company takes a random sample of 40 letters and rejects H_0 if the number of incorrectly addressed letters is more than 3.

(a) Find the size of this test.

(2)

(b) Find the probability of a Type II error in the case where p is in fact 0.10

(2)

Table 1 below gives some values, to 2 decimal places, of the power function of this test.

p	0.075	0.100	0.125	0.150	0.175	0.200	0.225
Power	0.35	s	0.75	0.87	0.94	0.97	0.99

Table 1

(c) Write down the value of s.

(1)

A visiting consultant uses an alternative system to test the same hypotheses. A sample of 15 letters is taken. If these are all correctly addressed then H_0 is accepted. If 2 or more are found to have been incorrectly addressed then H_0 is rejected. If only one is found to be incorrectly addressed then a further random sample of 15 is taken and H_0 is rejected if 2 or more are found to have been incorrectly addressed in this second sample, otherwise H_0 is accepted.

(d) Find the size of the test used by the consultant.

(3)

[Question continued on next page]

For your convenience Table 1 is repeated here

p	0.075	0.100	0.125	0.150	0.175	0.200	0.225
Power	0.35	s	0.75	0.87	0.94	0.97	0.99

Table 1

Figure 1 shows the graph of the power function of the test used by the consultant.

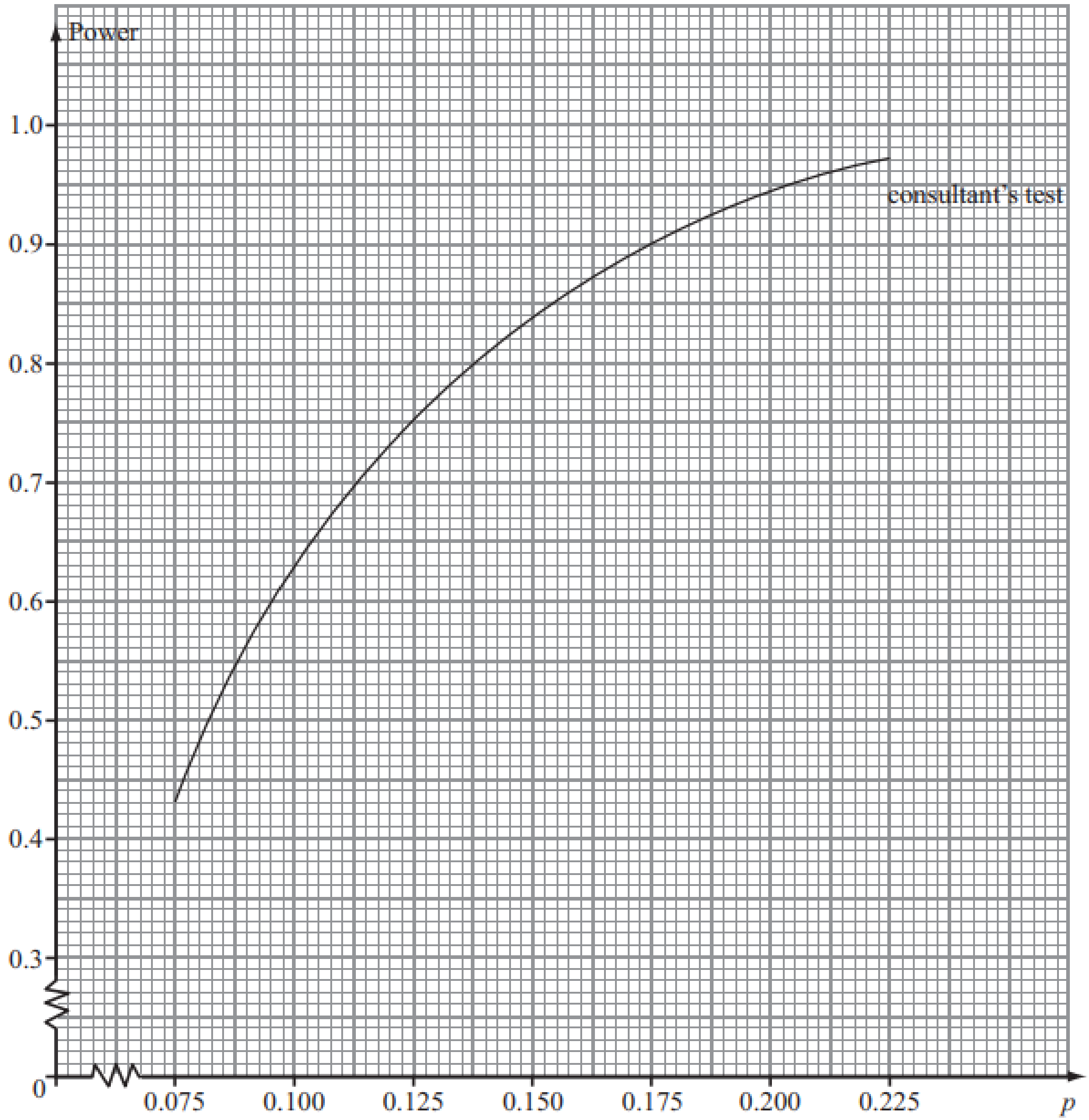


Figure 1

(e) On Figure 1 draw the graph of the power function of the manager's test.

(2)

(f) State, giving your reasons, which test you would recommend.

(2)

Q5, (Jun 2013, Q3)

The number of houses sold per week by a firm of estate agents follows a Poisson distribution with mean 2. The firm believes that the appointment of a new salesman will increase the number of houses sold. The firm tests its belief by recording the number of houses sold, x, in the week following the appointment. The firm sets up the hypotheses $H_0: \lambda = 2$ and $H_1: \lambda > 2$, where λ is the mean number of houses sold per week, and rejects the null hypothesis if $x \geqslant 3$

(a) Find the size of the test.

(2)

(b) Show that the power function for this test is

$$1 - \frac{1}{2}e^{-\lambda}(2 + 2\lambda + \lambda^2)$$

(3)

The table below gives the values of the power function to 2 decimal places.

λ	2.5	3.0	3.5	4.0	5.0	7.0
Power	0.46	r	0.68	s	0.88	0.97

Table 1

(c) Calculate the values of r and s.

(2)

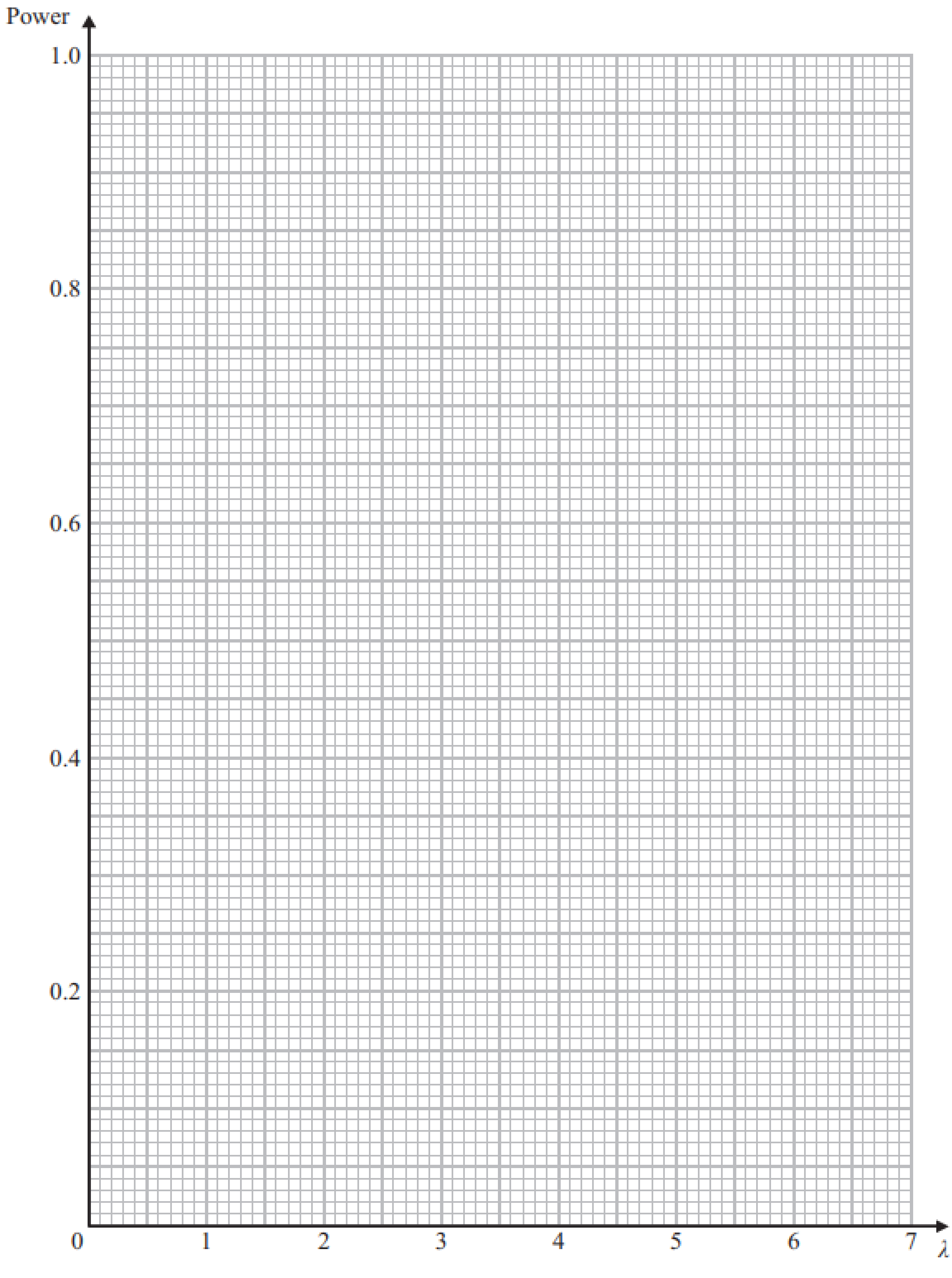
(d) Draw a graph of the power function on the graph paper provided on page 6

(2)

(e) Find the range of values of λ for which the power of this test is greater than 0.6

(1)

[Graph on Next Page]



Power
1.0
0.8
0.6
0.4
0.2
0
1
2
3
4
5
6
7
λ

Q6, (Jun 2013, Q5)

Water is tested at various stages during a purification process by an environmental scientist. A certain organism occurs randomly in the water at a rate of λ every 10 ml. The scientist selects a random sample of 20 ml of water to check whether there is evidence that λ is greater than 1. The criterion the scientist uses for rejecting the hypothesis that $\lambda = 1$ is that there are 4 or more organisms in the sample of 20 ml.

(a) Find the size of the test.

(2)

(b) When $\lambda = 2.5$ find P(Type II error).

(2)

A statistician suggests using an alternative test. The statistician's test involves taking a random sample of 10 ml and rejecting the hypothesis that $\lambda = 1$ if 2 or more organisms are present but accepting the hypothesis if no organisms are in the sample. If only 1 organism is found then a second random sample of 10 ml is taken and the hypothesis is rejected if 2 or more organisms are present, otherwise the hypothesis is accepted.

(c) Show that the power of the statistician's test is given by

$$1 - e^{-\lambda} - \lambda(1 + \lambda)e^{-2\lambda}$$

(4)

Table 1 below gives some values, to 2 decimal places, of the power function of the statistician's test.

λ	1.5	2	2.5	3	3.5	4
Power	0.59	0.75	0.86	r	0.96	0.97

Table 1

(d) Find the value of r.

(1)

[Question Continued on Next Page]

Figure 1 shows a graph of the power function for the scientist's test.

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

(2)

Given that it takes 20 minutes to collect and test a 20 ml sample and 15 minutes to collect and test a 10 ml sample

(f) show that the expected time of the statistician's test is slower than the scientist's test for $\lambda e^{-\lambda} > \frac{1}{3}$

(4)

(g) By considering the times when $\lambda = 1$ and $\lambda = 2$ together with the power curves in part (e) suggest, giving a reason, which test you would use.

(2)

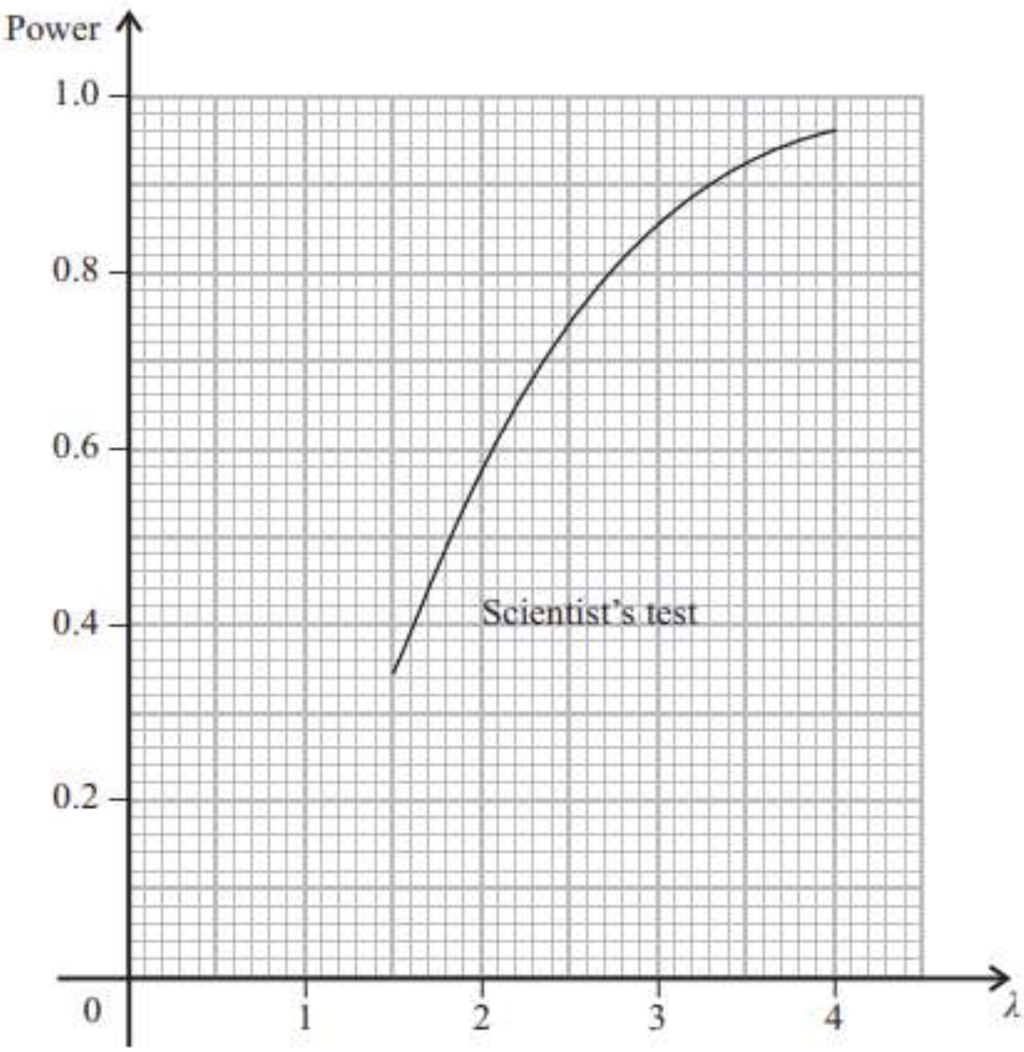


Figure 1

Q7, (Jun 2014, Q2)

(a) Define

(i) a Type I error,

(ii) a Type II error.

(2)

Rolls of material, manufactured by a machine, contain defects at a mean rate of 6 per roll.

The machine is modified. A single roll is selected at random and a test is carried out to see whether or not the mean number of defects per roll has decreased. The significance level is chosen to be as close as possible to 5%.

(b) Calculate the probability of a Type I error for this test.

(3)

(c) Given that the true mean number of defects per roll of material made by the machine is now 4, calculate the probability of a Type II error.

(2)

Q8, (Jun 2015, Q4)

A poultry farm produces eggs which are sold in boxes of 6. The farmer believes that the proportion, p, of eggs that are cracked when they are packed in the boxes is approximately 5%. She decides to test the hypotheses

$$H_0: p = 0.05 \quad \text{against} \quad H_1: p > 0.05$$

To test these hypotheses she randomly selects a box of eggs and rejects H_0 if the box contains 2 or more eggs that are cracked. If the box contains 1 egg that is cracked, she randomly selects a second box of eggs and rejects H_0 if it contains at least 1 egg that is cracked. If the first or the second box contains no cracked eggs, H_0 is immediately accepted and no further boxes are sampled.

(a) Show that the power function of this test is

$$1 - (1 - p)^6 - 6p(1 - p)^{11}$$

(3)

(b) Calculate the size of this test.

(2)

Given that $p = 0.1$

(c) find the expected number of eggs inspected each time this test is carried out, giving your answer correct to 3 significant figures,

(3)

(d) calculate the probability of a Type II error.

(2)

Given that $p = 0.1$ is an unacceptably high value for the farmer,

(e) use your answer from part (d) to comment on the farmer's test.

(1)