OCR Statistics 2

Past Paper Questions

Hypothesis testing

- Statistical hypothesis
- Critical regions
- Types of error.

Edited by K V Kumaran

Web: kumarmaths.weebly.com

Phone: 07961319548

Continuous Hypothesis Testing

- In *any* hypothesis test you will be testing a 'null' hypothesis H_0 against an 'alternative' hypothesis H_1. In S2, your H_0 will only *ever* be one of these three:

  $H_0 : p =$something

  $H_0 : \lambda =$something

  $H_0 : \mu =$something

  Don't deviate from this and you can't go wrong. Notice that it does *not* say $H_0 = p =$something.

- The book gives three approaches to continuous hypothesis testing, but they are all essentially the same. You always compare the probability of what you have seen (under H_0) and anything more extreme, and compare this probability to the significance level. If it is less than the significance level, then you reject H_0 and if it is greater, then you accept H_0.

- Remember we connect the real (X) world to the standard (Z) world using $Z = \frac{X-\mu}{\sigma}$.

- You can do this by:

  1. Calculating the probability of the observed value and anything more extreme and comparing to the significance level.
  2. Finding the critical Z-values for the test and finding the Z-value for the observed event and comparing. (e.g. critical Z-values of 1.96 and -1.96; if observed Z is 1.90 we accept H_0; if observed is -2.11 the reject H_0.)
  3. Finding the critical values for $\bar{X}$. For example critical values might be 17 and 20. If X lies between them then accept H_0; else reject H_0.

- Example: P111 Que 8. Using method 3 from above.

  Let X be the amount of magnesium in a bottle. We are told $X \sim N(\mu, 0.18^2)$. We are taking a sample of size 10, so $\bar{X} \sim N(\mu, \frac{0.18^2}{10})$. Clearly

$$\begin{aligned} H_0 &: \mu = 6.8 \\ H_1 &: \mu \neq 6.8. \end{aligned}$$

We proceed assuming H_0 is correct. Under H_0, $\bar{X} \sim N(6.8, \frac{0.18^2}{10})$. This is a 5% two-tailed test, so we need $2\frac{1}{2}\%$ at each end of our normal distribution. The critical Z values are (by reverse lookup) $Z_{\text{crit}} = \pm 1.960$. To find how these relate to $\bar{X}_{\text{crit}}$ we convert thus

$$\begin{aligned} Z_{\text{crit}} &= \frac{\bar{X}_{\text{crit}} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \\ 1.960 &= \frac{\bar{X}_{\text{crit}} - 6.8}{\sqrt{\frac{0.18^2}{10}}} \\ \text{and} \quad -1.960 &= \frac{\bar{X}_{\text{crit}} - 6.8}{\sqrt{\frac{0.18^2}{10}}} \end{aligned}$$

These solve to $\bar{X}_{\text{crit}} = 6.912$ and $\bar{X}_{\text{crit}} = 6.688$. The observed $\bar{X}$ is 6.92 which lies just outside the acceptance region. We therefore reject H_0 and conclude that the amount of magnesium per bottle is probably *different* to 6.8. [The book is in error in claiming that we conclude it is bigger than 6.8.]

Discrete Hypothesis Testing

- For any test with discrete variables, it is usually best to find the critical value(s) for the test you have set and hence the critical region. The critical value is the first value *at which you would reject* the null hypothesis.

- For example if testing $X \sim B(16, p)$ we may test (at the 5% level)

$$\begin{aligned} \mathrm{H}_0 &: p = \tfrac{5}{6} \\ \mathrm{H}_1 &: p < \tfrac{5}{6}. \end{aligned}$$

  We are looking for the value at the lower end of the distribution (remember the "$<$" acts as an arrow telling us where to look in the distribution). We find $\mathbb{P}(X \leqslant 11) = 0.1134$ and $\mathbb{P}(X \leqslant 10) = 0.0378$. Therefore the critical value is 10. Thus the critical region is $\{0, 1, 2 \ldots 9, 10\}$. So when the result for the experiment is announced, if it lies in the critical region, we reject H_0, else accept H_0.

- Another example: If testing $X \sim B(20, p)$ at the 10% level with

$$\begin{aligned} \mathrm{H}_0 &: p = \tfrac{1}{6} \\ \mathrm{H}_1 &: p \neq \tfrac{1}{6}. \end{aligned}$$

  Here we have a two tailed test with 5% at either end of the distribution. At the lower end we find $\mathbb{P}(X = 0) = 0.0261$ and $\mathbb{P}(X \leqslant 1) = 0.1304$ so the critical value is 0 at the lower end. At the upper end we find $\mathbb{P}(X \leqslant 5) = 0.8982$ and $\mathbb{P}(X \leqslant 6) = 0.9629$. Therefore

$$\begin{aligned} \mathbb{P}(X \geqslant 6) &= 1 - \mathbb{P}(X \leqslant 5) = 1 - 0.8982 = 0.1018 \\ \mathbb{P}(X \geqslant 7) &= 1 - \mathbb{P}(X \leqslant 6) = 1 - 0.9629 = 0.0371 \end{aligned}$$

  So at the upper end we find $X = 7$ to be the critical value. [Remember that at the upper end, the critical value is always one more than the upper of the two values where the gap occurs; here the gap was between 5 and 6 in the tables, so 7 is the critical value.] The critical region is therefore $\{0, 7, 8 \ldots 20\}$.

- There is a Poisson example in the 'Errors in hypothesis testing' section.

Errors In Hypothesis Testing

- A Type I error is made when a true null hypothesis is rejected.
- A Type II error is made when a false null hypothesis is accepted.
- For continuous hypothesis tests, the $\mathbb{P}$(Type I error) is just the significance level of the test. [This fact should be obvious; if not think about it harder!]
- For a Type II error, you must consider something like the example on page 140/1 which is superbly explained. From the original test, you will have discovered the acceptance and the rejection region(s). When you are told the real mean of the distribution and asked to calculate the $\mathbb{P}$(Type II error), you must use the new, real mean and the old standard deviation (with a new normal distribution; e.g. $N(\mu_{\text{new}}, \sigma^2_{\text{old}}/n)$) and work out the probability that the value lies within the old acceptance region. [Again, the book is *very* good on this and my explanation is poor.]
- For discrete hypothesis tests, the $\mathbb{P}$(Type I error) is not merely the stated significance level of the test. The stated value (e.g. 5%) is merely the 'notional' value of the test. The true significance level of the test (and, therefore, the $\mathbb{P}$(Type I error)) is the probability of all the values in the rejection region, given the truth of the null hypothesis.

  For example in a binomial hypothesis test we might have discovered the rejection region was $X \leqslant 3$ and $X \geqslant 16$. If the null hypothesis was "H_0: $p = 0.3$", then the true significance level of the test would be $\mathbb{P}(X \leqslant 3 \text{ or } X \geqslant 16 \mid p = 0.3)$.
- To calculate $\mathbb{P}$(Type II error) you would, given the true value for p (or λ for Poisson), calculate the probability of the *complementary* event. So in the above example, if the true value of p was shown to be 0.4, you would calculate $\mathbb{P}(3 < X < 16 \mid p = 0.4)$.
- Worked example for Poisson: A hypothesis is carried out to test the following:

$$\begin{aligned} \text{H}_0 &: \lambda = 7 \\ \text{H}_1 &: \lambda \neq 7 \\ \alpha &= 10\% \\ &\text{Two tailed test.} \end{aligned}$$

  Under H_0, $X \sim \text{Po}(7)$. We discover the critical values are $X = 2$ and $X = 13$. The critical region is therefore $X \leqslant 2$ and $X \geqslant 13$.

  Therefore $\mathbb{P}$(Type I error) and the true value of the test is therefore

$$\begin{aligned} \mathbb{P}(X \leqslant 2 \text{ or } X \geqslant 13 \mid \lambda = 7) &= \mathbb{P}(X \leqslant 2) + \mathbb{P}(X \geqslant 13) \\ &= \mathbb{P}(X \leqslant 2) + 1 - \mathbb{P}(X \leqslant 12) \\ &= 0.0296 + 1 - 0.9730 \\ &= 0.0566 = 5.66\%. \end{aligned}$$

1.

The height of sweet pea plants grown in a nursery is a random variable. A random sample of 50 plants is measured and is found to have a mean height 1.72 m and variance 0.0967 m^2.

(i) Calculate an unbiased estimate for the population variance of the heights of sweet pea plants. [2]

(ii) Hence test, at the 10% significance level, whether the mean height of sweet pea plants grown by the nursery is 1.8 m, stating your hypotheses clearly. [7]

(Q4, June 2005)

2.

A factory makes chocolates of different types. The proportion of milk chocolates made on any day is denoted by p. It is desired to test the null hypothesis $H_0 : p = 0.8$ against the alternative hypothesis $H_1 : p < 0.8$. The test consists of choosing a random sample of 25 chocolates. H_0 is rejected if the number of milk chocolates is k or fewer. The test is carried out at a significance level as close to 5% as possible.

(i) Use tables to find the value of k, giving the values of any relevant probabilities. [3]

(ii) The test is carried out 20 times, and each time the value of p is 0.8. Each of the tests is independent of all the others. State the expected number of times that the test will result in rejection of the null hypothesis. [2]

(iii) The test is carried out once. If in fact the value of p is 0.6, find the probability of rejecting H_0. [2]

(iv) The test is carried out twice. Each time the value of p is equally likely to be 0.8 or 0.6. Find the probability that exactly one of the two tests results in rejection of the null hypothesis. [4]

(Q6, June 2005)

3.

Alex obtained the actual waist measurements, w inches, of a random sample of 50 pairs of jeans, each of which was labelled as having a 32-inch waist. The results are summarised by

$$n = 50, \quad \Sigma w = 1615.0, \quad \Sigma w^2 = 52\,214.50.$$

Test, at the 0.1% significance level, whether this sample provides evidence that the mean waist measurement of jeans labelled as having 32-inch waists is in fact greater than 32 inches. State your hypotheses clearly. [10]

(Q6, Jan 2006)

4.

The random variable X has the distribution $N(\mu, 8^2)$. The mean of a random sample of 12 observations of X is denoted by $\overline{X}$. A test is carried out at the 1% significance level of the null hypothesis H_0: $\mu = 80$ against the alternative hypothesis H_1: $\mu < 80$. The test is summarised as follows: 'Reject H_0 if $\overline{X} < c$; otherwise do not reject H_0'.

(i) Calculate the value of c. [4]

(ii) Assuming that $\mu = 80$, state whether the conclusion of the test is correct, results in a Type I error, or results in a Type II error if:

(a) $\overline{X} = 74.0$, [1]

(b) $\overline{X} = 75.0$. [1]

(iii) Independent repetitions of the above test, using the value of c found in part **(i)**, suggest that in fact the probability of rejecting the null hypothesis is 0.06. Use this information to calculate the value of μ. [4]

(Q7, Jan 2006)

5.

(i) The random variable R has the distribution $B(6, p)$. A random observation of R is found to be 6. Carry out a 5% significance test of the null hypothesis H_0: $p = 0.45$ against the alternative hypothesis H_1: $p \neq 0.45$, showing all necessary details of your calculation. [4]

(ii) The random variable S has the distribution $B(n, p)$. H_0 and H_1 are as in part **(i)**. A random observation of S is found to be 1. Use tables to find the largest value of n for which H_0 is not rejected. Show the values of any relevant probabilities. [3]

(Q2, June 2006)

6.

A television company believes that the proportion of households that can receive Channel C is 0.35.

(i) In a random sample of 14 households it is found that 2 can receive Channel C. Test, at the 2.5% significance level, whether there is evidence that the proportion of households that can receive Channel C is less than 0.35. [7]

(ii) On another occasion the test is carried out again, with the same hypotheses and significance level as in part **(i)**, but using a new sample, of size n. It is found that no members of the sample can receive Channel C. Find the largest value of n for which the null hypothesis is not rejected. Show all relevant working. [4]

(Q7, Jan 2007)

7.

Three independent researchers, A, B and C, carry out significance tests on the power consumption of a manufacturer's domestic heaters. The power consumption, X watts, is a normally distributed random variable with mean μ and standard deviation 60. Each researcher tests the null hypothesis H_0: $\mu = 4000$ against the alternative hypothesis H_1: $\mu > 4000$.

Researcher A uses a sample of size 50 and a significance level of 5%.

(i) Find the critical region for this test, giving your answer correct to 4 significant figures. [6]

In fact the value of μ is 4020.

(ii) Calculate the probability that Researcher A makes a Type II error. [6]

(iii) Researcher B uses a sample bigger than 50 and a significance level of 5%. Explain whether the probability that Researcher B makes a Type II error is less than, equal to, or greater than your answer to part **(ii)**. [2]

(iv) Researcher C uses a sample of size 50 and a significance level bigger than 5%. Explain whether the probability that Researcher C makes a Type II error is less than, equal to, or greater than your answer to part **(ii)**. [2]

(v) State with a reason whether it is necessary to use the Central Limit Theorem at any point in this question. [2]

(Q7, June 2006)

8.

The quantity, X milligrams per litre, of silicon dioxide in a certain brand of mineral water is a random variable with distribution $N(\mu, 5.6^2)$.

(i) A random sample of 80 observations of X has sample mean 100.7. Test, at the 1% significance level, the null hypothesis $H_0 : \mu = 102$ against the alternative hypothesis $H_1 : \mu \neq 102$. [5]

(ii) The test is redesigned so as to meet the following conditions.

- The hypotheses are $H_0 : \mu = 102$ and $H_1 : \mu < 102$.
- The significance level is 1%.
- The probability of making a Type II error when $\mu = 100$ is to be (approximately) 0.05.

The sample size is n, and the critical region is $\overline{X} < c$, where $\overline{X}$ denotes the sample mean.

(a) Show that n and c satisfy (approximately) the equation $102 - c = \dfrac{13.0256}{\sqrt{n}}$. [3]

(b) Find another equation satisfied by n and c. [2]

(c) Hence find the values of n and c. [4]

(Q8, Jan 2007)

9.

The number of system failures per month in a large network is a random variable with the distribution $\mathrm{Po}(\lambda)$. A significance test of the null hypothesis $\mathrm{H}_0 : \lambda = 2.5$ is carried out by counting R, the number of system failures in a period of 6 months. The result of the test is that H_0 is rejected if $R > 23$ but is not rejected if $R \leqslant 23$.

(i) State the alternative hypothesis. [1]

(ii) Find the significance level of the test. [3]

(iii) Given that $\mathrm{P}(R > 23) < 0.1$, use tables to find the largest possible actual value of λ. You should show the values of any relevant probabilities. [3]

(Q5, June 2007)

10.

A random variable Y is normally distributed with mean μ and variance 12.25. Two statisticians carry out significance tests of the hypotheses $\mathrm{H}_0 : \mu = 63.0$, $\mathrm{H}_1 : \mu > 63.0$.

(i) Statistician A uses the mean $\overline{Y}$ of a sample of size 23, and the critical region for his test is $\overline{Y} > 64.20$. Find the significance level for A's test. [4]

(ii) Statistician B uses the mean of a sample of size 50 and a significance level of 5%.

(a) Find the critical region for B's test. [3]

(b) Given that $\mu = 65.0$, find the probability that B's test results in a Type II error. [4]

(iii) Given that, when $\mu = 65.0$, the probability that A's test results in a Type II error is 0.1365, state with a reason which test is better. [2]

(Q8, June 2007)

11.

The random variable G has the distribution $\mathrm{Po}(\lambda)$. A test is carried out of the null hypothesis $\mathrm{H}_0 : \lambda = 4.5$ against the alternative hypothesis $\mathrm{H}_1 : \lambda \neq 4.5$, based on a single observation of G. The critical region for the test is $G \leqslant 1$ and $G \geqslant 9$.

(i) Find the significance level of the test. [5]

(ii) Given that $\lambda = 5.5$, calculate the probability that the test results in a Type II error. [3]

(Q3, Jan 2008)

12.

Over a long period the number of visitors per week to a stately home was known to have the distribution $N(500, 100^2)$. After higher car parking charges were introduced, a sample of four randomly chosen weeks gave a mean number of visitors per week of 435. You should assume that the number of visitors per week is still normally distributed with variance 100^2.

(i) Test, at the 10% significance level, whether there is evidence that the mean number of visitors per week has fallen. [7]

(ii) Explain why it is necessary to assume that the distribution of the number of visitors per week (after the introduction of higher charges) is normal in order to carry out the test. [2]

(Q5, Jan 2008)

13.

The random variable U has the distribution $N(\mu, \sigma^2)$, where the value of σ is known. A test is carried out of the null hypothesis $H_0 : \mu = 50$ against the alternative hypothesis $H_1 : \mu > 50$. The test is carried out at the 1% significance level and is based on a random sample of size 10.

(i) The test is carried out once. The value of the sample mean is 53.0. The outcome of the test is that H_0 is not rejected. Show that $\sigma > 4.08$, correct to 3 significant figures. [4]

(ii) The test is carried out repeatedly. In each test the actual value of μ is 50. Find the probability that the first test to result in a Type I error is the fifth to be carried out. Give your answer correct to 2 significant figures. [3]

(Q4, June 2008)

14.

Wendy analyses the number of 'dropped catches' in international cricket matches. She finds that the mean number of dropped catches per day is 2. In a recent 5-day match she found that there was a total of c dropped catches. She tests, at the 5% significance level, whether the mean number of dropped catches per day has increased.

(i) State conditions needed for the number of dropped catches per day to be well modelled by a Poisson distribution. [2]

Assume now that these conditions hold.

(ii) Find the probability that the test results in a Type I error.

(iii) Given that $c = 14$, carry out the test.

[10]

(Q7, June 2008)

15.

A company sponsors a series of concerts. Surveys show that on average 40% of audience members know the name of the sponsor. As this figure is thought to be disappointingly low, the publicity material is redesigned.

(i) After the publicity material has been redesigned, a random sample of 12 audience members is obtained, and it is found that 9 members of this sample know the name of the sponsor. Test, at the 5% significance level, whether there is evidence that the proportion of audience members who know the name of the sponsor has increased. [7]

(ii) A more detailed 5% hypothesis test is carried out, based on a random sample of size 400. This test produces significant evidence that the proportion of audience members knowing the name of the sponsor has increased. Using an appropriate approximation, calculate the smallest possible number of audience members in the sample of 400 who know the name of the sponsor. [7]

(Q7, June 2008)

16.

A television company believes that the proportion of adults who watched a certain programme is 0.14. Out of a random sample of 22 adults, it is found that 2 watched the programme.

(i) Carry out a significance test, at the 10% level, to determine, on the basis of this sample, whether the television company is overestimating the proportion of adults who watched the programme. **[8]**

(ii) The sample was selected randomly. State what properties of this method of sampling are needed to justify the use of the distribution used in your test. **[2]**

(Q4, Jan 2009)

17.

The weight of a plastic box manufactured by a company is W grams, where $W \sim \mathrm{N}(\mu, 20.25)$. A significance test of the null hypothesis $\mathrm{H}_0 : \mu = 50.0$, against the alternative hypothesis $\mathrm{H}_1 : \mu \neq 50.0$, is carried out at the 5% significance level, based on a sample of size n.

(i) Given that $n = 81$,

(a) find the critical region for the test, in terms of the sample mean $\overline{W}$, **[5]**

(b) find the probability that the test results in a Type II error when $\mu = 50.2$. **[5]**

(ii) State how the probability of this Type II error would change if n were greater than 81. **[1]**

(Q6, Jan 2009)

18.

A motorist records the time taken, T minutes, to drive a particular stretch of road on each of 64 occasions. Her results are summarised by

$$\Sigma t = 876.8, \quad \Sigma t^2 = 12\,657.28.$$

(i) Test, at the 5% significance level, whether the mean time for the motorist to drive the stretch of road is greater than 13.1 minutes. **[11]**

(ii) Explain whether it is necessary to use the Central Limit Theorem in your test. **[1]**

(Q7, Jan 2009)

19.

An electronics company is developing a new sound system. The company claims that 60% of potential buyers think that the system would be good value for money. In a random sample of 12 potential buyers, 4 thought that it would be good value for money. Test, at the 5% significance level, whether the proportion claimed by the company is too high. **[7]**

(Q3, June 2009)

20.

In a large company the time taken for an employee to carry out a certain task is a normally distributed random variable with mean 78.0 s and unknown variance. A new training scheme is introduced and after its introduction the times taken by a random sample of 120 employees are recorded. The mean time for the sample is 76.4 s and an unbiased estimate of the population variance is 68.9 s^2.

(i) Test, at the 1% significance level, whether the mean time taken for the task has changed. **[7]**

(ii) It is required to redesign the test so that the probability of making a Type I error is less than 0.01 when the sample mean is 77.0 s. Calculate an estimate of the smallest sample size needed, and explain why your answer is only an estimate. **[4]**

(Q8, June 2009)

21.

80 randomly chosen people are asked to estimate a time interval of 60 seconds without using a watch or clock. The mean of the 80 estimates is 58.9 seconds. Previous evidence shows that the population standard deviation of such estimates is 5.0 seconds. Test, at the 5% significance level, whether there is evidence that people tend to underestimate the time interval. **[7]**

(Q4, Jan 2010)

22.

The number of customers arriving at a store between 8.50 am and 9 am on Saturday mornings is a random variable which can be modelled by the distribution Po(11.0). Following a series of price cuts, on one particular Saturday morning 19 customers arrive between 8.50 am and 9 am. The store's management claims, first, that the mean number of customers has increased, and second, that this is due to the price cuts.

(i) Test the first part of the claim, at the 5% significance level. **[7]**

(ii) Comment on the second part of the claim. **[1]**

(Q5, Jan 2010)

23.

The random variable R has the distribution $B(10, p)$. The null hypothesis H_0: $p = 0.7$ is to be tested against the alternative hypothesis H_1: $p < 0.7$, at a significance level of 5%.

(i) Find the critical region for the test and the probability of making a Type I error. **[3]**

(ii) Given that $p = 0.4$, find the probability that the test results in a Type II error. **[3]**

(iii) Given that p is equally likely to take the values 0.4 and 0.7, find the probability that the test results in a Type II error. **[2]**

(Q8, Jan 2010)

24.

The proportion of commuters in a town who travel to work by train is 0.4. Following the opening of a new station car park, a random sample of 16 commuters is obtained, and 11 of these travel to work by train. Test at the 1% significance level whether there is evidence of an increase in the proportion of commuters in this town who travel to work by train. **[7]**

(Q4, June 2010)

25.

A machine is designed to make paper with mean thickness 56.80 micrometres. The thicknesses, x micrometres, of a random sample of 300 sheets are summarised by

$$n = 300, \qquad \Sigma x = 17\,085.0, \qquad \Sigma x^2 = 973\,847.0.$$

Test, at the 10% significance level, whether the machine is producing paper of the designed thickness. **[11]**

(Q7, June 2010)

26.

The time T seconds needed for a computer to be ready to use, from the moment it is switched on, is a normally distributed random variable with standard deviation 5 seconds. The specification of the computer says that the population mean time should be not more than 30 seconds.

(i) A test is carried out, at the 5% significance level, of whether the specification is being met, using the mean $\bar{t}$ of a random sample of 10 times.

(a) Find the critical region for the test, in terms of $\bar{t}$. **[4]**

(b) Given that the population mean time is in fact 35 seconds, find the probability that the test results in a Type II error. **[3]**

(ii) Because of system degradation and memory load, the population mean time μ seconds increases with the number of months of use, m. A formula for μ in terms of m is $\mu = 20 + 0.6m$. Use this formula to find the value of m for which the probability that the test results in rejection of the null hypothesis is 0.5. **[4]**

(Q5, June 2010)

27.

The continuous random variable X has mean μ and standard deviation 45. A significance test is to be carried out of the null hypothesis H_0: $\mu = 230$ against the alternative hypothesis H_1: $\mu \neq 230$, at the 1% significance level. A random sample of size 50 is obtained, and the sample mean is found to be 213.4.

(i) Carry out the test. **[5]**

(ii) Explain whether it is necessary to use the Central Limit Theorem in your test. **[2]**

(Q4, Jan 2011)

28.

A temporary job is advertised annually. The number of applicants for the job is a random variable which is known from many years' experience to have a distribution Po(12). In 2010 there were 19 applicants for the job. Test, at the 10% significance level, whether there is evidence of an increase in the mean number of applicants for the job. **[7]**

(Q5, Jan 2011)

29.

The random variable X has the distribution $N(\mu, 5^2)$. A hypothesis test is carried out of H_0: $\mu = 20.0$ against H_1: $\mu < 20.0$, at the 1% level of significance, based on the mean of a sample of size 16. Given that in fact $\mu = 15.0$, find the probability that the test results in a Type II error. **[7]**

(Q3, June 2011)

30.

A pharmaceutical company is developing a new drug to treat a certain disease. The company will continue to develop the drug if the proportion p of those who have the disease and show a substantial improvement after treatment is greater than 0.7. The company carries out a test, at the 5% significance level, on a random sample of 14 patients who suffer from the disease.

(i) Find the critical region for the test. **[3]**

(ii) Given that 12 of the 14 patients in the sample show a substantial improvement, carry out the test. **[5]**

(iii) Find the probability that the test results in a Type II error if in fact $p = 0.8$. **[3]**

(Q9, Jan 2011)

31.

A travel company finds from its records that 40% of its customers book with travel agents. The company redesigns its website, and then carries out a survey of 10 randomly chosen customers. The result of the survey is that 1 of these customers booked with a travel agent.

(i) Test at the 5% significance level whether the percentage of customers who book with travel agents has decreased. **[7]**

(ii) The managing director says that "Our redesigned website has resulted in a decrease in the percentage of our customers who book with travel agents." Comment on this statement. **[1]**

(Q5, June 2011)

32.

Records show that before the year 1990 the maximum daily temperature T °C at a seaside resort in August can be modelled by a distribution with mean 24.3. The maximum temperatures of a random sample of 50 August days since 1990 can be summarised by

$$n = 50, \qquad \Sigma t = 1314.0, \qquad \Sigma t^2 = 36\,602.17.$$

(i) Test, at the 1% significance level, whether there is evidence of a change in the mean maximum daily temperature in August since 1990. **[11]**

(ii) Give a reason why it is possible to use the Central Limit Theorem in your test. **[1]**

(Q6, June 2011)

33.

The number of fruit pips in 1 cubic centimetre of raspberry jam has the distribution Po(λ). Under a traditional jam-making process it is known that $\lambda = 6.3$. A new process is introduced and a random sample of 1 cubic centimetre of jam produced by the new process is found to contain 2 pips. Test, at the 5% significance level, whether this is evidence that under the new process the average number of pips has been reduced. **[8]**

(Q6, Jan 2012)

34.

It is desired to test whether the average amount of sleep obtained by school pupils in Year 11 is 8 hours, based on a random sample of size 64. The population standard deviation is 0.87 hours and the sample mean is denoted by $\overline{H}$. The critical values for the test are $\overline{H} = 7.72$ and $\overline{H} = 8.28$.

(i) State appropriate hypotheses for the test, explaining the meaning of any symbol you use. **[3]**

(ii) Calculate the significance level of the test. **[4]**

(iii) Explain what is meant by a Type I error in this context. **[1]**

(iv) Given that in fact the average amount of sleep obtained by all pupils in Year 11 is 7.9 hours, find the probability that the test results in a Type II error. **[3]**

(Q6, Jan 2012)

35.

It is known that on average one person in three prefers the colour of a certain object to be blue. In a psychological test, 12 randomly chosen people were seated in a room with blue walls, and asked to state independently which colour they preferred for the object. Seven of the 12 people said that they preferred blue. Carry out a significance test, at the 5% level, of whether the statement "on average one person in three prefers the colour of the object to be blue" is true for people who are seated in a room with blue walls. **[7]**

(Q3, June 2012)

36.

A random variable has the distribution B(n, p). It is required to test H_0: $p = \frac{2}{3}$ against H_1: $p < \frac{2}{3}$ at a significance level as close to 1% as possible, using a sample of size $n = 8$, 9 or 10. Use tables to find which value of n gives such a test, stating the critical region for the test and the corresponding significance level. **[4]**

(Q1, Jan 2013)

37.

The random variable X has the distribution $N(\mu, 8^2)$. A test is carried out, at the 5% significance level, of H_0: $\mu = 30$ against H_1: $\mu > 30$, based on a random sample of size 18.

(i) Find the critical region for the test. **[4]**

(ii) If $\mu = 30$ and the outcome of the test is that H_0 is rejected, state the type of error that is made. **[1]**

On a particular day this test is carried out independently a total of 20 times, and for 4 of these tests the outcome is that H_0 is rejected. It is known that the value of μ remains the same throughout these 20 tests.

(iii) Find the probability that H_0 is rejected at least 4 times if $\mu = 30$. Hence state whether you think that $\mu = 30$, giving a reason. **[3]**

(iv) Given that the probability of making an error of the type different from that stated in part **(ii)** is 0.4, calculate the actual value of μ, giving your answer correct to 4 significant figures. **[4]**

(Q8, June 2012)

38.

Gordon is a cricketer. Over a long period he knows that his population mean score, in number of runs per innings, is 28, and the population standard deviation is 12. In a new season he adopts a different batting style and he finds that in 30 innings using this style his mean score is 28.98.

(i) Stating a necessary assumption, test at the 5% significance level whether his population mean score has increased. **[8]**

(ii) Explain whether it was necessary to use the Central Limit Theorem in part **(i)**. **[2]**

(Q6, Jan 2013)

39.

The random variable A has the distribution $B(30, p)$. A test is carried out of the hypotheses H_0: $p = 0.6$ against H_1: $p < 0.6$. The critical region is $A \leqslant 13$.

(i) State the probability that H_0 is rejected when $p = 0.6$. **[1]**

(ii) Find the probability that a Type II error occurs when $p = 0.5$. **[2]**

(iii) It is known that on average $p = 0.5$ on one day in five, and on other days the value of p is 0.6. On each day two tests are carried out. If the result of the first test is that H_0 is rejected, the value of p is adjusted if necessary, to ensure that $p = 0.6$ for the rest of the day. Otherwise the value of p remains the same as for the first test. Calculate the probability that the result of the second test is to reject H_0. **[5]**

(Q9, Jan 2013)

40.

The number of floods in a certain river plain is known to have a Poisson distribution. It is known that up until 10 years ago the mean number of floods per year was 0.32. During the last 10 years there were 6 floods. Test at the 1% significance level whether there is evidence of an increase in the mean number of floods per year. **[7]**

(Q4, June 2013)

41.

The random variable X denotes the yield, in kilograms per acre, of a certain crop. Under the standard treatment it is known that $\mathrm{E}(X) = 38.4$. Under a new treatment, the yields of 50 randomly chosen regions can be summarised as

$$n = 50, \qquad \Sigma x = 1834.0, \qquad \Sigma x^2 = 70027.37.$$

Test at the 1% level whether there has been a change in the mean crop yield. **[11]**

(Q6, June 2013)

42.

Past experience shows that 35% of the senior pupils in a large school know the regulations about bringing cars to school. The head teacher addresses this subject in an assembly, and afterwards a random sample of 120 senior pupils is selected. In this sample it is found that 50 of these pupils know the regulations. Use a suitable approximation to test, at the 10% significance level, whether there is evidence that the proportion of senior pupils who know the regulations has increased. Justify your approximation. **[11]**

(Q7, June 2013)

43.

The random variable R has the distribution $\mathrm{B}(14, p)$. A test is carried out at the α% significance level of the null hypothesis H_0: $p = 0.25$, against H_1: $p > 0.25$.

(i) Given that α is as close to 5 as possible, find the probability of a Type II error when the true value of p is 0.4. **[4]**

(ii) State what happens to the probability of a Type II error as

(a) p increases from 0.4,

(b) α increases, giving a reason. **[2]**

(Q8, June 2013)

44.

The random variable W has the distribution $\text{Po}(\lambda)$. A significance test is carried out of the null hypothesis H_0: $\lambda = 3.60$, against the alternative hypothesis H_1: $\lambda < 3.60$. The test is based on a single observation of W. The critical region is $W = 0$.

(i) Find the significance level of the test. **[2]**

(ii) It is known that, when $\lambda = \lambda_0$, the probability that the test results in a Type II error is 0.8. Find the value of λ_0. **[4]**

(Q8, June 2014)

45.

In a city the proportion of inhabitants from ethnic group Z is known to be 0.4. A sample of 12 employees of a large company in this city is obtained and it is found that 2 of them are from ethnic group Z. A test is carried out, at the 5% significance level, of whether the proportion of employees in this company from ethnic group Z is less than in the city as a whole.

(i) State an assumption that must be made about the sample for a significance test to be valid. **[1]**

(ii) Describe briefly an appropriate way of obtaining the sample. **[2]**

(iii) Carry out the test. **[7]**

(iv) A manager believes that the company discriminates against ethnic group Z. Explain whether carrying out the test at the 10% significance level would be more supportive or less supportive of the manager's belief. **[2]**

(Q6, June 2014)

46.

An examination board is developing a new syllabus and wants to know if the question papers are the right length. A random sample of 50 candidates was given a pre-test on a dummy paper. The times, t minutes, taken by these candidates to complete the paper can be summarised by

$$n = 50, \qquad \sum t = 4050, \qquad \sum t^2 = 329\,800.$$

Assume that times are normally distributed.

(i) Estimate the proportion of candidates that could not complete the paper within 90 minutes. **[6]**

(ii) Test, at the 10% significance level, whether the mean time for all candidates to complete this paper is 80 minutes. Use a two-tail test. **[7]**

(iii) Explain whether the assumption that times are normally distributed is necessary in answering

(a) part **(i)**,

(b) part **(ii)**. **[2]**

(Q7, June 2014)

47.

A continuous random variable is normally distributed with mean μ. A significance test for μ is carried out, at the 5% significance level, on 90 independent occasions.

(i) Given that the null hypothesis is correct on all 90 occasions, use a suitable approximation to find the probability that on 6 or fewer occasions the test results in a Type I error. Justify your approximation. **[5]**

(ii) Given instead that on all 90 occasions the probability of a Type II error is 0.35, use a suitable approximation to find the probability that on fewer than 29 occasions the test results in a Type II error. **[5]**

(Q4, June 2015)

48.

(i) State an advantage of using random numbers in selecting samples. **[1]**

(ii) It is known that in analysing the digits in large sets of financial records, the probability that the leading digit is 1 is 0.25. A random sample of 18 leading digits from a certain large set of financial records is obtained and it is found that 8 of the leading digits are 1s. Test, at the 5% significance level, whether the probability that the leading digit is 1 in this set of records is greater than 0.25. **[7]**

(Q5, June 2015)

49.

Records for a doctors' surgery over a long period suggest that the time taken for a consultation, T minutes, has a mean of 11.0. Following the introduction of new regulations, a doctor believes that the average time has changed. She finds that, with new regulations, the consultation times for a random sample of 120 patients can be summarised as

$$n = 120, \Sigma t = 1411.20, \Sigma t^2 = 18\,737.712.$$

(i) Test, at the 10% significance level, whether the doctor's belief is correct. **[11]**

(ii) Explain whether, in answering part **(i)**, it was necessary to assume that the consultation times were normally distributed. **[1]**

(Q6, June 2015)

50.

A large railway network suffers points failures at an average rate of 1 every 3 days. Assume that the number of points failures can be modelled by a Poisson distribution. The network employs a new firm of engineers. After the new engineers have become established, it is found that in a randomly chosen period of 15 days there are 2 instances of points failures.

(i) Test, at the 5% significance level, whether there is evidence that the mean number of points failures has been reduced. **[7]**

(ii) A new test is carried out over a period of 150 days. Use a suitable approximation to find the greatest number of points failures there could be in 150 days that would lead to a 5% significance test concluding that the average number of points failures had been reduced. **[6]**

(Q7, June 2015)

51.

The random variable S has the distribution $B(14, p)$. A significance test is carried out of the null hypothesis H_0: $p = 0.3$ against the alternative hypothesis H_1: $p > 0.3$. The critical region for the test is $S \geqslant 8$.

(i) Find the significance level of the test, correct to 3 significant figures. **[2]**

(ii) It is given that, on each occasion that the test is carried out, the true value of p is equally likely to be 0.3, 0.5 or 0.7, independently of any other test. Four independent tests are carried out. Find the probability that at least one of the tests results in a Type II error. **[5]**

(Q8, June 2015)

52.

55% of the pupils in a large school are girls. A member of the student council claims that the probability that a girl rather than a boy becomes Head Student is greater than 0.55. As evidence for his claim he says that 6 of the last 8 Head Students have been girls.

(i) Use an exact binomial distribution to test the claim at the 10% significance level. **[7]**

(ii) A statistics teacher says that considering only the last 8 Head Students may not be satisfactory. Explain what needs to be assumed about the data for the test to be valid. **[1]**

(Q5, June 2016)

53.

The random variable R has the distribution $\mathrm{Po}(\lambda)$. A significance test is carried out at the 1% level of the null hypothesis $H_0: \lambda = 11$ against $H_1: \lambda > 11$, based on a single observation of R. Given that in fact the value of λ is 14, find the probability that the result of the test is incorrect, and give the technical name for such an incorrect outcome. You should show the values of any relevant probabilities. **[6]**

(Q9, June 2016)

53.

It is known that the lifetime of a certain species of animal in the wild has mean 13.3 years. A zoologist reads a study of 50 randomly chosen animals of this species that have been kept in zoos. According to the study, for these 50 animals the sample mean lifetime is 12.48 years and the population variance is 12.25 years^2.

(i) Test at the 5% significance level whether these results provide evidence that animals of this species that have been kept in zoos have a shorter expected lifetime than those in the wild. **[7]**

(ii) Subsequently the zoologist discovered that there had been a mistake in the study. The quoted variance of 12.25 years^2 was in fact the sample variance. Determine whether this makes a difference to the conclusion of the test. **[5]**

(iii) Explain whether the Central Limit Theorem is needed in these tests. **[1]**

(Q8, June 2016)