M140

TMA 03 2020J

Covers Units 6, 7, 8 and 9 **Cut-off date 24 March 2021**

Please read the Assessment Guide on the module website before beginning work on this TMA. You can submit your TMA either by post or electronically using the University's online TMA/EMA service.

This TMA is marked out of 100. Your overall score for this TMA will be the sum of your marks for each question.

The marks allocated to each part of each question are indicated in brackets in the margin.

Guidance about how to answer TMA questions is given in Subsection 7.2 of Unit 1.

Note that the Minitab files that you require for this assignment should be downloaded from the 'Assessment' area on the module website.

Please note that you should round you answers to an appropriate level of accuracy.

If you have a disability that makes it difficult for you to attempt any of these questions, then please contact your Student Support Team or your tutor for advice.

*You should be able to answer Questions 1 and 2 after you have studied
Unit 6.*

*You should be able to answer Questions 3 and 4 after you have studied
Unit 7. You will need to use Minitab to answer Question 4.*

*You should be able to answer Questions 5 and 6 after you have studied
Unit 8. You will need to use Minitab to answer Question 6.*

*You should be able to answer Questions 7 and 8 after you have studied
Unit 9. You will need to use Minitab to answer Questions 7 and 8.*

Question 1 (Unit 6)   –   13 marks

(a) A researcher is interested in studying childhood obesity in the UK. She
downloads some data from the government website relating to
overweight children. This data was collected through the National Child
Measurement Programme (NCMP) and a child is defined as obese based
on their BMI measurement. The latest data that she finds is for the
period 2017 to 2018; this data is split by ethnicity and given for children
aged 4 to 5 years old and 10 to 11 years old. The number of children in
these age ranges that are not considered to be obese is also given.

The data is summarised and shown in Table 1.

Table 1

Ethnicity	Obese 4–5 yr olds	Not obese 4–5 yr olds	Obese 10–11 yr olds	Not obese 10–11 yr olds	Total
Asian	10828	47768	22327	34044	
Black		18337	13001	15875	
Mixed	6477	23670	9064	15977	
White	87875	303374	118458	248885	
Chinese or Other	3231	11288		8275	27901
Unknown	20587	69412	29931	55630	
Total	136586				

In this question we will consider the case where one child is selected at
random.

(i) Complete Table 1 by calculating the missing entries, and the
column, row and overall totals. Present these as part of a table.   [3]

(ii) Calculate the probability that the selected child is obese and is 4 to
5 years old.   [2]

(iii) Calculate the probability that the selected child is black, obese and
is 10 to 11 years old.   [2]

(iv) Calculate the probability that a 4 to 5 year old obese child is not
white. Show all of your working.   [3]

(b) Emily, Rebecca and Caroline are training for a marathon. They need to train by running several times a week over a six-week period. The probability that on a particular day Emily chooses to run is 0.45, for Rebecca this probability is 0.25, and for Caroline this probability is 0.3.

What is the probability that all three will choose to run on the same day?

State any important assumption that you make in order to calculate this probability.   [3]

Question 2 (Unit 6) – 12 marks

A scientist wants to investigate the effects of climate change in the UK. The rainfall in 12 regions of the United Kingdom is recorded for February 2008 and 2018. Table 2 gives the total rainfall, in mm, along with the differences between the rainfall in February 2008 and 2018. This data comes from the Met Office and is shown in Table 2.

Table 2

Region	Rainfall (mm) Feb 2008	Rainfall (mm) Feb 2018	Difference (mm) 2018–2008
England N and NE	33.3	46.7	13.4
England North	51.1	54.7	3.6
England South	30.4	39.5	9.1
Midlands	36.7	34.6	−2.1
England NW and N Wales	80.8	76.2	−4.6
England SW and S Wales	60.4	62.6	2.2
East Anglia	18.7	40.3	21.6
England SE and Central	26.5	40.9	14.4
Scotland West	176.8	116.7	−60.1
Scotland East	87.1	57.3	−29.8
Scotland North	206.7	98.4	−108.3
Northern Ireland	60.7	74.0	13.3

A sign test is to be performed to investigate whether the rainfall recorded in February 2018 differs from the rainfall recorded in February 2008.

(a) Write down the hypothesis to be tested. [2]

(b) Record the number of values lying above and the number lying below the hypothesised value. What is the value of the test statistic? [2]

(c) What is the appropriate critical value at the 5% significance level? [1]

(d) Decide whether or not you would reject the hypothesis at the 5% significance level. [1]

(e) Using Figure 6 of Unit 6 (Subsection 4.1), calculate the p-value given by the hypothesis test. [3]

(f) Looking at this p-value, and using Table 10 of Unit 6 (Subsection 5.1), what conclusion can be drawn from the hypothesis test? [1]

(g) How does this conclusion sit with the result of part (d)? [1]

(h) What is your overall conclusion in terms of the difference in rainfall in February 2008 and February 2018, based on these data? [1]

Question 3 (Unit 7) – 15 marks

(a) For the normal distribution shown in Figure 1, find approximate values for its mean and standard deviation. Explain how you obtained your answers. [3]
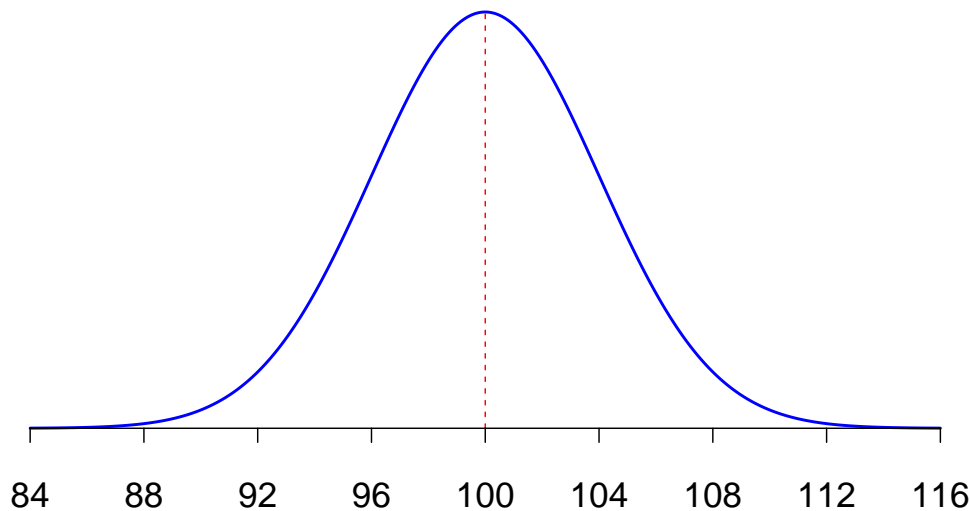

Figure 1

(b) (i) The normal distribution of a variable x has mean $\mu = 70$ and standard deviation $\sigma = 5$. Sketch this distribution. [3]

(ii) Write down the formula for z that converts each value of the variable x in part (b)(i) so that z follows a standard normal distribution. Your answer should contain both the general formula and the specific formula for this example. [2]

(iii) Calculate the value of z corresponding to $x = 61$. [1]

(iv) Interpret the value of z in terms of the number of standard deviations x is above or below its mean. [2]

(c) Suppose that the population distribution of weights of bags of cement produced in a factory has mean $10\,\text{kg}$ and standard deviation $0.05\,\text{kg}$.

(i) Find the standard deviation of the sampling distribution of the mean weight for samples of 50 bags of cement. [2]

(ii) Hence give the approximate distribution of the mean weight for samples of 50 such bags of cement. [2]

Question 4 (Unit 7) – 10 marks

You will need to use Minitab to answer this question.

A random sample of mobile phone prices from one store in 2019 were collected. The sample size, mean and standard deviation are presented in Table 3.

Table 3

Sample size	Mean price (£)	Standard deviation (£)
44	531.10	337.90

The overall population mean mobile phone price was £548.80.

A one-sample z-test is to be performed to investigate whether the mean mobile phone price in the store was the same as the population mobile phone price in all stores in 2019.

(a) Using appropriate notation, which you should define, specify the null and alternative hypotheses associated with the test. [2]

(b) Use Minitab to perform a one-sample z-test on the mobile phone price data. Include a copy of the relevant Minitab output in your answer. What is the value of the test statistic? [3]

(c) Give the p-value from the test performed in part (b), and say what may be concluded from the test in terms of evidence about the hypotheses. Relate your conclusions back to consideration of the mean mobile phone price from this store. [3]

(d) Suppose that in addition, a sample of size greater than 25 of phone prices was also obtained from a different store in 2019. The researchers proposed to use a two-sample z-test to make an appropriate comparison regarding the two stores. Using appropriate notation, which you should define, specify the null and alternative hypotheses that would be associated with this test. [2]

Question 5 (Unit 8) – 12 marks

Yuliya is training for a triathlon. She is confident with her cycling, so is concentrating on running and swimming while training. It turns out that:

- on 55% of the days she goes running

- of the days that she goes running, on 60% of these days she also goes swimming.

(a) Let R denote the event Yuliya goes running while training, and let S denote the event that Yuliya goes swimming.

Write the information given in the two bullet points in symbolic form. [2]

(b) Calculate the probability that on a randomly chosen day Yuliya chooses to go running and swimming while training. [3]

(c) Additional information is now given that Yuliya goes swimming on 50% of the days. Calculate the probability that on a randomly chosen day, while training, when Yuliya has chosen to go swimming that she also goes running. [3]

(d) What is the probability that Yuliya chooses to run or swim, or both while training? [2]

(e) For Yuliya, while training, are the events of running and swimming independent? Give a reason for your answer. [2]

Question 6 (Unit 8) – 13 marks

You will need to use Minitab to answer this question.

The Minitab file that you require for this question should be downloaded from the 'Assessment' area on the module website.

Table 4 shows NHS data relating to children and young people with eating disorders. Recorded are the waiting times before treatment is started for routine cases in a number of regions in England. The data covers the period 2018–19.

Table 4

Region name	0–1 week	>1–4 weeks	4 plus
North of England	114	287	90
Midlands and East of England	126	271	75
London	39	187	18
South West of England	61	146	79
South East of England	77	160	51

(a) Check whether this table is a contingency table, by checking that it meets all of the three criteria. Explain your reasons. [3]

(b) The data are to be analysed using a χ^2 test for contingency tables.

The Minitab worksheet **eating-disorders.mwx** contains this data. The columns are labelled `0-1`, `>1-4`, and `4 plus`. The rows are labelled `North`, `Midlands and East`, `London`, `South West` and `South East`.

(i) Specify the null and alternative hypotheses associated with the χ^2 test. [2]

(ii) Perform the χ^2 test using Minitab. Include a copy of the relevant Minitab output in your answer. [2]

(iii) Identify, from your Minitab output, the expected value for London where children and young people waited more than 4 weeks for treatment. [1]

(iv) Minitab gives the degrees of freedom associated with this χ^2 test as 8. Show how this value arises. [2]

(v) Interpret the results of the χ^2 test, giving your conclusions about the relationship between waiting times for treatment and region of England. [3]

Question 7 (Unit 9) – 9 marks

You will need to use Minitab to answer this question.

The Minitab file that you require for this question should be downloaded from the 'Assessment' area on the module website.

The Minitab worksheet **Uranium-Crudeoil.mwx** contains three columns. The first column, headed `year`, contains years 1979 to 2010. The second column, headed `Uranium`, contains values representing US uranium imports, measured in millions of barrels. The third column, headed `Crude Oil`, contains values representing US crude oil imports, measured in millions of pounds.

Run Minitab and open this worksheet.

(a) Use Minitab to make a scatterplot of `Uranium` as the Y variable against `Crude Oil` as the X variable. Include a copy of your plot in your answer.   [2]

(b) From your plot, is there a strong correlation between `Uranium` and `Crude Oil`? Justify your answer.   [2]

(c) Using Minitab, calculate the correlation coefficient between `Uranium` and `Crude Oil`. You do not need to include a copy of the Minitab output in your answer.   [2]

(d) Is there a positive or negative correlation between `Uranium` and `Crude Oil`? Justify your answer.   [2]

(e) Does a change in US crude oil imports necessarily cause a change in US uranium imports?   [1]

Question 8 (Unit 9) – 16 marks

You will need to use Minitab to answer this question.

The survey of mobile phone prices in Question 4 also has information available on the mobile phone prices of a different store, labelled Store 2. The mean price of mobile phones in the two stores, Store 1 and Store 2, are given in Table 5. Assume that this is a random sample of the mobile phone models available from each of these stores.

Table 5

	Sample size	Mean price (£)	Standard deviation (£)
Store 1	44	531.10	337.90
Store 2	32	448.00	301.00

(a) (i) Using the **1-Sample Z...** option in Minitab, calculate and report a 95% confidence interval for the population mean price of mobile phones in Store 1. Include a copy of the Minitab output in your answer. [3]

 (ii) Fill in the blanks in the following paragraph to interpret the confidence interval that you have just calculated, in terms of all possible random samples of mobile phones from Store 1.

 About _____ of the possible random samples we could select will give rise to an interval containing the population _____, while only about _____ of the possible random samples we could select will give rise to an interval that does not contain the population _____. [2]

 (iii) On the basis of the confidence interval from part (a)(ii), what would have been the outcome of a z-test of the null hypothesis that the population mean price of mobile phones at Store 1 was £450.00? Interpret the result of the test. [3]

(b) Now we want to compare the mean mobile phone prices between the two stores given in Table 4.

 Interest now centres on the difference between prices of mobile phones in Store 1 and Store 2.

 (i) Calculate by hand the value of the estimated standard error of the difference between the sample means. [3]

 (ii) Calculate a 95% confidence interval for the difference between the population mean prices in Store 1 and Store 2. [3]

 (iii) On the basis of the confidence interval from part (b)(ii), what would have been the outcome of a z-test of the null hypothesis that the population mean price of mobile phones in Store 1 was the same as the population mean price of mobile phones in Store 2? Interpret the result of the test. [2]