Question 1

The quality control inspector of a supermarket chain based in Sydney has the task of checking on the weights of various cuts of meat for sale in the freezer. They are labelled as ‘net weight 250 grams’ although they are cut by a machine that may not be entirely accurate. She has recently heard reports that several customers who shop in the chain are becoming increasingly dissatisfied and suspicious that some of the meat does not adequately meet the displayed weight. Although some leeway is tolerated in the weights, if they are too light (customer paying too much) or too heavy (store loses money) then serious problems may arise for the chain management. This means, of course, that some meat may be *under* specifications and some may be *over* specifications, but the *overall mean* weight of all such cuts of meat should be very close to the specified weight. Any inaccuracy is due to the fact that the cutting process itself is subject to variation.

The inspector assures the complaining customers that the weights of the meat will be checked, but she clearly does not have time to weigh them all. To get a preliminary idea on any possible problems, she selects a random sample of 60 cuts of meat from supermarkets in the chain from each of the four regions of Sydney (East, West, North and South) These meats are ideally supposed to have a weight of 250 grams. She weighs each cuts of meat in the sample and records the weight. These 240 weights are given in C1 of the *Minitab* worksheet *T3 2020 Cut Weight.mtw* that may be downloaded from the IDA website. Column C2 of the worksheet shows the region from which the meat was sampled.

The instructions from management is that she is to calculate, based on the sample, a *95% confidence interval* for the population mean weight of *all* cuts of meat on sale in the supermarket chain. This confidence interval must meet *all* *three* of the following requirements:

* It *must* contain 252.00 grams
* It must *not* contain 255.00 grams
* It must *not* contain 245.00 grams

If these requirements are *not* all met, the inspector must launch an immediate investigation by measuring the weight of *every* cut of meat on sale during the coming six months – a process that will be very time consuming and expensive for the supermarket chain. On the other hand, if the above requirements *are* all met, she is to report that the meat weights are within acceptable limits (although there may be some outside) and she will not pursue the matter further for the time being.

Answer the following questions.

**1**(a) Use *Minitab* to find the *mean, standard error of the mean, standard deviation, minimum, Q1,* *median, Q3 and maximum* of the weights of the 60 cuts of meat in the sample for each region **and** the overall in values for all 240 cuts in Sydney. Cut and paste the relevant Minitab output below. (1 mark)

1. Use Minitab to produce a histogram (with fit) of the distribution of cut weights for the entire Sydney region. Analyse the graph and comment on your findings. (1 mark)

The histogram follows a Normal distribution as mean is almost equivalent to the median of the population of data set.

1. What is the *percentage* of the cuts of meat *in the entire Sydney sample* that have a weight *of at least* 250 grams? (1 mark)
2. Based on the sample, find by hand a 95% confidence interval for the estimate of the ***percentage* of the population** of *all* such cuts of meat in Sydney that have a weight *less than* 250 grams. Would you believe that m*ore than 45%* of all such cuts of meat weigh less than 250 grams? Explain the reasons for your answer and show all working. (4 marks)
3. Find, by hand, a *95% confidence interval* for the estimate of the ***mean* *weight*** of *all* cuts of meat in the chain across Sydney. Indicate whether they meet all the requirements of the inspector and therefore whether she should pursue the matter. (5 marks)

(f) The inspector now wants to know if there is any difference in the mean weights of the meat across the four regions. She decides to use a *One-way ANOVA* to determine this. Answer the following questions.

* 1. Firstly , use *Minitab* to compare box plots of the data for each region. Insert it below and comment on what preliminary impressions you get regarding significant differences between the regions, and the presence of any outliers. Base your answer only on the boxplot output. (1 mark)
  2. Use a *Ryan-Joiner* Normality Test to determine whether it is reasonable to assume that the weights for Sydney came from a normal distribution. Insert the plot below and state the reasons for your conclusion. (1 mark)
  3. Use Minitab to perform a one-way ANOVA to determine if there is a difference in weights between the four regions. Use α = 0.01. Insert the relevant Minitab output below.

Be sure to write out the hypotheses and state your conclusions with reasons.

You should also test if there is anything that may invalidate your conclusion reached using the ANOVA test. (4 marks)

(iv) Use Tukey’s test with a 5% family error rate to determine where significant differences exist, if any. Insert the relevant *Minitab* output below. Write your conclusions clearly, with reasons. Summarise your results for any significant differences that exist. (2 marks)

Question 2

1. Cochlear are a listed Australian company that manufacture bionic ear implants that enable profoundly deaf people to hear. The following report appeared in the Australian Financial Review on Saturday 12th October, 2019.

**Broker watch - Hopes of an upgrade at Cochlear**

*Sarah Turner, AFR Sat Oct 12th 2019.*

*There is a 30 per cent probability that Cochlear will upgrade its 2019-20 profit guidance at its October 22 annual general meeting, according to broker Morgan Stanley, which ascribes an even higher probability to the company reasserting guidance.*

*Cochlear has forecast a net profit after tax for fiscal year 2019-20 between $290 million and $300 million, a 9 to 13 per cent increase on underlying net profit in 2018-19.*

*The broker believes there is a 30 per cent chance Cochlear will upgrade this to between $300 million and $310 million, a 60 per cent chance it will retain its current guidance and a 10 per cent chance it will downgrade to $280 million to $290 million.*

*“We think management commentary regarding short-term influences of unit growth will be prescriptive for the share price,” Morgan Stanley said.*

*An upgrade would add an estimated 10 per cent to Cochlear shares, and the stock stands to advance 5 per cent if guidance can be maintained. In the unlikely event of a downgrade the shares would fall 15 per cent.*

*The stock was trading at $208 on Friday.*

*A guidance downgrade might be possible depending on the progress of ongoing investment in development.*

*Cochlear stands to regain most or all of the market share that it lost in 2018-19. “Due to the low base for unit growth in fiscal year 2019, we expect very strong unit growth in FY20 – potentially leading to upward revision of guidance,” Morgan Stanley said.*

*Alternatively, the company may only say that while it is observing strong unit growth, competing product launches are a risk in the second half of 2019-20.*

Based on the information provided in the article, estimate the expected value of the Cochlear share price after the profit guidance is released on October 22nd and determine if you would recommend purchase at the current share price. (You can base your recommendation only on the profit guidance report and the expected change is share price.). (5 marks)

1. An on-line order firm considers three possible foul-ups in filling an order:

A: The wrong item is sent. This happens to 2% of orders.

B: The item is lost in transit. This happens to 1% of orders.

C: The item is damaged in transit. This happens to 4% of orders.

Assume that event A is independent of event B; that event A is independent of event C; and that events B and C are mutually exclusive.

Find the probability that at least one of these foul-ups occurs for a randomly chosen order. (2 marks)

1. The inwards goods receiving department of the firm randomly tests batches of electrical appliances before approving the batch for sale on their web page. Unknown to the inspector, an incoming batch of 100 appliances contains 2 defectives. The inspector selects two appliances at random. (3 marks)
   1. What is the probability that the first appliance is defective?
   2. What is the probability that the second appliance is defective if the first appliance proved not to be defective?

Question 3 (10 marks)

Airline records show that the total number of paper towels used by passengers in the washrooms on their Sydney-Hong Kong weekday *QF127* flights follows a normal distribution with a mean of 372 towels with a standard deviation of 32.0 towels.

1. On a particular weekday flight they carry 400 towels at the beginning of the flight. What is the probability they will run out of them? (2 marks)
2. Assuming that the use of towels on each flight are independent, calculate the probability (to two decimal places) that on each of four successive weekday flights there will be at least 300 towels used. (4 marks)
3. The airline wants to determine how many towels to take on the *QF127* flight on Wednesday. How many towels should they take so that there is at least a 99% chance that they will not run out? (2 marks)
4. The airline wants to determine how many towels to take on the *QF127* flight on Thursday. How many towels should they take so that there is at least a 90% chance that all the towels will be used? (2 marks)

Question 4 (8 marks)

A consultant has been hired to determine the amount of time customers spend queueing for various types of lottery tickets at a large suburban newsagent. The consultant selects a random sample of 85 customers at random times during the week and finds that the mean waiting time in the queue is 503 seconds with a standard deviation of 70 seconds. These times do not follow a normal distribution.

1. Construct a 95% confidence interval for the mean waiting time for a customer who queues and clearly interpret your findings. (Include the value of the parameters and the formula you use.) (4 marks)
2. The newsagent advertises that the mean time their customers wait in the queue to purchase tickets is less than 4 minutes. Based on the evidence collected by the consultant, do you think this is false advertising? Also provide the chance that you have made an error in your conclusion. (2 marks)
3. The consultant wants to estimate the proportion of customers who wait less than 4 minutes. If you can estimate this proportion with the data supplied in the question, then do so. If not, explain why. (2 marks)

Question 5

The number of males and females (in ‘000) employed full-time in Australia is shown for each month between January 1979 and April 2020 in a file named *T32020 Employment Data.mtw* that can be downloaded from the IDA website. (*Source*: ABS, 6202.0 - Labour Force, Australia, Table 01). Use this data to answer the following questions using *Minitab*.

1. Draw a scatterplot of the *Employed total-Males* v *Employed total-Female* data and insert the *Minitab* output **below**. Comment on the correlation. (1 mark)
2. Find the correlation coefficient between *Employed total-Males* and *Employed total-Female* along with the *p*-value. Insert the relevant *Minitab* output **below**. State your conclusions regarding the strength of the correlation with reasons. (1 mark)
3. Produce a single time series plot that contains both *Employed total-Males* and *Employed total-Females.* Comment on aspects of the data that it reveals.(2 marks)

Choose either d(i) or d(ii) below depending on your version of Minitab used.

1. (i) MINITAB – 16/17/18 - Create an **additive** **time series decomposition** to forecast the values for *Employed total-Males*. Comment on the overall accuracy of the model. (2 marks)

OR

(ii) MINITAB Express - Use **time series trend analysis** to forecast the values for *Employed total-Males.* Comment on the overall accuracy of the model.(2 marks)

(e ) Write down the Fitted Trend Equation for your time series model produced in part (d) and use it to answer the following questions.

1. What is the average monthly increase in number of *Employed total-Males.* (1 marks)
2. What is the forecast value ( to two decimal places) for the number of *Employed total-Males* in April 2020. Compare the forecast value with the actual value and comment on the accuracy of the forecast. (3 marks)

Question 6

The file *T32020 Cholesterol.mtw* contains data taken from a random sample of 21 males and 29 females who work in administration at University. Recorded, along with their gender, (1 = male, 0 = female) is their cholesterol level (mmol/litre), systolic blood pressure (mm), weight (kg), height (cm) and age (years). The aim is to find the best model that will predict the value of the cholesterol level based on the best combination of the five variables *Weight*, *Height*, *Age* and *Gender* and *Blood Pressure*.

Answer the following questions using *Minitab* and insert relevant output with comments below.

1. Construct and insert a *correlation matrix* of all six variables including their *p*-values. Comment on any variables you find to be significantly correlated at α = 0.01. Give reasons. (1 mark)
2. Run a best subsets regression of *Cholesterol level* using the five variables *Weight*, *Height*, *Age*, *Gender* and *Blood Pressure*. Select, with reasons, the best subset of variable(s) you would use in your final model. (2 marks)
3. Find the regression equation of the best fit regression line of *Cholesterol level* on the variables you selected in (b). [Do not remove outliers.] Comment on whether you feel the model is a good fit of the data, giving specific reasons for your assessment. (5 marks)
4. Explain clearly the impact and significance of each of the coefficients of the variables of your regression equation in (c). (3 marks)
5. Use your equation in (c) and *Minitab* to :
   1. predict the cholesterol level of a 38-year-old female who weighs 64kg, is 162cm tall and has a blood pressure of 125mm. Also find the 95% prediction interval and interpret. (2 marks)
   2. to predict the cholesterol level of a 59-year-old male who has a blood pressure of 147mm, weighs 84kg, is 179cm tall and has a blood pressure of 147mm. Also find the 95% prediction interval and interpret. (2 marks)

Question 7

Since the emergence of the Coronavirus COVID-19 in late 2019 the search has been on for an effective treatment. One drug trialled (and heavily promoted by some including President Donald Trump in the USA and Clive Palmer in Australia) has been hydroxychloroquine. This drug has also been used in combination with another drug azithromycin.

One [nationwide study](https://www.medrxiv.org/content/10.1101/2020.04.16.20065920v1.full.pdf) of patients treated at Veterans Administration medical centres in the USA has provided one of the largest dataset yet reported on the outcomes of COVID-19 patients treated with hydroxychloroquine, with or without azithromycin, anywhere in the world. (ref: <https://www.medrxiv.org/content/10.1101/2020.04.16.20065920v1.full.pdf>)

The results reported are summarised in the table below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Outcome | HC | HC+AV | No HC | Total |
| Death | 27 | 25 | 18 | 70 |
| Recovered | 70 | 88 | 140 | 298 |
| Total | 97 | 113 | 158 | 368 |

HC: hydroxychloroquine-treated

HC+AZ: hydroxychloroquine and azithromycin-treated

No HC: not treated with hydroxychloroquine or azithromycin ( i.e. Control Group)

1. Perform a Chi-sq test (α =0.01) to determine if the drug treatments have any effect. Be sure to state your hypothesis. (Do not use Minitab – manually calculate and use Chi-sq Tables to draw your conclusion). (5 Marks)
2. Make you recommendation on the preferred treatment method based on your findings. (2 Marks)

`