# Assessment Brief - Coursework

| | |
|---|---|
| **Academic Year** | **2023-24** |
| **Semester** | **2** |
| **Module Number** | **CM1606** |
| **Module Title** | **Computational Mathematics** |
| **Assessment Method** | **Coursework** |
| **Deadline (time and date)** | **7th April 2024, 11:59 p.m. (IST)** |
| **Submission** | **Assessment Dropbox in the Module Study Area in Campus Moodle.** |
| **Word Limit** | **N/A** |
| **Use of Generative Artificial Intelligence (AI) text** | **IS NOT authorised** |
| **Module Co-ordinator** | **Prashan Rathnayaka.** |

| What knowledge and/or skills will I develop by undertaking the assessment? |
|---|
| Data representation, analysis, and visualization; application of statistical models and hypothesis testing to real world problems. |
| **On successful completion of the assessment students will be able to achieve the following Learning Outcomes:**<br>1. Apply a range of statistical distribution models and hypothesis testing to real-world problems.<br>2. Represent, analyze and visualize data, in order to infer helpful insights about data collections. |
| **Please also refer to the Module Descriptor, available from the module Moodle study area.** |

## What is expected of me in this assessment?

### Task(s) – content

**Question 1:**

A bag contains 5 red balls, each labelled with a number 5, and 3 green balls, each labelled with a number 10. Another bag contains 2 blue balls, each labelled with a number 15, and 4 yellow balls, each labelled with a number 20. You randomly draw one ball from each bag. Let the random variable $X$ represent the sum of the numbers on the two balls:

a) Obtain the possible values of $X$.
b) Find the probability mass function (pmf) of $X$ in the usual format.
c) Calculate the expected value $E(X)$ and variance $Var(X)$.
d) Let $Y = 2X - 3$ be another random variable on the same sample space. Construct the pmf of $Y$.
e) Obtain the cumulative distribution function (cdf) of $Y$ and show in the usual format.
f) Calculate $P(Y = 37)$ using the cdf of $Y$.

**Question 2:**

Generate a random sample of 500 values from a normal distribution with a mean of 36 and a standard deviation (SD) of 8.

a) Create a histogram with 10 bins using the range (min to max).
b) Show the density curve over the histogram constructed above.
c) Comment on the histogram and the density curve with respect to the data generated.

**Question 3:**

Consider the following dataset.

| $X$ | 2 | 2.5 | 3 | 3.5 | 4 | 4.5 | 5 | 5.5 | 6 |
|---|---|---|---|---|---|---|---|---|---|
| $Y$ | 6 | 7.25 | 8 | 9.0625 | 10 | 11.0625 | 12.25 | 13.5625 | 15 |

a) Create a data frame using the above data and plot the data ($Y$ versus $X$).
b) Find the Pearson's correlation coefficient ( $r_{x,y}$).
c) Comment on the possible reasons for the value of ( $r_{x,y}$).
d) Find the Pearson's correlation coefficient ($R_{x,y}$)by considering only the last six pairs of data, and comment on the possible reasons why $r_{x,y}$ has changed.
e) Consider the new variables $X_1 = 2X - 1$ and $X_2 = X^2$ What can you say about the relationship between $r_{x,y}$ & $r_{x_1,y}$ and $r_{x,y}$ & $r_{x_2,y}$?

## What is expected of me in this assessment?

**Question 4:**

The built-in dataset "mtcars" provides information about various car models, including attributes like mpg (miles per gallon), hp (horsepower), and qsec (quarter-mile time in seconds). This dataset is a data frame with 32 rows and 11 variables.

a) Print the first ten rows of the "mtcars" dataset.
b) Print the 5-number summary of two variables: mpg (miles per gallon) and hp (horsepower).
c) Plot a scatter plot to visualize the association between mpg ($X$-axis) and hp ($Y$-axis). Comment on the type of association you observe.
d) Fit a linear regression model to predict mpg in terms of hp (horsepower), obtain the output and the fitted values, and show the regression line over the data.
e) Write down the fitted equation of the form $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$. By how many units does mpg increase or decrease when hp (horsepower) increases by one unit?
f) Plot the residuals versus $X$ (i.e., residual plot) to find out if the fitted straight line is adequate and to check for possible outliers. Comment on the plot.
g) Predict the mpg for a car with a horsepower of 110.

**Question 5:**

**Part A: Customer Spending at a Coffee Shop**

A local coffee shop has analysed its customer spending habits and found that the spending distribution follows a normal distribution with a mean of \$5.50 and a standard deviation of \$1.20. Calculate the following:

a) Find the $90^{th}$ percentile of customer spending.
b) Determine the $25^{th}$ percentile of customer spending.
c) Calculate the median value of customer spending.
d) What percentage of customers spend more than \$7.00?

**Part B: Disease Prevalence in a Population**

Suppose 5% of the population in a city is infected with a certain disease. Let $X$ be the number of infected individuals in a random sample of size 50 from this population.

a) Based on the problem setup, identify the most suitable probability distribution for $X$ and specify its name and parameters.
b) Calculate the probability that fewer than 3 individuals in the sample are infected.
c) Find the mean and variance of $X$.

## What is expected of me in this assessment?

d) If the infection rate in the city decreases to 2% and the sample size is increased to 200, identify the most suitable probability distribution for the new scenario. Provide reasoning for your choice.

**Question 6:**

You have a dataset of 40 exam scores: [82, 88, 75, 94, 90, 85, 78, 91, 86, 89, 92, 80, 87, 79, 84, 77, 83, 81, 76, 93, 88, 85, 89, 90, 82, 86, 75, 91, 79, 84, 78, 95, 88, 87, 93, 86, 82, 89, 90,80].

a) Estimate the sampling distribution of the sample mean for this dataset using bootstrap. Use 20,000 bootstrap samples.
b) Obtain the histogram of the bootstrap means.
c) Calculate a 90% bootstrap percentile confidence interval for the mean.
d) Check for the normality of the sampling distribution using a normal Q-Q plot.

***

**Task(s) - format**

***Instructions to Candidates:***

- *There are 6 questions in the assignment. Answer all questions.*
- *Each question is worth 10 marks with a breakdown as shown.*
- *The assignment should be written using R Markdown in RStudio and submitted as a PDF document. (i.e. write up your works as a .Rmd file, "knit" the results to a PDF file, and submit the PDF file to LMS.)*
- *Be sure to use R code for all your calculations and the statistical analysis.*
- *Between R chunks, you can use Latex to write up text or any mathematical equations (Using Latex is optional but can be a useful skill moving forward)*

## How will I be graded?

A grade will be provided for each criterion on the feedback grid which is specific to the assessment.

The overall grade for the assessment will be calculated using the algorithm below*.

| | |
|---|---|
| **A** | At least 50% of the subgrades to be at Grade A, at least 75% of the subgrades to be at Grade B or better, and normally 100% of the subgrades to be at Grade C or better. |
| **B** | At least 50% of the subgrades to be at Grade B or better, at least 75% of the subgrades to be at Grade C or better, and normally 100% of the subgrades to be at Grade D or better. |

| How will I be graded? | |
|---|---|
| **C** | At least 50% of the subgrades to be at Grade C or better, and at least 75% of the subgrades to be at Grade D or better. |
| **D** | At least 50% of the subgrades to be at Grade D or better, and at least 75% of the subgrades to be at Grade E or better. |
| **E** | At least 50% of the subgrades to be at Grade E or better. |
| **F** | Failing to achieve at least 50% of the subgrades to be at Grade E or better. |
| **NS** | Non-submission. |

*If the word count is above the specified word limit by more than 10% or the submission contains an excessive use of text within tables, the grade for the submission will be reduced to the next lowest grade.

| GRADE | A | B | • C | D | E | F |
|---|---|---|---|---|---|---|
| **DEFINITION / CRITERIA (WEIGHTING)** | **EXCELLENT** Outstanding Performance | **COMMENDABLE/VERY GOOD** Meritorious Performance | **GOOD** Highly Competent Performance | **SATISFACTORY** Competent Performance | **BORDERLINE FAIL** | **UNSATISFACTORY** Fail |

| GRADE | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| **DEFINITION / CRITERIA (WEIGHTING)** | **EXCELLENT** Outstanding Performance | **COMMENDABLE/VERY GOOD** Meritorious Performance | **GOOD** Highly Competent Performance | **SATISFACTORY** Competent Performance | **BORDERLINE FAIL** | **UNSATISFACTORY** Fail |
| **Question 1** | • The student provides a comprehensive list of possible values for random variable X and accurately explains how these values are obtained from the given information.<br>• The student accurately calculates the probability mass function (pmf) of X in the usual format, including the probabilities for each possible value of X.<br>• The student correctly calculates the expected value E(X) and variance Var(X) of random variable X.<br>• The student constructs the pmf of random variable Y = 2X - 3, accurately providing probabilities for each possible value of Y.<br>• The student accurately obtains the cumulative distribution function (cdf) of Y in the usual format and correctly shows it.<br>• The student correctly calculates P(Y = 37) using the cdf of Y and shows the correct result. | • The student provides a mostly accurate list of possible values for random variable X but may miss some values or explanations may be less clear.<br>• The student mostly accurately calculates the pmf of X but may make minor errors in probabilities or formatting.<br>• The student mostly correctly calculates E(X) and Var(X) but may make minor errors in calculations.<br>• The student constructs the pmf of Y with minor errors in probabilities or formatting.<br>• The student mostly accurately obtains the cdf of Y but may make minor errors in calculations or formatting.<br>• The student mostly correctly calculates P(Y = 37) but may make minor errors in calculations or formatting. | • The student lists some possible values for random variable X but may miss important values or explanations may lack clarity.<br>• The student attempts to calculate the pmf of X but makes significant errors in probabilities or formatting.<br>• The student attempts to calculate E(X) and Var(X) but makes significant errors in calculations.<br>• The student attempts to the pmf of Y but makes significant errors<br>• The student attempts to the the cdf of Y but makes significant errors in calculations or formatting.<br>• The student attempts to calculates P(Y = 37) but makes significant errors | • (Possible Values): The student attempts to identify possible values of X but makes much more errors, and explanations may more lack clarity.<br>• (pmf of X): The student attempts to calculate the pmf of X but makes much more errors in probabilities or formatting. The presentation is unclear, and the calculations are inaccurate.<br>• (E(X) and Var(X)): The student attempts to calculate E(X) and Var(X) but makes more errors in calculations. The explanation of these concepts may be vague or incorrect.<br>• (pmf of Y): The student attempts to construct the pmf of Y = 2X - 3 but makes more errors in probabilities or formatting. The response lacks clarity and accuracy.<br>• (cdf of Y): The student attempts to obtain the cdf of Y but makes more errors in calculations or presentation.<br>• (P(Y=37)): The student attempts to calculate P(Y=37) using the cdf of Y but makes more errors in calculations. | • (Possible Values): The student identifies some possible values of X but may miss some or provide incomplete explanations.<br>• (pmf of X): The student attempts to calculate the pmf of X but may make minor errors in probabilities or formatting.<br>• (E(X) and Var(X)): The student attempts to calculate E(X) and Var(X) but may make minor errors in calculations.<br>• (pmf of Y): The student attempts to construct the pmf of Y = 2X - 3 but may make minor errors in probabilities or formatting.<br>• (cdf of Y): The student attempts to obtain the cdf of Y but may make minor errors in calculations or presentation.<br>• (P(Y=37)): The student attempts to calculate P(Y=37) using the cdf of Y but may make minor errors in calculations | • (Possible Values): The student does not attempt to identify possible values of X.<br>• (pmf of X): The student does not attempt to calculate the pmf of X.<br>• (E(X) and Var(X)): The student does not attempt to calculate E(X) and Var(X).<br>• (pmf of Y): The student does not attempt to construct the pmf of Y = 2X - 3.<br>• (cdf of Y): The student does not attempt to obtain the cdf of Y.<br>• (P(Y=37)): The student does not attempt to calculate P(Y=37) using the cdf of Y. |

| GRADE | A | B | • C | D | E | F |
|---|---|---|---|---|---|---|
| **DEFINITION / CRITERIA (WEIGHTING)** | **EXCELLENT** Outstanding Performance | **COMMENDABLE/VERY GOOD** Meritorious Performance | **GOOD** Highly Competent Performance | **SATISFACTORY** Competent Performance | **BORDERLINE FAIL** | **UNSATISFACTORY** Fail |
| **Question 2** | • The student correctly generates a random sample of 500 values from a normal distribution with a mean of 36 and a standard deviation of 8.<br>• The student creates a histogram with 10 bins using the range (min to max) of the generated data and accurately displays it.<br>• The student accurately overlays the density curve over the histogram and provides appropriate labelling.<br>• The student provides insightful and comprehensive comments on the histogram and density curve, discussing their observations and insights about the generated data. | • The student generates a random sample of 500 values but may make minor errors in parameters or distribution type.<br>• The student creates a histogram with 10 bins but may make minor errors in displaying or labelling.<br>• The student overlays the density curve but may make minor errors in labelling or formatting.<br>• The student provides mostly insightful comments but may miss some key observations or insights. | • The student attempts to generate a random sample but makes significant errors in parameters or distribution type.<br>• The student attempts to create a histogram with 10 bins but makes significant errors in displaying or labelling.<br>• The student attempts to overlay the density curve but makes significant errors in labelling or formatting.<br>• . | • (Random Sample): The student generates a random sample but may make more errors in the process.<br>• (Histogram): The student attempts to create a histogram but may make more errors in binning, labelling, or presentation.<br>• (Density Curve): The student attempts to overlay the density curve but may make more errors in presentation.<br>• (Comments): The student provides basic comments on the histogram and density curve with limited insights. | • (Random Sample): The student attempt to generate a random sample but may make minor errors in the process.<br>• (Histogram): The student attempt to create a histogram but may make minor errors in binning or labelling.<br>• (Density Curve): The student attempt to create a density curve but may make minor errors in presentation.<br>• (Comments): The student attempt to provides comments on the histogram and density curve but may lack some insights or details. | • (Random Sample): The student does not attempt to generate a random sample.<br>• (Histogram): The student does not attempt to create a histogram.<br>• (Density Curve): The student does not attempt to overlay the density curve.<br>• (Comments): The student does not provide comments on the histogram and density curve. |

| GRADE | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| **DEFINITION / CRITERIA (WEIGHTING)** | **EXCELLENT** Outstanding Performance | **COMMENDABLE/VERY GOOD** Meritorious Performance | **GOOD** Highly Competent Performance | **SATISFACTORY** Competent Performance | **BORDERLINE FAIL** | **UNSATISFACTORY** Fail |
| **Question 3** | • The student correctly creates a data frame using the provided data and accurately plots the data (Y versus X) with appropriate labelling and formatting.<br>• The student correctly finds the Pearson's correlation coefficient ($r_{xy}$) for the given dataset and accurately presents the result.<br>• The student provides insightful comments on the possible reasons for the value of r_xy, discussing the relationship between the variables and the implications of the correlation coefficient.<br>• The student correctly finds the Pearson's correlation coefficient (R_xy) by considering only the last six pairs of data and accurately presents the result.<br>• The student provides insightful comments on the possible reasons why r_xy has changed when considering only the last six pairs of data, discussing the impact of data selection on the correlation coefficient.<br>• The student correctly explores the relationship between r_xy and the new variables X_1 = 2X - 1 and X_2 = X^2, providing insightful comments on how these transformations affect the correlation coefficient. | • The student creates a data frame and plots the data but may make minor errors in labelling or formatting.<br>• The student finds r_xy but may make minor errors in calculations or presentation.<br>• The student provides mostly insightful comments but may miss some key insights or explanations.<br>• The student finds R_xy but may make minor errors in calculations or presentation.<br>• The student provides mostly insightful comments but may miss some key insights or explanations.<br>• The student explores the relationship between r_xy, r_x1y, and r_x2y and provides mostly insightful comments but may miss some key insights or explanations. | • The student attempts to create a data frame and plot the data but makes significant errors in labelling or formatting.<br>• The student attempts to find r_xy but makes significant errors in calculations or presentation.<br>• The student provides basic comments on the possible reasons for r_xy with limited insights.<br>• The student attempts to find R_xy but makes significant errors in calculations or presentation.<br>• The student provides basic comments on the reasons for the change in r_xy with limited insights.<br>• The student explores the relationship between r_xy, r_x1y, and r_x2y with basic comments and limited insights. | • (Data Frame and Plot): The student attempts to create a data frame and plot the data (Y versus X) but makes more errors in formatting or presentation.<br>• (Pearson's correlation coefficient): The student attempts to calculate the Pearson's correlation coefficient (r_xy) but makes more errors in calculations.<br>• (Comments - r_xy): The student provides limited comments on the possible reasons for the value of r_xy with limited insights.<br>• (Pearson's correlation coefficient - Last six pairs): The student attempts to calculate the Pearson's correlation coefficient (R_xy) using only the last six pairs of data but may make significant errors in calculations.<br>• (Relationship - X_1, X_2): The student attempts to analyze the relationship between r_xy and r_x1y, as well as r_xy and r_x2y, but with limited accuracy or clarity. | • (Data Frame and Plot): The student attempts to create a data frame and plot the data (Y versus X) but may make minor errors in formatting or presentation.<br>• (Pearson's correlation coefficient): The student attempts to calculate the Pearson's correlation coefficient (r_xy) but may make minor errors in calculations.<br>• (Comments - r_xy): The student provides comments on the possible reasons for r_xy but may lack some insights.<br>• (Pearson's correlation coefficient - Last six pairs): The student attempts to calculate the Pearson's correlation coefficient (R_xy) using only the last six pairs of data but may make minor errors in calculations.<br>• (Relationship - X_1, X_2): The student attempts to analyze the relationship between r_xy and r_x1y, as well as r_xy and r_x2y, with some accuracy. | • (Data Frame and Plot): The student does not attempt to create a data frame or plot the data (Y versus X).<br>• (Pearson's correlation coefficient): The student does not attempt to calculate the Pearson's correlation coefficient (r_xy).<br>• (Comments - r_xy): The student does not provide comments on the possible reasons for r_xy.<br>• (Pearson's correlation coefficient - Last six pairs): The student does not attempt to calculate the Pearson's correlation coefficient (R_xy) using only the last six pairs of data.<br>• (Relationship - X_1, X_2): The student does not attempt to analyze the relationship between r_xy and r_x1y, as well as r_xy and r_x2y. |

| GRADE | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| **DEFINITION / CRITERIA (WEIGHTING)** | **EXCELLENT** Outstanding Performance | **COMMENDABLE/VERY GOOD** Meritorious Performance | **GOOD** Highly Competent Performance | **SATISFACTORY** Competent Performance | **BORDERLINE FAIL** | **UNSATISFACTORY** Fail |
| **Question 4** | • The student correctly prints the first ten rows of the "mtcars" dataset, providing the data accurately.<br>• The student correctly calculates and presents the 5-number summary (minimum, 1st quartile, median, 3rd quartile, maximum) for two variables: mpg (miles per gallon) and hp (horsepower) from the "mtcars" dataset.<br>• The student correctly plots a scatter plot to visualize the association between mpg (X-axis) and hp (Y-axis) from the "mtcars" dataset and provides insightful comments on the type of association observed.<br>• The student correctly fits a linear regression model to predict mpg in terms of hp (horsepower), obtains the output and fitted values, and shows the regression line over the data.<br>• The student writes down the fitted equation of the form $\hat{Y} = \beta_0 + \beta_1 X$, and correctly explains by how many units mpg increases or decreases when hp (horsepower) increases by one unit.<br>• The student correctly plots the residuals versus X (i.e., residual plot) and provides insightful comments on the adequacy of the fitted straight line and checks for possible outliers.<br>• The student correctly predicts the mpg for a car with a horsepower of 110. | • The student prints the first ten rows but may make minor errors in formatting or presentation.<br>• The student calculates the 5-number summary but may make minor errors in formatting or presentation.<br>• The student plots the scatter plot but may make minor errors in labelling or presentation and provides mostly insightful comments.<br>• The student fits the linear regression model but may make minor errors in output presentation or interpretation.<br>• The student writes down the fitted equation and explains how mpg changes with hp, but may make minor errors in explanation.<br>• The student plots the residuals and provides mostly insightful comments but may miss some key observations or insights.<br>• The student predicts the mpg for a car with a horsepower of 110 but may make minor errors in calculations or presentation. | • The student attempts to print the dataset but makes significant errors in formatting or presentation.<br>• The student attempts to calculate the 5-number summary but makes significant errors in formatting or presentation.<br>• The student attempts to plot the scatter plot but makes significant errors in labelling or presentation and provides basic comments.<br>• The student attempts to fit the linear regression model but makes significant errors in output presentation or interpretation.<br>• The student writes down the fitted equation and explains how mpg changes with hp, but makes vague or unclear explanations.<br>• The student attempts to plot the residuals and provides basic comments with limited insights.<br>• The student attempts to predict the mpg for a car with a horsepower of 110 but makes significant errors in calculations or presentation. | • The student attempts to print the first ten rows but makes major errors in formatting or presentation.<br>• The student attempts to calculate the 5-number summary but makes major errors in formatting or presentation.<br>• The student attempts to plot the scatter plot but makes major errors in labelling or presentation and provides vague or unclear comments.<br>• The student attempts to fit the linear regression model but makes major errors in output presentation or interpretation.<br>• The student attempts to write down the fitted equation and explain how mpg changes with hp but provides inadequate or incorrect explanations.<br>• The student attempts to plot the residuals and provides inadequate comments with no meaningful insights.<br>• The student attempts to predict the mpg for a car with a horsepower of 110 but makes significant errors in calculations or presentation. | • The student provides an incorrect or incomplete printout of the dataset.<br>• The student provides an incorrect or incomplete 5-number summary.<br>• The student provides an incorrect or incomplete scatter plot or does not comment on the association.<br>• The student provides an incorrect or incomplete linear regression model fitting.<br>• The student provides incorrect or vague explanations of how mpg changes with hp.<br>• The student provides incorrect or incomplete comments on the residuals.<br>• The student provides an incorrect or incomplete prediction of mpg for a car with a horsepower of 110. | • The student does not attempt to print the dataset.<br>• The student does not attempt to calculate the 5-number summary.<br>• The student does not attempt to plot the scatter plot or provide any comments.<br>• The student does not attempt to fit the linear regression model.<br>• The student does not attempt to write down the fitted equation or explain how mpg changes with hp.<br>• The student does not attempt to plot the residuals or provide any comments.<br>• The student does not attempt to predict mpg for a car with a horsepower of 110. |

| GRADE | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| **DEFINITION / CRITERIA (WEIGHTING)** | **EXCELLENT** Outstanding Performance | **COMMENDABLE/VERY GOOD** Meritorious Performance | **GOOD** Highly Competent Performance | **SATISFACTORY** Competent Performance | **BORDERLINE FAIL** | **UNSATISFACTORY** Fail |
| **Question 5** | • The student accurately calculates the 90th percentile, 25th percentile, and median of customer spending, providing correct values.<br>• The student correctly calculates the percentage of customers spending more than $7.00.<br>• The student identifies the probability distribution for X (number of infected individuals) as the binomial distribution, specifies its parameters, and accurately calculates the probability of fewer than 3 individuals in the sample being infected.<br>• The student correctly finds the mean and variance of X (number of infected individuals) and provides accurate results.<br>• The student identifies the new probability distribution for the decreased infection rate and larger sample size scenario (X ~ Binomial) and provides reasoning for this choice. | • The student calculates the percentiles and median but may make minor errors in calculations.<br>• The student calculates the percentage of customers spending more than $7.00 but may make minor errors in calculations.<br>• The student identifies the probability distribution for X and specifies its parameters but may make minor errors in calculations.<br>• The student finds the mean and variance of X but may make minor errors in calculations.<br>• The student identifies the new probability distribution for the changed scenario but may provide limited reasoning. | • The student attempts to calculate percentiles and median but makes significant errors in calculations.<br>• The student attempts to calculate the percentage of customers spending more than $7.00 but makes significant errors in calculations.<br>• The student attempts to identify the probability distribution for X but makes significant errors in calculations or specification.<br>• The student attempts to find the mean and variance of X but makes significant errors in calculations.<br>• The student provides vague or unclear reasoning for the choice of probability distribution in the changed scenario. | • The student attempts to calculate percentiles and median but makes major errors in calculations.<br>• The student attempts to calculate the percentage of customers spending more than $7.00 but makes major errors in calculations.<br>• The student attempts to identify the probability distribution for X but makes major errors in calculations or specification.<br>• The student attempts to find the mean and variance of X but makes major errors in calculations.<br>• The student provides vague or unclear reasoning for the choice of probability distribution in the changed scenario. | • The student attempts to calculate percentiles and median but makes significant errors in calculations.<br>• The student attempts to calculate the percentage of customers spending more than $7.00 but makes significant errors in calculations.<br>• The student attempts to identify the probability distribution for X but makes significant errors in calculations or specification.<br>• The student attempts to find the mean and variance of X but makes significant errors in calculations.<br>• The student provides incomplete or vague reasoning for the choice of probability distribution in the changed scenario. | • The student does not attempt to calculate percentiles, median, or any other statistics.<br>• The student does not attempt to calculate the percentage or provide any analysis.<br>• The student does not attempt to identify the probability distribution or provide any analysis.<br>• The student does not attempt to find the mean, variance, or provide any analysis.<br>• The student provides no reasoning or analysis for the choice of probability distribution in the changed scenario. |

| GRADE | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| **DEFINITION / CRITERIA (WEIGHTING)** | **EXCELLENT** Outstanding Performance | **COMMENDABLE/VERY GOOD** Meritorious Performance | **GOOD** Highly Competent Performance | **SATISFACTORY** Competent Performance | **BORDERLINE FAIL** | **UNSATISFACTORY** Fail |
| **Question 6** | • The student correctly estimates the sampling distribution of the sample mean using bootstrap with 20,000 bootstrap samples.<br>• The student accurately obtains the histogram of the bootstrap means.<br>• The student correctly calculates a 90% bootstrap percentile confidence interval for the mean.<br>• The student checks for normality of the sampling distribution using a normal Q-Q plot and provides an accurate interpretation. | • The student estimates the sampling distribution using bootstrap but may make minor errors.<br>• The student obtains the histogram of the bootstrap means but may make minor errors in presentation.<br>• The student calculates a 90% bootstrap percentile confidence interval but may make minor errors.<br>• The student checks for normality using a normal Q-Q plot but may provide limited interpretation. | • The student attempts to estimate the sampling distribution using bootstrap but makes significant errors.<br>• The student attempts to obtain the histogram but makes significant errors in presentation.<br>• The student attempts to calculate a confidence interval but makes significant errors.<br>• The student attempts to check for normality using a Q-Q plot but provides significant vague or not significant enough interpretation. | • The student attempts to estimate the sampling distribution using bootstrap but makes much more significant errors.<br>• The student attempts to obtain the histogram but makes much more significant errors in presentation.<br>• The student attempts to calculate a confidence interval but makes much more significant errors.<br>• The student attempts to check for normality using a Q-Q plot but provides unclear interpretation. | • The student provides an incomplete or inaccurate attempt to estimate the sampling distribution using bootstrap.<br>• The student provides an incomplete or inaccurate attempt to obtain the histogram.<br>• The student provides an incomplete or inaccurate attempt to calculate a confidence interval.<br>• The student provides an incomplete or inaccurate attempt to check for normality using a Q-Q plot. | • The student does not attempt to estimate the sampling distribution or provide any analysis.<br>• The student does not attempt to obtain the histogram or provide any analysis.<br>• The student does not attempt to calculate a confidence interval or provide any analysis.<br>• The student does not attempt to check for normality using a Q-Q plot or provide any analysis. |

***Coursework received late will be regarded as a non-submission (NS) and one of your assessment opportunities will be lost.***

ROBERT GORDON UNIVERSITY ABERDEEN

## What else is important to my assessment?

### What is the Assessment Word Limit Statement?

It is important that you adhere to the Word Limit specified above. The Assessment Word Limit Statement can be found in Appendix 2 of the RGU Assessment Policy. It provides detail on the purpose, setting and implementation of wordage limits; lists what is included and excluded from the word count; and the penalty for exceeding the word count.

**What's included in the word count?**

The table below lists the constituent parts which are included and excluded from the word limit of a Coursework; more detail can be found in the full Assessment Word Limit Statement. Images will not be allowed as a mechanism to circumvent the word count.

| Excluded | Included |
|---|---|
| Cover or Title Page | Main Text e.g. Introduction, Literature Review, Methodology, Results, Discussion, Analysis, Conclusions, and Recommendations |
| Executive Summary (Reports) or Abstract | Headings and subheadings |
| Contents Page | In-text citations |
| List of Abbreviations and/or List of Acronyms | Footnotes (relating to in-text footnote numbers) |
| List of Tables and/or List of Figures | Quotes and quotations written within "…" |
| Tables – mainly numeric content | Tables – mainly text content |
| Figures | |
| Reference List and/or Bibliography | |
| Appendices | |
| Glossary | |

**What are the penalties?**

The grade for the submission will be reduced to the next lowest grade if:

- The word count of submitted work is above the specified word limit by more than 10%.
- The submission contains an excessive use of text within Tables or Footnotes.

## What else is important to my assessment?

### What is plagiarism?

Plagiarism is "the practice of presenting the thoughts, writings or other output of another or others as original, without acknowledgement of their source(s) at the point of their use in the student's work. All materials including text, data, diagrams or other illustrations used to support a piece of work, whether from a printed publication or from electronic media, should be appropriately identified and referenced and should not normally be copied directly unless as an acknowledged quotation. Text, opinions or ideas translated into the words of the individual student should in all cases acknowledge the original source" (RGU 2022).

### What is collusion?

"Collusion is defined as two or more people working together with the intention of deceiving another. Within the academic environment this can occur when students work with others on an assignment, or part of an assignment, that is intended to be completed separately" (RGU 2022).

For further information please see Academic Integrity.

### What if I'm unable to submit?

- The University operates a Fit to Sit Policy which means that if you undertake an assessment then you are declaring yourself well enough to do so.
- If you require an extension, you should complete and submit a Coursework Extension Form. This form is available on the RGU Student and Applicant Forms page.
- Further support is available from your Course Leader.

### What additional support is available?

- RGU Study Skills provide advice and guidance on academic writing, study skills, maths and statistics and basic IT.
- RGU Library guidance on referencing and citing.
- The Inclusion Centre: Disability & Dyslexia.
- Your Module Coordinator, Course Leader and designated Personal Tutor can also provide support.

### What are the University rules on assessment?

The University Regulation 'A4: Assessment and Recommendations of Assessment Boards' sets out important information about assessment and how it is conducted across the University.