



National
Qualifications
ADDITIONAL SPECIMEN

S844/76/11

**Applications of Mathematics
Data booklet**

Date — Not applicable

Duration — 2 hours 30 minutes

Pre-release material

This booklet will be issued to centres in advance of the date of examination.

Candidates will be issued with a clean copy of this booklet. Copies will be issued at the start of the examination session and collected at the end of the session. Candidates must not take their own copies of this booklet into the examination.

Centres should ensure that candidates are familiarised with any contexts and information contained in this booklet in preparation for the examination.

Some examination questions will be based on this material.

This booklet contains some helpful R commands. There is no other material for this examination.



* S 8 4 4 7 6 1 1 *

Some helpful R commands

Entering data to R Studio

To read in data from an Excel csv file called `excel_data.csv` to R Studio and name it `mydata`, first use the drop down menus in R Studio **Session > Set Working Directory > Choose Directory** to indicate the location of `excel_data.csv` on your computer. The following code will then read the data in to R Studio:

```
mydata<-read.csv("excel_data.csv")
attach(mydata) — this adds the variable names
```

At the end of the analysis remember to use `detach(mydata)` to disassociate the variable names.

(a) Graphics

`hist(X, col="yellow", main="Histogram of X (units)")` — this produces a histogram of the variable named 'X'

`plot(X, Y, xlab="x-axis label", ylab="y-axis label", main="Scatterplot of Y on X", pch=21, bg="black")` — produces a scatterplot of X vs Y with black dots of the size specified by 'pch'

`pie(table(X), main="Title")` — this gives a simple pie chart of the categories in variable X with the specified title

`barplot(table(X), main="title", xlab="x-axis label", col="orange")` — this gives a bar chart of the categories in the variable X with the required title, axis labels and colour

`boxplot(Y)` — produces a boxplot of the numerical variable Y

(b) Descriptive Statistics

`mean(X)` — computes the mean of X

`sd(X)` — computes the standard deviation of X

`summary(X)` — computes the mean, median, minimum, maximum and upper and lower quartiles

`table(X)` — computes the number of observations in each level of the categorical variable X

`prop.table(table(X))` — returns the proportion of observations in each level of the categorical variable X

`prop.table(table(X))*100` — returns the percentage of observations in each level of the categorical variable X

`table(X, Y)` — produces a cross-tabulation between the two categorical variables X and Y

(c) Correlation and Regression

`cor.test(age, length)` — computes the correlation between X and Y and performs a test of the null hypothesis of zero correlation

`lm(Y~X)` — fits a linear regression line to the data (`lm` command stands for linear model)

`abline(lm(Y~X))` — produces a scatterplot with the least squares linear regression line superimposed on the data

`summary(lm(Y~X))` — displays the coefficient of determination (r-squared)

`predict(lm(Y ~ X), newdata=data.frame(X=C), interval = "pred")` — computes the predicted value of Y when X=C along with a 95% prediction interval

(d) Hypothesis Testing

`t.test(X, Y)` — performs a two sample t-test between X and Y

`t.test(X, Y, paired=TRUE)` — performs a paired t-test between X and Y

`prop.test(x = c(a, b), n = c(n1, n2))` — performs a 2-sample test for equality of proportions with continuity correction

[END OF DATA BOOKLET]