

# PPSAS205

## Introduction and Modelling in R

Duration: 2 Hours

Marks: 60

Note:

- 1) The candidate has the option to either attempt question 3A or question 3B. Rest all questions are mandatory.
- 2) Numbers on the right indicate full marks.
- 3) The candidate will be provided with the formula sheet and graph papers [if required] for the examination.
- 4) Use of approved scientific calculators is allowed.

Q1. Attempt the following:

15 Marks

- A. A machine that fills bottles is known to have a mean filling amount of 125 gms and a standard deviation of 20 gms. A quality control manager took a random sample of filled bottles and found the sample mean to be 130. The quality manager assumed the sample must not be representative. Validate whether the conclusion is correct or not.

```
sample_data <- c(129, 132, 127, 134, 131, 128, 130, 133, 126, 128,  
127, 129, 133, 132, 130, 129, 135, 128, 131, 129,  
130, 131, 132, 128, 126, 129, 130, 133, 130, 131)
```

5 Marks

- B. In a sample of 25 observations from a normal distribution with mean 98.6 and standard deviation 17.2

What is  $P(92 < X < 102)$

3 Marks

Find the probability if given a sample of 36

2 Marks

- C. A manufacturer of pet foods was wondering whether cat owners and dog owners reacted differently to premium pet foods. They commissioned a consumer survey that yielded the following data.

Pet	Owners Surveyed	Number Using Premium Food
Cat	280	152
Dog	190	81

Is it reasonable to conclude at significance level 0.02 that cat owners are more likely than dog owners to feed their pets premium food?

5 Marks

**Q2. Attempt the following:****15 Marks**

- A.** A study compared the effects of four 1-month point-of-purchase promotions on sales. The unit sales for five stores using all four promotions in different months follow:

Free Sample	78	87	81	89	85
One gift Pack	94	91	87	90	88
Cents Off	73	78	69	83	76
Refund by email	79	83	78	69	81

Check whether there is an effect on promotions on sales

- i. Create a matrix "sale" similar to the above table **1 Mark**
- ii. Print sales matrix **1 Mark**
- iii. Test the hypothesis to check the effect of promotions on sales **3 Marks.**

- B.** Create a contingency table for the below data:

Gender <- c("Male", "Female", "Male", "Female", "Male", "Male", "Female", "Female")

Smoking <- c("Yes", "No", "No", "Yes", "Yes", "No", "Yes", "No") **5 Marks**

- C.** For the below mentioned data, check whether the null hypothesis is true that there is a strong correlation between them? Perform a correlation test on x and y

x <- c(3, 8, 7, 6, 10, 5, 2, 9, 4, 1)

y <- c(5, 12, 11, 9, 15, 8, 3, 14, 6, 2)

**5 Marks**

**Q3 Attempt Question 3A or question 3B. Rest all questions**

**Q3 A.** Sales of major appliances vary with the new housing market: When new home sales are good so are the sales of dishwashers, washing machines, driers, refrigerators. A trade association compiled the following data ( in thousands of units) on major appliance sales and housing:

**30 Marks**

Housing	Appliance Sales
2	5
2.5	5.5
3.2	6
3.6	7
3.3	7.2
4	7.7
4.2	8.4
4.6	9
4.8	9.7
5	10

- i. Develop an equation for the relationship between appliances sales and housing **7 Marks.**
- ii. Interpret the slope of the regression line **3 Marks.**
- iii. Compute and interpret the standard error of estimate **8 Marks.**
- iv. Housing next year may be beyond recorded range, estimates as high as 8 million units have been predicted. Compute an approximate 90% prediction interval of appliance sales, based on previous data and new prediction of housing. **12 Marks**

**OR**

**Q3 B**

The mtcars dataset contains information on various car models. We are interested in exploring the relationship between miles per gallon (mpg) and several other variables. Answer the following sub-questions using a glm in R:

- i. Load the mtcars dataset into R and create a new binary variable high\_mpg that indicates whether a car has an mpg value greater than or equal to 20 (1 if yes, 0 if no). **10 Marks**
- ii. Fit a glm model to predict high\_mpg using the variables wt, hp, drat, and qsec. **5 Marks**
- iii. Interpret the coefficients of the model. **7 Marks**
- iv. Use the model to predict the probability that a car with a weight of 3.5, horsepower of 120, drat of 3.9, and qsec of 17 will have an mpg value greater than or equal to 20. **8 Marks**