Questions 3

Dr. Roman Belavkin

BIS4435

Question 1

What is a model and its error, and how can they be defined in terms of information?

Question 2

What are the linear models, and why are they called 'linear'?

Question 3

What are the data–driven models, and what does the data represent?

Question 4

Why plotting the data on a chart can be useful?

Question 5

How can the errors of the model be measured? Why is it important for the model?

Question 6

Compute linear mean-square model for the following set of data:

Monthly Income (£K)	Home Owner
2	0
1	0
6	1

Question 7

Suppose that a multiple regression model using $m = 6$ input variables should be created based on some data. What is the minimum number of cases the dataset should have?

Question 8

Consider the following data:

Monthly Income (£K)	Monthly Expenses (£K)	Home Owner	Credit Score
2	1	0	3
1	2	0	1
6	2	1	5
3	1	1	4
3	2	0	2

Suppose you are building a linear mean-square model based on this data to predict the customers' credit scores. Let us denote the three input variables (the first three columns of the table) as x_1, x_2 and x_3, while the output variable as y. Answer the following questions:

a) Does this dataset contain sufficient number of cases to build the model?

b) The regression coefficients for this model are $b_1 = 0,69$, $b_2 = -1,31$ and $b_3 = 0,56$. Write the linear equation for $f(x_1, x_2, x_3)$.

c) Using the model, compute the credit score for the following customer:

Monthly Income (£K)	Monthly Expenses (£K)	Home Owner	Credit Score
5	3	1	?

Question 9

How to test whether a model is good in forecasting?

Question 10

Compute correlation for the following set of data:

Monthly Expenses (£K)	Home Owner
1	0
2	1
1	1
2	0

What does the value of the correlation mean in this case?

Question 11

What are the outliers? Why is the mean-square model sensitive to outliers? What does it mean for the mean-square model?

Question 12

What are the main differences between the mean-square and the mean absolute models?