



MACHAKOS UNIVERSITY

University Examinations 2018/2019

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF PHYSICAL SCIENCES

APRIL SESSION EXAMINATION FOR

MASTER OF SCIENCE (CHEMISTRY)

SCU 801: SCIENTIFIC DATA ANALYSIS

DATE: 14/8/2019

TIME: 8:30-11:30AM

INSTRUCTIONS

Question ONE is compulsory. Answer any TWO questions.

QUESTION ONE (30 MARKS)

- a) Graphically, illustrate the physical meaning of convolution by using a slit function. (3 marks)
- b) Most statistical tests have the same sequence of events. List any FOUR steps used in statistical testing. (4 marks)
- c) Using an arbitrary assumption of detector responses (y-values) against mass or concentration of the analyte, show the various types of correlations graphically. (4 marks)
- d) Give the meaning of each of the following: White, Black and Grey Systems (3 marks)
- e) A proposed new method for the determination of sulfate in an industrial waste effluent is compared with an existing method, giving the following results. Is there a significant difference between the precisions of the two methods? (3 marks)

Method	Mean/g dm ⁻³	No. of replicates	No. of degrees of freedom	s/mg dm ⁻³
Existing	72	8	7	3.38
New	72	8	7	1.50

- f) Describe the application of standard deviation as a measure of precision in analytical measurements (4 marks)

- g) Distinguish between parametric and non-parametric tests in the context of scientific data analysis. (2 marks)
- h) An unbiased analytical system is used to measure an analyte whose true value for the parameter measured is 100 units. If the standard deviation of repeated measurements is 0.80 units, what fraction of measurements will be further than one unit from the true value? Note that this fraction will also be the probability that a single measurement will produce a test result which is either less than 99 or greater than 101, when the material is measured. What fraction of measurements will be further than 1.5 units from the true value? Give the answers to two decimal places. (3 marks)
- i) Distinguish between Internal standardization and Internal normalization (4 marks)

QUESTION TWO (15 MARKS)

- a) Discuss the following data smoothing and filtering methods (10 marks)
- Moving window average
 - Savitsky-Golay filter
- b) Give TWO examples of test statistics used in laboratory testing and their corresponding formulae. Clearly define the all the terms in the formulae. (5 marks)

QUESTION THREE (15 MARKS)

- a) The main aim of data compression is to represent information from a source instrument in the best way using minimal bits. Show how this is achieved using principal component analysis (PCA) as a tool chemometrics. (4 marks)
- b) A study was conducted to investigate the influence of various system parameters on the performance of a programmed temperature vaporizer (PTV) operated in the solvent split mode. The purpose was to develop a GC method of analyzing trace aroma compounds in wine, without the need for prior extraction. One phase of the investigation involved studying three factors each at two levels; A: Injection Volume (100 or 200 μL), B: Solvent Elimination Flow Rate (200 or 400 mL min^{-1}), C: PTV Initial Temperature (20 or 40 $^{\circ}\text{C}$). The eight design points were replicated twice, giving a total of 16 observations. The order in which the 16 runs were carried out was randomized. One of the responses studied was the total peak area given by a synthetic mixture of aroma compounds in ethanol– water (12:88 v/v). The results (absolute peak areas) are shown in standard order in Table below.

Table Design matrix and results for PTV development study

<i>Design point</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>Results</i>	
1	–	–	–	37.5	24.6
2	+	–	–	149.4	146.5
3	–	+	–	24.9	28.8
4	+	+	–	117.3	114.0
5	–	–	+	27.5	32.2
6	+	–	+	129.5	136.6
7	–	+	+	23.7	22.4
8	+	+	+	111.0	112.6

- i) Graphically show how average results may be displayed in a 2^3 -factorial design (3 marks)
- ii) Describe, with appropriate calculations, how various three-factor interactions, corresponding to all possible pairs of factors, may be estimated. (8 marks)

QUESTION FOUR (15 MARKS)

- a) Write down the expression used to determine the correlation coefficient, r , which indicates the degree of linearity between x and y variables in statistical data calibration (3 marks)
- b) A calibration graph was prepared as part of a validation procedure for a new method to determine an active constituent of a sun cream by UV spectrophotometry. The following data were obtained:

Analyte conc. (mg cm^{-3})	0	20	40	60	80	100	120
UV absorbance at 325 nm	0.095	0.227	0.409	0.573	0.786	0.955	1.123

 Check the data above for linearity (8 marks)
- c) Illustrate, using appropriate diagrams, how Fourier transformation is used in data compression (4 marks)

QUESTION FIVE (15 MARKS)

- a) Discuss the Kalman method as optimal linear recursive estimation as a data filtering procedure (4 marks)
- b) Describe using an example the steps followed in hypothesis. (5 marks)
- c) Discuss the role of Shewhart charts in quality assurance and describe how they are constructed (6 marks)