Silence Please

Do Not turn over this page until advised to by the Invigilator

CIT Semester 1 Examinations 2018/19

Note to Candidates:	Check the Programme Title and the Module Description to ensure that you have received the correct examination. If in doubt please contact an Invigilator.
Module Title:	**Statistical Quality Control**
Module Code:	**STAT8003**
Programme Title(s):	BSc Hons Anlytcl Chmstry QA Y4
Block Code(s):	**SCHQA_8_Y4**
External Examiner(s):	**Dr. Katarina Domijan**
Internal Examiner(s):	Dr. Catherine Palmer
Instructions:	Answer any three questions. All questions carry equal marks.
Duration:	2 Hours
Required Items:	Calculator, Murdoch & Barnes Tables, Log/Formulae Tables

1.

(a) The following two acceptance sampling plans are to be compared:

Plan A: take a random sample of 100 items from the batch. If the sample contains 4 or fewer defectives, accept the batch; otherwise reject it.

Plan B: take a sample of 50 items from the batch. If the sample contains 2 defectives or less, accept the batch. If it contains more than 4 defectives, reject the batch. Otherwise, take a second sample of 50 items. Accept the batch if the total number of defectives is 4 or less, otherwise reject it.

Which of the plans is more likely to accept a batch that is 2% defective?

(9 marks)

(b) The concentration of calcium in a particular brand of bottled water is normally distributed with a mean of 80 mg/L and a standard deviation of 20 mg/L.

(i) Find the percentage of bottles with a calcium concentration less than 115 mg/L.

(ii) 5% of bottles have a calcium concentration greater than x mg/L. What is the value of x?

(iii) Suppose a sample of 10 bottles is chosen at random and the mean concentration of calcium (mg/L) is calculated. What is the probability that the mean concentration of the sample is less than 75 mg/L?

(7 marks)

(c) The temperature of a solution is being monitored on an hourly basis. For each hour, a sample of five measurements is taken and the mean and range are calculated for each sample. After twenty hours, the following summary data are obtained: $\bar{\bar{x}} = 45.2$°C and $\bar{R} = 0.95$°C. What are the 3-sigma control limits for $\bar{x}$ and R?

(5 marks)

(d) The following table shows the number of defectives observed in samples of 100 products checked over 10 shifts. Use the data to calculate the 3-sigma control limits for the proportion of nonconforming items and comment on whether the process appears to be in control.

Sample Number	1	2	3	4	5	6	7	8	9	10
Number of Defectives	0	2	1	5	9	4	3	7	2	1

(4 marks)

2.

(a) The amount of protein in a particular brand of infant formula is to be determined. Ten replicate samples weighing 10 grams were selected at random and the amount of protein (grams) measured for each replicate is shown below.

1.40, 1.44, 1.46, 1.38, 1.39, 1.40, 1.47, 1.37, 1.48, 1.42

(i) Produce a 95% confidence interval for the mean amount of protein per gram and interpret the confidence interval.

(ii) Produce a 95% confidence interval for the variance of the amount of protein per gram.

(5 marks)

(b) A new chromatographic method for determining the amount of protein in a sample has been developed. A laboratory would like to compare the new method (method A) to the method currently in use (method B). Eight replicate samples are chosen at random and the amount of protein per gram for each replicate is measured using both method A and method B, the results are shown in the table below. Carry out appropriate tests of hypothesis to determine whether the mean results obtained by the methods differ significantly.

Method A (mg)	0.145	0.143	0.151	0.147	0.149	0.141	0.143	0.139
Method B (mg)	0.142	0.142	0.145	0.147	0.136	0.147	0.143	0.145

(13 marks)

(c) The laboratory investigating the new chromatographic method for determining the amount of protein in a sample will only consider using method A if the variance associated with the method is below 0.005mg. Using the data for method A (shown in part (b)) perform the necessary analysis and state your conclusions.

(7 marks)

3.

(a) The table below shows the results obtained in an investigation into the stability of a fluorescent reagent stored under three different conditions. The values given are the fluorescence signals in relative light units (RLU) from solutions of equal concentration. Four replicate measurements were made on each sample.

Replicate	Freshly prepared	Stored for 1 hour in the dark	Stored for 1 hour in bright light
1	102	100	93
2	101	97	91
3	105	95	94
4	99	97	93

Carry out a one-way analysis of variance on this set of data to determine whether there is a difference in the mean strength of fluorescence signal between the reagents that were stored in different conditions. (12 marks)

(b) A factorial experiment is carried out to determine how the concentration of alcohol in the output of steam from a distillation process is affected by the reboiler temperature (°F) and the condenser temperature (°F). Three replicates of a 2^2 design produced the following results.

Reboiler temp (°F)	Condenser temp (°F)	Alcohol concentration (%)
160	100	9.9, 10.2, 10.0
160	110	9.7, 9.4, 9.5
170	100	11, 11.5, 11.7
170	110	10.2, 10.8, 10.5

(i) Estimate the main effects and the interaction effect.

(ii) Test the significance of the effect estimates calculated in (i).

(iii) Draw the interaction plot, and comment. (13 marks)

4.

The concentration of phytic acid (mg/L) in 10 samples was determined by a new catalytic fluorimetric (CF) method and the results were compared with those obtained using an established extraction photometric (EP) method. The concentrations (mg/L) are shown in the table below.

Concentration using EP (mg/L) (x)	1.37	2.87	3.56	1.89	1.50	0.80	1.19	2.40
Concentration using CF (mg/L) (y)	1.51	2.65	3.62	1.95	1.57	0.74	1.25	2.32

$\sum x = 15.58$ $\sum y = 15.61$ $\sum xy = 36.2494$ $\sum x^2 = 36.4256$ $\sum y^2 = 36.1669$

(i) Construct a scatter plot for these data, and comment on the level of correlation. (5 marks)

(ii) Find the equation of the regression line of concentration using the CF method on concentration using the EP method. (6 marks)

(iii) Produce the analysis of variance table for the regression, and say what conclusions you draw from it. (7 marks)

(iv) Using relevant information from the ANOVA table, find the value of the correlation coefficient. (2 marks)

(v) Find 95% confidence limits for the *predicted* concentration of phytic acid using the new CF method when the concentration measured using the EP method is 1.5 mg/L. (5 marks)

Statistical Formulae

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

$$t_{n-1} = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

$$\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}};$$

$$\bar{X} \pm t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}.$$

$$z = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}}}$$

$$t_{n_1+n_2-2} = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s_p^{\ 2} = \frac{(n_1 - 1)s_1^{\ 2} + (n_2 - 1)s_2^{\ 2}}{n_1 + n_2 - 2}$$

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$$

$$\frac{s_1^{\ 2}}{s_2^{\ 2}} \sim F^{n_1 - 1}_{n_2 - 1}$$

$$\left(\frac{(n-1)s^2}{\chi^2_{n-1, \alpha/2}}, \frac{(n-1)s^2}{\chi^2_{n-1, 1-\alpha/2}} \right)$$

$$SE(p) = \sqrt{\frac{p_0(1 - p_0)}{n}}$$

ANOVA

1. **One-way model**: $y_{ij} = \mu + \tau_i + \varepsilon_{ij},\ i = 1,2,...,a,\ j = 1,2,....,n.$

Total SS: $\sum \sum y^2{}_{ij} - \frac{y_{..}^2}{an}$ or $\sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \bar{\bar{y}})^2$

Factor SS: $\sum_{i=1}^{a} \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{an}$ or $n \sum_{i=1}^{a} (\bar{y}_i - \bar{\bar{y}})^2$

2^k design, *n* replicates

Effect estimate given by $\frac{(Contrast)}{n.2^{k-1}}$

Effect SS given by $\frac{(Contrast)^2}{n.2^k}$

Variance of an effect estimate is $\frac{\sigma_e^2}{n.2^{k-2}}$

Regression

$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}$ $\quad$ $S_{xy} = \sum xy - \frac{(\sum x)(\sum y)}{n}$ $\quad$ $S_{yy} = \sum y^2 - \frac{(\sum y)^2}{n}$

$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$ $\quad$ $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

$s_e = \hat{\sigma} = \sqrt{\frac{S_{yy} - \widehat{\beta_1}^2 S_{xx}}{n-2}}$ $\quad$ $s_e = \hat{\sigma} = \sqrt{\frac{SSE}{n-2}}$

$S.E.(\hat{\beta}_1) = \frac{s_e}{\sqrt{S_{xx}}}$ $\quad$ $S.E.(\hat{\beta}_0) = s_e \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}}$

$SSR = \frac{(S_{xy})^2}{S_{xx}}$

$S.E.(\hat{\beta}_0 + \widehat{\beta_1} x_0) = s_e \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}$

$S.E.(individual\ \hat{\beta}_0 + \widehat{\beta_1} x_0) = s_e \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}$