CORK INSTITUTE OF TECHNOLOGY
INSTITIÚID TEICNEOLAÍOCHTA CHORCAÍ

Semester 2 Examinations 2010/2011

Module Title: Statistics and experimental design for Chemical Engineering

Module Code: **STAT 8004**

School: Mechanical and Process Engineering

Programme Title: Bachelor of Engineering (Honours) in Chemical and Process Engineering - Award

Programme Code: **ECPEN_8_Y4**

External Examiner(s): Mr. J. Reilly
Internal Examiner(s): Mr. D. O'Hare

Instructions: Answer **any three questions.** All questions carry equal marks.

Duration: 2 Hours

Sitting: Summer 2011

Requirements for this examination: Statistical Tables by Murdoch and Barnes.

Note to Candidates: Please check the Programme Title and the Module Title to ensure that you have received the correct examination paper.
If in doubt please contact an Invigilator.

Q1. (a) Components of a certain type are shipped to a supplier in batches of twenty. Three components are randomly selected from a batch and tested. If one such batch contains three defective components,

(i) what is the probability that none of the selected components is defective?

(ii) What is the probability that at least two of the selected components are defective?

(6 marks)

(b) A manufacturing process has 100 customer orders to fill. Each order requires one component part that is purchased from a supplier. However, typically, 2% of the components are identified as defective, and the components can be assumed to be independent.

(i) If the manufacturer stocks 100 components, what is the probability that all 100 orders can be filled without reordering components?

(ii) If 105 components are stocked, what is the probability that 100 orders can be filled without reordering components?

(5 marks)

(c) The Rockwell hardness of a particular alloy is normally distributed with a mean of 68 and a standard deviation of 3.

(i) If a specimen is acceptable only if its hardness is between 62 and 75, what is the probability that a randomly chosen specimen has an acceptable hardness?

(ii) 90% of specimens have a hardness value greater than *k*. What is the value of *k*?

(6 marks)

(d) The content, *X*, of magnesium in an alloy is a random variable, distributed according to the following probability density function

$$f(x) = \begin{cases} \dfrac{x}{18}, & \text{for } 0 \le x \le 6 \\ 0, & \text{otherwise} \end{cases}$$

(i) Verify that this is a well-defined probability density function.

(ii) Find the mean value of *X*. Is the mean greater than or less than the median in this distribution? Justify your answer.

(8 marks)

Q2. (a) This year a large manufacturing firm began a program of compensating its management personnel for sick days not used. The firm decided to pay each manager a bonus for each unused sick day. In past years, the number of sick days used per manager per year had a probability distribution with $\mu = 9.2$, and $\sigma^2 = 3.24$. To determine whether the compensation program has effectively reduced the mean number of sick days used, the firm randomly sampled 50 managers and recorded the number of sick days used by each at year's end.

(i) Assuming that that the compensation program was **not** effective in reducing the mean number of sick days used, find the probability that the mean number of sick days used by the sample of 50 managers is less than 8.50 days.

(ii) If you observe that the sample mean is less than 8.50, what conclusion would you draw about the effectiveness of the compensation program? Explain. (6 marks)

(b) The following are the lifetimes of a sample of eight components of a certain type. Establish a 95% confidence interval for the mean lifetime and a 95% upper confidence limit on the variance of lifetimes for this type of component.

265	283	294	312	360	380	244	295

(6 marks)

(c) The following data are gathered in relation to an investigation, where the researcher's hypothesis is that $\mu_1 > \mu_2$:

$$n_1 = 10, \bar{x}_1 = 83.2, {s_1}^2 = 16.67, n_2 = 12, \bar{x}_2 = 71.4, {s_2}^2 = 5.37.$$

Carry out the appropriate test(s), and state your conclusions. (8 marks)

(d) Two different types of polishing solutions are being evaluated for possible use in the manufacture of interocular lenses. 250 lenses were polished using the first solution, and of these 204 had no polishing-induced defects. 300 lenses were polished using the second solution, and 202 of these had no polishing-induced defects. Examine for significant difference between the two polishing solutions, and state your conclusions. (5 marks)

Q3. (a) In the following table, y is the purity of oxygen produced in a chemical distillation process, and x is the percentage hydrocarbons that are present in the main condenser of the distillation unit.

x	0.99	1.15	1.46	0.87	1.55	1.19	0.98
y	90.01	91.43	96.73	87.59	99.42	93.54	90.56

(i) Plot the data on a scatter diagram, and comment on the level of correlation.

(ii) The equation of the least squares line is `y = 74.1 + 16.0 x.` For each of the x values in the table, calculate the *predicted* value of purity and hence the set of residuals.

(iii) Using the residuals, calculate the error sum of squares. Produce the ANOVA table and say what conclusion you draw from this table. Find also the value of the correlation coefficient, and test its significance.

(17 marks)

(b) It is suggested that the following set of data comes from Poisson distribution. Carry out a goodness-of-fit test, and comment.

x	f
0	18
1	46
2	36
3	28
4	12
5	10

(8 marks)

Q4. (a) An experiment was carried out to investigate the effects of three bleaching chemicals on pulp brightness. The data obtained are as follows.

Chemical	Brightness			
A	77.2	74.5	92.7	76.2
B	80.5	79.3	81.2	80.3
C	78.0	78.4	77.6	74.2

Produce the analysis of variance table.

What conclusion do you draw from your calculated F ratio, and why?

(10 marks)

(b) An engineer is trying to improve the efficiency of a reaction that converts a raw material into a product. The response variable is called the degree of conversion. The engineer varies three factors: catalyst type, reactant concentration and reaction temperature. Two replications were made at each combination of factor levels. The data, in coded form, are presented below.

	Catalyst			
	Type 1		Type 2	
	Reactant Concentration			
Temperature	0.1%	0.5%	0.1%	0.5%
120°	25	26	2	8
	16	33	1	15
160°	29	27	4	27
	23	31	0	35

(i) Estimate the catalyst effect, the temperature effect and the catalyst-temperature interaction effect and test their significance.

(ii) Draw the catalyst-temperature interaction plot and comment. (15 marks)

Statistical Formulae

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

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

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

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

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

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

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

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

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^{\,2}}{n_1} + \frac{s_2^{\,2}}{n_2}}} \sim t_\nu$$

$$\nu = \frac{\left(\frac{s_1^{\,2}}{n_1} + \frac{s_2^{\,2}}{n_2}\right)^2}{\frac{(s_1^{\,2}/n_1)^2}{n_1 + 1} + \frac{(s_2^{\,2}/n_2)^2}{n_2 + 1}} - 2$$

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

$$\frac{s_1^{\,2} / \sigma_1^{\,2}}{s_2^{\,2} / \sigma_2^{\,2}} \sim F^{n_1 - 1}_{n_2 - 1}$$

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

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

ANOVA

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

Total SS: $\sum\sum y^2_{ij} - \frac{y^2_{..}}{an}$

Factor SS: $\sum_{i=1}^{a} \frac{y^2_{i.}}{n} - \frac{y^2_{..}}{an}$

2. **AXB factorial model**:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta_{ij}) + \varepsilon_{ijk},\ i = 1,2,\ldots,a;\ j = 1,2,\ldots,b;\ k = 1,2,\ldots,n.$$

Total SS: $\sum\sum\sum y^2_{ijk} - \frac{y^2_{...}}{abn}$

SSA: $\sum_{i=1}^{a} \frac{y^2_{i..}}{bn} - \frac{y^2_{...}}{abn}$

SSB: $\sum_{j=1}^{b} \frac{y^2_{.j.}}{an} - \frac{y^2_{...}}{abn}$

3. 2^k design, n replicates.

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

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

Standard error of an effect estimate is $\sqrt{\frac{\sigma_e^2}{n.2^{k-2}}}$, where σ_e^2 is the error variance.

Regression

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

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} \qquad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \qquad s_e = \hat{\sigma} = \sqrt{\frac{S_{yy} - \hat{\beta}_1^{\,2} S_{xx}}{n-2}}$$

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

$$S.E.(\hat{\beta}_0 + \hat{\beta}_1 x_0) = s_e \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}} \qquad S.E.(individual\,\hat{\beta}_0 + \hat{\beta}_1 x_0) = s_e \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}$$