CORK INSTITUTE OF TECHNOLOGY
INSTITIÚID TEICNEOLAÍOCHTA CHORCAÍ

Autumn Examinations 2017/2018

Module Title:	**Statistics for Engineering**
Module Code:	STAT 8005
School:	School of Science and Informatics
Programme Title(s):	BEng (Hons) in Structural Engineering BEng (Hons) in Biomedical Engineering BEng (Hons) in Mechanical Engineering
Programmes Code(s):	**CSTRU_8_Y3** **EBIOM_8_Y3** **EMECH_8_Y3**
External Examiner(s):	Prof. Michael Wallace
Internal Examiner(s):	Ms. Sarah Murphy
Instructions:	Answer three questions. All questions carry equal marks.
Duration:	2 hours
Sitting:	Autumn 2018

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

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

Question 1

(a) At a local company, 85% of staff members have children. It is found that 52% of the staff members who have children have health insurance and 36% of the staff members who do not have children have health insurance. A staff member is selected at random.

i. Calculate the probability that the chosen staff member does not have health insurance.
ii. John Smith is one member of staff who does not have health insurance. What is the probability that he has children?
iii. Are having children and having health insurance independent events or not? Explain your answer.

(8 marks)

(b) An insurance company offers a discount for early payments. It has been found that 25% of transactions made to this company receive the early payment discount. A sample of 20 transactions are selected at random.

i. State the random variable.
ii. What is the probability that one transaction receives the early payment discount?
iii. What is the probability that at least two transactions receive the early payment discount?

(5 marks)

(c) A discrete random variable X is distributed according to the following probability function:

$$P(x) = a4^x \qquad x = 0, 1, 2, 3.$$

i. Find the value of a.
ii. Find the values of $E(X)$ and $V(X)$, the mean and the variance of X.

(7 marks)

(d) A quality control inspector is examining steel beams for flaws as they are coming off a production line. The average number of flaws identified per hour is 2.4. Assuming the number of flaws identified follows a Poisson distribution, find the probability that in the next hour the inspector will find:

i. exactly three flaws
ii. less than two flaws
iii. What is the probability that the inspector will find at least two flaws in a 1.5 hour interval?

(5 marks)

Question 2

(a) A continuous random variable, X, has the following probability density function

$$f(x) = \begin{cases} 4e^{4x} & \text{for } x \leq 0 \\ 0 & \text{for } x > 0 \end{cases}$$

i. Show that this function is well defined.
ii. Find the probability that it will take on a value between -4 and -1
iii. Find the probability that it will take on a value less than -0.8.
iv. Draw a rough sketch of the function.

(9 marks)

(b) The weekly profit at an ice cream parlour is known to follow a Normal distribution with a mean of €840 and standard deviation of €95.

i. Find the probability that a randomly chosen week will have a profit between €750 and €1,050.
ii. Thirty-five percent of weeks have a profit of €k or more. What is the value of k?

(7 marks)

(c) A study is conducted on the resting heart rate (beats per minute) of female patients admitted to the Accident and Emergency department at a local hospital. The stem and leaf display below was generated in Minitab using the resting heart rates gathered from these patients.

Stem-and-leaf Display

1	4	8
4	5	069
9	6	01789
15	7	035678
22	8	0223469
(14)	9	11234555567789
9	10	124456679

Leaf Unit = 1

i. How many patients were involved in this study?
ii. State an appropriate key for this stem and leaf display?
iii. Comment on the shape of the stem and leaf display.
iv. Using the stem and leaf display, find the range and the interquartile range.
v. Without calculating the values, say whether you think the mean here is greater or less than the median, and why.

(9 marks)

Question 3

(a) Answer each of the following questions:

i. Explain what is meant by a type II error.
ii. What is a critical value?
iii. If you have a small sample size, what is the best plot to generate to check for normality?
iv. What is meant by the term "sampling distribution of $\bar{x}$ ". In your explanation, include important relationships between the parameters of this distribution and the parameters of the population from which the samples are taken.
v. What are the assumptions of an independent two sample test?

(8 marks)

(b) A machine produces metal rods used in an automobile suspension system. A random sample of 12 rods is selected and the diameter is measured in each case. The resulting data are:

6.61	6.62	6.58	6.57	6.62	6.54
6.59	6.54	6.55	6.57	6.58	6.59

It is known that the rod diameters follow a Normal distribution.

i. Find a 95% confidence interval for the true mean rod diameter and interpret this confidence interval.
ii. If you increased the confidence level from 95% to 99%, would the width of confidence interval increase or decrease? *Note:* You do not need to find the 99% confidence interval.

(6 marks)

(c) A company states that the average deflection of their 3-metre steel beams is 1.4mm. A construction contractor who purchases large quantities of steel beams suspects that the manufacturer misleads its customers and that the true deflection is, in fact, larger than the value stated. To see whether this suspicion is justified, the contractor selects 10 beams at random and determines their deflection with the following results:

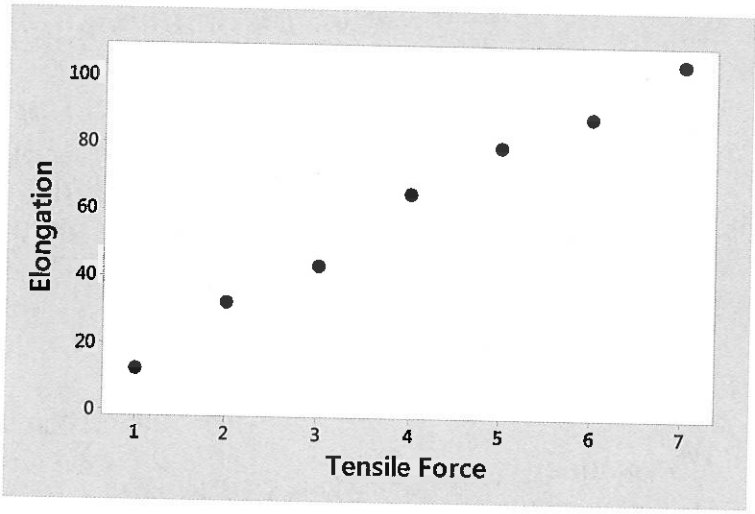
1.43	1.45	1.37	1.42	1.44	1.38	1.41	1.39	1.42	1.43

Is there evidence to support the contractor's suspicions? Carry out an appropriate test of hypothesis on the above data. Use a 1% level of significance and clearly state your hypotheses and conclusions.

(11 marks)

Question 4

(a) A study was conducted to examine the relationship between the tensile force applied to a steel specimen in thousands of pounds, and the resulting elongation, in thousandths of an inch. A sample of size seven yielded the following output.



$$\sum x = 28, \quad \sum y = 423, \quad \sum xy = 2{,}116, \quad \sum x^2 = 140, \quad \sum y^2 = 32{,}043$$

i. Using the scatterplot, comment on the relationship between the variables.
ii. Find the value of the correlation coefficient.
iii. Find the equation of the least squares regression line of elongation on tensile force.
iv. Using the equation that you obtained in part (iii) find the best predicted elongation if the tensile force is 4.7.

(9 marks)

(b) Two types of instruments for measuring the amount of sulphur monoxide in the atmosphere are being compared in an air pollution experiment. It is of interest to determine whether the two types of instruments yield measurements having the same mean. The following readings were recorded for the two instruments:

Instrument *A*	0.81	0.82	0.75	0.61	0.65	0.68	0.89	0.64	0.72
Instrument *B*	0.74	0.63	0.87	0.56	0.70	0.69	0.76	0.53	0.57

i. Assuming equal variances, carry out a suitable hypothesis test on the data using a 5% level of significance.
ii. State an approximate p-value for the test.
iii. Find a 95% confidence interval for the difference in means. Does this support your conclusion from part i?

(16 marks)

Statistical Formulae

Select formulae given below. See statistical tables or mathematical tables for other formulae.

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

$$P(X = r) = {}^{n}C_{r}\, p^{r}\, q^{n-r}$$

$$E(X) = \sum xP(x)$$

$$P(X = r) = \frac{\lambda^{r} e^{-\lambda}}{r!}$$

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

$$V(X) = E(X^2) - \left(E(X)\right)^2$$

Regression and Correlation:

$$\hat{y} = b_0 + b_1 x \qquad b_0 = \frac{\sum y - b_1 \sum x}{n} \qquad b_1 = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}} \qquad t_0 = r\sqrt{\frac{n-2}{1-r^2}}$$

Confidence Intervals and Hypothesis Testing:

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

$$(\bar{x}_1 - \bar{x}_2) \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

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

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

$$t_0 = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$Z_0 = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\frac{\alpha}{2}}\, s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$Z_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$