CORK INSTITUTE OF TECHNOLOGY
INSTITIÚID TEICNEOLAÍOCHTA CHORCAÍ

Semester 2 Examinations 2010/11

Module Title: Probability and Statistics

Module Code: **STAT 7001**

School: **School of Computing & Mathematics**

Programme Title: **Bachelor of Science (Honours) in Software Development**

Programme Code: **KSDEV_8_Y2**

External Examiner(s): **Mr. J. Reilly**

Internal Examiner(s): **Dr. Tadhg Creedon**

Instructions: **Answer any THREE questions. All questions carry equal marks. Pages 6, 7 and 8 contain formulae specific to this course. A booklet of statistical tables accompanies this paper. The following are also provided: Mathematical Tables and Murdoch and Barnes Statistical Tables.**

Duration: **2 HOURS**

Sitting: Summer 2011

Requirements for this examination:

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.

1. (a) For the sample below draw an ordered stem and leaf plot and construct a box plot for the data (identifying any outliers):

26	41	15	38	28
45	27	33	75	35
25	38	42	17	12

(7 marks)

(b) The frequency distribution for a sample of 100 distances (in km) driven in rental cars in a week was as follows:

Distance (km)	Number of cars
0-100	27
100-200	39
200-300	24
300-500	10

Calculate the mean, the sample standard deviation and the median.

(8 marks)

(c) Machines A and B make components. Of those made by machine A, 95% are reliable; of those made by machine B, 92% are reliable. Machine A makes 70% of the components with machine B making the rest. Calculate the probability that a component picked at random is

(i) made by machine A and is reliable;

(ii) made by machine B and is unreliable;

(iii) reliable.

(5 marks)

(d) A committee of four people is to be selected from a group of six lawyers and three engineers. All candidates are equally likely to be selected. Let X represents the number of lawyers on the committee.
Find the probability distribution of X. Find the expected value of X.

(5 marks)

2. (a) A continuous random variable X has probability density function given by

$$f(x) = \begin{cases} \dfrac{x^2}{9}, & 0 \le x \le 3 \\ 0, & \text{elsewhere} \end{cases}$$

(i) Find the mean $E(X)$ and the variance $V(X)$.

(ii) By finding the cumulative distribution function $F(x)$, obtain the median of this distribution.

(iii) Find the probability $P(0.5 < X < 2.2)$.

(9 marks)

(b) Transaction times at a service counter have a negative exponential distribution with mean transaction time 4 minutes.

(i) Find the percentage of transaction times between 3 and 5 minutes.

(ii) Find the interval containing the longest 10% of transactions.

(8 marks)

(c) A machine being used for packaging raisins has been set so that on average 150g of raisins will be packaged per box. The operations manager wishes to test the machine setting and selects a random sample of 10 raisin boxes filled during the production process. The sample mean is 150.4g and the sample standard deviation is 4.25.

Is there evidence that the mean weight per box is different from 150g?

Use a formal test with a 5% level of significance and state your conclusions clearly.

(8 marks)

3. (a) The probability that a machine needs correcting adjustments during a day's production run is 0.2. If there are 6 such machines running on a particular day find the probability that

(i) no machines need correcting;
(ii) just one machine needs correcting;
(iii) more than two machines need correcting.

(6 marks)

(b) Customers arrive randomly at a department store at an average rate of 3.4 per minute. Calculate the probability that

(i) no customers arrive in a particular minute;
(ii) exactly one customer arrives in a particular minute;
(iii) two or more customers arrive in a particular minte;
(iv) one or more customers arrive in a 30-second period.

(7 marks)

(c) A discrete random variable has the following probability distribution:

x	0	1	2	3
$P(X = x)$	0.1	0.4	0.3	0.2

(i) Find the mean and the variance of *X*.

(ii) Two independent random observations X_1 and X_2 are made on *X*. Let $Y = X_1 + X_2$. Construct the probability distribution for the random variable *Y*. What are the mean and variance of *Y*?

(12 marks)

4. (a) The weekly revenue at a restaurant is a normal random variable with mean € 2200 and standard deviation € 200. Find the percenatge of weeks for which the restaurant's revenue is

(i) more than € 2500;

(ii) between € 2000 and € 2300.

What is the probability that the restaurant's total revenue over the next two weeks exceeds € 5000?

(7 marks)

(b) A sample of 300 workers in an industry showed average annual overtime earnings of €3,120 with a standard deviation of € 844.

(i) Find a 95% confidence interval for the mean overtime earnings for the entire population of workers.

(ii) What sample size would be necessary to estimate the population mean to within € 25 with 95% confidence?

(iii) What effect would it have on your answers above if it were known that the number of workers in the population is 5,500?

(11 marks)

(c) The total number of households in an area is 5,523. A sample of 300 of these households revealed that 123 had internet access.

(i) Find a 99% confidence interval for the proportion of all the 5,523 households in the area with internet access.

(ii) What sample size would be necessary to estimate the population proportion to within 0.02 with 95% confidence?

(7 marks)

Formulae for DCOM2 Statistics Examinations 2011

Descriptive Statistics

Mean : $\bar{x} = \frac{\sum x}{n}$ or $\bar{x} = \frac{\sum fx}{\sum f}$

Mode = $L + \left(\frac{D_1}{D_1 + D_2} \right) C$

Median = $L_M + \frac{\left(\frac{n}{2} - F_{M-1} \right)}{f_M} . C_M$

Population standard deviation: $\sqrt{\frac{\sum \left(x - \bar{x}\right)^2}{n}} = \sqrt{\frac{\sum x^2}{n} - \left(\bar{x}\right)^2}$

Population standard deviation: $\sqrt{\frac{\sum f\left(x - \bar{x}\right)^2}{\sum f}} = \sqrt{\frac{\sum fx^2}{\sum f} - \left(\bar{x}\right)^2}$

Sample standard deviation $= s = \sqrt{\frac{\sum \left(x - \bar{x}\right)^2}{n-1}} = \sqrt{\frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1}}$

Sample standard deviation $= s = \sqrt{\frac{\sum fx^2 - \frac{\left(\sum fx\right)^2}{\sum f}}{\left(\sum f\right) - 1}}$

Probability

Addition Law: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Multiplication Law $P(A \cap B) = P(A \mid B) P(B)$

Bayes' Theorem: $P(B \mid A) = \frac{P(A \mid B) P(B)}{P(A)}$

Discrete Random Variables

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

$$E(X^2) = \sum x^2 P(X = x)$$

$$V(X) = E(X^2) - [E(X)]^2$$

Binomial $P(X = r) = {}^nC_r\, p^r (1-p)^{n-r}$

Poisson $P(X = r) = \dfrac{e^{-m} m^r}{r!}$

Hypergeometric $P(X = r) = \dfrac{{}^MC_r \; {}^{N-M}C_{n-r}}{{}^NC_n}$

Continuous Random Variables

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,dx$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx$$

$$V(X) = E(X^2) - [E(X)]^2$$

<u>Negative Exponential Distribution</u>

$f(x) = ae^{-ax}$, $F(x) = 1 - e^{-ax}$, $R(x) = e^{-ax}$, $E(X) = \dfrac{1}{a}$

Sampling Theory

X is $N(\mu, \sigma)$ or n large $\Rightarrow \bar{x}$ is $N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right)$

For n large, p_s is $N\left(P, \sqrt{\frac{P(1-P)}{n}}\right)$

	Mean	
	Large Sample	Small Sample
Confidence Intervals	$\bar{x} \pm Z \frac{s}{\sqrt{n}}$ $\bar{x} \pm Z \frac{s}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$ $n = \frac{Z^2 s^2}{E^2}$ $n_f = \frac{n}{1+\frac{n}{N}}$	$\bar{x} \pm t_{n-1} \frac{s}{\sqrt{n}}$
One Sample Test	$Z = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$	$t_{n-1} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$

	Proportion
	Large Sample
Confidence Intervals	$p_s \pm Z \sqrt{\frac{p_s(1-p_s)}{n}}$ $p_s \pm Z \sqrt{\frac{p_s(1-p_s)}{n}} \sqrt{\frac{N-n}{N-1}}$ $n = \frac{Z^2 p_s(1-p_s)}{E^2}$ or $n = \frac{1}{4}\frac{Z^2}{E^2}$ $n_f = \frac{n}{1+\frac{n}{N}}$
One Sample Test	$Z = \frac{p_s - P}{\sqrt{\frac{p_s(1-p_s)}{n}}}$