Answer ***all*** questions in Section A.

The marks for the ***best three*** answers in Section B will be used in the assessment.

Normal and Chi-squared tables are provided at the end of the paper

## SECTION A

1. The probability density function $f(y)$ of a continuous random variable $Y$ is given by

$$f(y) = \begin{cases} K\left(20y^3 - 15y^2\right) & \text{if } 1 \le y \le 2 \\ 0 & \text{otherwise.} \end{cases}$$

(i) Verify that $K = 1/40$. [2 marks]

(ii) Calculate $E[Y]$ [3 marks]

(iii) Calculate the variance of $Y$. [3 marks]

2. One group of patients with a particular disease were treated with drug $A$ whilst another were treated with drug $B$. Each patient was then assessed to see if they had responded to treatment or not. The results are given in the following table.

| | Response to treatment | |
|---|---|---|
| Drug | Yes | No |
| *A* | 20 | 20 |
| *B* | 12 | 28 |

Test the hypothesis that there is no association between the drug given and the response to treatment. [8 marks]

3. Daily wind speeds were recorded at a particular location over a consecutive 30-day period. The figure below shows a stem-and-leaf plot for these data.

| | | | | | | |
|---|---|---|---|---|---|---|
| 0 | 5 | 8 | 8 | | | |
| 1 | 1 | 1 | 4 | 6 | 8 | |
| 2 | 0 | 2 | 5 | 5 | 9 | 9 |
| 3 | 2 | 3 | 4 | 7 | 8 | |
| 4 | 4 | 6 | 9 | | | |
| 5 | 3 | 4 | 8 | | | |
| 6 | 0 | 7 | | | | |
| 7 | 1 | 5 | | | | |
| 8 | 2 | | | | | |

3 | 4 represents 34 miles per hour.

Describe the distribution of wind speed

(i) in terms of its shape [2 marks]

(ii) by calculating an appropriate measure of location [2 marks]

and (iii) by calculating an appropriate measure of spread. [4 marks]

4. A random variable $X$ is normally distributed with mean 8 and variance 9. Calculate:

(i) $P(X > 8)$ [1 mark]

(ii) $P(5 < X < 11)$ [2 marks]

(iii) the probability that $X$ is positive [2 marks]

(iv) the value of $a$ such that $P(X > a) = 0.166$ [3 marks]

5. The quality control division of a food manufacturing company is concerned about the contamination of a particular product with traces of nuts. In the past it has been found that 8% of packets of the product contain traces of nuts.

A random sample of 30 packets are taken from the production line. What is the

(i) expected number of packets containing nut traces? [2 marks]

(ii) probability that none of the packets contain traces of nuts? [2 marks]

(iii) probability that fewer than 5 packets contain traces of nuts? [4 marks]

## SECTION B

6. A gas company wishes to review its staffing for emergency callouts in a particular area. The company believes the number of emergency callouts per day has a Poisson distribution. They believe current staffing levels are appropriate provided the average number of emergency callouts per day is 1.

To assist its review, the company collected the following daily data over the period 01/06/01 - 31/07/01 inclusive.

| Number of emergency callouts | Number of days |
|---|---|
| 0 | 6 |
| 1 | 20 |
| 2 | 18 |
| 3 | 10 |
| 4 | 7 |
| ≥5 | 0 |

(i) Use a $\chi^2$ test, applying Cochran's rule, to decide whether the company needs to change its staffing levels. [7 marks]

(ii) Test the hypothesis that the data are drawn from a Poisson distribution with the mean determined from the data. Apply Cochran's rule as appropriate. [10 marks]

(iii) The company considers they may need to increase the staffing levels if the probability of more than 4 callouts in a day is higher then 0.05. Do they need to increase their staffing levels? Justify your answer. [3 marks]

7(a) For two events $A$ and $B$, prove that

$$p(B|A) = \frac{p(A|B)\,p(B)}{p(A|B)\,p(B) + p(A|\overline{B})\,p(\overline{B})}$$ [8 marks]

(b) Tomato weevil is a problem for tomato growers: 30% of plants are affected. The chance of a plant yielding edible tomatoes is 70% if the plant is affected and 95% if the plant is unaffected.

Given that a plant has produced edible tomatoes, what is the probability that the plant was affected? [6 marks]

If a plant has produced inedible tomatoes, what is the probability that the plant was unaffected? [6 marks]

8(a) If $X$ has a discrete Uniform distribution on 1, 2, …, 12, show that the mean and variance of $X$ are 6½ and $11^{11}/_{12}$ respectively. [10 marks]

(b) A company can make up to 6 items per day. Past experience shows that the daily demand follows the probability distribution below:

| Number of items sold, $y$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $P(Y = y)$ | 0.01 | 0.04 | 0.12 | C | 0.29 | 0.22 | 0.10 |

Find the value of $c$. [2 marks]

If demand on different days is independent, what is the probability that over a two-day period, exactly 5 items will be sold? [5 marks]

The profit made by the company depends on the number of items sold as shown below:

| Number of items sold, $y$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| Profit, in £s | -160 | -80 | 0 | 80 | 160 | 240 | 320 |

Calculate the expected daily profit made by the company. [3 marks]

9. The distribution of a critical dimension on auto engine crankshafts is thought to be approximately Normal with $\mu$ = 224mm and $\sigma$ = 0.03mm. Crankshafts with dimensions between 223.92mm and 224.08mm are acceptable. What percentage of all crankshafts produced are acceptable? [4 marks]

A quality engineer measured a sample of 200 of these crankshafts. The table below summarises the frequency distribution of these measurements.

| Dimension | Frequency |
|---|---|
| 223.00-223.30 | 2 |
| 223.30-223.50 | 4 |
| 223.50-223.70 | 9 |
| 223.70-223.80 | 23 |
| 223.80-223.90 | 32 |
| 223.90-224.00 | 49 |
| 224.00-224.10 | 34 |
| 224.10-224.20 | 18 |
| 224.20-224.30 | 12 |
| 224.30-224.50 | 8 |
| 224.50-224.70 | 6 |
| 224.70-225.00 | 3 |

Plot a histogram of these data. [8 marks]

Comment, with reasoning, on whether the histogram suggests the data are Normally distributed or not. [2 marks]

The sample mean, $\bar{x}$, and standard deviation, $s$, were calculated to be 223.98 and 0.42 respectively. Calculate a 95% confidence interval for the mean and interpret your result.

[6 marks]

10. A drug given to epileptic patients to help control their seizures is found to cause a skin rash in 3% of people. A doctor has recently reviewed her patients. She found that of 62 epilepsy patients to whom she gave this drug, 6 of them experienced a rash.

(a) Assuming the number of patients experiencing a rash has a Binomial distribution, calculate the probability of at least 6 of the patients experiencing a rash. [4 marks]

(b) Use the Poisson approximation to the Binomial distribution to estimate the probability that at least 6 of the patients would experience a rash. [7 marks]

(c) Use the Normal approximation to the Binomial distribution to estimate the probability that at least 6 of the patients would experience a rash. [7 marks]

(d) Which of the approximations (b) or (c) gives a result closer to (a)? Explain why this should be so. [2 marks]