7. The continuous random variable D is uniformly distributed over the interval $[x - 1, x + 5]$ where x is a constant.

A random sample of n observations of D is taken, where n is large.

(a) Use the Central Limit Theorem to find an approximate distribution for $\overline{D}$
Give your answer in terms of n and x where appropriate.

(3)

The n observations of D have a sample mean of 24.6

Given that the lower bound of the 99% confidence interval for x is 22.101 to 3 decimal places,

(b) find the value of n
Show your working clearly.

(5)

DO NOT WRITE IN THIS AREA

Leave blank

Answer All questions. Write your answers in the spaces provided.

1. A head teacher wants to find out the students' opinions about the length of lessons. The head teacher wishes to survey the students using the three groups in the table below.

	Number of students
Group 1 (Y7 – Y9)	432
Group 2 (Y10 – Y11)	360
Group 3 (Y12 – Y13)	108

She decides to take a stratified sample of 50 students.

(a) Explain how to select the students for this stratified sample.

(3)

(b) Give one advantage of carrying out this survey using stratified sampling, rather than taking a simple random sample of the whole school.

(1)

DO NOT WRITE IN THIS AREA

1 A machine fills bottles with mineral water.

The machine is checked every day to ensure that it is working correctly. On a particular day a random sample of 100 bottles is taken. The volume of water, x millilitres, for each bottle is measured and each measurement is coded using

$$y = x - 1000$$

The results are summarised below

$$\sum y = 847 \qquad \sum y^2 = 13\ 510.09$$

(a) (i) Show that the value of the unbiased estimate of the mean of x is 1008.47

(ii) Calculate the unbiased estimate of the variance of x

(4)

The machine was initially set so that the volume of water in a bottle had a mean value of 1010 millilitres.

Later, a test at the 5% significance level is used to determine whether or not the mean volume of water in a bottle has changed. If it has changed then the machine is stopped and reset.

(b) Write down suitable null and alternative hypotheses for a 2-tailed test.

(1)

(c) Find the critical region for $\bar{X}$ in the above test.

(4)

(d) Using your answer to part (a) and your critical region found in part (c), comment on whether or not the machine needs to be stopped and reset.
Give a reason for your answer.

(2)

(e) Explain why the use of $\sigma^2 = s^2$ is reasonable in this situation.

(1)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

5: The following table shows the number of male puppies born in 250 dog litters of size 5

Number of males	0	1	2	3	4	5
Number of litters	2	40	90	85	30	3

Jeff believes that a binomial distribution would be a suitable model for these data.

(a) Find the proportion of male puppies born in these litters of size 5

(2)

Jeff calculates expected frequencies, to 2 decimal places, as follows.

Number of males	0	1	2	3	4	5
Expected frequencies	r	41.92	79.91	76.16	36.30	6.92

(b) Find the value of r

(2)

The value of $\sum \frac{(O-E)^2}{E}$ for the given values in the table **excluding** $\boldsymbol{r}$ is 5.70 to 2 decimal places.

(c) Using a 5% significance level, test whether or not a binomial distribution is a suitable model for the number of male puppies born in these 250 litters.
You should state the hypotheses, the degrees of freedom and the critical value used.

(7)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA