4 **(a)** Given that $f(r) = \dfrac{r}{2^r}$, by considering $f(r+1) - f(r)$, find $\displaystyle\sum_{r=1}^{n} \frac{1-r}{2^{r+1}}$. **[3]**

(b) **(i)** Cauchy's root test states that a series of the form $\displaystyle\sum_{r=0}^{\infty} a_r$ (where $a_r > 0$ for all r) converges when $\displaystyle\lim_{n\to\infty} \sqrt[n]{a_n} < 1$, and diverges when $\displaystyle\lim_{n\to\infty} \sqrt[n]{a_n} > 1$. When $\displaystyle\lim_{n\to\infty} \sqrt[n]{a_n} = 1$, the test is inconclusive. Using the test and given that $\displaystyle\lim_{n\to\infty} \sqrt[n]{n^p} = 1$ for all positive p, explain why the series $\displaystyle\sum_{r=0}^{\infty} \frac{2^r r^x}{3^r}$ converges for all positive values of x. **[3]**

(ii) By considering $(1-y)^{-2} = 1 + 2y + 3y^2 + 4y^3 + \ldots$, evaluate $\displaystyle\sum_{r=0}^{\infty} \frac{2^r r^x}{3^r}$ for the case when $x = 1$. **[2]**

Section B: Probability and Statistics [60 marks]

5 Seng Ann Joo Cooperative sells granulated sugar in packets. These packets come in two sizes: standard and large. The masses, in grams, of these packets are normally distributed with mean and standard deviation as shown in the table below.

	Mean	Standard Deviation
Standard	520	8
Large	1030	11

(i) Find the probability that two standard packets weigh more than a large packet. **[3]**

(ii) Find the probability that the mean mass of two standard packets and one large packet of sugar is between 680g and 700g. **[3]**

6 A university drama club contains 3 Biology students, 4 History students, and 6 Literature students. 5 students are to be selected as the cast of an upcoming production.

(i) In how many ways can the 5 cast members be selected so that there are at most 2 Biology students? **[2]**

(ii) Find the probability that, amongst the cast members, the number of History students exceeds the number of Literature students, given that there are at most 2 Biology students. **[4]**

[Turn Over]

7 **(i)** The discrete random variable X takes values 1, 2, 3,… , n, where n is a positive integer greater than 1, with equal probabilities.

Find, in terms of n, the mean μ, and the variance, σ^2, of X. **[4]**

[You may use the result $\sum_{r=1}^{n} r^2 = \frac{n(n+1)(2n+1)}{6}$.]

Let $n = 6$. An observation of X is defined as an *outlier* if $|X - \mu| > \sigma$.

(ii) 20 observations of X are made. Find the probability that there are at least 8 observations that are outliers. **[4]**

8 In a game of chance, a player has to draw a counter from a bag containing n red counters and $(40-n)$ blue counters before throwing a fair die. If a red counter is drawn, she throws a six-sided die, with faces labelled 1 to 6. If a blue counter is drawn, she throws a ten-sided die, with faces labelled 1 to 10. She wins the game if the uppermost face of the die thrown shows a number that is a perfect square.

(i) Given that $n = 15$, find the exact probability that a player wins the game. Hence, find the probability that, when 3 people play this game, exactly 2 won. **[3]**

(ii) For a general value of n, the probability that a winning player drew a blue counter is denoted by $\mathrm{f}(n)$. Show that $\mathrm{f}(n) = a + \frac{b}{360+n}$, where a and b are constants to be determined. Without further working, explain why f is a decreasing function for $0 \le n \le 40$, and interpret what this statement means in the context of the question. **[5]**

9 Many different interest groups, such as the lumber industry, ecologists, and foresters, benefit from being able to predict the volume of a tree from its diameter. The following table of 10 shortleaf pines is part of the data set concerning the diameter of a tree, x, in inches and volume of a tree y, in cubic feet.

Diameter (x inches)	5.0	5.6	7.5	9.1	9.9	10.3	11.5	12.5	16.0	18.3
Volume (y cubic feet)	3.0	7.2	10.3	17.0	23.1	27.4	26.0	41.3	65.9	97.9

(Bruce and Schumacher, 1935)

(i) Draw a scatter diagram to illustrate the data, labelling the axes clearly. **[2]**

It is thought that the volume of trees with different diameters can be modelled by one of the formulae

[Turn Over]

$$y = a + bx \quad \text{or} \quad \ln y = c + d \ln x$$

where a, b, c and d are constants.

(ii) Find the value of the product moment correlation coefficient between

(a) y and x,

(b) $\ln y$ and $\ln x$.

Leave your answers correct to 5 decimal places **[2]**

(iii) Use your answers to parts **(i)** and **(ii)** to explain which of the models is the better model. **[1]**

(iv) It is required to estimate the value of y for which $x = 20$. Find the equation of a suitable regression line and use it to find the required estimate, correct to 1 decimal place. Explain whether your estimate is reliable. **[3]**

10 A factory manufactures a large number of erasers in a variety of colours. Each box of erasers contains 36 randomly chosen erasers. On average, 20% of erasers in the box are blue.

(i) State, in context, two assumptions needed for the number of blue erasers in a box to be well modelled by a binomial distribution. **[2]**

(ii) Find the probability that a randomly chosen box of erasers contain at most six blue erasers. **[1]**

200 randomly chosen boxes are packed into a carton. A carton is considered acceptable if at least 40% of the boxes contain at most six blue erasers each.

(iii) Find the probability that a randomly chosen carton is acceptable. **[3]**

The cartons are exported by sea. Over a one-year period, there are 30 shipments of 150 cartons each.

(iv) Using a suitable approximation, find the probability that the mean number of acceptable cartons per shipment for the year is less than 80. **[3]**

The owner decided to change the proportion of blue erasers to p. A box of erasers is chosen.

(v) Write down in terms of p, the probability that the box contains exactly one blue eraser. **[1]**

(vi) The probability that a box contains exactly one blue eraser is twice the probability that the box contains exactly two blue erasers. Write an equation in terms of p, and hence find the value of p. **[2]**

[Turn Over]

11 The time T seconds required for a computer to boot up, from the moment it is switched on, is a normally distributed random variable. The specifications for the computer state that the population mean time should not be more than 30 seconds. A Quality Control inspector checks the boot up time using a sample of 25 randomly chosen computers.

A particular sample yielded $\sum t = 802.5$ and $\sum t^2 = 26360.25$.

(i) Calculate the unbiased estimates of the population mean and variance. **[2]**

(ii) What do you understand by the term "unbiased estimate"? **[1]**

(iii) Test, at the 5% level of significance level, whether the specification is being met. Explain in the context of the question, the meaning of "5% level of significance". **[5]**

(iv) Find the range of values of $\bar{t}$ such that the specification will be met in the test carried out in part **(iii)**. **[1]**

(v) A new Quality Control policy is that when the specification is not met, all the computers will be sent back to the manufacturer for upgrading. The inspector tested a second random sample of 25 computers, and the boot up time, y seconds, of each computer is measured, with $\bar{y} = 32.4$. Using a hypothesis test at the 5% level of significance, find the range of values of the population standard deviation such that the computers will not be sent back for upgrading. **[3]**

End of Paper