Q1: A sample of n independent observations x_1, x_2, $\ldots$, x_n are obtained of a random variable having a Poisson distribution with mean μ. Show that the maximum likelihood estimate of μ is the sample mean $\bar{x} = \sum_{i=1}^{n} x_i/n$. Show that corresponding estimator $(\overline{X})$ is an unbiased estimator of μ, and has variance μ/n.

The scientists Rutherford and Geiger reported an experiment in which they counted the number of alpha particles emitted from a radioactive source during intervals of $7\frac{1}{2}$ seconds duration, for $2,612$ different intervals. A total of $10,126$ particles were counted. The data obtained are summarised in the table below.

Number	0	1	2	3	4	5	6	7	8	9	10	11	12	>12	Total
Frequency	57	203	383	525	532	408	273	139	49	27	10	4	2	0	2612

It has been suggested that the number of particles emitted in an interval may be adequately modelled by a Poisson distribution. Assuming this conjecture to be correct, find the maximum likelihood estimate of the mean of this distribution, and use this to estimate the expected frequencies corresponding to the observed frequencies given in the table. Comment informally on the extent of agreement between these observed and expected frequencies.

Q2: A random sample X_1, X_2, $\ldots$, X_n $(n > 2)$ of independent observations is taken from a random variable X which has a geometric distribution with parameter p $(0 < p < 1)$, whose probability function is given by

$$p(x) = p^{x-1}(1-p), \text{ for } x = 1, 2, 3, \ldots, \infty.$$

Write down the likelihood for p and hence show that the log likelihood for p is given by

$$l(p) = (\sum_{i=1}^{n} x_i - n)\log(p) + n\log(1-p),$$

and that the maximum likelihood estimate $\hat{p}$ of p is given by

$$\hat{p} = \frac{\sum_{i=1}^{n} x_i - n}{\sum_{i=1}^{n} x_i}.$$

Show that $P(X \geqslant y) = p^{y-1}$. Hence if a further m independent observations Y_1, Y_2, $\ldots$, Y_m are obtained where $X \geqslant Y$, show that the log likelihood for p is now given by

$$l(p) = (\sum_{i=1}^{n} x_i + \sum_{i=1}^{m} y_i - n - m)\log(p) + n\log(1-p),$$

and that the maximum likelihood estimate $\hat{p}$ of p is now given by

$$\hat{p} = \frac{\sum_{i=1}^{n} x_i + \sum_{i=1}^{m} y_i - n - m}{\sum_{i=1}^{n} x_i + \sum_{i=1}^{m} y_i - m}.$$

Information was obtained on the birth order of children in 7745 families. The records were examined to see when a family's first daughter was born. The birth order number for the first daughter and the data are given below.

Birth Order Number	1	2	3	4	$\geqslant 5$	Total
Frequency	3684	1964	1011	549	537	7745

It has been suggested that the birth order number of the first daughter may be adequately modelled by a geometric distribution. Assuming this conjecture to be correct, find the maximum likelihood estimate of p (using the second result above) and estimate the expected frequencies for the corresponding observed frequencies in the table. Comment informally on the extent of agreement between these observed and expected frequencies.

Q3: A random sample $X_1, X_2, \ldots, X_n$ of independent observations is taken from a *Gamma* distribution, whose density function is given by

$$f(x) = \begin{cases} \dfrac{x}{\theta^2} \exp\left(-\dfrac{x}{\theta}\right) & , \quad \text{for } x \geqslant 0 \ , \\ 0 & , \quad \text{for } x < 0 \ . \end{cases}$$

Show that the maximum likelihood estimator $\hat{\theta}$ of θ is given by $\hat{\theta} = \sum_{i=1}^{n} X_i/(2n)$.

Using the result that $E(X_i^r) = (r+1)!\theta^r$ deduce that $\hat{\theta}$ is an unbiased estimator of θ and has variance $\theta^2/(2n)$. What reason would you give for supposing that the distribution of $\hat{\theta}$ is approximately normally distributed for large n?