1. $X_1, X_2, ..., X_n$ are observations of a random sample of size n from the geometric distribution with probability distribution $f(x, \theta) = \theta(1-\theta)^x$ for $x = 0, 1, \ldots$.

   (a) **(1 mark)** Find the estimator of θ by the method of moment.

   (b) **(1 mark)** Find the estimator of θ by the method of maximum likelihood.

   (c) **(2 marks)** Find the Cramer-Rao Lower Bound for the variance of an unbiased estimator for θ.

   (d) **(2 marks)** Does the variance of any unbiased estimator for θ achieve this bound? Why? Explain in details.

   (e) **(3 marks)** Find the limiting distribution of the maximum likelihood estimator for θ by Central Limit Theorem and Delta method. What phenomenon do you observe?

   (f) **(1 mark)** Is the geometric distribution a member of exponential family? Hence or otherwise, find the minimal sufficient and complete statistic.

   (g) **(1 mark)** Define the minimal sufficient and complete statistic in part(f) to be S. Find its distribution.

   (h) **(1 mark)** Find the UMVUE of $\frac{1-\theta}{\theta}$.

   (i) **(5 marks)** Find the UMVUE of θ.

2. Two independent random samples, X_{ij} for $i = 1, 2; j = 1, .., n_i$, are normally distributed with mean μ_i and variance σ_i^2. Define

$$\begin{aligned}
\bar{X}_i &= \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij} \qquad \text{for } \ i = 1, 2 \\
S_i^2 &= \frac{1}{n_i - 1}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2 \qquad \text{for } \ i = 1, 2 \\
\bar{\bar{X}} &= \frac{1}{\sum_{i=1}^{2} n_i}\sum_{i=1}^{2}\sum_{j=1}^{n_i} X_{ij} \\
S_p^2 &= \frac{1}{\sum_{i=1}^{2}(n_i - 1)}\sum_{i=1}^{2}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2 \ .
\end{aligned}$$

   (a) **(4 marks)** Assume that $\mu_1 = \mu_2 = \mu \in \mathcal{R}$ and $\sigma_1^2 = \sigma_2^2 = \sigma^2 > 0$. Find the distributions of $\bar{X}_1$, $\bar{X}_2$ and $\bar{\bar{X}}$. Which of $\bar{X}_1$, $\bar{X}_2$ and $\bar{\bar{X}}$ is the most efficient estimator of μ? Why?

   (b) **(4 marks)** Assume that $\mu_1 = \mu_2 = \mu \in \mathcal{R}$ and $\sigma_1^2 = \sigma_2^2 = \sigma^2 > 0$. Find the distributions of S_1^2, S_2^2 and S_p^2. Which of S_1^2, S_2^2 and S_p^2 is the most efficient estimator of σ^2? Why?

(c) Assume that $\mu_1 \in \mathcal{R}$, $\mu_2 \in \mathcal{R}$, $\sigma_1^2 > 0$ and $\sigma_2^2 > 0$.

i. **(1 mark)** Find the set of minimal sufficient and complete statistics for the unknown parameters of μ_1, μ_2, σ_1^2 and σ_2^2.

ii. **(6 marks)** Find the UMVUE of σ_1/σ_2.

(d) Assume that $\mu_1 \in \mathcal{R}$, $\mu_2 \in \mathcal{R}$, $\sigma_1^2 = \sigma_2^2 = \sigma^2 > 0$.

i. **(1 mark)** Find the set of minimal sufficient and complete statistics for the unknown parameters of μ_1, μ_2 and σ^2.

ii. **(4 marks)** Using the fact that sample mean and sample variance are independent, find the UMVUE of $(\mu_1 - \mu_2)/\sigma$.

3. Suppose that we have two independent random samples: $X_1, \ldots, X_n$ are exponential(θ), with density

$$f(x|\theta) = \theta e^{-\theta x}, \quad x > 0$$

and $Y_1, \ldots, Y_m$ are exponential (μ).

(a) **(4 marks)** Find the UMP test for testing $H_0 : \theta \leq \theta_0$ against $H_1 : \theta > \theta_0$ at $\alpha = 0.05$.

(b) **(2 marks)** Based on the test derived in part (a), determine the minimum sample size n required to obtain a power of at least 0.95 when $\theta_0 = 10$, $\theta_1 = 25$ and $\alpha = 0.05$.

(c) **(4 marks)** Find the expression of likelihood ratio, $\lambda(X_1, \ldots, X_n, Y_1, \ldots, Y_m)$, for testing $H_0 : \theta = \mu$ against $H_1 : \theta \neq \mu$.

(d) **(4 marks)** Hence or otherwise, find the likelihood ratio test for testing $H_0 : \theta = \mu$ against $H_1 : \theta \neq \mu$ at $\alpha = 0.1$, $n = 10$ and $m = 15$.

(e) **(4 marks)** Derive the approximate large sample likelihood ratio test for testing $H_0 : \theta = \mu$ against $H_1 : \theta \neq \mu$ at the significance level of α and a large values of n and m. Make your conclusion at $\alpha = 0.05$ if $\sum x_i = 100$, $\sum y_i = 50$ and $n = m = 50$. Write down the value of test statistic and critical value clearly.

4. (a) Individuals were classified according to their answers of the question: "Did you get married before you were 25?" and according to which ethnic group they are, i.e.,

	Group A	Group B
Yes	x_{11}	x_{12}
No	x_{21}	x_{22}

Let $X = (X_{11}, X_{12}, X_{21}, X_{22}) \sim$ multinomial $(n, P_{11}, P_{12}, P_{21}, P_{22})$.

Suppose that 100 females sampled from each of two ethnic groups. 62 females and 29 females said "Yes" for group A and group B, respectively. Perform the following tests at

$\alpha = 0.05$. State clearly the value of test statistic, critical value and your conclusion for each test. There is no need to make the continuity correction for the Pearson's goodness of fit test.

i. (**6 marks**) Test whether the null hypothesis $H_0 : P_{11} = p^2, P_{12} = p(1-p), P_{21} = p(1-p), P_{22} = (1-p)^2$ is true by

A. Approximate large sample likelihood ratio test;

B. Pearson's goodness of fit test.

ii. (**6 marks**) Test whether the proportions of answering "Yes" for the two groups are equal at 0.05 level of significance by

A. z test;

B. Approximate large sample likelihood ratio test;

C. Pearson's goodness of fit test.

(b) Let $(X_1, ..., X_n)$ be a random sample from $U(0, \theta)$ with $\theta > 0$.

i. (**4 marks**) Find UMP test at the level of significance α for testing $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$.

ii. **(2 marks)** Based on the test derived in part (i), calculate the minimum sample size n such that the test for testing $H_0 : \theta \leq \frac{1}{2}$ versus $H_1 : \theta > \frac{1}{2}$ has a power of at least 0.98 at $\theta_1 = \frac{3}{4}$, where $\theta_1 \in \Theta_1$, when $\alpha = 0.05$.

iii. **(2 marks)** Based on the test derived in part (i), calculate the power at $\theta_1 = \frac{2}{3}$, where $\theta_1 \in \Theta_1$, for testing $H_0 : \theta \leq \frac{1}{2}$ versus $H_1 : \theta > \frac{1}{2}$ when $\alpha = 0.05$ and $n = 10$.

5. (**Bonus**) Let $(X_1, ..., X_n)$ be a random sample from a location distribution family

$$f(x; \theta) = \frac{1}{\theta} \exp\left(-\frac{x - \delta}{\theta}\right) I(x \geq \delta) .$$

(a) (**5 marks**) Suppose that θ is known. Find an exact likelihood ratio test at the level of significance α for testing $H_0 : \delta = \delta_0$ versus $H_1 : \delta \neq \delta_0$.

(b) (**5 marks**) When both θ and δ are unknown, find an exact likelihood ratio test at the level of significance α for testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.