1. A set of engine bearings are produced in a factory and an inspection device is adopted. For every n bearings produced ($n > 1$), the number of defective bearings is immediately reported. Let $r_1, r_2, \ldots, r_m$ be numbers reported among mn bearings produced. We are interested in the parameter p, the defective rate.
   a. (6%) Find the maximum likelihood estimator of p and give its limiting distribution as $m \to \infty$.
   b. (14%) Suppose that the inspection device reports r_i only if $r_i > 0$ and nothing is reported if $r_i = 0$. Upon observing $r_1, r_2, \ldots, r_m$ ($r_i > 0$ for all i), find the maximum likelihood estimator of p and derive its limiting distribution as $m \to \infty$.

2. Suppose you have a function $f(x, y)$ defined on $[0,1] \times [0,1]$. You pick values $x_1, \ldots, x_n$ independently from the uniform distribution on $[0,1]$, and values $y_1, \ldots, y_m$ also independently from the uniform distribution on $[0,1]$. You now evaluate the function f at the mn points: you obtain $f(x_i, y_j)$, for $i = 1, \ldots, m$, $j = 1, \ldots, n$. Let

$$\theta = \int_0^1 \int_0^1 f(x,y)\,dx\,dy,$$

   and let

$$\hat{\theta} = \frac{1}{mn} \sum_{i=1}^{n} \sum_{j=1}^{m} f(x_i, y_j).$$

   a. (5%) Prove that $\hat{\theta}$ is an unbiased estimate of θ.
   b. (11%) Derive $Var(\hat{\theta})$ and present the result in the form of integrals.
   c. (4%) Propose an unbiased estimate of $Var(\hat{\theta})$.

3. Consider the two-parameter exponential family with density

$$\frac{1}{\theta} e^{-(x-\mu)/\theta} 1_{\{x > u\}}.$$

   Suppose that n independent and identically distributed observations, $X_1, X_2, \ldots, X_n$ are available.
   a. (4%) If θ is known, show that $X_{(1)}$, the minimum order statistic, is sufficient for μ.
   b. (8%) For both θ and μ unknown, show that the maximum likelihood estimate of θ, $\hat{\theta}$ is the average of the X_i minus $X_{(1)}$.
   c. (8%) Show that $\hat{\theta}$ is a consistent estimator of θ.

4. Let $Y_i \sim N(\alpha_0 + \beta_0 x_i, \sigma^2)$, $i = 1, \ldots, n$, $Z_j \sim N(\alpha_1 + \beta_0 x_j, \sigma^2)$, $j = 1, \ldots, m$, be independent random variables.

   a. (4%) Derive the maximum likelihood estimates of α_0, α_1, and β_0.

   b. (8%) Suppose that $m \le n$. Give condition on $x_1, \ldots, x_n$ to guarantee that the maximum likelihood estimates of α_0, α_1, and β_0 are consistent.

   c. (8%) How would you test the hypothesis that α_0 and α_1 are equal? In your answer, you should state both hypotheses, specify the testing procedure with proper table to check, and explain why the test works.

5. (12%) Let X and Y be two random variables with finite variance. Determine a and b which minimize $E(Y - a - bX)^2$.

6. (8%) Among families with two children, let X be the score on a statistics test taken by the first child at age 21, and let Y be the income of the second child at age 40. Suppose the pair (X, Y) has a bivariate normal distribution, and $E(X) = 65$, $Var(X) = 100$, $E(Y) = 1$ million per year, $\sqrt{Var(Y)} = 200,000$ per year, and $corr(X, Y) = 1/2$. [All of this is fiction.] Among families in which the first child scores 75 on the statistics test at age 21, in what proportion of cases does the second child have an income of at least $1,250,000$ per year at age 40? (It is not necessary to give the exact numerical value. But the analytic detail should be given.)