1. A chemical engineer is investigating the yield of a process. There are \( k \) process variables, \( x_1, \ldots, x_k \), and each variable can be run at a low and a high level. Thus this engineer runs a \( 2^k \) full factorial design with \( n \) center points, and \( x_i, i = 1, \ldots, k \) are defined on a coded scale from \(-1\) to \(1\). To represent the results, this engineer decides to fit a main effects only model, say

\[
y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + \epsilon,
\]

where \( \epsilon \) is the random error. Find the relationship between the least-squares estimates of \( \beta_i, i = 1, \ldots, k \) and main effect estimates. (10 points)

2. Find the least-squares estimator of \( \beta \) in the model \( y = X\beta + \varepsilon \) subject to a set of equality constraints on \( \beta \), say \( T\beta = c \). (10 points)

3. Consider the single-factor fixed-effects analysis of variance with 3 treatments:

\[
y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad i = 1, 2, 3, \quad j = 1, \ldots, n,
\]

where \( \varepsilon_{ij} \) is an normal error with zero mean and variance \( \sigma^2 \). This model can be expressed in terms of the regression model with two indicate variables.

(1) Find the least-squares estimates of these regression coefficients and show the relationship with estimates of \( \mu \) and \( \tau_i \). (5 points)

(2) Show how this regression model could be used to test the hypothesis, \( H_0 : \tau_1 = \tau_2 = \tau_3 = 0 \), and write down the corresponding test statistic and its distribution under \( H_0 \). (10 points)

4. Consider a \( p \)-dimensional random vector, \( U \), with zero mean vector and covariance matrix, \( \Sigma \). Let \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \) be the eigenvalues of \( \Sigma \), and \( P_1, \ldots, P_p \) be the corresponding eigenvectors.

(1) Find a \( p \times p \) matrix \( W \) such that the components of \( Z = WU \) are uncorrelated and their variances are all equal to 1. (5 points)

(2) Let \( B \) be any \( p \times 1 \) vector such that \( \|B\| = 1 \). Show that

\[
\max_{B \perp P_1, \ldots, P_{i-1}} \text{Var}(B'U) = \lambda_i,
\]

and “max” is attained when \( B = P_i \). (5 points)

(3) Let \( E_p \) denote the \( p \)-dimensional Euclidean space. Show

\[
\min_{S_{i-1}} \max_{\|B\|=1, B \perp S_{i-1}} \text{Var}(B'U) = \lambda_i,
\]

where \( S_{i-1} \) is a space of \((i - 1)\) dimensions in \( E_p \). (10 points)
5. Let \( X = (X_1, \ldots, X_p)' \) come from a \( p \)-dimensional multivariate normal distribution with mean vector \( \mu \) and covariance matrix \( \Sigma \), and \( A \) be a \( p \times p \) symmetric matrix.

(1) Show that the covariance of \( X \) with \( X'AX \) is

\[
Cov(X, X'AX) = 2\Sigma A\mu. 
\]

(5 points)

(2) Prove that \( X'AX \) and \( BX \) are distributed independently if and only if \( B\Sigma A = 0 \). (10 points)

6. Consider the simple regression model is

\[
y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon,
\]

where \( E(\varepsilon) = 0 \) and \( Var(\varepsilon) = \sigma^2 \). Suppose that \( n > k \) uncorrelated observations are available, and let \( y_i \) denote the \( i \)th observed response and \( x_{ij} \) denote the \( i \)th level of regressor \( x_j \). Here inverse matrix of \( \left( X_n'X_n \right) \) is assumed to exist, where \( X_n \) is the model matrix of these \( n \) observations. Let \( \beta = (\beta_0, \beta_1, \ldots, \beta_k)' \) be the coefficient vector. Suppose the \((n+1)\)th observation is obtained. Compute the covariance matrix of current least-squares estimator, \( \hat{\beta}_{n+1} \), based on \( \left( X_n'X_n \right)^{-1} \). (15 points)

7. Consider the \( 2^{6-2} \) design with the generators \( I = ABCE \) and \( I = BCDF \).

(1) Construct this design. (5 points)

(2) Show the complete alias structure of this design. (10 points)