(1) The hiring manager Tom posed the following question in a job interview: Each of factors \( A \) and \( B \) separately causes significant changes on variable \( y \). However, when acting together, the collective effect of \( A \) and \( B \) on the change of \( y \) is not obvious at all. Why is that? As a statistician, please formulate the problem in an unambiguous statistical setting that includes (1a) well defined variables; (1b) a clear criterion under which people can compare different models and procedures; (1c) a reasonable answer to Tom’s question.

(2) Consider the regression model
\[
y_i = \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \epsilon_i, \quad i = 1, \ldots, n.
\]

(2a) In a controlled experiment, the covariates \( x_{ij} \) are assumed to be known, with unknown coefficients \( \beta_j, \ j = 1, \ldots, p; \) and \( \epsilon_i, \ i = 1, \ldots, n \) are iid random errors with mean 0 and an unknown variance \( \sigma^2 \). Suppose the sample size \( n \) is moderate and it is not possible to collect more data (due to some unexpected experimental difficulties). Propose a detailed procedure to obtain confidence intervals for \( \beta_1, \ldots, \beta_p \).

(2b) In an observational social study, the covariates \( \{x_{ij}\} \) are also included in the given data set together with \( \{y_i\} \), but cannot be treated as fixed values. What procedure would you propose to obtain confidence intervals for \( \beta_1, \ldots, \beta_p \) based on the observed data \( \{(y_i; x_{i1}, \ldots, x_{ip}) : i = 1, \ldots, n\}\)?

(3) Consider the regression model
\[
y = X\beta + \epsilon \quad \text{with observations} \quad y = (y_1, \ldots, y_4)^t, \quad \text{unknown coefficients} \quad \beta = (\beta_1, \beta_2, \beta_3)^t, \quad \text{iid errors} \quad \epsilon = (\epsilon_1, \ldots, \epsilon_4)^t \quad \text{with mean zero and unknown variance} \quad \sigma^2, \quad \text{and design matrix} \quad X = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 2 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix}.
\]

(3a) Find an unbiased estimate for \( \beta_2 \) based on \( y \)? Can you find the BLUE for \( \beta_2 \)? Explain.

(3b) Is there identifiability issue for this model? If so, how would you resolve it?

(3c) Is it possible to test \( H_0 : \beta_2 = 0 \)? If so, provide the test statistics. If not, explain.