1. **(5 points)** A large pharmaceutical company is interested in testing the uniformity (a continuous measurement that can be taken by a measurement instrument) of their film-coated pain-killer pills produced at various blending sites. Among the large pool of blending sites, four were randomly chosen. In each site, a large number of batches of such pills were produced regularly. A random sample of five batches were collected from each of the four blending sites. Three pills were assayed from each batch. Carefully explain how the experiment should be conducted and how the data collected from such experiment should be analyzed. Be specific about the ANOVA table.

This is a fully nested design in which factor A (blending site) has \( a = 4 \) levels and factor B (batch), which is nested within factor A, has \( b = 5 \) levels (batches). Each batch has \( n = 3 \) replicates (pills). Both factors are random. The effect model can be expressed as

\[
y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}, \quad i = 1, \cdots, 4, j = 1, \cdots, 5, k = 1, \cdots, 3,\]

where

- \( \mu \) = overall population mean
- \( \alpha_i \) = \( i \)th random effect of factor A
- \( \beta_{j(i)} \) = \( j \)th random effect of factor B nested within the \( i \)th level of factor A
- \( \varepsilon_{ijk} \) = error (\( N = 60 \) i.i.d. \( N(0, \sigma^2) \)).

The model assumptions are listed below:

- \( \alpha_1, \cdots, \alpha_4 \) form a random sample of \( N(0, \sigma^2_\alpha) \);
- \( \{\beta_{1(i)}, \cdots, \beta_{5(i)}\} \) form a random sample of \( N(0, \sigma^2_\beta) \) for each \( i = 1, \cdots, 4 \); all \( ab + abn = 20 + 60 = 80 \) r.v.’s are independent.

The sums of squares are

(a) \( SS_T = \sum_{i=1}^{4} \sum_{j=1}^{5} \sum_{k=1}^{3} y_{ijk}^2 - T^2 \cdot 60 \)

(b) \( SS_E = \sum_{i=1}^{4} \sum_{j=1}^{5} \sum_{k=1}^{3} y_{ijk}^2 - \sum_{i=1}^{4} \sum_{j=1}^{5} T_{ij}^2 / 3; MS_E = SS_E / 40. \)

(c) \( SS_A = \sum_{i=1}^{4} T_{i..}^2 / 15 - T^2 \cdot 60; MS_A = SS_A / 3. \)

(d) \( SS_{B(A)} = \sum_{i=1}^{4} \sum_{j=1}^{5} T_{ij}^2 / 3 - \sum_{i=1}^{4} T_{i..}^2 / 15; MS_{B(A)} = SS_{B(A)} / 16. \)
Hypotheses of interest:

\[ H_0^A : \sigma_a^2 = 0; \quad H_0^B : \sigma_b^2 = 0. \]

The ANOVA table is given by

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>E[MS]</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>( SS_A )</td>
<td>3</td>
<td>( MS_A )</td>
<td>( \sigma^2 + 3\sigma_b^2 + 15\sigma_a^2 )</td>
<td>( MS_A/MS_{B(A)} )</td>
</tr>
<tr>
<td>B( \omega ) A</td>
<td>( SS_{B(A)} )</td>
<td>16</td>
<td>( MS_{B(A)} )</td>
<td>( \sigma^2 + 3\sigma_b^2 )</td>
<td>( MS_{B(A)}/MS_E )</td>
</tr>
<tr>
<td>Residual</td>
<td>( SS_E )</td>
<td>40</td>
<td>( MS_E )</td>
<td>( \sigma^2 )</td>
<td>( \quad )</td>
</tr>
<tr>
<td>Total</td>
<td>( SS_T )</td>
<td>59</td>
<td>( \quad )</td>
<td>( \quad )</td>
<td>( \quad )</td>
</tr>
</tbody>
</table>

2. Two coded variables \( x_1 = (A - 90)/8, \ x_2 = (B - 17)/12 \) are examined, and the data below are obtained. It is known that the standard deviation of an observation is 2.

<table>
<thead>
<tr>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( y )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>-</td>
<td>16</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>25</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>8</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>13</td>
</tr>
</tbody>
</table>

(a) (3 points) Fit a first-order model (that is, planar model) and check the appropriateness of the model.
First of all, the standard error of effect is

\[ \text{std.error(effect)} = \frac{2}{\sqrt{n_f}} \sigma = \frac{2}{\sqrt{4}} \times 2 = 2 \]

and an effect can be judged as significant if its magnitude is greater than 4, twice the standard error of effect. The effects are

\[
\begin{align*}
A &= \frac{3 + 8}{2} - \frac{16 + 25}{2} = -15 \\
B &= \frac{25 + 8}{2} - \frac{16 + 3}{2} = 7 \\
AB &= \frac{16 + 8}{2} - \frac{3 + 25}{2} = -2 
\end{align*}
\]

Both main effects are significant. The interaction is insignificant. Moreover, the curvature effect can be checked:

\[
\bar{y}_f - \bar{y}_c = \frac{16 + 3 + 25 + 8}{4} - 13 = 0 
\]

with a standard error for the curvature effect:

\[
\text{std.error(curvature)} = \sigma \sqrt{\frac{1}{4} + 1} = 2 \times \sqrt{\frac{5}{4}} = 2.236.
\]

We have zero curvature effect. The planar model seems to be appropriate.

(b) (2 points) Disregard the appropriateness/inappropriateness of the first order model, use the first-order model in (a) to answer the question below:

Is the point \((A, B) = (60, 38)\) on the path of steepest ascent?

The first order model is given by

\[ \hat{y} = 13 - 7.5x_1 + 3.5x_2 \]

The steepest ascent direction is simultaneously moving a multiple of \(-1\) in \(x_1\) and \(\frac{3.5}{7.5} = \frac{7}{15}\) in \(x_2\). Note that \(x_1 = \frac{60 - 90}{8} = -3.75\) when \(A = 60\) and \(x_2 = \frac{38 - 17}{12} = 1.75 = 3.75 \times \frac{7}{15}\) when \(B = 38\). Hence, the point \((A, B) = (60, 38)\) is on the path of steepest ascent.

3. In the second chemical example by Montgomery (the data are reproduced here), focus on the response \textit{yield} only.
The fitted quadratic model is

\[ \hat{y}_1 = 79.4 + 0.99x_1 + 0.52x_2 - 1.38x_1^2 - 1.00x_2^2 + 0.25x_1x_2. \]

The estimated error is \( \hat{\sigma} = 0.266. \)

(a) **(2 points)** Check the adequacy of exact quadratic in \( x_1 \) direction.

The standard error is

\[ \text{std.error}(\hat{\beta}_{111} - \hat{\beta}_{122}) = 0.266 \times \sqrt{\frac{1}{8} + \frac{1}{8} + \frac{1}{4}} = 0.1881. \]

The exact quadratic in \( x_1 \) direction seems to be justified since the magnitude of the estimate is less than twice the standard error.

(b) **(2 points)** What type of response surface for \( \text{yield} \)? Why?  (Use the attached computer output to answer the question.)

The response surface in the design range is approximately a hill with simple maximum since the eigenvalues are both negative and comparable in magnitude and the stationary point (0.386624, 0.308328) is located within the design region.
4. **(3 points)** A 12-run composite design is performed. The observed responses from the design is shown on the figure below. Comment on the adequacy of the first-order model simply by inspecting this figure. Hint: How do you draw the contour of responses?

The first-order model is inappropriate since the response surface contour appears to be non-linear (that is, one can not capture the response contour with a system of parallel lines).

5. **(3 points)** Identify the design below using the notation $2^{k-p}_R$: basic defining relations where $k =$ # of factors, $p =$ # of basic words, and $R =$ Resolution. State the reasons of your answer.
This is a $2^{5-2}_{IV}$ design, reasoned below. Note that factors A, B, C, and D form a full factorial (though not displayed in standard order). Factor E takes on the reversed signs of BCD and factor F is completely confounded with ABC. The complete set of generators should also include EF = −AD. The complete set of defining relations is $I = −BCDE = ABCF = −ADEF$ with shortest word length of 4. Consequently, this is a $2^{5-2} = \frac{1}{4}$-fraction resolution IV design in $2^{6-2} = 2^4 = 16$ runs with 6 factors and 2 basic generators E = −BCD, F = ABC.
Computer Output for Problem 3(c)

==============================
= CANONICAL ANALYSIS =
==============================

<EIGENVALUES and STATIONARY POINT (in original coordinates)>

Stationary Row Eigenvalues Point
1 -1.41743 0.386624
2 -0.96257 0.308328

<EIGENVECTORS>:

Matrix Eigenvectors

0.957971 0.286865
-0.286865 0.957971

<CONSTANT IN TRANSFORMED COORDINATES>:

YSHAT 79.6715

The sorted absolute eigen values are:

|Eigenvalues|
0.96257 1.41743