Theoretical questions for the exam

1. **Confidence Intervals and Prediction Intervals**

   You have measured a random variable X, which is assumed to follow a Normal distribution, 20 times. The measurements yield the following results :
   — Sample mean (m) : 25
   — Sample standard deviation (s) : 3
   Answer the following questions :

   a) Determine the 95% confidence interval for the population mean (μ).

   b) If you want to predict the outcome of the next measurement with 68% confidence, what prediction interval should you use ?

2. **Geometric Interpretation of Linear Regression**

   Linear regression involves fitting a model to data in order to explain the relationship between variables. Using your understanding of geometry, provide a detailed interpretation of how linear regression can be visualized in a multidimensional space.

   In your explanation, address the following points :

   a) How can the data points, regression line, and residuals be represented geometrically ?

   b) What is the significance of minimizing the residuals in terms of distance or projection ?

   c) How does the dimensionality of the space affect the interpretation ?

3. **Structure and Interpretation of an ANOVA Table**

   The Analysis of Variance (ANOVA) is a powerful tool used to compare means and analyze variability in data.

   a) Describe the structure of a standard ANOVA table. What are the key components included in the rows and columns ?

   b) For each row of the table, explain its purpose and the type of variability it represents.

   c) For each column of the table, provide a clear interpretation of the values.

   Be precise in your explanation and highlight the relationships between the components.

4. **Calculating the Sum of Squares for Lack of Fit**

   Consider a one-factor linear model with 3 distinct levels of the factor, each replicated 3 times, resulting in a total of 9 observations.

   a) Explain the concept of lack of fit in a regression model. How does it differ from pure error ?

   b) Provide a step-by-step procedure to calculate the sum of squares for lack of fit (SS_{LoF}) in this example.

   c) Ensure to include the following steps :
      — Partitioning the total sum of squares (SS_{Total}) into components.
      — Computing the pure error sum of squares (SS_{PE}). Deriving the lack of fit sum of squares (SS_{LoF}) as a residual component.
      — Interpret the meaning of a significant lack of fit and its implications for the adequacy of the model.

5. **The Dual Meanings of "Contrast" in Experimental Design**

   The term "contrast" is used in different contexts within experimental design, particularly in :
   — The analysis of a constant coefficient model.
   — The construction and interpretation of fractional factorial designs.
   Explain the meaning of "contrast" in each situation, addressing the following points :

   a) Define the term "contrast" in the context of a constant coefficient model and describe its role in hypothesis testing or comparison of effects.

   b) Define the term "contrast" in the context of fractional factorial designs and explain its role in estimating effects or resolving aliasing.

   c) Highlight the differences and similarities in the use of the term in these two contexts.

6. **Sums of Squares Types I and II**

   a) Define sums of squares Type I and Type II in the context of Analysis of Variance (ANOVA).

   b) Explain the key differences in how they are calculated and interpreted.

7. **Evaluating the Quality of a Design Matrix E**

   a) What are the key criteria or elements that can be used to assess the quality of a design matrix$E$ in experimental design ?

   b) Define and explain the importance of each criterion in ensuring a robust and effective design.

   c) Discuss how these criteria are applied in practice to select or improve a design.

8. **Canonical Analysis**

   a) Define what canonical analysis is and explain its purpose in the context of response surface methodology (RSM).

   b) Describe the steps involved in performing a canonical analysis, emphasizing how it transforms a response surface into a simpler mathematical form.

   c) Provide a few examples of possible results from a canonical analysis and explain their interpretation in terms of optimization or decision-making.

9. **Scheffé Model in Mixture Problems**

   Mixture experiments involve designing and analyzing experiments where the factors are proportions of components that must sum to 1. In this context :

   a) Explain the purpose of using a Scheffé polynomial model in mixture problems. What are its key advantages when modeling response surfaces in constrained spaces (such as mixtures) ?

   b) Discuss the limitations or challenges of the Scheffé model. In which situations might it be less appropriate ?

   c) Describe two alternative modeling approaches for mixture problems.

10. **Concentration Constraints in a Ternary Diagram**

    In mixture experiments, ternary diagrams are commonly used to represent the proportions of three components. Answer the following :

    a) How can minimum and maximum concentration constraints for components be represented graphically on a ternary diagram ?

    b) Discuss how these constraints modify the feasible experimental region within the ternary diagram.

    c) What are the potential consequences of these constraints on the design of experiments, such as the choice of experimental points, the applicability of standard mixture models, and the complexity of the design ?

11. **Graeco-Latin Squares**

    a) Define what a Graeco-Latin square is and explain how it is constructed.

    b) Describe the conditions under which a Graeco-Latin square design is useful in experiments.

    c) Provide an example of how to use a Graeco-Latin square in a real experimental setup.

12. **Construction of a 2^{8-3} Fractional Factorial Design**

   a) Explain what a 2^{8-3} fractional factorial design is and describe its key features, including resolution and aliasing structure.

   b) Provide a step-by-step guide for constructing such a design, ensuring to include the following :
   — Choice of generators for the design.
   — Identification of the principal fraction.
   — Assignments of generators to factors

13. **Optimal Strategy for Experimentation with 9 Continuous Factors**

   Consider an experimental situation with the following characteristics :
   — Nine continuous factors (x_i) to study.
   — Each factor has a defined range (Δx_i) within a 9-dimensional experimental space.
   — It is likely that not all factors are of critical importance.
   Propose an optimal strategy, represented as a flowchart or logigram, to design an efficient experiment. The strategy should address the following :

   a) Identifying critical factors among the nine.

   b) Exploring the experimental space systematically while minimizing the number of experiments.

   c) Refining the exploration as more information about the system becomes available.

14. **Rechtschaffner Designs**

   a) Define what a Rechtschaffner design is, including its general structure.

   b) Explain the key characteristics that distinguish Rechtschaffner designs from other fractional factorial designs.

   c) Discuss the typical use cases for Rechtschaffner designs and their advantages in experimental settings.

15. **Doehlert Designs**

   a) Define what a Doehlert design is and explain its structure, particularly how it differs from other response surface designs.

   b) Describe the main characteristics of Doehlert designs, including their geometric representation and flexibility.

   c) Explain the advantages and limitations of using a Doehlert design.

   d) Provide an example of a practical scenario where a Doehlert design would be an optimal choice, and explain why.

16. **Simplex Lattice Designs**

   a) Define what a simplex lattice design is and explain its structure in the context of mixture experiments.

   b) Describe the key characteristics of simplex lattice designs, including how they differ from other designs like factorial or Doehlert designs.

   c) Discuss the typical use cases for simplex lattice designs and their advantages in mixture experiments.

   d) Provide an example of a practical scenario where a simplex lattice design would be the optimal choice and explain why.

17. **Generators of a Fractional Factorial Design**
   Given the generators for a fractional factorial design :

$$I = 1235 = -2346$$

   a) Explain the meaning of the generators and how they define the structure of the design.

   b) Set up the table of contrasts for the factors and interactions based on the given generators.

18. **Alias Matrix in a Plackett-Burman Design**
   Explain how the alias matrix A is calculated in a Plackett-Burman design and its purpose in identifying potential biases caused by first-order interactions on the main effects.

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

   a) Analyze the given alias matrix A to determine which main effects are potentially biased by interactions.

   b) Provide recommendations for interpreting the results and mitigating biases.