Design of Experiments

Yavuz Shahzad
Winter 2025
Prof. Alia Sajjad

Contents

Introduction

This document provides an overview of concepts in statistically designed experiments, following the text *Design of Comparative Experiments* by R.A. Bailey.

Stages in a Statistically Designed Experiment

Consultation

A statistician's role in an experiment ideally begins early in the design process. Effective consultation requires understanding the purpose, available resources, and timeline of the study.

Example 1.1 (Ladybirds): A pesticide company tested a new pesticide, a standard pesticide, and a control (no treatment). Misconceptions about randomization led to flawed conclusions. (Bailey, p.2)

Data Collection

A well-designed experiment ensures that data collection is structured for reliable analysis.

- Each observational unit should have its own row in the dataset.
- Treatments should be assigned in a structured manner.
- Data should be retained in raw form to prevent errors.

Data Scrutiny

After data collection, a statistician should inspect for anomalies.

Example 1.3 (Leafstripe): A data entry mistake caused an extreme outlier, highlighting the importance of careful data review.

Data Analysis

Planning data analysis prior to conducting the experiment ensures that appropriate statistical methods are used.

Key Experimental Concepts

Replication

Replication improves precision and generalizability. It reduces the standard error, increasing statistical power.

False Replication

Repeated measurements on the same experimental unit should not be treated as independent replications.

Local Control

Blocking is used to group similar experimental units to reduce variability, thus improving efficiency.

Orthogonality and ANOVA Assumptions

ANOVA Assumptions

Analysis of variance (ANOVA) requires that errors are independent, identically distributed, and have constant variance.

- Normality of residuals can be checked using the Shapiro-Wilk test.
- A residuals vs. fitted values plot can reveal heteroscedasticity.
- Bartlett's test assesses homogeneity of variances.

Orthogonality in Experimental Design

Orthogonal designs ensure that treatment effects can be independently estimated.

Definition: The treatment subspace V_T consists of vectors that are constant within each treatment group.

Projection and Estimation in Linear Models

Orthogonal Projection

If V is the space of experimental units and V_T is the treatment subspace, then V can be decomposed as:

$$V = V_T \oplus V_T^{\perp}$$

where $V_T^{\perp}$ is the orthogonal complement of V_T.

Estimating Treatment Effects

The best linear unbiased estimator (BLUE) for a treatment effect τ_i is given by the sample mean for that treatment:

$$\hat{\tau}_i = \frac{1}{r_i} \sum_{T=i} Y_i$$

Contrast and Hypothesis Testing

A contrast is a linear combination of treatment effects that sums to zero:

$$l_m = a_1\tau_1 + a_2\tau_2 + \cdots + a_t\tau_t \quad \text{where } \sum a_i = 0.$$

Projection in ANOVA

The projection matrix for a subspace W satisfies:

$$P_W^2 = P_W, \quad P_W^T = P_W, \quad \text{and } \mathrm{Tr}(P_W) = \dim(W).$$

Conclusion

Statistical design principles help ensure valid and efficient experiments. This document provides an introduction to key ideas such as replication, blocking, and orthogonality in experimental design.