

Experimental Design Presentation

Presented by

Jim Patrie MS

Department of Health Evaluation Sciences

Division of Biostatistics and Epidemiology

University of Virginia

Charlottesville, VA

email jp4h@virginia.edu

Presentation Outline

I) Background Information.

- randomized experiments.
- randomized comparative experiments.
- observational comparative experiments.

II) Experimental Design Terminology.

- Experimental Unit.
- Replications.
- Treatments.
- Factors
- Factor Levels.
- Fix Factors
- Random Factors.

III) The Three Major Components of Experimental Design.

- Treatment Structure.
- Design Structure.
- Randomization.

IV) Examples of Commonly Utilized Experimental Designs

- Single Factor Design
- Factorial Design.
- Randomized Complete Block Design.
- Latin Square Design.
- Split Plot Design

+Treatment and Design Structure Layout, Statistical Model, ANOVA Table, and a Typically Analysis Summary.

V) Sample Size Estimation.

- Information required to determine samples size.
- Sample size formula for a two sample comparison of means.

VI) Multiple Comparison Adjustment.

- Controlling Type I error rate.

VII) Experimental Design Check List.

VIII) Listing of Useful Introductory and Intermediate Text Books on Experimental Design.

Introduction

The goal of any randomized experiment is to arrive at a clear answer to the question(s) of interest while expending a minimum amount of resources.

The goal of experimental design is to attain maximum information, precision, and accuracy in the results by efficiently utilizing existing resources.

When selecting an experimental design, first and foremost base your choice of design on those designs that will allow you to address your research questions of interest. Having clearly defined objectives is of paramount importance to appropriate design selection.

Randomized Experiments (I)

A randomized experiment is typically conducted for one or more of the following reasons:

- to compare the response achieved at different settings of the controllable variables*.
- to determine the principle cause of variation in a measured response.
- to find the conditions that give rise to maximum or minimum response.
- to obtain a mathematical model in order to predict future response.

Randomized Comparative Experiment

The distinguishing features of a randomized comparative experiment include the following:

- conditions are controlled and manipulated by the investigator.
- treatments are assigned at **random** to the experimental units.
- confounding by non-experimental factors is controlled by the randomization process.
- inferential cause and effect relationships **can** be inferred.

Observational Comparative Study

The distinguishing features of an observational comparative study include the following:

- conditions influencing the response are not under the control of the investigator.
- exposure or treatment is **not** assigned at random to the observational units.
- associations between exposure and outcome may result from confounding by factors associated with exposure and outcome.
- inferential cause and effect relationships **cannot** be inferred.

Experimental Design Terminology (II)

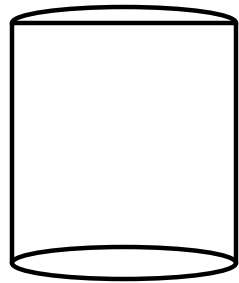
Experimental Unit

- the physical entity or subject that is exposed to the experimental treatment **independently** of other units.

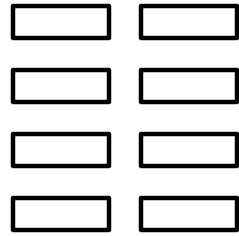
Replication

- an independent repetition of the experiment, in which an experimental unit upon exposure to the treatment constitutes a single replication of the experiment.

True Replication



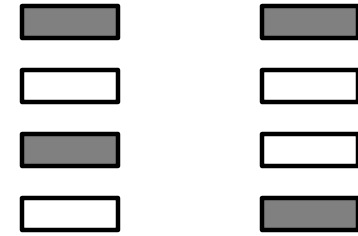
Batch
Alloy



Bars

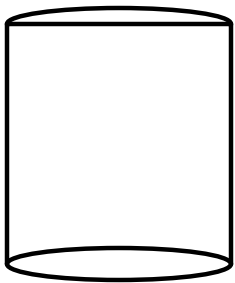


Treatment

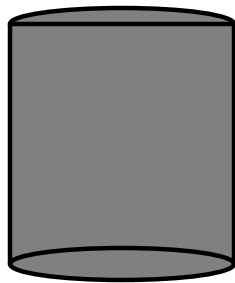


4 reps/Treatment
(E.U. Bar, n= 8)

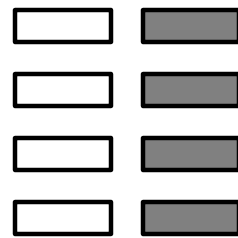
Pseudo Replication



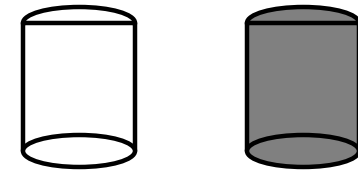
Batch
Alloy



Batch
Alloy



Bars



1 rep/Treatment
(E.U. Batch, n=2)

Treatment

- the set of circumstances created for the experiment in response to the research hypotheses.

Factor

- a particular group of treatments (e.g. passenger restraint).

Factor Levels

- the categories of the treatment factor (e.g. seat belt, air bag, seat belt plus air bag).

Fixed Factor

- the levels of the factor are assumed to be fixed (reproducible).
- inference is restricted to those levels of the factor that were included in the studied.
- the parameters of interest are the **treatment effects**.

Random Factor

- the levels of the factor are assumed to represent a random sample of the potential levels within a well defined population.
- inference is directed to the population.
- the parameters of interest are the **variance components**.

Major Components of Experimental Design (III)



Treatment Structure

The set of treatments or treatment combinations that the investigator has selected to study and/or compare.

Design Structure

The grouping of experimental units into homogenous groups; often referred to as the blocking criterion.

Randomization

The process of randomly assigning the treatments to the experimental units of the design structure.

Treatment Structure

- The treatment structure should be dictated by the research hypotheses, and constructed based on those factors to be compared as measured by their effect on the given response variable.
- The treatment structure can be a set of individual treatments, referred to as a one-way treatment structure, or a set of treatment combinations, referred to as a factorial treatment structure.
- The treatment structure can also be a hierarchical arrangement involving multiple size experiment units, in which the treatment levels of one or more factors occur within the levels of one or more of the remaining factors.

Design Structure

- The design structure of an experiment involves the grouping of experimental units such that the variability of the units within the groups is less than that among all units prior to grouping. The goal is to group the experimental units in such a way that the conditions under which the treatments are observed are as uniformed as possible.
- The design structure can range from a completely randomized design structure with no blocking criterion, to design structure which include multiple blocking criteria.

Randomization

- Randomization is the link between the treatment structure and the design structure.
- Randomization is the process of randomly assigning the treatments of the treatment structure to the experimental units of the design structure.
- Randomization functions to prevent systematic and personal biases from being introduced into the experiment.
- Randomization also functions to provide a valid estimate of error variance for justifiable statistical inference.

Experimental Designs Commonly Utilized In the Engineering Sciences (IV)

- Single Factor Designs
- Factorial and Fractional Factorial Designs
- Randomized Block and Incomplete Block Designs
- Latin Square Designs
- Split Plot Designs
- Response Surface Designs

Specific Examples

I) Designs

- Single Factor Design.
- Two Factor Factorial Design.
- Randomized Complete Block Design.
- Latin Square Design.
- Split-Plot Design.

II) Design Outline

- a description of the study protocol.
- a description of the design and treatment structure layout.
- a description of the cell means model.
- a description of the model assumptions.
- presentation of the ANOVA table.
- presentation of a typical analysis summary.

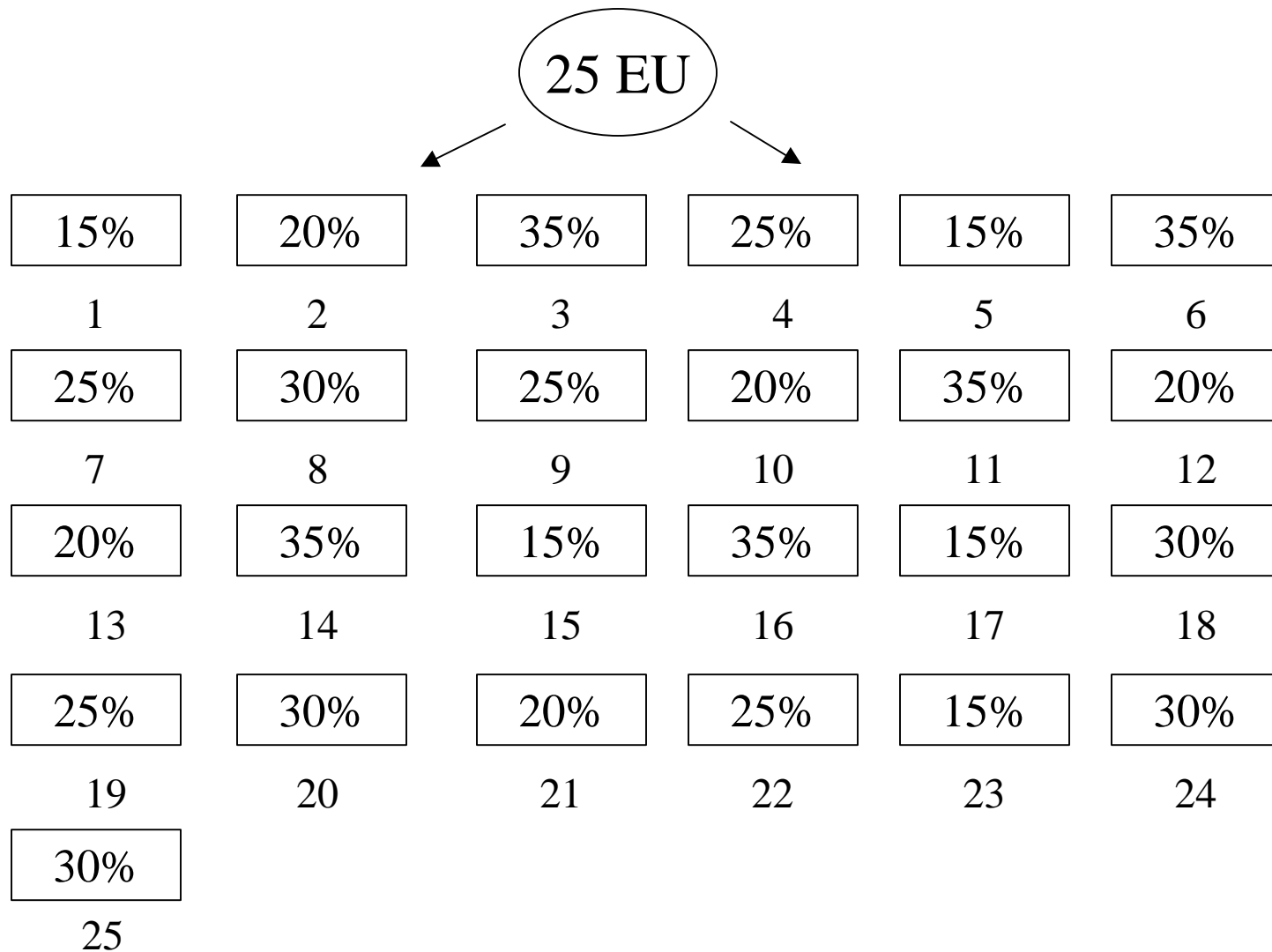
Single-Factor Design

A production development engineer is interested in maximizing a new synthetic fiber that will be used in men's clothing. The engineer knows from previous experience that the fiber's strength is affected by the percentage of cotton (c) in the fiber. Furthermore, he suspects that increasing the cotton content will increase the strength, at least initially. The engineer decides to test specimens at 5 levels of cotton percentage: 15 percent, 20 percent, 25 percent, 30 percent, and 35 percent. He also decides to test 5 specimens at each level. Twenty five machines were randomly selected to be used to produce fiber specimens at the 5 levels of percent cotton. Five machines were randomly assigned to each level. A specimen from each machine was selected at random to be evaluated for strength.

Example taken from "*Design and Analysis of Experiments*".
Montgomery, 1991. (Modified for this presentation)

Treatment and Design Structure

- one-way treatment structure, single fix factor (c), five treatment levels.
- design structure has no blocking factors.



Cell Means Model

$$y_{ij} = \mathbf{m} + \mathbf{t}_i + \mathbf{e}_{ij} \quad \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, r \end{cases}$$

Where

y_{ij} = the j th response within the i th level of % cotton (c).

\mathbf{m} = the grand mean.

\mathbf{t}_i = the effect for the i th level of % cotton (c)

\mathbf{e}_{ij} = the random error associated with the i th j th response.

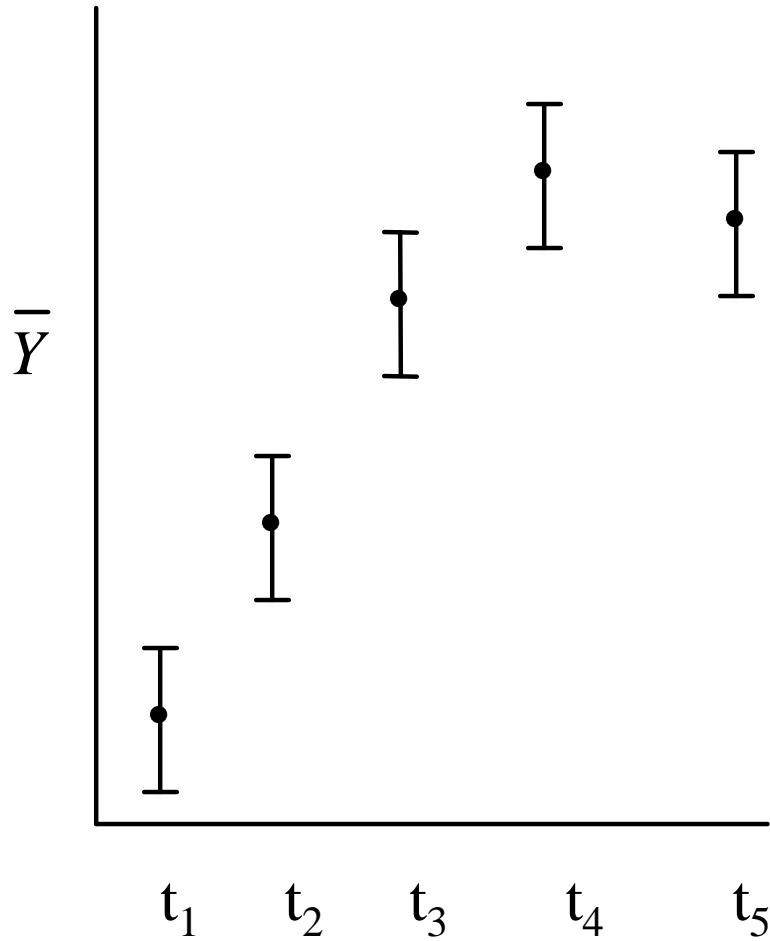
Assumptions

$$\sum_{i=1}^a \mathbf{t}_i = 0 \quad \mathbf{e}_{ij} \sim N(0, \mathbf{s})$$

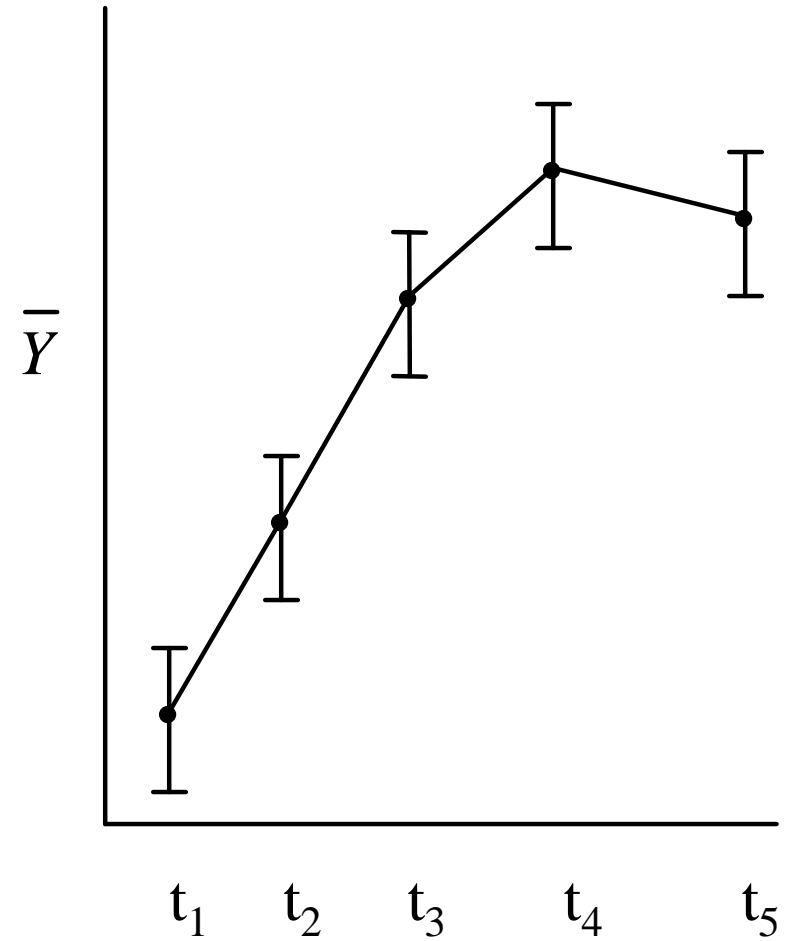
ANOVA Table

Sources	DF	DF Ex	MS	F-test
% Cotton	a-1	4	MSA	MSA/MSE
Error	a(r-1)	20	MSE	-----
Total	N-1	24		

Typical Analysis Summary



Qualitative



Quantitative

Factorial-Design

An engineer is designing a battery for use in a device that will be subjected to some extreme variations in temperature. The only design parameter that he can select at this point is the plate material (m) for the battery, and he has three possible choices. The engineer decides to test all three plate materials at three temperature (t) levels; 15°F, 70°F, and 125°F. Five batteries are tested at each combination of plate material and temperature.

Example taken from “*Design and Analysis of Experiments*”.
Montgomery, 1991.

Treatment and Design Structure

- two-way factorial treatment structure, two fix factors (m,t), with three levels per factor.
- design structure has no blocking factors.

	15°F	75°F	125°F
m ₁	m_1t_1 $n_{11}=5$	m_1t_2 5	m_1t_3 5
m ₂	m_2t_1 5	m_2t_2 5	m_2t_3 5
m ₃	m_3t_1 5	m_3t_2 5	m_3t_3 5

Cell Means Model

$$y_{ijk} = \mathbf{m} + \mathbf{t}_i + \mathbf{l}_j + (\mathbf{tl})_{ij} + \mathbf{e}_{ijk} \quad \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, b \\ k = 1, 2, \dots, r \end{cases}$$

Where

y_{ijk} = the k th response in the i th j th treatment group.

\mathbf{m} = the grand mean.

\mathbf{t}_i = the effect for the i th material (m)

\mathbf{l}_j = the effect for the j th level of temperature (t).

$(\mathbf{tl})_{ij}$ = the interaction between the i th level of (m) and j th level of (t).

\mathbf{e}_{ijk} = the random error associated with the i th j th, k th response.

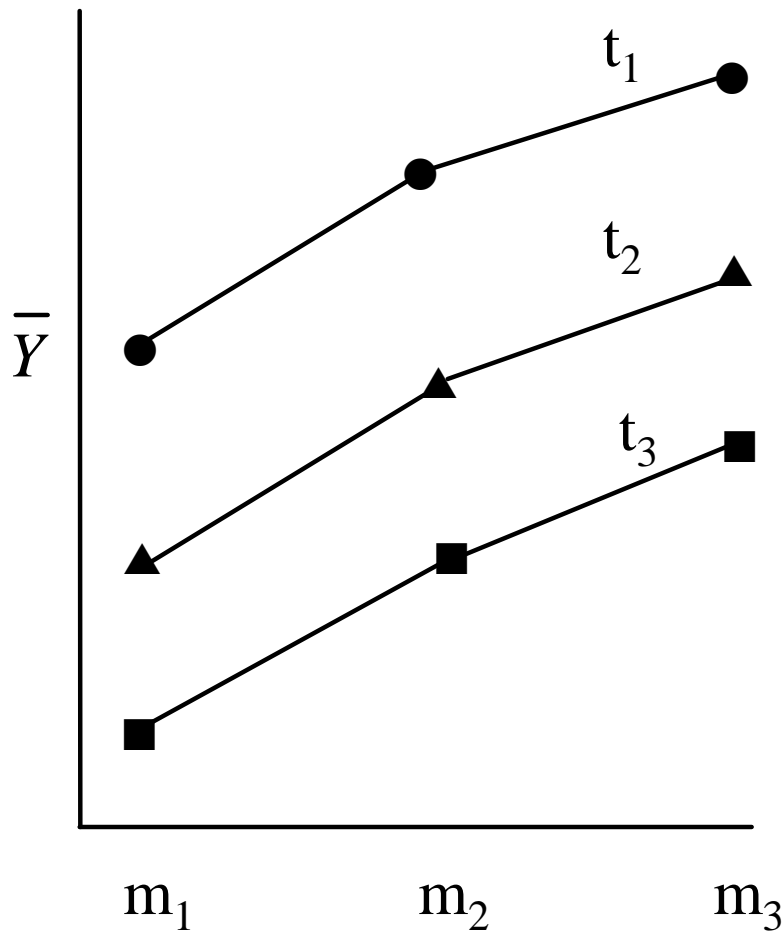
Assumptions

$$\sum_{i=1}^a \mathbf{t}_i = \sum_{j=1}^b \mathbf{l}_j = \sum_{i=1}^a \sum_{j=1}^b (\mathbf{tl})_{ij} = 0, \quad \mathbf{e}_{ijk} \sim N(0, \mathbf{s})$$

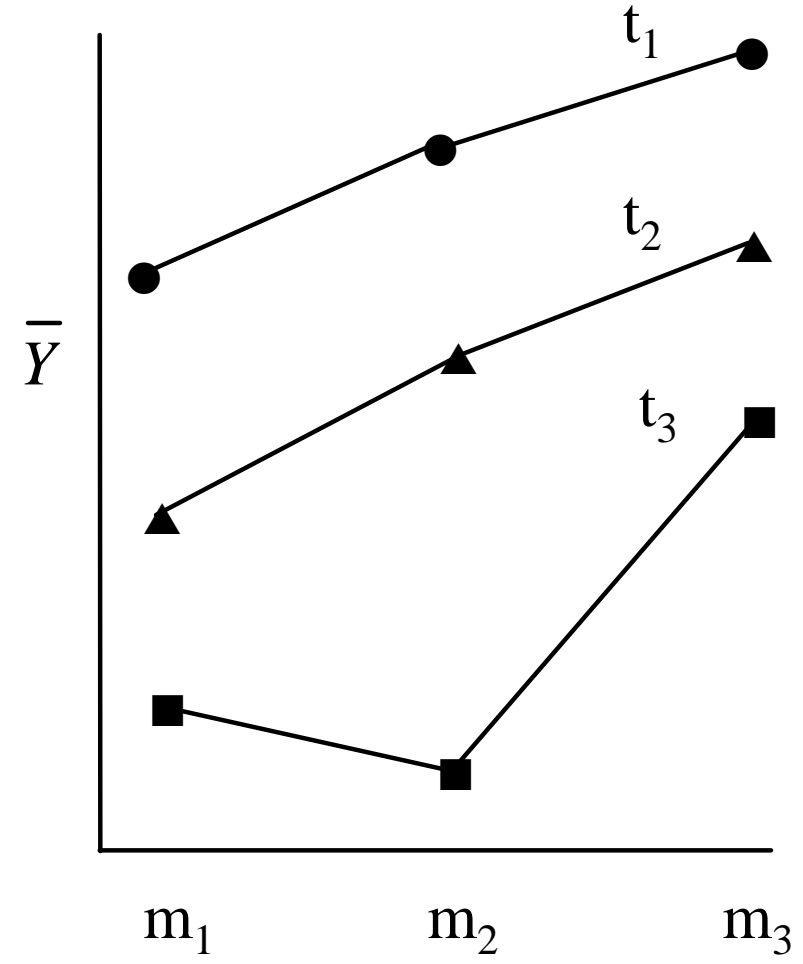
ANOVA Table

Source	DF	DF Ex	MS	F-test
Material (m)	a-1	2	MSA	MSA/MSE
Temp (t)	b-1	2	MSB	MSB/MSE
m x t	(a-1)(b-1)	4	MSAB	MSAB/MSE
Error	ab(r-1)	36	MSE	-----
Total	N-1	44		

Typical Analysis Summary



No Interaction



Interaction

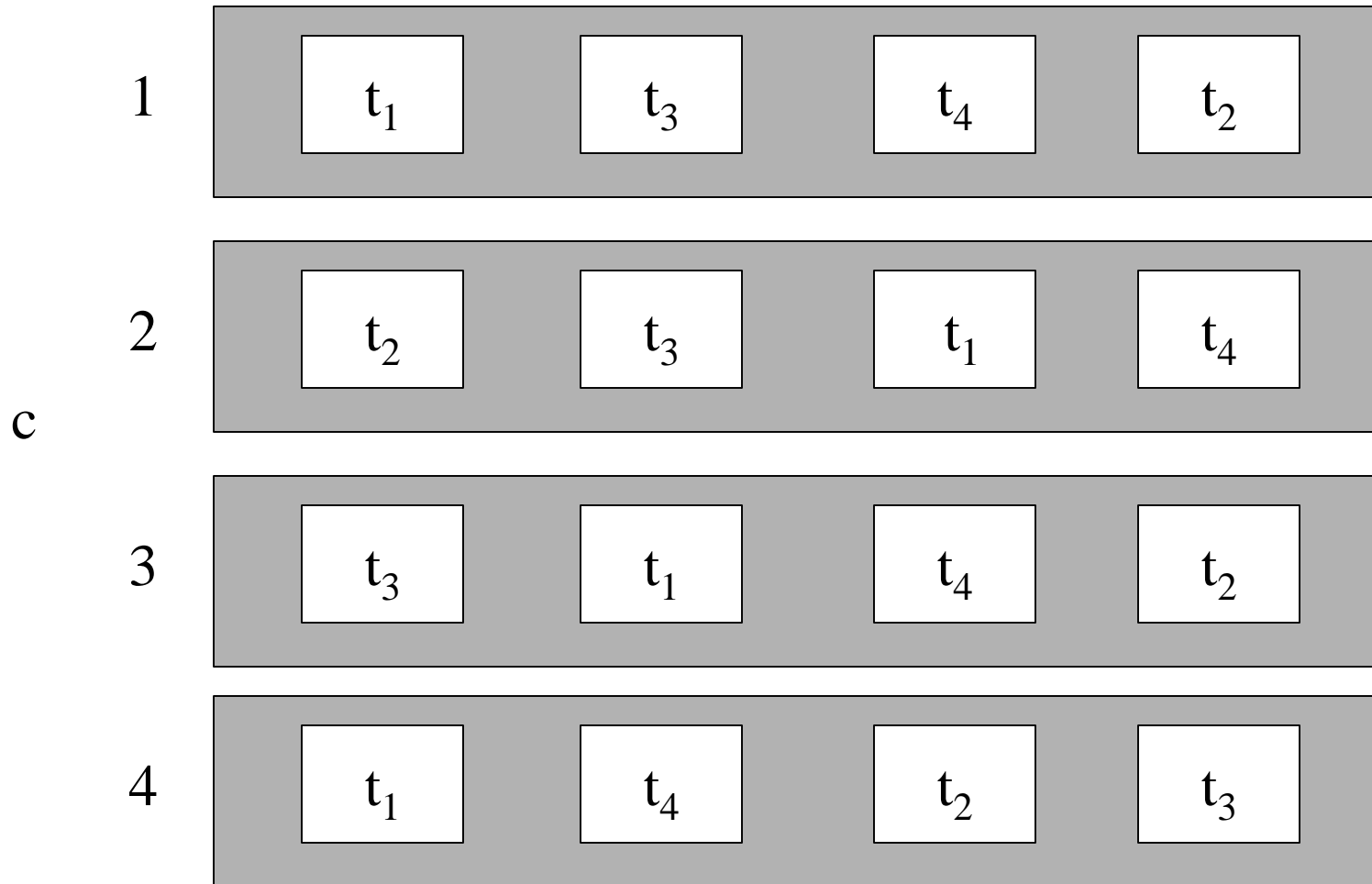
Randomized Block Design

An engineer wishes to determine whether or not four different tips (t) produce different readings on a hardness testing machine. The machine operates by pressing the tip into a metal test coupon, and from the depth of the resulting depression, the hardness of the coupon can be determined. The experimenter has decided to obtain four observations for each type of tip. To minimize coupon to coupon variability the experimenter decides to test each type of tip utilizing the same coupon, with a total of four coupons utilized in experiment.

Example taken from “*Design and Analysis of Experiments*”.
Montgomery, 1991.

Treatment and Design Structure

- one-way treatment structure, single fixed factor (t), four levels
- design structure; a single blocking factor (c).



Cell Means Model

$$y_{ij} = \mathbf{m} + \mathbf{r}_j + \mathbf{t}_i + \mathbf{e}_{ij} \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, b \end{cases}$$

Where

y_{ij} = the response for the i th tip type (t) in the j th block.

\mathbf{m} = the grand mean.

\mathbf{r}_j = the random effect of the j th block.

\mathbf{t}_i = the effect of the i th tip type (t).

\mathbf{e}_{ij} = the random error associated with the i th j th response.

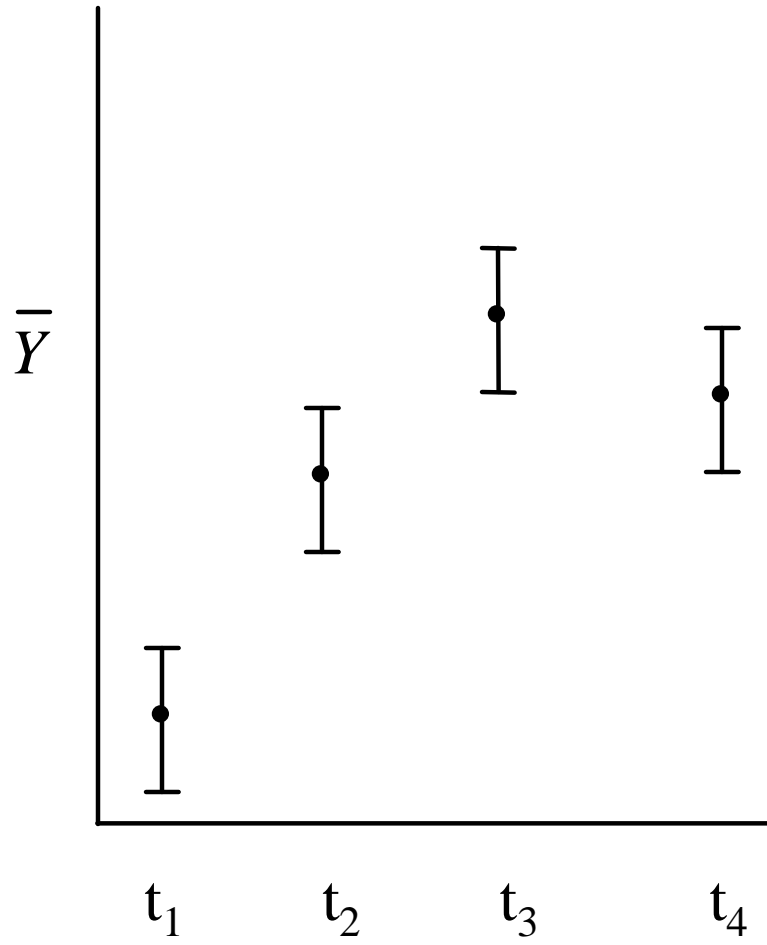
Assumptions

$$\sum_{i=1}^a \mathbf{t}_i = 0 \quad \mathbf{r}_j \sim N(0, \mathbf{S}_R) \quad \mathbf{e}_{ij} \sim N(0, \mathbf{S})$$

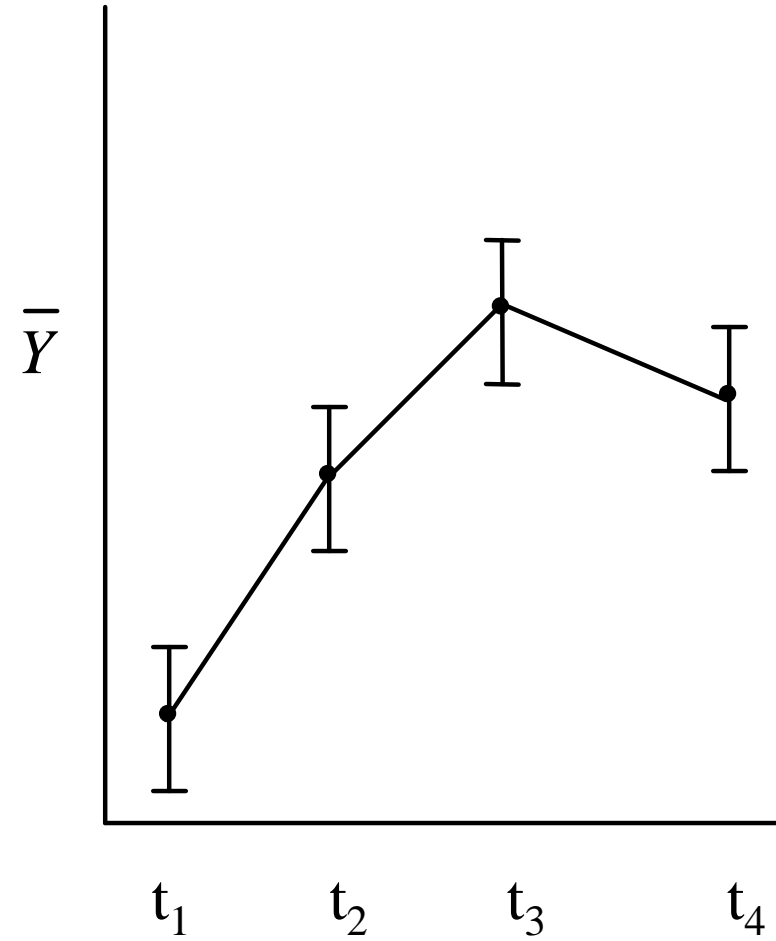
ANOVA Table

Sources	DF	DF Ex	MS	F-test
Block	b-1	3	MSB	-----
Tip	a-1	3	MSA	MSA/MSE
Error	(a-1)(b-1)	9	MSE	-----
Total	N-1	15		

Typical Analysis Summary



Qualitative



Quantitative

Latin Square Design

An engineer is studying the effects of five different formulations (f) of an explosive mixture used in the manufacture of dynamite on the observed explosive force. Each formulation is mixed from a batch of raw material that is only large enough for five formulations to be tested. Furthermore, the formulations are prepared by several operators, and there may be substantial differences in the skills and experience of the operators. The engineer wishes to minimize the variability in force that can be attributed to the batch to batch variability (b), as well as the operator to operator variability (o).

Example taken from “*Design and Analysis of Experiments*”.
Montgomery, 1991.

Treatment and Design Structure

- one-way treatment structure, single fixed factor (f), five levels.
- design structure; two blocking factors (b,o).

		Operator (o)				
		1	2	3	4	5
Batch (b)	1	f_1	f_2	f_3	f_4	f_5
	2	f_2	f_3	f_4	f_5	f_1
	3	f_3	f_4	f_5	f_1	f_2
	4	f_4	f_5	f_1	f_2	f_3
	5	f_5	f_1	f_2	f_3	f_4

Cell Means Model

$$y_{ijk} = \mathbf{m} + \mathbf{r}_i + \mathbf{g}_j + \mathbf{t}_k + \mathbf{e}_{ij} \quad \begin{cases} i = 1, 2, \dots, r \\ j = 1, 2, \dots, c \\ k = 1, 2, \dots, t \end{cases}$$

Where

y_{ijk} = the response for the k th formulation in the i th row, j th column.

\mathbf{m} = the grand mean.

\mathbf{r}_i = the effect of the i th batch.

\mathbf{g}_j = the effect of the j th operator.

\mathbf{t}_k = the effect of the k th formulation (f).

\mathbf{e}_{ij} = the random error associated with the i th, j th, k th response.

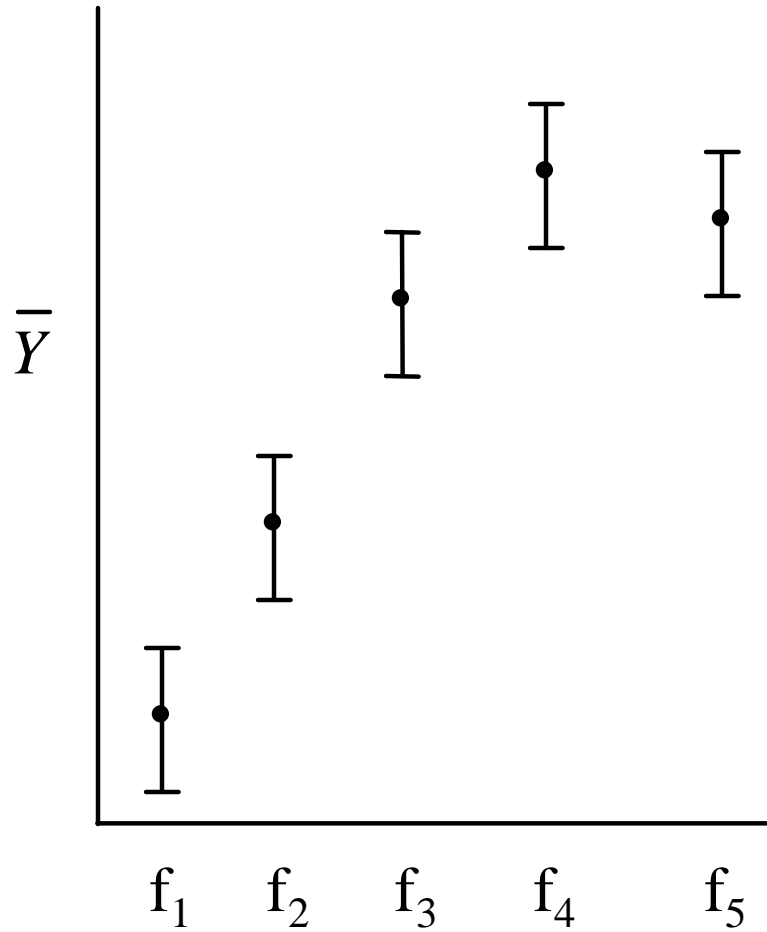
Assumptions

$$\sum_{k=1}^t \mathbf{t}_k = 0, \quad \mathbf{r}_i \sim N(0, \mathbf{s}_r), \quad \mathbf{g}_j \sim N(0, \mathbf{s}_g), \quad \mathbf{e}_{ij} \sim N(0, \mathbf{s})$$

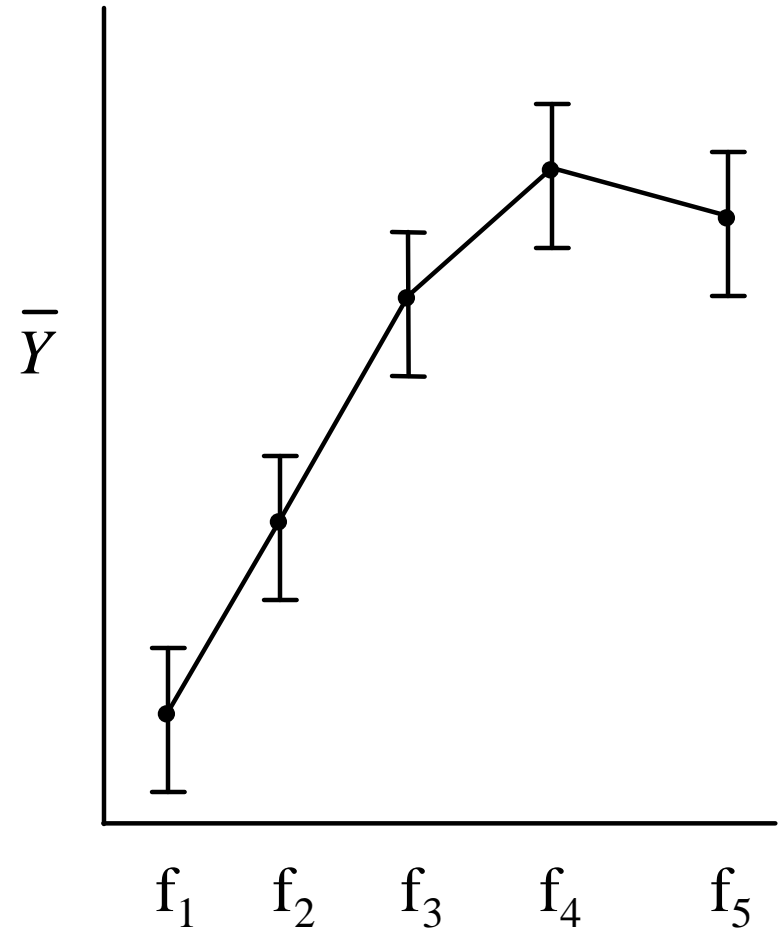
ANOVA Table

Source	DF	DF Ex	MS	F-test
Block(Batch)	$r-1$	4	MSR	-----
Block(Operator)	$c-1$	4	MSC	-----
Formulation (f)	$t-1$	4	MST	MST/MSE
Error	$(r - 1)(c-2)$	12	MSE	-----
Total	$(r^2 - 1)$	24		

Typical Analysis Summary



Qualitative



Quantitative

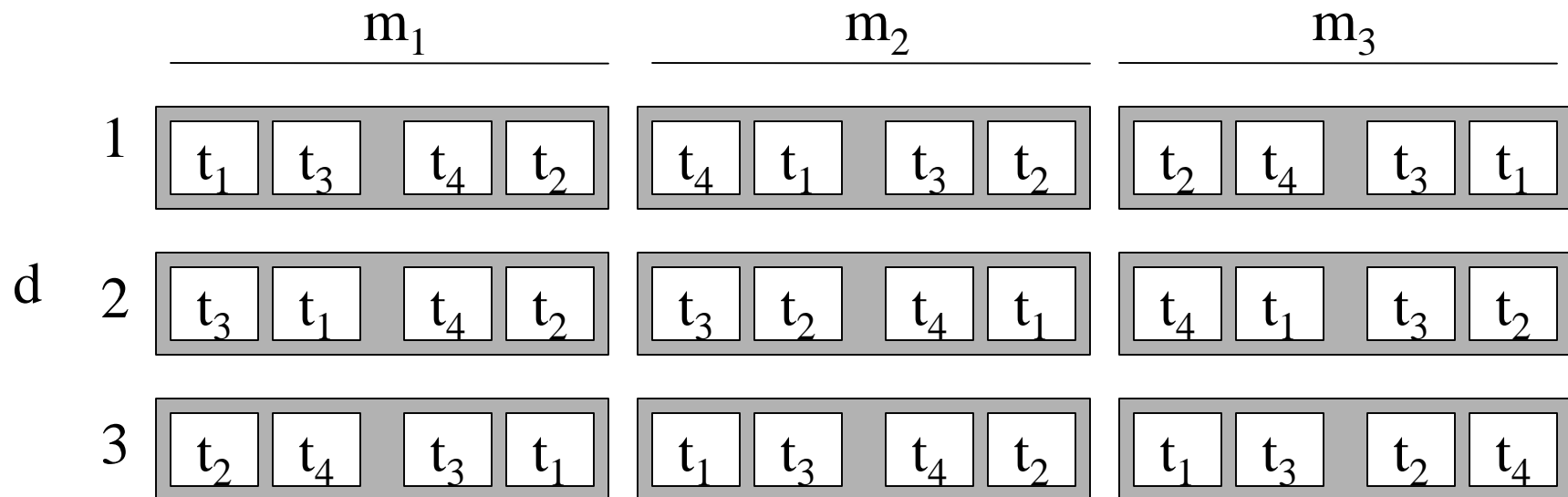
Split-Plot-Design

A paper manufacturer is interested in three different pulp preparation methods (m) and four cooking temperatures (t) for the pulp. The manufacturer wishes to study the effect of these two factors on the tensile strength of the paper. Each replication of a factorial requires 12 observations, and the experimenter has decided to run three replicates. However, the pilot plant is only capable of making 12 runs per day, so the experimenter decides to run one replicate on each of three days, with day considered as a blocking factor (d). On any day that the experiment is conducted a batch of pulp is produced by one of the three preparation methods. This batch is then divided into four samples and each sample is cooked at one of the four temperatures and the tensile strength of the ensuing product is determined.

Example taken from: “*Design and Analysis of Experiments*”.
Montgomery, 1991.

Design Structure

- two-way treatment structure, two fixed factors, with three levels of material (m) and four levels of temperature (t) randomized within the levels of material.
- design structure; single blocking factor (d).



Cell Means Model

$$y_{ijk} = \mathbf{m} + \mathbf{t}_i + \mathbf{r}_k + d_{ik} + \mathbf{l}_j + (\mathbf{tl})_{ij} + \mathbf{e}_{ijk} \quad \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, b \\ k = 1, 2, \dots, c \end{cases}$$

Where

y_{ijk} = the response for the j th level of temperature (t) within the i th level of material (m) in block k .

\mathbf{m} = the grand mean.

\mathbf{t}_i = the effect for the i th level of material (m).

\mathbf{r}_k = the random error associated with the k th block.

d_{ik} = random error associated with the i th level of material (m).

\mathbf{l}_j = the effect of the j th level of temperature (t).

$(\mathbf{tl})_{ij}$ = the interaction between i th level of m and j th level of t.

\mathbf{e}_{ijk} = random error associated with the i th, j th, k th response.

Assumptions

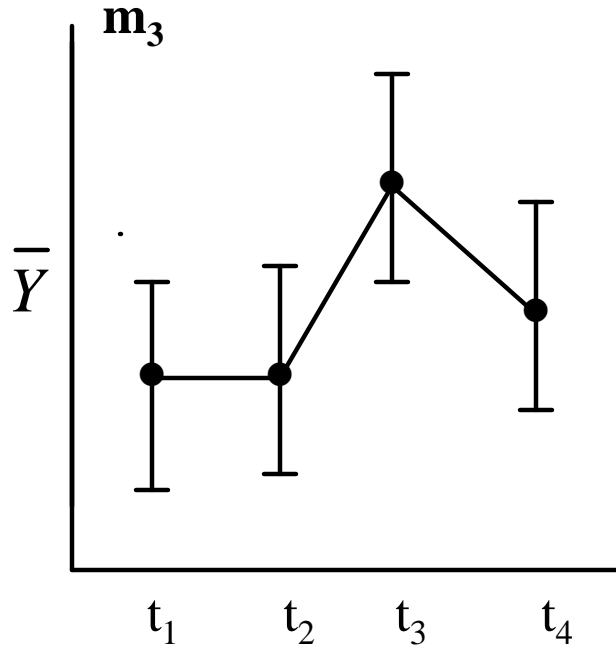
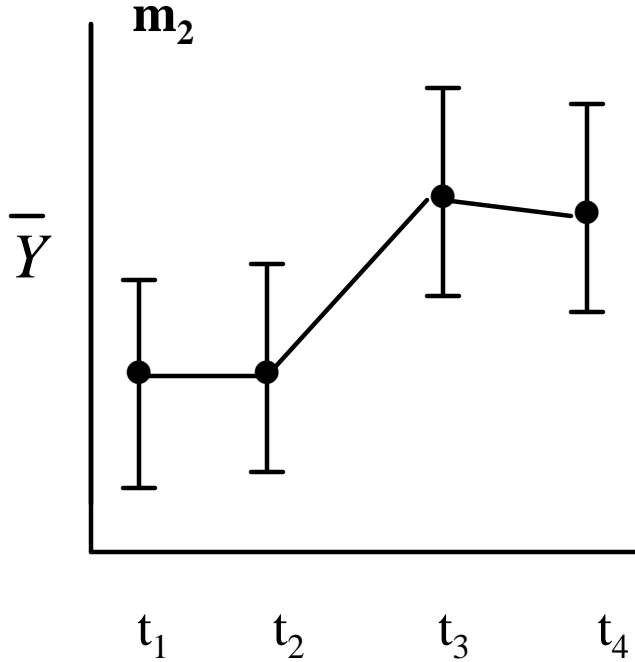
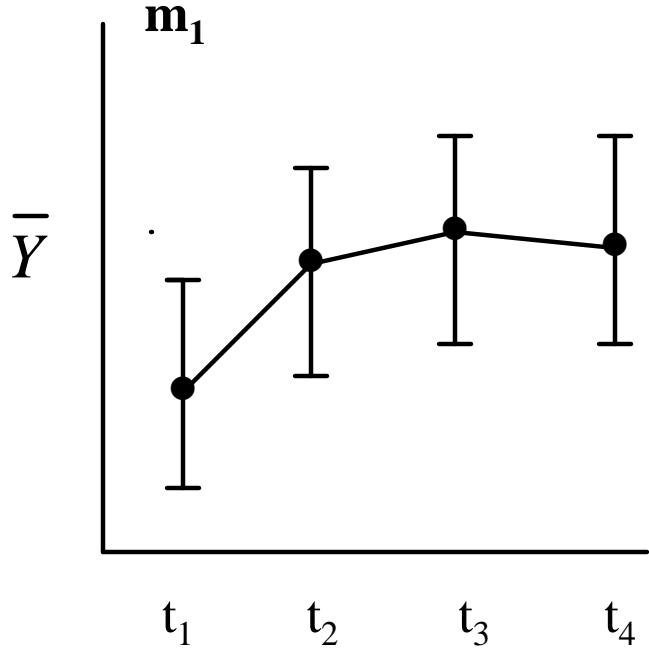
$$\sum_{i=1}^a t_i = 0, \quad \sum_{j=1}^b l_j = 0, \quad \sum_{i=1}^a \sum_{j=1}^b (tl)_{ij} = 0.$$

$$\mathbf{r}_k \sim N(0, \mathbf{S}_r), \quad d_{ik} \sim N(0, \mathbf{S}_d), \quad \mathbf{e}_{ij} \sim N(0, \mathbf{S}).$$

ANOVA Table

Sources	DF	DF Ex	MS	F-test
Block (d)	c-1	2	MSC	-----
Method (m)	a-1	2	MSA	MSA/MSE(a)
Error (a)	(a-1)(c-1)	4	MSE(a)	----
Temp (t)	b-1	3	MSB	MSB/MSE(b)
m x t	(a-1)(b-1)	6	MSAB	MSAB/MSE(b)
Error (b)	a(b-1)(c-1)	18	MSE(b)	----
Total	N-1	35		

Typical Analysis Summary



Sample Size Approximation (V)

The primary information required to estimate sample size:

- an estimate of variance σ^2 .
- the size of the difference between treatment means (δ) that has physical significant (critical difference).
- the significance level of the test (α), or the probability of a Type I error.
- the power of the test ($1-\beta$), or the probability of detecting δ , where β is the probability of a Type II error.

Sample Size Approximation

$$r \geq 2[Z_{\mathbf{a}/2} + Z_{\mathbf{b}}]^2 \left(\frac{\mathbf{s}}{\mathbf{d}} \right)^2$$

Where

r = the number of replication per treatment group.

$Z_{\mathbf{a}/2}$ = the standard normal variate exceeded with probability $\mathbf{a}/2$.

$Z_{\mathbf{b}}$ = the standard normal variate exceeded with probability \mathbf{b} .

\mathbf{s} = the estimate standard deviation of the response.

\mathbf{d} = the critical difference between the two treatment means

Multiple Comparison Adjustment (VI)

- Adjustment procedure to control the type I error rate for multiple hypotheses tests; where the type I error is defined as the probability of rejecting the null hypothesis when the null hypothesis is true.

Probability of at least k type I errors:

$$P(t \leq k) = \sum_{t=0}^k \frac{n!}{k!(n-k)!} \mathbf{a}_c^k (1 - \mathbf{a}_c)^{n-k}$$

$$P(0) = (1 - \mathbf{a}_c)^n$$

Where

n = the number of comparisons.

α_c = the comparison type I error rate.

Common Adjustments

- Dunnett's Method (comparison of all treatments to control, experimentwise error rate).
- MCB Method (single or combination of best treatments, experimentwise error rate).
- Fisher's Least Significant Difference Criterion (pair-wise comparisons, experimentwise error rate, least conservative).
- Tukey's Honest Significant Difference Criterion (all pair-wise comparisons, experimentwise error rate, moderately conservative).
- Bonferroni Criterion (pairwise comparisons, experimentwise error rate, most conservative).

Experimental Design Check List (VII)

- a) Define the objectives of the experiment.
- b) Identify all sources of variation, including:
 - treatment factors and their levels.
 - experimental units.
 - blocking factors, noise factors, and covariates.
- c) Choose appropriate rule for assigning the experimental units to the treatments.
- d) Specify the measurements to be made.
- e) Run a pilot study if possible*.
- f) Specify the model.
- g) Outline the analysis.
- h) Calculate the number of observations that need to be taken.

Reference Material (VIII)

- Barrentine, L.B. *An Introduction to Design of Experiments*. 1999. ASQ Quality Press Milwaukee, Wisconsin.
- Dean A., Voss D. *Design and Analysis of Experiments*. 1999. Springer. NY.
- Kuel R.O. *Design of Experiments: Statistical Principles of Research Design and Analysis. Second Edition*. 1996. Duxbury Press. Belmont, Cal.
- Montgomery, D.C. *Design and Analysis of Experiments: Fourth Edition*. 1997. John Wiley & Sons. NY.