STAT 850 Notes

by Langtian Ma & ChatGPT

Table of contents

1 Randomized Block Design

Problem Setting:

- One block factor with b levels. and one treatment factor with t levels.
- Treatments randomized with blocks.
- No replicate observations. Our observed data y_{ij} represents the data in ith block with jth factor.

	Treatment				
	A	B	C	D	Mean
Block 1	89	88	97	94	92
Block 2	84	77	92	79	83
Block 3	81	87	87	85	85
Block 4	87	92	89	84	88
Block 5	79	81	80	88	82
Mean	84	85	89	86	86

Table 1.

Model without block-treatment interactions:

$$y_{ij} = \mu + b_i + t_j + \epsilon_{ij} \quad i \in 1{:}\,B, j = 1{:}\,T$$

with $\epsilon_{ij} \overset{\text{iid}}{\sim} N(0, \sigma^2)$ and $\sum_j t_j = 0$. b_is and t_js indicate block and factor effects.

There are two settings for the block effects in the model:

- **Fixed block effects:** $\sum_{i=1}^{N} b_i = 0$ (sum to zero constraint for block effects)
- **Random block effects:** $b_1, \ldots, b_B \overset{\text{iid}}{\sim} N(0, \sigma_B^2)$

Data decomposition: the observed data can be decomposed to a summation of the (estimated) effects:

$$\begin{aligned} y_{ij} &= y_{\cdot\cdot} + (y_{i\cdot} - y_{\cdot\cdot}) + (y_{\cdot j} - y_{\cdot\cdot}) + (y_{ij} - y_{i\cdot} - y_{\cdot j} + y_{\cdot\cdot}) \\ &= \hat{\mu} + \hat{b}_i + \hat{t}_j + \hat{\epsilon}_{ij} \end{aligned}$$

ANOVA decomposition: the total sum of squares can be decomposed to:

$$\underbrace{\sum_i \sum_j (y_{ij} - y_{\cdot\cdot})^2}_{\text{Total sum of squares}} = \underbrace{t \sum_i (y_i - y_{\cdot\cdot})^2}_{\text{Blocks}} + \underbrace{b \sum_j (y_{\cdot j} - y_{\cdot\cdot})^2}_{\text{Treatments}} + \underbrace{\sum_i \sum_j (y_{ij} - y_{i\cdot} - y_{\cdot j} + y_{\cdot\cdot})^2}_{\text{Errors}}$$

Source	SS	df	$\mathbb{E}(\text{MS} = \text{SS/df})$
	Fixed block effects		
Blocks	$t \sum_i (y_{i\cdot} - y_{\cdot\cdot})^2$	$b-1$	$\sigma^2 + t\,(b-1)^{-1} \sum_i b_i^2$
Treatments	$b \sum_j (y_{\cdot j} - y_{\cdot\cdot})^2$	$t-1$	$\sigma^2 + b\,(t-1)^{-1} \sum_j t_j^2$
Error	$\sum_{ij} (y_{ij} - y_{i\cdot} - y_{\cdot j} + y_{\cdot\cdot})^2$	$(b-1)\,(t-1)$	σ^2
Total	$\sum_{ij} (y_{ij} - y_{\cdot\cdot})^2$	$b\,t-1$	
	Random block effects		
Blocks	$t \sum_i (y_{i\cdot} - y_{\cdot\cdot})^2$	$b-1$	$\sigma^2 + t\,\sigma_b^2$
Treatments	$b \sum_j (y_{\cdot j} - y_{\cdot\cdot})^2$	$t-1$	$\sigma^2 + b\,(t-1)^{-1} \sum_j t_j^2$
Error	$\sum_{ij} (y_{ij} - y_{i\cdot} - y_{\cdot j} + y_{\cdot\cdot})^2$	$(b-1)\,(t-1)$	σ^2
Total	$\sum_{ij} (y_{ij} - y_{\cdot\cdot})^2$	$b\,t-1$	

Table 2. ANOVA table for Randomized Block Design

2 Factorial Treatment Structure

2.1 Model Setting

Consider an experiment with two factors P and Q (***P* and *Q* may have interactions**) with levels $j = 1, 2, \ldots, p$, and $k = 1, 2, \ldots, q$, replicated r times ($l = 1, 2, \ldots, r$), with model

$$y_{jkl} = \mu_{jk} + \epsilon_{jkl}$$

Group effect parameters:

- Grand mean: $\mu_{\cdot\cdot} = (pq)^{-1} \sum_j \sum_k \mu_{jk}$
- Group means for factor P: $\mu_{j\cdot} = q^{-1} \sum_k \mu_{jk}$
- Group means for factor Q: $\mu_{\cdot k} = p^{-1} \sum_j \mu_{jk}$
- Effect of factor P: $p_j = \mu_{j\cdot} - \mu_{\cdot\cdot}$
- Effect of factor Q: $q_k = \mu_{\cdot k} - \mu_{\cdot\cdot}$

We have sum to zero constrains under this setting:

$$\sum_j p_j = 0, \quad \sum_k q_k = 0$$

Interaction effect parameters:

$$(pq)_{jk} = \mu_{jk} - (\mu_{\cdot\cdot} + p_j + q_k) = (\mu_{jk} - \mu_{\cdot k}) - (\mu_{j\cdot} - \mu_{\cdot\cdot})$$

Also, we have

$$\sum_j (pq)_{jk} = 0 \quad \text{for all } k, \quad \sum_k (pq)_{jk} = 0 \quad \text{for all } j$$

Then the model can be expand as:

$$\mu_{jk} = \mu_{\cdot\cdot} + p_j + q_k + (pq)_{jk}. \tag{1}$$

Remark 1. p and q without subscripts denote the number of levels for fator P and Q, while p_i and q_j denote the effect parameter for each level.

Example 1. (4×4 Design)

	μ_{jk}				$\mu_{j\cdot}$	p_j	$(pq)_{jk}$			
	4	10	20	30	16	-6	-3	-4	1	6
	14	20	40	50	31	9	-8	-9	6	11
	14	20	10	20	16	-6	7	6	-9	-4
	20	30	30	20	25	3	4	7	2	-13
$\mu_{\cdot k}$	13	20	25	30	22					
q_k	-9	-2	3	8						

Table 3. 4×4 design data table

2.2 Compare differences between treatments

To compare differences between treatments, we define **contrast** and **interaction contrast**:

Definition 1. *A contrast for the main effects of factor P is defined as*

$$C_P = \sum_{j=1}^{p} l_j \mu_{j\cdot},$$

where $l_1, \ldots, l_p$ are coefficients with $\sum_{j=1}^{p} l_j = 0$.

Example 2. (Simple Pairwise Comparison)

$$C_P = \mu_{1\cdot} - \mu_{2\cdot}$$

Definition 2. *An interaction contrast is defined as:*

$$C_{PQ} = \sum_{j=1}^{p} \sum_{k=1}^{q} l_j m_k \mu_{jk},$$

where $m_1, \ldots, m_q$ are also coefficients with $\sum_{k=1}^{q} m_k = 0$.

Example 3. Test whether the difference between levels of P depends on the level of Q.

$$C_{\mathrm{PQ}} = (\mu_{11} - \mu_{12}) - (\mu_{21} - \mu_{22})$$

Interpretation of main and interaction effects:

1. **Always start by checking main effects.** Interactions modify these effects and only make sense in that context.

2. **If interactions are negligible,** simplify the interpretation and focus on main effects**.**

3. **If 3 or higher order interactions are negligible**, but second-order interactions are significant, then we should focus on both main effects and second-order interactions.

4. If a two-factor interaction is **very important**, and its **mean square (MS) value is similar to the MS values for main effects**, then the best way to interpret results is by **looking at the mean values for two-factor combinations** rather than just reporting main effects.

5. If a **two-factor interaction is significant**, but **one or both main effects are much larger than the interaction**, then the interpretation should consider main effects first, with adjustments for interaction effects.

2.3 Least-squares estimation for an unreplicated 2×3 design

Consider a two-way factorial design with $p=2$ and $q=3$

$$\mu_{jk} = \mu_{..} + p_j + q_k + (p\,q)_{jk}$$

	1	2	3
1	y_{11}	y_{12}	y_{13}
2	y_{21}	y_{22}	y_{23}

Table 4.

$$\mathbf{y}=\begin{pmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{pmatrix} = \mathbf{X}\beta = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 0 & -1 & 0 \\ 1 & -1 & 0 & 1 & 0 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \mu.. \\ p_1 \\ q_1 \\ q_2 \\ (p\,q)_{11} \\ (p\,q)_{12} \end{pmatrix} + \epsilon$$

The columns of $\mathbf{X}$ w.r.t. different parameter groups are orthogonal. In this case, the columns with respect to p and columns with respect to q are orthogonal. Also, they are orthogonal to the columns w.r.t. pq. Therefore, $\mathbf{X}^T\mathbf{X}$ appears to be block diagonal:

$$\mathbf{X}'\mathbf{X}=\begin{pmatrix} 6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 & 0 \\ 0 & 0 & 2 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 0 & 2 & 4 \end{pmatrix}$$

$$(\mathbf{X}'\mathbf{X})^{-1}=\begin{pmatrix} 1/6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & -1/6 & 0 & 0 \\ 0 & 0 & -1/6 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & -1/6 \\ 0 & 0 & 0 & 0 & -1/6 & 1/3 \end{pmatrix}$$

The least square solution gives

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}=\begin{pmatrix} y.. \\ y_{1\cdot} - y.. \\ y_{\cdot 1} - y.. \\ y_{\cdot 2} - y.. \\ y_{11} - y_{1\cdot} - y_{\cdot 1} + y.. \\ y_{12} - y_{1\cdot} - y_{\cdot 2} + y.. \end{pmatrix}.$$

Note that the number of parameter equals to the number of observations, $\mathbf{X}$ is invertible, $\hat{\beta}$ is the solution to $\mathbf{X}\beta = \mathbf{y}$.

Now we consider an additive model without interactions:

$$\mu_{jk} = \mu_{..} + p_j + q_k,$$

the least square solution gives:

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}\mathbf{y} = (\ y_{..}\ \ y_{1\cdot} - y_{..}\ \ y_{\cdot 1} - y_{..}\ \ y_{\cdot 2} - y_{..}\ )^T,$$

which align with the solution of the model with interactions. This is due to the orthogonality.

2.4 Experiment with Replication

The model for an experiment with replication can be written as:

$$y_{jkl} = \mu + p_j + q_k + (pq)_{jk} + \epsilon_{jkl}, j = 1, \ldots, p; k = 1, \ldots q, l = 1, \ldots r,$$

where ϵ_{jkl} are i.i.d. $\mathcal{N}(0, \sigma^2)$. We can decompose data as:

$$\begin{aligned} y_{jkl} &= \hat{\mu} + \hat{p}_j + \hat{q}_k + \widehat{(p\,q)}_{jk} + \hat{\epsilon}_{jkl} \\ &= y_{...} + (y_{j\cdot\cdot} - y_{...}) + (y_{\cdot k\cdot} - y_{...}) + (y_{jk\cdot} - y_{j\cdot\cdot} - y_{\cdot k\cdot} + y_{...}) + (y_{jkl} - y_{jk\cdot}). \end{aligned}$$

The sum of squares are defined as:

$$\begin{aligned} S_P &= q\,r \sum_j (y_{j\cdot\cdot} - y_{...})^2 \\ S_Q &= p\,r \sum_k (y_{\cdot k\cdot} - y_{...})^2 \\ S_{PQ} &= r \sum_j \sum_k (y_{jk\cdot} - y_{j\cdot\cdot} - y_{\cdot k\cdot} + y_{...})^2 \\ S_R &= \sum_j \sum_k \sum_l (y_{jkl} - y_{jk\cdot})^2 \\ S_D &= \sum_j \sum_k \sum_l (y_{jkl} - y_{...})^2 \end{aligned}$$

Source	SS	df	MS	Ratio
Factor P	$S_P = 1.03301$	$p - 1 = 2$	$s_P^2 = 0.51651$	$s_P^2 / s_R^2 = 23.22$
Factor Q	$S_B = 0.92121$	$q - 1 = 3$	$s_Q^2 = 0.30707$	$s_Q^2 / s_R^2 = 13.81$
Interaction	$S_{PQ} = 0.25014$	$(p-1)(q-1) = 6$	$s_{PQ}^2 = 0.04169$	$s_{PQ}^2 / s_R^2 = 1.87$
Residual	$S_R = 0.80073$	$p\,q\,(r-1) = 36$	$s_R^2 = 0.02224$	
Total	$S_D = 3.00508$	$p\,q\,r - 1 = 47$		

Table 5. ANOVA table for two factors experiments with replications

2.5 Model Checking

1. Define the estimated value of μ_{jk} under the full model as $\hat{\mu}_{\text{jk}} = y_{\text{jk}\cdot}$. Let $\tilde{\mu}_{jk} = y_{j\cdot\cdot} + y_{\cdot k\cdot} - y_{...}$ denote the estimated value of μ_{jk} assuming no interactions.

2. To assess the homogeneity of variance in interactions, create a plot of the residuals $y_{jkl} - \hat{\mu}_{\text{jk}}$ against the fitted values $\hat{\mu}_{\text{jk}}$. A consistent spread of residuals across different values of $\hat{\mu}_{\text{jk}}$ suggests homogeneity, while a pattern or funnel shape may indicate variance issues.

3. To detect possible nonadditivity, plot $y_{jk} - \tilde{\mu}_{jk}$ against $\tilde{\mu}_{jk}$. If the plot exhibits a curvilinear pattern, this suggests the presence of transformable nonadditivity, meaning that a transformation of the response variable may be necessary for a better model fit.

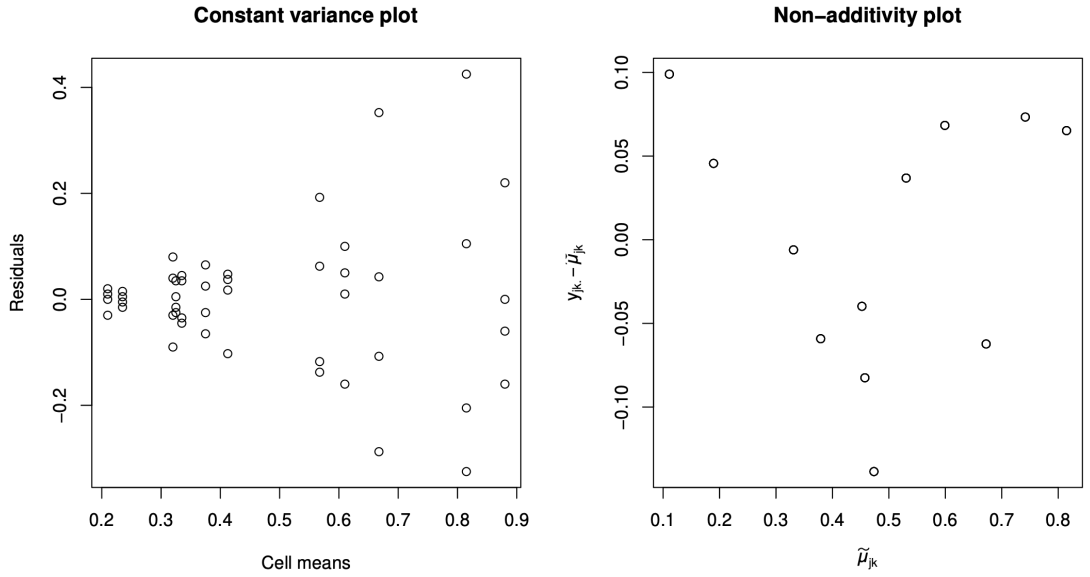

Figure 1. Model checking plots suggesting heteroscedasticity and non-additivity

2.6 Transformations

2.6.1 Taylor Power Transformation

Consider heterogeneous data $\text{Var}(y_{jkl}) = \sigma_{jk}^2$ and assume that $\sigma_{jk}^2 \propto \mu_{jk}^{\beta}$ for some β. We can use **Taylor power transformation** to deal with the heteroscedasticity (Assume $y_{jk\cdot} > 0$).

1. Calculate the sample mean $y_{jk\cdot}$ and sample standard deviation s_{jk} for each (j,k) cell.
2. Fit a linear regression on $\log s_{jk} \sim \log y_{jk}$.
3. The fitted slope $\hat{\beta}$ is an estimate of β.
4. Use the transformation $(y^{\lambda} - 1)/\lambda$ with $\lambda := 1 - \hat{\beta}$.
5. If $\lambda = 0$, apply the log transformation $y_{jkl}^{*} = \log(y_{jkl})$.

Remark 2. Simple power transformation uses y^{λ}, which doesn't smoothly transit to $\log y$ as $\lambda \to 0$.

Justification for the method:

Define

$$z = f_{\lambda}(y) = \begin{cases} (y^{\lambda} - 1)/\lambda, & \lambda \neq 0 \\ \log y, & \lambda = 0 \end{cases} \tag{2}$$

By Taylor expansion:

$$z_{jkl} = f_{\lambda}(y_{jkl}) \approx f_{\lambda}(\mu_{jk}) + f_{\lambda}'(\mu_{jk})\,(y_{jkl} - \mu_{jk}).$$

Since $f'(y) = y^{\lambda - 1}$,

$$\begin{aligned} \text{Var}(z_{jkl}) \approx & (f_{\lambda}'(\mu_{jk}))^2 \text{Var}(y_{jkl}) \\ = & \mu_{jk}^{2(\lambda-1)}\, \sigma_{jk}^2 \\ \propto & \mu_{jk}^{2(\lambda-1)}\, \mu_{jk}^{2\beta} \\ = & \mu_{jk}^{2(\lambda-1+\beta)}, \end{aligned}$$

and $\text{Var}(z_{jkl})$ becomes a constant if $\lambda = 1 - \beta$.

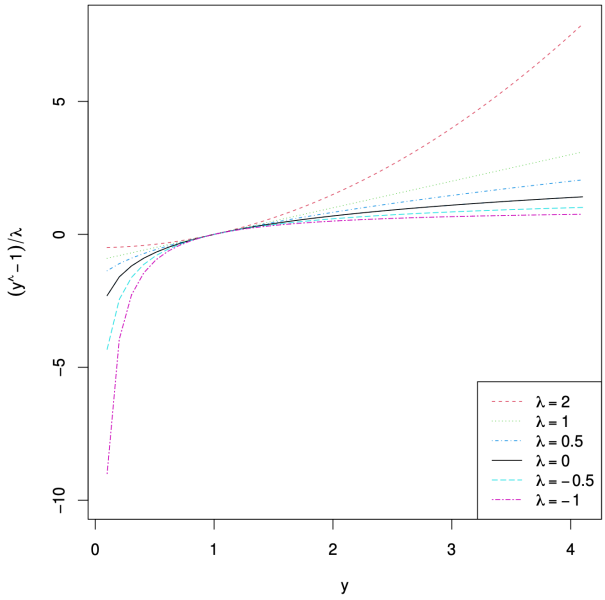

Figure 2. Power transformations

2.6.2 Box-Cox Transformation

Assumptions: There exists a λ such that $\{f_\lambda(y_i)\}_{i=1}^n$:

- are mutually independent
- are normally distributed
- have constant variance
- satisfy a linear model $f_\lambda(\mathbf{y}) = \mathbf{X}\beta + \boldsymbol{\epsilon}$

2.6.3 Maximum Likelihood Estimation of Box-Cox $\boldsymbol{\lambda}$

Assume $\mathbf{y}^{(\lambda)} = f_\lambda(\mathbf{y}) = \mathbf{X}\beta + \boldsymbol{\epsilon}$ where f_λ is defined in (2) with $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2\mathbf{I})$ for some λ. Then the likelihood function for the untransformed data follows:

$$L(\lambda, \beta, \sigma) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{(\mathbf{y}^{(\lambda)} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y}^{(\lambda)} - \mathbf{X}\boldsymbol{\beta})}{2\sigma^2}\right\} J(\lambda, \mathbf{y})$$

with Jacobian

$$J(\lambda, y) = \prod_{i=1}^n y_i^{\lambda-1}$$

We first find the LSEs of β and σ_λ for fixed λ:

$$\hat{\beta}_\lambda = (X^T X)^{-1} X^T \mathbf{y}^{(\lambda)}, \quad \hat{\sigma}_\lambda = \frac{(\mathbf{y}^{(\lambda)} - \mathbf{X}\hat{\boldsymbol{\beta}}_\lambda)^T (\mathbf{y}^{(\lambda)} - \mathbf{X}\hat{\boldsymbol{\beta}}_\lambda)}{n - p - 1},$$

then we have

$$\begin{aligned}\ell(\lambda, \hat{\boldsymbol{\beta}}_\lambda, \hat{\sigma}_\lambda) &= \frac{\exp\left(-(n-p-1)/2\right)}{(2\pi)^{n/2}\hat{\sigma}_\lambda^n} J(\lambda, \mathbf{y}) \\ &= \hat{\sigma}_\lambda^{-n} \prod_{i=1}^n y_i^{\lambda-1} \frac{\exp\left(-(n-p-1)/2\right)}{(2\pi)^{n/2}}\end{aligned}$$

Let $\ell(\lambda) = \log L(\lambda, \hat{\beta}_\lambda, \hat{\sigma}_\lambda)$, and find the MLE $\hat{\lambda}$ that maximizes $\ell(\lambda)$.

The **confidence interval** follows from the standard result that the log-likelihood ratio statistic follows a chi-square distribution with 1 degree of freedom

$$2(\ell(\hat{\lambda}) - \ell(\lambda)) \sim \chi^2_1.$$

Then the confidence interval is the root for $\ell(\lambda) = \ell(\hat{\lambda}) - 0.5\chi^2_{1,\alpha}$.

Remark 3. The `boxcox` function in `MASS` library gives the MLE of λ.

```
library(MASS)
bc <- boxcox(y ~ p+q)
title(paste("Without interaction, lambda =",round(bc$x[which.max(bc$y)],2)))
bc <- boxcox(y ~ p*q)
title(paste("With interaction, lambda =",round(bc$x[which.max(bc$y)],2)))
```

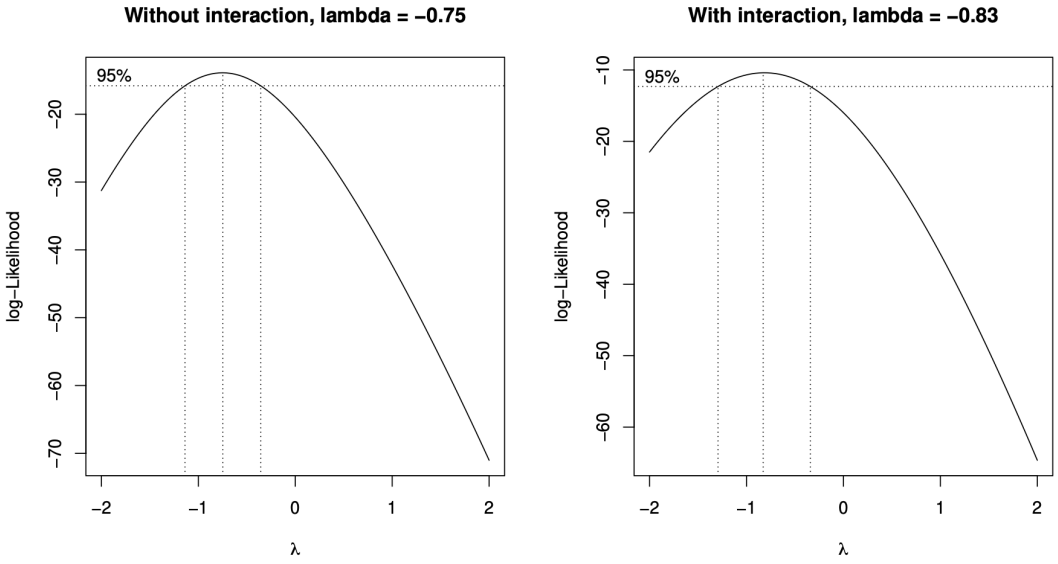

Figure 3. Likelihood of Box-Cox Transformation

Example 4. For the Poison data, we present the ANOVA table after the two transformations.

	Df	Sum Sq	Mean Sq	F value	P-value
Poison	2	34.877	17.4386	70.6302	5.17e−13
Treatment	3	20.414	6.8048	27.5610	2.48e−09
Interaction	6	1.571	0.2618	1.0603	0.4046
Residuals	35	8.643	0.2469		
Poison	2	11.926	5.9631	66.5525	1.18e−12
Treatment	3	7.158	2.3860	26.6295	3.76e−09
Interaction	6	0.486	0.0810	0.9040	0.5032
Residuals	35	3.136	0.0896		

Table 6. ANOVA table after simple Taylor power transformation ($y^{(\lambda)} = y^\lambda$ with $\lambda = -1$) (above) and Box-Cox transformation (below).

Remark 4. The residual Df is reduced by 1 to compensate for the estimation of λ.

2.7 Confidence Intervals

2.7.1 When interactions are not significant

Let $u_j = \mu + p_j$ and $v_k = \mu + q_k$. Let the unbiased estimators be $\hat{u}_j = y_{j\cdot\cdot}$ and $\hat{v}_k = y_{\cdot k\cdot}$ with $\text{Var}(\hat{u}_j) = \sigma^2/(pr)$, $\text{Var}(\hat{v}_k) = \sigma^2/(qr)$. The sample standard deviation is

$$s(\hat{u}_j) = S_R/\sqrt{pr}, s(\hat{v}_k) = S_R/\sqrt{qr}.$$

Then $100(1-\alpha)\%$ confidence interval for u_j is $\hat{u}_j \pm t_{\nu_R;\alpha/2} \times \text{se}(\hat{u}_j)$, where $\nu_R = pq(r-1)$ is the degree of freedom for S_R.

For simultaneous confidence intervals, we define the contrast of interest be

$$L = \sum_{j=1}^{p} c_j u_j, \quad \text{where } \sum_{j=1}^{p} c_j = 0$$

with its estimator

$$\hat{L} = \sum_{j=1}^{p} c_j y_{j\cdot\cdot} \quad \text{with} \quad s(\hat{L}) = S_R \sqrt{(qr)^{-1} \sum_{j=1}^{p} c_j^2}$$

Then a $100(1-\alpha)\%$ simultaneous confidence interval for L takes the form:

$$\hat{L} \pm T s(\hat{L}),$$

where T is a multiplier that depends on the type of the inference method used.

1. **Tukey's Method (Pairwise Comparisons):** Tukey's method is designed for simultaneous confidence intervals when comparing all possible pairwise differences between group means. The multiplier is:

   $$T = \frac{q(p, \nu_R; \alpha)}{\sqrt{2}}$$

   where $q(p, \nu_R; \alpha)$ is the studentized range statistic for p groups and residual degrees of freedom ν_R.

   When comparing all pairs, each pairwise difference is a contrast (with coefficients $c_j = 1$ for one group and $c_j = -1$ for the other, and 0 elsewhere). Each interval is given by:

   $$\text{CI for } (u_j - u_k)\text{: } A\,C \quad (y_{j\cdot\cdot} - y_k) \pm \frac{q\,(p, \nu_R; \alpha)}{\sqrt{2}}\, s_R \sqrt{\frac{1^2 + (-1)^2}{q\,r}}$$

   Here, the multiplier $\frac{q\,(p, \nu_R; \alpha)}{\sqrt{2}}$ is applied to each pairwise contrast.

2. **Scheffé's Method (All Contrasts):** Scheffés method is more conservative and applies to all possible contrasts, not just pairwise comparisons. The multiplier is:

   $$T = \sqrt{(p-1)\, F_{(p-1),\nu_R;\alpha}}$$

   where $F_{(p-1),\nu_R;\alpha}$ is the critical value from the F-distribution with $p-1$ and ν_R degrees of freedom.

For any contrast $L = \sum_j c_j u_j$, the simultaneous confidence interval is:

$$\hat{L} \pm \sqrt{(p-1)\, F_{(p-1),\nu_R;\alpha}}\, S_R \sqrt{\frac{1}{q\, r} \sum_j c_j^2}$$

This interval applies to every possible contrast you might form.

3. **Bonferroni's Method (For g Comparisons):** The Bonferroni method controls the familywise error rate by adjusting the significance level for multiple comparisons. For any set of g comparisons, the multiplier is:

$$T = t_{\nu_R;\alpha/(2\,g)}$$

where $t_{\nu_R;\alpha/(2g)}$ is the t-distribution critical value with residual degrees of freedom ν_R and a Bonferroni-adjusted significance level of $\alpha\,/\,(2\,g)$.

If you have a specific set of g comparisons (contrasts) you plan to test, each interval is:

$$\hat{L} \pm t_{\nu_R;\alpha/(2\,g)}\, S_R \sqrt{\frac{1}{q\, r} \sum_j c_j^2}.$$

Each of the g contrasts gets its own interval, with the critical value adjusted by dividing α by $2\,g$.

2.7.2 When interactions are significant

In this case, each combination of factor levels has its own mean:

$$\mu_{jk} = \mu + p_j + q_k + (pq)_{jk}$$

If we are interested in comparing the means of two specific treatment combinations, say $\mu_{j_1k_1}$ and $\mu_{j_2k_2}$, we are comparing two of the pq treatments.

1. **Tukey's Method:** for all possible pairs of treatment means,

$$(\hat{\mu}_{j_1k_1} - \hat{\mu}_{j_2k_2}) \pm \frac{q(p\,q, \nu_R; \alpha)}{\sqrt{2}}\, S_R \sqrt{\frac{2}{r}}$$

where $q(pq, \nu_R; \alpha)$ is the quantile of the studentized range statistic for pq treatments and ν_R is the residual degrees of freedom.

2. **Bonferroni's Method:** for g pairs of comparisons,

$$(\hat{\mu}_{j_1k_1} - \hat{\mu}_{j_2k_2}) \pm t_{\nu_R;\alpha/(2g)}\, s_R \sqrt{\frac{2}{r}},$$

where $t_{\nu_R;\alpha/(2g)}$ is the quantile from t-distribution.

3. **Scheffé's Method for General Contrasts**: the contrast takes the form:

$$L = \sum_{j=1}^{p} \sum_{k=1}^{q} c_{jk}\, \mu_{jk}, \quad \text{with} \quad \sum_{j,k} c_{jk} = 0$$

The confidence interval for the contrast L is given by:

$$\hat{L} \pm \sqrt{(p\,q-1)\, F_{(pq-1),\nu_R;\alpha}}\, s(\hat{L})$$

where:

- $\hat{L} = \sum_{j,k} c_{jk}\hat{\mu}_{jk}$,
- $s(\hat{L}) = s_R\sqrt{\sum_{j,k}\frac{c_{jk}^2}{r}}$ (assuming balanced replication),
- $F_{(pq-1),\nu_R;\alpha}$ is the critical value from the F-distribution with $pq-1$ and ν_R degrees of freedom.

2.8 Two-way Factorial with Blocks

2.8.1 No Replicates Within Blocks

Model:

$$y_{ijk} = \mu + b_i + t_{jk} + \varepsilon_{ijk} = \mu + b_i + p_j + q_k + (pq)_{jk} + \varepsilon_{ijk}$$

for $i = 1, \ldots, b$, $j = 1, \ldots, p$, and $k = 1, \ldots, q$, with sum-to-zero constrains:

$$\sum_i b_i = \sum_j p_j = \sum_k q_k = 0.$$

Data decomposition:

$$\begin{aligned} y_{ijk} =& y_{...} + (y_{i..} - y_{...}) + (y_{.jk} - y_{...}) + (y_{ijk} - y_{i..} - y_{.jk} + y_{...}) \\ =& y_{...} + (y_{i..} - y_{...}) + (y_{.j.} - y) + (y_{..k} - y_{...}) \\ &+ (y_{.jk} - y_{.j.} - y_{..k} + y_{...}) + (y_{ijk} - y_{i..} - y_{.jk} + y_{...}) \end{aligned}$$

Source	SS	df
Blocks	$p\,q\sum_i (y_{i..} - y_{...})^2$	$b-1$
P	$b\,q\sum_j (y_{.j.} - y_{...})^2$	$p-1$
Q	$b\,p\sum_k (y_{..k} - y_{...})^2$	$q-1$
PQ	$b\sum_j\sum_k (y_{.jk} - y_{.j.} - y_{..k} + y_{...})^2$	$(p-1)\,(q-1)$
Residual	$\sum_i\sum_j\sum_k (y_{ijk} - y_{i..} - y_{.jk} + y_{...})^2$	$(b-1)\,(p\,q-1)$
Total	$\sum_i\sum_j\sum_k (y_{ijk} - y_{...})^2$	$b\,p\,q-1$

Table 7. ANOVA table for two-way factorial with blocks and without replicates

2.8.2 With Replicates Within Blocks

Model:

$$y_{ijkl} = \mu + b_i + t_{jk} + \varepsilon_{ijk} = \mu + b_i + p_j + q_k + (pq)_{jk} + \varepsilon_{ijkl}$$

or $i = 1, \ldots, b$, $j = 1, \ldots, p$, $l = 1, \ldots, r$, and $k = 1, \ldots, q$, with sum-to-zero constrains:

$$\sum_i b_i = \sum_j p_j = \sum_k q_k = 0.$$

Data decomposition:

$$\begin{aligned} y_{ijkl} =& y_{....} + (y_{i...} - y_{....}) + (y_{.j..} - y_{....}) + (y_{..k.} - y_{....}) \\ &+ (y_{.jk.} - y_{.j..} - y_{...k} + y_{....}) + (y_{ijkl} - y_{i...} - y_{.jk.} + y_{....}) \end{aligned}$$

Source	SS	df
Blocks	$p\,q\,r\sum_i (y_{i\cdots} - y_{\cdots\cdot})^2$	$b-1$
P	$b\,q\,r\sum_j (y_{\cdot j\cdot\cdot} - y_{\cdots\cdot})^2$	$p-1$
Q	$b\,p\,r\sum_k (y_{\cdot\cdot k\cdot} - y_{\cdots\cdot})^2$	$q-1$
PQ	$b\sum_j\sum_k (y_{\cdot jk\cdot} - y_{\cdot j\cdot\cdot} - y_{\cdot\cdot k\cdot} + y_{\cdots\cdot})^2$	$(p-1)\,(q-1)$
Residual	$\sum_i\sum_j\sum_k\sum_l (y_{ijkl} - y_{i\cdots} - y_{\cdot jk\cdot} + y_{\cdots\cdot})^2$	$bpqr - b - pq + 1$
Total	$\sum_i\sum_j\sum_k\sum_l (y_{ijkl} - y_{\cdots\cdot})^2$	$b\,p\,q\,r - 1$

Table 8.

2.8.3 Advantages of Two-level Factorial Designs

1. **Efficiency in runs:** They require only a few experimental runs per factor, making them cost-effective and time-saving even when several factors are under study.

   Two-level factorial designs use only two settings per factor, which are chosen to maximize the contrast between high and low values. This strong contrast helps in estimating the main effects with greater clarity.

2. **Identifying Major Trends:** While they don't cover the entire range of possible factor values, these designs help reveal the main effects and trends, pointing toward areas that merit further investigation.

3. **Flexible Augmentation:** If more detailed local analysis is needed later, two-level designs can be expanded or combined with additional points (composite designs) to explore the factor space more thoroughly.

4. **Modular Design Building:** They serve as building blocks. By combining them appropriately, you can tailor the complexity of the experimental design to match the sophistication of the problem at hand.

5. **Precision in Effect Estimation:** With fewer runs and a clear structure, the estimates of the main effects of factors are generally precise.

6. They form the basis for two-level fractional factorial designs.

3 Factorial Design

3.1 Comparison of designs

3.1.1 OFAT v.s. Factorial

OFAT Experiments: An experimental approach in which only **one factor (independent variable)** is varied at a time while keeping all other factors constant. The goal is to observe the effect of that single factor on the response variable.

2^2 Factorial Design: an experimental design where there are **two factors**, each with **two levels** (e.g., "low" and "high"). This results in **$2^2 = 4$ experimental runs**, covering all possible combinations of factor levels.

Example 5. OFAT Design: An engineer designed an experiment to compare a standard and a new gas anneal process by varying pressure and temperature across three runs:

1. Standard pressure & standard temperature

2. Standard pressure & new temperature

3. New pressure & new temperature

A single lot of 48 wafers was used, with 16 wafers assigned to each run.

	Temperature	
Pressure	Standard	New
Standard	16 wafers	16 wafers
New	-	16 wafers

Table 9. OFAT Design

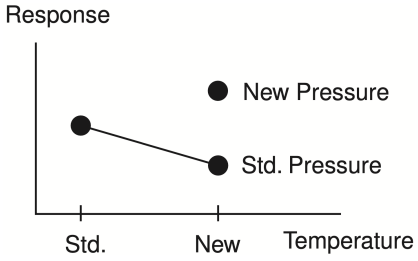

Figure 4. OFAT Design

2^2 Factorial Design:

	Temperature	
Pressure	Standard	New
Standard	12 wafers	12 wafers
New	12 wafers	12 wafers

Table 10. 2^2 Factorial Design

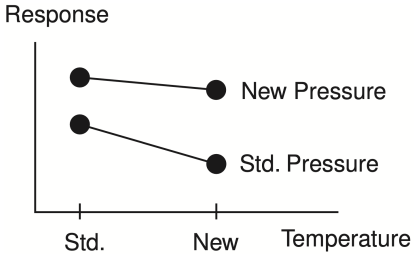

Figure 5. 2^2 Factorial Design

Advantages of 2^2 Factorial Design:

1. Estimates of the factor effects are more precise (all 48 wafers are used, while only 32 wafers in the OFAT design)

   a. Variance of each effect is $\sigma^2/12$ for the 2^2 design.

   b. Variance of each factor effect for the OFAT is $\sigma^2/8$(50% more).

2. Interaction effect can be estimated.

3. 2^2 design provides data over a broader factor space.

3.1.2 Central Composite Design

Example 6. Two engineers planned an experiment for a rapid thermal anneal process. They wanted to study the sensitivity of the response sheet resistance to two factors—time and temperature. The followings are 3 different designs

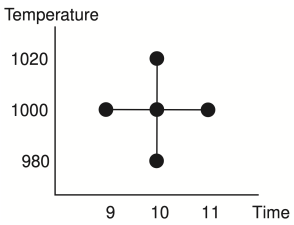

Figure 6. OFAT Design

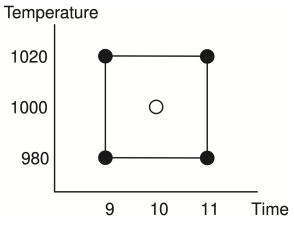

Figure 7. 2^2 Design

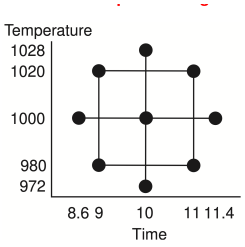

Figure 8. Central Composite

Central Composite Design (CCD): is a type of **Response Surface Methodology (RSM)** used for building a quadratic model in experimental designs. It helps **optimize processes** by exploring both linear and interaction effects, as well as curvature in factor responses. It contains the following components:

- **Factorial Points**: A full or fractional factorial design that represents the main effects and interactions.
- **Axial (Star) Points**: Points placed outside the factorial design to capture curvature and quadratic effects.
- **Center Points**: Replicated runs at the center of the design space to improve model accuracy and detect pure error.

Type	X_1	X_2
Factorial Points	$-1, +1$	$-1, +1$
Axial Points	$\pm\alpha$	$\pm\alpha$
Center Points	0	0

Table 11. Example of CCD with two factors

Advantages of Central Composite Design:

1. Estimates of factor effects are more precise
2. Interactions can be estimated
3. Curvature can be estimated in the entire space
4. Optimization is possible for the entire space
5. If center point is replicated 4 times, the design is rotatable, i.e., equal precision in estimation at all points equidistant from the center

3.1.3 2^3 Design

Example 7. 3-factor OFAT design & 2^3 experiment

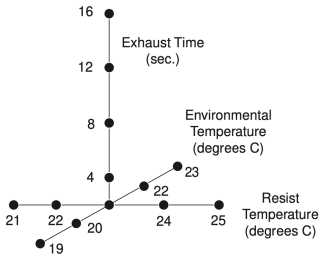

Figure 9. 3-factor OFAT in 15 runs

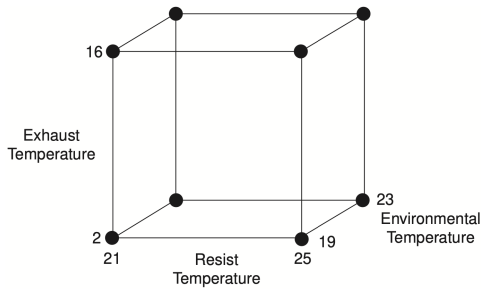

Figure 10. 2^3 experiment in 8 runs

Advantages of the 2^3 design:

1. Requires less resources: 8 runs instead of 15.

2. Estimates of factor effects are more accurate: all 8 runs are used in each estimate. But in OFAT, only 5 runs are used in each estimate.

3. Interaction effects can be estimated.

4. 2^3 design yields information over a larger region of the factor space

3.1.4 Box-Behnken design

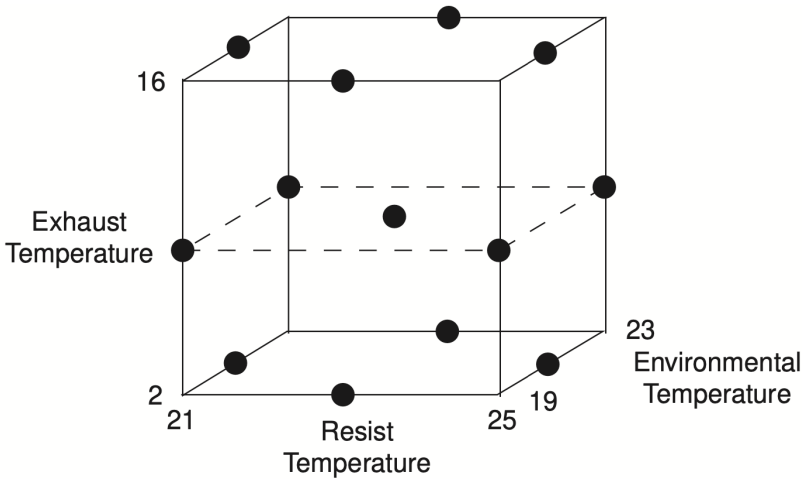

Figure 11. Box-Behnken design in 15 runs

Advantages of Box-Behnken Design:

1. Interactions can be estimated in Box-Behnken design.

2. Runs are more evenly spread out—hence more accurate predictions while OFAT can estimate curvature along 3 lines only.

3. Box-Behnken can estimate curvature in entire experimental region.

4. Box-Behnken design can be optimized over entire factor space while OFAT can be optimized along 3 lines only.

3.2 Analysis of Factorial Design

3.2.1 2^2 Design

Notation 1. *We use the following notations to represent factors and interactions in a* 2^2 *design:*

- *Two factors* A *and* B*, each at two levels.*
- *4 sets of experimental conditions. "Low" indicated by "−" and "high" indicated by "+".*
- *Use notation* (1)*,* a*,* b*,* ab *to denote the mean responses at the treatment combinations listed in figure 12.*

Run	A	B	Symbol
1	-	-	(1)
2	+	-	a
3	-	+	b
4	+	+	ab

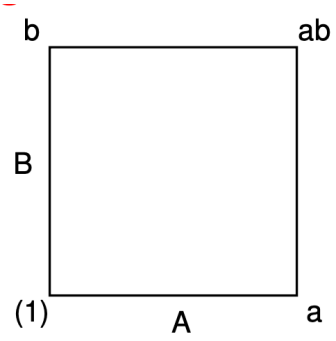

Figure 12. Notations in a 2^2 design

Model:

$$y_{ijk} = \eta + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \tag{3}$$

with sum-to-zero constrains:

$$\alpha_1 + \alpha_2 = \beta_1 + \beta_2 = (\alpha\,\beta)_{11} + (\alpha\,\beta)_{12} = (\alpha\,\beta)_{21} + (\alpha\,\beta)_{22} = (\alpha\,\beta)_{11} + (\alpha\,\beta)_{21} = 0.$$

Define $\alpha_2 = \alpha,\ \beta_2 = \beta$, then we can write

$$\alpha_1 = -\alpha, \beta_1 = -\beta, (\alpha\,\beta)_{21} = -(\alpha\,\beta), (\alpha\,\beta)_{11} = (\alpha\,\beta), (\alpha\,\beta)_{12} = -(\alpha\,\beta)$$

i	j	A	B	Observation	Without constraints	With constraints
1	1	-	-	(1)	$\eta + \alpha_1 + \beta_1 + (\alpha\,\beta)_{11}$	$\eta - \alpha - \beta + (\alpha\,\beta)$
2	1	+	-	a	$\eta + \alpha_2 + \beta_1 + (\alpha\,\beta)_{21}$	$\eta + \alpha - \beta - (\alpha\,\beta)$
1	2	-	+	b	$\eta + \alpha_1 + \beta_2 + (\alpha\,\beta)_{12}$	$\eta - \alpha + \beta - (\alpha\,\beta)$
2	2	+	+	ab	$\eta + \alpha_2 + \beta_2 + (\alpha\,\beta)_{22}$	$\eta + \alpha + \beta + (\alpha\,\beta)$

Table 12. Summary of 2^2 design model

Then we have the least square estimates of the parameters:

$$\begin{aligned}
\hat{\eta} &= [a\,b + a + b + (1)]/4 = y... \\
\hat{\alpha} &= [a\,b + a - b - (1)]/4 \\
\hat{\beta} &= [a\,b - a + b - (1)]/4 \\
\widehat{(\alpha\,\beta)} &= [a\,b - a - b + (1)]/4
\end{aligned}$$

Remark 5. The interaction coefficients cancel out because of the sum-to-zero constrains.

The **main effect of A** is defined as:

$$\begin{aligned}
A &= 2\,\hat{\alpha} = [a\,b + a - b - (1)]/2 \\
&= [a\,b + a]/2 - [b + (1)]/2 \\
&= (1/2)\,\{[a\,b - b] + [a - (1)]\}
\end{aligned}$$

The **interaction effect** is defined as:

$$\begin{aligned}
A\,B &= 2\,\widehat{(\alpha\,\beta)} \\
&= [a\,b - a - b + (1)]/2 \\
&= (1/2)\,\{[a\,b - b] - [a - (1)]\} \\
&= (1/2)\,\{[a\,b - a] - [b - (1)]\}
\end{aligned}$$

Notation 2. *We use upraight letters to denote the factors, and use italic letters to denote the effects defined above.*

3.2.2 Generalization to 2^3 Design

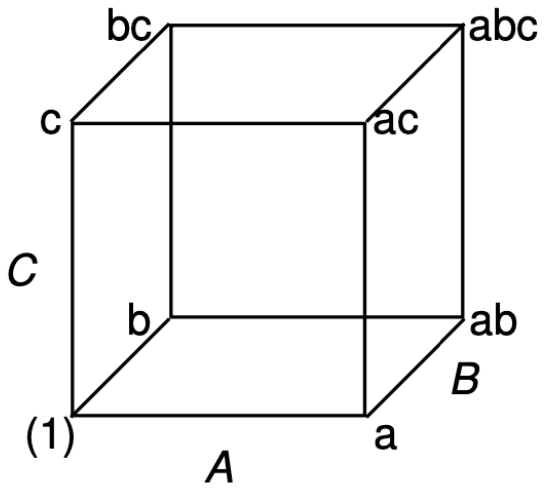

Figure 13. 2^3 design

Mean effect of A:

$$\begin{aligned}
A =& (a + a\,b + a\,c + a\,b\,c)/4 - ((1) + b + c + b\,c)/4 \\
=& \text{mean of high A} - \text{mean of low A} \\
=& (1/4)\,\{(a - (1)) + (a\,b - b) + (a\,c - c) + (a\,b\,c - b\,c)\} \\
=& \text{mean of 4univariate A effects} \\
=& (1/2)\,[\{a\,b\,c + a\,c - b\,c - c\}/2 + \{a\,b + a - b - (1)\}/2] \\
=& (1/2)\,\{\text{Effect in A-B design at high C+Aeffect in A-B design at low C}\}
\end{aligned}$$

Interaction Effects:

- The 2^3 consists of two 2^2 designs: one at high C and one at low C
- AB interaction at low C is $A\,B_1 = (a\,b + (1) - a - b)/2$
- AB interaction at high C is $A\,B_2 = (a\,b\,c + c - a\,c - b\,c)/2$
- $A\,B$ and $A\,B\,C$ effects for 2^3 are defined as

$$\begin{aligned}
A\,B &= (1/2)\,(A\,B_2 + A\,B_1) = (1/4)\,(a\,b\,c + c - a\,c - b\,c + a\,b + (1) - a - b) \\
A\,B\,C &= (1/2)\,(A\,B_2 - A\,B_1) = (1/4)\,(a\,b\,c + c - a\,c - b\,c - a\,b - (1) + a + b)
\end{aligned}$$

Alternative Expressions:

The LSEs

$$\begin{aligned}
\hat{\eta} &= [a\,b\,c + a\,b + a\,c + b\,c + a + b + c + (1)]/8 \\
A &= [a\,b\,c + a\,b + a\,c + a - b\,c - b - c - (1)]/4 \\
A\,B &= [a\,b\,c + a\,b - a\,c - b\,c - a - b + c + (1)]/4 \\
A\,B\,C &= [a\,b\,c - b\,c - a\,c - a\,b + a + b + c - (1)]/4
\end{aligned}$$

Combining the expressions, we have:

$$\begin{aligned}\hat{\eta} &= (a+1)\,(b+1)\,(c+1)/8\\ A &= (a-1)\,(b+1)\,(c+1)/4\\ AB &= (a-1)\,(b-1)\,(c+1)/4\\ ABC &= (a-1)\,(b-1)\,(c-1)/4\end{aligned}$$

- Variance of any effect estimation is $\text{var}(\text{effect}) = 4\,N^{-1}\,\sigma^2$, where N is total number of observations and $\sigma^2 = \text{var}(\epsilon)$
- Let s_i^2 be the estimate of σ^2 at the i th treatment combination ($i = 1, 2, \ldots, g$). Let $\nu_i = r_i - 1$ be the degrees of freedom for s_i^2, where r_i is the number of replicates in rth combination. Then the pooled estimate of σ^2 is:

$$s^2 = \frac{\nu_1\,s_1^2 + \cdots + \nu_g\,s_g^2}{\nu_1 + \cdots + \nu_g}$$

- Confidence intervals for estimated effects may be obtained using the Bonferroni method.
- Hidden replication property: When estimating the effect of a single factor (say, A) in a 2^3 design (with factors A, B, and C), the precision of the estimate for A is the same as it would be if you had conducted an experiment with only A at two levels with the same number of replications (for example, 4 runs at low A and 4 runs at high A).

We may use a **table of contrast** to summary the result of a 2^3 design. The "Dot product" in this table refers to the inner product between each column (e.g. A) and the response vector $\mathbf{y}$, and Dividing it by $N/2$ gives the estimate of the corresponding effect.

	A	B	C	AB	AC	BC	ABC	y
	−1	−1	−1	1	1	1	−1	60
	1	−1	−1	−1	−1	1	1	72
	−1	1	−1	−1	1	−1	1	54
	1	1	−1	1	−1	−1	−1	68
	−1	−1	1	1	−1	−1	1	52
	1	−1	1	−1	1	−1	−1	83
	−1	1	1	−1	−1	1	−1	45
	1	1	1	1	1	1	1	80
Dot product	92	−20	6.0	6.0	40	0	2.0	
Division by $N/2$	23	−5	1.5	1.5	10	0	0.5	

Table 13. Table of contrast example

Notation 3. ***(Notation summary)*** *We use upright letters (e.g.* A, AB*) to represent the factor or the corresponding columns in the table of contrast, and use italic letters (e.g. A, AB) to represent the corresponding effects defined in section 3.2.1.*

3.2.3 Regression Formulation

The model of 2^3 design can be written as:

$$\mathbb{E}(y_{ijk}) = \eta + \alpha_i + \beta_j + \gamma_k + (\alpha\,\beta)_{ij} + (\alpha\,\gamma)_{ik} + (\beta\,\gamma)_{jk} + (\alpha\,\beta\,\gamma)_{ijk}$$

For an observation y_l, Let $x_{il} = 1$ indicate the ith factor is at its high level and $x_{il} = -1$ indicate the ith factor is at its low level. Then we can write

$$\mathbb{E}(y_l) = \beta_0 + \beta_1\,x_{1l} + \beta_2\,x_{2l} + \beta_3\,x_{3l} + \beta_{12}\,x_{1l}\,x_{2l} + \beta_{13}\,x_{1l}\,x_{3l} + \beta_{23}\,x_{2l}\,x_{3l} + \beta_{123}x_{1l}x_{2l}x_{3l}.$$

The LSEs are:

$$\hat{\eta} = \hat{\beta}_0 = [a\,b\,c + a\,b + a\,c + b\,c + a + b + c + (1)]\,/\,8$$
$$A = 2\,\hat{\beta}_1 = [a\,b\,c + a\,b + a\,c + a - b\,c - b - c - (1)]\,/\,4$$
$$A\,B = 2\,\hat{\beta}_{12} = [a\,b\,c + a\,b - a\,c - b\,c - a - b + c + (1)]\,/\,4$$
$$A\,B\,C = 2\,\hat{\beta}_{123} = [a\,b\,c - b\,c - a\,c - a\,b + a + b + c - (1)]\,/\,4, \text{etc.}$$

The general 2^k design model can be written in a matrix form:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where $\mathbf{y}$ is the vector of observations and when $k = 3$:

$$\mathbf{X} = \left(\begin{array}{c|ccc|ccc|c} 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{array}\right), \quad \boldsymbol{\beta} = \begin{pmatrix} \eta \\ \alpha \\ \beta \\ \gamma \\ (\alpha\,\beta) \\ (\alpha\,\gamma) \\ (\beta\,\gamma) \\ (\alpha\,\beta\,\gamma) \end{pmatrix}$$

Note that $\mathbf{X}$ is an orthogonal matrix, then $\mathbf{X}^T\mathbf{X} = 2^k\mathbf{I}$ is an diagonal matrix, and we have

$$\hat{\boldsymbol{\beta}} = 2^{-k}\mathbf{X}^T\mathbf{y}.$$

3.3 Yates Algorithm for 2^k Experiments

Yates algorithm is a systematic computational procedure used to analyze data from a full 2^k factorial experiment. In these experiments, you have k factors, each at two levels (often labeled as "low" and "high"). The algorithm helps you calculate the main effects and interaction effects quickly and efficiently.

1. **Order the Data:**
   List the experimental data in the standard order (the order shown in the 1st column of table 14).

2. **Construct k New Columns:**
   For each new column (starting from the first column of responses), perform the following:

   - **Addition (First Half):**
     For the first 2^{k-1} entries, form each new entry by adding pairs of consecutive entries from the previous column. That is, for row i, compute:

     $$\text{New entry} = (\text{entry in row } 2\,i - 1) + (\text{entry in row } 2\,i).$$

   - **Subtraction (Second Half):**
     For the remaining 2^{k-1} entries, compute each new entry by subtracting the entry in row $2i - 1$ from the entry in row $2i$ in the previous column:

     $$\text{New entry} = (\text{entry in row } 2\,i) - (\text{entry in row } 2\,i - 1).$$

3. **Create a Scaling Column:**
   Add an additional column (the $(k+1)$th column) where all entries are 2^{k-1}, except the first entry, which is 2^k.

4. **Compute the Estimates:**
   Obtain the effect estimates by dividing the entries in the kth column by the corresponding entries in the $(k+1)$th (scaling) column.

5. **Identify the Effects:**
   Determine the identity of each effect by inspecting the signs of the factors along each row in the table.

A	B	C	y	(1)	(2)	(3)	Div.	Est.	Effect
-1	-1	-1	60	132	254	514	8	64.25	$\bar{y}$
1	-1	-1	72	122	260	92	4	23.00	A
-1	1	-1	54	135	26	-20	4	-5.00	B
1	1	-1	68	125	66	6	4	1.50	AB
-1	-1	1	52	12	-10	6	4	1.50	C
1	-1	1	83	14	-10	40	4	10.00	AC
-1	1	1	45	31	2	0	4	0.00	BC
1	1	1	80	35	4	2	4	0.50	ABC

Table 14. Example of data table created by Yates algorithm

3.4 Factorial Design with blocks

3.4.1 2^2 Design in 2 Blocks of Size 2

Let $x_1, x_2, z \in \{-1, 1\}$, where x_1, x_2 indicates levels of A and B and z indicates the block variables. The model is defined as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \delta z + \varepsilon$$

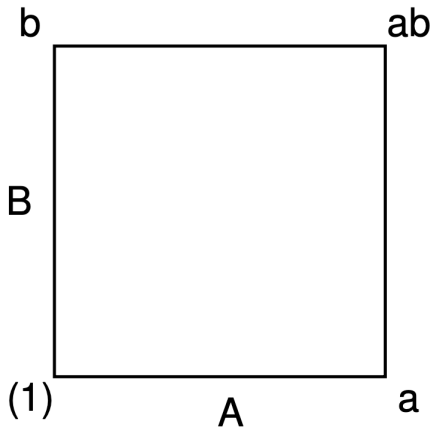

Figure 14. 2^2 design with blocks

Consider 3 possible designs:

- D1. (1), a in block 1; b, ab in block 2
- D2. (1), b in block 1; a, ab in block 2
- D3. (1), ab in block 1; a, b in block 2

In D1, B is confounded by the block effect because $z = x_2$ leads to:

$$\mathbb{E}(ab + b - a - (1)) = 4\beta_2 + 4\delta,$$

while A and AB are not confounded.

3.4.2 2^3 in 2 Blocks of Size 4

OFAT Block Design:

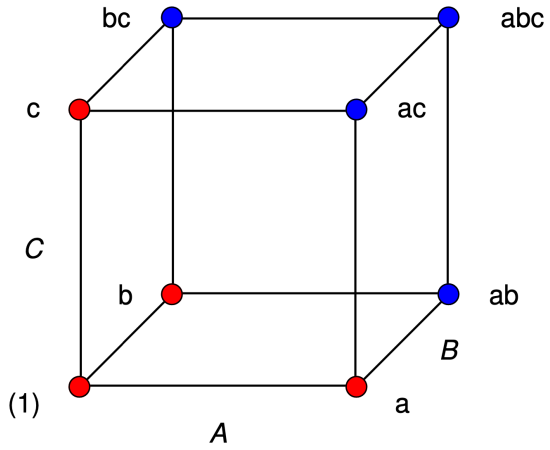

Figure 15. OFAT block design. Different colors indicate the experiment is conducted in different blocks.

	A	B	C	AB	AC	BC	ABC	Block
(1)	−	−	−	+	+	+	−	−
a	+	−	−	−	−	+	+	−
b	−	+	−	−	+	−	+	−
ab	+	+	−	+	−	−	−	+
c	−	−	+	+	−	−	+	−
ac	+	−	+	−	+	−	−	+
bc	−	+	+	−	−	+	−	+
abc	+	+	+	+	+	+	+	+

Table 15. Treatment table of OFAT block design

The corresponding model is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3 + \delta z + \epsilon \tag{4}$$

A is confounded in the design shown in Figure 15 and Table 15 because the expression

$$A = [a\,b\,c + a\,b + a\,c + a - b\,c - b - c - (1)]/4$$

cannot eliminate δ if we plug in (4).

We may change the design of the experiment to make the parameters of interest unfounded while some other parameters confounded. As shown in table 16, we can make the assignment of blocks exactly the same as the assignment of the 3-way interaction ABC. By doing this, only ABC is confounded in our data. However, we can never make all the parameters unconfounded if we have to impose 2 blocks of size 4.

	A	B	C	AB	AC	BC	ABC	Block
(1)	-	-	-	+	+	+	-	-
a	+	-	-	-	-	+	+	+
b	-	+	-	-	+	-	+	+
ab	+	+	-	+	-	-	-	-
c	-	-	+	+	-	-	+	+
ac	+	-	+	-	+	-	-	-
bc	-	+	+	-	-	+	-	-
abc	+	+	+	+	+	+	+	+

Table 16. Experimental design such that only ABC is confounded.

Remark 6. In a 2^k design in 2 blocks, we usually use the largest-order interaction column to define the assignment of blocks such that only the highest order interaction is confounded.

3.4.3 2^3 in 4 Blocks of Size 2

If we have 4 blocks, we need two block variables $X, Y \in \{-1, 1\}$ to represent the block combinations in the table.

	A	B	C	AB	AC	BC	ABC	X	Y
(1)	-	-	-	+	+	+	-	+	+
a	+	-	-	-	-	+	+	-	+
b	-	+	-	-	+	-	+	-	-
ab	+	+	-	+	-	-	-	+	-
c	-	-	+	+	-	-	+	+	-
ac	+	-	+	-	+	-	-	-	-
bc	-	+	+	-	-	+	-	-	+
abc	+	+	+	+	+	+	+	+	+

Table 17. 2^3 design with blocks of size 2

If we put the columns $X = AB, Y = \text{BC}$, then **all** the two-way interactions are confounded while the other parameters are not.

3.4.4 2^4 in 4 blocks

A	B	C	D	ABC	ABD	CD	Block
-	-	-	-	-	-	+	1
+	-	-	-	+	+	+	4
-	+	-	-	+	+	+	4
+	+	-	-	-	-	+	1
-	-	+	-	+	-	-	3
+	-	+	-	-	+	-	2
-	+	+	-	-	+	-	2
+	+	+	-	+	-	-	3
-	-	-	+	-	+	-	2
+	-	-	+	+	-	-	3
-	+	-	+	+	-	-	3
+	+	-	+	-	+	-	2
-	-	+	+	+	+	+	4
+	-	+	+	-	-	+	1
-	+	+	+	-	-	+	1
+	+	+	+	+	+	+	4

Table 18. 2^4 design with blocks of size 2

Use ABC and ABD to define blocks as shown in figure 18, then CD is confounded as well since $\text{ABC} \times \text{ABD} = \text{CD}$.

3.4.5 Partial Confounding

4 Fractional two-level designs

2^k designs include too many interactions and variables when k is large. But in most cases we are only interested in few of them. **Fractional factorial design** reduce the number of experiments by assuming many higher-order interactions are negligible and we only care about main effects and two-factor interactions.

4.1 Half-fraction of 2^3

Assume that all interactions are negligible, then the model can be written as

$$y = \mu + \alpha x_1 + \beta x_2 + \gamma x_3 + \varepsilon, \quad x_1, x_2, x_3 = \pm 1 \tag{5}$$

Now we investigate different experimental designs.

1. **Design 1 (OFAT):** Using runs (1), a, b and c, i.e. using the design matrix:

$$X = \begin{pmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix},$$

   and the covariance matrix for the estimation is:

$$\sigma^2(X^TX)^{-1} = \sigma^2 \begin{pmatrix} 1 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1/4 & 1/4 \\ 1/2 & 1/4 & 1/2 & 1/4 \\ 1/2 & 1/4 & 1/4 & 1/2 \end{pmatrix}$$

2. **Design 2:** Using runs a, b, c, and abc, i.e. using the design matrix:

$$X = \begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix},$$

   and the covariance matrix for the estimation is

$$\sigma^2(X^TX)^{-1} = \sigma^2 \begin{pmatrix} 1/4 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 \\ 0 & 0 & 1/4 & 0 \\ 0 & 0 & 0 & 1/4 \end{pmatrix}$$

4.2 Construction of 2^{5-1} design

1. Write a full 2^4 design in variables A, B, C and D
2. Add a column of signs for factor E defined by E = ABCD

Notation 4. *Product of upright letters (e.g.* ABCD*) is a shorthand for taking Hadamard (element-wise) product of the corresponding column vector in the design table.*

A	B	C	D	E	A	B	C	D	E
-	-	-	-	+	-	-	-	+	-
+	-	-	-	-	+	-	-	+	+
-	+	-	-	-	-	+	-	+	+
+	+	-	-	+	+	+	-	+	-
-	-	+	-	-	-	-	+	+	+
+	-	+	-	+	+	-	+	+	-
-	+	+	-	+	-	+	+	+	-
+	+	+	-	-	+	+	+	+	+

Table 19. E=ABCD design

The equation

$$E = ABCD$$

is called the generator of the design. Multiplying both sides of the generator by E gives:

$$\mathrm{I} = \mathrm{ABCDE}, \tag{6}$$

where I is a column of positive signs and (6) is called the **defining relation** of the design. We can easily obtain its confounding pattern by multiplying (6) on both sides by any main or interaction effect, e.g., A = BCDE, B = ACDE, AB = CDE, which means the effects on both sides of the equality is confounded with each other. The expected values of the estimations are:

$$\begin{aligned}
\mathbb{E}(\bar{y}_1) &= \eta + A\,B\,C\,D\,E/2 \\
\mathbb{E}(\hat{A}_1) &= A + B\,C\,D\,E \\
&\vdots \\
\mathbb{E}(\widehat{A\,B}_1) &= A\,B + C\,D\,E \\
&\vdots \\
\mathbb{E}(\widehat{A\,B\,C}_1) &= A\,B\,C + D\,E \\
&\vdots \\
\mathbb{E}(\widehat{A\,B\,C\,D}_1) &= A\,B\,C\,D + E
\end{aligned}$$

Combining two half-fractions. Suppose we are allowed to run another 2^{5-1} experiment with defining relation E = −ABCD. Let δ be the block effect of the second run, $\hat{A}_2, \ldots, \widehat{ABCD}_2$ be the estimations for the corresponding effects in the second run, then we have

$$\begin{aligned}
\mathbb{E}(\bar{y}_2) &= \eta + \delta - A\,B\,C\,D\,E/2 \\
\mathbb{E}(\hat{A}_2) &= A - B\,C\,D\,E \\
&\vdots \\
\mathbb{E}(\widehat{A\,B}_2) &= A\,B - C\,D\,E \\
&\vdots \\
\mathbb{E}(\widehat{A\,B\,C}_2) &= A\,B\,C - D\,E \\
&\vdots \\
\mathbb{E}(\widehat{A\,B\,C\,D}_2) &= A\,B\,C\,D - E.
\end{aligned}$$

Thus, we can obtain unbiased estimation of each effect by combining two estimations. e.g.

$$\begin{aligned}
\mathbb{E}[(\hat{A}_1 + \hat{A}_2)/2] &= A \\
\mathbb{E}[(\widehat{AB}_1 + \widehat{AB}_2)/2] &= AB,
\end{aligned}$$

while the grand mean is still confounded by the block effect δ.

4.3 Resolution

Definition 3. *The **resolution** of a two-level fractional design is length of the shortest non-identity word in the defining relation.*

Interpretation: A design with **Resolution R** means that **Any p-factor effect** (an effect that involves p factors, such as a main effect where $p = 1$ or an interaction where $p > 1$) is **not** confounded with any other effect that involves fewer than $R - p$ **factors**.

- A design with **Resolution III** means that while the mean effect are not aliased with one another, they may be aliased with two-factor interactions.

- A design with **Resolution IV** ensures that main effects are unconfounded with two-factor interactions while two-factor interactions might be aliased with each other.

- In a **Resolution V** design, main effects and two-factor interactions are not aliased with each other. In other words, any confounding (aliasing) involves interactions of order three or higher.

4.3.1 Half-fractional design for the highest resolution

To construct a half-fraction of the highest resolution (i.e., a 2^{k-1} design) with maximal clarity for estimating lower-order effects:

1. Start with a Full Factorial: Construct a full 2^{k-1} factorial design for the first $k-1$ factors.

2. Generate the k th Factor: Define the k th factor as the product of the first $k-1$ factors (i.e., $x_k = x_1 \times x_2 \times \cdots \times x_{k-1}$). This leads to the defining relation:

$$\mathrm{I} = \mathrm{F}_1\mathrm{F}_2 \ldots \mathrm{F}_k$$

   where F_i denotes the ith factor, and the defining relation has length k.

This approach produces the highest possible resolution for a 2^{k-1} design because a longer defining word (length k) ensures that any aliasing only involves high-order interactions (order k or more). As a consequence, main effects and low-order interactions (which are usually of primary interest) remain unconfounded, maximizing the interpretability of the experimental results.

Remark 7. If we remove any one column in a 2^{k-1} design, the remaining four columns form a complete 2^k factorial. A fractional factorial design of resolution R is constructed in such a way that every $R-1$ factors appears as a complete factorial.

4.3.2 Sequential use of fractional designs

It is generally advisable for an experimenter planning a 2^5 design in 32 runs to first conduct a half-fraction (16 runs), review the results, and then decide how to proceed. Here are the key points:

1. **Within-Fraction Randomization:**
   Randomize the order of runs within each half-fraction.

2. **Orthogonal Blocking When Combining Fractions:**
   If you choose to run another fraction later, structure them as randomized orthogonal blocks of the complete factorial design.

3. **Retention of Information:**
   No information is lost except for the interaction effect that becomes confounded with the block effect.

4. **Enhanced Precision:**
   Running the design as two randomized fractions can provide greater precision than a full random order because the block effect is removed.

4.3.3 Construction of Resolution III designs

Resolution III designs involving $2^k - 1$ variables can be constructed by "saturating" a 2^k factorial design with additional variables. For instance, let $k = 4$, a 2^{15-11}_{III} design (a design with 2^{15-11} runs and resolution III) may be obtained by the following procedure:

1. **Full Factorial Formation:**
   Start by constructing a full factorial design on variables 1, 2, 3, and 4.

2. **Association with Interaction Columns:**
   Assign the additional variables, numbered 5 through 15, to the 11 interaction columns corresponding to the interactions:

$$\pm F_1F_2,\ \pm F_1F_3,\ \pm F_1F_4,\ \pm F_2F_3,\ \pm F_2F_4,\ \pm F_3F_4,$$
$$\pm F_1F_2F_3,\ \pm F_1F_2F_4,\ \pm F_1F_3F_4,\ \pm F_2F_3F_4,\ \pm F_1F_2F_3F_4.$$

   In this context, each distinct assignment of the ± signs to these interaction terms defines one fraction of the design.

Example 8. (2^{7-4}_{III} design) Use generators F_4=F_1F_2, F_5=F_1F_3, F_6=F_2F_3 and F_7=$F_1F_2F_3$

1	2	3	4	5	6	7
-	-	-	+	+	+	-
+	-	-	-	-	+	+
-	+	-	-	+	-	+
+	+	-	+	-	-	-
-	-	+	+	-	-	+
+	-	+	-	+	-	-
-	+	+	-	-	+	-
+	+	+	+	+	+	+

Table 20. 2^{7-4}_{III} design

If we run another half-fractional design by switching the sign for 4, 5, 6, 7, then F_4=-F_1F_2, F_5=-F_1F_3, F_6=-F_2F_3 and F_7=-$F_1F_2F_3$ then the shortest defining relation becomes

$$I = (F_1F_2 \cdot F_4)(F_1F_3 \cdot F_5) = F_2F_4F_3F_5.$$

The resolution is 4 and we therefore have a 2^{7-3}_{IV} design.

4.3.4 Plackett-Burman Designs

Saturated fractional factorial designs are valued for their **orthogonality**: in any two columns of the design matrix, exactly half the runs in column A are "+," and among those, half are "+" and half "–" in column B; the same balance holds for the runs where column A is "–." Under the assumption that all interactions are negligible, this balance guarantees unbiased, minimum-variance estimates of all main effects for $k = N - 1$ factors in just N runs.

- **Classical limitation:** Regular 2^{k-p} designs exist only when N is a power of 2.
- **Plackett & Burman (1946):** Extended this orthogonality to any N that is a multiple of 4, allowing efficient main-effects screening with non–power-of-2 runs.
- **Example ($k = 11$, $N = 12$):**

  1. Write the first row of 11 signs (e.g. "+ − + + − ..."). If you have a Hadamard matrix of order N, remove the **1** column. Any one of the remaining rows of length $N - 1$ can be the first row. Standard choices of first rows are summarized in table 35.
  2. Generate each subsequent row by cyclically shifting the previous row one column to the right.
  3. Add a final 12th row of all "–" signs.

This simple "shifted-row" construction produces a 12×11 Plackett–Burman design that retains the key orthogonality property for unbiased estimation of main effects.

Run	1	2	3	4	5	6	7	8	9	10	11
1	+	-	+	-	-	-	+	+	+	-	+
2	+	+	-	+	-	-	-	+	+	+	-
3	-	+	+	-	+	-	-	-	+	+	+
4	+	-	+	+	-	+	-	-	-	+	+
5	+	+	-	+	+	-	+	-	-	-	+
6	+	+	+	-	+	+	-	+	-	-	-
7	-	+	+	+	-	+	+	-	+	-	-
8	-	-	+	+	+	-	+	+	-	+	-
9	-	-	-	+	+	+	-	+	+	-	+
10	+	-	-	-	+	+	+	-	+	+	-
11	-	+	-	-	-	+	+	+	-	+	+
12	-	-	-	-	-	-	-	-	-	-	-

Table 21. Plackett-Burman design for $k = 11$ and $N = 12$

4.3.5 Construction of Resolution IV Designs

You can obtain a 2_{IV}^{k-p} design from any resolution III fraction in three steps:

1. **Base Fraction (Resolution III):**
   Construct a $2_{\text{III}}^{(k-1)-p}$ design on the first $k-1$ factors.

2. **Add a "Pure" Column:**
   Append a kth factor column consisting entirely of "+" signs to that design.

3. **Foldover:**
   Create a second block of runs by flipping every sign in the first block. Together, the original block and its foldover form a resolution IV design on k factors.

Example 9. Constructing a 2_{IV}^{7-3} design from 2_{III}^{6-3} Design

1. **Base 2_{III}^{6-3} Design.** Define the generators:

$$\text{F}_4 = \text{F}_1\text{F}_2, \quad \text{F}_5 = \text{F}_1\text{F}_3, \quad \text{F}_6 = \text{F}_2\text{F}_3.$$

   The defining relation for the first $2^{6-3} = 8$ runs is:

$$\text{I} = \text{F}_1\text{F}_2\text{F}_4 = \text{F}_1\text{F}_3\text{F}_5 = \text{F}_2\text{F}_3\text{F}_6.$$

2. **Add the seventh factor F_7.** Assign $\text{F}_7 = +1$ in the first half, introducing the word

$$\text{I} = \text{F}_7.$$

3. **Fold-over to Resolution IV.** Create a second set of 8 runs by flipping every sign in the first half. Its defining relation becomes:

$$I = -\text{F}_1\text{F}_2\text{F}_4 = -\text{F}_1\text{F}_3\text{F}_5 = -\text{F}_2\text{F}_3\text{F}_6 = -\text{F}_7.$$

4. **Combine Both Halves.** When the two halves are pooled, each length-3 word appears once with "+" and once with "–," so they cancel. The shortest remaining aliasing words involve four factors, yielding a *resolution IV* design.

Example 10. Constructing a 2_{IV}^{7-3} design from 2_{III}^{7-4} Design.

1. **Base 2_{III}^{7-4} design**: Use generators

$$\mathrm{F}_4 = \mathrm{F}_1\mathrm{F}_2, \quad \mathrm{F}_5 = \mathrm{F}_1\mathrm{F}_3, \quad \mathrm{F}_6 = \mathrm{F}_2\mathrm{F}_3, \quad \mathrm{F}_7 = \mathrm{F}_1\mathrm{F}_2\mathrm{F}_3.$$

   Its defining relation is

$$\begin{aligned} \mathrm{I}_1 =& \mathrm{F}_1\mathrm{F}_2\mathrm{F}_4 = \mathrm{F}_1\mathrm{F}_3\mathrm{F}_5 = \mathrm{F}_2\mathrm{F}_3\mathrm{F}_6 = \mathrm{F}_1\mathrm{F}_2\mathrm{F}_3\mathrm{F}_7 \\ &= \mathrm{F}_2\mathrm{F}_3\mathrm{F}_4\mathrm{F}_5 = \mathrm{F}_1\mathrm{F}_3\mathrm{F}_4\mathrm{F}_6 = \mathrm{F}_3\mathrm{F}_4\mathrm{F}_6\mathrm{F}_7 \\ &= \mathrm{F}_1\mathrm{F}_2\mathrm{F}_5\mathrm{F}_6 = \mathrm{F}_2\mathrm{F}_5\mathrm{F}_6\mathrm{F}_7 = \mathrm{F}_1\mathrm{F}_5\mathrm{F}_6\mathrm{F}_7 \\ &= \mathrm{F}_1\mathrm{F}_2\mathrm{F}_3\mathrm{F}_4\mathrm{F}_5\mathrm{F}_6\mathrm{F}_7. \end{aligned}$$

2. **Second Fraction by Sign-Switching**: Flip every sign in all columns of the first fraction. Its defining relation becomes

$$\begin{aligned} \mathrm{I}_2 =& -\mathrm{F}_1\mathrm{F}_2\mathrm{F}_4 = -\mathrm{F}_1\mathrm{F}_3\mathrm{F}_5 = -\mathrm{F}_2\mathrm{F}_3\mathrm{F}_6 = \mathrm{F}_1\mathrm{F}_2\mathrm{F}_3\mathrm{F}_7 \\ &= \mathrm{F}_2\mathrm{F}_3\mathrm{F}_4\mathrm{F}_5 = \mathrm{F}_1\mathrm{F}_3\mathrm{F}_4\mathrm{F}_6 = -\mathrm{F}_3\mathrm{F}_4\mathrm{F}_6\mathrm{F}_7 \\ &= \mathrm{F}_1\mathrm{F}_2\mathrm{F}_5\mathrm{F}_6 = -\mathrm{F}_2\mathrm{F}_5\mathrm{F}_6\mathrm{F}_7 = -\mathrm{F}_1\mathrm{F}_5\mathrm{F}_6\mathrm{F}_7 \\ &= -\mathrm{F}_1\mathrm{F}_2\mathrm{F}_3\mathrm{F}_4\mathrm{F}_5\mathrm{F}_6\mathrm{F}_7. \end{aligned}$$

3. **Combine to Resolution IV**: Pooling both fractions cancels all length-3 and length-7 words, leaving only length-4 words as the shortest aliases. The final defining relation is

$$\mathrm{I} = \mathrm{F}_1\mathrm{F}_2\mathrm{F}_3\mathrm{F}_7 = \mathrm{F}_2\mathrm{F}_3\mathrm{F}_4\mathrm{F}_5 = \mathrm{F}_1\mathrm{F}_3\mathrm{F}_4\mathrm{F}_6 = \mathrm{F}_1\mathrm{F}_2\mathrm{F}_5\mathrm{F}_6 = \mathrm{F}_3\mathrm{F}_5\mathrm{F}_6\mathrm{F}_7 = \mathrm{F}_2\mathrm{F}_4\mathrm{F}_6\mathrm{F}_7 = \mathrm{F}_1\mathrm{F}_4\mathrm{F}_5\mathrm{F}_7.$$

Remark 8. (Selection of Resolution IV and V Designs)

- A *main effect* or *two-factor interaction* (2-fi) is said to be ***clear*** if none of its aliases is a main effect or two-factor interaction.
- A main effect or two-factor interaction is said to be ***strongly clear*** if none of its aliases is a main effect, two-factor interaction, or three-factor interaction (3-fi).

1. In any Resolution IV design, all main effects are clear.
2. In any Resolution V design, all main effects are strongly clear and all two-factor interactions are clear.
3. Among Resolution IV designs with fixed k and p, those maximizing the number of clear two-factor interactions are preferred.

4.4 Blocking Fractional Factorials

Blocking fractional factorials is a technique used in *fractional factorial experimental designs* to control *nuisance variables* (like time, batch, machine differences) without confounding them with the *primary treatment effects* you care about.

1. **Add "block" as a new factor**, but instead of randomizing fully across all runs, you restrict randomization *within each block*.
2. The blocks are **chosen based on defining relations** — you decide that certain interactions (e.g., AC) are "negligible," so you use them to split the design into blocks.

3. As a result, the block effects are **confounded** (aliased) with higher-order interactions, but **not** with the main effects or low-order interactions you're most interested in.

4.4.1 2_V^{5-1} in 2 or 4 blocks

Suppose we have a 2^{5-1} fractional factorial design, with defining relation:

$$\mathrm{I} = \mathrm{ABCDE}$$

If we want to run it in 2 **blocks** of 8 runs each. we might use AC as the **blocking factor**:

- Runs where $\mathrm{AC} = +1$ go into Block 1
- Runs where $\mathrm{AC} = -1$ go into Block 2

If we want to run it in **4 blocks of size 4**. We might use AC and CD as the blocking factors

- Runs with $\mathrm{AC} = +$ and $\mathrm{BD} = +$ $\rightarrow$ Block 1.
- Runs with $\mathrm{AC} = +$ and $\mathrm{BD} = -$ $\rightarrow$ Block 2.
- Runs with $\mathrm{AC} = -$ and $\mathrm{BD} = +$ $\rightarrow$ Block 3.
- Runs with $\mathrm{AC} = -$ and $\mathrm{BD} = -$ $\rightarrow$ Block 4.

A	B	C	D	E	AC
−	−	−	−	+	+
+	−	−	−	−	−
−	+	−	−	−	+
+	+	−	−	+	−
−	−	+	−	−	−
+	−	+	−	+	+
−	+	+	−	+	−
+	+	+	−	−	+
−	−	−	+	−	+
+	−	−	+	+	−
−	+	−	+	+	+
+	+	−	+	−	−
−	−	+	+	+	−
+	−	+	+	−	+
−	+	+	+	−	−
+	+	+	+	+	+

Table 22. 2_V^{5-1} in 2 blocks of size 8

A	B	C	D	E	AC	BC
−	−	−	−	+	+	+
+	−	−	−	−	−	+
−	+	−	−	−	+	−
+	+	−	−	+	−	−
−	−	+	−	−	−	−
+	−	+	−	+	+	−
−	+	+	−	+	−	+
+	+	+	−	−	+	+
−	−	−	+	−	+	+
+	−	−	+	+	−	+
−	+	−	+	+	+	−
+	+	−	+	−	−	−
−	−	+	+	+	−	−
+	−	+	+	−	+	−
−	+	+	+	−	−	+
+	+	+	+	+	+	+

Table 23. 2_V^{5-1} in 4 blocks of size 4

4.4.2 2_{IV}^{4-1} in 4 blocks

Consider a 2_{IV}^{4-1} design with defining relation

$$\mathrm{I} = \mathrm{ABCD}.$$

We introduce blocking factors X = −AC and Y = −AB to define the blocks. Then we get a **"fold-over"** design as shown in table 24 with following properties:

- Any fold-over design of resolution IV can be divided into blocks of size two.

  without losing its resolution IV property.

- Each block consists of a pair of complementary treatment combinations.

- The blocking procedure corresponds to taking the 2fi's as blocking factors.

A	B	C	D	X	Y
-	-	-	-	-	-
+	-	-	+	+	+
-	+	-	+	-	+
+	+	-	-	+	-
-	-	+	+	+	-
+	-	+	-	-	+
-	+	+	-	+	+
+	+	+	+	-	-

Table 24. 2^{4-1}_{IV} in 4 blocks

4.4.3 Blocking resolution III designs

In a **saturated Resolution III** design with 2^k runs and $2^k - 1$ factors:

- **Blocking one factor** (e.g., A) into two blocks leaves $2^k - 2$ treatment factors, with each block having 2^{k-1} runs.

- **Blocking two factors and their interaction** (e.g., A, B, and AB) creates **four blocks**, reducing to $2^k - 4$ treatment factors. Because AB would correspond to a third factor.

Example:

- Start with $2^4 = 16$ runs, factors A, B, C, D.

- Define blocking factors $X = \text{AD}$, $Y = \text{BC}$.

- Skip XY = ABCD (already determined by X and Y).

- Add new factors: $E = \text{AB}$, $F = \text{AC}$, $G = \text{BD}$, $H = \text{CD}$, $J = \text{ABC}$, $K = \text{ABD}$, $L = \text{ACD}$, $M = \text{BCD}$.

Result:

- A 2^{12-8}_{III} design: 12 factors, 4 blocks of 4 runs each, maintaining Resolution III.

5 Random Effects Model

5.1 General Case: Mixed Effects Model

Consider the following mixed-effects formulation for a three-way ANOVA structure:

$$y_{ijk} = \mu + \alpha_i + b_j + c_{ij} + \varepsilon_{ijk}, i = 1, \ldots, I,\ j = 1, \ldots, J,\ k = 1, \ldots, n,$$

where

- μ denotes the overall intercept.
- α_i are fixed primary-factor effects satisfying $\sum_{i=1}^{I} \alpha_i = 0$.
- b_j are random secondary-factor effects with $\mathbb{E}[b_j] = 0$ and $\mathrm{Var}(b_j) = \sigma_b^2$.
- c_{ij} are random interaction effects between levels i and j, with $\mathbb{E}[c_{ij}] = 0$, $\mathrm{Var}(c_{ij}) = \sigma_c^2$, and, for each j, $\sum_{i=1}^{I} c_{ij} = 0$.
- ε_{ijk} are residual errors satisfying $\mathbb{E}[\varepsilon_{ijk}] = 0$, $\mathrm{Var}(\varepsilon_{ijk}) = \sigma^2$, and independent of all random effects.

This model admits the compact matrix representation

$$y = X\,\beta + Z\,a + \varepsilon, \tag{7}$$

where

- y is the $(I\ J\ n) \times 1$ response vector.
- $\beta = (\mu,\ \alpha_1, \ldots, \alpha_I)^T$ is the $(I+1) \times 1$ vector of fixed effects.
- $a = (b_1, \ldots, b_J,\ c_{11}, \ldots, c_{\mathrm{IJ}})^T$ is the random-effects vector and

$$\mathrm{Cov}(b_j,\ c_{ii'}) = 0 \quad \forall\ i, i', j.$$

- X and Z are the corresponding design matrices for fixed and random effects, respectively.
- ε is the residual vector of length $I\ J\ n$.

We assume

$$\mathbb{E}[a] = 0, \mathbb{E}[\varepsilon] = 0, \mathrm{Cov}(a, \varepsilon) = 0,$$

$$\mathrm{Cov}(a) = V, \mathrm{Cov}(\varepsilon) = \sigma^2\ I,$$

with I the identity matrix, V typically block-diagonal and can be partitioned as

$$V = \begin{pmatrix} \underbrace{\sigma_b^2\ I_J}_{\mathrm{Var}(b)} & 0 \\ 0 & \underbrace{\sigma_c^2\ I_{I\ J}}_{\mathrm{Var}(c)} \end{pmatrix}.$$

Continuous-sum constraints $\sum_i \alpha_i = 0$ and $\sum_i c_{ij} = 0$ ensure identifiability of fixed and interaction effects.

5.1.1 Recap: estimable function and its properties

Now we recall the definition and some properties of estimable functions.

Definition 4. ***(Estimable Function)*** *A linear function of the parameter vector $\psi = c^T\beta$ is said to be* ***estimable*** *if there exists a vector $a \in R^n$ such that the linear statistic $a^T y$ is unbiased for ψ, i.e. $E[a^T y] = c^T \beta$ for all β.*

Theorem 1. ***(Characterization of Estimability)*** *The function $\psi = c^T\beta$ is estimable if and only if the row vector c^T lies in the row space of the design matrix X. Equivalently, there exists a vector $a \in R^n$ satisfying $c^T = a^T X$.*

Lemma 1. ***(Uniqueness via Projection)*** *Suppose $\psi = c^T\beta$ is estimable, and let $V_r = \mathrm{Col}(X)$ denote the column space of X. Then among all unbiased linear estimators of ψ, there is a unique one of the form $a^T y$ with $a \in V_r$. Moreover, if $b^T y$ is any unbiased estimator of ψ, then $a = \mathrm{Proj}_{V_r}(b)$.*

Theorem 2. ***(Gauss–Markov).***
Every estimable function $\psi = c^T\beta$ admits a unique linear unbiased estimator $\hat{\psi}$whose variance is minimal among all unbiased linear estimators. Concretely, if $\hat{\beta}$ is any (generalized) least-squares solution to $X\hat{\beta} = y$, then

$$\hat{\psi} = c^T \ \hat{\beta}$$

is the Best Linear Unbiased Estimator (BLUE) of ψ.

5.1.2 Estimation of contrasts

Under the linear mixed-effects framework, consider estimation of an estimable contrast $\psi = \lambda^T\beta$. We proceed as follows:

1. **Condition for estimability.**
   The scalar function $\psi = \lambda^T\beta$ is estimable if and only if the row vector λ^T lies in the row space of the fixed-effects design matrix X.

2. **Covariance structure.**
   Writing $\Sigma = \mathrm{Cov}(y) = ZVZ^T + \sigma^2 I$, where $V = \mathrm{Cov}(a)$ and $\sigma^2 I = \mathrm{Cov}(\varepsilon)$, the generalized least squares (GLS) estimator for the fixed-effects vector is

   $$\hat{\beta}_{\mathrm{GLS}} = (X^T\Sigma^{-1}X)^- X^T\Sigma^{-1}y,$$

   and hence the Best Linear Unbiased Estimator (BLUE) of ψ is

   $$\hat{\psi}_{\mathrm{GLS}} = \lambda^T\hat{\beta}_{\mathrm{GLS}} = \lambda^T(X^T\Sigma^{-1}X)^- X^T\Sigma^{-1}y.$$

3. **Balanced-data simplification.**
   In many applications—particularly when the design is orthogonal or balanced—the covariance matrix Σ is unknown but need not be explicitly estimated: one can show that the ordinary least squares estimator

   $$\hat{\psi}_{\mathrm{OLS}} = \lambda^T(X^TX)^- X^T y$$

   coincides with the GLS-based BLUE of ψ.

4. **Illustrative example.**
   For the mixed-effects model

   $$y_{ijk} = \mu + \alpha_i + b_j + c_{ij} + \varepsilon_{ijk},$$

   the fixed-effects component is $\mathbb{E}[y_{ijk}] = \mu + \alpha_i$. The contrast $\alpha_1 - \alpha_2$ is estimable (since its coefficient vector lies in the row space of X), and under balance the OLS estimator

   $$\hat{\alpha}_1 - \hat{\alpha}_2 = \bar{y}_{1\cdot\cdot} - \bar{y}_{2\cdot\cdot}$$

is therefore also the BLUE for the mixed model.

Best linear unbiased predictor of vector a.

Consider the mixed effect model (7), the predictor of a minimizing mean-squared error is the conditional expectation $\mathbb{E}[a|y]$.

If (a, y) is jointly Gaussian, then $\mathrm{Cov}(a, y) = V\,Z^T$, and the standard formula for the multivariate normal conditional expectation gives

$$E[a|\,y] = E[a] + \mathrm{Cov}(a, y)\ [\mathrm{Cov}(y)]^{-1}(y - E[y]) = V\,Z^T\ \Sigma^{-1}\ (y - X\ \beta).$$

More generally, $V\,Z^T\ \Sigma^{-1}\ (y - X\ \beta)$ is the **best linear unbiased predictor** (BLUP) of a. Because β is typically unknown, one replaces it by its generalized least-squares (or, in balanced designs, OLS) estimator

$$\hat{\beta} = (X^T\ \Sigma^{-1}\ X)^- X^T\ \Sigma^{-1}\ y\ = (X^T X)^- X^T\ y.$$

The resulting BLUP is

$$\hat{a} = V\,Z^T\ \Sigma^{-1}(y - X\ \hat{\beta}).$$

5.2 One-way Random Effects Model

Now we consider a special case: one-way random-effects model which is defined by

$$y_{ij}\ =\ \mu\ +\ a_i\ +\ \varepsilon_{ij}, i = 1, \ldots, t,\ \ j = 1, \ldots, n,$$

where:

1. The group effects a_i are independent and identically distributed with

$$E[a_i] = 0, \mathrm{Var}(a_i) = \sigma_a^2, \mathrm{Cov}(a_i, a_{i'}) = 0 \quad (i \neq i').$$

2. The within-group errors $\varepsilon_{\mathrm{ij}}$ are independent and identically distributed with

$$E[\varepsilon_{ij}] = 0, \mathrm{Var}(\varepsilon_{ij}) = \sigma^2,$$

and

$$\mathrm{Cov}(\varepsilon_{ij}, \varepsilon_{i'j'}) = 0 \text{whenever } i \neq i' \text{ or } j \neq j'.$$

3. The group effects and the errors are mutually independent:

$$\mathrm{Cov}(a_i, \varepsilon_{i'j'}) = 0 \text{ for all } i, i', j'.$$

Under these assumptions, the marginal moments of y_{ij} are

$$\mathrm{Var}(y_{\mathrm{ij}}) = \mathrm{Var}(a_i) + \mathrm{Var}(\varepsilon_{ij}) = \sigma_a^2 + \sigma^2, \forall\ i, j,$$

$$\mathrm{Cov}(y_{ij},\ y_{ik}) = \mathrm{Cov}(a_i + \varepsilon_{ij},\ a_i + \varepsilon_{ik}) = \mathrm{Var}(a_i) = \sigma_a^2, j \neq k,$$

$$\mathrm{Cov}(y_{ij},\ y_{i'j'}) = 0, i \neq i',$$

so that observations from distinct groups are uncorrelated, while observations within the same group share the common component a_i.

5.2.1 Estimation of random effects

Now we consider doing estimation and inference in such a model.

1. **Design matrices and variance components.**
   Writing the model in matrix form $y = X\,\beta + Z\,a + \varepsilon$, one finds that the ML (or GLS) estimator of the fixed parameters satisfies

   $$\hat{\beta} = y_{\cdot\cdot}\ ,$$

   i.e. the overall sample mean (since with two observations per group the normal equations collapse to an intercept only). We assume the random-effects covariance is

   $$V = \mathrm{Cov}(a) = \sigma_a^2\ I_3,$$

   and the marginal covariance of y is

   $$\Sigma = Z\ V\ Z^T\ +\ \sigma^2\ I_6 = \mathrm{blockdiag}(\sigma^2 I_2 + \sigma_a^2 J_2,\ \sigma^2 I_2 + \sigma_a^2 J_2,\ \sigma^2 I_2 + \sigma_a^2 J_2),$$

   where J_2 is the 2×2 matrix of ones.

2. **Inverse covariance.**
   A standard matrix-inverse identity for a compound-symmetric block gives

   $$(\sigma^2 I_2 + \sigma_a^2 J_2)^{-1} = \frac{1}{\sigma^2}\Big(I_2\ -\ \frac{\sigma_a^2}{\sigma^2 + 2\sigma_a^2}\ J_2 \Big).$$

   Hence

   $$\Sigma^{-1} = \frac{1}{\sigma^2}\ \mathrm{blockdiag}\Big(I_2 - \tfrac{\sigma_a^2}{\sigma^2 + 2\sigma_a^2}\ J_2,\ \ I_2 - \tfrac{\sigma_a^2}{\sigma^2 + 2\sigma_a^2}\ J_2,\ \ I_2 - \tfrac{\sigma_a^2}{\sigma^2 + 2\sigma_a^2}\ J_2 \Big).$$

3. **Derivation of the BLUP.**
   The Best Linear Unbiased Predictor of a is

   $$\mathrm{BLUP}(a) = V\ Z^T\ \Sigma^{-1}\ (y - X\ \hat{\beta}).$$

   Noting that Z^T concatenates the row-sums within each group (i.e.\ each block has a row of two ones in the corresponding position), one obtains after simplification

   $$\mathrm{BLUP}(a) = \frac{2\,\sigma_a^2}{\sigma^2 + 2\,\sigma_a^2} \begin{pmatrix} y_{1\cdot} - y_{\cdot\cdot} \\ y_{2\cdot} - y_{\cdot\cdot} \\ y_{3\cdot} - y_{\cdot\cdot} \end{pmatrix},$$

   where $\bar{y}_{i\cdot} = (y_{i1} + y_{i2})/2$ is the ith group mean and $\bar{y}_{\cdot\cdot}$ the overall mean.

4. **Plug-in estimator**
   Replacing σ_a^2 and σ^2 by their estimates $\hat{\sigma}_a^2$ and $\hat{\sigma}^2$ yields the empirical BLUP

This estimator shrinks each group deviation toward zero in proportion to the signal-to-noise ratio $2\sigma_a^2/(\sigma^2 + 2\sigma_a^2)$.

5.2.2 Estimation of variances

Under the one-way random-effects model, one may obtain method-of-moments estimators of the variance components by equating the observed mean squares to their theoretical expectations:

$$E[\mathrm{MS}_R]\ \ = \sigma^2, \quad E[\mathrm{MS}_T]\ \ = \sigma^2 + n\ \sigma_a^2.$$

Replacing the expectations with the corresponding sample mean squares,

$$\mathrm{MS}_R = \frac{S_R}{t(n-1)}, \mathrm{MS}_T = \frac{S_T}{t-1},$$

we obtain the moment estimators

$$\begin{aligned} \hat{\sigma}^2 &= \mathrm{MS}_R = \frac{S_R}{t(n-1)}, \\ \hat{\sigma}^2 + n\,\hat{\sigma}_a^2 &= \mathrm{MS}_T = \frac{S_T}{t-1}. \end{aligned}$$

Solving for $\hat{\sigma}_a^2$ yields

$$\hat{\sigma}_a^2 = \frac{1}{n}\,(\mathrm{MS}_T - \mathrm{MS}_R) = \frac{1}{n}\left[\frac{S_T}{t-1} - \frac{S_R}{t(n-1)}\right].$$

These estimators are unbiased for σ^2 and σ_a^2. However, in finite samples it is possible for $\hat{\sigma}_a^2$ to take a negative value, in which case one typically sets it to zero or employs alternative estimation methods (e.g. restricted maximum likelihood) to ensure nonnegativity.

Source	Sum of squares	DF	E(MS)
Treatment	$S_T = n\sum_{i=1}^{t}(y_{i\cdot} - y_{\cdot\cdot})^2$	$t-1$	$\sigma^2 + n\,\sigma_a^2$
Residual	$S_R = \sum_{i=1}^{t}\sum_{j=1}^{n}(y_{ij} - y_{i\cdot})^2$	$t\,(n-1)$	σ^2
Total	$S_D = \sum_{i=1}^{t}\sum_{j=1}^{n}(y_{ij} - y_{\cdot\cdot})^2$	$n\,t-1$	

Table 25. ANOVA Table for Random Effects Model

5.2.3 Hypothesis test

Suppose we want to test are there significant random effects in our data, it is equivalent to test

$$H_0 : \sigma_a^2 = 0.$$

We can derive hypothesis test procedure based on the following theorem.

Theorem 3. *Let* $y \sim N_{\mathrm{nt}}(\mu I_{\mathrm{nt}},\ \Sigma)$, *where*

$$\Sigma = \sigma^2\ I_{\mathrm{nt}} + \sigma_a^2\ A,$$

and A *is the block-diagonal matrix with* t *identical blocks of size* $n \times n$, *each equal to the matrix of ones* J_n*:*

$$A = \mathrm{blockdiag}(J_n, J_n, \ldots, J_n).$$

Define the between- and within-sum of squares by

$$S_T = n\sum_{i=1}^{t}(y_{i\cdot} - y_{\cdot\cdot})^2, S_R = \sum_{i=1}^{t}\sum_{j=1}^{n}(y_{ij} - y_{i\cdot})^2.$$

Then:

1. $\dfrac{S_T}{\sigma^2 + n\ \sigma_a^2} \sim \chi^2_{t-1}$.
2. $\dfrac{S_R}{\sigma^2} \sim \chi^2_{t(n-1)}$.
3. S_T *and* S_R *are independent.*

Under the full model, the ratio

$$\frac{S_T/[(t-1)\ (\sigma^2+n\ \sigma_a^2)]}{S_R/[t(n-1)\ \sigma^2]} \sim F_{\,t-1,\ t(n-1)}.$$

In particular, under H_0 (when $\sigma_a^2 = 0$), this simplifies to the test statistic

$$F = \frac{S_T/(t-1)}{S_R/\{t(n-1)\}} \sim F_{\,t-1,\ t(n-1)}.$$

Hence, at significance level α, one rejects H_0 if

$$F > F_{\,t-1,\ t(n-1);\ \alpha},$$

where $F_{\,\nu_1,\nu_2;\ \alpha}$ denotes the upper α–quantile of the F-distribution with ν_1 and ν_2 degrees of freedom.

Remark 9. This testing procedure coincides with the classical one-way ANOVA F-test for fixed effects. When H_0 is false (i.e. $\sigma_a^2 > 0$), the numerator mean square acquires a noncentral chi-square component, yielding a (scaled) noncentral Fdistribution.

6 Split-Plot Design

A **split-plot design** is a two-stage (or multi-stage) experimental design used when some factors are harder or more expensive to change than others. It was originally developed in agricultural experiments (e.g. for machine-controlled irrigation vs. hand-applied fertilizer), but now appears in many industrial and scientific settings.

6.1 Motivating Example: Corrosion Resistance

A split-plot experiment was conducted to assess the corrosion resistance (measured by yield) of steel bars under two factors whose levels differ markedly in how readily they can be randomized:

1. **Whole-plot factor:** Furnace temperature P at three levels,

   $$P \in \{360°C,\ 370°C,\ 380°C\}.$$

   Because heating the furnace to a new temperature requires a full day's operation, these treatments were applied to *whole plots* (i.e.\ each "day" of the experiment).

2. **Sub-plot factor:** Surface coating Q at four levels,

   $$Q \in \{C_1,\ C_2,\ C_3,\ C_4\}.$$

   Within each day's run, the four furnace positions (slots 1–4) received a random permutation of the four coatings; after the furnace reached its designated whole-plot temperature, bars in each slot were coated and then tested.

The experiment was replicated $b = 2$ times over six consecutive days (labelled D1–D6). Each replicate thus occupies three days (one per temperature level), and within each day the four coatings are randomized across the four positions. Table 26 summarizes the raw corrosion-resistance readings y_{ijk}, where $i = 1, 2$ indexes the replicate, $j = 1, 2, 3$ the temperature level, and $k = 1, \ldots,$ 4 the coating.

Day	Temp. (°C)	Slot 1	Slot 2	Slot 3	Slot 4
D1	360	73 (C_2)	83 (C_3)	67 (C_1)	89 (C_4)
D2	370	65 (C_1)	87 (C_3)	86 (C_4)	91 (C_2)
D3	380	147 (C_3)	155 (C_1)	127 (C_2)	212 (C_4)
D4	380	153 (C_4)	90 (C_3)	100 (C_2)	108 (C_1)
D5	370	150 (C_4)	140 (C_1)	121 (C_3)	142 (C_2)
D6	360	33 (C_1)	54 (C_4)	8 (C_2)	46 (C_3)

Table 26. Corrosion Resistance Data

Aggregating these observations yields the following cell means (two replicates per temperature × coating combination) and marginal means:

	C_1	C_2	C_3	C_4	Mean
360	67	73	83	89	78.00
370	65	91	87	86	82.25
380	155	127	147	212	160.25
360	33	8	46	54	35.25
370	140	142	121	150	138.25
380	108	100	90	153	112.75
Mean	94.67	90.17	95.67	124	101.125

Table 27.

6.2 Split-Plot Model

6.2.1 Model Formulation

Let

$$y_{ijk} = \text{yield for replicate } i = 1, 2, \text{ temperature } j = 1, 2, 3, \text{ coating } k = 1, \ldots, 4.$$

We decompose y_{ijk} into whole-plot and subplot components as follows:

1. **Whole-plot (day) model:**

$$m_{ij} = \mu + p_j + \epsilon^w_{ij},$$

where

- μ is the grand mean,
- p_j is the fixed main effect of temperature level j, with $\sum_j p_j = 0$,
- ϵ^w_{ij} is the whole-plot error for replicate i, day j, satisfying $\mathbb{E}[\epsilon^w_{ij}] = 0$ and $\mathrm{Var}(\epsilon^w_{ij}) = \sigma^2_w$.

2. **Sub-plot (coating) model within each day:**

$$y_{ijk} = m_{ij} + q_k + (pq)_{jk} + \epsilon^s_{ijk},$$

where

- q_k is the fixed main effect of coating k, $\sum_k q_k = 0$,
- $(pq)_{jk}$ is the fixed interaction between temperature j and coating k, subject to $\sum_j (pq)_{jk} = 0$ and $\sum_k (pq)_{jk} = 0$,

- ϵ^s_{ijk} is the subplot error with $\mathbb{E}[\epsilon^s_{ijk}] = 0$ and $\mathrm{Var}(\epsilon^s_{ijk}) = \sigma_s^2$, independent of $\epsilon^w_{i'j'}$.

Equivalently, the split-plot model can be written as

$$y_{ijk} = \underbrace{\mu + p_j + \epsilon^w_{ij}}_{\text{whole-plot (day) model}} + \underbrace{q_k + (pq)_{jk} + \epsilon^s_{ijk}}_{\text{subplot (coating) model}},$$

Because temperature assignments cannot be rerandomized more than once per day, the variability due to ϵ^w must be estimated from the variation among whole plots (days). Coating effects and their interactions are estimated from within-day comparisons and thus are tested against the finer-scale error σ_s^2. This split-plot structure both reflects the physical constraints of the furnace and provides efficient estimation of the subplot (coating) effects while properly accounting for the larger whole-plot variability.

6.2.2 ANOVA for split-plot design

Data Decomposition
Any observation may be decomposed into orthogonal components corresponding to overall mean, main effects, interaction, and residuals:

$$y_{\mathrm{ijk}} = y_{\cdots} + (y_{\cdot j\cdot} - y_{\cdots}) + (y_{\mathrm{ij}\cdot} - y_{\cdot j\cdot}) + (y_{\cdot\cdot k} - y_{\cdots}) + (y_{\cdot\mathrm{jk}} - y_{\cdot j\cdot} - y_{\cdot\cdot k} + y_{\cdots}) + (y_{\mathrm{ijk}} - y_{\mathrm{ij}\cdot} - y_{\cdot\mathrm{jk}} + y_{\cdot j\cdot}).$$

ANOVA Table
Let

$$\begin{aligned}
\mathrm{SS}_P &= b\,q\sum_{j=1}^{p}(y_{\cdot j\cdot} - y_{\cdots})^2, & \mathrm{df}_P &= p-1,\\
\mathrm{SS}_{\mathrm{WP}} &= q\sum_{i=1}^{b}\sum_{j=1}^{p}(y_{\mathrm{ij}\cdot} - y_{\cdot j\cdot})^2, & \mathrm{df}_{\mathrm{WP}} &= b\,p-p,\\
\mathrm{SS}_Q &= b\,p\sum_{k=1}^{q}(y_{\cdot\cdot k} - y_{\cdots})^2, & \mathrm{df}_Q &= q-1,\\
\mathrm{SS}_{\mathrm{PQ}} &= b\sum_{j=1}^{p}\sum_{k=1}^{q}(y_{\cdot\mathrm{jk}} - y_{\cdot j\cdot} - y_{\cdot\cdot k} + y_{\cdots})^2, & \mathrm{df}_{\mathrm{PQ}} &= (p-1)(q-1),\\
\mathrm{SS}_{\mathrm{SP}} &= \sum_{i,j,k}(y_{\mathrm{ijk}} - y_{\mathrm{ij}\cdot} - y_{\cdot\mathrm{jk}} + y_{\cdot j\cdot})^2, & \mathrm{df}_{\mathrm{SP}} &= b\,p\,(q-1),\\
\mathrm{SS}_{\mathrm{Total}} &= \sum_{i,j,k}(y_{\mathrm{ijk}} - y_{\cdots})^2, & \mathrm{df}_{\mathrm{Total}} &= b\,p\,q-1.
\end{aligned}$$

Then we have an ANOVA table

Source	SS	df	EMS
P	SS_P	$p-1$	$bq^{-1}\sigma_s^2 + q\,\sigma_w^2 + bq\,\dfrac{\sum_j p_j^2}{p-1}$
Whole-plot error	$\mathrm{SS}_{\mathrm{WP}}$	$p(b-1)$	$q\,\sigma_w^2 + \sigma_s^2$
Q	SS_Q	$q-1$	$\sigma_s^2 + bp\,\dfrac{\sum_k q_k^2}{q-1}$
$P \times Q$	$\mathrm{SS}_{\mathrm{PQ}}$	$(p-1)(q-1)$	$\sigma_s^2 + b\,\dfrac{\sum_{jk}(pq)_{jk}^2}{(p-1)(q-1)}$
Subplot error	$\mathrm{SS}_{\mathrm{SP}}$	$bp(q-1)$	σ_s^2
Total	$\mathrm{SS}_{\mathrm{Total}}$	$bpq-1$	—

Table 28. ANOVA Table of Split-plot Design

6.3 Split-plot with Blocks

Let

$$y_{ijk} = \text{response in block } i,\ \text{whole-plot treatment } j,\ \text{subplot treatment } k,$$

for $i = 1,\ldots,b,\ j = 1,\ldots,p,\ k = 1,\ldots,q$. Under the usual sum-to-zero constraints on all fixed effects, the mixed model is

$$y_{ijk} = \underbrace{\mu + b_i + p_j + \epsilon_{ij}^{(w)}}_{\substack{\text{grand mean } \mu,\\ \text{block effect } b_i,\\ \text{whole-plot (block}\times\text{P) interaction error } \epsilon_{ij}^{(w)} \sim (0,\sigma_w^2)}} + \underbrace{q_k + (pq)_{jk} + \epsilon_{ijk}^{(s)}}_{\substack{\text{subplot factor } q_k,\\ P\times Q \text{ interaction } (pq)_{jk},\\ \text{subplot error } \epsilon_{ijk}^{(s)} \sim (0,\sigma_s^2)}} .$$

Each observation decomposes into orthogonal contrasts:

$$\begin{aligned} y_{ijk} \ =& y_{...} + (y_{i..} - y_{...}) + (y_{.j.} - y_{...}) + (y_{ij.} - y_{i..} - y_{.j.} + y_{...}) \\ &+ (y_{..k} - y_{...}) + (y_{.jk} - y_{.j.} - y_{..k} + y_{...}) + (y_{ijk} - y_{ij.} - y_{.jk} + y_{.j.}), \end{aligned}$$

Define the whole-plot error sum of squares

$$\mathrm{SS}_{\mathrm{WP}} = \sum_{i=1}^{b} \sum_{j=1}^{p} (y_{\mathrm{ij}\cdot} - y_{i\cdot\cdot} - y_{\cdot j\cdot} + y_{\cdots})^2.$$

One shows by independence and zero-sum constraints that

$$\mathbb{E}[\mathrm{SS}_{\mathrm{WP}}] = (b-1)(p-1)\ (q^{-1}\sigma_s^2 + \sigma_w^2).$$

Thus the whole-plot error mean square satisfies

$$\mathbb{E}[\mathrm{MS}_{\mathrm{WP}}] = \frac{q}{(b-1)(p-1)}\ \mathbb{E}[\mathrm{SS}_{\mathrm{WP}}] \ = \ q\ \sigma_w^2 \ + \ \sigma_s^2,$$

exactly as in the block free split-plot design.

6.4 Split-Plot with Two Subplot Factors

Now suppose that within each whole-plot (block $\times$ P) we implement a full factorial in two subplot factors Q (with q levels) and R (with r levels). The observation index becomes y_{ijkl} for block i, whole-plot j, subplot levels $k$and l.

6.4.1 Model Specification

$$y_{ijkl} \ = \mu + p_j + \epsilon_{ij}^{(w)} + q_k + r_l + (qr)_{kl} + (pq)_{jk} + (pr)_{jl} + (pqr)_{jkl} + \epsilon_{ijkl}^{(s)},$$

with all effects subject to their appropriate sum-to-zero constraints. In particular, the three-factor interaction is defined by

$$(pqr)_{jkl} = \mu_{jkl} - (\mu + p_j + q_k + r_l + (pq)_{jk} + (pr)_{jl} + (qr)_{kl}),$$

and its unbiased estimate is

$$\widehat{(pqr)}_{jkl} = y_{\cdot jkl} - \{\ y_{\cdots\cdot} + (y_{\cdot j\cdot\cdot} - y_{\cdots\cdot}) + (y_{\cdot\cdot k\cdot} - y_{\cdots\cdot}) + (y_{\cdots l} - y_{\cdots\cdot})\}$$

$$-\{(y_{\cdot jk\cdot} - y_{\cdot j\cdot\cdot} - y_{\cdot\cdot k\cdot} + y_{\cdots\cdot}) + (y_{\cdot j\cdot l} - y_{\cdot j\cdot\cdot} - y_{\cdots l} + y_{\cdots\cdot}) + (y_{\cdot\cdot kl} - y_{\cdot\cdot k\cdot} - y_{\cdots l} + y_{\cdots\cdot})\}.$$

6.4.2 Full Orthogonal Decomposition

Analogous to the two-factor split-plot, one can write each y_{ijkl} as the sum of ten orthogonal components:

$$\begin{aligned} y_{ijkl} \;=& y_{\cdots} + (y_{\cdot j\cdot\cdot} - y_{\cdots}) + (y_{ij\cdot\cdot} - y_{\cdot j\cdot\cdot}) + (y_{\cdot\cdot k\cdot} - y_{\cdots}) + (y_{\cdots l} - y_{\cdots}) \\ &+ (y_{\cdot jk\cdot} - y_{\cdot j\cdot\cdot} - y_{\cdot\cdot k\cdot} + y_{\cdots}) + (y_{\cdot j\cdot l} - y_{\cdot j\cdot\cdot} - y_{\cdots l} + y_{\cdots}) \\ &+ (y_{\cdot\cdot \mathrm{kl}} - y_{\cdot\cdot k\cdot} - y_{\cdots l} + y_{\cdots}) + (y_{\cdot jkl} - \ldots + y_{\cdots}) \\ &+ (y_{ijkl} - y_{ij\cdot\cdot} - y_{\cdot jkl} + y_{\cdot j\cdot\cdot}), \end{aligned}$$

thus isolating the grand mean, three whole-plot terms, three two-factor interactions, the three-factor interaction, and the subplot error. This orthogonal decomposition underpins the construction of the corresponding ANOVA table and the derivation of each expected mean square for valid F-tests.

Source	**SS**	**DF**
P	$b\,q\,r\,\sum_j (\bar{y}_{\cdot j\cdot\cdot} - \bar{y}_{\cdots})^2$	$p-1$
Whole plot error	$q\,r\,\sum_{i,j} (\bar{y}_{ij\cdot\cdot} - \bar{y}_{\cdot j\cdot\cdot})^2$	$(b-1)\,p$
Whole plot total	$q\,r\,\sum_{i,j} (\bar{y}_{ij\cdot\cdot} - \bar{y}_{\cdots})^2$	$b\,p-1$
Q	$b\,p\,r\,\sum_k (\bar{y}_{\cdot\cdot k\cdot} - \bar{y}_{\cdots})^2$	$q-1$
R	$b\,p\,q\,\sum_l (\bar{y}_{\cdots l} - \bar{y}_{\cdots})^2$	$r-1$
$Q\,R$	—	$(q-1)\,(r-1)$
PQ	$b\,r\,\sum_{j,k} (\bar{y}_{\cdot jk\cdot} - \bar{y}_{\cdot j\cdot\cdot} - \bar{y}_{\cdot\cdot k\cdot} + \bar{y}_{\cdots})^2$	$(p-1)\,(q-1)$
PR	$b\,q\,\sum_{j,l} (\bar{y}_{\cdot j\cdot l} - \bar{y}_{\cdot j\cdot\cdot} - \bar{y}_{\cdots l} + \bar{y}_{\cdots})^2$	$(p-1)\,(r-1)$
$PQ\,R$	$b\,\sum_{j,k,l} (\bar{y}_{\cdot jkl} - \bar{y}_{\cdot jk\cdot} - \bar{y}_{\cdot j\cdot l} - \bar{y}_{\cdot\cdot kl} + \bar{y}_{\cdot j\cdot\cdot} + \bar{y}_{\cdot\cdot k\cdot} + \bar{y}_{\cdots l} - \bar{y}_{\cdots})^2$	$(p-1)\,(q-1)\,(r-1)$
Subplot error	$\sum_{i,j,k} (y_{ijkl} - \bar{y}_{ij\cdot\cdot} - \bar{y}_{\cdot jkl} + \bar{y}_{\cdot j\cdot\cdot})^2$	$(b-1)\,p\,(q\,r-1)$
Total	$\sum_{i,j,k} (y_{ijkl} - \bar{y}_{\cdots})^2$	$b\,p\,q\,r-1$

Table 29. ANOVA table for the split-plot experiment

6.5 Final Remarks on Split-Plot Design

Advantages

- **Enhanced subplot precision:** Subplot-factor comparisons use the smaller subplot-error MS, yielding more precise inferences than a fully randomized design.
- **Practical logistics:** Accommodates hard-to-change whole-plot treatments without complete rerandomization.

Disadvantages

1. **Fewer error df:** Two variance components (σ_w^2, σ_s^2) consume more degrees of freedom than a single residual term.
2. **Uneven precision trade-off:** Because whole-plot df $\ll$ subplot df, the loss in whole-plot precision often outweighs the gain in subplot precision.
3. **Reduced power for some interactions:** Tests involving whole-plot factors have less error df than in a completely randomized design.
4. **Multiple SEs required:** Up to four distinct MS estimates complicate graphical or tabular comparisons.
5. **Complex inference:** Exact distributions for whole-plot MS are not closed-form, making some tests and CIs more involved.

Take-Away
Split-plot designs suit experiments constrained by hard-to-change factors but sacrifice uniform precision. When all treatment contrasts require similar accuracy, alternative factorial layouts typically outperform.

7 Incomplete Block Designs

We wish to compare six treatments A, B, C, D, E, F using two blocks (indexed by $i = 1, 2$), each of size 5. The allocation is

$$\begin{array}{ll} \text{Block 1:} & A,\ C,\ D,\ E,\ F, \\ \text{Block 2:} & B,\ C,\ D,\ E,\ F. \end{array}$$

Let σ^2 denote the common within-block error variance. We denote by Y_{ij} the observed yield for treatment $j \in \{A, \ldots, F\}$ in block i, except that the pair $(i, j) = (1, B)$ and $(2, A)$ are unobserved by design.

7.1 Direct Comparisons

Comparisons among the treatments in both blocks. For any two treatments $j, k \in \{C, D, E, F\}$, each appears once in each block. Hence an unbiased estimator of the treatment difference $t_j - t_k$ is

$$\frac{(Y_{1j} - Y_{1k}) + (Y_{2j} - Y_{2k})}{2},$$

which has variance

$$\mathrm{Var}\big[\tfrac{1}{2}((Y_{1j} - Y_{1k}) + (Y_{2j} - Y_{2k}))\big] = \sigma^2.$$

Comparison of A versus B. Because A appears only in Block 1 and B only in Block 2, we must adjust for the block effect. Define the block-effect estimator

$$\hat{b} = \underbrace{\frac{1}{4} \sum_{j \in \{C,D,E,F\}} Y_{1j}}_{Y_{1\cdot}} - \frac{1}{4} \sum_{j \in \{C,D,E,F\}} Y_{2j},$$

which satisfies $\mathrm{Var}(\hat{b}) = \frac{\sigma^2}{2}$. Then an unbiased estimator of $t_A - t_B$ is

$$(Y_{1A} - Y_{2B}) \ - \ \hat{b},$$

with variance

$$\mathrm{Var}(Y_{1A} - Y_{2B} - \hat{b}) = 2\ \sigma^2 + \frac{\sigma^2}{2} = \frac{5\ \sigma^2}{2}.$$

Comparison of A versus C. Here C is observed in both blocks but A only in Block 1. One may form the estimator

$$Y_{1A} \ - \ \frac{Y_{1C} + Y_{2C} + \hat{b}}{2},$$

which is unbiased for $t_A - t_C$ and has variance

$$\mathrm{Var}\big[Y_{1A} - \tfrac{1}{2}(Y_{1C} + Y_{2C} + \hat{b})\big] = \sigma^2 + \frac{\sigma^2}{2} = \frac{3\ \sigma^2}{2}.$$

(See, however, the least-squares solution below for the exact efficiency factor.)

7.2 Least-Squares Analysis

We fit the standard one-way two-block model with treatment effects and block effects:

$$Y_{ij} = \mu + b_i + t_j + \epsilon_{ij}, \quad i = 1, 2,\ j \in \{A, \ldots, F\},$$

where $\epsilon_{ij} \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$, subject to the constraints

$$b_1 + b_2 = 0, \quad \sum_j t_j = 0.$$

Omitting the two empty cells $(1, B)$ and $(2, A)$, the normal-equations yield the following estimates.

Estimate of $t_A - t_B$.

$$\begin{aligned}\hat{t}_A - \hat{t}_B &= \frac{5}{4}\left[(Y_{1A} - Y_{2B}) - (Y_{1\cdot} - Y_{2\cdot})\right] \\ &= \frac{1}{4}\left(4\,Y_{1A} \;-\; 4\,Y_{2B} \;-\; \sum_{j \in \{C,D,E,F\}} Y_{1j} \;+\; \sum_{j \in \{C,D,E,F\}} Y_{2j}\right),\end{aligned}$$

with

$$\mathrm{Var}(\hat{t}_A - \hat{t}_B) = \frac{5\ \sigma^2}{2}.$$

Estimate of $t_A - t_C$.

$$\begin{aligned}\hat{t}_A - \hat{t}_C &= \frac{1}{8}\left[9\,Y_{1A} - Y_{2B} - 8\,Y_{\cdot C} - 5\,(Y_{1\cdot} - Y_{2\cdot})\right] \\ &= \frac{1}{8}\left(8\,Y_{1A} - \,5\,Y_{1C} - \,3\,Y_{2C} - \sum_{j \in \{D,E,F\}} (Y_{1j} - Y_{2j})\right),\end{aligned}$$

where $Y_{\cdot C} = Y_{1C} + Y_{2C}$ and $Y_{i\bullet} = \sum_j Y_{ij}$. Its variance is

$$\mathrm{Var}(\hat{t}_A - \hat{t}_C) = \frac{\sigma^2}{8^2}\,(8^2 + 1^2 + 5^2 + 3^2 + 6) = \frac{13\ \sigma^2}{8}.$$

Here the "6" arises from the two missing observations (one each in Blocks 1 and 2).

These formulas quantify the relative efficiencies of direct versus least-squares estimators in an incomplete-block setting. The least-squares approach automatically accounts for the unequal replication of A and B across blocks and yields the correct variance factors.

7.3 Balanced Incomplete Block Design

Let v be the number of treatments, b be the number of blocks, r be the replicate of each treatment, $k < v$ be the unites per block. A **balanced incomplete-block design** (BIBD) with parameters (v, b, r, k, λ), where λ means each pair of treatments occurs together in λ blocks, is a way of arranging v treatments into b blocks so that

1. Each block contains exactly k distinct treatments (so "incomplete" means $k < v$).

2. Each treatment appears in exactly r blocks.

3. Every unordered pair of distinct treatments occurs together in exactly λ blocks (the "balanced" condition).

These parameters are not independent but must satisfy two counting identities:

1. **Total-unit count**
   Counting "treatment-in-block" incidences in two ways gives

   $$v\ r = b\ k.$$

   On the left is "each of v treatments appears in r blocks," for a total of $v\ r$ incidences; on the right is "each of $b$blocks contains k treatments," for a total of $b\ k$.

2. **Pair-co-occurrence count**
   Fix any one treatment T. It appears in r blocks; in each of those blocks it co-occurs with $(k-1)$ other treatments, giving $r\ (k-1)$ ordered "T–other" incidences. On the other hand, each of the other $(v-1)$treatments co-occurs with T in exactly λ blocks, for a total of $\lambda\ (v-1)$ "T–other" incidences. Equating,

   $$r\ (k-1) = \lambda\ (v-1).$$

   Solving for λ yields the standard formula

   $$\lambda = \frac{r\ (k-1)}{v-1}.$$

Properties of a Balanced Incomplete Block Design

- **Estimability of treatment contrasts.**
  Every linear contrast among the treatment effects can be uniquely estimated under the standard BIBD model.

- **Uniform precision of pairwise comparisons.**
  The variance of the estimator for any difference between two treatments is the same for all $\binom{t}{2}$ pairs.

- **Minimized average confidence-interval length.**
  Over all pairwise contrasts, a BIBD tends to yield the shortest average length of $(1-\alpha)\%$-confidence intervals.

Limitation

- **Existence constraints.**
  Feasible BIBDs occur only for specific integer solutions (t, b, r, k, λ); for many parameter combinations no design exists.

Example 11. (BIBD(6, 6, 5, 5, 4)) 6 treatments with 30 units grouped in 6 blocks of 5 units.Below is the simplest balanced incomplete-block design (a BIBD) that meets your requirements:

- **Treatments**: $v = 6$ (call them A, B, C, D, E, F)
- **Blocks**: $b = 6$, each of size $k = 5$
- **Replications**: each treatment appears $r = 5$ times
- **Each pair of treatments occurs together in**

$$\lambda = \frac{r\ (k-1)}{v-1} = \frac{5 \cdot (5-1)}{6-1} = 4$$

blocks.

Because $vr = bk$ $(6 \cdot 5 = 6 \cdot 5)$, such a BIBD exists and is in fact the *complement* of the trivial one-treatment-per-block design: simply omit each treatment exactly once. Concretely, number your blocks $i = 1, \ldots, 6$, and let Block i contain all treatments except treatment i. In alphabetical labels:

Block	Treatments included	Omitted
1	B, C, D, E, F	A
2	A, C, D, E, F	B
3	A, B, D, E, F	C
4	A, B, C, E, F	D
5	A, B, C, D, F	E
6	A, B, C, D, E	F

Table 30. BIBD$(6, 6, 5, 5, 4)$

Example 12. (Incomplete–block design with $v = 6$, $b = 6$, $k = 4$, $r = 4$)

Six treatments $\{A, \ldots, F\}$, six blocks of size 4, each treatment in $r = 4$ blocks (omitted twice), never repeated within a block. A BIBD would require

$$\lambda = \frac{r\ (k-1)}{v-1} = \frac{4 \cdot 3}{5} = \frac{12}{5},$$

which is not an integer. Since λ must be a whole number, **no BIBD can exist for these parameters.**

General variance formula. For any contrast $\mathbf{c}'\hat{\beta}$,

$$\mathrm{Var}(\mathbf{c}'\hat{\beta}) = \sigma^2\, \mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}.$$

Design I (Missing pairs: $(A\,B), (A\,C), (B\,C), (D\,E), (D\,F), (E\,F)$).

- Notice this "splits" the six treatments into two triangles $\{A, B, C\}$ and $\{D, E, F\}$.
- As a result, for the three "within-triangle" pairs

  AB, AC, BC and DE, DF, EF

  each pair *never* appears together in exactly two blocks (those two in which the pair is omitted), and *always* appears in the other four. This gives them higher precision.
- The remaining nine pairs (one from each triangle across triangles, e.g. A vs. D, B vs.E, etc.) each co-occur in only three blocks, so are estimated with slightly larger SD.

Contrast	SD (Design I)
AB, AC, BC, DE, DF, EF	0.7303 σ
the other 9 pairs	0.7601 σ

Table 31. Variance of design I

Design II (Missing pairs: $(A\,B), (A\,C), (B\,D), (C\,E), (D\,F), (E\,F)$).

- This pattern "mixes" treatments more evenly across blocks.

- Consequently, six of the contrasts attain the best precision, another six are slightly worse, and the remaining three are just a hair larger.

Contrast	SD (Design II)
AB, AC, BD, CE, DF, EF	0.7321 σ
AD, BF, DE, FC, EA, CB	0.7579 σ
AF, BE, DC	0.7596 σ

Table 32. Variance of design II

Comparison. Design I gives maximum efficiency for its six primary contrasts; Design II spreads precision more evenly. Choose according to which pairs are of greatest interest.

Remark 10. Note if the observations were obtained from 4 complete blocks of size 6, the standard deviation would be $\sigma/\sqrt{2} \approx 0.707\sigma$

Switching from 4 RCBs of size 6 to 6 blocks of size 4 incurs only a small efficiency penalty (6–15%) in SD of treatment differences, while offering potential gains in reducing σ via more homogeneous blocking and greater logistical flexibility. The choice of how to omit treatments—Design I versus Design II—lets you "tune" which pairs receive the best precision.

Example 13. (Unbalanced Design) Construct an incomplete–block arrangement for six treatments A, B, C, D, E, F on $N = 25$ experimental units, grouped into $b = 5$ blocks of size $k = 5$. Each block omits exactly one treatment, so one treatment must appear in all blocks. Label the blocks $1, \ldots,$ 5, and choose A to be the "always–included" treatment. A convenient allocation is:

Block	I	II	III	IV	V
Treatments	A	A	A	A	A
	B	B	B	B	C
	C	C	D	D	D
	D	D	E	E	E
	E	F	F	F	F

Table 33. Unblanced Design

- $v = 6$ treatments, $b = 5$ blocks, each of size $k = 5$.
- Treatment A occurs in all 5 blocks; each of B, C, D, E, F occurs in exactly 4 blocks.
- Each block omits a different treatment, so that no block repeats.
- This design is *not* balanced since the replication numbers differ, but it may be useful when one treatment must serve as a control (appearing in every block) and each other treatment must be compared to it under as homogeneous block conditions as possible.

Example 14. (Unequal-Block Design)We have

- **Treatments**: A, B, C, D, E, F, G $(v = 7)$.
- **Total units**: 35.
- **Blocks**: 5 blocks of sizes $5, 6, 7, 7, 10$.
- **Replication**: each treatment must appear $r = 35/7 = 5$ times.

Because block V holds $10 > 7$ units, some treatments must be repeated there. One convenient allocation is:

Block	Size	Multiset of treatments
I	5	$\{A, B, C, D, E\}$
II	6	$\{A, B, C, D, F, G\}$
III	7	$\{A, B, C, D, E, F, G\}$
IV	7	$\{A, B, C, D, E, F, G\}$
V	10	$\{A, B, C, D, E, E, F, F, G, G\}$

Table 34. Unequal Block Design

- **Within-block replication** (in block V) is a legitimate way to attain the required overall replication when $k_i > v$.

- **Analysis** is still a simple two-way ANOVA with blocks and treatments; just build the design matrix X according to the above multiplicities.

This design thus achieves a perfectly balanced replication $r = 5$ for each treatment while respecting the pre-specified block-size pattern $(5, 6, 7, 7, 10)$.

7.4 Cyclic Balanced Incomplete Block Designs

- **Definition.**
  A *cyclic design* is a BIBD in which every block can be obtained by applying a fixed cyclic permutation (rotation) to the treatments in an initial "starter" block.

- **Construction procedure.**

  1. **Select an initial block** of size k from the t treatments.

  2. **Generate subsequent blocks** by cyclically rotating the labels of the treatments in the starter block through all t positions.

  3. **Validate balance.** The starter block must be chosen so that each unordered pair of treatments appears exactly λ times across the b rotated blocks.

- **Illustrative example ($t = b = 7,\ k = r = 3,\ \lambda = 1$).**

  1. **Inferior choice:** If the starter block is $(A\ B\ C)$, cyclic rotation yields

     $$(\mathrm{ABC}),\ (\mathrm{BCD}),\ (\mathrm{CDE}),\ (\mathrm{DEF}),\ (\mathrm{EFG}),\ (\mathrm{FGA}),\ (\mathrm{GAB}).$$

     In this arrangement some treatment-pairs occur multiple times while others occur only once, violating the balance requirement.

  2. **Valid BIBD starter:** If instead one uses $(A\ B\ D)$, cyclic rotation produces

     $$(\mathrm{ABD}),\ (\mathrm{BCE}),\ (\mathrm{CDF}),\ (\mathrm{DEG}),\ (\mathrm{EFA}),\ (\mathrm{FGB}),\ (\mathrm{GAC}),$$

     which satisfies

     $$t = 7, b = 7, k = 3, r = 3, \lambda = 1,$$

     and hence constitutes a proper cyclic BIBD.

7.5 Latin Squres

A **Latin-square design** provides a way to control two orthogonal nuisance factors simultaneously—commonly called "row" and "column" blocks—while comparing t treatments. Its key features are:

1. **Layout and blocking structure**
   - There are t^2 experimental units arranged in a $t \times t$ grid.
   - **Row blocks:** t horizontal strips, each containing one of each treatment.
   - **Column blocks:** t vertical strips, each also containing one of each treatment.
   - Thus each treatment appears exactly once in every row and once in every column.
2. **Connection with randomized-block designs**
   - Viewed marginally, the rows themselves form a randomized-complete-block design (RCBD) for the $t$treatments (one replicate per row).
   - Likewise, the columns form an independent RCBD.
3. **Canonical example ($t = 4$)**

$$\begin{array}{cccc} A & B & C & D \\ B & D & A & C \\ C & A & D & B \\ D & C & B & A \end{array}$$

   In this 4×4 Latin square, each of the treatments A, B, C, D appears exactly once in each row and each column.
4. **Statistical model**
   For the response y_{ij} observed in row i and column j, let $k(i, j)$ denote the treatment assigned there. The usual additive model is

$$y_{ij} = \mu + r_i + c_j + \tau_{k(i,j)} + \varepsilon_{ij},$$

   subject to the usual side-conditions
   $\sum_{i=1}^{t} r_i = 0,\ \sum_{j=1}^{t} c_j = 0,\ \sum_{h=1}^{t} \tau_h = 0,$ and $\varepsilon_{ij} \overset{\text{iid}}{\sim} N(0, \sigma^2)$.
5. **Analysis and degrees of freedom**
   - **Total:** $t^2 - 1$
   - **Row effects:** $t - 1$
   - **Column effects:** $t - 1$
   - **Treatment effects:** $t - 1$
   - **Error:** $(t - 1)(t - 2)$
     One then performs an ANOVA splitting sums of squares accordingly, testing treatment differences against the mean-square error.
6. **Advantages and limitations**
   - **Advantages:**
     - Controls two blocking factors simultaneously, thus reducing experimental error when both sources are important.

- Requires only one replicate per treatment (efficient in unit usage).

- **Limitations:**
  - Requires exactly t^2 units and equal block sizes.
  - Assumes no interaction among treatments and the two blocking factors.
  - Suitable only when both row and column blocks are genuinely irrelevant to treatment ordering (no ordinal or carry-over effects).

A Plackett–Burman Table

Variable	12	16	20	24	32
1	+	+	+	+	-
2	+	-	+	+	-
3	-	-	-	+	-
4	+	-	-	+	-
5	+	+	+	+	+
6	+	-	+	-	-
7	-	-	+	+	+
8	-	+	+	-	-
9	-	+	-	+	+
10	+	-	+	+	+
11	-	+	-	-	+
12	·	-	+	-	-
13	·	+	-	+	+
14	·	+	-	+	+
15	·	+	-	-	-
16	·	·	-	-	-
17	·	·	+	+	-
18	·	·	+	-	+
19	·	·	-	+	+
20	·	·	·	-	+
21	·	·	·	-	+
22	·	·	·	-	+
23	·	·	·	-	-
24	·	·	·	·	-
25	·	·	·	·	+
26	·	·	·	·	+
27	·	·	·	·	-
28	·	·	·	·	+
29	·	·	·	·	-
30	·	·	·	·	-
31	·	·	·	·	+

Table 35. Plackett–Burman Table