ST 5203: Experimental Design

(Semester 1, AY 2017/2018)

Text book: *Experiments: Planning, Analysis, and Optimization (2nd. edition)*

by Jeff Wu and Mike Hamada

Topic 4: Factorial Experiments at Two Levels

- Basic concepts and notations for 2^k experiments
- Estimation of effects and plots
- Fundamental principles in factorial experiments
- Estimation methods in factorial experiments
- Statistical inference of effects
- Comparisons with "one factor at a time" approach

Epitaxial Layer Growth Experiment

- Four factors each at two levels; 6 replicates for each of the 16 ($=2^4$) level combinations.

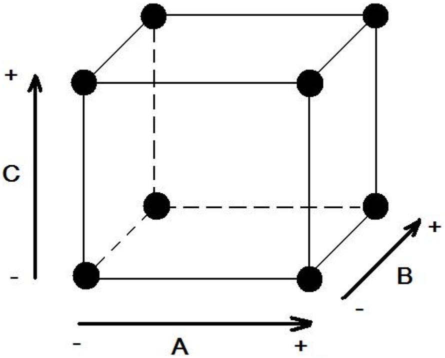
	Factor	Level +	Level −
A.	susceptor rotation method	continuous	oscillating
B.	nozzle position	2	6
C.	deposition temperature °C	1210	1220
D.	deposition time	low	high

- Response y: layer thickness (μm)
- Objective: Find the best factor level combinations, to either maximize the response y, or reduce variation of y around its target 14.5 μm.

2^k Experiments

- Suppose we have k factors, each at 2 levels, denoted by "$-$" and "$+$". Conventionally, $+$ represents the high level. $-$ represents the low level.
- The uppercase letters $A, B, C, \ldots$ represent factors, main effects and interactions. e.g. ABC denotes the three interaction of factors A, B and C.
- Use combinations of lowercase letters $a, b, c, \ldots$ to denote each run in the experiment. The lowercase letter is absent if the factor is at level $-$. Use "(1)" to denote the run such that all factors are at level $-$. For example, ac stands for the following run: A at $+$, B at $-$, C at $+$.

Cuboidal Representation of 2^3 Experiment



Cuboidal Representation of a 2^3 Design

Example: Truck Leaf Spring Manufacture

- Purpose of the experiment: investigate the effects of factors on a manufacturing process for leaf springs used on trucks.
- Three factors each at two levels are considered. (1). Furnace temperature (A); (2). Heating time in the furnace (B); (3). Transfer time between the furnace and the forming machine (C).
- The experimental design and outcomes are listed in the following table.

Design and Outcomes of Truck Leaf Spring Manufacture Experiment

Table 1: Design and Outcomes

Treatment	A	B	C	y	Furnace Temp. (°F)	Heating Time (sec)	Transfer Time (sec)
(1)	$-$	$-$	$-$	32	1840	23	10
a	$+$	$-$	$-$	35	1880	23	10
b	$-$	$+$	$-$	28	1840	25	10
ab	$+$	$+$	$-$	31	1880	25	10
c	$-$	$-$	$+$	48	1840	23	12
ac	$+$	$-$	$+$	39	1880	23	12
bc	$-$	$+$	$+$	28	1840	25	12
abc	$+$	$+$	$+$	29	1880	25	12

Clarification on Notations

- The uppercase letters ($A, B, AB, \ldots$) can have 3 meanings:
  - They denote the treatment factors.
  - They denote the main effects (A, B, C) and their interactions (AB, AC, BC, ABC).
  - They denote the contrast vectors in Table 1. For example, A represents the contrast vector $(-1, 1, -1, 1, -1, 1, -1, 1)^\top$.

- The lowercase letters ($a, b, ab, \ldots$) can have 2 meanings:
  - They denote the levels. The presence of a lower case letter (a) means level $+$ (for factor A); otherwise it means $-$ (for factor A).
  - They denote the group means. For example, a denotes the group mean of the treatment combination $A = \text{high}, B = \text{low}, C = \text{low}$; ab denotes the group mean of the treatment combination $A = \text{high}, B = \text{high}, C = \text{low}$

Abuse of notations: What they mean exactly can only be understood from the context.

Clarification on Notations (Cont.)

- We also use $\mu(\cdot)$ to denote the true group means. For example, in a 2^3 design, $\mu(A+, B+, C-)$ is the mean with $A = \text{high}, B = \text{high}, C = \text{low}$. So $\mu(A+, B+, C-) = ab$. $\mu(A+)$ is the mean with $A = \text{high}$. How do we express $\mu(A+)$ in terms of lowercase letters? (See next slide.)

- We use $\bar{y}.$ to denote the group sample averages. For example, $\bar{y}_{ab}$ is the sample average in the group with $A = \text{high}, B = \text{high}, C = \text{low}$.

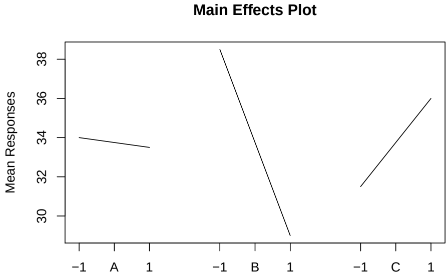
Main Effects and Plots

- Main effect of factor A:

$$A = ME(A) = \text{“Mean y at } A+\text{”} - \text{“Mean y at } A-\text{”}$$
$$= \mu(A+) - \mu(A-)$$

- For a 2^3 factorial experiment

$$A = \frac{abc + ab + ac + a}{4} - \frac{(1) + b + c + bc}{4}$$



Interaction Effects

- Conditional main effect of B when A at level "+" is

$$\begin{aligned} ME(B|A+) &= \text{“Mean y at } B+, A+\text{”} - \text{“Mean y at } B-, A+\text{”} \\ &= \mu(B+, A+) - \mu(B-, A+) \end{aligned}$$

- Similarly, we can define $ME(B|A-)$.
- Two-factor interaction effect between A and B (explain):

$$\begin{aligned} AB &= INT(A, B) = \frac{1}{2}\Big\{ME(B|A+) - ME(B|A-)\Big\} \\ &= \frac{1}{2}\Big\{ME(A|B+) - ME(A|B-)\Big\} \\ &= \frac{1}{2}\{\mu(A+, B+) + \mu(A-, B-)\} \\ &\quad -\frac{1}{2}\{\mu(A+, B-) + \mu(A-, B+)\} \end{aligned}$$

Interaction Effects for 2^3 Experiment

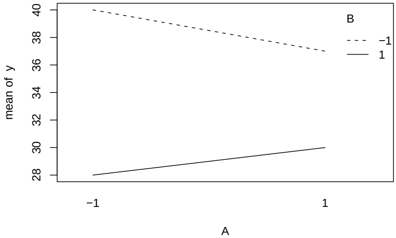
- For a 2^3 factorial experiment

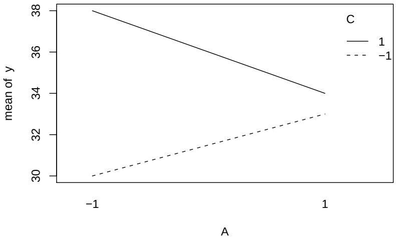
$$ME(A|B-) = \frac{ac + a}{2} - \frac{c + (1)}{2}$$

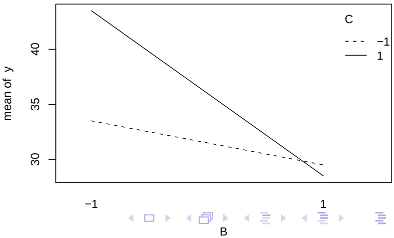
- AB interaction is

$$AB = \frac{abc + ab - bc - b}{4} - \frac{ac + a - c - (1)}{4}$$

Interaction Effect Plots







A General Formula for Interaction Effects

- For a general interaction (explain):

$$
\begin{aligned}
& A_1 A_2 \ldots A_k = INT(A_1 A_2 \ldots A_k) \\
= \; & \frac{1}{2} INT(A_1 A_2 \ldots A_{k-1} | A_k +) - \frac{1}{2} INT(A_1 A_2 \ldots A_{k-1} | A_k -)
\end{aligned}
$$

Planning Matrix and Model Matrix

- The planning matrix is the matrix composed of main effect columns in model matrix. See in the next slide, columns A, B and C together are the planning matrix of the experiment which determines the whole experiment.
- The model matrix is composed of 1 intercept column, k main effect columns and $2^k - k - 1$ interaction columns. Note that each interaction column can be constructed by the products of according main effect columns (explain).
- Any two columns in the model matrix are orthogonal.
- Any effect (main or interaction) can be estimated by the inner product of the corresponding column and the data column then divided by 2^{k-1}.

Planning Matrix and Model Matrix for a 2^3 Experiment

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

Regression Analysis of 2^k Factorial Experiments

- Consider the 2^3 design with factors A, B and C, whose columns are denoted by $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$ $(= 1$ or $-1)$. The AB, AC, BC and ABC columns are then equal to

$$\mathbf{x}_4 = \mathbf{x}_1\mathbf{x}_2, \mathbf{x}_5 = \mathbf{x}_1\mathbf{x}_3, \mathbf{x}_6 = \mathbf{x}_2\mathbf{x}_3, \mathbf{x}_7 = \mathbf{x}_1\mathbf{x}_2\mathbf{x}_3.$$

  The products above are component-wise.
- Consider the regression model:

$$\begin{aligned} y_i &= \mu + \alpha x_{i,1} + \beta x_{i,2} + \gamma x_{i,3} \\ &\quad + (\alpha\beta) x_{i,4} + (\alpha\gamma) x_{i,5} + (\beta\gamma) x_{i,6} \\ &\quad + (\alpha\beta\gamma) x_{i,7} + e_i \end{aligned}$$

- **Note**: In fact, the model above is the regression representation of three-way layout experiment (3 factors, each at two levels).

A General Statistical Model for 2^k Full Factorial Design

- For a general 2^k full factorial design with n replicates, the statistical model is

$$\bar{y} = \beta_0 + \sum_j \beta_j x_j + \sum_{j,\ell} \beta_{j\ell} x_j x_\ell + \sum_{j,\ell,m} \beta_{j\ell m} x_j x_\ell x_m + \dots$$
$$+ \beta_{12\dots,k} x_1 x_2 \dots x_k + \bar{e},$$

  where $\bar{e} \sim N(0, \sigma^2/n)$.
- This is a model about the group averages $\bar{y}$. The error term $\bar{e}$ is already averaged over n replicates in each treatment combination.

It can be shown that statistical inference based on this model of group averages is equivalent to the inference based on the raw data.

General Formula for Least Square Estimators

(Derivation on board.)

- For 2^k design, the general formulas of least square estimators are

$$\hat{\beta}_0 = \frac{\sum_{i=1}^{2^k} \bar{y}_i}{2^k} = \bar{\bar{y}}, \quad \hat{\beta}_j = \frac{\sum_{i=1}^{2^k} c_{ij}\bar{y}_i}{2^k},$$

where $\mathbf{c}_j = (c_{1j}, \ldots, c_{2^k j})^\top$ is the contrast vector for the jth effect.

- $\mathrm{Cov}(\hat{\boldsymbol{\beta}}) = \dfrac{\sigma^2}{2^k n}\mathbf{I}$, where $\mathbf{I}$ is the $2^k \times 2^k$ identity matrix. Let $N = 2^k n$ (total sample size). This implies that for all estimated coefficients $\hat{\beta}_j$,

$$\mathrm{Var}(\hat{\beta}_j) = \frac{\sigma^2}{N}.$$

Least Square Estimators of 2^3 Design

$$\hat{\beta}_0 = \frac{1}{8}(\bar{y}_{(1)} + \bar{y}_a + \bar{y}_b + \bar{y}_{ab} + \bar{y}_c + \bar{y}_{ac} + \bar{y}_{bc} + \bar{y}_{abc}),$$

$$\hat{\beta}_1 = \frac{1}{8}(-\bar{y}_{(1)} + \bar{y}_a - \bar{y}_b + \bar{y}_{ab} - \bar{y}_c + \bar{y}_{ac} - \bar{y}_{bc} + \bar{y}_{abc}),$$

$$\hat{\beta}_2 = \frac{1}{8}(-\bar{y}_{(1)} - \bar{y}_a + \bar{y}_b + \bar{y}_{ab} - \bar{y}_c - \bar{y}_{ac} + \bar{y}_{bc} + \bar{y}_{abc}),$$

$$\hat{\beta}_3 = \frac{1}{8}(-\bar{y}_{(1)} - \bar{y}_a - \bar{y}_b - \bar{y}_{ab} + \bar{y}_c + \bar{y}_{ac} + \bar{y}_{bc} + \bar{y}_{abc}),$$

$$\hat{\beta}_{12} = \frac{1}{8}(\bar{y}_{(1)} - \bar{y}_a - \bar{y}_b + \bar{y}_{ab} + \bar{y}_c - \bar{y}_{ac} - \bar{y}_{bc} + \bar{y}_{abc}),$$

$$\hat{\beta}_{13} = \frac{1}{8}(\bar{y}_{(1)} - \bar{y}_a + \bar{y}_b - \bar{y}_{ab} - \bar{y}_c + \bar{y}_{ac} - \bar{y}_{bc} + \bar{y}_{abc}),$$

$$\hat{\beta}_{23} = \frac{1}{8}(\bar{y}_{(1)} + \bar{y}_a - \bar{y}_b - \bar{y}_{ab} - \bar{y}_c - \bar{y}_{ac} + \bar{y}_{bc} + \bar{y}_{abc}),$$

$$\hat{\beta}_{123} = \frac{1}{8}(-\bar{y}_{(1)} + \bar{y}_a + \bar{y}_b - \bar{y}_{ab} + \bar{y}_c - \bar{y}_{ac} - \bar{y}_{bc} + \bar{y}_{abc}).$$

Treatment Means as Linear Combinations of Effects

$$
\begin{aligned}
(1) &= \mu - \frac{A}{2} - \frac{B}{2} - \frac{C}{2} + \frac{AB}{2} + \frac{AC}{2} + \frac{BC}{2} - \frac{ABC}{2},\\
a &= \mu + \frac{A}{2} - \frac{B}{2} - \frac{C}{2} - \frac{AB}{2} - \frac{AC}{2} + \frac{BC}{2} + \frac{ABC}{2},\\
b &= \mu - \frac{A}{2} + \frac{B}{2} - \frac{C}{2} - \frac{AB}{2} + \frac{AC}{2} - \frac{BC}{2} + \frac{ABC}{2},\\
ab &= \mu + \frac{A}{2} + \frac{B}{2} - \frac{C}{2} + \frac{AB}{2} - \frac{AC}{2} - \frac{BC}{2} - \frac{ABC}{2},\\
c &= \mu - \frac{A}{2} - \frac{B}{2} + \frac{C}{2} + \frac{AB}{2} - \frac{AC}{2} - \frac{BC}{2} + \frac{ABC}{2},\\
ac &= \mu + \frac{A}{2} - \frac{B}{2} + \frac{C}{2} - \frac{AB}{2} + \frac{AC}{2} - \frac{BC}{2} - \frac{ABC}{2},\\
bc &= \mu - \frac{A}{2} + \frac{B}{2} + \frac{C}{2} - \frac{AB}{2} - \frac{AC}{2} + \frac{BC}{2} - \frac{ABC}{2},\\
abc &= \mu + \frac{A}{2} + \frac{B}{2} + \frac{C}{2} + \frac{AB}{2} + \frac{AC}{2} + \frac{BC}{2} + \frac{ABC}{2}.
\end{aligned}
$$

Effects as Linear Combinations of Treatment Means

$$
\begin{aligned}
\mu &= \frac{1}{8}\left[(1) + a + b + ab + c + ac + bc + abc\right], \\
A &= \frac{1}{4}\left[-(1) + a - b + ab - c + ac - bc + abc\right], \\
B &= \frac{1}{4}\left[-(1) - a + b + ab - c - ac + bc + abc\right], \\
C &= \frac{1}{4}\left[-(1) - a - b - ab + c + ac + bc + abc\right], \\
AB &= \frac{1}{4}\left[(1) - a - b + ab + c - ac - bc + abc\right], \\
AC &= \frac{1}{4}\left[(1) - a + b - ab - c + ac - bc + abc\right], \\
BC &= \frac{1}{4}\left[(1) + a - b - ab - c - ac + bc + abc\right], \\
ABC &= \frac{1}{4}\left[-(1) + a + b - ab + c - ac - bc + abc\right].
\end{aligned}
$$

Estimated Effects in General 2^k Full Factorial Design

- In a 2^3 full factorial design with n replicates, the estimated effects are

$$\widehat{\mu} = \hat{\beta}_0, \quad \widehat{A} = 2\hat{\beta}_1, \quad \widehat{B} = 2\hat{\beta}_2, \quad \widehat{C} = 2\hat{\beta}_3,$$
$$\widehat{AB} = 2\hat{\beta}_{12}, \quad \widehat{AC} = 2\hat{\beta}_{13}, \quad \widehat{BC} = 2\hat{\beta}_{23}, \quad \widehat{ABC} = 2\hat{\beta}_{123}.$$

- In a general 2^k full factorial design, the estimated effects are always twice of the corresponding estimated coefficients. The estimated grand mean is equal to the estimated β_0.

ANOVA Table of 2^3 Full Factorial Design

Source	D.F.	Sum of Squares
A	1	SS_A
B	1	SS_B
C	1	SS_C
AB	1	SS_{AB}
AC	1	SS_{AC}
BC	1	SS_{BC}
ABC	1	SS_{ABC}
Error	$2^3(n-1)$	SSE
Total	$2^3 n - 1$	SST

- $SS_A = MS_A = N\hat{\beta}_1^2 = \frac{N}{4}\widehat{A}^2$, where $N = 2^3 n$. The same relation holds for all other main effects and interaction effects.
- F test for each individual effect can be performed using this ANOTA table.

Example: Travel Time from Bicycle Experiment

Box, Hunter, and Hunter (2005, pp. 215-217) describe a student project in which a student studied the effects of bicycle seat height, generator use, and tire pressure on the time taken to make a half-block uphill run. The study was done as a 2^3 design replicated twice ($n = 2$). The levels of the factors were as follows:

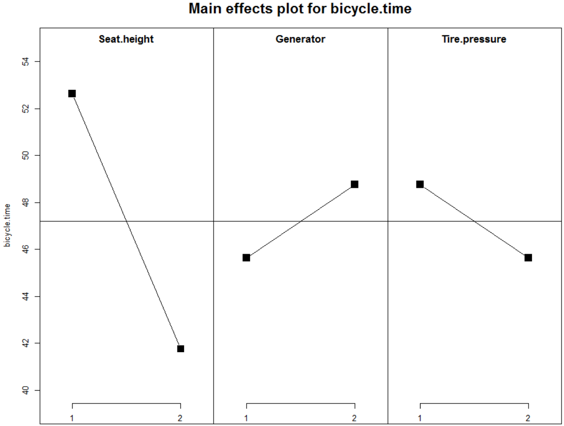
- Seat height (factor A): 26in ($-$), 30in ($+$).
- Generator (factor B): off ($-$), on ($+$).
- Tire pressure (factor C): 40psi ($-$), 55psi ($+$).

Example: Travel Time from Bicycle Experiment

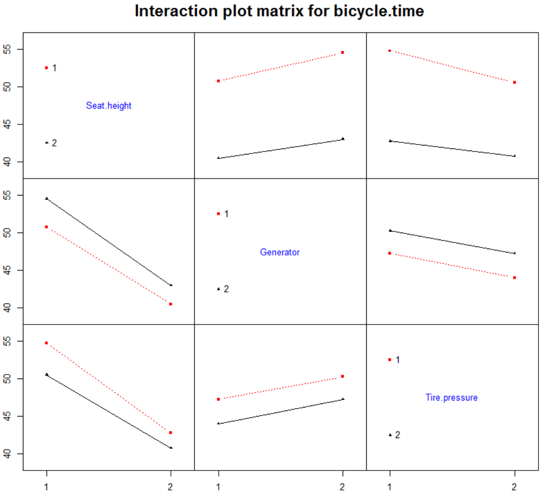
The data from the experiment are in the following table:

Factor			Time (sec.)		
A	B	C	Run 1	Run 2	Mean
$-$	$-$	$-$	51	54	52.5
$+$	$-$	$-$	41	43	42.0
$-$	$+$	$-$	54	60	57.0
$+$	$+$	$-$	44	43	43.5
$-$	$-$	$+$	50	48	49.0
$+$	$-$	$+$	39	39	39.0
$-$	$+$	$+$	53	51	52.0
$+$	$+$	$+$	41	44	42.5

Example: Main Effect Plots for Bicycle Data



Example: Interaction Effect Plots for Bicycle Data



Example: Estimation of Effects

```
> summary(bicycle.fit1)
...
Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                        47.1875     0.5116  92.238 2.13e-13 ***
Seat.height1                       -5.4375     0.5116 -10.629 5.37e-06 ***
Generator1                          1.5625     0.5116   3.054   0.0157 *
Tire.pressure1                     -1.5625     0.5116  -3.054   0.0157 *
Seat.height1:Generator1            -0.3125     0.5116  -0.611   0.5583
Seat.height1:Tire.pressure1         0.5625     0.5116   1.100   0.3035
Generator1:Tire.pressure1           0.0625     0.5116   0.122   0.9058
Seat.height1:Generator1:Tire...     0.4375     0.5116   0.855   0.4173
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 2.046 on 8 degrees of freedom
Multiple R-squared:  0.9436,    Adjusted R-squared:  0.8943
F-statistic: 19.14 on 7 and 8 DF,  p-value: 0.0002109
```

Estimates of effects are:

$$\widehat{A} = -5.4375 \times 2 = -10.875,\ \widehat{B} = 1.5625 \times 2 = 3.125,\ \widehat{C} = -1.5625 \times 2 = -3.125,$$

$$\widehat{AB} = -0.3125 \times 2 = -0.625,\ \widehat{AC} = 0.5625 \times 2 = 1.125,\ \widehat{BC} = 0.0625 \times 2 = 0.125,$$

$$\widehat{ABC} = 0.4375 \times 2 = 0.875.$$

Example: ANOVA Table for Bicycle Data

Effect	S.S.	D.F.	M.S.	F	p value
A	473.0625	1	473.0625	112.97	0.000
B	39.0625	1	39.0625	9.33	0.016
C	39.0625	1	39.0625	9.33	0.016
AB	1.5625	1	1.5625	0.37	0.558
AC	5.0625	1	5.0625	1.21	0.304
BC	0.0625	1	0.0625	0.01	0.906
ABC	3.0625	1	3.0625	0.73	0.417
Error	33.5000	8	4,1875		
Total	594.9375	15			

So all the main effects are significant at level 0.05, while all interaction effects are nonsignificant.

Example: Fitted Model for Bicycle Data

- From the R output, the fitted model for the bicycle data is

$$\hat{y} = 47.1875 - 5.4375x_1 + 1.5625x_2 - 1.5625x_3 - 0.3125x_1x_2 \\ + 0.5625x_1x_3 + 0.0625x_2x_3 + 0.4375x_1x_2x_3.$$

- If we drop the nonsignificant effects AB, AC, BC, ABC, then the fitted model becomes

$$\hat{y} = 47.1875 - 5.4375x_1 + 1.5625x_2 - 1.5625x_3.$$

- Each of x_1, x_2, x_3 takes values in $\{-1, 1\}$.

Yates Algorithm for a 2^k Experiment

- Display the data as a column in standard order.
- Create k new columns as follows.
  (a). The first 2^{k-1} entries in any subsequent column are obtained by addition. The entry in the ith row of a column is the sum of the $(2i-1)$th and the $(2i)$th entries in the previous column.
  (b). The 2nd 2^{k-1} entries are difference. The $(2^{k-1}+i)$th entry is obtained from the previous column by subtracting the $(2i-1)$th entry from the $(2i)$th entry.
- Create a $(k+1)$th column counting of entries all of which are equal to 2^{k-1} except for the 1st entry which is equal to 2^k.
- The estimated effects are obtained by dividing the kth column entries by the corresponding entries in the $(k+1)$th column.
- The identities of the effects are obtained by looking at the signs of the factors along each row.

Yates Algorithm: Truck Leaf Spring Manufacture Experiment

y	1	2	$k=3$	$2^{k-1}(2^k)$	Estimated	Effects	runs
32	57	116	260	8	32.5	μ	(1)
25	59	144	-12	4	-3.0	A	a
28	87	-4	-28	4	-7.0	B	b
31	57	-8	20	4	5.0	AB	ab
48	-7	2	28	4	7.0	C	c
39	3	-30	-4	4	-1.0	AC	ac
28	-9	10	-32	4	-8.0	BC	bc
29	1	10	0	4	0.0	ABC	abc

An Intuitive Formula for Computing Effects

- A general formula for 2^k factorial design parameter estimation:

$$\frac{(a \pm 1)(b \pm 1)(c \pm 1)(d \pm 1) \cdots}{2^{k-1}} \tag{1}$$

  Note that if the upper case letter is included in the estimated term, then the corresponding term in (1) is "$-$", otherwise "$+$". The numerator has exactly k multiplied terms.

- Example: Suppose $k = 5$, we are estimating BDE, then, "B", "D", "E" are included in the estimated term, but "A", "C" are not. Thus, $b - 1, d - 1, e - 1,\ a + 1, c + 1$ are in the formula. Totally, we have $k = 5$ terms in the numerator. Thus, we can have

$$BDE = \frac{(a+1)(b-1)(c+1)(d-1)(e-1)}{2^4}$$

- **Note:** This formula is not computed by plugging in values of "a", "b", ... One should expand the formula and plug in the corresponding running results (explain).

Fundamental Principles in Factorial Experiments

- **Effect Hierarchy Principle:** (1). Lower order effects are more likely to be important than higher order effects. (2). Effects of the same order are equally likely to be important.
- **Effect Sparsity Principle:** The number of relatively important effects in a factorial experiment is small.
- **Effect Heredity Principle:** In order for an interaction to be significant, at least one of its parent factors should be significant.

The Case of Single Replication

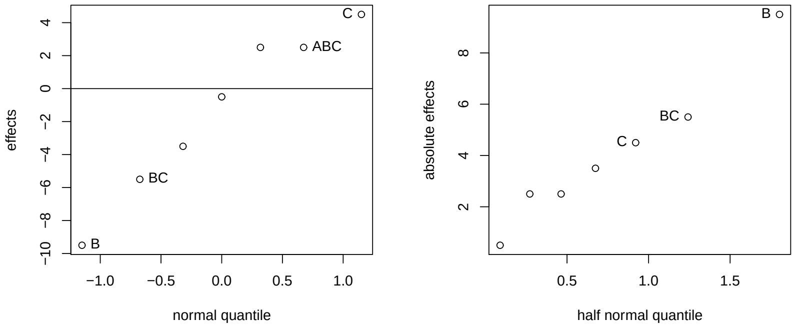
- For a 2^k experiment with single replication (i.e. each treatment combination has exactly 1 replicate), the data cannot be used to estimate the full model parameters and the error variance.
- If one would like to estimate all model parameters, one way is to assume no error terms, i.e. $e_i = 0$. In this case, all the parameters can be exactly computed from the model. (explain)
- If instead we want to estimate the error variance, then one possibility is to give up the estimation of some or all high-order interaction effects.

Use of Normal Plot to Detect Effect Significance

We consider the 2^k experiment with single replication (i.e. $n = 1$ and $N = 2^k$). We denote all the effects as $\hat{\theta}_i$, $i = 1, \ldots, 2^k - 1$.

- If we can assume that all the $2^k - 1$ effects are nonsignificant, then for any effect, $\hat{\theta}_i \sim N\left(0, \frac{4\sigma^2}{2^k}\right)$ (why?). Thus, the resulting normal plot of $\{\hat{\theta}_i : \; i = 1, \ldots, 2^k - 1\}$ should follow a straight line.
- By fitting a straight line through the origin in the normal plot, any effect whose corresponding point is far away from the line can be declared as significant (Daniel 1959).
- Unlike t or F test, no estimate of the error variance σ^2 is required. This method is particularly suitable for experiments with single replication. Daniel's idea is to use "normal" as the *reference* distribution.

Normal Plot and Half Normal Plot



[1]: Comparison of Normal and Half-Normal Plots

Normal Plot and Half Normal Plot

- Potential misuse of normal plot: Based on the left panel of Figure 1, one may declare C as significant because it "deviates" from the straight line. However, the magnitude of BC is even larger than C. This points to a potential visual misjudgement and misuse with the normal plot.
- The half normal plot is better than the normal plot because it only plots the absolute magnitude of effects ($|\hat{\theta}_i|$).
- In general, the half normal plot is more commonly used for detecting significant effects.

Half Normal Plot

- Order the absolute effects $\hat{\theta}_{(i)}$ values as $|\hat{\theta}_{(1)}| \leq \ldots \leq |\hat{\theta}_{(I)}|$ ($I = 2^k - 1$ for a 2^k full factorial experiment), and plot them on the positive axis of the normal distribution (i.e. the half normal distribution). This would avoid the potential misjudgement between the positive and negative values.

- The half normal probability plot consists of the points:

$$(z_{0.5+0.5[i-0.5]/I}, |\hat{\theta}_{(i)}|) \quad \text{for } i = 1, \ldots, I$$

- From Figure 1 (right panel), B deviates away from a fitted straight line. How do we decide whether C and BC are significant?

A Formal Test of Significance: Lenth's Method

The following Lenth's Method offers a formal test for significance of effects.

Lenth's Method:

1. Define an initial estimate of standard error (i.e. the SD of $\hat{\theta}_i$'s)

$$s_0 = 1.5 \cdot \text{median}(|\hat{\theta}_i|),$$

and the pseudo standard error

$$PSE = 1.5 \cdot \text{median}\{|\hat{\theta}_i| : |\hat{\theta}_i| \leq 2.5 s_0\}.$$

The step to discard those values larger than $2.5s_0$ is called trimming.

Reasoning: if $\theta_i = 0$ and error is normal, s_0 is a consistent estimate of the standard deviation of $\hat{\theta}_i$. Use of median gives "robustness" to outlying values of $\hat{\theta}_i$.

Lenth's Method (Cont.)

2. Compute the statistic

$$t_{PSE,i} = \frac{\hat{\theta}_i}{PSE} \quad \text{for each } i.$$

If $|t_{PSE,i}|$ exceeds the critical value given in the table of **Appendix H** of the textbook, $\hat{\theta}_i$ is declared as significant. Two versions of the critical values are considered.

Two Versions of Lenth's Method

- Consider the hypothesis H_0 : all θ_i's $= 0$.
- **Individual Error Rate (IER)**:

$$P(|t_{PSE,i}| > IER_\alpha | H_0) = \alpha \quad \text{for } i = 1, \ldots, I.$$

Note: Since under H_0, $\theta_i = 0$, $t_{PSE,i}$ has the same distribution for all i. That is, IER_α are the same for all i.

- **Experimental-wise Error Rate (EER)**

$$P(|t_{PSE,i}| > EER_\alpha \text{ for at least one } i, i = 1, \ldots, I | H_0)$$
$$= P(\max_{1 \le i \le I} |t_{PSE,i}| > EER_\alpha | H_0) = \alpha.$$

- EER controls the experimental-wise error rate, thus is a multiple comparison technique, but often gives conservative results (less powerful). IER is a individual test approach and in screening experiment, it is more powerful and preferable. The EER critical values can be inflated when there are many $\hat{\theta}_i$ values. (why?)

Example: Truck Leaf Spring Manufacture Experiment

- The seven effects are listed below:

A	B	C	AB	AC	BC	ABC
-0.5	-9.5	4.5	2.5	-3.5	-5.5	2.5

  median $|\hat{\theta}_i| = 3.5$, $s_0 = 1.5 \times 3.5 = 5.25$. Trimming constant $2.5s_0 = 13.125$, which doesn't eliminate any effects. $PSE = 1.5 \cdot \text{median}\{|\hat{\theta}_i| : |\hat{\theta}_i| \leq 2.5s_0\} = 1.5 \times 3.5 = 5.25$. The corresponding t_{PSE} values are listed below

A	B	C	AB	AC	BC	ABC
-0.095	-1.810	0.857	0.476	-0.667	-1.048	0.476

- For $\alpha = 0.05$, $IER_{0.05} = 2.3$, $EER_{0.05} = 4.87$ for $I = 7$. By comparing with the $|t_{PSE}|$, no significant effect declared at level 0.05 for both IER and EER. Consider $\alpha = 0.1$, $IER_{0.1} = 1.71$, $EER_{0.1} = 3.69$. Now, IER declares B as a significant effect, but EER still does not declare any significant effect.

Example: A Single Replicate of 2^4 Design

(Example 6-2 from Montgomery's book) A Single Replicate of the 2^4 design. A chemical product is produced in a pressure vessel. A factorial experiment is carried out in the pilot plant to study the factors thought to influence the filtration rate of this product. The four factors are temperature (A), pressure (B), concentration of formaldehyde (C), and stirring rate (D).

Example: Filtration Rate Experiment

Run Label	Factor A	B	C	D	Filtration rate (gal/h)
(1)	−	−	−	−	45
a	+	−	−	−	71
b	−	+	−	−	48
ab	+	+	−	−	65
c	−	−	+	−	68
ac	+	−	+	−	60
bc	−	+	+	−	80
abc	+	+	+	−	65
d	−	−	−	+	43
ad	−	−	−	+	100
bd	−	+	−	+	45
abd	+	+	−	+	104
cd	−	−	+	+	75
acd	+	−	+	+	86
bcd	−	+	+	+	70
$abcd$	+	+	+	+	96

Example: Estimation of Effects for Filtration Rate Experiment

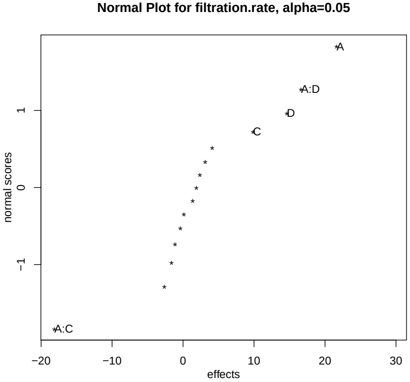
```
> summary(filtration.fit1)
...
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  70.0625         NA      NA       NA
A1           10.8125         NA      NA       NA
B1            1.5625         NA      NA       NA
C1            4.9375         NA      NA       NA
D1            7.3125         NA      NA       NA
A1:B1         0.0625         NA      NA       NA
A1:C1        -9.0625         NA      NA       NA
A1:D1         8.3125         NA      NA       NA
B1:C1         1.1875         NA      NA       NA
B1:D1        -0.1875         NA      NA       NA
C1:D1        -0.5625         NA      NA       NA
A1:B1:C1      0.9375         NA      NA       NA
A1:B1:D1      2.0625         NA      NA       NA
A1:C1:D1     -0.8125         NA      NA       NA
B1:C1:D1     -1.3125         NA      NA       NA
A1:B1:C1:D1   0.6875         NA      NA       NA
```

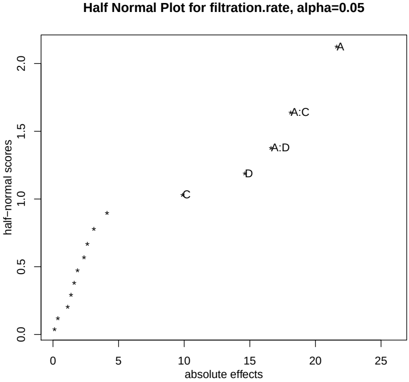
Estimates of effects are:

$\widehat{A} = 10.8125 \times 2 = 21.625$, $\widehat{B} = 1.5625 \times 2 = 3.125$, $\widehat{C} = 4.9375 \times 2 = 9.875$,

$\widehat{D} = 7.3125 \times 2 = 14.625$, $\widehat{AB} = 0.0625 \times 2 = 0.125, \ldots$

Example: Normal Plot and Half Normal Plot for Filtration Rate Experiment





Significant effects at level $\alpha = 0.05$ are marked out in R.

Example: Lenth Method for Filtration Rate Experiment

- In the filtration rate example, $I = 15$, $\alpha = 0.05$. From Appendix H, $IER_\alpha = 2.16$.
- From the estimated effects, $s_0 = 3.9375$, $PSE = 2.625$.
- The effects with $|t_{PSE,i}|$ exceeding 2.16 are: $t_{PSE,A} = 8.238$, $t_{PSE,C} = 3.762$, $t_{PSE,D} = 5.571$, $t_{PSD,AC} = -6.905$, $t_{PSE,AD} = 6.333$. Therefore, A, C, D, AC, AD are declared as the significant effects.

One Factor At a Time (ofat) Approach

Revisit our truck leaf spring manufacture experiment

Runs	Furnace Temp. (^{0}F) A	Heating Time (sec) B	Transfer Time (sec) C	y
1	1840	23	10	32
2	1880	23	10	35
3	1840	25	10	28
4	1880	25	10	31
5	1840	23	12	48
6	1880	23	12	39
7	1840	25	12	28
8	1880	25	12	29

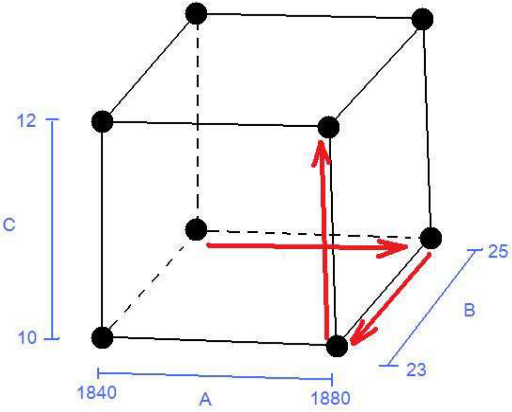
Ofat Steps

Step 1: Suppose we are starting with the standard condition A: 1840; B: 25; C: 10 and choose factor A as the factor of consideration. Then, fixing factors B and C at standard conditions (B: 25; C: 10), two levels of A at 1840 (run 3) and 1880 (run 4) are compared. Thus A is chosen at 1880 as it provides larger quality values.

Step 2: Next, if factor B is chosen as the factor of consideration. By fixing A at 1880 from Step 1 and C at 10 (standard condition), the two levels of B at 23 (run 2) and 25 (run 4) are compared. Here B at 23 is chosen.

Step 3: The last factor we should consider is C. Two levels of C are compared with A at 1880 and B at 23 based on the studies in Step 1 and 2 (say run 2 and run 6). The level C at 12 is chosen.

The Path of a Ofat Approach



The Path of a ofat Plan

Disadvantages of ofat Approach Relative to Factorial Approach

- It requires more runs for the same precision in effect estimation. In the example, the 2^3 design requires 8 runs. For ofat to have the same precision, each of the 4 corners on the ofat path needs to have 4 runs, totaling 16 runs. In general, to be comparable to a 2^k design, ofat would require 2^{k-1} runs at each of the $k+1$ corners on its path, totaling $(k+1)2^{k-1}$. The ratio is $(k+1)2^{k-1}/2^k = (k+1)/2$.
- It cannot estimate some interactions.
- Conclusions for analysis not as general.
- It can miss optimal settings.

Why Experimenters Continue to Use ofat

- Many physical laws are taught by varying one factor at a time. It is easier to focus on one factor each time.
- Experimenters often have good intuition about the problem when thinking in this mode.
- Experimenters lack exposure to statistical design of experiments.