Chapter 31

Explanatory Experiments—Two–Level Factorial and Fractional Factorial Designs

Two–level factorial and fractional factorial designs are useful in the early stages of an experiment, when many factors are investigated.

31.1 Two–Level Full Factorial Designs

Exercise 31.1 (Two–Level Full Factorial Designs)

1. *Different arrangements of same data.*
   Consider the effect of air temperature and noise on the rate of oxygen consumption of mice. The two factors are studied at the following low (-1) and high (1) levels.

temperature ↓ noise →	10 dB (low, $X_2 = -1$)	100 dB (high, $X_2 = 1$)
0^o F (low, $X_1 = -1$)	10.3	11.4
30^o F (high, $X_1 = 1$)	9.4	9.7

   could also be written as (*choose none, one or more*)

   (a) Data candidate A

treatment	temperature, X_1	noise, X_2
10.3	0^o F	10 dB
11.4	0^o F	100 dB
9.4	30^o F	10 dB
9.7	30^o F	100 dB

(b) Data candidate B

treatment	temperature, X_1	noise, X_2
10.3	-1	-1
11.4	-1	1
9.4	1	-1
9.7	1	1

as long as, in this second case, we remember[1]
$X_1 = -1$ means 0^o F, $X_1 = 1$ means 30^o F and that
$X_2 = -1$ means 10 dB, $X_2 = 1$ means 100 dB.

2. *The 2^2 full factorial design, replication.*
Consider the effect of air temperature and noise on the rate of oxygen consumption of mice. We assume
$X_1 = -1$ means 0^o F, $X_1 = 1$ means 30^o F and that
$X_2 = -1$ means 10 dB, $X_2 = 1$ means 100 dB.

treatment	temperature, X_1	noise, X_2
10.3	-1	-1
11.4	-1	1
9.4	1	-1
9.7	1	1
8.3	-1	-1
10.4	-1	1
9.5	1	-1
8.7	1	1

This experiment has been replicated (choose one) **once** / **twice**.

3. *The 2^2 full factorial design, interaction.*
Consider the effect of air temperature and noise on the rate of oxygen consumption of mice. We assume
$X_1 = -1$ means 0^o F, $X_1 = 1$ means 30^o F and that

[1] Notice that this coding scheme is essentially opposite of what we have done in the past where, for example,

$$X_1 = \begin{cases} 1, & \text{if case from treatment 1} \\ -1, & \text{if case from treatment 2,} \end{cases}$$

or, if more than two (r, say) levels of the factor,

$$X_i = \begin{cases} 1, & \text{if case from treatment 1} \\ -1, & \text{if case from treatment r} \\ 0, & \text{otherwise,} \end{cases}$$

and so on for $X_{i+1}, \ldots, X_{r-1}$.

$X_2 = -1$ means 10 dB, $X_2 = 1$ means 100 dB.
This 2^2 full factorial design, has one interaction term, X_{12}, (choose one)

(a) **temperature** $\times$ **noise**

(b) **temperature** $\times$ **temperature**

(c) **noise** $\times$ **noise**

and is represented by the following coding scheme,

treatment	temperature, X_1	noise, X_2	temp $\times$ noise, X_{12}
10.3	-1	-1	1
11.4	-1	1	-1
9.4	1	-1	-1
9.7	1	1	1

4. *The 2^2 full factorial model, including intercept X_0 and interaction X_{12}.* Consider the effect of air temperature and noise on the rate of oxygen consumption of mice. We assume $X_1 = -1$ means 0^o F, $X_1 = 1$ means 30^o F and that $X_2 = -1$ means 10 dB, $X_2 = 1$ means 100 dB.
Using the following regression version of the 2^2 full factorial model[2], match appropriately,

$$Y_i = \beta_0 X_{i0} + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_{12} X_{i12} + \varepsilon_i,$$

model	example
(a) Y_i	**(a)** error
(b) β_0	**(b)** interaction temp $\times$ noise effect
(c) β_1	**(c)** main noise effect
(d) β_2	**(d)** main temperature effect
(e) β_{12}	**(e)** (grand) average of all responses
(f) ε_i	**(f)** individual roc response

model	(a)	(b)	(c)	(d)	(e)	(f)
example						

where the *design* matrix is given by

[2]The notation has changed (simplified) from before, where, instead, we would have written

$$Y_{ijk} = \mu_{\cdot\cdot} + \alpha_1 X_{ijk1} + \beta_1 X_{ijk2} + (\alpha\beta)_{11} X_{ijk1} X_{ijk2} + \varepsilon_{ijk}$$

treatment	intercept, X_0	temperature, X_1	noise, X_2	temp × noise, X_{12}
10.3	1	−1	−1	1
11.4	1	−1	1	−1
9.4	1	1	−1	−1
9.7	1	1	1	1

5. *Standard order for a 2^2 full factorial design.*
The *standard order* of the treatments is a listing of the treatments such that the level of the first factor[3] changes *most frequently* and the level of the second factor changes second most frequently and so on.
Consider the effect of air temperature and noise
on the rate of oxygen consumption of mice. We assume
$X_1 = -1$ means 0^o F, $X_1 = 1$ means 30^o F and that
$X_2 = -1$ means 10 dB, $X_2 = 1$ means 100 dB.
The standard order of the design matrix of the main factors in this study is (choose one)

(a) Factor Candidate List A

treatment	temperature, X_1	noise, X_2
10.3	−1	−1
11.4	−1	1
9.4	1	−1
9.7	1	1

(b) Factor Candidate List B

treatment	temperature, X_1	noise, X_2
10.3	−1	−1
9.4	1	−1
11.4	−1	1
9.7	1	1

(c) Factor Candidate List C

treatment	temperature, X_1	noise, X_2
9.4	1	−1
9.7	1	1
10.3	−1	−1
11.4	−1	1

6. *A 2^3 full factorial design.*
Consider the effect of air temperature, noise and humidity on the rate of oxygen

[3] The first factor is X_1, not X_0, which is associated with the overall average.

consumption of mice. The three factors are studied at the following low and high levels.

factor	low level, −1	high level, 1
temperature, X_1	0^o F	30^o F
noise, X_2	10 dB	100 dB
humidity, X_3	50%	75%

The model is given by

$$\begin{aligned} Y_i = \beta_0 X_0 &+ \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} \\ &+ \beta_{12} X_{i12} + \beta_{13} X_{i13} + \beta_{23} X_{i23} \\ &+ \beta_{123} X_{i123} + \varepsilon_{ijk} \end{aligned}$$

where the complete design matrix $\boldsymbol{X}$ is

Y	X_0	X_1	X_2	X_3	X_{12}	X_{13}	X_{23}	X_{123}
9.8	1	−1	−1	−1	1	1	1	−1
10.2	1	1	−1	−1	−1	−1	1	1
10.4	1	−1	1	−1	−1	1	−1	1
8.5	1	1	1	−1	1	−1	−1	−1
11.1	1	−1	−1	1	1	−1	−1	1
10.7	1	1	−1	1	−1	1	−1	−1
10.7	1	−1	1	1	−1	−1	1	−1
9.9	1	1	1	1	1	1	1	1

There are 2^3 = (choose one) **8** / **16** / **32** treatments
and the design matrix (choose one) **is** / **is not** in standard order.
There are four interaction terms
(choose *four*) $\boldsymbol{\beta_1}$ / $\boldsymbol{\beta_2}$ / $\boldsymbol{\beta_{12}}$ / $\boldsymbol{\beta_{13}}$ / $\boldsymbol{\beta_{23}}$ / $\boldsymbol{\beta_{123}}$
If $X_1 = X_2 = X_3 = -1$, Y = (choose one) **9.8** / **10.7** / **11.1**
If $X_1 = X_2 = -1$ and $X_3 = 1$, Y = (choose one) **9.9** / **10.7** / **11.1**
If $X_1 = X_2 = X_3 = -1$, $X_{13} = X_1 \times X_3 = -1 \times -1$ = (choose one) **−1** / **1**
If $X_1 = X_2 = -1$ and $X_3 = 1$, $X_{13} = X_1 \times X_3 = -1 \times 1$ = (choose one) **−1** / **1**
If $X_1 = X_2 = -1$ and $X_3 = 1$, $X_{123} = -1 \times -1 \times 1$ = (choose one) **−1** / **1**
Furthermore, this experiment has been replicated (choose one) **once** / **twice**
and so will have no (zero) error variance and, as a consequence, it is
(choose one) **still** / **not** possible to calculate the p–value for the various effects.

7. *A 2^4 full factorial design.*
Consider the effect of air temperature, noise, humidity and air pressure on the rate of oxygen consumption of mice. The four factors are studied at the following low and high levels,

factor	low level, -1	high level, 1
temperature, X_1	0^o F	30^o F
noise, X_2	10 dB	100 dB
humidity, X_3	50%	75%
pressure, X_4	28.5 inches	30.5 inches

where the design matrix of the main factors and intercept are given by,

ROC response, Y	X_0	X_1	X_2	X_3	X_4
9.8	1	-1	-1	-1	-1
10.2	1	1	-1	-1	-1
10.4	1	-1	1	-1	-1
8.5	1	1	1	-1	-1
11.1	1	-1	-1	1	-1
10.7	1	1	-1	1	-1
10.7	1	-1	1	1	-1
9.9	1	1	1	1	-1
9.9	1	-1	-1	-1	1
10.4	1	1	-1	-1	1
10.2	1	-1	1	-1	1
9.2	1	1	1	-1	1
11.8	1	-1	-1	1	1
9.5	1	1	-1	1	1
10.7	1	-1	1	1	1
9.8	1	1	1	1	1

which (choose one) **is** / **is not** in standard order
Also, there are (choose one) **4** / **16** / **32** factors
and 2^4 = (choose one) **4** / **16** / **32** treatments
in this two–level full factorial design
If $X_1 = X_2 = -1$ and $X_3 = X_4 = 1$, Y = (choose one) **9.9** / **10.7** / **11.8**
If $X_1 = X_2 = X_3 = X_4 = -1$, $X_{14} = X_1 \times X_3 = -1 \times -1 =$
(choose one) $\mathbf{-1}$ / **1**
If $X_1 = X_2 = -1$ and $X_3 = X_4 = 1$, $X_{123} = X_1 \times X_3 = -1 \times 1 =$
(choose one) $\mathbf{-1}$ / **1**
If $X_1 = X_2 = -1$ and $X_3 = X_4 = 1$, $X_{134} = -1 \times -1 \times 1 =$
(choose one) $\mathbf{-1}$ / **1**
and which (choose one) **is** / **is not** replicated
and so it is not possible to calculate the p–values for the various effects.

8. *More* 2^4 *full factorial design.*
Consider the effect of air temperature, noise, humidity and air pressure on the rate of oxygen consumption of mice. The four factors are studied at the following low and high levels.

factor	low level, −1	high level, 1
temperature, X_1	0^o F	30^o F
noise, X_2	10 dB	100 dB
humidity, X_3	50%	75%
pressure, X_4	28.5 inches	30.5 inches

The model is given by

$$\begin{aligned} Y_i = \beta_0 X_0 &+ \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} \\ &+ \beta_{12} X_{i12} + \beta_{13} X_{i13} + \beta_{14} X_{i14} \\ &+ \beta_{23} X_{i23} + \beta_{24} X_{i24} + \beta_{34} X_{i34} \\ &+ \beta_{123} X_{i123} + \beta_{124} X_{i124} + \beta_{134} X_{i134} + \beta_{234} X_{i234} \\ &+ \beta_{1234} X_{i1234} + \varepsilon_{ijk} \end{aligned}$$

where

$$X_1 = \begin{cases} 1, & \text{if case from first level of factor 1} \\ -1, & \text{if case from second level of factor 1,} \end{cases}$$

$$X_2 = \begin{cases} 1, & \text{if case from first level of factor 2} \\ -1, & \text{if case from second level of factor 2,} \end{cases}$$

and so on and where, for example, $X_{12} = X_1 \times X_2 = X_1 X_2$.
So, X_{123} = (choose one) $\boldsymbol{X_1 X_2}$ / $\boldsymbol{X_1 X_2 X_3}$ / $\boldsymbol{X_2 X_3}$
We would first check if any of the interaction effects are significant.
If not, we are then free to test the main effects.
If so, we would then investigate the interaction effects further.

9. *A 2^7 full factorial design: exploratory design.*
Consider the effect of seven factors on the rate of oxygen consumption of mice. Each factor is studied at the following low and high levels.

factor	low level, −1	high level, 1
temperature, X_1	0^o F	30^o F
noise, X_2	10 dB	100 dB
humidity, X_3	50%	75%
pressure, X_4	28.5 inches	30.5 inches
lighting, X_5	50 watts	100 watts
living space, X_6	25 in^3	50 in^3
water, X_7	10 fluid ounces	30 fluid ounces

There are (choose one) **7** / **14** / **128**
treatments in this two–level full factorial design.
Most likely, this is an *exploratory* experiment where, to begin with, we are

simply interested, after checking the interactions, in identifying which of the seven factors are *active*,

$$H_0: \ \beta_q = 0 \quad vs \quad H_a: \ \beta_q \neq 0$$

Having identified the (two, three or four, say) active factor effects, we would then go on to analyze these active factor effects in greater detail using the previously discussed designs.

31.2 Analysis of Unreplicated Two–Level Studies

SAS program: att12-31-2-mice-23fullfactorial

In unreplicated two–level studies, there is *no* estimated error variance. Consequently, it is not possible test for active factor effects because the F statistics necessary to do this require the estimated error variance (MSE). There are two methods for obtaining estimated error variance

- pool higher order interactions and form the error variance from these interactions.
- add more observations (replications) called *center points* and form the error variance from these added replications.

Exercise 31.2 (Analysis of Unreplicated Two–Level Studies)
Consider the effect of air temperature, noise and humidity on the rate of oxygen consumption of mice. The three factors are studied at the following low and high levels.

factor	low level, -1	high level, 1
temperature, X_1	0^o F	30^o F
noise, X_2	10 dB	100 dB
humidity, X_3	50%	75%

where

ROC response, Y	X_0	X_1	X_2	X_3
9.8	1	-1	-1	-1
10.2	1	1	-1	-1
10.4	1	-1	1	-1
8.5	1	1	1	-1
11.1	1	-1	-1	1
10.7	1	1	-1	1
10.7	1	-1	1	1
9.9	1	1	1	1

and where the regression equation is

$$\begin{aligned} Y_i = \beta_0 &+ \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} \\ &+ \beta_{12} X_{i12} + \beta_{13} X_{i14} + \beta_{23} X_{i12} \\ &+ \beta_{123} X_{i123} + \varepsilon_{ijk} \end{aligned}$$

1. *A 2^3 full factorial design, one replication.*
   From the SAS output, although there is no (zero) error variance (and so no p–values), the coefficients of the various effects in the model are (choose one)

   (a) Candidate model A

coefficient	b_q	p–value
b_0	10.1625	
b_1	-0.3375	na
b_2	-0.2875	na
b_3	0.4375	na
b_{12}	-0.3375	na
b_{13}	0.0375	na
b_{23}	-0.0125	na
b_{123}	0.2375	na

   (b) Candidate model B

coefficient	b_q	p–value
b_0	13.1625	
b_1	-0.3375	na
b_2	-0.2875	na
b_3	0.4375	na
b_{12}	-0.6375	na
b_{13}	0.0375	na
b_{23}	-0.0125	na
b_{123}	0.2375	na

2. *A 2^3 full factorial design, pool higher order interactions.*
From the SAS output, match the following two different pooling methods (I or II) with the two estimated models (A or B).

Pooling Method	I	II
Coefficients	(A)/(B)	(A)/(B)

(I) Pooling Method I:
X_{123} is used alone

(II) Pooling Method II:
X_{12}, X_{13}, X_{23} and X_{123} are pooled together

(A) Candidate model A

coefficient	b_q	p–value
b_0	10.1625	0.0149
b_1	-0.3375	0.3904
b_2	-0.2875	0.4396
b_3	0.4375	0.3166
b_{12}	-0.3375	0.3904
b_{13}	0.0375	0.9003
b_{23}	-0.0125	0.9665

(B) Candidate model B

coefficient	b_q	p–value
b_0	10.16125	0.0001
b_1	-0.3375	0.1788
b_2	-0.2875	0.2377
b_3	0.4375	0.1024

In both pooling methods, *none* of the main or interaction effects are significant (have a p–value less than 0.05) except the intercept (the grand average); in other words, the mice roc is *not* influenced by any of these factors.

3. *A 2^3 full factorial design, replications at center point.*
Consider the effect of air temperature, noise and humidity on the rate of oxygen consumption of mice. The three factors are studied at the following low and high levels.

factor	low level, -1	high level, 1
temperature, X_1	0^o F	30^o F
noise, X_2	10 dB	100 dB
humidity, X_3	50%	75%

where

ROC response, Y	X_0	X_1	X_2	X_3
9.8	1	-1	-1	-1
10.2	1	1	-1	-1
10.4	1	-1	1	-1
8.5	1	1	1	-1
11.1	1	-1	-1	1
10.7	1	1	-1	1
10.7	1	-1	1	1
9.9	1	1	1	1
10.3	1	0	0	0
9.7	1	0	0	0
10.0	1	0	0	0

In this case, three additional center points are used to calculate error variance and are given by (choose one)

(a) Candidate Center Points A
10.3, 9.7, 10.0

(b) Candidate Center Points B
9.9, 10.3, 9.7

4. *More on a 2^3 full factorial design, replications at center point.*
The resulting coefficients and corresponding p–values are given below.

coefficient	b_q	p–value
b_0	10.11818	0.000
b_1	-0.3375	0.086
b_2	-0.2875	0.1134
b_3	0.4375	0.000
b_{12}	-0.3375	0.086
b_{13}	0.0375	0.7574
b_{23}	-0.0125	0.9169
b_{123}	0.2375	0.077

The p–values given in this table do NOT match the p–values given on the SAS output; they are calculated in a special way. For example, for coefficient b_1,

$$MSPE = \frac{\sum(Y_{0i} - \bar{Y}_{0\cdot})^2}{n_0 - 1} = \frac{(10.3-10)^2 + (9.7-10)^2 + (10.0-10)^2}{3-1} \approx 0.09$$

and so $s\{b_q\} = \sqrt{\frac{MSPE}{n_T}} = \sqrt{\frac{0.09}{8}} \approx 0.106$
and so $t = \frac{b_1}{s\{b_q\}} = \frac{-0.3375}{0.106} = -3.181$

and so p–value $= 2 \times P(T < -3.181; 2) \approx 0.086$
The other p–values are calculated in a similar way.
In this case, the *significant* effects at $\alpha = 0.10$ are (choose one)

(i) β_1

(ii) β_2

(iii) β_1 and β_3

(iv) β_3 and β_{12}

(v) β_1, β_2, β_3, β_{12} and β_{123}

5. *More investigation of* 2^3 *full factorial design, replications at center point.*

   (a) *Residual plot versus predicted*
   The residual plot, from the SAS output, indicates
   (choose one) **constant** / **nonconstant** variance
   In fact, there appears to be one bad outlier.

   (b) *Normal probability plot of the residuals*
   The normal probability plot, also obtained from the SAS output, indicates
   (choose one) **normal** / **non–normal** residuals.

   (c) *Test of correlation of normal probability plot of the residuals at* $\alpha = 0.05$
   $H_0:$ normal versus $H_a:$ not normal
   from SAS, $r =$ (choose one) **0.6556** / **0.7753** / **0.918**
   from table B.6[4], $r^* =$ (choose one) **0.6556** / **0.7753** / **0.918**
   since $r = 0.77530 < r^* = 0.918$
   reject null; that is, the residuals are *not* normal
   and so, really, we can*not* proceed

6. *Investigation of reduced* 2^3 *full factorial design, no center point replications but including only main factors 1 and 3.*
   We will include *only* main factors 1 and 3 in a *second* analysis because these main factors were the most significant quantities in the initial analysis. We will then check to see if this change improves our results.

   (a) *Residual plot versus predicted*
   The residual plot, from the SAS output, indicates
   (choose one) **constant** / **nonconstant** variance

   (b) *Normal probability plot of the residuals*
   The normal probability plot, also obtained from the SAS output, indicates
   (choose one) **normal** / **non–normal** residuals

[4] where $n - 1 = 11 - 1 = 10$ and $\alpha = 0.05$

(c) *Test of correlation of normal probability plot of the residuals at* $\alpha = 0.05$
H_0 : normal versus H_a : not normal
from SAS, $r =$ (choose one) **0.6556 / 0.7753 / 0.98627**
from table B.6[5], $r^* =$ (choose one) **0.6556 / 0.898 / 0.918**
since $r = 0.98627 > r^* = 0.898$
accept null; that is, the residuals *are* normal
(which confirms the findings in the normal probability plot)

(d) *Lack of Fit test*
From SAS, the ANOVA table is

Source	Sum Of Squares	Degrees of Freedom	Mean Squares
Regression	2.4425	2	1.22125
Error	2.03625	5	0.40725
Lack of Fit	0.01125	1	0.01125
Pure Error	2.025	4	0.50625
Total	4.47875	7	

H_0 : $\mu = \beta_0 + \beta_1 X_{i1} + \beta_3 X_{i3}$ versus
H_a : $\mu \neq \beta_0 + \beta_1 X_{i1} + \beta_3 X_{i3}$
(In other words, H_o : no lack of fit versus H_a : lack of fit)
The test statistic is

$$\begin{aligned} F^* &= \frac{SSE(R) - SSE(F)}{df_R - df_F} \div \frac{SSE(F)}{df_F} \\ &= \frac{SSE - SSPE}{df_E - df_{PE}} \div \frac{SSPE}{df_{PE}} \\ &= \frac{SSLF}{df_{LF}} \div \frac{SSPE}{df_{PE}} \\ &= \frac{0.01125}{1} \div \frac{2.025}{4} = \end{aligned}$$

(choose one) **0.02 / 0.7753 / 0.98627**
The critical value is $F(1-\alpha; df_{LF}, df_{PE}) = F(0.95; 1, 4) = 7.71$
since $F^* = 0.02 < F = 7.71$
accept null; that is, there is *no* lack of fit

(e) *Which coefficients are significant?*
H_0 : $\beta_q = 0$ versus H_a : $\beta_q \neq 0$
The significant effects[6] are (choose one)

(i) none

(ii) β_3

[5] where $n - 1 = 8 - 1 = 7$ and $\alpha = 0.05$

[6] Use the p–values as is on the SAS output; do not change them as was done above, compare with $\alpha = 0.05$.

(iii) β_1, β_3
(iv) β_1 and β_{34}
(v) β_1, β_3 and β_{34}

(f) *More graphical findings*
From SAS *main effect* plot graph output,
If temperature X_1 is high, (1), roc is (choose one) **low** / **high**
If humidity X_3 high, (1), roc is (choose one) **low** / **high**

(g) *Estimated model*
The estimated model chosen by this analysis is (choose one)

i. $\hat{Y} = 9.1625 - 0.3375X_1 + 0.4375X_3$
ii. $\hat{Y} = 10.1625 - 0.3375X_1 + 0.4375X_3$
iii. $\hat{Y} = 11.1625 - 0.3375X_1 + 0.4375X_3$

and so, for example, if temperature and humidity were both set at high levels, $X_1 = X_3 = 1$, then the roc[7] of mice would be

$$\hat{Y} = 9.1625 - 0.3375(1) + 0.4375(1) = 9.2625$$

31.3 Two–Level Fractional Factorial Designs

SAS program: att12-31-3-fracfactorial

Factional factorial designs allow the analysis of factorial designs with less (a fraction of the) data.

Exercise 31.3 (Two–Level Fractional Factorial Designs)

1. *Two–Level Fractional Factorial Designs.*
Consider the effect of air temperature, noise, humidity and air pressure on the rate of oxygen consumption of mice. The four factors are studied at the following low and high levels.

factor	low level, -1	high level, 1
temperature, X_1	0^o F	30^o F
noise, X_2	10 dB	100 dB
humidity, X_3	50%	75%
pressure, X_4	28.5 inches	30.5 inches

where, for a *full* 2^4 factorial design,

[7]One possible objective of the analysis would be to set the levels of temperature and humidity to achieve some "optimal" setting, such as 9.0, say.

ROC response, Y	X_0	X_1	X_2	X_3	X_4
9.8	1	−1	−1	−1	−1
10.2	1	1	−1	−1	−1
10.4	1	−1	1	−1	−1
8.5	1	1	1	−1	−1
11.1	1	−1	−1	1	−1
10.7	1	1	−1	1	−1
10.7	1	−1	1	1	−1
9.9	1	1	1	1	−1
9.9	1	−1	−1	−1	1
10.4	1	1	−1	−1	1
10.2	1	−1	1	−1	1
9.2	1	1	1	−1	1
11.8	1	−1	−1	1	1
9.5	1	1	−1	1	1
10.7	1	−1	1	1	1
9.8	1	1	1	1	1

where there are (choose one) **4** / **16** / **32** factors
and 2^4 = (choose one) **4** / **16** / **32** treatments.

However, an example of a *half fractional* factorial design is

ROC response, Y	X_0	X_1	X_2	X_3	X_4
9.8	1	−1	−1	−1	−1
10.2	1	1	−1	−1	−1
10.4	1	−1	1	−1	−1
8.5	1	1	1	−1	−1
11.1	1	−1	−1	1	−1
10.7	1	1	−1	1	−1
10.7	1	−1	1	1	−1
9.9	1	1	1	1	−1

where there are (choose one) **4** / **16** / **32** factors
but only 2^{4-1} = (choose one) **4** / **8** / **32** treatments

An example of another *fractional* factorial design is

ROC response, Y	X_0	X_1	X_2	X_3	X_4
9.8	1	−1	−1	−1	−1
10.2	1	1	−1	−1	−1
10.4	1	−1	1	−1	−1
8.5	1	1	1	−1	−1

where there are (choose one) **4** / **16** / **32** factors
but only 2^{4-2} = (choose one) **4** / **8** / **32** treatments
and so this is an example of a 2^{-2} = **half** / **quarter** / **eighth** fractional design.

An example of a *replicated quarter fractional* factorial design is

ROC response, Y	X_0	X_1	X_2	X_3	X_4
9.8	1	−1	−1	−1	−1
10.2	1	1	−1	−1	−1
10.4	1	−1	1	−1	−1
8.5	1	1	1	−1	−1
8.8	1	−1	−1	−1	−1
9.2	1	1	−1	−1	−1
10.6	1	−1	1	−1	−1
8.8	1	1	1	−1	−1

where there are (choose one) **4** / **16** / **32** factors
but only 2^{4-2} = (choose one) **4** / **8** / **32** treatments
and (choose one) **1** / **2** / **3** replications.

2. *Confounding (Aliasing).*
Consider the following 2^{4-2} fractional factorial design.

ROC response, Y	X_0	X_1	X_2	X_3	X_4
9.8	1	−1	−1	−1	−1
10.2	1	1	−1	−1	−1
10.4	1	−1	1	−1	−1
8.5	1	1	1	−1	−1

Notice that column vectors X_3 and X_4 are identical, $\mathbf{X}_3 = \mathbf{X}_4$. This means that in the model of this design is given by

$$\begin{aligned} Y &= \cdots + \beta_3 X_3 + \beta_4 X_4 + \cdots \\ &= \cdots + \beta_3 X_3 + \beta_4 X_3 + \cdots \\ &= \cdots + (\beta_3 + \beta_4) X_3 + \cdots \end{aligned}$$

and so we are not able to estimated factors 3 and 4 separately, but only their combined effects. Factors 3 and 4 are said to be *confounded* (or *aliased*). For the following 2^{4-1} fractional factorial design,

ROC response, Y	X_0	X_1	X_2	X_3	X_4
9.8	1	−1	−1	−1	−1
10.2	1	1	−1	−1	−1
10.4	1	−1	1	−1	−1
8.5	1	1	1	−1	−1
11.1	1	−1	−1	1	−1
10.7	1	1	−1	1	−1
10.7	1	−1	1	1	−1
9.9	1	1	1	1	−1

For this design, (choose one)

(a) factor 1 and factor 4 are confounded, $\mathbf{X}_1 = \mathbf{X}_4$.

(b) factor 1 and factor 2 are confounded, $\mathbf{X}_1 = \mathbf{X}_2$.

(c) factor 2 and factor 3 are aliased, $\mathbf{X}_2 = \mathbf{X}_3$.

(d) factor 3 and factor 4 are confounded, *3 = 4*.

(e) none of the *main* factors, 1, 2, 3 or 4 are confounded with one another (although this does not mean some of the main factors are confounded with some of the interaction factors).

3. *Notation of confounding.*
Consider the following design,

ROC response, Y	X_0	X_1	X_2	X_3	X_4
9.8	1	−1	−1	−1	−1
10.2	1	1	−1	−1	−1
10.4	1	−1	1	−1	−1
8.5	1	1	1	−1	−1
11.1	1	−1	−1	1	−1
10.7	1	1	−1	1	−1
10.7	1	−1	1	1	−1
9.9	1	1	1	1	−1

(a) **True / False**
Since $X_0 = 1$ and so multiplying each term in the X_0 vector by itself[8],

$$X_0 \times X_0 = X_0^2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = X_0$$

we can state this using the following more succinct notation, as

$$\mathit{0} \times \mathit{0} = \mathit{0}^{\mathit{2}} = \mathit{0}$$

(b) **True / False**
Since multiplying X_0 by any other column X_2 (say) gives X_2, in other words,

$$X_0 \times X_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \times \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = X_2$$

we can state this using the following more succinct notation, as
$\mathit{0} \times \mathit{2} = \mathit{2}$, or, in general, for any factor q

$$\mathit{0} \times q = q$$

(c) **True / False**
Since multiplying a column X_2 (say) by itself, gives a column of 1s, in other words,

$$X_2 \times X_2 = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} \times \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = X_0$$

[8] Notice that the multiplication does *not* involve transposing one of the vectors.

we can state this using the following more succinct notation, as
$2 \times 2 = 0$, or, in general, for any factor q

$$q \times q = 0$$

(d) **True / False**
Since multiplying one column X_1 (say) by another column (not X_0 and not itself) X_2 (say), gives a *new* column of numbers, which is neither X_0, X_1 or X_2

$$X_1 \times X_2 = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \times \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \neq X_1, \text{ or } X_2$$

we can state this using the following more succinct notation, as
$1 \times 2 = 12$, or, in general, for any two factors q and r

$$q \times r = qr$$

4. *More confounding*
$0 \times 0 =$ (choose one) 0 / 1 / 2
$0 \times 3 =$ (choose one) 0 / 1 / 3
$3 \times 3 =$ (choose one) 0 / 1 / 3
$3 \times 4 =$ (choose one) 0 / 3 / 34
$3 \times 3 \times 3 =$ (choose one) 0 / 1 / 3
$3 \times 3 \times 4 =$ (choose two) 0 / 04 / 4
$3 \times 4 \times 3 =$ (choose two) 0 / 04 / 4
$343 =$ (choose two) 0 / 04 / 4
$3^2 4 =$ (choose two) 0 / 04 / 4
$3^3 4 =$ (choose one) 0 / 04 / 34
$123^3 4 =$ (choose one) 123 / 124 / 1234

5. *More confounding.*
Match the confounding in the two columns.

Column I	Column II
(a) 2×1234	**(a)** 23
(b) 12×13	**(b)** 134
(c) 0×3	**(c)** 123
(d) 34×124	**(d)** 3

Column I	(a)	(b)	(c)	(d)
Column II				

6. *Confounding scheme.*
   Consider the following 2^{4-2} fractional factorial design.

X_0	X_1	X_2	X_3	X_4	X_{12}	X_{13}	X_{14}	X_{23}	X_{24}	X_{34}	X_{123}	X_{124}	X_{134}	X_{234}	X_{1234}
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

   For this design, (*choose none, one or more*)

   (a) *0* = *34*
   (b) *1* = *134*
   (c) *2* = *234*
   (d) *3* = *4*
   (e) *12* = *1234*
   (f) *13* = *14*
   (g) *23* = *24*
   (h) *123* = *124*

   In fact, this is the entire *confounding scheme* and relation *0* = *34* is said to be the *defining relation*.

7. *Defining relation generates confounding scheme.*
   True / **False** The entire confounding scheme can be found from the defining relation. For example, for the 2^{4-2} fractional factorial design with defining relation *0* = *34*,

$$0 \times 1 = 34 \times 1 = 134$$

   and all of the rest of the confounding scheme can be found in a similar way.

8. *More defining relation generates confounding scheme.*
   The entire confounding scheme for the 2^{4-2} fractional factorial design with defining relation *0* = *12*, is given by (*choose none, one or more*)

   (a) *0* = *12*
   (b) *1* = *2*
   (c) *3* = *123*
   (d) *4* = *124*

(e) *13 = 23*

(f) *14 = 24*

(g) *34 = 1234*

(h) *134 = 234*

9. *More defining relation generates confounding scheme.*
The entire confounding scheme for the 2^{4-1} fractional factorial design with defining relation *0 = 123*, is given by (*choose none, one or more*)

(a) *0 = 123*

(b) *1 = 23*

(c) *2 = 13*

(d) *3 = 12*

(e) *4 = 1234*

(f) *14 = 234*

(g) *24 = 134*

(h) *34 = 124*

10. *Good confounding schemes*
It is better that the main effects are confounded with high order interaction effects rather than either other main effects or low order interaction effects because it is often the case that the interaction effects, particularly high order interactions, are insignificant. So, although a main effect and an (assumed insignificant) interaction effect are confounded, we would be inclined to believe that the main effect, rather than the interaction effect, in the confounding relationship is the contributing significant factor. Consequently, the best confounding scheme of the three given above, is (choose one)

(a) *0 = 12*

(b) *0 = 123*

(c) *0 = 34*

because in this case the main effects are confounded with, at worst, second order confounding interactions, whereas in the other two confounding schemes, at least one main effect is confounded with another main effect.

11. *Defining Relation and Resolution.*
The number of factors in the lowest order effect in a defining relation is the called the *resolution*. For example, defining relation *0 = 12* has resolution II. In a similar way, defining relation *0 = 123* has resolution

(choose one) **I** / **II** / **III** / **IV**
In general, the higher the resolution, the better the design; higher resolution designs have confounding schemes where the main effects are confounded with higher order interactions.

12. *More resolution.*
From SAS, the resolution of a

2^{5-1} design, since *0 = 1234*, for example, is
(choose one) **I** / **II** / **III** / **IV** / **V**

2^{5-2} design, since lowest of *0 = 145* and *0 = 1234*, for example, is
(choose one) **I** / **II** / **III** / **IV** / **V**

31.4 Screening Experiments

Not covered.

31.5 Incomplete Block Designs for Two–Level Factorial Experiments

Not covered.