Fractional Factorial Designs

BIOE 498/598

2/24/2020

The problem with Factorial Designs

Factorial designs are the most efficient designs for estimating effects.

The problem with Factorial Designs

Factorial designs are the most efficient designs for estimating effects.

Their efficiency **grows** as the number of factors increases.

The problem with Factorial Designs

Factorial designs are the most efficient designs for estimating effects.

Their efficiency **grows** as the number of factors increases.

Unfortunately, the number of runs also grows. Quickly.

Factors (k)	Experiments (2^k)
4	16
5	32
6	64
7	128
8	256
9	512

How do we conduct experiments with lots of factors?

Most experimenters abandon factorial designs when the number of factors becomes large. Common strategies are to

- Resort to one-at-a-time designs
- Select only a subset of factors for a factorial design

How do we conduct experiments with lots of factors?

Most experimenters abandon factorial designs when the number of factors becomes large. Common strategies are to

- Resort to one-at-a-time designs
- Select only a subset of factors for a factorial design

In both cases we lose the efficiency and power of the factorial design.

A better method is to use a *fractional factorial design*.

Fractional Factorial Designs

A (full) factorial design with k factors, each with two levels, is called a 2^k design.

We can instead test k factors using only half of the runs of a 2^k design. This is called a 2^{k-1} fractional design.

Fractional Factorial Designs

A (full) factorial design with k factors, each with two levels, is called a 2^k design.

We can instead test k factors using only half of the runs of a 2^k design. This is called a 2^{k-1} fractional design.

For example:

- A 2^4 design tests four 2-level factors using 16 runs.
- A 2^{4-1} design tests four 2-level factors using 8 runs.
- A 2^3 design tests three 2-level factors using 8 runs.

Why do fractional factorial designs work?

Fractional designs are motivated by two guiding principles in statistical modeling:

1. The *effect sparsity principle* states that only a small proportion of the factors in an experiment will have significant effects.
2. The *hierarchical ordering principle* states that lower-order interactions (including primary effects) are more important that higher-order interactions.

Both principles become "more true" as the number of factors increases.

Why do fractional factorial designs work?

Fractional designs are motivated by two guiding principles in statistical modeling:

1. The *effect sparsity principle* states that only a small proportion of the factors in an experiment will have significant effects.
2. The *hierarchical ordering principle* states that lower-order interactions (including primary effects) are more important that higher-order interactions.

Both principles become "more true" as the number of factors increases.

Fractional designs rely on an assumption that

$$|\text{low-order effects}| \gg |\text{high-order effects}|$$

Example: the 2^{4-1} fractional design

We begin with a 2^3 full factorial design (the *base design*).

I	A	B	C	AB	AC	BC	ABC
$+$	$-$	$-$	$-$	$+$	$+$	$+$	$-$
$+$	$+$	$-$	$-$	$-$	$-$	$+$	$+$
$+$	$-$	$+$	$-$	$-$	$+$	$-$	$+$
$+$	$+$	$+$	$-$	$+$	$-$	$-$	$-$
$+$	$-$	$-$	$+$	$+$	$-$	$-$	$+$
$+$	$+$	$-$	$+$	$-$	$+$	$-$	$-$
$+$	$-$	$+$	$+$	$-$	$-$	$+$	$-$
$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$

Example: the 2^{4-1} fractional design

We begin with a 2^3 full factorial design (the *base design*).

I	A	B	C	AB	AC	BC	ABC
$+$	$-$	$-$	$-$	$+$	$+$	$+$	$-$
$+$	$+$	$-$	$-$	$-$	$-$	$+$	$+$
$+$	$-$	$+$	$-$	$-$	$+$	$-$	$+$
$+$	$+$	$+$	$-$	$+$	$-$	$-$	$-$
$+$	$-$	$-$	$+$	$+$	$-$	$-$	$+$
$+$	$+$	$-$	$+$	$-$	$+$	$-$	$-$
$+$	$-$	$+$	$+$	$-$	$-$	$+$	$-$
$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$

This design is orthogonal and the design matrix is full rank. We can't add a column for D without messing up these properties.

Confounding

If we choose to set D equal to an existing column in our design, we have *confounded* it. Since the factors vary together in our design we cannot estimate their effects separately.

For example, let D=ABC. Then

$$\beta_{\mathrm{D|ABC}} = \beta_{\mathrm{D}} + \beta_{\mathrm{ABC}}$$

However, by the hierarchical ordering principle we expect that $\beta_{\mathrm{ABC}} \approx 0 \ll \beta_{\mathrm{D}}$, so

$$\beta_{\mathrm{D|ABC}} = \beta_{\mathrm{D}}$$

The 2^{4-1} fraction design (with D=ABC)

We replace the highest interaction (ABC) with D and fill in the rest of the interactions.

I	A	B	C	AB	AC	BC	D= ABC	AD	BD	CD	ABC	BCD	ABD	ACD	ABCD
$+$	$-$	$-$	$-$	$+$	$+$	$+$	$-$	$+$	$+$	$+$	$-$	$-$	$-$	$-$	$+$
$+$	$+$	$-$	$-$	$-$	$-$	$+$	$+$	$+$	$-$	$-$	$+$	$+$	$-$	$-$	$+$
$+$	$-$	$+$	$-$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$-$	$+$	$+$
$+$	$+$	$+$	$-$	$+$	$-$	$-$	$-$	$-$	$-$	$+$	$-$	$+$	$-$	$+$	$+$
$+$	$-$	$-$	$+$	$+$	$-$	$-$	$+$	$-$	$-$	$+$	$+$	$-$	$+$	$-$	$+$
$+$	$+$	$-$	$+$	$-$	$+$	$-$	$-$	$-$	$+$	$-$	$-$	$+$	$+$	$-$	$+$
$+$	$-$	$+$	$+$	$-$	$-$	$+$	$-$	$+$	$-$	$-$	$-$	$-$	$+$	$+$	$+$
$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$

The 2^{4-1} fraction design (with D=ABC)

We replace the highest interaction (ABC) with D and fill in the rest of the interactions.

I	A	B	C	AB	AC	BC	D= ABC	AD	BD	CD	ABC	BCD	ABD	ACD	ABCD
+	−	−	−	+	+	+	−	+	+	+	−	−	−	−	+
+	+	−	−	−	−	+	+	+	−	−	+	+	−	−	+
+	−	+	−	−	+	−	+	−	+	−	+	−	−	+	+
+	+	+	−	+	−	−	−	−	−	+	−	+	−	+	+
+	−	−	+	+	−	−	+	−	−	+	+	−	+	−	+
+	+	−	+	−	+	−	−	−	+	−	−	+	+	−	+
+	−	+	+	−	−	+	−	+	−	−	−	−	+	+	+
+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

All of the variables are now confounded:

$$\begin{array}{ll} A + BCD & AB + CD \\ B + ACD & AC + BD \\ C + ABD & AD + BC \\ D + ABC & I + ABCD \end{array}$$

Generator Algebra

Filling in the entire design is impractical, especially for large designs. We can identify the *confounding pattern* (or *alias structure*) using a special type of algebra.

Generator Algebra Axioms

- $\mathrm{XX} = \mathrm{X}^2 = \mathrm{I}$ for any factor X.
- $\mathrm{IX} = \mathrm{X}$ for any factor X.
- Multiplication commutes, associates, and distributes.

Generating the 2^{4-1} design

We start with the *generator* of the design — the replacement we made to the base design.

$$\mathrm{D} = \mathrm{ABC}$$
$$\mathrm{D}^2 = \mathrm{ABCD}$$
$$\mathrm{I} = \mathrm{ABCD}$$

Generating the 2^{4-1} design

We start with the *generator* of the design — the replacement we made to the base design.

$$\begin{aligned} \mathrm{D} &= \mathrm{ABC} \\ \mathrm{D}^2 &= \mathrm{ABCD} \\ \mathrm{I} &= \mathrm{ABCD} \end{aligned}$$

This last statement (I=ABCD) is called the *defining relation* for the design with generator D=ABC.

Generating the 2^{4-1} design (continued)

With the defining relation (I=ABCD) we can compute the confounding for any variable.

Generating the 2^{4-1} design (continued)

With the defining relation (I=ABCD) we can compute the confounding for any variable.

For A:

$$\begin{aligned} \mathrm{A(I)} &= \mathrm{A(ABCD)} \\ \mathrm{A} &= \mathrm{A^2BCD} \\ &= \mathrm{IBCD} \\ &= \mathrm{BCD} \end{aligned}$$

Generating the 2^{4-1} design (continued)

With the defining relation (I=ABCD) we can compute the confounding for any variable.

For A:

$$\begin{aligned} \mathrm{A(I)} &= \mathrm{A(ABCD)} \\ \mathrm{A} &= \mathrm{A^2BCD} \\ &= \mathrm{IBCD} \\ &= \mathrm{BCD} \end{aligned}$$

For the interaction CD:

Generating the 2^{4-1} design (continued)

With the defining relation (I=ABCD) we can compute the confounding for any variable.

For A:

$$\begin{aligned} \mathrm{A(I)} &= \mathrm{A(ABCD)} \\ \mathrm{A} &= \mathrm{A^2BCD} \\ &= \mathrm{IBCD} \\ &= \mathrm{BCD} \end{aligned}$$

For the interaction CD:

$$\begin{aligned} \mathrm{CD(I)} &= \mathrm{CD(ABCD)} \\ \mathrm{AB} &= \mathrm{ABC^2D^2} \\ &= \mathrm{AB} \end{aligned}$$