Contents

7

Fractional factorials

Even at just two levels, studying k factors brings in 2^k levels. It is a curse of dimension. We cannot necessarily afford that. And maybe we don't have to. Fractional factorials look at k factors of 2 levels each using fewer than 2^k runs. We will study k factors using only $N = 2^{k-p}$ runs for some integer $1 \leqslant p < k$. These are called ***fractional factorials***. The name stems from the fact that they are fractional replicates of 2^k. It is as if we are doing n replicates for $n = 1/2$.

There are many good references for this topic, such as Wu and Hamada (2011) and Box et al. (1978) and Montgomery (1997). The last two examples are not the most recent editions of those books. I prefer the cited editions. For instance later versions of the book by Montgomery seem less suitable for a graduate level class, though they are better suited for their target audience than the cited edition.

We will study k binary inputs in 2^{k-r} runs. This necessarily involves aliasing/confounding relationships among the 2^k estimands of interest. When the phenomenon is dominated by main effects and low order interactions and the high order interactions are small, then aliasing the high order interactions with each other makes sense. We still get to learn about the important quantities at low cost.

One problem with a 2^k factorial experiment is that as k grows larger, most of the degrees of freedom are used up in interactions of order about $k/2$. We have much more interest in learning about main effects and low order interactions. There's a sense that they're more likely to be large and they are also going to be simpler to interpret if we can estimate them. It can be puzzling to wonder why degrees of freedom are so important. The explanation is that each extra degree of freedom costs us one more data point. See Table 7.1 which illustrates the point with $k = 7$ and runs costing \$1000 each. We end up spending \$8000

Interaction order:	0	1	2	3	4	5	6	7
Degrees of freedom:	$\binom{7}{0}$	$\binom{7}{1}$	$\binom{7}{2}$	$\binom{7}{3}$	$\binom{7}{4}$	$\binom{7}{5}$	$\binom{7}{6}$	$\binom{7}{7}$
Degrees of freedom:	1	7	21	35	35	21	7	1
Cost	\$1k	\$7k	\$21k	\$35k	\$35k	\$21k	\$7k	\$1k

Table 7.1: For a hypothetical 2^7 experiment where each run costs \$1000, the table shows how much of the budget goes to interactions of each size.

on the grand mean and main effects along with \$59,000 on interactions of order three and higher.

7.1 Half replicates

Let's do 2^{k-1} runs, i.e., half of the full 2^k experiment. We will get confounding. To make the confounding minimally damaging, we will confound "done" versus "not done" with the k-fold interaction.

For $k = 3$ we could do the 4 runs at the high level of ABC:

$$\begin{array}{c} \\ a \\ b \\ c \\ abc \end{array} \begin{array}{c} \begin{array}{cccccccc} I & A & B & AB & C & AC & BC & ABC \end{array} \\ \begin{bmatrix} + & + & - & - & - & - & + & + \\ + & - & + & - & - & + & - & + \\ + & - & - & + & + & - & - & + \\ + & + & + & + & + & + & + & + \end{bmatrix} \end{array}$$

We see right away that the intercept column I is identical to the one for ABC. We say that the grand mean is ***aliased*** with ABC. It is also confounded with ABC. We then get

$$\mathbb{E}(\hat{\mu}) = \mu + \alpha_{ABC}.$$

In our half replicate, $I = ABC$ holds. If we multiply both sides by A, we get $A \times I = A \times ABC$. Of course $A \times I = A$ and $A \times ABC = A^2BC = BC$. It follows that $A = BC$ so the main effect for A is aliased with the BC interaction. Therefore

$$\mathbb{E}(\hat{\alpha}_A) = \alpha_A + \alpha_{BC}.$$

By the same argument $B = AC$ and $C = AB$. If we do a 2^{k-1} experiment running everything at the high level of the k-fold interaction then every effect will be aliased with one other, the one that complements it's set of variables.

The set of all eight runs looks like this:

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

We could do either the top half, with $ABC = I$ or the bottom half with $ABC = -I$. That is $ABC = \pm I$ give a 2^{3-1} experiment. If we use the bottom half we get

$$\mathbb{E}(\hat{\alpha}_A) = \alpha_A - \alpha_{BC}.$$

If any of the k factors in our 2^{k-1} factorial experiment are "null" then we get a full 2^k factorial experiment in the remaining ones. Practially, null means "relatively null" in comparison to some larger effects.

Geometrically the sampled points of a 2^{3-1} experiment lie on 4 corners of the cube $\{-1, 1\}^3$ (or if you prefer $[0, 1]^3$). If we draw in the edges on each face of the cube, no pair of sampled points share an edge.

A 2^{k-1} design looks like one block of a blocked 2^k where the blocks are defined by high and low levels of the k-factor interaction. If we have just done the $ABC \cdots Z = I$ block we could analyze it as a 2^{k-1} design and perhaps decide that we have learned what we need and don't have to do the other block.

7.2 Catapult example

The lightning calculator Cat-100 catapult is described here `http://www.qualitytng.com/cat-100-catapult/`. In class we looked at data from a 2^{5-1} experiment in it. The data are shown in Table 7.2. Distance is in centimeters. It appears that most of the energy from the surgical tubing goes into moving the wooden arm, so the projectiles remain safe for classroom use (and are not suitable as siege weapon).

Table 7.3 shows those same data in standard (Yates) order. That is (1), a, b, ab, et cetera for four of the five variables. The reason to do this is that the data were analyzed by Yates' algorithm which is not really necessary on a problem this small.

Yates' algorithm pair off the data into $n/2$ consecutive pairs. It takes sums within those $n/2$ pairs and places them above differences. If we do that k times in a full factorial or $k-1$ times in a fractional factorial we get a column with main effects and interaction estimates in the standard order (scaled up by n for

	Front	Back	Fixed	Moving	Bucket	Dist
1	1	-1	-1	1	1	210.3
2	1	1	1	1	1	343.0
3	1	-1	-1	-1	-1	50.0
4	-1	-1	1	1	1	263.5
5	-1	-1	-1	1	-1	134.5
6	-1	-1	-1	-1	1	94.5
7	-1	1	-1	1	1	310.8
8	1	1	-1	-1	1	94.8
9	-1	1	-1	-1	-1	91.5
10	1	1	-1	1	-1	168.5
11	-1	1	1	-1	1	277.4
12	1	1	1	-1	-1	145.5
13	1	-1	1	-1	1	157.5
14	-1	1	1	1	-1	266.5
15	-1	-1	1	-1	-1	120.5
16	1	-1	1	1	-1	166.5

Table 7.2: Experimental output from a catapult experiment.

	Front	Back	Fixed	Moving	Bucket	Dist
6	-1	-1	-1	-1	1	94.5
3	1	-1	-1	-1	-1	50.0
9	-1	1	-1	-1	-1	91.5
8	1	1	-1	-1	1	94.8
15	-1	-1	1	-1	-1	120.5
13	1	-1	1	-1	1	157.5
11	-1	1	1	-1	1	277.4
12	1	1	1	-1	-1	145.5
5	-1	-1	-1	1	-1	134.5
1	1	-1	-1	1	1	210.3
7	-1	1	-1	1	1	310.8
10	1	1	-1	1	-1	168.5
4	-1	-1	1	1	1	263.5
16	1	-1	1	1	-1	166.5
14	-1	1	1	1	-1	266.5
2	1	1	1	1	1	343.0

Table 7.3: Catapult data in Yates' order.

the main effect and $n/2$ for the others). For $k = 2$ we get

$$\begin{array}{llllll} y_{(1)} & y_a + y_{(1)} & y_{ab} + y_b + y_a + y_{(1)} & \rightarrow & 4\hat{\mu} \\ y_a & y_{ab} + y_b & y_{ab} - y_b + y_a - y_{(1)} & \rightarrow & 2\hat{\alpha}_A \\ y_b & y_a - y_{(1)} & y_{ab} + y_b - y_a - y_{(1)} & \rightarrow & 2\hat{\alpha}_B \\ y_{ab} & y_{ab} - y_b & y_{ab} - y_b - y_a + y_{(1)} & \rightarrow & 2\hat{\alpha}_{AB} \end{array}$$

and for $k = 2^{30}$ it would only take 30 of these operations to compute 2^{30} effects of interest (most likely in a computer experiment). In a fractional factorial we need to account for the aliasing.

Here it is (or at least part of it) for $k = 3$:

$$\begin{array}{llll} y_{(1)} & y_a + y_{(1)} & y_{ab} + y_b + y_a + y_{(1)} & y_{ab} + y_b + y_a + y_{(1)} y_{ab} + y_b + y_a + y_{(1)} \\ y_a & y_{ab} + y_b & y_{abc} + y_{bc} + y_{ac} + y_c & y_{abc} - y_{bc} + y_{ac} - y_c + y_{ab} - y_b + y_a - y_{(1)} \\ y_b & y_{ac} + y_c & y_{ab} - y_b + y_a - y_{(1)} & \\ y_{ab} & y_{abc} + y_{bc} & y_{abc} - y_{bc} + y_{ac} - y_c & \text{et cetera} \\ y_c & y_a - y_{(1)} & y_{ab} + y_b - y_a - y_{(1)} & \\ y_{ac} & y_{ab} - y_b & y_{abc} + y_{bc} - y_{ac} - y_c & \\ y_{bc} & y_{ac} - y_c & y_{ab} - y_b - y_a + y_{(1)} & \\ y_{abc} & y_{abc} - y_{bc} & y_{abc} - y_{bc} - y_{ac} + y_c & \end{array}$$

The digression on Yates' algorithm demystifies the axis label in Figure 7.1. There we see that main effects for 'back stop', 'fixed arm', 'moving arm' and 'bucket' all have positive values clearly separated from the noise level. The effect for 'front stop' appears to be negative but is not clearly separated from the others. The reference line is based on a least squares fit to the 11 smallest effects (F and the interactions). Effects are named after their lowest alias.

In this instance there was a clear break between large effects and most small effects and some reasonable doubt as to whether F belongs with the large or the small ones. In many examples in the literature some main effects end up in the bulk of small effects and a handful of two factor interactions show large effects. Usually one or both of the interacting factors appears also as a large main effect.

One reason to be interested in statistical insignificance is that when it happens we clearly do not know the sign of the effect, even if we're certain that an exact zero cannot be true. Statistical significance makes it more reasonable that you can confidently state the sign of the effect. There can however be doubt about the sign if the confidence interval for the effect has an edge too close to zero. This could happen in a low power setting. See Owen (2017).

Here is a naive regression analysis of the data.

```
Coefficients:
              Value Std. Error t value
(Intercept)  180.96      7.441  24.318
      Front  -13.94      7.441  -1.874
       Back   31.29      7.441   4.205
      Fixed   36.59      7.441   4.918
```

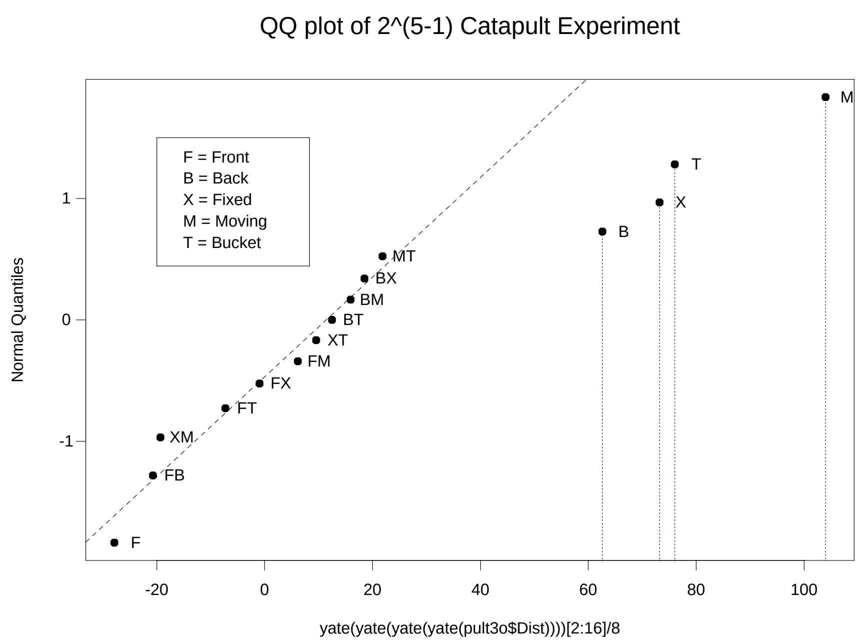

Figure 7.1: QQ plot of catapult experiment.

```
    Moving   51.99      7.441   6.987
    Bucket   38.02      7.441   5.109

Residual standard error: 29.77 on 10 degrees of freedom
Multiple R-Squared: 0.9233
```

It would be interesting to see how modern post-selective inference methods might work with factorial and fractional factorial models.

7.3 Quarter fractions

For a 2^{k-2} experiment we can set two combinations of effects equal to I. Or we could set one or both of them equal to $-I$. For $k = 6$, if we have set $ABCDEF = I$ we might first consider (bad idea) also setting $ABCDE = I$. This is bad because then the product of $ABCDEF$ and $ABCDE$ becomes $I^2 = I$ and this product is

$$ABCDEF \times ABCDE = F.$$

We certainly don't want to alias one of our main effects to the grand mean.

A better choice is to set $ABCD = I$ and $CDEF = I$. We then also get $ABCD \times CDEF = ABEF = I$. Then we find that

$$A = A \times (ABCD, CDEF, ABEF) = BCD = ACDEF = BEF$$

and so we get

$$\begin{aligned}
\mathbb{E}(\hat{\alpha}_A) &= \alpha_A + \alpha_{ABCD} + \alpha_{CDEF} + \alpha_{ABEF}, \quad \text{and similarly} \\
\mathbb{E}(\hat{\alpha}_{AB}) &= \alpha_{AB} + \alpha_{CD} + \alpha_{ABCDEF} + \alpha_{EF}, \quad \text{and} \\
\mathbb{E}(\hat{\alpha}_{ABC}) &= \alpha_{ABC} + \alpha_D + \alpha_{ABDEF} + \alpha_{CEF}.
\end{aligned}$$

In a quarter replicate, any effect that we compute estimates a combination of four effects. One is the desired effect. The other three are aliases. The aliases above all have a positive sign. If we had set $ABCD = -I$ then $\mathbb{E}(\hat{\alpha}_A)$ would include some aliased effects with negative signs.

Here are some of the aliasing patterns in this design:

$$\begin{aligned}
I &= ABCD = CDEF = ABEF \\
AB &= CD = ABCDEF = EF \\
AC &= BD = ADEF = CBEF \\
AD &= BC = AEF = BDEF \\
AE &= BCDE = ACDF = BF \\
AF &= BCDF = ACDE = BE
\end{aligned}$$

The cited books have tables of design choices for fractional factorial experiments. Those tables can run to several pages. Probably nobody actually memorizes or even uses them all but it is good to have those choices. One of the key quantities in those designs is the "resolution" that we discuss next.

7.4 Resolution

The ***resolution*** of a fractional factorial experiment is the number of letters in the shortest word in the defining relation. A defining relation is an expression like $ABCD = \pm I$. In this case the ***word*** is ABCD (not I!) and it has length 4. The quarter fraction example from the previous section had defining relations $ABCD = I$, $CDEF = I$ and $ABEF = I$, so the shortest one is 4. In a 2^{k-1} fraction with defining relation setting the k-factor interaction equal to $\pm I$ the shortest work length is k.

When there are k factors of two levels each and we have 2^{k-p} runs in a fractional factorial of resolution R, then the design is labeled 2_R^{k-p}. Resolution is conventionally given in Roman numerals. The three most important ones are III, IV and V.

For resolution $R = III$, no main effects are aliased/confounded with each other. Some main effects are confounded with two factor interactions.

For resolution $R = IV$, no main effects are confounded with each other, no main effects are confounded with any two factor interactions, and some two factor interactions are confounded with each other.

For esolution $R = V$, no main effects are confounded with each other, no main effects are confounded with two factor interactions and no two factor interactions are confounded with each other. However, some main effects are confounded with four factor interactions, and some two factor interactions confounded with three factor interactions.

Table 7.4 has an informal summary of what these resolutions require. Resolution V has the least confounding but requires the most expense. Resolution III has the least expense but could be misleading if we do not have enough factor sparsity. Resolution IV is a compromise but it requires careful thought. Some but not all of the two factor interactions may be aliased with others. We can check the tables or defining relations to see which they are. If we have good knowledge or guesses ahead of time we can keep the interactions most likely to be important unconfounded with each other. Similarly, after the experiment a better understanding of the underlying science would help us guess which interaction in a confounded pair might actually be most important.

For a specific pattern 2_R^{k-p} there can be multiple designs and they are not all equally good. For instance with $R = IV$ we would prefer a design with the smallest number of aliased $2FI$'s (***two factor interactions***). If there's tie we would break it in favor of the design with the fewest aliased $3FI$s. Carrying on until the tie is broken we reach a ***minimum aberration*** design. An investigator would of course turn to tables of minimum aberration designs constructed by researchers who specialize in this.

For resolution R	you need to be	because you need
III	lucky	few significant effects, all of low order
IV	smart	to untangle confounded two factor inter.s
V	rich	the largest sample size

Table 7.4: Informal synopsis of what we you need for the three most common resolutions.

For the 2^{k-1} experiment with the k-fold interaction aliased to $\pm I$ we find that $R = k$. So $k = 3$ gives 2^{3-1}_{III}, $k = 4$ gives 2^{4-1}_{IV} and $k = 5$ gives 2^{5-1}_{V}.

There is a projection property for resolution R. If we select any $R-1$ factors then we get all 2^{R-1} possible combinations the same number of times. That has to be $2^{k-p}/2^{R-1}$ times, that is $2^{k-p-R+1}$ times.

7.5 Overwriting notation

We need to figure out how to actually conduct our 2^{k-p} experiment. For a 2^{4-1} we can get a table of runs in factors A, B, and C. Then we replace/overwrite the ABC column by D, like this

$$
\begin{array}{c}
\\ a \\ b \\ c \\ abc \\ bc \\ ac \\ ab \\ (1)
\end{array}
\begin{array}{cccccccc}
I & A & B & AB & C & AC & BC & D{=}ABC \\
+ & + & - & - & - & - & + & + \\
+ & - & + & - & - & + & - & + \\
+ & - & - & + & + & - & - & + \\
+ & + & + & + & + & + & + & + \\
+ & - & + & - & + & - & + & - \\
+ & + & - & - & + & + & - & - \\
+ & + & + & + & - & - & - & - \\
+ & - & - & + & - & + & + & -
\end{array}
$$

Now $AD = BC$ as before (check that for yourself).

The more recent edition of Box Hunter and Hunter (Box et al., 2005) describes a ***nodal design***. It has $n = 16$ runs. You could analyze 15 effects, A, B, C, D, and interactions up to ABCD.

For 2^{5-1}_{V} they alias a fifth effect to ABCD. For 2^{8-4}_{IV} they alias four more effects as follows $L = ABC$, $M = ABD$, $N = ACD$ and $O = BCD$. For 2^{15-11}_{III} they alias effects as follows $E = AB$, $F = AC$, $G = AD$, $H = BC$, $J = BD$ and $K = CD$. They skip the letter I because it can mean 'intercept'.

7.6 Saturated designs

Using 2^{15-11}_{III} we can estimate a grand mean and 15 main effect, but no interactions and we get no degrees of freedom for error. Or if we like we could study only $15 - r$ effects and get r degrees of freedom for error.

There is a special setting where we can dispense with any concern over interactions. The classic example is when we are weighing objects. The weight of a pair of objects is simply the sum of their weights, with interactions being zero. Experimental designs tuned to a setting where interactions are known to be impossible are called ***weighing designs***.

A special kind of weighing design is known as ***Plackett-Burman*** designs. They exist when $n = 4m$ for (most) values of m, and so n does not have to be a power of 2. We will see them later as Hadamard designs.

7.7 Followup fractions

Suppose we have done a 1/4 fraction. Then we can follow up with a second 1/4 in more than one different way. We would ordinarily want to treat that second fraction as a block.

Suppose that factor A looks really important in the first 1/4 that we do. Maybe the aliasing pattern is

$$\mathbb{E}(\hat\alpha_A) = \alpha_A + \alpha_{BC} \pm \cdots .$$

In the second part of our experiment, we could flip the signs of the A assignments but leave everything else unchanged. In that second half we will have

$$\mathbb{E}(\hat\alpha_A) = \alpha_A - \alpha_{BC} \mp \cdots .$$

If we pool the two parts of our experiment, then

$$\begin{aligned}\hat\alpha_A &= \frac{1}{2}\bigl(\text{first expt } \hat\alpha_A + \text{second expt } \hat\alpha_A\bigr)\\ \mathbb{E}(\hat\alpha_A) &= \frac{1}{2}\bigl((\alpha_A + \alpha_{BC} + \cdots) + (\alpha_A - \alpha_{BC} - \cdots)\bigr) = \alpha_A.\end{aligned}$$

We get rid of any aliasing for A. Of course if B had looked important after the first 1/4 we could have flipped B. We would not have to know which one is important before doing the first 1/4 of the factorial.

A second kind of followup is the ***foldover***. We flip the signs of all the factors. (Note that we cannot and don't attempt to flip the intercept.) Foldovers deconfound an main effect from any 2FI that it was confounded with. To see this, suppose that $A = BC$ in the first experiment. Flipping signs of everything makes $(-A) = (-B)(-C)$. That means $-A = BC$ or $A = -BC$. Then

$$\mathbb{E}(\hat\alpha_A) = \frac{1}{2}\Bigl[(\alpha_A + \alpha_{BC} + \cdots) + (\alpha_A - \alpha_{BC} + \cdots)\Bigr] = \alpha_A + \cdots .$$

If however $A = BCD$ then flipping the signs makes $(-A) = (-B)(-C)(-D) = -BCD$ so A still equals BCD.

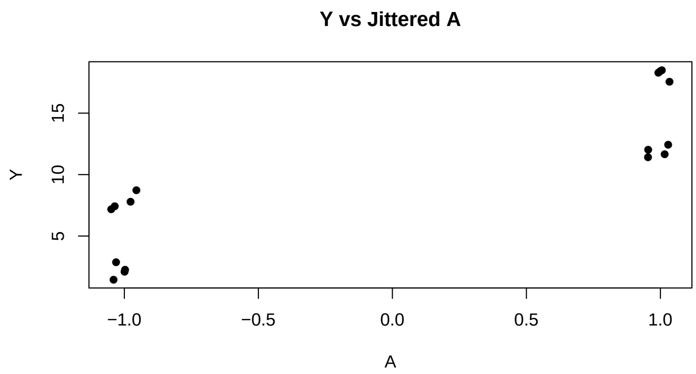

Figure 7.2: Exaggerated plot when A is a strong main effect and there is a second strong main effect.

7.8 More data analysis

Figure 7.2 shows what we might see in a plot of Y versus an important factor A. The bimodal pattern is what we would see if there were a second important factor. If A were a relatively very weak (almost null) factor then we would see almost the same point pattern at each side of the plot. Figure 7.3 shows what we might see of A has a strong interaction with one or more other factors. The variance of Y depends strongly on A. A graphical way to look for such patterns is to compute $F_A = \log(\text{var}(y \mid A = 1)/\text{var}(y \mid A = -1))$ and similar things for other main effects and interactions and produce a QQ plot of them. The outliers could be factors that are involved in many interactions. This could be fooled: it would be possible to have AB place a lot of variance at the low level of A while AC places a lot at the high level of A making F_A close to zero.

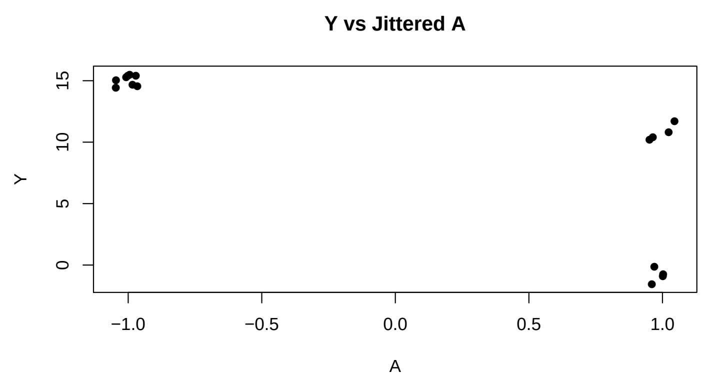

Figure 7.3: Exaggerated plot when A is a strong main effect and it has a strong interaction with one or more other factors.

Bibliography

Box, George, E., Hunter, J. S., and Hunter, W. G. (2005). *Statistics for experimenters: design, innovation, and discovery*. Wiley, New York.

Box, G. E., Hunter, W. H., and Hunter, S. (1978). *Statistics for experimenters*, volume 664. John Wiley and sons, New York.

Montgomery, D. C. (1997). *Design and analysis of experiments*. John wiley & sons, 4 edition.

Owen, A. B. (2017). Confidence intervals with control of the sign error in low power settings. Technical report, Stanford University.

Wu, C. J. and Hamada, M. S. (2011). *Experiments: planning, analysis, and optimization*, volume 552. John Wiley & Sons.