Chapter 6

Confounding

Confounding

- Confounding: A design technique for factorial experiments where block size is less than # treatments.
- A confounding variable is associated with the exposure and it affects the outcome, but it is not an intermediate link in the chain of causation between exposure and outcome.

- In confounding cannot completely replicate factorial design in each block
- Block size less than # treatment combinations in a replicate
- Design technique: Confound higher order interactions with blocking
- Examine simplified:
  - Given 2^k factorial design
    - Develop design with 2p (p < k) incomplete sets
    - Have 2, 4, 8, … incomplete blocks

- Different allocation confounds different effect
- Generally, if confound, then confound the highest-order interaction

  2-factor: confound ab

  3-factor: confound abc

  4-factor: confound abcd

  …

How to deal with confounding

At the design stage

1) Randomisation into groups should control for known and unknown confounders, used particularly in clinical trials

2) Restrict entry into the study

3) Match study participants on predicted confounding variables

At the analysis stage

1)Stratification. Measure the strength of the association at different levels of the confounder.

2)Mathematical modelling. Particularly useful if it is necessary to adjust for multiple confounders.

NOTE: we cannot adjust for confounding at the analysis stage if you have not collected data on the confounder

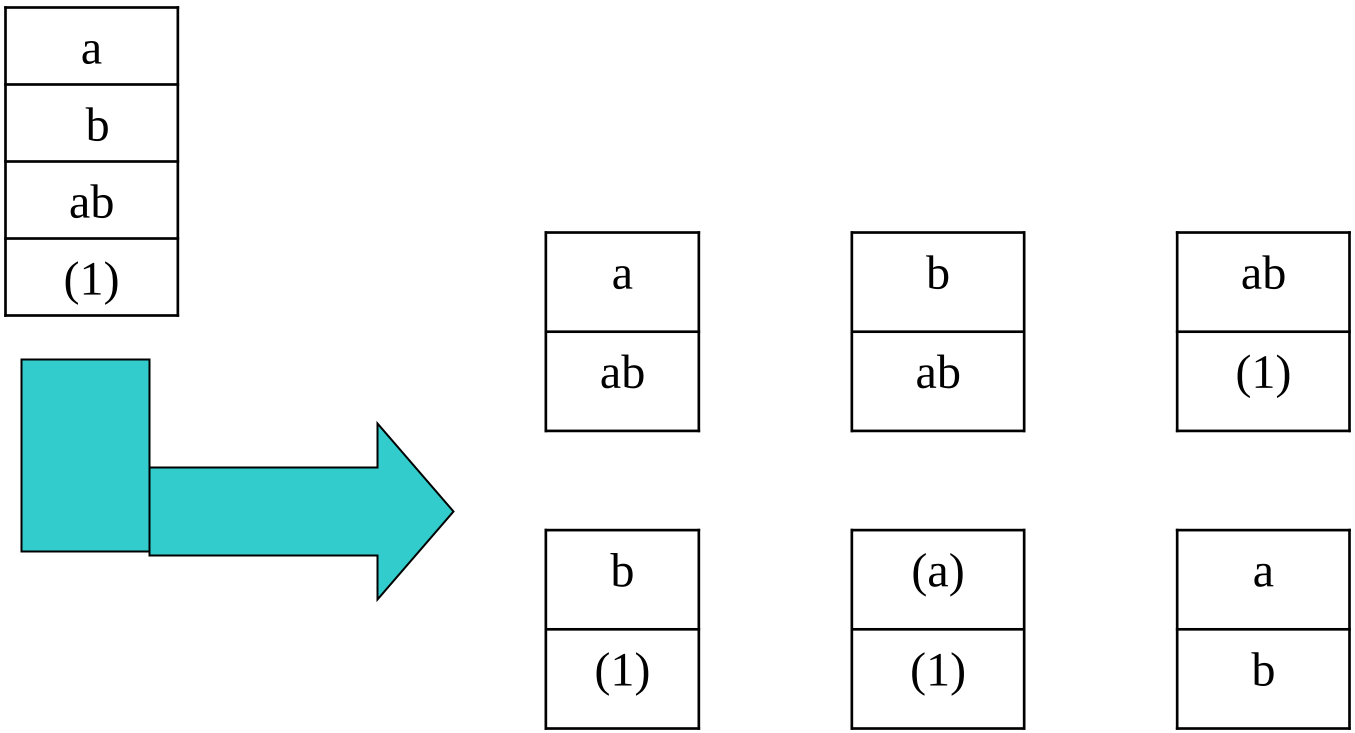
Example

Suppose we have 2X2experiment Conduct the Confounding

1)Numbers of factors =2 (A,B)
2)Numbers of levels=2 (A : a_0,a_1 B : b_0,b_1)
3)Numbers of Treat Combination (T.C.)

B \ A	b_0	b_1
a_0	a_0b_0 (1)	a_0b_1 (b)
a_1	a_1b_0 (a)	a_1b_1 (ab)

4) Numbers of effects =3 (A,B,AB)



Effect of A = (a-1)(b+1) = ab+a-b-1 = (a+ab)-(b+(1))

Effect of B = (a+1)(b-1)= ab –a+b-1 = (b+ab)-(a +(1))

Effect of AB = (a-1)(b-1) = ab –b-a+1= (ab+ (1))-(b+a)

Advantages

- conceptually straightforward
  - handles difficult to quantities variables
  - can also be used in analysis phase

Disadvantages

- may limit number of eligible subjects
- inefficient to screen subjects, then not enroll
- “residual confounding” may persist if restriction categories not sufficiently narrow (e.g. “20 to 30 years old” might be too broad)
- limits generalizability (but don’t worry too much about this)
- not possible to evaluate the relationship of interest at different levels of the restricted variable (i.e. cannot assess interaction)

Type of Confounding

- Complete Confounding
- Partial Confounding

Experiment can be repeated as many times If the built-in effect all times the same effect in this case is called (Complete Confounding), but if there is no confound of the same effect in each iteration is called the confound in this case (Partial Confounding)

Complete Confounding Example

a 2^3 experiment in blocks of size 4 (4 replicates). *ABC* interaction is **confounded** in **1st** replication. *ABC* interaction is **confounded** in **2nd** replication. *ABC* interaction is **confounded** in **3rd** replication. ABC interaction is confounded in 4th replication

Repli(1)	Repli(2)	Repli(3)	Repli(4)
c 920	ac 733	bc 799	a 835
abc 945	ab 851	(1) 861	c 656
a 991	bc 863	ab 806	b 764
b 1027	(1) 794	ac 702	abc 750

ab 793	b 884	a 881	(1) 706
ac 757	c 796	c 782	ac 715
(1) 642	abc 916	abc 813	ab 580
bc 734	a 635	b 915	bc 638

ABC	ABC	ABC	ABC

$$C.F = \frac{(25504)^2}{abcr = 2 \times 2 \times 2 \times 4}$$

$$\text{SStotal} = 642^2 + 794^2 + \ldots . + 750^2 - \frac{(25504)^2}{32}$$

Treat Comb.	1	2	3	4	Treat Total
(1)	642	794	781	706	3003
a	991	635	881	835	3342
b	1027	884	915	784	3610
c	920	796	782	656	3154
ab	793	851	806	590	3030
ac	757	733	702	715	2907
bc	734	863	799	638	3034
abc	945	916	813	750	3424
Repli Total	6809	6472	6559	5664	25504

Blocks	1	2	3	4	Total
1	3883	3231	3391	3025	13530
2	2926	3241	3168	2639	11974
Total	6809	6472	6559	5664	25504

$$SS(Blocks) = \frac{3883^2 + 3231^2 + ... + 2639^2}{4} - \frac{(25504)^2}{32}$$

$$= 23147.650$$

$$SS\,\mathrm{Re}\,pli. = \frac{6809^2 + 6472^2 + ... + 5664^2}{8} - \frac{(25504)^2}{32}$$

$$= 9214.225$$

$$SS(ABC = Block\ over\,\mathrm{Re}\,pli.) = \frac{11974^2 + 13530^2}{16} - \frac{(25504)^2}{32}$$

$$= 75660.750$$

$$SS(Block \times \mathrm{Re}\,pli.) = SS(Block\ /\,\mathrm{Re}\,pli.) - SS(\mathrm{Re}\,pli.) - SS(ABC = Block\ over\,\mathrm{Re}\,pli.)$$

$$= 23147.650 - 9214.225 - 75660.50$$

$$= 63673.70$$

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

$$SSA = \frac{1}{2K}[(a + ab + ac + abc) - ((1) + b + c + bc)]^2$$

$$= \frac{1}{2(16)}[(3342 + 3030 + 2907 + 3424) - (3003 + 3610 + 3154 + 3034)]^2$$

$$= 300.125$$

$$SSB = \frac{1}{2K}[(b + ab + bc + abc) - ((1) + a + c + ac)]^2$$

$$= \frac{1}{2(16)}[(13098) - (12406)]^2$$

$$= 14964.5$$

$$SSC = \frac{1}{2K}[(c + ac + bc + abc) - ((1) + a + b + ab)]^2$$

$$= \frac{1}{2(16)}[(12519) - (12985)]^2$$

$$= 6786.125$$

$$SSAB = \frac{1}{2K}[(c + ac + bc + abc) - ((1) + a + b + ab)]^2$$

$$= \frac{1}{2(16)}[(12519) - (12985)]^2$$

$$= 6786.125$$

$$SSAC = \frac{1}{2K}[((1) + b + ac + abc) - (a + ab + c + bc)]^2$$

$$= \frac{1}{2(16)}[(12944) - (12560)]^2$$

$$= 4608$$

$$SSBC = \frac{1}{2K}[((1) + a + bc + abc) - (b + ab + c + ac)]^2$$

$$= \frac{1}{2(16)}[(12803) - (12701)]^2$$

$$= 325.125$$

$$SSE = SST - SS(blocks) - SSA - SSB - SSC - SSAB - SSAC - SSABC$$

$$= 362276 - 231476.500 - 300.125 - 14964.5 - 6786.125 - 4608 - 325.125$$

$$= 101329.5$$

S.O.V	D.F	S.S	M.S
Blocks	7	231476.5	
Replicates	(3)	(92142.250)	
Blocks over reps=ABC	(1)	(75660.5)	
Blocks x Reps	(3)	(63673.750)	
A	1	300.125	300.125
B	1	14964.5	14964.5
C	1	6786.125	6786.125
AB	1	2485.125	2485.125
AC	1	4608	4608
BC	1	326.125	326.125
Error	18	101329.5	5629.472
Total	31	362276	

Partial Confounding Example

2^3 experiment in blocks of size 4 (4 replicates). *BC* interaction is **confounded** in **1st** replication. *AC* interaction is **confounded** in **2nd** replication. *ABC* interaction is **confounded** in **3rd** replication. AB interaction is confounded in **4th** replication

Repl.(1)	Repl(2)	Repli(3)	Repl(4)
ac 782	c 656	abc 916	bc 945
b 915	a 750	a 635	b 1027
c 813	ab 835	c 884	a 991
ab 881	bc 784	b 796	ac 920
bc 799	b 638	(1) 733	ab 757
a 861	abc 715	ac 794	(1) 642
abc 702	(1) 580	ab 851	abc 793
(1) 806	ac 706	bc 863	c 734
BC	AC	ABC	AB

Treat Comb.	1	2	3	4	Treat Total
(1)	806	580	733	642	2761
a	861	750	635	991	3237
b	915	638	796	1027	3376
c	813	656	884	734	3087
ab	881	835	851	757	3324
ac	782	706	794	920	3202
bc	799	784	863	945	3391
abc	702	715	916	793	3126
Repli Total	6559	5664	6472	6809	25504

$$C.F = \frac{(25504)^2}{abcr = 2 \times 2 \times 2 \times 4}$$

$$\text{SStotal} = 806^2 + 580^2 + \ldots + 793^2 - \frac{(25504)^2}{32}$$

Blocks	1	2	3	4
1	3391	3025	3231	3883
2	3168	2639	3241	2926
Total	6559	5664	6472	6809

$$SS(Blocks) = \frac{3391^2 + 3025^2 + ... + 3241^2}{4} - \frac{(25504)^2}{32}$$

$$= 23147650$$

$$SS\,\mathrm{Re}\,pli. = \frac{6559^2 + 5664^2 + 6472^2 + 6809^2}{8} - \frac{(25504)^2}{32}$$

$$= 9214225$$

$$SS(Block\ /\ \mathrm{Re}\,pli.) = SS(Blocks) - SS\,\mathrm{Re}\,pli$$

$$= 23147650 - 9214225$$

$$= 139334250$$

$$SSA = \frac{1}{2K}[(a + ab + ac + abc) - ((1) + b + c + bc)]^2$$

$$= \frac{1}{2(16)}[(3237 + 3324 + 3202 + 3126) - (2761 + 3376 + 3087 + 3391)]^2$$

$$= 2346125$$

$$SSB = \frac{1}{2K}[(b + ab + bc + abc) - ((1) + a + c + ac)]^2$$

$$= \frac{1}{2(16)}[(3376 + 3324 + 3391 + 3126) - (2761 + 3237 + 3087 + 3202)]^2$$

$$= 27028125$$

$$SSC = \frac{1}{2K}[(c + ac + bc + abc) - ((1) + a + b + ab)]^2$$

$$= \frac{1}{2(16)}[(3087 + 3202 + 3391 + 3126) - (2761 + 3237 + 3376 + 3324)]^2$$

$$= 364500$$

Treat Comb.	BC 2+3+4	AC 1+3+4	ABC 1+2+4	AB 1+2+3
(1)	1955	2181	2028	2119
a	2376	2487	2602	2246
b	2461	2738	2580	2349
c	2274	2431	2203	2353
ab	2443	2489	2473	2567
ac	2420	2496	2408	2282
bc	2592	2607	2528	2446
abc	2424	2211	2210	2333

$$SSAB = \frac{1}{2K}[(a + b + ac + bc) - ((1) + c + ab + abc)]^2$$

$$= \frac{1}{2(12)}[(2246 + 2349 + 2282 + 2446) - (2119 + 2353 + 2567 + 2337)]^2$$

$$= 100.041$$

$$SSAC = \frac{1}{2K}[(a + c + ab + bc) - ((1) + b + ac + abc)]^2$$

$$= \frac{1}{2(12)}[(2487 + 2431 + 2489 + 2607) - (2181 + 2738 + 2496 + 2411)]^2$$

$$= 1472.666$$

$$SSBC = \frac{1}{2K}[(b + c + ab + ac) - ((1) + a + bc + abc)]^2$$

$$= \frac{1}{2(12)}[(2461 + 2274 + 2443 + 2420) - (1955 + 2376 + 2592 + 2424)]^2$$

$$= 2625.041$$

$$SSABC = \frac{1}{2K}[(a + b + c + abc) - ((1) + ab + ac + bc)]^2$$

$$= \frac{1}{2(12)}[(2461 + 2274 + 2443 + 2420) - (1955 + 2376 + 2592 + 2424)]^2$$

$$= 1040.836$$

$$SSE = SST - SS(blocks) - SSA - SSB - SSC - SSAB - SSAC - SSABC$$

$$= 95822.836$$

S.O.V	D.F	S.S	M.S
Blocks	7	231476.500	
Replicates	(3)	(92142.250)	
Blocks / Reps	(4)	(139334.250)	
A	1	2346.125	2346.125
B	1	27028.125	27028.125
C	1	364.500	364.500
AB	1	100.041	100.041
AC	1	1472.666	1472.666
BC	1	2625.041	2625.041
ABC	1	1040.166	1040.166
Error	17	95822.836	5636.637
Total	31	362276	