LMA201/LMA521: Faktorförsök

Föreläsning 3

Anders Hildeman

Innehåll

- Reducerade försöksplaner
- Generatorer
- Definierande relationer
- Ord
- Upplösning

Reducerade försöksplaner

- Varje mätning kommer med en kostnad.
  I många fall är den kostnaden så dyr att man vill minimera antalet mätningar man behöver genomföra.
- Vi har sett att ifall vi vill analysera påverkan av K faktorer så behöver vi göra 2^K antal mätningar.
- Detta blir ett ganska stort antal mätningar om vi egentligen bara är intresserade av huvudeffekter och kanske tvåfaktor-samspelseffekter.
- Kan man komma undan med att göra ett mindre antal mätningar?

- Det finns ett sätt att göra detta på som kallas "**reducerad försöksplan**".
- Att använda en reducerad försöksplan kan alltså minska kostnader men man tvingas offra någonting också.
  Vissa skattade effekter kommer nu att blandas ihop så att den skattade effekten egentligen blir summan av effekten av flera olika faktorer. Det går alltså inte att särskilja vissa effekter från varandra.
- Låt oss se hur det här går till!

Exempel

Säg att vi vill minska antal mätningar i ett 2^3 försök (alltså ett försök med 3 faktorer).

Grupp nr	A	B	C	AB	BC	AC	ABC
1	-	-	-	+	+	+	-
2	+	-	-	-	+	-	+
3	-	+	-	-	-	+	+
4	+	+	-	+	-	-	-
5	-	-	+	+	-	-	+
6	+	-	+	-	-	+	-
7	-	+	+	-	+	-	-
8	+	+	+	+	+	+	+

Välj en effekt som vi skall "offra". En effekt som vi alltså inte kommer ha möjlighet att skatta i vårt reducerade försök. Låt oss välja ABC då ett trefaktorsamspel oftast är mindre intressant än lägre ordningens effekter.

Skugga grupperna motsvarande då ABC är “hög” (eller “låg” om du hellre vill, spelar ingen roll).

Grupp nr	A	B	C	AB	BC	AC	ABC
1	-	-	-	+	+	+	-
2	+	-	-	-	+	-	+
3	-	+	-	-	-	+	+
4	+	+	-	+	-	-	-
5	-	-	+	+	-	-	+
6	+	-	+	-	-	+	-
7	-	+	+	-	+	-	-
8	+	+	+	+	+	+	+

“Klipp bort” de skuggade raderna.

Grupp nr	A	B	C	AB	BC	AC	ABC
1	-	-	-	+	+	+	-
4	+	+	-	+	-	-	-
6	+	-	+	-	-	+	-
7	-	+	+	-	+	-	-

Nu är det bara fyra mätgrupper kvar. Vi har alltså reducerat antal mätningar med hälften.

Ta bort ABC-kolumnen då vi nu inte kan skatta dess effekt ändå. Numrera även om grupperna.

Grupp nr	A	B	C	AB	BC	AC
1	-	-	-	+	+	+
2	+	+	-	+	-	-
3	+	-	+	-	-	+
4	-	+	+	-	+	-

$$l_A = \frac{\bar{y}_2 + \bar{y}_3}{2} - \frac{\bar{y}_1 + \bar{y}_4}{2}$$

$$l_{BC} = \frac{\bar{y}_1 + \bar{y}_4}{2} - \frac{\bar{y}_2 + \bar{y}_3}{2} = -l_A$$

Den skattade effekten av l_A och l_{BC} kommer alltså vara helt beroende av varandra oavsett vilket system vi faktiskt valt att studera och oavsett vilka faktorer vi valt att kalla A, B och C.

Grupp nr	A	B	C	AB	BC	AC
1	-	-	-	+	+	+
2	+	+	-	+	-	-
3	+	-	+	-	-	+
4	-	+	+	-	+	-

- Det som hänt är att vi inte längre har en representant av varje nivå på alla andra faktorer då A är hög respektive låg. Istället så visar det sig att BC alltid är hög då A är låg och tvärtom.
- Våra skattade effekter kommer alltså vara summan av flera effekter. När vi försöker räkna ut l_A räknar vi egentligen ut $l_A - l_{BC}$. Vi kan inte separera hur mycket av effekten som berodde på vilken av de två effekterna.

- Man säger att BC är ett **alias** till A (och tvärtom).
- Om vi som här valde att ta bort en effekt helt och alltså minska antalet mätningar till hälften så säger vi att vi har en 2^{3-1}-plan. Det är alltså egentligen en 2^3 plan fast vi har reducerat den med en nivå.
- Hade vi reducerat med två nivåer (och sålunda bara haft två mätningar återstående) så hade vi kallat det en 2^{3-2} plan.
- Varje plan som reduceras en nivå kommer innebära att varje effekt kommer ha exakt ett alias. För en plan reducerad två nivåer så får istället varje effekt $2^2 - 1$ alias osv.

- Vi såg hur vi kunde få en reducerad försöksplan genom att välja bort någon valfri effekt.
- Nu vet vi att vi kommer få en 2^2-plan med ett alias för varje effekt om vi reducerar en 2^3-plan med en nivå. Vi kan därför konstruera vår 2^{3-1}-plan på ett lättare sätt.

Skapande av en reducerad försöksplan

Vi kan börja med att konstruera en 2^2-plan.

Grupp nr	A	B	AB
1	-	-	+
2	+	+	+
3	+	-	-
4	-	+	-

Faktorerna från den egentliga 2^3-planen kan sedan läggas till som alias om vi säger, t.ex. att kolumn AB skall vara samma som C. Detta kommer definiera hela 2^{3-1}-planen.

- Påståendet $C = AB$ kallar vi för en **generator** eftersom detta påstående kommer generera vår reducerade försöksplan. Uppenbarligen så är AB och C alias till varandra.
- Vi kan också se att om vi multiplicerar tecknen i A-kolumnen med tecknen i AB-kolumnen så får vi B-kolumnen. Men, AB kolumnen är ju identisk med C kolumnen. Alltså måste AC ha samma tecken som B! AC och B är alltså alias till varandra!
- Detta resonemang går att generalisera.

- Ta generatorn och multiplicera både vänster och höger led med bokstäverna i vänster led. Alltså i vårt fall:

$$C \cdot C = C \cdot AB.$$

- Varje gång en och samma bokstav multipliceras med sig själv så tar den ut sig själv (tänk att varje rad-värde i C kolumnen multiplicerat med sig själv kommer blir $+1$).
- Detta leder till att

$$CC = CAB \Leftrightarrow I = CAB.$$

I här betyder "identitet" och menas en kolumn där varje rad har värdet $+1$. Det innebär att vilken kolumn som helst multiplicerat med I kommer blir sig själv.

- $I = CAB$ kallas vår 2^{3-1}-plans **definierande relation**. Hade vi valt en annan generator så hade vi fått en annan definierande relation.
- Bokstavskombinationerna till höger om likhetstecknet i den definierande relationen kallas för ett **ord**.
- Tricket är nu att vi kan ta fram alias till vilken effekt som helst genom att multiplicera effekten med ordet.
  T.ex. visste vi ju redan att B och AC är alias till varandra. Vi ser nu även att $B \cdot CAB = CA = AC$.
  Vi kan också nu ta reda på att $A \cdot CAB = CB$. Alltså är A och CB alias.

- Då vi reducerar med mer än en nivå så kommer vi ha mer än en definierande relation.
- De första definierande relationerna räknar man med hjälp av de olika generatorerna (vi måste ju definiera en generator för varje nivå vi reducerar).
- De återstående definierande relationerna får vi genom att multiplicera orden från de redan uträknade definierade relationerna med varandra i alla tänkbara kombinationer.
- Reducerar vi med två nivåer får vi alltså tre definierande relationer. Två från generatorerna och den sista genom att multiplicera ihop orden från de andra två.
- Reducera med K nivåer ger då $\sum_{k=1}^{K} \binom{K}{k} = 2^K - 1$ olika definierande relationer och således $2^K - 1$ alias.

Exempel

- Antag att vi vill se hur 5 faktorer påverkar vår storhet, y.
- Vi har inte resurserna att utföra alla 32 mätningar som skulle behövas för ett fullständigt 2^5-plan.
  Istället så reducerar vi med 2 nivåer så att vi bara behöver göra 8 mätningar.
- Vi får alltså en 2^{5-2}-plan.

Börja som om det var en vanlig 2^3-plan.

Grupp nr	A	B	C	AB	AC	BC	ABC
1	-	-	-	+	+	+	-
2	+	-	-	-	-	+	+
3	-	+	-	-	+	-	+
4	+	+	-	+	-	-	-
5	-	-	+	+	-	-	+
6	+	-	+	-	+	-	-
7	-	+	+	-	-	+	-
8	+	+	+	+	+	+	+

- Vi väljer nu våra två generatorer, en för varje nivå som vi reducerar.

$$AB = D$$

$$AC = E$$

Vår plan blir nu:

Grupp nr	A	B	C	D	E	BC	ABC
1	-	-	-	+	+	+	-
2	+	-	-	-	-	+	+
3	-	+	-	-	+	-	+
4	+	+	-	+	-	-	-
5	-	-	+	+	-	-	+
6	+	-	+	-	+	-	-
7	-	+	+	-	-	+	-
8	+	+	+	+	+	+	+

- Vad blir de definierande relationerna?

Vår plan blir nu:

Grupp nr	A	B	C	D	E	BC	ABC
1	-	-	-	+	+	+	-
2	+	-	-	-	-	+	+
3	-	+	-	-	+	-	+
4	+	+	-	+	-	-	-
5	-	-	+	+	-	-	+
6	+	-	+	-	+	-	-
7	-	+	+	-	-	+	-
8	+	+	+	+	+	+	+

- Vi får de definierande relationerna

$$I_1 = DAB = ABD$$

$$I_2 = EAC = ACE$$

$$I_3 = ABD \cdot ACE = BCED$$

$I_1 = ABD$, $I_2 = ACE$, $I_3 = BCDE$

- Vilka alias har vi?

$I_1 = ABD$, $I_2 = ACE$, $I_3 = BCDE$

- Vi får alias:

Faktor	I_1	I_2	I_3
M:	ABD	ACE	$BCDE$
A:	BD	CE	$ABCDE$
B:	AD	$ABCE$	CDE
C:	$ABCD$	AE	BDE
D:	AB	$ACDE$	BCE
E:	$ABDE$	AC	BCD
BC:	ACD	ABE	DE
ABC:	CD	BE	ADE

Där M symboliserar medelvärdet, alltså en kolumn med bara $+1$, såsom när vi räknar ut l_0. En sådan kolumn är alltså identisk med identiteten, I.

- Kom ihåg att kolumnerna för alla huvudeffekter bör vara utskrivna i tabellen för den reducerade försöksplanen. Detta är nödvändigt för att man när man väl utför experimenten skall se hur man skall ställa in faktornivåerna i respektive mätgrupp.
- Vi ser alltså att alla effekter vi försöker skatta kommer att blandas ihop med tre andra effekter. Även medelvärdet!
- Ihopblandningen kan leda till att en effekt ser stor ut då alla fyra deleffekterna kanske är relativt små i sin egen rätt men summerar upp till något större.
- Ihopblandningen kan också leda till det motsatta (vilken kan vara även farligare). Vi kan tro att effekter inte existerar pga att de fyra effekterna tar ut varandra även om de var och en kanske är stora.

Upplösning

- Oftast är effekter av högre ordningen (trefaktorsamspel eller högre) inte så starka. Det är inte ovanligt att man antar att dessa inte existerar eller är så små att man inte bryr sig om dem.
- Man skulle därför vilja att huvudeffekter, och helst även tvåfaktorsamspel, inte hade alias som också var huvudeffekter eller tvåfaktorsamspel.
- Om vi tittar på våra alias ser vi att ingen huvudeffekt har en annan huvudeffekt som alias. Dock har de alla ett tvåfaktorsamspel. Vi ser också att medelvärdet inte har någon huvudeffekt eller tvåfaktorsamspel som alias.

- Ordningen av samspelen för aliasen är relaterade till orden i de definierande relationerna.
- Har vi ord som alla är större än 3 bokstäver så kan omöjligt en huvudeffekt få ett alias av mindre ordning än 3-faktorsamspel (eftersom en bokstav max kan dra bort en bokstav från orden).
- Vi ser alltså att det finns en fördel med att ha långa ord i våra definierande relationer!
- En kedja är ju aldrig starkare än sin svagaste länk och därför får man titta på det ord som är kortast för att se hur bra alias vi kan få. Antal bokstäver i det kortaste ordet kallas den reducerade provtagningsplanens **upplösning**.
- I vårt fall så var upplösningen 3 då både I_1 och I_2 hade ord som var 3 bokstäver långa.

- Vi har alltså en 2_{III}^{5-2}-plan. (Man skriver upplösningen med romerska bokstäver som index till 2:an)
- Hade vi kunnat få en bättre plan?

- Vi har alltså en 2^{5-2}_{III}-plan. (Man skriver upplösningen med romerska bokstäver som index till 2:an)
- Hade vi kunnat få en bättre plan?
- Nej, hade vi låtit de två första orden vara av längd 4 så hade produkten av dem haft 3 bokstäver gemensamt. Således hade 3:e ordet blivit av längd 2 vilket hade gett en sämre upplösning.
- Upplösning *III* är alltså det bästa vi kan åstadkomma för en 2^{5-2}-plan.

Sammanfattning

- En fullständig försöksplan kan kräva många mätningar och därför bli dyr.
- En reducerad försöksplan kan göras istället men har vissa nackdelar.
  I en reducerad försöksplan får skattade effekter alias med vilka de delar skattningen. Detta kan vara problematiskt om man har alias som har en stor effekt och "blandar sig i" den uppskattade effekten.
- Högre upplösning gör att mindre jobbiga effekter blandar sig med varandra.
- Man kan använda reducerade försöksplaner för en preliminär analys som sedan skall ge en svar på om det är värt att fortsätta med en mer grundlig fullständig analys.