

Försöksplanering – faktorförsök¹

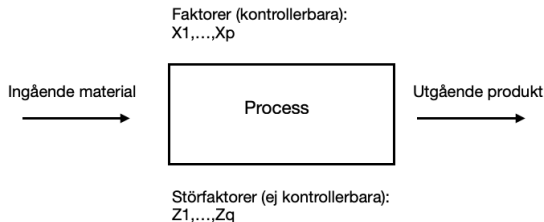
LMA521 & LKT325

Chalmers Tekniska Högskola

2022

¹Denna presentation innehåller bilder tagna från Ulla Blomqvist, Försöksplanering – Faktorförsök (2017) samt delar från Olof Zetterqvists föreläsningssanteckningar.

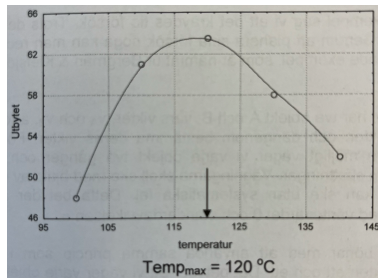
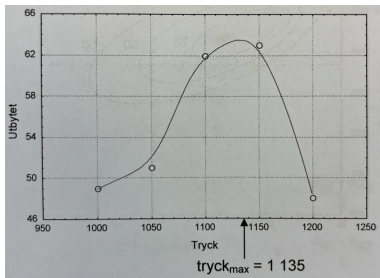
- Kan betraktas som ett sätt att generera fakta om en viss process/ett visst kontrollerat experiment.
- Man vill ta reda på om/hur ett utfall påverkas av att man förändrar invariabler, så kallade faktorer.
- Kan leda till kostnadsbesparingar, miljöförbättring etc.
- Motivering: produktutveckling (kvalitet), processutveckling (t.ex. besparingar), ökad förståelse.
- Vad får ändring av alla/vissa av faktorerna x_1, \dots, x_p för effekt på utfallet y ? Finns det kausala samband?



- Det finns olika metoder för att utföra försöksplanering, t.ex. bestämmande av stickprovsstorlek (för att säkra en viss bredd på ett konfidensintervall) och faktorförsök.
- En faktor är en kategorisk/kvalitativ variabel som kan anta ett fixt antal möjliga värden.
- Vi kommer bara att titta på faktorer som har två nivåer, vilket vi betecknar med (+) och (−) för "hög" respektive "låg" nivå. Notera att (+) inte nödvändigtvis representerar höga mätvärden – vi kan studera både kvalitativa och kvantitativa variabler.
- En enkel idé: Ändra en faktor i taget.

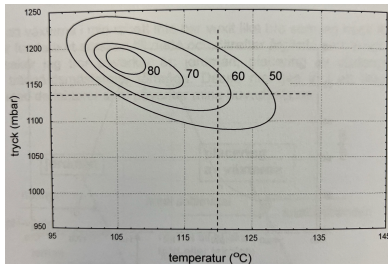
Exempel ("En i taget")

- Man tror att faktorerna tryck och temperatur påverkar utfallet (vikt i gram) som skall maximeras.
- Man fixerar först temperaturen till $120\text{ }^{\circ}\text{C}$ och mäter utfallet för 5 olika tryck. Tryck som maximerar utfallet: 1135 mbar.
- Kan vi få ett bättre resultat om vi varierar temperaturen med 1135 mbar fixerat? Temperatur som maximerar utfallet: $120\text{ }^{\circ}\text{C}$.
- Slutsats: Optimalt utfall med $120\text{ }^{\circ}\text{C}$ och 1135 mbar.



Exempel ("En i taget")

- Men...
- vi har erhållit ett lokalt maximum.
- Om man ritar upp konturen för den faktiska funktionen ser man att det globala maximumet ligger en bit ifrån.
- Slutsats: det finns samspelseffekter.
- Ett bättre sätt är faktorförsök vilket innebär att vi varierar flera faktorer samtidigt.



Fullständiga faktorförsök

Faktorförsök vs "En i taget":

- Man kan generera samma information med färre mätningar (resursoptimering).
- Man erhåller bättre skattningar av effekten av en viss faktor (besparing).
- Man får skattningar av samspelseffekterna.

Fullständiga faktorförsök

Faktorförsök vs "En i taget":

- Man kan generera samma information med färre mätningar (resursoptimering).
- Man erhåller bättre skattningar av effekten av en viss faktor (besparing).
- Man får skattningar av samspelseffekterna.

Vi börjar med att titta på fullständiga tvåfaktorsförsök (betecknas 2^k): var och en av de $k \geq 2$ faktorer som undersöks har två nivåer ($(+)$ och $(-)$).

- Det går även att studera 3^k -försök, 4^k -försök o.s.v. (de är uppbyggda på samma sätt men är svårare att illustrera grafiskt) men 2^k -försök är ett billigt sätt att komma ganska långt.

Fullständiga faktorförsök

Faktorförsök vs "En i taget":

- Man kan generera samma information med färre mätningar (resursoptimering).
- Man erhåller bättre skattningar av effekten av en viss faktor (besparing).
- Man får skattningar av samspelseffekterna.

Vi börjar med att titta på fullständiga tvåfaktorsförsök (betecknas 2^k): var och en av de $k \geq 2$ faktorer som undersöks har två nivåer ((+) och (-)).

- Det går även att studera 3^k -försök, 4^k -försök o.s.v. (de är uppbyggda på samma sätt men är svårare att illustrera grafiskt) men 2^k -försök är ett billigt sätt att komma ganska långt.
- Det totala antalet försök är $N = 2^k$, $k \geq 2$, t.ex. om vi har tre faktorer så kommer vi att behöva göra $2^3 = 8$ mätningar.

Fullständiga faktorförsök

Faktorförsök vs "En i taget":

- Man kan generera samma information med färre mätningar (resursoptimering).
- Man erhåller bättre skattningar av effekten av en viss faktor (besparing).
- Man får skattningar av samspelseffekterna.

Vi börjar med att titta på fullständiga tvåfaktorsförsök (betecknas 2^k): var och en av de $k \geq 2$ faktorer som undersöks har två nivåer ((+) och (-)).

- Det går även att studera 3^k -försök, 4^k -försök o.s.v. (de är uppbyggda på samma sätt men är svårare att illustrera grafiskt) men 2^k -försök är ett billigt sätt att komma ganska långt.
- Det totala antalet försök är $N = 2^k$, $k \geq 2$, t.ex. om vi har tre faktorer så kommer vi att behöva göra $2^3 = 8$ mätningar.
- Man brukar beteckna de olika faktorerna med stora bokstäver: A, B, C osv. Deras samspel/interaktioner betecknas med AB, AC, BC, ABC osv.

Fullständiga faktorförsök

Faktorförsök vs "En i taget":

- Man kan generera samma information med färre mätningar (resursoptimering).
- Man erhåller bättre skattningar av effekten av en viss faktor (besparing).
- Man får skattningar av samspelseffekterna.

Vi börjar med att titta på fullständiga tvåfaktorsförsök (betecknas 2^k): var och en av de $k \geq 2$ faktorer som undersöks har två nivåer ((+) och (-)).

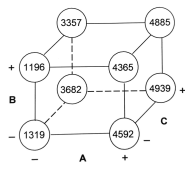
- Det går även att studera 3^k -försök, 4^k -försök o.s.v. (de är uppbyggda på samma sätt men är svårare att illustrera grafiskt) men 2^k -försök är ett billigt sätt att komma ganska långt.
- Det totala antalet försök är $N = 2^k$, $k \geq 2$, t.ex. om vi har tre faktorer så kommer vi att behöva göra $2^3 = 8$ mätningar.
- Man brukar beteckna de olika faktorerna med stora bokstäver: A, B, C osv. Deras samspel/interaktioner betecknas med AB, AC, BC, ABC osv.
- T.ex. A^+ och A^- betecknar faktor A på hög respektive låg nivå.

Designmatriser

- De olika försökssituationerna beskrivs mha en s.k. design-/försöksmatris och en graf.
- Man håller den s.k. standardordningen för hur matrisen struktureras (kolumn 1: +, -, +, -, +, ... osv.)

Exempel: y =poäng för visuell uppfattningsförmåga, A =exponeringstid ((-): 60, (+): 120 msec), B =bakgrundsfärg ((-): vit, (+): mönstrad), C ((-): 6, (+): 12 punkter).

Nr	A	B	C	AB	AC	BC	ABC	y
1	-	-	-	+	+	+	-	1 319
2	+	-	-	-	-	+	+	4 592
3	-	+	-	-	+	-	+	1 196
4	+	+	-	+	-	-	-	4 365
5	-	-	+	+	-	-	+	3 682
6	+	-	+	-	+	-	-	4 939
7	-	+	+	-	-	+	-	3 357
8	+	+	+	+	+	+	+	4 885



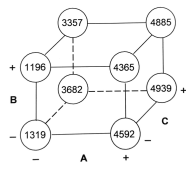
Sedan gör man mätningar y_1, \dots, y_N , $N = 2^k$, enligt designmatrisen:

Designmatriser

- De olika försökssituationerna beskrivs mha en s.k. design-/försöksmatris och en graf.
- Man håller den s.k. standardordningen för hur matrisen struktureras (kolumn 1: +, -, +, -, +, ... osv.)

Exempel: y =poäng för visuell uppfattningsförmåga, A =exponeringstid ((-): 60, (+): 120 msec), B =bakgrundsfärg ((-): vit, (+): mönstrad), C ((-): 6, (+): 12 punkter).

Nr	A	B	C	AB	AC	BC	ABC	y
1	-	-	-	+	+	+	-	1 319
2	+	-	-	-	-	+	+	4 592
3	-	+	-	-	+	-	+	1 196
4	+	+	-	+	-	-	-	4 365
5	-	-	+	+	-	-	+	3 682
6	+	-	+	-	+	-	-	4 939
7	-	+	+	-	-	+	-	3 357
8	+	+	+	+	+	+	+	4 885



Sedan gör man mätningar y_1, \dots, y_N , $N = 2^k$, enligt designmatrisen:

- Ordningen i vilken man utför mätningarna skall randomiseras: välj rader i designmatrisen slumpmässigt utan återläggning (se sid 17 i boken för förklaring).
- Samspelens nivåer bestäms av individuella faktorerens nivåer: I försök 1 får AB nivån $(-1) \cdot (-1) = (+1)$ och i försök 2 får ABC $[(+1) \cdot (-1)] \cdot (-1) = (-1) \cdot (-1) = (+1)$

Beräkning av effekter

Den allmänna formeln, i ett 2^k -faktorförsök, för effekten för en faktor F , som representeras av kolumn j i designmatrisen (t.ex. $j = 2$ när $F = B$), ges av

$$\ell_F = \ell(j) = \frac{1}{N/2} \sum_{i=1}^N \pm y_i$$

Vi finner $\pm y_i$ genom att multiplicera y -värdet på rad $i = 1, \dots, N = 2^k$ i designmatrisen med det tecken ((+) eller (-)) som svarar mot element (i, j) i designmatrisen.

Beräkning av effekter

Den allmänna formeln, i ett 2^k -faktorförsök, för effekten för en faktor F , som representeras av kolumn j i designmatrisen (t.ex. $j = 2$ när $F = B$), ges av

$$\ell_F = \ell(j) = \frac{1}{N/2} \sum_{i=1}^N \pm y_i$$

Vi finner $\pm y_i$ genom att multiplicera y -värdet på rad $i = 1, \dots, N = 2^k$ i designmatrisen med det tecken ((+) eller (-)) som svarar mot element (i, j) i designmatrisen. Vad vi gör här är att vi beräknar förändringarna vi får om vi går från F^- till F^+ , givet alla olika fixeringar av de andra faktorerna, och medelvärdesbildar över dessa förändringar.

Beräkning av effekter

Den allmänna formeln, i ett 2^k -faktorförsök, för effekten för en faktor F , som representeras av kolumn j i designmatrisen (t.ex. $j = 2$ när $F = B$), ges av

$$\ell_F = \ell(j) = \frac{1}{N/2} \sum_{i=1}^N \pm y_i$$

Vi finner $\pm y_i$ genom att multiplicera y -värdet på rad $i = 1, \dots, N = 2^k$ i designmatrisen med det tecken ((+) eller (-)) som svarar mot element (i, j) i designmatrisen. Vad vi gör här är att vi beräknar förändringarna vi får om vi går från F^- till F^+ , givet alla olika fixeringar av de andra faktorerna, och medelvärdesbildar över dessa förändringar.

I exemplet från sidan innan (där $N = 2^3 = 8$) får vi t.ex.:

- Huvudeffekt för A:

$$\begin{aligned} \ell_A &= \frac{-y_1 + y_2 - y_3 + y_4 - y_5 + y_6 - y_7 + y_8}{8/2} = \ell_{A^+} - \ell_{A^-} = \frac{y_2 + y_4 + y_6 + y_8}{8/2} - \frac{y_1 + y_3 + y_5 + y_7}{8/2} = \\ &= \frac{4592 + 4365 + 4939 + 4885}{8/2} - \frac{1319 + 1196 + 3682 + 3357}{8/2} = 4695.25 - 2388.5 = 2306.75 \end{aligned}$$

- Samspel AB:

$$\ell_{AB} = \ell_{A^+} - \ell_{A^-} = ((y_1 + y_4 + y_5 + y_8) - (y_2 + y_3 + y_6 + y_7)) / (8/2) = 41.75$$

Räknar vi ut alla effekter får vi:

Nr	A	B	C	AB	AC	BC	ABC	y
1	-	-	-	+	+	+	-	1 319
2	+	-	-	-	-	+	+	4 592
3	-	+	-	-	+	-	+	1 196
4	+	+	-	+	-	-	-	4 365
5	-	-	+	+	-	-	+	3 682
6	+	-	+	-	+	-	-	4 939
7	-	+	+	-	-	+	-	3 357
8	+	+	+	+	+	+	+	4 885

$$\ell_A = \frac{1}{4}(4592 + 4365 + 4939 + 4885) - \frac{1}{4}(1319 + 1196 + 3682 + 3357) = 2\,306.75$$

$$\ell_B = \frac{1}{4}(1196 + 4365 + 3357 + 4885) - \frac{1}{4}(1319 + 4592 + 3682 + 4939) = -182.25$$

$$\ell_C = \frac{1}{4}(3682 + 4939 + 3357 + 4885) - \frac{1}{4}(1319 + 4592 + 1196 + 4365) = 1\,347.75$$

$$\ell_{AB} = \frac{1}{4}(1319 + 4365 + 3682 + 4885) - \frac{1}{4}(4592 + 1196 + 4939 + 3357) = 41.75$$

$$\ell_{AC} = \frac{1}{4}(1319 + 1196 + 4939 + 4885) - \frac{1}{4}(4592 + 4365 + 3682 + 3357) = -914.25$$

$$\ell_{BC} = \frac{1}{4}(1319 + 4592 + 3357 + 4885) - \frac{1}{4}(1196 + 4365 + 3682 + 4939) = -7.25$$

$$\ell_{ABC} = \frac{1}{4}(4592 + 1196 + 3682 + 4885) - \frac{1}{4}(1319 + 4365 + 4939 + 3357) = 93.75$$

$\ell_F = \ell(j) = \frac{1}{N/2} \sum_{i=1}^N \pm y_i$, $j = 1, \dots, n = \sum_{i=1}^k \binom{k}{i}$, där $\binom{k}{i}$, $i \geq 2$, är hur många samspelseffekter av storlek i vi får ($i = 2$ svara mot AB, AC, BC i exemplet).

Diagram för huvudeffekter (t.ex. en linje från medelvärdet av alla y_i där A har låg nivå, ℓ_{A-} , till medelvärdet för alla y_i där A har hög nivå, ℓ_{A+}):

$$\ell_{A-} = \frac{1}{4}(y_1 + y_3 + y_5 + y_7) = \frac{1}{4}(1319 + 1196 + 3682 + 3357) = 2\,388.5$$

$$\ell_{A+} = \frac{1}{4}(y_2 + y_4 + y_6 + y_8) = \frac{1}{4}(4592 + 4365 + 4939 + 4885) = 4\,695.25$$

$$\ell_{B-} = \frac{1}{4}(y_1 + y_2 + y_5 + y_6) = \frac{1}{4}(1319 + 4592 + 3682 + 4939) = 3\,633$$

$$\ell_{B+} = \frac{1}{4}(y_3 + y_4 + y_7 + y_8) = \frac{1}{4}(1196 + 4365 + 3357 + 4885) = 3\,450.75$$

$$\ell_{C-} = \frac{1}{4}(y_1 + y_2 + y_3 + y_4) = \frac{1}{4}(1319 + 4592 + 1196 + 4365) = 2\,868$$

$$\ell_{C+} = \frac{1}{4}(y_5 + y_6 + y_7 + y_8) = \frac{1}{4}(3682 + 4939 + 3357 + 4885) = 4\,215.75$$

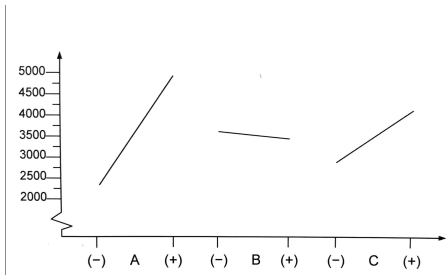


Diagram för huvudeffekter (t.ex. en linje från medelvärdet av alla y_i där A har låg nivå, ℓ_{A-} , till medelvärdet för alla y_i där A har hög nivå, ℓ_{A+}):

$$\ell_{A-} = \frac{1}{4}(y_1 + y_3 + y_5 + y_7) = \frac{1}{4}(1319 + 1196 + 3682 + 3357) = 2\,388.5$$

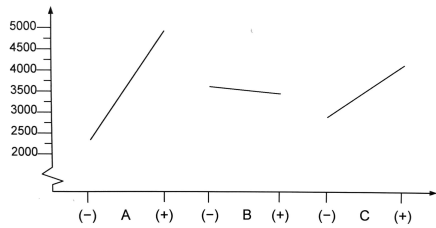
$$\ell_{A+} = \frac{1}{4}(y_2 + y_4 + y_6 + y_8) = \frac{1}{4}(4592 + 4365 + 4939 + 4885) = 4\,695.25$$

$$\ell_{B-} = \frac{1}{4}(y_1 + y_2 + y_5 + y_6) = \frac{1}{4}(1319 + 4592 + 3682 + 4939) = 3\,633$$

$$\ell_{B+} = \frac{1}{4}(y_3 + y_4 + y_7 + y_8) = \frac{1}{4}(1196 + 4365 + 3357 + 4885) = 3\,450.75$$

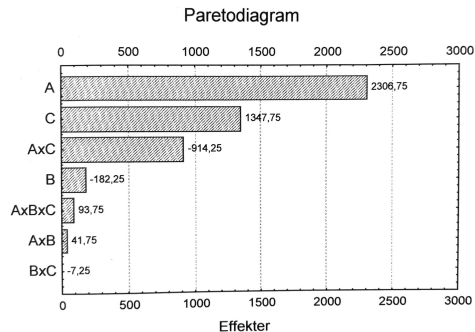
$$\ell_{C-} = \frac{1}{4}(y_1 + y_2 + y_3 + y_4) = \frac{1}{4}(1319 + 4592 + 1196 + 4365) = 2\,868$$

$$\ell_{C+} = \frac{1}{4}(y_5 + y_6 + y_7 + y_8) = \frac{1}{4}(3682 + 4939 + 3357 + 4885) = 4\,215.75$$



- Vi ser att, bland huvudeffekterna, är effekten av en nivåhöjning av faktor A störst, därefter den för faktor C.
- En nivåhöjning av faktor B ger minst förändring, sett till storleken, och ger dessutom upphov till en minskad effekt.
- Notera att vi bara ser linjära förändringar här (vi studerar bara linjära förändringar men icke-linjära samband kan finnas).

Låt oss titta på storleksordningen på alla effekter (inklusive samspelens effekter). Dessa illustreras typiskt med ett s.k. Paretdiagram², där man sorterar effekterna efter deras storleksordning:



$$\ell_A = \frac{1}{4}(4592 + 4365 + 4939 + 4885) - \frac{1}{4}(1319 + 1196 + 3682 + 3357) = 2\,306.75$$

$$\ell_B = \frac{1}{4}(1196 + 4365 + 3357 + 4885) - \frac{1}{4}(1319 + 4592 + 3682 + 4939) = -182.25$$

$$\ell_C = \frac{1}{4}(3682 + 4939 + 3357 + 4885) - \frac{1}{4}(1319 + 4592 + 1196 + 4365) = 1\,347.75$$

$$\ell_{AB} = \frac{1}{4}(1319 + 4365 + 3682 + 4885) - \frac{1}{4}(4592 + 1196 + 4939 + 3357) = 41.75$$

$$\ell_{AC} = \frac{1}{4}(1319 + 1196 + 4939 + 4885) - \frac{1}{4}(4592 + 4365 + 3682 + 3357) = -914.25$$

$$\ell_{BC} = \frac{1}{4}(1319 + 4592 + 3357 + 4885) - \frac{1}{4}(1196 + 4365 + 3682 + 4939) = -7.25$$

$$\ell_{ABC} = \frac{1}{4}(4592 + 1196 + 3682 + 4885) - \frac{1}{4}(1319 + 4365 + 4939 + 3357) = 93.75$$

Faktor A har störst effekt, därefter kommer faktor C och samspelet AC. Notera att ℓ_{AB} förklarar hur A och B tillsammans förstärker/motverkar den effekt som redan givits av ℓ_A och ℓ_B .

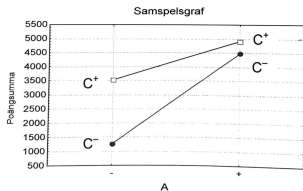
²Ett (liggande) histogram där man har sorterat staplarna/variablerna i storleksordning.

Vi tar nu fram en samspelsgraf för samspelet AB (man ritar in den faktor som har störst huvudeffekt på x-axeln).

- Allmänt: samspel saknas om de två linjerna är helt parallella och om de korsas så har vi tydlig indikation på samspel (ju större vinkel mellan dem desto större samspel)³.
- På låg nivå för A finns det en skillnad i utfallet y beroende på nivån på C ((+) ger högre y än (-)). När vi ändrar A till hög nivå minskar skillnaden, m.a.o. att höja A minskar effekten som en C-höjning har på y . A^+ och C^+ ger bäst effekt här.

$$A^-C^- = \frac{y_1+y_3}{2} = \frac{1319+1196}{2} = 1257.5 \quad A^-C^+ = \frac{y_5+y_7}{2} = \frac{3682+3357}{2} = 3519.5$$

$$A^+C^- = \frac{y_2+y_4}{2} = \frac{4592+4365}{2} = 4478.5 \quad A^+C^+ = \frac{y_6+y_8}{2} = \frac{4885+4935}{2} = 4910$$



Nr	A	B	C	AB	AC	BC	ABC	y
1	-	-	-	+	+	+	-	1 319
2	+	-	-	-	-	+	+	4 592
3	-	+	-	-	+	-	+	1 196
4	+	+	-	+	-	-	-	4 365
5	-	-	+	+	-	-	+	3 682
6	+	-	+	-	+	-	-	4 939
7	-	+	+	-	-	+	-	3 357
8	+	+	+	+	+	+	+	4 885

³För ytterligare tolkningar: <https://data.library.virginia.edu/understanding-2-way-interactions/>

Generell stokastisk modellstruktur

- Kom ihåg störfaktorerna som nämndes i början. De påverkar våra observerade y -värden och utan dem hade statistiska metoder inte behövts eftersom allt hade varit deterministiskt/perfekt förutsägbart.

Generell stokastisk modellstruktur

- Kom ihåg störfaktorerna som nämndes i början. De påverkar våra observerade y -värden och utan dem hade statistiska metoder inte behövts eftersom allt hade varit deterministiskt/perfekt förutsägbart.
- Vi modellerar dem m.h.a. en stokastisk variabel ε som vi antar har väntevärde 0 och varians σ^2 .

Generell stokastisk modellstruktur

- Kom ihåg störfaktorerna som nämndes i början. De påverkar våra observerade y -värden och utan dem hade statistiska metoder inte behövts eftersom allt hade varit deterministiskt/perfekt förutsägbart.
- Vi modellerar dem m.h.a. en stokastisk variabel ε som vi antar har väntevärde 0 och varians σ^2 .
- Som tidigare betonats så antar vi att utfallet beror linjärt på faktorerna.

Generell stokastisk modellstruktur

- Kom ihåg störfaktorerna som nämndes i början. De påverkar våra observerade y -värden och utan dem hade statistiska metoder inte behövts eftersom allt hade varit deterministiskt/perfekt förutsägbart.
- Vi modellerar dem m.h.a. en stokastisk variabel ε som vi antar har väntevärde 0 och varians σ^2 .
- Som tidigare betonats så antar vi att utfallet beror linjärt på faktorerna.
- Läger man ihop dessa observationer får man att ett utfall y i ett 2^2 -försök modelleras som en stokastisk variabel

$$Y = Y|x = \mu + \beta_A x_A + \beta_B x_B + \beta_{AB} x_A x_B + \varepsilon$$

när vi har två faktorer som vi betecknar $x = (x_A, x_B) \in \{-1, +1\} \times \{-1, +1\}$ ($x_A x_B$ betecknar samspelet mellan x_A och x_B). Detta kallas för en linjär modell/regression.

Generell stokastisk modellstruktur

- Kom ihåg störfaktorerna som nämndes i början. De påverkar våra observerade y -värden och utan dem hade statistiska metoder inte behövts eftersom allt hade varit deterministiskt/perfekt förutsägbart.
- Vi modellerar dem m.h.a. en stokastisk variabel ε som vi antar har väntevärde 0 och varians σ^2 .
- Som tidigare betonats så antar vi att utfallet beror linjärt på faktorerna.
- Läger man ihop dessa observationer får man att ett utfall y i ett 2^2 -försök modelleras som en stokastisk variabel

$$Y = Y|x = \mu + \beta_A x_A + \beta_B x_B + \beta_{AB} x_A x_B + \varepsilon$$

när vi har två faktorer som vi betecknar $x = (x_A, x_B) \in \{-1, +1\} \times \{-1, +1\}$ ($x_A x_B$ betecknar samspelet mellan x_A och x_B). Detta kallas för en linjär modell/regression.

- Varje observerat värde y_i med tillhörande faktornivåer $x = (x_A, x_B)$ antas alltså vara en realisering av $Y_i = Y|x$ ($x \Leftrightarrow$ i-te raden i designmatrisen); alla Y_i är oberoende.

Generell stokastisk modellstruktur

- Kom ihåg störfaktorerna som nämndes i början. De påverkar våra observerade y -värden och utan dem hade statistiska metoder inte behövts eftersom allt hade varit deterministiskt/perfekt förutsägbart.
- Vi modellerar dem m.h.a. en stokastisk variabel ε som vi antar har väntevärde 0 och varians σ^2 .
- Som tidigare betonats så antar vi att utfallet beror linjärt på faktorerna.
- Läger man ihop dessa observationer får man att ett utfall y i ett 2^2 -försök modelleras som en stokastisk variabel

$$Y = Y|x = \mu + \beta_A x_A + \beta_B x_B + \beta_{AB} x_A x_B + \varepsilon$$

när vi har två faktorer som vi betecknar $x = (x_A, x_B) \in \{-1, +1\} \times \{-1, +1\}$ ($x_A x_B$ betecknar samspelet mellan x_A och x_B). Detta kallas för en linjär modell/regression.

- Varje observerat värde y_i med tillhörande faktornivåer $x = (x_A, x_B)$ antas alltså vara en realisering av $Y_i = Y|x$ ($x \Leftrightarrow$ ite raden i designmatrisen); alla Y_i är oberoende.
- I ett 2^3 -försök, lägg till $\beta_C x_C + \beta_{AC} x_A x_C + \beta_{BC} x_B x_C + \beta_{ABC} x_A x_B x_C$ osv.

Vi får (mha kunskaper från första delen av kursen) att

$$Y = Y|x = \mu + \beta_A x_A + \beta_B x_B + \beta_{AB} x_A x_B + \varepsilon$$

har varians σ^2 och väntevärde $\mu + \beta_A x_A + \beta_B x_B + \beta_{AB} x_A x_B$.

Vi får (mha kunskaper från första delen av kursen) att

$$Y = Y|x = \mu + \beta_A x_A + \beta_B x_B + \beta_{AB} x_A x_B + \varepsilon$$

har varians σ^2 och väntevärde $\mu + \beta_A x_A + \beta_B x_B + \beta_{AB} x_A x_B$.

- μ motsvarar det genomsnittliga förväntade utfallet, sett till faktorernas nivåer:

$$\mu = \frac{\mathbb{E}[Y|(-1, -1)] + \mathbb{E}[Y|(+1, -1)] + \mathbb{E}[Y|(-1, +1)] + \mathbb{E}[Y|(+1, +1)]}{N}; \quad N = 4.$$

Vi får (mha kunskaper från första delen av kursen) att

$$Y = Y|x = \mu + \beta_A x_A + \beta_B x_B + \beta_{AB} x_A x_B + \varepsilon$$

har varians σ^2 och väntevärde $\mu + \beta_A x_A + \beta_B x_B + \beta_{AB} x_A x_B$.

- μ motsvarar det genomsnittliga förväntade utfallet, sett till faktorernas nivåer:

$$\mu = \frac{\mathbb{E}[Y|(-1, -1)] + \mathbb{E}[Y|(+1, -1)] + \mathbb{E}[Y|(-1, +1)] + \mathbb{E}[Y|(+1, +1)]}{N}; \quad N = 4.$$

- T.ex. koefficienten/parametern β_A motsvarar hur Y förväntas förändras då faktor A byter nivå om man medelvärdesbildar mellan båda nivåerna på övriga faktorer.

Vi får (mha kunskaper från första delen av kursen) att

$$Y = Y|x = \mu + \beta_A x_A + \beta_B x_B + \beta_{AB} x_A x_B + \varepsilon$$

har varians σ^2 och väntevärde $\mu + \beta_A x_A + \beta_B x_B + \beta_{AB} x_A x_B$.

- μ motsvarar det genomsnittliga förväntade utfallet, sett till faktorernas nivåer:

$$\mu = \frac{\mathbb{E}[Y|(-1, -1)] + \mathbb{E}[Y|(+1, -1)] + \mathbb{E}[Y|(-1, +1)] + \mathbb{E}[Y|(+1, +1)]}{N}; \quad N = 4.$$

- T.ex. koefficienten/parametern β_A motsvarar hur Y förväntas förändras då faktor A byter nivå om man medelvärdesbildar mellan båda nivåerna på övriga faktorer.

M.a.o. får vi:

$$\beta_A = \frac{\mathbb{E}[Y|(+1, -1)] + \mathbb{E}[Y|(+1, +1)]}{N/2} - \frac{\mathbb{E}[Y|(-1, -1)] + \mathbb{E}[Y|(-1, +1)]}{N/2}$$

$$\beta_B = \frac{\mathbb{E}[Y|(-1, +1)] + \mathbb{E}[Y|(+1, +1)]}{N/2} - \frac{\mathbb{E}[Y|(-1, -1)] + \mathbb{E}[Y|(+1, -1)]}{N/2}$$

$$\beta_{AB} = \frac{\mathbb{E}[Y|(+1, +1)] + \mathbb{E}[Y|(-1, -1)]}{N/2} - \frac{\mathbb{E}[Y|(-1, +1)] + \mathbb{E}[Y|(+1, -1)]}{N/2}$$

Skattning av koefficienterna

- Det kan ske att man gör flera utfallsmätningar för en och samma nivå, t.ex. $y_{i,(+1,+1)}$, $i = 1, \dots, n_{(+1,+1)}$, för $x = (x_A, x_B) = (+1, +1)$ (ordningen på alla mätningar skall randomiseras). Analog notation för övriga nivåkombinationer.

Skattning av koefficienterna

- Det kan ske att man gör flera utfallsmätningar för en och samma nivå, t.ex. $y_{i,(+1,+1)}$, $i = 1, \dots, n_{(+1,+1)}$, för $x = (x_A, x_B) = (+1, +1)$ (ordningen på alla mätningar skall randomiseras). Analog notation för övriga nivåkombinationer.
- Tänker vi tillbaka på att vi skattar väntevärden med medelvärden får vi t.ex. skattningen

$$\ell_A = \ell_{A^+} - \ell_{A^-} = \frac{\bar{y}_{(+1,-1)} - \bar{y}_{(+1,+1)}}{4/2} - \frac{\bar{y}_{(-1,-1)} - \bar{y}_{(-1,+1)}}{4/2}$$

av β_A , där t.ex. $\bar{y}_{(+1,-1)}$ är medelvärdet av alla $y_{i,(+1,-1)}$, $i = 1, \dots, n_{(+1,-1)}$. Vi skattar μ med $\ell_M = (\bar{y}_{(+1,-1)} + \bar{y}_{(+1,+1)} + \bar{y}_{(-1,-1)} + \bar{y}_{(-1,+1)})/4$.

Skattning av koefficienterna

- Det kan ske att man gör flera utfallsmätningar för en och samma nivå, t.ex. $y_{i,(+1,+1)}$, $i = 1, \dots, n_{(+1,+1)}$, för $x = (x_A, x_B) = (+1, +1)$ (ordningen på alla mätningar skall randomiseras). Analog notation för övriga nivåkombinationer.
- Tänker vi tillbaka på att vi skattar väntevärden med medelvärden får vi t.ex. skattningen

$$\ell_A = \ell_{A^+} - \ell_{A^-} = \frac{\bar{y}_{(+1,-1)} - \bar{y}_{(+1,+1)}}{4/2} - \frac{\bar{y}_{(-1,-1)} - \bar{y}_{(-1,+1)}}{4/2}$$

av β_A , där t.ex. $\bar{y}_{(+1,-1)}$ är medelvärdet av alla $y_{i,(+1,-1)}$, $i = 1, \dots, n_{(+1,-1)}$. Vi skattar μ med $\ell_M = (\bar{y}_{(+1,-1)} + \bar{y}_{(+1,+1)} + \bar{y}_{(-1,-1)} + \bar{y}_{(-1,+1)})/4$.

- Notera att vi i exemplet i början hade $n_{(-1,-1,-1)} = \dots = n_{(+1,+1,+1)} = 1$, dvs vi hade bara en mätning per nivåkombination, så t.ex.

$\bar{y}_{(-1,-1,-1)} = y|(-1, -1, -1) = y_1$. M.a.o. följer skattningen i detta fall den tidigare givna formen $\ell_F = \ell(j) = \frac{1}{N/2} \sum_{i=1}^N \pm y_i$.

Skattning av koefficienterna

- Det kan ske att man gör flera utfallsmätningar för en och samma nivå, t.ex. $y_{i,(+1,+1)}$, $i = 1, \dots, n_{(+1,+1)}$, för $x = (x_A, x_B) = (+1, +1)$ (ordningen på alla mätningar skall randomiseras). Analog notation för övriga nivåkombinationer.
- Tänker vi tillbaka på att vi skattar väntevärden med medelvärden får vi t.ex. skattningen

$$\ell_A = \ell_{A^+} - \ell_{A^-} = \frac{\bar{y}_{(+1,-1)} - \bar{y}_{(+1,+1)}}{4/2} - \frac{\bar{y}_{(-1,-1)} - \bar{y}_{(-1,+1)}}{4/2}$$

av β_A , där t.ex. $\bar{y}_{(+1,-1)}$ är medelvärdet av alla $y_{i,(+1,-1)}$, $i = 1, \dots, n_{(+1,-1)}$. Vi skattar μ med $\ell_M = (\bar{y}_{(+1,-1)} + \bar{y}_{(+1,+1)} + \bar{y}_{(-1,-1)} + \bar{y}_{(-1,+1)})/4$.

- Notera att vi i exemplet i början hade $n_{(-1,-1,-1)} = \dots = n_{(+1,+1,+1)} = 1$, dvs vi hade bara en mätning per nivåkombination, så t.ex.

$\bar{y}_{(-1,-1,-1)} = y|(-1, -1, -1) = y_1$. M.a.o. följer skattningen i detta fall den tidigare givna formen $\ell_F = \ell(j) = \frac{1}{N/2} \sum_{i=1}^N \pm y_i$.

- M.a.o. kan vi först räkna ut alla medelvärden och sedan stoppa in dem i designmatrisen (i y_i -värdenas kolumn).

- Låt oss titta närmare på egenskaper hos skattarna för β -koefficienterna, dvs de skattare som ger skattningarna ℓ_A , ℓ_B osv.

- Låt oss titta närmare på egenskaper hos skattarna för β -koefficienterna, dvs de skattare som ger skattningarna ℓ_A , ℓ_B osv.
- I ett 2^k -försök med en enda mätning per nivå-kombination får vi

$$\sigma_{\ell_F}^2 = \text{Var}(\ell_F) = \text{Var}\left(\frac{1}{N/2} \sum_{i=1}^N \pm Y_i\right) = \frac{1}{(N/2)^2} \sum_{i=1}^N (\pm 1)^2 \text{Var}(Y_i) = \frac{4\sigma^2}{N}$$
p.g.a. oberoendet; här är Y_i den stokastiska variabel som svarar mot den i te raden i designmatrisen. När vi istället gör n mätningar per kombination och ersätter Y_i med medelvärdet \bar{Y}_i av de mätningar som svara mot rad i så får skattaren istället variansen $\sigma_{\ell_F}^2 = 4\sigma^2/(nN)$.

- Låt oss titta närmare på egenskaper hos skattarna för β -koefficienterna, dvs de skattare som ger skattningarna ℓ_A , ℓ_B osv.
- I ett 2^k -försök med en enda mätning per nivå-kombination får vi

$$\sigma_{\ell_F}^2 = \text{Var}(\ell_F) = \text{Var}\left(\frac{1}{N/2} \sum_{i=1}^N \pm Y_i\right) = \frac{1}{(N/2)^2} \sum_{i=1}^N (\pm 1)^2 \text{Var}(Y_i) = \frac{4\sigma^2}{N}$$
p.g.a. oberoendet; här är Y_i den stokastiska variabel som svarar mot den i te raden i designmatrisen. När vi istället gör n mätningar per kombination och ersätter Y_i med medelvärdet \bar{Y}_i av de mätningar som svara mot rad i så får skattaren istället variansen $\sigma_{\ell_F}^2 = 4\sigma^2/(nN)$.
- Vi ser att denna varians är samma för alla faktorer och samspel.

- Låt oss titta närmare på egenskaper hos skattarna för β -koefficienterna, dvs de skattare som ger skattningarna ℓ_A , ℓ_B osv.
- I ett 2^k -försök med en enda mätning per nivå-kombination får vi
$$\sigma_{\ell_F}^2 = \text{Var}(\ell_F) = \text{Var}\left(\frac{1}{N/2} \sum_{i=1}^N \pm Y_i\right) = \frac{1}{(N/2)^2} \sum_{i=1}^N (\pm 1)^2 \text{Var}(Y_i) = \frac{4\sigma^2}{N}$$
 p.g.a. oberoendet; här är Y_i den stokastiska variabel som svarar mot den i te raden i designmatrisen. När vi istället gör n mätningar per kombination och ersätter Y_i med medelvärdet \bar{Y}_i av de mätningar som svara mot rad i så får skattaren istället variansen $\sigma_{\ell_F}^2 = 4\sigma^2/(nN)$.
- Vi ser att denna varians är samma för alla faktorer och samspel.
- Vi antar att varje skattare ℓ_F är en normalfördelad stokastisk variabel med varians $\sigma_{\ell_F}^2$ (notera att den antagna fördelningen styrs av den fördelning som vi antar för ε).

- Låt oss titta närmare på egenskaper hos skattarna för β -koefficienterna, dvs de skattare som ger skattningarna ℓ_A , ℓ_B osv.
- I ett 2^k -försök med en enda mätning per nivå-kombination får vi
$$\sigma_{\ell_F}^2 = \text{Var}(\ell_F) = \text{Var}\left(\frac{1}{N/2} \sum_{i=1}^N \pm Y_i\right) = \frac{1}{(N/2)^2} \sum_{i=1}^N (\pm 1)^2 \text{Var}(Y_i) = \frac{4\sigma^2}{N}$$
 p.g.a. oberoendet; här är Y_i den stokastiska variabel som svarar mot den i te raden i designmatrisen. När vi istället gör n mätningar per kombination och ersätter Y_i med medelvärdet \bar{Y}_i av de mätningar som svara mot rad i så får skattaren istället variansen $\sigma_{\ell_F}^2 = 4\sigma^2/(nN)$.
- Vi ser att denna varians är samma för alla faktorer och samspel.
- Vi antar att varje skattare ℓ_F är en normalfördelad stokastisk variabel med varians $\sigma_{\ell_F}^2$ (notera att den antagna fördelningen styrs av den fördelning som vi antar för ε).
- Vi kan skriva $\ell_F = \beta_F + \ell_X$, där β_F är den sanna effekten och $\ell_X \sim N(0, \sigma_{\ell_F}^2)$, vilken kallas för referensfördelningen.

- Låt oss titta närmare på egenskaper hos skattarna för β -koefficienterna, dvs de skattare som ger skattningarna ℓ_A , ℓ_B osv.
- I ett 2^k -försök med en enda mätning per nivå-kombination får vi
$$\sigma_{\ell_F}^2 = \text{Var}(\ell_F) = \text{Var}\left(\frac{1}{N/2} \sum_{i=1}^N \pm Y_i\right) = \frac{1}{(N/2)^2} \sum_{i=1}^N (\pm 1)^2 \text{Var}(Y_i) = \frac{4\sigma^2}{N}$$
 p.g.a. oberoendet; här är Y_i den stokastiska variabel som svarar mot den i te raden i designmatrisen. När vi istället gör n mätningar per kombination och ersätter Y_i med medelvärdet \bar{Y}_i av de mätningar som svara mot rad i så får skattaren istället variansen $\sigma_{\ell_F}^2 = 4\sigma^2/(nN)$.
- Vi ser att denna varians är samma för alla faktorer och samspel.
- Vi antar att varje skattare ℓ_F är en normalfördelad stokastisk variabel med varians $\sigma_{\ell_F}^2$ (notera att den antagna fördelningen styrs av den fördelning som vi antar för ε).
- Vi kan skriva $\ell_F = \beta_F + \ell_X$, där β_F är den sanna effekten och $\ell_X \sim N(0, \sigma_{\ell_F}^2)$, vilken kallas för referensfördelningen.
- Notera att referensfördelningen svarar mot fördelningen för ℓ_F när det i själva verket inte finns någon effekt (dvs när $\beta_F = 0$) och den representerar sannolikhetsfördelningen för summan av alla störfaktorer.

- Med tanke på att vi skattar modellens parametrar, som är väntevärden, mha medelvärden baserade på observerad data så kommer den observerade skattningen $\hat{\beta}_F = \ell_F$ att i de flesta fall vara skild från 0 när det kanske gäller att parametern β_F i själva verket är 0.

- Med tanke på att vi skattar modellens parametrar, som är väntevärden, mha medelvärden baserade på observerad data så kommer den observerade skattningen $\hat{\beta}_F = \ell_F$ att i de flesta fall vara skild från 0 när det kanske gäller att parametern β_F i själva verket är 0.
- Kan vi på något sätt statistiskt säkerställa att en observerad effekt är en faktisk effekt (med en given sannolikhet)?

- Med tanke på att vi skattar modellens parametrar, som är väntevärden, mha medelvärden baserade på observerad data så kommer den observerade skattningen $\hat{\beta}_F = \ell_F$ att i de flesta fall vara skild från 0 när det kanske gäller att parametern β_F i själva verket är 0.
- Kan vi på något sätt statistiskt säkerställa att en observerad effekt är en faktisk effekt (med en given sannolikhet)?
- Detta kallas att undersöka om en effekt är signifikant eller ej. Formellt representeras detta av ett s.k. hypotestest där den s.k. noll-hypotesen ges av H_0 : det finns ingen effekt ($\beta_F = \beta_F^0 = 0$). Alternativhypotesen ges av H_1 : det finns en effekt ($\beta_F \neq 0$).

- Med tanke på att vi skattar modellens parametrar, som är väntevärden, mha medelvärden baserade på observerad data så kommer den observerade skattningen $\hat{\beta}_F = \ell_F$ att i de flesta fall vara skild från 0 när det kanske gäller att parametern β_F i själva verket är 0.
- Kan vi på något sätt statistiskt säkerställa att en observerad effekt är en faktisk effekt (med en given sannolikhet)?
- Detta kallas att undersöka om en effekt är signifikant eller ej. Formellt representeras detta av ett s.k. hypotestest där den s.k. noll-hypotesen ges av H_0 : det finns ingen effekt ($\beta_F = \beta_F^0 = 0$). Alternativhypotesen ges av H_1 : det finns en effekt ($\beta_F \neq 0$).
- Väsentlig idé: Om den observerade skattningen är för extrem för vad den fördelning som H_0 svarar mot ($\beta_F = 0$) så förkastar vi H_0 till förmån för H_1 . Vi drar alltså slutsatsen att det finns en effekt.

- Med tanke på att vi skattar modellens parametrar, som är väntevärden, mha medelvärden baserade på observerad data så kommer den observerade skattningen $\hat{\beta}_F = \ell_F$ att i de flesta fall vara skild från 0 när det kanske gäller att parametern β_F i själva verket är 0.
- Kan vi på något sätt statistiskt säkerställa att en observerad effekt är en faktisk effekt (med en given sannolikhet)?
- Detta kallas att undersöka om en effekt är signifikant eller ej. Formellt representeras detta av ett s.k. hypotestest där den s.k. noll-hypotesen ges av H_0 : det finns ingen effekt ($\beta_F = \beta_F^0 = 0$). Alternativhypotesen ges av H_1 : det finns en effekt ($\beta_F \neq 0$).
- Väsentlig idé: Om den observerade skattningen är för extrem för vad den fördelning som H_0 svarar mot ($\beta_F = 0$) så förkastar vi H_0 till förmån för H_1 . Vi drar alltså slutsatsen att det finns en effekt.
- Man kan utföra hypotestestning mha konfidensintervall: om det skattade värdet på parametern hamnar utanför ett konfidensintervall centrerat kring det parametervärde som svarar mot H_0 (dvs $\beta_F = 0$) så förkastar vi H_0 .

- Med tanke på att vi skattar modellens parametrar, som är väntevärden, mha medelvärden baserade på observerad data så kommer den observerade skattningen $\hat{\beta}_F = \ell_F$ att i de flesta fall vara skild från 0 när det kanske gäller att parametern β_F i själva verket är 0.
- Kan vi på något sätt statistiskt säkerställa att en observerad effekt är en faktisk effekt (med en given sannolikhet)?
- Detta kallas att undersöka om en effekt är signifikant eller ej. Formellt representeras detta av ett s.k. hypotestest där den s.k. noll-hypotesen ges av H_0 : det finns ingen effekt ($\beta_F = \beta_F^0 = 0$). Alternativhypotesen ges av H_1 : det finns en effekt ($\beta_F \neq 0$).
- Väsentlig idé: Om den observerade skattningen är för extrem för vad den fördelning som H_0 svarar mot ($\beta_F = 0$) så förkastar vi H_0 till förmån för H_1 . Vi drar alltså slutsatsen att det finns en effekt.
- Man kan utföra hypotestestning mha konfidensintervall: om det skattade värdet på parametern hamnar utanför ett konfidensintervall centrerat kring det parametervärde som svarar mot H_0 (dvs $\beta_F = 0$) så förkastar vi H_0 .
- Om vi använder konfidensgraden $1 - \alpha$ så blir

$$\mathbb{P}(\text{typ I-fel}) = \mathbb{P}(\text{förkasta } H_0 \text{ när } H_0\text{-fördelningen var den sanna}) = \alpha$$

- I hypotestest kallas α , som vi vill ha liten, för signifikansnivån och om vi förkastar H_0 säger vi att vi har signifikans.

- I hypotestest kallas α , som vi vill ha liten, för signifikansnivån och om vi förkastar H_0 säger vi att vi har signifikans.
- Förkasta H_0 och konkludera signifikant effekt för faktor/samspel F om $\beta_F^0 = 0$ faller utanför konfidensintervallet $\ell_F \pm z_{1-\alpha/2}\sigma_{\ell_F} = \ell_F \pm z_{1-\alpha/2}2\sigma/\sqrt{nN}$ där vi minns att n är antalet mätningar vi gör per faktor-kombination (per rad i designmatrisen).

- I hypotestest kallas α , som vi vill ha liten, för signifikansnivån och om vi förkastar H_0 säger vi att vi har signifikans.
- Förkasta H_0 och konkludera signifikant effekt för faktor/samspel F om $\beta_F^0 = 0$ faller utanför konfidensintervallet $\ell_F \pm z_{1-\alpha/2}\sigma_{\ell_F} = \ell_F \pm z_{1-\alpha/2}2\sigma/\sqrt{nN}$ där vi minns att n är antalet mätningar vi gör per faktor-kombination (per rad i designmatrisen).
Är 0 inkluderat kan vi inte utesluta att effekten egentligen inte existerar och att det bara var "mätbrus" som gav oss ett värde skilt från 0.

- I hypotestest kallas α , som vi vill ha liten, för signifikansnivån och om vi förkastar H_0 säger vi att vi har signifikans.
- Förkasta H_0 och konkludera signifikant effekt för faktor/samspel F om $\beta_F^0 = 0$ faller utanför konfidensintervallet $\ell_F \pm z_{1-\alpha/2}\sigma_{\ell_F} = \ell_F \pm z_{1-\alpha/2}2\sigma/\sqrt{nN}$ där vi minns att n är antalet mätningar vi gör per faktor-kombination (per rad i designmatrisen). Är 0 inkluderat kan vi inte utesluta att effekten egentligen inte existerar och att det bara var "mätbrus" som gav oss ett värde skilt från 0.
- Ekvivalent: Förkasta H_0 om ℓ_F faller utanför det s.k. referensintervallet

$$0 \pm z_{1-\alpha/2}2\sigma/\sqrt{nN}$$

där vi som tidigare i kursen finner $z_{1-\alpha/2}$ i standard-normal-tabellen.

- I hypotestest kallas α , som vi vill ha liten, för signifikansnivån och om vi förkastar H_0 säger vi att vi har signifikans.
- Förkasta H_0 och konkludera signifikant effekt för faktor/samspel F om $\beta_F^0 = 0$ faller utanför konfidensintervallet $\ell_F \pm z_{1-\alpha/2}\sigma_{\ell_F} = \ell_F \pm z_{1-\alpha/2}2\sigma/\sqrt{nN}$ där vi minns att n är antalet mätningar vi gör per faktor-kombination (per rad i designmatrisen). Är 0 inkluderat kan vi inte utesluta att effekten egentligen inte existerar och att det bara var "mätbrus" som gav oss ett värde skilt från 0.
- Ekvivalent: Förkasta H_0 om ℓ_F faller utanför det s.k. referensintervallet

$$0 \pm z_{1-\alpha/2}2\sigma/\sqrt{nN}$$

där vi som tidigare i kursen finner $z_{1-\alpha/2}$ i standard-normal-tabellen.

- Precis som i tidigare konfidensintervall i kursen brukar α sättas till 0.05, 0.01 eller 0.001, beroende på vilken säkerhet vi vill ha (kom ihåg att mindre α betyder bredare intervall).

- I hypotestest kallas α , som vi vill ha liten, för signifikansnivån och om vi förkastar H_0 säger vi att vi har signifikans.
- Förkasta H_0 och konkludera signifikant effekt för faktor/samspel F om $\beta_F^0 = 0$ faller utanför konfidensintervallet $\ell_F \pm z_{1-\alpha/2}\sigma_{\ell_F} = \ell_F \pm z_{1-\alpha/2}2\sigma/\sqrt{nN}$ där vi minns att n är antalet mätningar vi gör per faktor-kombination (per rad i designmatrisen). Är 0 inkluderat kan vi inte utesluta att effekten egentligen inte existerar och att det bara var "mätbrus" som gav oss ett värde skilt från 0.
- Ekvivalent: Förkasta H_0 om ℓ_F faller utanför det s.k. referensintervallet

$$0 \pm z_{1-\alpha/2}2\sigma/\sqrt{nN}$$

där vi som tidigare i kursen finner $z_{1-\alpha/2}$ i standard-normal-tabellen.

- Precis som i tidigare konfidensintervall i kursen brukar α sättas till 0.05, 0.01 eller 0.001, beroende på vilken säkerhet vi vill ha (kom ihåg att mindre α betyder bredare intervall).
- Vad gör vi om vi inte har något känt värde på σ ? Vi tittar nu närmare på lite olika metoder som var och en har sitt eget användningsområde.

Metod 1:

- Även om vi inte känner till σ^2 så kan vi skatta den med punktskattaren S^2 vilket ger skattningen s^2 (vilket vi minns från föreläsningen om skattning).

Metod 1:

- Även om vi inte känner till σ^2 så kan vi skatta den med punktskattaren S^2 vilket ger skattningen s^2 (vilket vi minns från föreläsningen om skattning).
- Antag att vi har flera mätningar per rad i designmatrisen. Då får vi en variansskattning per rad: s_i^2 , $i = 1, \dots, N = 2^k$. Eftersom vi vill skatta den gemensamma variansen σ^2 så kombinerar vi dessa separata skattningar enligt ("pooled variance estimate")

$$s_p^2 = \frac{\sum_i (n_i - 1) s_i^2}{\sum_i (n_i - 1)}$$

för att få en skattning av σ^2 .

Metod 1:

- Även om vi inte känner till σ^2 så kan vi skatta den med punktskattaren S^2 vilket ger skattningen s^2 (vilket vi minns från föreläsningen om skattning).
- Antag att vi har flera mätningar per rad i designmatrisen. Då får vi en variansskattning per rad: s_i^2 , $i = 1, \dots, N = 2^k$. Eftersom vi vill skatta den gemensamma variansen σ^2 så kombinerar vi dessa separata skattningar enligt ("pooled variance estimate")

$$s_p^2 = \frac{\sum_i (n_i - 1) s_i^2}{\sum_i (n_i - 1)}$$

för att få en skattning av σ^2 .

- I enlighet med hur t -fördelningar kommer in i konfidensintervall för väntevärden när variansen är okänd och man skattar den med S^2 så får vi här referensintervallet

$$0 \pm t_{1-\alpha/2}(\nu) 2s_p / \sqrt{nN}$$

där antalet frihetsgrader är $\nu = \sum_i (n_i - 1)$ (det vi delar med i variansskattningen – jämför med frihetsgradsantalet i konfidensintervall för väntevärdet).

Exempel

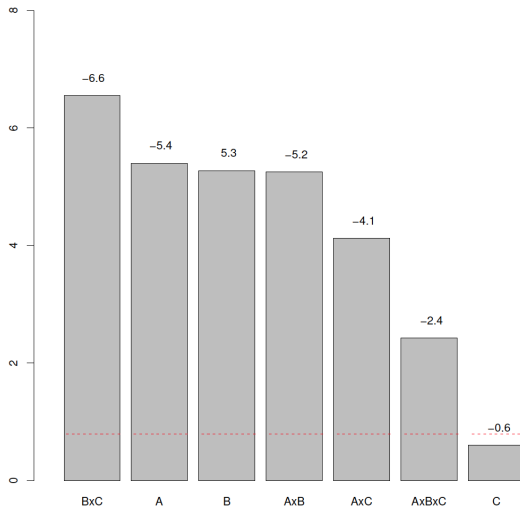
Nr	A	B	C	AB	BC	AC	ABC	Resultat	s^2
1	-	-	-	+	+	+	-	$y_1 = [3.7, 2.8]$	$s_1^2 = 0.405$
2	+	-	-	-	+	-	+	$y_2 = [4.8, 4.8]$	$s_2^2 = 0.0$
3	-	+	-	-	-	+	+	$y_3 = [18.7, 17.1]$	$s_3^2 = 1.28$
4	+	+	-	+	-	-	-	$y_4 = [13.5, 14.1]$	$s_4^2 = 0.18$
5	-	-	+	+	-	-	+	$y_5 = [10.1, 11.7]$	$s_5^2 = 1.28$
6	+	-	+	-	-	+	-	$y_6 = [8.8, 9.3]$	$s_6^2 = 0.125$
7	-	+	+	-	+	-	-	$y_7 = [17.7, 16.9]$	$s_7^2 = 0.320$
8	+	+	+	+	+	+	+	$y_8 = [0.4, -0.2]$	$s_8^2 = 0.18$

$$s_p^2 = \frac{1 \cdot s_1^2 + 1 \cdot s_2^2 + \dots}{8} = \frac{0.405 + 0 + 1.28 + \dots}{8} = 0.47125$$

$$t_{1-0.05/2}(\nu=8) \frac{2s_p}{\sqrt{8 \cdot 2}} = 2.306 \frac{2 \cdot \sqrt{0.47125}}{\sqrt{16}} = 2.306 \cdot 0.343 = 0.79233$$

Värdet på $t_{1-\alpha/2}(\nu)$ hittar vi i tabellen för t -fördelningar.

Paretodiagram för faktorer och samspel tillsammans med referenslinjen (värdet $t_{1-\alpha/2}(\nu = 8)2s_p/\sqrt{8 \cdot 2} = 0.79233$). De effekter vars staplar går över referenslinjen är signifikanta (kom ihåg att staplarna ger den absoluta storleken på effekterna).



Metod 2:

- Vanligtvis kostar mätningar både tid och pengar vilket kan medföra att vi inte har flera mätningar per nivå-kombination. Hur gör vi för att skatta σ^2 när man bara har tillgång till en mätning per nivå-kombination/rad?

Metod 2:

- Vanligtvis kostar mätningar både tid och pengar vilket kan medföra att vi inte har flera mätningar per nivå-kombination. Hur gör vi för att skatta σ^2 när man bara har tillgång till en mätning per nivå-kombination/rad?
- Antag att du känner till att några av effekterna egentligen inte existerar (typiskt handlar det om samspel av ordningen 3 och högre i en 2^k -design där $k \geq 3$; högre samspelseffekter som är skilda från 0 kan tolkas som utfall av bakgrundsvariabler/störfaktorer).

Metod 2:

- Vanligtvis kostar mätningar både tid och pengar vilket kan medföra att vi inte har flera mätningar per nivå-kombination. Hur gör vi för att skatta σ^2 när man bara har tillgång till en mätning per nivå-kombination/rad?
- Antag att du känner till att några av effekterna egentligen inte existerar (typiskt handlar det om samspel av ordningen 3 och högre i en 2^k -design där $k \geq 3$; högre samspelseffekter som är skilda från 0 kan tolkas som utfall av bakgrundsvariabler/störfaktorer).
- Sådana effektskattningar kan betraktas som utfall från referensfördelningen $N(0, 4\sigma^2/(nN))$ och kan således användas för skattning av σ^2 :

$$\frac{4s^2}{nN} = \frac{1}{N_J} \sum_{j \in J} \ell_j^2 \iff s^2 = \frac{nN}{4} \frac{1}{N_J} \sum_{j \in J} \ell_j^2$$

där J , med $N_J = \#J$, innehåller alla index på de effekter som antas inte existera.

Metod 2:

- Vanligtvis kostar mätningar både tid och pengar vilket kan medföra att vi inte har flera mätningar per nivå-kombination. Hur gör vi för att skatta σ^2 när man bara har tillgång till en mätning per nivå-kombination/rad?
- Antag att du känner till att några av effekterna egentligen inte existerar (typiskt handlar det om samspel av ordningen 3 och högre i en 2^k -design där $k \geq 3$; högre samspelseffekter som är skilda från 0 kan tolkas som utfall av bakgrundsvariabler/störfaktorer).
- Sådana effektskattningar kan betraktas som utfall från referensfördelningen $N(0, 4\sigma^2/(nN))$ och kan således användas för skattning av σ^2 :

$$\frac{4s^2}{nN} = \frac{1}{N_J} \sum_{j \in J} \ell_j^2 \iff s^2 = \frac{nN}{4} \frac{1}{N_J} \sum_{j \in J} \ell_j^2$$

där J , med $N_J = \#J$, innehåller alla index på de effekter som antas inte existera.

- Exempel 2.4 i boken illustrerar detta i en 2^4 -design, med $N = 2^4 = 16$, där $n = 1$ (en mätning per rad) och alla samspel av ordningen 3 och 4 antas tillhöra J (notera att de betecknar $\frac{4s^2}{nN}$ med s_{effekt}).

Metod 2:

- Vanligtvis kostar mätningar både tid och pengar vilket kan medföra att vi inte har flera mätningar per nivå-kombination. Hur gör vi för att skatta σ^2 när man bara har tillgång till en mätning per nivå-kombination/rad?
- Antag att du känner till att några av effekterna egentligen inte existerar (typiskt handlar det om samspel av ordningen 3 och högre i en 2^k -design där $k \geq 3$; högre samspelseffekter som är skilda från 0 kan tolkas som utfall av bakgrundsvariabler/störfaktorer).
- Sådana effektskattningar kan betraktas som utfall från referensfördelningen $N(0, 4\sigma^2/(nN))$ och kan således användas för skattning av σ^2 :

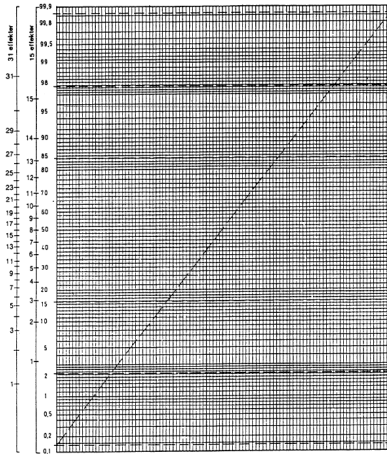
$$\frac{4s^2}{nN} = \frac{1}{N_J} \sum_{j \in J} \ell_j^2 \iff s^2 = \frac{nN}{4} \frac{1}{N_J} \sum_{j \in J} \ell_j^2$$

där J , med $N_J = \#J$, innehåller alla index på de effekter som antas inte existera.

- Exempel 2.4 i boken illustrerar detta i en 2^4 -design, med $N = 2^4 = 16$, där $n = 1$ (en mätning per rad) och alla samspel av ordningen 3 och 4 antas tillhöra J (notera att de betecknar $\frac{4s^2}{nN}$ med s_{effekt}).
- Fördel om man både kan ha upprepade mätningar ($n \geq 2$) och exkludera effekter.

Metod 3:

- Denna sista metod är en grafisk metod där vi använder oss av ett s.k. normalfördelningspapper (en s.k. Q-Q-plot för en normalfördelning).
- Tillåter oss att identifiera ifall mycket av datan ser ut att följa samma normalfördelning.
- Det går även att se vilka datapunkter som inte verkar följa samma fördelning.
- I fallet av våra effekter så kan vi således se vilka som verkar följa referensefördelningen (icke-signifikanta) samt vilka som inte verkar följa referensefördelningen (signifikanta).
- Fördelen är att vi inte behöver antaga vilka faktorer som är effektlösa (Metod 2) och att vi inte heller kräver många olika mätningar inom samma grupp (Metod 1).
- Nackdelen är att tolkningen av normalfördelningsdiagrammet är subjektiv: vad är tillräckligt långt från den linje som indikerar om en effekt tillhör referensfördelningen och hur skall man manuellt välja linjen så den passar så bra som möjligt ihop med punkterna?



(a) Normalfördelningspapper

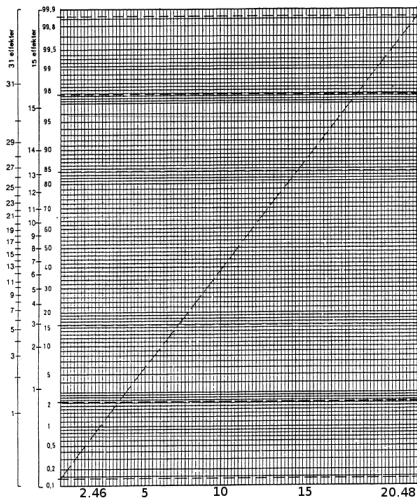
Diagram där man på y-axeln har skrivit ut sannolikheter med ett avstånd emellan som kommer ge en rak linje för en normalfördelning.

- Man kan använda ett sådant diagram för att avgöra om mätningar verkar komma från en normalfördelning eller inte. Man kan även skatta väntevärde samt standardavvikelse för normalfördelad data.
- Sortera mätningarna från minsta till största. Sortera deras värden som x-axeln i diagrammet, x_i blir värdet på den mätning som var nummer i i ordningen från minst till störst.
- Skatta empiriska fördelningsfunktionen genom att ge varje mätning en sannolikhet beroende på dess plats i den sorterade ordningen, $p_i = \frac{i-0.5}{k}$.
- Är datan normalfördelad skall punkterna (x_i, p_i) nu ligga (approximativt) på en linje.

- Vi kan använda ett normalfördelningsdiagram för att avgöra vilka skattade effekter som är signifikanta.
- Först skattar vi alla effekter. Sedan sorterar vi dessa effekter och räknar ut motsvarande sannolikhetsvärde så att vi kan rita in dem i normalfördelningsdiagrammet.
- Vi tittar sedan på vilka punkter som ser ut att ligga på en linje och drar ett streck genom dessa. De punkter som inte ligger nära linjen anser vi är signifikanta effekter då de inte ser ut att följa referensfördelningen.

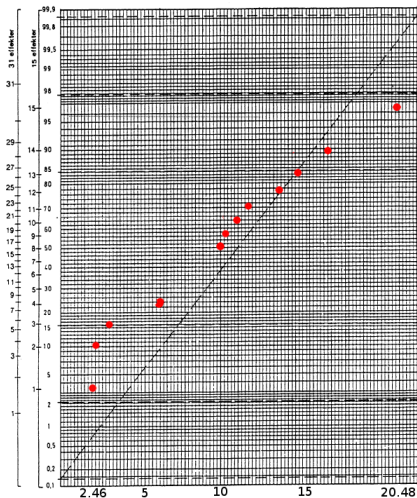
Exempel

i	x_i	p_i
1	2.46	0.033
2	2.52	0.100
3	3.46	0.167
4	5.99	0.233
5	7.45	0.3
6	9.02	0.367
7	9.14	0.433
8	9.98	0.500
9	10.18	0.567
10	11.21	0.633
11	12.34	0.70
12	13.44	0.767
13	14.23	0.833
14	15.98	0.90
15	20.48	0.967



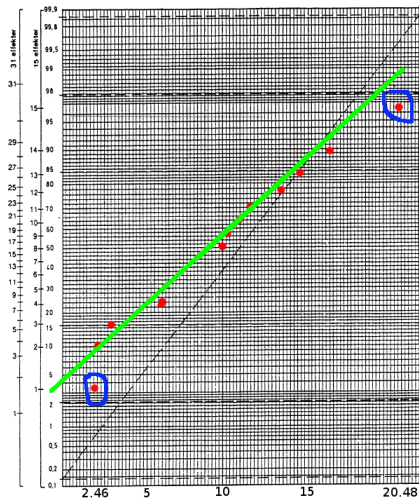
Exempel

i	x_i	p_i
1	2.46	0.033
2	2.52	0.100
3	3.46	0.167
4	5.99	0.233
5	7.45	0.3
6	9.02	0.367
7	9.14	0.433
8	9.98	0.500
9	10.18	0.567
10	11.21	0.633
11	12.34	0.70
12	13.44	0.767
13	14.23	0.833
14	15.98	0.90
15	20.48	0.967



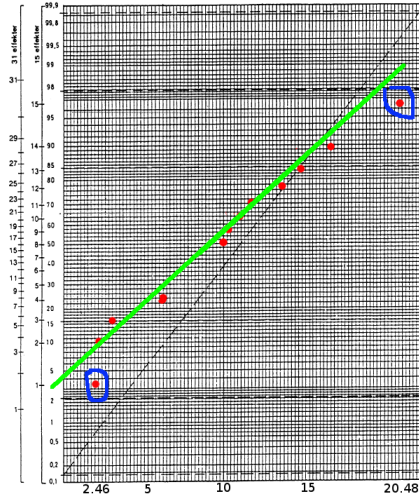
Exempel

i	x_i	p_i
1	2.46	0.033
2	2.52	0.100
3	3.46	0.167
4	5.99	0.233
5	7.45	0.3
6	9.02	0.367
7	9.14	0.433
8	9.98	0.500
9	10.18	0.567
10	11.21	0.633
11	12.34	0.70
12	13.44	0.767
13	14.23	0.833
14	15.98	0.90
15	20.48	0.967



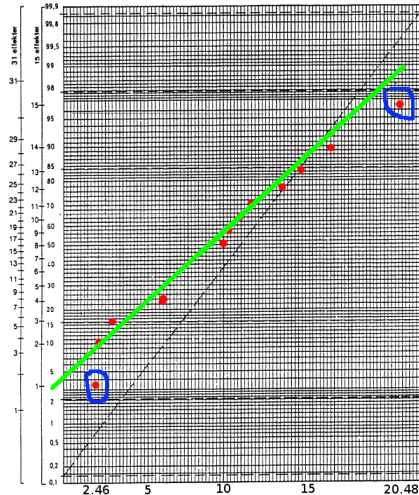
Exempel

Väntevärdet skattas genom att ta det x-värde på det gröna strecket som motsvarar y-värdet 0.5. I det här fallet motsvarar det ungefär 9. Datan i det här exemplet är genererad slumpmässigt. För en riktig referensfördelning borde väntevärdet vara nära 0.



Exempel

Standardavvikelsen
skattas genom att
ta skillnaden mellan
x-värdena där det gröna
strecket korsar de tjocka
streckade horisontella
linjerna och dela med
4. $s \approx \frac{19+1}{4} = 5$.



Vi vet att referensfördelningen skall ha väntervärdet 0. Vi kan använda den kunskapen när vi ritar ut vårt streck i normalfördelningsdiagrammet.

Tips: anpassa streck för referensfördelningen

- Det anpassade strecket skall gå igenom punkten som har x-värdet 0 och y-värdet 50%.
- Strecket skall passa "så bra som möjligt" till de skattade effekterna omkring $x = 0$ då dessa är minst och därför troligast tillhör referensfördelningen.