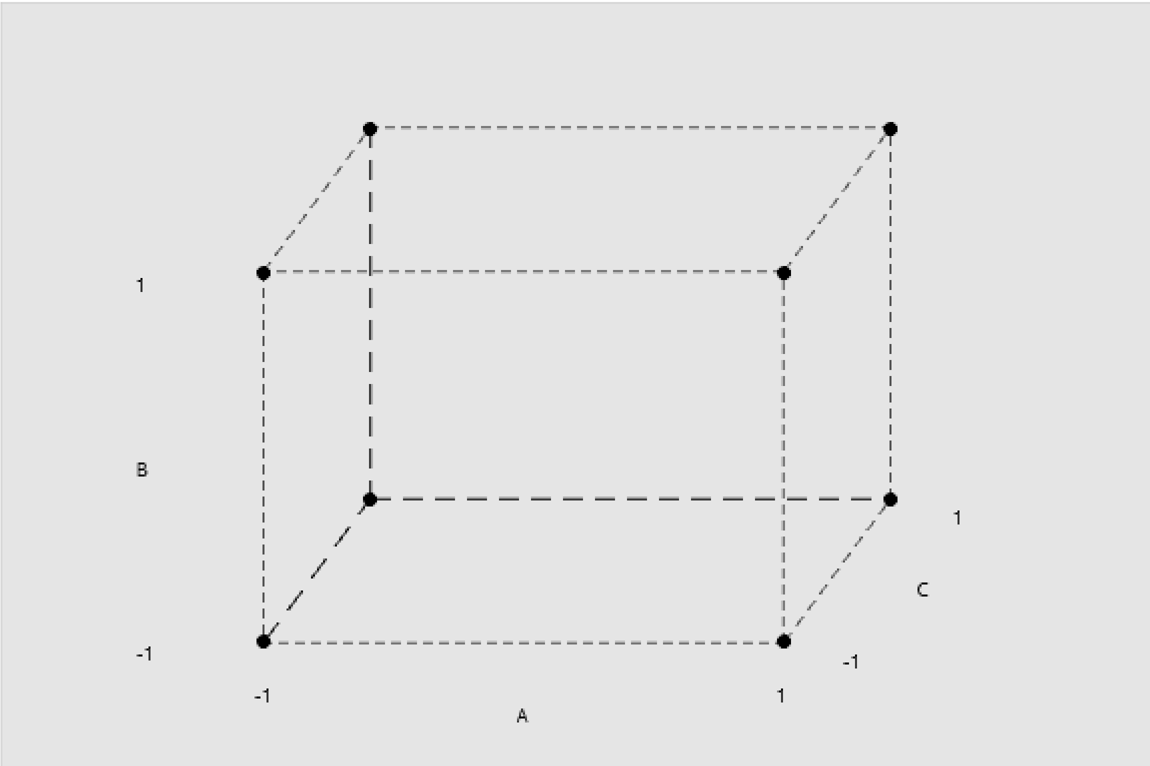
# To-nivå Faktorielle Forsøk

av

John Tyssedal



# Innleing

Experimentering er ein gammal disiplin, men moderne forsøksplanleggings tek utgangspunkt i pionerarbeidet til Ronald Aylmer Fisher (1890-1962) ved Rothamsted forsøksstasjon der han blei tilsett som statistikar i 1919. Rothamsted var eit forskningsinstitutt retta mot landbruksforskning. Fisher fekk fort erfare problema med å analysere tilfeldig innsamla data og innsåg fordelane det ville vere med å samle inn data på ein planlagt og kontrollert måte. I 1935 gav han ut den berømte boka si *Design of Experiments*. Ideane hans fann overraskande fort vegen inn i industrien, men der støtte dei på ei hindring for suksess. Eksperiment i landbruket er ofte store med fleire variable som kvar kan ha mange nivå og gjentak var ofte nødvendig. Dessutan kunne dei ta lang tid å gjennomføre. Eksperiment i industrien kan vere dyre, noko som gjer at kostnaden må takast med i vurderinga av kor mange forsøk ein skal utføre. På den andre sida, i motsetning til landbruket der ein sår om våren og haustar om hausten, gjev dei ofte øyeblikkeleg respons og nye forsøk kan planleggast og utførast neste veke. På slutten av 1940-talet oppdaga George Box (1919-2013) at sekvensiell eksperimentering, der ein i kvart steg utførte mindre forsøk med få nivå for kvar faktor, kunne bringe ein produksjonsprosess mykje raskare nær optimale operasjonsbetingelsar. Det som nå følgjer er ei lita innføring i det mest basale innanfor forsøksplanlegging der kvar faktor berre har to nivå.

# To-nivå forsøk

La oss starte med eit eksempel.

**<u>Eksempel 1. Fjærer med sprekkdanning</u>**

Ved produksjon av stål-karbon fjærer hadde ein store problem med sprekker i fjærene. Basis metallurgikunnskap tilsa at mogelege årsaker til dette kunne vere temperaturen i stålet før avkjøling og kor mykje karbon ein hadde i blandinga med stålet. Det blei utført 4 forsøk der ein nytta 2 verdiar (nivå) for kvar av faktorane: temperatur og mengde karbon i blandinga. Dette gjev 4 mogelege nivåkombinasjonar. For kvar nivåkombinasjon blei det registrert prosentvis kor mange fjærer som ikkje hadde sprekkdanning. Målet med forsøket var å finne verdiar for dei to faktorane der sprekkdanninga var minst.

<u>Tabell 1. Faktorar, nivå og data frå forsøket med stål-karbon fjærer</u>

| Faktor\Nivå | Lågt | Høgt |
|---|---|---|
| Temperatur | 1450F | 1600F |
| Mengde Karbon | 0.5% | 0.7% |

| Forsøksnr. | Temperatur | Mengde karbon i % | Prosentvis fjærer utan sprekkdanning |
|---|---|---|---|
| 1 | 1450 F | 0.5 | 67 |
| 2 | 1600 F | 0.5 | 79 |
| 3 | 1450 F | 0.7 | 61 |
| 4 | 1600F | 0.7 | 75 |

Når ein utfører eit forsøk skal enkeltforsøka utførast i randomisert rekkjefølgje. Randomisering er vår beste garanti for uavhengige observasjonar og gjer at utanforliggande faktorar får mindre sjanse til å påverke responsen og med det medvirke til feilaktige konklusjonar. Det er og viktig å stille inn alle nivåkombinasjonar på nytt mellom kvart forsøk for å sikre at alle observasjonane får mest mogeleg lik varians.

La oss no innføre faktorane

A = temperatur i stålet før avkjøling

B = prosentvis mengde karbon i stålet

Responsvariablane, som er tilfeldig variable, vil vi kalle $Y_i$, $i = 1, 2, \ldots n$. Observert verdi for $Y_i$ i forsøk i skal vi kalle $y_i$ og talet på observasjonar skal vi kalle *n*. I dette forsøket er $n = 4$. For å finne ut om nokon av faktorane påverkar prosentvis sprekkdanning kan ein estimere effektane av faktorane. I forsøk med to nivå for kvar faktor vil desse delast inn i hovedeffektar og samspelseffektar.

## Definisjon av effektar

**Definisjon hovedeffekt**:

*For to-nivå forsøk definerer ein hovedeffekten av ein faktor som: Forventa gjennomsnittsrespons når faktoren er på høgt nivå – forventa gjennomsnittsrespons når faktoren er på lågt nivå.*

Det er naturleg at estimatet for denne effekten då blir $\bar{y}_H - \bar{y}_L$, det vil sei gjennomsnittet på høgt nivå – gjennomsnittet på lågt nivå, som for temperatur blir:

$$\hat{A} = \frac{79+75}{2} - \frac{61+67}{2} = 13,$$

og for mengde karbon

$$\hat{B} = \frac{61+75}{2} - \frac{79+67}{2} = -5$$

Det er og mogeleg og finne ut om hovedeffektar avheng av nivået for andre effektar. Om slike effektar finst, blir dei kalla samspelseffektar og er definert slik

**Definisjon samspelseffekt mellom 2 faktorar**

*For to-nivå forsøk definerer ein samspelseffekten mellom to faktorar (tofaktorsamspelet) som: Halvdelen av hovedeffekten av ein faktor når den andre er på høgt nivå – halvdelen av hovedeffekten av faktoren når den andre er på lågt nivå.*

Samspelseffekten mellom temperatur og mengde karbon blir då estimert til:

$$\widehat{AB} = \frac{75-61}{2} - \frac{79-67}{2} = 1$$

Forsøket ovanfor med to faktorar kvar på 2 nivå, blir kalla eit $2^2$ forsøk. La oss no kode om faktorverdiane med følgjande lineære transfomasjon:

$$x_{\mathrm{A}} = \frac{\mathrm{A} - 1525}{75}$$
$$x_{\mathrm{B}} = \frac{\mathrm{B} - 0.6}{0.1}$$

Verdiane for faktorane blir først sentrert og deretter delt ned på halvparten av avstanden mellom høgt og lågt nivå. Dette gjev følgjande koda verdiar der vi har laga ein ekstra kolonne $x_{\mathrm{AB}} = x_{\mathrm{A}} \cdot x_{\mathrm{B}}$

| $x_{\mathrm{A}}$ | $x_{\mathrm{B}}$ | $x_{\mathrm{AB}} = x_{\mathrm{A}} \cdot x_{\mathrm{B}}$ |
|---|---|---|
| -1 | -1 | 1 |
| 1 | -1 | -1 |
| -1 | 1 | -1 |
| 1 | 1 | 1 |

Legg merke til at dei tre kolonnane er ortogonale.

Når vi skal analysere to-nivå forsøk, er det oftast mest praktisk å omkode faktorverdiane til 1 og –1 som ovanfor. Legg merke til at -1 svarar til lågt nivå og 1 til høgt nivå. Eit anna argument er at vi ofte har kvalitative faktorar som t. d: Vi skal teste ut 2 forskjellige merker eller vi vil teste ut kva som skjer med og utan ei behandling. I det siste tilfellet vil ein gjerne og ha eit mål på kva effekt behandlinga har på responsen.

## Forteikn for å rekne ut effektane

For to-nivå forsøk gjeld: Alle kvantitative nivå kan kodast om til –1 og 1. Alle kvalitative nivå kan naturleg setjast til dette. Om vi vedtek konvensjonen at høgt nivå av ein faktor svarar til 1 og lågt nivå til –1, ser vi at vi fekk dei estimerte effektane ved å legge saman responsverdiane med forteikn bestemt av forteikna i kolonnene for $x_{\mathrm{A}}$, $x_{\mathrm{B}}$ og $x_{\mathrm{AB}}$, for deretter å dele ned på halvparten av antall observasjonar. Difor lagar vi oss ofte berre ei forteiknsmatrise der dei naudsynte forteikna for utrekning av effektar står. I vårt eksempel blir denne:

| A | B | AB |
|---|---|---|
| - | - | + |
| + | - | - |
| - | + | - |
| + | + | + |

## Andre notasjonar for faktorkombinasjonar

Ein kan og kome bort i denne nivåkodinga.

Høgt nivå er markert med bokstaven for faktoren.
Lågt nivå er markert med 1.
-1 er utelaten dersom andre bokstavar er brukt.

Eit eksempel på ei slik nivåkoding for eit $2^2$ forsøk er gitt nedanfor.

| A | B | Nivåkode |
|---|---|---|
| - | - | 1 |
| + | - | a |
| - | + | b |
| + | + | ab |

## $2^3$ forsøk

Eit $2^k$ forsøk har k faktorar, kvar på 2 nivå. I eksempel 1 såg vi på eit $2^2$ forsøk og korleis vi kunne estimere effektane i eit slikt forsøk. Eigenleg var det i dette forsøket med ein tredje faktor som var temperaturen i oljen brukt til å kjøle ned fjærene med. Nivåa for denne var valgt som følgjer:

| Faktor\Nivå | Lågt | Høgt |
|---|---|---|
| Temperatur i kjølevæske | 70F | 120F |

Med tre faktorar som kvar er på to nivå, er det mogleg å lage 8 mogelege nivåkombinasjonar av høg og låg eller + og – om vi berre vil bruke forteikna etter omkoding. La A og B vere som før og la C vere temperaturen i oljen brukt til avkjøling. Då kan vi setje opp følgjande forteiknsmatrise utvida med nivåkode og responsverdiar.

Tabell 2. Det fullstendige $2^3$ forsøket for eksperimenteringa med stål-karbon fjærer sett opp med hovedeffekts- og samspelskolonnar.

| Forsøksnr. | A | B | C | AB | AC | BC | ABC | Nivåkode | y |
|---|---|---|---|---|---|---|---|---|---|
| 1 | - | - | - | + | + | + | - | 1 | 67 |
| 2 | + | - | - | - | - | + | + | a | 79 |
| 3 | - | + | - | - | + | - | + | b | 61 |
| 4 | + | + | - | + | - | - | - | ab | 75 |
| 5 | - | - | + | + | - | - | + | c | 59 |
| 6 | + | - | + | - | + | - | - | ac | 90 |
| 7 | - | + | + | - | - | + | - | bc | 52 |
| 8 | + | + | + | + | + | + | + | abc | 87 |

Legg merke til rekkjefølgja forteikna for dei tre faktorane er sett opp i. Dette kallast standardform. Ved utføring av eit førsøk skal ein randomisere rekkjefølgja ein gjer forsøka i. Forsøka kan likevel alltid setjast opp på standardform etter at dei er utført om ein ynskjer. Det er ofte på standardform ein møter forsøk i litteraturen.

Estimerte hovedeffektar blir:

$$\hat{A} = \frac{79+75+90+87}{4} - \frac{62+61+59+52}{4} = 23,$$
$$\hat{B} = \frac{61+75+52+87}{4} - \frac{67+79+52+90}{4} = -5,$$
$$\hat{C} = \frac{59+90+52+87}{4} - \frac{67+79+61+75}{4} = 1.5.$$

Når det gjeld samspelet mellom A og B, AB, skal vi finne hovedeffekten av A når B er på høgt nivå og trekkje frå hovedeffekten av A når B er på lågt nivå for deretter å dele på 2. Når vi estimerer effekten skal vi altså der B er på høgt nivå, leggje saman responsverdiane med forteikna i kolonnen til A (forsøk nr. 3,4,7 og 8), og der B er på lågt nivå skal vi leggje dei saman med motsette forteikn av det som står i kolonnen til A. Men dette svarar til å bruke forteikna i ein kolonne der forteikna i kolonnen A og kolonnen B er multiplisert saman rekkjevis, det vil sei forteikna i kolonnen for AB. Tilsvarande får ein forteikna i dei andre samspelskolonnane. Dersom eit tofaktorsamspel er avhengig av nivået på ein anna faktor har vi eit trefaktorsamspel. Forteikna i kolonnen for trefaktorsamspelet ABC får ein ved å multiplisere saman rekkjevis forteikna i kolonnane for A, B og C. Dette gjev:

$$\hat{AB} = \frac{67+75+59+87}{4} - \frac{79+61+90+52}{4} = 1.5,$$
$$\hat{AC} = \frac{67+61+90+87}{4} - \frac{79+75+59+52}{4} = 10,$$
$$\hat{BC} = \frac{67+79+52+87}{4} - \frac{61+75+59+90}{4} = 0,$$
$$\hat{ABC} = \frac{80+52+54+72}{4} - \frac{83+45+60+68}{4} = 0.5.$$

## Vurdering av signifikans av effektar

Vi observerer at nokre estimerte effektar er relativt store medan andre er relativt små. Korleis kan vi avgjere kven av effektane som er signifikante når vi ikkje har estimat for eller veit variansen til responsvariabelen? Dei mest vanlege teknikkane er å bruke normalplott, gjerne saman med Lenth's metode. La A vere ein vilkårleg faktor. Estimatoren for effekten til A er då gitt ved:

$$\hat{A} = \frac{\sum_{i=1}^{n} \delta_i Y_i}{n/2}$$

der *n* er talet på observasjonar og $\delta_i$ er anten 1 eller –1 avhengig av forteikna i kolonnen for den effekten vi reknar ut. Sidan alle $Y_i$ er uavhengige får vi:

$$Var\left(\hat{A}\right) = \sigma^2_{effekt} = \sum_{i=1}^{n} \frac{\delta_i^2 Var\left(Y_i\right)}{n^2/4} = \frac{4\sum_{i=1}^{n}\sigma^2}{n^2} = \frac{4\sigma^2}{n}.$$

Dersom alle effektane er 0, og dataane er normalfordelte med same varians, $\sigma^2$, vil alle estimatorane for effektane vere $\mathrm{N}\left(0,\ \frac{4\sigma^2}{n}\right)$. Ved å framstille dei estimerte effektane i eit normalplott, kan dette brukast til å anslå kva effektar som er dei viktigaste.

**Normalplott**

Normalplott kan lagast som følgjer. Gå ut i frå at vi har $n$ uavhengige observasjonar $y_1, y_2, \ldots, y_n$ som alle kjem frå same fordeling. La oss så ordne desse etter algebraisk storleik og la dei ordna verdiane vere: $y_{(1)}, y_{(2)}, \ldots, y_{(n)}$. Eit direkte estimat av fordelingsfunksjonen $F(y_i) = P\left(\mathrm{Y} \le y_{(i)}\right), i = 1, 2, \ldots, n$ er då gitt ved $i/n$. Teoretisk kan det visast at $F_i = \frac{i - 3/8}{n + 1/4}$ er eit betre val. For dei estimerte effektane frå $2^3$ forsøket i Tabell 2 kan vi konstruere følgjande tabell.

Tabell 3. Tabell for konstruksjon av normalplott.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $y_{(i)}$: | -5 | 0 | 0.5 | 1.5 | 1.5 | 10 | 23 |
| $F_i$ : | 0.086 | 0.224 | 0.362 | 0.500 | 0.638 | 0.776 | 0.914 |
| $\Phi^{-1}(F_i)$ | - 1.37 | -0.76 | -0.35 | 0 | 0.35 | 0.76 | 1.37 |

Motivasjonen for den tredje rekka er som følgjer: For data frå ei normalfordeling med forventing $\mu$ og varians $\sigma^2$, er $F_i \approx F\left(y_{(i)}\right) = \Phi\left(\frac{y_{(i)} - \mu}{\sigma}\right)$ der $\Phi(.)$ er fordelingsfunsksjonen i ei standard normalfordeling. Det er difor å forvente at $\Phi^{-1}(F_i) \approx \frac{y_{(i)} - \mu}{\sigma}$ slik at eit plott av $\Phi^{-1}(F_i) \approx \frac{y_{(i)} - \mu}{\sigma}$ mot $y_{(i)}$ tilnærma blir ei rett linje.

Dersom alle effektane ikkje er signifikant forskjellige frå 0, skulle alle i eit normalplott bli liggande på ei rett linje. Dei som fell av, kan vi anslå å vere signifikante. Å fastslå signifikans frå normalplottet er ei vurderingssak, for ein vil aldri ha det slik at alle dei effektane som ikkje er signifikante ligg nøyaktig på linja. Difor kan det vere nyttig å støtte seg på andre metodar, så som Lenth's metode. Litt bakgrunn for denne er gitt nedanfor

**Lenth's metode**

For ein tilfeldig variabel $Z \sim N\left(0, \tau^2\right)$, vil medianen til $|Z|$ tilnærma vere gitt ved $0.675\tau$ noko som gjev at $1.5 \cdot \left(\text{medianen of } |\mathrm{Z}|\ \right) \approx 1.5 \cdot 0.675\tau \approx 1.0125\tau \approx \tau$. La $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m$ vere uavhengige estimatorar for effektane og la variansen deira vere $\tau^2$. Ein estimator for $\tau$ er då gitt ved:

$$\hat{\tau} = 1.5 \cdot \text{Median}\left(\left|\hat{\theta}_i\right|, i = 1, 2, \ldots, m\right)$$

Ein meir bearbeida estimator der ein forsøker å ta bort effektar som ein må gå ut frå er signifikante er:

$$\hat{\tau}^* = 1.5 \cdot \text{Median}\left(\left|\hat{\theta}_i\right| : \left|\hat{\theta}_i\right| < 2.5\hat{\tau}\right)$$

$\tau^*$ blir på engelsk kalla «pseudo standard error», forkorta til PSE.

For å teste $H_0 : \theta_i = 0$ mot $H_1 : \theta_i \neq 0$ brukar vi at $T = \dfrac{\hat{\theta}_i}{PSE}$ er tilnærma t-fordelt med $m/3$ fridomsgrader. Ved simuleringar har det vist seg at dette blir vel konservativt og at dersom ein forkastar $H_0$ når $\left|\hat{\theta}_i\right| \geq c \cdot PSE$ og brukar 5% signifikansnivå, så kan verdiane i tabell 4 nedanfor fungere betre

Tabell 4. Verdiar av *c* for utrekning av kritiske verdiar ved bruk av Lenth's metode.

| *m* | *c* |
|---|---|
| 7 | 2.3 |
| 15 | 2.16 |
| 31 | 2.06 |
| 63 | 2.01 |

Dersom ein vil ha meir nøyaktige verdiar for *c*, kan ein finne dette i boka: Experiments, Planning, Analysis and Optimization av Hamada and Wu (2009), side 702.

**MINITAB**

MINITAB er ein brukarvennleg statistikkprogrampakke, opprinneleg laga for bruk i undervisning. Siste versjon finst både for Windows og Mac. Det er lett å analysere standard forsøksopplegg med MINITAB. Den er menystyrt og har eit regneark som ein lett kan legge inn data i. NTNU har universitetslisens på denne som mellom anna innber at alle studentar kan laste ned programvaren til si personlege datamaskin.

For å gjere ein analyse av eit utført forsøk i MINITAB gjer ein følgjande:
I menyen øverst til venstre gå til Stat og klikk vidare

Stat $\rightarrow$ DOE $\rightarrow$ Factorial $\rightarrow$ Create factorial design

I boksen som kjem opp, vel talet på faktorar som her er 3, klikk på «Designs», velg «Full factorial». Trykk OK. Deretter klikk på «Options» og i noverande situasjon klikk på avhaking (fjern) for «randomize runs». (MINITAB tilbyr alltid ei randomisert rekkjefølgje å utføre forsøka i). Klikk OK og deretter OK. Det skal no vere 7 kolonnar i regnearket. Dei 4 første har informasjon om rekkjefølgja forsøka er utført i, senterpunkt og blokkdeling. Dei 3 siste kolonnane inneheld koda nivåkombinasjonar for faktorane. Faktorane har standardnamn A, B og C. Sjølv om du skriv om med andre namn, følgjer desse bokstavane faktorane i alle plott. Du må no legge inn dei 8 responsverdiane i ein kolonne og er då klar til å analysere forsøket. Då går du til Stat og klikkar vidare:

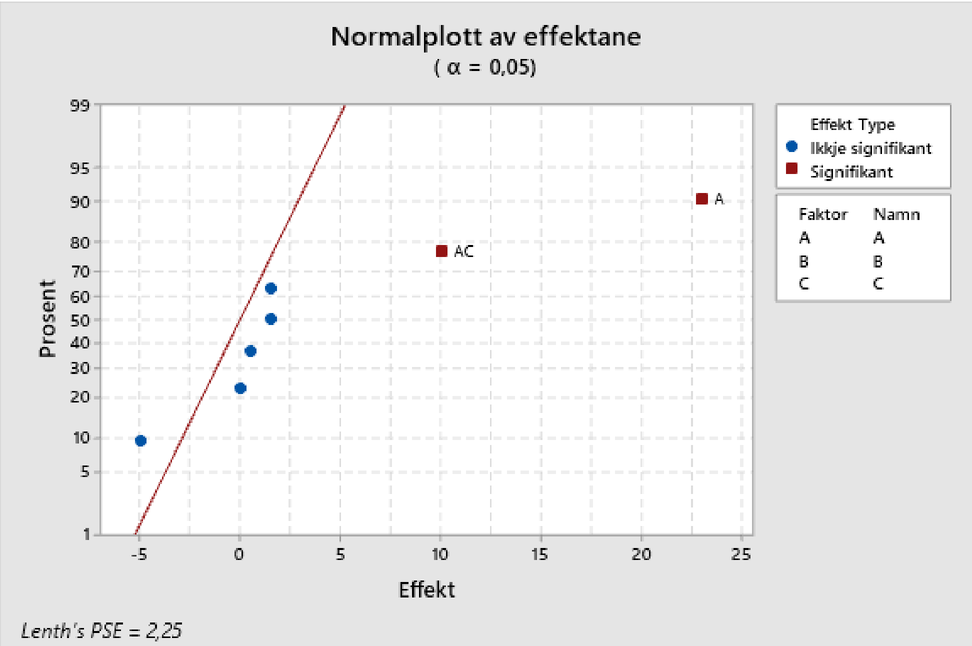
Stat $\rightarrow$ DOE $\rightarrow$ Factorial $\rightarrow$ Analyze factorial design

Du må bestemme respons. Klikk deretter på «Graphs» og hak av normalplott. Klikk OK og deretter OK. I den ferdige analysen som følgjer, er det her nok å sjå på den første tabellen der du finn dei estimerte effektane og dei to plotta, eit nomalplott, figur 1, og eit paretodiagram, figur 2, som baserer seg på Lenth's metode. Sjølv om ein forandrar variabelnamn i regnearket, vil plotta framleis sverge til standardnotasjonen. Dei har rett nok ei omkoding. Begge plotta peikar mot at det er temperaturen i stålet, samt sampelet mellom denne temperaturen og temperaturen i oljen brukt til avkjøling som har effekt. Dessutan blir dei effektane som er vurdert signifikante med Lenth's metode markert med rød firkantar i normalplottet. I paretodiagrammet, figur 2, finn vi ein kritisk verdi for absoluttverdien til effektane på 8.47. Den svarar til ei t-fordeling med 7/3 fridomsgrader og viser at MINITAB er heller konservativ i vurdering av signifikans med Lenth's metode. Men om ein hadde brukt $2.25 \cdot 2.3 = 5.18$ hadde ikkje hovedeffekten av karbon blitt signifikant likevel.

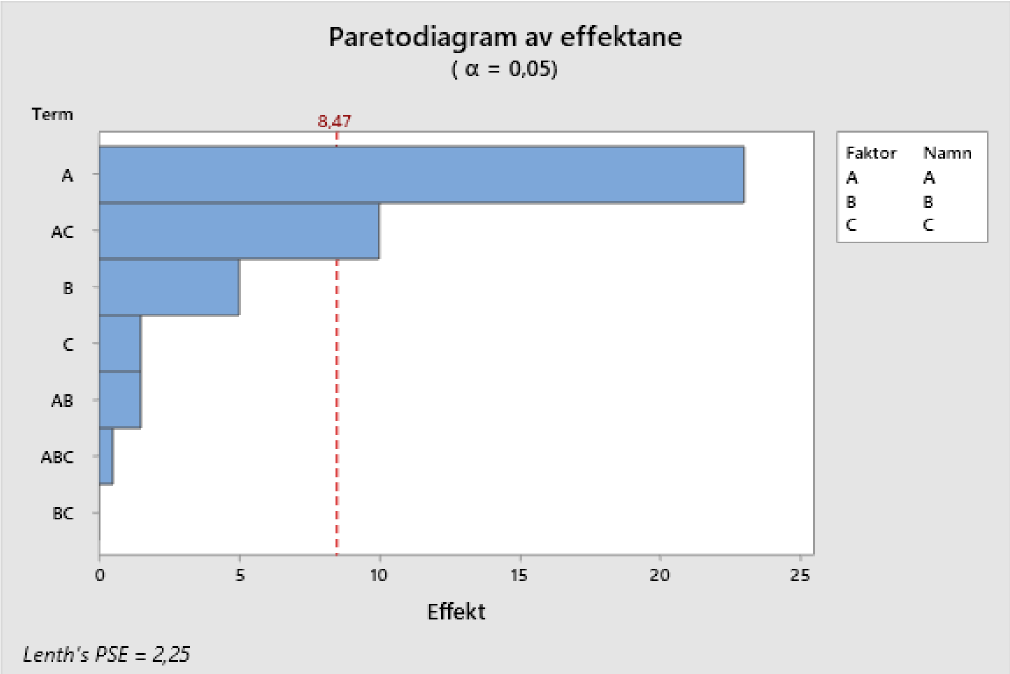
Om ein ser på y-aksen i normalplottet så stemmer ikkje den med verdien i tabell 3. For x = -5 skulle y vere -1,37 og y skulle vere -0.76 for x = 0. Istaden set MINITAB $F_1 = 0.086$ og $F_2 = 0.224$ for desse verdiane. Meir nøyaktig brukar MINITAB den standardiserte normalfordelinga og set 1% persentilen, -2.33, til 1%, 5% persentilen, -1,645, til 5%, 10% persentilen, -1,28, til 10% osb.

## Coded Coefficients

| Term | Effect | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|---|
| Constant | | 71,25 | * | * | * | |
| A | 23,00 | 11,50 | * | * | * | 1,00 |
| B | -5,000 | -2,500 | * | * | * | 1,00 |
| C | 1,5000 | 0,7500 | * | * | * | 1,00 |
| A*B | 1,5000 | 0,7500 | * | * | * | 1,00 |
| A*C | 10,000 | 5,000 | * | * | * | 1,00 |
| B*C | -0,000000 | -0,000000 | * | * | * | 1,00 |
| A*B*C | 0,5000 | 0,2500 | * | * | * | 1,00 |




Figur 1. Normalplott av effektane frå forsøket med stål-karbon fjærer.



Figur 2. Paretodiagram av effektane frå forsøket med stål-karbon fjærer.

## Bruk av effektar for å rekne ut signifikans

Eit forsøk kan ha mange høgare ordens effektar (trefaktorsamspel og høgare ordens samspel). Om ein går ut i frå at desse kan neglisjerast, kan ein bruke desse til å estimere variansen til effektane. I eit $2^4$ forsøk er det 4 trefaktorsamspel og eit firefaktorsamspel. Ved å ta gjennomsnittet av dei 5 kvadrerte trefaktor- og firefaktorsamspela får vi eit estimat for variansen til effektane med 5 fridomsgrader. Dette kan brukast til å undersøke signifikansen til effektane. Til dømes er hovedeffekten av ein faktor A signifikant dersom:

$$\left|\frac{\hat{A}}{s_{\hat{A}}}\right| \geq t_{\frac{\alpha}{2},5}\,.$$

**Eksempel 2**

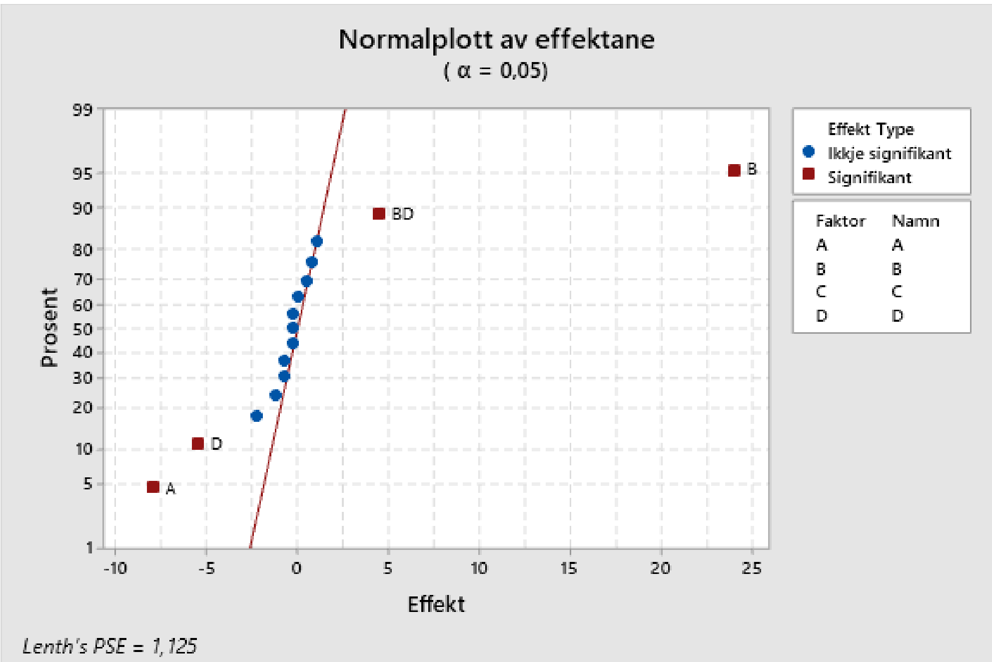
Eit $2^4$ forsøk i dei 4 faktorane A = katalysator ladning, B = temperatur, C = trykk og D = konsentrasjon blei utført i eit prosessutviklings-studie. Dei 16 nivåkombinasjonane saman med responsverdien er gitt nedanfor.

| A | B | C | D | Prosentvis omdanning |
|---|---|---|---|---|
| -1 | -1 | -1 | -1 | 71 |
| 1 | -1 | -1 | -1 | 61 |
| -1 | 1 | -1 | -1 | 90 |
| 1 | 1 | -1 | -1 | 82 |
| -1 | -1 | 1 | -1 | 68 |
| 1 | -1 | 1 | -1 | 61 |
| -1 | 1 | 1 | -1 | 87 |
| 1 | 1 | 1 | -1 | 80 |

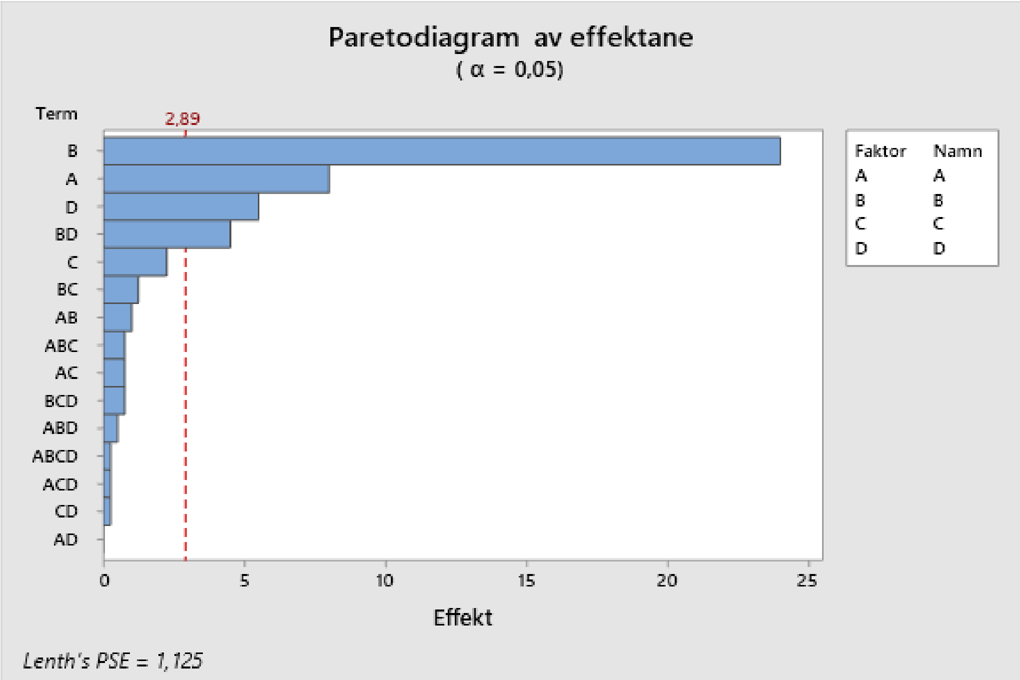
| | | | | |
|---|---|---|---|---|
| -1 | -1 | -1 | 1 | 61 |
| 1 | -1 | -1 | 1 | 50 |
| -1 | 1 | -1 | 1 | 89 |
| 1 | 1 | -1 | 1 | 83 |
| -1 | -1 | 1 | 1 | 59 |
| 1 | -1 | 1 | 1 | 51 |
| -1 | 1 | 1 | 1 | 85 |
| 1 | 1 | 1 | 1 | 78 |

## Coded Coefficients

| Term | Effect | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|---|
| Constant | | 72,25 | * | * | * | |
| A | -8,000 | -4,000 | * | * | * | 1,00 |
| B | 24,00 | 12,00 | * | * | * | 1,00 |
| C | -2,250 | -1,125 | * | * | * | 1,00 |
| D | -5,500 | -2,750 | * | * | * | 1,00 |
| A*B | 1,0000 | 0,5000 | * | * | * | 1,00 |
| A*C | 0,7500 | 0,3750 | * | * | * | 1,00 |
| A*D | -0,000000 | -0,000000 | * | * | * | 1,00 |
| B*C | -1,2500 | -0,6250 | * | * | * | 1,00 |
| B*D | 4,500 | 2,250 | * | * | * | 1,00 |
| C*D | -0,2500 | -0,1250 | * | * | * | 1,00 |
| A*B*C | -0,7500 | -0,3750 | * | * | * | 1,00 |
| A*B*D | 0,5000 | 0,2500 | * | * | * | 1,00 |
| A*C*D | -0,2500 | -0,1250 | * | * | * | 1,00 |
| B*C*D | -0,7500 | -0,3750 | * | * | * | 1,00 |
| A*B*C*D | -0,2500 | -0,1250 | * | * | * | 1,00 |




Figur 3. Normalplott for prosessutviklingsforsøket




Figur 4. Paretodiagram for effektane i prosessutviklingsforsøket

Frå plotta er effektane av A, B, D og BD vurdert til å vere signifikante, men vi veit at Lenth's metode implementert i MINITAB er noko konservativ. Dei estimerte trefaktor- og firefaktorsamspela er små. Dersom dei sanne verdiane for desse er 0, vil estimatorane for desse ha forventing 0 og varians $\sigma^2_{\text{effekt}}$ .

Dermed er ein estimator for $\sigma^2_{\text{effekt}}$ gitt ved: $\dfrac{\widehat{ABC}^2 + \widehat{ABD}^2 + \widehat{ACD}^2 + \widehat{BCD}^2 + \widehat{ABCD}^2}{5}$

For våre data blei estimatet: $s^2_{\text{effekt}} = \dfrac{(-0.75)^2 + (0.5)^2 + (-0.25)^2 + (-0.75)^2 + (-0.25)^2}{5} = 0.3$

og standardavviket til effektane blir estimert til $s_{\text{effekt}} = \sqrt{0.3} = 0.55$.

Ved vurdering av signifikans skal absoluttverdien av dei estimerte effektane samanliknast med: $s_{\text{effekt}} \cdot t_{0.025,5} = 0.55 \cdot 2.571 = 1.41$. Såleis er det rom for å undrast om hovedeffekten av trykk, det vil sei C, og er signifikant.

## Estimering av varians ved gjentak

Estimering av varians krev normalt at ein gjer gjentak av forsøket. Dersom eit gjentak er gjort, har ein to responsverdiar for kvar nivåkombinasjon som begge har same forventing.

I tabell 5 er det gitt data frå eit $2^3$ forsøk med gjentak. Dataane er frå eit forsøk utført for å forbetre metall-legeringa for ein komponent i flymotorar. Her er A = temperatur, B = innhaldet av titan og C = varmebehandlingsmetode. Responsen er lengda av sprekker i metall-legeringa målt i $\text{mm} \times 10^{-2}$. La $y_{11}$ og $y_{12}$ vere dei observerte responsverdiane for nivåkombinasjon 1. Eit estimat for variansen til observasjonane, $\sigma^2$, er då gitt ved:

$$\sum_{j=1}^{2}\left(y_{1j}-\bar{y}_1\right)^2=\left(y_{11}-\frac{y_{11}+y_{12}}{2}\right)^2+\left(y_{12}-\frac{y_{11}+y_{12}}{2}\right)^2$$
$$=\left(\frac{y_{11}-y_{12}}{2}\right)^2+\left(\frac{-y_{11}+y_{12}}{2}\right)^2=\frac{\left(y_{11}-y_{12}\right)^2}{2}$$

Normalt får ein $2^k$ slike estimat som kan brukast til å estimere $\sigma^2$ ved å ta gjennomsnittet av dei.

Tabell 5. $2^3$ forsøk med gjentak

| A | B | C | $y_{i1}$ | $y_{i2}$ | $y_{i1}-y_{i2}$ | $\frac{(y_{i2}-y_{i1})^2}{2}$ |
|---|---|---|---|---|---|---|
| - | - | - | 6.48 | 5.87 | 0.61 | 0.186 |
| + | - | - | 7.88 | 8.24 | -0.36 | 0.065 |
| - | + | - | 9.57 | 9.34 | 0.23 | 0.026 |
| + | + | - | 10.90 | 11.35 | -0.45 | 0.101 |
| - | - | + | 10.71 | 11.12 | -0.41 | 0.084 |
| + | - | + | 12.77 | 12.57 | 0.20 | 0.02 |
| - | + | + | 8.61 | 8.52 | 0.09 | 0.004 |
| + | + | + | 10.30 | 10.06 | 0.24 | 0.029 |
| Totalt | | | | | | 0.515 |

Estimatet for $\sigma^2$ blir då: $s^2=\frac{0.515}{8}=0.064$, og vi får

$$\sigma^2_{effekt}=\frac{4\sigma^2}{n}\Rightarrow s^2_{effekt}=\frac{4s^2}{16}=\frac{4\cdot 0.064}{16}=0.016.$$

Denne estimatoren har 8 fridomsgrader. Hovedeffekten av ein faktor A vil då vere signifikant dersom $\left|\frac{\hat{A}}{s_{effekt}}\right|\geq t_{0.025,8}$. Normalt gjer ein ikkje meir enn eit gjentak, men generelt har ein ved (m-1) gjentak (totalt m verdiar for kvar nivåkombinasjon) at $\sum_{j=1}^{m}\frac{(Y_{ij}-\bar{Y}_i)^2}{m-1}$ er ein estimator for $\sigma^2$ for kvar i. Ved å midle over desse får vi ein estimator for variansen. Denne har $(m-1)2^k$ fridomsgrader.

Dersom ein berre har kvantitative faktorar, kan ein og få estimert variansen ved å gjere forsøk i senterpunktet til forsøket. Gjer ein r slike forsøk vil ein frå variablene $Y_1, Y_2, \dots, Y_r$ få ein estimator for varians , $\frac{\sum_{i=1}^{r}\left(Y_i-\bar{Y}\right)^2}{r-1}$, med r-1 fridomsgrader.

Dersom ein gjer gjentak må ein etter kommandosekvensen:

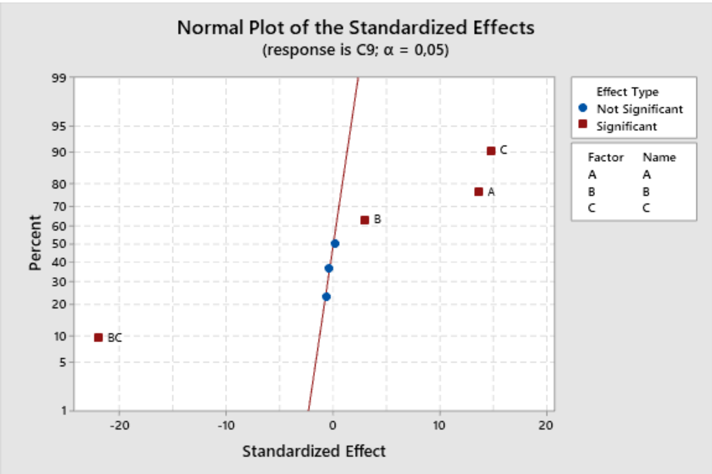
Stat $\rightarrow$DOE $\rightarrow$Factorial $\rightarrow$Create factorial design

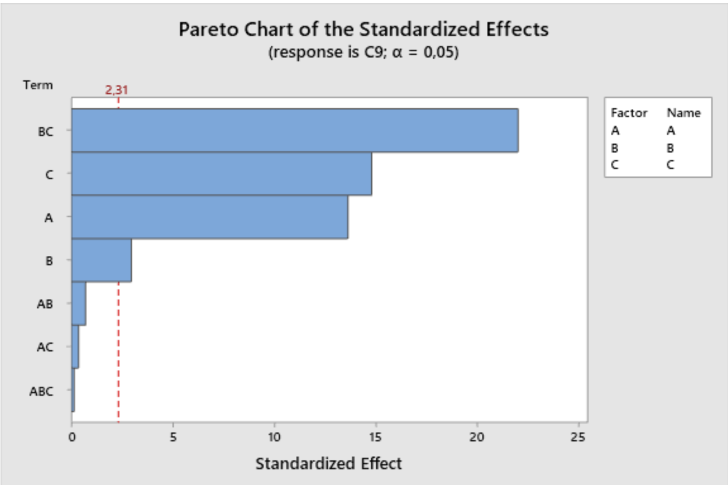
i tillegg til talet på faktorar og full factorial og velge talet på gjentak, her 2. Hak bort «randomize runs». No kan ein lese inn dei 16 dataane i ei kolonne og analysere forsøket.

Sidan det er gjentak får ein både standardavvik til koffisientane (halvparten av standard-avviket til effektane) og p-verdiar. Legg merke til at det er standardiserte effektar (effekt/ standardavvik som no er framstilt i plotta). Både p-verdiane, normalplottet og paretodiagrammet peikar mot at A, B, C og BC er signifikante. Plotta er direkte frå programpakken. Ein kan forandre tekst, både overskrift og namn på aksane ved å dobbeltklikke.

## Coded Coefficients

| Term | Effect | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|---|
| Constant | | 9,6431 | 0,0635 | 151,96 | 0,000 | |
| A | 1,7313 | 0,8656 | 0,0635 | 13,64 | 0,000 | 1,00 |
| B | 0,3763 | 0,1881 | 0,0635 | 2,96 | 0,018 | 1,00 |
| C | 1,8788 | 0,9394 | 0,0635 | 14,80 | 0,000 | 1,00 |
| A*B | -0,0887 | -0,0444 | 0,0635 | -0,70 | 0,504 | 1,00 |
| A*C | -0,0463 | -0,0231 | 0,0635 | -0,36 | 0,725 | 1,00 |
| B*C | -2,7963 | -1,3981 | 0,0635 | -22,03 | 0,000 | 1,00 |
| A*B*C | 0,0188 | 0,0094 | 0,0635 | 0,15 | 0,886 | 1,00 |



# Tolkning av effektar

Dersom ein faktor ikkje har samspel, tolkar vi estimert hovedeffekt som estimert forandring i forventa respons når vi går frå lågt til høgt nivå av faktoren. For faktorar som har samspel med andre faktorar blir tolkninga gjort ved hjelp av sampelsplott. I $2^3$ forsøket for å redusere talet på fjærer med sprekkdanning var estimatet for hovedeffekten $\hat{A}=23$ og $\hat{AC}$ blei estimert til 10. Vi kan no lage oss ein tabell der vi studerer kva som skjer for dei 4 nivåkombinasjonane av A og C. Ei grafisk framstilling av denne tabellen blir kalla eit samspelsplott.

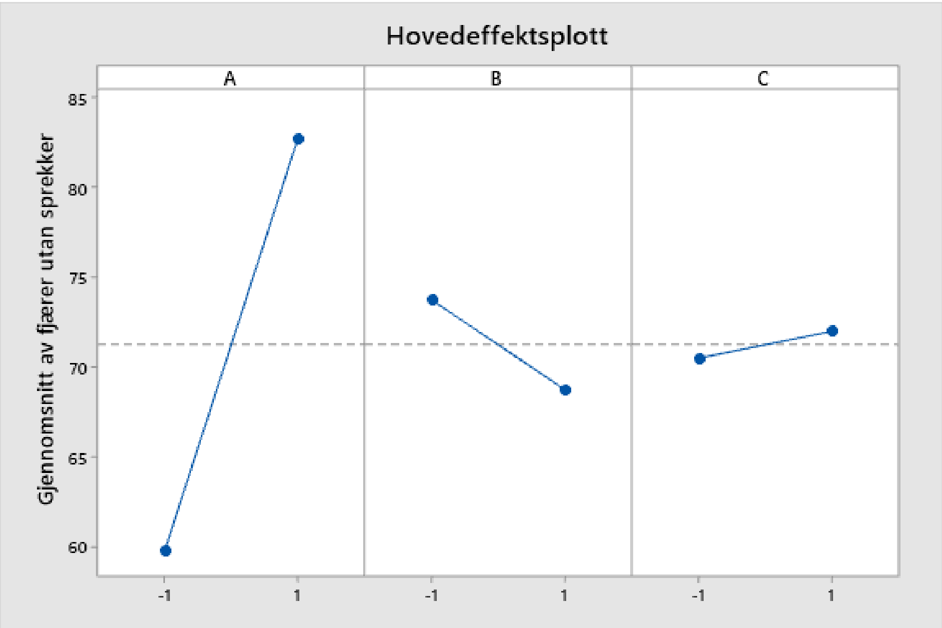
Tabell 6. Tabell som grunnlag for samspelsplott

| | C | |
|---|---|---|
| A | - | + |
| - | $\frac{67+61}{2}=64$ | $\frac{59+52}{2}=55.5$ |
| + | $\frac{79+75}{2}=77$ | $\frac{90+87}{2}=88.5$ |

Ved i kommandolinja å gå til Stat og klikke vidare

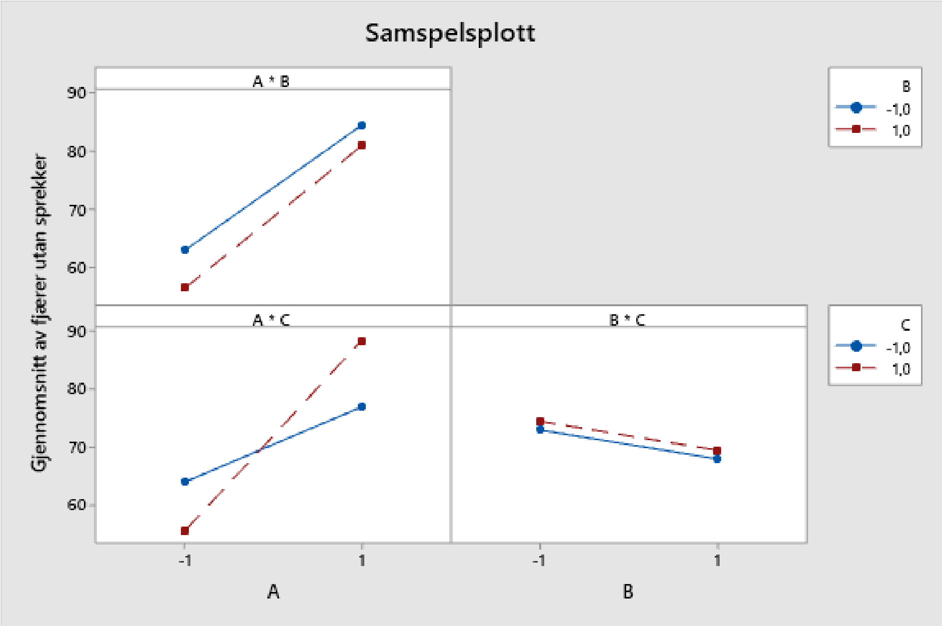
$$Stat \rightarrow DOE \rightarrow Factorial \rightarrow Factorial\ plots,$$

kan ein få laga hovedeffektsplott og samspelsplott. Dersom det ikkje er samspel mellom to faktorar er effekten av ein faktor den same uavhengig av nivået til den andre faktoren. Linjene i tofaktorsamspelsplotta skal då bli parallelle. I figur 6 ser vi at dette er langt på veg tilfelle for faktorane A og B og for B og C, men ikkje for A og C.




Figur 5. Hovedeffektsplott for forsøket med stål-karbon fjærer

Tolkninga av analysen blir då at effekten av å auke temperaturen på kjølevæska er negativ dersom ein har temperaturen av stålet på lågt nivå, men positiv dersom temperaturen av stålet er på høgt nivå. Minst sprekker får ein når begge temperaturane er på høgt nivå.



Figur 6. Samspelsplott for forsøket med stål-karbon fjærer

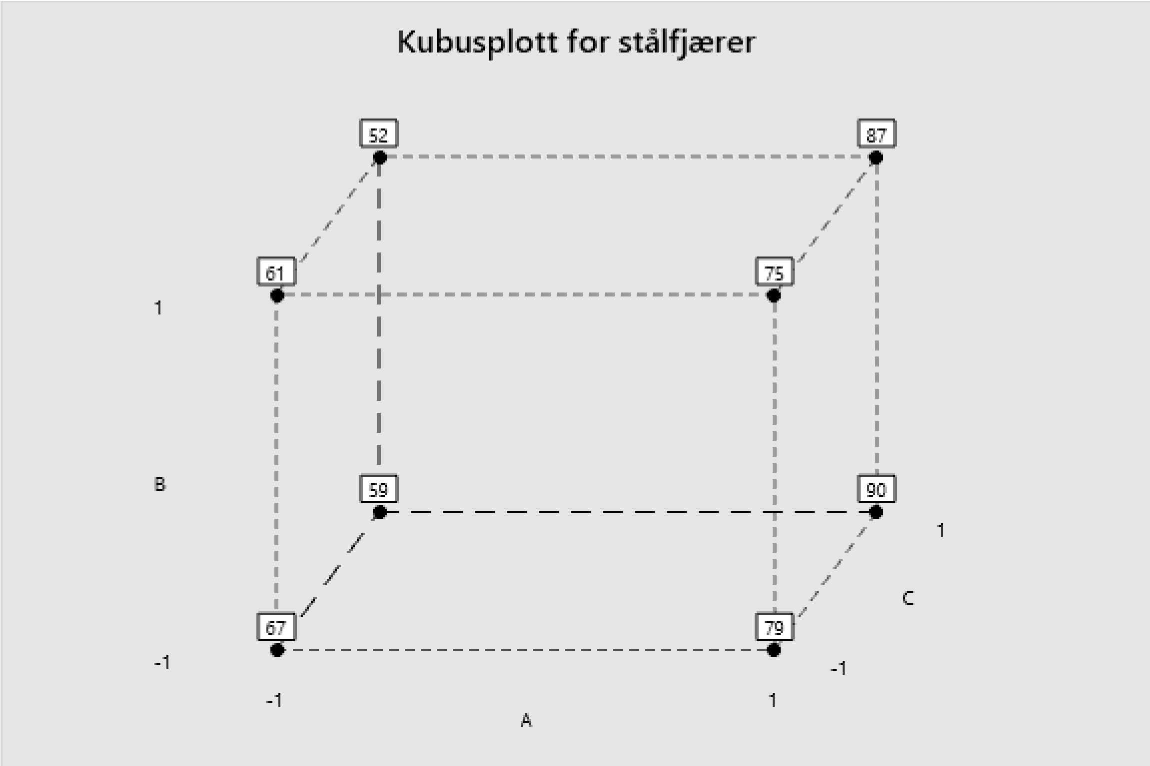
Om vi går tilbake til utskrifta for forsøket med stål-karbon fjærer, finn vi ein kolonne med koeffisientar. Dei fleste tala der er halvparten av effektane. Det skuldast at ein effekt måler kva som skjer med responsen når vi for ein faktor går frå lågt (-1) nivå til høgt (+1) nivå, det vil sei ein avstand på 2, medan ein koeffisient berre måler kva som skjer med responsen når vi forandrar ein faktor med 1, eller går frå 0 til 1. Første talet i kolonna for koeffisientane som svarar til konstantleddet er gjennomsnittet av alle obervasjonane. Det svarar til ein tenkt verdi for responsen der faktorane var innstilt på 0, det vil sei midt mellom lågt og høgt nivå. Vi kan bruke kolonnen for koeffisientane til å setje opp ein tilnærma modell for responsen. Sidan A og AC var dei signifikante effektane vil ein slik modell vere:

$$\hat{y} = 71.3 + 11.5\mathrm{A} + 5\mathrm{AC} \text{ eller } \hat{y} = 71.3 + 11.5x_{\mathrm{A}} + 5\mathrm{x}_{\mathrm{AC}}$$

Med A og C på høgt nivå eller $x_{\mathrm{A}} = x_{\mathrm{AC}} = 1$ er prosentvis estimat for fjærer utan sprekker lik 87.8. Eit kubusplott som vist i figur 7, gjev nyttig informasjon om kva nivåkombinasjonar som gjev best resultat. Dette får ein med kommandosekvensen:

$$Stat \rightarrow DOE \rightarrow Factorial \rightarrow Cube\ Plot$$

Verdiane ein får ved å rekne ut $e_i = y_i - \hat{y}_i$, $i = 1,2,\ldots,8$ blir kalla residual. Ved bruk av normalplott kan ein sjekke om det er rimeleg å anta at desse og dermed enkeltobservasjonane kjem frå ei normalfordeling. Eit plott av $e_i$ mot $\hat{y}_i$, er nyttig for å sjekke om det er rimeleg at alle observasjonane har lik varians. Når det gjeld normalfordeling, hjelper det analysen at estimatorane for effektane er basert på gjennomsnitt og har lik varians uansett. Det viktigaste ein får ut av desse plotta, er nok at ein kan oppdage avvikande observasjonar.




Figur 7. Kubusplott for dataane i forsøket med stål-karbon fjærer

## Blokkdeling i $2^k$ forsøk

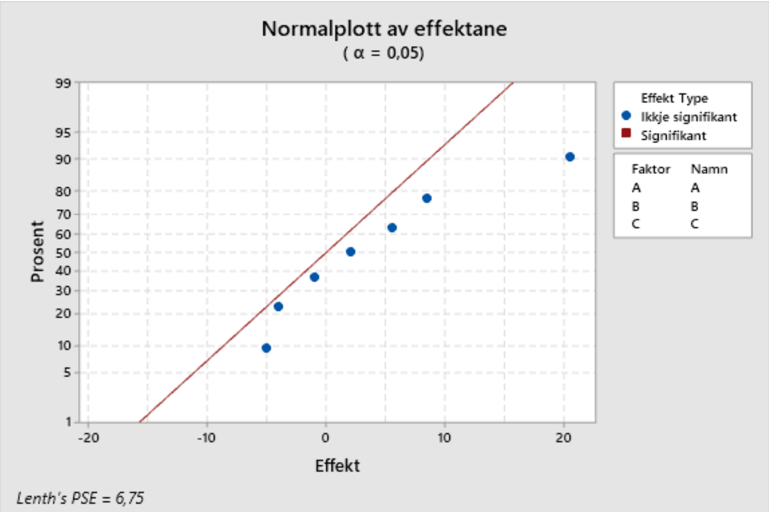
Dersom vi skal gjere mange forsøk vil det ofte vere slik at ytre forhold forandrar seg frå vi startar forsøket til vi er ferdige. Forandringar i ytre forhold kan påverke responsverdiane og gjere at vi feilestimerer effektane. Vi kan motverke dette ved å blokkdele forsøket. Stundom er det andre føringar, mangel på råmateriale t. d. som gjer at vi ynskjer å blokkdele eit forsøk. Når vi blokkdeler forsøket, skal randomiseringa skje i kvar blokk. Gå ut i frå at $2^3$ forsøket i eksempel 1 blei gjort på nytt og at ei randomisering i MINITAB gav følgjande rekkjefølgje dei 8 forsøka skulle utførast i. Det viste seg at det berre var mogeleg å gjere 4 forsøk på ein dag, slik at forsøka med nummer 5-8 måtte gjerast dagen etter. Forandringar i ytre forhold gjorde at responsen i desse forsøka fekk eit tillegg på 5.

| Forsøksnr. | A | B | C | Respons |
|---|---|---|---|---|
| 1 | + | - | - | 79 |
| 2 | + | + | + | 87 |
| 3 | + | + | - | 75 |
| 4 | - | - | + | 67 |
| 5 | - | + | + | 57 |
| 6 | - | + | - | 66 |
| 7 | + | - | + | 95 |
| 8 | - | - | + | 64 |

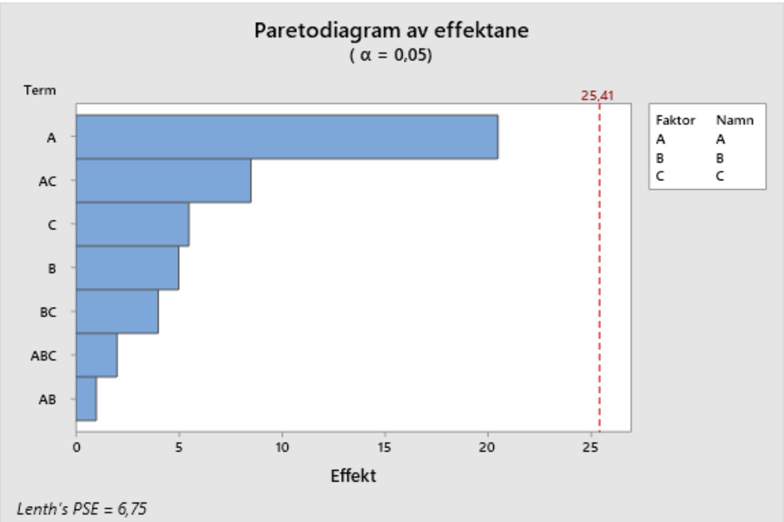
Ei utskrift med MINITAB gav følgjande estimerte effektar, med normalplott og paretodiagram. Lenth's metode viser no ingen signifikante effektar. Effekten av A er rett nok mykje større enn dei andre, og frå normalplottet vil det nok vere skjellig grunn til mistanke om at den kan påverke responsen.

## Coded Coefficients

| Term | Effect | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|---|
| Constant | | 73,75 | * | * | * | |
| A | 20,50 | 10,25 | * | * | * | 1,00 |
| B | -5,000 | -2,500 | * | * | * | 1,00 |
| C | 5,500 | 2,750 | * | * | * | 1,00 |
| A*B | -1,0000 | -0,5000 | * | * | * | 1,00 |
| A*C | 8,500 | 4,250 | * | * | * | 1,00 |
| B*C | -4,000 | -2,000 | * | * | * | 1,00 |
| A*B*C | 2,000 | 1,000 | * | * | * | 1,00 |




Figur 7. Normalplott for ei uheldig gjennomføring av forsøket med stål-karbon fjærer.




Figur 8. Paretodiagram for ei uheldig gjennomføring av forsøket med stål-karbon fjærer.

Det har no oppstått ein blokkeffekt i forsøket mellom dei to dagane. Denne påverkar estimeringa ev effektane og analysen etterpå. Dette hadde vore mogeleg å unngå om ein på førehand hadde planlagt å utføre forsøket over to dagar. I tabell 7 er det synt ei blokkdeling der ein har samla forsøka der trefaktorsamspelet ABC er på lågt nivå i ei blokk, Blokk 1, og dei andre i Blokk 2.

Tabell 7. Eit $2^3$ forsøk i 2 blokker kvar på 4 forsøk

| Forsøk | A | B | C | AB | AC | BC | | ABC |
|---|---|---|---|---|---|---|---|---|
| 1 | - | - | - | + | + | + | Blokk 1 | - |
| 4 | + | + | - | + | - | - | | - |
| 6 | + | - | + | - | + | - | | - |
| 7 | - | + | + | - | - | + | | - |
| | | | | | | | | |
| 2 | + | - | - | - | - | + | Blokk 2 | + |
| 3 | - | + | - | - | + | - | | + |
| 5 | - | - | + | + | - | - | | + |
| 8 | + | + | + | + | + | + | | + |

Vi observerer at om alle enkeltforsøka i blokk 2 får eit tillegg h, så vil utrekning av hovedeffektane og tofaktorsamspela vere upåverka fordi det er like mange – som + i kvar blokk. Såleis ville ein forvente at estimatet av alle desse effektane er uforandra og at ein endar opp med same konklusjon som om ein kunne utføre forsøket på ein dag. Dette gjeld ikkje for trefaktorsamspelet som blir konfundert (blanda saman) med blokkeffekten. Generelt kan ein velge ein vilkårleg av dei 7 kolonnane å blokkdele etter, men sidan effekten som er tilordna den kolonnen ein blokkdeler etter blir konfundert med ein potensiell blokkeffekt, har ein som hovedregel at om ein skal dele forsøket opp i to blokker, vel ein kolonnen for samspelet av høgaste orden.

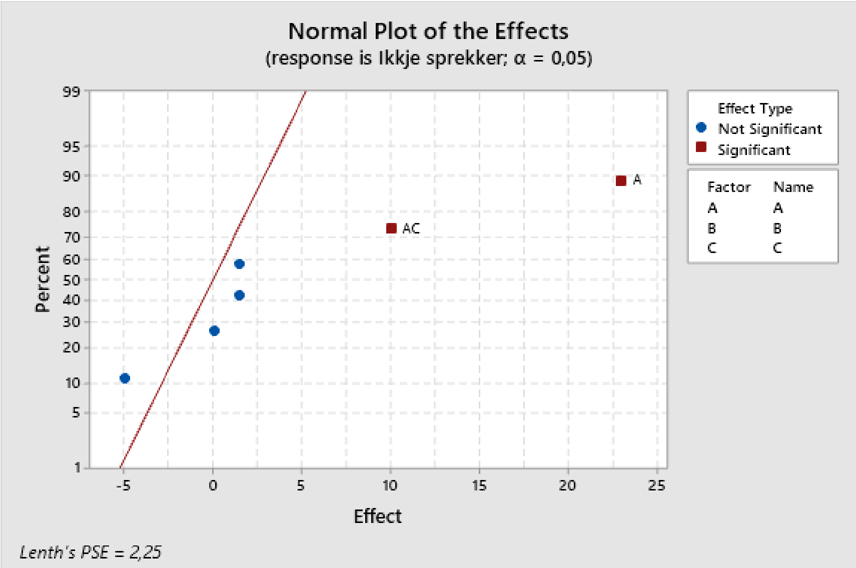
Etter kommandosekvensen:

$$Stat \rightarrow DOE \rightarrow Factorial \rightarrow Analyze\ factorial\ design$$

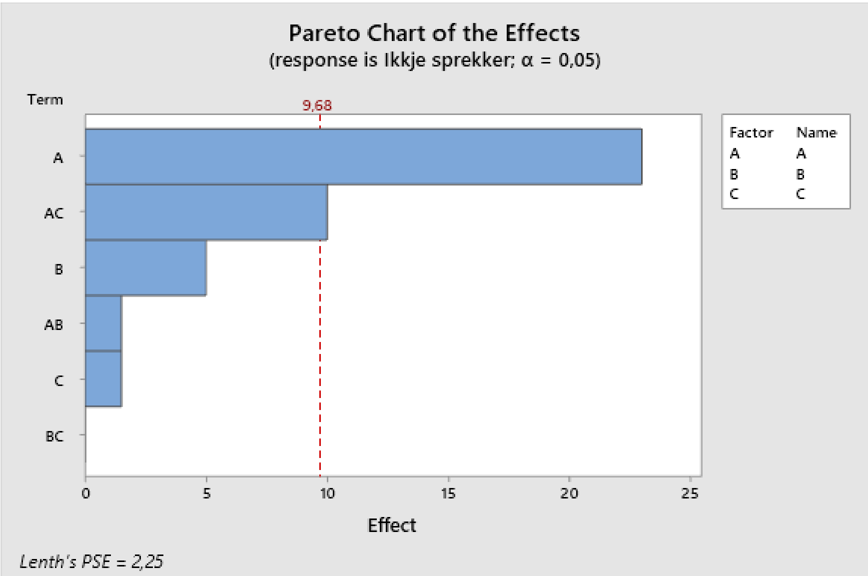
kan ein i tillegg til talet på faktorar og «Full factorial» og velge talet på blokker, her 2. Då får ein nivåkombinasjonane fordelt i to blokker som i vist i tabell 7. Om ein les inn respons til tilhøyrande nivåkombinasjon, får ein denne analysen:

## Coded Coefficients

| Term | Effect | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|---|
| Constant | | 71,25 | * | * | * | |
| Blocks | | | | | | |
| 1 | | -0,2500 | * | * | * | 1,00 |
| A | 23,00 | 11,50 | * | * | * | 1,00 |
| B | -5,000 | -2,500 | * | * | * | 1,00 |
| C | 1,5000 | 0,7500 | * | * | * | 1,00 |
| A*B | 1,5000 | 0,7500 | * | * | * | 1,00 |
| A*C | 10,000 | 5,000 | * | * | * | 1,00 |
| B*C | -0,000000 | -0,000000 | * | * | * | 1,00 |




Figur 9. Normalplott for det blokkdelte $2^3$ forsøket med stål-karbon fjærer.




Figur 10. Paretodiagram for det blokkdelte $2^3$ forsøket med stål-karbon fjærer.

Effekten av trefaktorsamspelet er ikkje med, men elles er estimerte effektar like og plotta nesten like. Ved analyse av eit blokkdelt forsøk, er det tryggast å fjerne effekten for den kolonnen evt. dei kolonnane som er konfunderte med blokkeffekten før ein lagar normalplott evt. paretodiagram. Særleg analysen frå paretodiagrammet kan bli veldig sårbar med blokkeffektar involvert sidan dei kan påverke medianen.

Ein kan dele eit forsøk inn i meir enn to blokker. Dersom ein skal dele det inn i 4 blokker må ein bruke to effektkolonnar, og helst slike som gjer at ein får estimert dei effektane ein trur påverkar responsen upåverka av blokkdelinga. For å dele inn i 4 blokker blokkdeler ein etter dei 4 mogelege nivåkombinasjonane ein har i to kolonnar. Desse er synt nedanfor

| Blokk 1 | Blokk 2 | Blokk 3 | Blokk 4 |
|---|---|---|---|
| (-  -) | (-  +) | (+  -) | (+  +) |

Dersom ein vil dele inn eit $2^4$ forsøk i 4 blokker, vil det beste vere å bruke to trefaktorsamspel, til dømes ACD og BCD som blokkgeneratorar. La $B_1$ = ABC og $B_2$ = ABD. Desse to trefaktorsamspela blir då konfunderte med blokkeffekten. Det same blir samspelet mellom dei $B_1B_2 = ABCABD = A^2B^2CD = CD$ . Legg merke til at her har vi brukt enkel algebra. Gangar vi saman forteikna i to like kolonner rekkjevis får vi ein kolonne med berre +. Vi skriv t. d. :

$$I = AA = BB = CC$$

der I er ein kolonne med berre + teikn. Ein slik kolonne påverkar ikkje rekkjevis multiplikasjon av forteikna og kan neglisjerast. I tabell 8 er det synt eit $2^4$ forsøk i 4 blokker.

Tabell 8. Eit $2^4$ forsøk i 4 blokker kvar på 4 forsøk

| A | B | C | D | AB | AC | AD | BC | BD | ACD | BCD | ABCD | | ABC | ABD | CD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| - | - | - | - | + | + | + | + | + | - | - | + | | - | - | + |
| + | + | - | - | + | - | - | - | - | + | + | + | Blokk 1 | - | - | + |
| + | - | + | + | - | + | + | - | - | + | - | - | | - | - | + |
| - | + | + | + | - | - | - | + | + | - | + | - | | - | - | + |
| | | | | | | | | | | | | | | | |
| + | - | + | - | - | + | - | - | + | - | + | + | | - | + | - |
| - | + | + | - | - | - | + | + | - | + | - | + | Blokk 2 | - | + | - |
| - | - | - | + | + | + | - | + | - | + | + | - | | - | + | - |
| + | + | - | + | + | - | + | - | + | - | - | - | | - | + | - |
| | | | | | | | | | | | | | | | |
| - | - | + | - | + | - | + | - | + | + | + | - | | + | - | - |
| + | + | + | - | + | + | - | + | - | - | - | - | Blokk 3 | + | - | - |
| + | - | - | + | - | - | + | + | - | - | + | + | | + | - | - |
| - | + | - | + | - | + | - | - | + | + | - | + | | + | - | - |
| | | | | | | | | | | | | | | | |
| + | - | - | - | - | - | - | + | + | + | - | - | | + | + | + |
| - | + | - | - | - | + | + | - | - | - | + | - | Blokk 4 | + | + | + |
| - | - | + | + | + | - | - | - | - | - | - | + | | + | + | + |
| + | + | + | + | + | + | + | + | + | + | + | + | | + | + | + |

# Fraksjonar av $2^k$ forsøk

Høgare ordens samspel ( >2 ) er mykje sjeldnare signifikante enn hovedeffektar og tofaktorsamspel. Då kan det vere nærliggande å tenkje at effektkolonnane for slike kan brukast til å inkludere fleire enn k faktorar i eit $2^k$ forsøk. Nedanfor er det synt eit forsøk med 4 faktorar i 8 forsøk. Ein fjerde faktor D er tilordna kolonnen for trefaktorsamspelet ABC. Ein seier då at D=ABC er ein *generator* for designet. Sidan D og ABC har same forteikn, får ein I=ABCD. Dette blir kalla *definerande relasjon* for forsøket. Sidan eit fullfaktorielt forsøk i 4 faktorar krev 16 forsøk, vil forsøket i 4 faktorar i 8 forsøk bli kalla ein halvfraksjon av eit $2^4$ forsøk. Ein skriv eit $2^{4-1}$ forsøk. I tabell 9 er det synt ein halvfraksjon av eit $2^4$ forsøk der dei 8 forsøka i eksempel 2 med ABCD = + er plukka ut.

Tabell 9. Ein halvfraksjon av eit $2^{4-1}$ forsøk

| A | B | C | D=ABC | Respons |
|---|---|---|---|---|
| - | - | - | - | 71 |
| + | - | - | + | 50 |
| - | + | - | + | 89 |
| + | + | - | - | 82 |
| - | - | + | + | 59 |
| + | - | + | - | 61 |
| - | + | + | - | 87 |
| + | + | + | + | 78 |

Denne halvfraksjonen kan i MINITAB analyserast ved å gå til Stat og klikke vidare:

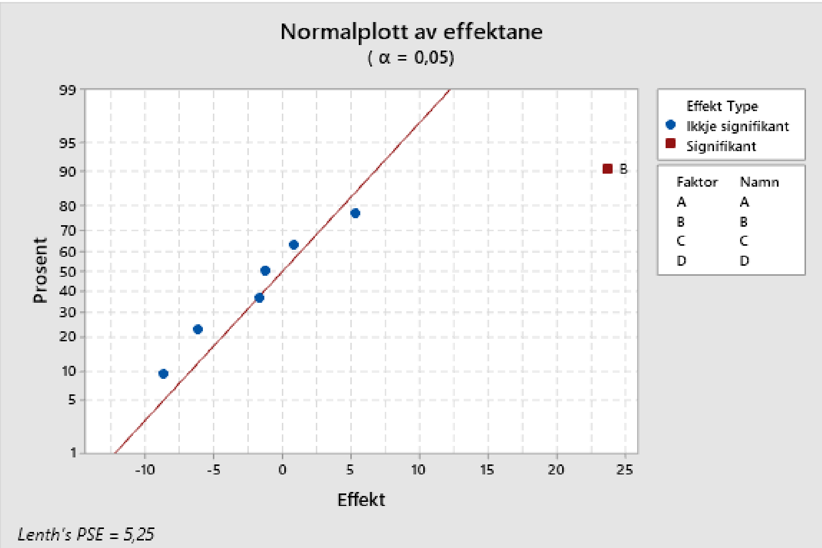
Stat $\rightarrow$ DOE $\rightarrow$ Factorial $\rightarrow$ Create factorial design

Velg 4 faktorar, «1/2 fraction» under «Designs» og fjern «randomize runs». Hak av for normalplott i

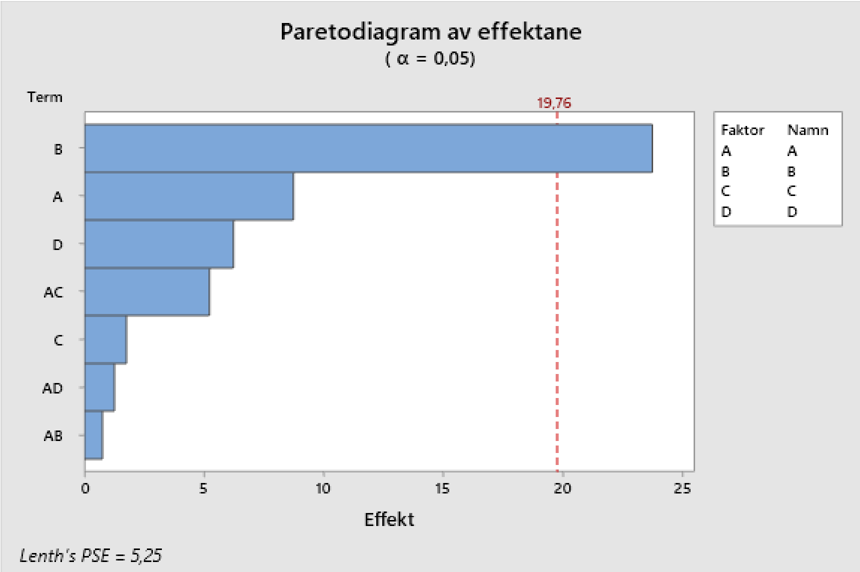
Stat $\rightarrow$ DOE $\rightarrow$ Factorial $\rightarrow$ Analyze factorial design og vi får:

## Coded Coefficients

| Term | Effect | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|---|
| Constant | | 72,13 | * | * | * | |
| A | -8,750 | -4,375 | * | * | * | 1,00 |
| B | 23,75 | 11,87 | * | * | * | 1,00 |
| C | -1,7500 | -0,8750 | * | * | * | 1,00 |
| D | -6,250 | -3,125 | * | * | * | 1,00 |
| A*B | 0,7500 | 0,3750 | * | * | * | 1,00 |
| A*C | 5,250 | 2,625 | * | * | * | 1,00 |
| A*D | -1,2500 | -0,6250 | * | * | * | 1,00 |




Figur 11. Normalplott av effektane i halvfraksjonen av $2^4$ forsøket for prosessutvikling



Figur 12. Paretodiagram for halvfraksjonen av $2^4$ forsøket for prosessutvikling

Tilynelatande er berre B signifikant, og sjølv om MINITAB er konservativ er det vanskeleg å konkludere med at A, D og AC og kanskje C og er signifikante som vi fekk i det fulle faktorielle forsøket. Vi ser at ein grunn til det er at PSE no er over 4 gonger så stor.

Dessutan må vi vere klar over at i fraksjonelle forsøk vil det alltid vere effektar vi ikkje kan skilje frå einannan. Slike effektar blir kalla aliasar. To slike aliasar kan begge vere signifikante men kansellere kvarandre. Dei kan og addere seg og lage ein signifikant effekt. Legg og merke til at i utskrifta frå MINITAB er det berre 3 tofaktorsamspel, medan det for fire faktorar er mogeleg å lage 6. Grunnen til det er at 2 og 2 tofaktorsamspel er aliasar. I denne fraksjonen er I=ABCD. Det gjev at A=BCD eller, sidan dei er fullstendige konfunderte, at $\hat{A} = \hat{A} + BC\hat{D}$. Tilsvarande er $\hat{B} = \hat{B} + AC\hat{D}$, $\hat{C} = \hat{C} + AB\hat{D}$ og $\hat{D} = \hat{D} + AB\hat{C}$. For tofaktorsamspela får vi AB=CD eller $A\hat{B} = A\hat{B} + C\hat{D}$, $A\hat{C} = A\hat{C} + B\hat{D}$ og $A\hat{D} = A\hat{D} + B\hat{C}$. Ein grei måte å finne ut kva effektar som er aliasar er å multiplisere ein effekt med den definerande relasjonen. Til dømes er:

$$ADI=ADABCD= A^2D^2BC = BC.$$

Dette eksempelet viser noko av problemet med Lenth's metode og prosentvis relativt mange signifikante effektar. Ein kan risikerer at PSE blir basert på signifikante effektar og dermed blir for stor. Sjølv om ein tek bort B og medianen blir basert på dei estimerte AC (BD) og C, er det frå analysen av det fullfaktorielle forsøket grunn til å tru at desse og er signifikante. Men er ein usikker på analysen av det fraksjonerte forsøket, kan ein utføre den andre halvfraksjonen og og utføre analysen på 16 forsøk istadenfor 8. Dei to fraksjonane blir no utført i to blokker med randomisering innan desse. Det einaste ein taper er estimeringa av firefaktorsamspelet.

Eit veldig godt forsøk er halvfraksjonen av eit $2^5$ forsøk, eller eit $2^{5-1}$ forsøk med generator E=ABCD.

Tabell 10. Ein halvfraksjon av eit $2^5$ forsøk

| A | B | C | D | E=ABCD |
|---|---|---|---|---|
| - | - | - | - | + |
| + | - | - | - | - |
| - | + | - | - | - |
| + | + | - | - | + |
| - | - | + | - | - |
| + | - | + | - | + |
| - | + | + | - | + |
| + | + | + | - | - |
| - | - | - | + | - |
| + | - | - | + | + |
| - | + | - | + | + |
| + | + | - | + | - |
| - | - | + | + | + |
| + | - | + | + | - |
| - | + | + | + | - |
| + | + | + | + | + |

Definerande relasjon blir I=ABCDE, og dermed er hovedeffektar aliasar med firefaktorsamspel og tofaktorsamspel aliasar med trefaktorsamspel. Dersom trefaktor- og firefaktorsamspela er neglisjerbare, får ein estimert 5 hovedeffektar og alle dei 10 tofaktorsamspela upåverka av andre effektar med dette designet.

Kva er det som gjer $2^{5-1}$ forsøket så bra? Grunnen er at dei viktigaste effektane berre er aliasar med høgare ordens samspel som normalt kan neglisjerast. Den definerande relasjonen fortel oss dette. For $2^{4-1}$ forsøket er I=ABCD og vi får problem med tofaktorsamspela som blir aliasar to og to. For $2^{5-1}$ forsøket er ingen tofaktorsamspel aliasar.

**Definisjon resolusjon**

*Resolusjonen, R, av eit forsøk er definert som lengda av det kortaste ordet i den definerande relasjonen.*

Dette er eit mål på kor godt forsøket er til å skilje lågare ordens samspel frå høgare ordens samspel og dess høgare resolusjon eit forsøk har, dess betre. For $2^{5-1}$ forsøket er R = 5, for $2^{4-1}$ forsøket er R=4 og dersom ein lagar seg eit $2^{3-1}$ forsøk med generator C=AB og I=ABC, er R=3. I eit R = 3 forsøk, er hovedeffektane aliasar med tofaktorsamspel.

Alt i alt er det alltid rom for å putte $2^k$ -1 faktorar inn i eit $2^k$ forsøk, og dersom ein kan stole på at ingen samspel har betydning, får ein estimert alle dei $2^k$ -1 hovedeffektane. I tabell 11 er det synt eit $2^{6-2}$ forsøk, eller ein kvartfraksjon av eit $2^6$ forsøk, der ein har putta 6 faktorar inn i 16 forsøk. Generatorane er E=ABC og F=ABD. Dermed blir I = ABCE = ABDF. Men I = $I^2$ = ABCEABDF = CDEF. Dermed blir den definerande relasjonen I = ABCE = ABDF = CDEF.

Tabell 11. Ein kvartfraksjon av eit $2^6$ forsøk

| A | B | C | D | E=ABC | F=ABD |
|---|---|---|---|---|---|
| - | - | - | - | - | - |
| + | - | - | - | + | + |
| - | + | - | - | + | + |
| + | + | - | - | - | - |
| - | - | + | - | + | - |
| + | - | + | - | - | + |
| - | + | + | - | - | + |
| + | + | + | - | + | - |
| - | - | - | + | - | + |
| + | - | - | + | + | - |
| - | + | - | + | + | - |
| + | + | - | + | - | + |
| - | - | + | + | + | + |
| + | - | + | + | - | - |
| - | + | + | + | - | - |
| + | + | + | + | + | + |

# Avsluttande kommentarar og litteratur for meir lesing

To-nivå forsøk kan brukast til å forbetre nesten alt, herunder og funksjonaliteten til maskinlæringsalgoritmer. Eit eksempel kan vere å finne ut kva hyperparametrar som har mest å seie og å finne gode innstillingar for desse. Men to-nivå forsøk kan ikkje brukast til å optimalisere. For å optimalisere ein prosess treng du å finne ein approksimativ modell for ei responsflate. Ein slik modell er ofte eit andre ordens polynom og du treng forsøksoppsett som kan estimere slike modellar. Dei kallast responsflate forsøk. Det mest kjende er CCD (Central Composite Designs) eller på norsk *Sentralt samansett forsøk* som legg til enkeltforsøk i senterverdiane til faktorane og i tillegg to på kvar akse. Avstanden frå senterverdiane er ofte $-\sqrt{k}$ og $\sqrt{k}$ der $k$ er talet på faktorar. Men i ei verd som forandrar seg dynamisk og som fokuserer på kontinuerleg forbedring, kan du få til mykje med berre to-nivå forsøk. Dei som ynskjer å lære meir om dette interessanne og utfordrande temaet kan gå til dei klassiske bøkene:

Box, G. E. P., Hunter, J. S. & Hunter, W. G.: Statistics for Experimenters (1978, 2005)
Montgomery, D. C.: Design and Analysis of Experiments (2007, 7th edition)
Wu, C. F. & Hamada, M.S.: Experiments, Planning, Analysis, and Optimization (2009, 2 th edition) .
Box, G. E. P. & Draper, N. R.: Response Surfaces, Mixtures, and Ridge Analyses (2007).
Myers, R. H. & Montgomery, D. C.: Response Surface Methodology (2001 eller seinare).