

Forsøksplaner, oppsett og datainnsamling, fra et statistisk ståsted



NIBIO

NORSK INSTITUTT FOR
BIOØKONOMI

12. mars 2019 9.00 – 10.30

Skypemøte 1 i NLR's kurs i forsøksarbeid 2019

Torfinn Torp

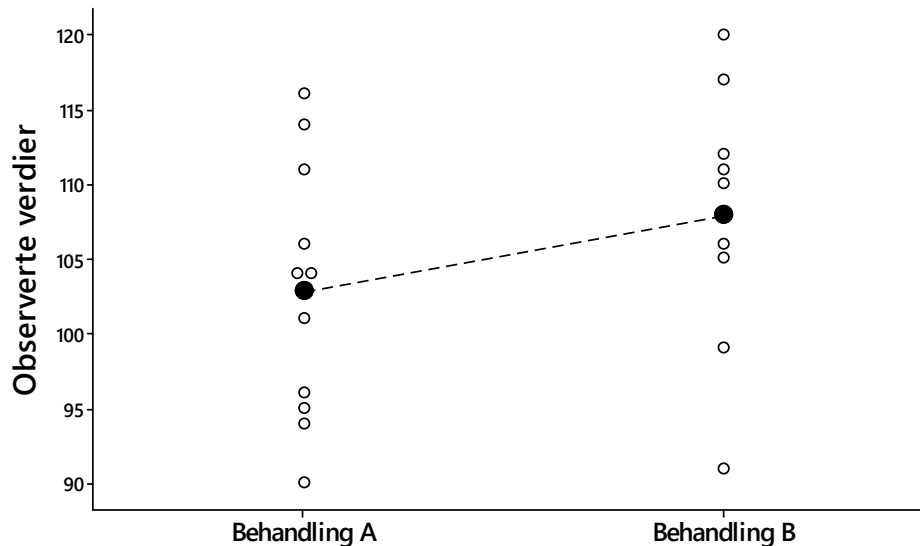
Temaer

- Sammenlignende forsøk.
- Hvorfor bruke blokker og randomisere, et eksempel.
- Noen grunnleggende begreper.
- Randomisert fullstendig blokkforsøk, et eksempel.
- Forsøksserier, over flere år og/eller flere steder.
- Kontrollbehandling i forsøk.
- Hvor mange gjentak trenger vi?
- Noen forsøksplaner.
- Noen videre begreper.
- Randomisert fullstendig blokkforsøk, et eksempel på forsøksplan og dataanalyse.
- Avslutning.

Sammenlignende forsøk

- Sammenligning av behandlinger, det være seg sorter, gjødslingsmengder, jordarbeidingsmetoder, sprøytekonsentrasjoner, dyrefor,
- Sammenligner variasjon/forskjell, i verdiene til en responsvariabel, mellom behandlingene med variasjon innen behandlingene.

Sammenligning av to behandlinger



Gjennomsnitt, behandling A = 102.8

Gjennomsnitt, behandling B = 107.9

Tester om forventet verdi i behandlingene er like:

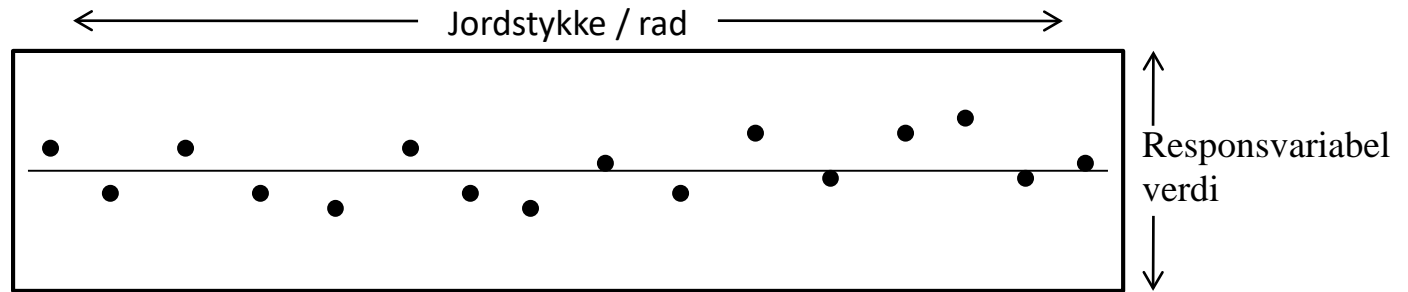
$$F = MS_{Beh} / MS_E = 131.44 / 72.03 = 1.82$$

$$P = 0.193$$

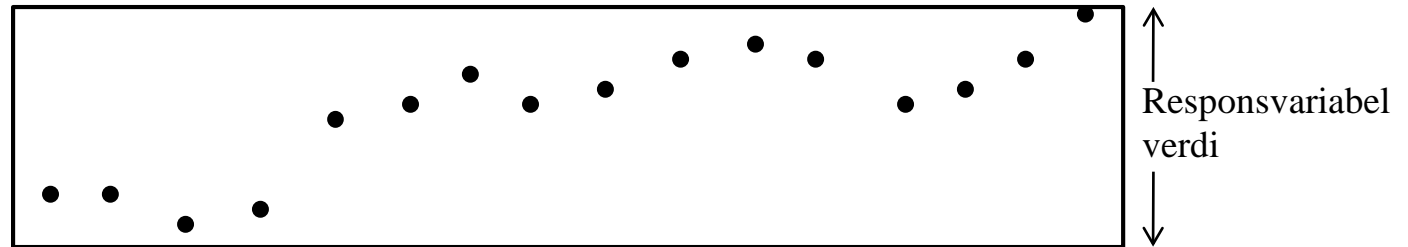
Eksempel

Fire hvetesorter, A, B, C og D skal sammenlignes (for eksempel hvor mye avling per enhet kan vi forvente at de gir). Hver sort skal brukes fire ganger. To situasjoner og to forslag til forsøksplaner:

En stasjonær
forstyrrelse/støy/
bakgrunnsstøy (S)



En ikke-stasjonær
forstyrrelse/støy/
bakgrunnsstøy (IS)



1. Et systematisk
forsøk



2. Et randomisert
blokk forsøk



Hvilken forsøksplan foretrekker du?

Noen grunnleggende begreper, I

Forsøksenhet (forsøksrute)

Den minste enheten som kan tildeles en enkelt behandling.

Forsøksfeil

Variasjon, i responsverdiene, mellom forsøksenheter gitt samme behandling og som i utgangspunktet er ment å være homogene.

Ønskes liten.

Gjentak

Gjentakelse av det grunnleggende forsøket, under samme forsøksbetingelser.

Gjør at forsøksfeilen kan estimeres.

Sammenligning av behandlingene blir mer nøyaktig.

Pseudogjentak / paralleller

Gjentak som skaffes ved å måle på enheter som er del av samme forsøksenhet.

Noen grunnleggende begreper, II

Blokker

En mengde homogene forsøksenheter, homogene før behandlingene brukes.

Bidrar til å redusere forsøksfeilen.

Sammenligning av behandlingene gjøres innen blokkene, det vil si under relativt like forhold.

Randomisering

Å bestemme tilfeldig hvilken behandling som skal gis de enkelte forsøksenheterne.

Bidrar til at den vanlige antagelsen om uavhengige feilledd etc. er oppfylt.

Bidrar til å nøytralisere innflytelsen av irrelevante og subjektive faktorer.

Eksempel, randomisert fullstendig blokkforsøk, I

Fire hvetesorter, A, B, C og D skal sammenlignes. 16 forsøksruter er tilgjengelig.

Forsøksplan og observerte verdier

Blokk 1:	B	42	A	40	D	36	C	38
Blokk 2:	A	49	D	42	C	47	B	53
Blokk 3:	D	34	B	39	A	39	C	34
Blokk 4:	A	46	C	42	B	48	D	37

Eksempel, randomisert fullstendig blokkforsøk, II

Kilde til variasjon	DF	SS	MS	F	P
Hvetesort:	3	158,25	52,75	23,44	0,000
Blokk:	3	293,25	97,75	43,44	0,000
Feil / Residual:	9	20,25	2,25		
Total:	15	471,75			

Ovenfor er sammenligningen av hvetesortene gjort innen blokkene.

Stor variasjon mellom blokkene og liten forsøksfeil.

Dersom vi ikke hadde tatt hensyn til variasjonen mellom blokkene og utelatt blokk fra modellen ville resultatet blitt (stor forsøksfeil):

Kilde til variasjon	DF	SS	MS	F	P
Hvetesort:	3	158,25	52,75	2,02	0,165
Feil / Residual:	12	313,50	26,125		
Total:	15	471,75			

Førsøksserier, over år og/eller sted

Norge har variable værforhold, jordforhold etc.

Resultater fra et feltforsøk gjort i et enkelt år på et sted kan sjelden generaliseres til framtidige år og til andre steder, som nettopp er situasjonen når resultatene skal brukes til rådgiving for kommende år på andre steder.

Dessverre trekkes for ofte bastante og generelle konklusjoner og generaliseringer, basert på et alt for tynt grunnlag.

Litt spissformulert: Hvis en ikke har ressurser til å utføre et forsøk skikkelig (randomisering, gjentak/blokker, flere år, flere steder etc.) er det kanskje bedre å la det være.

I planlagte flerårige forsøk kan det noen ganger være fristende å forandre forsøksplanen for etterfølgende år basert på for eksempel første års resultater. Ikke gjør det, et flerårig forsøk utføres jo nettopp fordi vi er klar over at resultatene kan variere fra år til år!

Men, det kan selvsagt planlegges med forandring, for eksempel fra år til år ved at randomisering av behandlingene på forsøksrutene innen (nye) blokker gjøres på nytt hvert år.

Kontrollbehandling (ingen behandling, 0-ledd, kontroll-ledd etc.)

Om en kontrollbehandling skal tas med i forsøket bør avgjøres av hensikten med forsøket.

Observasjoner fra kontrollbehandlingen kan ofte skille seg fra de andre behandlingene, både deres gjennomsnitt og deres variasjon.

Om kontrollbehandlingen skiller seg veldig fra de andre behandlingene kan det overskygge eventuelle forskjeller mellom de andre, mer interessante og reelle behandlingene.

Dersom kontrollbehandlingen skiller seg veldig fra de andre behandlingene bør det være tillatt å konkludere med at kontrollbehandlingen skiller seg fra de andre behandlingene uten å bruke en formell statistisk test. Deretter utelates kontrollbehandlingen fra den formelle statistiske modellen og analysen som brukes for å sammenligne de andre behandlingene.

Hvor mange gjentak trenger vi, en grov indikasjon, I?

Vi skal sammenligne k behandlinger i et *fullstendig randomisert blokk forsøk* med b blokker.

Problemet er å indikere nødvendig antall blokker, b .

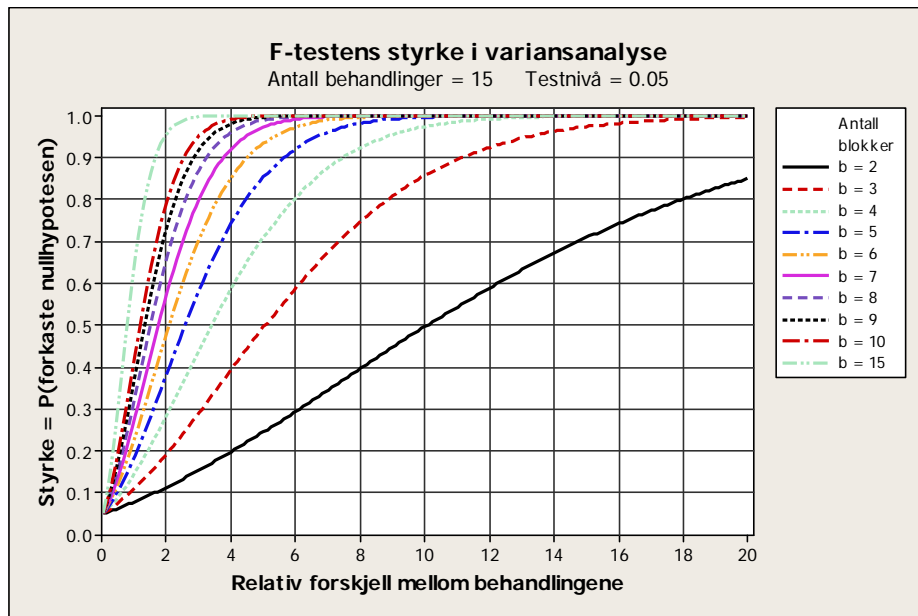
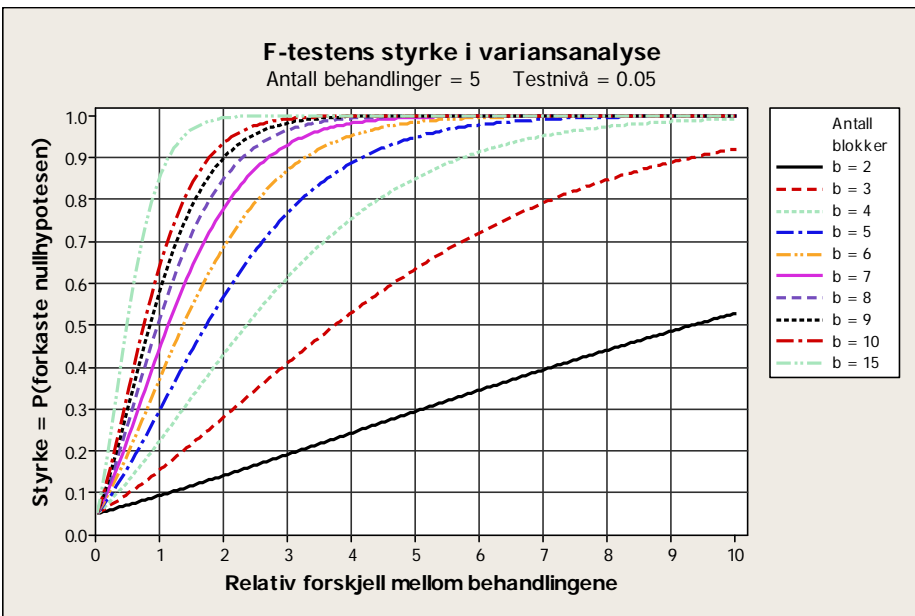
Vi må ta stilling til hva vi ønsker å oppnå. I dette tilfellet betyr det å anslå hvor stor sannsynlighet vi ønsker for å avsløre en gitt forskjell mellom behandlingene (sammenlignet med variasjonen/forskjellene innen behandlingene).

Grov skisse til framgangsmåte:

- 1 Gitt testnivå α .
- 2 Spesifiser (maksimal forskjell)/variasjon i forventet responsvariabelverdi mellom behandlingene, Δ .
- 3 Spesifiser standard avviket til responsvariabelverdiene, variasjonen innen behandlingene, σ .
- 4 Beregne relativ forskjell mellom behandlingene, $\Phi = \Delta / \sigma$.
- 5 Beregne b slik at sannsynligheten for å avsløre situasjonen gitt via Φ er tilstrekkelig stor.

Hvor mange gjentak trenger vi, en grov indikasjon, II?

To eksempler



Noen forsøksplaner, I - Randomisert fullstendig blokkforsøk.

Med en faktor, Sort

Blokk 1

Blokk 2

Blokk 3

Rute 101 D	Rute 201 G	Rute 301 B
Rute 102 A	Rute 202 D	Rute 302 E
Rute 103 C	Rute 203 A	Rute 303 F
Rute 104 G	Rute 204 B	Rute 304 D
Rute 105 F	Rute 205 E	Rute 305 C
Rute 106 B	Rute 206 C	Rute 306 A
Rute 107 E	Rute 207 F	Rute 307 G

Med to faktorer, Sort og Nitrogenmengde

Blokk 1

Blokk 2

Blokk 3

Rute 101 A N2	Rute 201 B N1	Rute 301 A N1
Rute 102 A N1	Rute 202 D N1	Rute 302 C N1
Rute 103 D N2	Rute 203 C N2	Rute 303 A N2
Rute 104 C N2	Rute 204 C N1	Rute 304 C N2
Rute 105 D N1	Rute 205 A N1	Rute 305 B N1
Rute 106 B N2	Rute 206 B N2	Rute 306 D N2
Rute 107 B N1	Rute 207 A N2	Rute 307 B N2
Rute 108 C N1	Rute 208 D N2	Rute 308 D N1

Ofte vern/grenseruter rundt hver blokk. Pseudogjentak?

Noen forsøksplaner, II - Split plot forsøk ("to faktorer")

	Jordarbeiding 3	Jordarbeiding 1	Jordarbeiding 4	Jordarbeiding 2	
<i>Småruter 1</i>	Gjødsling 3	Gjødsling 4	Gjødsling 5	Gjødsling 5	<i>Blokk 1</i>
<i>Småruter 2</i>	Gjødsling 1	Gjødsling 1	Gjødsling 3	Gjødsling 4	
<i>Småruter 3</i>	Gjødsling 2	Gjødsling 3	Gjødsling 1	Gjødsling 2	
<i>Småruter 4</i>	Gjødsling 5	Gjødsling 2	Gjødsling 2	Gjødsling 1	
<i>Småruter 5</i>	Gjødsling 4	Gjødsling 5	Gjødsling 4	Gjødsling 3	
	<i>Storrute 1</i>	<i>Storrute 2</i>	<i>Storrute 3</i>	<i>Storrute 4</i>	
	Jordarbeiding 4	Jordarbeiding 3	Jordarbeiding 1	Jordarbeiding 2	
<i>Småruter 1</i>	Gjødsling 3	Gjødsling 4	Gjødsling 3	Gjødsling 2	<i>Blokk 2</i>
<i>Småruter 2</i>	Gjødsling 2	Gjødsling 5	Gjødsling 5	Gjødsling 1	
<i>Småruter 3</i>	Gjødsling 1	Gjødsling 2	Gjødsling 1	Gjødsling 4	
<i>Småruter 4</i>	Gjødsling 4	Gjødsling 1	Gjødsling 2	Gjødsling 5	
<i>Småruter 5</i>	Gjødsling 5	Gjødsling 3	Gjødsling 4	Gjødsling 3	
	<i>Storrute 1</i>	<i>Storrute 2</i>	<i>Storrute 3</i>	<i>Storrute 4</i>	
	Jordarbeiding 2	Jordarbeiding 4	Jordarbeiding 3	Jordarbeiding 1	
<i>Småruter 1</i>	Gjødsling 1	Gjødsling 3	Gjødsling 5	Gjødsling 4	<i>Blokk 3</i>
<i>Småruter 2</i>	Gjødsling 5	Gjødsling 2	Gjødsling 2	Gjødsling 5	
<i>Småruter 3</i>	Gjødsling 2	Gjødsling 1	Gjødsling 1	Gjødsling 2	
<i>Småruter 4</i>	Gjødsling 4	Gjødsling 4	Gjødsling 3	Gjødsling 3	
<i>Småruter 5</i>	Gjødsling 3	Gjødsling 5	Gjødsling 4	Gjødsling 1	
	<i>Storrute 1</i>	<i>Storrute 2</i>	<i>Storrute 3</i>	<i>Storrute 4</i>	

Noen forsøksplaner, III - Split split plot forsøk ("tre faktorer")

Blokk 1

Sort 2			Sort 3			Sort 4			Sort 1		
Sort 2 K2 N2	Sort 2 K3 N3	Sort 2 K1 N1	Sort 3 K1 N2	Sort 3 K3 N2	Sort 3 K2 N2	Sort 4 K3 N1	Sort 4 K2 N2	Sort 4 K1 N3	Sort 1 K3 N3	Sort 1 K1 N3	Sort 1 K2 N1
Sort 2 K2 N1	Sort 2 K3 N1	Sort 2 K1 N3	Sort 3 K1 N1	Sort 3 K3 N3	Sort 3 K2 N1	Sort 4 K3 N3	Sort 4 K2 N3	Sort 4 K1 N2	Sort 1 K3 N1	Sort 1 K1 N2	Sort 1 K2 N2
Sort 2 K2 N3	Sort 2 K3 N2	Sort 2 K1 N2	Sort 3 K1 N3	Sort 3 K3 N1	Sort 3 K2 N3	Sort 4 K3 N2	Sort 4 K2 N1	Sort 4 K1 N1	Sort 1 K3 N2	Sort 1 K1 N1	Sort 1 K2 N3

Blokk 2

Sort 3			Sort 4			Sort 1			Sort 2		
Sort 3 K3 N3	Sort 3 K2 N1	Sort 3 K1 N1	Sort 4 K2 N3	Sort 4 K1 N2	Sort 4 K3 N2	Sort 1 K1 N2	Sort 1 K3 N2	Sort 1 K2 N3	Sort 2 K1 N1	Sort 2 K2 N3	Sort 2 K3 N2
Sort 3 K3 N1	Sort 3 K2 N3	Sort 3 K1 N2	Sort 4 K2 N2	Sort 4 K1 N3	Sort 4 K3 N3	Sort 1 K1 N1	Sort 1 K3 N3	Sort 1 K2 N1	Sort 2 K1 N3	Sort 2 K2 N1	Sort 2 K3 N3
Sort 3 K3 N2	Sort 3 K2 N2	Sort 3 K1 N3	Sort 4 K2 N1	Sort 4 K1 N1	Sort 4 K3 N1	Sort 1 K1 N3	Sort 1 K3 N1	Sort 1 K2 N2	Sort 2 K1 N2	Sort 2 K2 N2	Sort 2 K3 N1

Sort med 4 nivåer på storruter, kalium med 3 nivåer på småruter og nitrogen med 3 nivåer på små småruter.

Noen forsøksplaner, IV - Split blokk forsøk = Strip plot forsøk ("to faktorer")

3 grasarter (faktor 1)
2 klippehøyder (faktor 2)

Grasartene og klippehøydene
kombineres til en "radfaktor",
med nivåene 1, 2, 3, ..., 6.

Blokk 1

Radfaktor	Kolonnefaktor								
	5	8	1	9	3	6	2	7	4
3	3 5	3 8	3 1	3 9	3 3	3 6	3 2	3 7	3 4
4	4 5	4 8	4 1	4 9	4 3	4 6	4 2	4 7	4 4
2	2 5	2 8	2 1	2 9	2 3	2 6	2 2	2 7	2 4
1	1 5	1 8	1 1	1 9	1 3	1 6	1 2	1 7	1 4
5	5 5	5 8	5 1	5 9	5 3	5 6	5 2	5 7	5 4
6	6 5	6 8	6 1	6 9	6 3	6 6	6 2	6 7	6 4

3 mengder N-gjødsling (faktor 3)
3 mengder P-gjødsling (faktor 4)

N-gjødslingene og P-gjødslingene
kombineres til en "kolonnefaktor",
med nivåene 1, 2, 3, ..., 9.

Blokk 2

Radfaktor	Kolonnefaktor								
	2	7	4	9	6	3	8	5	1
5	5 2	5 7	5 4	5 9	5 6	5 3	5 8	5 5	5 1
2	2 2	2 7	2 4	2 9	2 6	2 3	2 8	2 5	2 1
3	3 2	3 7	3 4	3 9	3 6	3 3	3 8	3 5	3 1
6	6 2	6 7	6 4	6 9	6 6	6 3	6 8	6 5	6 1
1	1 2	1 7	1 4	1 9	1 6	1 3	1 8	1 5	1 1
4	4 2	4 7	4 4	4 9	4 6	4 3	4 8	4 5	4 1

Noen forsøksplaner, V - Latinske kvadrat forsøk ("en faktor")

Blokkdeling i to "retninger".

Eksempel 1

4 dyrefor A, B, C og D sammenlignes.

4 kyr og 4 tidsperioder.

Ku (blokk)	Tidsperiode (blokk)			
	1	2	3	4
1	C	B	A	D
2	B	D	C	A
3	A	C	D	B
4	D	A	B	C

Eksempel 2

4 sorter A, B, C og D sammenlignes.

4 blokker i nord – sør retning og
4 blokker i vest – øst retning.

Vest ↓ Øst blokk	Nord ↔ Sør blokk			
	1	2	3	4
1	D	A	C	B
2	A	D	B	C
3	B	C	D	A
4	C	B	A	D

Noen forsøksplaner, VI - Ufullstendige blokkforsøk

Forsøk der blokkene ikke kan gjøres store nok til at alle behandlingene kan brukes i hver blokk.

Det finnes flere typer ufullstendige blokkforsøk. De skiller seg på betingelser for hvordan behandlingene randomiseres på rutene innen blokkene (oftest ved hjelp av egnet programvare).

Noen typer ufullstendige blokkforsøk:

- Balanserte ufullstendige blokkforsøk.
- Delvis balanserte ufullstendige blokkforsøk.
- Youden kvadrat forsøk.
- Sykliske forsøk.
- Lattice forsøk.
- Alfa forsøk.
- ⋮

Eksempel

Balansert ufullstendig blokkforsøk (alle par av behandlingene opptrer i samme antall blokker, her opptrer alle parene AB, AC, AD, BC, BD og CD i 2 blokker).

Med en faktor, Sort

Blokk 1	Blokk 2	Blokk 3	Blokk 4
Rute 101 D	Rute 201 C	Rute 301 B	Rute 401 A
Rute 102 C	Rute 202 A	Rute 302 D	Rute 402 B
Rute 103 A	Rute 203 B	Rute 303 C	Rute 403 D

Noen videre begreper, I

Faste og tilfeldige faktorer

En faktor er fast hvis dens nivåer er de nivåene vi er interessert i.

Eksempel: sorter.

En faktor er tilfeldig hvis dens nivåer ikke har noen spesiell interesse og kan betraktes som et tilfeldig utvalg av mulige nivåer.

Eksempel: gårder.

Kryssede og nøstede faktorer

To faktorer, A og B, er kryssede hvis hvert nivå av A forekommer sammen med hvert nivå av B.

Eksempel: Alle kombinasjonene av sortene og nitrogenmengdene brukes i alle blokker.

En faktor B er nøstet innen en faktor A hvis hvert nivå av B forekommer sammen med bare et av nivåene til A.

Eksempel: A = Gjødsling med nivåene 1 og 2, B = Felt, 1, 2, 3 og 4. Gjødsling 1 brukes på feltene 1 og 2, gjødsling 2 brukes på feltene 3 og 4. På hvert felt tas n prøver.

Noen videre begreper, II

Balanserte, ubalanserte og ufullstendige forsøk

Et forsøk er balansert hvis det er like mange observasjoner for hver behandling.

Et forsøk er ubalansert hvis det ikke er like mange observasjoner for hver behandling.

Et forsøk er ufullstendig hvis det mangler observasjoner for noen behandlinger.

Eksempel (6 behandlinger $A_1B_1, A_1B_2, \dots, A_2B_3$, X angir en observasjon/forsøksenhet) :

	A_1	A_2	A_1	A_2	A_1	A_2
B_1	XXX	XXX	XXX	XX	XX	XXX
B_2	XXX	XXX	XXX	XXX		XX
B_3	XXX	XXX	XX	X	XXX	XX
	Balansert		Ubalansert		Ufullstendig	

Sammenblanding av effekter

Sammenblanding av effekter betyr at effektene av faktorer er blandet sammen slik at det ikke er mulig å bestemme hvilken av faktorene som har effekt.

Eksempel: To behandlinger skal sammenlignes på et laboratorium. Prøver fra den ene behandlingen analyseres på dag 1, prøver fra den andre behandlingen analyseres på dag 2. På dag 1 og dag 2 er det forskjellige laboranter, kanskje også forskjellig temperatur, luftfuktighet etc. Faktorene behandling og dag er blandet sammen.

Eksempel, I

Det er gjort et randomisert fullstendig blokkforsøk for å studere hvordan nitrogeninnhold i jorda påvirkes av fire forskjellige gjødslingsmiddel. To uker etter tilføring av gjødslingsmidlene ble det tatt prøver fra forsøksrutene der blant annet nitrogeninnholdet ble målt. Forsøksplan og målt nitrogeninnhold framgår nedenfor.

Blokk 1	Blokk 2	Blokk 3
Rute 101 Beh 2: 10.70	Rute 201 Beh 3: 9.35	Rute 301 Beh 4: 7.55
Rute 102 Beh 4: 10.95	Rute 202 Beh 1: 4.25	Rute 302 Beh 2: 9.15
Rute 103 Beh 1: 3.85	Rute 203 Beh 4: 11.80	Rute 303 Beh 1: 5.60
Rute 104 Beh 3: 8.60	Rute 204 Beh 2: 11.85	Rute 304 Beh 3: 8.40

	Gjennomsnitt	Standardavvik
Behandling 1:	4.57	0.92
Behandling 2:	10.57	1.36
Behandling 3:	8.78	0.50
Behandling 4:	10.10	2.25
Blokk 1:	8.53	3.29
Blokk 2:	9.31	3.57
Blokk 3:	7.68	1.53

Eksempel, II

En vanlig variansanalyse gir resultatene:

Source	DF	SS	MS	F-Value	P-Value
Behandling	3	67.147	22.382	12.66	0.005
Blokk	2	5.365	2.683	1.52	0.293
Error	6	10.605	1.767		
Total	11	83.117			

Konklusjonen er at det er forskjellig forventet nitrogeninnhold i jorda med de fire gjødslingsmidlene fordi forskjellene/variasjonene mellom gjødslingsmidlene er påfallende stor i forhold til forskjellene/variasjonene innen gjødslingsmidlene ($F = 12.66$, $P = 0.005$).

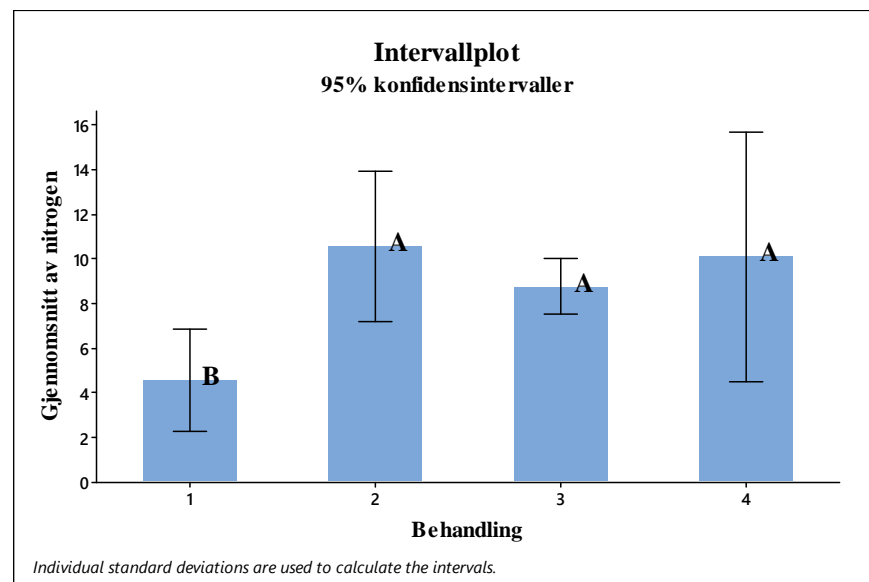
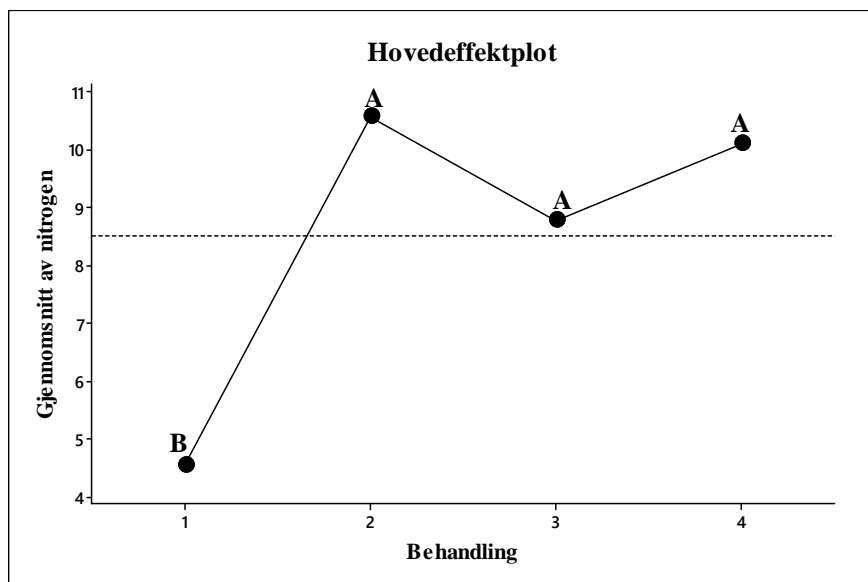
Hvilke gjødslingsmidler er forskjellige?

Eksempel, III

Svaret på hvilke gjødslingsmidler som er forskjellige kan gis ved å bruke en multipl sammenligningsmetode, for eksempel Tukey's parvise sammenligningsmetode. Resultat:

<u>Behandling</u>	<u>N</u>	<u>Mean</u>	<u>Grouping</u>
2	3	10.57	A
4	3	10.10	A
3	3	8.78	A
1	3	4.57	B

Means that do not share a letter are significantly different (95 % confidence).



Avslutning

Spesifiser problemstillingene/hypotesene før forsøket og datainnsamlingen starter.

Bruk enklest mulig forsøksplan i henhold til problemstillingene.

Utfør forsøket grundig og nøyaktig (i planlegging, i forhåndsbearbeiding, i uttak, registrering og bearbeiding av prøver,).

Ikke forandre forsøksplanen underveis basert på midlertidige resultater i forsøksperioden.

Forsøk gjennomført på en skikkelig måte (randomisering, gjentak/blokker, etc.) øker muligheten for å avdekke årsak-virknings forhold.

Å kunne generalisere resultater fra feltforsøk krever (nesten) alltid flerårige forsøk på flere steder.

Hver gang du planlegger et forsøk eller en innsamling av data, husk at dataene skal modelleres og analyseres statistisk den dagen de foreligger.