

Statistisk analyse av data fra planlagte forsøk



NIBIO

NORSK INSTITUTT FOR
BIOØKONOMI

19. mars 2019 9.00 – 10.30

Skypemøte 2 i NLR's kurs i forsøksarbeid 2019

Torfinn Torp

Temaer

- Noen sentrale begreper, framgangsmåte etc., via et eksempel.
- Noen retningslinjer ved planlegging, gjennomføring og analyse av forsøk.
- Randomisert fullstendig blokkforsøk, statistisk modell og analyse.
- Split plot forsøk, statistisk modell og analyse.
- Split split plot forsøk, statistisk modell.
- Split blokk forsøk (= strip plot forsøk), statistisk modell.
- Nøstede, faste og tilfeldige faktorer, manglende observasjoner, statistisk modell og analyse.
- Forsøksserier, over år og/eller sted.
- Programvare.
- Avslutning.

Noen sentrale begreper, framgangsmåte etc., via et eksempel, I

Dataene kommer fra et fullstendig randomisert forsøk (altså uten blokker) der maisavlingen er registrert på 24 homogene forsøksruter som har fått forskjellige gjødslingsmidler. Hvert av 4 gjødslingsmidler (= behandlingene) er fordelt tilfeldig på 6 forsøksruter.

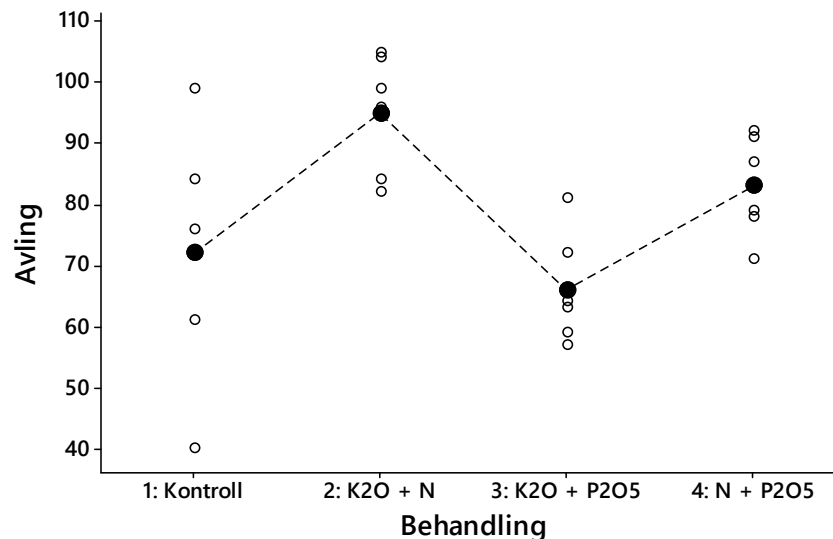
<u>Behandling</u>	<u>Forsøksrute</u>	<u>Avling</u>
1: Kontroll	12	99
1: Kontroll	4	40
1: Kontroll	2	61
1: Kontroll	6	72
1: Kontroll	10	76
1: Kontroll	14	84
2: K ₂ O + N	23	96
2: K ₂ O + N	19	84
2: K ₂ O + N	22	82
2: K ₂ O + N	24	104
2: K ₂ O + N	18	99
2: K ₂ O + N	11	105

<u>Behandling</u>	<u>Forsøksrute</u>	<u>Avling</u>
3: K ₂ O + P ₂ O ₅	21	63
3: K ₂ O + P ₂ O ₅	1	57
3: K ₂ O + P ₂ O ₅	13	81
3: K ₂ O + P ₂ O ₅	7	59
3: K ₂ O + P ₂ O ₅	16	64
3: K ₂ O + P ₂ O ₅	5	72
4: N + P ₂ O ₅	15	79
4: N + P ₂ O ₅	9	92
4: N + P ₂ O ₅	17	91
4: N + P ₂ O ₅	8	87
4: N + P ₂ O ₅	3	78
4: N + P ₂ O ₅	20	71

Noen sentrale begreper, framgangsmåte etc., via et eksempel, II

Behandling	Gjennomsnitt	Varians	Standardavvik	CV %
1: Kontroll	72.00	406.80	20.17	28.01
2: K ₂ O + N	95.00	97.60	9.88	10.40
3: K ₂ O + P ₂ O ₅	66.00	80.80	8.99	13.62
4: N + P ₂ O ₅	83.00	69.20	8.32	10.02

Totalt	79.00	163.60	12.79	16.19
--------	-------	--------	-------	-------



Sammenligner forskjellene/variasjonen mellom behandlingene med variasjonen innen behandlingene.

Noen sentrale begreper, framgangsmåte etc., via et eksempel, III

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Behandling	3	2940	980.0	5.99	0.004
Error	20	3272	163.6		
Total	23	6212			

F er et mål på forskjellene/variasjonen mellom behandlingene sammenlignet med variasjonen innen behandlingene, når vi tar hensyn til antall behandlinger og antall observasjoner både for hver behandling og totalt.

Følgende relasjon for observasjonenes totale variasjon gjelder (k behandlinger, $N = n_1 + n_2 + \dots + n_k$):

$$\underbrace{\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y})^2}_{SS_T} = \underbrace{\sum_{i=1}^k n_i (\bar{Y}_i - \bar{Y})^2}_{SS_B} + \underbrace{\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2}_{SS_E}$$

Total variasjon
 $N-1$ frihetsgrader

Variasjon mellom behandlingene
 $k-1$ frihetsgrader

Variasjon innen behandlingene
 $N-k$ frihetsgrader

Noen sentrale begreper, framgangsmåte etc., via et eksempel, IV

Antagelser:

- Modell for observasjonene: $Y_{ij} = \mu_i + \varepsilon_{ij}$.
- Alle ε_{ij} – ene er uavhengige.
- Alle ε_{ij} – ene er normalfordelte.
- Alle ε_{ij} – ene har forventet verdi 0.
- Alle ε_{ij} – ene har samme varians, σ^2 .
- $\mu_1, \mu_2, \dots, \mu_k$ og σ^2 er ukjente parametere hvis verdier estimeres ved hjelp av dataene.

Fra antagelsene og problemstillingen:

Alle observasjonene med behandling i har forventet verdi μ_i .

Vi tester hypotesen $H_0: \mu_1 = \mu_2 = \dots = \mu_k$ mot H_1 : ikke alle μ_i – ene er like.

Hypotesen testes ved å forkaste H_0 dersom F er påfallende stor. For å avgjøre om det er tilfelle brukes P -verdien knyttet til F .

Noen sentrale begreper, framgangsmåte etc., via et eksempel, V

Generelt om testing av hypoteser

Vår konklusjon basert på data	Sannheten (dessverre ukjent)	
	H_0 er sann	H_1 er sann
Forkast H_0 , det vil si påstå H_1	Type I feil	Riktig beslutning
Ikke grunnlag for å forkaste H_0	Riktig beslutning	Type II feil

Bestemmer et signifikansnivå (= testnivå) α (ofte velges $\alpha = 0.05 = 5\%$).

Krever $P(\text{Type I feil}) \leq \alpha$.

Desto mindre α desto mindre er $P(\text{Type I feil})$.

P -verdien er den minste α som gir forkastning av H_0 på grunnlag av dataene.

Noen sentrale begreper, framgangsmåte etc., via et eksempel, VI

Våre maisdata gir P -verdi = 0.004.

Går derfor videre med en multipl sammenligningsmetode:

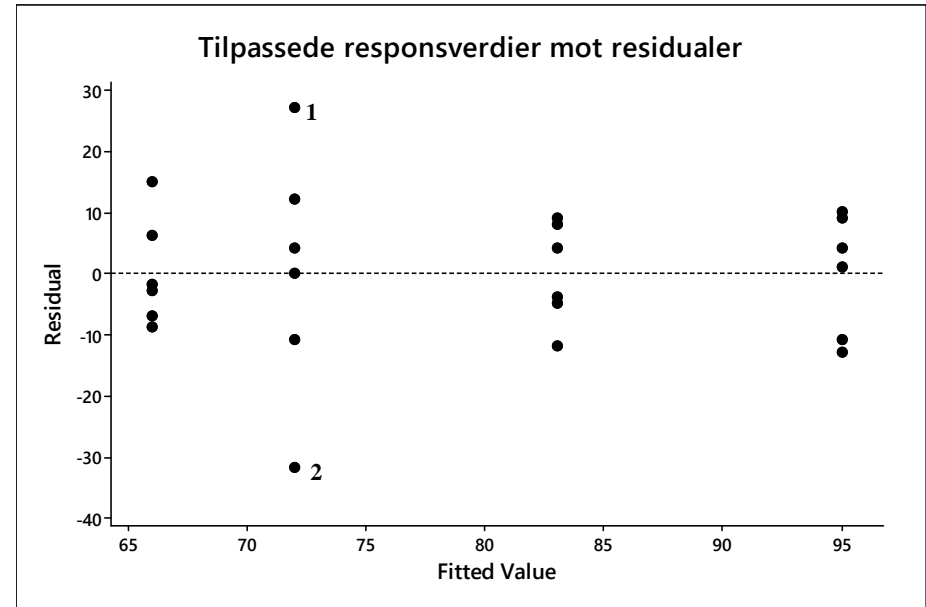
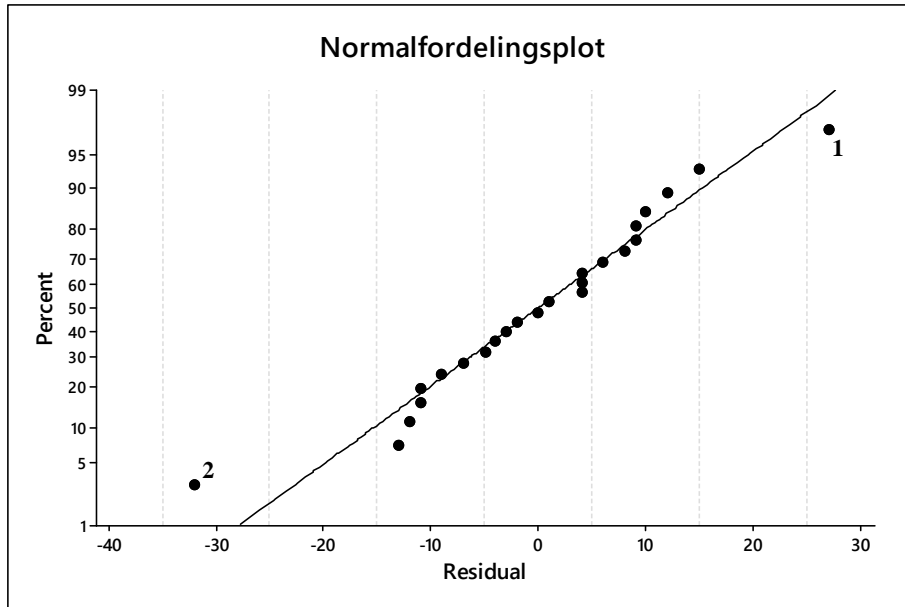
Gruppering ved bruk av Tukey's metode			
Behandling	N	Mean	Grouping
2: K20 + N	6	95.00	A
4: N + P205	6	83.00	A B
1: Kontroll	6	72.00	B
3: K20 + P205	6	66.00	B

Gruppering ved bruk av LSD metoden			
Behandling	N	Mean	Grouping
2: K20 + N	6	95.00	A
4: N + P205	6	83.00	A B
1: Kontroll	6	72.00	B C
3: K20 + P205	6	66.00	C

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

Noen sentrale begreper, framgangsmåte etc., via et eksempel, VII

Sjekking av antagelsene:



Kontrollbehandlingen skiller seg noe ut, spesielt ved større variasjon i responsverdiene og med de to observasjonene, 1 og 2 i datamaterialet, markert i figurene ovenfor.

Hva dersom kontrollbehandlingen utelates?

Noen sentrale begreper, framgangsmåte etc., via et eksempel, VIII

Resultater fra data uten kontrollbehandlingen:

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Behandling	2	2548	1274.00	15.44	0.000
Error	15	1238	82.53		
Total	17	3786			

Gruppering ved bruk av Tukey's metode

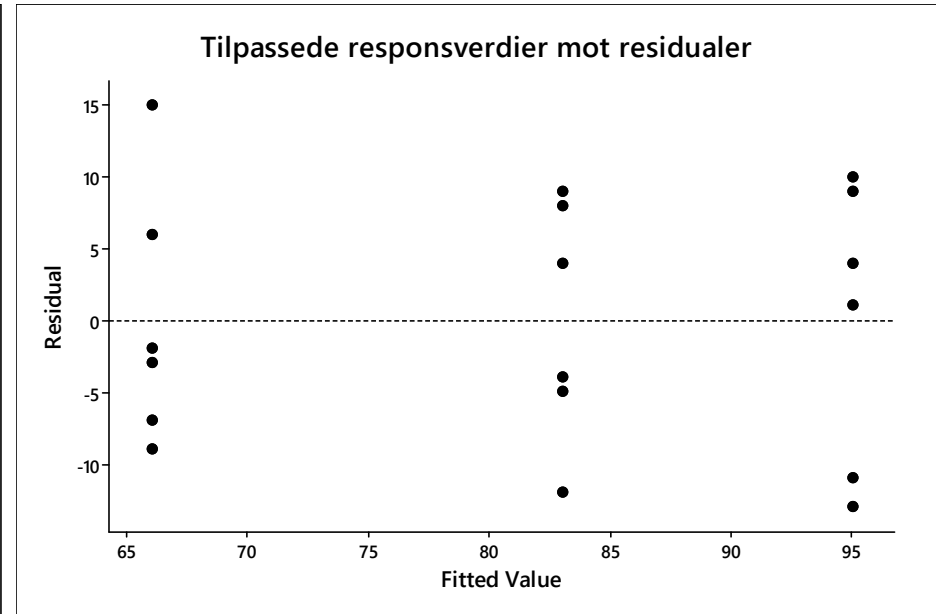
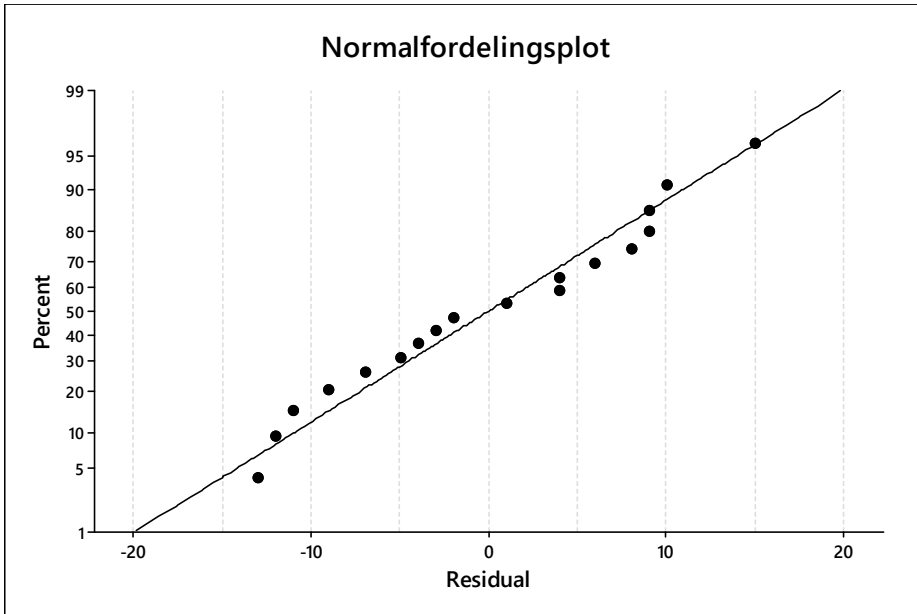
Behandling	N	Mean	Grouping
2: K20 + N	6	95.00	A
4: N + P205	6	83.00	A
3: K20 + P205	6	66.00	B

Gruppering ved bruk av LSD metoden

Behandling	N	Mean	Grouping
2: K20 + N	6	95.00	A
4: N + P205	6	83.00	B
3: K20 + P205	6	66.00	C

Noen sentrale begreper, framgangsmåte etc., via et eksempel, IX

Sjekking av antagelsene når data fra kontrollbehandlingen er utelatt:



Noen retningslinjer ved planlegging, gjennomføring og analyse av forsøk

- 1. Definer og spesifiser problemstillingene (hypotesene).
- 2. Velg responsvariabler.
- 3. Velg faktorer, deres nivåer og eventuelle kovariabler.
- 4. Velg forsøksplan.
- 5. Utfør forsøket og datainnsamlingen.
- 6. Modeller og analyser dataene statistisk (estimer parametere og test hypotesene).
- 7. Konkluder, rapporter og gi råd.

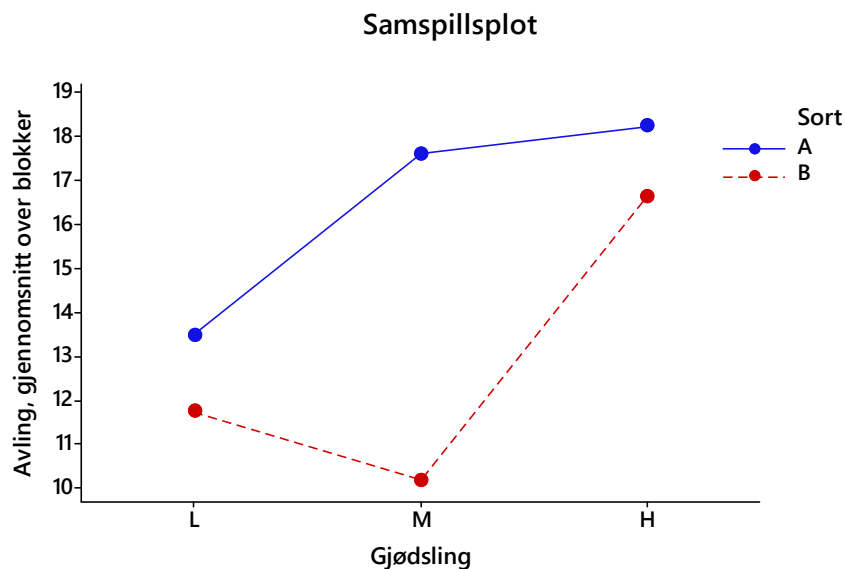
Noen punkter å ha med seg:

- Bruk mest mulig tilgjengelig kunnskap om problemstillingene.
- Utfør forsøket og analysene nøyaktig, grundig og enklest mulig.
- Husk forskjellen mellom praktisk signifikans og statistisk signifikans.
- Et forsøk gir neppe svar på alle problemstillingene. Forsøksvirksomhet er iterativ.

Randomisert fullstendig blokkforsøk, statistisk modell og analyse, I

Det er gjort et faktorielt randomisert fullstendig blokk forsøk med 4 blokker for å studere hvordan avlingen avhenger av sort (2 sorter, A og B) og gjødsling (3 nivåer, Lav, Middels og Høy).

Blokk 1	Blokk 2	Blokk 3	Blokk 4
Rute 101 (A M) 18.1	Rute 201 (A H) 18.2	Rute 301 (A M) 17.1	Rute 401 (B H) 16.6
Rute 102 (B H) 15.7	Rute 202 (B H) 17.5	Rute 302 (A H) 18.9	Rute 402 (A H) 18.2
Rute 103 (B L) 12.6	Rute 203 (A M) 17.6	Rute 303 (B H) 16.7	Rute 403 (A M) 17.6
Rute 104 (A H) 17.6	Rute 204 (A L) 14.5	Rute 304 (B M) 8.3	Rute 404 (B M) 9.1
Rute 105 (A L) 14.3	Rute 205 (B L) 11.2	Rute 305 (B L) 11.0	Rute 405 (A L) 13.6
Rute 106 (B M) 10.5	Rute 206 (B M) 12.8	Rute 306 (A L) 11.5	Rute 406 (B L) 12.1



Statistisk modell for dataene:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + B_k + \varepsilon_{ijk}$$

Y_{ijk} er observert avling for sort i ($i = A, B$), gjødsling j ($j = L, M, H$) i blokk k ($k = 1, 2, 3, 4$).

$$E(Y_{ijk}) = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} \quad \text{Var}(Y_{ijk}) = \sigma_B^2 + \sigma^2$$

Den tilfeldige blokkfaktoren tar høyde for at observasjoner fra samme blokk kan være korrelerte.

Tests of Fixed Effects

Term	DF Num	DF Den	F-value	P-value
Sort	1.00	15.00	72.46	0.000
Gjødsling	2.00	15.00	46.75	0.000
Sort*Gjødsling	2.00	15.00	20.65	0.000

Variance Components

Source	Var	% of Total	SE Var	Z-value	P-Value
Blokk	0.1539	12.59	0.2787	0.5522	0.290
Error	1.0682	87.41	0.3900	2.7386	0.003
Total	1.2221				

Gruppering ved bruk av Tukey's metode

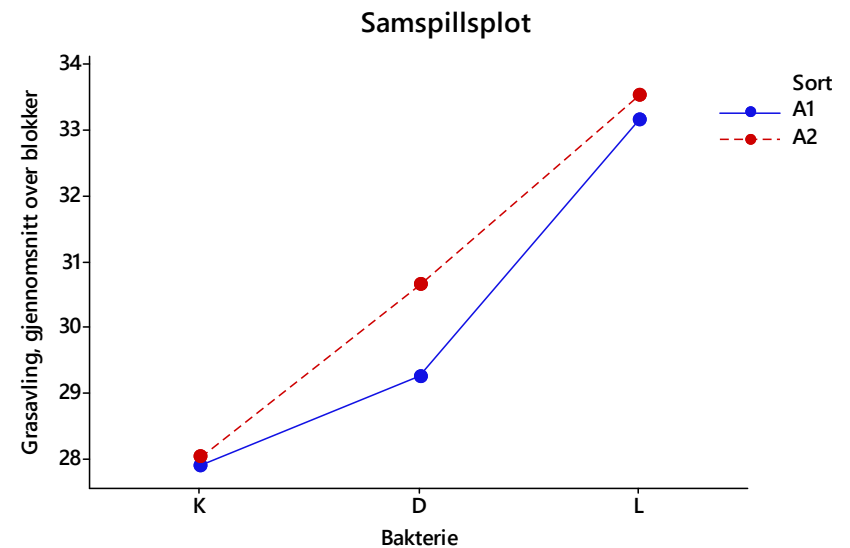
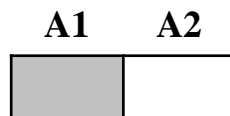
Sort*Gjødsling		N	Mean	Grouping
A	H	4	18.225	A
A	M	4	17.600	A
B	H	4	16.625	A
A	L	4	13.475	B
B	L	4	11.725	B C
B	M	4	10.175	C

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

Split plot forsøk, statistisk modell og analyse, I

Grasavling (tørrstoff) studeres. Storrutefaktoren er sort (2 sorter, A1 og A2). Smårutefaktoren er behandling med en bakterie (3 nivåer, Kontroll, Levende og Død). Forsøket ble gjentatt i 4 blokker.

Blokk 1	Blokk 2	Blokk 3	Blokk 4
K 29.4	D 28.7	D 29.7	K 26.7
L 34.4	L 33.4	K 28.6	L 31.8
D 32.5	K 28.9	L 32.9	D 28.9
K 27.4	L 36.4	K 27.2	D 28.6
L 34.5	D 32.4	L 32.6	L 30.7
D 29.7	K 28.7	D 29.1	K 26.8



Split plot forsøk, statistisk modell og analyse, II

Statistisk modell for dataene:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + B_k + (B\alpha)_{ik} + \varepsilon_{ijk}$$

Y_{ijk} er observert grasavling for sort i ($i = A1, A2$), bakterie j ($j = K, L, D$) i blokk k ($k = 1, 2, 3, 4$).

$$E(Y_{ijk}) = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} \quad \text{Var}(Y_{ijk}) = \sigma_B^2 + \sigma_{B\alpha}^2 + \sigma^2$$

De tilfeldige faktorene tar høyde for at observasjoner fra samme storrute i samme blokk kan være korrelerte og minst like korrelerte som observasjoner fra samme blokk, men fra forskjellige storruter.

Tests of Fixed Effects

Term	DF Num	DF Den	F-value	P-value
Sort	1.00	3.00	0.76	0.447
Bakterie	2.00	12.00	83.76	0.000
Sort*Bakterie	2.00	12.00	1.29	0.310

Variance Components

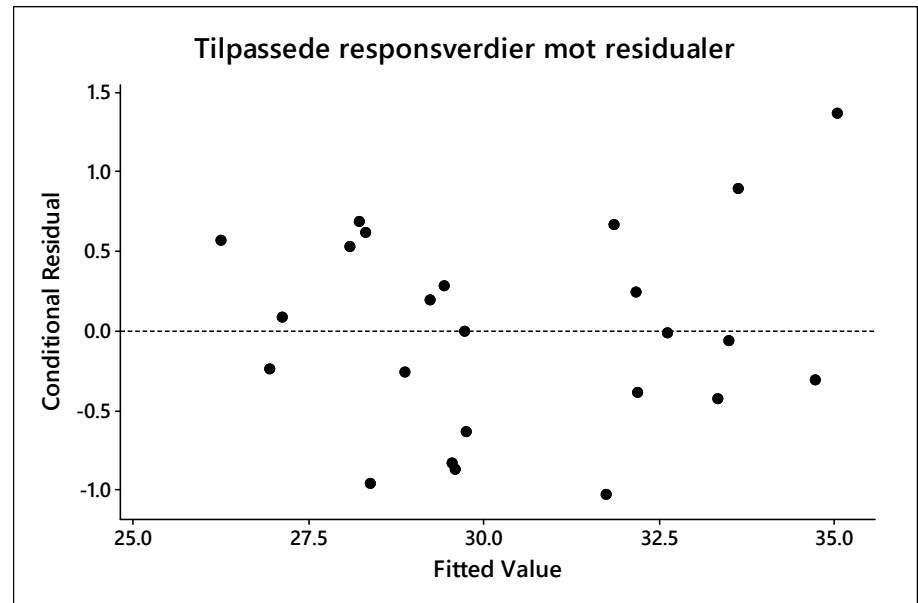
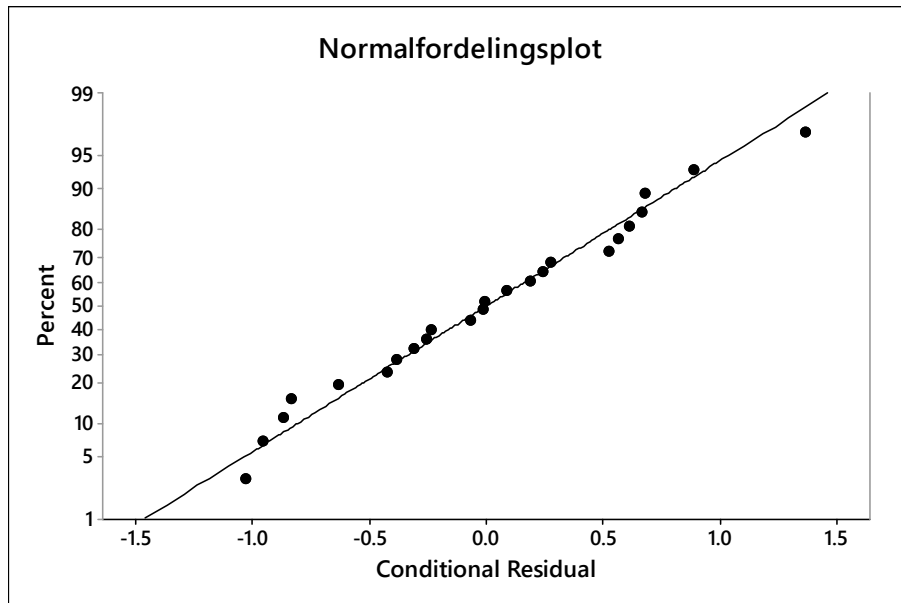
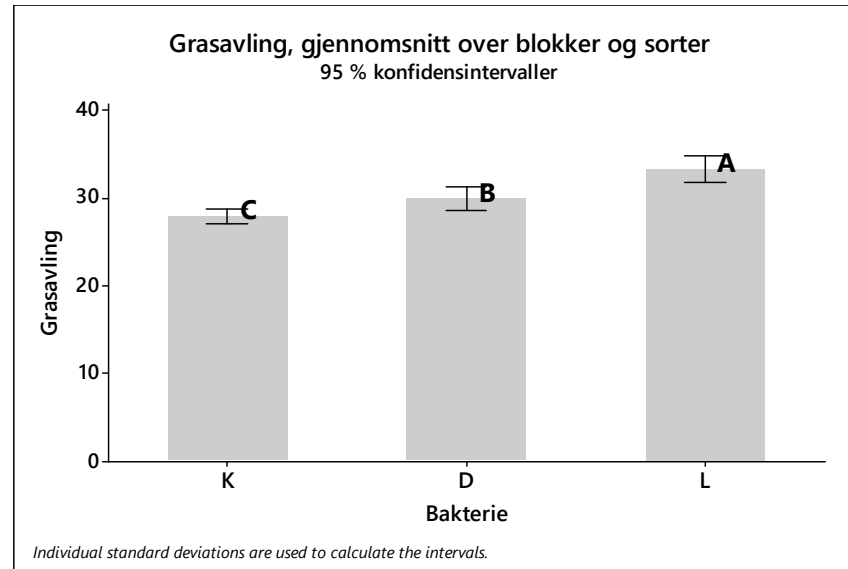
Source	Var	% of Total	SE Var	Z-value	P-Value
Blokk	0.8800	36.61	1.2264	0.7175	0.237
Blokk*Sort	0.8182	34.04	0.8654	0.9455	0.172
Error	0.7054	29.35	0.2880	2.4495	0.007
Total	2.4036				

Gruppering ved bruk av Tukey's metode

Bakterie	N	Mean	Grouping
L	8	33.3375	A
D	8	29.9500	B
K	8	27.9625	C

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

Split plot forsøk, statistisk modell og analyse, IV



Split split plot forsøk, statistisk modell

- I et split split plot forsøk inngår 3 faktorer, la oss kalle dem A (storrutefaktoren), B (smårutefaktoren) og C (små smårutefaktoren).
- Nivåene til A randomiseres på storrutene innen hver blokk.
- Nivåene til B randomiseres på smårutene innen hver storrute innen hver blokk.
- Nivåene til C randomiseres på små smårutene innen hver smårute innen hver storrute innen hver blokk.

Statistisk modell for dataene:

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + B_l + (B\alpha)_{il} + (B\alpha\beta)_{ijl} + \varepsilon_{ijkl}$$

Y_{ijkl} er observert verdi på responsvariabelen for nivå i av A, nivå j av B, nivå k av C i blokk l .

$$E(Y_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk}$$

$$\text{Var}(Y_{ijk}) = \sigma_B^2 + \sigma_{B\alpha}^2 + \sigma_{B\alpha\beta}^2 + \sigma^2$$

De tilfeldige faktorene tar høyde for eventuelle korrelasjonsstrukturer mellom observasjonene.

Split blokk forsøk (= strip plot forsøk), statistisk modell

- I et split blokk forsøk inngår 2 faktorer, la oss kalle dem A (radfaktoren) og B (kolonnefaktoren).
- Nivåene til A randomiseres på radene innen hver blokk.
- Nivåene til B randomiseres på kolonnene innen hver blokk.

Statistisk modell for dataene:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + B_k + (B\alpha)_{ik} + (B\beta)_{jk} + \varepsilon_{ijk}$$

Y_{ijk} er observert verdi på responsvariabelen for nivå i av A, nivå j av B i blokk k .

$$E(Y_{ijk}) = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$$

$$\text{Var}(Y_{ijk}) = \sigma_B^2 + \sigma_{B\alpha}^2 + \sigma_{B\beta}^2 + \sigma^2$$

De tilfeldige faktorene tar høyde for eventuelle korrelasjonsstrukturer mellom observasjonene.

3 økologiske gjødslingsmidler skal sammenlignes for dyrking av mangold. 12 felt er disponible og hvert gjødslingsmiddel ble brukt på 4 tilfeldig valgte felt. Fra hvert felt ble det tatt prøver fra 2 tilfeldig valgte mangoldplanter der innholdet av jern i planten ble målt.

Forsøksplan og data framgår nedenfor.

Gjødsling	Felt	Prøve(felt)	Jern
1	3	1	102.4
1	3	2	98.3
1	6	1	99.7
1	6	2	99.3
1	2	1	100.1
1	2	2	100.4
1	11	1	97.0
1	11	2	99.2
2	10	1	96.4
2	10	2	98.8
2	1	1	100.7
2	1	2	98.1
2	8	1	101.2
2	8	2	101.5
2	5	1	97.5
2	5	2	97.6
3	9	1	103.8
3	9	2	104.1
3	12	1	105.6
3	12	2	104.7
3	7	1	109.1
3	7	2	108.4
3	4	1	101.4
3	4	2	102.6

Statistisk modell for dataene:

$$Y_{ijk} = \mu + \alpha_i + B_{j(i)} + \varepsilon_{ijk}$$

Y_{ijk} er observert innhold av jern i prøve k ($k = 1, 2$) fra felt j ($j = 1, 2, \dots, 12$) innen gjødsling i ($i = 1, 2, 3$).

$$E(Y_{ijk}) = \mu + \alpha_i$$

$$\text{Var}(Y_{ijk}) = \sigma_{F(G)}^2 + \sigma^2$$

Den tilfeldige faktoren tar høyde for at observasjonene i prøvene fra samme felt (innen gjødsling) kan være korrelerte.

Tests of Fixed Effects

Term	DF Num	DF Den	F-value	P-value
Gjødsling	2.00	9.00	10.57	0.004

Variance Components

Source	Var	% of Total	SE Var	Z-value	P-Value
Felt(Gjødsling)	3.35	68.3	1.97	1.70	0.045
Error	1.56	31.7	0.64	2.50	0.007
Total	4.91				

Gruppering ved bruk av Tukey's metode

Gjødsling	N	Mean	Grouping
3	8	104.96	A
1	8	99.55	B
2	8	98.98	B

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

Hva hvis det mangler observasjoner fra noen felt?

Manglende verdier markert med * nedenfor.

Gjødsling	Felt	Prøve(felt)	Jern
1	3	1	102.4
1	3	2	98.3
1	6	1	*
1	6	2	99.3
1	2	1	100.1
1	2	2	100.4
1	11	1	97.0
1	11	2	99.2
2	10	1	96.4
2	10	2	98.8
2	1	1	100.7
2	1	2	98.1
2	8	1	*
2	8	2	*
2	5	1	97.5
2	5	2	97.6
3	9	1	103.8
3	9	2	104.1
3	12	1	*
3	12	2	104.7
3	7	1	*
3	7	2	*
3	4	1	101.4
3	4	2	102.6

Samme statistiske modell for disse ubalanserte og ufullstendige dataene som vi brukte for de balanserte og fullstendige dataene.

Tests of Fixed Effects

Term	DF Num	DF Den	F-value	P-value
Gjødsling	2.00	7.01	14.54	0.003

Variance Components

Source	Var	% of Total	SE Var	Z-value	P-Value
Felt(Gjødsling)	0.24	10.3	0.90	0.27	0.394
Error	2.14	89.9	1.02	2.09	0.018
Total	2.38				

Gruppering ved bruk av Tukey's metode

Gjødsling	N	Mean	Grouping
3	5	103.35	A
1	7	99.53	B
2	6	98.18	B

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

Forsøksserier, over år og/eller sted, I

- Bestem om sted skal modelleres som tilfeldig eller fast faktor.
- År modelleres nesten alltid som tilfeldig faktor.
- Som regel får vi mye informasjon ved å inkludere samspill i modellen, i alle fall i utgangspunktet. Det være seg samspill mellom behandlingene og årene, mellom behandlingene og stedene, og eventuelt mellom årene og stedene.

2 eksempler på modeller (på en svært skjematisk form):

- Behandling modelleres som fast.
- År og sted og samspill der år og sted inngår modelleres som tilfeldige.
- Forsøk modelleres som tilfeldig.
- Behandlingene kan være kombinasjoner av nivåene til to eller flere faktorer.

Forsøk fra flere år. Flere forsøk pr. år. Forsøk fra et sted.

$$Y = \textit{Behandling} + \textit{År} + \textit{År} \cdot \textit{Behandling} + \textit{Forsøk} + \textit{Error}$$

Forsøk fra flere år og flere steder. Flere forsøk pr. år på hvert sted.

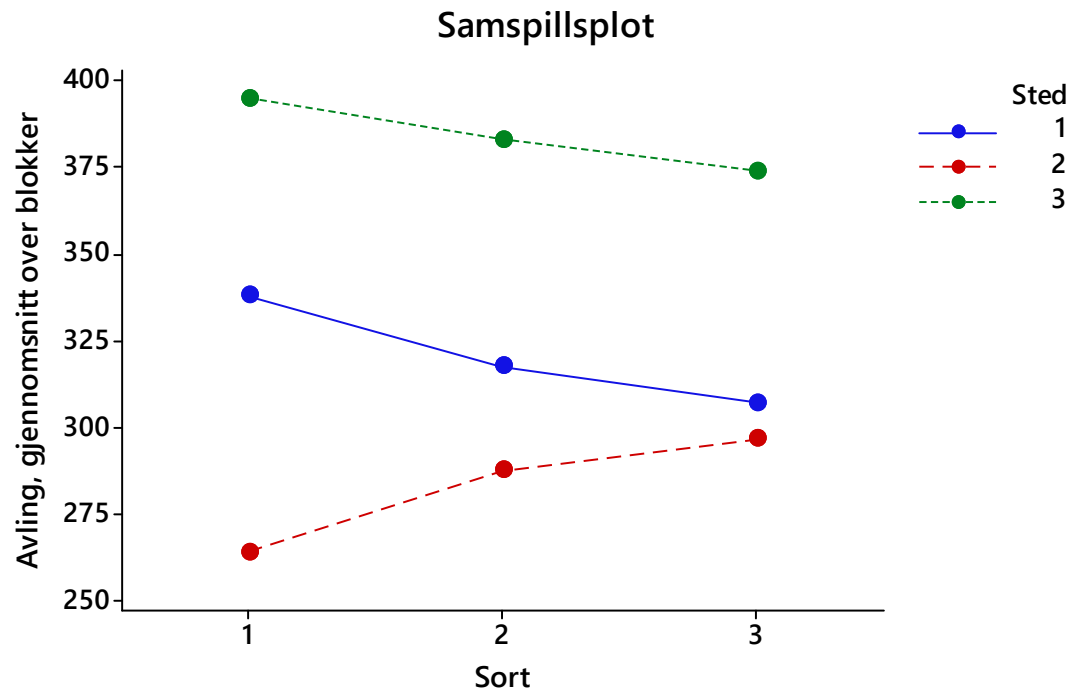
$$Y = \textit{Behandling} + \textit{År} + \textit{Sted} + \textit{År} \cdot \textit{Behandling} + \textit{Sted} \cdot \textit{Behandling} + \textit{År} \cdot \textit{Sted} + \textit{Forsøk} + \textit{Error}$$

Forsøksserier, over år og/eller sted, II

Eksempel.

For å sammenligne 3 sorter er det er gjort et randomisert blokk forsøk på 3 steder. Her er responsvariabelen avling. Data og noen resultater er gitt nedenfor.

Sted	Blokk	Sort	Rute	Avling
1	1	1	12	332
1	1	2	11	312
1	1	3	13	299
1	2	1	23	344
1	2	2	22	323
1	2	3	21	315
2	3	1	33	250
2	3	2	32	276
2	3	3	31	289
2	4	1	42	278
2	4	2	41	299
2	4	3	43	304
3	5	1	51	401
3	5	2	53	388
3	5	3	52	377
3	6	1	61	389
3	6	2	62	378
3	6	3	63	371



Forsøksserier, over år og/eller sted, III

Statistisk modell for dataene:

$$Y = \text{Sort} + \text{Sted} + \text{Blokk}(\text{Sted}) + \text{Sted} \cdot \text{Sort} + \text{Error}$$

Sted modelleres som *tilfeldig* faktor

Tests of Fixed Effects

Term	DF Num	DF Den	F-value	P-value
Sort	2.00	4.00	0.10	0.903

Variance Components

Source	Var	P-Value
Sted	2457.60	0.174
Blokk(Sted)	120.06	0.117
Sted*Sort	299.76	0.082
Error	9.89	0.042
Total	2887.31	

Sted modelleres som *fast* faktor

Tests of Fixed Effects

Term	DF Num	DF Den	F-value	P-value
Sted	2.00	3.00	42.47	0.006
Sort	2.00	6.00	6.42	0.032
Sted*Sort	4.00	6.00	61.63	0.000

Variance Components

Source	Var	P-Value
Blokk(Sted)	120.06	0.117
Error	9.89	0.042
Total	129.95	

Programvare

Krav:

- Tilfeldig trekning.
- Generelle lineære modeller (normalfordelte observasjoner).
- Generaliserte lineære modeller (ikke nødvendigvis normalfordelte observasjoner).
- Modellene må tillate faste og tilfeldige faktorer.

Programvare:

- Minitab.
- Sas.
- Statistica.
- Matlab.
- R.
- Jmp.
- Sigmaplot.
- Stata.
- Statgraphics.
- Spss.
- Systat.
- ⋮

Avslutning

Bruk tid på planlegging.

Gjør ting så enkelt som mulig.

Vær nøyaktig og grundig i planlegging, gjennomføring, datainnsamling, statistisk modellering, statistisk analyse, konklusjon, rapportering, rådgiving

Vær redelig i rapportering og konklusjoner, husk at det er data beheftet med usikkerhet som ligger til grunn.

Og , atter en gang:

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