

# Case 7: konfunderinger i ufuldstændigt blokforsøg

## Statistisk Dataanalyse 2

Anders Tolver

Institut for Matematiske Fag

Uge 7, torsdag d. 26/10-2017

## Case 7: forsøgsdesign og statistisk model (spm. a-c)

Forsøget er et  $2^3$ -forsøg med behandlingerne N, P og K hver på 2 niveauer. Hver behandlingskombination givet ved  $N \times P \times K$  forekommer 3 gange i forsøgsplanen.

Forsøgshederne er inddelt i 6 blokke med 4 parceller i hver.

Den statistiske analyse bør tage udgangspunkt i modellen

$$Y_i = \delta(N \times P \times K_i) + b(\text{blok}_i) + e_i, b(j) \sim N(0, \sigma_B^2), e_i \sim N(0, \sigma^2)$$

hvor vi lader blok indgå med tilfældig effekt.

Slutmodellen (efter reduktion af den systematiske del) bliver

$$Y_i = \alpha(N_i) + \beta(K_i) + b(\text{blok}_i) + e_i, b(j) \sim N(0, \sigma_B^2), e_i \sim N(0, \sigma^2).$$

Ved brug af reglen fra theorem 9.11 i kompendiet ses, at forsøgsplanen består af 3 gentagelser af et  $2^3$ -forsøg på blokke, hvor man i hver gentagelse har konfunderet trefaktorvekselvirkningen  $N \times P \times K$  med blok.

## Case 7: modelreduktion (spm. d)

```
> modelblock = lm(yield ~block + N*P*K)
> summary(modelblock)
Coefficients: (1 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  47.9250     2.7785   17.248  7.8e-10 ***
block2        3.9000     2.7785    1.404  0.18578
block3        0.4000     2.7785    0.144  0.88792
block4        7.3250     2.7785    2.636  0.02172 *
block5        6.2250     2.7785    2.240  0.04477 *
block6       10.6500     2.7785    3.833  0.00238 **
N2            9.8500     2.7785    3.545  0.00403 **
P2            0.4167     2.7785    0.150  0.88329
K2           -1.9167     2.7785   -0.690  0.50344
N2:P2        -3.7667     3.2084   -1.174  0.26317
N2:K2        -4.7000     3.2084   -1.465  0.16865
P2:K2         0.5667     3.2084    0.177  0.86275
N2:P2:K2      NA          NA      NA      NA
```

Manglende estimat for trefaktorvekselvirkningen!

## Case 7: estimation af vekselvirkninger (spm. e-i)

### blok som systematisk faktor

Estimat for effekt af vekselvirkning mellem N og P:  $-3.77$   $[-10.76, 3.22]$

Da  $N \times P \times K$  er konfunderet med blok kan effekten af trefaktorvirkning ikke estimeres, når vi inddrager blok som systematisk faktor i modellen.

### blok som tilfældig faktor

Estimat for effekt af vekselvirkning mellem N og P:  $-3.77$   $[-10.76, 3.22]$

Estimat for effekt af trefaktorvekselvirkning:  $9.93$   $[-15.75, 35.61]$

Selvom  $N \times P \times K$  er konfunderet med blok kan effekten af trefaktorvirkning godt estimeres, når vi inddrager blok som tilfældig faktor i modellen. Der er dog meget stor spredning på estimatet!

### Diskussion/forklaring

Hvis blok er systematisk kan effekten af  $N \times P \times K$  ikke skelnes fra blokeffekten.

Hvis blok er tilfældig kan effekten af  $N \times P \times K$  estimeres, og spredningen på estimatet er baseret på den resterende variation mellem blokke. Dette svarer til at estimere forskelle for helplotfaktoren i et split-plot forsøg.

## Case 7: forbedring af forsøgsplanen (spm. j-k)

Forsøgsplanen består af 3 gentagelser af et  $2^3$ -forsøg på blokke, hvor man i hver gentagelse har konfunderet trefaktorvekselvirkningen  $N \times P \times K$  med blok.

Vi kunne i stedet vha. theorem 9.11 have konfunderet 3 *forskellige* effekter med blok på hvert af de tre par af blokke. På denne måde ville det være muligt at teste og estimere alle hoved- og vekselvirkninger. Dette trick kaldes **partiell konfundering**.

Alternativt kunne man udvide forsøget til at omfatte 7 par af blokke, hvor man konfunderer hver af de 7 hoved- eller vekselvirkninger på et par af blokke.

Man kan eftervise, at det fremkomne design bliver et BIBD:

- ▶ antal behandlinger  $\nu_T = 8$
- ▶ blokstørrelse  $r_B = 4$
- ▶ antal blokke  $\nu_B = 14$
- ▶ hver beh. optræder  $r_T = 7$  gange i forsøgsplanen
- ▶ hver par af beh. optræder  $\lambda = 3$  gange inden for samme blok

## Case 7: LSD-værdi ved sammenlign. af beh. (spm. I)

Fra startmodellen i spm. b)

$$Y_i = \delta(N \times P \times K_i) + b(\text{blok}_i) + e_i, b(j) \sim N(0, \sigma_B^2), e_i \sim N(0, \sigma^2)$$

fås estimerterne

```
> finalmodel = lme(yield ~ N+K, random= ~1|block,method="REML")
> VarCorr(finalmodel) ## gives the variance estimates
block = pdLogChol(1)
              Variance StdDev
(Intercept) 13.28062 3.644259
Residual    15.53666 3.941657
```

Indsættelse i formel (9.11) s. 169 i kompendiet giver

$$\text{LSD}_{0.95} = 3.942 \cdot 2.030 \cdot \sqrt{\frac{2(4 \cdot 13.281 + 15.537)}{3 \cdot 8 \cdot 13.281 + 7 \cdot 15.537}} = 4.53$$

[– se detaljer om beregningen i R-program]