STAT 516 Lec 09

Randomized complete block designs

Karl Gregory

2026-11-04

Nitrogen fertilization data from Kuehl (2000)

Six nitrogen timing schedules in each of four plots in a field of wheat.

Response is the nitrate content in stem tissue samples.

Display 8.1 Arrangement of Experimental Plots for the Wheat Experiment in a Randomized Complete Block Design

Irrigation Gradient ↓

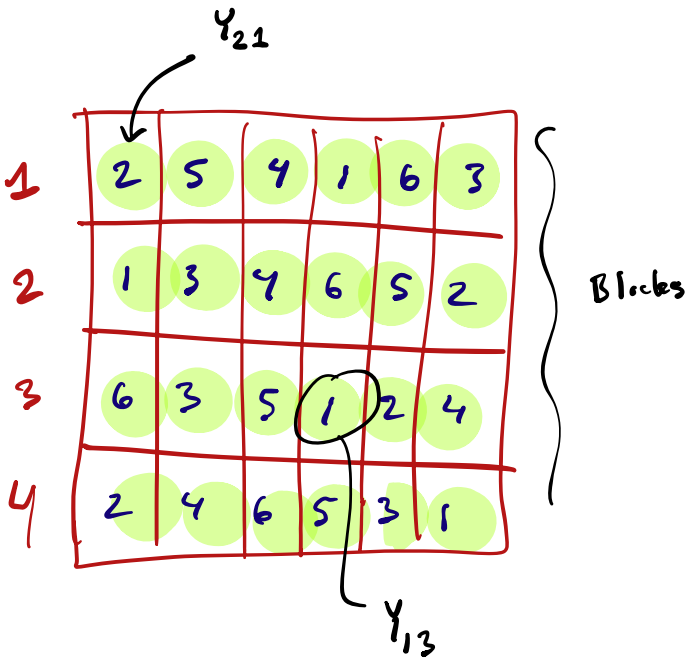
Block 1	2	5	4	1	6	3
	40.89	37.99	37.18	34.98	34.89	42.07
Block 2	1	3	4	6	5	2
	41.22	49.42	45.85	50.15	41.99	46.69
Block 3	6	3	5	1	2	4
	44.57	52.68	37.61	36.94	46.65	40.23
Block 4	2	4	6	5	3	1
	41.90	39.20	43.29	40.45	42.91	39.97

Source: Dr. T. Doerge, Department of Soil and Water Science, University of Arizona.

"Random assignment"

1	2	3	3	3	2	3
2	4	1	2	5	4	1
3	2	5	4	1	6	5
4	1	5	6	6	6	4

"Blocked"



R.C.B.D.

$$Y_{ij} = \mu + \tau_i + B_j + \varepsilon_{ij}$$

- Y_{ij}: fixed effect of trt i
- μ: baseline / overall mean
- τ_i: fixed effect of trt i
- B_j: random effect of block j
- ε_{ij}: noise

$$\varepsilon_{ij} \overset{ind}{\sim} N(0, \sigma^2_\varepsilon)$$

$$B_j \overset{ind}{\sim} N(0, \sigma^2_B)$$

One-way random effects model:

$$Y_{ij} = \mu + A_i + \varepsilon_{ij}$$

Nitrogen fertilization

```
ntr <- c(40.89,37.99,37.18,34.98,34.89,42.07,
         41.22,49.42,45.85,50.15,41.99,46.69,
         44.57,52.68,37.61,36.94,46.65,40.23,
         41.90,39.20,43.29,40.45,42.91,39.97)
block <- as.factor(c(1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3,4,4,4,4,4,4))
trt <- as.factor(c(2,5,4,1,6,3,1,3,4,6,5,2,6,3,5,1,2,4,2,4,6,5,3,1))
```

Randomized complete block design (RCBD)

- EUs belong to blocks—groups of EUs homogeneous in some way.
- Each EU in a block is randomly assigned to a fixed treatment.
- All treatments appear exactly once in each block.
- Purpose is to capture the between-block variability among the EUs.
- This helps us detect treatment effects with greater power.

Treatment effects model for the RCBD

Assume

$$Y_{ij} = \mu + \tau_i + B_j + \varepsilon_{ij}, \quad \text{for } i = 1, \ldots, a, \quad j = 1, \ldots, b,$$

where

- Y_{ij} is the response of the EU in block j receiving treatment i.
- the τ_i are the fixed effects of the treatment.
- the B_j are independent $\mathcal{N}(0, \sigma_B^2)$ random block effects.
- the ε_{ij} are independent $\mathcal{N}(0, \sigma_\varepsilon^2)$ error/interaction terms.
- μ is an overall or baseline mean.

Define the cell means as

$$\mu_i = \mu + \tau_i, \quad i = 1, \ldots, a.$$

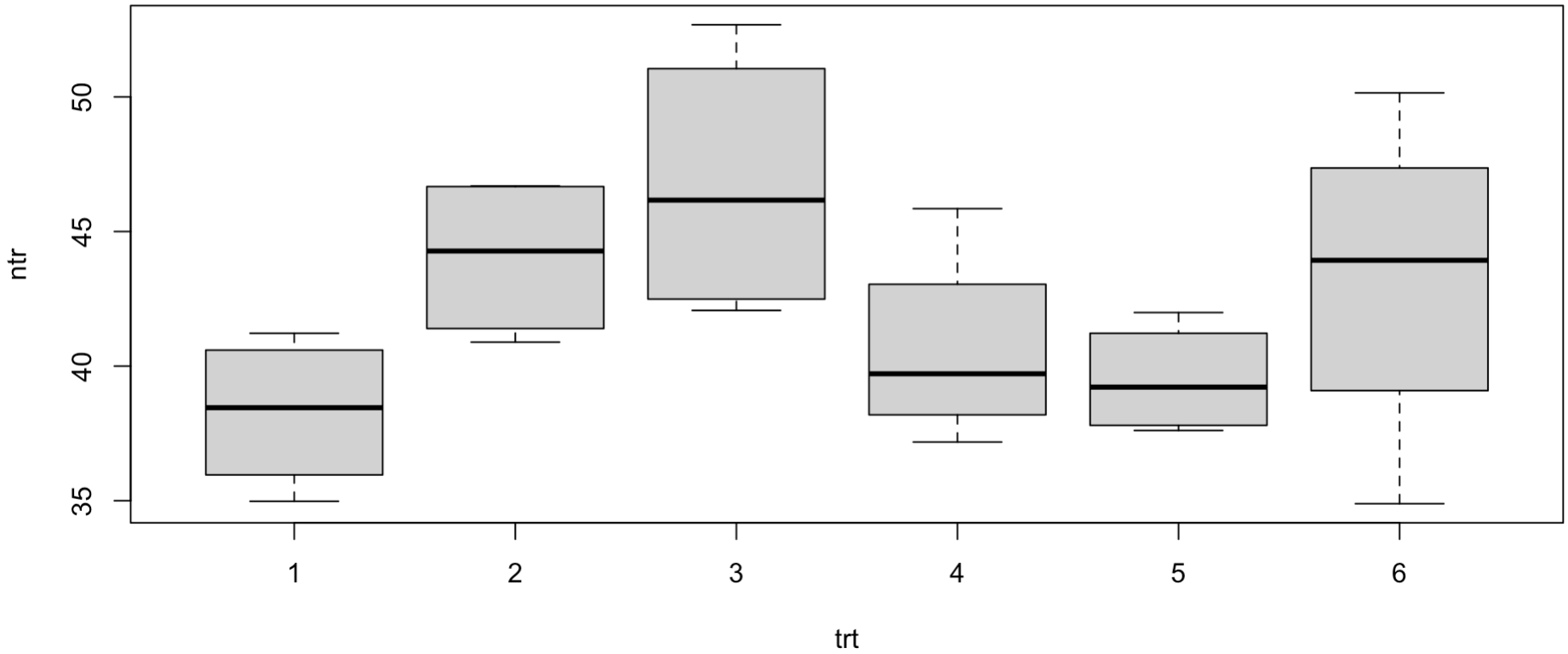
Goals in the RCBD

In the randomized complete block design we wish to

1. Visualize the data.
2. Decompose the variability in the Y_{ij} into its sources.
3. Estimate the variance components σ^2_B and σ^2_ε.
4. Test whether the treatment has any effect.
5. Make comparisons between treatment means.
6. Check whether the model assumptions are satisfied.

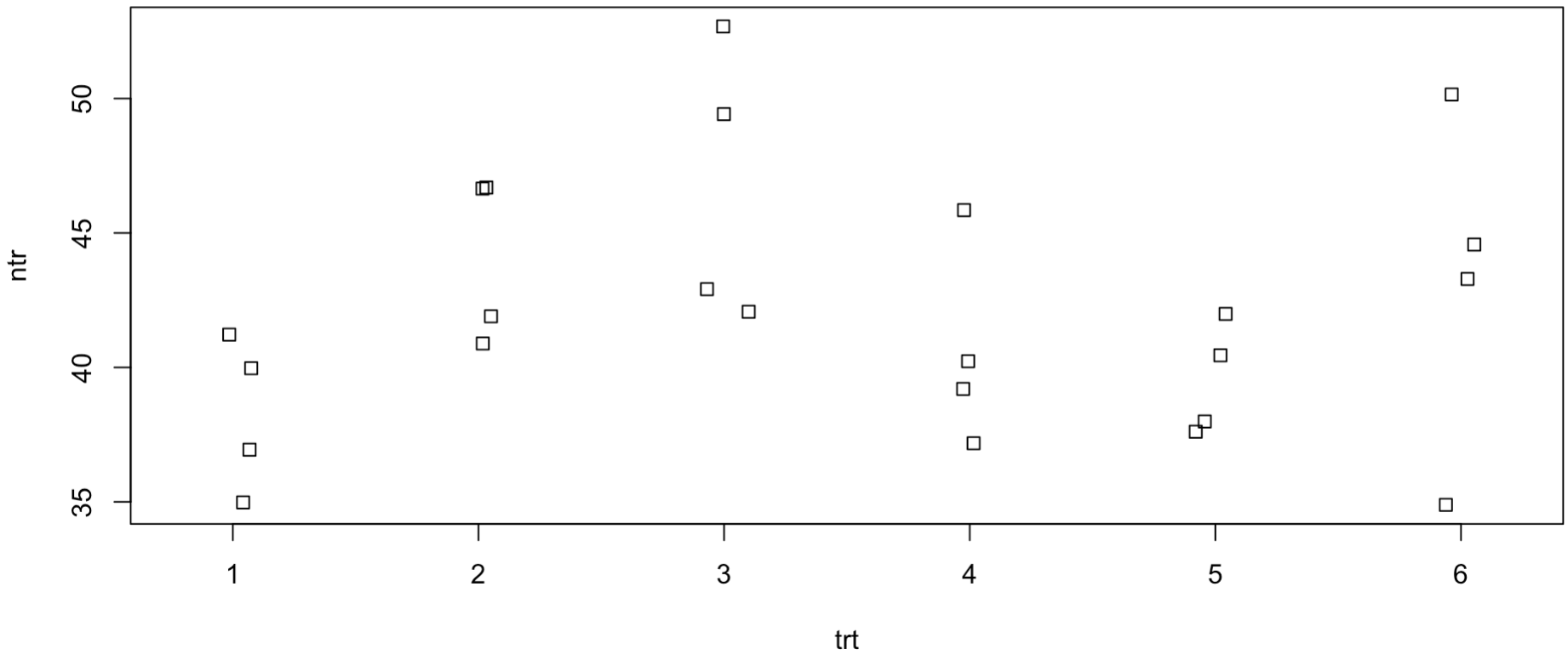
Nitrogen fertilization

```
1 boxplot(ntr~trt)
```



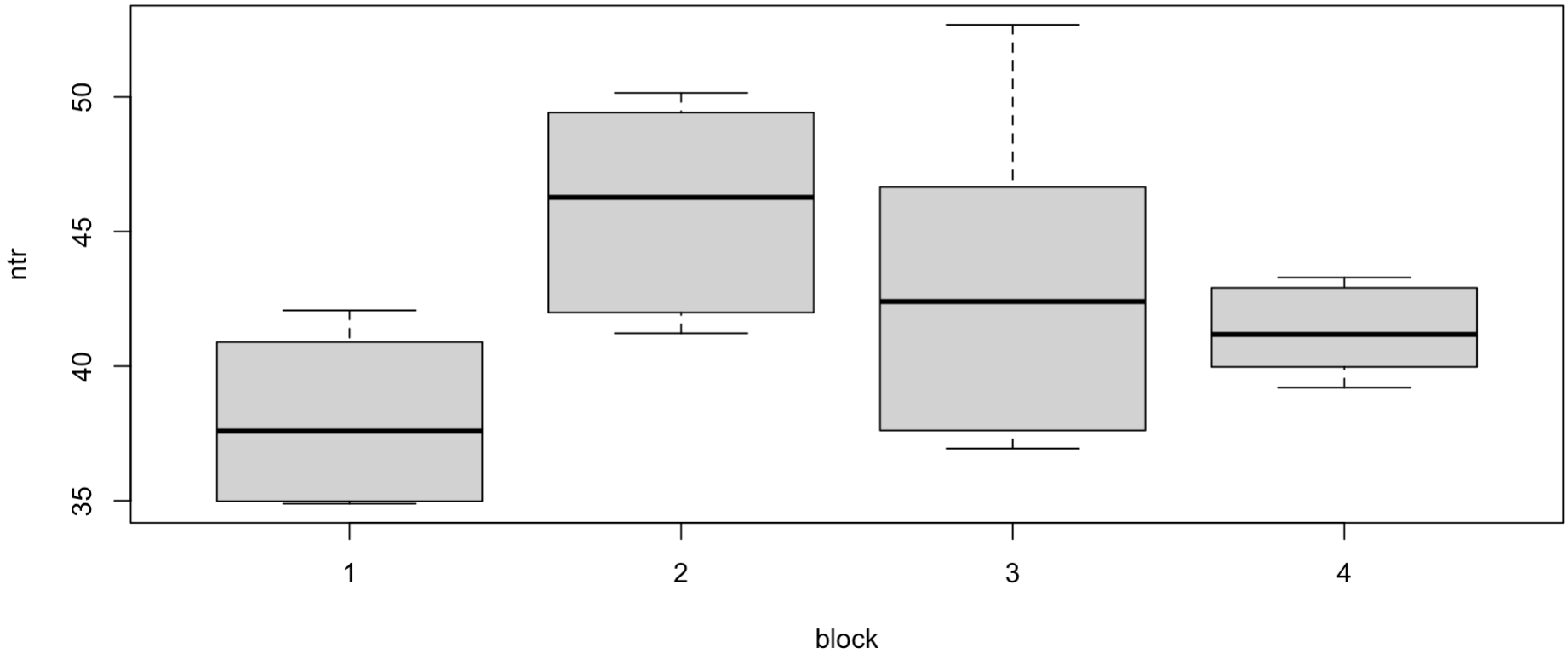
Nitrogen fertilization

```
1 stripchart(ntr~trt,vertical=T,method="jitter",xlab = "trt")
```



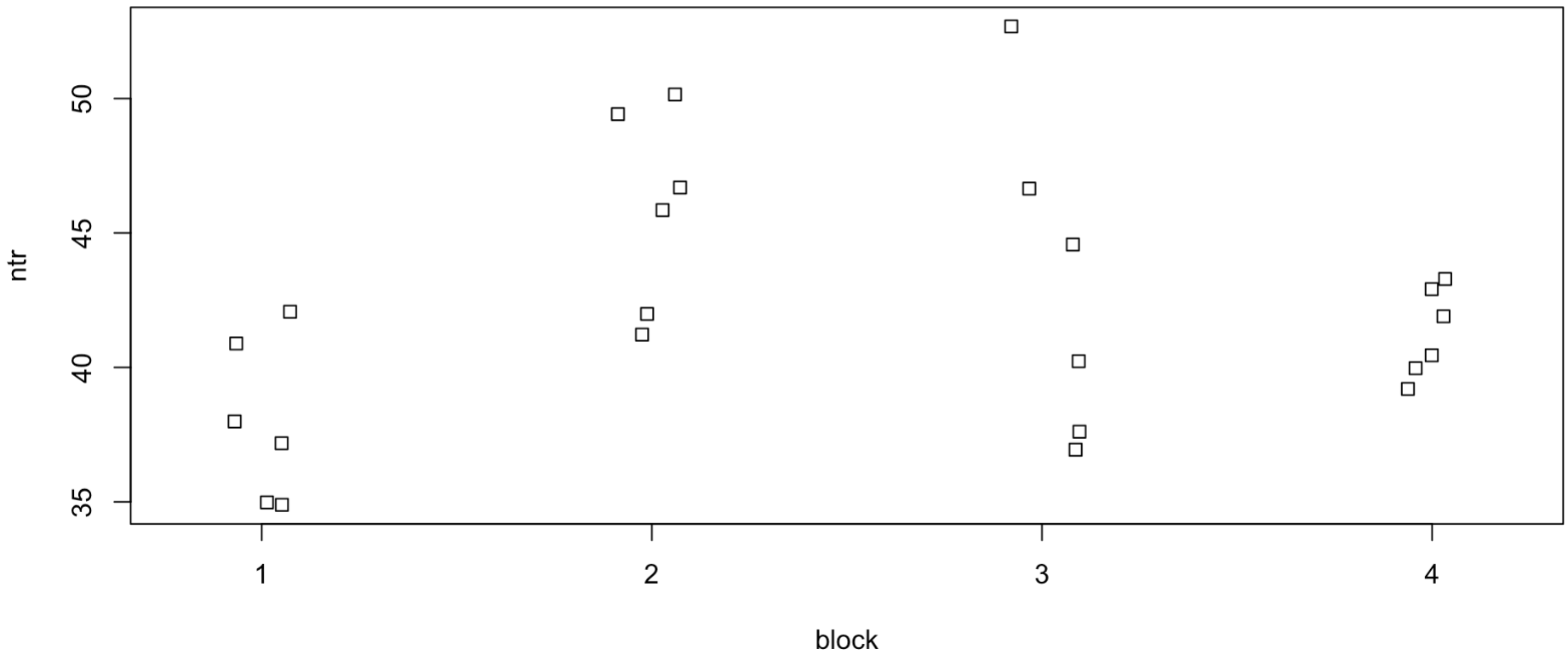
Nitrogen fertilization

```
1 boxplot(ntr~block)
```



Nitrogen fertilization

```
1 stripchart(ntr~block,vertical=T,method = "jitter",xlab = "block")
```



Sums of squares for the RCBD

Sum of squares	Symbol	Formula
Total	$\mathrm{SS}_{\mathrm{Tot}}$	$\sum_{i=1}^{a}\sum_{j=1}^{b}(Y_{ij}-\bar{Y}_{\cdot\cdot})^2$
Treatment	SS_{A}	$b\sum_{i=1}^{a}(\bar{Y}_{i\cdot}-\bar{Y}_{\cdot\cdot})^2$
Block	SS_{B}	$a\sum_{j=1}^{b}(\bar{Y}_{\cdot j}-\bar{Y}_{\cdot\cdot})^2$
Error	$\mathrm{SS}_{\mathrm{Error}}$	$\sum_{i=1}^{a}\sum_{j=1}^{b}(Y_{ij}-(\bar{Y}_{i\cdot}+\bar{Y}_{\cdot j}-\bar{Y}_{\cdot\cdot}))^2$

Same as before

- We can make the decomposition $\mathrm{SS}_{\mathrm{Tot}} = \mathrm{SS}_{\mathrm{A}} + \mathrm{SS}_{\mathrm{B}} + \mathrm{SS}_{\mathrm{Error}}$.
- The $\mathrm{SS}_{\mathrm{Error}}$ is really the interaction sum of squares $\mathrm{SS}_{\mathrm{AB}}$.
- But without replication, we cannot estimate an interaction.
- So the interaction serves as the error term.

"no replication"

$n = 1$

$Y_{ij} = \mu + \tau_i + B_j + \varepsilon_{ij}$

a = # levels of Factor A

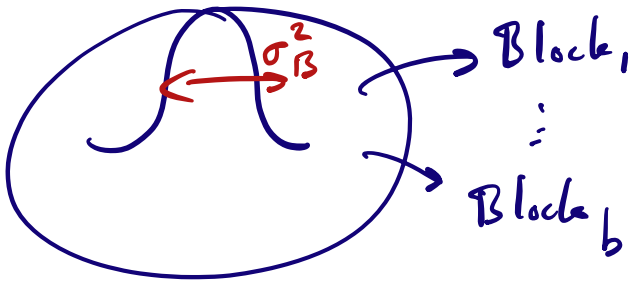
b = # blocks

ANOVA table for RCBD

Source	Df	SS	MS	F value
A	$a-1$	SS_A	MS_A	$F_\mathrm{A} = \mathrm{MS}_\mathrm{A} / \mathrm{MS}_\mathrm{Error}$
B	$b-1$	SS_B	MS_B	$F_\mathrm{B} = \mathrm{MS}_\mathrm{B} / \mathrm{MS}_\mathrm{Error}$
Error	$(a-1)(b-1)$	$\mathrm{SS}_\mathrm{Error}$	$\mathrm{MS}_\mathrm{Error}$	
Total	$ab-1$	SS_Tot		

$\tau_1 = \cdots = \tau_a$ (all trt means are the same)

1. Reject $H_0: \mu_1 = \cdots = \mu_a$ if $F_\mathrm{A} > F_{a-1,(a-1)(b-1),\alpha}$.
2. Reject $H_0: \sigma_B^2 = 0$ if $F_\mathrm{B} > F_{b-1,(a-1)(b-1),\alpha}$.



Nitrogen fertilization

Obtain ANOVA table “manually”:

```
y <- ntr
y.. <- predict(lm(ntr ~ 1))
yi. <- predict(lm(ntr ~ trt))
y.j <- predict(lm(ntr ~ block))

SST <- sum((y - y..)^2)
SSA <- sum((yi. - y..)^2)
SSB <- sum((y.j - y..)^2)
SSE <- SST - (SSA + SSB)

a <- 6
b <- 4

MSA <- SSA/(a-1)
MSB <- SSB/(b-1)
MSE <- SSE/((a-1)*(b-1))

FA <- MSA / MSE
FB <- MSB / MSE

pA <- 1 - pf(FA,a-1,(a-1)*(b-1))
pB <- 1 - pf(FB,b-1,(a-1)*(b-1))
```

Nitrogen fertilization

Source	Df	SS	MS	F value	p value
A	5	201.32	40.26	5.5917	0.0042
B	3	197.00	65.67	9.1198	0.0011
Error	15	108.01	7.20		
Total	23	506.33			

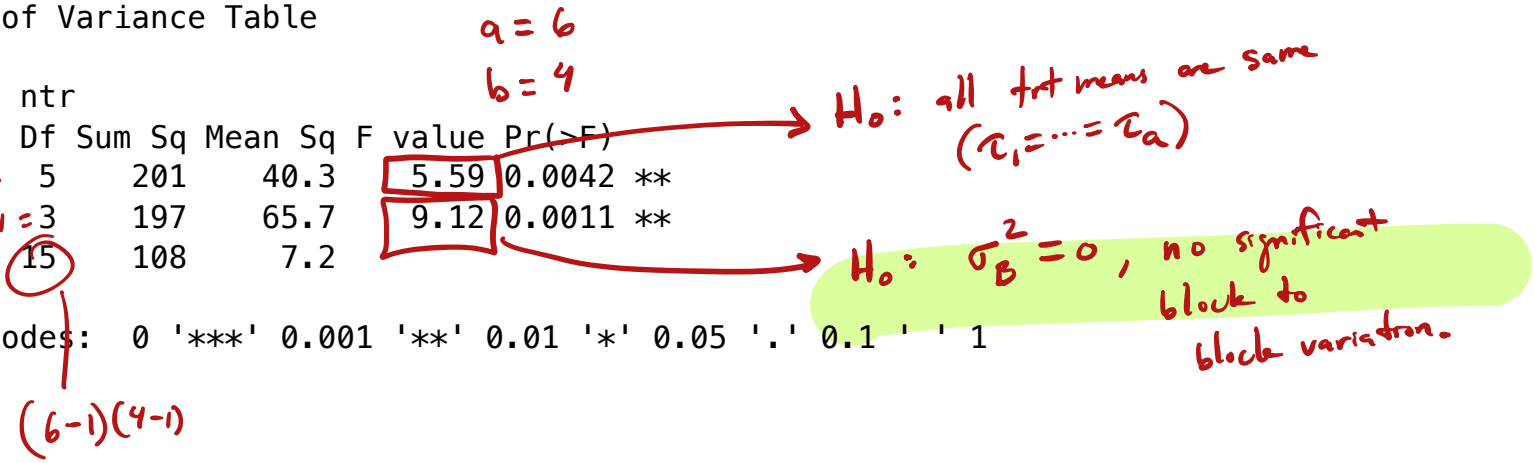
Nitrogen fertilization

Obtain ANOVA table with `anova()` function on `lm()` output:

```
lm_out <- lm(ntr ~ trt + block) # do not include the interaction
anova(lm_out)
```

```
Analysis of Variance Table

Response: ntr
          Df Sum Sq Mean Sq F value Pr(>F)
trt        5    201    40.3    5.59 0.0042 **
block      3    197    65.7    9.12 0.0011 **
Residuals 15    108     7.2
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



Expected mean squares in the RCBD

Source	Df	Expected mean square
A	$a-1$	$b\theta_A^2+\sigma_\varepsilon^2$
B	$b-1$	$a\sigma_B^2+\sigma_\varepsilon^2$
Error	$(a-1)(b-1)$	σ_ε^2

In the above $\theta_A^2=(a-1)^{-1}\sum_{i=1}^{a}(\mu_i-\bar{\mu}_\cdot)^2$.

Method of moments for variance components in RCBD

Equating MS_{B} and $\mathrm{MS}_{\mathrm{Error}}$ with their expectations gives

- $\dot{\sigma}^2_{\varepsilon} = \mathrm{MS}_{\mathrm{Error}}.$
- $\dot{\sigma}^2_{B} = \dfrac{\mathrm{MS}_{\mathrm{B}} - \mathrm{MS}_{\mathrm{Error}}}{a}$

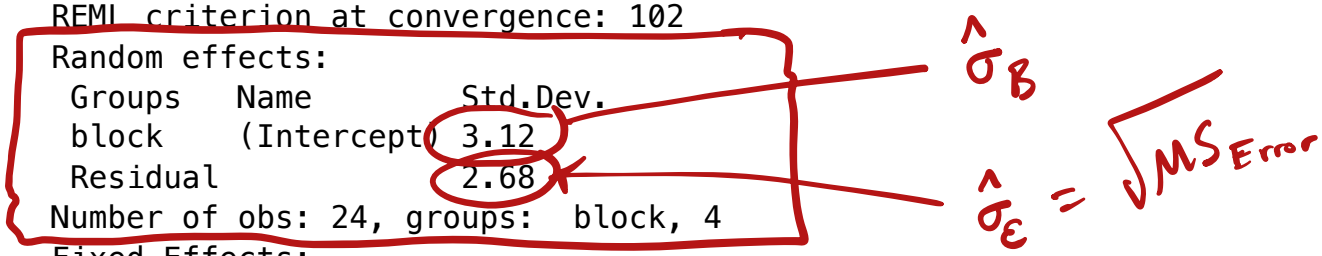
May obtain $\dot{\sigma}^2_{B} < 0$, so one should use REML estimation.

Nitrogen fertilization

Obtain REML estimators of σ_B^2 and σ_ε^2 on the fertilization data.

```
library(lmerTest) # first time run install.packages("lmerTest")
lmer_out <- lmer(ntr ~ trt + (1|block))
lmer_out
```

```
Linear mixed model fit by REML ['lmerModLmerTest']
Formula: ntr ~ trt + (1 | block)
REML criterion at convergence: 102
Random effects:
 Groups   Name        Std.Dev.
 block    (Intercept) 3.12
 Residual             2.68
Number of obs: 24, groups:  block, 4
Fixed Effects:
(Intercept)         trt2         trt3         trt4         trt5         trt6
      38.28         5.75         8.49         2.34         1.23         4.95
```



Nitrogen fertilization

Obtain MoMs estimators for σ_B^2 and σ_ε^2 on the fertilization data.

```
sg_B <- sqrt((MSB - MSE)/a)
sg_e <- sqrt(MSE)
```

We have $\dot{\sigma}_B = 3.122$ and $\dot{\sigma}_\varepsilon = 2.683$.

Variances of means and difference in means

Contrast	Variance	MoM variance estimator
$\bar{Y}_{i.}$	$\frac{1}{b}(\sigma_B^2 + \sigma_\varepsilon^2)$	$\frac{2}{ab}[\mathrm{MS_B} + (b-1)\,\mathrm{MS_{Error}}]$
$\bar{Y}_{i.} - \bar{Y}_{i'.}$	$\frac{2}{b}\sigma_\varepsilon^2$	$\frac{2}{b}\mathrm{MS_{Error}}$

Some (unadjusted) CIs in RCB split plot design

Target	$(1-\alpha)100\%$ **confidence interval**
μ_i	$\bar{Y}_{i.} \pm t_{\nu^*,\alpha/2}\sqrt{\mathrm{MS_B} + (b-1)\,\mathrm{MS_{Error}}}\sqrt{\frac{2}{ab}}$
$\mu_i - \mu_{i'}$	$\bar{Y}_{i.} - \bar{Y}_{i'.} \pm t_{(a-1)(b-1),\alpha/2}\sqrt{\mathrm{MS_{Error}}}\sqrt{\frac{2}{b}}$

focus on this

In the above $\nu^* = \dfrac{\mathrm{MS_B} + (b-1)\,\mathrm{MS_{Error}}}{\frac{\mathrm{MS_B}^2}{(b-1)} + \frac{(b-1)^2\,\mathrm{MS_{Error}}^2}{(a-1)(b-1)}}$ à la Satterthwaite.

can ignore

Nitrogen fertilization

Compute a confidence interval for the difference the mean nitrate content between timing schedules one and two (one minus two).

```
y1. <- mean(ntr[trt=="1"])
y2. <- mean(ntr[trt=="2"])
alpha <- 0.05
tval <- qt(1-alpha/2,(a-1)*(b-1))
me <- tval * sqrt(MSE) * sqrt(2/b)
lo <- y1. - y2. - me
up <- y1. - y2. + me
```

We obtain the confidence interval $[-9.799, -1.711]$.

Nitrogen fertilization

Unadjusted CIs for all pairwise differences in means with `ls_means()` from R package `lmerTest`:

```
1 ls_means(lmer_out, pairwise = TRUE)
```

```
Least Squares Means table:

           Estimate Std. Error df t value   lower   upper Pr(>|t|)
trt1 - trt2   -5.755      1.897 15   -3.03  -9.799  -1.711  0.00839 **
trt1 - trt3   -8.492      1.897 15   -4.48 -12.537  -4.448  0.00044 ***
trt1 - trt4   -2.337      1.897 15   -1.23  -6.382   1.707  0.23694
trt1 - trt5   -1.232      1.897 15   -0.65  -5.277   2.812  0.52580
trt1 - trt6   -4.947      1.897 15   -2.61  -8.992  -0.903  0.01980 *
trt2 - trt3   -2.738      1.897 15   -1.44  -6.782   1.307  0.16965
trt2 - trt4    3.418      1.897 15    1.80  -0.627   7.462  0.09182 .
trt2 - trt5    4.522      1.897 15    2.38   0.478   8.567  0.03080 *
trt2 - trt6    0.808      1.897 15    0.43  -3.237   4.852  0.67646
trt3 - trt4    6.155      1.897 15    3.24   2.111  10.199  0.00545 **
trt3 - trt5    7.260      1.897 15    3.83   3.216  11.304  0.00165 **
trt3 - trt6    3.545      1.897 15    1.87  -0.499   7.589  0.08138 .
trt4 - trt5    1.105      1.897 15    0.58  -2.939   5.149  0.56897
trt4 - trt6   -2.610      1.897 15   -1.38  -6.654   1.434  0.18916
trt5 - trt6   -3.715      1.897 15   -1.96  -7.759   0.329  0.06911 .
---
```

$$\binom{6}{2} = \frac{6!}{2!(6-2)!} = \frac{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{6 \cdot 5}{2} = 15$$

$$\binom{n}{2} = \frac{n(n-1)}{2}$$

Can always use "Bonferroni" adjustment if Tukey's or Dunnett's are not available.

Bonferroni: Divide α by the # of CIs

Nitrogen fertilization

Unadjusted CIs for all means with `ls_means()` from R package `lmerTest`:

```
1 ls_means(lmer_out)
```

```
Least Squares Means table:

     Estimate Std. Error  df t value lower upper Pr(>|t|)
trt1    38.28       2.06 6.8    18.6 33.38 43.18  4.5e-07 ***
trt2    44.03       2.06 6.8    21.4 39.13 48.93  1.8e-07 ***
trt3    46.77       2.06 6.8    22.7 41.87 51.67  1.2e-07 ***
trt4    40.62       2.06 6.8    19.7 35.72 45.51  3.0e-07 ***
trt5    39.51       2.06 6.8    19.2 34.61 44.41  3.6e-07 ***
trt6    43.23       2.06 6.8    21.0 38.33 48.12  2.0e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

  Confidence level: 95%
  Degrees of freedom method: Satterthwaite
```

Multiple comparisons in the RCBD

- Tukey's for comparing all pairs of means among $\mu_1, \ldots, \mu_a$:

$$\bar{Y}_{i\cdot} - \bar{Y}_{i'\cdot} \pm q_{a,(a-1)(b-1),\alpha} \sqrt{\mathrm{MS}_{\mathrm{Error}}} \sqrt{\frac{1}{b}}, \quad 1 \leq i < i' \leq a.$$

- Dunnett's for comparing $\mu_2, \ldots, \mu_a$ to a baseline μ_1:

$$\bar{Y}_{i\cdot} - \bar{Y}_{1\cdot} \pm d_{a,(a-1)(b-1),\alpha} \sqrt{\mathrm{MS}_{\mathrm{Error}}} \sqrt{\frac{2}{b}}, \quad i = 2, \ldots, a.$$

Nitrogen fertilization

Compare all pairs of fertilizers with Tukey's CIs for mean differences.

```r
alpha <- 0.05
a <- 6
b <- 4
se <- sqrt(MSE) * sqrt(2/b) # have compute MSE previously
me <- qtukey(1-alpha,a,(a-1)*(b-1)) / sqrt(2) * se
ntr_means <- aggregate(ntr, by = list(trt), mean)$x

CIs <- matrix(NA,choose(a,2),2)
comp <- numeric(choose(a,2))

k <- 1
for(i in 1:(a-1))
  for(j in (i+1):a){ # double loop takes us through all pairs

    dij <- ntr_means[i] - ntr_means[j]
    CIs[k,] <- c(dij - me, dij + me)
    comp[k] <- paste(i,"-",j)
    k <- k + 1

  }

colnames(CIs) <- c("lower","upper")
rownames(CIs) <- comp
```

Nitrogen fertilization

```
1 round(CIs,3)
```

```
       lower upper
1 - 2 -11.92  0.41
1 - 3 -14.66 -2.33
1 - 4  -8.50  3.83
1 - 5  -7.40  4.93
1 - 6 -11.11  1.22
2 - 3  -8.90  3.43
2 - 4  -2.75  9.58
2 - 5  -1.64 10.69
2 - 6  -5.36  6.97
3 - 4  -0.01 12.32
3 - 5   1.09 13.43
3 - 6  -2.62  9.71
4 - 5  -5.06  7.27
4 - 6  -8.78  3.56
5 - 6  -9.88  2.45
```

Nitrogen fertilization

Compare all fertilizers to fertilizer 1 using Dunnett's method.

```r
alpha <- 0.05
a <- 6
b <- 4
me <- 2.82 * sqrt(MSE) * sqrt(2/b) # value 2.82 from Dunnett's table
ntr_means <- aggregate(ntr, by = list(trt), mean)$x

CIs <- matrix(NA,a-1,2)

k <- 1
for(i in 2:a){

  di <- ntr_means[i] - ntr_means[1]
  CIs[k,] <- c(di - me,di + me)
  k <- k + 1

}

colnames(CIs) <- c("lower","upper")
rownames(CIs) <- paste(2:a,"- 1")
```

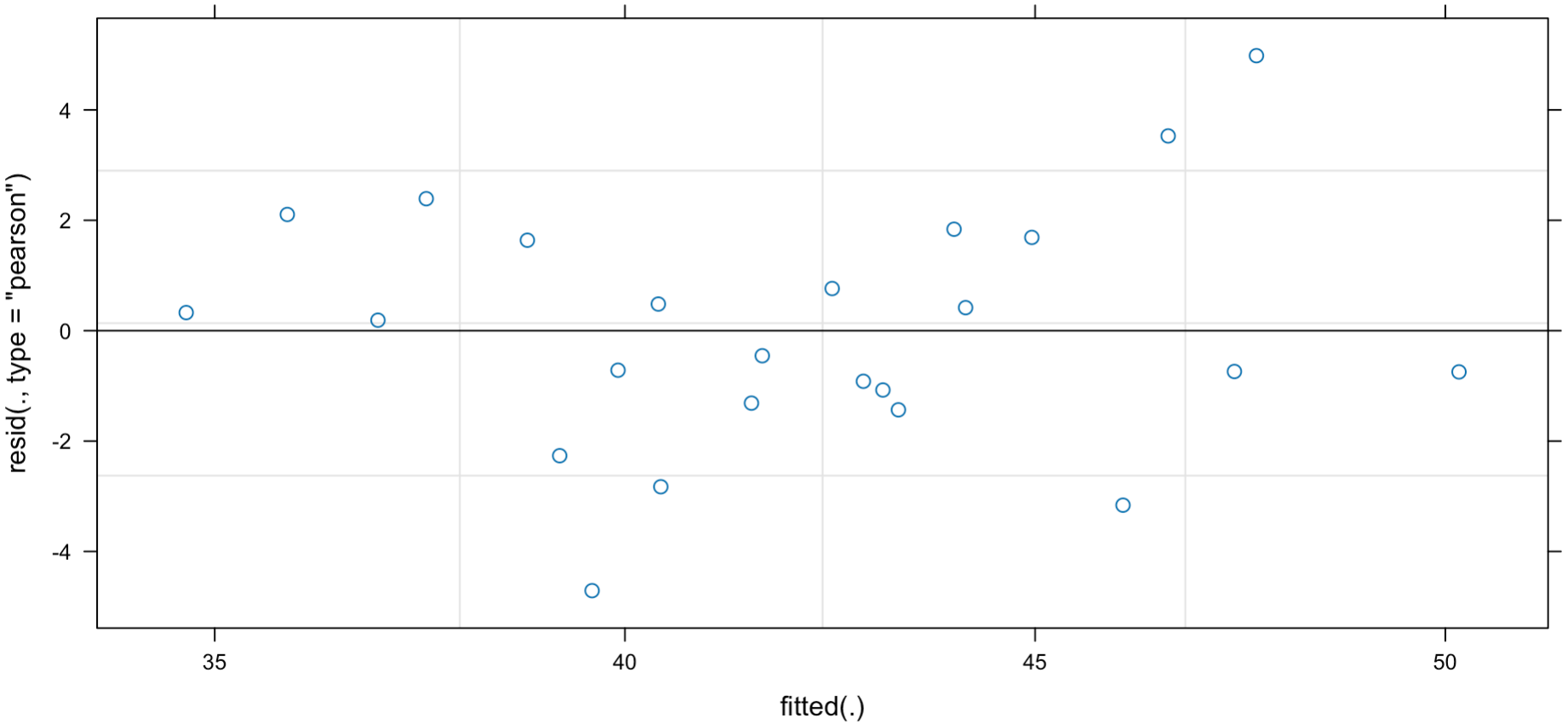
Nitrogen fertilization

```
round(CIs,3)
```

```
       lower upper
2 - 1  0.404 11.11
3 - 1  3.142 13.84
4 - 1 -3.013  7.69
5 - 1 -4.118  6.58
6 - 1 -0.403 10.30
```

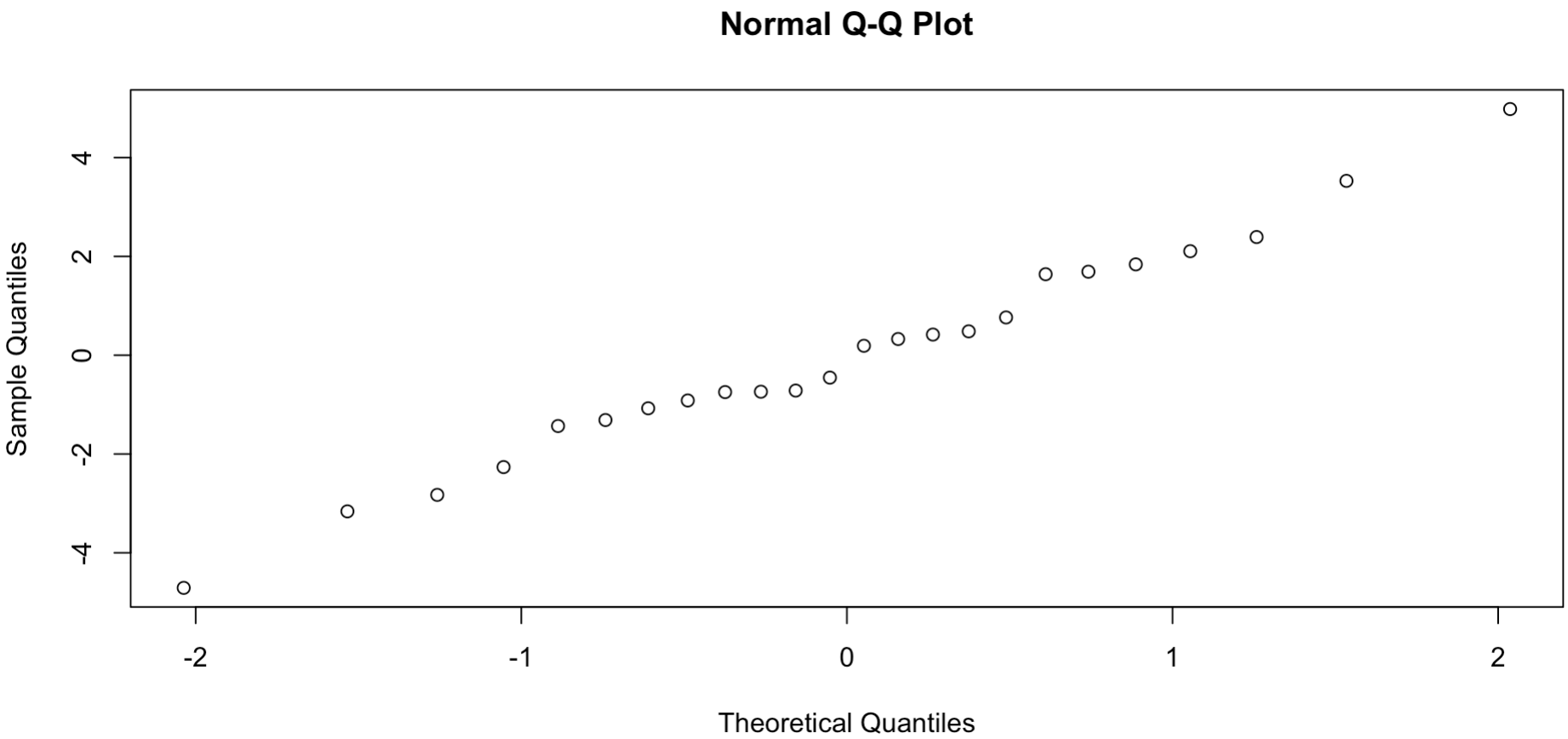
Nitrogen fertilization

```
1 plot(lmer_out)
```



Nitrogen fertilization

```
yhat <- predict(lmer_out)
ehat <- ntr - yhat
qqnorm(ehat)
```



Ignoring the blocks in the nitrogen data

If we ignore the blocks, the design looks like a one-way ANOVA:

Treatment	**1**	**2**	**3**	**4**	**5**	**6**
	34.98	40.89	42.07	37.18	37.99	34.89
	41.22	46.69	49.42	45.85	41.99	50.15
	36.94	46.65	52.68	40.23	37.61	44.57
	39.97	41.90	43.29	39.20	40.45	43.29

Suppose we fit $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$, $i = 1, 2, 3, 4, 5, 6$, $j = 1, 2, 3, 4$.

Nitrogen fertilization

We lose power to detect a treatment effect!

```
1 anova(lm(ntr ~ trt))
```

```
Analysis of Variance Table

Response: ntr
          Df Sum Sq Mean Sq F value Pr(>F)
trt        5    201    40.3    2.38   0.08 .
Residuals 18    305    16.9
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

End of Exam II material

Randomized complete block design

$$Y_{ij} = \mu + \tau_i + B_j + \varepsilon_{ij}$$

$i = 1, \ldots, a$

$j = 1, \ldots, b$

R.C.B.D. with replication

sex: τ_i — classroom: B_j — sex x classroom: $(\tau B)_{ij}$ — noise: ε_{ijk}

$$Y_{ijk} = \mu + \tau_i + B_j + (\tau B)_{ij} + \varepsilon_{ijk}$$

$i = 1, \ldots, a$ level of fixed factor

$j = 1, \ldots, b$ block

$k = 1, \ldots, n$ E.U.s within each block x factor level combination

Cabbage counts versus nitrogen, Kuehl (2000)

Heads of cabbage in subplots of two field plots.

Nitrogen	Block 1		Block 2	
0	104	114	109	124
50	134	130	154	164
100	146	142	152	156
150	147	160	160	163
200	133	146	156	161

Source: Dr. W.D. Pew, Department of Plant Sciences, University of Arizona.

Subplots in each plot randomly assigned to nitrogen levels.

$$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$$

$i = 1,2,3,4,5$ (Nitrogen)

$j = 1,2$ (block)

$k = 1,2$ $\quad n = 2$

Cabbage counts

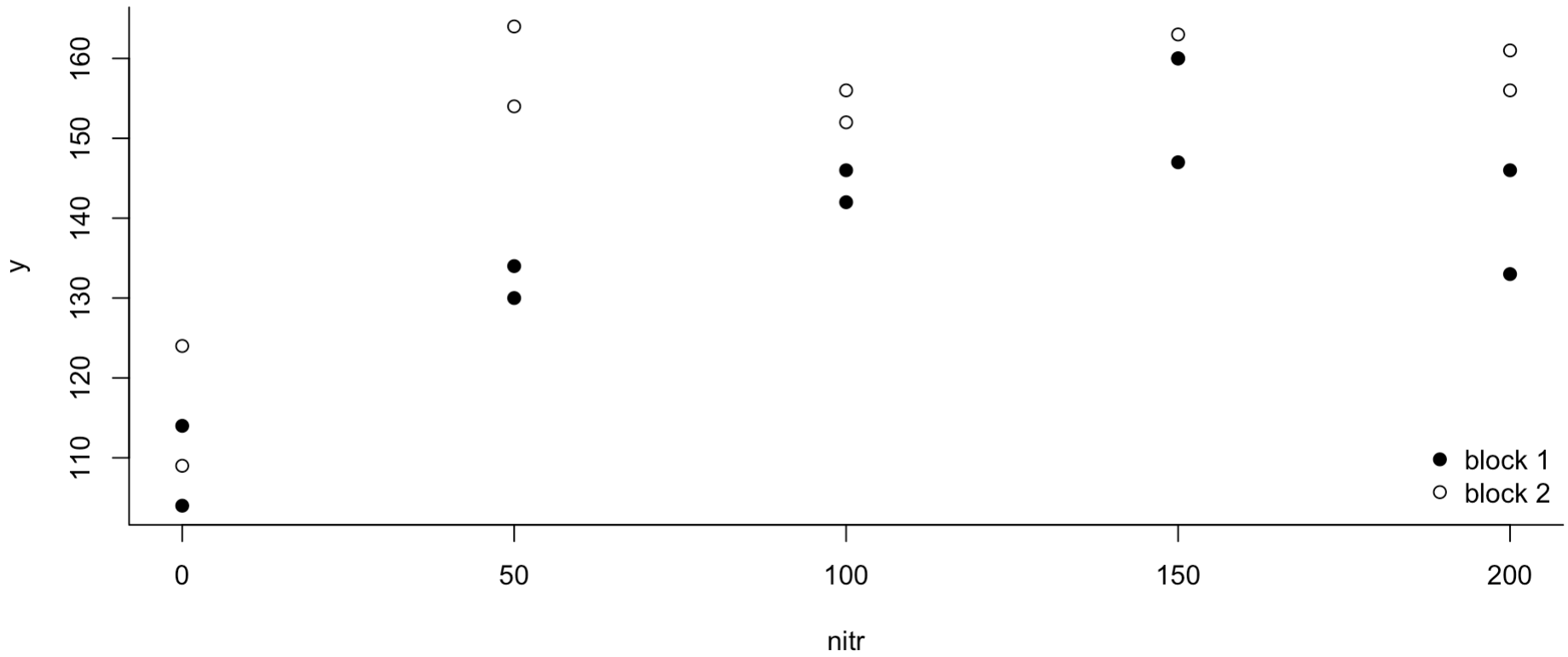
```
y <- c(104,114,109,124,
       134,130,154,164,
       146,142,152,156,
       147,160,160,163,
       133,146,156,161)

nitr <- as.factor(c(0,0,0,0,
                    50,50,50,50,
                    100,100,100,100,
                    150,150,150,150,
                    200,200,200,200))

blk <- as.factor(c(1,1,2,2,
                   1,1,2,2,
                   1,1,2,2,
                   1,1,2,2,
                   1,1,2,2))
```

Cabbage counts

```r
ylims <- range(y)
par(bty = "l")
stripchart(y~nitr, subset=blk==1, vertical=TRUE, pch=19, ylim=ylims,xlab="nitr")
stripchart(y~nitr, subset=blk==2, vertical=TRUE, pch=1, add=TRUE)
legend("bottomright",legend = c("block 1", "block 2"),pch = c(19, 1),bty="n")
```



Randomized CBD with replication

$$Y_{ijk} = \mu + \tau_i + B_j + (\tau B)_{ij} + \varepsilon_{ijk}$$

- τ_i: trt effect
- B_j: random block effect
- $(\tau B)_{ij}$: random interaction effect
- ε_{ijk}: noise

$i = 1, \ldots, a$ treatment level

$j = 1, \ldots, b$ block

$k = 1, \ldots, n$ assume **balanced**

Assume

$$B_j \overset{ind}{\sim} N(0, \sigma_B^2)$$

$$(\tau B)_{ij} \overset{ind}{\sim} N(0, \sigma_{AB}^2)$$

$$\varepsilon_{ijk} \overset{ind}{\sim} N(0, \sigma_\varepsilon^2)$$

Hypotheses of interest:

(1) $H_0: \tau_1 = \cdots = \tau_a$ (treatment makes no difference)

$H_1:$ at least one difference

$F_A =$

(2) $H_0: \sigma_B^2 = 0$

$H_1: \sigma_B^2 > 0$

$F_B =$

(3) $H_0: \sigma_{AB}^2 = 0$

$H_1: \sigma_{AB}^2 > 0$

$F_{AB} =$

Treatment effects model for RCBD with replication

Assume

$$Y_{ijk} + \mu + \tau_i + B_j + (\tau B)_{ij} + \varepsilon_{ijk}$$

for $i = 1, \ldots, a$, $j = 1, \ldots, b$, and $k = 1, \ldots, n_{ij}$, where

- Y_{ijk} is response of EU k in block j receiving treatment i.
- the τ_i are the fixed effects of the treatment.
- the B_j are independent $\mathcal{N}(0, \sigma_B^2)$ random block effects.
- the $(\tau B)_{ij}$ are indep. $\mathcal{N}(0, \sigma_{AB}^2)$ random interaction effects.
- the ε_{ijk} are independent $\mathcal{N}(0, \sigma_\varepsilon^2)$ error terms.
- μ is an overall or baseline mean.

Define the cell means as $\mu_i = \mu + \tau_i, \quad i = 1, \ldots, a.$

Assume for now a balanced design: $n_{ij} = n$ for all i, j.

Sums of squares for RCBD with replication

Sum of squares	Symbol	Formula
Total	$\mathrm{SS}_{\mathrm{Tot}}$	$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk}-\bar{Y}_{...})^2$
Treatment	SS_{A}	$nb\sum_{i=1}^{a}(\bar{Y}_{i..}-\bar{Y}_{...})^2$
Block	SS_{B}	$na\sum_{j=1}^{b}(\bar{Y}_{.j.}-\bar{Y}_{...})^2$
Block × Treatment	$\mathrm{SS}_{\mathrm{AB}}$	$n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{Y}_{ij.}-(\bar{Y}_{i..}+\bar{Y}_{.j.}-\bar{Y}_{...}))^2$
Error	$\mathrm{SS}_{\mathrm{Error}}$	$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk}-\bar{Y}_{ij.})^2$

- Can make the decomposition $\mathrm{SS}_{\mathrm{Tot}} = \mathrm{SS}_{\mathrm{A}} + \mathrm{SS}_{\mathrm{B}} + \mathrm{SS}_{\mathrm{AB}} + \mathrm{SS}_{\mathrm{Error}}$.

ANOVA table for RCBD with replication

Source	Df	SS	MS	F value
A	$a-1$	$\mathrm{SS_A}$	$\mathrm{MS_A}$	$F_\mathrm{A} = \mathrm{MS_A} / \mathrm{MS_{AB}}$
B	$b-1$	$\mathrm{SS_B}$	$\mathrm{MS_B}$	$F_\mathrm{B} = \mathrm{MS_B} / \mathrm{MS_{AB}}$
AB	$(a-1)(b-1)$	$\mathrm{SS_{AB}}$	$\mathrm{MS_{AB}}$	$F_\mathrm{AB} = \mathrm{MS_{AB}} / \mathrm{MS_{Error}}$
Error	$ab(n-1)$	$\mathrm{SS_{Error}}$	$\mathrm{MS_{Error}}$	
Total	$abn-1$	$\mathrm{SS_{Tot}}$		

interaction MS in denominator??

Just as in two-way factorial ANOVA

$\tau_1 = \cdots = \tau_a$

1. Reject $H_0: \mu_1 = \cdots = \mu_a$ if $F_\mathrm{A} > F_{a-1,(a-1)(b-1),\alpha}$.
2. Reject $H_0: \sigma_B^2 = 0$ if $F_\mathrm{B} > F_{b-1,(a-1)(b-1),\alpha}$.
3. Reject $H_0: \sigma_{AB}^2 = 0$ if $F_\mathrm{AB} > F_{(a-1)(b-1),ab(n-1),\alpha}$.

$$\frac{\chi^2_{df_1}/df_1}{\chi^2_{df_2}/df_2} \sim F_{df_1, df_2}$$

Expected MS in RCBD with replication

$$\frac{MS_A}{MS_{AB}} \sim F_{a-1,(a-1)(b-1)}$$

Source	Df	Expected mean square
A	$a-1$	$nb\theta_A^2 + n\sigma^2_{AB} + \sigma^2_\varepsilon$
B	$b-1$	$na\sigma^2_B + n\sigma^2_{AB} + \sigma^2_\varepsilon$
AB	$(a-1)(b-1)$	$n\sigma^2_{AB} + \sigma^2_\varepsilon$
Error	$ab(n-1)$	σ^2_ε

$H_0: \tau_1 = \cdots = \tau_a$

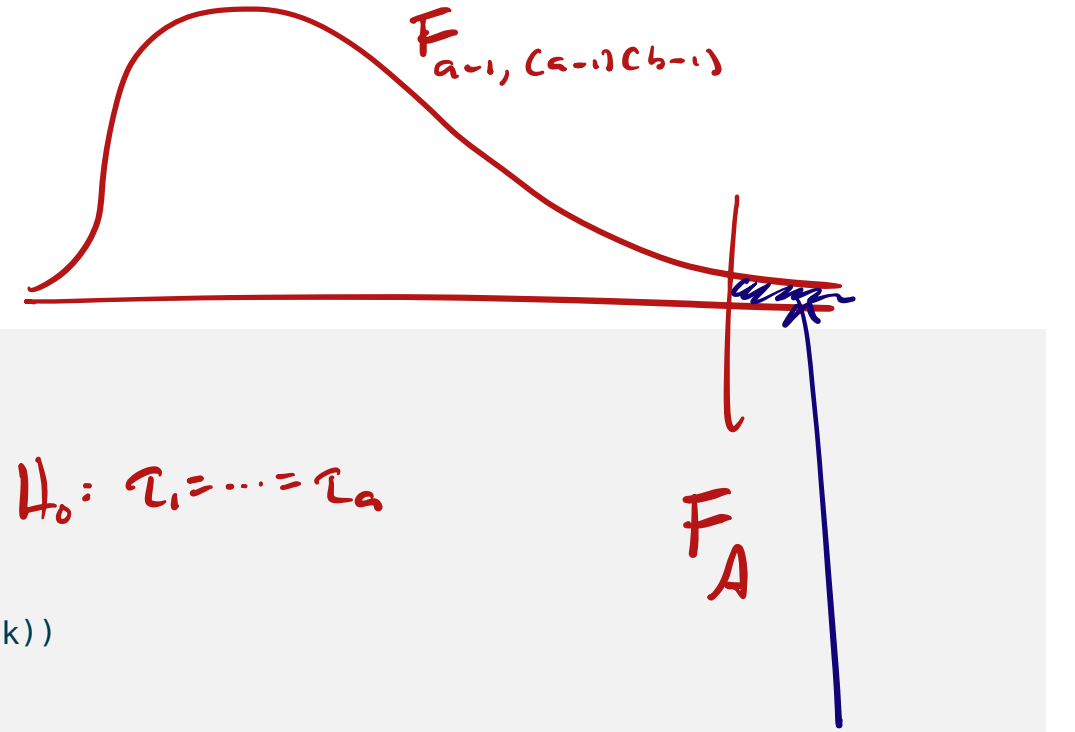
$H_0: \sigma^2_B = 0$

$$\frac{MS_B}{MS_{AB}} \sim F_{b-1,(a-1)(b-1)}$$

In the above $\theta^2_A = (a-1)^{-1}\sum_{i=1}^{a}(\mu_i - \bar{\mu}_\cdot)^2$.

Check expected MS values under each H_0 on previous slide.

Cabbage counts




```
a <- nlevels(nitr)
b <- nlevels(blk)
n <- 2

y... <- predict(lm(y ~ 1))
yi.. <- predict(lm(y ~ nitr))
y.j. <- predict(lm(y ~ blk))
yij. <- predict(lm(y ~ nitr + blk + nitr:blk))

SST <- sum((y - y...)**2)
SSA <- sum((yi.. - y...)**2)
SSB <- sum((y.j. - y...)**2)
SSAB <- sum((yij. - (yi.. + y.j. - y...))**2)
SSE <- sum((y - yij.)**2)

MSA <- SSA / (a-1)
MSB <- SSB / (b-1)
MSAB <- SSAB / ((a-1)*(b-1))
MSE <- SSE / (a*b*(n-1))

FA <- MSA / MSAB
FB <- MSB / MSAB
FAB <- MSAB / MSE
```

Cabbage counts

Source	Df	SS	MS	F value	p value
A	4	4813.00	1203.25	16.7234	0.0092
B	1	1022.45	1022.45	14.2106	0.0196
AB	4	287.80	71.95	1.7030	0.2253
Error	10	422.50	42.25		
Total	19	6545.75			

$\alpha = 0.05$

Correct numbers

Reject $H_0: \tau_1 = \cdots = \tau_a$

Reject $H_0: \sigma^2_{\tau\beta} = 0$

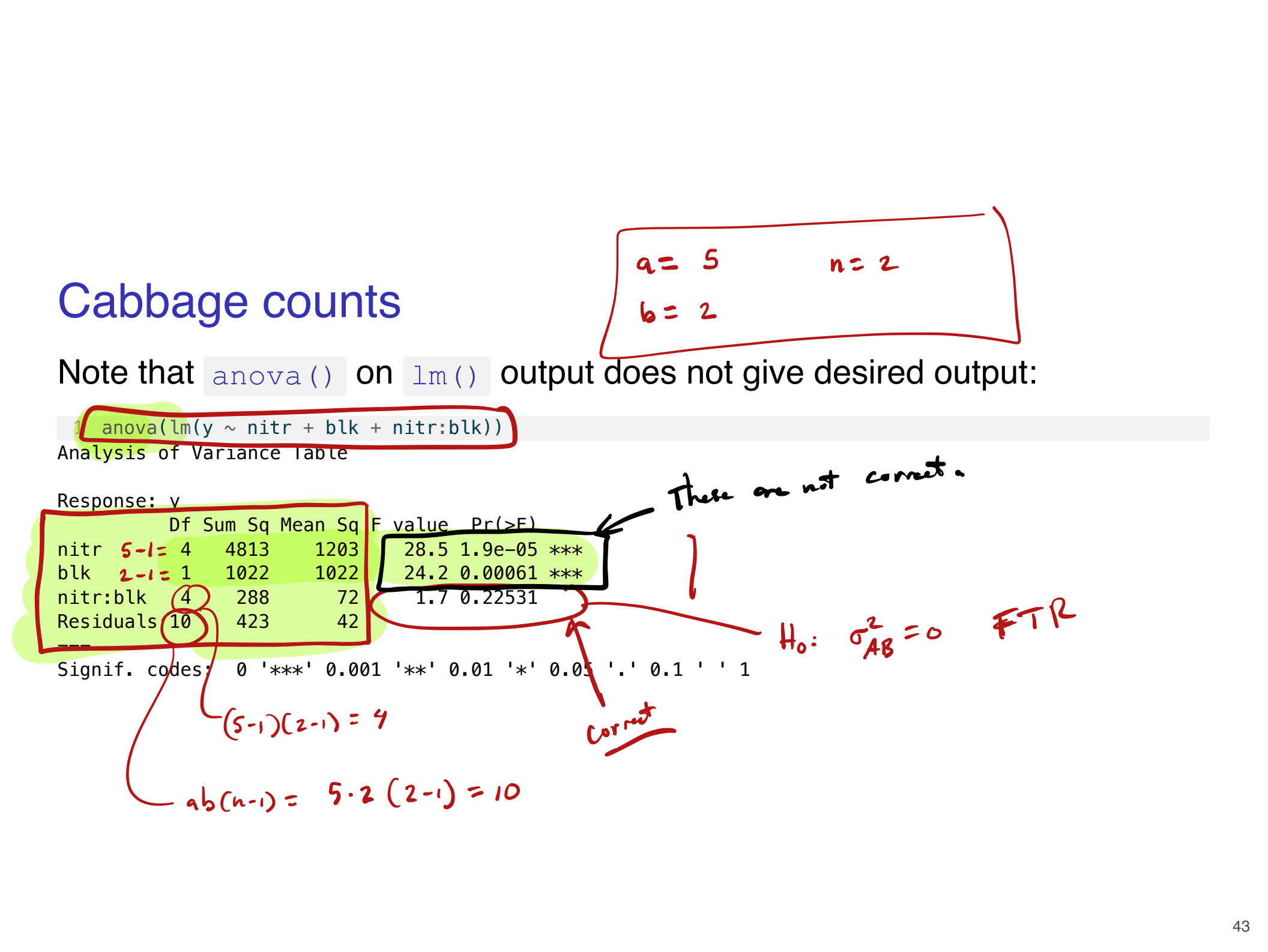
Cabbage counts

Note that `anova()` on `lm()` output does not give desired output:

```
anova(lm(y ~ nitr + blk + nitr:blk))
```

```
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value  Pr(>F)
nitr       4   4813    1203    28.5 1.9e-05 ***
blk        1   1022    1022    24.2 0.00061 ***
nitr:blk   4    288      72     1.7 0.22531
Residuals 10    423      42
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



Method of moments for estimating σ^2_B, σ^2_{AB}, and σ^2_ε

Set the mean squares equal to their expectations and solve:

$$\begin{aligned} \mathrm{MS_B} &\overset{\text{set}}{=} na\sigma^2_B + n\sigma^2_{AB} + \sigma^2_\varepsilon \\ \mathrm{MS_{AB}} &\overset{\text{set}}{=} n\sigma^2_{AB} + \sigma^2_\varepsilon \\ \mathrm{MS_{Error}} &\overset{\text{set}}{=} \sigma^2_\varepsilon \end{aligned}$$

Solve for σ^2_B, σ^2_{AB}, and σ^2_ε:

$$\begin{aligned} \dot{\sigma}^2_B &= \frac{1}{na}(\mathrm{MS_B} - \mathrm{MS_{AB}}) \\ \dot{\sigma}^2_{AB} &= \frac{1}{n}(\mathrm{MS_{AB}} - \mathrm{MS_{Error}}) \\ \dot{\sigma}^2_\varepsilon &= \mathrm{MS_{Error}} \end{aligned}$$

Possible to obtain negative values for $\dot{\sigma}^2_B$ and $\dot{\sigma}^2_{AB}$. Then use REML.

Cabbage counts

Obtain the method of moments estimates for the variance components.

```
sigma_B_sq <- (MSB - MSAB) / (n*a)
sigma_AB_sq <- (MSAB - MSE) / n
sigma_e_sq <- MSE

sigma_B <- sqrt(sigma_B_sq)
sigma_AB <- sqrt(sigma_AB_sq)
sigma_e <- sqrt(sigma_e_sq)
```

We obtain $\dot{\sigma}_B = 9.749$, $\dot{\sigma}_{AB} = 3.854$, and $\dot{\sigma}_\varepsilon = 6.500$.

Cabbage counts

Obtain the REML estimates for the variance components.

```
1 library(lmerTest) # first time run install.packages("lmerTest")
2 lmer_out <- lmer(y ~ nitr + (1|blk) + (1|nitr:blk))
3 lmer_out
```

```
Linear mixed model fit by REML ['lmerModLmerTest']
Formula: y ~ nitr + (1 | blk) + (1 | nitr:blk)
REML criterion at convergence: 111
Random effects:
 Groups   Name        Std.Dev.
 nitr:blk (Intercept) 3.85
 blk      (Intercept) 9.75
 Residual             6.50
Number of obs: 20, groups:  nitr:blk, 10; blk, 2
Fixed Effects:
(Intercept)       nitr50      nitr100      nitr150      nitr200
      112.8         32.7         36.2         44.7         36.2
```

$\hat{\sigma}_{AB}$ $\hat{\sigma}_{B}$ $\hat{\sigma}_{\varepsilon}$

CIs (unadjusted) for diffs in means in RCBD with rep.

- A $(1-\alpha)100\%$ confidence interval for $\mu_i - \mu_{i'}$ is

$$\bar{Y}_{i\cdot\cdot} - \bar{Y}_{i'\cdot\cdot} \pm t_{(a-1)(b-1),\alpha/2}\sqrt{\mathrm{MS}_{\mathrm{AB}}}\sqrt{\frac{2}{nb}}$$

  for $1 \leq i < i' \leq a$.

- If $(a-1)(b-1) = 1$, this interval will be frightfully wide.
- In this case one should drop the interaction term from the model!

CI for difference in means for cabbage count data

Compare the means of the nitrogen level 0 and 50 groups:

```
y0.. <- mean(y[nitr == "0"])
y50.. <- mean(y[nitr == "50"])
alpha <- 0.05
tval <- qt(1 - alpha/2,(a-1)*(b-1))
me <- tval * sqrt(MSAB) * sqrt(2/(n*b))
lo <- y0.. - y50.. - me
up <- y0.. - y50.. + me
```

We obtain the confidence interval $[-49.403, -16.097]$.

CIs (unadjusted) for differences in means in RCBD with replication

Can use `ls_means()` on output of `lmer()`:

```
1 ls_means(lmer_out,pairwise= T)
```

```
Least Squares Means table:

                  Estimate Std. Error df t value     lower     upper Pr(>|t|)
nitr0 - nitr50   -3.27e+01   6.00e+00  4   -5.46 -4.94e+01 -1.61e+01   0.0055
nitr0 - nitr100  -3.62e+01   6.00e+00  4   -6.04 -5.29e+01 -1.96e+01   0.0038
nitr0 - nitr150  -4.47e+01   6.00e+00  4   -7.46 -6.14e+01 -2.81e+01   0.0017
nitr0 - nitr200  -3.62e+01   6.00e+00  4   -6.04 -5.29e+01 -1.96e+01   0.0038
nitr50 - nitr100 -3.50e+00   6.00e+00  4   -0.58 -2.02e+01  1.32e+01   0.5909
nitr50 - nitr150 -1.20e+01   6.00e+00  4   -2.00 -2.87e+01  4.65e+00   0.1160
nitr50 - nitr200 -3.50e+00   6.00e+00  4   -0.58 -2.02e+01  1.32e+01   0.5909
nitr100 - nitr150 -8.50e+00   6.00e+00  4   -1.42 -2.52e+01  8.15e+00   0.2294
nitr100 - nitr200 -7.11e-15   6.00e+00  4    0.00 -1.67e+01  1.67e+01   1.0000
nitr150 - nitr200  8.50e+00   6.00e+00  4    1.42 -8.15e+00  2.52e+01   0.2294

nitr0 - nitr50    **
nitr0 - nitr100   **
nitr0 - nitr150   **
nitr0 - nitr200   **
nitr50 - nitr100
```

[-49.4, -16.1]

These CIs are *not* adjusted to achieve a familywise coverage probability!

CIs (unadjusted) for all means in RCBD with replication

Can use `ls_means()` on output of `lmer()`:

```
1 ls_means(lmer_out)
```

```
Least Squares Means table:

        Estimate Std. Error  df t value  lower  upper Pr(>|t|)
nitr0     112.75       8.09 1.6    13.9  68.46 157.04   0.0114 *
nitr50    145.50       8.09 1.6    18.0 101.21 189.79   0.0076 **
nitr100   149.00       8.09 1.6    18.4 104.71 193.29   0.0073 **
nitr150   157.50       8.09 1.6    19.5 113.21 201.79   0.0067 **
nitr200   149.00       8.09 1.6    18.4 104.71 193.29   0.0073 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

  Confidence level: 95%
  Degrees of freedom method: Satterthwaite
```

These CIs are *not* adjusted to achieve a familywise coverage probability!

These CIs are *not* adjusted to achieve a familywise coverage probability!

RCBD with replication *without* an interaction term

If we omit the interaction term in the RCBD with replication, we have

$$Y_{ijk} + \mu + \tau_i + B_j + \varepsilon_{ijk}$$

for $i = 1, \ldots, a$, $j = 1, \ldots, b$, and $k = 1, \ldots, n_{ij}$, where

- Y_{ijk} is response of EU k in block j receiving treatment i.
- the τ_i are the fixed effects of the treatment.
- the B_j are independent $\mathcal{N}(0, \sigma_B^2)$ random block effects.
- the ε_{ijk} are independent $\mathcal{N}(0, \sigma_\varepsilon^2)$ error terms.
- μ is an overall or baseline mean.

Assume for now a balanced design: $n_{ij} = n$ for all i, j.

Use this model if $a = b = 2$, under which $(a-1)(b-1) = 1$.

Sums of squares for RCBD with rep., no interaction

SS	Symbol	Formula
Total	$\mathrm{SS}_{\mathrm{Tot}}$	$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk}-\bar{Y}_{...})^2$
Treatment	SS_{A}	$nb\sum_{i=1}^{a}(\bar{Y}_{i..}-\bar{Y}_{...})^2$
Block	SS_{B}	$na\sum_{j=1}^{b}(\bar{Y}_{.j.}-\bar{Y}_{...})^2$
Error	$\mathrm{SS}_{\mathrm{Error}}$	$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk}-(\bar{Y}_{i..}+\bar{Y}_{.j.}-\bar{Y}_{...}))^2$

- Can make the decomposition $\mathrm{SS}_{\mathrm{Tot}} = \mathrm{SS}_{\mathrm{A}} + \mathrm{SS}_{\mathrm{B}} + \mathrm{SS}_{\mathrm{Error}}$.

ANOVA table for RCBD with rep., omitting interaction

Source	Df	SS	MS	F value
A	$a-1$	SS_{A}	MS_{A}	$F_{\mathrm{A}} = \mathrm{MS}_{\mathrm{A}} \,/\, \mathrm{MS}_{\mathrm{Error}}$
B	$b-1$	SS_{B}	MS_{B}	$F_{\mathrm{B}} = \mathrm{MS}_{\mathrm{B}} \,/\, \mathrm{MS}_{\mathrm{Error}}$
Error	$abn-a-b+1$	$\mathrm{SS}_{\mathrm{Error}}$	$\mathrm{MS}_{\mathrm{Error}}$	
Total	$abn-1$	$\mathrm{SS}_{\mathrm{Tot}}$		

1. Reject H_0: $\mu_1 = \cdots = \mu_a$ if $F_{\mathrm{A}} > F_{a-1,abn-a-b+1,\alpha}$.
2. Reject H_0: $\sigma_B^2 = 0$ if $F_{\mathrm{B}} > F_{b-1,abn-a-b+1,\alpha}$.

Expected MS in RCBD with rep., omitting interaction

Source	Df	Expected mean square
A	$a-1$	$nb\theta_A^2+\sigma_\varepsilon^2$
B	$b-1$	$na\sigma_B^2+\sigma_\varepsilon^2$
Error	$abn-a-b+1$	σ_ε^2

In the above $\theta_A^2=(a-1)^{-1}\sum_{i=1}^{a}(\mu_i-\bar{\mu}_\cdot)^2$.

Cabbage counts

ANOVA table, omitting the interaction:

```
a <- nlevels(nitr)
b <- nlevels(blk)
n <- 2

y... <- predict(lm(y ~ 1))
yi.. <- predict(lm(y ~ nitr))
y.j. <- predict(lm(y ~ blk))

SST <- sum((y - y...)**2)
SSA <- sum((yi.. - y...)**2)
SSB <- sum((y.j. - y...)**2)
SSE <- sum((y - (yi.. + y.j. - y...))**2)

MSA <- SSA / (a-1)
MSB <- SSB / (b-1)
MSE <- SSE / (a*b*n - a - b + 1)

FA <- MSA / MSE
FB <- MSB / MSE

pA <- 1 - pf(FA,a-1,a*b*n - a - b + 1)
pB <- 1 - pf(FB,b-1,a*b*n - a - b + 1)
```

Cabbage counts

Source	Df	SS	MS	F value	p value
A	4	4813.00	1203.25	23.7160	0.0000
B	1	1022.45	1022.45	20.1525	0.0005
Error	14	710.30	50.74		
Total	19	6545.75			

Cabbage counts

With the interaction term omitting, `anova()` on `lm()` output is correct:

```
1 anova(lm(y ~ nitr + blk))
```

```
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value  Pr(>F)
nitr       4   4813    1203    23.7 4.1e-06 ***
blk        1   1022    1022    20.1 0.00051 ***
Residuals 14    710      51
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

MoM for variance components when omitting interaction

Set the mean squares equal to their expectations and solve:

$$\begin{aligned} \mathrm{MS_B} &\overset{\text{set}}{=} na\sigma_B^2 + \sigma_\varepsilon^2 \\ \mathrm{MS_{Error}} &\overset{\text{set}}{=} \sigma_\varepsilon^2 \end{aligned}$$

Solve for σ_B^2 and σ_ε^2:

$$\begin{aligned} \dot{\sigma}_B^2 &= \frac{1}{na}(\mathrm{MS_B} - \mathrm{MS_{Error}}) \\ \dot{\sigma}_\varepsilon^2 &= \mathrm{MS_{Error}} \end{aligned}$$

Possible to obtain a negative value for $\dot{\sigma}_B^2$. In this case use REML.

Cabbage counts

MoM estimates when omitting interaction:

```
sigma_B_sq <- (MSB - MSE) / (n*a)
sigma_e_sq <- MSE

sigma_B <- sqrt(sigma_B_sq)
sigma_e <- sqrt(sigma_e_sq)
```

We obtain $\dot{\sigma}_B = 9.858$ and $\dot{\sigma}_\varepsilon = 7.123$.

Cabbage counts

REML estimates when omitting interaction:

```
1 library(lmerTest) # first time run install.packages("lmerTest")
2 lmer_out <- lmer(y ~ nitr + (1|blk))
3 lmer_out
```

```
Linear mixed model fit by REML ['lmerModLmerTest']
Formula: y ~ nitr + (1 | blk)
REML criterion at convergence: 111
Random effects:
 Groups   Name        Std.Dev.
 blk      (Intercept) 9.86
 Residual             7.12
Number of obs: 20, groups:  blk, 2
Fixed Effects:
(Intercept)       nitr50      nitr100      nitr150      nitr200
      112.8         32.8         36.3         44.8         36.3
```

CIs (unadjusted) for diffs in means in RCBD with rep. without interaction

- A $(1-\alpha)100\%$ confidence interval for $\mu_i - \mu_{i'}$ is

$$\bar{Y}_{i\cdot\cdot} - \bar{Y}_{i'\cdot\cdot} \pm t_{abn-a-b+1,\alpha/2}\sqrt{\mathrm{MS}_{\mathrm{Error}}}\sqrt{\frac{2}{nb}}$$

for $1 \leq i < i' \leq a$.

Cabbage counts

Compare the means of the nitrogen level 50 and 0 groups:

```
y50.. <- mean(y[nitr == "50"])
y0.. <- mean(y[nitr == "0"])
alpha <- 0.05
tval <- qt(1 - alpha/2,a*b*n - a - b + 1)
me <- tval * sqrt(MSE) * sqrt(2/(n*b))
lo <- y50.. - y0.. - me
up <- y50.. - y0.. + me
```

We obtain the confidence interval $[21.947, 43.553]$.

Cabbage counts

Can use `ls_means()` as below:

```
1 ls_means(lmer_out,pairwise=T)
```

```
Least Squares Means table:

                 Estimate Std. Error df t value  lower  upper Pr(>|t|)
nitr0 - nitr50      -32.75       5.04 14   -6.50 -43.55 -21.95  1.4e-05 ***
nitr0 - nitr100     -36.25       5.04 14   -7.20 -47.05 -25.45  4.6e-06 ***
nitr0 - nitr150     -44.75       5.04 14   -8.88 -55.55 -33.95  4.0e-07 ***
nitr0 - nitr200     -36.25       5.04 14   -7.20 -47.05 -25.45  4.6e-06 ***
nitr50 - nitr100     -3.50       5.04 14   -0.69 -14.30   7.30    0.498
nitr50 - nitr150    -12.00       5.04 14   -2.38 -22.80  -1.20    0.032 *
nitr50 - nitr200     -3.50       5.04 14   -0.69 -14.30   7.30    0.498
nitr100 - nitr150    -8.50       5.04 14   -1.69 -19.30   2.30    0.114
nitr100 - nitr200     0.00       5.04 14    0.00 -10.80  10.80    1.000
nitr150 - nitr200     8.50       5.04 14    1.69  -2.30  19.30    0.114
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

  Confidence level: 95%
  Degrees of freedom method: Satterthwaite
```

- Note that the `ls_means()` function will remove random effects with variance components estimated to be zero. So the CI from `ls_means()` may not match the one computed manually.

Skin response data, Mohr, Wilson, and Freund (2021)

Galvanic skin responses of five subjects under shock and noise stimuli.

		SUBJECT				
Noise	**Shock**	**1**	**2**	**3**	**4**	**5**
40	0.25	3	7	9	4	1
40	0.50	5	11	13	8	3
40	0.75	9	12	14	11	5
40	1.00	6	11	12	7	4
80	0.25	5	10	10	6	3
80	0.50	6	12	15	9	5
80	0.75	18	18	15	13	9
80	1.00	7	15	14	9	7

```
skin <- data.frame(resp = c(3,7,9,4,1,5,11,13,8,3,9,12,14,11,5,6,11,12,7,4,5,
                            10,10,6,3,6,12,15,9,5,18,18,15,13,9,7,15,14,9,7),
                   noise = as.factor(c(rep(40,20),rep(80,20))),
                   shock = as.factor(rep(c(rep(.25,5),rep(.5,5),
                                           rep(.75,5),rep(1,5)),2)),
                   subj = as.factor(rep(1:5,8)))
```

Skin response

```
1 head(skin,n=20)
```

	resp	noise	shock	subj
1	3	40	0.25	1
2	7	40	0.25	2
3	9	40	0.25	3
4	4	40	0.25	4
5	1	40	0.25	5
6	5	40	0.5	1
7	11	40	0.5	2
8	13	40	0.5	3
9	8	40	0.5	4
10	3	40	0.5	5
11	9	40	0.75	1
12	12	40	0.75	2
13	14	40	0.75	3
14	11	40	0.75	4
15	5	40	0.75	5
16	6	40	1	1
17	11	40	1	2
18	12	40	1	3

Treatment effects model for two-way factorial RCBD

Assume

$$Y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + C_k + \varepsilon_{ijk},$$

for $i = 1, \ldots, a$, $j = 1, \ldots, b$, and $k = 1, \ldots, c$, where

- Y_{ijk} is the response in block k under treatment combination $i \times j$.
- μ is an overall or baseline mean.
- the τ_i are fixed effects for factor A.
- the γ_j are fixed effects for factor B.
- the $(\tau\gamma)_{ij}$ are effects for the AB interaction.
- the C_k are independent $\mathcal{N}(0, \sigma_C^2)$ block effects.
- the ε_{ijk} are independent $\mathcal{N}(0, \sigma_\varepsilon^2)$ error terms.

Define the cell means as

$$\mu_{ij} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij}, \quad i = 1, \ldots, a, \quad j = 1, \ldots, b.$$

Sums of squares for the two-way factorial RCBD

Sum of squares	**Symbol**	**Formula**
Total	$\mathrm{SS}_{\mathrm{Tot}}$	$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}(Y_{ijk}-\bar{Y}_{...})^2$
A	SS_{A}	$bc\sum_{i=1}^{a}(\bar{Y}_{i..}-\bar{Y}_{...})^2$
B	SS_{B}	$ac\sum_{j=1}^{b}(\bar{Y}_{.j.}-\bar{Y}_{...})^2$
AB	$\mathrm{SS}_{\mathrm{AB}}$	$c\sum_{i=1}^{a}\sum_{j=1}^{b}(Y_{ij.}-(\bar{Y}_{i..}+\bar{Y}_{.j.}-\bar{Y}_{...}))^2$
C	SS_{C}	$ab\sum_{k=1}^{c}(\bar{Y}_{..k}-\bar{Y}_{...})^2$
Error	$\mathrm{SS}_{\mathrm{Error}}$	$\mathrm{SS}_{\mathrm{Tot}}-(\mathrm{SS}_{\mathrm{A}}+\mathrm{SS}_{\mathrm{B}}+\mathrm{SS}_{\mathrm{AB}}+\mathrm{SS}_{\mathrm{C}})$

ANOVA table for two-way factorial RCBD

Source	Df	SS	MS	F value
A	$a-1$	SS_{A}	MS_{A}	$F_{\mathrm{A}} = \mathrm{MS}_{\mathrm{A}} \,/\, \mathrm{MS}_{\mathrm{Error}}$
B	$b-1$	SS_{B}	MS_{B}	$F_{\mathrm{B}} = \mathrm{MS}_{\mathrm{B}} \,/\, \mathrm{MS}_{\mathrm{Error}}$
AB	$(a-1)(b-1)$	$\mathrm{SS}_{\mathrm{AB}}$	$\mathrm{MS}_{\mathrm{AB}}$	$F_{\mathrm{AB}} = \mathrm{MS}_{\mathrm{B}} \,/\, \mathrm{MS}_{\mathrm{Error}}$
C	$c-1$	SS_{C}	MS_{C}	$F_{\mathrm{C}} = \mathrm{MS}_{\mathrm{C}} \,/\, \mathrm{MS}_{\mathrm{Error}}$
Error	$(ab-1)(c-1)$	$\mathrm{SS}_{\mathrm{Error}}$	$\mathrm{MS}_{\mathrm{Error}}$	
Total	$abc-1$	$\mathrm{SS}_{\mathrm{Tot}}$		

1. Reject H_0: $\tau_1 = \cdots = \tau_a$ if $F_{\mathrm{A}} > F_{a-1,(ab-1)(c-1),\alpha}$.
2. Reject H_0: $\gamma_1 = \cdots = \gamma_b$ if $F_{\mathrm{B}} > F_{b-1,(ab-1)(c-1),\alpha}$.
3. R. H_0: $(\tau\gamma)_{11} = \cdots = (\tau\gamma)_{ab}$ if $F_{\mathrm{AB}} > F_{(a-1)(b-1),(ab-1)(c-1),\alpha}$.
4. Reject H_0: $\sigma_C^2 = 0$ if $F_{\mathrm{C}} > F_{c-1,(ab-1)(c-1),\alpha}$.

Skin response

```r
lm_out <- lm(resp ~ noise + shock + noise:shock + subj, data = skin)
anova(lm_out)
```

```
Analysis of Variance Table

Response: resp
            Df Sum Sq Mean Sq F value  Pr(>F)
noise        1     65    65.0   28.38 1.1e-05 ***
shock        3    219    73.1   31.90 3.6e-09 ***
subj         4    362    90.5   39.48 4.0e-11 ***
noise:shock  3     13     4.2    1.84    0.16
Residuals   28     64     2.3
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Skin response

```r
y <- skin$resp
y... <- predict(lm(resp ~ 1,data = skin))
yi.. <- predict(lm(resp ~ noise,data=skin))
y.j. <- predict(lm(resp ~ shock,data=skin))
y..k <- predict(lm(resp ~ subj,data=skin))
yij. <- predict(lm(resp ~ noise + shock + noise:shock,data=skin))

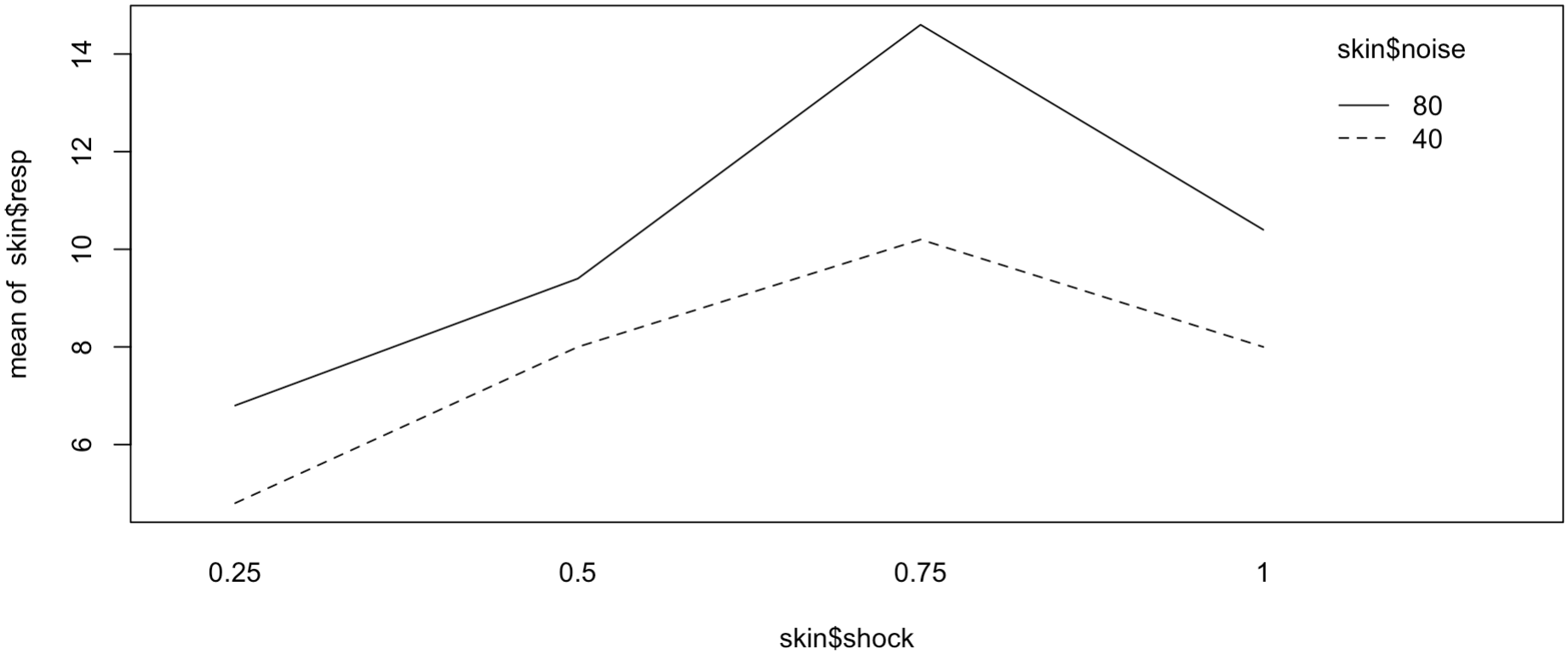
SSA <- sum((yi.. - y...)^2)
SSB <- sum((y.j. - y...)^2)
SSC <- sum((y..k - y...)^2)
SSAB <- sum((yij. - (yi.. + y.j. - y...))^2)
SST <- sum((y - y...)^2)
SSE <- SST - (SSA + SSB + SSC + SSAB)

a <- 2
b <- 4
c <- 5

MSA <- SSA/(a-1)
MSB <- SSB/(b-1)
MSC <- SSC/(c-1)
MSAB <- SSAB/((a-1)*(b-1))
MSE <- SSE/((a*b-1)*(c-1))
```

Skin response

```
1 interaction.plot(skin$shock,skin$noise,skin$resp)
```



Expected mean squares in factorial RCBD

Source	Df	Expected mean square
A	$a-1$	$bc\theta_A^2 + \sigma_\varepsilon^2$
B	$b-1$	$ac\theta_B^2 + \sigma_\varepsilon^2$
AB	$(a-1)(b-1)$	$c\theta_{AB}^2 + \sigma_\varepsilon^2$
C	$c-1$	$ab\sigma_C^2 + \sigma_\varepsilon^2$
Error	$(ab-1)(c-1)$	σ_ε^2

In the above

- $\theta_A^2 = (a-1)^{-1}\sum_{i=1}^{a}(\bar{\mu}_{i.} - \bar{\mu}_{..})^2$
- $\theta_B^2 = (b-1)^{-1}\sum_{j=1}^{b}(\bar{\mu}_{.j} - \bar{\mu}_{..})^2$
- $\theta_{AB}^2 = [(a-1)(b-1)]^{-1}\sum_{i=1}^{a}\sum_{j=1}^{b}(\mu_{ij} - (\bar{\mu}_{i.} + \bar{\mu}_{.j} - \bar{\mu}_{..}))^2$

MoMs for variance components in two-way factorial RCBD

- Equating MS_{C} and $\mathrm{MS}_{\mathrm{Error}}$ with their expectations gives

$$\dot{\sigma}_C^2 = \frac{\mathrm{MS}_{\mathrm{C}} - \mathrm{MS}_{\mathrm{Error}}}{ab} \quad \text{and} \quad \dot{\sigma}_\varepsilon^2 = \mathrm{MS}_{\mathrm{Error}}.$$

- May obtain $\dot{\sigma}_C^2 < 0$, so one should use REML estimation.

Skin response

Obtain REML estimators of σ_C^2 and σ_ε^2 on the skin response data.

```
lmer_out <- lmer(resp ~ noise + shock + noise:shock + (1|subj), data = skin)
lmer_out
```

```
Linear mixed model fit by REML ['lmerModLmerTest']
Formula: resp ~ noise + shock + noise:shock + (1 | subj)
   Data: skin
REML criterion at convergence: 145
Random effects:
 Groups   Name        Std.Dev.
 subj     (Intercept) 3.32
 Residual             1.51
Number of obs: 40, groups:  subj, 5
Fixed Effects:
      (Intercept)            noise80           shock0.5          shock0.75
              4.8                2.0                3.2                5.4
           shock1   noise80:shock0.5  noise80:shock0.75     noise80:shock1
              3.2               -0.6                2.4                0.4
```

Skin response

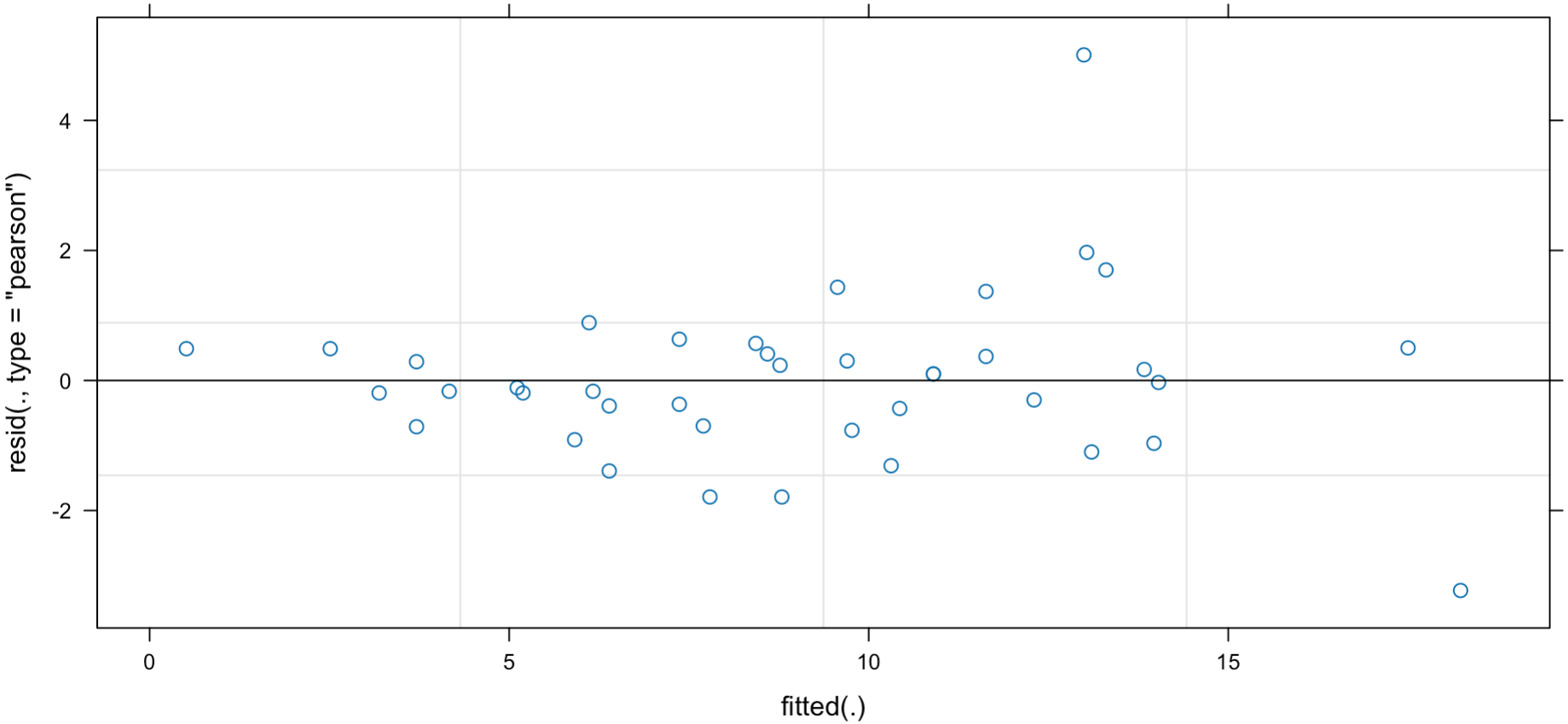
Obtain MoM estimates for σ_C^2 and σ_ε^2 on the skin response data.

```
sg_C <- sqrt((MSC - MSE)/(a*b))
sg_e <- sqrt(MSE)
```

We have $\dot{\sigma}_C = 3.320$ and $\dot{\sigma}_\varepsilon = 1.514$.

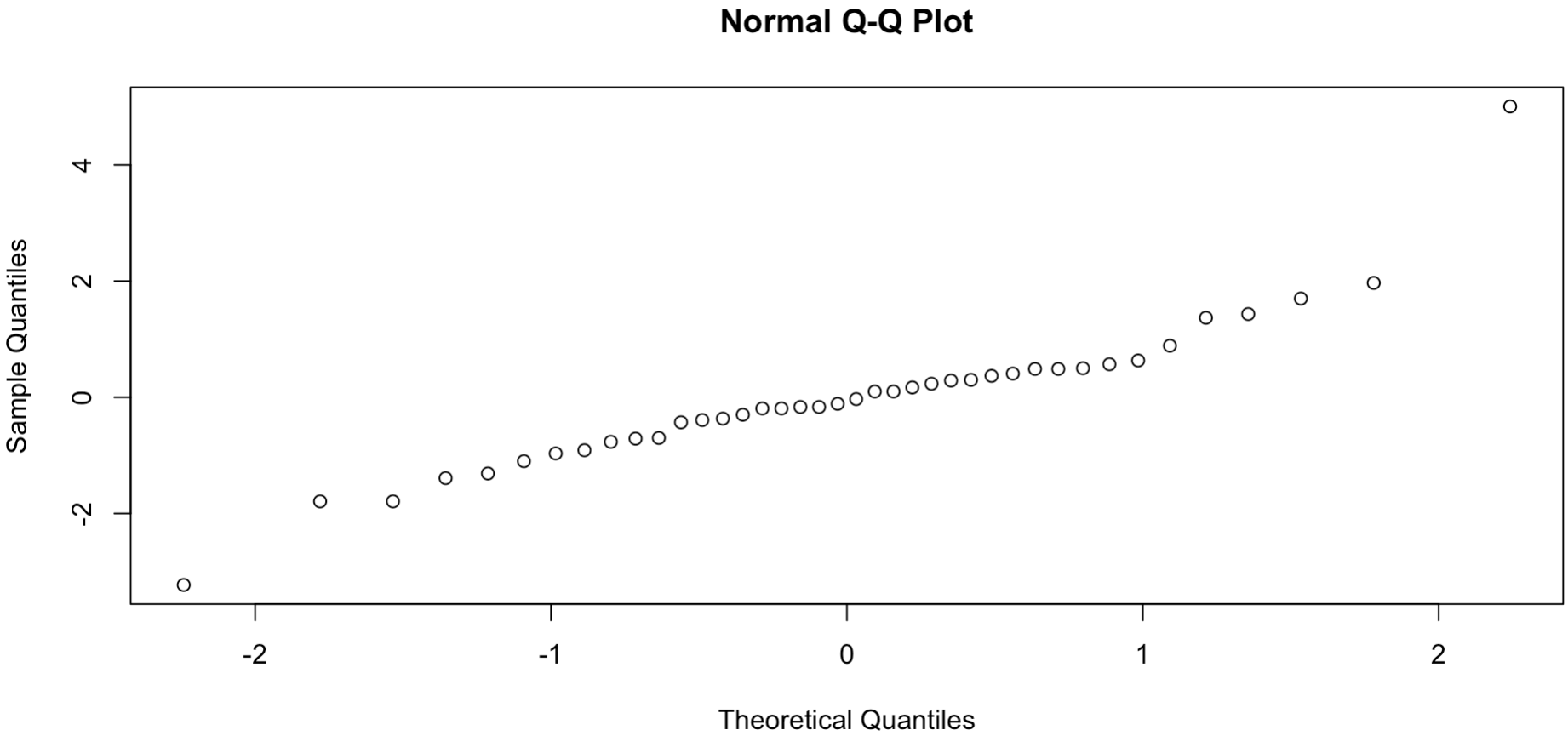
Skin response

```
1 plot(lmer_out)
```



Skin response

```
yhat <- predict(lmer_out)
ehat <- skin$resp - yhat
qqnorm(ehat)
```



References

Kuehl, R. O. 2000. *Design of Experiments: Statistical Principles of Research Design and Analysis*. Duxbury/Thomson Learning.

Mohr, Donna L, William J Wilson, and Rudolf J Freund. 2021. *Statistical Methods*. Academic Press.