STAT 516 Lec 09

Randomized complete block designs

experimental design

Karl Gregory

2026-04-06

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.

a = 6

b = 4

a*b = 24

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

Irrigation Gradient ↓

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

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

```r
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[1] in each block.
- Purpose is to capture the between-block variability among the EUs.
- This helps us detect treatment effects with greater power.

[1] Can have replication, but many RCBDs do not.

Recall 2-way factorial design

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

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

- Y_{ij}: i^{th} treatment / fixed factor level; j^{th} block
- B_j: effect of block j, $\sim N(0, \sigma_B^2)$
- ε_{ij}: error/noise, $\sim N(0, \sigma_\varepsilon^2)$

Really the interaction between τ_i and B_j but functions as the error term.

Treatment effects model for the RCBD

there is no little n, because no replication (yet)

Assume

n = 1

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

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 Normal$(0, \sigma_B^2)$ random block effects.
- the ε_{ij} are independent Normal$(0, \sigma_\varepsilon^2)$ error/interaction terms.
- μ is an overall or baseline mean.

Define the cell means as

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

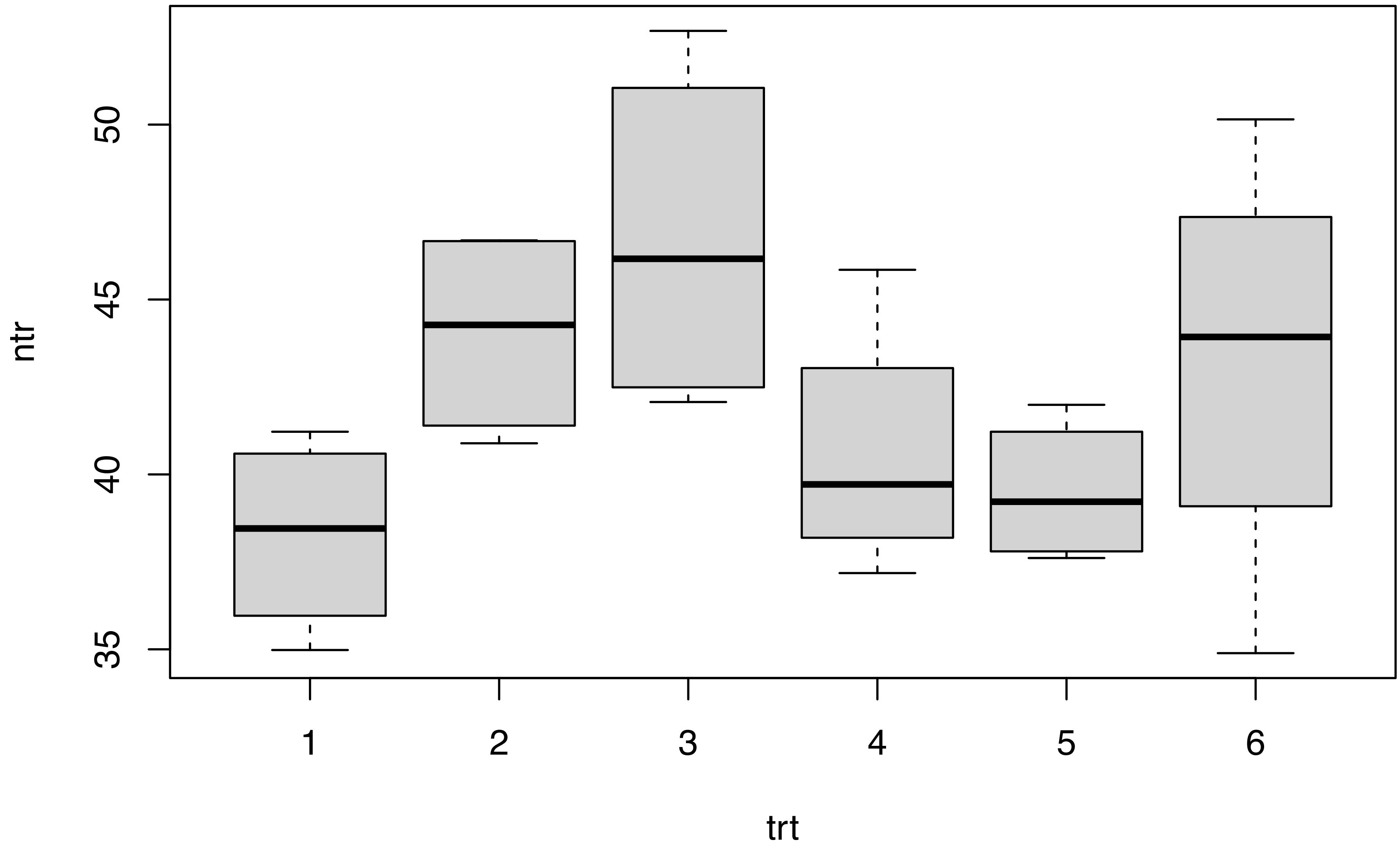
Goals in the RCBD

In the randomized complete block design we wish to

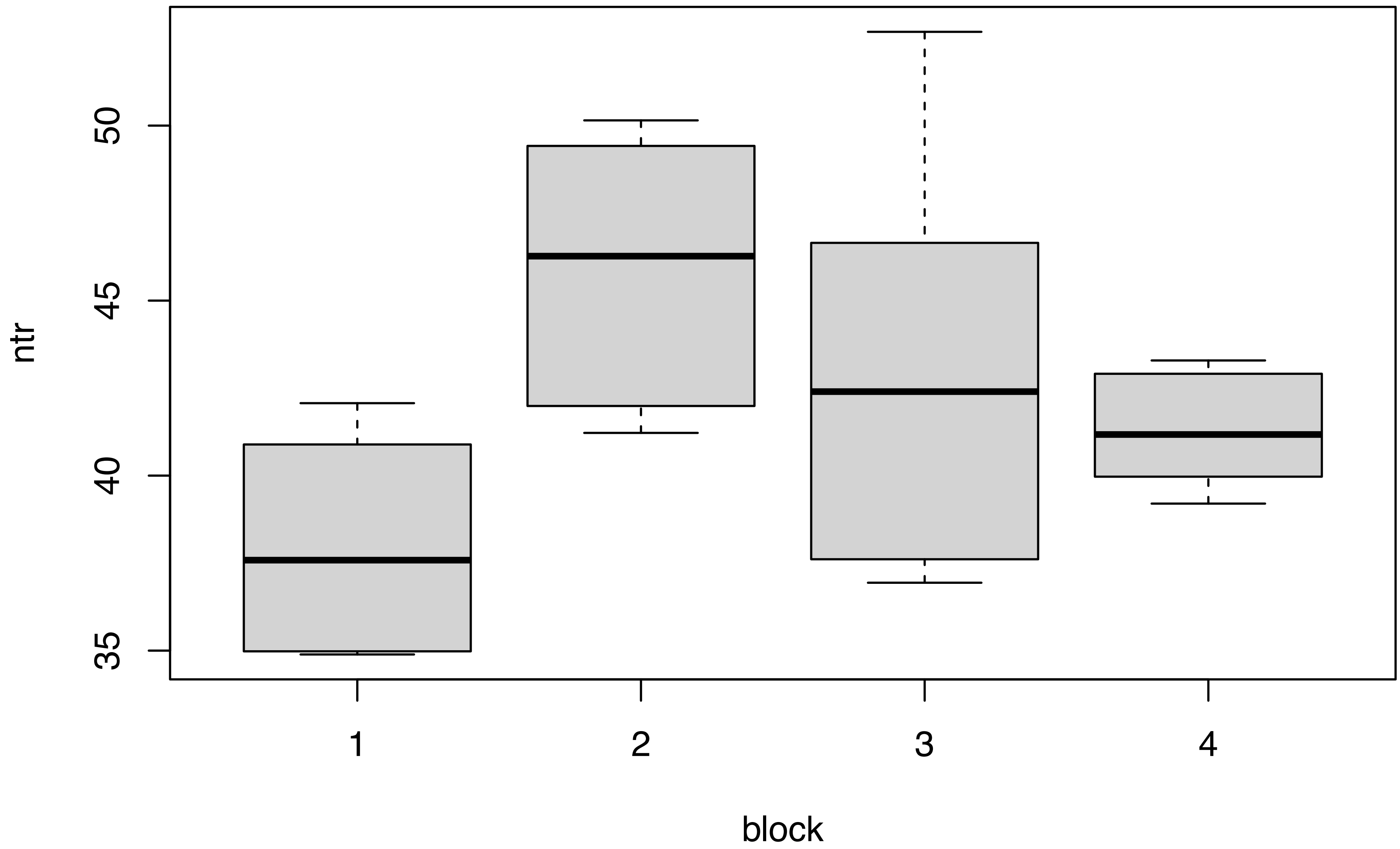
1. Visualize the data.
2. Decompose the variability in the Y_{ij} into its sources.
3. Estimate the variance components σ_B^2 and σ_ε^2.
4. Test whether the treatment has any effect.
5. Make comparisons between treatment means.
6. Check whether the model assumptions are satisfied.

Nitrogen fertilization data (cont)

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



```r
boxplot(ntr~block)
```



$$y_{ij} = \boxed{\mu + \tau_i + \beta_j} + \varepsilon_{ij}$$

Estimated mean of combination i,j. → $\boxed{\bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..})}$

$$= \bar{y}_{i.} + \bar{y}_{.j} - \bar{y}_{..}$$

SS for RCBD

$$SS_{Tot} = \sum_{i=1}^{a} \sum_{j=1}^{b} (y_{ij} - \bar{y}_{..})^2$$

a = # levels fixed factor

b = # blocks

$$SS_{Trt} = \sum_{i=1}^{a} b (\bar{y}_{i.} - \bar{y}_{..})^2$$

$$SS_{Blocks} = \sum_{j=1}^{b} a(\bar{y}_{.j} - \bar{y}_{..})^2$$

$$SS_{Error} = \sum_{i=1}^{a} \sum_{j=1}^{b} \left(y_{ij} - \underbrace{(\bar{y}_{i.} + \bar{y}_{.j} - \bar{y}_{..})}_{\text{estimated mean for } i,j \text{ group}}\right)^2$$

Sums of squares for the RCBD

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

- 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.

ANOVA table for RCBD

$a-1+b-1+(a-1)(b-1) = a-1+b-1+ab-a-b+1 = ab-1.$

add up to total

Really the interaction AB

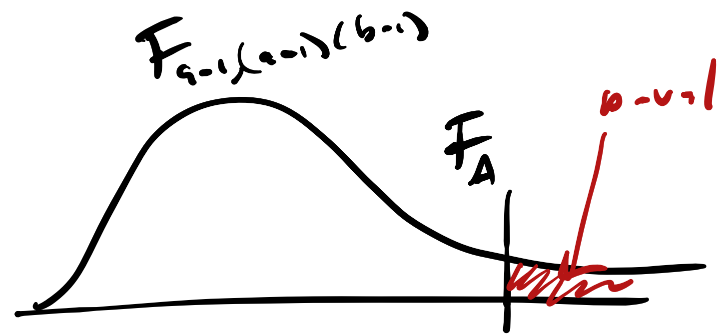
Source	Df	SS	MS	F value
A	$a-1$	$\mathrm{SS_A}$	$\mathrm{MS_A}$	$F_\mathrm{A} = \mathrm{MS_A} / \mathrm{MS_{Error}}$
B	$b-1$	$\mathrm{SS_B}$	$\mathrm{MS_B}$	$F_\mathrm{B} = \mathrm{MS_B} / \mathrm{MS_{Error}}$
Error	$(a-1)(b-1)$	$\mathrm{SS_{Error}}$	$\mathrm{MS_{Error}}$	
Total	$ab-1$	$\mathrm{SS_{Tot}}$		

obs.

pval: $1 - \mathrm{pf}(F_A, a-1, (a-1)(b-1))$

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

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



Nitrogen fertilization data (cont)

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

```
Analysis of Variance Table

Response: ntr
          Df Sum Sq Mean Sq F value   Pr(>F)
trt        5 201.32  40.263  5.5917 0.004191 **
block      3 197.00  65.668  9.1198 0.001116 **
Residuals 15 108.01   7.201
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

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

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))
```

Expected mean squares in the RCBD

Estimate σ_ε^2, σ_B^2

Source	Df	Expected mean square
A	$a-1$	$\mathbb{E}MS_A = b\theta_A^2 + \sigma_\varepsilon^2$
B	$b-1$	$\mathbb{E}MS_B = a\sigma_B^2 + \sigma_\varepsilon^2$
Error	$(a-1)(b-1)$	$\sigma_\varepsilon^2 = \mathbb{E}MS_{Error}$

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

$$\hat{\sigma}_B^2 = \frac{1}{a}\left[MS_B - MS_{Error}\right]$$

$$\hat{\sigma}_\varepsilon^2 = MS_{Error}$$

Method of moments for variance components in RCBD

Equating $\mathrm{MS_B}$ and $\mathrm{MS_{Error}}$ with their expectations gives

- $\dot{\sigma}_\varepsilon^2 = \mathrm{MS_{Error}}$.
- $\dot{\sigma}_B^2 = \dfrac{\mathrm{MS_B} - \mathrm{MS_{Error}}}{a}$

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

Nitrogen fertilization data (cont)

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: 101.5658
Random effects:
 Groups   Name        Std.Dev.
 block    (Intercept) 3.122
 Residual             2.683
Number of obs: 24, groups:  block, 4
Fixed Effects:
(Intercept)         trt2         trt3         trt4         trt5         trt6
     38.278        5.755        8.492        2.337        1.232        4.947
```

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

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[2].

[2] a degrees of freedom approximation when one has not exactly a t-distribution.

Nitrogen fertilization data (cont)

a = 6

$\binom{6}{2} = \frac{6!}{2!\,4!} = \frac{6 \cdot 5}{2} = 15$

Unadjusted CIs with `ls_means()` from R package `lmerTest`

```
ls_means(lmer_out)
```

```
Least Squares Means table:

     Estimate Std. Error  df t value   lower   upper  Pr(>|t|)
trt1  38.2775     2.0582 6.8  18.597 33.3789 43.1761 4.505e-07 ***
trt2  44.0325     2.0582 6.8  21.393 39.1339 48.9311 1.767e-07 ***
trt3  46.7700     2.0582 6.8  22.724 41.8714 51.6686 1.179e-07 ***
trt4  40.6150     2.0582 6.8  19.733 35.7164 45.5136 3.033e-07 ***
trt5  39.5100     2.0582 6.8  19.196 34.6114 44.4086 3.646e-07 ***
trt6  43.2250     2.0582 6.8  21.001 38.3264 48.1236 2.000e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

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

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

```
Least Squares Means table:

              Estimate Std. Error df t value    lower    upper  Pr(>|t|)
trt1 - trt2  -5.7550     1.8974 15 -3.0330  -9.7993  -1.7107 0.0083887 **
trt1 - trt3  -8.4925     1.8974 15 -4.4758 -12.5368  -4.4482 0.0004443 ***
trt1 - trt4  -2.3375     1.8974 15 -1.2319  -6.3818   1.7068 0.2369413
trt1 - trt5  -1.2325     1.8974 15 -0.6496  -5.2768   2.8118 0.5257996
trt1 - trt6  -4.9475     1.8974 15 -2.6075  -8.9918  -0.9032 0.0198026 *
trt2 - trt3  -2.7375     1.8974 15 -1.4427  -6.7818   1.3068 0.1696521
trt2 - trt4   3.4175     1.8974 15  1.8011  -0.6268   7.4618 0.0918200 .
trt2 - trt5   4.5225     1.8974 15  2.3835   0.4782   8.5668 0.0308031 *
trt2 - trt6   0.8075     1.8974 15  0.4256  -3.2368   4.8518 0.6764617
trt3 - trt4   6.1550     1.8974 15  3.2438   2.1107  10.1993 0.0054516 **
trt3 - trt5   7.2600     1.8974 15  3.8262   3.2157  11.3043 0.0016523 **
trt3 - trt6   3.5450     1.8974 15  1.8683  -0.4993   7.5893 0.0813776 .
trt4 - trt5   1.1050     1.8974 15  0.5824  -2.9393   5.1493 0.5689729
trt4 - trt6  -2.6100     1.8974 15 -1.3755  -6.6543   1.4343 0.1891597
trt5 - trt6  -3.7150     1.8974 15 -1.9579  -7.7593   0.3293 0.0691104 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

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

Multiple comparisons of treatment means in the RCBD

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

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

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

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

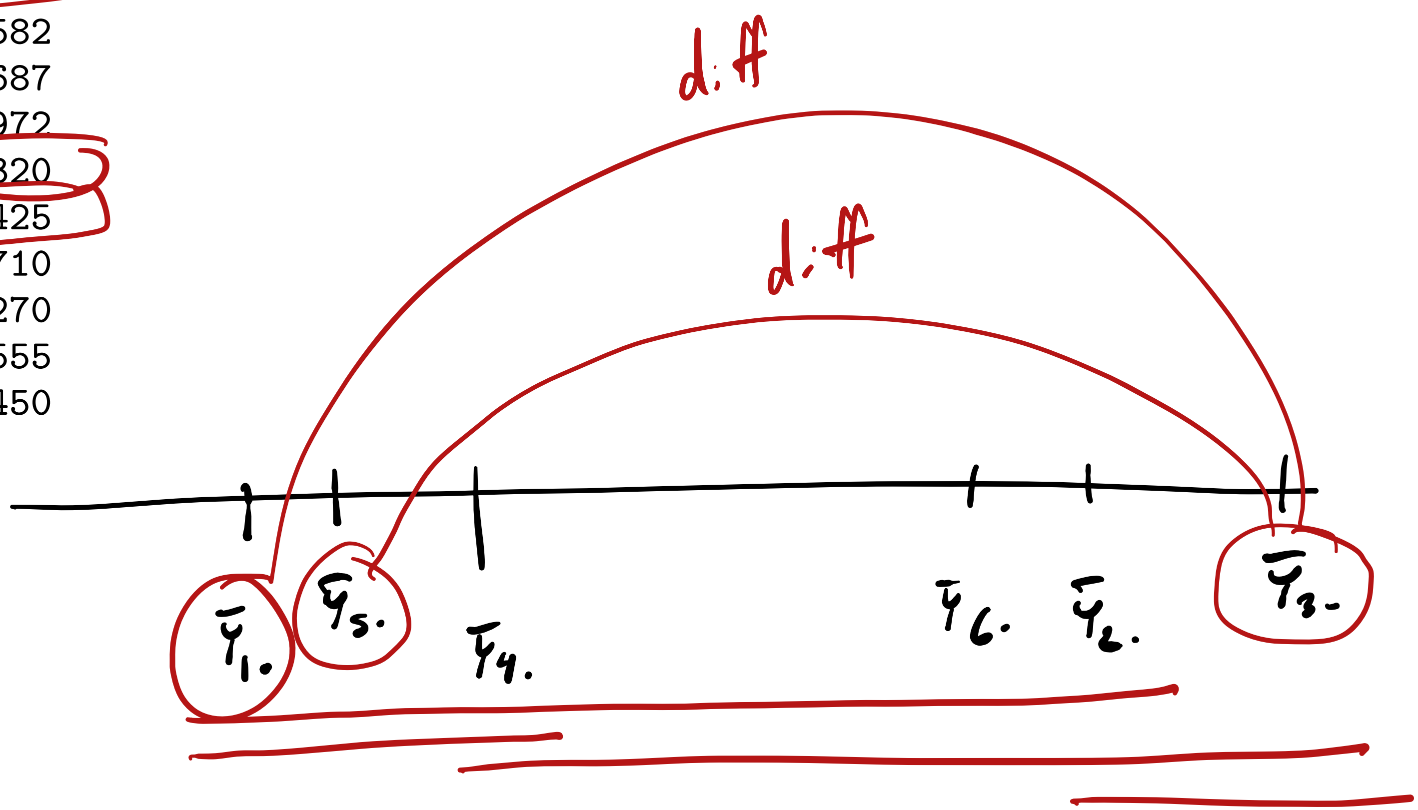
Nitrogen fertilization data (cont)

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

```
alpha <- 0.05
a <- 6
b <- 4
MSE <- anova_out$`Mean Sq`[3]
se <- sqrt(MSE) * sqrt(2/b)
me <- qtukey(1-alpha,a,(a-1)*(b-1)) / sqrt(2) * se
ntr_means <- aggregate(ntr, by = list(trt), mean)$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
```

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

```
         lower   upper
1 - 2 -11.920   0.410
1 - 3 -14.657  -2.328
1 - 4  -8.502   3.827
1 - 5  -7.397   4.932
1 - 6 -11.112   1.217
2 - 3  -8.902   3.427
2 - 4  -2.747   9.582
2 - 5  -1.642  10.687
2 - 6  -5.357   6.972
3 - 4  -0.010  12.320
3 - 5   1.095  13.425
3 - 6  -2.620   9.710
4 - 5  -5.060   7.270
4 - 6  -8.775   3.555
5 - 6  -9.880   2.450
```

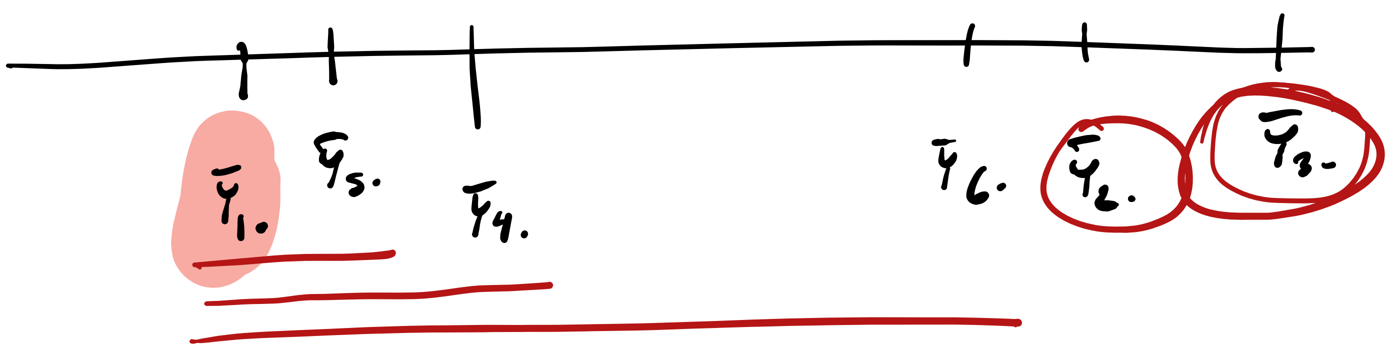


Compare all fertilizers to fertilizer 1 using Dunnett's method.

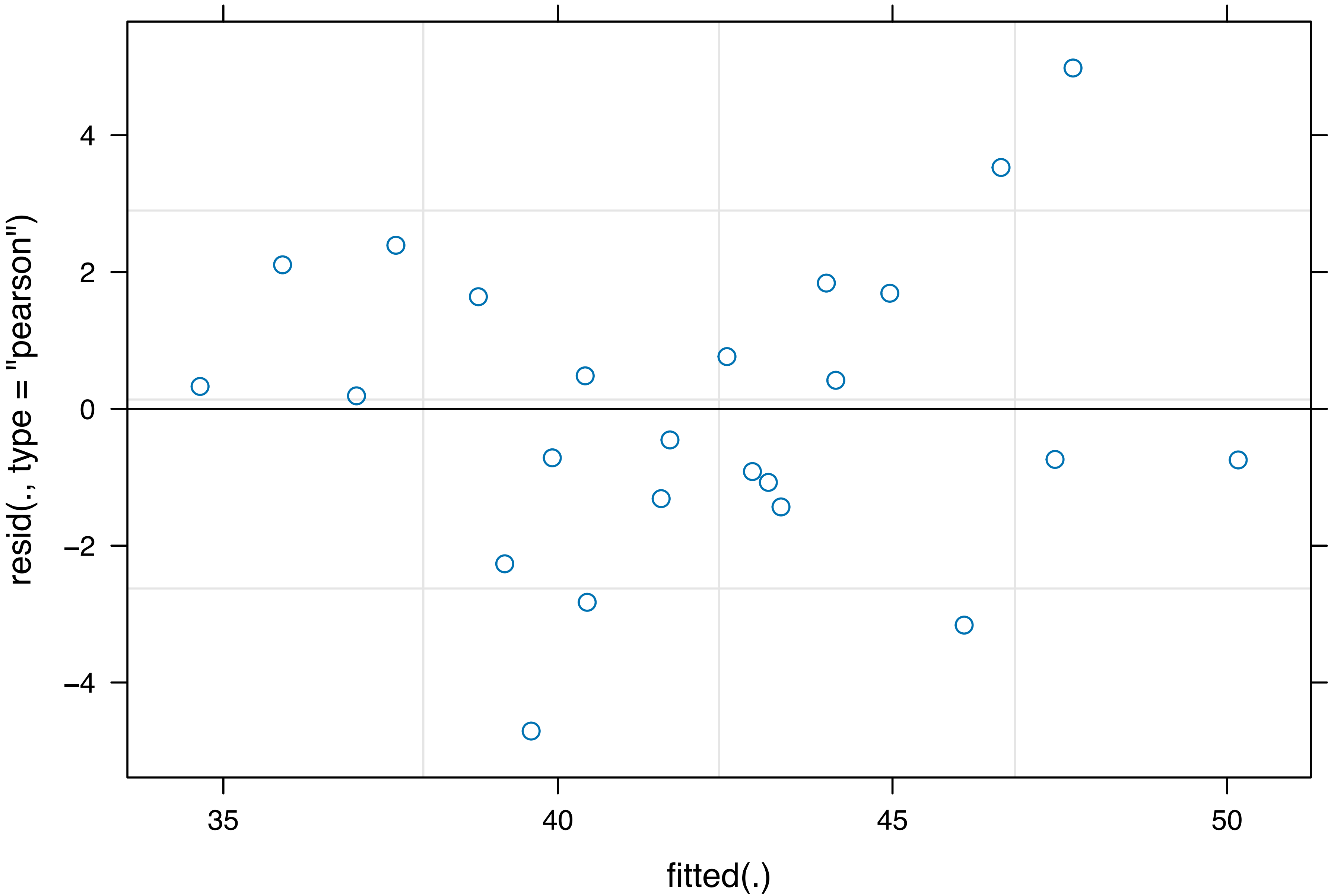
```
alpha <- 0.05
a <- 6
b <- 4
MSE <- anova_out$`Mean Sq`[3]
me <- 2.82 * sqrt(MSE) * sqrt(2/b) # value 2.82 from Dunnett's table
ntr_means <- aggregate(ntr, by = list(trt), mean)$x

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")
```

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

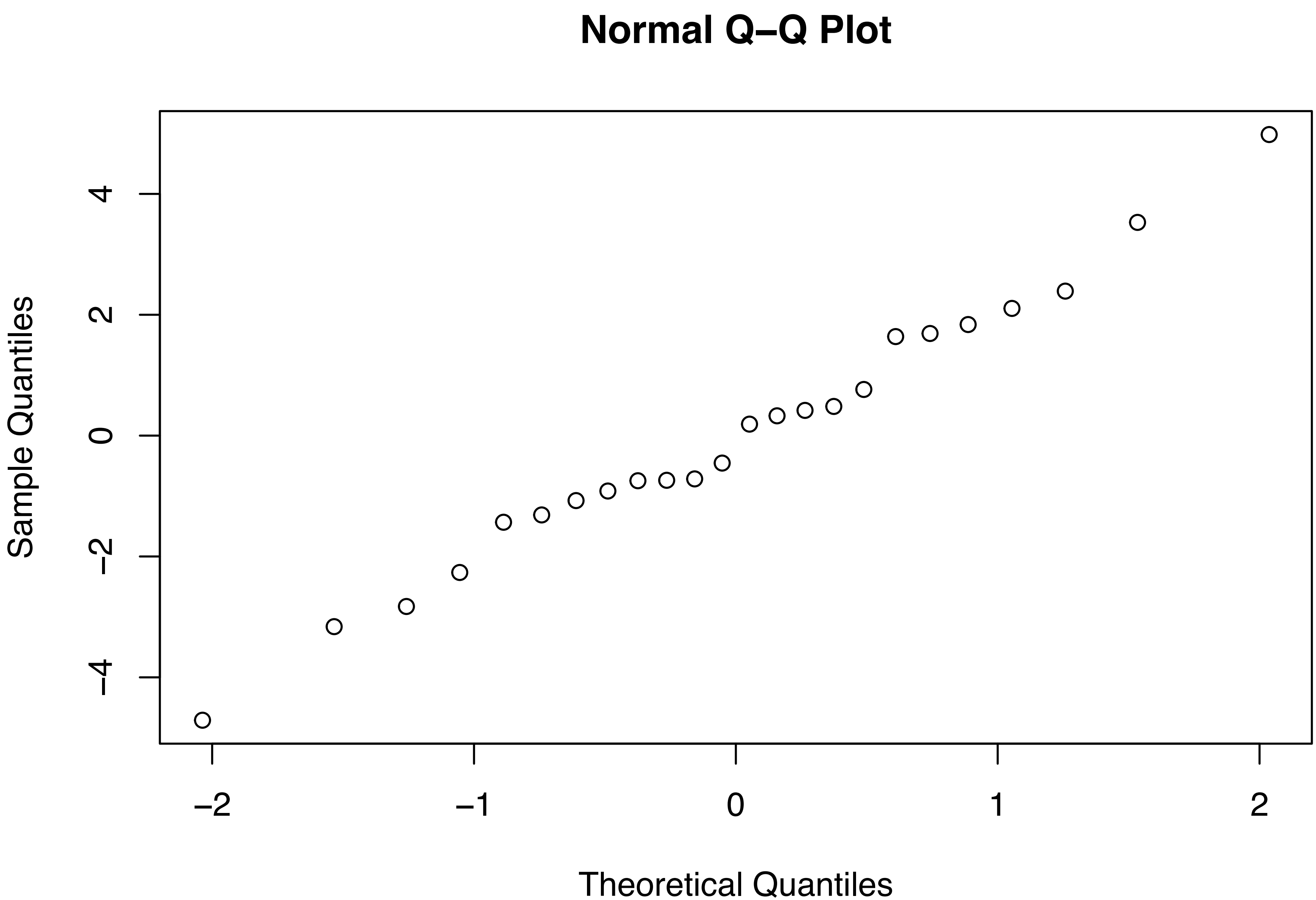
```
      lower  upper
2 - 1 0.404 11.106
3 - 1 3.142 13.843
4 - 1 -3.013  7.688
5 - 1 -4.118  6.583
6 - 1 -0.403 10.298
```



```r
plot(lmer_out)
```



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

We lose power to detect a treatment effect!

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

← ignore the block?

$H_0: \mu_1 = \dots = \mu_a$

```
Analysis of Variance Table

Response: ntr
          Df Sum Sq Mean Sq F value  Pr(>F)
trt        5 201.32  40.263  2.3761 0.08024 .
Residuals 18 305.01  16.945
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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.

$a = 5$

Subplots in each plot randomly assigned to nitrogen levels.

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

Block × trt interaction

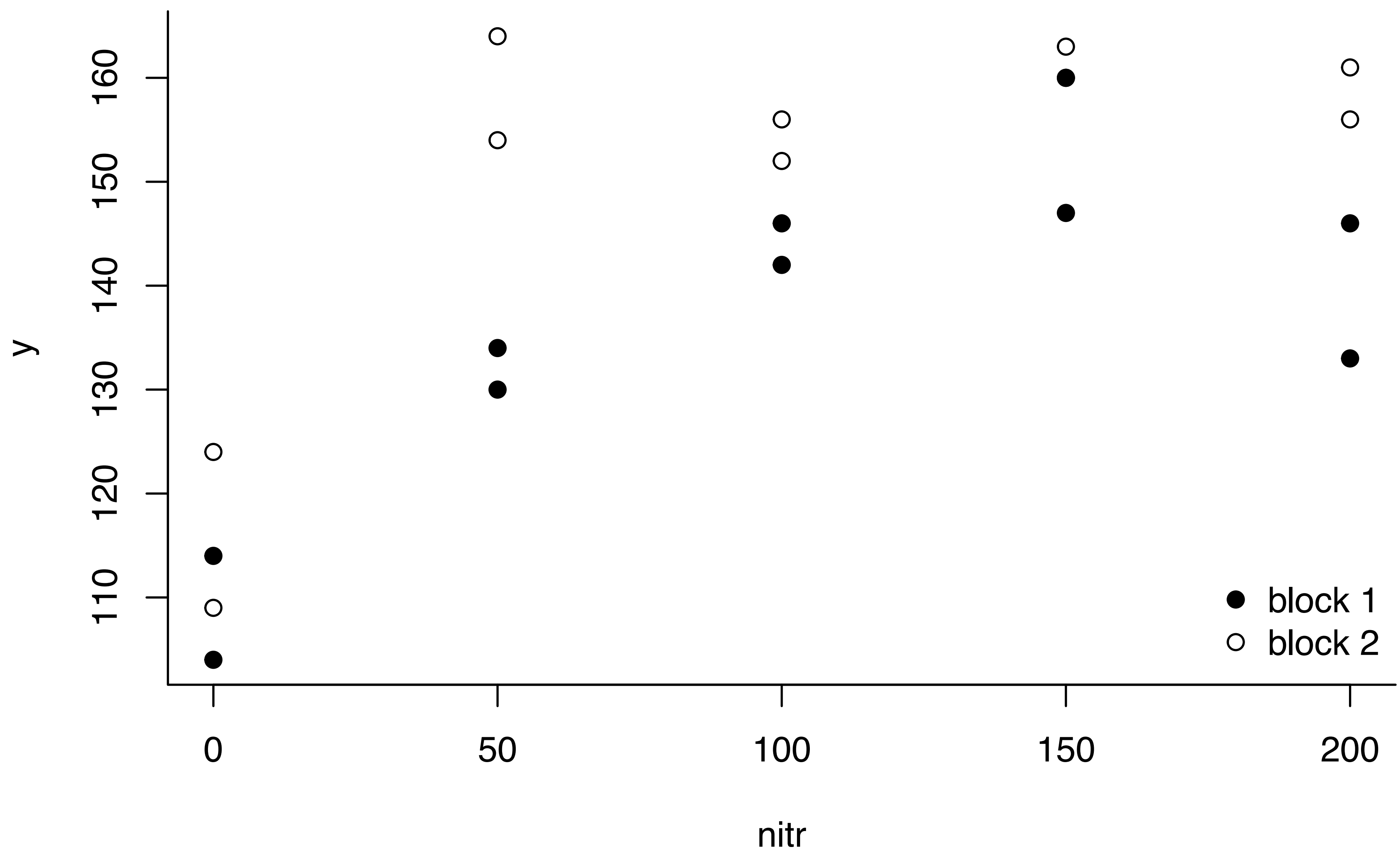
$i = 1, \ldots, 5$

$j = 1, 2$

$k = 1, 2$

```r
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))
```

```
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")
```



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 Normal$(0, \sigma_B^2)$ random block effects.
- the $(\tau B)_{ij}$ are indep. Normal$(0, \sigma_{AB}^2)$ random interaction effects.
- the ε_{ijk} are independent Normal$(0, \sigma_\varepsilon^2)$ 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

Exactly like the 2-way ANOVA for two fixed factors

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 $\times$ 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	$abn-1$	$\mathrm{SS_{Tot}}$		

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 $F_\mathrm{AB} > F_{(a-1)(b-1),ab(n-1),\alpha}$.

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

$\mu + \tau_i = \mu_i$, $B_j \sim N(0, \sigma_B^2)$, $(\tau B)_{ij} \sim N(0, \sigma_{AB}^2)$, $\varepsilon_{ijk} \sim N(0, \sigma_\varepsilon^2)$

Expected MS in RCBD with replication

$MS_A = SS_A/(a-1)$ $\quad MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$

$MS_B = SS_B/(b-1)$

Source	Df	Expected mean square
A	$a-1$	$nb\theta_A^2 + n\sigma_{AB}^2 + \sigma_\varepsilon^2$
B	$b-1$	$na\sigma_B^2 + n\sigma_{AB}^2 + \sigma_\varepsilon^2$ ← Under H_0
AB	$(a-1)(b-1)$	$n\sigma_{AB}^2 + \sigma_\varepsilon^2$
Error	$ab(n-1)$	σ_ε^2

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

Check expected MS values under each H_0 on previous slide.

$H_0: \sigma_B^2 = 0$. $\quad F_{test} = \frac{MS_B}{MS_{AB}}$

ANOVA table for cabbage count data

```
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

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

ANOVA table for cabbage counts data

$a=2$
$b=2$

$a=5$
$b=2$
$n=2$

$H_0: \mu_1 = \cdots = \mu_5$

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			

$5-1$

$2-1$

$(a-1)(b-1)$

$(5-1)(2-1) = 4$

$abn-1 = 5\cdot 2\cdot 2-1 = 19.$

$ab(n-1) = 5\cdot 2(2-1) = 10$

$H_0: \sigma_B^2 = 0$

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

ANOVA table for cabbage count data

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 4813.0 1203.25  28.479 1.915e-05 ***
blk        1 1022.5 1022.45  24.200 0.0006054 ***
nitr:blk   4  287.8   71.95   1.703 0.2253096
Residuals 10  422.5   42.25
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Incorrect when blk is a random effect.

Method of moments for estimating σ_B^2, σ_{AB}^2, and σ_ε^2

Set the mean squares equal to their expectations and solve:

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

Solve for σ_B^2, σ_{AB}^2, and σ_ε^2:

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

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

MOM estimates of cabbage data 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 $\dot{\sigma}_\varepsilon = 6.500$.

REML estimates of cabbage data variance components

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

```
Linear mixed model fit by REML ['lmerModLmerTest']
Formula: y ~ nitr + (1 | blk) + (1 | nitr:blk)
REML criterion at convergence: 110.9695
Random effects:
 Groups   Name        Std.Dev.
 nitr:blk (Intercept) 3.854
 blk      (Intercept) 9.749
 Residual             6.500
Number of obs: 20, groups:  nitr:blk, 10; blk, 2
Fixed Effects:
(Intercept)       nitr50      nitr100      nitr150      nitr200
     112.75        32.75        36.25        44.75        36.25
```

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

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

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

  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 50 and 0 groups:

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

```
[1] 16.09711 49.40289
```

CIs (unadjusted) for parameters in RCBD with replication

Use `confint()` on output of `lmer()`:

```
confint(lmer_out)
```

```
                 2.5 %      97.5 %
.sig01        0.000000   6.253288
.sig02        2.715017  29.859582
.sigma        4.422824   9.062168
(Intercept) 95.101400 130.398323
nitr50       23.245458  42.254543
nitr100      26.745458  45.754543
nitr150      35.245458  54.254543
nitr200      26.745458  45.754543
```

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 Normal$(0, \sigma_B^2)$ random block effects.
- the ε_{ijk} are independent Normal$(0, \sigma_\varepsilon^2)$ 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 $(a-1)(b-1) = 1$.

$$\mu_{ij} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} \qquad i = 1, 2, \quad j = 1, 2.$$

$$\left(\begin{array}{l} \tau_1 = 0 \\ \gamma_1 = 0 \end{array} \right.$$

$$\begin{bmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \end{bmatrix} = \begin{bmatrix} \mu & \mu + \gamma_2 \\ \mu + \tau_2 & \mu + \tau_2 + \gamma_2 + (\tau\gamma)_{22} \end{bmatrix}$$

Sums of squares for RCBD with rep., no interaction

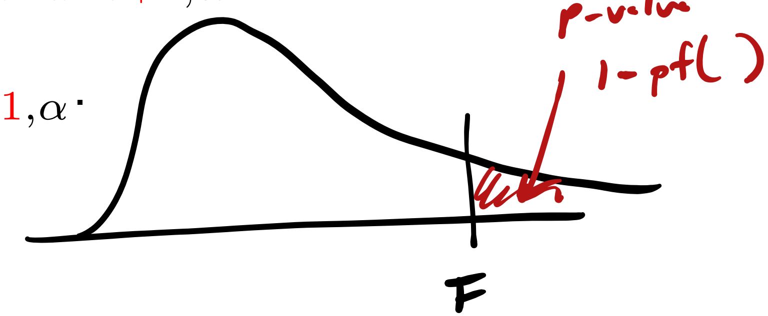
SS	Symbol	Formula
Total	SS_{Tot}	$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk}-\bar{Y}_{...})^2$
Treat-ment	SS_{A}	$nb\sum_{i=1}^{a}(\bar{Y}_{i..}-\bar{Y}_{...})^2$
Block	SS_{B}	$na\sum_{j=1}^{b}(\bar{Y}_{.j.}-\bar{Y}_{...})^2$
Error	SS_{Error}	$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk}-(\bar{Y}_{i..}+\bar{Y}_{.j.}-\bar{Y}_{...}))^2$

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

ANOVA table for RCBD with rep., omitting interaction

Source	Df	SS	MS	F value
A	$a-1$	$\mathrm{SS_A}$	$\mathrm{MS_A}$	$F_\mathrm{A} = \mathrm{MS_A} / \mathrm{MS_{Error}}$
B	$b-1$	$\mathrm{SS_B}$	$\mathrm{MS_B}$	$F_\mathrm{B} = \mathrm{MS_B} / \mathrm{MS_{Error}}$
Error	$abn-a-b+1$	$\mathrm{SS_{Error}}$	$\mathrm{MS_{Error}}$ ($\hat{\sigma}_\varepsilon^2$)	
Total	$abn-1$	$\mathrm{SS_{Tot}}$		

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 $F_\mathrm{B} > F_{b-1,abn-a-b+1,\alpha}$.



Previously:

A	$a-1$
B	$b-1$
AB	$(a-1)(b-1)$
error	$ab(n-1)$
total	$abn-1$

combine these

$(a-1)(b-1) + ab(n-1) = ab - a - b + 1 + abn - ab$

$= abn - a - b + 1$

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

ANOVA table for cabbage counts, omitting interaction

```r
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)
```

ANOVA table for cabbage counts, omitting interaction

$H_0: \mu_1 = \cdots = \mu_a$

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			

$H_0: \sigma_B^2 = 0$

ANOVA table for cabbage counts, omitting interaction

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

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

```
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
nitr       4 4813.0 1203.25  23.716  4.13e-06 ***
blk        1 1022.5 1022.45  20.152 0.0005097 ***
Residuals 14  710.3   50.74
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

M.o.M. for σ_B^2 and σ_ε^2 when omitting interaction

Set the mean squares equal to their expectations and solve:

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

Solve for σ_B^2 and σ_ε^2:

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

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

M.o.M.s for cabbage data 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 $\dot{\sigma}_\varepsilon = 7.123$.

REML for cabbage data when omitting interaction

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

```
Linear mixed model fit by REML ['lmerModLmerTest']
Formula: y ~ nitr + (1 | blk)
REML criterion at convergence: 111.4024
Random effects:
 Groups   Name        Std.Dev.
 blk      (Intercept) 9.858
 Residual             7.123
Number of obs: 20, groups:  blk, 2
Fixed Effects:
(Intercept)       nitr50      nitr100      nitr150      nitr200
     112.75        32.75        36.25        44.75        36.25
```

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

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

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

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

CI for difference in means for cabbage count data when omitting interaction

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
c(lo,up)
```

$t_{abn-a-b+1,\alpha/2}$

```
[1] 21.94746 43.55254
```

CI for difference in means for cabbage count data when omitting interaction

Compare above to `confint()` on output from `lmer()`.

```
confint(lmer_out)
```

```
                 2.5 %      97.5 %
.sig01        2.715987   29.862871
.sigma        4.678640    9.062171
(Intercept)  95.101978  130.398015
nitr50       23.558217   41.941783
nitr100      27.058217   45.441783
nitr150      35.558217   53.941783
nitr200      27.058217   45.441783
```

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

```r
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)))
```

```r
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
19    7    40     1    4
20    4    40     1    5
```

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 A×B interaction.
- the C_k are independent Normal$(0, \sigma_C^2)$ block effects.
- the ε_{ijk} are independent Normal$(0, \sigma_\varepsilon^2)$ 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	SS_{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	SS_{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	SS_{Error}	$SS_{\text{Tot}}-(SS_{\text{A}}+SS_{\text{B}}+SS_{\text{AB}}+SS_{\text{C}})$

- Then we have $SS_{\text{Tot}} = SS_{\text{A}} + SS_{\text{B}} + SS_{\text{AB}} + SS_{\text{C}} + SS_{\text{Error}}$.
- The error SS_{Error} is really the interaction sum of squares $SS_{\text{Trt}\times C}$.
- Again, without replication, we cannot estimate this interaction.
- So the interaction serves as the error term.

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: $\mu_{1.} = \cdots = \mu_{a.}$ if $F_{\mathrm{A}} > F_{a-1,(ab-1)(c-1),\alpha}$.
2. Reject H_0: $\mu_{.1} = \cdots = \mu_{.b}$ if $F_{\mathrm{B}} > F_{b-1,(ab-1)(c-1),\alpha}$.
3. R. H_0: $\mu_{ij} = \bar{\mu}_{i.} + \bar{\mu}_{.j} - \bar{\mu}_{..}$ $\forall ij$ if $F_{\mathrm{AB}} > F_{(a-1)(b-1),(ab-1)(c-1),\alpha}$.
4. Reject H_0: $\sigma_C^2 = 0$ if $F_{\mathrm{C}} > F_{c-1,(ab-1)(c-1),\alpha}$.

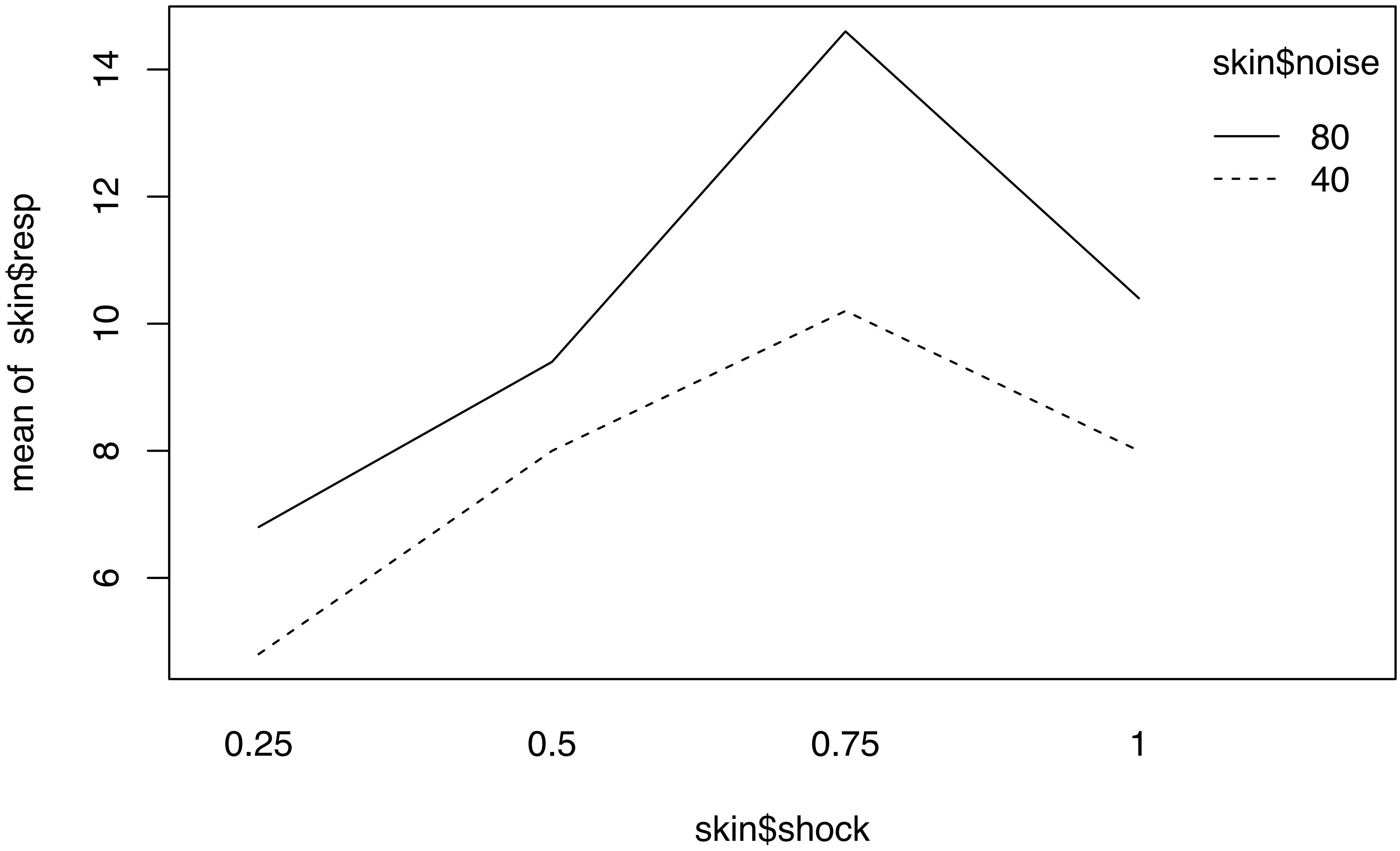
Skin response data (cont)

```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  65.02  65.025 28.3819 1.134e-05 ***
shock        3 219.27  73.092 31.9028 3.564e-09 ***
subj         4 361.85  90.463 39.4848 3.975e-11 ***
noise:shock  3  12.67   4.225  1.8441    0.1621    
Residuals   28  64.15   2.291                      
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```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))

FA <- MSA / MSE
FB <- MSB / MSE
FAB <- MSAB / MSE
FC <- MSC / MSE

pA <- 1 - pf(FA,a-1,(a*b-1)*(c-1))
pB <- 1 - pf(FB,b-1,(a*b-1)*(c-1))
pAB <- 1 - pf(FAB,(a-1)*(b-1),(a*b-1)*(c-1))
pC <- 1 - pf(FC,c-1,(a*b-1)*(c-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 $\mathrm{MS_C}$ and $\mathrm{MS_{Error}}$ with their expectations gives

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

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

Skin response data (cont)

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

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

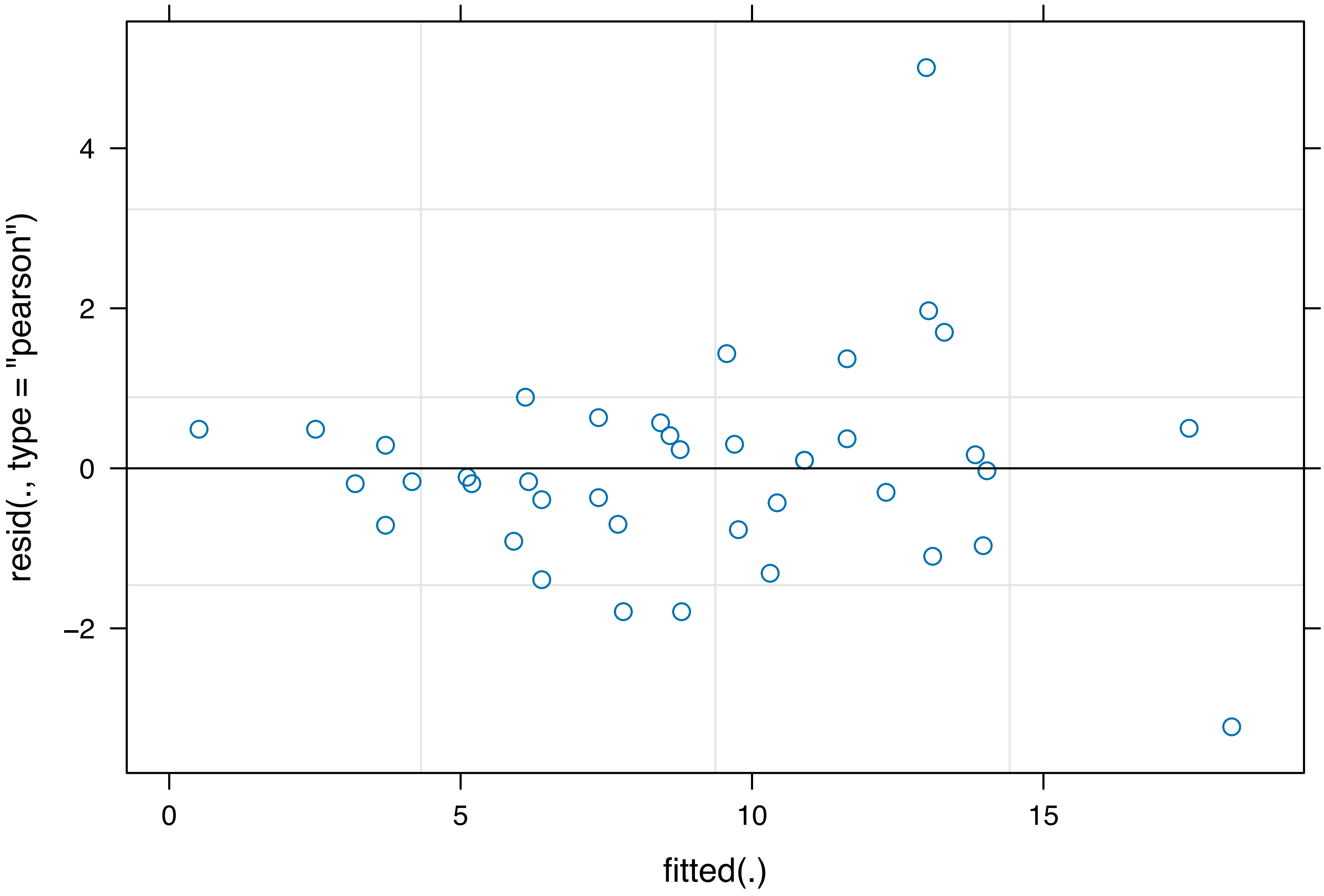
```
Linear mixed model fit by REML ['lmerModLmerTest']
Formula: resp ~ noise + shock + noise:shock + (1 | subj)
   Data: skin
REML criterion at convergence: 144.9199
Random effects:
 Groups   Name        Std.Dev.
 subj     (Intercept) 3.320
 Residual             1.514
Number of obs: 40, groups:  subj, 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
```

Obtain MoMs estimators for σ_C^2 and σ_ε^2 on the skin response data.

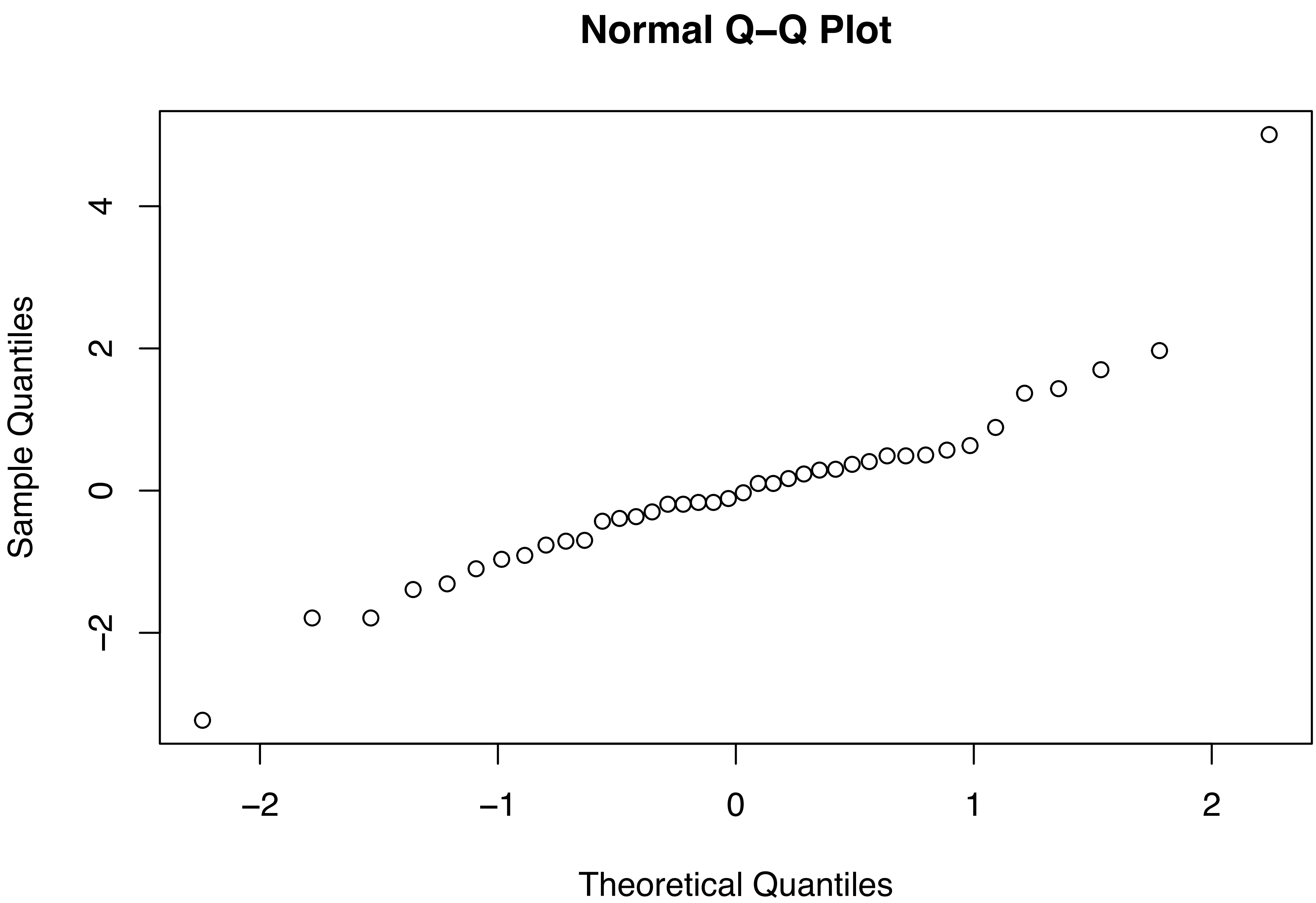
```
sg_C <- sqrt((MSC - MSE)/(a*b))
sg_e <- sqrt(MSE)
```

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

```
plot(lmer_out)
```



```r
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.