

STAT 516 Lec 09

Randomized complete block designs

Karl Gregory

2024-04-04

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
↓

| | | | | | | |
|----------------|------------|------------|------------|------------|------------|------------|
| <i>Block 1</i> | 2 40.89 | 5 37.99 | 4 37.18 | 1 34.98 | 6 34.89 | 3 42.07 |
| <i>Block 2</i> | 1 41.22 | 3 49.42 | 4 45.85 | 6 50.15 | 5 41.99 | 2 46.69 |
| <i>Block 3</i> | 6 44.57 | 3 52.68 | 5 37.61 | 1 36.94 | 2 46.65 | 4 40.23 |
| <i>Block 4</i> | 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.

```
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(1,1,1,1,1,1,2,2,2,2,2,3,3,3,3,3,3,4,4,4,4,4,4))  
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 once¹ in each block.
- ▶ Purpose is to capture the between-block variability among the EUs.
- ▶ This helps us detect treatment effects with greater power.

¹Can have replication, but many RCBDs do not.

Treatment effects model for the RCBD

Assume

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

where

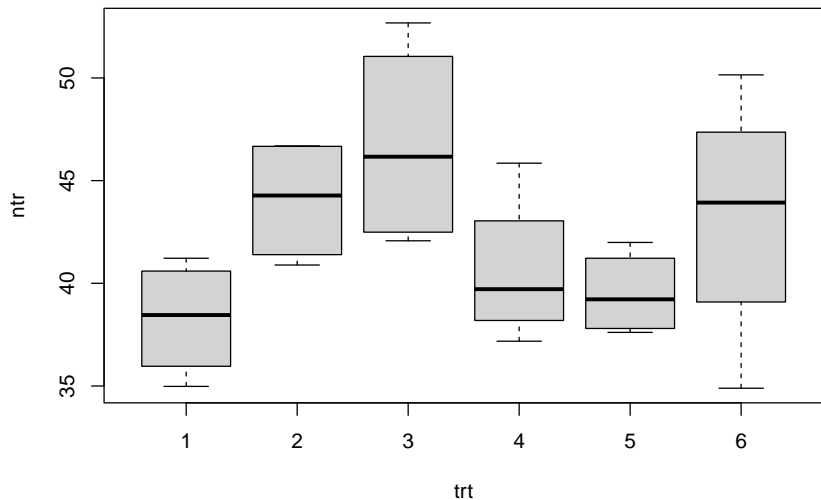
- ▶ Y_{ij} is the response of the EU in block j receiving treatment i .
- ▶ the τ_i are the fixed effects of the treatment.
- ▶ the B_j are independent $\text{Normal}(0, \sigma_B^2)$ random block effects.
- ▶ the ε_{ij} are independent $\text{Normal}(0, \sigma_\varepsilon^2)$ error/interaction terms.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

Define the cell means as

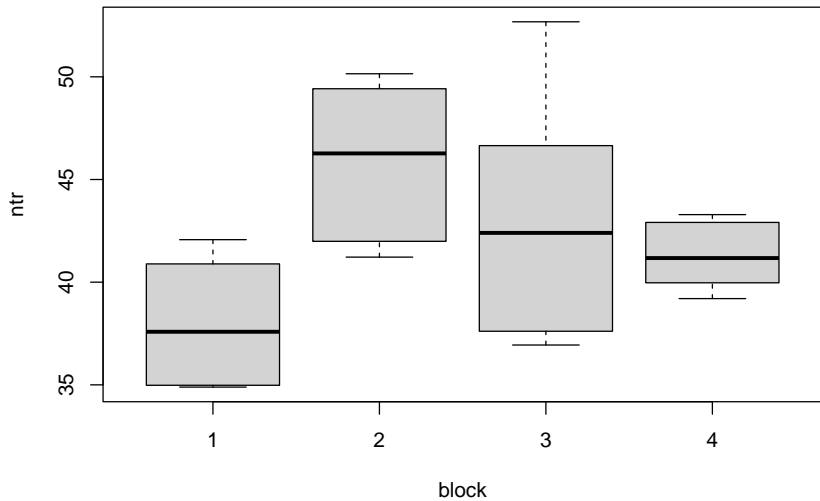
$$\mu_i = \mu + \tau_i, \quad i = 1, \dots, a.$$

Nitrogen fertilization data (cont)

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



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



Sums of squares for the RCBD

| Sum of squares | Symbol | Formula |
|----------------|---------------------|---|
| Total | SS_{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 | SS_{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 $SS_{\text{Tot}} = SS_A + SS_B + SS_{\text{Error}}$.
- ▶ The SS_{Error} is really the interaction sum of squares SS_{AB} .
- ▶ But without replication, we cannot estimate an interaction.
- ▶ So the interaction serves as the error term.

ANOVA table for RCBD

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

1. Reject $H_0: \mu_1 = \dots = \mu_a$ if $F_A > F_{a-1, (a-1)(b-1), \alpha}$.
2. Reject $H_0: \sigma_B^2 = 0$ if $F_B > F_{b-1, (a-1)(b-1), \alpha}$.

Nitrogen fertilization data (cont)

```
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

```
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

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

Method of moments for variance components in RCBD

Equating MS_B and MS_{Error} with their expectations gives

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

$$\blacktriangleright \hat{\sigma}_B^2 = \frac{MS_B - MS_{\text{Error}}}{a}$$

May obtain $\hat{\sigma}_B^2 < 0$, 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 $\hat{\sigma}_B = 3.122$ and $\hat{\sigma}_\varepsilon = 2.683$.

Variances of some means and difference in means

| Contrast | Variance | MoM variance estimator |
|-------------------------------|--|--|
| $\bar{Y}_{i.}$ | $\frac{1}{b}(\sigma_B^2 + \sigma_\varepsilon^2)$ | $\frac{2}{ab}[\text{MS}_B + (b - 1) \text{MS}_{\text{Error}}]$ |
| $\bar{Y}_{i.} - \bar{Y}_{i'.$ | $\frac{2}{b}\sigma_\varepsilon^2$ | $\frac{2}{b} \text{MS}_{\text{Error}}$ |

Some (unadjusted) CIs in RCB split plot design

| | |
|--------|---|
| Target | $(1 - \alpha)100\%$ confidence interval |
|--------|---|

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

$$\mu_i - \mu_{i'} \quad \bar{Y}_{i.} - \bar{Y}_{i'}. \pm t_{(a-1)(b-1), \alpha/2} \sqrt{MS_{\text{Error}}} \sqrt{\frac{2}{b}}$$

In the above $\nu^* = \frac{MS_B + (b-1) MS_{\text{Error}}}{\frac{MS_B^2}{(b-1)} + \frac{(b-1)^2 MS_{\text{Error}}^2}{(a-1)(b-1)}}$ à la Satterthwaite².

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

Nitrogen fertilization data (cont)

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

```
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 μ_1, \dots, μ_a :

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

- ▶ Dunnett's for comparing μ_1, \dots, μ_{a-1} to a baseline μ_a :

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

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

```
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

}

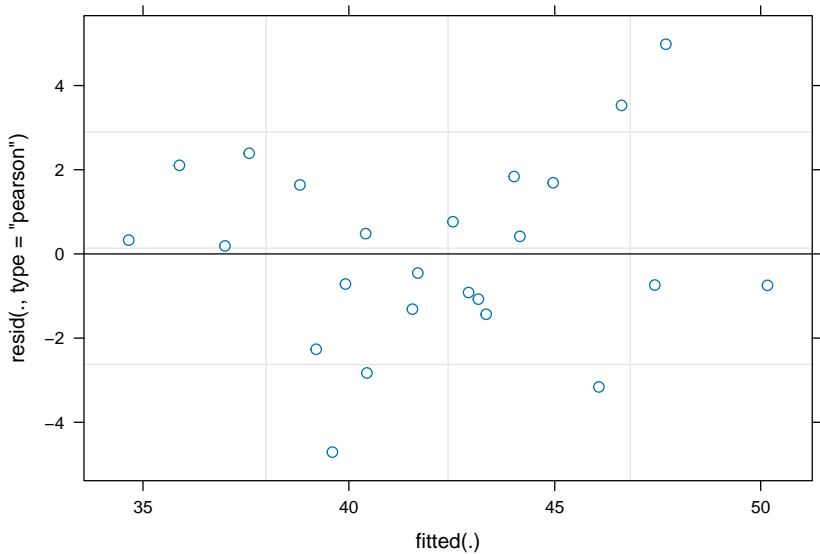
colnames(CIs) <- c("lower","upper")
rownames(CIs) <- paste(2:a,"- 1")
```

```
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
```

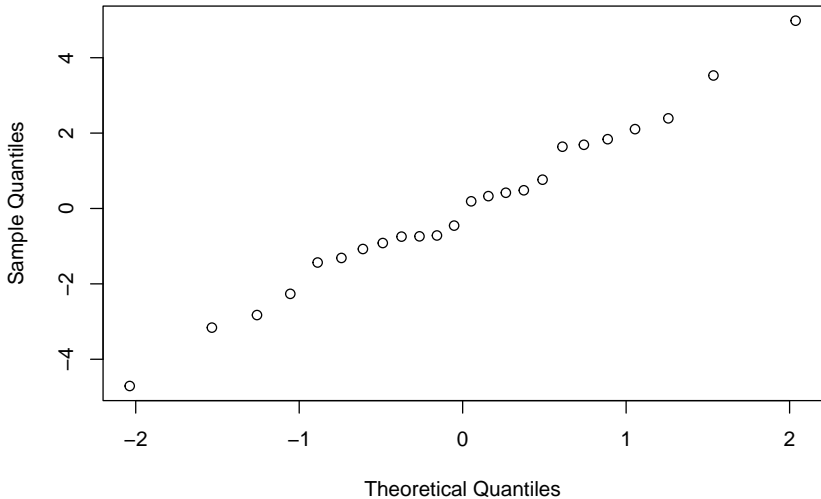


```
plot(lmer_out)
```



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

Normal Q-Q Plot



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))
```

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

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)))
```

```
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 $i = 1, \dots, a$, $j = 1, \dots, b$, and $k = 1, \dots, c$, 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 $\text{Normal}(0, \sigma_C^2)$ block effects.
- ▶ the ε_{ijk} are independent $\text{Normal}(0, \sigma_\varepsilon^2)$ error terms.

Define the cell means as

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

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} | $SS_{\text{Tot}} - (SS_A + SS_B + SS_{AB} + SS_C)$ |

- ▶ Then we have $SS_{\text{Tot}} = SS_A + SS_B + SS_{AB} + SS_C + SS_{\text{Error}}$.
- ▶ 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 | $F_A = MS_A / MS_{\text{Error}}$ |
| B | $b - 1$ | SS_B | MS_B | $F_B = MS_B / MS_{\text{Error}}$ |
| AB | $(a - 1)(b - 1)$ | SS_{AB} | MS_{AB} | $F_{AB} = MS_{AB} / MS_{\text{Error}}$ |
| C | $c - 1$ | SS_C | MS_C | $F_C = MS_C / MS_{\text{Error}}$ |
| Error | $(ab - 1)(c - 1)$ | SS_{Error} | MS_{Error} | |
| Total | $abc - 1$ | SS_{Tot} | | |

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

Skin response data (cont)

```
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

```
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)
```

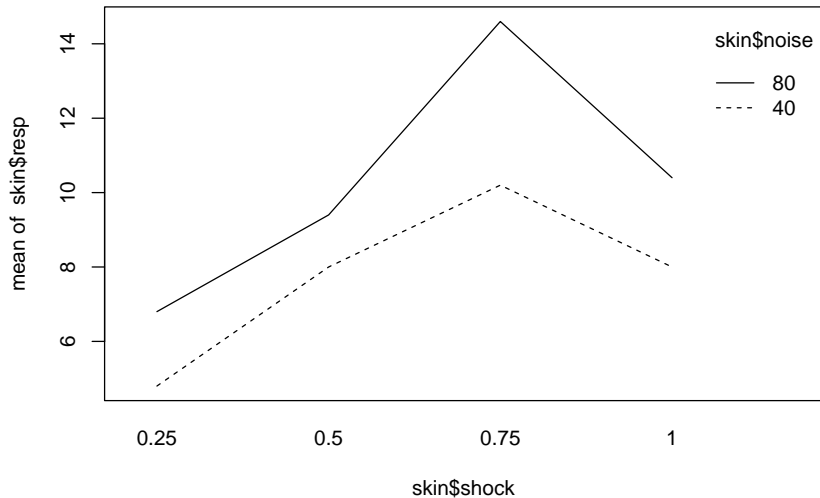
```
a <- 2
b <- 4
c <- 5
```

```
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

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

$$\blacktriangleright \theta_B^2 = (b - 1)^{-1} \sum_{j=1}^b (\bar{\mu}_{.j} - \bar{\mu}_{..})^2$$

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

MoMs for variance components in two-way factorial RCBD

- ▶ Equating MS_C and MS_{Error} with their expectations gives

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

- ▶ May obtain $\hat{\sigma}_C^2 < 0$, 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.

```
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:

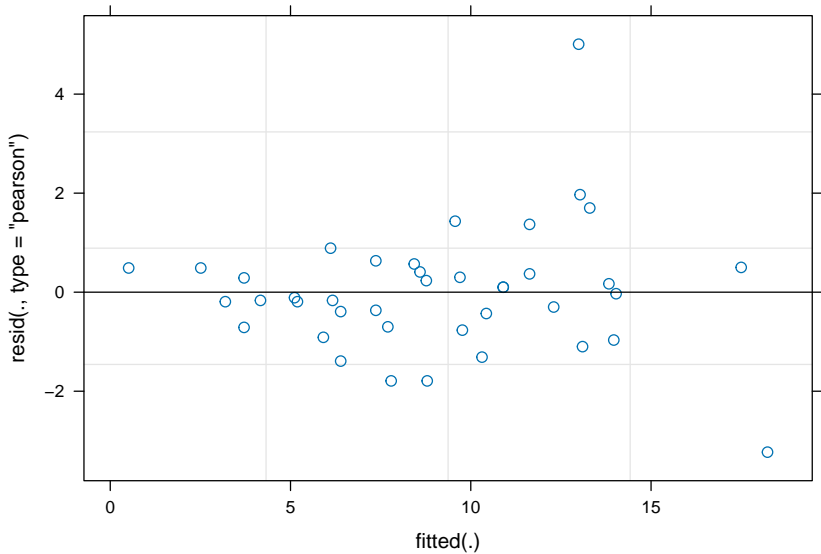
| | | | | |
|-------------|------------------|-------------------|----------|----------------|
| (Intercept) | | noise80 | shock0.5 | shock0.75 |
| 4.8 | | 2.0 | 3.2 | 5.4 |
| shock1 | noise80:shock0.5 | noise80:shock0.75 | | noise80:shock1 |
| 3.2 | -0.6 | 2.4 | | 0.4 |

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)
```

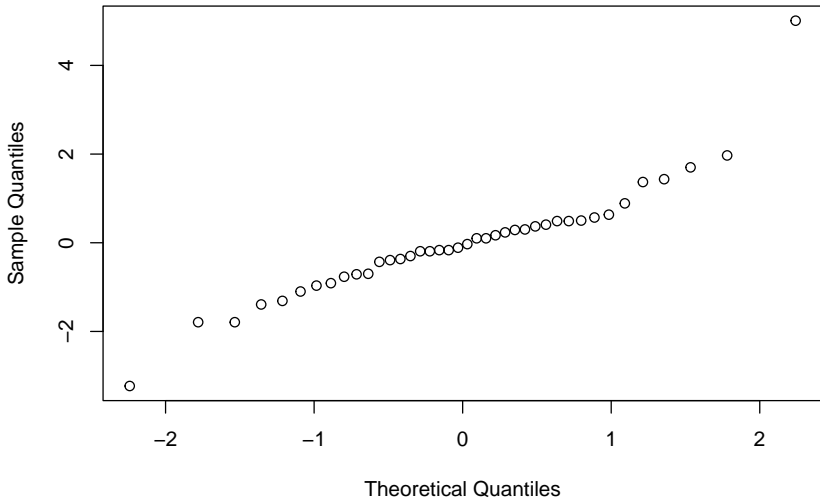
We have $\hat{\sigma}_C = 3.320$ and $\hat{\sigma}_\varepsilon = 1.514$.


```
plot(lmer_out)
```



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

Normal Q-Q Plot



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.