STAT 516 Lec 08

One- and two-way random effects models

Karl Gregory

2026-11-02

Fish net data from Dr. Longnecker's Notes

Strength of fish net material made from four randomly chosen machines.

	Strength Measurements: y_{ij}					
Machine	1	2	3	4	5	Mean: $\bar{y}_{i\cdot}$
1	128	127	129	126	128	127.6
2	121	120	123	122	125	122.2
3	126	125	127	125	124	125.4
4	125	126	129	128	127	127.0

One-way ANOVA:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij},$$

τ_i: treatment effect

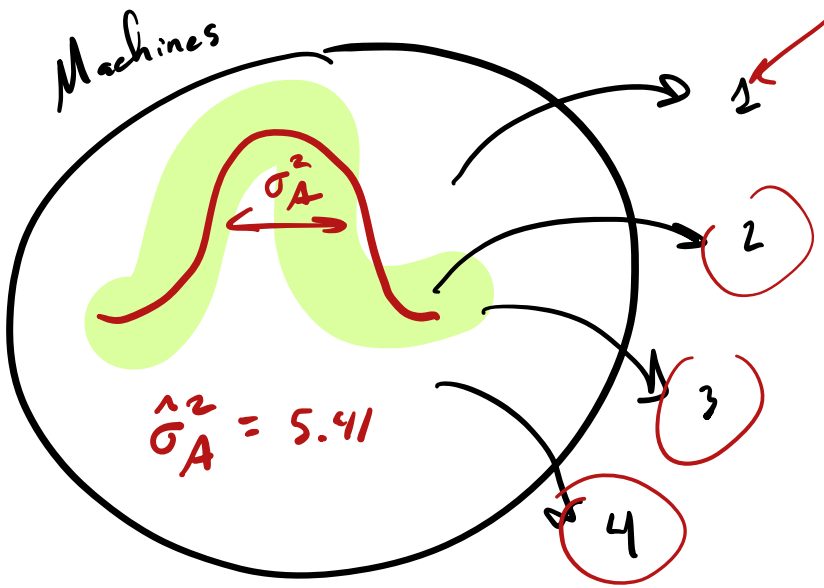
$$\varepsilon_{ij} \overset{ind}{\sim} N(0, \sigma^2)$$

τ_i are "fixed"

"fixed effects"

```
y <- c(128,127,129,126,128,121,120,123,122,125,
       126,125,127,125,124,125,126,129,128,127)
machine <- as.factor(c(1,1,1,1,1,2,2,2,2,2,3,3,3,3,3,4,4,4,4,4))
```

Is there significant machine-to-machine variability?



One-way random effects model:

$$y_{ij} = \mu + A_i + \varepsilon_{ij},$$

noise

treatment effect

overall mean

"variance components"

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

"random effects"

$$A_i \overset{ind}{\sim} N(0, \sigma_A^2)$$

One-way random effects model

Suppose

$$Y_{ij} = \mu + A_i + \varepsilon_{ij}, \quad i = 1, \ldots, a, \quad j = 1, \ldots, n_i$$

where

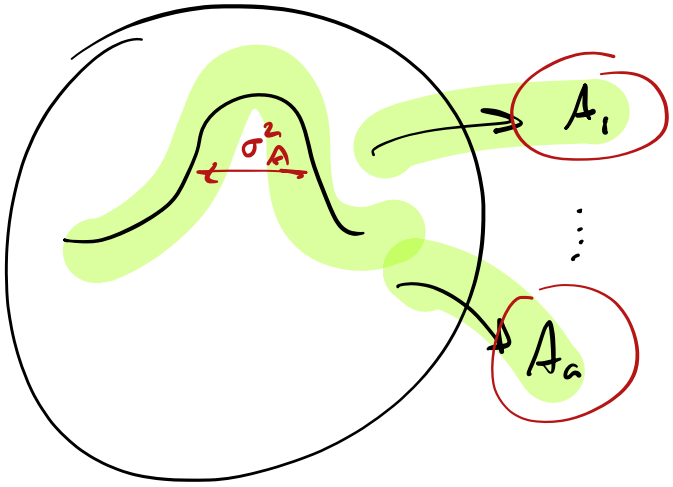
- Y_{ij} is the reponse of EU j in treatment group i.
- the A_i are independent $\mathcal{N}(0, \sigma_A^2)$ rvs called random effects.
- the ε_{ij} are independent $\mathcal{N}(0, \sigma_\varepsilon^2)$.
- the A_i and the ε_{ij} are independent of each other.
- μ is the overall mean
- we call σ_A^2 and σ_ε^2 variance components.

Until now we have studied fixed-effects models.

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

$A_i \sim N(0, \sigma_A^2)$ $\quad$ $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$

random effects



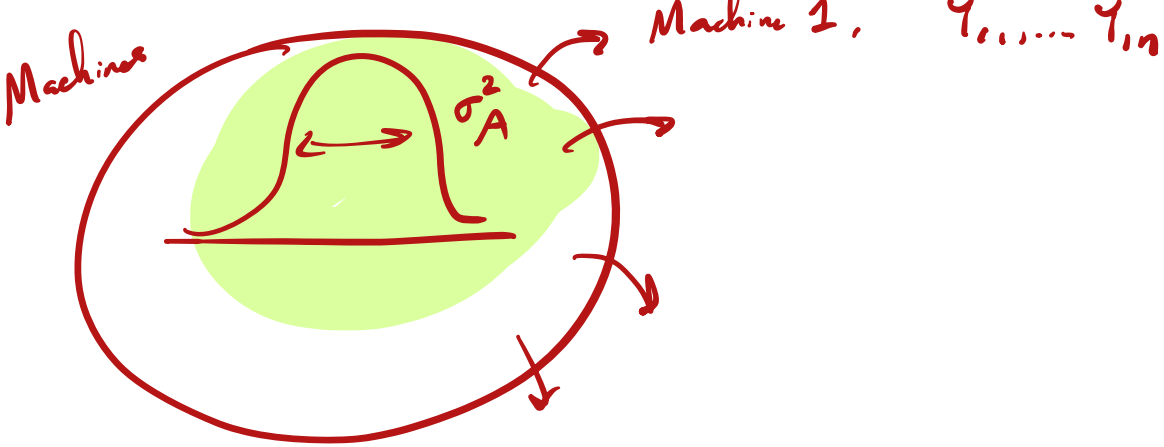
Fixed-effects model: $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$

"fixed"

Goals in the one-way random effects model

In the one-way random effects model we wish to

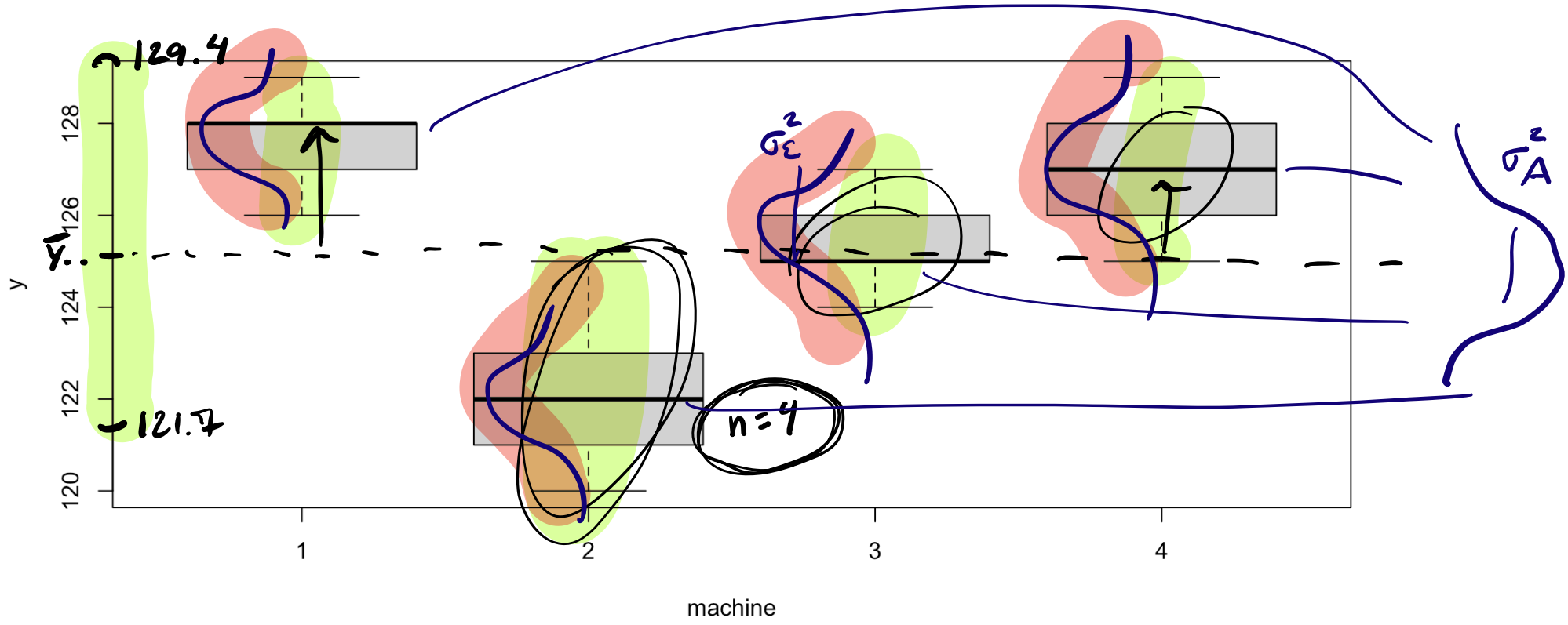
1. Visualize the data.
2. Decompose the variability in the Y_{ij} into its sources.
3. Test $H_0: \sigma_A^2 = 0$ versus $H_1: \sigma_A^2 > 0$.
4. Estimate the variance components σ_A^2 and σ_ε^2.
5. Estimate the overall mean μ.
6. "Predict" the realized values of $A_1, \ldots, A_a$.
7. Check whether the model assumptions are satisfied.



Fish net material

Side-by-side boxplots are a natural choice if the n_i are not too small.

```
1 boxplot(y~machine)
```

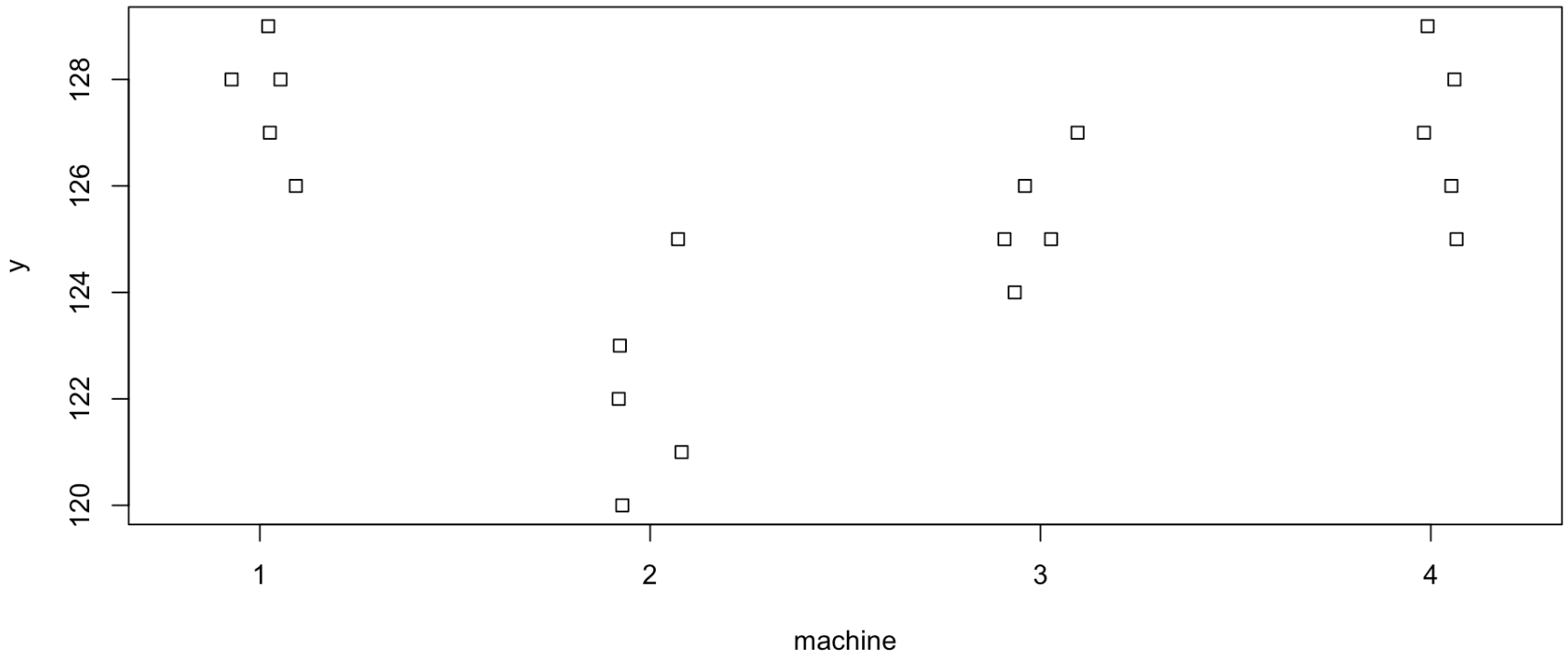


GeA $\hat{A}_i$

Fish net material

Dotplots are a better when the n_i are small.

```
1 stripchart(y~machine, vertical=T, method="jitter", xlab = "machine")
```



Sums of squares for one-way random effects model

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

Assuming $n_i = n$ for all $i = 1, \ldots, a$. (balanced design)

We have the decomposition $\mathrm{SS}_{\mathrm{Tot}} = \mathrm{SS}_{\mathrm{Trt}} + \mathrm{SS}_{\mathrm{Error}}$.

Fish net material

Easiest way to obtain sums of squares:

```
lm_out <- lm(y ~ machine)

y.. <- mean(y)
SStot <- sum((y - y..)**2)
SSerror <- sum(lm_out$residuals**2)
SStrt <- SStot - SSerror
```

We obtain $\mathrm{SS}_{\mathrm{Tot}} = 122.95$, $\mathrm{SS}_{\mathrm{Trt}} = 87.75$, and $\mathrm{SS}_{\mathrm{Error}} = 35.2$.

ANOVA table for one-way random effects model

Obtain the MS values by dividing the SS values by the Df values.

Source	Df	SS	MS	F value
Treatment	$a-1$	$\mathrm{SS}_{\mathrm{Trt}}$	$\mathrm{MS}_{\mathrm{Trt}}$	$\mathrm{F}_{\mathrm{test}} = \mathrm{MS}_{\mathrm{Trt}} \,/\, \mathrm{MS}_{\mathrm{Error}}$
Error	$a(n-1)$	$\mathrm{SS}_{\mathrm{Error}}$	$\mathrm{MS}_{\mathrm{Error}}$	
Total	$an-1$	$\mathrm{SS}_{\mathrm{Tot}}$		

We reject $H_0 \colon \sigma_A^2 = 0$ if $\mathrm{F}_{\mathrm{test}} > F_{a-1,a(n-1),\alpha}$.

The corresponding p-value is $P(F > \mathrm{F}_{\mathrm{test}})$, where $F \sim F_{a-1,a(n-1),\alpha}$.

Discuss: It can be shown that $\mathrm{SS}_{\mathrm{Trt}}$ and $\mathrm{SS}_{\mathrm{Error}}$ are independent and

$$\frac{\mathrm{SS}_{\mathrm{Trt}}}{n\sigma_A^2 + \sigma_\varepsilon^2} \sim \chi^2_{a-1} \quad \text{and} \quad \frac{\mathrm{SS}_{\mathrm{Error}}}{\sigma_\varepsilon^2} \sim \chi^2_{a(n-1)}.$$

Show that $\mathrm{F}_{\mathrm{test}} = \dfrac{\mathrm{MS}_{\mathrm{Trt}}}{\mathrm{MS}_{\mathrm{Error}}} \sim F_{a-1,a(n-1)}$ under $H_0\colon \sigma_A^2 = 0$.

extra

Fish net material

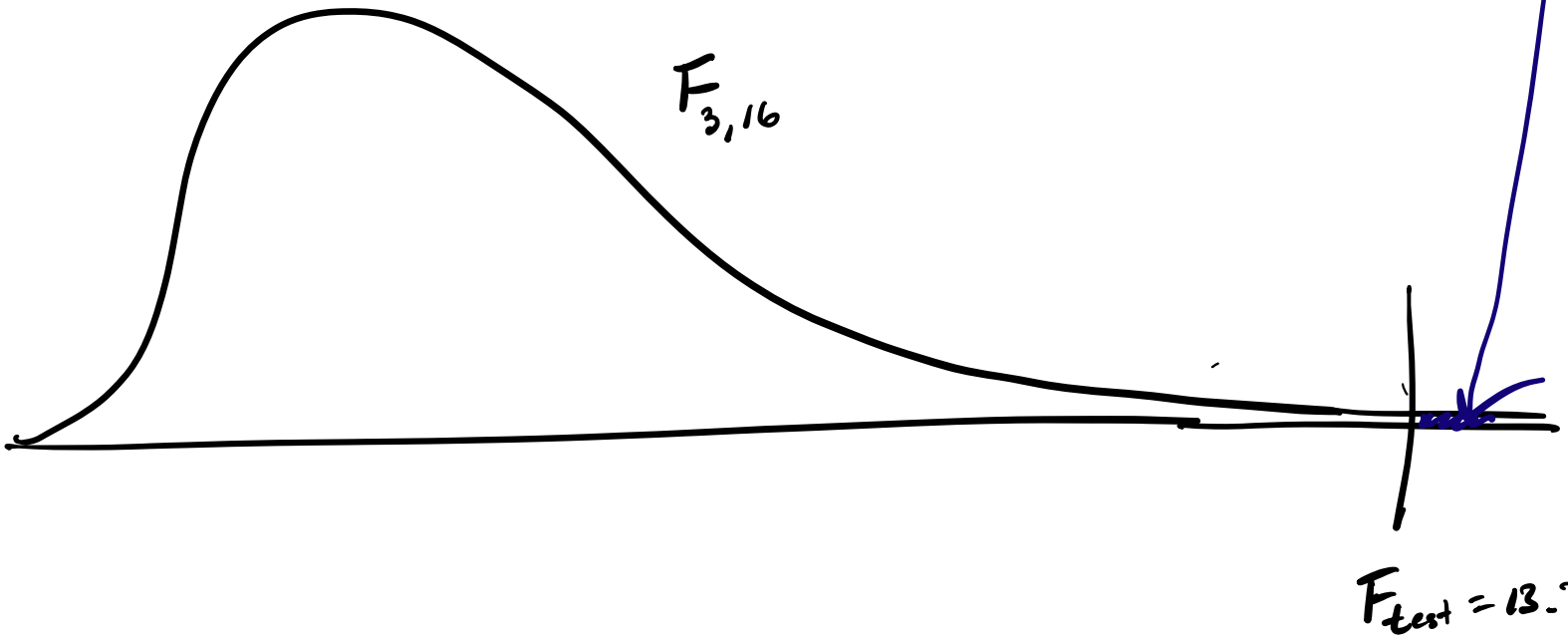
We obtain the ANOVA table just as we did in the fixed-effects case.

```
1 anova(lm_out)
```

```
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value  Pr(>F)
machine    3   87.7    29.2    13.3 0.00013 ***
Residuals 16   35.2     2.2
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The small p-value indicates strong evidence against $H_0: \sigma_A^2 = 0$.



Expected MS values in one-way random-effects model

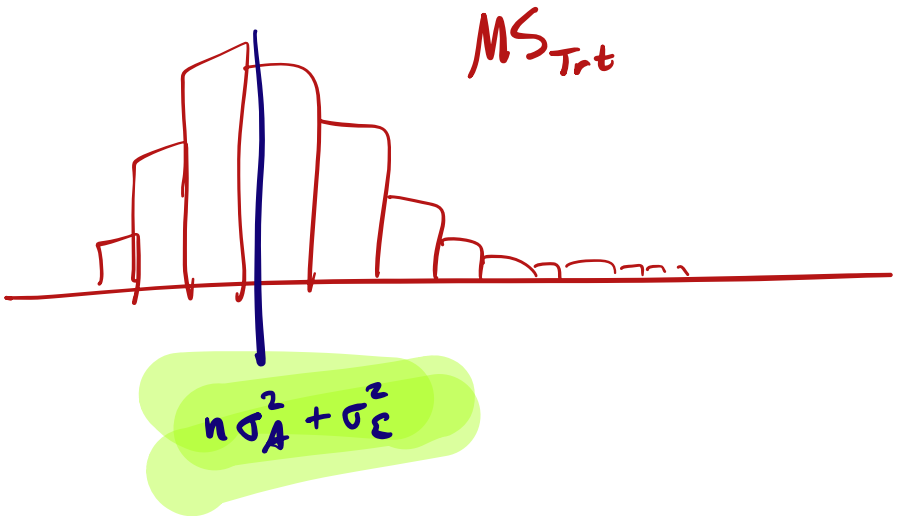
In models with random effects, we like to track how the variance components σ^2_A and σ^2_ε affect the expected values of the mean squares.

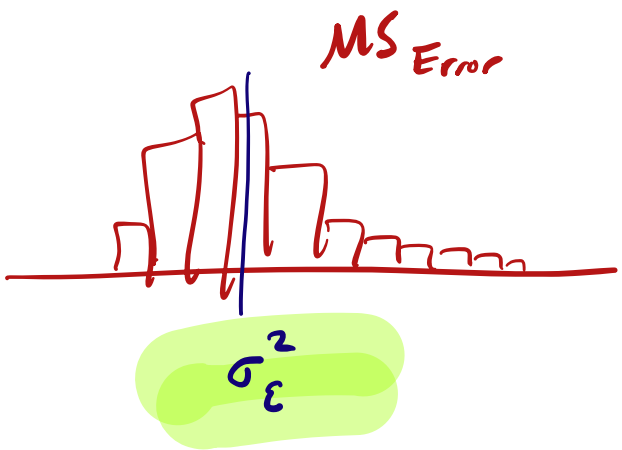
Source	Df	Expected mean square
Treatment	$a-1$	$n\sigma^2_A + \sigma^2_\varepsilon$
Error	$a(n-1)$	σ^2_ε

$\mathbb{E}\, MS_{Trt} = \mathbb{E}\left[\frac{SS_{Trt}}{a-1}\right]$

$\mathbb{E}\, MS_{Error}$

"$\mathbb{E}$" means "expected value" or the average over many repeats of the experiment.





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

$n\hat{\sigma}_A^2 + \hat{\sigma}_\varepsilon^2 = MS_{Trt}$

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

$\hat{\sigma}_A^2 = ?$

Solve for $\hat{\sigma}_A^2$. I get

$$\hat{\sigma}_A^2 = \frac{MS_{Trt} - MS_{Error}}{n}$$

Method of moments for estimating σ^2_A and σ^2_ε

The method of moments (MoM) is one way to estimate the variance components.

1. Set the observed mean squares equal to the expected mean squares:

$$\mathrm{MS}_{\mathrm{Trt}} \overset{\text{set}}{=} n\sigma^2_A + \sigma^2_\varepsilon \quad \text{and} \quad \mathrm{MS}_{\mathrm{Error}} \overset{\text{set}}{=} \sigma^2_\varepsilon$$

2. Solve these equations for σ^2_A and σ^2_ε, which gives

$$\dot{\sigma}^2_A = \frac{\mathrm{MS}_{\mathrm{Trt}} - \mathrm{MS}_{\mathrm{Error}}}{n} \quad \text{and} \quad \dot{\sigma}^2_\varepsilon = \mathrm{MS}_{\mathrm{Error}}.$$

MoM method is deprecated because it is possible to get $\dot{\sigma}^2_A < 0$.

Fish net material

Compute the MoM estimators $\dot{\sigma}_A^2$ and $\dot{\sigma}_\varepsilon^2$ for the fish net data.

```
n <- 5
a <- 4
MStrt <- SStrt / (a - 1)
MSerror <- SSerror / (a*(n-1))

sigma_e_mom <- sqrt(MSerror)
sigma_A_mom <- sqrt((MStrt - MSerror)/n)
```

We obtain $\dot{\sigma}_A^2 = 5.41$ and $\dot{\sigma}_\varepsilon^2 = 2.2$.

REML for estimating the variance components

- REML stands for restricted maximum likelihood.
- It is a recipe (beyond course) for estimating variance components.
- In the one-way random-effects model it results in

$$\hat{\sigma}_A^2 = \begin{cases} \dot{\sigma}_A^2 & \text{if } \dot{\sigma}_A^2 \geq 0 \\ 0 & \text{if } \dot{\sigma}_A^2 < 0 \end{cases} \text{ and } \hat{\sigma}_\varepsilon^2 = \begin{cases} \dot{\sigma}_\varepsilon^2 & \text{if } \dot{\sigma}_A^2 \geq 0 \\ \dfrac{\mathrm{SS}_{\mathrm{Tot}}}{an-1} & \text{if } \dot{\sigma}_A^2 < 0 \end{cases}$$

- Will always return a nonnegative estimate of σ_A^2.
- Same as MoM estimator if $\dot{\sigma}_A^2 > 0$.
- The “state of the art” is to use REML estimation.

Fish net material

Use R package `lmerTest` to compute the REML estimates of σ_A^2 and σ_ε^2.

```
# first time run install.packages("lmerTest")
library(lmerTest) # may need to run install.packages("Matrix") for this to work
lmer_out <- lmer(y ~ 1 + (1 | machine))
summary(lmer_out)
```

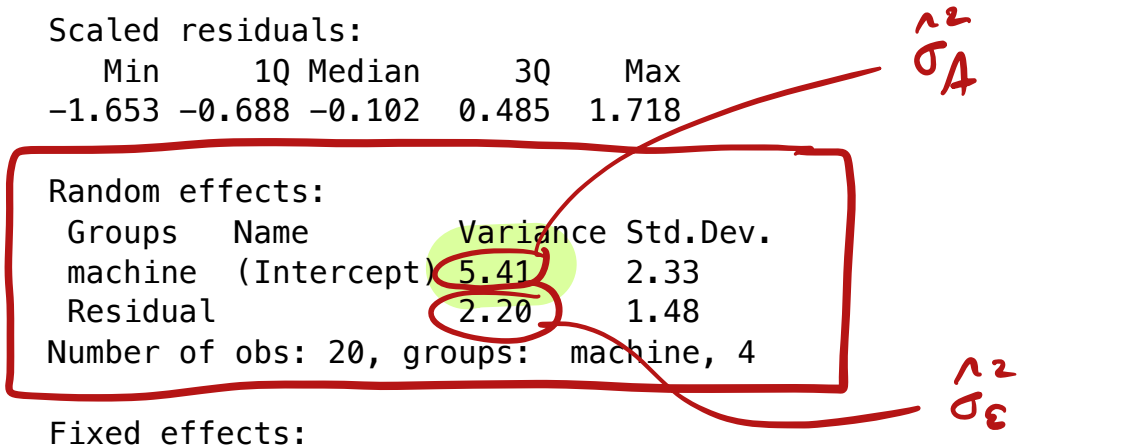
```
Linear mixed model fit by REML. t-tests use Satterthwaite's method [
lmerModLmerTest]
Formula: y ~ 1 + (1 | machine)

REML criterion at convergence: 79.7

Scaled residuals:
   Min     1Q Median     3Q    Max
-1.653 -0.688 -0.102  0.485  1.718

Random effects:
 Groups   Name        Variance Std.Dev.
 machine  (Intercept) 5.41     2.33
 Residual             2.20     1.48
Number of obs: 20, groups:  machine, 4

Fixed effects:
            Estimate Std. Error     df t value Pr(>|t|)
(Intercept)   125.55       1.21   3.00     104    2e-06 ***
```



Estimating μ and predicting $A_1, \ldots, A_a$.

- We estimate the overall mean μ with $\bar{Y}_{\cdot\cdot}$.
- We also wish to guess at the realized values of $A_1, \ldots, A_a$.
- Call this *prediction* instead of estimation, because $A_1, \ldots, A_a$ are random variables rather than parameters with fixed values.
- A recipe (beyond course) for the best linear unbiased predictor yields

$$\hat{A}_i = \left(\frac{n\hat{\sigma}_A^2}{n\hat{\sigma}_A^2 + \hat{\sigma}_\varepsilon^2} \right) (\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot}) \quad \text{for } i = 1, \ldots, a.$$

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

For large n, $\hat{A}_i \approx \bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot}$

"Fitted values": $\hat{Y}_{ij} = \hat{\mu} + \hat{A}_i = \bar{Y}_{\cdot\cdot} + \hat{A}_i$

$\hat{\mu} = \bar{Y}_{\cdot\cdot}$

Fish net material

Obtain the predicted values of A_1, A_2, A_3, and A_4 for the fish net data.

```
yi. <- aggregate(y, by = list(machine = machine), mean)$x
# use MoMs, since sA2_dot > 0
Ahat <- (n * sigma_A_mom**2)/(n * sigma_A_mom**2 + sigma_e_mom**2) * (yi. - y..)
Ahat
```

```
[1]  1.896 -3.098 -0.139  1.341
```

Fish net material

Can use `predict()` on the `lmer()` output to obtain the $\hat{A}_i$ values:

```
1  predict(lmer_out,random.only = TRUE) # gives estimates of the A_i for each obs
```

```
     1      2      3      4      5      6      7      8      9     10     11
 1.896  1.896  1.896  1.896  1.896 -3.098 -3.098 -3.098 -3.098 -3.098 -0.139
    12     13     14     15     16     17     18     19     20
-0.139 -0.139 -0.139 -0.139  1.341  1.341  1.341  1.341  1.341
```

In Oneway ANOVA with fixed effects

Fitted values $\hat{y}_{ij} = \bar{y}_{i\cdot}$

Residuals $\hat{\varepsilon}_{ij} = y_{ij} - \bar{y}_{i\cdot}$

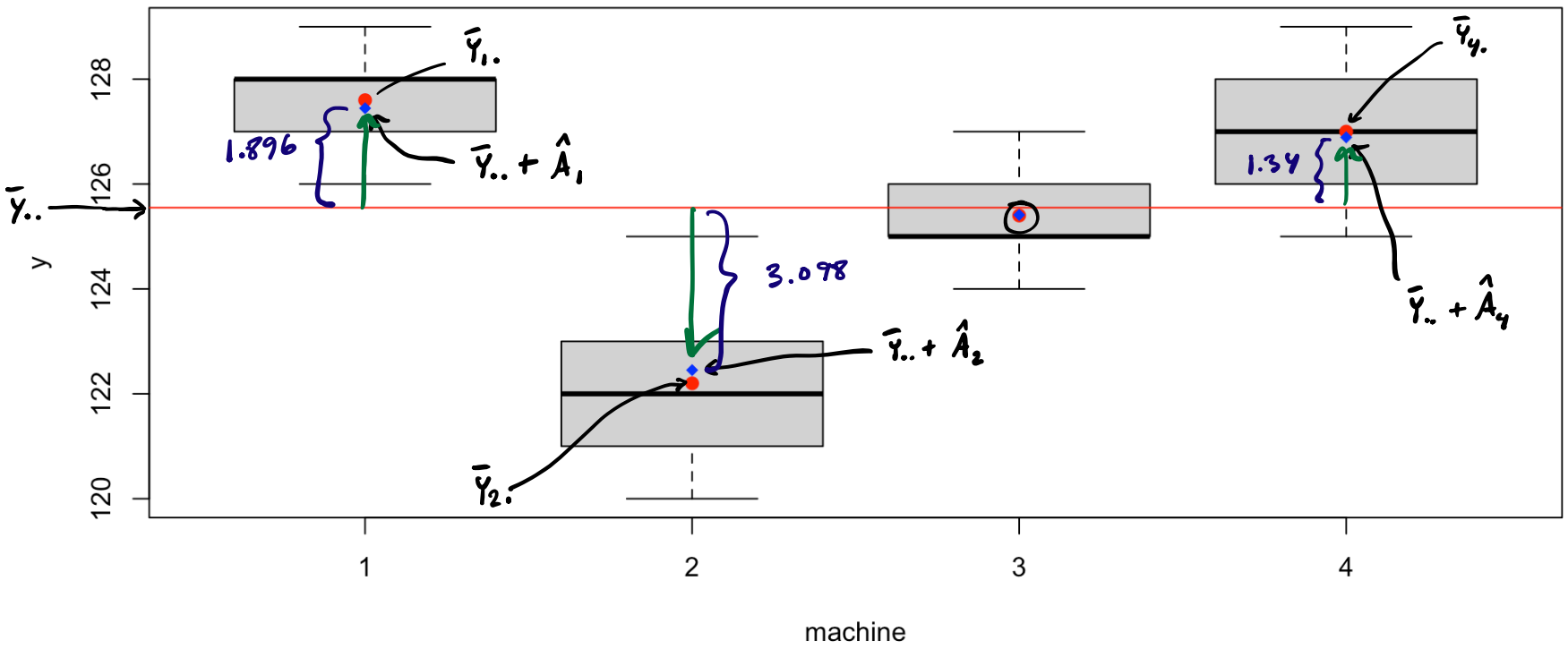
Oneway random effects model

Fitted value $\hat{y}_{ij} = \bar{y}_{\cdot\cdot} + \hat{A}_i$

Residuals $\hat{\varepsilon}_{ij} = y_{ij} - \hat{y}_{ij} = y_{ij} - \left(\bar{y}_{\cdot\cdot} + \hat{A}_i\right)$

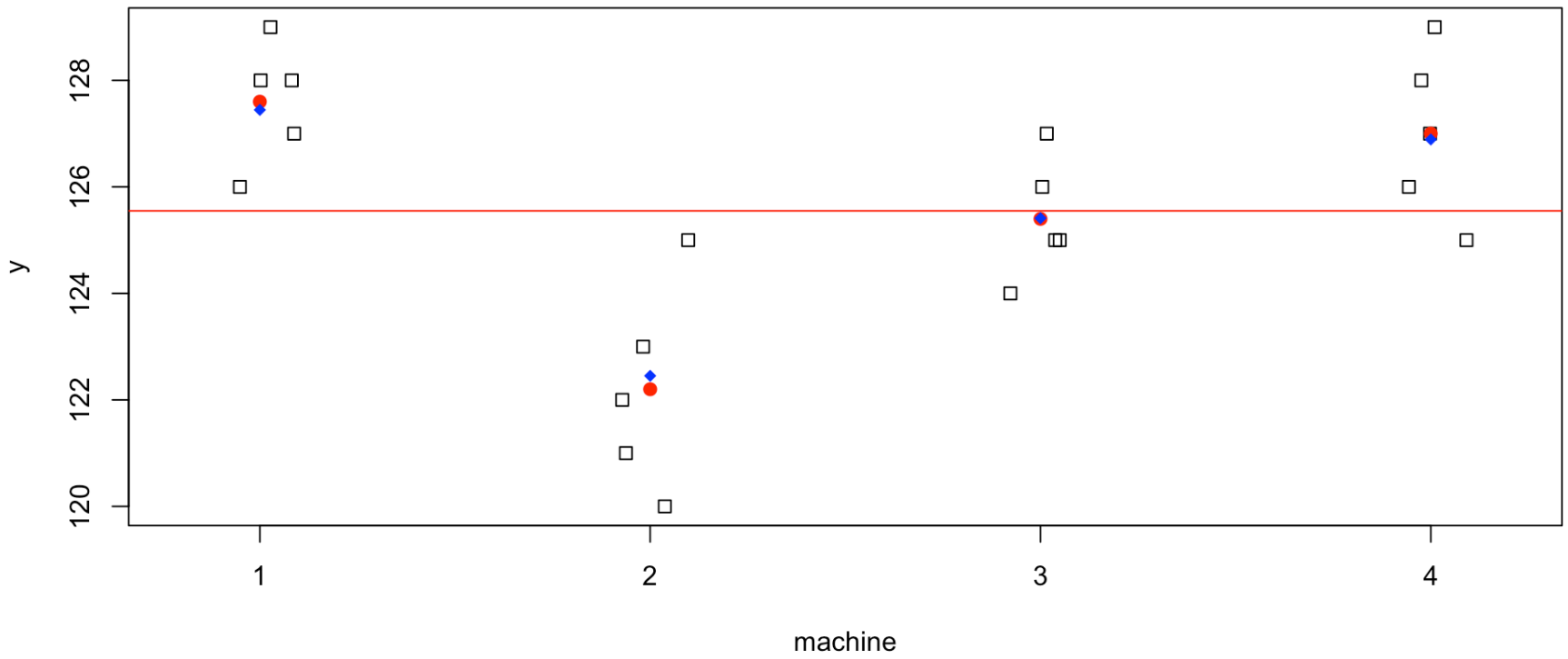
Fish net material

```
boxplot(y~machine)
points(yi.~as.factor(levels(machine)),col="red",pch=19)
points((y.. + Ahat)~as.factor(levels(machine)),col="blue",pch=18)
abline(h = y..,col = "red")
```



Fish net material

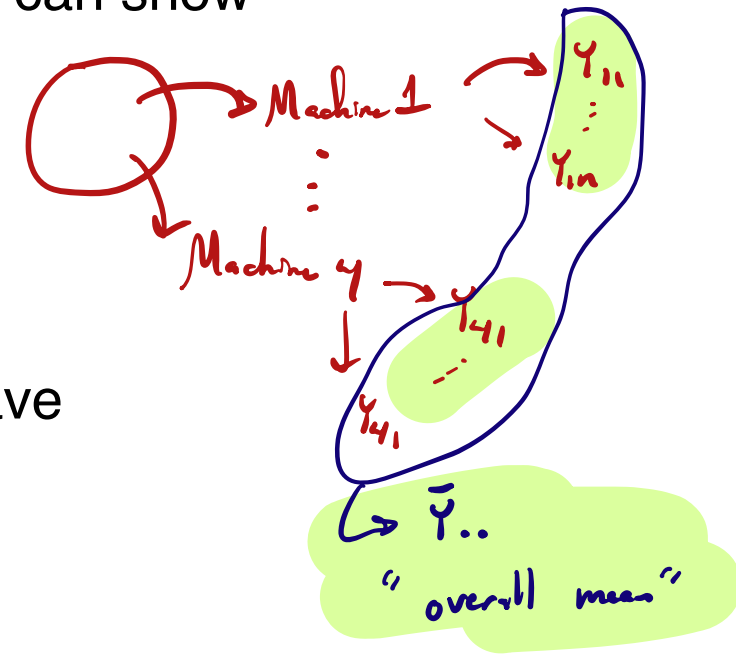
```
stripchart(y~machine, vertical = T, method = "jitter", xlab = "machine")
points(yi.~as.factor(levels(machine)),col="red",pch=19)
points((y.. + Ahat)~as.factor(levels(machine)),col="blue",pch=18)
abline(h = y..,col = "red")
```



Confidence interval for the overall mean

- Under the one-way random effects model, we can show

$$\bar{Y}_{\cdot\cdot} \sim \mathcal{N}\left(\mu, \frac{n\sigma_A^2 + \sigma_\varepsilon^2}{na}\right).$$



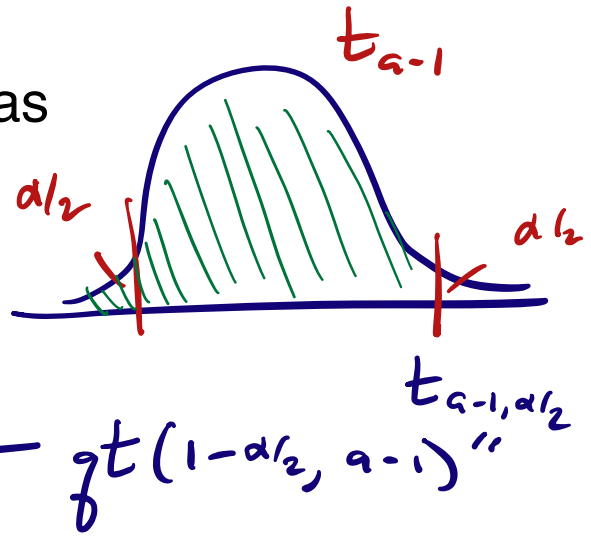
- From here, using $\dfrac{\mathrm{SS}_{\mathrm{Trt}}}{n\sigma_A^2 + \sigma_\varepsilon^2} \sim \chi^2_{a-1}$, we have

$$\frac{\bar{Y}_{\cdot\cdot} - \mu}{\sqrt{\mathrm{MS}_{\mathrm{Trt}}/(na)}} \sim t_{a-1}.$$

- Thus a $(1-\alpha)100\%$ CI for μ can be obtained as

$$\bar{Y}_{\cdot\cdot} \pm t_{a-1,\alpha/2}\sqrt{\frac{\mathrm{MS}_{\mathrm{Trt}}}{an}}.$$

ME



Fish net material

Construct a 95% confidence interval for μ in fish net example.

```
alpha <- 0.05
tval <- qt(1-alpha/2,a-1)
me <- tval * sqrt(MStrt/(a*n))
lo <- y.. - me
up <- y.. + me
```

We obtain the interval $[121.701, 129.399]$.

Fish net material

Default confidence intervals from `confint()` function.

Will be slightly different because `confint()` function on `lmer()` output uses another method.

```
confint(lmer_out)
```

```
              2.5 % 97.5 %
.sig01        0.968   5.09
.sigma        1.087   2.19
(Intercept) 122.890 128.21
```

Fish net material

Wald-type CI $Y_{\cdot\cdot} \pm z_{\alpha/2}\sqrt{\mathrm{MS}_{\mathrm{Trt}}/(na)}$ on fish net data.

```
zval <- qnorm(1-alpha/2)
me <- zval * sqrt(MStrt/(a*n))
lo <- y.. - me
up <- y.. + me
```

We obtain the interval $[123.18, 127.92]$.

Compare to:

```
confint(lmer_out,method="Wald")
```

```
            2.5 % 97.5 %
.sig01         NA     NA
.sigma         NA     NA
(Intercept)   123    128
```

Checking model assumptions

Validity of the methods in these slides depends on these assumptions:

1. The responses are normally distributed around the treatment means.

   (Check QQ plot of residuals)

2. The treatment means are normally distributed.

   (The number of groups a is often too small for this to be checked)

3. The response has the same variance in all treatment groups.

   (Check residuals vs fitted values plot)

4. The response values in each treatment group are independent.

   (No way to check; must trust experimental design)

Fitted values and residuals

- Build fitted values from $\hat{\mu}$ and the predictions $\hat{A}_1, \ldots, \hat{A}_a$ as

$$\hat{Y}_{ij} = \hat{\mu} + \hat{A}_i, \quad \text{for } i = 1, \ldots, a, \quad j = 1, \ldots, n.$$

- Our residuals then become

$$\hat{\varepsilon}_{ij} = Y_{ij} - \hat{Y}_{ij}, \quad \text{for } i = 1, \ldots, a, \quad j = 1, \ldots, n.$$

- These residuals are not simply the deviations of the responses around their treatment group means (as in fixed effects models)!
- So $\mathrm{SS}_{\mathrm{Error}}$ is not the sum of squared residuals here.

Fish net material

Use `predict()` on the `lmer()` output to obtain the $\hat{Y}_{ij} = \hat{\mu} + \hat{A}_i$.

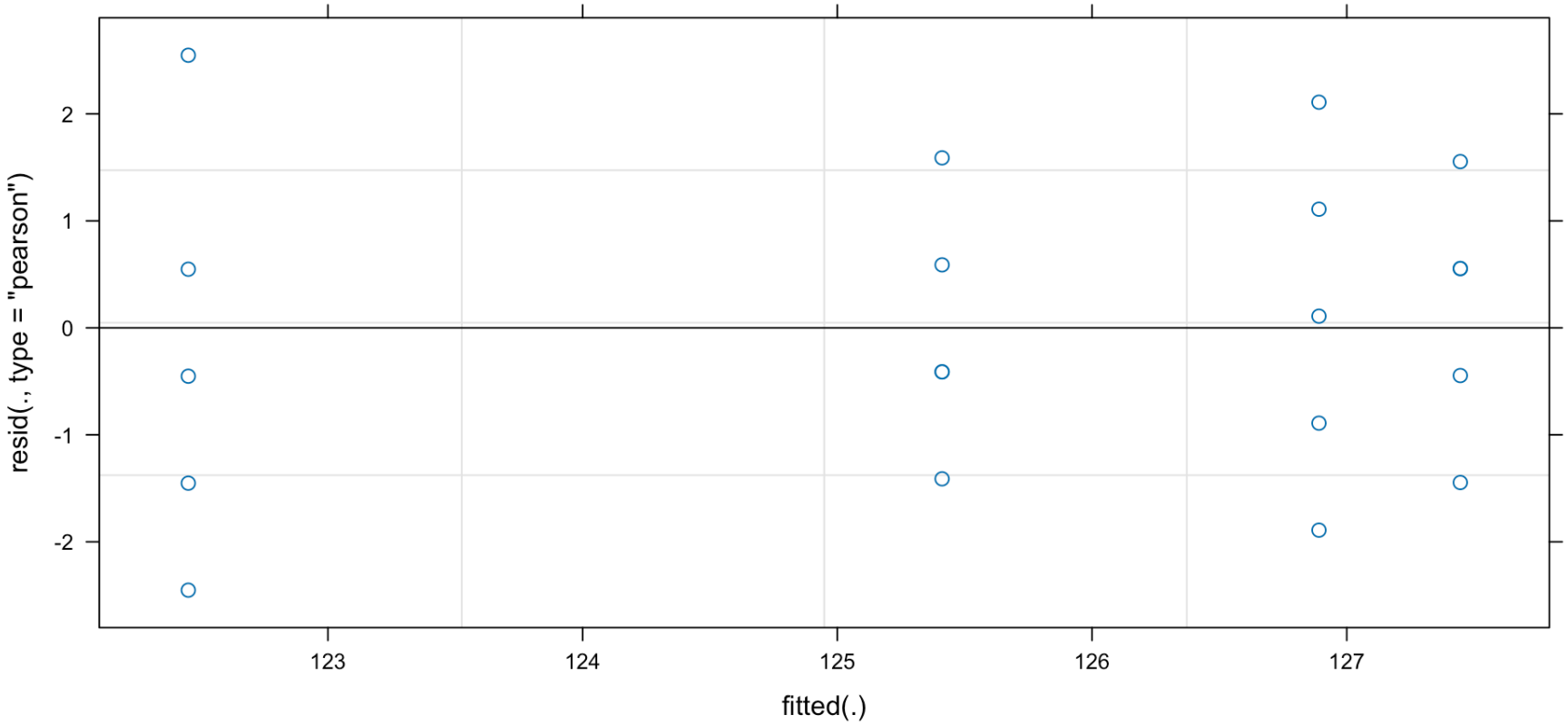
```
predict(lmer_out)
```

```
  1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
127 127 127 127 127 122 122 122 122 122 125 125 125 125 125 127 127 127 127 127
```

Fish net material

Check the residuals vs fitted values plot: Use `plot()` on `lmer()` output.

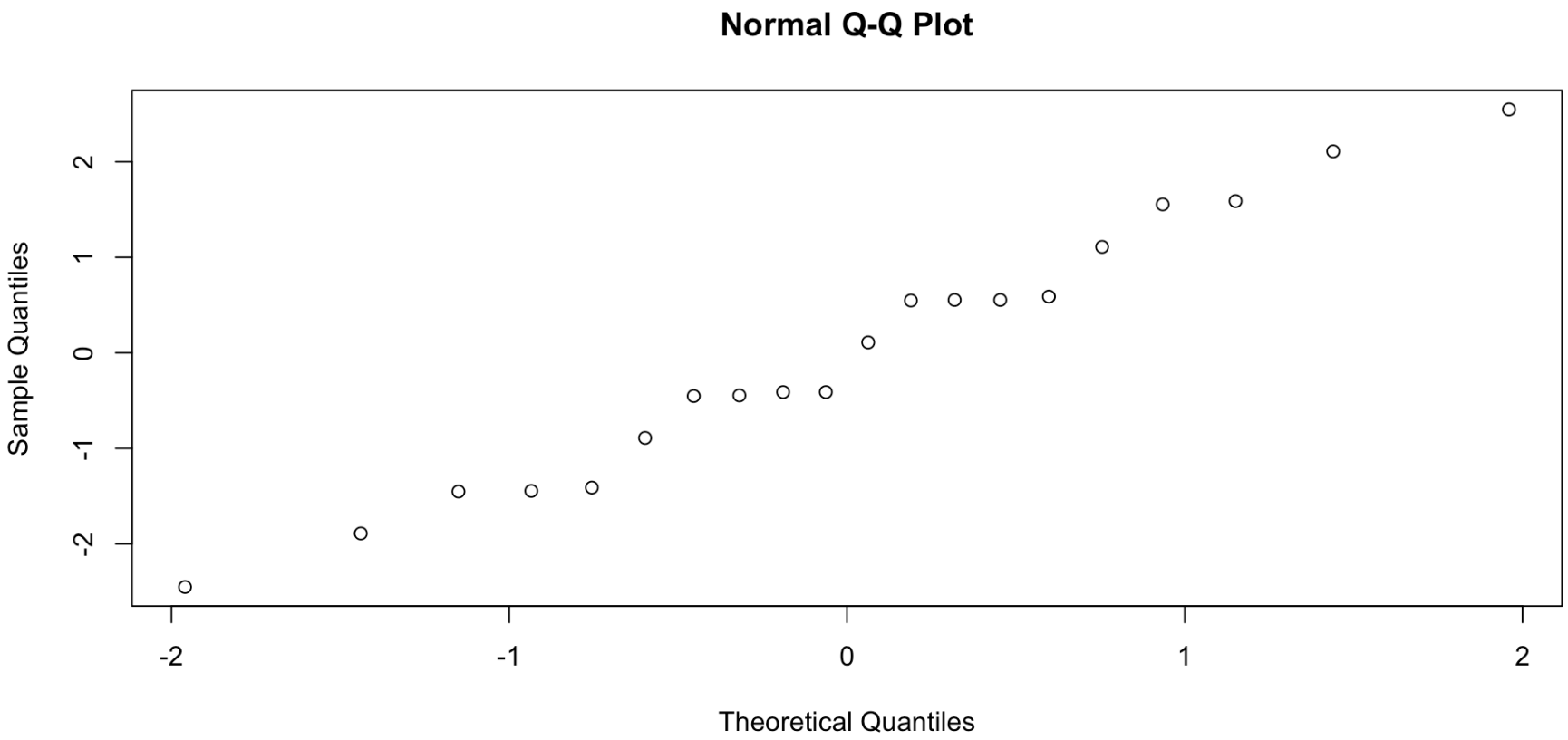
```
1 plot(lmer_out)
```



Fish net material

Check the normal q-q plot of the residuals:

```
1 yhat <- predict(lmer_out)
2 ehat <- y - yhat
3 qqnorm(ehat)
```



Below material optional.

Triglyceride level data from Kuehl (2000)

Triglyceride levels in serum samples measured with 4 machines on 4 days.

	Machine			
Day	1	2	3	4
1	142.3, 144.0	148.6, 146.9	142.9, 147.4	133.8, 133.2
2	134.9, 146.3	145.2, 146.3	125.9, 127.6	108.9, 107.5
3	148.6, 156.5	148.6, 153.1	135.5, 138.9	132.1, 149.7
4	152.0, 151.4	149.7, 152.0	142.9, 142.3	141.7, 141.2

Source: Dr. J. Anderson, Beckman Instruments, Inc.

```
tg <- c(142.3,144.0,148.6,146.9,142.9,147.4,133.8,133.2,
        134.9,146.3,145.2,146.3,125.9,127.6,108.9,107.5,
        148.6,156.5,148.6,153.1,135.5,138.9,132.1,149.7,
        152.0,151.4,149.7,152.0,142.9,142.3,141.7,141.2)
days <- as.factor(c(rep(1,8),rep(2,8),rep(3,8),rep(4,8)))
machine <- as.factor(c(1,1,2,2,3,3,4,4,1,1,2,2,3,3,4,4,
                       1,1,2,2,3,3,4,4,1,1,2,2,3,3,4,4))
```

Machines randomly sampled. Regard days as randomly sampled.

Two-way random effects model

Assume

$$Y_{ijk} = \mu + A_i + B_j + (AB)_{ij} + \varepsilon_{ijk}$$

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

- $A_i \overset{\text{ind}}{\sim} \mathcal{N}(0, \sigma^2_A)$
- $B_j \overset{\text{ind}}{\sim} \mathcal{N}(0, \sigma^2_B)$
- $(AB)_{ij} \overset{\text{ind}}{\sim} \mathcal{N}(0, \sigma^2_{AB})$
- $\varepsilon_{ijk} \overset{\text{ind}}{\sim} \mathcal{N}(0, \sigma^2_\varepsilon)$
- the A_i, B_j, and ε_{ijk} are all independent.
- μ is the overall mean.
- we call σ^2_A, σ^2_B, σ^2_{AB}, and σ^2_ε variance components.

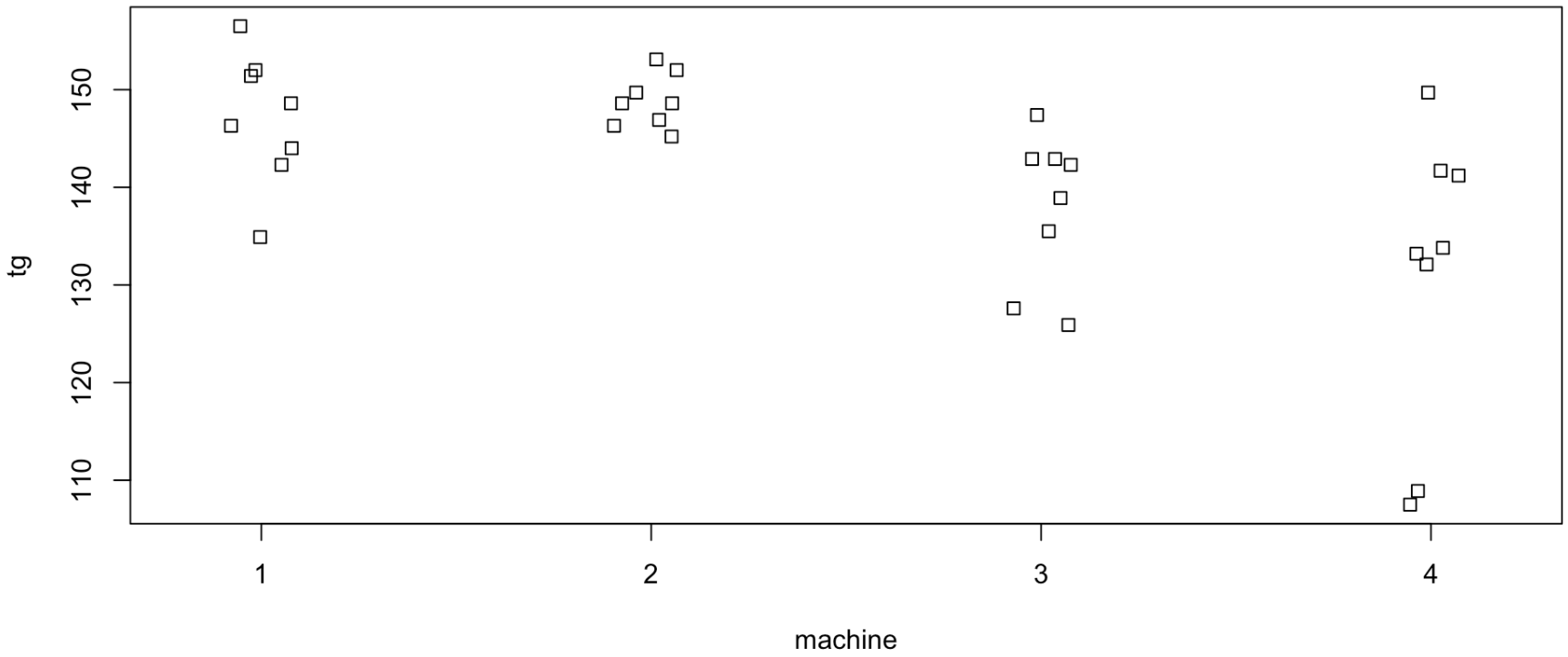
Goals in the two-way random effects model

In the two-way random effects model we wish to

1. Visualize the data.
2. Decompose the variability in the Y_{ijk} into its sources.
3. Test whether σ^2_A, σ^2_B, and σ^2_{AB} are equal to zero.
4. Estimate the variance components σ^2_A, σ^2_B, σ^2_{AB}, and σ^2_ε.
5. Estimate the overall mean μ.
6. “Predict” the realized values of the random effects.
7. Check whether the model assumptions are satisfied.

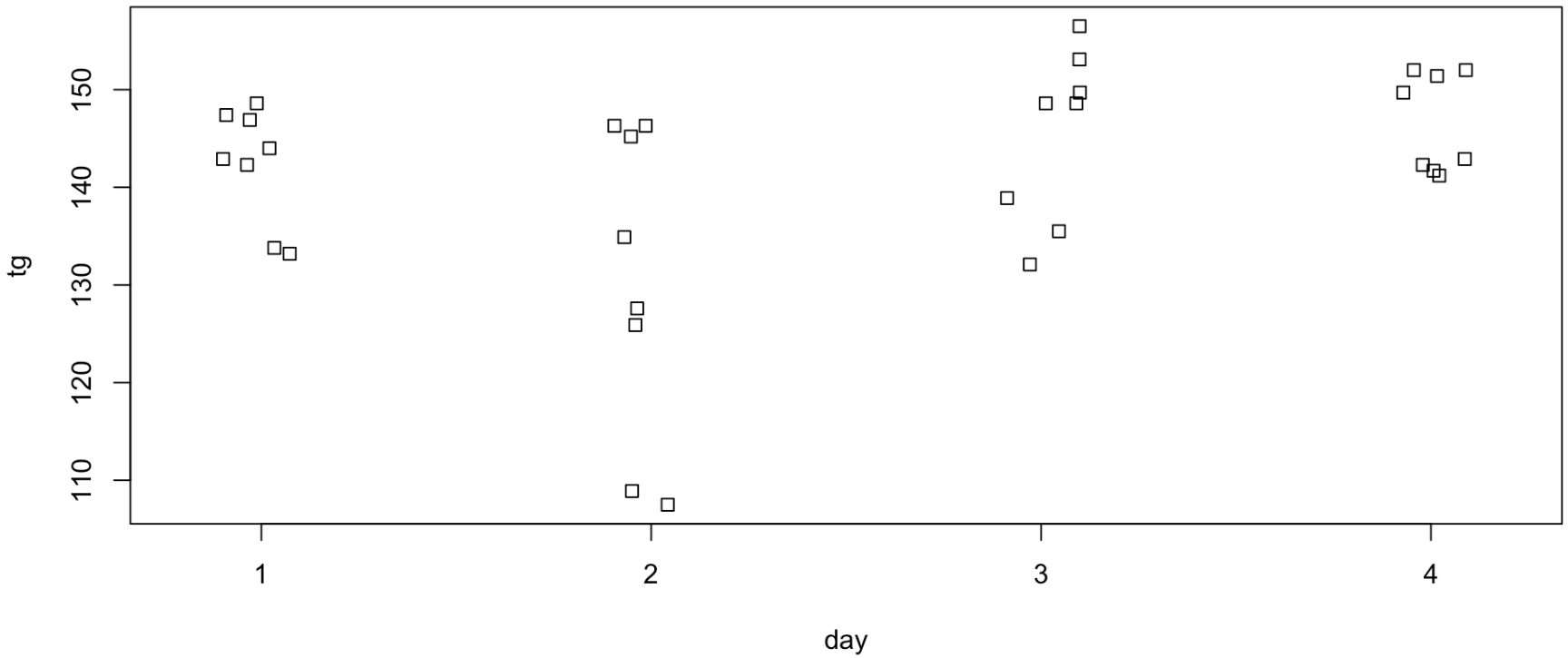
Triglyceride level

```
1 stripchart(tg~machine,vertical = T, method = "jitter", xlab = "machine")
```



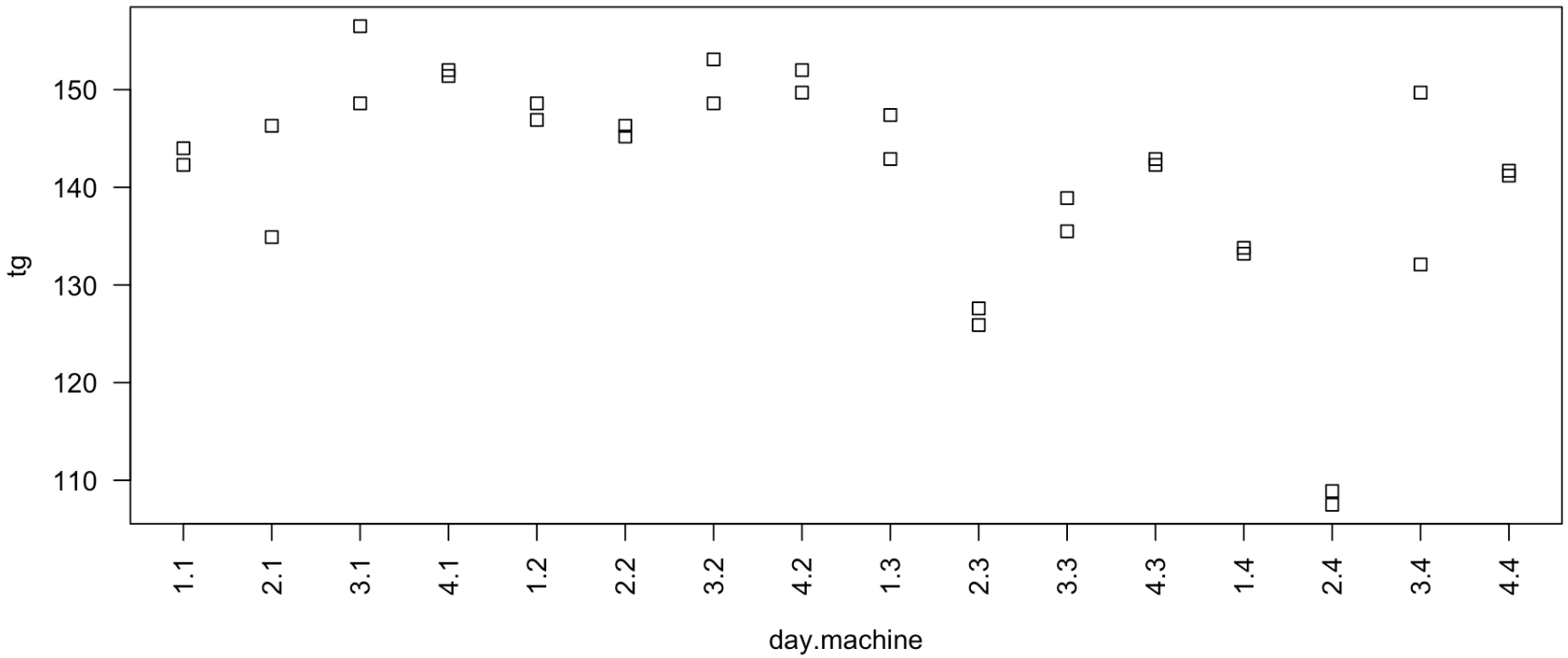
Triglyceride level

```
1 stripchart(tg~days,vertical=T, method = "jitter", xlab = "day")
```



Triglyceride level

```
1 stripchart(tg~days + machine,vertical = T, xlab = "day.machine",las = 2)
```



SS for the two-way random effects model

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

We have the decomposition $\mathrm{SS}_{\mathrm{Tot}} = \mathrm{SS}_{\mathrm{A}} + \mathrm{SS}_{\mathrm{B}} + \mathrm{SS}_{\mathrm{AB}} + \mathrm{SS}_{\mathrm{Error}}$.

Full ANOVA table for balanced two-way fact. design

Obtain the MS values by dividing the SS values by the Df values.

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

1. Reject H_0: $\sigma_A^2 = 0$ 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 $F_{\mathrm{AB}} > F_{(a-1)(b-1),ab(n-1),\alpha}$.

Note F_{A} and F_{B} are different from their fixed-effects counterparts!

Expected MS in two-way random-effects model

We again tabulate the expected values of the mean squares:

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

Discuss: Make F-stat by dividing by the MS having the same expectation under H_0.

Triglyceride level

Obtain p-values for testing whether the variance components are zero.

The `anova()` function on the `lm()` output only gives the correct F statistic and p-value for the interaction.

```
1 anova_out <- anova(lm(tg ~ days + machine + days:machine))
2 anova_out
```

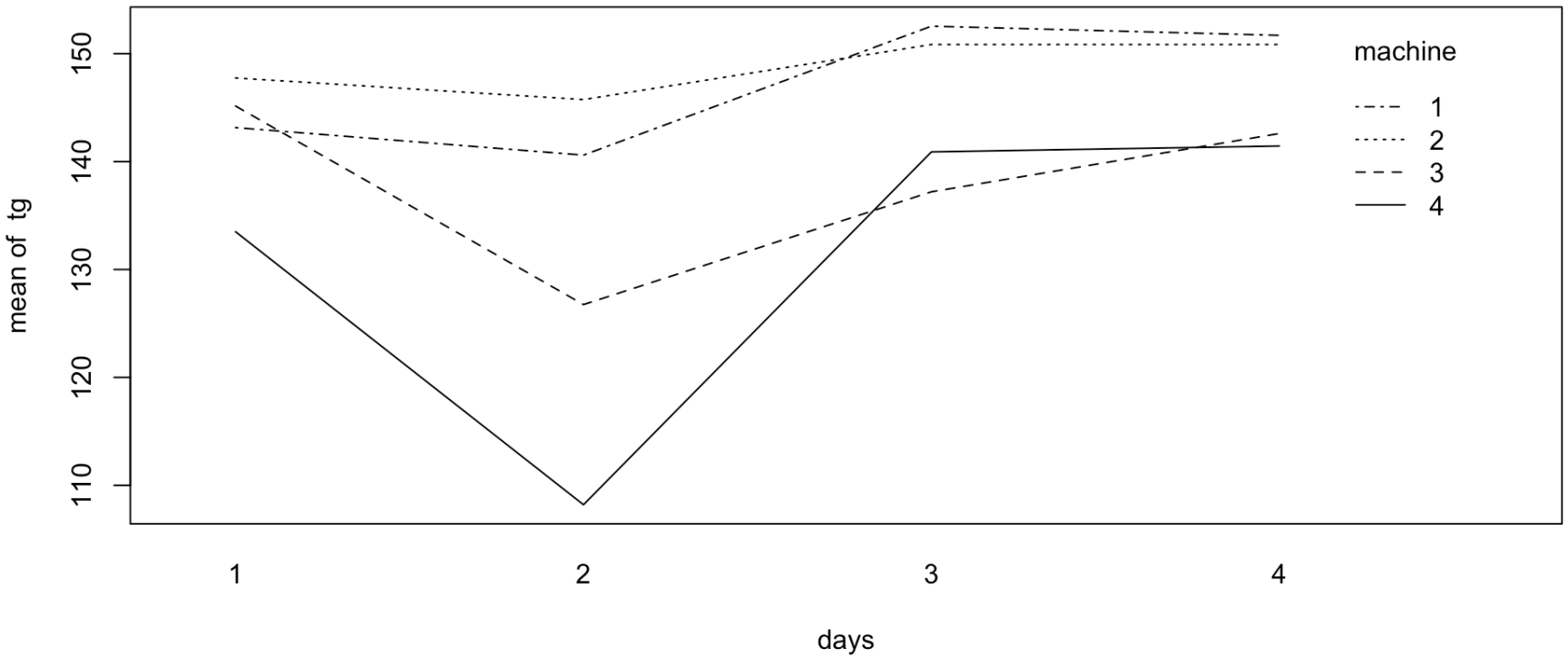
```
Analysis of Variance Table

Response: tg
             Df Sum Sq Mean Sq F value  Pr(>F)
days          3   1334     445   24.86 2.9e-06 ***
machine       3   1647     549   30.68 7.2e-07 ***
days:machine  9    786      87    4.88  0.0029 **
Residuals    16    286      18
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the p-value for H_0: $\sigma^2_{AB} = 0$, we conclude there is interaction between the day and the machine. So machine performance does not appear to be consistent across days.

Triglyceride level

```
1 interaction.plot(days,machine,tg)
```



Triglyceride level

```r
a <- 4
b <- 4
n <- 2
MSA <- anova_out$`Mean Sq`[1]
MSB <- anova_out$`Mean Sq`[2]
MSAB <- anova_out$`Mean Sq`[3]
FA <- MSA / MSAB
FB <- MSB / MSAB
pA <- 1 - pf(FA,a - 1,(a-1)*(b-1))
pB <- 1 - pf(FB,b - 1,(a-1)*(b-1))
```

For H_0: $\sigma_A^2 = 0$ we obtain a p-value of 0.0248.

For H_0: $\sigma_B^2 = 0$ we obtain a p-value of 0.0137.

There is significant machine-to-machine and day-to-day variation.

Estimation of σ^2_A, σ^2_B, σ^2_{AB}, and σ^2_ε

- Could use the MoM based on the table of expected mean squares.
- MoM can result is negative estimates of σ^2_A, σ^2_B, or σ^2_{AB}.
- REML estimation is prescribed.
- Estimate μ with the overall data mean $\bar{Y}_{\cdots}$.

Triglyceride level

Obtain the REML estimators of σ^2_A, σ^2_B, σ^2_{AB}, and σ^2_ε.

Use `summary()` on the `lmer()` output.

```
1 lmer_out <- lmer(tg ~ 1 + (1|days) + (1|machine) + (1|days:machine))
2 summary(lmer_out)
```

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method [
lmerModLmerTest]
Formula: tg ~ 1 + (1 | days) + (1 | machine) + (1 | days:machine)

REML criterion at convergence: 215

Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.8428 -0.3558  0.0348  0.2070  2.3177

Random effects:
 Groups       Name        Variance Std.Dev.
 days:machine (Intercept) 34.7     5.89
 machine      (Intercept) 57.7     7.60
 days         (Intercept) 44.7     6.68
 Residual                 17.9     4.23
Number of obs: 32, groups:  days:machine, 16; machine, 4; days, 4

Fixed effects:
```

Prediction of realized values of random effects

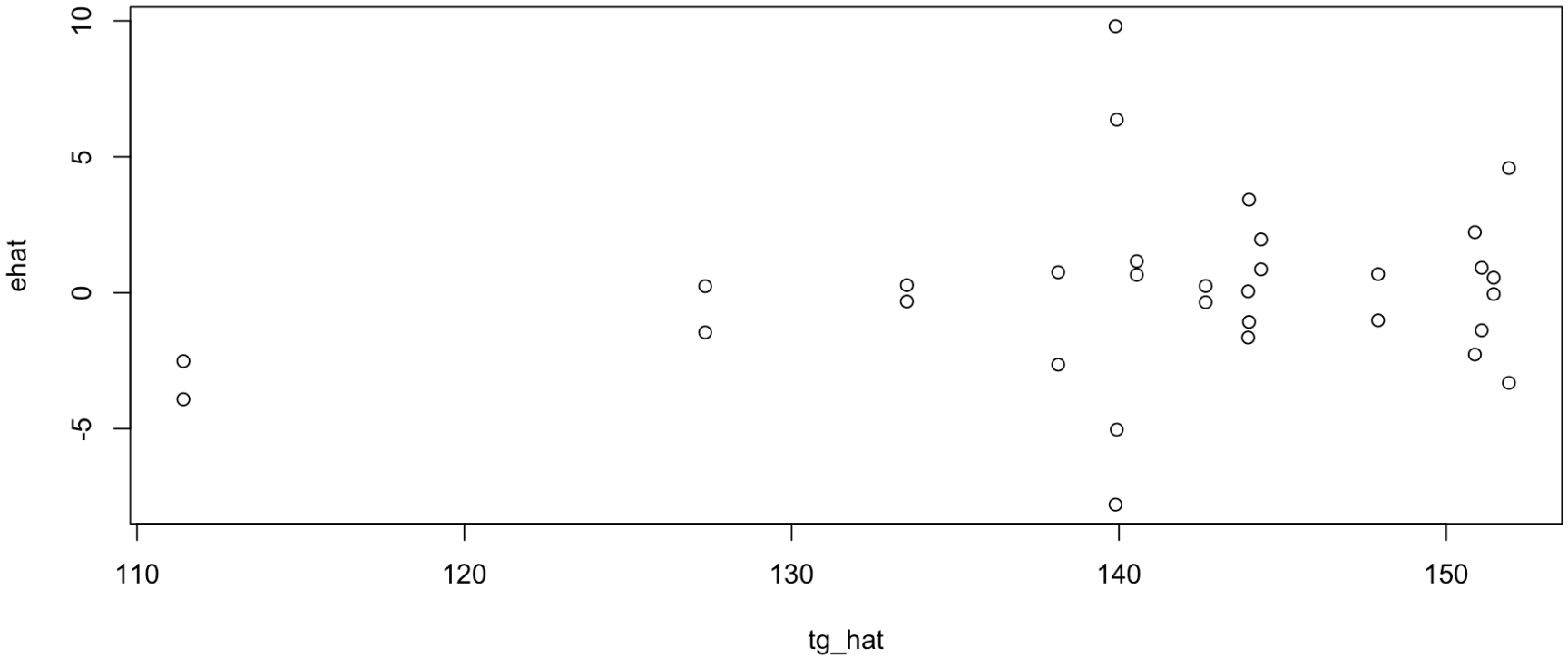
- We wish to guess at the values of A_i, B_j, and $(AB)_{ij}$ for all i, j.
- The formulas for the best linear unbiased predictors are complicated.
- The fitted values for the model are

$$\hat{Y}_{ijk} = \hat{\mu} + \hat{A}_i + \hat{B}_j + (\hat{AB})_{ij}, \qquad \text{for all } i, j, k$$

- Use `predict()` on the `lmer()` output to obtain these fitted values.
- The residuals are then $Y_{ijk} - \hat{Y}_{ijk}$ for all i, j, k.
- Can then check diagnostic plots.

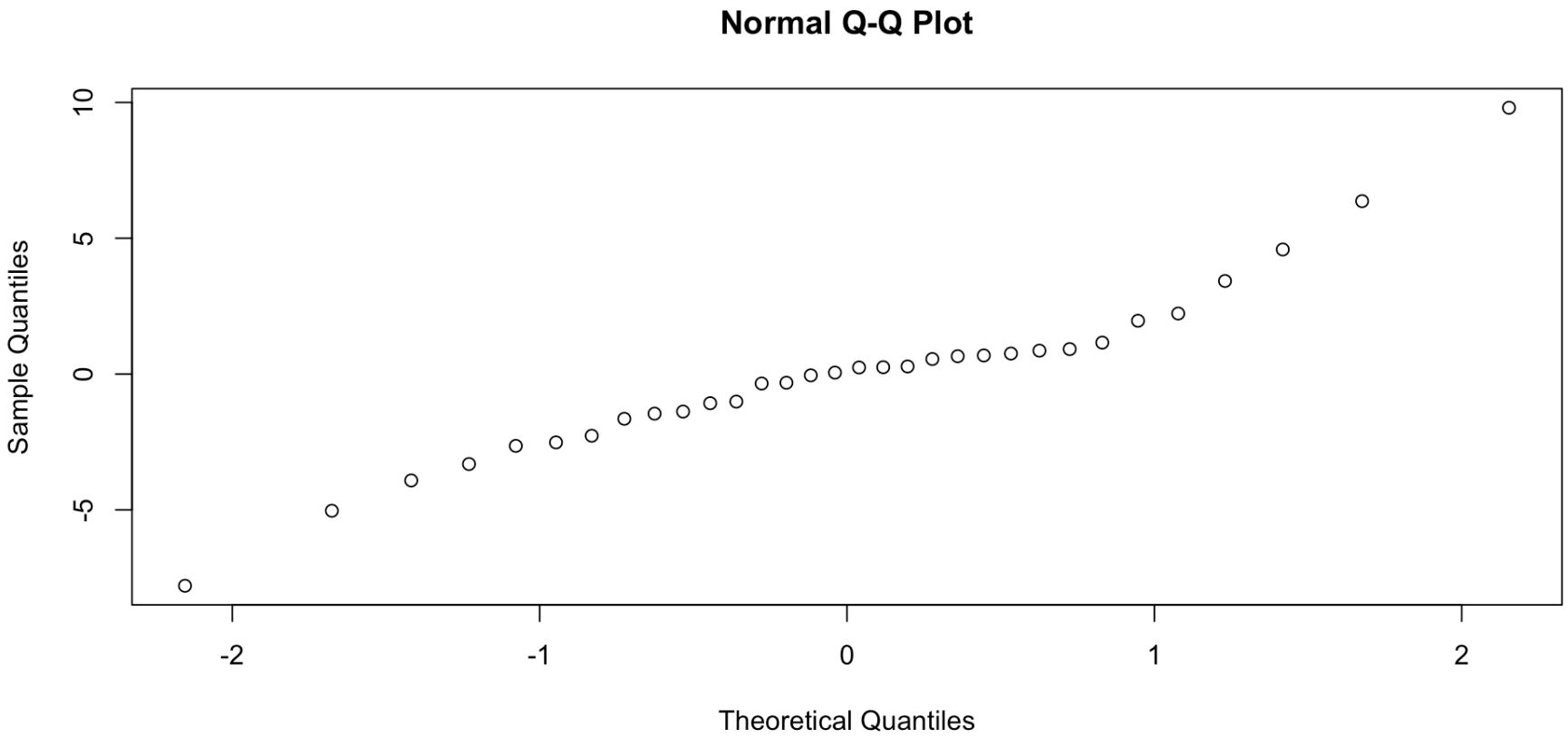
Triglyceride level

```
1 tg_hat <- predict(lmer_out)
2 ehat <- tg - tg_hat
3 plot(ehat ~ tg_hat)
```



Triglyceride level

```
1 qqnorm(ehat)
```



Both plots show some problems; interpret our results with caution.

References

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