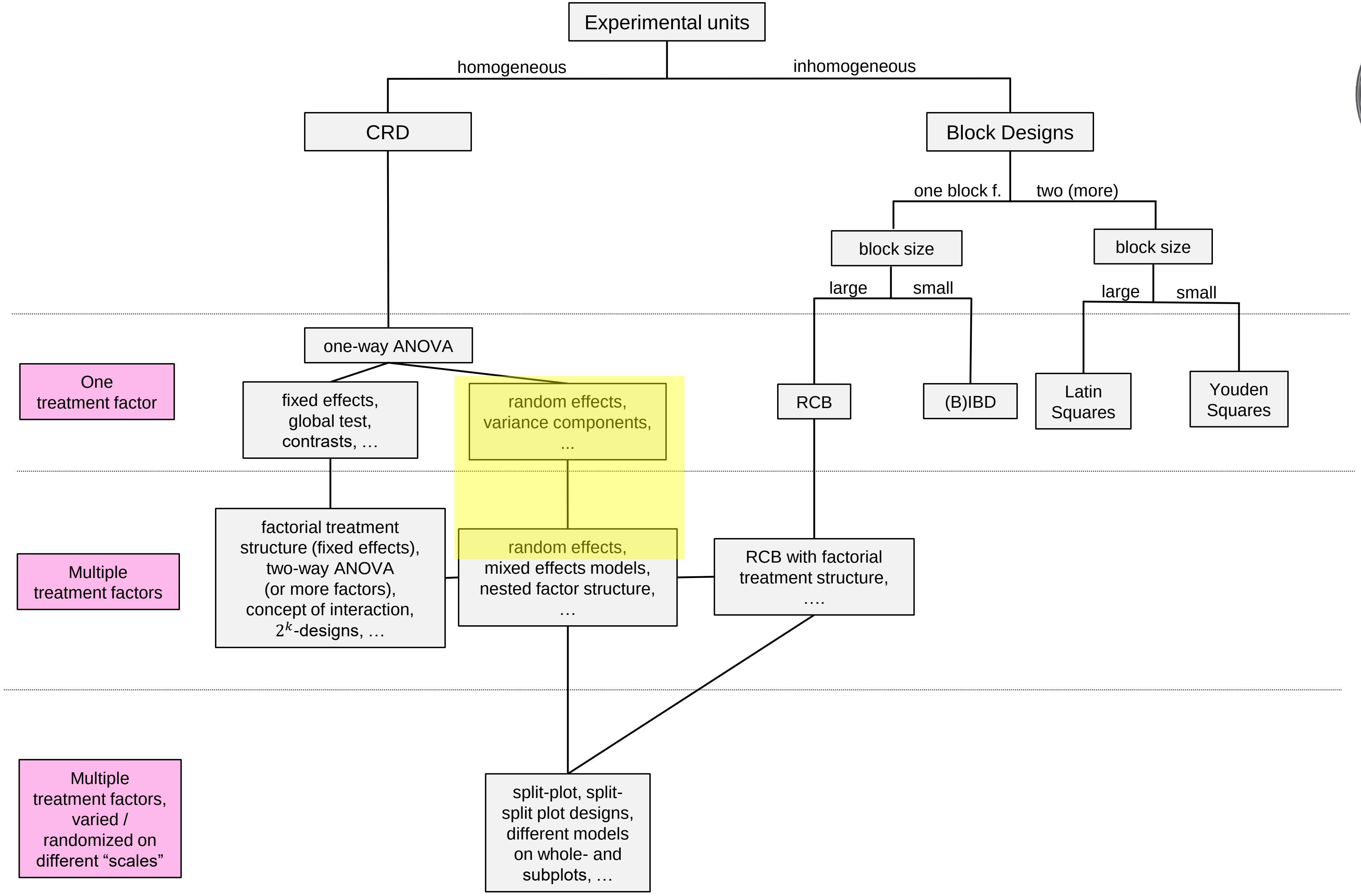
7

# Random Effects




*Similar to Lawson (2015)*

# New Philosophy…

- Up to now: Treatment effects were **fixed, unknown parameters** that we were trying to **estimate**.
- Such models are also called **fixed effects models**.
- Now: Consider the situation where treatments are **random samples** from a **large population** of potential treatments.


- Example: Effect of machine operators that were **randomly selected** from a large pool of operators.
- In this setup, treatment effects are **random variables** and therefore called **random effects**. The corresponding model will be a **random effects model**.

# New Philosophy…

- Why would we be interested in a random effects situation?
- It is a useful way of thinking if we want to make a statement (conclusion) about the **population of all treatments**.
- In the operator example, we **shift the focus away from the individual operators** (treatments) to the **population of all operators** (treatments).
- Typical questions:
  - What is the **variance** of the treatment population?
  - What is the **expected value** of the treatment population?

# Examples of Random Effects

| ***Randomly select …*** | ***… from …*** |
|---|---|
| clinics | … all clinics in a country. |
| school classes | … all school classes in a region. |
| investigators | … a large pool of investigators. |
| series in quality control | … all series in a certain time period. |
| … | … |

# Carton Experiment I (Oehlert, 2010)

- Company with 50 machines that produce cardboard cartons.
- Ideally, strength of the cartons shouldn't vary too much.
- Therefore, we want to have an idea about
  - "machine-to-machine" variation
  - "sample-to-sample" variation on the same machine.
- Perform experiment:
  - Choose 10 machines at **random** (out of the 50).
  - Produce 40 cartons on each machine.
  - Test resulting cartons for strength (= response).

# Carton Experiment I (Oehlert, 2010)

- Model so far:

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij},$$

  where $\alpha_i$ is the (**fixed**) effect of machine $i$ and $\varepsilon_{ij}$ are the errors with the usual assumptions.

- However, this model does **not** reflect the sampling mechanism from above.
- If we **repeat the experiment**, the selected machines **change** and therefore also the **meaning** of the **parameters**: They typically correspond to a **different** machine!
- Moreover, we want to learn something about the **population** of **all** machines and are **not** interested in a specific machine.

# Carton Experiment I (Oehlert, 2010)

- New: **Random effects model**:

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij},$$

effect of machine

Parameter

Random variable

with
  - $\alpha_i$ i.i.d. $\sim N(0, \sigma_\alpha^2)$ NEW
  - $\epsilon_{ij}$ i.i.d. $\sim N(0, \sigma^2)$

- This looks very similar to the old model, however the $\alpha_i$'s are now **random variables**!

- This small change will have a **large impact** on the properties of the model and on our way to analyze such kind of data.

# Properties of Random Effects Model

- $\text{Var}(Y_{ij}) = \sigma_\alpha^2 + \sigma^2$ **variance components**

- $$\text{Cor}(Y_{ij}, Y_{kl}) = \begin{cases} 0 & i \neq k \\ \sigma_\alpha^2/(\sigma_\alpha^2 + \sigma^2) & i = k, j \neq l \\ 1 & i = k, j = l \end{cases}$$

  different machines

  same machine

  **intraclass correlation**

  Reason: Observations from the same machine "share" the **same random value** $\alpha_i$ and are therefore correlated.

- Conceptually, we could also put all the correlation structure into the error term and forget about the $\alpha_i$'s, i.e.

$$Y_{ij} = \mu + \epsilon_{ij}$$

where $\epsilon_{ij}$ has the appropriate correlation structure from above. Sometimes this interpretation is a useful way of thinking.

# Random vs. Fixed: Overview

- **Comparison** between random and fixed effects models

| *Term* | *Fixed effects model* | *Random effects model* |
|---|---|---|
| $\alpha_i$ | fixed, unknown constant | $\alpha_i \text{ i.i.d.} \sim N(0, \sigma_\alpha^2)$ |
| Side constraint on $\alpha_i$ | needed | **not** needed |
| $E[Y_{ij}]$ | $\mu + \alpha_i$ | $\mu$, but $E\left[Y_{ij} \mid \alpha_i\right] = \mu + \alpha_i$ |
| $\text{Var}(Y_{ij})$ | $\sigma^2$ | $\sigma_\alpha^2 + \sigma^2$ |
| $\text{Corr}(Y_{ij}, Y_{kl})$ | $= 0 \ (j \neq l)$ | $= \begin{cases} 0 & i \neq k \\ \sigma_\alpha^2/(\sigma_\alpha^2 + \sigma^2) & i = k, j \neq l \\ 1 & i = k, j = l \end{cases}$ |

- A note on the **sampling mechanism**:
  - Fixed: Draw new random errors only, everything else is kept **constant**.
  - Random: Draw new "treatment effects" **and** new random errors (!) NEW

# Illustration of Sampling Mechanism and Correlation Structure

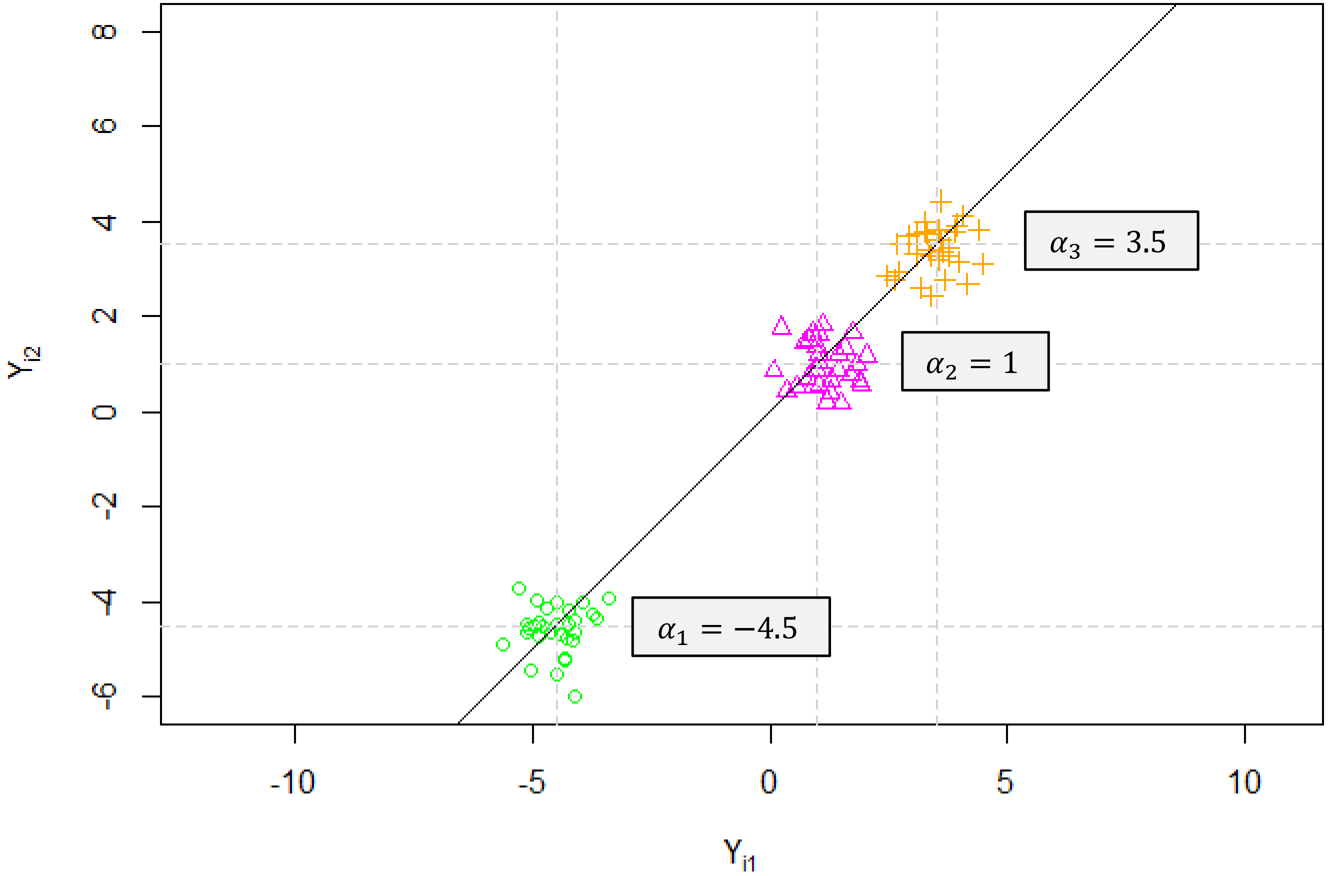
**Fixed case:** 3 different **fixed** treatment levels $\alpha_i$.

Think of 3 **specific** machines

We (repeatedly) sample 2 observations per treatment level:

Think of 2 carton samples

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$$



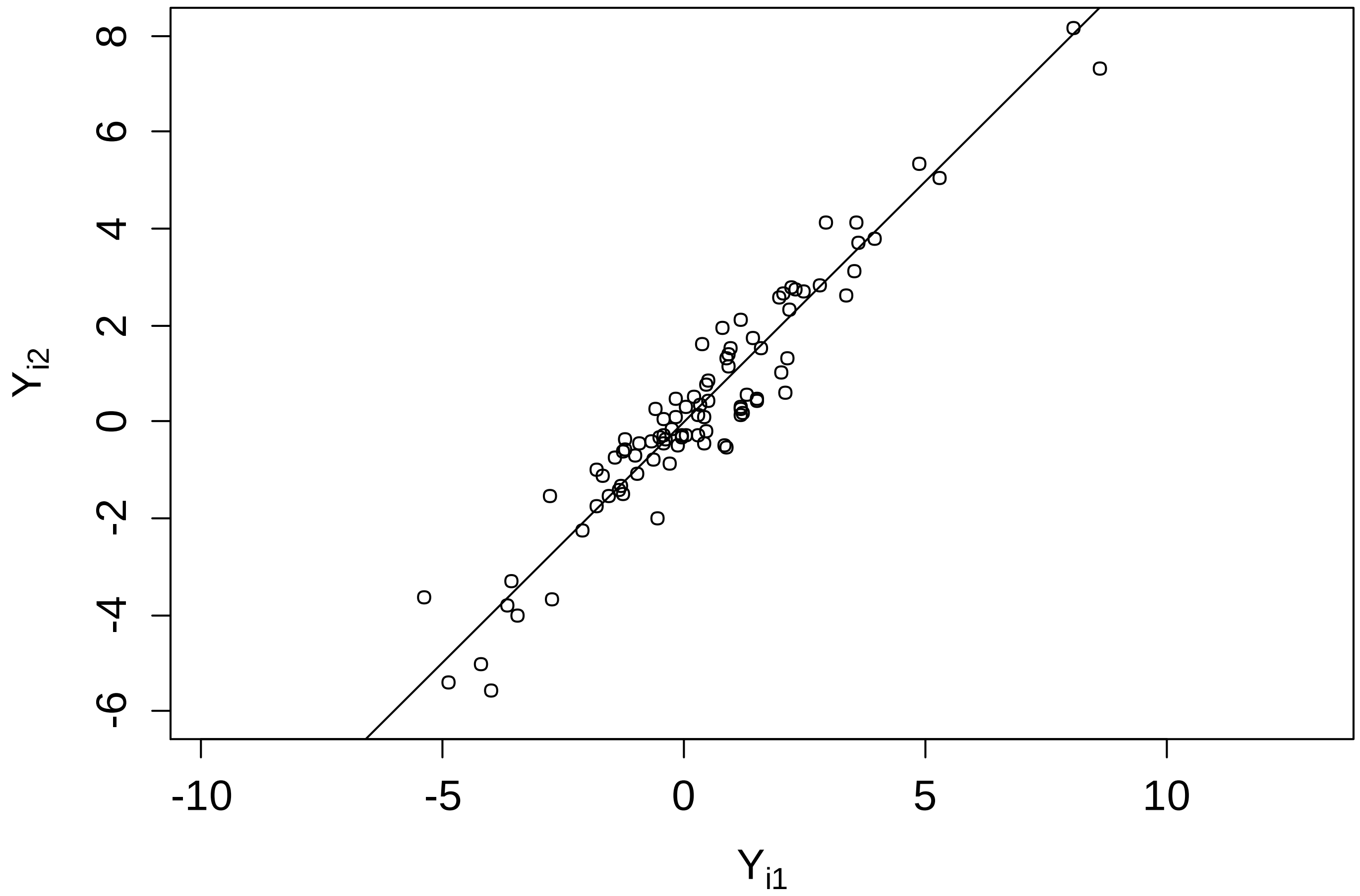
# Illustration of Sampling Mechanism and Correlation Structure

## Random case:

Think of 2 carton samples

Whenever we draw 2 observations $Y_{i1}$ and $Y_{i2}$, we first have to draw a **new** (common) random treatment effect $\alpha_i$.

Think of a **random** machine.



# Carton Experiment II (Oehlert, 2010)

- Let us **extend** the previous experiment.
- Assume that machine operators also influence the production process.
- Choose 10 operators at **random**.
- **Each operator** will produce 4 cartons on **each machine** (hence, operator and machine are **crossed** factors).
- All assignments are completely randomized.

# Carton Experiment II (Oehlert, 2010)

- Model:

$$Y_{ijk} = \mu + \underbrace{\overbrace{\alpha_i}^{N(0,\sigma_\alpha^2)}}_{\substack{\textbf{main effect}\\ \text{machine}}} + \underbrace{\overbrace{\beta_j}^{N(0,\sigma_\beta^2)}}_{\substack{\textbf{main effect}\\ \text{operator}}} + \underbrace{\overbrace{(\alpha\beta)_{ij}}^{N(0,\sigma_{\alpha\beta}^2)}}_{\substack{\textbf{interaction}\\ \text{machine}\times\text{operator}}} + \overbrace{\epsilon_{ijk}}^{N(0,\sigma^2)},$$

  with

  - $\alpha_i, \beta_j, (\alpha\beta)_{ij}, \epsilon_{ijk}$ independent and normally distributed.
  - $\mathrm{Var}(Y_{ijk}) = \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_{\alpha\beta}^2 + \sigma^2$ (different **variance components**).
  - Measurements from the same machine and / or operator are again **correlated**.
  - The more random effects two observations **share**, the larger the correlation. It is given by

$$\frac{\text{sum of } \textbf{shared} \text{ variance components}}{\text{sum of } \textbf{all} \text{ variance components}}$$

  - E.g., the correlation between two (different) observations from the **same operator** on **different machines** is given by

$$\frac{\sigma_\beta^2}{\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_{\alpha\beta}^2 + \sigma^2}$$

# Carton Experiment II (Oehlert, 2010)

- **Hierarchy** is typically less problematic in random effects models.
  1) What part of the variation is due to general machine-to-machine variation? → $\sigma_\alpha^2$
  2) What part of the variation is due to operator-specific machine variation? → $\sigma_{\alpha\beta}^2$

  Could ask question (1) even if interaction is present (question (2)).

- Extending the model to **more than two factors** is straightforward.

# Estimates of Variance Components: REML

- Could estimate variance using an “old-school approach” (see Appendix).
- More modern and much more flexible: **Restricted maximum-likelihood** estimator (**REML**).
- Think of a modification of the maximum likelihood principle that **removes bias** in the estimation of the variance components.
- Software implementation in R-package `lme4` (or `lmerTest`).
- `lme4` and `lmerTest` allow to fit so called **mixed effects models** (containing both random **and** fixed effects, more details later).
- Basically, `lmerTest` is the same as `lme4` but with some more features.

# Tests and Confidence Intervals for Variance Components

- General rule: Variances are "difficult" to estimate in the sense that you'll need **a lot** of observations to have some **reasonable accuracy**.
- If we do a study with random effects, it is good if we have **a lot of levels** of a random effect in order to estimate a variance component with high precision. Or in other words: Who wants to estimate a variance with only very few observations?
- **Approximate confidence intervals (or tests)** can be obtained by calling the function `confint`.
- **Exact tests** (simulation based) for variance components can be found in the package `RLRsim`.
- Typically, we are more interested in the accuracy of the variance component estimates (confidence intervals) than in tests of the form: $H_0: \sigma_\alpha^2 = 0$ vs. $H_A: \sigma_\alpha^2 > 0$.

# Residual Analysis

- We can also get "estimates" (more precisely: conditional means) of the random effects $\alpha_i$ with the function `ranef`.
- They can be used to do a residual analysis "as before".
- As a new addition, we also do QQ-plots of the $\alpha_i$'s (not only of the residuals).
- Let us have a look at an easy example.

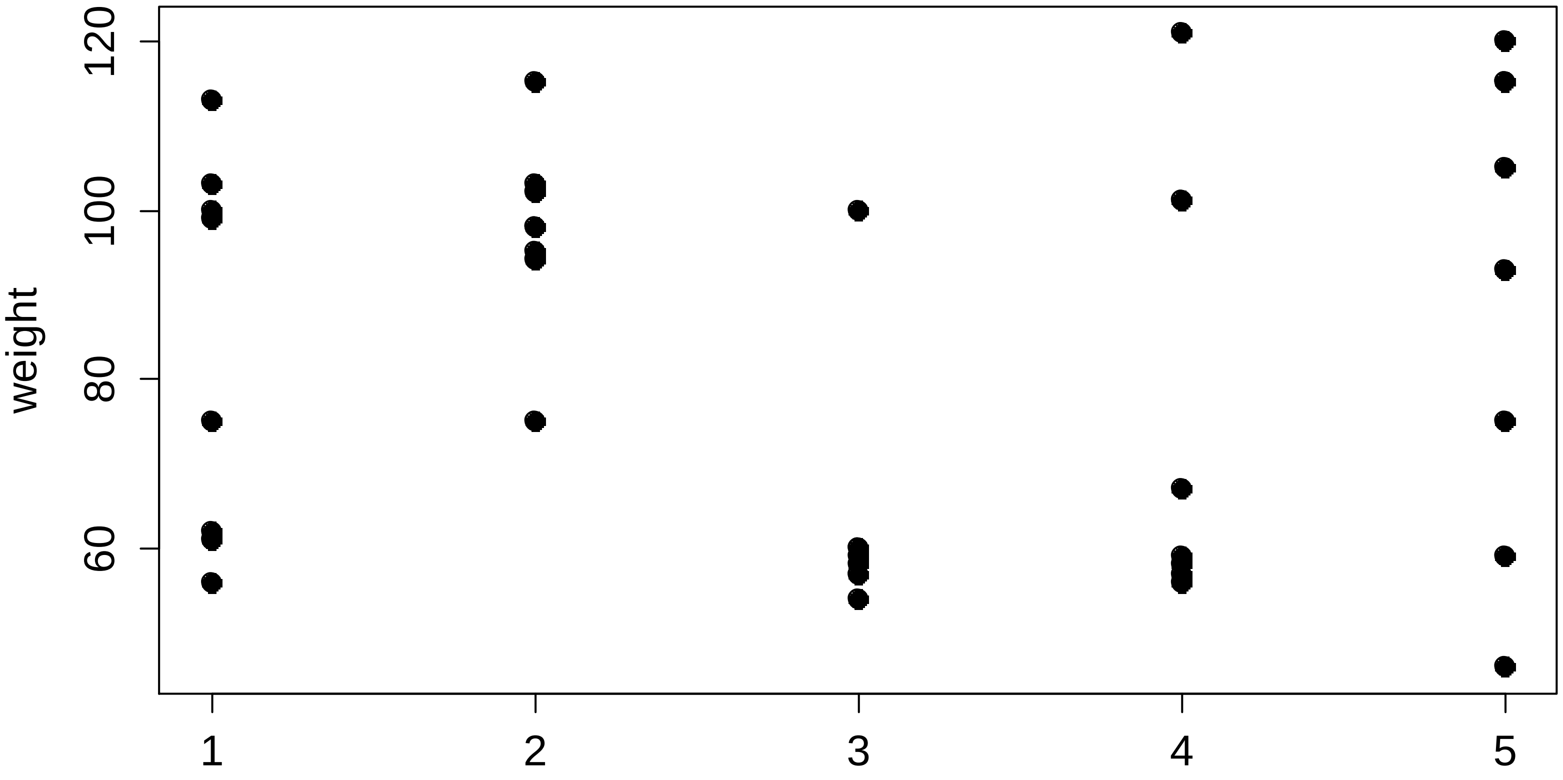
# Example: Genetics Study (Kuehl, 2000, Exercise 5.1)

- **Inheritance study** of **birth weights** (of beef animals).
- 5 **sires**, each mated to a separate group of **dams**.
- Birth weight of 8 male calves (from different dams) in each of the five sire groups:

| Sire | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| **1** | 61 | 100 | 56 | 113 | 99 | 103 | 75 | 62 |
| **2** | 75 | 102 | 95 | 103 | 98 | 115 | 98 | 94 |
| **3** | 58 | 60 | 60 | 57 | 57 | 59 | 54 | 100 |
| **4** | 57 | 56 | 67 | 59 | 58 | 121 | 101 | 101 |
| **5** | 59 | 46 | 120 | 115 | 115 | 93 | 105 | 75 |

- What part of variation can be explained by sire?
- Analyze data using a **random effect for sire**.

# Example: Genetics Study (Kuehl, 2000, Chapter 5, Ex. 1)



# Example: Genetics Study

- Model: $Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$, $\alpha_i$ i.i.d. $\sim N(0, \sigma_\alpha^2)$, $\epsilon_{ij}$ i.i.d. $\sim N(0, \sigma^2)$.
- In R using the function `lmer` in the package `lmerTest` (or `lme4`).

```
> fit.lme <- lmer(weight ~ 1 | sire, data = animals)
> summary(fit.lme)
Linear mixed model fit by REML ['lmerMod']
Formula: weight ~ 1 | sire
   Data: animals

REML criterion at convergence: 358.2

Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.9593 -0.7459 -0.1581  0.8143  1.9421

Random effects:
 Groups   Name        Variance Std.Dev.
 sire     (Intercept) 116.7    10.81
 Residual             463.8    21.54
Number of obs: 40, groups:  sire, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)   82.550      5.911   13.96
```

Meaning: a random effect **per** sire

$\sqrt{}$

$\hat{\sigma}_\alpha$ Estimate of **population standard deviation**!

$\hat{\sigma}$

Check if model was interpreted correctly

$\hat{\mu}$ Estimate of **population mean**!

# Example: Genetics Study

- The variance of $Y_{ij}$ is estimated as

$$\hat{\sigma}_\alpha^2 + \hat{\sigma}^2 = 116.7 + 463.8 = 580.5.$$

- We have

$$\frac{\hat{\sigma}_\alpha^2}{\hat{\sigma}_\alpha^2 + \hat{\sigma}^2} = \frac{116.7}{580.5} = 0.2.$$

- Hence, variation due to sire accounts for about 20% of the total variance.
- Or in other words: **Intraclass correlation** is 0.2.

# Example: Evaluating Machine Performance (Kuehl, 2000, Ex. 7.1)

- A manufacturer develops a new spectrophotometer for medical labs.
- The development is at pilot stage. Evaluate machine performance from assembly line production.
- Critical: **Consistency** of measurements from **day to day** among **different machines**.
- Design:
  - 4 (randomly selected) machines
  - 4 (randomly selected) days
- Per day: 8 serum samples (from the **same** stock reagent), randomly assign 2 samples to each of the 4 machines.

# Example: Evaluating Machine Performance

- Measure triglyceride levels (mg/dl) of the samples.
- Note: Always the **same technician** prepared the serum samples and operated the machines **throughout the experiment**.
- Fit random effects model with interaction with usual assumptions:

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

| $N(0, \sigma_\alpha^2)$ [day] | $N(0, \sigma_\beta^2)$ [machine] | $N(0, \sigma_{\alpha\beta}^2)$ [day × machine] | $N(0, \sigma^2)$ [error] |
|---|---|---|---|

# Example: Evaluating Machine Performance

- Using the function `lmer` in the package `lmerTest` (or `lme4`).

```
> fit.lme <- lmer(y ~ (1 | day) + (1 | machine) + (1 | machine:day), data = trigly)
> summary(fit.lme)
Linear mixed model fit by REML ['lmerMod']
Formula: y ~ (1 | day) + (1 | machine) + (1 | machine:day)
   Data: trigly

REML criterion at convergence: 215

Scaled residuals:
     Min       1Q   Median       3Q      Max
-1.84282 -0.35581  0.03484  0.20699  2.31766

Random effects:
 Groups      Name        Variance Std.Dev.
 machine:day (Intercept) 34.72    5.892
 machine     (Intercept) 57.72    7.597
 day         (Intercept) 44.69    6.685
 Residual                17.90    4.230
Number of obs: 32, groups:  machine:day, 16; machine, 4; day, 4

Fixed effects:
            Estimate Std. Error t value
(Intercept)  141.184      5.323   26.52
```

Meaning: a random effect per **day**, per **machine** and per **day×machine combination**

$\sqrt{}$

$\hat{\sigma}_{\alpha\beta}$

$\hat{\sigma}_{\beta}$

$\hat{\sigma}_{\alpha}$

$\hat{\sigma}$

Check if model was interpreted correctly

$\hat{\mu}$

# Example: Evaluating Machine Performance

- The total variance is $17.9 + 34.7 + 44.7 + 57.7 = 155$.
- Individual contributions:

| ***Source*** | ***Percentage*** | ***Interpretation*** |
|---|---|---|
| Day | $\frac{44.7}{155} = 29\%$ | Day to day operational differences (e.g., due to daily calibration). |
| Machine | $\frac{57.7}{155} = 37\%$ | Variability in machine performance. |
| Interaction | $\frac{34.7}{155} = 22\%$ | Variability due to inconsistent behavior of machines over days (calibration inconsistency within the same day?). |
| Error | $\frac{17.9}{155} = 12\%$ | Variation in serum samples. |

- Manufacturer now has to decide if some sources of variation are too large.

# Appendix

# ANOVA for Random Effects Models (balanced designs)

*OLD*

- Sums of squares, degrees of freedom and mean squares are being calculated as if the model were a **fixed effects** model (!)

- One-way ANOVA ($A$ random, $n$ observations per cell):

| Source | df | SS | MS | E[MS] |
|---|---|---|---|---|
| $A$ | $g-1$ | … | … | $\sigma^2 + n\sigma_\alpha^2$ NEW |
| Error | $N-g$ | … | … | $\sigma^2$ |

- Two-way ANOVA ($A, B, AB$ random, $n$ observations per cell):

| Source | df | SS | MS | E[MS] |
|---|---|---|---|---|
| $A$ | $a-1$ | … | … | $\sigma^2 + b \cdot n \cdot \sigma_\alpha^2 + n \cdot \sigma_{\alpha\beta}^2$ NEW |
| $B$ | $b-1$ | … | … | $\sigma^2 + a \cdot n \cdot \sigma_\beta^2 + n \cdot \sigma_{\alpha\beta}^2$ NEW |
| $AB$ | $(a-1)(b-1)$ | … | … | $\sigma^2 + n \cdot \sigma_{\alpha\beta}^2$ NEW |
| Error | $ab(n-1)$ | … | … | $\sigma^2$ |

# One-Way ANOVA with Random Effects

$\mathcal{OLD}$

- We are now formulating our null-hypothesis with respect to the parameter $\sigma_\alpha^2$.
- To test $H_0: \sigma_\alpha^2 = 0$ vs. $H_A: \sigma_\alpha^2 > 0$, we use the ratio

  $$F = \frac{MS_A}{MS_E} \sim F_{g-1,\, N-g} \text{ under } H_0.$$

  **Exactly** as in the fixed effect case!
- Why? Under the **old and the new** $H_0$, both models are the same!

# Two-Way ANOVA with Random Effects

OLD

- To test $H_0\!: \sigma_\alpha^2 = 0$, we need to find a term which has identical $E[MS]$ under $H_0$.
- Use $MS_{AB}$, i.e. $F = \frac{MS_A}{MS_{AB}} \sim F_{a-1,\,(a-1)(b-1)}$ under $H_0$. NEW
- Similarly for the test $H_0\!: \sigma_\beta^2 = 0$. NEW
- The interaction will be tested against the error, i.e. use

$$F = \frac{MS_{AB}}{MS_E} \sim F_{(a-1)(b-1),\,ab(n-1)}$$

  under $H_0\!: \sigma_{\alpha\beta}^2 = 0$.
- In the fixed effect case we would test **all effects** against the **error term** (i.e., use $MS_E$ instead of $MS_{AB}$ to build $F$-ratio)!

# Two-Way ANOVA with Random Effects

Didn't look at this column when analyzing factorials

*OLD*

- Reason: ANOVA table for fixed effects:

Shorthand notation for a term depending on $\alpha_i's$

| Source | df | E[MS] |
|---|---|---|
| $A$ | $a-1$ | $\sigma^2 + b \cdot n \cdot Q(\alpha)$ |
| $B$ | $b-1$ | $\sigma^2 + a \cdot n \cdot Q(\beta)$ |
| $AB$ | $(a-1)(b-1)$ | $\sigma^2 + n \cdot Q(\alpha\beta)$ |
| Error | $ab(n-1)$ | $\sigma^2$ |

- E.g, $SS_A$ $(MS_A)$ is being calculated based on row-wise means.
- In the **fixed effects model**, the expected mean squares do **not** "contain" any other components.

# Two-Way ANOVA with Random Effects

$\mathcal{OLD}$

- In a random effects model, a row-wise mean is “contaminated” with the average of the corresponding interaction terms.
- In a **fixed effects** model, the sum (or mean) of these interaction terms is **zero by definition**.
- In the **random effects** model, this is only true for the **expected value**, but **not for an individual realization**!
- Hence, we need to check whether the variation from “row to row” is larger than the term based on the error **and** interaction term.

# Point Estimates of Variance Components

$\mathcal{OLD}$

- We do not only want to **test** the variance components, we also want to have **estimates** of them.
- I.e., we want to determine $\hat{\sigma}_{\alpha}^2, \hat{\sigma}_{\beta}^2, \hat{\sigma}_{\alpha\beta}^2, \hat{\sigma}^2$ etc.
- Easiest approach: **ANOVA estimates** of the variance components.
- Use columns “MS” and “E[MS]” in the ANOVA table, solve the corresponding equations from bottom to top.
- Example: One-way ANOVA
  - $\hat{\sigma}^2 = MS_E$
  - $\hat{\sigma}_{\alpha}^2 = \frac{(MS_A - MS_E)}{n}$

# Point Estimates of Variance Components

$\mathcal{OLD}$

- Advantage: Can be done using **standard ANOVA functions** (i.e., no special software needed).
- Disadvantages:
  - Estimates can be negative (in previous example, if $MS_A < MS_E$). Set them to zero in such cases.
  - Not always as easy as here.
- This is like a method of moments estimator.
- More modern and much more flexible: **Restricted maximum-likelihood** estimator (**REML**).

# Example: Genetics Study

*OLD*

- Model: $Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$, $\alpha_i \text{ i.i.d.} \sim N(0, \sigma_\alpha^2)$, $\epsilon_{ij} \text{ i.i.d.} \sim N(0, \sigma^2)$

```
> fit <- aov(weight ~ sire, data = animals)
> summary(fit)
            Df Sum Sq Mean Sq F value Pr(>F)
sire         4   5591  1397.8   3.014 0.0309
Residuals   35  16233   463.8
```

- We reject $H_0: \sigma_\alpha^2 = 0$.
- We estimate $\sigma_\alpha^2$ by $\hat{\sigma}_\alpha^2 = \frac{1397.8 - 463.8}{8} = 116.75$.
- The variance of $Y_{ij}$ is estimated as

$$\hat{\sigma}_\alpha^2 + \hat{\sigma}^2 = 116.75 + 463.8 = 580.55.$$

- Variation due to sire accounts for about 20% of the total variance (= **intraclass correlation**).
- We fitted the model as if it was a fixed effects model and then “adjusted” the output for random effects specific questions.

# Example: Evaluating Machine Performance

$OLD$

- Fit random effects model with interaction with usual assumptions.

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

| $N(0, \sigma_\alpha^2)$ [day] | $N(0, \sigma_\beta^2)$ [machine] | $N(0, \sigma_{\alpha\beta}^2)$ [day × machine] | $N(0, \sigma^2)$ [error] |
|---|---|---|---|

- Classical approach:

```
> fit <- aov(y ~ day * machine, data = trigly)
> summary(fit)
            Df Sum Sq Mean Sq F value   Pr(>F)    
day          3 1334.5   444.8   24.86 2.91e-06 ***
machine      3 1647.3   549.1   30.68 7.19e-07 ***
day:machine  9  786.0    87.3    4.88  0.00294 ** 
Residuals   16  286.3    17.9
```

- Classical approach to estimate variance components:

$$\hat{\sigma}^2 = 17.9 \qquad \hat{\sigma}_\alpha^2 = \frac{444.8 - 87.3}{8} = 44.7$$

$$\hat{\sigma}_{\alpha\beta}^2 = \frac{87.3 - 17.9}{2} = 34.7 \qquad \hat{\sigma}_\beta^2 = \frac{549.1 - 87.3}{8} = 57.7$$

# Example: Evaluating Machine Performance

$\mathcal{OLD}$

Testing the variance components: “by hand”

- **Interaction**: $H_0\colon \sigma^2_{\alpha\beta} = 0.$

$$\frac{\mathrm{M}S_{AB}}{MS_E} = \frac{87.3}{17.9} = 4.9,\ F_{9,16}\text{-distribution}$$

```
> pf(87.3 / 17.9, 9, 16, lower.tail = FALSE)
[1] 0.002946051
```

→ reject

- **Main effect day**: $H_0\colon \sigma^2_{\alpha} = 0.$

$$\frac{\mathrm{M}S_{A}}{MS_{AB}} = \frac{444.8}{87.3} = 5.1,\ F_{3,9}\text{-distribution}$$

```
> pf(444.8 / 87.3, 3, 9, lower.tail = FALSE)
[1] 0.02477665
```

→ reject

- **Main effect machine**: $H_0\colon \sigma^2_{\beta} = 0.$

$$\frac{\mathrm{M}S_{B}}{MS_{AB}} = \frac{549.1}{87.3} = 6.3,\ F_{3,9}\text{-distribution}$$

```
> pf(549.1 / 87.3, 3, 9, lower.tail = FALSE)
[1] 0.01370686
```

→ reject