Random Effects Models

Jesse Wheeler

Overview

Objectives and overview

- This material is based Chapter 11–12 on the textbook "A First Course in Design and Analysis of Experiments" (Oehlert, 2000).
- Chapter outcomes include:
  - Students will be able to distinguish between a random effect and a fixed effect.
  - Students will be able to identify when random effects are appropriate.
  - Students will know how to implement random effects and mixed effects models.
  - Students will gain some familiarity with the `lme4` package in R.

Introduction

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

- So far, the treatment effects that we have been working with can be classified as fixed-effects: we wish to estimate the effect of specific (fixed) treatments.
- What if the treatments themselves are a random selection from a larger population?

μ, α_i's are fixed effects

Example: Farming Equipment

A farmer has 100 pieces of equipment, and would like to know how the equipment used may impact crop-yield. Rather than testing all 100 pieces of equipment, they *randomly select* g pieces of equipment, and performs a CRD with these equipment.

Introduction II

Random Effects

A random effect is appropriate to model treatments that are a *random sample* from a population of potential treatments.

Random effects are useful if the focus of inference is on the population, not the individual treatment effects.

- When *designing* for a random effect, the treatments should be a random draw from a larger population.

Models for Random Effects

Example: Carton factory

- We will build intuition of Random Effects models and designs using an example.

Carboard company

Suppose a company makes cardboard cartons for storing goods. They have 50 machines to make cartons, and want to understand the variation in quality for finished products.

They choose 10 machines at random from the 50, and make 40 cartons on each machine. The resulting cartons are tested for strength.

Example: Carton factory II

- This is a completely randomized design (CRD), with $g = 10$ treatments (the different machines), and $n_i = 40$, for a total of $N = 400$ EUs.
- The model we have been using for this type of data is:

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij}, \quad \epsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2).$$

- We are still going to use a very similar model. However, we are not as interested in the exact α_i for the 10 randomly chosen machines. Instead, we want to make inference on the larger population.

Example: Carton factory III

- We would get an entirely different set of machines if we repeated the experiment!
- Thus, we want to introduce a random effect: let the effect α_i represent a random draw from a population of possible effects:

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij}, \qquad \alpha_i \overset{iid}{\sim} N(0, \sigma^2_\alpha), \quad \epsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2).$$

- We no longer assume $\sum_i \alpha_i = 0$, rather $E[\alpha_i] = 0$.
- We have introduced a new parameter, σ^2_α, which describes the spread of the effects.
- We assume that the two random components are independent. That is, α_i and ϵ_{ij} are independent.

Example: Carton factory IV

- If they are independent, the variance of the observed data is increased: Independent RVs: $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

$$\mathrm{Var}(y_{ij}) = \mathrm{Var}(\mu + \alpha_i + \epsilon_{ij}) = \sigma_\alpha^2 + \sigma^2.$$

- The parameters σ_α^2 and σ^2 are called *variance components*.
- If observations belong to the same treatment group, then they now have non-zero correlation:

$$\mathrm{Cor}(y_{ij}, y_{ik}) = \begin{cases} 1 & j = k \\ \sigma_\alpha^2/(\sigma_\alpha^2 + \sigma^2) & j \neq k \end{cases}$$

- We call this within treatment group correlation the intraclass correlation. ICC

Example: Carton factory V

Key summary

- In a random effects model, the parameters are μ, σ^2, and σ^2_α.
- We are not interested in making inferences about the specific α_i and ϵ_{ij}, which are treated as random variables, not parameters.
- Inference wise, we are now interested in confidence intervals and hypothesis tests regarding μ and σ^2. σ^2_α

Example: Carton factory VI

Practical Advice

In fixed effects, we generally prefer larger n and smaller g; for random effects, it's the opposite!

- We really would like a minimum of $g = 10$ groups for reliable estimates; otherwise, confidence intervals will be very large / meaningless.
- $g = 15$ or more is preferred.

Hypothesis Testing

Hypothesis Testing

- If we are testing for population effects based on our random sample, there are multiple ways to do this formally.
- The hypothesis we are generally interested in testing in this scenario is:

$$H_0 : \sigma_\alpha^2 = 0, \quad \text{vs} \quad H_0 : \sigma_\alpha^2 > 0.$$

- Surprisingly, this leads to the same F-tests that we have been using before!

$\alpha_i \overset{iid}{\sim} N(0, \sigma_\alpha^2)$

If H_0 is true $\text{var}(\alpha_i) = 0$, $\alpha_1 = \alpha_2 = \cdots = \alpha_g = 0$

$H_0 : \sigma_\alpha^2 = 0$ is true IF and only IF $H_0 : \alpha_i = 0$ is true.

So testing one way random effect is equivalent to a one-way ANOVA.

we want to estimate σ_α^2.

Example data in R

- The following example uses data synthetically generated with $\alpha_i \sim N(0, 1.37^2)$, $\epsilon_{ij} \sim N(0, 0.5^2)$, and $\mu = 10$.

```
synthData <- read.csv("synthData.csv")
mod <- aov(response ~ factor(trt), data = synthData)
summary(mod)
```

```
             Df Sum Sq Mean Sq F value   Pr(>F)
factor(trt)   3  94.21  31.404   40.66 2.46e-10 ***
Residuals    28  21.63   0.772
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Example data in R II

Above, we see a significant (random) treatment effect, rejecting the hypothesis that $\sigma_\alpha^2 = 0$.

- Note in this problem, the F statistic that is used matches the R output:

  $$F_{DF_A, DF_E} = \frac{MS_A}{MS_E}.$$

  This will *not* be the case for factorial models or if the data are unbalanced.
- For factorials models, the denominator changes. We *can* do it by hand, but the output of `aov` does *not* match!

Estimation variance components

Estimation of Variance Components

In perfectly balanced data, the `aov()` output can be used to obtain estimates for σ^2_α; this follows a similar approach to estimating σ^2, like we did back in module 3.

- Recall: $\hat{\sigma}^2 = \text{MS}_\text{E}$. $= 0.772$
- If the data are balanced, then we can subtract off the error variance, and divide by the number of observations (per-group): EMS algorithm.

$$\hat{\sigma}^2_\alpha = \frac{MS_\text{A} - MS_\text{E}}{n \text{ per group}}$$

true value used to create data.

$\sigma^2_\alpha = 4$

$$= \frac{31.404 - 0.772}{8}$$

$$\hat{\sigma}^2_\alpha = 3.829$$

parameters	estimates
μ, α_i (if fixed effects)	$\hat{\mu}$, $\hat{\alpha}_i$
σ^2_α (if random)	$\hat{\sigma}^2_\alpha$, $\hat{\sigma}^2$
σ^2	
We never see these in real-world.	We only can estimate.

Estimation: REML and LRT

Beyond ANOVA: Likelihood Based Inference

- For factorial random effects models, or unbalanced data, the default `aov()` output will not work!
- Classic solutions rely on Expected Mean Squares (EMS) (See Oehlert (2000, Sections 11.3–11.6)). These approaches can be computed by hand.
- Unfortunately, these approaches have their weaknesses.
- We will shift from the textbook and rely on likelihood-based methods, which is in-line with modern practices.
- For the type of models considered so far (fixed-effects, or balanced random effect), the two approaches are the same.

Beyond ANOVA: Likelihood Based Inference II

- **Maximum Likelihood (ML)** finds the parameter estimates (mean and variances) that make the observed data most probable; confidence intervals are obtained using likelihood ratio tests.

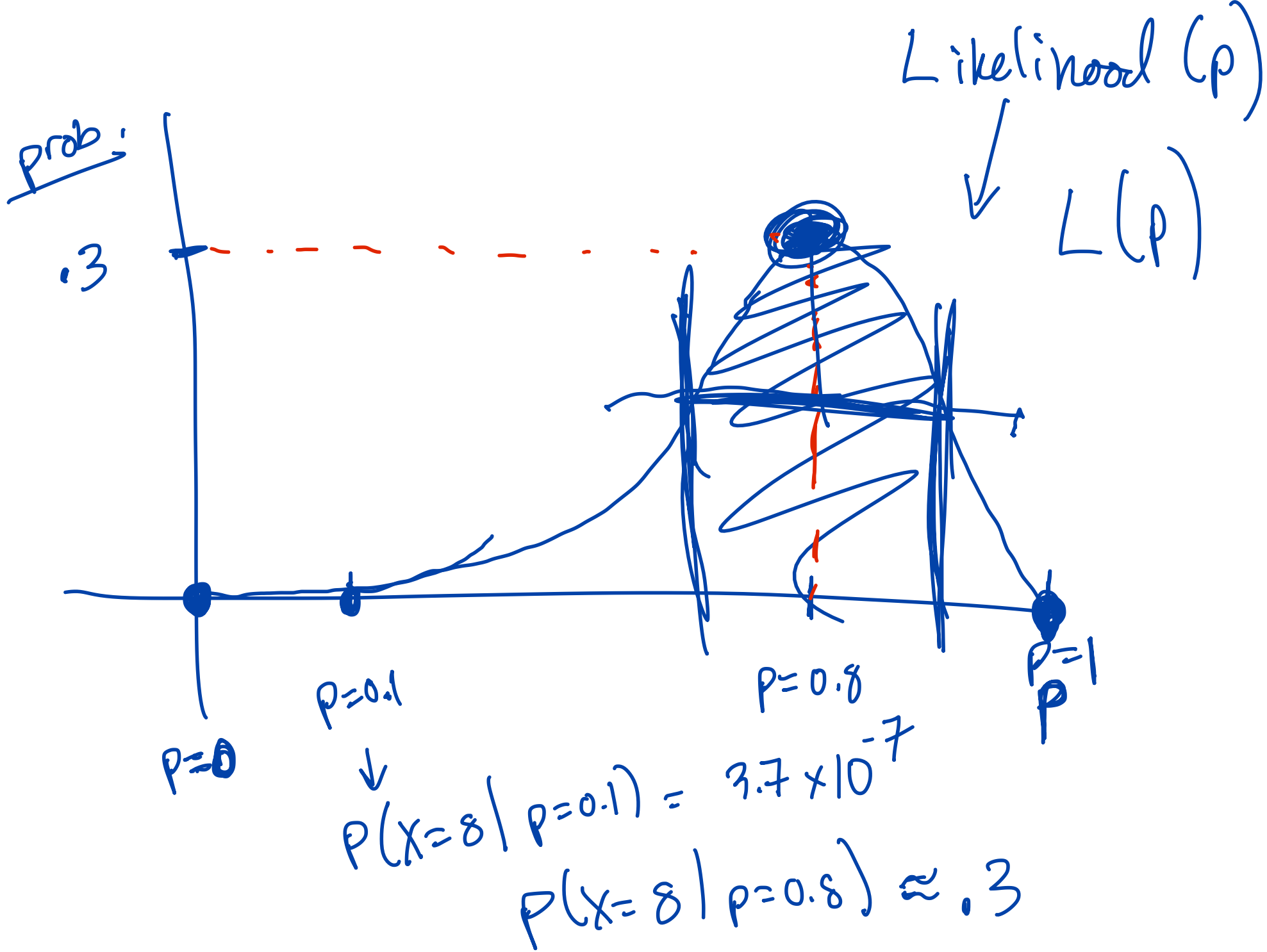
Flipping a coin $n=10$ times.

X is # of heads. $X \sim Bin(\overset{n=10}{10}, p)$

prob of heads.

$X = 8$ heads observed.

Q: what does this tell us about p?


Likelihood (p)
L(p)
prob:
.3
p=0
p=0.1
p=0.8
p=1
p
$P(X=8 \mid p=0.1) = 3.7 \times 10^{-7}$
$P(X=8 \mid p=0.8) \approx .3$

the point $p = 0.8$ maximizes the likelihood.

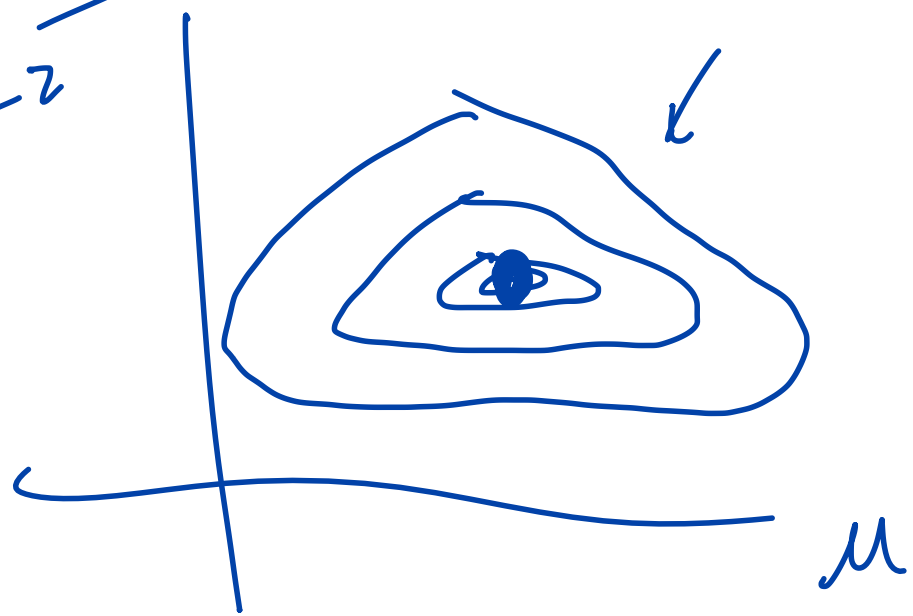
$\hat{p} = 0.8$ is called the "Maximum likelihood estimate".

MLE

Any probability model has a likelihood function.

$\theta = (\mu, \sigma^2)$

$X \sim N(\mu, \sigma^2)$ $\quad L(\theta)$



Recap 11/19

$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$, $\quad \alpha_i \overset{iid}{\sim} N(0, \sigma_\alpha^2)$

$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\alpha^2)$.

For balanced data, $\quad H_0: \sigma_\alpha^2 = 0 \quad$ vs

$H_1: \sigma_\alpha^2 > 0$

is the same as

$H_0: \alpha_i = 0 \quad$ vs

$H_1: \alpha_i \neq 0$.

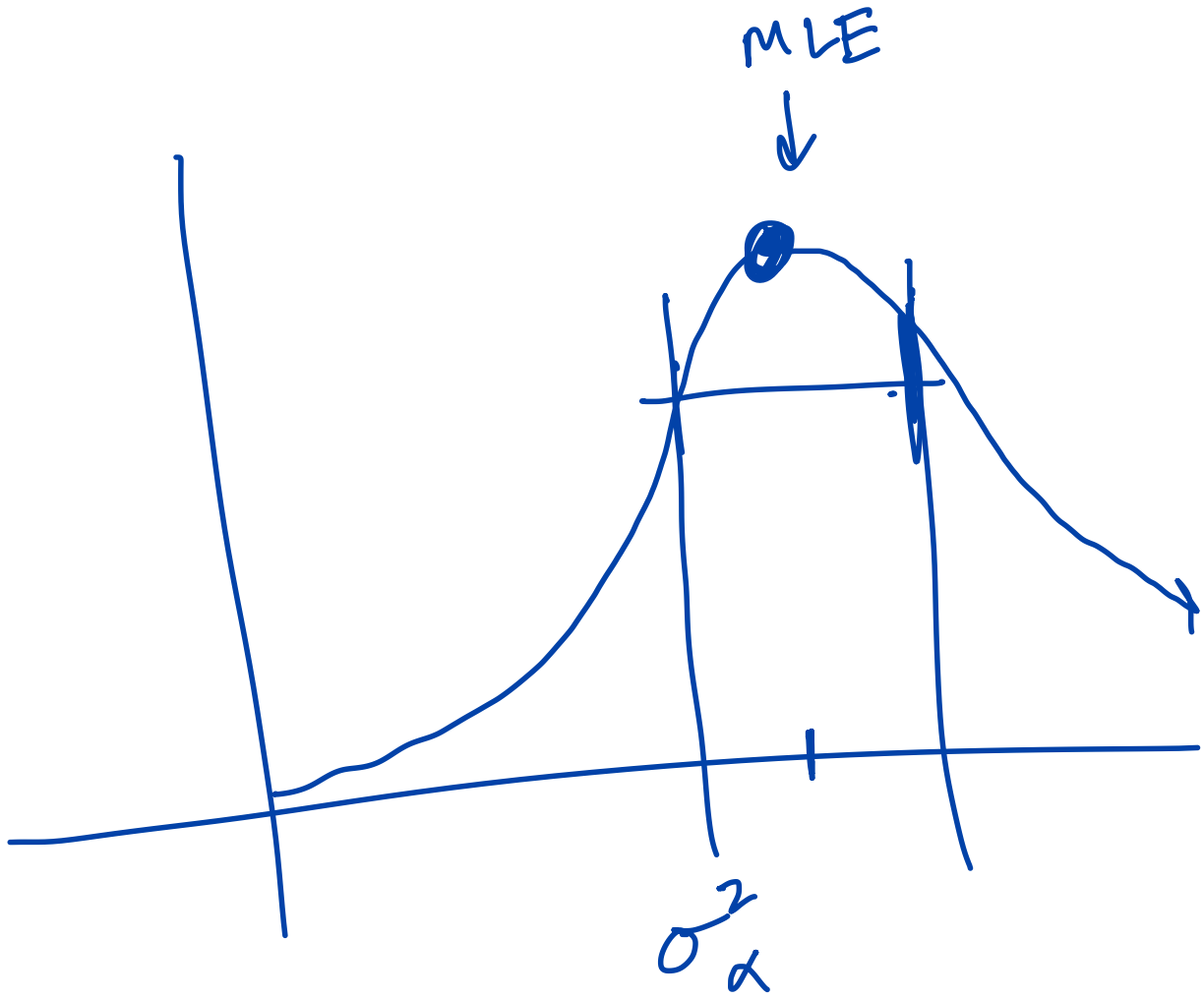
Could do F-test: $F = \frac{MS_{trt}}{MS_E}$ →

estimate σ_α^2. → $\hat{\sigma}_\alpha^2$.

with balanced data, we can compute (same for all treatments).

$$\hat{\sigma}_\alpha^2 = \frac{MS_{trt} - MS_E}{n}$$

likelihood - based inference



$\hat{\sigma}_\alpha$

Restricted likelihood (REML)

(Re)stricted (Maximum (L)ikelihood)

$\mu, \sigma_\alpha^2, \sigma^2$

$$X_i \overset{iid}{\sim} N(\mu, \sigma^2)$$

Variance components.

REML

Likleihood

$$\hat{\mu}_{MLE} = \bar{X}$$

$$\hat{\sigma}^2_{MLE} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

$$\hat{\mu}_{REML} = \bar{X}$$

$$\hat{\sigma}^2_{REML} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Great news!

for balanced data,

ANOVA-Table (EMS) $\hat{\sigma}_\alpha^2 = \frac{MS_{trt} - MS_E}{n}$

$=$ $\updownarrow$

(REML) $\hat{\sigma}_\alpha^2 =$ REML.

Restricted likelihood

- However, standard ML has a flaw when estimating variance components: it does not account for the degrees of freedom used to estimate the fixed effects (like the grand mean).
- This makes standard ML variance estimates *biased* (they underestimate the true variance).

aov → lme4::lmer

(EMS) (REML)

Restricted Maximum Likelihood (REML)

- To fix the bias in standard ML, modern software uses Restricted Maximum Likelihood (REML).
- REML accounts for the degrees of freedom lost in estimating fixed effects, REML provides *unbiased* estimates of σ^2 and σ^2_α.
- Additionally, REML handles unbalanced data effortlessly and prevents the impossible "negative variance" estimates that can occur when using classical EMS methods.

$$\sigma^2_\alpha = \frac{MS_{trt} - MS_E}{n} < 0$$

if $MS_E > MS_{trt}$.

Estimating variance components in R

- For variance components (random effects) We will shift from our `aov()` workflow and use the `lmer` function from the `lme4` package to do this in R:

aov(response ~ randomEffect, data=data)

```
# install.packages("lme4")
library(lme4)
lmer(response ~ 1 + (1 | randomEffect), data = data)
```

On the next slide, we will use this for the same `synthData` dataset we used the aov function for.

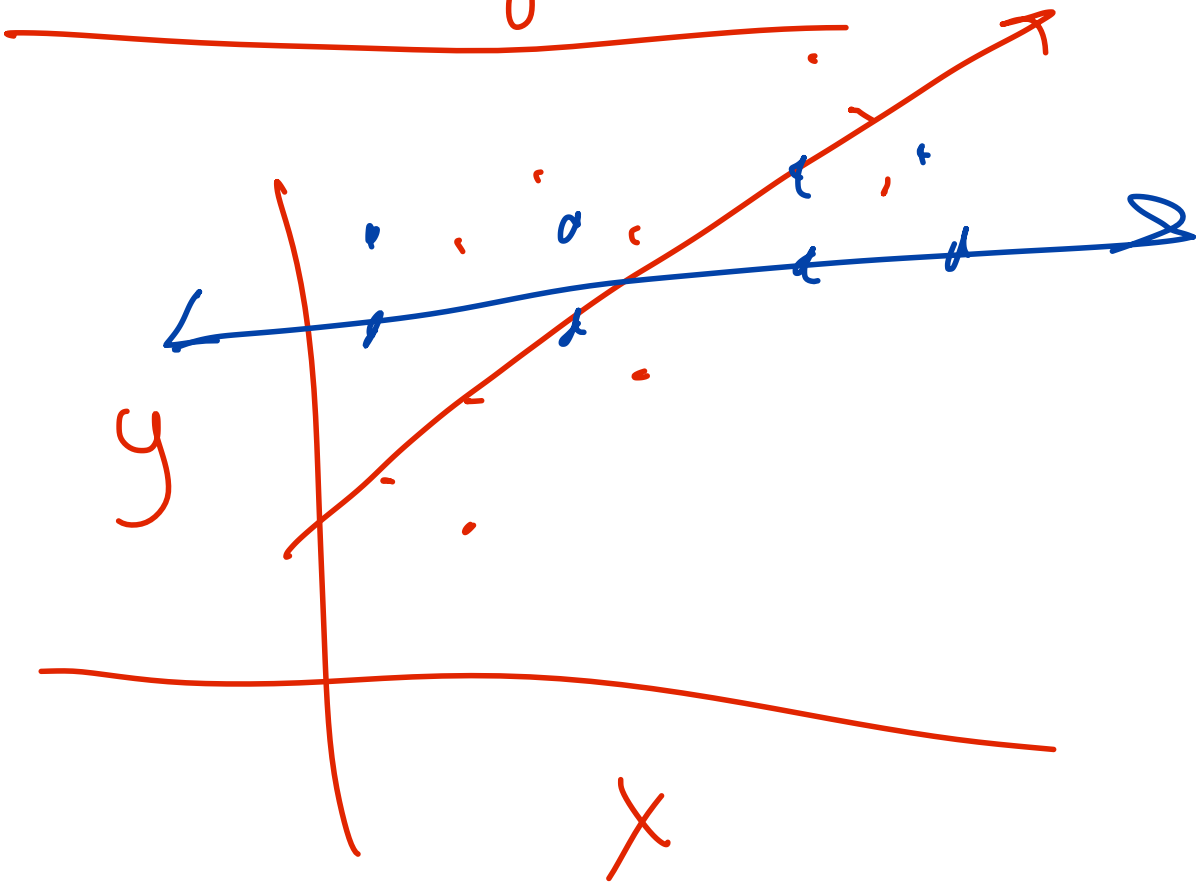
Example: Synthetic data

```
# Old Code:
# aov(response ~ factor(trt), data = synthData)
synthData$trt <- factor(synthData$trt)
lmer(response ~ (1 | trt), data = synthData)

Linear mixed model fit by REML ['lmerMod']
Formula: response ~ (1 | trt)
   Data: synthData
REML criterion at convergence: 94.5478
Random effects:
 Groups   Name        Std.Dev.
 trt      (Intercept) 1.9568
 Residual             0.8788
Number of obs: 32, groups:  trt, 4
Fixed Effects:
(Intercept)
      12.78
```

$\leftarrow \sigma_a$, not σ_a^2 (square this)

$\leftarrow \sigma$, not σ^2 (square this)

5100: Regression



$y \sim x$

$y \sim (x \mid \text{rand} \mathbb{E})$

Example: Penicillin

- We will first consider the built-in `Penicillin` data in R, which contains the reduced diameter of bacterial growth, after penicillin application. Each measurement is taken on a specific `plate`, which is our random effect.

```
data("Penicillin")

output <- lmer(
  diameter ~ 1 + (1 | plate),
  data = Penicillin, REML = TRUE   # default
)
```

Example: Penicillin II

- What is the notation here doing? `1 + (1 | plate)`.
  - The model we are fitting is $y_{ij} = \mu + \alpha_i + \epsilon_{ij}$.
  - The `1` in the beginning is not strictly needed, but represents the grand mean μ.
  - The `(1 | plate)` represents the random effect. In `lme4`, the syntax `(A | B)` implies that B is the random effect grouping variable, and A is the effect that is random. For a simple random effects model, the mean (`1`) is always the thing that is random.
- Now let's look at the output.

Example: Penicillin III

```
summary(output)

Linear mixed model fit by REML ['lmerMod']
Formula: diameter ~ 1 + (1 | plate)
   Data: Penicillin

REML criterion at convergence: 613.3

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.41583 -0.55523  0.01212  0.59232  1.95241

Random effects:
 Groups   Name        Variance Std.Dev.
 plate    (Intercept) 0.09509  0.3084
 Residual             4.03333  2.0083
Number of obs: 144, groups:  plate, 24
```

Comparing to `aov`

- Since the data are balanced, the estimated parameters of this new `lmer` function will match what we can compute with `aov`:

```
# Compute (MS_plate - MS_E) / n, n = 6
summary(aov(diameter ~ plate, data = Penicillin))
```

```
            Df Sum Sq Mean Sq F value Pr(>F)
plate       23  105.9   4.604   1.141  0.313
Residuals  120  484.0   4.033
```

g-1

g = 24

matches REML

N - g = 120

N = 144

144/24 = 6

$\hat{\sigma}^2_\alpha = \frac{4.604 - 4.033}{6} = 0.95$

$F = \frac{4.604}{4.033}$

Hypothesis testing

- The `lmer` function uses REML to get parameter estimates, i.e., $\hat{\sigma}^2_\alpha$.
- If we want to perform a hypothesis test for $\sigma^2_\alpha > 0$, we have two main options:
  1. Refit the model with `aov()`, and use the built-in F-test.
     - **Advantage:** Exact F-tests. These are more powerful.
     - **Disadvantage:** Only works if data are balanced, and the output P-value only matches if there is a single random factor in the model (otherwise, can compute by hand).
  2. Use likelihood based inference procedures (shown below).
     - **Advantage:** Works whether or not data is balanced, and output P-values always "correct". More robust to violations of assumptions.

Hypothesis testing II

- **Disadvantage:** Based on large-sample sizes, so less powerful if samples are small.

- We have already seen analysis of 1-factor random effects with `aov`, now let's do the second approach.

Likelihood Ratio Tests (LRT)

- The profile likelihood confidence intervals are based on the Likelihood Ratio Test.
- Instead of using exact F-tests derived from ratios of Mean Squares, modern mixed modeling often uses the Likelihood Ratio Test (LRT) for hypothesis testing.
- The LRT explicitly compares two competing models:
  1. A **Full Model** that includes the random effect.
  2. A **Reduced Model** (Null Model) that removes the random effect.

Likelihood Ratio Tests (LRT) II

- The test measures how much the likelihood (fit) of the data drops when the random effect is removed.

$$\Lambda = \frac{\text{Likelihood}_{\text{reduced}}}{\text{Likelihood}_{\text{Full}}}, \quad \Longrightarrow \quad -2\log\Lambda \sim \chi^2_d$$

Likelihood Ratio Tests (LRT) III

- Using properties of logarithms, we often consider differences in log-likelihood:

$$\ell_{\text{full}} - \ell_{\text{reduced}} \sim \frac{1}{2}\chi^2_d,$$

where d is the difference in number parameters between the models.

Likelihood Ratio Tests (LRT) IV

- The test can be extended for restricted likelihood ratio tests (RLRT); the idea is similar: fit two models (full vs reduced) using REML, compare differences in the restricted log-likelihoods.
- Rather than fitting two models manually and comparing, we will use the `lmerTest` package to do this automatically for us.

Likelihood Ratio Tests (LRT) V

```
# install.packages("lmerTest")
lmerTest::ranova(output)  # Will do RLRT or LRT, depending on ob

ANOVA-like table for random-effects: Single term deletions

Model:
diameter ~ 1 + (1 | plate)
             npar  logLik    AIC     LRT Df Pr(>Chisq)
<none>          3 -306.63 619.26
(1 | plate)     2 -306.71 617.43 0.17402  1     0.6766
```

H_0: reduced model is as good as full model.

Confidence intervals

confint

- We can get confidence intervals for our estimates using the `confint` function.
- The two primary methods will be `profile` (default) and `boot`.
- This computes profile likelihood confidence intervals, which are effectively the same as a **likelihood ratio test** (LRT).

Confidence intervals II

```
set.seed(123)
confint(output, signame = FALSE, level = 0.95, method = "boot")
```

```
                            2.5 %     97.5 %
sd_(Intercept)|plate  0.000000  0.8138926
sigma                 1.771368  2.2218709
(Intercept)          22.656619 23.3378052
```

conf int for σ^2_α

```
confint(output, signame = FALSE)
```

```
                            2.5 %     97.5 %
sd_(Intercept)|plate  0.000000  0.8458197
sigma                 1.778727  2.2792231
(Intercept)          22.614951 23.3294931
```

Factorial Random Effects Models

Factorial models

- Just like with fixed effects, we may want to consider random effects with multiple random treatment factors.
- Again, we use random effects when we aren't interested in the specific treatment effects chosen, but the larger population of treatment effects.

Example: Carton Factory (continued)

- Let's expand our motivating example of the Carton Factory.

Two-way random effects model

Suppose the carton factory has multiple machines and would like to know the variance in the effect of these machines, but there are also different machine operators, and their effect may be random.

- We randomly select machines and operators, and have each operator work on each each machine and produce $n = 4$ cartons each.
- We have a two-way, completely randomized factorial design.

Example: Carton Factory (continued) II

- In the design proposed above, a two-way random effects model is natural:

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

with

$$\alpha_i \overset{iid}{\sim} N(0, \sigma^2_\alpha) \quad \beta_j \overset{iid}{\sim} N(0, \sigma^2_\beta)$$

$$(\alpha\beta)_{ij} \overset{iid}{\sim} N(0, \sigma^2_{\alpha\beta}) \quad \epsilon_{ijk} \overset{iid}{\sim} N(0, \sigma^2)$$

- The variance of each observation is

$$\mathrm{Var}(y_{ijk}) = \sigma^2_\alpha + \sigma^2_\beta + \sigma^2_{\alpha\beta} + \sigma^2.$$

Example: Carton Factory (continued) III

- This idea once again can generalize to higher-order factorial models

Parameter estimation

- The parameter estimation will still happen in a similar way: we can use ML or REML.
- If we are primarily interested in the error components, REML is unbiased, and generally preferred.

```
carton <- read.csv('carton.csv')
output <- lmer(
  # thickness ~ 1 + (1 | Machine) + (1 | Operator) + (1 | Machin
  thickness ~ (1 | Machine * Operator),
  data = carton,
  REML = TRUE
)
```

Shortcuts still work!

$$A * B = A + B + A:B$$

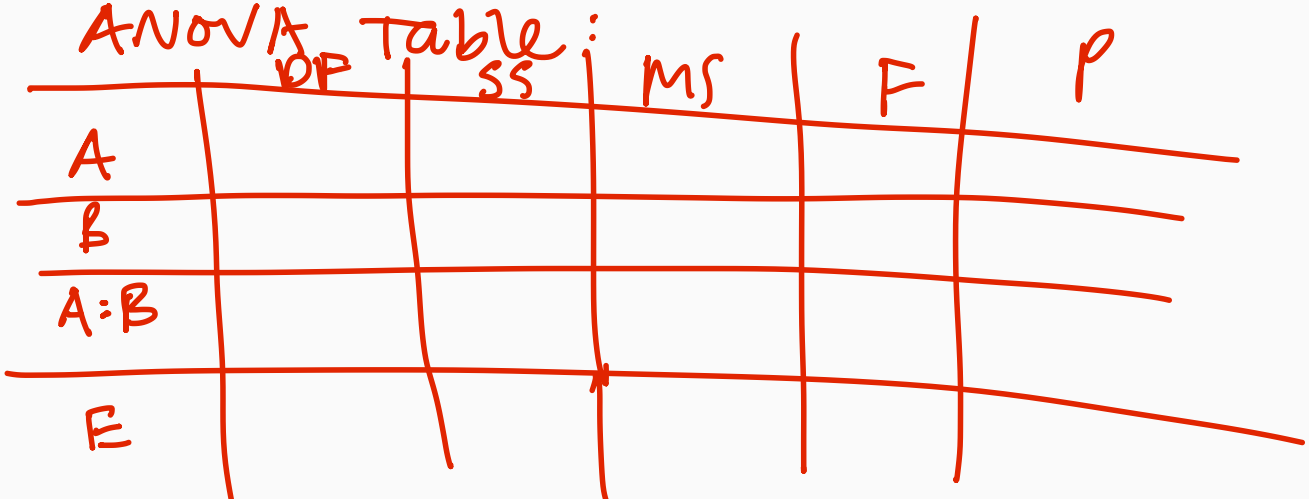
$$(1 \mid A * B) = (1 \mid A) + (1 \mid B) + (1:A:B)$$

```
summary(output)

Linear mixed model fit by REML ['lmerMod']
Formula: thickness ~ (1 | Machine * Operator)
   Data: carton

REML criterion at convergence: 317.2

Scaled residuals:
     Min       1Q   Median       3Q      Max
-1.58412 -0.78134 -0.07557  0.72636  2.15223

Random effects:
 Groups           Name        Variance Std.Dev.
 Machine:Operator (Intercept)  3.782   1.945
 Machine          (Intercept)  2.083   1.443
 Operator         (Intercept) 11.024   3.320
 Residual                      1.465   1.210
```

```
Number of obs: 80, groups:
Machine:Operator, 20; Machine, 5; Operator, 4

Fixed effects:
            Estimate Std. Error t value
(Intercept)   46.671      1.838   25.39
```

Hypothesis Testing (F-test)

- Recall that the F-test is exact, and valid for testing the hypothesis $\sigma^2_A, \sigma^2_B, \sigma^2_{AB} > 0$.
- To do this, we would re-fit the model with `aov()`.
- HOWEVER: we must be extremely careful. The output P-values are INCORRECT, because of the factorial structure.
- Exact P-values can be computed by hand (visit office hours if you're interested). However, we will instead use the likelihood based inference methods.



Confidence intervals

- We can obtain confidence intervals the same way as before. (can be slow).

```
set.seed(444)
# Or method = 'boot'
confint(output, method = 'profile', signame = FALSE)
```

```
                                    2.5 %    97.5 %
sd_(Intercept)|Machine:Operator  1.296337  3.204244
sd_(Intercept)|Machine           0.000000  3.800128
sd_(Intercept)|Operator          1.418039  7.402798
sigma                            1.022308  1.464003
(Intercept)                     42.675508 50.666713
```

Hypothesis testing (likelihood based)

- We can also do hypothesis testing, but it can get tricky due to the interaction effect.
- The `lmerTest::ranova` function just drops the term from the model, and compares REML (or ML). This may not be meaningful for all variables in the presence of an interaction effect.
- Instead, we often will rely on the confidence intervals.

Hypothesis testing (likelihood based) II

```
# May not be meaningful for main (random) effects if interaction

lmerTest::ranova(output)

ANOVA-like table for random-effects: Single term deletions

Model:
thickness ~ 1 + (1 | Machine) + (1 | Operator) + (1 | Machine:Op
                        npar  logLik    AIC    LRT Df
<none>                     5 -158.62 327.24
(1 | Machine)              4 -159.67 327.35  2.105  1
(1 | Operator)             4 -164.39 336.77 11.531  1
(1 | Machine:Operator)     4 -180.10 368.19 42.952  1
                        Pr(>Chisq)
<none>
(1 | Machine)            0.1468521
(1 | Operator)           0.0006846 ***
```

Summary

- We fit a random effects model if we are interested in the variance components specifically, rather than specific treatment effects.

- With random effect, more treatment groups, not more observations per trt.

* g groups, n per group

→ we want more g: $g \geq 10$

Mixed Effects Models

- One of the most common ways we see a random effect is in the context of a mixed effects model.

Mixed Effects Modes: Motivation

'**Idea:** We wish to estimate the fixed effects of specific treatments, while accounting for nuisance or contextual factors that are randomly drawn from a larger population.

Linear Mixed effect models

LME LMM

Mixed Effects: an introduction II

- Mixed effects models are common for block designs:
  - We have a treatment factor A of direct interest (**fixed effect**).
  - We have a blocking factor B (**random effect**).
- Appropriate if blocks are randomly drawn from a population of possible blocks.

Model Formulation

- Consider a Randomized Complete Block Design (RCBD) with a fixed treatments and b randomly selected blocks:

$$y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$$

- Here, α_i are **fixed parameters** (we assume $\sum_i \alpha_i = 0$), while β_j and ϵ_{ij} are **independent random variables**:

$$\beta_j \overset{iid}{\sim} N(0, \sigma^2_\beta), \qquad \epsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2).$$

A has g-levels
B has r-levels
$g \cdot r = N$ total observations.
+1 mean, +g for treatment, +1 for σ^2_β, +1 for σ^2

Model Formulation II

- Because β_j is random, observations within the same block share a common random shift, inducing correlation:

$$\mathrm{Var}(y_{ij}) = \sigma_b^2 + \sigma^2, \qquad \mathrm{Cor}(y_{ij}, y_{ij'}) = \frac{\sigma_b^2}{\sigma_b^2 + \sigma^2} \quad (j \neq j').$$

- **Key advantage:** Treating blocks as random allows our conclusions about the fixed treatments (α_i) to generalize to the broader population of blocks, rather than being restricted to the specific blocks in our experiment.

Inference in Mixed Effects Models

- When our model contains both fixed and random effects, we separate our inference workflow into two distinct steps using `lme4` and `lmerTest`:

lmer_model ← lmer(...)

- **1. Fixed Effects (Treatments):** Use F-tests via `anova(lmer_model)`.
  - `lmerTest` uses REML to estimate variances and applies the Satterthwaite approximation to calculate fractional denominator degrees of freedom.

- **2. Random Effects (Blocks):** Use Restricted Likelihood Ratio Tests via `lmerTest::ranova(model)`.
  - If your model includes a random interaction (e.g., Block $\times$ Treatment), test the interaction first.

Inference in Mixed Effects Models II

- If the interaction is significant, retain it and stop. Do not test the main random effect.
- If non-significant, you can test the main random block effects.

Example: Ergonomic Stools (RCBD)

- Consider the `ergoStool` dataset (from the `nlme` package).
- The data come from a Randomized Complete Block study, with the treatment being stool-type ($g = 4$ levels), and the block the test subject ($r = 9$).
- We measure the physical effort required to arise from 4 stool types (**fixed treatments**) across 9 randomly selected subjects (**random blocks**).

Example: Ergonomic Stools (RCBD) II

```r
library(lme4)
library(lmerTest)
data(ergoStool, package = "nlme")

# Fit mixed model: Subject is random, Type is fixed
mod_ergo <- lmer(effort ~ Type + (1 | Subject), data = ergoStool
```

Testing Fixed vs. Random Effects

- First, we test the fixed treatment effect (Type) using the Satterthwaite approximated F-test:

```
anova(mod_ergo, type = 3)

Type III Analysis of Variance Table with Satterthwaite's method
     Sum Sq Mean Sq NumDF DenDF F value    Pr(>F)
Type 81.194  27.065     3    24  22.356 3.935e-07 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Testing Fixed vs. Random Effects II

- Next, we may be interested in the estimated variance component for the subject-to-subject variation:

Testing Fixed vs. Random Effects III

```
# Either summary, or just inspect the object
mod_ergo

Linear mixed model fit by REML ['lmerModLmerTest']
Formula: effort ~ Type + (1 | Subject)
   Data: ergoStool
REML criterion at convergence: 121.1308
Random effects:
 Groups   Name        Std.Dev.
 Subject  (Intercept) 1.332
 Residual             1.100
Number of obs: 36, groups:  Subject, 9
Fixed Effects:
(Intercept)       TypeT2       TypeT3       TypeT4
     8.5556       3.8889       2.2222       0.6667
```

$\leftarrow \sigma_B = 1.332$

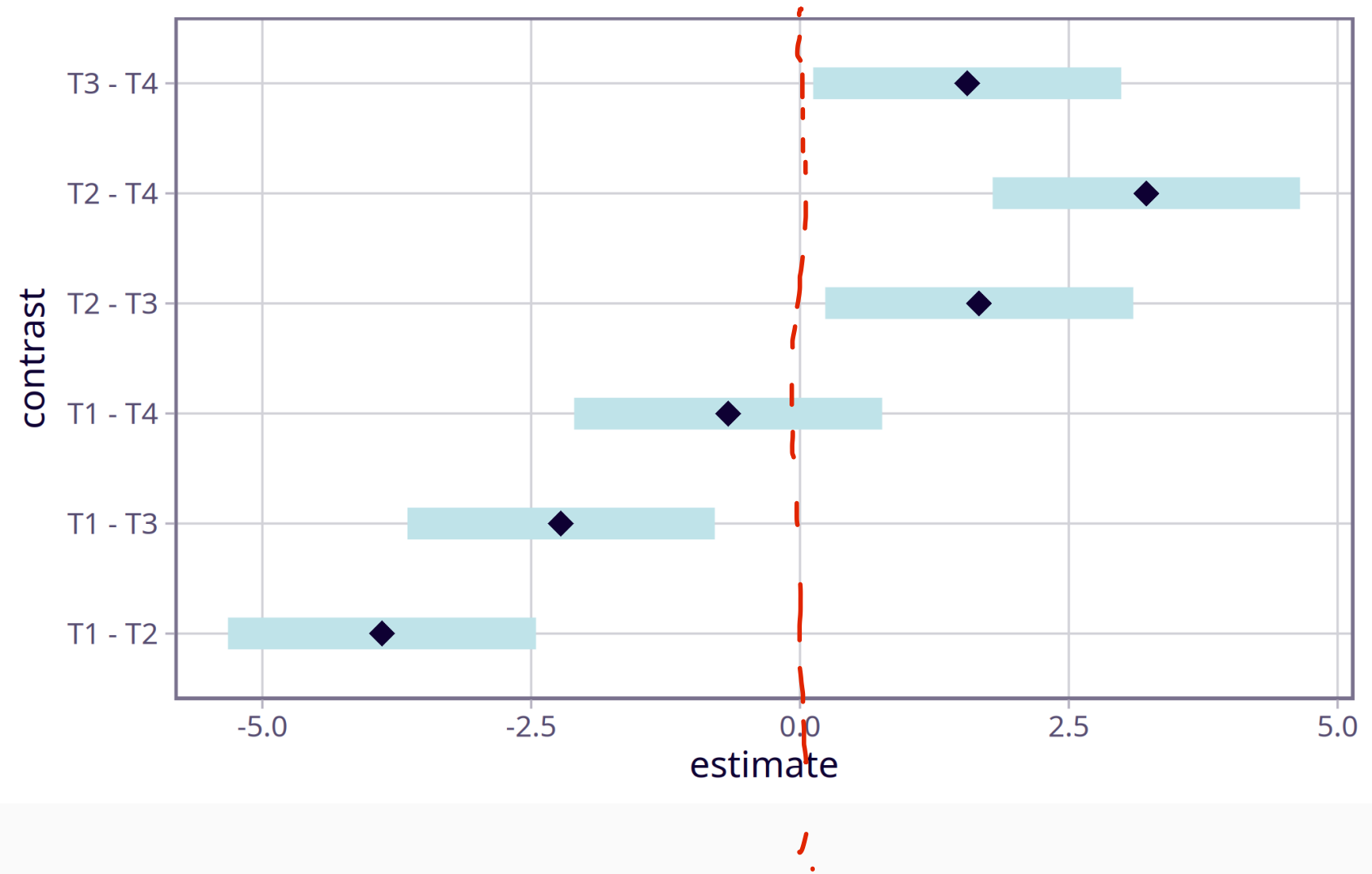
- Next, we test whether the baseline random variance between

Post-Hoc Comparisons with Mixed Models

- Just like with fixed-effects models, we can use the `emmeans` package on an `lmer` object to estimate marginal means and perform pairwise comparisons for our fixed treatments.
- `emmeans` automatically incorporates the random effect variance into the standard errors and uses the Satterthwaite degrees of freedom.

```
library(emmeans)
em_mod <- emmeans(mod_ergo,  ~ Type)
plot(pairs(em_mod))
```

Post-Hoc Comparisons with Mixed Models II



Factorial Mixed Models: The Interaction Rule

- If we had replicated data (multiple trials per subject per stool), we would include a Subject $\times$ Type interaction.
- Because Subject is random, the interaction is also a random effect:

$$y_{ijk} = \mu + \alpha_i + b_j + (\alpha b)_{ij} + \epsilon_{ijk}.$$

- **In R:** `lmer(y ~ Type + (1 | Subject) + (1 | Type:Subject), data = df)`
- Using `anova(lmer_mod)` automatically tests the fixed main effect against the random interaction variance (using fractional degrees of freedom)—avoiding complex manual Expected Mean Squares algebra entirely!

Incorrect: Type + (1 | Subject * Type)

References and Acknowledgements

Oehlert GW (2000). *A First Course in Design and Analysis of Experiments*. W. H. Freeman. ISBN 0-7167-3510-5.

- Compiled on November 13, 2026 using R version 4.6.2.
- Licensed under the Creative Commons Attribution-NonCommercial license. Please share and remix non-commercially, mentioning its origin.
- We acknowledge students and instructors for previous versions of this course / slides.