# Two-Way ANOVA vs. Mixed Model

## A Tale of Two Thieves

Statistical Consulting

November 11, 2025

# Traditional Two-Way ANOVA (Both Fixed)

**Model Specification:**

$$\text{ASSAY}_{ijk} = \mu + \beta_i \cdot \text{METHOD}_i + \alpha_j + \gamma_{ij} + \varepsilon_{ijk}$$

**Where:**

- $\mu$ = Grand mean
- $\beta_i$ = Fixed METHOD effect (INTM or UNIT)
- $\alpha_j$ = Fixed LOCATION effect ($j = 1, \ldots, 6$)
- $\gamma_{ij}$ = Fixed METHOD $\times$ LOCATION interaction
- $\varepsilon_{ijk}$ = Random error, $\varepsilon_{ijk} \sim N(0, \sigma^2)$

**Key Assumption:**

- Inference applies only to these 6 specific locations
- Cannot generalize to other locations

# Mixed Model (METHOD Fixed, LOCATION Random)

**Model Specification:**

$$\text{ASSAY}_{ijk} = \mu + \beta \cdot \text{METHOD} + b_i + \varepsilon_{ijk}$$

**Where:**

- $\mu$ = Grand mean
- $\beta$ = Fixed METHOD effect (applies to any method)
- $b_i$ = Random LOCATION intercept, $b_i \sim N(0, \sigma_b^2)$
- $\varepsilon_{ijk}$ = Error term, $\varepsilon_{ijk} \sim N(0, \sigma_i^2)$

**Key Feature: Heterogeneous Variance**

$$\sigma^2_{\text{INTM}} = 0.7069 \neq \sigma^2_{\text{UNIT}} = 3.4001 \quad (4.81\times \text{ difference})$$

**Generalizability:**

- Inference generalizes to any location within the blender

# Key Differences: Side-by-Side Comparison

| Aspect | Two-Way ANOVA | Mixed Model |
|---|---|---|
| METHOD effect | $\beta_i$ (fixed) | $\beta$ (fixed) |
| LOCATION effect | $\alpha_j$ (fixed) | $b_i \sim N(0, \sigma_b^2)$ |
| Interaction | Estimable | Absorbed in $b_i$ |
| Variance | Homogeneous: $\sigma^2$ | Heterogeneous: $\sigma_i^2$ |
| Generalization | Limited to 6 LOCs | Beyond 6 LOCs |
| Variance decomp. | Not explicit | 31% LOC + 2.3% METHOD |

# Random Slope Model: Allowing METHOD Effects to Vary by Location

**Model Specification:**

$$\mathsf{ASSAY}_{ijk} = \mu + (\beta + b_{i,\mathsf{METHOD}}) \cdot \mathsf{METHOD} + b_i + \varepsilon_{ijk}$$

**Where:**

- $\beta$ = Global METHOD effect (fixed)
- $b_i$ = Location-specific random intercept
- $b_{i,\mathsf{METHOD}}$ = Location-specific METHOD effect (random slope)
- $\varepsilon_{ijk}$ = Random error term

**Distributional Assumptions:**

$$b_i \sim N(0, \sigma_b^2)$$
$$b_{i,\mathsf{METHOD}} \sim N(0, \sigma_{b,\mathsf{METHOD}}^2)$$
$$\varepsilon_{ijk} \sim N(0, \sigma_i^2)$$

# Three Models Comparison

| Feature | Two-Way ANOVA | Mixed Model | Random Slope |
|---|---|---|---|
| METHOD effect | Fixed $\beta_i$ | Fixed $\beta$ | Fixed $\beta$ |
| LOCATION effect | Fixed $\alpha_j$ | Random $b_i$ | Random $b_i$ |
| Interaction | Fixed $\gamma_{ij}$ | Absorbed | Random $b_{i,\text{METHOD}}$ |
| Variance | Homogeneous | Heterogeneous | Heterogeneous |
| Flexibility | Low | Medium | High |
| Complexity | Simple | Moderate | Complex |
| Use when: | LOCs fixed | LOCs random | LOCs random + METHOD varies |

**In Our Data:**

- Likelihood Ratio Test: $\chi^2 = 1.21$, $p = 0.546$ (not significant)
- Conclusion: Random Slope NOT needed $\rightarrow$ basic Mixed Model is adequate

**1. Sampling Design**

### Key Question

Are these 6 locations *fixed* or a *random sample*?

**Answer:** The 6 locations represent a random sample from all possible locations within the V-blender.

- We didn't pre-select "important" locations
- They were chosen to represent the overall blender
- We want to infer about *any* future location

**Conclusion:** LOCATION should be random, not fixed

# Why Mixed Model? (2/4)

**2. Generalizability**

**Research Goal**

Understand blender uniformity for future production batches

- Two-Way ANOVA: "Are these 6 locations different?"
- Mixed Model: "Does location systematically affect content?"

**Example:**

- ANOVA results apply only to Locations 1-6
- Mixed Model generalizes to any location in any future batch

# Why Mixed Model? (3/4)

**3. Precision Comparison**

**Question:** Do INTM and UNIT have different measurement precisions?

**Heterogeneous Variance Structure:**

$$\sigma^2_{\text{INTM}} = 0.7069 \quad \text{vs} \quad \sigma^2_{\text{UNIT}} = 3.4001$$

- INTM is $4.81\times$ more precise than UNIT
- This difference cannot be detected with homogeneous variance
- Precision matters for quality control!

# Why Mixed Model? (4/4)

**4. Explicit Variance Decomposition**

Mixed model partitions total variation into three components:

**Variance Decomposition Formula:**

$$\mathrm{Var}(\mathrm{ASSAY}_{ijk}) = \mathrm{Var}(b_i) + \mathrm{Var}(\varepsilon_{ijk}) + \mathrm{Var}(\beta \cdot \mathrm{METHOD})$$

$$= \underbrace{0.9977}_{\text{Location}} + \underbrace{2.0535}_{\text{Within-Location}} + \underbrace{0.0731}_{\text{METHOD}}$$

**Percentage Decomposition:**

- Between-Location Variance: $\frac{0.9977}{3.1243} = 31.9\%$
  - Random intercepts: $b_i \sim N(0, 0.9977)$
  - Range: 35.01 to 37.98 mg/100mg (2.96 span)
- METHOD Effect: $\frac{0.0731}{3.1243} = 2.3\%$
  - Fixed effect: $\beta \cdot \mathrm{METHOD}$
- Residual Error: $\frac{2.0535}{3.1243} = 65.7\%$
  - Heterogeneous: $\sigma^2_{\mathrm{INTM}} = 0.7069$, $\sigma^2_{\mathrm{UNIT}} = 3.4001$

**Key Insight:** Location effect (31.9%) dominates, not method choice

# Summary: Mixed Model Choice

## When to Use Mixed Model

- Some factors are fixed (METHOD) $\rightarrow$ conditional inference
- Other factors are random (LOCATION) $\rightarrow$ population inference
- Want to partition variance explicitly
- Need heterogeneous variance structure

## Our Case

- METHOD: Fixed (only these 2 types matter)
- LOCATION: Random (represents any location in blender)
- Variance: Heterogeneous (INTM vs UNIT differ $4.81\times$)
- Insight: Location effect (31%) dominates METHOD (2.3%)