Module 6: Random Effects & Linear Mixed Models

A Completely Randomized Design with Random Treatment Effects

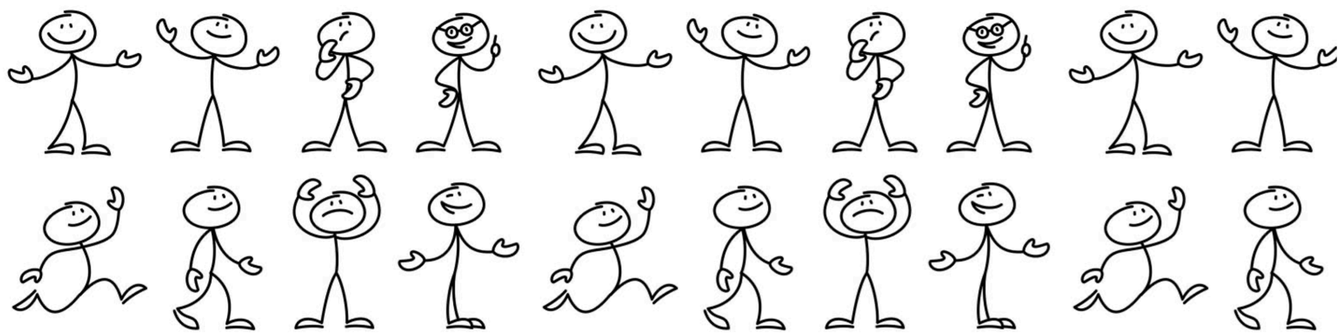
Example 6.2: Fast Food Company

> A company employs many personnel officers.
>
> - Five officers were selected at random.
> - Each officer rated 4 candidates.
>
> Ratings are from 0–100.

The company wants to know:

- The overall mean rating
- The extent of variability in ratings among officers

Study Blueprint

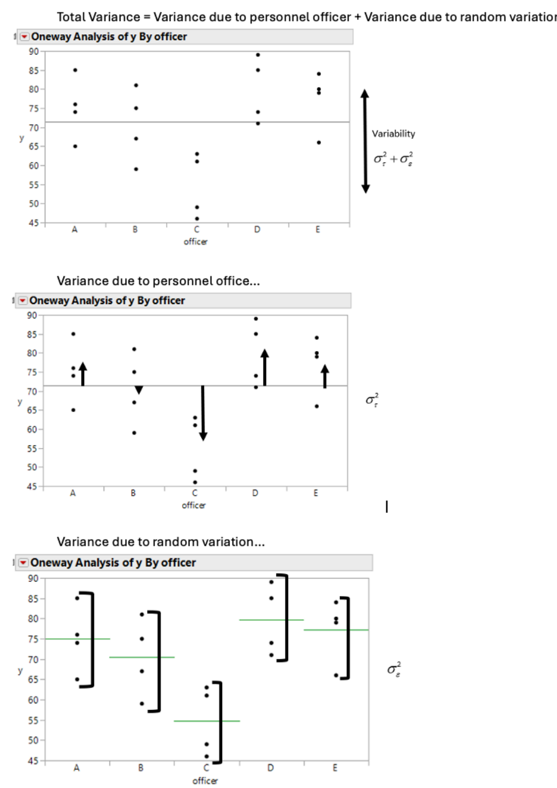


Treatment structure: One-way: Officer (5 levels – officer A, B, C, D, E)

Experimental structure: Officers were randomly assigned to interview r = 4 candidates (e.u.) in a CRD. The rating is recorded for each candidate (m.u.).

Officer = Random Treatment Effect

- We are not comparing the mean ratings of officer A vs B vs C.
- We are estimating variability among officers.



The Data

```r
library(tidyverse)
personnel_data <- read_csv("data/06-personnel-data.csv") |>
  mutate(across(officer:candidate, as.factor))
personnel_data
```

```
# A tibble: 20 × 3
   officer candidate rating
   <fct>   <fct>      <dbl>
 1 A       1             76
 2 A       2             65
 3 A       3             85
 4 A       4             74
 5 B       1             59
 6 B       2             75
 7 B       3             81
 8 B       4             67
 9 C       1             49
10 C       2             63
11 C       3             61
12 C       4             46
13 D       1             74
14 D       2             71
15 D       3             85
16 D       4             89
17 E       1             66
18 E       2             84
19 E       3             80
20 E       4             79
```

Statistical Random Effects Model

$$y_{ij} = \mu + t_i + \epsilon_{ij} \text{ for } i = 1, 2, 3, 4, 5; j = 1, 2, 3, 4$$

with:

- $t_i \sim$ independent $N(0, \sigma_t^2)$
- $\epsilon_{ij} \sim$ independent $N(0, \sigma_\epsilon^2)$
- t_i is independent ϵ_{ij}

where:

- y_{ij} is the rating of the j^{th} candidate by the i^{th} personnel officer
- μ is the overall grand mean rating
- t_i is the effect of the i^{th} personnel officer who is selected at random
- ϵ_{ij} is the error associated with the j^{th} candidate being rated by the i^{th} personnel officer

REML (Residual Maximum Likelihood) Estimation

Goal: estimate variance components

- σ_t^2 = variability **among officers**
- σ_ϵ^2 = variability **within officer**

Why REML?

- Ordinary ML tends to underestimate variance components (especially with small samples)
- REML adjusts for estimating fixed effects (here, μ) before estimating variances

R: Fitting the Random Effects Model

```r
library(lme4)
personnel_mod <- lmer(rating ~ (1 | officer),
                      REML = TRUE,
                      data = personnel_data)
summary(personnel_mod)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: rating ~ (1 | officer)
   Data: personnel_data

REML criterion at convergence: 145.2

Scaled residuals:
    Min      1Q  Median      3Q     Max 
-1.3841 -0.8901  0.2620  0.6496  1.2605 

Random effects:
 Groups   Name        Variance Std.Dev.
 officer  (Intercept) 80.41    8.967   
 Residual             73.28    8.561   
Number of obs: 20, groups:  officer, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)   71.450      4.444   16.08
```

Estimating Variances

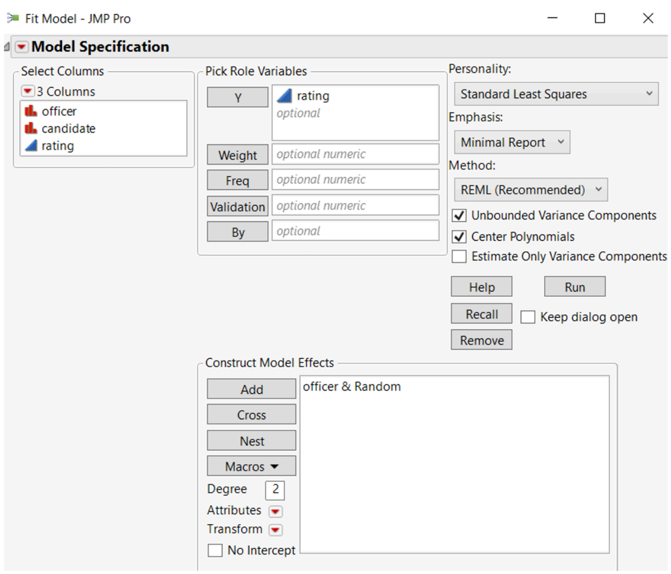
- Total Variance = $\hat{\sigma}_t^2 + \hat{\sigma}_\epsilon^2 =$
- What proportion of the total variance is due to personnel officer? $\frac{\hat{\sigma}_t^2}{\hat{\sigma}_t^2+\hat{\sigma}_\epsilon^2} =$

Estimating the Overall Mean

- $\hat{\mu} =$

JMP: Fitting the Random Effects Model

- `Analyze` > `Fit Model`
- Highlight officer, then `Attributes` ▾ > `Random Effect`
- Set Method: `REML`



REML Variance Component Estimates

Random Effect	Var Ratio	Var Component	Std Error	95% Lower	95% Upper	Wald p-Value	Pct of Total
officer	1.0972538	80.410417	70.133327	-57.04838	217.86921	0.2516	52.319
Residual		73.283333	26.75929	39.989606	175.53909		47.681
Total		153.69375	72.64099	72.600381	514.33185		100.000

Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|
Intercept	71.45	4.443675	4	16.08	<.0001*