SEM-Based Univariate Meta-Analysis

```
Loading required package: OpenMx

OpenMx may run faster if it is compiled to take advantage of multiple cores.

"SLSQP" is set as the default optimizer in OpenMx.

mxOption(NULL, "Gradient algorithm") is set at "central".

mxOption(NULL, "Optimality tolerance") is set at "6.3e-14".

mxOption(NULL, "Gradient iterations") is set at "2".
```

Structural Equation Modeling (SEM) Based Univariate Meta-Analysis

---

1. Introduction

> **i** Key Conceptual Foundations
>
> - **SEM-Meta Integration**:
>   Treat studies as "subjects" in SEM frameworks, where:
>   - **Observed variables** = Reported effect sizes (e.g., SMD, odds ratios).
>   - **Latent variables** = True population effects (modeled as unobserved constructs).

- **Advantages**:
  - Handle missing data via **Full Information Maximum Likelihood (FIML)**.
  - Fix known sampling variances using **definition variables**.
  - Visualize models with path diagrams (e.g., latent heterogeneity).

---

2. Computing Effect Sizes

Standardized Mean Difference (SMD)

Equations:

$$y_{\text{SMD}} = \frac{\bar{X}_1 - \bar{X}_2}{S_{\text{pooled}}}, \quad S_{\text{pooled}} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

Sampling Variance:

$$v_{\text{SMD}} = \frac{n_1 + n_2}{n_1 n_2} + \frac{y_{\text{SMD}}^2}{2(n_1 + n_2)}$$

R Code:

```r
compute_SMD <- function(m1, m2, sd1, sd2, n1, n2) {
  pooled_sd <- sqrt(((n1 - 1)*sd1^2 + (n2 - 1)*sd2^2) / (n1 + n2 - 2))
  smd <- (m1 - m2) / pooled_sd
  v_smd <- (n1 + n2)/(n1 * n2) + smd^2/(2*(n1 + n2))
  return(data.frame(y = smd, v = v_smd))
}

# Example: Compute SMD for two groups
compute_SMD(m1 = 10, m2 = 8, sd1 = 2, sd2 = 1.5, n1 = 50, n2 = 50)
```

```
         y      v
1 1.131371 0.0464
```

---

3. Fixed-Effect Model

Model Specification

Equation:

$$y_i = \beta_F + e_i, \quad e_i \sim N(0, v_i)$$

SEM Representation: - No latent variables.
- Fixed parameter: β_F (common effect).
- Known sampling variance v_i fixed via definition variables.

R Code:

```r
data_fixed <- data.frame(
  y = c(0.5, 0.7, 0.3),   # Effect sizes
  v = c(0.1, 0.15, 0.2)   # Sampling variances
)

fixed_model <- meta(y = y, v = v, data = data_fixed, model.name = "Fixed Effect")
summary(fixed_model)
```

```
Call:
meta(y = y, v = v, data = data_fixed, model.name = "Fixed Effect")

95% confidence intervals: z statistic approximation (robust=FALSE)
Coefficients:
             Estimate  Std.Error     lbound     ubound z value Pr(>|z|)
Intercept1 5.1538e-01 2.1472e-01 9.4549e-02 9.3622e-01  2.4003  0.01638 *
Tau2_1_1   1.0000e-10         NA         NA         NA      NA       NA
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Q statistic on the homogeneity of effect sizes: 0.4615385
Degrees of freedom of the Q statistic: 2
P value of the Q statistic: 0.7939227

Heterogeneity indices (based on the estimated Tau2):
                              Estimate
Intercept1: I2 (Q statistic)        0

Number of studies (or clusters): 3
```

```
Number of observed statistics: 3
Number of estimated parameters: 2
Degrees of freedom: 1
-2 log likelihood: 0.1660267
OpenMx status1: 5 ("0" or "1": The optimization is considered fine.
Other values may indicate problems.)
```

```
Warning in print.summary.meta(x): OpenMx status1 is neither 0 or 1. You are advised to 'rerun
```

Interpretation: - `Estimate` = $\hat{\beta}_F$ (common effect).
- `Std.Error` = standard error of $\hat{\beta}_F$.

---

4. Random-Effects Model

Model Specification

Equation:

$$y_i = \beta_R + u_i + e_i, \quad u_i \sim N(0, \tau^2), \quad e_i \sim N(0, v_i)$$

SEM Representation: - **Latent variable**: $f_i \sim N(\beta_R, \tau^2)$ (true effect).
- **Observed variable**: $y_i = f_i + e_i$, with $e_i \sim N(0, v_i)$.

Key Metrics: - τ^2: Between-study variance.
- $I^2 = \frac{\tau^2}{\tau^2 + \tilde{v}}$: Proportion of total variance due to heterogeneity.

R Code:

```r
random_model <- meta(y = y, v = v, data = data_fixed, model.name = "Random Effects")
summary(random_model)
```

```
Call:
meta(y = y, v = v, data = data_fixed, model.name = "Random Effects")

95% confidence intervals: z statistic approximation (robust=FALSE)
Coefficients:
             Estimate  Std.Error     lbound     ubound z value Pr(>|z|)
Intercept1 5.1538e-01 2.1472e-01 9.4549e-02 9.3622e-01  2.4003  0.01638 *
Tau2_1_1   1.0000e-10         NA         NA         NA      NA       NA
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Q statistic on the homogeneity of effect sizes: 0.4615385
Degrees of freedom of the Q statistic: 2
P value of the Q statistic: 0.7939227

Heterogeneity indices (based on the estimated Tau2):
                             Estimate
Intercept1: I2 (Q statistic)        0

Number of studies (or clusters): 3
Number of observed statistics: 3
Number of estimated parameters: 2
Degrees of freedom: 1
-2 log likelihood: 0.1660267
OpenMx status1: 5 ("0" or "1": The optimization is considered fine.
Other values may indicate problems.)

Warning in print.summary.meta(x): OpenMx status1 is neither 0 or 1. You are advised to 'reru
```

```
# Calculate I²
I2 <- random_model$I2.values
#cat("I² =", round(I2, 2))
```

---

5. Mixed-Effects Model

Model Specification

Equation:

$$y_i = \beta_0 + \beta_1 x_i + u_i + e_i, \quad u_i \sim N(0, \tau^2)$$

SEM Representation: - **Latent variable**: $f_i \sim N(\beta_0 + \beta_1 x_i, \tau^2)$.
- **Observed variable**: $y_i = f_i + e_i$, with $e_i \sim N(0, v_i)$.

Interpretation: - β_1: Change in effect size per unit increase in x_i.
- $R^2 = \frac{\tau^2_{\text{without } x} - \tau^2_{\text{with } x}}{\tau^2_{\text{without } x}}$: Variance explained by x_i.

R Code:

```r
# Add moderator
data_mixed <- data.frame(
  y = c(0.5, 0.7, 0.3),
  v = c(0.1, 0.15, 0.2),
  year = c(2010, 2015, 2020)  # Moderator
)

mixed_model <- meta(y = y, v = v, x = year, data = data_mixed, model.name = "Mixed Effects")
summary(mixed_model)
```

```
Call:
meta(y = y, v = v, x = year, data = data_mixed, model.name = "Mixed Effects")

95% confidence intervals: z statistic approximation (robust=FALSE)
Coefficients:
              Estimate   Std.Error      lbound      ubound z value Pr(>|z|)
Intercept1  2.7367e+01  1.0769e+02 -1.8370e+02  2.3843e+02  0.2541   0.7994
Slope1_1   -1.3333e-02  5.3473e-02 -1.1814e-01  9.1472e-02 -0.2493   0.8031
Tau2_1_1    1.0000e-10          NA          NA          NA      NA       NA

Q statistic on the homogeneity of effect sizes: 0.4615385
Degrees of freedom of the Q statistic: 2
P value of the Q statistic: 0.7939227

Explained variances (R2):
                       y1
Tau2 (no predictor)     0
Tau2 (with predictors)  0
R2                      0

Number of studies (or clusters): 3
Number of observed statistics: 3
Number of estimated parameters: 3
Degrees of freedom: 0
-2 log likelihood: 0.1044882
OpenMx status1: 5 ("0" or "1": The optimization is considered fine.
Other values may indicate problems.)

Warning in print.summary.meta(x): OpenMx status1 is neither 0 or 1. You are advised to 'rerun
```

```r
# Calculate R²
tau2_without_x <- meta(y = y, v = v, data = data_mixed)$tau2
tau2_with_x <- mixed_model$tau2
R2 <- (tau2_without_x - tau2_with_x) / tau2_without_x
cat("R² =", round(R2, 2))
```

```
R² =
```

---

6. Conceptual Deep Dive

> **Why SEM for Meta-Analysis?**
>
> 1. **Latent Variables**:
>   - Separate true effects (f_i) from sampling error (e_i).
>   - Example: If $\tau^2 = 0$, all variability is due to sampling error (fixed-effect model).
> 2. **Definition Variables**:
>   - Fix known sampling variances (v_i) as constants per study.
> 3. **Missing Data**:
>   - FIML retains studies with incomplete data, unlike traditional listwise deletion.

---

7. Summary

Model	Equation	SEM Component	R Function
Fixed-Effect	$y_i = \beta_F + e_i$	No latent variables	`meta(y, v)`
Random-Effects	$y_i = \beta_R + u_i + e_i$	Latent $f_i \sim N(\beta_R, \tau^2)$	`meta(y, v)`
Mixed-Effects	$y_i = \beta_0 + \beta_1 x_i + u_i + e_i$	Latent $f_i \sim N(\beta_0 + \beta_1 x_i, \tau^2)$	`meta(y, v, x)`

Advantages of SEM-Based Meta-Analysis

- **Flexibility**: Extend to multivariate/multilevel models.
- **Precision**: Directly model heterogeneity as latent variance.
- **Robustness**: Integrate with SEM's estimation tools (e.g., FIML, constraints).