

Modelos para Meta-Análise

Revisão Sistemática e Meta-Análise

Marcelo Weber & Nicholas Marino

github.com/nacmarino/maR

Recapitulando

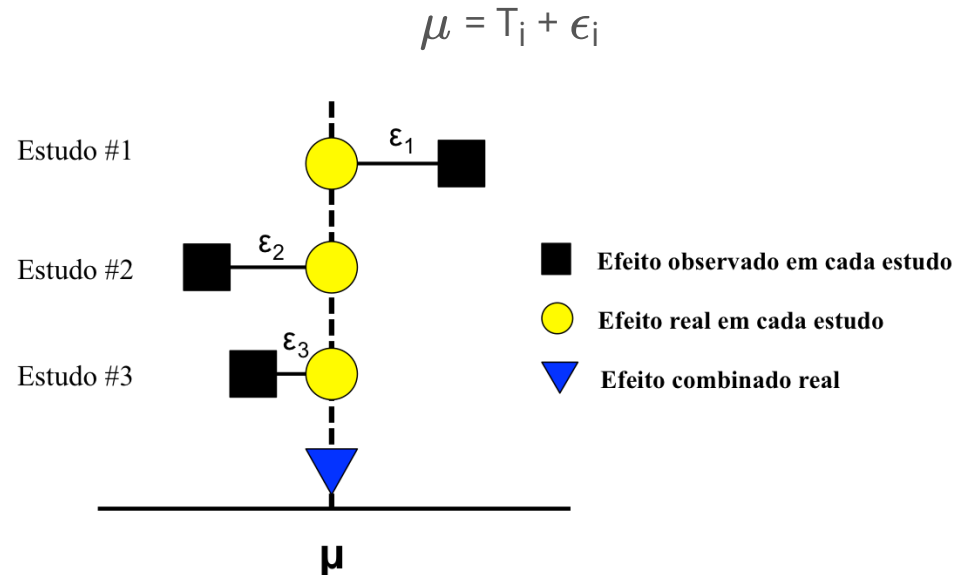
- **Meta-Análise:** "é a análise estatística de uma ampla coleção de resultados de estudos com o propósito de integrar a evidência disponível". (*Glass, 1976*)
- É essencial extrairmos uma métrica de tamanho de efeito e também a sua variância, para que o peso de cada estudo seja proporcional à sua precisão.
- Com estas informações, queremos chegar a um valor que defina a direção e/ou magnitude de um efeito/padrão/processo, a incerteza associada a este valor e fontes que expliquem esta heterogeneidade.

Modelos para Meta-Análise

- Existem diversos tipos de modelo para serem usados, cuja escolha:
 - **Reflete o objetivo da meta-análise:** estimar um efeito vs explorar heterogeneidade;
 - **Define como as estimativas de cada estudo serão combinadas:** todos os estudos são iguais vs estudos diferem em sua 'qualidade';
 - **Descreve as fontes de heterogeneidade entre estudos:** erro amostral vs variação entre estudos;
 - **Determina o tipo de inferência que pode ser feita:** população de estudos analisada vs todos os estudos, até os desconhecidos.
- Desta forma, a escolha do modelo reflete o(s) objetivo(s) da sua meta-análise;
- No geral, os modelos de meta-análise se dividem em dois grandes grupos: *fixed-effects models* e *random-effects models*.

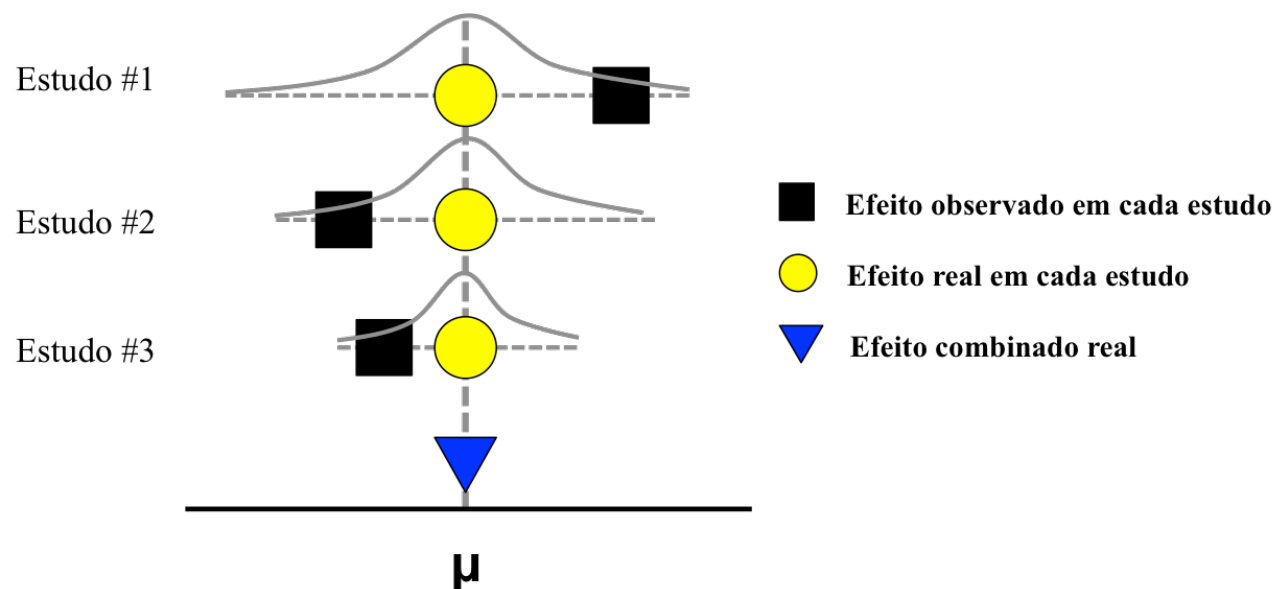
Fixed-effects model

- Usado para estimar a natureza de um efeito;
- Baseado no pressuposto de normalidade;
- Em um fixed-effects model, nós assumimos que:
 - Todos os estudos pertencem a uma mesma população e estão medindo a mesma coisa.
 - A variabilidade existente é explicada unicamente por erro na amostragem em cada estudo;
 - Variabilidade causada por diferenças entre estudos é ínfima ou inexistente.
- Exemplo de fixed-effects model:



Fixed-effects model

- A incerteza na estimativa de cada estudo está contida dentro da variabilidade do efeito comum que estes estão medindo.
- Generalizações são restritas aos estudos incluídos na meta-análise.



- Cada estudo então contribui com uma estimativa do efeito real, que deve ser ponderado pela sua imprecisão ao medir este efeito.

Fixed-effects models

- No geral, o peso de cada estudo (W_i) em uma meta-análise é proporcional ao inverso de sua variância (s_i): quanto maior variância de um estudo, menor o seu peso ao estimar o efeito real.

$$W_i = \frac{1}{s_i}$$

- Neste sentido, o modelo de meta-análise é semelhante à uma *weighted regression*: o efeito real é calculado com base em cada medida de effect size ($ES_i: d_i, LRR_i, r_i, \dots$) e o peso associado à ele (W_i).

$$\mu = \frac{\sum ES_i W_i}{\sum W_i} \text{ (onde } \mu, \text{ é o efeito comum entre todos os estudos)}$$

- A variância em torno de μ (V_T) é dada por:

$$V_T = \frac{1}{\sum W_i}$$

- Com isto, é possível calcular o intervalo de confiança ao redor de μ , assim como realizar testes de significância com o valor estimado de μ (que segue uma distribuição normal):

$$95\% \text{ CI} = \mu \pm 1.96 \sqrt{V_T}$$

$$z = \frac{\mu}{\sqrt{V_T}}$$

No R

- Vamos usar novamente o pacote `metafor`, e explorar o conjunto de dados `dat.bcg`.

```
library(metafor)
dat <- escalc(measure="RR", ai=tpos, bi=tneg, ci=cpos, di=cneg, data=dat.bcg)
dat
```

trial	author	year	tpos	tneg	cpos	cneg	ablat	alloc	yi	vi
1	Aronson	1948	4	119	11	128	44	random	-0.88931130	3.255848
2	Ferguson & Simes	1949	6	300	29	274	55	random	-1.58538870	1.945811
3	Rosenthal et al	1960	3	228	11	209	42	random	-1.34807310	4.153680
4	Hart & Sutherland	1977	62	13536	248	12619	52	random	-1.44155120	0.0200100
5	Frimodt-Moller et al	1973	33	5036	47	5761	13	alternate	-0.21754730	0.0512102
6	Stein & Aronson	1953	180	1361	372	1079	44	alternate	-0.78611560	0.0069056
7	Vandiviere et al	1973	8	2537	10	619	19	random	-1.62089820	0.2230172
8	TPT Madras	1980	505	87886	499	87892	13	random	0.0119523	0.0039616
9	Coetzee & Berjak	1968	29	7470	45	7232	27	random	-0.46941760	0.0564342
10	Rosenthal et al	1961	17	1699	65	1600	42	systematic	-1.37134480	0.0730248
11	Comstock et al	1974	186	50448	141	27197	18	systematic	-0.33935880	0.0124122
12	Comstock & Webster	1969	5	2493	3	2338	33	systematic	0.4459134	0.5325058
13	Comstock et al	1976	27	16886	29	17825	33	systematic	-0.01731390	0.0714047

No R

- A função `rma` é utilizada para ajustar os modelos de meta-análise no `metafor`.
- `method = "FE"` diz para o `metafor` que você quer usar um fixed-effects model.

```
rma(yi = yi, vi = vi, data = dat, method = "FE")
```

```
##  
## Fixed-Effects Model (k = 13)  
##  
## Test for Heterogeneity:  
## Q(df = 12) = 152.2330, p-val < .0001  
##  
## Model Results:  
##  
## estimate      se      zval      pval      ci.lb      ci.ub      ***  
## -0.4303    0.0405 -10.6247    <.0001    -0.5097    -0.3509      ***  
##  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


Random-effects model

- Usado para estimar a natureza de um efeito;
- Baseado no pressuposto de normalidade;
- Em um random-effects model:
 - Estudos pertencem a diferentes subpopulações que compõem uma população maior: cada estudo está estimando o verdadeiro efeito em sua subpopulação.
 - A variabilidade existente é explicada por erro na amostragem em cada estudo e por diferenças entre estudos.
- Em um random-effects model, temos dois componentes de variância:
 - Variância dentro dos estudos: ϵ_i (**within-study variance**)
 - Variância entre os estudos: τ_i^2 (**between-study variance**)
- Você é capaz de explorar onde existe maior variância nos resultados: dentro ou entre estudos - em outras palavras: *consistência* ou *contingência*.

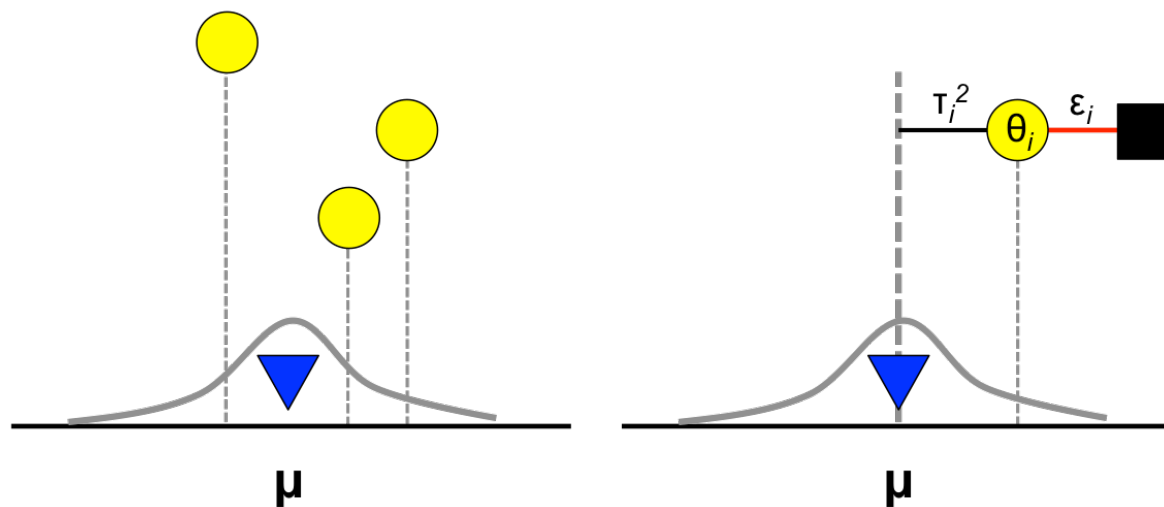
Random-effects model

- Em um random-effects model:

$$T_i = \theta_i + \epsilon_i \text{ (efeito real em cada estudo)}$$

$$\mu = T_i + \tau_i^2 \text{ (efeito combinado real)}$$

■ Efeito observado em cada estudo ● Efeito real em cada estudo ▼ Efeito combinado real

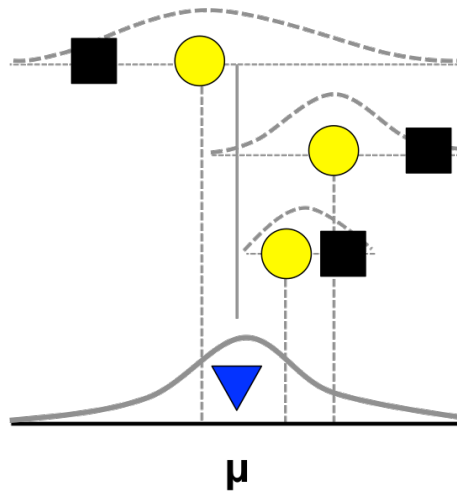


Random-effects model

- O efeito comum que os estudos estão medindo contém a variabilidade entre estudos e também dentro de cada estudo.
- Isto afeta diretamente a estimativa do peso de cada estudo no modelo da meta-análise:

$$W_i = \frac{1}{s_i + \tau^2}$$

- Aqui, τ^2 representa o componente da heterogeneidade total que é originária entre estudos (resolução dos cálculos para chegar a este valor está no Box 9.2, Capítulo 9 do livro de Meta-Análise).
- Generalizações podem ser feitas aos estudos incluídos na meta-análise e também aqueles que são desconhecidos.



No R

- Usando novamente o conjunto de dados `dat.bcg` do pacote `metafor`, vamos ajustar um random-effects model também com a função `rma`.
- Neste caso, você não precisa determinar o argumento `method`: por padrão, a função usa `REML`.

```
rma(yi = yi, vi = vi, data = dat)
```

```
##
## Random-Effects Model (k = 13; tau^2 estimator: REML)
##
## tau^2 (estimated amount of total heterogeneity): 0.3132 (SE = 0.1664)
## tau (square root of estimated tau^2 value):      0.5597
## I^2 (total heterogeneity / total variability):    92.22%
## H^2 (total variability / sampling variability):   12.86
##
## Test for Heterogeneity:
## Q(df = 12) = 152.2330, p-val < .0001
##
## Model Results:
##
## estimate      se      zval      pval      ci.lb      ci.ub
## -0.7145    0.1798   -3.9744    <.0001   -1.0669   -0.3622      ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Fixed- vs Random-effects model

- Qual modelo utilizar?
 - RE é mais realístico à meta-análises em ecologia, uma vez que é impossível replicar perfeitamente os experimentos e sistemas de estudos são naturalmente diferentes;
 - No entanto, RE requer a estimativa de um parâmetro adicional relacionado à variância entre estudos, que pode ter estimativa enviesada por pequeno tamanho amostral ou número de estudos - uso de REML reduz um pouco este problema;
 - Por conta disso, intervalos de confiança estimados com RE models tendem a ser mais amplos do que aqueles estimados com FE models.

Modelo	Estimate	SE	CI.Inferior	CI.Superior
Fixed-effects model	-0.43	0.04	-0.51	-0.35
Random-effects model	-0.71	0.18	-1.07	-0.36

- Note que caso o τ^2 seja 0 (ou muito próximo à zero), a estimativa do peso de cada estudo é reduzida àquela de um fixed-effects model;
- Por conta disso, se você precisa, o ideal é tentar um random-effects model e caso o valor de τ^2 seja baixo aí sim adotar um fixed-effects model.

Mixed-effects models

- É utilizado quando queremos não só estimar um efeito, mas também explorar a heterogeneidade ao redor dele;
- Acomoda uma fração aleatória da heterogeneidade (random-effects) - causada por variância entre- e dentro- dos estudos -, assim como uma fração fixa (característica dos estudos que queremos explorar);
- Neste sentido, ao invés de 'jogar' toda a heterogeneidade entre estudos para um termo de variância, parte dela é atribuída a *moderadores* (veremos mais sobre isso na próxima aula);
- No R, basta usarmos a função `rma` no `metafor` e especificar a variável que queremos analisar com o argumento `mods`;
- Você também pode buscar explorar a heterogeneidade ao redor da estimativa do efeito utilizando um fixed-effects model: basta especificar o argumento `methods = "FE"` na mesma estrutura do modelo abaixo.

```
dat3 <- rma(yi = yi, vi = vi, data = dat, mods = ~alloc)
```

No R

```
##
## Mixed-Effects Model (k = 13; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):      0.3615 (SE = 0.2111)
## tau (square root of estimated tau^2 value):             0.6013
## I^2 (residual heterogeneity / unaccounted variability): 88.77%
## H^2 (unaccounted variability / sampling variability):    8.91
## R^2 (amount of heterogeneity accounted for):             0.00%
##
## Test for Residual Heterogeneity:
## QE(df = 10) = 132.3676, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3):
## QM(df = 2) = 1.7675, p-val = 0.4132
##
## Model Results:
##
##              estimate      se      zval      pval      ci.lb      ci.ub
## intrcpt          -0.5180  0.4412  -1.1740  0.2404  -1.3827  0.3468
## allocrandom       -0.4478  0.5158  -0.8682  0.3853  -1.4588  0.5632
## allocsystematic    0.0890  0.5600   0.1590  0.8737  -1.0086  1.1867
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Outros modelos mais complexos

- Existem outros tipos de modelos de meta-análise que você pode encontrar por aí:
 - **Meta-Regressão:** é mais uma terminologia para representar os modelos em que a heterogeneidade é explicada por variáveis contínuas (embora algumas pessoas também usem este termo quando usam variáveis categóricas);
 - **Modelo Hierárquico ou Multivariado:** usado quando estrutura é similar à uma ANOVA aninhada ou quando existe correlação entre as observações. Em alguns casos, você acaba fazendo este tipo de modelo de forma artificial para explorar heterogeneidade em subgrupos dentro dos grupos testados (no R: `rma.mv`).
 - **Modelo Fatorial:** usado quando a meta-análise segue um desenho fatorial.
 - **Modelo via Generalized Linear (Mixed-Effects) Models:** similar aos fixed-, random- e mixed-effects model, mas permite usar outras famílias de distribuição de variáveis aleatórias (no R: `rma.glmm`).

Observações

- No R, o `metafor` encara que cada linha pertence à um estudo distinto - o que muitas vezes não é o caso (um estudo com múltiplas observações). Se você quiser especificar que a *hierarquia* das observações está no nível do estudo (e não do da observação) você pode usar a função `rma.mv` e especificar o argumento `random = ~ 1|estudo`:

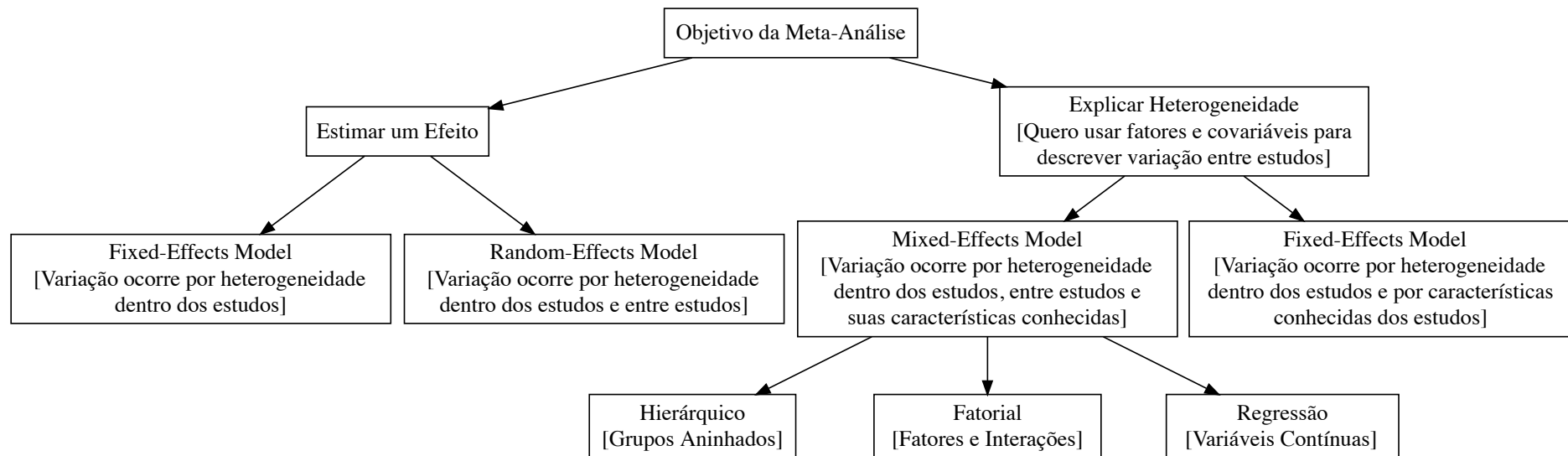
```
rma.mv(yi = yi, V = vi, mods = ~alloc, random = ~1|trial, data = dat)
```

- Você também pode adicionar uma estrutura de correlação às duas observações utilizando o argumento `random` na função `rma.mv`; neste caso, o nível que define a estrutura da correlação vem no lado esquerdo do argumento:

```
rma.mv(yi = yi, V = vi, mods = ~alloc, random = ~author|trial, data = dat)
```

- Você também pode rodar os modelos meta-analíticos usando as funções `lme` (pacote `nlme`) e `lmer` (pacote `lme4`), utilizando o argumento `weights` para determinar o peso de cada observação/estudo.

Só para ficar didático



Resumindo

- Existem diversos tipos de modelos que podem ser usados em uma meta-análise, cada um com um tipo de característica e uso próprio;
- Não existe o melhor modelo para a sua meta-análise, mas sim o modelo que descreve melhor seus objetivos e perguntas;
- Independente do modelo que você use, é boa prática ponderar cada observação/estudo pelo inverso de sua variância.

Literatura Recomendada

1. Nakagawa & Santos, 2012, *Evol Ecol*, Methodological issues and advances in biological meta-analysis
2. Harrison, 2011, *Methods Ecol Evol*, Getting started with meta-analysis
3. Mengersen et al, 2013, Statistical models and approaches to inference, In: *Handbook of meta-analysis in ecology and evolution* (Capítulo 8)
4. Rosenberg, 2013, Moment and least-squares based approaches to meta-analytic inference, In: *Handbook of meta-analysis in ecology and evolution* (Capítulo 9)
5. Mengersen & Schmid, 2013, Maximum likelihood approaches to meta-analysis, In: *Handbook of meta-analysis in ecology and evolution* (Capítulo 10)