

# Effect Size

Revisão Sistemática e Meta-Análise

Marcelo Weber & Nicholas Marino  
[github.com/nacmarino/maR](https://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 importante determinar a sua pergunta e objetivos com clareza: **P**opulation, **I**ntervention, **C**omparison, **O**utcome - (**PICO**).
- A extração de dados precisa ser muito bem planejada e o processo muito bem conduzido.
- A qualidade da meta-análise depende do que você coloca nela: *garbage in, garbage out*.

# Testes de Significância

- É o método mais utilizado para determinar a significância e importância de um efeito: "a ditadura dos valores de p".
  - A hipótese nula raramente é realista ou plausível;
  - Dá importância demais à uma única hipótese que pode ser consistente com os dados (quando na verdade, podem existir múltiplas);
  - Trabalha com aceite ou rejeição de uma hipótese, ao invés de sua plausibilidade;
  - Sofre forte influência do tamanho amostral e de dados aberrantes.

<u>P-VALUE</u>	<u>INTERPRETATION</u>
0.001	HIGHLY SIGNIFICANT
0.01	
0.02	
0.03	
0.04	SIGNIFICANT
0.049	
0.050	OH CRAP. REDO CALCULATIONS.
0.051	ON THE EDGE OF SIGNIFICANCE
0.06	
0.07	HIGHLY SUGGESTIVE, SIGNIFICANT AT THE P<0.10 LEVEL
0.08	
0.09	
0.099	HEY, LOOK AT THIS INTERESTING SUBGROUP ANALYSIS
≥0.1	

# Tamanho do Efeito

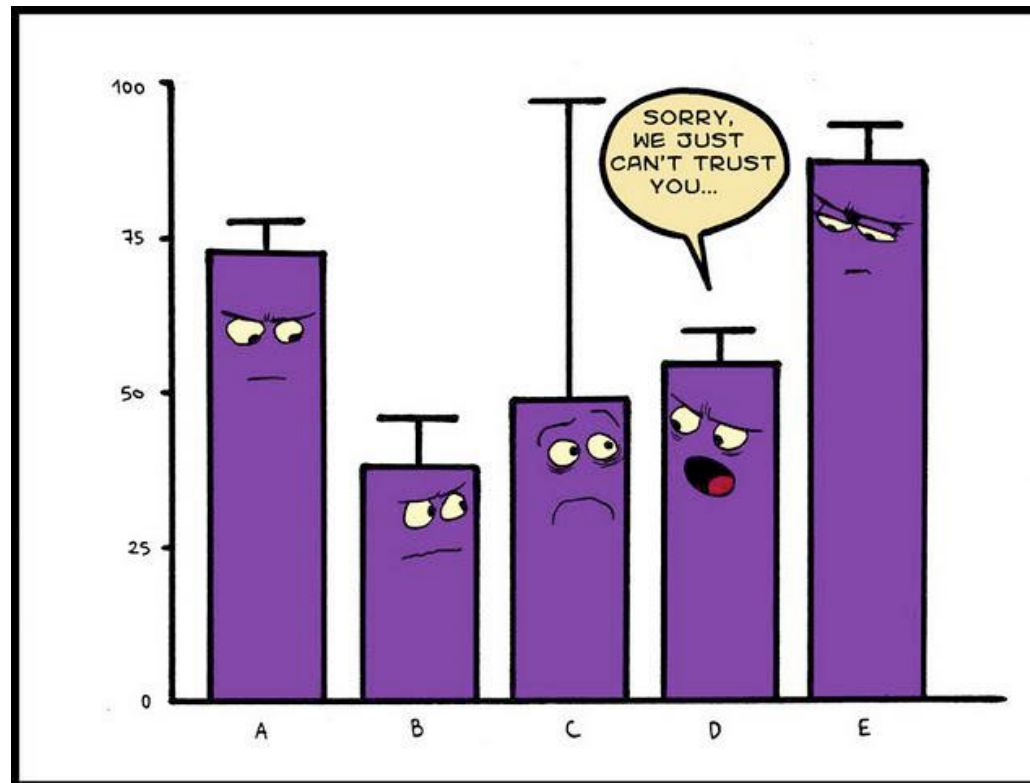
- Tamanho do Efeito vs Significância do Efeito: perspectiva da relevância de um resultado.
- Você já usa métricas de tamanho do efeito sem saber:
  - Quando você estima o efeito de uma variável em uma regressão;
  - Quando você compara os resultados de diferentes tratamentos;
  - Quando você avalia o efeito de diferentes níveis do mesmo tratamento.
- Em uma meta-análise, o tamanho do efeito é um parâmetro estatístico que pode ser usado para comparar, em uma mesma escala, os resultados de diferentes estudos no qual um efeito de interesse tenha sido medido.
  - Mesma resposta, medida de formas diferentes;
  - Mesma medida, expressa em unidades diferentes.

# Intervalos de Confiança

- Interpretado como um envelope dentro do qual um parâmetro de interesse é muito plausível de estar.
  - Se uma população for amostrada  $n$  vezes, o parâmetro de interesse vai estar contido no intervalo de confiança de  $x\%$  em  $z\%$  das amostragens.
  - Combina uma estimativa de uma característica da população em um ponto, com a variabilidade associada à esta estimativa.
- Dificilmente (ou nunca) um teste estatístico será significativo quando o cálculo de um intervalo de confiança sugerir que o valor do parâmetro de interesse pode ser zero.
- Deveria ser o foco principal ao fazermos qualquer inferência, mas não é:
  - Efeito significativo, mas impreciso: a intenção de votos é de 42 pontos, com um desvio de 41 pontos para mais ou para menos.
  - Efeito significativo, mas muito preciso: a intenção de votos é de 42 pontos, com um desvio de 2 pontos para mais ou para menos.

# Pensamento Meta-analítico

- Todo estudo estima o parâmetro correto que descreve a população, mas com precisão diferente.
- Portanto, o valor de cada estudo é dado pelo parâmetro estimado e pela incerteza ao redor dele.



# Pensamento Meta-analítico

- Em uma meta-análise precisamos estimar dois parâmetros:
  - A métrica de effect size *per se*, que é a medida do tamanho do efeito de interesse.
  - Uma métrica que descreva a incerteza sobre a estimativa do effect size - a *variância*.
- Estes dois parâmetros são necessários para rodarmos qualquer modelo de meta-análise:
  - Métrica de effect size é a variável resposta (como já esperado);
  - O inverso da variância é usado para ponderar cada observação: observações mais precisas têm mais peso do que as observações com baixa precisão.
- A falta de algum dos dados pode limitar a escolha da métrica de effect size e também o uso adequado dos modelos.

# Tipos de Medida

- Medidas de tamanho de efeito podem ser postas em dois grupos:
  - **Binárias:** tipo de resposta  $x$  e de resposta  $y$ , resposta  $z$  em uma população  $k, \dots$
  - **Contínuas:** valor médio de  $a$ , força da relação entre  $b$  e  $c, \dots$
- Medidas de efeito também podem ser usadas com diferentes objetivos:
  - **Determinar a magnitude e/ou direção de um fenômeno:** coeficientes de correlação, slopes, valores de média, predominância de um efeito...;
  - **Comparar grupos:** comparação entre médias.
- Em ecologia, a maior parte das meta-análises envolve comparações entre grupos, mas a determinação da magnitude/direção de um fenômeno ecológico também é bastante marcante.
- Para o cálculo de qualquer métrica de effect size e sua variância precisamos:
  1. De uma estimativa de ponto (médias, correlações, slopes);
  2. De uma estimativa de erro (intervalo de confiança, erro padrão, desvio padrão, variância);
  3. De uma estimativa de tamanho da amostra.
- No R, o pacote *metafor* oferece a função `escalc` para o cálculo de diversas métricas de effect size.



# Hedge's $d$

- É um tipo muito comum de métrica em meta-análises na ecologia.
- Principal objetivo é comparar grupos através de suas médias (medida contínua).
- Também conhecido como *Standardized Mean Difference*.
- Estimativa da diferença entre a média de dois grupos (normalmente um tratamento e um controle), padronizado pelo desvio padrão agrupado dos dois grupos e tamanho amostral.
- Limitação: para o cálculo do valor de  $d$  você precisa de estimativas de média, erro e tamanho da amostra para todos os grupos, caso contrário você não consegue calcular esta métrica de effect size.
- Vantagem: pode ser usado com valores de média negativas.

# Hedge's $d$

- O cálculo desta métrica de *effect size* é feito com a fórmula:

$$d = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}} J$$

- Onde  $\bar{Y}$  é a média de cada um dos grupos,  $s$  é o desvio padrão de cada grupo e  $n$  é o número de réplicas em cada grupo (1 e 2).
- $J$  é um fator de correção para viés causado por baixa replicagem (o que é muito comum em estudos ecológicos), e é calculado como:

$$J = 1 - \frac{3}{4(n_1 + n_2 - 2) - 1}$$

# Hedge's $d$

- Uma vez calculada a estimativa do effect size, podemos calcular a variância associada à esta medida:

$$v_d = \frac{n_1 + n_2}{n_1 n_2} + \frac{d^2}{2(n_1 + n_2)}$$

- $d$  pode ter qualquer valor entre  $-\infty$  e  $+\infty$ , e segue uma distribuição normal (no geral).
- Note que esta métrica é descrita como diferença em termos de desvio padrão:
  - $|d| \leq 0.2$ : efeito fraco;
  - $0.2 > |d| < 0.8$ : efeito moderado;
  - $|d| \geq 0.8$ : efeito forte.

# No R

- Vamos carregar um conjunto de dados e dar uma olhada nele.

```
library(metafor)
dat <- get(data(dat.normand1999))
dat
```

##	study	source	nli	mli	sdli	n2i	m2i	sd2i
## 1	1	Edinburgh	155	55	47	156	75	64
## 2	2	Orpington-Mild	31	27	7	32	29	4
## 3	3	Orpington-Moderate	75	64	17	71	119	29
## 4	4	Orpington-Severe	18	66	20	18	137	48
## 5	5	Montreal-Home	8	14	8	13	18	11
## 6	6	Montreal-Transfer	57	19	7	52	18	4
## 7	7	Newcastle	34	52	45	33	41	34
## 8	8	Umea	110	21	16	183	31	27
## 9	9	Uppsala	60	30	27	52	23	20

# No R

- Vamos calcular o valor de  $d$  para cada estudo neste conjunto de dados com o `escalc`.
- "**SMD**" especifica o calculo de  $d$ : **S**tandardized **M**ean **D**ifference.
- Incorpora a correção para amostras pequenas.

```
escalc(measure="SMD", mli=mli, sdli=sdli, nli=nli, m2i=m2i, sd2i=sd2i, n2i=n2i, data=dat)
```

```
##   study      source  nli mli sdli n2i m2i sd2i      yi      vi
## 1     1      Edinburgh 155  55  47 156  75   64 -0.3552 0.0131
## 2     2  Orpington-Mild  31  27   7  32  29    4 -0.3479 0.0645
## 3     3 Orpington-Moderate  75  64  17  71 119   29 -2.3176 0.0458
## 4     4  Orpington-Severe  18  66  20  18 137   48 -1.8880 0.1606
## 5     5   Montreal-Home    8  14   8  13  18   11 -0.3840 0.2054
## 6     6 Montreal-Transfer  57  19   7  52  18    4  0.1721 0.0369
## 7     7      Newcastle  34  52  45  33  41   34  0.2721 0.0603
## 8     8          Umea  110  21  16 183  31   27 -0.4246 0.0149
## 9     9       Uppsala   60  30  27  52  23   20  0.2896 0.0363
```

# Log Response Ratio

- Outro tipo muito comum de métrica em meta-análises na ecologia, que comparar grupos através de suas médias (medida contínua).
- Estimativa da diferença pela razão entre a média de um tratamento e o controle.
- O logarítmo natural é aplicado para normalizar a razão.
- Pode assumir qualquer valor entre  $-\infty$  e  $+\infty$ .
- Limitação: não pode ser usado com valores de média negativas.
- Vantagem: para o cálculo do valor de LRR você precisa somente de estimativas de média.

$$\ln R = \ln\left(\frac{\overline{Y}_1}{\overline{Y}_2}\right) = \ln \overline{Y}_1 - \ln \overline{Y}_2.$$

$$v_{\ln R} = \frac{s_1^2}{n_1 \overline{Y}_1^2} + \frac{s_2^2}{n_2 \overline{Y}_2^2}.$$

# Log Response Ratio

- Lajeunesse (2015) demonstrou que esta formulação do LRR sofre um viés quando tamanho amostral do estudo é pequeno, e pode fornecer estimativas de variância erradas quando a escala dos parâmetros de estudo é próxima a zero (isto é, os valores de média são muito próximos a zero).
- Sugere a utilização de dois outros estimadores para o LRR e sua variância:
  - $LRR^{\Delta}$  (baseado no Método Delta);
  - $LRR^{\Sigma}$  (baseado na regra de Expectativa de Linearidade).
- $LRR^{\Delta}$  fornece estimativas um pouco melhores do que  $LRR^{\Sigma}$ , mas ambos são pouco eficientes quando valores da média beiram o zero.

$$RR^{\Delta} = RR + \frac{1}{2} \left[ \frac{(SD_T)^2}{N_T \bar{X}_T^2} - \frac{(SD_C)^2}{N_C \bar{X}_C^2} \right] \quad \text{var}(RR^{\Delta}) = \text{var}(RR) + \frac{1}{2} \left[ \frac{(SD_T)^4}{N_T^2 \bar{X}_T^4} + \frac{(SD_C)^4}{N_C^2 \bar{X}_C^4} \right]$$

$$\text{var}(RR^{\Sigma}) = 2 \times \text{var}(RR)$$

$$RR^{\Sigma} = \frac{1}{2} \ln \left[ \frac{\bar{X}_T^2 + N_T^{-1} (SD_T)^2}{\bar{X}_C^2 + N_C^{-1} (SD_C)^2} \right] \quad - \ln \left[ 1 + \text{var}(RR) + \frac{(SD_T)^2 (SD_C)^2}{N_T N_C \bar{X}_T^2 \bar{X}_C^2} \right]$$

# No R

- Modificações de Lajeunesse (2015) não foram implementadas (e acho que não serão).
- Vamos calcular o valor de LRR para cada estudo neste conjunto de dados com o `escalc`.
- "ROM" especifica o calculo do LRR: **R**atio of **M**eans.
- O resultado pode ser expresso em %, caso você tire o exponencial do LRR.

```
escalc(measure="ROM", mli=mli, sdli=sdli, nli=nli, m2i=m2i, sd2i=sd2i, n2i=n2i, data=dat)
```

```
##   study          source  nli  mli  sdli  n2i  m2i  sd2i      yi      vi
## 1     1      Edinburgh  155   55   47  156   75   64 -0.3102 0.0094
## 2     2  Orpington-Mild   31   27    7   32   29    4 -0.0715 0.0028
## 3     3 Orpington-Moderate  75   64   17   71  119   29 -0.6202 0.0018
## 4     4  Orpington-Severe  18   66   20   18  137   48 -0.7303 0.0119
## 5     5   Montreal-Home    8   14    8   13   18   11 -0.2513 0.0695
## 6     6 Montreal-Transfer  57   19    7   52   18    4  0.0541 0.0033
## 7     7      Newcastle   34   52   45   33   41   34  0.2377 0.0429
## 8     8          Umea  110   21   16  183   31   27 -0.3895 0.0094
## 9     9       Uppsala   60   30   27   52   23   20  0.2657 0.0280
```



# Correlação

- É valor numérico que mede o grau de associação entre duas variáveis.
- Para o cálculo desta métrica de effect size você só precisa do coeficiente de correlação e do tamanho amostral.
- Muito cuidado deve ser tomado quando usar esta métrica, pois:
  1. Ela se baseia no pressuposto de linearidade entre as duas variáveis independentes;
  2. Pressupõem que não haja problemas com dados aberrantes na correlação.
- Quando os valores de correlação são muito próximos à -1 ou +1, a distribuição dos dados tende a ficar deslocada. Neste sentido, uma opção é realizar a normalização dos dados é a conversão do valor do coeficiente de correção  $r$  para o  $z$  score de Fisher.

$$z = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right), \quad v_z = \frac{1}{n-3}.$$

- Para trazer o valor de  $z$  de Fisher de volta para o coeficiente de correção, basta usar a fórmula:

$$r = \tanh(Z_r) = \frac{e^{2Z_r} - 1}{e^{2Z_r} + 1}.$$

# No R

- Vamos usar outro conjunto de dados, próprio para esta finalidade: `dat.mcdaniel1994`.
- No `escalc`, "COR" especifica o calculo do effect size baseado no valor bruto da correlação, enquanto "ZCOR" o faz baseado no valor do coeficiente de correlção transformado.
- A estimativa da variância não é afetada, uma vez que seu calculo não depende do valor de  $r$ .

```
dat <- dat.mcdaniel1994
str(dat)
escalc(measure="ZCOR", ri=ri, ni=ni, data=dat)
```

```
##      study   ni   ri type struct      yi      vi
## 1         1  123 0.00    j      s 0.0000 0.0083
## 2         2   95 0.06    p      u 0.0601 0.0109
## 3         3   69 0.36    j      s 0.3769 0.0152
## 4         4 1832 0.15    j      s 0.1511 0.0005
## 5         5   78 0.14    j      s 0.1409 0.0133
## 6         6  329 0.06    j      s 0.0601 0.0031
## 7         7  153 0.09    j      s 0.0902 0.0067
## 8         8   29 0.40    j      s 0.4236 0.0385
## 9         9   29 0.39    s      s 0.4118 0.0385
## 10        10  157 0.14    s      s 0.1409 0.0065
```

# Slopes

- Outra medida relevante que pode ser usada como *effect size*;
- Mede a taxa de mudança em um variável dependente de acordo com uma variável independente;
- No R (ou qualquer outro programa estatístico), este valor é representado pelo **estimate** em uma análise;
- Muito cuidado deve ser tomado quando usar esta métrica, pois:
  1. Ela se baseia no pressuposto de linearidade;
  2. Pressupõem que não haja problemas com dados aberrantes;
  3. Quando em um contexto de regressão múltipla, valor do *slope* não é o mesmo do quando quando em uma regressão simples.
- O valor do effect size é o slope em si, e o valor da variância é o valor do erro associado ao slope, elevado ao quadrado;
- Caso a variância não seja fornecida, podemos calcular ela de outras formas (ver Capítulo 6 do livro de Meta-análise para uma explicação mais detalhada).

# Outras Métricas

- Em essência, qualquer estimativa de um dado efeito pode ser usada para o cálculo do *effect size*;
- Existem diversas outras métricas disponíveis:
  - Odds Ratio: define a razão entre a chance de um evento ocorrer em um grupo vs em um segundo grupo;
  - Incidence Rate Ratio: razão entre o número de eventos específicos e o número total de eventos em uma população;
  - Estimativas de Padrões: valores que descrevam algum padrão/processo ecológico (emissão de um gás, densidade de plantas em áreas alagadas de um determinado tipo,...).
- Para este último caso...

```
dat <- dat.senn2013
str(dat)
escalc(measure="MN", mi = mi, sdi=sdi, ni=ni, data=dat)
```

```
##           study  ni treatment    mi    sdi      yi      vi
## 1 De Fronzo (1995) 213 metformin -1.70 1.459 -1.7000 0.0100
## 2 De Fronzo (1995) 209  placebo  0.20 1.446  0.2000 0.0100
## 3 Lewin (2007) 431 metformin -0.74 1.106 -0.7400 0.0028
## 4 Lewin (2007) 144  placebo  0.08 1.004  0.0800 0.0070
```

# Transformando entre métricas

- Alguns estudos podem não te fornecer os valores de média, coeficientes de correlação, slopes e estimativas de erro de forma direta.
- Uma das opções é realizar conversões baseadas nos dados que estão disponíveis (lista completa no Capítulo 13 do livro de Meta-Análise):
  - $t$  para  $d$
  - $r$  para  $d$  (e vice-versa)
  - $F$  para  $d$
  - $z$  para  $d$  (e vice-versa)
  - $\chi^2$  para  $d$
  - Slope para  $r$  (e vice-versa)
  - $t$  para  $r$
  - $F$  para  $r$
  - $z$  para  $r$  (e vice-versa)
  - $\chi^2$  para  $r$
  - $t$  para  $z$

# Resumindo

- Em qualquer trabalho que fizemos, é importante considerar tanto o tamanho do efeito do que estamos medindo, quanto a incerteza existente nesta estimativa.
- Este tipo de pensamento é essencial para passarmos uma estatística frequentista baseada em uma única hipótese, para aquela que contempla múltiplas hipóteses igualmente válidas.
- Em uma meta-análise, é essencial calcularmos ou extrairmos uma métrica de tamanho de efeito e também a sua variância, para que o peso de cada estudo seja proporcional à sua precisão.
- Existem vários tipos de métricas de effect size, mas a escolha de qual delas usar depende em grande parte da sua pergunta.
- Alguns métodos específicos existem para lidar para o cálculo de effect size e sua variância quando houverem dados faltantes ou incompletos.

# Literatura Recomendada

1. Nakagawa & Cuthill, 2007, Biol Rev, Effect size, confidence interval and statistical significance - a practical guide for biologists
2. Lajeunesse, 2015, Ecology, Bias and correction for the log response ratio in ecological meta-analysis
3. Rosenberg et al, 2013, Effect Sizes: Conventional choices and calculations, In: Handbook of meta-analysis in ecology and evolution (Capítulo 6)
4. Mengersen & Gurevitch, 2013, Using other metrics of effect size in meta-analysis, In: Handbook of meta-analysis in ecology and evolution (Capítulo 7)
5. Lajeunesse, 2013, Recovering missing or partial data from studies: a survey of conversions and imputations for meta-analysis, In: Handbook of meta-analysis in ecology and evolution (Capítulo 13)