

# Janaina Silva Fortirer



## Mestranda em Bioinformática, USP.

Pesquisa sobre mudanças climáticas e os efeitos nos parâmetros ecofisiológicos na vegetação brasileira. Mediante a estudo de revisão sistemática, meta-análise e experimentos em laboratório.

Além disso, tenho interesses em educação popular, recursos tecnológicos à educação, *Big Data*, *Machine Learning*, *Blockchain*, empreendedorismo social...

## Meus exercícios

Clique [aqui](#) para acessar meus exercícios

## Trabalho final

---

### Plano A:

#### Contextualização

A meta-análise é um conjunto de métodos estatísticos para comparar quantitativamente resultados de diferentes estudos e que abordam uma mesma questão<sup>1)</sup>. Também determina se o efeito de interesse (média da população) existe em uma condição particular (tratamento), comparando-o com as diferentes pesquisas realizadas. Além de auxiliar na síntese e fornecer um panorama dos resultados sobre determinado tema.

O pesquisador precisa realizar uma etapa anterior a meta-análise, que é a [revisão sistemática](#). Para procurar os trabalhos já realizados que abordam a pergunta que se quer responder. O resultado dessa pesquisa é organizado em uma tabela ([exemplo de tabela](#)). A partir dos dados anotados na tabela, é possível verificar o efeito de interesse de cada estudo. Esses efeitos combinados podem estimar uma grande média do tamanho do efeito e um intervalo de confiança, e por fim testar se esse efeito geral é diferente de zero.

**Para facilitar a visualização dos diferentes resultados obtidos, a proposta da função é gerar um gráfico com a média e calcular o intervalo de confiança para cada estudo.**

[Veja as fórmulas utilizadas para calcular a média e o intervalo de confiança na função.](#)<sup>2)</sup>



## Planejamento da função:

### Description

A função cria um gráfico com a média e calcula o intervalo de confiança.

### Usage

metaR (data, namoc, namot, mc, mt, sdc, sdt, measure = mean, replace = TRUE)

### Arguments

- data = data.frame contendo as variáveis dos argumentos.
- namoc = integer, com o número réplicas do grupo controle.
- namot = integer, com o número réplicas do grupo tratamento.
- mc = numeric, média do grupo controle.
- mt = numeric, média do grupo tratamento.
- sdc = numeric, desvio padrão do grupo controle.
- sdt = numeric, desvio padrão do grupo tratamento.
- measure = logical, indica que a estatística de interesse é a média entre os grupos.
- replace = logical, indica que todos os valores serão substituídos.

### Verificando os parâmetros

- data é um data.frame Se não, escreve: "O argumento data precisa ser um data.frame."
- namoc é um número inteiro e maior que 3? Se não, escreve: "O argumento namoc precisa ser um número inteiro e > 3."
- namot é um número inteiro e maior que 3? Se não, escreve: "O argumento namot precisa ser um número inteiro e > 3."
- mc é do tipo numérico? Se não, escreve: "O argumento mc precisa ser numérico."
- mt é do tipo numérico? Se não, escreve: "O argumento mt precisa ser numérico."
- sdc é do tipo numérico? Se não, escreve: "O argumento sdc precisa ser numérico."
- sdt é do tipo numérico? Se não, escreve: "O argumento sdt precisa ser numérico."
- measure é TRUE? Se não, escreve: "O argumento measure precisa ser TRUE, pois a estatística de interesse dessa função é a média."
- replace é TRUE? Se não, escreve: "O argumento replace precisa ser TRUE."

### Pseudo-código:

1. Leia o data.frame.
2. Escreva: "Os argumentos sdc e sdt são os desvio padrão da amostra? Digite S para sim ou N para não"
3. Se (N), então escreva: "É necessário transformar o erro padrão em desvio padrão."
4. Calcule o desvio padrão a partir do erro padrão.<sup>3)</sup>
5. Indexar o resultado em uma nova coluna do objeto data [i]
6. Escreva: " O erro padrão foi transformado em desvio padrão."
7. Caso a resposta seja (S).
8. Remova NA das amostras

9. Escreva: "Data.frame com NAs omitidos."
10. Para (for) reamostragem com reposição.
11. Cria um vetor com distribuição normal com  $N = 5000$ , média e o desvio padrão de acordo com cada amostra.
12. Faça o bootstrap para cada amostra com a função sample utilizando o  $N$  de cada amostra e o vetor criado com distribuição normal.
13. Guarde o resultado do bootstrap num objeto.
14. Calcule o intervalo de confiança (C.I. 95%).
15. Cria uma matriz.
16. Cria  $x$  linhas e  $x$  colunas na matriz criada (ainda não defini a quantidade de linhas e colunas).
17. Cria legenda no eixo  $x$  (side 1).
18. Cria legenda no eixo  $y$  (side 2) com nome do autor, ano de publicação e  $N$  amostral.
19. Cria uma linha vertical na matriz.
20. Insira média por resultado.
21. Cria segmento com o valor do intervalo de confiança.
22. Salva gráfico

### Saída:

- Gráfico com a média e intervalo de confiança para cada pesquisa.
- Data frame com os valores de intervalo de confiança calculados.

Exemplo de output (gráfico):



## Plano B:

### Contextualização

Vocês já repararam que as lojas de materiais de construção estão sempre lotadas, parecendo shopping? A maioria das pessoas que frequentam esses lugares estão em busca de dar uma repaginada no visu da casa, dar uma modernizada, fazer um puxadinho, ou decidem fazer uma reforma simplesmente porque querem e podem. Ah! Tem os casos específicos, mães, as que todo ano pintam a casa inteira, pois acreditam que passar o ano novo sem pintar a casa dá azar (suspeito que foi algum vendedor de tinta que criou essa lenda). E casos "raros" de pessoas que tem tics por derrubar parede, trocar piso, reformar a sala, mudar isso ou aquilo...

Num mundo ideal, a melhor escolha seria contratar um profissional e realizar um orçamento da quantidade de todos os materiais necessários para fazer a reforma. Porém, de pedreiro e louco todo mundo tem um pouco. Claro, que as pessoas decidem ir comprando ao longo da reforma sem nenhum planejamento de custo e estimativa da quantidade dos materiais a serem utilizados. Então, para auxiliar as pessoas que irão fazer uma reforma, mas que querem economizar e fazer um planejamento dos custos. A proposta de função, é criar um algoritmo que otimize, calcula a quantidade de material a ser utilizado por  $m^2$ , e cria uma tabela com todos os gastos para cada tipo de material escolhido.

## Planejamento da função

### Description

A função irá calcular a quantidade de cada material a ser utilizado por m2, retorna um data frame com os custos e quantidade dos materiais.

### Usage

reformaR (tijolo, m2tijolo, typetijolo, piso, m2piso, typepiso, telhado, m2telha, typetelha, tinta, m2parede, m2teto, typetinta, areia, m3areia, cimento, m2cim, brita, m3brita, data)

### Arguments

- tijolo = numeric, preço do tijolo por m2.
- m2tijolo = numeric, m2 a ser construído.
- typetijolo = string de caracteres com o tipo do tijolo.
- piso = numeric com o preço do m2.
- m2piso = numeric, m2 a ser colocado o piso.
- typepiso = string de caracteres com o tipo de piso.
- telhado = numeric com o preço do m2.
- m2telha = numeric com o m2 da area a ser coberta.
- typetelha = string de caracteres com o tipo da telha.
- tinta = numeric com o preço do m2.
- m2parede = numeric com o m2 total da parede.
- m2teto = numeric com o m2 total do teto.
- typetinta = string de caracteres com o tipo de tinta.
- areia = numeric, preço por m3.
- m3areia = numeric, m2 total da area.
- cimento = numeric, preço por unidade (saco).
- m2cim = numeric com m2 total da area.
- brita = numeric com m2 total da area.
- m3brita = numeric com m2 total da area.
- data = data frame, opcional, o usuario pode inserir um data frame com todos os argumentos acima.

**Observação:** a metodologia com os parâmetros para calcular a quantidade minima por m2 será de acordo com TISAKA, 2006.<sup>4)</sup>

### Verificando os parâmetros:

- Para os argumentos, faça:
- tijolo, m2tijolo, piso, m2piso, telhado, m2telha, tinta, m2parede, m2teto, areia, m3areia, cimento, m2cim, brita, m3brita é numérico e  $> 0$  ? Se não, escreva: " O argumento[i] precisa ser do tipo numérico e  $> 0$  ."
- typetijolo, typepiso, typetelha, typetinta é string e umas das opções pré determinadas ? Se não, escreva; " O argumento [i] precisa ser do tipo string e umas das opções listadas."
- data é um data frame ? Se não, escreva: " O argumento data precisa ser do tipo data.frame."

### Pseudo-código:

1. Se tiver objeto data. Então, leia data.frame com os argumentos para cada coluna
2. Se não, leia os argumentos e cria objeto X com os dados.
3. Se, existe argumento tijolo. Então, mostre na tela as opções para escolher o tipo de tijolo (usar função switch para criar lista de opções).
4. Leia a opção selecionada e atribua no objeto typetijolo [i]
5. Se, existe argumento piso. Então, mostre na tela as opções para escolher o tipo de piso.
6. Leia a opção selecionada e atribua no objeto typepiso [i]
7. Se, existe argumento telha. Então, mostre na tela as opções para escolher o tipo de telha.
8. Leia a opção selecionada e atribua no objeto typetelha [i]
9. Se, existe argumento tinta. Então, mostre na tela as opções para escolher o tipo de tinta.
10. Leia a opção selecionada e atribua no objeto typetinta [i]
11. Para todos os argumentos de entrada do input, faça.
12. Calcule a quantidade necessária por m2.
13. Calcule a quantidade necessária por m2 e por tipo de (tijolo, piso, telha, tinta).
14. Calcule o preço por quantidade necessária
15. Calcule qual é o tipo de (tijolo, piso, telha, tinta) com o menor até o maior custo e a quantidade total.
16. Calcule o gasto total para cada uma das opções
17. Fim para
18. Retorne um data.frame com os itens dos argumentos e os respectivos valores calculados. E qual a melhor escolha com os menores preços até os maiores valores de acordo com a quantidade de cada item.

### Saída:

Um data frame com os valores e quantidade necessária para cada material de construção por m2.

### Comentários Danilo

Janaína, seu plano A parece super legal e interessante, só recomendo fortemente que vc explique melhor o que a função vai calcular usando bootstrap. Isso vai ser especialmente importante na hora de escrever o help!

Sugiro fazer o plano A mesmo, apesar do texto do plano B ser maravilhoso.

*Danilo*

### Resposta Janaina

Oi Danilo, obrigada pela sugestão 😊

Realmente preciso definir o cálculo usando bootstrap, essa parte ainda está meio obscura 🤔

*Janaina*

## Links para o trabalho final

Decidimos (Eu, Danilo e o acaso 😊) seguir com a proposta A.

Como observado, a parte do bootstrap não estava clara na função. E como nos estudos com meta-análise, geralmente não se tem o acesso aos dados brutos. Optei, por não fazer o bootstrap e calcular o intervalo de confiança a partir de fórmulas estatísticas encontradas na literatura (veja as fórmulas na descrição do plano A). Também, optei por mudar os argumentos da função. Antes, o usuário teria que inserir cada valor manualmente nos argumentos da função. O que dispensaria tempo. Então, deixei como argumento da função um data.frame de entrada que conterá todos os dados necessários para a função

**Portanto, a função calcula a média padronizada de cada estudo e o intervalo de confiança. Cria um gráfico, estilo forest plot e retorna um data.frame ao usuário.**

Além, do que foi acordado, a função cria novas colunas no data.frame com os valores que podem ser utilizados para análise estatística, como: fator de correção (J) , grau de liberdade (df), estatística (Z), variância (var.d), erro padrão (SE.d).

- Função: [metaR](#)
- Precisa de um: [Help ?](#)
- Exemplo de [output](#) da função metaR.

### Limitações da função:

Para que a função execute é necessário que o usuário coloque exatamente os nomes pré estabelecidos para o cabeçalho. O que restringe o usuário a usar apenas esses nomes. Porém, ao desenvolver a função quis ter a certeza que o cálculo realizado na função seria dos valores correspondente a cada coluna. Por isso, restringir o nome para ter certeza que os valores serão inseridos corretamente. Como, quero utilizar esta função pretendo resolver posteriormente essa limitação. Outra limitação, o titulo da legenda do lado esquerdo e direito não ficam alinhados com a descrição da legenda.

### Referências:

- 1) Koricheva, J., Gurevitch, J., Mengersen, K.(2013) Handbook of Meta-analysis in Ecology and Evolution. Princeton: Princeton University Press.
- 2) Curtis, P.S. (1996). A meta-analysis of leaf gas exchange and nitrogen in trees grown under elevated carbon dioxide. Plant, Cell and Environment
- 3)  $\text{desvio padrão (sd)} = \text{erro padrão} * \sqrt{n}$
- 4) Tisaka, M.(2006) Orçamento na construção civil : consultoria, projeto e execução.São Paulo : Editora Pini.

From:

<http://ecor.ib.usp.br/> - **ecoR**

Permanent link:

[http://ecor.ib.usp.br/doku.php?id=05\\_curso\\_antigo:r2018:alunos:trabalho\\_final:psjana:start](http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2018:alunos:trabalho_final:psjana:start) 

Last update: **2020/08/12 06:04**