



# Gustavo A. Agudelo-Cantero, B.Sc.

Estudante de Mestrado em Fisiologia Geral

Laboratório de Ecofisiologia e Fisiologia Evolutiva (LEFE)

Instituto de Biociências

Universidade de São Paulo

## Meus exercícios

Exercícios 1 - Introdução ao R [exercicio1\\_ga.txt](#)

Exercícios 2 - Funções Matemáticas [exercicio2\\_ga.txt](#)

Exercícios 3 - Leitura e Manipulação de Dados [exercicio3\\_ga.txt](#)

Exercícios 4 - Análises Exploratórias [exercicio4\\_ga.txt](#)

Exercícios 5 - Criação de Gráficos [exercicio5\\_ga.txt](#)

Exercícios 6 - Teste de Hipótese e Simulação [exercicio6\\_ga.txt](#)

Exercícios 7 - Regressões Lineares Simples e Múltiplas [exercicio7\\_ga.txt](#)

Exercícios 8 - Construção de Funções [exercicios\\_8\\_ga.txt](#)

## Trabalho Final

### Proposta A

#### Tolerância ao Aquecimento

Deutsch et al. (2008) definiram a tolerância ao aquecimento de um animal ectotérmico como “a diferença entre sua temperatura crítica máxima e a temperatura climatológica atual do habitat do organismo” (Warming Tolerance,  $WT = CT_{m\acute{a}x} - Th_{ab}$ ), que indica “a quantidade média de aquecimento ambiental que um ectotermo pode tolerar antes que a performance caia para níveis fatais”. Embora este conceito é útil e envolve medições da temperatura do habitat, o  $CT_{m\acute{a}x}$  utilizado na sua estimativa tem sido medido em taxas de aquecimento agudo (e.g.  $\Delta T = 1^{\circ}C \text{ min}^{-1}$ ) que não refletem as taxas de aquecimento no habitat e, portanto, a tolerância térmica no campo (Ribeiro et al. 2012), o que pode tornar irreal a estimativa do WT. Por isso, a proposta é expor e avaliar um modelo

alternativo onde WT não está baseado em valores absolutos para  $CT_{\text{máx}}$  ou  $Th_{\text{ab}}$ , mas em estes parâmetros numa taxa de aquecimento específica para cada um. Esse modelo será descrito pela função WT no R.

## Proposta B

### Curvas de Performance Termal

A temperatura corporal é talvez a variável ecofisiológica mais importante que afeta profundamente a ecologia dos ectotermos, pela sua influência na fisiologia e comportamento (Huey e Stevenson, 1979; Angilletta et al., 2002). Em primeiro lugar, a temperatura de um organismo é uma medida da energia cinética de seus reações bioquímicas e de seus processos fisiológicos (Lutterschmidt e Hutchinson, 1997; Angilletta, 2009). Em segundo lugar, a capacidade funcional do organismo (performance) e sua aptidão ou fitness também são afetados pela temperatura (Angilletta, 2009; Rezende et al., 2011). A relação entre a performance e a temperatura corporal é descrita por uma função assimétrica que exibe três parâmetros: (1) o ótimo térmico ( $T_o$ ), a temperatura corporal em que a performance é máxima; (2) a amplitude da performance térmica ( $B\%$ ), o rango de temperaturas corporais onde a performance é maior ou igual do que um nível arbitrário, geralmente expressado em percentagem do máximo nível; (3) os limites térmicos, a temperatura corporal mínima ( $CT_{\text{min}}$ ) e máxima ( $CT_{\text{max}}$ ) onde a performance é possível (Huey e Stevenson, 1979; Angilletta et al., 2002; Lutterschmidt e Hutchinson, 1997; Angilletta, 2009). A função  $P_t$  vai construir Curvas de Performance Termal a partir de dados da performance dos organismos (alguma atividade relevante para o fitness, como máxima velocidade de locomoção, taxas de crescimento, etc.) ao longo de um amplo espectro de temperaturas corporais. O modelo de ajuste da curva será efetuado segundo Huey e Stevenson (1979), Angilletta (2006) e Deutsch et al. (2008).

Gustavo, suas propostas estão confusas. Por favor reescreva deixando de lado a fisiologia, pois o importante para nós é o que a função que vc propõe vai fazer, quais tipos de dados ela recebe e qual saída ela gera, assim fica mais fácil de a gente te ajudar. Sua proposta, A, por exemplo, me parece mais uma proposta de pesquisa do que uma função do R. Na sua proposta B a função vai ajustar um modelo não-linear, é isso? Veja na página [Trabalho final](#) um modelo do que sua proposta deve ter.

--- *Vitor*

## Mais Comentários

**Comentários** Concordo com os comentários do Vitor. Fiquei

sem entender o que a função irá executar e ao mesmo tempo com a impressão que é muito específica (em ambas as propostas). Siga os conselhos do Vito e veja algumas dicas de como estruturar a proposta que incluímos no wiki:

[http://ecologia.ib.usp.br/bie5782/doku.php?id=bie5782:01\\_curso\\_atual:alunos:trabalho\\_final:start#a\\_proposta](http://ecologia.ib.usp.br/bie5782/doku.php?id=bie5782:01_curso_atual:alunos:trabalho_final:start#a_proposta)

Aguardamos suas correções! Avise o Vitor quando terminar via email acima para ele ou eu darmos uma outra olhada. —  
*Alexandre Adalardo de Oliveira* 2014/04/25 17:02

## Propostas Ajustadas

### Proposta A: Teste G de bondade de ajuste

A minha função irá realizar um teste G de bondade de ajuste (ou teste da razão de verossimilhança para bondade de ajuste). Os dados a serem tratados serão contagens do número de observações em cada categoria da variável categórica comparadas com as contagens esperadas, que são calculadas utilizando algum tipo de expectativa teórica (como uma proporção de sexos de 1:1 ou uma proporção 1:2:1 um cruzamento genético). O argumento de entrada da função será um dataframe de 3 colunas (Categorias, Números observados, Proporções esperadas) e duas ou mais filas, e ela vai retornar uma lista com o valor do estimador G e o valor p de significância estatística. Além disso, a correção de Yates será feita quando tiver só duas categorias da variável categórica.

### Proposta B: Taxa intrínseca de crescimento populacional (r) a partir de dados demográficos

Esta função vai calcular a taxa intrínseca de crescimento populacional através da construção de uma tabela de fertilidade ou fecundidade. O argumento de entrada da função será um dataframe de três colunas e n filas. Os dados contidos nas colunas serão: x, as classes de idade em que a população foi subdividida; nx, número de indivíduos vivos no começo de cada intervalo de idade; e mx, número de descendentes produzidos por uma fêmea de idade x durante o intervalo de tempo x a x+1. A saída da função será o objeto r contendo o valor da taxa intrínseca de crescimento, mas também terá um objeto dentro da função contendo a tabela de fertilidade completa.

## Comentários das Propostas Reformuladas

**Comentários** Caro Gustavo,

Ambas as reformulações estão mais gerais e são interessantes. O teste de “bondade de ajuste” (gosto mais de “teste de aderência”) é bastante geral e o calculo do G é simples. Para deixar a função mais interessante poderia

calcular as frequências esperadas a partir de um vetor de razões ou proporções e tirar a coluna de entrada de expectativa teórica. Note que o vetor de razão ou proporções deve ter o mesmo tamanho do numero de categorias dos dados, e uma coisa interessante é testar se os tamanhos são realmente iguais e retornar uma mensagem de erro caso não sejam, logo no começo da função.

A proposta 2 parece tb. interessante, mas não entendi se os dados de entrada são referentes a uma coorte. Nesse caso sugiro que além do r a função faça tb. um gráfico da projeção da população para um intervalo de tempo definido pelo usuário e de uma população com um N0 para cada classe de idade definida pelo usuário tb. — [Alexandre Adalardo de Oliveira](#) 2014/05/01 11:39

## Teste G de aderência estatística

### Página de Ajuda: Help

Teste.G package:unknown R Documentation

Teste G de aderência estatística

#### Description:

O Teste G de aderência estatística é usado para comparar contagens de observações de uma variável categórica com contagens esperadas que podem ser calculadas a partir dos dados ou com alguma expectativa teórica ('a priori'). A hipótese nula estatística é que o número de observações em cada categoria é igual ao número esperado pelo modelo, e a hipótese alternativa é que o numero de observações é diferente ao esperado. Se só existem duas categorias da variável, McDonald (2008) recomenda fazer a correção de continuidade do Yates.

#### Usage:

```
Teste.G <- function(x, pe=1)
```

#### Arguments:

x dataframe. Pode conter 2 o 3 colunas. A primeira coluna deve conter as categorias da variável, a segunda o número de observações e, se tem 3 colunas, a terceira deve incluir as proporções esperadas segundo a hipótese extrínseca de expectativa teórica.

pe caráter. Proporções esperadas. Default: 1. Indica que as proporções esperadas serão calculadas a partir de um vetor de razões ou proporções derivado dos dados observados. Neste caso, todas as categorias têm a mesma probabilidade de ocorrer. Se  $pe = 2$ , indica que as proporções esperadas estão contidas na coluna 3 do dataframe (x) e correspondem a alguma expectativa teórica.

#### Details:

Se o número de categorias de observações são somente duas, então a função faz a correção de continuidade do Yates.

Para mais detalhes, visite: <http://udel.edu/~mcdonald/statgtestgof.html>

#### Value:

Teste.G retorna uma lista com os componentes seguintes:

comp1 : Tabela. Contem o dataframe original (x) com as colunas correspondentes a todos os cálculos do teste G.

comp2 : G. Apresenta o valor do estimador G.

comp3 : valor.p. Apresenta o valor p de significância estatística para tomar a decisão com respeito à hipótese nula.

#### Warning:

A função pára se os vetores de proporções esperadas e de categorias dos dados não têm o mesmo tamanho

#### Author(s):

Gustavo A. Agudelo C.  
gustavo.a.agudelo@gmail.com  
gustavo.agudelo@ib.usp.br

#### References:

McDonald, J.H. 2008. Handbook of Biological Statistics. University of Delaware, Maryland. USA  
<http://udel.edu/~mcdonald/statintro.html>

#### Examples:

```
sexo <- c("machos", "fêmeas")  
f.observadas <- c(20, 24)
```

```
x1 <- data.frame(sexo, f.observadas)
Teste.G(x1)

fenótipo <- c("Alta e Roxa", "Alta e Branca", "Baixa e Roxa", "Baixa e Branca")
f.observadas <- c(108, 35, 46, 11)
p.esperadas <- c(9, 3, 3, 1)
x2 <- data.frame(fenótipo, f.observadas, p.esperadas)

Teste.G(x2, pe=2)
```

## Código da função

```
Teste.G <- function(x, pe=1)
{
  if(pe==1) #Se as frequências esperadas serão calculadas a partir de
um vetor de razões ou proporções derivado dos dados observados
  #Neste caso, todas as categorias têm a mesma probabilidade
de ocorrer
  {
    p.esperada <- 1/(length(x[,2])) #Divide 1, que es a máxima
probabilidade, entre o número de categorias de observações
    x$p.esperadas <- rep(p.esperada, times=length(x[,2])) #Repete a
probabilidade de ocorrência tantas vezes quanto o número de categorias
#de observações, e cria a
coluna de proporções esperadas dentro da tabela
    x$f.esperadas <- x$p.esperadas*sum(x[,2]) #Cria a coluna de
frequências esperadas multiplicando cada fila da coluna de proporções
esperadas
#por o número total de observações
  }

  if(pe==2) #Se o dataframe contem uma terceira coluna de proporções
esperadas, derivada 'a priori' de alguma expectativa teórica
  {
    x$f.esperadas <- ((x[,3])/(sum(x[,3])))*(sum(x[,2])) #Cria a
coluna de frequências esperadas dividindo cada fila da coluna de proporções
#esperadas entre o valor
total da mesma coluna e multiplicando depois pelo número
#total de observações
  }

  if(length(x[,3])!=length(x[,1])) stop("O vetor de proporções não tem
o mesmo tamanho do número de categoria dos dados") #A função pára
#se os vetores de proporções esperadas e de categorias dos dados não
têm o mesmo tamanho

  if(length(x[,1])==2) #Se o número de categorias de observações são
somente duas, então a função faz a correção de continuidade do Yates
```

```
{
  if(x[1,2]>x[1,4]) #Se a primeira frequência observada é maior do
que a primeira frequência esperada
  {
    x$yates.correction[1] <- x[1,2] - 0.5 #Subtrai-se 0.5 para a
primeira frequência observada e coloca o resultado na primeira
#fila de uma nova coluna chamada
de yates.correction
  }
  if(x[1,2]<=x[1,4]) #Se a primeira frequência observada é menor
ou igual do que a primeira frequência esperada
  {
    x$yates.correction[1] <- x[1,2] + 0.5 #Suma-se 0.5 para a
primeira frequência observada e coloca o resultado na primeira
#fila da coluna yates.correction
  }
  if(x[2,2]>x[2,4]) #Se a segunda frequência observada é maior do
que a segunda frequência esperada
  {
    x$yates.correction[2] <- x[2,2] - 0.5 #Subtrai-se 0.5 para a
segunda frequência observada e coloca o resultado na segunda
#fila da coluna yates.correction
  }
  if(x[2,2]<=x[2,4]) #Se a segunda frequência observada é menor ou
igual do que a segunda frequência esperada
  {
    x$yates.correction[2] <- x[2,2] + 0.5 #Suma-se 0.5 para a
segunda frequência observada e coloca o resultado na segunda
#fila da coluna yates.correction
  }
  x$taxa <- x$yates.correction/x$f.esperadas #Calcula-se a taxa
das frequências para cada categoria. Neste caso, divide-se cada
#frequência corrigida pelo Yates
entre cada frequência esperada, e coloca-se o
#resultado numa nova coluna chamada
de taxa
  x$obs.log.taxa <- (x$yates.correction)*(log(x$taxa)) #Calcula-se
o logaritmo natural (base e) para cada taxa corrigida e multiplica-se
#por cada frequência
corrigida pelo Yates. O resultado é colocado numa nova coluna
#chamada de obs.log.taxa
}
if(length(x[,2])>2) #Se o número de categorias de observações são
mais do que duas
{
  x$taxa <- x[,2]/x[,4] #Calcula-se a taxa das frequências para
cada categoria. Neste caso, divide-se cada frequência observada
#entre cada frequência esperada, e coloca-se o resultado
#numa nova coluna chamada de taxa
  x$obs.log.taxa <- (x[,2])*(log(x$taxa)) #Calcula-se o logaritmo
natural (base e) para cada taxa e multiplica-se por cada frequência
```

```
                                #observada. O resultado é colocado
numa nova coluna chamada de obs.log.taxa
    }

    G <- (2)*(sum(x$obs.log.taxa)) #Calcula-se o valor do estimador G a
partir da somatória da coluna obs.log.taxa multiplicada por 2
    df <- (length(x[,1]))-1 #Calcula-se o valor dos graus de liberdade
(df = n-1, onde n = número de categorias da variável)
    valor.p <- pchisq(G,df, lower.tail=F) #Calcula-se o valor p de
significância estatística comparando o valor do estimador G com graus de
liberdade

                                #df numa distribuição qui-quadrado

    resultado <- list(Tabela=x, G=G, valor.p=valor.p) #Cria uma lista
com a tabela completa de todos os cálculos, o valor do estimador G e o valor
p

                                #de significância estatística do
teste
    return(resultado)
}
```

## Script: Teste G

[teste.g\\_ga.txt](#)

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

Permanent link:  
[http://ecor.ib.usp.br/doku.php?id=05\\_curso\\_antigo:r2014:alunos:trabalho\\_final:gustavo.agudelo:start](http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2014:alunos:trabalho_final:gustavo.agudelo:start)

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

