


Introdução ao uso do



em Ciências da Saúde

2. Noções básicas para o trabalho em R
b. Números aleatórios e cálculos estatísticos

Fúlvio Borges Nedel

Departamento de Saúde Pública – SPB

Centro de Ciências da Saúde – CCS

Universidade Federal de Santa Catarina – UFSC

Grups de Recerca d'Amèrica i Àfrica Llatines – GRAAL
<http://graal.uab.cat>

19 de dezembro de 2017

Introdução ao R
2. Noções básicas (b)
2/7

Fúlvio Nedel
SPB/UFSC

Exemplos

Exercício

1 Exemplos

2 Exercício

Introdução ao R

2. Noções básicas (b)

3/7

Fúlvio Nedel
SPB/UFSC

Exemplos

Exercício

Às vezes precisamos gerar números aleatórios

- Seleção de observações para uma amostra
- Exemplos e exercícios de estatística
- Palpite para jogar na loteria :D

Usaremos uma série de funções como exemplo para:

- 1 Selecionar seis números para apostar na "Mega-Sena"
- 2 Definir um vetor com 30 observações de uma distribuição normal e calcular seu desvio-padrão
- 3 Criar uma função personalizada

Lembre que para saber mais sobre cada função usada, basta digitar ?<nome da função>

Introdução ao R
2. Noções básicas (b)
4/7

Fúlvio Nedel
SPB/UFSC

Exemplos

Exercício

1. Seis números para a Mega-Sena

```
sort(as.integer(runif(6, 1, 60)))
```

```
[1] 6 13 19 37 45 54
```

```
?sort
```

```
?as.integer
```

```
?runif
```

2. Calcular o desvio-padrão de uma amostra: $s = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$

```
head(obs <- round(rnorm(n=30, mean=20, sd=5), 1))  
[1] 24.8 13.5 14.9 25.5 24.1 22.8  
(m <- mean(obs)) # calcula a média da distribuição  
[1] 19.83333  
head(d <- obs - m, 4) # calcula a distância das observações à média  
[1] 4.966667 -6.333333 -4.933333 5.666667  
head(qd <- d^2, 4) # eleva a distância ao quadrado  
[1] 24.66778 40.11111 24.33778 32.11111  
(sqd <- sum(qd)) # soma o quadrado das distâncias  
[1] 977.2867  
(n <- length(obs)) # conta o número de observações  
[1] 30  
(gl <- n-1) # graus de liberdade na amostra  
[1] 29
```

s

```
sqrt(sqd/gl)  
[1] 5.805131
```

s, calculado em uma única linha

```
sqrt(sum((obs-mean(obs))^2)/(length(obs)-1))  
[1] 5.805131
```

3. Criar uma função



Introdução ao R
2. Noções básicas (b)
6/7

Fúlvio Nedel
SPB/UFSC

Exemplos

Exercício

O usuário pode criar funções para automatizar cálculos e outras ações, segundo sua conveniência

```
dp <- function(x){  
  # Calcula o desvio-padrão  
  x <- na.omit(x)  
  m <- mean(x)  
  gl <- length(x)-1  
  sqrt(sum((x-m)^2)/gl)  
}  
dp(obs)  
[1] 5.805131
```

Salvar a função em arquivo

```
save(dp, file="funcao_dp.r")
```

Para ativar a função, em futuras sessões de trabalho neste diretório, digite:

```
load('funcao_dp.r')
```

3. Criar uma função



Introdução ao R
2. Noções básicas (b)
6/7

Fúlvio Nedel
SPB/UFSC

Exemplos

Exercício

O usuário pode criar funções para automatizar cálculos e outras ações, segundo sua conveniência

```
dp <- function(x){  
  # Calcula o desvio-padrão  
  x <- na.omit(x)  
  m <- mean(x)  
  gl <- length(x)-1  
  sqrt(sum((x-m)^2)/gl)  
}  
dp(obs)  
[1] 5.805131
```

Salvar a função em arquivo

```
save(dp, file="funcao_dp.r")
```

Para ativar a função, em futuras sessões de trabalho neste diretório, digite:

```
load('funcao_dp.r')
```

O R tem uma função para calcular o DP.

A função sd

```
sd(obs)  
[1] 5.805131
```

3. Criar uma função



Introdução ao R
2. Noções básicas (b)
6/7

Fúlvio Nedel
SPB/UFSC

Exemplos

Exercício

O usuário pode criar funções para automatizar cálculos e outras ações, segundo sua conveniência

```
dp <- function(x){  
  # Calcula o desvio-padrão  
  x <- na.omit(x)  
  m <- mean(x)  
  gl <- length(x)-1  
  sqrt(sum((x-m)^2)/gl)  
}  
dp(obs)  
[1] 5.805131
```

Salvar a função em arquivo

```
save(dp, file="funcao_dp.r")
```

Para ativar a função, em futuras sessões de trabalho neste diretório, digite:

```
load('funcao_dp.r')
```

O R tem uma função para calcular o DP.

A função sd

```
sd(obs)  
[1] 5.805131
```

Se um dos valores fosse missing

```
obs[3] <- NA  
head(obs, 4)  
[1] 24.8 13.5 NA 25.5  
dp(obs) # a nossa  
[1] 5.831288  
sd(obs) # a do R  
[1] NA  
sd(obs, na.rm = T)  
[1] 5.831288
```


Noções básicas

■ Funções úteis:

```
> demo ( )  
?  
?demo  
> demo(graphics)  
> demo(lm.glm)
```

Exercício

Supondo uma população infinita com distribuição normal de peso (média=74 Kg, DP=14,6 Kg) e altura (média=1,6 m, DP=9 cm),

- 1) No RStudio, crie um novo arquivo de sintaxe ("R Script", extensão ".R") e nele escreva os comandos para:
- 2) Criar um vetor com 10 valores aleatórios* com um decimal, para representar uma amostra de medidas de peso;
- 3) Criar um vetor com 10 valores aleatórios* com duas casas decimais, para representar uma amostra de medidas de altura;
- 4) Criar um vetor com o resultado do cálculo do Índice de Massa Corporal (IMC, Kg/m^2);
- 5) Combinar os vetores criados em um banco de dados (*data frame*), com cada vetor em uma coluna e cada registro em uma linha;
- 6) Nomear as variáveis: peso, altura, imc (se ainda não o fez);
- 7) Calcular a média e o desvio-padrão do IMC; e se um valor extremo fosse missing?

Exercício: funções utilizadas

* Para gerar os números aleatórios, utilize a função `rnorm()` com os parâmetros definidos de peso e altura. Para ajuda, digite `?rnorm` no console.

O arredondamento pode ser feito de forma automática com a função `round()`.

Use também a função `set.seed()`, para permitir a reprodutibilidade dos resultados. `set.seed()` deve ser usada antes de `rnorm()`.

`set.seed(<seu nº de ordem na lista de inscrição no curso>)`

* As funções podem ser combinadas, como em

`x <- t(cbind(a,b,c,...))`

Exercício – exemplo

Por exemplo, com a “semente” (v. ?set . seed) 52, o banco de dados resultado do exercício é o seguinte:

```
> banco
  peso altura   imc
1  59.6   1.61 22.99294
2  68.2   1.52 29.51870
3  93.1   1.62 35.47478
4  87.9   1.55 36.58689
5  42.9   1.71 14.67118
6  89.6   1.56 36.81788
7  54.9   1.68 19.45153
8  82.0   1.77 26.17383
9  59.2   1.56 24.32610
10 85.4   1.59 33.78031
```

IMC: média = 28,0 Kg/m²; sd = 7,72 Kg/m²

```
> banco
  peso altura   imc
1  59.6   1.61 22.99294
2  68.2   1.52 29.51870
3  93.1   1.62 35.47478
4  87.9   1.55 36.58689
5  42.9   1.71 NA
6  89.6   1.56 36.81788
7  54.9   1.68 19.45153
8  82.0   1.77 26.17383
9  59.2   1.56 24.32610
10 85.4   1.59 33.78031
```

IMC: média = 29,5 Kg/m²; sd = 6,51 Kg/m²