

Apostila sobre gráficos

DESENVOLVENDO ATIVIDADES DE ENSINO COM O AUXÍLIO DO SOFTWARE R

Título: Treinando habilidades de elaboração de gráficos
com o software R

Aluna: Paola da Silva Martins

Orientadora: Luciane Ferreira Alcoforado

Apresentação

Este trabalho é o resultado da experiência e aprendizado durante o período em que a monitora atuou junto à disciplina de Métodos Computacionais I. Foi elaborado com o objetivo de oferecer material de apoio tanto para a oficina de gráficos realizada durante a agenda acadêmica 2011 como para a disciplina em questão.

Pode ser útil a todos aqueles que desejam aprender a construir gráficos adequados utilizando o programa R de forma prática e rápida. Supõe-se alguma familiaridade no R no que diz respeito à criação de objetos do tipo vetor, matriz e tabela de dados.

O material está organizado de acordo com os diversos tipos de gráficos que são abordados nesta apostila. Procurou-se estabelecer uma linha mestra quanto à escolha do gráfico mais adequado à variável de interesse. Organizou-se uma seção para cada tipo de gráfico onde é explicado como inserir cada elemento necessário a um gráfico informativo como título, legenda, nomes nos eixos, escala dos eixos, aplicação de cores, tamanho de fonte, dentro outros objetos gráficos. Ao final de cada seção há uma série de exercícios permitindo ao leitor a fixação do conteúdo estudado.

A sequência de comandos aqui descrita pode ser adaptada para situações semelhantes em que seja necessária a construção de um dos gráficos apresentados. Espera-se que após o treinamento fornecido por esta apostila o leitor possa construir, com todos os elementos necessários, um gráfico adequado.

Sumário

1.	Classificação das Variáveis	1
1.1	Variáveis Qualitativas:	1
1.2	Variáveis Quantitativas:.....	1
2	Os gráficos	2
2.1	Gráfico de barras:	2
2.1.1	Exercícios.....	10
2.2	Gráfico de setor.....	12
2.2.1	Exercícios:.....	16
2.3	Histograma.....	17
2.3.1	Exercícios:.....	21
2.4	Boxplot.....	22
2.4.1	Exercícios.....	26
2.5	Gráfico de dispersão.....	27
2.5.1	Exercícios.....	30
2.6	Gráfico de linhas.....	32
2.6.1	Exercícios.....	35
2.7	Gráfico da distribuição normal.....	37
2.7.1	Exercícios.....	41
3	Referência Bibliográfica	42
4	Resposta dos exercícios.....	43

LISTA DE FIGURAS E TABELAS

FIGURA 2.1: GRÁFICO DE BARRAS ILUSTRANDO O NÚMERO DE RECÉM NASCIDOS EM TRÊS HOSPITAIS.	2
TABELA 2.1: DADOS DOS PROFESSORES E ALUNOS MATRICULADOS EM ESCOLAS DO MUNICÍPIO DE NITERÓI EM 2009.	2
FIGURA 2.2: GRÁFICO DE BARRAS RESULTADO DO COMANDO BARPLOT SEM TÍTULO	3
FIGURA 2.3: GRÁFICO DE BARRAS COM INCLUSÃO DE TÍTULO.....	4
FIGURA 2.4: GRÁFICO DE BARRAS COM TÍTULO E NOME NOS EIXOS X E Y.	5
FIGURA 2.5: GRÁFICO DE BARRAS COM TÍTULO, NOME NOS EIXOS E SUBTÍTULO	6
FIGURA 2.6: GRÁFICO DE BARRAS NA POSIÇÃO HORIZONTAL	7
FIGURA 2.7: GRÁFICO DE BARRAS DE DUAS SÉRIES DE DADOS.....	8
FIGURA 2.8: GRÁFICO DE BARRAS DE DUAS SÉRIES COM BARRAS LADO A LADO	9
FIGURA 2.9: GRÁFICO DE BARRAS DE DUAS SÉRIES COM BARRAS LADO A LADO, TÍTULO E LEGENDA	10
FIGURA 2.10: GRÁFICO DE SETORES ILUSTRANDO PREFERÊNCIA POR MODALIDADES ESPORTIVAS.....	12
TABELA 2.2: FROTA DE PRINCIPAIS TIPOS DE VEÍCULO DO MUNICÍPIO DE NITERÓI EM 2009.	12
FIGURA 2.11: GRÁFICO DE SETORES PARA OS DADOS DA FROTA DE VEÍCULOS DE NITERÓI EM 2009 SEM ACABAMENTO.	13
FIGURA 2.12: GRÁFICO DE SETORES PARA OS DADOS DA FROTA DE VEÍCULOS DE NITERÓI EM 2009 COM TÍTULO E LEGENDA.....	14
FIGURA 2.13: GRÁFICO DE SETORES PARA OS DADOS DA FROTA DE VEÍCULOS DE NITERÓI EM 2009 COM TÍTULO E RÓTULOS DE PORCENTAGEM.....	15
FIGURA 2.14: HISTOGRAMA ILUSTRATIVO.....	17
FIGURA 2.15: HISTOGRAMA DO CONSUMO DE MILHAS POR GALÃO	18
FIGURA 2.16: HISTOGRAMA DA FREQUENCIA RELATIVA DO CONSUMO DE MILHAS POR GALÃO.....	18
FIGURA 2.17: HISTOGRAMA COM LINHAS DE SOMBREAMENTO	19
FIGURA 2.18: HISTOGRAMA COM TÍTULO PRINCIPAL E NOS EIXOS.	20
FIGURA 2.19: BOX PLOT ILUSTRATIVO.....	22
FIGURA 2.20: BOX PLOT SIMPLES	23
FIGURA 2.21: BOX PLOT COM TÍTULO	23
FIGURA 2.22: BOX PLOT PARA VÁRIOS GRUPOS	24
FIGURA 2.23: BOX PLOT HORIZONTAL.....	25
FIGURA 2.24: DIAGRAMA DE DISPERSÃO	27
FIGURA 2.25: DIAGRAMA DE DISPERSÃO DOS DADOS DO PÉ DE LARANJEIRA.....	28
FIGURA 2.26: DIAGRAMA DE DISPERSÃO E A RETA DE REGRESSÃO	28
FIGURA 2.27: DIAGRAMA DE DISPERSÃO LADO A LADO.....	30
FIGURA 2.28: GRÁFICO DE DUAS FUNÇÕES	32
TABELA 2.3: TEMPERATURA MÉDIA MENSAL DE ÍTHACA-NY	33
FIGURA 2.29: GRÁFICO DE LINHA DA SÉRIE DE TEMPERATURA	33
FIGURA 2.30: GRÁFICO DE LINHA E PONTOS DA SÉRIE DE TEMPERATURA	34
FIGURA 2.31: GRÁFICO DE LINHA E PONTOS DE DUAS SÉRIES NO MESMO PERÍODO DO TEMPO	35
FIGURA 2.32: GRÁFICO DA DISTRIBUIÇÃO NORMAL DE MÉDIA 0 E DESVIO PADRÃO 1.....	37
FIGURA 2.33: HISTOGRAMA DE UMA VARIÁVEL COM DISTRIBUIÇÃO NORMAL.....	38
FIGURA 2.34: GRÁFICO DE HISTOGRAMA COM A CURVA DA NORMAL	38
FIGURA 2.35: GRÁFICOS DE DUAS NORMAIS COM MÉDIAS DIFERENTES	39
FIGURA 2.36: GRÁFICO DE DUAS NORMAIS COM MÉDIAS DIFERENTES E LEGENDA	40
FIGURA 2.37: GRÁFICO DE DUAS NORMAIS COM VARIÂNCIAS DIFERENTES	41

1. Classificação das Variáveis

1.1 Variáveis Qualitativas:

Características que dizem respeito a qualidade ou atributos. Se as categorias da variável possuem uma ordem natural, ela é dita ser **qualitativa ordinal**; caso contrário, **qualitativa nominal**.

Exemplos:

- **Variáveis Qualitativas Nominais:**

Sexo: Feminino, Masculino

Cor de olhos: Pretos, Castanhos, Azuis, Verdes

Tipo de Veículo: Carro, Moto, Ônibus

- **Variáveis Qualitativas Ordinais:**

Estado de Saúde: Ruim, Regular, Bom

Tipo de Acidente: Leve, Moderado, Grave

Cargo na empresa: diretor, vice-presidente, presidente

Quais gráficos podem ser usados?

Setor e barras

1.2 Variáveis Quantitativas:

Características que dizem respeito a quantidades. Se os valores da variável se referem a uma contagem, ela é dita ser **quantitativa discreta**. Caso se refira a uma mensuração (comunidade de medida), ela é dita ser **quantitativa contínua**.

Exemplos:

- **Variáveis Quantitativas Discretas:**

Nº de pessoas na família

Nº de acidentes na BR101 em 2009

Nº de medicações já utilizadas no paciente

- **Variáveis Quantitativas Contínuas:**

Peso (Kg)

Idade (anos)

Duração do efeito da medicação (horas)

Quais gráficos podem ser usados?

Histograma, boxplot, dispersão, linha.

2 Os gráficos

2.1 Gráfico de barras:

O gráfico de barras é composto por duas linhas ou eixos, um vertical e outro horizontal. No eixo vertical são construídas as barras que representam a variação de um fenômeno ou de um processo de acordo com sua intensidade. Essa intensidade é indicada pela altura da barra. No eixo horizontal especifica-se as categorias da variável. As barras devem sempre possuir a mesma largura e a distância entre elas deve ser constante.

Para ilustrar este tipo de gráfico, veja a figura 2.1.

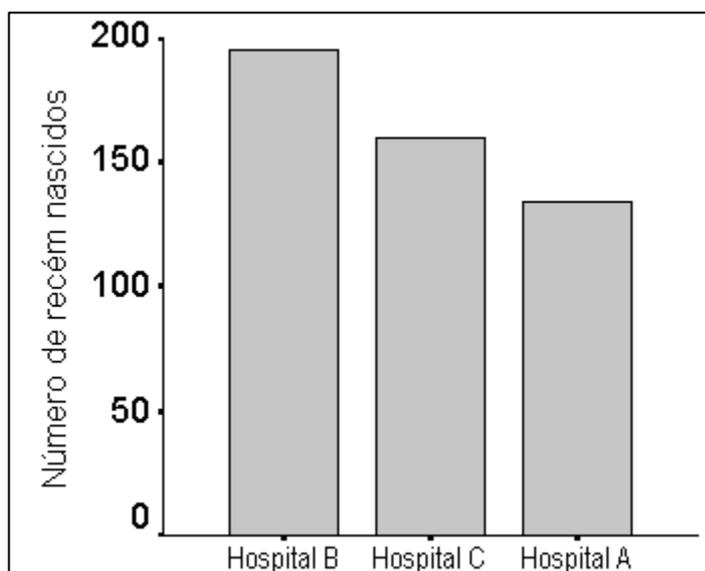


FIGURA 2.1: GRÁFICO DE BARRAS ILUSTRANDO O NÚMERO DE RECÉM NASCIDOS EM TRÊS HOSPITAIS.

Veja a seguir como construir este tipo de gráfico no R. Utilizaremos os dados da tabela 2.1 referente a matrícula de professores.

TABELA 2.1: DADOS DOS PROFESSORES E ALUNOS MATRICULADOS EM ESCOLAS DO MUNICÍPIO DE NITERÓI EM 2009.

Escola de Ensino Fundamental	Matrículas de professores	Matrícula de alunos
Privada	1751	25280
Pública Estadual	1186	21328
Pública Municipal	947	18432
Pública Federal	29	280

Fonte:www.ibge.com.br

O primeiro passo na construção do gráfico é ter os dados armazenados em objeto apropriado. No caso de gráfico de barras é necessário que os dados estejam armazenados em um vetor ou matriz.

```
prof<-c(1751,1186,947,29)
```

Vamos nomear cada barra:

```
escola<-c("privada","estadual","municipal","federal")
```

```
barplot(prof, names.arg=escola)
```

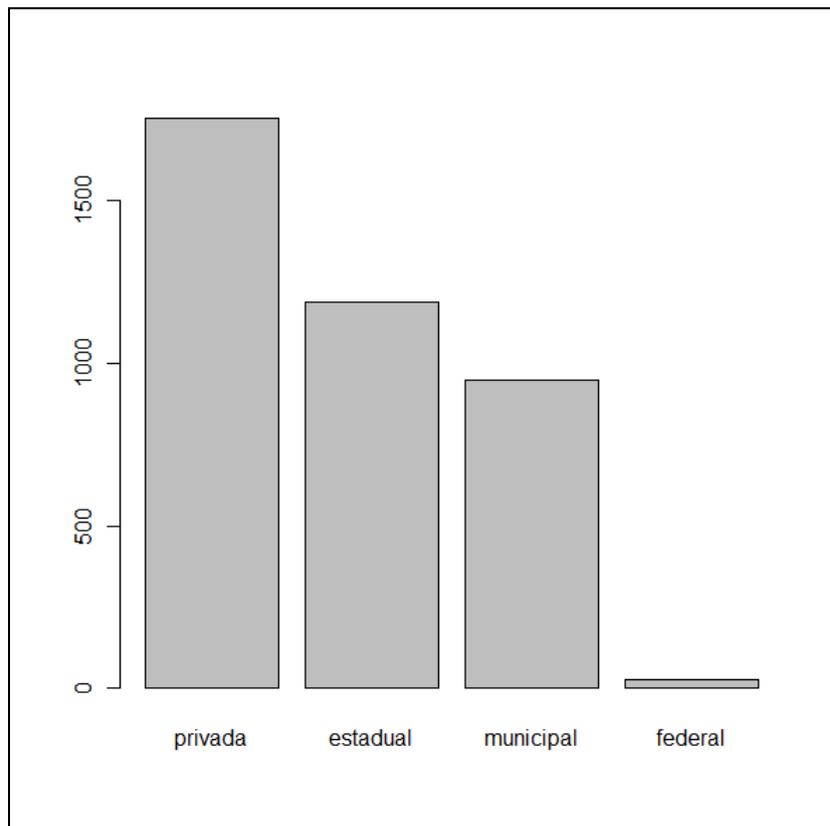


FIGURA 2.2: GRÁFICO DE BARRAS RESULTADO DO COMANDO BARPLOT SEM TÍTULO

Outra possibilidade de gerar o gráfico da figura 2.2 é nomear as posições do vetor prof através do comando names():

```
names(prof)<-c("privada","estadual","municipal","federal")
```

prof #observe que os dados agora possuem um nome em cada posição.

```
privada estadual municipal federal
```

```
1751 1186 947 29
```

Uma vez nomeada as posições do vetor, basta fazer

```
barplot(prof) # e o gráfico é o mesmo da figura 2.2
```

A partir de agora podemos omitir o argumento names.arg da função barplot que obteremos o mesmo resultado.

Como colocar título no gráfico:

```
barplot(prof, names.arg=escola, main="Distribuição de professores do ensino fundamental em Niterói, 2009")
```

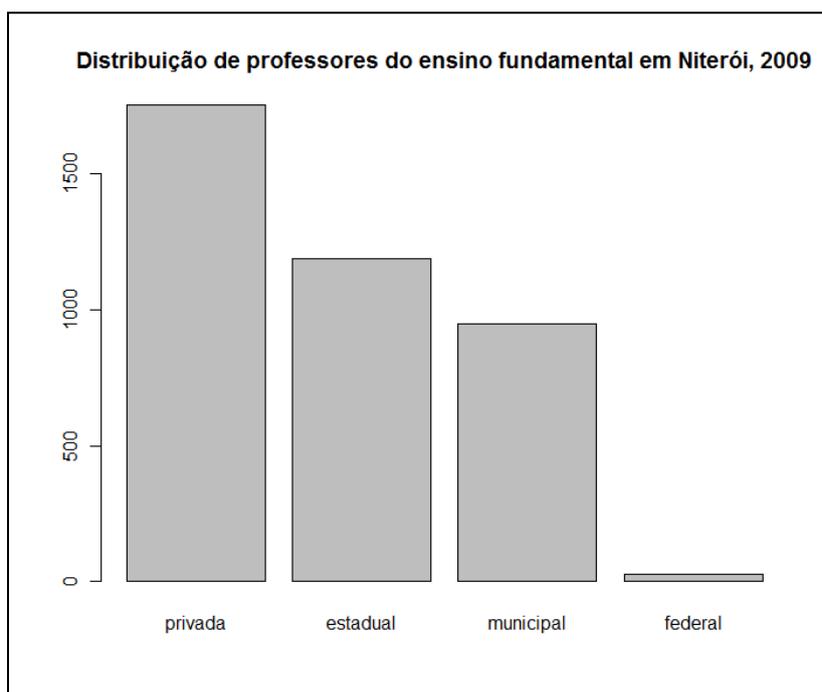


FIGURA 2.3: GRÁFICO DE BARRAS COM INCLUSÃO DE TÍTULO.

Se necessitar mudar o tamanho da fonte do título, acrescente o argumento `cex.main=valor`. No R o valor padrão é 1, para aumentar use valor maior do que 1; para diminuir use valor menor do que 1. Confira o resultado da diminuição/aumento do título reproduzindo os comandos abaixo.

```
barplot(prof,main="Professores de escolas de ensino fundamental em 2009 no município de Niterói",cex.main=0.9)
```

```
barplot(prof, main="Professores de escolas de ensino fundamental em 2009 no município de Niterói",cex.main=1.5)
```

Analogamente, para mudar o tamanho dos números do eixo y utilize o argumento `cex.axis=valor`.

```
barplot(prof,cex.axis=0.9)
```

Colocando nome nos eixos x e y:

```
barplot(prof, names.arg=escola, main="Distribuição de professores do ensino fundamental em Niterói, 2009",ylab="matrículas",xlab="escolas")
```

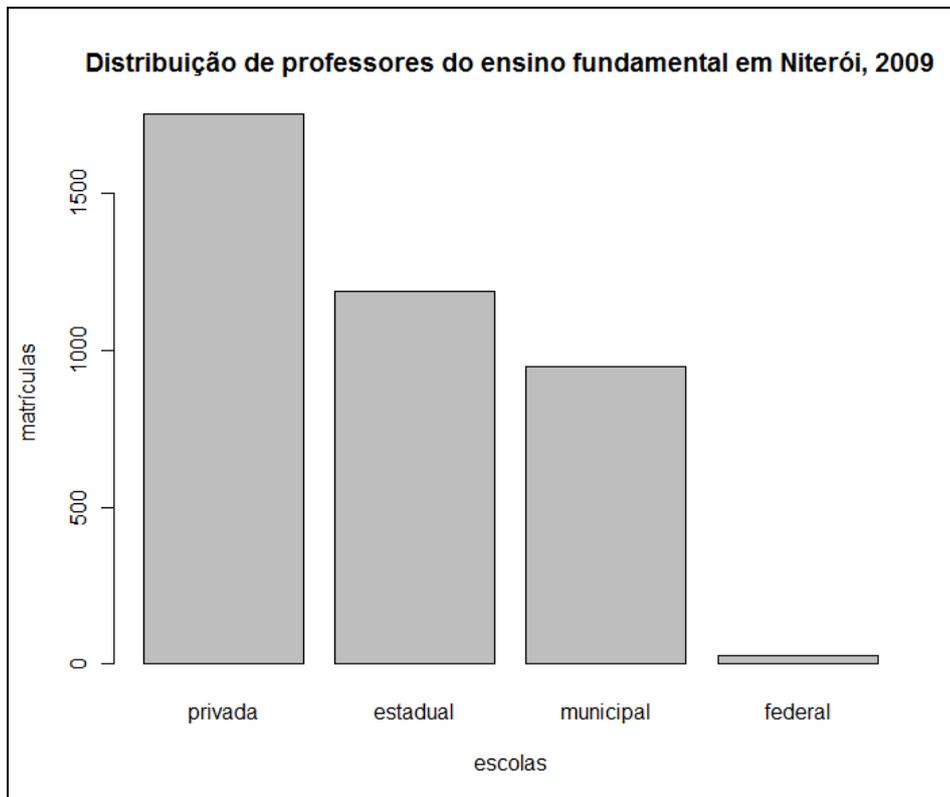


FIGURA 2.4: GRÁFICO DE BARRAS COM TÍTULO E NOME NOS EIXOS X E Y.

Para mudar o tamanho da fonte dos nomes dos eixos, utilize o argumento `cex.lab=valor`.

```
barplot(prof,ylab="matrículas",xlab="escolas",cex.lab=0.9)
```

Usando o `cex.`, podemos mudar o tamanho da fonte de diversas partes do gráfico.

Podemos ainda retirar os eixos do gráfico, mas no nosso caso isso não é interessante:

```
barplot(prof,axes=FALSE)
```

É possível incluir subtítulo no gráfico incluindo o argumento `sub = "texto"`.

```
barplot(prof, names.arg=escola, main="Distribuição de professores do ensino fundamental em Niterói, 2009",ylab="matrículas",xlab="escolas",sub="fonte:www.ibge.com.br")
```



FIGURA 2.5: GRÁFICO DE BARRAS COM TÍTULO, NOME NOS EIXOS E SUBTÍTULO

Para mudar o tamanho da fonte dos nomes dos eixos, utilize o argumento `cex.lab=valor`.

```
barplot(prof,ylab="matrículas",xlab="escolas",cex.lab=0.9)
```

Usando o `cex.`, podemos mudar o tamanho da fonte de diversas partes do gráfico.

Podemos ainda retirar os eixos do gráfico, mas no nosso caso isso não é interessante:

```
barplot(prof,axes=FALSE)
```

É possível incluir subtítulo no gráfico incluindo o argumento `sub = "texto"`.

```
barplot(prof, main="Distribuição de professores do ensino fundamental em Niterói, 2009",ylab="matrículas",xlab="escolas",sub="fonte:www.ibge.com.br")
```

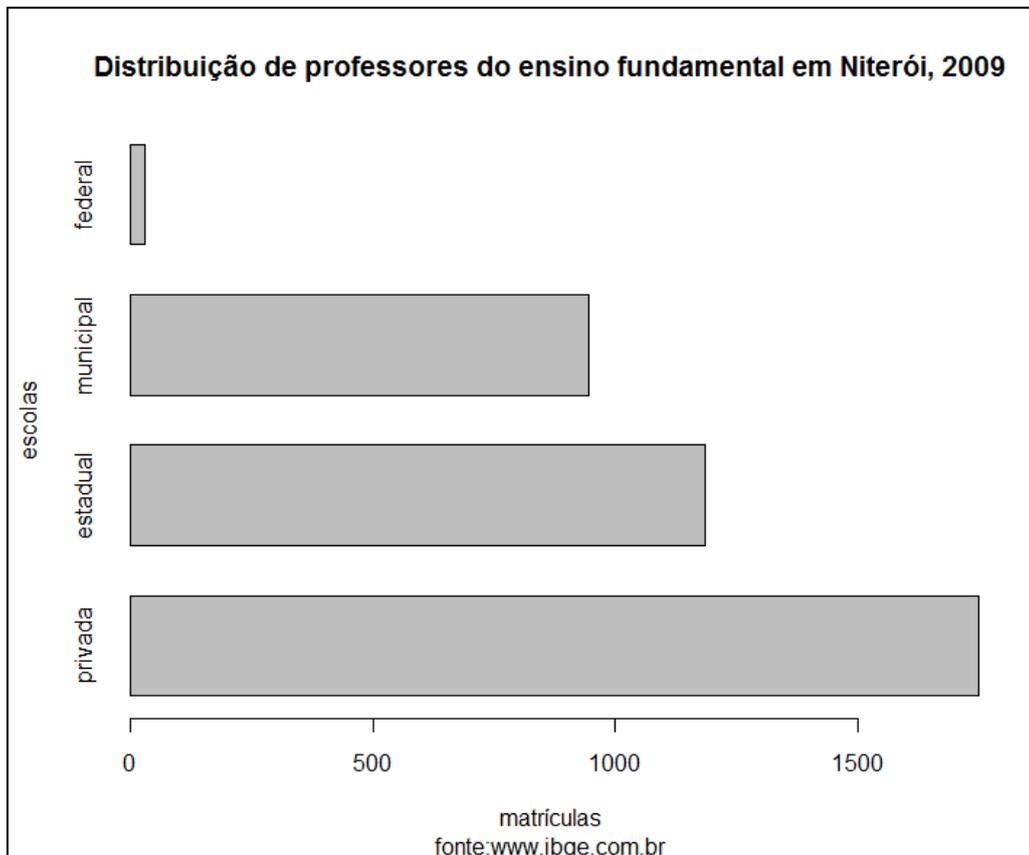


FIGURA 2.6: GRÁFICO DE BARRAS NA POSIÇÃO HORIZONTAL

Podemos preencher as barras com linhas de sombreamento ou com cor uniforme. Os gráficos acima foram preenchidos na cor cinza que é o padrão do R. Para escolher a linha de sombreamento das barras utilizamos `density=valor`, este valor representa a densidade de linhas por polegada e é um número positivo.

```
barplot(prof,density=40)
```

E se ainda quisermos escolher qual a angulação destas linhas de sombreamento utilizamos o `angle=valor`, este valor variando de 0 a 360 no sentido anti horário.

```
barplot(prof,density=40,angle=90)
```

Para preencher as barras com cores uniformes diferentes do padrão estabelecido, usamos o comando `col = "nome da cor"` ou `col = valor`, este valor representa o número da cor.

```
barplot(prof,col=c("blue","red","green","orange"))
```

```
barplot(prof,col=c(1,2,3,4))
```

Podemos excluir a linha das bordas das barras dessa forma:

```
barplot(prof,border=NA)
```

Ou ainda, se for utilizado a linha de sombreamento, a cor da borda será de acordo com a da linha de sombreamento, assim:

```
barplot(prof,density=50,border=TRUE)
```

Também podemos criar o gráfico de barras de duas variáveis, um ao lado do outro, na mesma janela gráfica. Para isso, vamos imaginar que os dados estejam armazenados em um objeto do tipo matriz, conforme comando a seguir. Estes dados são os da tabela 2.1:

```
alunosprof<- matrix(c(1751,1186,947,29,25280,21328,18432,280), nrow=4, ncol=2)
```

```
alunosprof
      [,1] [,2]
[1,] 1751 25280
[2,] 1186 21328
[3,]  947 18432
[4,]   29  280
```

Para colocarmos nome nas linhas e colunas da matriz, utilizamos o argumento `dim.names`.

```
alunosprof<- matrix(c(1751,1186,947,29,25280,21328,18432,280), nrow=4, ncol=2,
dimnames=list(c("privada", "estadual", "municipal", "federal"), c("professores", "alunos")))
```

Agora usamos o comando `barplot` para fazer o gráfico:

```
barplot(alunosprof)
```

Este comando exibirá o seguinte gráfico em janela separada:

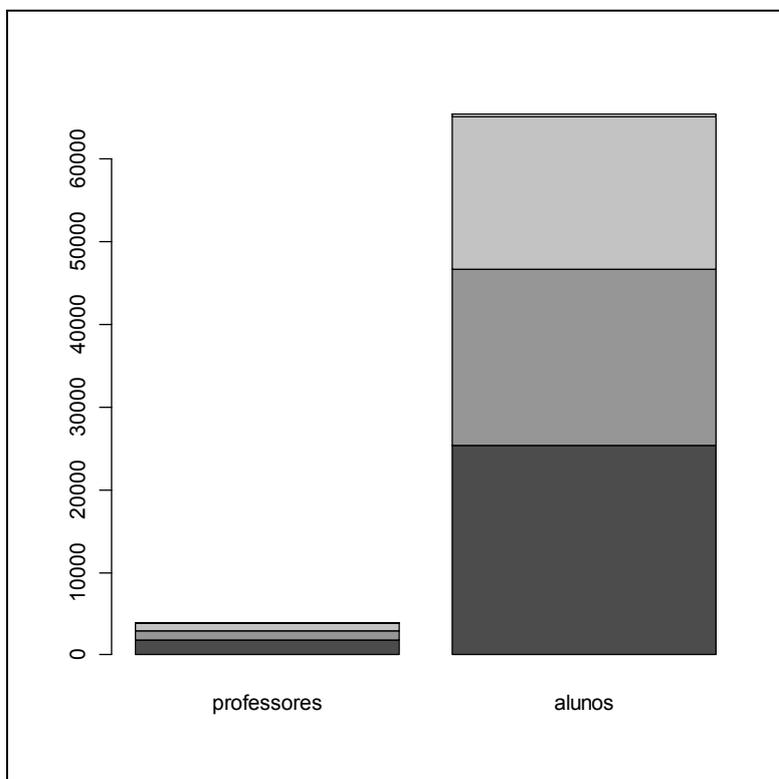


FIGURA 2.7: GRÁFICO DE BARRAS DE DUAS SÉRIES DE DADOS

O default desse comando é colocar todas as categorias uma em cima da outra, acumulando os resultados. Às vezes, essa maneira não é a mais adequada para visualizar o que se deseja, então poderemos usar o seguinte argumento para obtermos barras lado a lado:

```
barplot(alunosprof, beside=TRUE)
```

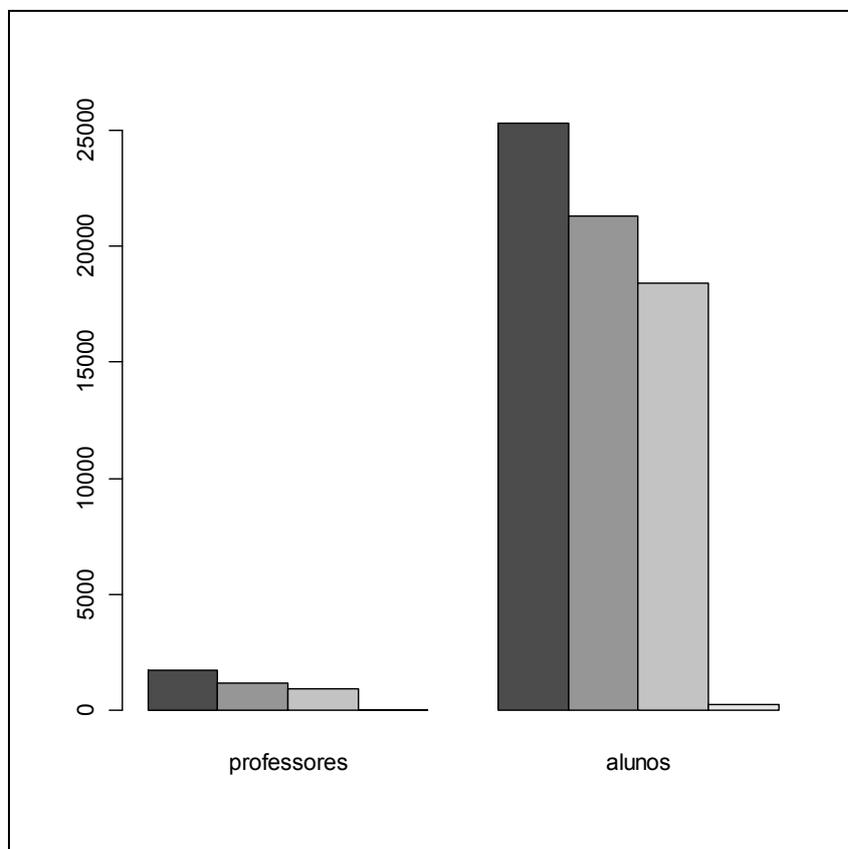


FIGURA 2.8: GRÁFICO DE BARRAS DE DUAS SÉRIES COM BARRAS LADO A LADO

Para finalizarmos a apresentação do gráfico, adicionaremos legenda, título, nomes nos eixos e subtítulo citando a fonte, invertemos a ordem dos dados, colocando primeiro a sequência de alunos e depois a de professores resultando numa disposição mais apresentável do ponto de vista estético.

```
barplot(alunosprof[,2:1],beside=TRUE,legend.text=rownames(alunosprof), main="Distribuição de matrícula de alunos e professores do ensino fundamental em Niterói, 2009",ylab="n. de matrículas",xlab="matrícula",sub="fonte:www.ibge.com.br")
```

O resultado deste comando pode ser visualizado na figura 2.9.

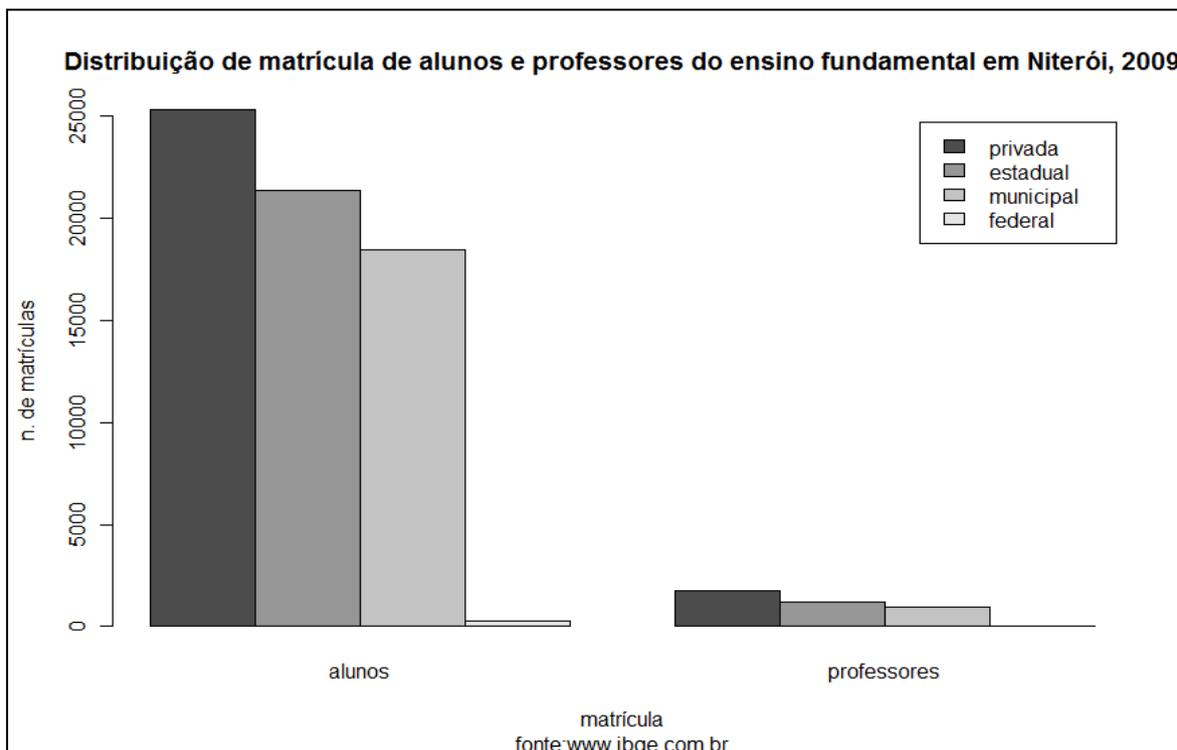


FIGURA 2.9: GRÁFICO DE BARRAS DE DUAS SÉRIES COM BARRAS LADO A LADO, TÍTULO E LEGENDA

2.1.1 Exercícios

1. Com base nos dados da tabela 2.1 construa um gráfico de barras que mostra a relação de alunos/professor em cada tipo de escola. Neste caso será necessário obter os valores da nova variável que é o resultado da coluna 3 dividido pela coluna 2. Coloque título adequado, legenda, e nomes nos eixos.
2. Com base na questão 1, refaça o gráfico de barras com as barras no sentido horizontal.
3. Refaça o gráfico da questão 1 utilizando sombreamento de densidade 60.
4. Refaça o gráfico da questão 1 utilizando sombreamento de densidade 50 e ângulo das linhas de sombreamento de 75 graus.
5. O vetor peso indica o peso médio de pintinhos com 2,4,6,8 e 10 dias de nascido respectivamente.

peso=(42,51,59,64,76).

Considere cada tempo de nascido como sendo uma categoria. Estabeleça um gráfico de barras para o peso médio dos pintinhos em cada categoria. Coloque título, legenda e nomes nos eixos.

6. Refaça o gráfico da questão 5 utilizando o seguinte esquema de cores: "blue", "pink", "yellow", "green", "red".
7. Refaça o gráfico da questão 6 utilizando barras horizontais.
8. Numa central telefônica de uma grande empresa, havia a sensação de saturação do sistema utilizado. Para melhor representar o que ocorria foi realizado um acompanhamento com as telefonistas que teriam que responder aos problemas em que

números ocorriam e lançá-los na **Lista de Verificação**. Tab3 resume os dados desta lista. Apresente um gráfico de barras para estes dados com barras verticais com sombreamento de densidade 80 e ângulo das linhas de 120 graus.

Tab3

Tipo de defeito	n. ocorrência
linha ruidosa	250
linha aberta	110
alarme	85
não responde	45
não toca	25

9. Considere os dados de Tab4 sobre a frequência para um grupo de fumantes de cigarro e um grupo de não-fumantes em diversos níveis séricos de cotinina. Construa um gráfico de barras semelhante ao da figura 2.9

Tab4

Nível de cotinina (ng/ml)	Fumantes	Não-fumantes
0 – 13	78	3300
14 – 49	133	72
50 – 99	142	23
100 – 149	206	15
150 – 199	197	7
200 – 249	220	8
250 – 299	151	9
300 +	412	11

2.2 Gráfico de setor

Os gráficos de setor (ou pizza) são representados por círculos divididos proporcionalmente de acordo com os dados do fenômeno ou do processo a ser representado. Os valores são expressos em números ou em percentuais (%).

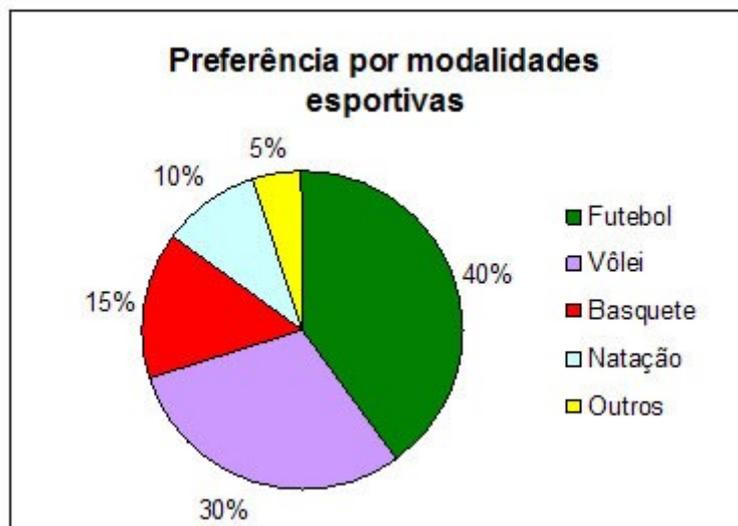


FIGURA 2.10: GRÁFICO DE SETORES ILUSTRANDO PREFERÊNCIA POR MODALIDADES ESPORTIVAS

Os dados da tabela 2.2 , são referentes a frota no município de Niterói no ano de 2009. Utilizaremos estes dados para fazermos um gráfico de pizza.

TABELA 2.2: FROTA DE PRINCIPAIS TIPOS DE VEÍCULO DO MUNICÍPIO DE NITERÓI EM 2009.

Tipo de veículo	Frota
Automóveis	151882
Motocicletas	19627
Caminhonetes	8058
Motonetas	3201
Onibus	1969
Caminhões	1895
Outros	1268

O primeiro passo na construção do gráfico é ter os dados armazenados em objeto apropriado nomeando cada posição do vetor. Os nomes de cada posição aparecerão no gráfico conforme pode ser visto na figura 2.11.

```
frota<-c(151882, 19627, 8058,1895, 3201, 1969,1268)
```

```
names(frota)<-
```

```
c("automóveis","motocicletas","caminhonetes","motonetas","ônibus","caminhões", "outros")
```

```
pie(frota)
```

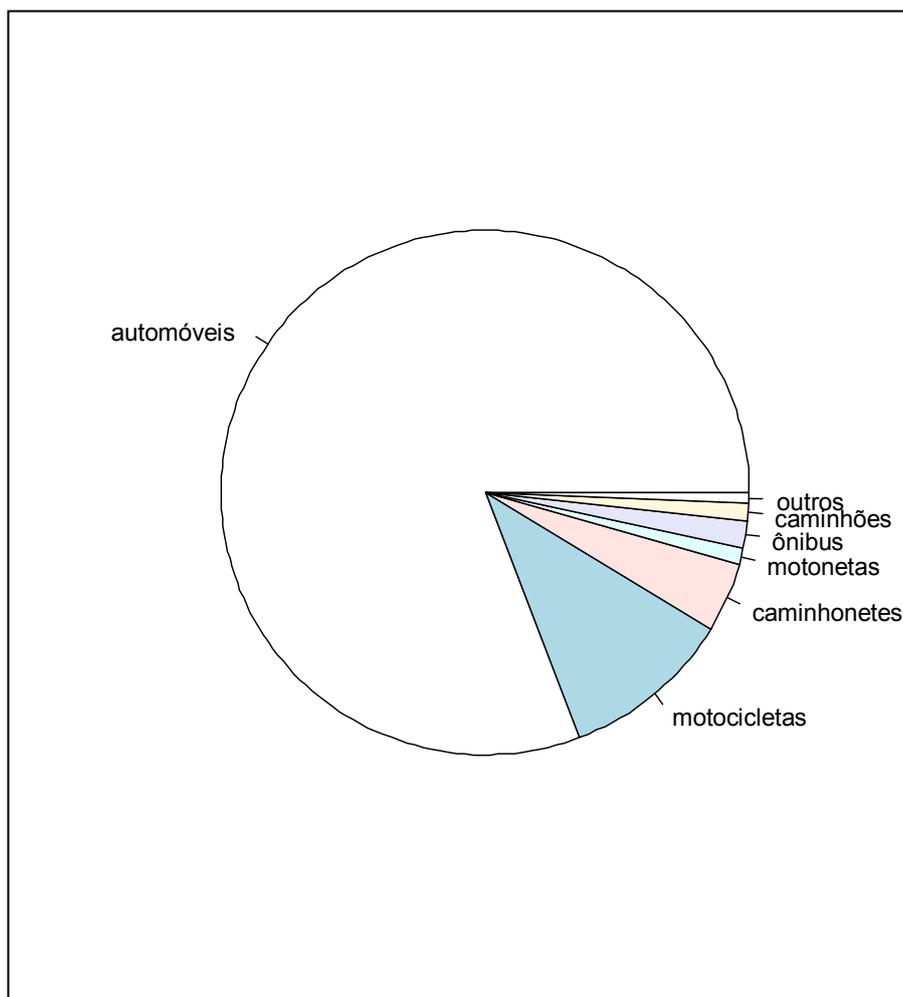


FIGURA 2.11: GRÁFICO DE SETORES PARA OS DADOS DA FROTA DE VEÍCULOS DE NITERÓI EM 2009 SEM ACABAMENTO.

O gráfico da figura 2.11 não é informativo, pois o leitor não sabe do que se trata. Para que se torne compreensível é necessário realizar diversas melhorias como colocação de título, e os percentuais que representa cada setor.

O título pode ser adicionado após a janela gráfica ser exibida com o gráfico da figura 2.11. Neste caso utilizamos o comando **title("texto")** na sequência de comandos do gráfico. Note aqui que o gráfico já está construído, apenas acrescentamos o título.

```
title("Frota 2009 - Niterói_RJ")
```

Para colocar as porcentagens utilizamos o argumento `labels`. Observe que nesta sequência o título é colocado como argumento da função `pie`. Aqui é necessário que se coloque uma legenda para que o leitor compreenda o gráfico.

```
porc<-round(frota*100/sum(frota),2) #arredonda a porcentagem para 2 dígitos significativos)
```

```
rotulos<-paste("(",porc,"%)",sep="")
```

```
pie(frota, main="Frota 2009 - Niterói_RJ",labels=rotulos, col=rainbow(7))
```

```
legend(1,1,names(frota),col = rainbow(7),pch=rep(20,6))
```

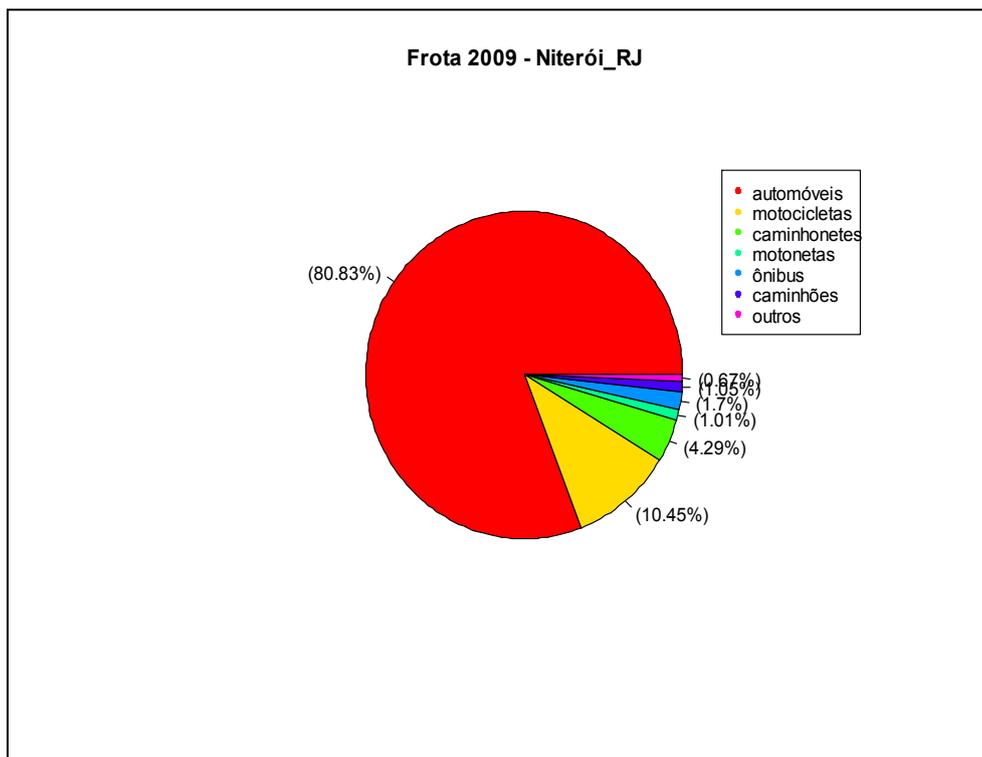


FIGURA 2.12: GRÁFICO DE SETORES PARA OS DADOS DA FROTA DE VEÍCULOS DE NITERÓI EM 2009 COM TÍTULO E LEGENDA.

Ou se desejar, o nome e a porcentagem utilizem a sequencia de comandos a seguir, diminuindo o tamanho da fonte e alterando o esquema de cores para melhor adequação:

```
rotulos<-paste(names(frota),",",porc,"%",sep="")
```

```
pie(frota, main="Frota 2009 - Niterói_RJ",labels=rotulos, cex=0.7, col=rainbow(7))
```

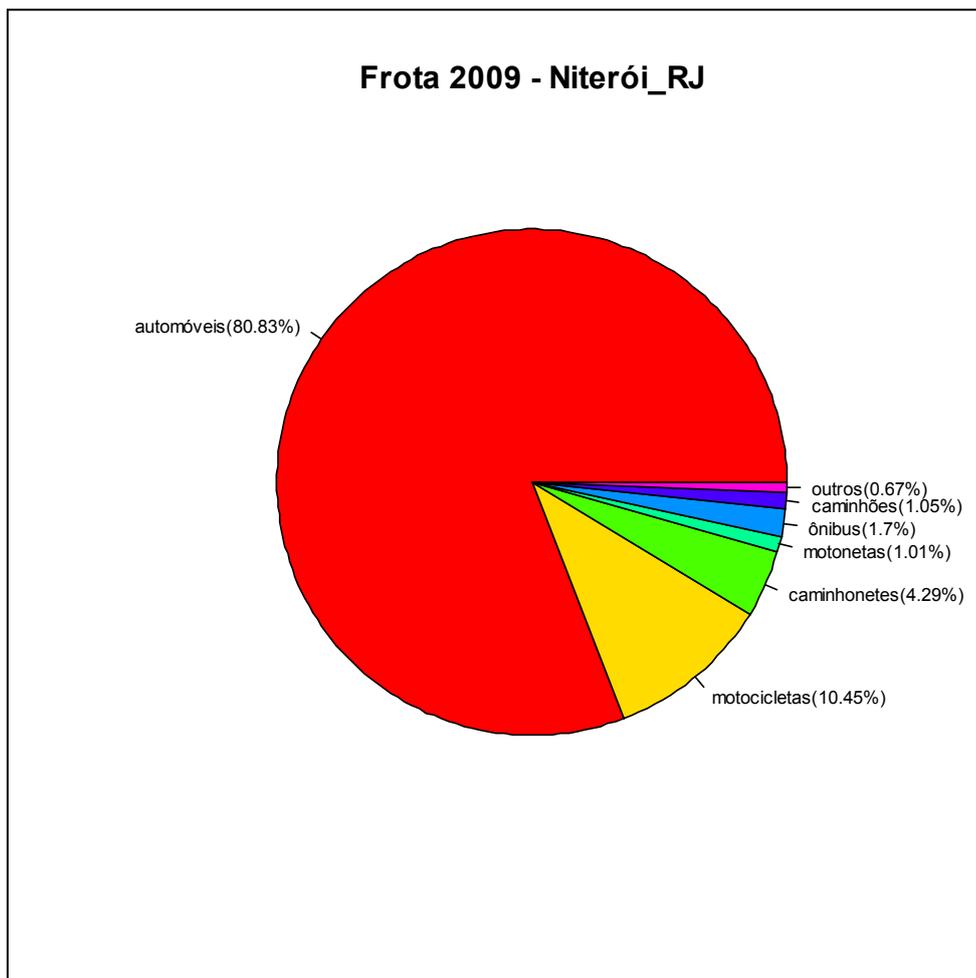


FIGURA 2.13: GRÁFICO DE SETORES PARA OS DADOS DA FROTA DE VEÍCULOS DE NITERÓI EM 2009 COM TÍTULO E RÓTULOS DE PORCENTAGEM

Podemos colocar os nomes manualmente, mas só é bom utilizarmos este recurso no caso de algum nome ficar sobreposto a outro, pois dá um pouco de trabalho, dessa forma:

```
pie(frota, main="Frota 2009 - Niterói_RJ", init.angle=180)
```

```
text(locator(length(names(frota))),rotulos)
```

Após este último comando é necessário que se clique na janela do gráfico onde se deseja que os rótulos sejam impressos. O gráfico só ficará pronto depois de clicar 7 vezes, pois este é o número de setores do gráfico. Pode-se repetir este procedimento até obter o gráfico desejado.

Alguns ajustes podem ser necessários. Para mudar o tamanho da fonte do título use `cex.main=valor`

```
pie(frota, main="Frota 2009 - Niterói_RJ",cex.main=1.2)
```

Para mudar o tamanho da fonte dos nomes das variáveis use `cex=valor`

```
pie(frota,cex=0.9)
```

Para girar o gráfico, use `init.angle=valor`

```
pie(frota, main="Frota 2009 - Niterói_RJ", init.angle=180)
```

2.2.1 Exercícios:

1. Numa central telefônica de uma grande empresa, havia a sensação de saturação do sistema utilizado. Para melhor representar o que ocorria foi realizado um acompanhamento com as telefonistas que teriam que responder aos problemas em que números ocorriam e lançá-los na **Lista de Verificação**. Tab3 resume os dados desta lista. Apresente um gráfico de setores para estes dados semelhante ao gráfico da figura 2.12.

Tab3

Tipo de defeito	n. ocorrência
linha ruidosa	250
linha aberta	110
alarme	85
não responde	45
não toca	25

2. Refaça o gráfico anterior semelhante ao apresentado na figura 2.13
3. Os dados a seguir nos informam as despesas pessoais dos Estados Unidos (em bilhões de dólares), com Comida e cigarro, domésticos, remédios e saúde, cuidados pessoais e educação privada, respectivamente. Faça o gráfico de setor semelhante ao da figura 2.12.

```
despesas=c(22200,10500,3530,1040,341)
```

4. Refaça o gráfico anterior semelhante ao apresentado na figura 2.13.

2.3 Histograma

Na estatística, um histograma é uma representação gráfica da distribuição de frequências de uma massa de medições, normalmente um gráfico de barras verticais. Tal gráfico é composto por retângulos justapostos em que a base de cada um deles corresponde ao intervalo de classe e a sua altura à respectiva frequência.

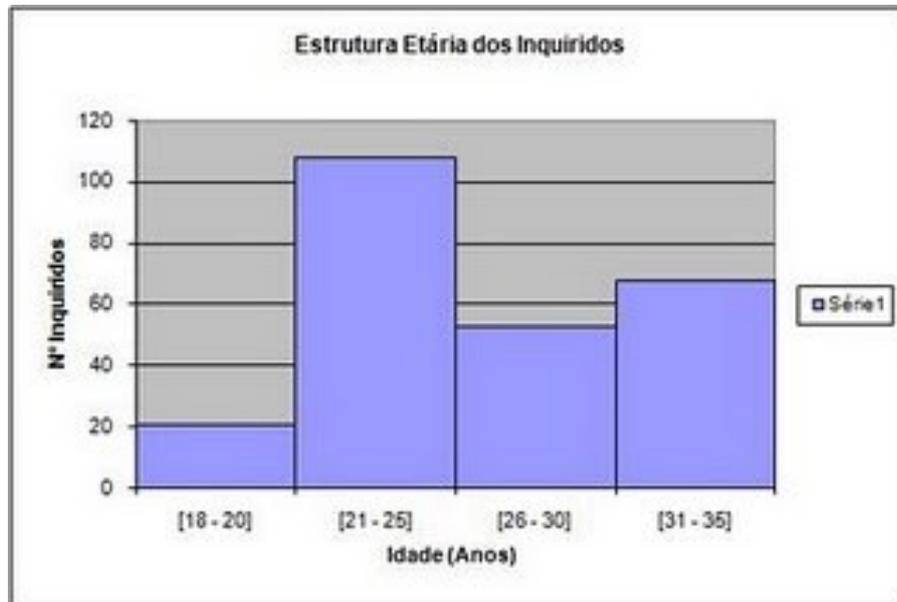


FIGURA 2.14: HISTOGRAMA ILUSTRATIVO

O primeiro passo na construção do gráfico é ter os dados armazenados em objeto apropriado. Por praticidade utilizaremos dados disponíveis no R `data(mtcars)`

```
data(mtcars)
```

```
mtcars[1:6,] #visualizando parte dos dados
```

```
      mpg cyl disp  hp drat   wt  qsec vs am gear carb
Mazda RX4      21.0  6  160 110 3.90 2.620 16.46 0  1   4   4
Mazda RX4 Wag  21.0  6  160 110 3.90 2.875 17.02 0  1   4   4
Datsun 710     22.8  4  108  93 3.85 2.320 18.61 1  1   4   1
Hornet 4 Drive 21.4  6  258 110 3.08 3.215 19.44 1  0   3   1
Hornet Sportabout 18.7  8  360 175 3.15 3.440 17.02 0  0   3   2
Valiant       18.1  6  225 105 2.76 3.460 20.22 1  0   3   1
```

Vamos explorar os dados da variável `mpg` (primeira coluna) que são os dados do consumo de milhas por galão de diversos tipos de carros.

Para fazermos o histograma usamos o comando `hist()`

```
attach(mtcars)
```

```
hist(mpg)
```

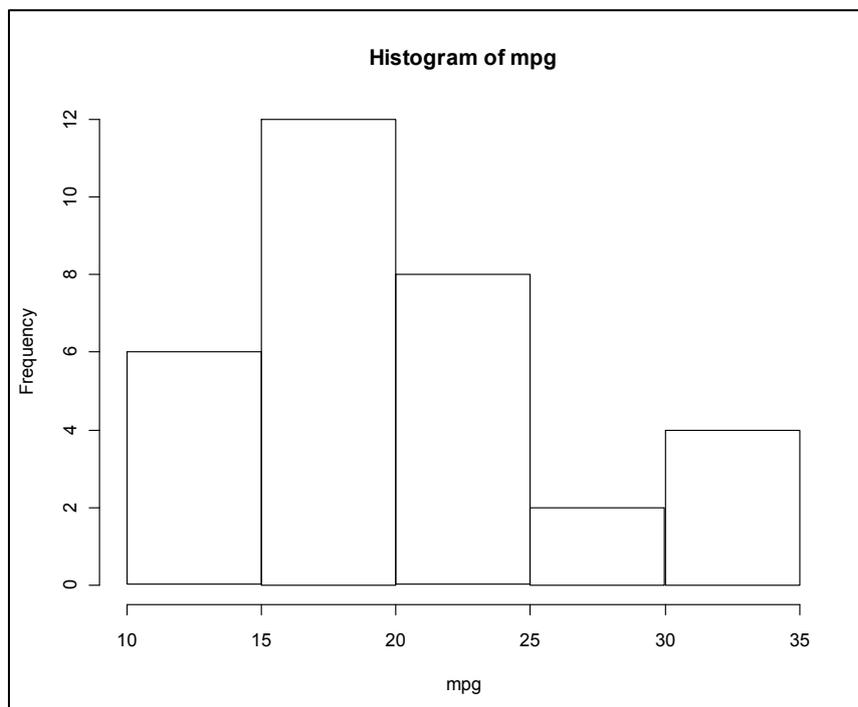


FIGURA 2.15: HISTOGRAMA DO CONSUMO DE MILHAS POR GALÃO

Podemos colocar nosso gráfico em frequência relativa:

```
hist(mpg,freq=FALSE)
```

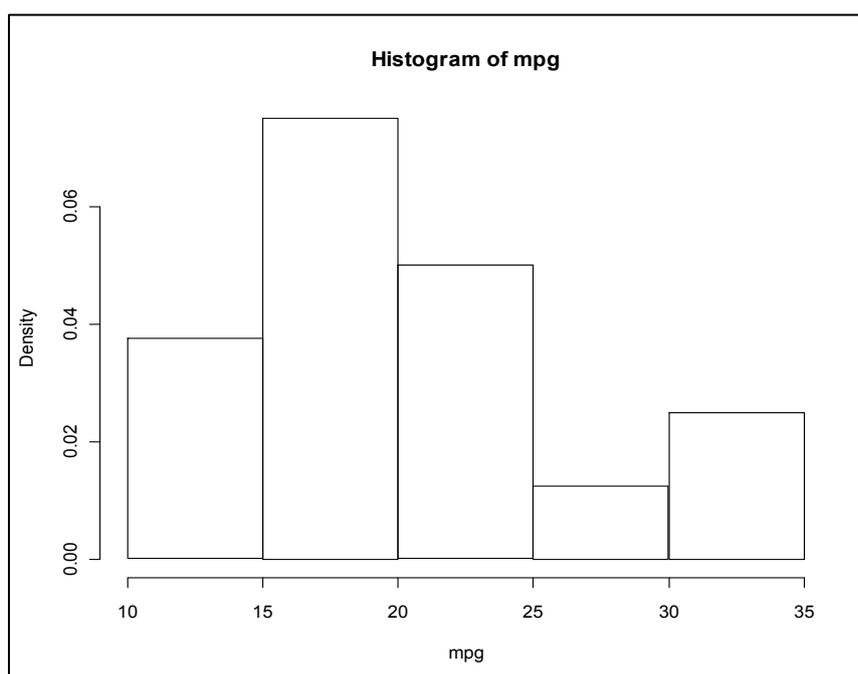


FIGURA 2.16: HISTOGRAMA DA FREQUENCIA RELATIVA DO CONSUMO DE MILHAS POR GALÃO

Colocar linhas de sombreamento:

```
hist(mpg,density=30)
```

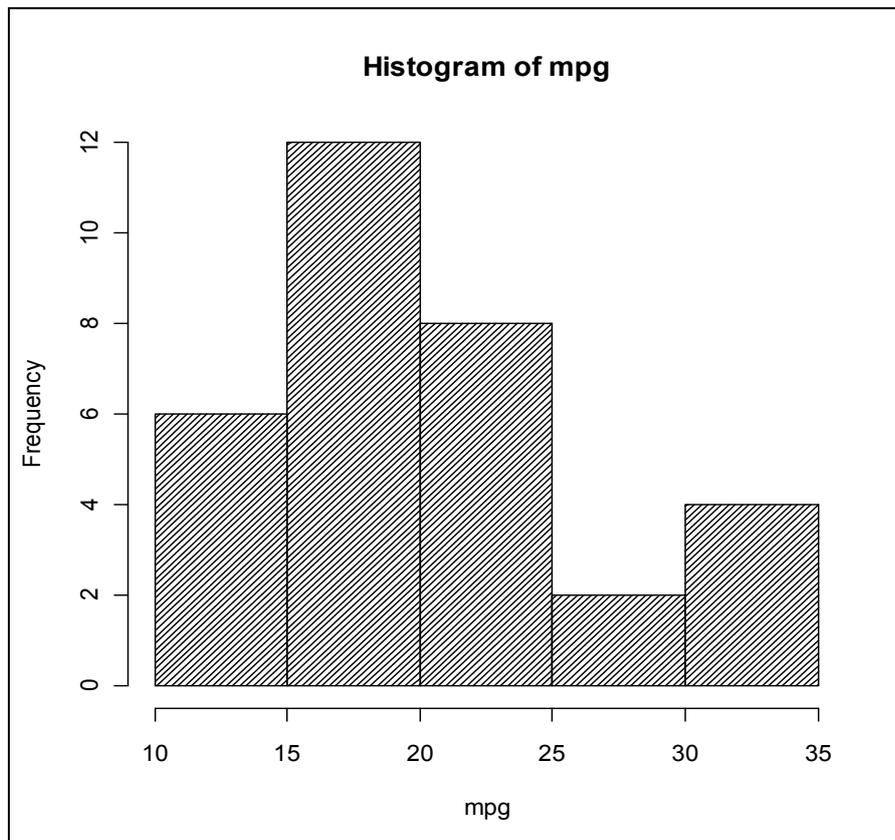


FIGURA 2.17: HISTOGRAMA COM LINHAS DE SOMBREAMENTO

Podemos mudar a angulação dessa linha:

```
hist(mpg,density=1,angle=80)
```

Para colocar barras com cores:

```
hist(mpg,col=c("blue","red","orange","green","pink"))
```

Podemos alterar as cores das bordas:

```
hist(mpg,border=c("blue","red","orange","green","pink"))
```

Ou retirar as bordas:

```
hist(mpg,col=c("blue","red","orange","green","pink"),border=FALSE)
```

O histograma, já vem com um título e nomes nos eixos, mas nós também podemos modificá-los:

```
hist(mpg,main="Histograma de milhas por galão",xlab="MPG",ylab="frequencia")
```

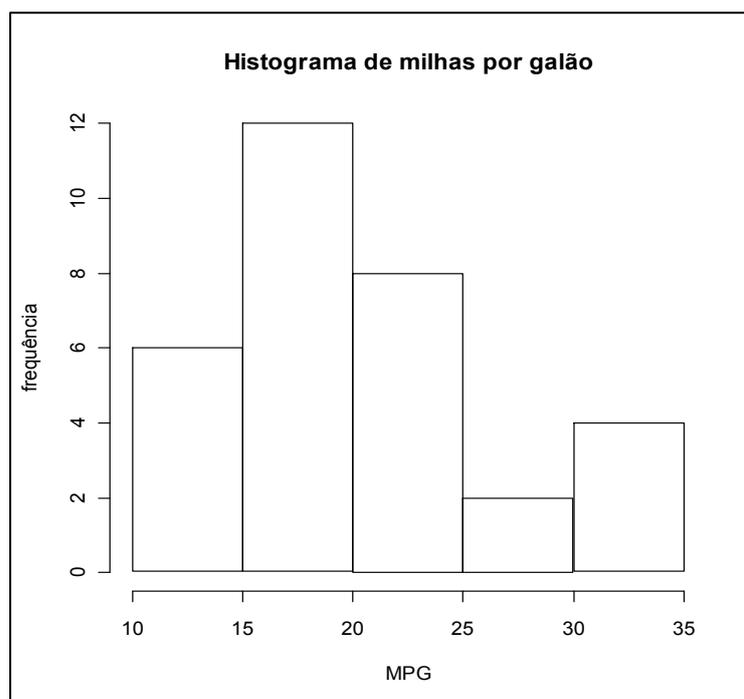


FIGURA 2.18: HISTOGRAMA COM TÍTULO PRINCIPAL E NOS EIXOS.

Podemos definir a quantidade de classes do gráfico:

```
hist(mpg,breaks=3)
```

E também fazer o ramo de folhas:

```
stem(mpg)
```

The decimal point is at the |

10 | 44

12 | 3

14 | 3702258

16 | 438

18 | 17227

20 | 00445

22 | 88

24 | 4

26 | 03

28 |

30 | 44

32 | 49

2.3.1 Exercícios:

1. Os dados são referentes às temperaturas diárias do mês de maio e setembro, em Fahrenheit, na cidade de Nova York em 1973.

tempm=c(67,72,74,62,56,66,65,59,61,69,74,69,66,68,58,64,66,57,68,62,59,73,61,61,57,58,57,67,81,79,76)

temps=c(91,92,93,93,87,84,80,78,75,73,81,76,77,71,71,78,67,76,68,82,64,71,81,69,63,70,77,75,76,68)

- a. Faça o histograma das temperaturas do mês de maio. Coloque título e linhas de sombreamento de densidade 30.
- b. Faça o histograma das temperaturas do mês de setembro. Coloque título e cor = "Violet".
- c. Converta as temperaturas do mês de maio para graus Celsius através da expressão $^{\circ}\text{C} = (^{\circ}\text{F} - 32) / 1,8$. Faça o histograma da frequência relativa, coloque título, sombreamento de densidade 25 e cor = "dark blue"

2.4 Boxplot

O boxplot é um gráfico que possibilita representar a distribuição de um conjunto de dados com base em alguns de seus parâmetros descritivos, quais sejam: a mediana (q_2), o quartil inferior (q_1), o quartil superior (q_3) e do intervalo interquartil ($IQR = q_3 - q_1$).

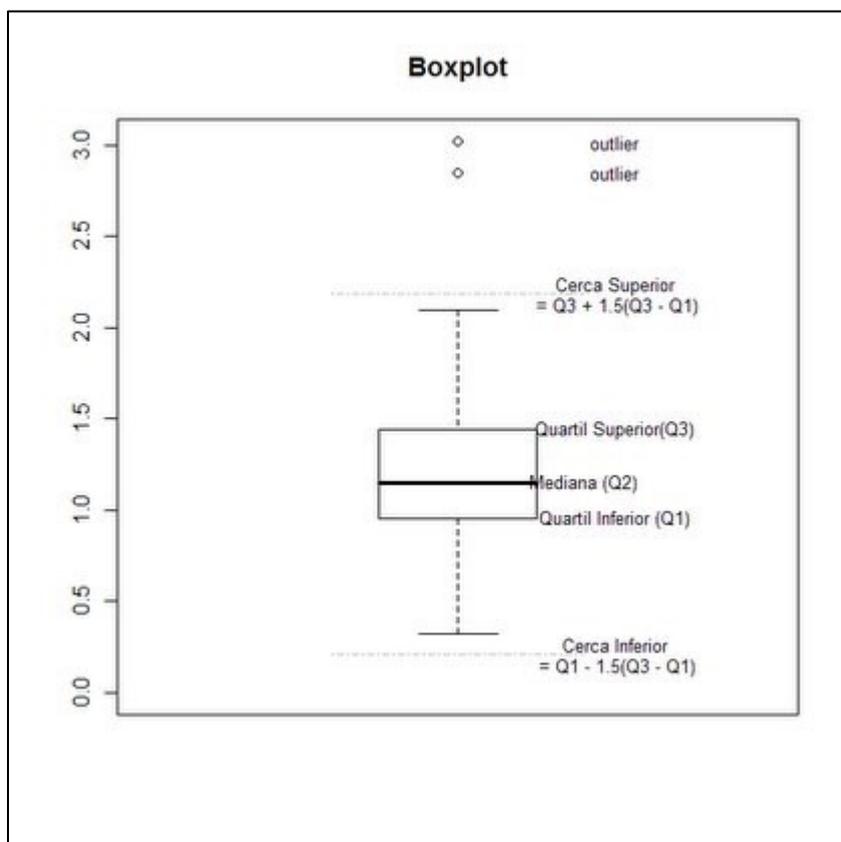


FIGURA 2.19: BOX PLOT ILUSTRATIVO.

Vamos usar os dados do próprio R, sobre uma amostra de 48 pedras de uma reserva de petróleo, onde temos a área, perímetro, formato e permeabilidade. Para exemplificarmos, usaremos a variável formato (shape) que é dada pelo perímetro dividido pela raiz quadrada da área da pedra.

```
data(rock)
```

```
attach(rock)
```

O comando pra fazer um gráfico boxplot é o seguinte:

```
boxplot(shape)
```

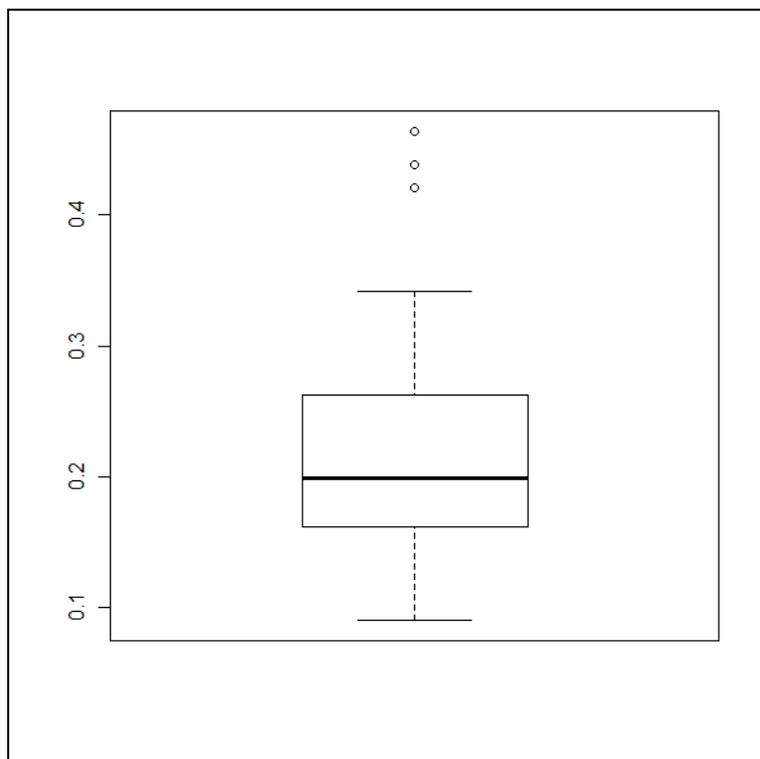


FIGURA 2.20: BOX PLOT SIMPLES

Adicionando título ao gráfico

`boxplot(shape, main="BoxPlot do formato da pedra" , ylab="perimetro/sqrt(area)")`

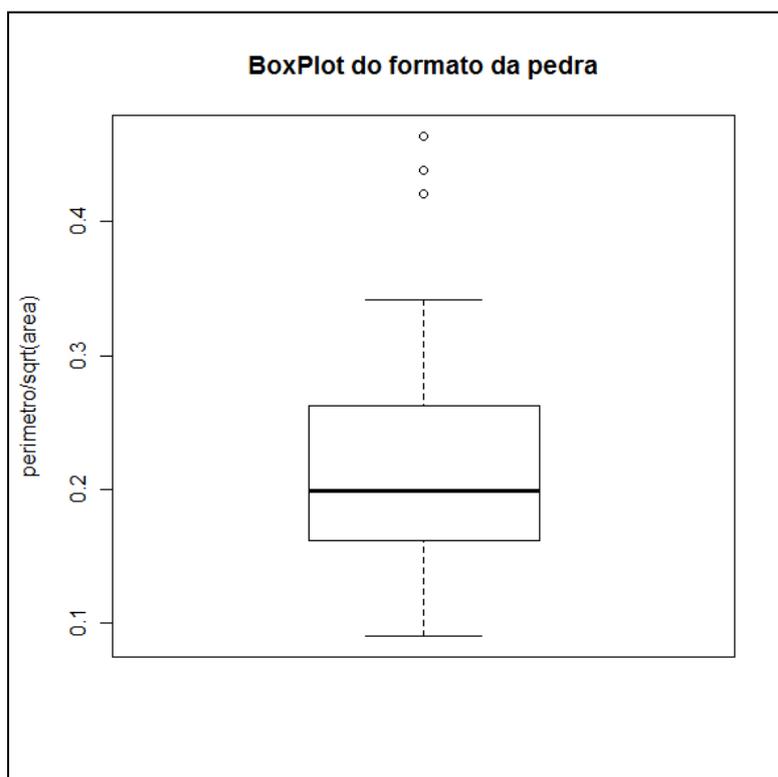


FIGURA 2.21: BOX PLOT COM TÍTULO

Se quisermos que o outlier não apareça no gráfico, utilizamos o seguinte comando:

```
boxplot(shape,outline=FALSE)
```

Podemos fazer o boxplot para vários grupos de observação, facilitando a comparação entre os grupos. Vamos utilizar os dados de `data(PlantGrowth)`

Estes dados são resultados de um experimento para comparar rendimentos (medido pelo peso seco de plantas) obtidos em um controle e duas condições de tratamento diferentes.

```
data(PlantGrowth)
```

```
attach(PlantGrowth)
```

```
boxplot(weight~group)
```

```
title("Boxplot para rendimento de plantas segundo o tratamento", xlab = "tratamento", ylab = "peso")
```

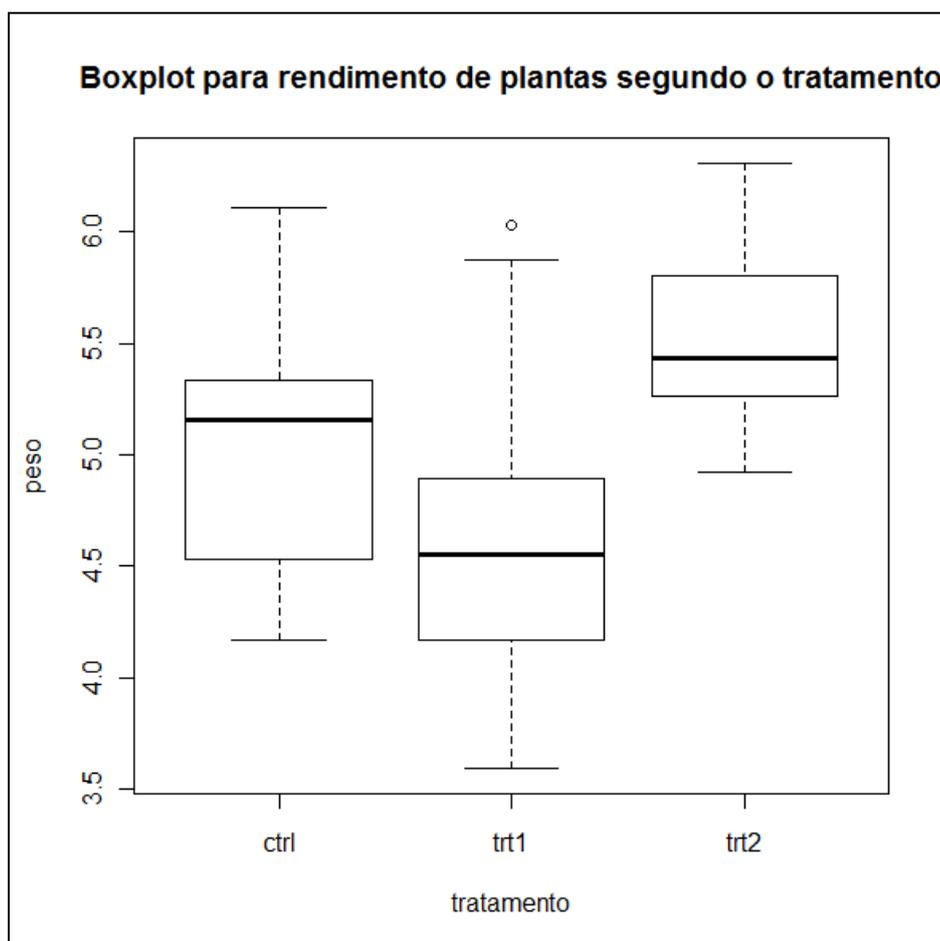


FIGURA 2.22: BOX PLOT PARA VÁRIOS GRUPOS

Aprimoramentos:

Se `varwidth=TRUE`, as caixas são desenhadas com larguras proporcionais à raiz quadrada do número de observações. Este caso se aplica a grupos de tamanhos diferentes.

```
x1<-c(2,3,4,4,5,6,8,9,7,7,7,7,6)
```

```
x2<-c(1,1,1,1,2,3,4,3,2,3,4,5,6,3,4,3,2,2,2,3,4,5,6,3,2,1,1,1,2,3,4)
```

```
boxplot(x1,x2,varwidth=TRUE)
```

Se desejamos nomear os grupos, utilizamos o argumento names:

```
boxplot(x1,x2,names=c("grupo1","grupo2"))
```

Para modificar a largura da caixa para mais estreita utilizando o argumento boxwex

```
boxplot(x1,x2,boxwex=0.3)
```

Para ajustar o tamanho das linhas limites

```
boxplot(x1,x2,staplewex=0.1)
```

Para colocar cor nas bordas:

```
boxplot(shape,border="red")
```

Para preencher a cor de dentro da caixa:

```
boxplot(shape,col="blue")
```

Podemos colocar o gráfico no sentido horizontal:

```
boxplot(shape,horizontal=TRUE,main="BoxPlot do formato da pedra",xlab="valor")
```

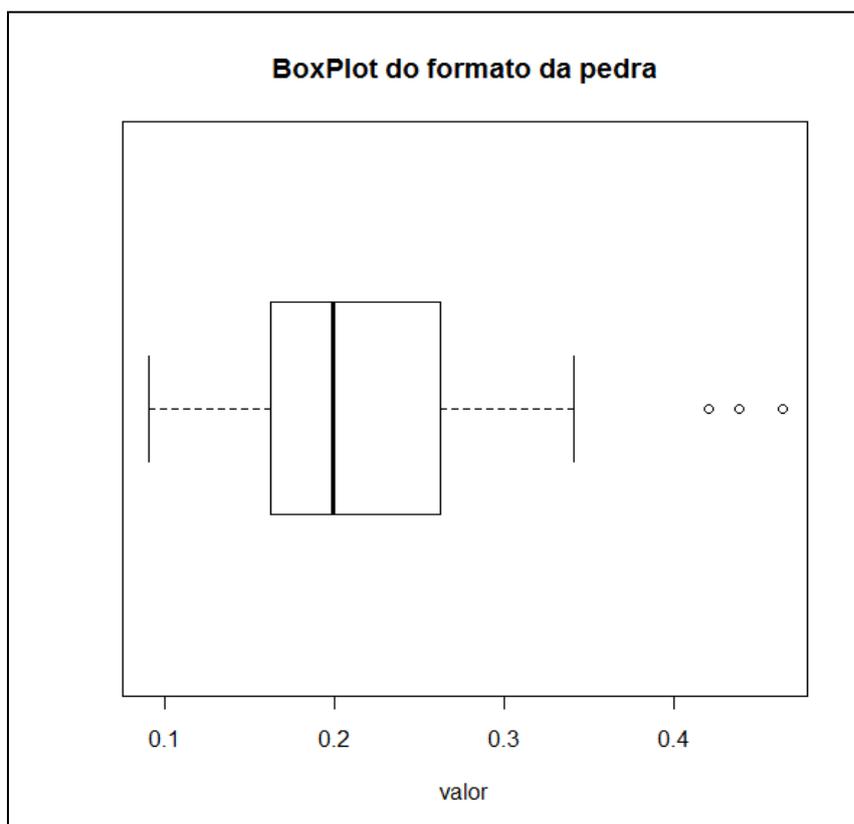


FIGURA 2.23: BOX PLOT HORIZONTAL

Também podemos verificar todos os cálculos utilizados para a formação do gráfico, seus limites, outliers, etc:

```
boxplot.stats(shape)
```

```
$stats
```

```
[1] 0.0903296 0.1621295 0.1988620 0.2626890 0.3412730
```

```
$n
```

```
[1] 48
```

```
$conf
```

```
[1] 0.1759291 0.2217949
```

```
$out
```

```
[1] 0.438712 0.464125 0.420477
```

2.4.1 Exercícios

1. Utilize os dados de contagens de insetos em unidades experimentais agrícolas tratados com inseticidas diferentes disponível no R em `InsectSprays` e construa um boxplot para os 6 tipos de inseticidas. Coloque título adequado, e caixas preenchidas com a cor `lightgray`.
2. Refaça o gráfico anterior sem a presença dos outliers.
3. Refaça o primeiro gráfico reduzindo a largura das caixas.
4. Utilize os dados disponíveis no R em `rock` (veja detalhes sobre estes dados na pag 21) e defina dois vetores:

`x` é o vetor de perímetros cujas pedras possuam área menor ou igual à área média, ou seja,
`x<-peri[area<=mean(area)]`

`y` é o vetor de perímetros cujas pedras possuam área maior à área média, ou seja, `y<-peri[área>mean(area)]`

Faça o boxplot de `x` e `y` semelhante ao da figura 2.22

5. Refaça o gráfico anterior utilizando o argumento `varwidth=TRUE`.

2.5 Gráfico de dispersão

Os diagramas de dispersão são representações de duas variáveis que são organizadas em um gráfico, para observar o padrão de relacionamento entre as mesmas. É um método gráfico que permite verificar a existência ou não de relação entre duas variáveis de natureza quantitativa, ou seja, variáveis que podem ser medidas ou contadas.

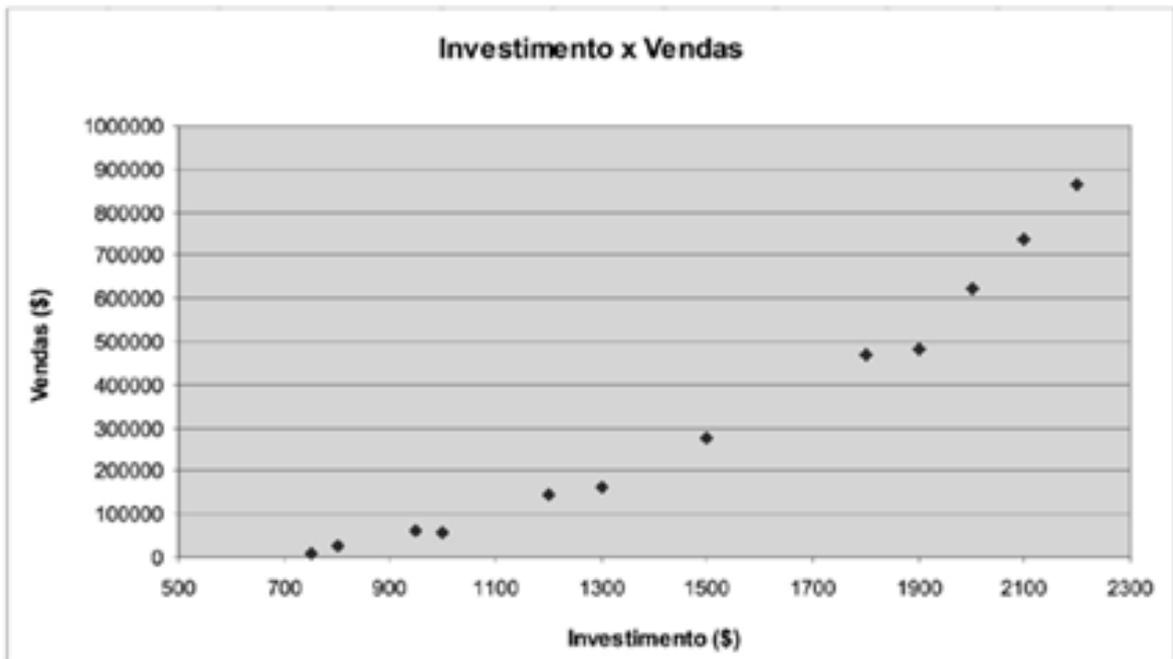


FIGURA 2.24: DIAGRAMA DE DISPERSÃO

Os dados que iremos analisar são medidas da circunferência em mm do tronco de pés de laranjeiras e sua respectiva idade em dias. Com esses dados, criaremos o gráfico de dispersão para observar se há relação entre o tamanho da circunferência com a idade da árvore.

```
data(Orange)
```

```
attach(Orange)
```

Vamos considerar os 7 primeiros valores que correspondem a uma única árvore:

```
plot(age[1:7],circumference[1:7],xlab="idade em dias",ylab="circunferência em mm",main="Dispersão entre idade e circunferência")
```

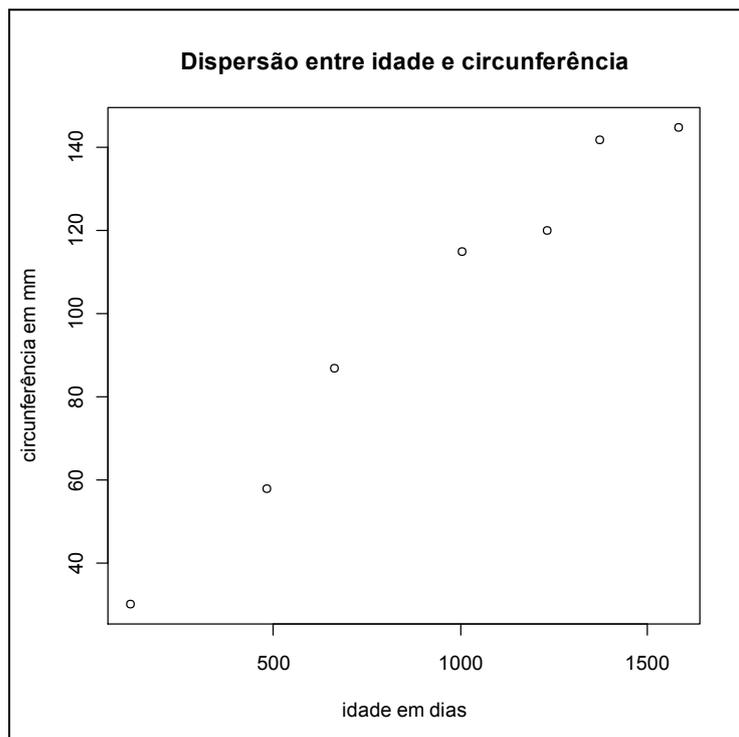


FIGURA 2.25: DIAGRAMA DE DISPERSÃO DOS DADOS DO PÉ DE LARANJEIRA

É interessante colocarmos a reta de regressão linear no mesmo gráfico. Para isso fazemos:

```
plot(age[1:7],circumference[1:7],xlab="idade em dias",ylab="circunferência em mm",main="Dispersão entre idade e circunferência do tronco de pés de laranjeiras")
```

```
abline(lm(circumference[1:7]~ age[1:7]),col=2) #acrescenta a reta de regressão na cor vermelha (col=2).
```

```
text(450,100,"reta de regressão") # acrescenta o texto entre aspas na posição x=450 e y=100.
```

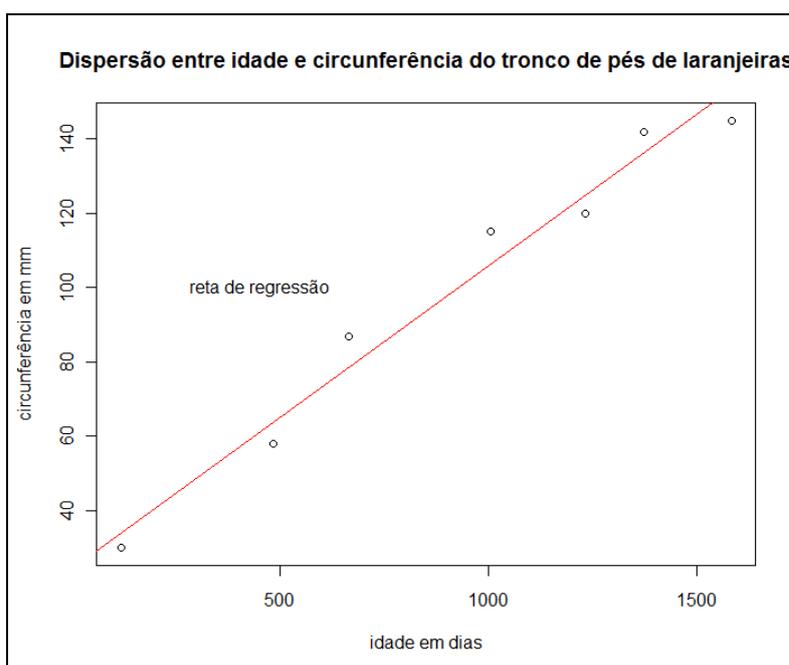


FIGURA 2.26: DIAGRAMA DE DISPERSÃO E A RETA DE REGRESSÃO

A largura da linha que contorna os pontos do gráfico pode ser mudada com o argumento `lwd`,

```
plot(age[1:7],circumference[1:7],lwd=4) #Teste vários números para ver a diferença.
```

Podemos mudar as bolas por outros símbolos. Para cada número temos uma forma diferente:

```
plot(age[1:7],circumference[1:7],pch=20). #Teste vários números para ver a diferença.
```

Podemos colocar vários símbolos num mesmo gráfico:

```
plot(age[1:7],circumference[1:7],pch=1:7)
```

Podemos criar alguns símbolos, por exemplo:

```
plot(age[1:7],circumference[1:7],pch="@")
```

```
plot(age[1:7],circumference[1:7],pch("&"))
```

Colocando cores:

```
plot(age[1:7],circumference[1:7],lwd=4,col="orange")
```

Colocando subtítulo:

```
plot(age[1:7],circumference[1:7], sub="Desde 1968")
```

E criar um texto também na posição escolhida (400,80)

```
text(400,80,"Isso aqui pode ser escrito aqui")
```

Lembrando que para colocar

título principal : `main = "texto"`

título nos eixos: `xlab="texto"` ou `ylab="texto"`

subtítulo: `sub="texto"`

Podemos ainda mudar os limites dos eixos:

```
plot(age[1:7],circumference[1:7], xlim=range(100:1600),ylim=range(20:180))
```

Ainda podemos colocar mais de um gráfico numa mesma janela:

```
par(mfrow=c(2,3)) #assim podemos fazer 6 gráficos em 2 linhas e 3 colunas.
```

```
plot(age[Tree==1],circumference[Tree==1],xlab="idade",ylab="circunferência", sub="árvore1")
```

```
plot(age[Tree==2],circumference[Tree==2],xlab="idade",ylab="circunferência", sub="árvore2")
```

```
plot(age[Tree==3],circumference[Tree==3],xlab="idade",ylab="circunferência", sub="árvore3")
```

```
plot(age[Tree==4],circumference[Tree==4],xlab="idade",ylab="circunferência", sub="árvore4")
```

```
plot(age[Tree==5],circumference[Tree==5],xlab="idade",ylab="circunferência", sub="árvore5")
```

```
plot(age,circumference,xlab="idade",ylab="circunferência", sub="árvores")
```

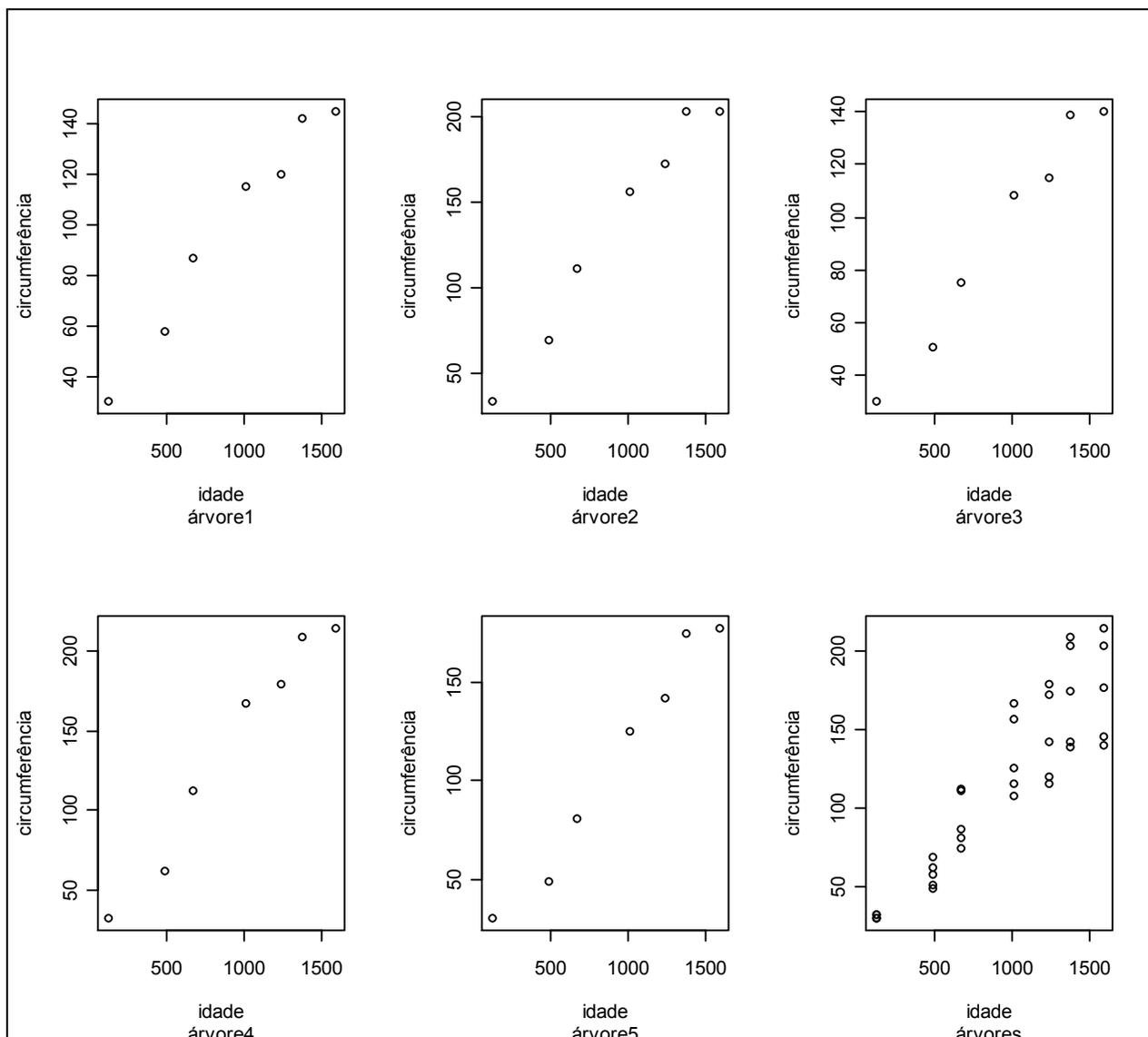


FIGURA 2.27: DIAGRAMA DE DISPERSÃO LADO A LADO

2.5.1 Exercícios

1- Um biólogo foi ao campo e contou o número de sapos em 20 locais. Ele também anotou a umidade e a temperatura em cada local. Faça dois gráficos de dispersão para mostrar a relação do número de sapos com as variáveis temperatura e umidade. Use a função par() para dividir a janela em duas.

Os dados são:

sapos 6-5-10-11-26-16-17-37-18-21-22-15-24-25-29-31-32-13-39-40

umid 62-24-21-30-34-36-41-48-56-74-57-46-58-61-68-76-79-33-85-86

temp 31-23-28-30-15-16-24-27-18-10-17-13-25-22-34-12-29-35-26-19

2- Um biólogo interessado em saber se o número de aves está relacionado ao número de uma determinada espécie de árvore, realizou amostras em 10 locais. Os valores obtidos foram:

aves<-c(22,28,37,34,13,24,39,5,33,32)

arvores<-c(25,26,40,30,10,20,35,8,35,28)

Faça um gráfico que mostra a relação entre o número de aves e o número de árvores. Um colega coletou mais dados sobre aves e árvores, em outra área, que podemos aproveitar. Os dados são:

arvores2<-c(6,17,18,11,6,15,20,16,12,15)

aves2<-c(7,15,12,14,4,14,16,60,13,16)

Inclua estes novos pontos no gráfico com um símbolo diferente e cor azul.

3 – Repita o gráfico do exercício 2 acima e faça as seguintes modificações.

Coloque um título no gráfico

Use Bolinhas cheias e azuis como símbolo.

Coloque as legendas no eixo-x e no eixo-y.

O padrão do R é fazer uma —caixa entorno do gráfico, faça uma alteração para que apareça apenas as linhas do eixo-x e do eixo-y. Veja bty em ?par

4-Foram levantados os tempos (em horas por semana) que 20 alunos de uma escola da rede pública gastam na Internet e praticando atividades físicas. Os resultados deste levantamento são apresentados a seguir:

Aluno 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Internet 30 8 20 15 12 4 40 25 20 14 22 2 15 18 30 20 24 15 8 8

At. Físicas 2 10 5 5 8 15 0 4 5 10 2 15 6 4 2 6 4 10 12 15

Construa um gráfico de linhas para esse par de variáveis. Com base neste gráfico, você acredita que há relação entre as horas dedicadas às duas atividades?

2.6 Gráfico de linhas

Este gráfico é útil para plotar uma série de dados ao longo do tempo ou para ligar pontos por linhas.

Vamos ilustrar uma situação onde desejamos plotar o gráfico de duas funções, x ao cubo e $-x$ ao cubo no mesmo gráfico. Neste caso, definiremos inicialmente uma janela gráfica usando o argumento `type="n"`. Com este argumento um "gráfico em branco" é criado, são ajustadas apenas as margens e eixos do gráfico e o restante é deixado em branco. A seguir adicionam-se linhas e pontos desejados. Você deve fornecer coordenadas x e y que cubram a amplitude de valores de todos os elementos que você deseja adicionar ao gráfico.

```
x <- 0:20
```

```
y <- x**3
```

```
plot(c(0,20),c(-8000,8000),type='n',xlab=NA,ylab=NA)
```

```
lines(x,y)
```

```
lines(x,-y, col='red')
```

```
title("Gráfico de duas funções",xlab="valores de x", ylab="valores de y")
```

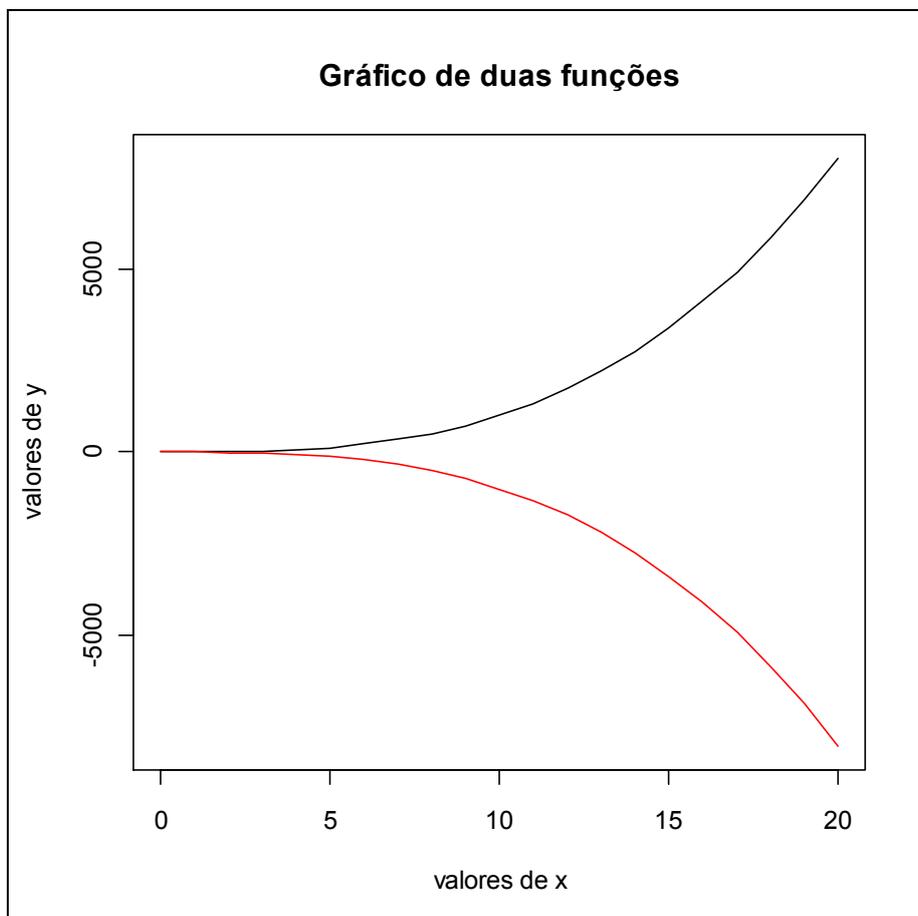


FIGURA 2.28: GRÁFICO DE DUAS FUNÇÕES

Vamos mostrar agora o gráfico de uma série de valores ao longo do tempo.

Considere os dados da tabela 2.3, correspondentes à temperatura média obtida em Ithaca (NY):

TABELA 2.3: TEMPERATURA MÉDIA MENSAL DE ITHACA-NY

mês	Temperatura °C
1	-5.44
2	-5.17
3	0.11
4	6.89
5	12.67
6	17.94
7	20.44
8	19.5
9	15.67
10	9.72
11	4.06
12	-2.56

```
mes<-1:12
```

```
temperatura<-c(-5.44,-5.17,0.11,6.89,12.67,17.94,20.44,19.5,15.67,9.72,4.06,-2.56)
```

```
plot(mes,temperatura,type='l')
```

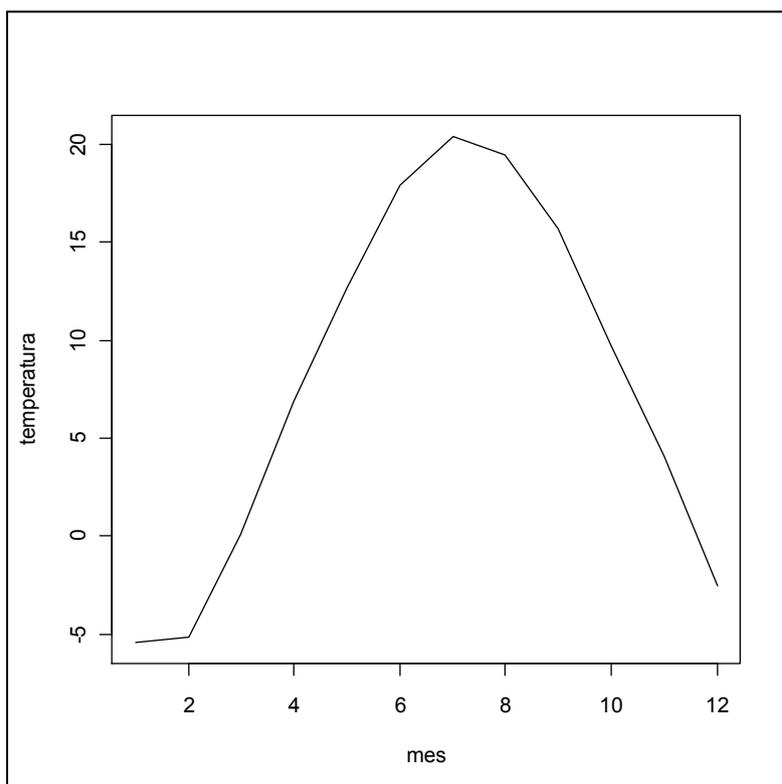


FIGURA 2.29: GRÁFICO DE LINHA DA SÉRIE DE TEMPERATURA

ou se quiser marcar os pontos substitua o comando acima por:

```
plot(mes,temperatura,type='b')
```

```
title("Gráfico da temperatura ao longo dos meses")
```

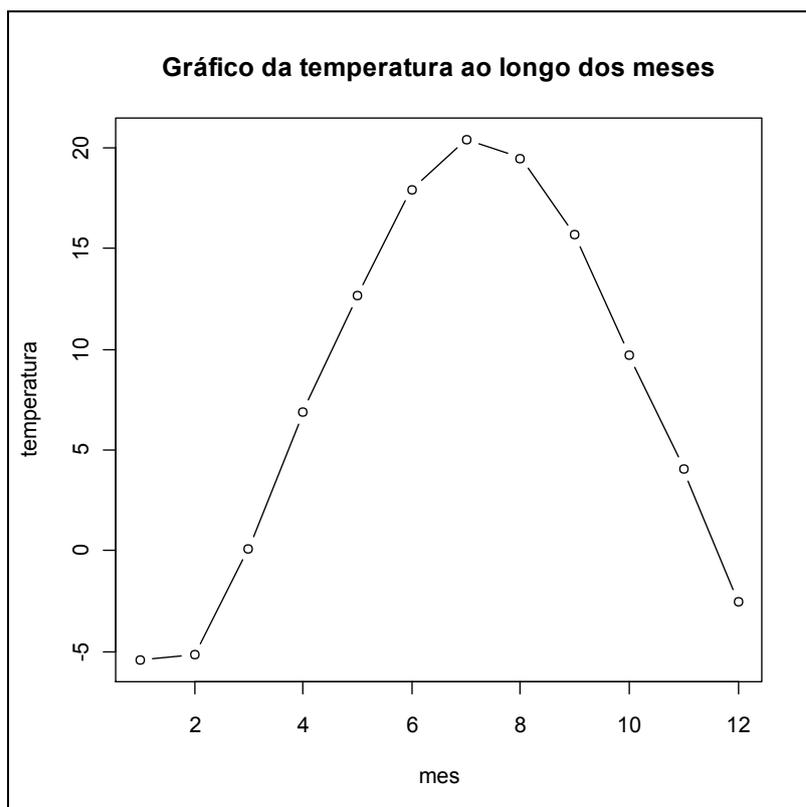


FIGURA 2.30: GRÁFICO DE LINHA E PONTOS DA SÉRIE DE TEMPERATURA

Podemos utilizar comandos já vistos anteriormente para adicionar título ao eixos, mudar tamanho da fonte, mudar cor, símbolos, adicionar subtítulo, mudar os limites dos eixos, etc.

`main="texto"` adiciona título ao gráfico

`xlab` ou `ylab` = "texto" adiciona título ao eixo x ou y

`cex` = número altera o tamanho da fonte

`type` = "p" para pontos,
 "l" para linhas,
 "b" para pontos e linhas,
 "c" para linhas descontínuas nos pontos,
 "o" para pontos sobre as linhas,
 "n" para nenhum gráfico, apenas a janela.

`col` = "nome" ou número da cor

`pch` = número (0 a 25)

lwd= número (controla a espessura da linha)

Podemos ainda grafar duas ou mais séries no mesmo gráfico:

```
ano<-2001:2009
```

```
tri1<-c(72.8,66.2,69.2,65.9,62.4,67.8,61.3,68.5,70.4)
```

```
tri2<-c(60.6,53.7,55.3,56.7,56.4,57.8,57.5,59.8,63.3)
```

```
plot(ano, tri1,type="l",main="Taxa de ocupação por trimestre dos hotéis - Município do Rio de Janeiro",xlab="ano",ylab="Taxa de ocupação %",col="blue",ylim=c(50,80))
```

```
lines(ano, tri2,col="red")
```

É sempre possível mudar o tipo de linha usando o argumento type=:

```
plot(ano, tri1,type="b",main="Taxa de ocupação por trimestre dos hotéis - Município do Rio de Janeiro",xlab="ano",ylab="Taxa de ocupação %",col="blue",ylim=c(50,80))
```

```
lines(ano, tri2,col="red", type="b")
```

```
legend(2007,80,c("1o.trim","2o.trim"),col =c("blue","red"), pch=rep(20,2))
```

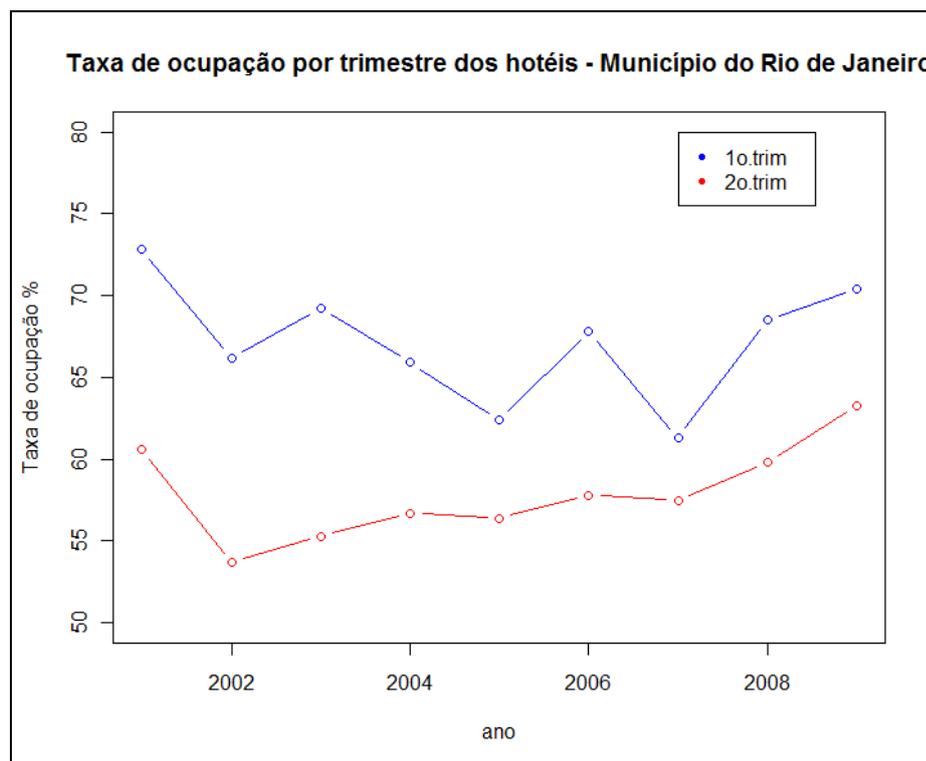


FIGURA 2.31: GRÁFICO DE LINHA E PONTOS DE DUAS SÉRIES NO MESMO PERÍODO DO TEMPO

2.6.1 Exercícios

- Os dados (já retirados do R) mostram a velocidade em que o carro está e a distancia que ele demora a parar. Crie um gráfico de linhas, colocando cor na linha, título e nome nos eixos.

cars

attach(cars)

2. Os seguintes dados são medidas de viscosidade para um produto químico observado de hora em hora (leia de cima para baixo e da esquerda para a direita).

47,9 48,8 48,6

47,9 48,1 48,0

48,6 48,3 47,9

48,0 47,2 48,3

48,4 48,9 48,5

48,1 48,6 48,1

48,0 48,0 48,0

48,6 47,5 48,3

Construa um gráfico de série temporal para esses dados.

3. De acordo com Ministério da Educação a quantidade e alunos matriculados no ensino de 1º grau no Brasil nos de 1990 a 1996 em milhares de alunos, são: 19.720 – 20.567 – 21.473 – 21.887 – 20.598 – 22.473 – 23.564. Faça um gráfico de série temporal para apresentar esses dados.

2.7 Gráfico da distribuição normal

A distribuição normal é uma das mais importantes distribuições da estatística, conhecida também como Distribuição de Gauss ou Gaussiana. É inteiramente descrita por seus parâmetros de média e desvio padrão, ou seja, conhecendo-se estes se consegue determinar qualquer probabilidade em uma distribuição Normal.

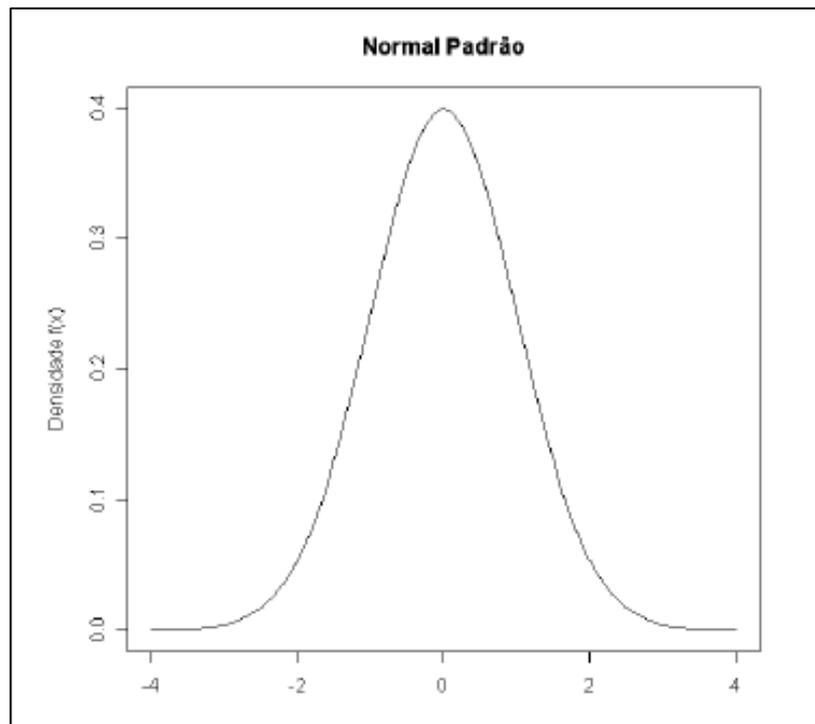


FIGURA 2.32: GRÁFICO DA DISTRIBUIÇÃO NORMAL DE MÉDIA 0 E DESVIO PADRÃO 1.

Geração de números aleatórios

`rnorm(50,mean=2,sd=sqrt(3))` # amostra de 50 número extraídos de uma distribuição normal com média = 2, variância = 3.

`rnorm(20)` # amostra de tamanho 20 da distribuição normal padrão, média = 0 e variância =1.

Histograma da freqüência relativa de uma amostra de 500 números extraídos de uma distribuição normal com média = 2, variância = 4.

`hist(rnorm(500,mean=2,sd=sqrt(4)),freq=FALSE)`

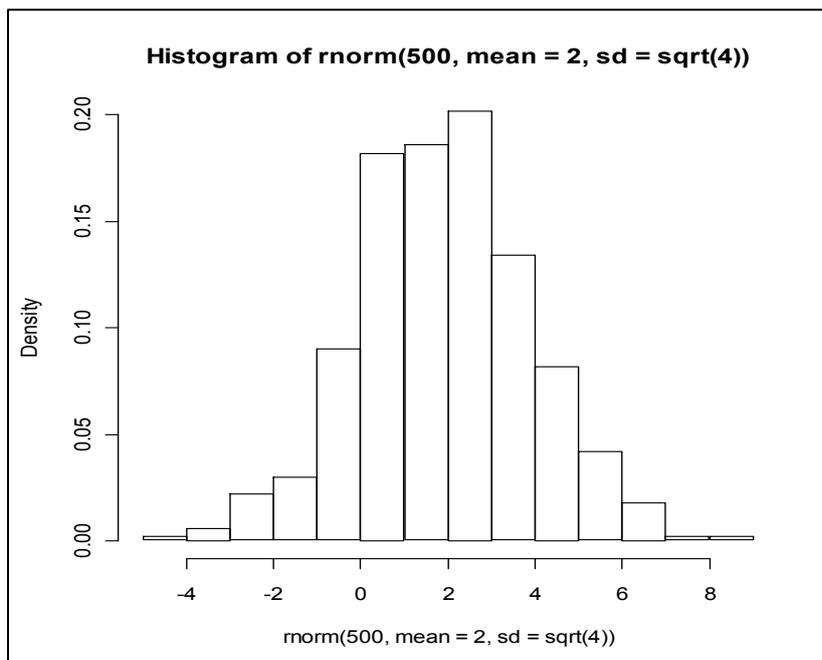


FIGURA 2.33: HISTOGRAMA DE UMA VARIÁVEL COM DISTRIBUIÇÃO NORMAL

Podemos traçar no mesmo gráfico a curva da função de densidade da distribuição normal com média 2 e variância 4.

Observe que `col=2` fará o traçado vermelho, `lty=2` fará linha tracejada e `lwd=2` se refere a espessura da linha e `add=TRUE` sobrepõe a curva ao gráfico anterior.

`curve(dnorm(x,mean=2,sd=sqrt(4)),col=2,lty=2,lwd=2,add=TRUE)`

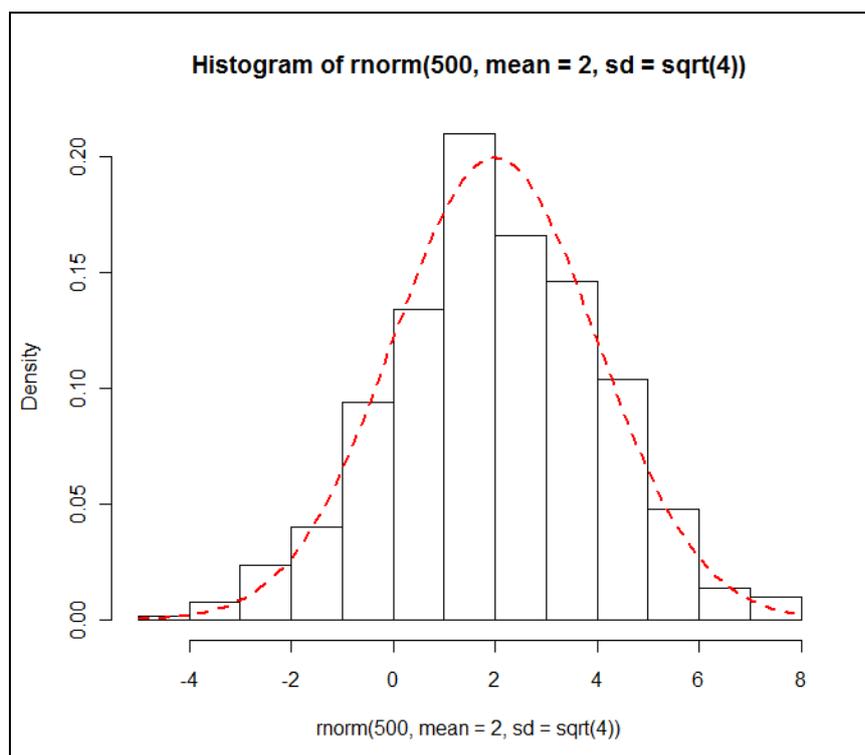


FIGURA 2.34: GRÁFICO DE HISTOGRAMA COM A CURVA DA NORMAL

Vamos ilustrar graficamente o que acontece com os parâmetros da função de densidade normal.

Fixando a variância e variando a média):

```
curve(dnorm(x,mean=1,sd=sqrt(2)),lwd=2,from=-6,to=17)
```

```
curve(dnorm(x,mean=10,sd=sqrt(2)),col=2,lwd=2,add=T)
```

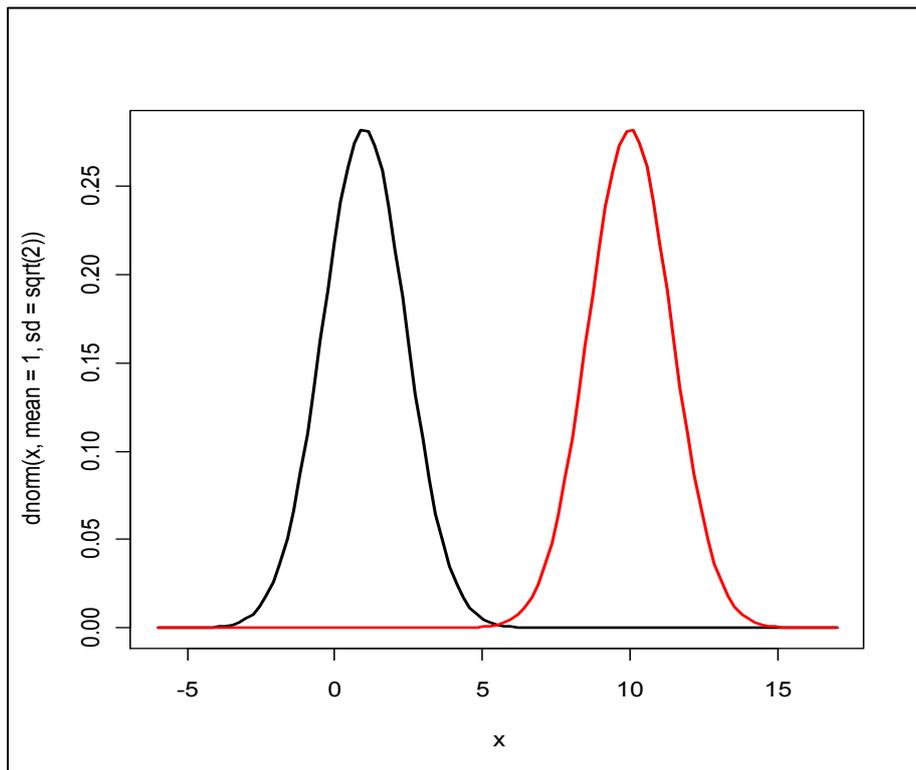


FIGURA 2.35: GRÁFICOS DE DUAS NORMAIS COM MÉDIAS DIFERENTES

Adicionando legenda e título ao gráfico:

```
legend('topright',legend=c(expression(mu==1),expression(mu==10)),text.col=c(1,2),cex=1.5)
```

```
title("Comparando a distribuição normal com médias diferentes")
```

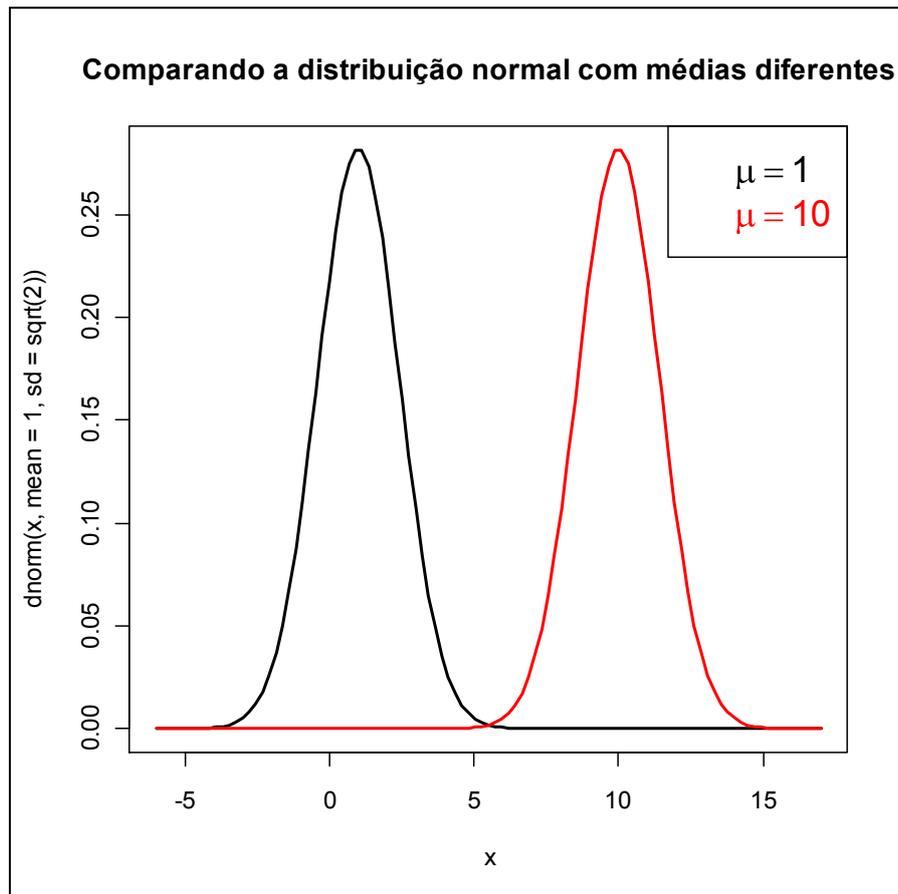


FIGURA 2.36: GRÁFICO DE DUAS NORMAIS COM MÉDIAS DIFERENTES E LEGENDA

Fixando a média e variando a variância:

```
curve(dnorm(x,mean=4,sd=sqrt(3)),lwd=2,from=-15,to=25)
```

```
curve(dnorm(x,mean=4,sd=sqrt(20)),col=2,lwd=2,add=T)
```

```
legend('topright',legend=c(expression(sigma^2==3),expression(sigma^2==20)),text.col=c(1,2),cex=1.5)
```

```
title("Comparando a distribuição normal com variâncias diferentes")
```

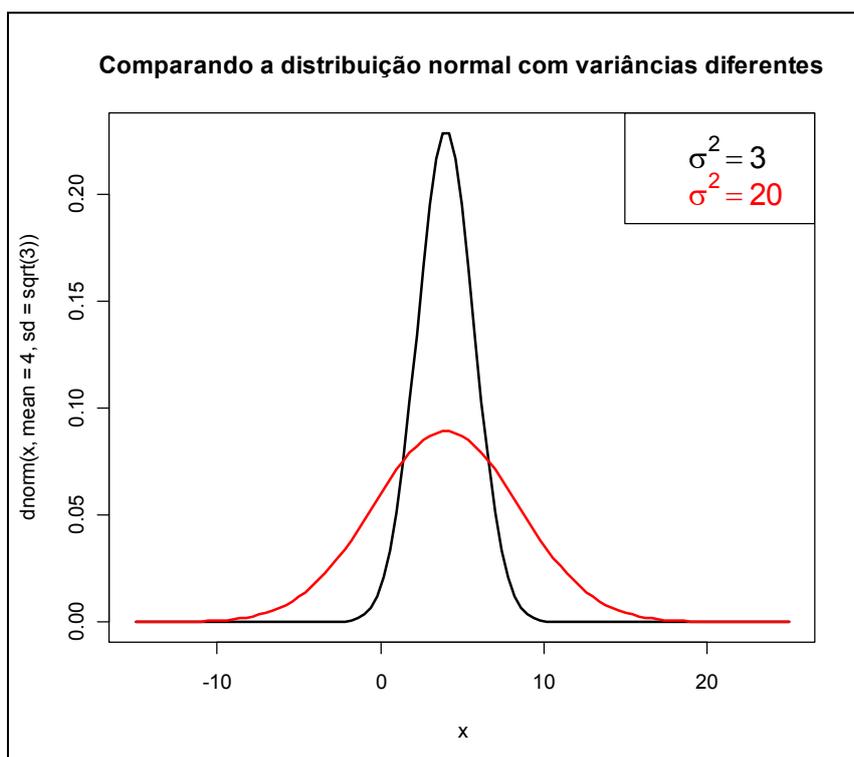


FIGURA 2.37: GRÁFICO DE DUAS NORMAIS COM VARIÂNCIAS DIFERENTES

2.7.1 Exercícios

1- Dez observações do tempo efetivo de vida de um catalisador usado em reações químicas produziram os resultados: 1176, 1191, 1214, 1220, 1205, 1192, 1201, 1190, 1183 e 1185. Supondo que estes tempos sigam a distribuição normal, calcule a média e o desvio padrão para construir o gráfico de uma normal com estes parâmetros.

2- Construa o gráfico da distribuição normal variando a média no intervalo $[-1,1]$ com tamanho de passo igual a 0.5. Fixe o desvio padrão em 1. Apresente uma legenda e cores diferentes para cada curva.

3- Construa o gráfico da distribuição normal variando o desvio padrão no intervalo $[0,2]$ com tamanho de passo igual a 0.5. Fixe a média em 0. Apresente uma legenda e cores diferentes para cada curva.

4- Considere o peso de uma amostra de 15 pacotes de açúcar da marca X.

pesos=

`c(0.9475,0.9705,0.9770,0.9775,0.9860,0.9960,0.9965,0.9975,1.0050,1.0075,1.0100,1.0175,1.0180,1.0200,1.0250)`

Construa o histograma destes pesos e no mesmo gráfico insira a curva de uma distribuição normal com média 1 e desvio padrão 0.05 na cor azul e outra com média igual a média dos pesos e desvio padrão igual ao desvio padrão dos pesos na cor vermelha. Coloque legenda para cada curva.

3 Referência Bibliográfica

Arriaza Gómez, A.J. et al. *Estatística Básica con R y R-commander*. Version Febrero 2008. Servicio de Publicaciones de La Universidad de Cadiz, 2008. 128p.

Cavalcanti, C.V.; Alcoforado, L.F. *UTILIZAÇÃO DO R NO ENSINO DA ESTATÍSTICA BÁSICA I*. Material de apoio, GET/UFF, 2010, disponível em http://www.professores.uff.br/luciane/images/stories/Arquivos/ap_mci2010.pdf

Crawley, M.J. *The R Book*. Ed. John Wiley Professio, 2007, 942p.

IBGE. Dados sobre Niterói, disponível em <http://www.ibge.gov.br/cidadesat/>

Morettin, P.A.; Bussab, W.O. *Estatística Básica*. 5.ed. Ed. Saraiva, 2009, 526p.

R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2009, disponível em <http://www.R-project.org>

4 Resposta dos exercícios

Exercício 2.1.1

1- `alunosprof<-c(25280/1751,21328/1186,18432/947,280/29)`

`escola<-c("Privada","Estadual","Municipal","Federal")`

`barplot(alunosprof,main="Quatidade de alunos por professor",names.arg=escola,ylab="Quatidade",xlab="Escolas De Ensino Fundamental")`

2- `barplot(alunosprof,main="Quatidade de alunos por professor",names.arg=escola,xlab="Quatidade",ylab="Escolas De Ensino Fundamental",horiz=TRUE)`

3- `barplot(alunosprof,main="Quatidade de alunos por professor",names.arg=escola,ylab="Quatidade",xlab="Escolas De Ensino Fundamental",density=60)`

4- `barplot(alunosprof,main="Quatidade de alunos por professor",names.arg=escola,ylab="Quatidade",xlab="Escolas De Ensino Fundamental",density=50,angle=75)`

5- `peso=c(42,51,59,64,76)`

`dias=c("Segundo","Quarto","Sexto","Oitavo","Décimo")`

`barplot(peso,main="Peso médio dos pintinhos recém nascidos",names.arg=dias,ylab="Peso",xlab="Dias")`

6- `barplot(peso,main="Peso médio dos pintinhos recém nascidos",names.arg=dias,ylab="Peso",xlab="Dias",col=c("blue","pink","yellow","green","red"))`

7- `barplot(peso,main="Peso médio dos pintinhos recém nascidos",names.arg=dias,xlab="Peso",ylab="Dias",col=c("blue","pink","yellow","green","red"),horiz=TRUE)`

8- `ocorrencia=c(250,110,85,45,25)`

`defeito=c("linha ruidosa","linha aberta","alarme","não responde","não toca")`

`barplot(ocorrencia,main="Defeitos em linhas Telefonicas",names.arg=defeito,xlab="tipo de de defeito",ylab="Número de ocorrências",density=80,angle=120)`

9-

`fumaenao=matrix(c(78,133,142,206,197,220,151,412,3300,72,23,15,7,8,9,11),nrow=8,ncol=2,dimnames=list(c("0 - 13","14 - 49","50 - 99","100 - 149","150 - 199","200 - 249","250 - 299","300 +"),c("Fumantes","Não Fumantes")))`

`barplot(fumaenao[,2:1],beside=TRUE,legend.text=rownames(fumaenao), main="Distribuição de nível de cotinina em fumantes e não fumantes",ylab="nível de cotinina",xlab="pessoas",sub="fonte:www.google.com.br")`

Exercício 2.2.1

1- `ocorrencia=c(250,110,85,45,25)`

```

names(ocorrencia)=c("linha ruidosa","linha aberta","alarme","não responde","não toca")
porc<-round(ocorrencia*100/sum(ocorrencia),2)
rotulos<-paste("(",porc,"%)",sep="")
pie(ocorrencia, main="Defeitos em linhas Telefonicas",labels=rotulos, col=rainbow(7))
legend(1,1,names(ocorrencia),col = rainbow(7),pch=rep(20,6))
2- rotulos<-paste(names(ocorrencia),"(",porc,"%)",sep="")
pie(ocorrencia, main="Defeitos em linhas Telefonicas",labels=rotulos, cex=0.8, col=rainbow(7))
3- despesas=c(22200,10500,3530,1040,341)
names(despesas)=c("comida e cigarro", "domésticos", "remédios e saúde", "cuidados pessoais",
"educação privada")
porc<-round(despesas*100/sum(despesas),2)
rotulos<-paste("(",porc,"%)",sep="")
pie(despesas, main="Despesas pessoais dos Estados Unidos (em bilhões de dólares)",labels=rotulos,
col=rainbow(7))
legend(1,1,names(despesas),col = rainbow(7),pch=rep(20,6))
4-rotulos<-paste(names(despesas),"(",porc,"%)",sep="")
pie(ocorrencia, main="Despesas pessoais dos Estados Unidos (em bilhões de dólares)",labels=rotulos,
cex=0.8, col=rainbow(7))

```

Exercicio 2.3.1

```

1-a)
tempm=c(67,72,74,62,56,66,65,59,61,69,74,69,66,68,58,64,66,57,68,62,59,73,61,61,57,58,57,67,81,79,
76)
hist(tempm,main="Temperaturas diárias do mês de maio",density=30)
b)
temps=c(91,92,93,93,87,84,80,78,75,73,81,76,77,71,71,78,67,76,68,82,64,71,81,69,63,70,77,75,76,68)
hist(temps,main="Temperaturas diárias do mês de setembro",col="violet")
c) graus = (tempm - 32) / 1.8
hist(graus,freq=FALSE,main="Frequência relativa das temperaturas diárias do mês de
maio",density=25,col="dark blue")

```

Exercicio 2.4.1

```

1- data(InsectSprays)

```

```
attach(InsectSprays)
```

```
boxplot(count~spray,main="Boxplot para contagem de insetos tratados com inseticidas",col="lightgray",xlab="tipo de inseticida")
```

```
2- boxplot(count~spray,main="Boxplot para contagem de insetos tratados com inseticidas",col="lightgray", xlab="tipo de inseticida",outline=FALSE)
```

```
3- boxplot(count~spray,main="Boxplot para contagem de insetos tratados com inseticidas",col="lightgray", xlab="tipo de inseticida",boxwex=0.3)
```

```
4- attach(rock)
```

```
x<-peri[area<=mean(area)]
```

```
y<-peri[area>mean(area)]
```

```
peri<-c(x,y)
```

```
area<-c(rep("<=área média",length(x)),rep(">área média",length(y)))
```

```
dados<-data.frame(peri,area)
```

```
boxplot(peri~area,main="Boxplot para perímetro de pedras",xlab="area da pedra")
```

```
5- boxplot(peri~area,main="Boxplot para perímetro de pedras",xlab="area da pedra",varwidth=TRUE)
```

Exercício 2.5.1

```
1- sapos=c(6,5, 10,11,26,16,17,37,18,21,22,15,24,25,29,31,32,13,39,40)
```

```
umid=c(62,24,21,30,34,36,41,48,56,74,57,46,58,61,68,76,79,33,85,86)
```

```
temp=c(31,23,28,30,15,16,24,27,18,10,17,13,25,22,34,12,29,35,26,19)
```

```
par(mfrow=c(1,2))
```

```
plot(umid,sapos)
```

```
plot(temp,sapos)
```

```
2- aves<-c(22,28,37,34,13,24,39,5,33,32)
```

```
arvores<-c(25,26,40,30,10,20,35,8,35,28)
```

```
plot(arvores,aves)
```

```
arvores2<-c(6,17,18,11,6,15,20,16,12,15)
```

```
aves2<-c(7,15,12,14,4,14,16,60,13,16)
```

```
points(arvores2,aves2,pch=3,col="blue")
```

```
3- plot(arvores,aves,main="Dispersão entre aves e árvores",xlab="árvores",ylab="aves", col="dark blue",bty='n')
```

```

points(arvores2,aves2,pch=20,col="blue")
4- internet=c(30, 8, 20, 15, 12, 4, 40, 25, 20, 14, 22, 2, 15, 18, 30, 20, 24, 15, 8, 8)
at.fisicas=c(2, 10, 5, 5, 8, 15, 0, 4, 5, 10, 2,15, 6, 4, 2, 6, 4, 10, 12, 15)
plot(internet,at.fisicas,main="Gráfico das horas na internet e praticando atividades
física",ylab="Atividade Física",xlab="Internet",col="red")

```

Exercício 2.6.1

```

1- data(cars)
attach(cars)
plot(c(4,25),c(2,120),type="n",main="Gráfico da velocidade e o tempo que o carro demora para
parar",ylab="Distancia",xlab="Velocidade")
lines(speed,dist,col="green")
2- visc=c(47.9, 48.8, 48.6,47.9, 48.1, 48.0,48.6, 48.3, 47.9,48.0, 47.2, 48.3,48.4, 48.9, 48.5,48.1, 48.6,
48.1,48.0, 48.0, 48.0,48.6, 47.5, 48.3)
plot(c(0,24),c(47,49),type='n',main="Gráfico de viscosidade para um produto químico",xlab="horas do
dia",ylab="viscosidade")
lines(visc,col=5)
3-alunos=c(19720,20567,21473,21887,20598,22473,23564)
anos<-1990:1996
plot(anos,alunos,type="b",main="Matriculas em milhares de alunos, 1990-
1996",ylab="Alunos",xlab="Anos",col="purple")

```

Exercício 2.7.1

```

1- x<-c(1176, 1191, 1214, 1220, 1205, 1192, 1201, 1190, 1183, 1185)
mi<-mean(x)
sigma<-sd(x)
curve(dnorm(x,mean=mi,sd= sigma),lwd=2,from=mi-3*sigma,to=mi+3*sigma)
2- curve(dnorm(x,mean=-1,sd=1),lwd=2,from=-9,to=9)
curve(dnorm(x,mean=-0.5,sd=1),col=2,lwd=2,from=-9,to=9,add=TRUE)
curve(dnorm(x,mean=0,sd=1),col=3,lwd=2,from=-9,to=9,add=TRUE)
curve(dnorm(x,mean=0.5,sd=1),col=4,lwd=2,from=-9,to=9,add=TRUE)

```

```
curve(dnorm(x,mean=1,sd=1),col=5,lwd=2,from=-9,to=9,add=TRUE)
```

```
legend('topright',legend=c(expression(mi==1),expression(mi==0.5),
expression(mi==0),expression(mi==0.5)),text.col=c(1,2,3,4,5),cex=.85)
```

```
title("distribuição normal com variância 1 e médias variando")
```

```
3- curve(dnorm(x,mean=0,sd=0),lwd=2,from=-3,to=3)
```

```
curve(dnorm(x,mean=0,sd=0.5),col=2,lwd=2,add=TRUE)
```

```
curve(dnorm(x,mean=0,sd=1),col=3,lwd=2, add=TRUE)
```

```
curve(dnorm(x,mean=0,sd=1.5),col=4,lwd=2,add=TRUE)
```

```
curve(dnorm(x,mean=0,sd=2),col=5,lwd=2, add=TRUE)
```

```
legend('topright',legend=c(expression(sigma==0),expression(sigma==0.5),
expression(sigma==1),expression(sigma==1.5),
expression(sigma==2)),text.col=c(1,2,3,4,5),cex=.85)
```

```
title("distribuição normal com média 0 e desvio padrão variando")
```

```
4- pesos<-
```

```
c(0.9475,0.9705,0.9770,0.9775,0.9860,0.9960,0.9965,0.9975,1.0050,1.0075,1.0100,1.0175,1.0180,1.0200,1.0250)
```

```
hist(pesos, freq=F, xlim=c(0.9,1.1))
```

```
curve(dnorm(x,mean=1,sd=0.05),col=4,lwd=2, add=TRUE)
```

```
curve(dnorm(x,mean=mean(pesos),sd=sd(pesos)),col=2,lwd=2, add=TRUE)
```

```
legend('topright',legend=c(expression(mi==1),expression(sigma==0.05),expression(mi==0.9968),expression(sigma==0.0217)),text.col=c(4,4,2,2),cex=.85)
```