CE-003: Estatística II - Turma: AMB, 1^a Prova (25/04/2012)

1. Uma série de características químicas foram medidas em diferentes vinhos. Os gráficos a seguir mostram quatro delas. Discuta os gráficos e suas interpretações utilizando conceitos e princípios de análise estatística descritiva/exploratória de dados. Inclua na sua discussão possíveis tratamentos dos dados.

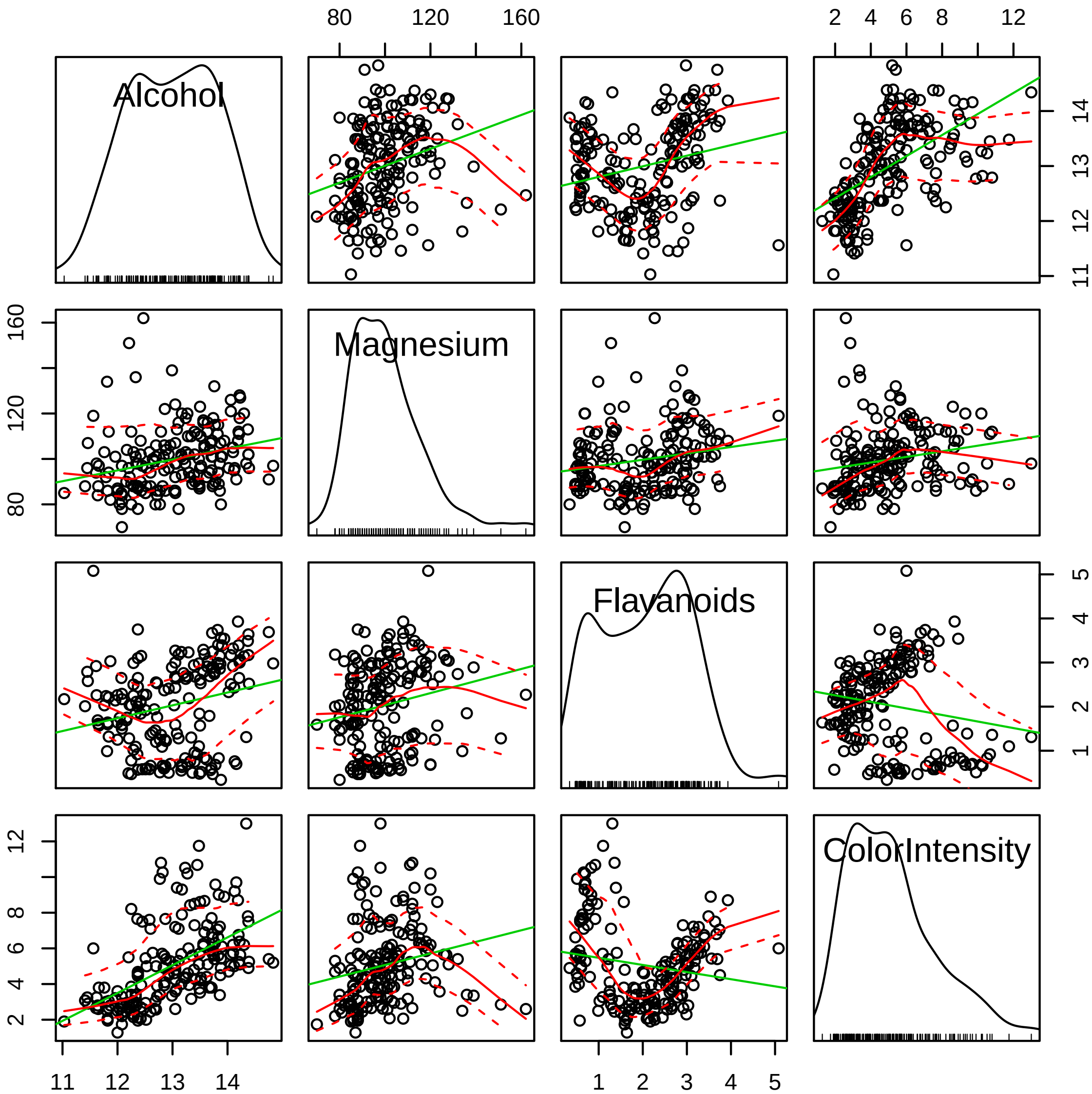


Figura 1: Algumas características de amostras de vinhos.

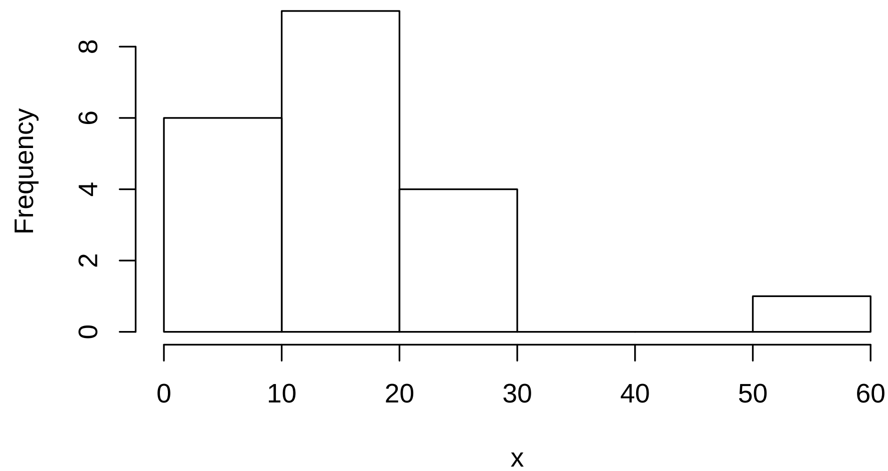
2. Foram registrados o tempo de execução (em segundos) de rotinas enviadas por vinte programadores.

```
10.4 13.8 51.0 17.6 18.5 23.8  5.2  9.3 11.5 22.7

18.2 10.5 24.4  2.2 28.4 11.3  6.4 14.0  6.7  8.9
```

(a) faça um histograma dos dados
(b) faça um diagrama ramo-e-folhas
(c) faça um gráfico *boxplot*
(d) obtenha a média e desvio padrão
(e) obtenha o coeficiente de variação
(f) obtenha a amplitude e a amplitude interquartílica
(g) caracterize/discuta a distribuição dos dados

(a) `> hist(x, main="")`

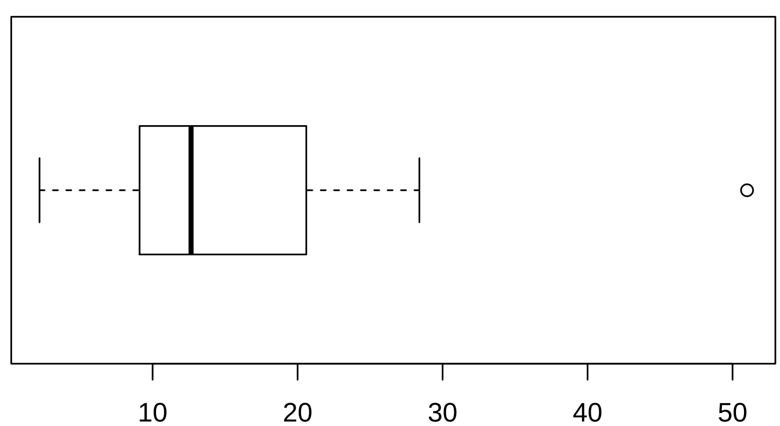


(b) `> stem(x)`

```
The decimal point is 1 digit(s) to the right of the |

0 | 256799
1 | 011244889
2 | 3448
3 |
4 |
5 | 1
```

(c) `> boxplot(x, horizontal=T)`



(d) obtenha a média e desvio padrão

```
> c(media=mean(x), desvio=sd(x))
```

```
 media desvio
 15.74  10.91
```

(e)
```
> 100*sd(x)/mean(x)
```

```
[1] 69.31
```

(f)
```
> range(x)
```

```
[1]  2.2 51.0
```

```
> range(fivenum(x)[c(2,4)])
```

```
[1]  9.1 20.6
```

(g) comentar sobre posição e dispersão, assimetria e presença de dados discrepantes.

---

3. Mostre que as funções a seguir são funções de densidade de probabilidade (f.d.p.) para algum valor de k e determine o valor de k.

- $f_1(x) = k_1(1+2x)$ para $0 < x < 2$.
- $f_2(x) = k_2x^2$ para $0 < x < 4$.

Obtenha para cada uma das distribuições:

- $F(x)$
- $\mathrm{E}[X]$
- $P[X > 1]$

- $\int_0^2 f_1(x)dx = 1 \longrightarrow \int_0^2 k_1(1+2x)dx = 1 \longrightarrow k_1 = 1/6$
- $\int_0^4 f_2(x)dx = 1 \longrightarrow \int_0^4 = k_2x^2dx = 1 \longrightarrow k_2 = 3/64$

Obtenha para cada uma das distribuições:

- $$F_1(x) = \int_0^x f_1(x)dx = (x+x^2)/6$$
  $$F_2(x) = \int_0^x f_2(x)dx = x^3/64$$
- $$E_1(x) = \int_0^2 xf_1(x)dx = 1.222$$
  $$E_2(x) = \int_0^4 xf_2(x)dx = 0.188$$
- $$P_1[X > 1] = 1 - F_1(1) = 0.667$$
  $$P_2[X > 1] = 1 - F_2(1) = 0.984$$

---

4. O tempo de montagem de um determinado mecanismo é uma v.a. uniforme mo intervalo de 30 a 40 segundos. Determine:

- as expressões de $f(x)$ e $F(x)$;
- a probabilidade da montagem ser feita em menos que 33 segundos;
- a probabilidade da montagem ser feita em menos que 38 segundos, sabendo que foi maior que 35 segundos;
- o tempo abaixo do qual 80% das montagens são feitas;
- o tempo esperado para montagem de 5.000 peças;

- o custo esperado para montagem de 5.000 peças sabendo que montagens abaixo de 33 segundos tem custo de R$1,00, entre 33 e 38 segundos o custo é de R$1,50 e acima de 38 segundos o custo é de R$3,00.

$$\begin{aligned}
X &: \text{ tempo de montagem} \\
X &\sim \mathrm{U}[30,40] \\
\text{a)} f(x) &= \frac{1}{40-30} = \frac{1}{10} \\
F(x) &= \frac{x-30}{40-30} = \frac{x-30}{10} \\
\text{b)} P[X<33] = F(33) &= \frac{33-30}{10} = 0,3 \\
\text{c)} P[X<38|X>35] &= \frac{P[35<X<38]}{P[X>35]} = \frac{F(38)-F(35)}{1-F(35)} = 0,6 \\
\text{d)} P[X<x] = 0,80 &\longrightarrow x = F^{-1}(0,80) \longrightarrow x = 38 \\
\text{e)} 5000 \cdot E[X] &= 5000 \cdot 35 = 170.000 \text{ segundos} = 48 \text{ hrs}, 36 \text{ min e } 40 \text{ seg}
\end{aligned}$$

f)Y : custo

Y	1,00	1,50	3,00
P[Y=y]	P[X<33]	P[33<X<38]	P[X>38]
P[Y=y]	0,3	0,5	0,2

$\text{CustoTotal} = 5000 \cdot E[Y] = \sum y_i P[Y=y_i] = 1,00 \cdot 0,3 + 1,50 \cdot 0,5 + 3,00 \cdot 0,2 = 8250$

---

5. Um lote de 120 containers de laranjas contém 25 que estão contaminados. São selecionados ao acaso e sequencialmente 15 containers para inspeção. Quais as probabilidades de que:

   (a) nenhum contaminado seja selecionado?

   (b) mais que 2 contaminados sejam encontrados?

   (c) de que o segundo inspecionado esteja contaminado sendo que o primeiro era contaminado?

   (d) de que os cinco primeiros não sejam contaminados?

$$\begin{aligned}
X &: \text{número de containers contaminados} \\
X &\sim \mathrm{HG}(N=120, n=15, r=25) \\
P[X=x] &= \frac{\binom{r}{x}\binom{N-r}{N-x}}{\binom{N}{n}} = \frac{\binom{25}{x}\binom{120-25}{15-x}}{\binom{120}{15}}
\end{aligned}$$

(a) $P[X=0] = \frac{\binom{25}{x}\binom{120-25}{15-x}}{\binom{120}{15}} = 0.023$

(b) $P[X>2] = 1 - P[X=0] - P[X=1] - P[X=2] = 0.647$

(c) $P[C_2|C_1] = \frac{24}{124} = 0.194$

(d) $P[\bar{C}_1 \cap \bar{C}_2 \cap \bar{C}_3 \cap \bar{C}_4 \cap \bar{C}_5] = P[\bar{C}_1] \cdot P[\bar{C}_2|\bar{C}_1] \cdot P[\bar{C}_3|\bar{C}_1\bar{C}_2] \cdot P[\bar{C}_4|\bar{C}_1\bar{C}_2\bar{C}_3] \cdot P[\bar{C}_5|\bar{C}_1\bar{C}_2\bar{C}_3\bar{C}_4] = \frac{95}{120} \cdot \frac{94}{119} \cdot \frac{93}{118} \cdot \frac{92}{117} \cdot \frac{91}{116} = 0.304$

ou $X \sim \mathrm{HG}(N=120, n=5, r=25) \longrightarrow P[X=0] = \frac{\binom{25}{x}\binom{120-25}{15-x}}{\binom{120}{5}} = 0.304$

---

6. Em um grupo de estudantes 45% são do curso A, 25% do curso B o restante do curso C. A proporção de mulheres em cada curso um dos cursos é de 20, 50 e 75%, respectivamente. Se um estudante é sorteado qual a probabilidade de:

   (a) seja homem;

   (b) seja do curso A, sabendo que foi sorteada uma mulher;

   (c) seja do curso C sabendo que foi sorteado um homem.

$$P[A] = 0,45 \ \ ; \ \ P[B] = 0,25 \ \ ; \ \ P[C] = 0,30$$

$$P[M|A] = 0,20 \ \ ; \ \ P[M|B] = 0,50 \ \ ; \ \ P[M|C] = 0,75$$

$$P[H|A] = 0,80 \ \ ; \ \ P[H|B] = 0,50 \ \ ; \ \ P[H|C] = 0,25$$

(a) $P[H] = P[H|A] \cdot P[A] + P[H|B] \cdot P[B] + P[H|C] \cdot P[C] = (0.2 \times 0.45) + (0.5 \times 0.25) + (0.75 \times 0.3) = 0.44$

(b) $P[A|M] = \frac{P[M|A] \cdot P[A]}{P[M]} = \frac{P[M|A] \cdot P[A]}{1 - P[H]} = \frac{0,45 \times 0,20}{1 - 0.44} = 0.161$

(c) $P[C|H] = \frac{P[H|C] \cdot P[C]}{P[H]} = \frac{0,25 \times 0,30}{0.44} = 0.17$