

CE-003: Estatística II - Turma K/O  
Avaliações Semanais (2º semestre 2014)

Semana 3 (av-01)

1. Um modelo simplificado do sistema de tipo sanguíneo humano possui quatro tipos de sangue:  $A$ ,  $B$ ,  $AB$  e  $O$ . Existem dois antígenos, anti- $A$  e anti- $B$  que reagem com o sangue de uma pessoa de diferentes formas de acordo com o tipo sanguíneo do indivíduo. Anti- $A$  reage com tipos sanguíneos  $A$  a  $AB$ , mas não com  $B$  e  $O$ . Anti- $B$  reage com tipos sanguíneos  $B$  a  $AB$ , mas não com  $A$  e  $O$ . Suponha que uma amostra de sangue de uma pessoa é coletada e testada com os dois antígenos. Denote por  $A$  o evento que o sangue reage com anti- $A$  e por  $B$  o evento que o sangue reage com anti- $B$ .
- (a) Classifique os possíveis tipos de sangue da pessoa usando a notação de eventos  $A$  e  $B$  e seus complementares. Suponha agora que, para uma determinada pessoa, a probabilidade de possuir o tipo  $O$  é de 0,50, a probabilidade do tipo  $A$  é 0,34 e a probabilidade do tipo  $B$  é de 0,12.
- (b) Encontre a probabilidade de que ambos antígenos sejam reagentes com o sangue da pessoa.
- (c) Encontre a probabilidade de que cada um dos antígenos vá reagir com o sangue da pessoa.
- (d) Sabendo que o sangue de uma pessoa reagiu com anti- $A$ , qual a probabilidade de que a pessoa tenha sangue do tipo  $AB$ ?
- (e) Os eventos  $A$  e  $B$  são mutuamente exclusivos? Justifique.
- (f) Os eventos  $A$  e  $B$  são independentes? Justifique.

**Solução:**

Notação:

$A$  : a amostra reage a anti- $A$

$B$  : a amostra reage a anti- $B$

(a)

Tabela 1: Tipos sanguíneos, eventos que os definem e probabilidades de cada tipo.

Tipo sanguíneo	Evento	Probabilidade
$O$	$A^c \cap B^c$	0,50
$A$	$A \cap B^c$	0,34
$B$	$A^c \cap B$	0,12
$AB$	$A \cap B$	0,04

(b)

$$P[AB] = P[A \cap B] = 1 - P[O] - P[A] - P[B] = 0,04$$

(c)

$$P[A] = P[A \cap B^c] + P[A \cap B] = 0,34 + 0,04 = 0,38$$
$$P[B] = P[A^c \cap B] + P[A \cap B] = 0,12 + 0,04 = 0,16$$

(d)

$$P[(A \cap B)|P(A)] = \frac{P[(A \cap B) \cap (A)]}{P[A]} = \frac{P[(A \cap B)]}{P[A]} = \frac{0,04}{0,38} = 2/19 = 0.105$$

(e) Não, pois  $P[A \cap B] = 0,04 \neq 0$

(f) Não, pois  $P[A \cap B] = 0,04 \neq P[A] \cdot P[B] = 0,38 \cdot 0.16 = 0.0608$

Semana 4 (av-02)

1. Três algoritmos diferentes são utilizados, de forma independente, para tentar encontrar a solução ótima de um problema. Sabe-se que o primeiro algoritmo tem uma taxa de acerto de 75%, o segundo tem 60% e o terceiro tem 85%.

- Qual a probabilidade do problema ser resolvido?
- Qual a probabilidade de que a solução ótima seja encontrada por mais de um algoritmo?
- Qual a probabilidade de não seja encontrada a solução ótima de um problema submetido aos três algoritmos?

Procure usar a notação de eventos e operação de eventos no desenvolvimento de sua solução.

### Solução:

(a)

$A$  : o primeiro resolve o problema  $P(A) = 0,75$   $P(\bar{A}) = 0,25$

$B$  : o segundo resolve o problema  $P(B) = 0,60$   $P(\bar{B}) = 0,40$

$C$  : o terceiro resolve o problema  $P(C) = 0,85$   $P(\bar{C}) = 0,15$

$$P(A \cup B \cup C) = 1 - P(\bar{A} \cap \bar{B} \cap \bar{C}) \stackrel{ind}{=} 1 - P(\bar{A}) \cdot P(\bar{B}) \cdot P(\bar{C}) = 1 - (1 - 0,75)(1 - 0,60)(1 - 0,85) = 0,985$$

(b)  $P(A \cap B \cap \bar{C}) + P(A \cap \bar{B} \cap C) + P(\bar{A} \cap B \cap C) + P(A \cap \bar{B} \cap C) \stackrel{ind}{=} P[A] \cdot P[B] \cdot P[\bar{C}] + P[A] \cdot P[\bar{B}] \cdot P[C] + P[\bar{A}] \cdot P[B] \cdot P[C] + P[A] \cdot P[\bar{B}] \cdot P[C] = 0,75 \cdot 0,60 \cdot 0,15 + 0,75 \cdot 0,40 \cdot 0,85 + 0,25 \cdot 0,60 \cdot 0,85 + 0,75 \cdot 0,60 \cdot 0,85 = 0,832$

(c)  $P(\bar{A} \cap \bar{B} \cap \bar{C}) \stackrel{ind}{=} P(\bar{A}) \cdot P(\bar{B}) \cdot P(\bar{C}) = (1 - 0,75)(1 - 0,60)(1 - 0,85) = 0,015$

2. Sabe-se a partir de históricos que 25% das requisições de transferências de dados em um sistema são recusadas. Há interesse em se estudar possíveis padrões de falhas o que é feito contando-se o número de falhas consecutivas até o sucesso no envio.

- Qual a probabilidade de que ocorram três ou mais falhas consecutivas em um tentativa de envio?
- Identifique neste problema a variável aleatória de interesse e seus possíveis valores
- Mostre as probabilidades associadas a alguns dos valores desta variável.
- Identifique uma equação que forneça os valores das probabilidades para os diferentes valores da variável.

### Solução:

Eventos:

$S$  : sucesso na transferência

$F \equiv \bar{S}$  : falha na transferência

**OBS:** supõem-se independência entre as tentativas.

(a)

$$P[3 \text{ ou mais falhas}] = 1 - P[2 \text{ ou menos falhas}] = 1 - \{P[0 \text{ falhas}] + P[1 \text{ falha}] + P[2 \text{ falhas}]\} \\ = 1 - \{P[S] + P[F \cap S] + P[F \cap F \cap S]\} = 1 - [0,75 + 0,25 \cdot 0,75 + 0,25^2 \cdot 0,75] = 0,0156$$

(b)

$X$  : número de falhas até conseguir o envio

$$x \in \{0, 1, 2, 3, 4, \dots\}$$

(c) 

$x$	0	1	2	3	...
$P[X=x]$	0,75	0,25 · 0,75	0,25 <sup>2</sup> · 0,75	0,25 <sup>3</sup> · 0,75	...

(d)  $P[X = x] = 0,25^x \cdot 0,75$

De forma mais geral a variável segue a distribuição chamada de *geométrica* com probabilidade  $p$  de sucesso:

$X$  : número de “falhas” até o primeiro “sucesso”

$$x \in \{0, 1, 2, 3, 4, \dots\}$$

$$X \sim G(p)$$

$$P[X = x] = (1 - p)^x p$$

### Semana 5 (av-03)

1. Seja a função de densidade de probabilidade dada por  $f(x) = Cx^2 I_{[0,4]}(x)$ . Obtenha:

(a) o valor de  $C$ ,

(b)  $P[X > 0,5]$ ,

(c)  $P[X > 0,7 | X > 0,5]$ ,

(d) Calcule o valor que resulta de  $\int_0^4 x \cdot f(x) dx$ . Você consegue associar alguma interpretação a este valor?

(e) Calcule o valor de  $x$  para o qual  $P[X < x] = P[X > x] = 0,5$ .

**Solução:**

$$f(x) = Cx^2 I_{[0,4]}(x)$$

(a)

$$\int_0^4 f(x) dx = 1$$

$$\int_0^4 Cx^2 dx = 1$$

$$C = \frac{1}{\int_0^4 x^2 dx}$$

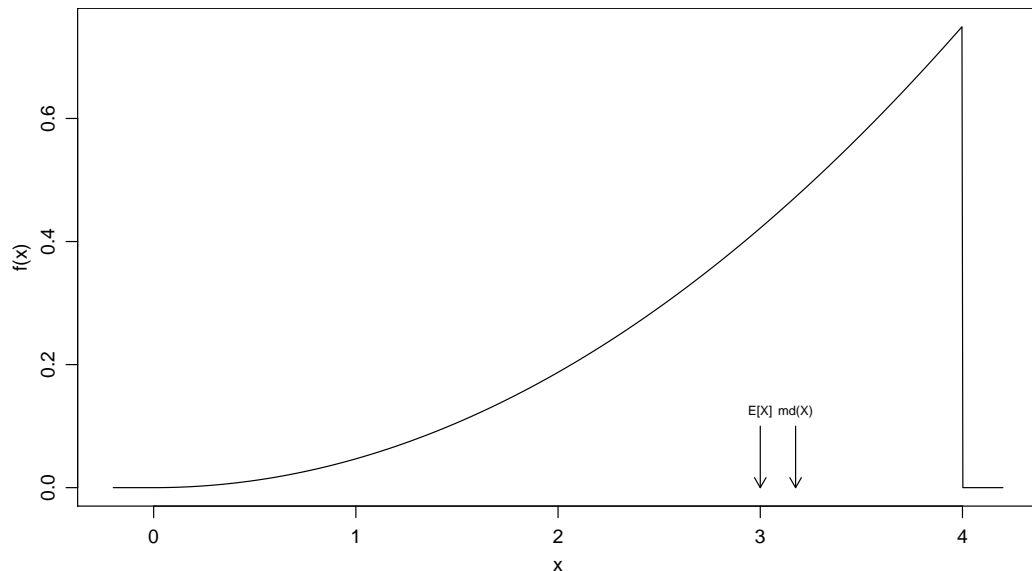
$$C = 3/64$$

$$(b) P[X > 0,5] = \int_{0,5}^4 Cx^2 dx = 0.998$$

$$(c) P[X > 0,7 | X > 0,5] = \frac{P[X > 0,7 \cap X > 0,5]}{P[X > 0,5]} = \frac{P[X > 0,7]}{P[X > 0,5]} = \frac{\int_{0,7}^4 Cx^2 dx}{\int_{0,5}^4 Cx^2 dx} = 0.997$$

$$(d) E(X) = \int_0^4 x \cdot f(x) dx = (3/64) \int_0^4 x^3 dx = \dots = 3$$

$$(e) \text{md}(x) : \int_0^{\text{md}(x)} f(x) dx = 0,50 \rightarrow \text{md}(x) = (64 \cdot 0,50)^{1/3} = 3.17$$



## Semana 6 (av-04)

1. Nos problemas a seguir identifique a variável aleatória, seus possíveis valores, a distribuição de probabilidades e calcule as quantidades pedidas, fazendo suposições adequadas quando necessário.
  - (a) Sabe-se que em um determinado lote de 50 peças de *hardware* 15 delas são defeituosas. Se três peças serão escolhidas ao acaso, qual a probabilidade de que (i) nenhuma seja defeituosa? (ii) todas sejam defeituosas?
  - (b) Em um serviço de televendas potenciais clientes são abordados sequencialmente e 12% das chamadas resultam em uma compra. Qual a probabilidade de se conseguir fazer duas vendas em dois contatos consecutivos? (ou seja, feita uma venda, qual a probabilidade de fazer uma outra venda na tentativa seguinte?)
  - (c) Considerando ainda este serviço de televendas, qual a probabilidade de se efetuar duas ou mais vendas em dez contatos com clientes?
  - (d) Um servidor web recebe em média 7,4 requisições por hora. (i) Qual a probabilidade de que este servidor receba mais do que cinco requisições em uma determinada hora? (ii) Qual a probabilidade de que o servidor não receba requisições em um intervalo qualquer de 15 minutos?

(a)

$X$  : número de defeituosas entre as 3 peças retiradas

$$x \in \{0, 1, 2, 3\}$$

$$X \sim \text{HG}(N = 50, n = 3, k = 15)$$

$$P(X = x) = \frac{\binom{15}{x} \binom{35}{3-x}}{\binom{50}{3}}$$

$$P[X = 0] = 0.334$$

$$P[X = 3] = 0.0232$$

(b)

$X$  : número de não vendas após efetuar uma venda

$$x \in \{0, 1, 2, \dots\}$$

$$X \sim G(p = 0, 12)$$

$$P(X = x) = (1 - 0, 12)^x \cdot 0, 12$$

$$P[X = 0] = 0.12$$

(c)

$X$  : número de vendas em 10 contatos

$$x \in \{0, 1, 2, \dots, 10\}$$

$$X \sim B(n = 5, p = 0, 12)$$

$$P(X = x) = \binom{10}{x} 0, 12^x 0, 88^{10-x}$$

$$P[X \geq 2] = 1 - P[X = 0] - P[X = 1] = 0.342$$

(d)

$X$  : número de requisições

$$x \in \{0, 1, 2, \dots\}$$

$$X_i \sim P(\lambda = 7, 4)$$

$$P(X_i = x) = \frac{e^{-7,4} 7, 4^x}{x!}$$

$$P[X_i > 5] = 1 - P[X_i \leq 5] = 0.608$$

$$X_{ii} \sim P(\lambda = 7, 4/4)$$

$$P(X_{ii} = x) = \frac{e^{-7,4/4} (7, 4/4)^x}{x!}$$

$$P[X_{ii} = 0] = 0.157$$

Soluções computacionais (em linguagem R)

```
> qa1 <- dhyper(0, m=15, n=35, k=3)
> qa2 <- dhyper(3, m=15, n=35, k=3)
> qb <- dgeom(0, prob=0.12)
> qc <- pbinom(1, size=10, prob=0.12, lower=FALSE)
> qd1 <- ppois(6, lambda=7.4, lower=FALSE)
> qd2 <- dpois(0, lambda=7.4/4)
```

---

Semana 7 (av-05)

1. Assume-se que o tempo entre acessos a um blog tem uma distribuição com média de 1,5 segundos. Assumindo a distribuição exponencial responda os itens a seguir.
  - (a) Quais as características do problema que fazem com que a distribuição exponencial seja uma escolha razoável?
  - (b) Qual a probabilidade de haver duas conexões com intervalo inferior a 1,5 segundos?
  - (c) Qual a probabilidade de se passarem 5 segundos sem conexão alguma?
  - (d) Tendo havido uma conexão, qual a probabilidade da próxima conexão ocorrer entre 0,5 e 2,5 segundos?
  - (e) Se já se passou 1 segundo sem conexão, qual a probabilidade de se passar mais 0,5 segundos adicionais sem conexão?
  - (f) Qual a probabilidade do intervalo entre conexões não superar 3,5 segundos se já se passaram 2 segundos sem conexão?

**Solução:** $X$  : intervalo de tempo entre conexões (segundos) $X \sim \text{Exp}(\lambda = 1/1, 5 = 2/3)$ 

$$f(x) = \frac{2}{3}e^{-2x/3} \mathbb{I}_{(0,\infty)}(x) \quad F(x) = 1 - e^{-2x/3}$$

(a) A distribuição exponencial é razoável considerando-se que: (i) que devem ser valores positivos, (ii) pela possibilidade de cálculos com as informações fornecidas, (iii) espera-se que a probabilidade de intervalo entre conexões seja maior para pequenos intervalos e torne-se menor com o aumento do intervalo de tempo.

(b)  $P[X < 1, 5] = \int_0^{1,5} f(x)dx = F(1, 5) = 0.63$

(c)  $P[X > 5] = \int_5^{\infty} f(x)dx = 1 - F(5) = 0.036$

(d)  $P[0, 5 < X < 2, 5] = \int_{0,5}^{2,5} f(x)dx = F(2, 5) - F(0, 5) = 0.53$

(e)  $P[X > 1, 5 | X > 1] = \frac{\int_{1,5}^{\infty} f(x)dx}{\int_1^{\infty} f(x)dx} \stackrel{\bar{1}}{=} P[X > 0, 5] = 1 - F(0, 5) = 0.72$

(f)  $P[X < 3, 5 | X > 2] = \frac{\int_2^{3,5} f(x)dx}{\int_{2}^{\infty} f(x)dx} = \frac{F(3,5) - F(2)}{1 - F(2)} \stackrel{\bar{1}}{=} P[X < 1, 5] = F(1, 5) = 0.63$

2. O peso de um tênis de corrida sofisticado é normalmente distribuído com média de 12 onças (onça é uma unidade de peso) e desvio padrão de 0,5 onças.

(a) qual a probabilidade de um tênis pesar mais que 13,2 onças?

(b) qual a probabilidade de um tênis pesar entre 11,6 e 12,7 onças?

(c) quanto deveria ser o desvio padrão para que 99,9% dos tênis tenham menos que 13 onças?

(d) se o desvio padrão se mantiver em 0,5, quanto deveria ser a média para que 99,9% dos tênis tenham menos que 13 onças?

**Solução:**

$$X \sim N(12, 0.5^2)$$

(a)  $P[X > 13, 2] = P[Z > \frac{13,2-12}{0,5}] = P[Z > 2.4] = 0.0082$

(b)  $P[11, 6 < X < 12, 7] = P[\frac{11,6-12}{0,5} < Z < \frac{12,7-12}{0,5}] = P[-0.8 < Z < 1.4] = 0.707$

(c)

$$X \sim N(12, \sigma^2)$$

$$P[X < 13] = 0,999 \quad ; \quad \sigma = ?$$

$$P[Z < \frac{13 - 12}{\sigma}] = 0,999$$

$$z = 3.09$$

$$\frac{13 - 12}{\sigma} = 3.09$$

$$\sigma = 0.324$$

(d) se o desvio padrão se mantiver em 0,5, quanto deveria ser a média para que 99,9% dos tênis tenham menos que 13 onças?

$$X \sim N(\mu, 0.5^2)$$

$$P[X < 13] = 0,999 \quad ; \quad \mu = ?$$

$$P[Z < \frac{13 - \mu}{0,5}] = 0,999$$

$$z = 3.09$$

$$\frac{13 - \mu}{0,5} = 3.09$$

$$\mu = 13 - 0,5(3.09)$$

$$\mu = 11.455$$

1. Os dados a seguir são das notas obtidas por um grupo de estudantes em uma disciplina. Com estes dados obtenha as análises pedidas a seguir.

61 77 51 29 55 77 33 70 56 41 61 28 87 23 22 86 63 99 38 25 90  
 59 87 53 85 86 87 75 50 59 77 77 71 99 78 70 93 78 93 94

- Agrupe os dados em classes e obtenha uma tabela com frequências absolutas e relativas.
- Faça um histograma das observações.
- Calcule a média e mediana *a partir dos dados originais*.
- Calcule a média e mediana *a partir dos dados agrupados na tabela de frequências*.
- Existem diferenças entre os resultados dos dois ítems anteriores? Justifique.
- Calcule (usando os dados originais) ao menos duas medidas de dispersão dos dados.
- Faça um diagrama ramo-e-folhas dos dados.
- Descreva textualmente em um parágrafo o desempenho do grupo, baseando-se nas análises dos dados.

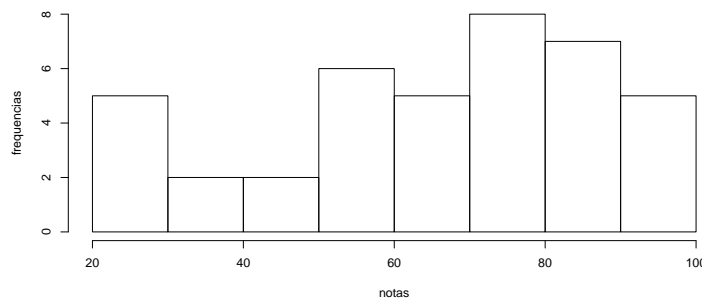
Considere agora que os dados na primeira linha são da *TURMA – A* enquanto os da segunda são da *TURMA – B*. Faça análises baseadas em gráficos e medidas que permitam comparar os desempenhos das duas turmas. Discuta os resultados destacando e comparando as características do desempenho dos dois grupos.

**Solução:**

- (a)

	Freq	FreqAc	FreqRel	FreqRelAc
[20,30]	5	5	0.12	0.12
(30,40]	2	7	0.05	0.17
(40,50]	2	9	0.05	0.23
(50,60]	6	15	0.15	0.38
(60,70]	5	20	0.12	0.50
(70,80]	8	28	0.20	0.70
(80,90]	7	35	0.17	0.88
(90,100]	5	40	0.12	1.00

- (b)



- (c)

$$\bar{x} = 66.1 \quad ; \quad \text{md}(x) = 70.5$$

- (d)

$$\bar{x}_{ag} = 65.2 \quad ; \quad \text{md}(x) = 70$$

- (e) Sim, pode haver, devido ao *erro de agrupamento*, ou seja, no segundo caso considera-se que a média dos dados de cada classe é igual ao ponto médio da classe, o que pode não ser verdadeiro.

- (f)

$$\text{Amplitude} = x_{max} - x_{min} = 77 \quad ; \quad \text{Amplitude Interquartílica} = Q_3 - Q_1 = 34$$

Outras medidas:

$$\text{Variância} = S^2 = 66.1 \quad ; \quad \text{desvio padrão} = S = 19.1 \quad ; \quad \text{desvio médio} = DM = 66.1$$

- (g) `> stem(notas)`

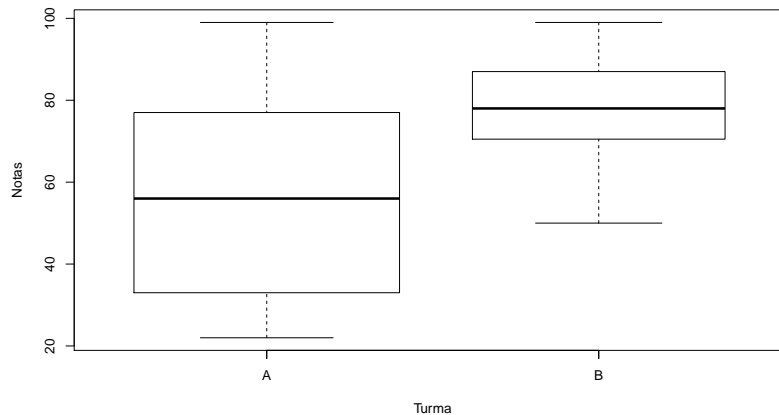
The decimal point is 1 digit(s) to the right of the |

2 | 23589  
 3 | 38  
 4 | 1  
 5 | 0135699  
 6 | 113  
 7 | 0015777788  
 8 | 566777  
 9 | 033499

(h) ...

Resultados para comparar os grupos:

	Media	Min	Q1	Md	Q3	Max	S2	S	CV	DM
1	55.81	22.00	33.00	56.00	77.00	99.00	593.26	24.36	43.64	20.29
2	77.42	50.00	70.50	78.00	87.00	99.00	203.92	14.28	18.44	11.14



Semana 11 (av-07)

- Um conjunto de dados (*Womentf*) disponível no pacote **car** do programa estatístico **R** possui registros de condições relacionadas ao trabalho de 263 mulheres no Canadá. Os atributos são: o tipo de trabalho (externo em tempo integral *fulltime*, externo em tempo parcial *parttime*, não trabalha fora de casa *not.work*), o salário do marido (*hincome*), se possui ou não filhos (com filho(s) *present*, sem filho(s) *absent*), a região do país (dividido em cinco regiões). Uma última coluna na tabela de dados apresenta uma categorização do salário do marido. Abaixo são mostrados: um extrato da tabela com os 10 primeiros registros e um resumo univariado de todos os 263 dados.

Discuta quais seriam as possíveis questões de interesse com este conjunto de dados relacionando (ao menos) duas variáveis. Verifique, interprete e discuta os resultados fornecidos a seguir.

```

partic hincome children region classHI
1 not.work 15 present Ontario (10,15]
2 not.work 13 present Ontario (10,15]
3 not.work 45 present Ontario (20,50]
4 not.work 23 present Ontario (20,50]
5 not.work 19 present Ontario (15,20]
6 not.work 7 present Ontario (5,10]
7 not.work 15 present Ontario (10,15]
8 fulltime 7 present Ontario (5,10]
9 not.work 15 present Ontario (10,15]
10 not.work 23 present Ontario (20,50]

```

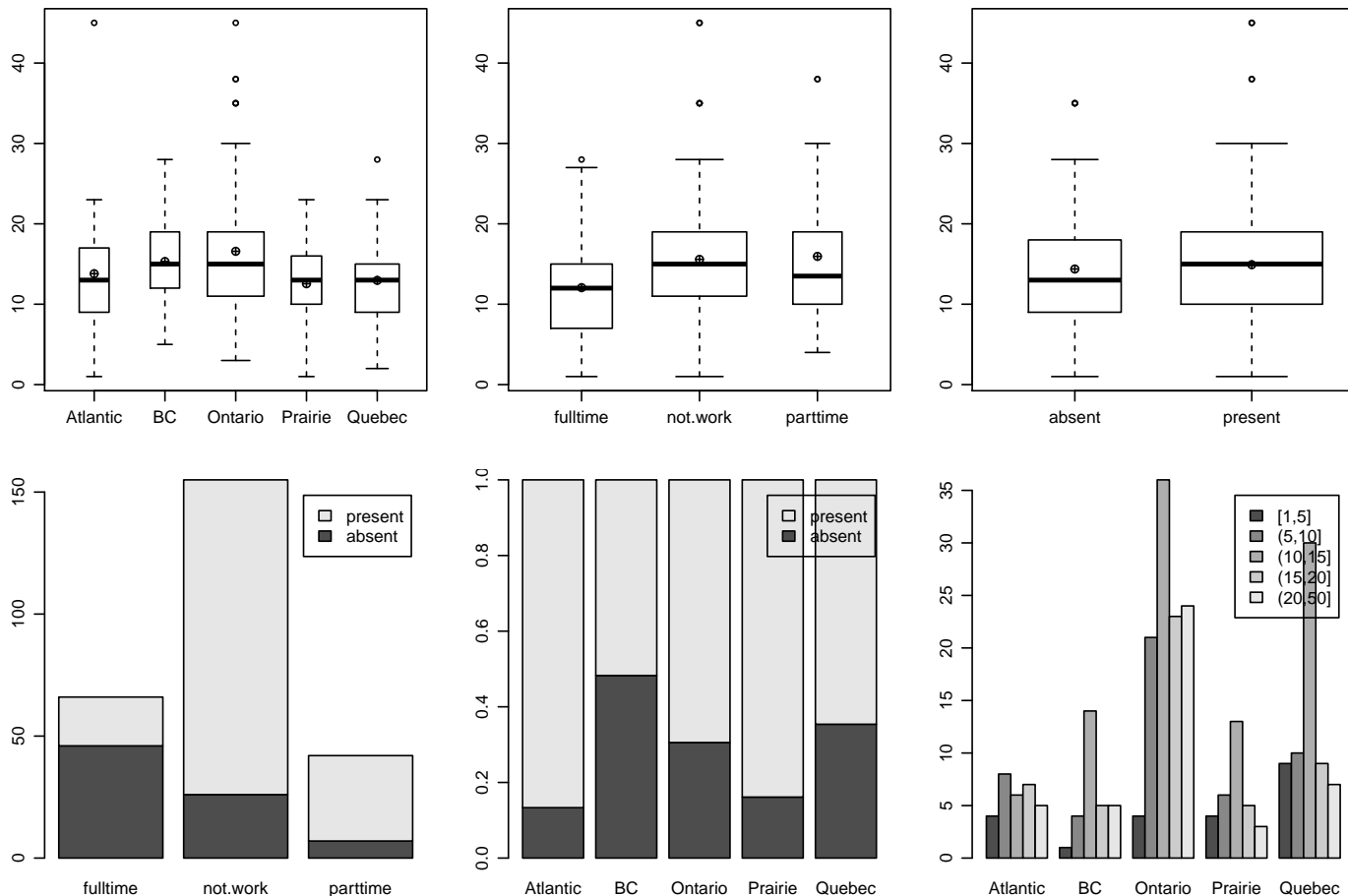
```

partic      hincome      children      region      classHI
fulltime: 66  Min.    : 1.0  absent : 79  Atlantic: 30  [1,5] :22
not.work:155  1st Qu.:10.0  present:184  BC       : 29  (5,10] :49
parttime: 42  Median :14.0  Ontario :108 (10,15] :99
              Mean  :14.8  Prairie  : 31  (15,20] :49
              3rd Qu.:19.0  Quebec  : 65  (20,50] :44
              Max.  :45.0

```

	fulltime	not.work	parttime	Sum
absent	46.00	26.00	7.00	79.00
present	20.00	129.00	35.00	184.00
Sum	66.00	155.00	42.00	263.00

	Atlantic	BC	Ontario	Prairie	Quebec	Sum
absent	0.05	0.18	0.42	0.06	0.29	1.00
present	0.14	0.08	0.41	0.14	0.23	1.00
Sum	0.19	0.26	0.83	0.20	0.52	2.00



Pearson's Chi-squared test

```
data: with(Womenlf, table(partic, region))
```

X-squared = 5.438, df = 8, p-value = 0.7098

Pearson's Chi-squared test

```
data: with(Womenlf, table(children, partic))
```

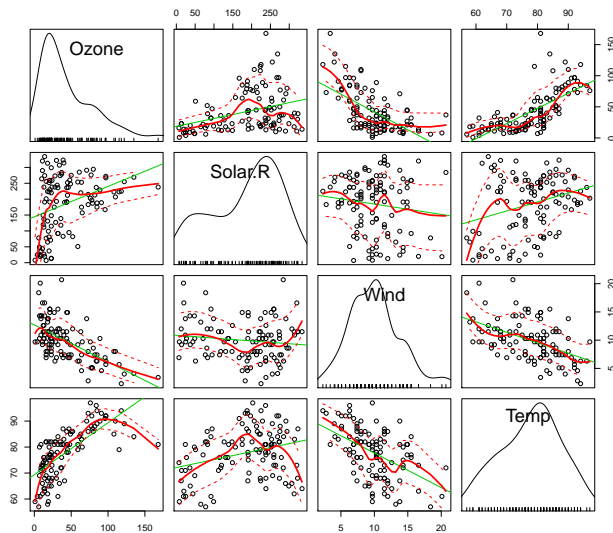
X-squared = 65.95, df = 2, p-value = 4.788e-15

## Semana 12 (av-08)

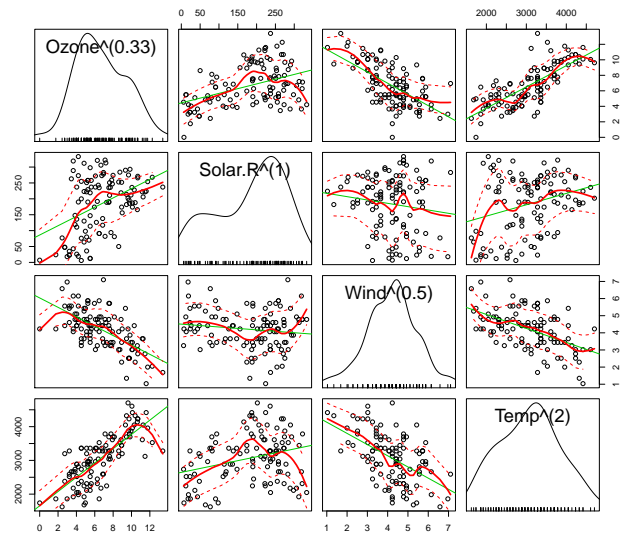
1. A figura a seguir mostra relações de medidas diárias de qualidade do ar em Nova York coletadas entre Maio e Setembro de 1973. Foram medidos: nível de **Ozônio** (*Ozone*), a **radiação solar** (*Solar.R*), a velocidade do **vento** (*Wind*) e a **temperatura** (*Temp*). Discuta as características de cada variável individualmente e relação das variáveis tomadas duas a duas, indicando como qual(ais) medida(s) pode(m) ser calculada(s) para melhor refletir a associação em cada caso. Aponte quais seriam as relações (duas a duas) mais "fortes" e mais "fracas" entre as variáveis
2. Descreva uma situação (hipotética/inspirada em um estudo real) em que uma amostra é tomada de uma população para avaliar uma característica de interesse. Neste contexto especifique: a população ("informal" e no sentido estatístico), o(s) atributo(s) de interesse, o parâmetro de interesse, a forma como a amostra poderia ser tomada, se a forma de amostragem sugerida constitui uma amostra aleatória simples, algum estimador que poderia ser utilizado.



	[1,5]	(5,10]	(10,15]	(15,20]	(20,50]	Sum
Atlantic	0.18	0.16	0.06	0.14	0.11	0.66
BC	0.05	0.08	0.14	0.10	0.11	0.48
Ontario	0.18	0.43	0.36	0.47	0.55	1.99
Prairie	0.18	0.12	0.13	0.10	0.07	0.61
Quebec	0.41	0.20	0.30	0.18	0.16	1.26
Sum	1.00	1.00	1.00	1.00	1.00	5.00



(a) Dados originais



(b) Dados transformados

Semana 14 (av-09)

- (B. & M.) Se uma amostra com 36 observações for tomada de uma população. qual deve ser o tamanho de uma outra amostra para que o desvio padrão da média desta amostra seja 2/3 do desvio padrão da média da primeira?

População:  $X \sim (E[X] = \mu, \text{Var}[X] = \sigma^2)$

Amostra 1:  $\bar{X}_1 \approx N(E[\bar{X}_1] = \mu, \text{Var}[\bar{X}_1] = \sigma^2/36)$

Amostra 2:  $\bar{X}_2 \approx N(E[\bar{X}_2] = \mu, \text{Var}[\bar{X}_2] = \sigma^2/n_2)$

$$\frac{\sigma}{\sqrt{n_2}} = \frac{2}{3} \frac{\sigma}{\sqrt{36}}$$

$$n_2 = 81$$

- Em um experimento foram feitas 24 tentativas (independentes) de resolver um problema até que a terceira solução correta fosse encontrada. Encontre os estimadores de (i) métodos dos momentos; (ii) de máxima verossimilhança do parâmetro  $p$  que representa a probabilidade de obter a solução correta em uma tentativa qualquer.

$X$  : número de "falhas" até obter a terceira solução correta

$X : \sim \text{BN}(r = 3, p)$

$$P[X = x] = \binom{x+r-1}{r-1} p^3 (1-p)^x$$

amostra :  $x = 21$

método dos momentos :

$$E[X] = r \frac{(1-p)}{p} = \bar{x}$$

$$3 \frac{(1-\hat{p})}{\hat{p}} = 21$$

$$\hat{p} = \frac{3}{24} = \frac{1}{8} = 0.125$$

método da máxima verossimilhança :

$$L(p) = \binom{31+3-1}{3-1} p^3 (1-p)^{21}$$

$$l(p) = \log[L(p)] = \log \left( \binom{31+3-1}{3-1} \right) + 3 \log(p) + 21 \log(1-p)$$

$$\frac{dl(p)}{dp} = \frac{3}{(\hat{p})} + \frac{21}{1-\hat{p}} (-1) = 0$$

$$\hat{p} = \frac{3}{24} = \frac{1}{8} = 0.125$$

---