

CE-003: Estatística II - Turma: AMB, Avaliações Semanais 1º semestre/2012

- Foram feitas medições dos teores de um poluente em duas regiões (A e B), representadas nos gráficos da figura a seguir.
 - Indique qual *boxplot* da figura à direita correspondente cada curva da figura à esquerda. Justifique sua resposta.
 - Em uma das regiões a média foi de 44.6 e a mediana 40.6, enquanto que em outra a média foi 49.5 e a mediana 49.2. Quais valores correspondem a cada região? Justifique sua resposta.
 - Interprete e discuta cada um dos gráficos, comparando as regiões.

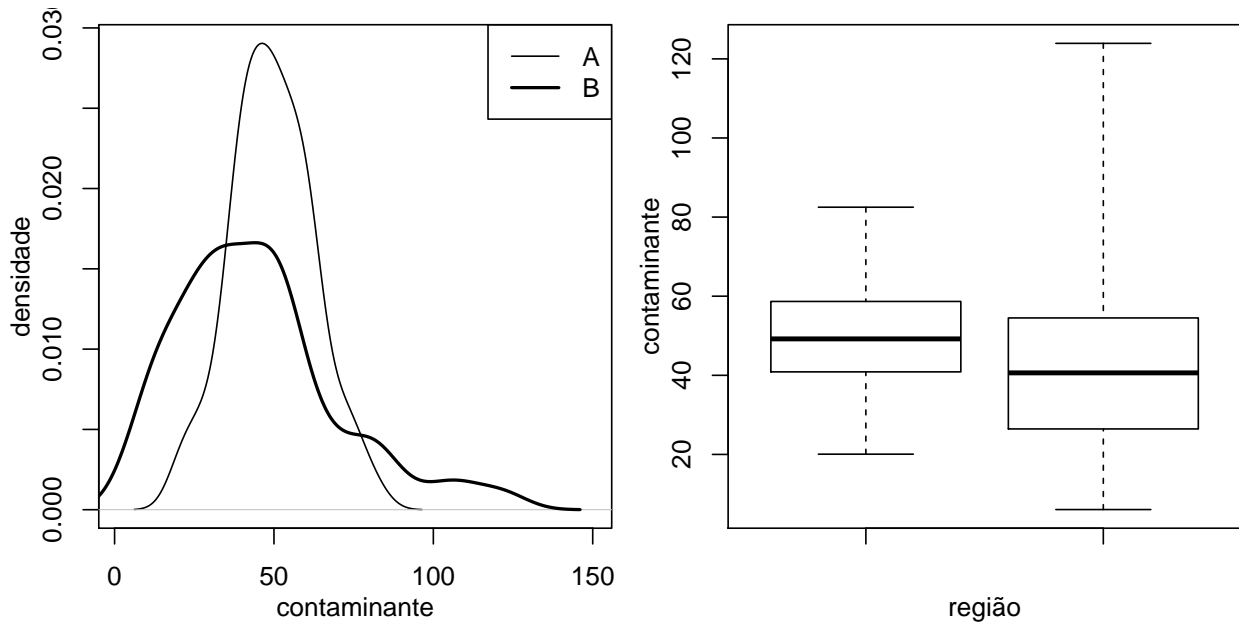


Figura 1: Teores de poluente medidos em amostras tomadas em duas regiões.

- Foram feitas medições de índices de qualidade da água em 20 locais e os dados coletados foram:

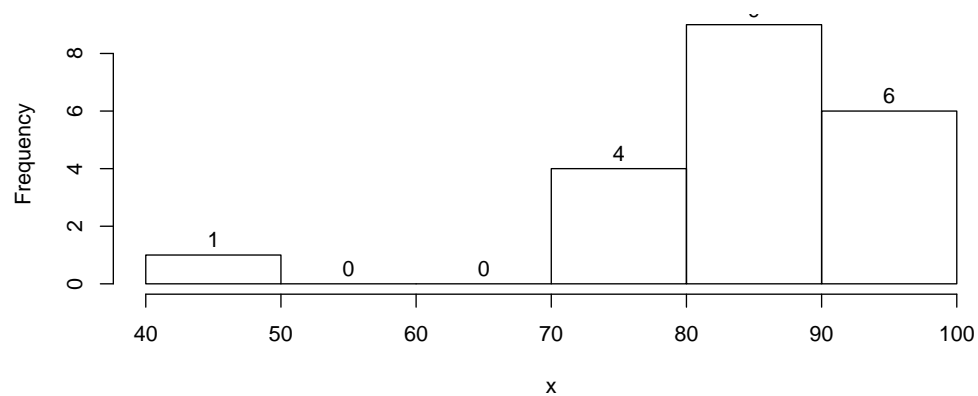
89.6 86.2 49.0 82.4 81.5 76.2 94.8 90.7 88.5 77.3

81.8 89.5 75.6 97.8 71.6 88.7 93.6 86.0 93.3 91.1

- faça um histograma dos dados
- faça um diagrama ramo-e-folhas
- faça um gráfico *boxplot*
- obtenha a média e desvio padrão
- obtenha o coeficiente de variação
- obtenha a amplitude e a amplitude interquartílica
- caracterize a distribuição dos dados

Solução:

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

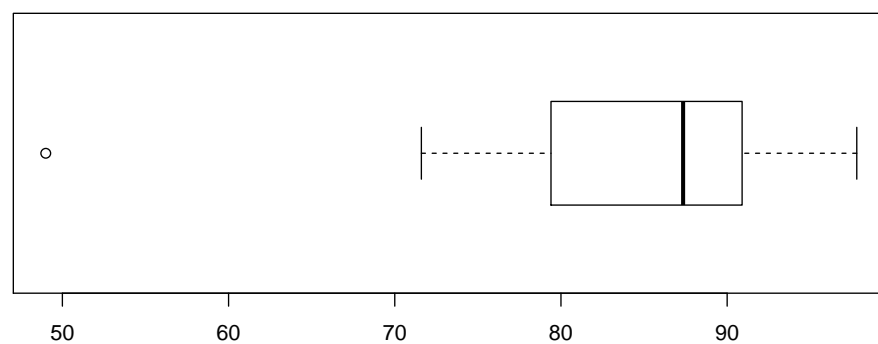


(b) `> stem(x)`

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

```
4 | 9
5 |
6 |
7 | 2667
8 | 2226699
9 | 00113458
```

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



(d) `> c(media=mean(x), desvioP = sd(x))`

```
media desvioP
84.26  10.91
```

(e) obtenha o coeficiente de variação

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

```
[1] 12.95
```

(f) `> range(x) ; diff(range(x))`

```
[1] 49.0 97.8
```

[1] 48.8

> *fivenum(x)[c(2,4)]; diff(fivenum(x)[c(2,4)])*

[1] 79.4 90.9

[1] 11.5

(g) Comentar sobre: posição, variabilidade, assimetria e dados discrepantes

3. Um estudo procurou relacionar medidas de um índice de poluição (PM10) com atendimentos hospitalares por doenças respiratórias. Foram anotados dados em vários períodos e em cinco capitais.

Discuta estratégias para investigar a relação desejada a partir dos dados. Mencione que tipos de análises estatísticas descritivas poderiam ser feitas, os possíveis cenários (resultados) e como seriam interpretados. Comente sobre o que deveria ser levado em consideração nas análises.

4. (B. & M.) Um empreiteiro apresentou orçamentos separados para a execução da parte elétrica e da parte de encanamento de um edifício. Ele acha que a probabilidade de ganhar e concorrência da parte elétrica é de $1/2$. Caso ganhe a parte elétrica, a chance de ganhar a parte de encanamento é de $3/4$; caso contrário, essa probabilidade é de $1/3$.

(a) Qual a probabilidade de ele:

- ganhar os dois contratos?
- ganhar apenas um dos contratos?
- não ganhar nenhum contrato?

(b) os eventos "ganhar o contrato elétrico" e "ganhar o contrato hidráulico"

- são independentes? (justifique)
- são mutuamente exclusivos? (justifique)

Solução:

G_1 : ganhar concorrência da parte elétrica

G_2 : ganhar concorrência do encanamento

$$P[G_1] = 1/2 \quad ; \quad P[G_2|G_1] = 3/4 \quad ; \quad P[G_2|\bar{G}_1] = 1/3$$

$$P[\bar{G}_1] = 1/2 \quad ; \quad P[\bar{G}_2|G_1] = 1/4 \quad ; \quad P[\bar{G}_2|\bar{G}_1] = 2/3$$

- (a)
- $P[G_1 \cap G_2] = P[G_1] \cdot P[G_2|G_1] = (1/2) \cdot (3/4) = 3/8 = 0.375$
 - $P[G_1 \cap \bar{G}_2] + P[\bar{G}_1 \cap G_2] = P[G_1] \cdot P[\bar{G}_2|G_1] + P[\bar{G}_1] \cdot P[G_2|\bar{G}_1] = (1/2) \cdot (1/4) + (1/2) \cdot (1/3) = 7/24 = 0.292$
 - $P[\bar{G}_1 \cap \bar{G}_2] = P[\bar{G}_1] \cdot P[\bar{G}_2|\bar{G}_1] = (1/2) \cdot (2/3) = 1/3 = 0.333$
- (b)
- Não, pois $P[G_1 \cap G_2] = 3/8 \neq P[G_1] \cdot P[G_2] = 13/48$,
em que $P[G_2] = P[G_2 \cap G_1] + P[G_2 \cap \bar{G}_1] = (3/8) + (1/6) = 13/24$
 - Não pois $P[G_1 \cap G_2] \neq 0$

5. Considere o lançamento de uma moeda 10 vezes.

(a) Se voce lançar a probabilidade de obter a face "cara" em todos os lançamentos?

(b) Considere agora que 1.000 pessoas fazem o mesmo. Qual a probabilidade de que alguém obtenha 10 "caras"?

(c) Qual(ais) a(s) suposição(ções) feita(s) nos cálculos?

(d) Discuta e interprete os resultados.

Solução:

(a)

$$(1/2)^{10} = 1/1024 = 0.00098$$

ou

 X : número de caras em 10 lançamentos

$$X \sim B(n = 10, p = 1/2)$$

$$P[X = 10] = \binom{10}{10} (1/2)^{10} (1 - 1/2)^{10-10} = 0.00098$$

(b)

$$1 - (1 - (1/2)^{10})^{1000} = 0.62358$$

ou

 Y : número pessoas que obtém 10 caras

$$Y \sim B(n = 1000, p = (1/2)^{10})$$

$$P[Y \geq 1] = 1 - P[Y = 0] = 1 - \binom{1000}{0} [(1/2)^{10}]^0 [1 - (1/2)^{10}]^{1000-0} = 0.62358$$

(c) Independência entre lançamentos, independência entre resultados de diferentes pessoas, probabilidades constantes de lançamentos de pessoas obterem 10 caras.

(d)

6. Registros de um laboratório mostram que 1 a cada 20 amostras de um determinado material são perdidas por contaminação. Responda cada um dos seguintes itens declarando a variável aleatória e a sua distribuição.

- (a) Se forem feitas 15 análises qual a probabilidade de que no máximo uma seja contaminada.
- (b) Em um teste para avaliar a contaminação análises serão feitas sequencialmente até que a primeira contaminada seja encontrada. Quantas análises espera-se fazer? Como voce calcularia a probabilidade de que o esse número de análises não chegue a 5?
- (c) O teste anterior foi repetido porém até que a terceira análise mostrasse contaminação. Em um particular ensaio foram feitas 10 análise desta forma. Qual a probabilidade desta ocorrência?
- (d) Um lote contendo 40 amostras das quais 15 eram contaminadas foi enviado para teste em outro laboratório na qual 12 amostras foram selecionadas ao acaso para testes. Qual a probabilidade de encontrar 3 ou mais contaminadas entre as selecionadas?
- (e) Considere agora que o laboratório faz um grande número de análises por mês e registra uma média de 2,5 casos de contaminação grave. Qual a probabilidade de que em um determinado mês não se registre nenhuma contaminação? E de que seja registradas mais do que 5 contaminações?

Solução:

$$p = 1/20$$

(a)

 X : número de contaminadas em 15 análises

$$X \sim B(n = 15, p = 1/20)$$

$$P[X \leq 1] = P[X = 0] + P[X = 1] = 0.829$$

(b)

X : número de análises **não contaminadas** até encontrar a primeira contaminada

$$X \sim G(p = 1/20) \quad P[X = x] = p \cdot (1 - p)^x, x = 0, 1, \dots$$

$$E[X] + 1 = \frac{1 - p}{p} + 1 = \frac{1}{p} = 20$$

$$P[X < 4] = \sum_{i=0}^4 P[X = i] = P[X = 0] + P[X = 1] + P[X = 2] + P[X = 3] + P[X = 4] = \sum_{i=0}^4 (1/20)(1 - 1/20)^i =$$

(c)

X : número de análises **não contaminadas** até encontrar a terceira contaminada

$$X \sim BN(r = 3, p = 1/20) \quad ; \quad P[X = x] = \binom{x + 3 - 1}{3} (1/20)^3 \cdot (1 - 1/20)^x, x = 0, 1, \dots$$

10 análises \rightarrow 7 não contaminadas

$$P[X = 7] = 0.0031$$

(d)

X : número de amostras contaminadas entre as 12

$$X \sim HG(N = 40, r = 15, n = 12) \quad ; \quad P[X = x] = \frac{\binom{25}{12-x} \binom{15}{x}}{\binom{40}{12}}$$

$$P[X \geq 3] = 1 - P[X \leq 2] = 1 - \sum_{i=0}^2 P[X = i] = 1 - P[X = 0] - P[X = 1] - P[X = 2] = 0.9257$$

(e)

X : número de análises contaminadas em um mês

$$X \sim P(\lambda = 2, 5) \quad P[X = x] = \frac{e^{-2,5} 2,5^x}{x!}, x = 0, 1, \dots$$

$$P[X = 0] = 0.0821$$

$$P[X > 5] = 1 - P[X \leq 5] = 1 - (P[X = 0] + P[X = 1] + \dots + P[X = 5]) = 1 - \sum_{i=0}^5 \frac{e^{-2,5} 2,5^i}{i!} = 0.042$$

7.

- (a) Conchas de mexilhões de uma certa espécie possuem, em uma certa região, possuem uma relação altura/comprimento com distribuição normal de média 0,5 e desvio padrão de 0,025. Os animais serão classificados de modo que 20% sejam considerados pequenos, 50% como médios e os restantes 30% como grandes. Quais os valores da relação altura/comprimento que definirão as classes de tamanho desejadas?

Solução:

$$X \sim N(\mu = 0,5 \quad ; \quad \sigma^2 = (0,025)^2)$$

$$P(X < x_1) = 0,20 \implies P(Z < \frac{x_1 - 0,5}{0,025}) = 0,20 \implies z_1 = -0.842 \implies x_1 = \mu + z_1 \cdot \sigma = 0,5 + (-0.842) \cdot 0,025 = 0,479$$

$$P(x_1 < X < x_2) = 0,50$$

$$P(X < x_1) = 0,20$$

$$P(Z < \frac{x_2 - 0,5}{0,025}) = 0,70 \implies z_2 = 0.524 \implies x_2 = \mu + z_2 \cdot \sigma = 0,5 + 0.524 \cdot 0.025 = 0.513$$

- (b) A durabilidade de um tipo de filtro é descrita por uma variável aleatória com distribuição normal de média 60.000 hrs de funcionamento e desvio padrão de 9.000 hrs.
- Se o fabricante garante a duração dos filtros pelas primeiras 47.500 hrs, qual a proporção de filtros que devem ser trocados pela garantia?
 - O que aconteceria com a proporção do item anterior se a garantia fosse para as primeiras 45.000 hrs?
 - Qual deveria ser a garantia (em hrs) de forma a assegurar que o fabricante trocava sob garantia no máximo 4% dos filtros?
 - Se uma indústria comprar cinco (5) filtros, qual será a probabilidade de utilizar a garantia (de 45.000 hrs) para trocar ao menos um (1) dos filtros?

Solução:

$$X \sim N(60.000, 9.000^2)$$

i.

$$P[X < 47500] = P\left[Z < \frac{47500 - 60000}{9000}\right] = P[Z < -1.389] = 0.082$$

$$100 \cdot P[X < 47500] = 8.24\% \text{ dos filtros.}$$

ii.

$$P[X < 45000] = P\left[Z < \frac{45000 - 60000}{9000}\right] = P[Z < -1.667] = 0.048$$

$$100 \cdot P[X < 45000] = 4.78\% \text{ dos filtros.}$$

iii.

$$P[X < t] = 0,04 ; t = ?$$

$$P\left[Z < \frac{t - 60000}{9000}\right] = 0,04$$

$$z = -1.751$$

$$\frac{t - 60000}{9000} = -1.751$$

$$t = 60000 + 9000(-1.751)$$

$$t = 44243.825$$

iv.

Y : número de trocados sob garantia dentre 5 comprados

$$Y \sim B(n = 5, p = P[X < 45000] = 0.048)$$

$$P[Y \geq 1] = 1 - P[Y = 0] = 0.217$$

8. Para efetuar o monitoramento de poluentes em uma determinada área foram coletadas amostras. Os valores obtidos para um determinado elemento poluente são fornecidos a seguir.

[1] 39.7 33.9 43.3 36.3 37.6 27.2 34.6 36.6 35.2 37.8 34.7 37.9

- Caracterize o nível de poluição (deste elemento) na área através de um resumo estatístico adequado dos dados.
- Qual valor voce escolheria para estimar o nível de poluição na área?
- Qual a estimativa deste valor? Como voce representaria a incerteza sobre este estimativa?
- A legislação afirma que se o teor estiver acima de 35 unidades a área deve ser considerada contaminada e sujeita a intervenção de controle. Baseando-se nos dados, voce indicaria a interveção na área?