

Universidade Federal Fluminense (UFF)

Instituto de Matemática e Estatística (IME)

Departamento de Estatística (GET)

Monitora de Bioestatística: Michelle Jeronimo de Souza¹

Orientador: Dr. José Rodrigo de Moraes² (GET/UFF)

Projeto Permanente: Projeto de Monitoria de Bioestatística - 2018

Título do Projeto: EXERCÍCIOS DE BIOESTATÍSTICA: UMA ABORDAGEM USANDO O PROGRAMA R

1ª Questão: Classifique as seguintes variáveis em qualitativas (nominal ou ordinal) ou quantitativa (continua ou discreta):

- a) Tipo sanguíneo (A, B, O, AB)
- b) Índice de massa corporal (Kg/m^2)
- c) Avaliação do plano de saúde (muito bom, bom, regular, ruim, muito ruim)
- d) Uso de prótese dentária (sim, não)
- e) Fratura óssea (presente, ausente)
- f) Pressão arterial sistólica (mmHg)
- g) Utilização de recurso para auxiliar a locomoção (sim, não)
- h) Peso corporal (kg)
- i) Grau de limitação das atividades habituais devido a deficiência visual (nenhum, um pouco, moderado, intenso, muito intenso)
- j) Realização de exame preventivo de câncer do colo do útero (sim, não)
- k) Avaliação da saúde bucal (muito boa, boa, regular, ruim, muito ruim)
- l) Motivo da consulta médica (acidente ou lesão, continuação de tratamento ou terapia, consulta pré-natal, exame médico periódico, problema de saúde mental, doença ou outro problema de saúde, outro motivo)
- m) Altura (metros)
- n) Sexo (masculino, feminino)
- o) Perímetro braquial (milímetros)
- p) Local do último atendimento odontológico (cidade de residência, outra cidade)
- q) Número de medicamentos de uso contínuo

¹ Aluna da Graduação em Estatística da UFF

² Estatístico (ENCE), Mestre em Estatística Social (ENCE) e Doutor em Saúde Coletiva (IESC/UFRJ). Professor Adjunto IV do Departamento de Estatística da UFF.

- r) Nível de bilirrubina no sangue (mg/dl)
- s) Região de residência (Norte, Nordeste, Sudeste, Sul, Centro-Oeste)

2ª Questão: A partir da tabela 1, determine as probabilidades solicitadas, representando os eventos de interesse:

Tabela 1: Distribuição dos adultos de 18 anos ou mais de idade que escovam os dentes pelo menos duas vezes por dia, por cor/raça, segundo a área de localização do domicílio - 2013

Área de localização do domicílio	Raça/Cor			Total
	Branca (B)	Preta (P)	Parda (D)	
Urbana (U)	57.003	10.419	44.293	111.715
Rural (R)	5.694	1.422	8.181	15.297
Total	62.697	11.841	52.474	127.012

Fonte: IBGE, Diretoria de Pesquisas, Coordenação de Trabalho e Rendimento, Pesquisa Nacional de Saúde 2013.

- a) Calcule a probabilidade de selecionar um adulto da área urbana.
- b) Calcule a probabilidade de selecionar um adulto da raça/cor parda.
- c) Calcule a probabilidade de selecionar um adulto da área urbana e da raça/cor parda.
- d) Calcule a probabilidade de selecionar um adulto da área urbana dado que é da raça/cor parda.
- e) Calcule a probabilidade de selecionar um adulto da raça/cor parda dado que reside na área urbana.
- f) Calcule a probabilidade de selecionar um adulto da área rural dado que não é da raça/cor parda.
- g) Calcule a probabilidade de selecionar um adulto da área rural e da raça/cor preta.
- h) Calcule a probabilidade de selecionar um adulto da área rural ou da raça/cor preta.

3ª Questão: Os Boxplots apresentam as distribuições dos percentuais (%) de escolares do sexo masculino e feminino do 9º ano do ensino fundamental que tiveram episódio de asma alguma vez na vida segundo as capitais das unidades da federação (UFs).

Tabela 2: Distribuição dos percentuais (%) de escolares do sexo masculino e feminino do 9º ano do ensino fundamental que tiveram episódio de asma alguma vez na vida.

UF	Sexo	
	Masculino	Feminino
Porto Velho	20,7	17,8
Rio Branco	20,9	20,9
Manaus	25,9	24,8
Boa Vista	20,5	23,4
Belém	25,1	25,3
Macapá	26,1	21,3
Palmas	17,0	15,2
São Luís	16,4	17,8
Teresina	15,6	17,3
Fortaleza	17,5	18,5
Natal	15,6	18,1
João Pessoa	15,9	16,4
Recife	22,1	29,4
Maceió	18,2	19,5
Aracaju	14,2	15,6
Salvador	12,8	17,2
Belo Horizonte	23,3	20,9
Vitória	25,6	24,2
Rio de Janeiro	15,0	16,8
São Paulo	15,0	14,6
Curitiba	18,9	19,1
Florianópolis	16,0	20,3
Porto Alegre	32,6	27,7
Campo Grande	12,6	15,3
Cuiabá	9,5	12,2
Goiânia	14,5	14,7
Brasília	18,7	18,6

Fonte: IBGE, Diretoria de Pesquisas, Coordenação de População e Indicadores Sociais, Pesquisa Nacional de Saúde do Escolar, 2015.

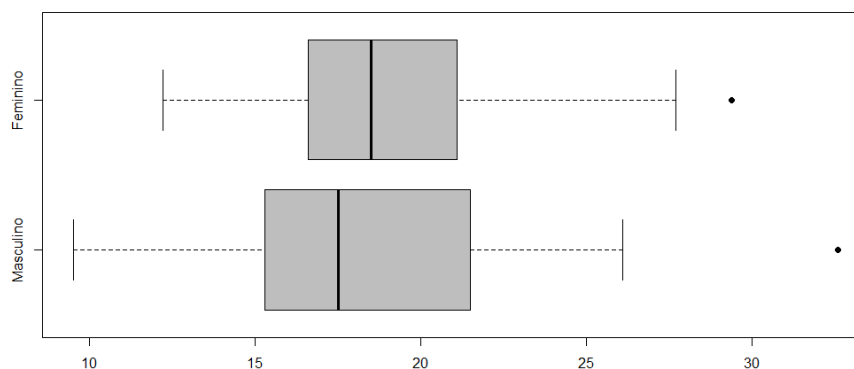


Figura 1: Distribuição dos percentuais (%) de escolares do sexo masculino e feminino do 9º ano do ensino fundamental que tiveram episódio de asma alguma vez na vida.

- Calcule a mediana, a média e a variância para o percentual de escolares que tiveram episódio de asma em cada grupo de sexo.
- Calcule o coeficiente de variação para o percentual de escolares que tiveram episódio de asma em cada grupo de sexo.
- Qual é a UF com maior percentual de escolares do sexo masculino que tiveram episódio de asma? Este valor é um *outlier*?
- Qual é a UF com maior percentual de escolares do sexo feminino que tiveram episódio de asma? Este valor é um *outlier*?
- Qual é o tipo de assimetria da distribuição do percentual de escolares do sexo feminino? Justifique a sua resposta.
- Em qual grupo de sexo existe maior percentual mediano de escolares que tiveram episódio de asma? Justifique a sua resposta.

4ª Questão: Para uma amostra de $n=8$ adultos de 18 anos ou mais de idade, obteve-se os seguintes medidas antropométricas (Tabela 3).

- Construa o gráfico de dispersão entre a altura e o peso corporal e faça a análise gráfica.
- Calcule o coeficiente de correlação linear de Pearson entre a altura e o peso corporal, e avalie o sentido e o grau da relação entre estas variáveis.

Tabela 3: Altura (cm) e Peso (Kg) de $n=8$ adultos.

Adulto	Altura (cm)	Peso corporal (Kg)
1	171	62
2	160	58
3	166	54
4	168	75
5	175	72
6	181	93
7	171	72
8	171	62
Total	1.192	486

5ª Questão: Suponha um estudo realizado com uma amostra aleatória de 32 mulheres selecionada dentre as pacientes que procuraram o ambulatório de Ginecologia de um hospital público universitário. O objetivo do estudo é estimar o percentual de mulheres com bexiga hiperativa (BH). Para tal, perguntou para cada mulher se algum médico ou profissional de saúde disse que ela tem bexiga hiperativa (1) ou outro tipo de incontinência urinária (0). Os dados amostrais são fornecidos a seguir:

1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1

Construa o intervalo de 95% de confiança para a proporção de mulheres com bexiga hiperativa (BH).

6ª Questão: Considere ainda a amostra de 32 mulheres selecionadas do ambulatório de Ginecologia de um hospital público universitário: 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1. Verifique a afirmação de que a proporção de mulheres com bexiga hiperativa (BH) é maior do que 0,15. Utilize o nível de significância de $\alpha=0,05$. **OBS:** É

necessário definir as seguintes etapas: 1) Hipóteses a serem testadas, 2) Valor observado da Estatística de teste, 3) Região crítica (RC) e, 4) Tomada de decisão, inclusive no contexto do problema.

7ª Questão: O IMC é um indicador quantitativo de avaliação nutricional, calculado pela fórmula $\text{peso}/\text{altura}^2$, podendo ser utilizado para classificar os indivíduos em diferentes categorias de estado nutricional. Para uma amostra aleatória de 7 adultos, mediu-se o peso (Kilogramas) e a altura (metros), obtendo os seguintes valores de IMC, expresso em kg/m^2 : 21,2; 22,7; 19,6; 26,6; 23,5; 28,4; 24,6.

Supondo que os valores do IMC seguem uma distribuição aproximadamente normal, calcule:

- a) uma estimativa pontual para o valor médio do IMC.
- b) o intervalo de confiança para o valor médio do IMC, considerando o nível de confiança de 95%.

8ª Questão: Ainda utilizando a amostra aleatória de 7 adultos da questão anterior, para os quais foram mensurados os valores de IMC (em kg/m^2), utilize um teste estatístico de hipóteses para avaliar a afirmação de que os adultos possuem em média valores de IMC que diferem de $24,99 \text{ kg}/\text{m}^2$ (limite máximo de normalidade do estado nutricional). Suponha que os valores do IMC seguem uma distribuição aproximadamente normal e utilize o nível de significância de 95%.

OBS: *É necessário definir as seguintes etapas: 1) Hipóteses a serem testadas, 2) Valor observado da Estatística de teste, 3) Região crítica (RC) e, 4) Tomada de decisão, inclusive no contexto do problema.*

9ª Questão: Uma amostra de cem ($n=100$) pré-escolares (de 2 a 6 anos de idade), de um determinado município, que interromperam precocemente o aleitamento materno exclusivo ($\text{AME}<6$ meses) foi selecionada aleatoriamente, obtendo o resultado de que 42 deles possuem excesso de peso. Teste a afirmação de que 40% dos pré-escolares que interromperam precocemente a amamentação exclusiva ($\text{AME}<6$ meses) apresentavam excesso de peso. **OBS:** *É necessário definir as seguintes etapas: 1) Hipóteses a serem testadas, 2) Valor observado da Estatística de teste, 3) Região crítica (RC) e, 4) Tomada de decisão, inclusive no contexto do problema.*

10ª Questão: Uma amostra de sete pré-escolares com 5 anos de idade foi selecionada aleatoriamente, obtendo as seguintes medidas de perímetro braquial (em centímetros):

16,5; 16,6; 16,9; 17,2; 17,5; 17,8; 18,0.

Considerando que as medidas de perímetro braquial seguem uma distribuição aproximadamente normal com variância conhecida e supostamente igual a $0,25 \text{ cm}^2$, teste a um nível de significância de 5% a afirmação de que o perímetro braquial médio dos pré-escolares difere de 16,7 cm (nível crítico de perímetro braquial para a idade em questão, isto é, abaixo deste valor indica desnutrição). **OBS:** *É necessário definir as seguintes etapas: 1) Hipóteses a serem testadas, 2) Valor observado da Estatística de teste, 3) Região crítica (RC) e, 4) Tomada de decisão, inclusive no contexto do problema.*

Respostas Finais da Lista de Exercícios
“Monitoria de Bioestatística”:

1ª Questão:

- a) Variável qualitativa nominal
- b) Variável quantitativa contínua
- c) Variável qualitativa ordinal
- d) Variável qualitativa nominal
- e) Variável qualitativa nominal
- f) Variável quantitativa contínua
- g) Variável qualitativa nominal
- h) Variável quantitativa contínua
- i) Variável qualitativa ordinal
- j) Variável qualitativa nominal
- k) Variável qualitativa ordinal
- l) Variável qualitativa nominal
- m) Variável quantitativa contínua
- n) Variável qualitativa nominal
- o) Variável quantitativa contínua
- p) Variável qualitativa nominal
- q) Variável quantitativa discreta
- r) Variável quantitativa contínua
- s) Variável qualitativa nominal

2ª Questão:

- a) $P(U) = 0,880$
- b) $P(D) = 0,413$
- c) $P(U \cap D) = 0,349$
- d) $P(U|D) = 0,844$
- e) $P(D|U) = 0,396$
- f) $P(R|\bar{D}) = 0,095$
- g) $P(R \cap P) = 0,011$
- h) $P(R \cup P) = 0,202$

3ª Questão:

a) Feminino:

$$\bar{x} = 19,37$$

$$Md = 18,5$$

$$S^2 = 17,69$$

Masculino:

$$\bar{x} = 18,75$$

$$Md = 17,5$$

$$S^2 = 26,90$$

Comandos do R:

Masculino=c(20.7,20.9,25.9,20.5,25.1,26.1,17.0,16.4,15.6,17.5,15.6,15.9,22.1,18.2,14.2,12.8,23.3,25.6,15.0,15.0,18.9,16.0,32.6,12.6,9.5,14.5,18.7)

Feminino=c(17.8,20.9,24.8,23.4,25.3,21.3,15.2,17.8,17.3,18.5,18.1,16.4,29.4,19.5,15.6,17.2,20.9,24.2,16.8,14.6,19.1,20.3,27.7,15.3,12.2,14.7,18.6)

Feminino

média=mean(Feminino);média

mediana=median(Feminino);mediana

variância=var(Feminino);variância

Masculino

média=mean(Masculino);média

mediana =median(Masculino);mediana

variância=var(Masculino);variância

b) Feminino

$$CV = 21,72\%$$

Masculino

$$CV = 27,66\%$$

Comandos do R:

CV para os percentuais de escolares do sexo feminino (expresso em termos percentuais)

CV= sd(Feminino)/ mean(Feminino)*100; round(CV,2)

CV para os percentuais de escolares do sexo masculino (expresso em termos percentuais)

CV=sd(Masculino)/ mean(Masculino)*100; round(CV,2)

- c) Porto Alegre (32,6% de escolares do sexo masculino com episódio de asma). Sim, é um valor atípico (*outlier*), pois apresenta um grande afastamento dos valores observados para os demais estados, no grupo masculino.

Comando do R:

máximo=max(Masculino);máximo

- d) Recife (29,4% de escolares do sexo feminino com episódio de asma). Sim, é um valor atípico (*outlier*), pois apresenta um grande afastamento dos valores observados para os demais estados, no grupo feminino.

Comando do R:

máximo=max(Feminino);máximo

- e) Possui uma assimetria positiva, pois a média é maior que a mediana ($\bar{x} > Md$), ou visualmente com base no Box-plot (Figura 1) observa-se que: $Q_3 - Md > Q_1 - Md$. Os quartis (Q_1 , Md e Q_3) podem ser calculados usando o comando “summary (Feminino)”.
f) No grupo Feminino, comparando os boxplots nota-se que a mediana dos percentuais de escolares com asma no grupo feminino é maior do que a do grupo masculino.

4ª Questão:

- a) Observando-se o gráfico, pode-se dizer que há uma relação positiva entre as variáveis, isto é, conforme a altura aumenta o peso também aumenta.

Comandos do R:

Altura=c(171,160,166,168,175,181,171)

Peso=c(62, 58,54,75,72,93,72)

plot(Altura, Peso, pch=19, xlab="Altura (m)", ylab="Peso (Kg)", xlim=c(150, 190), ylim=c(50,100))

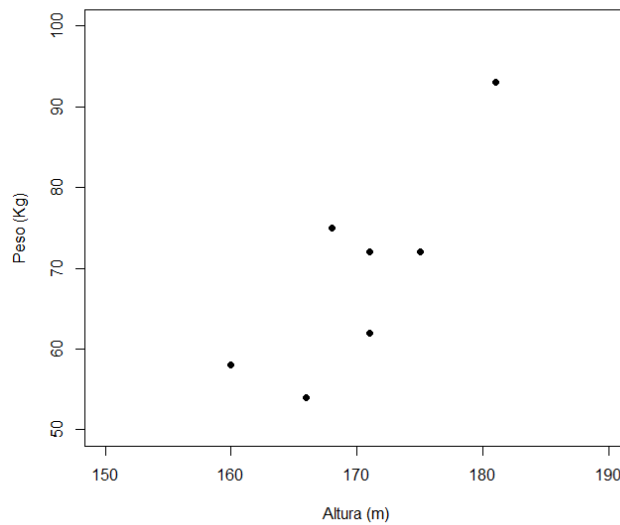


Figura: Gráfico de dispersão entre a altura (m) e o peso corporal (kg).

b) Coeficiente de correlação linear de Pearson: $r=0,826$

Relação positiva ($r>0$) e forte ($|r| \geq 0,70$)

- Quanto ao sentido da correlação: se o coeficiente for positivo (“+”) a relação é positiva (ou crescente), e quando for negativo (“-”) a relação é negativa (ou decrescente). Quanto ao grau da correlação (Siqueira & Tibúrcio, 2011)³: $0,7 \leq |r| \leq 1$ indica uma correlação forte; $0,4 \leq |r| < 0,7$ indica uma correlação moderada; $0 \leq |r| \leq 0,4$ indica uma correlação fraca.

Comando do R:

`r=cor(Altura, Peso);r`

5ª Questão:

$$\hat{p} = 18,75\%; \quad IC_{p,95\%} = [5,23\%; 32,27 \%]$$

Comandos do R:

Instalando e carregando o pacote

`install.packages("TeachingDemos")`

`library(TeachingDemos)`

Estimativa pontual da proporção populacional

³ SIQUEIRA, A.L.; TIBÚRCIO, J.D. **Estatística na área de saúde: conceitos, metodologia, aplicações e prática computacional**. Belo Horizonte: Coopmed, 2011.

```
x=c(1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1)
```

```
z.test(x, stdev=sqrt(0.1875*(1-0.1875)), conf.level = 0.95)$estimate
```

```
# IC de 95% para a proporção populacional
```

```
z.test(x, stdev=sqrt(0.1875*(1-0.1875)), conf.level = 0.95)$conf.int
```

OBS: "stdv" representa o desvio-padrão da população, que no caso da estimação de proporção é:
 $stdev = \sqrt{p \cdot q}$

6ª Questão:

$z_{obs} \cong 0,594$; valor crítico: $+ z_{0,05} = + 1,64$; Como z_{obs} não pertence a região crítica ($z_{obs} \notin RC$), conclui-se que não há evidências estatísticas para rejeitar H_0 ao nível de significância de 5%, ou seja, não há evidências de que a proporção de mulheres com BH seja maior do que 0,15.

Comandos do R:

```
# Instalando e carregando o pacote
```

```
install.packages("TeachingDemos")
```

```
library(TeachingDemos)
```

```
# Teste Z unilateral superior para a proporção populacional p
```

```
z.test(x, mu=0.15, stdev=sqrt(0.15*(1-0.15)), alternative="greater")
```

```
# Valor critico arredondado para 2 casas - Pode ser obtido na Tabela da Normal Padrão  $Z \sim N(0,1)$ 
```

```
zcritico=qnorm(0.95);round(zcritico,2)
```

7ª Questão:

a) $\bar{x} = 23,8 \text{ kg/m}^2$

Comandos do R:

```
imc=c(21.2, 22.7, 19.6, 26.6, 23.5, 28.4, 24.6)
```

```
mean(imc)
```

b) $IC_{\mu,95\%} = [20,99 \text{ kg/m}^2; 26,61 \text{ kg/m}^2]$

Comando do R:

```
# IC de 95% para a média populacional  $\mu$ , usando a distribuição t-Student
```

```
t.test(imc,conf.level = 0.95)$conf.int
```

8ª Questão:

$t_{obs} \cong -1,037$; valores críticos: $\pm t_{0,025;6} = \pm 2,447$; Como t_{obs} não pertence a região crítica ($t_{obs} \notin RC$), conclui-se que não há evidências estatísticas para rejeitar H_0 ao nível de significância de 5%, ou seja, o valor médio do IMC não difere significativamente de 24,99 kg/m^2 .

Comandos do R

```
# Teste t-Student (bilateral) para a média populacional  $\mu$ 
```

```
t.test(imc, mu=24.99, alternative="two.sided", conf.level= 0.95)
```

```
# Valores críticos obtidos na distribuição t-Student (aproximados para 3 casas decimais)
```

```
tc1=qt(0.025,6);round(tc1,3)
```

```
tc2=qt(0.975,6);round(tc2,3)
```

9ª Questão:

$z_{obs} = 0,408$; valores críticos: $\pm z_{0,025} = \pm 1,96$; Como z_{obs} não pertence a região crítica ($z_{obs} \notin RC$), conclui-se que não há evidências estatísticas para rejeitar H_0 ao nível de significância de 5%, ou seja, não há evidências de que a proporção de pré-escolares com excesso de peso, dentre os que interromperam o AME, seja significativamente diferente de 40%.

Comandos do R:

```
# Teste Z (bilateral) para a proporção populacional  $p$ 
```

```
a=42 # Número de elementos com a característica de interesse (sucessos)
```

```
n=100 # Tamanho da amostra
```

```
pchapeu = a/n;pchapeu # Proporção amostral estimada
```

```
p0 = 0.40 # Valor hipotético para  $\mu$ ,  $\mu_0=0,40$  (40%)
```

```
zobs = (pchapeu-p0)/(sqrt(p0*(1-p0)/n));zobs #valor observado da estatística de teste
```

```
# Valores críticos da distribuição normal padrão (aproximados para 2 casas decimais)
```

```
zcritico1=qnorm(0.025);round(zcritico1,2)
```

```
zcritico2=qnorm(0.975);round(zcritico2,2)
```

P-valor do teste

pvalor=2*pnorm(zobs, lower.tail = FALSE);pvalor # p-valor=2*P(Z>=zobs/ H0 verdadeira)

10ª Questão:

$z_{obs} \cong 2,72$; valores críticos: $\pm z_{0,025} = \pm 1,96$; Como z_{obs} pertence a região crítica ($z_{obs} \in RC$), conclui-se que há evidências estatísticas para rejeitar H_0 ao nível de significância de 5%, ou seja, o perímetro braquial médio difere significativamente de 16,7 cm.

Comando do R:

Instalando e carregando o pacote

```
install.packages(TeachingDemos)
```

```
library(TeachingDemos)
```

```
pb= c(16.5, 16.6, 16.9, 17.2, 17.5, 17.8, 18.0)
```

Teste Z (bilateral) para a média populacional μ (inclui o p-valor do teste)

OBS: stdev=Desvio-padrão da população: $\sigma = \sqrt{\sigma^2} = \sqrt{0,25} = 0,5$

```
z.test(pb, mu=16.7, stdev=0.5, conf.level = 0.95)
```

Valores críticos da distribuição normal padrão

```
zcritico1=qnorm(0.025);round(zcritico1,2) # -1,96
```

```
zcritico2=qnorm(0.975);round(zcritico2,2) # 1,96
```