



Leia:

*Para os cálculos, respostas e decisões nesta lista de exercícios o aluno pode e deve usar o R (ou R-Studio) ou EpiInfo. Todas as linhas de comando utilizadas (no caso de usar o R) ou prints de janelas de comando (no caso do EpiInfo) devem ser anexadas nas respostas – serão usadas na correção.*

*Os dois softwares têm bons tutoriais e vídeos explicativos, especialmente o R. O documento enviado com as respostas deve ser fotografado (em caso de está manuscrito) ou em PDF (ou .doc, em caso de ser digitado).*

*A média das notas desta e da próxima lista será a primeira nota a ser lançada no sig@.*

## Primeira Lista de Exercícios de Epidemiologia – 2022.1

Aluno: \_\_\_\_\_

### Resumindo os dados

#### Gráficos

1º) Na tabela abaixo são listados os gastos com os cuidados com a saúde per capita em 1989 para 23 das 24 nações que constituem a Organização para Cooperação Econômica e Desenvolvimento.

Nação	Gastos Per Capita (US\$)
Austrália	1032,00
Áustria	1093,00
Bélgica	980,00
Inglaterra	836,00
Canadá	1683,00
Dinamarca	912,00
Finlândia	1067,00
França	1274,00
Alemanha	1232,00
Grécia	371,00
Islândia	1353,00
Irlanda	658,00
Itália	1050,00
Japão	1035,00
Luxemburgo	1193,00
Países Baixos	1135,00
Nova Zelândia	820,00
Noruega	1234,00

<b>Portugal</b>	464,00
<b>Espanha</b>	644,00
<b>Suécia</b>	1361,00
<b>Suíça</b>	1376,00
<b>Estados Unidos</b>	2354,00

- Ordene esses países de acordo com os gastos per capita com a saúde.
- Construa um histograma para os valores dos gastos per capita.
- Descreva a forma do histograma..

2º) A tabela abaixo categoriza 10.614.000 visitas ao consultório de especialistas de doenças cardiovasculares nos estados unidos por duração de cada visita. Uma duração de 0 (zero) minuto implica que o paciente não teve contato direto com o especialista.

<b>Duração (minutos)</b>	<b>Número de visitas (milhares)</b>
<b>0</b>	390
<b>1-5</b>	227
<b>6-10</b>	1023
<b>11-15</b>	3390
<b>16-30</b>	4431
<b>31-60</b>	968
<b>61+</b>	185
<b>Total</b>	10614

Pode-se fazer a afirmação de que as visitas a consultórios de especialistas de doenças cardiovasculares têm duração mais frequente entre 16 e 30 mim. Você concorda com essa afirmação? Por quê? Justifique.

### Medidas de posição e de dispersão

3º) Na tabela abaixo estão taxas de glicose em miligramas por 100ml de sangue em ratos machos da raça Wistar com 30 dias de idade, que serão usados em um experimento para o teste de determinada droga. Ache a média e a mediana. A moda seria uma informação importante? Por quê?

Tabela 01 - Taxas de glicose em miligramas por 100ml de sangue em ratos machos da raça Wistar com 30 dias de idade

<b>Nº do rato</b>	<b>Taxa de glicose</b>
<b>1</b>	101
<b>2</b>	98
<b>3</b>	97
<b>4</b>	104

<b>5</b>	95
<b>6</b>	105

4º) Com os dados apresentados na tabela que segue, calcule o número médio de dentes cariados, para cada sexo.

Tabela 02 – escolares de 12 anos, segundo o número de dentes cariados e o sexo

Nº de dentes cariados	Sexo	
	Masculino	Feminino
<b>0</b>	16	13
<b>1</b>	2	5
<b>2</b>	3	3
<b>3</b>	2	2
<b>4</b>	2	2

5º) Para estudar o tempo de latência de um sonífero usado em ratos de laboratório, um pesquisador administrou o sonífero a 10 ratos e determinou o tempo que eles demoravam para dormir. Dos 10 ratos, dois demoraram meio minuto, quatro demoraram 1 minuto, três demoraram 1 minuto e meio e um rato não dormiu. Calcule o tempo médio da latência.

6º) Para calcular dois programas de treinamento para executar um serviço especializado, foi feito um experimento. Dez homens foram selecionados ao acaso para serem treinados pelo método A e outros 10 para serem treinados pelo método B. Terminado o treinamento, todos os homens fizeram o serviço e foi registrado o tempo em que cada um desempenhou a tarefa. Os dados estão na tabela a seguir. Encontre para cada método: mínimo, primeiro quartil, mediana, terceiro quartil e máximo. Compare e discuta.

Tabela 03 – Tempo, em minutos, despendido em executar o serviço, segundo o método de treinamento.

Método	
<b>A</b>	<b>B</b>
15	23
20	31
11	13
23	19
16	23
21	17
18	28
16	26
27	25
24	28

7º) Responda às questões:

- a) O valor do desvio padrão pode ser maior do que o valor da média?
- b) O valor do desvio padrão pode ser igual ao valor da média?
- c) O valor do desvio padrão pode ser negativo?
- d) Quando o desvio padrão é igual a zero?

8º) Considere a situação de uma loja de equipamentos de som. Em dez ocasiões durante os três últimos meses, a loja usou comerciais de televisão de fins de semana para promover as vendas. Os gerentes querem verificar se pode ser demonstrada uma relação entre o número de comerciais mostrados e as vendas na loja durante a semana seguinte. Os dados para as 10 semanas com vendas em centenas de dólares estão mostrados na tabela abaixo. Construa o diagrama de dispersão considerando o número de comerciais como sendo o eixo (x) e as vendas o eixo (y). Faça comentários sobre a relação entre as duas variáveis.

Semana	Nº de Comerciais (x)	Volume de vendas (por R\$ 100)
1	2	50
2	5	57
3	1	41
4	3	54
5	4	54
6	1	38
7	5	63
8	3	48
9	4	59
10	2	46

9º) Um estudo foi conduzido para investigar o prognóstico a longo prazo de crianças que sofreram um episódio agudo de meningite bacteriana. Abaixo estão listados os tempos para o ataque apoplético de 13 crianças que tomaram parte no estudo. Em meses, as medidas foram:

0,10   0,25   0,20   4   12   12   24   24   31   36   42   55   96

- a) Obtenha as seguintes medidas-resumo numéricas dos dados
  - i. Média
  - ii. Mediana
  - iii. Moda
  - iv. Amplitude
  - v. Intervalo interquartil
  - vi. Desvio padrão
- b) Mostre que  $\sum_{i=1}^{13}(x_i - \bar{x})$  é igual a zero.

10º) Em Massachusetts, oito indivíduos sofreram um episódio inexplicável de intoxicação por vitamina D que exigiu hospitalização; pensou-se que essas ocorrências extraordinárias pudessem resultar de uma excessiva suplementação de leite. Os níveis de cálcio e albumina - um tipo de proteína - no sangue para cada indivíduo no momento da internação no hospital são mostrados abaixo:

Cálcio (mmol/l)	Albumina (g/l)
2,92	43
3,84	42

<b>2,37</b>	42
<b>2,99</b>	40
<b>2,67</b>	42
<b>3,17</b>	38
<b>3,74</b>	34
<b>3,44</b>	42

- Obtenha a média, a mediana, o desvio-padrão e a amplitude dos níveis de cálcio registrados.
- Calcule a média, a mediana, o desvio-padrão e a amplitude para os dados de níveis de albumina.
- Para indivíduos saudáveis, o intervalo normal de valores de cálcio é de 2,12 até 2,74 mmol/l, enquanto o intervalo de níveis de albumina é de 32 até 55 g/l. Você acredita que os pacientes que sofreram intoxicação por vitamina D tinham níveis normais de cálcio e de albumina no sangue?

11º) Considere os dados da Questão 10 e faça o que se pede:

- construa um intervalo de confiança unilateral de 95% - um limite inferior – para o nível médio verdadeiro de cálcio de indivíduos que sofreram a intoxicação de vitamina D
- construa um intervalo de confiança unilateral inferior de 95% para o nível médio verdadeiro de albumina desse grupo.
- Para indivíduos saudáveis, o intervalo normal de valores de cálcio é de 2,12 a 2,74 mmol/l e o intervalo de níveis de albumina é de 32 a 55 g/l. Você acredita os pacientes que sofrem de intoxicação de vitamina D têm níveis normais de cálcio e de albumina no sangue?

### Sensibilidade e especificidade

12º) Um estudo registrou que a sensibilidade da mamografia como teste de triagem para detecção do câncer de mama é 0,85, enquanto sua especificidade é 0,80.

- Qual a probabilidade de um resultado de teste falso negativo?
- Qual a probabilidade de resultado falso positivo?
- Na população na qual a probabilidade de que uma mulher tenha câncer de mama é 0,0025, qual a probabilidade de que tenha câncer, se sua mamografia for positiva?

13º) O Instituto Nacional de Segurança Ocupacional e de Saúde desenvolveu uma definição de casos de síndrome de túnel carpal – uma doença do punho – que incorpora três critérios: sintomas de envolvimento do nervo, história de fatores de risco ocupacional e a presença de materiais de exames físicos. A sensibilidade dessa definição como um teste para a síndrome de túnel carpal é de 0,67; sua especificidade é de 0,58.

- Em uma população cuja prevalência da síndrome de túnel carpal esta estimada em 15%, qual o valor preditivo de um resultado positivo de teste?
- Como esse valor preditivo se modifica, se a prevalência for somente 10%? E se for 5%?

14º) Os dados seguintes são tomados de um estudo que investiga o uso de uma técnica chamada ventriculografia radionuclídica como teste de diagnóstico para se detectar doença da artéria coronária.

Teste	Doença		Total
	Presente	Ausente	
Positivo	302	80	382
Negativo	179	372	551
Total	481	452	933

- Qual a sensibilidade da ventrículografia radionuclídica nesse estudo? Qual a sensibilidade?
- Para uma população cuja prevalência da doença da artéria coronária seja 0,10 calcule a probabilidade de que um indivíduo tenha a doença, sendo que ele apresenta resultado positivo usando a ventrículografia radionuclídica.
- Qual o valor preditivo de um teste negativo.

### Intervalo de confiança, teste de hipótese

15º) Os seguintes dados foram coletados para uma amostra de uma população normal: 10, 8, 12, 15, 13, 11, 6, 5.

- Qual a estimativa pontual da média populacional?
- Qual a estimativa pontual do desvio padrão da população?
- Qual o intervalo de confiança com 95% para a média da população?

16º) Considere o seguinte teste de hipótese.

$$H_0: \mu \geq 10$$

$$H_0: \mu < 10$$

Uma amostra com  $n=50$  fornece uma média de amostra de 9,46 e um desvio padrão da amostra de 2.

- Com  $\alpha = 0,05$ , qual o valor crítico para  $z$ ? Qual a regra de rejeição?
- Calcule o valor da estatística de  $z$ . Qual a sua conclusão?

### Teste Qui-quadrado

17º) Os dados seguintes vêm de um estudo concebido para investigar problemas de bebidas entre os estudantes universitários. Em 1983, foi perguntado a um grupo se já dirigira um automóvel depois de beber. Em 1987, depois de atingida a idade legal para consumo de bebidas alcoólicas, foi feita a mesma questão a outro grupo universitário.

Dirigia enquanto bebia	Ano		
	1983	1987	Total
Sim	1250	991	2241
Não	1387	1666	3053
Total	2637	2657	5294

- Use o teste qui-quadrado para avaliar a hipótese nula de que as proporções de estudantes da população que dirigia enquanto bebia são as mesmas nos dois anos.

- b) O que você conclui sobre o comportamento desses estudantes?

### Correlação e Regressão

18º) Cinco observações tomadas para duas variáveis são apresentadas a seguir:

$x_i$	4	6	11	3	16
$y_i$	50	50	40	60	30

- Desenvolva um diagrama de dispersão com x no eixo horizontal.
- O que o diagrama de dispersão, desenvolvido no item (a) indica sobre a relação entre as duas variáveis?
- Calcule e interprete a covariância da amostra.
- Calcule e interprete o coeficiente de correlação da amostra.

19º) Dada cinco observações para duas variáveis, x e y.

$x_i$	1	2	3	4	5
$y_i$	3	7	5	11	14

- Desenvolva o diagrama de dispersão par esses dados.
- O que o diagrama de dispersão, desenvolvido no item (a) indica sobre a relação entre as duas variáveis?
- Desenvolva uma equação de regressão estimada calculando os valores de  $\widehat{\beta}_0$  e  $\widehat{\beta}_1$  usando qualquer método de estimação e cite o método que usou.
- Use a equação de regressão estimada para determinar o valor de y quando x=4.

### Regressão logística

20º) Em um estudo que investiga os fatores de risco maternas para sífilis congênita, a doença é tratada como uma variável resposta dicotômica, na qual 1 representa a presença da doença em um recém nascido e 0 sua ausência. Os coeficientes estimados do modelo de regressão logística que contém as variáveis explicativas uso de cocaína ou crack, *status* marital, número de consultas pré-natais, uso de álcool e nível educacional estão listados abaixo. O intercepto não é dado.

Variável	Coeficiente
Uso de cocaína/crack	1,354
Status marital	0,779
Número de consultas pré-natais	-0,098
Uso de álcool	0,723
Nível educacional	0,298

- conforme o número de consulta pré-natais aumente, o que acontece com a probabilidade de uma criança nascer com sífilis?
- O *status* marital representa uma variável dicotômica, na qual o valor 1 indica que a mulher não é casada e 0 que ela é casada. Qual a chance relativa de que um recém nascido tenha sífilis para mães não casadas *versus* as mães casadas?

c) O uso de cocaína ou de crack também é uma variável aleatória dicotômica; o valor 1 indica que uma mulher usou drogas durante a gravidez e 0 que não usou. Qual a razão de chances estimada de uma criança nascer com sífilis para as mulheres que usaram cocaína ou crack *versus* as que não usaram?

Boa prova!