

# ME110 - NOÇÕES DE ESTATÍSTICA

## PROVA 1

HILDETE PINHEIRO - DEPTO. DE ESTATÍSTICA - UNICAMP

- Você deve colocar seu NOME e RA na primeira página do arquivo submetido com a solução da prova
- Você pode submeter até 4 arquivos: a solução da prova, o arquivo com seus dados e um ou dois com códigos ou cálculos, se achar necessário. Os arquivos devem conter nomes que identifiquem o(a) aluno(a), como o login da DAC. Por exemplo:
  - arquivo com a solução da prova (h902287-P1ME110.pdf),
  - arquivo dos dados (h902287ME110.csv),
  - arquivo com o código (h902287-progP1ME110.R) e
  - arquivo excel com cálculos (h902287-calcP1ME110.xls).
- O número de páginas (incluindo gráficos e tabelas) do seu arquivo .pdf não deve exceder 15 páginas.
- Você pode consultar livros, notas de aula, usar o computador, calculadora etc. No entanto, a prova é individual e você não pode trocar ideias ou conversar com seus colegas.

**Questão 1** (2,5 pontos)

Para os dados obtidos para gerar a Tabela 1 das instruções para coleta de dados, faça um teste exato de Fisher para a hipótese de igualdade de proporções de óbitos nos dois estados. Assuma que os dados foram coletados a partir de uma amostra estratificada por estado.

**Tabela 1:** Número total de casos e de óbitos de COVID-19 no período de XX-XX-XX-XX a XX-XX-XXXX (ou no dia XX-XX-XXXX) nos estados XX e YY.

	Número de		Total de
	óbitos	não óbitos	casos
Estado 1	$n_{11}$	$n_{12}$	$n_{1.}$
Estado 2	$n_{21}$	$n_{22}$	$n_{2.}$
Total	$n_{.1}$	$n_{.2}$	$n$

(a) Coloque a tabela com os dados coletados e os estados correspondentes. (0,25 pontos)

(b) Descreva suas hipóteses nula e alternativa. (0,5 pontos)

(c) Explique passo a passo como é feito o teste e escolha o nível de significância de 5%. (0,75 pontos)

(d) Calcule o p-valor do teste. (0,5 pontos)

(e) Interprete os resultados. (0,5 pontos)

**Questão 2** (2,5 pontos)

Utilize os dados obtidos para gerar a Tabela 2 das instruções para coleta de dados e responda aos itens abaixo.

**Tabela 2:** Número de óbitos com suspeita ou confirmação de COVID-19 por sexo e faixa etária para o estado XXXX no período de 01/02/2020 até XX/05/2020.

Faixa etária	Sexo		Total
	Masc	Fem	
< 39 anos	$n_{11}$	$n_{12}$	$n_{1.}$
40-79 anos	$n_{21}$	$n_{22}$	$n_{2.}$
$\geq 80$ anos	$n_{31}$	$n_{32}$	$n_{3.}$
Total	$n_{.1}$	$n_{.2}$	$n$

**Para cada item, explique como você chegou nos resultados. Não coloque somente um número com o resultado final.**

- (a) Quais as estimativas de probabilidades de óbito para cada sexo? (0,5 pontos)
- (b) Quais as estimativas de probabilidades de óbito para cada faixa etária? (0,5 pontos)
- (c) Entre as pessoas do sexo masculino, qual a probabilidade de óbito na faixa de 40-79 anos? (0,5 pontos)
- (d) Qual a estimativa da probabilidade de óbito para uma pessoa do sexo Feminino com mais de 80 anos? (0,5 pontos)
- (e) O que você pode dizer sobre as probabilidades de óbito de acordo com sexo e faixa etária? Interprete os resultados encontrados. (0,5 pontos)

**Questão 3** (2,5 pontos) Para o conjunto de dados obtido a partir do arquivo *ministSaude.csv*:

- (a) Calcule as estatísticas descritivas para as variáveis (*casosNovos*) e (*obitosNovos*) por região e por estado. Qual a interpretação para a média? Entre as suas regiões, qual a que está em pior situação? Qual(is) estado(s) está(ão) em pior(es) situação(ões)? Interprete os resultados de um modo geral. (0,5 pontos)
- (b) Faça um diagrama de caixa da variável *casosNovos* ao longo do período para cada estado. Por exemplo, coloque os diagramas de caixa de cada estado (lado a lado) para a região 1 num gráfico e depois faça outro gráfico com os dados dos estados da região 2. Interprete os resultados. Lembre-se de escrever quem são os eixos X e Y nos gráficos e colocar os títulos dos gráficos de forma clara para o leitor. (1,0 ponto)
- (c) Agora use apenas *casosAcumulados* e *obitosAcumulados* para cada um dos estados de suas regiões nos dias 04/04/2020 e 04/05/2020 e compare os resultados. (0,5 pontos)
- (d) Usando os *obitosAcumulados* e *casosAcumulados* do dia 04/05/2020, calcule a estimativa da taxa de morte em cada estado. Compare os resultados. (0,5 pontos)

**Questão 4** (2,5 pontos)

(a) Para o conjunto de dados obtido a partir do arquivo *ministSaude.csv*, faça um gráfico de dispersão do dia (*date*)  $\times$  número de casos (*casosNovos*) para cada estado. Se quiser, pode ligar os pontos do gráfico. Se for possível, coloque os 4 estados no mesmo gráfico com legenda. Se não for possível, pode fazer 4 gráficos. Lembre-se de escrever quem são os eixos X e Y no(s) gráfico(s) e colocar o(s) título(s) do(s) gráfico(s) de forma clara para o leitor. Interprete os resultados. (0,75 pontos)

(b) Para o conjunto de dados obtido a partir do arquivo *minstSaude.csv*, faça um gráfico de dispersão do dia (*date*)  $\times$  número de óbitos (*obitosNovos*) para cada estado. Se quiser, pode ligar os pontos do gráfico. Se for possível, coloque os 4 estados no mesmo gráfico com legenda. Se não for possível, pode fazer 4 gráficos. Lembre-se de escrever quem são os eixos X e Y no(s) gráfico(s) e colocar o(s) título(s) do(s) gráfico(s) de forma clara para o leitor. Interprete os resultados. (0,75 pontos)

(c) Imagine que está sendo desenvolvida uma Droga para combater COVID-19 e você é o estatístico responsável para planejar esse ensaio clínico. Como você faria o planejamento desse experimento? (1,0 ponto)