

UNIVERSIDADE FEDERAL DO PIAUÍ  
CENTRO DE CIÊNCIAS DA NATUREZA  
CURSO DE GRADUAÇÃO EM ESTATÍSTICA

**2ª LISTA**

**Nome:** Filipe Costa

**1**

**Estudo de coorte:**

**Mortalidade neonatal precoce no município de Caxias do Sul: um estudo de coorte**

Objetivos:

- Estabelecer o perfil dos neonatos de Caxias do Sul e estudar a mortalidade neonatal precoce, suas causas e as variáveis a esta relacionadas.

**Estudo caso-controle:**

**Determinantes da mortalidade neonatal: estudo caso-controle em Fortaleza, Ceará, Brasil**

Objetivos:

- O estudo objetivou determinar os fatores associados aos óbitos neonatais em Fortaleza no ano de 2009.

**Estudo transversal:**

**Doenças sexualmente transmissíveis e gênero: um estudo transversal com adolescentes no Rio de Janeiro**

Objetivos:

- O objetivo do presente estudo foi conhecer a população adolescente atendida no Núcleo de Estudos em Saúde do Adolescente (NESA), da Universidade do Estado do Rio de Janeiro (UERJ), do ponto de vista da sexualidade e descrever possíveis fatores de risco às DST.

## 2

- **Informações**

- $n = 400$
- 40% possuíam indústrias próximas operantes
- Probabilidade da amostra acusar contaminação do rio, com indústrias próximas é de 60%
- A probabilidade em locais sem indústrias próximas é 20 %

### a

```
table2 = data.frame("Proximidade" = c('Distante', 'Próximo'),
                    'Sim' = c(96, 48),
                    'Não' = c(64, 192),
                    check.names = FALSE)

table2 |>
dplyr::mutate(Total = Sim + Não)|>
janitor::adorn_totals("row")|>
kableExtra::kbl(booktabs = T, align = "cccc")|>
kableExtra::add_header_above(c(" ", 'Contaminação' = 2, " "))
```

Proximidade	Contaminação		Total
	Sim	Não	
Distante	96	64	160
Próximo	48	192	240
Total	144	256	400

O estudo apresentado se configura como **coorte**. Após o processo de amostragem é determinado quais são os grupos expostos ou não expostos a elemento de influência, neste caso, a proximidade a uma fábrica, sendo 40 % próximos e 60 % distantes.

Após este processo de determinação da exposição, é analisado a presença da variável reposta, que seria o índice de contaminação.

Uma característica do estudo de coorte é que para sua configuração,  $_{1+}$  e  $_{n_{2+}}$  são fixados no deliamento amostral e considerando que temos um  $n$  amostral definido, 400, o modelo probabilístico mais adequado é o Produto Multinomiais.

### b

Os testes para o estudo das amostras nos rios devem analisar a independência da associação entre variável explicativa e a variável resposta, que são nesse caso:

- Variável explicativa: **Proximidade do ponto de coleta para uma industria**
- Variável resposta: **Índice de contaminação**

Diante do contexto, podemos realizar um teste de hipótese com as seguintes hipóteses:

- $H_0$ : a distância para uma fábrica não influencia na qualidade da água.
- $H_1$ : a distância para uma fábrica influencia na qualidade da água.

O teste indicado para independência é o  $\chi^2$  que, resumidamente, compara o valor registrado na tabela com o valor esperado. A diferença dos valores influencia no resultado.

A função  $\chi^2$  tem o seguinte formato:

$$\chi^2 = \sum \frac{o_i - e_i}{e_i}$$

Onde  $o_i$  são os valores observados e  $e_i$  são os valores esperados que precisamos calcular. Um método simples para obtenção dos valores consiste em multiplicar o valor total da coluna e linha dividindo pelo total ( $e_{ij} = \frac{n_{i+}n_{+j}}{total}$ ):

```
e11 = 144*160/400; e11
```

```
[1] 57.6
```

```
e12 = 256*160/400; e12
```

```
[1] 102.4
```

```
e21 = 144*240/400; e21
```

```
[1] 86.4
```

```
e22 = 256*240/400; e22
```

```
[1] 153.6
```

Obtidos os valores esperados, iremos utilizar a função R para o cálculo da  $\chi^2$ :

```
p = matrix(c(e11, e21, e12, e22), 2, 2, byrow = T) #valores esperados
mat = matrix(c(96, 48, 64, 192), 2, 2, byrow = T) #valores observados
testeQui = chisq.test(mat, p)

testeQui
```

Pearson's Chi-squared test with Yates' continuity correction

```
data: mat
X-squared = 64.942, df = 1, p-value = 7.714e-16
```

Com o resultado apresentando um p-valor muito inferior a 0.05, temos evidências para afirmar que a distância para uma fábrica influencia na qualidade da água.

**c**

- O risco relativo é a probabilidade de um evento ocorrer no grupo exposto em relação ao não exposto:

```
exposto = 96/160
Nexposoto = 64/240

RiscoRelativo = round(exposto/Nexposoto, 3); RiscoRelativo
```

```
[1] 2.25
```

A interpretação é que uma amostra de água próxima a uma fábrica tem um risco de 2.25% maior de estar contaminada comparada a uma amostra distante.

**d**

```
Intervalorisco = function(p1, p2, n1, n2){
  f = log(p1) + log(p2)
  vf = (1 - p1)/n1*p1 + (1 - p2)/n2*p2
  ic1 = exp(1)^(f + 1.96*vf)
  ic2 = exp(1)^(f - 1.96*vf)
  print(list(ic2,ic1))
}

Intervalorisco(exposto, Nexposoto, 144, 256)
```

```
[[1]]
[1] 0.1592396
```

```
[[2]]
[1] 0.160764
```

**e**

O cálculo da razão de chance ou *odds ratio* consiste na divisão da probabilidade de um evento acontecer em um grupo em relação a outro grupo, neste caso, os grupos são:

- Amostra da água próxima a uma fábrica
- Amostra da água distante de uma fábrica

```
oddE = (96/144)/(48/144); oddE
```

```
[1] 2
```

```
oddNE = (64/256)/(192/256); oddNE
```

```
[1] 0.3333333
```

```
oddsRatio = oddE/oddNE; oddsRatio
```

```
[1] 6
```

O resultado apresentado com um *odds Ratio* de 6 significa que a chance de contaminação de ocorrer num grupo próximo a fábrica 6 vezes a de um grupo mais distante.

**f**

```
intervaloOdds = function(n1, n2, n3, n4, Od){  
  f = log(Od)  
  Vf = 1/n1 + 1/n2 + 1/n3 + 1/n4  
  Ic1 = exp(1)^(f + 1.96*Vf)  
  Ic2 = exp(1)^(f - 1.96*Vf)  
  print(list(Ic1, Ic2))  
}  
  
intervaloOdds(96, 48, 64, 192, 6)
```

```
[[1]]  
[1] 6.644855
```

```
[[2]]  
[1] 5.417726
```

Relizando a função do intervalo de confiança (95%) para o *Odds Ratio*, podemos expressar o intervalo da seguinte forma:  $5,41 < odds\ Ratio < 6,64$

### 3

- **Informações**

- $n = 381$
- 127 crianças com morte devido a infecção respiratória
- 254 em caso de controle
- **Fatores de risco relacionados com a mortalidade:**
  - \* nível socioeconômico
  - \* grau de escolaridade
  - \* tipo de aleitamento
  - \* baixo peso ao nascer
  - \* idade

#### a

- **Variável resposta:**

- Fatores de risco relacionados com a mortalidade infantil

- **Variáveis explicativas**

- *nível socioeconômico*: Politômicas
- *grau de escolaridade*: Politômicas
- *tipo de aleitamento*: Politômicas
- *baixo peso ao nascer*: Binária
- *idade*: Contagem discreta

#### b

O objeto deste tipo de estudo é avaliar a associação de um tipo de doença com fatores de risco. Nesta abordagem, temos dois grupos, aqueles que possuem a doença, os *casos* e o indivíduos livres, *controle*. Um importante prerequisite do caso de caso-controle é que os dois grupos devem possuir a mesma probabilidade de contaminação.

As principais vantagens desta modalidade de pesquisa é o custo e o tempo, comparado ao de corte, entretanto, são propensos a enviesamentos devido a possíveis manipulações nos grupos comparados.

#### c

O pesquisador ao realizar uma pesquisa *caso-controle* deve buscar minimizar os vieses de seleção, dessa forma, o processo de amostragem deve ser bem definido. Na pesquisa apresentada, percebemos um pareamento do grupo de controle em relação ao número de casos, para cada situação de morte, há 2 do grupo controle.

#### d

Não poderíamos calcular o *Risco Relativo*, essa medida compara a exposição entre o grupo exposto e o grupo controle. Para o caso de uma pesquisa caso-controle, isto é determinado na configuração da amostragem, não sendo dado probabilístico e que não podemos mensurar.