

2^a LISTA

Nome: Filipe Costa

1

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("Distância de uma indústria " = c('Próximo a indústria',  
                                                    'Longe da indústria'),  
                    'S' = c(96, 48),  
                    'N' = c(64, 192),  
                    check.names = FALSE)  
  
table2 |>  
  dplyr::mutate(Total = S + N)|>  
  janitor::adorn_totals("row")|>  
  kableExtra::kbl(booktabs = T, align = "cccc")|>  
  kableExtra::add_header_above(c(" ", 'Chance de contaminação' = 2, " "))
```

Distância de uma indústria	Chance de contaminação		Total
	S	N	
Próximo a indústria	96	64	160
Longe da indústria	48	192	240
Total	144	256	400

O estudo apresentado pode ser indicado como deliniameto **transversal**, a característica deste procedimento é captação dos dados no pontos de coleta num respectivo período no tempo e, aparetamente, consiste no que foi realizado na pesquisa, selecionado pontos para retirada de amostras de água.

Não há indicação de restrição do tamanho da amostra, foram selecionados 400 pontos independente da proximidade de insdustria ou não. O cálculo da distância para uma fábrica foi medida depois. Pela **liberdade** da amostragem, temos um modelo probabilístico baseado no **Produto Poisson**

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 indústria**
- 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 χ^2 que, resumidamente, compara o valor registrado na tabela com o valor esperado. A diferença dos valores influencia no resultado.

Usando a função em R, temos:

```
mat = matrix(c(96, 48, 64, 192), 2,2, byrow = T )
chisq.test(mat)
```

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 apresentado, temos evidências para afirmar que a distância para uma fábrica influencia na qualidade da água.

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