

Sumário

| | Página |
|---|--------|
| 1 Desenvolvimento | 2 |
| 1.1 Estimação da proporção média menor que 75% de participação | 2 |
| 1.2 As notas de Língua Portuguesa e Matemática são normalmente distribuídas | 2 |
| 1.3 Verificando se existe associação entre algumas variáveis | 3 |
| 1.3.1 Associação entre região e categoria administrativa | 3 |
| 1.3.2 Associação entre tamanho da escola e tamanho do município | 5 |

pacote 'xlsx' desempacotado com sucesso e somas MD5 verificadas

Os pacotes binários baixados estão em

C:\Users\felip\AppData\Local\Temp\RtmpyUOYJZ\downloaded_packages

| Estatística | Valor |
|---------------|--------|
| Média | 187,08 |
| Desvio Padrão | 23,91 |
| Variância | 571,70 |
| Mínimo | 129,97 |
| 1º Quartil | 170,76 |
| Mediana | 184,64 |
| 3º Quartil | 206,14 |
| Máximo | 251,18 |

1 Desenvolvimento

1.1 Estimação da proporção média menor que 75% de participação

Esta análise tem como intenção estimar a proporção de todas as escolas que obtiveram uma participação média menor que 95% em todas as provas realizadas. Para isso foi utilizada a variável “Participação”, sendo esta quantitativa contínua. Para a construção do intervalo de confiança, foi suposto que ambas as notas seguem uma distribuição normal.

O intervalo de confiança de 95% obtido para a proporção média menor que 75% de participação foi $p \in (0,04; 0,11)$ para a amostra com 200 observações e $p \in (0; 0,16)$ para a amostra com 50 observações.

1.2 As notas de Língua Portuguesa e Matemática são normalmente distribuídas

Nessa análise tem o intuito descobrir se as notas de Língua Portuguesa e Matemática aderem ao modelo de distribuição normal. Foram utilizados para a confecção desta análise as variáveis “Nota Língua Portuguesa” e “Nota Matemática”, ambas quantitativas contínuas. Para isso foi necessário estimar a média e o desvio padrão de ambas as matérias, sendo estas, 187,08 de média e 23,91 de desvio padrão para Língua Portuguesa para a amostra de 200 observações; 186,97 de média e 26,37 de desvio padrão para Língua Portuguesa para a amostra de 50 observações; 206,78 de média e 27,97 de desvio padrão para Matemática para a amostra de 200 observações; 208,23 de média e 29,60 de desvio padrão para Matemática para a amostra de 50 observações. Para isso foi feito os testes de hipótese: 1º) A Nota de Língua Portuguesa adere a distribuição normal?; 2º) A Nota de Matemática adere a distribuição normal?

Primeiramente serão analisadas as amostras com 200 observações.

Hipóteses primeiro teste: H_0) Nota de Língua Portuguesa segue o modelo de distribuição normal; H_1) Nota de Língua Portuguesa não segue o modelo de distribuição normal.

Estatística do Teste: $X^2 = \sum_{i=1}^6 ((O_i - E_i)^2) \div E_i$, Qui-Quadrado com 5 graus de liberdade.

Região Crítica para $\alpha=5\%$: $X_1^2 < 0,8312$ e $X_2^2 > 12,8325$

Conclusão: Como o valor obtido pela estatística do teste foi de $X^2 = 5,2896$, H_0 não será rejeitada. Utilizando o p-valor, a conclusão é a mesma, obtendo um valor de 0,3816, ainda não rejeitando H_0 .

Hipóteses segundo teste: H_0) Nota de Matemática segue o modelo de distribuição

normal; H_1) Nota de Matemática não segue o modelo de distribuição normal.

Estatística do Teste: $X^2 = \sum_{i=1}^7 ((O_i - E_i)^2) \div E_i$, Qui-Quadrado com 6 graus de liberdade.

Região Crítica para $\alpha=5\%$: $X_1^2 < 1,2373$ e $X_2^2 > 14,4494$

Conclusão: O valor obtido para estatística do teste foi de $X^2 = 4,1033$, H_0 não será rejeitada. O p-valor também aponta para a não rejeição de H_0 , com o valor de 0,6627.

Será analisado de forma semelhante as amostras de 50 observações.

Hipóteses primeiro teste: H_0) Nota de Língua Portuguesa segue o modelo de distribuição normal; H_1) Nota de Língua Portuguesa não segue o modelo de distribuição normal.

Estatística do Teste: $X^2 = \sum_{i=1}^5 ((O_i - E_i)^2) \div E_i$, Qui-Quadrado com 4 graus de liberdade.

Região Crítica para $\alpha=5\%$: $X_1^2 < 0,4844$ e $X_2^2 > 11,1433$

Conclusão: O valor obtido para estatística do teste foi de $X^2 = 3,5847$, H_0 não será rejeitada. O p-valor também aponta para a não rejeição de H_0 , com valor de 0,4651.

Hipóteses segundo teste: H_0) Nota de Matemática segue o modelo de distribuição normal; H_1) Nota de Matemática não segue o modelo de distribuição normal.

Estatística do Teste: $X^2 = \sum_{i=1}^6 ((O_i - E_i)^2) \div E_i$, Qui-Quadrado com 5 graus de liberdade.

Região Crítica para $\alpha=5\%$: $X_1^2 < 0,8312$ e $X_2^2 > 12,8325$

Conclusão: A estatística do teste obteve um valor de $X^2 = 1,6651$, assim não rejeitando H_0 . O p-valor também aponta para a não rejeição de H_0 , com um valor de 0,8933.

1.3 Verificando se existe associação entre algumas variáveis

Esta análise tem como intenção verificar se existe associação entre algumas variáveis presentes na amostra. Sendo assim, esta análise será segmentada em duas a fim de facilitar a compreensão e a leitura.

1.3.1 Associação entre região e categoria administrativa

A fim de verificar a existência de uma associação entre as variáveis “Região”, qualitativa nominal, e “Categoria administrativa”, qualitativa nominal, foram feitos gráficos a seguir.

Figura 1: Gráfico de colunas da região pela categoria administrativa para amostra de 200

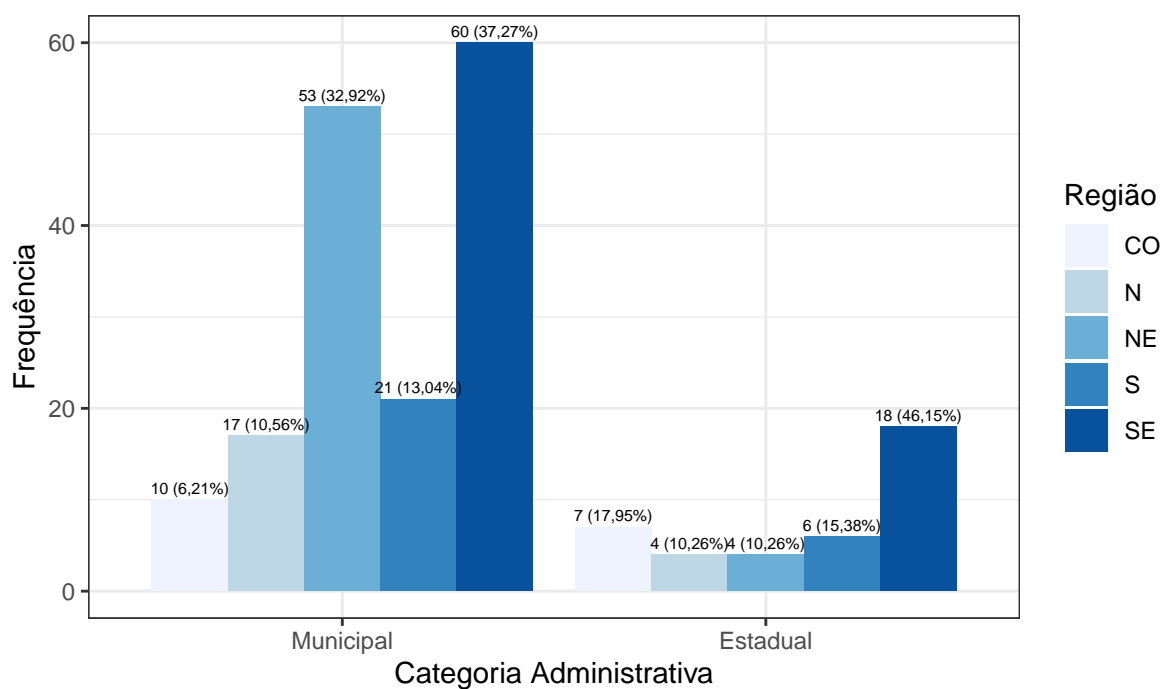
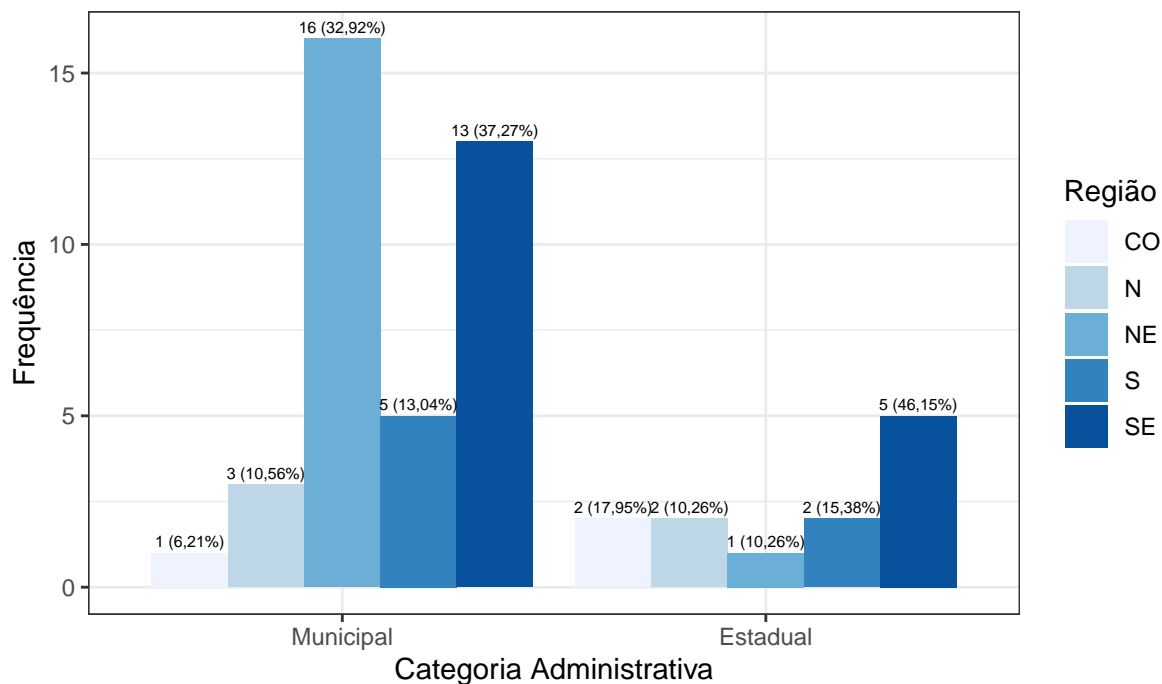


Figura 2: Gráfico de colunas da região pela categoria administrativa para amostra de 50



Como pode ser observado na **Figura 1** e **Figura 2**, existem mais observações presentes na categoria “Municipal”, totalizando 161 observações, mais que o quádruplo da categoria “Estadual”. Vale a pena ser ressaltado que ambas categorias, a região

“Sudeste” é a que apresenta mais observações. Na categoria “Municipal”, pode-se observar uma grande quantidade de observações na região “Nordeste” com um total de 53. Para verificar a associação entre as variáveis foi utilizado o Coeficiente de Contingência Modificado, este que possui um valor entre 0 e 1 com valores próximos de 0 demonstrando uma associação fraca e valores próximos de 1 uma associação forte. Na análise em questão, obteve um valor de $C^*=0,33$ para a amostra de 200, apresentando uma associação de fraca a moderada entre as variáveis e $C^*=0,26$ para a amostra de 50, apresentando uma associação de fraca a moderada entre as variáveis, mas menor que a amostra de 200.

1.3.2 Associação entre tamanho da escola e tamanho do município

A fim de verificar a existência de uma associação entre as variáveis “Tamanho do Município”, qualitativa ordinal, e “Tamanho da Escola”, qualitativa ordinal, foram feitos os gráficos a seguir.

Figura 3: Gráfico de colunas do tamanho da escola pelo tamanho do município para amostra de 200

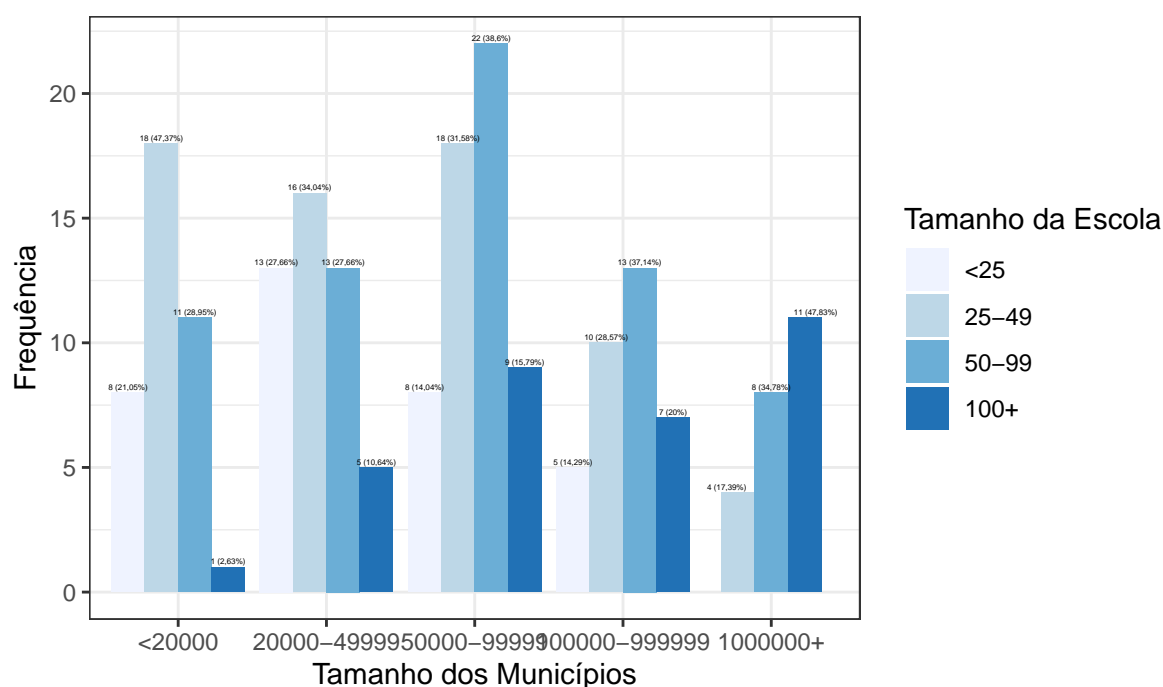
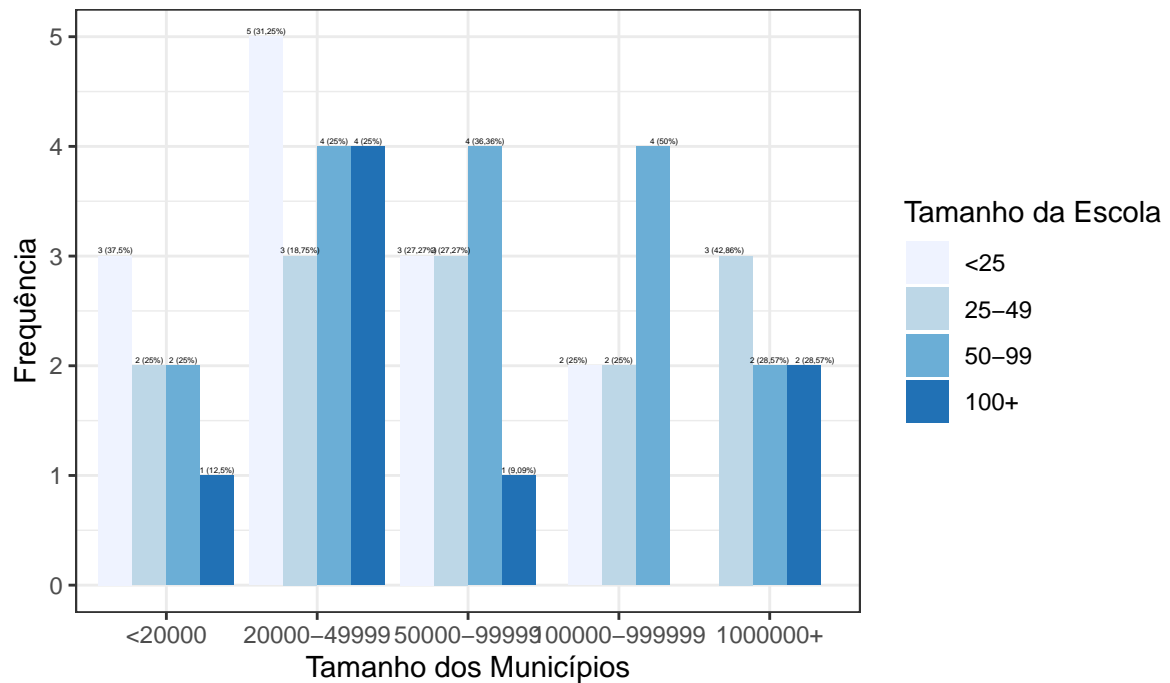


Figura 4: Gráfico de colunas do tamanho da escola pelo tamanho do município para amostra de 50



Observando a **Figura 3** e a **Figura 4** é possível observar existem mais escolas com tamanho “50000-99999” na amostra de 200, com um total 57 observações. Contudo, na amostra de de 50, existem mais escolas com tamanho “20000-49999”, com um valor de 16. Para verificar a associação entre as variáveis foi utilizado o Coeficiente de Contingência Modificado, este que possui um valor entre 0 e 1 com valores próximos de 0 demonstrando uma associação fraca e valores próximos de 1 uma associação forte. Na análise em questão, obteve um valor de $C^*=0,44$ para a amostra de 200, apresentando uma associação a moderada entre as variáveis e $C^*=0,23$ para a amostra de 50, apresentando uma associação de fraca a moderada entre as variáveis, sendo menor que a amostra de 200.