

ANÁLISE INFERENCIAL DE DADOS: ENTENDENDO ALGUNS TESTES ESTATÍSTICOS

Marco Antonio Paes de Souza¹;

Universidade Estadual do Sudoeste da Bahia (PPGEd/UESB), Vitória da Conquista, Bahia.

<https://orcid.org/0000-0003-1054-897X>

Vanilda Batista Ribeiro²;

Universidade Estadual do Sudoeste da Bahia (PPGEd/UESB), Vitória da Conquista, Bahia.

<https://orcid.org/0000-0002-7858-3248>

Zildete Soares Aranha Azevêdo³;

Universidade Estadual do Sudoeste da Bahia (PPGEd/UESB), Vitória da Conquista, Bahia.

<https://orcid.org/0000-0001-8702-2912>

Cláudio Pinto Nunes⁴;

Universidade Estadual do Sudoeste da Bahia (PPGEd/UESB), Vitória da Conquista, Bahia.

<https://orcid.org/0000-0003-1514-6961>

Berta Leni Costa Cardoso⁵;

Universidade do Estado da Bahia (PPGEDuF/UNEB), Guanambi, Bahia.

Universidade Estadual do Sudoeste da Bahia (PPGEd/UESB), Vitória da Conquista, Bahia.

<https://orcid.org/0000-0001-7697-0423>

Luiz Humberto Rodrigues Souza⁶.

Universidade do Estado da Bahia (UNEB), Guanambi, Bahia.

Universidade Estadual do Sudoeste da Bahia (PPGEd/UESB), Vitória da Conquista, Bahia.

<https://orcid.org/0000-0001-9237-3928>

RESUMO: A análise quantitativa é o processo de aplicação de técnicas estatísticas e lógicas para avaliar informações obtidas a partir de um conjunto de dados, e assim tomar decisões mais assertivas. Nesse sentido, o objetivo do estudo foi descrever os conceitos básicos da análise inferencial e apresentar um roteiro para a execução de alguns testes estatísticos. Foi realizada uma pesquisa descritiva em que a organização das informações se baseou nos princípios da pesquisa bibliográfica. Inicialmente, foi apresentado o conceito de distribuição normal, que é um pressuposto teórico de muitos testes estatísticos. Em seguida, foi apresentada a diferença entre os testes paramétricos e não paramétricos.

Na sequência, foi necessário definir hipótese nula (previsão de que a manipulação experimental não tem efeito ou que as variáveis do estudo não se relacionam entre si), nível de significância (probabilidade de rejeitar uma hipótese nula quando ela é verdadeira) e valor-p (probabilidade de que o teste estatístico tenha um valor igual ou mais extremo em relação ao valor observado na amostra, assumindo que a hipótese nula é verdadeira). Essas informações permitiram a interpretação da tomada de decisão dos testes estatísticos. Os três testes estatísticos exemplificados foram: teste-t de *Student* para amostras independentes, correlação de Pearson e teste qui-quadrado.

PALAVRAS-CHAVE: Educação. Pesquisa Quantitativa. Testes Estatísticos.

INFERENCEAL DATA ANALYSIS: UNDERSTANDING SOME STATISTICAL TESTS

ABSTRACT: Quantitative analysis is the process of applying statistical and logical techniques to evaluate information obtained from a set of data, and thus make more assertive decisions. In this sense, the objective of the study was to describe the basic concepts of inferential analysis and present a guide for executing some statistical tests. A descriptive study was carried out in which the organization of the information was based on the principles of bibliographic research. Initially, the concept of normal distribution was presented, which is a theoretical assumption of many statistical tests. Next, the difference between parametric and nonparametric tests was presented. Next, it was necessary to define the null hypothesis (prediction that the experimental manipulation has no effect or that the study variables are not related to each other), the significance level (probability of rejecting a null hypothesis when it is true) and p-value (probability that the statistical test has a value equal to or more extreme in relation to the value observed in the sample, assuming that the null hypothesis is true). This information allowed the interpretation of the decision-making of the statistical tests. The three statistical tests exemplified were: Student's t-test for independent samples, Pearson's correlation and chi-square test.

KEY-WORDS: Education. Quantitative Research. Statistical Tests.

INTRODUÇÃO

Conforme Gatti (2004), na área de pesquisa educacional, com exceção às análises de dados das avaliações de rendimento escolar, poucos estudos utilizam metodologias quantitativas. Contudo, ainda segundo a autora, há problemas educacionais que precisam ser qualificados através de dados quantitativos.

Outro problema trazido por Gatti (2004) é que:

Os estudos quantitativos em educação, especialmente os que se utilizam de técnicas de análise mais sofisticadas, mais flexíveis e mais robustas, não são realizados por educadores, mas por pesquisadores de outras áreas que se debruçam sobre o objeto da educação. Com isto, interpretações e teorizações nem sempre incorporam as discussões em pauta no campo das reflexões sobre a educação (Gatti, 2004, p. 14).

Ademais, Pereira e Ortigão (2016) salientaram que é de interesse geral que os profissionais da educação se apoderem da capacidade de lidar com dados quantitativos, pois esse método pode ser muito útil na compreensão de diversos problemas educacionais.

A análise quantitativa é o processo de aplicação de técnicas estatísticas e lógicas para avaliar informações obtidas a partir de um conjunto de dados, e assim tomar decisões mais assertivas (Thomas; Nelson; Silverman, 2012). A pesquisa científica é o documento acadêmico que o pesquisador transforma os dados em informação e responde ao questionamento que deu início à pesquisa. Portanto, uma análise de dados bem realizada garante a qualidade e o rigor científico de um estudo, assegurando conclusões condizentes com a realidade pesquisada. Nesse sentido, o objetivo do estudo foi descrever os conceitos básicos da análise inferencial e apresentar um roteiro para a execução de alguns testes estatísticos.

METODOLOGIA

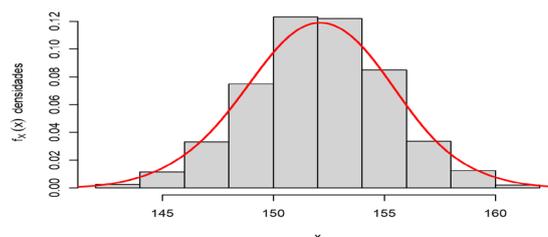
Trata-se de uma pesquisa descritiva em que a organização das informações se baseou nos princípios da pesquisa bibliográfica, usando como referência as obras de Bussab e Morettin (2017), Field (2009) e Vieira (2012) e os artigos científicos de Hirakata, Mancuso e Jezus Castro (2019) e Mukaka (2012).

RESULTADOS

Distribuição normal

A distribuição de frequências ou histograma é um gráfico contendo os valores observados no eixo horizontal, com barras mostrando quantas vezes cada valor ocorreu no conjunto de dados. Em um contexto ideal, os dados estariam distribuídos simetricamente em volta do centro de todos os escores. Isso é conhecido como distribuição normal e é caracterizado por uma curva em forma de sino (Figura 1). Essa forma basicamente indica que a maioria dos escores está em torno do centro da distribuição, e à medida que se distancia do centro, as barras ficam menores, sugerindo que sua frequência diminuiu (Field, 2009).

Figura 1: Distribuição normal.



Fonte: <https://www.inf.ufsc.br/~andre.zibetti/probabilidade/normal.html>

Testes paramétricos e não paramétricos

Os testes paramétricos são métodos estatísticos em que os dados numéricos seguem uma distribuição normal (Vieira, 2012). Além disso, a maioria dos testes paramétricos baseados na distribuição normal deve preencher quatro hipóteses: dados normalmente distribuídos (assume-se que os dados foram obtidos de uma ou mais populações normais), homogeneidade das variâncias (as variâncias devem ser as mesmas para as diferentes populações consideradas), dados por intervalo (os dados devem ser mensurados pelo menos em nível ordinal) e independência (os dados de participantes diferentes são independentes, ou seja, o comportamento de um participante não influencia o comportamento de outro) (Field, 2009).

Por outro lado, os testes não paramétricos são métodos estatísticos que não assumem uma distribuição específica para os dados. Trata-se de testes mais flexíveis que podem ser usados quando os dados não seguem uma distribuição normal ou quando os dados são ordinais ou categóricos (Vieira, 2012).

Hipóteses

A literatura descreve, em regra, duas hipóteses para os estudos quantitativos inferenciais: a hipótese nula e a hipótese alternativa, também conhecida como hipótese experimental. A hipótese nula corresponde à previsão de que a manipulação experimental não tem efeito ou que as variáveis do estudo não se relacionam entre si, enquanto a hipótese alternativa preconiza que a manipulação experimental tem efeito ou que as variáveis do estudo se relacionam entre si (Field, 2009). Em síntese, a hipótese nula afirma que não existe efeito sobre um fenômeno investigado. O critério mais utilizado para rejeitar ou não uma hipótese nula é o valor-p.

Nível de significância

Na estatística, o nível de significância de um estudo é representado pelo símbolo grego alfa (α), e é definido como a probabilidade da estimativa de um parâmetro estatístico em uma população estar fora do intervalo de confiança (Bussab; Morettin, 2017). Simplificando:

o nível de significância é a probabilidade de rejeitar uma hipótese nula quando ela é verdadeira. Mais adiante, veremos que o nível de significância representa a probabilidade de se cometer o erro tipo I.

O nível de significância marca o limite para determinar se um resultado de uma pesquisa é estatisticamente significativo ou não, de modo que se o valor-p for menor que o nível de significância, o resultado é considerado estatisticamente significativo e a hipótese nula deve, portanto, ser rejeitada. A literatura científica, na maioria das situações, adota o nível de significância de 5%, ou seja, $\alpha = 0,05$.

O valor-p

O valor-p é a probabilidade de que o teste estatístico tenha um valor igual ou mais extremo em relação ao valor observado na amostra, assumindo que a hipótese nula é verdadeira (Hirakata; Mancuso; Jesus Castro, 2019). Em outras palavras, ele indica a probabilidade do acaso no resultado encontrado, ou seja, quanto menor o valor-p, menor é a probabilidade do acaso explicar seus resultados. Também, podemos dizer que o valor-p apenas informa o quão improvável os dados seriam se a hipótese nula fosse verdadeira ou que o valor-p apenas indica se os dados fornecem evidências para rejeitar a hipótese nula.

Dito isso, caminhemos para a regra de decisão após obter o resultado de um teste estatístico:

- Se o valor-p < nível de significância (alfa) → a hipótese nula deve ser rejeitada.
- Se o valor-p \geq nível de significância (alfa) → a hipótese nula não deve ser rejeitada.

Tomada de decisão em testes estatísticos

O Quadro 1 apresenta as possíveis decisões que podem ser tomadas após a realização de um teste estatístico.

Quadro 1: Decisão do teste.

| Decisão | Realidade | |
|------------------------------|--------------------------|--------------------------|
| | Hipótese nula verdadeira | Hipótese nula falsa |
| Não rejeitar a hipótese nula | Decisão correta | Erro tipo II (β) |
| Rejeitar a hipótese nula | Erro tipo I (α) | Decisão correta |

Fonte: adaptado de Bussab e Morettin (2017).

Mesmo que pareça repetitivo, optamos em descrever as informações do Quadro 1. O erro tipo I ocorre quando acreditamos que existe um efeito genuíno na nossa população, mas de fato, ele não existe, ou seja, rejeitamos a hipótese nula quando ela é verdadeira.

Seguindo este raciocínio, o erro tipo II ocorre quando acreditamos que não existe um efeito na população, mas, na verdade, ele existe, ou seja, não rejeitamos a hipótese nula quando ela é falsa.

Teste-t de Student para amostras independentes

É um teste estatístico paramétrico utilizado para comparar se duas médias numéricas coletadas de amostras independentes diferem de forma significativa (Field, 2009). Por exemplo: foi utilizada a versão abreviada do questionário Whoqol (Fleck *et al.*, 2000) para avaliar a qualidade de vida de vinte e um estudantes de um programa de mestrado acadêmico em educação. Além disso, foi perguntado se essas pessoas faziam o uso de medicamentos de forma contínua. A variável “qualidade de vida”, neste caso, tem natureza numérica devido ao instrumento utilizado, enquanto a variável “uso de medicamentos” é categórica. Assim, queremos saber se o uso de medicamentos exerceu efeito na qualidade de vida desses estudantes.

O próximo passo foi a definição das hipóteses do estudo:

Hipótese nula: a média da qualidade de vida dos estudantes que não usavam medicamentos é igual à média da qualidade de vida dos estudantes que usavam medicamentos ($Média_{\text{grupo não medicamento}} = Média_{\text{grupo medicamento}}$).

Hipótese alternativa: a média da qualidade de vida dos estudantes que não usavam medicamentos é diferente da média da qualidade de vida dos estudantes que usavam medicamentos ($Média_{\text{grupo não medicamento}} \neq Média_{\text{grupo medicamento}}$).

Por meio do *software* estatístico Jamovi, foram geradas a Figura 2 e a Figura 3. A Figura 2 apresenta as medidas descritivas (média, mediana, desvio padrão e erro padrão) da variável contínua “qualidade de vida” (QV) dos estudantes que não usavam medicamentos (Não) e dos estudantes que usavam medicamentos (Sim).

Figura 2: Medidas descritivas.

| Teste t para amostras independentes | | | | |
|-------------------------------------|--------------|-------------|------|-------|
| | | Estatística | gl | p |
| QV | t de Student | 2.20 | 19.0 | 0.040 |

Nota. $H_2: \mu_{\text{Não}} \neq \mu_{\text{Sim}}$

| Descritivas de Grupo | | | | | | |
|----------------------|-------|----|-------|---------|---------------|-------------|
| | Grupo | N | Média | Mediana | Desvio-padrão | Erro-padrão |
| QV | Não | 10 | 67.7 | 69.2 | 10.8 | 3.40 |
| | Sim | 11 | 56.3 | 54.7 | 12.8 | 3.87 |

Fonte: elaborado pelos autores a partir do *software* Jamovi.

Trocando em miúdos, verificamos que o grupo de estudantes que não usava medicamentos possuía uma média e desvio padrão de $67,7 \pm 10,8$ pontos, enquanto o grupo de estudantes que usava medicamentos possuía uma média e desvio padrão de $56,3 \pm 12,8$ pontos. Isso representa uma diferença de 11,4 pontos entre os grupos. E agora? Essa diferença é estatisticamente significativa? A Figura 3 apresenta o valor do teste estatístico utilizado (Estatística) e o valor-p encontrado (p).

Figura 3: Teste-t de *Student* para amostras independentes.

| Teste t para amostras independentes | | | | |
|-------------------------------------|--------------|-------------|------|-------|
| | | Estatística | gl | p |
| QV | t de Student | 2.20 | 19.0 | 0.040 |

Nota. $H_0: \mu_{\text{Não}} = \mu_{\text{Sim}}$

| Descritivas de Grupo | | | | | | |
|----------------------|-------|----|-------|---------|---------------|-------------|
| | Grupo | N | Média | Mediana | Desvio-padrão | Erro-padrão |
| QV | Não | 10 | 67.7 | 69.2 | 10.8 | 3.40 |
| | Sim | 11 | 56.3 | 54.7 | 12.8 | 3.87 |

Fonte: elaborado pelos autores a partir do *software* Jamovi.

Para finalizar o raciocínio, é preciso lembrar algumas informações: primeiro, o nível de significância adotado foi de 5% ($\alpha = 0,05$); e por último, o valor-p deve ser usado para tomar uma decisão em relação à hipótese nula do estudo. Na Figura 3, notou-se que o valor-p foi de 0,040 ou 0,04; assim, como o valor-p foi menor que o nível de significância (0,05), a hipótese nula deve ser rejeitada. Portanto, essa análise (teste-t de *Student* para amostras independentes) sugeriu que a qualidade de vida entre os grupos foi diferente, de tal modo que aqueles estudantes que não usavam medicamentos tiveram melhor qualidade de vida ($67,7 \pm 10,8$ pontos) do que aqueles que usavam medicamentos ($56,3 \pm 12,8$ pontos).

Correlação bivariada de Pearson

É um teste estatístico paramétrico utilizado para verificar se há correlação entre duas variáveis numéricas (Field, 2009). Sendo assim, este teste usa o coeficiente de correlação de Pearson (r), que é uma medida padronizada da força do relacionamento entre duas variáveis, para auxiliar a interpretação dos resultados. O coeficiente de correlação de Pearson (r) pode ter qualquer valor de -1 (à medida que uma variável muda, a outra muda na direção oposta pela mesma quantia) passando por 0 (à medida que uma variável muda, a outra não muda) até +1 (à medida que uma variável muda, a outra muda na mesma direção pela mesma quantia) (Field, 2009).

A literatura disponibilizou alguns pontos de corte para auxiliar a interpretação do “r”. Neste estudo, usamos os valores sugeridos por Mukaka (2012).

Quadro 2: Regra prática para interpretar a força de um coeficiente de correlação.

| FORÇA DA CORRELAÇÃO | INTERPRETAÇÃO |
|-------------------------|-----------------------------------|
| 0,9 a 1,0 (-0,9 a -1,0) | indica uma correlação muito forte |
| 0,7 a 0,9 (-0,7 a -0,9) | indica uma correlação forte |
| 0,5 a 0,7 (-0,5 a -0,7) | indica uma correlação moderada |
| 0,3 a 0,5 (-0,3 a -0,5) | indica uma correlação fraca |
| 0 a 0,3 (0 a -0,3) | indica uma correlação desprezível |

Fonte: adaptado de Mukaka (2012).

No exemplo da análise anterior, vamos considerar que as horas de sono também foram informadas pelos estudantes. Neste caso, queremos saber se as horas de sono estão relacionadas com a qualidade de vida dos estudantes. Assim, foi preciso definir as hipóteses do estudo:

Hipótese nula: as horas de sono não se relacionam com a qualidade de vida dos estudantes.

Hipótese alternativa: as horas de sono se relacionam com a qualidade de vida dos estudantes.

Por meio do *software* estatístico Jamovi, foi gerada a Figura 4, que apresentou o coeficiente de correlação de Pearson entre as duas variáveis e o valor-p.

Figura 4: Correlação bivariada de Pearson.

| | | Horas de sono por noite | QV |
|-------------------------|--------------|-------------------------|----|
| Horas de sono por noite | R de Pearson | — | |
| | p-value | — | |
| QV | R de Pearson | 0.508 | — |
| | p-value | 0.019 | — |

Fonte: elaborado pelos autores a partir do *software* Jamovi.

Mais uma vez, é preciso lembrar que o nível de significância adotado foi de 5% ($\alpha = 0,05$) e o valor-p foi usado para tomar uma decisão em relação à hipótese nula do estudo. Na Figura 4, nota-se que o valor-p foi de 0,019; assim, como o valor-p foi menor que o nível de significância (0,05), a hipótese nula deve ser rejeitada. Portanto, a análise de correlação de Pearson sugeriu que as horas de sono se relacionaram significativamente com a qualidade de vida dos estudantes. Em outras palavras, os estudantes que dormiram

mais durante a noite também foram aqueles que apresentaram maior qualidade de vida. Essa interpretação foi possível pois o valor “r” foi positivo, ou seja, as grandezas possuíram relação diretamente proporcional. Quanto à interpretação do coeficiente de correlação de Pearson, podemos concluir que o valor de 0,508 indicou uma relação moderada entre as variáveis analisadas (olhe os valores de referência no quadro 2).

Teste qui-quadrado de Pearson

É um teste não paramétrico utilizado para examinar se duas variáveis categóricas dispostas em uma tabela de contingência estão associadas. Trata-se de uma estatística baseada na ideia de comparar as frequências observadas em certas categorias com as frequências esperadas nessas mesmas categorias (Vieira, 2012).

Nas tabelas de contingência 2x2, todas as frequências esperadas devem ser maiores do que 5, pois senão o teste pode falhar em detectar um efeito genuíno. Quando isso acontece, é recomendado usar o teste exato de Fisher.

Vamos continuar com o exemplo em que o uso de medicamentos e a qualidade de vida dos estudantes de mestrado foram avaliados. No entanto, utilizamos a mediana do conjunto de dados para categorizar os estudantes em duas classes: os estudantes que obtiveram a qualidade de vida inferior à mediana pertencem à classe de “baixa qualidade de vida” e os estudantes que obtiveram a qualidade de vida igual ou superior à mediana pertencem à classe de “alta qualidade de vida”. Portanto, agora há duas variáveis categóricas que podem ser testadas à luz do teste qui-quadrado para saber se há associação entre elas.

Como dito, a estatística qui-quadrado testa se as duas variáveis categóricas são dependentes. Se o valor-p for menor que o nível de significância, rejeitamos a hipótese nula de que as variáveis são independentes e aceitamos a hipótese de que elas estão associadas. Ou podemos definir as hipóteses como de costume:

Hipótese nula: não há associação entre as variáveis (ou seja, elas são independentes).

Hipótese alternativa: há associação entre as variáveis (ou seja, elas são dependentes).

Por meio do *software* estatístico Jasp, foi gerada a Figura 5 que sumarizou os valores observados (*count*), valores esperados (*expected count*) e os valores residuais ajustados (*standartized residuals*) na tabela de contingência.

Figura 5: Tabela de contingência 2x2.

| Medicamento | | QVida | | Total |
|-------------|------------------------|--------|--------|--------|
| | | Baixa | Alta | |
| Sim | Count | 8.000 | 3.000 | 11.000 |
| | Expected count | 5.238 | 5.762 | 11.000 |
| | Standardized residuals | 2.416 | -2.416 | |
| Não | Count | 2.000 | 8.000 | 10.000 |
| | Expected count | 4.762 | 5.238 | 10.000 |
| | Standardized residuals | -2.416 | 2.416 | |
| Total | Count | 10.000 | 11.000 | 21.000 |
| | Expected count | 10.000 | 11.000 | 21.000 |

Chi-Squared Tests

| | Value | df | p |
|----------------|-------|----|-------|
| X ² | 5.838 | 1 | 0.016 |
| N | 21 | | |

Fonte: elaborado pelos autores a partir do *software* Jasp.

De maneira descritiva, sabe-se que oito pessoas usavam medicamentos e apresentaram qualidade de vida baixa, três pessoas usavam medicamentos e apresentaram qualidade de vida alta, duas pessoas não usavam medicamentos e apresentaram qualidade de vida baixa e oito pessoas não usavam medicamentos e apresentaram qualidade de vida alta. E agora? Existe associação entre o uso/não uso de medicamentos e alta/baixa qualidade de vida? Embora em uma célula (categoria) foi observada frequência esperada menor que 5, vamos seguir a apresentação dos resultados com o teste qui-quadrado e não o teste exato que Fisher (seria a alternativa adequada).

A Figura 6 apresentou o valor do teste qui-quadrado (X²) e o valor-p para auxiliar a tomada de decisão. Observou-se que o valor-p (0,016) foi menor que o nível de significância (0,05), sugerindo que as variáveis foram associadas (a hipótese nula foi rejeitada).

Figura 6: Teste qui-quadrado de Pearson.

| Medicamento | | QVida | | Total |
|-------------|------------------------|--------|--------|--------|
| | | Baixa | Alta | |
| Sim | Count | 8.000 | 3.000 | 11.000 |
| | Expected count | 5.238 | 5.762 | 11.000 |
| | Standardized residuals | 2.416 | -2.416 | |
| Não | Count | 2.000 | 8.000 | 10.000 |
| | Expected count | 4.762 | 5.238 | 10.000 |
| | Standardized residuals | -2.416 | 2.416 | |
| Total | Count | 10.000 | 11.000 | 21.000 |
| | Expected count | 10.000 | 11.000 | 21.000 |

Chi-Squared Tests

| | Value | df | p |
|----------------|-------|----|-------|
| X ² | 5.838 | 1 | 0.016 |
| N | 21 | | |

Fonte: elaborado pelos autores a partir do *software* Jasp.

Mas, quais classes (células ou categoria) estiveram associadas? É necessário olhar os valores residuais ajustados (*standartized residuals*) na Figura 5. Eles sinalizam uma relação diretamente proporcional (sinal positivo) entre uso de medicamentos e qualidade de vida baixa e entre o não uso de medicamentos e qualidade de vida alta.

Há mais informações que podem ser extraídas dos valores residuais ajustados. No teste qui-quadrado de Pearson também é possível realizar outras subanálises. Mas isso, extrapolaria o objetivo desse capítulo de livro.

CONCLUSÃO

Este estudo apresentou conceitos importantes sobre análise inferencial de dados, como distribuição normal, testes paramétricos e não paramétricos, hipótese nula, nível de significância e valor-p. Além disso, foi apresentado o conceito de três testes estatísticos (teste-t de *Student* para amostras independentes, correlação de Pearson e teste qui-quadrado) para auxiliar a interpretação dos resultados e a tomada de decisão.

DECLARAÇÃO DE INTERESSES

Nós, autores deste artigo, declaramos que não possuímos conflitos de interesses de ordem financeira, comercial, político, acadêmico e pessoal.

REFERÊNCIAS

- BUSSAB, W. O.; MORETTIN, P. A. **Estatística básica**. 9. ed. São Paulo: Saraiva, 2017.
- FIELD, A. **Descobrimo a estatística usando o SPSS**. 2. ed. Porto Alegre: Artmed, 2009.
- FLECK, M. P.; LOUZADA, S.; XAVIER, M.; CHACHAMOVICH, E.; VIEIRA, G.; SANTOS, L.; PINZON, L. Aplicação da versão em português do instrumento abreviado de avaliação da qualidade de vida WHOQOL-bref. **Revista de Saúde Pública**, v. 34, n. 2, p. 178-83, 2000.
- GATTI, B. A. Estudos quantitativos em educação. **Educação e Pesquisa**, São Paulo, v.30, n.1, p. 11-30, 2004.
- HIRAKATA, V. N.; MANCUSO, A. C. B.; JESUS CASTRO, S. M. Teste de hipóteses: perguntas que você sempre quis fazer, mas nunca teve coragem. **Clinical and Biomedical Research**, v. 39, n. 2, 2019.
- MUKAKA, M. M. A guide to appropriate use of correlation coefficient in medical research. **Malawi Medical Journal**, v. 24, n. 3, p. 69-71, 2012.
- PEREIRA, G.; ORTIGÃO, M. I. R. Pesquisa quantitativa em educação: algumas considerações. **Periferia**, v. 8, n. 1, p. 66-79, 2016.
- THOMAS, J. R.; NELSON, J. K.; SILVERMAN, S. J. **Métodos de pesquisa em atividade física**. 6. ed. Porto Alegre: Artmed, 2012.
- VIEIRA, S. **Estatística básica**. São Paulo: Cengage Learning, 2012.