UNIVERSIDADE FEDERAL DO AMAPÁ
PRÓ-REITORIA DE ENSINO DE GRADUAÇÃO
CURSO DE LICENCIATURA EM MATEMÁTICA

KELLEN CRISTIANE SILVA DE LIMA

ESTATÍSTICA BÁSICA APLICADA À EDUCAÇÃO: UMA ABORDAGEM SOBRE O ENSINO E APRENDIZAGEM NA CONSTRUÇÃO DE ÂNGULOS

Macapá
2013

KELLEN CRISTIANE SILVA DE LIMA

ESTATÍSTICA BÁSICA APLICADA À EDUCAÇÃO: UMA ABORDAGEM SOBRE O ENSINO E APRENDIZAGEM NA CONSTRUÇÃO DE ÂNGULOS

Trabalho de Conclusão de Curso apresentado ao corpo docente do Curso de Licenciatura Plena em Matemática da UNIFAP, como requisito parcial para a obtenção da graduação de Licenciatura em Matemática.

Orientador: Prof. Espec. João Socorrro Pinheiro Ferreira.

Macapá
2013

ESTATÍSTICA BÁSICA APLICADA À EDUCAÇÃO: UMA ABORDAGEM SOBRE O ENSINO E APRENDIZAGEM NA CONSTRUÇÃO DE ÂNGULOS

por

LIMA, Kellen Cristiane Silva de

Este Trabalho de Conclusão de Curso, foi julgado e aprovado, pelo Corpo Docente do Curso de Licenciatura em Matemática - UNIFAP, como requisito parcial para a obtenção da Graduação de Licenciatura em Matemática.

Macapá, 25 de janeiro de 2013

Banca Examinadora

Orientador: Prof. Espec. João Socorro Pinheiro Ferreira.
Universidade Federal do Amapá - UNIFAP

Membro: Prof. Dr. Gúzman Eulálio Isla Chamilco.
Universidade Federal do Amapá - UNIFAP

Membro: Prof. Dr. José Walter Cárdenas Sotil.
Universidade Federal do Amapá - UNIFAP

Dedico este trabalho a minha família pelo apoio incondicional, mesmo nos momentos de ausência, amigos e ao meu Professor Orientador João Socorro Pinheiro Ferreira, pela forma profissional com que conduziu as orientações para a elaboração do trabalho.

AGRADECIMENTOS

A Deus:

Essa vitória Senhor, jamais poderia ser só minha, é acima de tudo Tua, e por isso mesmo peço que o teu amor me ilumine e me dê força para que exerça a minha profissão com dignidade.

"A aprendizagem significativa, parte do compartilhamento de experiências e não apenas da transmissão de conhecimentos."

Saint-Simon (Século XVII)

RESUMO

Este Trabalho de Conclusão de Curso foi desenvolvido sobre o ensino e aprendizagem de construção de ângulos agudos, retos e obtusos, com o auxílio de transferidor. A pesquisa de campo foi realizada em uma escola estadual de ensino fundamental, localizada no município de Macapá, capital do Estado do Amapá. Participaram das atividades propostas, trinta alunos do sétimo ano (antiga sexta série), organizados em quinze duplas, além dos autores desta pesquisa - a acadêmica e o seu orientador de TCC - e do professor de Matemática da turma. Foi realizada uma intervenção pedagógica em setembro de 2012, em que consistia à aplicação de quatro atividades, sendo que a primeira era sobre a construção de ângulos agudos e a segunda à construção de ângulos obtusos, ambas com o auxílio de um transferidor. A terceira e quarta atividades apresentadas as duplas foram extraídas do livro texto adotado pela turma. A terceira era para identificar se os ângulos desenhados no livro seriam: agudo, reto ou obtuso. A quarta atividade consistia em responder perguntas também sobre os mesmos tipos de ângulos, quanto a sua abertura. Procuraram-se neste trabalho correlacionar as duas primeiras atividades, sendo que a primeira foi designada de variável X (ângulos agudos) e a segunda de variável Y (ângulos obtusos). A partir dessas definições, fez-se uma abordagem estatística aprofundada e após exaurir-se todas as análises dos dados, chegou-se à conclusão de que o aprendizado sobre ângulo obtuso (variável dependente) a partir do conhecimento científico sobre ângulo agudo (variável independente) foi regular, isto porque os resultados obtidos apontam que cinquenta e cinco e meio por cento das duplas apresentam um desempenho insatisfatório. Esta conclusão foi corroborada em função do coeficiente de correlação linear de Pearson obtido entre as variáveis estudadas ser de aproximadamente 0.76 e segundo fontes bibliográficas indicadas no texto está classificado como forte. A terceira atividade (variável W) foi analisada estatisticamente independente das demais e seus resultados estão registrados no trabalho. A quarta e última atividade (variável Z), foi correlacionada com as duas primeiras, ocorrendo então uma análise estatística bidimensional, e os seus resultados são relativamente satisfatórios. Com isso, pretende-se destacar a importância da estatística básica como indicador da assimilação de aprendizagem de geometria plana e poder contribuir para a melhoria de qualidade de ensino na turma pesquisada.

Palavras chaves: Educação Matemática. Ângulos. Educação Estatística. Qualidade de Ensino.

ABSTRACT

This Labor Completion of Course was developed on the teaching and learning of building sharp angles, straight and obtuse, with the help of protractor. The field research was conducted in a public school elementary school located in the city of Macapá, capital of the state of Amapá. Participants proposed activities thirty seventh year students (former sixth grade), organized into fifteen pairs, besides the authors of this research - and his academic advisor TCC - and the math teacher's class. Educational intervention was performed in September 2012, that consisted in the application of four activities, the first of which was about building and second acute angles, obtuse angles of the construction, both with the aid of a protractor. The third and fourth doubles the activities presented were taken from the textbook adopted by the class. The third was to identify whether the angles drawn in the book are: acute, obtuse or straight. The fourth activity was to answer questions also about the same kinds of angles, as its opening. Sought in this work to correlate the first two activities, the first of which was designated variable X (sharp angles) and the second variable Y (obtuse angles). From these definitions, we did a thorough statistical approach and after exhausting themselves all analyzes of the data, it was concluded that learning about obtuse angle (dependent variable) from the scientific knowledge about acute angle (independent variable) is regular, ie because the results show that fifty-five and a half percent of the pairs have unsatisfactory performance. This conclusion was corroborated on the coefficient of Pearson linear correlation obtained between variables to be $r_{XY}0.76$ seconds and bibliographical sources indicated in the text is classified as strong. The third activity (W variable) was analyzed statistically independent of the others and their results are recorded in the job. The fourth and last activity (variable Z), was correlated with the first two, then going on a statistical analysis of two-dimensional, and its results are relatively satisfactory. This is intended to highlight the importance of basic statistics as an indicator of assimilation learning plane geometry and to contribute to the improvement of teaching quality in the class researched.

Keywords: Mathematics Education. Angles. Education Statistics. Teaching Quality.

LISTA DE FIGURAS

LISTA DE TABELAS

SUMÁRIO

INTRODUÇÃO

Neste trabalho, realizou-se um estudo sobre o ensino e aprendizagem de ângulos, particularmente a sua construção com o auxílio de transferidor, cujo objetivo geral seria aferir o quanto a turma asssimilou o conteúdo ministrado em sala de aula. Para atingir tal meta, utilizou-se como ferramenta principal a estatística básica, destacando a sua importância como indicador do grau de assimilação de conteúdo ministrado para alunos do sétimo ano do ensino fundamental.

No primeiro capítulo, fez-se uma abordagem aos conceitos básicos que dão suporte científico à validade desta importante Ciência. Quando fez-se necessário exemplificar algumas definições, utilizo-se as respostas fornecidas pelas duplas como dados para construir tabelas, plotar gráficos - com auxílio do *Excel* - e efetuar cálculos estatísticos (tabulação).

No segundo capítulo, são apresentadas as atividades que foram aplicadas a turma, sendo que a primeira era para construir três exemplos de ângulos agudos, a segunda para construir três exemplos de ângulos obtusos, a terceira para identificar e classificar os ângulos desenhados no livro texto da turma, se os mesmos eram agudos, retos ou obtusos e a quarta atividade era para responder sobre os tipos de ângulos. Para analisar as quatro atividades, adotamos a seguinte metodologia: as duas primeiras foram correlacionadas entre si, a terceira atividade foi analisada individualmente e a quarta, correlacionamo-a, separadamente, com a primeira e depois com a segunda, com o objetivo de saber com qual delas a correlação seria mais significativa.

No decorrer do texto são apresentadas tabelas, gráficos construídos com o auxílio do *winplot* e também da planilha *Excel* e exemplos como forma de torná-lo o mais didático possível.

Ao final consegui-se atingir os objetivos propostos neste trabalho, utilizando como base de dados, a pesquisa de campo realizada na escola. Serviu também para aplicar os conhecimentos adquiridos ao cursar a disciplina Estatística e as demais do currículo de Matemática.

Capítulo 1

CONCEITOS BÁSICOS

Neste capítulo iremos descrever os principais conceitos fundamentais de estatística básica, pois acredita-se que os mesmos são de suma importância para o desenvolvimento deste texto. Se por acaso, alguma definição deixar de ser incluída nele, far-se-á quando necessário.

1.1 Tipos de variáveis

Algumas variáveis, como sexo, educação, estado civil, apresentam como possíveis realizações uma qualidade (ou atributo) do indivíduo pesquisado, são denominadas *qualitativas*, ao passo que outras, como número de filhos, salário, idade, numero de acertos de questões de uma prova, apresentam como possíveis realizações números resultantes de uma contagem ou mensuração, são chamadas de *quantitativas*.

Neste trabalho, a coleta de dados ocorreu durante uma intervenção pedagógica sobre o ensino e aprendizagem de geometria plana, mais especificamente, na construção de ângulos agudos, reto e obtusos, com a utilização de transferidor, desenvolvido por alunos do sétimo ano, antiga sexta série, do ensino fundamental, da Escola Estadual Predicanda Lopes, localizada em Macapá.

A análise dos dados obedecerá aos propósitos deste trabalho que são identificar o nível de aprendizado da turma quanto ao conteúdo ministrado em sala de aula pelo professor. Para isso, a variável número de acerto de uma questão proposta poderá ser tanto do tipo *quantitativa*, por se tratar de "possíveis valores formarem um conjunto finito ou enumerável de números, e que resultam, frequentemente, de uma contagem, como por exemplo *número de acertos de uma questão proposta (0, 1, 2, 3)*", [2] (grifo nosso), portanto as mesmas serão *discretas*. Mas ao mesmo tempo, vamos analisar a qualidade de ensino e aprendizagem do conteúdo abordado em sala de aula, daí estamos diante de uma análise estatística *qualitativa*, até porque colocamos em ordem crescente os possíveis números de acertos de uma das quatro atividades realizadas em sala de aula, portanto uma variável *ordinal*.

Em síntese, quanto ao modo de abordagem das variáveis, estudaremos o comportamento do número de acertos utilizando tanto as *variáveis quantitativas discretas* quanto as *qualitativas ordinais*, decorrentes das atividades propostas. Sendo assim, realizaremos uma abordagem estatística *qualiquantitativa*.

1.2 Distribuição de frequências

Quando se estuda uma variável, o maior interesse do pesquisador é conhecer o *comportamento* desta, analisando a ocorrência de suas possíveis realizações.

Em geral, segundo [16], para facilitar os estudos estatísticos sobre determinados fenômenos é comum construir uma tabela para atingir tal meta, por isso será disposto na Tabela 1.1 , a distribuição da frequência e as respectivas porcentagens do número de acertos da primeira atividade proposta à turma de estudantes da escola pesquisada, cujos dados foram retirados das folhas de resoluçao das duplas de alunos.

Tabela 1.1 *Frequência e porcentagem de acertos da* 1^a *Atividade proposta.*

Variável X x_i	Frequência n_i	Proporção f_i	Porcentagem $100f_i$
0	2	0.13	13
1	2	0.13	13
2	2	0.14	14
3	9	0.60	60
$\sum$	**15**	**1.00**	**100**

Fonte:Folhas de resolução das atividades, conforme mostrada na Figura 2.1.

Na primeira atividade proposta, será designada de *variável* X, a quantidade de acertos realizados pela dupla de estudantes. Observando os dados da Tabela 1.1, verifica-se que duas duplas não acertaram ou não resolveram a primeira atividade, que consistia em "Construir 3 exemplos de ângulos agudos.". Esta atividade foi escrita no quadro pela autora deste trabalho. Duas duplas acertaram somente um exemplo; duas duplas, dois

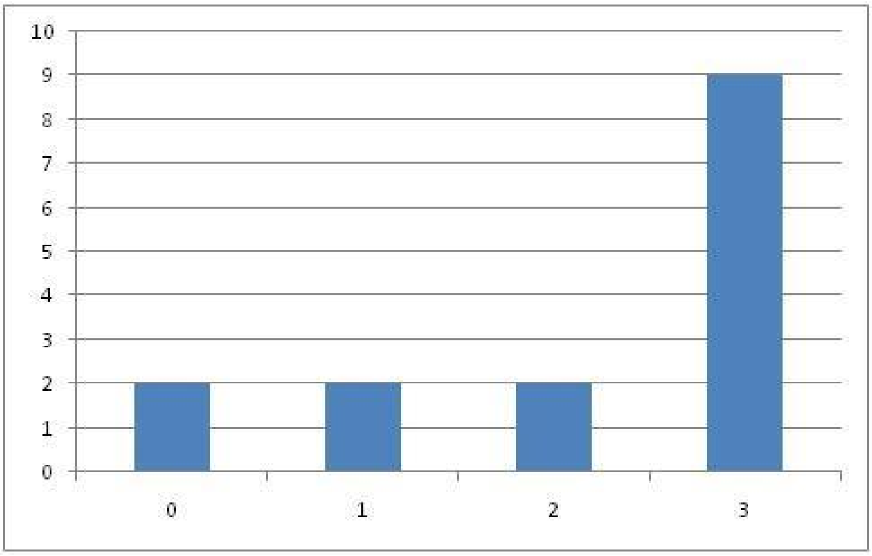

Figura 1.1 *Gráfico em barras verticais para a variável X: número de acertos da* 1^a *Atividade.*

Fonte: Folhas de resolução das atividades.Ver Figura 2.1.

exemplos e nove duplas acertaram corretamente os três ângulos, correspondente a sessenta por cento da turma.

Um estudo completo da intervenção pedagógica efetuada em sala de aula será apresentado no Capítulo 2.

1.3 Representação gráfica

A esse respeito [17], diz que para as variáveis quantitativas podemos considerar uma variedade maior de gráficos estatísticos, dentre os mais utilizados são: gráficos de barras (horizontal e vertical), de dispersão unidimensional e histograma. Para construir os gráficos mostrados neste trabalho utilizamos a planilha eletrônica *Excel*.

Observe que a Figura 1.1, mostra o gráfico de barras verticais, sendo que na legenda a frequência (n_i) é a absoluta, representadas pelas barras na cor azul. No semieixo horizontal estão as possibilidades de acertos da variável X, $(0, 1, 2$ ou $3)$ e no semieixo vertical estão as respectivas frequências absolutas $(2, 2, 2, 6)$, com isto é possível formar pares ordenados, (x_i, n_i), ou seja, $(0, 2)$, $(1, 2)$, $(2, 2)$ e $(3, 6)$.

Um gráfico de dispersão de número de acertos e frequência absoluta é mostrado na Figura 1.2, onde foram determinados os parâmetros alfa ($\alpha = 2,1$) e ($\beta = 0,6$), da reta de regressão, sendo que os mesmos serão abordados oportunamente.

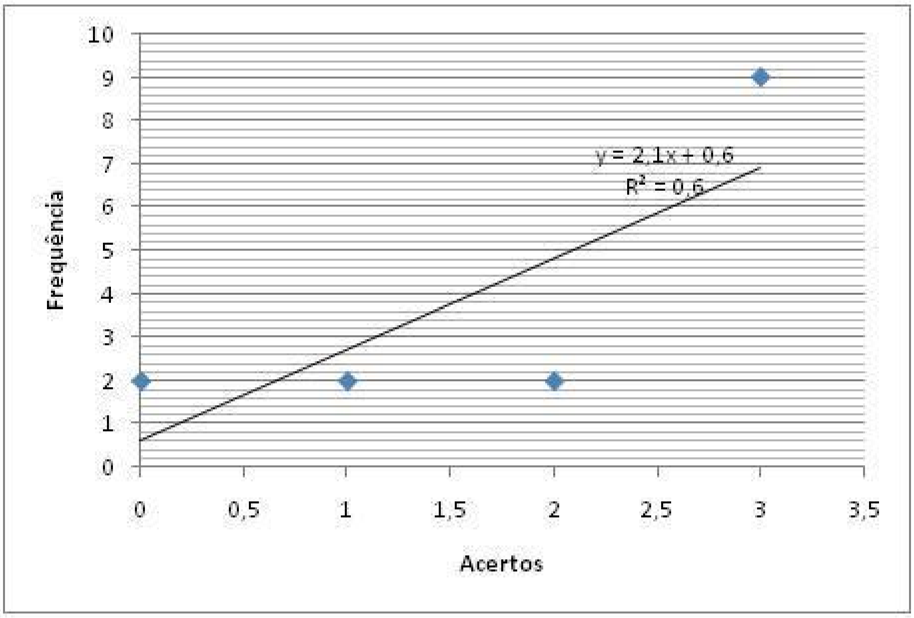

Figura 1.2 *Gráfico de dispersão de acertos e frequência absoluta da variável X.*
Fonte: Folhas de resolução das atividades. Ver a Figura 2.1.

1.4 Medidas de resumo

1.4.1 Medidas de posição

Estudamos que o resumo de dados por meio de tabelas de frequência fornece muito mais informações sobre o comportamento de uma variável do que a própria tabela original de dados. Muitas vezes, queremos resumir ainda mais estes dados, apresentando um ou alguns valores que sejam *representativos* da série toda. Quando usamos um só valor, obtemos uma redução drástica dos dados. Usualmente, emprega-se uma das seguintes medidas de posição (ou localização) central: média, mediana ou moda.

Os autores [12] definem a seguir as três medidas de posição:

A **moda** é definida como a realização mais frequente do conjunto de valores observados. Por exemplo, considere a variável quantitativa discreta Y, que representa o número de acertos da segunda atividade, resumida na Tabela 1.2. Vemos que a moda é zero, correspondente à realização com maior frequência absoluta, neste caso 7, significando que sete duplas não acertaram os itens propostos. Neste caso, temos uma moda, ou seja, a distribuição dos valores é denominada modal. Mas existem casos em que a distribuição apresenta duas (bimodal) ou mais modas, podendo ser denominada de trimodal, polimodal etc.

Tabela 1.2 *Frequência e porcentagem de acertos da* 2^a *Atividade proposta.*

Variável X x_i	Frequência n_i	Proporção f_i	Porcentagem $100f_i$
0	7	0.47	47
1	0	0.00	00
2	3	0.20	20
3	5	0.33	33
$\sum$	**15**	**1.00**	**100**

Fonte:Folhas de resolução das atividades, conforme mostrada na Figura 2.3.

A **mediana** é a realização que ocupa a posição central da série de observações, quando estão ordenadas em ordem crescente ou decrescente. Assim, de acordo com a Tabela 1.2, como as quinze observações da variável Y: 0, 0, 0, 0, 0, 0, 0, 2, 2, 2, 3, 3, 3, 3 e 3, a mediana é o valor 2 (dois), correspondente à oitava observação. Quando o número de observações for par, usa-se como mediana, a média aritmética das duas observações centrais.

Exemplo 1.4.1. *Usando os dados da Tabela 1.2, encontramos a moda da variável* Y*, é zero . Para a mediana, constatamos que é* 2*, porque é o termo central da série. A média aritmética é* 1.4*. A média aritmética pode também pode ser calculada por*

$$(0 \cdot 7 + 1 \cdot 0 + 2 \cdot 3 + 3 \cdot 5)/15 = (0 + 0 + 6 + 15)/15 = 21/15 = 1.4.$$

que é igual ao valor encontrado anteriormente.

Neste exemplo, as três medidas tem valores próximos e qualquer uma delas pode ser usada como *representativo* da série toda. A média aritmética é, talvez, a mais usada. Contudo, ela pode conduzir a erros de interpretação. Em muitas situações, a mediana é uma medida mais adequada.

Vamos formalizar os conceitos introduzidos anteriormente. Se x_1, ..., x_n são os n valores (distintos ou não) de qualquer variável, denotaremos por variável X, a média aritmética, ou simplesmente média de X e pode ser obtida por

$$\bar{x} = \frac{x_1 + ... + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{1.1}$$

agora, se tivermos n observações da variável X, das quais n_1 são iguais a x_1 , n_2 são iguais a x_2 etc. , n_n iguais a x_n , então a média de X pode ser escrita

$$\bar{x} = \frac{n_1 x_1 + ... + n_n x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} n_i x_i \tag{1.2}$$

Se $f_i = n_i/n$ representar a frequência relativa da observação x_i, então a Equação 1.2 também pode ser escrita

$$\bar{x} = \sum_{i=1}^{n} f_i x_i \tag{1.3}$$

Consideremos, agora, as observações ordenadas crescentemente. Vamos denotar a menor observação por $x_{(1)}$, a segunda por $x_{(2)}$, e assim por diante, obtendo-se

$$x_{(1)} \leq x_{(2)} \leq x_{(3)} \leq \ldots \leq x_{(n)} \tag{1.4}$$

Por exemplo, se $x_1 = 3$, $x_2 = -2$, $x_3 = 6$, $x_4 = 1$, $x_5 = 3$, então $-2 \leq 1 \leq 3 \leq 3 \leq 6$, de modo que $x_{(1)} = -2$, $x_{(2)} = 1$, $x_{(3)} = 3$, $x_{(4)} = 3$ e $x_{(5)} = 6$.

As observações ordenadas como na Equação 1.4 são chamadas *estatísticas de ordem.*

Com esta notaçao, a mediana de uma variável qualquer X pode ser definida como

$$md(X) = \begin{cases} x_{\frac{n+1}{2}} & se \quad n \quad impar \\ \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} & se \quad n \quad épar \end{cases} \tag{1.5}$$

Estas condições limitam bastante o cálculo de medidas de resumo para as variáveis qualitativas. Para as variáveis nominais somente podemos trabalhar com a moda. Para as variáveis ordinais, além da moda, podemos usar também a mediana. Devido a esse fato, iremos apresentar daqui em diante medidas de resumo para variáveis quantitativas, que permitem o uso de operações aritméticas com seus valores.

Por exemplo, a mediana do número de acertos da variável W é três, porque corresponde ao número central de acertos e escreve-se $Me(W) = 3$, isto é, o termo central da distribuição é o oitavo, cujo valor vale tres. Utilizando-se a Equação 1.5 e considerando que $n = 15$, temos:

$$Me(W) = x_{\frac{(15+1)}{2}} = x_{\frac{16}{2}} = x_8 = 3.$$

Abaixo, mostramos a distribuição de acertos da variável W.

0,0,0,0,0,0,0,3,3,3,4,5,6,6,6

Com isso, dividimos a distribuição em dois grupos: o primeiro com sete duplas que não acertaram a resposta ou não resolveram e o segundo, também com sete duplas que acertaram 3, 4, 5 ou 6 itens da terceira atividade.

1.4.2 Medidas de dispersão

Para [14], o resumo de um conjunto de dados por uma única medida representativa de posiçao esconde toda a informação sobre a variabilidade do conjunto de observações. Por exemplo, suponhamos que quinze duplas de alunos submeteram-se a um teste (intervenção estatística), composto por quatro questões, sendo que neste trabalho são designadas por atividades, obtendo as seguintes quantidades de acertos possíveis (0, 1, 2 ou 3), para as 1ª, 2ª e 4ª atividades, e (0 , 1, 2, 3, 4, 5 ou 6), para a 3ª atividade, conforme mostrado a seguir:

1ª Atividade (Variável X): 2 , 2 , 2 , 9
2ª Atividade (Variável Y): 7 , 0 , 3 , 5
3ª Atividade (Variável W): 7 , 0 , 0 , 3, 1, 1, 3
4ª Atividade (Variável Z): 8 , 1 , 0 , 6

Os autores destacam que as médias aritméticas das variáveis X, Y, W e Z, respectivamente, 2.2, 1.4, 2.4 e 1.3, pouco informam sobre suas diferentes variabilidades. Percebemos, então, a conveniência de serem criadas medidas que sumarizem a variabilidade de um conjunto de observações e que nos permita, por exemplo, comparar conjuntos diferentes de valores, como os dados acima, segundo algum critério estabelecido conforme [1].

Um critério usa do para tal fim é aquele que mede a dispersão dos dados em torno de sua média, e duas são as mais usadas: desvio médio e variância.

Para calcularmos o desvio médio, precisamos primeiramente calcular o desvio (d_i) de cada frequência absoluta em relação à média aritmética de cada variável. Por exemplo, para a 1ª Atividade acima os desvios $d_i = x_i - \bar{x}$ são: -2.2, -1.2, -0.2 e 0.8 (Ver Tabela 1.1).

Nestas condições, a soma dos desvios $[2 \times (-2.2) + 2 \times (-1.2) + 2 \times (-0.2) + 9 \times (0.8) = -4.4 - 2.4 - 0.4 + 7.2 = 0$. *Para qualquer conjunto de dados, a soma dos desvios é igual a zero.*], não é uma boa medida de dispersão para o conjunto de valores de qualquer distribuição, como por exemplo da 1ª Atividade.

Duas opções são:

(a) considerar o total dos desvios em valor absoluto;

(b) considerar o total dos quadrados dos desvios.

Para a 1ª Atividade, teríamos, respectivamente,

$$\sum_{i=1}^{4} n_i \, |x_i - \bar{x}| = 2 \times 2.2 + 2 \times 1.2 + 2 \times 0.2 + 9 \times 0.8 = 14.4,$$

$$\sum_{i=1}^{4} = n_i (x_i - \bar{x})^2 = 2 \times 4.84 + 2 \times 1.44 + 2 \times 0.04 + 9 \times 0.64 = 18.4,$$

para i = 1, 2, 3 e 4.

Para a 3ª Atividade, teríamos, respectivamente, conforme a Tabela 1.3:

$$\sum_{i=1}^{7} = n_i \, |w_i - \bar{w}| = 7 \times 2.4 + 0 \times 1.4 + 0 \times 0.4 + 3 \times 0.6 + 1 \times 1.6 + 1 \times 2.6 + 3 \times 3.6 = 33.6,$$

$$\sum_{i=1}^{7} = n_i (w_i - \bar{w})^2 =$$
$$7 \times 5.76 + 0 \times 1.96 + 0 \times 0.16 + 3 \times 0.36 + 1 \times 2.56 + 1 \times 6.76 + 3 \times 12.96 = 89.6,$$

para $i = 1, 2, ..., 7$.

O uso desses totais pode causar dificuldades quando comparamos conjuntos de dados com números diferentes de observações, como as 1ª e 3ª Atividades acima. Desse

Tabela 1.3 *Tabulação de dados da terceira atividade.*

Acerto w_i	Frequência absoluta n_i	Desvio relativo $w_i - \bar{w}$	Desvio absoluto $\lvert w_i - \bar{w}\rvert$	Quadrado dos desvio $(w_i - \bar{w})^2$	$n_i \cdot \lvert w_i - \bar{w}\rvert$	$n_i \cdot (w_i - \bar{w})^2$
0	7	-2.4	2.4	5.76	16.8	40.32
1	0	-1.4	1.6	1.96	0	0
2	0	-0.4	0.4	0.16	0	0
3	3	0.6	0.6	0.36	1.8	1.08
4	1	1.6	1.6	2.56	1.6	2.56
5	1	2.6	2.6	6.76	2.6	6.76
6	3	3.6	3.6	12.96	10.8	38,88
$\sum$	15				**33.6**	**89.6**

Fonte: Folhas de resolução das atividades.

modo, é mais conveniente exprimir as medidas como médias, isto é, o desvio médio e a variância são definidos por

$$dm(X) = \frac{\sum_{i=1}^{n} n_i \left| x_i - \bar{x} \right|}{n} \tag{1.6}$$

$$var(X) = \frac{\sum_{i=1}^{n} n_i \left(x_i - \bar{x} \right)^2}{n} \tag{1.7}$$

respectivamente.

Para a 1ª Atividade temos

$$dm(X) = 14.4/15 = 0.96,$$

$$var(X) = 18.4/15 = 1.23 \ ,$$

enquanto para a 3ª Atividade temos

$$\text{dm(W)} = 55.2/15 = 3.68,$$

$$\text{var(W)} = 89.6/15 = 5.97 \ ,$$

Podemos afirmar, então, que, de acordo com o desvio médio e a variância, a 1ª

Atividade é mais homogênea que a 3ª Atividade que apresentou valores muito afastado da unidade.

Sendo a variância uma medida de dimensão igual ao quadrado da dimensão dos dados (por exemplo, se os dados são expressos em metros (m), a variância será expressa em m^2), pode causar problemas de integração. Costuma-se usar então, o *desvio padrão*, que é definido como raiz quadrada positiva da variância. Para a 1ª Atividade o desvio padrão é

$$dp(X) = \sqrt{var(X)} = \sqrt{1.23} \approx 1.11,$$

que está próximo de um (1), o ideal neste caso seria próximo de zero. Para a 3ª Atividade o desvio padrão é

$$dp(W) = \sqrt{5.97} \approx 2.44.$$

Ambas as medidas de dispersão, dm e dp, indicam em média qual será o "erro" (desvio) cometido ao tentar substituir cada observação pela medida resumo do conjunto de dados (no caso, a média). Lembremos que as médias são: $\bar{x} = 2.2$ e $\bar{w} = 2.4$.

É de suma importância registrar que para a 3ª Atividade, ver Figura 2.3, o desvio padrão está muito próximo da média aritmética, mostrando com isso, que ocorreu um baixo nível de aprendizagem sobre a classificação de ângulos utilizando o transferidor.

Suponhamos que observemos n_1 vezes os valores x_i etc., n_k vezes o valor x_k da variável X. Então,

$$dm(X) = \frac{\sum_{i=1}^{k} n_i \left| x_i - \bar{x} \right|}{n} = \sum_{i=i}^{k} f_i \left| x_i - \bar{x} \right| \tag{1.8}$$

$$var(X) = \frac{\sum_{i=1}^{k} n_i \left(x_i - \bar{x} \right)^2}{n} = \sum_{i=1}^{k} f_i (x_i - \bar{x}) \tag{1.9}$$

$$dp(X) = \sqrt{var(X)} \tag{1.10}$$

Tanto a variância como o desvio médio são medidas de dispersão calculadas em relação à média das observações.

Assim como a média, a variância (ou o desvio padrão) são uma boa medida se a distribuição dos dados for aproximadamente normal.

Poderíamos considerar uma medida que seja calculada em relação à mediana. O desvio absoluto mediano é um exemplo e é mais resistente que o desvio padrão.

Algumas vezes, uma maneira computacionalmente mais eficiente de calcular a variância é

$$var(X) = \sum \frac{x_i^2}{n} - (\bar{x})^2 \tag{1.11}$$

e, no caso de observações repetidas,

$$var(X) = \sum f_i x_i^2 - (\bar{x})^2 \tag{1.12}$$

para $i = 1, ..., n$.

1.4.3 Frequência acumulada

Uma outra medida muito usada para descrever dados quantitativos é a *frequência acumulada*, que indica quantos elementos, ou que porcentagem deles, estão abaixo de um certo valor.

Na Tabela 1.4, a terceira e a sexta colunas indicam respectivamente a frequência absoluta acumulada e a proporção (porcentagem) acumulada. Assim, analisando a Tabela 1.4 podemos afirmar que 53% das duplas não acertaram nenhuma das três respostas da 4ª Atividade; 60% das duplas acertaram até duas respostas e 100% acertaram em até três respostas da quarta atividade.

Para um tratamento estatístico mais rigoroso das variáveis quantitativas, costuma-se usar uma definição mais precisa para a distribuição das frequências acumuladas.

Definição 1.4.1. *Dadas n observações de uma variável quantitativa e um número x real qualquer, indicar-se-á por $N(x)$ o número de observações menores ou iguais a x, e chamar-se-á de função de distribuição empírica (f.d.e) a função $F_n(x)$ ou $F_n(x)$*

$$F_e(x) = F_n(x) = \frac{N(x)}{n} \tag{1.13}$$

Tabela 1.4 *Frequência e porcentagem acumuladas de acertos da* 4^a *Atividade .*

Variável Z z_i	Frequência n_i	Frequência Acumulada N_i	Proporção f_i	Porcentagem $100f_i$	Porcentagem Acumulada
0	8	8	0.53	53	53.00
1	1	8	0.07	07	60.00
2	0	9	0.00	00	60.00
3	6	15	0.40	40	100.00
$\sum$	**15**		**1.00**	**100**	

Fonte: Folhas de resolução das atividades, conforme mostrada na Figura 2.1.

onde n é o número de elementos da distribuição.

1.5 Quantis

Tanto a média como o desvio padrão podem não ser medidas adequadas para representar um conjunto de dados, pois:

(a) são afetadas, de forma exagerada, por valores extremos;

(b) apenas com estes dois valores não temos ideia da simetria ou assimetria da distribuição dos dados.

Para contornar esses fatos, outras medidas tem de ser consideradas.

Vimos que a mediana é um valor que deixa metade dos dados abaixo dela e metade acima (ver Equação 1.5). De modo geral, podemos definir uma medida, chamada *quantil de ordem p* ou *p-quantil*, indicado por *q(p)*, onde p é uma proporção qualquer, $0 < p < 1$, tal que 100% das observações sejam menores do que *q(p)*.

Indicamos, abaixo, alguns quantis e seus nomes particulares.

$q(0,25)$: 1º Quartil = 25º Percentil

q(0.50): Mediana = 5º Decil = 50º Percentil

q(0.75): 3º Quartil = 75º Percentil

q(0.40): 4º Decil

q(0.95): 95º Percentil

Dependendo do valor de p, há dificuldades ao se calcular os quantis. Isso é ilustrado no exemplo a seguir.

Exemplo 1.5.1. *De acordo com a Tabela 1.4, temos os seguintes valores de uma variável Z (número de acerto, por dupla, da 4^a Atividade proposta):*

$$0, 0, 0, 0, 0, 0, 0, 0, 1, 3, 3, 3, 3, 3, 3.$$

Por estatística de ordem temos: $x_{(1)} = 0$, $x_{(2)} = 0$, $x_{(3)} = 1, ..., x_{(15)} = 3$, ou seja teremos

$$0 = 0 = 0 = 0 = 0 = 0 = 0 = 0 < 1 < 3 = 3 = 3 = 3 = 3 = 3.$$

Usando a definição de mediana, teremos que $md = q(0.5) = 0$. Suponhamos que queiramos calcular $q(0.65)$, ou seja, aquele valor que deixa 65% das observações a sua esquerda; mas 65% das observações correspondem a 9.75 observações. Qual valor devemos tomar? 1, que é a nona observação, ou 3, que é a décima observação, ou um valor entre 1 e 3?

Para responder a essa questão poderíamos usar a função de distribuição acumulada (fda) ou empírica (Fe). Essa função fornece, para cada número real x, a proporção das observações menores ou iguais a x. No nosso exemplo, temos na terceira coluna da Tabela 1.4 as frequências absolutas acumuladas N_i

$$N_0(z) = 8; N_1(z) = 9; N_2(z) = 9; N_3 = 15.$$

Por exemplo, para escrever a função densidade acumulada da variável Z, $F_n(z)$, utilizamos as informações da Tabela 1.4 e os dispomos na Equação **??**ção 13Equação 13, obtendo a função por partes a seguir

$$F_e(z) = \begin{cases} 0 & z < 0 \\ \frac{8}{15} & 0 \leq z < 1 \\ \frac{9}{15} & 1 \leq z < 2 \\ \frac{9}{15} & 2 \leq z < 3 \\ 1 & z \leq 3 \end{cases} \tag{1.14}$$

Como não há possibilidade do número de acertos ser negativo logo para $z < 0$ o valor correspondente da frequência acumulada relativa (far) igual a zero; como oito duplas não acertaram nenhuma das três peguntas propostas, logo no intervalo $0 \leq z < 1$, a far é 8/15, isto representa 53% das duplas; na terceira faixa, incluindo as duas anteriores, a far representa nove duplas que acertaram pelo menos uma resposta, 9/15; como nenhuma dupla acertou exatamente duas respostas, logo a far é a mesma da faixa anterior; e finalmente, quinze duplas respondenram ou não pelo menos as três perguntas, logo a far é 15, equivalente a cem por cento, representado por um (1).

Para melhores esclarecimentos esboçamos o gráfico de $F_n(z)$ conforme a 1.3.

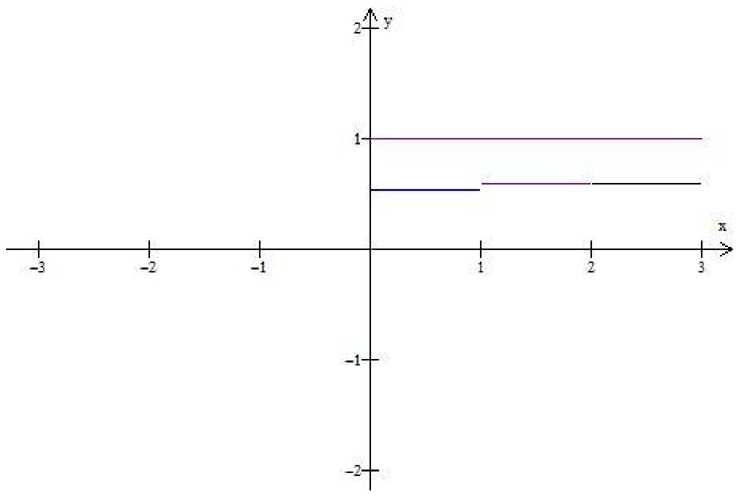

Figura 1.3 *Gráfico da função distribuição acumulada empírica da variável Z.*
Fonte: A partir de 1.14.

No semieixo horizontal a letra “x”representa a variável z. Observe que a semirreta correspondente a frequência acumulada relativa 1 abarca todos os possíveis números de acertos da quarta atividade.

Capítulo 2

ANÁLISE E DISCUSSÃO DA INTERVENÇÃO PEDAGÓGICA

A intervenção pedagógica sobre o ensino de geometria plana, que está servindo de objeto de estudos para este trabalho, ocorreu no dia vinte e um de setembro de 2012, na Escola Estadual Predicanda Lopes, localizada no Bairro Santa Rita, em Macapá. Em seu relatório de pesquisa, a acadêmica Kellen Cristiane Silva de Lima, acompanhada por seu professor orientador de TCC (João Ferreira) e do professor de Matemática da turma 711($7^{o}ano$), relatou que estavam presentes na sala de aula 30 alunos, destacando que a receptividade foi muito boa pelos pesquisados e pelo professor da disciplina. O conteúdo da aula foi sobre classificação dos ângulos quanto a sua abertura, isto é, se o ângulo desenhado nas atividades propostas, tanto as do livro texto adotado como a proposta pela pesquisadora, seria: agudo, reto ou obtuso (os estudantes utilizaram o transferidor para resolver estas atividades). Também foram utilizados como instrumentos auxiliares de instrução para esta atividade os seguintes materiais: apagador, geoplano, livro didático, pincel para quadro branco (lousa).

O desenvolvimento da aula transcorreu de forma muito dinâmica, porque a turma foi divida em duplas. Foram aplicadas quatro atividades - sendo que as duas primeiras foram escritas no quadro, pela aluna Kellen e as outras duas extraídas do livro didático (ver [8]), conforme mostrada nas Figura 2.1, Figura 2.3 e Figura 2.5; registrando que as três primeiras atividades os alunos teriam que utilizar o transferidor para resolvê-las e a quarta somente para responder.

Em relação aos alunos, percebemos que aproximadamente a totalidade da turma se interessou em participar da aula e consequentemente colaborar com os estudos da autora desta monografia.

Na realidade, procuramos mostrar aos alunos como é fácil utilizar um instrumento de medição, neste caso o transferidor, para resolver os problemas sobre construção de ângulos, enfatizando que este, na antiguidade, era uma ferramenta tecnológica e de grande importância para a solução de problemas de Matemática, tal qual, por exemplo, uma calculadora nos dias de hoje. Não se deve esquecer que ainda nos dias de hoje a construção e classificação de ângulos ainda é muito utilizado nas atividades científicas e as do cotidianas da humanidade. As 9h30min concluímos a pesquisa em sala de aula e agradecendo aos participantes pela dedicação e colaboração.

Nesta intervenção pedagógica, a primeira questão era a seguinte: "Construir 3 exemplos de ângulos agudos.", a segunda questão propunha: "Construir 3 exemplos de ângulos obtusos.", a terceira questão dizia o seguinte: "$A\hat{O}B$ é um ângulo reto, agudo ou obtuso? Use o transferidor, se for necessário.", e a quarta e última questão fazia uma pergunta com três subitens, a saber: "Qual é o maior: a) Um ângulo reto ou um ângulo obtuso? Resposta: ______________________; b) Um ângulo reto ou um ângulo agudo? Resposta: ______________________; c) Um ângulo obtuso ou ângulo agudo? Resposta:______________________".

Vamos começar analisando a primeira atividade estatisticamente. Em seguida faremos a análise e interpretação da segunda questão, para posterior comparação do desempenho da turma em relação as duas (correlação). Designaremos por X à variável independente correspondente ao número de acertos da primeira questão e por Y a variável independente correspondente ao número de acertos da segunda questão.

2.1 Análise estatística das 1^{a} e 2^{a} atividades

Faremos agora uma análise estatística das questões resolvidas pela turma com o objetivo de conhecermos mais detalhadamente o nível de aprendizagem dos participantes da pesquisa.

2.1.1 Análise bidimensional das variáveis X e Y

Como afirmou [3], é de costume para facilitar o processamento dos dados obtidos em uma pesquisa de campo, os mesmo serem colocados em uma tabela estatística, por isso, os dados da Atividade 1, foram tabulados na Tabela 1.1. Para melhorar o entendimento da

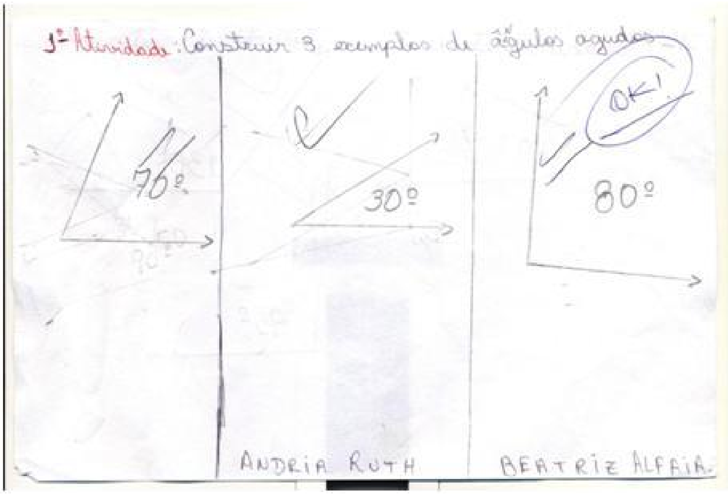

Figura 2.1 1^a *Atividade proposta à turma de* 7^o *ano.*
Fonte: Folhas de resolução das atividades.

proposta de estudos, a Figura 2.2 foi construída; no semieixo horizontal estão os possíveis números de acertos (zero (0) significa que a dupla não acertou nenhum dos três exemplos ou devolveu a atividade em branco, um (1)significa que a dupla acertou apenas um dos três exemplos, dois (2) significa que a dupla acertou somente dois dos três exemplos e três (3) significa que a dupla acertou todos três exemplos) e no semieixo vertical estão as respectivas frequências absolutas de possibilidade de acertos.

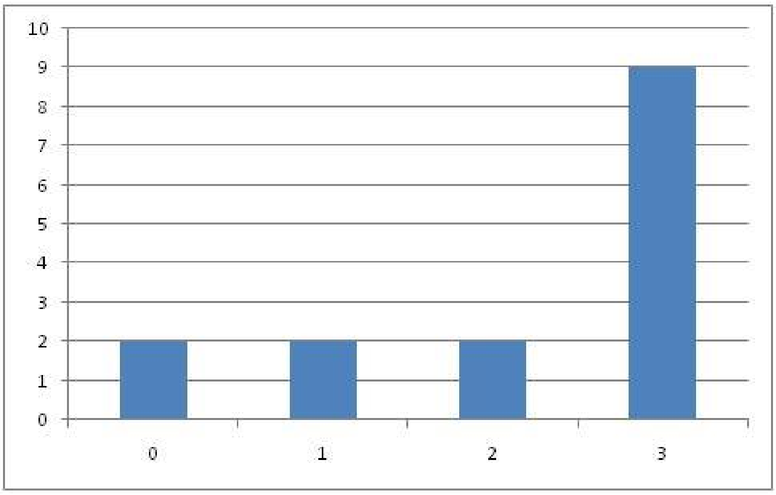

Figura 2.2 *Gráfico de frequências absolutas de acertos da* 1^a *Atividade.*
Fonte: Folhas de resolução das atividades.

Na Tabela 2.1, a sigla n_i significa a frequência absoluta de acertos de duplas, f_i, significa a frequência relativa, $100f_i$, a porcentagem, d_i, o desvio da frequência absoluta em relação à média aritmética,$\bar{x}$, a média aritmética simples, da, o desvio absoluto, $n_i da$, serve para se calcular o desvio médio (conforme Equação 1.8), isto é, 0.96; $(di)^2$, são os desvios ao quadrado, $n_i(di)^2$ são os desvios ao quadrado multiplicado pelas respectivas frequências absolutas. Utilizada para calcular a variância, com a Equação 1.9, cujo valor é 1.23 e o desvio padrão, neste caso de aproximadamente 1.1.

Tabela 2.1 *Tabulação de dados da* 1ª *Atividade: Construir* 3 *exemplos de ângulos agudos.*

Variável X x_i	Freq. absol. n_i	Freq. relat. f_i	Proporção $100f_i$	Desvio d_i	$n_i d_i$	Desvio absoluto da	$n_i da$	d_i^2	$n_i d_i^2$
0	2	0.13	13	-2.2	-4.4	2.2	4.4	4.84	9.68
1	2	0.13	13	-1.2	-2.4	1.2	2.4	1.44	2.88
2	2	0.14	14	-2.4	-0.4	0.2	0.4	0.44	0.08
3	9	0.60	60	0.8	7.2	0.8	7.2	0.64	5.76
$\sum$	**15**	**1.00**	**100**		**0.0**		**14.4**		**18.4**

Fonte:Folhas de resolução das atividades, conforme mostrada na Figura 2.1.

Analisando-se os resultados alcançados da variável X (acertos da primeira questão) conforme à Tabela 2.1, a moda do conjunto de acertos é três ($Mo(X) = 3$), isso porque 9 duplas responderam os três exemplos solicitados na primeira questão corretamente, equivalente a sessenta por cento (60%), significando que os alunos aprenderam corretamente como se controi ângulos agudos com o auxílio de um transferidor; a mediana também é três ($Me(X) = 3$), ou seja, separando-se as respostas em dois subgrupos, sete duplas responderam abaixo de três acertos e sete responderam corretamente acima de três acertos, mostrando com isso temos um equilíbrio no nível de conhecimento sobre ângulo agudo; a média aritmética ($\bar{x}$) é 2.2, indicando que a quantidade de acertos estão acima de cinquenta por cento (50%) do total de acertos. O desvio médio é $dm(X) = 0.96$, a variância dos acertos é $var(X) = 1.23$ e consequentemente o desvio padrão é 1.1. Com essas três medidas de dispersão ainda não se pode afirmar que a turma tem um bom nível de conhecimento sobre a construção de ângulos agudos utilizando o transferidor, para isso analisaremos então a segunda atividade.

Agora, vamos analisar a segunda questão e depois correlacioná-la com a primeira, com o objetivo de aprofundarmos os nossos conhecimentos sobre a turma no que refere aos estudos de construção de ângulos agudos e obtusos. Designaremos por Y a variável que representa os acertos da segunda atividade.

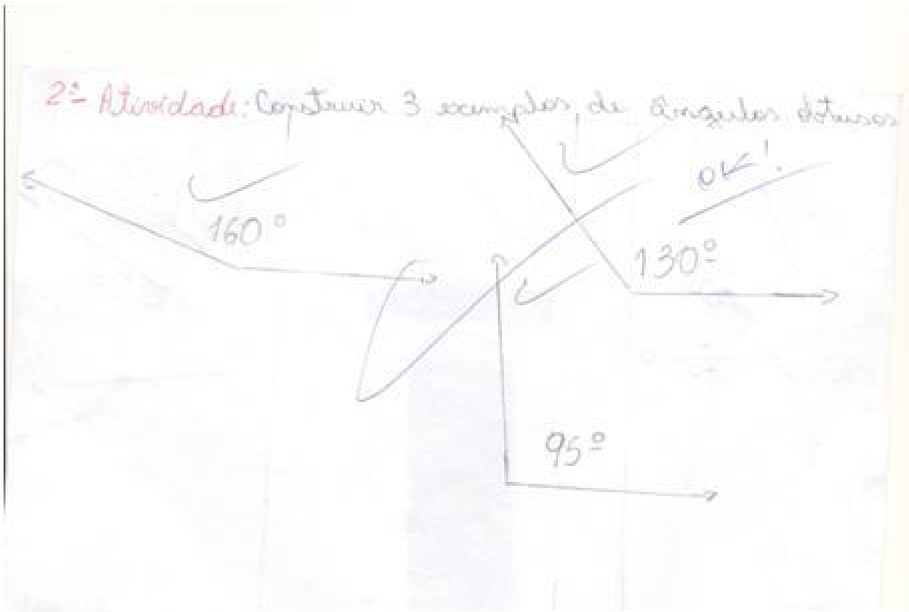

Figura 2.3 *Segunda atividade desenvolvida em sala de aula pelos estudantes.*
Fonte: Folhas de resolução das atividades.

A média aritmética da variável Y será calculada pela Equação 1.3, a seguir:

$$\bar{y} = (7 \times 0 + 0 \times 1 + 3 \times 2 + 5 \times 3)/15 = 21/15 = 1.4.$$

Tabela 2.2 *Tabulação da 2ª Atividade: Construir* 3 *exemplos de ângulo obtusos.*

Variáv. Y y_i	Freq. absol. n_i	Freq. relat. f_i	Prop. $100f_i$	Desv. d_i	n_id_i	Desv. abs. (da)	n_ida	d_i^2	$n_id_i^2$
0	7	0.47	47	-1.4	-9.8	1.4	9.8	1.96	13.72
1	0	0.00	00	-0.4	0.0	0.4	0.0	0.16	0.00
2	3	0.20	20	0.6	1.8	0.6	1.8	0.36	1.08
3	5	0.33	33	1.6	8.0	1.6	8.0	2.56	12.80
$\sum$	**15**	**1.00**	**100**		**0.0**		**19.6**	**5.04**	**27.60**

Fonte:Folhas de resolução das atividades, conforme mostrada na Figura 2.3.

A partir dos dados da Tabela 2.2, a média aritmética da variável Y (acertos da Atividade 2) é $\bar{y} = 1.4$, isso mostra que a média de acertos foi abaixo de 50% do total de exemplos solicitados. Temos uma moda: zero, porque sete duplas não acertaram nenhum dos três exemplos. A mediana é igual a dois acertos. A variância, que é a média aritmética do somatório dos quadrados dos desvios é igual a 1.84 e o desvio padrão é 1.36.

2.1.2 Discussão dos resultados das variáveis X e Y

Comparando as medidas de resumo das variáveis X e Y, conforme a Tabela 2.3, percebe-se que a 2ª Atividade, apresentou as três medidas de posição inferiores as da primeira, enquanto que as medidas de dispersão são superiores. Com esses resultados recomenda-se que seja desenvolvido um reforço escolar no processo de ensino e aprendizagem sobre construção de ângulos obtusos com o auxílio de transferidor.

Tabela 2.3 *Resultado das variáveis X e Y.*

Variável	Moda	Mediana	Média	Desvio médio	Variância	Desvio padrão
X	3	3	2.2	0.96	1.23	1.1
Y	0	2	1.4	1.306	1.84	1.4

Fonte:Tabulação das variáveis X e Y, de acordo com cálculos constantes neste texto.

2.1.3 Correlação linear simples entre as variáveis X e Y

Para os professores [2], é importante "[...] definir o coeficiente de correlação (linear) entre duas variáveis, que é uma medida do grau de associação entre elas e também da proximidade dos dados a uma reta.". Para calcular o coeficiente de correlação entre variáveis independentes quantitativas discretas (número de acertos) deve-se utilizar a Equação 2.1. Isto serve como forma de se verificar o grau de confiabilidade de uma pesquisa de campo, em outras palavras se o instrumento avaliativo pode ou não ser objeto de estudo.

$$corr(X,Y) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{dp(X)}\right)\left(\frac{y_i - \bar{y}}{dp(Y)}\right) \tag{2.1}$$

onde n é o número possível de respostas corretas da amostra, x_i e y_i são as quantidade $i-ésima$ das variáveis X e Y, $\bar{x}$ e $\bar{y}$ são respectivamente as médias aritméticas da variáveis X e Y e $dp(X)$ e $dp(Y)$ são os desvios padrões das variáveis X e Y. Para facilitar o cálculo do coeficiente de correlação, construiremos a Tabela 2.4.

Tabela 2.4 *Cálculo do coeficiente de correlação das variáveis X e Y.*

Variável X	Variável Y	$x_i - \bar{x}$	$y_i - \bar{y}$	$z_x = \frac{x_i - \bar{x}}{dp(X)}$	$z_y = \frac{y_i - \bar{y}}{dp(Y)}$	$z_x \cdot z_y$
0	0	-2.2	-1.4	-2	-1	2.00
0	0	-2.2	-1.4	-2	-1	2.00
1	0	-1.2	-1.4	-1.1	-1	1.10
1	0	-1.2	-1.4	-1	-1	1.10
2	0	-0.2	-1.4	-0.18	-1	0.18
2	0	-0.2	-1.4	-0.18	-1	0.18
3	0	0.8	-1.4	0.73	-1	-0.73
3	2	0.8	0.6	0.73	0.43	0.3139
3	2	0.8	0.6	0.73	0.43	0.3139
3	2	0.8	0.6	0.73	0.43	0.3139
3	3	0.8	1.6	0.73	1.14	0.8322
3	3	0.8	1.6	0.73	1.14	0.8322
3	3	0.8	1.6	0.73	1.14	0.8322
3	3	0.8	1.6	0.73	1.14	0.8322
3	3	0.8	1.6	0.73	1.14	0.8322
			0.0			**10.9327**

Fonte:Folhas de resolução das atividades, conforme mostrada nas Figuras 2.1 e 2.3.

Substituindo-se os valores da Tabela 2.4 na Equação 2.1, temos:

$$corr(X, Y) = \frac{1}{15} \times 10.9327 = 0.73$$

O resulta obtido (0.73) mostra que a distribuição está com a inclinação um pouco afastada do padronizado, que seria a unidade (1). Segundo os autores,[1], a correlação varia entre -1 e $+1$ (p. 84) e o ideal deve ser um resultado mais próximo de 1; mas mesmo assim é considerado um resultado forte, indicando com isto que a análise está atingindo a proposta desta pesquisa.

A Equação 2.1, pode ser operacionalizada de modo mais conveniente pelas seguintes fórmulas:

$$corr(X,Y) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{dp(X)}\right)\left(\frac{y_i - \bar{y}}{dp(Y)}\right) = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sqrt{\left(x_i^2 - n\bar{x}\right)\left(y_i^2 - n\bar{y}\right)}} \tag{2.2}$$

2.1.4 Teste para o coeficiente de correlação

Para os autores [7], em pesquisas, frequentemente, procura-se verificar se existe relação entre duas ou mais variáveis, isto é, saber se as alterações sofridas por uma das variáveis são acompanhadas por alterações nas outras. Por exemplo, peso versus idade, consumo versus renda, altura versus peso, de um indivíduo ou o número de acertos de uma atividade geométrica versus outra atividade geométrica proposta, sobre o mesmo conteúdo. O termo correlação significa relação em dois sentidos (co + relação), e é usado em estatística para designar a força que mantém unidos dois conjuntos de valores. A verificação da existência e do grau de relação entre as variáveis é o objeto de estudo da correlação.

Uma vez caracterizada esta relação, destaca [16], procura-se descrevê-la sob forma matemática, através de uma função. A estimação dos parámetros dessa função matemática é o objeto da regressão.

Os pares de valores das duas variáveis poderão ser colocados num diagrama cartesiano chamado “diagrama de disperção”, conforme Figura 2.4. A vantagem de construir um diagrama de dispersão está em que, muitas vezes sua simples observação já nos dá uma ideia bastante boa de como as duas variáveis se relacionam.

Uma medida do grau e do sinal da correlação é dada pela covariância entre as duas variáveis X e Y que é uma medida numérica de associação linear existente entre elas, e definida pela Equação 2.7 .

É mais conveniente usar para medida de correlação, o coeficiente de correlação linear de Pearson, como estimador de ρ_{xy}, definido na Equação 2.9.

A partir das variáveis X e Y são determinadas todas as somas necessárias para este cálculo, conforme mostrado na Tabela 2.4.

O coeficiente de correlação r_{xy} linear é um número puro que varia de -1 a $+1$, como dissemos anteriormente, e sua interpretação dependerá do valor numérico e do sinal, como segue:

$r_{xy} = -1 \Rightarrow$ correlação perfeita negativa;
$-1 < r_{xy} < 0 \Rightarrow$ correlação negativa;
$r_{xy} = 0 \Rightarrow$ correlação nula;
$0 < r_{xy} < 1 \Rightarrow$ correlação positiva;
$r_{xy} = -1 \Rightarrow$ correlação perfeita positiva;
$0,2 < r_{xy} < 0.4 \Rightarrow$ correlação fraca *;
$0,4 < r_{xy} < 0.7 \Rightarrow$ correlação moderada;
$0,7 < r < 0.9 \Rightarrow$ correlação forte;

O coeficiente de correlação linear de Pearson obtido a partir da Equação 2.9, foi $r \approx 0,76$, que de acordo com a classificação acima, a correlação entre as duas variáveis é forte. A correlação calculada pela planilha Excel é de aproximadamente 0.75.

2.1.5 Regressão linear simples

Segundo [4], chamamos de modelo linear à Equação 2.3, pois esta representa uma reta.

$$y(x) = \alpha + \beta x \tag{2.3}$$

Todavia, em casos mais gerais, o termo *linear* refere-se ao modo como os parâmetros entram no modelo, ou seja, de forma linear. Por exemplo, o modelo

$$y(x) = \alpha + \beta x + \gamma x^2 \tag{2.4}$$

embora graficamente represente uma parábola, é um *modelo linear* em α, β e γ. Por outro lado,

$$y(x) = \alpha e^{\beta x} \tag{2.5}$$

é um modelo não linear em α e β. Na realidade, o modelo acima, Equação 2.5, é definido como *semiexponencial*, contudo não será abordado nesta monografia.

* NA: Possui o mesmo significado para os casos negativos ou positivos.

Na Figura 2.4, com o auxílio da planilha *Excel*, plotou-se as duas curvas de tendência linear, sendo que uma é do tipo da Equação 2.3 e a outra do tipo de Equação 2.4.

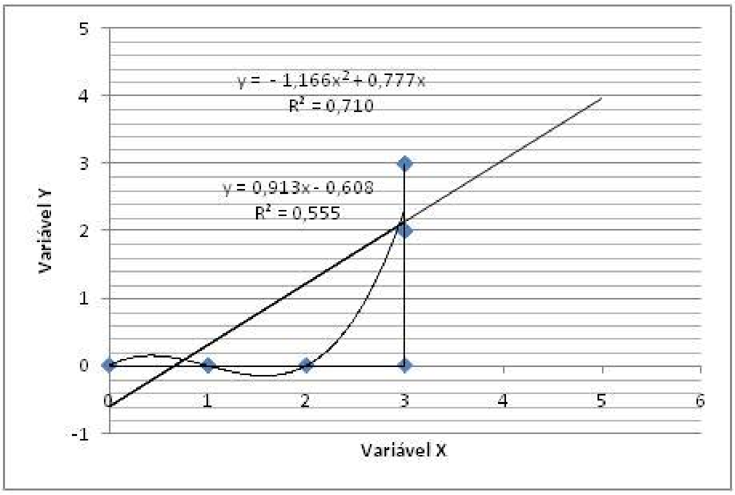

Figura 2.4 *Gráfico de dispersão entre as variáveis X e Y com as linhas de tendências linear e polinomial.*

Fonte: Folhas de resolução das atividades.

Para estimar os parâmetros α e β, vamos adotar o critério que consiste em encontrar os valores de α e β que minimizam a soma dos erros, dados por

$$e_i = y_i - (\alpha + \beta x_i) \tag{2.6}$$

$i = 1, ..., n.$

Analisando os gráficos da Figura 2.4, existem duas curvas de regressão linear: a primeira correspondente ao modelo apresentado na Equação 2.3, cujos interceptos são $\alpha = -0.608$ e $\beta = 0.913$. Embora o β esteja muito próximo da unidade (1), o parâmetro mais importante que devemos considerar é o *coeficiente de determinação* R^2 cujo valor é 0.555, ele indica a proporção de variação da variável dependente (Y) a partir da variação da variável independente (X), ou seja, é uma ferramenta que avalia a qualidade do ajuste. A sua definição e o seu intervalo de variação são

$$R^2 = r^2_{XY} \quad \text{e} \quad 0 \leq R^2 \leq 1.$$

Quanto mais próximo da unidade o R^2 estiver, melhor a qualidade do ajuste. O seu valor fornece a proporção da variável Y (conhecimentos sobre ângulo obtuso) explicada pela variável X (conhecimentos sobre ângulo agudo) através da função ajustada. Neste caso, $R^2 = 0.555 \approx 55,5\%$. É a proporção que Y é explicada por X, ou seja, 55.5% da variação do número de acertos da 2ª Atividade é explicada pela resolução da 1ª Atividade.

A segunda curva de tendência linear corresponde ao modelo apresentado na Equação 2.4, cujos parâmetros são $\alpha = -1.116$, $\beta = 0.777$, $\gamma = 0$ e $R^2 = 0.710$, neste caso foi maior do que o obtido no modelo discutido anteriormente. Sendo assim, para este modelo temos uma proporção de aprendizagem sobre ângulos obtusos em função de conhecimentos sobre ângulos agudos de 71%. Portanto, o segunda curva de tendência é a que mais se ajusta aos estudos realizados sobre a construção de ângulos agudos e obtusos, para os alunos pesquisados.

2.1.6 Covariância entre as variáveis X e Y

O numerador da Equação 2.2, que mede o total da concentração dos pontos pelos quatro quadrantes, dá origem a uma medida bastante usada chamada de *covariância* entre as duas variáveis. Para calcularmos a covariância utiliza-se a seguinte fórmula

$$cov\left(X,Y\right) = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{n} \tag{2.7}$$

ou seja, a média dos produtos dos valores centrados das variáveis. Para determinarmos a covariância vamos construir a Tabela 2.5.

Substituindo-se os valores da Tabela 2.5 na Equação 2.7, temos:

$$cov\left(X,Y\right) = \tfrac{19.6}{15} \approx 1.3$$

Vale registrar que o valor calculado pela planilha *Excel* é de aproximadamente 1.12.

Com a definição da Equação 2.7, o *coeficiente de correlação* pode ser escrito como

$$corr(X,Y) = \frac{cov(X,Y)}{dp(X)dp(Y)} \tag{2.8}$$

Sendo assim, temos:

$$corr(X,Y) = \tfrac{1.3}{1.1 \times 1.36} = \tfrac{1.3}{1.496} \approx 0.79$$

Tabela 2.5 *Cálculo da covariância.*

Variável X	Variável Y	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})(y_i - \bar{y})$
0	0	-2.2	-1.4	3,08
0	0	-2.2	-1.4	3,08
1	0	-1.2	-1.4	3,08
1	0	-1.2	-1.4	3,08
2	0	-0.2	-1.4	0,28
2	0	-0.2	-1.4	0,28
3	0	0.8	-1.4	-1,12
3	2	0.8	0.6	0,48
3	2	0.8	0.6	0,48
3	2	0.8	0.6	0,48
3	3	0.8	1.6	1,28
3	3	0.8	1.6	1,28
3	3	0.8	1.6	1,28
3	3	0.8	1.6	1,28
3	3	0.8	1.6	1.28
$\sum_{i=15}^{15}$				**19.6**

Fonte:Da autora (2012).

Comparando-se o resultado acima com o obtido pela Equação 2.9, verifica-se que estão muito próximos um do outro. O valor da correlação calculada com o Excel foi de aproximadamente 0.75.

2.1.7 Coeficiente de correlação de K. Pearson

Esta é uma forma simplificada do Coeficiente de correlação linear de Pearson, como estimador de ρ_{xy}, e é descrita pela fórmula:

$$r_{xy} = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left[\sum x^2 - \frac{(\sum x)^2}{n}\right]\left[\sum y^2 - \frac{(\sum y)^2}{n}\right]}} \tag{2.9}$$

Para colaborar na obtenção do mesmo, vamos preencher a Tabela 2.6.

Substituindo-se os valores da Tabela 2.6 na Equação 2.9, obtemos:

Tabela 2.6 *Cálculo do coeficiente de Pearson.*

X	Y	X^2	Y^2	X · Y
0	0	0	0	0
0	0	0	0	0
1	0	1	0	0
1	0	1	0	0
2	0	4	0	0
2	0	4	0	0
3	0	9	0	0
3	2	9	4	6
3	2	9	4	6
3	2	9	4	6
3	3	9	9	9
3	3	9	9	9
3	3	9	9	9
3	3	9	9	9
3	3	9	9	9
33	**21**	**91**	**57**	**63**

Fonte:Da autora (2012).

$$r_{xy} = \frac{63-\frac{33\times 21}{15}}{\sqrt{\left[91-\frac{(33)^2}{15}\right]\left[57-\frac{(21)^2}{15}\right]}} = \frac{63-\frac{693}{15}}{\sqrt{\left[91-\frac{1098}{15}\right]\left[57-\frac{441}{15}\right]}} = \frac{63-46.2}{\sqrt{[91-73.2][57-29.4]}}$$

$$\frac{16.8}{\sqrt{[17.8][27.6]}} = \frac{16.8}{\sqrt{491,28}} = \frac{16.8}{22.16} \approx 0.76$$

Este resultado está próximo dos valores encontrados anteriormente. Na prática estes três coeficientes, correlação, covariância e Pearson, são sinônimos, ou seja, representam o mesmo valor e interpretação.

Tabela 2.7 *Coeficientes das variáveis X e Y.*

$Cor(X,Y)$	$Cov(X,Y)$	Coeficiente de Pearson (r_{XY})	Coeficiente de determinação R^2
0.73	**1.3**	**0.76**	**0,5776**

Fonte:Tabulação das variáveis X e Y, de acordo com cálculos constantes neste texto.

2.2 Análise estatística da 3ª atividade

A terceira atividade (variável W) será analisada estatisticamente de forma independentemente das outras, por apresentar seis itens a serem respondidos pelas duplas, conforme mostrado na Figura 2.5.

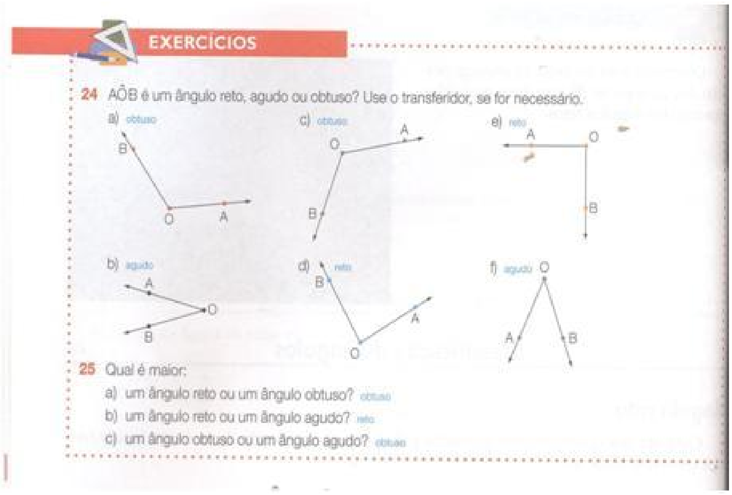

Figura 2.5 *Terceira e quarta atividades propostas.*
Fonte: ([8], 2009, p. 80).

A partir dos dados da Tabela 2.8, temos: a moda , $Mo(W) = 0$, ou seja, é o valor mais repetido ou que ocorre com maior frequência na distribuição; a mediana, $Me(W) = 3$ é o número central da distribuição; a média aritmética é $\bar{w} = 2.4$; o desvio médio, $dm(W) = 2.24(33.6/15)$, ou seja, esta é a média dos desvios absolutos dos pontos de dados a partir de sua média; a variância , $var(W) = 5.97$ e o desvio padrão, $dp(W) = 2.44$.

Analisando-se os resultados obtidos no parágrafo anterior, percebemos que os valores das informações são bastantes elevados, principalmente a variância e o desvio padrão. Mas isso se deu em detrimento de que 46% das duplas não acertarem ou não responderam pelo menos três respostas dos seis itens propostos, isto é, 7 duplas devolveram a folha de resolução em branco ou responderam os seis itens errados.

A Figura 2.6, mostra a distribuição dos pares (w_i, n_i) no gráfico de dispersão com a linha de tendência poligonal passando pelos sete pontos e também a curva de dispersão, apresentando inclinação negativa, isto ocorre porque o número de acertos aumenta e frequência absoluta diminui.

Tabela 2.8 *Tabulação de dados da terceira atividade.*

Acerto w_i	Frequência absoluta n_i	Frequência relativa f_i	Porcentagem $100f_i$	Desvio relativo d_i	Produto $n_i d_i$	Desvio absoluto $\|d_i\|$	Produto $n_i\|d_i\|$
0	7	0.46	46	-2.4	-16.8	2.4	16.8
1	0	0.00	00	-1.4	0.0	1.4	00.0
2	0	0.00	00	-0.4	0.0	0.4	00.0
3	3	0,20	20	0.6	1.8	0.6	1.8
4	1	0.07	7	1.6	1.6	1.6	1.6
5	1	0.07	7	2.6	2.6	2.6	2.6
6	3	0.20	20	3.6	10.8	3.6	10.8
$\sum$	**15**	**1,00**	**100**		**0**	**19.8**	**33.6**

Fonte: Folhas de resolução das atividades.

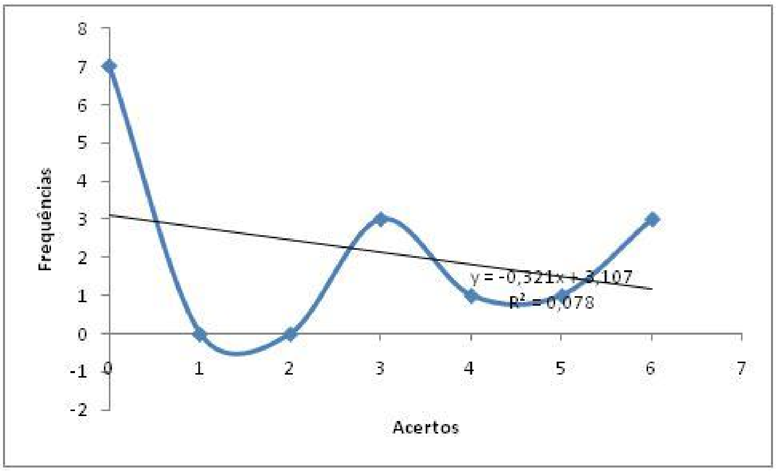

Figura 2.6 *Gráfico de dispersão da 3ª Atividade com a sua curva de tendência linear.*
Fonte: Ver Tabela 2.8.

Os coeficientes, dos conjuntos de dados acertos (w_i) versus frequências absolutas (n_i) são: $cor(w, n) = -0,28016$ (correlação fraca negativa), $cov(w, n) = -1,28571$, coeficiente de Pearson, $r_{wn} = -0,28016$ e consequentemente $R^2 = 0,078$, cujo valor indica a associação linear existente entre as variáveis w e n, conforme mostrado na Figura 2.6. Este resultado corrobora as informações alcançadas com as respostas dadas pelas duplas ao resolver a terceira atividade.

Tabela 2.9 *Coeficientes da variável W .*

$Cor(W,n)$	$Cov(W,n)$	Coeficiente de Pearson (r_{Wn})	Coeficiente de determinação R^2
-0.28	**-1.3**	**-0.28**	**0,078**

Fonte:Tabulação da variável W.

2.3 Análise estatística da 4^a atividade

A quarta atividade (variável Z) será analisada inicialmente de forma individual. Posteriormente faremos uma comparação dos resultados com os das variáveis X e Y por apresentarem o mesmo número de respostas, conforme mostrado na Figura 2.5. Os dados estão dispostos na Tabela 2.10.

Tabela 2.10 *Frequência e porcentagem de acertos da 4^a Atividade proposta.*

Acertos Variável Z z_i	Frequência absoluta n_i	Frequência relativa f_i	Porcentagem $100f_i$
0	8	0.53	53
1	1	0.07	07
2	0	0.00	00
3	6	0,40	40
$\sum$	15	1,00	100

Fonte: Folhas de resolução da quarta atividade.

A partir da Tabela 2.10, obtivemos as seguintes informações: a moda ,$Mo(Z) = 0$, ou seja, é o valor mais repetido ou que ocorre com maior frequência na distribuição; a mediana, $Me(Z) = 0$ é o número central da distribuição; a média aritmética é $\bar{z} = 1.3$; o desvio médio, $DM(Z) = 1.39$, ou seja, esta é a média dos desvios absolutos dos pontos de dados a partir de sua média; a variância , $var(Z) = 2.21$ e o desvio padrão, $dp(Z) = 1.5$.

Agora vamos comparar as medidas de posição e de dispersão das variáveis X, Y e Z, com o intuito de resumirmos o máximo possível as informações a respeito das atividades aplicadas as duplas, reunidos na Tabela 2.11.

As medidas de posição moda, mediana e média são decrescentes, enquanto que as medidas de dispersão, variância e desvio padrão são crescentes, indicando com isto, que

Tabela 2.11 *Comparação dos resultados das variáveis X, Y e Z.*

Variáveis	Moda Mo	Mediana Me	Média aritmética	Desvio médio	Variância var	Desvio padrão
X	3	3	2.2	0.96	1.31	1.15
Y	0	2	1.4	1.31	1.84	1.4
Z	0	0	1.3	1.39	2.21	1.5

Fonte: Tabulação deste trabalho.

as duplas apresentaram um rendimento decrescente em relação ao conteúdo estudado. É perceptível que os alunos começaram as atividades de forma dedicada e interessada, mas à medida que avançavam nas resoluções as respostas foram sendo dadas de maneira erradas ou entregaram as atividades em branco.

2.3.1 Análise bidimensional das variáveis Z e X

Vamos determinar o coeficiente de correlação entre as variáveis Z (dependente) e X (independente), para conhecermos como a turma conseguiu resolver a quarta atividade em função do conhecimento da primeira; para isto utilizaremos a Equação 2.2.

$$corr(X,Z) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{dp(X)}\right)\left(\frac{z_i-\bar{z}}{dp(Z)}\right) \approx 0.64$$

Este resultado aponta uma correlação moderada entre as variáveis Z e X. Portanto, o fato de saber responder a primeira atividade necessariamente não implica em saber responder a quarta atividade. O coeficiente de correlação de Pearson é de aproximadamente $r_{ZX} \approx 0.64$.

O coeficiente de determinação R^2 é de aproximadamente 0.41, indicando que apenas 41% dos alunos conseguiram correlacionar as duas atividades. Lembrando que a primeira era para construir ângulos agudos e a quarta é para responder a comparação entre ângulos agudos, retos e obtusos.

A covariância da variáveis, $cov(Z,X) = 1.013$. O sinal na covariância indica o tipo de relação que as duas variáveis tem. Um sinal positivo indica que elas movem juntas e um negativo que elas movem em direções opostas.

2.3.2 Análise bidimensional das variáveis Z e Y

Agora, vamos determinar o coeficiente de correlação entre as variáveis Z (dependente) e Y (independente). Para conhecermos como a turma conseguiu resolver a quarta atividade em função do conhecimento da segunda; para isto utilizaremos a Equação 2.2.

$$corr(Z,Y) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i-\bar{y}}{dp(Y)}\right)\left(\frac{z_i-\bar{z}}{dp(Z)}\right) \approx 0.903$$

Este resultado aponta uma correlação muito forte entre as variáveis Z e Y. Portanto, o fato de não saber responder a segunda atividade esta associada a não saber responder a quarta atividade. O coeficiente de correlação de Pearson é de aproximadamente $r_{ZY} \approx 0.903$.

O coeficiente de determinação R^2 é de aproximadamente 0.82, o dobro do coeficiente de determinação entre as variáveis Z e Y, indicando que agrande maioria, 82% dos alunos, não conseguiu correlacionar as duas atividades.

A covariância da variáveis é $cov(Z,Y) = 1.76$. O sinal na covariância indica o tipo de relação que as duas variáveis possuem. Um sinal positivo indica que elas movem juntas e um negativo que elas movem em direções opostas.

Os resultados descritos na Tabela 2.12, monstram que a não resolução da quarta atividade a partir do não conhecimento segunda é muito forte. Esta leitura é obtida a partir do coeficiente de determinação 0.82 correspondente a 82% das respostas correlacionadas. Por sinal, esta correlação é o dobro das variáveis Z e X.

Tabela 2.12 *Comparação dos coeficientes entre as variáveis* (Z,X) *e* (Z,Y).

Variáveis	Correlação	Covariância	Pearson	R^2
(Z,X)	0.64	1.013	0.64	0.41
(Z,Y)	0.903	1.76	0.903	0.82

Fonte: Tabulação deste trabalho.

CONCLUSÃO

No estudo ficou evidenciado que as medidas de resumo das variáveis X e Y, de acordo com a Tabela 2.3, apresentam possíveis constatações referentes as duas primeiras atividades: a segunda atividade, apresentou as três medidas de posição inferiores as da primeira, enquanto que as medidas de dispersão foram superiores. Com esses resultados recomenda-se que sejam desenvolvidos reforços escolares no processo de ensino e aprendizagem sobre construção de ângulos obtusos com o auxílio de transferidor.

Para aprofundar os estudos com o objetivo de se interpretar os resultados preliminares, recorreu-se a análise bidimensional das variáveis e o resultado da correlação linear entre as variáveis X e Y, é de aproximadamente 0.73, isso sugere que a distribuição está com a inclinação um pouco afastada do esperado, que segundo a literatura pertinente seria de uma unidade (1).

Aparentemente, os estudos sugerem que com os três coeficientes calculados, haja uma forte correlação linear entre as duas primeiras atividades (variáveis X e Y), com isso, conjetura-se como satisfatória as informações e fornecem a noção do quanto a turma aprendeu sobre ângulos agudos e obtusos.

Na seção 2.2, constatou-se que em virtude do grande número de duplas que não resolveram ou resolveram errado a atividade (somente três duplas responderam corretamente os seis itens), isto influenciou demais nos valores da variância e do desvio padrão, portanto sugere-se que sejam repetidos os estudos e refeita uma atividade similar.

A quarta e última atividade foi estudada de maneira diferente das demais, isso porque tratava-se de perguntas e respostas, daí essa possibilidade de se aplicar um metodologia alternativa. Possivelmente os resultados encontrados e dispostos nas Tabela 2.11 e Tabela 2.12 indicam que correlação entre as variáveis Z e Y são mais fortes do que entre as variáveis Z e X. Essa ideia surgiu a partir do indicador R^2_{ZY} estar muito próximo de 1; em termos percentuais é de aproximadamente 82%, oportunizando uma especulação de que o aprendizado da primeira implica na resolução da segunda.

Tabela 2.13 *Medidas de resumo das quatro variáveis.*

Variável	Moda	Mediana	Média	Desvio médio	Variância	Desvio padrão
X	3	3	2.2	0.96	1.23	1.1
Y	0	2	1.4	1.306	1.84	1.4
W	0	3	2.4	2.24	5.97	2.44
Z	0	0	1.3	1.39	2.21	1.5

Fonte: De acordo com os cálculos constantes no texto.

A Tabela 2.13 apresenta as medidas de resumo das quatro atividades, sendo que a variável X é aquela que apresenta a melhor média de acertos. A variável Y influenciou na resolução da quarta atividade, pois de acordo com a Tabela 2.12, o coeficiente de determinação monstra que a não resolção da quarta atividade a partir do pouco conhecimento segunda é muito forte.

Espera-se que este Trabalho de Conclusão de Curso tenha alcançado os seus objetivos e propósito delineados no decorrer dos estudos realizados. E que possivelmente sirva de referencial teórico para educadores e estudiosos que se debrucem sobre o tema de melhoria da qualidade de ensino.

BIBLIOGRAFIA

BOTELHO, Elisa Maria Diniz; MACIEL, Antonio José **Estatística descritiva**: um curso introdutório. Minas Gerais: Universidade Federal de Viçosa, 1992.

BUSSAB, Wilton de O.; MORETTIN, Pedro A. **Estatística básica**. 5. ed. 6. tir. rev. atual. São Paulo: Editora Saraiva, 2006.

CLARK, Jeffrey. **Estatística aplicada**. São Paulo: Saraiva, 2006.

COSTA NETO, Pedro Luiz de Oliveira. **Estatística**. São Paulo: Edgar Blucher, 1977.

CRESPO, Antono Arnot. **Estatística fácil**. São Paulo: Saraiva, 1999.

CUNHA, S. Ezequiel da. **Estatística descritiva** : na psicologia e na educação. Rio de Janeiro: Forense Universitária, 1978.

FONSECA, Jairo Simon da; MARTINS, Gilberto de Andrade. **Curso de estatística**. São Paulo: Atlas, 1996.

IEZZI, Gelson; DOLCE, Osvaldo; MACHADO, Antonio. **Matemática e realidade**. 7. ano. 6. ed. SÃ£o Paulo: Atual, 2009.

MOREIRA, Jairo Simon. **Estatística Aplicada**. São Paulo: Atlas, 1995.

MOREIRA, José dos Santos. **Elementos de Estatística**.São Paulo: Atlas, s.d.

NAZARETH, Helenalda Resende de Souza. **Curso básico de estatística**. São Paulo: Ática, 1999.

OVALLE, Ivo Izidoro; TOLEDO, Geraldo Luciano. **Estatística Básica**. São Paulo: Atlas, 1995.

ROBERTS, Harry V.; WALIS, N. Allen. **Curso de estatística**. Rio de Janeiro: Fundo de Cultura, 1969.

SILVA, Nelson Pereira da. **Estatística autoexplicativa**. São Paulo: Erica, 1998.

SPIEGEL, Murrai R. **Estatística**. São Paulo: Pearson Education do Brasil, 1994.

TRIOLA , Mario. **Introdução a estatística**. Rio de Janeiro: LTC, 1999.

VESSEREAU, Andre. **A estatística**. São Paulo: Difusão Europeia do Livro, 1965.