



UNIVERSIDADE FEDERAL DO PIAUÍ.
CENTRO DE CIÊNCIAS DA NATUREZA
PÓS-GRADUAÇÃO EM MATEMÁTICA
MESTRADO EM MATEMÁTICA

**A importância do Ensino de Regressão Linear
Simples no Ensino Médio: um estudo com alunos do
3º Ano do Ensino Médio - IFMA - Imperatriz**

Paulo Cardoso Jales

Teresina - 2014

Paulo Cardoso Jales

Dissertação de Mestrado:

**A importância do Ensino de Regressão Linear Simples no
Ensino Médio: um estudo com alunos do 3º Ano do Ensino
Médio - IFMA - Imperatriz**

Dissertação submetida à Coordenação do Programa de Pós-Graduação em Matemática, da Universidade Federal do Piauí, como requisito parcial para obtenção do grau de Mestre em Matemática.

Orientador:

Prof. Dra. Valmária Rocha da Silva Ferraz

Teresina - 2014

FICHA CATALOGRÁFICA
Universidade Federal do Piauí
Biblioteca Comunitária Jornalista Carlos Castello Branco
Serviço de Processamento Técnico

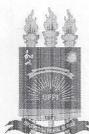
J26i Jales, Paulo Cardoso.
A Importância do ensino de Regressão Linear Simples no
Ensino Médio : um estudo com alunos do 3º ano do Ensino
Médio IFMA- Imperatriz / Paulo Cardoso Jales. -- Teresina,
2014.
46 f.

Dissertação (Mestrado em Matemática) – Universidade
Federal do Piauí, 2014.

Orientação: Profª. Drª. Valmária Rocha da Silva Ferraz.

1. Regressão Linear - Matemática. 2. Matemática -
Estatística. 3. Ensino Médio. 4. Matemática. I. Título.

CDD 519.5



UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS DA NATUREZA
CENTRO DE EDUCAÇÃO ABERTA E À DISTÂNCIA
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL

Dissertação de Mestrado submetida à coordenação Acadêmica Institucional, na Universidade Federal do Piauí, do Programa de Mestrado Profissional em Matemática em Rede Nacional para obtenção do grau de **mestre em matemática** intitulada:
A importância do Ensino de Regressão Linear Simples no Ensino Médio: um estudo com alunos do 3º ano do Ensino Médio IPMA - Imperatriz defendida
por Paulo Cardoso Sales em
13.10.2014 e aprovada pela banca constituída pelos professores:

Valmaio R. S. Ferraz
Presidente da Banca Examinadora

José Carlos de Oliveira Neto
Examinador

Luiza Rosalva Oliveira dos Santos
Examinador Externo

Dedico esse trabalho à minha esposa Aline, aos meus pais, Sabino e Antonia e aos meus irmãos.

Agradecimentos

A Deus, que por sua presença e força sempre me abençoa e capacita para tudo aquilo que Ele me destina.

À minha esposa Aline e a toda minha família que, com muito carinho e apoio, não mediram esforços para que eu chegasse a esta etapa de minha vida.

À professora Dr^a. Valmária Rocha da Silva Ferraz pela paciência na orientação e incentivo que tornaram possível a conclusão desta dissertação.

A todos os professores do curso, que foram tão importantes nessa etapa da minha vida e no desenvolvimento deste trabalho.

À diretoria e servidores do IFMA/Imperatriz, por ter autorizado e apoiado o desenvolvimento deste projeto.

Aos alunos do IFMA/ Imperatriz pela cooperação e interesse em fazer parte do trabalho.

Aos amigos e colegas, pelo incentivo e pelo apoio constantes.

Agradeço à CAPES e ao IMPA por acreditarem na formação de professores através do PROFMAT.

A todos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

“A sabedoria, porém, lá do alto é, primeiramente, pura; depois, pacífica, indulgente, tratável, plena de misericórdia e de bons frutos, imparcial, sem fingimento”.

(Tiago 3.17).

Resumo

O ensino da estatística nas escolas auxilia a formação de um cidadão integral, com a capacidade de relacionar o saber escolar com a prática social. Este trabalho tem por objetivo apresentar uma abordagem geral acerca dos conceitos de Correlação e Regressão Linear Simples e sua aplicabilidade no cotidiano dos educandos e avaliar a compreensão dos conteúdos abordados e a capacidade de aplicar tais assuntos em situações-problemas como também de interrelacionar com as diversas áreas do saber. Para o mesmo foi realizado um minicurso com ministração de aulas expositivas e práticas no laboratório de informática. Participaram deste 19 alunos do 3º ano do ensino médio integrado do Instituto Federal do Maranhão campus Imperatriz (IFMA) no período de 17 a 21 de março de 2014. Com base neste estudo, conclui-se que os alunos tiveram um bom aproveitamento do conteúdo ministrado, não apresentaram dificuldades na utilização dos softwares e na aplicação do conteúdo no cotidiano.

Palavras- Chave: Correlação, Regressão Linear, Estatística.

Abstract

The teaching of statistics in schools assists the formation of a full citizen, with the ability to relate school knowledge with social practice. This research aims to present a general approach on the concepts of correlation and linear regression and its applicability in everyday life of students and assess the comprehension of the subject the was explained and the ability to apply such matters in problem situations as well as to interrelate with the various areas of knowledge. At the same research a short course o with presentation of lectures and practices in the computer lab were performed. The participants were 19 students of the 3rd year of integrated high school's Instituto Federal do Maranhão campus Imperatriz (IFMA) for the period 17 to 21 March 2014. Based on this study, it is concluded that the students had a good use of the subject taught, they had no difficulties in using the software and the applicability of the subjects in daily life.

Key Words: Correlation, Regression, Statistics.

Sumário

Resumo	iv
Abstract	v
Lista de tabelas	viii
Lista de figuras	ix
Introdução	1
1 Fundamentação Teórica	5
1.1 Correlação Linear	6
1.2 Regressão Linear Simples	9
2 Proposta Metodológica	22
2.1 Descrição do Projeto	23
3 Resultados	27
3.1 Perfil do aluno	27
3.2 Avaliação do Aluno sobre o Projeto	30
3.3 Avaliação dos alunos	32
3.3.1 Resultados da 1ª etapa	32
3.3.2 Resultados da 2ª etapa	35
3.4 Resultado Geral	35
4 Conclusão	37
Referências Bibliográficas	39

Lista de Tabelas

1.1	Valores de propaganda e vendas	12
1.2	Tabela de valores	14
1.3	População e homicídios	16
4.1	Tabela: Valores Críticos do Coeficiente de Correlação r de Pearson	45

Lista de Figuras

1.1	Correlação Positiva	6
1.2	Correlação Negativa	6
1.3	Correlação Positiva Perfeita	7
1.4	Correlação Negativa Perfeita	7
1.5	Ausência de Correlação	7
1.6	Correlação Curvilínea	7
1.7	Gráfico Preço x venda	11
1.8	Gráfico de dispersão	13
1.9	Reta de regressão linear simples	15
1.10	Microsoft office excel 2013	17
1.11	Inserir linha de tendência	17
1.12	Reta ajustada	18
1.13	Reta e os parâmetros estimados de regressão	19
1.14	Altura e massa	20
1.15	Gráfico de dispesão	20
1.16	Reta de regressão linear simples	21
3.1	Gênero dos alunos participantes	28
3.2	Divisão quanto ao curso	28
3.3	Pesquisa	29
3.4	Área de formação	29
3.5	Contribuição dos softwares utilizados	30
3.6	Aplicação no cotidiano	31
3.7	Avaliação dos discentes	34

Introdução

Conforme [25] a palavra Estatística, do latim “status”(estado, situação), por um longo período foi concebida como ciência dos negócios do Estado. Enquanto que [23] acrescentam ainda que a Estatística pode ser entendida como um conjunto de técnicas que permite, de forma sistemática, organizar, descrever, analisar, e interpretar dados oriundos de estudos ou experimentos, realizados em qualquer área do conhecimento.

Na antiguidade, os governos tinham interesse de registrar o número de habitantes, de nascimentos, de óbitos, e também de estimar as riquezas individual e social, utilizando assim, noções de Estatística. Outra finalidade era para a exigência de impostos [25].

Embora muitas vezes as origens das informações fossem distorcidas e muitas se perdessem no perpassar da História, há vestígios do uso da Estatística na China a datar da antiguidade e também por civilizações indígenas como os incas, maias e astecas. Os egípcios investigavam a população e as riquezas materiais de seu país antes de iniciarem a construção de alguma pirâmide .

Ao passo que na contemporaneidade, usufrui-se de muitos recursos construídos ao longo de sua história com alguns aperfeiçoamentos e com eventos que surgiram posteriormente.

A exemplo disso temos os censos que são realizados até os dias atuais. Com os dados do censo o governo e a sociedade em geral pode discernir os locais onde é mais considerável a aplicação de capitais nos setores de saúde, educação, habitação, transportes etc.; constatar em que lugares se carecem de programas de fomento ao progresso econômico, como estabelecimento de polos industriais; definir onde instalar suas fábricas, shopping centers, supermercados, escolas etc. [17].

Acrescenta-se a isso, o fato de que a estatística também cumpre diversas atribuições, tais como no assessoria na operacionalização das hipóteses da pesquisa, na caracterização da população a ser estudada e das variáveis, da mesma maneira que na coleta de dados

e análise. Nessa proposta, a estatística constitui-se uma ferramenta imprescindível na construção de “estudantes-cientistas”, dado que tal ciência estende-se para várias áreas do conhecimento humano.

Na Educação Básica, o ensino de Estatística é papel dos educadores de Matemática. No PCN [6] de matemática para o Ensino Médio há uma compartimentação em três blocos: Álgebra - Números e Funções, Geometria e Medidas e Análise de dados (inclui Contagem, Probabilidade e Estatística). Comparando com o Ensino Fundamental no Ensino Médio, o estudo da Estatística é mais minucioso e o estudante deverá conhecer seus tópicos e, além disso, habilitado a interpretá-los criticamente tomando decisões [4].

Dentre as diversas áreas da estatística, a análise de regressão é uma das técnicas estatísticas mais empregadas no auxílio a pesquisa modelando a relação existente entre duas ou mais variáveis, pois averigua a existência e o grau de dependência estatística entre as variáveis aleatórias, ou seja, aquelas que têm distribuição de probabilidade. Essa técnica pode ser definida, segundo [10] como modelo que descreve a relação entre duas variáveis através do gráfico e da equação da reta que melhor representa a relação.

Por outro lado, a análise de coeficiente de correlação fundamenta-se na mensuração do grau ou intensidade de associação entre duas variáveis. Na ocasião em que se pode constatar que a variação de uma variável está de alguma maneira associada com a variação da outra, assim pode-se afirmar que as duas variáveis estão correlacionadas. Por coeficiente de correlação entende-se que é uma medida numérica da força da relação entre duas variáveis que representam dados quantitativos. [28].

Tendo em vista os fatos supracitados, este trabalho abordou, por meio de um minicurso, os seguintes temas da área de estatística: coeficiente de correlação linear e regressão linear simples. Assim, a pretensão desse projeto não é explicar a estatística por completo, mas visualizar situações cotidianas dos alunos para fazer uso do conhecimento teórico de regressão linear simples, esclarecer aquele fenômeno mediante a teoria, além de instrumentalizar os participantes com as ferramentas estatísticas disponíveis na planilha de dados e o aplicativo estatístico R. O público alvo foi alunos de 3º ano do ensino médio integrado do Instituto Federal do Maranhão campus Imperatriz (IFMA) no período de 17 a 21 de março de 2014.

JUSTIFICATIVA

Atualmente, diversas pessoas adquirem, diariamente, uma vasta quantidade de informações e, geralmente, utilizam técnicas estatísticas para correlacionar dados e, a partir destes, tirar conclusões. Ademais, outras áreas do conhecimento, tais como a Geografia, Biologia, Química, etc, se desempenham com a utilização da linguagem estatística.

Para [8] discutir sobre a importância da Estatística nos cursos de Exatas é redundante, mas ressalta que sua relevância se mostra não menos destacada nas áreas de Ciências Sociais e também nas Biomédicas e de Saúde, visto que a Estatística apresenta-se como disciplina obrigatória nos diversos campos de formação acadêmica, nas Áreas de Ciências Exatas, Humanas e Biológicas. Desse modo, como afirma [22], vislumbramos o ensino da estatística assumindo um papel de instrumento de operacionalização, de integração entre diversas disciplinas e mesmo entre variadas temáticas dentro da própria Matemática.

Isto posto, observa-se que existe uma interdisciplinariedade de saberes digna de ser mantida e estimulada pelos educadores. Segundo [6], a interdisciplinaridade não tem o anseio em criar novas disciplinas, mas utilizar os conhecimentos de várias para resolver problemas ou conceber fenômenos sob diferentes pontos de vista, tendo assim uma função “instrumental”, valendo-se de um saber diretamente útil e utilizável para responder questões e problemas sociais. Pode ser vista, também, como sendo mais um artifício para ampliar as incontáveis possibilidades de interação entre disciplinas.

Por conseguinte, o ensino da estatística nas escolas auxilia a formação de um cidadão integral, com a capacidade de relacionar o saber escolar com a prática social. Deve-se ressaltar, que a escola tem o encargo de formar cidadãos, e observa-se que esse cuidado quanto à formação de um cidadão para adquirir subsídios para uma boa convivência em sociedade foi suscitada pelos autores das Propostas Curriculares Nacionais (PCN, 1998), justificando que grande parte das informações trazidas pela mídia, geralmente, é apresentada em forma de tabelas e gráficos, portanto, para que todo cidadão seja letrado estatisticamente, deverá ter alguma instrução em Estatística, podendo, assim, se aprimorar cognitivamente numa sociedade cada vez mais tecnológica e informatizada.

Deve-se agregar a ideia de que o aluno também pode aplicar os temas abordados em estatística na produção científica tanto na escolar como no Ensino Superior, além da busca por questionamentos acerca de questões cotidianas. Tais aspectos muitas vezes não são incitados, pois como afirma [15], a relevância que se atribui aos temas e conteúdos que

necessitam ser findados no desenrolar dos anos letivos, e o esforço pelos resultados produz um professor “transmissor” desses conteúdos, ao invés de estimular os questionamentos, à elaboração de perguntas pelo aluno e a instrumentação pela pesquisa.

Soma-se a isto, que a pesquisa provoca no estudante maturidade e estimula sua criticidade além de promover uma maior responsabilidade do mesmo com relação ao seu ambiente e com o mundo. Para [11] a pesquisa produz “desenvolvimento da autonomia intelectual”.

Considerando ainda os pressupostos da LDB, (Lei nº 9.394, de 20 de dezembro de 1996) que dentre os objetivos da educação básica destacam-se: o desenvolvimento do educando, a formação comum indispensável para o exercício da cidadania e fornecimento de meios para progredir no trabalho e em estudos posteriores.

Dessa forma compreende-se a relevância de tal projeto, pois servirá também como instrumento e incentivo ao aluno a pesquisa. Visto que tais conhecimentos reproduzirão a compreensão dos fundamentos científico-tecnológicos dos processos produtivos, relacionando a teoria com a prática.

OBJETIVOS

Geral

- Apresentar uma abordagem geral acerca dos conceitos de Correlação e regressão linear simples e sua aplicabilidade no cotidiano dos educandos.

Específicos

- Introduzir a utilização das técnicas de regressão para descrição de dados, estimação de parâmetros, previsão, ajuste de modelo de regressão linear simples.
- Auxiliar no reconhecimento da aplicabilidade da regressão linear em situações cotidianas;
- Avaliar a compreensão dos conteúdos abordados e a capacidade de aplicar tais assuntos em situações-problemas e interrelacionar com as diversas áreas do saber;

Capítulo 1

Fundamentação Teórica

Segundo [19] O termo regressão foi criado por Francis Galton. Em um artigo famoso, Galton verificou que, embora existisse uma tendência de que pais altos tivessem filhos altos e pais baixos tivessem filhos baixos, a estatura média das crianças nascidas de pais com uma dada altura tendia a mover-se ou “regredir” à altura média da população em geral, ou seja, a altura dos filhos de pais mais altos ou mais baixos que o padrão tende a mover-se no sentido da altura média da população. A lei da regressão universal de Galton foi ratificada por Karl Pearson, que coletou mais de mil registros de altura de membros de grupos familiares. Ele constatou que a altura média dos filhos de um grupo de pais altos era menor do que a de seus pais e que a altura média de um grupo de filhos de pais baixos era maior do que a de seus pais; portanto, filhos de pais altos e baixos “regrediam” igualmente à altura média de todos os homens. Nas palavras de Galton, isso era uma “regressão à mediocridade”.

Contudo, a interpretação moderna da regressão é bastante diferente. De modo geral, podemos afirmar: A análise de regressão diz respeito ao estudo da dependência de uma variável, a variável dependente, em relação a uma ou mais variáveis, as variáveis explanatórias, visando estimar e/ou prever o valor médio (da população) da primeira em termos dos valores conhecidos ou fixados (em amostragens repetidas) das segundas.

A análise de regressão tem por objetivo traduzir as relações entre as variáveis através de modelos matemáticos produzidos a partir dos dados que resultam da análise dessas variáveis, que possibilitam a previsão do comportamento de uma variável em função de outra.

Chamamos de correlação simples se a relação se verificar entre duas variáveis e se for

entre mais de três variáveis diz-se múltipla.

Como exemplos podemos citar a relação entre massa e altura dos indivíduos, o preço de venda de carros seminovos e o ano de fabricação, homicídios e tamanho da população, a quantidade de fertilizante usada e a produtividade de determinada área agrícola. [26].

1.1 Correlação Linear

A correlação linear (r) é mensuração numérica da intensidade da relação entre duas variáveis que exibem dados quantitativos. Há uma correlação entre duas variáveis na ocasião em que uma delas está relacionada com a outra de alguma maneira. Empregando dados de r e a seguir, usam esse valor para comprovar que há (ou não) uma relação entre duas variáveis [28].

Por intermédio do diagrama de dispersão que é um gráfico cartesiano em que cada um dos eixos corresponde às variáveis correlacionadas podemos conceber visualmente se esses valores seguem um padrão. A variável dependente (Y) situa-se no eixo vertical e o eixo das abscissas é designado para a variável independente (X). Os pares ordenados formam uma nuvem de pontos. A configuração geométrica do diagrama de dispersão pode estar associada a uma linha reta (correlação linear), uma linha curva (correlação curvilínea) ou, ainda, ter os pontos esparsos de maneira que não definam nenhuma configuração linear; neste último caso, não há correlação.

Figura 1.1: Correlação Positiva

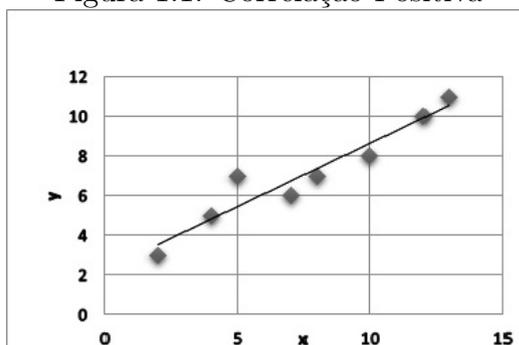


Figura 1.2: Correlação Negativa

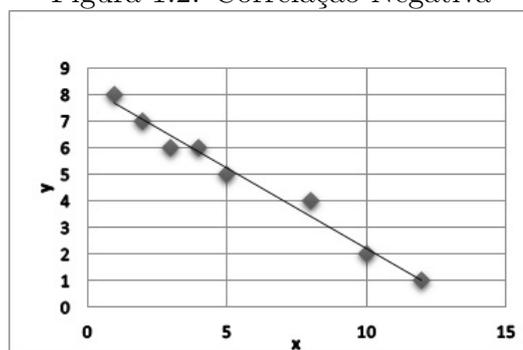


Figura 1.3: Correlação Positiva Perfeita

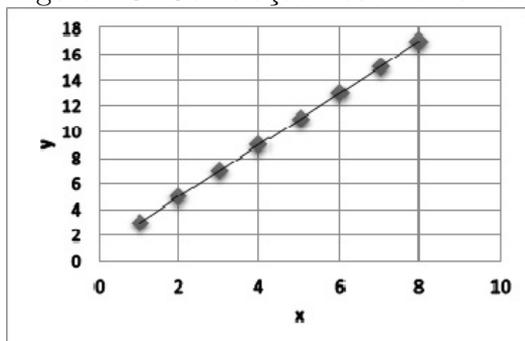


Figura 1.4: Correlação Negativa Perfeita

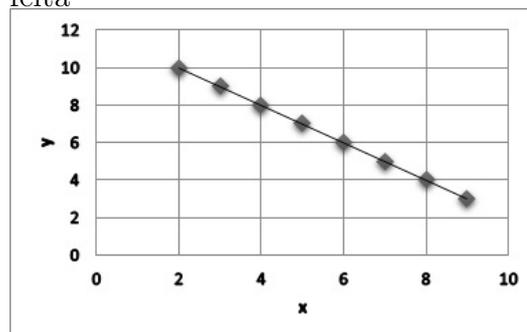


Figura 1.5: Ausência de Correlação

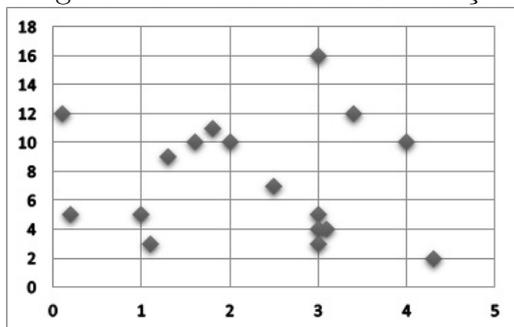
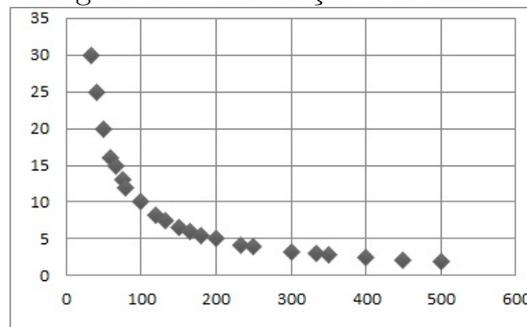


Figura 1.6: Correlação Curvilínea



Como a apuração visual de diagramas de dispersão é altamente subjetiva, necessitamos de medidas mais precisas e objetivas. Usamos o coeficiente de correlação linear r , que é proveitoso para identificar padrões lineares.

Exposição: o coeficiente de correlação linear r afere o grau da relação linear entre os valores quantitativos emparelhados x e y em uma amostra. Seu valor é calculado usando-se a fórmula

$$r = \frac{\sum (x_i)(y_i) - n(\bar{x})(\bar{y})}{\sqrt{(\sum (x_i)^2 - n\bar{x}^2)(\sum (y_i)^2 - n\bar{y}^2)}} \quad (1.1)$$

onde $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ e $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$

O coeficiente de correlação linear é, por vezes, denominado de coeficiente de correlação do produto de momento de Pearson, em homenagem a Karl Pearson que o explicou originalmente [28]. Como o coeficiente de correlação linear r é calculado com o uso de dados amostrais, ele é uma estatística amostral usada para medir a intensidade da correlação linear entre x e y . Se tivéssemos todos os pares de valores populacionais x e y , o resultado

da Fórmula 1.1 seria um parâmetro populacional, representado por ρ .

Interpretação de r : Variação Explicada

Se deduzimos que há uma correlação linear entre x e y , podemos encontrar uma equação linear que expresse y em termos de x , e essa equação pode ser usada para antever valores de y para valores dados de x . Entretanto um valor previsto de y não será inevitavelmente o resultado exato, porque, além de x , há outros fatores que afetam y , tais como variação aleatória e outras características não abrangidas no estudo. O valor de r^2 é a proporção da variação em y que é elucidada pela relação linear entre x e y . Se r estiver muito próximo de 0, concluímos que não há correlação linear significativa entre x e y , mas se r estiver próximo de -1 ou de +1 concluímos que há uma relação linear significativa entre x e y .

Observando apenas o sinal do coeficiente de Pearson, $r > 0$, indica que as duas variáveis variam no mesmo sentido (quando uma aumenta a outra também aumenta), enquanto que $r < 0$, indica que as duas variáveis variam em sentidos contrários (quando uma aumenta a outra variável diminui). Interpretações de próximo de 0, ou 1, ou -1 são vagas, de modo que usamos o seguinte critério de decisão bem específico.

Teste de Hipótese

Em [28] encontra-se dois métodos para o teste de hipótese do coeficiente de correlação linear são eles:

O Método 1 usa a distribuição t de Student com uma estatística de teste que tem a forma

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

O critério de decisão é rejeitar a hipótese nula ($\rho=0$) se o valor absoluto da estatística de teste ultrapassar ($|t| > \text{valor crítico}$) os valores críticos; a rejeição de $\rho=0$ significa que há premissa considerável para respaldar a afirmativa de uma correlação linear entre as duas variáveis. Se a estatística de teste não ultrapassar os valores críticos, então deixamos de rejeitar $\rho=0$; ou seja, não há indício razoável para se concluir que haja uma correlação linear entre as variáveis.

Enquanto que método 2 compara o valor do coeficiente de correlação encontrado pela fórmula 1.1 com os valores da Tabela de Valores Críticos do Coeficiente de Correlação r de Pearson (Tabela 4.1 anexo III). O critério de decisão é rejeitar a hipótese nula ($\rho=0$)

se o valor absoluto do coeficiente de correlação linear r ultrapassar os valores críticos; a rejeição de $\rho=0$ exprime que há prova satisfatória para sustentar a afirmativa de uma correlação linear entre as duas variáveis. Se o valor absoluto do coeficiente de correlação linear r não exceder o valor crítico, então considera-se a hipótese nula ($\rho=0$); isto é, não há evidência suficiente para se concluir que haja uma correlação linear entre as variáveis.

Teste de hipótese para correlação

$H_0: \rho=0$ (Não há correlação linear)

$H_1: \rho \neq 0$ (Há correlação linear)

- Método 1: A Estatística de Teste é t

$$\text{Estatística de teste } t: t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

Valores críticos: usa-se a tabela t de Student (anexo II) com $n-2$ graus de liberdade

Valor P : usa-se a tabela t de Student (anexo II) com $n-2$ graus de liberdade. Con-

clusão: se $|t| >$ valor crítico da tabela, rejeite H_0 e conclui-se que há uma correlação linear. Se $|t| <$ valor crítico, deixe de rejeitar H_0 ; não há comprovação suficiente para se excluir que haja uma correlação linear.

- Método 2: A Estatística de Teste r : Compara-se o próprio coeficiente de correlação de Pearson obtido pela fórmula 1.1 com os valores críticos encontrados na Tabela 4.1 (anexo III)

Estatística de teste: r ; Valores críticos: Consultar a Tabela de Valores Críticos (anexo III). Conclusão: se $|r| >$ valor crítico da Tabela 4.1 (anexo III), rejeite H_0 e conclui-se que há uma correlação linear. Se $|r|$ valor crítico, aceita-se H_0 ; não há evidência suficiente para se excluir que haja uma correlação linear.

1.2 Regressão Linear Simples

Explicar a relação entre duas variáveis por intermédio do gráfico e da equação da reta que melhor representa a relação. Esta reta é chamada reta de regressão, e sua equação é chamada equação de regressão. Dados valores amostrais emparelhados, obteremos valores estimados do intercepto b_0 e da inclinação b_1 , de modo a podermos identificar uma reta com equação $\hat{y} = b_0 + b_1x$. Sob condições propícias, essa equação pode ser usada para se fazerem prognósticos.

Em algumas situações, duas variáveis se relacionam de uma maneira determinística, o que implica dizer que, dado um valor de uma variável, o valor da outra variável fica automaticamente determinado sem qualquer erro. Por exemplo, o custo total y de um item com preço de tabela de x e com taxa de venda de 5% pode ser obtido utilizando-se a equação determinística $y = 1,05x$. Se um item tem preço de 50,00 reais, seu custo total será 52,50 reais. Tais funções são consideradas extensivamente em cursos de álgebra. Nos modelos probabilísticos uma variável não é inteiramente determinada por outra variável. A equação da regressão expressa uma relação entre x , chamada variável explanatória ou variável previsora ou variável independente e y , chamada de variável resposta ou variável dependente. A equação típica de uma reta $y = bx + a$ é expressa na forma de $\hat{y} = b_0 + b_1x$, onde, b_0 é o intercepto e b_1 é a inclinação. A notação dada mostra que b_0 e b_1 são estatísticas amostrais usadas para estimarem-se os parâmetros populacionais b_0 e b_1 . Usaremos dados amostrais emparelhados para determinar a equação de regressão. Usando apenas dados amostrais, não podemos achar os valores exatos dos parâmetros populacionais b_0 e b_1 , mas podemos usar os dados amostrais para estimá-los com b_0 e b_1 , que são encontrados com o uso das fórmulas [28].

$$b_1 = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n (x_i)^2 - (\sum_{i=1}^n x_i)^2} \text{ e } b_0 = \bar{y} - b_1 \bar{x} \quad (1.2)$$

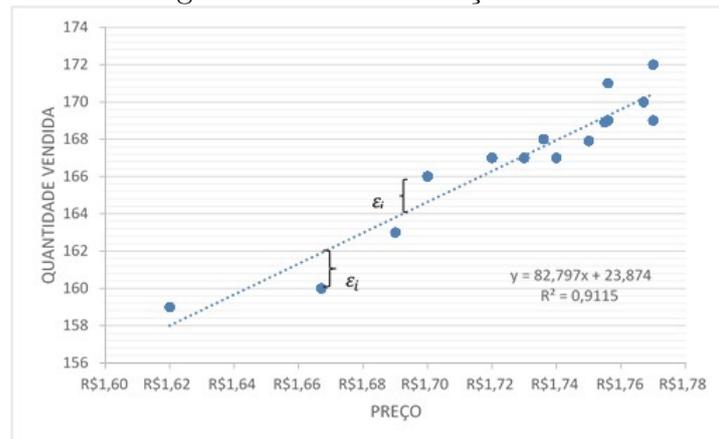
Na prática, os pontos experimentais terão uma variação em torno da linha representativa dessa função, em razão da existência de uma variação aleatória adicional, chamada de variação residual. Essa função de regressão, por consequência, nos dá o valor médio de uma das variáveis em função da outra. Consideremos que a linha de regressão seja uma reta e tentamos determinar a regressão y em função de x . A função pretendida será da forma: $Y = a + bx$. Os parâmetros a e b da reta teórica podem ser estimados através dos pontos experimentais fornecidos por uma amostra, obtendo uma reta estimativa da forma: $\hat{y} = b_0 + b_1x$. Onde b_0 é a estimativa do parâmetro a e b_1 também chamado de coeficiente de regressão linear, é a estimativa do parâmetro b . O símbolo \hat{y} é utilizado para uma diferenciação dos valores dados pela reta estimativa, das ordenadas dos pontos experimentalmente obtidos.

Para verificar a qualidade do ajustamento calcula-se o coeficiente de determinação, r^2 bastando para tal elevar ao quadrado o valor do coeficiente de correlação.

O emprego do princípio de ajustamento linear induz ao procedimento de mínimos quadrados, segundo o qual a reta a ser adotada deverá ser aquela que torna mínima a soma

dos quadrados das distâncias da reta aos pontos experimentais, medidas no sentido da variação aleatória, que, como estamos considerando somente a variável y como aleatória, essas distâncias serão medidas na direção vertical. Abaixo temos um gráfico de dispersão que relaciona o preço de certo produto e a quantidade vendida [26].

Figura 1.7: Gráfico Preço x venda



A soma de quadrados dos desvios das observações em relação à linha de regressão é:

$Y_i = a + bx_i + \epsilon_i$ os termos de erro ($\epsilon_1, \epsilon_2, \epsilon_3, \dots, \epsilon_n$) são variáveis aleatórias independentes; $E\epsilon_i = 0$; $V\epsilon_i = \sigma^2$; e ϵ_i tem distribuição normal ($i = 1, 2, \dots, n$).

$$\text{resíduo} = \epsilon$$

$$\epsilon = y_{\text{observado}} - y_{\text{previsto}}$$

$$\epsilon = y - \hat{y}.$$

Uma reta satisfaz a propriedade dos mínimos quadrados se a soma dos quadrados dos resíduos é a menor possível. Devemos procurar a reta para a qual se consiga minimizar:

Devemos procurar a reta para a qual se consiga minimizar:

$$\sum_{i=1}^n \epsilon_i^2 \tag{1.3}$$

$$Q = \min \sum_{i=1}^n (\epsilon_i^2) = \min \sum_{i=1}^n (y_i - a - bx_i)^2 \tag{1.4}$$

Os estimadores de mínimos quadrados (EMQ) de a e b denotados por b_0 e b_1 devem satisfazer as seguintes equações:

$$\frac{\partial Q}{\partial a} |_{b_0 b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i)$$

$$\frac{\partial Q}{\partial b} |_{b_0 b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) x_i$$

Resolvendo o sistema acima obtemos:

$$b_1 = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \tag{1.5}$$

$$b_0 = \bar{y} - b_1 \bar{x} \tag{1.6}$$

Exemplo 1: O diretor de vendas de uma rede de varejo nacional precisa averiguar a relação entre o investimento em propaganda e as vendas da empresa. O objetivo é dispor de uma equação matemática que propicia realizar projeções de vendas a partir de investimentos em propaganda. O departamento de vendas preparou a tabela seguinte com as vendas em milhões de reais e os investimentos em propaganda em milhões reais dos dez últimos anos [21].

Tabela 1.1: Valores de propaganda e vendas

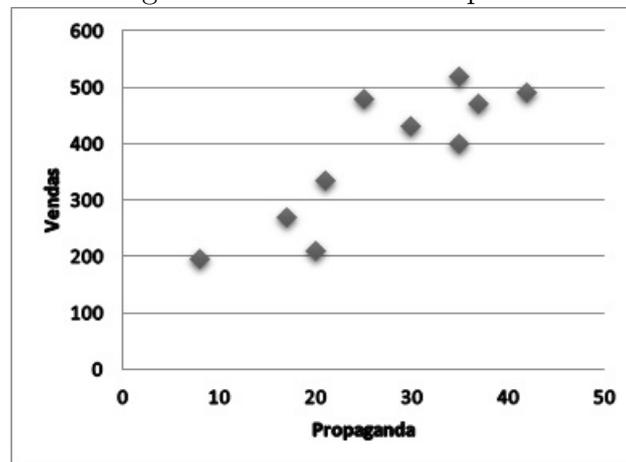
Propaganda	30	21	35	42	37	20	8	17	35	25
Vendas	430	335	520	490	470	210	195	270	400	480

- Represente os dados num diagrama de dispersão
- Indique a variável independente e a variável dependente
- A partir da observação do diagrama de dispersão, classifique a correlação existente entre o propagandas e vendas.
- Determine a equação da reta de regressão e represente a reta no diagrama de dispersão.
- Classifique a correlação linear entre as variáveis apresentadas, com base no valor do coeficiente de Pearson.
- Calcule e interprete o coeficiente de determinação.
- Interprete os parâmetros e estime a projeção de vendas a partir de um investimento de 12,5 milhões em propaganda.

Resolução

a)

Figura 1.8: Gráfico de dispersão



b) Vendas: Variável dependente

Propaganda: Variável independente

c) Observando o diagrama de dispersão e a tabela de dados, constatamos que de fato, à medida que aumenta o investimento em propaganda ocorre um aumento nas vendas.

d) Para obter a equação da reta de regressão que traduz a relação entre as duas variáveis determinam-se os parâmetros, b_0 e b_1 , que minimizam o quadrado dos erros através das expressões

$$b_0 = \bar{y} - b_1 \bar{x} \text{ e } b_1 = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

Para colaborar no cálculo do coeficiente de correlação linear construímos a seguinte tabela baseada na tabela de valores da questão.

Dos dados temos $\bar{x} = 27$ e $\bar{y} = 380$

Tabela 1.2: Tabela de valores

Propaganda (X)	Venda Y	$x_i y_i$	x_i^2	y_i^2
30	430	12900	900	184900
21	335	7035	441	112225
35	520	18200	1225	270400
42	490	20580	1764	240100
37	470	17390	1369	220900
20	210	4200	400	44100
8	195	1560	64	38025
17	270	4590	289	72900
35	400	14000	1225	160000
25	480	12000	625	230400
$\sum x_i = 270$	$\sum y_i = 3800$	$\sum x_i y_i = 112455$	$\sum x_i^2 = 108302$	$\sum y_i^2 = 1573950$

Vimos que os valores apresentados na tabela e o diagrama de dispersão sugerem que uma reta pode ser utilizada para representar o crescimento proporcional do valor investido em propagandas e o valor das vendas. Para obtermos essa reta, precisamos calcular as estimativas b_0 e b_1 .

$$b_0 = \bar{y} - b_1 \bar{x} \text{ e } b_1 = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

$$b_1 = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

$$b_1 = \frac{10 \cdot 112455 - 270 \cdot 3800}{10 \cdot 8302 - 270^2}$$

$$b_1 = 9,7381$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

$$b_0 = 3809,7381 \cdot 27$$

$$b_0 = 117,0713$$

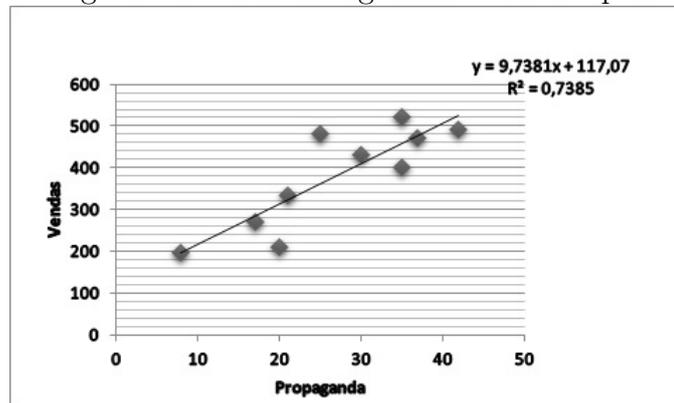
Portanto dado $X = x_i$, a reta ajustada fornece valores \hat{Y}_i , dados por:

$$\hat{Y}_i = b_0 + b_1 x_i$$

$$\hat{Y} = 117,0713 + 9,7381x_i$$

A figura a seguir mostra os dados originais e a reta ajustada. O gráfico propõe que o modelo de regressão linear simples e o ajuste parece adequado aos dados.

Figura 1.9: Reta de regressão linear simples



e) Concordante com apresentado anteriormente o coeficiente de correlação linear entre propaganda X e as vendas Y é calculado por: $r = \frac{\sum (x_i)(y_i) - n(x_i)(y_i)}{\sqrt{(\sum (x_i)^2 - n\bar{x}^2)(\sum (y_i)^2 - n\bar{y}^2)}}$.
 Cálculo do coeficiente de Correlação Linear (r):

$$r = \frac{\sum (x_i)(y_i) - n(x_i)(y_i)}{\sqrt{(\sum (x_i)^2 - n\bar{x}^2)(\sum (y_i)^2 - n\bar{y}^2)}} \Rightarrow$$

$$r = \frac{112455 - 10.27.380}{\sqrt{(8302 - 10.27^2).(1573950 - 10.380^2)}} \Rightarrow$$

$$r = 0,859366125$$

Para fazer o teste de hipóteses vamos comparar o valor de r encontrado com o valor crítico da tabela 4.1 (com $\alpha = 0,05$) e encontramos o valor crítico de r é 0,632.

H_0 : não há correlação

H_1 : há correlação.

Temos $0,859 > 0,632$ portanto $r > r_{\text{crítico}}$. Logo devemos rejeitar H_0 . A vista disso, vemos que as vendas são sensivelmente influenciada pela variação do investimento em

propaganda.

f) Calculando o coeficiente de determinação, $r^2 = 0,738$ significa que 73,8% da variação total da variável dependente é explicada pela reta.

g) A interpretação dos parâmetros da equação da reta é feita da seguinte forma: o valor em vendas esperado quando não há investimento x é 117,07 (obtido substituindo $x_i = 0$ na equação calculada acima); de outro modo, o aumento de 1 milhão de reais em propaganda x implica em um ganho médio esperado de R \$ 9,7381 milhões.

Para estimar o valor em vendas, substitui-se x por 12,5, na equação da reta de regressão e obtém-se:

$$\hat{Y} = 117,0713 + 9,7381x_i$$

$$\hat{Y} = 117,0713 + 9,7381.12,5$$

$$\hat{Y} = 238,79755$$

Exemplo 2: A tabela a seguir lista os números de assassinatos e os tamanhos das populações (em centenas de milhares) em grandes cidades americanas durante um anos recente (com base em dados do New York Times). Com o auxílio do *Microsoft Office Excel* Estime a equação da reta e ache a melhor previsão para o tamanho da população de uma cidade com 120 assassinatos.

Tabela 1.3: População e homicídios

Homicídios	258	264	402	253	111	648	288	654	256	60	590
População	4	6	9	6	3	29	15	38	20	6	81

Solução: Para termos uma reta estimada vamos seguir os passos

1. Construir a tabela no *Microsoft Office Excel* (homicídios x população) e em seguida inserir gráfico de dispersão (figura 1.10).
2. Na botão Elementos do gráfico escolhemos adicionar linha de tendência(figura 1.11).

3. e por fim na opção formatar linha de tendência marque as opções exibir equações no gráfico e exibir o valor de R-quadrado no gráfico (figura 1.12).

Figura 1.10: Microsoft office excel 2013

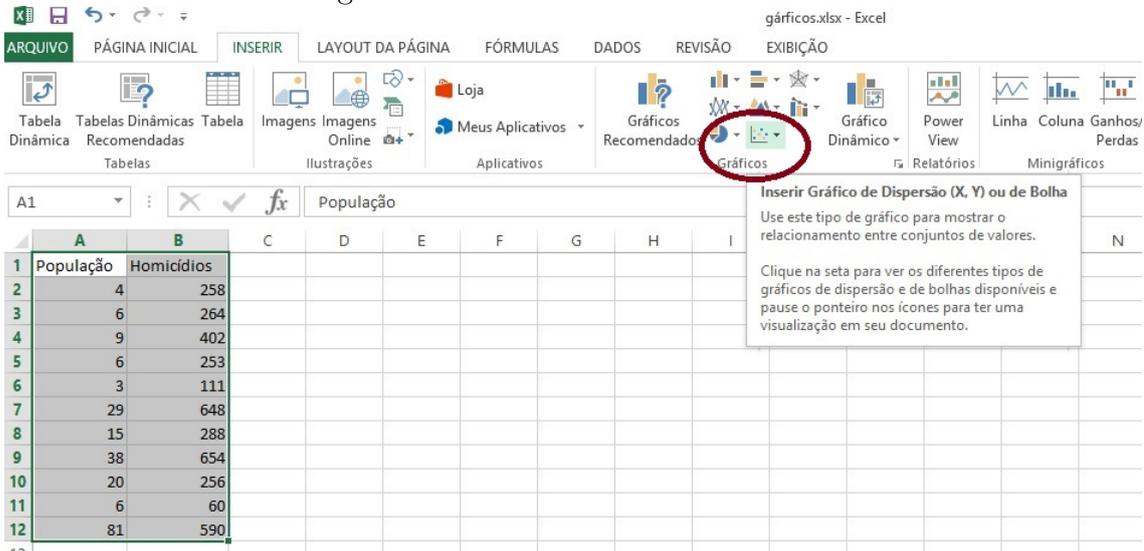


Figura 1.11: Inserir linha de tendência

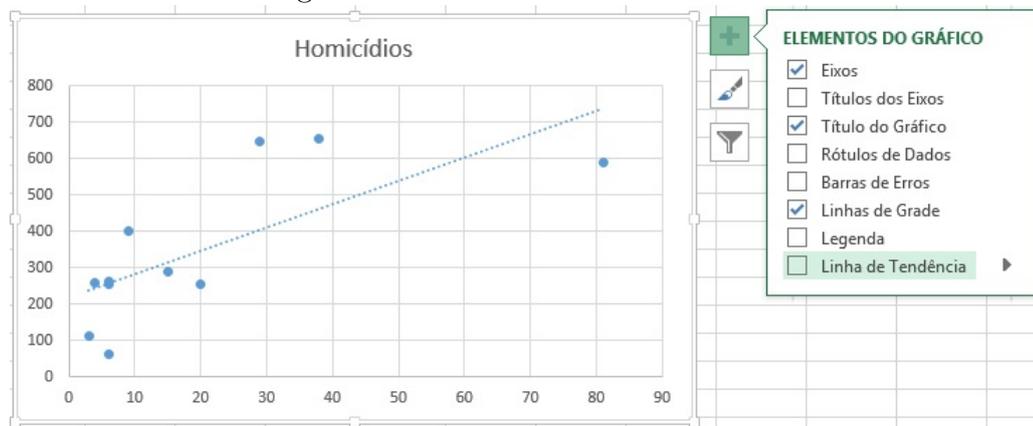
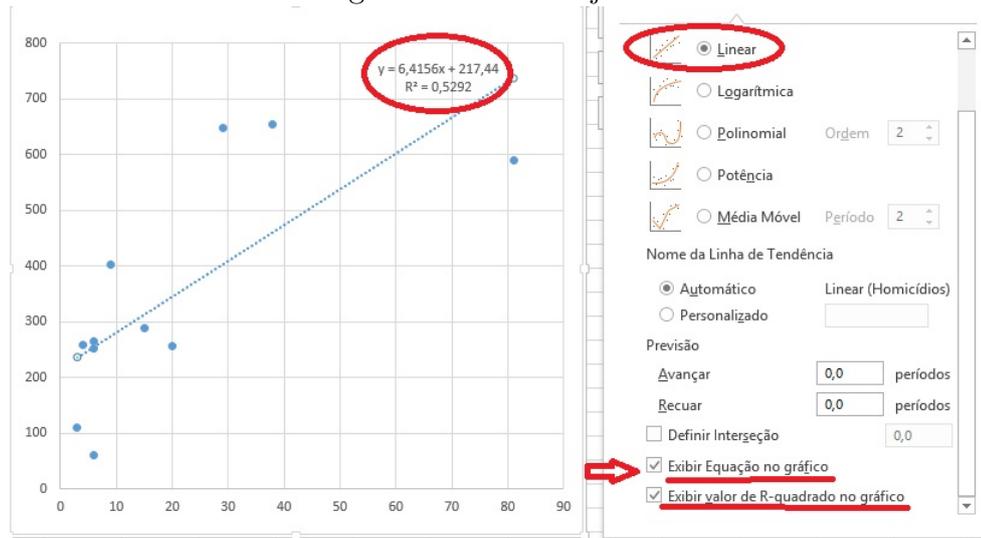


Figura 1.12: Reta ajustada



Com a equação apresentada pelo comando linha de tendência o leitor poderá representar a quantidade de homicídios em função da população. Na figura acima temos o valor de $r^2 = 0,5292$ o que indica que o valor de r é $0,7274$. Para fazer o teste de hipóteses vamos comparar o valor de r encontrado com o valor crítico da tabela 4.1 (com $\alpha = 0,05$) e encontramos o valor crítico de r é $0,602$.

H_0 : não há correlação

H_1 : há correlação.

Temos $0,7274 > 0,602$ (valor crítico encontrado na Tabela 4.1) portanto $r > r_{\text{crítico}}$. Logo devemos rejeitar H_0 ao nível de significância de 5%. A vista disso, vemos que a quantidade de homicídios é sensivelmente influenciada pela quantidade da população.

Podemos determinar os parâmetros da reta e o grau de adequação do modelo.

Utilizando as funções:

- INTERCEPÇÃO (Val_Conhecidos_y; Val_Conhecidos_x): dá o coeficiente de regressão b_0 da reta de regressão linear $\hat{y} = b_0 + b_1x$ considerando os valores das amostras informadas nos argumentos Val_Conhecidos_y; Val_Conhecidos_x.
- INCLINAÇÃO (Val_Conhecidos_y; Val_Conhecidos_x): dá o coeficiente de regressão b_1 da reta de regressão linear $\hat{y} = b_0 + b_1x$ considerando os valores das amostras informadas nos argumentos Val_Conhecidos_y; Val_Conhecidos_x.

- PREVISÃO (x;Val_Conhecidos_y;Val_Conhecidos_x): dá o valor projetado para o valor registrado no argumento x considerando a reta de regressão linear simples $\hat{y} = b_0 + b_1x$ considerando os valores das amostras informadas nos argumentos Val_Conhecidos_y;Val_Conhecidos_x.
- RQUAD (Val_Conhecidos_y;Val_Conhecidos_x): dá o coeficiente de determinação r^2 da reta de regressão linear $\hat{y} = b_0 + b_1x$ considerando os valores das amostras informadas nos argumentos Val_Conhecidos_y;Val_Conhecidos_x.

Teremos então as seguintes fórmulas:

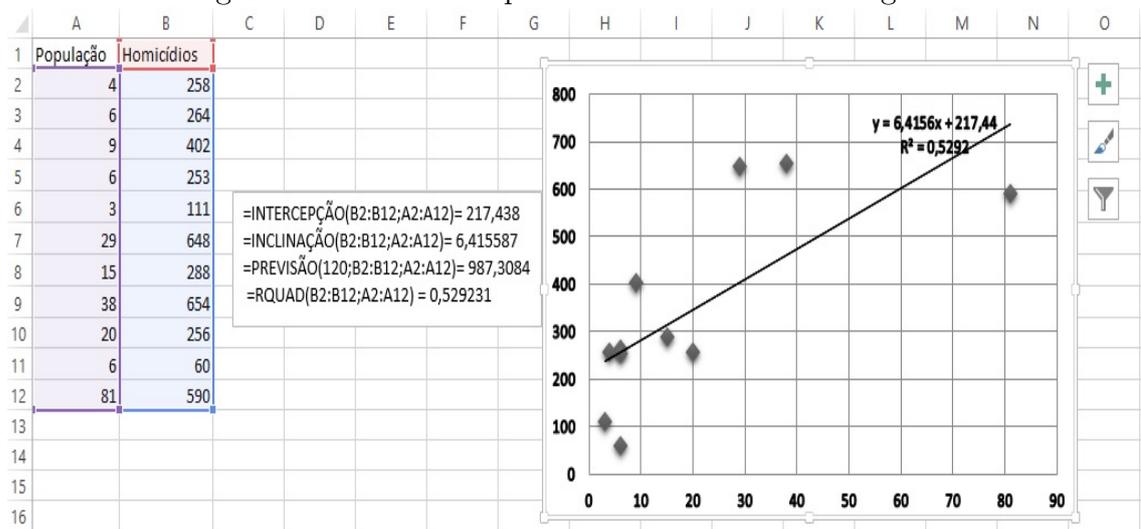
=INTERCEPÇÃO(B2:B12;A2:A12) = 217,4379605

=INCLINAÇÃO(B2:B12;A2:A12) = 6,415587256

=PREVISÃO(120;B2:B12;A2:A12) = 435,5679272

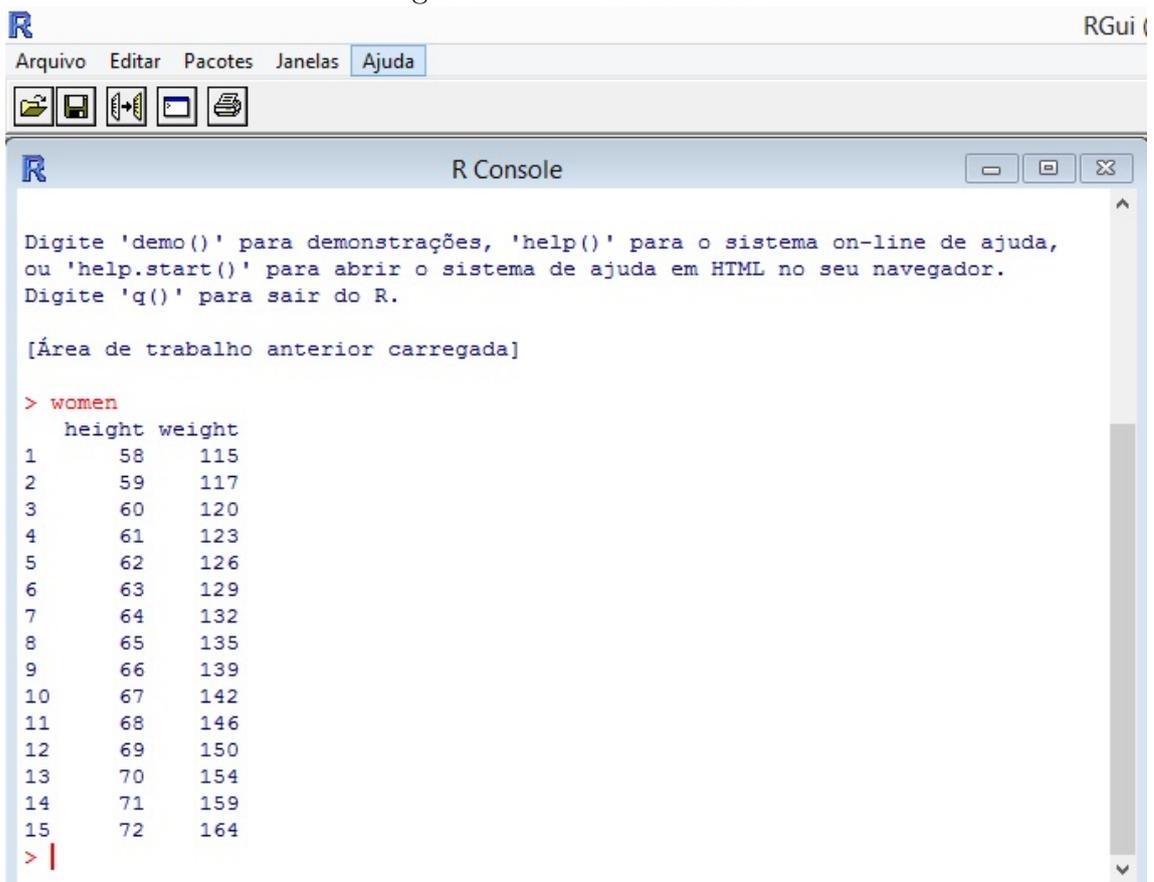
=RQUAD (B2:B12;A2:A12) = 0,529231226

Figura 1.13: Reta e os parâmetros estimados de regressão



Para exemplificar como fazer uma regressão no R-project vamos usar dados do banco de dados do próprio programa. Abordaremos uma questão sobre a altura e massa de uma amostra de 15 mulheres. Usaremos o R-project para encontrar o gráfico de dispersão, calcular o coeficiente de correlação e os parâmetros da reta. O estudo de regressão pode ser feito, inicialmente com a definição do modelo. Para isso vamos visualizar os pontos plotados em um diagrama de dispersão. Usaremos os seguintes comandos no R

Figura 1.14: Altura e massa

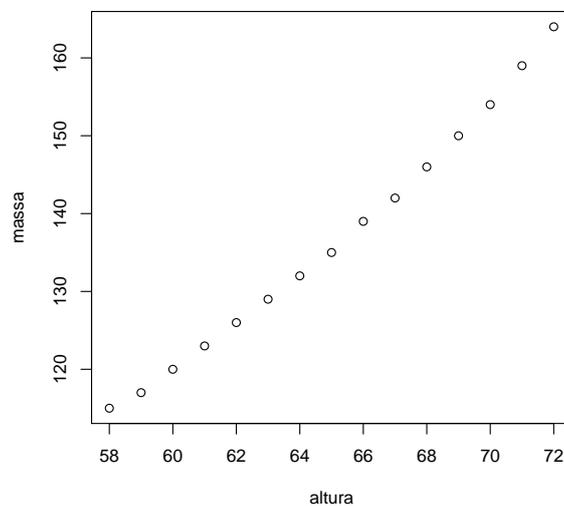


```
Altura = women[,1]
```

```
Massa = women[,2]
```

```
plot(Altura,Massa) # a variável independente deve vir primeiro
```

Figura 1.15: Gráfico de dispersão



O diagrama sugere uma tendência linear dos dados. Então vamos montar um modelo de regressão linear simples (simples, pois existe apenas uma variável independente “altura” relacionada a variação da variável dependente “massa”). Montando o modelo:

```
> reglin <- lm(Massa ~ Altura)
```

```
> reglin
```

```
lm(formula = Massa ~ Altura)
```

```
Coefficients:
```

```
(Intercept)      Altura
```

```
-87.52           3.45
```

Da saída acima podemos tirar duas informações: o valor do intercepto e o valor que representa um coeficiente de relação entre o massa e a altura, ou seja, quanto a massa irá variar para cada variação unitária da altura. Esses valores são comumente chamados de b_0 e b_1 respectivamente. Logo podemos concluir que o modelo matemático/estatístico desta regressão é: $\hat{Y} = -87,52 + 3.45x$

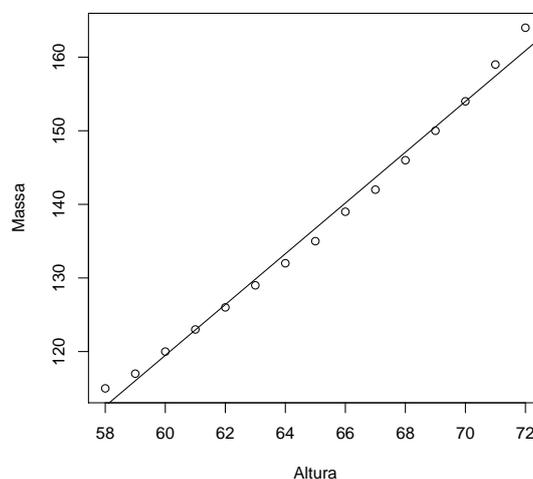
Podemos obter os valores estimados (preditos) pelos valores tabelados de “altura” da função com o comando: `predict(reglin)` e usamos o comando abaixo para encontrar a figura 2.16

```
plot(x,y,xlab="Altura",ylab="Massa")
```

```
fit=lm(y~x)
```

```
fit
```

Figura 1.16: Retas de regressão linear simples



Capítulo 2

Proposta Metodológica

Conforme já registrado, muitas entraves no que tange a iniciação científica são comuns no cotidiano escolar, além do esporádico uso de técnicas de ensino que fujam dos métodos tradicionais e proporcionem ao aluno uma incitação à utilização do conhecimento adquirido na escola com o cotidiano. Tendo em vista tal obstáculo, este projeto de extensão foi desenvolvido com base na Modelagem Matemática pois presume-se que a Modelagem Matemática seja uma ferramenta importante no aperfeiçoamento do ensino aprendizagem, pois fomenta nos alunos a busca por formas e soluções para os problemas. A palavra modelagem advém do latim *modelus* e descreve a típica forma com que os seres humanos copiam a realidade [27].

Segundo [1] a modelagem consiste, essencialmente, na arte de transformar situações da realidade em problemas matemáticos cujas soluções devem ser interpretadas na linguagem usual.

A modelagem matemática é uma das atividades científicas e pedagógicas que favorecem essa prática interdisciplinar, possibilitam a colaboração entre matemáticos e especialistas de outras áreas do conhecimento e incentiva um ensino mais significativo da matemática. A modelagem matemática, além de ser matemática é, também, epistemologia, uma vez que os modelos matemáticos, por meio de um processo investigativo que não se limita ao senso comum, visam entender e explicar fatos e fenômenos observados na realidade, [9].

O método de desenvolvimento de uma atividade de modelagem matemática abrange algumas etapas fundamentais. A princípio, deve-se designar um tema central para ser desenvolvido pelos alunos, e recolher dados gerais e quantitativos que consigam colaborar com o levantamento de hipóteses com propósito de elaborar problemas conforme interesse

dos grupos de alunos.

Depois de selecionar as variáveis fundamentais envolvidas nos problemas e formular as hipóteses, se faz necessário à sistematização dos princípios que serão adotados na resolução dos modelos e a apreciação da solução (analítica, e se possível, graficamente). Para findar, dependendo do objetivo, estabelecer a validação dos modelos, comparando os resultados obtidos com os dados coletados.

Dessa forma compreende-se que a relevância de tal projeto pois servirá de instrumento e incentivo ao aluno a pesquisa. Visto que tais conhecimentos reproduzirão a compreensão dos fundamentos científico-tecnológicos dos processos produtivos, relacionando a teoria com a prática.

O projeto ocorreu em sala de aula como também no laboratório de informática. No primeiro momento o professor apresentou os objetivos e explanou o conteúdo. Depois, no laboratório de informática com o auxílio dos softwares *Microsoft Office Excel* e *R Project*, inicialmente foram instruídos como utilizar os softwares, e a partir de então os alunos passaram a resolver exercícios com o auxílio dos mesmos.

Nesse sentido, viu-se então como é relevante a realização de uma pesquisa com o uso da informática em sala de aula na intenção de contribuir com uma investigação e avaliação que proporcione melhorias significativas no desenvolvimento da prática docente com o uso da informática como elemento facilitador do processo ensino-aprendizagem.

Diante do exposto, este trabalho abordou os seguintes temas da área de estatística: coeficiente de correlação linear e regressão linear simples. Assim, a pretensão desse projeto não é explanar a estatística por completo, mas visualizar situações cotidianas dos alunos para fazer uso do conhecimento teórico de regressão linear simples esclarecer aquele fenômeno mediante a teoria, além de instrumentalizar os participantes com as ferramentas estatísticas disponíveis no planilha de dados e o aplicativo estatístico R. O público alvo foi alunos de 3º ano do ensino médio integrado do Instituto Federal do Maranhão campus Imperatriz (IFMA) no período de 17 a 21 de março de 2014.

2.1 Descrição do Projeto

Como referimos anteriormente o método utilizado foi a Modelagem Matemática, por se tratar de um procedimento que objetiva a acepção e percepção dos mais diversos

fenômenos do nosso cotidiano. A Modelagem pode ser aplicada em praticamente todas as temáticas relacionadas à Matemática. Para o ensino da Estatística, torna-se um poderoso mecanismo, visto que podemos aliar os conceitos formais estudados aos assuntos de interesse dos alunos. No momento em que esse interesse vem à tona, o real aprendizado passa a efetuar-se e ser apreciável.

O trabalho de pesquisa foi realizado no Instituto Federal de Educação, Ciência e Tecnologia do Maranhão (IFMA)/ Campus Imperatriz-Ma, localizado na rua Newton Belo, S/N no bairro Vila Maria, no período de 17 a 21 de março de 2014 com carga horária total de 10 horas/aulas (O projeto foi realizado de forma presencial). A divulgação do projeto em redes sociais e nos murais do instituto e a seleção dos participantes do minicurso foi feita por meio do lançamento de um edital com informações sobre o mesmo e com os critérios de participação do minicurso.

Para atender a proposta da pesquisa, o trabalho foi realizado com a participação de alunos de 3ª séries do Cursos Técnicos Integrado em Informática e Segurança do Trabalho. Para tanto, se apresentaram 21 alunos que voluntariamente participaram das atividades; todas realizadas no período vespertino, visto que esses alunos tem suas atividades regulares no período matutino. Para o experimento dessa etapa foram elaboradas atividades específicas para cada aula desenvolvida em sala de aula, laboratório de informática. No primeiro caso, o conteúdo teórico foi transmitido por meio de aulas expositivas. Algumas aulas foram dedicadas à resolução orientada de listas de exercícios, visando a assimilação conteúdo e no segundo a utilização de softwares (*Microsoft Office Excel* e *R Project*) no laboratório de informática. É importante destacar que a turma escolhida foi bem receptiva ao trabalho apresentado, se mostrando interessada em participar de todas as atividades propostas. O pesquisador é professor titular da turma selecionada. Todo o trabalho desenvolvido teve o apoio da direção da escola, coordenadores, técnicos do laboratório de informática.

Descrição do 1º encontro

Inicialmente, foi entregue pelo professor a cada participante da pesquisa um termo de consentimento e livre esclarecido na qual se comprometia em manter em sigilo o nome verdadeiro de cada dos participantes e acesso aos resultados da pesquisa. Nessa primeira abordagem foi realizada uma apresentação inicial acerca da proposta, destacando a importância do estudo de Correlação e Regressão Linear Simples e sua aplicação prática,

com o objetivo principal de mostrar aos educandos, as diversas formas de aplicar a Estatística no nosso dia-a-dia. Logo em seguida, foi apresentado por meio de slides gráficos de dispersão mostrando se era possível visualizar ou não uma possível relação entre as duas variáveis de cada tabela.

Por conseguinte o professor deu instruções de como achar o gráfico de dispersão no *Microsoft Office Excel* e *R Project* como exemplos foram usado números do banco de dados do *R Project*. Depois das instruções iniciais foi entregue a cada aluno uma folha com tabelas (com duas variáveis) para que fizessem no *Microsoft Office Excel* e *R Project*, gráficos de dispersão. A partir daí, os alunos foram questionados sobre os gráficos se os mesmos seguiam um padrão (direção linear para cima ou para abaixo ou curvilínea).

Descrição do 2º encontro

No segundo encontro foi trabalhado acerca do conceito do coeficiente de correlação linear r e suas propriedades. Logo em seguida solicitou-se aos alunos que calculassem (utilizando apenas calculadoras) o coeficiente de correlação linear das tabelas que receberam no encontro anterior.

Em seguida foi utilizado o programa estatístico *R Project* para calcular o valor do coeficiente de correlação e comparado com o valores que foram encontrados. No final do encontro foi apresentado dois testes de hipótese para correlação o teste t e o teste r .

Descrição do 3º encontro

Trabalhou-se a definição de regressão linear simples e de como se encontrar sua equação. Notou-se grande interesse no entendimento dos recursos utilizados no tocante a aplicação dos métodos em situações de suas respectivas áreas de formação técnica e cotidianas em seguida os alunos usaram o laboratório de informática do Instituto para pesquisarem situações e/ou trabalhos na área de interesse contendo o uso de regressão linear simples ou nos quais seria possível sua utilização.

À vista disso, utilizaram os trabalhos pesquisados para calcular correlação - fazer estudo da correlação, inserir gráfico de dispersão exibindo a regressão e o coeficiente de determinação (R^2) e a equação de regressão linear.

Ao final desse encontro foram discutidos novamente os conceitos e definições trabalhados anteriormente.

Descrição do 4º encontro

O Professor intermediou uma discussão sobre o tema com base nas pesquisas reali-

zadas pelos alunos (como mencionado no tópico anterior) como forma de motivar a sua participação na execução de projetos de pesquisa.

Foi realizada uma atividade em sala com resolução de exercícios investigativos sobre regressão linear nas quais fizeram o gráfico de dispersão; escolheram (pelo menos) duas variáveis quantitativas discretas cuja finalidade foi: calcular correlação, fazer estudo da correlação, inserir gráfico de dispersão exibindo a regressão e o coeficiente de determinação. Para atingir tal desfecho, realizaram os cálculos por meio dos softwares e interpretaram os resultados obtidos.

Ao final do quarto encontro, todos os envolvidos já demonstravam condições satisfatórias para a realização de cálculos nas diversas situações que seriam propostas a posteriori.

Descrição do 5º encontro

Nessa aula, foi feita uma avaliação do projeto por parte dos alunos através de um Questionário de Avaliação (Anexo IV) e em seguida foi aplicado uma avaliação com o objetivo de mensurar o conhecimento dos alunos sobre o conteúdo abordado no projeto.

Capítulo 3

Resultados

Consoante ao que foi elucidado, para seleção dos alunos participantes do projeto foi lançado um edital no qual disponibilizava 25 vagas e o período de inscrição se estendeu do dia 10 de março de 2014 ao dia 14 de março de 2014 e nesse período houve 25 inscrições (todos preenchiam os pré-requisitos informados no edital). Embora tenham se apresentados na primeira aula 21 alunos o trabalho foi realizado com 19 alunos pois esses tiveram uma frequência regular durante o projeto.

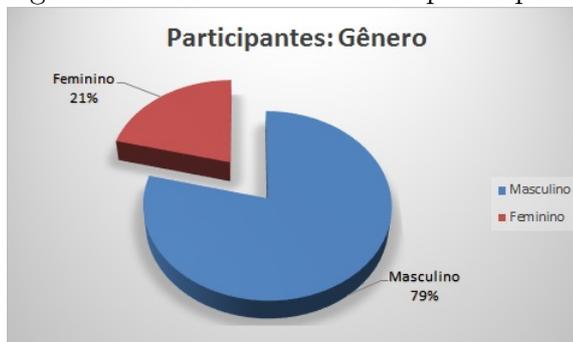
A transcrição dos resultados da pesquisa é resultado dos dados coletados no final do projeto por meio de questionário de avaliação do projeto (Anexo IV), teste de avaliação do discente, da observação no decorrer da aplicação do projeto na sala de aula e laboratório de informática. O questionário de avaliação do projeto (Anexo IV) foi aplicado aos 19 alunos que participaram integralmente do mesmo (todos os discente se dispuseram a responder o questionário).

3.1 Perfil do aluno

Sobre o gênero dos participantes temos que dos 19 alunos que frequentaram o minicurso 21%(4 alunos) eram do sexo feminino enquanto 79%(15) eram do sexo masculino (Figura 3.1). Tais resultados são esperados pela maioria daqueles que lecionam na área de exatas, pois observa-se uma predominância do sexo masculino. De acordo com [13] a Matemática como ciência sempre têm sido estereotipadas como campo masculino. Tradicionalmente, os matemáticos são homens e a maioria dos professores de matemática na escola secundária e média são do sexo masculino. Além do mais, a matemática como um

campo de estudo é frequentemente ligada a empregos masculinos em carreiras militares e nas Engenharias.

Figura 3.1: Gênero dos alunos participantes



Sobre Curso de formação técnica integrado ao Ensino Médio que os alunos estão matriculados observou-se que 42% (8) são do curso de Informática e 58% (11) são do curso de segurança do trabalho (Figura 3.2).

Figura 3.2: Divisão quanto ao curso



Em relação ao vínculo com pesquisa, sete alunos (37%) têm/tiveram algum vínculo com pesquisa desses 4 são/foram bolsistas e 12 alunos (63%) ainda não realizaram pesquisa (figura 3.3).

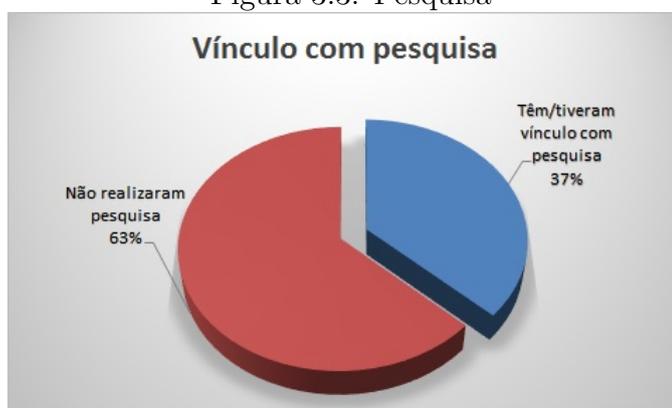
Para os alunos que declararam vínculo com alguma pesquisa foi feita a seguinte pergunta: “Com base no que foi estudado, você acha que aplicaria os conhecimentos adquiridos em seu projeto/pesquisa?” para tal indagação 71% responderam sim enquanto 29% responderam que não era possível aplicar em seus projetos.

Segundo [14] o Brasil ainda não possibilita aos seus adolescentes jovens um programa

sistematizado de fomento à prática da pesquisa científica. O que ocorre são iniciativas isoladas, e não políticas concretas de incentivo à pesquisa

No Brasil, o processo de estabelecimento de cientistas revela -se tardio, uma vez que as instituições de nível superior são caminhos predominantes de engajamento, triagem e treino dos futuros pesquisadores. Os indivíduos vocacionados para pesquisa são escolhidos durante a graduação e a pós-graduação, sendo que esta última se estabeleceu no Brasil, etapa obrigatória de formação do pesquisador.

Figura 3.3: Pesquisa



No que concerne ao interesse em fazer um curso superior todos responderam “sim”. Alia-se a isto os resultados obtidos no que tange a área de interesse em seguir carreira por parte dos educandos. Pode-se observar na figura 3.4 que 11% (2) pretendem fazer graduação na área de Humanas, 21% (4) na área de saúde e 68% (13) intencionam seguir na área de exatas e/ou tecnológicas

Figura 3.4: Área de formação



3.2 Avaliação do Aluno sobre o Projeto

No que tange ao usos dos softwares (*R Project* e *Microsoft Office Excel*) e o uso do computador nas aulas ministradas observou-se que houve contribuição dos softwares (*R-Project* e *Excel*) e quinze (15) alunos admitiram que ajudou muito no aprendizado o uso do computador no projeto desenvolvido enquanto três (3) classificaram que contribuiu razoavelmente no seu aprendizado e um (1) aluno considerou auxiliou um pouco no seu aprendizado.(figura 3.5)

Figura 3.5: Contribuição dos softwares utilizados



O uso do computador pode assessorar grandemente no progresso cognitivo dos alunos, proporcionando a prática de novas categorias de atividades e de novas maneiras de raciocinar e proceder.

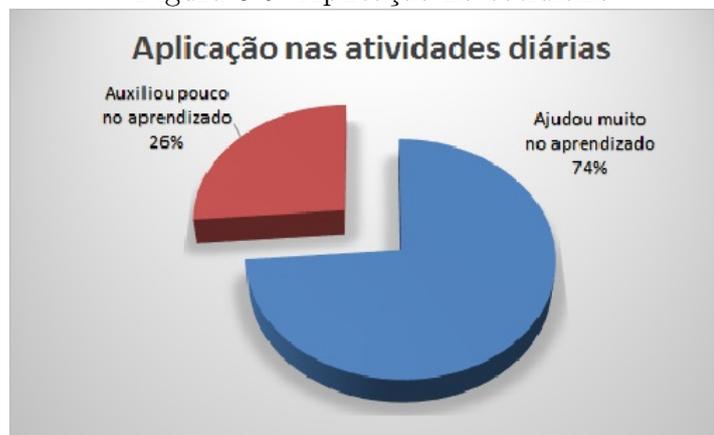
Contudo, este potencial ainda não tem sido apropriadamente experimentado e incorporado ao cotidiano da prática escolar, ficando restringido a discussões teóricas e acadêmicas. Para as escolas e para muitos professores, as tecnologias perdura a ser um corpo estranho, que acarreta sobretudo incomodidade, para outros, falta de recursos. O receio de ficar para trás tem levado as escolas a investirem na compra de equipamentos, muitas vezes deixando para segundo plano o ensino das novas tecnologias. [12].

Em pleno século XXI com o uso da informática e todo avanço tecnológico, nota-se ainda na prática docente algumas debilidades nessa área, onde o educador, provavelmente na sua formação, não teve a possibilidade de acesso à instrução necessária para o uso desta ferramenta. Vale salientar que os alunos já possuem uma maior facilidade de acolhimento e uso destas tecnologias, pois já nasceram nesta era digital e desde cedo já experimentam o uso destas, fazendo com que o auxílio do computador lhe seja mais aprazível e simples de ser utilizado [18].

No tocante a aplicabilidade de Correlação e Regressão Linear Simples em situações

cotidianas os dezenove (19) alunos responderam que o assunto estudado podem lhes auxiliar em situações cotidianas. Quanto à inclusão de situações-problemas no dia a dia nas atividades aplicadas pelo professor, 79% (14) alunos consideraram que auxilia muito o aprendizado, 21% cinco (5) consideraram auxilia um pouco.(figura 3.6)

Figura 3.6: Aplicação no cotidiano



Dessa forma, Rocha (2001) afirma que o ensino de Matemática contribui para a construção de juízo e conclue que podemos dominar e debater economia e política, podemos constatar e replicar às ilegalidades, relacionar as diferenças salariais, julgar os índices e os gráficos difundidos na imprensa. Outrossim, a Matemática pode nos assessorar na tomada de decisões e na compreensão do uso da tecnologia.

Sucedesse desse comportamento que os conhecimentos matemáticos, em suma, devem servir e ter uma magnitude de experiência social mais qualificada, pois as peculiaridades e a amplitude dos saberes matemáticos impelem uma maior percepção e discernimento na interação com a cultura que transpassa, conferindo às pessoas maior autonomia e desprendimento social.

Ao passo que quando indagados sobre o conhecimento prévio do conteúdo explanado, apenas um aluno declarou que já dominava algo a respeito de correlação e regressão linear simples. O fato de a Estatística já ser desenvolvida nos primeiros anos de escolaridade é um elemento positivo, pois quando o indivíduo chegar no Ensino Superior, poderá ampliar ainda mais seus conhecimentos, estudando outras variedades de análise e testes estatísticos que hoje não são exercidos devido à falta de conhecimentos básicos acerca da Estatística [2].

Segundo [16] o desenvolvimento de calculadoras cada vez mais completas, computadores e softwares cada vez mais desenvolvidos, faz com que a análise estatística esteja

eclodindo, e seu estudo tem sido elemento basilar para educação cedida nas escolas.

3.3 Avaliação dos alunos

Para aferir o conhecimento adquirido foi aplicada uma avaliação de 4 questões onde os alunos responderam em duas etapas: 1ª etapa: o aluno podia responder com auxílio de calculadoras, porém, sem o uso do computador; 2ª etapa: responderam as mesmas questões usando os programas *R- Project* e/ou *Microsoft Office Excel 2013* Explanaremos primeiramente os resultados da 1ª etapa

3.3.1 Resultados da 1ª etapa

A 1ª avaliação aconteceu em sala de aula no dia 21 de março de 2014, abaixo temos a explanação dos resultados e o gráfico de resultados na figura 3.7:

- QUESTÃO 1

Abaixo estão listados os orçamentos (em milhões de dólares) e as bilheterias (em milhões de dólares) de filmes selecionados aleatoriamente (com base em dados da Motion Picture Association of America).

ORÇAMENTO	62	90	50	35	200	100	90
RECEITA	65	94	48	57	601	146	47

- Encontre o coeficiente de correlação linear de Pearson e classifique a correlação linear entre as variáveis apresentadas.
- Qual a proporção da variação no valor da receita pode ser explicada pela variação do orçamento?

Para a resolução dessa questão é necessário que os alunos saibam identificar uma possível relação entre as variáveis e classificar e interpretar essa relação com base no coeficiente de correlação de Pearson. A primeira análise, feita com base no problema relacionado a correlação linear. Sendo esta satisfatória, pois demonstrou que somente 1 aluno errou a questão; 2 acertaram parcialmente e 16 alunos acertaram totalmente a questão, como mostra o gráfico.

• QUESTÃO 2

Num estudo sobre o efeito da exposição prolongada a fumaça na saúde de bombeiros, foi observada uma amostra de 7 bombeiros e recolhidos dados relativos à incidência de doenças respiratórias. Os dados seguintes referem-se ao tempo de serviço como bombeiro (em anos) e ao tempo de internamento hospitalar (em dias) em consequência de problemas respiratórios.

anos de serviço	15	8	12	4	9	10	7
Tempo de internamento	21	7	9	1	13	14	4

Supondo que a relação entre as duas variáveis pode ser traduzida por um modelo linear, A partir da amostra estimou-se, através dos métodos dos mínimos quadrados, o modelo de regressão linear simples que traduz a relação entre as duas variáveis.

Sendo x o tempo de serviço em anos e y a quantidade de dias de internamento obteve-se: $\sum x = 65$; $\sum y = 69$; $\sum x^2 = 679$; $\sum y^2 = 953$; $\sum xy = 768$

Sendo $\hat{Y} = b_0 + 1,688x$ determine o valor de b_0 e Estime o número de dias de internamento hospitalar em consequência de problemas respiratórios de um bombeiro com 5 anos de serviço?

Na questão 2 em que o aluno precisava fazer estimativa de previsão, 18 alunos acertaram e somente 1 aluno errou.

• QUESTÃO 3

Para verificar se existe relação entre a renda familiar em salários mínimos e o número de filhos, foi coletada uma amostra de 8 famílias em uma cidade. Os resultados obtidos estão na tabela a seguir:

Renda familiar	12	14	15	17	23	27	34	43
Filhos	3	2	2	1	1	0	0	0

Há uma correlação linear entre renda familiar e quantidade de filhos? Explique fazendo o teste de hipóteses $\alpha = 5\%$

Na 3ª questão foi dada uma situação para que o aluno fizesse o teste de hipóteses para correlação. Dentre os 19 alunos, 17 alcançaram o objetivo e apenas 2 falharam.

• QUESTÃO 4

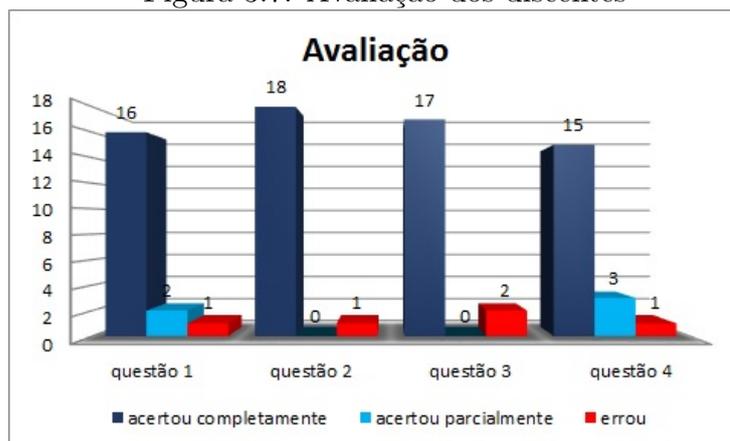
A tabela seguinte apresenta o bônus (R\$) recebido mensalmente pelos funcionários de uma empresa e o respectivo tempo de serviço (em meses). Com base nos valores da tabela:

- a) Indique a variável independente e a variável dependente
- b) A partir da observação do diagrama de dispersão, classifique a correlação existente entre o Bônus mensal e o Tempo de serviço.
- c) Determine a equação da reta de regressão e represente a reta no diagrama de dispersão.

Bolsa mensal	147	125	160	118	149	128	150	145	115	140	152	155
Tempo de serviço	56	42	72	36	63	47	55	49	38	42	68	60

Houve também um bom resultado na resolução do problema de Estimação da equação da reta pois apenas 1 deles errou o problema e 3 alunos acertaram parcialmente enquanto os demais (15) acertaram totalmente.

Figura 3.7: Avaliação dos discentes



3.3.2 Resultados da 2ª etapa

A segunda etapa da avaliação foi realizada no laboratório de informática. Os alunos tiveram de resolver as mesmas questões, contudo, com a utilização de softwares. Pôde-se constatar que todos os alunos conseguiram responder corretamente a todas as questões propostas. Vale ressaltar que os nossos alunos já possuem uma maior facilidade de aceitação e uso destas tecnologias, pois já nasceram nesta era digital e desde cedo já vivenciam o uso destas, fazendo com que os recursos do computador lhe sejam mais agradáveis e fáceis de serem utilizados. Assim, não demonstraram nenhuma dificuldade na resolução com computadores.

“Não se pode negar o impacto provocado pela tecnologia de informação e comunicação na configuração da sociedade atual. Por um lado, tem-se a inserção dessa tecnologia no dia-a-dia da sociedade, a exigir indivíduos com capacitação para bem usá-la; por outro lado, tem-se nessa mesma tecnologia um recurso que pode subsidiar o processo de aprendizagem da Matemática. É importante contemplar uma formação escolar nesses dois sentidos, ou seja, a Matemática como ferramenta para entender a tecnologia, e a tecnologia como ferramenta para entender a Matemática.” [7]

Os softwares podem ser considerados uma ferramenta que produz eficácia e agilidade, todavia, conforme afirma [20] o que é demasiadamente compreensível é desmotivante para os aprendizes por não proporcionar desafio. Dificultar a aprendizagem intencionalmente e com objetivos bem definidos pode compelir os estudantes a realizar processamentos complementar que contribuirão na formação de uma representação mais elaborada na memória. Dessa forma, facilitar as atividades a ponto de refrear o erro, tem a inconveniência de não dar aos estudantes oportunidade de eles fazerem reestruturações do conhecimento a partir de seus erros.

3.4 Resultado Geral

Para analisar esses momentos, recorreremos nas observações realizadas em sala de aula até o presente momento, bem como no desempenho dos alunos nas provas e exercícios realizados. Durante o desenvolvimento do minicurso, procuramos criar um ambiente participativo e de constante diálogo, entre os alunos-alunos, alunos-professor, para que assim pudesse haver a troca de ideias, participação entre os componentes e uma melhor aprendizagem.

Desta forma, em alguns momentos os próprios alunos trouxeram situações problemas com possibilidade de uso da correlação simples e regressão linear simples e foram desenvolvidas atividades práticas no intuito de promover a criticidade, a criatividade, a tomada de decisões e a autonomia.

Durante a realização do projeto percebeu-se, pelo diálogo entre alunos, exemplos de situações em que poderíamos trabalhar no cotidiano e que tinham relação com as suas respectivas formações técnicas.

O assunto abordado despertou grande interesse dos alunos em aplicar seus conhecimentos sobre a correlação e regressão linear simples. Observou-se também o interesse por parte deles em continuar o estudo de outras correlações e regressões. Foi o que ocorreu, por exemplo, com alunos do curso técnico em informática que ambicionam desenvolver um software para calcular correlação e regressão.

Em virtude das situações supracitadas podemos afirmar que o Ensino de Correlação Simples e Regressão Linear Simples apresenta-se como um excelente e importante instrumento para o aluno de Ensino Médio por diversos motivos, entre eles podemos citar o fato de proporcionar aos alunos a oportunidade de instrumentalizar os alunos para pesquisas e conseqüentemente, analisar os resultados da mesma trabalhar com dados estatísticos reais, e discutir, não só com o professor, mas também com os colegas, assuntos pertinentes ao seu dia-a-dia.

Outro fato importante foi a utilização do *Microsoft Office Excel* e do *R- Project* no cálculo de correlação e regressão linear simples. Tanto o uso do computador na construção dos gráficos, tanto o uso da Internet para fins de pesquisa aproximou o conteúdo trabalhado em sala de aula do dia-a-dia dos alunos, pois quase todos utilizam a Internet, e conseqüentemente o computador, para navegar em site de relacionamentos. Soma-se a isto que pouco utilizavam essa ferramenta para pesquisa visto que ainda não conheciam todas as ferramentas do *Microsoft Office Excel* e do estatístico *R Project*.

Capítulo 4

Conclusão

Pode-se concluir que este trabalho, desenvolvido através do minicurso, proporcionou um ambiente participativo e de constante diálogo, entre os alunos, entre os alunos e o professor, que possibilitou troca de ideias e uma melhor participação entre os componentes contribuindo para uma melhor aprendizagem.

Conclui-se também que foi possível estabelecer uma prática onde a interdisciplinaridade, a aprendizagem significativa e a metodologia de pesquisa foram usadas com o intuito de aproximar a disciplina estatística, através do ensino de Correlação e Regressão Linear Simples, das disciplinas da área técnica, isto é, procurando verificar a aplicabilidade da Estatística nas disciplinas da área técnica.

Foi possível constatar um melhor resultado na resolução das questões com a utilização de softwares. Não detectou-se nenhum problema na aplicação de tal ferramenta, pois os alunos facilmente se adaptaram. Da mesma forma a escola conta com uma estrutura de laboratório, o que facilitou o trabalho.

A repercussão do projeto executado também despertou interesse entre os alunos do período vespertino que não puderam participar do minicurso, pois este foi realizado no mesmo horário das aulas regulares. Além disto, a direção da escola solicitou que o minicurso pudesse ser reproduzido em outras oportunidades para que outros alunos também pudessem participar. Observa-se assim, que a instituição além descomplicar a implantação de tais trabalhos, incentivou a realização de outros.

A partir do que foi exposto, podemos concluir que este trabalho contribuiu de maneira positiva para uma formação do educando no que se refere às competências e habilidades, além de adquirir mais conhecimento, em particular, nas questões referentes ao cotidiano.

Espera-se que outros trabalhos incentivem o ensino e a utilização dos conteúdos de Estatística, pois esta contribui para experiências que liguem o concreto ao abstrato e tornem a sua aprendizagem numa aventura associada a uma realidade mais vasta.

Referências Bibliográficas

- [1] BASSANEZI, R. C. **Ensino-aprendizagem com modelagem matemática:** uma nova estratégia. São Paulo: Contexto, 2002.
- [2] BAYER, A.; ECHEVEST,S. **Estatística na escola:** importância dos conteúdos de estatística no ensino fundamental e médio. Acta Scientiae v.5 n.1 jan./jun. 2003.
- [3] BRASIL, Lei nº 9.394, de 20 de dezembro de 1996.
- [4] BRASIL, Ministério da Educação. Secretaria de Educação Fundamental. **Parâmetros curriculares nacionais:** Ciências Naturais /Secretaria de Educação Fundamental. Brasília : MEC / SEF, 1998.
- [5] BRASIL, Ministério da Educação. **Parâmetros Curriculares Nacionais Ensino Médio.** Brasília: Secretaria de Educação Média e Tecnológica/MEC, 1999.
- [6] BRASIL, Ministério da Educação. **Programa Nacional de apoio às feiras de ciências da Educação Básica FENACEB/** Secretaria de Educação Básica-Brasília: Ministério da Educação, Secretaria de Educação Básica, 2006.
- [7] BRASIL, Ministério da Educação. Secretaria de Educação Fundamental. **Parâmetros curriculares nacionais :** introdução aos parâmetros curriculares nacionais / Secretaria de Educação Fundamental. Brasília : MEC/SEF, 1997.
- [8] CAMPOS, C. R. **A educação estatística:** uma investigação acerca dos aspectos relevantes à didática da estatística em cursos de graduação. 2007. 242 f. Tese (Doutorado) - Instituto de Geociências e Ciências Exatas, Universidade Estadual Paulista, Rio Claro, 2007.

- [9] CIFUENTES, J. C.; NEGRELLI, L. G. **Uma Interpretação Epistemológica do Processo de Modelagem Matemática:** implicações para a matemática. *Bolema*, Rio Claro (SP), v. 26, n. 43, p. 791-815, ago. 2012
- [10] CHARNET,R., FREIRE,C.A.L., CHARNET,E.M.R., BONVINO,H., **Análise de Modelos de Regressão Linear**, Campinas, SP: Ed. UNICAMP, 2008.
- [11] DEMO, P. **Educar pela pesquisa**. 6. ed. Campinas: Autores Associados, 2003.
- [12] DULLIUS, M. M., *et al.* **O ensino e a aprendizagem da Matemática em ambientes informatizados**. Relatório final de Pesquisa, Lajeado, 2007
- [13] FERNANDES, M.C.V. **A inserção e vivência da mulher na docência de Matemática:** uma questão de gênero. 2006. 108f. Universidade Federal Da Paraíba Centro De Educação Programa De Pós-Graduação Em Educação Mestrado. João Pessoa, 2006.
- [14] FIALHO, J.F; MOURA, M.A. **A formação do pesquisador juvenil**. *Perspect. ciênc. inf.*, Belo Horizonte, v.10 n.2, p. 194-207, jul./dez. 2005.
- [15] FREIBERGER, Regiane Müller; BERBEL, Neusi Aparecida Navas. **A importância da pesquisa como princípio educativo na atuação pedagógica de professores de educação infantil e ensino fundamental**. *Cadernos de Educação*.Pelotas [37]: 207 - 245, setembro/dezembro 2010. Disponível em: <<http://periodicos.ufpel.edu.br/ojs2/index.php/caduc/article/viewFile/1587/1472>> Acesso em 25 Fevereiro de 2014.
- [16] FRIEDMAN, H., HALPERN, N. e SALB, D. **Teaching Statistics using Humorous Anecdotes**. *The Mathematics Teacher*, Vol 92, Abril, 1999.
- [17] GOMES, V. A. A. **A estatística em foco no município de Barão dos Cocais**. 2013. 77f. Dissertação (Mestrado)- Universidade Federal de Viçosa. Minas Gerais, Brasil, 2013.
- [18] GUEDES, P.C.C.Aplicação do software GeoGebra ao ensino da geometria analítica .2013.69f. Dissertação (Mestrado)- Universidade Federal do Espírito Santo.

- [19] GUAJARATI, D. N.; PORTER, D. C. 2011. **Econometria**. 5ª edição. Porto Alegre: McGraw-Hill.
- [20] KINTSCH, E., FRANZKE, M., HALEY, P., KINTSCH W. **Principles of learning in multimedia educational systems**. Boulder: University of Colorado, 1995.
- [21] LIMA, C. M.N.; RODRIGUES, D. T.; OLIVEIRA, L. R. 2013. **Estatística Básica no Excel e Uso da Calculadora Científica**. UFPI.
- [22] LOPES, C.E. **O ensino da estatística e da probabilidade na educação básica e a formação dos professores**. Cad. CEDES vol.28 no.74 Campinas Jan./Apr. 2008. Disponível em: < [http://www.scielo.br/scielo.php?pid= S0101 – 32622008000100005&script=sci_arttext](http://www.scielo.br/scielo.php?pid=S0101-32622008000100005&script=sci_arttext)> Acesso em 20 de junho de 2014.
- [23] MAGALHÃES, M. N.; LIMA, A. C. P. (2010), **Noções de Probabilidade e Estatística**, Editora USP: São Paulo.
- [24] MOURA, Dácio G. & BARBOSA, Eduardo F. **Trabalhando com Projetos Planejamento e Gestão de Projetos Educacionais**, Ed. Vozes, 4ª. Edição, Petrópolis-RJ, 2008.
- [25] MARÇOLA, Alice Ludovice Wandrowelzti. **Estatística e tratamento da informação**. Dezembro de 2006. Disponível em:< http://www.feg.unesp.br/extensao/teia/trab_finais/TrabalhoAlice.pdf > Acesso em: 25 Fevereiro de 2014.
- [26] SANTOS, Maria Lopes da Silva Afonso. 2010. **Estatística Descritiva** - Manual de Auto-aprendizagem. 2ª ed. Lisboa: Edições Silabo, LTDA.
- [27] SEGALIN, Tatiana Daou. **Modelos físicos: uma análise de influências no Processo projetual em design industrial** Universidade Do Estado De Santa Catarina. 2011. Disponível em:<<http://www.pergamum.udesc.br/dados-bu/000000/000000000013/000013FF.pdf>> Acesso em: 24 Fevereiro de 2014.
- [28] TRIOLA, Mario F. 2010. **Introdução à estatística**. 10ª ed. Rio de Janeiro: LTC

Anexos

Anexo I - Exercícios

1. O diretor de vendas de uma rede de varejo nacional precisa averiguar a relação entre o investimento em propaganda e as vendas da empresa. O objetivo é dispor de uma equação matemática que propicia realizar projeções de vendas a partir de investimentos em propaganda. O departamento de vendas preparou a tabela seguinte com as vendas em milhões de reais e os investimentos em propaganda em milhões reais dos dez últimos anos (LIMA *et al.*,)

Propaganda	30	21	35	42	37	20	8	17	35	25
Vendas	430	335	520	490	470	210	195	270	400	480

- a) Represente os dados num diagrama de dispersão
 - b) Indique a variável independente e a variável dependente
 - c) A partir da observação do diagrama de dispersão, classifique a correlação existente entre o propagandas e vendas.
 - d) Determine a equação da reta de regressão e represente a reta no diagrama de dispersão.
 - e) Classifique a correlação linear entre as variáveis apresentadas, com base no valor do coeficiente de Pearson.
 - f) Calcule e interprete o coeficiente de determinação.
 - g) Interprete os parâmetros e estime a projeção de vendas a partir de um investimento de 12,5 milhões em propaganda.
2. A tabela a seguir lista os números de assassinatos e os tamanhos das populações(em centenas de milhares) em grandes cidades americanas durante um anos recente

(com base em dados do New York Times). Com o auxílio do *Microsoft Office Excel* Estime a equação da reta e ache a melhor previsão para o tamanho da população de uma cidade com 120 assassinatos.

Homicídios	258	264	402	253	111	648	288	654	256	60	590
População	4	6	9	6	3	29	15	38	20	6	81

3. Preocupada com os custos que as campanhas eleitorais envolvem, a direção de um certo partido político pretende determinar um modelo que traduza a relação entre custo diário de campanha eleitoral e o número de pessoas envolvidas. Registrados os valores das duas variáveis, durante 12 dias, os resultados obtidos foram os seguintes: $\sum x = 2100$; $\sum y = 1275$; $\sum x^2 = 396800$; $\sum y^2 = 137975$; $\sum xy = 230350$
- Supondo que a relação entre as duas variáveis pode ser traduzida por um modelo linear:
- Identifique a variável independente e a variável dependente ;
 - A partir dos dados fornecidos, determine a equação da reta de mínimos quadrados;
 - Determine e interprete os coeficientes de correlação e determinação
4. Os dados seguintes referem-se ao número de funcionários e ao número de acidentes de trabalho verificados no ultimo ano, numa amostra de 9 clientes da Max-Segur Higiene e Segurança no Trabalho.

Nº de funcionários	510	55	536	73	102	35	29	46	409
Nº acidentes	12	2	6	2	3	1	1	3	8

- Determine o número médio de acidentes do trabalho.
- Determine a equação da reta de regressão linear.
- Estime o número de acidente numa empresa com 200 funcionários.

Anexo II - Tabela t de Student

Distribuição t de Student (Bicaudal e Unicaudal)

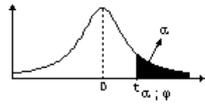


TABELA - Distribuição t de Student (Unicaudal e Bicaudal)

φ = graus de liberdade

φ	α	25%	10%	5%	2,5%	1%	0,5%	φ	α	25%	10%	5%	2,5%	1%	0,5%
1		1,0000	3,0777	6,3138	12,7062	31,8207	63,6574	46		0,6799	1,3002	1,6787	2,0129	2,4102	2,6870
2		0,8165	1,8856	2,9200	4,3027	6,9646	9,9248	47		0,6797	1,2998	1,6779	2,0117	2,4083	2,6846
3		0,7649	1,6377	2,3534	3,1824	4,5407	5,8409	48		0,6796	1,2994	1,6772	2,0106	2,4066	2,6822
4		0,7407	1,5332	2,1318	2,7764	3,7469	4,6041	49		0,6795	1,2991	1,6766	2,0096	2,4049	2,6800
5		0,7267	1,4759	2,0150	2,5706	3,3649	4,0322	50		0,6794	1,2987	1,6759	2,0086	2,4033	2,6778
6		0,7176	1,4398	1,9432	2,4469	3,1427	3,7074	51		0,6793	1,2984	1,6753	2,0076	2,4017	2,6757
7		0,7111	1,4149	1,8946	2,3646	2,9980	3,4995	52		0,6792	1,2980	1,6747	2,0066	2,4002	2,6737
8		0,7064	1,3968	1,8595	2,3060	2,8965	3,3554	53		0,6791	1,2977	1,6741	2,0057	2,3988	2,6718
9		0,7027	1,3830	1,8331	2,2622	2,8214	3,2498	54		0,6791	1,2974	1,6736	2,0049	2,3974	2,6700
10		0,6998	1,3722	1,8125	2,2281	2,7638	3,1693	55		0,6790	1,2971	1,6730	2,0040	2,3961	2,6682
11		0,6974	1,3634	1,7959	2,2010	2,7181	3,1058	56		0,6789	1,2969	1,6725	2,0032	2,3948	2,6665
12		0,6955	1,3562	1,7823	2,1788	2,6810	3,0545	57		0,6788	1,2966	1,6720	2,0025	2,3936	2,6649
13		0,6938	1,3502	1,7709	2,1604	2,6503	3,0123	58		0,6787	1,2963	1,6716	2,0017	2,3924	2,6633
14		0,6924	1,3450	1,7613	2,1448	2,6245	2,9768	59		0,6787	1,2961	1,6711	2,0010	2,3912	2,6618
15		0,6912	1,3406	1,7531	2,1315	2,6025	2,9467	60		0,6786	1,2958	1,6706	2,0003	2,3901	2,6603
16		0,6901	1,3368	1,7459	2,1199	2,5835	2,9208	61		0,6785	1,2956	1,6702	1,9996	2,3890	2,6589
17		0,6892	1,3334	1,7396	2,1098	2,5669	2,8982	62		0,6785	1,2954	1,6698	1,9990	2,3880	2,6575
18		0,6884	1,3304	1,7341	2,1009	2,5524	2,8784	63		0,6784	1,2951	1,6694	1,9983	2,3870	2,6561
19		0,6876	1,3277	1,7291	2,0930	2,5395	2,8609	64		0,6783	1,2949	1,6690	1,9977	2,3860	2,6549
20		0,6870	1,3253	1,7247	2,0860	2,5280	2,8453	65		0,6783	1,2947	1,6686	1,9971	2,3851	2,6536
21		0,6864	1,3232	1,7207	2,0796	2,5177	2,8314	66		0,6782	1,2945	1,6683	1,9966	2,3842	2,6524
22		0,6858	1,3212	1,7171	2,0739	2,5083	2,8188	67		0,6782	1,2943	1,6679	1,9960	2,3833	2,6512
23		0,6853	1,3195	1,7139	2,0687	2,4999	2,8073	68		0,6781	1,2941	1,6676	1,9955	2,3824	2,6501
24		0,6848	1,3178	1,7109	2,0639	2,4922	2,7969	69		0,6781	1,2939	1,6672	1,9949	2,3816	2,6490
25		0,6844	1,3163	1,7081	2,0595	2,4851	2,7874	70		0,6780	1,2938	1,6669	1,9944	2,3808	2,6479
26		0,6840	1,3150	1,7056	2,0555	2,4786	2,7787	71		0,6780	1,2936	1,6666	1,9939	2,3800	2,6469
27		0,6837	1,3137	1,7033	2,0518	2,4727	2,7707	72		0,6779	1,2934	1,6663	1,9935	2,3793	2,6459
28		0,6834	1,3125	1,7011	2,0484	2,4671	2,7633	73		0,6779	1,2933	1,6660	1,9930	2,3785	2,6449
29		0,6830	1,3114	1,6991	2,0452	2,4620	2,7564	74		0,6778	1,2931	1,6657	1,9925	2,3778	2,6439
30		0,6828	1,3104	1,6973	2,0423	2,4573	2,7500	75		0,6778	1,2929	1,6654	1,9921	2,3771	2,6430
31		0,6825	1,3095	1,6955	2,0395	2,4528	2,7440	76		0,6777	1,2928	1,6652	1,9917	2,3764	2,6421
32		0,6822	1,3086	1,6939	2,0369	2,4487	2,7385	77		0,6777	1,2926	1,6649	1,9913	2,3758	2,6412
33		0,6820	1,3077	1,6924	2,0345	2,4448	2,7333	78		0,6776	1,2925	1,6646	1,9908	2,3751	2,6403
34		0,6818	1,3070	1,6909	2,0322	2,4411	2,7284	79		0,6776	1,2924	1,6644	1,9905	2,3745	2,6395
35		0,6816	1,3062	1,6896	2,0301	2,4377	2,7238	80		0,6776	1,2922	1,6641	1,9901	2,3739	2,6387
36		0,6814	1,3055	1,6883	2,0281	2,4345	2,7195	81		0,6775	1,2921	1,6639	1,9897	2,3733	2,6379
37		0,6812	1,3049	1,6871	2,0262	2,4314	2,7154	82		0,6775	1,2920	1,6636	1,9893	2,3727	2,6371
38		0,6810	1,3042	1,6860	2,0244	2,4286	2,7116	83		0,6775	1,2918	1,6634	1,9890	2,3721	2,6364
39		0,6808	1,3036	1,6849	2,0227	2,4258	2,7079	84		0,6774	1,2917	1,6632	1,9886	2,3716	2,6356
40		0,6807	1,3031	1,6839	2,0211	2,4233	2,7045	85		0,6774	1,2916	1,6630	1,9883	2,3710	2,6349
41		0,6805	1,3025	1,6829	2,0195	2,4208	2,7012	86		0,6774	1,2915	1,6628	1,9879	2,3705	2,6342
42		0,6804	1,3020	1,6820	2,0181	2,4185	2,6981	87		0,6773	1,2914	1,6626	1,9876	2,3700	2,6335
43		0,6802	1,3016	1,6811	2,0167	2,4163	2,6951	88		0,6773	1,2912	1,6624	1,9873	2,3695	2,6329
44		0,6801	1,3011	1,6802	2,0154	2,4141	2,6923	89		0,6773	1,2911	1,6622	1,9870	2,3690	2,6322
45		0,6800	1,3006	1,6794	2,0141	2,4121	2,6896	90		0,6772	1,2910	1,6620	1,9867	2,3685	2,6316
								100		0,677	1,290	1,660	1,984	2,364	2,626
								120		0,677	1,289	1,658	1,980	2,358	2,617
								∞		0,674	1,282	1,645	1,960	2,326	2,576

Anexo III

Tabela 4.1: Tabela: Valores Críticos do Coeficiente de Correlação r de Pearson

n	$\alpha = 0,05$	$\alpha = 0,01$
4	0,950	0,999
5	0,878	0,959
6	0,811	0,917
7	0,754	0,875
8	0,707	0,834
9	0,666	0,798
10	0,632	0,765
11	0,602	0,735
12	0,576	0,708
13	0,553	0,684
14	0,532	0,661
15	0,514	0,641
16	0,497	0,623
17	0,482	0,606
18	0,468	0,590
19	0,456	0,575
20	0,444	0,561
25	0,396	0,505
30	0,361	0,463
35	0,335	0,430
40	0,312	0,402
45	0,294	0,378
50	0,279	0,361
60	0,254	0,330
70	0,236	0,305
80	0,220	0,286
90	0,207	0,269
100	0,196	0,256

Anexo IV

QUESTIONÁRIO APLICADO AOS ALUNOS **PERFIL DO ALUNO E AVALIAÇÃO DO CURSO**

1- Sexo:

F () M ()

2- Qual seu curso?

- () Ensino Médio Integrado em Informática
() Ensino Médio Integrado Segurança do Trabalho
() Outro? Qual? _____

4- Após o término do Ensino Médio, você pretende iniciar Ensino Superior?

() Não () Sim

Caso afirmativo, em qual área?

- () Humanas/ Línguas
() Exatas/ Tecnologia
() Saúde
() Outra. Qual? _____

3- Você tem/teve algum vínculo com pesquisa científica?

() Não () Sim. Qual tema? _____

Com base no que foi estudado, você acha que aplicaria os conhecimentos adquiridos em seu projeto?

() sim () não

5- Já conhecia o tema abordado - Regressão Linear?

() Não () Sim

6- O curso realizado:

- () proporcionou-me novos conhecimentos sobre o assunto
() não proporcionou-me novos conhecimentos além dos já possuídos

7- Com que frequência você utiliza o computador para aprendizado de Matemática/ Estatística?

a) sempre b) às vezes c) Teve apenas uma oportunidade d) Nunca utilizou.

8- Você considera que o mini curso lhe auxiliou positivamente na aquisição de conhecimento?

- () não contribuiu () contribuiu razoavelmente
() contribuiu pouco () contribuiu muito

9- Em relação à contribuição do software e o uso do computador, você considera que:

- () não interferiu no seu aprendizado.
() auxiliou um pouco no seu aprendizado
() auxiliou razoavelmente no seu aprendizado.
() auxiliou muito no seu aprendizado.

10- Você considera que a Regressão Linear Simples pode lhe auxiliar em situações cotidianas?

() Sim () Não