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

Análise dos resultados da preparação e da avaliação externa SPAECE utilizando regressão linear simples

Francisco Egilberto Faustino

Teresina - 2016

Francisco Egilberto Faustino

Dissertação de Mestrado:

Análise dos resultados da preparação e da avaliação externa SPAECE utilizando regressão linear simples

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:
Profa. Dra. Valmária Rocha da Silva Ferraz

Teresina - 2016

PROFMAT

SBM

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: **Análise dos resultados da preparação e da avaliação externa SPAECE utilizando modelo de regressão linear simples**, defendida por **Francisco Egilberto Faustino** em **24/06/2016** e aprovada pela banca constituída pelos professores:

Valmaria R. B. Ferraz

Profa. Dra. Valmária Rocha da Silva Ferraz (UFPI) – **Presidente**

Fernando F. Nascimento

Prof. Dr. Fernando Ferraz do Nascimento (UFPI) – **examinador interno**

Francisco Adriano Gomes Bezerra

Prof. Ms. Francisco Adriano Gomes Bezerra (SEDUC - CE) - **examinador externo**

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca Setorial do CCN

F268a Faustino, Francisco Egilberto.
Análise dos resultados da preparação e da avaliação externa SPAECE utilizando regressão linear simples / Francisco Egilberto Faustino. – Teresina, 2016.
68f. il.

Dissertação (Mestrado Profissional) – Universidade Federal do Piauí, Centro de Ciências da Natureza, Pós-Graduação em Matemática, 2016.
Orientadora: Profa. Dra. Walmária Rocha da Silva Ferraz.

1. Estatística – Ensino de Matemática. 2. Modelo de Regressão Simples. I. Título

CDD 001.422-021

Minha Família.

Agradecimentos

Agradeço a Deus por ter me dado a oportunidade de viver e fazer esse mestrado, a meus pais, Francisco de Assis Faustino e Luciana Ferreira Faustino que mesmo com pouco estudo mas muita sabedoria me incentivaram a estudar, e me fizeram acreditar que é através dos estudos que conseguimos ser alguém na vida. Agradeço também a minha esposa Maria Catiania de Souza, por estar comigo dando apoio e tendo paciência nos momentos difíceis durante o período do mestrado, já que este exigiu muita dedicação, e às minhas filhas: Ana Érika Souza Faustino e Ana Káren Souza Faustino que fazem me esforçar cada vez mais para lhes dar uma vida melhor.

Agradeço aos colegas: Kelson, Itaércio, Jelves, José Clades, Dayssi, Fabricio da turma do PROFMAT 2014 UFPI, pelos momentos de estudo, caronas, amizade vividos nestes dois anos de curso, e aos demais colegas por algum momento de estudo e esclarecimentos de dúvidas. A todos os meus professores, em especial aos meus professores da UFPI em nome da minha orientadora Prof. Dra. Valmária Rocha da Silva Ferraz, que com tanto empenho me orientou neste trabalho e que também tenho como exemplo de pessoa, uma mulher forte, guerreira e caráter exemplar, obrigado Valmária.

Agradeço ao colégio de Ensino Médio Wilebaldo Aguiar em nome da atual diretora Ana Georgele Olímpio Frota pela a oportunidade para executar meu trabalho e fornecer dados para realização deste trabalho e de forma muito especial a coordenadora e amiga Eva Maria Silva Costa, que me ajudou de forma significativa para fazer esta dissertação. Agradeço (a CAPES, ao CNPq, a FAPEPI) pelo apoio financeiro.

"A Matemática apresenta invenções tão sutis que poderão servir não só para satisfazer os curiosos como, também para auxiliar as artes e poupar trabalho aos homens".

Descartes.

Resumo

Lecionar Matemática no Ensino Médio, em escola pública principalmente, é um desafio, visto que os alunos ingressam com dificuldades básicas de aprendizagem. Para saber se a escola realizou um bom trabalho, o governo estadual do Ceará aplica uma avaliação denominada SPAECE (Sistema Permanente de Avaliação da Educação Básica do Ceará). A preparação para essa prova é composta por um conjunto de projetos desenvolvidos na E.E.M. Wilebaldo Aguiar em Massapê-CE, tendo como base uma matriz de referência especifica dessa avaliação, através de métodos simples aplicados no dia a dia dos alunos e professores. Neste sentido, esse trabalho analisa os resultados da preparação para essa avaliação externa e os resultados da avaliação em si, utilizando um modelo de regressão linear simples.

Palavras-chave: Ensino de Matemática; Estatística; Modelo de Regressão Simples.

Abstract

Teaching mathematics in high school , in public school primarily is a challenge , as students enter with basic learning difficulties. To find out if the school held a good job , the state government of Ceara applies an evaluation called SPAECE (Permanent System of Evaluation of Basic Education of Ceará) . Preparation for this test consists of a set of projects developed in E.E.M. Wilebaldo Aguiar in Massapê -EC , based on a reference matrix specifies that assessment, through simple methods applied in the daily lives of students and teachers. In this sense, this paper analyzes the preparation of the results for this external evaluation and assessment of the results themselves , using a simple linear regression model.

Keywords: mathematics education; Statistics; Simple Regression Model.

Lista de Figuras

Lista de Tabelas

Sumário

Introdução

Para verificar como está o ensino de Português e Matemática no estado do Ceará o governo realiza uma avaliação externa denominada SPAECE (Sistema Permanente de Avaliação do Ensino Básico do Ceará). Daí, vem o questionamento: Será que é possível ter uma previsão das notas dos alunos numa avaliação externa sabendo suas notas durante a preparação? Este trabalho visa verificar, utilizando ferramentas estisticas, se existe relação entre a proficiencia no SPAECE e as notas da preparação, bem como, quantificar essa relção e principalmente ajustar uma reta de regressão linear simples, utilizando a avaliação da preparação como uma variável explicativa da nota da avaliação externa. Além disso, neste trabalho, também foi feito a previsão, baseado nos parâmetros de regressão estimatimados.

Nos dois primeiros capítulos foram explanados o processo de ensino aprendizagem matemática, e o sistema educacional do estado do Ceará, dando ênfase a componentes pedagógicos. A importância do primeiro capítulo em mencionar a Educação Matemática se dá ao fato de que não se pode falar em processo de ensino aprendizagem matemática sem entender e nem sem compreender alguns pontos históricos como a evolução do ensino de matemática no mundo, que foi abordado de forma sucinta e objetiva, com a intenção de iniciar as discussões sobre como ensinar matemática nas escolas de Ensino Médio atualmente no Brasil. No segundo capítulo, o assunto abordado foi o Ensino Médio no estado do Ceará, com abordagens aos seus projetos, dando maior ênfase ao SPAECE.

No Capítulo 3 é feita uma abordagem sobre a teoria de Regressão Linear Simples e no Capítulo 4 faz-se a descrição da metodologia aplicada na preparação e a aplicação da análise de regressão utilizando os dados obtidos e, assim, verificar se a escola está no caminho certo.

Capítulo 1

Ensino de Matemática

A matemática no mundo é uma disciplina de muita importância para um país, o desenvolvimento dela pode contrubuir para diferenciar um local do outro, pois os conhecimentos matemático estão presentes na tecnologia, na economia, na engenharia e em vários outros ramos do cotidiano das pessoas, Logo é uma disciplina que caracteriza um país. Segundo ([1], p.148):

> [...] A matemática tem sido comparada a um jogo, como o xadrez, no qual, mediante a utilização de regras fixas, parte-se de uma posição inicial para chegar a conclusões bem determinadas. Isto porém é uma visão muito parcial, no máximo, pode ser considerado como uma metáfora para exemplificar o método dedutivo. Na verdade, a matemática é muito mais que esse método. Usa-o como um poderoso auxiliar porém seu âmbito é mais extenso, suas ambições são mais variadas e elevadas, suas vitorias e conquistas bem mais profundas e significativas para a humanidade [...]

Neste capítulo será apresentado um pouco da evolução do Ensino de Matemática no século XX e XXI, seus desafios, dificuldades e vitorias, também será relatado como os alunos brasileiros veem o estudo da disciplina, assim como as características de um bom professor de matemática.

1.1 Um pouco da história do Ensino de Matemática

O ensino de matemática ao longo do século passado passou por muitas modificações, discriminações e grandes evoluções, iremos nos deter a explanar alguns tipos e faces do ensino de matemática que influenciaram direto ou indiretamente no sistema de ensino atualmente adotado pelo Brasil. Conforme ([1],p.156)

> [...] através dos tempos ocorreram mudanças radicais na organização do ensino. Muitas disciplinas desapareceram ou perderam importância. Mas a matemática, com todas as dificuldades ligadas à sua aprendizagem, perdurou como elemento fundamental da estrutura do ensino [...]

No decorrer da história de evolução existem três exemplos recentes que marcaram o ensino de Matemática e vale apena explaná-los, são investidas na tentativas de mudanças do ensino de matemática nas escolas de Ensino Básico que marcaram o processo, a primeira a ser comentada é a Matemática Moderna, em segundo lugar fica a Matemática dos Computadores e a última a Geometria da União Soviética.

A matemática moderna que se iniciou na década de 60, que utilizava padrões da década de 20, tendo como base dos estudos matemáticos a abstração, que segundo ([1],p.150):

> [...] as consequências desse movimento em nosso pais foram desastrosas, em que pese o fato de que algumas das práticas propostas eram realmente aconselháveis. Acontece que, tradicionalmente, desde nossos dias de colônia, estamos acostumados a seguir a moda que nos ditam os países mais desenvolvidos. E, em geral, imitamos o que é fácil, superficial e frívolo. Nossa imitação da matemática moderna resultou em abandono da geometria e dos cálculos numéricos, substituídos por exageros conjuntivistas e um pseudo-formalismo vazio e desligado da realidade [...]

Já na matemática dos computadores, que foi liderada pelo Japão, tinha a ilusão de inserir no processo ensino aprendizagem como ferramenta principal o computador, depois de perceber as dificuldades dos professores em trocar seus métodos tradicionais de ensinar pela utilização direta desta tecnologia, o pais com depois de vários estudos chegou a sensata conclusão, que esta inclusão dos computadores só seria viável nas series mais avançadas do ensino básico, depois que os alunos assimilassem os conhecimentos fundamentais da matemática. Desta forma ([1],p.150) lembra que:

> [...]este exemplo serve como advertência aos dirigentes de nosso pais, os quais, na ânsia de uma modernidade ilusória e em busca de uma publicidade fácil,

> colocam a aquisição de maquinas acima do aperfeiçoamento, da melhoria das condições de trabalho e da remuneração adequada dos professores [...]

Na Geometria da União Soviética, durante aproximadamente 70 anos o estudo era baseado no livro Geometria Elementar, que era baseado em descobertas cientificas, que desenvolveram o conhecimento geométrico daquele país. Com o passar dos anos sentiu-se a necessidade de atualização do currículo incluindo novos conteúdos compatíveis a época, assim como menciona ([1],p.154).

> [...] finalmente, o exemplo soviético nos mostra que nem sempre o caminho nitidamente melhor para o matemático (ou mesmo o cientista) é melhor para o professor e para o aluno. Que a formação continuada dos professores é uma tarefa inadiável e permanente [...]

Analisando estes três exemplos de tentativas de acertos no processo verifica-se que este questionamento do Ensino de Matemática é complexo e estará sempre em constante modificações, sendo de competência de todos que compõem o processo de ensino da matemática nunca parar de pensar em como a matemática deve contribuir para a vida e como ela deve ser ensinada nas escolas.

1.2 A Matemática do Ensino Médio no Brasil

O Ensino Básico no Brasil é dividido em duas etapas: Ensino Fundamental e Ensino Médio, sendo o primeiro com duração de 9 anos e o segundo com 3 anos. O principal objetivo do Ensino Fundamental, é ensinar os conhecimentos fundamentais para sua convivência como um cidadão na sociedade, no Ensino Médio o objetivo é aprofundar os conhecimentos necessários para que os alunos possam ingressar no mercado de trabalho e nas Universidades. E por este motivo o ensino médio ocorre de formas distintas em diferentes países, assim como mostra ([1],p.172):

> [...] Correspondentemente, o ensino médio pode, em linhas gerais, ser organizado de três modos diferentes: a) O mesmo currículo, portanto a mesma escola para todos os alunos. Este é o sistema brasileiro atual. b) A mesma escola, porém com currículos diferenciados, de acordo com as preferencias e os

> objetivos individuais. Este é o sistema adotado nos Estados Unidos. c) Escolas diferentes, separadas conforme os objetivos dos alunos e seus rendimentos nos estudos anteriores. Este é o sistema vigente na Alemanha [...]

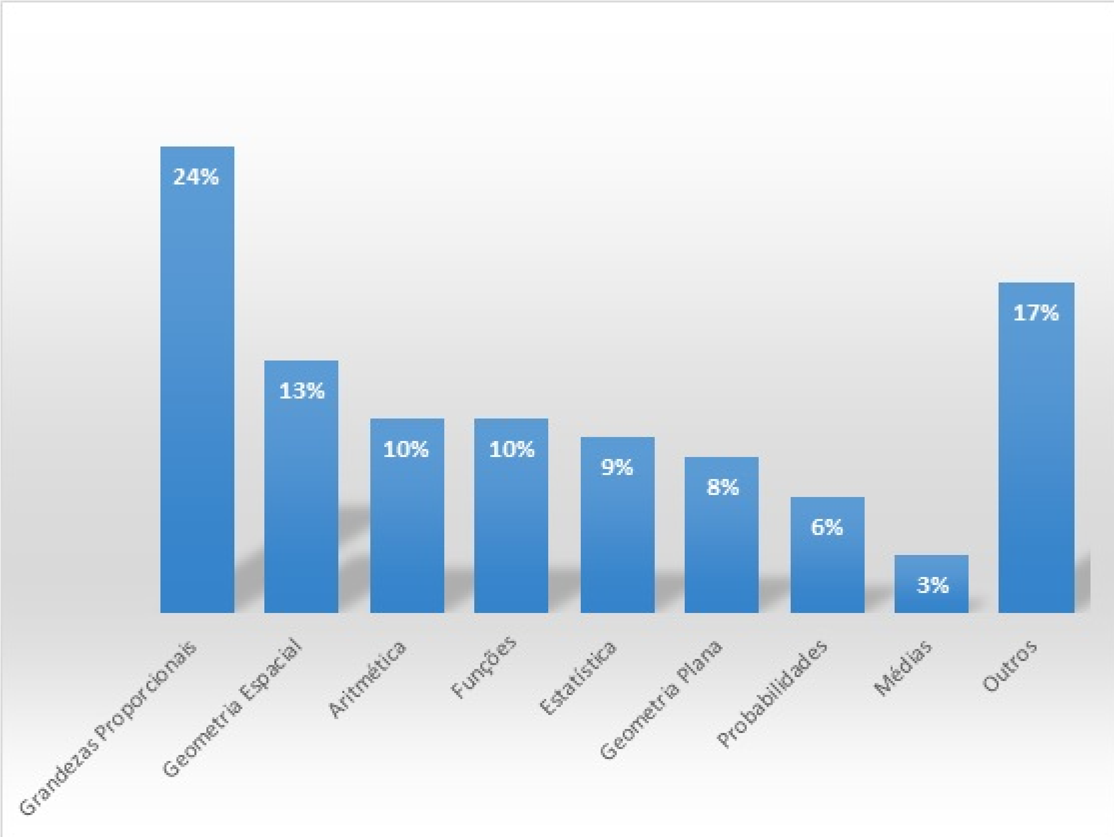
Na atualidade o Brasil está fazendo investimentos no Ensino Profissional com escolas de tempo integral que oferecem aulas do Ensino Médio regular e aulas profissionalizantes em áreas diversas, dependendo da demanda local, onde o aluno ao terminar o Ensino Médio sai também com um diploma de um Curso Técnico.

Em relação ao Ensino de Matemática nas escolas do país no Ensino Médio, pode se verificar que este é baseado nas competências e habilidades da Matriz de Referência do Exame Nacional do Ensino Médio (Enem), que proporciona o ingresso dos estudantes nas maioria das Universidades públicas do país.

Para falar de Ensino Médio atualmente no Brasil é preciso mencionar o Enem, este que possui uma Matriz de Referência baseada em contextualização e interdisciplinaridade, e especificamente na Matemática pode-se verificar por análise dos últimos exames e estudos das habilidades, os principais conteúdos abordados na avaliação que estão descritos no gráfico da figura 1.1.

Além da Mariz de Referência do Enem, existem os Parâmetros Curriculares Nacionais-PCN que são orientações para a construção do Projeto Curricular nas escolas, e outras avaliações externas que possuem uma matriz curricular que muitas das vezes norteia o projeto curricular das escolas, no Ceará é realizado o SPAECE- Sistema Permanente de Avaliação do Ensino Básico do Ceará, que possui grande influência no desenvolvimento do ensino de matemática do estado como veremos nos capítulos seguintes.

Uma outra avaliação de relevância não só para o Brasil assim como para o mundo todo é o PISA- Programa Internacional de Avaliação de Estudantes, que coordenado pela Organização para Cooperação e Desenvolvimento Econômico (OCDE), e é aplicado em vários países com avaliações de matemática, leitura e ciências, para alunos na faixa de 15 anos que na maioria dos países já concluiu o ensino básico e o caso do Brasil o Ensino fundamental, podemos verificar no gráfico da figura 1.2 o desempenho do Brasil ao longo dessas últimas 5 participações.



Fonte: Elaborado pelo autor

Figura 1.1: principais conteúdos de matemática que são abordados no ENEM.

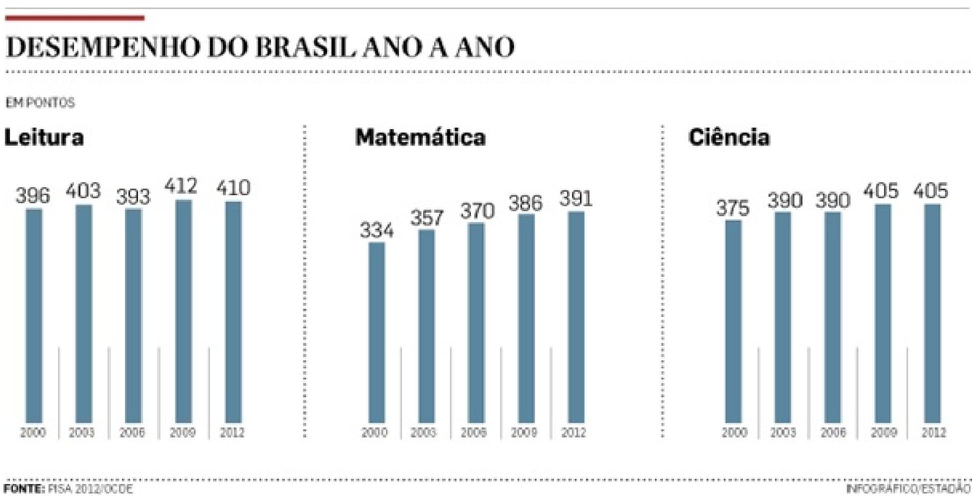


Figura 1.2: Resultados do Brasil no Pisa.

Capítulo 2

O Ensino Médio no estado do Ceará

O estado do Ceará vem melhorando na educação nos últimos anos em decorrência dos investimentos nas unidades escolares, com reformas e estruturação dos equipamentos de tecnologia, e incentivos financeiros aos professores e funcionários e premiação para alunos destaques. Além disso, existe alguns projetos de incentivo e melhoria do Ensino Básico, como mostra a Tabela 2.1.

Este capítulo relata sobre os dois últimos projetos citados na Tabela 2.1, que são o Jovem de Futuro e o SPAECE, pois são esses dois que estão em quase todas as escolas do Ensino Médio cearense, e que são os projetos de maiores investimentos técnicos e financeiros do estado.

O Jovem de Futuro e o SPAECE caminham unidos pois um depende do avanço do outro, o Jovem de Futuro é um projeto que busca a melhoria da qualidade do ensino médio, e esse avanço é baseado nos resultados do SPAECE que é um sistema de avaliação externa que premia os alunos que atingirem a meta do estado. Segundo ([2], p.6):

> [...] o Ceará foi um dos primeiros estados a realizar avaliação externa do sistema educacional. Com efeito, em 1992, num momento em que o MEC divulgava os primeiros resultados do SAEB para o País, as regiões e os estados, também aqui, já realizávamos nossa primeira experiência de avaliação de aprendizagem dos alunos [...]

No Ranking da Educação do Ensino Médio dos estados brasileiros, o Ceará se encontra na 5 posição juntamente com o estado de Rondônia e Mato Grosso do Sul, com uma nota de 3,6, como mostra a Tabela 2.2. Esses dados foram retirados do site do INEP-

Instituto Nacional de Estudos e Pesquisas, e são referentes as notas do IDEB-Índice de Desenvolvimento da Educação Básica.

Percebe-se avanços nos investimentos na educação do Ensino Médio cearense, como os concursos para quase 8.000 professores, nesses 7 últimos anos, aumento de salários, criação de mais de 100 escolas de tempo integral com ensino técnico, mas os resultados de aprendizagem do ensino médio ainda não são significativos. Mas sabe-se que a educação não se pode mudar em 3 anos nem em 20, educação e cultura são costumes e hábitos da comunidade, que podem levar várias gerações para ser mudados, o Brasil como um todo não tem hábito de estudar e de ler livros, os investimentos ainda são muito tímidos e tudo depende de uma boa base na Educação Infantil e Fundamental, além do apoio familiar que está totalmente atrelada ao desenvolvimento sócio econômico do local, desta forma podemos perceber que a educação depende de todos os aspectos da sociedade, como fala ([1],p.2):

> [...] os países ricos, aqueles onde o povo tem uma vida mais confortável, são precisamente aqueles em que as pessoas têm acesso a uma educação de melhor qualidade. Isso significa escolas bem equipadas e professores competentes. Esse quadro resulta da conscientização, arraigada na cultura nacional, de que a educação, além de ser a única porta para o bem-estar, é um direito do cidadão e um dever do Estado [...]

2.1 SPAECE: Uma avaliação externa

SPAECE significa Sistema Permanente de Avaliação da Educação Básica do Ceará, é uma avaliação em larga escala aplicada no Ensino Fundamental e Médio do estado, com questões de Português e Matemática, anualmente em toda rede público, escolas municipais e estaduais, com o objetivo central de acompanhar e incentivar a evolução da aprendizagem dos alunos.

Segundo ([3],p.50) a primeira edição do SPAECE ocorreu em 1992 no Ensino Fundamental e no Ensino Médio só surgiu em 2001, sendo que em 2000 não foi realizado o SPAECE, assim em 2015 foi aplicado a vigésima quarta edição do sistema no Ensino Fundamental e a décima quarta no Médio.

Projetos	Objetivos
PAIC: Programa Aprendizagem na Idade Certa.	Tem a finalidade de alfabetizar os alunos até o final do segundo ano do ensino fundamental.
Peteca	Prevenção da violência no ambiente escolar.
Mais Educação	Busca aumentar a oferta educativa nas escolas públicas por meio de atividades extras.
Premio escola nota 10	É uma premiação pata até 150 escolas públicas que atingirem as metas de aprendizagem do estado.
Diretor de turma	Nas escolas estaduais cada sala de aula possui um professor diretor, que é responsável pelo acompanhamento das atividades escolares e sociais dos alunos.
Geração da paz	Tem como objetivo promover e desenvolver estratégias de aproximação da escola e a comunidade.
Enem chego junto chego bem	Tem como finalidade mobilizar e preparar os estudantes para o Enem.
Jovem de futuro	As escolas recebem capacitações e assessorias técnicas para planejar, executar, acompanhar e avaliar uma proposta de melhoria de seus resultados além da ajuda financeira.
SPAECE	É uma avaliação externa aplicada nas escolas públicas que incentiva e direciona os estudos no Ensino Fundamental e Médio.

Tabela 2.1: Os principais projetos de educação do estado do Ceará.

Posição	Estado	Nota
1º	São Paulo	4,1
2º	Distrito Federal	4,0
2º	Goiás	4,0
2º	Rio de Janeiro	4,0
2º	Santa Catarina	4,0
3º	Rio Grande do Sul	3,9
4º	Pernambuco	3,8
4º	Minas Gerais	3,8
4º	Espírito Santo	3,8
4º	Paraná	3,8
5º	Rondônia	3,6
5º	Ceará	3,6
5º	Mato Grosso do Sul	3,6
6º	Acre	3,4
6º	Roraima	3,4
7º	Tocantins	3,3
7º	Piauí	3,3
8º	Amazônas	3,2
8º	Sergipe	3,2
9º	Rio Grande do Norte	3,1
10º	Amapá	3,0
10º	Maranhão	3,0
10º	Alagoas	3,0
10º	Bahia	3,0
10º	Mato Grosso	3,0
11º	Pará	2,9

Tabela 2.2: Ranking da educação do Ensino Médio no Brasil.

No Ensino Médio a avaliação é aplicada, hoje em dia, geralmente no mês de Novembro, de forma censitário no 1º ano e amostral para os alunos de 2º e 3º anos, a prova é composta de 52 questões de múltipla escolha sendo 26 questões de Português e 26 de Matemática, e um questionário sócio econômico.

Não fugindo a regra das avaliações externas o SPAECE provocou uma mudança no processo de ensino aprendizagem das escolas públicas, pois a maioria começaram a se preocupar em projetar metas e a comparar resultados com outras escolas, e assim crescer com bons exemplos. E para o governo os resultados servem para nortear as Políticas Públicas do estado assim como foi mencionado em ([4], p.07)

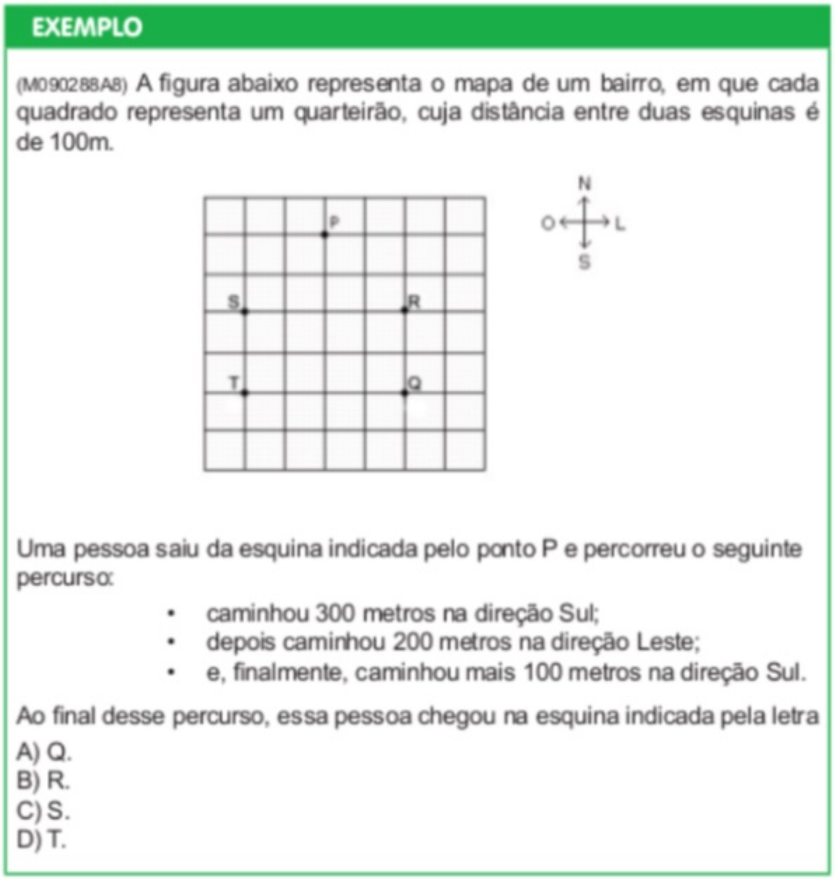
> [...] Essa avaliação produz um diagnóstico do nível de desempenho de cada aluno das três séries avaliadas, bem como da evolução desse desempenho ao longo da trajetória escolar, possibilitando a formulação e reformulação de pol íticas que visam a correção das distorções evidenciadas e por conseguinte a melhoria dos padrões de desempenho do Ensino Médio [...]

2.1.1 Proficiência e Descritores

As questões da avaliação do SPAECE são baseadas em uma Matriz de Referência que por sua vez está dividida em 4 grupos de descritores, onde descritor significa uma forma de descrever as habilidades desejadas para compreender as competências. No grupo de descritores estão comtemplados as habilidades que são enfocados no SAEB- Sistema de Avaliação da Educação Básica e nos PCN- Parâmetros Curriculares Nacionais, é interessante ressaltar que no teste cada questão comtempla apenas um descritor.

Os descritores são divididos em 4 domínios matemáticos que são: Números e Funções, Espaços e Formas (Geometria), Grandezas e medidas, e Tratamento de informação. Assim como mostra o quadro da Figura 2.2.

A avaliação externa SPAECE é avaliada em uma escala que vai de 0 a 500, que é a escala de proficiência de aprendizagem, as questões são classificadas com grau de proficiência em mais ou menos complexas conforme os avanços de compreensão matemática pedido, pode se verificar o seguinte exemplo na Figura 2.1.


EXEMPLO

(M090288A8) A figura abaixo representa o mapa de um bairro, em que cada quadrado representa um quarteirão, cuja distância entre duas esquinas é de 100m.

Uma pessoa saiu da esquina indicada pelo ponto P e percorreu o seguinte percurso:

- caminhou 300 metros na direção Sul;
- depois caminhou 200 metros na direção Leste;
- e, finalmente, caminhou mais 100 metros na direção Sul.

Ao final desse percurso, essa pessoa chegou na esquina indicada pela letra

A) Q.
B) R.
C) S.
D) T.

Figura 2.1: Um modelo de questão do SPAECE

Esta questão é considerada segundo ([4]) como:

> [...] A cor amarelo-escura, 250 a 300 pontos na escala, indica um novo grau de complexidade dessa competência. Nesse intervalo, os alunos mostram-se capazes de associar uma trajetória representada em um mapa à sua descrição textual. Por exemplo: dada uma trajetória entre duas localidades no mapa, o aluno verifica qual a descrição textual representa esse deslocamento, e vice-versa [...]

Já o nível mais avançado desta habilidade matemático seria segundo ([4]),

> [...] No intervalo acima de 375 pontos, representado pela cor vermelha, os alunos localizam figuras geométricas por meio das coordenadas cartesianas de seus vértices, utilizando a nomenclatura abscissa e ordenada [...]

Para ficar mais claro sobre os níveis de proficiência criou-se uma escala onde são mostrados os diferentes graus de dificuldades, quanto mais escura for a tonalidade maior é o grau de dificuldade de resolução da questão, como foi mostrado o exemplo acima, assim pode se entender que na avaliação as questões possuem pontuações diferencias conforme o gráfico de proficiência em anexo.

MATRIZ DE REFERÊNCIA PARA AVALIAÇÃO EM MATEMÁTICA - 1ª SÉRIE DO ENSINO MÉDIO SISTEMA PERMANENTE DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA DO CEARÁ - SPAECE	
TEMA I: INTERAGINDO COM OS NÚMEROS E FUNÇÕES	
Nº DESCRITOR	DESCRITOR
D11	Ordenar ou identificar a localização de números racionais na reta numérica.
D16	Estabelecer relações entre representações fracionárias e decimais dos números racionais.
D17	Resolver situação problema utilizando porcentagem.
D18	Resolver situação problema envolvendo a variação proporcional entre grandezas direta ou inversamente proporcionais.
D19	Resolver problema envolvendo juros simples.
D22	Identificar a localização de números reais na reta numérica.
D23	Resolver situação problema com números reais envolvendo suas operações.
D28	Reconhecer a representação algébrica ou gráfica da função polinomial de 1º grau.
D29	Resolver situação problema envolvendo função polinomial do 1º grau.
D30	Reconhecer a representação algébrica ou gráfica da função polinomial de 2º grau.
D31	Resolver situação problema envolvendo função quadrática.
D32	Resolver situação problema que envolva os pontos de máximo ou de mínimo no gráfico de uma função polinomial do 2º grau.
D33	Reconhecer a representação algébrica ou gráfica da função exponencial.
D34	Resolver situação problema envolvendo função exponencial.
D35	Reconhecer a representação algébrica ou gráfica da função logarítmica.
D37	Resolver situação problema envolvendo inequações do 1º ou 2º graus.
D39	Resolver situação problema envolvendo propriedades de uma progressão aritmética ou geométrica (termo geral ou soma).
D44	Analisar crescimento/decrescimento e/ou zeros de funções reais apresentadas em gráficos.
TEMA II: CONVIVENDO COM A GEOMETRIA	
Nº DESCRITOR	DESCRITOR
D49	Resolver problemas envolvendo semelhança de figuras planas.
D53	Resolver situação problema envolvendo as razões trigonométricas no triângulo retângulo (seno, cosseno, tangente).
D57	Identificar a localização de pontos no plano cartesiano.
TEMA III: VIVENCIANDO AS MEDIDAS	
Nº DESCRITOR	DESCRITOR
D65	Calcular o perímetro de figuras planas numa situação-problema.
D67	Resolver problema envolvendo o cálculo de área de figuras planas.
TEMA IV: TRATAMENTO DA INFORMAÇÃO	
Nº DESCRITOR	DESCRITOR
D75	Resolver problema envolvendo informações apresentadas em tabelas ou gráficos.
D76	Associar informações apresentadas em listas e/ ou tabelas aos gráficos que as representam, e vice-versa.

Figura 2.2: Descritores da Matriz de Referência do SPAECE para o 1º ano do Ensino Médio

2.1.2 Uma ferramenta de aprendizagem

O foco principal do SPAECE e dimensionar a proficiência de aprendizagem dos alunos, ou seja o nível de aprendizagem dos estudantes em Português e Matemática (Figura 2.3). Desde 2002 o governo incentiva os estudantes com uma premiação equivalente a um computador, para os melhores de cada crede, hoje em dia a premiação é um notebook para todos os alunos que atingirem o nível satisfatório nas duas disciplinas.

A cada ano que se realiza a avaliação os resultados são enviados para as escolas, com o objetivo de que gestores e professores possam verificar o desempenho de suas escolas e discutirem o nível de aprendizagem de cada aluno, e assim poder traçar metas, articular projetos e direcionar da melhor forma o processo de ensino e aprendizagem da instituição como disse ([3])

> [...]As avaliações externas fazem parte de uma realidade bastante comum dentro das escolas brasileiras, porém as discussões sobre os resultados obtidos precisam ainda ser objeto de reflexão e direcionamento de ações. No que se refere ao ensino de Matemática, as informações divulgadas sobre os resultados encontrados indicam que, de um modo geral, o desempenho dos estudantes está abaixo do esperado [...]

Muito Crítico	Crítico	Intermediário	Adequado
<250	250 – 300	300 - 350	≥350

Figura 2.3: Classificação dos níveis de aprendizagem do SPAECE.

Sendo assim é necessário uma análise mais criteriosa destes dados da matemática para que os avanços no ensino sejam significativos conforme ([2],p.6)

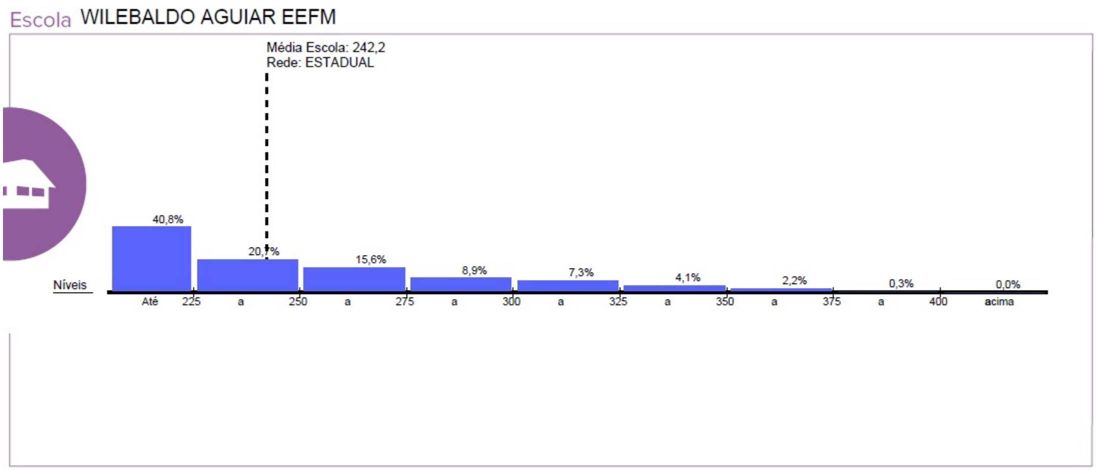
> [...] a avaliação educacional com foco nos resultados de aprendizagem dos alunos é um recurso imprescindível para identificarmos e corrigirmos distorções,

> tendo papel central como estratégia de gestão num sistema que tem por objetivo a oferta universal de ensino de qualidade [...]

Na Escola Wilebaldo Aguiar do município de Massapê-CE, que é a escola onde foi aplicada a pesquisa relatado neste trabalho, nestes últimos 5 anos os resultados do SPAECE são estudados minuciosamente, e todos os projetos de aprendizagem são baseados nestes dados produzidos pelo SPAECE.

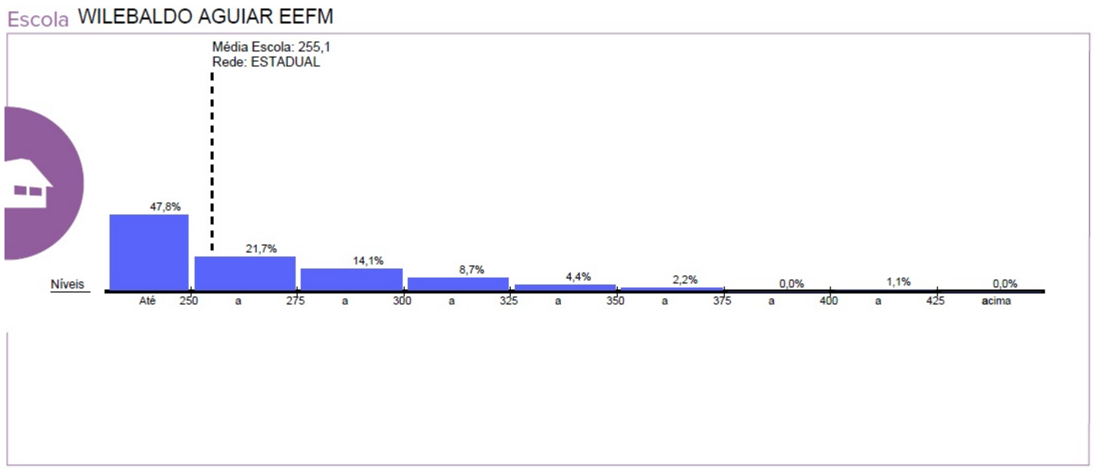
A Escola Wilebaldo Aguiar nos anos de 2011, 2012, 2013 foi uma das escolas regulares da Crede 06 que mais ganhou premiações do SPAECE, motivando assim toda a comunidade escolar a melhorar a cada ano os projetos e objetivos de aprendizagem dos alunos.

Desta forma a escola acredita que os descritores são a base necessária que o aluno precisa aprender para terem êxito nas competências e habilidades do Enem, assim a escola trabalha com foco, no 1º ano no SPAECE e nos 2º e 3º no ENEM que é o sistema que dará acesso as universidades dos sonhos de muitos jovens.



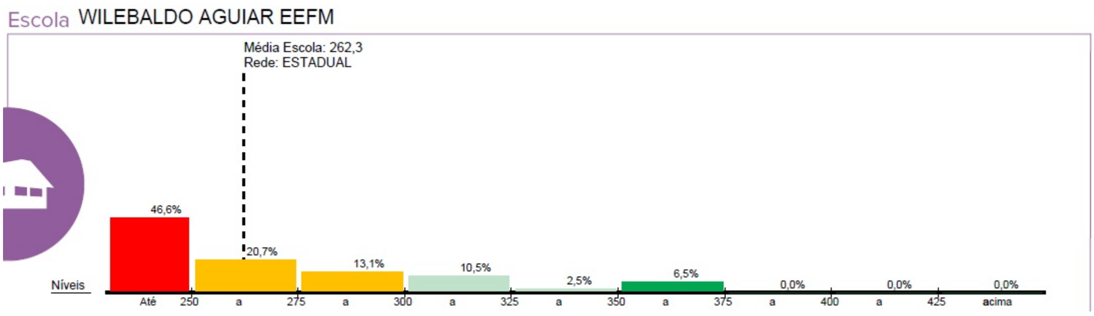
Fonte: http://resultados.caedufjf.net/resultados/publicacao/publico/escola.jsf

Figura 2.4: Proficiência da Escola Wilebaldo Aguiar no SPAECE 2014 1º ano.



Fonte: http://resultados.caedufjf.net/resultados/publicacao/publico/escola.jsf

Figura 2.5: Proficiência da Escola Wilebaldo Aguiar no SPAECE 2014 2º ano.



Fonte: http://resultados.caedufjf.net/resultados/publicacao/publico/escola.jsf

Figura 2.6: Proficiência da Escola Wilebaldo Aguiar no SPAECE 2014 1º ano.

2.2 Projeto Jovem de Futuro

O projeto Jovem de Futuro é uma parceria dos Governos Estaduais, com o Federal e o Instituto Unibanco que surgiu no ano de 2008 nas escolas estaduais do Rio Grade do Sul, Minas Gerais, São Paulo e Rio de Janeiro, e tem como objetivo melhorar a aprendizagem dos alunos nas disciplinas base, Português e Matemática, através de uma poio técnico e financeiro.

Com o sucesso obtido inicialmente dos estados envolvidos na primeira etapa, o projeto foi aderido pelo Ceará e outros estados, no ano de 2011 contemplando 100 escola e no ano de 2013 a Escola Wilebaldo Aguiar iniciou seus trabalhos com o Jovem de Futuro.

O Jovem de Futuro tem o foco o futuro dos jovens do Ensino Médio, com o desenvolvimento dos conteúdos de Português e Matemática, e também com a finalidade de nortear os jovens na sociedade com as Metodologias Pedagógicas de Mobilização e Articulação listadas na Figura 2.7. O projeto tem duração de 3 anos mais um ano de base, com a intenção de oferecer o suporte neste período e depois deixar a escola autônoma e capacitadas para desenvolver seus projetos.

Figura 2.7: Quadro das metodologias do Jovem de Futuro.

As principais metas são baseadas na aprendizagem e no combate ao índice de abandono da escola. Com uma meta ousada de reduzir 40 por cento dos alunos que estão no nível mais baixo de proficiência, o Muito Crítico, diagnosticado no SPAECE e reduzir em 50 por cento o abandono na escola.

A contra partida do Estado no projeto é incentivar o trabalho dos monitores e tutores na metodologia Entre Jovens, que tem como finalidade, como o nome já revela, jovens aprenderem com outros jovens. Os monitores são alunos da própria escola e desenvolvem ações de resgate do conhecimento do Ensino Fundamental juntamente com os tutores que são alunos universitários de preferência dos cursos de licenciatura em letras e/ou matemática, conforme a lei 15.190 de 19 de julho de 2012, eles desenvolvem os projetos em uma carga horário de 8h semanais e recebem uma bolsa de 100,00 ao mês para os monitores e 300,00 para os tutores.

Capítulo 3

Estatística Descritiva para Variáveis Discretas

As medidas de estatística descritiva para variáveis discretas que foram utilizadas na interpretação dos dados na aplicação foram as medidas de tedência central: Média aritmética, mediana e moda que são medidas que tendem está no centro e representam um conjunto de dados, e as medidas de dispersão: Variância e desvio padrão que são medidas para verificar a dispersão dos dados em relação a tendência de média.

Todos os gráficos e cálculos de estatística descritiva, de correlação e de regressão linear simples foram feitos usando o software excel cujas fórmulas estão no Anexo C, exceto os gráficos das Figuras 5.6 e Figura 5.7 que foram feitas usando o software R.

3.1 Medidas de Tendência Central

- **Média Aritmética**: A média aritmética que geralmente é representada por $\overline{x}$ de uma amostra de n observações $x_1, x_2, ..., x_n$ segundo [10] é dada por:

$$\overline{x} = \frac{\sum\limits_{i=1}^{n} x_i}{n}$$

.

- **Mediana**: Colocando os dados em ordem crescente, a mediana é o termo central se a quantidade de dados for par e será a média aritmética dos dois termos centrais se a quantidade for ímpar.

- **Moda**: É o elemento do conjunto de dados com maior frequência.

3.2 Medidas de Dispersão

- **Variância**: Mede a dispersão dos dados x_i $i = 1, 2, ..., n$ em relação a média, se os desvios $d_i = (x_i - \overline{x})$ forem baixos, teremos pouca dispersão, caso contrário teremos elevada dispersão. Geralmente a variância é representada por σ^2 e segundo [10] é calculada por:

$$\sigma^2 = \frac{\sum\limits_{i=1}^{n}(x_i - \overline{x})^2}{n-1}$$

.

- **Desvio Padrão**: Ele mostra o quanto de variação existe em relação a média, se o desvio padrão for baixo os dados tendem está próximo da média, caso contrário os dados estão espalhados por valores próximos e distantes da média. O desvio padrão é dado segundo [10] por:

$$\sigma = \sqrt{\sigma^2}$$

Capítulo 4

Correlação e Regressão Linear Simples

Procurar verificar se existe uma relação (ou associação) entre duas ou mais variáveis é muito comum, e esta relação é objeto de estudos da correlação e da regressão. Por exemplo, o rendimento escolar dos alunos pode estar relacionado com o número de horas de estudo por dia de cada aluno; o número de horas de treinamento de um funcionário pode estar relacionado com o número de acidentes envolvendo esse funcionário; o peso de uma pessoa pode estar relacionado com a sua idade; a renda das pessoas pode estar relacionada com o número de anos de estudo de cada pessoa; o consumo das famílias pode estar relacionado com sua renda; as vendas de uma empresa podem estar relacionadas com os gastos com propagandas promocionais; etc.

Um mapa de dispersão pode ser usado para determinar se existe dependência ou uma correlação (estudaremos a linear) entre duas variáveis. Em um mapa de dispersão, os pares ordenados (x, y) representam pontos em um plano cartesiano. A variável independente (quem explica) X é medida sobre o eixo horizontal, e a variável dependente (quem é explicada) Y sobre o eixo vertical.

Segundo [5]. para medir o **grau de dependência** entre duas variáveis aleatórias (uma variável aleatória é uma variável quantitativa, cujo resultado (valor) depende de fatores aleatórios) o pesquisador usa a medida chamada de **covariância**. Ela é calculada pelo valor médio do produto dos desvios da variável aleatória X e da variável aleatória Y em relação às suas respectivas médias:

$$\mathsf{Cov}(X, Y) = E[(X - \mu_x)(Y - \mu_y)] = E(XY) - E(X)E(Y)$$

onde $E(XY) = \sum_i \sum_j x_i y_j p(x_i, y_j)$ se x e y forem discretos e $\int_i \int_j x_i y_j f(x_i, y_j) dx dy$ se x e y forem contínuos.

Se Cov(X, Y) = 0, dizemos que as variáveis aleatórias X e Y são não correlacionadas.

Exemplo: Considere a distribuição conjunta de duas variáveis aleatórias:

X/Y	0	1	2	p(x)
1	3/20	3/20	2/20	8/20
2	1/20	1/20	2/20	4/20
3	4/20	1/20	3/20	8/20
p(y)	8/20	4/20	7/20	1

Tabela 4.1: distribuição conjunta de duas variáveis aleatórias

Neste caso, E(X) = 2 e E(Y) = 0,95 e E(XY) = 1,9. Logo, Cov(X, Y) = 2.0,95-1,9 = 0. Ou seja, as variáveis são não-correlacionadas. Sempre que duas variáveis aleatórias forem independentes sua covariância será igual a zero. Mas a recíproca não é verdadeira. Note que no exemplo acima as variáveis não são independentes, mas sua covariância é nula, pois duas variáveis são independentes quando $p(x, y) = p(x).p(y)$ e se tomarmos $x = 1$ e $y = 2$ no exemplo acima temos $p(1, 2) = \frac{2}{20}$ e $p(1).p(2) = \frac{8}{20}.\frac{7}{20} = \frac{7}{50}$, logo $p(1, 2) \neq p(1).p(2)$, portanto as variáveis não são independentes.

Verificar que a covariância depende da unidade de medida das variáveis X e Y é importante. Além disso, para verificar se existe uma relação (ou associação) entre duas ou mais variáveis que é algo extremamente usual sem levar em consideração essas unidades de medida das variáveis envolvidas, usamos o coeficiente de correlação.

Segundo [5] e [9] o coeficiente de correlação entre X e Y é definido por: $\rho(x, y) = \frac{\mathsf{Cov}(X, Y)}{\sigma_x \sigma_y}$.

A verificação da **existência** e do **grau de relação** entre variáveis é objeto do estudo da correlação. Uma vez caracterizada, procura-se descrever uma relação sob forma matemática, através de uma função (é o que chamamos de modelo matemático de previsão).

A **estimação dos parâmetros desse modelo matemático** é o objeto de estudo da regressão.

4.1 Correlação Linear Simples

O estudo da correlação tem o objetivo de medir e avaliar o grau de relação existente entre duas variáveis aleatórias. O diagrama de dispersão nos indica se existe uma tendência de relacionamento entre duas variáveis em uma amostra bivariada, quando uma amostra consiste de mais de uma variável.

A correlação linear simples objetiva medir a relação entre essas variáveis através da disposição dos pontos (X,Y) da amostra (diagrama de dispersão) em torno de uma reta. Segundo [5] e [9] o instrumento de medida da correlação linear para uma amostra é dado pelo coeficiente de correlação de Pearson (variando de -1 a 1):

$$r_{x,y} = \frac{\Sigma x_i y_i - \frac{1}{n}(\Sigma x_i)(\Sigma y_i)}{\sqrt{(\Sigma x_i^2 - \frac{1}{n}(\Sigma x_i)^2)(\Sigma y_i^2 - \frac{1}{n}(\Sigma y_i)^2)}}$$

Observação: O valor $r_{x,y}$, é calculado para uma amostra. Se tivermos acesso a todos os pontos de uma população teremos $\rho_{x,y}$ que é o coeficiente de correlação populacional.

4.1.1 Tipos de correlação linear simples

1. Correlação Linear Positiva: $0 < r_{x,y} < 1$. Nesse caso, valores crescentes (ou decrescentes) de X estão associados a valores crescentes (ou decrescentes) de Y.

2. Correlação Linear Perfeita Positiva: $r_{x,y} = 1$. A correlação é diretamente proporcional e os pontos (X,Y) estão perfeitamente alinhados em uma reta.

3. Correlação Negativa: $-1 < r_{x,y} < 0$. Neste caso, valores crescentes (ou decrescentes) da variável X estão associados a valores decrescentes (ou crescentes) da variável Y.

4. Correlação Perfeita Negativa: $r_{x,y} = -1$. A correlação é inversamente proporcional e os pontos (X,Y) estão perfeitamente alinhados em uma reta.

5. Ausência de Correlação: Quando não houver relação entre as variáveis X e Y, ou seja, quando as variações de X e Y ocorrem independentemente, então não existe

correlação entre elas, então $r_{x,y} = 0$.

As relações são descritas como **tendências** nos diagramas de dispersão e não como **causa e efeito**: a correlação por si mesma não pode provar que existe causa e efeito, ou seja, ela não pode provar que a alteração no valor de uma variável causou uma alteração na outra variável. Uma correlação forte pode ter sido produzida pelo acaso ou por outras variáveis (chamada de correlação espúria).

Considere o seguinte exemplo: Um estudo foi realizado para avaliar os efeitos da temperatura ambiente no consumo de energia elétrica de uma indústria química. Outros fatores foram mantidos constantes e os dados foram coletados de uma fábrica experimental piloto. Verifique qual o valor do coeficiente de correlação para essas variáveis.

Consumo de energia (BTU):	250	285	320	295	265	298	267	321
Temperatura (ºF):	27	45	72	58	31	60	34	74

Tabela 4.2: Consumo de energia em função da temperatura

Solução: r = 0,9896. Logo, valores **crescentes** da temperatura ambiente implicam **fortemente** em valores **crescentes** no consumo de energia, ou vice-versa.

4.1.2 Teste da Existência de Correlação

Em uma amostra bivariada, é possível determinar se as duas variáveis aleatórias são realmente relacionadas ou não, com nível de significância α. Se as variáveis aleatórias X e Y forem normalmente distribuídas, dispomos de um teste bastante simples pra verificar se elas estão associadas. Por [5] o teste será bilateral e a hipótese nula será rejeitada se $|t_c| > t_{\frac{\alpha}{2},n-2}$. Daí pode-se afirmar que a correlação existe e é significativa com risco α. As hipóteses serão:

$$\begin{cases} H_0 : \rho_{x,y} = 0 \\ H_1 : \rho_{x,y} \neq 0 \end{cases} \quad \text{e a estatística do teste é dada por: } t = \frac{r_{x,y}\sqrt{n-2}}{\sqrt{1-r_{x,y}^2}} \backsim t_{(n-2)}$$

Ex: No caso da indústria química, $t_{calc} = 16,847$ e para 5% de significância pela Tabela da Distribuição t de Student que está no Anexo A temos $t_{tab} = 2,447$ com 6 graus de liberdade. Logo, rejeitamos a hipótese de que não há correlação entre a temperatura ambiente e o consumo de energia, com risco de 5% e podemos afirmar que a teperatura influência no consumo.

4.2 Regressão Linear Simples

Muitos fenômenos podem ser explicados por meio de modelos matemáticos. Inicialmente, um diagrama de dispersão indica que tipo de relação existe entre duas variáveis. Essa relação pode ser dada como uma reta, como uma parábola, como uma função polinomial, como uma função exponencial, etc. Nesse caso, precisamos definir qual é a variável que explica (X) o comportamento da outra (Y). A variável X é também chamada de variável independente e a variável Y é chamada de dependente.

4.2.1 Como se ajusta uma reta de Regressão Linear Simples

Se o fenômeno em estudo é <u>bem descrito por uma reta</u> é preciso determinar então a melhor reta que representa essa relação. No entanto, não podemos esquecer que essa reta é apenas uma aproximação para o verdadeiro fenômeno. A equação que descreve como a variável aleatória Y está relacionada com X e com um erro é chamada de <u>modelo de regressão</u>. O modelo de regressão usado em uma regressão linear simples é apresentado a seguir:

$$Y = \alpha + \beta X + \varepsilon$$

.

No modelo de regressão linear simples, Y é uma função linear de X mais um erro ε. Os valores α e β são denominados parâmetros do modelo, ou seja, são os parâmetros desconhecidos de inclinação e intercepto, respectivamente, e ε é uma variável aleatória definida como o termo erro (ou resíduo). O termo erro mede a variabilidade em Y que não pode ser explicada pela relação linear entre X e Y. Uma das suposições para o modelo de regressão linear simples é que o valor esperado de ε é zero. A presença deste erro aleatório evita que o modelo se torne simplesmente uma equação determinística. Se o modelo for bem escolhido, os erros positivos e negativos em torno da reta serão razoáveis.

Geralmente, os valores dos parâmetros não são conhecidos na prática e devem ser estimados usando os dados da amostra, então a equação de regressão estimada (também chamada de reta ajustada) será:

$$\hat{Y} = a + bX, \text{ ou seja, } e = Y - \hat{Y}$$

sendo os valores de a e b as estimativas de α e β; O valor de a é a intersecção de Y, ou seja, o ponto onde a reta ajustada corta o eixo da variável Y, b é a inclinação, $\hat{Y}$ é o valor estimado de Y para um dado valor de X e e é a estimativa do erro ε.

Um bom critério para determinar essa reta é minimizando o quadrado desses erros entre o valor real de Y e o valor da sua estimativa, usando o método dos mínimos quadrados, ou seja, determinamos os valores a e b que minimizam a função $\sum\limits_{i=1}^{n}(Y - \hat{Y})^2$.

Par tanto basta derivá-la com relação aos parâmetros e igualar essas derivadas a zero. Os valores de a e b serão a solução do sistema. Utilizando este procedimento podemos deduzir que as estimativas de mínimos quadrados da intersecção e da inclinação no modelo de regressão linear simples são:

$$a = \frac{1}{n}\left(\sum_{i=1}^{n} Y_i - b\sum_{i=1}^{n} X_i\right) = \overline{y} - b\overline{x}, \text{ e, } b = \frac{\Sigma x_i y_i - \frac{1}{n}(\Sigma x_i)(\Sigma Y_i)}{\Sigma x_i^2 - \frac{1}{n}(\Sigma x_i)^2} = \frac{\sum\limits_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum\limits_{i=1}^{n}(x_i - \overline{x})^2}$$

O valor de b equivale ao aumento médio (ou diminuição média) de Y que ocorrerá a cada aumento de uma unidade em X. O valor de a pode ser interpretado como um valor inicial de Y ou como o valor da variação de Y que não depende de X.

Exemplo: a) Para dos dados do exemplo acima, determine os estimadores da reta de regressão ajustada pelo método dos mínimos quadrados que explica consumo de energia baseada na temperatura ambiente.

Solução: $a = 218,255$ e $b = 1,384$. Logo, para cada aumento de um grau na temperatura ambiente, espera-se um aumento de 1,384 BTU?s no consumo de energia. A interseção 218,255 de Y pode ser vista como a estimativa de consumo quando a temperatura ambiente for de 0º F ou como representativa da parcela do consumo de energia que varia em função de outros fatores que não seja a temperatura ambiente.

b) (Previsões) Qual é a previsão de consumo de energia quando a temperatura ambiente for de 48ºF?

Solução: Se $a = 218,255$ e $b = 1,384$ então o modelo de regressão será: $\hat{Y}_i = 218,255 + 1,384x_i$.

Logo, para $x = 48$ teremos $\hat{Y}_i = 284,687$. Ou seja, para uma temperatura ambiente igual a 48ºF teremos em média um consumo de 284,687BTUs de energia elétrica.

c)Qual a temperatura para um consumo de 300BTUs?

Solução: Ao substituirmos $y = 300$ no modelo de regressão $\hat{Y}_i = 218,255 + 1,384x_i$, temos $x = 59,06$, ou seja, para consumir em média 300BTUs de energia elétrica a teperatura ambiente deve ser 59,06°F.

OBSERVAÇÕES IMPORTANTES:

1. O método dos mínimos quadrados busca traçar a melhor reta através dos pontos amostrais, ou seja, aquela que torna mínima distância entre os pontos amostrais e a reta de regressão;

2. Sempre é possível obter a equação de uma reta que passa por um conjunto de pontos, mas isto não significa que o modelo seja adequado. Para verificar a adequação do modelo, emprega-se a metodologia da Análise de Variância (ANOVA) e também é recomendável fazer uma análise de resíduos;

3. Ao utilizar um modelo de regressão para fins de previsão, é importante que você considere somente o intervalo relevante (engloba todos os valores de X utilizados no desenvolvimento do modelo de regressão, desde o menor até o maior observado) da variável independente. Você não deve extrapolar além do intervalo dos valores referentes a X.

Algumas suposições se fazem necessárias:

a) A relação entre X e Y é linear (os acréscimos em X produzem acréscimos ou decréscimos proporcionais em Y e a razão de crescimento é constante).

b) Os valores de X são fixados arbitrariamente (X **não** é uma variável aleatória).

c) Y é uma variável aleatória que depende entre outras coisas (erros) dos valores de X.

d) ε_i é o erro aleatório, portanto é uma variável aleatória com distribuição normal, com média zero e variância comum σ^2, ou seja, $\varepsilon_i \sim N(0, \sigma^2)$, $\forall i = 1, 2, \ldots, n$. Essa suposição é importante para que possamos realizar testes de hipóteses e estimação do intervalo de previsão para Y. Cada ε_i representa a variação de Y que não é explicada pela variável independente X.

e) Os erros são considerados independentes.

4.2.2 A Tabela de Análise de Variância

Obtida a reta de regressão estimada, é necessário determinar na sua precisão, isto é, verificar se ela é útil na representação da tendência dos dados observados, e um método chamado análise de variância pode ser usado para testar a significância da regressão. Segundo [5] o procedimento divide a variância total na variável de resposta em componentes significantes como base para o teste.

Note que: $y_i - \hat{y} = (y_i - \overline{y}) - (\hat{y} - \overline{y}) \Rightarrow y_i - \overline{y} = (y_i - \hat{y}_i) + (\hat{y}_i - \overline{y})$.

Se elevarmos ambos os membros ao quadrado e os somarmos $\forall i$, teremos:

$$\sum_{i=1}^{n}(y_i - \overline{y})^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2.$$

Analisando a expressão acima, vemos que a variação total das observações em torno de sua média, dada por $\sum_{i=1}^{n}(y_i - \overline{y})^2$, se decompõe em duas parcelas: soma dos quadrados dos resíduos $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$, que mede a variação em torno da reta de reta de regressão (variação "não-explicada"), e soma dos quadrados dos desvios dos valores estimados (preditos) em relação à sua média (variação "explicada" pela regressão de Y em X) medida por $\sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2$. Resumindo, temos:

$$SQTotal = SQR + SQReg,$$

em que:

$$SQTotal = \left(\sum_{i=1}^{n} y_i^2\right) - n\overline{y}^2 \qquad\qquad SQReg = b^2\sum_{i=1}^{n}(x_i - \overline{x})^2$$

$$SQR = \Sigma y_i^2 - a\,(\Sigma y_i) - b\,(\Sigma x_i y_i) \text{ (ou por subtração: SQTotal - SQReg)}$$

Ou seja, **variação total = variação não explicada pelo modelo + variação explicada pelo modelo**

Essa decomposição associada aos graus de liberdade de cada termo da variação e a uma estatística de teste (F) é sintetizada em uma tabela, chamada Tabela de Análise de Variância (ANOVA).

Se o modelo proposto é correto, $QMR = SQR/gl$ estima σ^2 (variância dos erros). Por isso, é muitas vezes representado por s^2 ou por S_e^2 (sua raiz é chamada de **Erro Padrão da Estimativa**). Contudo, se o modelo proposto não é correto, S_e^2 superistima σ^2. Ela medirá não só a variação aleatória de Y em torno de sua previsão, mas também o mau ajustamento dos dados ao modelo escolhido (a reta), o que é denominado de falta de ajuste (aderência).

Causa de Variação (CV)	g.l	SQ	QM	F
Regressão	1	SQReg	$QMReg = \dfrac{SQReg}{g.l}$	$F = \dfrac{QMReg}{QMR}$
Resíduos	n-2	SQR	$QMR = \dfrac{SQR}{g.l}$	
Total	n-1	SQTotal		

Tabela 4.3: ANOVA

Com esta tabela de Análise de Variância podemos testar a seguinte hipótese $\begin{cases} H_0 : \beta = 0 \\ H_1 : \beta \neq 0. \end{cases}$

Ou seja, podemos testar se a regressão é significativa, pois se a inclinação for nula, então os valores de X não interferem nos valores de Y. Para isso fazemos a comparação dos valores do QMR e QMReg. Se eles forem muito afastados ou, equivalentemente, se $F = \dfrac{QMReg}{QMR}$ for significativamente maior que 1, rejeita-se H_0 com risco α, ou seja, a regressão é significativa. Resta decidir quando considerar F significativamente maior que 1. É possível mostra que, sob o modelo de erro normal, F tem distribuição F de Fisher com 1 grau de liberdade no numerador e (n-2) graus de liberdade no denominador, e H_0 será rejeitada se $F > F_{(\alpha,1,n-2)}$ (concentra α à direita). Segue a tabela da F com 1 g.l. no numerador:

gl den	1	2	3	4	5	6	7	8	9	10	11
F_{tab}	647,793	38,506	17,443	12,218	10,007	8,813	8,073	7,571	7,209	6,937	6,724
gl den	12	13	14	15	16	17	18	19	20	21	22
F_{tab}	6,554	6,414	6,298	6,199	6,115	6,042	5,978	5,922	5,871	5,827	5,785
gl den	23	24	25	26	27	28	29	30	40	50	100
F_{tab}	5,75	5,717	5,686	5,659	5,633	5,61	5,588	5,568	5,424	5,34	5,179

Tabela 4.4: Distribuição F-Snedecor $P(F > F_{tab}) = 0,025$

Exemplo: Nesse caso temos 1 g.l. no numerador e 6 no denominador. Logo, com $2,5\%$ de significância $F_{sup} = 8,813$.

Causa de Variação (CV)	g.l	SQ	QM	F
Regressão	1	4.586,91	4.586,91	283,81
Resíduos	6	96,97	16,16	
Total	7	4.683,88		

Tabela 4.5: Anova consumo de energia em função da temperatura

Como são 8 observações no exemplo acima,temos $283,81 > 8,813$, logo rejeitamos a hipótese nula e afirmamos que a regressão é significativa, ou seja, os valores da temperatura ambiente influenciam no consumo de energia elétrica com 2,5 por cento de risco. Nesse exemplo, $S_e = 4,02$. Ou seja, em média o valor do consumo de energia elétrica estimados através da reta de regressão $\hat{Y}_i = 218,255 + 1,384x_i$ estão 4BTUs distantes dos valores reais do consumo de energia.

Observação: Também podemos fazer o teste t para a inclinação. Neste caso as hipóteses continuam sendo as mesmas dadas acima, mas a estatística de teste passa a ser:
$t = \dfrac{b - \beta}{S_b} \sim t_{(n-2)}$, onde $S_b = \dfrac{S_e}{\sqrt{SQX}} = \dfrac{S_e}{\sqrt{\Sigma(x_i - \overline{x})^2}}$ é o erro padrão da inclinação (β).

Coeficiente de Determinação (poder explicativo do modelo) (R^2)

O coeficiente de determinação tem o objetivo avaliar a "qualidade" do ajuste de um modelo de regressão. Seu valor fornece a proporção da variação total da variável Y que é explicada pela variável X através da função ajustada. Por [5] e [6] podemos expressar R^2 por:

$$R^2 = (r_{x,y})^2 \text{ ou } R^2 = \frac{SQReg}{SQTotal}$$

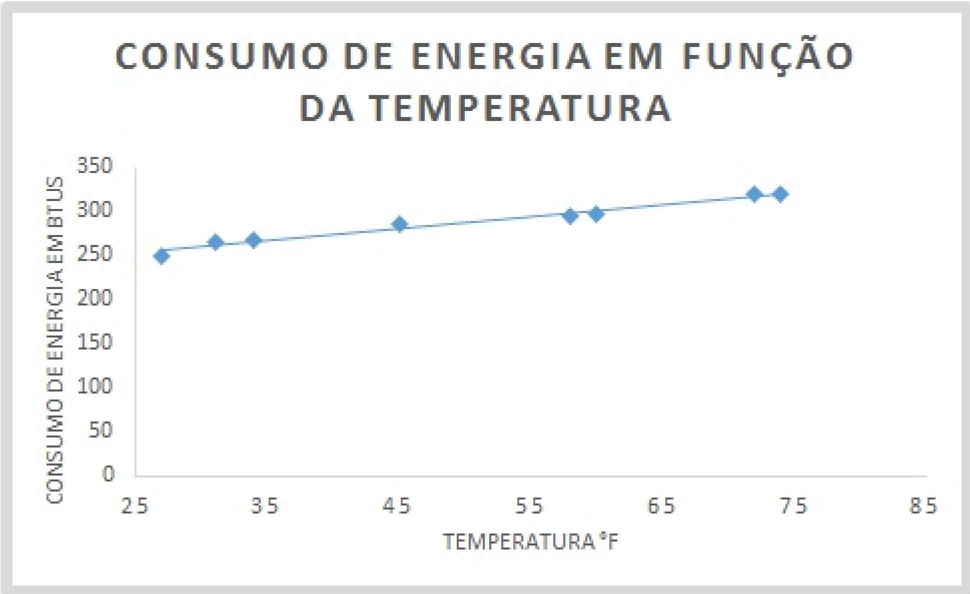
sendo $0 \leqslant R^2 \leqslant 1$.

Quando $R^2 = 0$, a variação explicada de Y é zero, a reta ajustada é paralela ao eixo de X. Se $R^2 = 1$, a reta ajustada explicará toda a variação de Y (todos os pontos estão sobre a reta de regressão).

Exemplo: No exemplo anterior: $R^2 = 0,9793$. Logo, $97,93\%$ da variação do consumo de energia elétrica é explicada pela temperatura ambiente e $2,07\%$ da variação do consumo de energia é atribuída ou à aleatoriedade ou a outras variáveis.

4.2.3 Estratégias para evitar as armadilhas de regressão

1. Sempre iniciar com um gráfico de dispersão para observar a possível relação linear entre X e Y no intervalo estudado. E no exemplo anterior pela Figura 4.1 a relação entre a temperatura e o consumo de energia se aproximam de uma reta, o que verifica essa condição.

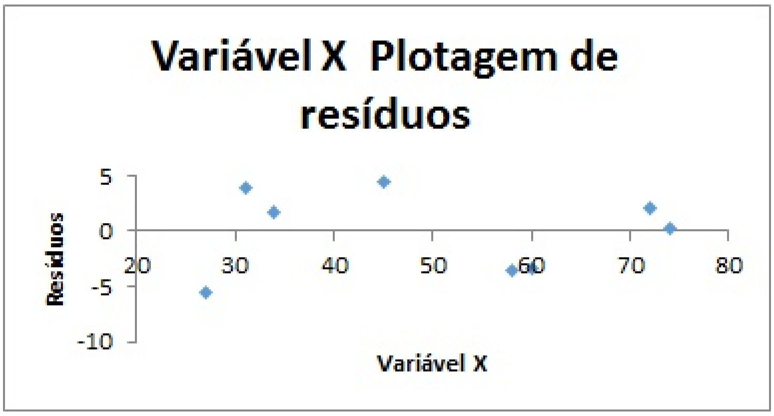


Fonte: Elaborado pelo autor

Figura 4.1: Gráfico de dispersão entre temperatura e consumo de energia.

2. Verificar as premissas da regressão depois que o modelo de regressão tenha sido ajustado, antes de prosseguir e utilizar os resultados do modelo. No exemplo anterior a relação entre X e Y é linear conforme Figura 4.1, onde os valores de X foram fixados arbitrariamente e Y é uma variável aleatória que depende entre outras coisas dos valores de X.

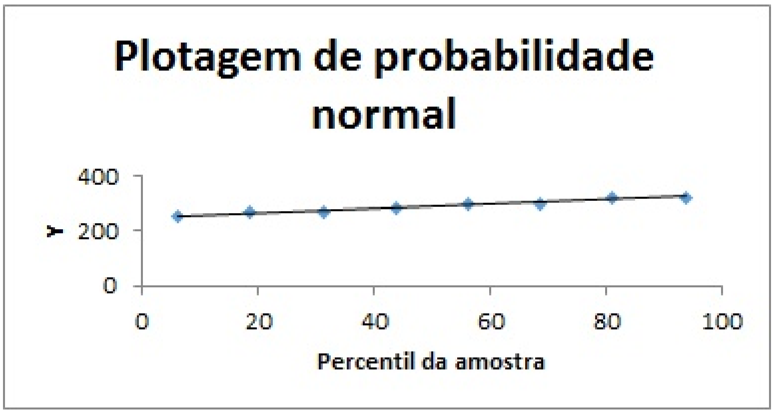
3. Fazer o gráfico dos resíduos em relação a variável independente. Isso permitirá que você não só determine se o ajuste do modelo aos dados é apropriado, bem como verifique visualmente alguma violação da premissa da homoscedasticidade. E o gráfico dos resíduos dados do exemplo mostra que os resíduos estão em torno de zero conforme Figura 4.2, verificando o ajuste do modelo.



Fonte: Elaborado pelo autor

Figura 4.2: Gráfico dos resíduos entre temperatura e consumo de energia.

4. Utilizar um histograma, uma disposição ramo-e-folha, um box-plot ou um gráfico da normal de probabilidade dos resíduos para avaliar, de forma gráfica, se a premissa da normalidade foi seriamente violada. Pela a Figura 4.3 percebe-se que essa condição é satisfeita para os dados do exempo.



Fonte: Elaborado pelo autor

Figura 4.3: Gráfico da normal de probabilidade

Para confirmar a normalidade dos resíduos dos dados faz-se o teste de normalidade como, por exemplo, o teste de Kolmogorov - Smirnov. Este teste analisa as seguintes hipóteses:

$$\begin{cases} H_0\text{: Os resíduos seguem uma distribuição normal.} \\ H_1\text{: Os resíduos não seguem uma distribuição normal.} \end{cases}$$

Em que os resíduos são colocados em ordem crescente e verifica-se qual a máxima diferença absoluta entre a função de distribuição acumulada assumida para os dados, no caso a Normal, e a função de distribuição empírica dos dados. Como critério, compara-se esta diferença com um valor crítico conforme Figura 4.4, para um dado nível de significância.

Segundo [11] a função de distribuição acumulada é definida por:

$$F(x_i) = P(X \leqslant x_i).$$

Em que o valor de $P(X \leqslant x_i) = P(Z_i \leqslant \frac{x_i - \overline{x}}{s})$ é encontrado na tabela de distribuição normal padrão que está no Anexo B, onde $\overline{x}$ é a média aritmética dos dados e s é o desvio padrão. E a função de distribuição empírica é dada por:

$$F_n(x) = \begin{cases} 0, \text{ se } x < x_1 \\ \dfrac{k}{n}, \text{ se } x_k \leqslant x < x_{k+1} \\ 1, \text{ se } x > x_n \end{cases}$$

A estatística de teste é dada por:

$$D_n = sup_x |F(x) - F_n(x)|$$

. A função de distribuição empírica F_n é descontínua e a função de distribuição hipotética é contínua, assim devemos considerar duas outras estatísticas:

$$D_1 = sup(x_i)|F(x_i) - F_n(x_i)|$$

e,

$$D_2 = sup(x_i)|F(x_i) - F_n(x_{i-1})|$$

para calcular a estatística de kolmogorov-Smirnov. Essas estatísticas medem as distâncias (vertical) entre os gráficos das duas funções, teórica e empírica, nos pontos $x_{i-)}$ e x_i. Com isso, podemos utilizar como estatística de teste

$$D_n = max(D_1, D_2).$$

Se D_n é maior que o valor crítico, rejeitamos a hipótese de normalidade dos dados com $(1 - \alpha)100\%$ de confiança. Caso contrário, não rejeitamos a hipótese de normalidade. O resumo da setatística de teste está na Tabela 4.6.

x(ordenado)	$F_n(x_i)$	$F(x_i) = P(Z_i \leqslant \frac{x_i - \bar{x}}{s})$	$\mid F(x_i) - F_n(x_i) \mid$	$\mid F(x_i) - F_n(x_{i-1}) \mid$
x_1	$\frac{1}{n}$	$F(x_1) = P(Z_1 \leqslant \frac{x_1 - \bar{x}}{s})$	$\mid F(x_1) - F_n(x_1) \mid$	$\mid F(x_1) - 0) \mid$
x_2	$\frac{2}{n}$	$F(x_2) = P(Z_2 \leqslant \frac{x_2 - \bar{x}}{s})$	$\mid F(x_2) - F_n(x_2) \mid$	$\mid F(x_2) - F_n(x_1) \mid$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_{n-1}	$\frac{n-1}{n}$	$F(x_{n-1}) = P(Z_{n-1} \leqslant \frac{x_{n-1} - \bar{x}}{s})$	$\mid F(x_{n-1}) - F_n(x_{n-1}) \mid$	$\mid F(x_{n-1}) - F_n(x_{n-2}) \mid$
x_n	$\frac{n}{n}$	$F(x_n) = P(Z_n \leqslant \frac{x_n - \bar{x}}{s})$	$\mid F(x_n) - F_n(x_n) \mid$	$\mid F(x_n) - F_n(x_{n-1}) \mid$

Tabela 4.6: Resumo da estatística de Kolmogorov - Smirnov.

Tamanho da amostra (N)	Nível de significância para $D_{crít} = máx\left|F_{obs}(X) - F_{esp}(X)\right|$				
	0,20	0,15	0,10	0,05	0,01
1	0,900	0,925	0,950	0,975	0,995
2	0,684	0,726	0,776	0,842	0,929
3	0,565	0,597	0,642	0,708	0,828
4	0,494	0,525	0,564	0,624	0,733
5	0,446	0,474	0,510	0,565	0,669
6	0,410	0,436	0,470	0,521	0,618
7	0,381	0,405	0,438	0,486	0,577
8	0,358	0,381	0,411	0,457	0,543
9	0,339	0,360	0,388	0,432	0,514
10	0,322	0,342	0,368	0,410	0,490
11	0,307	0,326	0,352	0,391	0,468
12	0,295	0,313	0,338	0,375	0,450
13	0,284	0,302	0,325	0,361	0,433
14	0,274	0,292	0,314	0,349	0,418
15	0,266	0,283	0,304	0,338	0,404
16	0,258	0,274	0,295	0,328	0,392
17	0,250	0,266	0,286	0,318	0,381
18	0,244	0,259	0,278	0,309	0,371
19	0,237	0,252	0,272	0,301	0,363
20	0,231	0,246	0,264	0,294	0,356
25	0,21	0,22	0,24	0,27	0,32
30	0,19	0,20	0,22	0,24	0,29
35	0,18	0,19	0,21	0,23	0,27
Mais de 35	$\frac{1,07}{\sqrt{N}}$	$\frac{1,14}{\sqrt{N}}$	$\frac{1,22}{\sqrt{N}}$	$\frac{1,36}{\sqrt{N}}$	$\frac{1,63}{\sqrt{N}}$

Fonte:https://www.google.com.br/search?q=tabela+do+teste+de+kolmogorov-smirnov.

Figura 4.4: Valores críticos para o teste de Kolmogorov - Smirnov.

Para os dados do exemplo temos $D_n = 0,225663191$, e pela tabela da Figura 4.4 para $n = 8$ e $\alpha = 0.05$ tem-se valor crítico 0.457. Como $D_n < 0.457$ não rejeita-se a hipótese de normalidade, logo os resíduos seguem uma distribuição normal.

5. Caso a avaliação realizada nos itens 3 e 4 indique uma violação em relação às premissas, utilize métodos alternativos em relação à regressão dos mínimos quadrados, ou modelo de mínimos quadrados alternativos (ou seja, regressão quadrática ou regressão múltipla), dependendo do que indicado pela avaliação.

6. Caso a avaliação realizada nos itens de 3 e 4 não indiquem violações em relação às premissas, então os aspectos de inferência da analise de regressão podem ser adotados. Podem ser realizados testes em relação aos coeficientes de significância e de regressão e desenvolvidos intervalos de confiança e de previsão. Como os itens 3 e 4 não indicaram violação com os dados do exemplo anterior, logo a relação estimada pode ser adotada.

Capítulo 5

Aplicação

A aplicação foi feita com os dados da preparação e da avaliação externa (SPAECE) de 301 alunos de 1º ano da escola estadual Wilebaldo Aguiar do município de Massapê no Ceará, onde nesse estabelecimento leciona-se somente Ensino Médio. O número de alunos por turma estão na Tabela 5.1, em que os primeiros A, B,C e D são do turno manhã e os primeiros E, F, G e H são do turno tarde.

TURMA	TURNO	Nº DE ALUNOS
1º A	Manhã	41
1º B	Manhã	40
1º C	Manhã	42
1º D	Manhã	45
1º E	Tarde	38
1º F	Tarde	31
1º G	Tarde	34
1º H	Tarde	30

Tabela 5.1: Número de alunos por turma.

Os resultados das notas da preparação sofreram influencias das matrizes de referencias do SPAECE e ENEM, pois todo o currículo escolar é elaborado baseado nestes eixos norteadores. Além disso, teve o projeto Jovem de Futuro propondo as metodologias utilizadas nas aulas, e dando apoio financeiro e pedagógico para o desenvolvimento das seguintes atividades: Avaliação diagnóstica e monitoramento, turma máster, projeto atletas da matemática, apostila do SPAECE, gincana de matemática, recreio cultural, tutoria e monitoria que são o conjunto das ações desenvolvidas na fase de preparação.

A avaliação Diagnóstica e de Monitoramento é uma avaliação composta de 40 questões de múltipla escolha, sendo 20 de português e 20 de matemática, com duração de 3 horas. As avaliações objetivas contemplam os descritores, habilidades, da Matriz de referência do SPAECE, a partir dos resultados das avaliações os alunos são divididos em quatro níveis, muito crítico, crítico, intermediário e adequado. E daí, os muito críticos e críticos que tiveram como ir para a escola no contra turno receberam apoio de monitores e tutores e os adequados participaram da sala máster.

Na Turma Máster os alunos receberam aulas de aprofundamento matemático, fora do seu turno regular, alunos que eram da manhã assistiram aula no turno da tarde e vice e versa, desta forma foram formadas duas turmas, uma pela manhã e uma pela tarde com 25 alunos cada. As aulas foram ministradas por 4 tutores que eram universitários contratados pelo projeto Jovem de Futuro, eles se dividiam em duplas, sendo sempre dois por sala, pois o principal objetivo da Turma Máster era trabalhar com turmas pequenas que os tutores pudessem acompanhar melhor e dar atenção necessária a cada aluno. As aulas possuíam um roteiro, de explicação e resolução de exercícios em grupo, levando o aluno a aprender seja com o colega do lado ou com um dos tutores que ficavam na sala orientando. Essas aulas ocorreram durante 3 meses dois dias na semana totalizando 8h semanais.

No Projeto Atletas da Matemática as aulas são ministradas com conteúdos avançados de preparação para a OBMEP- Olimpíada Brasileira de Matemática das Escola Públicas. Essas aulas eram ministradas também no contra turno para alunos que foram pré-selecionados, de todas as series, e os que quisessem tirar dúvidas ou acrescentar seu conhecimentos matemáticos e , assim, melhorar suas proficiências no SPAECE.

Uma apostila foi elaborada exclusiva para o trabalho de preparação para o Spaece, com 400 questões divididas entre os 26 descritores de aprendizagem matemática. Todos os alunos receberam a apostila no mês de Setembro, onde os professores trabalharam com ela nas aulas regulares e também na gincana, os alunos podiam levar para casa e estudar em grupo fora do horário de aula.

A estratégia da reta final de revisão dos conteúdos de forma dinâmica e não cansativa, foi a realização de uma Gincana de Matemática, onde as turmas foram divididos em equipes encabeçadas pelos alunos da Turma Master, sendo de 6 a 8 alunos por equipe. Em cada etapa os alunos tinham que estudar as questões da apostila conforme os conteúdos de cada fase, na hora da gincana a coordenação organizou uma mesa de jurados, extra escola, o professor da sala sorteava questões para as equipes irem responder no quadro, assim eram geradas notas para cada equipe. O período da gincana foi dividido em uma semana de estudo domiciliar e em grupo e outra de gincana em sala, durante dois meses Outubro e Novembro, e a melhore equipe foi premiada com medalha e uma viagem.

No mesmo período da Gincana de Matemática os recreios eram realizado de forma diferente, no auditório da escola foi montado um ambiente de gincana geral onde todos os alunos poderiam participar concorrendo a vários prêmios, as questões eram apresentadas em um telão pelos professores para os alunos responderem, no final de cada participação dos alunos os professores da mesa de jurados acrescentavam comentários tornando assim um recreio de aprendizagem.

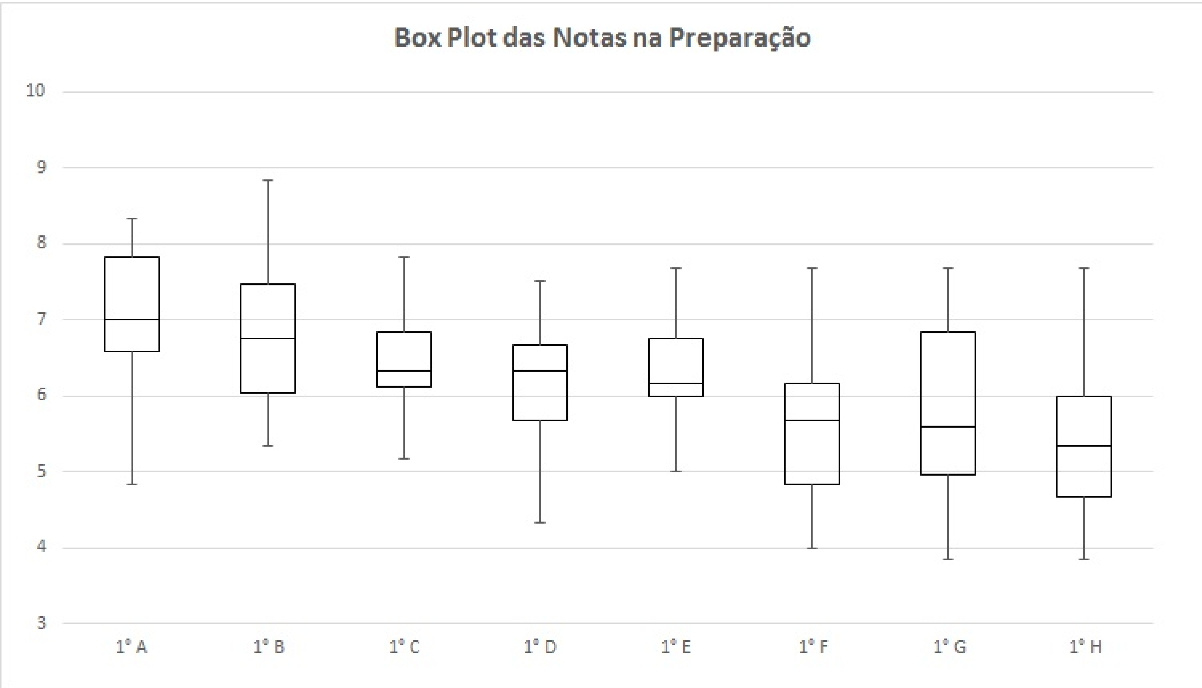
A cada 50 dias letivos durante essa preparação os alunos receberam uma nota representado seu desempenho, e como a prova do SPAECE é geralmente no mês de novembro os alunos tinham três notas. Tomando a média aritmética dessas três notas, denominadas notas na preparação, e a proficiência obtida no SPAECE foram obtidos os dados a seguir.

5.1 Análise descritiva dos dados

Os resultados da análise descritiva dos dados dos alunos na preparação que podem ser visualizados na Tabela 5.2 e no box plot Figura

Turma	Média	Moda	Mediana	Desvio Padrão	Variância
1º A	7	6,7	7	0,8	0,7
1º B	6,8	6	6,5	0,9	0,8
1º C	6,4	6,3	6,3	0,7	0,5
1º D	6,2	6,3	6,3	0,8	0,6
1º E	6,4	6	6,2	0,9	0,8
1º F	5,6	6	5,7	0,8	0,7
1º G	5,7	5	5,6	1,1	1,2
1º H	5,3	6	5,3	0,6	0,9

Tabela 5.2: Análise descritiva dos dados da preparação.



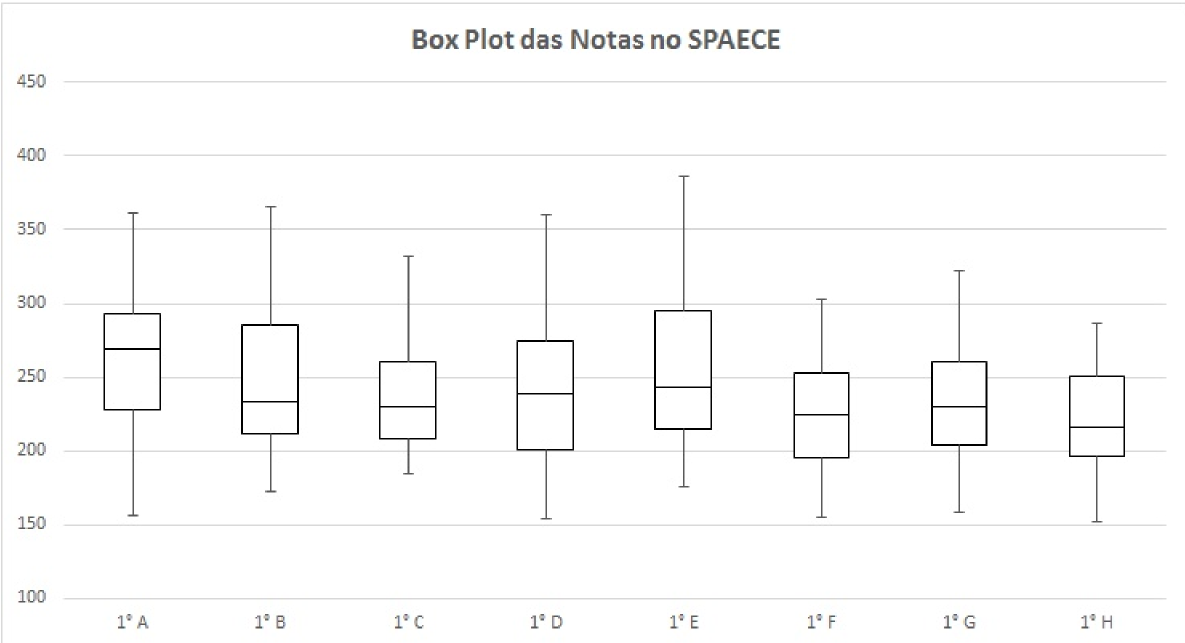
Fonte: Elaborado pelo autor.

Figura 5.1: Box plot das notas na preparação

Quanto aos resultados no SPAECE, o 1º A obteve a maior média e o 1º H a menor. O Desvio Padrão e a Variância nos mostram que os dados estão muito dispersos, com valores muito acima da média ou valores muito abaixo da média, ver Tabela 5.3 e box plot 5.4.

Turma	Média	Mediana	Desvio Padrão	Variância
1º A	265	268,9	47,9	2290,5
1º B	249,9	233,2	50,6	2560,4
1º C	237,8	229,6	37,7	1423,9
1º D	242,3	238,9	52,9	2795,8
1º E	254,8	243,1	53,5	2868,1
1º F	226,6	224,5	37,8	1426,9
1º G	231,4	230,2	42,3	1785,9
1º H	222,6	215,3	44,1	1943,4

Tabela 5.3: Análise descritiva dos dados no SPAECE.



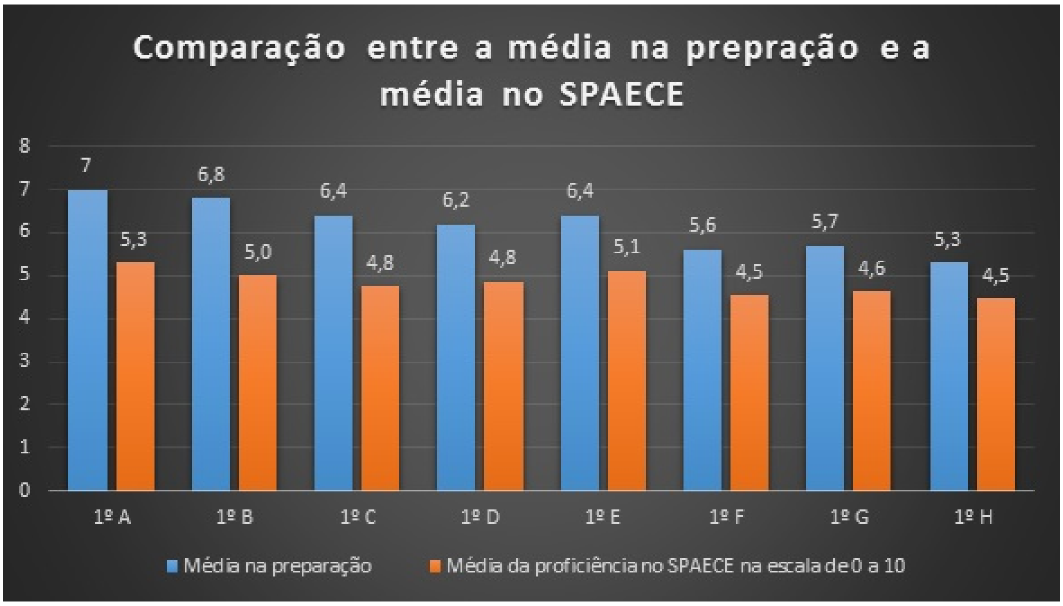
Fonte: Elaborado pelo autor.

Figura 5.2: Box plot das notas no SPAECE.

Como as notas da preparação e do SPAECE não estavam na mesma escala, foi necessário fazer uma transformação nos dados do SPAECE usando regra de três simples, pois as notas variam de 0 à 500 e para fazer a comparação com as notas da preparação fez-se necessário estarem na escala de 0 a 10. Assim, comparando a média das turmas na preparação e no SPAECE, percebe-se que a turma que teve maior média na preparação também teve maior média no SPAECE, ver Tabela 5.4 e Figura 5.3.

Turma	Média na preparação	Média da proficiência no SPAECE	Média da proficiência no SPAECE na escala de 0 a 10
1º A	7	265	5,3
1º B	6,8	249,9	5,0
1º C	6,4	237,8	4,8
1º D	6,2	242,3	4,8
1º E	6,4	254,8	5,1
1º F	5,6	226,6	4,5
1º G	5,7	231,4	4,6
1º H	5,3	222,6	4,5

Tabela 5.4: Comparação entre a média na preparação e a média no SPAECE.

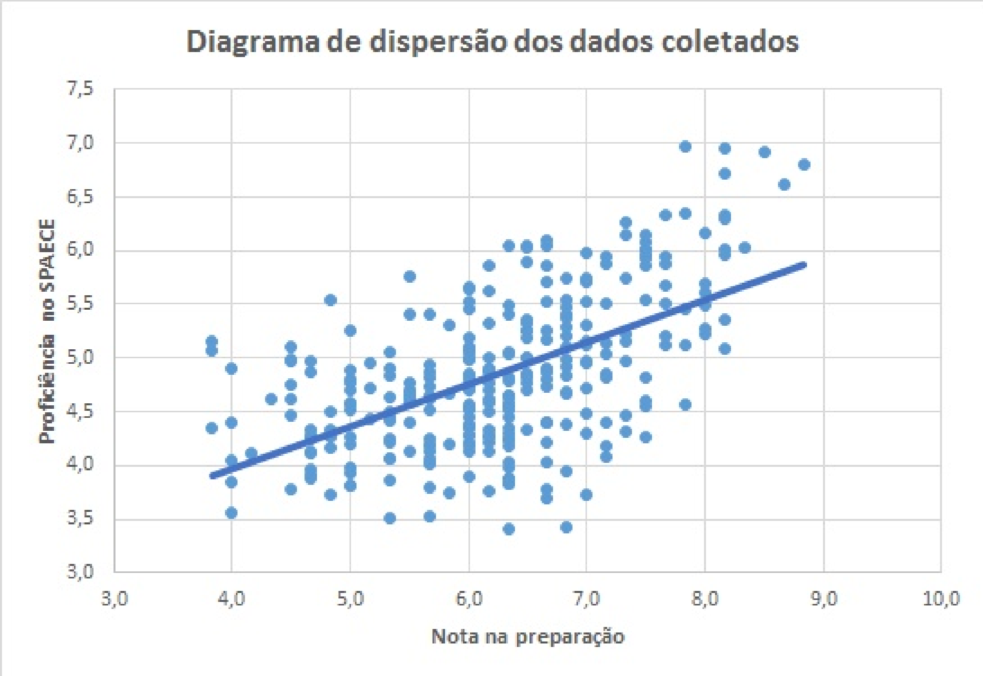


Fonte: Elaborado pelo autor.

Figura 5.3: Comparação entre a média na preparação e a média no SPAECE.

5.2 Ajuste da reta de regressão

Comforme os dados coletados, temos o diagrama de dispersão que se encotra na Figura 5.4, onde mostra que os dados aproximam-se de uma reta, sendo que as notas da preparação representam a variável independente X e as proficiências no SPAECE representa a variável dependente Y.



Fonte: Elaborado pelo autor.

Figura 5.4: Gárfico de dispersão dos dados coletados

Usando o excel verifica-se que o coeficiente de correlação de Pearson entre X e Y é $r_{x,y} = 0,56$. E fazendo o teste da existêcia de correlação, onde a hipótese nula é $H_0 : r_{x,y} = 0$ e a hipótese alternativa é $H_1 : r_{x,y} \neq 0$, temos $t_{observac\tilde{a}o} = 11,68$ e para 5% de significância com $gl = 299$ pela Tabela de Distribuição t de Student que está no Anexo A tem-se $t_{tab} = 1,96$. Como $t_{observac\tilde{a}o} > t_{tab}$, Rejeitamos a hipótese nula que é a não existência de corelação entre as variaveis X e Y. Portanto, as variaveis X e Y estão correlacionadas.

Pelos dados coletados temos a seguinte relação

$$\widehat{y} = 2,4 + 0,4x$$

para o modelo de regressão linear simples cuja anova está na Tabela 5.5 e Tabela 5.6, onde $a = 2,4$ e $b = 0,4$.

	gl	SQ	MQ	F	F de significação
Regressão	1	47,66768	47,66768	138,4845	1,58E-26
Resíduo	299	102,9186	0,344209		
Total	300	150,5863			

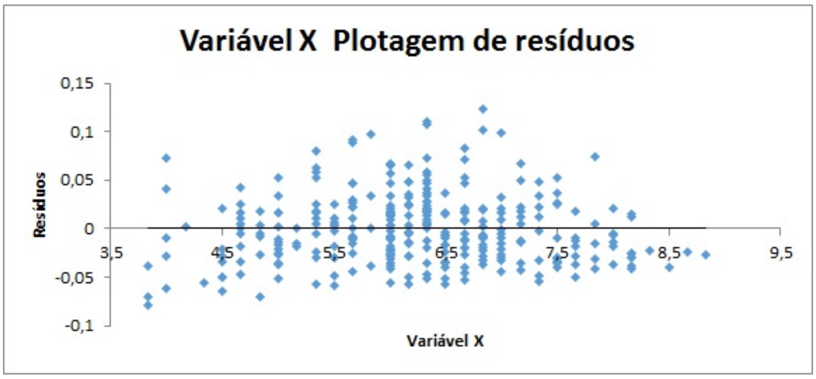
Tabela 5.5: ANOVA dos dados coletados

	Coeficientes	Erro padrão	valor-P	95% inferiores	95% superiores	Inferior 95,0%	Superior 95,0%
Interseção	2,407793572	0,210371	2,13E-25	1,993798	2,821789	1,993798	2,821789
Variável X	0,391418853	0,033261	1,58E-26	0,325963	0,456875	0,325963	0,456875

Tabela 5.6: Informações sobre coeficientes

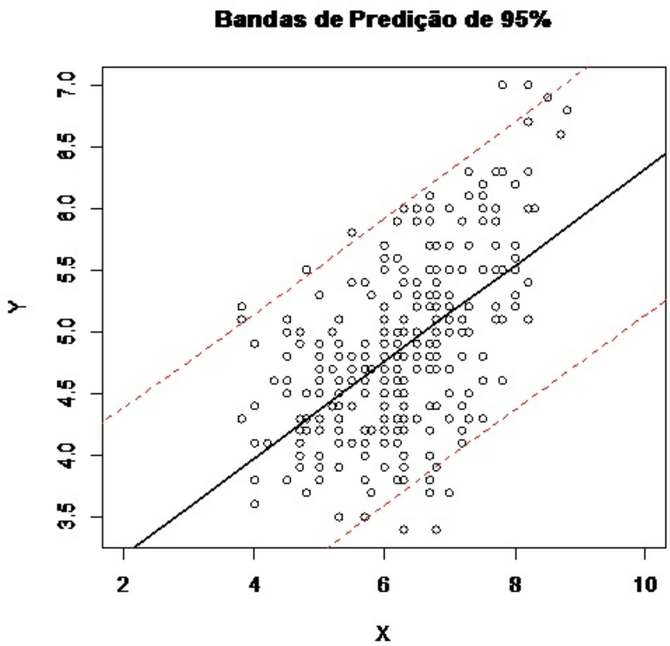
A partir da análise dos dados da Tabela 5.5 e da Tabela 5.6, verifica-se que a regressão é significativa, ou seja, as notas da preparação influenciam na proficiência do SPAECE. Além disso, $S_e = 0,59$, isto é, em média a nota da proficiência estimada através da reta de regressão $\hat{y_i} = 2,4 + 0,4x_i$, i = 1, ... ,301 estão 0,59 distantes dos valores reais da proficiencia no SPAECE.

A adequação do modelo também verifica-se na análise dos resíduos dos dados em análise, conforme Figura 5.5 e Figura 5.6, em que os resíduos estão ditribuídos aleatoriamente em torno do zero, sem nenhuma observação muito discrepante e seguem uma distribuição normal conforme Figura 5.7. Na Figura 5.6 tem-se o gráfico do intervalo de predição, onde observa-se que apenas 12 das 301 observações estão fora do intervalo de predição, ou seja, menos de 5% das observações.



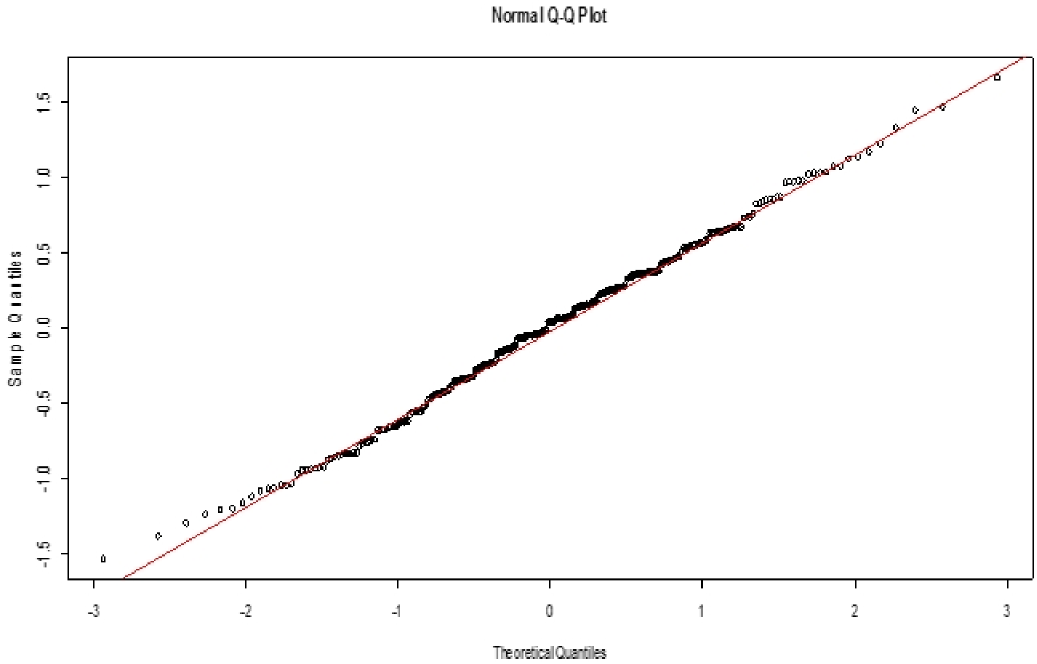
Fonte: Elaborado pelo autor.

Figura 5.5: Resíduos dos dados coletados



Fonte: Elaborado pelo autor.

Figura 5.6: Bandas de predição



Fonte: Elaborado pelo autor.

Figura 5.7: Gráfico da normalidade dos resíduos

Fazendo o teste de normalidade de Kolmogorov-Smirnov para os resíduos conforme descrito no Capítulo 3, tem-se $D_n = 0,028617211$ e pela tabela da Figura 4.4 com $\alpha = 0.05$ e $n = 301$ o valor crítico é 0.0783. Como $D_n < 0.0783$ não rejeita-se a hipótese de normalidade dos resíduos, logo pela Figura 5.7 e por este teste os resíduos dos dados observados seguem o critério de normalidade.

Comforme análise das ANOVAS que estão no apêndice, os coeficientes da reta de regressão linear simples de cada turma com seus respectivos erros, estão na Tabela 5.7.

Turma	a	e	b	e
1º A	76,4	56,1	26,7	7,9
1º B	-4,3	49,5	37,4	7,2
1º C	64,4	46,6	27	7,2
1º D	61,1	59,7	29,9	9,6
1º E	-10,2	46,1	41,3	7,1
1º F	120	43,8	19	7,7
1º G	183,8	38,8	8,3	6,6
1º H	80,9	38,2	26,7	7,1

Tabela 5.7: Ajuste da reta de regressão.

5.3 Inferência Teste "t"

O teste será o teste T pareado feito com os dados dos 301 alunos observados. Vamos testar as seguintes hipóteses:

H_0: Nota na preparação = proficiência no SPAECE;

H_1: Nota na preparação > proficiência no SPAECE;

Ou seja, se $T_{observacão} > T_{tabela}$ com nível de significância $\alpha = 0.05$, rejeitamos H_0, caso contrário aceitamos H_0.

Segundo ([6], pág. 202) a estatísica t para dados pareados é definida por:

$$t = \frac{\overline{D} \times \sqrt{n}}{S_D}$$

onde,

n: Tamanho das amostras, que, neste caso, corresponde ao número de pares observados;

$\overline{D}$: Média das diferenças internas dos pares;

S_D: Desvio padrão das diferenças internas dos pares;

Dos dados em análise temos $\overline{D} = 1,4$ e $S_D = 0.8$. Daí,

$$t = \frac{1,4 \times \sqrt{301}}{0,8} = 30,4.$$

Logo, Pela tabela t de Student (que está em anexo) para $gl = 300$ conforme amostra, temos $t = 1,6$ menor do que o valor do t da observação que é $t = 30,4$. Portanto, rejeitamos H_0, e afirmamos que as notas na preparação são melhores que os resultados no SPAECE.

5.4 Estimação

Pelo exposto anteriormente, o modelo de regressão linear simples é adequado para estimar as notas no SPAECE e verificar se a preparação para essa avaliação externa foi boa ou não. No caso dos dados em análise a média no SPAECE foi 242,5, o que é considerado conforme Figura 2.3 nível critico.

Uma vez que grande parte da variação da variável resposta é explicada pelas variáveis regressoras, podemos utilizar o modelo estimado para predizer valores de Y correspondentes a valores das variáveis independentes que fazem ou não parte do conjunto de dados. Asssim, por exemplo, podemos fazer uma estimação da proficiência no SPAECE de um aluno que obteve nota 10 na preparação.

A análise utilizando regressão linear simples mostrou que se a escola mantiver os mesmos procedimentos e metodologias na preparação para o SPAECE, provavelmente sua proficiencia será baixa. Pois segundo a relação de regressão $\widehat{y} = 2,4 + 0,4x$ se o aluno obtiver nota 10 na preparação, ele obterá $\widehat{y} = 6,4$ o que corresponde aproximadamente no SPAECE uma proficiencia de 320 pontos e que de acordo com a Figura 2.3 corresponde ao nível intermeiario, assim, a escola atingiria no máximo o nível intermediario.

Portanto, O modelo de regressão linear simples pode contribuir de forma significativa para a escola, pois possibilita prever resultados e, assim, fazer com que a instituição de ensino faça a adequação de suas ações de acordo os números previstos. E em particular, o professor também pode utilizar esse método de previsão para a média bimestral ou anual de seus alunos, pois segue os mesmos critérios da preparação para o SPAECE.

Referências Bibliográficas

[1] LIMA, E.L. - *Matemática e Ensino*. 3. ed. Rio de Janeiro: SBM, 2007.

[2] CEARA, S. - *Sistema Permanente de Avaliação da Educação Básica do Ceará:SPACE 2006: relatório técnico-pedagógico de matemática*. 1. ed. Rio de Janeiro: Fundação Cesgranrio, 2007.

[3] SANTOS, S. - *Impactos gerados pelo Sistema Permanente de Avaliação da Educação Básica do Ceará-SPACE na melhoria do ensino aprendizagem do ensino medio*. 1. ed. Fortaleza: UFC, 2010.

[4] CEARA, S. - *Sistema Permanente de Avaliação da Educação Básica do Ceará:SPACE 2008: Boletim Pedagógico de Avaliação: Matemática, Ensino Médio*. 1. ed. Juiz de Fora: CAED, 2008.

[5] CHARNET, R. ,FREIRE C.A.L. , CHARNET, E.M.R. BONVINO, H. - *Análise de Modelos de Regressão Linear com Aplicações*. 2. ed. Editora da Unicamp, 2008.

[6] BARBETTA, Pedro Alberto. - *Estatística aplicada ás ciências sociais*. 8. ed. rev. -Florianópolis: Ed. da UFSC, 2012.

[7] SILVA COSTA, EVA MARIA. - *Uma análise estatística dos resultados de um projeto de aprendizagem matemática no Ensino Médio*. Teresina: [s.n.], 2015.

[8] BERNAL, R.; SILVA, N. - *O uso do EXCEL para análises estatística*. 1. ed. São Paulo: USP, 2012.

[9] BUSSAB, WILTON de O. MORENTTIN, PEDRO A. - *Estatística Básica*. 5. ed. São Paulo: Saraiva,2004.

[10] MARTINS, G. - *Estatística geral e aplicada*. 3. ed. São Paulo: Atlas, 2010.

[11] Portal Action. - *Teste de Kolmogorov-Smirnov.* Disponível em:¡http://www.portalaction.com.br/inferencia/62-teste-de-kolmogorov-smirnov¿. Acesso em 02 de Julho de 2016.

Apêndice

ANOVA

	gl	SQ	MQ	F	F de significação
Regressão	1	20840,74304	20840,74	11,4835326	0,001617341
Resíduo	39	70778,65386	1814,837		
Total	40	91619,3969			

	Coeficientes	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores	Inferior 95,0%	Superior 95,0%
Interseção	76,41628734	56,03830493	1,363644	0,18049955	-36,93188321	189,7644579	-36,93188321	189,7644579
Variável X	26,75014609	7,893841544	3,388736	0,00161734	10,78334447	42,7169477	10,78334447	42,7169477

Apêndice A - Anova 1º A

ANOVA

	gl	SQ	MQ	F	F de significação
Regressão	1	41290,25	41290,25	26,79058	7,65311E-06
Resíduo	38	58566,45	1541,222		
Total	39	99856,7			

	Coeficientes	rro padrã	Stat t	valor-P	95% inferiores	95% superiores	Inferior 95,0%	Superior 95,0%
Interseção	-4,253401538	49,48897	-0,08595	0,93196	-104,4385932	95,93179013	-104,4385932	95,93179013
Variável X 1	37,39499471	7,224743	5,175962	7,65E-06	22,76926776	52,02072166	22,76926776	52,02072166

Apêndice B - Anova 1º B

ANOVA

	gl	SQ	MQ	F	F de significação
Regressão	1	15163,68	15163,68	14,03572	0,000565889
Resíduo	40	43214,56	1080,364		
Total	41	58378,24			

	Coeficientes	rro padrã	Stat t	valor-P	95% inferiores	5% superiores	Inferior 95,0%	Superior 95,0%
Interseção	64,37285087	46,56046	1,382565	0,174469	-29,72934344	158,4750452	-29,72934344	158,4750452
Variável X 1	27,00626797	7,20854	3,746427	0,000566	12,43726594	41,57527	12,43726594	41,57527

Apêndice C - Anova 1º C

ANOVA

	gl	SQ	MQ	F	F de significação
Regressão	1	21965,45	21965,45	9,3468	0,00383245
Resíduo	43	101052,2	2350,05		
Total	44	123017,6			

	Coeficiente:	rro padrãc	Stat t	valor-P	95% inferiores	95% superiores	Inferior 95,0%	Superior 95,0%
Interseção	61,11709	59,70003	1,023736	0,311685	-59,27949741	181,5136819	-59,27949741	181,5136819
Variável X 1	29,43293	9,627244	3,057254	0,003832	10,01773875	48,84811577	10,01773875	48,84811577

Apêndice D - Anova 1º D

ANOVA

	gl	SQ	MQ	F	F de significação
Regressão	1	51251,27	51251,27	33,62698	1,29049E-06
Resíduo	36	54868,03	1524,112		
Total	37	106119,3			

	Coeficiente:	rro padrãc	Stat t	valor-P	95% inferiores	95% superiores	Inferior 95,0%	Superior 95,0%
Interseção	-10,1916	46,14087	-0,22088	0,826434	-103,7696025	83,38644121	-103,7696025	83,38644121
Variável X 1	41,30383	7,122729	5,798877	1,29E-06	26,85826467	55,74939137	26,85826467	55,74939137

Apêndice E - Anova 1º E

ANOVA

	gl	SQ	MQ	F	F de significação
Regressão	1	7370,2369	7370,237	6,031251	0,020284546
Resíduo	29	35438,229	1222,008		
Total	30	42808,466			

	Coeficientes	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores	Inferior 95,0%	Superior 95,0%
Interseção	120,04756	43,831782	2,738825	0,010432	30,40150059	209,6936188	30,40150059	209,6936188
Variável X 1	18,998518	7,7359918	2,455861	0,020285	3,176637949	34,82039733	3,176637949	34,82039733

Apêndice F - Anova 1º F

ANOVA

	gl	SQ	MQ	F	F de significação
Regressão	1	2742,116896	2742,117	1,561588	0,220495821
Resíduo	32	56191,36778	1755,98		
Total	33	58933,48467			

	Coeficientes	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores	Inferior 95,0%	Superior 95,0%
Interseção	183,7712657	38,76930353	4,740123	4,22E-05	104,8007786	262,7417527	104,8007786	262,7417527
Variável X 1	8,31507004	6,653999164	1,249635	0,220496	-5,238682725	21,8688228	-5,238682725	21,8688228

Apêndice G - Anova 1º G

ANOVA

	gl	SQ	MQ	F	F de significação
Regressão	1	18933,7339	18933,73	14,16501	0,000788994
Resíduo	28	37426,3383	1336,655		
Total	29	56360,0723			

	Coeficientes	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores	Inferior 95,0%	Superior 95,0%
Interseção	80,95783	38,2251748	2,117919	0,043192	2,657108849	159,2585509	2,657108849	159,2585509
Variável X 1	26,699467	7,09404771	3,763644	0,000789	12,16796864	41,2309646	12,16796864	41,2309646

Apêndice H - Anova 1º H

Anexo A

Distribuição t de Student												
	Área contida nas duas caudas laterais (bicaudal) da distribuição t de Student											
gl/q	0,990	0,980	0,975	0,950	0,900	0,800	0,200	0,100	0,050	0,025	0,020	0,010
	Área contida na cauda superior ou inferior (unicaudal) da distribuição t de Student											
graus de liberdade	0,995	0,990	0,9875	0,975	0,950	0,900	0,100	0,050	0,025	0,0125	0,010	0,005
1	0,0157	0,0314	0,0393	0,0787	0,1584	0,3249	3,0777	6,3138	12,7062	25,4517	31,8205	63,6567
2	0,0141	0,0283	0,0354	0,0708	0,1421	0,2887	1,8856	2,9200	4,3027	6,2053	6,9646	9,9248
3	0,0136	0,0272	0,0340	0,0681	0,1366	0,2767	1,6377	2,3534	3,1824	4,1765	4,5407	5,8409
4	0,0133	0,0267	0,0333	0,0667	0,1338	0,2707	1,5332	2,1318	2,7764	3,4954	3,7469	4,6041
5	0,0132	0,0263	0,0329	0,0659	0,1322	0,2672	1,4759	2,0150	2,5706	3,1634	3,3649	4,0321
6	0,0131	0,0261	0,0327	0,0654	0,1311	0,2648	1,4398	1,9432	2,4469	2,9687	3,1427	3,7074
7	0,0130	0,0260	0,0325	0,0650	0,1303	0,2632	1,4149	1,8946	2,3646	2,8412	2,9980	3,4995
8	0,0129	0,0259	0,0323	0,0647	0,1297	0,2619	1,3968	1,8595	2,3060	2,7515	2,8965	3,3554
9	0,0129	0,0258	0,0322	0,0645	0,1293	0,2610	1,3830	1,8331	2,2622	2,6850	2,8214	3,2498
10	0,0129	0,0257	0,0321	0,0643	0,1289	0,2602	1,3722	1,8125	2,2281	2,6338	2,7638	3,1693
11	0,0128	0,0256	0,0321	0,0642	0,1286	0,2596	1,3634	1,7959	2,2010	2,5931	2,7181	3,1058
12	0,0128	0,0256	0,0320	0,0640	0,1283	0,2590	1,3562	1,7823	2,1788	2,5600	2,6810	3,0545
13	0,0128	0,0256	0,0319	0,0639	0,1281	0,2586	1,3502	1,7709	2,1604	2,5326	2,6503	3,0123
14	0,0128	0,0255	0,0319	0,0638	0,1280	0,2582	1,3450	1,7613	2,1448	2,5096	2,6245	2,9768
15	0,0127	0,0255	0,0319	0,0638	0,1278	0,2579	1,3406	1,7531	2,1314	2,4899	2,6025	2,9467
16	0,0127	0,0255	0,0318	0,0637	0,1277	0,2576	1,3368	1,7459	2,1199	2,4729	2,5835	2,9208
17	0,0127	0,0254	0,0318	0,0636	0,1276	0,2573	1,3334	1,7396	2,1098	2,4581	2,5669	2,8982
18	0,0127	0,0254	0,0318	0,0636	0,1274	0,2571	1,3304	1,7341	2,1009	2,4450	2,5524	2,8784
19	0,0127	0,0254	0,0318	0,0635	0,1274	0,2569	1,3277	1,7291	2,0930	2,4334	2,5395	2,8609
20	0,0127	0,0254	0,0317	0,0635	0,1273	0,2567	1,3253	1,7247	2,0860	2,4231	2,5280	2,8453
21	0,0127	0,0254	0,0317	0,0635	0,1272	0,2566	1,3232	1,7207	2,0796	2,4138	2,5176	2,8314
22	0,0127	0,0254	0,0317	0,0634	0,1271	0,2564	1,3212	1,7171	2,0739	2,4055	2,5083	2,8188
23	0,0127	0,0253	0,0317	0,0634	0,1271	0,2563	1,3195	1,7139	2,0687	2,3979	2,4999	2,8073
24	0,0127	0,0253	0,0317	0,0634	0,1270	0,2562	1,3178	1,7109	2,0639	2,3909	2,4922	2,7969
25	0,0127	0,0253	0,0317	0,0633	0,1269	0,2561	1,3163	1,7081	2,0595	2,3846	2,4851	2,7874
26	0,0127	0,0253	0,0316	0,0633	0,1269	0,2560	1,3150	1,7056	2,0555	2,3788	2,4786	2,7787
27	0,0127	0,0253	0,0316	0,0633	0,1268	0,2559	1,3137	1,7033	2,0518	2,3734	2,4727	2,7707
28	0,0126	0,0253	0,0316	0,0633	0,1268	0,2558	1,3125	1,7011	2,0484	2,3685	2,4671	2,7633
29	0,0126	0,0253	0,0316	0,0633	0,1268	0,2557	1,3114	1,6991	2,0452	2,3638	2,4620	2,7564
30	0,0126	0,0253	0,0316	0,0632	0,1267	0,2556	1,3104	1,6973	2,0423	2,3596	2,4573	2,7500
31	0,0126	0,0253	0,0316	0,0632	0,1267	0,2555	1,3095	1,6955	2,0395	2,3556	2,4528	2,7440
32	0,0126	0,0253	0,0316	0,0632	0,1267	0,2555	1,3086	1,6939	2,0369	2,3518	2,4487	2,7385
33	0,0126	0,0253	0,0316	0,0632	0,1266	0,2554	1,3077	1,6924	2,0345	2,3483	2,4448	2,7333
34	0,0126	0,0253	0,0316	0,0632	0,1266	0,2553	1,3070	1,6909	2,0322	2,3451	2,4411	2,7284
35	0,0126	0,0252	0,0316	0,0632	0,1266	0,2553	1,3062	1,6896	2,0301	2,3420	2,4377	2,7238
36	0,0126	0,0252	0,0316	0,0631	0,1266	0,2552	1,3055	1,6883	2,0281	2,3391	2,4345	2,7195
37	0,0126	0,0252	0,0316	0,0631	0,1265	0,2552	1,3049	1,6871	2,0262	2,3363	2,4314	2,7154
38	0,0126	0,0252	0,0315	0,0631	0,1265	0,2551	1,3042	1,6860	2,0244	2,3337	2,4286	2,7116
39	0,0126	0,0252	0,0315	0,0631	0,1265	0,2551	1,3036	1,6849	2,0227	2,3313	2,4258	2,7079
40	0,0126	0,0252	0,0315	0,0631	0,1265	0,2550	1,3031	1,6839	2,0211	2,3289	2,4233	2,7045
45	0,0126	0,0252	0,0315	0,0631	0,1264	0,2549	1,3006	1,6794	2,0141	2,3189	2,4121	2,6896
48	0,0126	0,0252	0,0315	0,0630	0,1263	0,2548	1,2994	1,6772	2,0106	2,3139	2,4066	2,6822
50	0,0126	0,0252	0,0315	0,0630	0,1263	0,2547	1,2987	1,6759	2,0086	2,3109	2,4033	2,6778
55	0,0126	0,0252	0,0315	0,0630	0,1262	0,2546	1,2971	1,6730	2,0040	2,3044	2,3961	2,6682
60	0,0126	0,0252	0,0315	0,0630	0,1262	0,2545	1,2958	1,6706	2,0003	2,2990	2,3901	2,6603
63	0,0126	0,0252	0,0315	0,0630	0,1262	0,2544	1,2951	1,6694	1,9983	2,2962	2,3870	2,6561
70	0,0126	0,0252	0,0315	0,0629	0,1261	0,2543	1,2938	1,6669	1,9944	2,2906	2,3808	2,6479
75	0,0126	0,0252	0,0314	0,0629	0,1261	0,2542	1,2929	1,6654	1,9921	2,2873	2,3771	2,6430
80	0,0126	0,0251	0,0314	0,0629	0,1261	0,2542	1,2922	1,6641	1,9901	2,2844	2,3739	2,6387
85	0,0126	0,0251	0,0314	0,0629	0,1260	0,2541	1,2916	1,6630	1,9883	2,2818	2,3710	2,6349
90	0,0126	0,0251	0,0314	0,0629	0,1260	0,2541	1,2910	1,6620	1,9867	2,2795	2,3685	2,6316
95	0,0126	0,0251	0,0314	0,0629	0,1260	0,2541	1,2905	1,6611	1,9853	2,2775	2,3662	2,6286
99	0,0126	0,0251	0,0314	0,0629	0,1260	0,2540	1,2902	1,6604	1,9842	2,2760	2,3646	2,6264
100	0,0126	0,0251	0,0314	0,0629	0,1260	0,2540	1,2901	1,6602	1,9840	2,2757	2,3642	2,6259
120	0,0126	0,0251	0,0314	0,0628	0,1259	0,2539	1,2886	1,6577	1,9799	2,2699	2,3578	2,6174
100000	0,0125	0,0251	0,0313	0,0627	0,1257	0,2533	1,2816	1,6449	1,9600	2,2414	2,3264	2,5759

Anexo A - Tabela t de Student

Anexo B

z_0	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4967	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Anexo B - Tabela de Distribuição Normal

Anexo C

FÓRMULAS DO EXCEL PARA ESTATÍSTICA

ESTATÍSTICA DESCRITIVA

Média: =MÉDIA(A1:A20)

Mediana: =MED(A1:A20)

Variância: = VAR.A(A1:A20)

Desvio Padrão: =DESVPAD.A(A1:A20)

CORRELAÇÃO

Coefiente de correlação de Pearson: =PEARSON(matriz x;matriz y)

REGRESSÃO LINEAR SIMPLES

Segue-se os passos:

1. Selecione DADOS na barra de menus;
2. Análise de dados;
3. Na caixa de diálogo selecione REGRESSÃO;

OBTER VALOR TABELA DISTRIBUIÇÃO NORMAL

= DIST.NORM.N(x;média;desvio padrão;cumulativo)

Anexo C - Fórmulas Excel