



UEPB

UNIVERSIDADE ESTADUAL DA PARAÍBA

Centro de Ciências Humanas e Exatas
Campus de Monteiro

ÉDERSON DA SILVA BEZERRA

**O USO DE TECNOLOGIAS NO ENSINO DE
ESTATÍSTICA: UMA ANÁLISE DE CORRELAÇÃO
E REGRESSÃO**

Monteiro – PB
2016

ÉDERSON DA SILVA BEZERRA

**O USO DE TECNOLOGIAS NO ENSINO DE
ESTATÍSTICA: UMA ANÁLISE DE CORRELAÇÃO
E REGRESSÃO**

Trabalho de Conclusão de Curso de Graduação apresentado ao Centro de Ciências Humanas e Exatas – CCHE da Universidade Estadual da Paraíba – UEPB, Campus Monteiro, em cumprimento às exigências legais para obtenção do título de grau de Licenciado em Matemática.

Orientadora: Profa. Ma. Patrícia Melo Rocha

Monteiro – PB
2016

É expressamente proibida a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano da dissertação.

B574u Bezerra, Éderson da Silva.

O uso de tecnologias no ensino de Estatística [manuscrito] :
uma análise de correlação e regressão / Ederson da Silva Bezerra.
2016.

43 p. : il.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Matemática)
- Universidade Estadual da Paraíba, Centro de Ciências Humanas e
Exatas, 2016.

"Orientação: Profa. Ma. Patrícia Melo Rocha, Departamento
de Exatas".

1. Ensino de Estatística. 2. Correlação estatística. 3.
Regressão estatística. 3. Tecnologia digital. I. Título.

21. ed. CDD 519.5

ÉDERSON DA SILVA BEZERRA

O USO DE TECNOLOGIAS NO ENSINO DE ESTATÍSTICA: UMA ANÁLISE
DE CORRELAÇÃO E REGRESSÃO

Trabalho de Conclusão de Curso de
Graduação apresentado ao Centro de
Ciências Humanas e Exatas – CCHE da
Universidade Estadual da Paraíba – UEPB,
Campus Monteiro, em cumprimento às
exigências legais para obtenção do título de
grau de Licenciado em Matemática.

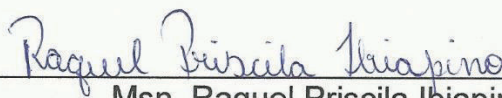
Orientadora: Profa. Ma. Patrícia Melo Rocha

Aprovado em 22/12/2016

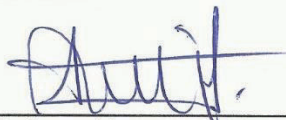
BANCA EXAMINADORA



Ma. Patrícia Melo Rocha
Orientadora



Msn. Raquel Priscila Ibiapino
Examinadora



Dr. Roger Ruben Huaman Huanca
Examinador

Dedico este trabalho a minha mãe e a minha esposa que sempre estiveram ao meu lado me dando forças para seguir em frente e chegar até aqui.

AGRADECIMENTOS

Agradeço em primeiro lugar a Deus, pois sem ele eu não teria forças para essa longa jornada.

Agradeço a minha família, especialmente, a minha mãe Elinete que teve papel fundamental desde a minha infância até este momento, pois sempre se preocupou com minha educação, de forma incansável, acreditando sempre no melhor que havia em mim.

A minha orientadora, a Professora Ma. Patrícia Rocha pela sua paciência, dedicação e incentivo que tornaram possível a conclusão deste trabalho.

Aos professores do Curso de Licenciatura Plena em Matemática, do Campus VI da Universidade Estadual da Paraíba - UEPB, em especial, aos professores membros da banca, Dr. Roger Huanca e Msn. Raquel Priscila Ibiapino que contribuíram para conclusão deste trabalho e, por terem aceito o convite para fazer parte dessa banca

A UEPB, em especial ao Campus Monteiro, por ter sido minha segunda casa, me dando o suporte necessário para me qualificar profissionalmente.

Enfim, agradeço a todos que direta ou indiretamente fizeram parte da minha formação acadêmica.

“A menos que modifiquemos a nossa maneira de pensar, não seremos capazes de resolver os problemas causados pela forma como nos acostumamos a ver o mundo”.

Albert Einstein.

RESUMO

O objetivo deste trabalho foi contribuir para o Ensino de Estatística fazendo uso do Excel no estudo de Correlação e Regressão. As aplicações estatísticas fazendo uso de tecnologias podem ser inseridas e adaptadas a qualquer nível de ensino, permitindo o tratamento de temas diversos de interesse dos alunos, dando-lhes a autonomia de escolha e decisão, exercitando o questionamento, a argumentação e a reflexão, que constituem as capacidades essenciais para a formação plena do indivíduo para o exercício da sua cidadania crítica e reflexiva. A presente pesquisa tem caráter bibliográfico e traz uma alternativa de como aprender Correlação e Regressão, tendo o computador como aliado e desencadeador da aprendizagem. Verificamos que o ensino de Regressão e Correlação fazendo uso do Excel, proporciona autonomia nos alunos de determinar os resultados que o computador lhe passa, mas para isso, eles precisam estar conscientes do que estão buscando, saber analisar os resultados obtidos e utilizá-los em suas interpretações, o que subentende a existência de conhecimentos prévios, discernimento e compreensão sobre o conteúdo que está sendo abordado.

Palavras-chave: Ensino de Estatística. Correlação e Regressão. Tecnologia Digital.

ABSTRACT

The purpose of this paper was to contribute to the teaching of statistic by using Excel in the study of Correlation and Regression. Statistical applications making use of technologies can be inserted and adapted to any level of education, allowing the treatment of diverse subjects of interest of the students, giving them the autonomy of choice and decision, exercising the questioning, the argumentation and the reflection, that constitute the essential capacities for the full formation of the individual for the exercise of his critical and reflective citizenship. The present research has bibliographic character and brings an alternative of how to learn Correlation and Regression, having the computer as ally and trigger of the learning. We checked that Regression and Correlation teaching using Excel gives students the autonomy to determine the results that the computer gives them, but for this, they need to be aware of what they are looking for, to know how to analyze the results obtained and to use them in Their interpretations, the existence of previous knowledge, insight and understanding about the content being addressed.

Keywords: Statistics Teaching. Correlation and Regression. Digital Technology.

SUMÁRIO

1 INTRODUÇÃO	09
2 O USO DA TECNOLOGIA COMO UMA FERRAMENTA NO ENSINO DE ESTATÍSTICA	13
2.1 UMA BREVE HISTÓRIA SOBRE A ESTATÍSTICA.....	13
2.2 O ENSINO DE ESTATÍSTICA E A ANÁLISE DE DADOS.....	14
2.3 A TECNOLOGIA E A EDUCAÇÃO.....	18
3 APLICAÇÃO DO SOFTWARE EXCEL COMO UM AUXÍLIO PARA O ENSINO DE CORRELAÇÃO E REGRESSÃO	21
3.1 O COMPUTADOR E O USO DE SOFTWARES.....	21
3.2 O SOFTWARE EXCEL.....	23
3.3 REGRESSÃO E CORRELAÇÃO.....	24
3.4 APLICANDO O EXCEL.....	33
4 CONSIDERAÇÕES FINAIS	38
REFERÊNCIAS	40

1 INTRODUÇÃO

A Estatística está presente no dia a dia do homem, até mesmo em suas atividades mais rotineiras e vem conquistando crescente importância na sociedade atual. Através das estatísticas pode-se fazer previsões, determinar tendências, auxiliar na tomada de decisões e, portanto, fornecer um planejamento e uma conclusão com maior precisão.

Sabe-se que o desenvolvimento tecnológico e conseqüentemente, o volume de informações disponibilizadas para sociedade exigem uma grande demanda por análises de informações e tomada de decisões, que são baseadas em estatísticas.

Esta ideia está sustentada nas características que distinguem essa ciência de outras. Por meio de ferramentas de representação e análise inferencial, a Estatística possibilita a compreensão e o estudo de fenômenos que envolvem grande quantidade de dados, variabilidade e incertezas, fenômenos estes cada vez mais presentes no cotidiano das pessoas (ESTEVAM; KALINKE, 2013).

Assim, o ensino de Estatística poderia ser aperfeiçoado ao se utilizar softwares como uma ferramenta de apoio, uma vez que seu uso pode favorecer a tomada de consciência a compreensão de determinados conceitos estatísticos.

O conhecimento de conceitos estatísticos, aliado ao uso de tecnologias, é uma competência essencial na formação profissional. O profissional em formação, pertencente à geração digital, tem se apresentado um pesquisador que aceita os desafios e a imprevisibilidade da época para se aprimorar.

Silva e Schimiguel (2013) dizem que, vivemos em um mundo em constante evolução, onde a todo instante surgem novas tecnologias em diversas áreas do conhecimento. Esta evolução não acontece de modo diferente na área da Educação. Segundo os autores, ao longo do tempo surgiram distintas metodologias, bem como correntes teóricas e também equipamentos diversos, sempre tendo como foco, propiciar melhorias significativas no processo de ensino e aprendizagem. Desde o pergaminho da antiguidade aos e-books de hoje, do quadro negro à lousa interativa, da prática conservadora à progressista, não há limites às inovações na esfera educacional.

Mas, considerando a educação atual complexa, ou seja, onde praticamente se depende de uma enorme quantidade de componentes de alta

tecnologia, esperava-se que todos estivessem preparados para interpretar corretamente quaisquer medidas estatística, como também tabelas, e gráficos, e compreender facilmente a inferência estatística. Entretanto, sabemos que não é algo tão simples quando se trata do confronto direto com os números, onde parece que o ensino de Estatística continua afastado da tecnologia. Isto porque, grande parte da população ainda se vê angustiada pelo medo de não conseguir entender os conceitos e formulações comuns a esta ou outras áreas do conhecimento que se relacione, direta ou indiretamente, com a Matemática, com os cálculos, com os gráficos etc.

Assim, o presente trabalho traz uma alternativa para o aprendizado da análise estatística de dados, mais precisamente o ensino de Correlação e a Análise de Regressão Linear Simples, tendo o computador como seu aliado e desencadeador da aprendizagem.

Dessa forma, visando atender a um novo perfil de aluno, futuros profissionais, entende-se que o ensino precisa ser reformulado, possibilitando o surgimento de outras formas de ensino que se adaptem melhor às constantes mudanças da sociedade, cujo conhecimento é continuamente e rapidamente atualizado.

Quanto aos procedimentos de investigação, a pesquisa tem caráter bibliográfico, onde foram analisados alguns livros, capítulos de livros e artigos que tratam sobre Estatística, Correlação e Regressão e Tecnologias.

Neste contexto, esta pesquisa visa contribuir para o Ensino de Estatística, mais precisamente no estudo da Correlação e Regressão Linear Simples, tendo o uso de tecnologias como aliado e desencadeador da aprendizagem.

Considerando-se nesta pesquisa as Tecnologias Digitais como sendo os computadores e suas interfaces, como softwares, por exemplo, o Microsoft Excel.

Justificativa da pesquisa

O motivo pelo qual escolhi a Licenciatura em Matemática vem desde a minha infância, pois diferentemente da maioria das crianças sempre gostei de números, cálculos e de resolver problemas, de modo que minhas melhores notas do Ensino Fundamental ao Ensino Médio sempre foram em Matemática. Então em 2009 fiz o vestibular para o Curso de Licenciatura Plena em Matemática no

Campus VI da UEPB, conseguindo a 3^o colocação e iniciando uma longa, porém, maravilhosa jornada de conhecimento.

A escolha deste tema de trabalho se deu também pela admiração que tenho pela Estatística e por toda a sua contribuição para as diversas áreas de conhecimento. Esse interesse cresceu quando estava estudando para o concurso do Instituto Brasileiro de Geografia e Estatística – IBGE, que é a maior referência em Estatística no Brasil.

Mas foi no nono e último período do curso que veio o contato com a disciplina Estatística e Probabilidade e todo esse mundo de coleta e interpretação de dados, entre outras coisas, me encantou. Foi todo esse envolvimento com a Estatística que me levou a tomar a decisão de realizar meu Trabalho de Conclusão de Curso nessa área que é tão fortemente presente na sociedade contemporânea.

Vale destacar que em se tratando de um Trabalho de Conclusão de Curso, a escolha não foi fácil, até porque tivemos apenas uma disciplina de Estatística, o que não faz da escolha desse tema algo tão comum. Já com relação ao uso de tecnologia, o interesse veio porque é algo muito presente em nossas vidas, então pensei na possibilidade de fazer um estudo relacionado com a Estatística, daí veio a ideia de fazer um estudo sobre o uso de Tecnologias no Ensino de Estatística, mais precisamente na Análise de Correlação e Regressão.

Objetivos

O principal objetivo deste trabalho é contribuir para o Ensino de Estatística fazendo uso do Excel no estudo de Correlação e Regressão.

A fim de trazer compreensões a respeito do objetivo geral acima, apresentamos alguns objetivos específicos:

- Descrever como usar a ferramenta Correlação e Regressão Linear Simples do Excel.
- Contribuir na aprendizagem da Estatística e no processo de construção do conhecimento a partir da Tecnologia Digital.

Organização da pesquisa

O nosso trabalho estrutura-se em quatro capítulos, apresentando, no primeiro uma introdução sobre o tema trabalhado, ressaltando a justificativa, os objetivos e por fim a organização da pesquisa.

O capítulo dois aborda uma breve história sobre a estatística, passando por discussões sobre o Ensino de Estatística e a Análise de Dados, e posteriormente sobre Tecnologia e Educação.

O capítulo três concentra-se no uso de softwares no ensino de Correlação e Regressão, assim, veremos conceitos de Correlação e Regressão e uma aplicação de análise de regressão com o Excel.

Por fim, o capítulo quatro trata das considerações finais do nosso trabalho.

2 O USO DA TECNOLOGIA COMO UMA FERRAMENTA NO ENSINO DE ESTATÍSTICA

Este capítulo foi elaborado com objetivo de apresentar uma breve história sobre a estatística e discutir sobre o Ensino de Estatística e a Análise de Dados. Além disso, discutiremos sobre Tecnologia e Educação.

2.1 UMA BREVE HISTÓRIA SOBRE A ESTATÍSTICA

De fato, não é fácil saber quando se originou a história de qualquer ramo do conhecimento, pois isso vai depender do conceito que fizermos dele e que, naturalmente, evoluirá no decorrer do tempo. “A história da Estatística bem confirma esta asserção” (MEMÓRIA, 2004, p. 10).

Até os dias atuais, o conceito da palavra Estatística é associado apenas a dados numéricos apresentados em tabelas e gráficos, geralmente publicados na mídia, referentes a fatos demográficos ou econômicos. Memória (2004), diz que a etimologia da palavra, do latim *status* (estado), era usada para designar a coleta e a apresentação de dados quantitativos de interesse do Estado, o que bem reflete sua origem.

Entretanto, o autor afirma que a mera coleta de dados assim apresentados está longe de ser o que entendemos, hoje, por Estatística. Na verdade, Memória (2004) diz que a feição essencial da Estatística é a de ser um conjunto de métodos (métodos estatísticos), especialmente apropriado ao tratamento de dados numéricos afetados por uma multiplicidade de causas. Esses métodos fazem uso da Matemática, particularmente do cálculo de probabilidades, na coleta, apresentação, análise e interpretação de dados quantitativos.

Desde remota antiguidade, os governos têm se interessado por informações sobre suas populações e riquezas, tendo em vista, principalmente, fins militares e tributários. O registro de informações perde-se no tempo. Confúcio relatou levantamentos feitos na China, há mais de 2000 anos antes da era cristã. No antigo Egito, os faraós fizeram uso sistemático de informações de caráter estatístico, conforme evidenciaram pesquisas arqueológicas. Desses registros também se utilizaram as civilizações pré-colombianas dos maias, astecas e incas. É conhecido de todos os cristãos o recenseamento dos judeus, ordenado pelo Imperador Augusto (MEMÓRIA, 2004, p. 11).

No entanto, o estudo da Estatística não se resume apenas a contagem da população. Alguns matemáticos importantes contribuíram para essa

abrangência da Estatística, como por exemplo, Laplace e Gauss. Esses matemáticos, no início do século XIX transformaram a concepção da Estatística. Este ramo da ciência passou a não somente designar interesses para o Estado, mas também a aceitar os erros de uma observação.

Mas existiram outros matemáticos que fizeram grandes contribuições, a exemplo de Francis Galton que foi um dos pesquisadores que deu maior impulso à aplicação da Estatística, introduzindo o conceito de Correlação e Regressão. E juntamente com o matemático Pearson, Galton desenvolveu uma das fórmulas mais utilizadas para o cálculo de Correlação, a prova de Qui-Quadrado e outras contribuições.

Em meados do século XX, difundiram-se os métodos multivariados para analisar simultaneamente conjuntos de variáveis. Suas origens devem ser buscadas na generalização da correlação e da regressão. Alguns desses métodos são: a análise fatorial, análise multivariada da variância, a regressão múltipla, a análise discriminante, a correlação canônica, a análise de correspondências, os modelos log-lineares, as escalas multidimensionais, a análise das equações estruturais, etc. Os métodos multivariados permanecem no terreno da estatística matemática devido à complexidade dos cálculos que requer. Contudo, com a recente proliferação e uso de computadores, esses métodos são cada vez mais utilizados em pesquisa educativa (BISQUERRA; SARRIERA; MARTINEZ, 2004, p. 16-17).

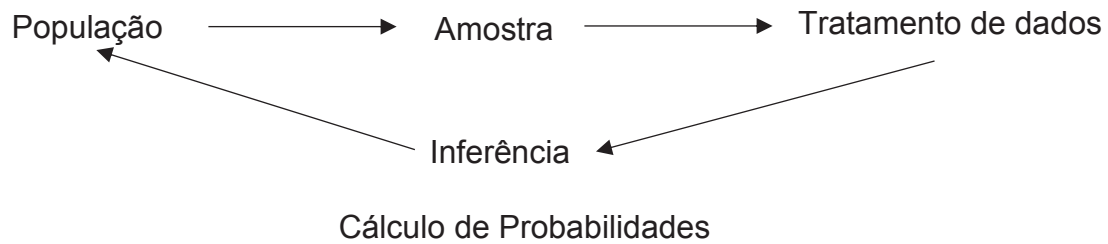
2.2 O ENSINO DE ESTATÍSTICA E A ANÁLISE DE DADOS

Até pouco tempo atrás a Estatística, segundo Botter et al. (1996, p. 1) “era considerada pelo leigo como uma sucessão de tabelas e gráficos associados a algum tipo de pesquisa”. Do mesmo modo, Crespo (2009) diz que, atualmente, o público leigo posiciona-se em dois extremos divergentes e igualmente errôneos com relação às análises e conclusões estatísticas. Ou eles creem fielmente ou afirmam que a estatística nada prova. Segundo esse autor, “os que assim pensam ignoram os objetivos, o campo e o rigor do método estatístico; ignoram a Estatística, quer teórica quer prática, ou a conhecem muito superficialmente” (CRESPO, 2009, p. 2).

Mas, em cada século seguinte mais e mais áreas foram se incorporando ao conjunto das áreas que faziam uso da estatística.

Nas últimas décadas, com o avanço das tecnologias, houve um avanço significativo das áreas de estatística e probabilidade pois, com pacotes específicos ou mesmo linguagens mais poderosas, o pesquisador dispõe de muitas ferramentas alternativas para seu trabalho (BOTTER et al., 1996, p. 2).

Nesse sentido, Botter et al. (1996) sintetizaram as etapas de uma pesquisa estatística e apresentam um esquema.



Para os autores, cada uma dessas passagens merece um estudo aprofundado e tem características próprias.

- População – é o conjunto de interesse final para a pesquisa. Em geral é o conjunto do qual as amostras são retiradas. Dizemos em geral, pois nem sempre tal população, que é a população alvo está disponível para o pesquisador.
- Amostra – chamaremos de amostra a qualquer subconjunto da população de interesse, quer dados tenham sido coletados de um estudo observacional, quer sejam provenientes de um experimento realizado sob certas condições de controle.
- Tratamento dos dados – conjunto de técnicas usadas para descrever os dados observados.
- Inferência – conjunto de métodos que permitem inferir o comportamento de uma população a partir do conhecimento da amostra.
- Cálculo das Probabilidades – ferramenta matemática que deduz, a partir de um modelo, as propriedades de um fenômeno aleatório.

Botter et al. (1996, p. 3) dizem que, “cada uma das etapas tem características próprias e poderia ser encarada como um tema a ser discutido exhaustivamente”. Ainda segundo os autores, essa tarefa é parcialmente cumprida nos bacharelados em estatística, que é um curso com ferramentas matemáticas apropriadas. Quando eles dizem parcialmente, estão se referindo que mesmo em 4 ou 5 anos de curso, não se consegue absorver todo o material possível. Eles ainda fazem um questionamento em relação a um curso introdutório de apenas um semestre, que é o que acontece na maioria dos cursos

de Graduação, que oferecem apenas uma disciplina de Estatística em sua grade curricular.

Para Botter et al. (1996), ao se oferecer um curso introdutório de estatística pretende-se dar ao aluno das diversas áreas do conhecimento, noções básicas, de forma a facilitar sua interação com a comunidade estatística ou mesmo com a literatura de sua área específica que use análise estatística.

Um dos componentes de uma análise estatística é a exploração de dados, comumente chamada de Estatística descritiva.

Esta análise serve como um primeiro guia ao pesquisador, fornecendo informações sobre a qualidade de seus dados e indicando algumas tendências (se existirem) e, em geral, não tem um fim em si própria, exceto os casos de censos. Estamos certos de que a “intimidade” com os dados que esta primeira análise proporciona é o caminho adequado para um prosseguimento de análise numa linha inferencial (BOTTER et al., 1996, p.11).

Mas essa “intimidade” deveria ter início desde o Ensino Básico. Porém o fato é que em algumas universidades no Brasil, o licenciado do curso de Matemática estuda em sua graduação apenas uma disciplina sobre Estatística. E são esses profissionais que atuam na Educação Básica lecionando o conteúdo Estatística.

Sendo esses conhecimentos teóricos de apenas um semestre que servirão como base para as aulas deste assunto. De modo geral, é dada maior ênfase aos aspectos técnicos e operacionais do que aos problemas reais.

Ter noções de Estatística vai além de saber a teoria. Calcular medidas de tendência central (média, moda e mediana), construir gráficos e tabelas, coletar dados quantitativos não é suficiente para um estudante. O conhecimento estatístico vai além disso, é necessário saber interpretar dados e tirar conclusões.

Campos, Wodewotzki e Jacobini (2011) dizem que da mesma forma que há a preocupação em relacionar a Matemática e o cotidiano, desejável em todos os níveis escolares, faz-se necessário abordar os conteúdos estatísticos na direção de uma formação ampla do aluno. Nesse sentido, a evolução da estatística e o debate sobre sua inserção nos currículos dos primeiros anos da escola resultou na inclusão de seus conteúdos no programa da disciplina de matemática pelos Parâmetros Curriculares Nacionais (PCN) em 1997. Os PCN constituem-se em referência curricular nacional para o Ensino Fundamental. No

documento elaborado, de uma forma geral, destaca-se a importância da educação para a cidadania, da geração de um aluno crítico capaz de interagir na sociedade em que vive, com habilidades para trabalhar com as informações que o rodeia como indica os Parâmetros Curriculares Nacionais, fazendo crescer a presença desses conteúdos no Ensino Básico.

Os conteúdos do bloco Análise de dados e Probabilidade têm sido recomendados para todos os níveis do Ensino Básico, em especial para o Ensino Médio.

Uma das razões desse ponto de vista reside na importância das ideias de incerteza e de probabilidade, associadas aos chamados fenômenos aleatórios, presentes de forma essencial nos mundos natural e social. O estudo desse bloco de conteúdo possibilita aos alunos ampliarem e formalizarem seus conhecimentos sobre o raciocínio combinatório, probabilístico e estatístico. Para dar aos alunos uma visão apropriada da importância dos modelos probabilísticos no mundo de hoje, é importante que os alunos tenham oportunidade de ver esses modelos em ação. Por exemplo, é possível simular o que ocorre em certa pesquisa de opinião estimando, com base em uma amostra, a fração de balas de determinada cor em uma caixa (BRASIL, 2006, p.78).

Assim, o estudo da estatística viabiliza a aprendizagem da formulação de perguntas que podem ser respondidas com uma coleta de dados, organização e representação. Dessa forma, ao longo do Ensino Médio, os alunos devem aprimorar as habilidades adquiridas no Ensino Fundamental no que se refere à coleta, à organização e à representação de dados.

As Orientações Curriculares Nacionais recomendam um trabalho com ênfase na construção e na representação de tabelas e gráficos mais elaborados, analisando sua conveniência e utilizando tecnologias, quando possível (BRASIL, 2006, p. 78).

Durante o ensino médio, os alunos precisam adquirir entendimento sobre o propósito e a lógica das investigações estatísticas, bem como sobre o processo de investigação. Deve-se possibilitar aos estudantes o entendimento intuitivo e formal das principais ideias matemáticas implícitas em representações estatísticas, procedimentos ou conceitos (BRASIL, 2006, p. 79).

Os Parâmetros Curriculares Nacionais ainda orientam que o ensino da estatística por meio de situações-problema envolve os alunos, que podem dedicar mais tempo à construção de estratégias e se sentir estimulados a testar suas hipóteses e interpretar resultados de resolução se dispuserem de calculadoras para efetuar cálculos, geralmente muito trabalhosos. Para isso

também há softwares interessantes, como os de planilhas eletrônicas, os que permitem construir diferentes tipos de gráfico (BRASIL, 1998, p. 85).

O fato é que a Estatística ao ser desenvolvida desde o Ensino Básico, é um fator positivo, pois quando o aluno chega ao Ensino Superior poderá aprofundar ainda mais seus conhecimentos, estudando outras técnicas de análise e testes estatísticos que hoje não são trabalhados devido à falta de conhecimentos básicos em relação à Estatística quando estes ingressam na Universidade.

2.3 A TECNOLOGIA E A EDUCAÇÃO

As tecnologias invadem as nossas vidas, ampliam a nossa memória, garantem novas possibilidades de bem-estar e fragilizam as capacidades naturais do ser humano. Somos muito diferentes de nossos antepassados, nos acostumamos com alguns confortos tecnológicos (KENSKY, 2010, p. 19).

Assim, Kensky (2010) diz que, a evolução tecnológica não se restringe apenas aos novos usos de determinados equipamentos e produtos. Ela altera comportamentos. Ainda segundo a autora, a ampliação e a banalização do uso de determinada tecnologia impõem-se à cultura existente e transformam não apenas o comportamento individual, mas o de todo grupo social.

Na atualidade, o surgimento de um novo tipo de sociedade tecnológica é determinado principalmente pelos avanços das tecnologias digitais de comunicação e informação e pela microeletrônica. Essas novas tecnologias - assim consideradas em relação às tecnologias anteriormente existentes -, quando disseminadas socialmente, alteram as qualificações profissionais e a maneira como as pessoas vivem cotidianamente, trabalham, informam-se e se comunicam com outras pessoas e com todo o mundo (KENSKY, 2010, p. 22).

Nesse sentido, tecnologias e educação são indissociáveis. Segundo o dicionário *Aurélio*, a educação diz respeito ao “processo de desenvolvimento da capacidade física, intelectual e moral do ser humano, visando à sua melhor integração individual e social”. Mas para que essa integração ocorra Kensky (2010, p. 43) expõe que, “é preciso que conhecimentos, valores, hábitos, atitudes e comportamentos do grupo sejam ensinados e aprendidos, ou seja, que se utilize a educação para ensinar sobre as tecnologias”. A autora ainda completa dizendo que as tecnologias estão na base da identidade e da ação do grupo, e que se faça o uso delas para ensinar as bases dessa educação.

Não há dúvida de que o uso de tecnologias trouxe mudanças consideráveis e positivas para a educação, a saber:

Vídeos, programas educativos na televisão e no computador, sites educacionais, softwares direcionados transformam a realidade da aula tradicional, dinamizam espaço de ensino-aprendizagem, onde, anteriormente, predominava a lousa, o giz, o livro e a voz do professor. Para que as tecnologias possam trazer alterações no processo educativo, no entanto, elas precisam ser compreendidas e incorporadas pedagogicamente. Isso significa que é preciso respeitar as especificidades do ensino e da própria tecnologia para poder garantir que o seu uso, realmente, faça diferença. Não basta usar a televisão ou o computador, é preciso saber usar de forma pedagogicamente correta a tecnologia escolhida (KENSKY, 2010, p. 46).

Assim, não há como deixar de lado os avanços tecnológicos, principalmente quando tratamos das relações de ensino. Neste caso, é preciso incorporar a tecnologia à sala de aula, aos métodos de ensino e ao cotidiano.

Nesse sentido, a autora diz que, para que isso se concretize, é preciso olhá-los de uma nova perspectiva. Até aqui, os computadores e a internet têm sido vistos, sobretudo, como fontes de informação e como ferramentas de transformação dessa informação. Mais do que o caráter instrumental e restrito do uso das tecnologias para realização de tarefas em sala de aula, é chegada a hora de alargar os horizontes da escola e de seus participantes, ou seja, de todos (KENSKY, 2010, p. 66).

Já Prensky (2001), acredita que houve uma mudança radical em nossos alunos e considera que o nosso sistema educacional não foi criado para os alunos de hoje, pois estes estão extremamente integrados à chegada e à rápida difusão da tecnologia digital ocorrida no final do século XX. O autor denomina os alunos da atualidade de “Nativos Digitais”, pois são todos “falantes nativos” desta linguagem digital presente no mundo dos computadores, dos vídeos, dos games e da internet.

Por outro lado, Prensky (2001) denomina de “Imigrantes Digitais” aqueles que não nasceram no mundo digital, por terem afinidade, fascínio ou por terem adotado as novidades tecnológicas, sempre sendo comparados em contraponto aos Nativos Digitais. Receber informações de forma rápida é algo rotineiro para os Nativos Digitais. Realizar tarefas múltiplas e processar mais de uma coisa por vez é habitual. Visualizar gráficos antes do texto é preferível ao invés do contrário. Preferem acessar aleatoriamente os conteúdos, como hipertextos, por

exemplo, e só trabalham melhor quando estão conectados a uma rede de contato (PRENSKY, 2001).

De fato, é necessário promover a integração, a criatividade, a criticidade e a cidadania. O que nos faz repensar os conceitos, métodos, ambientes e os próprios currículos dos cursos, buscando inserir os avanços da era digital ao contexto educacional, assimilando os benefícios da tecnologia à formação profissional.

Para Ponte, Brocardo e Oliveira (2013), as tecnologias têm exercido um papel de grande influência no ensino de “Estatística”, uma vez que possibilitam a realização de cálculos e facilitam o uso de variadas formas de representação. Os autores enfatizam que as tecnologias “permitem o tratamento de dados reais, ao invés de trabalhar apenas com amostras de pequena dimensão, com valores escolhidos artificialmente de modo a proporcionar cálculos simples”. Dessa forma, acreditamos que seja fundamental não só o uso, mas, primordialmente, o domínio de tecnologias que possibilitem identificar de forma rápida questões que possam influenciar positivamente ou negativamente a tomada de decisões.

3 APLICAÇÃO DO SOFTWARE EXCEL COMO UM AUXÍLIO PARA O ENSINO DE REGRESSÃO E CORRELAÇÃO

Neste capítulo nos propusemos a aplicar o software Excel no ensino de Regressão e Correlação. Para melhor estruturar o capítulo, iniciamos falando do computador e do uso de software. Posteriormente vamos conhecer um pouco sobre o Software Excel. Depois sobre a Regressão e Correlação e por fim ilustraremos a análise de regressão com o Excel.

3.1 O COMPUTADOR E O USO DE SOFTWARES

Bisquerra, Sarriera e Martinez (2004) dizem que, foi a partir do início dos anos 80, que houve uma generalização do uso dos computadores.

Este processo se iniciou nos Estados Unidos no início dos anos 70, mas chegou até nós muito mais tarde. Hoje, praticamente todas as pesquisas realizam suas análises de dados mediante procedimentos informáticos. Isto permite falar de uma revolução na estatística aplicada. Entre as características que mais se destacam nesta revolução está o fato de que o cálculo deixou de ter importância, já que o computador o faz. Agora podem-se propor análises que antes eram impensáveis, e o tempo investido nesta fase diminui de forma considerável (BISQUERRA; SARRIERA; MARTINEZ, 2004, p. 28).

Segundo esses autores, a informática nasceu do ramo da lógica matemática (matemática discreta) e se desenvolveu com o crescimento da tecnologia eletrônica, mais especificamente com a microeletrônica, ramo da engenharia em sua época e passou a ser a ciência da informação, como já é chamada por Engenheiros e Matemáticos. Assim, Bisquerra, Sarriera e Martinez (2004) concluem que a informática é um conjunto de atividades, de procedimentos, de técnicas e de ferramentas, ajudadas pelo computador, que tratam a informação.

Os primeiros computadores (calculadoras programáveis capazes de armazenar os programas) surgiram na Inglaterra e nos Estados Unidos em 1945. No Brasil, o movimento da informática aplicada à educação tem sua origem na década de 1970, no setor administrativo das escolas, visando informatizar as secretarias das Escolas.

O termo "Informática na educação refere-se à inserção do computador no processo de aprendizagem dos conteúdos curriculares de todos os níveis e modalidades de educação". Como se vê, o computador é um recurso que pode indicar múltiplos caminhos no processo de ensino-aprendizagem. Mas, para que isso aconteça, o professor deve assumir o papel de mediador dos conhecimentos e utilizar esses recursos de forma potencial em sua prática. O computador já faz parte do cotidiano

escolar e, atualmente, representa para o trabalho docente um recurso indispensável, já que oferece um fazer educativo problematizador, ativo e protagonizante (GOMES; MOITA, 2016, p. 156).

Para Kenski (2010), um dos grandes desafios da docência do século XXI consiste em encontrar a melhor maneira de utilizar o computador e conseqüentemente a tecnologia no processo de ensino e aprendizagem, que possibilite a reconfiguração do papel do professor e do aluno neste novo cenário, proporcionando-lhes uma formação mais adequada à realidade atual. Mas, segundo a autora, a simples inserção tecnológica no Ensino da Estatística por si só não basta para a real promoção da aprendizagem.

Sallum, Cavalari Júnior e Schimiguel (2011) dizem que as Tecnologias se tornaram um elemento primordial na evolução da relação ensino e aprendizagem, oferecendo a esta relação novas possibilidades para a educação.

Para os autores, os softwares podem ser criados com finalidades educacionais, comerciais ou atendendo a ambas, como é o caso do Microsoft Excel. Considerado como uma planilha eletrônica, o Excel reúne diversas ferramentas como gráficos e tabelas dinâmicas, bem como possibilita o uso de fórmulas matemáticas, lógicas, financeiras e estatísticas. É, em suma, a planilha eletrônica mais conhecida em todo o mundo, mas que, entretanto, poucos sabem utilizar eficientemente como uma ferramenta auxiliar no tratamento estatístico de dados e construção de gráficos estatísticos.

Segundo Silva e Schimiguel (2013), o uso de planilhas eletrônicas como o Excel pode ser de grande utilidade para ampliar as aplicações da estatística, assim como facilita o aprendizado estatístico, como é o caso desta pesquisa. Os autores, ainda ressaltam que, por se tratarem de softwares populares e de fácil acesso, as planilhas eletrônicas possibilitam aos usuários uma maior interação com as funções elementares da Matemática e da Estatística do que os softwares estatísticos em geral.

Dessa forma, a assimilação dos conceitos estatísticos também é beneficiada pela possibilidade de simulação, que permite alterar os dados e analisar o novo comportamento destes ao observar as alterações nas tabelas de distribuição de frequência ou em gráficos gerados, comparando o antes e o depois.

3.2 O SOFTWARE EXCEL

De forma breve vamos conhecer um pouco sobre o Software Excel que é uma das peças fundamentais do presente trabalho. O Microsoft Office Excel é desenvolvido pela Microsoft desde 1987 e nada mais é do que um editor de planilhas para computadores que utilizam o sistema operacional Microsoft Windows, além de computadores Macintosh da Apple Inc. e dispositivos móveis como o Windows Phone, Android ou o iOS.

Os recursos do Excel incluem uma interface intuitiva e capacidades ferramentas de cálculo e de construção de gráficos que, juntamente com marketing agressivo, o tornaram um dos mais populares aplicativos de computador até hoje.

Segundo Miranda (2010), o Excel é muito utilizado para cálculos, estatísticas, gráficos, relatórios, formulários e entre outros requisitos das rotinas empresariais, econômicas, administrativas e domésticas.

A possibilidade de usar fórmulas é o que diferencia um programa de planilha de uma calculadora. Quando é colocado uma fórmula em uma célula, é dito que o conteúdo dessa célula deve ser calculado em função dos valores contidos em outras células, podendo ser transferidos e interligados com outras planilhas salvas como arquivos. (MIRANDA, 2010, p. 27).

Ainda de acordo com esse autor, funções de uma planilha são comandos compactos e rápidos para executar as fórmulas, se mostrando um meio descomplicado de efetuar operações complexas. Miranda (2010) cita as principais funções do Excel que são:

- A. Funções financeiras - Para calcular juros, rendimento de aplicações, depreciação de ativos etc.
- B. Funções matemáticas e trigonométricas - Permitem calcular raiz quadrada, fatorial, seno, tangente etc.
- C. Funções estatísticas - Para calcular a média de valores, valores máximos e mínimos de uma lista, desvio padrão, correlação, regressão distribuições etc.
- D. Funções lógicas - Possibilitam comparar células e apresentar valores que não podem ser calculados com fórmulas tradicionais.

O Excel foi o primeiro programa de seu tipo a permitir ao usuário definir a aparência das planilhas (fontes, atributos de caracteres e aparência das células). Também, introduziu recomputação inteligente de células, na qual apenas células dependentes da célula a ser modificada são atualizadas (programas anteriores recomputavam tudo o tempo todo ou aguardavam um comando específico do usuário). O Excel tem capacidades avançadas de construção de gráficos onde os mais utilizados são os de Pizza, Dispersão, Linhas, Colunas dentre outros.

De forma resumida conhecemos um pouco mais desse software amplamente utilizado pelo mundo nos mais diversos setores, seja de forma comercial como por exemplo para organizar as folhas de pagamento de uma determinada empresa ou usado de forma acadêmica para organizar as notas e calcular as médias de determinados alunos de uma turma de uma faculdade qualquer, como também para servir como estudo para o presente trabalho.

3.3 REGRESSÃO E CORRELAÇÃO

Martins e Donaire (2006, p. 203) dizem que, um dos maiores problemas para o investigador de fenômenos humanos ou físicos é “o estabelecimento de um modelo matemático que descreva e explique o fenômeno ocorrido na vida real, com boa aproximação”.

Para os autores, a busca de uma relação funcional entre as variáveis observadas que descrevem o fato é uma tarefa de muitos cientistas em qualquer área de estudo. Por exemplo,

O pediatra tem interesse em estabelecer uma relação funcional entre o peso e a altura dos bebês; o economista busca o estabelecimento de uma função que explique o comportamento das vendas, em unidades de um produto, em função do preço; o administrador precisa de uma função que descreva os custos de um produto, quando as quantidades variam; o engenheiro quer saber a relação funcional entre a resistência do concreto e a razão água/cimento; o médico tem interesse em relacionar através de uma função o volume do plasma sanguíneo e a superfície dos corpos dos pacientes; o psicólogo deseja a função que explique o QI (quociente de inteligência) etc. (MARTINS; DONAIRE, 2006, p. 203).

Entretanto, se for possível obter os dados, existem procedimentos estatísticos que podem ser usados para demonstrar como as variáveis se relacionam.

Nesse sentido Loesch (2014, p. 168) diz que, “Regressão é uma técnica estatística que permite inferir a relação de uma variável dependente (variável

resposta) com variáveis independentes específicas (variáveis explicatórias)". Ainda de acordo com o autor, a análise de Regressão pode ser usada como um método descritivo da análise de dados (por exemplo, o ajustamento de curvas) sem serem necessárias quaisquer suposições acerca dos processos que permitiram gerar os dados.

Já a correlação, segundo Loesch (2014), é medida através de um coeficiente de correlação, se refere à medida da relação entre duas variáveis e a direção do relacionamento, embora a correlação não implique causalidade. Assim, existem vários coeficientes que medem o grau de correlação, adaptados à natureza dos dados, dentre os quais o coeficiente de correlação de Pearson é o mais conhecido.

- Regressão: responde à pergunta de como variáveis quantitativas se relacionam, ou seja, qual é a equação de relacionamento entre as variáveis;
- Correlação: responde à pergunta em que grau as variáveis quantitativas se relacionam.

Regressão

Segundo Sweeney, Williams e Anderson (2015, p. 498), “na terminologia da análise de regressão, a variável que é prevista é dita variável dependente”. Já a variável ou variáveis usadas para prever o valor da variável dependente são denominadas variáveis independentes.

Diz-se regressão de Y em (sobre) X. Se a relação funcional entre elas é expressa por uma equação do 1º grau, cuja representação geométrica é uma linha reta, a regressão é dita linear.

Modelo de regressão e equação de regressão linear simples

A equação que descreve como y está relacionado com x e com uma parcela de erro é chamada modelo de regressão. O modelo de regressão usado na regressão linear simples é o seguinte:

$$y = \beta_0 + \beta_1 x + \epsilon$$

Onde β_0 e β_1 se referem aos parâmetros do modelo, e ϵ (a letra grega épsilon) é uma variável aleatória que se denomina erro aleatório ou parcela de erro. A

parcela de erro é responsável pela variabilidade em y que não pode ser explicada pela relação linear entre x e y .

Assim, cada distribuição dos valores de y tem sua própria média ou valor esperado. A equação que descreve como o valor esperado de y – designado por $E(y)$ – está relacionado com x é chamada equação de regressão. A equação de regressão para regressão linear simples é a seguinte:

$$y = \beta_0 + \beta_1 + \epsilon$$

O gráfico da equação de regressão linear simples é uma linha reta; β_0 é o ponto onde a linha (ou reta) de regressão intercepta o eixo y , β_1 é a inclinação. $E(y)$ é a média ou valor esperado de y para determinado valor de x .

Exemplos de possíveis retas de regressão são mostrados abaixo.

Gráfico A: Relação Linear Positiva

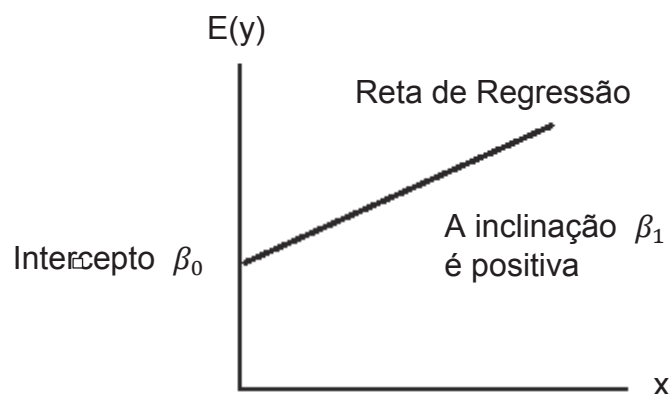


Gráfico B: Relação Linear Negativa

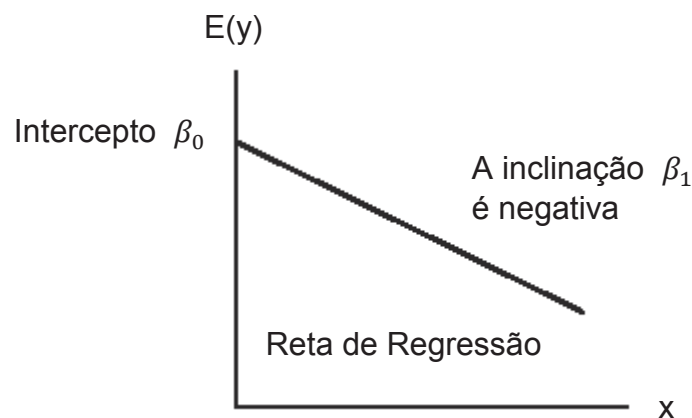
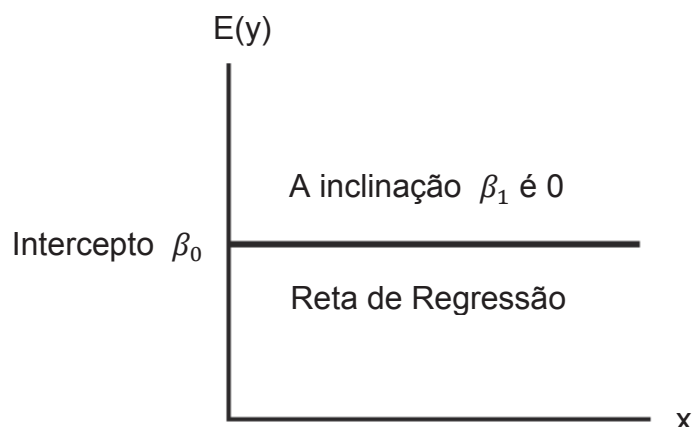


Gráfico C: Sem Relação



A reta de regressão do gráfico A mostra que o valor médio de y está relacionado positivamente com x , e valores maiores de $E(y)$ estão associados a valores maiores de x . A reta de regressão do gráfico B mostra que o valor médio de y está relacionado negativamente com x , e valores menores de $E(y)$ estão associados a valores maiores de x . A reta de regressão do gráfico C apresenta o caso em que o valor médio de y não está relacionado com x ; ou seja, o valor médio de y é o mesmo para todo valor de x (SWEENEY; WILLIAMS; ANDERSON, 2015).

Equação de regressão estimada

Se os valores dos parâmetros β_0 e β_1 fossem conhecidos, poderíamos usar a equação de regressão linear simples para calcular o valor médio de y para dado valor x . Na prática, os valores dos parâmetros não são conhecidos e precisam ser estimados utilizando-se os dados amostrais. As estatísticas amostrais (designadas por b_0 e b_1) são calculadas como estimativas dos parâmetros populacionais β_0 e β_1 . Substituindo os valores das estatísticas amostrais b_0 e b_1 no lugar de β_0 e β_1 na equação da regressão, obtemos a equação de regressão estimada. A equação de regressão linear simples estimada é a seguinte:

$$\hat{y} = b_0 + b_1X$$

O gráfico da equação de regressão linear simples estimada é denominado reta de regressão estimada; b_0 é o ponto de intersecção com o eixo y e b_1 é a

inclinação. Em geral, \hat{y} é o estimador pontual de $E(y)$, o valor médio de y para um dado valor de x .

Assim, uma vez que o valor de \hat{y} fornece uma estimativa pontual de $E(y)$ para um dado valor de x , e uma previsão do valor de y para um valor de x dado, iremos chamá-lo simplesmente de valor estimado de y .

Nesse sentido, a análise de regressão não pode ser interpretada como um procedimento para estabelecer uma relação de causa e efeito entre as variáveis. Ela somente é capaz de indicar como ou em que graus as variáveis estão associadas entre si. Quaisquer conclusões sobre causa e efeito devem basear-se no julgamento das pessoas que têm o melhor conhecimento da aplicação.

Método dos mínimos quadrados

É um procedimento que usa dados amostrais para encontrar a equação de regressão estimada.

O método de mínimo quadrados fornece uma equação de regressão estimada que minimiza a soma de desvios quadráticos entre os valores estimados da variável dependente y_i e os valores estimados da variável dependente \hat{y}_i . Além disso, é usado para escolher a equação que fornece o melhor ajuste. Se algum outro critério fosse utilizado por exemplo, minimizar a soma dos desvios absolutos entre y_i e \hat{y}_i uma equação diferente seria obtida. Na prática, o método dos mínimos quadrados é amplamente o mais usado (SWEENEY; WILLIAMS; ANDERSON, 2015).

Dessa forma, esse método utiliza dados amostrais para produzir os valores b_0 e b_1 que minimizam a soma dos quadrados dos desvios entre os valores observados da variável dependente y_i e os valores estimados da variável dependente \hat{y}_i . O critério utilizado no método dos mínimos quadrados é dado pela Equação,

$$\text{Min } \sum (y_i - \hat{y}_i)^2$$

Em que

y_i = valor observado da variável dependente para a i -ésima observação

\hat{y}_i = valor estimado da variável dependente para a i -ésima observação

Inclinação e intersecção da equação de regressão estimada

$$b_1 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2}$$

$$b_0 = \bar{Y} - b_1\bar{X}$$

x_i = valor da variável independente para a i -ésima observação

y_i = valor da variável dependente para a i -ésima observação

\bar{x} = valor médio da variável independente

\bar{y} = valor médio da variável dependente

n = número total de observações

Coeficiente de determinação

O coeficiente de determinação nos dá uma medida da qualidade de ajuste da equação da regressão estimada.

Em relação a i -ésima observação, a diferença entre o valor observado da variável dependente, y_i e o valor estimado da variável dependente, \hat{y}_i , é denominada i -ésimo resíduo. O i -ésimo resíduo representa o erro de utilizarmos \hat{y}_i para estimar y_i . Dessa forma, para a i -ésima observação, o resíduo é $y_i - \hat{y}_i$. A soma dos quadrados desses resíduos ou erros é a quantidade que é minimizada pelo método dos mínimos quadrados. Essa quantidade, também conhecida como a soma dos quadrados dos resíduos, é designada por SQRes.

$$\text{SQRes} = \sum(y_i - \hat{y}_i)^2$$

O valor da SQRes é uma medida do erro de usar a equação de regressão estimada para calcular os valores da variável dependente na amostra.

Soma dos quadrados total

A soma dos quadrados correspondente é denominada soma dos quadrados total, designada por SQTot.

$$\text{SQTot} = \sum(y_i - \bar{y})^2$$

Podemos imaginar a SQTot como uma medida de quão adequadamente as observações se agrupam nas proximidades na reta \hat{y} .

Para medir quanto os valores na reta de regressão estimada se afastam de \bar{y} , outra soma de quadrados é calculada. Essa soma de quadrados, denominada soma dos quadrados da regressão.

Soma dos quadrados da regressão

$$SQReg = \sum(\hat{y}_i - \bar{y})^2$$

Relação entre SQTot, SQReg e SQRes

A partir dessas somas, devemos esperar que SQTot, SQReg e SQRes estejam relacionadas. De fato, a relação entre essas somas de quadrados produz um dos resultados mais importantes da estatística.

$$SQTot = SQReg + SQRes$$

em que

SQTot = soma dos quadrados total

SQReg = soma dos quadrados de regressão

SQRes = Soma dos quadrados dos resíduos

A equação mostra que a soma dos quadrados total pode ser dividida em dois componentes: a soma dos quadrados da regressão e a soma dos quadrados dos resíduos. Portanto, se os valores de quaisquer dessas somas de quadrados de regressão forem conhecidos, a terceira soma de quadrados poderá ser calculada facilmente.

Essas três somas de quadrados podem ser usadas para fornecer uma medida da qualidade de ajuste da equação de regressão estimada. A equação de regressão estimada forneceria um ajuste perfeito se todo valor da variável dependente y_i se situasse na reta de regressão estimada. Nesse caso, $y_i - \hat{y}_i$ seria igual a zero para cada observação, resultando em $SQRes = 0$. Uma vez que $SQTot = SQReg + SQRes$, notamos que para haver um ajuste perfeito $SQReg$ deve igualar-se a $SQTot$, e a razão ($SQReg / SQTot$) deve ser igual a 1. Ajustes inadequados resultarão em valores maiores para $SQRes$. Resolvendo $SQRes$ na equação, notamos que $SQRes = SQTot - SQReg$. Portanto, o maior valor para $SQRes$ (e, daí, o pior ajuste) ocorre quando $SQReg = 0$ e $SQRes = SQTot$.

O quociente SQ_{Reg}/SQ_{Tot} , o qual assumirá valores entre zero e um, é usado para avaliar a qualidade de ajuste da equação de regressão estimada. Esse quociente é chamado de coeficiente de determinação e é designado por r^2 (SWEENEY; WILLIAMS; ANDERSON, 2015).

Coeficiente de determinação

$$r^2 = \frac{SQ_{Reg}}{SQ_{Tot}}$$

Quando expressamos o coeficiente de determinação como uma porcentagem, r^2 pode ser interpretado como a porcentagem da soma dos quadrados total que pode ser explicada usando-se a equação de regressão estimada.

Suposições do modelo

Ao realizarmos uma análise de regressão, começamos fazendo uma suposição acerca do modelo apropriado para a relação entre a variável dependente e a(s) independente(s). Para o caso de uma regressão linear simples, o modelo de regressão suposto é:

$$Y = \beta_0 + \beta_1 X + \epsilon$$

Então, o método dos mínimos quadrados é usado para obter valores para b_0 e b_1 , os quais são as estimativas dos parâmetros do modelo β_0 e β_1 , respectivamente. A equação de regressão estimada é:

$$\hat{y} = b_0 + b_1 X$$

Vimos que o valor do coeficiente de determinação (r^2) é uma medida da qualidade de ajuste da equação de regressão estimada. Entretanto, mesmo com um valor grande r^2 , a equação de regressão estimada não deve ser usada enquanto não se fizer uma análise adicional da adequabilidade do modelo suposto. Uma etapa importante para determinar se o modelo suposto é apropriado envolve testar a significância da relação. Os testes de significância na análise de regressão baseiam-se nas seguintes suposições sobre a parcela de erro ϵ .

Correlação linear

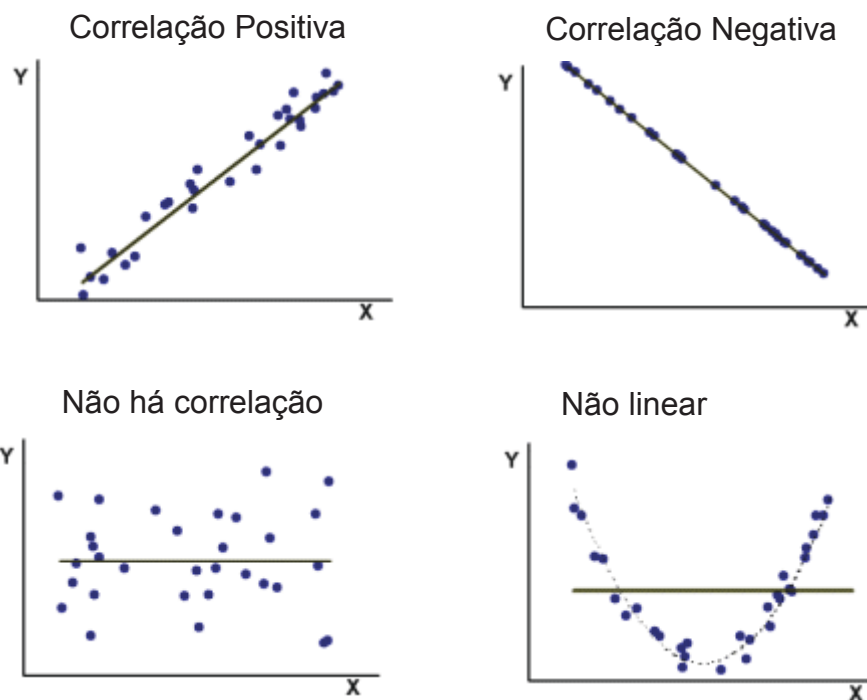
Sendo a relação entre as variáveis de natureza quantitativa, a correlação é o instrumento adequado para descobrir e medir essa relação.

Os pontos obtidos, vistos em conjunto, formam em elipse em diagonal. Podemos imaginar que, quanto mais fina for a elipse, mais ela se aproximará de uma reta. Dizemos então, que a correlação de forma elíptica tem como “imagem” uma reta, sendo, por isso denominada de correlação linear.

É possível verificar que cada correlação está associada como “imagem” uma relação funcional. Por esse motivo, as relações funcionais são chamadas relações perfeitas.

A correlação entre duas variáveis pode ter vários graus, dependendo da maior ou menor dependência entre elas. Se a variável y for, totalmente dependente da variável x e, não houver nenhum outro fato que interfira entre elas, teremos então, uma correlação perfeita (CRESPO, 2009).

No caso, porém, de existirem outras variáveis que interfiram na relação x , y , a correlação vai diminuindo, podendo até deixar de existir.



Assim, uma correlação é:

- Linear positiva se os pontos do diagrama têm como “imagem” uma reta ascendente;
- Linear negativa se os pontos têm como “imagem” uma reta descendente;
- Não há correlação se os pontos apresentam-se dispersos, não oferecendo uma “imagem” definida.
- Não linear se os pontos têm como “imagem” uma curva.

Coefficiente de correlação

O instrumento empregado para a medida da correlação linear é o coeficiente de correlação. Esse coeficiente deve indicar o grau de intensidade da correlação.

$$r = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{[n \sum x_i^2 - (\sum x_i)^2][n \sum y_i^2 - (\sum y_i)^2]}}$$

Os valores do coeficiente de correlação estão sempre entre -1 e +1. Um valor +1 indica que as duas variáveis x e y estão perfeitamente relacionadas em um sentido linear positivo. Ou seja, todos os pontos de dados estão em uma linha reta que tem uma inclinação positiva. Um valor -1 indica que x e y estão perfeitamente relacionadas em um sentido linear negativo, com todos os pontos de dados em uma linha reta que tem uma inclinação negativa. Valores do coeficiente de correlação próximos de zero indicam que x e y não estão linearmente relacionados.

3.4 APLICANDO O EXCEL

Ilustraremos como a ferramenta Regressão do Excel pode ser usada para realizar os cálculos de análise de regressão do problema dos restaurantes Armand’s Pizza Parlors retirado do livro Estatística Aplicada à Administração e Economia de Sweeney, Williams e Anderson, 2015.

A Armand’s Pizza Parlors é uma rede de restaurantes de comida italiana localizada em cinco Estados norte-americanos. As localizações mais bem-sucedidas dos restaurantes estão próximas a campus universitários. Os gestores acreditam que as vendas trimestrais nesses restaurantes (designadas y) estão

relacionadas positivamente com o tamanho da população estudantil (designado x), ou seja, os restaurantes próximos aos campus universitários que contam com uma grande população estudantil tendem a gerar mais vendas que aqueles que estão localizados próximos a campus que contam com uma pequena população estudantil. Usando a análise de regressão, podemos determinar uma equação que mostra como a variável dependente y está relacionada com a variável independente x .

Dados sobre a população de estudantes e as vendas trimestrais em restaurantes Armand's Pizza Parlors		
Restaurante	População de Estudantes (em milhares)	Vendas Trimestrais (em milhares de dólares)
I	x_i	y_i
1	2	58
2	6	105
3	8	88
4	8	118
5	12	117
6	16	137
7	20	157
8	20	169
9	22	149
10	26	202

A figura abaixo mostra os dados e o resumo dos resultados que descrevemos nas etapas envolvidas.

Os Rótulos Restaurante, População e Vendas são inseridos nas células A1:C1 da planilha. Para identificar cada uma das dez observações, inserimos os números 1 a 10 nas células A2:A11.

	A	B	C	D	E	F	G	H	I
1	RESTAURANTE	POPULAÇÃO	VENDAS						
2	1	2	58						
3	2	6	105						
4	3	8	88						
5	4	8	118						
6	5	12	117						
7	6	16	137						
8	7	20	157						
9	8	20	169						
10	9	22	149						
11	10	26	202						
12									
13	RESUMO DOS RESULTADOS								
14									
15	<i>Estatística de regressão</i>								
16	R múltiplo	0,95012296							
17	R-Quadrado	0,90273363							
18	R-quadrado aju	0,89057533							
19	Erro padrão	13,8293167							
20	Observações	10							
21									
22	ANOVA								
23		<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>			
24	Regressão	1	14200	14200	74,24837	2,54887E-05			
25	Resíduo	8	1530	191,25					
26	Total	9	15730						
27									
28		<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 99,0%</i>	<i>Superior 99,0%</i>
29	Interseção	60	9,2260348	6,5033355	0,000187	38,72472558	81,27527442	29,04307968	90,95692032
30	POPULAÇÃO	5	0,5802652	8,6167492	2,55E-05	3,661905962	6,338094038	3,052985371	6,947014629
31									

Os dados amostrais são inseridos nas células B2:C11. As etapas a seguir descrevem como usar o Excel para produzir os resultados de regressão.

Etapas 1. Selecione o menu **Ferramentas**

Etapas 2. Escolha a opção **Análise de Dados**

Etapas 3. Escolha **Regressão** na lista de ferramentas de Análise

Etapas 4. Dê um clique em **OK**

Etapas 5. Quando a caixa de diálogo Regressão aparecer:

Digite C1:C11 na caixa **Intervalo Y de Entrada**

Digite B1:B11 na caixa **Intervalo X de Entrada**

Selecione **Rótulos**

Selecione **Nível de Confiança**

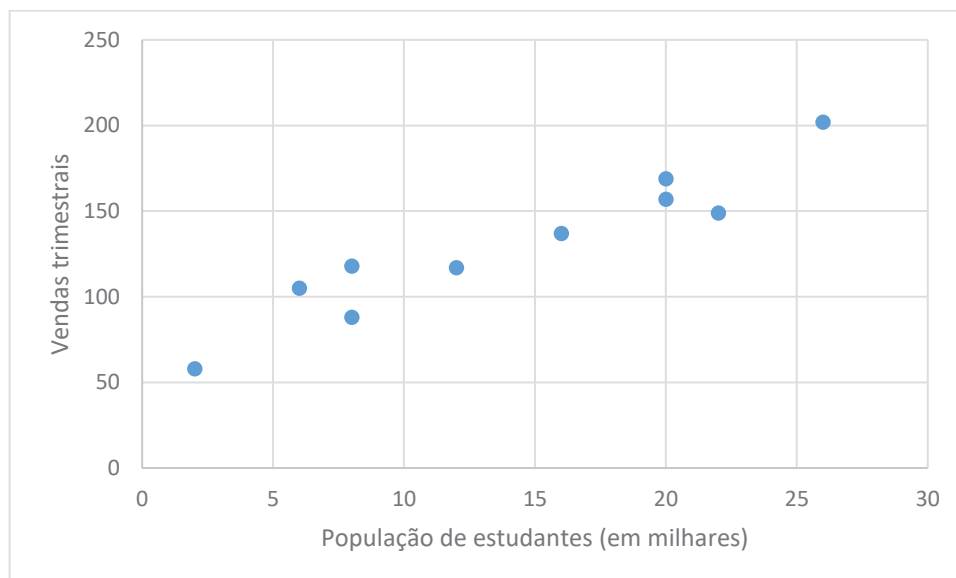
Digite 99 na caixa **Nível de Confiança**

Selecione **Intervalo de Saída**

Digite A13 na caixa **Intervalo de Saída**

(Qualquer célula do canto superior esquerdo que indique onde a saída deve ser iniciada pode ser inserida aqui.)

Dê um clique em **OK**



A primeira seção de saída de dados, intitulada *Estatísticas da Regressão*, apresenta um resumo estatístico, por exemplo, o coeficiente de determinação (R-Quadrado). A segunda seção da saída, intitulada ANOVA, contém a tabela de análise de variância. A última seção da saída, a qual não tem título, contém o coeficiente de regressão estimado e as informações correspondentes. Iniciaremos nossa discussão da interpretação da saída de regressão com informação contida nas células A28:I30.

Interpretação da Saída de Dados da Equação de Regressão Estimada

O ponto em que a reta de regressão estimada intercepta o eixo y, $b_0 = 60$, é mostrado na célula B29, e a inclinação da reta de regressão estimada $b_1 = 5$ é mostrada na célula B30. O rótulo Intercepto na célula A29 e o rótulo População na célula A30 são usados para identificar esses dois valores.

A informação nas células F28:I30 pode ser usada para desenvolver estimações por intervalo de confiança da intercepção com o eixo y e a inclinação da equação de regressão estimada. O Excel sempre apresenta os limites mínimo e máximo de um intervalo de confiança 95%. Na etapa 4, selecionamos Nível de Confiança e inserimos 99 na caixa Nível de Confiança. Em consequência, a ferramenta Regressão do Excel também fornece os limites mínimo e máximo de um intervalo de confiança de 99%.

Interpretação da Saída de Dados ANOVA

A informação nas células A22:F26 é um resumo dos cálculos da análise de variância. As três fontes de variação são rotuladas de Regressão, Resíduo e Total. O rótulo *gl* na célula B23 refere-se a “graus de liberdade”, o rótulo *SQ* Na célula C23 corresponde à soma dos quadrados e o rótulo *MQ* na célula D23 refere-se à medida quadrática.

Interpretação da Saída de Dados da Estatística de Regressão

O coeficiente de determinação, 0,9027, aparece na célula B:17; o rótulo correspondente, R-Quadrado, é indicado na célula A17. A raiz quadrada do coeficiente de determinação fornece o coeficiente de correlação amostral 0,9501, mostrado na célula B16. O Excel usa o rótulo R-Múltipla (célula A16) para identificar esse valor. Na célula A19, o rótulo Erro Padrão é usado para identificar o valor do erro padrão da estimativa exposta na célula B19. Desse modo, o desvio padrão da estimativa é 13, 8293. É necessário ter em mente que, na saída do Excel, o rótulo Erro padrão aparece em dois lugares diferentes.

Pode-se fazer muito mais aplicações estatísticas com o Excel, no qual só foi apresentado apenas um exemplo para o auxílio no ensino de Regressão e Correlação. Todo o conhecimento nesta área da Estatística, por meio das tecnologias precisa de uma certa dedução lógica e este tipo de metodologia leva o aluno a raciocinar. Nesse sentido, o uso de tecnologias, a exemplo do software Excel, pode surgir como um método inovador, para auxiliar na aprendizagem dos alunos, pois se trata de um software de fácil utilização, instalação e acesso, o que tende a ser um facilitador no ensino da Estatística.

4 CONSIDERAÇÕES FINAIS

Este trabalho buscou ressaltar e salientar a importância de se usar tecnologia na sala de aula, dentre estas tecnologias podemos destacar os softwares educativos, como o Excel. Este software pode auxiliar o professor em sala de aula, a ensinar a maioria dos conteúdos de Estatística, fazendo assim, com que os alunos aprendam cada vez mais.

O uso do software facilita a explicação desses, de modo que os professores possam adaptar conteúdos de Regressão e Correlação, estabelecendo uma melhor interação entre eles.

Constatamos que a construção de gráficos no software Excel, pode ajudar no ensino e na aprendizagem da Estatística na sala de aula. Assim, o foco deste trabalho foi apresentar o Excel como uma alternativa para melhorar a aprendizagem dos alunos e ensinar Regressão e Correlação. Sempre visando a análise de Regressão e Correlação por meio da observação e da construção gráfica, realizados com este software, onde foi introduzido uma aplicação para ficar fácil o entendimento e então permitir aos alunos enxergar o gráfico como um conjunto de variáveis.

O Excel se trata de apenas um dos instrumentos que podem ser utilizados pelos professores e vem a estimular o interesse dos profissionais de educação pela utilização de novas técnicas, como da tecnologia a favor da educação em âmbito escolar. O uso do software surge como método inovador e dinâmico, de fácil utilização que auxilia no entendimento dos conteúdos pelos alunos.

Destacamos, também, que os softwares podem proporcionar aos alunos uma maior integração com a Estatística e com o a Regressão e Correlação, proporcionando uma maior consistência com as análises, os gráficos e suas respectivas fórmulas. Inúmeras são as possibilidades do professor em sala de aula, quando se tem como auxílio à ferramenta Excel, que é um passo no desenvolvimento da inteligência que é alcançada pela exploração das capacidades visuais e interpretativa dos alunos. Para realização de atividades o software Excel, pode ser uma grande ferramenta, onde os alunos podem articular a passagem de uma representação gráfica para realização da análise e conclusão das informações.

Dessa forma, é necessário o aperfeiçoamento dos profissionais de educação quanto à utilização desse tipo de programa, pois eles serão os responsáveis por apresentar o software aos alunos e explanarem as suas funções. Portanto, caso não haja um pleno entendimento do docente a respeito do Excel, esse se tornará mais um obstáculo para a aprendizagem dos alunos em relação aos conteúdos de Regressão e Correlação. Vale concluir que o conhecimento estatístico é fundamental para o desenvolvimento e formação do cidadão para a sociedade em que vivemos. E o software Excel serve para melhorar o aprendizado dos alunos nas etapas de sua formação e que possa contribuir para esclarecimentos de alguns conteúdos estatísticos, devendo o professor sempre buscar formas mais eficazes de transmitir o conhecimento.

REFERÊNCIAS BIBLIOGRÁFICAS

BISQUERRA, R.; SARRIERA, J. C.; MARTINEZ, F. Introdução à Estatística: Enfoque informático com o pacote estatístico SPSS. Tradução de Fátima Murad. Porto Alegre: Artmed, 2004.

BOTTER, D. A.; PAULA, G.A.; LEITE, J. G.; CORDANI, L. K. Noções de Estatística: com apoio computacional. Versão preliminar, editora: Instituto de Matemática e Estatística – USP. São Paulo, 1996.

BRASIL. Parâmetros Curriculares Nacionais: Matemática – 3º e 4º ciclos. Brasília: MEC, 1998. 148p.

_____. Ministério da Educação – Secretaria da Educação Básica. Orientações curriculares para o Ensino Médio: Ciências da natureza, matemática e suas tecnologias. Brasília, 2006. v.2, p.69-98.

_____. Ministério da Educação. Conselho Nacional de Educação. Resolução. Parecer CNE/CP 09/2001. Institui as diretrizes curriculares nacionais para a formação de professores da educação básica, em cursos de nível superior, curso de licenciatura, de graduação plena. Brasília: MEC, 2002.

_____. Secretaria de Educação Média e Tecnológica. PCN+ Ensino Médio: Orientações Educacionais complementares aos Parâmetros Curriculares Nacionais. Ciências da Natureza, Matemática e suas tecnologias. Brasília: MEC ; SEMTEC, 2002.

CAMPOS, C. R.; WODEWOTZKI, M. L. L.; JACOBINI, O. R. Educação Estatística: teoria e prática em ambientes de modelagem matemática. Belo Horizonte: Autêntica Editora, 2011.

CRESPO, A. A. Estatística Fácil. 19 ed. São Paulo: Saraiva, 2009.

ESTEVAM, J. G. E; KALINKE, M. A. Recursos Tecnológicos e Ensino de Estatística na Educação Básica: um cenário de pesquisas brasileiras. Revista Brasileira de Informática na Educação, v. 21, n. 2, p. 104-117, 2013.

GOMES, L. L.; MOITA, F. M. G. S. C. O uso do laboratório de informática educacional: partilhando vivências do cotidiano escolar. In: Robson Pequeno de Sousa; Carolina Cavalcanti Bezerra; Eliane de Moura Silva, et al. (Organizadores). Teorias e Práticas em Tecnologias Educacionais. Campina Grande: Eduepb, 2016.

KENSKI, V. M. Aprendizagem Mediada pela Tecnologia. Revista Diálogo Educacional, Curitiba, 4(10): 47-56, 2003.

LOESCH, C. Probabilidade e Estatística. Rio de Janeiro: LCT 2014.

MARTINS, G. A.; D, D. Princípios de Estatística. 4 ed. São Paulo: Atlas, 2006.

MEMÓRIA, J. M. P. Breve história da estatística. Brasília, DF : Embrapa Informação Tecnológica, 2004,111p.

MIRANDA, E. G. de S. Uma aplicação prática em Excel na análise de projetos de viabilidade econômica. Trabalho de Conclusão de Curso (Graduação em Bacharelado em Ciências Econômicas) - Universidade Federal de Santa Catarina, Florianópolis, 2010.

PONTE, J. P.; BROCARD, J.; OLIVEIRA, H. Investigações matemáticas na sala de aula. Belo Horizonte: Editora Autêntica, 2013, 152p.

PRENSKY, M. Digital natives, digital immigrants. On the Horizon, NCB University Press, v. 9, n. 5, 2001. Disponível em: <<http://www.marcpresnky.com/writing/default.asp>>. Acesso em: 6 nov. 2016.

SALLUM, W. G.; CAVALARI, J. O.; SCHIMIGUEL, J. Concepções de objetos de aprendizagem na matemática: de Jean Piaget a David Wiley. In: LOPES, C. E.; ALLEVATO, N. S. G. Matemática e tecnologias. São Paulo: Terracota, 2011. p. 107125.

SILVA, J. F.; SCHIMIGUEL, J. O uso das TICs no ensino superior: a integração de diferentes tecnologias à educação estatística. Revista de Produção Discente em Educação Matemática, São Paulo, v.2, n.1, p. 51-60, 2013. Disponível em: <<http://revistas.pucsp.br/index.php/pdemat/index>>. Acesso em: 5 nov. 2016.

SWEENEY, D. J; WILLIAMS, T. A.; ANDERSON, D. R. Estatística Aplicada à Administração e Economia. 3 ed. Cengage Learning, São Paulo, 2015.