



RELAÇÃO ENTRE AVALIAÇÕES ATRAVÉS DA INFERÊNCIA ESTATÍSTICA NO PROCESSO DE ENSINO E APRENDIZAGEM

Luz Delicia Castillo Villalobos (UTFPR) luz_delicia@yahoo.com.br
Milagros Noemi Quintana (UDESC) milla.quicas@gmail.com

Resumo

O objetivo da pesquisa é testar se as primeiras avaliações que o professor aplica em seus alunos tem alguma relação com a última avaliação, isto é, se esta seria uma resultante das anteriores, de modo que serviria como referência para verificar se o processo de ensino e aprendizagem está sob controle. Como consequência desta análise, o professor poderá tomar decisões sobre o método de ensino aplicado numa matéria específica em uma determinada turma e se é possível manter o método de ensino em turmas futuras. Para atender satisfatoriamente a estas questões, se faz necessário acudir à inferência estatística através da utilização de um modelo de regressão múltipla e dos gráficos de controle de resíduos. Para aplicar o método, foi selecionada uma turma de Cálculo 1 da universidade, pertencente ao primeiro período letivo. Se obteve como resultado, com modelo de regressão múltipla, a existência de relação estatisticamente significativa entre as primeiras e a última avaliação. Nos gráficos de controle se obteve que para um desvio padrão temos 13 alunos fora de controle e para dois desvios padrões a variável resultante está sob controle em toda a turma.

Palavras-Chaves: Ensino e aprendizagem; Avaliação; Regressão; Gráficos de controle.

1. Introdução

O objetivo do presente trabalho é apresentar uma metodologia que ajude ao professor de alguma forma a verificar se o método que está aplicando em sala de aula está sendo aproveitado pelos seus alunos. Para tanto, ele precisa aplicar avaliações periódicas na turma de modo que possa medir quanto eles assimilaram o conhecimento passado em sala de aula.

Para que o processo de ensino e aprendizagem, bem como o de avaliação, seja eficaz, deve-se levar em consideração a inter-relação entre as diferentes avaliações que o professor aplica, de modo a corrigir e prevenir situações extremas de reprovação no final do processo de ensino da matéria correspondente.



O presente trabalho pretende passar uma metodologia que permita visualizar em forma quantitativa esta inter-relação. Para isso serão utilizados métodos estatísticos, cujo objetivo é conseguir uma análise consistente dos dados que serão utilizados no processo de estudo. Segundo Samá e Fonseca (2019), a inter-relação entre os conceitos estatísticos auxilia na escolha das medidas estatísticas utilizadas para resumir a informação sobre os dados coletados, bem como destaca a complexidade das possibilidades de interpretação sobre a temática investigada no Projeto de Aprendizagem e o conhecimento que o aluno pode alcançar.

As inferências e decisões que se fazem a partir de um conjunto de dados são cada vez melhores à medida que se tem uma descrição clara destes. É necessário, portanto, ter índices apropriados, significativos e práticos dos dados que se dispõe (VELIZ, 1993).

Essas abordagens favorecem o desenvolvimento de competências para a tomada de decisões em situações incertas, as quais demandam compreensão, explicação e quantificação sobre a variabilidade nos dados (ESPASANDIN *et al.*, 2018). No presente trabalho será utilizado como indicadores as notas obtidas nas avaliações em sala de aula. Nas matérias de matemática geralmente as avaliações são a base de provas escritas cujas perguntas estão relacionadas com conteúdo lecionado em sala de aula. A avaliação torna-se um dos componentes das estratégias de ensino para compreender se o aluno sabe ou não sabe, o que possibilita realizar intervenções pedagógicas apropriadas que tendem a gerar melhoria nas suas aprendizagens.

As ferramentas estatísticas são muito eficientes nas análises dos dados nas diferentes áreas, dando ao pesquisador segurança no caminho que está seguindo. Assim segundo Ribeiro *et al.*, (2019) a Modelagem Matemática se torna coerente com os pressupostos da Educação Estatística ao conjugar a ideia de aprender Estatística por meio do estudo, investigação, análise, interpretação, crítica e discussão de situações do cotidiano do aluno.

Conseqüentemente, os métodos de avaliação que são aplicados durante o processo de ensino e aprendizagem são ferramentas que permitem de alguma forma conhecer em forma individual cada aluno que compõe a turma.

São diversas as formas como os professores avaliam seus alunos, todas com o intuito de verificar se o ensino que ele aplicou foi assimilado. Mas também é importante procurar formas de testar se os métodos de avaliação de alguma forma explicam a aprendizagem do aluno.

No presente trabalho apresenta-se uma metodologia como ferramenta do estudo das relações que guardam as avaliações aplicadas pelo professor, a qual se baseia num modelo de regressão múltipla que é construído a partir das informações de sala de aula. Este modelo de

regressão proporciona os elementos necessários para construir os gráficos de controle de erros, que são as diferenças entre o aprendizado real e o aprendizado que o professor estimou ao aplicar os métodos de ensino.

2. Procedimentos metodológicos

A metodologia consiste em definir as seguintes etapas:

- a) Selecionar as turmas que precisam ser controladas. Geralmente são os períodos iniciais onde são lecionadas matérias básicas como cálculo, álgebra, probabilidade etc. Estas matérias contribuem com a segurança, criatividade e fortalecimento do aprendizado do aluno facilitando o desenvolvimento das matérias seguintes;
- b) Definir as variáveis que servirão como indicadores no processo de análise. Estes indicadores podem ser: notas de provas, frequências, exercícios, trabalhos etc. ou qualquer avaliação que o professor considerar; informações sobre a situação socioeconômica do aluno e outras, mas que seja possível de quantificar;
- c) Estabelecer relações entre os indicadores. Para a presente pesquisa será construído um modelo de regressão múltipla de modo que defina a relação entre as diferentes avaliações aplicadas pelo professor, por conseguinte o modelo a ser definido é:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \mu_i, \quad i = 1, 2, \dots, n; \quad (1)$$

Onde:

Y , variável dependente ou resultante;

n , número de observações;

k , número de parâmetros;

X , variáveis que precisam ser controladas;

β , representam os parâmetros do modelo;

μ , termo de perturbação ou erro.

O modelo (1) pode ser escrito em forma matricial para ajudar sua operacionalidade da seguinte forma:

$$Y = X\beta + \mu \quad (2)$$

Para estimar os parâmetros β é utilizado o princípio dos mínimos quadrados cujo objetivo é calcular a coluna de estimadores $\{\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k\}$. Pode-se então escrever que:

$$Y = X\hat{\beta} + e \quad (3)$$

Isolando e se tem:

$$e = Y - X\hat{\beta} \quad (4)$$

Onde, e é um estimador de μ , que é um vetor coluna dos resíduos das n observações, cuja soma de quadrados se define como:

$$\sum_{i=1}^n e^2 = e'e \quad (5)$$

Calculando o valor esperado da soma de quadrados dos resíduos (SQR), se tem:

$$SQR = E(e'e) = (n - k)\sigma^2 \quad (6)$$

Onde σ^2 , é a variância de perturbação, cujo estimador não tendencioso é o quadrado médio dos resíduos MQR :

$$MQR = \frac{e'e}{n-k} = \frac{(Y-X\hat{\beta})(Y-X\hat{\beta})}{n-k} \quad (7)$$

- d) Construção dos gráficos de controle da regressão múltipla. Aqui está se entrando em uma nova análise, concentrando-se no estudo da variável dependente Y frente a sua estimação \hat{Y} . Esta análise se inicia com o cálculo do desvio padrão do estimador da variância de perturbação que é a raiz do quadrado médio dos resíduos (\sqrt{MQR}). Portanto, os gráficos de controle se baseiam no comportamento dos resíduos, isto é, $e = Y - X\hat{\beta}$. Segundo o objetivo do presente trabalho, um aluno será considerado fora de controle quando a diferença entre o rendimento esperado $\hat{Y} = X\hat{\beta}$ e o rendimento real Y ultrapassa os limites do gráfico de controle dos resíduos apresentados a seguir:

Limite superior de controle: $Lsc = E[Y - \hat{Y}] + tS$;

Média de controle: $Lmc = E[Y - \hat{Y}]$;

Limite inferior de controle: $Lic = E[Y - \hat{Y}] - tS$.

O valor de t pode tomar valores múltiplos de S , mas no máximo 3, que significa uma alta confiança para uma um número grande de observações.

O valor esperado $E[Y - \hat{Y}] = E(e) = 0$, é uma das suposições da regressão, portanto, os limites de controle ficam:

$$Lsc = +tS$$

$$Lmc = 0$$

$$Lic = -tS$$

Os alunos cujo erro de previsão está fora de controle serão observados cuidadosamente para detectar se as notas previstas pelo professor condizem com as notas reais obtidas pelos alunos ao final do processo para verificar as causas que colocaram fora de controle. Num processo de ensino e aprendizagem a observação em análise é um aluno, portanto, se está fora de controle você não pode ignorar essa observação como acontece em outro tipo de processos, mas sim corrigir. Maiores informações em Kamenta (1990); Montgomery e Peck (1992); Johnston (1975).

3. Aplicação da metodologia

A metodologia apresentada foi aplicada em uma turma da matéria de Cálculo 1. Foram analisados 38 alunos regularmente matriculados com frequência dentro dos parâmetros estabelecidos. Os alunos foram identificados por números por serem dados confidenciais. O modelo selecionado foi submetido a testes estatísticos necessários para assegurar que as suposições da regressão não foram violadas. Foram consideradas 3 avaliações que consistiam cada uma em uma prova e um trabalho relacionada com os conteúdos passados em sala de aula e que seriam cobrados na prova correspondente.

O objetivo do estudo foi medir se as duas avaliações anteriores têm relação com a terceira e última avaliação à qual servirá para a construção dos gráficos de controle dos erros. Então descrevendo o modelo, as variáveis independentes são as duas primeiras avaliações representadas pelas variáveis X e a variável dependente que representa a terceira avaliação é representada pela variável Y , tendo como resultante o modelo mostrado a seguir:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \mu_i, \quad i = 1, \dots, 38,$$

Onde:

Y_i , é a variável que representa a resultante ou consequência das duas anteriores avaliações do i-ésimo aluno;

β_0 , β_1 e β_2 , são os parâmetros que se precisa estimar;

X_{i1} , é a primeira avaliação do aluno i-ésimo;

X_{i2} , é a segunda avaliação do aluno i-ésimo;

μ_i , é o erro.

Para estimar os parâmetros β é utilizado o princípio dos mínimos quadrados, cujo objetivo é calcular os estimadores $\{\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2\}$.

Depois de obter os estimadores pode-se então escrever que o modelo estimado é:

$$\hat{Y}_i = -1,49233557 + 0,46524507X_{i1} + 0,64015280X_{i2}$$

Os valores calculados de $\hat{\beta}_1$ e $\hat{\beta}_2$ explicam o grau de relação positiva entre as variáveis X_1 e X_2 respectivamente com a variável dependente Y , em outras palavras existe relação das duas primeiras avaliações com a última avaliação. Mas existe a necessidade de verificar estatisticamente se efetivamente o modelo estimado explica essa relação, para tanto será usado o teste seguinte:

$$H_0: \beta_1 = \beta_2 = 0 ;$$

$$H_a: \text{Não } H_0$$

Onde, H_0 é a hipótese nula e H_a é a hipótese alternativa. Considerando o 95% de confiança, pode-se usar então como medida de significância estatística o p-valor que deve ser menor de 0,05. ($p < 0,05$).

Os resultados deste teste são apresentados na tabela ANOVA:

1 – Tabela ANOVA

	Graus de liberdade (gl)	Soma de quadrados (SQ)	Quadrado médio (MQ)	F	F de significação
Regressão	2	240,36125209	120,18062605	21,57214578	0,00000079
Resíduo	35	194,98857247	5,57110207		
Total	37	435,34982456			

Fonte: O autor (2023)

Como o F de significação é igual a 0,00000079 e este é menor que o p , então rejeitamos a hipótese nula e conclui-se que os β_1 e β_2 são diferentes de zero, e, portanto, as duas primeiras avaliações explicam com antecedência sobre a terceira avaliação. Em outras palavras, existe relação entre as variáveis.

Em seguida se precisa obter os valores estimados de Y representado por \hat{Y} com objetivo de calcular o erro estimado que é:

$e_i = Y_i - \hat{Y}_i$, $i = 1, \dots, 38$. Assim, podemos obter o valor esperado da soma de quadrados dos resíduos estimados:

$$SQR = E(e'e) = 194,98857247$$

Uma vez calculado o valor esperado quadrado dos erros estimados, calcula-se o estimador da variância de perturbação ou quadrado médio do resíduo usando a expressão (7):

$$MQR = \frac{E(e'e)}{n - k} = 5,57110207$$

A raiz do quadrado médio dos resíduos é $\sqrt{MQR} = 2,36032$, e este valor será utilizado para o cálculo dos gráficos de controle.

O valor esperado $E[Y - \hat{Y}] = E(e) = 0$; é uma das suposições da regressão, portanto os limites de controle para $t = 1$ ficam:

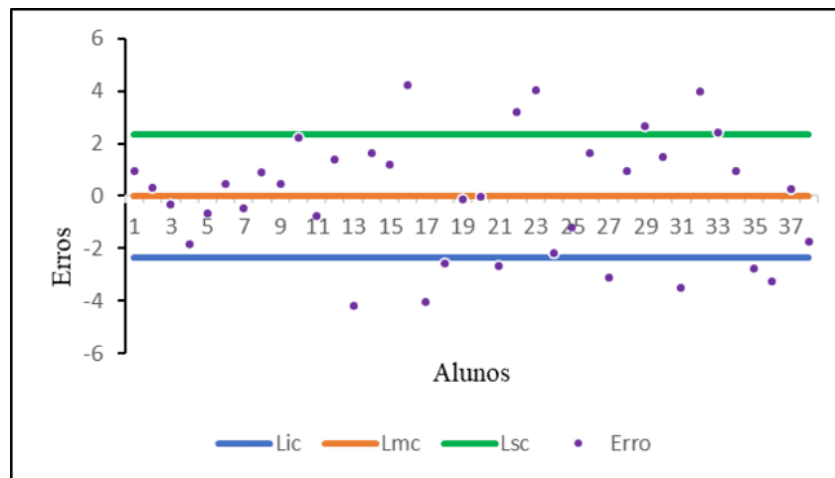
$$Lsc = +t\sqrt{MQR} = 2,360318214$$

$$Lmc = 0$$

$$Lic = -t\sqrt{MQR} = -2,360318214$$

Por conseguinte, o gráfico de controle de resíduos neste caso pode ser visto na Figura 1.

Figura 1 – Gráfico de controle de resíduos para $t = 1$



Fonte: O autor (2023)

Segundo a Figura 1, os pontos representam os erros dos 38 alunos, que é a diferença entre a aprendizagem prevista pelo professor ao início do semestre letivo e o resultado do processo.

Também se observa na Figura 1 que 13 alunos estão fora de controle, dos quais 5 alunos estão acima do limite superior de controle (Lsc). Nesta situação o professor fez uma previsão muito menor sobre o rendimento acadêmico que o aluno obteve ao final, ou seja, o aluno superou as expectativas do professor conforme o passar do tempo. Conseqüentemente, neste caso não representa preocupação no processo. Os 8 alunos restantes que estão abaixo do limite inferior de controle (Lic) representam situação inversa à anterior e o professor faz previsões favoráveis a estes alunos, mas eles não obtiveram o resultado esperado. Eles representam problemas ao processo de ensino e aprendizagem e, portanto, não foi alcançada a expectativa do professor, já que suas notas foram inferiores as estimadas.

Estes alunos que estão fora dos limites de controle representam erros de previsão, mas só foi considerado um desvio com relação à média, o que significa uma reduzida área de aceitação dentro dos limites definidos. Se são considerados dois desvios na área de aceitação, ou seja, que o valor para $t = 2$ os limites de controle ficariam da seguinte forma:

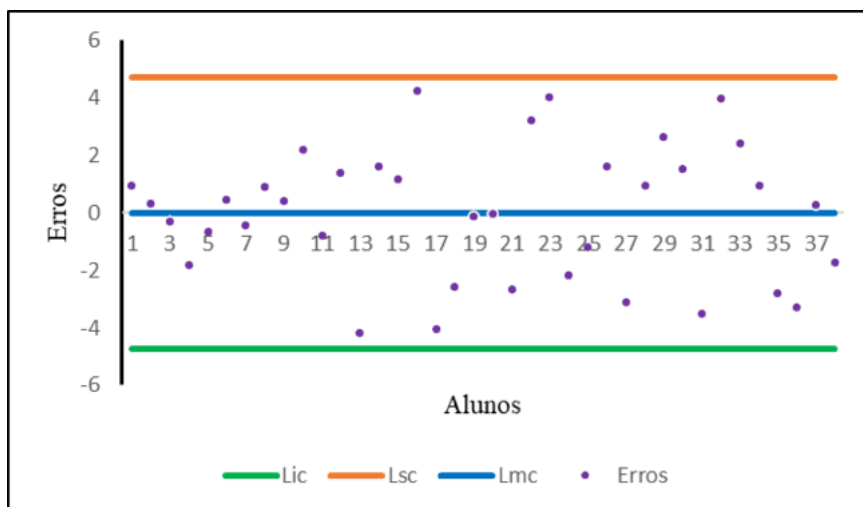
$$Lsc = +t\sqrt{MQR} = 4,720636428$$

$$Lmc = 0$$

$$Lic = -t\sqrt{MQR} = -4,720636428$$

O gráfico da distribuição de erros segundo estes limites é apresentado na Figura 2.

Figura 2 - Gráfico de controle de resíduos para $t = 2$



Fonte: O autor (2023)

Na Figura 2 se observa que todos os pontos estão dentro dos limites de controle. Neste caso podemos dizer que as previsões foram acertadas. Segundo este critério, podemos seguir aplicando os procedimentos de ensino e aprendizagem em futuras turmas.

4. Conclusões

A metodologia apresentada foi satisfatória e atingiu os objetivos propostos. Demonstrou que é necessário relacionar as diferentes variáveis que envolvem o processo de ensino e aprendizagem através da construção de um modelo de regressão através da relação entre as diferentes avaliações que são aplicadas nos alunos de uma determinada turma.

O modelo de regressão múltipla estimado (1), mostra que as primeiras avaliações têm relação positiva com a terceira e última, e esta relação é significativa como explica o teste F .

A partir deste modelo de regressão pode-se construir os gráficos de controle, de onde foi utilizado o desvio padrão da variância de perturbação ou raiz do quadrado médio dos resíduos.

Os resultados destes gráficos de controle mostram que para que um desvio seja $t = 1$ se tem 13 alunos fora de controle. Os alunos que estão fora dos limites de controle podem servir ao professor como indicadores para inferir sobre suas observações iniciais em relação a seu método de ensino e seus procedimentos de avaliação aplicados a uma determinada turma. Estes alunos que estão fora dos limites de previsão representam erros de previsão, mas só foi considerado um desvio com relação à média, significando uma reduzida área de aceitação



dentro dos limites definidos. Quando utilizados dois desvios com relação à média, isto é, ampliando-se a área de aceitação, todos os alunos estiveram dentro dos limites de controle, significando que as previsões do modelo estimado foram acertadas, como mostra a Figura 2. A metodologia de gráficos de controle de regressão permite até 3 desvios na construção dos gráficos de controle, portanto, dar valor de $t = 2$ ou seja dois desvios estase dentro dos limites permitido.

Assim sendo, o modelo de regressão juntamente com os gráficos de controle assegura que a avaliação final ou variável dependente Y é resultante das primeiras avaliações X . Desse modo estas primeiras avaliações devem ser controladas, isso explica que os alunos devem ser observados desde o início do processo.

REFERÊNCIAS

ESPASANDIN, Celi Lopes.; DE SOUZA, Antonio Carlos.; OLIVEIRA, Leandro de Souza; DE OLIVEIRA, Luzinete Mendoza. **O Campo de pesquisa da Educação Estatística Brasileira demarcado pela diversidade temática**. REnCiMa, v. 9, n. 2, p. 1-4, 2018.

JOHNSTON. J. **Métodos de Econometria**. Editora Vicens-Vives, Barcelona, 1975.

KAMENTA, J. **Elementos de Econometria**. Vol.2, 2^o ed. Atlas S.A. São Paulo 1990.

MONTGOMERY, D. C.; PECK, E. A. **Introduction to linear Regression Analysis**. 2^a ed. Jhon Wiley e Sons, Inc., New York, 1992.

RIBEIRO, Celso Campos; QUEIROZ, Cileida; SILVA, Coutinho. **A Modelagem Matemática e o letramento estatístico no ensino de gráficos**. REVEMAT, Florianópolis (SC), v.14, edição especial, p.1-20, 2019.

SAMÁ, Suzi; FONSECA, Laerte. **Projetos de aprendizagem sob as lentes da neurociência cognitiva: possibilidade para a construção de conceitos Estatísticos**. REVEMAT, Florianópolis SC. v.14, edição especial p.1-16, 2019.

VELIZ, Carlos Capuñay. **Estadística Aplicaciones**. 2^o Ed. Lima-Peru, 1993.