UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE MATEMÁTICA

RELAÇÃO DA NOTA NO PROEB/2012 COM NOTAS DE ANOS ANTERIOS E PERCENTUAL DE FREQUENCIA

CELIA RODRIGUES PIMENTEL

Trabalho de conclusão de curso apresentado ao curso em Matemática a Distância - PARFOR da Universidade Federal de Uberlândia (UFU) como parte dos requisitos para a conclusão do Curso de Licenciatura em Matemática.

UBERLÂNDIA
2016

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE MATEMÁTICA

RELAÇÃO DA NOTA NO PROEB/2012 COM NOTAS DE ANOS ANTERIOS E PERCENTUAL DE FREQUENCIA

CELIA RODRIGUES PIMENTEL

Orientador: Prof. Dr. LÚCIO BORGES DE ARAÚJO

Trabalho de conclusão de curso apresentado ao curso em Matemática a Distância - PARFOR da Universidade Federal de Uberlândia (UFU) como parte dos requisitos para a conclusão do Curso de Licenciatura em Matemática.

UBERLÂNDIA
2016

AGRADECIMENTOS

Gostaria de agradecer a Deus, pelo milagre da vida, por este sonho realizado, ele que é fonte infinita da sabedoria e promotor de toda realização humana.

Aos meus professores pela paciência e conhecimento repassado a mim, aos meus amigos que fiz no curso, pela disponibilidade durante esta trajetória.

A minha família pela compreensão durante o período dedicado a realização deste trabalho.

Enfim, de forma muito especial ofereço este trabalho ao meu orientador Lucio que, tenho uma grande admiração e carinho. Agradeço pelo conhecimento repassado a mim, pela atenção, dedicação e disponibilidades durante a construção deste trabalho acadêmico.

RESUMO

No estado de Minas Gerais o governo utiliza o SIMAVE para fazer uma avaliação do ensino público. Um dos programas do SIMAVE é o PROEB (Programa de Avaliação da Rede Pública de Educação Básica). O objeto deste trabalho foi utilizar análise de correlação e regressão é avaliar relação entre as notas em matemática no PROEB/2012 com as notas obtidas no ano 2010, 2011 e o percentual de alunos que fizeram prova no ano de 2012, de escolas estaduais da SER de Unaí-MG. Foi possível concluir que concluir que as notas de 2012 no PROEB estão relacionadas as notas dos anos anteriores. Já a variável percentual de alunos presentes na prova de 2012 não apresentou relação com a nota de 2012.

Palavras-chave: Desempenho escolar, correlação, regressão, PROEB.

Sumário

1 INTRODUÇÃO

O Sistema Mineiro de Avaliação da Educação Pública (SIMAVE) é composto por três programas de avaliação da Educação no Estado de Mina Gerais. Estes programas são (SIMAVE, 2016): o Programa de Avaliação da Rede Pública de Educação Básica (PROEB), o Programa de Avaliação da Aprendizagem Escolar (PAAE) e o Programa de Avaliação da Alfabetização (PROALFA). Conforme visualização no site da educação do Estado de Minas Gerais descreve “PROEB” como sendo uma Avaliação externa e censitária que busca diagnosticar a educação pública do estado de Minas Gerais (SECRETARIA EDUCAÇÃO DO ESTADO DE MINAS GERAIS, 2008).

O objetivo deste sistema de avaliação tem o intuito de direcionar ao governo seja no âmbito Municipal ou Estadual subsídios capazes de auxiliar no momento de decisões relativas a políticas públicas voltada para a área educacional e auxilio as escolas no tocante a área pedagógica (SECRETARIA EDUCAÇÃO DO ESTADO DE MINAS GERAIS, 2008).

O PROEB visa testar os conhecimentos dos alunos ao final de cada etapa da educação, no ensino fundamental e aplicado no início 5º ano e ao final no 9º ano e em relação ao ensino médio e aplicado aos alunos do 3º ano. Até o ano de 2015 era aplicado anualmente e de forma censitária, onde os alunos eram avaliados e classificados de acordo com seu desempenho, sendo dividido em três padrões a saber: baixo, intermediário e recomendável nas disciplinas de Português e Matemática e os resultados apresentados em uma escala de 0 a 500 (DEFATOONLINE, 2013).

No tocante a disciplina matemática o resultado de 2014 foi conforme a seguir: No PROEB, as mudanças também não foram significativas e os resultados permanecem insatisfatórios. Em Matemática os percentuais são ainda menores, exceto para o 5º ano, em que a proficiência média foi de 239,3, com 61,5% em padrão recomendado; no 9º ano, 265,5 de proficiência média e 23,2% encontram-se no padrão recomendado. No 3º ano do Ensino Médio, a proficiência média foi de 283,4 e apenas 4,4% estão em padrão recomendado de Matemática. (SECRETARIA EDUCAÇÃO DO ESTADO DE MINAS GERAIS, 2015).

Após o resultado do PROEB de 2014 a Secretaria de Educação de Minas Gerais apresentou novidades com a finalidade de ajudar as escolas em projetos pedagógicos mais direcionados. Inicialmente pretende se criar instrumentos capazes de possibilitar aos alunos as dificuldades encontradas nessas avaliações. Assim, com o objetivo de estimular os gestores escolares a utilizar os dados educacionais como ferramenta no projeto pedagógico (SECRETARIA EDUCAÇÃO DO ESTADO DE MINAS GERAIS, 2015). A primeira das mudanças e aplicação da avaliação aos alunos do 7º ano do ensino fundamental e 1º ano do ensino médio, sendo esta uma novidade porque as provas eram aplicadas tão somente aos alunos do 5º e 9º anos do ensino fundamental e bem como o 3º ano do ensino médio.

Essa medida possibilita aos educadores que ao identificar o problema possa realizar intervenções pedagógicas antes que o aluno termine o ensino fundamental ou médio. Uma novidade e que a partir do presente ano (2016) os estudantes do 5º e 9º anos do ensino fundamental e 3º ano do ensino médio somente serão avaliados nos anos que sejam pares. A partir deste e em todos os anos impares irão realizar a avaliação os alunos do 7º ano do ensino fundamental do 1º e 3º ano do ensino médio (SECRETARIA EDUCAÇÃO DO ESTADO DE MINAS GERAIS, 2015).

Outra mudança e no tocante aos padrões que será acrescentado mais um, sendo assim dividido em: abaixo do básico, básico, adequado e avançado permitindo atuações mais efetivas e pontuais no setor pedagógico (SECRETARIA EDUCAÇÃO DO ESTADO DE MINAS GERAIS, 2015).

Os objetivos deste trabalho de conclusão de curso é avaliar relação entre as notas em matemática no PROEB/2012 com as notas obtidas no ano 2010, 2011 e o percentual de alunos que fizeram prova no ano de 2012, das escolas estaduais da Secretaria regional de Ensino (SER) de Unaí-MG.

2 METODOLOGIA

2.1 DESCRIÇÃO DOS DADOS

Os dados serão obtidos no site do SIMAVE (http://www.simave.caedufjf.net/o-programa/). Foram coleadas as seguintes informações de escolas estaduais da Secretaria Regional de Educação (SER) de Unaí:

- Nota em matemática no PROEB/2012;
- Percentual de alunos que fizeram a prova no ano de 2012;
- Nota em matemática no PROEB/2011;
- Nota em matemática no PROEB/2010;

Estes dados refere-se a avaliação do terceiro ano do ensino média. Todas análises serão feitos utilizando o software BIOESTAT (2007) e nível de significância de 5%.

2.2 CORRELAÇÃO

Frequentemente é necessário estudar o relacionamento entre duas ou mais variáveis (LORI, 2016). Ao estudo do relacionamento entre duas ou mais variáveis denominamos de correlação e regressão. Se o estudo tratar apenas de duas variáveis tem-se a correlação e a regressão simples, se envolver mais do que duas variáveis, tem-se a correlação e a regressão múltiplas. A regressão e a correlação tratam apenas do relacionamento do tipo linear entre duas variáveis.

A análise de correlação fornece um número que resume o grau de relacionamento linear entre as duas variáveis (MONTYGOMERY, RUGER; 2003). Já a análise de regressão fornece uma equação que descreve o comportamento de uma das variáveis em função do comportamento da outra variável (MONTYGOMERY, RUGER; 2003).

Para ter uma ideia do tipo de relacionamento que possa existir entre as duas variáveis, os valores das variáveis são colocadas no que é denominado de diagrama de dispersão (TRIOLA, 1999). Uma das variáveis (X) é representada no eixo horizontal e a outra variável (Y) no eixo vertical,

Uma olhada rápida no diagrama de dispersão mostra a existência de um relacionamento entre as variáveis, com altos valores de uma das variáveis associados a altos valores da outra variável (TRIOLA, 1999). Se não houvesse relacionamento entre elas, os pontos estariam distribuídos ao acaso no gráfico sem mostrarem alguma tendência.

Apesar do diagrama de dispersão apresentar uma ideia do tipo e extensão do relacionamento entre duas variáveis X e Y, seria altamente desejável ter um número que medisse esta relação. Esta medida existe e é denominada de coeficiente de correlação (MONTYGOMERY, RUGER; 2003).

A suposição básica sobre o coeficiente de correlação é que o relacionamento entre as duas variáveis seja linear (VIALI, 2016). As duas variáveis podem estar perfeitamente relacionadas, mas se não for de forma linear o valor do coeficiente pode ser zero ou próximo de zero.

Uma segunda hipótese é que as variáveis envolvidas sejam aleatórias e que sejam medidas no mínimo em escala de intervalo. Ele não se aplica a variáveis em escala nominal ou ordinal ou quando uma das variáveis é manipulada experimentalmente, pois neste caso, a escolha dos valores experimentais vai influenciar o valor de r obtido.

Uma terceira hipótese é que as duas variáveis tenham uma distribuição conjunta normal bivariada. Isto é equivalente a dizer que para cada x dado a variável y é normalmente distribuída.

Suponha-se que existam apenas duas variáveis X e Y. Uma amostra da variável "X", assumindo os valores particulares $X_1, X_2, ..., X_n$ e uma amostra da variável "Y" assumindo os valores particulares $Y_1, Y_2, ..., Y_n$ são obtidas e suponha-se ainda que o objetivo é saber se existe algum tipo de relacionamento linear entre estas duas variáveis (LEVIN, FOX; 2004). Isto poderá ser medido pelo coeficiente e correlação que fornece o grau de relacionamento linear entre duas variáveis.

Assim dadas duas amostras, uma da variável X e outra da variável Y, o coeficiente de correlação amostral poderá ser calculado através da seguinte expressão (MONTYGOMERY, RUGER; 2003):

$$r = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{[\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2]^{1/2}}$$

As propriedades mais importantes do coeficiente de correlação são:

1. O intervalo de variação vai de -1 a +1.

2. O coeficiente de correlação é uma medida adimensional, isto é, ele é independente das unidades de medida das variáveis X e Y.
3. Quanto mais próximo de +1 for “r”, maior o grau de relacionamento linear positivo entre X e Y, ou seja, se X varia em uma direção Y variará na mesma direção.
4. Quanto mais próximo de -1 for “r”, maior o grau de relacionamento linear negativo entre X e Y, isto é, se X varia em um sentido Y variará no sentido inverso.
5. Quanto mais próximo de zero estiver “r” menor será o relacionamento linear entre X e Y. Um valor igual a zero, indicará ausência apenas de relacionamento linear.

2.3 REGRESSÃO

Uma vez constatado que existe correlação linear entre duas variáveis, pode-se tentar prever o comportamento de uma delas em função da variação da outra (MONTGOMERY, RUNGER, 2003).

Para tanto será suposto que existem apenas duas variáveis. A variável X (denominada variável controlada, explicativa ou independente) com valores observados $X_1, X_2, ..., X_n$ e a variável Y (denominada variável dependente ou explicada) com valores $Y_1, Y_2, ..., Y_n$. Os valores de Y são aleatórios, pois eles dependem não apenas de X, mas também de outras variáveis que não estão sendo representadas no modelo. Estas variáveis são consideradas no modelo através de um termo aleatório denominado “erro”. A variável X pode ser aleatória ou então controlada.

Desta forma pode-se considerar que o modelo para o relacionamento linear entre as variáveis X e Y seja representado por uma equação do tipo (MONTGOMERY, RUNGER, 2003):

$$Y = \alpha + \beta X + U,$$

em que “U” é o termo erro, isto é, “U” representa as outras influências na variável Y além da exercida pela variável “X”.

Esta equação permite que Y seja maior ou menor do que $\alpha + \beta X$, dependendo de “U” ser positivo ou negativo. De forma ideal o termo “U” deve ser pequeno e independente de X, de modo que se possa modificar X, sem modificar "U", e determinar o que ocorrerá, em média, a Y, isto é:

$$E(Y/X) = \alpha + \beta X$$

Os dados $\{(X_i, Y_i), i = 1, 2, ..., n\}$ podem ser representados graficamente marcando-se cada par (X_i, Y_i) como um ponto de um plano (VIALI, 2016). Os termos U_i são iguais a distância vertical entre os pontos observados (X_i, Y_i), e os pontos calculados $(X_i, \alpha + \beta X)$. Isto está ilustrado na Figura 1.

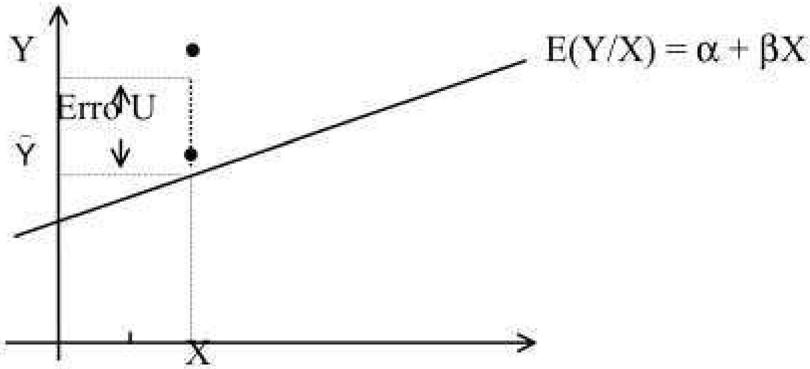

Figura 1: Representação dos valores observados, reta ajustada e erro (VIALI, 2016).

Um modelo de regressão consiste em um conjunto de suposições sobre a distribuição dos termos "erro" e as relações entre as variáveis X e Y.

Estas suposições são (VIALI, 2016):

i. $E(U_i) = 0$;
ii. $Var(U_i) = \sigma^2$

Na suposição (i) o que se está supondo é que os U_i são variáveis aleatórias independentes com valor esperado igual a zero e na (ii) que a variância de cada U_i é a mesma e igual a σ^2, para todos os valores de X.

Existem alguns métodos para ajustar uma linha entre as variáveis X e Y e o mais utilizado é o denominado método dos mínimos quadrados (MMQ). A reta obtida através deste método, não é necessariamente, o "melhor" ajustamento possível, mas possui muitas propriedades estatísticas que são desejáveis.

Sejam a e b estimadores de α e β e $E_i = Y_i - a - bX_i$ o desvio observado em relação a reta ajustada, isto é, E_i é um estimador do termo U_i. O método dos mínimos quadrados exige que os estimadores a e b sejam escolhidos de tal forma que a soma dos quadrados dos desvios dos mesmos em relação à reta de regressão ajustada seja mínima, isto é (MONTGOMERY, RUNGER, 2003):

$$L = \sum_{i=1}^{n} E_i^2 = \sum_{i=1}^{n} (Y_i - a - bx_i)^2$$

Para tornar mínima esta soma em relação aos estimadores a e b, é necessário derivar a expressão parcialmente em relação aos valores a e b. Após algumas simplificações obtém-se:

$$\sum_{i=1}^{n} Y_i = na + b \sum_{i=1}^{n} X_i \qquad \text{(i)}$$
$$\sum_{i=1}^{n} Y_i X_i = a \sum_{i=1}^{n} X_i + b \sum_{i=1}^{n} {X_i}^2 \qquad \text{(ii)}$$

que são denominadas de equações normais da regressão, onde “n” é o número de pares de observações.

Dividindo-se a equação (i) por “n” e isolando o valor de a vem:

$$a = \bar{Y} - b\bar{X}$$

levando-se este resultado na equação (ii) tem-se:

$$b = \frac{\sum_{i=1}^{n} Y_i X_i - \frac{(\sum_{i=1}^{n} Y_i)(\sum_{i=1}^{n} X_i)}{(n)}}{\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} Y_i\right)^2}{n}}$$

A reta estimada de regressão será então:

$$\hat{Y} = a + bX$$

2.4 DECOMPOSIÇÃO DA SOMA DOS QUADRADOS

Pela Figura 2, pode-se perceber que o desvio em relação a Y (desvio total), isto é, $Y_i - \overline{Y}$ pode ser decomposto em dois outros desvios (BARBETA, REIS, BORNIA; 2004). O desvio explicado pela linha de regressão, isto é, $\widehat{Y_\iota} - \overline{Y}$ e o desvio não-explicado (resíduos) pela linha de regressão, isto é, $Y - \widehat{Y_i}$.

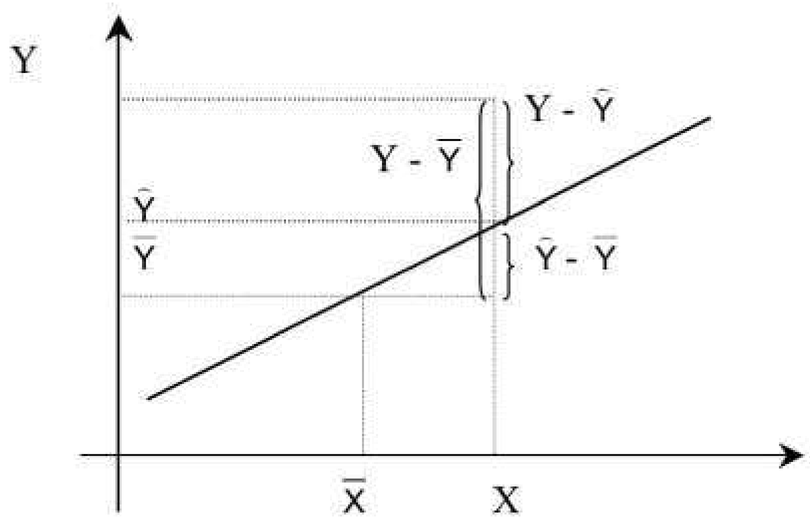

Figura 2: Desvios na Regressão Linear Simples. (VIALI, 2016)

Assim, percebe-se que:

$$\sum(Y_i - \bar{Y})^2 = \sum\left(\hat{Y}_i - \bar{Y}\right)^2 + \sum\left(Y_i - \hat{Y}\right)^2$$

em que:

$\sum(Y_i - \bar{Y})^2$ é a variação total

$\sum\left(\hat{Y}_i - \bar{Y}\right)^2$ é a variação explicada pela equação de regressão

$\sum\left(Y_i - \hat{Y}\right)^2$ é a variação não explicada pela equação de regressão

Cada soma de quadrados está associada a um certo número de graus de liberdade (BARBETA, REIS, BORNIA; 2004). Os desvios de cada observação em relação às estimativas de E(Y) têm graus de liberdade a n subtraído do número de parâmetros estimados em E(Y). Assim, os desvios $Y_i - \bar{Y}$ têm $n -$ 1 graus de liberdade; e os desvios $Y_i - \hat{Y}_i$ tem $n -$ 2 graus de liberdade.

A soma de quadrados dividida pelo correspondente grau de liberdade fornece o quadrado médio ou variância. E a razão entre o quadrado médio da regressão e o quadrado médio do erro resulta na chamada razão F (veja Tabela 1).

Tabela 1: Esquema da Análise de variância (Anova) para regressão linear simples.

Fonte de variação	*gl*	SQ	QM	F
Regressão	1	SQR = $\sum(\hat{y}_i - \bar{y})^2$	QMR = SQR/1	*f* = QMR/QME
Erro	n – 2	SQE = $\sum(y_i - \hat{y}_i)^2$	QME = SQE/n – 2	-
Total	n – 1	SQT = $\sum(y_i - \bar{y})^2$	QMT = SQT/n – 1	-

2.5 COEFICIENTE DE DETERMINAÇÃO

Um indicador que fornece elementos para a análise do modelo adotado é o coeficiente de determinação ou de explicação, definido por (BARBETA, REIS, BORNIA; 2004):

$$R^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = \frac{\text{variação explicada pela regressão}}{\text{variação total}}$$

O coeficiente de determinação indica quantos por cento a variação explicada pela regressão representa sobre a variação total. Deve-se ter:

$$0 \leq R^2 \leq 1$$

Se R^2 for igual a 1, isto significa que todos os pontos observados se situam “exatamente” sobre a reta de regressão. Tendo-se, neste caso, um ajuste perfeito. As variações da variável Y são 100% explicadas pelas variações da variável X, não ocorrendo desvios em torno da função estimada. Por outro lado, se $R^2 = 0$, isto quer dizer que as variações de Y são exclusivamente aleatórias e explicadas pelas variações de outros fatores que não X (BARBETA, REIS, BORNIA; 2004).

3 RESULTADOS

Diante da análise de dados, nos anos de 2010 a 2012, verificou-se que uma possível diferença entre tais anos como pode ser observado na Tabela 2, podendo, portanto, haver alguma relação de dependência entre as notas de 2012 e as notas em 2010 ou 2011 ou a frequência em 2012. A correlação entre essas variáveis são observadas na Tabela 3 e nota-se que as notas em 2012 estão correlacionadas significativamente com as notas em 2011 (p = 0,049) e com as notas em 2010 (p = 0,001), mas não apresentou significância (p = 0,206) com o percentual de alunos que participaram da prova em 2012. Nas Figura 3, Figura 4 e Figura 5, são apresentados os diagramas de dispersão variáveis preditoras (nota em 2010, em 2011 e percentual de frequência) versus a vaiável resposta (nota em 2012).

Tabela 2: Estatísticas descritivas das variáveis notas nos anos 2010, 2011, 2012 e percentual de frequência em 2012.

Variável	Média	Mediana	Desvio Padrão	Mínimo	Máximo
nota 2012	264,53	262,25	19,05	228,40	304,50
nota 2011	265,52	266,10	17,82	231,80	298,50
nota 2010	260,00	266,70	58,20	0,00	325,20
Percentual	86,13	87,65	12,26	52,50	100,00

Tabela 3: Estatísticas descritivas das variáveis notas nos anos 2010, 2011, 2012 e percentual de frequência em 2012.

	nota 2010	nota 2011	percentual
Correlação	0,60	0,40	0,26
valor p	0,001	0,049	0,206

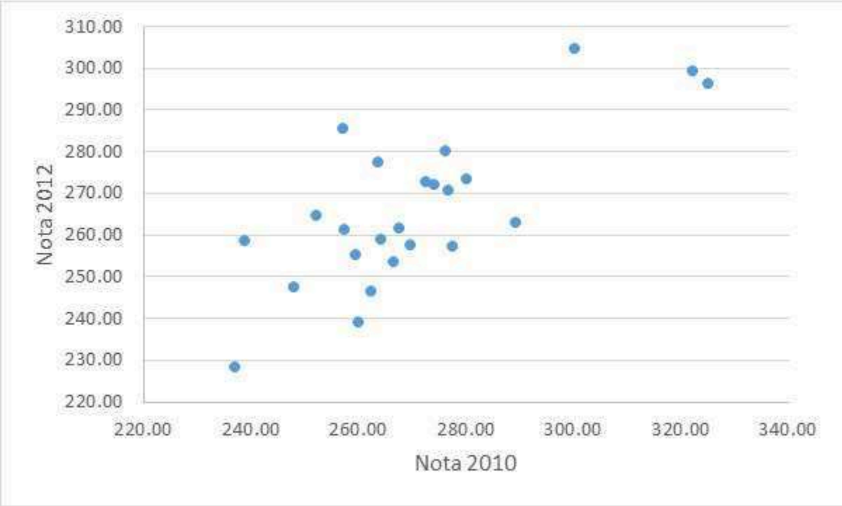

Figura 3: Gráfico de dispersão das variáveis nota em 2010 versus nota em 2012.

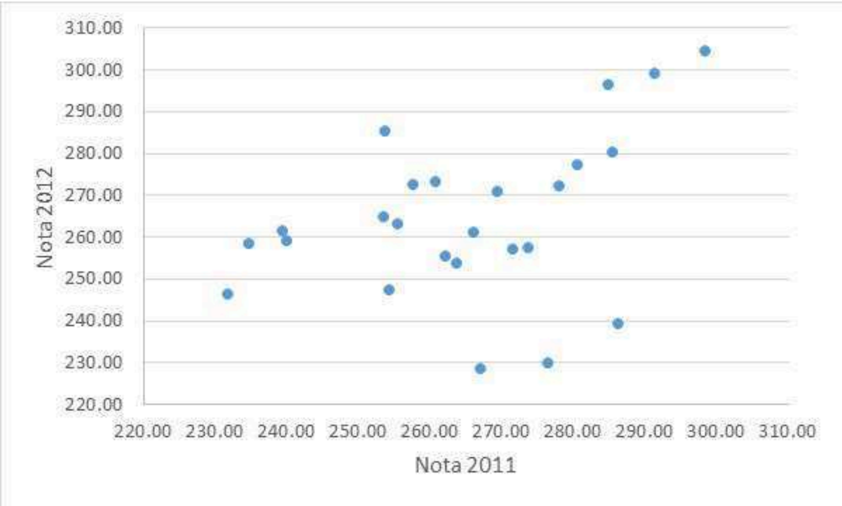

Figura 4: Gráfico de dispersão das variáveis nota em 2011 versus nota em 2012.

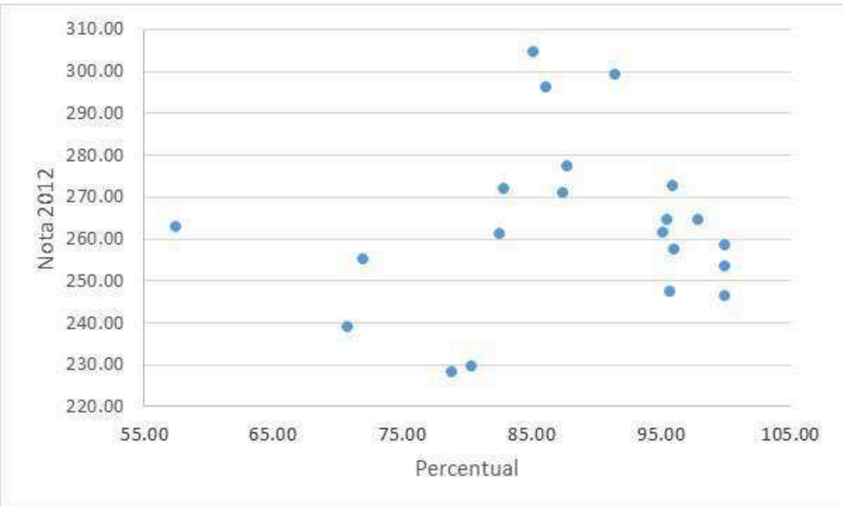

Figura 5: Gráfico de dispersão das variáveis percentual de frequência em 2012 versus nota em 2012.

As análises de variância da modelos de regressão linear simples (da nota em 2012 em função das demais variáveis) são apresentados na Tabela 4.

Tabela 4: Análise de variância da regressão linear simples da variável nota em 2012 em função da das variáveis nota em 2010, nota em 2011 ou percentual de frequência em 2012.

		2010			2011			Percentual		
	GL	QM	F	valor p	QM	F	valor p	QM	F	valor p
Regressão	1	3305,85	13,19	0,001	1431,92	4,31	0,049	595,14	1,69	0,207
Resíduo	23	250,60			332,08			353,10		
Total	24									
R^2		0,36				0,16				0,07

Ao analisar a nota 2010 como variável explicativa da nota em 2012, tem-se um R^2 = 0,36, ou seja, 36% da variação da nota em 2012 e explicada pela nota de 2010 (Tabela 4), que é um valor moderado. Percebe-se que este modelo é significativo (p = 0,001) e as estimativas de mínimos quadrados dos parâmetros (Tabela 5) são a= 212,10 e b = 0,20, sendo que o valor de b indica para 1 ponto no ano de 2010 espera-se que aumente 0,20 pontos na nota do ano de 2012. Dessa foram a equação para estimar a nota em 2012 (*N*2012) em função da nota de 2010 (*N*2010) é:

$$\widehat{N2012} = 212{,}10 + 0{,}20(N2010)$$

Tabela 5: Estimativa dos coeficientes da regressão linear simples variável nota em 2012 em função das variáveis nota em 2010, nota em 2011 ou percentual de frequência em 2012.

Coeficientes	2010	2011	Percentual
a	212.10	149.40	230.26
b	0.20	0.43	0.40

Considerando a nota 2011 tem-se um R^2 = 0,16, ou seja, 16% da variação da nota em 2012 e explicada pela nota de 2011 (Tabela 4). Percebe-se que este modelo é significativo (p = 0,049), mas com baixo poder de explicação e as estimativas de mínimos quadrados dos parâmetros (Tabela 5) são a= 149,40 e b = 0,43, sendo que o valor de b indica para 1 ponto no ano de 20101 espera-se que aumente 0,43 pontos na nota do ano de 2012. Dessa foram a equação para estimar a nota em 2012 ($N2012$) em função da nota de 2011 ($N2011$) é:

$$\widehat{N2012} = 149{,}40 + 0{,}43(N2011)$$

O baixo pode predito deste modelo pode ser justificado pelo fato que teve 3 escola que tiveram notas relativamente altas em 2011 (acima de 267 pontos) e em 2012 foram relativamente baixas (abaixo de 239,1 pontos), como pode ser observado na Figura 4.

Vale ressaltar que tentou-se ajustar um modelo de regressão múltipla, considerando como variável independente a nota em 2010 e nota em 2011, mas ao fazer a seleção de variáveis o modelo selecionado incluía somente a nota em 2010.

Considerando o percentual de presença no ano de 2012 tem-se um $R^2 = 0{,}07$, ou seja, apenas 7% da variação da nota em 2012 e pelo percentual de alunos presente na prova no ano de 2012 (Tabela 4). Percebe-se que este modelo não é significativo ($p = 0{,}207$). Portanto está variável não é adequada para estudar a nota da escola.

4 CONCLUSÃO

Com este trabalho, utilizando a análise de correlação e regressão, foi possível concluir que as notas de 2012 no PROEB estão relacionadas as notas dos anos anteriores, sendo que em o ano de 2010 se mostrou melhor preditor que 2011. Já a variável percentual de alunos presentes na prova de 2012 não apresentou relação com a nota de 2012. Portanto cabe aos gestores das escolas das escolas estaduais da SRE de Unaí avaliar as notas das escolas no PROEB, pois escolas notas baixos nos anos anteriores tendem a permanecer com notas e escolas notas altas a ter notas altas nas próximas avaliações.

5 BIBLIOGRAFIA

BARBETTA, P.A.; REIS, M.M.; BORNIA, A.C. *Estatística para os cursos de engenharia e informática*. São Paulo: Atlas, 2004. 410p.

BIOESAT. Versão5.3 Instituto Mamiraua 2007.

DEFATOONLINE. *Alunos das escolas estaduais de Minas melhoram desempenho em Português e Matemática*. Disponivel em: < http://www.defatoonline.com.br/noticias/ultimas/23-04-2013/alunos-das-escolas-estaduais-de-minas-melhoram-desempenho-em-portugues-e-matematica> acesso em 05 de maio de 2016.

LEVIN, J.; FOX, J.A. *Estatística para Ciências Humanas*. 9.ed. São Paulo: Prentice – Hall, 2004

MONTGOMERY, D.C., RUNGER, G.C. *Estatística Aplicada e Probabilidade Para Engenheiros*, 2 ed., Rio de Janeiro: LTC - Livros Técnicos e Científicos, 2003. 463p

SECRETARIA EDUCAÇÃO DO ESTADO DE MINAS GERAIS. *Avaliação da Rede Pública de Educação Básica – PROEB*. 2008. Disponível em: <https://www.educacao.mg.gov.br/ajuda/page/297-proeb >. Acesso em: 05 de maio de 2016.

SECRETARIA EDUCAÇÃO DO ESTADO DE MINAS GERAIS. *Secretaria de Estado de Educação apresenta resultados de avaliações e novidades em sistema de gestão escolar. 2015*. Disponível em: < https://www.educacao.mg.gov.br/leis/story/7184-secretaria-de-estado-de-educacao-apresenta-resultados-de-avaliacoes-e-novidades-em-sistema-de-gestao-escolar>. Acesso em: 07 de maio de 2016.

SIMAVE. *O Simave*. 2016 disponivel em: < http://www.simave.caedufjf.net/o-programa/>. Acesso em: 05 de maio de 2016.

TRIOLA, M.F. *Introdução à estatística*. 7. ed. Rio de Janeiro: LTC – Livros Técnicos e Científicos, 1999. 410p.

VIALI, L. *Correlação e Regressão: Série estatística básica*, 2016, disponível em: http://www.pucrs.br/famat/viali/graduacao/engenharias/material/apostilas/Apostila_5.pdf > Acesso em: 18 de agosto de 2016.