



COBENGE
2021

XLIX Congresso Brasileiro
de Educação em Engenharia
e IV Simpósio Internacional
de Educação em Engenharia
da ABENGE

28 a 30 de SETEMBRO

Evento Online

"Formação em Engenharia:
Tecnologia, Inovação e Sustentabilidade"

REGRESSÃO LINEAR MÚLTIPLA COMO ESTRATÉGIA DE PLANEJAMENTO INDUSTRIAL: UM EXEMPLO DE TRANSFERÊNCIA DE CONHECIMENTO ENTRE UNIVERSIDADE E INDÚSTRIA

DOI: 10.37702/2175-957X.COBENGE.2021.3646

Natal Junio Pires - natal.fg@gmail.com
Centro Federal de Educação Tecnológica de Minas Gerais
Rua das Candeias 235
38181-330 - Araxá - MG

Antônio Lenito Soares Júnior - alsoaresjr@gmail.com
Centro Federal de Educação Tecnológica de Minas Gerais
Av. Ministro Olavo Drummond 25
38180-510 - Araxá - MG

Rogério Leonardo Oliveira - rogerio.eminas@gmail.com
Mosaic Fertilizantes
Rua Antonio Conselheiro 190
38181-429 - Araxá - MG

Tiago Alves - tiago.alves.net@gmail.com
Mosaic Fertilizantes do Brasil
Rua Elisabete Cintia Gervasio 30
38181-778 - Araxá - MG

Manuel Cardoso Junior - manuel.junior@mosaicco.com
Mosaic Fertilizantes
Rua Maria Helena Carneiro 70
38182-652 - Araxá - MG

Resumo: Este trabalho apresenta como um problema industrial emergido durante a oferta de uma disciplina num curso de Mestrado Profissional em Engenharia de Minas, pôde ser resolvido a partir da interação sinérgica entre docentes da disciplina e alunos oriundos de uma empresa de mineração na cidade de Araxá-MG. O problema proposto aqui colocava a necessidade de se obter um modelo matemático que relacionasse um conjunto de variáveis predictoras com a

Promoção:



Realização:





COBENGE
2021

XLIX Congresso Brasileiro
de Educação em Engenharia
e IV Simpósio Internacional
de Educação em Engenharia
da ABENGE

28 a 30 de SETEMBRO

Evento Online

"Formação em Engenharia:
Tecnologia, Inovação e Sustentabilidade"

precificação da tonelada de minério explorado pela empresa, de modo a melhorar a qualidade das decisões tomadas durante o planejamento e gerenciamento do orçamento anual. Neste sentido, um conjunto de dados históricos foi obtido, processado estatisticamente e a partir das estatísticas, média e desvio-padrão, realizou-se a simulação de um total de 20.000 cenários diferentes em função de 11 variáveis regressoras. A partir deste conjunto de cenários, empregou-se o "método dos mínimos quadrados" para obter a equação de regressão através do software Minitab®. Com o modelo obtido e sua tabela de ANOVA foi possível, dentre outras coisas, quantificar a contribuição de cada variável preditora sobre a variável resposta (precificação do minério), bem como utilizá-lo para fazer previsões e tomadas de decisão acerca do planejamento, produtividade e custo nos projetos da referida indústria. Por fim, esta relação se mostrou extremamente produtiva e indicadora da importância da aproximação entre a universidade e a indústria, como possibilidade de produzir interessantes soluções dentro da própria empresa.

Palavras-chave: Interação universidade-indústria, regressão linear, mineração.

Promoção:



Realização:



REGRESSÃO LINEAR MÚLTIPLA COMO ESTRATÉGIA DE PLANEJAMENTO INDUSTRIAL: UM EXEMPLO DE TRANSFERÊNCIA DE CONHECIMENTO ENTRE UNIVERSIDADE E INDÚSTRIA

1 INTRODUÇÃO

Darcy Ribeiro em seu livro intitulado *"A universidade necessária"*, já evidenciava, dentro de um amplo conjunto de levantamento teórico, a importância da universidade como força de transformação da sociedade e aceleração evolutiva. Não basta à universidade apenas gerar conhecimento, mas também é importante que ela capacite os discentes para a aplicação deste conhecimento no sentido de superar problemas de ordem diversa, ou seja, em suma qualificar uma força de trabalho que garanta o desenvolvimento autônomo de um país (RIBEIRO, 1969). Passados 46 anos da publicação anterior, Renato Dagnino, engenheiro e estudioso brasileiro, escreveu um artigo cujo título é: *"Como é a universidade de que o Brasil precisa?"* (DAGNINO, 2015). Como se pode ver, a questão levantada por Ribeiro (1969) permanece em aberto, não porque nada tenha sido feito durante este tempo, mas sim pela complexidade do tema em se definir qual a universidade necessária ou como a universidade deve ser. Isto se relaciona com a grande dinamicidade dos setores que da universidade dependem, os quais vão desde o social até o industrial.

Nessa infinidade de situações complexas, uma que ganha destaque é aquela ligada à relação universidade-empresa. Conforme salienta Dagnino (2019), várias têm sido, ao longo do tempo, as sugestões de aproximação entre os setores produtivo e educacional, porém até hoje a quantidade de situações exitosas é muito pequena em nosso país (CRUZ, 2018). A isso se atribui uma série de razões regionais, econômicas e o fato de que empresas brasileiras, em geral, não têm atividades próprias de pesquisa avançada, conforme exposto por Dagnino em entrevista concedida ao jornal Sul21 (DAGNINO, 2019).

Goldemberg (2018) enfatiza que a universidade tem um papel central no desenvolvimento regional e global de uma nação. Principalmente quando se trata de interiorização geográfica da universidade, uma das principais funções dela é elevar a qualidade de vida nas regiões em que ela se instala. Dessa forma, uma vez implantada, ela não fica alheia e passa a se envolver em questões de ordem tecnológica, econômica e social, que uma vez tratadas contribuem para o desenvolvimento da sociedade.

Neste sentido, entende-se que um dos espaços mais privilegiados em termos de problemas a serem tratados são as empresas/indústrias, pois com a crescente exigência por inovação e otimização de processos, já há muito tempo se sabe que tais empresas não podem depender apenas de competências internas para abordar seus problemas. A articulação com agentes externos é de fundamental importância para que se atinja um nível de desempenho que permita à empresa sobreviver no mercado, e assim a universidade é um local estratégico para elas se apoiarem (PINHO, 2018).

Neste cenário, o processo de capacitação técnica do corpo produtivo e continuidade dele, passa pela universidade, dado que esta detém um corpo técnico altamente especializado e apto a interagir, dispostos assim a sanar problemas de natureza diversa. Entendendo essa necessidade de aproximação com o setor produtivo, a nossa instituição decidiu implantar no ano de 2018, o seu programa de pós-graduação em Engenharia de Minas – Mestrado Profissional, com vistas a atender uma ampla gama de empresas/indústrias da região, as quais atuam direta ou indiretamente no campo da mineração. Desde então, tem se frutificado uma série de parcerias com empresas diversas

no sentido de receberem formação técnica e em contrapartida estas financiam projetos, laboratórios e pesquisas na instituição.

Neste trabalho é apresentado um problema industrial que emergiu durante a oferta de uma disciplina intitulada "Estatística, Planejamento e Análise de Experimentos" no curso de Mestrado Profissional supracitado. O problema pôde ser resolvido a partir da interação sinérgica entre docentes da disciplina e alunos/profissionais de uma empresa de mineração na cidade de Araxá-MG. Com este exemplo, espera-se contribuir para o incentivo de mais interações em que tanto a empresa quanto a universidade saiam recompensadas.

2 APRESENTAÇÃO DO PROBLEMA

Segundo Borges (2013), no ramo da mineração os investimentos ultrapassam facilmente a cifra de centenas de milhões de dólares. Portanto, a otimização dos gastos/investimentos é crucial para se garantir a viabilidade do processo de produção. Um dos grandes desafios é o de gerir o custo do seu processo produtivo, e isto é ainda mais desejável em indústrias com alto capital mobilizado em ativos, como é o caso da mineração (NETO, 2018).

Segundo Thums e Mareth (2020), a redução dos custos de produção sempre foi uma das metas ligadas a estudos e planejamentos de produção, uma vez que a razão da existência das empresas é a geração de riqueza.

Diante dessas questões, o objetivo deste trabalho foi o de se utilizar de uma metodologia que conseguisse precificar o custo do minério lavrado em função de variáveis que nele influem, e assim servir como ferramenta para a tomada de decisões, bem como buscar o aumento da produtividade e diminuir os custos no processo de produção. Este problema trata-se de uma situação real de planejamento numa indústria envolvida com o campo de mineração na cidade de Araxá-MG.

3 METODOLOGIA

A partir de um extenso banco de dados obtido de empresas envolvidas em operações de mina, foram calculados a média e desvio-padrão para as variáveis que influenciam na precificação do minério lavrado, considerando-se gastos fixos e variáveis. Foram avaliados os impactos do tempo fixo de operação dos equipamentos, distância de transporte, velocidade média de transporte, carga média transportada, proporção estéril \times minério, disponibilidade física e utilização física de frota. Com esses dados realizou-se uma simulação variando simultânea e aleatoriamente os índices dos fatores selecionados em 20.000 cenários diferentes, usando-se para isto o Microsoft Excel®, com a fórmula **INV.NORM.N (ALEATÓRIO(); "média"; "desvio padrão")** nos índices avaliados. Esta fórmula retorna o inverso da distribuição cumulativa normal para a média específica e o desvio-padrão, apresentando um panorama geral de uma rotina de operação. As estimativas de preços tais como, diesel, manutenção, mão de obra, pneus, materiais de desgaste e outros, foram baseadas em valores médios de mercado. Os dados utilizados não se referem a uma operação específica da empresa, pois foram selecionados e estimados nas bases de dados de forma a preservar dados confidenciais das organizações cedentes, porém sem perder a representação do processo de produção real.

De posse dos dados em todos os cenários citados, procedeu-se a uma regressão linear múltipla de forma a obter um modelo matemático que relacionasse as variáveis predictoras e o preço(custo) do minério lavrado (variável resposta).

Uma vez obtido o modelo de regressão linear múltipla pôde-se, a partir da tabela de ANOVA, mensurar a contribuição percentual de cada variável preditora no preço do minério lavrado, o que nos apresenta assim, de forma inequívoca, qual(is) variável(is) mais influencia(m) sobre o custo do minério lavrado.

Em seguida usando a ferramenta "Otimizador de Resposta" do Minitab® foram realizadas diversas variações de possibilidades das variáveis preditoras e analisou-se o seu impacto no custo da produção. Ao estabelecer uma meta para o orçamento, com base no modelo de regressão, foi calculado o intervalo de confiança (IC) com nível de significância de 5% para estimativa da precificação do minério lavrado.

4 RESULTADOS

Com base em conhecimento prévio, pôde-se indicar com bastante segurança e assertividade quais são os fatores que influenciam a precificação do minério lavrado dentro de um cenário específico qualquer. Esta indicação foi utilizada para acessar as informações armazenadas nos bancos de dados, possibilitando assim, a partir do histórico, a simulação de cenários diversos. Nas subseções seguintes serão apresentadas, na ordem, a discriminação das variáveis levantadas, o processo de regressão linear múltipla a partir dos dados levantados para variáveis preditoras e variável resposta, e por fim, serão exibidas as análises e consequências delas no planejamento de orçamento da empresa.

4.1 Variáveis preditoras e variável resposta

Dentro da operação da empresa¹ a precificação do minério (variável resposta) é influenciada por vários fatores/variáveis. Foram listados aqueles que, ao longo do tempo, se mostraram mais importantes e que têm relação direta com a variável resposta. A seguir expõe-se cada um desses fatores.

REM ⇨ relação estéril/minério, ou seja, é a proporção de estéril que necessita ser deslocada para acessar e movimentar uma certa quantidade de minério. Por exemplo, uma REM igual a 2, significa que temos que deslocar duas toneladas de estéril para acessar uma tonelada de minério. Dessa forma, quanto maior for a REM, mais elevado será o custo para se obter 1 tonelada de minério.

CM ⇨ carga média em toneladas que os caminhões transportam em cada viagem. Quanto maior a carga média, maior será a produtividade da frota, pois ela vai transportar mais material por hora trabalhada, e assim tem-se menor custo.

VM ⇨ velocidade média, em km/h, com que o transporte da carga é feito. Neste sentido, quanto mais próximo da velocidade máxima permitida para o trajeto, maior será a produtividade.

DMT ⇨ distância média de transporte. É a distância média entre o ponto de partida e o ponto de destino do carregamento. Aqui, quanto menor for a distância média de transporte, menor será o custo de produção.

TF-TR ⇨ tempo fixo de transporte. Trata-se dos tempos que teoricamente não variam, tais como, tempo de manobra do caminhão durante a carga e descarga, tempo de basculamento e carregamento e tempo de fila na carga e descarga. Quanto menor for esse tempo, menor o custo de produção, pois estes tempos fazem parte do ciclo normal de operação, porém são tempos de espera e não tempo de deslocamento.

¹ A identidade da empresa será mantida em sigilo por exigência de sua política interna. Vale ressaltar que, conforme mencionado anteriormente, isto em nada interfere na análise, bem como nas conclusões obtidas.

TC-CG ⇨ tempo de ciclo de carga. Compreende o tempo de manobra do caminhão para se posicionar diante da máquina de carga juntamente com o tempo que a máquina gasta para carregá-lo. Neste caso, quanto menor for o tempo, menor será o custo de produção.

OC ⇨ tempo de ociosidade (atraso) no ciclo dos equipamentos. Pode ser pensado como o tempo que um caminhão fica aguardando a disponibilidade de uma máquina de carga ou o tempo que a máquina de carga perde entre um carregamento e outro para acessar o próximo caminhão. É medido como percentual de horas ociosas. Um percentual maior resulta em custo de produção mais elevado.

DF-CG ⇨ disponibilidade física da frota de carga. Trata-se do percentual de horas que o equipamento de carga não se encontra em manutenção.

DF-TR ⇨ disponibilidade física da frota de transporte. Trata-se do percentual de horas que o equipamento de transporte não se encontra em manutenção.

UF-CG ⇨ utilização física dos equipamentos de carga, aqui expressa em percentual de horas que o equipamento foi utilizado (não estando em manutenção).

UF-TR ⇨ utilização física dos equipamentos de transporte, aqui expressa em percentual de horas que o equipamento foi utilizado (não estando em manutenção).

O Quadro 1 traz uma listagem dos fatores dispostos anteriormente, sua sigla e sua breve definição.

Quadro 1 – Fatores (variáveis preditoras) e resposta usados no estudo, sua respectiva sigla de abreviação seguido de uma breve definição.

	Fatores	Abreviação	Definição
Variáveis preditoras ou fatores	A	REM	Relação estéril/minério
	B	CM(ton)	Carga média transportada
	C	VM(km/h)	Velocidade média de transporte da carga
	D	DMT(km)	Distância média de transporte
	E	TF-TR	Tempo fixo de transporte
	F	TC-CG	Tempo de ciclo de carga
	G	OC(%)	Ociosidade
	H	DF-CG(%)	Disponibilidade física da frota de carga
	I	DF-TR(%)	Disponibilidade física da frota de transporte
	J	UF-CG(%)	Utilidade física dos equipamentos de carga
	L	UF-TR(%)	Utilidade física dos equipamentos de transporte

Fonte: Elaborado pelos autores

A variável resposta, a qual será denominada no modelo como y , é a precificação por tonelada de minério (R\$/ton).

Uma vez elencado o grupo de fatores que influem a resposta, utilizou-se a base de dados da empresa para dar prosseguimento ao trabalho. Ao nível de significância de 5% foi verificado que todos os dados dos fatores elencados anteriormente, seguem a distribuição normal. Esses testes de normalidade podem ser facilmente realizados com a maioria dos pacotes estatísticos disponíveis. Neste trabalho foi utilizado o Minitab® para fazer esta e outras análises. Sabendo que todos os fatores seguem a distribuição normal, foram obtidas as estatísticas, média e desvio-padrão, ao longo do tempo, para cada um dos fatores considerados. Com esses dados em mãos foi realizada uma simulação para ampliar o conjunto de dados, e assim, dada a quantidade maior de cenários obtidos, alcançar maior qualidade nas conclusões e decisões a serem tomadas.

Em seguida, usando o Microsoft Excel®, procedeu à simulação de 20.000 dados para cada um dos fatores com base em sua média e desvio-padrão obtidos. Com isso, foi

possível combinar 20.000 dados aleatórios de cada fator com cada outro, acessando então a precificação por tonelada de minério para cada um dos cenários oriundos dessas combinações. Ou seja, a partir de um conjunto de dados em que se tem cada um dos fatores em 20.000 níveis, obteve-se 20.000 respostas dessas combinações. A partir daí, essa enorme quantidade de dados/informação foi resumida através da obtenção de um modelo matemático/estatístico que correlacionasse cada um dos 11 fatores com a resposta de interesse. Com o modelo robusto em mãos, pôde-se então usá-lo para todo um espectro de análises e tomada de decisões, as quais impactam diretamente no orçamento da empresa. A regressão linear múltipla foi utilizada para estimar esse modelo, sendo ela uma das técnicas matemáticas/estatísticas mais importantes que há nos mais diversos campos da ciência.

4.2 O modelo de regressão linear múltipla

Segundo Montgomery e colaboradores (2021), a regressão linear é uma das técnicas mais importantes para se tratar dados obtidos nos mais diversos campos do conhecimento. Sua utilidade se baseia num processo conceitualmente lógico de usar uma equação matemática para expressar a relação entre uma ou mais variáveis de interesse (respostas) e um conjunto de variáveis preditoras. Além disso, a regressão linear traz em seu bojo uma ampla, densa e elegante teoria estatística, a qual nos permite quantificar a incerteza e elevar a qualidade de nossas decisões. Como acontece com a maioria das análises estatísticas, o objetivo da regressão linear é resumir os dados observados da maneira mais simples, útil e eficaz possível (WEISBERG, 2014).

Modelos de regressão múltipla são frequentemente denominados de modelos empíricos ou funções aproximadas. Ou seja, a verdadeira relação funcional entre a variável(is) resposta e seus preditores é desconhecida, porém sobre certos intervalos, as variáveis regressoras do modelo de regressão é uma aproximação adequada da função verdadeira e desconhecida. Box (1979) resumiu tal fato de forma muito sagaz dizendo que "todos os modelos estão errados, mas alguns são úteis".

Em geral a variável resposta (y) em um modelo de regressão relacionada a k regressores ou variáveis preditoras pode ser escrita na Equação (1).

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (1)$$

Na Equação (1) os parâmetros β_j , com $j = 0, 1, \dots, k$; são denominados de coeficientes da regressão. Este modelo representa o que denominamos de hiperplano num espaço de dimensão k das variáveis regressoras x_j . Aqui β_j é interpretado como sendo a taxa de variação esperada na resposta y por unidade de variação em x_i quando todas as outras variáveis regressoras $x_i (i \neq j)$ são mantidas constantes. Por esta razão os parâmetros β_j são frequentemente denominados de coeficientes de regressão parcial.

Os coeficientes β_j são obtidos empregando-se o método dos mínimos quadrados. Sua lógica é expressa como sendo uma técnica que "escolhe" valores de β_j que minimizem a soma dos quadrados (do inglês **Sum of Squares - SS**) dos resíduos deixados pelo modelo em relação a cada ponto experimental. Essa quantidade é expressa de acordo com o somatório apresentado na Equação (2):

$$SS = \sum_{i=1}^n \varepsilon_i^2 \quad (2)$$

onde ε_i é uma medida da discrepância entre o valor previsto e o valor medido (experimental) para o i -ésimo ponto medido (BINGHAM & FRY, 2010).

Um modo mais conveniente e compacto de tratar a obtenção de modelos de regressão linear múltipla é empregar a representação matricial. Isso permite uma exibição relativamente otimizada do modelo, dados e resultados. Em notação matricial, o modelo é expresso segundo a Equação (3).

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{10} & x_{11} & x_{22} & \cdots & x_{1k} \\ x_{20} & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n0} & x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad (3)$$

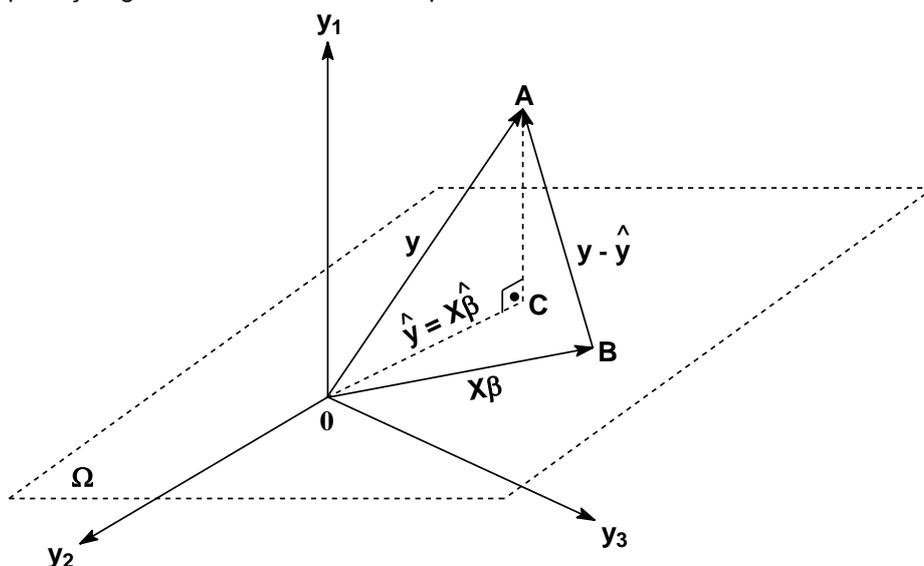
Esta representação matricial da Equação (3) é escrita de forma compacta conforme apresentado na Equação (4).

$$y = X\beta + \varepsilon \quad (4)$$

em que y é um vetor $n \times 1$ das observações realizadas, X é uma matriz $n \times p$ dos níveis das variáveis regressoras, β é um vetor $p \times 1$ dos coeficientes de regressão, e ε é um vetor $n \times 1$ dos erros aleatórios, tal como expresso pela Equação (3).

Em se tratando de regressão linear, uma interpretação geométrica do método dos mínimos quadrados é esclarecedora e útil. Considere que o vetor das observações y define um vetor que sai da origem e vai até o ponto A, mostrado na Figura 1.

Figura 1 – Interpretação geométrica dos mínimos quadrados.



Fonte: Elaborada pelos autores

Neste caso, y constitui as coordenadas de um espaço amostral com n -dimensões, e no caso específico da Figura 1, trata-se de um espaço amostral tridimensional. A matriz X consiste de $p(n \times 1)$ colunas de vetores e cada uma dessas colunas define um vetor desde a origem no espaço amostral. Esses p vetores formam um subespaço de p -dimensões denominado de espaço de estimação (Ω). Na representação da Figura 1 tem-se $p = 2$. É possível representar qualquer ponto neste subespaço como uma combinação linear dos vetores x_0, x_1, \dots, x_k . Sendo assim, qualquer ponto neste subespaço de estimação tem a forma $X\beta$. Deixa-se o vetor $X\beta$ determinar o ponto B em nossa representação.

Matematicamente teremos um sistema linear que não pode ser resolvido exatamente e para o qual é necessário obter alguma solução aproximada. Esse tipo de sistema é muito comum em engenharia e costuma aparecer em aplicações nas quais erros de medição "perturbam" os coeficientes de um sistema consistente a tal ponto que o sistema passa a ser inconsistente. Como não é possível encontrar alguma solução exata, vamos procurar (estimar) um vetor $\hat{\beta}$ que chegue "tão perto quanto possível" de ser uma solução. O ponto de partida será aquele em que o erro nessa aproximação seja o mínimo possível. Conforme já mencionado, isso é conhecido como ajuste linear pelo método dos mínimos quadrados. A seguir será dada uma interpretação e análise geométrica deste método.

Na Figura 1, é possível constatar que o quadrado da distância de **B** até **A** pode ser dado pela Equação (5),

$$SS = (y - X\beta)'(y - X\beta) \quad (5)$$

e, portanto, o procedimento dos mínimos quadrados se aplica quando se escolhe o ponto no espaço de estimação que minimiza essa distância ao quadrado. O ponto em que ocorre a igualdade entre os vetores $\mathbf{BA} = \mathbf{CA}$ é o melhor resultado a ser alcançado, ou seja, quando \mathbf{BA} for perpendicular a Ω . Sendo o ponto **C** definido pelo vetor $y = X\beta$, assim, é possível expressar a condição de mínimo quando se aplica a condição da Equação (6).

$$y - \hat{y} = y - X\hat{\beta} \quad (6)$$

Como $y - \hat{y}$ é perpendicular ao espaço de estimação (Ω), o produto ortogonal entre X e $y - \hat{y}$ é nulo, e conseqüentemente, obtém-se, algebricamente, a Equação (7):

$$X'(y - X\hat{\beta}) = 0 \quad \Rightarrow \quad X'X\hat{\beta} = X'y \quad \Rightarrow \quad \hat{\beta} = (X'X)^{-1}X'y \quad (7)$$

que é a equação normal do método dos mínimos quadrados.

Aplicando a técnica dos mínimos quadrados, representada pela Equação (7), aos dados simulados a partir do levantamento do banco de dados empregado, foi obtido o modelo de regressão linear múltipla apresentado na Equação (8).

$$y = 22,133 + 5,24426 A - 0,302323 B - 0,50498 C + 2,95896 D + 0,60697 E \\ + 0,87004 F + 2,692 G - 0,0671 H - 1,2156 I - 0,5130 J - 4,4887 L \quad (8)$$

Num primeiro momento, a análise do modelo obtido é bastante simples, ou seja, aqueles fatores que possuem coeficientes $\hat{\beta}'_s$ estimados com sinal positivo, ao terem seus valores aumentados, elevam a resposta y . A título de exemplo, vamos considerar o fator A, o qual no Quadro 1 equivale à REM. Caso o nível/valor de A aumenta, este aumento provocará um aumento na resposta y , ou seja, aumenta o custo por tonelada de minério.

Tome agora, como exemplo o fator L, o qual equivale a UF-TR(%) que trata da utilização física dos equipamentos de transporte, expressa em percentual de horas que o equipamento é utilizado. Podemos perceber pelo sinal negativo do coeficiente que antecede o termo L na Equação (8) que, quanto maior o valor de L, menor será o custo por tonelada de minério. Todos os outros coeficientes podem ser analisados de modo similar.

Juntamente com a equação de regressão foi obtido também a tabela de análise de variância (ANOVA) que apresentaremos e discutiremos a seguir.

4.3 Análise de variância (ANOVA) da regressão linear múltipla

A análise de variância (ANOVA) na obtenção e avaliação de modelos estatísticos tornou-se uma ferramenta amplamente utilizada e desempenha um papel fundamental na maior parte da aplicação da estatística atualmente. Em particular, o uso da ANOVA envolvendo efeitos aleatórios encontra ampla aplicação em projetos experimentais numa variedade de campos que requerem quantificação da variabilidade, com especial destaque em controle de qualidade e na engenharia (HIROTSU, 2017). A Tabela 1 apresenta a ANOVA para o modelo que foi exibido na subseção anterior.

Tabela 1 – Análise de Variância (ANOVA) para o modelo de precificação de minério.

Fonte	GL	SQ Seq	Contribuição	SQ (Aj.)	QM (Aj.)	Valor F	Valor-P
Regressão	11	37996,7	98,74%	37996,7	3454,2	142113,95	0,000
A	1	11863,9	30,83%	12288,5	12288,5	505569,52	0,000
B	1	7065,8	18,36%	7193,6	7193,6	295955,90	0,000
C	1	5232,1	13,60%	5109,2	5109,2	210199,90	0,000
D	1	7819,6	20,32%	7865,6	7865,6	323606,34	0,000
E	1	4944,1	12,85%	910,9	910,9	37475,24	0,000
F	1	889,2	2,31%	895,4	895,4	36837,62	0,000
G	1	9,1	0,02%	8,2	8,2	337,14	0,000
H	1	0,1	0,00%	0,0	0,0	1,48	0,224
I	1	11,2	0,03%	11,7	11,7	483,07	0,000
J	1	2,4	0,01%	2,1	2,1	85,56	0,000
L	1	159,1	0,41%	159,1	159,1	6546,61	0,000
Erro	19988	485,8	1,26%	485,8	0,0		
Total	19999	38482,5	100,00%				

Fonte: Elaborada pelos autores

A Tabela 1 traz várias grandezas estatísticas interrelacionadas, porém serão destacados apenas alguns resultados de ordem prática e que aqui são importantes. O primeiro deles diz respeito à coluna que contém de Valor-P. Nela nota-se que todos os fatores, à exceção do valor (em vermelho) correspondente ao fator H, são significativos para o modelo de regressão ao nível de significância de 5%. Este resultado de não significância de H não quer dizer que não seja uma variável importante na precificação, mas tão somente que este fator não provoca variabilidade significativa na resposta, ou seja, não está provocando variação no preço de produção de minério. O que isto está querendo nos informar é que a disponibilidade física da frota de carga não sofre variação ao longo do tempo, isto é, ela é historicamente regular.

Outras informações importantes podem ser obtidas a partir da coluna de "Contribuição" na Tabela 1. Algumas de delas são:

1ª: o valor de contribuição da regressão (também conhecido como R^2) é de 98,74%, o que significa que 98,74% (em cor violeta) da variabilidade da resposta é explicada a partir dos fatores de A até L;

2ª: o complementar desse valor, 1,26% (em cor violeta), está atribuído a erros aleatórios inerentes ao processo;

3ª: os 98,74% de explicação da variabilidade é então dividido entre os fatores de A até L de acordo com a sua importância na resposta. Desses 98,74% de variação explicada, 30,83% (em cor azul), o maior valor, é atribuído ao fator A (REM na Tabela 1), sendo esta a variável mais importante na precificação de produção de minério.

4.4 Otimização da resposta para fins de planejamento

A partir do modelo de regressão, usamos a ferramenta “Otimizador de Resposta” do Minitab® para obtenção de um gráfico de otimização em que é possível atuar de forma dinâmica sobre ele. Durante a otimização, pode-se estabelecer como objetivo sobre a resposta do modelo: minimizar, maximizar ou atingir um “Valor alvo”. Neste estudo, buscou atingir um valor alvo numericamente igual ao orçamento que se tinha disponível. Sendo assim, supondo que o teto do orçamento fosse de R\$20,00 por tonelada de minério² (valor alvo), foi alcançado pelo processo de otimização (com 95% de confiança) que a meta será atingida. O gráfico de otimização é apresentado na Figura 2.

Figura 2 – Gráfico de otimização obtido através do modelo de regressão no Minitab®.



Fonte: Elaborada pelos autores a partir do Minitab®

Na parte superior, em vermelho, da Figura 2 é determinado o nível médio esperado para cada um dos fatores necessários ao atingimento do valor de orçamento do período estipulado como valor alvo — de A a L — com base nas informações do modelo. Abaixo e acima do valor em vermelho, são descritos os valores mínimo e máximo esperados, respectivamente, para cada fator de acordo com os dados. Na parte inferior a cada fator é possível, visualmente, avaliar a influência relativa de cada fator na resposta. Quanto maior a inclinação da reta em relação à linha pontilhada, maior a influência sobre o custo. Por outro lado, quanto mais próxima da horizontal, menor é a influência do fator na variabilidade da resposta. A inclinação, positiva ou negativa, está diretamente relacionada ao sinal do coeficiente para o fator na Equação (8), já discutido na subseção 4.2.

Além dessas informações, o gráfico no Minitab® é dinâmico. É permitido ajustar instantaneamente o nível de cada fator (dígitos em vermelho), e assim, verificar qual o impacto esta mudança terá no orçamento (resposta). Essa informação então é prontamente utilizada para avaliação e tomada de decisão no sentido de intervir sobre o processo de modo a ajustá-lo, caso necessário, buscando sempre atingir a meta limitada pelo orçamento.

² Os dados reais foram “descharacterizados” por se constituírem em informações que a política da empresa exige que sejam mantidas em sigilo. Porém, os ajustes foram realizados pelos próprios autores deste trabalho de forma a garantir aqui as mesmas conclusões que foram tiradas quando da realização primeira do estudo.

Por fim, a partir da análise de otimização, obtém-se o intervalo de confiança para a variável y (custo por tonelada), que neste caso foi de (19,99; 20,01), com 95% de confiança, ou seja, se está com 95% confiança de que a verdadeira média para o orçamento, com base nas informações e modelo levantados, se encontra entre 19,99 e 20,01. Isto dá uma segurança importante quanto ao aspecto gerencial do processo, bem como de atingimento das metas de produção estipuladas e esperadas.

5 CONSIDERAÇÕES FINAIS

Neste estudo resulta demonstrado o quão importante e benéfica pode ser para a indústria, a aproximação com a universidade no sentido de encontrar soluções apropriadas para problemas presentes no seu dia a dia. A sistemática proposta e aplicada a partir da apresentação do problema se mostrou adequada e eficaz para resolver o problema de planejamento e administração do orçamento na empresa de mineração. Vale ressaltar que, se já na oferta de uma disciplina num programa de Mestrado em Engenharia surgiram resultados tão positivos, muito mais ainda pode ser alcançado em termos de diminuição de custos e aumento da produtividade quando da aproximação sinérgica universidade-empresa.

Diante da constatação de que a aproximação academia-indústria traz *insights* e resultados tão interessantes de aplicação do conhecimento, espera-se que ações como esta sejam determinantes para estabelecer uma "consciência empresarial" no sentido de que vale à pena estar mais próximo da academia, não só para formar seus quadros de mão-de-obra, mas também como uma possibilidade de que haja, por parte da indústria, contribuições no sentido de financiar projetos, auxiliar a aquisição de equipamentos, ampliar espaços, aumentar capacidade de atendimento das instituições etc., resultando em benefícios que seguramente em curto e longo tempo, poderão ser revertidos para a própria indústria, o que assim acaba por beneficiar ambas as partes.

Por fim, espera-se que a prática já há muito tempo comum na indústria que é a de adaptar ou comprar a solução para um determinado problema, seja pelo menos em parte substituída pelo desejo de produzir respostas próprias às suas questões. Neste sentido, a universidade com seus centros de excelência tem muito a contribuir numa configuração de parceria intelectual.

Agradecimentos

Os autores agradecem ao Programa de Pós-Graduação em Engenharia de Minas – Mestrado Profissional de nossa instituição, pelo apoio ao desenvolvimento deste trabalho e à indústria mineradora pela concordância em publicar este trabalho.

REFERÊNCIAS

BINGHAM, Nicholas; FRY, John. **Regression: Linear Models in Statistics**. London: Springer-Verlag, 2010.

BORGES, Thiago, C. **Análise dos Custos Operacionais de Produção no Dimensionamento de Frotas de Carregamento e Transporte em Mineração** 2013. 98 f. Dissertação (Mestrado em Engenharia de Minas) – Universidade Federal de Ouro Preto, Ouro Preto, 2013.

BOX, George. Robustness in the strategy of scientific model building. In: R. Launer & G. Wilkinson (Eds.). **Robustness in statistics**. New York: Academic Press, 1979. p. 201–235.

CRUZ, Carlos Henrique de Brito. Indicadores sobre Interação Universidade-Empresa em Pesquisa em São Paulo. In: MARCOVITCH, Jacques (Org.). **REPENSAR A UNIVERSIDADE: Desempenho Acadêmico e Comparações Internacionais**. São Paulo: Com-Arte; Fapesp, 2018. p. 187-200.

DAGNINO, Roberto. Como é a universidade de que o Brasil precisa? **Avaliação**, Campinas; Sorocaba, SP, v. 20, n. 2, p. 293-333, jul. 2015

DAGNINO, Roberto. **Future-se não vai dar certo porque a empresa brasileira não está interessada na universidade**. [Entrevista cedida a] Luís Eduardo Gomes. Sul21, Porto Alegre, set. 2019. Disponível em: <https://www.sul21.com.br/entrevistas-2/2019/09/future-se-nao-vai-dar-certo-porque-a-empresa-brasileira-nao-esta-interessada-na-universidade/>. Acesso em: 19 abr. 2021.

GOLDEMBERG, José. Ciência, Desenvolvimento e Universidade. In: MARCOVITCH, Jacques (Org.). **REPENSAR A UNIVERSIDADE: Desempenho Acadêmico e Comparações Internacionais**. São Paulo: Com-Arte; Fapesp, 2018. p. 21-29.

HIROTSU, Chihiro. **Advanced Analysis of Variance**. New Jersey: John Wiley and Sons, Inc., 2017.

MONTGOMERY, Douglas C.; PECK, Elizabeth A.; VINING, Geoffrey. **Introduction to Linear Regression Analysis**. 5th. ed. New York: John Wiley and Sons, Inc., 2021.

NETO, Arnaldo. T. T. **MÉTODOS DE CUSTEIO: Identificação do Método utilizado pela Companhia Brasileira de Mineração e Metalurgia** 2019. 83 f. Monografia (Ciências Contábeis) – Pontifca Universidade Católica de Minas Gerais, Belo Horizonte, 2019.

PINHO, Marcelo. Mais do que se supõe, menos do que se precisa: relações entre universidades e empresas no Brasil. In: GARCIA, Renato et. al. (Org.). **Estudos de caso da interação universidade-empresa no Brasil**. Belo Horizonte: FACE/UFMG, 2018. p. 35-57.

RIBEIRO, Darcy. **A universidade necessária**. 2. ed. Rio de Janeiro: Paz e Terra, 1969.

THUMS, Alice; MARETH, Taciana. **Contribuição das Práticas de Engenharia Reversa para a Gestão Estratégica de Custos**. XXVII Congresso Brasileiro de Custos – Associação Brasileira de Custos, 2020.

WEISBERG, Sandford. **Applied Linear Regression**. 4th. ed. New York: John Wiley and Sons, Inc., 2014.

**MULTIPLE LINEAR REGRESSION AS AN INDUSTRIAL PLANNING STRATEGY: AN
EXAMPLE OF KNOWLEDGE TRANSFER BETWEEN UNIVERSITY AND INDUSTRY**

Abstract: *This work presents as an industrial problem that emerged during the offer of a discipline in a Professional Master's course in Mining Engineering, it could be solved from the synergistic interaction between teachers of the discipline and students from a mining company in the city of Araxá-MG. The problem proposed here posed the need to obtain a mathematical model that related a set of predictive variables with the pricing of the ton of ore explored by the company, in order to improve the quality of the decisions made during the planning and management of the annual budget. In this sense, a set of historical data was obtained, processed statistically and from the statistics, mean and standard deviation, a simulation of a total of 20,000 different scenarios was carried out according to 11 regressor variables. From this set of scenarios, the "least squares method" was used to obtain the regression equation using the Minitab® software. With the model obtained and its ANOVA table it was possible, among other things, to quantify the contribution of each predictor variable on the response variable (ore pricing), as well as use it to make predictions and decision-making about planning, productivity, and cost in the projects of that industry. Finally, this relationship proved to be extremely productive and an indication of the importance of bringing universities and industry closer together, as a possibility of producing interesting solutions within the company itself.*

Keywords: *University-industry interaction, linear regression, mining.*