



Uso de regressão linear para investigação de baixo rendimento de lotes produzidos em uma indústria farmacêutica

João Cláudio Pinheiro Batista

Introdução

Modelos de regressão são modelos matemáticos que relacionam o comportamento de uma variável Y com outra X. Quando a função f que relaciona duas variáveis é do tipo $f(X) = a + bX$ temos o modelo de regressão simples. A variável X é a variável independente da equação enquanto $Y = f(X)$ é a variável dependente das variações de X. O modelo de regressão é chamado de simples quando envolve uma relação causal entre duas variáveis. O modelo de regressão é multivariado quando envolve uma relação causal com mais de duas variáveis. Isto é, quando o comportamento de Y é explicado por mais de uma variável independente X_1, X_2, \dots, X_n . [1].

O modelo matemático de regressão linear simples foi aplicado em um site de produção de insumos farmacêuticos, objetivando o aumento no rendimento nos lotes produzidos de determinada enzima de natureza proteolítica. O processo de produção fermentativo era constituído de duas fases distintas, onde a primeira fase era a obtenção de massa celular como controle de determinadas condições de cultivo e a segunda fase era a obtenção da enzima proteolítica também como controle de determinados parâmetros de controle. Em um período da produção desta enzima, o rendimento esperado dos lotes estava abaixo da média histórica, desta forma a aplicação do modelo de regressão linear foi capaz de corrigir o desvio sem necessidade de investimento na planta de produção.

Material e métodos

Os dados amostrais para construção do modelo de regressão linear simples, foram obtidos através do banco de dados dos registros de lotes produzidos na unidade fabril de insumos farmacêuticos de uma indústria localizada na cidade de Montes Claros, estado de Minas Gerais. Os dados foram provenientes da produção de uma determinada enzima obtida por processo fermentativo (produção de proteína recombinante), onde foram avaliados os parâmetros pH do processo fermentativo, concentração enzimática ao término do processo fermentativo, parâmetros de crescimento de biomassa.

Toda a análise estatística desta pesquisa foi realizada com o auxílio do software estatístico Minitab®, na versão 14.0 [2]. Durante a análise dos dados, os parâmetros de controle do processo fermentativo que não influenciavam estatisticamente a produção da enzima, foram retirados do modelo matemático.

Ao término da análise estatística dos dados amostrais e construção de um modelo de regressão linear, os novos parâmetros de controle das condições de cultivo e obtenção da enzima proteolítica foram informados para a linha de produção, onde os novos lotes foram produzidos na rotina de produção.

Resultados

A Fig. 1 apresenta os dados históricos da concentração enzimática obtida ao término do processo fermentativo. O gráfico Produção – período estável, mostra os dados históricos da produção de enzima em um período de boa estabilidade do processo. A média histórica dos lotes produzidos era de 22,03 AE/ml (atividade enzimática por mililitro). O gráfico Produção período instável apresenta os dados de produção da enzima em um período de perda da estabilidade do processo, nota-se que a média obtida na produção da enzima era de 17,84 AE/ml.

A Fig. 2 apresenta o modelo de regressão linear para o parâmetro que estatisticamente influenciou o rendimento dos lotes produzidos no período de boa estabilidade do processo. Nota-se que o modelo proposto foi capaz de responder a 70% de ótimos resultados produzidos. Cada ponto na figura representa um lote produzido.

O parâmetro de processo que foi estatisticamente significativo para o aumento na concentração da atividade enzimática está associado com valores de pH na fase de obtenção da enzima. O modelo de regressão linear mostrou que valores de pH entre 7,20 a 7,50.

A Fig. 3 apresenta o gráfico comparativo da produção antes e depois da aplicação do modelo de regressão linear. Observa-se que houve um aumento de cerca de 25% no rendimento de cada lote.



Discussão

A faixa de pH onde os melhores resultados da concentração enzimática foram obtidos é explicado pelo valor de pH ótimo para promover a formação da cadeia políptica da proteína na sua conformação nativa [3]. Ao comparar as médias das concentrações da atividade enzimática antes e depois da aplicação do modelo de regressão linear, é possível notar um aumento de cerca de 25% no rendimento dos lotes. Nota-se que pelo coeficiente de determinação, a variabilidade variável pH é responsável por 70% da variabilidade da variável atividade enzimática.

O modelo de regressão satisfaz a relação pH e atividade enzimática, pois o nota-se que o mesmo é significativo ao nível de 5%.

Os maiores rendimentos obtidos foram observados quando o parâmetro pH compreende a faixa de 7,25 a 7,50, conforme observado na Fig.2.

O aumento de 25% no rendimento dos lotes após a nova configuração de valores de pH conforme foi evidenciado na Fig.2 é estatisticamente diferente. Na Fig.3 esta afirmação fica evidente ao comparar os dados da produção para ambos os períodos. Observa-se um deslocamento da média e uma diminuição na variabilidade dos resultados indicado pelas linhas vermelhas que indicam os limites inferior e superior de controle estatístico.

Ao término da produção de lotes após a fixação de nova faixa de trabalho para o parâmetro pH, um teste de hipótese foi realizado sendo confirmado a diferença estatística significante.

Considerações finais

O resultado da pesquisa mostrou o que parâmetro pH exerce forte influência na produção da enzima pelo processo fermentativo. Modernas técnicas de produção de proteínas recombinantes podem ser melhoradas com uso de modelagem matemática, onde é possível realizar melhorias em processos com baixo custo de investimento.

A análise de regressão linear se mostrou eficiente ao analisar os efeitos de uma variável X sobre uma variável Y [4].

Referências

- [1] MONTGOMERY, Douglas C. Introdução ao controle estatístico da qualidade. 4. ed. Rio de Janeiro: LTC, 2001
- [2] CAMPOS, Marco Siqueira. Desvendando o MINITAB – versão 14, Rio de Janeiro, RJ: Qualitymark, 2009.
- [3] LEHNINGER, A. L. Princípios de bioquímica. São Paulo: Sarvier, 2011.
- [4] UNIVERSIDADE DE SÃO PAULO, USP [online]. Notas de aula da disciplina estatística aplicada. Homepage: http://www.usp.br/fau/cursos/graduacao/arg_urbanismo/disciplinas/aut0516/Apostila_Regressao_Linear.pdf. Acesso em 13/08/2015

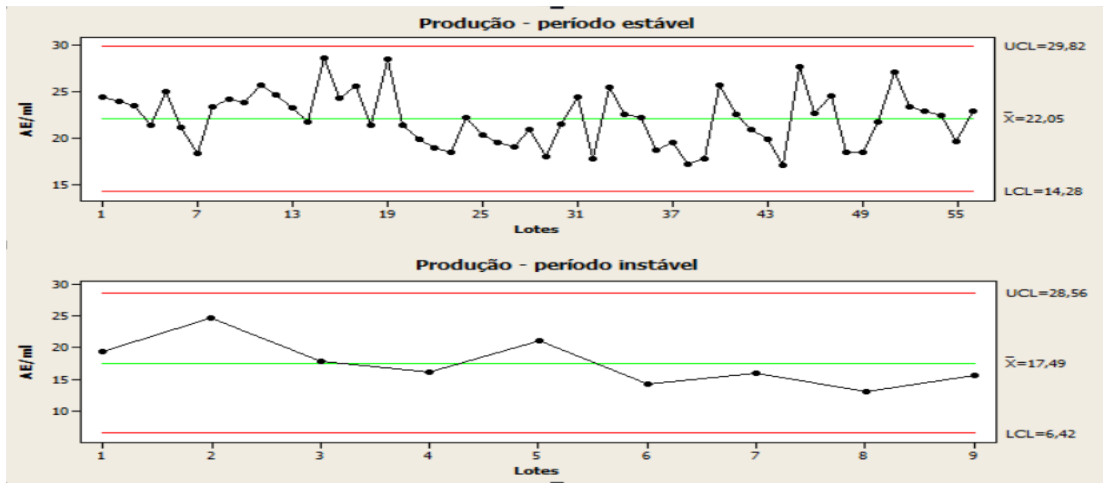


Figura 1- Produção nos períodos: estável e instável respectivamente.

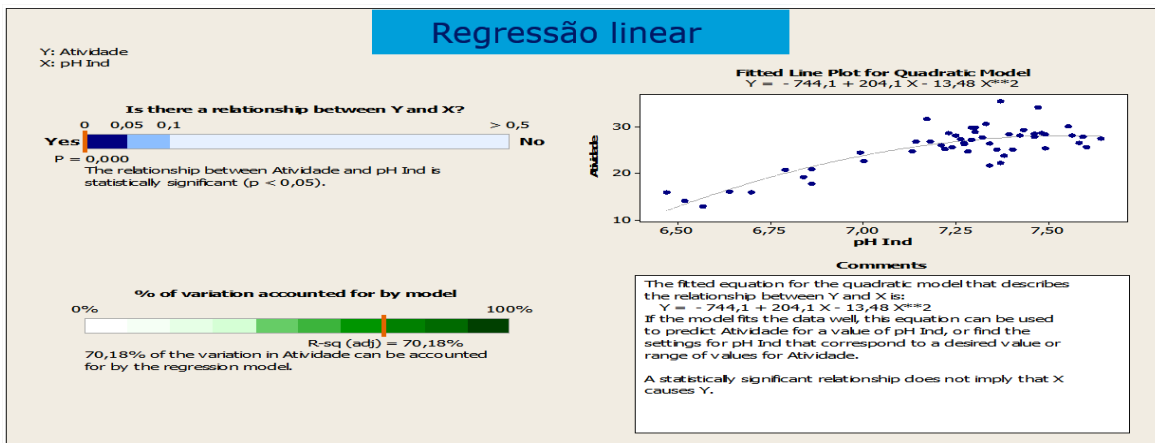


Figura 2 – modelo de regressão linear simples para o parâmetro pH e atividade enzimática

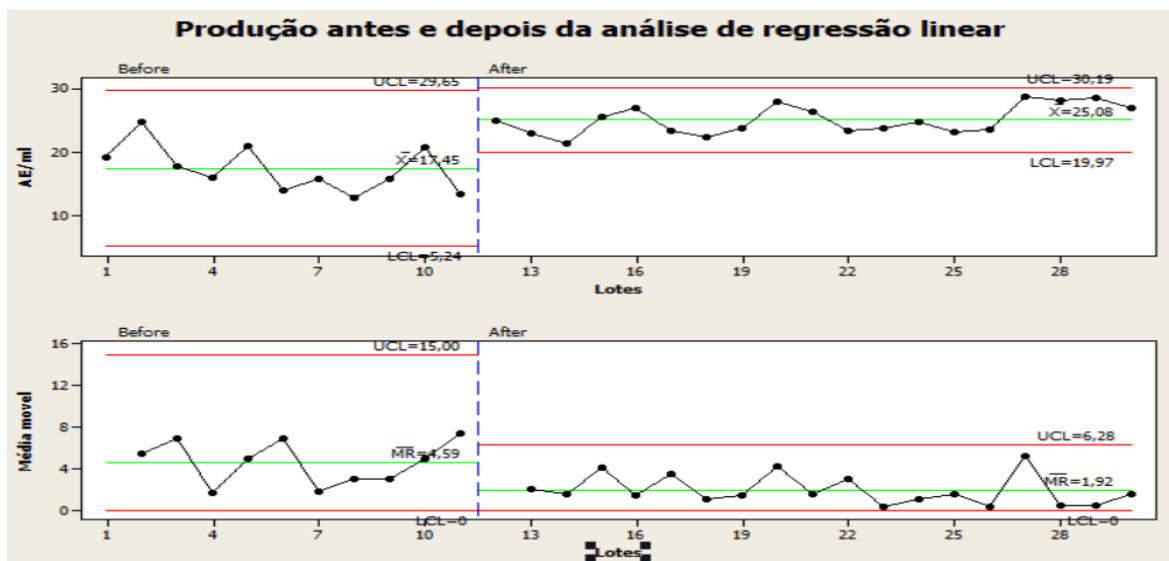


Figura 3 – Produção antes e depois da modelagem, note-se que a media dos resultados foi aumentada em 25%.