

INFLUÊNCIA DOS ERROS DE MEDIDAS NA CONSTRUÇÃO DOS LIMITES DE CONTROLE DE UM GRÁFICO DE CALIBRAÇÃO FUNCIONAL, COM A VARIÂNCIA DOS ERROS SUPOSTAMENTE CONHECIDA

Manoel DOMINGOS FILHO¹
Robert Wayne SAMOHYL²
Eustáquio José MACHADO¹
Raimundo Nonato Castro da SILVA¹
Altemir da Silva BRAGA¹

- RESUMO: O objetivo deste artigo é verificar a influência que os erros de medidas exercem na obtenção dos limites de controle, quando se constrói um gráfico para monitorar o comportamento ou o relacionamento simultâneo entre duas variáveis correlacionadas linearmente. Particularmente será utilizado nesse artigo, o gráfico do controle de calibração linear funcional, com a suposição de que a variância dos erros de medida seja conhecida. Um modelo de calibração nada mais é do que um modelo de regressão inversa. No modelo de calibração funcional temos como principal característica o fato da variável independente X ser considerada uma constante fixa ou variável controlada. Porém, os valores dessa variável X poderão ou não conter erros de mensuração. Essas duas situações serão consideradas, além de se considerar também a magnitude da variância desses erros de medida.
- PALAVRAS-CHAVE: Calibração; estimadores; gráfico de controle.

1 Introdução

Em controle estatístico de processos, quando se deseja monitorar duas ou mais características da qualidade, os gráficos de controle univariados são geralmente

¹Universidade Federal do Acre – UFAC, Rio Branco, AC, Brasil. E-mail: *manoelufsc@yahoo.com.br* / *chefinhoo@globo.com* / *nonato@ufac.br* / *altemirbraga@ufac.br*

²Universidade Federal de Santa Catarina – UFSC, Departamento de Engenharia de Produção e Sistemas, Trindade, SC, Brasil. E-mail: *samohyl@deps.ufsc.br*

ineficientes e ineficazes, quanto à capacidade de perceber mudanças que ocorrem simultaneamente nas variáveis. Os gráficos de controle multivariados constituem-se então numa ferramenta muito eficaz em captar as alterações que podem ocorrer numa ou mais variáveis. Um dos problemas na construção desse tipo de gráfico é quando temos mais de duas variáveis a serem monitoradas conjuntamente, o que torna difícil sua construção. Outra dificuldade que pode ocorrer é quando as variáveis são correlacionadas. Nesse caso, o cálculo da probabilidade de se ter um alarme falso também não é simples de se obter. Uma alternativa muito interessante para monitorar o comportamento de duas ou mais variáveis correlacionadas, é a construção de gráficos de controle de regressão. Quando se têm duas variáveis correlacionadas linearmente, então podemos usar o gráfico de controle de regressão apresentado por DiPaola (1945). Porém, o gráfico de controle de regressão, não considera os possíveis erros de medidas na variável independente X . Nesse artigo, vamos modelar os gráficos de controle utilizando uma função de regressão inversa, também chamada de calibração. Nesse tipo de regressão a variável resposta a ser determinada de forma indireta é a variável independente X , enquanto que os valores da variável dependente Y são medidos diretamente.

Dessa forma, conduziu-se esse trabalho, com o objetivo de propor modificações nos os gráficos de controle utilizando uma função de regressão inversa, também chamada de calibração e avaliar seus desempenhos por meio de simulação Monte Carlo.

2 Metodologia

Para obter os resultados apresentados, neste estudo, os erros de medidas ε e δ foram simulados de uma distribuição normal com média zero ($\mu = 0$) e variâncias σ_ε^2 e σ_δ^2 . Para isso, definiu-se a regressão inversa ou modelo de regressão, inicialmente e, em seguida, foi apresentado os estimadores de máxima verossimilhança (definido detalhadamente no item 2.1), para os parâmetros do modelo de calibração, além dos estimadores clássicos da quantidade desconhecida da variável independente e, ainda, foi apresentado a estimativa da variância e do erro padrão para a variável resposta. Os limites de controle foram construídos de forma analítica e para avaliar seus desempenhos realizou-se simulação para diferentes valores da variância dos erros.

2.1 Método de estimação de máxima verossimilhança

Dentre os vários métodos de estimação disponíveis na literatura estatística, tem-se um método bastante conhecido denominado “método de máxima verossimilhança”. Esse método consiste, basicamente, em retirar uma amostra aleatória X_1, X_2, \dots, X_n , de uma população que possui uma função densidade de probabilidade (fdp) associada a ela designada por $f(x, \theta)$, a qual depende somente dos dados amostrais e do vetor de parâmetros denominado por θ . A partir daí, uma função de densidade de probabilidade conjunta é definida

pelo produto das densidades de cada uma das observações amostrais, ou seja, $f(x_1, \theta) \cdot f(x_2, \theta) \cdot \dots \cdot f(x_n, \theta)$ ou simplesmente, $\prod_{i=1}^n f(x_i, \theta)$ em que θ representa um vetor de parâmetros que é fixo e X_i é considerada uma variável aleatória.

A função de densidade de probabilidade conjunta é conhecida como função de verossimilhança e usualmente representada por: $L(x_i, \theta) = \prod_{i=1}^n f(x_i, \theta)$.

O método de máxima verossimilhança consiste em obter um vetor $\hat{\theta}$ que maximiza a função de verossimilhança. Observa-se que antes da retirada da amostra, cada observação é considerada como uma variável aleatória sendo sua fdp igual a fdp da população de onde se retirou a amostra. Nesta situação, o vetor de parâmetros θ é considerado fixo, enquanto que X_i é uma variável aleatória.

Porém, após a retirada da amostra, os valores de X_i passam a ser conhecidos, ou seja, são valores fixos e por conseguinte a fdp conjunta passa a depender somente do vetor de parâmetros θ . Neste caso, θ passa a ser um conjunto de variáveis. Como $\hat{\theta}$ maximiza a função de verossimilhança, deve-se ter:

$$L(x_i, \hat{\theta}) > L(x_i, \hat{\gamma}),$$

em que $\hat{\gamma}$ é um estimador obtido por qualquer outro método de estimação.

2.2 Regressão inversa

Nesse tipo de regressão a variável resposta a ser determinada de forma indireta é a variável independente X, enquanto que os valores da variável dependente Y são medidos diretamente. Formalmente, o modelo de calibração pode ser definido da seguinte forma:

$$Y_i = \alpha + \beta X_i + \varepsilon_i, \quad i = 1, \dots, n \quad (1)$$

$$Y_{0i} = \alpha + \beta X_0 + \varepsilon_{0i}, \quad i = n + 1, \dots, n + k, \quad (2)$$

em que os erros $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ e $\varepsilon_{1n+1}, \varepsilon_{1n+2}, \dots, \varepsilon_{1n+k}$, são mutuamente independentes e identicamente distribuídos com média zero e variância σ^2 . X_1, X_2, \dots, X_n , são consideradas constantes conhecidas enquanto que α, β, X_0 e σ^2 são parâmetros desconhecidos. Neste caso, os valores de X são obtidos supostamente sem erros de mensuração. Entretanto, se considerarmos que os valores de X contenham erros de medidas, então o modelo de calibração linear pode ser definido pelas expressões:

$$\begin{aligned} y_i &= \alpha + \beta U_i + \varepsilon_i, & i = 1, 2, \dots, n \\ X_i &= U_i + \delta_i, & i = 1, 2, \dots, n \\ Y_{0i} &= \alpha + \beta U_0 + \varepsilon_{0i}, & i = n + 1, \dots, n + k, \end{aligned} \quad (3)$$

em que os erros $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ e $\varepsilon_{1n+1}, \varepsilon_{1n+2}, \dots, \varepsilon_{1n+k}$ são considerados mutuamente independentes e identicamente distribuídos com média zero e variância σ_ε^2 . Os erros $\delta_1, \delta_2, \dots, \delta_n$ também são mutuamente independentes e identicamente distribuídos com média zero e variância σ_δ^2 .

As duas primeiras equações do modelo (3) assim como a Equação 1, dizem respeito à primeira etapa de calibração, enquanto que a última equação de (3) e a Equação 2 só dizem respeito à segunda etapa de calibração.

O modelo definido pela expressão (3) pode ser considerado como um modelo de calibração funcional, que se caracteriza essencialmente, pelo fato das quantidades U_i , $i = 1, 2, \dots, n$, serem consideradas como constantes fixas ou variáveis controladas. Há que se perceber que no modelo de calibração funcional, o número de parâmetros cresce com o tamanho da amostra n e os $n + 4$ parâmetros do modelo são: $\alpha, \beta, \sigma_\varepsilon^2, \sigma_\delta^2, U_1, U_2, \dots, U_n$ em que $\alpha, \beta, \sigma_\varepsilon^2, \sigma_\delta^2$ são os parâmetros estruturais do modelo e os n últimos são os parâmetros incidentais. Estas nomenclaturas foram dadas por Neyman e Scott (1951), em que eles analisaram situações as quais o número de parâmetros aumenta com o tamanho da amostra.

Quando os dados possuem distribuição normal, o modelo de calibração funcional tem o inconveniente de que o logaritmo da função de verossimilhança é ilimitado, como mostrado em Bolfarine *et al.* (1999). Assim, quando o logaritmo da função verossimilhança se torna ilimitado, significa que não se pode encontrar um valor máximo para a função, chamado de máximo global. Desse modo, o que se consegue são os pontos de sela denominados máximos locais. Portanto, torna-se necessário fazer suposições adicionais ao modelo para torná-lo factível, isto é, para que se consiga fazer com que a função log-verossimilhança seja limitada e, portanto, admita um único valor que a maximiza. As suposições adicionais mais utilizadas na literatura são as seguintes:

- (a) Conhecer a razão das variâncias $\lambda = \sigma_\varepsilon^2 / \sigma_\delta^2$;
- (b) Supor que σ_ε^2 e σ_δ^2 sejam conhecidas;
- (c) Supor que σ_ε^2 e σ_δ^2 sejam conhecidas.

Dentre as diversas suposições já contempladas na literatura, será utilizada nesse artigo, a condição de se conhecer a variância dos erros de medida σ_δ^2 .

Considerando inicialmente os modelos (1) e (2) com a suposição de que os dados sejam normalmente distribuídos, tem-se que a função de verossimilhança é dada por:

$$L(G) = \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{n+k}{2}} \exp -\frac{1}{2\sigma^2} \left\{ \sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2 + \sum_{i=n+1}^{n+k} (Y_{0i} - \alpha - \beta X_0)^2 \right\},$$

em que $L(G)$ representa $L(\alpha, \beta, \sigma^2, X_0/X_1, \dots, X_n, Y_1, \dots, Y_{0n+k})$.

Igualando-se a zero as derivadas parciais do logaritmo desta função em relação aos parâmetros, obtemos os estimadores de máxima verossimilhança os quais são:

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X},$$

$$\hat{\beta} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2},$$

$$\hat{\sigma}^2 = \frac{1}{n+k} \left[\sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta}X_i)^2 + \sum_{i=n+1}^{n+k} (Y_{0i} - \bar{Y}_0)^2 \right],$$

$$\hat{X}_{0C} = \frac{\bar{Y}_0 - \hat{\alpha}}{\hat{\beta}}, \hat{\beta} \neq 0;$$

em que $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$; $\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n}$ e $\bar{Y}_0 = \frac{\sum_{i=n+1}^{n+k} Y_{0i}}{k}$.

O estimador da quantidade desconhecida X_0 denotado por \hat{X}_{0C} é conhecido na literatura como estimador clássico, e sua variância segundo Lima (1996) é dada por:

$$Var(\hat{X}_{0C}) = \frac{\sigma^2}{\beta^2} \left[\frac{1}{n} + \frac{1}{k} + \frac{3\sigma^2}{kA\beta^2} + \frac{(\bar{X} - X_0)^2}{A} \right]. \quad (4)$$

A metodologia pela qual os gráficos de controle são construídos demanda a existência de três elementos essenciais os quais são: a linha central do gráfico que corresponde à média do processo; uma linha inferior e paralela à linha central que corresponde ao limite inferior de controle e uma linha superior e paralela à linha central que corresponde ao limite superior de controle. Essas linhas paralelas são construídas, de forma equidistantes à linha central, sendo que essa distância é medida em quantidade de desvios padrão do estimador de interesse. Vale ressaltar ainda que no caso dos gráficos de controle que utilizam modelos de regressão linear como proposto inicialmente por DiPaola (1945) e Mandel (1969), não possuem média constante, pois a linha central do gráfico que representa a média segue a nuvem de pontos distribuídos num diagrama de dispersão.

Portanto, para construir um gráfico de controle que utilize o modelo de calibração linear funcional, necessitamos agora tão somente obter o erro padrão do estimador \hat{X}_{0C} , o que pode ser feito, tomando a raiz quadrada da expressão da variância que está representada pela Equação 4. Assim, os elementos necessários à construção do gráfico de controle são dados por:

$$\begin{aligned} LSC_{\hat{X}_{0C}} &= \hat{X}_{0C} + q \sqrt{\frac{\sigma^2}{\beta^2} \left[\frac{1}{n} + \frac{1}{k} + \frac{3\sigma^2}{kA\beta^2} + \frac{(\bar{X} - X_0)^2}{A} \right]} \\ LM_{\hat{X}_{0C}} &= \hat{X}_{0C} \\ LIC_{\hat{X}_{0C}} &= \hat{X}_{0C} - q \sqrt{\frac{\sigma^2}{\beta^2} \left[\frac{1}{n} + \frac{1}{k} + \frac{3\sigma^2}{kA\beta^2} + \frac{(\bar{X} - X_0)^2}{A} \right]}, \end{aligned} \quad (5)$$

sendo que q representa a quantidade de desvios padrão desejada pelo pesquisador.

Contudo, os estimadores necessitam satisfazer algumas propriedades estatísticas, para que se possam oferecer resultados mais robustos e confiáveis. Dentre as propriedades existentes na literatura estatística, é sempre bom ter estimadores que sejam não-viesados e que sejam consistentes (Bussab e Morettin, 2002). Às vezes o estimador não satisfaz às duas propriedades supracitadas. Não satisfazer a propriedade de não viés é algo que pode ser contornado, controlando o viés do estimador. Contudo, a consistência do estimador é algo que o pesquisador não deve abrir mão, e deve se esforçar para obtê-la. Assim, considerando o modelo definido pelas equações dadas em (3), com as suposições

$$\begin{pmatrix} \delta_i \\ \varepsilon_i \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} \sigma_\delta^2 & 0 \\ 0 & \sigma_\varepsilon^2 \end{pmatrix} \right), i = 1, \dots, n, \quad (6)$$

$$\varepsilon_{0i} \sim N(0, \sigma_\varepsilon^2), i = n + 1, \dots, n + k, \quad (7)$$

e que σ_δ^2 seja conhecida, tem-se que os estimadores de máxima verossimilhança dos parâmetros α e β , são estimadores fortemente consistentes, α e β são dados por:

$$\begin{aligned} \hat{\alpha}_C &= \bar{Y} - \hat{\beta}_C \bar{X} \\ \hat{\beta}_C &= \frac{S_{xy}}{S_{xx} - \sigma_\delta^2}. \end{aligned}$$

Além desses dois estimadores fortemente consistentes, temos ainda os seguintes estimadores consistentes, de máxima verossimilhança:

$$\begin{aligned} \hat{\sigma}_\varepsilon^2 &= \frac{1}{2}(S_{yy} + S_{y_0y_0} - \hat{\beta}S_{xy}) \\ \hat{\sigma}_u^2 &= S_{xx} - \sigma_\delta^2 \\ \hat{U}_{0C} &= \frac{\bar{Y}_0 - \hat{\alpha}_C}{\hat{\beta}_C}, \end{aligned}$$

sendo este último é o estimador clássico consistente da quantidade desconhecida X_0 . A variância assintótica do estimador clássico consistente acima, segundo Lima (1996) é dada pela expressão:

$$\hat{Var}(\hat{U}_{0C}) = \frac{\hat{\varphi}}{n\hat{\beta}_C^2} + \frac{\hat{\sigma}_\varepsilon^2}{k\hat{\beta}_C^2} + \frac{3\hat{\sigma}_\varepsilon^4(1+l)}{nk\hat{\beta}_C^4\sigma_\delta^2l^2} + \frac{\hat{\sigma}_\varepsilon^2(2-3l)}{nk\hat{\beta}_C^2l^2} + \Gamma, \quad (8)$$

em que $\Gamma = (\hat{\sigma}_\varepsilon^2(1+l) + \hat{\beta}_C^2\sigma_\delta^2(2+l))\frac{(\hat{U}_{0C}-\bar{X})^2}{n\hat{\beta}_C^2l^2\sigma_\delta^2} + O(n^{-2})$. Sendo que o erro padrão de \hat{U}_{0C} será denotado por $EP(\hat{U}_{0C})$, que é a raiz quadrada da expressão

da variância assintótica apresentada anteriormente. Nesta expressão tem-se as seguintes estatísticas:

$$\begin{aligned}\hat{\phi} &= \hat{\sigma}_\varepsilon^2 + \hat{\beta}_C^2 \sigma_\delta^2 \\ l &= \frac{\sum_{i=1}^n (U_i - \bar{U})^2}{n\sigma_\delta^2},\end{aligned}\tag{9}$$

sendo k o número de observações obtidas na segunda etapa de calibração e $O(n^{-2})$ é um erro de segunda ordem.

Portanto, se desprezarmos os erros de segunda ordem e trabalharmos apenas com valores de primeira ordem, podemos construir o gráfico de controle com os limites assintóticos. Ressaltando que na construção do gráfico de controle, a linha central, bem como os limites, inferior e superior de controle utilizando modelos de calibração com erros nas variáveis X e Y utilizando o estimador clássico consistentes de X_0 são dadas pelas expressões:

$$\begin{aligned}LSC_{\hat{U}_{0C}} &= \hat{U}_{0C} + qEP(\hat{U}_{0C}) \\ LM_{\hat{U}_{0C}} &= \hat{U}_{0C} \\ LIC_{\hat{U}_{0C}} &= \hat{U}_{0C} - qEP(\hat{U}_{0C}).\end{aligned}\tag{10}$$

Observa-se que a linha central do gráfico de controle, denotada por $LM_{\hat{U}_{0C}}$, é igual ao próprio estimador consistente e que a largura das bandas de controle ou região de ação do gráfico depende exclusivamente do erro padrão do estimador.

2.3 Simulação Monte Carlo

Para simular o processo, foi gerada inicialmente, uma amostra aleatória de 20 pares de observações (X, Y) , com X e Y relacionando-se linearmente. Para tanto, os valores de foram considerados constantes fixas, definidas por: $X = 1, 2, 3, \dots, 20$. Os vetores: $\varepsilon = \varepsilon_1, \dots, \varepsilon_{20}, \varepsilon_{0n+1}, \dots, \varepsilon_{0n+k}$, e $\delta = \delta_1, \dots, \delta_{20}$ foram gerados a partir de uma distribuição normal com média zero e variância um, com amostra de tamanho $n = 20$. O vetor $U = U_1, \dots, U_n$ foi gerado a partir da diferença dos valores de X com os valores de δ , sendo considerados, também, como valores fixos, para satisfazer a condição do modelo funcional. Os valores de Y foram obtidos a partir da expressão $Y = \alpha + \beta U + \varepsilon$. Em cada par ordenado, os valores de Y , que foram utilizados nas expressões do modelo funcional, correspondem à média aritmética de 1000 replicações. Como uma das condições adicionais, que fazem com que o modelo de calibração funcional seja factível é de se supor que a variância dos erros δ seja conhecida, utilizou-se aqui o valor $\sigma_\delta^2 = 1$.

2.4 Software

Os cálculos em todo o trabalho de simulação foram realizados a partir de rotinas desenvolvidas no software R (R Development Core Team, 2007).

3 Resultados e discussão

Os resultados da simulação do processo, utilizando vários valores para a variância dos erros de medida δ , são mostrados nas figuras a seguir.

Na Figura 1 é mostrado o gráfico do processo simulado sob controle estatístico, com $\alpha = 0$, $\beta = 1$, $n = 20$ e $\sigma_\delta^2 = 1$.

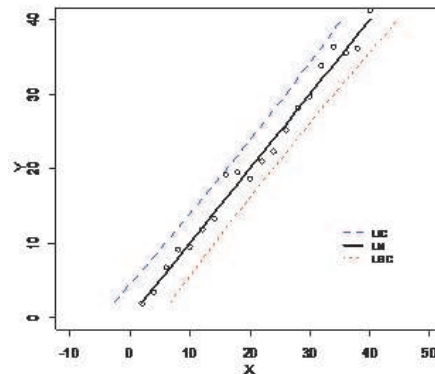


Figura 1 - Gráfico com processo sob controle estatístico com $\alpha = 0$, $\beta = 1$, $n = 20$ e $\sigma_\delta^2 = 1$.

Como pode ser observado, o programa gerou aleatoriamente todos os pontos dentro dos limites de controle, construídos com três desvios padrão. Do total das 20000 observações geradas, o programa capturou onze delas fora dos limites de controle, sendo que seis observações ficaram acima ou à esquerda do limite inferior e cinco ficaram abaixo ou a direita do limite superior de controle. Isto confirma que o programa realmente gerou um processo sob controle estatístico, pois nesta situação se admitiria até 54 observações fora dos limites que ainda assim o processo estaria sob controle estatístico. Nas figuras a seguir são mostrados os gráficos de controle, correspondentes a outros valores da variância dos erros.

Na Figura 2 observa-se que ao aumentar o valor da variância dos erros δ de uma para duas unidades, os limites praticamente deixaram de ser retas paralelas à linha central e continuaram paralelos, porém agora, ficaram um pouco curvados, à medida que os valores de X e Y se afastam dos valores centrais. Nesta situação, o quarto ponto observado próximo ao limite superior de controle provavelmente estaria fora da área de ação do gráfico, se os limites continuassem linhas retas.

Na Figura 3 observa-se que a variância de δ foi aumentada em mais uma unidade, os limites ficaram mais curvos ainda, e nesta situação percebemos que pelo menos dois pontos estariam à direita do limite superior de controle, isto é, foram da área de atuação do gráfico. Notamos também que pelo menos um ponto estaria à esquerda do limite inferior de controle, caso as linhas paralelas à linha central fossem linhas retas e não curvas acentuadas como é o que está acontecendo.

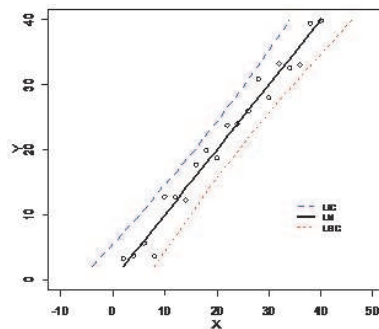


Figura 2 - Gráfico com processo sob controle estatístico com $\alpha = 0$, $\beta = 1$, $n = 20$ e $\sigma_{\delta}^2 = 2$.

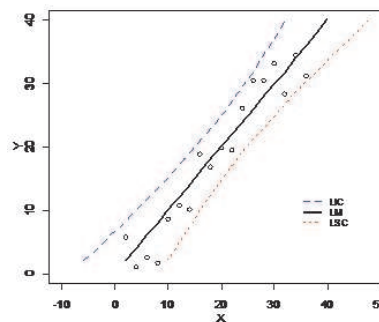


Figura 3 - Gráfico com processo sob controle estatístico com $\alpha = 0$, $\beta = 1$, $n = 20$ e $\sigma_{\delta}^2 = 3$.

A Figura 4 já mostra pontos fora dos limites de controle. Na verdade, um ponto à direita do limite superior e dois pontos em cima dos limites, inferior e superior. Mostra também que mais dois pontos poderiam estar em cima ou até mesmo fora da área de atuação do gráfico, mais precisamente a direita do limite superior.

A Figura 5 mostra um bem extremo, a variância assume o valor, ou seja, a distribuição dos erros continua com média zero, mas agora com variação de unidades. Note que além de pontos fora dos limites de controle, temos também pontos em cima das linhas e pontos bem próximos às linhas, principalmente nas regiões mais extremas. As linhas que representam os limites de controle agora estão realmente bem curvadas, mostrando assim que a variabilidade nos erros de mensuração afeta profundamente o poder de atuação do gráfico de controle. Isto deixa claro, a necessidade que o pesquisador ou o responsável pelo laboratório do processo de medição deve ter, em monitorar e controlar o processo das medições da variável independente X.

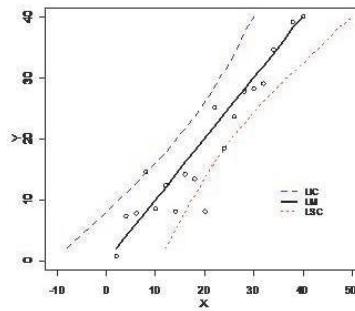


Figura 4 - Gráfico com processo sob controle estatístico com $\alpha = 0$, $\beta = 1$, $n = 20$ e $\sigma_{\delta}^2 = 4$.

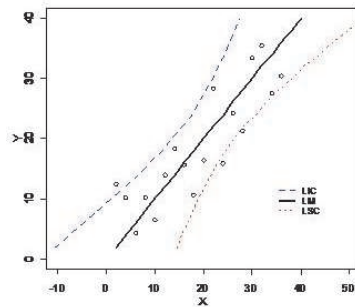


Figura 5 - Gráfico com processo sob controle estatístico com $\alpha = 0$, $\beta = 1$, $n = 20$ e $\sigma_{\delta}^2 = 5$.

Aplicação aos dados referentes a produção de trigo

Para ilustrar a teoria apresentada será utilizado um conjunto de dados que se encontra na literatura (Fuller, 1987), referente à produção de trigo e à determinação de nitrogênio disponível no solo. Para esses dados faz-se necessário considerar que as estimativas do nitrogênio no solo contêm erros de medidas, sendo que a variância destes erros tem sido estimada por $\hat{\sigma}_{\delta}^2 = 57$. Além disso, é necessário verificar se a relação entre as variáveis: produção de trigo e nitrogênio disponível no solo é linear. Os dados estão dispostos na Tabela 1.

Tabela 1 - Valores da produção de trigo e da quantidade de nitrogênio no solo representados por (Y) e (X), respectivamente

Y	86	115	90	86	110	91	99	96	99	104	96
X	70	97	53	64	95	64	50	70	94	69	51

Fonte: Fuller (1987).

A verificação da condição de linearidade entre as variáveis pode ser realizada por meio da construção de um gráfico de dispersão, como mostrado a seguir. A seguir, mostra-se os gráficos de controle para as variáveis em questão, considerando inicialmente que as estimativas do nitrogênio são obtidas sem erros de medida e, posteriormente, serão consideradas estimativas com erros de medidas, assumindo os valores 57 e 10 para a variância dos erros de medidas do Nitrogênio no solo.

Como pode ser observado nas Figuras 6, 7 e 8, os erros de medida na variável independente, produzem mudanças substanciais nos limites dos gráficos de controle quando utilizamos modelos de calibração. Veja, por exemplo, que na Figura 6 onde o gráfico foi construído levando em consideração que a variável nitrogênio no solo é obtida sem erros de mensuração, os limites são retas paralelas à linha central, enquanto que nas Figuras 7 e 8, onde utilizamos a variável nitrogênio no solo com erros de mensuração para construir tais gráficos, os limites paralelos à linha central ficaram curvados, tal como já mostrados no processo de simulação. Isto significa dizer que se faz necessário obter medidas do nitrogênio, de forma que a variância seja a menor possível.

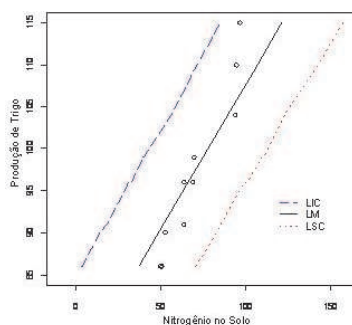


Figura 6 - Gráfico de controle de calibração, sem erros de medidas.

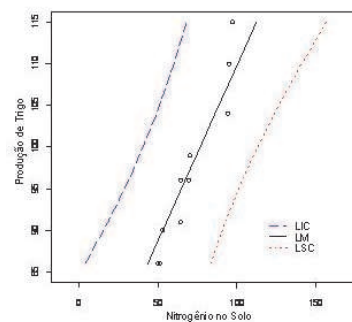


Figura 7 - Gráfico de controle de calibração com $\sigma_0^2 = 10$.

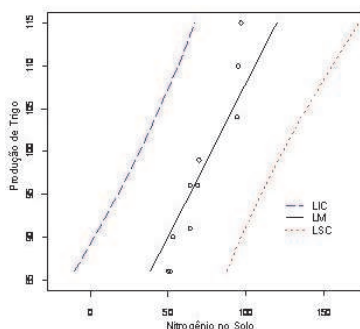


Figura 8 - Gráfico de controle de calibração com $e \sigma_0^2 = 57$.

Conclusões

Nesse artigo, obsevou-se que o monitoramento de um processo envolvendo duas variáveis correlacionadas linearmente, pode ser feito através do gráfico de controle de regressão inversa, conhecida na literatura como calibração. Particularmente, foi utilizado o modelo de calibração funcional que considera a variável independente como uma constante fixa, porém, obtida com erros de mensuração. Os limites de controle obtidos analiticamente através dos estimadores dos parâmetros do modelo e através do estimador da quantidade desconhecida X_0 , nos fornecem resultados assintóticos (resultados aproximados), com os erros de segunda ordem sendo desprezados.

Como o objeto era a investigação no formato do gráfico de controle de calibração, devido aos erros de mensuração, o processo de simulação nos mostrou que os erros de mensuração na variável independente produzem alterações significativas no formato do gráfico de controle, uma vez que à medida que os erros aumentam, os limites inferior e superior de controle tornam-se encurvados em suas extremidades, produzindo assim, menor rigor quanto à atuação do gráfico.

Nota-se que à medida que se aumentam a magnitude da variância dos erros, os limites de controle além de se tornarem mais curvados para valores mais afastados do centro da distribuição de Y , ficam também mais largos. Esse alargamento acontece tanto no centro da distribuição dos dados, como em suas extremidades, levando a um afrouxamento do gráfico de controle. Esse afrouxamento leva a uma redução do número de alarmes falsos, uma vez que pontos que deveriam cair fora dos limites, ficam dentro da área de atuação do gráfico, levando assim a um engano crucial e não desejado pelos gerentes de processos. Dessa forma, torna-se imprescindível obter medições das características da qualidade onde os erros sejam os menores possíveis.

DOMINGOS FILHO, M.; SAMOHYL, R. W.; MACHADO, E. J.; SILVA, R. N. C.; BRAGA, A. S. Influence of measurement errors in the construction of control limits of a calibration graph functional, with the variance of the errors allegedly known. *Rev. Bras. Biom.*, São Paulo, v.29, n.1, p.1-13, 2011.

- **ABSTRACT:** *The objective of this article is to verify the influence that the measurement errors exert on the obtaining of the control limits, when a graph is constructed to monitor the behavior or simultaneous relationship between two variables linearly correlated. In this article specifically, a graph of a functional linear calibration control will be used, supposing that the variance of measurement errors are known. A calibration model is nothing but an inverted regression model. In the functional calibration model we find as main characteristic the fact that the independent variable X is considered a rigid constant or a controlled variable. Nevertheless, the values for this variable X can or cannot present errors in its measuring. These two situations will be taken into account, besides the necessity to consider the magnitude of variance of these measurement errors as well.*
- **KEYWORDS:** *Calibration; estimators; control graph.*

Referências

- BOLFARINE, H.; LIMA, C. R. O. P.; SANDOVAL, M. C. Linear calibration in functional regression models with one the variances known. *South African Statist. J.*, Bloemfontein, v.33, p.95-116, 1999.
- BUSSAB, W. O.; MORETTIN, P. A. *Estatística básica*. 5.ed. São Paulo: Saraiva, 2002. p.291-296. (Coleção métodos quantitativos).
- CHENG, C. L.; Van NESS, J. W. *Statistical regression with measurement error*. New York: Arold, 1999. 262p.
- DiPAOLA, P. P. Use of correlation in quality control. *Ind. Qual. Control*, Washington, v.2, n.1, p.10-14, 1945.
- FULLER, W.A. *Measurement error models*. New York: Wiley, 1987. 440p.
- LIMA, C.R.O.P. *Calibração absoluta com erros nas variáveis*. 1996. Tese (Doutorado em Estatística) – Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 1996.
- MANDEL, B. J. The regression control chart. *J. Qual. Technol.*, Milwaukee, v.1, n.1, p.1-9, 1969.
- NEYMAN, J.; SCOTT, E. On certain methods of estimating the linear structural relation. *Ann. Math. Stat.*, Cornell, v.22, p.352-361, 1951.
- R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. R: Foundation for Statistical Computing. Vienna, Austria. Disponível em: < <http://www.R-project.org> >. Acesso em: 20 dez. 2007.
- SHUKLA, G.K. On the problem of calibration. *Technometrics*, London, v.14, n.3, p.547-553, 1972.

Recebido em 11.10.2010.

Aprovado após revisão em 23.03.2011.