**RT-MAE 2006-18**

# HOMOSCEDASTIC CONTROLLED CALIBRATION MODEL

**Betsabé G. Blas Achic, Mônica C. Sandoval**
**and**
**Olga Satomi Yoshida**

**Palavras-Chave: Regression model, linear calibration model, measurement error model, Berkson model**
**Classificação AMS: 62J05,62F12.**

**- Novembro 2006 -**

# Homoscedastic controlled calibration model

Betsabé G. Blas Achic[a], Mônica C. Sandoval [b] and Olga Satomi Yoshida[c]

[a,b]Departamento de Estatística, Universidade de São Paulo, São Paulo, Brasil
[c] Centro de Metrologia de Fluidos, Instituto de Pesquisas Tecnológicas, São Paulo, Brasil

## Abstract

In the context of the usual calibration model, we consider the case in which the independent variable is unobservable, but a pre-fixed value on its surrogate is available. Thus, considering controlled variables and assuming that the measurement errors have equal variances we propose a new calibration model. Likelihood based methodology is used to estimate the model parameters and the Fisher information matrix is used to construct a confidence interval for the unknown value of the regressor variable. A simulation study is carried out to asses the effect of the measurement error on the estimation of the parameter of interest. This new approach is illustrated with an example.

*Keywords:* Regression model, linear calibration model, measurement error model, Berkson model.

# 1 Introduction

In the first stage of a calibration problem, a pair of data sample $(x_i, Y_i)$, $i = 1, 2, \cdots n$ is observed. In the second stage, it is observed one or more values, which are the responses corresponding to a single unknown value of the regressor variable, $X_0$. The first and second stage equations of the usual linear calibration model are defined, respectively, as

$$
\begin{aligned}
Y_i &= \alpha + \beta x_i + \epsilon_i, \quad i = 1, 2 \cdots, n, & (1.1)\\
Y_{0i} &= \alpha + \beta X_0 + \epsilon_i, \quad i = n+1, n+2, \cdots, n+k. & (1.2)
\end{aligned}
$$

It is considered the following assumptions:

- $x_1, x_2, \cdots, x_n$ take fixed values, which are considered as true values.
- $\epsilon_1, \epsilon_2, \cdots, \epsilon_{n+k}$ are independent and normally distributed with mean 0 and variance $\sigma_\epsilon^2$.

The model parameters are $\alpha, \beta, X_0$ and $\sigma_\epsilon^2$ and the main interest is to estimate the quantity $X_0$.

The maximun likelihood estimators of the usual calibration model are given by

$$
\begin{aligned}
\hat{\alpha} &= \bar{Y} - \hat{\beta}\bar{x}, \qquad \hat{\beta} = \frac{S_{xY}}{S_{xx}}, \qquad \hat{X}_0 = \frac{\bar{Y}_0 - \hat{\alpha}}{\hat{\beta}}, & (1.3)\\
\sigma_\epsilon^2 &= \frac{1}{n+k}[\sum_{i=1}^{n}(Y_i - \hat{\alpha} - \hat{\beta}x_i)^2 + \sum_{i=n+1}^{n+k}(Y_{0i} - \bar{Y}_0)^2], & (1.4)
\end{aligned}
$$

where

$$
\begin{aligned}
\bar{x} &= \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i, \quad S_{xY} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(Y_i - \bar{Y}),\\
S_{xx} &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2, \quad \bar{Y}_0 = \frac{1}{n}\sum_{i=n+1}^{n+k} Y_{0i}.
\end{aligned}
$$

In [10] an approximate expression is derived for the variance of the estimator $\hat{X}_0$, which is derived through the propagation error law. Another approximation for the variance of $\hat{X}_0$ is given by the Fisher information of $\theta = (\alpha, \beta, X_0, \sigma_\epsilon^2)$ which, after some length algebraic manipulations, it can be shown to be given by

$$
I(\theta) = \frac{1}{\sigma_\epsilon^2}\begin{pmatrix} n+k & kX_0 + n\bar{x} & k\beta & 0 \\ kX_0 + n\bar{x} & kX_0^2 + \sum_{i=1}^{n} x_i^2 & \kappa\beta X_0 & 0 \\ k\beta & \kappa\beta X_0 & k\beta^2 & 0 \\ 0 & 0 & 0 & \frac{n+k}{2\sigma_\epsilon^2} \end{pmatrix}. \qquad (1.5)
$$

The maximum likelihood estimator of $\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{X}_0, \hat{\sigma_\epsilon^2})$ has approximately normal distribution with mean $\theta$ and covariance matrix $I(\theta)^{-1}$, when $k = qn$, $q \in Q^+$ and $n \longrightarrow \infty$. Thus, the approximation of order $n^{-1}$ for the variance of $\hat{X}_0$ is given by

$$V_1(\hat{X}_0) = \frac{\sigma_\epsilon^2}{\beta^2}\left[\frac{1}{k} + \frac{1}{n} + \frac{(\bar{X} - X_0)^2}{nS_{xx}}\right]. \tag{1.6}$$

On the other hand, in [4] the size $k$ of the second stage is considered fixed, so that expanding $\hat{X}_0$ in Taylor series around the point $(\alpha, \beta)$ and ignoring terms of order less than $n^{-2}$, we can find the following approximations for the bias and variance of $\hat{X}_0$, respectively,

$$Bias(\hat{X}_0) = \frac{\sigma_\epsilon^2(X_0 - \bar{x})}{n\beta^2 S_{xx}}, \tag{1.7}$$

$$V_2(\hat{X}_0) = \frac{\sigma_\epsilon^2}{\beta^2}\left[\frac{1}{k} + \frac{1}{n} + \frac{(\bar{X} - X_0)^2}{nS_{xx}} + \frac{3\sigma_\epsilon^2}{nk\beta^2 S_{xx}}\right]. \tag{1.8}$$

In order to construct a confidence interval for $X_0$, we consider that

$$\frac{\hat{X}_0 - X_0}{\sqrt{\hat{V}(\hat{X}_0)}} \xrightarrow{D} N(0,1), \tag{1.9}$$

where $\hat{V}(\hat{X}_0)$ is the estimated variance computed according to (1.6) or (1.8). Hence, the approximated confidence interval for $X_0$ with a confidence level $(1 - \alpha)$, is given by

$$\left(\hat{X}_0 - z_{\frac{\alpha}{2}}\sqrt{\hat{V}(\hat{X}_0)}, \hat{X}_0 + z_{\frac{\alpha}{2}}\sqrt{\hat{V}(\hat{X}_0)}\right), \tag{1.10}$$

where $z_{\frac{\alpha}{2}}$ is the quantile of order $(1 - \frac{\alpha}{2})$ of the standard normal distribution.

The usual calibration problem has been discussed in the literature for several decades (see [1]-[6]). An illustration of this model is presented for example in [7]. We can find a review of the literature on statistical calibration in [8], where some approaches to the solution of the calibration problem are summarized.

This model encounters applications in different areas, but it is not well suited in some instances as, for example, in chemical analysis, where the preparation process of standard solutions are subject to measurement error ([10]).

There exists some situations, as mentioned above, where the independent variable, $x_i$, is measured with error. In this case, [11] defines two types of observations: *controlled* and *uncontrolled*.

In the *uncontrolled* situation, the usual procedure to obtain the true value of the independent variable $x_i$ generates an error and the observed value is

$$X_i = x_i + \delta_i, \quad i = 1, \cdots, n. \tag{1.11}$$

We have that $x_i$ is an unknown quantity, $\delta_i$ is a measurement error and $X_i$ is a random variable. Assuming that $x_i$ is a parameter the model defined by (1.1) and (1.11) is named as functional model ([12]). In this case there exists correlation between the model error and the variable $X_i$. Assuming that $x_i$ is a random variable the model (1.1) and (1.11) is called as structural model ([12]). On the other hand, the model defined by (1.1), (1.2) and (1.11) is called as the functional or structural calibration model if $x_i$ is assumed as a parameter or a random variable, respectively ([13]).

The *controlled* observation is defined by a pre-fixed value $X_i$ according to the experimenter convenience and a procedure is established in order to attain the pre-fixed value. The experiment gives the unobserved $x_i$ and it is such that

$$x_i = X_i - \delta_i, \quad i = 1, \cdots, n. \tag{1.12}$$

In this case, the fixed quantity is $X_i$, the measurement error is $\delta_i$ and $x_i$ is the random variable. The model (1.1) and (1.12) is known as Berkson regression model ([9]). Notice that the model error and the quantity $X_i$ are independent. The model defined by (1.1), (1.2) and (1.12) has not been considered before in the measurement error literature and in this work it will be called as the *controlled calibration model*.

In the calibration model defined by (1.1), (1.2) and (1.11), the values of the regressor, $X_i$, from the first stage are randomly generated, whereas in the controlled calibration model, (1.1), (1.2) and (1.12), they are assumed as pre-fixed by the experimenter.

This work is organized as follows. In Section 2, we derive the maximum likelihood estimators of the homoscedastic controlled calibration model by considering both cases: $\sigma_\delta^2$ *unknown* and *known*. In Section 3, a simulation study is undertaken to investigate the sensitivity of parameter estimates of the proposed model. In Section 4, an example is presented to illustrate our new approach. In Section 5, the concluding remark is presented.

# 2 Parameter estimation

In this section we study the controlled calibration model. From the equations (1.1), (1.2) and (1.12) we can write

$$\begin{aligned} Y_i &= \alpha + \beta X_i + (\epsilon_i - \beta\delta_i), \quad i = 1, 2 \cdots, n, & (2.1) \\ Y_{0i} &= \alpha + \beta X_0 + \epsilon_i, \quad i = n+1, n+2, \cdots, n+k. & (2.2) \end{aligned}$$

with the following assumptions for the random errors

- $\epsilon_i$ are independent $N(0, \sigma_\epsilon^2)$ random variables.
- $\mathrm{E}(\delta_i)$=0, $\mathrm{V}(\delta_i)$= $\sigma_{\delta_i}^2$.
- $\mathrm{cov}(\delta_i, \delta_j)$=0 for any $i \neq j$.

- $\text{cov}(\epsilon_i, \delta_j)=0$ for all $i, j$.

Some comments are in order here. The variable $X_i$ in (2.1) is controlled and the error model $(\epsilon_i - \beta\delta_i)$ is independent of $X_i$. The error model in (2.2) is only in function of error measure $\epsilon_i$ related to $Y_{0i}$, this model assume that there is not error in the preparation sample related to parameter $X_0$. We define the homoscedastic controlled calibration model by considering that the errors $\delta_i$ are independent and normally distributed with mean 0 and constant variance, $\sigma_\delta^2$. The study of this model is carried out following similar analysis to the usual calibration model as summarized above.

The maximum likelihood estimator for the homoscedastic controlled calibration model is derived in the following. The logarithm of the likelihood function is given by:

$$\begin{aligned} l(\alpha, \beta, X_0, \sigma_\epsilon^2, \sigma_\delta^2) \;\propto\; & -\frac{n}{2}log(\sigma_\epsilon^2 + \beta^2\sigma_\delta^2) - \frac{k}{2}log(\sigma_\epsilon^2) - \\ & \frac{1}{2}\Big[\frac{1}{\sigma_\epsilon^2 + \beta^2\sigma_\delta^2}\sum_{i=1}^{n}(Y_i - \alpha - \beta X_i)^2 + \\ & \frac{1}{\sigma_\epsilon^2}\sum_{i=n+1}^{n+k}(Y_{0i} - \alpha - \beta X_0)^2\Big]. \end{aligned} \tag{2.3}$$

Solving $\partial l/\partial\alpha = 0$ and $\partial l/\partial X_0 = 0$ we have the maximum likelihood estimator of $\alpha$ and $X_0$, which are given, respectively, by

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} \quad \text{and} \quad \hat{X}_0 = \frac{\bar{Y}_0 - \hat{\alpha}}{\hat{\beta}}. \tag{2.4}$$

From (2.3) and (2.4), it follows that the likelihood for $(\beta, \sigma_\epsilon^2, \sigma_\delta^2)$ can be written as

$$\begin{aligned} l(\beta, \sigma_\epsilon^2, \sigma_\delta^2) \propto & - \frac{n}{2}log(\sigma_\epsilon^2 + \beta^2\sigma_\delta^2) - \frac{k}{2}log(\sigma_\epsilon^2) - \frac{1}{2}\Big[\frac{1}{\sigma_\epsilon^2 + \beta^2\sigma_\delta^2}\sum_{i=1}^{n}[(Y_i - \bar{Y}) \\ & - \beta(X_i - \bar{X})]^2 + \frac{1}{\sigma_\epsilon^2}\sum_{i=n+1}^{n+k}(Y_{0i} - \bar{Y}_0)^2\Big]. \end{aligned} \tag{2.5}$$

Next, we consider two cases for $\sigma_\delta^2$. Firstly, we obtain the maximum likelihood estimator of $\beta$, $\sigma_\epsilon^2$ and $\sigma_\delta^2$ from (2.5). In the second case we assume that the variance $\sigma_\delta^2$ is known and obtain the maximum likelihood estimators for $\beta$ and $\sigma_\epsilon^2$.

## Case 1: *unknown* variance $\sigma_\delta^2$

Taking the partial derivative of (2.5) with respect to $\beta$, $\sigma_\epsilon^2$ and $\sigma_\delta^2$ and equating to zero we obtain, respectively,

$$\hat{\beta}\hat{\sigma}_\delta^2(\hat{\sigma}_\epsilon^2 + \hat{\beta}^2\hat{\sigma}_\delta^2 - S_{YY} + \hat{\beta}S_{XY}) = (S_{XY} - \hat{\beta}S_{XX})\hat{\sigma}_\epsilon^2, \tag{2.6}$$

$$
\begin{aligned}
\hat{\sigma}_\epsilon^2 + \hat{\beta}^2\hat{\sigma}_\delta^2 &= S_{YY} - 2\hat{\beta}S_{XY} + \hat{\beta}^2 S_{XX}, && (2.7)\\
\frac{kS_{Y_0Y_0}}{(\hat{\sigma}_\epsilon^2)^2} - \frac{k}{\hat{\sigma}_\epsilon^2} &= \frac{n}{\hat{\sigma}_\epsilon^2 + \hat{\beta}^2\hat{\sigma}_\delta^2} - \frac{n(S_{YY} - 2\hat{\beta}S_{XY} + \hat{\beta}^2 S_{XX})}{(\hat{\sigma}_\epsilon^2 + \hat{\beta}^2\hat{\sigma}_\delta^2)^2}, && (2.8)
\end{aligned}
$$

where $S_{XX} = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$, $S_{XY} = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$, $S_{YY} = \frac{1}{n}\sum_{i=1}^n (Y_i - \bar{Y})^2$ and $S_{Y_0Y_0} = \frac{1}{k}\sum_{i=n+1}^{n+k} (Y_{0i} - \bar{Y}_0)^2$, and the relevant estimator notation has been introduced. From (2.6) and (2.7) we have the following equations:

$$(\hat{\beta}S_{XX} - S_{XY})(S_{YY} - 2\hat{\beta}S_{XY} + \hat{\beta}^2 S_{XX}) = 0,$$

hence

$$
\begin{aligned}
\hat{\beta}S_{XX} - S_{XY} &= 0 \quad \text{or} && (2.9)\\
S_{YY} - 2\hat{\beta}S_{XY} + \hat{\beta}^2 S_{XX} &= 0. && (2.10)
\end{aligned}
$$

Therefore, from (2.9), we have that $\hat{\beta} = S_{XY}/S_{XX}$. But, according to the Cauchy-Schwarz inequality, $S_{XX}S_{YY} \geq S_{XY}^2$, hence (2.10) has real roots if and only if $Y_i = cX_i$, where $c$ is a constant.

The estimator of $\sigma_\delta^2$ can be obtained from the equation (2.7)

$$\hat{\sigma}_\delta^2 = \frac{(S_{YY} - 2\hat{\beta}S_{XY} + \hat{\beta}^2 S_{XX}) - \hat{\sigma}_\epsilon^2}{\hat{\beta}^2}.$$

Likewise, from equations (2.7) and (2.8) we obtain the estimator of the variance $\sigma_\epsilon^2$

$$\hat{\sigma}_\epsilon^2 = S_{Y_0Y_0}. \tag{2.11}$$

In order to find the variance of $\hat{X}_0$, we need to derive the Fisher information matrix of $\theta = (\alpha, \beta, X_0, \sigma_\delta^2, \sigma_\epsilon^2)$, which can be shown to be given by

$$
I(\theta) = \begin{pmatrix}
\frac{n}{\gamma} + \frac{k}{\sigma_\epsilon^2} & \frac{n\bar{X}}{\gamma} + \frac{kX_0}{\sigma_\epsilon^2} & \frac{k\beta}{\sigma_\epsilon^2} & 0 & 0\\
\frac{n\bar{X}}{\gamma} + \frac{kX_0}{\sigma_\epsilon^2} & \frac{\sum_{i=1}^n X_i^2}{\gamma} + \frac{2n\beta^2\sigma_\delta^4}{\gamma^2} + \frac{kX_0^2}{\sigma_\epsilon^2} & \frac{k\beta X_0}{\sigma_\epsilon^2} & \frac{n\beta^3\sigma_\delta^2}{\gamma^2} & \frac{n\beta\sigma_\delta^2}{\gamma^2}\\
\frac{k\beta}{\sigma_\epsilon^2} & \frac{k\beta X_0}{\sigma_\epsilon^2} & \frac{k\beta^2}{\sigma_\epsilon^2} & 0 & 0\\
0 & \frac{n\beta^3\sigma_\delta^2}{\gamma^2} & 0 & \frac{n\beta^4}{2\gamma^2} & \frac{n\beta^2}{2\gamma^2}\\
0 & \frac{n\beta\sigma_\delta^2}{\gamma^2} & 0 & \frac{n\beta^2}{2\gamma^2} & \frac{n}{2\gamma^2} + \frac{k}{2\sigma_\epsilon^4}
\end{pmatrix},
$$

where

$$\gamma = \beta^2\sigma_\delta^2 + \sigma_\epsilon^2. \tag{2.12}$$

When $k = qn$, $q \in Q^+$ and $n \longrightarrow \infty$, the estimator $\hat{\theta}$ is approximately normally distributed with mean $\theta$ and variance $I(\theta)^{-1}$, thus we have that the approximate variance to order $n^{-1}$ for $\hat{X}_0$ is given by

$$V_1(\hat{X}_0) = \frac{\sigma_\epsilon^2}{\beta^2}\left[\frac{1}{k} + \frac{\gamma}{n\sigma_\epsilon^2} + \frac{\gamma}{\sigma_\epsilon^2}\frac{(\bar{X} - X_0)^2}{nS_{XX}}\right]. \tag{2.13}$$

Considering $k$ fixed and expanding $\hat{X}_0$ in a Taylor series around $(\alpha, \beta)$ and ignoring terms of order less than $n^{-2}$, it can be shown that the bias and variance of $\hat{X}_0$ (the proof is given in Appendix A), are given by

$$
\begin{aligned}
Bias(\hat{X}_0) &= \frac{\gamma(\bar{X} - X_0)}{n\beta^2 S_{XX}}, & (2.14)\\
V_2(\hat{X}_0) &= \frac{\sigma_\epsilon^2}{\beta^2}\left[\frac{1}{k} + \frac{\gamma}{n\sigma_\epsilon^2} + \frac{\gamma(\bar{X} - X_0)^2}{n\sigma_\epsilon^2 S_{XX}} + \frac{3\gamma}{nk\beta^2 S_{XX}}\right]. & (2.15)
\end{aligned}
$$

We can observe that the estimator of $X_0$ is biased, but it is asymptotically unbiased.

With relation to the variance of the estimator $\hat{X}_0$, let us notice that when $k = qn,\ q \in Q^+$, and ignoring the terms of order less than $n^{-1}$ the variance in (2.15) coincide with the variance given in (2.13), which was found through the Fisher information. Equation (2.13) consider large sample sizes in the first and second stage ($n$ and $k$), whereas (2.15) consider large sample sizes in the first stage and a fixed sample size in the second stage.

Notice that when $\sigma_\delta^2 = 0$, (2.13) and (2.15) coincide with (1.6) and (1.8) of the usual model, respectively.

## Caso 2: *known* variance $\sigma_\delta^2$

Assuming now that $\sigma_\delta^2$ is known and equating to zero the partial derivative of (2.5) with respect to the parameters $\beta$ and $\sigma_\epsilon^2$, we have the following equations, respectively,

$$
\begin{aligned}
\hat{\beta}\sigma_\delta^2(\hat{\sigma}_\epsilon^2 + \hat{\beta}^2\sigma_\delta^2 - S_{YY} + \hat{\beta}S_{XY}) &= (S_{XY} - \hat{\beta}S_{XX})\hat{\sigma}_\epsilon^2 \quad \text{and} & (2.16)\\
\frac{kS_{Y_0Y_0}}{(\hat{\sigma}_\epsilon^2)^2} - \frac{k}{\hat{\sigma}_\epsilon^2} &= \frac{n}{\hat{\sigma}_\epsilon^2 + \hat{\beta}^2\sigma_\delta^2} - \frac{S_{YY} - 2\hat{\beta}S_{XY} + \hat{\beta}^2 S_{XX}}{(\hat{\sigma}_\epsilon^2 + \hat{\beta}^2\sigma_\delta^2)^2}. & (2.17)
\end{aligned}
$$

The estimates of $\beta$ and $\sigma_\epsilon^2$ are obtained using some iterative method to solve (2.16) and (2.17).

Similarly, as in Case 1, the Fisher information matrix of $\theta = (\alpha, \beta, X_0, \sigma_\epsilon^2)$ is given by

$$
I(\theta) = \begin{pmatrix}
\frac{n}{\gamma} + \frac{k}{\sigma_\epsilon^2} & \frac{n\bar{X}}{\gamma} + \frac{kX_0}{\sigma_\epsilon^2} & \frac{k\beta}{\sigma_\epsilon^2} & 0\\
\frac{n\bar{X}}{\gamma} + \frac{kX_0}{\sigma_\epsilon^2} & \frac{\sum_{i=1}^n X_i^2}{\gamma} + \frac{2n\beta^2\sigma_\delta^4}{\gamma^2} + \frac{kX_0^2}{\sigma_\epsilon^2} & \frac{k\beta X_0}{\sigma_\epsilon^2} & \frac{n\beta\sigma_\delta^2}{\gamma^2}\\
\frac{k\beta}{\sigma_\epsilon^2} & \frac{k\beta X_0}{\sigma_\epsilon^2} & \frac{k\beta^2}{\sigma_\epsilon^2} & 0\\
0 & \frac{n\beta\sigma_\delta^2}{\gamma^2} & 0 & \frac{n}{2\gamma^2} + \frac{k}{2\sigma_\epsilon^4}
\end{pmatrix}, \quad (2.18)
$$

where $\gamma$ is defined in (2.12).

The large sample variance of $\hat{X}_0$ follows by inverting the Fisher information matrix and is given by

$$
V(\hat{X}_0) = \frac{\sigma_\epsilon^2}{\beta^2}\left[\frac{1}{k} + \frac{\gamma}{n\sigma_\epsilon^2} + \frac{\gamma}{\sigma_\epsilon^2}E\right], \quad (2.19)
$$

where,

$$E = \frac{nX_0^2\sigma_\epsilon^4 + kX_0^2\gamma^2 - 2nX_0\bar{X}\sigma_\epsilon^4 - 2kX_0\bar{X}\gamma^2 + n\bar{X}^2\sigma_\epsilon^4 + k\bar{X}^2\gamma^2}{(n\sigma_\epsilon^4 + k\gamma^2)\sum_{i=1}^n X_i^2 + 2nk\beta^2\gamma\sigma_\delta^4 - n^2\bar{X}^2\sigma_\epsilon^4 - nk\bar{X}^2\gamma^2}.$$

Notice that if $\sigma_\delta^2 = 0$, the expression (2.19) is reduced to (1.6).

To construct a confidence interval for $X_0$, for both cases $\sigma_\delta^2$ *unknown* and *known*, we consider the interval (1.10), where $\hat{V}(\hat{X}_{0C})$ is the estimated variance that follows from (2.13), (2.15) or (2.19).

# 3 Simulation study

In this section we present a simulation study for both cases of the homoscedastic controlled calibration model: $\sigma_\delta^2$ known and unknown. The objetive of this section is to study the performance of the estimators of the proposed model (Proposed-M) and verify the impact by considering erratically the usual model (Usual-M).

It was considered 5000 samples generated from the homoscedastic controlled calibration model. In all samples, the value of the parameters $\alpha$ and $\beta$ were 0.1 and 2, respectively. The range of values for the controlled variable was [0,2]. The fixed values for the controlled variable were $x_1 = 0$, $x_i = x_{i-1} + \frac{2}{n-1}$, $i = 2, \cdots, n$, and the parameter values $X_0$ were 0.01 (extreme inferior value), 0.8 (near to the central value) and 1.9 (extreme superior value). It was considered $\sigma_\epsilon^2 = 0.04$ and the parameter values of $\sigma_\delta^2$ were 0.01 and 0.1, which are named, respectively, as small and large variances. For the first and second stages we consider the sample of sizes $n = 5, 20, 100$ and $k = 2, 20, 100$, respectively.

The empirical mean bias is given by $\sum_{j=1}^{5000}(\hat{X}_0 - X_0)/5000$ and the empirical mean squared error (MSE) is given by $\sum_{j=1}^{5000}(\hat{X}_0 - X_0)^2/5000$. The mean estimated variance of $\hat{X}_0$ is given by $\sum_{j=1}^{5000} \hat{V}(\hat{X}_0)/5000$, with $\hat{V}(\hat{X}_0) = \hat{V}_1(\hat{X}_0)$ or $\hat{V}_2(\hat{X}_0)$, where $\hat{V}_1(\hat{X}_0)$ is the estimated variance of (1.6), (2.13) or (2.19) and $\hat{V}_2(\hat{X}_0)$ is the estimated variance of (2.15). The theoretical variances of $\hat{X}_0$ denoted as $V_1(\hat{X}_0)$ and $V_2(\hat{X}_0)$, are referred, respectively, to the expressions (1.6), (2.13) or (2.19) and (2.15) evaluated on the relevant parameter values. In Appendix B it is presented the simulation results.

Tables B1, B2, B5 and B6 present the empirical bias, the empirical mean squares error, the theoretical variance and the estimated variance of $X_0$. In these tables, it is considered only the variance (1.6) of the usual model, because based on a simulation study in [15] it was shown that the variances (1.6) and (1.8) give similar results.

Tables B3, B4 and B7 present the covering percentages and the confidence interval amplitudes constructed with a 95% confidence level for the parameter $X_0$. In Table B3, the covering percentages $\%_1$ and $\%_2$ and amplitudes $A_1$ and

$A_2$ are referred to the confidence intervals constructed using the equations (2.13) and (2.19).

Tables B1-B4 consider the homoscedastic controlled calibration model assuming that $\sigma_\delta^2$ is unknown.

In Table B1 the empirical bias and MSE of $\hat{X}_0$ are little and an addition in the size of the variance $\sigma_\delta^2$, described in Table B2, causes an increasing in the bias and MSE. Moreover, we have that the bias and MSE of $\hat{X}_0$ are smaller when $X_0$ is near to the center value of the variation interval of the variable $X$. These tables show that for all $n, k$ and $X_0$, the theoretical variances obtained using the expressions (2.19) and (2.15) are equal. This fact occurs also for the mean estimated variances. We verify also that when $n \geq 20$ and $k \geq 20$ the theoretical variances and the mean estimated variances from the proposed model are approximately equal. Observing these tables, we can also notice that there exists differences between the mean estimated variances of the usual and proposed models.

Analyzing Tables B3 and B4, we observe that for all $n$ and $X_0$ when it is adopted erratically the usual model, the amplitudes decrease very much as the size of $k$ increases. This causes the covering percentage to decrease moving away from 95%. Whereas, adopting the proposed model it is observed that when $k$ increases the confidence interval amplitude decreases, but the covering percentages increase approaching 95%. Notice that the covering percentage $\%_1$ and $\%_2$ and the amplitudes $A_1$ and $A_2$ are approximately equal, the amplitudes are very small for $X_0 = 0.8$. In these tables, we observe that when $k = 20$ or 100 and when $n$ increases the amplitudes of the intervals decrease and the covering percentages approaches 95%. In most cases, the covering percentage obtained through the proposed model are greater than that for the usual model results and are close to 95%.

Tables B5 and B7 describe the results for the controlled homoscedastic calibration model with $\sigma_\delta^2$ known. The iterative method Quasi-Newton [14] has been used.

In Tables B5 and B6 we have that the empirical bias and SME decrease as the size of $n$ or $k$ increase and they are small when $X_0$ is near to the central value of the variation interval, $X_0 = 0.8$. When $\sigma_\delta^2$ is small (Table B5), for all $n$ and $k$, the empirical values of MSE from the usual and proposed model are close to the theoretical variance, but only the mean estimated variance from the proposed model is close to the theoretical variance. When $\sigma_\delta^2$ is large (Table B6), in general, the empirical MSE and the mean estimated variance from the usual and proposed model are different, but the values supplied by the proposed model are very close to the theoretical variance.

Analyzing Table B7, we can make similar comments to the ones we made about Tables B3 and B4.

# 4 Aplication

In this section we test our model, considering both cases $\sigma_\delta$ known and unknown, using the data supplied by the chemical laboratory of the "Instituto de Pesquisas Tecnológicas (IPT)" - Brasil. We also consider the usual model in order to observe the performance of the proposed model. Our main interest is to estimate the unknown concentration value $X_0$ of two samples A and B of the chemical elements cromo and cadmium.

Tables 1 and 4 present the fixed values of concentration of the standard solutions and the corresponding intensities for the cromo and cadmium element, respectively, which are supplied by the plasma spectrometry method. This data is referred to as the first stage of the calibration model.

Tables 2 and 5 present the intensities corresponding to 3 sample solutions from the sample A and B. This data is referred to as the second stage of the calibration model.

Tables 3 and 6 describe the estimates of $\alpha$, $\beta, X_0$, $V(\hat{X}_0)$, $\sigma_\delta^2$ and the confidence interval amplitude $U(X_0)$ from the homoscedastic controlled calibration model of the samples A and B for the chemical elements cromo and cadmium. The values of the variance $\sigma_\delta^2$ considered as known are obtained from an external study carried out by the IPT, which are $\sigma_\delta^2 = 2,5865E-06$ for the cromo element and $\sigma_\delta^2 = 0.0017E+02$ for the cadmium element. As seen in Section 2, in order to obtain the estimates of the parameters $\beta$ and $\sigma_\epsilon^2$ of the proposed model when $\sigma_\delta^2$ is known, iterative methods are required. In order to solve the system of equations (2.16) and (2.17) it was used the Quasi-Newton iterative method. It is also presented the estimates from the usual model. The estimates of the variance of $\hat{X}_0$ are computed using the relevant expressions (1.6), (2.13) or (2.19). The amplitude $U(X_0)$ is given by the product of the squared root of the estimated variance of $\hat{X}_0$ and 1.96.

In Tables 3 and 6 we can observe that the estimates of $\alpha$ and $\beta$ supplied by the usual model is equal to the proposed model when $\sigma_\delta^2$ is unknown and they are equal for samples A and B, this occurs because the expression of the estimators $\hat{\alpha}$ and $\hat{\beta}$ of both models are equal and they only depend on the first stage of the calibration model. These estimates are slightly different when compared with the estimates from the proposed model when $\sigma_\delta^2$ is known. With respect to the estimate of $X_0$, we observe that there is no difference of the estimates supplied by the usual and the proposed models of the cromo and cadmium element in both samples A and B, respectively. The estimates of the concentration of the sample A, of the elements cromo and cadmium, are outside of the variation range of the standard solution concentrations. We verify that, except to the sample B of the cadmium element, the estimates of the variance of $\hat{X}_0$ and the amplitude $U(X_0)$ from the usual model are greater than the estimates supplied by the both proposed models.

Table 1: Concentration ($mg/g$) and intensity of the standard solutions of cromo element.

| $X_i$ | Intensity |
|---|---|
| 0,05 | 6455,900 |
| 0,11 | 13042,933 |
| 0,26 | 32621,733 |
| 0,79 | 97364,500 |
| 1,05 | 129178,100 |

Table 2: Intensity of the sample solutions A and B of cromo element.

| Intensity | |
|---|---|
| Sample A | Sample B |
| 1465,0 | 10173,6 |
| 1351,0 | 10516,9 |
| 1495,6 | 10352,2 |

Table 3: Estimates of $\alpha$, $\beta$, $X_0$, $V(\hat{X}_0)$ and the confidence interval amplitude $U(X_0)$ from the usual and proposed model for the samples A and B of cromo element.

| Parameters | Sample A | | | Sample B | | |
|---|---|---|---|---|---|---|
| | Usual-M | Proposed-M | | Usual-M | Proposed-M | |
| | | unknown $\sigma_\delta^2$ | known $\sigma_\delta^2$ | | unknown $\sigma_\delta^2$ | known $\sigma_\delta^2$ |
| $\alpha$ | 123,574 | 123,574 | 123,889 | 123,574 | 123,574 | 124,021 |
| $\beta$ | 1,23E+05 | 1,23E+05 | 1,23E+05 | 1,23E+05 | 1,23E+05 | 1,23E+05 |
| $X_0$ | 0,011 | 0,011 | 0,011 | 0,083 | 0,083 | 0,083 |
| $V(\hat{X}_0)$ | 9,80E-07 | 9,15E-07 | 1,35E-06 | 1,16E-06 | 1,13E-06 | 1,71E-06 |
| $\sigma_\delta^2$ | - | 1,60E-06 | - | - | 5,48E-07 | - |
| $U(X_0)$ | 2,55E-03 | 2,46E-03 | 2,99E-03 | 2,77E-03 | 2,73E-03 | 3,36E-03 |

Table 4: Concentration ($mg/g$) and intensity of the standard solutions of cadmium element.

| $X_i$ | Intensity |
|---|---|
| 0.05 | 4,89733 |
| 0.10 | 9,706 |
| 0,25 | 23,41333 |
| 0.73 | 69,73 |
| 1.01 | 96,85667 |

Table 5: Intensity of the sample solutions A and B of cadmium element.

| Intensity | |
|---|---|
| Sample A | Sample B |
| 0,679 | 5,066 |
| 0,6837 | 5,027 |
| 0,6846 | 5,085 |

Table 6: Estimates of $\alpha$, $\beta$, $X_0$, $V(\hat{X}_0)$ and the confidence interval amplitude $U(X_0)$ from the usual and homoscedastic models for the samples A and B of cadmium element.

| Parameters | Sample A | | | Sample B | | |
|---|---|---|---|---|---|---|
| | Usual-M | Proposed-M | | Usual-M | Proposed-M | |
| | | unknown $\sigma_\delta^2$ | known $\sigma_\delta^2$ | | unknown $\sigma_\delta^2$ | known $\sigma_\delta^2$ |
| $\alpha$ | -0,156 | -0,156 | -0,158 | -0,156 | -0,156 | -0,158 |
| $\beta$ | 95,828 | 95,828 | 95,831 | 95,828 | 95,828 | 95,831 |
| $X_0$ | 8,75E-03 | 8,75E-03 | 8,77E-03 | 0,054 | 0,054 | 0,054 |
| $V(\hat{X}_0)$ | 4,06E-06 | 3,72E-06 | 1,26E-06 | 3,81E-06 | 3,32E-06 | 1,17E-06 |
| $\sigma_\delta^2$ | - | 8,31E-06 | - | - | 8,24E-06 | - |
| $U(X_0)$ | 5,18E-03 | 4,96E-03 | 2,89E-03 | 5,02E-03 | 4,68E-03 | 2,78E-03 |

# 5 Concluding remarks

In general, the simulation study reveals that the proposed model is sensible to the presence of error related to the independent variable and gives better results in contrast to the usual model results. It was noticed that when the error variance $\sigma_\delta^2$ increases, the mean estimated variance of $\hat{X}_0$ obtained using the usual model moves away from the theoretical value. In the example above, the confidence interval amplitude from the proposed models are supplied by the incorporation of error due to the lecture of equipment and the preparation of the standard solutions. It is observed that despite the classical model only considers the error originated from the lecture of the equipment, the amplitude is greater than the obtained by the new approach.

**Acknowledgements**

Betsabé G. Blas achic has been partially supported by IPT (São Paulo).

# Appendix

## A Bias and variance for the maximum likelihood estimator

In the following we derive the bias (2.14) and the variance (2.15) of the estimator $\hat{X}_0$ from the homoscedastic controlled calibration model when $\sigma_\delta^2$ is known.

Considering the model (2.1) and (2.2), the estimator $\hat{X}_0 = (\bar{Y}_0 - \hat{\alpha})/\hat{\beta}$ can be expressed as

$$\hat{X}_0 = \bar{X} + \frac{\beta(X_0 - \bar{X}) + \bar{\epsilon}_0 - \bar{\phi}}{\hat{\beta}}, \tag{A.1}$$

where $\bar{\epsilon}_0 = \sum_{i=n+1}^{n+k} \epsilon_i/k$ and $\bar{\phi} = \sum_{i=1}^{n}(\epsilon_i - \beta\delta_i)/n$.

Considering $k$ fixed, expanding $1/\hat{\beta}$ in a Taylor series around $\beta$ and ignoring terms of order less than $n^{-2}$, we obtain the expected value of (A.1), given by

$$E(\hat{X}_0) = X_0 + \frac{\gamma(\bar{X} - X_0)}{n\beta^2 S_{XX}}. \tag{A.2}$$

From this last equation we get the bias (2.14).

To derive the variance (2.15) we take the variance of (A.1), which is given by

$$V(\hat{X}_0) = \beta^2(X_0 - \bar{X})^2 V(\frac{1}{\hat{\beta}}) + V(\frac{\bar{\epsilon}_0}{\hat{\beta}}) + V(\frac{\bar{\phi}}{\hat{\beta}}). \tag{A.3}$$

We call attention to the fact that (A.3) is only expressed as a function of the related variances because the corresponding covariances are zero. The variances $V(1/\hat{\beta}), V(\bar{\epsilon}_0/\hat{\beta})$ and $V(\bar{\phi}/\hat{\beta})$ can be obtained by expanding $1/\hat{\beta}$, $\bar{\epsilon}_0/\hat{\beta}$ and $\bar{\phi}/\hat{\beta}$ in a Taylor series around $\beta$ and ignoring terms of order less than $n^{-2}$. They are given by

$$V(1/\hat{\beta}) = \frac{V(\hat{\beta})}{\beta^4}, \tag{A.4}$$

$$V(\bar{\epsilon}_0/\hat{\beta}) = \frac{\sigma_\epsilon^2}{k\beta^2} + 3\frac{\sigma_\epsilon^2}{k\beta^4}V(\hat{\beta}), \tag{A.5}$$

$$V(\bar{\phi}/\hat{\beta}) = \frac{\gamma}{n\beta^2}. \tag{A.6}$$

Substituing (A.4), (A.5) and (A.6) in (A.3), then, the variance (2.15) is obtained.

# B Tables

Table B1. Empirical bias and mean squared error, theoretical variance and the mean estimated variance of $\hat{X}_0$, for $\sigma_\delta^2 = 0,01$ and unknown.

| $X_0$ | $n$ | $k$ | Empirical | | Theoretical | | Mean of $\hat{V}(\hat{X}_0)$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | Proposed-M | | Usual-M | Proposed-M | |
| | | | Bias | MSE | $V_1(X_0)$ | $V_2(X_0)$ | $V_1(X_0)$ | $V_1(X_0)$ | $V_2(X_0)$ |
| 0,01 | 5 | 2 | -0,0060 | 0,0180 | 0,0170 | 0,0170 | 0,0120 | 0,0100 | 0,0100 |
| | | 20 | -0,0087 | 0,0130 | 0,0120 | 0,0120 | 0,0072 | 0,0120 | 0,0120 |
| | | 100 | -0,0060 | 0,0130 | 0,0120 | 0,0120 | 0,0065 | 0,0120 | 0,0120 |
| | 20 | 2 | -0,0038 | 0,0086 | 0,0087 | 0,0087 | 0,0120 | 0,0052 | 0,0053 |
| | | 20 | -0,0028 | 0,0043 | 0,0042 | 0,0042 | 0,0033 | 0,0040 | 0,0040 |
| | | 100 | -0,0032 | 0,0038 | 0,0038 | 0,0038 | 0,0022 | 0,0036 | 0,0036 |
| | 100 | 2 | -0,0023 | 0,0058 | 0,0058 | 0,0058 | 0,0100 | 0,0027 | 0,0027 |
| | | 20 | -0,0002 | 0,0013 | 0,0013 | 0,0013 | 0,0016 | 0,0012 | 0,0012 |
| | | 100 | -0,0007 | 0,0008 | 0,0009 | 0,0009 | 0,0007 | 0,0009 | 0,0009 |
| 0,8 | 5 | 2 | -0,0011 | 0,0091 | 0,0093 | 0,0094 | 0,0079 | 0,0045 | 0,0046 |
| | | 20 | -0,0034 | 0,0050 | 0,0048 | 0,0048 | 0,0029 | 0,0045 | 0,0046 |
| | | 100 | -0,0007 | 0,0047 | 0,0044 | 0,0044 | 0,0024 | 0,0045 | 0,0045 |
| | 20 | 2 | 0,0005 | 0,0063 | 0,0061 | 0,0061 | 0,0095 | 0,0028 | 0,0029 |
| | | 20 | 0,0007 | 0,0016 | 0,0016 | 0,0016 | 0,0015 | 0,0015 | 0,0015 |
| | | 100 | -0,0001 | 0,0012 | 0,0012 | 0,0012 | 0,0008 | 0,0012 | 0,0012 |
| | 100 | 2 | 0,0005 | 0,0050 | 0,0052 | 0,0052 | 0,0099 | 0,0021 | 0,0021 |
| | | 20 | -0,0001 | 0,0007 | 0,0007 | 0,0007 | 0,0011 | 0,0007 | 0,0007 |
| | | 100 | -0,0003 | 0,0003 | 0,0003 | 0,0003 | 0,0003 | 0,0003 | 0,0003 |
| 1,9 | 5 | 2 | 0,0041 | 0,0160 | 0,0150 | 0,0160 | 0,0120 | 0,0093 | 0,0094 |
| | | 20 | 0,0026 | 0,0110 | 0,0110 | 0,0110 | 0,0065 | 0,0110 | 0,0110 |
| | | 100 | 0,0076 | 0,0110 | 0,0110 | 0,0110 | 0,0058 | 0,0110 | 0,0110 |
| | 20 | 2 | 0,0006 | 0,0079 | 0,0082 | 0,0082 | 0,0110 | 0,0049 | 0,0049 |
| | | 20 | 0,0040 | 0,0039 | 0,0037 | 0,0037 | 0,0030 | 0,0035 | 0,0035 |
| | | 100 | 0,0008 | 0,0033 | 0,0033 | 0,0033 | 0,0019 | 0,0031 | 0,0031 |
| | 100 | 2 | 0,0020 | 0,0057 | 0,0057 | 0,0057 | 0,0100 | 0,0025 | 0,0025 |
| | | 20 | 0,0003 | 0,0012 | 0,0012 | 0,0012 | 0,0015 | 0,0011 | 0,0011 |
| | | 100 | 0,0003 | 0,0008 | 0,0008 | 0,0008 | 0,0006 | 0,0008 | 0,0008 |

Table B2. Empirical bias and mean squared error, theoretical variance and the mean estimated variance of $\hat{X}_0$, for $\sigma_\delta^2 = 0,1$ and unknown.

| $X_0$ | $n$ | $k$ | Empirical | | Theoretical Proposed-M | | Mean of $\hat{V}(\hat{X}_0)$ Usual-M | Mean of $\hat{V}(\hat{X}_0)$ Proposed-M | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Bias | MSE | $V_1(X_0)$ | $V_2(X_0)$ | $\hat{V}_1(\hat{X}_0)$ | $\hat{V}_1(\hat{X}_0)$ | $\hat{V}_2(\hat{X}_0)$ |
| 0,01 | 5 | 2 | -0,0510 | 0,1000 | 0,0700 | 0,0710 | 0,0770 | 0,0680 | 0,0690 |
| | | 20 | -0,0500 | 0,0950 | 0,0660 | 0,0660 | 0,0220 | 0,0660 | 0,0660 |
| | | 100 | -0,0510 | 0,0950 | 0,0650 | 0,0650 | 0,0130 | 0,0730 | 0,0730 |
| | 20 | 2 | -0,0180 | 0,0280 | 0,0250 | 0,0250 | 0,0670 | 0,0240 | 0,0240 |
| | | 20 | -0,0160 | 0,0230 | 0,0210 | 0,0210 | 0,0140 | 0,0210 | 0,0210 |
| | | 100 | -0,0170 | 0,0230 | 0,0200 | 0,0200 | 0,0055 | 0,0210 | 0,0210 |
| | 100 | 2 | -0,0016 | 0,0094 | 0,0093 | 0,0093 | 0,0580 | 0,0069 | 0,0069 |
| | | 20 | -0,0026 | 0,0048 | 0,0048 | 0,0048 | 0,0082 | 0,0048 | 0,0048 |
| | | 100 | -0,0033 | 0,0043 | 0,0044 | 0,0044 | 0,0029 | 0,0044 | 0,0044 |
| 0,8 | 5 | 2 | -0,0084 | 0,0370 | 0,0290 | 0,0290 | 0,0460 | 0,0250 | 0,0250 |
| | | 20 | -0,0095 | 0,0270 | 0,0240 | 0,0240 | 0,0071 | 0,0200 | 0,0200 |
| | | 100 | -0,0072 | 0,0290 | 0,0240 | 0,0240 | 0,0037 | 0,0200 | 0,0200 |
| | 20 | 2 | -0,0030 | 0,0120 | 0,0110 | 0,0110 | 0,0530 | 0,0085 | 0,0086 |
| | | 20 | -0,0040 | 0,0068 | 0,0066 | 0,0066 | 0,0061 | 0,0064 | 0,0064 |
| | | 100 | -0,0031 | 0,0063 | 0,0062 | 0,0062 | 0,0017 | 0,0060 | 0,0060 |
| | 100 | 2 | -0,0011 | 0,0063 | 0,0062 | 0,0063 | 0,0550 | 0,0038 | 0,0038 |
| | | 20 | -0,0015 | 0,0017 | 0,0017 | 0,0017 | 0,0057 | 0,0017 | 0,0017 |
| | | 100 | -0,0011 | 0,0014 | 0,0013 | 0,0013 | 0,0013 | 0,0013 | 0,0013 |
| 1,9 | 5 | 2 | 0,0430 | 0,1090 | 0,0630 | 0,0630 | 0,0830 | 0,0750 | 0,0750 |
| | | 20 | 0,0450 | 0,0860 | 0,0580 | 0,0580 | 0,0210 | 0,0650 | 0,0650 |
| | | 100 | 0,0410 | 0,0860 | 0,0580 | 0,0580 | 0,0110 | 0,0600 | 0,0600 |
| | 20 | 2 | 0,0160 | 0,0260 | 0,0230 | 0,0230 | 0,0650 | 0,0210 | 0,0210 |
| | | 20 | 0,0140 | 0,0190 | 0,0180 | 0,0180 | 0,0130 | 0,0180 | 0,0180 |
| | | 100 | 0,0170 | 0,0200 | 0,0180 | 0,0180 | 0,0048 | 0,0180 | 0,0180 |
| | 100 | 2 | 0,0050 | 0,0088 | 0,0087 | 0,0088 | 0,0570 | 0,0063 | 0,0063 |
| | | 20 | 0,0030 | 0,0043 | 0,0042 | 0,0042 | 0,0078 | 0,0042 | 0,0042 |
| | | 100 | 0,0020 | 0,0039 | 0,0038 | 0,0038 | 0,0026 | 0,0038 | 0,0038 |

Table B3. Covering percentage (%) and amplitude (A) of the intervals with a 95% confidence level for the parameter $X_0$, when $\sigma_\delta^2 = 0,01$ and unknown.

| $X_0$ | $n$ | $k$ | Usual-M | | Proposed-M | | | |
|---|---|---|---|---|---|---|---|---|
| | | | % | $A$ | $\%_1$ | $A_1$ | $\%_2$ | $A_2$ |
| 0,01 | 5 | 2 | 83,01 | 0,40 | 79,34 | 0,36 | 79,37 | 0,36 |
| | | 20 | 84,95 | 0,32 | 91,15 | 0,41 | 91,15 | 0,41 |
| | | 100 | 83,72 | 0,31 | 92,24 | 0,42 | 92,24 | 0,42 |
| | 20 | 2 | 96,46 | 0,42 | 83,48 | 0,28 | 83,52 | 0,28 |
| | | 20 | 90,19 | 0,22 | 92,47 | 0,24 | 92,47 | 0,24 |
| | | 100 | 86,16 | 0,18 | 92,78 | 0,23 | 92,78 | 0,23 |
| | 100 | 2 | 98,84 | 0,40 | 73,16 | 0,19 | 73,16 | 0,19 |
| | | 20 | 96,71 | 0,16 | 94,14 | 0,14 | 94,14 | 0,14 |
| | | 100 | 91,90 | 0,11 | 94,68 | 0,12 | 94,68 | 0,12 |
| 0,8 | 5 | 2 | 85,43 | 0,32 | 74,33 | 0,24 | 74,39 | 0,24 |
| | | 20 | 85,16 | 0,21 | 91,34 | 0,26 | 91,34 | 0,26 |
| | | 100 | 85,01 | 0,19 | 92,50 | 0,25 | 92,50 | 0,25 |
| | 20 | 2 | 97,90 | 0,38 | 73,55 | 0,20 | 73,55 | 0,20 |
| | | 20 | 93.55 | 0,15 | 93,89 | 0,15 | 93,89 | 0,15 |
| | | 100 | 86,53 | 0,11 | 93,12 | 0,13 | 93,12 | 0,13 |
| | 100 | 2 | 99,41 | 0,39 | 65,05 | 0,16 | 65,05 | 0,16 |
| | | 20 | 98,51 | 0.13 | 91,05 | 0,10 | 94,05 | 0,10 |
| | | 100 | 94.56 | 0,07 | 94,86 | 0,07 | 94,86 | 0,07 |
| 1,9 | 5 | 2 | 82,47 | 0,39 | 78,05 | 0,35 | 78,13 | 0,35 |
| | | 20 | 84.50 | 0,31 | 90,92 | 0,39 | 90,92 | 0,39 |
| | | 100 | 84.75 | 0,29 | 92,83 | 0,39 | 92,83 | 0,39 |
| | 20 | 2 | 96.79 | 0,41 | 83,09 | 0,26 | 83,11 | 0,26 |
| | | 20 | 91,32 | 0,21 | 93,29 | 0,23 | 93,31 | 0,23 |
| | | 100 | 86,43 | 0,17 | 93,32 | 0,22 | 93,32 | 0,22 |
| | 100 | 2 | 98,88 | 0,10 | 73,06 | 0,19 | 73,06 | 0,19 |
| | | 20 | 97,12 | 0,15 | 94,31 | 0,13 | 94,31 | 0,13 |
| | | 100 | 92,56 | 0.10 | 94,74 | 0,11 | 94,74 | 0,11 |

Table B4. Covering percentage (%) and amplitude (A) of the intervals with a 95% confidence level for the parameter $X_0$, when $\sigma_\delta^2 = 0,1$ and unknown.

| $X_0$ | $n$ | $k$ | Usual-M | | Proposed-M | | | |
|---|---|---|---|---|---|---|---|---|
| | | | % | $A$ | $\%_1$ | $A_1$ | $\%_2$ | $A_2$ |
| 0,01 | 5 | 2 | 84,89 | 0,94 | 80,73 | 0,85 | 80,79 | 0,86 |
| | | 20 | 64,10 | 0,51 | 82,75 | 0,86 | 82,75 | 0,86 |
| | | 100 | 52,09 | 0,39 | 82,42 | 0,86 | 82,42 | 0,86 |
| | 20 | 2 | 99,40 | 0,99 | 91,04 | 0,58 | 91,10 | 0,58 |
| | | 20 | 87,10 | 0,45 | 92,60 | 0,55 | 92,62 | 0,55 |
| | | 100 | 65,70 | 0,28 | 92,16 | 0,55 | 92,18 | 0,55 |
| | 100 | 2 | 100,00 | 0,94 | 87,60 | 0,32 | 87,66 | 0,32 |
| | | 20 | 98,82 | 0,36 | 94,88 | 0,27 | 94,88 | 0,27 |
| | | 100 | 89,12 | 0,21 | 94,84 | 0,26 | 94,84 | 0,26 |
| 0,8 | 5 | 2 | 90,27 | 0,73 | 80,63 | 0,52 | 80,75 | 0,52 |
| | | 20 | 64,39 | 0,31 | 82,68 | 0,51 | 82,74 | 0,51 |
| | | 100 | 50,57 | 0,23 | 84,08 | 0,50 | 84,08 | 0,50 |
| | 20 | 2 | 99,92 | 0.88 | 87,30 | 0,35 | 87,50 | 0,35 |
| | | 20 | 91,80 | 0,30 | 92,78 | 0,31 | 92,82 | 0,31 |
| | | 100 | 67,46 | 0,16 | 92,38 | 0,30 | 92,38 | 0,30 |
| | 100 | 2 | 100,00 | 0.91 | 77,88 | 0,22 | 77,98 | 0,22 |
| | | 20 | 99,91 | 0,29 | 94,98 | 0,16 | 95,04 | 0,16 |
| | | 100 | 94,60 | 0.14 | 94,92 | 0,14 | 94,92 | 0,14 |
| 1,9 | 5 | 2 | 85,63 | 0.91 | 81,42 | 0,81 | 81,42 | 0,81 |
| | | 20 | 61.71 | 0.49 | 81,89 | 0,82 | 81,89 | 0,82 |
| | | 100 | 50.71 | 0,37 | 81,24 | 0,81 | 81,24 | 0,81 |
| | 20 | 2 | 99.54 | 0,97 | 91,12 | 0,54 | 91,18 | 0,55 |
| | | 20 | 88,10 | 0,13 | 92,74 | 0,52 | 92,76 | 0,52 |
| | | 100 | 66,44 | 0,26 | 92,86 | 0,51 | 92,86 | 0,51 |
| | 100 | 2 | 100.00 | 0,93 | 86,38 | 0,30 | 86,38 | 0,30 |
| | | 20 | 99.12 | 0.35 | 94,76 | 0,25 | 94,76 | 0,25 |
| | | 100 | 89.40 | 0.20 | 95,00 | 0,24 | 95,00 | 0,24 |

Table B5. Empirical bias and mean squared error, theoretical variance and the mean estimated variance of $\hat{X}_0$, for $\sigma_\delta^2 = 0,01$ and known.

| $X_0$ | $n$ | $k$ | Empirical | | | | Theoretical | Mean of $V(X_0)$ | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Usual-M | | Proposed-M | | Proposed-M | Usual-M | Proposed-M |
| | | | Bias | MSE | Bias | MSE | $V(X_0)$ | | |
| 0,01 | 5 | 2 | -0,0290 | 0,0210 | -0.0280 | 0,0210 | 0,0170 | 0,0180 | 0,0160 |
| | | 20 | -0,0290 | 0,0140 | -0.0320 | 0,0140 | 0,0120 | 0,0081 | 0,0130 |
| | | 100 | -0,0240 | 0,0140 | -0.0270 | 0,0140 | 0,0120 | 0,0070 | 0,0130 |
| | 20 | 2 | -0,0081 | 0,0091 | -0,0064 | 0,0090 | 0,0086 | 0,0130 | 0,0076 |
| | | 20 | -0,0060 | 0,0043 | -0,0072 | 0,0043 | 0,0041 | 0,0034 | 0,0041 |
| | | 100 | -0,0038 | 0,0038 | -0,0060 | 0,0038 | 0,0037 | 0,0022 | 0,0038 |
| | 100 | 2 | -0,0011 | 0,0056 | -0.0005 | 0.0056 | 0,0058 | 0,0100 | 0,0053 |
| | | 20 | -0,0002 | 0,0013 | -0,0001 | 0,0013 | 0,0013 | 0,0016 | 0,0012 |
| | | 100 | -0,0009 | 0,0009 | -0.0012 | 0.0009 | 0,0009 | 0,0007 | 0,0009 |
| 0,8 | 5 | 2 | -0,0074 | 0,0110 | -0.0072 | 0,0100 | 0,0093 | 0,0120 | 0,0085 |
| | | 20 | -0,0046 | 0,0051 | -0.0051 | 0,0052 | 0,0048 | 0,0032 | 0,0049 |
| | | 100 | -0,0076 | 0,0048 | -0.0082 | 0,0018 | 0,0044 | 0,0025 | 0,0046 |
| | 20 | 2 | -0,0034 | 0,0063 | -0.0031 | 0,0063 | 0,0061 | 0,0100 | 0,0052 |
| | | 20 | 0,0001 | 0,0016 | 0.0000 | 0,0016 | 0,0016 | 0,0015 | 0,0016 |
| | | 100 | -0,0009 | 0,0012 | -0.0013 | 0.0012 | 0,0012 | 0,0008 | 0,0012 |
| | 100 | 2 | 0,0000 | 0,0053 | 0.0002 | 0,0053 | 0,0052 | 0,0099 | 0,0048 |
| | | 20 | 0,0000 | 0,0007 | 0.0000 | 0.0007 | 0,0007 | 0,0011 | 0,0007 |
| | | 100 | -0,0001 | 0,0003 | -0.0002 | 0.0003 | 0,0003 | 0,0003 | 0,0003 |
| 1,9 | 5 | 2 | 0,0200 | 0,0180 | 0.0200 | 0,0180 | 0,0150 | 0,0170 | 0,0140 |
| | | 20 | 0,0240 | 0,0120 | 0.0260 | 0,0130 | 0,0110 | 0,0071 | 0,0120 |
| | | 100 | 0,0200 | 0,0130 | 0.0230 | 0.0130 | 0,0100 | 0,0062 | 0,0110 |
| | 20 | 2 | 0,0037 | 0,0082 | 0.0021 | 0,0081 | 0,0082 | 0,0120 | 0,0071 |
| | | 20 | 0,0059 | 0.0037 | 0.0066 | 0.0037 | 0,0037 | 0,0030 | 0,0036 |
| | | 100 | 0,0033 | 0,0032 | 0.0051 | 0,0032 | 0,0032 | 0,0020 | 0,0033 |
| | 100 | 2 | 0,0020 | 0,0058 | 0.0015 | 0,0057 | 0,0057 | 0,0100 | 0,0053 |
| | | 20 | 0,0003 | 0,0012 | 0,0002 | 0.0012 | 0,0012 | 0,0015 | 0,0011 |
| | | 100 | 0,0003 | 0,0008 | 0.0006 | 0,0008 | 0,0008 | 0,0006 | 0,0008 |

Table B6. Empirical bias and mean squared error, theoretical variance and the mean estimated variance of $\hat{X}_0$, for $\sigma_\delta^2 = 0,1$ and known.

| $X_0$ | $n$ | $k$ | Empirical | | | | Theoretical | Mean of $\hat{V}(X_0)$ | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Usual-M | | Proposed-M | | Proposed-M | Usual-M | Proposed-M |
| | | | Bias | MSE | Bias | MSE | $V(X_0)$ | | |
| 0,01 | 5 | 2 | -0,4330 | 0,5590 | -0.3830 | 0,4580 | 0,0590 | 0,3650 | 0,2310 |
| | | 20 | -0,1140 | 0,1310 | 0.4140 | 1,0880 | 0,0540 | 0,0250 | 0,0760 |
| | | 100 | -0,1890 | 0,1540 | -0,0290 | 0,7730 | 0,0540 | 0,0200 | 0,1250 |
| | 20 | 2 | -0,0930 | 0,0430 | -0,0770 | 0,0360 | 0,0210 | 0,0950 | 0,0380 |
| | | 20 | -0,0490 | 0,0210 | -0.0510 | 0,0210 | 0,0160 | 0,0150 | 0,0180 |
| | | 100 | -0,0430 | 0,0190 | -0.0250 | 0,0640 | 0,0150 | 0,0058 | 0,0170 |
| | 100 | 2 | -0,0200 | 0,0110 | -0.0160 | 0,0097 | 0,0084 | 0,0630 | 0,0110 |
| | | 20 | -0,0077 | 0,0041 | -0.0085 | 0,0039 | 0,0037 | 0,0084 | 0,0038 |
| | | 100 | -0,0066 | 0.0037 | -0.0087 | 0,0037 | 0,0033 | 0,0029 | 0,0033 |
| 0,8 | 5 | 2 | -0,0500 | 0,0620 | -0.0130 | 0,0550 | 0,0280 | 0,1150 | 0,0620 |
| | | 20 | -0,0450 | 0,0380 | 0.0820 | 0,0810 | 0,0240 | 0,0086 | 0,0320 |
| | | 100 | -0,0530 | 0,0510 | -0.0069 | 0.0980 | 0,0230 | 0,0055 | 0,0360 |
| | 20 | 2 | -0,0280 | 0,0150 | -0.0240 | 0,0140 | 0,0110 | 0,0740 | 0,0210 |
| | | 20 | -0,0079 | 0,0068 | -0,0085 | 0,0069 | 0,0064 | 0,0065 | 0,0067 |
| | | 100 | -0,0055 | 0,0065 | -0.0030 | 0,0079 | 0,0060 | 0,0018 | 0,0062 |
| | 100 | 2 | -0,0030 | 0,0067 | -0.0021 | 0,0065 | 0,0062 | 0,0600 | 0,0091 |
| | | 20 | -0,0018 | 0,0018 | -0.0018 | 0,0017 | 0,0017 | 0,0058 | 0,0017 |
| | | 100 | -0,0021 | 0.0013 | -0.0025 | 0.0013 | 0,0013 | 0,0013 | 0,0013 |
| 1,9 | 5 | 2 | 0,3480 | 0,3770 | 0.3070 | 0.3000 | 0,0530 | 0,2850 | 0,1760 |
| | | 20 | 0,1400 | 0,1630 | -0.3310 | 0.9370 | 0,0490 | 0,0300 | 0,0780 |
| | | 100 | 0,0410 | 0.0180 | 0.0270 | 0,0510 | 0,0140 | 0,0051 | 0,0150 |
| | 20 | 2 | 0,0970 | 0,0430 | 0.0790 | 0.0350 | 0,0190 | 0,0930 | 0,0350 |
| | | 20 | 0,1400 | 0,1630 | -0.3310 | 0.9370 | 0,0490 | 0,0300 | 0,0780 |
| | | 100 | 0,1670 | 0,1300 | -0.0005 | 0.7090 | 0,0480 | 0,0160 | 0,1120 |
| | 100 | 2 | 0,0200 | 0,0093 | 0.0160 | 0,0087 | 0,0080 | 0,0620 | 0,0110 |
| | | 20 | 0,0120 | 0,0039 | 0.0130 | 0,0037 | 0,0033 | 0,0080 | 0,0035 |
| | | 100 | 0,0072 | 0,0031 | 0.0089 | 0.0031 | 0,0029 | 0,0027 | 0,0030 |

Table B7. Covering percentage (%) and amplitude (A) of the intervals with a 95% confidence level for the parameter $X_0$, when $\sigma_\delta^2 = 0,01$. and 0,1 and known.

| $X_0$ | $n$ | $k$ | $\sigma_\delta^2 = 0,01$ | | | | $\sigma_\delta^2 = 0,1$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Usual-M | | Proposed-M | | Usual-M | | Proposed-M | |
| | | | % | A | % | A | % | A | % | A |
| 0,01 | 5 | 2 | 92,10 | 0.51 | 91,20 | 0,48 | 95,06 | 1,89 | 92,40 | 1,49 |
| | | 20 | 87,32 | 0,34 | 95,18 | 0,44 | 63,38 | 0,59 | 64,47 | 1,09 |
| | | 100 | 81.50 | 0.32 | 95,19 | 0,44 | 52,69 | 0,48 | 89,24 | 1,18 |
| | 20 | 2 | 97,46 | 0,43 | 90,12 | 0,33 | 99,87 | 1,17 | 96,27 | 0,71 |
| | | 20 | 91,00 | 0,23 | 94,03 | 0,25 | 92,76 | 0,48 | 94,48 | 0,52 |
| | | 100 | 86.73 | 0,19 | 95,21 | 0,24 | 70,38 | 0,30 | 90,35 | 0,51 |
| | 100 | 2 | 97,01 | 0,43 | 90,00 | 0,33 | 100,00 | 0,98 | 93,41 | 0,40 |
| | | 20 | 92,34 | 0.23 | 95,26 | 0,25 | 99,77 | 0,36 | 95,09 | 0,24 |
| | | 100 | 85,93 | 0,19 | 94,55 | 0,24 | 92,70 | 0,21 | 94,42 | 0,23 |
| 0,8 | 5 | 2 | 94,33 | 0,41 | 89,28 | 0,35 | 98,77 | 1,31 | 97,54 | 0,98 |
| | | 20 | 86,84 | 0,22 | 95,03 | 0,27 | 61,85 | 0,34 | 81,25 | 0,69 |
| | | 100 | 85.65 | 0,19 | 95,56 | 0,27 | 49,29 | 0,27 | 87,82 | 0,69 |
| | 20 | 2 | 98,35 | 0,39 | 88,41 | 0,27 | 100,00 | 1,03 | 94,09 | 0,53 |
| | | 20 | 93.84 | 0,15 | 91,41 | 0,15 | 94,44 | 0,32 | 95,11 | 0,32 |
| | | 100 | 85,98 | 0,11 | 91,32 | 0,14 | 70,38 | 0,17 | 92,41 | 0,31 |
| | 100 | 2 | 98,08 | 0,39 | 88.22 | 0,27 | 100,00 | 0,95 | 92,25 | 0,34 |
| | | 20 | 92,75 | 0,15 | 93.96 | 0.15 | 99,89 | 0,30 | 94,89 | 0,16 |
| | | 100 | 87.17 | 0.11 | 91.60 | 0.14 | 95,00 | 0,14 | 94,88 | 0,14 |
| 1,9 | 5 | 2 | 92.30 | 0,50 | 90,61 | 0,46 | 96,26 | 1,78 | 92,14 | 1,45 |
| | | 20 | 86,96 | 0.33 | 95,13 | 0,42 | 63,58 | 0,56 | 67,24 | 1,03 |
| | | 100 | 85.84 | 0.30 | 91,73 | 0,42 | 49,30 | 0,46 | 83,75 | 1,21 |
| | 20 | 2 | 97,61 | 0,43 | 89.92 | 0,32 | 100,00 | 1,17 | 96,27 | 0,70 |
| | | 20 | 91.60 | 0,21 | 91.27 | 0,23 | 94,26 | 0,46 | 94,95 | 0,49 |
| | | 100 | 86,16 | 0,17 | 91.99 | 0,23 | 72,44 | 0,28 | 94,18 | 0,48 |
| | 100 | 2 | 97.04 | 0,43 | 89.70 | 0.32 | 100,00 | 0,97 | 94,27 | 0,39 |
| | | 20 | 91,21 | 0.21 | 93.85 | 0.23 | 99,59 | 0,35 | 95,18 | 0,23 |
| | | 100 | 87.04 | 0.17 | 95.11 | 0.23 | 93,38 | 0,20 | 94,98 | 0,21 |

## References


[1] Berkson, J. Estimation of a linear function for a calibration line: consideration of a recent proposal. *Technometrics* **11** (1969) 649.

[2] Williams, E.J. A note on regression methods in calibration. *Technometrics* **11** (1969) 192.

[3] Lwin, T. Discussion of Hunter and Lamboy's paper. *Technometrics* **14** (1981) 339.

[4] Shukla, G.K. On the problem of calibration. *Technometrics* **14** (1972) 547.

[5] Brown, P.J. Measurement, regression and calibration. Oxford: Oxford University Press.

[6] Tellinghuisen, J. Inverse vs. classical calibration for small data sets. *Fresenius J Anal Chem* 368 (2000) 585.

[7] Aitchison, J. and Dunsmore, I. (1975). Statistical Prediction Analysis. Cambridge University Press.

[8] Osbourne, C. Statistical calibration: A review. *International Statistical Review* 59 (1991) 309.

[9] Cheng, Chi-Lun, Van Ness J. W. (1999). Statistical regression with measurement error. London; New York : Arnold: Oxford University Press.

[10] Sociedade Brasileira de Metrologia (SBM) (2002). Determinando a Incerteza em Medicão Analítica. 2. ed. Brasileira - Traducão do Guia EURACHEM /CITAC.

[11] Berkson, J. Are there two regression?. *Journal of the American Statistical Association* 45 (1950) 164.

[12] Fuller, W. A. (1987). Measurement error models. John Wiley Sons, New York.

[13] Lima, C.R.P (1996). Calibracão absoluta com erro nas variáveis. São Paulo: IME-USP. Tese de Doutorado.

[14] Polak, E. (1997). Optimization Algorithms and Consistent Approximations. New York: Springer-Verlag.

[15] Blas, B. G.(2005). Calibracão controlada aplicada à química analítica. São Paulo: IME-USP. Dissertacão de Mestrado.

## ÚLTIMOS RELATÓRIOS TÉCNICOS PUBLICADOS

**2006-01 - SILVA, M.F., FERRARI, S.L.P., CRIBARI-NETO, F.** Improved likelihood inference for the shape parameter in Weibull regression. 2006. 24p. (RT-MAE 2006-01)

**2006-02 - LACHOS, V.H., BOLFARINE, H., ARELLANO-VALLE, R. B., MONTENEGRO, L. C.** Inference in multivariate skew-normal regression models. 2006. 21p. (RT-MAE 2006-02)

**2006-03 - BUENO, V.C.** Redundancy allocation for a coherent system of dependent components. 2006. 16p. (RT-MAE 2006-03)

**2006-04 - ESPINHEIRA, P.L., FERRARI, S. L. P., CRIBARI-NETO, F.** On Beta regression residuals. 2006. 17p. (RT-MAE 2006-04)

**2006-05 - BAZÁN, J. L., BOLFARINE, H., BRANCO, M. D.** A generalized skew probit class link for binary regression. 2006. 22p. (RT-MAE 2006-05)

**2006-06 - MIRANDA, J. C. S. DE; MORETTIN, P. A.** On the estimation of the intensity of point processes via wavelets. 2006. 45p. (RT-MAE 2006-06)

**2006-07 - DUARTE, D.; GALVES, J.; GARCIA, N. L.** Markov approximation and consistent estimation of unbounded probabilistic suffix trees. 2006. 12p. (RT-MAE 2006-06)

**2006-08 - RIFO, L. R.; WECHSLER, S.** Regularly varying densities for St. Petersburg envelopes. 2006. 8p. (RT-MAE 2006-08)

**2006-09 - FERRARI, S. L. P.; CYSNEIROS, A. H. M. A.** Skovgaad's adjustment to likelihood ratio tests in exponential family nonlinear models. 2006. 15p. (RT-MAE 2006-09)

**2006-10 - POLPO, A., PEREIRA, C. A. B.,** Parallel and series systems: rates, distributions and sub-distributions. 2006. 13p. (RT-MAE 2006-10)

**2006-11 - LACHOS, V. H., BOLFARINE, H., GENTON, M.G.** Bayesian inference for measurement error models with flexible skew-symmetric distributions. 2006. 29p. (RT-MAE 2006-11)

**2006-12 - BUENO, V.C.** A Martingale Estimator for a Structural Importance Measure of a Coherent System's Component Through its Critical Level.

**2006-13 - BUENO, V.C.** Estimating system´s component reliability importance under dependence conditions. 2006. 09p. (RT-MAE 2006-13)

**2006-14 - PANZIERI, A., BELITSKY, V., ROCHA, P. R. M., ALMEIDA PRADO, F. P.** An implicit measure of tail dependence and its application for identification of the charge of dependence between assets in market booms and crashes. 2006. 29p. (RT-MAE-2006-14)

**2006-15 - ALMEIDA, A. E. ELIAN, S. N., NOBRE, S. J.** Modificações e Alternativas aos Teste de Levene e de Brown e forsythe para Igualdade de variâncias e Médias. 2006. 26p. (RT-MAE-2006-15)

**2006-16 - ROMEO, J.S., TANAKA, N.T., LIMA, A.C.P.** Functional Central Limit Theorem for Copulas in Survival Analysis. 2006. 24p. (RT-MAE-2006-16)

**200-17 - ESPINHEIRA, P.L., FERRARI, S.L.P., CRIBARI-NETO, F.** Influence Diagnostics ina a Class of Econometric Regression Models for Rates and Proportions. 2006. 19p. (RT-MAE-2006-17)

**The complete list of "Relatórios do Departamento de Estatística", IME-USP, will be sent upon request.**

*Departamento de Estatística*
*IME-USP*
*Caixa Postal 66.281*
*05311-970 - São Paulo, Brasil*