Accepted Manuscript

Statistical calibration and exact one-sided simultaneous tolerance intervals for polynomial regression

Yang Han, Wei Liu, Frank Bretz, Fang Wan, Ping Yang

Please cite this article as: Han, Y., Liu, W., Bretz, F., Wan, F., Yang, P., Statistical calibration and exact one-sided simultaneous tolerance intervals for polynomial regression. J. Statist. Plann. Inference (2015), http://dx.doi.org/10.1016/j.jspi.2015.07.005

Corresponding Author: Dr. W. Liu,

Corresponding Author's Institution:

First Author: W. Liu

Order of Authors: W. Liu; Yang Han, PhD; Frank Bretz, PhD; Fang Wan, PhD; Ping Yang, PhD student

Abstract: Statistical calibration using linear regression is a useful statistical tool having many applications.
Calibration for infinitely many future y-values requires the construction of simultaneous tolerance intervals (STI's). As calibration often involves only two variables x and y and polynomial regression is probably the most frequently used model for relating y with x, construction of STI's for polynomial regression plays a key role in statistical calibration for infinitely many future y-values. The only exact STI's published in the statistical literature are provided by Mee {\it et al.} (1991) and Odeh and Mee (1990). But they are for a multiple linear regression model, in which the covariates are assumed to have no functional relationships. When applied to polynomial regression, the resultant STI's are conservative. In this paper, exact one-sided STI's have been constructed for a polynomial regression model over any given interval. The available computer program allows the exact methods developed in this paper to be implemented easily. Real examples are given for illustration.

Statistical calibration and exact one-sided simultaneous tolerance intervals for polynomial regression

Yang Han[a,b], Wei Liu[a,*], Frank Bretz[c,d], Fang Wan[a], Ping Yang[e]

[a]*S3RI and School of Mathematics, University of Southampton, Southampton SO17 1BJ, UK*
[b]*University of Exeter Medical School, Exeter, EX1 2LU, UK*
[c]*Novartis Pharma AG, Basel, 4002, Switzerland*
[d]*Shanghai University of Finance and Economics, Shanghai, 200433, China*
[e]*Department of Statistics, The Chinese University of Hong Kong, Hong Kong*

Abstract

Statistical calibration using linear regression is a useful statistical tool having many applications. Calibration for infinitely many future y-values requires the construction of simultaneous tolerance intervals (STI's). As calibration often involves only two variables x and y and polynomial regression is probably the most frequently used model for relating y with x, construction of STI's for polynomial regression plays a key role in statistical calibration for infinitely many future y-values. The only exact STI's published in the statistical literature are provided by Mee *et al.* (1991) and Odeh and Mee (1990). But they are for a multiple linear regression model, in which the covariates are assumed to have no functional relationships. When applied to polynomial regression, the resultant STI's are conservative. In this paper, one-sided exact STI's have been constructed for a polynomial regression model over any given interval. The available computer program allows the exact methods developed in this paper to be implemented easily. Real examples are given for illustration.

Keywords: Confidence level, Linear regression, Polynomial regression, Quantile line, Simultaneous confidence band, Simultaneous tolerance intervals, Statistical simulation

---

*Corresponding author.
 Email address: `w.liu@maths.soton.ac.uk` (Wei Liu)

1. Introduction

Statistical calibration using linear regression has a rich history going back to Eisenhart (1939). The problem involves a quantity of interest x which is expensive or difficult to measure, a surrogate quantity y which is cheaper or easy to measure, and the assumption that y and x are related by a linear regression model. For example, x is the true concentration of radon, $^{222}R_n$, while y is the concentration reading on an alpha track detector (ATD), at a place, or x is the true alcohol level in blood stream while y is the reading on a breathalyzer, of a driver. In order to use an observed y to infer the corresponding but unobserved x, a calibration experiment is carried out to measure y_{0i} corresponding to a known x_{0i} for $i = 1, \cdots, n$. A regression model of y on x is then fitted by using the training data $\mathcal{E} = \{(x_{0i}, y_{0i}), i = 1, \cdots, n\}$ and used to infer the x-values corresponding to infinitely many y-values to be observed in future. The inference for the x-value corresponding to one single future y-value is considered by Eisenhart (1939), Brown (1982) and Smith and Corbett (1987) among others, and the relevant literature is reviewed in Osborne (1991) and Brown (1993).

This paper focuses on inference for infinitely many future y-values. Specifically, a confidence set $C(y_x)$ for the unknown x corresponding to each observed future y_x is constructed and the infinite sequence of confidence sets $C(y_x)$ corresponding to an infinite sequence of observed future y_x-values has the property: *with confidence level γ, with respect to the randomness in the training data $\mathcal{E}$, that the proportion of confidence sets $C(y_x)$ containing the corresponding true x-values is at least β, where $0 < \gamma, \beta < 1$ are pre-specified constants.* This *property* can be expressed as

$$P_{\mathcal{E}} \left\{ \liminf_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} I_{\left\{ x_i \in C(y_{x_i}) \right\}} \geq \beta \right\} \geq \gamma \tag{1}$$

where I_A denotes the indicator function of the set A and hence $\frac{1}{N} \sum_{i=1}^{N} I_{\left\{ x_i \in C(y_{x_i}) \right\}}$ is the proportion of the N confidence sets that contain the true x-values. It is argued by Lieberman *et al.* (1967), Scheffé (1973), Aitchison (1982), Mee *et al.* (1991), Mee and Eberhardt (1996), Mathew and Zha (1997), Mathew *et al.*

2

(1998) and Krishnamoorthy and Mathew (2009, Chapter 3) among others that this *property* is highly desirable in many applications, and overwhelming majority publications on infinite many calibrations aim to guarantee this *property*. Other properties that may have useful applications are discussed in Mee and Eberhardt (1996).

One standard way to construct the confidence sets $C(y_x)$ having this *property* is to use the (β, γ)-simultaneous tolerance intervals (STI's). Assume *a priori* that the unknown x-values corresponding to all the future y_x's are in a given interval $[a, b]$. For example, the true blood alcohol level of any driver cannot be lower than $a = 0$ or higher than some upper threshold b. The (β, γ)-STI's $[L(x; \mathcal{E}), U(x; \mathcal{E})]$ over the interval $x \in [a, b]$ satisfy

$$P_{\mathcal{E}} \left\{ P_{y_x} \left\{ L(x; \mathcal{E}) < y_x < U(x; \mathcal{E}) \,|\, \mathcal{E}, x \right\} \geq \beta \ \text{ for all } \ x \in [a, b] \right\} \geq \gamma \qquad (2)$$

where y_x denotes a future y-value corresponding to x and is independent of the training data $\mathcal{E}$, the probability P_{y_x} is with respect to y_x and conditional on $\mathcal{E}$, and the probability $P_{\mathcal{E}}$ is with respect to $\mathcal{E}$. Then for each future y_x the confidence set $C(y_x)$ for the corresponding x is defined as

$$C(y_x) = \left\{ x \in [a, b] : \ L(x; \mathcal{E}) \leq y_x \leq U(x; \mathcal{E}) \right\}. \qquad (3)$$

It is shown in Scheffé (1973, Appendix B) that these confidence sets $C(y_x)$ have the *property* in (1).

Numerical results in Mee and and Eberhardt (1996) and Lee (1999) lead to the conjecture that the property in (1) is guaranteed by using the pointwise tolerance intervals (PTI's) instead of the STI's in (3). We have constructed counter examples to show that the property in (1) does not hold in general if the STI's are replaced by the PTI's in the construction of $C(y_x)$ in (3). The counter examples are not given here to save space but available from the authors. Hence the STI's are central to the construction of $C(y_x)$'s in order to guarantee the property in (1).

Construction of (β, γ)-STI's is considered first by Lieberman and Miller (1963) for simultaneous predictions, and there are three construction methods

3

available in the literature. The first is the construction of *central* (β, γ)-STI's by Lieberman and Miller (1963), Lieberman *et al.* (1967) and Scheffé (1973). Note that the central (β, γ)-STI's are two-sided and even the exact central (β, γ)-STI's are conservative as (β, γ)-STI's. The aforementioned papers only provide conservative central (β, γ)-STI's. The second is the probability set method by Wilson (1967) and Liman and Thomas (1988). Similar to the confidence set construction method of Rao (1973, pp.473), this method hinges on a γ level confidence set for the unknown parameters of the regression model and the resultant (β, γ)-STI's are also conservative and two-sided. The third is an exact method by Mee *et al.* (1991) for two-sided (β, γ)-STI's and Odeh and Mee (1990) for one-sided (β, γ)-STI's. Since the first two methods are conservative while Mee *et al.*'s (1991) method is exact, the two-sided (β, γ)-STI's of Mee *et al.* (1991) are usually narrower and so better than the conservative (β, γ)-STI's, as demonstrated numerically in Mee *et al.* (1991, Section 3).

In this paper, we focus on polynomial regression. A calibration problem often involves only two quantities y and x (or their suitable transformations), and a polynomial regression model is a simple yet probably the most frequently used model to relate two quantities. Exact one-sided (β, γ)-STI's will be constructed for a polynomial model of any order $p-1$ over any given covariate interval $x \in [a, b]$. While the construction method of this paper is also applicable to other regression models, such as the fractional polynomials (cf. Royston and Altman, 1994), the key step of maximizing $K(\boldsymbol{x})$ in the expression (9) below may require a different optimization method depending on the specific form of the regression model considered.

The upper (β, γ)-STI's have $L(x; \mathcal{E}) = -\infty$ in (2), and the lower (β, γ)-STI's have $U(x; \mathcal{E}) = \infty$ in (2). The confidence set $C(y_x)$ corresponding to the upper STI's often takes the form of a lower confidence limit, which is most relevant for the example of blood alcohol level since the police want to catch those drivers whose blood alcohol levels are above the legal limit by using the lower confidence limits (cf. Krishnamoorthy *et al.*, 2001). The confidence set $C(y_x)$ corresponding to the lower STI's often takes the form of an upper confidence

4

limit, which is most relevant for the example of ATD since the company wants to monitor that the radon concentrations are not above the safety threshold set by government agency by using the upper confidence limits.

Note that the exact (β, γ)-STI's of Mee *et al.* (1991) and Odeh and Mee (1990) are for a multiple linear regression model, in which the covariates are assumed to have no functional relationships, over a special covariate region only. These STI's become conservative when applied to a polynomial regression model of order two (i.e. quadratic regression) or above. Even for the simple linear regression (i.e. polynomial regression of order one), these STI's are only over a covariate set that is symmetric about $\bar{x}$, the mean of the observed covariate values in $\mathcal{E}$. See Section 2.2 for more details.

The layout of this paper is as follows. Section 2 deals with the construction of exact one-sided (β, γ)-STI's for a polynomial regression model over a given covariate interval. It also shows why the exact one-sided (β, γ)-STI's for a multiple linear regression model of Odeh and Mee (1990) can be used to produce only conservative STI's for a polynomial regression model of order two or higher. Section 3 provides two examples to illustrate the new STI's and to demonstrate the conservativeness of the STI's of Odeh and Mee (1990) when used for polynomial regression. The notations are adapted largely from Mee *et al.* (1991).

2. Exact one-sided STI's for polynomial regression

Assume that y and the only covariate x are related by a polynomial regression model of order $p - 1 \, (\geq 1)$:

$$y = \alpha_0 + \alpha_1 x + \cdots + \alpha_{p-1} x^{p-1} + \epsilon = \boldsymbol{x}^T \boldsymbol{\alpha} + \epsilon$$

where $\boldsymbol{x} = (1, x, \cdots, x^{p-1})^T$, $\boldsymbol{\alpha} = (\alpha_0, \cdots, \alpha_{p-1})^T$ is the vector of unknown regression coefficients, and the ϵ's are independent $N(0, \sigma^2)$ errors with unknown variance $\sigma^2 > 0$. The observed training data $\mathcal{E}$ can be represented in the usual matrix form $\boldsymbol{y} = X\boldsymbol{\alpha} + \boldsymbol{\epsilon}$, where the $n \times p$ design matrix X is assumed to be

5

of full column-rank without loss of generality. The usual estimators of $\boldsymbol{\alpha}$ and σ are denoted as $\hat{\boldsymbol{\alpha}}$ and $\hat{\sigma}$.

Note that the upper (β, γ)-STI's are just a γ level upper simultaneous confidence band (SCB) for the β-quantile line of y_x, $\boldsymbol{x}^T\boldsymbol{\alpha} + z_\beta\sigma$, where z_β denotes the β-quantile of the standard normal distribution $N(0, 1)$. Specifically, by setting $L(x; \mathcal{E}) = -\infty$ in (2), the upper (β, γ)-STI's $U(x; \mathcal{E})$ satisfy

$$P_{\mathcal{E}} \left\{ \boldsymbol{x}^T\boldsymbol{\alpha} + z_\beta\sigma < U(x; \mathcal{E}) \ \text{ for all } \ x \in [a, b] \right\} \geq \gamma. \tag{4}$$

Similarly, the lower (β, γ)-STI's $L(x; \mathcal{E})$ satisfy

$$P_{\mathcal{E}} \left\{ L(x; \mathcal{E}) < \boldsymbol{x}^T\boldsymbol{\alpha} + z_{1-\beta}\sigma \ \text{ for all } \ x \in [a, b] \right\} \geq \gamma. \tag{5}$$

If $U^-(x; \mathcal{E})$ denotes the upper (β, γ)-STI's for regressing $-y$ on x, then it is straightforward to show that $L(x; \mathcal{E}) = -U^-(x; \mathcal{E})$. Hence the focus in this section is on the construction of $U(x; \mathcal{E})$. It is noteworthy that two-sided SCB's for a quantile line in linear regression have been studied by several researchers; see e.g. Han *et al.* (2015) and the reference therein. In particular, exact two-sided SCB's for a quantile line that are uniformly narrower than previously published SCB's are provided in Han *et al.* (2015).

2.1. Exact STI's

Exact upper STI's of the form

$$U(x; \mathcal{E}) = \boldsymbol{x}^T\hat{\boldsymbol{\alpha}} + \lambda\hat{\sigma}\left[z_\beta + \sqrt{(p+2)\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}}\right] \tag{6}$$

will be constructed, where $\lambda > 0$ is a critical constant chosen so that

$$P_{\mathcal{E}} \left\{ \boldsymbol{x}^T\boldsymbol{\alpha} + z_\beta\sigma < U(x; \mathcal{E}) \ \text{ for all } \ x \in [a, b] \right\} = \gamma \tag{7}$$

for given $0 < \beta, \gamma < 1$. This form is taken from Odeh and Mee (1990) and similar to what is used in Scheffé (1973) and Mee *et al.* (1991) for the case of two-sided STI's. It warrants further research to try other forms in the hope of finding smaller upper limits.

6

To find the critical constant λ, the probability in (7) is written as

$$P_{\mathcal{E}} \left\{ \boldsymbol{x}^T \boldsymbol{\alpha} + z_\beta \sigma < \boldsymbol{x}^T \hat{\boldsymbol{\alpha}} + \lambda \hat{\sigma} \left[z_\beta + \sqrt{(p+2)\boldsymbol{x}^T (X^T X)^{-1} \boldsymbol{x}} \right] \text{ for all } x \in [a,b] \right\}$$

$$= P \left\{ \max_{x \in [a,b]} \frac{\boldsymbol{x}^T (\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}})/\sigma + z_\beta}{(\hat{\sigma}/\sigma) \left[z_\beta + \sqrt{(p+2)\boldsymbol{x}^T (X^T X)^{-1} \boldsymbol{x}} \right]} < \lambda \right\} = P\{Q < \lambda\} \qquad (8)$$

where

$$Q = \max_{x \in [a,b]} K(\boldsymbol{x}) \quad \text{with} \quad K(\boldsymbol{x}) = \frac{\boldsymbol{x}^T \boldsymbol{Z} + z_\beta}{u \left[z_\beta + \sqrt{(p+2)\boldsymbol{x}^T (X^T X)^{-1} \boldsymbol{x}} \right]}, \qquad (9)$$

where $\boldsymbol{Z} = (\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}})/\sigma \sim \boldsymbol{N}(\boldsymbol{0}, (X^T X)^{-1})$, $u = \hat{\sigma}/\sigma \sim \sqrt{\chi_\nu^2/\nu}$ with $\nu = n - p$, and $\boldsymbol{Z}$ and u are independent. It is clear from the expression of Q in (9) that the probability in (8) has nothing to do with the unknown parameters $\boldsymbol{\alpha}$ and σ. Furthermore, λ is simply the γ-quantile of the random variable Q, depending only on $p, \nu, \beta, (X^T X)^{-1}, \gamma$ and the interval $[a, b]$.

For a general $p \geq 2$ the following simple simulation-based method for finding λ fast and accurately is used. A large number R of independent replicates of $Q : Q_1, \cdots, Q_R$ are simulated, and the γ-quantile of the sample $Q_1, \cdots, Q_R$ is used as λ. It is well known that this approximation approaches λ almost surely as R approaches infinity (cf. Serfling, 1980). This approach of using sample quantile to approximate the population quantile has been used successfully in solving many otherwise difficult problems; see, for example, Edwards and Berry (1987) and Liu *et al.* (2004, 2005, 2010).

In each simulation of Q, independent $\boldsymbol{Z} \sim \boldsymbol{N}(\boldsymbol{0}, (X^T X)^{-1})$ and $u \sim \sqrt{\chi_\nu^2/\nu}$ are simulated first and Q is then computed using the expression in (9). Note that Q is equal to the maximum of $K(\boldsymbol{x})$ over the finite set $\mathcal{F}$ of x-values that contains only a, b and all the stationary points of $K(\boldsymbol{x})$ in the interval $x \in [a, b]$. All the stationary points of $K(\boldsymbol{x})$ can be solved from $K'(\boldsymbol{x}) = 0$, where $f'(\boldsymbol{x})$ denotes the derivative of $f(\boldsymbol{x})$ with respect to x. Now a few lines of calculus show that $K'(\boldsymbol{x}) = 0$ is equivalent to

$$- 4 (z_\beta)^2 d(\boldsymbol{x}) (q'(\boldsymbol{x}))^2 + (p+2) (q(\boldsymbol{x})d'(\boldsymbol{x}) - 2q'(\boldsymbol{x})d(\boldsymbol{x}))^2 = 0 \qquad (10)$$

7

where $d(\boldsymbol{x}) = \boldsymbol{x}^T (X^T X)^{-1} \boldsymbol{x}$ and $q(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{Z} + z_\beta$. Since the left-side of the equation in (10) is a polynomial function of x (of order $6(p-1)-4$), all the roots of the equation in (10) can quickly be found numerically by using available algorithm; the Matlab built-in function `roots` is used in our computation. Hence the set $\mathcal{F}$ and so Q can be computed efficiently. A Matlab program for computing λ is available from the authors. This technique of computing $\max_{x \in [a,b]} K(\boldsymbol{x})$ has been used successfully in several other problems; see e.g. Liu *et al.* (2008) and Wan *et al.* (2015).

Each λ in Examples 3.1-3.2 of Section 3 is computed using $R = 1,000,000$ simulations and took about 500 seconds on an ordinary Window's PC (Core(TM2) Due CPU P8400@2.26GHz). These critical constants are accurate to at least two decimal places from our experiments with different random seeds. For example, when $\beta = 0.95$, $\gamma = 0.99$, $p = 2$, $\tau = 2$ and $n = 40$, the critical value 1.2675 given in Odeh and Mee (1990, Table 1.3) is computed to be 1.2671 by our Matlab program. It is worth emphasizing that the λ, computed accurately once, will be used for calculating all future $C(y_x)$'s.

2.2. The STI's of Odeh and Mee (1990)

To be specific, assumed that y and the $(p-1)$ covariates $\boldsymbol{x}_{(0)} = (x_1, \cdots, x_{p-1})^T$ are related by a multiple linear regression model:

$$y = \alpha_0 + \alpha_1 x_1 + \cdots + \alpha_{p-1} x_{p-1} + \epsilon = \boldsymbol{x}^T \boldsymbol{\alpha} + \epsilon$$

where $\boldsymbol{x}^T = (1, x_1, \cdots, x_{p-1}) = (1, \boldsymbol{x}_{(0)}{}^T)$, $\boldsymbol{\alpha} = (\alpha_0, \cdots, \alpha_{p-1})^T$, and the ϵ's are independent $N(0, \sigma^2)$ errors with unknown variance $\sigma^2 > 0$. As in Section 2.1, the observed training data $\mathcal{E}$ can be represented in matrix notation $\boldsymbol{y} = X\boldsymbol{\alpha} + \boldsymbol{\epsilon}$, where the $n \times p$ design matrix X is assumed to be of full column-rank without loss of generality.

Odeh and Mee (1990) consider exact upper (β, γ)-STI's

$$U_{OM}(\boldsymbol{x}_{(0)}; \mathcal{E}) = \boldsymbol{x}^T \hat{\boldsymbol{\alpha}} + \lambda \hat{\sigma} \left[z_\beta + \sqrt{(p+2)\boldsymbol{x}^T (X^T X)^{-1} \boldsymbol{x}} \right] \qquad (11)$$

8

for $\boldsymbol{x}_{(0)}$ in the covariate region $\mathcal{D} = \mathcal{D}_{\tau_2} - \mathcal{D}_{\tau_1}$, where $0 \leq \tau_1 < \tau_2$ are given constants and

$$\mathcal{D}_\tau = \left\{ \boldsymbol{x}_{(0)} : \ \left(\boldsymbol{x}_{(0)} - \bar{\boldsymbol{x}}_{(0)} \right)^T \boldsymbol{V}^{-1} \left(\boldsymbol{x}_{(0)} - \bar{\boldsymbol{x}}_{(0)} \right) \leq \tau^2 \right\} \tag{12}$$

with $\bar{\boldsymbol{x}}_{(0)} = (\bar{x}_1, \cdots, \bar{x}_{p-1})^T$ being the sample mean and $\boldsymbol{V}$ the sample covariance matrix of the observed values on the $p-1$ covariates $\boldsymbol{x}_{(0)}$ in the training data $\mathcal{E}$. Tables of critical constant λ in (11) over $\boldsymbol{x}_{(0)} \in \mathcal{D} = \mathcal{D}_\tau$ are provided in Odeh and Mee (1990) for several τ values.

When $p = 2$, $\mathcal{D}_\tau$ is an interval symmetric about the mean $\bar{x}_1$ of the only covariate. Hence if STI's over the interval $[a, b] = [\bar{x}, \bar{x} + \delta]$, for a given $\delta > 0$, for example, are required then the exact STI's of Odeh and Mee (1990) over the interval $[\bar{x} - \delta, \bar{x} + \delta]$ have to be used as conservative STI's over the required interval $[a, b] = [\bar{x}, \bar{x} + \delta]$. Some numerical information on the conservativeness of the λ from the tables in Odeh and Mee (1990) in comparison with the exact λ over $[\bar{x}, \bar{x} + \delta]$ from the method in Section 2.1 is given in Example 3.1 below.

When $p \geq 3$, we first show how the λ from the tables in Odeh and Mee (1990) can be used as a conservative value for a polynomial regression model of order $p - 1$ over the given covariate interval $x \in [a, b]$. For ease of exposition, we assume $p = 3$. In this case, in order to use the λ from the tables in Odeh and Mee (1990), one has to determine the $\mathcal{D}_\tau$ in the following way.

First, determine the centre of $\mathcal{D}_\tau$, which is given by $(\bar{x^1}, \bar{x^2})$, where $\bar{x^k} = (1/n) \sum_{i=1}^n x_{0i}^k$ since the x_{0i}^k $(i = 1, \cdots, n)$ are the observed value on the k-th covariate $(k = 1, 2)$ in the multiple regression model. Second, determine the region $\mathcal{D}^* = \{(x, x^2) : x \in [a, b]\}$, which is a (one-dimensional) curve in the two-dimensional (x_1, x_2)-plane. Finally, choose the smallest value of τ so that $\mathcal{D}_\tau$ just includes the curve $\mathcal{D}^*$.

Now the critical value λ from Odeh and Mee (1990) over $\mathcal{D}_\tau$ is used as a conservative critical value of STI's for the quadratic polynomial over $x \in [a, b]$. It is clear that the λ from Odeh and Mee (1990) is the γ-quantile of $\max_{\boldsymbol{x}_{(0)} \in \mathcal{D}_\tau} K(\boldsymbol{x})$, while the λ of our exact method is the γ-quantile of $\max_{\boldsymbol{x}_{(0)} \in \mathcal{D}^*} K(\boldsymbol{x})$, where $K(\boldsymbol{x})$ is given in (9) but with $\boldsymbol{x} = (1, x_1, x_2)^T = (1, \boldsymbol{x}_{(0)}{}^T)^T$. Since $\mathcal{D}^*$ is

9

a one-dimensional region that is strictly contained in the two-dimensional region $\mathcal{D}_\tau$, the λ from Odeh and Mee (1990) is larger than the λ from our exact method. This observation holds for a general $p \geq 3$ even when the interval $[a, b]$ is symmetric about the mean of the observed values on the only covariate in polynomial regression. Example 3.2 below provides numerical information on the conservativeness of the λ from Odeh and Mee (1990) when used for polynomial regression.
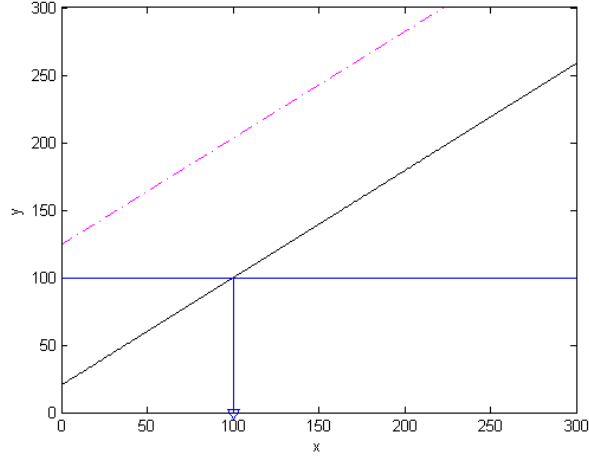
3. Examples

3.1. Radon Example Mee and Eberhardt (1996, Table 3) provide data from a calibration experiment of ATD's which are used to measure indoor concentrations of radon. In the experiment, the ATD's were exposed to known levels of radon x in a laboratory. The response variable y is an optical count of number of damage tracks, caused by alpha radioactive decays, over a specific area of the film. After the usual model diagnosis, the $n = 40$ observations are fitted by a simple linear regression model in Mee and Eberhardt (1996), with the fitted least squares line $\hat{y} = 124.4 + .789x$, $\hat{\sigma} = 41.26$, $R^2 = 0.93$, $\nu = n-p = 38$, $\min(x_{0,1}, \cdots, x_{0,40}) = 50$, $\max(x_{0,1}, \cdots, x_{0,40}) = 4241$, $\bar{x} = 683.3$ and $S_x = \sum_{i=1}^{n}(x_{0i} - \bar{x})^2 = 5.717 \times 10^7$.

In order to construct upper confidence bounds on the level of radon exposures x corresponding to future observed ATD readings y_x, lower STI's are required. Set $\beta = 0.95$ and $\gamma = 0.99$. For the STI's on $[a, b] = [0, 3074]$ considered in Mee and Eberhardt (1996), one finds the critical constant $\lambda = 1.2675$ from Odeh and Mee (1990, Table 1.3, with $\tau = 2$ and $n = 40$) and the lower STI's are given by

$$L(x; \mathcal{E}) = \boldsymbol{x}^T\hat{\boldsymbol{\alpha}} - \lambda\hat{\sigma}\left[z_\beta + \sqrt{(p+2)\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}}\right] \quad \text{for all } x \in [a, b]. \quad (13)$$

Note that these STI's are actually on the covariate interval $\bar{x} \pm \tau\sqrt{S_x/n} = [-1708, 3074]$, which is the smallest $\mathcal{D}_\tau$ for $p = 2$ that contains the interval $[a, b] = [0, 3074]$. The method of Section 2.1 can be used to computed the exact STI's on $[0, 3074]$, and the critical constant λ is computed to be 1.2557, which is

10

Figure 1: *The fitted regression line (dashdot line), the STI's (13) with $\lambda = 1.2557$ (solid curve), and the corresponding upper confidence bound on x for $y_x = 100$ (the arrow)*


smaller, though only marginally in this case, than the conservative $\lambda = 1.2675$ from Odeh and Mee (1990) as expected.

The exact STI's with $\lambda = 1.2557$ are plotted in Figure 1 by the solid curve. If a future $y_x = 100$, for example, is observed then the upper bound on x is 100.3, which is the x-coordinator of the point where the level line $y = 100$ intersects the STI's.

As pointed out above, the λ from the tables in Odeh and Mee (1990) is for interval $[a, b]$ that is symmetric about $\bar{x}$. Hence if $[a, b] = [\bar{x}, \bar{x} + \delta]$, for a given $\delta > 0$, then the λ from Odeh and Mee (1990) for the interval $[\bar{x} - \delta, \bar{x} + \delta]$ has to be used as a conservative value for the desired interval $[a, b] = [\bar{x}, \bar{x} + \delta]$. To get some idea on the degree of conservativeness of the λ from Odeh and Mee (1990) in comparison with the exact λ for $[\bar{x}, \bar{x} + \delta]$ from Section 2.1, several δ values have been tried for this particular example. When $\delta = 4782200$ for example, the λ from Odeh and Mee (1990) is 1.3848 and the exact λ is 1.3016. Hence the conservative $\lambda = 1.3848$ is about $(1.3848 - 1.3016)/1.3016 = 6.4\%$ larger than the exact $\lambda = 1.3016$. The conservativeness of λ from Odeh and Mee (1990) when used for a quadratic polynomial can be more pronounced, however, as

11

Figure 2: *The fitted regression model (dashdot curve), upper STI's (6) (solid line), and the lower confidence bound on x, 9.34(indicated by the arrow), for $y_x = 137.2$ (the level line)*
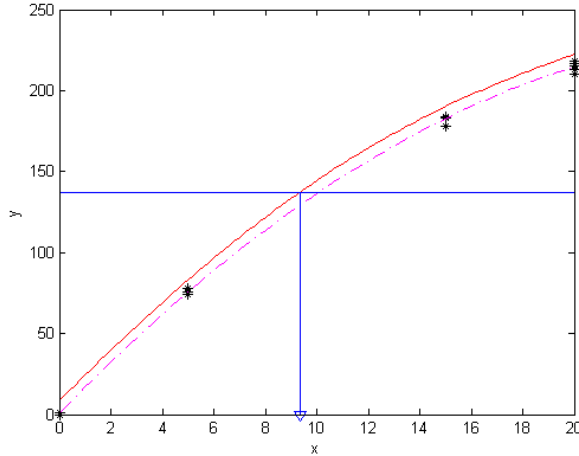
demonstrated in Example 3.2 below.

3.2. Spectroscopy Example Graphite furnace atomic absorption spectroscopy (GFAAS) is an analytical technique for the determination of trace metal concentrations in different types of samples. When a sample is atomized, an absorbency signal is obtained with a stripchart recorder. The variation of (peak) absorbency y (in mm) with concentration x (parts per billion) is established by atomizing samples of known concentrations and the obtained calibration curve is used to infer the concentrations of future observed disturbances. Lundberg and De Maré (1980, Table 1) provide data from a calibration experiment of GFAAS. After the usual model diagnosis, a quadratic regression model is recommended in Lundberg and De Maré (1980); the fitted model is given by $\hat{y} = 0.729 + 16.44x - 0.287x^2$ and fits very well the $n = 21$ pairs of observations at four different concentrations $x = 0, 5, 15$ and 20 from the calibration experiment.

Now suppose that one wants to construct lower confidence bounds on x for future observed y_x's. For this one can use the upper STI's in (6). For $\beta = 0.95$,

12

Figure 3: *The curve $\mathcal{D}^*$ (dotted curve) and the region $\mathcal{D}_\tau$ for $\tau = 1.6652$ (ellipse)*
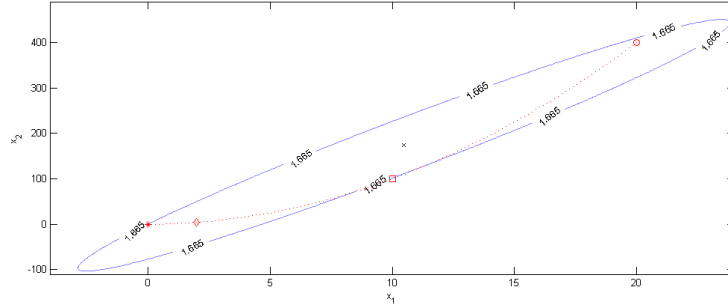
$\gamma = 0.99$ and $[a, b] = [0, 20]$, the exact critical constant is computed by our program to be $\lambda = 1.4213$. Figure 2 plots the training data points (stars), the fitted quadratic model (dashdot curve) and the upper STI's (6) (solid curve). From the upper STI's, if one observe $y_x = 137.2$, for example, then the lower confidence bound on x is 9.34 as indicated by the arrow.

As pointed out in Section 2.2, in order to use the λ tabulated in Odeh and Mee (1990) as a conservative value in this example with $p = 3$, one needs to find the smallest τ so that $\mathcal{D}_\tau$ in (12) contains $\mathcal{D}^* = \{(x, x^2) : x \in [a, b]\}$. For $[a, b] = [0, 20]$, $\mathcal{D}^*$ is plotted in Figure 3 by the dotted curve with one end marked by the star and the other end by the small circle, and the corresponding $\mathcal{D}_\tau$ is bounded by the ellipse with $\tau = 1.6652$. With $\beta = 0.95$, $\gamma = 0.99$, $\tau = 1.6652$ and $n = 21$, the critical constant λ from Odeh and Mee (1990) is $\lambda = 1.4589$, which is larger than the exact $\lambda = 1.4213$ given above as expected.

Now note that the point (x, x^2) with $x = a = 0$, marked by the star in Figure 3, is on the boundary of $\mathcal{D}_\tau$ and so the same $\mathcal{D}_\tau$ has to be used for $\mathcal{D}^* = \{(x, x^2) : x \in [a, b]\}$ with $a = 0$ and any $b \in (0, 20]$; the point (x, x^2) with $x = 10$ is marked by the small square, and the point (x, x^2) with $x = 2$ is marked by the small diamond. Hence the same conservative $\lambda = 1.459$ from Odeh and Mee (1990) is used for interval $[a, b]$ with $a = 0$ and any $b \in (0, 20]$. On the other hand, the exact critical constant λ decreases as b decreases to $a = 0$. For example, the exact λ is equal to 1.366 for $b = 10$, 1.308 for $b = 2$,

13

and 1.251 for $b = 0.001$. This serves to demonstrate the potential gain and so importance of using the exact λ provided in this paper, even though, for this particular example, it is unlikely that a value of b less than 10 will be used given the observed covariate range of $[0, 20]$. It is expected that the λ from Odeh and Mee (1990) can potentially be even more conservative than the exact λ for $p = 4$, since in this case $\mathcal{D}^* = \{(x, x^2, x^3) : x \in [a, b]\}$ is still a one-dimensional region while $\mathcal{D}_\tau$ is a three-dimensional region containing $\mathcal{D}^*$.

Acknowledgments

Reference

[1] Aitchison, T.C. (1982). Discussion of Brown's paper. *JRSS(B)*, 44(3), 309-310.

[2] Brown, P.J. (1982). Multivariate Calibration. *JRSS(B)*, 44(3), 287-321.

[3] Brown, P.J. (1993). *Measurement, Regression and Calibration.* New York: Oxford University Press.

[4] Edwards, D. and Berry, J.J. (1987). The efficiency of simulation-based multiple comparisons. *Biometrics*, 43, 913-928.

[5] Eisenhart, C. (1939). The interpretation of certain regression methods and their use in biological and industrial research. *The Annals of Mathematical Statistics*, 10, 162-186.

[6] Han, Y., Liu, W., Bretz, F. and Wan, F. (2015). Simultaneous confidence bands for a percentile line in linear regression. *Computational Statistics and Data Analysis*, 81, 1-9.

[7] Krishnamoorthy, K. and Mathew, T. (2009). *Statistical Tolerance Regions: Theory, Applications and Computation.* New York: Wiley.

[8] Krishnamoorthy, K., Kulkarni, P.M. and Mathew, T. (2001). Multiple use one-sided hypotheses testing in univariate linear calibration. *Journal of Statistical Planning and Inference*, 93, 211-223.

[9] Lee, Y. (1999). *Tolerance Regions in Univariate and Multivariate Linear Models.* PhD dissertations, Dept of Mathematics and Statistics, University of Maryland Baltimore County, USA.

[10] Lieberman, G.J. and Miller, R.G., Jr. (1963). Simultaneous tolerance intervals in regression. *Biometrika*, 50, 155-168.

[11] Lieberman, G.J., Miller, R.G., Jr. and Hamilton, M.A. (1967). Simultaneous discrimination intervals in regression. *Biometrika*, 54, 133-145 (Correction: 58, p.687).

[12] Limam, M.M.T. and Thomas, D.R. (1988). Simultaneous tolerance intervals for the linear regression model. *J. Amer. Statist. Assoc.*, 83, 801-804.

[13] Liu, W. (2010). *Simultaneous Inference in Regression.* Chapman and Hall.

[14] Liu, W., Jamshidian, M. and Zhang, Y. (2004). Multiple comparison of several linear regression models. *J. Amer. Statist. Assoc.*, 99, 395-403.

[15] Liu, W., Jamshidian, M., Zhang, Y. and J. Donnelly (2005). Simulation-based simultaneous confidence bands in multiple linear regression with predictor variables constrained in intervals. *Journal of Computational and Graphical Statistics*, 14(2), 459-484.

[16] Liu, W., Wynn, H.P. and Hayter, A.J. (2008). Statistical inferences for linear regression models when the covariates have functional relationships: polynomial regression. *Journal of Statistical Computation and Simulation*, 78(4), 315-324.

[17] Lundberg, E. and De Maré, J. (1980). Interval estimates in the spectroscopy calibration problem. *Scandinavia Journal of Statistics*, 7, 40-42.

15

[18] Mathew, T., Sharma, M.K. and Nordstrom, K. (1998). Tolerance regions and multiple-use confidence regions in multivariate calibration. *The Annals of Statistics*, 26, 1989-2013.

[19] Mathew, T. and Zha, W. (1997). Multiple use confidence regions in multivariate calibration. *J. Amer. Statist. Assoc.* 92, 1141-1150.

[20] Mee, R.W. and Eberhardt, K.R. (1996). A Comparison of uncertainty criteria for calibration. *Technometrics*, 38, 221-229.

[21] Mee, R.W., Eberhardt, K.R. and Reeve, C.P. (1991). Calibration and simultaneous tolerance intervals for regression. *Technometrics*, 33, 211-219.

[22] Odeh, R.E. and Mee, R.W. (1990). One-sided simultaneous tolerance limit for regression. *Comm. Statist.– Simulation Comput.*, 19, 663-680.

[23] Osborne, C. (1991). Statistical calibration: a review. *Internat. Statist. Rev.*, 59, 309-336.

[24] Rao, C. R. (1973). *Linear Statistical Inference and Its Application, 2nd ed.* Oxford University Press.

[25] Royston P. and Altman, D.G. (1994). Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. *Applied Statistics*, 36, 429-467

[26] Scheffé, H. (1973). A statistical theory of calibration. *Annals of Statistics*, 1, 1-37.

[27] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley.

[28] Smith, R.L. and Corbett, M. (1987). Measure marathon courses: an application of statistical calibration theory. *Applied Statistics*, 36, 283-295.

[29] Wan, F., Liu, W., Bretz, F. and Han, Y. (2015). An exact confidence set for a maximum point of a univariate polynomial function in a given interval. *Technometrics*, to appear.

16

[30] Wilson, A.L. (1967). An approach to simultaneous tolerance intervals in regression. *The Annals of Mathematical Statistics*, 38, 1536-1540.

This paper provides, for the first time, exact one-sided STI's for a polynomial regression model of any order over any given interval.